CHAPTER 4

Classification and Quantification of Individual Gases/Odors Using Dynamic Responses of Thick Film Sensor Array

This chapter deals with classification and simultaneous quantification of individual gases/odors using dynamic responses of thick film gas sensor array. A new method called average slope multiplication (ASM) has been proposed for classification and simultaneous quantification of gases/odors. The published dynamic responses of thick film gas sensor array [Chaturvedi *et al.* (1999)] have been used as input data for the present work. The instantaneous values extracted from the dynamic response/recovery plots for various test gases viz., LPG, CCl_4, CO, and C_3H_7OH were correlated to its neighboring response plots by the use of proposed ASM technique. It has been demonstrated that the proposed method offers excellent results for classification and simultaneous quantification of individual gases/odors using the dynamic responses of thick film gas sensor array along with the use of BPNN classifier. Principal component analysis (PCA) has been further used wherever needed for data preprocessing, dimensionality reduction and for performance comparison purpose.

4.1 Introduction

Data preprocessing is generally one of the ubiquitous steps in the field of pattern recognition (PARC). In gas sensing field, preprocessing the raw data may drastically improve the results of the subsequent neural network used for identification and/or quantification. Moreover, the complexity of the whole gas recognition module can vary depending on the complexity of the chosen data preprocessing technique. The technique to be chosen for PARC system also depends on whether the given data is in steady state or dynamic form or their mixture and hence the complexity and accuracy of the gas sensing system also varies accordingly.

Neural Network techniques are being widely used for the pattern recognition and classification purposes due to their ability of handling highly non-linear data with massive parallel processing ability. A large number of neural network techniques are reported in the literature such as back propagation neural network [Yaqun (2010)],

radial basis function neural network [Lippman (1989); Kumar *et al.* (2009(b))], support vector machines [Nick *et al.* (2010)], self organizing map [Kohonen (1990); Marco (1998)], adaptive resonance theory and so on [Carpenter and Grossberg (1988)]. Each of these techniques has its own advantages over the other.

In literature, both steady states as well as dynamic responses have been used for classification and quantification of gases/odors. In the present chapter, only dynamic responses have been used as input data for classification and simultaneous quantification of gases/odors.

Many authors have reported their effort for classification of gases/odors/volatile organic compounds (VOCs) using steady state and/or dynamic responses of sensor array. Llobet *et al.* (1997) have reported successful gases/odors classification tasks using the combination of steady state as well as dynamic response of a thick film gas sensor array. Nakamura *et al.* (1997) have done gases/odors classification using the dynamic response of the chemical sensor array using the Kalman Filter algorithm for data transformation and linear vector quantization (LVQ) for classification purpose. Sobanski *et al.* (2006) have reported the improvement in classification accuracy by using the discrete wavelet transform (DWT) together with k-nearest neighbor (kNN) classification algorithm and probabilistic neural network (PNN). Szczureka *et al.* (2009) reported the identification of various gases based on the dynamic response of sensor array with emphasis on processing of discrete measurement data using data profiling scheme for the gas identification purpose. Kumar *et al.* (2011(c)) obtained 100% correct identification of gases/odors by processing the dynamic responses of a gas sensor array using highly complex technique of wavelet transformation and multi-scale principal component analysis (MSPCA).

The quantification of harmful gases is required before they reach to their explosive and/or harmful limit. A lot of work has been reported in literature on quantification of gases/odors and/or volatile organic compounds (VOCs). Llobet *et al.* (1997) have reported qualitative and quantitative analysis of VOCs using transient as well as steady state responses of a thick film gas sensor array with principal component analysis (PCA) and artificial neural network approach. Lee *et al.* (2000) have reported quantification of explosive gases using nine discrete sensors and PCA, ANN approach. Daki and Wei (2007) have reported identification and quantification of odor using function estimation model ensemble. Kumar *et al.* (2010(b)) have

reported classification and quantification of alcohols and alcoholic beverages using steady state responses of thick film gas sensor array using fuzzy-neuro concept. Thus, various schemes have been adopted for quantification of gases/odors and VOCs.

It can therefore be observed that much complex and intricate schemes have been employed for the identification and simultaneous quantification of gases/odors using dynamic responses of the sensor array. In this work, a new and simpler scheme has been proposed for the classification and simultaneous quantification of gases/odor using the dynamic response of thick film gas sensor array. It has been found that the slope of response/recovery characteristics are different for different test gases/odors and their concentrations for a particular sensor.

The instantaneous values extracted from the dynamic response/recovery plots for test gases/odors have been correlated to its neighboring response/recovery plots by the use of proposed ASM technique. The gases/odors' classification and simultaneous quantification results obtained with the proposed method are found to be more promising. The backpropagation neural network is used on transformed data for the classification purpose. The BPNN is chosen because it is relatively simple and is quite suitable for implementing multi-class classification [Osuna (2002)].

This chapter has been divided into 4 main sections. Section 4.2 deals with the experimental background and methods followed. Section 4.3 includes the results and discussion. Section 4.4 concludes the reported experiment and its findings.

4.2 Materials and Methods

The classification and simultaneous quantification scheme followed in this work can be described as per the block diagram shown in Fig. 4.1. Its first block represents the data generation block consisting of the gas sensor array which generates the dynamic response/recovery data for various samples of the considered test gases/odors. This dynamic response/recovery raw dataset is then fed into the succeeding data preprocessing blocks which transformed the raw data into respective ASM transformed data and PCA preprocessed data. The raw data, the ASM transformed data and their PCA preprocessed versions are finally fed to BPNN classifier for training and validation purpose.

