CHAPTER 4

Results and Discussion

In this chapter investigation about the performance of various nose sensors in classifying two different types of sample used in pharmaceutical and in petroleum industry has been discussed.

4.1 Introduction

The analysis has been performed in three steps that can be broadly defined to classify the implementation of work. These three broad categories are explained below:

**Step 1: Simulation modeling using simulink:** MATLAB/Simulink have been used in this step to design sensor model with consideration of humidity and temperature.

**Step 2 Classification algorithm performance analysis using simulated model:** In this step the above mentioned sensor model is used to design the sensor array. The generated signal of these sensor array are used further to classify mixture of two gases using PCA (Section 3.2.4) based classification analysis. During classification the efficiency of PCA classification have been checked over the different signal preprocessing technique (Section 3.2.4).
Step 3 Applications and comparison of PCA and PLS technique on TGS and MICS sensors: After justifying the capability of PCA classification, preprocessing the algorithm have been used on the real time applications instead of using simulated sensor outputs. For this purpose the data have been collected from two different scenarios having applications related to medical herb Glycyrrhiza Glabra and another applications based on classification of organic compound named as ethanol, acetone and propane.

Above mentioned steps are described in further section of this chapter consequently along with the results for each step that can describe the variation of sensor response and its utilization in classifying the desired type of sample. In step 3 two types of technique for multivariate analysis classification have been used, these technique are PCA and PLS. The description of data and sensors used is also provided in the appendix A and B that is used during the development of algorithm.

4.2 Development in Simulation of Sensor Models

Sensors, based on conducting polymer resistors, are very attractive for vapour/odour sensing applications because of the wide range of available polymer combinations and their ease of deposition, their ability to operate at room temperature (i.e. power consumption of small device), and compassion to a wide range of unstable organic compounds. A polymeric chemoresistor is proposed as novel parametric model for the simulation of a smart gas sensor. The resistive principle, in which a change in the sensor resistance $\Delta R_S$ is monitored while the sensor is exposed to the gas, is the most commonly used principle within the field of vapour/odour sensing. Here the model for the simulation and design of resistive sensors employing carbon-black polymer composite films as the set of gas
responsive material has been described. The batch to-batch variation in baseline resistance and its large temperature and humidity coefficients are some disadvantages commonly associated with polymeric chemoresistors. For instance, transient and dynamic sensor responses $[2,115–118]$ can be monitored and used to extract information that can improve the gas recognition performance. Development of parametric models for polymeric chemoresistors can help in the design of new devices with improved characteristics and in the interpretation of experimental data.

4.2.1 Model Description

A parametric model of a polymeric chemoresistor developed for use in the design and simulation of smart gas/odor sensor systems is being presented in this thesis. The Cadence software is used for implementation of this model, thus allowing both the implementation of resistive elements in smart sensor design and the simulation of the chemical static response and chemical step response of a polymeric chemoresistor to a mixture of different gases. The temperature and humidity effects are also taken into account in this model and simulate the noise present in polymer sensors, such as flicker or $1/f$ noise $[126–128]$ and Johnson $[125]$. Sample delivery system (SDS) parameters, such as the volumetric flow rate and volume of the sensor chamber that can be customized to specific experimental set-ups are also taken into account in the new model. To demonstrate the practical application of the new model in the design of polymeric chemoresistors and to simulate their behaviour Polymer-carbon black composite films are used here as gas sensitive materials. Lewis at Caltech firstly reported the films, which consist of conducting black nano carbon spheres dispersed into a non-conducting polymer base film. The polymer within the composite film absorbs the vapour and swells reversibly when exposed to gases. This swelling causes the
distances between the conductive carbon black particles to increase and thus induces a resistance change in the composite film [109,111]. Arrays of several sensors made up by different polymer composite films have the pattern response of a ‘fingerprint’ to classify different gas or chemical mixture [2,113].

![Figure 4.1: Basic block diagram of E Nose](image)

### 4.2.2 Sensor Simulation Model Design and Results Validation:

After verification of sensor simulation model which is described in this section it was used to simulate the behavior of polymer composite sensors when exposed to a single gas or mixture of gases like carbon black polymer. The complete cell comprising a chemoresistors model, schematic representation and a layout design was used. In this simulation a d.c. voltage of 2.4 V is applied to the input while the output signal is the voltage difference across the sensor terminal R1 & R2. To convert the sensor resistance into a voltage that can be further amplified, processed or interfaced to other devices, some gas sensor manufacturers recommend to use a voltage...
divider configuration. The variation in resistance of the gas-sensitive device (typically 1–50%), subsequent to exposure to a gas of concentration $C_{gas}$, can then be measured through the change in the output voltage $V_{out}$ which is a function of the sensor resistance $R_s$.

\[ V_{out} = V_{ref} \cdot \frac{R_s(C)}{R_{ref}+R_s(C)} \]  
\[ \text{4.1} \]

The sensitivity $S$ of the potential divider is defined as

\[ S = \frac{dV_{out}}{dR_s} = \frac{V_{ref} \cdot R_{ref}}{(R_{ref}+R_s(C))^2} \]  
\[ \text{4.2} \]

Which has a maximum for $R_{ref}=R_s$. The reference resistance value was chosen to be equal to the sensor baseline value (10 kΩ) to maximize sensitivity for small changes of $R_s$. Gases injection and sensor temperature are simulated by voltage sources.

