2009-01-01

Hybrid Approach For Site Selection Using Impact Assessment And Principal Component Analysis

Ugandhar Reddy Kondamadugula
University of Texas at El Paso, urkondamadugula@miners.utep.edu

HYBRID APPROACH FOR SITE SELECTION USING
IMPACT ASSESSMENT AND PRINCIPAL COMPONENT ANALYSIS

UGANDHAR REDDY KONDAMADUGULA

Department of Industrial Engineering

APPROVED:

Tzu-Liang Bill Tseng, Ph.D., Chair

Jianmei Zhang, Ph.D.

Tim Yao, Ph.D.

Patricia D. Witherspoon, Ph.D.
Dean of the Graduate School

This Thesis is dedicated to my Mentors, Family and Friends.

HYBRID APPROACH FOR SITE SELECTION USING IMPACT ASSESSMENT AND PRINCIPAL COMPONENT ANALYSIS

by

UGANDHAR REDDY KONDAMADUGULA, B.E.

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Department of Industrial Engineering

THE UNIVERSITY OF TEXAS AT EL PASO

December 2009

ACKNOWLEDGEMENTS

I will be profusely grateful to my parents and few people for constant encouragement and valuable suggestions. Without their support this thesis wouldn't have been a possibility. First of all my whole hearted thanks to my thesis advisor, Dr. Bill Tseng for giving me a valuable opportunity to work in his research team His faith in my abil

ity coupled with his positive outlook was the biggest source of strength at times of despair. I am also grateful to him for introducing me to new and interesting technologies in Industrials development. I am proud to be his one of thesis students at UTEP and hope I can make him feel the same. Finally, I would like to sincerely thank him for all the financial support that he offered during the course of this program.

I would like to express my deepest gratitude to my parents, who inculcated the art of learning in any domain with pure, unselfish and honest passion this fundamental quality made me grow and appreciate the world in the way I see.

I would also like to thank my Brother Mr. Srikanth Reddy Kondamadugula, friends Mr. Bharat Reddy Gurrala, Mr. Sunil Kumar Varlaparla, Mr. Prashanth Deveram, Mr. Aditya Chilukuri and other colleagues of Intelligence Systems Engineering Laboratory for creating a friendly ambience around me and providing me all the necessities to accomplish this thesis. I also want to thanks the Industrial Engineering faculty at The University of Texas at El Paso for guiding me throughout my master's program and giving me the right knowledge and experience to excel in my areas of interest.

I would also like to extend my gratitude to my thesis committee members, Dr. Jianmei Zhang, and Dr. Tim Yao for the dedication of their time and participation in the thesis.

Finally, I thank the *Almighty* showering all the blessings necessitated for the successful completion of this thesis work.

.

ABSTRACT

Principle component analysis (PCA) is used to analyze week data of emission of particulate material from the Residential community. In this Thesis we selected five sites maps where sensors are arranged to collect particulate materials PM0.25, PM 10.Impact Assessment on the site areas are done in the first case. The results Impact Assessment on the site shows that the Global warming potential is really high at the site when the analysis is run for hundred years the Global warming potential has a very high impact on Environment leading to various health hazards. Secondly Principal Component Analysis is used to find out if there is any correlation among the emission of particulate materials collected at different locations or not. PC loading Indicates there is significant correlation between the site maps while collecting the data for area 1 area4 and area3and area2 has correlation in case of PM 10 and In case PM 0.25 area3, areae4 and area1, area2 and area5 area correlated respectively. As there is correlation among the site data we can suggest them that they chose different site maps for collecting the data or we can suggest them to avoid the any one of the site which is having strong correlation with another site.

TABLE OF CONTENTS

ACKNOWLEDGEMENT…………..…………………………………………………...…..iv

ABSTRACT ……………………………………………………………………………….vi

TABLE OF CONTENTS...…………………………………………………………………vii

LIST OF TABLES………...…………………………………………………………..…x

LIST OF FIGURES……………..…………………………………………………..…..xi

CHAPTERS

1. INTRODUCTION…………………………………………………………………….1

1.1 Motivation of Research…………………………………………………………2

1.2 Thesis Overview……………………………………………………………...2

1.3 Problem Statement……………………………………………………………...3

1.4 Thesis Organization…………………………………………………………......3

2. LITERATURE REVIEW……………………………………………………………4

2.1 Pollutants Effecting Environment…………………………………..……………6

2.1.1 Ozone…………………………………………………………………………6

2.1.2 Particulate Matter……………………………………………………………6

2.1.3 Carbon Monoxide……………………………………………………...….8

2.1.4 Lead…………………………………………………………………………..8

2.1.5 Sulphur Dioxide…………………………………………………………...…8

2.1.6 Nitrogen Oxide………………………………………………………………9

2.2 Environmental Impact Assessment……………………………………………9

2.3 Global warming and Global warming Potential…….…………...……………11

2.4 Monitor Site…………………………………………………………..…….14

2.5 Principal component Analysis……………………………………………………..14

2.5.1 Geometric Approach………………….……………………………………15

2.5.1.1 Rotation of Axes……….…………………………………………..17

2.5.2 Algebraic Approach……….………………………………………………..17

2.5.3 Principal Components………………………………………………...20

2.5.4 Method1……………………………………………………………………..20

2.5.5 Method2……………….……………………………………………………20

2.5.6 Method3……………….……………………………………………………21

2.5.7 Method4……………….…………………………………………………..…22

3. METHODOLOGY……………………………………………………………..24

3.1 Approach towards the Problem..……………………………………....…..…24

3.2 Example of Principal Component Analysis…………………………..……...27

3.2.1 Interpretation From graph………………………………………………...29

3.2.2 Analysis from Principal Components……………………………………..30

3.3 Experimental Setup……………………...……………………………….…..30

4. RESULTS AND CONCLUSION……………………...………………...……….34

4.1 Impact Assessment…………………..……………………..………..……...34

4.1.1 Impact Assessment using GaBi……………………………………...…..35

4.1.2 Results from Environmental Impact Assessment…………………….…….38

4.2 Principal Component Analysis…………………………....………….....………39

4.2.1 Results from Principal Component Analysis………………………………49

4.3 Conclusions……………………………………………………………….……..50

5. REFERENCES…………………………………...………………………………51

CIRRCULUM VITAE…………………………………………………………………….....61

List of Tables

Tab 2.1: EPA Air Quality Index..……………………………………..…………..…7

Tab 2.2: Health Effects due to Sulphur Dioxide……………………...……..…………8

Table 2.3: Shows the global warming gases and their effects on environment…..….13

Tabl2 2.4 Applications of Principal Component Analysis…………………………...23

Tab3.1: Survey data is shown in the table…………………………………………….27

Tab4.1: Shows the PM 10 data ……………………………………………………..…39

Tab4.2: Shows the PM 25 data…………………………………………………..….44

List of figures

Fig 2.1: Life Cycle Assessment………………………………………………………...5

Fig 2.2: Stages of Environmental Impact Assessment……………………….………10

Fig 2.3: Shows Atmospheric concentratio0n of important long lived green gases over last 2000 years……………………………………………………………………………….11

Fig 2.4: Principal Component transformation…………………………..…………...16

Fig 2.5: Redundancy of Variables…………………………………………..……16

Fig 2.6: General Model of Principal Component Analysis…………………………..…19

Fig 2.7: Graph between Eigen value and Eigen size…………………….……..…21

Fig3.1 Shows which components are to be retained for analysis…………………………26

Fig 3.2: Screen shot of the results from Mini Tab………………...……………..………28

Fig 3.3: Screen Shot of graph between Component number and Eigenvalue……..….…29

Fig3.4: Shows the experimental setup for collecting the data………...…….……………..32

Fig 3.5: Conceptual Framework…………………………………………………………33

Fig4.1: Shows the flow of the Air Pollutants inside the Site……………………..………..35

Fig 4.2: Input and output flow of the whole site map……………………………………..36

Fig 4.3: Graph of the Green Warming Potential…………...…………………………...…37

Fig 4.4: Shows the screen shot of output and input parameters of site………………….…38

Fig4.5: shows the screen shot of results for Principal Component Analysis……………..40

Fig4.6: Shows Graph shows the plot between the component number and Eigen values…………………………………………………………………………………41

Fig4.7: Plot shows the scatter of data of different site……..................................………42

Fig4.1: Shows the screen shot of results for Principal Component Analysis…………….37

Fig 4.9: plot shows the loading of first and second components...................................…....45

Fig4.10: Shows Graph shows the plot between the component number and Eigen values…………………………………………………………………………………46

Fig 4.11: Plot shows the scatter of data of different sites……………………………..………47

Fig 4.12: plot shows the loading of first and second components..…………..…….……..48

1. INTRODUCTION

The analysis of the monitoring sites has been difficult especially when there are lots of measuring variables in a dataset. In general the monitoring sites are setup for analyzing the amount of pollutants in particular area where the analysis is intended to be done. For this kind of analysis it is difficult to analyze the whole data collected from different monitoring sites and there might be some redundancy in data which can be eliminated or ignored while doing analysis. There is much complexity associated, when they are associated with large number of variables. The classification, modeling and interpretation of monitoring data are important for quality assessment. [1] The site values in the site map may interact with each other and also it is difficult to interpret all parameter patterns in combinations (s hinab, 1993). There are different techniques which are in use for reducing the dimensions of the data set. Multivariate analysis techniques are largely used for analysis of data with large number of variables which gives easily interpretable results. The selection Multivariate analysis depends on the nature of data. If the data variables arise on equal footing, multivariate regression analysis methodology is to be considered for the analysis. When there are two variables for the analysis, correlation analysis is sufficient. [1] Data mining tools are used for classification of data, Association rule extraction and Clustering. PCA (SVD) based methods, PLS are presently being used for the dimensional reduction of data. One advantage of Principle component Analysis (PCA) [2] is that it reduces the dimensions of the data and retains the variables which accounts for maximum variation in the data. Principle component analysis is used by many environmental researchers for interpreting the quality parameters, finding the characterization of pollutants in a site not only in environmental data but also in various other cases where there are lots of variables involved. Principle component analysis is used for data reduction in cases like face recognition. [3]

PCA is also used for pattern recognition for summarizing micro array research in biomedical field. **[4]** Principle component analysis has wide range of applications from computer vision to neuroscience **[5, 6, 7, and 8].** Principle component Analysis methodology is selected in this research as PCA reduces the dimensions of the data which contains large variables and also gives variables which accounts to maximum variability of dataset and produces output without losing the critical information contained in the whole dataset (i.e. without losing the main information in the dataset).

1.1 Motivation of this Research:

This research aims at proving that Impact Assessment and PCA can be used as a tool for analyzing data collected at different locations i.e. Impact Assessment is used on the site to check the Global warming potential is high at the site when the analysis is run for hundred years, If the Global warming potential is very high impact on Environment leading to various health hazards. Then Principal Component Analysis is used to find out if there is any correlation between the data collected at different areas (site) or not. This helps the environmental investigators analyze the current locations for any correlation in the data and to select the best set of sites which gives them uncorrelated data. This also gives an opportunity to know the cause of the initial situation. (It helps researches to find out the cause for the correlation of parameters and also helps them in finding the different locations where the data is uncorrelated.)

1.2 Thesis Overview:

The overall objective of the research is reduce the number of sensors which are used for acquiring the data from different sites by using principal component analysis (PCA) and Impact Assessment suggesting that both PCA & Impact Assessment can be used as tool for analyzing the data and reducing the number of site maps where sensors are located. Further it can be an informative analysis for suggesting the site maps. So that most of the redundancy can be avoided that helps in saving costs involved in the setup of sensors. This study deals with the experimental data collected by the environmentalist at a community site using the sensors aimed at finding out the concentration of particulate material (PM10 & PM25). Interpretation of data collected was done after performing the Principal component analysis. It is suggested to reduce the number of site maps which in turn reduces the number of sensors required. Results prove that PCA and Impact Assessment can be used as tool for determining the site maps for environmental data collection.

1.3 Problem Statement: Environment Impact assessment plays a vital role in figuring whether you have to go with assessment on site or not and Environmental data collection involves large number of data collecting sites where the data is collected from the sensors situated at different sites. These sensors are used to test the concentration of pollutants there by helps in analyzing how environmental changes can affect the pollutants' concentration in the community. To do this you have to collect or chose the data wisely avoiding redundancy. If the data is redundant or correlated it is just waste of collecting the data from all sites and is also a costlier affair to set up all the sensors.

1.4 Thesis Organization:

This thesis is organized into five chapters. Chapter 1 gives the introduction and thesis overview. A formal literature review of the pollutants, Global Warming and Global warming Potential and the Principal Component analysis was presented in chapter 2. Experimental setup and methodology were explained in chapter 3. Results from the analysis and conclusions are presented in chapter 4.

2. LITERATURE REVIEW

In the past few years with constant rise in environmental issues like Global warming, Pollution to counter these engineers and Scientist are putting their effort on Environment conscious Manufacturing, Green engineering, and Green Environment there is a big challenges in sustainability of current methods and practices that we are following. We have many issues like Air pollution, water pollution, sewage waste, Energy. The companies and Industries all around the world are making constant effort to reduce the Environmental waste which affects the health and environment of the Individuals who living around the area. Environment is the main concerned while performing any activity or producing goods.

Environmental Conscious Manufacturing which is termed as (ECM) Manufacturing system has many manufacturing strategies like improving the business, profitability and to have competitive edge [9]. Apart from these Environmental issue is also an main concern for the enterprises now they are developing goods which can be recycled and can be redesigned ,which produces less waste, environmental friendly and non hazardous while operating [10].The main topics in ECM is **Life cycle assessment** which is also called as life cycle Analysis. This considers creating a product from raw material stage, Usages of the product in the real world scenario and Recycling of the product which comes after Usage stage. Fig 1 shows the life cycle assessment. The companies are using life cycle assessment to evaluate the overall impact of the product on the environment and to create greener product which is desired by the consumer. This analysis also helps the company in efficiency improvement by avoiding the waste treatment cost and for reducing green house gas emissions. [11]

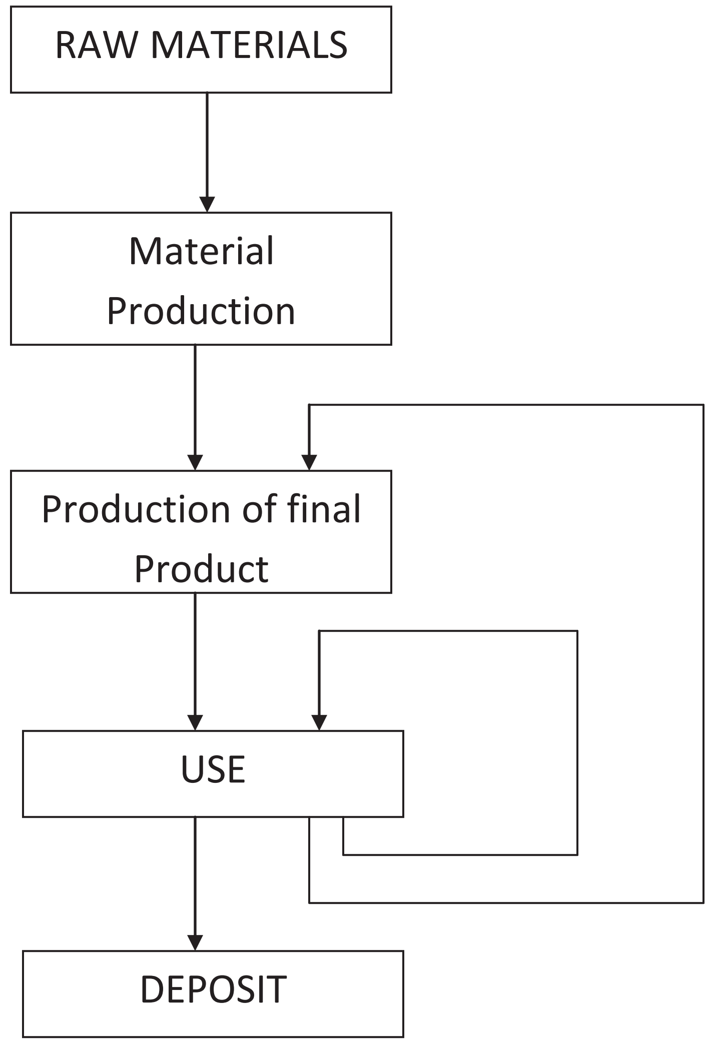


Fig 2.1: shows the flow of the Life cycle of a product

2.1: Pollutants Effecting Environment:

2.1.1 Ozone:

It is composed with three oxygen atoms .Ozone is produced when the reaction takes place between the oxides of nitrogen and volatile organic compounds (voc) in presence of sunlight. Motor vehicles, gasoline vapors and solvents of chemical are good source for the Nitrogen Oxide which in turn produces Ozone gas. Ozone gas absorbs the heat energy and doesn't allow the energy to reflect back from the earth surface which causes global warming that is why Ozone is considered as a green house gas. Ozone gas is poisonous to living beings on earth .An exposure level of 0.1 ppm Ozone will result in ageing in man if it he is exposed for long time.

