Proceedings of the Institute of Acoustics

APPLICATION OF A NEURAL NETWORK TO ARRAY POSITION PROCESSING

A. A. Pitman

Underwater Studies, EASAMS Limited, Lyon Way, Frimley, Camberley, GU16 5EX

1. INTRODUCTION

1.1 Basis and Scope

The results presented on this paper are based upon investigations carried out by EASAMS into the use of Towed Seismic Arrays. Presentation of the results concentrates on the qualitative aspects of using Neural Networks for Array Position Processing.

This work has been carried out with the support of Procurement Executive, Ministry of Defence.

1.2 Background

Interest in the use of large aperture sensors derives from the higher gains and resolutions achievable with them. The use of towed arrays allows the construction of aperture sizes that would otherwise be impossible. However, the relative positions of the sensors must be known sufficiently accurately, typically to within one fifth of the wavelength of the sensed energy, to achieve a useful performance gain.

The typical sideways beam-forming scenario is shown in figure 1. Figure 1a shows a variety of factors which can influence the shape of a Towed Seismic Array in the lateral plane, which was the plane of interest for these investigations. Figure 1b shows a typical beam pattern that might be achieved in this situation if the senor locations were known precisely. This quickly degrades even with small errors. Figure 1c shows the beam pattern that might actually be achieved if no Array Position Processing were used when the array is disturbed.

Approaching the result shown in figure 1b requires knowledge of the Array Sensor Positions to within a tiny fraction of one percent of the array length. In addition, a variety of engineering constraints in the Seismic Array application required that Array Position Processing must be based upon indirect indications of position, generally heading sensor information. This is the problem that has been investigated.

1.3 Structure of Paper

In the section that follows the basic theory on which this investigation is based is presented. Section 3 describes the qualitative aspects of the results obtained. Chapter 4 seeks to explain these results in terms of the theory presented in section 2. Finally conclusions are drawn as to the relative merits of the approaches considered.

2. BASIC NETWORK THEORY

2.1 Approaches Considered

The characteristics of three different approaches were compared. The three approaches were:

- a linear Kalman Filter formulation. The basis of this approach is well known (e.g. [2]) and is not described further.

- a Multi-Layer Back-Propagation Network ([3] and [4]).

- a so-called Pi-Sigma Network. Pi-Sigma is a general term coined to cover the types of network described in [5] and [6], i.e. which share the common features of having one layer of weights and performing a polynomial expansion of the input vector (see later).

Both the Neural Networks investigated were feed-forward networks (see fig.s 2a and 2b) in which there is a uni-directional flow of information from input to output as opposed to Recurrent Networks (see fig. 2c) which are characterised by iteration in information flow.

Although different in formulation, the operation of both feed-forward networks is the same in that they seek to approximate a non-linear function between input and output. This is attempted by optimising the system of weights contained in the network using an optimisation method, which in neural network terms is called a learning or training rule.

It can be shown ([3], [4], [5] and [6]) that, provided there is a unique mapping between input and output (i.e. a given input always gives the same output), both these types of network can always approximate that function. The precision of the approximation is correlated with the network size in both cases (i.e. bigger networks are generally more precise).

2.2 Multi-Layer Back-Propagation Networks

As can be seen in figure 2a, the network output is calculated as the sum of a system of non-linear functions (in this case sigmoids) of the input.

Optimisation of the weights using the Back Propagation training rule allows the network to 'learn' a non-linear function between input and output. The principal features of these networks are described in [3] and [4] but the most noteworthy for the moment is that Back Propagation suffers from local optima in the training process.

2.3 Pi-Sigma Networks

Pi-Sigma Networks have one layer of inputs (a (possibly incomplete) polynomial expansion of the inputs of order N, i.e. multiplied together in combinations of 1 to N), one output layer and one layer of weights connecting every input to every output. 'Training' of such a network simply involves optimisation of the weights to produce a (multi-dimensional) polynomial curve fit to the non-linear input-output relationship. The main differentiations of this approach from Back Propagation are that training can be effected by the (simpler) Delta Rule and that the network is not subject to local optima [5].

3. REPORTING OF RESULTS

3.1 Basis of Investigations

Figure 3a shows a 'black box' view of the 3 approaches considered. The only difference between Neural Network and Kalman Filter black boxes is that the Neural Network generates an additional 'Novelty' output (which is not subsequently used as an input). This output gives an indication of the Network's 'familiarity' with the current input conditions.

These 'black boxes' were tested using a three dimensional hydrodynamic towed array model. The training data for the neural networks were also obtained from this model.

