

MULTIVARIATE CHEMOMETRIC ANALYSIS OF RADON CONCENTRATION IN  
GEOHERMAL FLUIDS FROM HOMA MOUNTAIN, HIGH BACKGROUND  
RADIATION AREA (HBRA)

GEOFFREY NYANDIKO MAYAKA B.Ed (Sc).

I56/CE/11985/07

A thesis submitted in partial fulfillment of the requirements for the award of the degree of Master of Science in physics in the School of Pure and Applied Sciences of Kenyatta University.

March, 2014.

**DECLARATION**

This thesis is my original work and has not been presented for the award of a degree or any other award in any University.

Signature

Date

.....  
**Geoffrey N. Mayaka**  
Department of Physics  
Kenyatta University.

.....

This thesis has been submitted with our approval as University Supervisors

Signature

Date

.....  
**Dr. N.O. Hashim**  
Department of Physics  
Kenyatta University.

.....

Signature

Date

.....  
**Dr. H.K. Angeyo**  
Department Physics  
University of Nairobi.

.....

**DEDICATION**

*This thesis is dedicated to my friend and dear wife Naom and our daughter Esther.*

## ACKNOWLEDGEMENTS

Foremost, I would like to sincerely thank my advisors: Dr. N. O. Hashim and Dr. H. K. Angeyo for their guidance and support during the research period that enabled me to reach the research objectives with strong motivation and dedication. I would also like to thank the National Council for Science and Technology for funding this project.

I am very grateful to all the physics department technical staff led by Mr. Simon Njuguna who lent me apparatus, and generally did things that were beyond the call of duty. I thank the Radiation Protection Board laboratory staff for allowing me to use their liquid scintillation counter.

I thank my parents for their unconditional love and support.

I would like to thank my dear wife, Naom who has been a source of unlimited love, friendship, laughter, support and motivation.

Above all I want to thank the Almighty God who has taken me this far. I confess that the strides I have made were only possible because he was my strength and motivator.

**TABLE OF CONTENTS**

|                                    |      |
|------------------------------------|------|
| Title                              | i    |
| Declaration                        | ii   |
| Dedication                         | iii  |
| Acknowledgement                    | iv   |
| Table of contents                  | v    |
| List of tables                     | ix   |
| List of figures                    | x    |
| List of abbreviations and acronyms | xii  |
| Abstract                           | xiii |

**CHAPTER 1****INTRODUCTION**

|  |    |
|--|----|
| 1.1 Background to the study                                  | 1  |
| 1.2 Radon  | 3  |
| 1.3 Sources of radon in the environment                      | 5  |
| 1.4 Radon gas in geothermal exploration                      | 6  |
| 1.5 Geothermal fields lying in HBRA                          | 7  |
| 1.6 Methods of measuring radon                               | 8  |
| 1.7 Multivariate chemometric analysis of radon concentration | 9  |
| 1.8 Statement of the research problem                        | 10 |

|                             |    |
|-----------------------------|----|
| 1.9 Objectives              | 11 |
| 1.9.1 Main objective        | 11 |
| 1.9.2 Specific objectives   | 11 |
| 1.10 Rationale of the study | 12 |

## **CHAPTER 2**

### **LITERATURE REVIEW**

|   |    |
|---|----|
| 2.1 Comparison of Lucas cell and LSC methods for measuring radon in water | 13 |
| 2.2 Studies of radon in high background radiation areas                   | 13 |
| 2.3 Studies of radon in geothermally active HBRA's                        | 16 |
| 2.4 Multivariate chemometric analysis of environmental data               | 18 |

## **CHAPTER 3**

### **PRINCIPLES OF LIQUID SCINTILLATION SPECTROMETRY AND MULTIVARIATE CHEMOMETRICS**

|  |    |
|--|----|
| 3.1 Introduction   | 22 |
| 3.2 Comparison of alpha/beta LSC and conventional beta LSC detectors | 23 |
| 3.3 Secular equilibrium of radon and its daughter radionuclides      | 23 |
| 3.4 Principal components analysis (PCA)                              | 25 |
| 3.5 Hierarchical cluster analysis (HCA)                              | 27 |

|  |    |
|--|----|
| 3.6 Soft independent modeling of class analogy (SIMCA) | 30 |
| 3.7 Cooman's plot                                      | 31 |

## CHAPTER 4

### MATERIALS AND METHODS

|  |    |
|--|----|
| 4.1 Sample collection for radon measurement  | 33 |
| 4.2 Liquid scintillation spectrometer  | 35 |
| 4.3 Liquid scintillation counting vials  | 37 |
| 4.4 Experimental procedures  | 38 |
| 4.4 Calibration and normalization of liquid scintillation counter                            | 38 |
| 4.4.1 Sample preparation for radon measurement   | 39 |
| 4.4.2 Background determination   | 39 |
| 4.4.3 Minimum detectable activity of LSC method  | 40 |
| 4.5 Radon measuring principle  | 41 |
| 4.6 Multivariate chemometric data analysis   | 42 |
| 4.7 Data pre-processing  | 43 |
| 4.8 Principal components analysis of the activity concentration of radon, pH and temperature | 44 |
| 4.9 Hierarchical cluster analysis of the activity concentration of radon                     | 45 |
| 4.10 SIMCA classification  | 46 |
| 4.10.1 Modeling power of the activity concentration of radon, pH and temperature             | 47 |

## CHAPTER 5

### RESULTS AND DISCUSSION

|  |    |
|--|----|
| 5.1 Activity concentration of radon                                      | 48 |
| 5.2 Statistical distribution of radon concentration                      | 56 |
| 5.3 Principal components analysis of the activity concentration of radon | 57 |
| 5.4 Hierarchical cluster analysis of the activity concentration of radon | 63 |
| 5.5 SIMCA classification of geothermal water samples                     | 68 |

## CHAPTER 6

### CONCLUSIONS AND RECOMMENDATIONS

|  |    |
|--|----|
| 6.1 Conclusions  | 75 |
| 6.2 Recommendations and future perspective   | 79 |
| REFERENCES   | 80 |
| APPENDIX I Uranium decay series.   | 87 |
| APPENDIX II Water samples analyzed in this work.   | 88 |
| APPENDIX III Mean centered results of radon, pH and temperature of spring, river and pond water samples analyzed in this work. | 89 |
| APPENDIX IV Liquid scintillation counter used to measure the activity concentration of radon in this work.                     | 91 |
| APPENDIX V A photograph of Homa mountain in western Kenya where the geothermal waters were sampled.                            | 92 |

## LIST OF TABLES

|                    |  |    |
|--------------------|--|----|
| <b>Table 1.1:</b>  | <sup>222</sup> Rn-decay and its progeny scheme (adapted from Wilkening, 1990).   | 4  |
| <b>Table 5.1:</b>  | The activity concentration of radon in spring waters from geothermally active HBRA.  | 48 |
| <b>Table 5.2:</b>  | The activity concentration of radon in spring waters from non-geothermally active HBRA.  | 48 |
| <b>Table 5.3:</b>  | The activity concentration of radon in spring waters from non-geothermally active non-HBRA.  | 49 |
| <b>Table 5.4:</b>  | The activity concentration of radon in river waters from geothermally active HBRA.   | 49 |
| <b>Table 5.5:</b>  | The activity concentration of radon in river waters from non-geothermally active HBRA.   | 49 |
| <b>Table 5.6:</b>  | The activity concentration of radon in river waters from non-geothermally active non-HBRA.   | 49 |
| <b>Table 5.7:</b>  | The activity concentration of radon in pond waters from non-geothermally active HBRA.  | 50 |
| <b>Table 5.8:</b>  | The activity concentration of radon in pond waters from non-geothermally active non-HBRA.  | 50 |
| <b>Table 5.9:</b>  | The activity concentration of radon in water samples from geothermally active HBRA, non geothermally active HBRA and non-geothermally active non-HBRA. | 54 |
| <b>Table 5.10:</b> | The skewness and kurtosis of the activity concentration of radon in spring, river and pond waters.   | 56 |
| <b>Table 5.11:</b> | SIMCA classification of geothermal water samples based on the activity concentration of radon.   | 69 |
| <b>Table 5.12:</b> | The classification of geothermal water samples based on the activity concentration of radon (the Cooman's plot).                                       | 71 |

## LIST OF FIGURES

|                    |   |    |
|--------------------|---|----|
| <b>Figure 1.1:</b> | Sources of ionizing radiation in the environment (UNSCEAR, 2008).   | 1  |
| <b>Figure 3.1:</b> | An illustration of the sequence of events in liquid scintillation process.  | 22 |
| <b>Figure 3.2:</b> | Parent and daughter radionuclides at secular equilibrium.   | 25 |
| <b>Figure 3.3:</b> | Dendrogram showing how HCA partitions data into non overlapping clusters.   | 30 |
| <b>Figure 3.4:</b> | An illustration of the Cooman's plot.   | 32 |
| <b>Figure 4.1:</b> | Sampling map and profile drawn using GPS data from sampling points.   | 34 |
| <b>Figure 4.2:</b> | Schematic diagram of the components of a contemporary liquid scintillation analyzer.  | 35 |
| <b>Figure 5.1:</b> | The bar graphs of the activity concentration of radon in spring waters from geothermally active HBRA, non-geothermally active HBRA and non-geothermally active non-HBRA.          | 51 |
| <b>Figure 5.2:</b> | The bar graphs of the activity concentration of radon in river waters from geothermally active HBRA, non-geothermally active HBRA and non-geothermally active non-HBRA.           | 51 |
| <b>Figure 5.3:</b> | The bar graphs of the activity concentration of radon in pond waters from geothermally active HBRA, non-geothermally active HBRA and non-geothermally active non-HBRA.            | 52 |
| <b>Figure 5.4:</b> | The bar graphs of the average activity concentrations of radon in water samples from geothermally active HBRA, non-geothermally active HBRA and non-geothermally active non-HBRA. | 54 |
| <b>Figure 5.5:</b> | PCA score plots of the activity concentration of radon in spring water samples from Homa mountain.  | 58 |

|                     |  |    |
|---------------------|--|----|
| <b>Figure 5.6:</b>  | PCA loading plots of the activity concentration of radon in spring water samples from Homa mountain. | 59 |
| <b>Figure 5.7:</b>  | PCA score plots of the activity concentration of radon in river water samples from Homa mountain.    | 60 |
| <b>Figure 5.8:</b>  | PCA loading plots of the activity concentration of radon in river water samples from Homa mountain.  | 60 |
| <b>Figure 5.9:</b>  | PCA score plots of the activity concentration of radon in pond water samples from Homa mountain.     | 61 |
| <b>Figure 5.10:</b> | PCA loading plots of the activity concentration of radon in pond water samples from Homa mountain.   | 62 |
| <b>Figure 5.11:</b> | Dendrogram of the activity concentration of radon in spring water samples from Homa mountain area.   | 64 |
| <b>Figure 5.12:</b> | Dendrogram of the activity concentration of radon in river water samples from Homa mountain area.    | 66 |
| <b>Figure 5.13:</b> | Dendrogram of the activity concentration of radon in pond watersamples from Homa mountain area.      | 67 |
| <b>Figure 5.14:</b> | The Cooman's plot of the test samples to Bala and Kanam models.                                      | 70 |
| <b>Figure 5.15:</b> | The Cooman's plot of the test samples to Bala and Kanjera models.                                    | 70 |
| <b>Figure 5.16:</b> | The Cooman's plot of the test samples to Kanam and Kanjera models.                                   | 71 |
| <b>Figure 5.17:</b> | The modeling power of the activity concentration of radon, pH and temperature of geothermal waters.  | 73 |
| <b>Figure 6.1:</b>  | Graphical relation between the activity concentrations of radon with pH.                             | 76 |
| <b>Figure 6.2:</b>  | Graphical relation between the activity concentrations of radon with temperature.                    | 76 |

**LIST OF ABBREVIATIONS AND ACRONYMS**

|         |   |
|---------|---|
| ADC     | Analog to Digital Converter.  |
| ANN     | Artificial Neural Network.  |
| EPA     | Environmental Protection Agency.  |
| GPS     | Geographical Positioning System.  |
| HBRA    | High Background Radiation Area.   |
| IAEA    | International Atomic Energy Agency.                                     |
| IPA     | Instrument Performance Assessment.                                      |
| kNN     | k-Nearest Neighbor.   |
| LDA     | Linear Discriminant Analysis.   |
| LSA     | Liquid Scintillation Analyzer.  |
| LSC     | Liquid Scintillation Counter.   |
| MCA     | Multichannel Analyzer.  |
| NORMS   | Naturally Occurring Radioactive Materials.                              |
| PCA     | Principal Component Analysis.   |
| PGNAA   | Prompt Gamma-Ray Neutron Activation Analysis.                           |
| PMT     | Photomultiplier Tube.   |
| SAS     | Statistical Analysis System.  |
| SIMCA   | Soft Independent Modeling of Class Analogy.                             |
| SNC     | Self Normalization and Calibration.                                     |
| SSNTDs  | Solid State Nuclear Track Detectors.                                    |
| UNSCEAR | United Nations Scientific Committee on the Effects of Atomic Radiation. |
| USEPA   | United States Environmental Protection Agency.                          |

## ABSTRACT

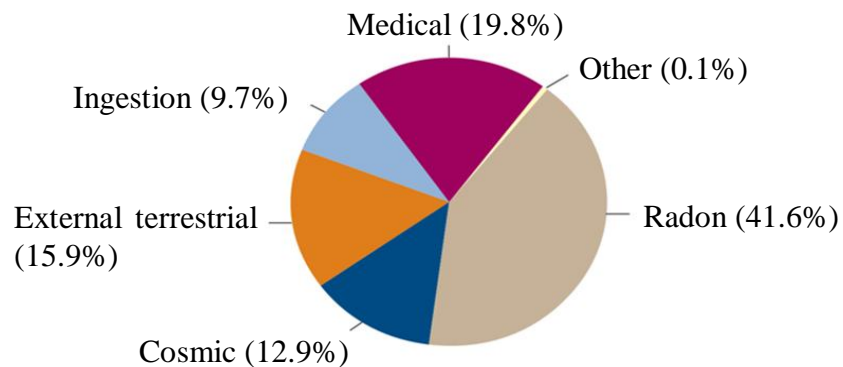
The activity concentration of radon in geothermal waters from Homa mountain region of south-western Kenya has been measured. Homa mountain is one of the regions in Kenya with elevated background radiation. This region is covered with geothermal grass and has some cold and hot springs. The activity concentration of radon in spring waters, river waters and pond waters from (i) geothermally active HBRA (ii) non-geothermally active HBRA and (iii) non-geothermally active non-HBRA areas of this region was measured using Liquid Scintillation Counting (LSC) method. The average concentrations of radon in spring waters, river waters and pond waters were above USEPA contamination limit of 11 Bq/L. Multivariate chemometric techniques namely Principal Components Analysis (PCA), Hierarchical Cluster Analysis (HCA) and Soft Independent Modeling of Class Analogy (SIMCA) were applied on the activity concentration of radon in the geothermal water samples. Principal components analysis was used to decipher patterns within the data set based on the activity concentrations of radon, pH and temperature. The score plots displayed distinct clusters in relation to the water source. Hierarchical cluster analysis was used to partition the radon concentrations into non-overlapping clusters; the dendrograms revealed that radon concentration in geothermal fluids form clusters which can be related to the water source as displayed by the PCA score plots. The formation of these clusters can be regarded to be controlled by geothermal activities in this region. Diagnostics of the geothermal potential in a typical high background radiation area based on the radon signatures was performed using **SIMCA** approach, in order to identify local models for geothermal waters from Homa mountain and to predict a probable class membership for new observations from the same region. Test samples were successfully classified into respective models after the Cooman's plot with 64% efficiency. The modeling power of the activity concentration of radon in relation to pH and temperature in this study was more than 0.3. Therefore, these variables were important in constructing SIMCA models in this study. In order to increase this classification efficiency, other factors such as chemicals and isotopic components in geothermal waters need to be investigated and SIMCA models constructed on the basis of these chemicals and isotopic components and the activity concentration of radon in a geothermally active HBRA.

## CHAPTER 1

### INTRODUCTION

#### 1.1 Background to the study

Radiation constitutes an important part of the environment. The main sources of radiation in the environment are primordial, cosmic and artificial radionuclides. Primordial radionuclides are leftovers from when the world and the universe were created and originate mainly from the interior of stars. Primordial radionuclides include  $^{235}\text{U}$ ,  $^{238}\text{U}$ ,  $^{232}\text{Th}$ ,  $^{226}\text{Ra}$ ,  $^{222}\text{Rn}$  and  $^{40}\text{K}$ . Radon gas and its progeny contribute almost half of the dose exposure reaching the general public (Hussein, 2008; Seyis *et al.*, 2010). Cosmogenic radionuclides ( $^{14}\text{C}$ ,  $^3\text{H}$ ,  $^7\text{Be}$ ,  $^{36}\text{Cl}$ ,  $^{32}\text{Si}$  and  $^{22}\text{Na}$ ) are continually formed in the atmosphere due to cosmic ray interactions (USCEAR, 2008). Artificial radiations are enhanced or formed due to human actions. These sources of artificial radiations include: medical sources, industrial sources, nuclear explosions, nuclear power and nuclear radiation accidents (USCEAR, 2008). Figure 1.1 shows the percentage contribution of various radiation sources to the environment.



**Figure 1.1:** Sources of ionizing radiations to the environment (UNSCEAR, 2008).

A geothermal source is a natural heat from the earth which can be exploited economically if it is close to the surface of the earth (Hussein, 2008). Geothermal energy is limited to the exploitation of hydrothermal-convection system. Hydrothermal-convective systems tend to occur in geologic zones of tectonic, volcanic, or orogenic activity and contain natural fluids which circulate and transfer heat. Convection results when fluids heated at the base of a system expand and rise due to buoyant forces. Hot water springs and vapor, fumaroles and geysers normally suggest prospecting geothermal energy sources (Stoker and Kruger, 1975).

Geochemical methods are useful in evaluating the potential of a geothermal resource. For example, chlorine content can be used to distinguish between hot-water and vapour dominated systems. Dissolved constituents like  $\text{SiO}_2$  and Na-K-Ca relations are used as geothermometers. Ratios of oxygen and hydrogen isotopes may lead to the identification of the source of water in the reservoir (Lambrakis and Stamatis, 2008). These methods are expensive, labour intensive and time consuming. In addition the drilling of deep wells for flow testing and to determine the extent of the geothermal source is expensive (Hameed *et al.*, 1990). Therefore there is need to choose the drilling targets carefully. This is achieved by carrying out surface exploration to obtain indirect information on the extent of the geothermal reservoir, the surface temperatures and permeability control. During this exploration phase, subsurface temperatures are estimated using isotope and chemical geothermometers as well as mixing models. The source of the water is also

identified using isotopic techniques (IAEA, 2000). These techniques are expensive and labour intensive. Therefore, there is need to try less expensive techniques such as radon mapping to determine the potential drilling targets.

Although radon and geothermal activity have little relationship in terms of origin, radon buried deep into the earth crust finds an easy passage to reach the top surface through faults normally associated with geothermal sources (Hameed *et al.*, 1990). Therefore, radon distribution in geothermal field fluids could exhibit unique patterns which are a characteristic of a geothermal field. To identify these patterns, multivariate chemometric techniques such as principal components analysis (PCA) and hierarchical cluster analysis (HCA) need to be applied on the activity concentrations of radon. Soft independent modeling of class analogy (SIMCA) could for instance be suitable in classifying and predicting class membership of samples on the basis of activity concentration of radon (Dragovic and Onjia, 2007).

## **1.2 Radon**

Radon is a colorless, odorless, tasteless, chemically inert and radioactive gas produced in rocks and soils through  $\alpha$ -decay of radium-226 with some atoms escaping to groundwater and air (Badhan *et al.*, 2010; Bonotto and Caprioglio, 2002). It has three isotopes,  $^{222}\text{Rn}$  (radon),  $^{220}\text{Rn}$  (thoron) and  $^{219}\text{Rn}$  (actinone).  $^{222}\text{Rn}$  results from the uranium-238 decay

series which has 99.28 % relative abundance of uranium by weight;  $^{220}\text{Rn}$  results from the  $^{232}\text{Th}$  chain and  $^{219}\text{Rn}$  results from the  $^{235}\text{U}$  chain which has 0.71 % of the relative abundance of uranium by weight. The most important isotope of radon is  $^{222}\text{Rn}$ .  $^{222}\text{Rn}$  has a half-life of 3.82 days, decays by emitting 5.49 MeV  $\alpha$  particle, and produces radioactive progeny (Badhan *et al.*, 2010; Sulekha and Sengupta, 2009).  $^{222}\text{Rn}$  and two of its daughters,  $^{218}\text{Po}$  and  $^{214}\text{Po}$ , are alpha emitters, while  $^{214}\text{Pb}$  and  $^{214}\text{Bi}$  are beta/gamma emitters (Kolarz *et al.*, 2009). However,  $^{220}\text{Rn}$  and  $^{219}\text{Rn}$  are ignored in these studies because they have short half-lives ( $T_{1/2}=55.3$  seconds and  $T_{1/2}=3.96$  seconds respectively) and due to the fact that they are generally lower in concentration than  $^{222}\text{Rn}$  in geological material. Table 1.1 shows the decay scheme of  $^{222}\text{Rn}$  and the major decay progenies and properties.

