

# **Regression Analysis of Multivariate Fluid Geochemical Data representing two distinct systems operative at the geothermal fields of Extra-Peninsular and Peninsular India**

**Amitabha Roy**

*Ex-Senior Director, Geological Survey of India*

---

## **Abstract**

*The present study deals with the Regression Analysis of multivariate fluid geochemical data to substantiate the findings of earlier Exploratory Factor Analysis in distinguishing two distinct systems of geotherms operative in Extra-Peninsular Himalayan mountain chains and relatively stable landmass or shield of Peninsular India. The regression analysis establishes. relationship between one Dependent Variable (DV) and one or more Independent Variables (IV). Factor analysis aids in the selection of significant fewer independent variables from many insignificant ones. Two different sets of Multivariate geochemical data – one from the tectonically active Extra-Peninsular Himalayan region and the other from Proterozoic belts in the Central Highland region of otherwise stable landmass or shield of Peninsular India, were subjected to multiple regression analysis to find out the genesis of geothermal hot springs spread over these areas conspicuously associating with the respective tectonic zones of different degree of severity. The regression analysis revealed two statistically significant suites of fluid geochemistry – 1. The overall salt assemblage and concentration of Cl-HCO<sub>3</sub>-SO<sub>4</sub>-Na-F or Chloride rich water suggestive of existence of a deep-seated hydrothermal magmatic system in operation at geothermal fields in Extra-Peninsular India and 2. Peninsular springs of K-Na-HCO<sub>3</sub> Bicarbonate rich waters with low SO<sub>4</sub>- content and higher contents of HCO<sub>3</sub> compared to other anions SO<sub>4</sub>, Cl, and F suggestive of a shallower non-magmatic origin.*

---

Date of Submission: 05-03-2023

Date of Acceptance: 18-03-2023

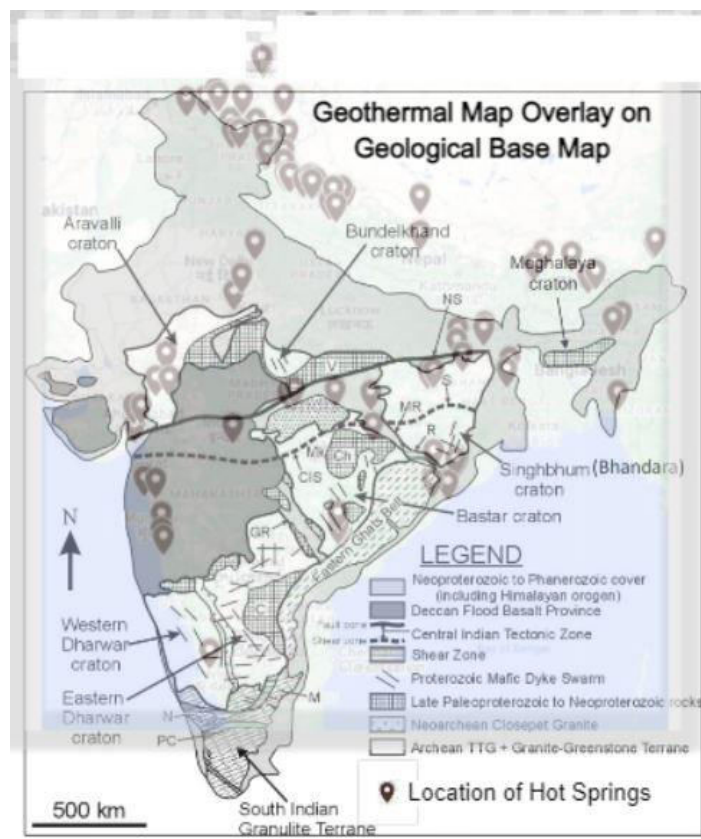
---

## **Introduction**

The surface manifestations of geothermal fields are the volcanoes, fumaroles, geysers, steaming grounds, and hot springs. There are about 340 hot springs spread over the Peninsular and Extra-Peninsular regions. The first attempt to list the hot springs in India was carried out by Schlagintweit in 1852 when he prepared an Inventory of 99 thermal springs. The GSI initiated its geothermal exploration with the launching of 'PUGA PROJECT' in Jammu and Kashmir. The Geological Survey of India has brought out a special Publication titled 'Geothermal Atlas of India' (Ravi Shankar et al., 1991) based on the data compiled from all sources of information both published and unpublished data on geothermal activities in India.

