Journal of Advanced Research in Geo Sciences & Remote Sensing
Volume 12, Issue 1&2 - 2025, Pg. No. 34-50
Peer Reviewed Journal

Research Article

ANOVA-Based Evaluation of Radium, Thorium, and Potassium Levels in Soil Samples of Rajpur Region, Dehradun District, Uttarakhand, India

Ganga Sharan[1], Lalit Mohan Upadhyaya[2], Sudhanshu Aggarwal[3]

[1]Associate Professor, Department of Physics, Municipal Post Graduate College, Mussoorie, Dehradun, Uttarakhand, India
[2]Associate Professor, Department of Mathematics, Municipal Post Graduate College, Mussoorie, Dehradun, Uttarakhand, India
[3]Assistant Professor, Department of Mathematics, National Post Graduate College, Barhalganj, Gorakhpur, Uttar Pradesh, India
DOI: https://doi.org/ 10.24321/2455.3190.202501

I N F O

Corresponding Author:
Lalit Mohan Upadhyaya, Department of Mathematics, Municipal Post Graduate College, Mussoorie, Dehradun, Uttarakhand, India
E-mail Id:
lmupadhyaya@rediffmail.com
How to cite this article:
Sharan G, Upadhyaya L M, Aggarwal S. ANOVA-Based Evaluation of Radium, Thorium, and Potassium Levels in Soil Samples of Rajpur Region, Dehradun District, Uttarakhand, India. *J Adv Res Geo Sci Rem Sens* 2025; 12(1&2): 34-50.

Date of Submission: 2025-05-26
Date of Acceptance: 2025-06-11

A B S T R A C T

Based on a recent study on radionuclide concentrations in soils by Tushar Kandari et al. (Journal of Radioanalytical and Nuclear Chemistry, 2021, 330:1545–1557, https://doi.org/10.1007/s10967-021-07988-2), we analyze the average concentrations of Radium-226 (^{226}Ra), Thorium-232 (^{232}Th), and Potassium-40 (^{40}K) in the soil samples from the Rajpur region of Dehradun district, Uttarakhand, India. Our analysis employs a one-way repeated measures ANOVA model technique to evaluate the concentrations of these radionuclides.

Keywords: Radium (^{226}Ra), Thorium (^{232}Th), Potassium (^{40}K), radionuclide, One-Way Repeated Measures ANOVA, Friedman's test, Wilks' lambda, Pillai's trace, Roy's largest root, Hotelling-Lawley trace 2020

Mathematics Subject Classification: 62J02, 62J05, 62J10, 62J99, 62P35, 62P99, 86-05, 86-08, 86-10, 86-11, 92C05.

1. Introduction

Radiation is a form of energy emitted by atoms in the form of electromagnetic waves or moving subatomic particles. It can be classified into ionizing and non-ionizing radiation. Ionizing radiation possesses enough energy to remove tightly bound electrons from atoms, effectively ionizing them.[1,2] This type of radiation includes high-energy ultraviolet light, X-rays, gamma rays, alpha particles, beta particles, neutrons, high-energy protons, and charged atomic nuclei from cosmic rays. In contrast, non-ionizing radiation does not have sufficient energy to ionize atoms or molecules, though it can still excite them, causing electrons to move to higher energy levels.

Natural radiation comes from radioactive materials that are naturally present in the environment, such as in soil, water, and air. The building material including bricks, stone, cement, marble, sand, limestone etc. is the primary source of background radiation due to the presence of Uranium and Thorium and their decay products (like radionuclide Potassium K-40).[2] Other most significant source is radon, a naturally occurring gas that emanates from rock and soil and can affect health severely.[7]

Both ionizing and non-ionizing radiations pose serious health risks, with the severity depending largely on the dose received. The higher the dose, the greater the potential for adverse health effects.[1,2] At very high doses, radiation can impair tissue and organ function, causing acute symptoms such as nausea, vomiting, skin redness, hair loss, acute radiation syndrome, and localized radiation injuries, commonly referred to as radiation burns.[3,4]

The contribution of radionuclides present in the earth crust to annual dose of radiation received by humans is maximum and affects their health the most.[1] In the radiation produced by radionuclides the contribution of gamma ray, as compared to the alpha- and beta- radiations, is significant thus highlighting its role as a potential health hazard radiation. Radon gas emanation in environment is another serious health hazard. Exposure to this gas is reported to cause lung cancer and the chances of non-smokers and smokers both being afflicted with lung cancer due to radon exposure are almost equal and the US Environmental Protection Agency recommends fixing of homes where the radon level is 4 picocuries per liter or higher.[5] In order to educate people about the risk of lung cancer due to radon exposure, the US Environmental Protection Agency recommends a protocol for testing of radon levels in the air and water supplies of the house they live in.[6]

With growing public awareness of the health risks posed by naturally occurring radioactive elements, including radon, we analyzed radionuclide concentrations of Ra-226, Th-232, and K-40 at various locations in the Dehradun district of Uttarakhand. To accomplish this we employ a One Way Repeated Measures ANOVA technique. The paper is organized as follows: Experimental data of radionuclide concentrations (sourced from T. Kandari et.al.) is presented in section- 2 as Table 1(a) and Table 1(b). Section-3 contains a detailed discussion on method of One Way Repeated measures ANOVA technique employing the data of Table 1(b). The conclusion of the study is presented in section-4.

Table 1(a). Radium (^{226}Ra), Thorium (^{232}Th) and Potassium (^{40}K) concentrations as measured at eighteen different sites of Rajpur region of Dehradun

Regions of Rajpur	*Radium Ra (Bq/kg)*	*Thorium Th (Bq/kg)*	*Potassium K (Bq/kg)*
R1	57	71	971
R2	100	69	1011
R3	51	65	1013
R4	70	76	662
R5	120	80	877
R6	80	118	1397
R7	50	75	435
R8	67	94	834
R9	54	83	564
R10	69	102	587
R11	68	99	1498
R12	56	108	1428
R13	88	87	1292
R14	59	98	384
R15	72	86	672
R16	91	101	942
R17	76	89	761
R18	67	86	756

Source: [8, Table 1, p. 1551]

ISSN: 2455-3190
DOI: https://doi.org/ 10.24321/2455.3190.202501

Table 1(b). Natural logarithmic concentrations of Radium (^{226}Ra), Thorium (^{232}Th) and Potassium (^{40}K) at eighteen different sites of Rajpur region of Dehradun

Regions of Rajpur	*Radium Ra (Bq/kg)*	*Thorium Th (Bq/kg)*	*Potassium K (Bq/kg)*
R1	4.043051	4.26268	6.878326
R2	4.605170	4.234107	6.918695
R3	3.931826	4.174387	6.920672
R4	4.248495	4.330733	6.495266
R5	4.787492	4.382027	6.776507
R6	4.382027	4.770685	7.242082
R7	3.912023	4.317488	6.075346
R8	4.204693	4.543295	6.726233
R9	3.988984	4.418841	6.335054
R10	4.234107	4.624973	6.375025
R11	4.219508	4.595120	7.311886
R12	4.025352	4.682131	7.264030
R13	4.477337	4.465908	7.163947
R14	4.077537	4.584967	5.950643
R15	4.276666	4.454347	6.510258
R16	4.510860	4.615121	6.848005
R17	4.330733	4.488636	6.634633
R18	4.204693	4.454347	6.628041

Source: [Authors' own computation]

2. Data for the study

Table 1(a) displays the secondary data used for this study and shows the experimental concentrations of Radium (^{226}Ra), Thorium (^{232}Th) and Potassium (^{40}K) as reported by T. Kandari et al.[8] at eighteen sites of Rajpur region in Dehradun. As a model requirement, the data of Table 1(a) is transformed into its natural logarithmic values and is tabulated as Table 1(b).