**Table 1.1:**  $^{222}\text{Rn}$ -decay and its progeny scheme (adapted from Wilkening, 1990; Lee and Kim, 2006)

| Isotope      | Mode of decay | Half-life         | Main Energies (KeV) |
|--------------|---------------|-------------------|---------------------|
| Radon-222    | $\alpha$      | 3.82 days         | 5490                |
| Polonium-218 | $\alpha$      | 3.05 minutes      | 6000                |
| Lead-214     | $\beta$       | 26.8 minutes      | 352                 |
| Bismuth-214  | $\beta$       | 18.8 minutes      | 609                 |
| Polonium-214 | $\alpha$      | 163 $\mu\text{S}$ | 7690                |

### 1.3 Sources of radon in the environment

Radon is a decay product of  $^{226}\text{Ra}$  which is a daughter product of  $^{238}\text{U}$  in soil, rocks and water. It is transported mainly upwards from deep layers of the earth by underground waters and diffusion through soils and cracks of rocks under the bedrock in the ground to the surface (Ioannides *et al.*, 2003). The rate at which radon is transported is affected by the distribution of radium in the soil and bedrock, the porosity of the soil, humidity, micro cracks, granulation, surface winds, temperature, rainfall and pressure (Al-Tamimi and Abumurad, 2001; Baykara *et al.*, 2007; Papp *et al.*, 2008). Secondary contributors of radon include natural gas, geothermal fluids, volcanic gases, ventilation from caves and mines, and combustion of coal (Ganesh *et al.*, 2008). Radon may be generated in geothermal areas due to the presence of small quantities of radioactive eruptive rocks containing uranium that lies in the path of the passage of geothermal waters and it may be transported along with an influx of magnetic gases such as  $\text{CO}_2$ , He,  $\text{H}_2$ , \ and  $\text{CH}_4$  from a deeply burned magma chamber (Ghose *et al.*, 2003; Das *et al.*, 2009; Katsanou *et al.*, 2010).

#### 1.4 Radon gas in geothermal exploration

Radium-223 and radium-224 have potential for being used to investigate the extent of geothermal fluid-rock interaction. Their levels are indicative of the mixing of different sources of fluids for example injection and primary geothermal fluids. However, field measurement and sampling of  $^{223}\text{Ra}$  and  $^{224}\text{Ra}$  are hampered by the fact that  $^{223}\text{Ra}$  and  $^{224}\text{Ra}$  are trace radioisotopes of high reactivity (IAEA, 2000).

Along with radium isotopes, radon level and occurrence can indicate geothermal fluid source and age (IAEA, 2000). The concentration of radon in geothermal fluids reaching the surface not only depends on geologic properties of the formation media, but also the radon emanation and the volumetric properties of the porous or fractured media. The flow rate determines the time available for accumulation or decay of radon in a unit volume of fluid before it reaches the surface through a well or a fissure. Radon and radium have long been associated with hot springs; and radon has been observed in conjunction with some geothermal phenomena such as volcanoes, geysers, hot springs, mud pots and fumaroles. Radon occurs in geothermal reservoir fluids that are mostly associated with high enthalpy or recent discharge (IAEA, 2000).

Although radon has been studied for a long time, its characteristics in geothermal reservoirs are still not well understood; hence its present application in geothermal investigations is limited (IAEA, 2000). This is because; this gas has normally been

analyzed using univariate statistical techniques. Univariate techniques t-test, ANOVA and multiple regression are limited in analyzing environmental data because environmental data is multivariate in nature. Therefore, there is need to apply multivariate techniques to assess the utility of this gas in geothermal field diagnostics. Applying multivariate chemometric techniques such as PCA and HCA can reveal patterns which are a characteristic of geothermal activities.

### **1.5 Geothermal fields lying in HBRA**

Geothermally active areas with elevated radioactivity around the world are found in Yangjiang in China, Kerala and Orissa in India, Guarapari in Brazil and Ramsar in Iran (Giassi-nejad *et al.*, 2002; Baranwal *et al.*, 2006; Elham *et al.*, 2012). The high background radiation associated with the geothermal springs in Ramsar is due to the presence of high levels of radium and its decay products in the area. Groundwater in such places is heated by subsurface geologic activity and passes through uraniferous igneous rocks. Radium is dissolved from the rocks by hot groundwater but uranium is not dissolved because ground water is anoxic and uranium is insoluble in anoxic waters (Giassi-nejad *et al.*, 2002).

Homa mountain in south western Kenya among other areas in the country such as Kerio valley is a geothermal field that is associated with elevated background radiation from

naturally occurring radioactive materials (NORMs) (Mustapha *et al.*, 1997).

The mountain is located in Karachuonyo constituency, Homa Bay county and is bound by latitude 00° 30' N and 00° 20'N and longitude 33° 26' E and 34° 34'E. The 1754-m-high flat-topped summit of the volcano rises about 600 m above the lake and has a series of cone sheets of carbonatite alkaline rocks and ijolite (alkaline igneous rocks) which have been documented to have high radiation levels (Mangala, 1987; Patel, 1991). Geothermal activity in Homa mountain occurs at the surficial environment through a few widespread hot/warm springs, gas release in the springs and the *Fimbristylis exilis* 'geothermal grass' which is widespread in the area and is also common in other geothermal areas within the main Kenya Rift Valley. The soils of these areas are mainly clay and in some parts clay mixed with sand. Radon concentration in geothermal fluids from this region is above USEPA acceptable contamination limit of 11Bq/L (Badhan *et al.*, 2010). Due to this elevated radiation and geothermal activities in this region, it is important to understand the relationship between geothermal activity and radon signatures.

### **1.6 Methods of measuring radon**

Radon in water may be measured using several methods. These methods and techniques are: Aeration of water containing radon using a carrier gas (nitrogen) and continuous and discontinuous measurement in an ionization chamber (e.g scintillation chamber, Lucas cell or ATMOS 12 DPX); gamma ray spectrometry using gamma rays from the  $^{222}\text{Rn}$

daughters  $^{214}\text{Pb}$  and  $^{214}\text{Bi}$ ; radon diffusion chamber equipped with solid state nuclear track detectors (SSNTDs), and electrets ion chambers based on the use of E-perm (electrets-passive environment). These methods require large sample volume and they are time consuming (Amrani and Cherouati, 1999).

In this study, the liquid scintillation counting (LSC) technique was used to quantify the concentration of radon in water samples because radon is highly soluble in organic solvents which are used in liquid scintillation cocktails. In LSC beta and alpha particles are brought into close proximity to liquid scintillator and produce detectable scintillations without self-absorption losses which occur in many other methods because of the short range of beta particles. This technique has high sensitivity and requires small sample volume of 10 ml (Vesterbacka *et al.*, 2010).

### **1.7 Multivariate chemometric analysis of radon concentration**

Chemometrics is the process of analyzing information from data based on many variables with various numerical techniques in order to extract useful information. Mathematical, statistical and other logic methods are applied on the data with the aim:

- (i) To evaluate and to interpret data.
- (ii) To optimize and to model processes and experiments.

- (iii) To extract a maximum of chemical (and hence analytical) information from experimental data (Einax *et al.*, 1997).

In this research, principal components analysis (PCA) was used for data dimensionality reduction and deciphering patterns of radon concentration within the set of data. Hierarchical cluster analysis (HCA) was used to partition the data into non overlapping clusters in order to assess the consistence of the patterns displayed by PCA. Soft independent modeling of class analogy (SIMCA) was used to identify local models for possible groups of water samples and to predict a probable class membership for new observations from the same geological area.

### **1.8 Statement of the research problem**

Radon signals have been measured and analyzed using univariate statistics in geothermal fields lying in HRRA and used for locating geothermal energy sources in some countries such as New Zealand, Mexico and USA (Hssein, 2008). However, there are no diagnostic models for geothermally active high background radiation areas based on radon signatures. Therefore there is need to analyze the activity concentration of radon from geothermal fields in Homa mountain using multivariate techniques because environmental data is multivariate in nature and construct models which can be used to identify possible groups of geothermal waters and to predict a probable class membership

for new observations from the same geological area.

## **1.9 Objectives**

### **1.9.1 Main objective**

The goal of this work was to determine the activity concentration of radon in geothermal fluids associated with high background radiation area using liquid scintillation spectrometry and to perform multivariate chemometric analysis on the activity concentration of the radon in order to explore and understand the underlying patterns, relationships and trends of radon activity concentration in relation to geothermal activity in typical HBRA.

### **1.9.2 Specific objectives**

- (i) To sample surface and ground waters from Kanjera, Kanam and Bala areas which are geothermally active in a high background radiation area and to determine the activity concentration of radon gas in these waters using liquid scintillation counting technique.
- (ii) To perform multivariate chemometric analysis on the activity concentrations of radon, pH and temperature of geothermal water.
- (iii) To use the above information to perform simple geothermal potential diagnostics in a typical high background radiation area based on the radon signatures.

### **1.10 Rationale of the study**

Radon's unique properties as a natural radioactive inert gas have led to its use as geophysical tracer. Liquid scintillation counting technique is used to measure the activity concentration of radon because it has high sensitivity, requires small sample volume and evaluates large number of samples per day (Vesterbacka *et al.*, 2010). LSC also has advantageous  $4\pi$  measuring geometry and the beta/alpha particles are brought in close proximity to the liquid scintillator and produce a detectable scintillation without self-absorption losses (Pates and Mullinger, 2007).

Radon levels in geothermal fields cannot be explained using univariate statistics. This is because radon activity concentrations in geothermal fields are governed by a number of factors including temperature and pH. Therefore, application of multivariate chemometric techniques on radon activity concentrations, temperature and pH of geothermal waters sampled in HBRA can reveal patterns which can be regarded as a characteristic of geothermal signature which can of great interest to the public and the Ministry of Energy. The results of this work can act as a guiding tool to the Ministry of Health and the Kenya Radiation Protection Board in formulating regulatory policy because drinking water in Homa mountain is from underground sources.

## CHAPTER 2

### LITERATURE REVIEW

#### **2.1 Comparison of Lucas cell and LSC methods for measuring radon in water**

Prichard and Gassel, 1977, measured radon in water using both Lucas cell and LSC methods. The results obtained by the Lucas cell method was plotted against the results obtained by the LSC method. The slope of the graph obtained was 0.934 indicating that the results were similar. In this study, LSC method was preferred to Lucas cell method because LSC technique combines the advantages of minimum sample preparation time (1 min/sample), small sample size (10 mL), and high sensitivity.

#### **2.2 Studies of radon in high background radiation areas (HBRAs)**

There are areas in the world with background radiation levels of more than 3 mSv/year (Pao-Shan and Chien-Li, 1994; Tavera *et al.*, 1999; Baranwal *et al.*, 2006). Malanca *et al.*, 1993, reported such levels have been found in Brazil, Jordan, Taiwan, China, India, France, Italy, Greece and Austria. These areas have been under study for many years in order to determine the utility of radon in fault mapping and the exposure level of radon to the public. In Kenya these areas include Jombo (Mrima) in the Mombasa-kwale area of southern Kenya, Ruri and Homa mountain in western Kenya, Usaki near Homa bay in

south Nyanza and Kerio valley in Baringo (Mustapha *et al.*, 1997; Ogwari *et al.*, 2010).

Radon and thoron levels in soil, water and indoor atmosphere of Budhakedar were measured in Garhwal Himalaya, India, which is a high background radiation area (Ganesh *et al.*, 2008). Radon emanometer was used to measure radon in groundwater and solid state nuclear track detector was used to measure indoor radon. Radon measured in groundwater varied from  $8 \pm 1$  Bq/L -  $3.05 \pm 0.02$  kBq/L. The high values of radon concentrations were as a result of presence of uranium mineralization in the area. Areas with granite rocks had higher contents of uranium hence high levels of radon than those with volcanic rocks which had low uranium content.

Radon gas was monitored in five active fault sites in northern and northwestern Greece using CR-39 solid state nuclear track detectors (SSNTDs) (Ioannides *et al.*, 2003). It was observed that radon levels are connected to the local tectonic structure, hence radon signature can be used to detect and map active fault zones.

Radon was measured in limestone quarry areas of Jordan using SSNTDs CR-39 in order to determine the suitability of radon levels to confirm fault zones and to help in radon mapping for epidemiological studies in Jordan (Al-Tamimi and Abumurad, 2001;). The results showed anomalous radon levels ( $50 \text{ kBq/m}^3$ ) in the sampled stations near fault zones. The anomalous levels were a characteristic of fault zones hence the radon gas proved to be a good tool for fault zones detections.

The activity concentrations of radon in drinking water and indoor air in various locations in Kenya were measured using liquid scintillation counting technique and passive integrated devices respectively (Mustapha *et al.*, 2002). The mean and maximum values of radon in water were 37 Bq/L and 410 Bq/L; and in air 100 Bq/m<sup>3</sup> and 1160 Bq/m<sup>3</sup> respectively. In some areas the levels of radon were higher than internationally recommended reference levels (11.1Bq/L).

Radon concentration and terrestrial gamma-radiation dose rates were assessed in the seismically active area of East Anatolian Active Fault System (EAFS), Turkey) by Baykara *et al.*, 2008. Soil <sup>222</sup>Rn concentration was measured using time integrated passive dosimeter, CR-39 detector, and prompt measurements system constructed by silicon detector and air pump to suck the soil air. The mean <sup>222</sup>Rn concentrations were found to lie between 1.2 - 3.6 kBq/m<sup>3</sup> using time integrated passive dosimeter and CR-39 detector and between 2 - 70 kBq/m<sup>3</sup> using passive and prompt methods, respectively. These results were a characteristics of the bedrock and soil covering the area. The terrestrial gamma-dose rate correlated positively with radon concentration throughout the fault line. The result supported the idea that the dissolved uranium and its daughter nuclide may easily precipitate close to the ground surface or fracture zones.

The influence of variable stress on underground radon concentrations was studied using a natural laboratory under a reservoir (Kies *et al.*, 2002). There was a significant change in

radon concentration in boreholes under water reservoir of a pumped storage station induced by variable water levels. Radon activity concentration was observed to exhibit a direct negative correlation with water levels. Long-scale interactions through a microscale network of clustered cracks and fractures and percolation defects may be responsible for these observations.

Radon concentration in ground water was measured using RAD7 and the average annual dose was assessed in the environs of NITJ, Punjab, India (Badhan *et al.*, 2010). The radon in drinking water was found to vary from 2560 - 7750 Bq/m<sup>3</sup> with an average of 5143.33 Bq/m<sup>3</sup>. No correlation was observed between pH value and radon concentration from ground water. The calculated indoor radon concentration values varied from 74 - 190 Bq/m<sup>3</sup> with an average values of 124.50 Bq/m<sup>3</sup>. The calculated values for the absorbed dose were in the range 1.26-3.24 mSv.y<sup>-1</sup>, which were below the action level.

### **2.3 Studies of radon in geothermally active HBRAs**

The activity of geological structures using radon mapping was investigated in Los Azufres geothermal energy field (Tavera *et al.*, 1999). Radon measurements were carried out by means of 240 samplers, each composed of a thin LR-115 plastic detector. From the investigation, it was found that low radon concentration values (< 20 Bq/L) were measured from sites with hot springs and high values (> 20 Bq/L) were measured from

sites without hot springs. The observed results are consistent with those obtained in this work, that is, the activity concentration of radon measured from sites with hot springs was lower than in sites without hot springs. The variation in radon concentration in the two sites can be attributed to temperature variation. Radon is susceptible to loss at high temperature leading to low concentration of radon in the sites with hot springs.

Radon in Tainese geothermal waters and cold springs was measured using an electret ion chamber (Pao-Shan and Chien-Li, 1994). Radon concentration ranged from 0.44 - 9.54 Bq/L. It was observed that the highest radon concentrations were measured from volcanic regions. This can be attributed to high distribution of radium in the bedrock.

Radon concentration was measured in groundwaters and sand stone samples from Guarany aquifer and Parana sedimentary basin in south America using liquid scintillation counting method (Bonotto and Caprioglio, 2002). The research was aimed at the application of dissolved radon data for hydrochemical prospecting. The measured radon concentration ranged between 0.111 – 122.211 Bq/L. It was observed that high radon concentration does not only depend on the accumulation of high uranium levels in the sites of measurement but also many factors such as fluid advection, diffusion, emanation and permeability of the rocks.

Radon was monitored during the flow test of wells at Chingshui geothermal field to investigate mechanisms for interpreting radon behaviour (Chen *et al.*, 2011). Radon was

measured using SARAD RTM 2100 detector and ranged from  $54 \pm 29$  Bq/m<sup>3</sup> -  $983 \pm 65$  Bq/m<sup>3</sup>. It was observed that radon concentration increased as a step function of cumulative production during the flow test. This behaviour implied a radial composite system which is induced by the formation damage with carbonate scales near the wells.

#### **2.4 Multivariate chemometric analysis of environmental data**

The application of multivariate chemometric analysis on data is aimed at: compression of large data set, extraction of information, elimination of 'noise' and redundancy, visualization of multidimensional data sets and hypothesis formation (Kowalik, C. and Einax, J.E., 2006). These techniques may be applied on environmental data with the aim of improved interpretation of the data which cannot be achieved by only univariate statistical methods. In this work, PCA, HCA and SIMCA approaches have been used. These techniques have been applied on some environmental studies elsewhere and they have interpreted data better than univariate statistical methods.

Trace elements from Nigerian coal were analyzed using hierarchical cluster analysis (Ewa, 2004). The analysis was aimed at investigating existing similarities for different types of coals mined in Nigeria based on the concentration of trace elements in each coal. HCA could recognize patterns useful for interpretation of coalification histories and prediction of fuel ranking for Nigeria coal.

Petrographic and radiometric features were investigated using multivariate analysis of volcanic activities of the island of Lipai, Italy in order to characterize the volcanic products and highlight some evolutionary aspects (Lanzo, 2010). Lavas and pyroclastic products were sampled and their geochemical elements and primordial radionuclides composition were determined by XRF and gamma ray spectrometry via HPge detector. Petrographic (Tas and  $K_2O$  vs  $SiO_2$ ) was performed using the geochemical data toolkit software. Multivariate analysis via PCA and HCA was used as an independent check of geochemical classification. Both PCA and HCA grouped the data into clusters (geochemical groups) coherently with the classification shown in the petrographic plots.

Multivariate statistical analysis was used to study groundwater flow in a regional aquifer (Klaus *et al.*, 1999). Major cation and trace element concentrations were measured from 11 springs in Ash Meadows National Wildlife Refuge. PCA was performed on the major cation and trace element concentrations. The PCA results indicated that multivariate statistical techniques can be used to assess possible groundwater flow regimes. Finally, PCA of hydrochemical data could be used to support hydrological and isotopic interpretations of ground water flow, and possibly the results of groundwater flow modeling.

Materials were classified for explosives from prompt gamma spectra using principal components analysis (PCA) (Hee-jung *et al.*, 2009). The study was aimed at applying

PCA on elemental composition data and gamma spectra obtained from prompt gamma-ray neutron activation analysis (PGNAA) for effective pattern recognition to discriminate explosives from non explosives. The PCA results obtained did not provide a perfect separation of explosives from non explosive materials because materials containing nitrogen such as silk and acryl were found in both explosive and non explosive groups; hence SIMCA of PCA was adopted for better classification. It was demonstrated that pattern recognition can be applied effectively for discrimination of explosives from non explosive materials based on elemental composition data and gamma spectra.

Multivariate chemometric techniques were used to recognize and classify soils of unknown geographic origin, based on radionuclide ( $^{226}\text{Ra}$ ,  $^{238}\text{U}$ ,  $^{235}\text{U}$ ,  $^{40}\text{K}$ ,  $^{134}\text{Cs}$ ,  $^{137}\text{Cs}$ ,  $^{232}\text{Th}$  and  $^7\text{Be}$ ) activities obtained using gamma ray spectrometry. Linear discriminant analysis (LDA), k-nearest neighbour (kNN), SIMCA and artificial neural network (ANN) were used. The classification ability of LDA, kNN, SIMCA and ANN were 82.8%, 88.6%, 60% and 92.1% respectively (Dragovic and onjia, 2007).

The method of using radon signals for locating geothermal energy sources has been used in countries such as New Zealand, Mexico and USA (Hussein, 2008). Although some success have been met, radon has not been used as a diagnostic tool of geothermal signature because the source of radon is not geothermal activity although the levels of radon are affected by geothermal activities (IAEA, 2000). Homa mountain in south western Kenya is one of the areas which is geothermally active and with elevated

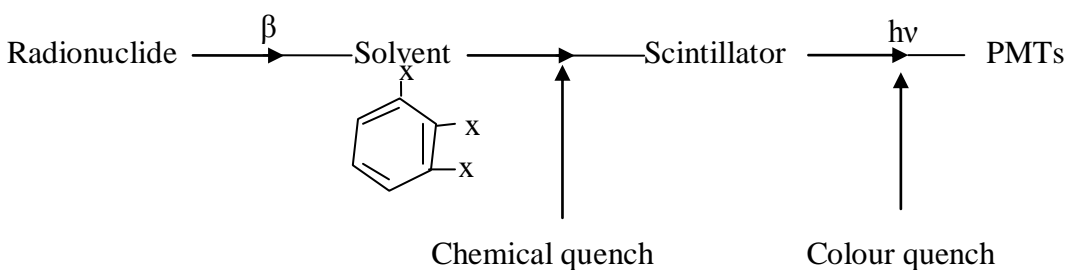
background radiations. Due to this, radon was measured in this area and analyzed using multivariate chemometric techniques in order to understand underlying patterns which are a characteristic of geothermal activity.