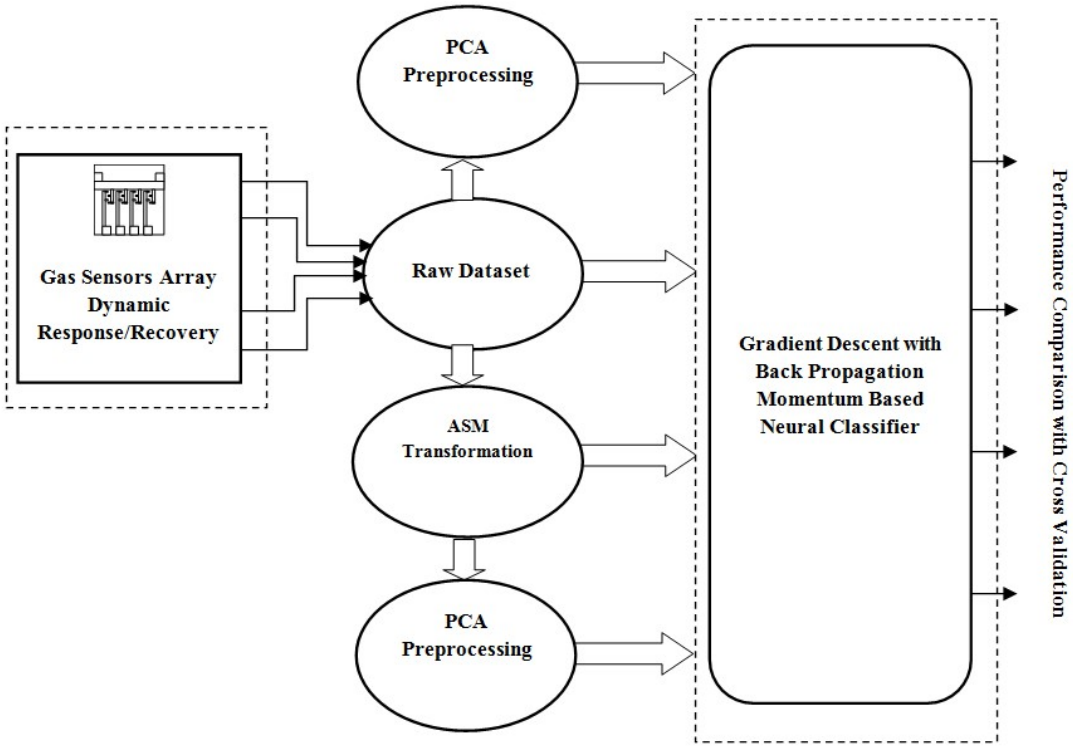

Fig. 4.1 The method followed for the gases/odors classification using ASM

4.2.1 Experimental Background

The gas sensor array referred in this work is a thick-film tin-oxide based sensor array consisting of six sensor elements fabricated and characterized in our own research lab by previous researchers [Chaturvedi *et al.* (1999)]. The schematic structure of the fabricated array is shown in Fig. 4.2. Its sensor elements were doped with 1% Pd, Pt, Cd, CuO, ZnO and pure SnO_2 (by weight), respectively. The sensor was also exposed to oxygen plasma (generated by low pressure and RF power at 13.56 MHz) for 15 min during the process of its fabrication. The response/ recovery characteristics of the array were measured in a test chamber with a provision to inject the desired volume of the test gases. The response of the array in terms of change in resistance of all the 6 sensors for the four gases/odors viz. LPG, CCl_4, CO and C_3H_7OH were recorded *w.r.t.* time and has been reproduced in Fig. 4.3. Further, details of the experiment may also be referred in [Chaturvedi *et al.* (1999)].

4.2.2 Data Extraction and Interpretation

The dynamic response/recovery raw dataset for its use in this chapter has been extracted from the dynamic characteristics curves of Fig. 4.3 using a graph extracting tool named 'Precision Image Digitizer (PID)' ver. 1.3.0.0' [Internet Resource (IR8)].

The characteristics of Cd doped sensor element have not been considered due to its poor response.

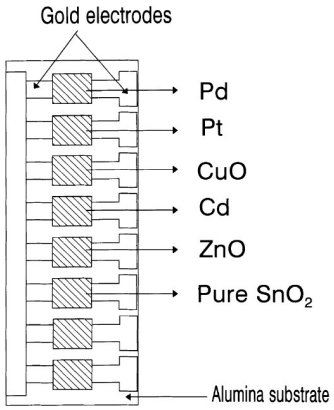

Fig. 4.2 The thick film gas sensor array device used for generating dynamic responses [Chaturvedi *et al.* (1999)]

For a given sensor, response and recovery curves are sampled at 6 different intervals for each concentration band which result in 12 data points. There are four concentration bands viz. 25ppm, 50 ppm, 75 ppm and 100 ppm, respectively (Fig. 4.3). This makes 12 x 4=48 data points for a single gas. For five chosen sensors, it makes a 48 x 5 dimensional matrix for a single gas. Now, there are four test gases, which makes (48 x 4) x 5= 192 x 5 dimensional matrix.

Taking all considered sensors at a time, the composite raw data were plotted on different 2-D scatter graphs for each concentration and to graphically explore the inherent data distribution. One such 2-D scatter plot (for 25 ppm conc. band) is shown in Fig. 4.4 (symbols S1, S2, S3, S4 and S5 correspond to pure SnO_2, Pd-doped, Pt-doped, CuO-doped and ZnO-doped sensor respectively). Similarly, Considering 3 sensors at a time (out of 5), the composite dataset was plotted on 10 different 3-D scatter graphs to graphically explore the inherent data distribution. One such 3-D scatter plot is shown in Fig. 4.5. It is observed that the raw data belonging to cluster of respective gases/odors are overlapping in its current form.

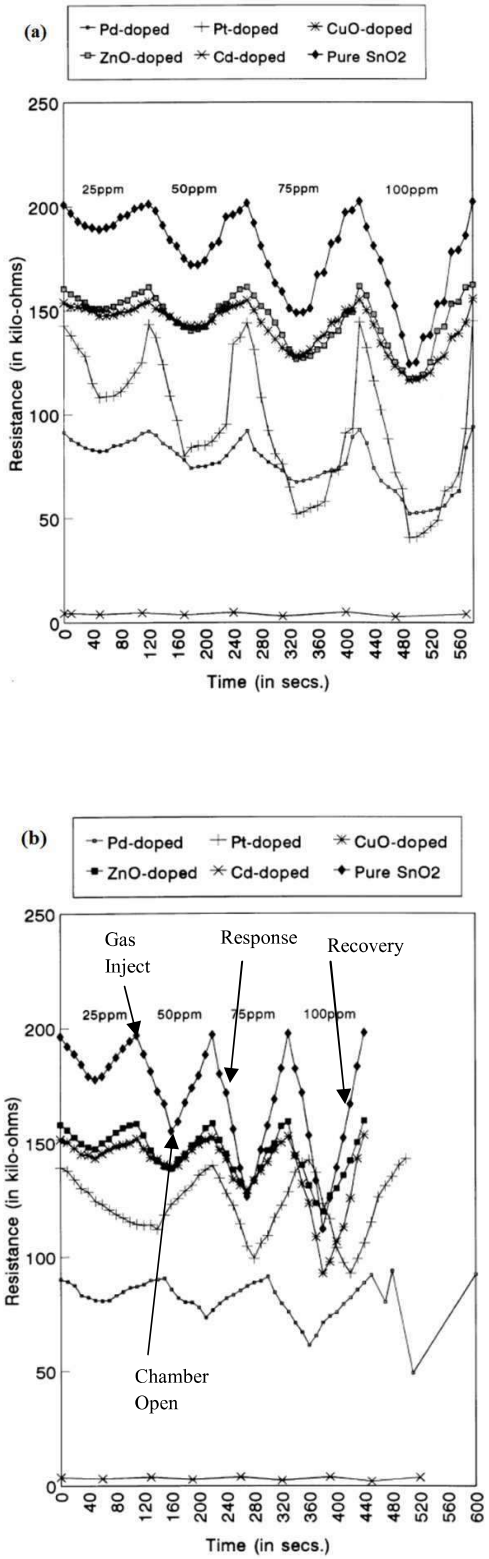

Fig. 4.3 Response–recovery plot of the sensor array for different concentrations of (a) LPG and (b) CCl_4 responses [Chaturvedi *et al.* (1999)]