**Figure 4.2: Sensor Simulation Model**

In figure 4.2 it is being shown an overview of the developed simulation model and its detail description using the simulink tool is given below. In this thesis it is being considered that the voltage response of sensor varies with change in the sensor resistance in presence of a gas. This variation
is written here by following generic power law function followed in most of the sensors having polymer composite film:

\[ R = R_O[1+KC\exp(K_S/T)] \]…………………4.3

Where \( R_O \) is the baseline sensor resistance (measured in the presence of a reference gas, generally clean air), \( k \) is a sensitivity coefficient, \( C \) is the gas concentration expressed in ppm, gamma is the power law exponent, \( K_S \) is a temperature coefficient and \( T \) is the temperature in degrees Kelvin. The sensitivity coefficient can be positive or negative depending on the nature of the gas and the polymer used, producing an increase or decrease of the sensor resistance after the gas is introduced. It should be noticed that an increase in temperature results in a reduction of the sensor resistance when \( K_S \) is positive. The above resistance model can be extended to a gas mixture by adding the effect of the individual component as separate inputs assuming that there is no interaction between them. This assumption is valid for low concentrations of volatile organic compounds [129,130]. The resistance model can also be expanded to include the independent additive effect of humidity:

\[ R=R_O[1+K_GC_G\exp(K_{SG}/T)+K_HC_H\exp(K_{SH}/T)] \]…………………………4.4

Where the subscripts \( G \) and \( H \) corresponds to the gas and the water vapour respectively. The above model is implemented in Simulink through a combination of functional blocks, performing basic operations (e.g. adder block, exponential block). This is represented in the schematic view shown in Fig.4.3. The same result could be achieved by the definition of the model using a Verilog-A or an HDL script [131,132].
Figure 4.3: Simulation Model for Gas Sensor Response
4.2.3 Simulation Results

In the figure 4.4 shown above we can see the response of dynamic resistance and static resistance when humidity is at 0Kppm and Temperature is kept at 40°C.

In the figure 4.5 shown below it can be seen that the dynamic resistance changes with respect to humidity at constant temperature. (H=0 and 10 Kppm, T=40°C)
Figure 4.5 Change in Rdyn w.r.t humidity at constant temperature T=40°C

Figure 4.6: Change in Rdyn w.r.t humidity at constant temperature T=50°C
Figure 4.7: Change in Rdyn w.r.t temperature at humidity H=0 K ppm

Figure 4.8: Change in Rdyn w.r.t temperature at humidity H=10 K ppm
In the fig 4.6 it can be seen that Rdyn not changes at constant temperature 50°C even the humidity changes from 0 to 10 K ppm. In fig 4.7 it can be seen that the Rdyn changes at different temperature and at constant humidity (T=40°C & 50°C, H=0 Kppm) whereas in fig 4.8 it can be seen that when H=10 K ppm and temperature varies from 40°C to 50°C then also Rdyn changes.

The above figures shows change in Dynamic Resistance without sample delivery system (SDS) at constant H and different T.

In the above figure 4.9 it can be seen that with sample delivery system the dynamic resistance changes with respect to humidity and at constant temperature. (H=0 and 10 K ppm, T=40°C).
Whereas in figure 4.10 it can be seen that the $R_{\text{dyn}}$ changes when temperature is kept constant at 40°C and humidity changes from 0 to 10 Kppm. In the above figure it can be seen that with sample delivery system the dynamic resistance changes largely with respect to temperature and at constant humidity. Figure 4.11 has (H=0 K ppm T=40°C & 50°C). Whereas in figure 4.12 it can be seen that the $R_{\text{dyn}}$ changes when humidity is kept constant at 10 Kppm and temperature changes from 40°C to 50°C. These results show that variation in R is much sensitive to change in temperature and it shows very small variation with respect to change in humidity.

The chemoresistor sensor used here are Carbon black/poly vinyl pyrrolidone (PVP) composite (20 wt % carbon black) (Aldrich). For experimental validation test was carried out in a closed gas station with independent control of the concentrations of water and methanol vapors.
Figure 4.11: Change in Dynamic Resistance with SDS at constant H and different T. 
H=0 K ppm, T=40° & 50° C

Figure 4.12: Change in Dynamic Resistance with SDS at constant H and different T. 
H=10 K ppm, T=40° & 50° C
4.3 Use of Array of Sensors

Electronic nose systems utilize a series of sensors to provide a fingerprint response to a specified odor, and prototype recognition software then performs odor recognition and discrimination. The electronic nose is a price-effective solution to the problems connected with sensory panels and by chromatographic and mass-spectrometric method and can put up real time presentation in the meadow when implemented in moveable form. Continuous real monitoring of odor is done at specific sites in the field over hours, days, weeks or even months. An electronic machine can also avoid many other troubles linked with the employ of human panels. Each and every variability, adaptation (becoming minimum sensitive during extended exposure), exhaustion, infections, mental condition, subjectivity, and revelation to hazardous compounds all come to mind.

