KJMS 2022 VOL.4 (1): 95-105

KDU Journal of Multidisciplinary Studies (KJMS)
Volume 4. Issue 1 July 2022
DOI: http://doi.org/10.4038/kjms.v4i1.43

ASSESSMENT OF GROUNDWATER QUALITY USING MULTIVARIATE STATISTICAL METHODS AT NAGADEEPA AREA IN BADULLA DISTRICT, SRI LANKA

Senadeera K.P.G.W[1], Piyasiri S[1 & 2], Nandalal K.D.W.[3], Chandrajith R[4], Rantunga K[1],

Department of Zoology, Faculty of Applied Sciences
University of Sri Jayewardenepura, Sri Lanka[1]
[2]Faculty of Graduate Studies, Sir John Kotelawala Defence University of Sri Lanka[2]
Department of Civil Engineering, Faculty of Engineering[3]
University of Peradeniya, Sri Lanka
Department of Geology, Faculty of Science[4],
University of Peradeniya, Sri Lanka

ABSTRACT

The study area, Nagadeepa is situated in Badulla District within the dryer parts of the country. Groundwater is the main drinking water source in the area. The study was carried out from 2013 to 2016 and the main objective of the study was to examine the chemical status and pollution levels by examining of 25 water quality parameters namely Na^+, K^+, Mg^{2+}, Ca^{2+}, Pb Mo, Cr, Cd, Mg, Hg, As, Al, Zn, Se Cl^-, F^-, NO_3^-, HCO_3^-, SO_4^{2-}, PO_4^{3-}, DO, pH, temperature, electrical conductivity and water hardness in 28 wells. The results of the factor analysis indicated that conductivity, alkalinity, hardness, Calcium, Fluoride, Mo, Cr and Cd have a strong impact on the water quality compared to the categories of moderate and weak loadings. The parameters namely; Mg, Na, As and Al, Zn, Se and Water Level, Se, Cl, pH and DO, Pb have a moderate impact on water quality based on the moderate loadings indicated by the PC analysis. The weak impact on water quality is indicated by Cl^-, SO_2^{2-} and K, PO_4, Hg, Se and Hg, Se, Al and Temperature, Cl^-, pH and PO_4^{3-}. Analysis indicated that; (i) significant and positive correlation with alkalinity, hardness, Ca^{2+} and Mg^{2+}, (ii) positive and significant correlation only with hardness and Ca^{2+}, (iii) Hardness shows positive and significant correlation only with Ca^{2+}, (iv) Fluoride has a positive and significant correlation with Mo^{2+} and (v) Na^+ has a positive and significant correlation only with Mg^{2+}. Cluster analysis indicated, cluster 1 with less polluted water in 9 wells, cluster 2 with moderately polluted water in 11 wells, cluster 3 with highly polluted water in 7 wells, and cluster 4 with very highly polluted water in one well. Factor analysis indicated that pH, conductivity, alkalinity, hardness, dissolved oxygen, chloride, and sulfate are the major factors that affect the quality of the water. In the majority of the sampling wells, the correlation coefficient revealed a geogenic nature. Cluster analysis revealed that 29% of wells are unsuitable for drinking purposes.

KEYWORDS*: Groundwater quality, Multivariate Statistical Methods*

Corresponding Author: Senadeera K.P.G.W Email: *wasantha.senadeera@gmail.com*

1. INTRODUCTION

Water and land are the vital resources of life that are progressively being polluted due to poor land-use management, population growth, industrialization, agricultural activities and anthropogenic impacts on the area (Olajire and Imeokparia, 2001). Groundwater is the main water source for human consumption, specifically in the dry areas where the surface water sources are unavailable. Analysis of water pollutants of groundwater is a very useful tool which provides important information for water quality management (Pazand et al., 2011). 80% of diseases in developing countries are directly related to poor drinking water (Olajire and Imeokparia, 2001).

Major chemical factors such as Na^+, K^+, Mg^{2+}, Ca^{2+}, Cl^-, F^-, NO_3^-, HCO_3^-, SO_4^{2-}, PO_4^{3-} pH values, temperature, electrical conductivity and hardness play an important role in the groundwater quality. Several parameters and their correlation are necessary to determine the water quality characteristics in the study area.

