IOSR Journal of Applied Geology and Geophysics (IOSR-JAGG)
e-ISSN: 2321–0990, p-ISSN: 2321–0982.Volume 13, Issue 3 Ser. I (May. – June. 2025), 51-62
www.iosrjournals.org

Trend Surface Mapping of Principal Component Factor Scores to Assess the Geothermal Geochemistry of India's Peninsular and Extra-Peninsular Hot Springs

Amitabha Roy[1], Ex-Senior Director, Geological Survey of India;
Sk Mafizul Haque[2], Assistant Professor, Cartography, RS, and GIS,
Geography Department, University of Calcutta, India

Abstract
This research paper presents a comprehensive analysis of the geothermal geochemistry of hot springs in India, focusing on two distinct regions: the tectonically active Extra-Peninsular Himalayan region and the relatively stable Late-Precambrian mobile belts of Peninsular India. Utilizing multivariate statistical techniques, particularly Principal Component Analysis (PCA) and Trend Surface Analysis (TSA), the study investigates approximately 340 hot springs by examining their geochemical data. The PCA results identify five significant factors that account for 84.037% of the total variance in the geothermal data, facilitating the classification of geochemical components into distinct groups. The TSA further enhances the understanding of spatial patterns in the geothermal geochemistry by providing a graphical representation of the computed PCA factor scores, revealing the distribution of geochemical variables across the study area. The findings of this study contribute valuable insights into the geothermal potential of India, highlighting the interplay between geological settings and geothermal characteristics, and serve as a foundation for future geothermal energy development initiatives in the region.
Keywords: *PCA analysis, Factor scores, Trend surface analysis, geothermal geochemistry, Peninsular and Extra-Peninsular India*

Date of Submission: 03-06-2025 Date of Acceptance: 13-06-2025

I. Introduction

This study expands on previous research that examined spatially dependent multivariate geothermal hot spring data from two regions with diverse geologic-tectonic settings: a 2,400 km-long arc in the tectonically active Extra-Peninsular Himalayan region and the relatively stable Late-Precambrian or Proterozoic mobile belts in Peninsular India's Central Highlands. employing a variety of robust multivariate statistical techniques. to spatially dependent multivariate geothermal geochemistry data from hot springs in two regions with. There are around 340 hot springs located throughout India, including the Peninsular and Extra-Peninsular regions. Schlagintweit compiled the first list of hot springs in India in 1852. In 1991, the Geological Survey of India produced a 'Geothermal Atlas of India' (Ravi Shanker), and the Indian government established a 'Hot Spring Committee' to investigate the development of geothermal power plants. Projects in the Puga Valley and Parvati aim to harvest 5000 MWh of geothermal energy, which is enough to run a 20 MWe power plant (Jonathan Craig, 2013). GTHERMIS is a computerised geothermal database system developed by the GSI (A. Roy, 1994).

The current study explores the multivariate geospatial fluid geochemistry of geothermal hot spring data in two ways: statistically and graphically using the GIS approach.

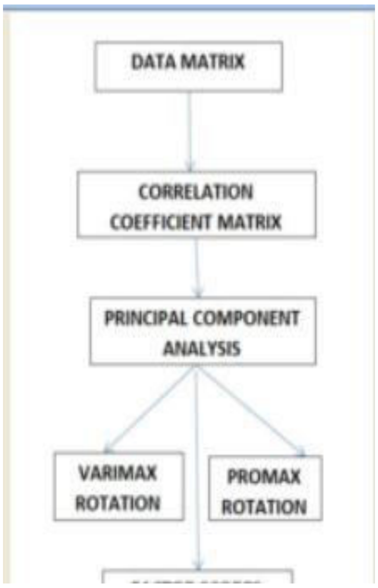
Software choice

Principal component analysis (PCA) was performed in Excel using the XLSTAT statistical software's 14-day trial add-on feature located under the Analyse Data menu. Dr Sk Maffizul Haque performed interpolation by polynomial trend surface analysis of statistically computed data using ArcGIS software in his Cartography, RS, GIS laboratory at Calcutta University's Geography Department.

II. PCA Mathematics

Principal Component Analysis (PCA)
Observations/variables table: Workbook = GFACTREV.xlsx / 181 rows and 9 columns
PCA type: Pearson Correlation
Filter factors: Maximum number = 5
Standardisation: (n)

Table I. Computational strategy / Flow chart of Principal Component (PCA)



Algorithm 1 Principal Component Analysis

```
1: function PCA(X)
2:   Compute the correlation matrix by:
```

$$r_{xy} = \frac{\sum(x_i - x)(y_i - y)}{\sqrt{\sum(x_i - \bar{x})^2 \sum(v_i - \bar{v})^2}}$$

```
3:   Compute the eigenvalues and eigenvectors of Σx
4:   Define the transformation matrix T = [w1, w2, ..., wd] with the d eigenvectors associated to the d
     largest eigenvalues.
5:   Project the data X into the PCA subspace:
       yi = Txi   for   i = 1, 2, ..., n
6:   return Y
7: end function
```

