RGNZ/MGPB

Multivariate Data Analysis to Assess Groundwater Hydrochemical Characterization in Rawadanau Basin, Banten Indonesia

Priyo Hartanto[1]; Rachmat Fajar Lubis[2]; Boy Yoseph C.S.S. Syah Alam[3]; Yoga Andriana Sendjaja[4]; Ismawan Ismawan[5]; T.Yan W.M. Iskandarsyah[6]; Hendarmawan Hendarmawan[7]

[1] *Research Center for Limnology and Water Resources, National Research and Innovation Agency (BRIN) Cibinong Science Center, Jl. Raya Bogor Km.46, Cibinong 16911, Indonesia., https://orcid.org/0000-0001-7759-6947*

[2] *Research Center for Limnology and Water Resources, National Research and Innovation Agency (BRIN) Cibinong Science Center, Jl. Raya Bogor Km.46, Cibinong 16911, Indonesia. https://orcid.org/0000-0002-5430-7169*

[3] *Faculty of Geological Engineering, Universitas Padjadjaran, Jatinangor, Sumedang 45363, Indonesia*

[4] *Faculty of Geological Engineering, Universitas Padjadjaran, Jatinangor, Sumedang 45363, Indonesia, https://orcid.org/0000-0001-9606-3046*

[5] *Faculty of Geological Engineering, Universitas Padjadjaran, Jatinangor, Sumedang 45363, Indonesia,*

[6] *Faculty of Geological Engineering, Universitas Padjadjaran, Jatinangor, Sumedang 45363, Indonesia, https://orcid.org/0000-0002-7583-9244*

[7] *Faculty of Geological Engineering, Universitas Padjadjaran, Jatinangor, Sumedang 45363, Indonesia, https://orcid.org/0000-0003-0900-3331*

Abstract

A multivariate statistical technique of principal component analysis (PCA) and hierarchical cluster analysis (HCA) has been applied to identify and classify the various water sources that comprise the Rawadanau Basin. The data collection includes 60 samples taken during the dry (29 samples) and the rainy season (31 samples) in tropical regions. Sources of sampled water include dug wells, rivers, cold springs, and hot springs. Water chemistry measurable variables include field data (T, pH, EC), major ions (Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^-, HCO_3^-, SO_4^{2-}), SiO_2, Fe_{total}, Mn, and stable isotopes of water (δ^2H, and $\delta^{18}O$). The correlation of the concentration of water chemistry shows changes in the rainy season to Fe_{total} and Mn. Interpretation based on HCA using the dendrogram based on the chemical elements of water produces two clusters. Cluster A reflects an unconfined aquifer and bicarbonate type. Meanwhile, cluster B is a chloride type from the confined aquifer and does not change in different seasons. The PCA results show that the three-component matrix accounts for 86.12% of the data structure describing the Rawadanau Basin water sources that volcanic rocks affect and strongly correlate with Na^+, K^+, Ca^{2+}, and Mg^{2+}. PC1 has a high positive value for hydrochemical composition, indicating that lithology influences the kind of water. PC2 has a positive value for the stable isotope ($\delta^{18}O$ and δ^2H), meaning it is the main water source in Rawadanau. PC3 has a positive value for elevation and a negative for longitude, indicating a recharge area influenced by geological factors and is correlated with geothermal influences and volcanic rocks. This multivariate analysis can identify components and clusters of hydrochemical variables that have not been determined in previous studies.

Keywords:

HCA; PCA; water chemistry; Rawadanau; volcanic rocks

1. Introduction

Sustainable development in all aspects of life depends on the stability of water resources. Water availability includes quantity and quality to fulfill the increasing needs of households, industry, and agriculture. This availability is influenced by many factors, including climate, geology, vegetation, and anthropogenic factors (**Hem, 1985**). Some minerals, such as carbonates and evaporites, dissolve quickly and change the composition of the water more quickly, while other minerals, such as silicates, dissolve more slowly and have a less pronounced effect on the composition of the water. In addition, temperature also plays a role in controlling freshwater bodies' chemical and biological composition (**Hem, 1985; Yidana et al., 2008**).

The need for clean water in the Cilegon industrial area and its surroundings has been fulfilled by Rawadanau, a mountain lake supplied by the Cidanau River, which empties into the Sunda Strait. The Rawadanau Basin is the leading supplier of water reserves from cold springs, hot springs, rivers, and rainwater. Quality and quantity should be maintained for the sustainability of the water supply. So far, the water supply to Rawadanau is believed to come from rainwater manifested as sur-

---

Corresponding author: Priyo Hartanto

e-mail address: priy004@brin.go.id

face water and springs. In the hydrologic cycle, water can interact with the rock and the types of soil traverses to trace its origins using hydrochemical properties. Previous research has confirmed the Rawadanau Basin water potential from thermal and non-thermal sources from meteoric water (**Syah et al., 2019; Hartanto et al., 2021, 2022**). Meanwhile, hot springs produce various types of water, with heating sourced from beneath Mount Parakasak (**Syah et al., 2019; Hartanto et al., 2021, 2022**). Springs with temperatures above 36.7°C and exceeding the average ambient area are called hot springs, and those below are considered cold springs (**Alfaro & Wallace, 1994**).

Identifying and understanding the sources of contaminants in groundwater requires the application of methods to identify and quantify the sources of influence and their outcomes (**Scheiber et al., 2023; Solano et al., 2023**). The result of recharge can vary in type and quality, which is influenced by the composition of the solute. Groundwater contaminants and hydrochemistry data can be effectively elaborated, interpreted, and represented using multivariate statistical techniques like cluster analysis (CA) and factor analysis (FA) (**Brown, 1998; Lee et al., 2003; Singh et al., 2004; Hartmann et al., 2005; Kim et al., 2005; Yidana et al., 2008; Tanasković et al., 2012; Blake et al., 2016; Jolliffe & Cadima, 2016; Trabelsi & Zouari, 2019; Sudhakaran et al., 2020; Kale et al., 2021; Scheiber et al., 2023**). Multivariate analysis helps characterize the spatial variability of heterogeneity in nature, complex interaction issues, and the joint occurrence of different parameters. Factor analysis, a multivariate statistical method, creates general relationships between measured variables by revealing multivariate patterns that can help classify and interpret the original data (**Tanakovic et al., 2012; Trabelsi & Zouari, 2019; Kale et al., 2021**). Hierarchical cluster analysis (HCA) is a semi-quantitative helpful technique for classifying individuals based on similarities into relatively homogeneous groups and emphasizing genetic relationships called clusters. The grouping is based on the similarity of the main parameters used to guide the formulation of a suitable distance measure is used as the cluster threshold (**Tanasković et al., 2012; Trabelsi & Zouari, 2019; Singh et al., 2004; Kim et al., 2005; Scheiber et al., 2023**). The multivariate statistical technique used to identify significant components or factors that explain the majority of the variance in the original system is called principal component analysis (PCA). The PCA approach produces a linear combination that represents the type of measurement, which is a better descriptor than the original measurement provided that the number of observations must be greater than the number of variables (**Helena et al., 2000; Tanasković et al., 2012; Jolliffe & Cadima, 2016; Trabelsi & Zouari, 2019; Kale et al., 2021**). PCA and HCA statistical methods were performed on 60 water samples in this investigation.

This area needed research to analyze groundwater conditions and assess groundwater quality in two seasons. Understanding of the components that influence groundwater quality is important in order to help conserve these water resources for the future. This paper will discuss and emphasize the water sources that fill the Rawadanau Basin and the temporal effects on groundwater. In addition to the development of previous research, which revealed the origin of water during the dry season and the characteristics of hot springs with heating sourced from Mount Parakasak's sub-surface (**Hartanto et al., 2021, 2022**), this study aims to apply multivariate statistics to characterize groundwater and predict what components influence it using hydrochemical parameters and stable isotopes of water. Furthermore, it contributes to understanding the hydrochemical characteristics of two seasons in volcanic product areas and the hydrogeochemistry that controls the water quality. For this purpose, data on water chemistry was collected for the dry and rainy seasons, while some were taken for two seasons from rivers, wells, and springs.

2. Methods

2.1. Area description

2.1.1. Location

Indonesia is a region that has two seasons, namely, the rainy and dry seasons. The study area is located on the west side of Java Island, precisely in the Serang Banten Province of Indonesia, at coordinates between 105° 51` 30" - 106° 03` 00" E and 06° 07` 10" - 06° 19` 10" S (see **Figure 1**). The study area is on a hilly plateau between two volcanoes that get an average of 2,806 millimeters of rain per year during the dry and wet seasons (**Hartanto et al., 2019**), while the average monthly air temperature is between 26.2 and 27.7°C. Between mid-November and May, there is a rainy season with monthly precipitation above 200 millimeters, with a peak rainy season from December to February (> 300 millimeters/month). The dry season lasts from late May to mid-November, while the peak is from August to September when the intensity is less than 50 mm/month. The morphology of the study area is between 100 and 1350 m above sea level (asl), with a recharge area at an elevation of 440 to 500 m asl (**Djiono et al., 2011**).