The multivariate statistical method of factor analysis is one of the tools used to identify the major factors responsible for groundwater contamination as discussed in Balasuriya et al (2021), Senanayake, Indunil et al. (2021) and Olajire and Imeokparia, 2001). In addition to that Cluster Analysis (CA), Descriptive Analysis (DA), and Principal Component Analysis (PCA) were used in this research work. These techniques can be used to obtain relationships and interactions between parameters and sampling locations, to identify the factors and sources influencing groundwater quality and to suggest useful tools for both management of water resources and monitoring of groundwater quality (Nosrati and Eekhault, 2012). CA was used to examine the spatial groupings of the sampling wells. It is a common method to classify variables into clusters (Massart and Kaufmann, 1983). CA and PCA are commonly supported by DA as a confirmation for CA and PCA, and they are usually referred to as pattern recognition techniques (Adams, 1998). The application of different pattern recognition techniques to reduce the complexity of a large data set has proven to give a better interpretation and understanding of water quality data (Brown et al., 1980). This method is widely applied to study the interrelation and correlation of variables in hydrology, geology, environmental science, etc. (Liu et al., 2003; Singh et al., 2004). Factor analysis acts as a tool that could measure and observe the variables and can reduce large data sets to fewer latent variables and share a common variance and are unobservable (Bartholomew, Knott and Moustaki, 2011). The unobservable factors are not directly measured, but it is essential to construct hypothetically to represent the variables (Cattell, 1973 cited in Yong and Pearce, 2013). Further, factor analysis is useful for easy interpretation of a large data set reducing to a smaller set by considering common key factors placing variables into meaningful categories (Rummel, 1970, cited in Yong and Pearce, 2013).

The eigenvalues and scree test (i.e., scree plot) are used to determine how many factors to retain. One criterion that can be used to determine the number of factors to retain is Kaiser's criterion, which is a rule of thumb. Therefore, it is suggested to use the scree test in conjunction with the eigenvalues to determine the number of factors to retain. The scree test consists of eigenvalues and factors (Cattell, 1978). The number of factors to be retained is the data points that are above the break (i.e., point of inflexion).

1.1 The objectives

The main objective of the present study is to examine the chemical status and pollution levels of the groundwater of Nagadeepa study area with respect to Na^+, K^+, Mg^{2+}, Ca^{2+}, Cl^-, F^-, NO_3, HCO_3^-, SO_4^{2-}, PO_4^{3-}, hardness and pH values, temperature, electrical conductivity and hence to ascertain the status of these elements in the water.

Many studies already carried out in the past exists (Chandrajith et al., 2011a; Jayasekara et al., 2013; Wanigasuriya et al., 2011; Bandara et al., 2010; Illeperuma et al., 2009; Wanigasuriya, 2007; Hittarage, 2004) on ground water quality in the study area. However, based on the multivariate statistical methods such as Factor Analysis, Cluster Analysis (CA), Descriptive Analysis (DA), and Principal Component Analysis (PCA) of such water quality parameters used in clustering the ground water sources which could be used for consumption is not available. The present study is to fill that research gap and to help people to use groundwater from clustered areas which are safer to use for consumptive purposes. The findings of this study will be significant for detecting such water sources for the people in the area to use for consumption. Particularly due to the Chronic Kidney disease with uncertain etiology (CKDu) reported in the area, this research will provide valuable information to use safer drinking water sources for consumptive purposes.

The problem statement of this research was to find out whether there are interconnected activities among the water quality parameters. The research question is to find out the suitability of the groundwater for drinking purposes. The significance of the research is important to find out the effect of such parameters and their composite effects on groundwater quality and to fill the information gap on water quality in the study area.

1.2 Study Area

The study area of Nagadeepa is situated in Badulla District, which falls within the dryer parts of the country. A small part of the Nagadeepa study area is irrigated by the Nagadeepa tank and the majority survives by receiving water from the monsoons. Groundwater is the main drinking water source for the people in that area. The Nagadeepa study area falls within 16 GN divisions in the Rideemaliyadda DS division of the Badulla district with the geographical positions of 81.072º longitude, 7.355º latitude and 81.145º longitude, 7.228º latitude and the total population is 19103 (Census, 2012), the study area receives rainfall mainly from the northeast monsoon.

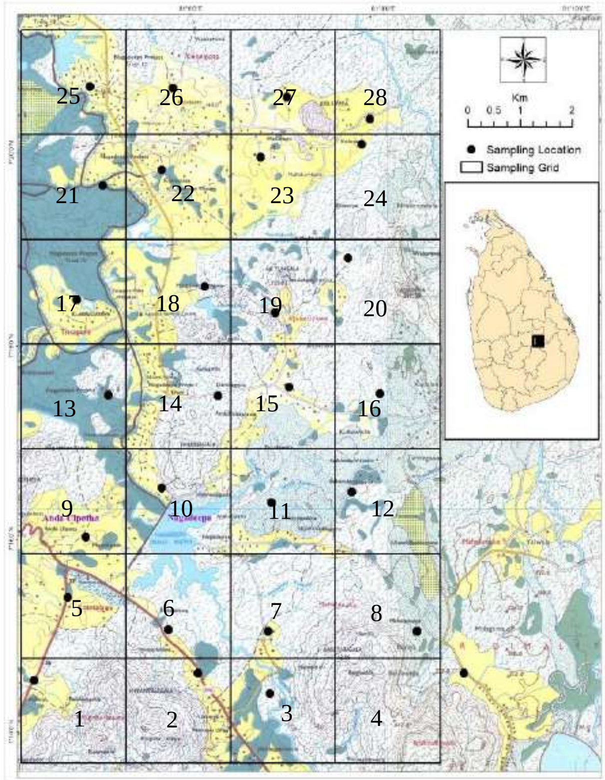

Figure. 1. Sampling locations

2. METHODOLOGY

Nagadeepa study area map (1:50,000) was divided into 28 squares to mark the approximate sampling points in each area (Fig. 1). The approximate middle point of each square was selected as the groundwater sampling location and GPS points were taken to locate the sampling points from which samples were collected.