2.1.2 Particulate material:

Particulate material (PM) is formed from the mixture of solid particles and liquid vapor in the air these are produced from different sources like power plants, Industrial process, and Diesel motor vehicles. It consists of coarse and fine particles. In general these particulate matters have aerodynamic diameter between 2.5µ-10µ m. Formed mainly due to suspended dust particles, dust from the road, Industry, agriculture, and construction.

PM 2.5 are called fine particulate matter are composed of various combinatio0ns of sulfate , carbon, ammonium, hydrogen ion, organic compounds, some metals and vapour.PM 2.5 comes from fossil fuels combustion, vegetation burning, and the smelting and processing of metals. The average life time is from days to week and has the travelling range from 100s to>1000s km.

PM 10 is classified as coarse particles mainly formed due to some heavy disturbances (e.g. crushing, grinding, and abrasion of surface) evaporation of sprays, and suspension of dust. It comprises of aluminum silicate and other oxides of crustal elements. [12]

These are some of the ranges of the Air quality levels which are given by EPA .The table shows the effect of the on human health and the precautions to be taken by the people when the particulate matter is present in the environment.

Table2.1: Air Quality Index [12]

Epa air quality	**Levels of health concern**	**PM 2.5**	**PM 10**
0-50	Good	None	None
51-100	Moderate	None	None
101-150	Unhealthy for sensitive groups	People with respiratory or heart disease should not stay there for long time	People with respiratory disease such as asthma, should limit outdoor activities
151-200	Unhealthy	People with respiratory or heart disease ,the elderly and children should avoid prolonged stay along with others	People with respiratory disease, such as asthma, should avoid outdoor activities
201-300	Unhealthy	People with respiratory diseases ,the eldery,and children should avoid any outdoor activity; everyone else should avoid prolong exposure	People with respiratory disease ,the eldery,and children should avoid any outdoor activity; everyone else should avoid outdoor activities
301-400	Hazardous	Everyone should avoid outdoor activities, persons having heart diseases should remain indoor	Everyone should avoid outdoor activities, persons having heart diseases should remain indoor

2.1.3 Carbon Monoxide: Carbon Monoxide is a byproduct of combustion mainly from vehicle, fuel burners, charcoal burners, Snow blowers etc. When it inhaled by the human beings it reacts with the hemoglobin of the blood to form Carboxyhemoglobin (COHb) when it form COHb the blood no longer can carry oxygen to heart and this leads to suffocation an death of the person if a person is exposed to this pollutant for longer period of time. If the COHb percentile level is more that 50% in the blood it leads to death

2.1.4 Lead: It is toxic metal generally emitted into air by vehicles and industries around the world. Sometimes it comes through water .The general health effects it cause are behavioral problems and learning disabilities, to seizures and death. Children health is mostly affected by the Lead presence.

2.1.5 Sulphur Dioxide: It is a colorless gas with pungent smell. It is mainly emitted from smelter and general utilities like electrical generation, Steel mill, Petroleum refineries etc.

Table2.2: Shows the Health Effects of Different Levels Caused By So2 [13]

Category		Aqi	So2
	Very good	0-15	No health effects are expected in healthy people
	Good	16-31	Damage some vegetation in combination with ozone
	Moderate	32-49	Damage some vegetation
	Poor	50-99	Odor, increasing vegetation damage
	Very poor	100 over	Increasing sensitivity for asthmatics and people with bronchitis

2.1.6 Nitrogen Oxide*:* It is formed when fuel burns at high temperatures mainly from the motor engines, electrical utilities, and other industries these are the prime sources of the nitrogen oxide in the world. Effects of NOx are as follows It causes acid rain when it gets reacted with sulfur dioxide which causes fog, snow and rain effecting the forests, historical monuments etc. NOx when reacted with ammonia, moisture and other compounds will affect the human health by damaging the lung tissue. Global warming Nitrogen oxide is considered as green gas accumulates in the atmosphere and there by increases the temperature of the earth which will lead to human health risk. Nitrate particles can obstruct the light which reduces the visibility. [13]

2.2 Environmental Impact Assessment:

Environmental Impact Assessment is a planning tool that its main purpose is to give the environment its due place in the decision making process by clearly evaluating the environmental consequences of a proposed activity before action is taken. The concept has ramifications in the long run for almost all development activity because sustainable development depends on protecting the natural resources which is the foundation for further development"[14] .The key elements of an EIA are (a) Scoping: identify key issues and concerns of interested parties; (b) Screening: decide whether an EIA is required based on information collected; (c) Identifying and evaluating alternatives: list alternative sites and techniques and the impacts of each; (d) Mitigating measures dealing with uncertainty: review proposed action to prevent or minimize the potential adverse effects of the project; and (e) Issuing environment statements.

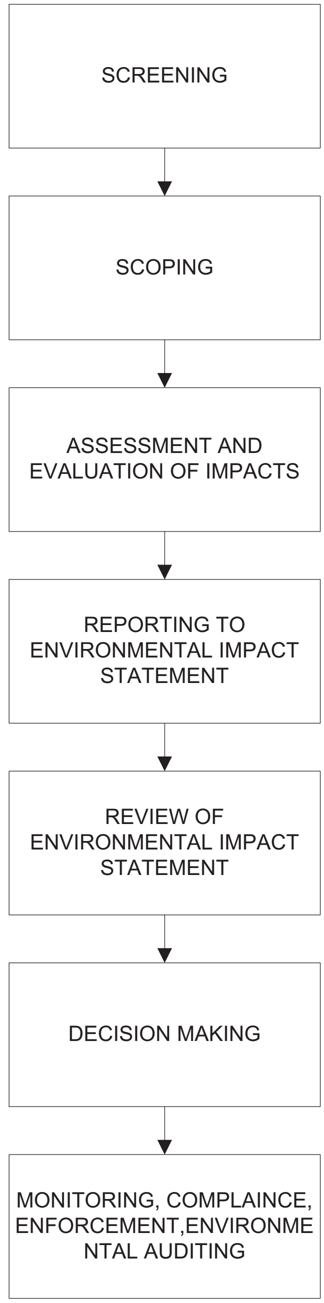


Fig 2.2: Stages of Environmental Impact Assessment [41]

2.3 Global warming and Global Warming Potentials:

Global warming' is a phrase that refers to the effect on the climate of human activities, in particular the burning of fossil fuels (coal, oil and gas) and large-scale deforestation, which cause emissions to the atmosphere of large amounts of 'greenhouse gases', of which the most important is carbon dioxide. [42] Natural temperature control system that enables the Earth to sustain average surface temperatures in the region of 15C. This is what sustains life – makes the Earth inhabitable. Main cause for Global warming is due to emission of greenhouse gasses and Energy consumption is believed to be the main responsibility for Green House Gas emission. It is believed that these GHG gas can be reduced by reducing the utilization energy in different ways [38].

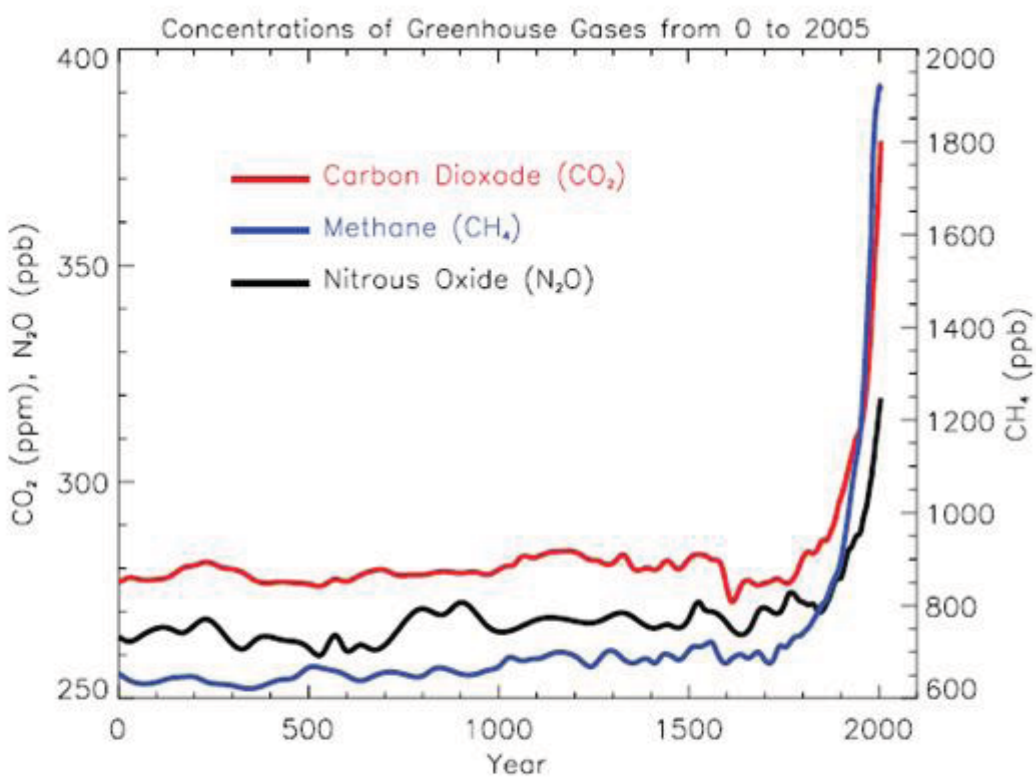


Fig 2.3 shows Atmospheric concentrations of important long-lived green gases over the last 2000 years. Increasing since about 1750 are attributes to human activities in industrial era. [41]

GWPs are indexes that are intended to compare the relative effects on climate of equal units of emissions of different gases. In practice they are calculated as the ratio of the radiative forcing resulting from an additional unit of emissions of a gas, integrated over a particular time horizon, relative to the same quantity calculated for an emission of a unit of CO_2.[41]

Global Warming Potentials (GWPs) are intended as a quantified measure of the globally averaged relative radioactive forcing impacts of a particular greenhouse gas. It is defined as the cumulative radioactive forcing both direct and indirect effects integrated over a period of time from the emission of a unit mass of gas relative to some reference gas (IPCC 1996). Carbon dioxide (CO_2) was chosen as this reference gas. Direct effects occur when the gas itself is a greenhouse gas. Indirect radioactive forcing occurs when chemical transformations involving the original gas produce a gas or gases that are greenhouse gases, or when a gas influences other radioactively important processes such as the atmospheric lifetimes of other gases. The relationship between giga grams (Gg) of a gas and

Tg CO_2 Eq. can be expressed as follows:

$$\text{Tg } CO_2 \text{ Eq} = (\text{Gg of gas}) \times (\text{GWP}) \times 1{,}000\text{Gg}$$

Where:

Tg CO_2 Eq. = Tera grams of Carbon Dioxide Equivalents

Gg = Giga grams (equivalent to a thousand metric tons)

GWP = Global Warming Potential

Tg = Tera grams

GWP values allow policy makers to compare the impacts of emissions and reductions of different gases.

Table 2.3: Shows the Global warming gases and their effects on environment

Green house gas	Primary Sources	Present concentration in atmosphere	% Annual Increase	Effective residence time in atmosphere	Sinks and reservoir
Carbon Dioxide	Production of commericial energy deforestation other biomass burning	360	0.40%	50-200 years	Atmospheric reservoir , ocean up take by N.Hemishere forest growth. Transfer to soils and to the deep ocean.
Methane	Natural gas production transmission; enteric fermentation ; rice cultivation landfill emissions, deforestation	1.7	0.50%	12.5 years	Main removal process: troposphere hydroxyl radical stratosphere; soils
Halocarbons Most abundant are CFC-11 and CFC-12	Solely of human origin: used in industrial processes and end-use products like air-conditioners and refrigerators(as coolants and insulation	CFC-11=27 CFC-12=500	Falling due to ban on use.(HCFCs and HFCs) are Showing increasing	Range from a few years to few thousand years	Atmospheric reservoir; removed mainly through breaking down by sunlight in the stratosphere
Nitrous Oxide	Mainly from use of fertilizer and fossil fuel combustion	315	0.25	120 years	Removedmainly through breaking by sunlight in the stratosphere

2.4 Monitoring Sites:

In general we collect lots of data from the different site maps to get the clear cut information of the environment site maps or for analysis. There are different cases in which we require data to determine the pollution or pollutants affecting the community living in that place or to find the levels of traffic in different places. But the question is selecting the site properly and aptly? To find out this PCA and Impact Assessment is being used in my thesis.

2.5 Principal Component Analysis:

Principal component analysis which is considered as the data reduction technique especially for reducing the environmental when data sets and the variables are large. (Henry and Hidy (1979)) recognized the advantage of PCA for data reduction and interpretation. (1979) [14] Now it is being used as tool for classification and reduction of the data in exploratory analysis when it is comprises of large data sets.PCA can be calculated using a correlation matrix, covariance matrix, Singular value decomposition and Eigen value decomposition for the analysis of a data. [15]. PCA is used to find out the underlying structure of the data being collected and gives the deep insights into the data. In general the results of PCA are represented in scores and loadings. In every application a decision has to be made on how many principal components are to be retained or considered in order to summarize the data. The goal of PCA is to reduce the variables, if ‘n’ is the number of data variables then by using the principle component analysis we get ‘x’ variables which is less than the ‘n’ without losing the total originality of the data. First principal component gives the linear combination with maximum variance. Second component is linear combination with maximal variance in orthogonal to first principal component analysis. [4] Singular value decomposition, Principal component analysis are two techniques which are used commonly in the multivariate analysis.[16] Apart from them

MDS (Multi dimensional Scaling), ICA (Independent component analysis are also used for analysis of data.

There are two approaches towards any problem using principal component analysis

1. Geometric Approach

2. Algebraic Approach

2.5.1 Geometric Approach:

Principal component analysis deals with a single sample of n observation vectors $y_{1}, y_{2}, \ldots \ldots y_{n}$ that forms a points in p- dimensional space. PCA can be applied to any distribution of y, easier to visualize when it is ellipsoidal. When variables are correlated, the ellipsoidal swarm of points is not oriented parallel to any of axes represented by all variables. [17]

This is done by translating the origin to one variable and then rotating the axes. After rotation so that axes become the natural axes of the ellipsoid the new variables will be uncorrelated.

