# THE APPLICATION OF NEURAL NETWORKS TO ANODIC STRIPPING VOLTAMMETRY TO IMPROVE TRACE METAL ANALYSIS

HOWARD STEPHEN MANWARING

A thesis submitted in partial fulfilment of the requirements of the University of Hertfordshire for the degree of Doctor of Philosophy

This programme of research was carried out in the School of Information Sciences, University of Hertfordshire

in collaboration with

the School of Natural Sciences, University of Hertfordshire

and

West Herts College

**June 1995**

## ACKNOWLEDGEMENTS

The assistance of West Herts College is gratefully acknowledged for providing me with the time to carry out this research in addition to my other duties and for contributing towards the cost of the work. Thanks are also due to Brian Woodget and the Analytical Services staff at University of Hertfordshire for making laboratory facilities available .

Particular thanks are due to the following people.

Professor Bhinda PhD. CEng. FIMechE. Mem ASME. MRAes, Professor Emeritus, University of Hatfield, for convincing me that I could do this research by part time study.

Dr Barry Pepper, Vice Principal (Academic), West Herts College, for his interest, his time and for behaving more like my friend than my vice principal.

Ray Frank for his ideas, for his managing of the administration connected with this work and for making an effort to understand the chemistry.

Finally my wife, Marie Baker, for her support, her mathematical knowledge and for still recognising me when the computer saw more of me than she did.

---

"After a year's research, one realises that it could have been done in a day"
[Sir] William H Bragg

*In*: Ewald P. P., ed. Fifty years of X ray Diffraction, 1962, Utrecht.

# Abstract

This thesis describes a novel application of an artificial neural network and links together the two diverse disciplines of electroanalytical chemistry and information sciences.

The artificial neural network is used to process data obtained from a Differential Pulse Anodic Stripping (DPAS) electroanalytical scan and produces as an output, predictions of lead concentration in samples where the concentration is less than 100 parts per billion.

A comparative study of several post analysis processing techniques is presented, both traditional and neural. Through this it is demonstrated that by using a neural network, both the accuracy and the precision of the concentration predictions are increased by a factor of approximately two, over those obtained using a traditional, peak height calibration curve method. Statistical justification for these findings is provided

Furthermore it is shown that, by post processing with a neural network, good quantitative predictions of heavy metal concentration may be made from instrument responses so poor that, if using tradition methods of calibration, the analytical scan would have had to be repeated.

As part of the research the author has designed and built a complete computer controlled analytical instrument which provides output both to a graphical display and to the neural network. This instrument, which is fully described in the text, is operated via a mouse driven user interface written by the author.

# Contents

# LIST OF TABLES

# LIST OF ILLUSTRATIONS

# 1. Introduction and Project Aims

## 1.1. Introduction

### 1.1.1.Overview of the work

Environmental pollution, whether accidental or deliberate, has far reaching consequences and the determination of trace quantities of heavy metals in such places as foodstuffs, ground water and soils is an increasingly important issue. Determination of such levels in a laboratory remote from the source is time consuming and costly and so the trend is towards on site determination using portable instrumentation.

Many analytical techniques can be used to assess concentration levels of such substances as lead, cadmium and thallium. The electrochemical technique used in this work, Differential Pulse Anodic Stripping at a Thin Film Mercury Electrode, is one of a class of techniques collectively called Voltammetry. Using a special rotating electrode, it is fast, requires equipment that can easily be made portable and yet is capable of determining the concentration of electroactive species in solution down to the parts per billion level.

Artificial neural networks (the word artificial being used to distinguish them from biological neural networks) are computational devices, loosely based around the operation of the brain, which have the capability to learn. They have existed in the academic world of computing for some time but are only now receiving growing

attention as tools in real world applications. They are usually written in software but some hardware implementations do exist. Recent years have seen a steady increase in the number of application areas where they may be usefully employed.

A neural network has many interesting properties. One of these is the ability to form an arbitrarily complex mapping between some presented input data and some corresponding desired output data. Others include the ability to process noisy data and the ability to generalise, that is produce meaningful output when presented with input data that does not fall into one of its learnt mappings.

The research reported here links together these two rather different disciplines of analytical chemistry and artificial neural networks. The author, who cannot claim to have a background in chemistry, has applied the pattern recognition and generalisation power of artificial neural networks to the response obtained from a particular method of trace metal analysis. The detailed aim of the work is given in the next section but in general terms the work is an investigation into whether neural networks can help to provide a more accurate and precise prediction of trace metal concentration than is currently available using Differential Pulse Anodic Stripping with existing post analysis data processing techniques.

There are two areas where it is claimed improvements might be made in the existing technique. One is the use of neural networks to extract more information from the analytical data so obtained than is currently done, by using a neural network to post process the response obtained from the analytical instrument. A second area is the optimisation of the analytical parameters, in advance of an analysis, by using a neural network operating in a control loop. It will be shown that improvement in this latter area is not as necessary as it first appears given the results of investigations into post processing by neural network.

The work follows a traditional approach. As it was desired to investigate the control of the analysis itself it was necessary to be familiar with the operation of the equipment which it was my intention to control. To this end it was decided that rather than be bound by the constraints of a commercial instrument, all the necessary equipment to perform the chemical analysis and obtain the required data would be constructed by the author prior to any study of neural network architectures and their particular application to this area.

Accordingly the following items specific to the task were designed, built and tested by the author

1. The rotating Thin Film Mercury Electrode and analytical cell arrangement.

2. The electronic hardware providing potentiostatic control of the electrode and the current measuring facilities.

3. The electronic hardware required to provide remote computer control of the potentiostat and its numerous operational features.

4. A suite of computer software which can be used to set the operational parameters, carry out the required analytical functions, store the results and display and process the resulting voltammogram. This was written as both an interactive, mouse driven user interface and as a file oriented, batch processing interface which allowed automatic running.

A photograph showing the complete arrangement is reproduced in figure 1. Further photographs showing details of various parts are given in chapter 4.

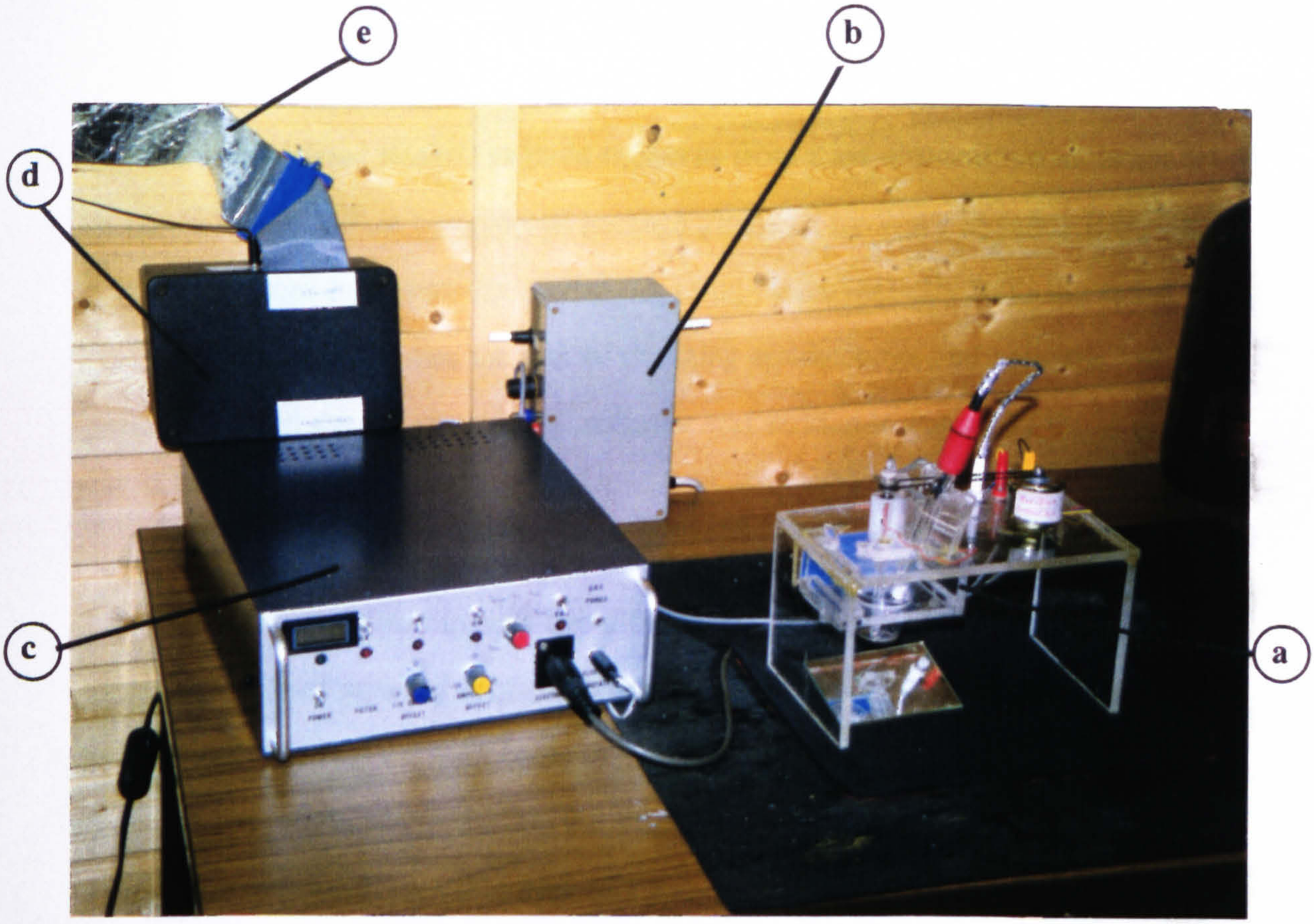



Figure 1. Photograph of the equipment built by the author for this study

Key: (a) analytical cell and working electrode holder, (b)Relay for electrode motor and electrically operated valve for gas purge, (c) Interfacing, control and potentiostat electronics, (d) 48 Opto isolators used to isolate (c) from the computer, (e) 50 Way ribbon cable leading to the computer

Following the testing of the analytical system the focus of the work shifted to a study of neural networks and their application to this area. Data drawn from various sources were used: synthetic data, real data obtained from the analysis of lead in curry powder and real data from prepared lead standards. Neural networks and traditional techniques were used with data obtained using the computerised instrument to predict lead concentrations in the parts per billion region.

Some of the neural networks used in this work to process the voltammograms obtained from analytical runs were written by the author in Pascal. Others were

written using Visual Basic© to build networks from a set of primitive tools available in NeuroWindows™, a commercial package produced by Ward Systems Inc[1]. The bulk of the work however was carried out using NeuroShell 2™ from the same company, which simplifies the task of changing architectures and network parameters. All the pre-processing software was written in Pascal by the author.

### 1.1.2. Overview of this thesis

As this thesis encompasses two different disciplines it was felt appropriate to offer a number of areas of background study. Readers with an information science background will find relevant background material on voltammetry and its development in Chapter 2, section 2.1 whereas chemists requiring information on neural networks are directed to section 2.2.

Chapter 3 brings these two areas together with a study of the existing applications of artificial intelligence in general and neural networks in particular, to the field of analytical chemistry. Sections 3.2 and 3.3 in this chapter follow this up with a justification of the aims of this work in the light of the current research.

Chapter 4 begins by describing the construction of the special electrode needed and the 'wet chemistry' analytical apparatus. The construction of the electronics which controls this apparatus is described next in section 4.2 followed in section 4.3 by a description of the interfacing software that enables a remote computer to drive the instrument.

The work relating to the major part of this study, that of post processing the instrument response by neural network begins in chapter five. This chapter details the experimental methods and electrochemical parameters used, the neural network architectures considered and the methods of data analysis employed. It first covers the preliminary work with curry powder, presents the results of this and outlines the

reasons why this line of research was curtailed. The latter part of the chapter deals with the important work on processing data from lead standards by neural network and gives the quality metrics by which the comparative study is carried out.

The results of the lead standards study are presented in chapter six, providing evidence in both graphical and tabular form. Each method studied is presented in turn. The chapter concludes with a summary of all the results obtained from this part of the work.

Chapter seven contains a discussion of the results obtained and the observations that may be made from them. Arguments regarding error analysis are given.

Finally, chapter eight presents the conclusions reached by the author and links these to the original aims of this work. It is claimed that the application of a neural network to this branch of electroanalytical chemistry would indeed improve the technique and statistical justification for this claim is given. The concluding section of this chapter discusses further research arising out of this study.

The appendices reproduce the author's published papers, contain details of the testing of the electronics and outline other avenues of research that were investigated during the course of this work.

This study cannot claim to make broad inroads into the leading edge of neural network technology research nor has a new electroanalytical technique been developed. Instead this thesis brings together the work of two very different disciplines and presents a novel application of neural networks to an area that has so far received little attention. Furthermore, it shows the application of neural networks to trace metal analysis by differential pulse stripping voltammetry to have immediate advantages and future promise.

## 1.2. The aims of the project

The central hypothesis of this work is that a neural network can be used to advantage to process data obtained from an electrochemical analysis using differential pulse anodic stripping and provide a prediction of the concentration of electroactive material present. Several proposals may be made.

- It is proposed that a neural network could be trained to provide a prediction of toxic metal concentration with an accuracy and precision exceeding that available using traditional calibration curve techniques.

- It is further proposed that by encoding knowledge of the instrumental parameters in force at the time, such predictions of concentration might be made over a greater range than currently possible .

- Finally it is proposed that a neural network may also be used as a controller in order to obtain optimum analytical conditions within a small range of operational parameters.

The overall aim is therefore to develop the above hypothesis by building hardware and software appropriate to the task, by qualitatively investigating the feasibility of post analytical data processing by neural network and by quantitatively comparing traditional and proposed methods.

# 2. Background Studies

## 2.1. Background studies - Voltammetry

### 2.1.1. Overview of the theory

Table 1, Page 24. lists some of the more common methods that have been used at one time or another, together or alone as a means of obtaining analytical information. Voltammetry, the technique used in this work, is the name given to a class of electroanalytical methods whereby a time varying potential is applied between two electrodes placed in a solution and the resulting changing current is measured to provide information about the chemical species present in the solution. Although the technique was originally called Polarography, nowadays the term Polarography is normally reserved for the special case of voltammetry where the electrode in question is formed by a continuously renewable drop of mercury at the end of a fine capillary tube.

The birth of voltammetry (actually using a polarographic method) is usually attributed to Heyrovský[2,3] in 1924, although he had been working on an associated problem for four years before this. In 1924, together with his colleague Shikata, he developed an automatic instrument to change continuously the potential applied to a mercury drop electrode and plot the resulting current - potential curve. They called this instrument a Polarograph[4]. The early historical development of Polarography may be found in several sources[5,6,7,8,9,10] and it is not the intention to elaborate further in any great detail here.

| Response Measured | Electrical Variable Controlled | Name or Description of Method |
|---|---|---|
| E | i (maintained at 0) | Potentiometry |
| E versus volume of reagent | i (maintained at 0) | Potentiometric titrations |
| i vs. E<br>or<br>Δi vs. E | E | Voltammetry / Polarography<br>(Linear scan)<br>(Cyclic voltammetry)<br>(Pulse methods)<br>(Stripping voltammetry) |
| Weight of deposit | i or E | Electrogravimetry |
| I/R (conductance) versus volume of reagent | E | Conductometric titrations |
| i versus volume of reagent | E | Amperometric titrations |
| Coulombs (current x time) | E or i | Coulometry |

Table 1. Summary of electrochemical methods of analysis

The general principles behind the technique are as follows. When a pair of electrodes are immersed in a solution three electrochemical processes may occur, reduction, oxidation or non Faradic. Reduction results in the plating of a metal onto the electrode from solution, Oxidation causes the stripping of metal from the electrode back into solution. A general term for these processes is a redox reaction. They both produce a flow of electrons across the electrode-solution interface, are governed by Faraday's laws and are appropriately known as Faradic processes. The third process, non-Faradic, does not result in a net flow of electrons across the interface and is due to changes in the structure of the electrode-solution interface or the electrical environment of the electrode.

### 2.1.1.1. Linear Scan Voltammetry

If the potential of an electrode is changed, an electroactive species in solution will begin a Faradic plating or stripping process once the applied potential exceeds what is known as the decomposition potential. This is determined by the formal potential of the species in question. If the potential applied to the electrode is changed still further in the same direction once the reaction has begun, the current due to the redox reaction will depend upon this potential as this determines how fast the reaction can proceed. At sufficiently high potentials (up to about 2V) the rate of reaction is so high that all of the electroactive material arriving at the electrode undergoes a redox reaction. The Faradic current therefore reaches a limiting value which depends upon how fast the species can reach the electrode.

If the electrode is rotated then a forced convection is set up with a net flow of solution towards the electrode, thereby increasing the quantity of electroactive material available for reaction. Details of the effects of this rotation are discussed in Fried[11]. When the electrode is stationary however, the rate of flow of electroactive species depends upon its rate of diffusion which itself depends upon the concentration of this species in the solution. This means that if a linear potential sweep is applied to an electrode, a diffusion controlled reaction will cause a maximum, limiting current that depends upon the species concentration. Hence this limiting current is a candidate for an analytical method. Further descriptions of special forms of voltammetry can be found in Anderson[12] and Wightman[13]

### 2.1.1.2. Pulse Voltammetry

The electrode-solution interface contains a small region of solution close to the electrode, called the double layer, with a reduced number of free ions due to adsorption in the electrode. Therefore there exists, at this interface, a capacitor with

the electrode and the solution a short distance away acting as the plates and the double layer, behaving as the dielectric. An important non Faradic process is the charging current that flows into the electrode due to the charging of this 'capacitor' when the potential of the electrode is changed. With typical solid electrode the capacitance is very small and the time for which the exponentially decreasing charging current is appreciable is very short, of the order of micro seconds.

In pulse voltammetry, described in many papers by Osteryoung[14,15] and also by Murphy[16], use is also made of the fact that when a pulse is applied to the electrode, so charging the double layer, the capacitive current decays exponentially with time according to

$$i_c = \frac{E}{R} e^{-t/RC} \qquad \{1\}$$

Where E is the potential applied between the working electrode and the reference electrode, R is the resistance of the solution in the double layer and C is the capacitance of the double layer. Immediately after the pulse however, the Faradic current decays as the electroactive species immediately in the vicinity of the electrode becomes exhausted. This Faradic current, however, decays more slowly than the non Faradic, capacitive current.

Therefore if a current measurement is made at a short time after the application of the pulse the majority of the current measured will be due to the Faradic process and the detection limit, relative to a linear scan method, will be lowered.

The waveform for normal pulse voltammetry where the pulses increase in amplitude, is shown in fig 2.

This waveform is little used these days as the current-potential response obtained has a sigmoidal shape rather than the more easily processed peak shape obtained with differential pulse techniques. Preference is therefore given to differential pulse wave forms.

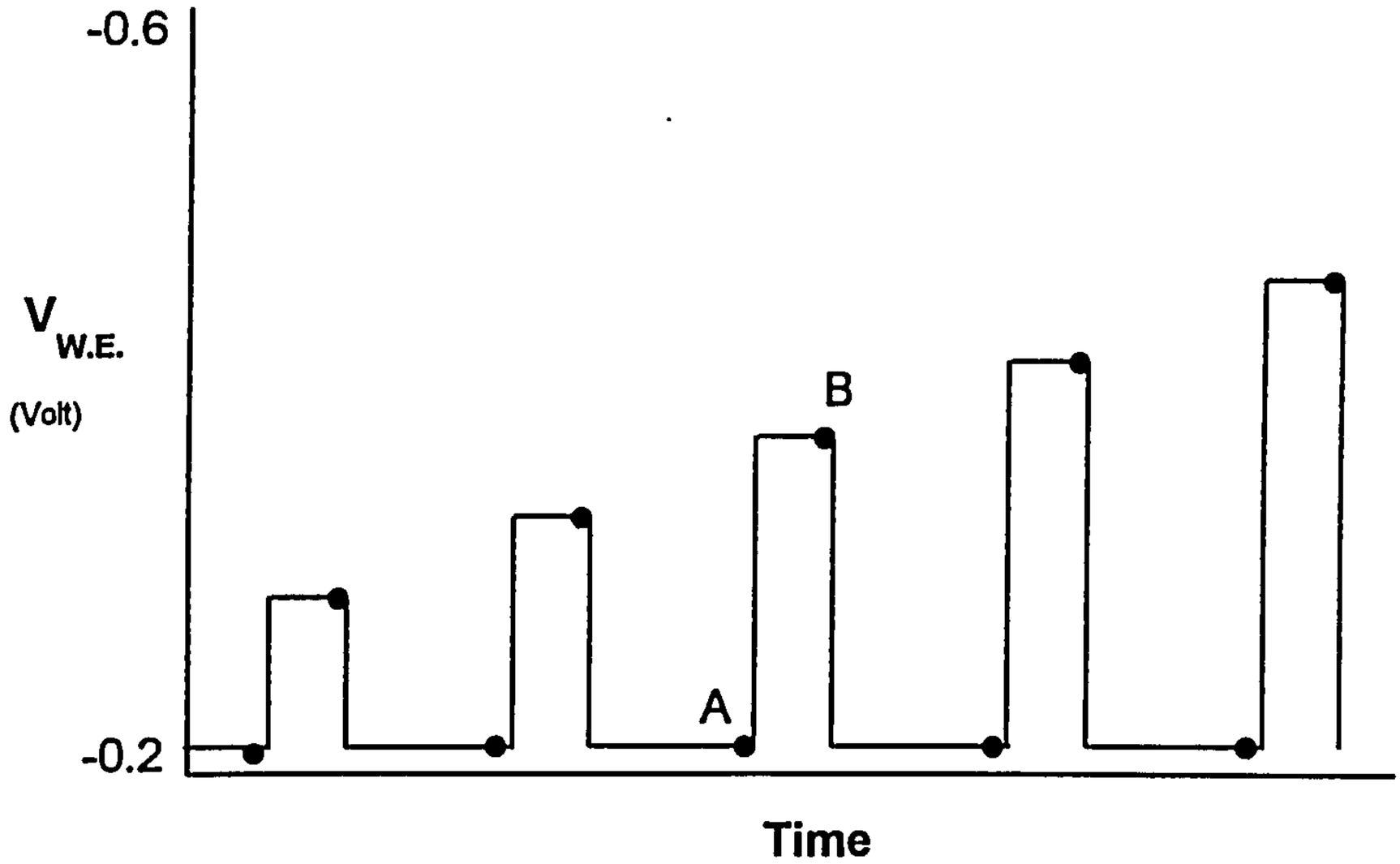



Figure 2. Typical input waveform for normal pulse voltammetry

## 2.1.1.3. Differential Pulse Voltammetry

In differential pulse voltammetry the pulses are superimposed either on a rising potential ramp or on a stepped ramp, the latter being more suited to digital computer control. Figure 3 shows both of these waveforms on the same potential axis with the stepped ramp, as used for this work, shown in the upper of the two waveforms. The current is measured before and after the pulse and the two values are subtracted. This results in a peak shaped response of $\Delta i$ against applied potential (figure 4) as the electrode potential sweeps though the region of the Faradic process for the species in question. In both fig 2 and fig 3, A and B are the two points

where the current is measured and subtracted for each pulse. The use of a differential pulse waveform with computerised instrumentation is described in Bond[17] using a polarographic technique. Bond thought his work unusual since he used pulses with a width similar to the time interval between the pulses , unlike previous workers who used pulse widths ten or more times smaller that the pulse interval. In this work the author has also found that using a simple pulse with a mark space ratio of 1:1 provides perfectly good results.