3.2 Qualitative Network Results

In figure 3b, the following qualitative characteristics of the Neural Network performance can be seen:

- The absolute output error is below that required for beam-forming for significant periods of time.

- The output error not obviously correlated with array shape.

- The output error is correlated with Novelty output which appears to set an upper bound to the output error.

- The output error does not increase monotonically with time.

3.3 Qualitative Comparison with Kalman Filter

In figure 3b, the following qualitative comparisons can be made between the performance of the Kalman Filter and Neural Networks:

- Across a wide range of array shapes absolute error magnitude for the Neural Network is generally smaller than for the Kalman Filter. However, at or near the straight array condition there is no significant difference.

- Kalman Filter performance deteriorates as network shape deviates from assumptions for which it is optimised, i.e. for small deviations from straight. Similarly, Neural Network performance sometimes deteriorates when network shape deviates from the shapes on which it was 'trained' (i.e. optimised).

3.4 Qualitative Comparison of Networks

Qualitatively the results for the two network types were very similar. However, certain differences were observed both in 'training' and in operation:

- Pi-Sigma trained significantly faster than Back Propagation Network.

- The performance obtained from the Back Propagation from training was not consistent, different start conditions leading to different results. Also during 'training' Back Propagation Network frequently encountered plateaus in performance, at levels worse than those eventually achieved, thereby prolonging training. These effects were not observed with the Pi-Sigma networks.

- In operation, Pi-Sigma networks tended to be more accurate than Back Propagation networks (on average).

4. DISCUSSION OF RESULTS

4.1 Network Results

Neural Networks appear a practical possibility as accuracies achieved are sufficient for beam-forming. The theoretical explanation for this result is that the network has learnt a non-linear function which allows it to map inputs to outputs in the way required to allow it to interpolate between and extrapolate from the training data used to derive it.

The non-linear function that has been learnt is of comparable complexity to the 'real-world' situation with no a priori knowledge of either the 'real world' or the simulation model. As such, these simulation results seem likely to be more than usually relevant to the practical situation as it seems quite possible that similar training might be similarly successful in the real world case.

Correlation of Network error with the Novelty network output rather than array shape appears to indicate that the limit of array shape that can be handled will be a function of available 'training' data rather than array shape.

Information content of measurements used appears to be sufficient to prevent 'drift' (i.e. errors building with time) as the Network error can be seen (figure 3b) to decrease without position update information.

4.2 Comparison with Kalman Filter

The operation of Neural Networks appears analogous to that of the Kalman Filter in that performance is best near the conditions at which they were optimised. The principal differences appear to be:

- The neural network can be optimised for many sets conditions rather than just one and can then interpolate and extrapolate from those optimisation points.

- The neural network is a non-linear formulation derived from empirical results without a priori knowledge of the system being estimated. In contrast the Kalman Filter is based (in this case at least) on a linear a priori model.

Additionally, the Novelty output from the Network is analogous to inspection of Kalman Covariance terms for determining the level of confidence in system output that is appropriate.

4.3 Comparison between Networks

The differences observed between the two types of network seem explicable in terms of:

- the different characteristics of network training error surface. Back Propagation is well known for suffering from local optima in its training, corresponding to the observed plateaus and occasional sub-optimal performance. Pi-Sigma Networks do not suffer from this [5].

- the differences in topology. Pi-Sigma Networks are really a set of networks mapping all the available inputs to each output independently. It seems it may be this independence between outputs which (in this application) allows it to out perform the Back Propagation Network.

5. CONCLUSIONS

The results of a simulation based investigation of Neural Networks as an Array Position Processing technique are presented. It is felt that these results are (if anything) more practically relevant than normal simulation studies as the Neural Network approach affords no possibility of 'biasing' the solution to the characteristics of a simulation model.

Neural networks (of the types described in this paper) appear to offer a capability analogous to that of the Kalman Filter in the Array Position Processing application. Their operation appears to be based upon the derivation of an optimal, non-linear process model from example system responses. In operation, Neural Networks appear to be able to interpolate and extrapolate from the example cases on which they are based to give, in most circumstances, better performance than the Kalman Filter.

The results obtained are broadly in keeping with those that would be expected from the theoretical foundations of the approaches considered. It is felt this indicates that Neural Networks may be considered in the mainstream of signal processing rather than as an area of 'black magic'.

6. REFERENCES

[1] D. Cragg, 'Report on Neural Networks Applied to Array Position Processing', EASAMS Internal Report L9019/NEURNET_APP/DC.

[2] A. Gelb (Ed.), 'Applied Optimal Estimation', The MIT Press 1974.