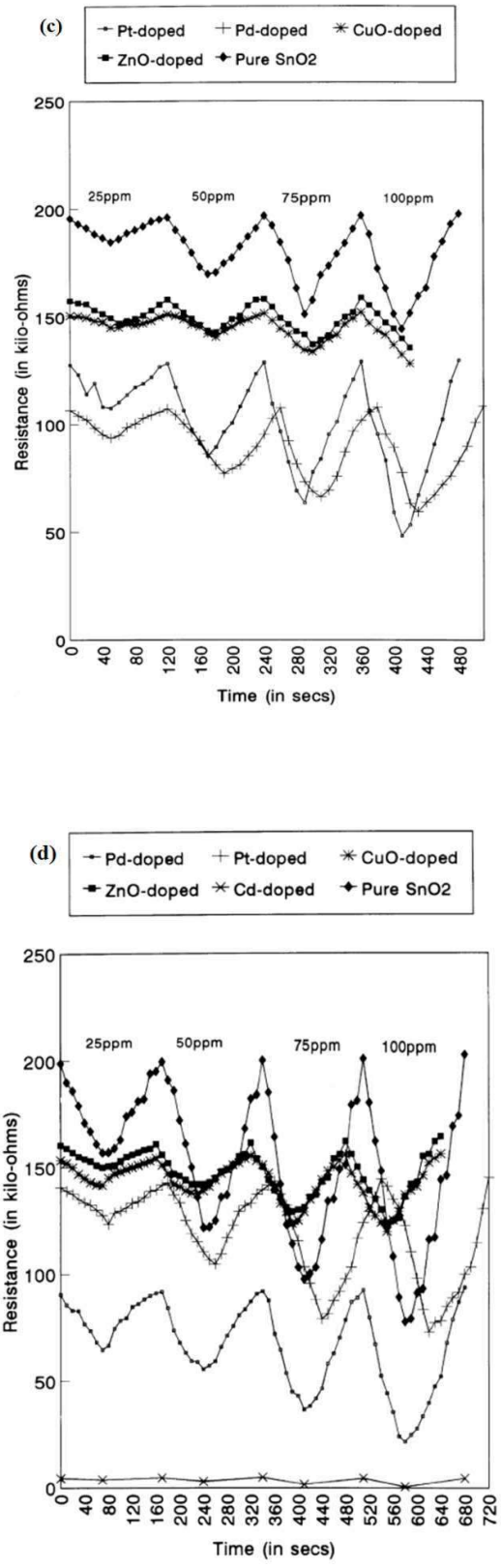

Fig. 4.3 Response–recovery plot of the sensor array for different concentrations of (c) CO, (d) C_3H_7OH responses [Chaturvedi *et al.* (1999)].

A transformation of the dataset has been done using a newly proposed method of average slope multiplication (ASM) to obtain less overlapped clusters of respective gases/odors. PCA technique has been used for further preprocessing the ASM data to investigate any improvements in the results. The raw data, ASM transformed data and its PCA processed data version have been used to train and test the BPNN classifier.

4.2.3 Average Slope Multiplication (ASM) Method

The ASM approach is based on multiplying the response and/or recovery data by the magnitude of the respective average slope value. The average slope method has some important properties which can make it suitable for dynamic data preprocessing. After a close observation of the linear region of the dynamic response/recovery curves (Fig. 4.3), it has been found that for a particular concentration band and for a given sensor, the slope of all response/recovery curves have different values for different gases/odors. Actually, this slope value reflects the response/recovery behavior of the gas/odors for particular sensor. Further, since the method is based on the average slope values, the inherent time inclusion can make this method suitable for dynamic data preprocessing. The ASM method thus utilizes the variation of the average slope value of the linear region of response/recovery characteristics of a test gas/odor and its concentration for a given sensor.

The average slope can be calculated by taking the difference between two extreme data points of the linear region of a response or a recovery curve and dividing it by the respective time difference. Suppose, there are two points $A = (t_A, y_A)$ and $B = (t_B, y_B)$ as shown in Fig. 4.6. The average slope (AS) of the curve between A and B can be defined as mentioned in (4.1) [Stran (1991)].

$$AS = \left| \frac{y_A - y_B}{t_B - t_A} \right| = \left| \frac{\Delta y}{\Delta t} \right| \tag{4.1}$$

Fig. 4.7 shows the magnitudes of average slope values plotted on the 2-D graph for all the five considered sensors for different gases/odors for 25-ppm concentration band (the plots for other concentration bands are not shown for the sake of simplicity). The graph shows that there is separation among the magnitude of average slope values for different gases for a given sensor. This idea was the motivation to develop ASM method for gases/odor classification using dynamic responses.

Similarly, ASM method was applied to recovery curve and hence for the complete dynamic characteristics to obtain a composite dataset of all gases/odors for

each sensor. This composite dataset was further used for preprocessing and classification. Fig. 4.8 2-D shows scatter plot of ASM transformed data for 25-ppm concentration band. Fig. 4.9 show the 3-D scatter plot of ASM transformed data the for same 25-ppm concentration band.

It is obvious from comparison of Fig. 4.5 and Fig. 4.9 that the ASM method provides more separation among different clusters of gases/odors data as compared to the raw data. Thus, the neural classifier is expected to provide better classification results with the ASM transformed data.

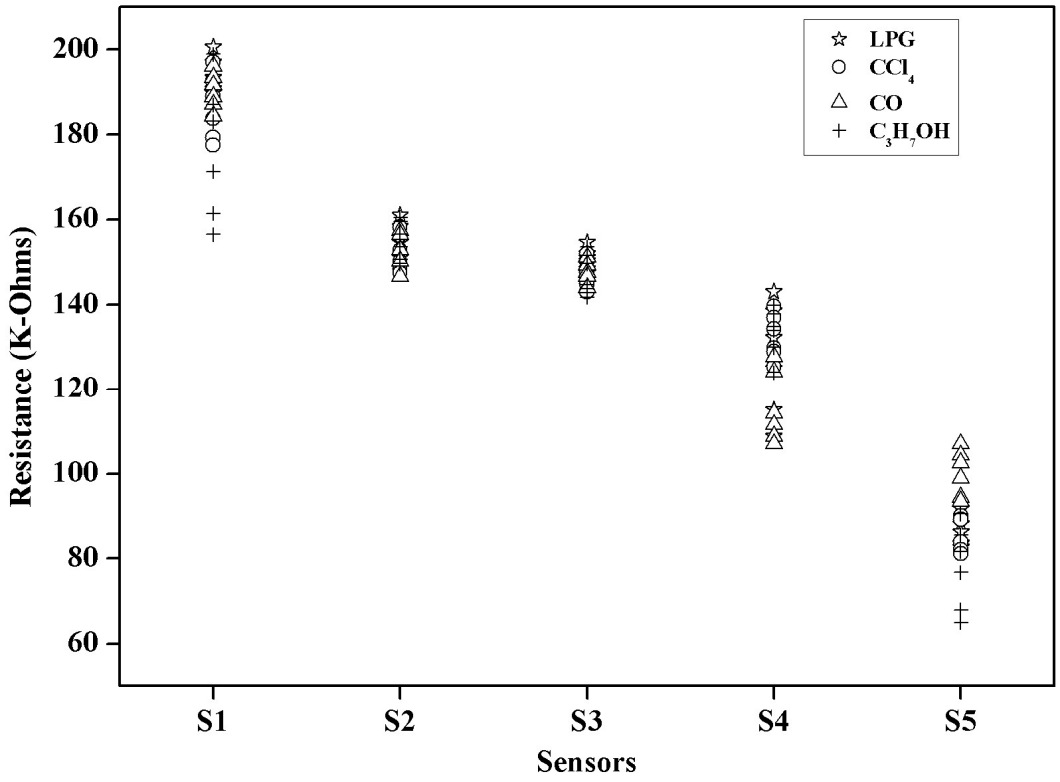

Fig. 4.4 2-D scatter plot of raw data extracted from dynamic responses for 25 ppm concentration band