Sample collection and storage:

The samples from dug wells were collected in two different sampling bottles. (i) 500 ml HDPE bottles to collect samples to analyze chloride, sulfate, fluoride, nitrate and phosphate. (ii) Samples were collected in 100 ml sample bottles and 2 ml trace metal nitric acid was added to each sample to analyze trace metals.

Water quality analysis:

pH, DO, Electrical Conductivity (EC), Temperature, Alkalinity and Hardness were analyzed in the field, using HACH senslON portable kits. In addition, the water levels of all the wells were noted. Hardness, chloride and alkalinity were determined using HACH digital titrator. Analyses of fluoride, nitrate, phosphate, sulfate, and iron were carried out using HACH DR 2700 Spectrophotometer.

Determination of the trace elements:

Atomic Absorption Spectrophotometer (AAS Thermo Scientific, iCE 3000 series) was used for the analysis of Ca^{+2} while Inductively Coupled Plasma Mass Spectrophotometer (ICP-MS; Thermo Fisher iCAP Q) was used for the analysis of Na^{+}, K^{+}, Mg^{+2}, Zn^{+2}, Pb^{+2}, Cd^{+2}, As, Cr, Al^{+3} Hg, Mo.

Data Analysis and Interpretation of the results:

In this study, SPSS Version 16 and Minitab Version 16 software packages were used to analyze the water quality data.

Multivariate Statistical Analysis:

Correlation Coefficient and Multivariate data analytical methods of Factor analysis and Cluster analysis were used to interpret the underlying pattern of the data sets. These techniques were used to identify different groups of variables and their relationships.

4. RESULTS & DISCUSSION

Multivariate Statistical Analysis on water quality parameters of the study area

Multivariate analysis of the statistical methods such as Factor Analysis, Cluster Analysis and Correlation Coefficient is widely used, and it is capable of analyzing a data set to get an insight into the underlying pattern of such data set. These techniques are useful to identify different groups of variables and their relationships.

In the present study, initial factor loadings were calculated by **1st step** of the Factor Analysis based on the extraction method of Principal Component Analysis (PCA). The **2nd step** of the analysis was the Factor rotation performed by the most common orthogonal method called varimax rotation to explain the number of characteristics of factors.

Table 1. Total Variance Explained by Principal Component Analysis Extraction Method in Nagadeepa

Component No.	Initial Eigenvalues		
	Eigenvalues	% of Variance	Cumulative %
1	4.122	15.852	15.852
2	3.383	13.013	28.865
3	2.855	10.982	39.847
4	2.156	8.292	48.139
5	1.634	6.286	54.426
6	1.243	4.782	59.208
7	1.151	4.427	63.635
8	.996	3.831	67.466
9	.936	3.600	71.066
10	.895	3.440	74.507
11	.820	3.155	77.661
12	.684	2.630	80.291
13	.640	2.461	82.752
14	.570	2.192	84.944
15	.553	2.126	87.070
16	.530	2.039	89.109
17	.470	1.809	90.918
18	.434	1.669	92.587
19	.389	1.497	94.084
20	.336	1.293	95.377
21	.313	1.204	96.582
22	.264	1.015	97.597
23	.198	.763	98.360
24	.173	.666	99.027
25	.131	.504	99.531
26	.122	.469	100.000

Extraction Method: Principal Component Analysis.

Table 1 illustrates the initial factor loadings which exhibit the eigenvalues, percentage of variance and cumulative percentage of variance associated in Nagadeepa study area. Twenty-six parameters in sampling locations have been considered for this analysis. The analysis generated 7 factors for the total variance. Eigenvalues greater than 1 were considered as a criterion for the extraction method of principal parameter components vital for explaining the sources of variance in the data.

As illustrated in Table 1, the results revealed that 7 factors have eigenvalue greater than 1 and these factors explain more than 63% of the total variance. Therefore, 7 factors namely; pH, conductivity, alkalinity, hardness, dissolved oxygen, chloride and sulfate are considered for explaining the major factors that affect the water quality in Nagadeepa study area. The seven factors account for 15.9, 13.0, 10.9, 8.3, 6.3, 4.8 and 4.4 of the percentage of total variance respectively. Fig. 2 illustrates the Eigenvalues vs Parameter components.

Table 2 illustrates the results of the verimax rotation. The significance of the factors classified according to Liu et al. (2003) are summarized in Table 3.