## **Location Map of Geothermal springs on Google My Map**

In the present study, 62 geothermal fields or hot springs spread over both Extra-Peninsular, and Peninsular regions of India have been considered. The GPS coordinates of each of the hot springs have been plotted on Google MY Map (AmitabhaRoy, 2023). For brevity, the Lat-Long i.e GPS coordinates have not been reflected here



Geologic-Tectonic Setting - Peninsular and Extra-Peninsular India

The Peninsular plateau is in the shape of a vast inverted triangle, bounded on the west by the Arabian Sea, on the east by the Bay of Bengal, and on the north by the Vindhya and Satpura ranges.

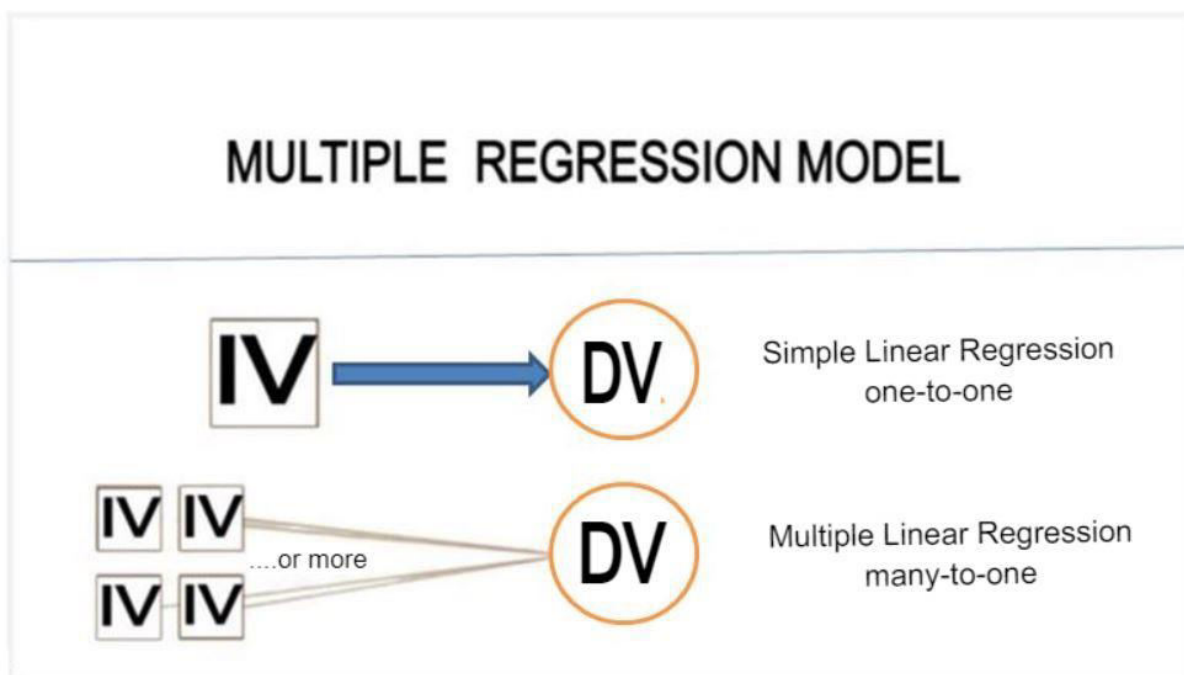
The Peninsular shield is a mosaic of Early to middle Precambrian stable landmass called cratons made up of very old igneous and metamorphic rocks (Archaean Gneisses and Schists/Dharwar system), Late Precambrian Proterozoic mobile belts or suture zones made up of oldest higher grade metamorphic rocks (quartzite-carbonate-phyllite) and Deccan flood basalts of Late Mesozoic to Tertiary age. The three main mobile belts are Satpura, Eastern Ghat, and Aravalli-Delhi belts. The major prominent rifts of the Proterozoic age that separate the southern and northern blocks of the Peninsular shield, is the Central India Tectonic Zone (CITZ) or Narmada-Son-Tapi (SONATA) or Saihadri-Satpura Lineament.

The Extra-Peninsular or the Himalayan region, a 2400 km long arcuate belt spreading over Jammu-Kashmir-Ladakh in the Northwest through Himachal Pradesh- Uttarakhand in the middle section to Sikkim-Arunachal Himalaya in the northeastern end represents a tectonically active orogeny associated with a high degree of seismicity in comparison to that of Peninsular India.