**CHAPTER 3**

**PRINCIPLES OF LIQUID SCINTILLATION SPECTROMETRY AND  
MULTIVARIATE CHEMOMETRICS**

**3.1 Introduction**

In liquid scintillation, the beta/alpha particles are brought in close proximity to the liquid scintillator and produce a detectable scintillation without self-absorption losses (Pates and Mullinger, 2007). LSC also has advantageous  $4\pi$  measuring geometry. Scintillation involves the transfer of kinetic energy from the beta particles or alpha particles emitted by the radionuclide to solvent molecules. The excited solvent molecules then transfer their energy to scintillator molecules. The scintillator molecules become excited, and return to their stable energy state by emitting photons with certain intensity. The photons are converted into photoelectrons at the photomultiplier photocathode and the resulting pulse is amplified and counted by multichannel analyzer (MCA) (Makinen, 1995).



**Figure 3.1:** An illustration of the sequence of events in liquid scintillation process (Makinen and Pekka, 2001).

### **3.2 Comparison of alpha/beta LSC and conventional beta LSC detectors**

There are two types of liquid scintillation counting detectors; that is, alpha/beta LSC and conventional beta LSC. Spaulding and Noakes (1993) determined radon concentration from private wells using both the alpha/beta LSC and the conventional beta LSC detectors. The results obtained were almost similar although the results from the alpha/beta detector were more accurate. This is because all five particles ( $3\alpha$  and  $2\beta$ ) emitted due to radon were detected by alpha/beta LSC detector whereas only two particles ( $2\beta$ ) were detected by the beta LSC detector.

In Kenya there was no alpha/beta LSC detector when the samples in this work were measured. Therefore, in this study, beta LSC detector was used because alpha/beta LSC was not available.

### **3.3 Secular equilibrium of radon and its daughter radionuclides**

In order to determine the time required for radon to equilibrate with daughter radionuclides after geothermal water samples had been mixed with Ultima gold scintillation cocktail, it was necessary to determine secular equilibrium of radon and its daughter radionuclides. Secular equilibrium is a steady-state condition that the activity of a daughter radionuclide is equal to that of the parent radionuclide. It occurs when the parent half-life is much longer than that of the daughter. When a parent decays into a

daughter, the rate of change of the number of daughter atoms must be the difference between the rate of growth from the parent and the rate of decay of the daughter. Therefore the idea of radioactive equilibrium can be explained using the decay law as shown in equation 3.1.

$$\frac{dN_D}{dt} = \lambda_P N_P - \lambda_D N_D \quad (3.1)$$

where  $\lambda_P$  and  $\lambda_D$  are the decay constants of parent radionuclide and daughter radionuclide respectively, while  $N_P$  and  $N_D$  are number of atoms in parent radionuclide and daughter radionuclide respectively.

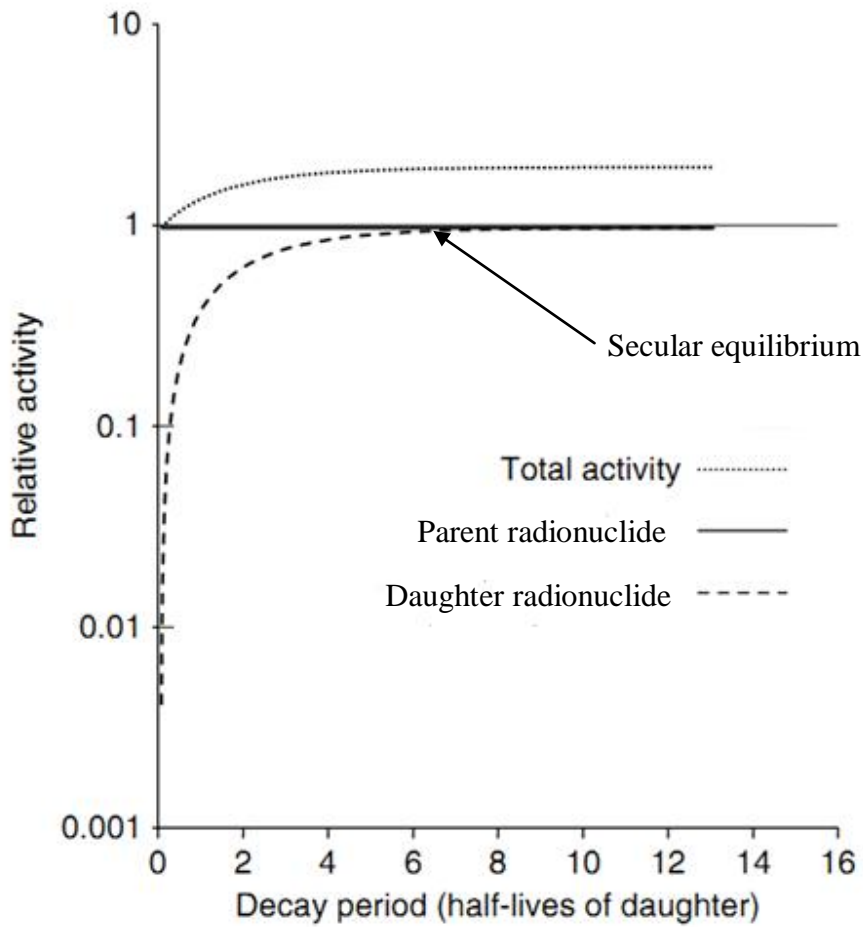
The decay constant is related to half-life as follows

$$\lambda = \frac{\ln 2}{t_{1/2}} \quad (3.2)$$

Secular equilibrium occurs when the activity of the daughter radionuclide is equal to that of the parent radionuclide

$$N_P = \frac{\lambda_D}{\lambda_P} N_D \quad (3.3)$$

as illustrated on Figure 3.2.



**Figure 3.2:** Parent and daughter radionuclides at secular equilibrium (Gordon, 2008). The activity levels are relative to the parent activity.

### 3.4 Principal components analysis (PCA)

Principal components analysis is a bilinear modeling method which gives an interpretable overview of the main information in a multivariate data. The correlated information carried by the original variables is projected onto principal components which are

themselves non correlated and orthogonal to each other. The first principal component covers the most variation in the original data. The second principal component covers the most remaining variation, and so on (Brai *et al.*, 2006; Lanzo *et al.*, 2010). A plot of principal components will display patterns which will enable one to view interrelationships between different variables, and detect and interpret sample, groupings, similarities or differences (Brereton, 2000; Bonifazzi *et al.*, 2000; Vinas *et al.*, 2007).

The goal of performing PCA in this work was to find and interpret any latent patterns which are characteristic of geothermal signature based on the activity concentration of radon, pH and temperature according to equation 3.4 (Brai *et al.*, 2006; Seddeek *et al.*, 2009). The central point of PCA analysis is to reduce the original (m , n)-data matrix X to the factor scores T and factor loadings P (Van der Graaf *et al.*, 2007; Kulahci and Sen, 2008).

$$X = T \cdot P^t + E \quad (3.4)$$

$$\begin{pmatrix} x_{11} & - & x_{1n} \\ x_{21} & - & x_{2n} \\ - & - & - \\ x_{m1} & - & x_{mn} \end{pmatrix} = \begin{pmatrix} t_{11} & - & t_{1s} \\ t_{21} & - & t_{2s} \\ - & - & - \\ t_{m1} & - & t_{ms} \end{pmatrix} \begin{pmatrix} p_{11} & - & p_{1n} \\ p_{21} & - & p_{2n} \\ - & - & - \\ p_{s1} & - & p_{sn} \end{pmatrix} + \begin{pmatrix} e_{11} & - & e_{1n} \\ e_{21} & - & e_{2n} \\ - & - & - \\ e_{m1} & - & e_{mn} \end{pmatrix}$$

where X is the data matrix, T is the factor scores, P is the factor loadings, P<sup>t</sup> is the transposed matrix of P, E is residuals as a result of reduction of dimensionality and s is the number of factors (Jurgen *et al.*, 1997).

The loadings describe the weight of each feature in each factor. Features with low loadings have only slight influence on the factor; features with high positive or negative loadings essentially determine the factor. From the loadings factors, one can conclude which features have a common influence on the objects. Common influences are present if more than one feature has high loadings in one factor. In geometrical terms a loading is the cosine of the angle between the variable and the current PC. The smaller the angle (that is the higher the link between the variable and the PC), the larger the loadings, which often range within  $\pm 1$ .

The score plots describe data structure in terms of sample patterns, and more generally they show sample differences and similarities (Yongming *et al.*, 2006). Samples with close scores along the same PC are similar (they have close values for the corresponding variables). Conversely, samples with high score differences are quite different from each other with respect to those variables.

### **3.5 Hierarchical cluster analysis (HCA)**

The purpose of HCA is to organize a set of data into clusters in a way that observations within a cluster are more similar to each other than they are to observations belonging to a different cluster (Martinez Angel and Martinez Wendy, 2005). Hierarchical Cluster Analysis examines the inter point distances between the measured variables and displays

that information in the form of a dendrogram. The dendrogram gives a visual representation of the distance at which clusters are combined. HCA is useful where statistical methods cannot satisfactorily interpret existing trends in data. The technique produces patterns that allow for a more resolved visualization of the similarities existing between the data (Ewa, 2004).

Similarity measure is usually determined using either Manhattan distance, Mahalanobis distance or squared Euclidean distance. In this work, the squared Euclidean distance was used according to equation 3.5. In this method, the distance between two objects is not affected by the addition of new objects to the analysis which may be outliers (Kulahci and Sen, 2008).

$$d_{ij} = \sum_{k=1}^n (x_{ik} - x_{jk})^2 \quad (3.5)$$

where,  $k$  is the number of repetitive measurements between the clusters  $i$  and  $j$ .

given two clusters (objects)  $i$  and  $j$ , the similarity distance  $d_{ij}$  exists only if  $d_{ij}=d_{ji}< 0$  where  $d_{ji}=1$  when  $x_i = x_j$ .  $x_i$  and  $x_j$  are the row vectors in the original measured data matrix  $X$ .

The Ward's method gives a minimum distance in hierarchical clustering. The method minimizes the sum of squares of the calculated Euclidean distances between any two

clusters using agglomerative strategies. The Ward's formula is given by equation 3.6 (Kulahci and Sen, 2008).

$$d_{mk} = \frac{n_i + n_k}{n + n_k} d_{ik} + \frac{n_j + n_k}{n + n_k} d_{jk} - \frac{n_k}{n + n_k} d_{ij} \quad (3.6)$$

where,

m is the new formed object or cluster.

i, j and k are the previously clustered objects.

$$n = n_i + n_j$$

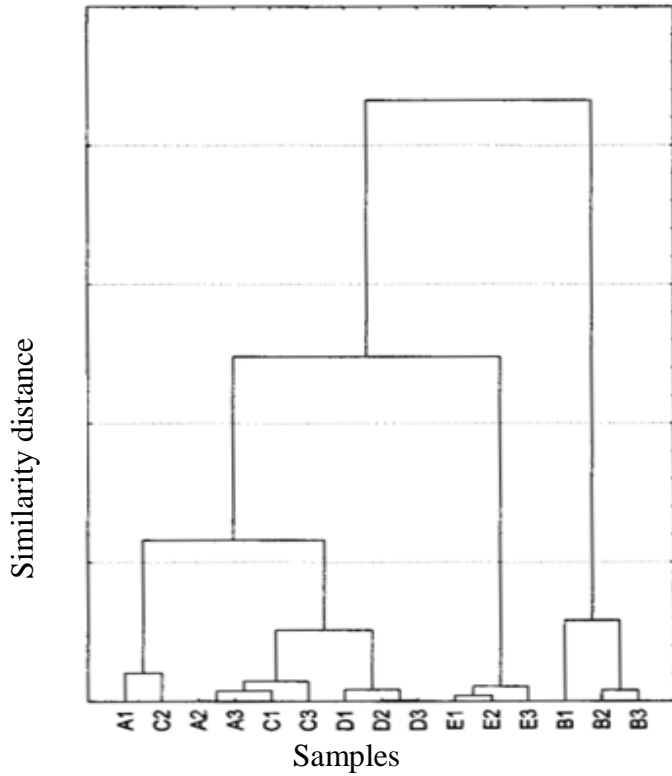
$n_i$  is the number of objects in cluster i

$n_j$  is the number of objects in cluster j

$d_{ik}$ ,  $d_{jk}$  and  $d_{ij}$  are the squared Euclidean distances.

In this method, nearest clusters are linked together until only one cluster remains (Simeonova *et al.*, 2010).

Figure 3.3 is an example of a dendrogram generated using HCA. This dendrogram is read from left to right and the vertical lines show joined clusters.



**Figure 3.3:** Dendrogram showing how HCA partitions data into non-overlapping clusters (Einax *et al.*, 1997).

### 3.6 Soft independent modeling of class analogy (SIMCA)

SIMCA is a classification method based on constructing PCA model on each class. It models the similarities between members of the same class. For each class  $i$  a PCA is performed according to equation 3.7, which leads to a principal component model (disjoint class model).

$$X_i = T_i \cdot P_i^t + E_i \quad (3.7)$$

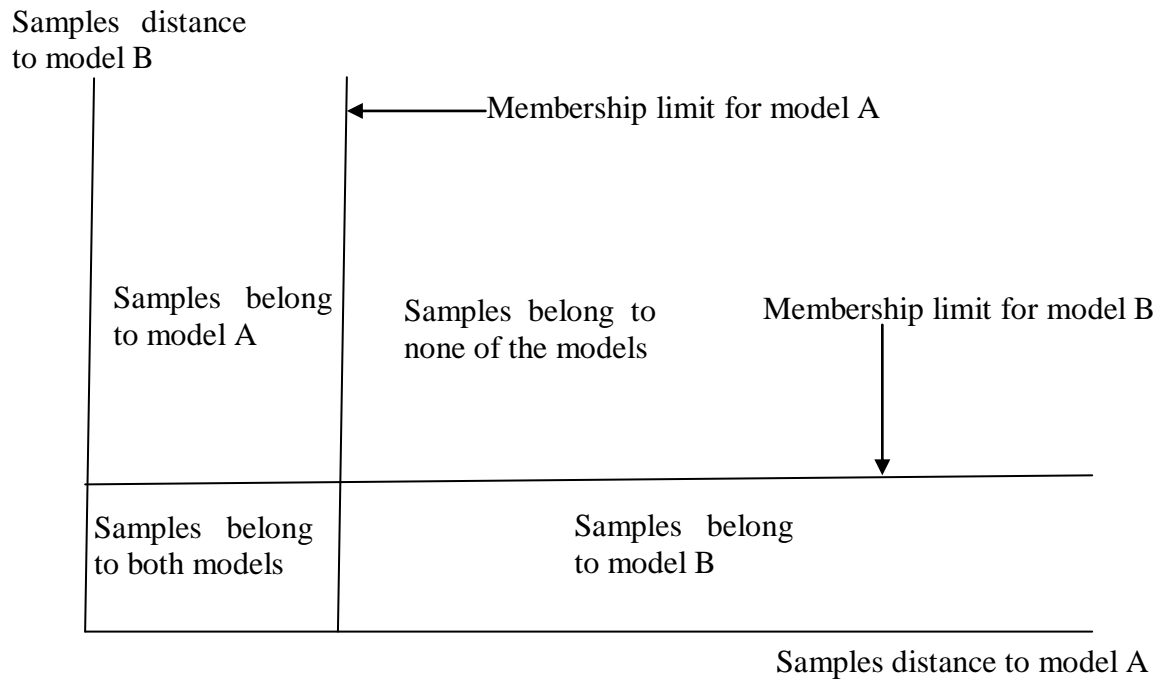
where  $X_i$  is the mean centered  $X$  matrix of class  $i$ ,  $P_i$  is the loading matrix in class  $i$ ,  $P_i^t$  the  $P_i$  transposed matrix in class  $i$ ,  $T_i$  is the score matrix in class  $i$  and  $E_i$  is the residual matrix in class  $i$  (Dragovic *et al.*, 2007; Hee-Jung *et al.*, 2009).

Before constructing PCA models for each class, the samples are divided randomly into the training set and test set. The training set performs the task of calibrating the model and the test set performs the role of unknown sample. The model size is determined by cross validation (Dragovic *et al.*, 2007). In cross validation, one sample from the training set is temporarily left out and a PCA model constructed from the remaining samples. The left out sample is classified using the generated model and its residue recorded. The left out sample is then returned to the training set, another sample is excluded, a new model is made and a new classification and residual generated. This process is repeated until every sample is left out once. After all classes are modeled, test samples are projected onto each model for classification. New samples similar to the members in the model are recognized as a member of a class; else it will be rejected.

### 3.7 Cooman's plot

The correct model in which the samples classified as false positive in the SIMCA classification table belong to can be determined by performing Cooman's plot. In the

Cooman's plot, the sample to model distance is plotted against each other for two models and it includes the class membership for both models. The display of the Cooman's plot for models A and B is presented in figure 3.4.



**Figure 3.4:** An illustration of the Cooman's plot indicating how samples are classified.

## CHAPTER 4

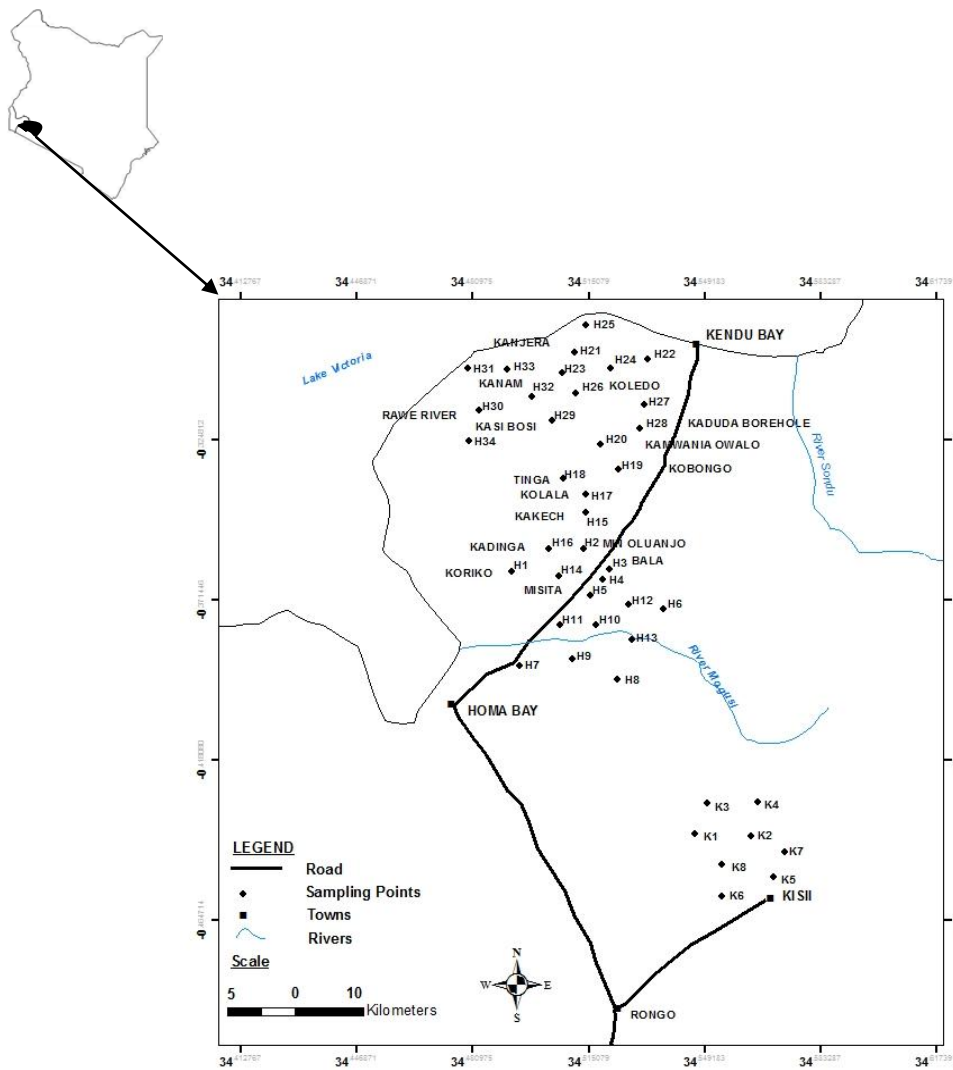
### MATERIALS AND METHODS

#### 4.1 Sample collection for radon measurement

Geothermal water samples were collected randomly around Homa mountain area in south western Kenya. Spring waters and river waters were sampled from Kanjera, Kanam and Bala which are geothermally active areas in a high background radiation area. Spring waters, river waters and pond waters were also sampled from Kamwania, Rawe river, Kaduda, Koriko, Min Oluanjo, Misita, Kakech, Kadinga, Kolala, Tinga, Kobongo, Koledo, and Kasibosi which are non geothermally active areas but in high background radiation area.

Similarly, spring waters, river waters and pond waters were also sampled from Kisii town and Itumbe; these two regions are non HBRA and non geothermally active areas but have the same geology as the above areas (Mustapha *et al.*, 1997). This was done in order to investigate the similarities and differences in radon signatures from these areas, these are control samples. The temperature and pH of the water samples from each sampling point were recorded. Temperature and pH were measured in order to investigate how they affect the activity concentration of radon. The sampling points were randomly selected but their locations were determined using the global positioning system (GPS). The sampling map is shown in figure 4.1.