2.1.2. Geological setting

Rawadanau, which is the study area, was formed from the Late Pleistocene to Holocene of Quaternary volcanic rocks produced by the Karang-Parakasak Volcano (south) and Garung Volcano (north) and is controlled by two fault systems, namely Wangun and Garung with the NE-SW pattern (**Van Bemmelen, 1949; Rusmana et al., 1991; Santosa et al., 1991; Suryadarma & Fauzi, 1991**). Litho-stratigraphically, it is composed of sedi-
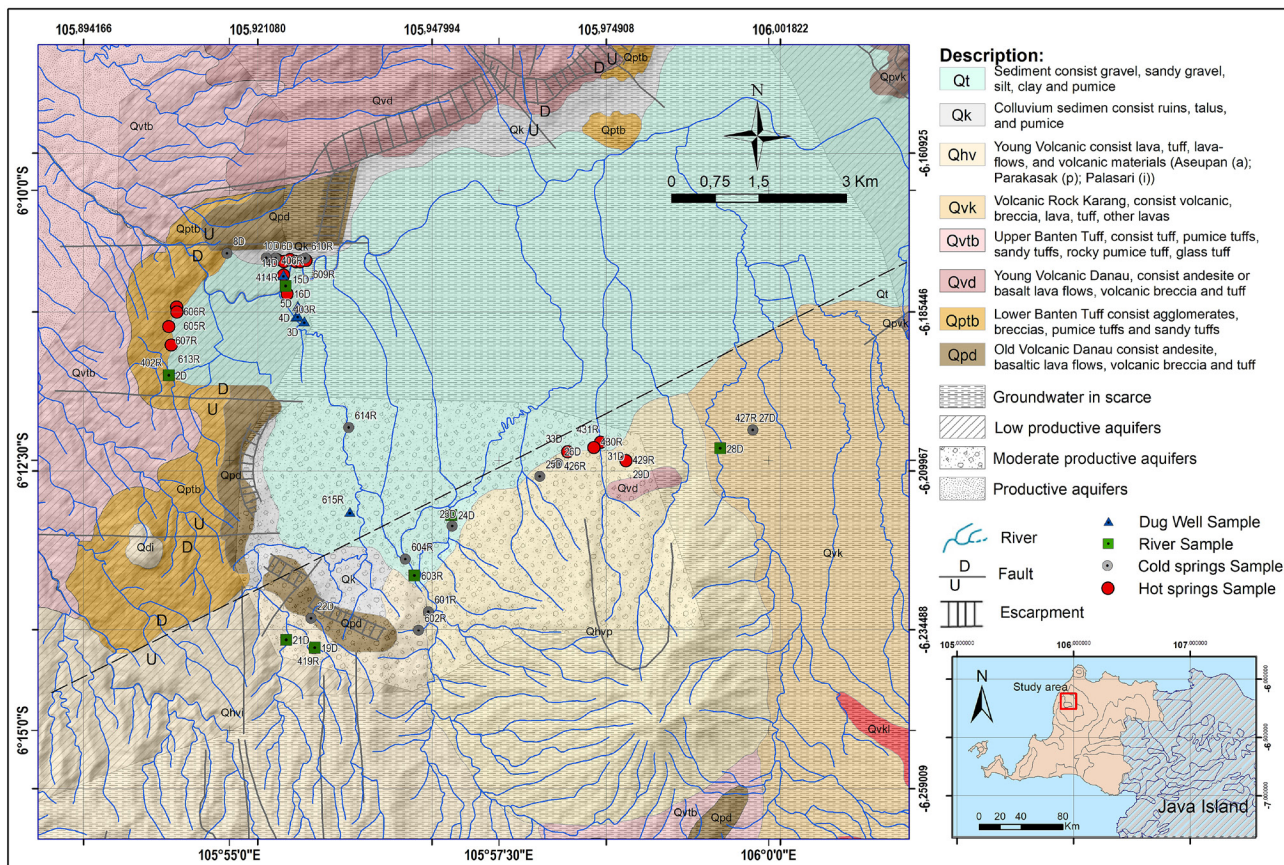
Figure 1: The map shows the study area's geological and hydrogeological setting (Modified after (**Santosa et al., 1991; Suryaman, 1999; Van Bemmelen, 1949**)). The coordinates are in UTM 48S.

mentary rocks (Qt, Qk) and volcanic products of Karang-Parakasak and Garung. Volcanic rock products include Karang Volcanic (Qvk), Young Volcanic (Qhv), Upper Banten Tuff Deposit (Qvtb), Young Volcanic Danau (Qvd), Lower Banten Tuff (Qptb), and Old Volcanic Danau Rock (Qpd) (see **Figure 1**).

Based on hydrogeology, especially aquifer capacity and productivity, the Rawadanau Basin is divided into extensive and highly productive aquifers, low to moderately productive aquifers, and scarcity (**Suryaman, 1999**). Groundwater flows through cracks and spaces between grains and rock fractures. Productive aquifer areas are characterized by a discharge of >5 litres/second, while moderate productive aquifers have a discharge of < 5 litres/second (see **Figure 1**); both have degree transmissivity between 400 and 2,000 m²/day (**Hartanto et al., 2022**). Meanwhile in scarce areas, water is only available in the rainy season.

2.2. Water analysis

2.2.1. Water sampling and analysis

Data on water chemistry was collected for the dry and rainy seasons, and some were taken for two seasons from rivers, wells, and springs. Field data were collected in April, September, and November 2020: dry season samples from September (labelled as D) and rainy season samples from April and November (labelled as R). The location map is presented in **Figure 1**. For the dry season, 29 samples were collected for the data sets while there were 31 samples collected during the rainy season. In comparison, 16 samples of the same ID were taken during the two seasons, differentiated by R or D (for example, 2D and 2R) (see **Table 1**). Sources of sampled water include dug wells, rivers, cold springs, and hot springs. A Water Quality Checker Toac 22EP was used to measure temperature (T), pH, and electrical conductivity (EC) immediately on the field. Information on water sampling for each season and the source of the water is summarized in **Table 1**.

The samples were placed in a 500-milliliter polyethylene bottle for major ions (anions and cations) based on APHA (**APHA, 2005**) and analyzed at the Research Center Water Laboratory BRIN using atomic absorption spectrophotometry (AAS). The major ions comprise sodium (code: 3500-Na$^+$. B) and potassium (code 3500-K$^+$), calcium (code 3500-Ca^{2+}. B), magnesium (code 3500-Mg^{2+}. B), sulfate (code: 4500-SO$_4^{2-}$ E), chloride (code: 4500-Cl$^-$. B), and silica SiO$_2$ (code: 4500-SiO$_2$ C). Bicarbonate was analyzed directly in the field after sampling (code 2320 B). In contrast, total dissolved solids (TDS) were computed from EC by multiplying a factor

Table 1: Data from field measurements (D = dry season, R = Rainy season, DW = dug well, CS= cold spring, HS= hot spring)

No.	ID	T (°C)	pH	EC(µS/cm)	Type		No.	ID	T (°C)	pH	EC(µS/cm)	Type
1	1D	27.7	7.65	222.0	River		31	604R	27.6	5.73	118.0	CS
2	1R	27.1	7.15	69.7	River		32	23D	27.6	6.55	140.0	CS
3	2D	26.2	7.11	64.0	River		33	25D	27.7	5.70	168.9	CS
4	2R	26.7	6.92	31.5	River		34	26D	27.9	5.84	170.0	CS
5	15R	29.2	6.93	239.0	River		35	26R	27.2	5.59	96.3	CS
6	19D	23.5	7.56	97.0	River		36	27D	25.8	6.22	151.0	CS
7	19R	24.7	7.66	67.4	River		37	27R	25.4	6.24	148.0	CS
8	21R	25.0	7.10	73.0	River		38	614R	27.9	6.39	251.0	CS
9	603R	29.2	8.22	158.0	River		39	610R	30.5	6.32	468.0	CS
10	24R	27.7	7.21	141.0	River		40	606R	41.0	6.90	1380.0	HS
11	608R	27.9	7.56	166.0	River		41	605R	38.5	7.00	1600.0	HS
12	28R	29.7	7.61	159.0	River		42	613R	33.1	7.16	1330.0	HS
13	3D	26.8	6.16	262.0	DW		43	607R	34.4	7.32	1190.0	HS
14	3R	26.9	4.93	69.7	DW		44	11D	44.0	6.78	1298.0	HS
15	4D	29.7	6.00	269.0	DW		45	11R	45.0	6.69	1405.0	HS
16	4R	26.1	6.53	194.3	DW		46	7D	42.6	6.75	1240.0	HS
17	5D	27.7	6.69	473.0	DW		47	7R	42.0	6.71	1278.0	HS
18	14D	27.2	6.76	505.0	DW		48	611R	34.4	7.03	1810.0	HS
19	14R	26.3	6.66	214.0	DW		49	612R	37.7	6.91	1880.0	HS
20	16D	30.5	6.69	612.0	DW		50	12D	41.4	6.88	1760.0	HS
21	16R	29.9	6.47	478.0	DW		51	609R	39.1	6.83	1540.0	HS
22	615W	28.2	6.10	478.0	DW		52	17D	31.2	6.63	1576.0	HS
23	6D	30.4	6.49	210.0	CS		53	13D	41.3	6.62	2410.0	HS
24	6R	28.7	6.18	194.9	CS		54	13R	41.6	6.54	1540.0	HS
25	8D	23.6	6.50	187.0	CS		55	29D	49.1	7.27	1155.0	HS
26	10D	29.5	6.43	221.0	CS		56	29R	47.1	7.68	1099.0	HS
27	20D	25.1	5.95	116.0	CS		57	30R	54.6	6.16	4460.0	HS
28	22D	25.3	6.11	46.0	CS		58	31D	49.7	6.32	2790.0	HS
29	601R	28.6	7.70	220.0	CS		59	31R	50.6	6.22	3870.0	HS
30	602R	27.2	7.46	217.0	CS		60	33D	51.8	6.48	2850.0	HS

of 0.65 (**Hem, 1985**). The accuracy of the analysis results is calculated by error balance with the error range within 5% (**Appelo & Postma, 2005**).