There is a striking similarity between the Peninsular and Extra-Peninsular regions in the distribution of geothermal fields. In both these regions there observed a sharp correlation between the tectonic zones and the disposition of geothermal hot springs as seen in the map given above.

In the Extra-Peninsular region, geochemical data were drawn from hot springs representing the Puga geothermal field in NW Himalaya, Ladakh district, Jammu & Kashmir, and the Tuting-Tidding Suture Zone (TTSZ) in NE Himalaya, Arunachal, India situated near the junction of the Indian and Asian plates characterized by volcanic sedimentary assemblages of rocks and Uttarakhand –Himachal Pradesh geothermal fields are situated within the Middle or Lesser Himalayan Crystalline (LHC) Zone. The hot springs of Peninsular India are restricted to the Proterozoic mobile belts in the Central Highlands leaving the southern more stable Deccan plateau completely devoid of any hot springs.

### Multiple Regression Analysis



### Multiple Regression Model

Multiple regression is an extension of linear regression. The Objective of Regression analysis is to explain variability in a Dependent Variable (DV) by means of one or more of Independent or control variables (IV). A multiple regression model is used when there is more than one independent variable affecting a dependent variable. While predicting the outcome variable, it is important to measure how each of the independent variables moves in their environment and how their changes will affect the output or target or dependent variable

The linear regression model describes the dependent variable with a straight line that is defined by the equation

$$Y = a + b \times X, \text{ where } a \text{ is the } y\text{-intersect of the line, and } b \text{ is its slope.}$$

Regression line for a multivariate data

$$Y = a + b_1 \times X_1 + b_2 \times X_2 + \dots + b_n \times X_n,$$

Where

Y = dependent variable

X<sub>i</sub> = independent variables

a = constant (y-intersect)

b<sub>i</sub> = regression coefficient of the variable X<sub>i</sub>

Going back to the Factor analysis (Amitabha Roy, 2023) the objective was to reduce a large number of variables into fewer numbers of factors or in other words to separate significant few from insignificant many. In factor analysis choice of the number of factors out of as many as the variables was a baffling issue. Here in multiple regression analysis too, choice of relevant variable (IV) out of many is an issue. One should never enter all the available variables at the same time. Carefully consider which independent variable is distinct or whether relevant to the problem. The first and the most reliable option is to use factor analysis which creates a smaller number of factors that account for, most of the original variables' information in them but which are mutually uncorrelated.

### **Assumptions**

- 1) Is the sample size sufficient or the chosen samples are representative of the population
- 2) Do the DV and IV show variation
- 3) Is the DV interval or ratio scaled
- 4) Is linearity or linear relationship between IV and DV exist
- 5) Multivariate normality i.e. approximately normally distributed (with a mean of zero)
- 6) No or little multicollinearity (occurs when the IVs are too highly correlated with each other)
- 6) No auto-correlation
- 7) Homoskedasticity vs Heteroskedasticity: The scatter plot is a good way to check whether the data are homoscedasticity meaning the residuals are equal around the regression line.

### **Multiple Regression Analysis Output**

- R (the multiple correlation coefficient),
- R squared (the coefficient of determination),
- Adjusted R-squared,
- The standard error of the estimate.

These statistics help one to figure out how well a regression model fits the data. The ANOVA table in the output would give you the p-value and f-statistic (MiroslawGrzesik, 2022)