3. ANOVA model for the dataset of Table 1(b)

Using the ANOVA technique, often used by mathematicians and statisticians,[10-14] we develop a One Way Repeated measures ANOVA model for the dataset of Table 1(b) containing the data of logarithmic-concentrations of Radium, Thorium and Potassium at eighteen different sites in the Rajpur region of Dehradun district, Uttarakhand. By treating the natural logarithmic concentrations of Ra, Th and K as our response and Regions of data as subjects we develop our model. The analysis summary of the model is shown in Table 2, from where we see that with a P-Value of 0.00 at the 5% level of significance, there exist statistically significant differences in the mean response amongst the three levels of natural logarithmic concentrations of Ra, Th and K. It is important to note that the F-test performed by us at this stage and it's computed P-Value is meaningful only when the data exhibits sphericity.[15–17] Fig.1 displays the Profile Plot of the dataset of Table 1(b), which shows the concentrations of the three different radionuclides in different regions of Rajpur against their names. The uppermost pink colored line of this plot corresponds to the region R11 with the highest logarithmic concentration of potassium (ln(K)) at 7.31189, followed by the green colored line of the region R12 with an ln(K) – concentration of 7.26403. The large difference between the values of ln(K), ln(Ra) and ln(K) is evident from the plot.

For the validity of the F-test performed in Table 2 above, we present our results of the Mauchly's Sphericity Test [16, 17] in Table 3. As the P-Value of the Mauchly's test is higher than 0.05, therefore, we do not reject the assumption of sphericity at the 5% significance level. Thus, sphericity is present in the sample of Table 1(b), accordingly, we do not apply the correction factor epsilon to the F-tests performed in Table 2. The small P-Values

ISSN: 2455-3190
DOI: https://doi.org/ 10.24321/2455.3190.202501

of the four Multivariate Analysis of Variance (MANOVA) tests Wilks' lambda, Pillai trace, Hotelling-Lawley trace and Roy's largest root[18–24] and the P-Values of the differences between the logarithmic concentrations of Ra, Th and K clearly indicate that statistically significant differences exist at the 5% significance level between the means of these three variables in the Rajpur region of Dehradun.

Table 4 shows the mean values for the logarithmic concentrations of Ra, Th and K for the eighteen regions of Rajpur with their 95% confidence intervals and standard errors. These mean values along with their 95% Fisher's Least Significant Difference (LSD) intervals are shown in Fig. 2. It is clear from Fig. 2 that none of the pairs of intervals for ln(Ra), ln(Th) and ln(K) overlap vertically, showing that there exists a statistically significant difference in their means at the 95% confidence level.

Table 2.Analysis Summary of the procedure

Oneway Repeated Measures ANOVA
Regions of Rajpur: ln(Ra) - ln(K)
Subjects: Region
Analysis of Variance for Logarithmic concentrations of Ra, Th and K

Source	*Sum of Squares*	*Df*	*Mean Square*	*F-Ratio*	*P-Value*
Model	69.6175	19	3.66408	57.82	0.0000
Residual	2.15466	34	0.0633723		
Total (Corr.)	71.7721	53			

Factor Sums of Squares

Source	*Sum of Squares*	*Df*	*Mean Square*	*F-Ratio*	*P-Value*
Logarithmic concs. of radionuclides	67.7216	2	33.8608	534.32	0.0000
Region	1.89584	17	0.11152	1.76	0.0789
Error(Regions of Rajpur)	2.15466	34	0.0633723		
Total (corrected)	71.7721	53			

Standard Error of Est. = 0.251739

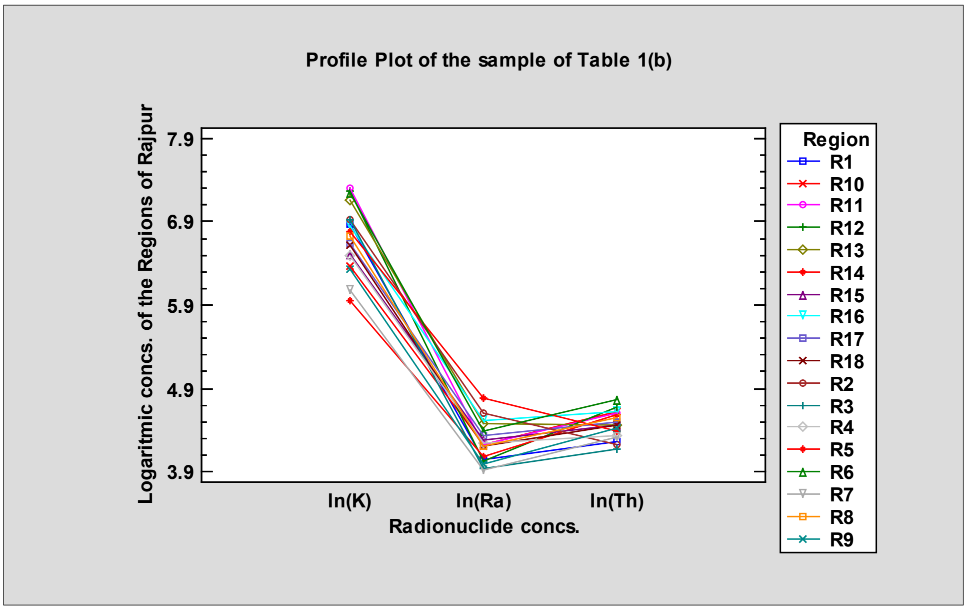

Figure 1.Profile Plot of the sample of Table 1

Table 3.Sphericity Tests and Adjustments

Sphericity Tests and Adjustments

Mauchly's Sphericity Test

W	*Chi-square*	*D.f.*	*P-value*
0.835477	2.87604	2.0	0.237398

Epsilon

Huynh-Feldt[25]	*Greenhouse-Geisser*[26]	*Lower-bound*
0.945979	0.858721	0.5

Tests of Within-Region Effects

Source	*Sphericity Correction*	*Sum of Squares*	*Df*	*Mean Square*	*F-Ratio*
Logarithmic Concs. of Radionuclides in Regions of Rajpur	None	67.7216	2.0	33.8608	534.32
	Huynh-Feldt	67.7216	1.89196	35.7945	534.32
	Greenhouse-Geisser	67.7216	1.71744	39.4317	534.32
	Lower-bound	67.7216	1.0	67.7216	534.32
Error(Logarithmic Concs. of Radionuclides in Regions of Rajpur)	None	2.15466	34.0	0.0633723	-
	Huynh-Feldt	2.15466	32.1633	0.0669913	
	Greenhouse-Geisser	2.15466	29.1965	0.0737985	
	Lower-bound	2.15466	17.0	0.126745	

Source	*P-Value*
Logarithmic Concs. of Radionuclides in Regions of Rajpur	0.0000
	0.0000
	0.0000
	0.0000
Error(Logarithmic Concs. of Radionuclides in Regions of Rajpur)	-

Multivariate Tests

Test	*Value*	*F*	*Hypothesis D.F.*	*Error D.F.*	*P-Value*
Wilks' lambda	0.0215979	362.406	2.0	16.0	0.0000
Pillai trace	0.978402	362.406	2.0	16.0	0.0000
Hotelling-Lawley trace	45.3007	362.406	2.0	16.0	0.0000
Roy's largest root	45.3007	362.406	2.0	16.0	0.0000

Table 4.Table of means for the three levels of logarithmic concentrations of Ra, Th and K in the Regions of Rajpur

Table of Means for Logarithmic concentrations of Ra, Th and K with 95.0 Percent Confidence Intervals