Figure 4.13 (a) shows sensor A, B, C and D are designed at different parameters. Figure 4.13(b) shows the block diagram of the designed sensor in which the sensor parameter used for feature extraction is sensor element conductance Gs. Feature evaluation Gs is evaluated using following equation:

\[
G_s = G_0^{T} e^{-\frac{E_A^{0} K}{T}} + K_1^{T} e^{-\frac{E_A^{1} K}{T}} C_1^{n_1 K T} + K_2^{T} e^{-\frac{E_A^{2} K}{T}} C_2^{n_2 K T} + \ldots K_{mix}^{T} e^{-\frac{E_A^{mix} K}{T}} C_1^{n_1 K T} C_2^{n_2 K T}
\] 4.5

Where \( K = \) Boltzmann’s constant, \( T = \) Absolute Temperature, \( C_1 \) & \( C_2 = \) Concentration of gas1 & gas2, \( K_1 T \) &\( K_2 T = \) pre exponential factors of species 1 & 2, \( n_1 \) & \( n_2 = \) pre factors of power law exponent for oxides, \( K_{mix} = \) pre exponential factor of the change in conductance caused by the interaction of the two species, \( E_A = \) activation energies
The response model described by equation no 4.5 assumes a quasi linear interaction between the gases in the binary mixture, which is the case when gases are chemically independent. Block diagrams of the e-nose simulation tool and a general PSPICE sensor replica are shown in fig. 4.14. The experimental data set contains calibration data which were collected as a piece of earlier study that developed a home appliance for measuring methane in the company of ethanol [133].
Figure 4.13(b) Simulink model for mixture of two gases for each sensor
These facts were used to regulate the values of the parameters of the MATLAB sensor models, by applying an average fitting process in Matlab. Table 4.1 shows the values of the stationary model parameters for the four tin oxide sensors. Formerly the sensor parameters of models were set and the simulation process was as follows.
Table 4.1: Values of Simulink static model parameter for the four Tin oxide gas sensors in the presence of ethanol, methane and their binary mixtures

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Sensor 1</th>
<th>Sensor 2</th>
<th>Sensor 3</th>
<th>Sensor 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_0$</td>
<td>0.16</td>
<td>0.32</td>
<td>0.12</td>
<td>0.23</td>
</tr>
<tr>
<td>$E_{a0}$</td>
<td>5*10^-3</td>
<td>4.3*10^-3</td>
<td>3.9*10^-3</td>
<td>5.6*10^-3</td>
</tr>
<tr>
<td>$K_1$</td>
<td>6.14*10^-4</td>
<td>3.14*10^-3</td>
<td>8.3*10^-3</td>
<td>1.7*10^-3</td>
</tr>
<tr>
<td>$EA_1$</td>
<td>5.31*10^-3</td>
<td>2.31*10^-2</td>
<td>2.1*10^-2</td>
<td>2.23*10^-2</td>
</tr>
<tr>
<td>$n_1$</td>
<td>13.95</td>
<td>11.03</td>
<td>5.97</td>
<td>10.26</td>
</tr>
<tr>
<td>$K_2$</td>
<td>9.8*10^-5</td>
<td>2.8*10^-4</td>
<td>1.6*10^-4</td>
<td>1.72*10^-3</td>
</tr>
<tr>
<td>$EA_2$</td>
<td>6.2*10^-3</td>
<td>9.21*10^-3</td>
<td>3.11*10^-3</td>
<td>1.02*10^-3</td>
</tr>
<tr>
<td>$n_2$</td>
<td>6.83</td>
<td>7.55</td>
<td>6.54</td>
<td>4.05</td>
</tr>
<tr>
<td>$K_{mix}$</td>
<td>-5.2*10^-6</td>
<td>-1.2*10^-6</td>
<td>-2.5*10^-5</td>
<td>-1*10^-5</td>
</tr>
<tr>
<td>$EA_{mix}$</td>
<td>3.31*10^-3</td>
<td>2.87*10^-3</td>
<td>4.35*10^-2</td>
<td>2.19*10^-3</td>
</tr>
</tbody>
</table>

1. Random and systematic errors were applied to change the values of the parameters in the PSpice models. Random errors were unspecified to have a Gaussian distribution.
2. By iteratively executing PSpice simulations by changed model parameters a huge response data set was achieved. Characteristic were refined from the responsive sensor.

3. These properties were key in to a PARC identification engine (Principal Component Analysis), which was simulated by MATLAB as to discuss system performance with relation to the electronic nose discrimination skill.

Two different sample delivery system Cgas1 (Ethanol) & Cgas2 (Methane) has been designed. Both samples are passed through all four sensors. The duration of sample contact is taken as 500 sec to 1500 sec.

![Figure 4.15 Gas Delivery System](image-url)
Temperature and humidity is controlled as shown in fig 4.17. Temperature is kept at 50°C and randomly changed as:

\[ V_h = V_h (1 + K_h \alpha_h) \]………………..4.5

Where \( V_h \) indicates heating voltage, \( K \) deviation and alpha random value. Similar to variation in T we also randomly input the humidity.

Figure 4.16 Response of Sensor A in presence of both gases
The response of sensor gives different variations in delG. These responses are for three different cases.

Case 1: In presence of ethanol ie Cg1=10 ppm and Cg2=0 ppm.

Case 2: In presence of methane ie Cg1=0 ppm and Cg2=100 ppm.

Case 3: In presence of both ethanol and methane ie Cg1=10 ppm and Cg2=100 ppm.

The concentration of gas varied by following relation:

\[ C_{\text{eth}}' = C_{\text{eth}} (1 + K_{\text{eth}} \alpha_{\text{eth}}) \] \[ C_{\text{met}}' = C_{\text{met}} (1 + K_{\text{met}} \alpha_{\text{met}}) \]

K=0.1 and alpha random value with deviation 1.
For all 3 cases all the 8 parameters of temperature and humidity are varied.

These parameters are: Go, Eao, Kn, EA1, n1, K2, EA2, n2.

Hence 10x3 response for training the classifier model using value of delGoA, delGoB, delGoC & delGoD has been generated and the response of delG for case one for error model with variation in temperature is shown in fig 4.18.

Figure 4.18: Response of delG for case one for error model with variation in temp.
Figure 4.19 Response of delG for case one for error model with variation in humidity.

Figure 4.20 Response of delG for case one for error model with variation in KsH.
Figure 4.21 Cluster Plot obtained for 3 cases for 10 error models at random error of 10%.