Table 3 illustrates Rotated Component Matrix in the Nagadeepa study area, in factor 1 that accounts for 15.85% of the total variance. Strong positive loading was observed in conductivity, alkalinity, hardness and calcium while it shows that the moderate loading of Mg^{2+} and weak loading of Cl, SO_2^{2-} and K^+. In factor 2, there were strong loadings in F and Mo while Na, As and Al were moderately loaded and PO_4^{3-} was weakly loaded. In the 3rd factor, there was strong loading only on Cr, moderate loading in Zn, Se and water level, and Hg was weakly loaded. Four parameters were observed in factor 4 where SO_4^{2-} and NO_3^- were with moderate loading and Se and Hg with weak loading.

Cd indicated a strong loading in factor 5 and Cl^- was moderate and Se, Al and Temperature were observed as weakly loaded.

Table 2. Variables and factor loadings after varimax rotation for the study period in Nagadeepa

Rotated Component Matrix[a]							
Parameter	1	2	3	4	5	6	7
pH	.028	-.006	-.264	-.071	.041	.690	.396
Conduc.	.911	.159	-.008	.082	-.024	.018	.088
Alkalinity	.888	.055	-.139	.047	.078	-.014	.055
Hardness	.855	-.157	-.198	.131	.234	-.079	.025
DO	-.270	.082	.205	-.128	-.078	.699	-.030
Cl	.341	-.150	-.185	.133	.521	.346	-.323
SO4	.352	.269	-.077	.674	-.071	.003	.177
F	.233	.778	-.206	-.097	.185	.085	.136
NO3	.080	-.045	.125	.671	-.087	.058	-.070
PO4	-.141	.400	-.338	.212	.100	-.021	.315
Na	.048	.686	.030	.114	-.479	.070	.093
K	.476	.019	.100	-.604	-.222	.004	.285
Ca	.753	-.283	.144	.073	-.074	-.013	-.112
Mg	.621	.209	.230	-.347	-.235	-.014	.127
Fe	-.120	-.067	-.103	-.265	-.088	-.617	.221
Zn	-.027	-.056	.637	.073	-.198	.217	.161
Pb	.129	.055	.153	-.030	.045	-.011	.715
Cd	-.016	.173	-.155	-.055	.755	-.033	.160
As	-.138	.704	.061	.030	.018	.015	.027
Cr	-.007	.119	.750	-.160	-.118	-.057	.022
Se	.008	-.088	.565	.454	.329	-.059	.069
Al	-.157	.647	.171	-.184	.316	.002	-.188
Hg	-.075	-.085	.412	.311	-.028	-.265	.140
Mo	.076	.829	.030	.035	-.030	.000	.024
Tem.	-.030	.296	-.155	.116	.327	.168	-.098
WLevel	-.059	.045	.576	.153	-.451	.158	-.296

Extraction Method: Principal Component Analysis
Rotation Method: Varimax with Kaiser Normalization

a. Rotation converged in 9 iterations

Table 3: Summary of classified factor loadings as strong, moderate and weak in Ngadeepa as indicated in Liu et al. (2003) derived from Table 2

	Factor Loadings			
Factor	Strong loading- >0.75	Moderate loading- 0.75 – 0.5	Weak loading- 0.5 – 0.3	% of Variance
1	conductivity, alkalinity, hardness, calcium	Mg	Cl, SO2 and K	15.852
2	Fluoride, Mo	Na, As and Al	PO_4	13.013
3	Cr	Zn, Se and Water Level	Hg	10.982
4	None	Se	Se and Hg	8.292
5	Cd	Cl	Se, Al and Tempera ture	6.286
6	None	pH and Do	Cl	4.782
7	None	Pb	pH and PO_4	4.427

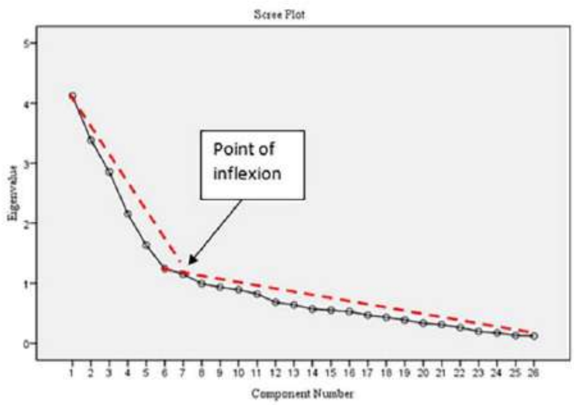

Figure: 2 Scree plot showing Eigenvalues vs. component number (parameters) in Nagadeepastudy area

In factor 6, pH and DO were with moderate loading and Cl^- was observed a weakly loading effect. Pb indicated a moderate loading in factor 7 while pH and PO_4^{3-} were weakly loading.