**Data Set - I : Extra-Peninsular Geothermal Data**

NUM	TEMPC	pH	SPCMHO/cm	HCO3 mg/L	Cl mg/L	SO4 mg/L	TotHard	Ca mg/L	Mg mg/L	Na mg/L	K mg/L	F mg/L	B mg/L	SiO2 mg/L	TDS mg/L
1	59	8.1	1271	300	163	62	0	14	5	210	13	12	5	80	800
2	96	7.7	827	170	133	36	131	44	15	88	19	0.8	33	60	514
3	59	7.1	5260	490	855	1244	1214	342	87	600	109	3.6	138	30	4072
4	24	8.2	795	210	102	83	136	30	15	110	19	1.2	25	25	488
5	56	7.6	1280	342	232	26	72	26	1	260	16	10	10	107	870
6	44	7.6	2015	303	200	340	302	103	11	260	45	6	13	87	1280
7	50	8.3	525	173	45	28	40	13	2	103	5	10	3	23	363
8	90	7.9	1045	276	170	33	0	52	12	135	27	3	10	83	874
9	73	7.5	410	145	30	55	0	38	13	30	7	1	3	50	378
10	52	7.8	25	15	2	0	12	3	1	1	0	0.2	0	2	20
11	50	8.2	700	248	72	48	40	13	2	140	6	5	3	65	480
12	55	7.7	400	272	10	14	0	56	24	8	5	0.4	0	68	366
13	54	7.6	845	445	35	0	0	50	52	50	10	1.2	0	41	536
14	55	6.8	2630	112	1485	22	230	70	13	490	37	1.6	19	115	1630
15	25	7.7	139	103	8	29	0	45	44	24	10	0.7	2	30	442
16	81	8.1	315	117	15	30	0	34	3	30	5	1.6	0	22	245
17	62	8	1460	861	48	14	72	14	99	290	43	3	5	91	1015
18	59	7.8	465	278	12	27	214	42	26	15	8	0.5	1	69	360
19	32	6.7	95	38	5	0	0	6	7	2	1	0.4	0	11	42
20	32	8.3	2000	953	86	0	0	0	47	80	83	0	0	18	0
21	68	6.4	1239	734	12	5	200	64	10	180	38	2	1	130	860
22	56	6.9	770	439	41	21	215	40	23	163	15	4	2	34	510
23	76	7.7	720	254	13	99	32	13	0	135	6	12.5	2.8	101	570
24	28	7.1	770	363	17	66	132	40	8	120	7	10	2	53	575
25	66	7.7	2030	1610	85	57	34	10	2	580	48	10	8	130	840
26	40	7.2	3641	259	11	1484	1352	504	22	200	6	2.5	1	35	2557
27	12	7	1060	233	58	383	536	169	28	10	2	0.2	0	18	834
28	12	7.1	63	32	3	0	24	9	0	2	0	0.4	0	6	42
29	68	6.9	386	112	30	72	40	14	1	56	4	6	1	35	235
30	9	8	178	0	6	12	52	15	3	9	3	1	0.9	9	114
31	18	8.3	205	0	7	2	88	27	5	6	2	0.2	0.9	9	123
32	34	7.4	2668	415	596	16	190	41	21	370	30	3	8	20	1399
33	40	7.6	469	264	13	10	186	44	18	19	10	0.3	0	28	279
34	35	6.6	446	49	104	6	20	7	1	75	3	7	1	28	260
35	52	7.2	8160	435	10	28	113	27	11	133	10	2.1	0	80	565
36	38	7.4	16630	362	154	370	396	127	19	150	17	1	0	60	1017
37	38	7.7	9800	353	35	36	158	54	5	86	9	0	0	40	638

**Data Set - II : Peninsular Geothermal Data**

NUM	Temp_C	pH	SPCMHO/	HCO3 mg/	Cl mg/L	SO4 mg/L	TotHard	Ca mg/L	Mg mg/L	Na mg/L	K mg/L	F mg/L	Bmg/L	SiO2 mg/l	TDS mg/L
1	32	7.5	5090	154	1375	210	872	204	88	660	18	0.7	0	18	2981
2	44	7.3	1115	339	165	24	270	82	16	110	6	0.4	0.9	46	500
3	40	7.7	1010	315	130	33	0	110	12	70	25	0	0	58	630
4	0	7.3	1410	390	195	75	0	65	40	210	5	0	0.1	40	830
5	26	7.8	1050	500	140	5	340	70	40	130	2	1	1.2	30	650
6	35.5	7.1	550	290	50	5	230	60	20	30	1	0.3	1.2	30	320
7	40	7.1	28980	190	1347	5	0	390	250	6810	55	0	0	16	0
8	0	7.2	960	410	110	25	0	45	15	95	2	0	0	4.5	0
9	43.5	7.4	8350	150	2725	10	0	105	40	1900	30	0.2	3	31	4790
10	99	0	0	1534	2428	672	0	9	8	1167	145	0	0	0	5744
11	40	7.4	4190	195	1485	0	0	90	40	875	14	0	0	25	0
12	39	7.6	550	183	71	33	160	40	21	40	2	0	0	22	328
13	54	7.5	13620	13	4800	185	4680	186	10	955	13	0	0.4	87	1614
14	43	9	2950	11	850	130	432	170	0.1	368	7	2	0.4	50	1868
15	64	8.6	4950	14	1210	144	890	348	0.2	391	8.5	7.2	0	65	2704
16	35	8.3	883	18	78	242	109	40	15	155	2	2.5	0	60	563
17	35	8	1917	71	426	107	100	32	6	292	4	1.5	1	5	965
18	61	7.6	1457	30	375	100	147	56	1.8	231	7.8	4	0.4	122	955
19	0	8	0	63	265	108	210	80	44	148	6	0.1	0	60	188
20	91	0	0	177	67	70	0	3	1	133	0	3	0.5	96	511
21	0	7.5	0	364	30	8	100	35	3	110	16	0.3	0	57	0
22	0	7.4	0	99	457	128	100	42	2	360	19	0.5	0	70	0
23	0	8	0	366	257	55	530	96	70	98	15	0.2	0	45	855
24	33	7.8	765	171	50	120.6	0	50	7.9	95	7.4	4	0	35	484.5
25	29	7.6	1077	128.6	166	182	0	20	13.4	208	4	5	0	28	756