Level	*Count*	*Mean*	*Stnd. Error*	*Lower Limit*	*Upper Limit*
GRAND MEAN	54	5.14657	0.0342573	5.07696	5.21619
Logarithmic concs. Of Radionuclides					
ln(K)	18	6.72526	0.0593354	6.60467	6.84584
ln(Ra)	18	4.24781	0.0593354	4.12722	4.36839
ln(Th)	18	4.46666	0.0593354	4.34607	4.58724

ISSN: 2455-3190
DOI: https://doi.org/ 10.24321/2455.3190.202501

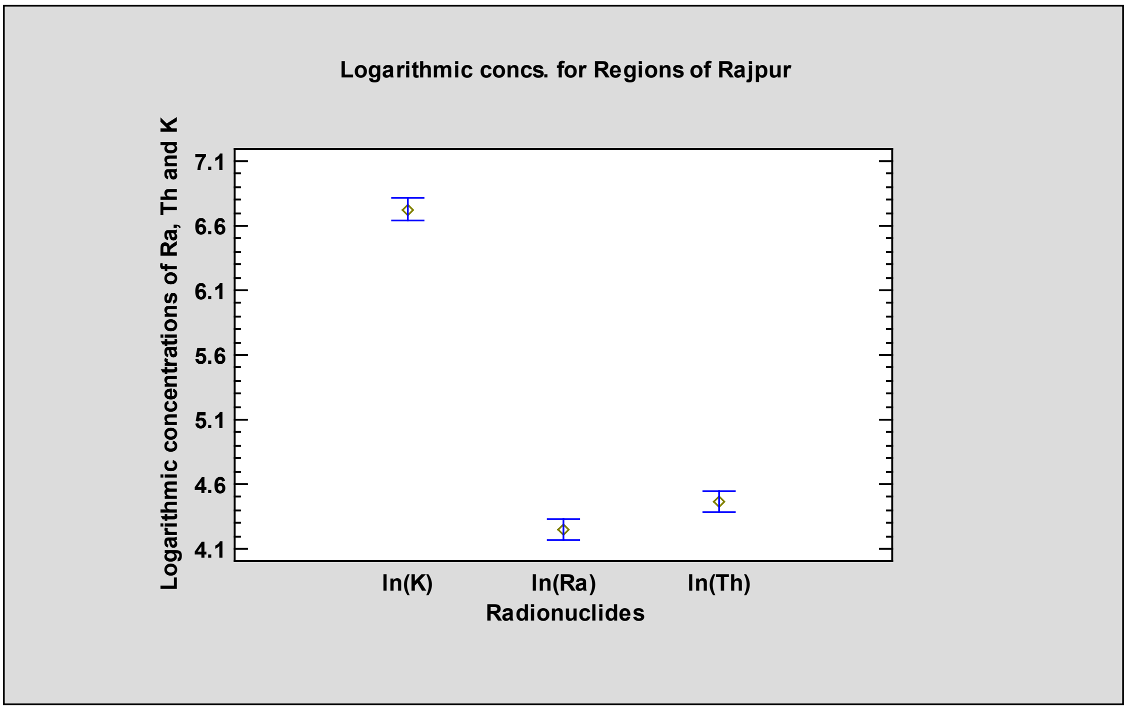

Figure 2.Means and 95% Fisher's LSD intervals for the sample of Table 1

The results of the Table of Means (Table 4) are examined in more detail by looking at Multiple Range Tests, recorded in Table 5, where means of logarithmic concentrations of three radionuclides in various regions of Rajpur are compared. In the first sub-table of Table 5, the last column, there are three identified homogeneous- groups found using the columns of X's. We note that every variable, ln(Ra), ln(Th) and ln(K), belongs of its own different group and this shows that according to Fisher's LSD procedure at the 5% significance level, there exists a statistically significant difference in means of logarithmic concentrations of radionuclides at different sites. These facts are further elaborated, numerically, in the lower sub-table of Table 5, where an asterisk is placed against the differences of means of the three pairs ln(K)-ln(Ra), ln(K)-ln(Th) and ln(Ra)-ln(Th) (red colored entries of 2.47745, 2.2586, -0.218847 respectively) which all exceed the +/- LSD means limit of 0.170532. This implies that all the three pairs exhibit a statistically significant difference of means for their logarithmic concentrations at 5% significance level.

In the Condition-Plot of the samples of Table 1(b) (Fig. 3), we have plotted the observed logarithmic concentrations of the three radionuclides as a function of sampling sites in the Rajpur region. To explore various possible trends in these concentrations, we use orthogonal polynomials and find different types of trends amongst the means of the logarithmic concentrations of Ra, Th and K. Since there are three levels of repeated measures, we can use a polynomial of at the most degree-two in this case. Table 6 shows the results of linear and quadratic trends for the logarithmic concentrations of Ra, Th and K means. The P-Values of both these trends are less than 0.05, and since the highest order trend with P-value less than 0.05 is quadratic, our simple conclusion is that a quadratic trend is necessary to model the means of the logarithmic concentrations of Ra, Th and K. The corresponding quadratic trend plot, i.e., a polynomial of order two fitted to the means of the logarithmic concentrations of these three radionuclides, is displayed in Fig. 3 with its equation. This plot also justifies the order of the conditions – ln(K), ln(Ra) and ln(Th) and the pattern of their means as shown in the Profile Plot of this sample in Fig. 1.

Table 7 records the fitted values of ln(Ra), ln(Th) and ln(K) as per the model developed by us. Their residuals, Studentized residuals, standardized error for forecast with 95% confidence intervals for the mean response and 95% confidence intervals for the forecast are also shown in the table. The table also includes model generated values corresponding to seven different hypothetical regions of Rajpur and they are highlighted with asterisks named as *R19, *R20,………, and *R25. The values of three independent variables used for this purpose were hypothesized by us.

Table 5.Multiple Range Tests for comparison of logarithmic concentrations of radionuclides in different regions of Rajpur

Multiple Comparisons for Logarithmic concentrations of Ra, Th and K by Regions of Rajpur
Method: 95.0 percent LSD

Radionuclides	*Count*	*LS Mean*	*LS Sigma*	*Homogeneous Groups*
ln(Ra)	18	4.24781	0.0593354	X
ln(Th)	18	4.46666	0.0593354	X
ln(K)	18	6.72526	0.0593354	X

Contrast	*Sig.*	*Difference*	*+/- Limits*
ln(K) - ln(Ra)	*	2.47745	0.170532
ln(K) - ln(Th)	*	2.2586	0.170532
ln(Ra) - ln(Th)	*	-0.218847	0.170532

*denotes a statistically significant difference.

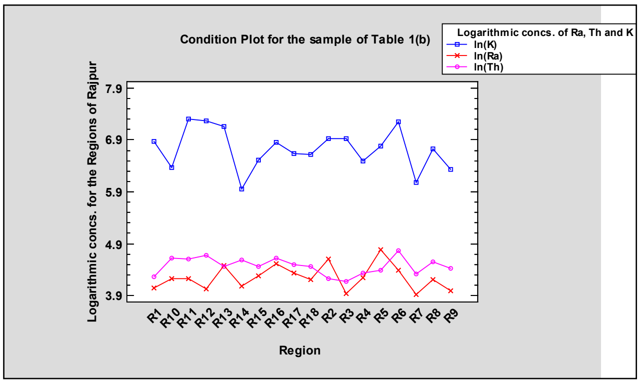

Figure 3.Condition Plot for the sample of Table 1

Table 6.Trend Analysis for the sample of Table 1

Trend Analysis

Trend	*Contrast*	*Stnd. error*	*t*	*P-value*
linear	-1.59707	0.0593354	-26.92	0.0000
quadratic	1.10076	0.0593354	18.55	0.0000

