IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

Research paper Volume 12, Iss 1, 2023

Statistical Examination of Coastal Groundwater Quality through Multiple Linear Regression: A Comprehensive Analysis

Sunandana Reddy Machireddy[1]

[1] *Rajeev Gandhi Memorial College of Engineering & Technology (Autonomous), Nandyal, AP, India*

Abstract

Groundwater, a vital natural resource for the survival of ecosystems and human existence, necessitates careful management and protection for the benefit of future generations. This study involved the examination of groundwater from 41 wells in the research area, assessing the impact of various chemical compositions. Utilizing the water quality index, the study evaluated the suitability of the samples for drinking purposes. Employing a regression model, the research aimed to establish connections between multiple chemical characteristics influencing groundwater quality, resulting in an empirical relationship derived from the selected data. Correlation analysis and scatter plots revealed a strong association, especially during premonsoon and postmonsoon seasons, between parameters like electrical conductivity, total dissolved solids, and other variables. Detailed descriptive analysis presented comprehensive information on the lowest, maximum, average, and changes in 11 key water quality metrics.

Keywords: Groundwater, Natural ecosystems, Human existence, Regression model

1. Introduction

Groundwater is an essential natural resource for sustaining natural ecosystems. Groundwater has grown into an important natural resource to meet all freshwater needs like domestic, agricultural, and industrial demands countries [C. Singaraja et al. 2017; Dipankar Saha et al. 2019]. India is one of the top groundwater extraction countries in the world. [Anirbid Sircar et al. 2021]. Groundwater quality assessment [G. Gyananath et al. 2001] will provide knowledge about the existing state of groundwater quality in a particular area. A Water Quality Index (WQI) is a measuring system to measure the quality of water in a certain area [Charan Kumar et al. 2019]. It provides a simple and quantitative technique for assessing water quality based on a set of parameters such as physical, chemical, and biological characteristics of water.

Application of statistical methods [Simpson et al. (2001)] to groundwater studies can help to know the current condition of the groundwater and also to adopt suitable strategies to improve the resource. manage natural studies if the samples are significantly different from each other and also with respect to the time changes [Charan Kumar et al.2020]. The General Linear

ISSN PRINT 2319 1775 Online 2320 7876
Research paper Volume 12, Iss 1, 2023

Model (GLM) is a statistical application [Gary et al. (2006)] that enables the investigation of the interrelationships between one or more independent variables and one or more dependent variables. Based on this, we can identify utility in water quality analysis to explore the linkages between various water quality parameters with respect to time, location, and land use. Regression analysis comes in a variety of forms, including basic linear regression, multiple linear regression, and nonlinear regression, all of these methods can be used for the study of water quality analysis. Here an attempt is made to examine the relationship between several explanatory and a dependent variables. By applying these methods, identification and quantification of these relationships can identify easily to develop more effective strategies for protecting and improving water quality.

A scatterplot matrix provides valuable perceptions into the relationships between different water quality parameters, thereby aiding the identification of potential sources of pollution or other water quality concerns. By scrutinizing these relationships, we can formulate strategies to enhance water quality and preserve our precious water resources.

2. Study area

To carry out this study, the ground water quality samples have been collected from Visakhapatnam which is positioned along the east coast of Andhra Pradesh in India at latitude $17^{o}45^{1}$ North and longitude $83^{o}16^{1}$ East. The study area and locations of the wells are shown in Figure 1. The rapid growth of population, industry, and agricultural practice has increased the significant diversion of surface water. To meet the requirements, dependence on groundwater is increasing. The exploitation of groundwater leads to the scarcity & resulting in the deterioration of groundwater quality. A total of 41 groundwater well samples from three Mandals, i.e., Bheemunipatnam, Visakhapatnam rural, and Visakhapatnam urban areas in Visakhapatnam, are collected for this study.

