Linear Correlation study and regression analysis of drinking water quality in Mokokchung town, Nagaland, India

Daniel Kibami, Chubaakum Pongener, K. Rao, Dipak Sinha

Abstract— **A statistical regression analysis method of drinking water from 12 different water sources around Mokokchung town was carried out. The water quality parameters considered in the present study were temperature, pH, electrical conductivity, total dissolved solids, alkalinity, total hardness, chloride, fluoride, dissolved oxygen, and the presence of calcium, and magnesium, Sodium, Potassium, nitrate, sulphate and phosphate. Correlation coefficients were determined to identify the highly correlated and interrelated water quality parameters. The results revealed a systematic correlation and regression with a significant linear relationship among different pairs of water quality parameters.**

Index Terms— Regression analysis, Water quality parameters, Correlation coefficient

INTRODUCTION

The quality of drinking water is a powerful environmental determinant of health [1]. According to the World Health Organization drinking water must be free of chemicals and microbial contaminations which are risk to human health. Good drinking water quality is essential for the well being of all people [1, 5]. It has been reported that increasing population and its necessities have lead to the deterioration of surface and sub surface water [2]. It is therefore vital to regularly monitor the quality of water. In the state of Nagaland, the predominant sources of water are surface water from rivers, streams, ponds and natural springs and subsurface water occurring as ground water. Information of water quality in Nagaland is very scanty, as there is very little documentation on the state of rivers/ water bodies and thus monitoring is a very recent phenomenon and so far has been taken up on a very limited basis [3]. This study was considered important since Mokokchung is a fast developing town of Nagaland, where rapid urbanization is leading to rampant pollution of water sources and hence it was thought that a preliminary study of surface and ground water quality would be of value for developing management strategies for maintaining potable water quality.

The statistical regression analysis has been found to be a highly useful tool for correlating different parameters. Correlation analysis measures the closeness of the relationship between chosen independent and dependent variables [6-8]. The analysis has been put to use in water quality parameters for predicting the river water quality management.

Manuscript received June 10, 2014

Daniel Kibami, Department of Chemistry, Nagaland University, Lumami, Nagaland, India
Chubaakum Pongener, Department of Chemistry, Fazl Ali College Mokokchung, Nagaland, India

Materials and Methods

Study area

Mokokchung with a Latitude of 26.33° N and Longitude of 94.53° E is situated in the north eastern region of Nagaland state. It is located 1325 meters above sea level and receives an annual rainfall of around 200 cm on an average [4].The present study was planned by selecting(one sample from each colony/ward) different ground and surface water sources around Mokokchung town.

.

Experimental methods

The water samples were collected from the various sources and analyzed as per standard procedures. The various water quality parameters are selected the methods of chemical analysis are listed in the Table 1. The detailed procedures employed for the analysis of different water quality parameters are as follows:

Sl.No.	Parameters	Methods of determination
1.	Temperature	Thermometer
2.	Hydrogen ion Concentration (pH)	pH metry
3.	Electrical conductivity (EC)	Conductometry
4.	Total dissolved Solids (TDS)	Gravimetric method
5.	Alkalinity (ALK) as $CaCO_3$	Titrimetry

6.	Total Hardness (TH)as $CaCO_3$	EDTA –Titrimetry
7.	Calcium(Ca)	EDTA –Titrimetry
8.	Magnesium (Mg)	EDTA –Titrimetry
9.	Sodium (Na)	Flamephotometry
10.	Potassium (K)	Flamephotometry
11.	Chloride (CI)	(Argentometric method) Titrimetry
12.	Nitrate(NO_3)	Spectrophotometry
13.	Sulphate(SO_4)	Spectrophotometry
14.	Phosphate	Spectrophotometry
15.	Fluoride	SPADNS method
16.	Dissolved Oxygen (DO)	Titrimetry