For stable isotope analysis of water ($\delta^{18}O$ and $\delta^{2}H$), the sample was placed in a 100 ml polythene bottle. The stable isotope analyses were done using the Picarro L-2130-i analyzer, at the Hydrochemical Laboratory of the Center for Environmental Management Geology at the Geological Agency. The result of stable isotope analysis expressed by ‰ values and reported to δ notation relative to deviation value of the Standard Mean Oceanic Water (SMOW) ratio (**Hamed, 2014; Hartanto et al., 2022**) based on the **Equation 1**:

$$\delta = [(R_{sample}/R_{SMOW})-1] \times 1{,}000 \qquad (1)$$

where, R_{Sample} = the isotope ratio ($\delta^{2}H$ or $\delta^{18}O$) of the sample; R_{SMOW} = the isotope ratio ($\delta^{2}H$ or $\delta^{18}O$) of the SMOW.

2.2.2. Statistical analysis of geochemical data

Investigating many variables allows one test rather than many univariate and bivariate tests (**Brown, 1998**). HCA classifies samples or "hydrochemical facies" based on statistical similarities, and these clusters can be used to identify hydrological processes. Besides hydrochemical data, the variables used for cluster analysis are pH, EC, and temperature. Hierarchical cluster analysis aims to separate groups based on measured characteristics and maximize the distance between groups (**Brown, 1998**). Cluster analysis can locate a common group of observation locations with some characteristics. The nearest linkage method resulting from data grouping is used to identify 'outliers' with chemistry different from other samples. Based on this assessment, the nearest samples are excluded so as not to cause bias. The following assessments are then grouped using Ward's linkage method, one of the most used methods because it mini-

mizes group variance. The results of HCA classification are visualized as a dendrogram representing the differences between objects (systems) by the heights of connected branches. The dendrogram is made based on the variables T, EC, pH, and hydrochemical parameters. The complete linkage of Euclidean distance, the distance between two points, is used, i.e. the maximum dissimilarity between objects within a cluster, which is considered. The significant Euclidean distance was used using the complete linkage method to derive clusters from standardized data from the dry and rainy seasons.

The Principal Component Analysis (PCA) technique was used in this study to explain the inter-correlated hydrochemical parameters and their association (**Brown, 1998**). A principal component analysis (PCA) using Varimax with Kaiser Normalizations was conducted to determine the chemical-physical parameters of water sources (**Kaiser, 1958**). Prior to PCA, the Bartlett sphericity test and KMO test were performed to assess the adequacy of sampling and the validity. The KMO is an index for comparing the magnitude of the observed correlation coefficient and the partial correlation coefficient (**Brown, 1998**). The dataset adequacy for PCA analysis was calculated using the Kaiser - Meyer - Olkin Test (KMO) and Bartlett's test; a KMO value of 0.5 is considered the smallest acceptable value (**Kaiser, 1958; Likas et al., 2003**). The PCA technique was carried out to reduce the dimensions of the dataset by finding correlations between different variables and hydrochemical parameters (**Brown, 1998; Kim et al., 2005; Kopić et al., 2016; Solano et al., 2023**). The multivariate dataset is used in PCA to create a covariance matrix and obtain its eigenvalues and eigenvectors. The principal components are extracted by projecting the initial variables onto the eigenvectors (also known as loadings), and their variance is expressed in the eigenvalues. The uncorrelated variables in the dataset are represented by the principal components, which are orthogonal to one another (**Helena et al., 2000; Singh et al., 2004**). The PCA transforms a set of potentially correlated variables into a set of linearly uncorrelated variable values called principal components, separating the uncorrelated variables. The correlation is the degree of relationship between two scaled quantities and is always between 1 and -1 (**Brown, 1998**), categorized as strong (>0.75), moderate

(0.50-0.75), and weak (<0.50) (**Brown, 1998; Tanasković et al., 2012; Kale et al., 2021**).

A preliminary step to avoid differences in variance ranges and misclassification from using different units of measurement, parameter analysis results (fields and laboratory data) require standardization to a z-scale as shown in **Equation 2** (**Brown, 1998; Jollife & Cadima, 2016; Trabelsi & Zouari, 2019**). Based on the standardized data, R-mode FA, including PCA, was carried out to reveal the interaction among the variables.

$$Z_i = \frac{X_i - X_m}{S_i} \qquad (2)$$

where Z_i is a standardized variable, X_i is the raw variable, X_m is the mean, and S_i is the standard deviation.

The study was conducted based on field and laboratory observation results, and water chemical analysis was carried out, analyzed, and modelled using SPSS v.26.

3. Results and discussion

3.1. Hydrochemical characteristics

A box and whisker plot are used to determine the dominance of the elements, the dominance of the major cations and anions, as shown in **Figure 2**. The abundance of the hydrochemical analyses results of all samples is reflected in the boxplot, the anions $HCO_3^- > Cl^- > SO_4^{2-} > CO_3^-$, and the cations $Ca^{2+} > Na^+ > Mg^{2+} > K^+$, respectively (see **Figures 2a**, **2b**, and **2c**).

Statistical summary of the result calculated physico-chemical parameter values for each season, as shown in **Table 2**. In the observations of water temperatures, there are generally no changes in the dry or rainy seasons. Water temperatures taken in different seasons ranged from 23.5-51.8°C (dry), between 24.7 - 54.6°C (rainy), and between 23.5 – 54.6 (two seasons). The pH ranged between 5.7 - 7.7 (mean 6.7) and 4.9 - 8.2 (mean 6.7), respectively, and 4.9 – 8.2 during dry, rainy, and two seasons. The average pH measurements result in the two seasons almost did not change, by only 0.1 (see **Table 2**). A change of 0.1 increased from pH 6.6 to 6.7 occurred in the dry season at locations 1R, 2R, 16W, 6S, 31S, and 26S. Meanwhile, pH increased during the rainy
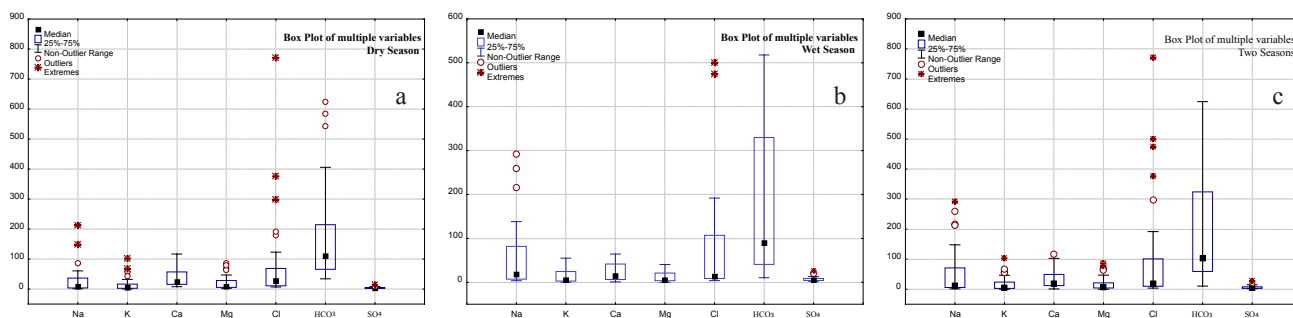

Figure 2: Boxplot major ion at dry season (a), rainy season (b), and two seasons (c)
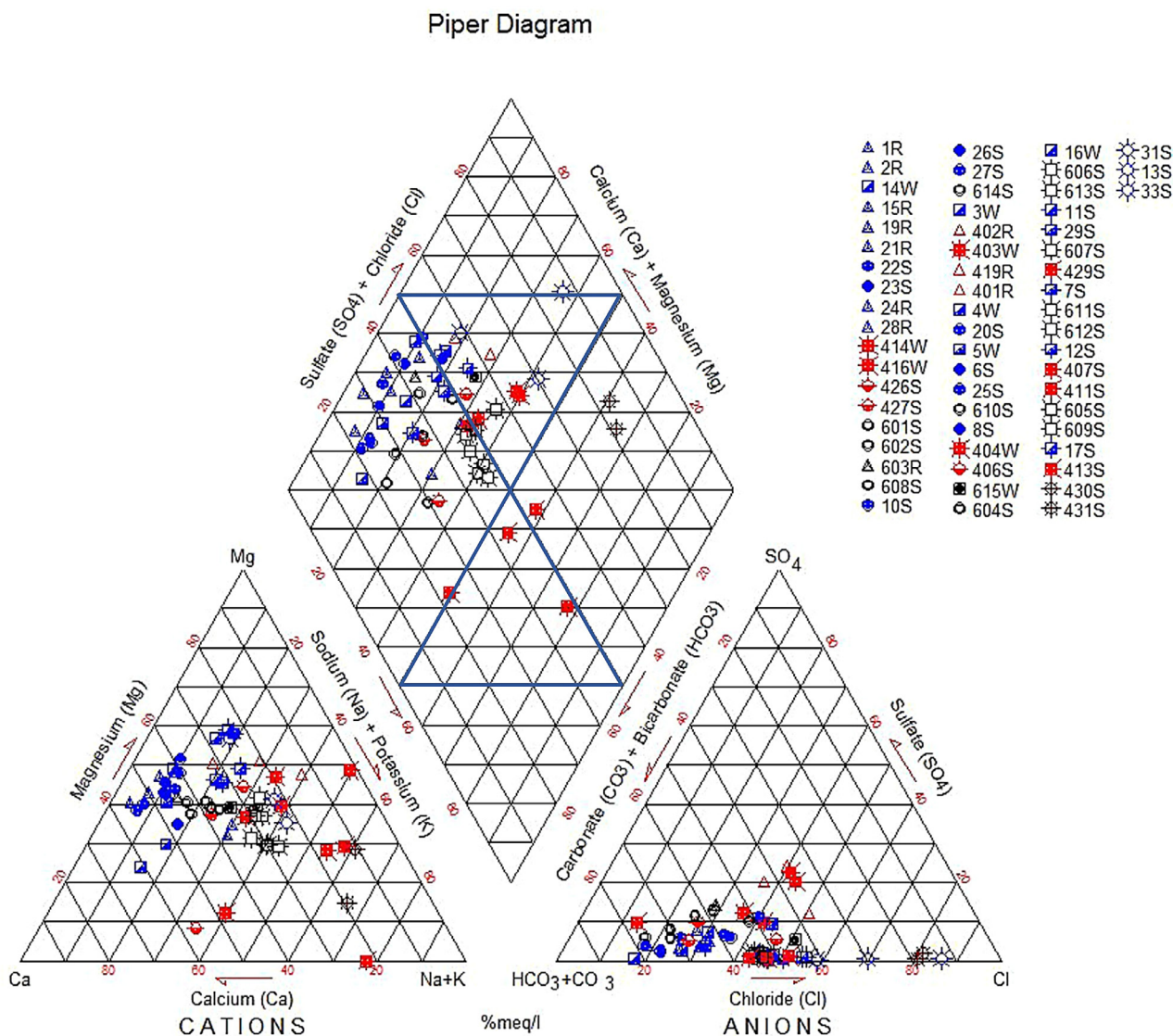
Figure 3: Piper diagram showing major ion compositions of river (R), well (W), spring (S) at Rawadanau. All data in meq/L normalized to 100%.