ISSN PRINT 2319 1775 Online 2320 7876

 Volume 12, Iss 1, 2023

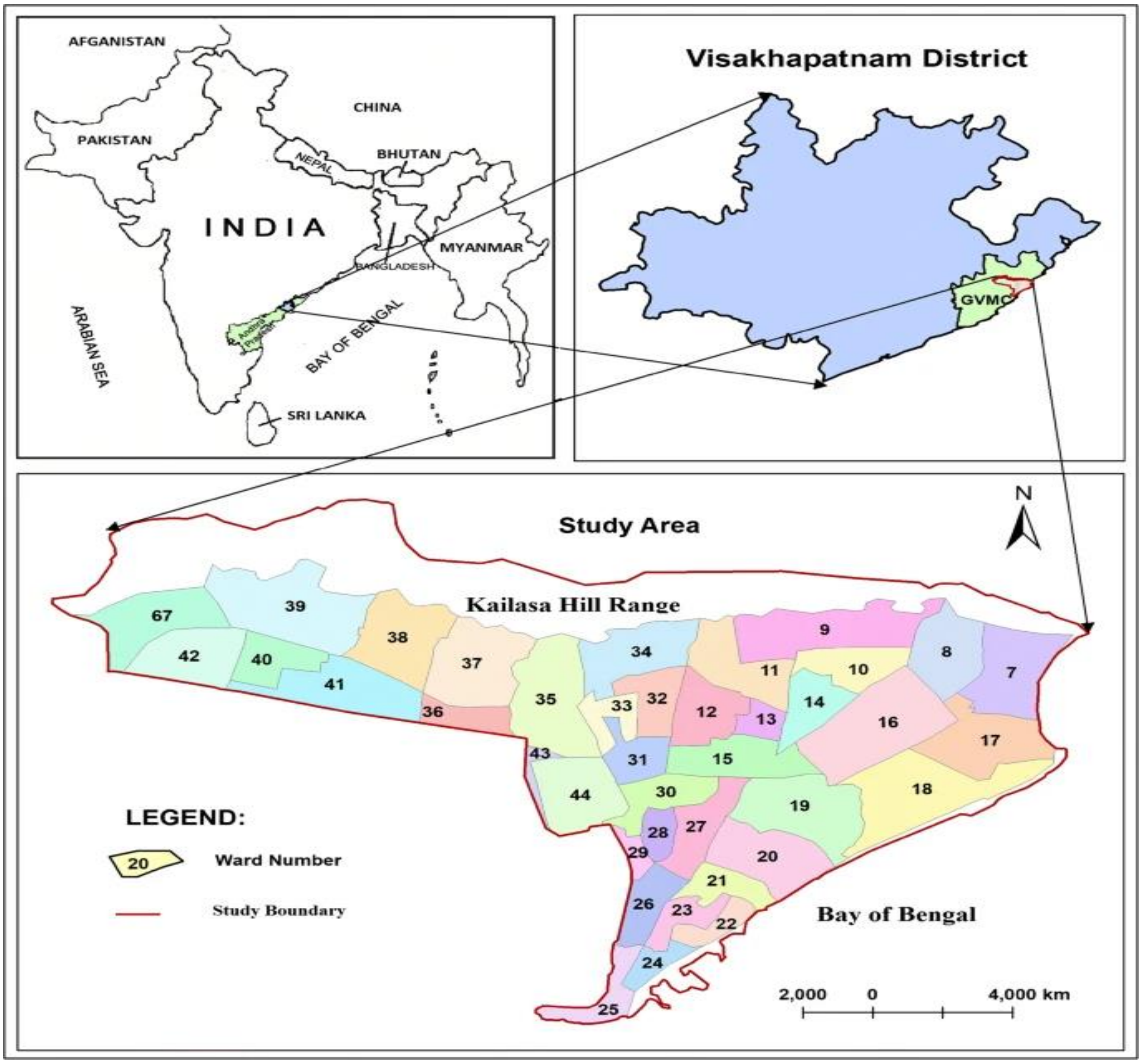

Figure 1: Location map of the study area's ground water well stations

3. Methodology

During the pre-monsoon (PRMS) and post-monsoon (POMS) samples have collected from '41' wells during the year '2021' in different places around the study area to evaluate the groundwater quality. Standard procedures have been applied to analyse the data samples [Narsimha Adimalla et al. 2019; Satyajit K. et al. 2020]. Physicochemical parameters such as p^H and electrical conductivity (EC) were measured by using a p^H meter (ELICO L1617) and EC meter (ELICO CM180) respectively. $CO_3^{\,-}$ and $HCO_3^{\,-}$ anions were determined by the titration method. Cations like Ca^{2+}, Na^+, Mg^{2+}, and K^+ and anions like Cl^-, $SO_4^{\,2-}$ and NO_3^- were identified by 930 compact ion flex chromatograph as per 4100B of APHA (American Public Health Association) methods.