Figure 4.22 Cluster Plot obtained for 3 cases for 10 error models at random error of 20%.
Figure 4.23 Change in gas concentration (Blue methane, green ethanol)

Figure 4.24 Response of sensors due to change in gas concentration.

Figure 4.25 Cluster Plot obtained for 3 cases for 10 error models on varying dynamic response coeff. (L, C and R) by 10%
From the above results shown in fig. 4.19, 4.20, 4.21, 4.22, 4.23, 4.24, 4.25, 4.26 it is shown in the cluster plot that even we add 10% and 20% of error in all 10 parameters even though the clusters for all three cases are easily discriminable. But in case of 20% deviation of parameter model dispersion has increased in principal component values. In the case of variation in dynamic response coefficient, the dispersion of cluster is increased in both 10% and 20% of deviation in parameters but still the clusters were discriminable.

In the meanwhile, mandarin fruit ripeness procedure on the tree, the respiration lessens, meaning the generated vapours gets lower down, which fumes reach in a fewer quantity, the mean signal of sensor series decrease. Vapour associated with fruit also contains methane. Above sensor array is sensitive to methane among the presence of different gas hence the
lowering in response can be taken as signal for picking time of mandarin fruit. Presence of methane indicates the natural gas in the coal mines and petroleum industries accidental leakage of gases can be easily detected as rise in conductance of sensor response. Ethanol is present in alcoholic beverages hence this sensor array can be used to check the drivers and other people using respiration or during breathing. Beverages industries check the Ethanol contents for quality measurement of wines etc.

4.4 Applications and comparison of PCA and PLS technique on TGS and MICS sensors

This is the final part of the thesis, in this part real time data have been considered which is recorded by international laboratories with high quality standard. There are two different kinds of data base D1 and D2. Their brief description is given below:

D1: It consists MATLAB structure array having file named as data, it consists of sensor voltage response in an array having size 18*241*12. It means there are 17 different kinds of samples taken at 241 times sample discrete interval for eleven different kind of MOS sensor. The plot of this data is shown in figure 4.27(a) where X axis correspond to the time 0-240 samples. Each colour band represents a different sensor. The data is recorded for three different kinds of glycyrrhiza glabra herbs odor. The type of glycyrrhiza glabra are named as good, over dried and bad. Where over dried samples are the sample of glycyrrhiza glabra herb that are kept at longer time at high temperature for removing moisture and drying of the herb. The bad samples are those samples that have complaints in their test similar to a burned taste of glycyrrhiza glabra.
D2: This data set consists of data recorded by MOS gas sensors array collected from website www.mrpt.org. It has data record of seven sensors recorded for 300-400 sec for different types of organic compounds named as acetone, ethanol and propane. The response curve of these data record is shown in figure 4.27 (b).

Figure 4.27 (a): Time response of Twelve sensors, all based on Metal Oxide Semiconductor (MOS) technologies for three different glycyrrhiza glabra samples each having six data records for data base D1.
4.4.1 Classification using PCA Analysis for quality assessment of Licorice data base D1:

The algorithms consist of concepts related to pattern recognition technique. It requires some particular type of scaling in order to get proper results. If scaling is not performed, it is desirable to apply the normalization of data due to the different nature of some parameters (humidity, temperature, conductance, etc). Two linear methods PCA and PLS with scaling and normalization have been used. PCA and PLS give different plots by using them one can get measurements in a form of cluster groups. These cluster scattering and their area boundary helps the electronic nose by providing enough resolution to discriminate samples and in determining how variables related to each other.
The clusters have to be manually identified and subjectively conclusions are drawn over the efficiency of classification algorithm. Prior to PCA/ PLS classification we have applied three different kinds of preprocessing technique using the voltage values obtained at different instance. From the voltage response of each gas sensor, many parameters are extracted using usual static parameters include: initial voltage \( (V_i) \), final voltage \( (V_f) \), Voltage increment \( \Delta V = V_f - V_i \) and normalized voltage increment \( V_n = (V_f - V_i) / V_i \).

4.4.2 PCA Analysis of D1 data base using data difference preprocessing:

In this section PCA has been applied to the D1 data base having 12 MOS sensors and 6*3 odor response of *Glycyrrhiza glabra* representing bad, good and fabricated bad groups of data sample. Figure 4.28 is the loading plot for the sensors performance clusters for considering difference in the readings of voltage response (\( \Delta V \)). This plot has 12 different scattering points with numbering 1 to 12.

![Figure 4.28 Loading plot of sensors classification using PCA for \( \Delta V \) preprocessing](image-url)
In this plot sensor \{2\}, \{3\} has totally different performance and \{1,6,8,9,10,11\},\{4,5\} are responding in a similar manner. The loading plot represents correlation between the variables and score plot represent the correlation between the observations. In this way for considering difference data preprocessing the response of sensor set \{1,6,8,9,10,11\} are correlated or similar same analysis is also true for sensor set \{4,5\}. The response of sensor \{3\} opposite of all other sensors and most significant role in classification is dominated by set \{7\}, \{11\}, \{12\}.

Figure 4.29 shows the score plot for analyzing the relation between observations of data difference. In this plot the clusters related to a particular kind of sensor has been bounded. There are three categories of good quality (good- green boundary), fabricated bad (FBAD- red boundary) and bad (BAD- Purple boundary). From the figure it can be clearly seen that PCA is able to discriminate all these three clusters by separate boundary. Hence data difference PCA classification is performing well for detecting three different kinds of sample qualities.

