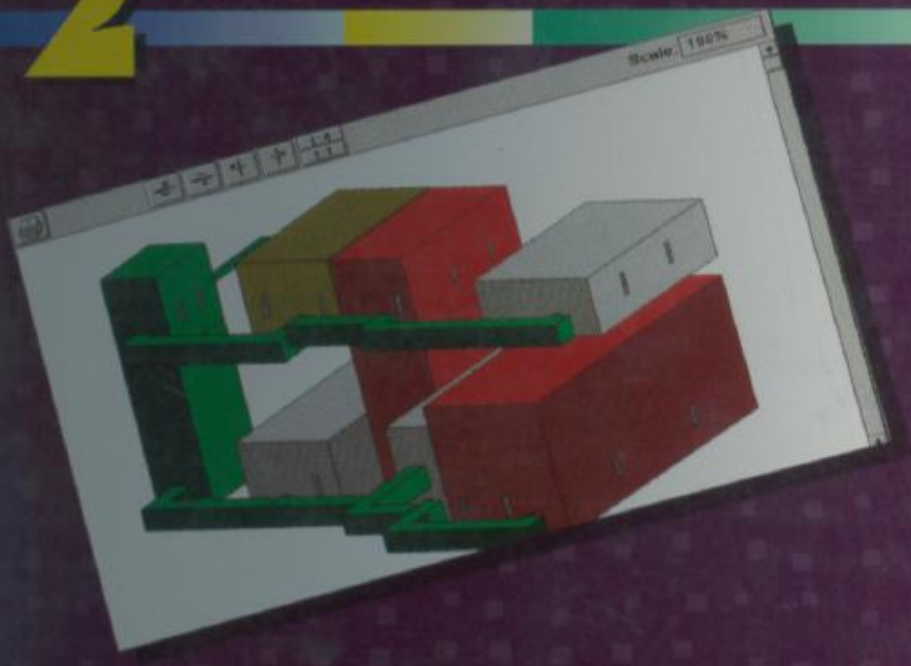
Edited by B. Braunschweig and B.A. Bremdal

ARTIFICIAL INTELLIGENCE IN THE PETROLEUM INDUSTRY

SYMBOLIC AND COMPUTATIONAL APPLICATIONS

2



Series on Artificial Intelligence in the Petroleum Industry

ÉDITIONS TECHNIP

Chapter 4

NEURAL NETWORK EXPERIMENTS ON SYNTHETIC SEISMIC DATA

P.F.M. de Groot[1]

Abstract. This article describes experiments with artificial neural networks aimed at predicting reservoir properties from seismic data. The data is generated with a rule-based Monte Carlo simulation algorithm which is capable of simulating realistic one-dimensional (1D) stratigraphic profiles with attached physical properties (de Groot, 1995). The acoustic properties of these 'pseudo-wells' are used to generate synthetic seismograms. The simulated data contains information on stratigraphic entities at different scale levels, attached physical properties and related seismic signals. Neural networks are trained on different datasets representing different geological settings to establish relationships between seismic response and underlying well properties. The influence of architecture and type of activation functions on network performance are examined for different paradigms.

In the first set of experiments the network design, i.e. number nodes in the input and hidden layer and the type of activation function, is varied for multi-layer perceptron (MLP) and radial basis functions (RBF) networks. After each variation, the network performance on the test data set is measured. In the second set of experiments, the network is fixed, but the geological model is made more complex by introducing new variables that affect the seismic response. In the third set of experiments the seismic bandwidth is varied. In the final experiment it is demonstrated how network performance can be increased by feeding the network with additional (non-seismic) information.

In the following section the network paradigms and activation funtions used in the experiments are decribed. This is followed by a description of the initial geological model. The experiments are presented next, followed by a discussion of the results and the conclusions.

[1]de Groot-Bril Earth Sciences BV, Boulevard 1945 nr. 24, 7511 AE Enschede, The Netherlands.

4.1 NETWORK PARADIGMS AND ACTIVATION FUNCTIONS

4.1.1 Multi-Layer Perceptrons

The MLP networks in this article are fully connected three- or four-layer networks with different types of activation functions used in the nodes of the hidden layer. For a discussion on MLP's and the so-called 'back-propagation' learning algorithm see e.g. Werbos (1974), LeCun (1985), Parker (1985) Rumelhart et.al. (1986) or Fahlman (1988). We will now describe the activation functions used in this article.

The mathematical expression of the perceptron (Rosenblatt, 1962) can be written as an activation function A applied to a weighting function W, defined as:

$$W(\mathbf{y}) = \sum_{i=0}^{L} w_i y_i, \tag{4.1}$$

where:
$\mathbf{y}$ is the neural network input vector written as y_i with $i = 1,\ldots,L$ and weighting vector w_i with $i = 1,\ldots,L$.

The activation function of the classical perceptron (Fig. 4.1a) can now be written in the following form:

$$A(W) = \begin{cases} 1 & W > 0 \\ 0 & W \leq 0 \end{cases}. \tag{4.2}$$

In MLP's the binary activation function is often replaced by a continuous function. The most widely used activation function is the sigmoid function (Fig. 4.1b). This function has the following form:

$$A(W) = \frac{2}{1+\exp(-W)} - 1. \tag{4.3}$$

Other activation functions used in this article are the linear, ramp and tangent hyperbolic functions. The linear function (Fig. 4.1c) is defined as:

$$A(W) = W. \tag{4.4}$$

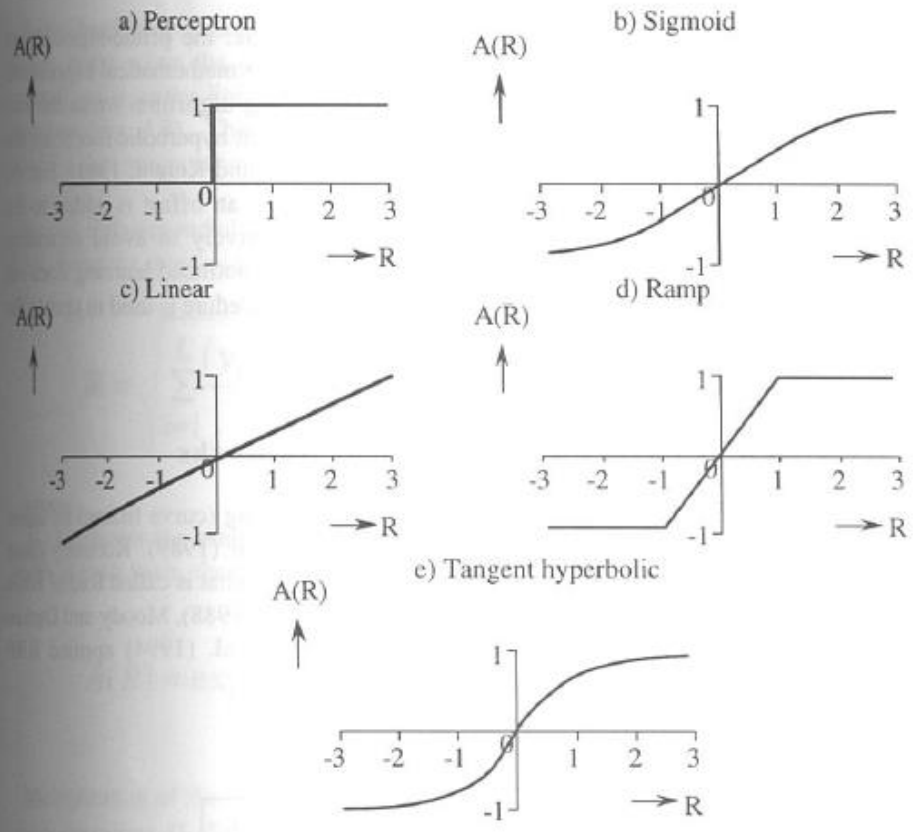


Fig. 4.1 Different activation functions for MLP networks. The Linear, Ramp, Sigmoid and Tangent hyperbolic functions have been used in this article.

The ramp function (Fig. 4.1d) is given by:

$$A(W) = \begin{cases} -1 & W < -1 \\ W & -1 \le W \le 1. \\ 1 & W > 1 \end{cases} \tag{4.5}$$

The tangent hyperbolic function (Fig. 4.1e) is written as:

$$A(W) = \frac{\exp(W) - \exp(-W)}{\exp(W) + \exp(-W)}. \tag{4.6}$$

Two other activation functions are used in this article: the prime-sigmoid an prime-tangent hyperbolic. These functions have the same mathematical expressio as equations (4.3) and (4.6), respectively. The training algorithm treats the tw types of functions differently. For the sigmoid and tangent hyperbolic functions, th derivative is used to update the weighting vector (Rich and Knight, 1991). For th prime-sigmoid and prime-tangent hyperbolic functions an offset is added to th absolute value of the derivative. This is done exclusively to avoid saturatio problems during learning, where saturation means that continued learning does no lead to improved network performance. This modified procedure is used to update th weighting vector.

4.1.2 Radial Basis Function Neural Networks

Radial basis functions have been used for data modelling (curve fitting) by many researchers, e.g. Powell (1987) and Poggio and Girossi (1989). Recently thes functions have been put in a neural network paradigm in what is called Radial Basi Function (RBF) Neural Networks (Broomhead and Lowe (1988), Moody and Darke (1988), Lee and Kil (1988), Platt (1991)). Schultz et.al. (1994) applied RBF networks in a seismic reservoir characterisation study.

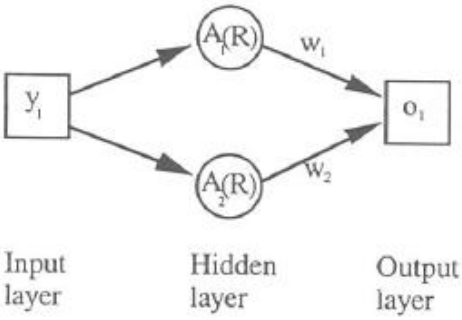


Fig. 4.2 Schematic representation of a Radial Basis Function network for the case of a single input variable, two basis functions and one output variable.