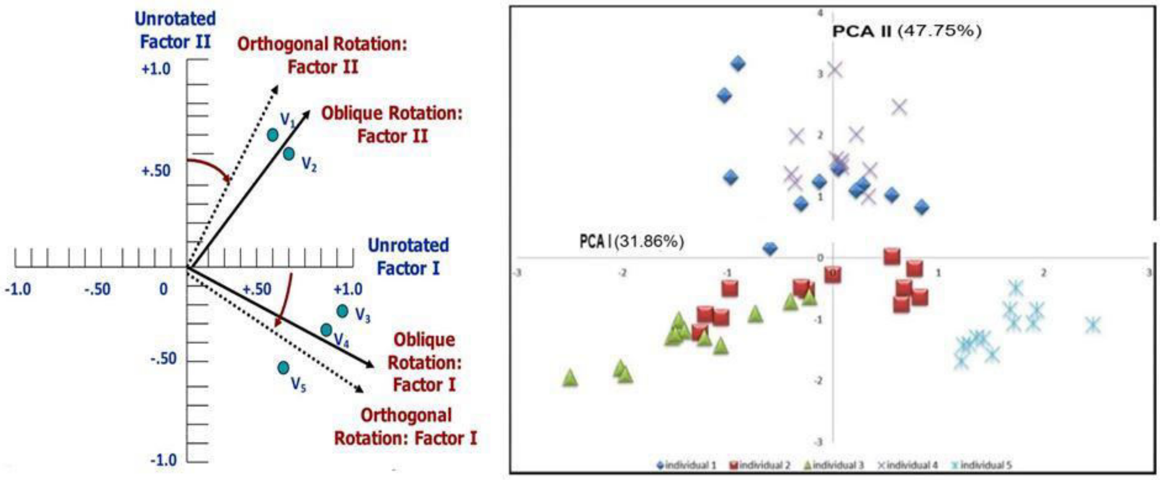
F

FACTOR SCORES (F1.F2.F3.F4.F5)

↓

TREND SURFACE ANALYSIS

Fig.1.Schematicdiagramof unrotated PCA in relation to rotated Varimax&Promax



Principal component analysis (PCA) is the foundation of a statistical approach for data reduction, or 'dimension reduction' of a dataset with multiple variables as input. It generates a dataset in a lower subspace that is a new reduced orthogonal or uncorrelated variable known as principal components, which are obtained as linear combinations of the original variables, thereby extracting the principal components containing the most significant information and rejecting noise or insignificant data while preserving all critical details. Recomputing a set of factor scores involves assigning values to these new variables based on observations. The correlation coefficient matrix is the initial stage in PCA analysis (Tables 2-6).

Unrotated Principal Component Analysis (PCA) brings about a linear orthogonal transformation of m original variables, geochemical elements in the m-dimensional measurement space to m new statistically independent variables or principal components where each new variable is a linear combination of the old. The PC analysis extracts m-eigenvectors (principal component axes) and corresponding m-eigenvalues (the variance measured along the eigenvector) from m x m symmetrical matrix of correlation. The results of PC analysis is given in Table-4B.3, and 4B.4. The columns of this PCT matrix are all orthogonal and hence inter-column correlations are near Zero. These columns represent eigenvectors (directions) and eigenvalues (coefficients attached to eigenvectors) from the correlation coefficient matrix. The eigenvalues account for all of the original data variances in the decreasing order such that each has variance or eigenvalue less than previous ones. The principal components are then converted into factors scores by multiplying each element of the principal components or eigenvectors (V) by the square root of the corresponding eigenvalues (that is, $H=\lambda^{1/2} V$). These, besides the direction, also represent the variances.

In the Exploratory Factor Analysis the most baffling issue is to determine the number of factors to retain significant few from many insignificant ones. There are different criteria suggested by different researchers namely,

● The Eigenvalue > 1

● From the inflection point of scree or Elbow curve plot i.e. a graph of the eigenvalues (Y-axix) against the factors (X-axis) listed in the descending order (Fig. 4B.2).

● Communalities for each of the variables (somewhat like R2 from Regresion analysis) at least 0.5

● The factor loadings for each variable should be ≥0.6 (Awang, 2014).

In the present study, the scree plot was used for the selection of five factors which accounts for 84.037% of the total variance of the original varables “load”on a factor of Principal component is retained.

Presenting the output results of PCA

Table 2. Correlation matrix (Pearson (n):

Variables	HCO3 mg/L	Cl mg/L	SO4 mg/L	Ca mg/L	Mg mg/L	Na mg/L	K mg/L	F mg/L	Temp_C
HCO3 mg/L	**1**	0.047	0.143	-0.033	0.113	0.065	0.542	0.021	0.057
Cl mg/L	0.047	**1**	0.188	0.344	0.168	0.429	0.311	-0.159	0.066
SO4 mg/L	0.143	0.188	**1**	0.670	0.126	0.058	0.377	-0.052	0.009
Ca mg/L	-0.033	0.344	0.670	**1**	0.489	0.434	0.286	-0.152	-0.069
Mg mg/L	0.113	0.168	0.126	0.489	**1**	0.757	0.287	-0.092	-0.014
Na mg/L	0.065	0.429	0.058	0.434	0.757	**1**	0.292	-0.098	0.010
K mg/L	0.542	0.311	0.377	0.286	0.287	0.292	**1**	-0.011	0.020
F mg/L	0.021	-0.159	-0.052	-0.152	-0.092	-0.098	-0.011	**1**	0.216
Temp_C	0.057	0.066	0.009	-0.069	-0.014	0.010	0.020	0.216	**1**