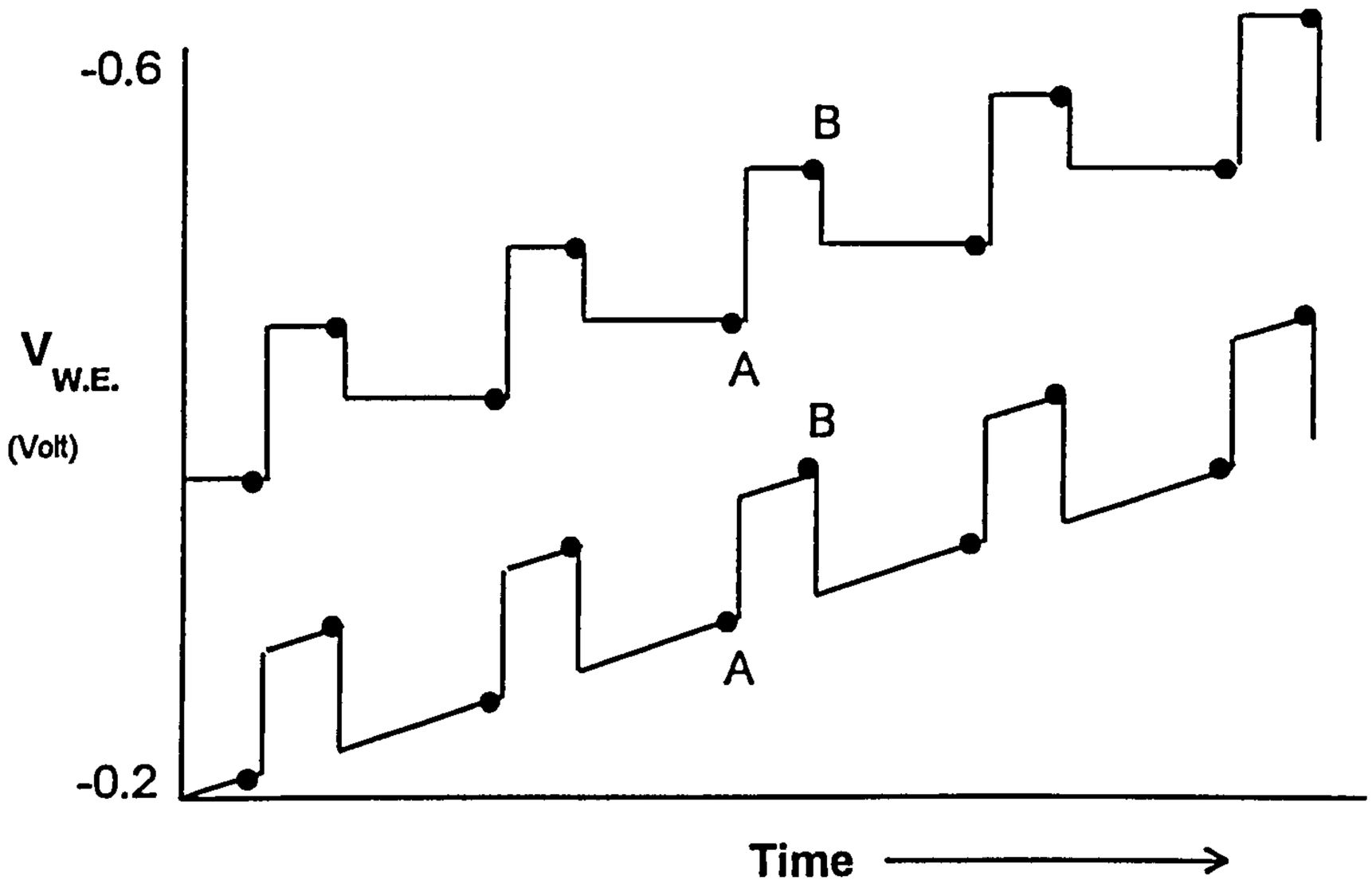



Figure 3. Typical input waveforms for differential pulse voltammetry

## 2.1.1.4. Differential Pulse Stripping Voltammetry

Differential pulse stripping voltammetry is described in Wang,[18] Kopanica[19] and Brainina[20]. The electrode used is mercury, either in the form of a drop or as a thin film that is plated on to a solid substrate. In common with the work described in this thesis, Graabæk[21] used a thin film of mercury on a glassy carbon substrate

for his electrode. When using this type of electrode the substrate is first electroplated with mercury, usually whilst rotating. Then it is placed in the solution under test and the electroactive species is allowed to accumulate into the mercury by the application of a constant potential that causes the species to plate the mercury and so form an amalgam with it. During the final, analysis, phase the differential pulse waveform is applied in such a direction that the species is stripped out of the mercury, producing the Faradic reaction.

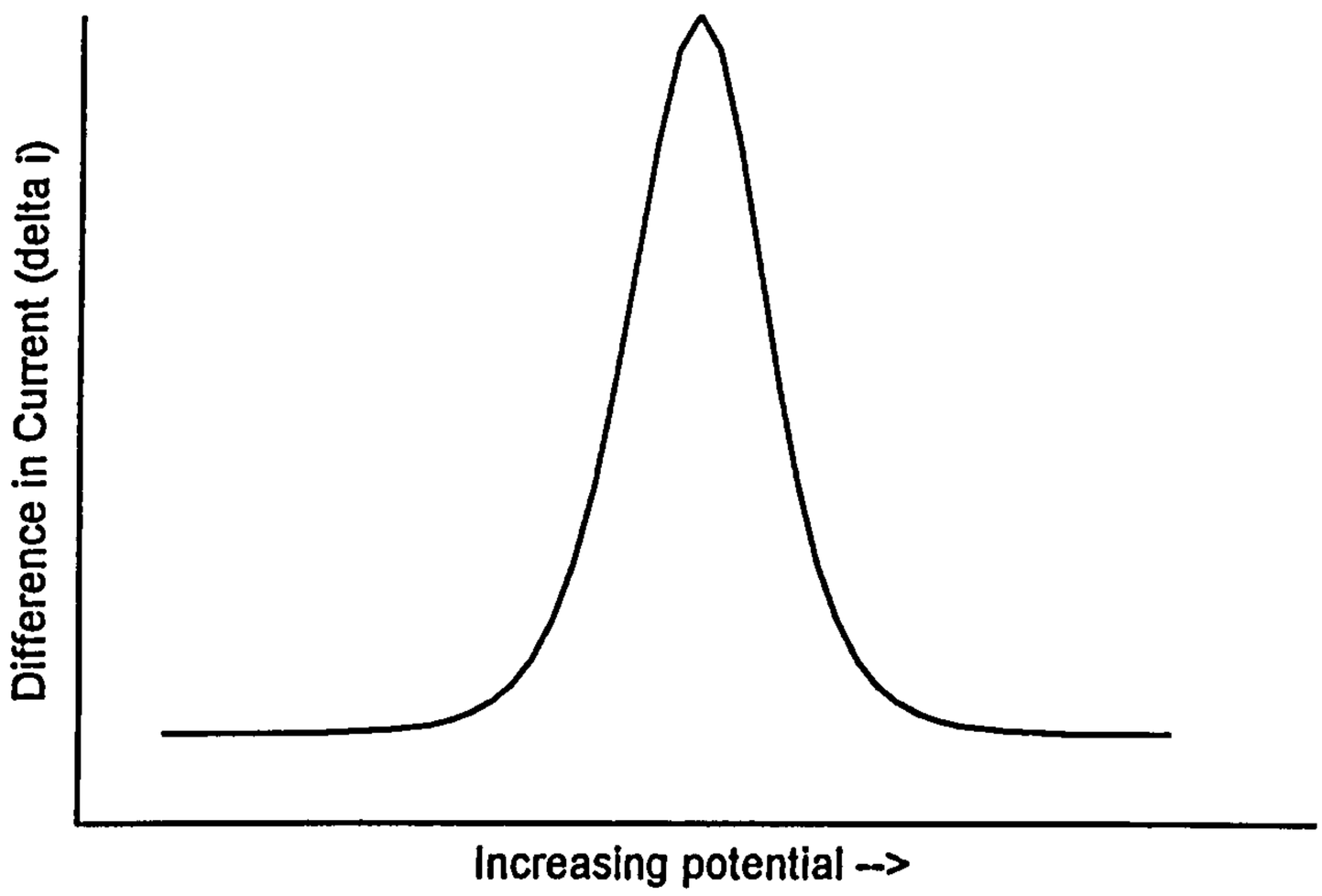



Figure 4. Ideal response from differential pulse voltammetry

The benefit of the stripping method is that since, due to the previous accumulation, the concentration of the species in the mercury is very much higher than in the analyte, this method provides considerably higher sensitivity than non stripping methods.

It should be noted that as an alternative to pre plating the electrode with mercury, it may be simultaneously plated with mercury and the species in question.

This technique, used for the bulk of this current work, is more fully described on page 146.

In addition Pomeroy[22] contains useful information on the technique of using thin film mercury electrodes. He includes a schematic diagram of a three electrode potentiostat circuit and gives full details of the electrode preparation and analytical procedure. The use of this electrode with differential pulse voltammetry is also mentioned in DeAngelis[23], who achieved a detection limit of 10 ng $ml^{-1}$ for lead in organic materials, although his electrode plating times were of the order of several minutes.

Additional material on voltammetry may be found in Copeland[24,25] and Edwards[26], both of whom give an overview but with Copeland offering a more detailed analysis. The related technique of staircase voltammetry is discussed by Svensmark[27] and Eisner[28], both of whom continued to rotate the electrode during the analytical scan. Eisner confirms the theoretical relationships, published several times elsewhere by his co-author Osteryoung, between the response and the analytical parameters. He also reports that the sensitivity of staircase voltammetry was about the same as the differential pulse method.

Square wave anodic stripping, which uses a slightly different waveform to that used in differential pulse stripping has been reported by Anderson[29], Mlakar[30] and Feldman[31] with the latter employing a continuously rotating electrode, similar to Eisner.

A more general discursive treatment of the instrumentation involved in the above techniques may be found in Christie[32], who investigated instrumental artifacts, Beebe[33], who looks at process analytical chemistry and Bersier[34,35] who compares several alternative methods. More detailed comparisons are found in

Bond,[36] Anderson[37] and later Willard[38], all of whom used polarography with a basic form of computerised instrumentation. Very readable accounts of the effects of the pulse shape on response may be found in Parry[39] and Rifkin[40] , the former dealing with normal pulse at a mercury drop electrode and the latter with differential pulse at a thin film electrode.

### 2.1.2. Computer control of equipment

A simple overview of the techniques involved in interfacing is given in Gates[41] and Okamura[42]. Gates in particular gives 8086 assembly code for interfacing and data collection and shows, in rather primitive BASIC, how Savitzky-Golay type smoothing, discussed further on page 33, might be applied. Interestingly, given the source of the article, an example is given of data collection from a polarographic analysis of lead in gasoline.

More thorough background material may be found in He[43], Betteridge[44] and Tacussel[45], all of whom present a discussion of recent progress in microprocessor instrumentation. Betteridge offers a very thorough overview of microcomputer applications in chemistry, whilst Tacussel concentrates upon voltammetry and polarography. Gunasingham[46] and Osteryoung[47] give a good introduction to the instruments used, including computers.

Keller and Osteryoung[48] produced an early computerised instrument in 1971 to perform pulse polarography and smooth the data by ensemble averaging and a year later Landowne[49] worked on a similar system for gas chromatograms. By 1980 Granéli[50] was using a microcomputer system with a fixed waveform for stripping analysis. Brown[51] and Gustine[52] investigated the possibility of a more versatile instrument whilst Paul[53] looked at the use of a 6502 processor controlling twin

working electrodes in order to reduce inter metallic compound formation by plating the interfering metal on one electrode and the metal under test on the other.

The data processing techniques available were enhanced by Kalcher[54] in 1986 who investigated synthetic baselines and by both Imaino[55] and Lysaght[56] in 1991who looked at the constraints imposed by laptop computers.

An interesting parallel development appeared in 1984 when Ziegler[57] studied the possibility of producing a command language for data processing. This was followed up in 1989 by Miller[58] and Thomas[59] who produced a language for performing voltammetric analysis and by Gerth[60] who in 1992 extended the concept into an object oriented system.

### 2.1.3. Smoothing and noise filtering

A general procedure for noise filtering is given in Carr[61] who describes useful electronic methods of noise filtering. The effect of smoothing on such parameters as peak shape and peak height were investigated by Bromba[62] in 1983 but not until some years later were proposals put forward on optimum sampling techniques by Deutscher[63]. DeNoyer[64] in 1990 used a maximum likelihood method of smoothing data which appears to be very successful and a year later Lee[65] used principal components analysis to do the same thing.

Fourier transforms, an overview of which is given in Aubanel[66] and which were used in this work as a means of reducing the dimension of the data vectors, were used by Horlick[67] as a means of smoothing spectra. The same technique was used by Hayes[68] for polarographic data and the advantages and disadvantages of this method over floating least squares discussed. Felinger[69] describes an advanced windowing method with the fourier transform and claims a significant improvement of the signal to noise ratio.

The Savitzky-Golay[70] filter is the method used in this work to smooth and differentiate the data displayed by the graphics module of the analytical software. The work of these authors elegantly calculates and tabulates the coefficients of the polynomials that should be convoluted as a sliding window with serial data in order to produce the desired amount of smoothing or differentiation. The smoothing can then be done with the minimum of computational expense.

It must be noted however that Savitzky's original work, published in 1964 contains errors in some of the tabulated results and that one of the equations given is incorrect. These errors were not corrected until eight years later when in 1972 Steiner[71] recalculated the values using the correct equation. It is the values tabulated by Steiner that are used in this work. Interestingly, a recent work by Bromba[72], who gives hints on the application of Savitzky-Golay filters, makes no mention of the original errors or the subsequent corrections.

### 2.1.4. Curve fitting and digital simulation of voltammetric curves

There have been many papers over the last thirty years which have addressed the equations describing voltammetric responses to various analytical procedures. The equations described in some of these have been used in this work to simulate responses in order to create a database of model responses for neural network training. This part of the research is described more fully in appendix two. A general background of curve fitting in this area of chemistry may be found in several references[73,74,75,76].

Two papers by De Vries[77,78] should, however, be seen as the starting point for any further investigation. Following a paper in 1964 presenting an approximate theory of anodic stripping voltammetry suitable for high rates of potential change and thick mercury films (e.g. 2 $Vmin^{-1}$ and 100μm), De Vries presented an exact

treatment in 1965. These papers are very theoretical however and are not suitable for direct translation into code for modern computers. This work was followed up in 1967 by a third paper[79] which concentrated more on linear sweep voltammetry and a more refined approximate theory was developed. Osteryoung[80] investigated modelling linear scan voltammetry by a series of single potential steps. He found good agreement with the theory of De Vries. This was taken up by Rifkin[81] in a very comprehensive and significant paper in which he investigated the effects of pulse duration and inter pulse time and by Dillard[82,83] who looked at the effect of modulation amplitude and who, three years later looked at the same problem from a finite differences perspective[84]. Ruzic[85] presented an overview of the emerging theory in 1978 and clearly stated the confusing nomenclature in use regarding the term differential pulse (fig 3), used to describe the type of waveform used in this work and derivative pulse (fig 2), used when the same current sampling technique is used on normal pulse wave forms. Heijne[86] also derived the equation by a theoretical method, was criticised for a discontinuity at zero pulse height by Kies[87] and developed a new model in a paper published two years later[88].

Since that time a number of publications have looked at curve fitting applied to this area of analytical chemistry many of which may hold potential for future workers in neural network applications associated with this branch of chemistry. Bond[89] used a quadratic to fit the baseline and remove the unwanted background current. This same technique was also used throughout the simulations performed in this work, further discussed in appendix two. Mellado[90] developed a rather primitive set of programs written in BASIC to approximate differential pulse polarograms.

Several researches have studied applications of the Kalman filter[91] to analytical chemistry. The principles of this algorithm are quite well covered in Rutan[92] and also in Yongnian[93], the latter using the filter to improve curve resolution in

differential pulse polarography. Seelig[94,95] used the algorithm to optimise the signal in noise corrupted returns from anodic stripping voltammetry and also made comparisons between this and non recursive estimation methods. Poulisse[96] looked at multicomponent analysis using spectrometric data whilst Wentzell[97] compared the use of the FFT in three different non electrochemical analytical fields. However, little further work appears to have been done recently in this area.

### 2.1.5. Resolution enhancement and mathematical post processing

This has been the subject of many papers since Perone[98] described using second order derivative techniques as a means of enhancing the sensitivity of peak position measurements when analysing polarograms. Later Grushka[99] took up this topic and looked at the limits of the technique when applied to overlapping peaks and presented useful results which showed the range of the method as a function of signal to noise ratio, separation and peak height. A similar topic was investigated by Oppenheimer[100] who looked at ways of determining the lowest limit for reliable measurement. Mitchell[101] published an interesting technique which involved fitting a series of least squares regression equations and then using the equation displaying the narrowest confidence limits as a means of measuring an unknown sample.

Various mathematical techniques have been investigated for the resolution of overlapping curves. Gutknecht[102] fitted computer generated 'standard' curves to the polarogram, Bond[103] also used standard curves which were successively subtracted from the polarogram until only a single peak remained. Grotch[104,105] used a form of table lookup against a library of curves and Caruana[106] resolved the spectra into a series of gaussians using a least squares fit. Several authors[107,108,109,110] have used principal components analysis to resolve separate influences on a complex instrument response resulting in spectral information and novel modifications to the algorithms have also recently been proposed[111,112,113].

Both Caruana and Glajch[114] presented criteria for goodness of fit of a given estimated mixture, the latter also offering a simple method to quantify peak separation by measuring the valley depth from a line fitted at a tangent to two adjacent peaks. Recently, Toft[115] has also published a novel resolution parameter but one which may be calculated prior to curve resolution.

Not surprisingly, many workers have used the Fast Fourier Transform (FFT) as an aid to resolution enhancement. A mathematical overview of the FFT is given in Elliot[116]. Grabaric[117] used it to sharpen polarographic peaks mathematically and Skarjune[118] developed an enhanced version of the algorithm using less multiplications. More recently, Jackson[119] compared the application of the fourier transform and maximum likelihood theories to curve fitting with particular emphasis to overlapped bands in spectroscopy.

Engblom[120,121] used the FFT to find peak parameters such as position and width as well as to enhance the resolution. This was done using the assumption that voltammetric peaks may be approximated by the function

$$p(x)=\frac{a}{\cosh^2[b(x-c)]} \qquad \{2\}$$

where a = height, b = a width parameter and c = the peak position and x represents the potential. This equation was also used in this current work to generate synthetic curves for training neural networks and also for the calculation of an objective function when investigating resolution enhancement using a genetic algorithm ( see chapter 9). Genetic algorithms are very well described in Davis[122] and Goldberg[123] .

# 2.2. Background studies - Artificial neural networks

## 2.2.1. Introduction

Neural networks are computational models which model the relationship between a set of inputs and a set of outputs. In this work the inputs are values of current taken from the voltammetric peak and the output is the concentration of a chemical species. Once the internal model of the data has been formed novel inputs can generate outputs, sometimes themselves novel, which fit the internal model. Unlike algorithmic prediction processes however, the model used by the network does not have to be determined a priori but can be derived, or learnt, by the neural network through the computational process known as training. The model thus created may be very complex and contain relationships that were hitherto unknown in the input - output data.

## 2.2.2. General Neural Network Structure

A neural network could be considered, in some elementary sense, as computationally equivalent to a neuron in the brain. Like the brain a neural network consists of many simple units connected together by a large number of weighted links, with each unit receiving input from many sources. Indeed the weights are often called synaptic weights, to increase further the correspondence between the real and artificial neuron. The input units obtain their input from external sources whilst other units obtain their input from the single output that each unit produces. The path of this output is split and terminates, via the weighted link, at the input of the receiving unit. Note that although the output *path* is split, the output *value* is transmitted unchanged to all receiving units. Some units, called the output units,

send their output back to the real world to be converted into appropriate responses. Figure 5 shows a typical connection schematic for the units in a multi layer neural network having three inputs and two outputs organised into three layers of units. Note that in this work, when the number of network layers is mentioned, this should be taken to include the untrained input layer. Hence the three layer network shown below has two trainable synaptic weight matrices.

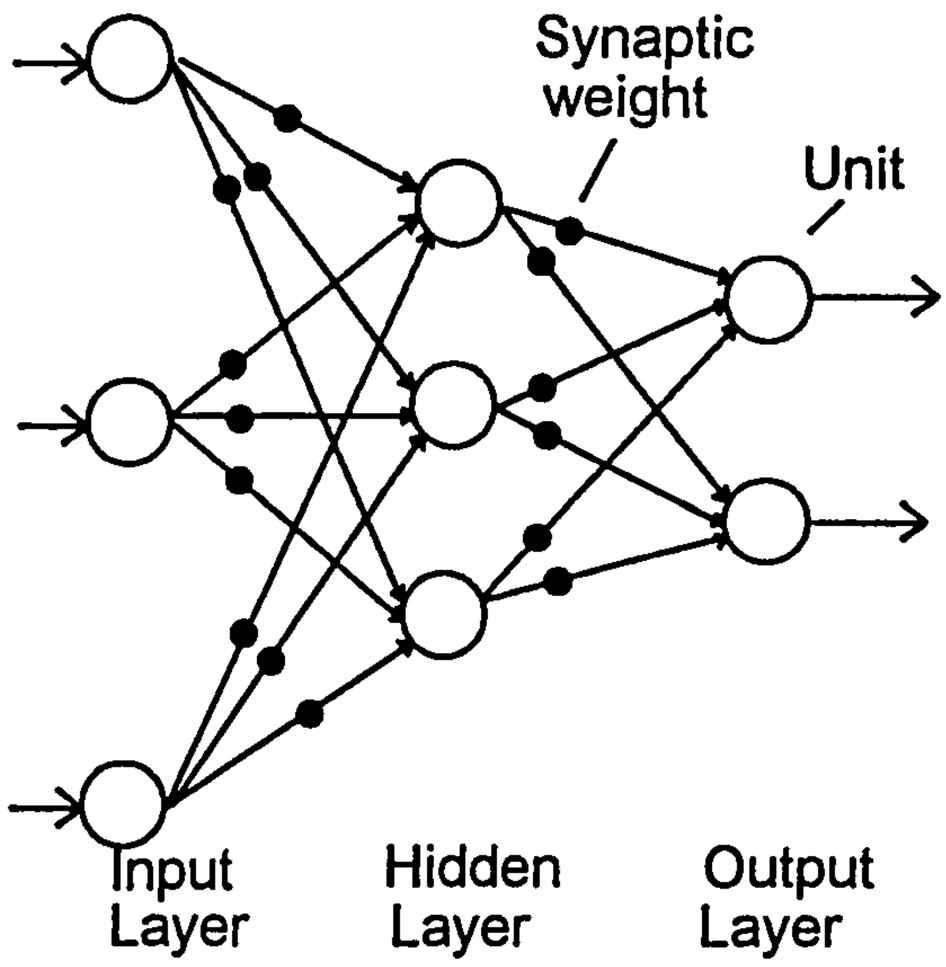



Figure 5. Schematic of a 3 layer neural network

Since each unit receives many inputs, and the entire network may contain many output units, the input and output data are usually considered to be vectors of real numbers which are mapped to the real world parameters corresponding to each element of the vector. The processing carried out by each unit to convert the vector of input values into an output, i.e. the transfer function is quite simple and is described below.