Figure 4.29 Score plot of sensors classification using PCA for $\Delta V$ preprocessing
Figure 4.30 Bi plot of sensors classification using PCA for ΔV preprocessing

Figure 4.30 consist of Bi plot which shows both score and loading plot together. In this plot both observations and variables scatter plot are displayed. The cluster on the right side are the sensor principal components where points below the origin is sensor 3 and remaining sensor are above origin and the observation of Good, FBAD and BAD samples are on the left side.

4.4.3 PCA Analysis of D1 data base using data fractional preprocessing:

In this section PCA has been applied to the D1 data base having 12 MOS sensors and 6*3 odor response of *Glycyrrhiza glabra* representing bad, good and fabricated bad groups of data sample. Figure 4.31 is the loading plot for the sensors performance clusters for considering fraction in the readings of voltage response. This plot has 12 different scattering points with numbering 1 to 12.
Figure 4.31 Loading plot of sensors classification using PCA for fractional preprocessing

In this plot sensor {1}, {2}, {7}, {9}, {11}, {12} has totally different performance and {3,6}, {4,5}, {8,10} are responding in a similar manner. The loading plot represents correlation between the variables and score plot represent the correlation between the observations. In this way for considering fractional data preprocessing the response of sensor set {3,6} are correlated or similar and same analysis is also true for sensor set {4,5}, {8,10}. The response of sensor {7,8,9,10,11,12} are opposite of all other sensors i.e 2,3,4,5 and 6 and most significant role in classification is dominated by sensor set{12}, {7}, {11}.

Figure 4.32 shows the score plot for analyzing the relation between observations of data fraction. In this plot the clusters related to a particular kind of sensor are being bounded. There are three categories of good quality (good- green boundary), fabricated bad (FBAD- red boundary) and bad (BAD- Purple boundary). From the figure it can be clearly seen that PCA is able to discriminate two clusters by separate boundary it is good and bad but clusters of FBAD are widely scattered and they are overlapping with
cluster of BAD quality. Hence data fraction PCA classification is performing well for detecting different GOOD and BAD kinds of sample qualities but not able to discriminate BAD and FBAD sample qualities.

Figure 4.32 Score plot of sensors classification using PCA for fraction preprocessing

Figure 4.33 Bi plot of sensors classification using PCA for fraction preprocessing
Figure 4.33 consist of Bi plot which shows both score and loading plot together. In this plot both observations and variables scatter plot are displayed. The cluster on the right side are the sensor principal components where points right of the origin is sensor 7 to 12 and remaining sensor are left of the origin and the observation of Good, FBAD and BAD samples are on the left side. In the Bi plot the sensor component are widely scattered as they do not giving a similar performance for fractional data and the observation point are too much aggregated due to this it is tough to classify the odor by fractional method.

**4.4.4 PCA Analysis of D1 data base using data relative preprocessing:**

In this section PCA is being applied to the D1 data base having 12 MOS sensors and 6*3 odor response of *Glycyrrhiza glabra* representing bad, good and fabricated bad groups of data sample. Figure 4.34 is the loading plot for the sensors performance clusters for considering relative in the readings of voltage response. This plot has 12 different scattering points with numbering 1 to 12.

In this plot sensor {2} has slightly different performance and {3,6},{4,5}, {1,7,8,9,10,11,12}are responding in a similar manner. The loading plot represents correlation between the variables and score plot represent the correlation between the observations. In this way, for considering relative data preprocessing the response of sensor set{3,6} are correlated or similar and same analysis is also true for sensor set {4,5},{1,7,8,9,10,11,12}. The response of sensor {1,7,8,9,10,11,12} are opposite of all other sensors and most significant role in classification is dominated by set {7,8,9,10,11,12}. 
Figure 4.34 Loading plot of sensors classification using PCA for relative preprocessing

Figure 4.35 shows the score plot for analyzing the relation between observations of data relative. In this plot the clusters related to a particular kind of sensor has been bounded. There are three categories of good quality (good- green boundary), fabricated bad (FBAD- red boundary) and bad (BAD- Purple boundary). From the figure it can be clearly seen that PCA is able to discriminate two clusters by separate boundary it is good and bad but clusters of FBAD are widely scattered and they are overlapping with both cluster of GOOD and BAD quality. Hence data relative PCA classification is not performing well for detecting different GOOD and FBAD kinds of sample qualities and also not able to discriminate BAD and FBAD sample qualities.
Figure 4.35 Score plot of sensors classification using PCA for relative preprocessing

Figure 4.36 Bi plot of sensors classification using PCA for relative preprocessing

Figure 4.36 consist of Bi plot which shows both score and loading plot together. In this plot both observations and variables scatter plot are displayed. The cluster on the right side are the sensor principal components where points right of the origin is sensor 7 to 12 and remaining sensor are left of the origin and the observation of Good, FBAD and BAD
samples are on the left side. In the Bi plot the sensor component are widely scattered as they do not giving a similar performance for relative data and the observation point are too much aggregated due to this it is tough to classify the odor by relative method.

4.4.5 PLS Analysis of D1 data base using data difference preprocessing:

In this section PLS technique is applied to the D1 data base having 12 MOS sensors and 6*3 odor response of *Glycyrrhiza glabra* representing bad, good and fabricated bad groups of data sample. Figure 4.37 is the loading plot for the sensors performance clusters for considering difference in the readings of voltage response (ΔV). This plot has 12 different scattering points with numbering 1 to 12.