[3] A. Lapedes and R. Farber, 'How Neural Networks Work', Los Alamos National Laboratory, Internal Reference LA-UR-88-418.

[4] A. Lapedes and R. Farber, 'Non-Linear Signal Processing: Prediction and System Modelling', Los Alamos National Laboratory, Internal Reference LA-UR-87-2662.

[5] M.R. Lynch and P.J. Rayner, 'The Properties and Implementation of the Non-Linear Vector Space Connectionist Model', Proceedings of IEE Conference on Neural Networks, October 1989.

[6] C.L. Giles and T. Maxwell, 'Learning, Invariance and Generalisation in High-Order Neural Networks', Applied Optics Volume 26 Number 23, 1st December 1987.

APPLICATION OF A NEURAL NETWORK TO ARRAY POSITION PROCESSING

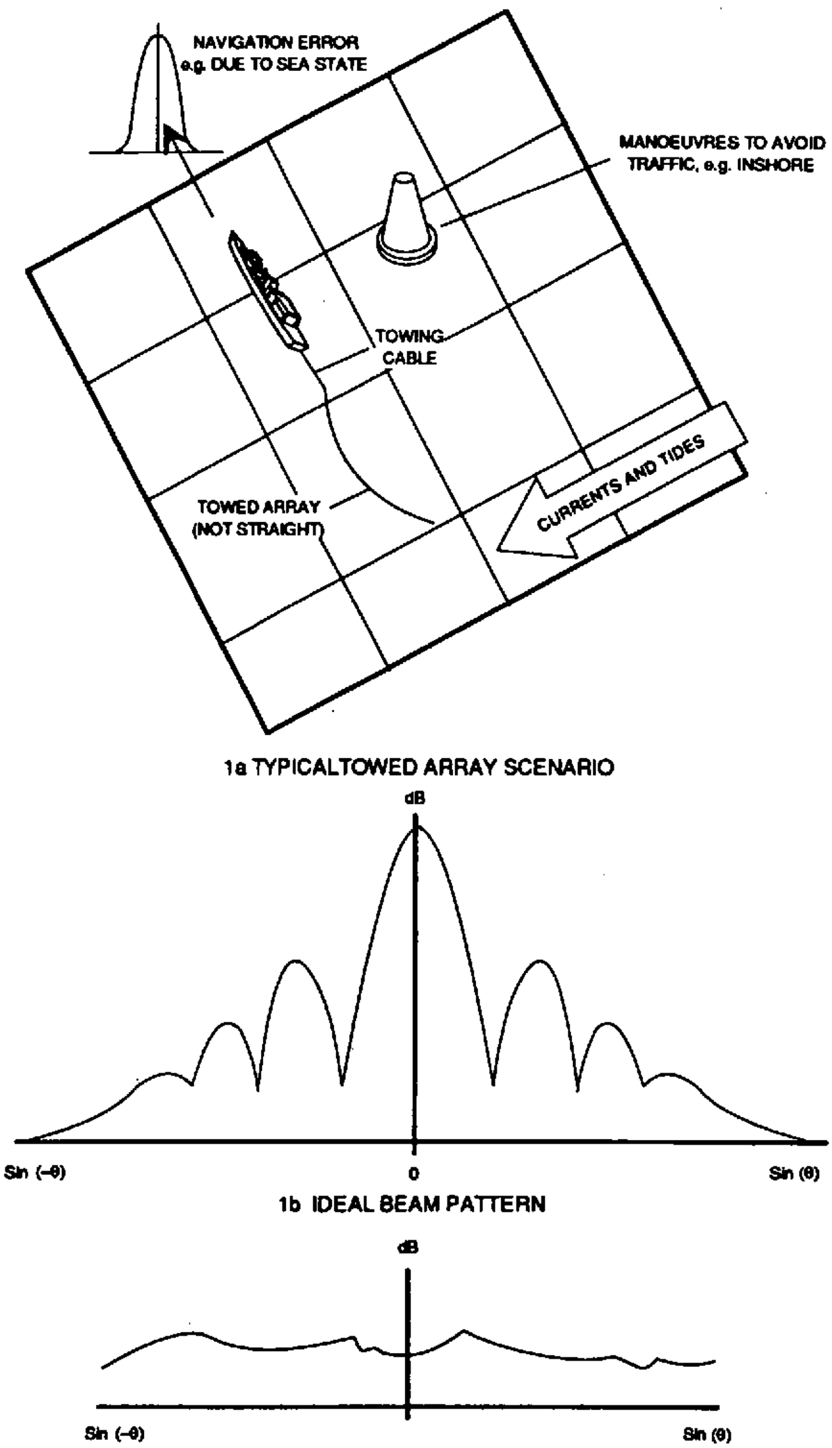
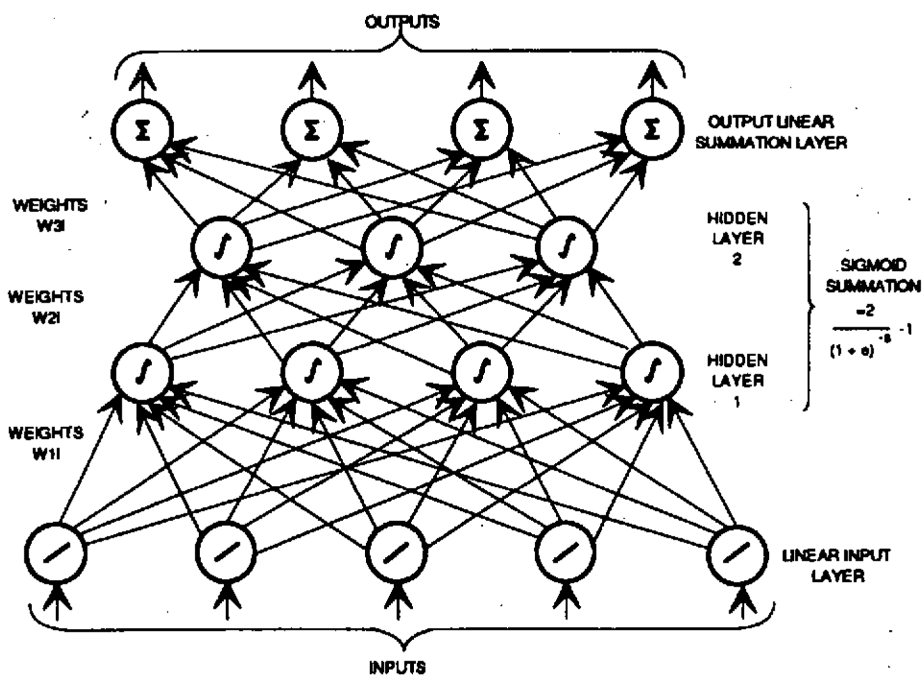