Principal Component Analysis:

Table 3. Eigenvalues:

	F1	F2	F3	F4	F5	F6	F7	F8	F9
Eigenvalue	2.867	1.430	1.239	1.111	0.917	0.707	0.359	0.201	0.170
Variability (%)	31.860	15.886	13.766	12.339	10.185	7.852	3.990	2.229	1.891
Cumulative %	31.860	47.747	61.513	73.852	84.037	91.890	95.880	98.109	100.000

Fig. 2. Scree plot

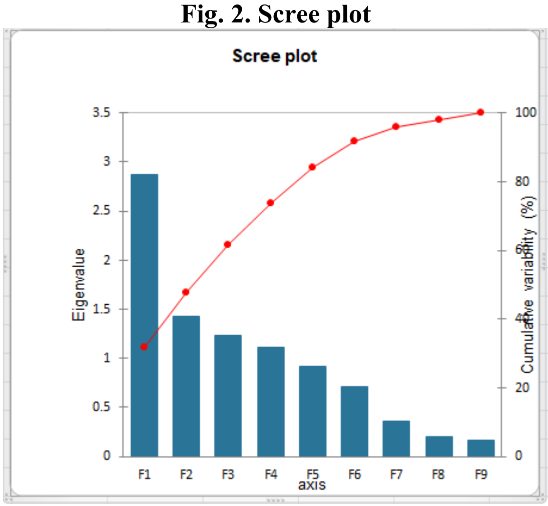


Table 4. Eigenvectors:

	F1	F2	F3	F4	F5
HCO3 mg/L	*0.176*	0.624	-0.079	-0.423	-0.053

Cl mg/L	*0.335*	-0.068	-0.046	0.050	0.711
SO4 mg/L	*0.335*	0.193	0.566	0.349	-0.118
Ca mg/L	*0.461*	-0.169	0.308	0.305	-0.153
Mg mg/L	*0.426*	-0.224	-0.370	-0.061	-0.366
Na mg/L	*0.439*	-0.261	-0.441	-0.059	-0.046
K mg/L	*0.379*	0.473	0.020	-0.221	0.026
F mg/L	*-0.121*	0.335	-0.315	0.485	-0.416
Temp_C	*-0.009*	0.294	-0.381	0.559	0.380

Table 5. Factor loadings:

	F1	F2	F3	F4	F5
HCO3 mg/L	*0.297*	0.746*	-0.088	-0.446	-0.051
Cl mg/L	*0.568*	-0.081	-0.052	0.053	0.681*
SO4 mg/L	*0.567*	0.230	0.630*	0.367	-0.113
Ca mg/L	*0.780**	-0.203	0.343	0.321	-0.146
Mg mg/L	*0.721**	-0.267	-0.412	-0.064	-0.351
Na mg/L	*0.743**	-0.312	-0.491	-0.062	-0.044
K mg/L	*0.642**	0.566	0.023	-0.233	0.025
F mg/L	*-0.204*	0.401	-0.351	0.511	-0.398
Temp_C	*-0.015*	0.352	-0.424	0.589*	0.363

Note: * ≥ 0.6

Table 6. Factor scores:

X_LONG	Y_LAT	F1	F2	F3	F4	F5
87.25	23.52	-1.119	-0.560	0.051	-0.062	0.469
87.25	23.52	-0.870	-0.107	-0.182	0.032	0.326
84.862	19.2647	-0.117	1.206	-1.322	1.228	-0.930
842350	19.2905	-1.452	0.679	-1.050	1.483	-0.395
84.0212	23.4515	-0.931	-0.473	0.245	0.000	0.645
84.294	23.492	-0.789	-0.497	0.037	-0.168	0.354
85.25	25.01	-1.391	-0.787	0.083	0.008	0.608
85.41	24.09	-1.038	-0.611	0.183	0.030	0.669
93.15	28.2	-0.618	0.100	-0.055	-0.552	0.370
93.26	28.25	0.609	0.038	1.176	0.177	-0.218
93.25	28.273	-0.626	-0.236	0.382	-0.617	0.227
77.1502	32.3295	-0.605	-0.750	0.913	-1.639	-0.239
77.3473	32.0278	-1.143	0.938	-0.896	1.369	-2.812
78.37	32.0345	5.859	1.859	2.681	1.704	-0.916
77.47	31.314	-0.467	-0.267	0.243	-0.384	0.115
77.582	31.342	-0.538	-0.147	0.201	-0.758	0.535
78.105	31.3	0.640	0.372	1.326	-0.412	-1.205
78.223	31.35	-1.439	0.802	-1.123	1.683	-0.650
78.23	31	-0.291	0.696	-0.803	0.920	1.240
78.25	30.572	-0.884	-0.077	-0.312	0.682	0.132
78.3336	30.5502	-1.413	-1.051	0.444	-0.474	0.015
78.4012	30.5325	-1.247	0.750	-0.888	1.253	-0.467