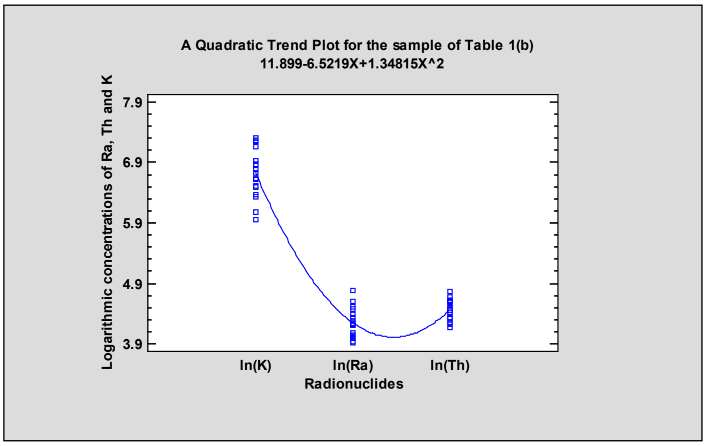

Figure 4.A quadratic trend plot for the sample of Table 1

ISSN: 2455-3190
DOI: https://doi.org/ 10.24321/2455.3190.202501

Table 7.Model Results for logarithmic concentrations of Ra, Th and K for different regions of Rajpur

Model Results for Logarithmic concentrations of Ra, Th and K

Row	*Region*	*Radionuclides*	*Observed Value*	*Fitted Value*	*Residual*	*Studentized Residual*	*Stnd. Error for Forecast*
1	R1	ln(Ra)	4.04305	4.18336	-0.140312	-0.642917	0.31619
2	R2	ln(Ra)	4.60517	4.37467	0.230502	1.06415	0.31619
3	R3	ln(Ra)	3.93183	4.13097	-0.199142	-0.916577	0.31619
4	R4	ln(Ra)	4.2485	4.14684	0.101658	0.464835	0.31619
5	R5	ln(Ra)	4.78749	4.43735	0.350138	1.64256	0.31619
6	R6	ln(Ra)	4.38203	4.58694	-0.204908	-0.943615	0.31619
7	R7	ln(Ra)	3.91202	3.89029	0.0217251	0.0991212	0.31619
8	R8	ln(Ra)	4.20469	4.28008	-0.0753883	-0.344359	0.31619
9	R9	ln(Ra)	3.98898	4.0363	-0.0473183	-0.215975	0.31619
10	R10	ln(Ra)	4.23411	4.20004	0.0340684	0.155462	0.31619
11	R11	ln(Ra)	4.21951	4.49751	-0.278005	-1.29055	0.31619
12	R12	ln(Ra)	4.02535	4.44584	-0.420495	-1.99814	0.31619
13	R13	ln(Ra)	4.47734	4.49107	-0.0137349	-0.0626621	0.31619
14	R14	ln(Ra)	4.07754	3.99306	0.0844817	0.38602	0.31619
15	R15	ln(Ra)	4.27667	4.20243	0.0742351	0.339078	0.31619
16	R16	ln(Ra)	4.51086	4.44667	0.0641884	0.293098	0.31619
17	R17	ln(Ra)	4.33073	4.27334	0.0573884	0.262	0.31619
18	R18	ln(Ra)	4.20469	4.2177	-0.0130116	-0.0593618	0.31619
19	*R19	ln(Ra)	3.97873	3.84933	0.129402	0.592538	0.31619
20	*R20	ln(Ra)	4.62597	4.3227	0.303268	1.41259	0.31619
21	*R21	ln(Ra)	4.02864	4.32639	-0.297748	-1.38582	0.31619
22	*R22	ln(Ra)	3.94927	4.29473	-0.345458	-1.61937	0.31619
23	*R23	ln(Ra)	4.23817	4.09245	0.145718	0.667922	0.31619
24	*R24	ln(Ra)	4.63824	4.418	0.220242	1.01572	0.31619
25	*R25	ln(Ra)	4.38482	4.16631	0.218505	1.00754	0.31619
26	R1	ln(Th)	4.26268	4.39546	-0.13278	-0.608125	0.31619
27	R2	ln(Th)	4.23411	4.58677	-0.352656	-1.65505	0.31619
28	R3	ln(Th)	4.17439	4.34307	-0.16868	-0.774419	0.31619
29	R4	ln(Th)	4.33073	4.35894	-0.0282096	-0.128716	0.31619
30	R5	ln(Th)	4.38203	4.64945	-0.26742	-1.23977	0.31619
31	R6	ln(Th)	4.77068	4.79904	-0.0283563	-0.129386	0.31619
32	R7	ln(Th)	4.31749	4.10239	0.215097	0.991492	0.31619
33	R8	ln(Th)	4.54329	4.49218	0.0511137	0.233318	0.31619
34	R9	ln(Th)	4.41884	4.2484	0.170444	0.782623	0.31619

ISSN: 2455-3190
DOI: https://doi.org/ 10.24321/2455.3190.202501

35	R10	ln(Th)	4.62497	4.41214	0.21283	0.980829	0.31619
36	R11	ln(Th)	4.59512	4.70961	-0.114493	-0.523845	0.31619
37	R12	ln(Th)	4.68213	4.65794	0.0241871	0.110357	0.31619
38	R13	ln(Th)	4.46591	4.70317	-0.237263	-1.09615	0.31619
39	R14	ln(Th)	4.58497	4.20516	0.379814	1.79087	0.31619
40	R15	ln(Th)	4.45435	4.41453	0.0398171	0.181711	0.31619
41	R16	ln(Th)	4.61512	4.65877	-0.0436496	-0.199216	0.31619
42	R17	ln(Th)	4.48864	4.48544	0.0032004	0.0146004	0.31619
43	R18	ln(Th)	4.45435	4.4298	0.0245504	0.112015	0.31619
44	*R19	ln(Th)	4.21962	4.06143	0.158194	0.72572	0.31619
45	*R20	ln(Th)	4.59374	4.5348	0.0589404	0.269096	0.31619
46	*R21	ln(Th)	4.68923	4.53849	0.150744	0.691187	0.31619
47	*R22	ln(Th)	4.41954	4.50683	-0.0872863	-0.398878	0.31619
48	*R23	ln(Th)	4.43462	4.30455	0.13007	0.595623	0.31619
49	*R24	ln(Th)	4.28592	4.6301	-0.344176	-1.61302	0.31619
50	*R25	ln(Th)	4.56438	4.37841	0.185967	0.854963	0.31619
51	R1	ln(K)	6.87833	6.60524	0.273091	1.26695	0.31619
52	R2	ln(K)	6.9187	6.79655	0.122155	0.559125	0.31619
53	R3	ln(K)	6.92067	6.55285	0.367821	1.73066	0.31619
54	R4	ln(K)	6.49527	6.56872	-0.0734488	-0.335478	0.31619
55	R5	ln(K)	6.77651	6.85923	-0.0827188	-0.37794	0.31619
56	R6	ln(K)	7.24208	7.00882	0.233265	1.07722	0.31619
57	R7	ln(K)	6.07535	6.31217	-0.236822	-1.09407	0.31619
58	R8	ln(K)	6.72623	6.70196	0.0242745	0.110756	0.31619
59	R9	ln(K)	6.33505	6.45818	-0.123125	-0.563599	0.31619
60	R10	ln(K)	6.37502	6.62192	-0.246899	-1.14188	0.31619
61	R11	ln(K)	7.31189	6.91939	0.392498	1.85498	0.31619
62	R12	ln(K)	7.26403	6.86772	0.396308	1.87433	0.31619
63	R13	ln(K)	7.16395	6.91295	0.250998	1.16138	0.31619
64	R14	ln(K)	5.95064	6.41494	-0.464295	-2.2271	0.31619
65	R15	ln(K)	6.51026	6.62431	-0.114052	-0.521816	0.31619
66	R16	ln(K)	6.84801	6.86855	-0.0205388	-0.0937078	0.31619
67	R17	ln(K)	6.63463	6.69522	-0.0605888	-0.276634	0.31619
68	R18	ln(K)	6.62804	6.63958	-0.0115388	-0.0526421	0.31619
69	*R19	ln(K)	5.98361	6.27121	-0.287595	-1.33673	0.31619
70	*R20	ln(K)	6.38237	6.74458	-0.362209	-1.70261	0.31619
71	*R21	ln(K)	6.89527	6.74827	0.147005	0.673874	0.31619
72	*R22	ln(K)	7.14935	6.71661	0.432745	2.06153	0.31619
73	*R23	ln(K)	6.23854	6.51433	-0.275789	-1.2799	0.31619
74	*R24	ln(K)	6.96381	6.83988	0.123935	0.567328	0.31619
75	*R25	ln(K)	6.18372	6.58819	-0.404472	-1.91592	0.31619