The axes can be rotated by multiplying each variable by an orthogonal matrix A. Where the orthogonal was denoted by

$$Z_i = Ay_i$$

$$\text{Since A is orthogonal } A^{/} A = I$$

$$\mathbf{z_i{}^{/} z_i = (A y_i)^{/} (A y_i) = Y_i^{/} A^{/} A\, Y_i = Y_i\, Y^{/}_i}$$

Thus an orthogonal matrix transforms that is same as distance from the origin, and axes are effectively rotated.

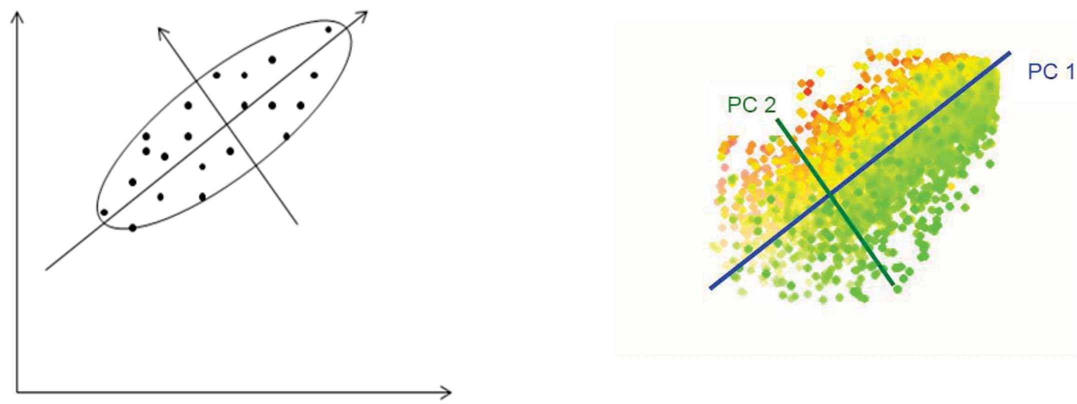


Fig 2.4 Principal component transformation

High dimensional data set is reduced to two dimensional graphs where the maximum variability of the data is caused by only few variables which are represented in two dimensional principle components which is show in the figure.

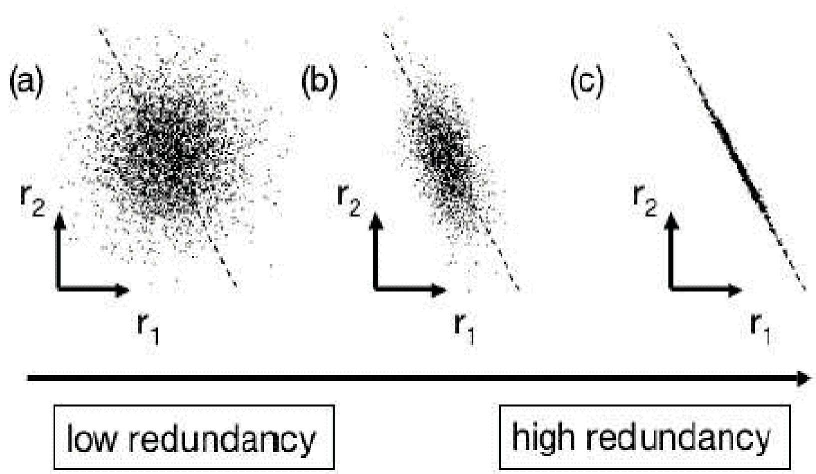


Fig2.5 shows the redundancy of variables [46]

Principal component analysis gives the smaller set of variables with less redundancy, which gives the good representation of the whole data set when considering the less number of the data variables. Principal component analysis is used to remove the redundancy variables are measured in correlations.

2.5.1.1 Rotation of axes:

The principal components are initially obtained by rotating axes in order to line with extension of the system, the new variables are uncorrelated and reflects the direction of maximum variance. If the result doesn't have any interpretation the axes are again rotated. To improve rotation we can consider the factor analysis which in turn doesn't destroy any of the properties.

2.5.2 Algebraic Approach:

In principal component linear combination with maximal variance is expected to be extracted.

The principal components of variables $X_1 \ldots X_n$ are linear combinations $A_1X \ldots\ldots\ldots A_n X$ such that A_1= arg Max V [AX]

.

.

$$A_k = \text{arg Max V [A X]}$$

$$A: \|A\|=1,$$

$$A \perp A_1 \ldots A_k$$

Solution is found through the Eigen problem for covariance matrix.

Find λ, $v \neq 0$ s.t.

$$\mathbf{\Sigma}\, V = \lambda\, v$$

2.12 GENERAL MODEL OF PCA:

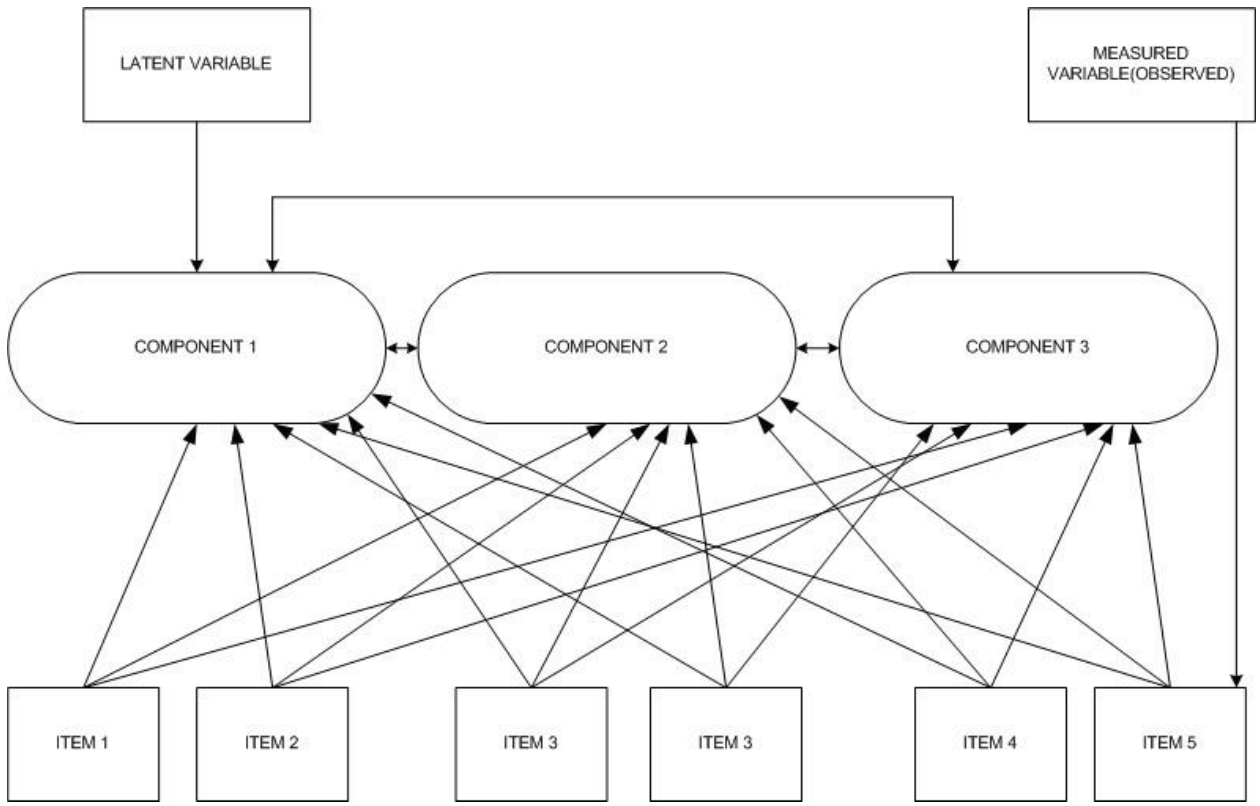


Fig2.5 Indicates general Principal component Model [18]

The Principal Component Analysis interacts with all the variables in the dataset and the principle components give the underlying structure in the data set. Fig 2.4 shows that the principal Component Analysis structure in which the principle components are interacting with all the 6 variables or items.

2.6 Principal components:

The decision must be made to retain the principal component analysis for effectively summarize the data. There are four rules to retain the principal components

1. Retain sufficient components to account for specified percentage of maximum variance.
2. The components whose eigenvalues are greater than average of eigenvalues. For correlation matrix.
3. Screen graph, a plot between λ_i versus i and look for natural break between the large eigenvalues and smaller eigenvalues.
4. Test for significance of large components that is components corresponding to the larger eigenvalues.

2.6.1 Method 1:

Selecting an appropriate threshold value percentage is a difficult part in this method. If we aim high we may include sample specific or variable specific. Sample specific components may not generalize to the population. Variable specific the component may not contain summary of all variables.[48]

2.6.2 Method 2:

This is used in many software packages like Minitab, stat soft follow this method. This gives the average variance of the individual variables. This method retains the components that account more variance than average variance of the variables.[48]

2.6.3 Method 3:

The screen graph between the Eigen number and the Eigenvalue size gives the components to be retained. The recommendation is to retain those eigenvalues in the steep curve before the first one on the straight line. The turning point between steep curve and the straight line may not be as distinct as this or there may be more discernible bend. In such cases this may not valid conclusion.[48]

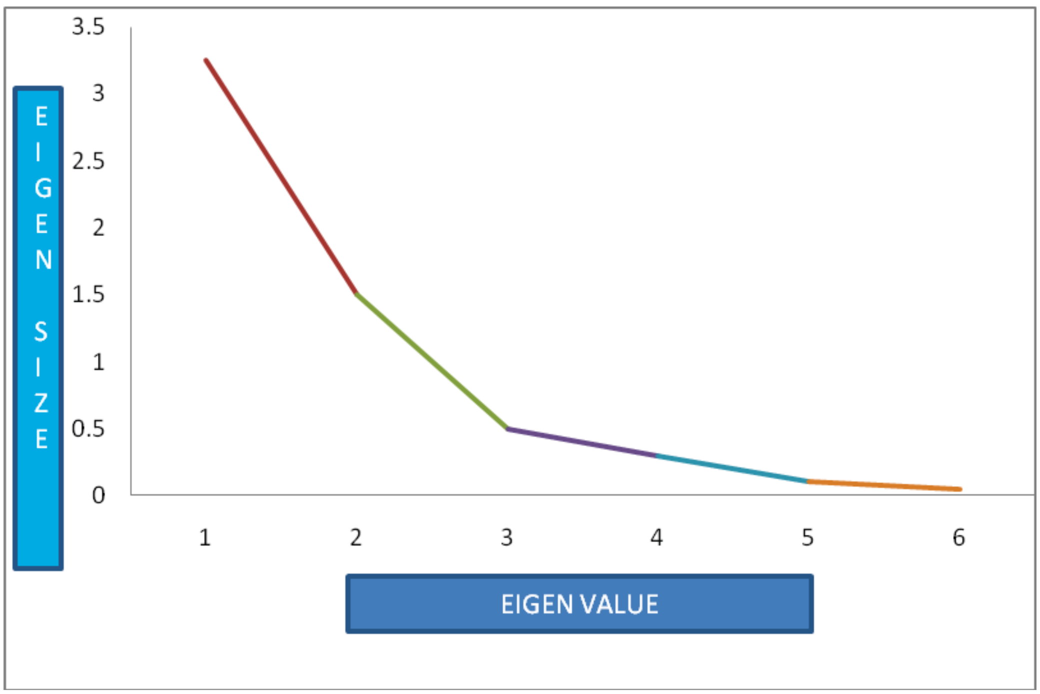


Fig 2.7 shows the graph between Eigen value and Eigen size

2.6.4 Method 4:

Test for significance, preliminary test of independence of variables are to be done .If the variables are independent there is no point in performing principal component. To test significance of large components we test the hypothesis that the last k population eigenvalues are small and equal.[48]

Tab2.4 Applications of Principal Component Analysis Applications

APPLICATIONS	RESEARCHERS
Quality Control	Feng Zhang et al./2007
Principal component analysis on Bio informatics	[xi chen lily wang " jonathan d.smith and bing zhang,], [Leif E. Peterson][19,20]
Principal component analysis in environment and Transportation	[Chock et al (1975) and Roch and pellerin(1982)],[Nagendra,Mukesh Khare [21,22]
Principal Component Analysis in Process Engineering	[Ricardo Dunia [1], S. Joe Qin [2 *], Thomas F. Edgar [2], Thomas J. McAvoy [3]] [23]
Principal Component Analysis in Face Recognition	[Kwang In Kim, Keechul Jung, and Hang Joon Kim], [M. Black and A. Jepson.], [B. Moghaddam and A. Pentland], [T. Cootes, G. Edwards, and C. Taylor.][N. Oliver, B. Rosario, and A. Pentland.] [24-28]
Principal Component Analysis in Manufacturing field	[WafikHachichaa,FaouziMasmoudm and Mohamed Haddar] [29]
Principle Component Analysis on Online Banking	[ChienBruceHoa,DeshengDashWub] [30]
Variable selection in large data set using PCA	[Jacquelynne R. King and Donald AJackson] [31]
Fault detection and Isolation with Principal Component Analysis	[Yvon Tharrault, Gilles Mourot, José Ragot, Didier Maquin] [32]
Process Fault detection and diagnosis by PCA	[Tao He,Wei-Rong Xie,Qing-Hua Wu,Tie-Lin Shi] [33]
Image Compression	[luminita state, dept. of computer science, university of pitesti, pitesti, romania, catalina lucia cocianu][34]

3. METHODOLOGY

Impact Assessment is used as a technique to choose the environmental site for analysis using the GaBi 4 software the analysis will be done on the site considering all the pollutants affecting the area. The Global Warming Potential is taken as criteria for selecting the site. If global warming potential near the site is high the site is considered for the further analysis. Principal component analysis is a multivariate technique in which a number of related variables are transformed to (hopefully, a smaller) set of uncorrelated variables. This is a tool used to reduce the interrelated variable dimensionality. PCA is used in for compression of data, in processing of Images, visualization, used in analysis of data and time series prediction.[35] The principle component analysis will help us in finding out the redundancy in large set of data. It gives the relation between the variables. PCA reduces the dimensionality of the data without effecting the total information contained in the data.

3.1 Approach towards the problem:

The data collected from different site maps are first analyzed using Impact Assessment using Gabi Software using global warming potential will determines whether we have do principal

The data collected at different sites from PEM may contain redundancies; data collected at one site might be equal to the data collected at different site. First the means from each data are subtracted and the variance and covariance between each site are calculated. Normalize the covariance by square root variances to form correlation matrix. This matrix is used to forms the basis for PCA.

Eigen vector have a dimension equals to 5 (which are the number of the sites in this case). First eigenvector represents the largest variance across the data.