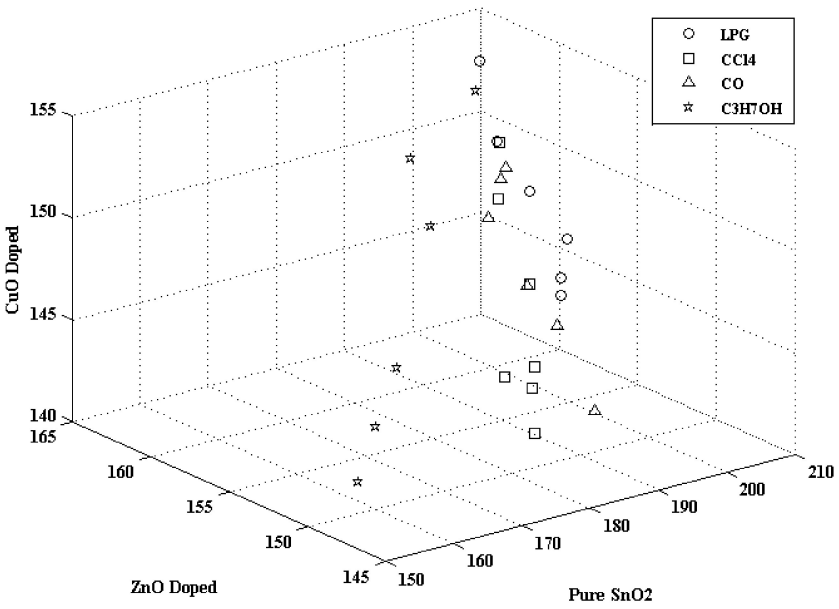

Fig. 4.5 3-D scatter plot of raw data extracted from dynamic responses for 25 ppm concentration band

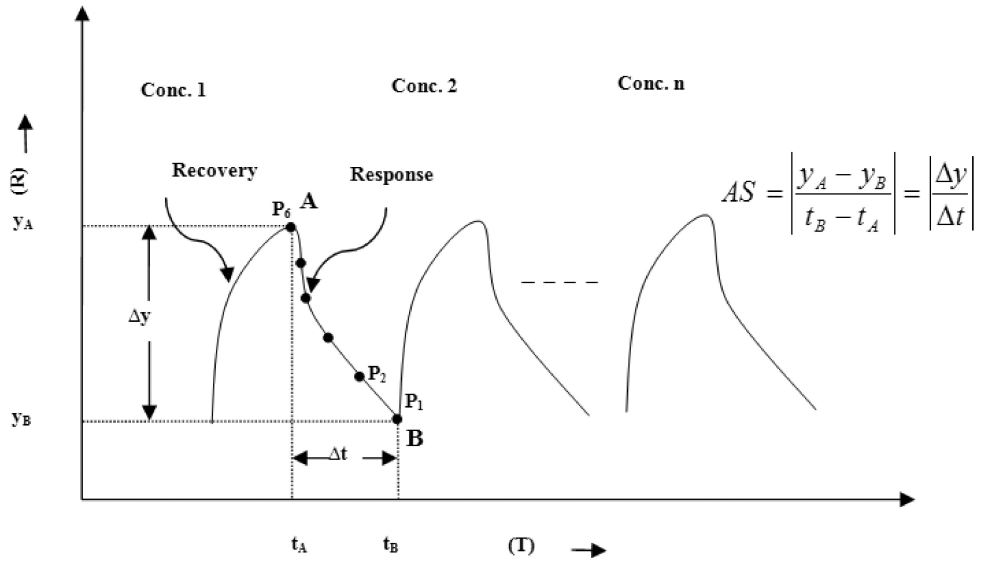

Fig. 4.6 Average slope calculation method used for dynamic responses

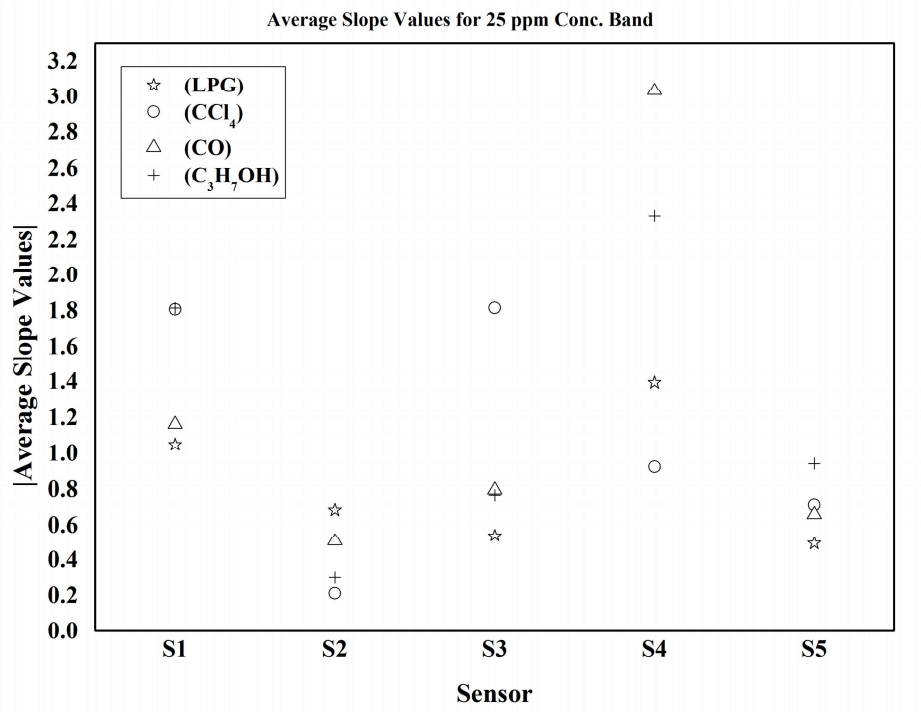

Fig. 4.7 Average slope values of response of each sensor for different gases/odors plotted on 2-D scatter graph for 25-ppm conc. Band