ISSN: 2455-3190
DOI: https://doi.org/ 10.24321/2455.3190.202501

Row	Lower 95.0% CL for Forecast	Upper 95.0% CL for Forecast	Lower 95.0% CL for Mean	Upper 95.0% CL for Mean
1	3.54762	4.81911	3.85627	4.51045
2	3.73892	5.01041	4.04758	4.70176
3	3.49523	4.76672	3.80388	4.45806
4	3.5111	4.78259	3.81975	4.47393
5	3.80161	5.0731	4.11026	4.76444
6	3.95119	5.22268	4.25985	4.91403
7	3.25455	4.52604	3.56321	4.21738
8	3.64433	4.91582	3.95299	4.60717
9	3.40055	4.67204	3.70921	4.36339
10	3.5643	4.83579	3.87295	4.52713
11	3.86177	5.13326	4.17043	4.8246
12	3.8101	5.08159	4.11876	4.77293
13	3.85533	5.12682	4.16399	4.81816
14	3.35731	4.6288	3.66597	4.32015
15	3.56669	4.83818	3.87535	4.52952
16	3.81093	5.08242	4.11958	4.77376
17	3.6376	4.90909	3.94625	4.60043
18	3.58196	4.85345	3.89061	4.54479
19	3.21358	4.48507	3.52224	4.17642
20	3.68696	4.95845	3.99561	4.64979
21	3.69064	4.96213	3.9993	4.65348
22	3.65898	4.93047	3.96764	4.62182
23	3.45671	4.7282	3.76536	4.41954
24	3.78225	5.05374	4.09091	4.74509
25	3.53057	4.80206	3.83923	4.4934
26	3.75972	5.0312	4.06837	4.72255
27	3.95102	5.22251	4.25968	4.91385
28	3.70733	4.97881	4.01598	4.67016
29	3.7232	4.99468	4.03185	4.68603
30	4.01371	5.28519	4.32236	4.97654
31	4.16329	5.43478	4.47195	5.12612
32	3.46665	4.73814	3.77531	4.42948
33	3.85643	5.12792	4.16509	4.81926
34	3.61265	4.88414	3.92131	4.57548
35	3.7764	5.04788	4.08505	4.73923
36	4.07387	5.34536	4.38253	5.0367
37	4.0222	5.29369	4.33086	4.98503
38	4.06743	5.33892	4.37609	5.03026

ISSN: 2455-3190
DOI: https://doi.org/ 10.24321/2455.3190.202501

39	3.56941	4.8409	3.87807	4.53224
40	3.77879	5.05028	4.08745	4.74162
41	4.02303	5.29451	4.33168	4.98586
42	3.8497	5.12118	4.15835	4.81253
43	3.79406	5.06554	4.10271	4.75689
44	3.42568	4.69717	3.73434	4.38851
45	3.89906	5.17054	4.20771	4.86189
46	3.90274	5.17423	4.2114	4.86557
47	3.87108	5.14257	4.17974	4.83391
48	3.66881	4.94029	3.97746	4.63164
49	3.99435	5.26584	4.30301	4.95718
50	3.74267	5.01416	4.05133	4.7055
51	5.9695	7.24098	6.27815	6.93233
52	6.1608	7.43229	6.46946	7.12363
53	5.91711	7.18859	6.22576	6.87994
54	5.93298	7.20446	6.24163	6.89581
55	6.22349	7.49497	6.53214	7.18632
56	6.37307	7.64456	6.68173	7.3359
57	5.67643	6.94792	5.98508	6.63926
58	6.06621	7.3377	6.37487	7.02904
59	5.82243	7.09392	6.13109	6.78526
60	5.98618	7.25766	6.29483	6.94901
61	6.28365	7.55514	6.5923	7.24648
62	6.23198	7.50347	6.54063	7.19481
63	6.27721	7.5487	6.58586	7.24004
64	5.77919	7.05068	6.08785	6.74202
65	5.98857	7.26006	6.29722	6.9514
66	6.23281	7.50429	6.54146	7.19564
67	6.05948	7.33096	6.36813	7.02231
68	6.00384	7.27532	6.31249	6.96667
69	5.63546	6.90695	5.94412	6.59829
70	6.10884	7.38032	6.41749	7.07167
71	6.11252	7.38401	6.42118	7.07535
72	6.08086	7.35235	6.38952	7.04369
73	5.87859	7.15007	6.18724	6.84142
74	6.20413	7.47562	6.51279	7.16696
75	5.95245	7.22394	6.2611	6.91528

* denotes the seven hypothetical regions of Rajpur introduced by us with hypothetical values of ln(Ra), ln(Th) and ln(K) to calculate the predicted values of these variables according to the model developed by us.

ISSN: 2455-3190
DOI: https://doi.org/ 10.24321/2455.3190.202501

Table 8.Comparison of the observed and the fitted values of the three radionuclide concentrations by the proposed model region wise

Region of Rajpur	*Observed Ra (Bq/kg)*	*Fitted Ra (Bq/kg)*	*ObservedTh (Bq/kg)*	*Fitted Th (Bq/ kg)*	*Observed K (Bq/kg)*	*Fitted K (Bq/kg)*
R1	56.99993	65.58585	71.00001	81.08192	971.0034	738.9572
R2	99.99998	79.41363	69.00024	98.176806	1011.005	894.7551
R3	51.00022	62.23826	65.00018	76.943393	1012.998	701.2399
R4	70.00033	63.23387	75.99975	78.174226	662.0029	712.4573
R5	119.9998	84.55059	80.00027	104.52748	877.0026	952.6333
R6	80.00027	98.1935	117.9995	121.39382	1396.997	1106.348
R7	49.99985	48.92507	75.00014	60.484673	435.0017	551.2398
R8	66.99982	72.24622	93.99955	89.315943	833.9972	813.9997
R9	53.99978	56.61647	82.99995	69.993333	563.9976	637.899
R10	69.00024	66.689	101.9997	82.445709	586.9972	751.3864
R11	68.00016	89.79327	99.00001	111.00886	1498.006	1011.703
R12	55.99991	85.27148	107.9999	105.4187	1428	960.7555
R13	88.00028	89.21686	87.00016	110.29626	1292.004	1005.208
R14	59.00015	54.22055	98.00025	67.031322	383.999	610.9041
R15	72.00028	66.84858	86.00023	82.64299	672.0011	753.1843
R16	91.00004	85.34228	100.9999	105.50623	942.0045	961.5533
R17	75.99975	71.76092	89.00032	88.715977	760.9974	808.5318
R18	66.99982	67.87719	86.00023	83.914632	755.999	764.7737
*R19	53.44911	46.96159	68.00764	58.057274	396.8705	529.1172
*R20	102.1018	75.39191	98.86349	93.204872	591.3275	849.4423
*R21	56.18445	75.67062	108.7694	93.549434	987.5923	852.5825
*R22	51.89747	73.31242	83.05807	90.634053	1273.278	826.0126
*R23	69.28095	59.88643	84.32008	74.035892	512.1103	674.7417
*R24	103.3623	82.93026	72.66937	102.52432	1057.656	934.377
*R25	80.22378	64.47709	96.00305	79.711192	484.792	726.4648

Fig. 5 shows a plot of observed values of logarithmic concentrations for three radionuclides versus corresponding fitted values. Though the data points lie roughly along the diagonal line, our proposed model does not predict the pattern of the observed concentrations of the radionuclides as accurately as we expected. We remark here that after examining a number of different possibilities and carrying out various possible procedures, this procedure was one of the successful procedures for the data under study. Since the variability of the data of Table 1 is very high and with our experience on various computations with this dataset, the limited success of the proposed ANOVA model for this sample is not a surprise to us.