First the weighted inputs are combined to form a single value, often by simply forming a weighted sum of the inputs using

$$I_i = \sum_{j=1}^{n} w_{ij} x_j \qquad \{3\}$$

Here $I_i$ is the combined weighted input to unit *i*, $w_{ij}$ is the weight on the link from unit *j* to unit *i* and $x_j$ is the unweighted signal coming from unit *j* (this assumes that the output of unit *j* is sent to the input of unit *i*). The weights may be positive or negative.

Once the combined weighted input has been calculated the unit calculates the value of its output signal. In early work, such as Rosenblatt's Perceptron[124], described in more detail later, this was simply set to +1 or -1 depending upon the value of $I_i$ relative to some threshold value. It is more common nowadays to use a non linear activation function such as the logistic sigmoid function,

$$f(I) = \frac{1}{1+e^{-I}} \qquad \{4\}$$

whose shape is shown in figure 6.

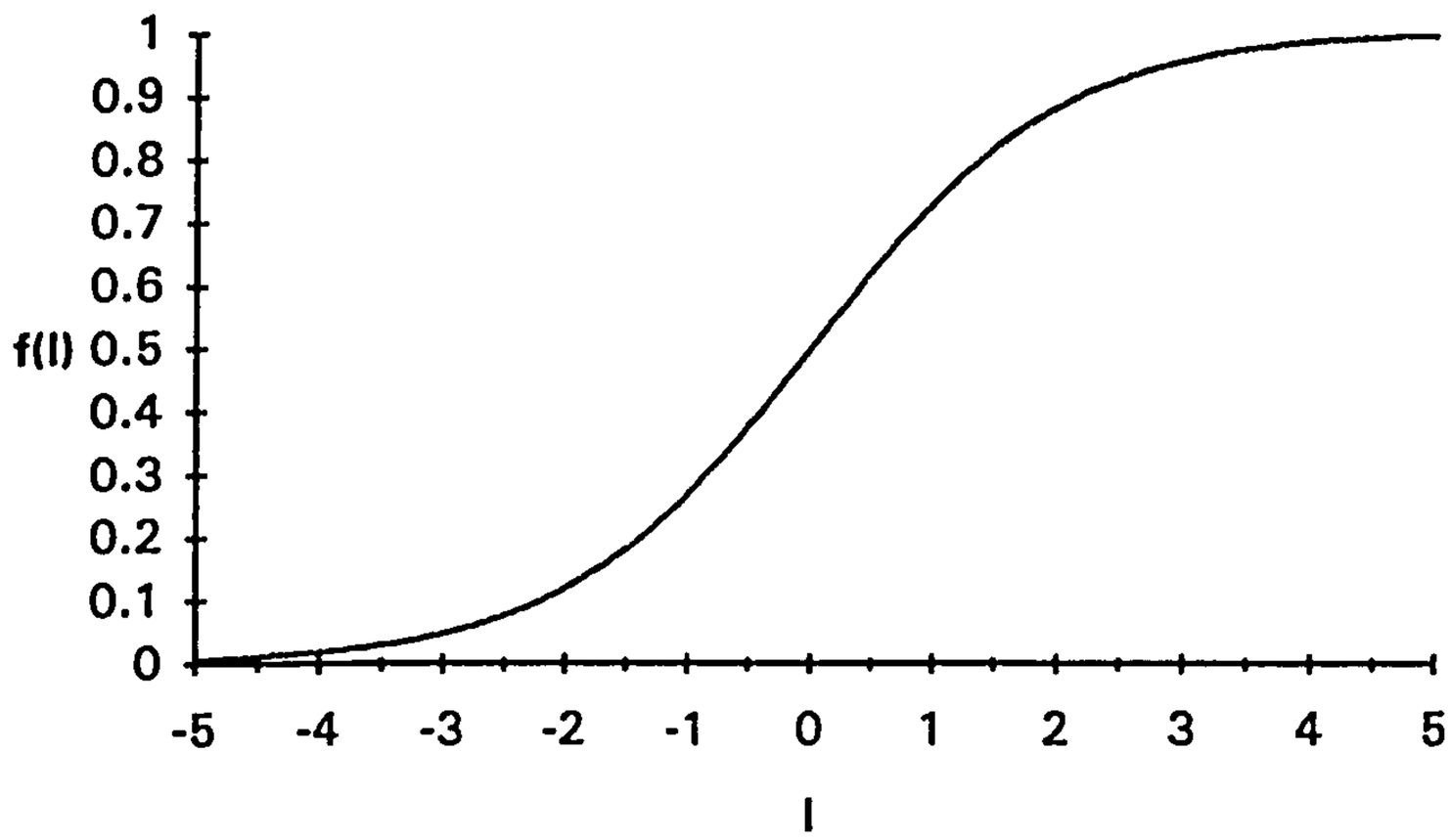



Figure 6. Shape of logistic function

It has been shown[125] that using this function, which varies between 0 and 1 as $I$ varies between $\pm \propto$ a network can form much richer mappings between input and output. It also has the advantage that its derivative is easy to compute, being given by.

$$\frac{df(I)}{dI} = f(I) \times \left(1 - f(I)\right) \qquad \{5\}$$

This gives a speed advantage during training when the derivative must be calculated many times. The threshold concept is still retained in some networks by having a simulated additional input to each unit with an activity fixed at at +1 but, like other inputs, having a weight that can modify this activity to increase or decrease the overall input to the unit.

The processing in the unit can therefore be described graphically by figure 7.

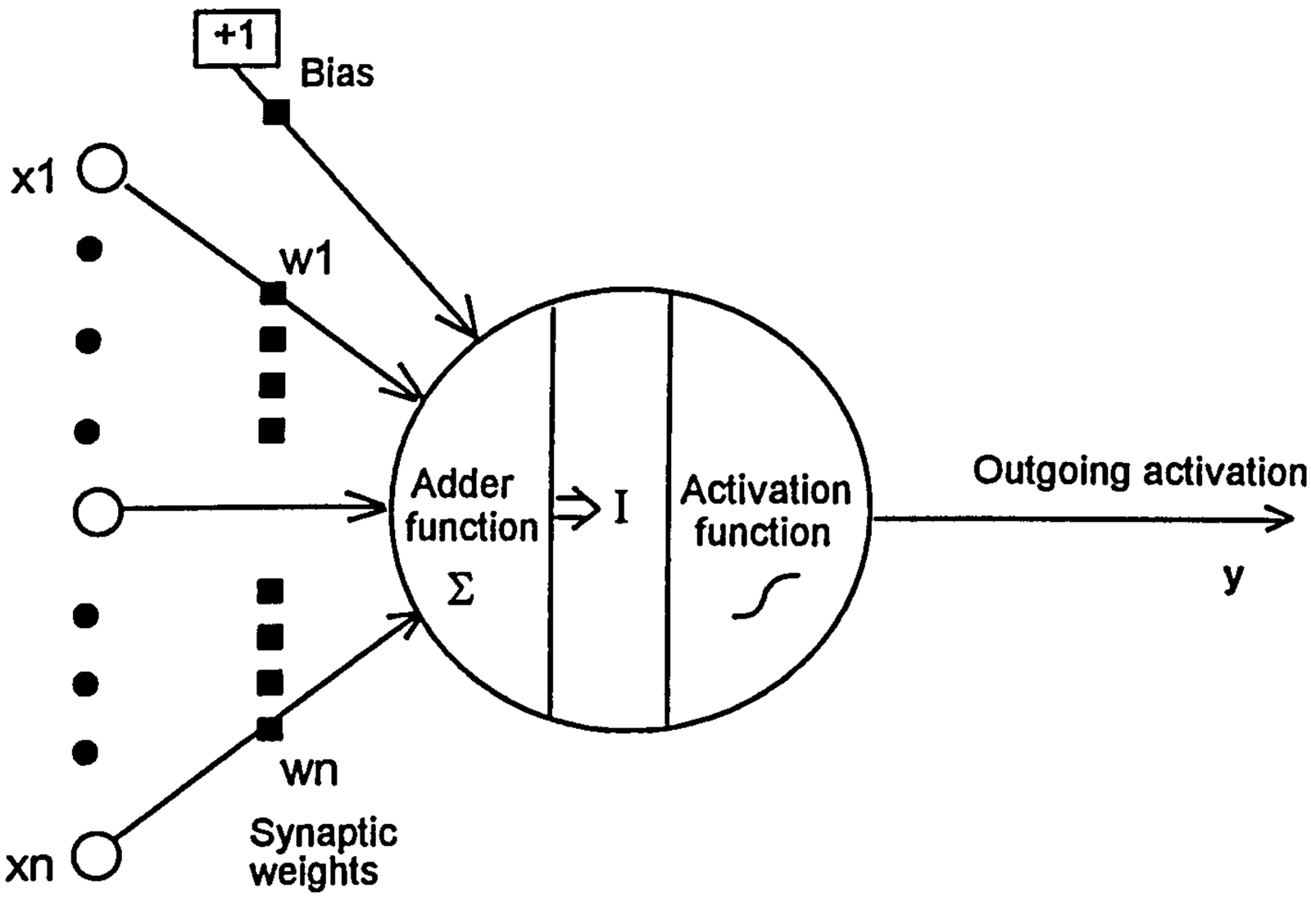



Figure 7. Simple processing unit

Starting from an initial set of random weights, the algorithm used to train the units changes the synaptic weights during the learning phase in order to produce the desired output for each input pattern. Supervised networks are those which learn by being repeatedly presented with a training set of data containing many examples of input vectors and the corresponding output vector. From this the actual error in the network's output can be determined and the weights adjusted accordingly. Unsupervised networks are those which are not given any target output for the training set and instead change the weights so as to classify the input vectors into clusters, based upon features discovered in the training set, which in some way preserve the natural topology present in the original data set.

### 2.2.3. Development

The first theorists to conceive of the fundamentals of neural computing as opposed to analytical neural modelling or neural-physiological research were McCulloch and Pitts[126]. In 1943 they devised a digital neuron with pre-set weights, which was unable to learn, but which could classify its inputs into one of two classes, so effectively performing a Boolean operation.

#### 2.2.3.1. The perceptron

In 1949 Hebb[127] proposed that learning in the brain was achieved by changes in the connection strength between neurons, i.e. the weights, that were proportional to the activation of the neuron undergoing a learning process. Rosenblatt[124] employed this idea in the first trainable, two layer network which he called a Perceptron. This was trained by presenting a set of input data to the network and calculating the weighted sum as previously described. The output $y$ was then calculated from this weighted sum using the simple activation function

$$y = \begin{pmatrix} +1 \text{ if } I > \theta \\ 0 \text{ if } -\theta \le I \le \theta \\ -1 \text{ if } I < -\theta \end{pmatrix} \qquad \{6\}$$

$\theta$ = some threshold value

If the Perceptron produced an incorrect answer, the vector of weights on the input were altered according to the following.

$$\mathbf{w}_{new} = \mathbf{w}_{old} + t\mathbf{x} \qquad \{7\}$$

where

$t$ = desired output, +1 or -1
$\mathbf{w}_{new}$ = new vector of weights
$\mathbf{w}_{old}$ = original vector of weights
$\mathbf{x}$ = vector of input values

Rosenblatt also proposed a three layer network with internal, hidden units but was unable to devise a method of training it as he could not see a way of obtaining the error in these hidden units.

Widrow and Hoff[128] developed an improvement to the Perceptron training algorithm in 1960 which is referred to as "minimum error", "least mean squares" or "delta rule" learning. Using this rule the output of the unit is found as before, with the threshold set at zero. The error $E$ in the output is then found by simply subtracting the actual output from the desired output for that input pattern. Finally the weights are adjusted according to;

$$\mathbf{w}_{new} = \mathbf{w}_{old} + \beta E\mathbf{x} \qquad \{8\}$$

where $\beta$ is a variable learning rate between zero and one. This process is repeatedly applied until the unit no longer makes an error. This resulted in a unit that was able

to generalise, that is produce a sensible output for an input that it had not previously seen during training.

In 1969 however, Minsky and Papert[129] published their famous book 'Perceptrons...' which showed the limitation of the Perceptron to be that it could only make classifications for problems that were linearly separable, i.e. one where, with n inputs, the decision boundary is a hyperplane in n dimensions. This meant that it could not even learn the two input XOR function where, as shown in figure 8, the decision boundary separating the two classes of output in two dimensions is not a line but a closed figure.

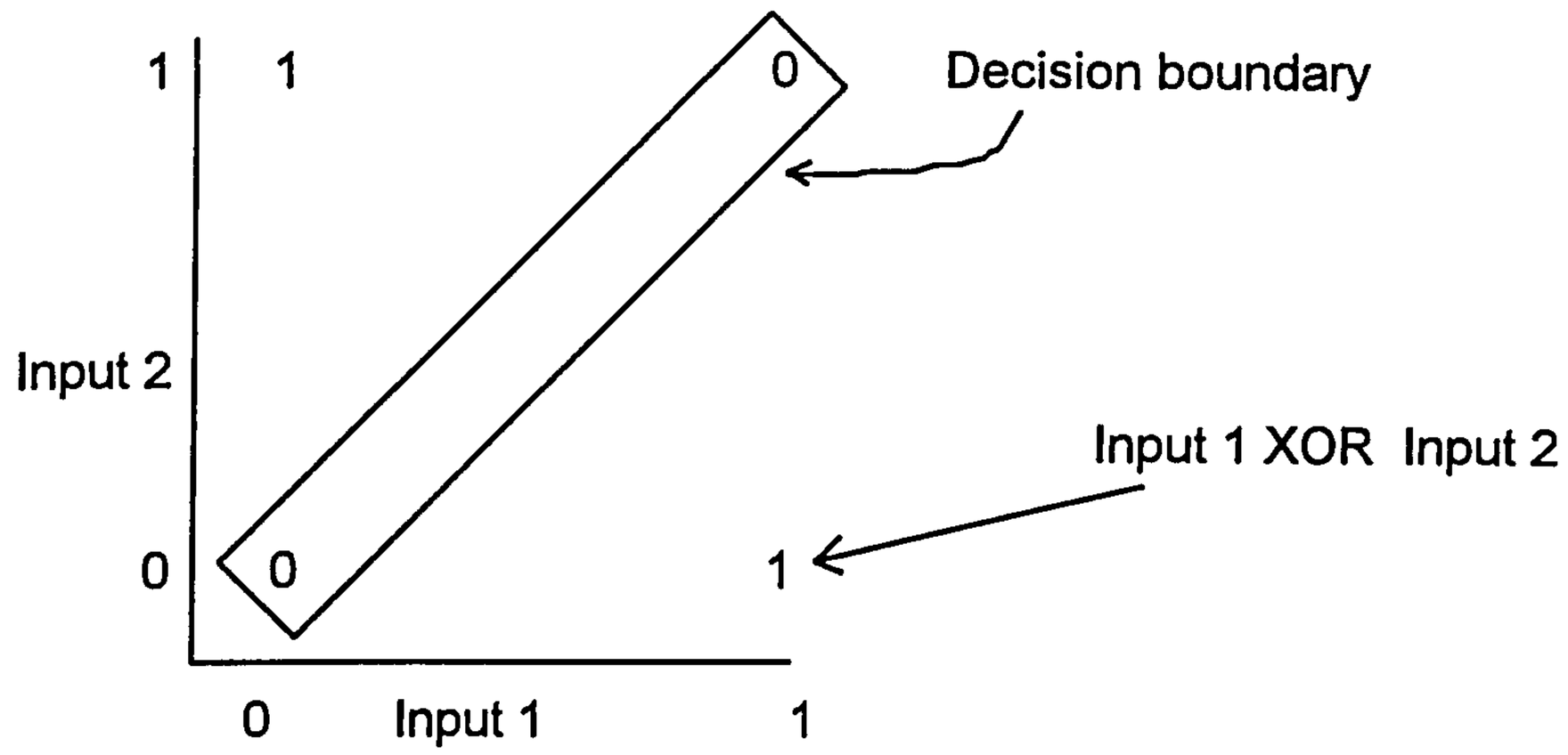



Figure 8. Decision boundary for the two input XOR function

This proof and the lack of a suitable algorithm for updating weights in a multi layer network so damped research enthusiasm that activity was very low for the next five years.

### 2.2.3.2. Higher Order Networks

Higher order networks appear to overcome the linear separability problem however. This term implies a threshold logic unit similar to a Perceptron that uses an input vector with an increased dimension, formed by including the raw data elements and also higher order terms obtained by combinations of the raw data elements or functions of them. This type of network was also used in this current work where the input dimensionality was increased by taking not only the real inputs, $x_1 \ldots x_n$, representing value of voltammetric current, but also the product of each pair of input values. i.e.

$$(x_1),(x_1 \times x_2),(x_1 \times x_3),\ldots(x_1 \times x_n),(x_2),(x_2 \times x_3),(x_2 \times x_4),\ldots(x_2 \times x_n),\ldots(x_n). \qquad \{9\}$$

A study of high order networks was made by in 1987 by Giles[130] who found them superior to simple two layer networks for character recognition.

### 2.2.3.3. Multi Layer Networks and Back Propagation of Error

During the years following Minsky and Papert's work, other network models were developed by workers such as Kohonen[131], who built an 'associative memory', Anderson[132] with his 'brain state in a box', Grossberg[133], (who has 146 publications listed in Klimasauskas[134]) and Carpenter[135], who developed adaptive resonance theory. However, it was not until the independent formulation of the backpropagation algorithm, first by Werbos[136] in 1974 without much publicity and subsequently by Parker[137] in 1982 that work on the multi layer Perceptron (MLP) developed further. The algorithm they produced was refined and widely published in 1986 by McClelland and Rumelhart[138] in a book that has since become a standard reference text.

The backpropagation algorithm is a variation of the minimum error algorithm discussed above. For each pattern in the training set the input activity is processed

as previously described, using a sigmoidal activation function, through to the output units. The error in each output unit $j$ is then found by subtraction.

$$E_j^{\text{output}} = y_j^{\text{desired}} - y_j^{\text{actual}} \qquad \{10\}$$

This error is then propagated backwards through the middle, or hidden, layers of units to the input layer being weighted at each unit by the same weights used during the forward activation pass. That is, the error in a hidden unit $i$ is the weighted sum of the errors in all the units $j$ to which it sends its output. The significance of the error is related to the current sensitivity of the unit, i.e. how much a small change in $I_i$ influences the output $y$. In order to take this into account the weighted sum of the errors is finally multiplied by the derivative of the activation function giving

$$E_i^{\text{hidden}} = \frac{\mathrm{d}f\left(I_i^{\text{hidden}}\right)}{\mathrm{d}I} \sum_{j=1}^{n} w_{ij} E_j^{\text{output}} \qquad \{11\}$$

This is repeated for previous layers with the meaning of the suffix 'output' being the layer immediately following the hidden layer.

The training process, which may take some time and is not guaranteed to succeed, consists of repeatedly presenting each input - output vector pair in the training data set to the network until the error in the network output falls to an acceptable level.

A multi layer Perceptron (MLP) trained using the back propagation algorithm is capable of classifications in non linear problems and can, for example, learn to model the two input XOR function.

Good introductions to neural computing may be found in Lippmann[139] and Kohonen[140] with elementary practical advice on building neural networks in

Kosko[141] and Nelson[142]. Haykin[143] provides a comprehensive coverage with numerous references and Blum,[144] Freeman[145] and Pao[146] offer in depth practical advice on building neural networks, including computer code. Müller[147] and Fausett[148] offer a mathematical description whilst McCord[149] describes a method of building a neural network in Lotus 123© . Zupan[150] gives a sound overview of the fundamental techniques concentrating upon the chemical problems to which neural networks are currently applied.

A study of learning speeds in back propagation was carried out by Fahlman[151] in 1987 and a faster version of the back propagation algorithm called quickprop described. Other workers have investigated various aspects of the back propagation MLP. Hirose[152] investigated varying the number of hidden units, but was only able to provide rough guidelines on the optimum number to use for a given problem. (Consequently, in the current work the author used an empirical method to find the optimum number, see page 178).

Weir[153] looked at automatically changing the learning rate during learning and Rigler[154] scaled the errors that are back propagated. Gaussian activation functions in the hidden layer were investigated by Weymaere[155] who also suggested training the network first on a subset of the training data and then retraining using the complete training set but starting with weights obtained on the subset.

### 2.2.3.4. The General Regression Neural Network

The general regression neural network, fully described in Specht[156], also belongs to the class of supervised networks but does not train using back propagation. It has the advantage of being able to train quickly on sparse data sets. There are as many neurons in the hidden layer as there are training patterns. It works by measuring how far in N dimensional space a new input pattern is from the

patterns in the training set, where N is the number of inputs in the pattern. When a new pattern is presented to the network the output that is predicted by the network is a proportional amount of all of the outputs in the training set. This proportion is based upon how far the new pattern is from the known patterns in the training set.

In essence, the network's estimate for an output associated with a particular test set input can be thought of as a weighted average of all observed training set output values, where each observed value is weighted according to the distance between its input vector and the unknown input vector whose output it is to determine. Two ways of measuring the distance between two vectors can be used, the Euclidean distance, a straight line in N dimensional space and City Block, the distance measured along the N axes.

A smoothing parameter $\sigma$ is used which affects the network's ability to generalise. A very large value tends to cause the estimated output to become the sample mean of the training outputs. As $\sigma$ approaches zero the output for an unknown vector assumes the output value of the training vector closest to the unknown vector. At intermediate values all values of training output are taken into account but those for training vectors closer to the test vector are given more weight. The search for the optimum value of $\sigma$ is described later.

### 2.2.3.5. Temporal Processing

Temporal processing was investigated by Lang[157] and Kämmerer[158], both of whom constructed training sets from time delayed windows of speech in the frequency domain. Networks incorporating temporal processing internally however were first formulated by Jordan[159] by feeding the output of the network for one input pattern presentation back to the input on the next pattern presentation, i.e. a recurrent network. This work was enhanced by Elman[160,161,162] who also fed back

outputs from the hidden layer to the input layer. Torreele[163] used a genetic algorithm to build a recurrent network for the difficult problem of a temporal multiplexor. This system supplied bit strings one bit at a time and classified the strings by using the first n bits as an unsigned address pointing to one of the remaining, temporarily later, bits which was taken to be the target category. Torreele claimed the genetic algorithm approach produced a smaller network than would have previously been thought necessary.

Some interesting ways of describing and analysing a temporal waveform using a string representation of the waveform are to be found in Fu[164] and Pavilidis[165], both of whom describe a syntactic method of analysing such patterns and give algorithms for peak detection and pattern matching. The relevant part of the work of Pavilidis was applied to electrocardiograms and carotid pulse wave forms. The shape of these wave forms is very similar to that found in a multicomponent voltammogram. Recently Roberts and Tarassenko[166,167] used electrocardiograms in conjunction with neural networks to determine, successfully, the various sleep states in a human subject based upon the peaks in the waveform. Eberhart[168] used three and four layer neural networks to successfully detect spikes in electroencephalograms, another waveform that is similar to a complex polarogram.

### 2.2.3.6. Process Control

Process control by neural networks was the subject of the work by Levin[169] who used a simulated plant and Jang[170] who looked at the inverted pendulum problem. Lister,[171] as well as applying an MLP to the field of plasma control in Tokamak research, looked at the problem of de convoluting synthetic overlapping gaussians by modelling the two gaussians using a three layer MLP. The work of these authors seems to suggest that the non linear modelling capabilities of neural networks made them strong candidates for real time control. Appendix two contains

details of work similar to that of Lister, that was carried out by the author of this thesis using the $sech^2$ curve to model overlapping voltammetric peaks.