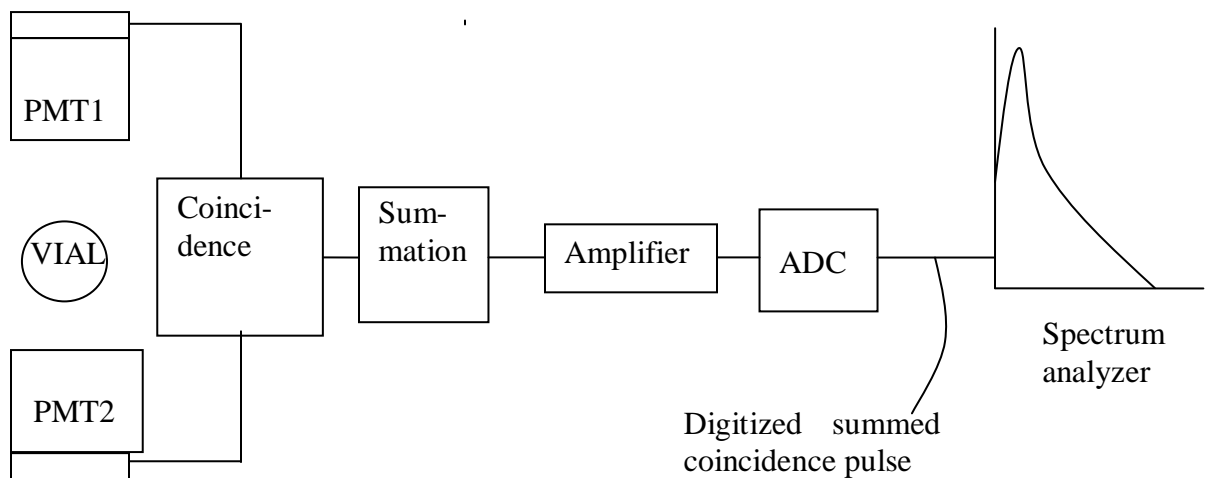
Ground waters and surface waters from each sampling point were taken directly into a standard EPA type water collection bottles with a volume of 100 ml. The bottles have rubber-teflon septa which prevents radon leaking from the bottle. The water was filled to the brim because radon is more soluble in air than in water, hence radon did not partition between air and water.



**Figure 4.1:** Sampling map and profile drew using GPS data from sampling points.

## 4.2 Liquid scintillation spectrometer

The LS counter has two photomultiplier tubes connected in a coincidence circuit as shown in figure 4.2. In order to quantify radioactivity in the sample, the sample was loaded into the counting chamber using downloading elevator mechanism. The downloading mechanism has the basic advantage of being able to prevent any external light from entering the counting chamber by using a double light seal mechanism. The double light seal is implemented by automatic loading of the sample vial from the sample chamber deck to a holding area, where the sample is sealed from external light. The sample was subsequently moved into the counting chamber, which is below the holding area (L'annunziata, 2007). Once the sample was loaded into a light tight chamber, the light was detected using two PMTs. In this work, QuantaSmart™ for the TriCarb® liquid scintillation spectrometer was used to determine the activity concentration of radon in water samples.



**Figure 4.2:** Schematic diagram of the components of contemporary liquid scintillation spectrometer.

The PMTs convert the light photons emitted from the liquid scintillation vial to electrons when the light photons hit the biakale photocathode located inside the face of the PMT. The electrons produced at the PMT photocathode are amplified through a series of positively charged dynodes, each dynode having an increasing positive voltage along the series. The increasing voltage accelerates the initial photoelectrons produced in a pulse amplification at the anode of the PMT. Light produced in the scintillation vial, is thereby converted to a corresponding electronic signal (L'annunziata, 2007).

The two PMTs permit coincidence light detection and coincidence pulse summation required for the LSC to be able to detect low energy radionuclides and to distinguish instrument background from true nuclear events. This is called coincidence counting. The principle behind the coincidence counting is that, when a nuclear decay event occurs in the scintillation vial, the light produced is isotropic. Since the decay process and resultant scintillation process produce multiphoton events, light is emitted in all directions from the scintillation vial and rapidly. If a signal is detected in both PMTs within a coincidence resolving time of 18 ns, it is accepted as a true nuclear decay event. The PMT signal that is sent to the coincidence circuit is an analog signal with pulse height that reflects, and is proportional to the original decay energy (Makinen, 1995; L'annunziata, 2007).

The summation circuit has dual purpose. The first is to reassemble the original two coincident signals into an individual signal with the summed intensity. This helps to

optimize the signal-to-noise ratio in the instrument. The second purpose is to compensate for the light intensity variations due to the position of nuclear decay in the vial that would occur when samples containing color are counted.

Subsequent to pulse summation in the LSC, the signal is further amplified and sent to the analog-to-digital converter (ADC). The ADC converts the signal from an analog signal, which is a pulse with a certain height, to a single number that represents its pulse height or intensity. The digital pulses are fully sorted on the basis of their magnitude or pulse height number by a multichannel analyzer (MCA) (L'annunziata, 2007).

The MCA has a series of bins or slots, where different pulse height magnitudes are placed once they have been detected. Two types of MCAs are commonly used: linear and logarithmic. The linear MCA provides data with pulse height calibrated to represent decay energy in keV on a linear scale. The logarithmic MCA displays the pulse height in channels plotted along a logarithmic scale.

### **4.3 Liquid scintillation counting vials**

In this study plastic vials with the following specifications were used: height with cap: 60.8 mm; diameter: 27.0 mm; diameter of opening: 17.5 mm; diameter of cap: 24.7 mm; wall thickness: 1.0 mm; normal volume: 20.0 ml; maximum volume: 24.0 ml and temperature resistance up to 80°C.

These vials were preferred because, they have: lower background level than glass vials; higher counting efficiency than glass vials; easier waste disposal and no solvent permeation with safer, high-flash point cocktails such as Ultima Gold and Optiphase Hi-safe families (Mullinger and Pates, 2007).

#### **4.4 Calibration and normalization of liquid scintillation counter**

Before the samples were counted, the system was calibrated and normalized (self normalization and calibration-SNC) for beta detection and the performance of the instrument (instrument performance assessment-IPA) assessed. During the SNC, the voltage applied to each of the two PMTs was adjusted until the tubes had been synchronized in their response to a carbon-14 standard. Calibration involved loading an unquenched  $^{14}\text{C}$  standard source into the counting chamber; the high voltage for each PMT was set to obtain the 100 %  $^{14}\text{C}$  efficiency, and the  $^{14}\text{C}$  end point energy of the standard was fixed at 156 keV.

During the IPA, the instrument measured the background, counting efficiency, sensitivity and the reproducibility of the sample result (chi-square test). The calibration, normalization and IPA procedures occurred by loading the SNC and IPA cassettes (containing the  $^{14}\text{C}$  calibration standard, unquenched  $^3\text{H}$  standard and background standard) on the counting deck and running SNC protocol flag.

#### **4.4.1 Sample preparation for radon measurement**

The Ultima Gold™ cocktail was measured and 10 ml placed in 24 ml plastic vials. A needle of a 20 ml hypodermic syringe was inserted below the water surface and several milliliters of water withdrawn and discarded. This rinse was repeated several times. 12-15 ml of water was withdrawn slowly to minimize air bubbles. The syringe was inverted to eject any air bubbles and retain 10 ml of water. The syringe needle was placed under the surface of 10 ml of Ultima Gold™ cocktail contained in a plastic scintillation vial and water ejected slowly from the syringe into the cocktail. The vial was tightly capped and vigorously shaken. During shaking almost all  $^{222}\text{Rn}$  is extracted from water to the cocktail, while other radionuclides, elements or minerals remain in the water (Salonen, 2010). The above procedure was repeated three times for each of the 42 samples. The samples were left for four hours for radon to equilibrate with its daughter nuclide before they were measured by LSC.

#### **4.4.2 Background determination**

Background count rates were determined by counting the first vial containing 10 ml of distilled water and 10 ml of the scintillation cocktail. The background pulse was stored in the computer memory and was then subtracted automatically from the sample count rates to provide a net count rate according to equation 4.1.

$$N = N_s - N_b \quad (4.1)$$

where  $N$  is the net sample count rates,  $N_s$  is the sample count rates and  $N_b$  is the background count rate.

#### 4.4.3 Minimum detectable activity of LSC method

The minimum detectable activity (MDA) of radon was calculated on basis of the standard deviation of the background rate with 99% confidence according to Currie (1968) using the equation 4.2 (Mullinger and Pates, 2007; Kiliari and Pashalidis, 2008) and it was found to be  $0.019 \text{ BqL}^{-1}$

$$\text{MDA (Bq/L)} = \frac{2.71 + 4.65\sqrt{N_b t}}{t \times E \times V \times 60} \quad (4.2)$$

where  $t$  (60 minutes) is the counting time,  $E$  (200%) is the counting efficiency (2  $\beta$ , assuming 100 % detection efficiency for each),  $V$  (0.01 litres) is the sample volume and  $N_b$  (14 counts per minute) is the background counts rates. Although,  $\beta$ -LSC and smaller sample volume was used in this work, the value of MDA obtained compare favourably with the value of  $0.016 \text{ Bq/L}$  obtained by Mullinger and Pates, 2007 for Ultima Gold scintillation cocktail who used  $\alpha/\beta$ -LSC and 580 ml of the sample.

#### 4.5 Radon measuring principle

To measure radon using LSC, the following steps were followed:

- (i) Calibration and normalization of the instrument.
- (ii) Selection of cpm assay.
- (iii) Definition of the assay parameters for radon with appropriate energy.  
Lower limit at 25 keV and upper limit at 900 keV.
- (iv) Association of the assay parameters to a radon protocol.
- (v) Loading the sample vials into the cassettes with the first vial as the background. The measurement started after four hours from the time the samples are prepared. This is because radon is in equilibrium with direct daughter nuclides after four hours storage time (Salonen, 2010; Gudjonsson and Theodorsson, 2000).
- (vi) Attaching the correct protocol flag to the first cassette to be counted and loading the cassettes with samples to the instrument.
- (vii) Starting sample counting. After all the vials were counted, cpm of the sample were printed automatically by the instrument.

Radon concentration was calculated using equation 4.3 (Pates and Mulinger, 2007;

Salonen, 2010).

$$\text{RnC} = \frac{100(N)\exp(\lambda t)}{60 \times 2 \times 0.951} \quad (4.3)$$

where RnC is  $^{222}\text{Rn}$  concentration at the time of sample collection ( $\text{Bq L}^{-1}$ ); N is the sample total count rate ( $\text{count min.}^{-1}$ ); t is the elapsed time between sample collection and counting ( $\text{min.}$ );  $\lambda$  is  $^{222}\text{Rn}$  decay constant ( $1.26 \times 10^{-4} \text{ min.}^{-1}$ ); 100 is a conversion factor of 10 ml to per liter ( $\text{L}^{-1}$ ); 60 is conversion factor from min. to sec. ( $\text{s. min.}^{-1}$ ); 2 (200 %) is the number of emissions per disintegration of  $^{222}\text{Rn}$  (2  $\beta$ , assuming 100 % detection efficiency for each); and 0.951 is the fraction of  $^{222}\text{Rn}$  in the Ultima Gold cocktail in a vial of 24 ml total capacity.

#### 4.6 Multivariate chemometric data analysis

Multivariate chemometric data analysis was performed using Unscrambler Camo 9.7 software. This software provided tools to help analyze multivariate data. This means finding variations, co-variations and other internal relationships in data matrices. The Unscrambler Camo 9.7 software can be used to do the following:

- (i) Design experiments, analyze the effects and find optima.
- (ii) Re-format and pre-process data to enhance future analyses.

- (iii) Find relevant variation in data matrices.
- (iv) Find relationship between data matrices (X and Y).
- (v) Validate multivariate models with uncertainty testing.
- (vi) Predict the unknown values of response variable.
- (vii) Classify unknown samples into various possible categories.

#### **4.7 Data pre-processing**

Chemical measurements are often inherently multivariate. Therefore multivariate chemometrics techniques were performed on radon concentrations measured in this work. Before performing these techniques on radon concentrations, pH and temperature of water samples, there was need to pre-process the data because environmental data are not normally distributed because of environmental variability and experimental/sampling errors. Data pre-processing means introducing changes in the values of the variables so as to make them better suited for analysis. This is necessary since the first step in PCA analysis is the computation of a correlation coefficient matrix which requires normal distribution of all variables. In this research three methods were used.

- (i) Taking the logarithm of all variables to improve the distribution of skewed variables.

- (ii) Smoothing the variables using the median filter in order to remove noise from the variables.
- (iii) Mean centering so that the result is interpretable in terms of variations around the mean.

#### **4.8 Principal component analysis of the activity concentration of radon, pH and temperature**

The PCA plot was performed on the activity concentration of radon, pH and temperature of spring waters, pond waters and river waters using Unscrambler Camo 9.7 software. The data was transformed into a product of two matrices according to equation 3.4, one of which contains information about the samples and the other about the variables. PCA technique was used to obtain patterns in radon levels which are characteristic of geothermal signatures.

The PCA biplot showed the variance, the scores and the loadings in the same plot. The variance revealed how much of the information in the data table was described by the model. The loadings described the data structure in terms of variable correlation such that each variable had a loading on each PC. The PCs reflected both how much the variables contributed to them and how well the variations of those variables were taken into

account over the data points.

The scores described data structure in terms of sample patterns, and more generally they showed sample differences and similarities. Each sample had a score on each PC and the score reflected the sample location along that PC. The information carried by the PC was interpreted with the help of the loadings and the score of the sample along that PC was used to characterize that sample. The scores described the major features of the sample, relative to the variables with large loadings on the same PC. Samples with close scores along the same PC were similar (they had close values for the corresponding variables). Conversely, samples with score differences much higher were quite different from each other with respect to those variables (Seddeek *et al.*, 2009).

#### **4.9 Hierarchical cluster analysis of the activity concentration of radon**

The HCA plot was performed on the activity concentration of radon in spring, pond and river waters using Matlab software. This multivariate technique examined interpoint distances between samples and radon levels and presented the information in the form of a dendrogram. The goal of HCA was to present the data in the form that facilitates the use of human pattern-recognition abilities by forming similar groups.

#### **4.10 SIMCA classification**

The SIMCA classification was performed on the activity concentration of radon data using Unscrambler Camo 9.7 software. Validated PCA models were built for each class, (waters: spring, river, pond; places: Bala, Kanam and Kanjera) which describe the structure of those classes as well as possible. The data set in each class was randomly divided into two parts, calibration samples (spring = H3, H5, H8, H12, H21, H23, H25, H29 and H33; river = H4, H6, H13, H25, H26, H30 and H34; pond = H2, H14, H15, H17, H19 and H20) and test samples (spring = H7, H10, H22, and H31; river = H9, H11, H24 and H32; pond = H1, H16 and H18). The test samples were used to perform the role of unknown samples.

The cooman's plot was performed on false positive test samples using Unscrambler Camo 9.7 software. This plot shows the orthogonal distances from the new objects to the two different models at the same time. This helps to classify the sample properly if it is classified as false positive. The class membership limit used in the interpretation of a significance level is 0.05 (or 5%). This limit helps to decide whether a sample is likely to belong to one class, or both, or none. The newly classified samples from the test samples were displayed in green color, while the calibration samples for the two models were displayed in blue and red.

#### 4.10.1 Modeling Power of the activity concentration of radon, pH and temperature

The modeling power of the activity concentration of radon in relation to pH and temperature of geothermal water samples from Bala, Kanam and Kanjera was determined. This was aimed at determining the influence of radon levels, pH and temperature on Bala, Kanam and Kanjera classification models. Modeling power was calculated using equation (4.4).

$$mp = 1 - \sqrt{\frac{\sigma_r^2}{\sigma_t^2}} \quad (4.4)$$

where  $\sigma_r^2$  and  $\sigma_t^2$  are residual and total variances respectively .

Variables with modeling power of more than 0.3 were considered to be more important for that model; variables with modeling power less than 0.3 were considered to be less important for the model and they should be left out when constructing that model (Esbensen, 2002).

## CHAPTER 5

### RESULTS AND DISCUSSION

#### 5.1 Activity concentration of radon

The activity concentrations of radon in spring waters, river waters and pond waters from geothermally active HBRA, non geothermally active HBRA and non geothermally active non HBRA are measured. The results are presented in Tables (5.1-5.8).

**Table 5.1:** The activity concentration of radon in spring waters from geothermally active HBRA (n=11)

| SAMPLE | <sup>222</sup> Rn (Bq/L) | PH  | TEMPERATURE (°C) |
|--------|--------------------------|-----|------------------|
| H3     | 31.657±1.884             | 7.6 | 57               |
| H5     | 25.532±1.887             | 7.5 | 60               |
| H7     | 29.267±0.714             | 7.7 | 81               |
| H8     | 25.583±5.718             | 7.7 | 83               |
| H10    | 20.244±0.716             | 7.7 | 86               |
| H12    | 39.657±1.239             | 7.5 | 43               |
| H21    | 17.456±1.951             | 7.6 | 38               |
| H22    | 15.339±3.835             | 7.3 | 53               |
| H23    | 15.349±3.837             | 7.4 | 55               |
| H31    | 21.294±6.690             | 8.1 | 74               |
| H33    | 14.798±1.994             | 8.3 | 88               |

**Table 5.2:** The activity concentration of radon in spring waters from non-geothermally active HBRA (n=2)

| SAMPLE | <sup>222</sup> Rn (Bq/L) | PH  | TEMPERATURE (°C) |
|--------|--------------------------|-----|------------------|
| H27    | 43.039±3.451             | 7.5 | 30               |
| H29    | 43.931±0.753             | 7.5 | 27               |

**Table 5.3:** The activity concentration of radon in spring waters from non-geothermally active non-HBRA (n=2)

| SAMPLE | <sup>222</sup> Rn (Bq/L) | PH  | TEMPERATURE (°C) |
|--------|--------------------------|-----|------------------|
| K3     | 93.240 ± 4.937           | 5.6 | 25               |
| K6     | 120.166 ± 7.433          | 5.6 | 22               |

**Table 5.4:** The activity concentration of radon in river waters from geothermally active HBRA measured (n=10)

| SAMPLE | <sup>222</sup> Rn (Bq/L) | PH  | TEMPERATURE (°C) |
|--------|--------------------------|-----|------------------|
| H4     | 9.465 ± 0.713            | 8.6 | 43               |
| H6     | 10.298 ± 0.713           | 8.8 | 46               |
| H9     | 9.493 ± 0.715            | 8.5 | 65               |
| H11    | 9.095 ± 0.716            | 8.6 | 59               |
| H13    | 8.275 ± 0.717            | 9.0 | 44               |
| H24    | 9.393 ± 0.390            | 7.6 | 47               |
| H25    | 4.281 ± 0.741            | 9.2 | 38               |
| H30    | 10.885 ± 0.754           | 9.1 | 26               |
| H32    | 8.698 ± 0.753            | 8.9 | 42               |
| H34    | 9.580 ± 0.754            | 8.9 | 39               |

**Table 5.5:** The activity concentration of radon in river waters from non geothermally active HBRA (n=2)

| SAMPLE | <sup>222</sup> Rn (Bq/L) | PH  | TEMPERATURE (°C) |
|--------|--------------------------|-----|------------------|
| H26    | 21.276 ± 0.752           | 8.3 | 32               |
| H28    | 20.967 ± 0.941           | 8.2 | 33               |

**Table 5.6:** The activity concentration of radon in river waters from non-geothermally active-non HBRA (n=2)

| SAMPLE | <sup>222</sup> Rn (Bq/L) | PH  | TEMPERATURE (°C) |
|--------|--------------------------|-----|------------------|
| K2     | 16.666 ± 0.704           | 6.7 | 25               |
| K8     | 16.565 ± 0.717           | 5.5 | 32               |

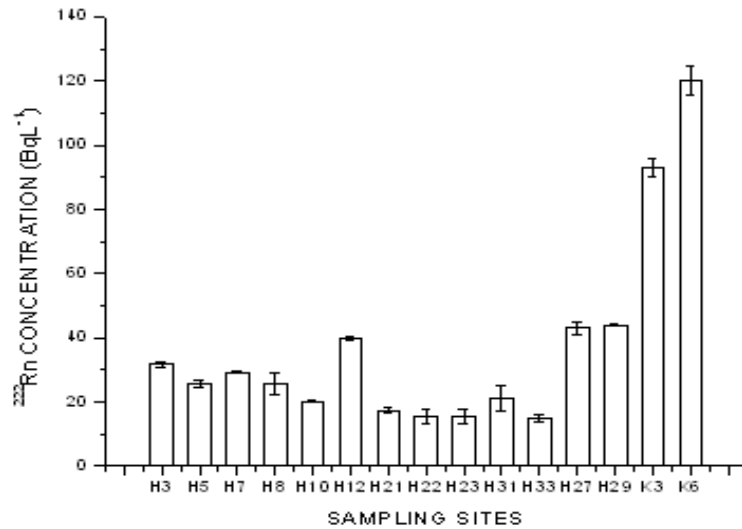
**Table 5.7:** The activity concentration of radon in pond waters from non-geothermally active HBRA (n=9)

| SAMPLE | <sup>222</sup> Rn (Bq/L) | PH  | TEMPERATURE (°C) |
|--------|--------------------------|-----|------------------|
| H1     | 13.798 ± 0.703           | 7.0 | 25               |
| H2     | 11.811 ± 0.705           | 7.1 | 27               |
| H14    | 13.336 ± 0.722           | 8.2 | 41               |
| H15    | 12.545 ± 1.255           | 8.6 | 40               |
| H16    | 12.148 ± 0.726           | 10  | 44               |
| H17    | 10.498 ± 0.727           | 8.1 | 44               |
| H18    | 11.095 ± 0.759           | 9.5 | 38               |
| H19    | 13.955 ± 1.269           | 8.6 | 38               |
| H20    | 14.371 ± 0.732           | 7.5 | 41               |