The Residual plot of the Studentized residuals of the formulated model versus the predicted logarithmic concentrations of the radionculides is drawn in Fig. 6. The values of the Studentized residuals beyond the range of ±2, show a departure of the residuals from the normality. In Table 7 we highlight two values of such residuals which correspond to the concentrations of ln(K) in the regions R14 and *R22.

By excluding the hypothetical regions *R19 to *R25, included by us in the analysis, we draw the Normal Probability Plot for Studentized residuals of the model for actual regions R1 to R18 in Fig. 7. This figure (Fig. 7) shows that the residuals lie very close to the reference diagonal line superimposed on the plot by the method of least squares indicating that the Studentized residuals of proposed model follow a normal distribution very closely.

The auto-correlation plot of residuals of the postulated model is shown in Fig. 8. We draw this plot at lag 10, which computes the autocorrelation coefficients between the residuals at time t and t-10. The red colored lines in Fig. 8 show the 95% probability limits around zero. Since all the ten computed autocorrelation coefficients in this figure lie well within these bounds, this indicates that the series may be completely random.

Table 9 lists all unusual residuals of the dataset of Table 1(b), and contains only those observations of Table 1(b) which have Studentized residuals exceeding 2 in numerical magnitude. These points refer to the observations of Table 1(b) for which the logarithmic concentrations of Ra, Th and K were more than two standard deviations away from the fitted model. There are no such observations, recorded in Table 9, for which the Studentized residuals are greater than 3 in magnitude and which could be outliers and ought to be removed from the data. From the summary statistics of the three variables, shown in Table 10, it is clear that the standardized skewness and the standardized kurtosis for all the data points are well within the range expected of a sample drawn from a normal distribution. This fact is further elucidated in the Normal Probability Plots of variables ln(Ra), ln(Th) and ln(K), shown in Figs. 9-11 along with P-Values of the corresponding Shapiro-Wilk's tests performed by us for them, and their 90th percentiles with 95% confidence intervals relevant for these percentiles.

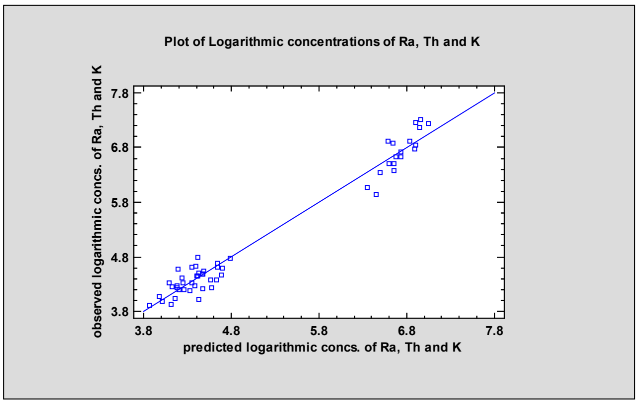

Figure 5.Plot of the observed and fitted values of logarithmic concentrations of Ra, Th and K

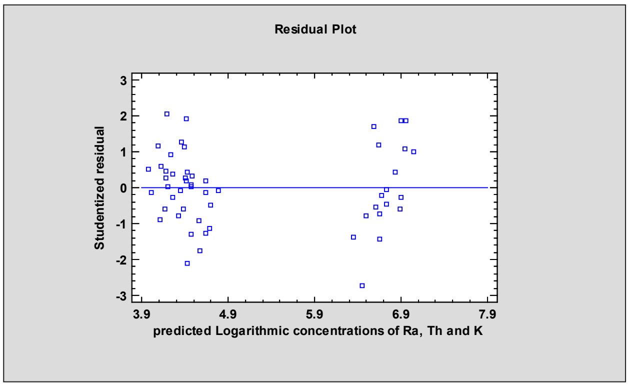

Figure 6.Residual Plot of the sample of Table 1

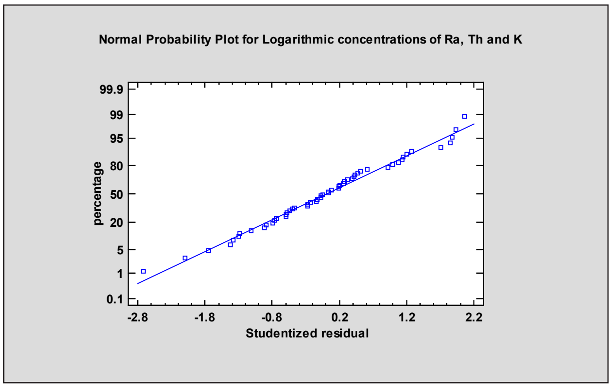

Figure 7.Normal Probability Plot of the Studentized residuals of the model for the regions R1 to R18 of Rajpur

ISSN: 2455-3190
DOI: https://doi.org/ 10.24321/2455.3190.202501

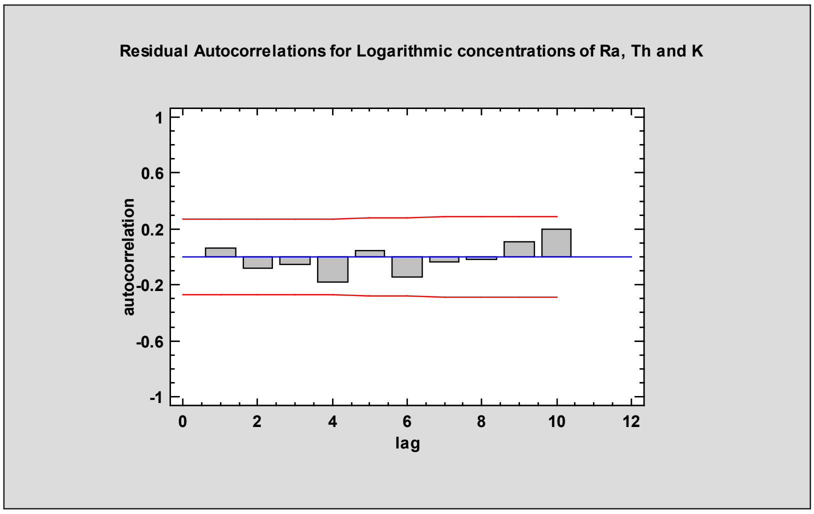

Figure 8.Residual Autocorrelation Plot at lag 10 of the residuals of the postulated model

Table 9.Unusual Residuals for the sample of Table 1(b) for the fitted model

Unusual Residuals for Logarithmic concentrations of Ra, Th and K

Row	*Region*	*Logarithmic Concs.*	*Y*	*Predicted Y*	*Residual*
12R12	ln(Ra)	4.02535	4.42507	-0.399721	-2.10
32R14	ln(Th)	4.58497	4.19113	0.393839	2.06
50R14	ln(K)	5.95064	6.44973	-0.499094	-2.72