These results indicated that conductivity, alkalinity, hardness, calcium, Fluoride, Mo, Cr and Cd have a strong impact on the water quality of Nagadeepa study area compared to the categories of moderate and weak loadings. The parameters namely; Mg, Na, As and Al, Zn, Se and Water Level, Se, Cl, pH and DO, Pb have a moderate impact on water quality based on the moderate loadings indicated by the PC analysis. The weak impact on water quality is indicated by Cl^-, SO_2^{2-} and K, PO_4, Hg, Se and Hg, Se, Al and Temperature, Cl^-, pH and PO_4.

Correlation coefficient

The results of the correlation analysis are considered in the subsequent interpretation.

High correlation coefficient of nearly 1 or -1 expresses a good relationship between two variables, and a correlation coefficient around zero means no relationship.

The variables having a coefficient value (r) > 0.5 or < -0.5 are considered significant. Positive values indicate a positive relationship while negative values of r indicate an inverse relationship. The correlation coefficients (r) among various water quality parameters of Nagadeepa study area were calculated and the values of correlation coefficients (r) are given in Table 4.

Correlation coefficients of parameters in the study area

The results are illustrated in Table 4 that pH has no significant correlation with other parameters, but it shows a weak positive and negative association with all. Negative correlations were observed in F^-, Na^+, Mg^{2+}, Fe^{2+}, Cd, As, Cr, Se and water level. Following ions act as a negative association, NO_3^-, Na^+, Ca^{2+}, Fe^{2+}, Zn^{2+}, Cd^{2+}, Cr, Se^{2+}, Al^{3+}, Hg and water level. The results of conductivity indicated a significant and positive correlation with alkalinity, hardness, Ca^{2+} and Mg^{2+}.

Table 4 Correlation Coefficient in Nagadeepa study area –
Correlations: pH, Conduc, Alk, Hard, Do, Cl, SO4, F,

	pH	**Conduc**	**Alk**	**Hard**	**Do**	**Cl**	**SO4**	**F**
Conduc	0.241							
Alk	0.141	0.823						
Hard	0.164	0.771	0.803					
Do	0.201	-0.154	-0.182	-0.216				
Cl	0.268	0.261	0.286	0.362	0.008			
SO4	0.103	0.447	0.380	0.349	-0.003	0.007		
F	0.186	0.360	0.348	0.101	0.026	0.077	0.296	
NO3	-0.105	0.070	0.028	0.070	0.022	0.071	0.250	-.152
PO4	0.186	-0.030	-0.027	-0.102	-0.106	0.082	0.123	0.354
Na	-0.004	0.182	0.096	-0.188	0.060	-0.270	0.320	0.474
K	0.156	0.410	0.341	0.260	-0.029	-0.074	-0.148	0.154
Ca	-0.007	0.605	0.535	0.627	-0.116	0.227	0.208	-.215
Mg	0.055	0.537	0.499	0.390	-0.087	-0.067	0.046	0.326
Fe	-0.104	-0.152	-0.148	-0.148	-0.025	-0.136	-0.061	-.103
Zn	-0.000	-0.050	-0.149	-0.216	0.108	-0.172	-0.015	-.073
Pb	0.109	0.129	0.107	0.072	-0.037	-0.106	0.099	0.102
Cd	-0.285	-0.072	0.115	0.086	0.050	-0.065	0.051	0.092
As	0.047	0.011	-0.095	-0.150	0.104	-0.165	0.089	0.341
Cr	-0.085	0.032	-0.080	-0.126	0.122	-0.207	-0.006	-.104
Se	-0.082	0.030	-0.031	0.012	0.001	0.144	0.208	-.141
Al	-0.026	-0.072	-0.094	-0.194	0.147	0.006	0.019	0.415
Hg	-0.163	-0.036	-0.069	-0.076	-0.066	-0.121	0.107	-.141
Mo	0.011	0.165	0.080	-0.086	0.045	-0.153	0.239	0.587
Tem	0.159	0.063	0.078	0.002	-0.049	0.227	0.022	0.234
Water L	-0.153	-0.044	-0.048	-0.097	0.225	-0.117	0.081	-.162

	NO3	**PO4**	**Na**	**K**	**Ca**	**Mg**	**Fe**	**Zn**
PO4	0.021							
Na	0.018	0.190						
K	-0.258	-0.060	0.111					
Ca	0.081	-0.288	-0.188	0.258				
Mg	-0.065	-0.037	0.249	0.596	0.262			
Fe	0.042	-0.006	-0.092	0.181	-0.093	-0.046		
Zn	0.037	-0.058	0.093	0.063	-0.028	0.071	-0.027	
Pb	-0.040	0.085	0.041	0.162	0.115	0.096	-0.005	0.107
Cd	0.018	-0.072	0.061	-0.094	-0.201	-0.020	0.069	-.035
As	0.089	0.247	0.323	-0.018	-0.244	0.055	-0.046	.016
Cr	0.075	-0.162	0.069	0.169	0.025	0.170	0.032	.237
Se	0.247	-0.051	-0.086	-0.191	0.108	-0.098	-0.103	0.292
Al	-0.113	0.151	0.243	-0.064	-0.226	-0.064	0.022	0.014
Hg	0.056	-0.090	0.053	-0.130	-0.014	-0.026	0.013	0.151
Mo	-0.066	0.219	0.561	0.004	-0.050	0.182	-0.111	-.072
Tem	-0.008	0.182	0.045	-0.123	-0.169	-0.036	-0.118	-.101
Water L	0.243	-0.169	0.206	-0.093	0.079	0.158	-0.033	0.313