78.433	30.5325	-0.758	0.037	-0.455	0.431	-0.587
79.0135	30.3905	-0.356	0.721	-1.364	0.741	-1.255
79.293	30.4445	0.649	0.344	-1.126	1.323	2.282
79.352	30.4158	-0.384	-0.416	-0.719	0.646	0.303
79.313	36.325	-1.012	-0.523	0.156	0.067	0.526
79.481	30.36	1.150	1.357	-1.793	-0.996	-1.217
79.373	30.293	-0.547	-0.251	-0.197	-0.104	0.248
80.5313	30.0527	-1.300	-0.630	-0.147	0.252	0.462
80.0856	29.58	0.977	2.455	-0.923	-1.731	-0.024
80.2023	30.0853	0.102	0.845	0.310	-1.667	-0.209
80.3384	29.515	-0.362	0.634	-0.871	0.420	0.603
773240	34.465	-0.995	0.258	-0.221	0.535	0.377
77.333	34.4525	-0.646	0.100	-0.058	-0.018	0.481
77.2825	34.564	0.742	2.962	-0.207	-2.953	0.010
75.0345	33.162	3.965	-0.289	5.407	3.080	-1.600
75.443	33.431	0.493	-0.066	0.743	1.530	-0.079
75.523	33.303	-1.493	-0.328	-0.334	0.721	0.248
76.562	33.204	-1.491	0.995	-1.292	2.358	-0.429
78.195	33.13	-1.565	-0.185	-0.349	0.864	-0.972
78.21	33.22	-1.525	0.266	-1.021	1.866	-0.349
76.0425	32.421	0.198	0.256	-0.364	-0.574	0.415
76.105	32.0755	-0.624	-0.806	0.632	-1.234	-0.373
76.431	32.071	-1.195	-1.203	0.694	-0.938	-0.009
77.04	28.15	2.322	-0.583	-1.087	1.824	0.204
77.04	28.15	-0.348	-0.777	0.711	-1.122	-0.423
76.193	26.353	-0.016	-0.444	0.834	-1.007	-0.190
76.53	27	-0.017	-0.299	-0.186	-0.304	-0.482
73.411	24.11	-0.063	-0.329	-0.196	-0.736	-0.113
73.424	24.13	-0.848	-0.022	-0.371	0.351	-0.779
72.1544	21.4053	9.517	-4.682	-5.473	-0.689	-2.277
72.1544	21.4053	-0.589	-0.197	0.035	-0.494	0.252
72.12	22.15	2.686	-1.547	-1.170	-0.335	2.630
72.41	22.14	4.631	4.078	1.452	-3.438	1.704
71.4331	23.4329	1.129	-1.320	-0.410	-0.568	1.003
73.56	23.2	-0.699	-0.797	0.221	-0.419	0.051
72.51	19.423	3.115	-1.792	0.340	0.827	5.059
72.543	19.4105	0.391	-1.774	1.588	-0.517	0.303
73.05	19.293	1.488	-1.494	1.395	1.238	1.125
73.27	18.04	-0.508	-0.130	-0.286	1.732	1.179
73.24	17.43	-0.530	-0.903	0.362	-0.076	0.702
73.35	17.15	-0.511	-0.928	0.545	0.002	0.658
76.1666	21.4196	0.074	-0.985	-0.068	0.361	0.262
74.25	23.1435	-1.077	0.505	-0.827	1.295	0.976
80.4315	17.55	-0.925	0.948	-0.825	0.842	-0.086
80.563	17.383	-0.301	-0.415	0.325	0.040	0.384
80.4425	17.5545	0.655	-0.554	-0.298	-0.503	-0.593
80.43	17.56	-0.740	0.095	-0.163	0.895	0.139