The work reported in this thesis investigated the use of three types of neural network. The two layer perceptron, the multi layer perceptron and the general regression neural network.

# 3. Review of Previous Work and Justification of Aims

## 3.1. Use of Computerised Learning Machines and Artificial Neural Networks in Analytical Chemistry

An early mention of using computerised pattern matching in analytical chemistry is by Jurs[172] et al. This team investigated the predictive ability of binary classifiers modelled along the lines of the Adaline but cascaded in a tree formation. This enabled them to make useful decisions about the presence or absence of a particular atom in low resolution petroleum mass spectrometry data obtained from the American Petroleum Institute Research Project 44 . Although this required 300 training patterns, with each one consisting of 155 peak positions, they achieved an average of 90.6% in recognising the presence of an oxygen atom in a test set of spectra.

Using the same data Kowalski[173] extended this work by modifying the algorithms in order to produce a continuous real valued output where different values of output corresponded to different oxygen numbers. For example, an output of 1 might represent 1 atom and output of 2 might represent 3 atoms so that an intermediate value of 1.7 might represent 2 atoms. They found that the interpolation of output values was not linear but nevertheless a simple form of multi component classification could be performed, producing a classification of one out of several possible categories. Both of the previous authors went on to apply similar

techniques to infrared spectra[174] and a combination of infrared spectra, melting point and boiling point[175]

The previous work was carried out using pure spectra, that is spectra derived from only one compound. Wangen[176] artificially combined seventeen 14 MeV neutron induced gamma ray spectra in order to produce mixtures containing various known percentages of each element. These were then subjected to pattern classification using a binary classifier. The binary target was a value indicating whether or not the pattern contained a given percentage, 0.1%, 1.0% or 10.0%, of the element in question. The classifier thus performed semiquantitative classification. Prediction percentages were as high as 99.7% but it must be noted that, whilst this was not using pure spectra, it was achieved on different mixtures of the same set of spectra in the absence of noise and statistical deviation.

An architecture resembling modern neural networks was later investigated by Jurs[177], again using the petroleum mass spectrometry data. In addition to applying transformations to the input data of binary classifiers, Jurs adapted the architecture to create a 'committee machine' using two layers of normal binary threshold logic units. The first layer of units consisted of multiple sets of ordinary binary classifiers, each of which was presented with the input pattern and formed an output corresponding to the most probable classification. Instead of using this output directly however, Jurs fed it into a second classifier which used as its input pattern, the vector of outputs from the previous layer. This classifier then produced a final classification based upon the one decided upon by the majority of the first level units.

An interesting approach to the production of the training and test sets was used by Jurs. A subset of the entire data set was used. Training patterns that participated in training were retained and the rest removed from the training set. This was

repeated for different subsets until only 172 training patterns were required to produce 100% classification on all 600 original patterns. However, since at one time or another all patterns had influenced the training, either by being selected to be in the training set or just as importantly, by not being selected, there no longer existed an independent set of patterns by which the efficacy of the classifier could be measured.

Up to this point however, no serious consideration had been made of the experimental deviations present in real analytical data, nor had the classification results been evaluated on anything other than a statistical basis. In 1971 Sybrandt[178] chose stationary electrode polarography as the source of experimental data used to investigate the qualitative analysis of mixtures involving overlapping instrumental signals.

Sybrandt used a similar process to Wangen to produce the pattern vectors, starting with data from a vector of polarographic current measurements at 2mV intervals produced by the separate analysis of single component Cd(II), In(III)and Sb(III) samples each at $5\times10^{-5}M$. In all 36 vectors were produced containing the expected statistical variations in peak position, peak height and background noise.

Sets of three vectors, one for each species, were selected and each one multiplied by a random number. The three vectors were then added to produce semi synthetic three component mixtures at various relative concentrations. The classifier's task was to qualitatively identify one particular species in the presence of peaks due to other species and, in the case of Cd(II) and In(III) with peak potentials separated by approximately 40 mV, in the presence of overlapping peaks.

In spite of this overlap, Cd(II) was identified correctly up to a 1:10 Cd:In concentration ratio, although the ability on previously unseen test set data as well as for some of the training set data fell to a 1:4 ratio.

A later development of this work by Sybrandt[179] saw the first use of parameterised input vectors being presented to a classifier, in this case using features extracted from the zero, first and second derivatives of the polarograms. Classification accuracy improved to 90% for 20:1 peak height ratios.

Synthetic mixtures of real, pure, components were also used by Tunnicliff[180], in an effort to train a learning machine to predict the properties of samples from their mass spectrographs. Also included in the training set were synthetic NMR spectra, refractive index and density. The machine was tested using real experimental data. Whilst the work showed some degree of success, the synthetic NMR data was generated without any additional noise. Thus too much weight was given to this part of the input and the resulting performance was not as good as it could have been if real NMR data had been used.

Over the next few years a growing number of attempts were made to use pattern recognition procedures in analytical chemistry. Thomas[181] investigated parameter optimisation in Anodic Stripping Voltammetrty (ASV), Byers[182] looked at structural characterisation of herbicides using cyclic staircase voltammograms as the source of his data, processing these using the Fourier transform and a k nearest classifier and Kaplan[183] embedded the software into a complete, although rather primitive, computer driven instrument.

The processing of analytical chemical data by Partial Least Squares Regression began to appear around 1983 although the technique grew out of work by Wold[184] in the late sixties. Between 1981and 1983 Wold used this method on multivariate

problems in econometrics and the social sciences[185,186]. This use of PLS concentrated upon so called 'mode B' PLS i.e. a formative model of the systems. The chemometric version of PLS, presented in 1983 by Wold et al[187] fell under the so called 'mode A' or reflective PLS model. The difference between mode A and mode B is primarily that in Mode B only one major contributing factor is extracted from the multivariate data in order to explain the relationships between several blocks of data. Mode A on the other hand is concerned with extracting as many factors as are necessary in order to adequately model the relationship between just two blocks of variables, the known X 'spectrum' and the unknown Y giving rise to the spectrum.

The technique was adopted by Lindberg[188] in 1983 who, in addition to presenting an algorithm for the process, showed that even severely overlapped spectra could be resolved by this method. However it was claimed in 1986 by Geladi and Kowalski[189] , correctly in this author's opinion, that much of the literature on PLS did not present the method in a complete and easily understandable manner. Geladi and Kowalski attempted to correct this in a tutorial on the method which remains one of the few such papers to graphically describe the process. Since then the technique has been extended by Wold[190,191], first in 1989 and again in 1992 to incorporate non linear PLS modelling. In this latter paper he also lays down a challenge of comparing non linear PLS with a neural network approach, drawing a parallel between the data processing requirements of the two techniques.

Comparisons between PLS and other multivariate techniques when applied to chemical analysis were made by Haarland and Thomas in 1988[192,193] who used both simulated and real infrared data from glasses to quantify components of the glass related to quality control. Whilst their work indicated that PLS compared favourably with other techniques, for example principal components regression, they were

unable to fully justify their findings due to the restricted number of samples used. De Weijer[194] compared PLS with neural networks in the identification of yarns from their physical properties and concluded that neural networks were superior.

The use of true neural networks in analytical chemistry made their first appearance in 1990 when Long[195] trained a back propagation network to estimate the protein content of wheat from its near infrared spectrum and investigated the effect of changing the network parameters. He concluded that the linear transfer function performed better than the sigmoid in this application. His results using a neural network were as good as, or very slightly better than using principal component regression. Long went on to use back propagation to classify jet fuel using a 'one out of n' output encoding system applied to chromatographic data[196].

Since these papers the use of neural networks and associated techniques in chemistry has increased significantly. Wythoff[197] used a simple back propagation network to identify signal peaks in infrared spectra. The system was essentially interactive and so heavily dependant on the users choice of input data but an important conclusion in this work was that including in the input data points corresponding to noise improved the network's ability to differentiate between a noise peak and a true peak. Many other researchers[198,199,200,201,202] have also used or compared combinations of principal component analysis and neural networks, although most applied the techniques to infrared spectra.

A three layer neural network was used by Gemperline[203] who compared its use with PCR when processing both simulated and real UV spectra. He concluded that the network performed better in the presence of noise than did PCR. The question of noise was taken up by Allanic[204] who used a neural network to identify fluorescence spectra. He found that the neural network could cope well with an artificially added 70 % noise. Borggaard[205] used an interesting combination involving using partial

least squares (PLS) and neural networks applied to near infrared spectra of latex spheres in water, where the percentage of latex was required and in those obtained from meat when determining the fat content. He found that networks could perform much better than PLS and PCA in this situation.

Fluorescence spectra were used by McAvoy[206] who found that the backpropagation MLP out performed linear techniques in deconvoluting fluorescence spectra obtained from mixtures of amino acids. They were also used by Blank[207,208] who accepted the challenge of Wold and used both real and synthetic data to investigate non-linear forms of PLS and PCR compared to a Back Propagation neural network. He concluded that the neural networks only performed better than the non connectionist methods when the sizes of the training sets used are sufficiently large to constrain the model yet extensive enough to provide adequate sampling of the mapping surface.

Data from Pyrolysis Mass Spectrometry has been used by Goodacre and processed using neural networks when analysing amino acids in glycogen[209], investigating bioprocesses[210] and for detecting the adulteration of olive oils by other seed oils[211]. A sigmoidal transfer function in a back propagation network was used for this work and the results, although impressive, do show marked non linearity at the two extremities of the calibration range due to the compression afforded by the sigmoid. The majority of the networks trained by the author in the development of this thesis use linear transfer functions instead on the output layer and the resulting networks do not suffer from the above problem.

The problem of deconvoluting overlapping peaks was investigated using a neural network by Gallant[212] who used a neural network architecture originally designed as an adaptive controller for robots. Synthetic chromatograms were processed by feature extraction from the zero and second derivatives. The network

was found to provide rapid de convolutions over a range of peak heights, widths and resolutions.

Recent sources relating to current applications of neural networks in chemistry can be found in Gasteiger[213], who discusses network architecture's and applications in chemical reactivity, process control, elucidation of chemical structure from spectra and secondary structure of proteins and De Weiger who discusses the use of neural networks to perform peak fitting. In this paper he uses a Pearson VII function to model the peaks. This method has also been used with good results by this author (see appendix two).

Kateman[214] on the other hand, challenges the acceptance of neural networks in analytical chemistry and offers guidelines on validation techniques necessary for general acceptance.

## 3.2. Observations from previous work

Nearly all of the literature on the use of neural networks in processing analytical chemistry data used infrared spectra as its starting point. As yet, apart from the author's own work, nothing appears to have been published concerning their use in qualitatively processing data from voltammetric analysis. This is significant as the nature of IR spectra tends to lend itself to neural network processing as each input may simply be devoted to a particular frequency. By analogy it may appear that when processing data from differential pulse stripping voltammetry each network input may be devoted to a particular potential. Whilst in broad terms this is correct, and is the approach taken in this work, it must be remembered that the position of the stripping peak along that potential axis will vary with scan number (see chapter 5), so complicating the issue.

The works comparing PLS and neural nets also appears to have been targeted on IR spectra, except where simulated data is used. In addition the general consensus to this point is that there is little difference between the performance of PLS and neural networks. One exception is the work by Næs[215] who showed that whilst the equations for the regressions formed by PLS and feed forward, backpropagation type neural networks were essentially the same, the neural networks performed better than the traditional methods which tended to over fit the data. Næs concluded that if neural networks are used "in the right way" then they could be a good method for prediction. However he felt that they were more difficult to understand and that their results were more difficult to visualise.

## 3.3. Justification of the project aims

Few papers address the problem of automating the calibration process or discuss the merits of sophisticated calibration over the method of standard addition. A notable exception is Bond[216], who compared linear calibration and standard addition methods for the quantification of trace level of cadmium in sea water, concluding that direct calibration methods may be used. Williams[217] also discusses this problem when he describes a hand held instrument for environmental monitoring of cadmium, copper and lead using ASV and a linear peak height calibration method. Palys[218] used a computer to design, run, collect and analyse data from electrochemical experiments involving staircase voltammetry but the system used in this case was an expert system not a neural network and the object was to gain information regarding the reaction mechanisms. Finally Smits[219] briefly mentions the use of instrumental parameters as an additional network input when classifying algae using flow cytometer data operating at different laser power settings.

This work, in addition to applying neural networks in a field that has so far received little attention, will show that neural networks may be applied to quantitative chemical analysis in a way that is easy both to understand and visualise. It will also show that the error and confidence in the concentrations made are superior to those obtained by other methods.

# 4. Construction of Experimental Apparatus

## 4.1. Construction of the working electrode and analytical cell

### 4.1.1. The working electrode

The working electrode itself, onto which the mercury film is plated, consists of a small disc of glassy carbon that has been polished to a mirror finish. Figure 9 shows the general arrangement and a photograph of the actual electrode manufactured by the author is shown in figure 10.

A sleeve was constructed from a 2.5 cm diameter rod of PTFE that was drilled to accept a 2 mm diameter silver steel rod and then turned down to an outside diameter of 6 mm. A short rod of glassy carbon was then inserted as a tight push fit with a small phosphor bronze spring between it and the silver steel rod. The other end of the silver steel rod was then arranged on small ball bearings in a holder, also turned from PTFE, so that the steel rod, the 6 mm PTFE sleeve and the glassy carbon electrode could rotate about their mutual axis.

A brass pulley and a silver slip ring arrangement were manufactured and fitted to the top of the rod so that a drive belt could turn the electrode whilst an electrical connection could be made via the slip ring to the silver steel rod and hence to the spring and glassy carbon electrode. The drive motor for this was salvaged from a

computer disc drive and, although not yet utilised in this work, has a tachogenerator attached which could be used for accurate speed control.

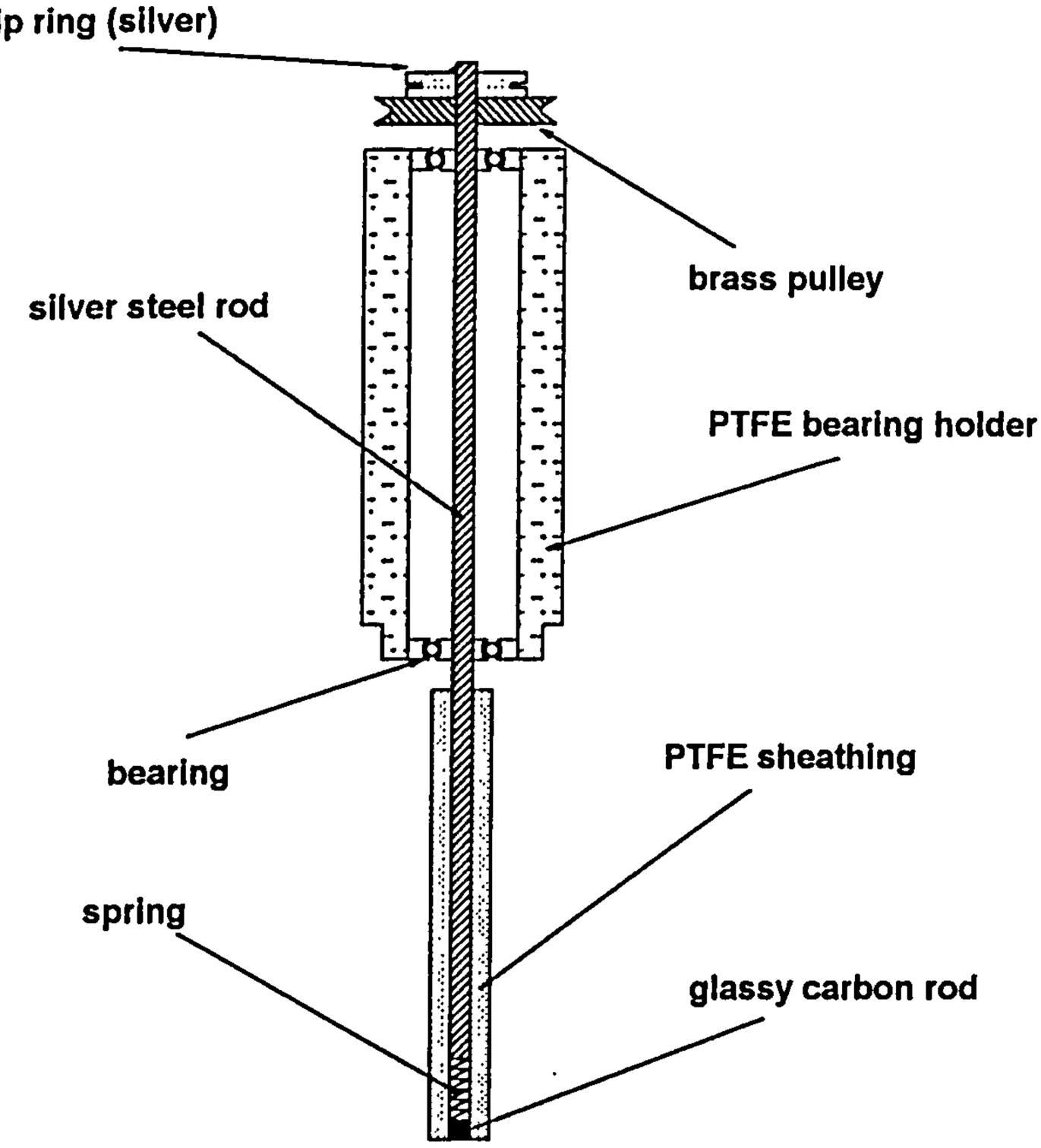



Figure 9. Cross section of the working electrode manufactured for this work

Once manufactured, the end of the electrode was polished until the glassy carbon rod had a mirror finish and the surface of it lay in the same plane as the surrounding, polished PTFE.

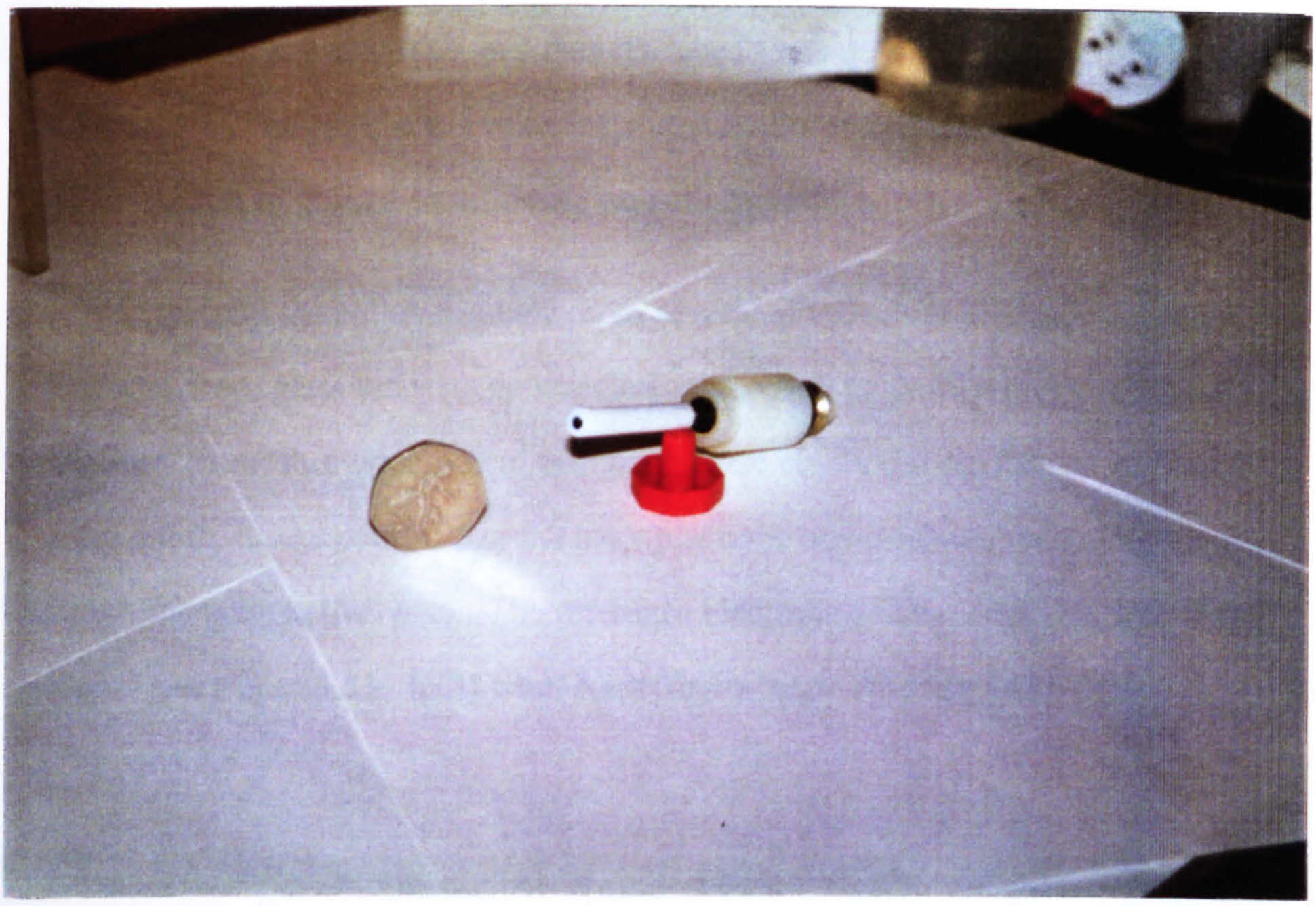


Figure 10. Photograph of the working electrode manufactured for this work

### 4.1.2. The analytical cell arrangement

The cup which contains the analyte to be studied is a disposable 20 ml polystyrene beaker. The author constructed a holder arrangement out of perspex which allowed the cup to be easily positioned and removed whilst ensuring that the relative positions of the cup and the three electrodes remained the same every time it was used. A photograph of the arrangement is shown in figure 11.

The working electrode could also be removed easily for cleaning and polishing as the slip ring brushes were mounted on a swinging arm which could be moved out of contact with the slip ring.

The auxiliary electrode was manufactured from a short length of platinum wire sealed into a glass capillary. Following the example of Svensmark[27], this electrode was connected to ground to minimise capacitance effects.

Originally the author manufactured several silver-silver chloride reference electrodes from glass tubes using standard methods. The porous plugs for these were reclaimed from fibre tipped pens. Although these electrodes performed quite well in earlier work, it was noticed that the peak positions obtained were not consistent enough for quantitative work. The reference electrode finally used was a refillable calomel electrode (3.5M KCl) from Kent Taylor (part Number 1431-210).