season at locations 29S, 24R, and 4W. Changes in pH may be associated with the dissolution of volcanic rocks (**Solano et al., 2023**). Electrical conductivity (EC) is between 46.0 - 2850.0 μS/cm (dry) and 31.5 – 4460.0 μS/cm during dry and rainy seasons, respectively. The EC range reflects the variety of groundwater originating from hot springs, cold springs, and rivers influenced by rainwater. Meanwhile, EC changes tend to rise in the dry season and occur at hot springs and wells (16W, 14W, 3W, and 4W). The increase in the dissolved element is probably the evaporation of water in the dry season, causing an increase in the value of EC (**Appelo & Postma, 2005**).

The plot in the Piper diagram produces five dominant water types: Ca-HCO$_3$ (N=32), Na-HCO$_3$ (N=14), Mg-HCO$_3$ (N=6), Na-Ca-HCO$_3$ (N=4) and Na-Cl (N=4), respectively (see **Figure 3**). The Ca-HCO$_3$ water type pointed to the dilution effect of rain or surface water in-

filtration (**Trabelsi & Zouari, 2019**). Meanwhile, the dominance of the calcium-magnesium bicarbonate type reflects the geological conditions of the study area, which is composed of volcanic rocks. The interaction of groundwater with rocks increases alkalinity (**Hartanto et al., 2022**). The ions exchange between Na$^+$+K$^+$ ions in water with Ca^{+2} and Mg^{+2} during the water in the rock. The accumulation of Na$^+$ and K$^+$ in groundwater could result from the albite and K-Feldspar decomposition in igneous rocks; this indicates the ion exchange occurs between Ca^{2+} or Mg^{2+} and Na$^+$ in the aquifer (**Appelo & Postma, 2005; Hartanto et al., 2022**).

The stable isotope of water (δ^{18}O and δ^{2}H), as shown in **Figure 4**, determine the Rawadanau area that all of the water sources originate from meteoric water (**Hartanto et al., 2022**). The local meteoric water line (LMWL) refers to the Padarincang area (**Djiono et al., 2011**), and the global meteoric water line (GMWL)
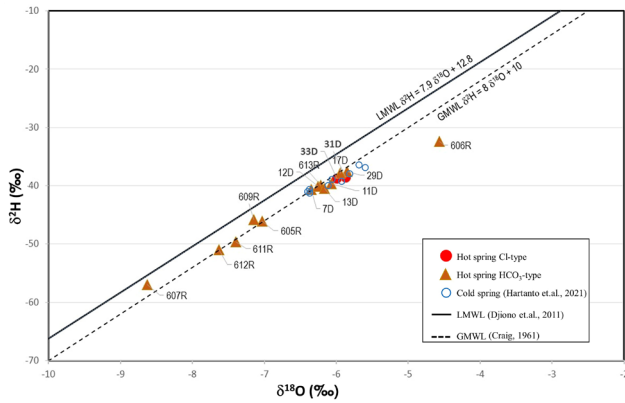
Figure 4: Plots of $\delta^{18}O$ versus δ^2H values of water source in the study area (**Craig, 1961; Djiono et al., 2011; Hartanto et al., 2022**)

(**Craig, 1961; Hartanto et al., 2022**) found that water samples taken in two seasons did not show different results. Nearly all water sources lie on the GMWL indicating a meteoric origin. The samples taken were not similar to LMWL (**Djiono et al., 2011**). The 606R mix-type water sample probably confirms the stable isotope condition affected by the geothermal steam of the contains of juvenile water (**Craig, 1963**) or kinetic fractionation effects during precipitation (**Bagheri et al., 2020**). Stable isotopes record sources and processes and can play an important role in hydrogeological investigations. Thermal springs usually contain CO_2 from geothermal systems, making the HCO_3-type heavier, probably due to the interaction of oxygen from meteoric water with CO_2 (**Bagheri et al., 2020; Nyakairu et al., 2023**). Also, the cold springs have heavier isotopic values than thermal waters (see **Figure 4**). The elevation of the recharge zone, evaporation, vapour separation, and mixing are the main processes that can increase the $\delta^{18}O$ in low-temperature resources (**Craig, 1963**; **Bagheri et al., 2020**).

3.2. Hierarchical cluster analysis

The purpose of a hierarchical cluster analysis is to separate groups based on measured characteristics and maximize the distance between groups (**Brown, 1998**). Cluster analysis can locate a common group of observation locations with some characteristics. A dendrogram of the clusters produced by the analysis is shown in **Figure 5a**, while the distribution map is shown in **Figure 5b**. Based on the variables temperature, electrical conductivity (EC), and chemical concentrations, two main clusters were identified, including clusters A (subclusters A1 and A2) and B (see **Figure 5a**). The significant Euclidean distance was used using the complete linkage method to derive three clusters from dry and rainy season standardized data.

Cluster A

The water samples from cluster A (A1 and A2) are characterized by high K^+/Na^+ and HCO_3^-/Cl^- ratios (see

Table 2: Descriptive analysis of 60 groundwater samples at dry (n=29), rainy (n=31), two seasons (n=60)