80.4	18.06	-0.511	-1.222	1.170	-1.119	-0.397
83.6839	23.6988	-0.152	2.609	-1.334	-0.258	1.463
90.43	27.51	2.776	-0.082	2.713	-0.419	-0.418
85.23	27.67	0.370	-0.957	0.760	-1.424	-0.082
81.8661	21.2787	-0.020	1.121	0.502	-2.968	0.204
75.5	33.291	-1.299	-0.523	0.017	0.063	-0.076
84.4881	20.8453	0.357	0.831	-0.040	-0.574	-0.719
85.2066	20.9573	1.636	0.971	1.402	-2.196	-0.658
85.5625	20	-0.778	-0.968	1.173	-1.421	-0.420
84.294	23.492	-1.014	-0.726	0.529	-0.186	0.185
85.19	24.55	-1.364	-1.461	0.949	-1.278	-0.267
85.294	25.0015	-1.370	-1.458	0.964	-1.270	-0.254
73.8413	18.4916	-0.553	-0.974	1.063	-1.448	-0.574
78.37	32.0345	5.806	2.093	2.427	2.085	-0.967
77.47	31.314	1.620	0.635	-0.457	0.385	0.586
78.002	31.322	-0.283	0.775	-0.330	0.229	-0.195
78.0525	31.312	-0.750	1.444	-1.110	1.782	-0.266
78.105	31.3	0.631	0.990	0.559	0.756	-0.458
78.1945	31.363	-1.340	0.482	-0.646	1.289	-0.626
78.203	313615	-0.947	-0.780	0.618	-0.536	-0.187
78.223	313500	-1.368	0.494	-0.791	1.183	-0.576
78.23	31	-0.350	0.813	-0.903	1.061	0.880
78.263	30.584	6.171	1.713	2.608	0.341	-1.394
78.25	30.572	-0.803	-0.130	-0.329	0.691	0.676
78.1415	30.53	0.248	-0.162	0.222	-0.867	-0.016
78.2855	30.504	-0.185	-0.302	0.157	-1.132	-0.739
78.41	30.56	-0.339	2.279	-1.249	0.610	-0.339
78.41	30.542	-0.420	1.735	-0.853	0.019	-0.446
78.4012	30.5325	-1.077	0.204	-0.345	0.426	0.006
78.3912	30.5145	-0.661	1.005	-0.579	-0.189	-0.216
78.3505	33.453	0.394	0.458	-0.510	-1.145	-0.643
78.43	30.5325	-0.575	-0.336	-0.157	-0.041	0.251
78.5412	30.4447	-0.161	0.922	-0.510	-0.568	-0.480
79.0135	30.3905	-0.049	-0.052	-0.675	-0.344	-0.181
79.293	30.4445	0.710	-0.259	-0.392	0.237	1.835
79.352	30.4158	-0.354	-1.092	0.146	-0.625	-0.497
79.313	36.325	-1.071	-0.001	-0.473	0.999	0.869
79.2902	30.2902	0.439	0.514	-0.684	-0.645	-0.430
80.013	30.434	1.447	2.557	0.015	-2.317	0.066
79.481	30.36	-0.005	1.838	-0.592	-1.039	0.099
80.5313	30.0527	-1.264	-1.001	0.307	-0.419	0.171
79.5556	29.593	0.037	-0.608	-0.363	-1.394	-0.840
80.0752	29.5832	-0.885	-0.088	0.478	-1.226	-0.024
80.1509	30.0806	-0.553	1.493	-0.854	0.139	-0.177
80.173	30.0308	0.830	0.317	-0.464	-0.607	-1.097
80.404	30.034	-0.869	-0.610	0.520	-0.478	0.031
80.414	29.5825	-0.231	-0.800	0.578	-0.870	-0.505

80.3384	29.515	-0.392	0.506	-0.618	0.128	-0.122
77.324	34.465	-1.319	1.332	-1.297	2.174	-0.476
77.333	34.4525	-0.952	0.566	-0.348	0.467	-1.155
77.2825	34.564	0.406	4.315	-1.642	-0.789	-0.507
74.243	33.143	-0.360	0.423	-0.410	-0.051	0.059
742940	33.123	-0.726	-0.538	-0.142	-0.084	-0.154
74.294	33.123	-0.699	-0.881	0.588	-0.876	-0.254
74.48	33.11	0.655	1.287	-0.471	-0.611	0.099
75.0345	33.162	3.878	0.051	5.049	3.622	-1.750
75.0345	33.162	-0.427	-1.014	1.437	-0.556	-0.404
75.072	33.17	1.127	1.395	-0.407	-0.079	0.809
75.443	33.431	0.577	-0.988	1.879	-0.148	-0.845
75.5	33.291	-1.177	-1.114	0.683	-0.833	-0.100
75.523	33.303	-1.354	-1.256	0.750	-0.892	-0.102
76.562	33.204	-1.265	0.174	-0.441	1.073	0.043
76.0425	32.421	0.184	0.221	-0.303	-0.661	0.253
76.105	32.0755	-0.639	-0.420	0.134	-0.503	0.102
76.431	32.071	-1.425	-0.278	-0.286	0.541	-0.371
76.193	26.353	-0.018	-0.166	0.467	-0.470	0.215
76.52	27	0.098	-1.037	0.670	-1.580	-0.727
73.4118	24.11	-0.080	-0.453	-0.012	-1.000	-0.438
73.424	24.13	-0.624	-0.639	0.207	-0.539	0.003
72.1544	21.4053	-0.646	-0.869	0.254	-0.466	0.101
72.1544	21.4053	9.537	-4.639	-5.552	-0.578	-2.057
71.05	21	-0.098	0.171	-0.289	0.785	0.493
72.41	22.14	4.590	5.413	-0.278	-0.901	3.426
72.1553	22.4818	-0.282	-0.431	-0.048	-0.654	0.952
72.2157	23.495	-0.404	-0.281	0.104	-0.638	0.309
71.4331	23.4329	1.157	-1.225	-0.567	-0.344	1.362
73.56	23.2	-0.691	-0.771	0.178	-0.356	0.156
72.51	19.423	3.109	-1.603	0.095	1.186	5.303
72.543	19.4105	0.306	-1.008	0.662	0.855	0.820
73.05	19.293	1.210	-0.686	0.572	2.424	0.512
73.23	18.05	-0.167	-0.538	0.612	0.819	0.328
73.27	18.04	-0.565	-0.760	0.613	0.428	-0.223
73.113	17.564	0.218	-0.828	0.279	0.987	1.103
73.24	17.43	-0.578	-0.830	0.318	-0.002	0.442
73.24	17.35	0.637	-0.866	-0.020	1.625	1.366
73.38	16.38	-0.823	-0.179	0.143	-0.639	0.358
80.4315	17.55	-0.667	-0.558	0.898	-1.728	-0.402
80.563	17.383	-0.219	-1.073	1.112	-1.127	0.007
80.53	17.51	-0.694	-0.112	0.199	0.499	-0.170
80.53	17.51	-0.720	-0.088	-0.071	0.432	-0.408
80.4425	17.5545	-0.793	0.095	0.219	0.385	-0.262
80.4	17.58	-0.723	-0.445	-0.289	-0.030	-0.672
80.4	18.06	-0.693	-0.373	0.239	0.286	-0.454
80.4	17.554	-0.466	0.157	-0.026	-0.484	-0.446