	Pb	**Cd**	**As**	**Cr**	**Se**	**Al**	**Hg**	**Mo**
Cd	-0.001							
As	0.034	-0.042						
Cr	0.112	-0.215	0.054					
Se	0.055	0.009	-0.023	0.104				
Al	-0.024	0.225	0.449	0.165	-0.029			
Hg	0.023	-0.123	-0.037	0.207	0.297	-0.090		
Mo	0.149	0.124	0.539	0.079	-0.015	0.410	-0.055	
Tem	0.060	-0.203	0.073	-0.069	0.112	0.163	-0.101	0.233
Water L	-0.088	0.062	0.006	0.383	0.119	0.004	0.181	0.049

	Tem
Water L	-0.172

Cell Contents: Pearson correlation

It indicated a negative correlation with Do, phosphate, Fe^{2+}, Zn^{2+}, Cd^{2+}, Al^{3+}, Hg and water level and others were in weak positive correlation. The result of measured alkalinity in the study area shows a positive and significant correlation only with hardness and Ca^{2+} and a negative correlation with Do, phosphate, Fe^{2+}, Zn^{2+}, As, Cr^{2+}, Se^{2+}, Al^{3+}, Hg and water level, and other 12 ions indicated weak positive correlations. Total hardness has a positive and significant correlation only with Ca^{2+} but it shows a negative correlation with 11 ions of Do, phosphate, Na^{+}, Fe^{2+}, Zn^{2+}, As, Cr^{2+}, Al^{3+}, Hg, Mo^{2+} and water level. The rest of the ions were in a weak positive correlation. Dissolved Oxygen (DO) has no positive and significant correlation with measured parameters. But it indicated a negative correlation with sulfate, phosphate, K^{+}, Ca^{2+}, Mg^{2+}, Fe^{2+}, Pb^{2+}, Hg and temperature and others show a weak positive correlation. It was observed that chloride does not indicate with significant correlation with any parameter but indicated a negative correlation with Na^{+}, K^{+}, Mg^{2+}, Fe^{2+}, Zn^{2+}, Pb^{2+}, Cd^{2+}, As, Cr^{2+}, Al^{3+}, Hg and water level, and others indicated weak positive correlation. Sulfate (SO_4^{-2}) revealed that there is no significant correlation with any parameter but there is a negative indication with K^{+}, Fe^{2+}, Zn^{2+} and Cr^{2+}. Others indicated a weak positive correlation. Fluoride (F^{-}) has a positive and significant correlation with Mo^{2+} and Na^{+} and Al^{3+} indicating marginal positive and significant values of 0.474 and 0.415 respectively. That indicates a weak positive correlation with others. Nitrate (NO_3^{-}) has only a weak positive correlation with some parameters, and a negative correlation with K^{+}, Mg^{2+}, Pb^{2+}, Al^{3+}, Mo^{2+} and temperature. Phosphate (PO_4^{3-}) has no positive and significant correlation and it has a weak correlation with some parameters. It has a negative correlation with K^{+}, Ca^{2+}, Mg^{2+}, Fe^{2+}, Zn^{2+}, Cd^{2+}, Se^{2+}, Hg and water level. Sodium (Na^{+}) has a positive and significant correlation only with Mg^{2+} and a negative correlation with Ca^{2+}, Fe^{2+} and Se. That indicates a weak positive correlation with other parameters. Potassium (K^{+}) has a positive and significant correlation with Mg^{2+}; a negative correlation with Cd^{2+}, As, Al^{3+}, Hg, temperature and water level and a weak positive correlation with other parameters in Nagadeepa. The result of Calcium (Ca^{+2}) has a negative correlation with Fe^{2+}, Zn^{2+}, Cd^{2+}, As, Al^{3+}, Hg, Mo and temperature, and weak positive correlation with other parameters. Magnesium (Mg^{+2}) revealed a negative correlation with Fe^{2+}, Cd^{2+}, Se^{2+}, Al^{3+}, Hg and temperature; and a weak positive correlation with other parameters. Iron (Fe^{2+}) has a negative correlation with Zn^{2+}, Pb^{2+}, As, Se^{2+}, Mo, temperature and water level; and a weak positive correlation with other parameters in Nagadeepa. Zinc (Zn) in the area has a negative correlation with Cd^{2+}, Mo and temperature and a weak correlation with other parameters. Lead (Pb) indicates a negative correlation with Cd^{2+}, Al^{3+}, water level and a weak positive correlation with As, Cr^{2+}, Se^{2+}, Hg, Mo and temperature. The result of Cadmium (Cd) revealed a negative correlation with As, Cr^{2+}, Hg and temperature. Arsenic shows a negative correlation with Se^{2+} and Hg and a weak correlation with Cr^{2+}, Al^{3+}, Mo, temperature and water level. Chromium (Cr^{2+}) has a negative correlation only with temperature and a weak positive correlation with Se^{2+}, Al^{3+}, Hg, Mo and water level. Selenium (Se) has a negative correlation with Al^{3+} and Mo and a weak correlation with Hg, temperature and water level. Al^{3+} has a negative correlation with Hg and a weak positive correlation with Mo, temperature and water level in Nagadeepa. Mercury (Hg) has a negative correlation with Mo and temperature and a weak positive correlation with the water level. Molybdenum (Mo) has a weak positive correlation with temperature and water level. The temperature has a weak positive correlation with the water level in the study area.