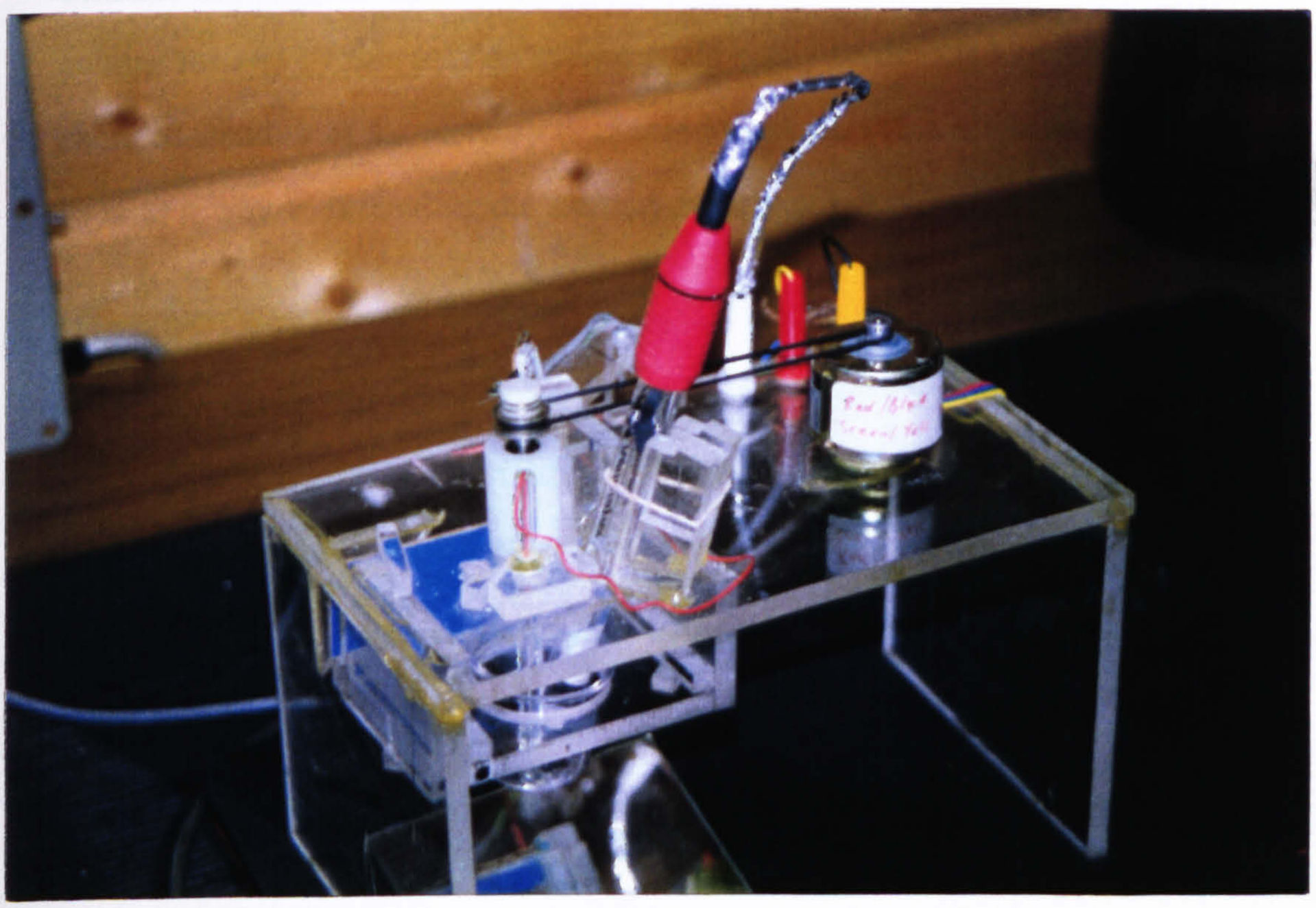


Figure 11. Photograph of the analytical cell arrangement built by the author

A mirror fitted below the electrode was used whilst the electrode was rotating to confirm that the plating of the electrode with mercury was proceeding normally. The relatively thick films formed during pre-plating could be seen as a dull finish slowly appearing on the previously shiny electrode surface.

The films formed during in situ plating are too thin to dull the electrode appreciably but an novel alternative way of viewing the progress was used. A bright light was reflected from the mirror onto the rotating glassy carbon in such a way that when the electrode was viewed in the mirror, the filament of the lamp could be seen reflected again from the surface of the glassy carbon. As the mercury plates the electrode, light being reflected from the mercury droplets travels a shorter distance than that being reflected from the carbon surface. The thin film interference pattern formed by this path difference could easily be detected as colours which changed steadily as the mercury thickness increased. (This might also have been used as a way of measuring the mercury thickness.)

## 4.2. Construction of the electronic hardware

### 4.2.1. Brief review of relevant literature

From Ohm's law, it can be stated that the potential difference $V$ across a resistor $R$ carrying a current $I$ is given by

$$V = IR \qquad \{12\}$$

All systems that supply a current possess internal resistance and this causes an internal potential drop, which increases as the current increases. This manifests itself as a reduction in the output potential available over the component being supplied with current. In voltammetry the apparent resistance of the cell changes as the redox reactions take place resulting in a changing current through the internal resistance of the supply. Potentiostats are electronic circuits that, within a certain range, overcome this effect and maintain an accurate and constant pre-defined output potential regardless of the demand for current.

A reasonable source of potentiostatic applications can be found in von Fraunhofer[220], which, although rather out of date, does give the theoretical background and schematic circuit diagrams for potentiostats. Blanchet[221] gives the circuit for a potentiostat and the schematic for a computer controlled interface. Stabilisation is a problem in potentiostats where high gain amplifiers are used with very high input impedances. Cath[222] gives a rather simplistic overview of current measurement procedures whilst over several years Booman[223], Brown[224],

Kanazawa[225] and Meyer[226] independently investigated stabilisation methods with potentiostats with a view to maximising the stability of the circuit.

More sophisticated applications are given in Tallman[227], who discusses a computerised potentiostat capable of measuring current to a full scale deflection of 100 pA, Harrington[228,229] who considered multiple electrode potentiostats and Candela[230] who reported on a digital pulse generator capable of producing 10 ns resolution pulses with a duration of 330 ns to several hours. However, although having a finer resolution, this pulse generator was less versatile than the function generator produced by Bond[231] ten years earlier.

### 4.2.2. Overview of the hardware

The requirements of the control hardware were that it should be capable of;

- maintaining the potential of an electrode at a value dictated by the control software
- measuring the small current passing through this electrode
- be able to store one measurement, take another and find the difference between the two values
- be able to operate with a wide range of currents and current differences
- be able to offset the current measurements by a variable amount

The system should be capable of performing the above operations under the control of the microcomputer as far as possible so that easily reproducible parameters could be obtained. The control electronics were therefore built in such a way that control signals on various lines (called Address lines, A1 to An, in this

work) could control the operation of analogue switches and multiplexors used to alter the operational parameters of the hardware.

The microcomputer chosen to run the software and to interface with the external control hardware was an IBM AT type machine. This machine is widely available in laptop and desktop versions and so the system would lend itself to portability. The microprocessor in the development system was an Intel 80386 running at a speed of 40 MHz. A 40 MHz IIT-3C87 maths co-processor was installed and the system had 8 Mb of memory.

### 4.2.4. The Programmable Input Output Card

The card used to interface the computer to the analytical electronics was a 48 line programmable I/O card (PIO-48, from Blue Chip Technology). This card uses two 8255 chips. These each appear to the software as four contiguous bytes, the first three bytes are the data ports which perform the communication with the external device and the last byte is a control register which controls the function and operation of the other three ports.

Each chip refers to its data ports as A, B, C. Since this card uses two chips, the ports are referred to in this work as A, B and C for the first three data ports (lowest address in the memory map) and A', B' and C' for the last three ports. The card was installed at a base address of 300hex. Thus the relative addresses of the various ports and control registers can be found from table 3 on page 69.

The 8255 can be configured in various ways, setting each port individually to eight bits input or eight bits output. The exception to this is port C (and port C' in the case of this board). This port can be configured so that the four high and four low bits can be independently set as outputs or inputs. Advantage was taken of this to give twelve bits wide input and output by assigning Port A and the lower four bits

of port C as outputs (data bits 0 to 7 through A and bits 8 to 11 through C) and port B together with the upper four bits of port C as inputs (bits 0 to 7 through port B and bits 8 to 11 through port C).

All the ports in the second chip (A', B' and C') were configured as outputs in order to control various functions on the electrode control board. The various bits in these ports are referred to in this work as 'Address Lines' and given the designation A0 to A23.

| Port | Bit | Function | Assignment |
|---|---|---|---|
| Port A | Bit 0 | Output | D/A DB0 |
| Port A | Bit 1 | Output | D/A DB1 |
| Port A | Bit 2 | Output | D/A DB2 |
| Port A | Bit 3 | Output | D/A DB3 |
| Port A | Bit 4 | Output | D/A DB4 |
| Port A | Bit 5 | Output | D/A DB5 |
| Port A | Bit 6 | Output | D/A DB6 |
| Port A | Bit 7 | Output | D/A DB7 |
| Port B | Bit 0 | Input | A/D DB0 |
| Port B | Bit 1 | Input | A/D DB1 |
| Port B | Bit 2 | Input | A/D DB2 |
| Port B | Bit 3 | Input | A/D DB3 |
| Port B | Bit 4 | Input | A/D DB4 |
| Port B | Bit 5 | Input | A/D DB5 |
| Port B | Bit 6 | Input | A/D DB6 |
| Port B | Bit 7 | Input | A/D DB7 |
| Port C | Bit 0 | Output | D/A DB8 |
| Port C | Bit 1 | Output | D/A DB9 |
| Port C | Bit 2 | Output | D/A DB10 |
| Port C | Bit 3 | Output | D/A DB11 |
| Port C | Bit 4 | Input | A/D DB8 |
| Port C | Bit 5 | Input | A/D DB9 |
| Port C | Bit 6 | Input | A/D DB10 |
| Port C | Bit 7 | Input | A/D DB11 |

Table 2. Programmers model of PIO data ports

These assignments are summarised in table 2 on page 68 which shows the programmer's model of the data lines and table 3 on page 69 which shows the

programmer's model of the control 'Address Lines'. The need for the control functions shown will be explained shortly.

| Port | Bit | Ribbon Pin No. | Centronics Pin No. | Function | Address Line No. |
|---|---|---|---|---|---|
| A' | 0 | 2 | 26 | I/E Current Offset Range | 0 |
| A' | 1 | 4 | 27 | ---- " ------" ------" --- | 1 |
| A' | 2 | 6 | 28 | I/E Sensitivity | 2 |
| A' | 3 | 8 | 29 | ---- " ------" ------" --- | 3 |
| A' | 4 | 10 | 30 | ---- " ------" ------" --- | 4 |
| A' | 5 | 12 | 31 | ---- " ------" ------" --- | 5 |
| A' | 6 | 14 | 32 | ---- " ------" ------" --- | 6 |
| A' | 7 | 16 | 33 | ---- " ------" ------" --- | 7 |
| | | | | | |
| B' | 0 | 18 | 34 | Delta Amp Gain | 8 |
| B' | 1 | 20 | 35 | ---- " ------" ------" --- | 9 |
| B' | 2 | 22 | 36 | RLY4 D-Amp input pin 19 Gnd (1) / from SH2 (0) | 10 |
| B' | 3 | 24 | 37 | RLY5 Gnd WE (1) / not Gnd (0) | 11 |
| B' | 4 | 26 | 38 | S/H 1 | 12 |
| B' | 5 | 28 | 39 | S/H 2 | 13 |
| B' | 6 | 30 | 40 | R/C for A/D Converter | 14 |
| B' | 7 | 32 | 41 | | 15 |
| | | | | | |
| C' | 0 | 34 | 42 | Gas Purge on (RLY6) | 16 |
| C' | 1 | 36 | 43 | Electrode Rotate On (RLY7) | 17 |
| C' | 2 | 38 | 44 | | 18 |
| C' | 3 | 40 | 45 | | 19 |
| C' | 4 | 42 | 46 | I/E Offset | 20 |
| C' | 5 | 44 | 47 | ---- " ------" ------" --- | 21 |
| C' | 6 | 46 | 48 | ---- " ------" ------" --- | 22 |
| C' | 7 | 48 | 49 | ---- " ------" ------" --- | 23 |

Table 3. Programmers model of PIO control 'address lines'

## 4.2.5. The Electrode Control and Current Measuring Hardware

An interface circuit between the PIO board and the electroanalytical cell was constructed by the author. The entire circuit performed several discrete functions.

The individual functions were;

1. Conversion of the digital signal from the computer into an analogue potential.

2. The application of this potential to the working electrode via a potentiostat.

3. Filtering the resulting cell current through a suitable filter to remove unwanted noise.

4. The measurement of the instantaneous cell current by converting it to a potential using one of several sensitivity levels, with the facility to offset any large dc component.

5. Storing a previous converted measurement whilst making another and calculating the difference between them, with amplification if necessary.

6. Converting the final potential to a digital signal for transmission to the CPU.

These functions were designed and implemented in hardware as separate modules. The relationships between the various modules is shown in figure 14 on page 74. Their function, operation and construction is discussed separately below with further details in appendix 6.

Photographs of the equipment are reproduced in figures 12 and 13.

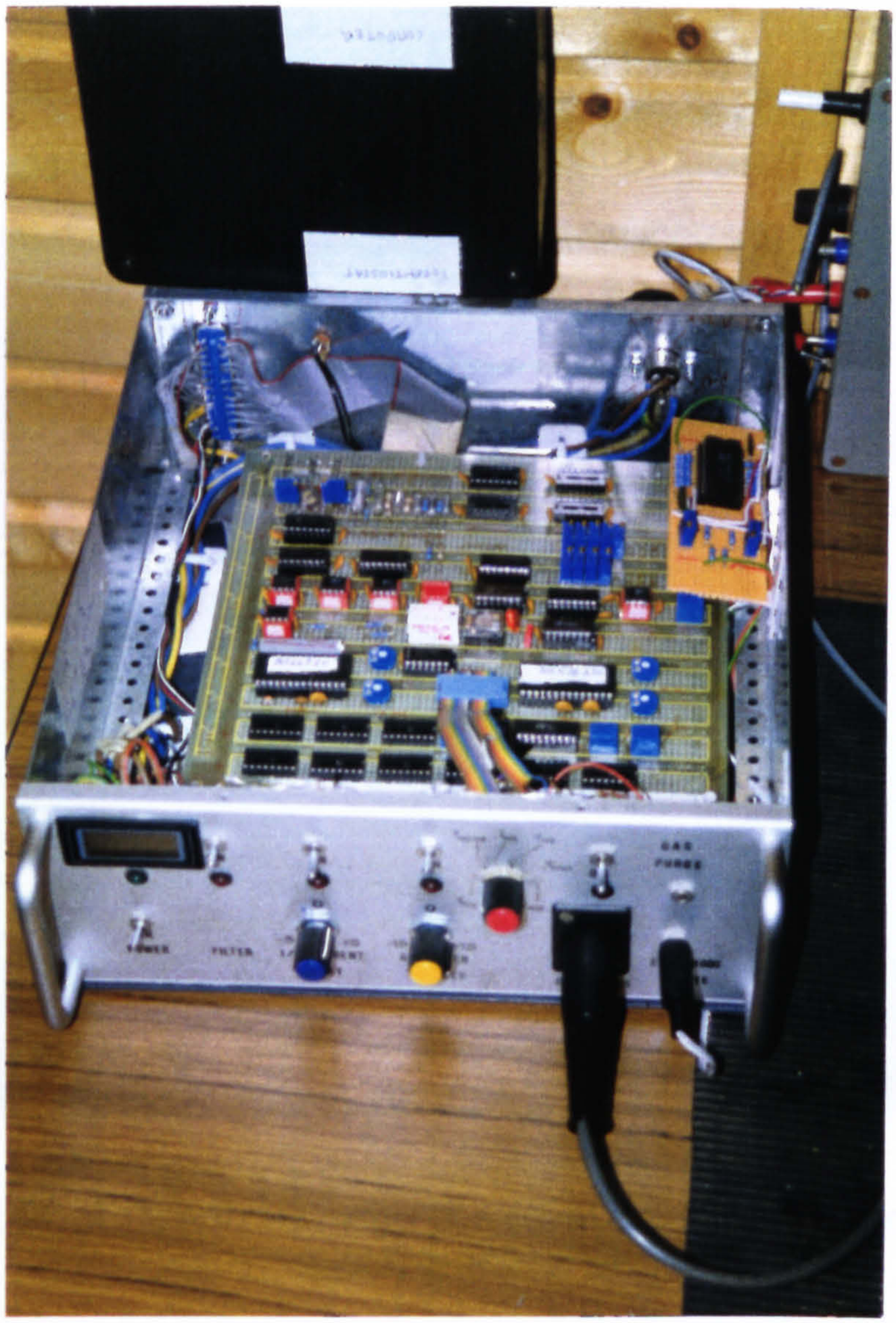


Figure 12. Photograph of the potentiostat and control electronics built by the author

Figure 12 shows the internals of the potentiostat, interfacing and control electronics. The front panel of this equipment is detailed in figure 13, where the digital display and various manual controls may be seen.

The entire circuit was built on a prototyping expansion card originally supplied as part of some Nascom 3 Z80 microcomputer development systems. When building was complete this card was fitted into a metal box with plug and socket

connections for all external devices. A smoothed power supply providing +15v, -15v, +5v, -5v and 0 v was used.


Figure 13. Photograph showing the front panel of the control electronics

The wiring system used was point to point using fine insulated wire. This made subsequent modifications and corrections to the board easier and also allowed the elimination of earth loop problems by taking each ground pin back individually to a common ground point on the A/D converter chip. Power supplies for each device were decoupled at the point of connection to the chip by grounding through a 100nF capacitor.

Switches have been provided on the front panel for switching the filter circuit in or out, for switching in buffered offsets to the I/E converter and the differential amplifier and for isolating all three electrodes from the rest of the circuit. After the

prototyping stage the I/E offset switch and its associated potentiometer circuit were disconnected and this function incorporated on the main board and controlled by the computer. ( see fig 24 on page 88 for the circuit).

In order to facilitate testing, the interfaces between the modules were brought out to a 16 way DIL socket on the board. A separate circuit was built and housed in its own case and fitted with a 16 way DIL plug and a 50 way Centronics plug. This 'test box' circuit, when plugged into the 16 way DIL and the Centronics plug ordinarily used for the computer interface, allowed manual setting of any 12 bit digital input signal, visual confirmation on LED's of any 12 bit output digital signal and a number of other functions. After initial testing was complete the 16 way DIL socket was used as a means of picking up signals which were taken, via a multipole switch, to a digital meter on the front panel. This allows the user to observe either the output potential of the DAC, the potential of the working electrode with respect to the reference electrode, the current flowing in the working electrode, the output potential of the I to E converter or the potential held in the first sample and hold device.

## 4.2.6. The Digital to Analogue Converter and Potentiostat Module

### 4.2.6.1. Construction

The circuit for this module is shown in figure 15 on page 76. The D/A converter was built using the AD 667 digital to analogue converter chip[232]. For this application it was configured to have a 12 bit unlatched input bus and an output range of ±2.5v by grounding all latch address decode pins (12,13,14,15), connecting output range pins 2 to 9 and 1 to 3 together and connecting pins 7 and 4 to pin 6 via 100Ω trimmers. The msb of the12 bit input bus was passed through an

inverter. This enabled two's complement encoding of the digital input to be used, thus simplifying the driving software.

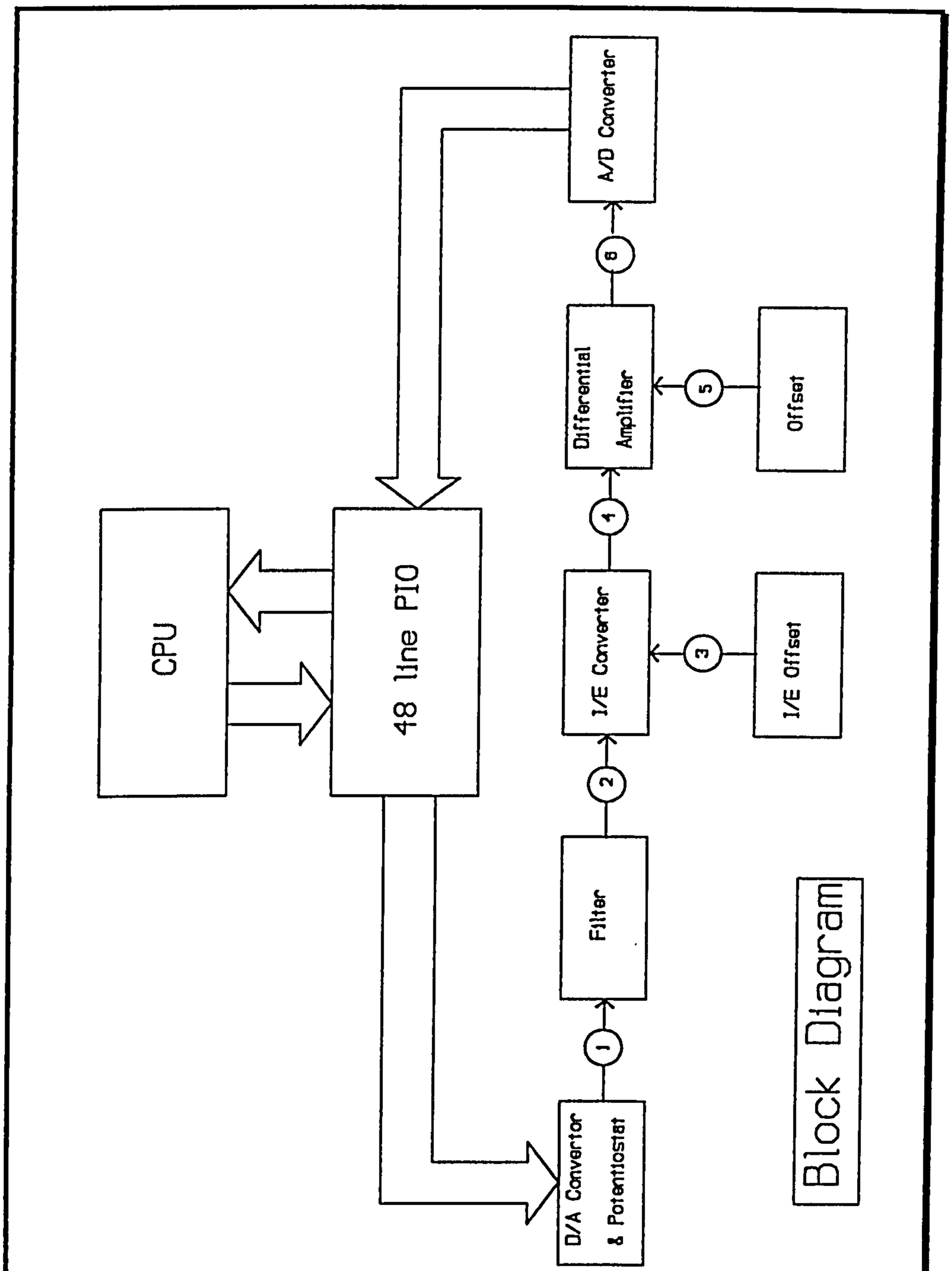



Figure 14. Block diagram showing module relationships

The potentiostat is a simple two op amp design after von Fraunhofer[220] and Jayaweera[233]. A NE5534N was chosen as the control amplifier as it has very low noise characteristics combined with a high output drive capability. The feedback op amp is a AD645JN precision, low noise, FET input device with an input bias current of 3pA. The gain of this device was fixed at unity.

#### 4.2.6.2. Operation

Ignoring capacitive effects, the equivalent circuit of a three electrode electrochemical cell may be considered to be two resistors in series due to the analyte resistance. These are $R_S$, the 'series' resistor, between the Auxiliary Electrode (A.E.) and the Reference Electrode (R.E.) and $R_C$, the 'control' resistor, between R.E. and the Working Electrode (W.E.). The potentiostat attempts to keep the potential difference between R.E. and W.E. constant and equal to the applied potential

Assume that initially no current is flowing and that by changing the digital signal on the DAC the potential on the inverting input of OA1 is made more positive with respect to ground. This results in the output of OA1 becoming more negative. The A.E., connected to this output, will now also become more negative with respect to the W.E. which is connected to virtual ground. i.e. the W.E. will be polarised anodically and stripping may take place from its surface. This will alter the electrical characteristics of the analyte near the W.E. hence changing Rc. If this change is to reduce Rc (as it would be if more ions were now present near the W.E.) then the potential drop across it - and hence between R.E. and W.E. - would decrease.