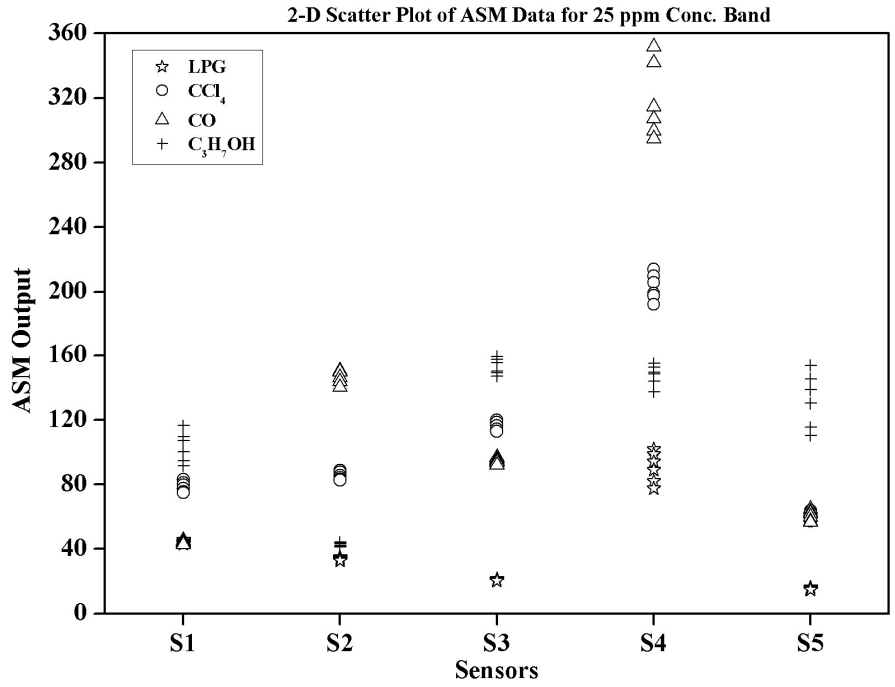

Fig. 4.8 2-D scatter plot of ASM transformed data for 25-ppm conc. band

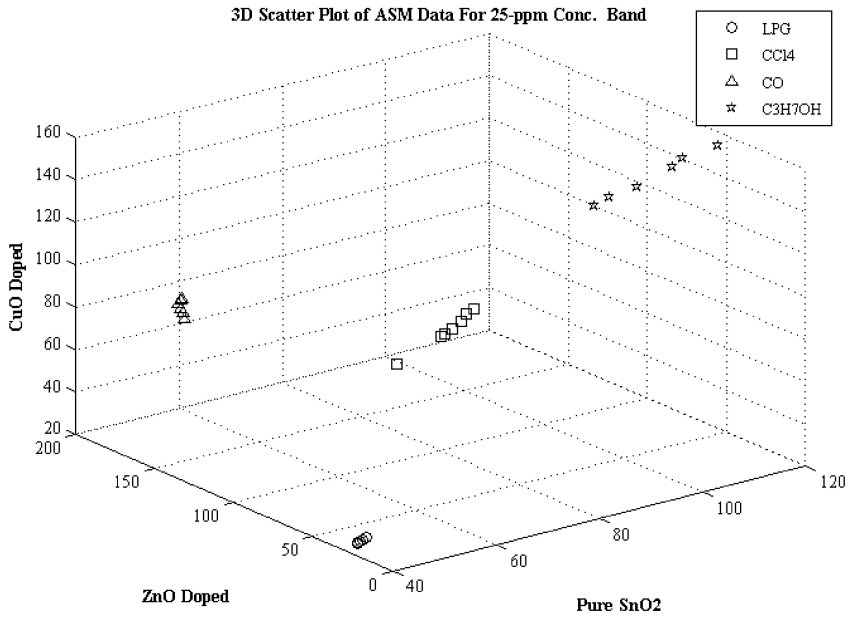

Fig. 4.9 3-D scatter plot of ASM transformed data for 25-ppm conc. band

The PCA technique was also applied to the raw data set of gases/odors. Dimensionality reduction was done from 5 to 3, as first three principal components were covering almost 99% of the variance for the combined data. PCA was applied to ASM transformed dataset as well to seek further improvement in the classification results. For quantification, the results were accurate even without PCA preprocessing.

4.3 Results and Discussion

Multilayer perceptron (MLP) with backpropagation (BP) algorithms is the basic neural network techniques used for pattern recognition problems. MLP constitutes an important class of neural network which typically consist of a set of sensory units called source node which form the input layer. It consists of one or more hidden layers of computation nodes having different threshold functions. Typical multilayer feed-forward neural network architecture is shown in Fig**.** 4.10.

As discussed in the previous chapter, in MLP networks the input signal propagates through the network in a forward direction, on a layer by layer basis. The MLP with error back propagation has been successfully applied to many problems in pattern recognition by training in a supervised manner. In the present study, gradient descent with momentum back propagation (GDMBP) algorithm has been used as a training algorithm. The GDMBP is considered as standard back propagation

algorithm and is widely used for multi layered artificial neural network because of its computational simplicity, ease of implementation and good generality [Rumelhart *et al.* (1986)]. The weight and bias values are updated by two training control parameters viz. learning rate (η) and momentum (α). The learning rate is the step size along the error surface while the momentum is a proportion of the previously calculated weight change. The weights are changed in according to equation (4.2) [Packianather and Drake (2005)].

$$\Delta W_{ji} = \eta \times \delta_j \times o_i + \alpha \times \Delta W_{ji}(old) \tag{4.2}$$

where,

$$\delta_j = f'(x_j)(t_j - o_j) \quad \text{for output layer neurons} \tag{4.3}$$

$$\delta_j = f'(x_j)\sum_k \delta_k W_{kj} \quad \text{for neurons in the hidden layer} \tag{4.4}$$

where, o_j denotes the actual outputs and t_j denotes the desired outputs.

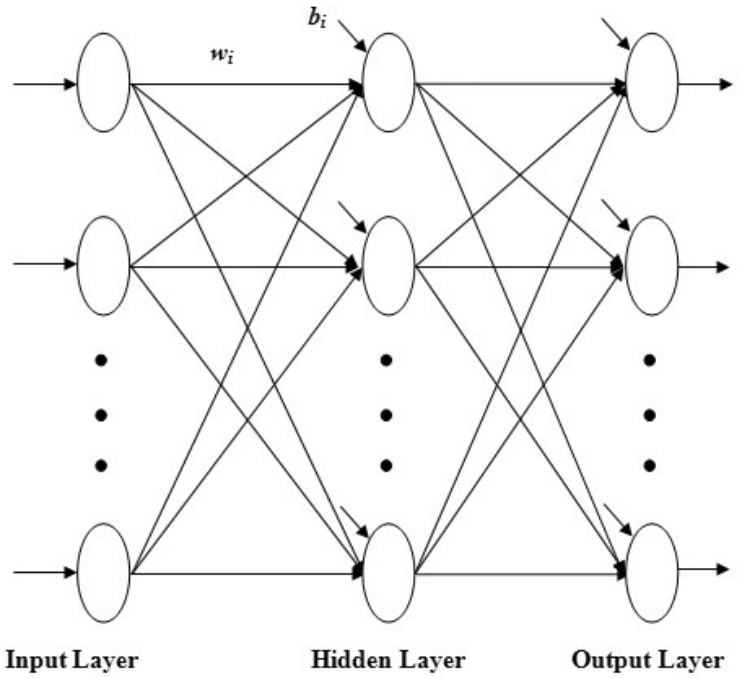

Fig. 4.10 Typical multilayer feed forward neural network used as classifier

The learning rate (η) and momentum (α) play very important role on the learning speed of BP algorithm. Faster convergence is obtained with larger values of learning rate but may lead to network oscillations while its small value leads to slower convergence. Also, the higher values of momentum improve the convergence

avoiding the local minima [Chaturvedi (2008); Haykin (2008)]. In the present work, these two parameters along with the number of hidden layers (*H*) and the number of neurons in each hidden layer (*NH*) were tuned for the best classification performance of GDMBP neural network separately with raw data, ASM transformed data and their PCA preprocessed versions.

4.3.1 Classification of Individual Gases/Odors with ASM Using Combined (Response and Recovery) Dataset

The average slope multiplication (ASM) technique was applied to the raw data, sampled from the dynamic responses of thick-film tin-oxide gas sensor array, as mentioned in section 4.2. The transformed data was plotted on 3-D scatter graph with different combinations of gas sensors taking 3 sensors at a time. The 3-D scatter graphs were plotted using MATLAB.