Table1: Determination of water quality parameters

Linear Regression Model

Determination of correlation is important in finding out the strength of the relationship between the two inter dependent variables. There is a direct correlation between two parameters if change in one of the variable affects the other in a linear way. The Correlation coefficient is defined only if both of the standards are finite and both are non zero, where its absolute value cannot exceed 1. The correlation coefficient is symmetric: corr(X, Y) = corr(Y, X). If we have a series of 'n' measurements of X and Y written as *x* and *y*

The sample correlation coefficient is written as:

$$r = \frac{n.\sum x.y - \sum x.\sum y}{\sqrt{\left(n\sum x^2 - \left(\sum x\right)^2\right)\left(n\sum y^2 - \left(\sum y\right)^2\right)}}$$

Where x and y are the sample means of X and Y, To determine the straight linear regression, following equation of straight line can be used:

$y = ax + b$

Where, y and x are the dependent and independent variable respectively. a is the slope of line, b is intercept on y axis. The slope, a, and y intercept, b can be determined using the following equations.

$$a = \frac{n.\sum x.y - \sum x.\sum y}{n\sum x^2 - \left(\sum x\right)^2}$$

$$b = \frac{\sum y - b\sum x}{n}$$

Results and Discussion

The results of physico-chemical parameters of the water samples collected from 12 different places in and around Mokokchung town is describing as below.

Sl.No.	Water source	Ward/Colony	Temperature° C	pH	Electrical Conductivity μScm^{-1}	TDS mg/L	Total alkalinity (TA) mg/L
1.	Tap water	F.A.C campus	22.2	7.69	322	287	111.5
2.	stream	Alisunkum	22.3	6.93	55.1	49.9	109
3.	Bore well	Tondentsunyoung	22.6	6.50	332	297	106
4.	stream	Alongmen	22.5	7.52	262	234	107.5
5.	pond	Medical	22.4	7.01	204	185	112.5
6.	pond	Majakong	22.6	6.55	214	180	111
7.	Bore well	Salangtem	23.4	5.66	794	728	114.5
8.	pond	Dilong	24.6	5.05	547	495	115.5
9.	Bore well	Aongza	23.7	6.35	390	351	109.5
10.	stream	Alempang	24.1	5.51	484	530	117.5
11.	pond	Arkong	24.3	5.10	375	342	117
12.	Bore well	Yimyu	22.3	7.16	215	189	113.5

Table2: Results of Physico – Chemical analysis water samples

Sl No.	Dissolved oxygen (DO) mg/L	Total hardness (TH) mg/L	Ca mg/L	Mg mg/L	Cl mg/L	NO_3 mg/L	SO_4 mg/L	K mg/L	Na mg/L	F mg/L	PO_4 mg/L
1	7.00	125	15.23	2.92	14.21	0.194	7.83	11.00	8.00	1.76	0.75
2	7.80	100	8.01	4.87	21.31	0.113	7.79	1.00	3.00	0.43	0.41
3	9.00	145	29.65	15.10	42.63	0.010	14.01	1.00	10.00	1.24	0.30
4	13.00	130	20.04	6.33	35.52	0.016	17.46	5.00	17.00	0.41	0.80
5	6.90	55	18.43	5.84	32.66	0.080	16.60	3.00	20.00	0.57	0.70
6	12.00	60	16.03	4.38	11.36	0.191	9.86	3.00	24.00	1.90	0.84
7	12.50	85	26.45	5.36	63.90	0.005	12.76	3.00	53.00	0.15	0.42
8	14.30	95	16.83	12.18	61.06	0.021	14.45	5.00	47.00	0.18	0.55
9	15.20	125	17.86	9.49	14.20	0.028	16.54	2.00	22.00	1.35	0.40
10	8.60	150	20.84	2.90	56.80	0.015	9.67	3.00	20.00	1.95	0.65
11	14.50	70	8.21	6.81	51.12	0.141	8.47	2.00	39.00	0.19	0.59
12	12.35	60	9.61	2.43	39.76	0.098	12.67	4.00	9.00	0.83	0.45