80.4	18.03	-1.027	-0.554	0.361	0.169	-0.401
81.2	29.6	-1.133	-0.078	-0.222	0.327	-0.574
85.3	28.3	-0.754	0.386	-0.601	1.105	-0.276
80.6	29.9	-1.318	-0.448	0.004	0.127	-0.273
80.6	29.7	-1.419	-0.429	-0.316	0.688	0.889
83.7	29.8	0.293	-0.061	0.071	0.697	-1.816
82.3	29.3	-1.207	-0.096	0.221	0.053	-0.542
82.1	29.2	-1.249	-0.472	0.187	0.580	-0.028
83.3	28.22	-0.003	1.381	-1.132	0.540	-0.187
83.57	27.56	-0.381	0.159	-0.436	0.865	-0.760
82.2	27.55	-0.231	-0.525	-0.645	0.262	-0.077
84.2	28.4	-0.628	-0.260	0.276	0.397	0.314
82.15	27.47	-0.681	0.409	-0.855	0.366	-1.466
83.7	28.5	1.412	2.073	-0.444	0.618	-1.062
77.524	34.798	-0.926	-0.015	0.093	0.122	0.293
74.45	24.7	-1.514	-0.520	-0.101	0.301	-0.386
89.363	27.41	-0.874	0.113	-0.142	-0.333	-0.041
90.5	27.6	-0.229	2.009	-0.684	-0.596	-0.305

III. Trend Surface Analysis

More recent studies furnish a geostatistical explanation of spatially dependent multivariate geothermal data representing two spatially distinctive regions of diverse geologic-tectonic settings: one from a 2400 km long arcuate belt of the tectonically active Extra-Peninsular Himalayan region and the other from Late-Precambrian or Proterozoic mobile belts in the Central Highland in an otherwise stable landmass or shield of Penininsular India. Geothermal hot springs spread over these areas, conspicuously coinciding with the respective tectonic zones of different degrees of severity. Peninsular India's hot springs are restricted to the Proterozoic mobile belts of the Central Highlands, leaving the more stable southern Deccan plateau nearly devoid of any hot springs. In the present study, the two sets of data representing Peninsular and Extra-Peninsular India in entirety were considered for trend surface analysis or trend surface mapping. While mapping, polynomial power was set at 3 or cubic, which gives a better recognisable pattern than the original raw data.

Mathematical consideration

A trend surface model is a particular case of a bivariate regression model with two independent variables, the coordinates X and Y, and a dependent variable, the thematic variable Z, to be modelled. In other words, the principle of a trend surface model is a regression function that estimates the variable value Zi at any location based on the Xi and Yi coordinates of this location. The general function is:

$$Z_i = f\,(Y_i, X_i)$$

Where:

z_i = variable value at location i

Y_i, X_i = geographic coordinate values at location i

f = regression function

One can select a linear regression function (first order) or, if the spatial distribution is more complex, a polynomial function (2nd, 3rd,..., or nth order). The modelled surface will correspond to a flat, oriented plane or a curved surface with an increasing number of polynomial order.The trend surface analysis technique is applied to separate the raw data into two components: the regional or background trend (regression) and the local anomalous values or residuals (deviations). It is in reality a mathematical curve fitting technique that accomplishes "best-fitting' a series of polynomial surfaces by the least squares criterion of multiple regression and determines a value about which the variance is a minimum. The goodness-of-fit of the appropriateness of trend surface curves to the data can be tested by the analysis of variance. A progressive higher-order trend surface fit to the data can be represented by the polynomial equation of arbitrary degree n:

$$Y_i = \sum_{r=0}^{n} a^r X_i^r$$

Where Y_i represents the calculated value in the i-th observation, a^r (r = 0,1,2,...n) indicates arbitrary complex coefficients, and X_i indicates the independent variable in the i-th observation. The 'best' is obtained when the sums of squares of differences between the observed values (y_i) for a particular application (in the present case, factor scores derived from PCA analysis) and the corresponding calculated values (Y_i) are as small as possible. That is