![Loading plot of sensors classification using PLS for difference (ΔV) preprocessing](image-url)
In this plot sensor \{2\},\{3\} has totally different performance and \{1,6,8,9,10,11\} are responding in a similar manner as \{4,5\}, \{7,12\}. The loading plot represents correlation between the variables and score plot represent the correlation between the observations. In this way for considering difference data preprocessing the response of sensor set \{1,6,8,9,10,11\} are correlated or similar and same analysis is also true for sensor set \{4,5\}, \{7,12\}. The response of sensor \{3\} are opposite of all other sensors and most significant role in classification is dominated by set \{3\},\{5\},\{4\}.

Figure 4.38 shows the score plot for analyzing the relation between observations of data relative. The clusters related to a particular kind of sensor are being bounded in this plot. There are three categories of good quality (good- green boundary), fabricated bad (FBAD- red boundary) and bad (BAD- Purple boundary). From the figure it can be clearly seen that PLS is able to discriminate all these three clusters by separate boundary. Hence data difference PLS classification is performing well for detecting three different kinds of sample qualities i.e. GOOD, BAD and FBAD.

Figure 4.39 consist of Bi plot which shows both score and loading plot together. In this plot both observations and variables scatter plot are displayed. The cluster on the left side are the sensor principal components where points below the origin is sensor 3 and remaining sensor are above origin and the observation of GOOD, FBAD and BAD samples are on the right side.
Figure 4.38 Score plot of sensors classification using PLS for differential preprocessing

Figure 4.39 Bi plot of sensors classification using PLS for differential preprocessing
4.4.6 PLS Analysis of D1 data base using data fraction preprocessing:

In this section PLS is applied to the D1 data base having 12 MOS sensors and 6*3 odor response of *Glycyrrhiza glabra* representing bad, good and fabricated bad groups of data sample. Figure 4.40 is the loading plot for the sensors performance clusters for considering difference in the readings of voltage response (ΔV). This plot has 12 different scattering points with numbering 1 to 12.

![Loading plot](image)

**Figure 4.40 Loading plot of sensors classification using PLS for fractional preprocessing**

In this plot sensor{2},{3},{6},{7},{9}, {11},{12} has totally different performance and sensor set {1,4,5} and sensor set{8,10} are responding in a similar manner. The loading plot represents correlation between the variables and score plot represent the correlation between the
observations. In this way for considering fractional data preprocessing the response of sensor set \(\{1,4,5\}\) are correlated or similar and same analysis is also true for sensor set \(\{8,10\}\). The response of sensor \(\{2,3,4,5,7,12\}\) are opposite of all other sensors and most significant role in classification is dominated by set \(\{12\},\{7\},\{11\}\).

Figure 4.41 shows the score plot for analyzing the relation between observations of data fraction. In this plot the clusters related to a particular kind of sensor has been bounded. There are three categories of good quality (GOOD- green boundary), fabricated bad (FBAD- red boundary) and bad (BAD- Purple boundary). From the figure it can be clearly seen that PLS is able to discriminate two clusters by separate boundary it is good and bad but clusters of FBAD are widely scattered and they are overlapping with cluster of BAD quality. Hence data fraction PLS classification is performing well only for detecting GOOD sample qualities but not able to discriminate BAD and FBAD sample qualities.

Figure 4.42 consist of Bi plot which shows both score and loading plot together. In this plot both observations and variables scatter plot are displayed. The cluster on the left side are the sensor principal components where points right of the origin is sensor 7 to 12 and remaining sensor are left of the origin and the observation of Good, FBAD and BAD samples are on the left side. In the Bi plot the sensor component are widely scattered as they do not give a similar performance for fractional data and the observation point are too much aggregated, due to this it is tough to classify the odor by fractional method.
Figure 4.41 Score plot of sensors classification using PLS for fractional preprocessing

Figure 4.42 Bi plot of sensors classification using PLS for fractional preprocessing
4.4.7 PLS Analysis of D1 data base using data relative preprocessing:

In this section PLS technique has been applied to the D1 data base having 12 MOS sensors and 6*3 odor response of Glycyrrhiza glabra representing bad, good and fabricated bad groups of data sample. Figure 4.43 is the loading plot for the sensors performance clusters for considering relative in the readings of voltage response (ΔV). This plot has 12 different scattering points with numbering 1 to 12.

![Loading plot](image)

**Figure 4.43 Loading plot of sensors classification using PLS for relative preprocessing**

In this plot sensor{2} has totally different performance and sensor set {1,7,8,9,10,11,12} and sensor set {3,6}, {4,5} are responding in a similar manner. The loading plot represents correlation between the variables and score plot represent the correlation between the observations. In this way for considering relative data preprocessing the response of sensor set
{1,7,8,9,10,11,12} are correlated or similar and same analysis is also true for sensor set {3,6}, {4,5}. The response of sensor {1,7,8,9,10,11,12} are opposite of all other sensors and most significant role in classification is dominated by set {6},{3},{4}.

Figure 4.44 shows the score plot for analyzing the relation between observations of data relative. In this plot no clusters related to a particular kind of sensor are made. There are three categories of good quality (GOOD), fabricated bad (FBAD) and bad (BAD). From the figure it can be clearly seen that PLS is unable to discriminate these clusters by separate boundary as the clusters are widely scattered and they are overlapping with cluster of each quality. Hence data relative PLS classification is not able to perform well in discriminating the three categories of GOOD, BAD and FBAD sample qualities.
Figure 4.45 Bi plot of sensors classification using PLS for relative preprocessing

Figure 4.45 consist of Bi plot which shows both score and loading plot together. In this plot both observations and variables scatter plot are displayed. The cluster on the left side are the sensor principal components where points right of the origin is sensor 2 to 6 and remaining sensor are left of the origin and the observation of Good, FBAD and BAD samples are on the right side. In the Bi plot the sensor component are widely scattered as they do not give a similar performance for relative data and the observation point are too much aggregated, due to this it is tough to classify the odor by relative method.
4.4.8 PCA Analysis of Organic Gases and Compounds D2 data base using data difference preprocessing:

In this section PCA is applied to the organic gases and compounds D2 data base having 6 MOS sensors and 3*3 odor response of acetone (A), ethanol (E) and propane (P) representing bad, good and fabricated bad groups of data sample. Figure 4.46 is the loading plot for the sensors performance clusters for considering difference in the readings of voltage response. This plot has 6 different scattering points with numbering 1 to 6.

