Annals of Pure and Applied Mathematics
Vol. 15, No. 1, 2017, 123-131
ISSN: 2279-087X (P), 2279-0888(online)
Published on 11 December 2017
www.researchmathsci.org
DOI: http://dx.doi.org/10.22457/apam.v15n1a12

Annals of Pure and Applied Mathematics

Assessment by Multivariate Statistical Analysis of Ground Water Physicochemical Parameters of Cauvery Delta Region, Tamil Nadu, India

P.Selvavathi **and** ***R. Sophia Porchelvi***

Department of Mathematics, A.D.M College for Women (Autonomous)
Nagapattinam, India

Received 5 November 2017; accepted 4 December 2017

Abstract. Water is an essential resource for all life on the planet. Water should be purified for a better life style. It is the basic duty of every individual to conserve water resources. The present study attempts to bring an acute awareness among the people about the quality of ground water by taking water samples from the Cauvery Delta Region, Tamil Nadu (India) for analysis. The sampling areas are Nagore, Nagapattinam, Kizhvelur, Adiyakamangalam, Tiruvarur, Koradacherry, Needamangalam, Kamalapuram, Thirumakottai and Thanjavur. The experiment analyses its various physicochemical parameters such as Electrical Conductivity, TDS, Turbidity, pH, Chloride, Sulphate, Calcium, Magnesium, Phosphate, Nitrate, Nitrite, Zinc, Copper, Iron, Manganese and Lead. The results are compared to standards of the WHO (World Health Organization) and Surface Water Quality Standards (as per IS: 2296). A systematic correlation and regression study showed a significant linear relationship among different pairs of water quality parameters.

Keywords: Ground water,Pearson correlation, Regression analysis, Multivariatestatistical method.

1. Introduction

According to a Central Ground Water Board report as many as 650 cities and towns lie along polluted rivers, which contaminate groundwater in India. About half of India's groundwater is contaminated, according to a Central Groundwater Board report. As many as 276 districts have high levels of fluoride, 387 districts report nitrates above safe levels and 86 districts have high levels of arsenic, said the report. On an average, contaminated water caused 10 million cases of diarrhoea, 740,000 cases of typhoid and 150,000 viral-hepatitis cases between 2007 and 2011, the groundwater board said. As many as 650 cities and towns lie along polluted rivers, which contaminate groundwater,the report said."Poor environmental management systems" in industries lead to toxic and organic waste discharges of water, the report said. This has resulted in "pollution of surface and groundwater sources from which water is drawn for irrigation and domestic use". Even this source of water – contaminated as it is – is dying.

In nine states - in the south, west and central India- groundwater levels are now described as "critical", according to a 2016 Parliament committee report on water resources. "Critical" implies a state where 90 per cent of groundwater has been extracted, with significant decline in recharge capability. As of December 2015, of 6,607 units (blocks, mandals, talukas) assessed, 1,071 in 16 states and two in Union Territories were categorized as "over-exploited", which means 100% of groundwater has been drawn, with little chance to recharge. Particularly, in rural areas, it accounts for 88% of the drinking water (Kumar A., 2004). The quality of water is a vital concern for mankind since it is directly linked to human welfare and it is required to conserve water resources (Jothivenkatachalam, et.al. 2010). Drinking water quality is affected by the presence of different soluble salts (Sonawane and Khole, 2010). The major problem with the groundwater is that once contaminated, it is difficult to restore its quality. Hence, there is a need and concern for the protection and management of groundwater quality.

In recent years, an easier and simpler approach based on statistical correlation, has been developed using mathematical relationship for comparison of physicochemical parameters (Iyer, C.S et al., 2003; SarkarMitali et al., 2006), for finding the water quality. The developed regression equations for the parameters having significant correlation coefficients can be successfully used to estimate the concentration of other constituents. A systematic study of correlation and regression coefficients of the water quality parameters not only helps to assess the overall water quality, but also to quantify relative concentration of various pollutants in water and provide necessary care for the implementation of rapid water quality management programs**.** In the present study, an attempt has been made to study the interactions between different components of groundwater and their relationship with total dissolved solids, using Pearson correlation matrix and multiple linear regression technique. This paper comprising of statistical analysis of groundwater quality parameters is the first exploration in the study area of the Cauvery Delta Region, Tamil Nadu, India.