Comparison of these plots (Fig. 4.5 and Fig. 4.9) clearly shows the highly improved separation in transformed data as compared to the raw data. This shows the improved performance of ASM transformed data over the raw data. PCA preprocessing was also applied on ASM transformed data to seek more classification accuracy for various gases/odors.

Further, the experiments were performed with backpropagation neural network with GDMBP training algorithm using the raw data, the transformed ASM data and with their PCA processed versions separately. The classification results thus obtained have been compared. The *k*-fold cross validation scheme explained in previous chapter [Rodriguez (2010)] was adopted here also for the all experiments. The value of *k* was chosen to be 8.

The network was tuned for number of hidden layers (*H*), number of neurons in the hidden layer (*NH*), the learning rate (η) and momentum (α) and number of epochs. In all cases, the single hidden layer produced the best results. It was found that the classification accuracy was only 86% (Fig. 4.11) using the GDMBP neural network with the raw data. The raw data after PCA preprocessing and dimensionality reduction from 5 to 3 provided 90% classification accuracy.

The same neural network was then trained and tested with the ASM transformed data which provided the classification accuracy of 97% (Fig. 4.12). Finally, the ASM transformed data was preprocessed with PCA and the dimensionality reduction was again done from 5 to 3.

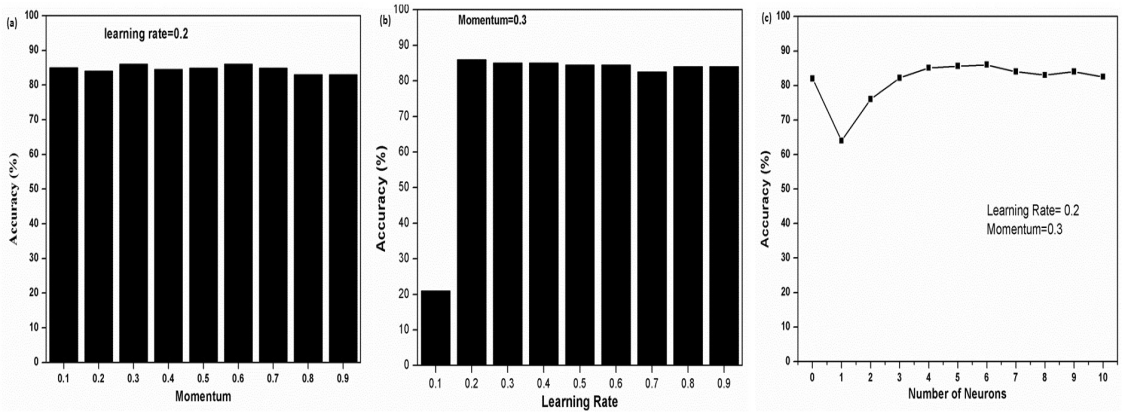

Fig. 4.11 Variation in classification accuracy of raw data with (a) Momentum (b) Learning rate (c) Number of neurons in the hidden layer

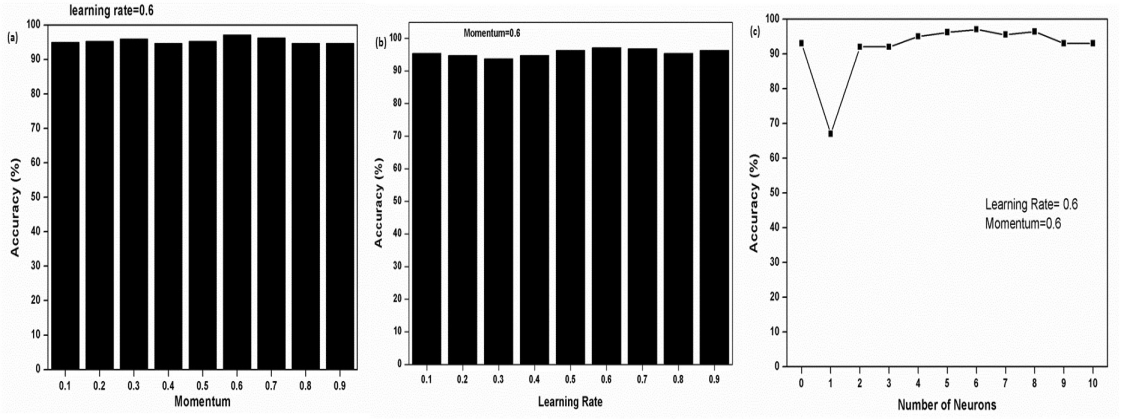

Fig. 4.12 Variation in classification accuracy with ASM data with (a) Momentum (b) Learning rate (c) Number of neurons in the hidden layer

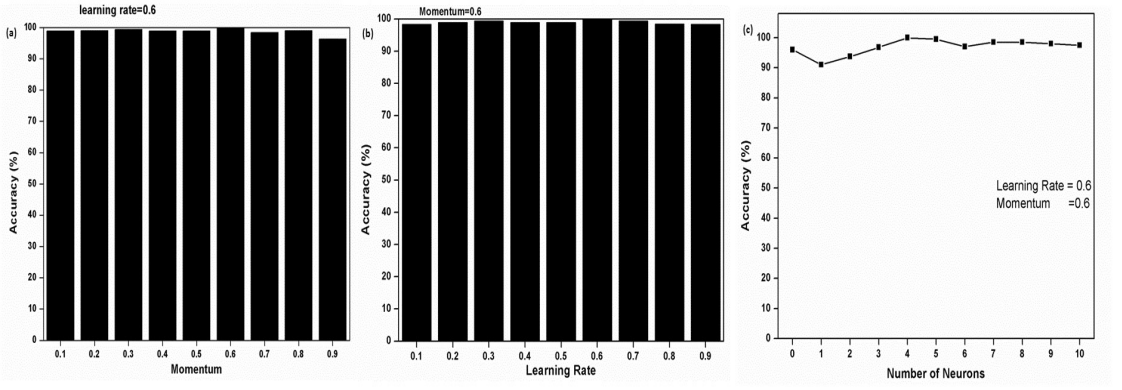

Fig. 4.13 Variation in classification accuracy with PCA preprocessed ASM data with (a) Momentum (b) Learning rate (c) Number of neurons in the hidden layer

This PCA preprocessed ASM data was fed to the GDMBP neural network. When the unseen test samples of PCA preprocessed ASM data were given to the duly trained neural classifier, each time it recognized the entire unseen sample perfectly with 100% classification accuracy (Fig. 4.13).

Table 4.1 summarizes the classification results obtained with raw data, ASM data and PCA preprocessed ASM data using BPNN with optimized values of different parameters like learning rate, momentum and number of epochs etc.

TABLE 4.1 Summary of classification results for combined dataset using ASM with BPNN

Parameters/ Data Type	(Raw) Data	(PCA) Data	(ASM) Data	PCA After ASM Transformation
No. of Neurons in Single Hidden Layer	6	6	5	5
Learning Rate (η)	0.2	0.2	0.6	0.6
Momentum (α)	0.3	0.3	0.6	0.6
Classification Accuracy	86%	90%	97%	100%
Number of Epochs	1508	1526	2071	2023

4.3.2 Simultaneous Quantification using ASM Method

Simultaneous quantification of gases/odors using proposed ASM method has also been performed using the gradient descent with momentum backpropagation (GDMBP) algorithm described previously. Different network architectures could be adopted for simultaneous quantification of the considered gases/odors. The first options was to perform classification of each of the gases/odors separately and then apply the classification results to the quantification network as reported in [Llobet (1997)] and is shown in Fig. 4.14 which is referred as neural network architecture-1 (NNA-1) in rest of this chapter.