**Presenting the results and Visualizing the results in a graph**

**Extra-Peninsular India**

Dependant Variable =Cl Independent variables = HCO<sub>3</sub>, SO<sub>4</sub>, Na F

R	R <sup>2</sup>	Adjusted R <sup>2</sup>	Standard error of the estimate
0.92	0.84	0.82	120.07

**ANOVA**

Model	df	F	p
Regression	4	42.1	<.001

**Coefficients**

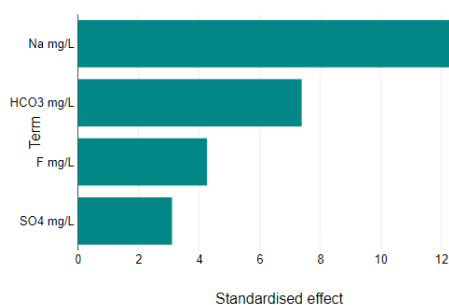
Model	Unstandardize d Coefficients	Standardized Coefficients	Standar d error	t	p	95% confidence interval for B	
	B	Beta				lower bound	upper bound
(Constant)	103.2		32.45	3.18	.003	37.08	169.32
HCO <sub>3</sub> mg/L	-0.6	-0.66	0.08	-7.38	<.001	-0.77	-0.44
SO <sub>4</sub> mg/L	-0.22	-0.25	0.07	-3.09	.004	-0.37	-0.08
Na mg/L	2.36	1.29	0.19	12.61	<.001	1.97	2.74
F mg/L	-24.81	-0.33	5.83	-4.25	<.001	-36.7	- 12.92

*B*= This value represents the slope of the line between the predictor variable and the dependent variable; *SE B*= standard error for the unstandardized beta, similar to the standard deviation for a mean. The larger the number, the more spread out the; (*β*)= the standardized beta similar to a correlation coefficient, ranging between 0 and ±1; *t*= the *t*test statistic calculated for the individual predictor variable and used to calculate the *p* value; *p* = the probability level to tell whether or not an individual variable significantly predicts the dependent variable points.





Pareto diagram of standardized effects



**Peninsular India**

Dependant variable = HCO<sub>3</sub> Independent variable - Na , K

R	R <sup>2</sup>	Adjusted R <sup>2</sup>	Standard error of the estimate
0.87	0.76	0.74	156.16