3.1. Water Quality Index (WQI)

ISSN PRINT 2319 1775 Online 2320 7876
Research paper Volume 12, Iss 1, 2023

The WQI is a mathematical model to analyse quality of water characteristics excellent, good, poor, very poor and unsuitable for drinking based on various parameters present in the collected groundwater samples [Shekhar Gupta et al. 2009; Tejasvi Hora et al. 2019]. To analyse the WQI (Asadi S S, 2019) first step is to assign the weights to each physicochemical parameter based on the relative importance of the water quality of drinking water. The weights of the parameters TDS, pH, EC, Ca^{2+}, Na^{+}, Mg^{2+}, K^{+}, Cl^{-}, $SO_4{}^{2-}$, $NO_3{}^{-}$, and total hardness are assigned 1 to the maximum weight of 5. The relative weight Wi of each parameter was calculated by the equations given below.

Wi= wi i=0nwi

(1)

where, w_i is the weight assigned to each parameter and n is the number of parameters.

After calculating relative weight, the next is to calculate the quality classification. The quality rating scale qi of each element is calculated by dividing the analysed concentration of each parameter by the standard value and multiplying by 100.

qi = CiSi*100 (2)

where, 'c_i' is analysed the concentration of chemical parameters of each sample and the 'S_i' Standard value of the parameter.

The final step of the WQI is to calculate the sub-index of i^{th} parameter '*Sli*' for each parameter. It is calculated by multiplying the relative weight and quality rating scale.

$Sli = W_i*qi$ (3)

The water quality index 'WQI' is given by

WQi=i=1nSli (4)

The WQI has classified into five classes excellent, good, poor, very poor water, and unsuitable for drinking in the selected region.

3.2. Statistical Analysis

In the present study two sets of data for pre-monsoon and post-monsoon have been considered. R-studio has been used for carry out the analysis. In the first step descriptive statistics of the data has been presented and also compare the data of two seasons using z-test. The test statistic z- is used to know difference between the means of the two groups, divided by the standard

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved, Volume 12, Iss 1, 2023**

error of the difference. This test statistic follows a normal distribution based on the sample sizes of the two groups. A scatterplot matrix is a graphical method used to visualize relationships between multiple pairs of variables in a dataset. In the present study it has been used to examine different water quality parameters. To develop a water quality scatterplot matrix, plot each pair of variables against each other in a matrix format. The general linear model (GLM) can be used to analyse complex relationships between water quality parameters and other factors that may affect water quality like seasons, location and time. The significance of relationships between dependent and independent variables will be tested using GLM, it will help us to understand the contribution of each independent variable to present the variation in the dependent variable (WQI). Application of statistical tools such as Metamodelling and simulation are used to observe the variations in the data as well as the impact of influencing factors. These models define linear function as a model outcome. The graphical representation of the data is well presented using Minitab (2010).

Suppose that there is a simulation output response variable (Y) and is related to 'n' independent variables say $a_1, a_2, \ldots, a_k$. The dependent variable (Y) is a random variable and the independent variables $a_1, a_2, \ldots, a_k$ are called design variables are usually subject to control. The true relationship between the variables Y and a_i ($1 \le i \le n$) is represented by the simulation model. Our goal to approximate the relationship by a simple mathematical function called a Metamodels. Regression analysis is one method for estimating the parameters. The functional relationship of several independent variables influence the response variable is determine by the technique of regression analysis

$$Y=a0+a1x1+a2x2+\ldots+aixi, \text{ where } i=1, 2, \ldots, n \tag{5}$$

Here mainly focused on the WQI as response variable (Y), which is influenced by the different variables such as the concentrations of pH (x_1), EC (x_2), TDS (x_3), HCO3 (x_4), Cl (x_5), NO3 (x_6), SO4(x_7), Na^+ (x_8), K^+ (x_9), Ca(x_{10}), Mg(x_{11}) are as regressors. All these factors are affecting groundwater quality and its index. Here a_0 is the intercept and $a_1, a_2, \ldots a_n$, are the coefficients of the model input parameters.