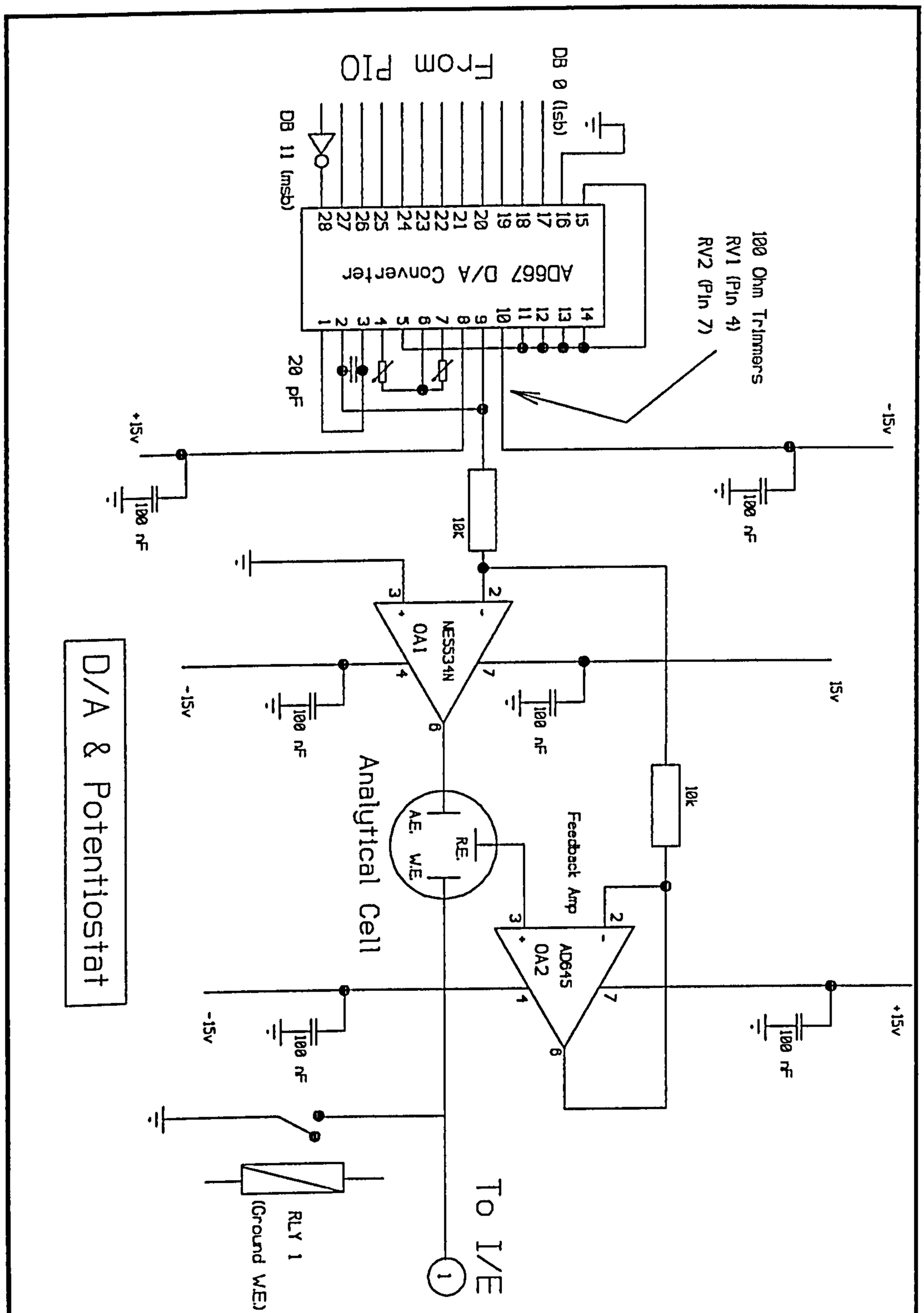



Figure 15. D/A converter and potentiostat module

However, since the W.E. is connected to virtual ground the only way that this can occur is for R.E. to become more positive. This increased positive potential is buffered through the very high input impedance OA2 and applied to the inverting input of OA1. The output of OA1 (and hence of the A.E.) will now become more negative. This increases the current flowing through Rc, increasing the potential drop across it and restoring the potential difference between R.E. and W.E.

In this way the potential of the W.E. with respect to the R.E. is kept constant regardless of changes in the cell impedance due to chemical reactions.

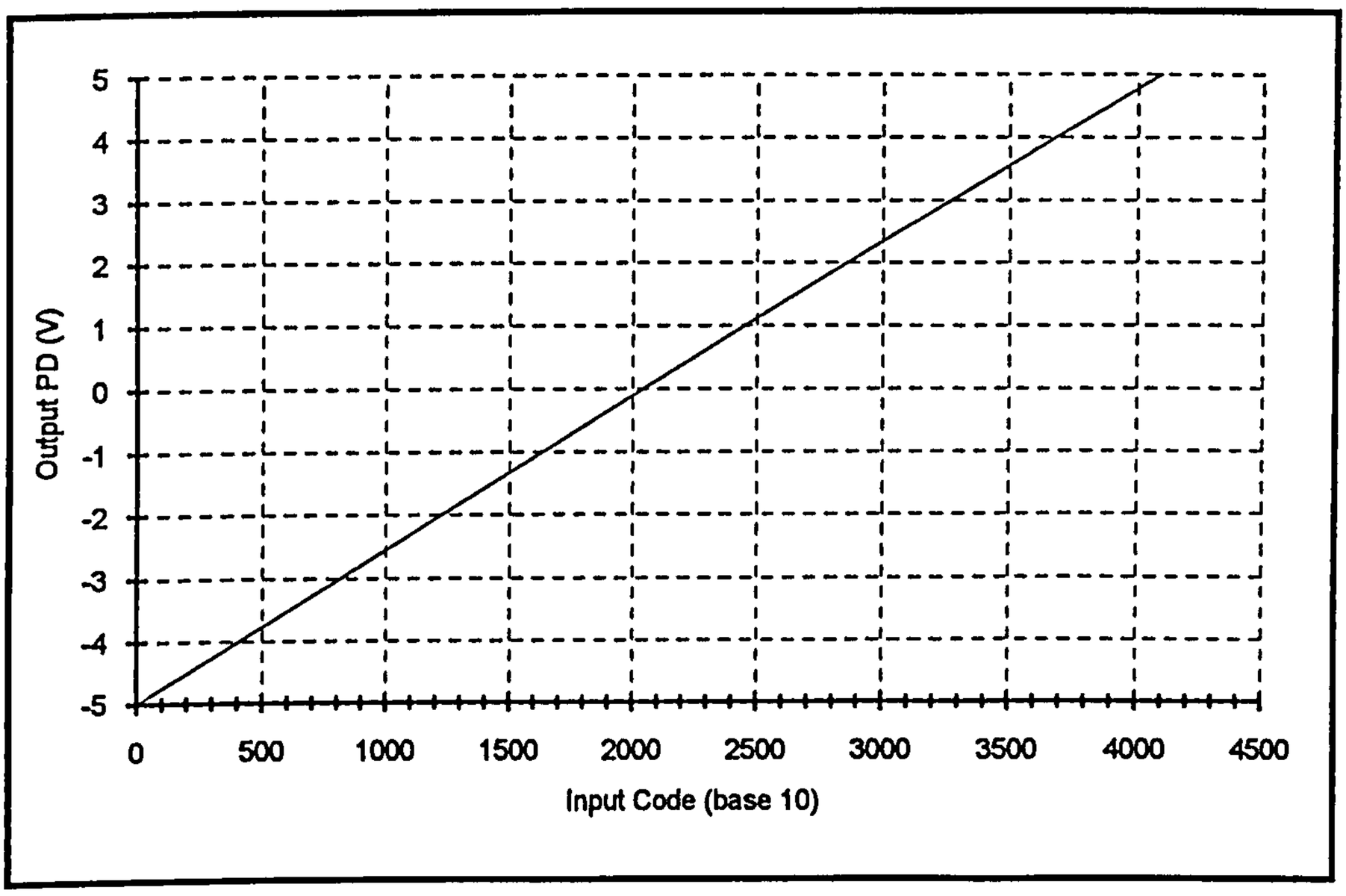



Figure 16. D/A converter calibration

## 4.2.6.3. Performance

The AD667 was calibrated using the test box described earlier. An input code of 0000_0000_0000 was applied and RV1 adjusted until the output was -5.0V. Then an input code of 1111_1111_1111 was applied and RV2 adjusted to give an output

of +4.997V. Finally the output was measured using a digital voltmeter for various input codes and the calibration chart shown in figure 16 on page 77 produced.

The operation of the potentiostat was tested using the dummy cell shown in figure 17 on page 78. The computer software was used to apply a potential of 0.298 V (wrt R.E.) to the working electrode and this potential was measured with a digital voltmeter while the resistance of the potentiometer between the W.E. and the R.E. was slowly reduced. No change in this potential was observed until a critical value of resistance was reached, after which the potential fell rapidly as the potentiostat saturated and was no longer able to maintain the desired potential. Subsequent calculations indicated that, until saturation, the current in the working electrode was steadily increasing from about 3.0 nA whilst the potential was remaining constant.

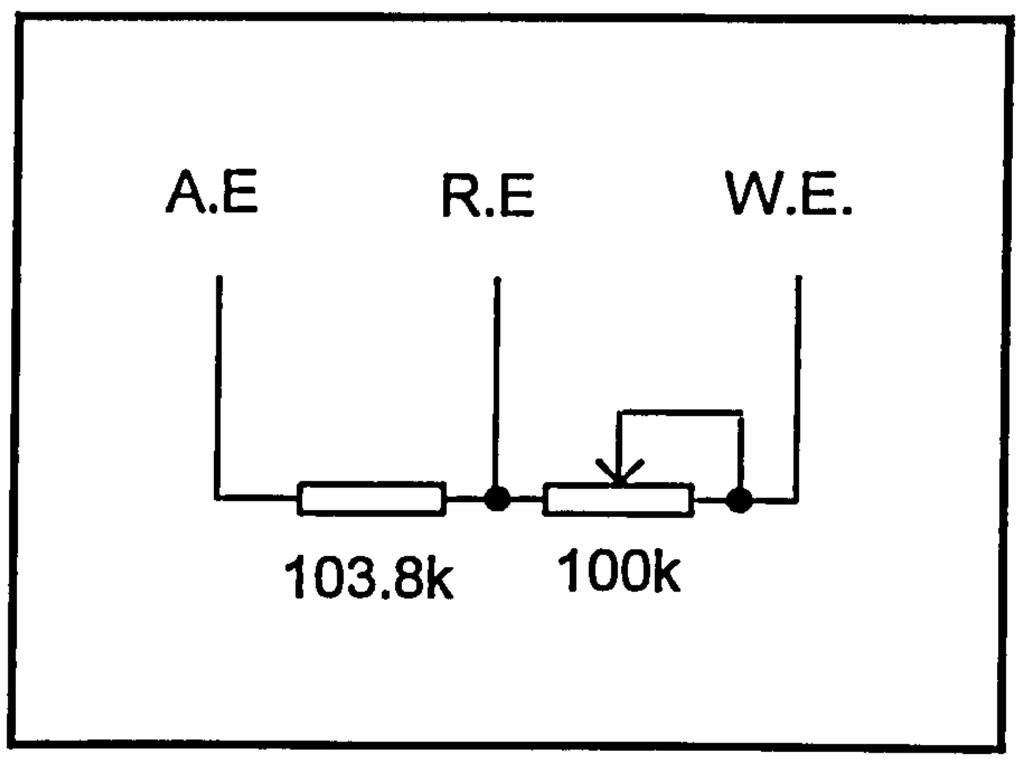



Figure 17. Dummy cell used to test the potentiostat

The operation of the potentiostat was also tested under working conditions by plating lead from a solution of 1ppm lead in 0.1M $KNO_3$ onto a rotating glassy

carbon electrode that had been plated with mercury (using 0.01g/dm$^3$ lead nitrate in 0.1M nitric acid) for 30s. The plate potential applied by the software was changed in 100 mV steps from zero to -900 mV whilst the actual potential of the W.E. was measured wrt the R.E. using a digital voltmeter. During this time the cell characteristics are known to change, the cell current being higher around -500 mV when the lead plates the electrode. Figure 18 shows that the actual potential of the W.E. did indeed follow that applied by the software in spite of the changes in cell characteristics, thus indicating the correct operation of the potentiostat within this range of W.E. current.

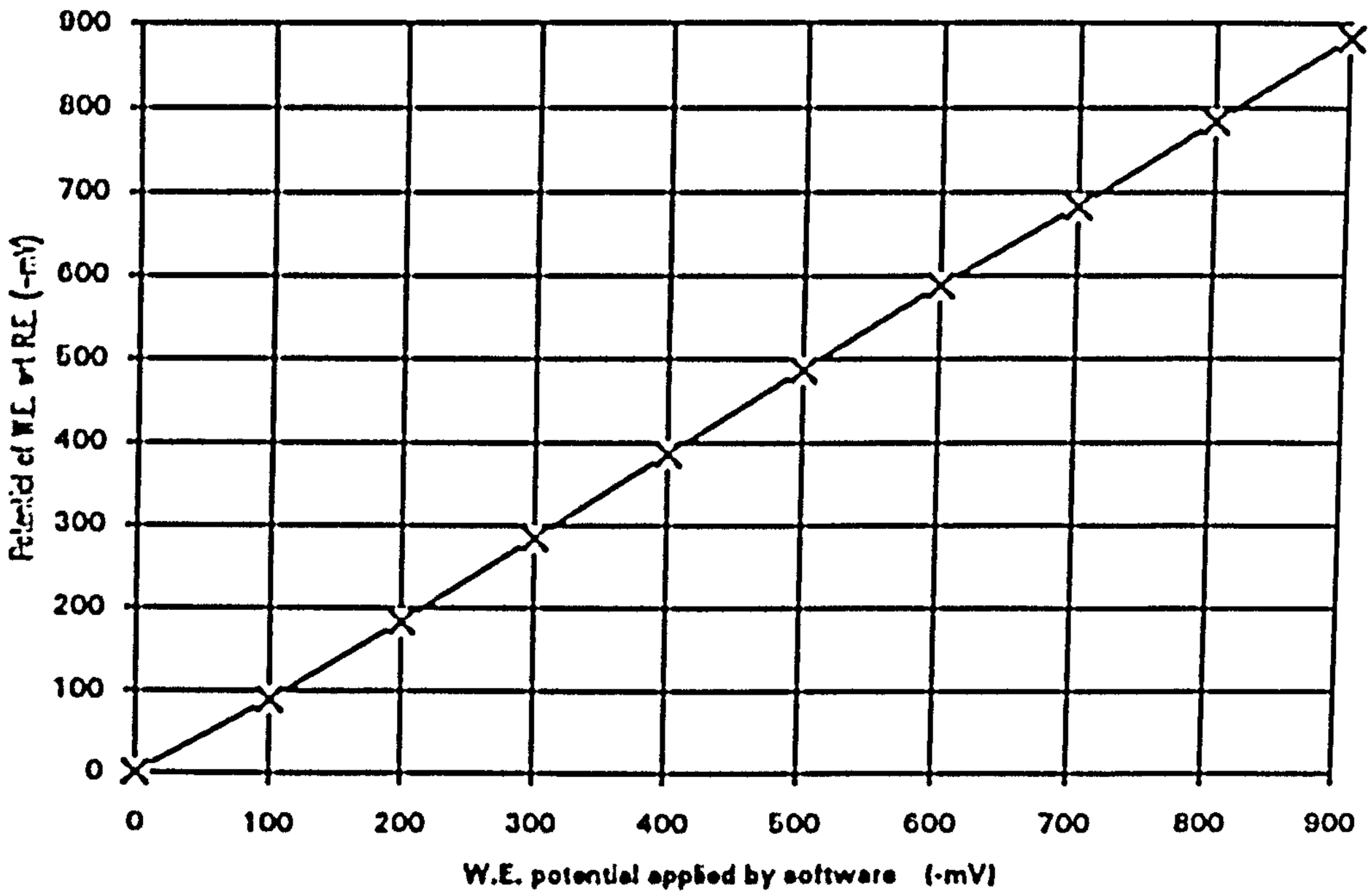



Figure 18. Plot confirming correct operation of the potentiostat.

### 4.2.7.The Filter Module

This part of the circuit was built so that the filter could be excluded via a switch on the front panel and the signal from the potentiostat passed directly to the I/E converter. However it was found that adequate signals could be obtained without filtering. Therefore although provision for its inclusion was available, the filter circuit was not completed and the switch left permanently in the 'filter bypass' position. This part of the circuit is shown in fig 19 on page 81.

### 4.2.8.The Current To Voltage Converter Module

#### 4.2.8.1. Construction

This classical design, shown in figure 20 on page 83, uses a single op amp, the 741 acting as a current follower. Švestka[234] (1986) showed that an improved current to voltage converter with better stability could be achieved by slightly modifying the basic circuit. This involved isolating the input to the op amp via a resistor. In the present circuit this is done using a 4.7kΩ resistor.

The feedback resistor employed can be set to one of six different values, i.e. conversion sensitivities, using one of two AD7590 analogue switches [235]. These packages each contain four, single pole switches, controlled by a separate input lines. The value of each resistor, the sensitivity that it produces and the address line pattern controlling it are given in table 4 on page 82.

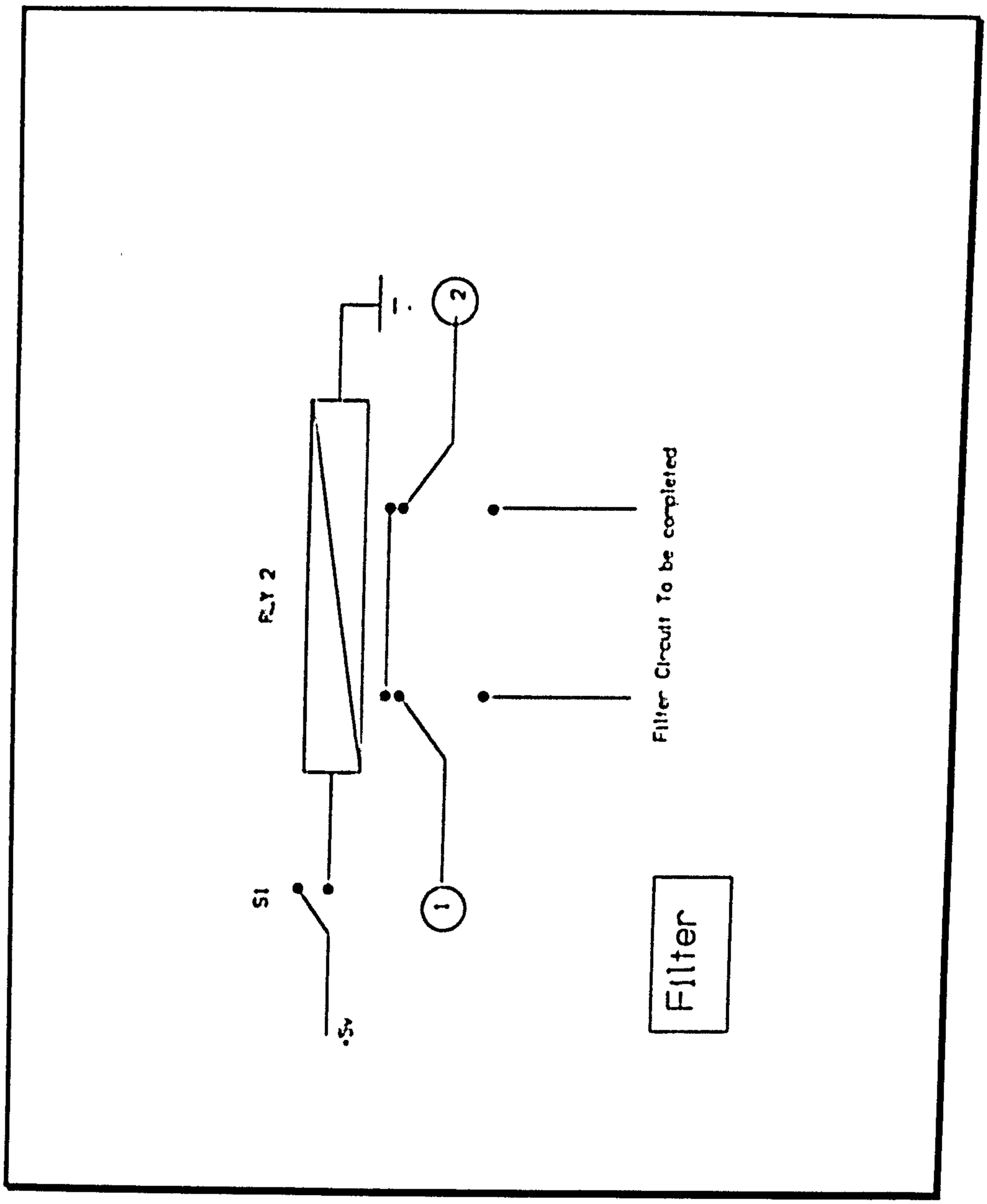



Figure 19. Filter module

| Address Line pattern | | | | | | Feedback Resistance | Nominal Sensitivity nA/V |
|---|---|---|---|---|---|---|---|
| A7 | A6 | A5 | A4 | A3 | A2 | | |
| 0 | 0 | 0 | 0 | 0 | 1 | 120kΩ | 8000 |
| 0 | 0 | 0 | 0 | 1 | 0 | 240kΩ | 4000 |
| 0 | 0 | 0 | 1 | 0 | 0 | 500kΩ | 2000 |
| 0 | 0 | 1 | 0 | 0 | 0 | 1.0MΩ | 1000 |
| 0 | 1 | 0 | 0 | 0 | 0 | 2.2MΩ | 500 |
| 1 | 0 | 0 | 0 | 0 | | 4.7MΩ | 250 |

Table 4. Current to voltage converter sensitivities

Davis[236] discussed the assumption that no current flows into the op amp of a current to voltage converter, i.e. it is at virtual ground. In fact this can never be the case as the Op Amp needs some input current in order to work. In reality the input appears as an inductor whose impedance varies with frequency. With the step inputs expected in the present work this would cause the cell current to overshoot and 'ring'. This would appear as a voltage over the double layer in excess of that applied by the potentiostat. In this case the current to voltage converter output may saturate momentarily.

A method of reducing this apparent inductive input is to shunt the feedback resistor with a capacitor[224,225]. Bezman[237] shows how the value for this capacitor can be deduced. The values used in figure 20 appear to be adequate.