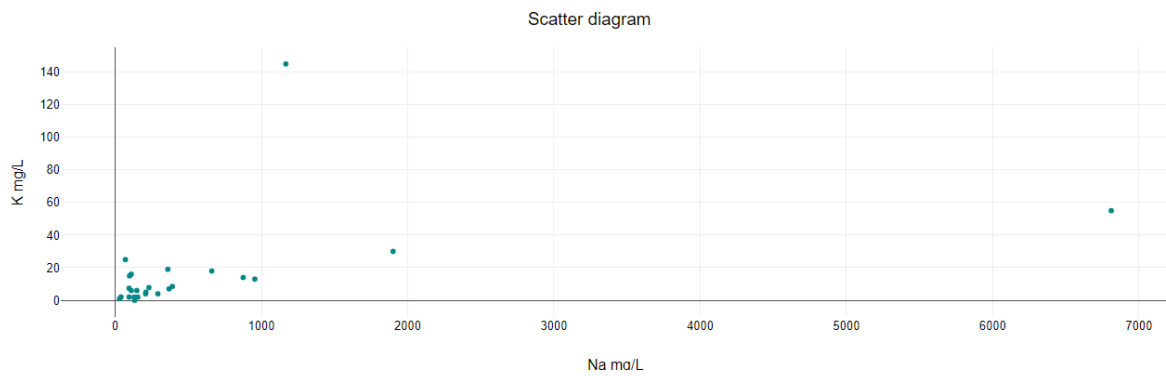
**ANOVA**

Model	df	F	p
Regression	2	34.39	<.001

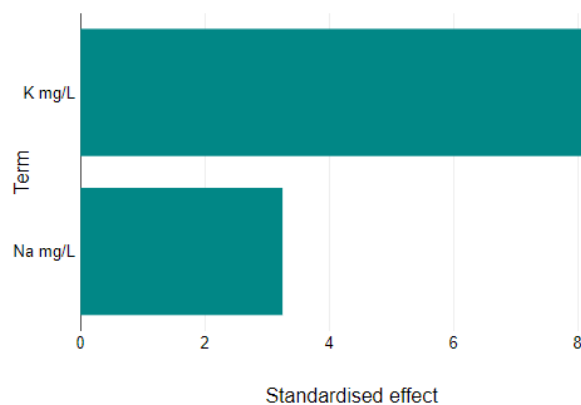
**Coefficients**

Model	Unstandardize	Standardized	Standar	t	p	95% confidence	
	d	Coefficients				error	interval for B
	B	Beta	d error			lower	upper
(Constant)	134.62		36.9	3.65	.001	58.1	211.15
Na mg/L	-0.08	-0.38	0.03	-3.25	.004	-0.14	-0.03
K mg/L	9.94	0.96	1.2				

*B*= This value represents the slope of the line between the predictor variable and the dependent variable; *SE B*= standard error for the unstandardized beta, similar to the standard deviation for a mean. The larger the number, the more spread out the; ( $\beta$ )= the standardized beta similar to a correlation coefficient, ranging between 0 and  $\pm 1$ ; *t* = the *t*test statistic calculated for the individual predictor variable and used to calculate the *p* value; *p*= the probability level to tell whether or not an individual variable significantly predicts the dependent variable points.



Pareto diagram of standardized effects





## **Interpretation of the results**

### **Extra-Peninsular India**

A multiple linear regression analysis was performed to examine the influence of the variables HCO<sub>3</sub> mg/L, SO<sub>4</sub> mg/L, Na mg/L, and F mg/L on the variable Cl mg/L.

#### **Model Summary**

The regression model showed that the variables HCO<sub>3</sub> mg/L, SO<sub>4</sub> mg/L, Na mg/L, and F mg/L explained 84.03% of the variance from the variable Cl mg/L. An ANOVA was used to test whether this value was significantly different from zero. Using the present example, it was found that the effect was significantly different from zero (meaning statistically significant),  $F= 42.1$ ,  $p= <.001$ ,  $R^2=0.84$ .

#### **Regression Coefficients**

The following regression model is obtained:

$$\text{Cl mg/L} = 103.2 - 0.6 \cdot \text{HCO}_3 \text{ mg/L} - 0.22 \cdot \text{SO}_4 \text{ mg/L} + 2.36 \cdot \text{Na mg/L} - 24.81 \cdot \text{F mg/L}$$

When all independent variables are zero, the value of the variable Cl mg/L is 103.2

If the value of the variable HCO<sub>3</sub> mg/L changes by 1 unit the value of the variable Cl mg/L changes by -0.6.

If the value of the variable SO<sub>4</sub> mg/L changes by 1 unit the value of the variable Cl mg/L changes by -0.22.

If the value of the variable Na mg/L changes by 1 unit the value of the variable Cl mg/L changes by 2.36.

If the value of the variable F mg/L changes by 1 unit the value of the variable Cl mg/L changes by -24.81.

#### **Standardized regression coefficients**

The standardized coefficients beta is independent of the measured variable and is always between -1 and + 1. The larger the amount of Beta, the greater the contribution of the respective independent variable to explain the dependent variable Cl mg/L. In this model, the variable Na mg/L has the greatest influence on the variable Cl mg/L.