RBF networks have the same feed-forward layered architecture as MLP networks but the weighting function W and the activation function A are different. With RBF networks, there are only weights between output layer and hidden layer (Fig. 4.2). Each node in the hidden layer has a unique function, called the basis function. For the simple network of Fig. 4.2 with a single input, single output and two basis functions, the output is given by the sum of the two basis functions, each multiplied with its own weighting factor. In principle, any type of function can be used to act as basis function. For example, spline functions are used (Kavli, 1992), but the identification RBF network, applies only if radial basis functions are used.

Radial basis functions give local support to data points. The output of the hidden nodes, peaks when the input is near the centroid of the node, and then falls of symmetrically as the Euclidean distance between input and the centroid of a node increases (Fig. 4.3). The consequence of this behaviour is that RBF networks are good for data interpolation, but not good for data extrapolation.

Several different radial basis functions are in use, with the Gaussian function (Fig. 4.3a), being the most widely used. If the radial basis centre R is defined as:

$$R = \sqrt{\sum_{i=1}^{L} \frac{(y_i - \mu_i)^2}{\sigma_i^2}}, \tag{4.7}$$

where:
μ_i represents the centre location of each basis and σ_i indicates a scaling of the width of each basis, then the Gaussian activation function is given by:

$$A(R) = \exp\left(-\frac{R^2}{2}\right). \tag{4.8}$$

Multiplication of the activation function $A(R)$ with a weighting factor w then yields the output o (Fig. 4.2).

Another widely used RBF function is the so-called Inverse Multi-Quadratic Equation (IMQE, Fig. 4.3b), defined as:

$$A(R) = \frac{1}{\sqrt{R + k^2}}, \tag{4.9}$$

where:
k is an empirically determined smoothing factor. In this article a value of 0.5 has been used for k.

Note, that the widths in RBF functions are specified independently from each input dimension, making the functions elliptic rather than spherical. Note as well, that unlike the activation functions for MLP's no bias is included in the RBF functions.

Centre locations are typically determined by randomly selecting training examples from a large set of training data. The smoothing parameters and the number of nodes are typically adjusted empirically during training. RBF neural networks and MLP's

have been compared by many workers. Kavli (1992) reported consistently better performance of RBF networks in five independent experiments. Another important aspect when comparing RBF networks and MLP's is the training speed. RBF networks can be trained within a fraction of the time that is required for training MLP's. RBF networks, however, generally require more nodes to obtain similar performances.

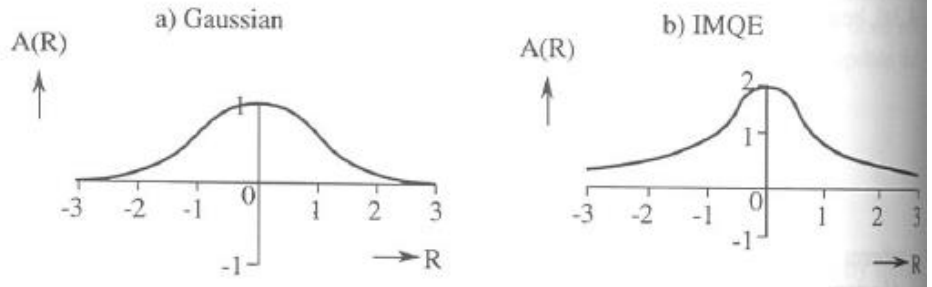


Fig. 4.3 Activation functions used in this article for RBF networks. The Gaussian function has a μ of 0 and a σ of 1. The IMQE function has a μ of 0, a σ of 1 and a k of 0.5.

The training algorithm used in this article for RBF's is the so-called HSOL algorithm (Lee and Kill, 1989, Carlin, 1992). HSOL uses standard back propagation for updating the function parameters, but this learning algorithm also dynamically allocates new nodes in the hidden layer during training.

4.2 INITIAL MODEL

The starting model represents a gas field, consisting of a sealing shale with constant acoustic properties overlying a carbonate reservoir (Fig. 4.4). The acoustic properties of this reservoir vary due to changing porosities and fluid content. The framework, i.e. the definition of the geological entities at different scale levels, of this model is given in Table 4.1.

Table 4.1

The integration framework for the initial model. The framework defines the geological entities at different scale levels.

Unit	Sub-unit	Lithology	Type	Code
Top	Marine	Shale	Seal	top.mar.shl
Bottom	Marine	Carbonate	Reservoir	bot.mar.car

The simulation specifications for this model are given in Table 4.2. The simulation algorithm generated 200 wells. The acoustic properties were used to create reflectivity logs. These were converted into synthetic seismograms by depth-time conversion, anti-alias filtering to 4 ms and convolution with a 30 Hz Ricker wavelet (Fig. 4.8). Sample impedance logs and corresponding synthetic seismic traces are shown in Fig. 4.5.

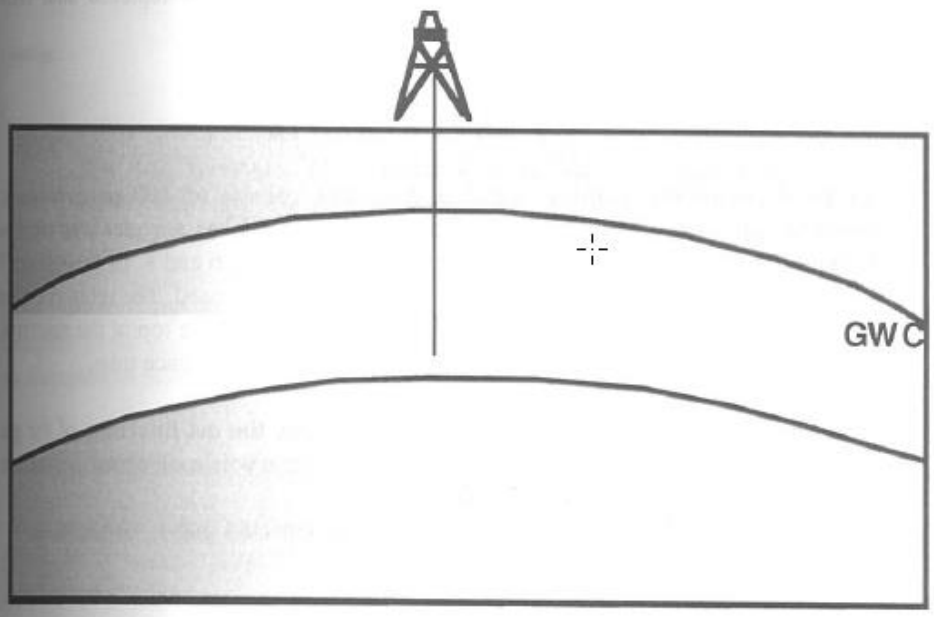


Fig. 4.4 Cross-section through the simulated field.

Table 4.2

Initial model simulation specifications. Probability density functions are specified as normal distributions with a mean and a standard deviation (n value1 value2) or as constants (c value).

Code	Thickness (m)	Sonic (us/m) *	Density (kg/m^3) *
top.mar.shl	c 91.4	c 377	c 2500
bot.mar	c 91.4		
bot.mar.car		n 278.9 11.5	n 2280 50
bot.mar.car.gas'		n 295.3 16.4	n 2100 100
Gas column "	n 15.2 15.2		

* Sonic and density distributions are correlated negatively (cor. coefficient=-1), the sonic distribution of the gas filled carbonate is correlated positively (cor. coefficient=1) with the sonic distribution of the brine filled carbonate.

' The acoustic properties of the carbonate reservoir depend on the fluid content.

" Maximum thickness = 45.6, minimum = 0. Values are repicked until these constraints are met.

4.3 EXPERIMENTS

In all experiments training and test data sets consist of 100 patterns each. Sampling rate of the seismic data is 4 ms. A 30 Hz Ricker wavelet was used to generate the synthetic seismic traces in experiments 1 until 6 and 8. In experiment 7 different Ricker wavelets (20,30,40,50 Hz frequency) were used. The reference time for selecting the seismic data was the time corresponding to the top of the reservoir. Selected seismic samples were interpolated relative to this reference time.

In all experiments, networks are trained to estimate the net thickness of the gas column and the average density of the gas-filled reservoir rock from the seismic response. The average density is calculated as:

$$\rho = \frac{\sum_{i=1}^{n} \rho_i \lambda_i}{\sum_{i=1}^{n} \lambda_i}, \tag{4.10}$$

where:

ρ is the density, λ the layer thickness, i the layer index and n the number of layers.

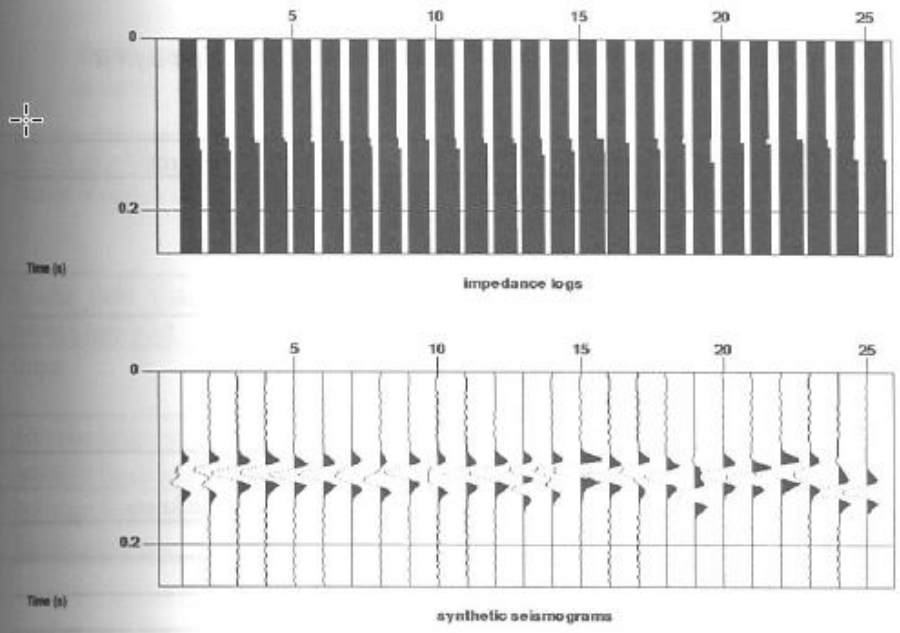


Fig. 4.5 Example of impedance logs and corresponding synthetic seismic traces of the initial model.

The network specifications are presented per set of experiments in a table. The performances of the test dataset are summarised in Table 4.12 for all experiments. Complete performance statistics and graphics are given in Appendix III. The results are discussed in Section 4.4.