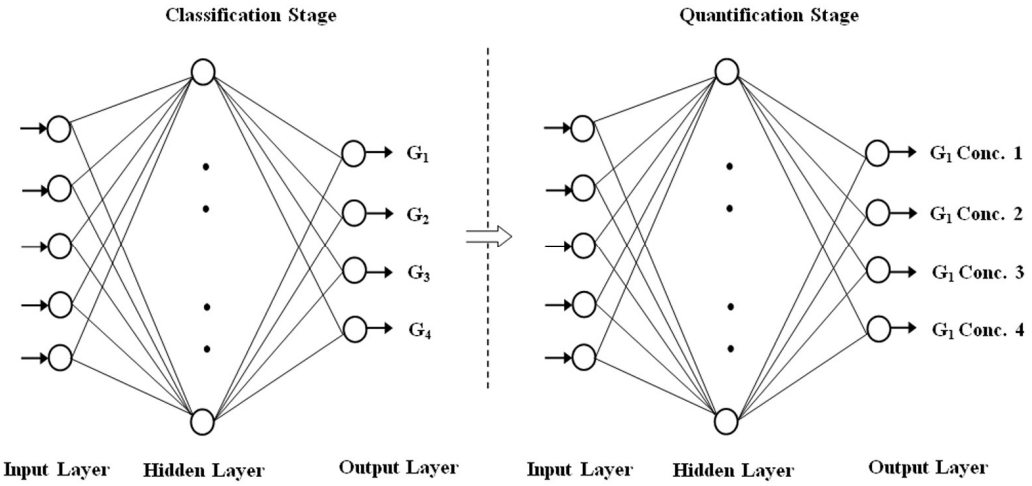

Fig. 4.14 Neural network architecture-1 (NNA-1) adopted for quantification

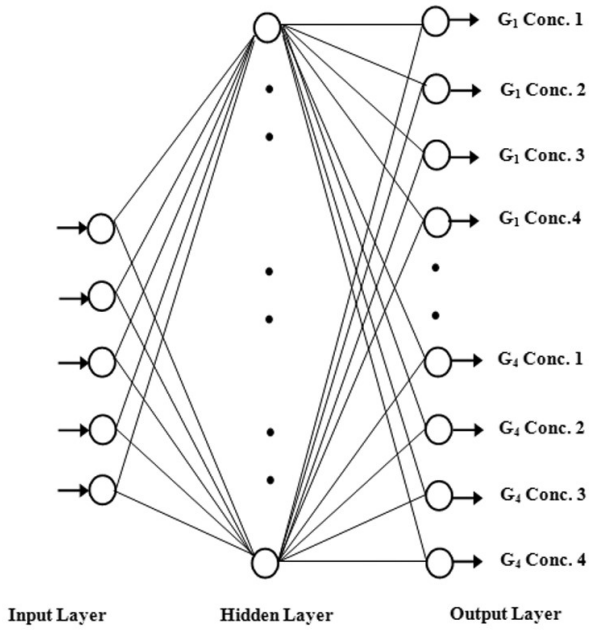

Fig. 4.15 Neural network architecture-2 (NNA-2) adopted for quantification

Another network architecture option was to have fully connected architecture having 5 inputs corresponding to 5 sensor input and 16 outputs where each 4 outputs correspond to 4 concentrations of each particular gas as shown in Fig. 4.15 which is referred as neural network architecture-2 (NNA-2) in rest of this chapter. These two mentioned network architectures were trained and tested using GDMBP with raw data, PCA data as well as with ASM data. The experiments were performed with target of minimum training error of 0.0001. The quantification results thus obtained are presented and compared here.

Firstly, the raw data were trained and tested using neural network architectures viz. NNA-1 and NNA-2 respectively with the GDMBP algorithm using MATLAB. Both the networks provided average quantification accuracy of 69% and 63% respectively with single hidden layer. For PCA data, improvement was noted and the average quantification accuracy was 74% for NNA-1. The NNA-2 showed 67% average quantification accuracy with PCA data but with single hidden layer of 11 neurons. Finally, both the network architectures were trained and tested with ASM transformed data which provided 100% quantification accuracy with single hidden layer. Also, as mentioned in the Table 4.2, the total number of epochs decreased to a large extent when ASM data were used to train the networks.

Thus, the best quantification results have been obtained with ASM transformed data. Table 4.2 summarizes the quantification results obtained with neural network architectures NNA-1 and NNA-2 using raw data PCA data and ASM data using GDMBP with best possible values of different parameters like learning rate, momentum and number of epochs etc.

4.3.3 Experiment with Exposure and Recovery Transients Data Individually

Role of exposure and recovery datasets individually along with the use of proposed ASM method have also been studied for classification and simultaneous quantification of gases/odors. The way experiment has been performed can be understood form the schematic of the Fig. 4.16.

By extracting numerical values of the data points as present in the published graphs [Chaturvedi *et al.* (1998)], 24 samples of exposure and 24 samples of recovery for each gas were obtained, (in our case four gases are there viz. LPG, CO, CCl_4, and C_3H_7OH). Accordingly, the composite data set in its raw form consisted of a total of 96 data samples for exposure and 96 samples of recovery dataset were prepared.

Table 4.2 Summary of simultaneous quantification results for combined dataset using ASM method

NNA-1	**Raw Data**					**PCA Data**					**ASM Data**				
	NH	**A_C (%)**	**(η)**	**(α)**	**Epochs**	**NH**	**A_C (%)**	**(η)**	**(α)**	**Epochs**	**NH**	**A_C (%)**	**(η)**	**(α)**	**Epochs**
LPG	4	70	0.3	0.9	436	4	73	0.3	0.3	475	4	100	0.3	0.2	81
CCl_4	5	68	0.3	0.2	520	4	80	0.2	0.2	434	4	100	0.3	0.2	28
CO	4	71	0.1	0.3	504	4	75	0.1	0.2	468	4	100	0.3	0.2	73
C_3H_7OH	4	67	0.2	0.3	565	4	68	0.2	0.3	498	4	100	0.3	0.2	78
NNA-2															
Combined Data	10	63	0.3	0.2	544	11	67	0.3	0.2	520	5	100	0.3	0.2	492

A_C= average accuracy, η= learning rate, α= momentum, *NH*= number of neurons in single hidden layer (if any)