Table 10.Summary statistics of the logarithmic concentrations of Ra, Th and K

Summary Statistics

	ln(Ra)	*ln(Th)*	*ln(K)*
Count	18	18	18
Average	4.24781	4.46666	6.72526
Standard deviation	0.239026	0.163492	0.39294
Coeff. of variation	5.62705%	3.66027%	5.84274%
Minimum	3.91202	4.17439	5.95064
Maximum	4.78749	4.77068	7.31189
Range	0.87547	0.59629	1.36125
Stnd. skewness	1.04634	-0.0785218	-0.499584
Stnd. kurtosis	0.0138555	-0.572045	-0.398692

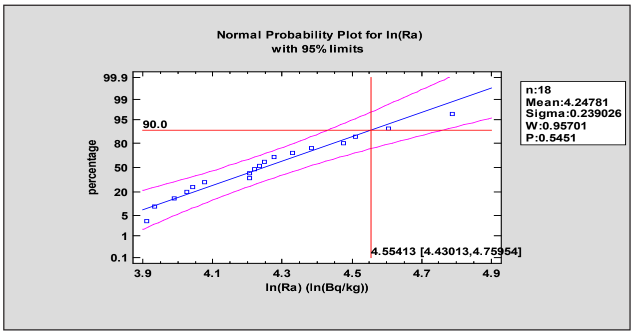

Figure 9.Normal Probability Plot for ln(Ra)

ISSN: 2455-3190
DOI: https://doi.org/ 10.24321/2455.3190.202501

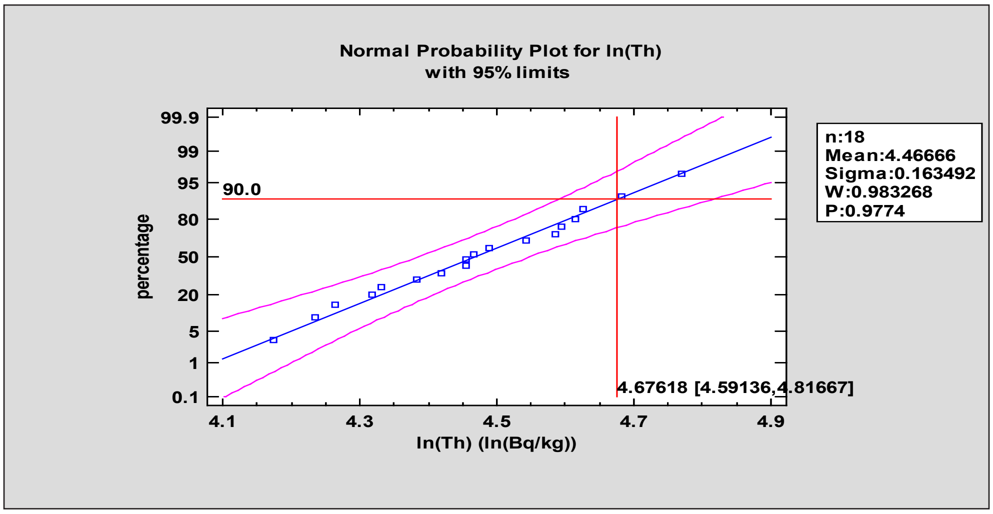

Figure 10.Normal Probability Plot for ln(Th)

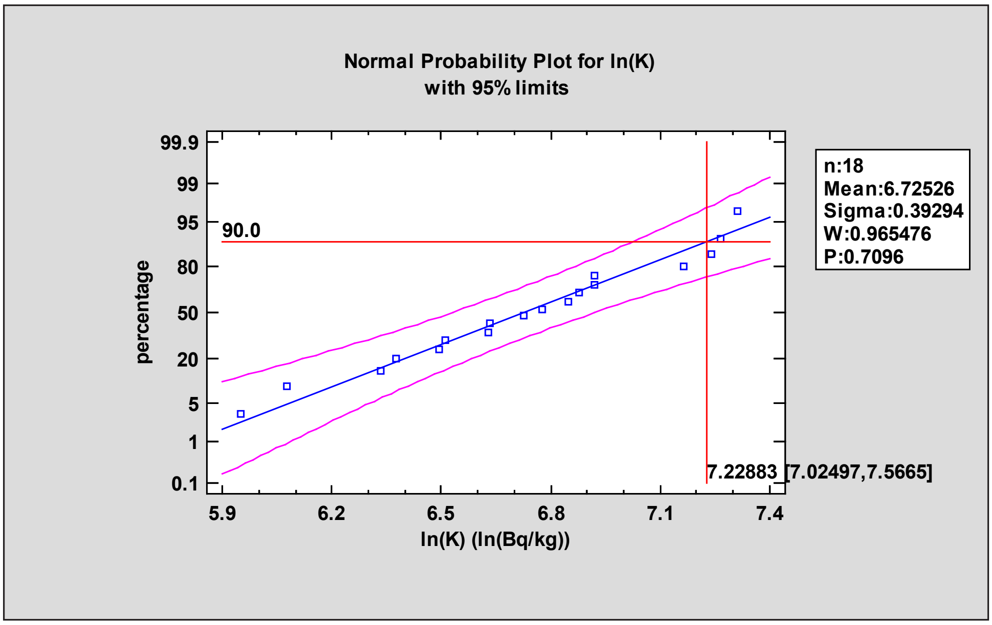

Figure 11.Normal Probability Plot for ln(K)

For the applicability of the above model developed by us to those sample concentrations of ln(Ra), ln(Th) and ln(K) shown in Table 1(b), which do not come from a normal population, we performed a Friedman's test[27-28] procedure for the dataset in hand. The Friedman's test gives a nonparametric option to the F test described in Table 2. The null hypothesis to be tested here through this test is that the medians within each of the three factor levels ln(Ra), ln(Th) and ln(K) are the same. The extremely small P-Value of the Friedman's test from Table 11 speaks of statistically significant difference amongst the means ranks of the factor levels at the 95% confidence level. The lower part of Table 11 shows that statistically significant differences between the means ranks exist between the pairs of levels ln(K)-ln(Ra) and ln(K)-ln(Th) by the Bonferroni procedure at the 95.0% family confidence level. Lastly, in Fig. 12 we draw the Plot of Mean Ranks of the three levels together with their 95% Bonferroni intervals, from where it is clear that the interval corresponding to ln(K) does not vertically overlap with the intervals of ln(Ra) and ln(Th), justifying the conclusion drawn from the lower part of Table 11. This completes our procedure of One-Way Repeated Measures ANOVA for the sample of Table 1(b).

ISSN: 2455-3190
DOI: https://doi.org/ 10.24321/2455.3190.202501

Table 11.Friedman's test for the sample of Table 1(b)

Friedman Test

	Sample Size	*Average Rank*
ln(K)	18	3.0
ln(Ra)	18	1.16667
ln(Th)	18	1.83333

Test statistic = 31.0   P-value =1.85539E-7

95.0 percent Bonferroni intervals

Contrast	*Sig.*	*Difference*	*+/- Limits*
ln(K) - ln(Ra)	*	1.83333	0.797995
ln(K) - ln(Th)	*	1.16667	0.797995
ln(Ra) - ln(Th)		-0.666667	0.797995

* denotes a statistically significant difference.