1.1. Study Areas

Cauvery Delta Zone (CDZ) lies in the eastern part of Tamil Nadu between 10.00-11.30, North latitude and between 78.15 – 79.45 longitude. It is bounded by the Bay of Bengal on the East and the Palk straight on the South, Trichy district on the west, Perambalur, Ariyalur districts on the north west, Cuddalore district on the North and Puddukkottai district in the South West. The present investigation was carried out on Tiruvarur, Thanjavur and Nagapattinam town and their adjacent areas in the Cauvery Delta region in Tamilnudu.

Figure 1: Map showing the study area of the Cauvery Delta Region

Sampling places	Sampling Point number
Nagore	S_1
Nagapattinam	S_2
Kizhvelur	S_3
Adiyakamangalam	S_4
Tiruvarur	S_5
Koradacherry	S_6
Needamangalam	S_7
Kamalapuram	S_8
Thirumakottai	S_9
Thanjavur	S_{10}

Table 1: Sampling Points of Cauvery delta region

Parameters	Method	Reference
pH	pH –metric	IS : 3025 (Part II)
Turbidity	Nephelo metric method	IS : 3025 (Part 10)
Electrical Conductivity	Conduct metric	IS 3025 : 1964
Total Dissolved Solid	Gravimetric method	IS : 3025 (Part 16)
Chloride	Argento metric method	IS : 3025 (Part 32)
Sulphate	Turbidity method	IS : 3025 (Part 24)

Calcium	EDTA Titrimetric method	IS : 3025 (Part 40)
Magnesium	EDTA Titrimetric method	IS : 3025 (Part 46)
Manganese	Colour comparison	IS 3025 : 1964
Nitrate	Chromo tropic acid method	IS : 3025 (Part 34)
Iron	1,10-Phenanthroline method	IS : 3025 (Part 53)
Nitrite	Spectrophotometric method	IS : 3025 (Part 34)
Copper	Neocuprine method	IS : 3025 (Part 42)
Phosphate	Spectrophotometric method	IS 3025 : 1964
Zinc	Spectrophotometric method	IS 3025 : 1964
Lead	Spectrophotometric method	IS 3025 : 1964

Table 2: Methods used for estimation of various physicochemical parameters

2. Materials and methods

2.1. Statistical analysis

2.1.1. Pearson correlation

The Pearson correlation coefficient is commonly used to measure and establish the strength of a linear relationship between two variables or two sets of data. It is a simplified statistical tool to show the degree of dependency of one variable to the other (Belkhiri et al., 2010). The Pearson correlation coefficient (r_{xy}) is computed by using the formula as given (Patil and Patil., 2010; Jothivenkatachalam, et.al., 2010; Kumar and Sinha., 2010):

$$\boldsymbol{r_{xy}} = \frac{\boldsymbol{n}\sum(\boldsymbol{x_i y_i}) - (\sum \boldsymbol{x_i})\,(\sum \boldsymbol{y_i})}{\sqrt{[\boldsymbol{n}\sum \boldsymbol{x_i^2} - (\sum \boldsymbol{x_i})^{\wedge}2] - [\boldsymbol{n}\sum \boldsymbol{y_i^2} - (\sum \boldsymbol{y_i})^{\wedge}2]}} \quad \textbf{(1)}$$

where the variables x and y represents two different water quality parameters; n= number of data points/ number of groundwater samples.

The interdependence of different water quality parameters on each other was evaluated on the basis of r_{xy} from the equation (1). The correlation coefficient is always between -1 and +1. A correlation closer to +/- 1implies that the association is closer to a perfect linear relation. Interpretation of the Pearson correlation coefficients, adopted in the present study are: r = -1 to -0.7 (strong negative association); r = + 0.7 to + 1.0 (strong positive association); r = -0.7 to -0.3 (weak negative association); r = + 0.3 to + 0.7 (weak positive association); r = - 0.3 to + 0.3 (negligible or no association). Thus, for the 16 water quality parameters, the possible correlations between every pair were computed using Microsoft excel and arranged into a correlation matrix. Precisely, a correlation matrix is a table of all possible correlation coefficients between a set of variables.

2.1.2. Multiple linear regressions

In this study, we have applied the multiple linear regression approach to develop a relationship between several independent / predictor variables and a dependent/predicted variable. This method is successfully used by different authors to establish a statistical model (Ghasemi and Saaidpour., 2007). Multiple linear regression is based on least squares : the model is fitted such that the sum-of-squares of differences of observed and predicted values is minimized.