#### **P-value**

The calculated regression coefficients refer to the sample used for the calculation of the regression analysis, therefore it is of interest whether the individual coefficients only deviate from zero in the population. To test this, the null hypothesis is made for each Coefficient that it is equal to zero in the population.

The standard error now indicates how much the respective Coefficient will scatter on average when the regression analysis is calculated for a further sample.

The test statistic t is then calculated from the standard error and the coefficient.

The p-value for the coefficient of HCO<sub>3</sub> mg/L is  $< .001$ . Thus the p-value is smaller than the significance level of 0.05 and the null hypothesis that the coefficient of HCO<sub>3</sub> mg/L is zero in the population is rejected. Thus it is ensured that the coefficient for the variable age in the population is different from zero.

The p-value for the coefficient of SO<sub>4</sub> mg/L is  $.004$ . Thus the p-value is smaller than the significance level of 0.05 and the null hypothesis that the coefficient of SO<sub>4</sub> mg/L is zero in the population is rejected. Thus it is ensured that the coefficient for the variable age in the population is different from zero.

The p-value for the coefficient of Na mg/L is  $< .001$ . Thus the p-value is smaller than the significance level of 0.05 and the null hypothesis that the coefficient of Na mg/L is zero in the population is rejected. Thus it is ensured that the coefficient for the variable age in the population is different from zero.

The p-value for the coefficient of F mg/L is  $< .001$ . Thus the p-value is smaller than the significance level of 0.05 and the null hypothesis that the coefficient of F mg/L is zero in the population is rejected. Thus it is ensured that the coefficient for the variable age in the population is different from zero.

### **Peninsular India**

A multiple linear regression analysis was performed to examine the influence of the variables Na mg/L and K mg/L on the variable HCO<sub>3</sub> mg/L.

#### **Model Summary**

The regression model showed that the variables Na mg/L and K mg/L explained 75.77% of the variance from the variable HCO<sub>3</sub> mg/L. An ANOVA was used to test whether this value was significantly different from zero using the present example, it was found that the effect was significantly different from zero (meaning statistically significant),  $F= 34.39$ ,  $p= <.001$ ,  $R^2=0.76$ .

#### **Regression Coefficients**

The following regression model is obtained:

$$\text{HCO}_3 \text{ mg/L} = 134.62 - 0.08 \cdot \text{Na mg/L} + 9.94 \cdot \text{K mg/L}$$

When all independent variables are zero, the value of the variable HCO<sub>3</sub> mg/L is 134.62

If the value of the variable Na mg/L changes by 1 unit the value of the variable HCO<sub>3</sub> mg/L changes by -0.08

If the value of the variable K mg/L changes by 1 unit the value of the variable HCO<sub>3</sub> mg/L changes by 9.94.

#### **Standardized regression coefficients**

The standardized coefficients beta is independent of the measured variable and is always between -1 and + 1. The larger the amount of Beta, the greater the contribution of the respective independent variable to explain the dependent variable HCO<sub>3</sub> mg/L. In this model, the variable K mg/L has the greatest influence on the variable HCO<sub>3</sub> mg/L.

P-value

The calculated regression coefficients refer to the sample used for the calculation of the regression analysis, therefore it is of interest whether the individual coefficients only deviate from zero in the population. To test this, the null hypothesis is made for each Coefficient that it is equal to zero in the population.

The standard error now indicates how much the respective Coefficient will scatter on average when the regression analysis is calculated for a further sample.

The test statistic t is then calculated from the standard error and the coefficient.

The p-value for the coefficient of Na mg/L is .004. Thus, the p-value is smaller than the significance level of 0.05 and the null hypothesis that the coefficient of Na mg/L is zero in the population is rejected. Thus it is ensured that the coefficient for the variable age in the population is different from zero.

The p-value for the coefficient of K mg/L is < .001. Thus the p-value is smaller than the significance level of 0.05 and the null hypothesis that the coefficient of K mg/L is zero in the population is rejected. Thus it is ensured that the coefficient for the variable age in the population is different from zero.

### Conclusion

From Multiple regression analysis two distinct geothermal statistically significant suites, emerge:

1. The overall salt assemblage and concentration of Cl- HCO<sub>3</sub>-SO<sub>4</sub> - Na -F or, chloride rich water suggest the existence of the hydrothermal magmatic system operating in geotherms of Extra-Pensular India.