	Symbol	Unit		Dry season					Rainy season				Two seasons		
		N	Max	Min	Mean	Std. Dev	N	Max	Min	Mean	Std. Dev	N	Max	Min	Mean	Std. Dev
Z	meter	29	296.5	23.1	132.1	59.3	31	296.5	23.1	125.3	48.0	60	296.5	23.1	128.5	53.3
T	°C	29	51.8	23.5	32.0	8.5	31	54.6	24.7	33.4	8.2	60	54.6	23.5	32.8	8.3
pH		29	7.7	5.7	6.6	0.5	31	8.2	4.9	6.7	0.7	60	8.2	4.9	6.7	0.6
EC	µS/cm	29	2850.0	46.0	695.1	854.4	31	4460.0	31.5	903.9	1074.3	60	4460.0	31.5	804.8	973.6
TDS	mg/l	29	1852.5	29.9	451.8	555.4	31	2899.0	20.5	587.6	698.3	60	2899.0	20.5	523.1	632.8
Na^+	mg/l	29	212.2	1.1	30.6	48.5	31	292.0	3.6	59.8	77.3	60	292.0	1.1	46.0	66.3
K^+	mg/l	29	102.4	0.6	15.9	24.3	31	55.0	0.4	13.5	14.4	60	102.4	0.4	14.6	19.6
Ca^{2+}	mg/l	29	116.8	8.0	38.0	29.8	31	64.4	1.0	24.6	20.8	60	116.8	1.0	30.9	26.1
Mg^{2+}	mg/l	29	86.7	2.5	21.5	25.3	31	40.6	0.0	12.0	11.5	60	86.7	0.0	16.5	19.7
Cl^-	mg/l	29	771.0	6.9	88.8	162.3	31	500.6	3.8	77.7	121.8	60	771.0	3.8	82.9	141.3
HCO_3^-	mg/l	29	624.7	34.3	186.7	170.6	31	517.9	10.4	189.1	164.9	60	624.7	10.4	187.9	166.2
CO_3^-	mg/l	29	58.9	0.0	6.5	15.1	31	0.0	0.0	0.0	0.0	60	58.9	0.0	3.1	10.8
SO_4^{2-}	mg/l	29	14.9	2.1	4.6	3.7	31	26.9	2.4	7.5	5.2	60	26.9	2.1	6.1	4.8
Fe_{total}	mg/l	29	4.4	0.0	0.6	1.1	15	0.9	0.0	0.3	0.3	43	4.4	0.0	0.5	0.9
Mn	mg/l	29	0.9	0.0	0.1	0.2	15	0.5	0.0	0.2	0.2	43	0.9	0.0	0.1	0.2
SiO_2	mg/l	29	157.1	34.2	83.5	36.8	15	167.9	40.2	117.1	47.2	43	167.9	34.2	95.2	43.3
Σ Cat	mg/l	29	510.1	14.2	106.0	121.6	31	420.4	6.7	109.9	115.7	60	510.1	6.7	108.0	117.5
Σ An	mg/l	29	1027.9	43.5	286.6	286.8	31	824.8	18.4	274.2	256.8	60	1027.9	18.4	280.1	269.2
σ	%	29	4.3	-5.0	-3.9	1.7	31	2.8	-5.0	-3.1	2.1	60	4.3	-5.3	-3.4	2.0
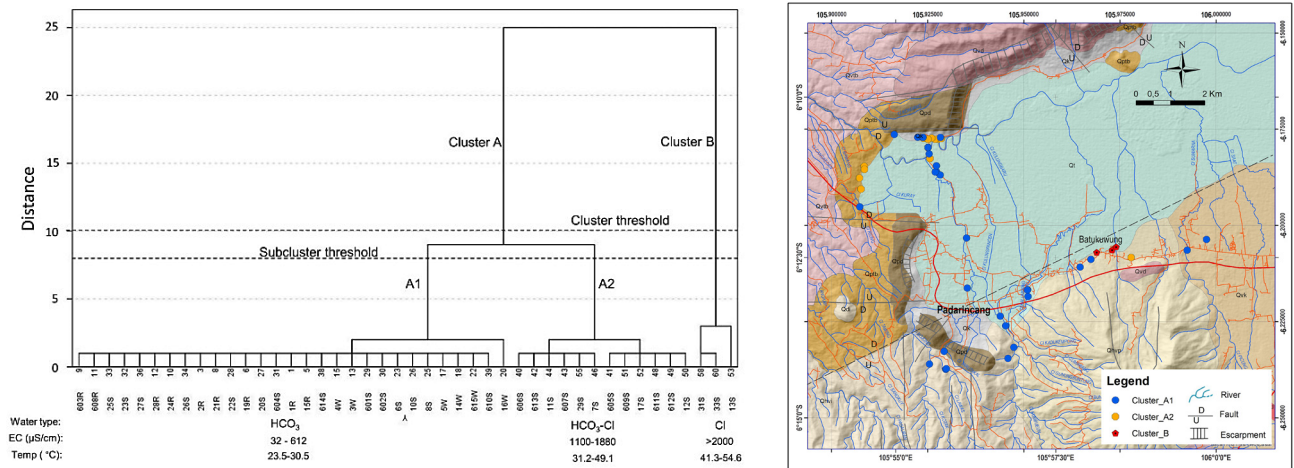
Figure 5: The dendrogram for the HCA shows two clusters and three subclusters (see **Figure 5a**) and a cluster distribution map (see **Figure 5b**). The dendrogram analyses 60 water samples, including 12 river samples, 10 well samples, and 38 spring samples. Geological information refers to **Figure 1**.
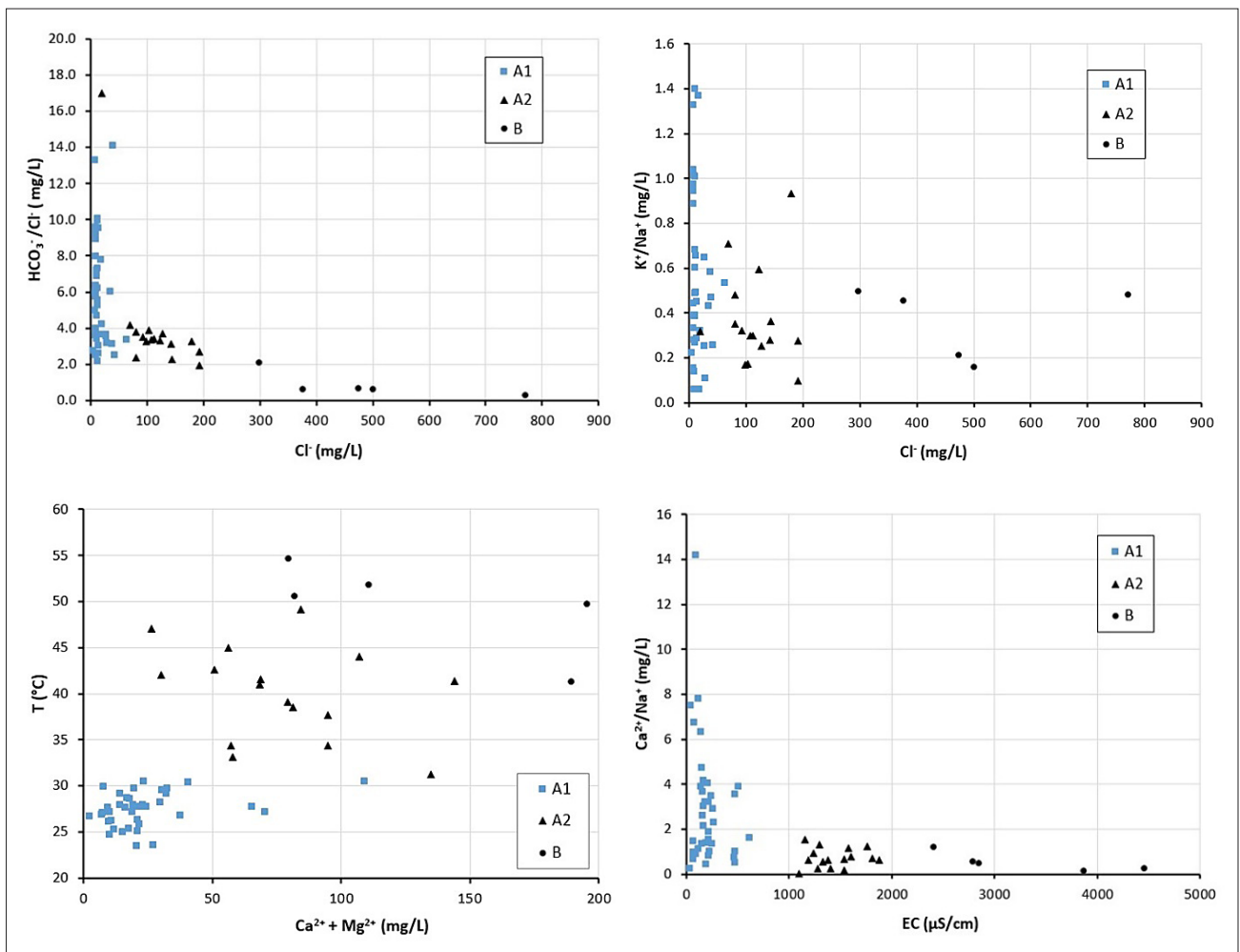

Figure 6: Bivariate plots of ion concentrations (mg/L unless otherwise stated) for water samples collected from Rawadanau: 6a. [Cl$^-$ versus (HCO$_3^-$/Cl$^-$)], 6b. [Cl$^-$ versus (K$^+$/Na$^+$)], 6c. [(Ca^{2+}+Mg^{2+}) versus T] and 6d. [EC versus (Ca^{2+}/Na$^+$)].

Figures 6a and **6b**). Subcluster A1 consists of surface water, river, dug well, and cold spring samples. This cluster is characterized by a temperature <30 °C and EC <600 µS/cm and a low total of Ca^{2+}+Mg^{2+} but a relatively high Ca^{2+}/Na$^+$ ratio (see **Figures 6c** and **6d**). This sub-cluster water type is bicarbonate, characterized by low silica,

Table 3: Correlations analysis groundwater samples at dry (*), rainy (**), two seasons (***)

	T	pH	EC	Na⁺	K⁺	Ca²⁺	Mg²⁺	Cl⁻	HCO₃⁻	CO₃⁻	SO₄²⁻	Fe$_{Total}$	Mn	SiO₂
T	1													
pH * pH ** pH ***	0.09 0.02 0.06	1												
EC * EC ** EC ***	**0.89** **0.90** **0.88**	-0.02 -0.02 0.02	1											
Na⁺ * Na⁺ ** Na⁺ ***	**0.80** **0.90** **0.81**	-0.07 -0.02 -0.01	0.93 0.99 0.94	1										
K⁺ * K⁺ ** K⁺ ***	**0.84** **0.90** **0.85**	-0.05 0.02 -0.03	**0.92** **0.98** **0.90**	0.97 0.98 0.93	1									
Ca²⁺ * Ca²⁺ ** Ca²⁺ ***	**0.77** **0.89** **0.79**	-0.03 0.01 -0.03	**0.90** **0.99** **0.91**	0.86 0.98 0.86	0.86 0.97 0.87	1								
Mg²⁺ * Mg²⁺ ** Mg²⁺ ***	*0.70* **0.85** *0.70*	0.00 -0.05 -0.04	**0.88** **0.97** **0.84**	**0.80** **0.94** **0.76**	**0.82** **0.89** **0.82**	**0.94** **0.96** **0.92**	1							
Cl⁻ * Cl⁻ ** Cl⁻ ***	*0.73* **0.89** *0.73*	-0.10 -0.03 -0.09	**0.88** **0.98** **0.84**	**0.98** **0.97** **0.92**	**0.94** **0.97** **0.94**	**0.83** **0.99** **0.82**	**0.78** **0.94** **0.79**	1						
HCO₃⁻ * HCO₃⁻ ** HCO₃⁻ ***	*0.57* **0.89** *0.63*	0.08 0.00 0.09	*0.67* **0.99** **0.76**	0.49 **0.99** *0.62*	*0.55* **0.96** *0.60*	**0.81** **0.98** **0.83**	**0.86** **0.97** **0.81**	0.41 **0.96** 0.44	1					
CO₃⁻ * CO₃⁻ ** CO₃⁻ ***	*0.62* -0.08 *0.55*	-0.02 -0.22 -0.09	*0.68* 0.09 *0.54*	*0.69* 0.06 *0.54*	*0.60* -0.06 *0.57*	*0.50* 0.09 0.45	0.45 0.18 0.46	*0.62* 0.03 *0.61*	0.18 0.09 0.12	1				
SO₄⁻² * SO₄⁻² ** SO₄⁻² ***	0.31 0.47 0.24	-0.10 0.00 -0.04	0.17 0.21 0.17	0.23 0.25 0.21	0.22 0.28 0.15	0.27 0.16 0.01	0.03 0.20 0.01	0.22 0.17 0.16	-0.02 0.26 0.02	0.32 -0.23 0.21	1			
Fe$_{Total}$ * Fe$_{Total}$ ** Fe$_{Total}$ ***	0.01 *0.55* 0.04	-0.08 -0.28 -0.08	0.02 *0.63* 0.02	0.07 *0.60* 0.07	0.01 *0.58* 0.03	0.16 *0.63* 0.16	0.14 *0.66* 0.16	-0.03 *0.67* -0.02	0.33 *0.60* 0.28	0.01 0.14 0.03	0.07 0.12 0.02	1		
Mn * Mn ** Mn ***	0.21 **0.78** 0.27	-0.08 0.38 -0.05	0.40 **0.81** 0.47	0.26 **0.81** 0.37	0.23 **0.84** 0.25	0.46 **0.82** 0.45	*0.52* **0.75** 0.44	0.25 **0.78** 0.24	*0.56* **0.83** *0.57*	-0.06 -0.26 -0.13	-0.11 0.18 0.04	0.22 0.27 0.15	1	
SiO₂ * SiO₂ ** SiO₂ ***	**0.88** -0.20 *0.74*	-0.03 **-0.79** 0.26	**0.83** -0.24 **0.78**	*0.74* -0.23 *0.73*	0.79 -0.25 *0.66*	*0.73* -0.27 *0.64*	*0.72* -0.22 *0.53*	*0.68* -0.20 *0.53*	*0.59* -0.26 *0.65*	*0.60* -0.07 0.30	0.10 -0.01 0.09	-0.09 0.30 -0.07	0.17 *-0.56* 0.31	1