4.3.1 Network design

In experiment 1 the network design is varied. The size of the input layer is progressively reduced from 25 in experiment A to 13, 7 and 1, in experiments B, C and D, respectively.

Table. 4.3
Experiment 1 network specifications. The size of the input layer is varied.

Network paradigm	Multi-Layer-Perceptron
Number of nodes: input-hidden-output	A) 25-3-2 B) 13-3-2 C) 7-3-2 D) 1-3-2
Seismic time gate relative to the reference time-pick (ms)	A) -25 - 75 B) -25 - 25 C) -12 - 12 D) 0 - 0
Output	average density net gas column thickness
Activation function input layer	none
Activation function hidden layer	sigmoid
Activation function output layer	linear
Training algorithm	backpropagation

In experiment 2 the size of the hidden layer is varied. In experiment A, the size of the hidden layer is first increased to 9. Then in experiment B, the size is reduced to 1 node. In experiment C, the network size is increased again with the introduction of a second hidden layer. The results of these experiments can be compared with the results of experiment 1A where three nodes were used in the hidden layer.

In experiment 3 Radial Basis Function networks are tested. In experiment A, a Gaussian activation function is used in the hidden layer and in experiment C the IMQE activation function is used. Training is stopped after 50.000 patterns. This is identical to other experiments, described in this chapter. Training is continued for another 150.000 patterns, because the RBF networks are still learning. The results of these prolonged training sessions for experiments A and C are reported in experiments B and E, respectively. In the RBF experiments, the HSOL training algorithm of Lee and Kill (1989) and Carlin (1992) is used.

In experiment 4, the final design experiment, different activation functions of the hidden layer are tested for MLP networks. In experiment A, a tangent hyperbolic function is used, followed by a prime tangent hyperbolic, a prime sigmoid, a ramp and a linear function, in experiments B, C, D and E, respectively. The results can be compared with the results of experiment 2A, where a sigmoid activation function is used in the hidden layer in a similar network configuration.

Table. 4.4

Experiment 2 network specifications. The size of the hidden layer is varied.

Network paradigm	Multi-Layer-Perceptron
Number of nodes: input-hidden-output	A) 25-9-2 B) 25-1-2 C) 25-9-3-2
Seismic time gate relative to the reference time-pick (ms)	-25 to 75 ms
Output	average density net gas column thickness
Activation function input layer	none
Activation function hidden layer	sigmoid
Activation function output layer	linear
Training algorithm	backpropagation

Table. 4.5

Experiment 3 network specifications. Radial Basis Functions; activation functions are varied along with the number of patterns trained.

Network paradigm	Radial Basis Functions
Number of nodes: input-hidden-output	25-3-2 (end of training 25-3-2)
Seismic time gate relative to the reference time-pick (ms)	-25 - 75
Output	average density net gas column thickness
Activation function input layer	none
Activation function hidden layer	A) Gaussian B) Gaussian C) IMQE D) IMQE
Activation function output layer	linear
Training algorithm	HSOL A) 50.0000 patterns trained B) 200.000 patterns trained C) 50.0000 patterns trained D) 200.000 patterns trained

Table. 4.6
Experiment 4 network specifications. MLP network with varying activation functions in the hidden layer.

Network paradigm	Multi-Layer-Perceptron
Number of nodes: input-hidden-output	25-9-2
Seismic time gate relative to the reference time-pick (ms)	-25 to 75 ms
Output	average density net gas column thickness
Activation function input layer	none
Activation function hidden layer	A) tangent hyperbolic B) prime tangent hyperbolic C) prime sigmoid D) ramp E) linear
Activation function output layer	linear
Training algorithm	backpropagation

4.3.2 Increasing the geological complexity

In the following experiments the complexity of the geological model is increased by introducing new variables that affect the seismic response. The network specifications are kept constant in these experiments (Table 4.8).

In the first geological complexity experiment (Experiment 5), shale intercalations are introduced into the carbonate reservoir. The framework is given in Table 4.7 and the simulation specification in Table 4.9. Examples of impedance logs with corresponding seismic responses are shown in Fig. 4.6.

Table 4.7
The integration framework for the carbonate-shale model.

Unit	Sub-unit	Lithology	Type	Code
Top	Marine	Shale	Seal	top.mar.shl
Bottom	Marine	Carbonate	Reservoir	bot.mar.car
		Shale	Waste	bot.mar.shl

Table 4.8
Network specification experiments 5, 6 and 7.

Network paradigm	Multi-Layer-Perceptron
# of nodes: input-hidden-output	25-9-2
Input time gate	-25 to 75 ms
Output	average density net gas column thickness
Activation function input layer	none
Activation function hidden layer	tangent hyperbolic
Activation function output layer	linear
Training algorithm	backpropagation

Table 4.9
Carbonate-shale simulation specification. Probability density functions are specified as normal distributions with a mean and a standard deviation (n value1 value2) or as constants (c value).

Code	Thickness (m)	Sonic (us/m) *	Density (kg/m^3) *
top.mar.shl	c 91.4	c 377.3	c 2500
bot.mar	c 91.4		
bot.mar.car	n 9.14 3.05	n 278.9 11.5	n 2280 50
bot.mar.car.gas '		n 295.3 16.4	n 2100 100
bot.mar.shl	n 1.5 1.5	n 360.9 6.6	n 2550 50
Gas column "	n 15.2 15.2		

* Sonic and density distributions are correlated negatively (cor. coefficient=-1); the sonic distribution of the gas filled carbonate is correlated positively (cor. coefficient=1) with the sonic distribution of the brine filled carbonate.

' The acoustic properties of the carbonate reservoir depend on the fluid content.

" Maximum thickness = 45.6, minimum = 0. Values are repicked until these constraints are met.

In experiment 6 the complexity of the model is increased even further. Not only are shale layers affecting the seismic response, also the acoustic properties of the overburden are varied. The variation in acoustic properties of the overburden is even much larger than the variations within the target zone (Table 4.10). Examples of acoustic impedance logs with corresponding seismic responses are shown in Fig. 4.7B.

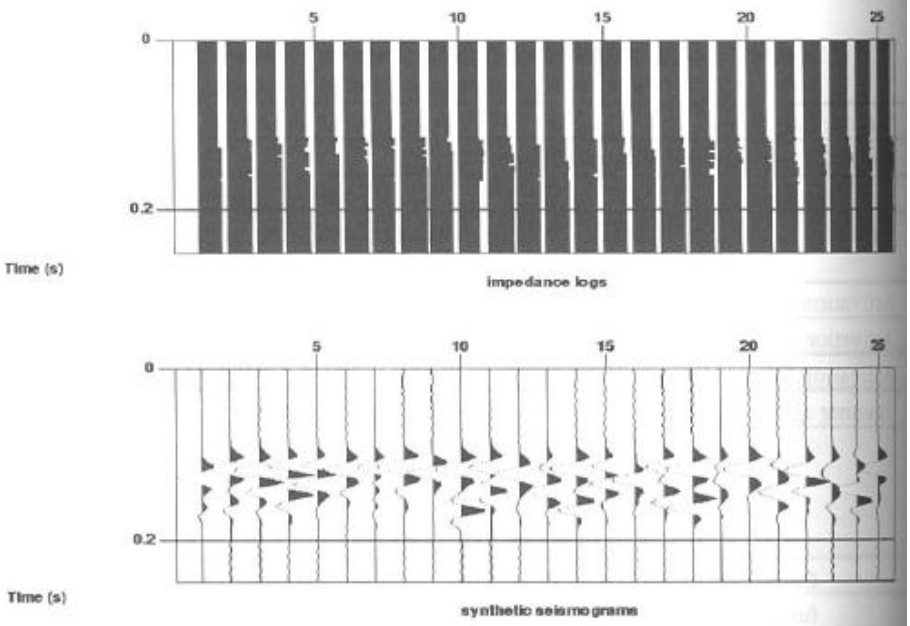


Fig. 4.6 Examples of acoustic impedance logs and corresponding synthetic seismic traces of the carbonate-shale model.

Table 4.10

Overburden model simulation specifications. Probability density functions are specified as normal distributions with a mean and a standard deviation (n value1 value2) or as constants (c value).

Code	Thickness (m)	Sonic (us/m) *	Density (kg/m^3) *
top.mar.shl	c 91.4	n 377.3 16.4	n 2500 500
bot.mar	c 91.4		
bot.mar.car	n 9.14 3.05	n 278.9 11.5	n 2280 50
bot.mar.car.gas '		n 295.3 16.4	n 2100 100
bot.mar.shl	n 1.5 1.5	n 360.9 6.6	n 2550 50
Gas column "	n 15.2 15.2		

* Sonic and density distributions are correlated negatively (cor. coefficient=-1); the sonic distribution of the gas filled carbonate is correlated positively (cor. coefficient=1) with the sonic distribution of the brine filled carbonate.

' The acoustic properties of the carbonate reservoir depend on the fluid content.

" Maximum thickness = 45.6, minimum = 0. Values are repicked until these constraints are met.