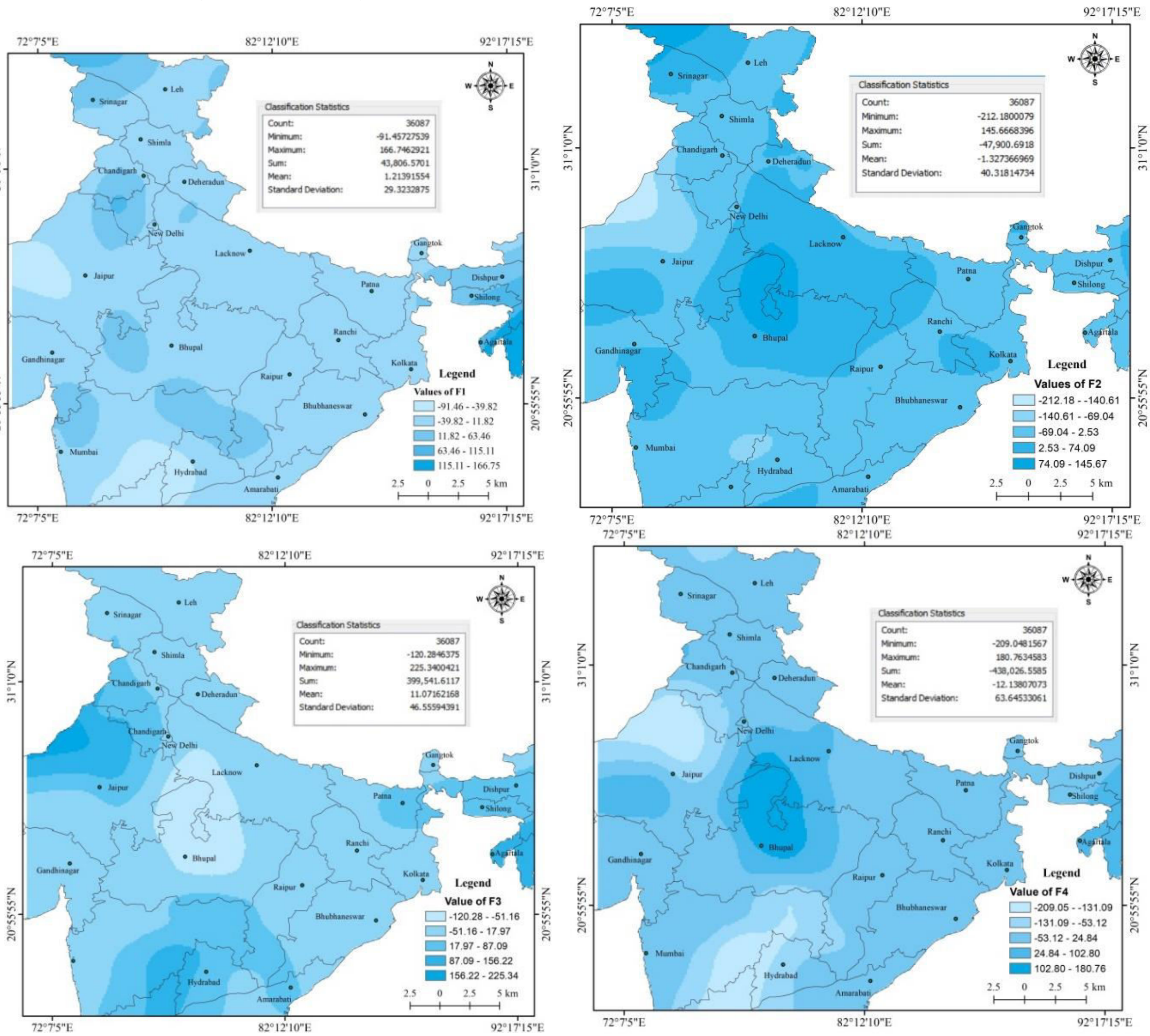
$$\sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} (Y_i - y_i)^2 = \text{minimum}$$

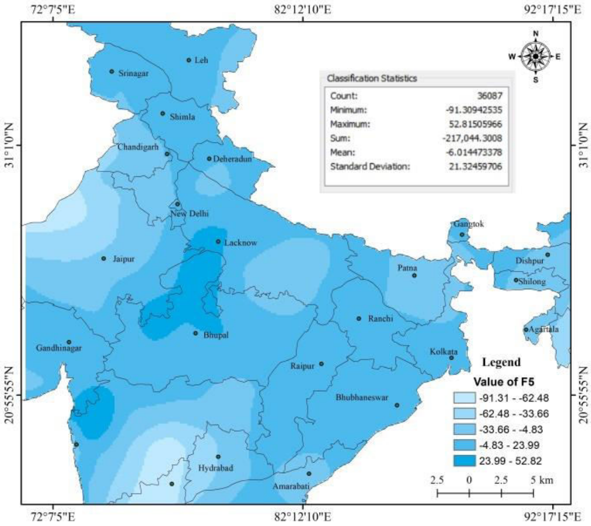
The differences (d_i) are called residuals. Moreover, in the case of two-dimensional trend surface mapping, (X_i) is defined as a function of the independent variables, such as mutually perpendicular geographic coordinates X_i (the easting) and Y_i (the northing). The trend is therefore a linear function of the geographic coordinates of a set of observations, minimising squared deviations from the trend.