'Strong' correlation coefficient > 0.75 (bold, statistically significant at 0.05 level).
'Moderate' correlation coefficient 0.75–0.50. (bold italic, statistically significant at the 0.05 level).
'Weak' correllation coefficient 0.50–0.36.

Fe$_{total}$, and Mn content, <100 mg/l, <0.56 mg/L, and <0.06 mg/L, respectively. This cluster of groundwater may result from mixing with rainwater and recharging with topsoil and quick flow (**Hartanto et al., 2022**). The high and significant correlation of Na⁺ and K⁺ (see **Table 3**) indicates the possibility of the weathering of albite and K-feldspar (orthoclase and microcline) contributing Na⁺ and K⁺ ions to groundwater (**Appelo & Postma, 2005; Hartanto et al., 2022**). Sub-Cluster A1 is relatively fresh water, EC < 600 µS/cm, typical for bicarbonate-type water, the main elements containing HCO₃⁻, Ca²⁺, and Mg²⁺. Sub-cluster A2 contains all hot springs, with temperatures between 33 and 49°C, and EC < 1880 µS/cm, respectively. The water type in this sub-cluster is a mix-type characterized by silica content (100 to 163 mg/L), Fe$_{total}$ (<91 mg/L), and Mn <0.41 mg/L. Although hot springs, this type of water flows up, mixes with shallow groundwater, and indicates water from an unconfined aquifer system (**Hartanto et al., 2022**). The two sub-clusters (A1 and A2) are clearly distinguished by the tem-

perature and EC factors that characterize groundwater sources. Subcluster A1 has samples from rivers, cold springs, and wells; subcluster A2 is related to water from hot springs. The EC can be seen as a measure of salinity and chloride between cations and anions, as shown in **Figure 6** (**6a**, **6b**, and **6d**), and is strongly to very strongly positive correlated (see **Table 3**).

Cluster B

Samples assigned to cluster B are characterized by low K^+/Na^+ and HCO_3^-/Cl^- and high Cl^- content (see **Figures 6a** and **6b**). Relatively high Cl^- concentrations, Fe_{total}, Mn, temperatures between 41 – 55 °C, and high EC between 2410 to 4460 µS/cm indicate water from a confined aquifer system (**Hartanto et al., 2022**). Cluster B water samples are type chloride hot springs in two seasons. As shown in **Figure 6c**, cluster B has a high total $Ca^{2+}+Mg^{2+}$, while the Ca^{2+}/Na^+ ratio is low (see **Figure 6d**).

The correlation matrix between variables in dry, rainy, and two seasons is presented in **Table 3**. In every season, almost all variables have strong correlations. A strong positive correlation (0.75) was shown by T, EC, Na^+, K^+, Ca^{2+}, and Mg^{2+}, with other variables, indicating that these influence each other. The strong positive correlation between the Mn and Fe_{total} variables and all other variables indicates the dissolution of soil and weathering of volcanic rocks during the rainy season (see **Table 3**).

Meanwhile, variable Silica (SiO_2) strongly correlates during the dry season. Variables that do not offer a cor-

relation include pH and sulphate (SO_4^{2-}) in **Table 3**. The negative and weak correlation between pH and all variables suggests the increasing concentration of other ions and the aggressiveness of the acidic medium towards the soil and host rock (**Helena et al., 2000**).

The strong positive correlation in dry, rainy, and two seasons shown in the variable temperature with all variables except pH and sulphate (SO_4^{2-}). Temperature has positive correlation with CO_3 (dry), Fe_{total} (rainy), Mn (rainy), and silica (SiO_2) (dry). Strong positive correlations were shown for EC in dry, rainy, and two seasons with temperature, Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^-, HCO_3^-, CO_3^- (dry), Fe_{total} (rainy), Mn (rainy), and silica (dry). A strong positive correlation is shown by the variables sodium (Na^+) with K^+, Ca^{2+}, Mg^{2+}, Cl^-, HCO_3^- (except dry), CO_3^- (except rainy), Fe_{total} (dry), Mn (dry), and silica (except rainy). Calcium and magnesium ions are also positively correlated with each other, suggesting that rock weathering is the main source of cations (**Kale et al., 2021**).

3.3. Principal component analysis

The Kaiser-Meyer-Olkin test (KMO) and Barlett test results are interpreted according to the guidelines (**Kaiser, 1958**). PCA analysis applies hydrochemical data taken from both seasons; it can be grouped into three main components: hydrochemical data (type of water), the stable isotopes of water ($\delta^{18}O$ and δ^2H) and indications of meteoric origin and locations (longitude and elevation). The value of 0.698 indicates the adequacy of the data for PCA, as represented in **Table 4**. The Bartlett's test of sphericity results (chi-square statistic = 829.568 for 78° of freedom; p<0.001) shows that there is enough correlation between the variables in the dataset. Furthermore, the KMO test (0.698) indicates that the variables are correlating with one to another, suggesting that PCA can be used to minimize the dataset's dimensionality. The results of the analyzed parameters, from 13 components, are about 86.12% of the existing variables, resulting in three main components that affect the parameters. Based on the eigenvalue greater than or

Table 4: Kaiser–Meyer–Olkin and Barlett's test

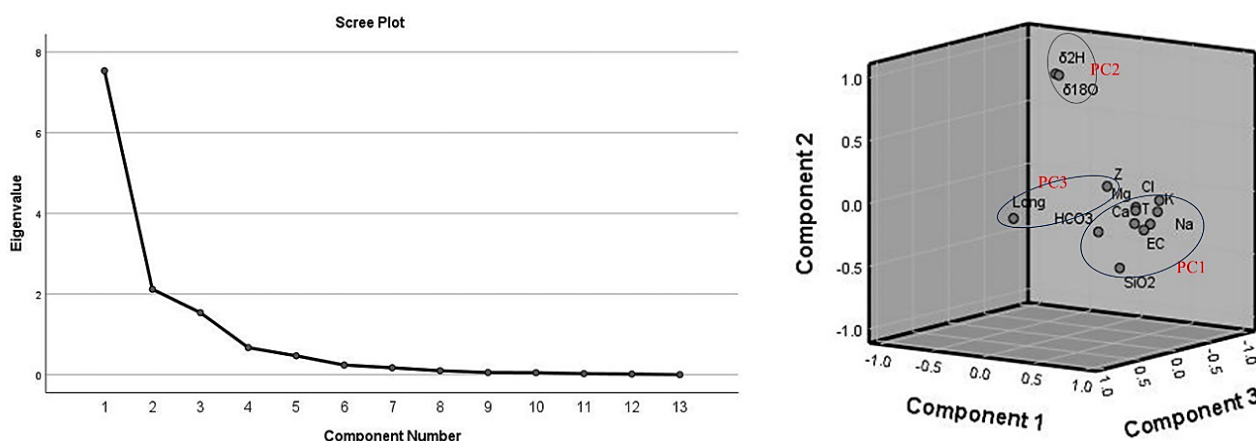| KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.698
Bartlett's Test of Sphericity	Approx. Chi-Square	829.568
	df	78
	Sig.	0

Figure 7: Scree plots (a) and rotated space (b) of the variance represented by each principal component for each data matrix.

Table 5: Eigenvalue and the number of variances in some component.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cum. %	Total	% of Variance	Cum. %	Total	% of Variance	Cum. %
1	7.537	57.98	57.98	7.54	57.98	57.98	6.91	53.17	53.14
2	2.119	16.30	74.28	2.12	16.30	74.28	2.20	16.87	70.04
3	1.539	11.84	86.12	1.54	11.84	86.12	2.09	16.08	86.12
4	0.672	5.17	91.29						
5	0.477	3.62	94.90						
6	0.238	1.83	96.73						
7	0.172	1.33	98.06						
8	0.10	0.77	98.83						
9	0.056	0.43	99.25						
10	0.051	0.39	99.64						
11	0.027	0.21	99.85						
12	0.017	0.13	99.98						
13	0.02	0.02	100						

Extraction Method: Principal Component Analysis.

equal to 1.0, the components were selected to represent the hydrogeochemical processes of the groundwater without the loss of significant information (**Brown, 1998; Helena et al., 2000**). The results of the analyzed parameters, from 13 components, are about 86.12% of the existing variables, resulting in three main components that affect the parameters. The results of the PCA of the initial variables and eigenvalues are presented in **Table 5**. All the variables were also plotted in rotated space, and a scree plot representing all the PCs that were extracted during PCA is shown in **Figure 7** (**7a & 7b**). In comparison, the respective factor relationships were categorized into strong (>0.75), moderate (0.50-0.75), and weak (<0.50) (**Brown, 1998, Tanasković et al., 2012, Kale et al., 2021**).