![Loading plot](image-url)
In this plot sensor \{1\}, \{2\}, \{3\} \{4\}, \{5\}, \{6\} has totally different performance. The loading plot represents correlation between the variables and score plot represent the correlation between the observations. In this way for considering difference data preprocessing zero number of sensor set has the response which are correlated. The response of sensor \{3\}, \{4\} opposite of all other sensors and most significant role in classification is dominated by sensor \{3\}, \{2\}, \{4\}.

Figure 4.47 shows the score plot for analyzing the relation between observations of data difference. In this plot the clusters related to a particular kind of sensor has been bounded. There are three categories of organic gases acetone (A), ethanol (E) and propane (P). From the figure it can be clearly seen that PCA is able to discriminate all these three clusters by separate boundary. Hence data difference PCA classification is performing well for detecting three different kinds of organic gases qualities.

Figure 4.48 consist of Bi plot which shows both score and loading plot together. In this plot both observations and variables scatter plot are displayed. The cluster on the left side are the sensor principal components where points below the origin is sensor 4 and remaining sensor are above origin and the observation of acetone, ethanol and propane samples are on the right side.
Figure 4.47 PCA on organic gasses Difference feature score plot

Figure 4.48 PCA on organic gasses Difference feature Bi Plot
4.4.9 PCA Analysis of Organic Gases and Compounds D2 data base using data fractional preprocessing:

In this section PCA has been applied to the organic gases and compounds D2 data base having 6 MOS sensors and 3*3 odor response of organic gases representing acetone, ethanol and propane groups of data sample. Figure 4.49 is the loading plot for the sensors performance clusters for considering fraction in the readings of voltage response. This plot has 6 different scattering points with numbering 1 to 6.

In this plot sensor {1}, {2},{3}, {4} has totally different performance and{5,6} are responding in a similar manner. The loading plot represents correlation between the variables and score plot represent the correlation between the observations. In this way for considering fractional data preprocessing the response of sensor set {5, 6} are correlated or similar analysis. The response of sensor {3}, {4}, {5}, {6} are opposite to other sensors and most significant role in classification is dominated by sensor {2}, {3}, {4}.

Figure 4.50 shows the score plot for analyzing the relation between observations of data fraction. In this plot there are bounded clusters related to a particular kind of sensor. There are three categories of organic gases acetone (A), ethanol (E) and propane (P).From the figure it can be clearly seen that PCA is able to discriminate all the three clusters by separate boundary i.e. acetone, propane and ethanol. Hence data fraction PCA classification is performing well for discriminating different acetone, propane and ethanol kinds of sample qualities.
Figure 4.49 PCA on organic gases fractional feature loading plot

Figure 4.51 consist of Bi plot which shows both score and loading plot together. In this plot both observations and variables scatter plot are displayed. The cluster on the left side are the sensor principal components where points below the origin is sensor 3 and 4 and remaining sensor are above origin and the observation of acetone, ethanol and propane samples are on the right side. In the Bi plot the sensor component are widely scattered as they do not giving a similar performance for fractional data and the observation point are too much aggregated due to this it is tough to classify the odor by fractional method.
Figure 4.50 PCA on organic gases fractional feature score plot

Figure 4.51 PCA on organic gases fractional feature Bi Plot
4.4.10 PCA Analysis of Organic Gases and Compounds D2 data base using data relative preprocessing:

In this section PCA is being applied to the D2 data base of organic gases having 6 MOS sensors and 3*3 odor response of organic gases representing acetone, ethanol and propane groups of data sample. Figure 4.52 is the loading plot for the sensors performance clusters for considering relative in the readings of voltage response. This plot has 6 different scattering points with numbering 1 to 6.

In this plot sensor {1}, {2}, {3}, {4} has totally different performance and {5,6} are responding in a similar manner. The loading plot represents correlation between the variables and score plot represent the correlation between the observations. In this way for considering relative data preprocessing the response of sensor set {5,6} are correlated or similar analysis. The response of sensor {2} is opposite to other sensors and most significant role in classification is dominated by sensor {2}, {3}, {4}.

Figure 4.53 shows the score plot for analyzing the relation between observations of data relative. In this plot the clusters related to a particular kind of sensor are being bounded. There are three categories of organic gases acetone (A), ethanol (E) and propane (P). From the figure it can be clearly seen that PCA is able to discriminate all three clusters by separate boundary it is acetone, ethanol and propane. Hence data relative PCA classification is performing well for detecting different organic gases.
Figure 4.52 PCA on organic gasses relative feature loading plot

Figure 4.53 PCA on organic gasses relative feature score plot
Figure 4.54 consist of Bi plot which shows both score and loading plot together. In this plot both observations and variables scatter plot are displayed. The cluster on the left side are the sensor principal components where points above of the origin is sensor 1,3,4,5 and 6 and sensor 2 is below of the origin and the observation of acetone, ethanol and propane samples are on the right side. In the Bi plot the sensor component are widely scattered as they do not give a similar performance for relative data and the observation point are too much aggregated, due to this it is tough to classify the odor by relative method.