IV. Data selection

Two types of data are required for trend surface analysis: geographical location, such as data point coordinates (latitude and longitude), and continuous multivariate numeric variables, which characterise each data point's value and variable measurements. The original dataset, or observations/variables table, or workbook = GTHREV.xlsx, used in this study comprises 181 rows and 11 columns, including two for LAT_LONG and seven for multivariate data. The reduced subspace, TSA_FSCORES.xlsx, obtained from PCA analysis of multivariate geothermal geochemical variables, has two columns for LAT_LONG and five for calculated factor scores data (Table- 6).

Fig. 3. Cubic Polynomial Trend Surface Mapping of PCA Factor Scores





V. Interpretation of results

Outcomes of principal component analysis (pca) and graphical visualization of calculated component scores applying trend surface analysis

PCA Analysis

Based on the scree plot (Fig. 2), we retain five factors that account for 84.037% of the total variance in the original variables "load on a factor of principal component". PCA factors are orthogonal, thus they are uncorrelated; nonetheless, correlation occurs between the variables inside each factor. We regard eigenvalues ≥ 0.6 as significant, indicating the relevance of variables for factors (Table- 5).

F1 (Ca, Mg, Na, K) = Basic/alkaline
F2 (HCO3) = Both acidic and basic (Amitabha Roy, 2024)
F3 (SO4) = Acidic
F4 (Temp▫C)
F5 (Cl) = Acidic

Trend Surface Analysis (TSA)

Figures (Fig. 3) show the real distribution of fluid geochemical variable values within a research region. Outside of the measured region, trend surface estimates become unrealistic. This is referred to as "edge effects." These are the places to avoid. In the current investigation, the polynomial of power 3, i.e., cube, trend surfaces from computed PCA factor scores F1, F2, F3, F4, and F5 (Table-6) provide a more recognisable pattern than the original raw data.

The basic/alkaline group F1 = Ca, Mg, Na, and K appears to be common, although not intense, in India's peninsular hot springs. Hot springs in India's peninsular and extra-peninsular locations have the acid-base group F2 = HCO3 in various degrees of intensity. The intensity of the group F3 = SO4 is rather low in the middle of the central highlands, but it climbs to relatively high levels in western and southern India. The group F4 = Temp▫C displays a 'high' that corresponds to exceptionally low SO_4 concentrations in the central highlands. Further study is needed to understand the inverse relationship between SO_4 and temperature in the central highlands region. The group F5 = Cl is equally prevalent across India's extra-peninsular and peninsular hot springs. Arcuate TSA contours with southeasterly convexity in the western region, particularly for factor scores attributed with HCO3, SO4, Cl, and TEMP, represent a conspicuous resemblance of arcate hills in the Kirther-Sulaiman hills of Pakistan.

VI. Conclusions

In conclusion, the trend surface analysis and principal component analysis (PCA) of the geothermal geochemistry data from India's Peninsular and Extra-peninsular hot springs revealed significant spatial patterns and geochemical relationships.The PCA identified five significant factors, each associated with specific geochemical elements, while the trend surface analysis provided a graphical representation of the spatial distribution of fluid geochemical variables. The findings not only enhance our understanding of the interplay between geological settings and geothermal characteristics but also lay a crucial foundation for advancing sustainable energy initiatives in India.

Acknowledgements

Special thanks to Prof. Sumana Banerjee, Head of the Geography Department, University of Calcutta, Ballygunge Science College, Kolkata, India, for giving access to cutting-edge facilities and resources that were critical in carrying out this research. The authors are also grateful to Dr Pablo Mendez, Customer Support Specialist, Lumivero, and support staff Ashutosh Sharma, NAPL, New Delhi, for giving a 14-day trial license for the installation of Advanced XLSTAT for PCA analysis.

References

[1]. Amitabha Roy, 2024. The book: Geostatistics applied to fluid geochemistry of geothermal fields in peninsular and extra-peninsular India. White Falcon Publishing, Chandigarh, India. 2024. Pp. 1-144. ISBN: 979-8-89222-356-0

[2]. A.Roy, 1994. GTHERMIS – An information management and analysis system for geothermal data of India, A Field season report (1993-94).

[3]. Awang, Z, 2014. Research Methodology And Data Analysis ...

[4]. Jonathan Craig, 2013. The hot springs and the geothermal energy potential of Jammu and Kashmir State Northwest Himalaya, India.

[5]. Ravi Shankar et al., 1991. Geothermal Atlas of India, GSI Spec Publ.