Table 6 shows that the PC1, the highest loading of 57.98%, has a strong positive correlation (>0.75) to the physical parameters of water chemistry, which includes water temperature (T), EC, Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^-, HCO_3^-, and SiO_2, as shown in **Table 3**. The results of the analysis suggest that hydrochemistry, which is PC1, is the main component and characteristic of the Rawadanau area, which probably reflects the tropics. Although it has a weak correlation (0.3 - 0.5) to elevation (Z) and longitude, it also influences hydrochemistry. This condition is typical of areas composed of volcanic rocks resulting from the dissolution of minerals that make up igneous rocks (**Helena et al., 2000**). Major ions include Na^+, K^+, Ca^{2+}, and Mg^{2+}, probably resulting from the release of albite minerals, K-feldspar, which are the main constituents of rocks in the study area (**Hartanto et al., 2022**). This main component produces water types of $Ca-HCO_3$ (N=32), $Na-HCO_3$ (N=14), $Mg-HCO_3$ (N=6), and $Na-Ca-HCO_3$, which are characteristic of areas composed of volcanic rocks. PC1 is dominated by the bicar-

bonate type, where the constituent cations are influenced by various minerals, depending on the source rock.

The second component matrix (PC2), 16.30%, has a strong positive correlation (>0.75), with the stable isotopes of water ($\delta^{18}O$ and δ^2H). They are indicating meteoric water origin, the main source of water in Rawadanau. Meanwhile, a weak negative correlation (0.3-0.5) with silica (SiO_2) indicates it is slightly related to thermal water. Finally, the third component matrix (PC3), with an amount of 11.84%, has a strong positive and negative correlation (>0.75) with the sampling position, that is, elevation and longitude (see **Table 6**). The water chemistry of the upland is strongly influenced by vol-

Table 6: Component matrix and results of rotation between physico-chemical parameters of groundwater.

Variables	Component Matrix			Rotated Component Matrix		
	1	2	3	1	2	3
Long	0.439	0.152	**-0.818**	0.17	-0.017	**0.925**
Elevation	-0.407	-0.21	0.751	-0.165	-0.053	**-0.863**
T	**0.885**	0.012	0.057	**0.856**	-0.091	0.212
EC	**0.973**	-0.023	0.082	**0.945**	-0.134	0.209
Na	**0.937**	-0.008	0.174	**0.941**	-0.101	0.115
K	**0.928**	0.073	0.259	**0.965**	-0.01	0.046
Ca	**0.948**	0.146	0.026	**0.917**	0.028	0.284
Mg	**0.885**	0.144	0.076	**0.872**	0.041	0.217
Cl	**0.857**	0.129	0.334	**0.925**	0.065	-0.036
HCO₃	**0.8**	-0.006	-0.243	0.68	-0.139	0.466
SiO₂	**0.78**	-0.368	-0.028	**0.701**	-0.463	0.198
δ¹⁸O	-0.118	**0.966**	0.097	-0.007	**0.977**	0.036
δ²H	-0.177	**0.959**	0.107	-0.06	**0.979**	0.008

canic rocks, unlike alluvial, which is the morphology of the lowlands. The components affecting this PC3 include volcanic rocks consisting of Basal-Andesitic, Andesitic, and Andesitic-Lavas (**Hartanto et al., 2022**). The recharge area is 440 to 500 m asl (**Djiono et al., 2011**), hence distance and the groundwater flow enrich the minerals from rocks and soil.

The positive correlation between Na^+ and Cl^- indicates that they are influenced by the deep geothermal waters found in the Rawadanau area (**Hartanto et al., 2022**), probably due to rock salt and sylvite; this can occur during crystallization from dissolved liquids (**Scott, 1987**). In addition, the strong positive correlation between Ca^{2+}, Mg^{2+}, and Na^+ revealed the possibility of ion exchange processes from the dissolution of silicate minerals. Hence, it is believable that the dissolution of silicate minerals dominated the anions and cations of water in the Rawadanau area. Additionally, a moderate to strong positive correlation shown by Fe_{total} and Mn during the rainy season indicated the dissolution of soil (**Helena et al., 2000**).

4. Conclusion

This study demonstrates the utility of multivariate analysis (HCA and PCA), a powerful tool in interpreting hydrogeochemical data to evaluate the main factors influencing groundwater composition. The study area has complex geographical characteristics, and many factors influence groundwater chemistry. The sequence of the dominant major cations and anions is $HCO_3^- > Cl^- > SO_4^{2-} > CO_3^-$, and $Ca^{2+} > Na^+ > Mg^{2+} > K^+$, respectively. The composition of groundwater in the Rawadanau Basin is influenced by various factors that produce two groups: a bicarbonate type of unconfined aquifer and a chloride type of confined aquifer. The types of water produced can be grouped into Ca-HCO_3, Na-HCO_3, Mg-HCO_3, Na-Ca-HCO_3, and Na-Cl hydrogeochemical facies. The weathering of silicate minerals controls the chemistry of the major ions, i.e. sodium, potassium, calcium, and magnesium. The increase in the concentration of iron in groundwater in the area is probably due to the oxidation of pyrite and the leaching of iron-carrying minerals that occur during the rainy season. The results obtained from multivariate analyses indicate that the parameters responsible for groundwater quality variations are mainly related to the presence and dissolution of rocks, natural processes, thermal water and water-rock interaction in mixed water type, and water-rock interactions under the influence of thermal water.

Three major groups were identified by the HCA method applied to 60 groundwater samples based on major ions. Cluster A1 is influenced by meteoric water and shallow groundwater. The contribution of thermal water interacting with shallow groundwater (unconfined aquifer) is reflected in cluster A2. The two subclusters (A1 and A2) differ significantly in the temperature and EC factors characterizing the groundwater sources. The contribution of deep thermal water is reflected in cluster B, characterized by relatively high Cl^-, Fe_{total}, Mn, temperatures, and high EC. The results of the PCA, the initial variables and eigenvalues accounts for 86.12% of the 13 influencing variables, and the Barlett sphericity test and KMO test are about 0.689. PC1 has a high positive value for most of the physicochemical composition, reflecting the type of water influenced by its constituent rocks and geological factors; this also reflects the research area, a tropical region. The PC2 is a positive value for the stable isotopes of water ($\delta^{18}O$ and δ^2H), indicating a water origin. Meteoric water is a factor that characterizes the filling and water sources in Rawadanau. The PC3 has a high positive charge for elevation and a high negative charge for longitude, which indicates the location of the water recharge and source, which is influenced by geological factors and constituent rocks and is associated with geothermal and volcanic rock influences. As groundwater flows to the surface (producing subcluster A2), the main factor governing groundwater quality gradually changes; water interacts between the unconfined aquifer and meteoric water.

5. References

Alfaro, C., Wallace, M., (1994): Origin and classification of springs and historical review with current applications. *Environ. Geol.* 24, 112–124. https://doi.org/10.1007/BF00 767884

APHA. (2005): Standard Methods for the Examination of Water and Wastewater. APHA. 2005. *Standard Methods for the Examination of Water and Wastewater. 21st Ed. American Public Health Association, Washington DC, 1220p.* https://doi.org/10.2105/SMWW.2882.216

Appelo, C. A. J., & Postma, D. (2005): Geochemistry, groundwater and pollution. 2nd. *Ed. Balkema, Rotterdam.*

Bagheri, R., Karami, G. H., Jafari, H., Eggenkamp, H., & Shamsi, A. (2020): Isotope hydrology and geothermometry of the thermal springs, Damavand volcanic region, Iran. *Journal of Volcanology and Geothermal Research*, *389*, 106745.

Blake, S., Henry, T., Murray, J., Flood, R., Muller, M. R., Jones, A. G., & Rath, V. (2016): Compositional multivariate statistical analysis of thermal groundwater provenance: a hydrogeochemical case study from Ireland. *Applied Geochemistry*, *75*, 171–188.

Brown, C. E. (1998): *Applied Multivariate Statistics in Geohydrology and Related Sciences [electronic resource]*. 247.

Craig, H. (1961): Isotopic variations in meteoric waters. *Science*, *133*(3465), 1702–1703. https://doi.org/10.1126/science.133.3465.1702

Craig, H. (1963): The isotopic geochemistry of water and carbon in geothermal areas. *Nuclear geology of geothermal areas*.

Djiono, D., Abidin, A., Sidauruk, P., Satrio, S., Pratikno, B., & Prasetio, R. (2011): Application of Isotope and Radiation Technology Development in Water Resources Management in Banten. *Scientific Seminar on Isotope and Radiation Research Results, BATAN*, 21p.

Hamed, Y., (2014): Stable isotope ratios in meteoric waters in El Kef Region, Northwestern Tunisia: implications for changes of moisture sources. J. Earth Sci. Clim. Change 5. https://doi.org/10.4172/2157-7617.1000203.

Hartanto, P., Alam, B. Y. C. S. S. S., Lubis, R. F., & Hendarmawan, H. (2021): The Origin and Quality of the Groundwater of the Rawadanau Basin in Serang Banten, Indonesia. *Rudarsko-Geološko-Naftni Zbornik (The Mining-Geological-Petroleum Bulletin), 36*(2 SE-Geology), 11–24. https://doi.org/10.17794/rgn.2021.2.2

Hartanto, P., Alam, B. Y. C. S. S. S., Lubis, R. F., Ismawan, I., Iskandarsyah, T. Y. W. M., Sendjaja, Y. A., & Hendarmawan, H. (2022): The application of hydrogeochemical and stable isotope data to decipher the origin and evolution of hot springs in the Rawadanau Basin, Indonesia. *Geothermics*, *105*, 102506. https://doi.org/10.1016/j.geothermics.2022.102506

Hartanto, P., Delinom, R. M., & Hendarmawan, H. (2019): Type and Quality of Water at the Peak of Dry Season Based on the Major Elements of Water Chemical at Rawa Danau Serang Regency. *RISET Indonesian Journal of Geology and Mining*, *29*(1), 13–25. https://doi.org/10.14203/risetgeotam2019.v29.1021

Hartmann, J., Berner, Z., Stüben, D., & Henze, N. (2005): A statistical procedure for the analysis of seismotectonically induced hydrochemical signals: a case study from the Eastern Carpathians, Romania. *Tectonophysics*, *405*(1–4), 77–98.