2.2. Model equation

The model expresses the value of a dependent variable as a linear function of one or more independent variables and an error term:

$$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} \dots\dots\dots\dots + \beta_n x_{i,n} + e_i \quad (2)$$

$x_{i,n}$= value of n^{th} predictor
β_0= regression constant
β_n= coefficient on the n^{th}predictor
n = total number of predictors
y_i = predictant
e_i = error term

2.3. Predicted equation

The model (1) is estimated by least squares, which yields parameter estimates such that the sum of squares of the errors is minimized. The resulting prediction equation is

$$\ddot{y}_i = \ddot{\beta}_0 + \ddot{\beta}_1 \ddot{x}_{i,1} + \ddot{\beta}_2 \ddot{x}_{i,2} + \cdots \dots . + \ddot{\beta}_n \ddot{x}_{i,n} \quad (3)$$

where variables are defined as in (2). "**..** " denotes estimated values.

2.4. Residuals

The error term in equation (2) is unknown because the true model is unknown. Once the model has been estimated, the regression residuals are defined as

$$\ddot{e}_i = y_i - \ddot{y}_i$$

y_i = observed value of predictant; $\ddot{y}_i$ = Predicted value of predictand.

3. Results and discussion

The minimum and maximum concentration, i.e. the range of the different physicochemical parameters of water quality constituents such as Electrical Conductivity (EC), total dissolved solid (TDS), Turbidity, pH, Chloride (cl^-), Sulphate (SO_4^{2-}), Calcium(Ca), Magnesium(Mg), Phosphate(PO_4^{2-}), Nitrate($N0_3^-$), Nitrite (NO_2^-), Zinc (Zn), Copper (Cu), Iron (Fe), Manganese (Mno) and Lead (Pb) in the study area are given in Table 2 along with Mean, Standard Deviation, Variance and Standard Error.

Similar descriptive statistics of the analyzed groundwater quality parameters are presented in Table 3.

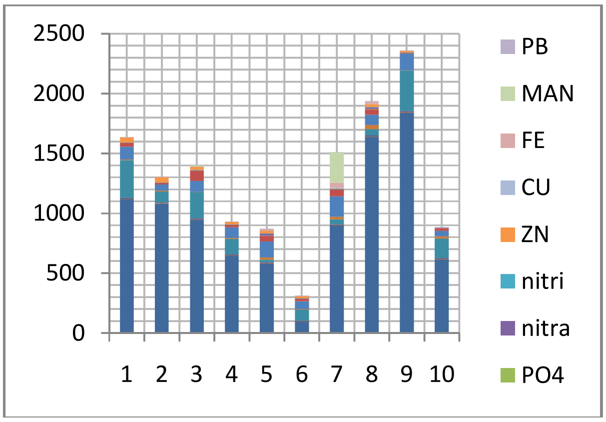

Figure 2: Physicochemical parameter variation of the study area

S.No	Parameter	Min.(mg/l)	Max.(mg/l)	Mean	Standard Deviation	Variance
1.	EC	0.58	2.62	1.887	1.2533	1.5707
2.	TDS	493	1638	985.4	453.5287	2568833022
3.	TUR	0.8	1.3	0.91	0.3478	0.121
4.	p^H	6.93	7.90	7.53	0.3323	0.1104
5.	Cl	25	339.90	146.926	11.6249	12460.12
6.	SO_4^{2-}	1.6	22.8	12.44	11.1575	124.49
7.	Ca	43.1	138.28	97.283	40.5557	1644.77
8.	Mg	6.09	89.18	35.973	25.0896	629.4896
9.	PO_4^{3-}	0.12	2.40	0.855	0.6604	0.4779
10.	NO^{3-}	2.47	18	6.017	5.8043	33.6899
11.	NO^{2-}	0.0012	0.0035	0.54328	0.8458	0.7166
12.	Zn	1.57	48.53	22.0755	15.5093	243.0587
13.	Cu	BDL	6.94	1.5062	2.3804	4.8105
14.	Fe	0.29	37.39	6.6731	12.1357	142.2754
15.	Mno	BDL	255	68.2025	80.1418	6422.7
16.	Pb	0.2	6.90	1.875	2.1708	4.7126

Table 3: Basic statistics of groundwater in the study area

Regression analysis was conducted to investigate the relationships between TDS and other water properties, which shows significant correlation with TDS (r= 0.7 to1) using Microsoft Excel. The [EC], [Cl] , [SO_4^{2-}], [Ca], [Mg], [NO_3^-], [Zn], [Mno], [Pb] were considered as independent variables and TDS as a dependent variable. An analysis of residuals was developed and R^2 values were studied. Among all candidate equations, the equation where this ratio was closer to 1 was selected. The descriptors and the regression coefficient of this model are presented in Table 4. The positive sign of the beta coefficients pertaining to these variables indicates that there is a positive relationship between TDS and elements of ground water properties [EC], [Cl] , [SO_4^{2-}], [Ca], [Mg], [NO_3^-], [Zn], [Mno], [Pb].