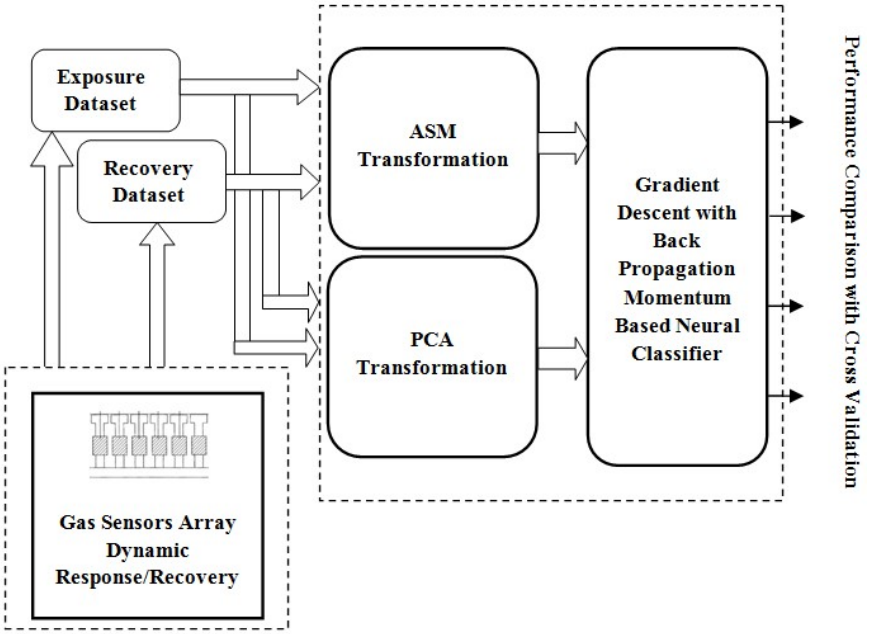

Fig. 4.16 The method followed for the gases/odors classification cum quantification for response and recovery data individually

Five considered sensors' responses were used for four different gases/odors and the exposure and recovery dataset individually consisted of (96 x 5) dimensional matrix. Principal component analysis was used for data preprocessing and dimensionality reduction. Only first three dimensions of PCA preprocessed dataset were used for analysis as they were covering almost 98% of variance. The raw datasets were transformed into respective ASM form with the method explained in Section 4.2 previously. The raw data, the PCA data and the ASM data were used for training and validation of GDMBP classifier. Separate training and validation sets were fed to the neural classifier with *k*-fold cross validation scheme where the value of *k* was optimized and found to be 8 for the present experiment.

(i) Classification of Exposure and Recovery Transients Data individually

For raw data, neural network consisted of five input nodes and four output nodes with GDMBP algorithm. The training error limit was set to 0.0001 for all analysis. Firstly, exposure dataset in raw form was fed to the neural classifier, and control parameters (learning rate and momentum) were tuned along with the number of hidden layer (*H*) and number of neurons in the hidden layer (*NH*). It was found that the exposure dataset in the raw form led to 81.2% classification accuracy with value of learning

rate (η = 0.8), and momentum (α = 0.1) with single hidden layer (H = 1) having five neurons (NH = 5). For PCA preprocessed dataset, the network consisted of three input nodes and four output nodes. The PCA preprocessed version of dataset provided 85.4% of classification accuracy with value of learning rate (η = 0.3), momentum (α = 0.2) with single hidden layer having five neurons. For the ASM dataset, the neural classifier architecture again consisted of five nodes in the input layer and four nodes in the output layer with one hidden layer of 4 neurons. This optimized network provided classification accuracy of 93.7%. Values of different control parameters obtained after optimization are shown in Table 4.3. The average accuracy (A_C) is the average of percentage of accuracy obtained in all cross validation folds. It was found that ASM transformed datasets turned into less complex neural architectures required for discrimination (single hidden layer with four neurons) with less number of epochs. The ASM data were optimized for different mentioned control parameters.

Similarly, recovery dataset was used to train and test the neural classifier in raw form, in PCA form and ASM form. The recovery dataset in raw form provided 84.3% classification accuracy, while the PCA preprocessed data showed 87.5% accuracy. The ASM transformation of recovery data showed best classification results with 95.8% accuracy. It was found that ASM transformation resulted into best accuracy with relatively least number of epochs for exposure data. Values of different control parameters obtained after optimization are shown in Table 4.3. The epoch, defined here is a single presentation of each of the *I/O* patterns in the training set. It is rather a convenient unit to measure learning time for the benchmarks [Fahlman (1998)].

(ii) Simultaneous Concentration Estimation for Exposure and Recovery Transients Data individually

The proposed ASM technique was also used for simultaneous concentration estimation of the individual gases. For this purpose, the exposure/recovery data were divided into different concentration bands (for 25, 50 ppm etc.) for each gas/odor. The neural network architecture adopted for the simultaneous quantification is shown in Fig. 4.15. The GDMBP algorithm explained previously was adopted for concentration estimation. The results obtained for simultaneous concentration estimation with GDMBP using raw, PCA, and ASM dataset, respectively, are shown in Table 4.4.

Table 4.3 Summary of classification results for exposure and recovery transients individually

Data Type	**(η)**	**(α)**	**(A_c)**	***H***	***NH***	**Epochs**
Exposure Data						
Raw	0.8	0.1	81.2	1	5	1248
PCA	0.3	0.2	85.4	1	4	1166
ASM	0.2	0.3	93.7	1	4	493
Recovery Data						
Raw	0.7	0.3	84.3	1	5	1154
PCA	0.2	0.3	87.5	1	4	1123
ASM	0.2	0.3	95.8	1	4	417

(η) = Learning rate, (α) = Momentum, A_c = Average accuracy (%) in all cross validation folds, *H*= Number of hidden layer, *NH*=Number of neurons in the single hidden layer

Table 4.4 Simultaneous concentration estimation results summary for response and recovery data analyzed individually

Data Type	**(η)**	**(α)**	**(A_c)**	***H***	***NH***	**Epochs**
Exposure Data						
Raw	0.8	0.1	81.2	1	5	1248
PCA	0.3	0.2	85.4	1	4	1166
ASM	0.2	0.3	93.7	1	4	493
Recovery Data						
Raw	0.7	0.3	84.3	1	5	1154
PCA	0.2	0.3	87.5	1	4	1123
ASM	0.2	0.3	95.8	1	4	417

It is obvious from the results shown in Table 4.4 that with ASM transformed data the simultaneous concentration estimation capability of the classifier cum quantifier for the individual gases has improved to a great extent in term of accuracy and number of epochs. Results obtained with recovery data after ASM transformations are found relatively more promising in terms of accuracy and number of epochs required to train the network.

4.4 Conclusion

The data preprocessing plays an important role in the identification problem for different gases/odors. Different data preprocessing techniques exist in literatures which were found to be very complex and sophisticated. In this chapter, a new concept of average slope multiplication (ASM) for classification and simultaneous quantification of gases/odors using dynamic characteristics of a gas sensor array has been proposed for the first time as per the available literature. The raw dataset extracted from the dynamic responses of thick film tin oxide gas sensors array with different gases was showing large overlapping among different classes, making it a difficult task to recognize each gas clearly. When this raw dataset was transformed using ASM technique, the separation was improved to a higher extent showing less overlapping between different clusters. This confirms that ASM is a useful tool for the data preprocessing for the gases/odors discrimination using dynamic response of gas sensor array. The PCA can be further used for preprocessing and dimensionality reduction to obtained more accurate classification results with reduced complexity of neural network classifier.

It was also found that exposure dataset and recovery dataset separately showed promising classification and simultaneous quantification accuracy with ASM method. In the present study, the recovery dataset was found to have more discrimination capability as compared to exposure data. This suggests that only recovery or exposure data along with proposed ASM preprocessing technique can be used for enhancing the response time of e-nose system by reducing the complexity of the e-nose system at the expense of little classification accuracy. Thus, the proposed ASM technique can be used for accurate classification and quantification of gases/odors using dynamic response/recovery characteristics of gas sensor array.