Table3: Results of Physico – Chemical analysis water samples

Correlation analysis

A high correlation coefficient (nearly 1 or -1) means a good relationship between two variables, and a correlation coefficient around zero means no relationship. Positive value in the range of +0.8 to 1.0 characterizes a strong correlation between the parameters, while negative values of r indicate an inverse relationship. From the observations of the results of all samples we can predict some important linear equations ($y=ax + b$) between the water quality parameters which are in good agreement with experimental data. These linear equations are very much useful to calculate the water quality parameters. The larger the numerical value of the correlation coefficient ($r > 0.8$), the greater is the extent to which correlation holds between the two variables. The correlation coefficients (r) among various water quality parameters were calculated and the values indicated high positive relation is observed between EC and TDS (0.99013) , EC and Na(0.8084). From these values it can be suggested that, EC is highly dependent on the amount of dissolved solids (such as salt) in the water, and Na^+ ions contributing a major role in electrical conductivity. Interrelationship studies between different variables are very helpful tools in promoting research and opening new frontiers of knowledge which. The study of correlation reduces the range of uncertainty associated with decision making.

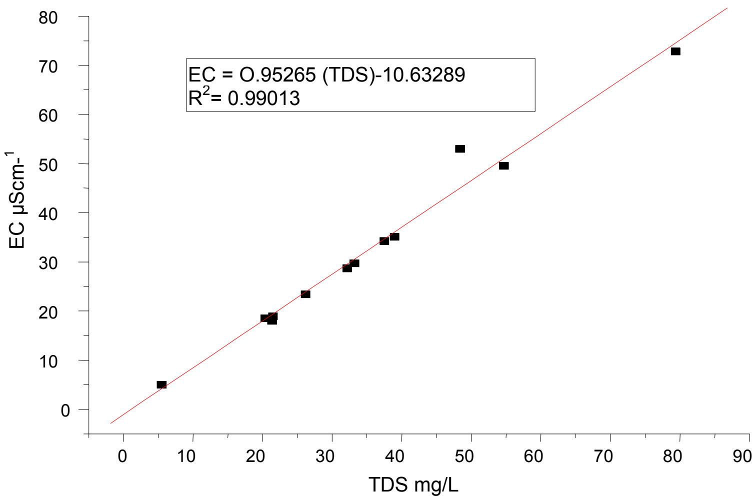

Fig: 1 Correlation between EC and TDS

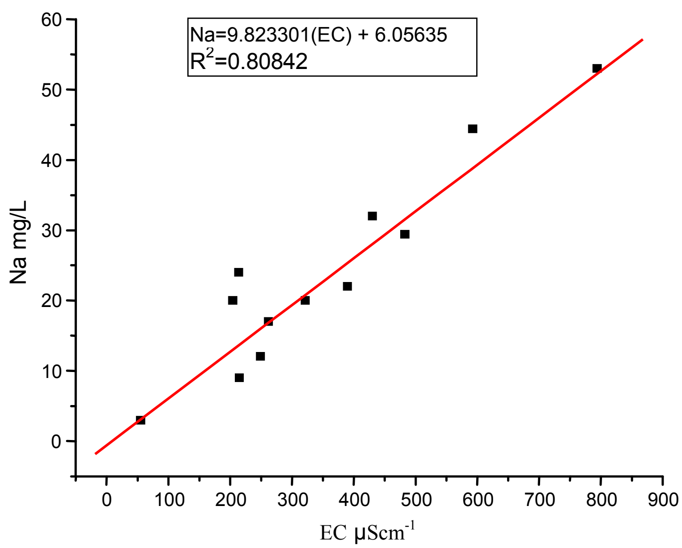

Fig: 2 Correlation between EC and Na

Pairs of perimeter	R value (n=12)	Regression coefficient		Regression equation
		a	B	
pH-EC	0.651	-0.00306	7.48735	pH=-0.00306(EC) + 7.48735
pH-TDS	0.6713	-0.00328	7.47484	pH= -0.00328(TDS)+7.47484
pH-TA	0.6649	-0.16415	24.81727	pH=-0.16415(TA) + 24.81727
pH-Cl^-	0.6956	-0.03347	7.659098	pH=-0.03347(Cl^-) +7.659098
TDS-EC	0.99013	0.95265	-10.63289	TDS= 0.95265 (EC) -10.63289
EC-Cl^-	0.696	7.138751	-85.05925	EC=7.138751(Cl^-) -85.05925
Na^+-EC	0.8084	9.823301	6.05635	Na^+=9.823301(EC) + 6.05635
EC-Ca^{2+}	0.5364	15.48502	-82.14652	EC=15.48502(Ca^{2+}) - 82.14652

TDS-TA	0.53812	0.010659	108.6476	TDS =0.010659(TA) + 108.6476
TDS-Cl^-	0.72151	0.073113	13.4781	TDS= 0.073113(Cl^-) + 13.4781
TDS-Ca^{2+}	0.5344	0.019243	11.06333	TDS= 0.019243(Ca^{2+}) + 11.06333
TDS-Na^+	0.7686	0.065744	1.475791	TDS= 0.065744(Na^+)+ 1.475791
TA-Cl^-	0.58660	0.114669	-107.8355	TA=0.114669(Cl^-) - 107.8355
TA-Na^+	0.5821	0.134814	109.0276	TA=0.134814 (Na^+) + 109.0276