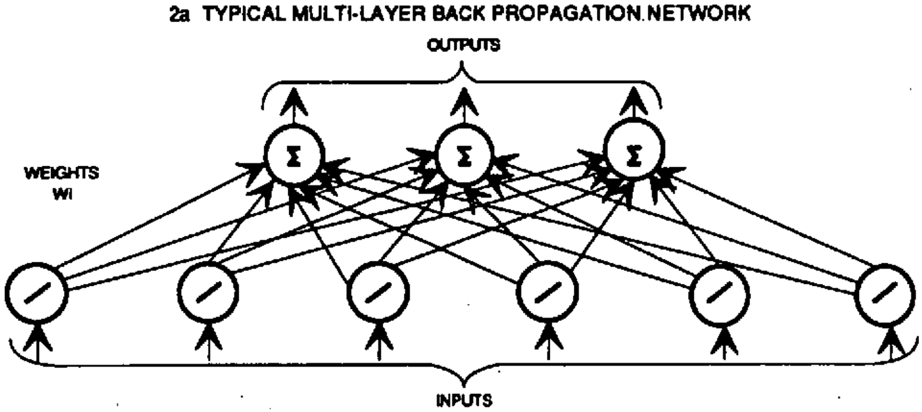
Fig.1 ARRAY POSITION PROCESSING PROBLEM

APPLICATION OF A NEURAL NETWORK TO ARRAY POSITION PROCESSING

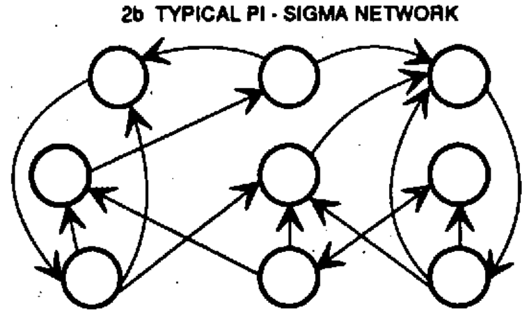
CG275/3



2a TYPICAL MULTI-LAYER BACK PROPAGATION NETWORK



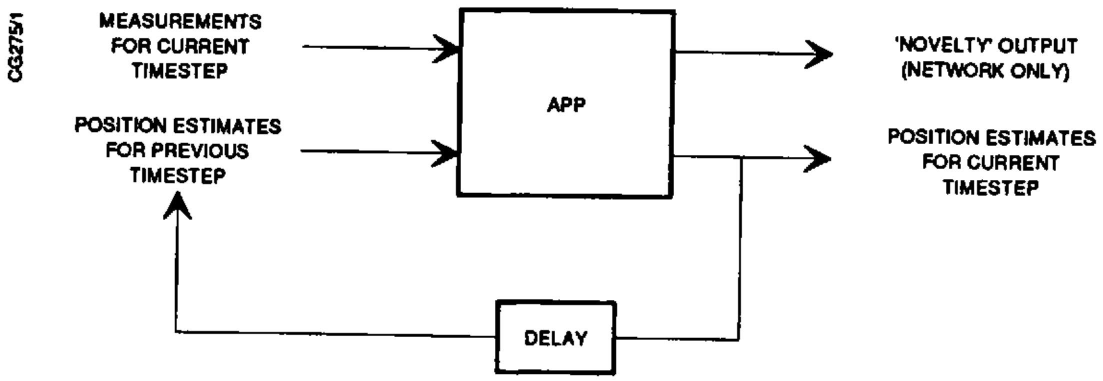