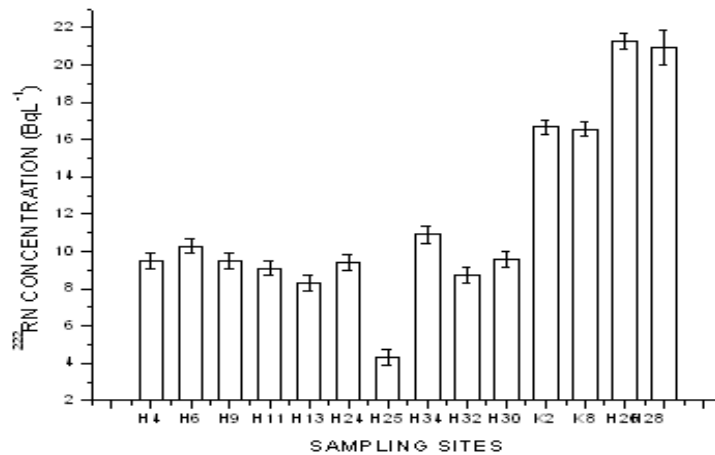
**Table 5.8:** The activity concentration of radon in pond waters from non-geothermally active non-HBRA (n=2)

| SAMPLE | <sup>222</sup> Rn (Bq/L) | PH  | TEMPERATURE (°C) |
|--------|--------------------------|-----|------------------|
| K1     | 24.909 ± 0.415           | 5.6 | 28               |
| K7     | 24.845 ± 0.242           | 5.7 | 30               |

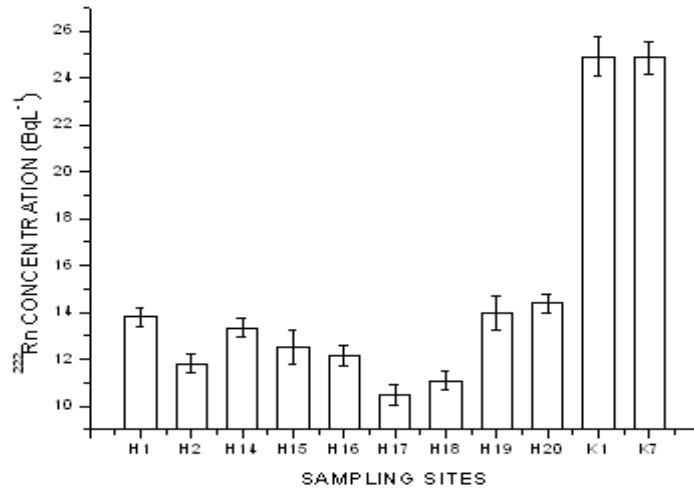
The bar graphs of the activity concentrations of radon in spring waters, river waters and pond waters from geothermally active HBRA, non geothermally active HBRA and non geothermally active non HBRA are shown in Figures (5.1-5.3).



**Figure 5.1:** The bar graphs of the activity concentration of radon in spring waters from geothermally active HBRA, non-geothermally active HBRA and non-geothermally active non-HBRA. The numbering of sampling sites is listed in Appendix II.



**Figure 5.2:** The bar graphs of the activity concentration of radon in river waters from geothermally active HBRA, non-geothermally active HBRA and non-geothermally active non-HBRA. The numbering of sampling sites is listed in Appendix II.



**Figure 5.3:** The bar graphs of the activity concentration of radon in pond waters from geothermally active HBRA, non-geothermally active HBRA and non-geothermally active non-HBRA. The numbering of sampling sites is listed in Appendix II.

It was observed that the activity concentrations of radon in water samples from geothermally active HBRA were higher in spring water samples than in river water samples. This is because the sources of these spring water samples are in contact with rock matrix containing radium which is the source of radon; hence more radon atoms will dissolve into the spring water samples. When the spring waters reach the earth's surface, they make the sources of rivers and the radon in the waters partitions between the water and air. The river waters have low concentrations of radon because radon is four times soluble in air than in water (Pates and Mullinger, 2007; Salonen, 2010).

The activity concentrations of radon in water samples from non geothermally active

HBRA were also high in spring water samples, followed by in river water samples and lowest in pond water samples. The high concentration of radon in spring waters is due to direct contact of the water samples and the rock matrix containing radium. The lower concentrations of radon in river water samples than spring water samples is as a result of the partitioning of radon between water and air when the water reaches the surface. The lower concentrations of radon in pond water samples compared to river water samples is as a result of large volume of water in the ponds which exert pressure on the basement and walls of the ponds thereby reducing the rate of radon emanation recoiling into the pond waters (Kies *et al.*, 2002).

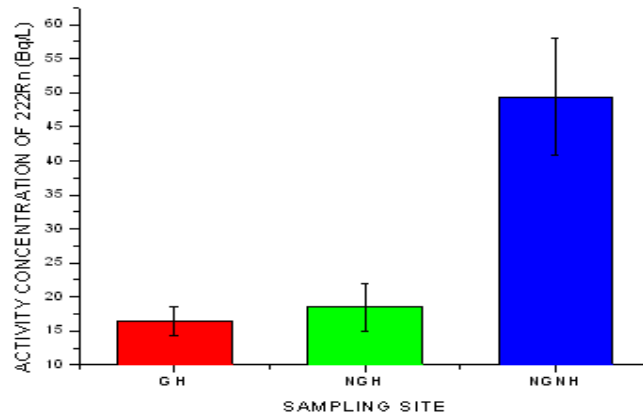
In non geothermally active non HBRA, the activity concentrations of radon in pond water samples were higher than the concentrations of radon in river water samples. This is because the volume of water in ponds from this region was low and the surface area for radon emanation was large.

The activity concentrations of radon in water samples from geothermally active HBRA, non geothermally active HBRA and non geothermally active non HBRA are computed. The results are presented in Table 5.9.

**Table 5.9:** The activity concentration of radon in water samples from geothermally active HBRA, non-geothermally active HBRA and non-geothermally active non-HBRA

| SITE                             | MINIMUM ACTIVITY CONCENTRATIONS OF $^{222}\text{Rn}$ (Bq/L) | MAXIMUM ACTIVITY CONCENTRATIONS OF $^{222}\text{Rn}$ (Bq/L) | AVERAGE ACTIVITY CONCENTRATIONS OF $^{222}\text{Rn}$ (Bq/L) |
|----------------------------------|---|---|---|
| GEOTHERMALLY ACTIVE HBRA         | $4.281 \pm 0.741$   | $39.657 \pm 1.239$  | $16.459 \pm 2.033$  |
| NON GEOTHERMALLY ACTIVE HBRA     | $10.498 \pm 0.727$  | $43.931 \pm 0.753$  | $18.484 \pm 3.462$  |
| NON GEOTHERMALLY ACTIVE NON HBRA | $16.666 \pm 0.704$  | $120.166 \pm 7.433$   | $49.399 \pm 8.513$  |

The bar graphs of the average activity concentrations of radon in water samples from geothermally active HBRA, non geothermally active HBRA and non geothermally active non HBRA are shown in figure 5.4.



**Figure 5.4:** The bar graphs of the average activity concentrations of radon in water samples from geothermally active HBRA, non-geothermally active HBRA and non-geothermally active non-HBRA.

## KEY

|      |                                  |
|------|----------------------------------|
| GH   | Geothermally active HBRA         |
| NGH  | Non Geothermally active HBRA     |
| NGNH | Non Geothermally active Non HBRA |

Although, it was expected that the levels of radon could be higher in geothermally active than non geothermally active areas because geothermally active areas have many fault lines due to geothermal activity, which provides escape routes for radon gas and non geothermally active areas are non porous providing an environment to trap radon gas, it was observed that the activity concentration of radon is high in non geothermally active non HBRA sites.

This indicates that among other factors affecting the concentration of radon in water such as distribution of radium in soil and bedrock, the porosity of soil, permeability of rocks and micro cracks, temperature and pH also play an important role in determining the concentration of radon in water samples. Temperature and pH of the water samples from non geothermally active non HBRA are lower than those of water samples from geothermally active HBRA and non geothermally active HBRA. The temperature and pH appear to have a simultaneous negative correlation with the concentration of radon in all water samples. Although the rocks in geothermally active HBRA have many fault lines which are as a result of highly altered rocks due to high temperature, the concentration of

radon is lower than the concentration of radon in non geothermally active non HBRA because radon is susceptible to loss at high temperature (Liu *et al.*, 1984).

## 5.2 Statistical distribution of radon concentration

The skewness and kurtosis of the activity concentration of radon in spring, river and pond waters were calculated and recorded in table 5.10. This was aimed at determining whether there is need to pre-process the data before performing PCA.

**Table 5.10:** The skewness and kurtosis of the activity concentration of radon in spring, river and pond waters

|          | Spring waters | River waters | Pond waters |
|----------|---------------|--------------|-------------|
| Skewness | 1.673         | 0.885        | 1.244       |
| Kurtosis | 1.626         | -0.115       | -0.188      |

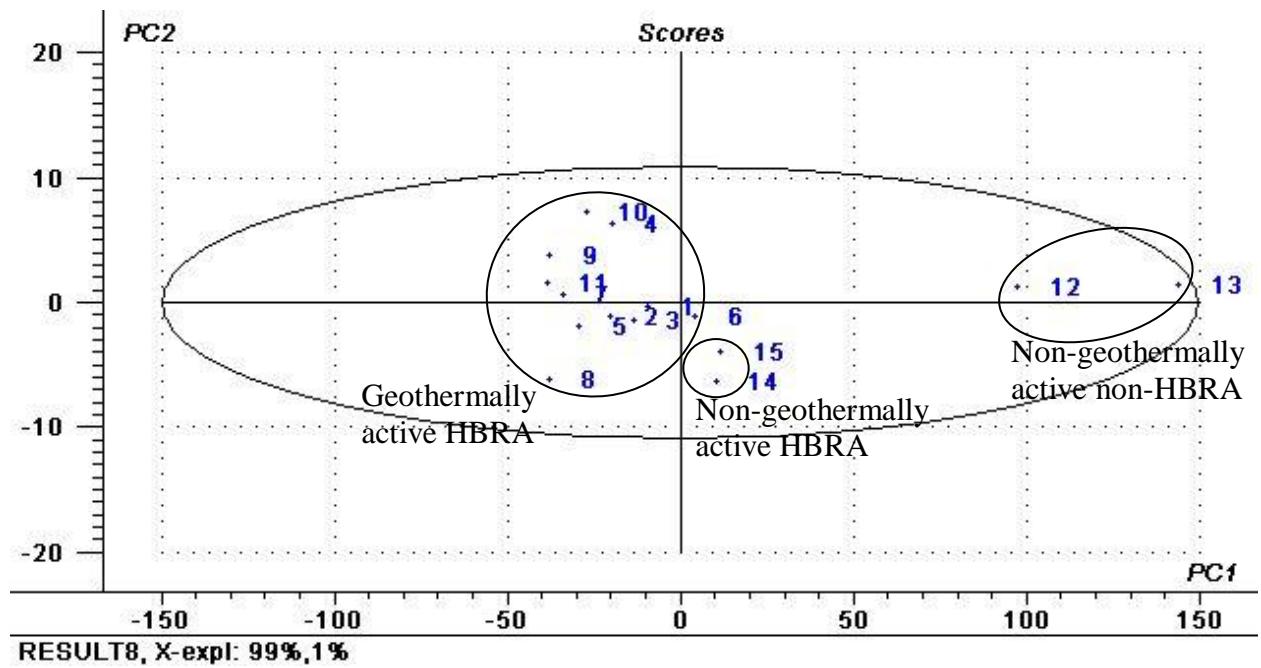
The result was compared with the values of skewness and kurtosis for normal distribution and it was found that the data was not normally distributed. Therefore, before performing PCA there is need to pre-process the data.

### 5.3 Principal components analysis of the activity concentration of radon

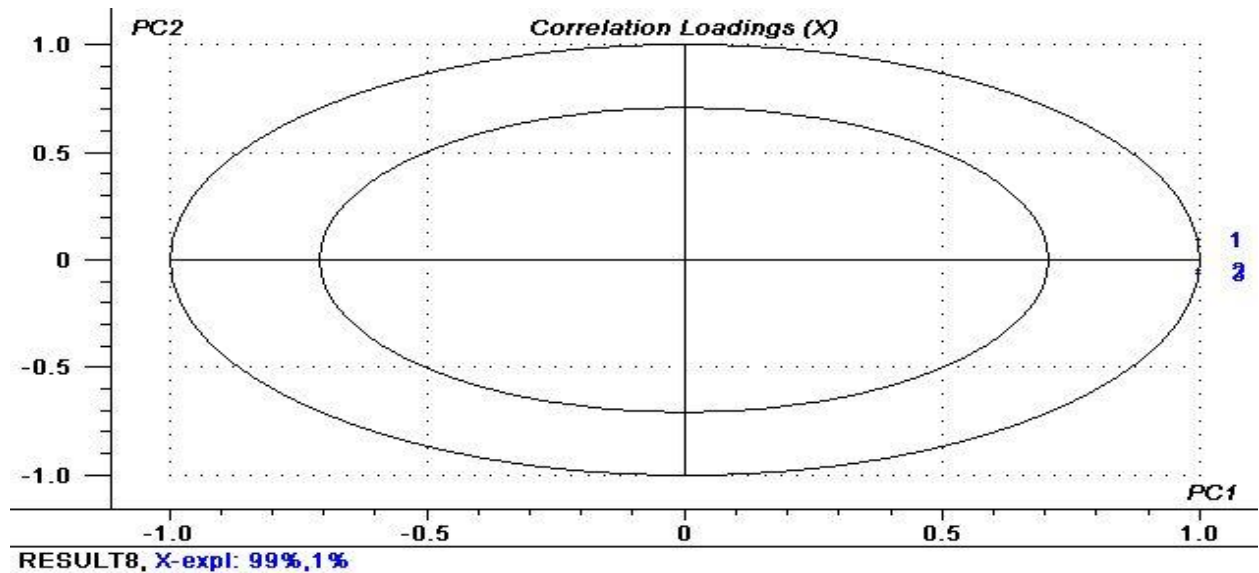
The Unscrambler Camo 9.7 software was used to perform principal components analysis on the activity concentration of radon in spring waters, river waters and pond waters in order to assess the similarity and dissimilarity of radon levels in the samples in relation to the pH and temperature at each site. Before the PCA was performed, the data was pre-processed through taking the logarithms of the variables so that the distribution of skewed variables is improved, smoothing the variables using median filter in order to remove noise from the variables and mean centering of the variables so that the result is interpretable in terms of variation around the mean. The mean centered results are shown in Appendix III. The scores and loadings plots obtained from the PCA results are presented in Figures (5.5-5.10). The numbering in the loadings plots refer to the activity concentration of radon (1), pH (2) and temperature (3).

Figure 5.5 indicates that two principal components in the score plots capture 100% of the variation contained in the original data. PC1 explains 99% and PC2 explains only 1% of the variation. It is observed that the spring water samples are classified into three distinct clusters (geothermally active HBRA, non geothermally active HBRA and non geothermally active non HBRA). From the loadings plot in Figure 5.6, it is observed that radon concentration, pH and temperature load heavily to PC1. This shows that the distinction of the three clusters is based on the levels of radon concentration, pH and temperature of the spring water samples. It should be noted that the temperature of spring

waters was highest in geothermally active HBRA followed by non-geothermally active HBRA and lowest in non-geothermally active non-HBRA water samples.

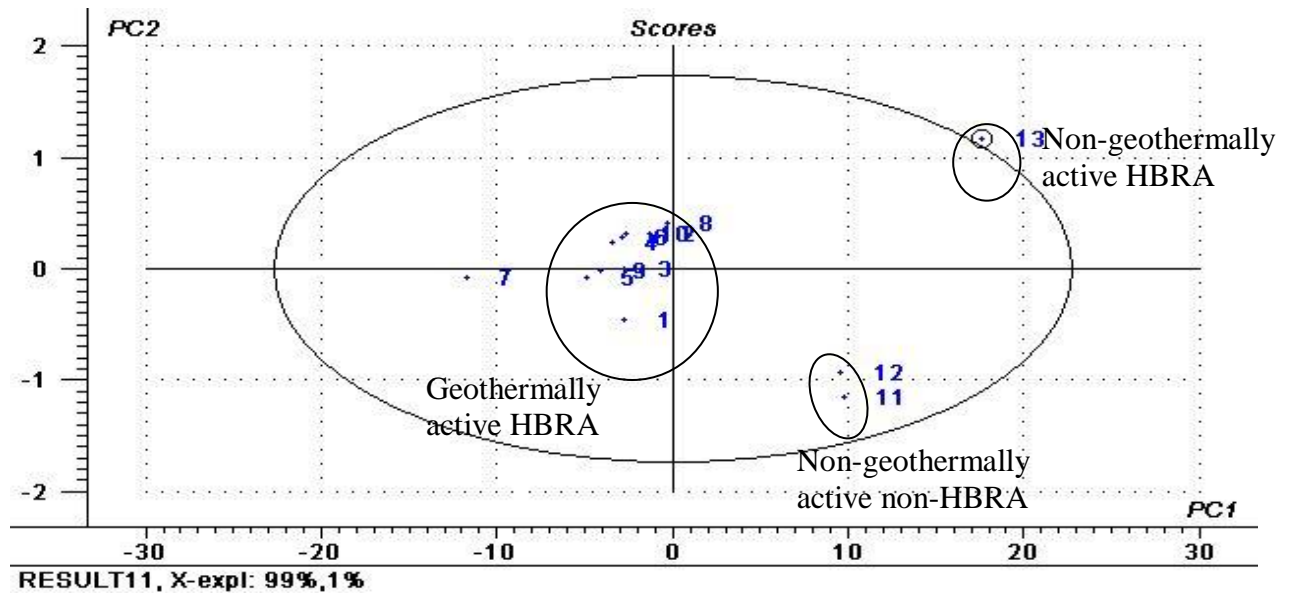


**Figure 5.5:** PCA score plots of the activity concentration of radon of spring water samples from Homa mountain area. Note that the clustering is according to the sampling sites. The numbering of sampling sites is listed in Appendix II.

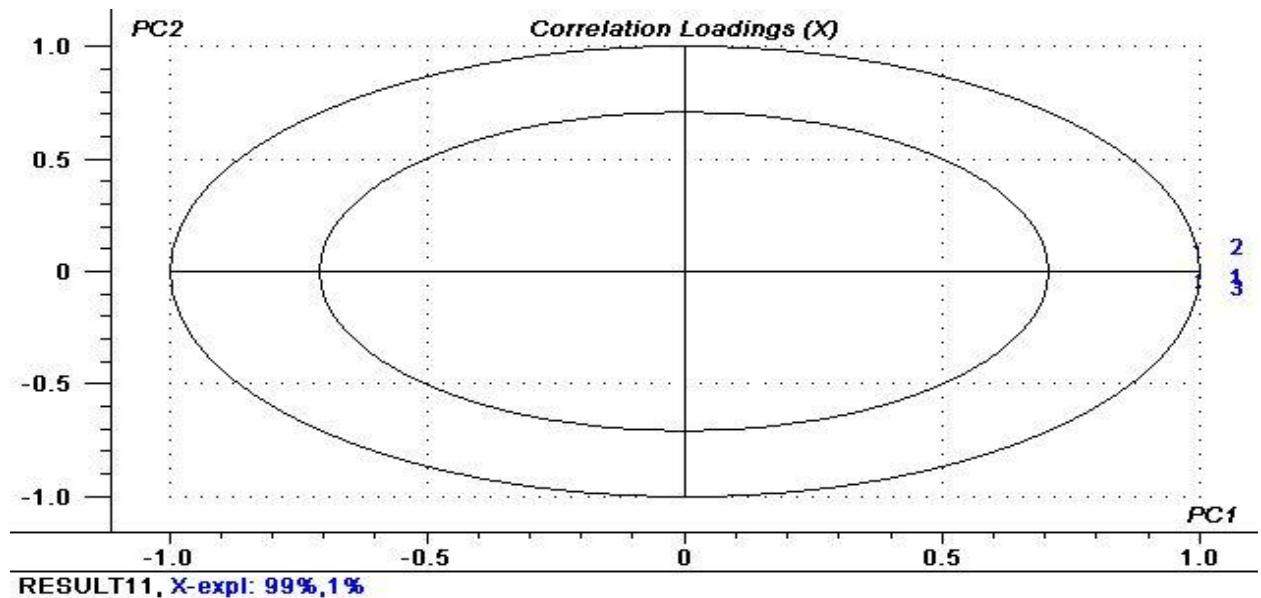


**Figure 5.6:** PCA loading plots of the activity concentration of radon of spring water samples from Homa mountain area.