Na^+-Cl^-	0.6426	0.761351	19.7868	Na^+= 0.761351(Cl^-) + 19.7868
Na^+-DO	0.5921	0.1138	8.516378	Na^+= 0.1138(DO) + 8.516378
PO_4^{3-} –K^+	0.52361	0.93359	7.90131	PO_4^{3-} = 0.93359(K^+) +7.90131
NO_3^- – $SO4^{2-}$	0.6551	0.01307	0.237364	NO_3^- = 0.01307($SO4^{2-}$) + 0.237364

Table 4: The Correlation coefficient (r) values for different water samples

	pH	EC	TDS	TH	TA	F	Cl	DO	NO_3	SO_4	PO_4	Ca	Mg	Na	K
pH		0.651	0.6713	0.04995	0.6649	0.24930	0.6956	-0.4709	0.3312	0.1177	0.2590	-0.093	-0.332	-0.775	0.3971
EC			0.99013	0.17178	0.4895	-0.17653	0.696	0.37731	-0.516	0.1290	-0.209	0.5364	0.1976	0.8084	0.0744
TDS				0.2287	0.53812	-0.11751	0.72151	0.3201	-0.5397	0.0899	-0.189	0.5344	0.1465	0.7686	0.0548
TH					-0.3245	0.378271	0.03117	-0.1715	-0.4594	0.0602	-0.165	0.4824	0.3011	-0.284	0.1231
TA						-0.122	0.58660	0.18271	0.05419	-0.334	0.1542	-0.270	-0.363	0.5821	0.0960
F							-0.4931	-0.3412	0.30289	-0.279	0.2805	0.1611	-0.229	-0.466	0.2351
Cl								0.23548	-0.6054	0.10985	-0.273	0.3095	0.2144	0.6426	-0.190
DO									-0.2376	0.31925	-0.127	-0.103	0.2781	0.5921	-0.200
NO_3										0.6551	0.4666	-0.639	-0.498	-0.296	0.37195
SO_4											-0.089	0.4691	0.4645	0.1473	-0.1516
PO_4												-0.17	-0.480	-0.017	0.52361
Ca													0.4701	0.2133	-0.0865
Mg														0.2168	-0.3427
Na															-0.0954
K															

Table 5: Linear correlation coefficient R and regression equation for some pairs Of parameters which have significant value of correlation

Conclusion

A linear regression analysis technique has been proven to be a very useful tool for monitoring drinking water and has a good accuracy between various physicochemical water parameters and it may be treated as one step ahead towards the drinking water quality management. A systematic correlation and regression in this study show that there is a significant linear relationship among different pairs of water quality parameters. The linear correlation is very useful to get fairly accurate idea of quality of the ground water by determining a few parameters experimentally. The significant correlation of pH with EC, TDS, TA and CI indicates that the value of pH is dependent upon these parameters. Whereas EC with TDS, CI and Na indicate the presence of chloride in sodium contributing to the total dissolved solids as electrolytes. Compared with the other water quality parameters, pH, EC and TDS are easily determinable. Hence on knowing this value exactly, we can calculate the other parameters, which are in good agreement with experimental data. So, this correlation determination provides quick monitoring of the quality of ground water.