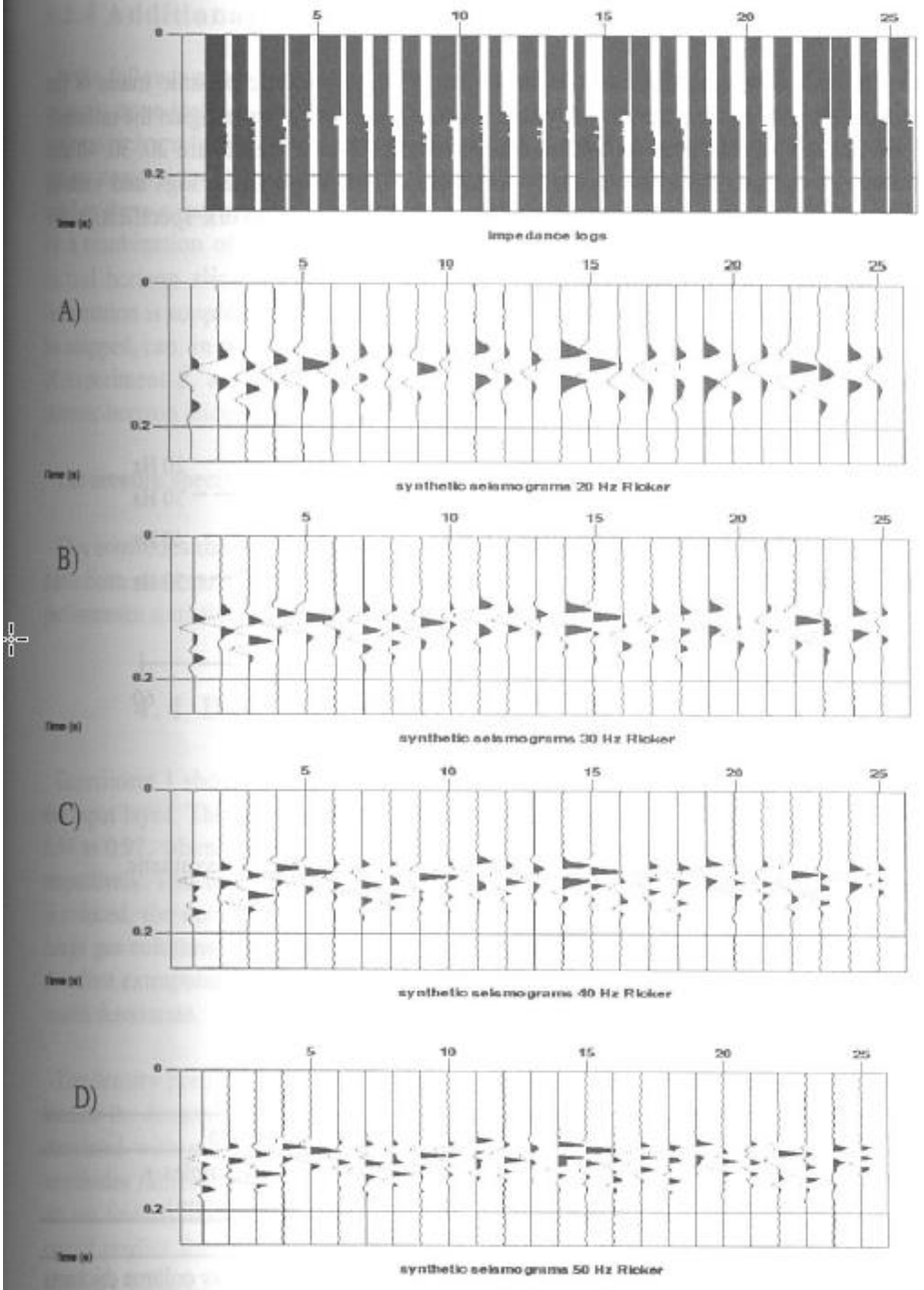


Fig. 4.7 Impedance logs of the overburden model with corresponding synthetic seismograms (A=20, B=30, C=40 and D=50 Hz Ricker wavelets).

4.3.3 Seismic bandwidth variations

In the following experiment (Experiment 7), the synthetic seismic traces of th overburden model are convolved with different wavelets to investigate the influenc of the seismic bandwidth on the inversion results. The wavelets are 20, 30, 40 an 50 Hz zero-phase Ricker wavelets (Fig. 4.8). The impedance logs and variou synthetic seismic traces are presented in Fig. 4.7. The network specification ar given in Table 4.8.

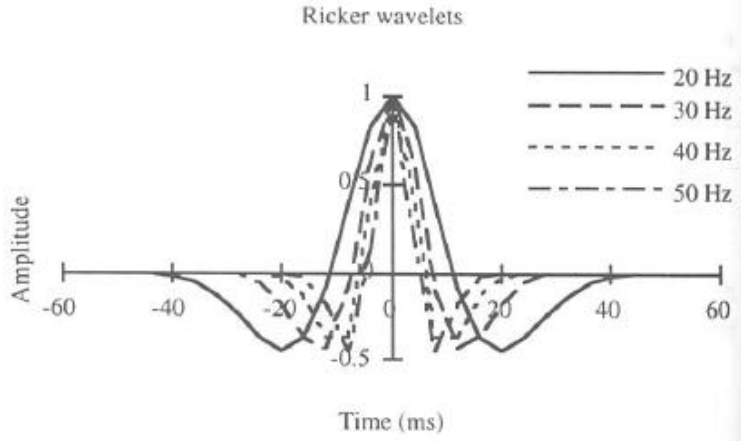


Fig. 4.8 Ricker type wavelets used to generate synthetic seismograms.

Table 4.11

Network specifications experiment 8.

Network paradigm	Multi-Layer-Perceptron
# of nodes: input-hidden-output	26-9-2 (25 seismic samples + gross gas-column)
Input time gate	-25 to 75 ms
Output	average density, net gas column thickness
Activation function input layer	none
Activation function hidden layer	tangent hyperbolic
Activation function output layer	linear
Training algorithm	backpropagation

4.3.4 Additional information

In the following experiment a carbonate-shale model with a constant overburden acoustic impedance is used (Table 4.9). The network input consists of 26 input nodes which are fed by 25 seismic samples and a gross gas-column thickness. There are no constraints to the design of networks in the GeoProbe[2] system which was used in these experiments. It is possible to design (and train) networks that are fed by a combination of seismic and well data. Application of such networks to the factual horizon slice, is possible as well. The only condition is, that the well information is supplied in the form of a XYZ grids. Therefore, any property that can be mapped, can, in principle, be used to constrain the inversion process. In the case of experiment 8, a gross gas-column grid should be supplied together with the seismic horizon slice.

The network specification of experiment 8 is presented in Table 4.11.

This concludes the description of the experiments with simulated data. The results have been summarised in Table 4.12 and will be discussed hereafter. Additional performance statistics are presented in Apendix III.

4.4 DISCUSSION OF THE RESULTS

Experiment 1 shows that gas column thickness prediction depends on the size of the input layer. The normalised RMS error increases from 0.33 through 0.44 and 0.54 to 0.97, when the number of input nodes is reduced from 25 to 13, 7 and 1, respectively. This behaviour can be explained as follows: when the number of nodes is reduced, the network is offered a smaller time-gate around the reference time. Large gas columns will have a seismic effect outside this gate. The network must therefore extrapolate, rather than interpolate the data. Consequently the prediction results deteriorate.

The density prediction depends much less on the size of the input layer. This is because the density variations (= impedance variations, since density and sonic were correlated with a -1 correlation coefficient in the simulation) affect the seismic amplitudes rather than the waveform. Because of tuning effects, seismic amplitudes are not linearly related to density, however. For this reason, a one-node network cannot predict the density as well as the larger-size networks, normalised RMS errors are 0.28, 0.29, 0.32 and 0.51 for 25, 13, 7 and 1 input nodes.

---

[2] de Groot-Bril Earth Sciences' proprietary seismic reservoir characterisation software package.

In experiment 2 the size of the hidden layer was varied. These results can be compared with the results of experiment 1A. It is clear that with one node in the hidden layer, experiment 2B, two variables cannot be predicted simultaneously. The training performance graph (App. I) shows that training initially gives a reasonable result for the density prediction. When training is continued, the gas column prediction improves at the expense of the density prediction performance. The best results are obtained with 9 nodes in the hidden layer, with normalised RMS errors of 0.24 for density and 0.32 for gas column prediction. Adding another hidden layer, experiment 2C, deteriorates the prediction results to 0.31 and 0.39 for density and the gas column, respectively.

Table 4.12

Normalised RMS errors on the test datasets for the average density of the gas-filled carbonate rock and the net gas-column thickness.

Experiment	Norm. RMS Density	Norm. RMS Gas column
1A	0.28	0.33
1B	0.29	0.44
1C	0.32	0.54
1D	0.51	0.97
2A	0.24	0.32
2B	0.93	0.48
2C	0.31	0.39
3A	0.46	0.37
3B	0.33	0.30
3C	0.45	0.31
3D	0.29	0.19
4A	0.18	0.25
4B	0.18	0.27
4C	0.22	0.36
4D	0.13	0.44
4E	0.13	0.43
5	0.29	0.58
6	0.83	0.60
7A	0.87	0.61
7B	0.83	0.60
7C	0.85	0.66
7D	0.83	0.60
8	0.28	0.21

Radial Basis Functions networks were tested in experiment 3. These networks are normally applied to lower dimensional problems (up to 5 dimensions), where a strong correlation exists between the input variables. In these experiments a strong

correlation between the input variables does exist, but the dimension of the problem is much larger than in standard RBF applications. Still RBF networks give good results, especially for the prediction of the gas column thickness where a normalised RMS error of 0.19 is reached after prolonged training using the IMQE activation function. In comparison with MLP networks, RBF networks take longer to converge. The IMQE activation function performs better than the Gaussian function in these experiments.

In experiment 4 different activation functions in the hidden layer were tested. These experiments can be compared with experiment 2A, where the sigmoid function was used. Comparing experiments 2A and 4A shows that the tangent hyperbolic activation function performs better than the sigmoid function. Normalised RMS errors for density and gas column are 0.18 and 0.25 for tangent hyperbolic and 0.24 and 0.32 for sigmoid functions, respectively. Also the prime tangent hyperbolic function scores better (0.18 and 0.27) than the sigmoid and prime sigmoid (0.22 and 0.36). The ramp and linear functions score very well when predicting the density (both 0.13), but perform less well when predicting the more difficult gas column thickness (0.44 and 0.43, respectively). The good performance on the density can be explained by the fact that seismic amplitudes outside the tuning range are related linearly to the density.

In experiments 5 and 6, the geological model was made more complex. These results are compared with the results of experiment 4A.

Shale intercalations introduced in experiment 5 have a considerable effect on the gas column prediction. The normalised RMS error decreases from 0.25 in experiment 4A to 0.58 in experiment 5. This result can be explained by the fact that non-unique solutions exist in the carbonate-shale model space. The introduction of shale layers into the reservoir, will have a large effect on the net column thickness and the seismic response. These effects, however, are not necessarily related, and, depending on the random selections made in the Monte Carlo simulations, completely different seismic signals might be related to similar net gas columns. The seismic amplitudes are less affected by the introduction of shale intercalations. Moreover, the average density is corrected for thickness of individual layers [Eq. 4.10]. Therefore, the decrease in performance of the density prediction is not so significant (from 0.18 in experiment 4A to 0.29 in experiment 5).