Cluster Analysis

Q-mode Cluster Analysis was applied to the water quality data set to evaluate the spatial variability among the monitoring wells (sampling locations). This analysis resulted in the grouping of monitoring wells into four groups in the study area as shown in Fig. 3. This analysis is a useful method for grouping groundwater wells into homogenous groups according to their water quality. The wells were grouped into four cluster categories namely; less polluted, moderately polluted, highly polluted and as very highly polluted.

Hierarchical cluster analysis

Dendrogram using Ward Method

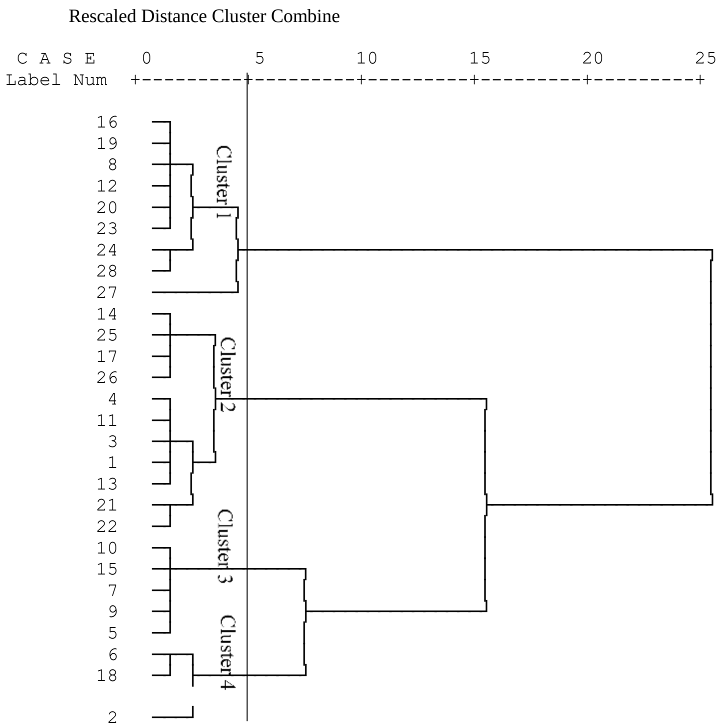

Figure. 3 Dendrogram of Cluster analysis in Nagadeepa study area

The results obtained exhibit; cluster 1 includes well location numbers; 8, 12, 16, 19, 20, 23, 24, 27 and 28 classified as less polluted. Cluster 2 includes well location Nos. of 1, 3, 4, 11, 13, 14, 17, 21, 22, 25 and 26 classified as moderately polluted, cluster 3 consists of sampling Nos. 5, 6, 7, 9, 10, 15 and 18 classified as highly polluted and cluster 4 includes only one well with location No. 2 as very highly polluted. The water quality categories based on the cluster analysis revealed that the marginal to very high are distributed in the study area. However, the good quality water category was absent.

5. CONCLUSIONS

The study exhibits the application of multivariate statistical methods in assessing the hydro geochemical characteristics in the area. Factor analysis revealed that 7 factors of pH, Conductivity, Alkalinity, Hardness, Dissolved Oxygen, Chloride and Sulfate are the major factors that affect the water quality.

The correlation coefficient (r) among various water quality parameters of Nagadeepa study area indicated positive and significant correlation between Conductivity, Alkalinity, Hardness, Sulfate, Na^+, K^+ and Fluoride with other parameters of Ca^{2+}, Mg^{2+}, Mo^{2+} and Po_4^{3-} in most of sampling wells indicating their geogenic nature.

According to the cluster analysis, 32% wells in the study area indicated less polluted waters. 39% represented moderate type of water quality, and the remaining 29% indicated polluted conditions water that is unsuitable for drinking purposes.

The geological and soil conditions of the area and the usage of chemical fertilizers and agrochemicals are the major deterioration factors since some wells are in paddy fields or close to paddy fields.