Figure 5.7 indicates that two PCs in the score plots explain 100% of the variation contained in the original data. PC1 explains 99% of the variation and PC2 explains only 1% of the variations in the original data. Sample 7 is classified as an outlier. This could be as a result of contamination during sample preparation. From figure 5.8, it is revealed that the activity concentrations of radon, pH and temperature load heavily to PC1. There are also three distinct clusters (geothermally active HBRA, non-geothermally active HBRA and non-geothermally active non-HBRA) in these score plots which are based on the activity concentrations of radon, pH and temperature of the river water samples. The clusters are similar to those in figure 5.5.

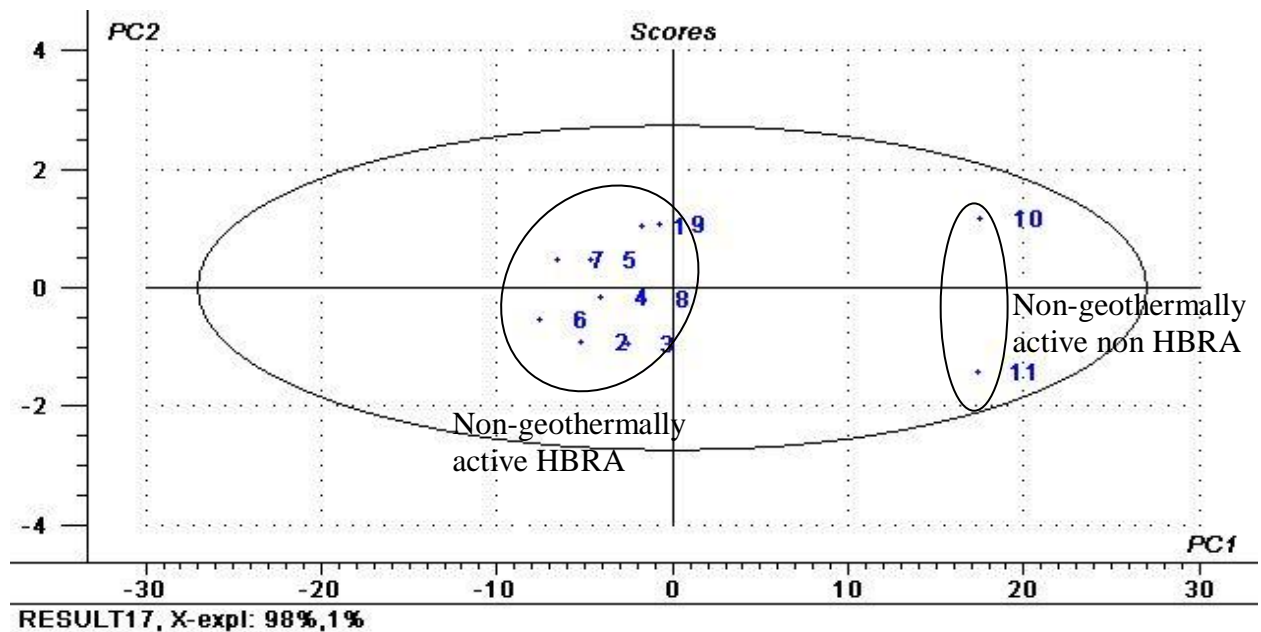


**Figure 5.7:** PCA score plots of the activity concentration of radon of river water samples from Homa mountain area. Note that the clustering is according to the sampling sites. The numbering of sampling sites is listed in Appendix II.

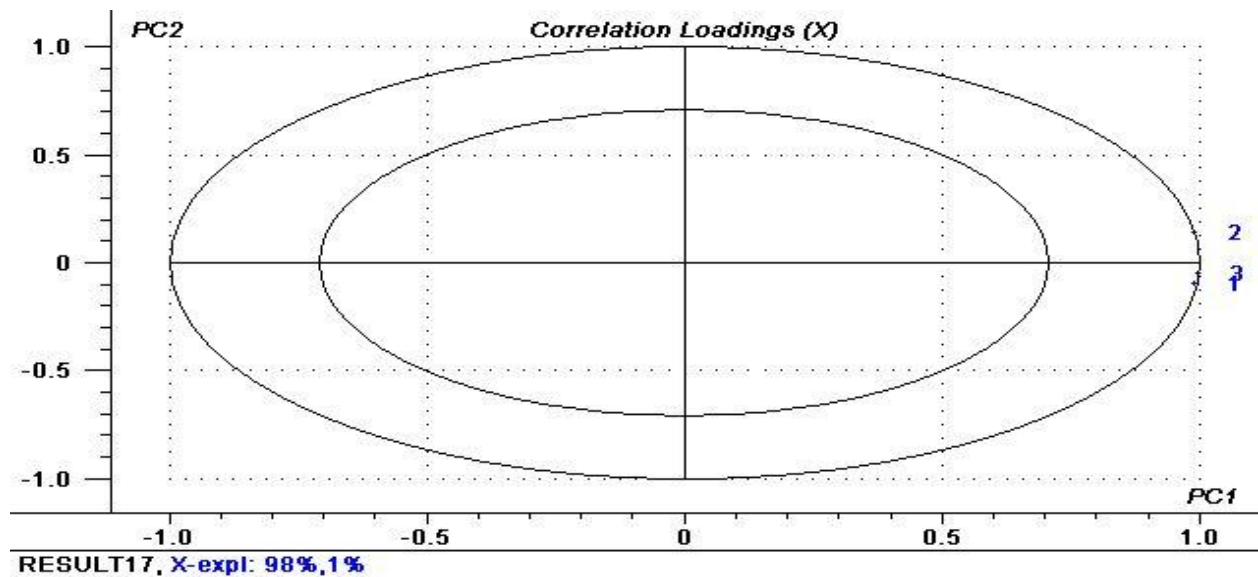


**Figure 5.8:** PCA loading plots of the activity concentration of radon of river water samples from Homa mountain area.

Figure 5.9 indicates that two PCs in the score plots explains 99% of the original data with PC1 explaining 98% and PC2 explaining only 1%. There are two distinct clusters in these score plots (non-geothermally active HBRA and non-geothermally active non-HBRA); there were no ponds in geothermally active HBRA. Figure 5.10 reveals that the activity concentrations of radon, pH and temperature load heavily to PC1. Therefore, this clear distinction of the clusters is based on the levels of radon, pH and temperature of the pond water samples.



**Figure 5.9:** PCA score plots of the activity concentration of radon of pond water samples from Homa mountain area. Note that the clustering is according to the sampling sites. The numbering of sampling sites is listed in Appendix II.



**Figure 5.10:** PCA loading plots of the activity concentration of radon of pond water samples from Homa mountain area.

It is observed that PCA is a powerful tool in revealing patterns which are not only distinct in relation to geological sites but also in relation to the state of the sites as defined by radon levels, pH and temperature of the water. The clusters which are a characteristic of the sampling site are easily distinguished. In this study the patterns which are a characteristic of geothermal activities are clearly distinguished from those ones which are as a result on non geothermal activities using only PC1. Only PC1 was considered because in the three cases (spring water, river water, and pond water samples) it explains more than 95% of the variation in the original data.

The levels of radon in geothermally active HBRA were observed to be lower than the

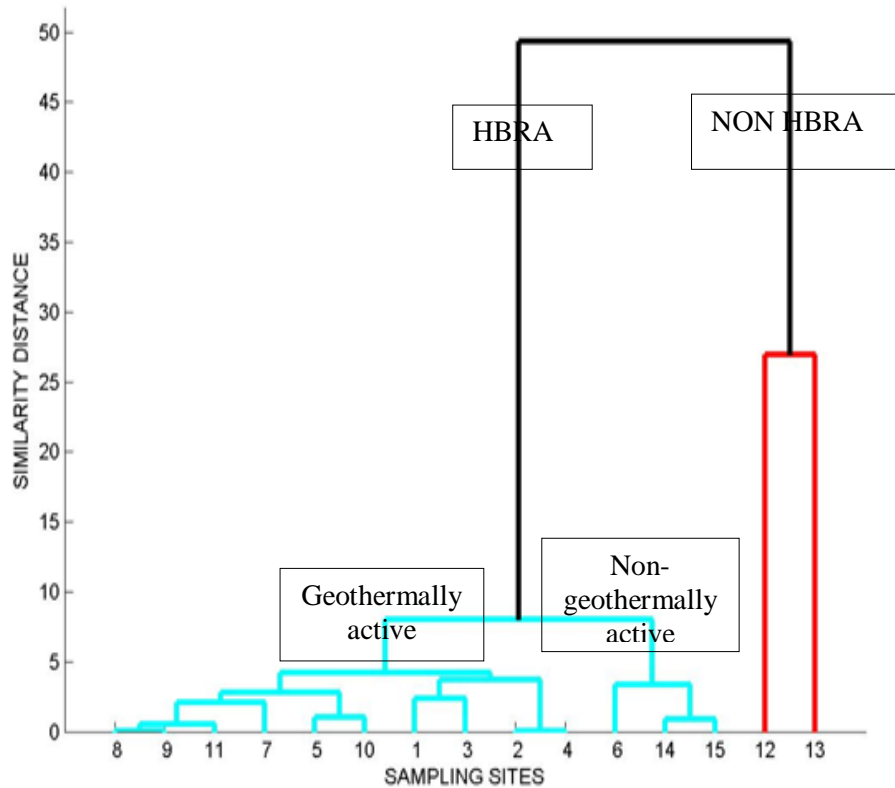
radon levels in non geothermally active non HBRA, although this does not necessarily suggest lower distribution of radium in the bedrock and lower fault lines which provide escape routes for radon gas in geothermally active HBRA. The variation in the levels of radon is as a result of the difference in temperature and pH of the two geological sites. The temperature and pH of geothermally active HBRA were higher than those of non geothermally active non HBRA. Therefore, the high temperature and alkaline pH of the waters lowers the solubility of radon in water (Liu *et al.*, 1984).

#### **5.4 Hierarchical cluster analysis of the activity concentration of radon**

Hierarchical cluster analysis was performed on only the activity concentration of radon in spring water, river water and pond water samples in order to assess whether the samples will cluster in relation to the geological sites in a similar way displayed by the PCA results. The HCA results obtained are presented in figures (5.11-5.13).

Two major clusters are observed from figure 5.11. These clusters are HBRA cluster and non HBRA cluster. These clusters are formed on the basis of the pH of spring water samples. The HBRA cluster is alkaline whereas the non HBRA cluster is acidic. The HBRA cluster has two sub clusters (geothermally active, non-geothermally active) which are formed on the basis of the temperature of the spring waters. The geothermally active sub cluster has higher temperature due to geothermal activities than the non-geothermally active sub cluster. The geothermally active sub cluster has two sub-sub clusters (8, 9, 11,

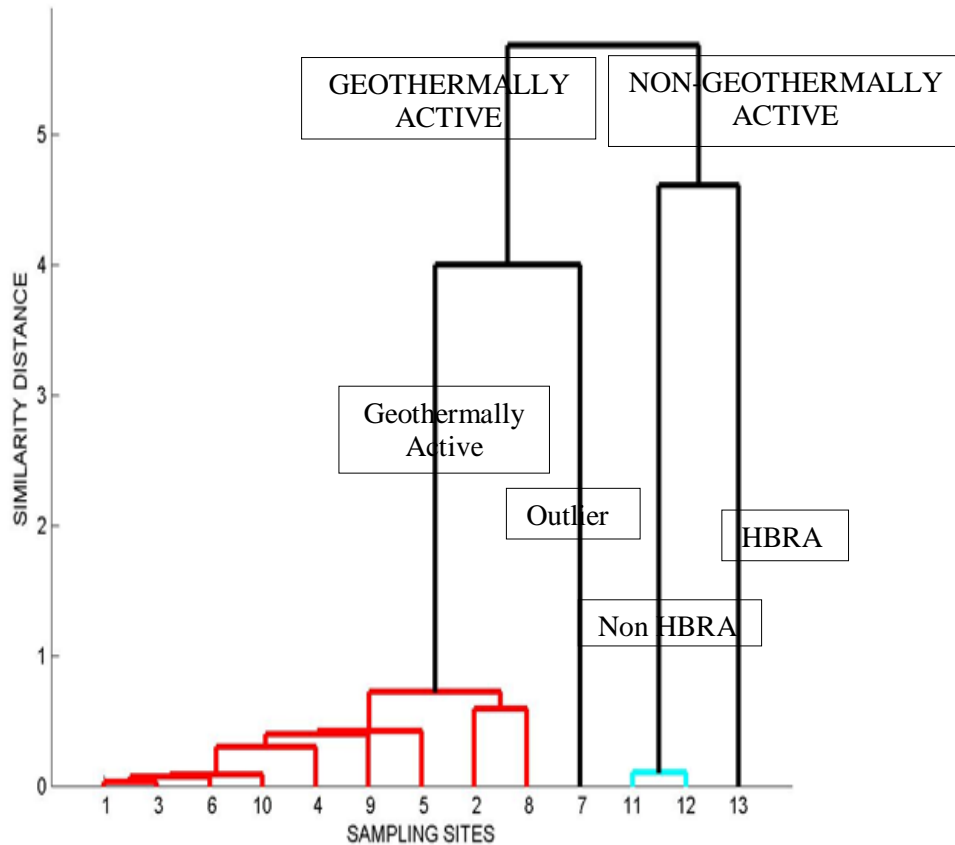
7, 5 and 10; 1, 3, 2 and 4). These sub-sub clusters are formed in relation to both temperature and pH of the spring waters. Therefore the clustering of the spring waters in relation to both temperature and pH, supports the fact that the activity concentrations of radon in spring waters is controlled by the temperature and pH of the spring water samples among other factors mentioned earlier.



**Figure 5.11:** Dendrogram of the activity concentration of radon in spring water samples from Homa mountain area. Note that the clustering is with regard to the sampling sites. The numbering of sampling sites is listed in Appendix II.

Figure 5.12 reveals two major clusters in river water samples. These clusters are geothermally active cluster and non-geothermally active cluster. The non-geothermally active cluster has two sub clusters, that is, HBRA and non HBRA. It was observed that the formation of major clusters is based on temperature variation and the formation of sub-clusters is based on pH variation. This is because the heated spring waters are the sources of rivers in geothermally active HBRA. Hence these samples had the highest temperature among the river water samples. The non-geothermally active HBRA and non-geothermally active non-HBRA had almost the same temperature, hence grouped in the same cluster even though their pH is different.

The pH of non-geothermally active HBRA was alkaline while the pH of non-geothermally active non-HBRA was acidic. Due to this variation in pH, the cluster non geothermally active is segregated into two sub clusters (HBRA and non HBRA) on the basis of the pH of the river water samples. Geothermally active cluster has an outlier (7) and two sub clusters. Sample 7 has similar temperature and pH to other samples in the same cluster but its activity concentration of radon deviates highly from other similar samples. This could be as a result of contamination during sample preparation; hence it is classified as an outlier. The sub clusters in geothermally active cluster are formed on the basis of both pH and temperature of the river waters. This reveals that the activity concentrations of radon in river waters are simultaneously affected by temperature and pH among other factors.

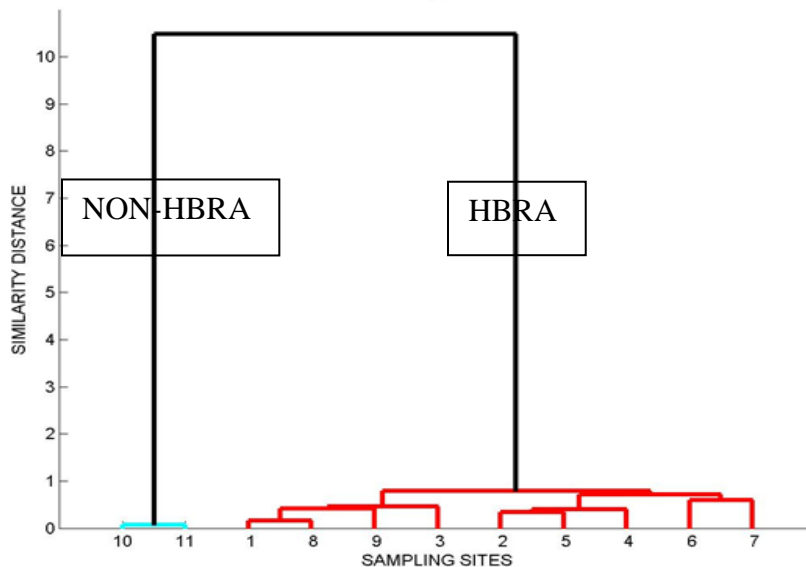


**Figure 5.12:** Dendrogram of the activity concentration of radon in river water samples from Homa mountain area. Note that the clustering is with regard to the sampling sites. The numbering of sampling sites is listed in Appendix II.

Figure 5.13 reveals two major clusters. The clusters are HBRA and non-HBRA. This is because the samples were only collected from non-geothermally active HBRA and non-geothermally non-HBRA due to absence of ponds in geothermally active HBRA. The major clusters are based on the pH of the pond water samples. The sub clusters on the HBRA cluster do not appear to be based on only the temperature and pH of the pond

water samples but also other factors could be controlling the levels of radon in these pond water samples. One of these factors could be pressure variation as a result of variation in volume of water in the ponds.

The HCA results were similar to the PCA results and it was shown that we can apportion the source of water to spring, river and pond based on the activity concentration of radon. In addition, the HCA results revealed the specific factors responsible for the clustering at each level. Therefore, the success of the HCA result is that, it proved that the activity concentration of radon could be used to classify geological sites by successfully segregating dissimilar data in terms of radon as shown in this study.



**Figure 5.13:** Dendrogram of the activity concentration of radon in pond water samples from Homa mountain area. Note that the clustering is with regard to the sampling sites. The numbering of sampling sites is listed in Appendix II.

### **5.5 SIMCA classification of geothermal water samples**

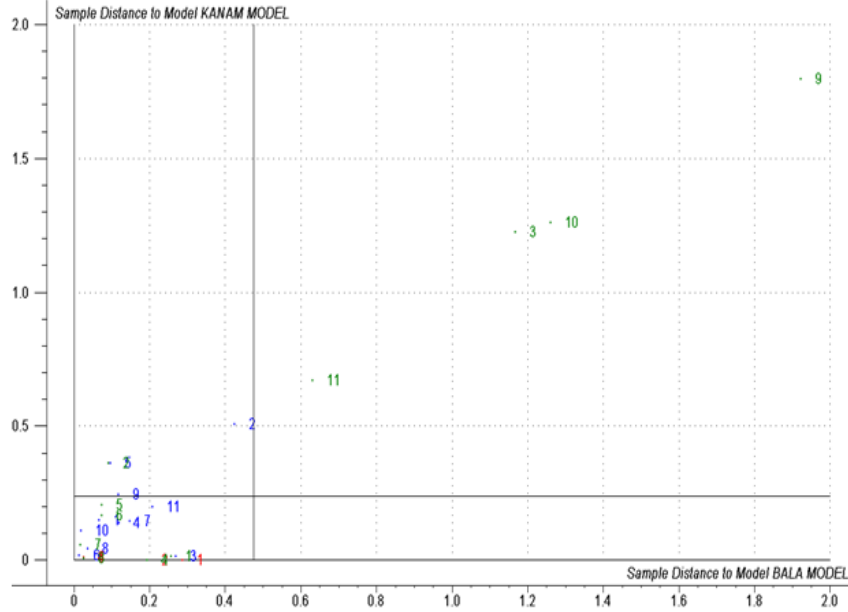
From the PCA and HCA results, it was observed that geothermal water samples exhibit distinct clusters based on the activity concentrations of radon. These clusters are a characteristic of either geothermally active HBRA, non-geothermally active HBRA or non geothermally active non HBRA. Therefore there was need to make SIMCA classification models based on the activity concentrations of radon in water samples which can be used to predict and classify samples into probable class membership. This was aimed at identifying the type of water (spring, river, pond) in a typical geothermal field based on the activity concentrations of radon.

In order to do this, Unscrambler Camo 9.7 software was used to construct PCA models for spring, river and pond water samples. The PCA models for water samples from Bala, Kanam and Kanjera were also constructed. Before these models were constructed, the water samples were divided randomly into training set and test set. The models were validated using cross validation approach. Test samples were projected onto each model. The samples were classified as members if they were similar enough in terms of the activity concentration of radon to the corresponding model. The results obtained are presented in Table 5.11.