In experiment 6 the geological model complexity was increased by introducing a variation in the acoustic impedance properties of the overburden. Such variations affect the seismic amplitudes and hence the prediction performance of the reservoir density. Performance decreases from 0.29 in experiment 5 to 0.83 in experiment 6. Please note, however, that the standard deviation of the acoustic properties of the overburden is almost a factor of 10 higher than the standard deviations of the acoustic properties of the carbonate rock (Table 4.6). Changes in seismic amplitude are, therefore, primarily caused by changes in the overburden, which are completely

independent of the property of interest. The performance of the gas column prediction is hardly affected by the variations in the overburden; a slight decrease in normalised RMS errors was observed: from 0.58 to 0.60.

In experiment 7 the seismic bandwidth was varied. The results show that the frequency content of the seismic data does not affect the performance of the density prediction. The normalised RMS errors are 0.87, 0.83, 0.85 and 0.83 for 20 Hz, 30 Hz, 40 Hz and 50 Hz Ricker wavelets, respectively. Since the density prediction depends primarily on the seismic amplitudes, this result is in line with the expectations. The performance of the gas column prediction was expected to be affected by the frequency content. The vertical resolution increases with increasing bandwidth, hence an increase in gas column prediction performance was anticipated with increasing frequency. The results do not show this increase. Instead the gas column prediction performance is independent of the frequency content. The normalised RMS errors are 0.61, 0.60, 0.66 and 0.60 for 20 Hz, 30 Hz, 40 Hz and 50 Hz Ricker wavelets, respectively. A possible explanation for this behaviour is that the property of interest, i.e. the gas column thickness, was calculated at the well data scale, while the inversion works on the seismic scale. In the up-scaling from well data to seismic data, the impedance logs were resampled to the seismic sampling rate. This non-linear transformation determines whether, or not, the seismic response is related to a well property.

In experiment 8 the gross gas-column was supplied to the network in addition to 25 seismic samples. These results are compared with the results of experiment 5 where the network was trained on seismic samples only. As expected the average density is predicted equally well by both networks (normalised RMS errors are 0.29 and 0.28 for experiment 5 and 8, respectively. Also, as can be expected, the prediction of the net gas-column thickness is far better for experiment 8 then for experiment 5 (normalised RMS errors are 0.21 and 0.58, respectively.

4.5 CONCLUSIONS

From the aforementioned experiments, the following conclusions are drawn:

- In order to avoid extrapolation of results, the seismic time-gate to be analysed must cover the response of the largest thickness, for thickness-related inversions.
- The size of the hidden layer should not be chosen too small.
- A one node hidden layer can predict one variable only.
- One hidden layer is sufficient.

- Performance of RBF networks is comparable to MLP networks. RBF networks performed slightly better on the thickness inversion and slightly worse on the density inversion.

- Convergence of RBF networks is slower than that of MLP networks.

- The tangent hyperbolic activation function gives the best overall prediction performance for MLP network. For RBF networks the IMQE activation function gave a better performance than the Gaussian function.

- Prediction performance of the linear and ramp activation functions is good for the density because this is a linear problem outside the tuning range.

- Network performance deteriorates when new variables are introduced which affect the seismic response and the target variable independently. In other words; performance deteriorates when non-unique solutions are introduced in the training set.

- Variations in the impedance of the overburden affect the seismic amplitudes and therefore the prediction performance of impedance-related properties of the target level.

- The introduction of new layers affect the seismic waveform and therefore the prediction performance of thickness-related properties of the target level.

- Performance of the predictions of density and gas-column thickness are independent of the seismic band-width.

- Performance of the prediction can be increased by supplying additional (non-seismic) information to the network.

REFERENCES

Broomhead, D.S. and Lowe, D. 1988. Multivariable functional interpolation and adaptive networks, *Complex Systems*, 2:231-355, 1988

Carlin, M., 1992. Radial Basis Function Networks and Nonlinear Data Modelling. *Proceedings of Neuro-Nimes'92, Neural Networks and their Application*, EC2, France, 1992, pp.623-633.

de Groot, P.F.M., 1995. Seismic reservoir characterisation employing factual and simulated wells. *PhD. thesis*, Delft University Press.

de Groot, P.F.M., Bril, A.H., Floris F.J.T, Campbell, A.E. Monte Carlo simulation of wells, *Geophysics*, in press.

Fahlman, 1988. An Empirical Study of Learning Speed in BackPropagation Network *Technical Report* CMU-CS-88-162, 1988.

Kavli, T.Ø., 1992. Learning Priciples in Dynamic Control. *PhD.Thesis* University Oslo, ISBN no. 82-411-0394-8.

LeCun, Y., 1985. Une procedure d'apprentissage pour réseau à seauil asymétrique (learning procedure for asymmetric threshold networks). *Proceedings of Cognitiv 85*, 599-604, Paris.

Lee, S. and Kil, R.M., 1988. Multilayer feedforward potential function network. *IEE International Conference on Neural Networks*, I-161 - I-171, San Diego, 1988.

Lee, S. and Kil, R.M., 1989. Bidirectional Continuous Associator Based On Gaussia Potential Function Network. *International Joint Conference on Neural Network* vol.1, 1989, pp.45-53.

Moody, J., and Darken, C.J., 1988. Learning with localized receptive fields, i *Proceedings of the 1988 Connectionist Models Summer School.* pp. 133-143 editors: Touretzky et.al., Morgan-Kaufman.

Parker, D.B., 1985. Learning-Logic, *Tech.Rep.*TR-47. MIT Center for Computational Research in Economic and Management Science, Cambridge, MA.

Platt, J., 1991. A resource-allocating network for function interpolation. *Neural Computation*, 3(2):213-225, 1991.

Poggio, T. and Girosi, F., 1989. A theory of networks for approximation and learning *Technical report*, Artificial Intelligence Laboratory, Massachusets Institute of Technology, Jul. 1989.

Powel, M.J.D., 1987. Radial basis functions for multivariable interpolation: A review, in *Algorithms for Approximation*. editors: Mason, J.C., and Cox, M.G., Clarendon Press, London.

Rich, E. and Knight, K., 1991. *Artificial Intelligence* second edition. McGraw-Hill, Inc.

Rosenblatt, F., 1962. *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Washington D.C., Spartan Books.

Rumelhart, D.E., Hinton, G.E., and Williams, R.J., 1986. Learning internal representations by error propagation, *Parallel Distributed Processing*. Editors: Rumelhart, D.E., McClelland, J.L. and the PDP Research group, 318-362. Cambridge, MA, MIT Press.

Schultz et.al.,1994. Seismic-guided estimation of log properties, Part 1: A data-driven interpretation methodology. *The Leading Edge*, May 1994; Part 2: Using artificial neural networks for nonlinear attribute calibration. The Leading Edge, June 1994; Part 3: A controlled study. The Leading Edge, July 1994.

Werbos, P.J., 1974. Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. *PhD thesis*, Harvard University, Cambridge, MA.

APPENDIX I

The results of the experiments are presented in this appendix. The performance statistics on training and test datasets are presented in one table per experiment. The performance on the test dataset, is presented in three figures, for each experiment. The left-hand figure shows the normalised RMS error as a function of training patterns. The solid line indicates the training performance of the average density variable, the dashed line indicates the net gas-column thickness. The middle- and right-hand figure show the network estimated values versus the target values of the test variables (average density and net-gas column thickness, respectively). In the experiments 1 until 5, these values have been scaled to a range between -1 and +1. In experiment 8 unscaled values are shown.

1A	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.28	13.47 kg/m^3	9.63 kg/m^3	51.48 kg/m^3
Gas column Training set	0.33	3.67 m	2.26 m	16.56 m
Density Test set	0.28	14.29 kg/m^3	11.20 kg/m^3	38.65 kg/m^3
Gas column Test set	0.33	3.16 m	2.17 m	13.40 m

1B	Normalised RMS	RMS	Mean Absolute	Max Absolu
Density Training set	0.29	14.33 kg/m^3	10.53 kg/m^3	57.95 kg/m^3
Gas column Training set	0.46	5.16 m	3.12 m	22.48 m
Density Test set	0.29	14.53 kg/m^3	10.92 kg/m^3	42.16 kg/m^3
Gas column Test set	0.44	4.22 m	2.66 m	18.54 m

1C	Normalised RMS	RMS	Mean Absolute	Max Absolu
Density Training set	0.32	14.53 kg/m^3	14.53 kg/m^3	14.53 kg/m^3
Gas column Training set	0.56	6.23 m	4.17 m	24.84 m
Density Test set	0.32	16.32 kg/m^3	12.10 kg/m^3	48.73 kg/m^3
Gas column Test set	0.54	5.19 m	3.60 m	21.09 m

1D	Normalised RMS	RMS	Mean Absolute	Max Absolu
Density Training set	0.54	26.15 kg/m^3	20.04 kg/m^3	104.10 kg/m^3
Gas column Training set	0.98	10.95 m	9.03 m	28.47 m
Density Test set	0.51	25.69 kg/m^3	20.49 kg/m^3	75.35 kg/m^3
Gas column Test set	0.97	9.36 m	8.08 m	20.31 m

2A	Normalised RMS	RMS	Mean Absolute	Max Absolu
Density Training set	0.22	10.91 kg/m^3	7.95 kg/m^3	42.51 kg/m^3
Gas column Training set	0.31	3.43 m	2.09 m	15.79 m
Density Test set	0.24	12.10 kg/m^3	9.10 kg/m^3	33.50 kg/m^3
Gas column Test set	0.32	3.06 m	1.92 m	12.62 m

2B	Normalised RMS	RMS	Mean Absolute	Max Absolu
Density Training set	0.92	44.85 kg/m^3	35.35 kg/m^3	114.50 kg/m^3
Gas column Training set	0.44	4.92 m	3.56 m	19.08 m
Density Test set	0.93	46.94 kg/m^3	38.34 kg/m^3	119.23 kg/m^3
Gas column Test set	0.48	4.60 m	3.51 m	16.49 m

2C	Normalised RMS	RMS	Mean Absolute	Max Absolu
Density Training set	0.30	14.71 kg/m^3	10.65 kg/m^3	53.70 kg/m^3
Gas column Training set	0.41	4.59 m	2.84 m	18.86 m
Density Test set	0.31	15.65 kg/m^3	12.45 kg/m^3	41.98 kg/m^3
Gas column Test set	0.39	3.81 m	2.51 m	16.08 m