4. Results and Discussion:

4.1. Analysis of groundwater

Groundwater samples collected from '41' wells in Visakhapatnam coastal area during PRMS and POMS, i.e., June 2021 and November 2021, were analysed for EC, pH, major anion, and cations. Table (1) shows the minimum, maximum, mean, and standard deviation

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, Volume 12, Iss 1, 2023**

values of the important chemical parameters and comparison of pre and post monsoon data has presented through figure (1).

4.2. Groundwater quality analysis for drinking purposes

Based on the analysis of Table (1), the WQI parameters show the minimum index value is 42.47, with a maximum of 1054.405 and an average of 381.07 for pre-monsoon data as well as the minimum index value is 47.00, maximum index is 1024.475 and an average of 354.56 for post monsoon data. As per the WQI values, samples were categorized into five classes shown in Table (3). From the analysed values of the water quality index classification, approximately 4.88% of the water samples fell under excellent in both seasons. 26.83% of the samples came under good water, 36.58% from poor water, 4.88% from very poor water, and 26.83% were not unsuitable for drinking during PRMS. Similarly, 17.08% of the samples come under good water, 41.46% under poor water, 4.88% are very, very poor water, and 31.70% are not unsuitable for drinking during POMS. Figure 2 shows WQI of '41' well stations have represented through line graph for both pre and post monsoon periods.

The significant difference between water quality indices of selected well stations between pre-monsoon and post-monsoon have studied with the help of paired t-test. Here we notified that the calculated value of 't' is 0.64125 and its p-calculated value is 0.525. As per the hypothesis testing procedure, there is no much difference between the pre and post monsoon data of selected stations. As per the calculation of the correlation, we noted that the WQI of selected well stations for the period pre and post monsoon exhibited the similar pattern.

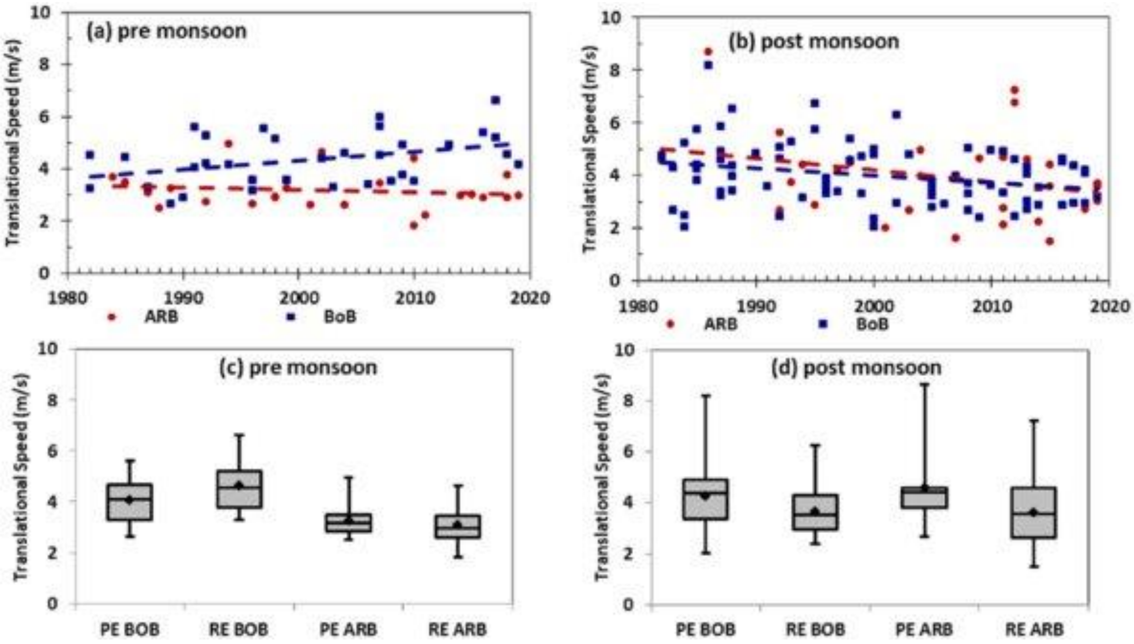

Figure 2: The correlation between premonsoon data's individual parameters

The relation of one variable with the other variable is represented through the Scatter plot matrix of Figure (3) and Figure (4). By observing the graphs, some variables are more strongly related to other variables. These variables can cause the affect of quality characteristics in measuring the WQI.