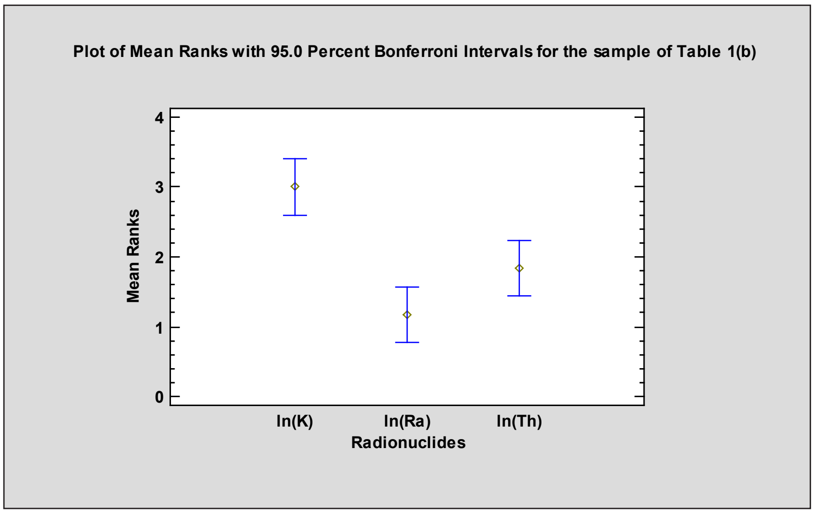

Figure 12.Plot of Mean Ranks for the sample of Table 1(b).

4. Conclusion

We presented the details of our model using a One-Way Repeated Measures ANOVA technique and applied it to predict the concentrations of Ra, Th, and K in eighteen soil samples from the Rajpur region of Dehradun district, Uttarakhand, India. The model analysis was performed using the data of radionuclide concentrations of soil samples as reported by Tushar Kandari et al.8. Additionally, we generated radionuclide concentrations at seven hypothetical locations (*R19, *R20,….,*R25) using this model. Our numerical experiments indicate that the concentrations of Ra, Th, and K in the soils-samples behave as purely random variables with minimal correlation and are nearly independently distributed. Consequently, the accuracy of the model-generated radionuclide concentrations is limited due to this inherent randomness. Our future work in this direction will focus on related issues concerning the concentrations of these radioactive nuclides in some other regions of Dehradun.

Acknowledgments

The authors of this paper gratefully acknowledge and sincerely thank all the learned authors of the paper[8] from where this data, as listed in Table 1(a), is sourced. The authors also express their gratitude to the Publishers of the Journal[8,9] and the copyright owners Akadémiai Kiadó, Budapest, Hungary of this work.[9] The first and the second authors of the paper would also like to thank Dr. Tushar Kandari[8] for discussions on this work. The critical comments and feedback of the anonymous referees are thankfully acknowledged.

References

1. United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR). Sources and effects of ionizing radiation. Vol. I. New York: United Nations; 2000.
2. United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR). Exposure from natural

sources of radiation. New York: United Nations; 1993.
3. World Health Organization. Ionizing radiation, health effects and protective measures. Geneva: WHO; 2016.
4. Ng K-H. Non-ionizing radiation – sources, biological effects, emission and exposures. In: Proceedings of the International Conference on Non-Ionizing Radiation at UNITEN ICNIR2003; 2003 Oct; Malaysia.
5. United States Environmental Protection Agency. Consumer's guide to radon reduction: how to fix your home. EPA 402/K-10/005. Washington DC: USEPA; 2016. Available from: https://www.epa.gov/sites/default/files/201612/documents/2016_consumers_guide_to_radon_reduction.pdf
6. United States Environmental Protection Agency. A citizen's guide to radon: the guide to protecting yourself and your family from radon. EPA 402/K-12/002. Washington DC: USEPA; 2016. Available from: https://www.epa.gov/sites/default/files/2016-12/documents/2016_a_citizens_guide_to_radon.pdf
7. Al-Zoughool M, Krewski D. Health effects of radon: a review of the literature. Int J Radiat Biol. 2009;85(1):57–69.
8. Kandari T, Singh P, Semwal P, Kumar A, Bourai AA, Ramola RC. Evaluation of background radiation level and excess lifetime cancer risk in Doon Valley, Garhwal Himalaya. J Radioanal Nucl Chem. 2021;330:1545–1557. https://doi.org/10.1007/s10967-021-07988-2
9. Kandari T, Singh P, Semwal P, Kumar A, Bourai AA, Ramola RC. Correction to: Evaluation of background radiation level and excess lifetime cancer risk in Doon valley, Garhwal Himalaya. J Radioanal Nucl Chem. 2022;330:1135. https://doi.org/10.1007/s10967-022-08200-9
10. Hand DJ, Taylor CC. Multivariate analysis of variance and repeated measures: a practical approach for behavioral scientists. New York: Chapman and Hall; 1987.
11. Wikipedia. Multivariate analysis of variance [Internet]. Available from: https://en.wikipedia.org/wiki/Multivariate_analysis_of_variance
12. Krzanowski WJ. Principles of multivariate analysis: a user's perspective. Oxford: Oxford University Press; 1988.
13. Warne RT. A primer on multivariate analysis of variance (MANOVA) for behavioral scientists. Pract Assess Res Eval. 2014;19(17):1–10. Available from: https://scholarworks.umass.edu/pare/vol19/iss1/17/\
14. Girden ER. ANOVA: repeated measures. Newbury Park, CA: Sage; 1992.
15. Sphericity in repeated measures analysis of variance [Internet]. Available from: http://oak.ucc.nau.edu/rh232/courses/EPS625/Handouts/RM-ANOVA/Sphericity.pdf
16. Mauchly JW. Significance test for sphericity of a normal n-variate distribution. Ann Math Stat. 1940;11:204–209.
17. Wikipedia. Mauchly's sphericity test [Internet]. Available from: https://en.wikipedia.org/wiki/Mauchly%27s_sphericity_test
18. Wikipedia. Wilk's lambda distribution [Internet]. Available from: https://en.wikipedia.org/wiki/Wilks%27s_lambda_distribution
19. Rao CR. An asymptotic expansion of the distribution of Wilks' criterion. Bull Inst Int Stat. 1951;33:177–180.
20. Das Gupta S. Pillai's trace test. In: Armitage P, Colton T, editors. Encyclopedia of Biostatistics. Hoboken: Wiley; 2005. https://doi.org/10.1002/0470011815.b2a13067
21. Zhang Q, Hu J, Bai Z. Modified Pillai's trace statistics for two high-dimensional sample covariance matrices. J Stat Plan Inference. 2020;207:255–275. https://doi.org/10.1016/j.jspi.2020.01.002
22. Zhu T, Zhang J-T, Cheng M-Y. One-way MANOVA for functional data via Lawley–Hotelling trace test. J Multivar Anal. 2022;192:105095. https://doi.org/10.1016/j.jmva.2022.105095
23. Nadler B, Johnstone IM. On the distribution of Roy's largest root test in MANOVA and in signal detection in noise. Tech Rep No. 2011-04. Stanford, CA: Dept of Statistics, Stanford University; 2011. Available from: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=d98b6ed2192512777ba367b522f3dc2f17b97b7c
24. Kuhfeld WF. A note on Roy's largest root. Psychometrika. 1986;51:479–481. https://doi.org/10.1007/BF02294069
25. Huynh H, Feldt LS. Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. J Educ Stat. 1976;1:69–82.
26. Greenhouse SW, Geisser S. On methods in the analysis of profile data. Psychometrika. 1959;24:95–112.
27. Wikipedia. Friedman's test [Internet]. Available from: https://de.wikipedia.org/wiki/Friedman-Test_(Statistik)
28. Fotouhi S, Asadi S, Kattan MW. A comprehensive data level analysis for cancer diagnosis on imbalanced data. J Biomed Inform. 2019;90:103089. https://doi.org/10.1016/j.jbi.2018.12.003

ISSN: 2455-3190
DOI: https://doi.org/ 10.24321/2455.3190.202501