6. ACKNOWLEDGMENT

The financial assistance provided by the World Class University Grant, University of Sri Jayewardenepura, Sri Lanka under the research grant: PhD/10/2012 for a period of 3-years is acknowledged.

7. REFERENCES

Adam M. J (1998). The principle of multivariate data analysis in P. R Ashurst & M. J Dennis (Eds). Analytical methods of food authentication p. 350. Blackie Academic professional.

Balasooriya B. M. J. K., Chaminda G. G. T., Weragoda, S. K Champika Ellawala.Kankanamge, Tomonori Kawakami (2021): Assessment of Groundwater Quality in Sri Lanka Using Multivariate Statistical Techniques, Published in: Contaminants in Drinking and Wastewater Sources, Springer

Bandara J M R S, Wijewardena H V P, Liyanage J, Upul M A, Bandara J M U A,.(2010). Chronic renal failure in Sri Lanka caused by elevated dietary cadmium: Trojen horse of the green revolution. Toxicology Letters:198(1): pp. 33-39.

Bartholomew, D.J., Knott, M. and Moustaki, I., (2011). *Latent variable models and factor analysis: A unified approach* vol. 904 John Wiley & Sons.

Brown, S. D, Skogerboe, R. K, & Kowalski, B. R. (1980). Pattern recognition assessment of water quality data: Coal strip mine drainage. Chemosphere, 9: pp. 265-276. doi: 10.1016/ 0045-6535(80)90003-X.

Cattell, R.B. (1973). Factor analysis. Westport, CT: Greenwood Press.

Cattell, R. B. (1978). The scientific use of factor analysis in behavioral and life sciences. New York: Plenum Press.

Chandrajith R, Nanayakkara S, Itai K, Aturaliya T N C, Dissanayake C B, Abeysekara T, Harada K, Watanabe T, Koizumi A, (2011a), Chronic kidney disease of uncertain etiology (CKDue) in Sri Lanka: geographic distribution and environmental implications, Environmental Geochemical Health:33: pp. 267-278.

Hittarage A (2004)., Chronic Renal Disease in north Central Province of Sri Lanka. Anuradhapura Medical Journal, Sri Lanka, pp. 3-5.

Ileperuma O A, Dharmagunawardene H A, Herath K P R P, (2009), Dissolution of aluminium from sub-standard utensils under high fluoride stress: a possible risk factor for Chronic renal failure in the North-Central Province, Journal of National Science Foundation, Sri Lanka, 37 (3) : pp. 219-222.

Jayasekara J M K B, Dissanayake D M, Adhikari S B, Bandara P.,(2013) Geographical distribution of chronic kidney disease of unknown origin in North Central Region of Sri Lanka, Ceylon Medical journal:vol 58: pp. 6- 10.

Liu, C. W., Lin, K. H., and Kuo, Y. M. (2003). Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. *Science of the Total Environment*, *313*(1): pp. 77-89.

Massart, D. L., & Kaufman, L. (1983), the interpretation of chemical data by the use of cluster analysis. Wiley.

Nosrati, K., & Van Den Eeckhaut, M. (2012). Assessment of groundwater quality using multivariate statistical techniques in Hashtgerd Plain, Iran. *Environmental Earth Sciences*, *65*(1): pp. 331-344.

Olajire, A. A., & Imeokparia, F. E. (2001). Water quality assessment of Osun River: studies on inorganic nutrients. *Environmental monitoring and assessment*, *69*(1): pp. 17-28.

Pazand, K., Hezarkhani, A., Ataei, M., & Ghanbari, Y. (2011). Application of multifractal modeling technique in systematic geochemical stream sediment survey to identify copper anomalies: a case study from Ahar, Azarbaijan, Northwest Iran. *Geochemistry*, *71*(4): pp. 397-402.

Rummel, R.J. (1970). Applied factor analysis. Evanston, IL: Northwestern University Press.

Senanayke, Indunil; Piyasiri, Swarna; Chandrajith, Rohana; Nandalal, Wasantha; Ranatunga, Kamal (2021), Assessment of Groundwater Quality Using Multivariate Statistical Analysis in the Medawachchiya Area and the Huruluwewa Areas in Anuradhapura District, KDU Journal of Multidisciplinary studies, 03(1):,: pp. 80-93.

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): pp. 3980-3992.

Wanigasuriya K, Peiris-John, Wickremasinghe R, Hittarage A, (2007), Chronic Renal Failure in North Central Province of Sri Lanka; an environmentally induced disease, Trans R Soc Trap Med Hyg, 101: pp. 1013-1017.

Wanigasuriya K P, Peiris-John J, et al, (2011), Chronic Kidney Disease of uncertain aetiology in Sri Lanka: is cadmium a likely cause?, Biomedcenter Nephrology, 12: p.32

Yong, A.G. and Pearce, S., (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, *9*(2): pp.79-94.