3A	Normalised RMS	RMS	Mean Absolute	Max Absolu
Density Training set	0.46	22.51 kg/m^3	16.24 kg/m^3	70.51 kg/m^3
Gas column Training set	0.45	4.97 m	3.05 m	21.44 m
Density Test set	0.46	23.28 kg/m^3	17.59 kg/m^3	85.49 kg/m^3
Gas column Test set	0.37	3.62 m	2.36 m	15.30m

3B	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.32	15.48 kg/m^3	11.22 kg/m^3	66.20 kg/m^3
Gas column Training set	0.32	3.58 m	2.39 m	21.52 m
Density Test set	0.33	16.49 kg/m^3	11.55 kg/m^3	78.70 kg/m^3
Gas column Test set	0.30	2.86 m	1.88 m	15.58 m

3C	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.47	22.75 kg/m^3	17.22 kg/m^3	84.07 kg/m^3
Gas column Training set	0.35	3.92 m	2.56 m	20.88 m
Density Test set	0.45	22.67 kg/m^3	17.75 kg/m^3	64.95 kg/m^3
Gas column Test set	0.31	3.04 m	2.30 m	12.37 m

3D	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.29	14.05 kg/m^3	10.48 kg/m^3	43.84 kg/m^3
Gas column Training set	0.23	2.54 m	1.44 m	19.50 m
Density Test set	0.29	14.48 kg/m^3	11.15 kg/m^3	40.34 kg/m^3
Gas column Test set	0.19	1.81 m	1.33 m	7.86 m

4A	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.17	8.14 kg/m^3	6.29 kg/m^3	29.49 kg/m^3
Gas column Training set	0.21	2.33 m	1.53 m	10.28 m
Density Test set	0.18	9.02 kg/m^3	7.08 kg/m^3	21.72 kg/m^3
Gas column Test set	0.25	2.45 m	1.55 m	14.09 m

4B	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.18	8.55 kg/m^3	6.72 kg/m^3	25.97 kg/m^3
Gas column Training set	0.24	2.72 m	1.92 m	9.15 m
Density Test set	0.18	9.01 kg/m^3	7.04 kg/m^3	25.95 kg/m^3
Gas column Test set	0.27	2.63 m	1.74 m	16.71 m

4C	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.21	10.13 kg/m^3	7.57 kg/m^3	40.13 kg/m^3
Gas column Training set	0.35	3.93 m	2.46 m	14.96 m
Density Test set	0.22	10.97 kg/m^3	8.20 kg/m^3	28.49 kg/m^3
Gas column Test set	0.36	3.49 m	2.18 m	15.23 m

4D	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.13	6.32 kg/m^3	4.96 kg/m^3	17.33 kg/m^3
Gas column Training set	0.39	4.30 m	3.12 m	15.91 m
Density Test set	0.13	6.34 kg/m^3	5.01 kg/m^3	16.59 kg/m^3
Gas column Test set	0.44	4.24 m	2.87 m	26.92 m

4E	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.12	5.85 kg/m^3	4.73 kg/m^3	13.91 kg/m^3
Gas column Training set	0.40	4.42 m	3.05 m	17.76 m
Density Test set	0.13	6.51 kg/m^3	5.17 kg/m^3	18.18 kg/m^3
Gas column Test set	0.43	4.14 m	2.71 m	27.58 m

5	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.22	11.69 kg/m^3	8.70 kg/m^3	41.43 kg/m^3
Gas column Training set	0.50	5.17 m	3.93 m	14.40 m
Density Test set	0.29	14.47 kg/m^3	11.41 kg/m^3	41.64 kg/m^3
Gas column Test set	0.58	5.73 m	4.20 m	20.08 m

6	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.82	37.95 kg/m^3	29.19 kg/m^3	122.35 kg/m^3
Gas column Training set	0.53	5.76 m	4.54 m	14.85 m
Density Test set	0.83	41.73 kg/m^3	33.36 kg/m^3	112.32 kg/m^3
Gas column Test set	0.60	6.52 m	4.75 m	23.54 m

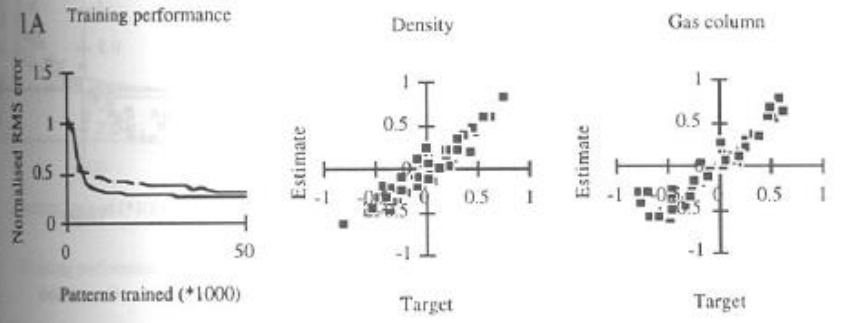
7A	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.80	37.11 kg/m^3	29.27 kg/m^3	118.03 kg/m^3
Gas column Training set	0.54	5.83 m	4.53 m	14.21 m
Density Test set	0.87	43.63 kg/m^3	34.53 kg/m^3	128.14 kg/m^3
Gas column Test set	0.61	6.66 m	4.84 m	26.51 m

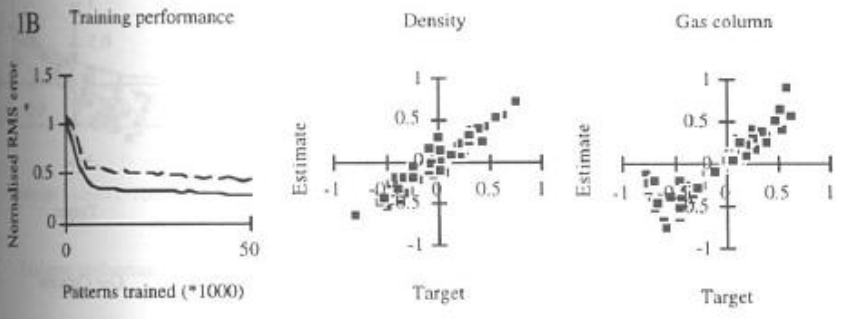
7B	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.82	37.95 kg/m^3	29.19 kg/m^3	122.35 kg/m^3
Gas column Training set	0.53	5.76 m	4.54 m	14.85 m
Density Test set	0.83	41.73 kg/m^3	33.36 kg/m^3	112.32kg/m^3
Gas column Test set	0.60	6.52 m	4.75 m	23.54 m

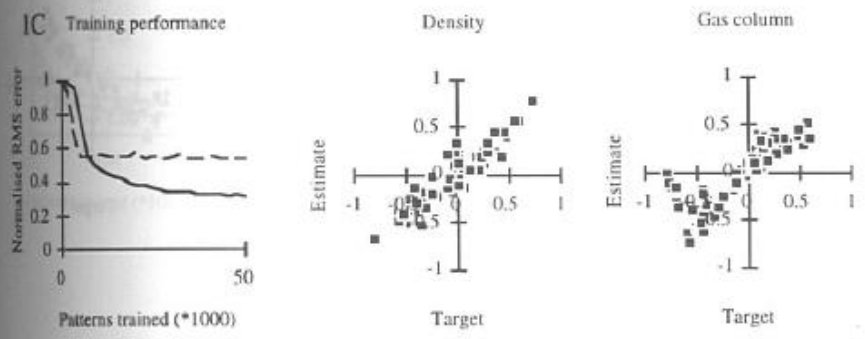
7C	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.82	37.93 kg/m^3	29.88 kg/m^3	114.71 kg/m^3
Gas column Training set	0.59	6.37 m	4.93 m	14.89 m
Density Test set	0.85	42.72 kg/m^3	34.26 kg/m^3	125.99 kg/m^3
Gas column Test set	0.66	7.23 m	5.38 m	22.21 m

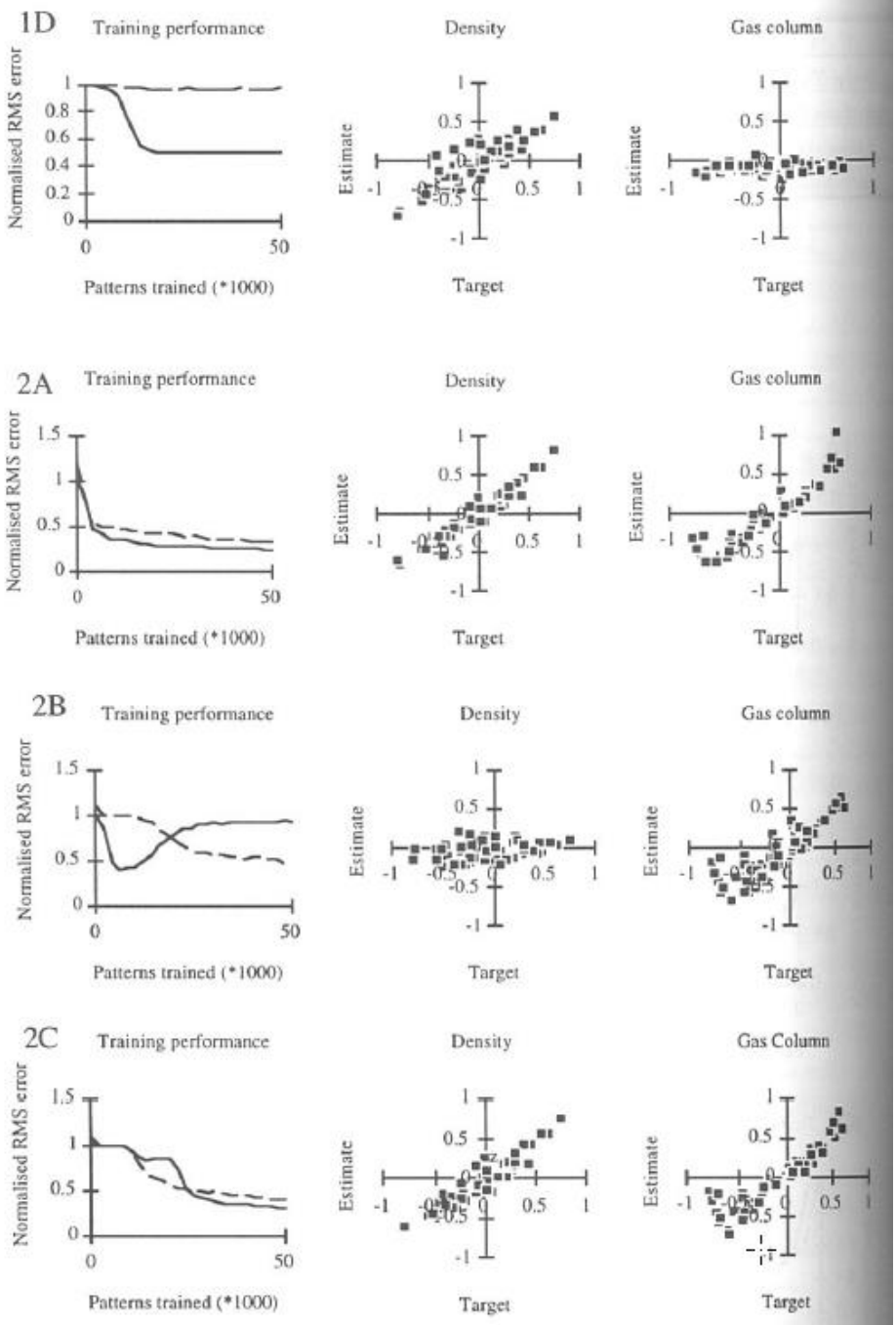
7D	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.82	37.95 kg/m^3	29.19 kg/m^3	122.35 kg/m^3
Gas column Training set	0.53	5.76 m	4.54 m	14.85 m
Density Test set	0.83	41.73 kg/m^3	33.36 kg/m^3	112.32 kg/m^3
Gas column Test set	0.60	6.52 m	4.75 m	23.54 m