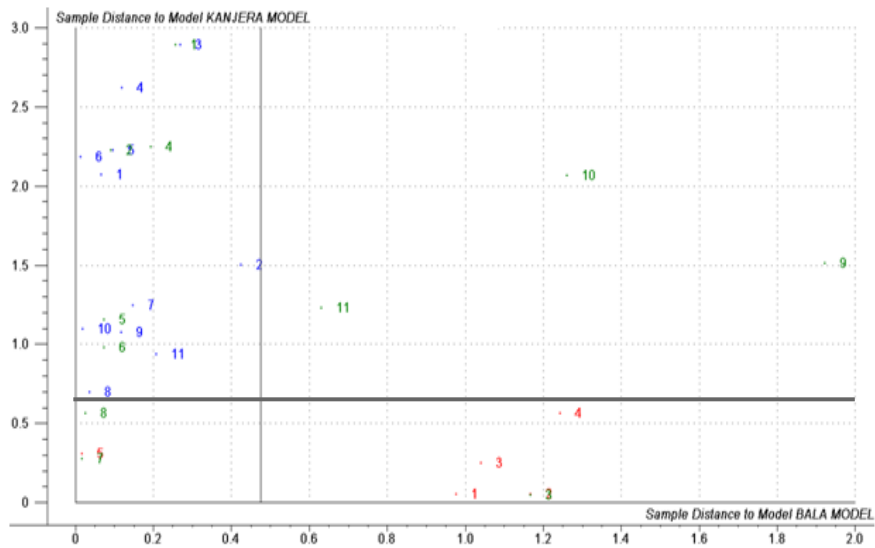
**Table 5.11:** SIMCA classification of geothermal water samples based on the activity concentration of radon. \* indicates where the sample is classified

| Sample | SPRING | RIVER | POND | BALA | KANAM | KANJERA |
|--------|--------|-------|------|------|-------|---------|
| H7     | *      |       |      | *    |       |         |
| H10    | *      |       |      | *    |       |         |
| H22    | *      |       |      |      |       | *       |
| H31    | *      |       |      | *    | *     |         |
| H9     |        | *     |      | *    | *     |         |
| H11    |        | *     |      | *    | *     | *       |
| H25    |        | *     |      | *    | *     | *       |
| H32    |        | *     |      | *    | *     | *       |
| H1     |        |       | *    |      |       |         |
| H16    |        |       | *    |      |       | *       |
| H18    |        |       | *    |      |       | *       |

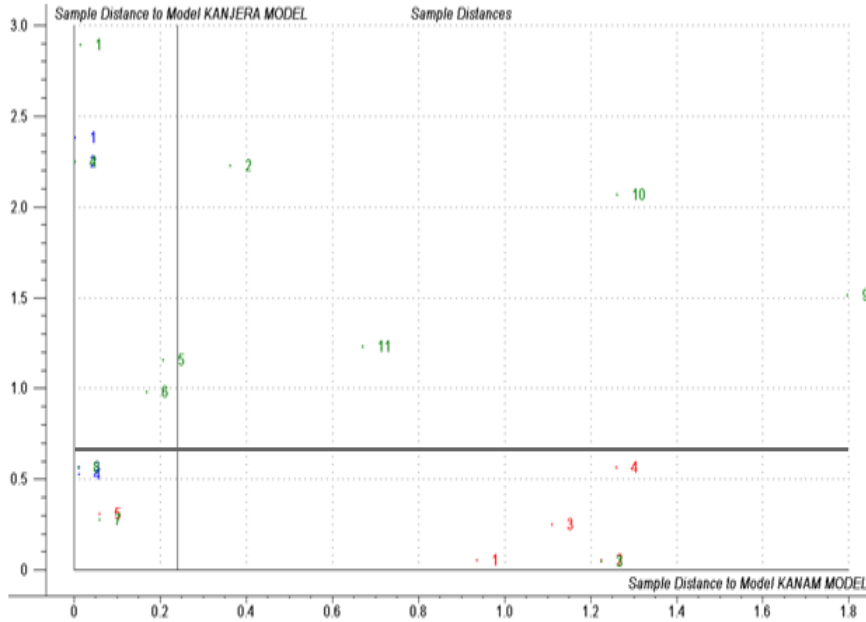
Samples H1, H7, H10 and H22 were classified as true positive meaning that they were correctly classified while samples H9, H11, H16, H18, H25, H31 and H32 were classified as false positive meaning that they were classified wrongly. In order to determine the model in which the false positive samples belong to, the Cooman's plot was performed on Bala, Kanam and Kanjera models. The Cooman's plot results obtained are presented in Figures (5.14-5.16).



**Figure 5.14:** Cooman's plot of test samples to Bala and Kanam models. Note that the blue and red numbering training samples in each model and green numbering are test samples. The numbering of samples is listed in Appendix II.



**Figure 5.15:** Cooman's plot of test samples to Bala and Kanjera models. Note that the blue and red numbering training samples in each model and green numbering are test samples. The numbering of samples is listed in Appendix II.



**Figure 5.16:** Cooman's plot of test samples to Kanam and Kanjera models. Note that the blue and red numbering training samples in each model and green numbering are test samples. The numbering of samples is listed in Appendix II.

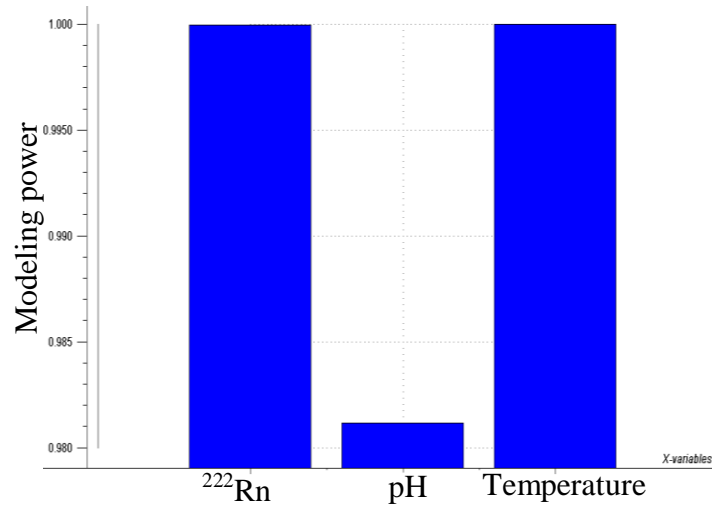
Table 5.12 shows the result obtained after the Cooman's plot.

**Table 5.12:** Classification of geothermal water samples based on the activity concentration of radon after the Cooman's plot. \* indicates where the sample is classified

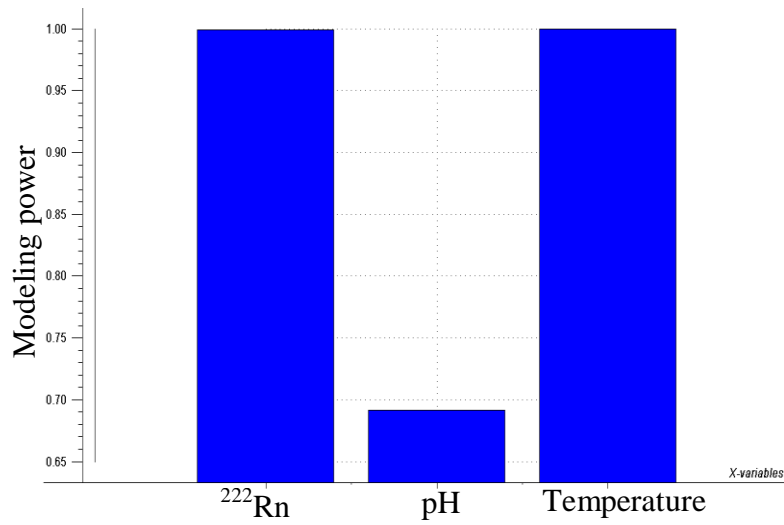
| Sample | SPRING | RIVER | POND | BALA | KANAM | KANJERA |
|--------|--------|-------|------|------|-------|---------|
| H7     | *      |       |      | *    |       |         |
| H10    | *      |       |      | *    |       |         |
| H22    | *      |       |      |      |       | *       |
| H31    | *      |       |      | *    |       |         |
| H9     |        | *     |      | *    | *     |         |
| H11    |        | *     |      | *    | *     | *       |
| H25    |        | *     |      | *    | *     | *       |
| H32    |        | *     |      | *    | *     | *       |
| H1     |        |       | *    |      |       |         |
| H16    |        |       | *    |      |       |         |
| H18    |        |       | *    |      |       |         |

It was observed from Table 5.12 that after the Cooman's plot, only samples H16, H18 and H31 are now additionally assigned to correct models. Samples H9, H11, H25 and H32 were incorrectly assigned to one or two wrong classes that were not always neighbors. This misidentification of the water samples could indicate that other variables also influence the levels of radon in geothermal waters.

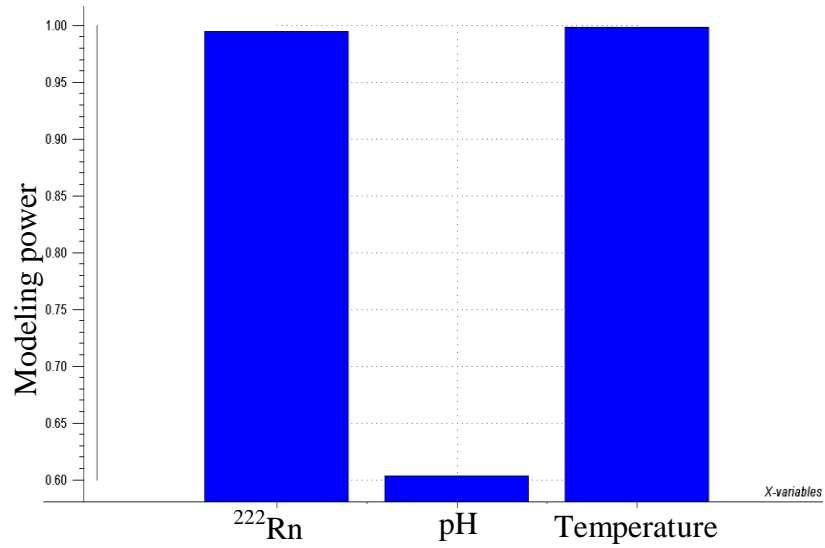
The models of the geothermally active HBRA were diagnosed using the modeling power plot of the activity concentrations of radon, pH and temperature. This was aimed at assessing how the radon levels, pH and temperature influence the classification of Bala, Kanam and Kanjera models. The results obtained are presented in figures (5.17 a-5.17 c). It was observed that the activity concentrations of radon, pH and temperature in the three models have a modeling power of more than 0.3. This means that the three variables are important in describing the models. This suggests that none of the variables should be deleted from the models. The only chance to improve on the classification of samples in order to diagnose geothermal potential in a typical HBRA based on the activity concentrations of radon is to measure some additional variables such as SiO<sub>2</sub>, B, Na, K, Ca, Mg, Fe, Al, CO<sub>2</sub>, SO<sub>2</sub>, H<sub>2</sub>S, Cl, F, Li and NH<sub>3</sub> which are chemicals and isotopic components in geothermal waters (IAEA, 2000). Then classification models are constructed based on the activity concentrations of radon and, chemicals and isotopic components in geothermal waters.



**Figure 5.17 a:** The modeling power of the activity concentration of radon, pH and temperature of geothermal water samples from Kanam.



**Figure 5.17 b:** The modeling power of the activity concentration of radon, pH and temperature of geothermal water samples from Bala.



**Figure 5.17 c:** The modeling power of the activity concentration of radon, pH and temperature of geothermal water samples from Kanjera.

## CHAPTER 6

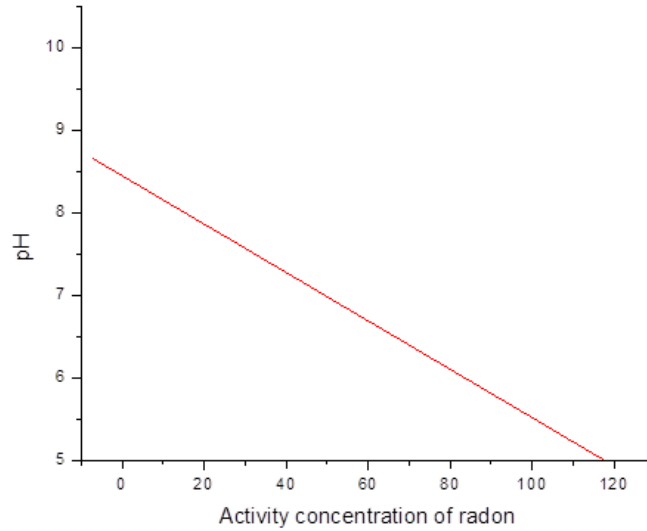
### CONCLUSIONS AND RECOMMENDATIONS

#### 6.1 Conclusions

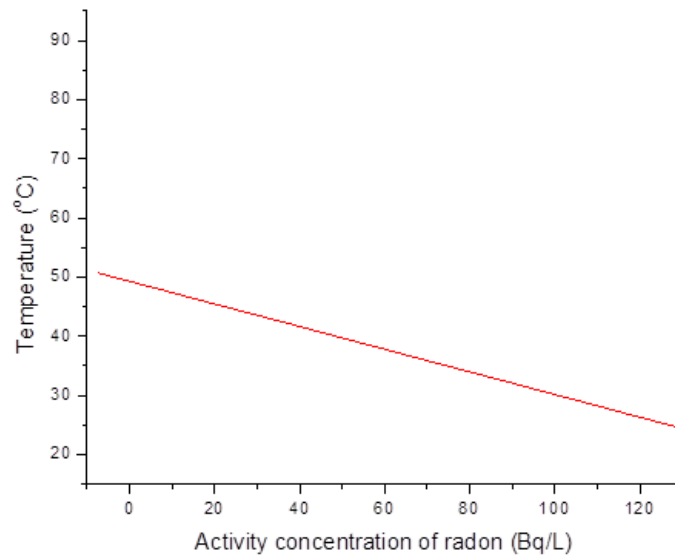
Liquid scintillation counting method was used to measure the activity concentration of radon in spring waters, river waters and pond waters from geothermally active HBRA, non-geothermally active HBRA and non-geothermally active non-HBRA. The mean values of the activity concentration of radon in water samples from geothermally active HBRA, non geothermally active HBRA and non geothermally active non HBRA are above USEPA contamination limit.

The results obtained in this study indicate that the concentrations of radon in water samples from Homa mountain, HBRA are lower than those from Kisii town and Itumbe, non HBRA. This is due to differences in pH of water samples from the two geological sites. The pH of water samples from HBRA was alkaline and that from non HBRA was acidic. This observation confirms that the solubility of radon in water is inversely proportional to the pH of the water. It was also observed that the concentrations of radon in geothermally active HBRA is lower than those in non geothermally active HBRA. This is due to differences in temperature of water samples from the two sites. The temperature of water samples from geothermally active HBRA was higher than that from non geothermally active HBRA. This observation confirms that the solubility of radon in

water is inversely proportional to the temperature of the water.



**Figure 6.1:** Graphical relation between the activity concentrations of radon with pH.



**Figure 6.2:** Graphical relation between the activity concentrations of radon with temperature.

It was observed that the concentrations of radon in spring waters are generally higher than in river and pond waters. This observation was consistent with the result obtained by Mustapha (2002) which indicated that the highest values of radon concentrations were obtained from groundwater sources. Radon is produced in rocks and soils through decay of  $^{226}\text{Ra}$  and can diffuse from mineral grains into the pore spaces and dissolve any water occupying those pore spaces making the underground water sources to have high levels of radon concentrations. The passing of air across the river and pond waters induces fast degassing of a large quantity of radon making the surface waters to have lower levels of radon concentrations than ground waters. It was further observed that the concentration of radon in geothermally active HBRA is above USEPA contamination limit (11.1 Bq/L) implying that the region has high uranium content. The high concentration of uranium may be one of the possible heat source together with the normal geothermal gradient for hot springs present in the Homa mountain area.

Principal components analysis (PCA) of spring, river and pond water samples showed distinct clusters based on the activity concentration of radon, pH and temperature which are a characteristic of the geological sites. Water samples from geothermally active HBRA, non geothermally active HBRA and non geothermally active non HBRA were segregated into distinct clusters on the basis of radon levels, pH and temperature. Hierarchical cluster analysis (HCA) results were similar to those of PCA even though these results were based on the activity concentration of radon only. The additional value

of the HCA result is that it was used as a basis for re-affirming group consistency based on the activity concentration of radon in relation to pH and temperature of the water samples. The consistency of clustering by PCA and HCA proved how successful multivariate techniques are in segregating data which are collected from different geological sites.

Soft independent modeling of class analogy (SIMCA) was used to construct classification models aimed at classifying samples according to geological sites based on the activity concentration of radon. The models predicted some samples correctly but others were classified as false positive. The prediction and classification efficiency of SIMCA was 64% indicating that the models are unreliable; hence more variables are required in the modeling in order to improve the prediction and classification ability of the models. The modeling power of the activity concentration of radon, pH and temperature of geothermally active HBRA models was more than 0.3 indicating that radon levels, pH and temperature have high influence and are relevant to Bala, Kanam and Kanjera models. Therefore, there is need to measure some additional variables such as chemicals and isotopic components in geothermal waters and use them together with the activity concentrations of radon to construct models which can diagnose geothermal potential based on radon levels.

## 6.2 Recommendations and future perspective

To improve the prediction and classification ability of SIMCA models in this work, future measurements of radon should include gas and steam condensates from fumaroles and seasonal variability should be considered as they are not considered in the present work. The modeling power of the activity concentration of radon, pH and temperature was more than 0.3 revealing that there is need to measure additional variables such as chemical and isotopic components of geothermal waters, which their levels depend on geothermal activities so that they are included in the chemometrics modeling.

Further scrutiny in areas with radon levels above USEPA contamination limit should be conducted. High concentrations of radon in ground waters indicate the presence of radon's parent nuclides,  $^{238}\text{U}$  and  $^{226}\text{Ra}$  in water-rock systems which are known health risks when ingested in drinking water. Ingestion of excessive radon levels in water is associated with lung cancer and risk of tumors of the stomach while ingestion high uranium content is identified as a nephrotoxin by world health organization (WHO) (USEPA, 1999; Skeppström and Olofsson, 2007). Therefore, the knowledge of the levels of radon and uranium content in the underground water is necessary to enable the Ministry of Health and the Radiation Protection Board to take appropriate remedial actions to protect populations from the consequences of excessive exposure to these radiations.

**REFERENCES**

- Al-Tamimi, M. and Abumurad, K., (2001). Radon anomalies along faults in north of Jordan. *Radiation Measurements* **34**: 397-400.
- Amrani, D. and Cherouati, D.E., (1999). Health effects from radon-222 in drinking water in Algiers. *Radiological Protection* **19(3)**: 275-279.
- Badhan, K., Mehra, R. and Sonkawade, R.G., (2010). Measurement of radon concentration in ground water using RAD& and assessment of average annual dose in the environs of NITJ, Punjab, India. *Indian Journal of Pure and Applied Physics* **48**: 508-511.
- Baranwal, V.C., Sharma, S.P., Senupta, D., Sandilya, M.K., Bhaumik, B.K., Guin, R. and Saha, S.K., (2006). A new background radiation area in the geothermal region of Eastern Ghats Mobile Belt (EGMB) of Orissa, India. *Radiation Measurements* **41**: 602-610.
- Baykara, O., Inceoz, M., Kulahci, F., Dogr, M. and Aksoy, E., (2008). Assessment of  $^{222}\text{Rn}$  Concentration and Terrestrial Gamma-radiation Dose Rates in the Seismically Active Areas. *Radioanalytical and Nuclear Chemistry* **278(1)**: 59-63.
- Bonifazzi, C., Didomenico, G., Lodi, E., Maino, G. and Tartari, A., (2000). Principal component analysis of large layer density in Compton scattering measurements. *Applied Radiation and Isotopes* **53**: 571-579.
- Bonotto, D. and Caprioglio, L., (2002). Radon in ground waters from Guarany aquifer, south America: environmental and exploration implication. *Applied Radiation and Isotopes* **57**: 931-940.
- Brai, M., Bellia, S., Hauser, S., Puccio, P., Rizzo, S., Basile, s. and Marale, M., (2006). Correlation of radioactivity measurements, air kerma rates and geological features of Sicily. *Radiation Measurements* **41**: 461-470.
- Chen, Y., Kuo, T., Fan, K., Liang, H., Tsai, C., Chiang, C. and Su, C., (2011). Radon measurements at IC-09 well of Chingshui geothermal field (Taiwan). *Radiation Measurements* **46**: 270-276.
- Currie, L.A., (1968). Limits for qualitative detection and quantitative determination. *Analytical Chemistry* **40(3)**: 586-592.
- Das N.K., Sen P., Bhandari R.K. and Sinha B., (2009). Nonlinear response of radon and its progeny in spring emission, *Applied Radiation and Isotopes* **67**: 313-318.