2b TYPICAL PI - SIGMA NETWORK



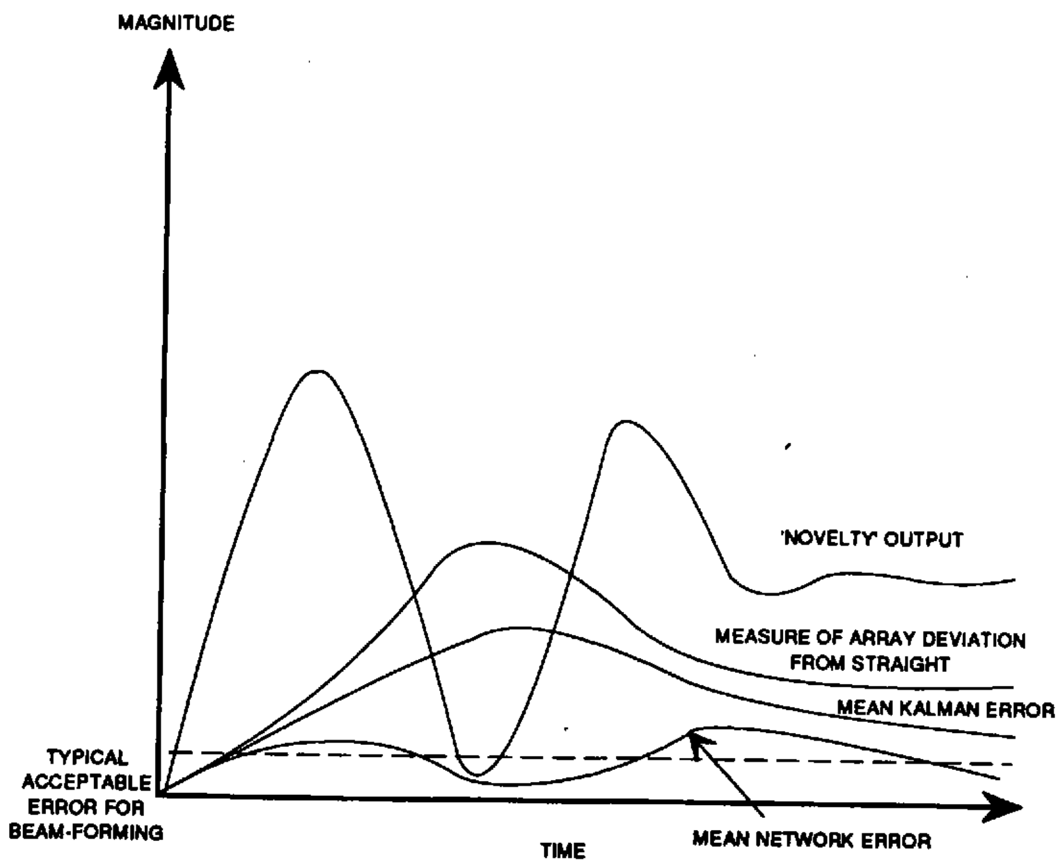
2c EXAMPLE RECURRENT NETWORK

Fig.2 EXAMPLE NETWORKS

APPLICATION OF A NEURAL NETWORK TO ARRAY POSITION PROCESSING



3a BLACK BOX VIEW OF APP



3b TYPICAL APP TIME HISTORY

Fig.3 QUALITATIVE RESULTS