If p is variable and n is number of samples. First principle component (X1) is gives the linear combination of the variables $Y_{1,} Y_2\ldots, Y_p$

$$X_1=b_{11}Y1+b_{12}Y_2+\ldots\ldots\ldots+ b_{1p}Y_p$$

In a matrix notation

$$X_1=b_1^T Y$$

The first principle component is computed such that it accounts for greatest possible variance in data set. The variance of X1 can be made as large as possible by choosing larger value for the weights $b_{I1,}$ $b_{I2\ldots}$ $b_{rp.}$ The weights are calculated with the constraint that their sum of squares is I.

$$b_{11}^2+ b_{12}^2 +\ldots\ldots.+ b_{1p}^2 =1$$

Second component is linear combination with maximal variance in orthogonal to first principle component analysis.

$$X_2=b_{21}Y1+b_{22}Y_2+\ldots\ldots\ldots+ b_{2p}Y_p$$

This continued until total of p principal components have been calculated equal to number of variables. Sum of the variances of al principle components will be equal the sum of variances of all the variables. Collectively, all of these transformations of the original variables to principle components are [36]

$$X=AY$$

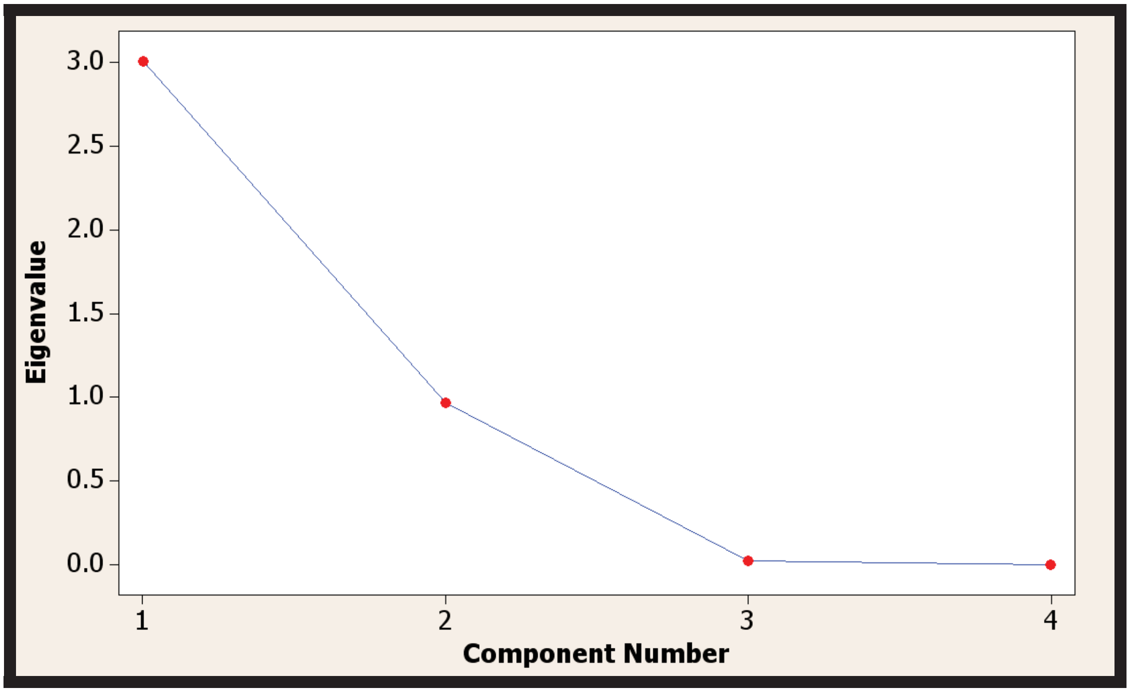


Fig3.1 Shows which components are to be retained for analysis

The Principle Components are selected based on the Eigen value if Eigen value is greater than 1 the principle component is retained otherwise it is omitted from the dataset. Actually the uncorrelated data in the original scene is pushed off into the high PC components. There will be no correlation and no more redundancy in the new data set which is obtained in the PC. [37]

3.2 Example of PCA:

This example shows how PCA can be used for analysis of the dataset which contains two variables and how it can be used for interpreting the results.

Consider this example: For suppose if we take two variables A and B which are consumed by the people of different age groups. This analysis gives the idea how PCA can be calculated.

Tab3.1: Survey data is shown in the table [47]

X	Y
1.5	2
2.6	1.1
3.1	3.4
1.8	1.3
1	1.5
2	1.9
1.9	2.8
3.1	1.4
1.6	2.1
1.9	1.7

Step 1: First the data is adjusted by subtracting the mean from A and B

X	Y	X-$\bar{X}$	Y-$\bar{Y}$
1.5	2	.55	.02
2.6	1.1	-.49	.92
3.1	3.4	-.99	-1.38
1.8	1.3	.31	.72
1	1.5	1.11	.52
2	1.9	.11	.12
1.9	2.8	.21	-.78
3.1	1.4	-.99	.62
1.6	2.1	.51	-.08
1.9	1.7	.21	.32

$\bar{X} = \frac{\Sigma x}{n}$=2.05; $\bar{Y} = \frac{\Sigma y}{n} = 1.92$

Step 2: Calculation of Covariance Matrix:

(X-$\bar{X}$)*(Y-$\bar{Y}$)	(X-$\bar{X}$)^2	(Y-$\bar{Y}$)^2
0.011	0.3025	0.0004
-0.4508	0.2401	0.8464
1.3662	0.9801	1.9044
0.2232	0.0961	0.5184
0.5772	1.2321	0.2704
0.0132	0.0121	0.0144
-0.1638	0.0441	0.6084
-0.6138	0.9801	0.3844
-0.0408	0.2601	0.0064
0.0672	0.0441	0.1024

$$\text{cov}_{XY} = \frac{\sum (X - M_X)(Y - M_Y)}{n}$$

Covariance of XY= .0.109867

$M_{X}=\frac{\sum x}{n}$=2.05; $M_y = \frac{\sum y}{n} = 1.92$

n=N-1=9

Covariance Matrix:

$$\text{Cov}=\begin{bmatrix} .465711 & .109867 \\ .0109867 & .517333 \end{bmatrix}$$

You can observe that diagonal element is positive that means x increases with y

Step3: Calculating Eigen values and Eigen Vector of Covariance matrix

$$\text{Cov}=\begin{bmatrix} .465711 & .109867 \\ .0109867 & .517333 \end{bmatrix}$$

From covariance matrix Eigen values and Eigen vectors are derived

$$\det \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} - \begin{bmatrix} .465711 & .109867 \\ .0109867 & .517333 \end{bmatrix} = 0$$

$$\det \begin{bmatrix} c - .465711 & -.109867 \\ -.09867 & c - .517333 \end{bmatrix} = 0$$

(C-.465711)*(C-.517333)-(-.109867)$^{\wedge 2}$=0

C^2-.983044C+.228856=0

C=0.378666, 0.60438

Eigen values:

$$\begin{pmatrix} 0.378666 \\ 0.60438 \end{pmatrix}$$

Eigen Vector corresponding to C=0.378666 is given by the following equation:

(A-cI) * x = 0

$$\begin{bmatrix} .465711 & .109867 \\ .0109867 & .517333 \end{bmatrix} - 0.378666 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} * x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Solving value for x we get the $\begin{pmatrix} -0.7838006 \\ 0.621006 \end{pmatrix}$

Eigen Vectors:

$$\begin{bmatrix} -.783806 & -.621006 \\ .621006 & -.783806 \end{bmatrix}$$

From the above Eigen vectors we have to chose the significant one which our principal component.

$$\begin{pmatrix} -.621006 \\ -.783806 \end{pmatrix}$$

This Eigen vector explains the whole relation between X and Y i.e if one increases other increases and if one decreases other decreases.

Consider this example: For suppose if we take survey on Drinks like Pepsi, Coke, Fanta and Sprite which are consumed by the people of different age groups. This survey gives a good idea of how people consume the drinks.

Tab3.1: Survey data is shown in the table

PEPSI	COKE	FANTA	SPRITE
23	20	55	12
45	43	34	23
34	32	56	34
64	66	78	67
32	32	12	42
76	76	56	42
45	45	52	23
56	102	67	67
39	78	87	32
12	24	45	78
39	39	34	45
20	40	68	98
29	29	90	89
56	23	78	12

Performing the PCA on this data gives the following results using Mini Tab Software:

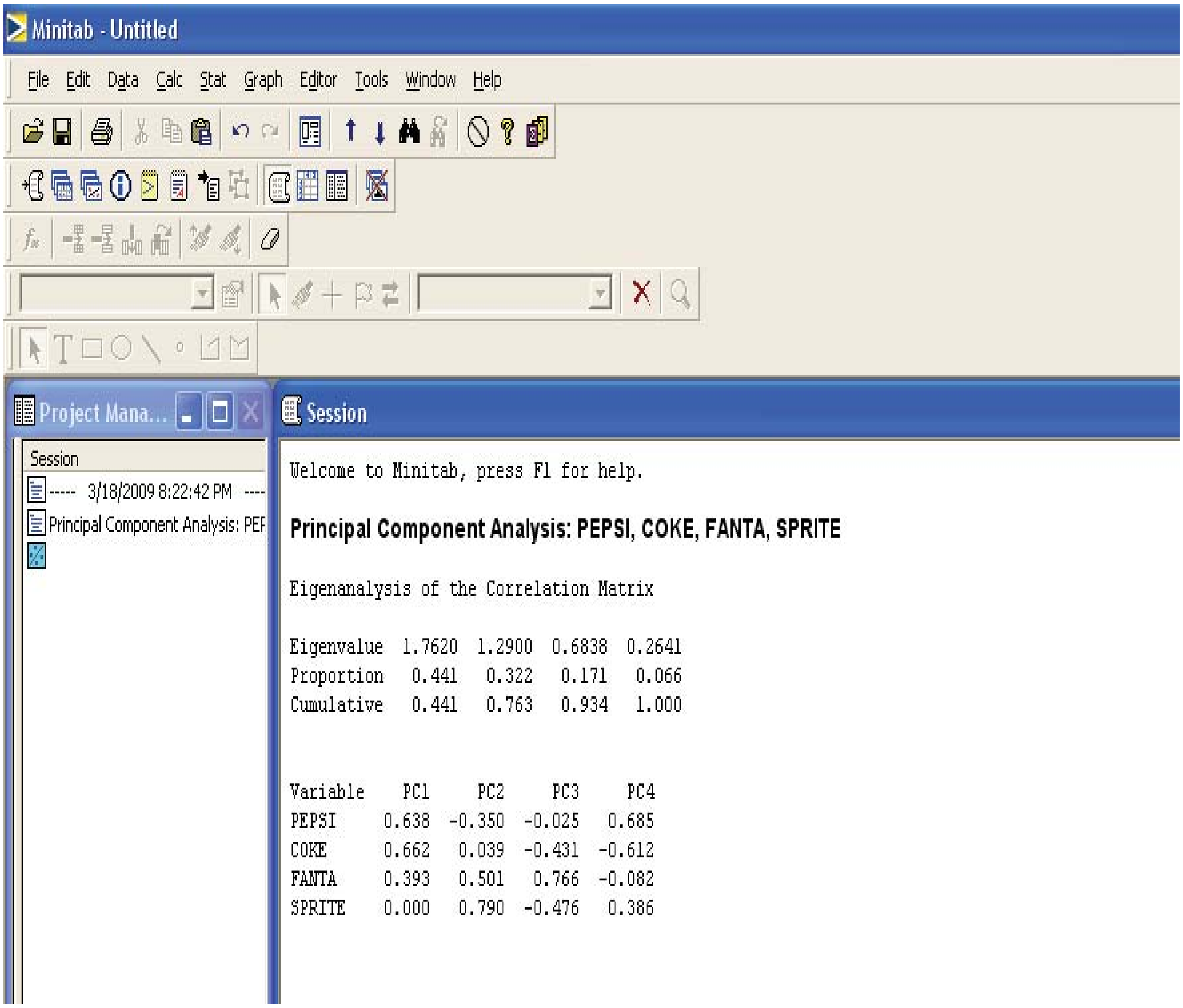


Fig 3.2: Screen shot of the results from Mini Tab

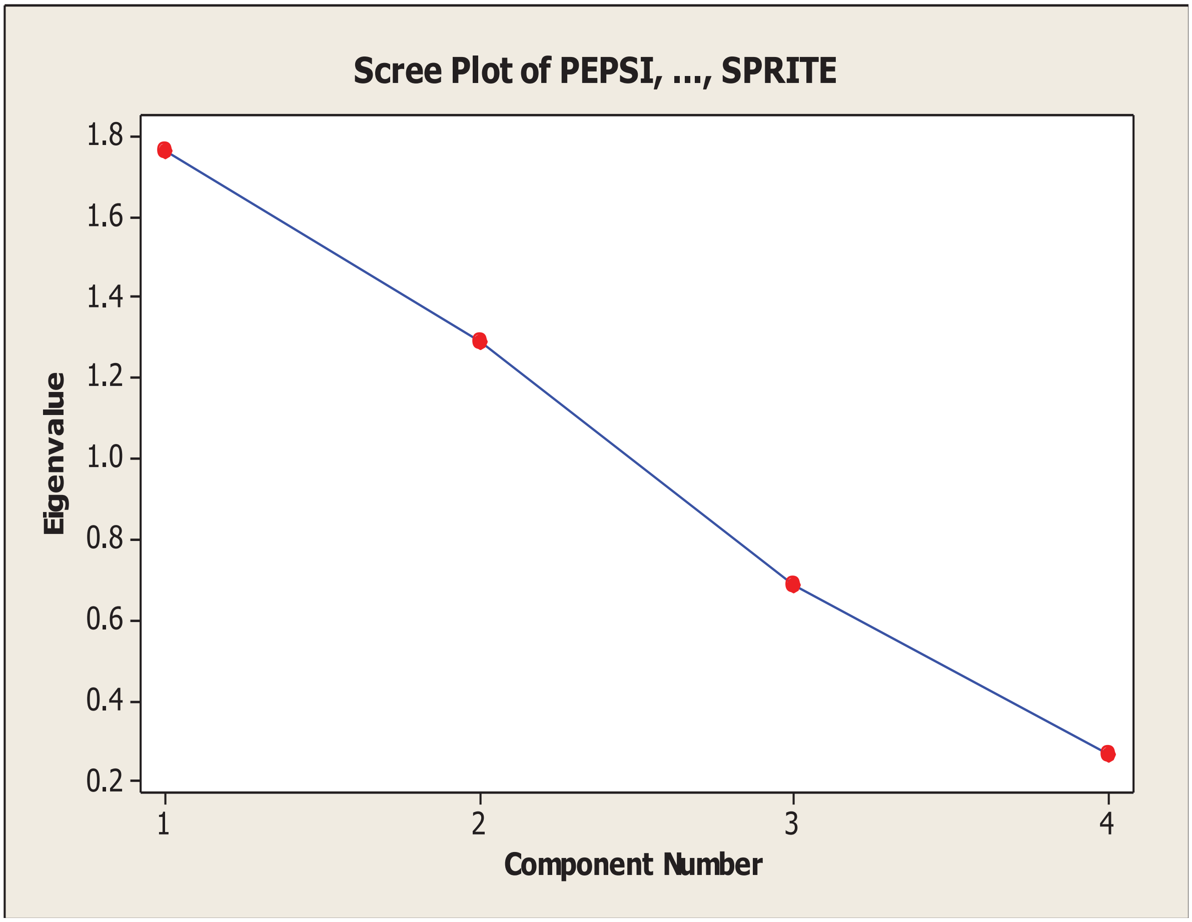


Fig 3.3 Screen shot of graph between Component number and Eigenvalue

3.2.1 Interpretation from graph:

The components which have Eigenvalue greater than one should be selected for analysis. The graph shows that the first two components have Eigenvalue greater than one in this case.

3.2.2 Analysis from Principal components:

The Dimensionality of the variables is reduced as you can see the four variables are explained in two principal components. The two principle components explain about 76% of total variability associated with total variable dataset. This suggests that we can capture most of variability in data with less than half of the original dimensions. The loadings of principal components also shows that that the consumption of Pepsi and coke by costumers are correlated this shows that amount of Pepsi and coke consumption are almost same and In the second principal component loadings shows explains about the Fanta and Sprite consumption.