The functional relationship of the form of several independent variables influences the response variable is

$$Y=a0+a1x1+a2x2+...+aixi, \text{ where } i = 1, 2, \ldots,11 \quad (6)$$

The significance of relationships between the WQI and chemical parameters have been tested in GLM, then by the equations, it can be understood how much each chemical parameter contributes to variations in the WQI.

The regression equation for pre-monsoon season is,

$$WQI = - 581 +80.5\ pH +126EC - 196TDS -0.157\ HCO3 - 0.240\ Cl$$

$$-15.4\ NO3 – 3.10\ SO4 + 1.89Na + 0.64\ K +5.09\ Ca + 1.99\ Mg \quad (7)$$

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 12, Iss 1, 2023

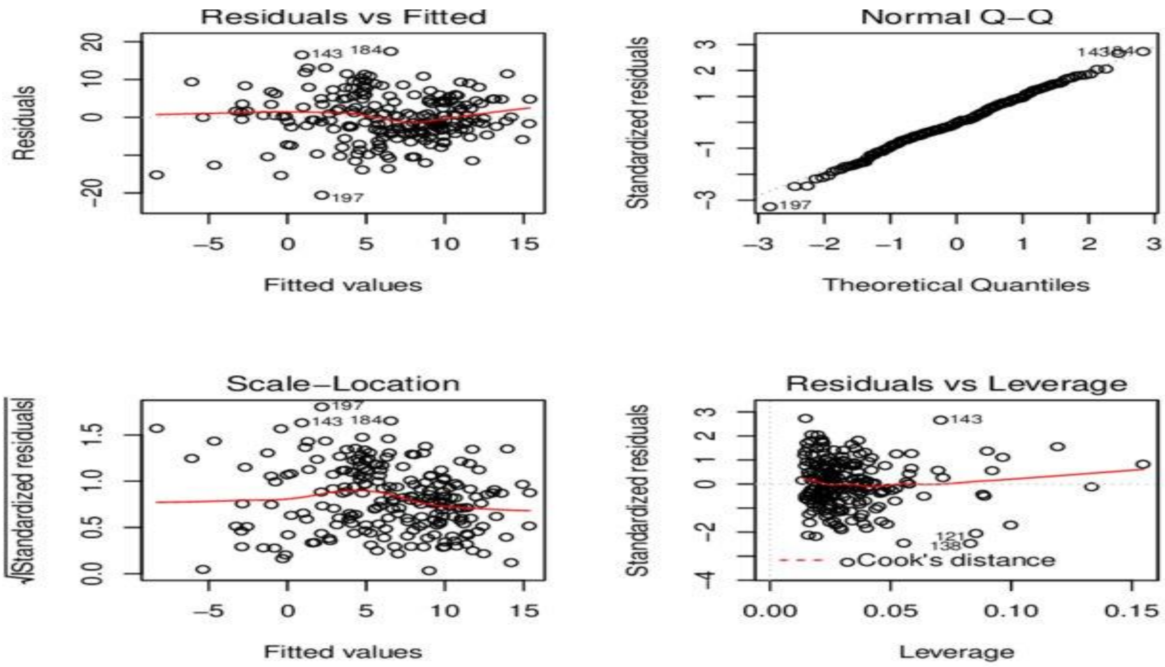

Figure 3: PRMS Residual Model Diagnosis

Marikavalasa denotes an observation with a large standardized residual Chippada denotes an observation value gives the large influence on PRMS.

In the post-monsoon study, TDS is highly correlated with other X variables, so TDS has been removed from the equation existing data is fitted to the equation number 8.
The regression equation for post-monsoon season is,

$$WQI = - 580 + 116\,PH + 0.380\,EC - 0.374\,HCO3 - 0.046\,Cl + 0.133\,NO3 + 0.385\,SO4 - 4.29\,Na + 0.56\,K - 7.34\,Ca + 6.69\,Mg \quad (8)$$

Sivajipalema and Pandurangapuram denotes the observations with a large standardized residual and Chippada and Kothooru denotes the observations value gives the large influence on POMS. Figures (5) and (6) shows the Residual Model Diagnosis of PRMS and POMS respectively.