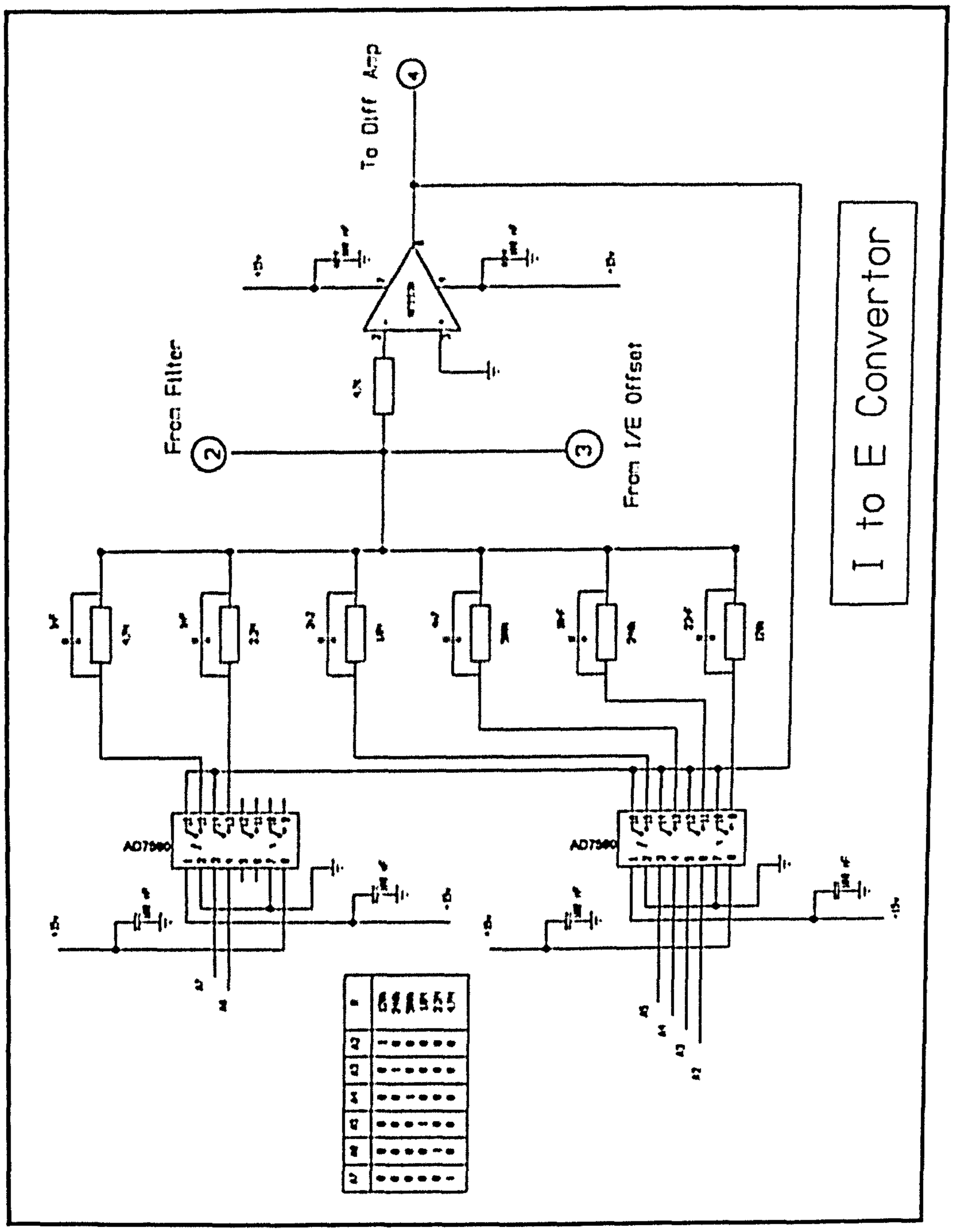



Figure 20. Current to voltage converter module

### 4.2.8.2. Operation

Once the appropriate signals are set on address lines A2 to A7, a particular feedback resistor and impedance matching capacitor combination are selected. The current $I_{w.e.}$ flowing from the cell is summed with the offset current from the offset module and is presented to the inverting input of the 741 Op Amp via the 4.7kΩ resistor. The 741 produces an output voltage over the feedback resistor which attempts to balance that on the input. This produces a current equal to but in the opposite sense to that applied, producing the virtual earth. Ignoring the inductive effect, the potential at the output is given by

$$E_{out} = R_{fb} \times I_{in} \qquad \{13\}$$

where $R_{fb}$ is the feedback resistor and $I_{in}$ is the cell current.

### 4.2.8.3. Performance

The performance of the Current to Voltage converter was investigated by using the dummy cell shown in figure 21.

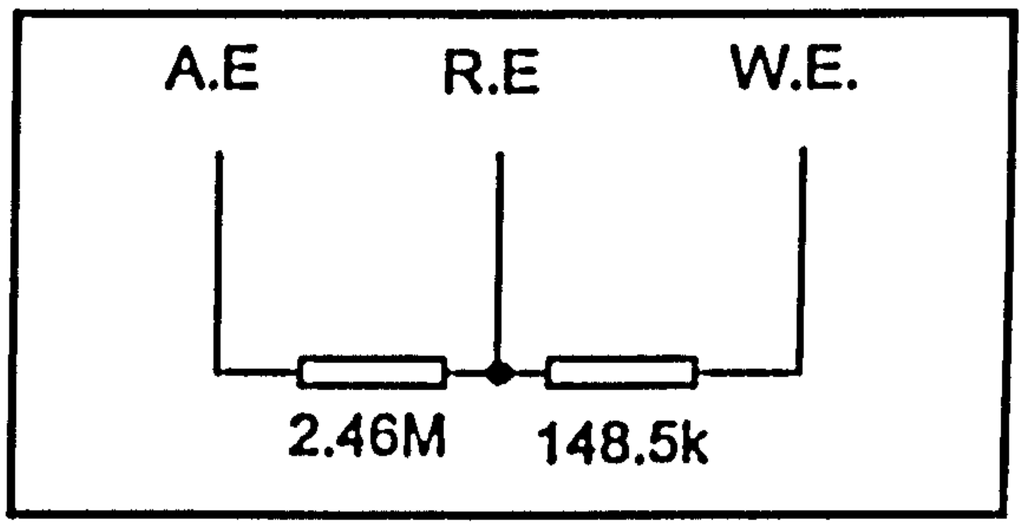



Figure 21. Dummy cell used to test current to voltage converter

Using the computer software to be described later, a linear scan was performed from -100 mV to +100 mV at a rate of 330 $mVs^{-1}$. The potential

applied and the measured current though the W.E. was recorded. This was performed at each of six current to voltage sensitivities and each of four differential amplifier gains. No current offset was used.

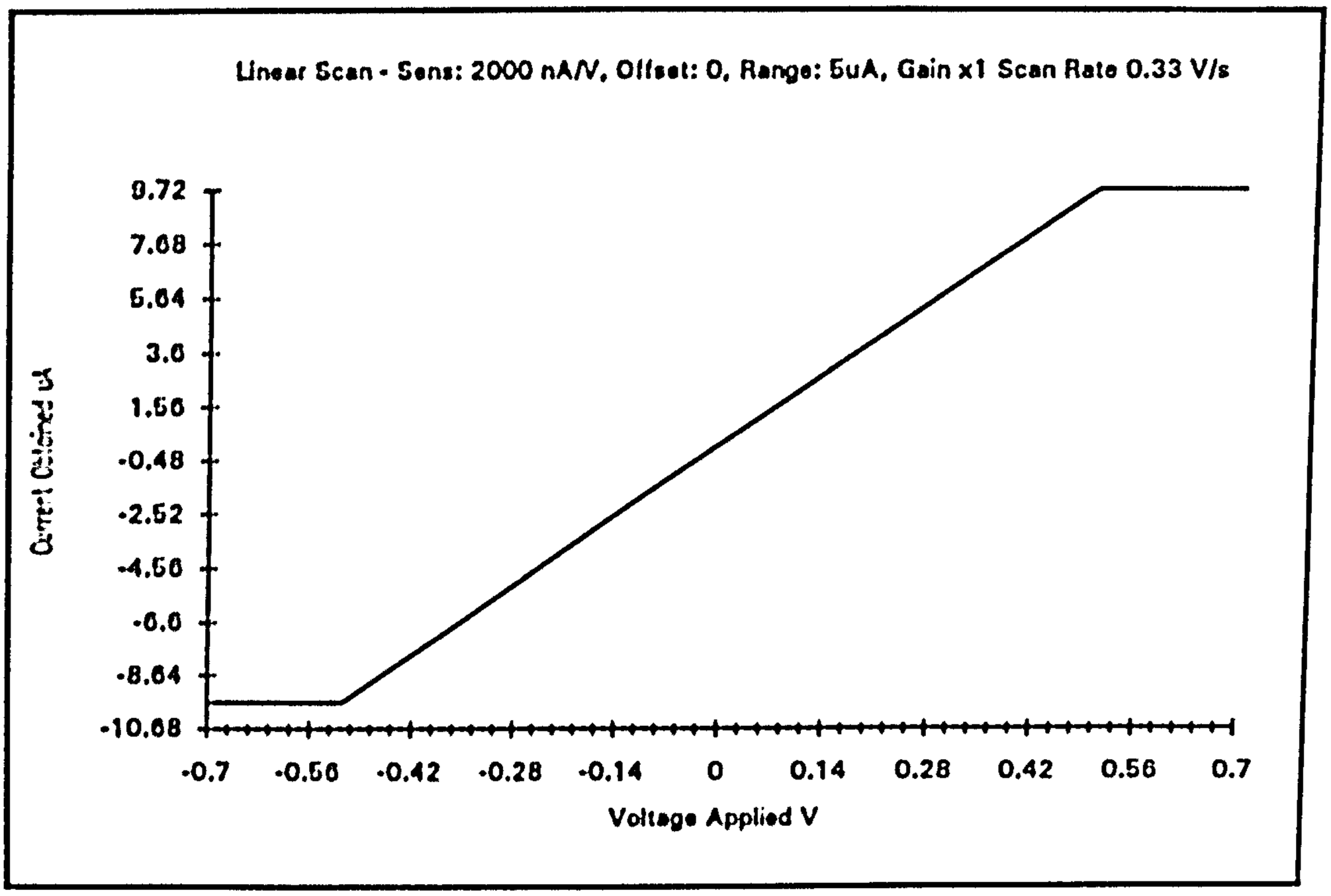



Figure 22. Results of a linear scan used during the testing of the I/E converter

Since there was a 148.5 Ω resistor between the W.E. and the R. E., the resulting plot of current against applied W.E. potential showed the expected linear relationship. With some sensitivities and gains however, saturation of the current to voltage converter was observed, resulting in an indication that current was not changing in spite of a change in PD. The results of a typical run indicating

saturation are shown in figure 22 on page 85. The proportional part of each graph was used to calculate a value for the resistance between the R.E. and the W.E. If the switching between different sensitivities and gains does not introduce an error this should be always be equal to 148.5 Ω

The results of this operation are shown in figure 23 on page 86 where the percentage error in calculating the resistance is plotted against the sensitivity for all four gains. Here the value for I has been calculated from the resulting A/D output assuming the sensitivities in table 4.

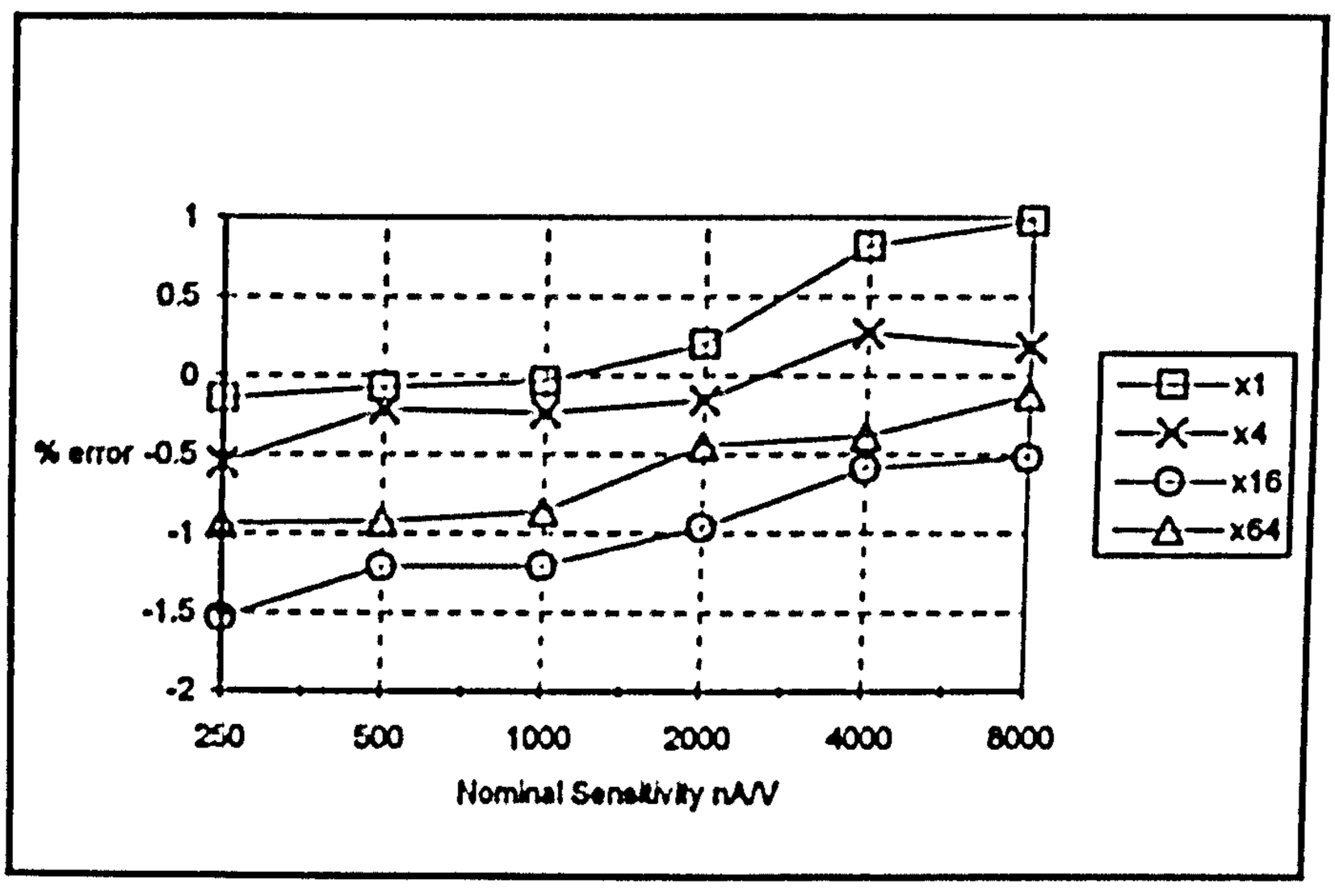



Figure 23. Graph of error in I to E converter with sensitivity at different gains

These results show that within an acceptable degree of error, changing the sensitivity and / or the gain during the lifetime of a single experimental run is possible without introducing undue distortion into the results obtained.

The values of W.E. current causing saturation at each sensitivity and gain setting were also recorded. These results are given in appendix 5.

### 4.2.9. The Current to Voltage Converter Offset Module

The current to voltage converter offset module is shown in figure 24 on page 88. The purpose of this module was to allow the 'backing off' of any large current component present in the response. It will offset the current $I_{w.e.}$ by a positive or negative amount, the value of which is determined by the settings of offset range and offset number. Offset numbers from -7 to +8 are available, each at two ranges. It should be noted that offset number does not refer to an absolute offset current value but is used to give an indication of the approximate value where settings -8 and +7 correspond to approximately -1.244 μA to +1.213 μA for the lower offset range and from -12.500 μA to +12.120 μA for the higher range. Details of each offset value that may be obtained are given in appendix 5, table 44 on page 286 and in figure 25 on page 91.

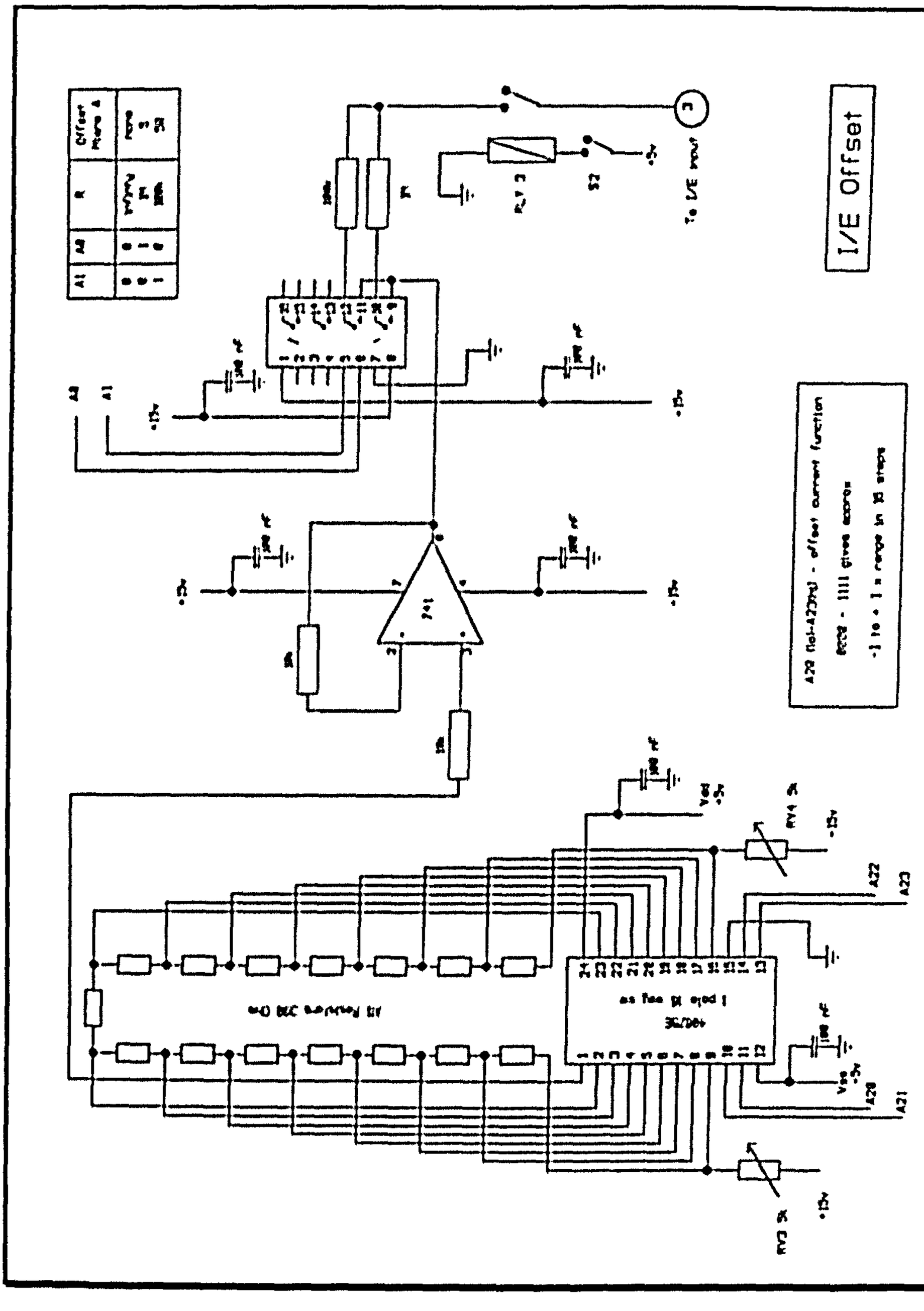



Figure 24. Current to Voltage Converter Offset Module

### 4.2.9.1. Construction

The circuit consists of a potentiometer that can tap off any one of 16 points in a nominal range of ±5 V. A 6067BE 1pole, 16 way analogue switch, is used with a resistor chain of 15 x 390 Ω resistors to provide a digitally controlled potentiometer. The ends of the resistor chain are taken to +15V and -15V respectively via 5kΩ trimmers adjusted to give approximately 10V over the resistor chain in equal steps with 0V at the centre tap. The output of this potentiometer is buffered via a 741 and fed via either 1MΩ or 100kΩ resistor to the input of the Current to Voltage converter. The value of this resistor is switched by a AD7590 four pole analogue switch and sets the range of offset to give approximately ±5μA or ±50μA respectively.

### 4.2.9.2. Operation

The range and the offset number can be controlled by the microcomputer by setting bits on the appropriate address lines. These address lines are indicated in table 46, Page 288 but further details of the switching codes and offset currents obtained are given in appendix 5 in table 42, page 284 and table 43 on page 285.

Once the settings have been determined, the potential at the appropriate tap point will appear at the common pin, pin 1, of the 4067BE 16 way switch. This potential is then buffered by the 741 and fed into pins 9 and 11 of the AD7590 analogue switch. Depending upon the signals on pins 5 and 6 this potential then appears on pin 10 or 12. This causes a current to flow through either the 100kΩ or the 1MΩ resistor, the remote end of which is held at virtual earth at the current to voltage converter. The offset current can be isolated from this Op Amp by the a switch on the front panel which controls sub miniature relay RLY 3.

#### 4.2.9.3. Performance

The output from the appropriate pin of the AD7590 offset range analogue switch (i.e. Pin 10 when using the 5 μA range and Pin 12 when using the 50 μA range) was measured with respect to ground using a high impedance digital voltmeter. The potential of Pin 2 on the I/E converter Op Amp was also measured with respect to ground. These two readings gave the potential across the appropriate resistor (100kΩ or 1MΩ). From this the actual offset current flowing for each setting could be calculated. This is shown in figure 25 on page 91 (and tabulated in appendix 5, table 44, page 286). The sensitivity setting in force is 1000 nA/V.

The curves shown here indicate that the relationship of current to offset number is not linear for the higher values of offset. This is particularly obvious when on the 50μA range. Whilst this does not constitute not a problem, the offset current being relative and not absolute, the reason for it is indicated in figure 26 on page 91, which shows the potential over to the 100kΩ offset resistor ($V_{pot}$) and the potential of the I/E Op Amp input Pin 2 ($V_{i/e}$) both plotted independently against the offset number for the 5μA range.

It can be seen that whilst the offset potentiometer resistor chain is producing a linear output ($V_{pot}$), the potential of the I/E Op Amp input Pin 2 ($V_{i/e}$) begins to rise at about offset number ±3. This is the point when saturation begins and the input to the I/E Op Amp can no longer be held at virtual ground.