- Dragovic, S. and Onjia, A., (2007). Classification of soil samples according to geographic origin using gamma-ray spectrometry and pattern recognition methods. *Applied Radiation and Isotopes* **65**: 218-224.
- Einax, J.W., Zwanziger, H.W. and Geibs, S., (1997). Chemometrics in Environmental Analysis. Willy-VCH Verlag GmbH Ltd, German, pp162.
- Elham B., Masoud V., Asad B. and Nasrin F., (2012). Analytical study of radionuclide concentration and radon exhalation rate in market available building materials of Ramsar. *Theoretical and Applied Physics* **6**: 5-11.
- Esbensen K. H., Guyot D., Westad F. and Houmoller L. P., (2002). Multivariate data analysis-in practice 5<sup>th</sup> edition. Aalborg university, Esbjerg. 350-360.
- Ewa, I.O.B., (2004). Data evaluation of trace elements determined in Nigeria coal using cluster procedures. *Applied Radiation and Isotopes* **65**: 92-103.
- Fleischer, R.L. and Raabe, O.G. (1978). Recoiling alpha emitting nuclei. Mechanisms for uranium series disequilibrium. *Geochimica Cosmochimica Acta* **42**: 973-978.
- Fournier, R.O. (1981). Application of Water Chemistry to Geothermal Exploration and Reservoir Engineering in Geothermal Systems, *Principles and Case Histories*, John Wiley, Chichester, 109-143.
- Ganesh, P., Yogesh, P.m Gusain, G.S. and Ramola, R.C., (2008). Measurement of radon and thoron levels in soil, water and indoor atmosphere of Budhakedar in Garhwal Himalaya, India. *Radiation Measurements* **43**: 375-379.
- Ghiassi-nejad, M., Mortazavi, S.M.J., Cameron, J.R., Niroomand-rad, A. and Karam, P.A., (2002). Very high background radiation areas of Ramsar, Iran: preliminary biological studies. *Health Physics* **82(1)**: 87-93.
- Ghose, D., Debasis, P. and Sastri, R.C., (2003). Radon as a tracer for helium exploration in geothermal areas. *Radiation Measurements* **36**: 375-377.
- Giammanco, S., Imme, G., Mangano, G., Morelli, D. and Nevi, M., (2009). Comparison between different methodologies for detecting radon in soil along an active fault. The case of the Pernicana fault system, Mt. Etna (Italy). *Applied Radiation and Isotopes* **67**: 178-185.
- Giggenbach, W.F. (1995). Variation in the chemical and isotopic composition of fluid discharged over the Tampo volcanic zone. *Volcanol Geoth Res.* **68**: 89-45.
- Gordon, R.G., (2008). Practical Gamma-Ray spectrometry (2<sup>nd</sup> edition). John Willy and

sons Ltd, UK, pp 18.

Gudjonsson, G.I. and Theodorsson, P., (2000). A compact automatic low-level liquid scintillation system for radon in water measurement by pulse pair counting. *Applied Radiation and Isotopes* **53**: 377-380.

Hameed, A.K., Muhammad, T. and Aziz, A.Q., (1990). Locating geothermal energy sources. *Islamic Academy of Sciences* **3**: 3 229-231.

Hee-Jung, I., Byoung, C.S., Yong, J.P. and Kyusek, S., (2009). Classification of materials for explosives from prompt gamma spectra by using principal components analysis. *Applied Radiation and Isotopes* **67**: 1458-1462.

Hussein A.S., (2008). Radon in the environment. Proceedings of the 3<sup>rd</sup> environmental physics conference 19-23 Feb. 2008, Aswan Egypt.

IAEA, (2000). Isotopic and Chemical Techniques in Geothermal Exploration, Development and Use. (Edited, 2000) Stefan Arnorsson, IAEA, Austria.

Ioannides, K., Papachristodoulou, C., Stamoulis, K., Karamanis, D., Pavlides, S., Chatzipetros, A., and Karakala, E., (2003). Soil gas radon: a tool for exploring active fault zones. *Applied Radiation and Isotopes* **59**: 205-213.

John, (2004). Principles and applications of liquid scintillation counting. National Diagnostics, USA.

Jurgen, W.E., Heinz, W.Z. and Sabine, G., (1997). Chemometrics in Environmental Analysis. VCH 164-165.

Katsanou K., Stratikopoulous K., Zagana E. and Lambrakis N., (2010). Radon changes along main faults in the broader Aigion region, NW Peloponnese. Bulletin of the geological society of Greece, proceedings of the 12<sup>th</sup> international congress.

Kies, A., Massen, F. and Tosheva, Z., (2002). Influence of variable stress on underground radon concentrations. *Geofisica International* **41(3)**: 325-329.

Kiliari, T. and Pashalidis, I., (2008). Determination of aquatic radon by liquid scintillation counting and airborne radon monitoring system. *Radiation Measurements* **43**: 1463-1466.

Klaus, J.S., Irene, M.F., Vernon, F.H., and Kevin, H.J., (1999). Using multivariate statistical analysis of groundwater major cation and trace element concentrations to evaluate groundwater flow in a regional aquifer. *Hydrological Processes* **13**: 2655-2673.

Kolarz, P.M., Filipovic, D.M. and Marinkovic, B.P., (2009). Daily variation of indoor air-ion and radon concentrations. *Applied Radiation and Isotopes* **67**: 2062-2067.

Kowalik, C. and Einax, J.W., (2006). Modern chemometric data analysis-methods for objective evaluation of load in river system. *Acta hydrochim. hydrobiol.* **34**: 425-436.

Kramer, R., (1998). *Chemometric Techniques for Quantitative Analysis*; Marcel Dekker, INC New York, U.S.A.

Kulahci, F. and Sen, Z., (2008). Multivariate analyzes of artificial radionuclides and heavy metals contamination in deep mud of Keban dam lake, Turkey. *Applied Radiation and Isotopes* **66**: 236-246.

Lambrakis N. J. and Stamatis G. N., (2008). Contribution to the study of thermal waters in Greece: chemical patterns and origin of thermal of thermal waters in the thermal springs of Lesvos. *Hydrological processes* **22**: 171-180.

Lanzo, G., Basile, S., Brai, M. and Rizzo, S., (2010). Volcanic products of Lipari (Aeolian islands, Italy): Multivariate analysis of petrographic and radiometric data. *Radiation Measurements* **45**: 816-822.

Le Bas, M.J., (1977). *Carbonatite-Nephelinite Volcanism. An African Case History*. John Wiley Limited, UK, 347.

Lee J. and Kim G., (2006). A simple and rapid method for analyzing radon in coastal and ground waters using a radon-in-air monitor. *Environmental radioactivity* **89**: 219-228.

Liu, K.K., Yui, T.F., Yeh, Y.H., Tsai, Y.B. and Teng, T.L., (1984). Variation of radon content in groundwaters and possible correlation with seismic activities in Northern Taiwan. *Pageoph* **122**: 231-244.

L'Annunziata, M.F., (2007). *Handbook of Radioactive Analysis*, 2<sup>nd</sup> Edition. Washington, USA.

Makinen, P.O., (1995). *Handbook of liquid scintillation counting*. Turku institute of technology, department of telecommunications. Turku.

Malanca, A., Pessina, V. and Dallara, G., (1993). Assessment of natural radioactivity in the Brazillian state of Rio Grande. *Health Physics* **65(3)**: 298-302.

Mangala, J.M., (1987). A multichannel X-ray fluorescence analyses of fluorspar ore and rocks from Mrima hills, Kenya. Msc thesis. University of Nairobi.

Martinez Angel and Martinez Wendy, (2005). Exploratory Data Analysis with Matlab®. A CRC Press Company, Boca Raton London New York Washington D.C. pg 127.

Matsuda, S. and Koike, K., (2004). Three dimensional analysis of temperature distribution in a geothermal area using well – logging data set. *Geoinformatics* **15**: 15-24.

Mustapha, A. O., Narayana, D. G. S., Patel, J. P. and Otwoma, D., (1997). Natural radioactivity in some building materials in Kenya and the contribution of indoor external doses. *Radiation Protection Dosimetry* **71(1)**: 65-69.

Mustapha, A.O., Patel, J.P. and Rathore, I.V.S., (2002). Preliminary report on radon concentration in drinking water and indoor air in Kenya. *Environmental Geochemistry and Health* **24**: 387-396.

Martens, H. and Naes, T. (1996). Multivariate Calibration. John Wiley and Sons Ltd; Chichester.

Nicholson, K., (1994), Soil geochemistry in exploration for low-enthalpy geothermal resources; ammonia and boron surveys. *Geothermics in Europe*, BRGM **230**: 75-81.

Nikbakht, A.M., Hshjin, T.T., Malekfar, R. and Gobadian, B., (2011). Nondestructive Determination of Tomato Fruit Quality Parameters Using Raman Spectroscopy. *Agri. Sci.Tech* **13**: 517-526.

Ogwari P. O., Angeyo H. K., Mustapha A. O. and Mangala J. M., (2010). Trace elements profiles of geothermal field matrices associated with high background radiation area (HBRA) analyzed using chemometric assisted XRF spectroscopy. Conference of radiation protection association, Nairobi. 12-17 September, 2010 (unpublished).

Pao-Sham, W. and Chien-Li, L., (1994). Radon concentrations in spa water taken from hot and cold springs in Taiwan. *Applied Radiation and Isotopes* **46(5)**: 293-295.

Papp B., Deak F., Horvath A., Kiss A., Rajnai G. and Szabo Cs., (2008). A new method for the determination of geophysical parameters by radon concentration measurements in bore-hole. *Environmental radioactivity* **99**: 1731-1735.

Passo, C.J.Jr. and Floeckher, J.M., (1991). The Liquid scintillation approach to radon counting in air and water. *Liquid Scintillation Counting and Organic Scintillators*: 375-384.

Patel, J.P., (1991). Environmental radiation survey of the area of high natural radioactivity of Mrima hill of Kenya. *Discovery and Innovation*, **3**: 31-36.

- Pates, J.M. and Mullinger, N.J., (2007). Determination of  $^{222}\text{Rn}$  in fresh water: Development of a robust method of analysis by  $\alpha/\beta$  separation liquid scintillation spectrometry. *Journal of Applied Radiation and Isotopes* 65: 92-103.
- Planinic, J., Radolic, V., Lazanin, Z., (2002). Temporal variation of radon in soil related to earthquakes. *Applied Radiation and Isotopes*, **55**: 267-272.
- Prichard, H.M. and Marien, K., (1985). A passive diffusion  $^{222}\text{Rn}$  sampler based on carbon adsorption. *Health Physics*, **48(6)**: 797-803.
- Prichard, H.M., Gesell, T.F. and Meyer, C.P., (1980). Liquid Scintillation analyses for Radium-226 and radon-222 in portable waters. *In liquid scintillation counting in recent applications and development. Volume II sample preparation applications*, ed. Peng, C., Horrocks, D.L., and Alpen, E.L. London, Academic Press.
- Prichard, H.M. and Gesell, T.F., (1977). Rapid measurements of  $^{222}\text{Rn}$  in drinking water with a commercial liquid scintillations counter. *Health Physics* **33**: 577-581.
- Richard G. Brereton, (2007). Applied Chemometrics for Scientists. University of Bristol UK 208-211.
- Salonen, L., (2010). Comparison of two direct LS methods for measuring  $^{222}\text{Rn}$  in drinking water using  $\alpha/\beta$  liquid scintillation spectrometry. *Applied Radiation and Isotopes*, 68: 1970-1979.
- Seddeek, M.K., Kozae, A.M., Sharshar, T. and Badran, H.M., (2009). Reduction of dimensionality and comparative analysis of multivariate radiological data. *Applied Radiation and Isotopes* **67**: 1721-1728.
- Seyis C., Inan S. and Streil T., (2010). Ground and indoor radon measurements in a geothermal area. *Acta geophysica* **58(5)**: 939-946.
- Simeonova, P., Lovichnov, V., Dimitrov, D. and Kadulov, I. (2010). Environmetric approaches for lake pollution assessment. *Environ Monit Assess.* **164**: 233-248.
- Simeonov, V., Kalina, M., Tsakovski, S. and Puxbaum, X., (2003). Multivariate statistical study of simultaneously monitored cloud water, aerosol and rain water data from different elevation levels in an alpine valley (Achenkirch, Tyrol, Austria), *Talanta* **61**: 519-528.
- Skeppström, K. and Olofsson, B., (2007). Uranium and radon in ground water. *European Water* **17/18**: 51-62.
- Spaulding, J.D. and Noakes, J.E., (1993). Determination of  $^{222}\text{Rn}$  in drinking water using

an alpha/ beta liquid scintillation counter. *Journal of radiocarbons*: 373-381.

Stoker, A.K. and Kruger, P., (1979). Radon measurement in geothermal systems. Stanford Geothermal Program, Stanford univ.stanford, California.

Sulekha, N.R. and Sengupta, D., (2009). Seasonal levels of radon and thoron in dwellings along Southern Coastal Orissa, Eastern India. *Applied Radiation and Isotopes* **68**: 28-32.

Tavera, L., Balcazar, M., Camacho, M., Chavez, A., Perez, H. and Gomez, J., (1999). Radon studies for extending Los Asufres geothermal energy field in Mexico. *Radiation Measurements* **31**: 367-370.

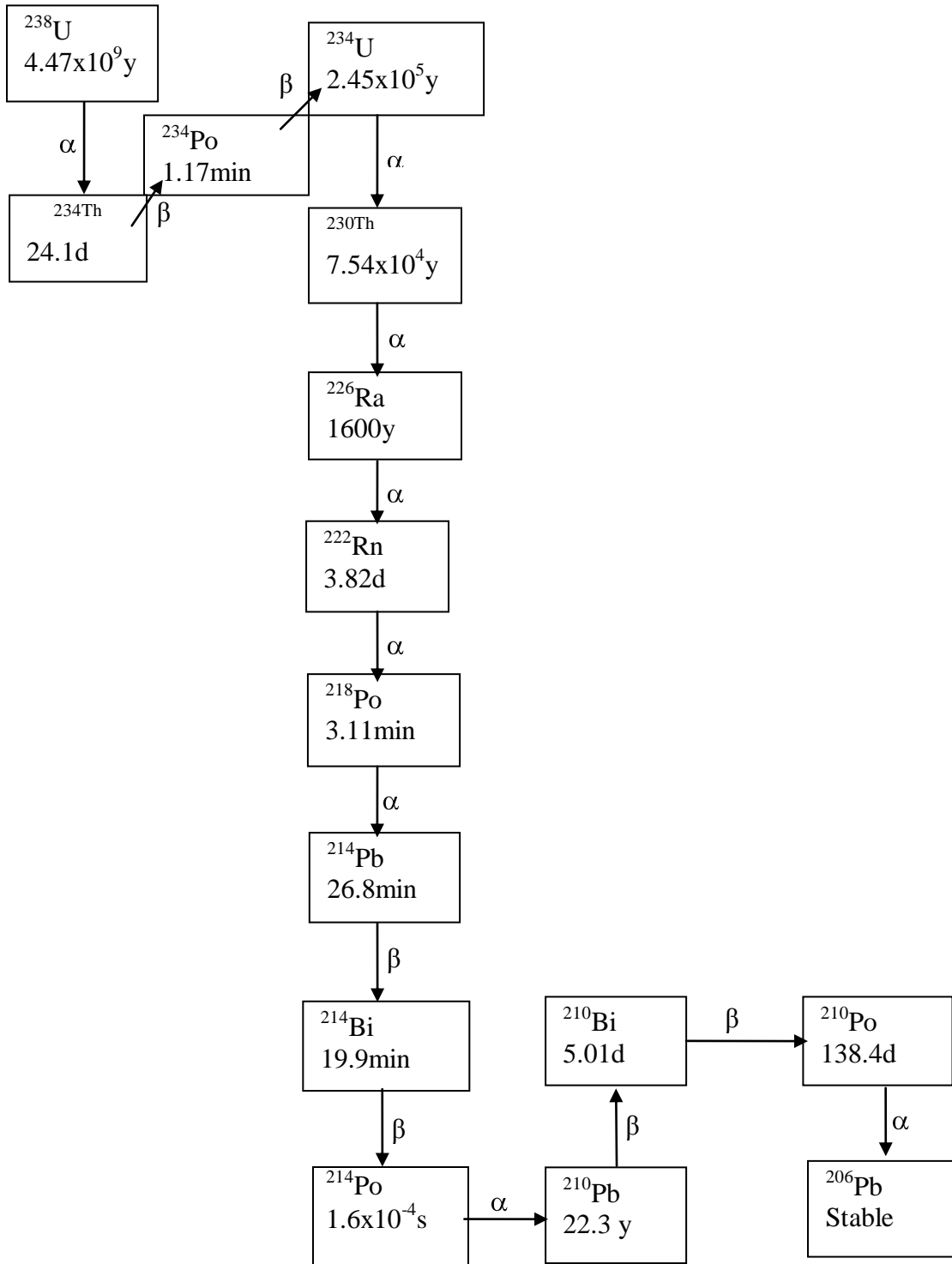
UNSCEAR, (2008). Sources and effects of ionizing radiation. Report to the general assembly with scientific Annexes Volume 1. United nations Newyork, 2010

USEPA, (1999). Radon in drinking water health risk reduction and cost analysis. *Washington, Federal Register* **64**: 9559-9599.

Van der Graaf, E.R., Koomans, R.L., Limbury, J. and De Vries, K., (2007). In situ radiometric mapping as a proxy of sediment contamination: Assessment of the underlying geochemical and geophysical principles. *Applied Radiation and Isotopes* **65**: 619-633.

Vesterbacka, P., Pettersson, H., Hanste, U.M., Jakobson, E., Kolstad, T., Roos, P. and Östergren, I., (2010). Intercomparison of radon-222 determination from groundwater. *Applied Radiation and Isotopes* **68**: 214-218.

Vinas, R., Eff-Darwich, A., Soler, V., Martin-Luis, M.C., Quesada, M.L. and Julio de Nuez, (2007). Processing of radon time series in underground environments: Implications for volcanic surveillance in the island of Tenerife, canary islands, Spain. *Radiation Measurements* **42**: 101-115.

**APPENDIX I: Uranium decay series**


**APPENDIX II: Water samples analyzed in this work**

|    |     |
|----|-----|
| 1  | H3  |
| 2  | H5  |
| 3  | H7  |
| 4  | H8  |
| 5  | H10 |
| 6  | H12 |
| 7  | H21 |
| 8  | H22 |
| 9  | H23 |
| 10 | H31 |
| 11 | H33 |
| 12 | K3  |
| 13 | K6  |
| 14 | H27 |
| 15 | H29 |

(a) Spring

|    |     |
|----|-----|
| 1  | H4  |
| 2  | H6  |
| 3  | H9  |
| 4  | H11 |
| 5  | H13 |
| 6  | H24 |
| 7  | H25 |
| 8  | H34 |
| 9  | H32 |
| 10 | H30 |
| 11 | K2  |
| 12 | K8  |
| 13 | H26 |
| 14 | H28 |

(b) River

|    |     |
|----|-----|
| 1  | H1  |
| 2  | H2  |
| 3  | H14 |
| 4  | H15 |
| 5  | H16 |
| 6  | H17 |
| 7  | H18 |
| 8  | H19 |
| 9  | H20 |
| 10 | K1  |
| 11 | K7  |

(c) Pond

|    |     |
|----|-----|
| 1  | H7  |
| 2  | H10 |
| 3  | H22 |
| 4  | H31 |
| 5  | H9  |
| 6  | H11 |
| 7  | H25 |
| 8  | H32 |
| 9  | H1  |
| 10 | H16 |
| 11 | H18 |

(d) Test samples used in SIMCA classification.

**APPENDIX III: Mean centered results of spring, river and pond water samples analyzed in this work**

(i) Spring waters

| Samples | <sup>222</sup> Rn | pH     | Temperature |
|---------|-------------------|--------|-------------|
| H3      | -5.447            | 0.227  | 2.200       |
| H5      | -11.572           | 0.127  | 5.200       |
| H7      | -7.836            | 0.327  | 26.200      |
| H8      | -11.520           | 0.327  | 28.200      |
| H10     | -16.859           | 0.327  | 31.200      |
| H12     | 2.554             | 0.127  | -11.800     |
| H21     | -19.647           | 0.227  | -16.800     |
| H22     | -21.765           | -0.073 | -1.800      |
| H23     | -21.755           | 0.027  | 0.200       |
| H31     | -15.809           | 0.727  | 19.200      |
| H33     | -22.306           | 0.927  | 33.200      |
| K3      | 56.137            | -1.773 | -29.800     |
| K6      | 83.062            | -1.773 | -32.800     |
| H27     | 5.936             | 0.127  | -24.800     |
| H29     | 6.828             | 0.127  | -27.800     |

(ii) Pond waters

| Samples | <sup>222</sup> Rn | pH     | Temperature |
|---------|-------------------|--------|-------------|
| H1      | -1.048            | -0.809 | -11.000     |
| H2      | -3.036            | -0.709 | -9.000      |
| H14     | -1.510            | 0.391  | 5.000       |
| H15     | -2.301            | 0.791  | 4.000       |
| H16     | -2.698            | 2.191  | 8.000       |
| H17     | -4.348            | 0.291  | 8.000       |
| H18     | -3.751            | 1.691  | 2.000       |
| H19     | -0.891            | 0.791  | 2.000       |
| H20     | -0.476            | -0.309 | 5.000       |
| K1      | 10.063            | -2.209 | -8.000      |
| K7      | 9.999             | -2.109 | -6.000      |

## (ii) River waters

| Samples | <sup>222</sup> Rn | pH     | Temperature |
|---------|-------------------|--------|-------------|
| H4      | -2.315            | 0.321  | 2.214       |
| H6      | -1.484            | 0.521  | 5.214       |
| H9      | -2.288            | 0.221  | 24.214      |
| H11     | -2.686            | 0.321  | 18.214      |
| H13     | -3.506            | 0.721  | 3.214       |
| H24     | -2.389            | -0.679 | 6.214       |
| H25     | -7.501            | 0.921  | -2.786      |
| H34     | -0.896            | 0.821  | -14.786     |
| H32     | -3.083            | 0.621  | 1.214       |
| H30     | -2.201            | 0.621  | -1.786      |
| K2      | 4.885             | -1.579 | -15.786     |
| K8      | 4.784             | -2.779 | -8.786      |
| H26     | 9.494             | 0.021  | -8.786      |
| H28     | 9.186             | -0.079 | -7.786      |

**APPENDIX IV: Liquid scintillation counter used to measure the activity concentration of radon in this work**



**APPENDIX V: A photograph of Homa mountain in south western Kenya where geothermal waters were sampled.**