8	Normalised RMS	RMS	Mean Absolute	Max Absolute
Density Training set	0.25	13.62 kg/m^3	10.37 kg/m^3	35.86 kg/m^3
Gas column Training set	0.13	1.34 m	0.91 m	5.69 m
Density Test set	0.28	14.26 kg/m^3	11.20 kg/m^3	40.19 kg/m^3
Gas column Test set	0.21	2.07 m	1.52 m	6.84 m

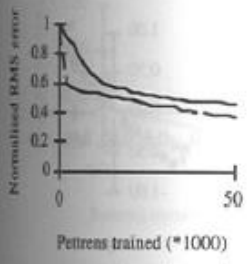




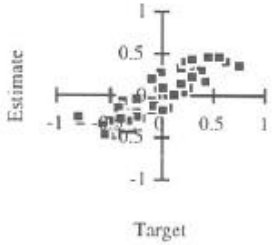



1D
Training performance
Normalised RMS error
Patterns trained (*1000)
Density
Estimate
Target
Gas column
Estimate
Target
2A
Training performance
Normalised RMS error
Patterns trained (*1000)
Density
Estimate
Target
Gas column
Estimate
Target
2B
Training performance
Normalised RMS error
Patterns trained (*1000)
Density
Estimate
Target
Gas column
Estimate
Target
2C
Training performance
Normalised RMS error
Patterns trained (*1000)
Density
Estimate
Target
Gas Column
Estimate
Target