3.4 The Experimental set up:

The air Exposure screening study is community based participatory research as well as a project that will provide relevant field monitoring data. Which collects data of ambient air pollutants at different locations ,these locations are selected based on multiple factors 1) availability of equipment 2) Wind Direction 3) accessibility to the sample site 4) relative distance of the residential areas from the center of quadrants 5) distance from major traffic roads. Accordingly five different locations are selected and they are divided into five quadrants namely Q1, Q2, Q3, and Q4 to collect the particulate material which is obtained from the sampler placed to collect different samples like carbon Monoxide (Co), Oxides of Nitrogen (NO/NO2/NOx) Sulfur Dioxide (SO2), Ozone (O3),Volatile Organic Compounds ,Formaldehydes/Acetaldehydes and Particulate matter (PM2.5 and PM10) and Heavy Metals are collected using Personal Environmental Monitor (PEM) device in all quadrants each PEM unit consists of air pump with pre-weighted Teflon filters in order to collect PM in the ambient air .PM2.5 and PM10 will be collected by separate PEMs. The filter will be replaced every 24 hours .Collected samples will be weighed at contract laboratory before XRF testing for heavy metal analysis. Each sample is collected after 24 hour for 25 days both PM2.5 and PM10.The data is collected at different locations i.e. Q1, Q2, Q3, Q4 and Q5 and the final analysis is done on the data collected and interpretations are done each monitoring site like how the PM is affecting the community.

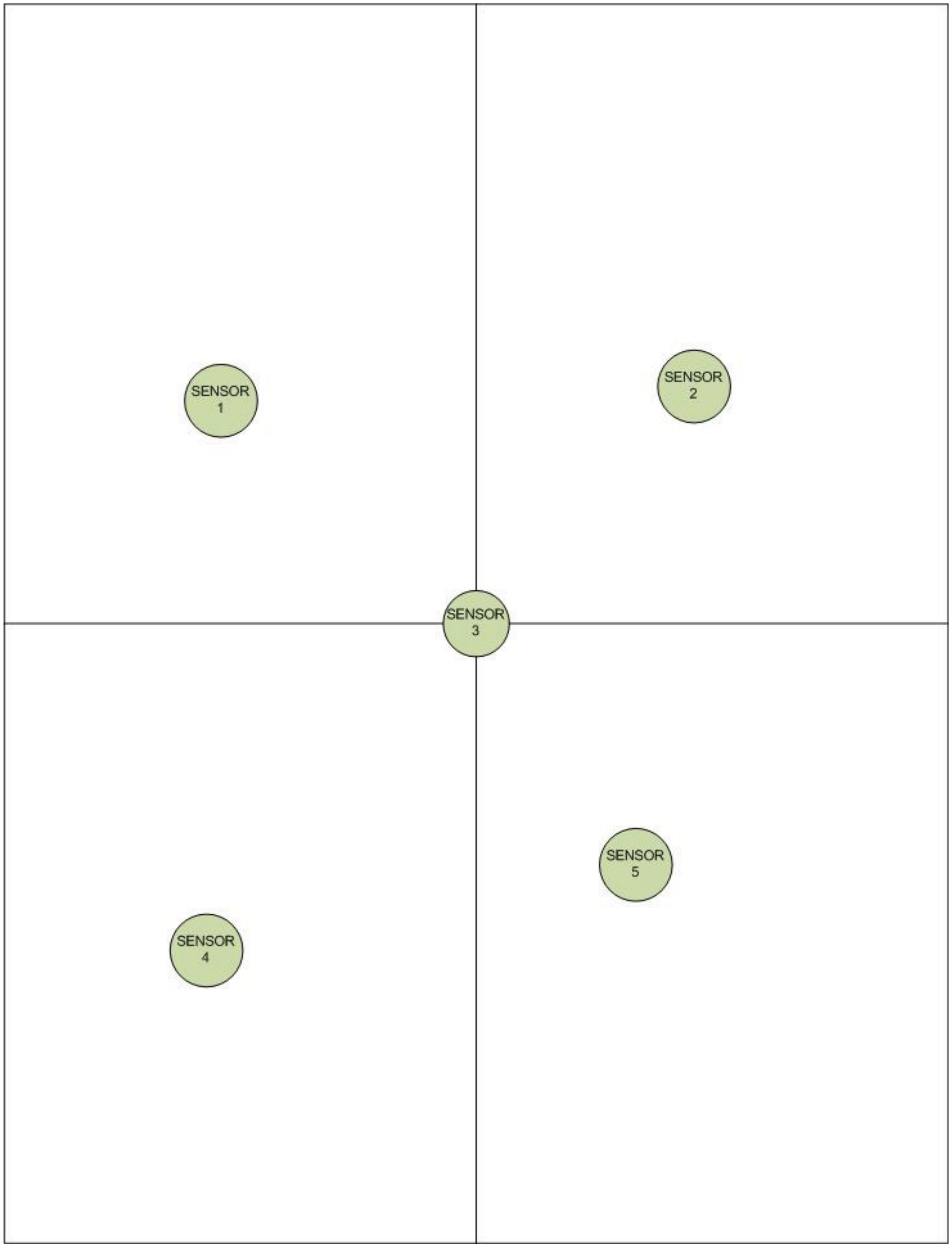


Fig3.4 shows the experimental setup for collecting the data

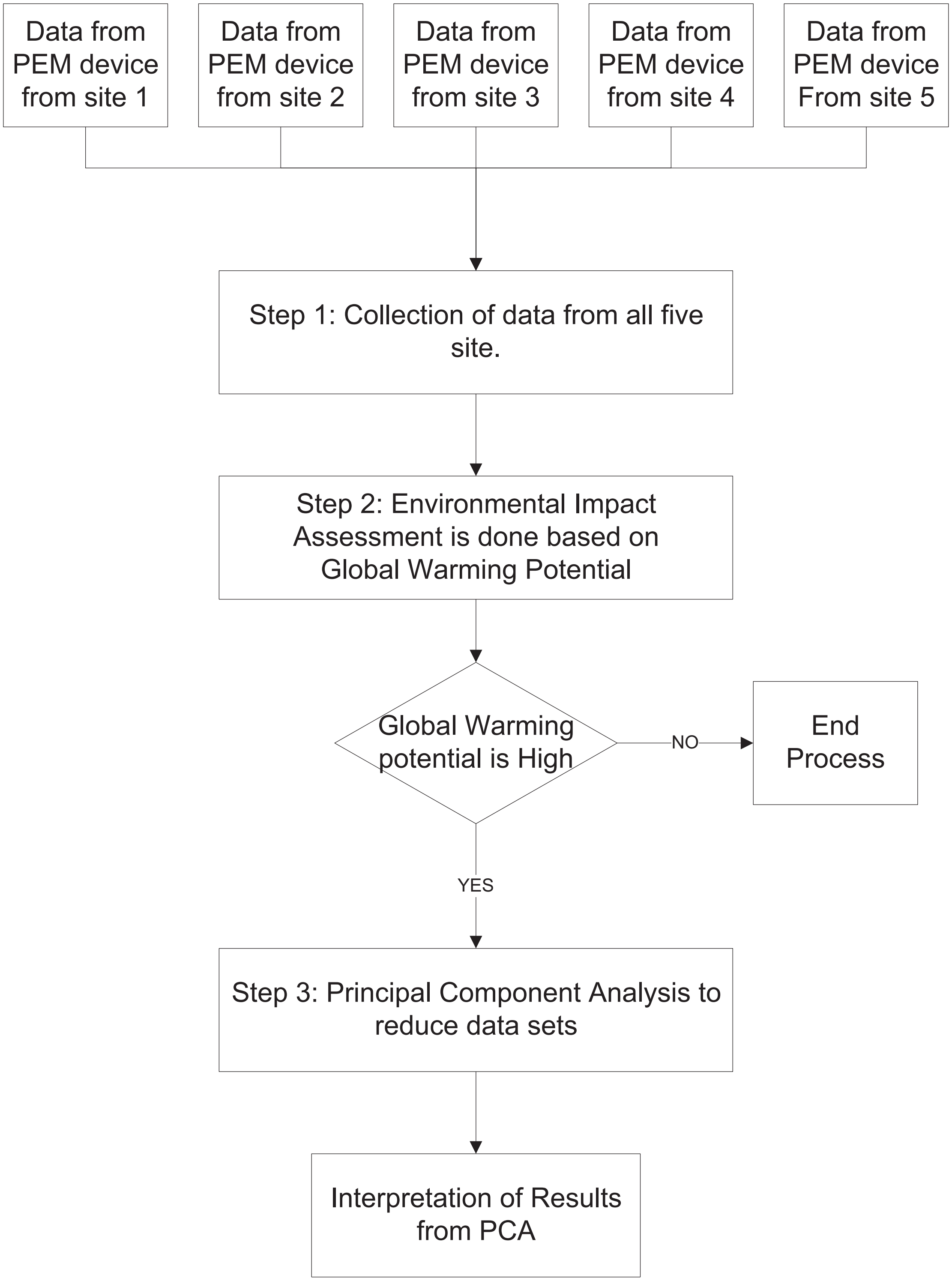


Fig 3.5 Conceptual Framework

4. RESULTS AND CONCLUSIONS

4.1 Impact Assessment:

For selecting a site map one need to do Impact Assessment first .In this Research the site maps are selected by performing Environmental Impact Assessment on the site area using Impact Assessment software called GaBi 4.The site maps contains large number of pollutants which are mention above and the source from which the pollutants are entering into atmosphere are taken into consideration while performing Impact assessment.

Impact Assessment is done on the site by taking the following things into consideration. The site map has the following type of pollutants Particulate material PM 10 and PM 25. While performing Impact Assessment the plants surrounding that area and the pollutants emitted from these plants are taken into consideration. For example in this site there is a steel plant there are different process involved in making of steel and each process emits different pollutants. These are mentioned in this Gabi Model the main pollutants such as Nitrogen Oxide, Sulphur Dioxide, Carbon Monoxide and Lead. These are mentioned in form of flows which are going into the site area which are mentioned in the Gabi software. The vehicle pollution is also taken into consideration while doing this Impact Assessment in this carbon monoxide gas is considered as flow or emission into the atmosphere. Global warming potential is calculated while considering all the pollutants from different process.

4.1.1Environmental Impact Assessment:

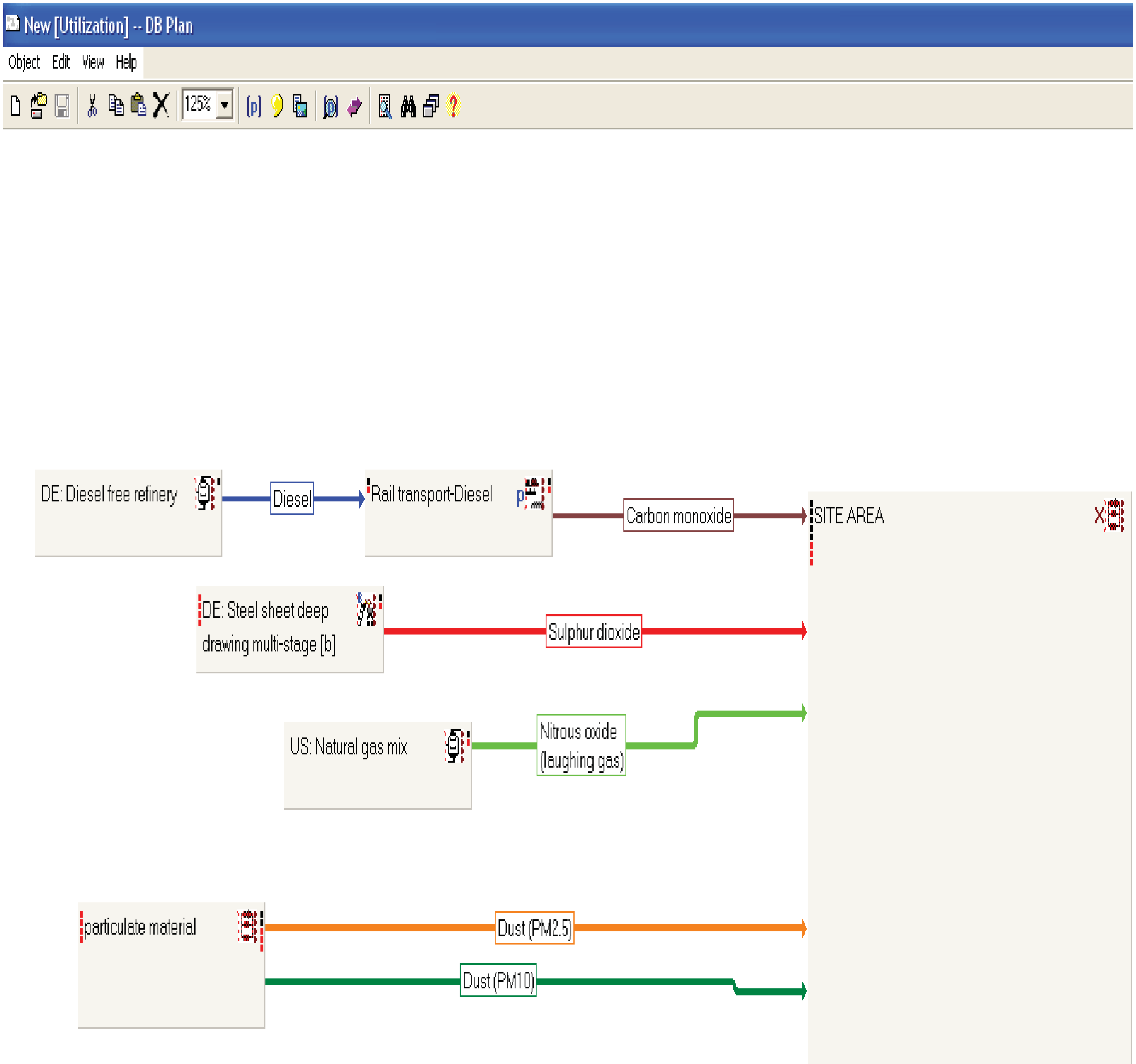


Fig4.1: Shows the flow of the Air Pollutants inside the Site

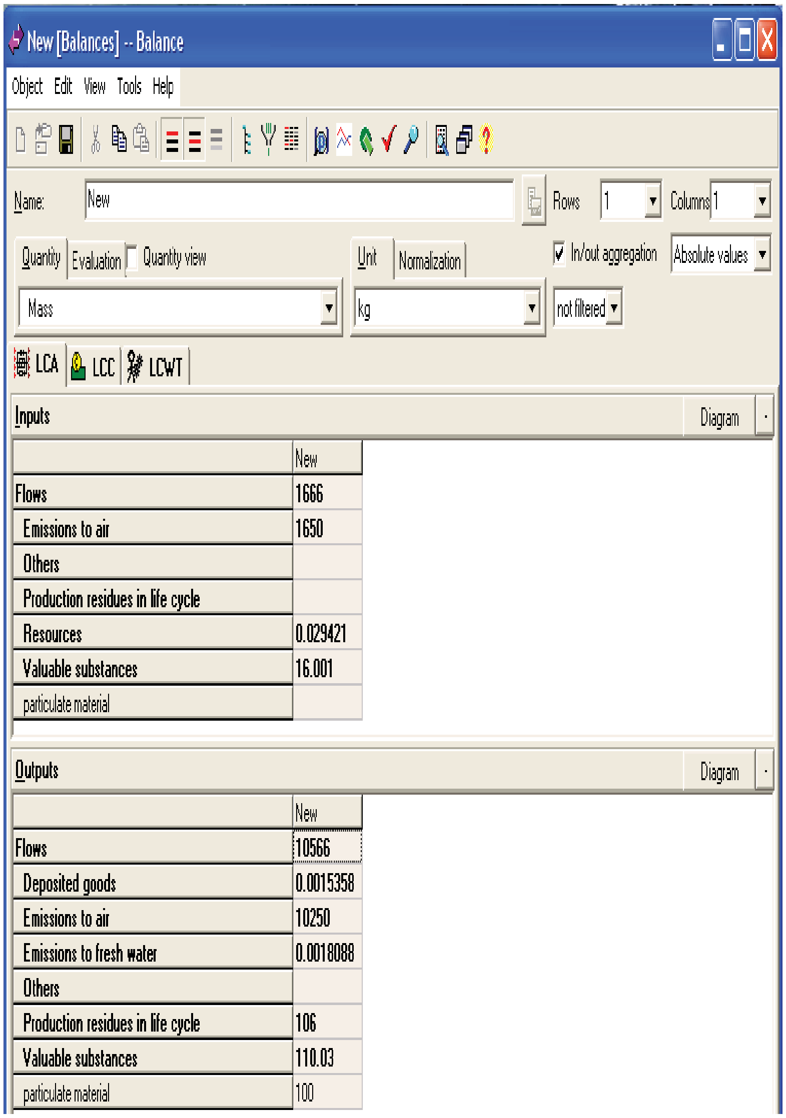


Fig 4.2: Input and output flow of the whole site map

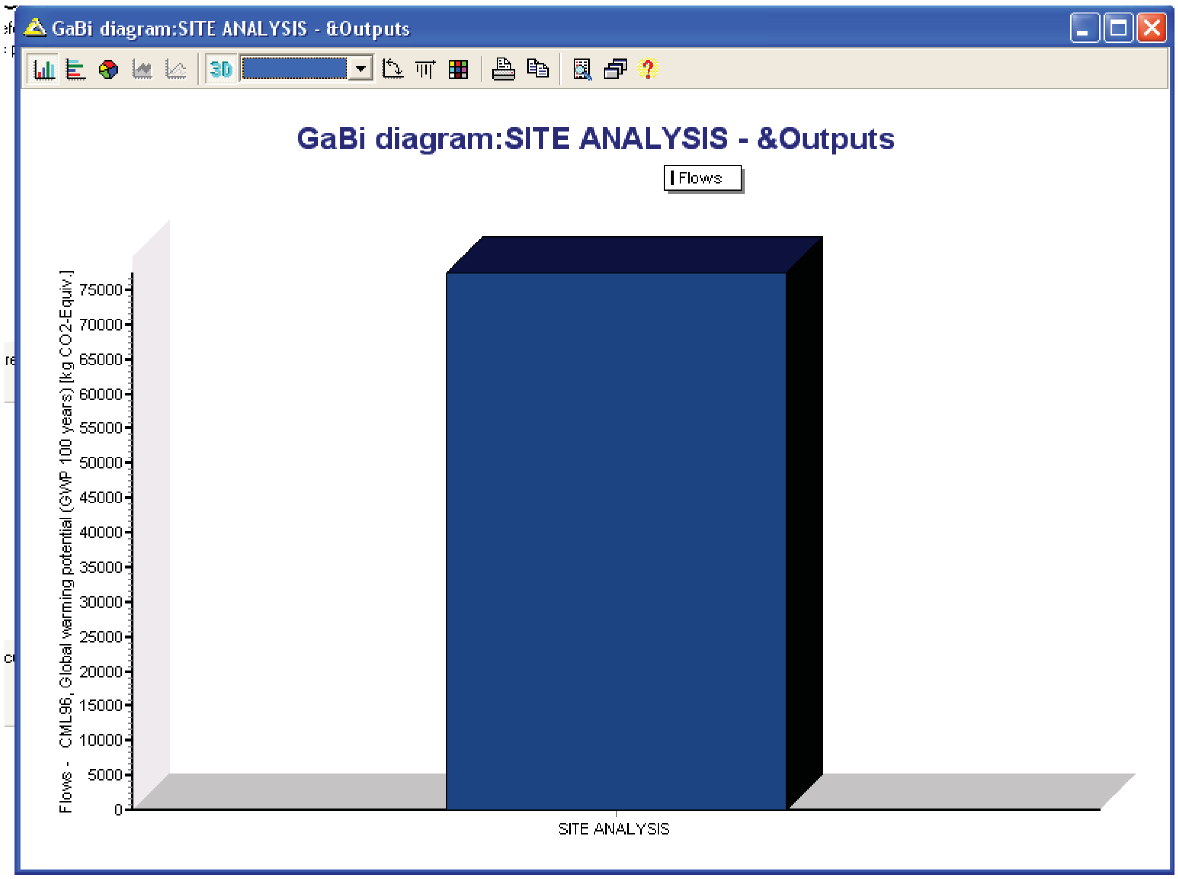


Fig 4.3: Graph of the Green Warming Potential

It was found that rise in carbon dioxide proportions to more than 450 parts per million will lead to increase in temperature by 2 degrees centigrade. In this graph the results from all the pollutants are mentioned in Kg equivalent to CO_2. The CO_2 content was really high in this case and further monitoring sites should be placed for future analysis.

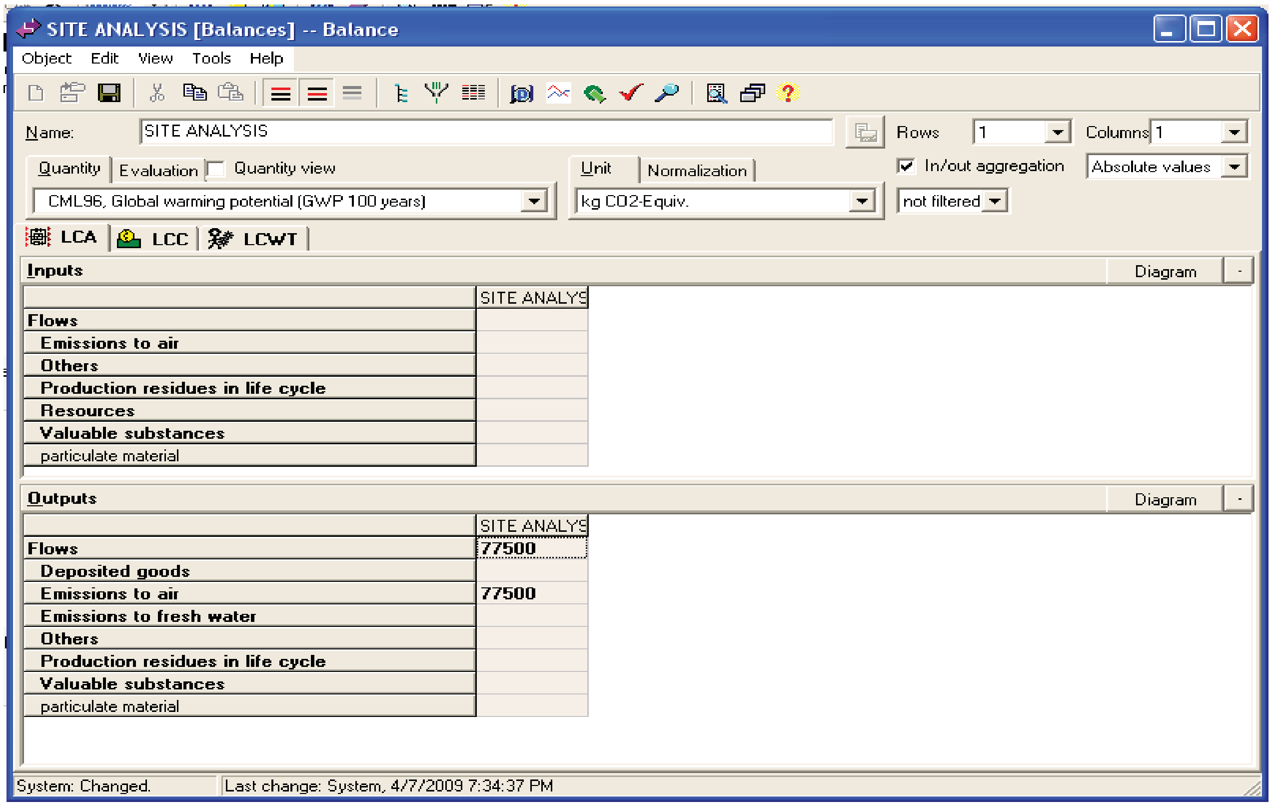


Fig 4.4: Shows the screen shot of output and input parameters of site

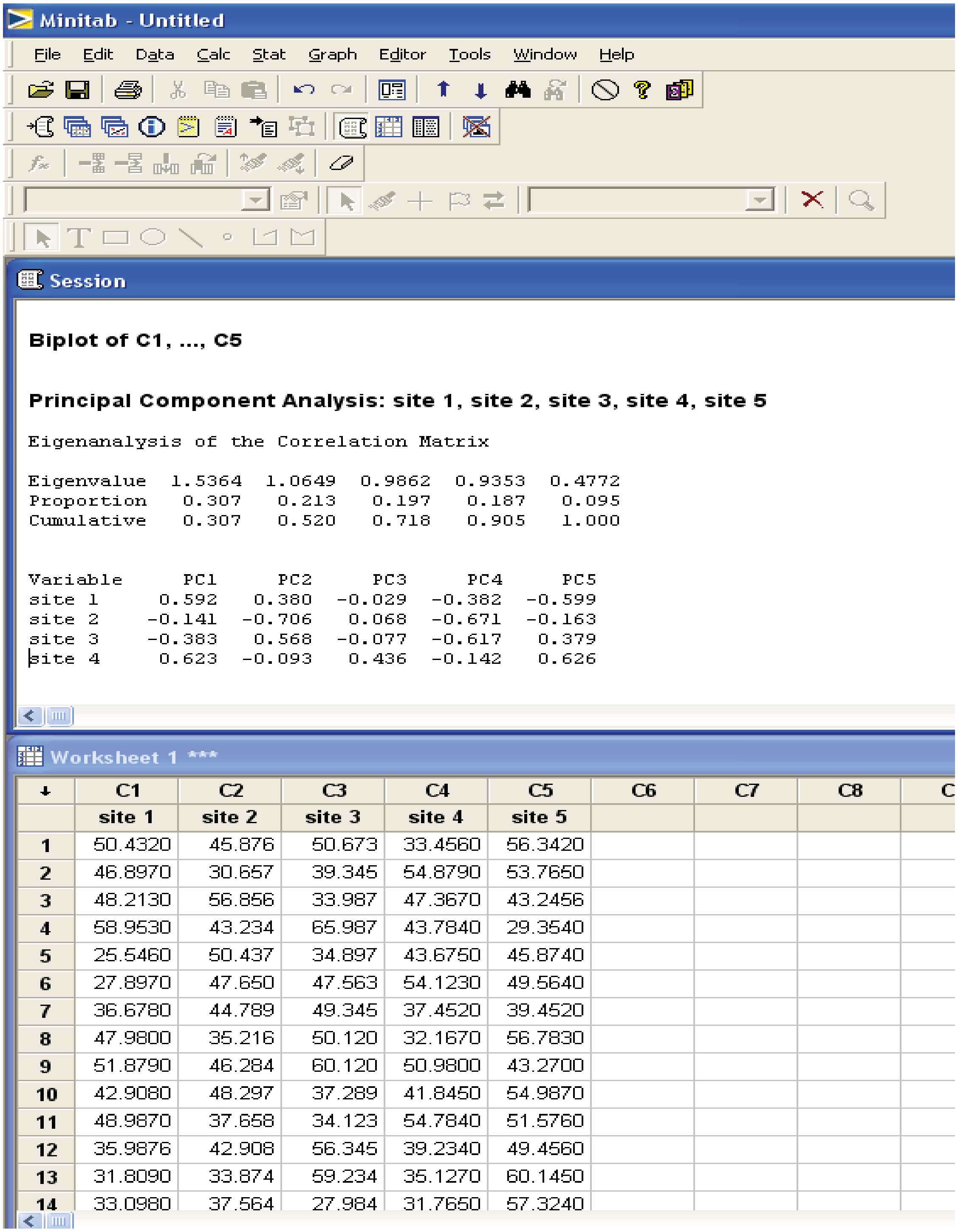
4.1.1 Results from Environmental Impact Assessment:

After Impact Assessment on the site it is clear that the Global warming potential is really high at the site and running the analysis for hundred years the Global warming potential has a very high impact on Environment leading to various health hazards. So it is really necessary to have the monitoring assessment on these site maps.

4.2 Principal Component Analysis:

Tab4.1: shows the data samples of PM 10 data

PM 10				
SITE 1	**SITE 2**	**SITE 3**	**SITE 4**	**SITE 5**
50.432	45.876	50.673	33.456	56.342
46.897	30.657	39.345	54.879	53.765
48.213	56.856	33.987	47.367	43.2456
58.953	43.234	65.987	43.784	29.354
25.546	50.437	34.897	43.675	45.874
27.897	47.65	47.563	54.123	49.564
36.678	44.789	49.345	37.452	39.452
47.98	35.216	50.12	32.167	56.783
51.879	46.284	60.12	50.98	43.27
42.908	48.297	37.289	41.845	54.987
48.987	37.658	34.123	54.784	51.576
35.987	42.908	56.345	39.234	49.456
31.809	33.874	59.234	35.127	60.145
33.098	37.564	27.984	31.765	57.324
37.98	29.876	40.267	35.478	30.245
42.765	39.876	43.876	54.563	55.873
24.143	47.854	54.893	41.34	27.987
56.98	50.439	31.265	56.7345	44.876
27.549	58.982	48.347	38.569	53.876
42.985	59.543	45.783	46.176	55.987
43.9123	43.98	78.987	38.972	59.123
53.912	23.765	45.234	55.432	49.987
49.469	56.432	50.376	43.965	56.926
36.983	39.756	41.64	40.678	48.872
25.123	43.875	58.345	38.178	23.158



Minitab - Untitled

File Edit Data Calc Stat Graph Editor Tools Window Help

Session

Biplot of C1, ..., C5

Principal Component Analysis: site 1, site 2, site 3, site 4, site 5

```
Eigenanalysis of the Correlation Matrix

Eigenvalue  1.5364  1.0649  0.9862  0.9353  0.4772
Proportion   0.307   0.213   0.197   0.187   0.095
Cumulative   0.307   0.520   0.718   0.905   1.000


Variable     PC1     PC2     PC3     PC4     PC5
site 1     0.592   0.380  -0.029  -0.382  -0.599
site 2    -0.141  -0.706   0.068  -0.671  -0.163
site 3    -0.383   0.568  -0.077  -0.617   0.379
site 4     0.623  -0.093   0.436  -0.142   0.626
```

Worksheet 1 ***

↓	C1	C2	C3	C4	C5	C6	C7	C8	C
	site 1	site 2	site 3	site 4	site 5				
1	50.4320	45.876	50.673	33.4560	56.3420				
2	46.8970	30.657	39.345	54.8790	53.7650				
3	48.2130	56.856	33.987	47.3670	43.2456				
4	58.9530	43.234	65.987	43.7840	29.3540				
5	25.5460	50.437	34.897	43.6750	45.8740				
6	27.8970	47.650	47.563	54.1230	49.5640				
7	36.6780	44.789	49.345	37.4520	39.4520				
8	47.9800	35.216	50.120	32.1670	56.7830				
9	51.8790	46.284	60.120	50.9800	43.2700				
10	42.9080	48.297	37.289	41.8450	54.9870				
11	48.9870	37.658	34.123	54.7840	51.5760				
12	35.9876	42.908	56.345	39.2340	49.4560				
13	31.8090	33.874	59.234	35.1270	60.1450				
14	33.0980	37.564	27.984	31.7650	57.3240				