Helena, B., Pardo, R., Vega, M., Barrado, E., Fernandez, J. M., & Fernandez, L. (2000): Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *Water Research*, *34*(3), 807–816.

Hem, J. D. (1985): *Study and interpretation of the chemical characteristics of natural water* (Vol. 2254). Department of the Interior, US Geological Survey.

Jolliffe, I. T., & Cadima, J. (2016): Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Kaiser, H. F. (1958): The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*(3), 187–200.

Kale, A., Bandela, N., Kulkarni, J., Sahoo, S. K., & Kumar, A. (2021): Hydrogeochemistry and multivariate statistical analysis of groundwater quality of hard rock aquifers from Deccan trap basalt in Western India. *Environmental Earth Sciences*, *80*, 1–24.

Kim, J., Kim, R., Lee, J., Cheong, T., Yum, B., & Chang, H. (2005): Multivariate statistical analysis to identify the major factors governing groundwater quality in the coastal area of Kimje, South Korea. *Hydrological Processes: An International Journal*, *19*(6), 1261–1276.

Kopić, J., Loborec, J., & Nakić, Z. (2016): Hydrogeological and hydrogeochemical characteristics of a wider area of the regional well field Eastern Slavonia – Sikirevci. *Rudarsko-Geološko-Naftni Zbornik*, *31*(3 SE-Geology). https://hrcak.srce.hr/ojs/index.php/rgn/article/view/4092

Lee, S. M., Min, K. D., Woo, N. C., Kim, Y. J., & Ahn, C. H. (2003): Statistical models for the assessment of nitrate contamination in urban groundwater using GIS. *Environmental Geology*, *44*(2), 210–221. https://doi.org/10.1007/s00254-002-0747-0

Likas, A., Vlassis, N., & Verbeek, J. J. (2003): The global k-means clustering algorithm. *Pattern Recognition*, *36*(2), 451–461.

Nyakairu, G. W. A., Kasule, J., Ouma, O., & Bahati, G. (2023): Origin and hydrogeochemical formation processes of geothermal fluids from the Kibiro area, Western Uganda. *Applied Geochemistry*, *152*, 105648.

Rusmana, E., Suwitodirdjo, K., & Suharsono. (1991): . *Geology of the Serang Quadrangle, Jawa, Scale 1: 100,000, explanatory notes and geological map. Geological Research and Development Centre. Bandung.*

Santosa, S., Sutrisno, T., Turkandi, T., Ratman, N., & Sukanta, U. (1991): Geology of the Anyer Quadrangle, West Jawa, 1:100.000, explanatory notes and geological map. Geological Research and Development Centre. Bandung, Indonesia. In *Geological Agency of Indonesia, Bandung, (unpublished reports)* (Vol. 1, Issue 100,000).

Scheiber, L., Jurado, A., Pujades, E., Criollo, R., & Suñé, E. V. (2023): Applied multivariate statistical analysis as a tool for assessing groundwater reactions in the Niebla-Posadas aquifer, Spain. *Hydrogeology Journal*, *31*(2), 521–536. https://doi.org/10.1007/s10040-022-02580-8

Scott, G. L. (1987): The origin of chlorine in the fluids of the Tongonan geothermal system, Philippines. *Proceedings, 9th New Zealand Geothermal Workshop*, 92–95.

Singh, K. P., Malik, A., Mohan, D., & Sinha, S. (2004): Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) – a case study. *Water Research*, *38*(18), 3980–3992.

Solano, C., Vargas-Azofeifa, I., Castillo-Muñoz, R., & RP, S. H. (2023): Multivariate data analysis applied to groundwater geochemical characterization, Central Pacific, Costa Rica. *Applied Geochemistry*, *151*, 105599.

Sudhakaran, S., Mahadevan, H., Arun, V., Krishnakumar, A. P., & Krishnan, K. A. (2020): A multivariate statistical approach in assessing the quality of potable and irrigation water environs of the Netravati River basin (India). *Groundwater for Sustainable Development*, *11*, 100462.

Suryadarma, & Fauzi, A. (1991): Hydrothermal Alteration of The Garung Banten Geothermal Area West Java. *Proc.13th*

New Zealand Geothermal Workshop, 193–197. https://www.geothermal-energy.org/pdf/IGAstandard/NZGW/1991/Suryadarma.pdf

Suryaman, M. (1999). Hydrogeological Map of Indonesia, Scale 1:100.000. In *Directorate of Environmental Geology, Bandung. Indonesia. (unpublished reports)*.

Syah, B. Y. C. S. S., Itoi, R., Taguchi, S., Saibi, H., & Yamashiro, R. (2019): Hydrogeochemical and isotope characterization of geothermal waters from the Cidanau geothermal field, West Java, Indonesia. *Geothermics*, *78*, 62–69.

Tanasković, I., Golobocanin, D., & Miljević, N. (2012): Multivariate statistical analysis of hydrochemical and radio-

logical data of Serbian spa waters. *Journal of Geochemical Exploration*, *112*, 226–234.

Trabelsi, R., & Zouari, K. (2019): Coupled geochemical modeling and multivariate statistical analysis approach for the assessment of groundwater quality in irrigated areas: A study from North Eastern of Tunisia. *Groundwater for Sustainable Development*, *8*, 413–427.

Van Bemmelen, R. W. (1949): The Geology of Indonesia. *Government Printing Office, Den Haag*, *Vol I, IA*, 732 p.

Yidana, S. M., Ophori, D., & Banoeng-Yakubo, B. (2008): A multivariate statistical analysis of surface water chemistry data – The Ankobra Basin, Ghana. *Journal of Environmental Management*, *86*(1), 80–87.

SAŽETAK

Multivarijatna analiza podataka u svrhu hidrokemijske karakterizacije podzemne vode u slivu Rawadanau, Banten, Indonezija

Multivarijantna analiza glavnih komponenata (PCA) i hijerarhijska klasterska analiza (HCA) primijenjene su za identifikaciju i klasifikaciju različitih izvora vode koji se pojavljuju u slijevu Rawadanau. Podatci uključuju 60 uzoraka uzetih tijekom sušne (29 uzoraka) i kišne sezone (31 uzorak) u tropskim regijama. Izvori zahvaćene vode obuhvaćaju iskopane zdence, rijeke te hladne i tople izvore. Korišteni kemijski parametri vode uključuju parametre izmjerene na terenu (T, pH, EC), koncentracije glavnih iona (Na$^+$, K$^+$, Ca^{2+}, Mg^{2+}, Cl$^-$, HCO$_3^-$, SO$_4^{2-}$), SiO$_2$, Fe$_{ukupno}$, Mn i stabilne izotope kisika i vodika u vodi (δ^{2}H i δ^{18}O). Korelacija glavnih iona pokazuje promjene u kišnoj sezoni u odnosu na Fe$_{ukupno}$ i Mn koncentracije. Interpretacija temeljena na HCA-u definirala je dva klastera. Klaster A uključuje otvoreni vodonosnik bikarbonatnoga tipa. S druge strane, klaster B predstavlja zatvoreni vodonosnik kloridnoga tipa čiji se kemijski sastav ne mijenja u različitim godišnjim dobima. Rezultati PCA-a pokazali su da trokomponentnu matricu čini 86,12 % strukture podataka koji opisuju izvore vode u slijevu Rawadanau na koje utječu vulkanske stijene i koji su u snažnoj korelaciji s ionima Na$^+$, K$^+$, Ca^{2+} i Mg^{2+}. PC1 ima visoku pozitivnu vrijednost za hidrokemijski sastav vode, što upućuje na snažan litološki utjecaj na kemiju vode. PC2 ima pozitivnu vrijednost za stabilne izotope (δ^{18}O i δ^{2}H), što upućuje na glavni izvor vode u Rawadanauu. PC3 ima pozitivan predznak za nadmorsku visinu te negativan za zemljopisnu dužinu, što upućuje na područje prihranjivanja, na koje utječu geološki čimbenici, što je dodatno povezano s geotermalnim utjecajima i vulkanskim stijenama. Ova multivarijantna analiza omogućuje identifikaciju komponenti i klastera hidrokemijskih varijabli koje nisu utvrđene u prethodnim studijama.

Ključne riječi:
HCA, PCA, kemija vode, Rawadanau, vulkanske stijene

Authors' contribution

Priyo Hartanto (Dr, senior researcher, hydrogeology) was in charge of the manuscript writing, sampling, measurements, and data analysis and gave possible conclusions for this research. **Rachmat Fajar Lubis** (Dr.Sci., senior researcher, hydrogeology) was in charge of the manuscript writing and participated in the sampling, measurements and data analyses. **Boy Yoseph CSS Syah Alam** (Dr., associate professor, hydrochemical) was in charge of hydrochemical data analyses. **Yoga Adriana Sendjaja** (Dr., associate professor, geochemist) was in charge of geochemical data analyses. **Ismawan Ismawan** (Dr., associate professor, geological structure) was in charge of geology structure data analyses. **Tengku Yan WM Iskandarsyah** (Dr., associate professor, applied geology) was in charge of geology environmental data analyses. **Hendarmawan Hendarmawan** (Professor of Hydrogeology) was responsible for the isotope analyses and overall manuscript structure.