2. Peninsular springs are K-Na-HCO<sub>3</sub> bi-carbonate rich waters with low SO<sub>4</sub>- content and relatively higher contents of HCO<sub>3</sub> compared to other anions SO<sub>4</sub>, Cl, and F suggestive of a non-volcanic origin (F. Tassi et al.,2010; H. Baioumy, 2015)

3. Extra Peninsular springs are magmatic-hydrothermal manifestations, a phenomenon of magma progressively degassing in their decreasing order of solubility CO<sub>2</sub> < SO<sub>2</sub> <HCl< HF i.e."CO<sub>2</sub>-first till HF-last" (Giggenbach 1987).

4. In the case of the non-magmatic thermal springs of Peninsular India, the water is heated by convective circulation: groundwater percolating downward through the fracture, faults reaching great depths of a kilometer or more where the temperature of rocks is high because of the normal temperature gradient of the Earth's crust—about 30 °C (54 °F) per kilometer in the first 10 km.

5. Both Exploratory Factor (AmitabhaRoy, 2023) and Multiple Regression analysis corroborate each other in deciphering the origin of these two suites of fluid geochemistry. The regression analysis revealed two statistically significant suites of fluid geochemistry – 1. The overall salt assemblage and concentration of Cl-HCO<sub>3</sub>-SO<sub>4</sub>-Na-F or chloride-rich acidic waters suggestive of the existence of the deep-seated hydrothermal magmatic-system operative in the geothermal fields in Extra-Peninsular India and 2. Peninsular springs of K-Na-HCO<sub>3</sub> bi-carbonate-rich alkaline waters with low SO<sub>4</sub>- content and high HCO<sub>3</sub> content- compared to other anions SO<sub>4</sub>, Cl, and F suggestive of a shallower non-magmatic origin.

6. **Literature Review:** At the backdrop of the cited literature in relevance to the present research topic, I may say while the objective may be the same, the gaps in knowledge and unresolved problems that are lacking in their studies have been addressed in my research by adopting a definitive approach of geostatistical/mathematical model study giving an insight into arriving at distinguishing two suites of the geotherms. The present technique differs from the conventional methods adopted by the earlier researchers, making comparison simpler and easier to follow between the fluid geochemistry inherent in two distinctive geologic-tectonic environs - one very tectonically active Alpine-Himalayan Extra-Peninsular and the other ancient Proterozoic mobile belt of the Central Highland regions in otherwise very stable landmass or Peninsular shield of India.

7.**Bio-medical Impact:** The bio-medical impact lies in the fact that the chloride and bicarbonate concentrations in geothermal waters share an inverse reciprocal acid-base relationship during either acidosis or alkalosis. These relationships are, in fact, due to the red cell chloride shift. Chloride-rich mineral waters are used as curative agents for digestive disorders like dyspepsia, irritable bowel syndrome, and constipation by breaking down food as a component of hydrochloric acid. Bi-carbonate neutralizes stomach acid, which makes it an effective antacid. On the flip side, bi-carbonate raise the pH of the water and cause havoc in soil and plants.

### References cited

- [1]. Amitabha Roy, 2023.Comparative Statistical Study of Geochemistry of Geothermal Fields of Peninsular and Extra-Peninsular India. J. Appl. Geol. &Geophys.(ISOR-JAGG), v.17,Issue 1,Ser. II, pp. 32-44.
- [2]. A.Roy, 1984.A Computer-Based Factor Model to Elucidate Secondary Trace-Element Distribution Patterns around the Sargipalli Lead-Zinc Sulphide Deposit, Sundergarh District, Orissa (India), J. Geol. Soc.Ind., Volume 25, Issue 6.
- [3]. F. Tassi et al., 2010. Fluid geochemistry of hydrothermal systems in the Arica-Parinacota, Tarapaca, and Antofagasta regions (northern Chile). J. Volcanol.Geotherm.Res.
- [4]. H. Baioumy, 2015. Geochemistry and geothermometry of non-volcanic hot .... J.Volcanol. Geotherm.Res.
- [5]. MiroslawGrzesik, 2022. Table containing the values measured and calculated from the regression equation as well as the statistical tests on the regression equation and its coefficients.Institute of Chemical Engineering, Polish Academi of Sciences.
- [6]. Ravi Shankar et al, 1991. Geothermal Atlas of India, GSI Spec Publ.
- [7]. WF Giggenbach et al, 1987.The effects of hydrothermal processes on the chemistry of some recent volcanic gas.