Fig4.5 shows the screen shot of results for Principal Component Analysis

Fig4.6: Shows Graph shows the plot between the component number and Eigenvalues

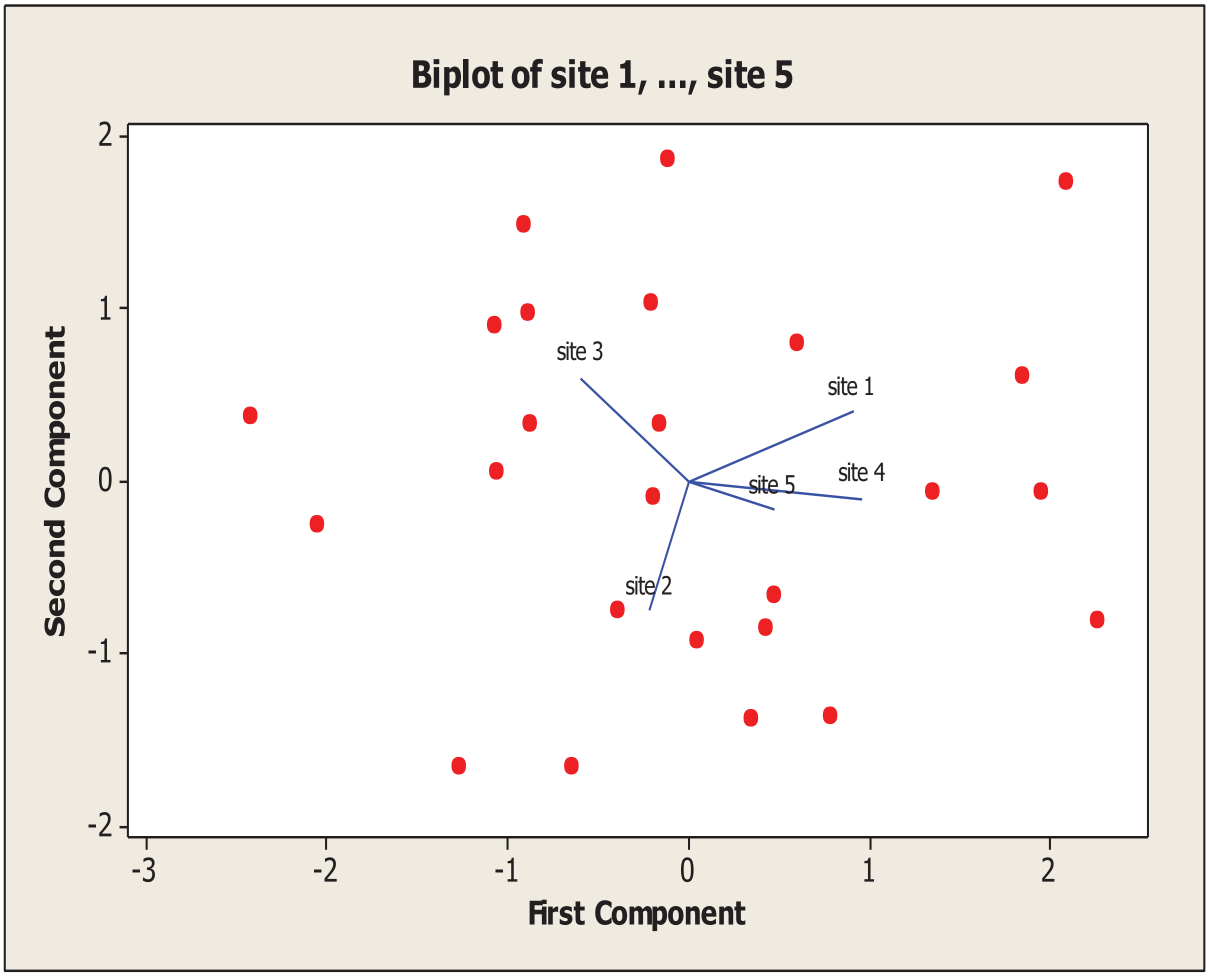


Fig 4.7: Plot shows the scatter of data of different sites

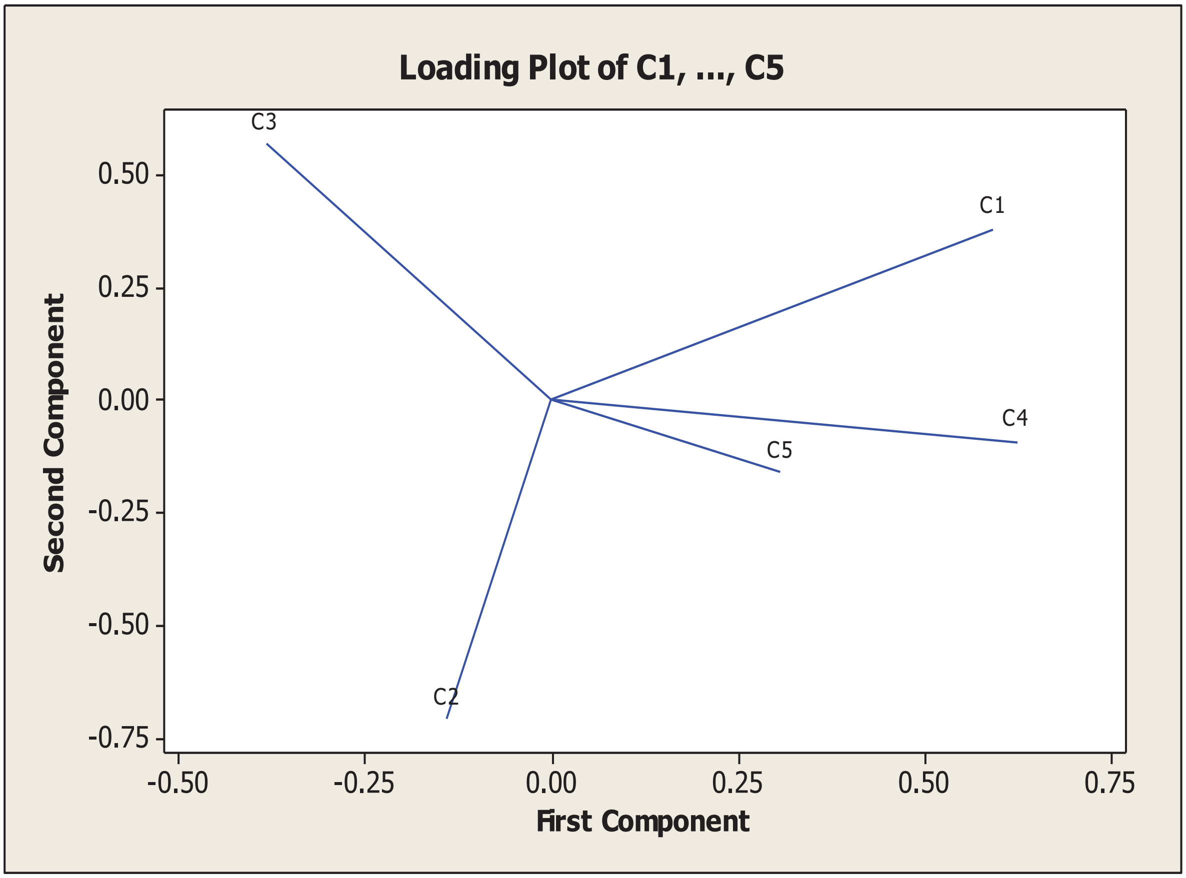


Fig 4.8: plot shows the loading of first and second components

Tab4.2: Shows the data samples of PM 25 data

PM 25				
SITE 1	SITE 2	SITE 3	SITE 4	SITE 5
56.12	35.9573	34.562	34.409	70.787
34.502	22.75	27.708	40.494	25.507
25.207	43.868	17.389	21.717	18.902
24.611	25.319	22.236	20.532	24.475
41.912	60.162	37.285	32.172	34.254
66.504	40.273	57.342	38.616	66.511
27.987	51.984	47.461	42.737	54.188
48.341	57.487	37.382	46.076	47.625
55.364	55.982	37.319	53.737	31.664
33.764	39.985	55.564	54.793	27.684
41.324	44.567	23.662	41.589	39.832
58.654	50.765	57.505	34.528	37.732
38.948	40.275	62.814	27.708	45.784
30.413	51.938	34.408	47.389	34.409
47.88	29.517	44.897	58.982	40.494
28.897	45.958	49.564	45.403	56.734
56.654	34.965	34.765	57.987	38.765
41.983	41.392	33.875	45.765	42.873
39.764	46.554	45.879	39.654	50.122
26.874	39.596	47.214	52.134	55.673
31.453	28.456	43.123	29.345	50.456
35.123	51.178	56.168	49.156	37.746
43.678	52.874	48.945	49.267	58.932
42.89	39.457	32.987	37.546	55.216
38.954	32.874	40.129	43.125	32.874