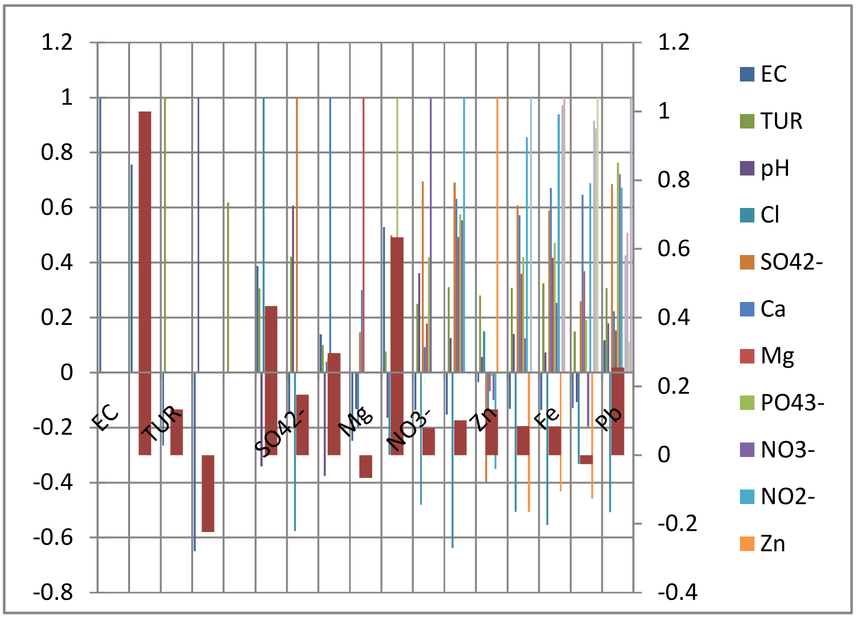

Figure 3: Graph showing correlation between TDS and other parameters

The selected regression equation is given by:

TDS= -1037.69+ 268.6805 EC+2.5435 Cl+38.0587 SO_4^{2-} + 2.9595 Ca+ 0.8993 Mg – 10.0923 NO_3^- + 16.1034 Zn + 0.391422 Mno + 1.9988 Pb.

Regression	Coefficients
(β_0)	-1037.69
EC (β_1)	268.6805
Cl (β_2)	2.5435
SO_4^{2-} (β_3)	38.0587
Ca (β_4)	2.9595
Mg (β_5)	0.8993
$NO_3^-$$(\beta_6)$	-10.0923
Zn (β_7)	16.1034
Mno (β_8)	0.3914
Pb (β_9)	1.9988

Table 4: Summary of the multiple linear regressions predicting TDS

Correlation	Coefficients
Multiple R	1.000
R Square	1.000
Adjusted R Square	-1.4E-09
Standard Error	0

Table 5: Summary of the multiple linear regressions predicting TDS

4. Conclusion

The quality of ground water sample collected from ten different locations of Cauvery Delta Region is analysed and studied. Multivariate statistical method used in this study (Pearson correlation coefficients and multiple linear regressions) help to find statistically important factors in data variability and thus improve conclusions in environmental impact studies. Pearson correlation matrix was applied to all the collected water samples for identifying the possible statistical relationship between different pairs of ground water quality parameters. A highly strong correlation was observed between Fe and Cu, which gives us an idea about the total hardness of water. A multiple linear regression was used to establish relationship between TDS and other chemical water properties. The positive sign of the regression coefficients indicates that there is a positive relationship between TDS and elements of ground water properties: [EC], [Cl] , [SO_4^{2-}], [Ca], [Mg], [NO_3^-], [Zn], [Mno], [Pb]. It can be concluded that the total dissolved solids is an important physicochemical water quality parameters, because they are correlated with most of the elements in the groundwater. The study further revealed that the water is not safe for drinking but suitable for aquatic life and also for irrigation purpose in some extend. The human beings of the Cauvery Delta Region are suffering with various diseases like gastrointestinal irritations, skin irritation, vomiting, dehydration, liver damage, renal damage, impaired neurological and motor development, etc. Hence, rapid and reliable monitoring measures are essential for keeping a close watch on water quality and health environment.