References

1. World Health Organization, Guidelines for drinking water quality, 4th edn, WHO, Geneva, Switzerland. 2011.
2. USEPA. Handbook for analytical quality in water and wastewater. Environmental monitoring and support laboratory, Cincinnati, Ohio, U.S.A. 1979
3. http://npcbngl.nic.in/laboratory.htm
4. http://nagaland.nic.in
5. Anonymous, Report on UN Conf. On Environ. & Development, A/CONF. 151/26., 1, 277. 1992.
6. Jain, C.K. and Sharma, M.K. Regression analysis of ground water quality data of Malprabha river basin, Karnataka. J. Indian Water Resources Society.Vol. 22(1): 30-35. 2002.
7. Sharma, M.K. and Jain, C.K. Fluoride modeling using Best Subset Procedure .Journal of Applied Hydrology. Vol. XVIII: 12-22. 2005.
8. Singanan, M., Rao, K.S. and Rambabu, C. A correlation study on physico-chemical characteristics of ground water in Rameswaram Island. Indian J. Env. Prot. Vol.15(3): 212-217 .1995.
9. APHA. Standard methods for the examination of water and waste water. 21st edn. Amer. Pub. Health Assoc.Inc.Washington D.C. 2005.
10. Anil Kumar De. Environmental Chemistry, 2nd Edition. Wiley Eastern Limited, New Delhi. 1989.
11. Jain, C. K., and Sharma, M. K "Relationship among water quality parameters of Groundwater of Jammu District", Pollution Research, Vol 16(4), 241-246 Issue 4. 1997.
12. Joarder, M. A., Raihan, F., Alam, J. B., Hasanuzzaman, S. "Regression analysis of Groundwater quality data of Sunamjang District, Bangladesh", International Journal of Environmental Research.Vol. 42 (4); 159-168. 2008.
13. Jothivenkatachalam, K., Nithya, A., and Chandra Mohan, S. "Correlation analysis of drinking
14. water quality in and around Perur block of Coimbatore District, Tamil Naidu India", Rasayan Journal Chemistry.Vol -3 No.(4), 615-824. 2010
15. Kumar, A., Garg, S. N., and Gopal, K. "Relationship between electrical conductivity and Cationic and anionic contents in groundwater", Indian Journal of Environmental Protection.14(7); 595-603. 1994.
16. ISI, Indian Standard specification for drinking water, IS10500, ISI, New Delhi, 1993.
17. Somasekhara Rao K. and Someswara Rao B.,Correlations among water quality parameters of ground waters of Musunur mandal, Krishna District,Indian J. Environ. Prot. Vol.14(2) 128-132. 1994.
18. Kaza Somasekhara Rao, Correlations among water quality parameters of ground waters of Nuzvid town and Nuzivid mandal. Indian J. Environ. Prot.Vol.(13) 4 , 261-266. 1993.

19 Kumar, N., and Sinha, D. K.. "Drinking water quality management through correlation studi among various physicochemical parameters: a case study", International Journal of Environmental Sciences,Vol. 1(2), 253-259. 2010.

20 Rao, K. S., and Rao, B. S. "Correlation among water quality parameters of groundwaters of Musunur Mandal, Krishna district", Indian Journal Environmental Protection.Vol.14 (7), 528-532. 1994.

21 Kalyanaraman,S.B. & Geetha,G., Correlation analysis and prediction of characteristics parameters and water quality index of ground water, Pollutio Research.Vol. 24 (1), 197-200. 2005