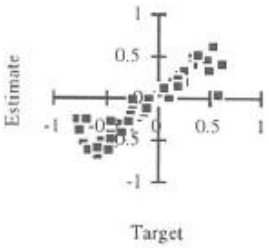
3A Training performance



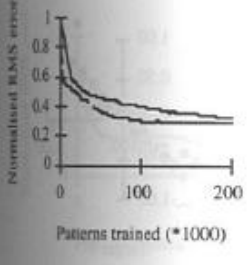
Density



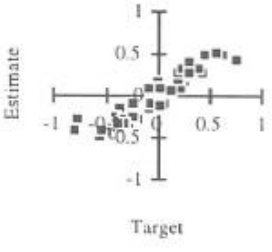
Gas column



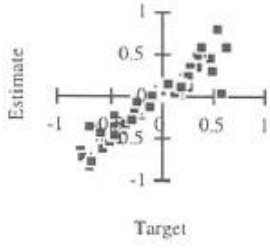
3B Training performance



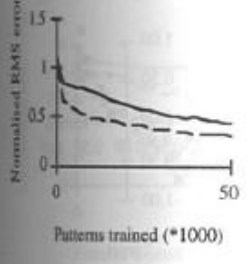
Density



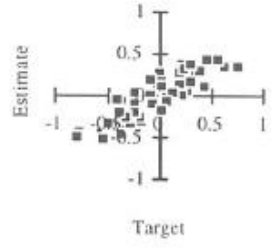
Gas column



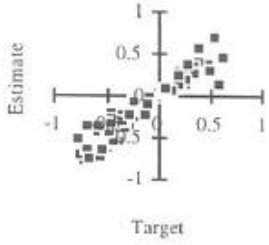
3C Training performance



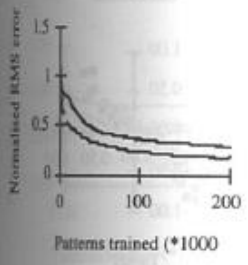
Density



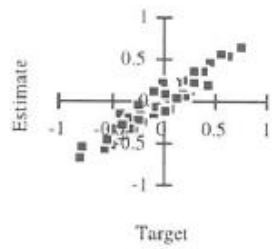
Gas column



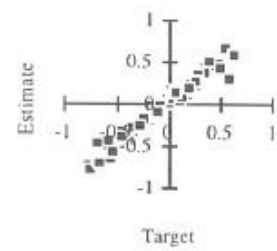
3D Training performance

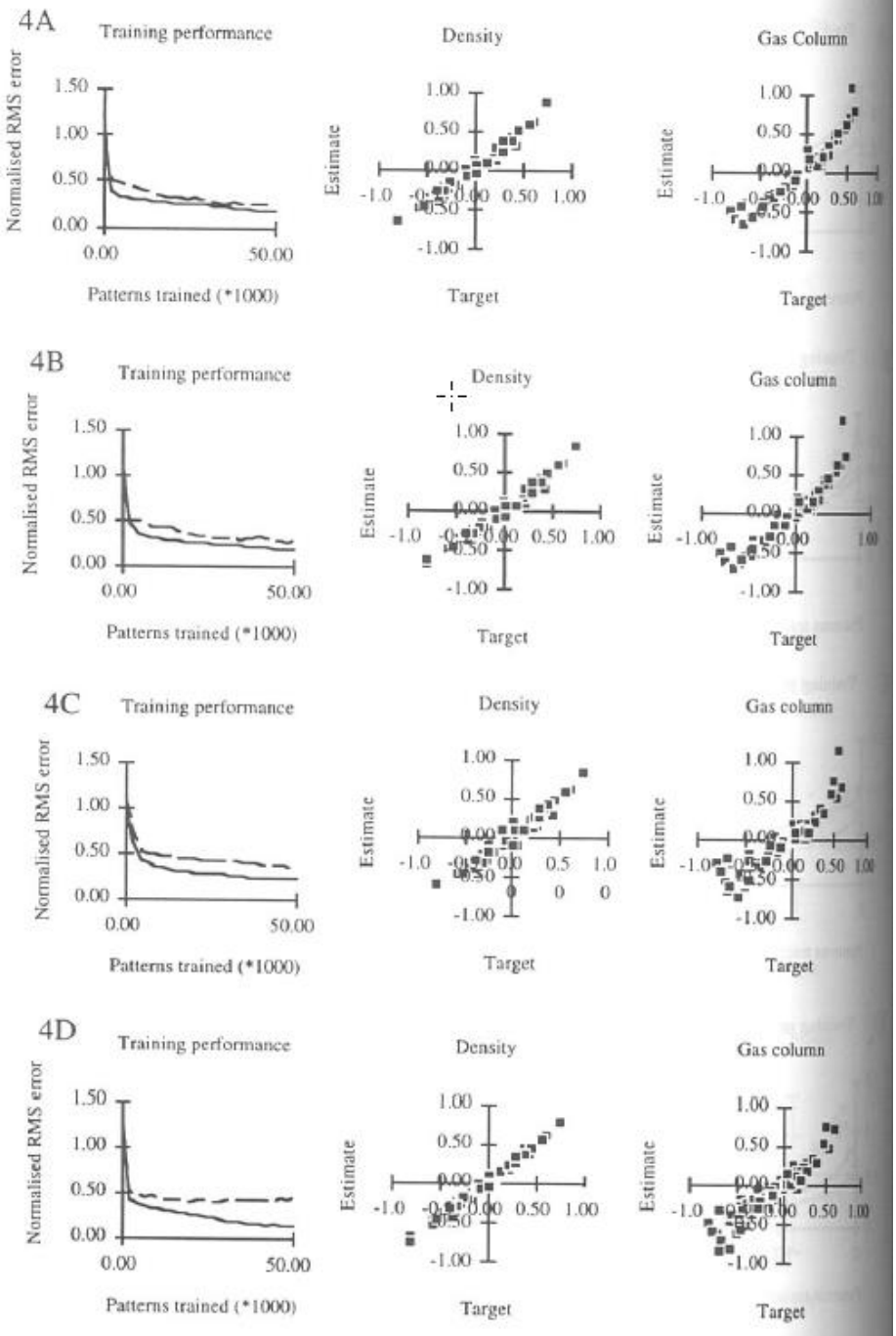


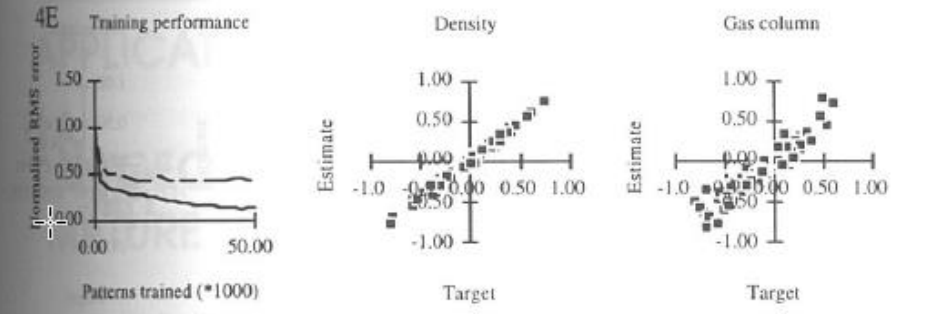
Density

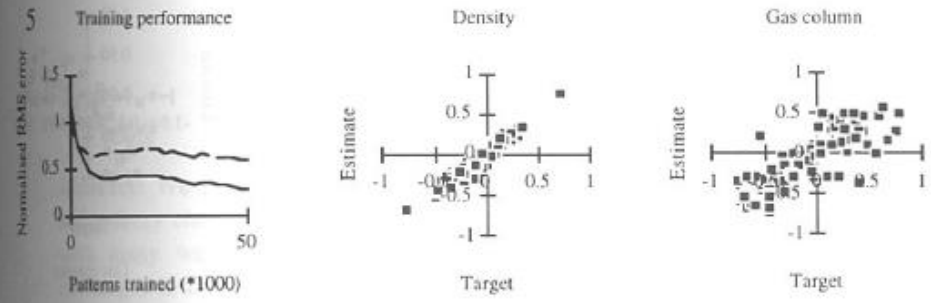


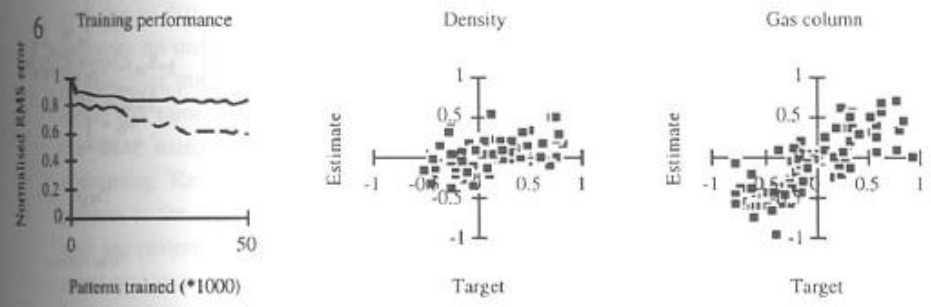
Gas column

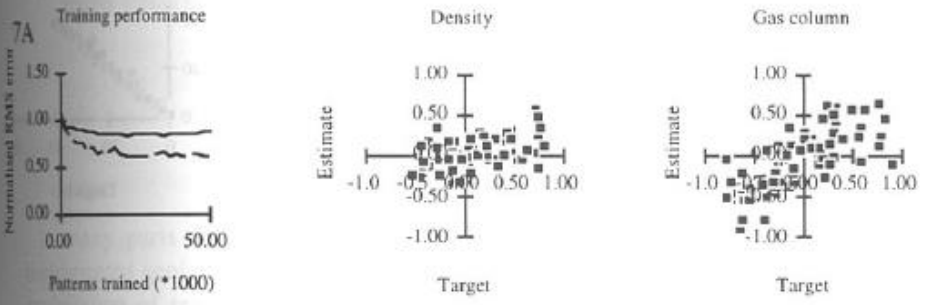


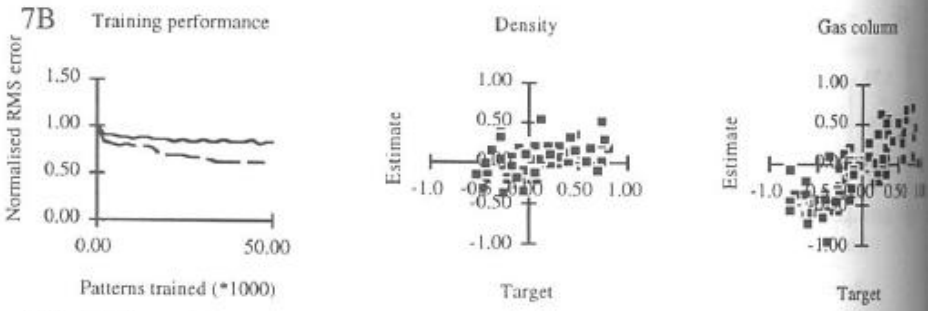

4A
Training performance
Normalised RMS error
1.50
1.00
0.50
0.00
0.00
50.00
Patterns trained (*1000)
Density
Estimate
1.00
0.50
0.00
-0.50
-1.00
-1.0
0.50
1.00
Target
Gas Column
Estimate
Target
4B
Training performance
Normalised RMS error
Patterns trained (*1000)
Density
Estimate
Target
Gas column
Estimate
Target
4C
Training performance
Normalised RMS error
Patterns trained (*1000)
Density
Estimate
Target
Gas column
Estimate
Target
4D
Training performance
Normalised RMS error
Patterns trained (*1000)
Density
Estimate
Target
Gas column
Estimate
Target

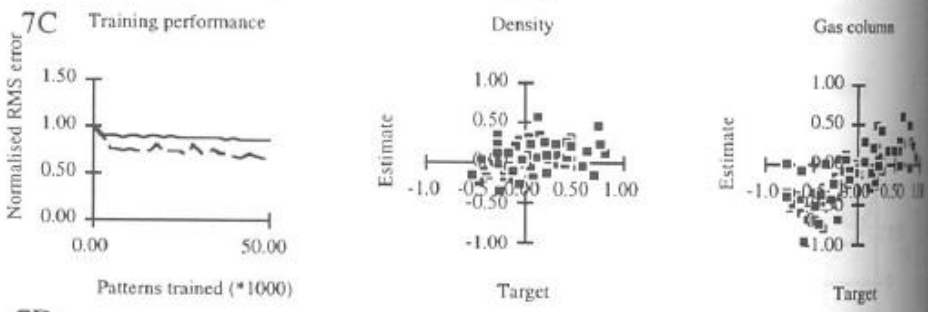

4E
Training performance
Normalised RMS error
1.50
1.00
0.50
0.00
0.00
50.00
Patterns trained (*1000)
Density
Estimate
1.00
0.50
0.00
-0.50
-1.00
-1.0
0.50
1.00
Target
Gas column
Estimate
1.00
0.50
0.00
-0.50
-1.00
-1.0
0.50
1.00
Target

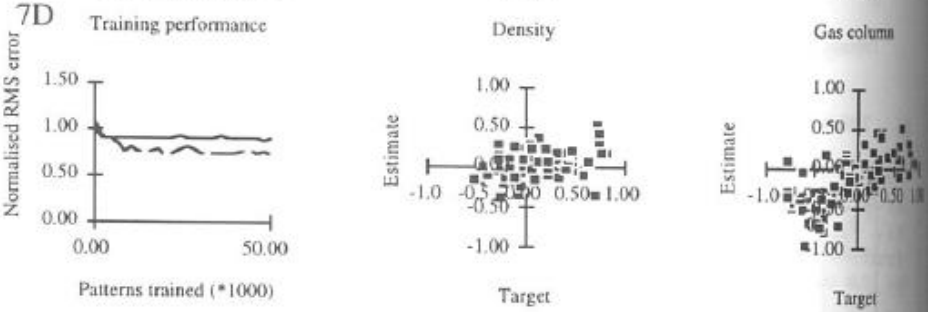

5
Training performance
Normalised RMS error
1.5
1
0.5
0
0
50
Patterns trained (*1000)
Density
Estimate
1
0.5
0
-0.5
-1
-1
0.5
1
Target
Gas column
Estimate
1
0.5
0
-0.5
-1
-1
0.5
1
Target

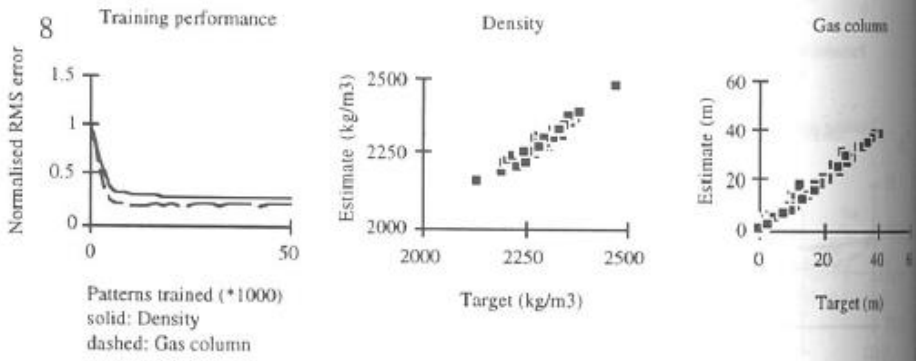

6
Training performance
Normalised RMS error
1
0.8
0.6
0.4
0.2
0
0
50
Patterns trained (*1000)
Density
Estimate
1
0.5
0
-0.5
-1
-1
0.5
1
Target
Gas column
Estimate
1
0.5
0
-0.5
-1
-1
0.5
1
Target


7A
Training performance
Normalised RMS error
1.50
1.00
0.50
0.00
0.00
50.00
Patterns trained (*1000)
Density
Estimate
1.00
0.50
0.00
-0.50
-1.00
-1.0
0.50
1.00
Target
Gas column
Estimate
1.00
0.50
0.00
-0.50
-1.00
-1.0
0.50
1.00
Target


7B
Training performance
Normalised RMS error
1.50
1.00
0.50
0.00
0.00
50.00
Patterns trained (*1000)
Density
Estimate
1.00
0.50
-0.50
-1.00
-1.0
0.50
1.00
Target
Gas column
Estimate
1.00
0.50
-1.00
Target


7C
Training performance
Normalised RMS error
1.50
1.00
0.50
0.00
0.00
50.00
Patterns trained (*1000)
Density
Estimate
1.00
0.50
-0.50
-1.00
-1.0
0.50
1.00
Target
Gas column
Estimate
1.00
0.50
-1.00
Target


7D
Training performance
Normalised RMS error
1.50
1.00
0.50
0.00
0.00
50.00
Patterns trained (*1000)
Density
Estimate
1.00
0.50
-0.50
-1.00
-1.0
0.50
1.00
Target
Gas column
Estimate
1.00
0.50
-1.00
Target


8
Training performance
Normalised RMS error
1.5
1
0.5
0
0
50
Patterns trained (*1000)
solid: Density
dashed: Gas column
Density
Estimate (kg/m3)
2500
2250
2000
2000
2250
2500
Target (kg/m3)
Gas column
Estimate (m)
60
40
20
0
0
20
40
Target (m)