Acknowledgement. The first author thankful to UGC for supporting this research work through Rajiv Gandhi National Fellowship.

REFERENCES

1. APHA. Standard methods for examination of water and wastewater, 19th edn.,*American Public Health Association,* Washington, D. C. (1995).
2. Arul Antony, *Indian Journal of Science and Technology*.(2008), 1(6), 1.
3. Bajpayee. et al, Assessment by multivariate statistical analysis of ground water geochemical data of Bankura, India, *International Journal Of Environmental Sciences,* 3(2) (2012)
4. L.Belkhiri, A. Boudoukha, L. Mouni, A multivariate Statistical Analysis of Groundwater Chemistry Data, *International Journal of Environmental Research*, 5(2) (2010) 537-544.
5. BIS Specifications for Drinking Water, IS:10500:1991, *Bureau of Indian Standard New Delhi*, India.
6. J.R. Dash., P.C. Dash, and H. K. Patra, *IJEP,* 26(6) (2006) 550.
7. C. Gajendran, and P. Thamarai, *Poll Res*., 27(4) (2008) 679.

8. J. Ghasemi, S. Saaidpour, Quantitative structure-property relationship study of n-octanol-water partition coefficients of some diverse drugs using multiple linear regression, *AnaliticaChimicaActa*, 604 (2) (2007) 99-106.
9. B.K. Honda, Hydro chemical Zones of India, Proc. *Seminar on Ground water Development, Roorkee*, (1986) 339-450.
10. Iyer, Padmanabha C.S., Sindhu, M., Kulkarni, S., Tambe, S., Kulkarni, D., *Journal of Environmental Monitoring*, 5 (2003) 324-327.
11. C.K. Jain, M.K. Sharma, Relationship among water quality parameters of groundwater of Jammu District, *Pollution Research*, 16(4) (1997) 241-246.
12. C.K. Jain, M.K. Sharma, Regression analysis of groundwater quality data of Sagar District Madhya Pradesh, *Indian Journal of Environmental Health,* 42(4) (2000) 159-168.
13. H.P. Jarvie., B. A. Whitton, and C. Neal, *Sci. Total Environ*., 79 (1998) 210-211.
14. M.A.M.Joarder, F. Raihan, J. B. Alam, and S. Hasanuzzaman, *Int. J. Environ. Res*, . 2(3**)** (2008) 291.
15. M. Joarder, A., Raihan, F., Alam, J. B., Hasanuzzaman, S., Regression analysis of groundwater quality data of sunamjang district, Bangladesh, *International Journal of Environmental Research,* 2(3) (2008) 291-296.
16. K. Jothivenkatachalam, A. Nithya, S. Chandra Mohan, Correlation analysis of drinking water quality in and around Perur block of Coimbatore District, Tamil Nadu, India, *Rasayan Journal Chemistry,* 3(4) (2010) 649-654.
17. A.Kumar, Water Pollution (2004), *Nisha Enterprises, New Delhi, 1*.
18. A. Kumar, S.N. Garg, K. Gopal, Relationship between electrical conductivity and Cationic and anionic contents in groundwater, *Indian Journal of Environmental Protection,* 14(7) (1994) 595-603.
19. N.Kumar, D.K.Sinha, Drinking water quality management through correlation studies among various physicochemical parameters: a case study, *International Journal of Environmental Sciences,* 1(2) (2010) 253-259.
20. S.W. Liao, H. S. Gau, W.L. Lai, J. J. Chen, and C. G. Lee, *J. Environ. Manag*., 88 (2) (2007) 286.
21. H.J Mahvi. Nouri, A. A. Babiei, and R. Nabizadeh, *Int. J. Environ. Sci. Tech.,* 2 (1) (2005) 41.
22. N. Manivasagam, Physico-chemical examination of water, sewage and industrial effluents, Pragati.
23. Mayur C. Shah, *Poll Res*., 25(3) (2006) 549.
24. S. Moscowand K. Jothivenkatachalam ,*J. of. Env. Sci*., communicated**.**