From the above analysis, it has been observed that, HCO_3 has the strong association with other variables such as EC, TDS; Chloride with EC, TDS and HCO_3; Sodium and Magnesium with EC and TDS in PRMS. Similarly, P Cl, Na and Mg have the strong association with other variables such as EC, TDS in POMS. In both the seasons EC, TDS are strongly correlating with other factors.

Conclusions

ISSN PRINT 2319 1775 Online 2320 7876
Research paper Volume 12, Iss 1, 2023

In the coastal region of Visakhapatnam, Andhra Pradesh, a comprehensive analysis of groundwater chemical quality was conducted. Forty-one samples were collected during premonsoon (PRMS) and postmonsoon (POMS) periods, assessing parameters like pH, EC, TDS, cations, and anions. pH levels ranged from 6.18 to 8.9 (PRMS) and 6.03 to 8.56 (POMS), reflecting an acidic to alkaline nature. TDS concentrations varied from 211.20 to 3392 mg/l (PRMS) and 153.60 to 4492.80 mg/l (POMS), suggesting potential seawater intrusion along the coast. Water quality index classification revealed 26.83% and 31.70% of samples unsuitable for drinking during PRMS and POMS, respectively. Specific areas, including Chukkavanipalem, Nagarampalem, Marikavalasa, Rushikonda, Yendada, Appugar, MVP, Sivajipalem, Pandurangapuram, and Arilova, consistently exhibited water unsuitability for drinking in both seasons. Statistical analysis provided a comprehensive overview, offering suggestions to enhance groundwater quality in the region.

References

1. Suryanarayana, C., & Mahammood, V. (2019). Groundwater-level assessment and prediction using realistic pumping and recharge rates for semi-arid coastal regions: a case study of Visakhapatnam city, India. Hydrogeology Journal, 27(1).
2. Deshpande, M., Singh, V. K., Ganadhi, M. K., Roxy, M. K., Emmanuel, R., & Kumar, U. (2021). Changing status of tropical cyclones over the north Indian Ocean. Climate Dynamics, 57, 3545-3567.
3. Asadi SS, Padmaja V, Anjireddy M (2005). Water quality index-based assessment of ground water quality in municipal corporation of Hyderabad, India, *J Ind Assoc Envir Mana*.; 32: 166-170.
4. Linn, K. A., Laber, E. B., & Stefanski, L. A. (2015). iqLearn: Interactive Q-learning in R. Journal of statistical software, 64(1).
5. Dipankar Saha and Ranjan K (2019). Groundwater Resources of India: Potential, Challenges and Management, *Groundwater Development and Management*, 19-42.
6. Charan Kumar G, Shoba Latha G, Rajyalakshmi K (2020). Identification of Influencing Parameters of Water Quality and its Analysis by using Spatial Graphs and Autocorrelation Function. *Journal of critical reviews*, 7 (5), 1-7.
7. Singaraja, C (2017). Relevance of water quality index for groundwater quality evaluation: Thoothukudi District, Tamil Nadu, India, *Applied Water Science*, 7, 2157–2173.
8. Gyananath, G, Islam SR and Shewdikar, S V (2001). Assessment of Environmental Parameter on Groundwater Quality, *Indian Journal of Environmental Protection*, 21(4), 289-294.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, Volume 12, Iss 1, 2023**

9. Saikrishna, K, Purushothamb, V. Sunithac, Y. Sudharshan Reddy, D. Linga and B. Kiran Kumar, (2020). Data for the evaluation of groundwater quality using water quality index and regression analysis in parts of Nalgonda district, Telangana, Southern India. *Data in Brief*, (32), 106235.
10. Minitab 17 Statistical Software [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com), 2010.
11. Narsimha Adimalla & Jianhua Wu. (2019). Groundwater quality and associated health risks in a semi-arid region of south India: Implication to sustainable groundwater management. *Human and Ecological Risk Assessment*, 1549-7860).
12. Satyajit K. Gaikwad, Ajay Kumar K. Kadam, Ritish R. Ramgir, Akanksha S. Kashikar, Vasant M. Wagh, Avinash M. Kandekar, Suryakant P. Gaikwad, Ramdas B. Madale, Namdev J. Pawar, Kishor and D. Kamble, (2020). Assessment of the groundwater geochemistry from a part of west coast of India using statistical methods and water quality index. *Hydro Research*, 48-60.