![Bi-plot](image)
4.4.11 PLS Analysis of Organic Gases and Compounds D2 data base using data difference preprocessing:

In this section PLS technique is being applied to the D2 data base having 6 MOS sensors and 3*3 odor response of organic gases representing acetone, ethanol and propane groups of data sample. Figure 4.55 is the loading plot for the sensors performance clusters for considering difference in the readings of voltage response (ΔV). This plot has 6 different scattering points with numbering 1 to 6.

![Loading plot](image)

Figure 4.55 PLS on organic gasses Difference feature loading plot

In this plot sensor {1}, {2}, {3}, {4}, {5}, {6} has totally different performance as they are uncorrelated sensor set. The loading plot represents correlation between the variables and score plot represent the
correlation between the observations. In this way for considering difference data preprocessing none of the sensor set has the response which are correlated. The response of sensor \{3\}, \{4\} opposite of all other sensors and most significant role in classification is dominated by sensor \{3\}, \{2\}, \{4\}.

Figure 4.56 shows the score plot for analyzing the relation between observations of data difference. In this plot the clusters related to a particular kind of sensor are being bounded. There are three categories of organic gases acetone (A), ethanol(E) and propane (P). From the figure it can be clearly seen that PLS is able to discriminate all these three clusters by separate boundary. Hence data difference PLS classification is performing well for detecting three different kinds of organic gases qualities.

![Score plot](image)

**Figure 4.56 PLS on organic gases Difference feature score plot**
Figure 4.57 consist of Bi plot which shows both score and loading plot together. In this plot both observations and variables scatter plot are displayed. The cluster on the left side are the sensor principal components where points below the origin is sensor 4 and remaining sensor are above origin and the observation of acetone, ethanol and propane samples are on the right side.

![Bi-plot](image)

**Figure 4.57 PLS on organic gasses Difference feature Bi Plot**

### 4.4.12 PLS Analysis of Organic Gases and Compounds D2 data base using data fraction preprocessing:

In this section PLS is applied to the D2 data base having 6 MOS sensors and 3*3 odor response of organic gases representing acetone, ethanol and propane groups of data sample. Figure 4.58 is the loading plot for the
sensors performance clusters for considering fraction in the readings of voltage response ($\Delta V$). This plot has 6 different scattering points with numbering 1 to 6.

In this plot sensor {1}, {2}, {3}, {4} has totally different performance and {5,6} are also responding in a different manner. The loading plot represents correlation between the variables and score plot represent the correlation between the observations. In this way for considering fraction data preprocessing the response of sensor set{5,6} are correlated. The response of sensor {3}, {4}, {5}, {6} opposite of all other sensors and most significant role in classification is dominated by sensor {2}, {4}, {3}.

![Figure 4.58 PLS on organic gasses fractional feature loading plot](image)
Figure 4.59 shows the score plot for analyzing the relation between observations of data fraction. In this plot the clusters related to a particular kind of sensor are being bounded. There are three categories of organic gases acetone (A), ethanol (E) and propane (P). From the figure 4.59 it can be clearly seen that PLS is able to discriminate all these three clusters by separate boundary. Hence data fraction PLS classification is performing well for detecting all three different kinds of organic gases qualities.

![Figure 4.59 PLS on organic gasses fractional feature score plot](image)

Figure 4.60 consist of Bi plot which shows both score and loading plot together. In this plot both observations and variables scatter plot are displayed. The cluster on the left side are the sensor principal components where points below the origin is sensor {1}, {2} and remaining sensor are...
above origin and the observation of acetone, ethanol and propane samples are on the right side.

![Bi-plot](image)

Figure 4.60 PLS on organic gasses fractional feature Bi Plot

### 4.4.13 PLS Analysis of Organic Gases and Compounds D2 data base using data relative preprocessing:

PLS technique has been applied in this section to the D2 data base having 6 MOS sensors and 3*3 odor response of organic gases representing acetone, ethanol and propane groups of data sample. Figure 4.61 is the loading plot for the sensors performance clusters for considering
relative in the readings of voltage response ($\Delta V$). This plot has 6 different scattering points with numbering 1 to 6.

In this plot sensor {1}, {2}, {3}, {4} has totally different performance and {5,6} are also responding in a different manner. The loading plot represents correlation between the variables and score plot represent the correlation between the observations. In this way for considering relative data preprocessing the response of sensor set {5,6} are correlated. The response of sensor {3}, {4}, {5}, {6} opposite of all other sensors and most significant role in classification is dominated by sensor {2}, {3}, {4}.

![Figure 4.61 PLS on organic gases relative feature loading plot](image-url)
Figure 4.62 PLS on organic gasses relative feature score plot

Figure 4.63 PLS on organic gasses relative feature Bi Plot
Figure 4.62 shows the score plot for analyzing the relation between observations of data relative. In this plot the clusters related to a particular kind of sensor are being bounded. There are three categories of organic gases acetone (A), ethanol (E) and propane (P). From the figure it can be clearly seen that PLS is able to discriminate all these three clusters by separate boundary. Hence data relative PLS classification is performing well for detecting all three different kinds of organic gases qualities.

Figure 4.63 consist of Bi plot which shows both score and loading plot together. In this plot both observations and variables scatter plot are displayed. The cluster in the left side are the sensor principal components where points below the origin is sensor {1}, {2} and remaining sensor are above origin and the observation of acetone, ethanol and propane samples are in the right side.