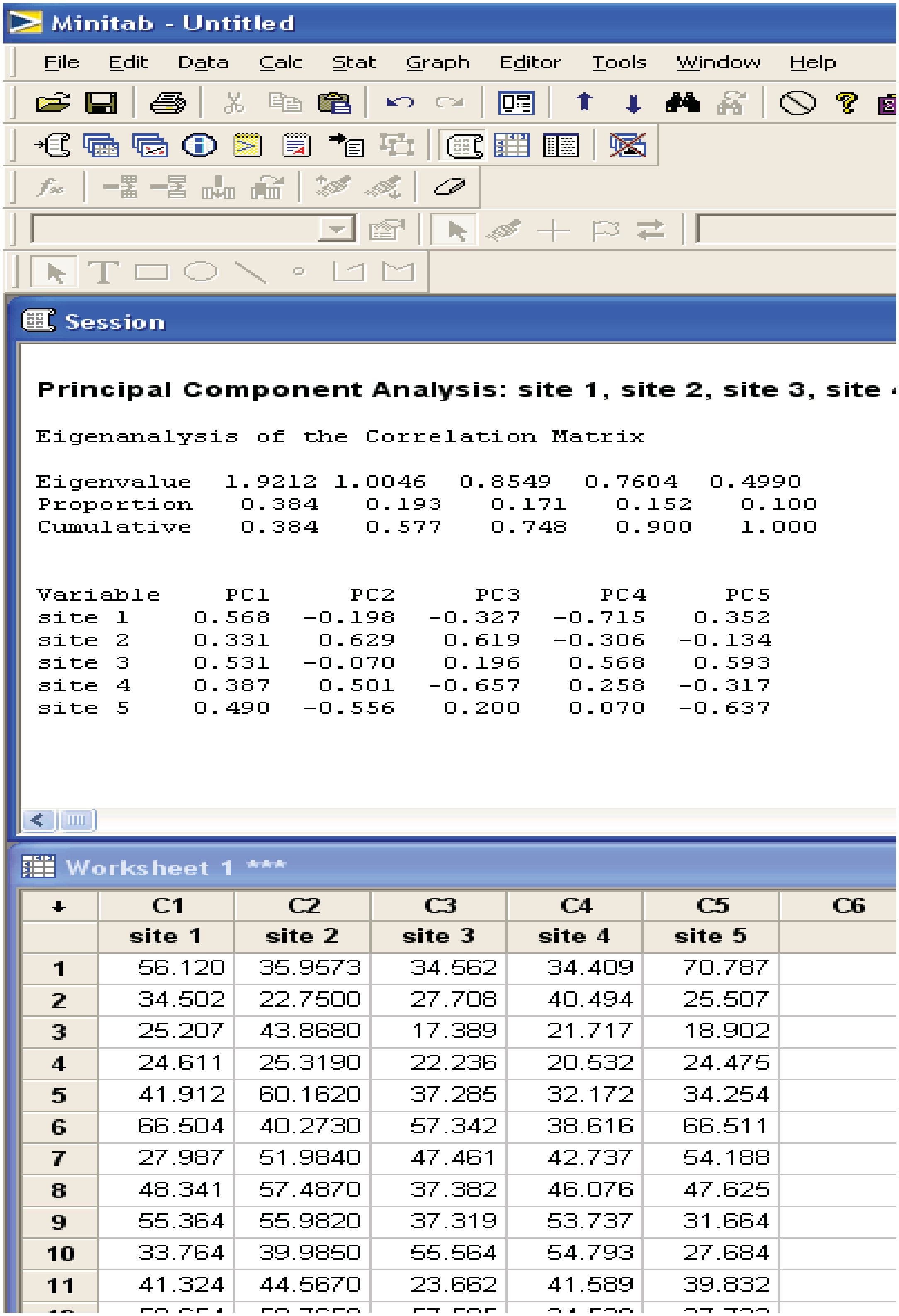


Fig4.9 shows the screen shot of results for Principal Component Analysis

Fig4.10: Shows Graph shows the plot between the component number and Eigenvalues

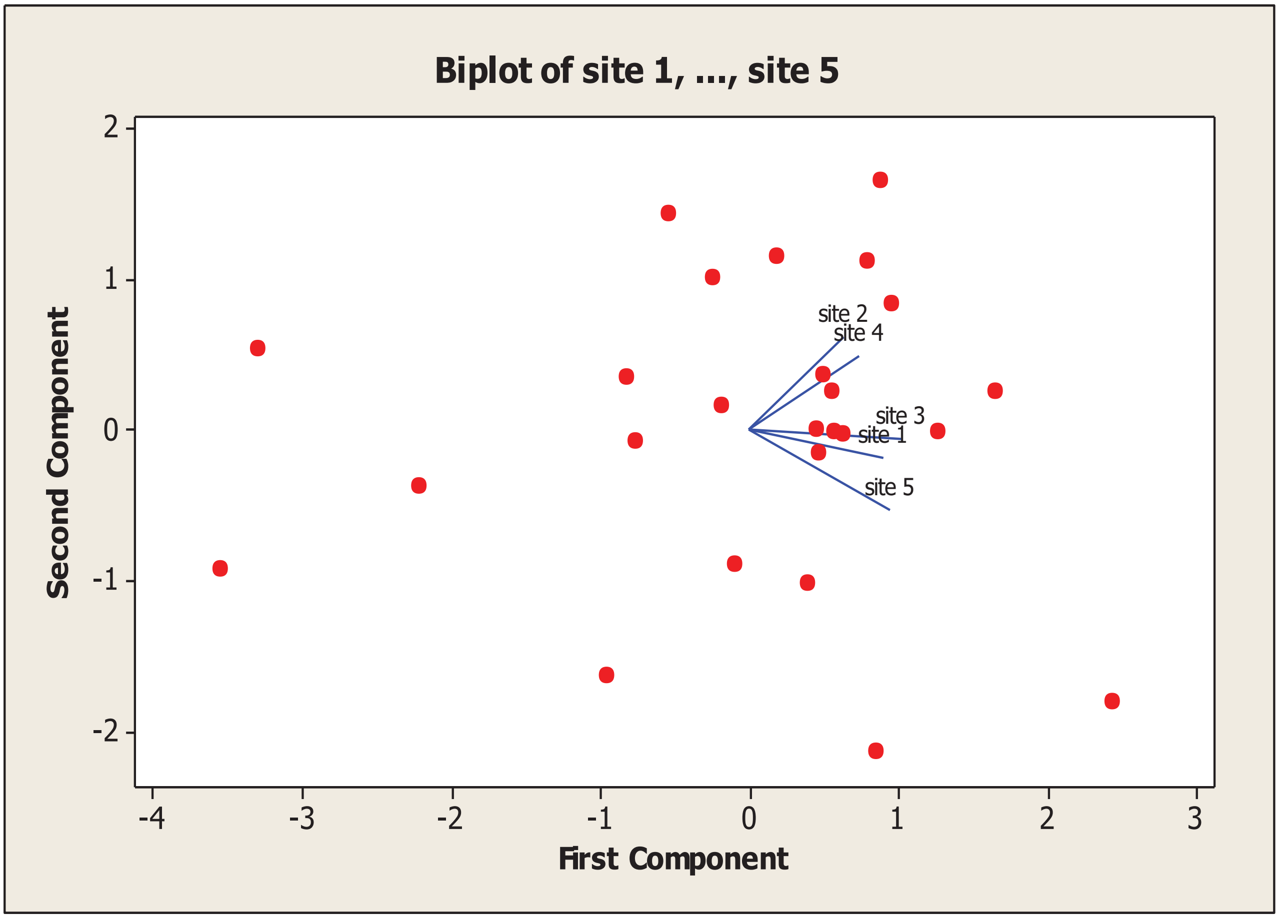


Fig 4.11: Plot shows the scatter of data of different sites

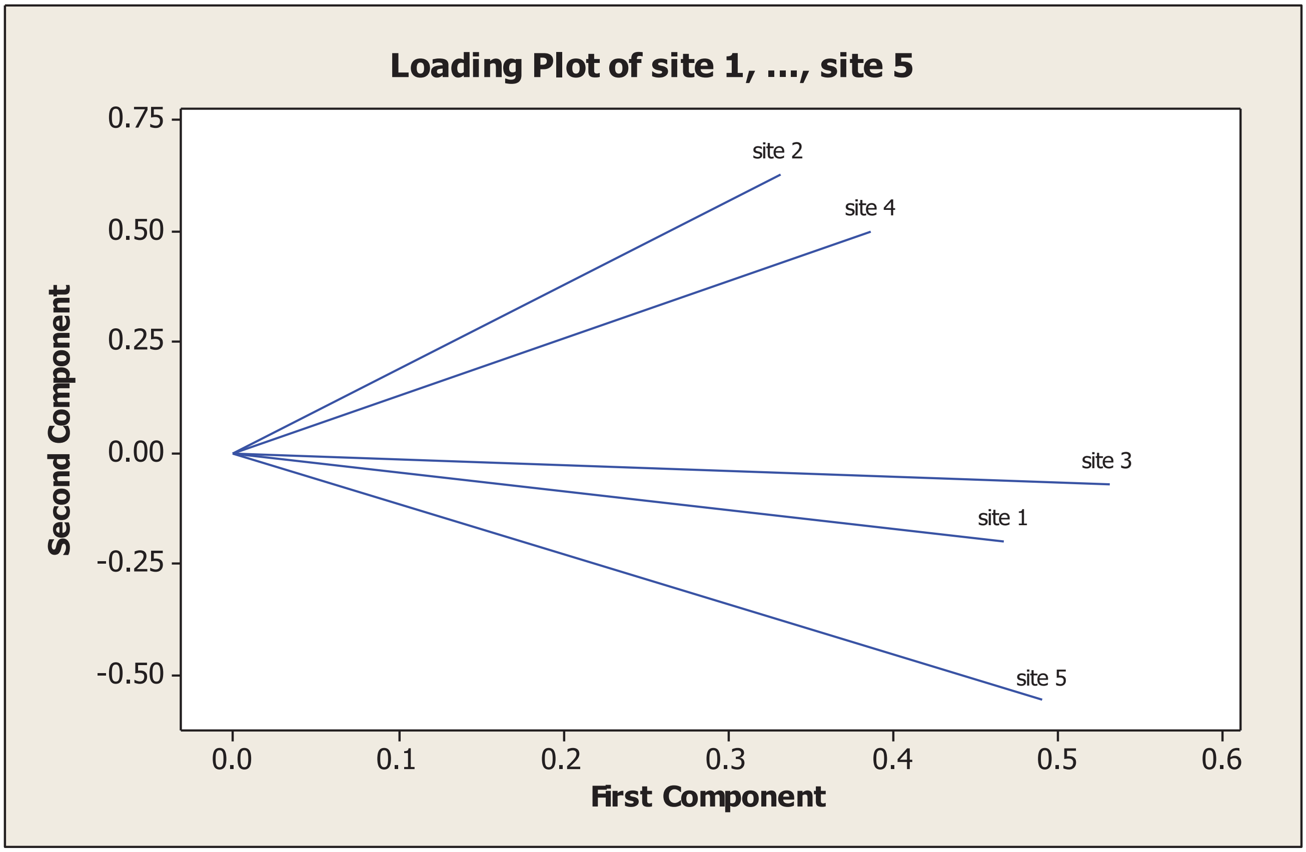


Fig 4.12: plot shows the loading of first and second components

4.2.1 Result From Principal Component Analysis:

Results of PCA on PM10 suggest that Principle component one and two are sufficient to explain the variability of data and further gives the insight into the data set. The dimensional reduction of the data set is achieved by performing PCA instead of considering all five variables we can consider two principle components to explain the whole data.

PC 1 gives that the site sensors1 and Sensor 4 are correlated and PC2 gives Senor 2 and Sensor3 are negatively correlated. That means sensor 1 and sensor 4 can be replaced by one sensor. Sensor 2 and sensor 3 can be replaced by one sensor for next readings in that area next time.

Results of PCA on PM .25 suggest that Principle component one and two are sufficient to explain the variability of data and further gives the insight into the data set. The dimensional reduction of the data set is achieved by performing PCA instead of considering all five variables we can consider two principle components to explain the whole data.

PC 1 gives that the site sensors3 and Sensor 4 are correlated and PC2 gives Senor 2 , sensor 4and Sensor5 are negatively correlated.

4.3 Conclusions:

With both Impacts assessment and Principal Component Analysis we can chose the site area precisely. From Impact Assessment results we can conclude that Monitoring stations are need to setup at these site maps.

From the results of Principal Component Analysis we can reduce number of monitoring sites by eliminating the monitoring sites which collect the redundant Information and there by the cost invested on the monitoring site setups can be reduced.

With Impact Assessment and Principal Component Analysis, it is possible to select the best Monitoring sites for analysis of the Pollutants. There by reducing the sensor setup costs.

REFERENCES

[1] Interpretation of water quality parameters for Tigris river by using principle component analysis by Mus'ab A.Al-tamir Mazien N. Al-sanjari, pp 1- 5.

[2] Jollife.Principal Component Analysis. Springer-Verlag, 1986

[3] Principal Component Analysis of Face Properties by Samarasena Buchala[1], Neil Davey[1], Tim Gale[1,2], Ray Frank[11] School of Computer Science, University of Hertfordshire, College Lane, Hatfield, AL10 9AB, UK[2] Department of Psychiatry, QEII Hospital, Welwyn Garden City, AL7 4HQ, UK),(Turk and Pentland (1991), pp 1-22.

[4] Principal components analysis to summarize microarray experiments: application to sporulation time series Soumya Raychaudhuri*, Joshua M. Stuart*, and Russ B. Altman*Stanford* Medical InformaticsStanford University, 251 Campus Drive, MSOB X-215, Stanford CA 94305-5479{sxr, stuart, altman} @smi.stanford.edu), pp 1-5.

[5] M.J. Black and A.D. Jepson. Eigen tracking: Robust matching and tracking of articulated objects using a view-based representation. In *ECCV*, pages 329–342, 1996

[6] M. Turk and A. Pentland. Face recognition using eigenface. In *CVPR*, pages 586– 591, 1991.

[7] T.F. Cootes, G.J. Edwards, and C.J. Taylor. Active appearance models. In *ECCV*, pages II 484–498, 1998.

[8] H. Murase and S.K. Nayar. Visual learning and recognition of 3d objects from appearance. *IJCV*, 14(1):5–24, June 1995, pp 1-10.

[9] R.G.Schroeder and T.N. Lahn. Development of a manufacturing strategy: A proven process. In: J.E. Ettlie, M.C. Burstein and A. Fiegenbaum (eds.), Manufacturing Strategy. Kluwer, Boston, MA, 1990 (RGS), pp 1-5

[10] S.H.Weissman and J.C.Sekutowski. Environmentally conscious manufacturing. A T& T Technical Journal, (1991), pp 1-5.

[11] LCA, uncertainties can be evaluated through Monte Carlo and interval analysis (Guinée et al., 2002 J.B. Guinée, M. Gorree and R. Heijungs, Handbook on Life Cycle Assessment an Operational Guide to the ISO Standard, Kluwer Academic, London (2002).[Guinée et al., 2002]

[12] http://www.epa.gov/ttn/oarpg/naaqsfin/pmfact.html

[13] http://pubweb.epa.gov/air/urbanair/nox/noxfldr.pdfa.

[14] Alan Gilpin (1995) Environmental Impact Assessment-Cutting Edge for the twenty-first century, Cambridge University Press. pp 1-10

[15] Principle component analysis of urban traffic characteristics and meteorological data S.M. Shiva Nagendra, Mukesh Khare pp 1-25

[16] Parallel analysis: method for determining significant principal components by Franklin,Scott B. Gibson ,David J. Robertson, Philip A., Pohlmann, jhonT.& Fralish, James S.

[17] Wall, Michael E., Andreas Rechtsteiner, Luis M. Rocha."Singular value decomposition and principal component analysis". in *A Practical Approach to Microarray Data Analysis*. D.P. Berrar, W. Dubitzky, M. Granzow, eds. pp. 91-109, Kluwer: Norwell, MA (2003). LANL LA-UR-02-4001

[18] Methods of Multivariate Analysis second edition by Alvin C. Rencher, pp 1-22

[19] Applied Multivariate Research By Lawrence S. Meyers, Glenn Gamst, A. J. Guarino

[20] Supervised principal component analysis for gene set enrichment of microarray data with continuous or survival outcomes xi chen [1,*,†], lily wang [2,†], jonathan d. smith [3] and bing zhang[4]

[21] Principal component analysis of gene expression profiles Leif E. Peterson, Ph.D. Dept. of Medicine; Dept. of Molec. & Human Genetics Baylor College of Medicine Chock et al (1975) Roch and pellerin (1982), pp 10-15.

[22] Principle component analysis of urban traffic characteristics and metrological data by Nagendra, Mukesh Khare, pp 15-18.

[23] Process systems engineering identification of faulty sensors using principal component analysis by Ricardo Dunia [1], S. Joe Qin [2 *], Thomas F. Edgar [2], Thomas J. McAvoy [3] **1**Fisher-Rosemount Systems, Austin, TX 78754 [2]Dept. of Chemical Engineering, University of Texas at Austin, Austin, TX 78712 [3]Dept. of Chemical Engineering, University of Maryland, College Park, MD 20742 Correspondence to S. Joe Qin, Dept. of Chemical Engineering, University of Texas at Austin, Austin, TX 78712

[24] Face Recognition Using Kernel Principal Component Analysis Kwang In Kim, Keechul Jung, and Hang Joon Kim, pp 12-15.

[25] M. Black and A. Jepson. Eigen tracking: Robust matching and tracking of objects using view-based representation. *ECCV*, pp. 329–342, 1996.

[26] B. Moghaddam and A. Pentland. Probabilistic visual learning for object detection. *ICCV*, 1995.

[27] T. Cootes, G. Edwards, and C. Taylor. Active appearance models. *5th ECCV*, 1998.

[28] N. Oliver, B. Rosario, and A. Pentland. A Bayesian computer vision system for modeling human interactions. *ICVS. Gran Canaria, Spain*, Jan. 1999.pp 1-3.

[29] Principal component analysis model for machine-part cell formation problem in group technology Wafik Hachichaa , Faouzi Masmoudi a,b and Mohamed Haddar a,b

a Unité de recherche de Mécanique, Modélisation et Production,(U2MP).

b Département de génie mécanique, Ecole Nationale d ingénieurs de Sfax, B.P. W, 3038 Sfax, Tunisia

[30] Online Banking performance evaluation using data envelopment and PCA by ChienTaBruceHoa,DeshengDashWub,aInstituteofElectronicCommerce,NationalChungHsingUniversity,Taiwan b Risk Lab, UniversityofToronto,19BordenST,Toronto,ON,Canada.pp 5-15

[31] Variable selection in large Environmental data sets using Principal Component Analysis by Jacquelynne R. King and Donald A. Jackson. PP 16-17

[32] Fault Detection And Isolation With Robust Principal Component Analysis By Yvon Tharrault, Gilles Mourot, José Ragot, Didier Maquin Centre De Recherche En Automatique De Nancy (Cran) Umr 7039, Nancy Université, Cnrs 2, Avenue De La Forêt De Haye, F-54 516 Vandoeuvre-Lès-Nancy, France.

[33] Process fault detection and diagnosis based on principal component analysis tao he1, 2, 3, wei-rong xie1, 2, qing-hua wu1, 2, 3, tie-lin shi1, 2, 3 1sch. Of mech. Engin., hubei univ. Of tech., wuhan 430068, china 2hubei key lab of modern manufacture quality engineering, wuhan 430068, china 3sch. Of mech. Sci. & engin., huazhong univ. Of sci. & tech., wuhan 430074, china e-mail: hetao_wh@163.com, kinghuawu@sohu.com

[34] Neural Network for Principal Component Analysis with Applications in Image compression Luminita STATE Dept. of Computer Science, University of Pitesti, Pitesti, Romania Catalina Lucia Cocianu Dept. of Computer Science, Academy of Economic Studies, Bucharest, Romania.pp 21-22

[35] Principle component analysis with Missing Data and outliers by Haifeng Chen electrical and computer engineering Department. pp 20-24

[36] Legendre, P., and L. Legendre, 1998.Numerical Ecology. Elsevier: Amsterdam, 853 p. Swan, A.R.H., and M. Sandilands, 1995.Introduction to Geological Data Analysis. Blackwell Science: Oxford, 446 p.

[37] A Tutorial on Principle Component Analysis by Lindsay I. Smith (Cornell University) www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf The ENVI software manual Introduction to Linear Algebra (Chapter on Eigenvalues) some textbooks on Remote Sensing have short Chapters on PCA

[38] Are global warming and economic growth compatible? Evidence from five OPEC countries Ramazan Sari *, Ugur Soytas1 (1) . pp 1-17.

[39] Global Warming and green house gas effect by Kevin Gurney.

[40] Atmospheric increasing and sinks and reservoirs from Intergovernmental Panel on climate change, Climate Change 1995 (Cambridge university press 1996).

[41] http://www.global-greenhouse-warming.com/greenhouse-gas.html.

[42] Explicit calculation of indirect global warming potentials for halons using atmospheric models by D. Youn1, 2, K. O.Patten1, J.-T. Lin 3 and D. J. Wuebbles1

[43] Bradford, D. F.: 2001, 'Time, Money, and Tradeoffs', Nature **410**, 649–650.

[44] Godal O. and Fuglestdedt, J.: 2002, 'Testing 100-Year Global Warming Potentials: Impacts on Compliance Costs and Abatement Profile Climate Change **52** (1), 93–127.

[45] Godal O: 2003, 'The IPCC's Assessment of Multidisciplinary Issues: The Case of Greenhouse Gas Indices Climate Change **58**, 243–249.

[46] A Tutorial on Principal Component Analysis by Jonathon Shlens_Systems Neurobiology Laboratory, Salk Insitute for Biological Studies La Jolla, CA 92037 andInstitute for Nonlinear Science, University of California, San Diego La Jolla, CA 92093-0402

[47] A tutorial on Principal Components Analysis by Lindsay I Smith pp

[48] Multivariate Statistical Analysis by Johnson, Richard Arnold .Pearson Prentice Hall, c2007

CURRICULUM VITAE

Ugandhar Reddy Kondamadugula was born on August 5, 1985 in India. He got his bachelor's degree in Mechanical Production and Industrial Engineering in April 2006 from Andhra University, Andhra Pradesh, India. During these years, he has been an active participant in Industrial trainings. After graduation he worked as an Engineer trainee in Afcons Infrastructure Ltd. In spring 2007 he came to The University of Texas at El paso to pursue his MS in Industrial Engineering. During his course of study he worked as research assistant under Dr. Bill Tseng. He also worked as an Industrial Engineer at Arcelor Mittal Company. He is honor student and member of the Industrial Engineering honor society Alpha Pi Mu. He is enthusiastic in pursuing a career in industrial engineering.

Permanent Address: MIG 32 Sector- 1,

M.v.p Colony, Visakhapatnam 530017,

Andhra Pradesh, India.