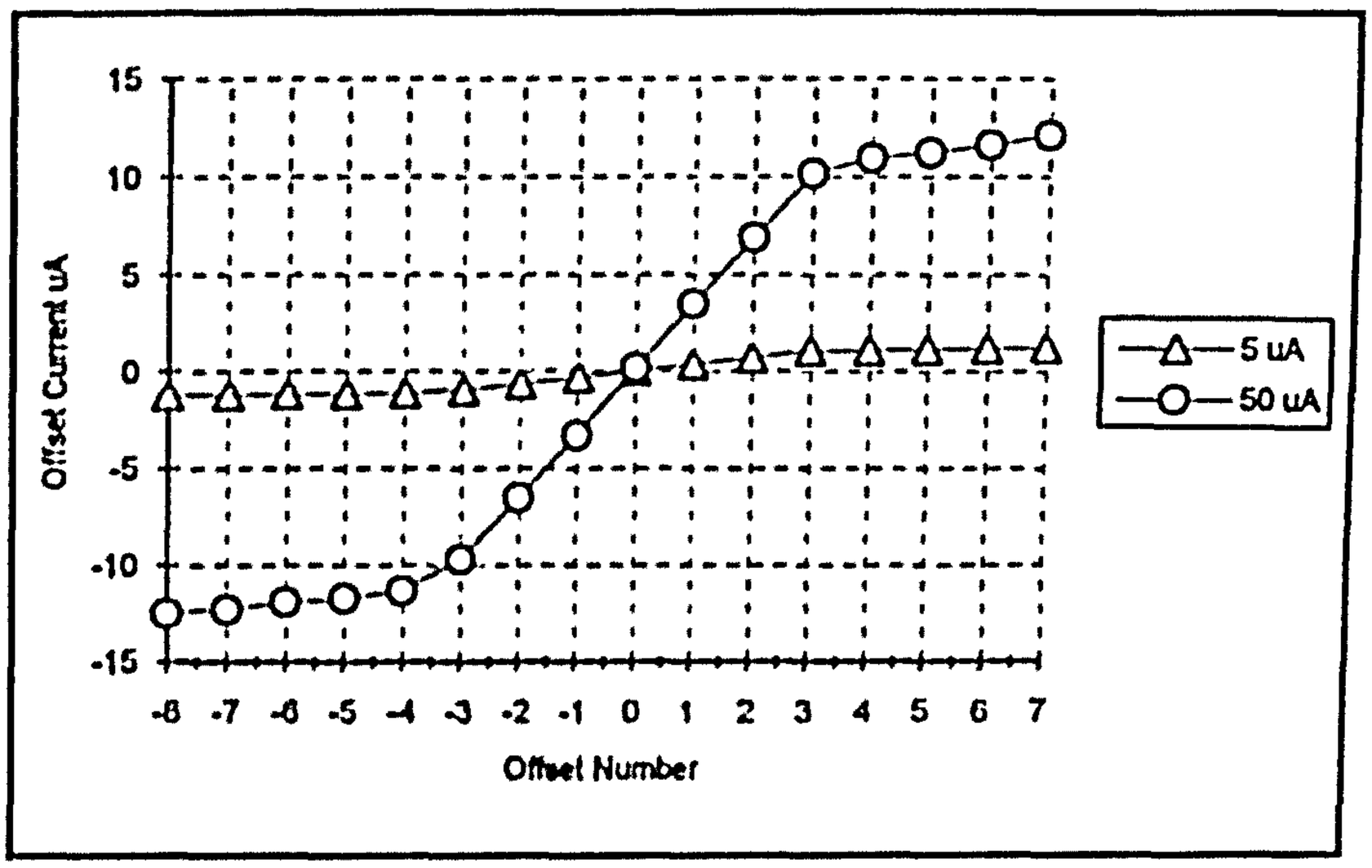



Figure 25. I/E Offset current vs. offset number

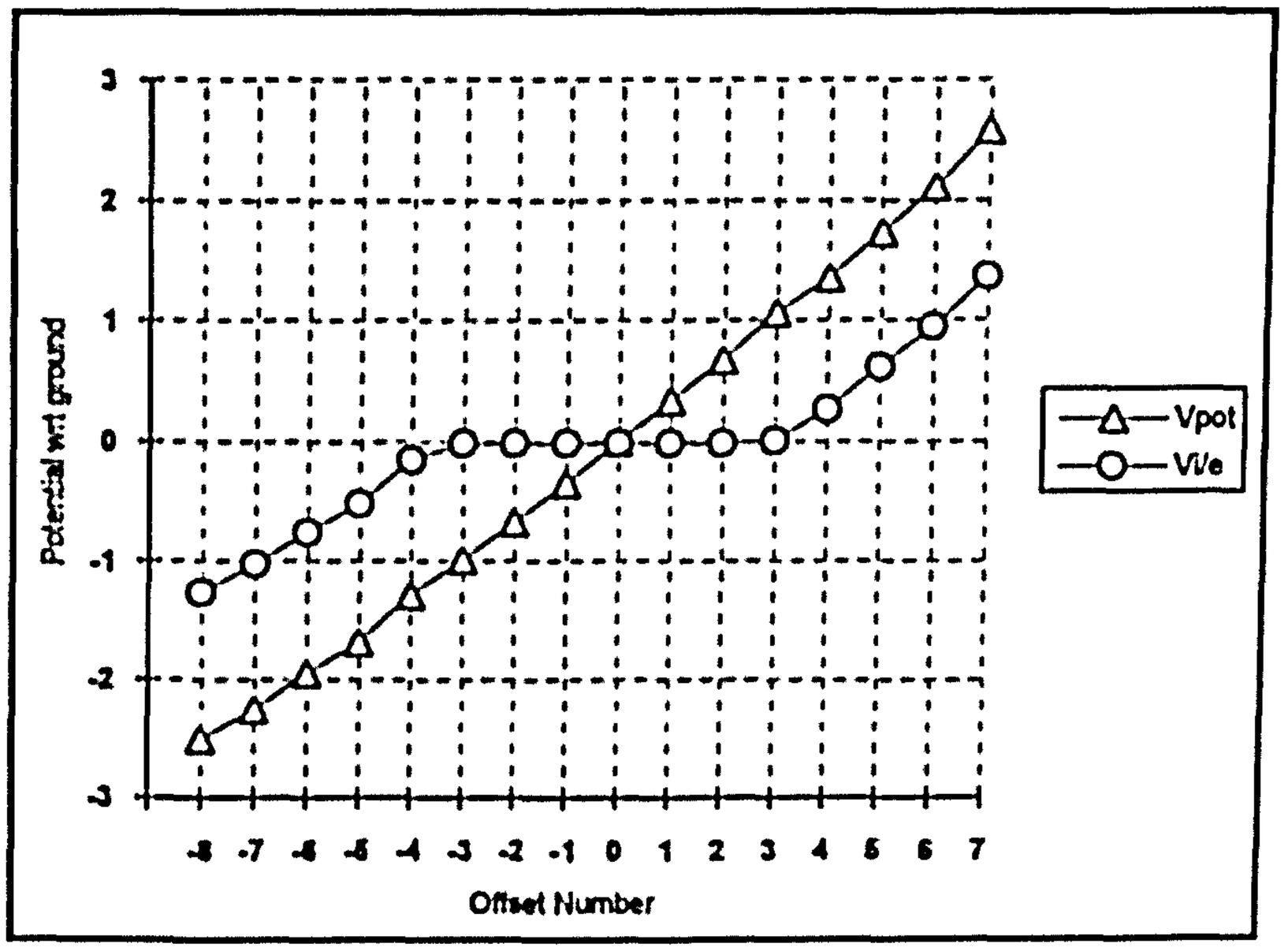



Figure 26. I/E Offset potentiometer output and I/E input potential vs. Offset Number

### 4.2.4.6. The Differential Amplifier Module

An AD625 programmable gain instrumentation amplifier[238] was used as a follow on stage to the I/E converter. This module is necessary for differential pulse analytical methods where the difference in current between two consecutive readings is required, not the absolute value of either. This module allows a voltage representation of the current at the previous reading to be stored and subtracted from the present reading. This difference may be amplified if the two reading would otherwise result in a small peak height.

### 4.2.10. Construction of the Differential Amplifier Module

The AD625 amplifier was programmed by a resistor network and a multiplexor following guidelines in the databook[238] to provide a software programmable gain of x1, x4, x16 and x64. The final circuit is shown in figure 27 on page 93.

A 7509 dual 1 to 4 multiplexor was used to switch in pairs of resistors into the sense/drive circuit of the AD625. This multiplexor is controlled by software using a binary code on address lines 8 and 9.

The two inputs to the AD625 differential amplifier were each fed from AD585 sample and hold IC's. These took as their input the output of the I/E converter at the time the hold signal was applied. Thus the voltage representation of the current held by S/H 2 was subtracted from the representation of the current held by S/H 1.

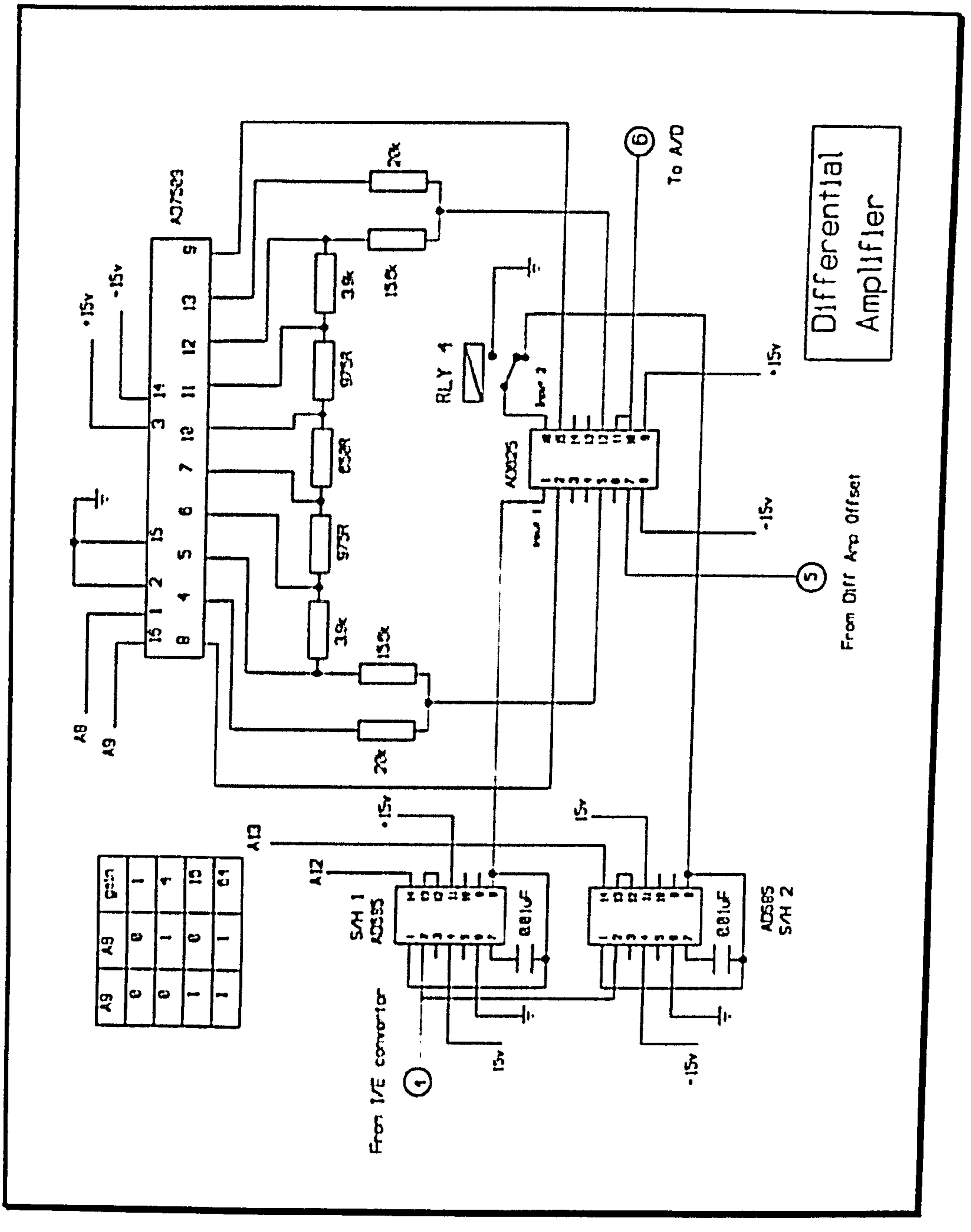



Figure 27. Differential amplifier module

Provision was made for grounding the inverting input via relay 4 under software control on address line 10. This enabled the amplifier to be used as a simple, non differential amplifier during linear scan techniques by sampling using S/H 1 only, connected to the non inverting input.

The output of this module was taken directly to the Analogue to Digital Converter.

### 4.2.10.1. Operation

A potential with respect to ground is presented to pin 2 of both sample and holds from the output of the I/E converter. This potential represents a current flowing through the electrochemical cell.

In the case of a linear scan technique the inverting input to the amplifier would be grounded. A sample signal is sent to S/H 1 on address line 12, followed a short time later by a hold signal. Address lines 8 and 9 have selected appropriate taps on the resistor chain so as to provide the correct gain and feedback resistors for the required amplified gain. The held potential is amplified by the differential amplifier which produces an output to the A/D converter module.

In the case of a differential pulse technique S/H 1 and 2 will be given a sample signal on address lines 12 and 13 respectively. Just before the pulse S/H 2 is set to hold and the pulse applied by changing the potential applied to the potentiostat. Just before the end of the pulse S/H 1 is set to hold. The amplifier then amplifies the difference between the values of the two held voltages and both sample and holds are set to sample again.

### 4.2.10.2. Performance

The amplifier was tested using the dummy cell circuit shown in fig 17 on page 78. A potential, measured with a digital voltmeter, of 0.039 V was applied at the non inverting input of the amplifier with the inverting input grounded. The output of the amplifier was then measured at each of the gains using both a digital voltmeter and the software written by the author. The results are shown in table 5 on page 95. These indicate that the differential amplifier circuit is producing the gains required.

| Set Gain | Output PD using Digital Voltmeter V | Output PD using Software V | Calculated Gain Using Voltmeter | Calculated Gain using Software |
|---|---|---|---|---|
| x1 | 0.039 | 0.038 | 1.00 | 0.97 |
| x4 | 0.158 | 0.160 | 4.05 | 3.95 |
| x16 | 0.657 | 0.665 | 16.85 | 16.00 |
| x64 | 2.55 | 2.582 | 65.38 | 62.40 |

Table 5. Gain calibration of differential amplifier

### 4.2.11. The Differential Amplifier Offset Module

The differential amplifier offset module is shown in fig 28 on page 96. This simple circuit uses a 741 to create a buffered potentiometer and can be used to offset the amplifier output by a potential in either the positive or the negative direction to the offset pin 7 on the differential amplifier. Relay (RLY 5) allows for zero offset by grounding pin 7.

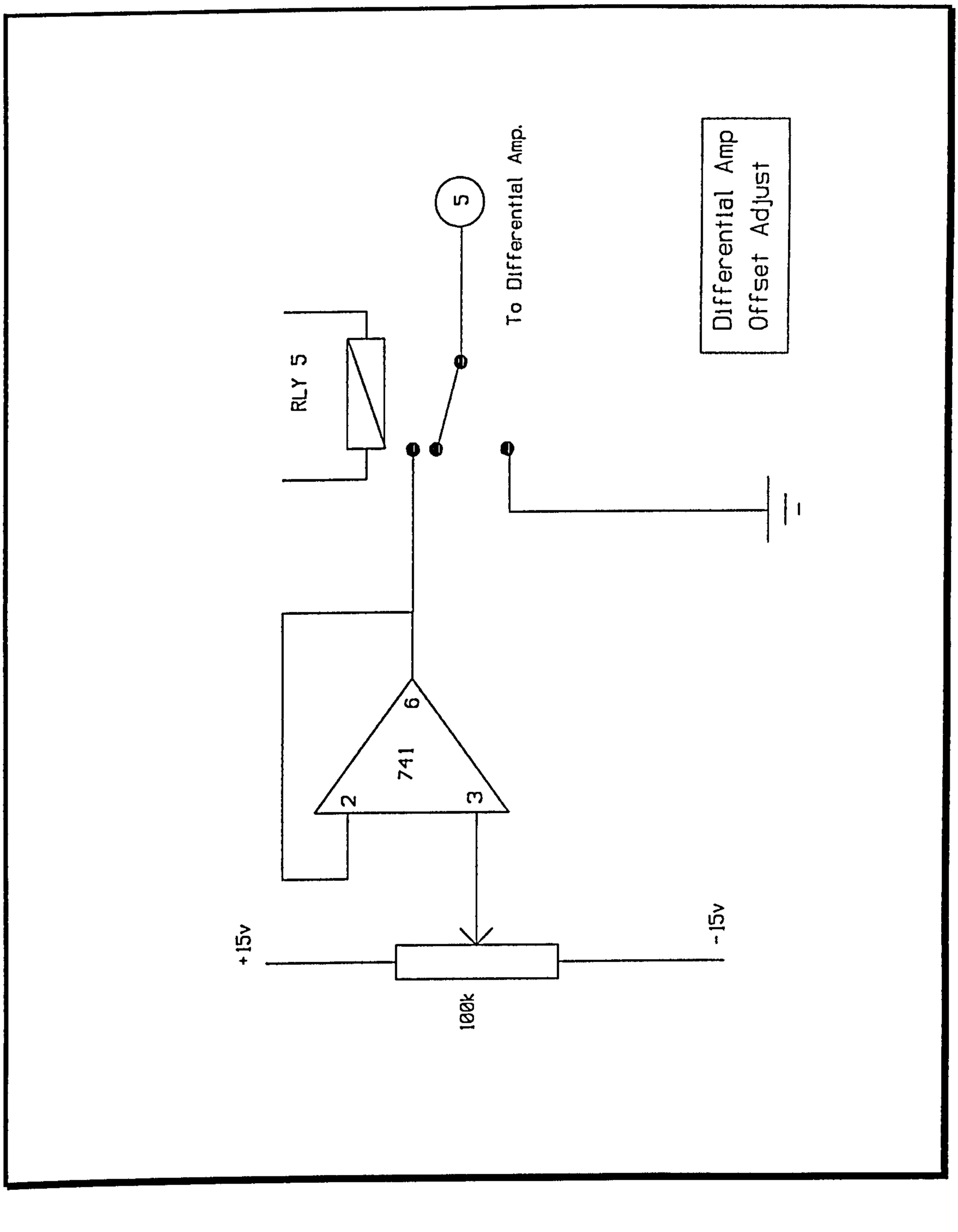



Figure 28. Differential amplifier offset module

Since it was not anticipated that this operation would be required except whilst setting up the apparatus, this control was brought out to the front panel and not supplied as a computer controlled function. A simple modification could however provide this function under computer control.

## 4.2.12. The A / D Converter Module

### 4.2.12.1. Construction

The analogue to digital converter module is shown in fig 29 on page 98. This uses a AD574A, single chip 12 bit D/A to convert the output of the differential amplifier into a digital value for transmission to the computer.

Following the guidelines in the databook[239], the device was connected in bipolar mode by connecting two 100Ω trimmers between pin 12 (bipolar offset) and pin 8 (ref out) and also between pins 10 (ref in) and 8. The input span was selected as 10v by connecting the output of the differential amplifier to pin 13 (range select). The output was selected as full 12 bit by connecting pin 2 (12/not 8 bit) to +5v and pin 4 ($A_0$) to 0v.

The above connections gave the chip the capability to convert analogue potentials in the range ±5.0 V with a resolution of 1.22 mV in full 12 bit mode.

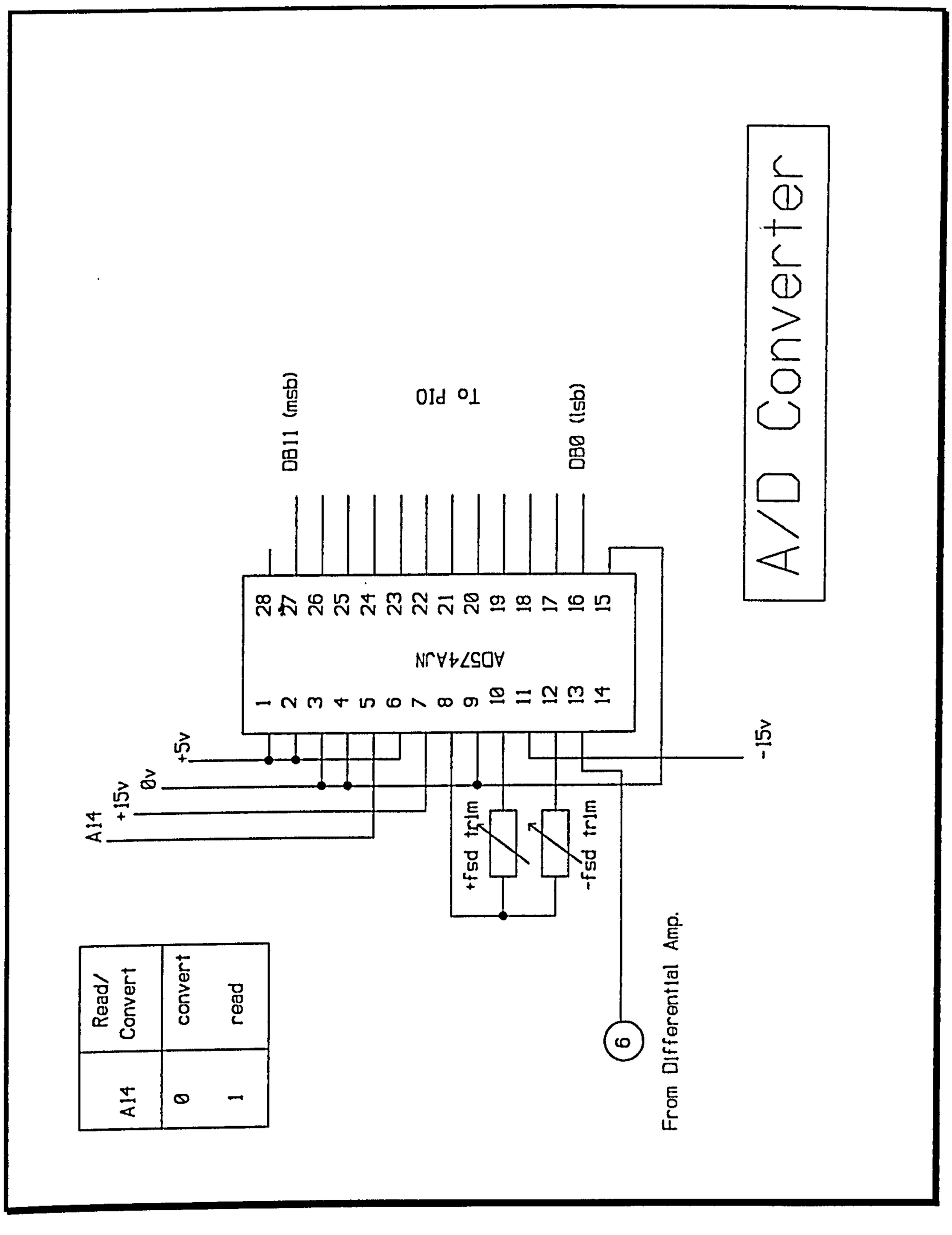



Figure 29. A / D converter module

#### 4.2.12.2. Operation

When it is required to perform a conversion the software takes pin 5 (read / convert) low. This starts the conversion during which any further changes on the input will be ignored. At the completion of the conversion the pin 28 (status) goes high. This pin can be polled or used to provide an interrupt. However since the maximum conversion time is guaranteed to be only 35μs, in this work a simple software delay loop is used to ensure that at least 35μs passes between starting the conversion and reading the data from the output.

Address line A14 is used for the read / convert signal. The software therefore sends a read signal (logic 1)on address line A14 whilst the exciting potential is sent to the potentiostat. When the measured current is required to be read by the software a convert signal (logic 0) is placed on A14 and a short delay loop entered while the conversion is carried out. At the end of this time the data is read from the port and the read signal once again placed on A14.

#### 4.2.12.3. Performance of the A/D Converter

The two 100Ω trimmers were used to calibrate the chip using the test box described previously. Using this box a potential of -4.998v (½ LSB above negative full scale) was applied and the trimmer -fsd trimmed in order to obtain the first transition from 0000_0000_0000 to 0000_0000_0001 on the output. This was detected via the 12 leds on the test box. Next a potential of +4.996 (1½ LSB below positive full scale) was applied and trimmer +fsd trimmed in order to obtain the last transition from 1111_1111_1110 to 1111_1111_1111. After this calibration the D/A was tested at several potentials between these ranges and found to convert correctly within 1.22 mV. A calibration graph is shown in fig 30 on page 100.

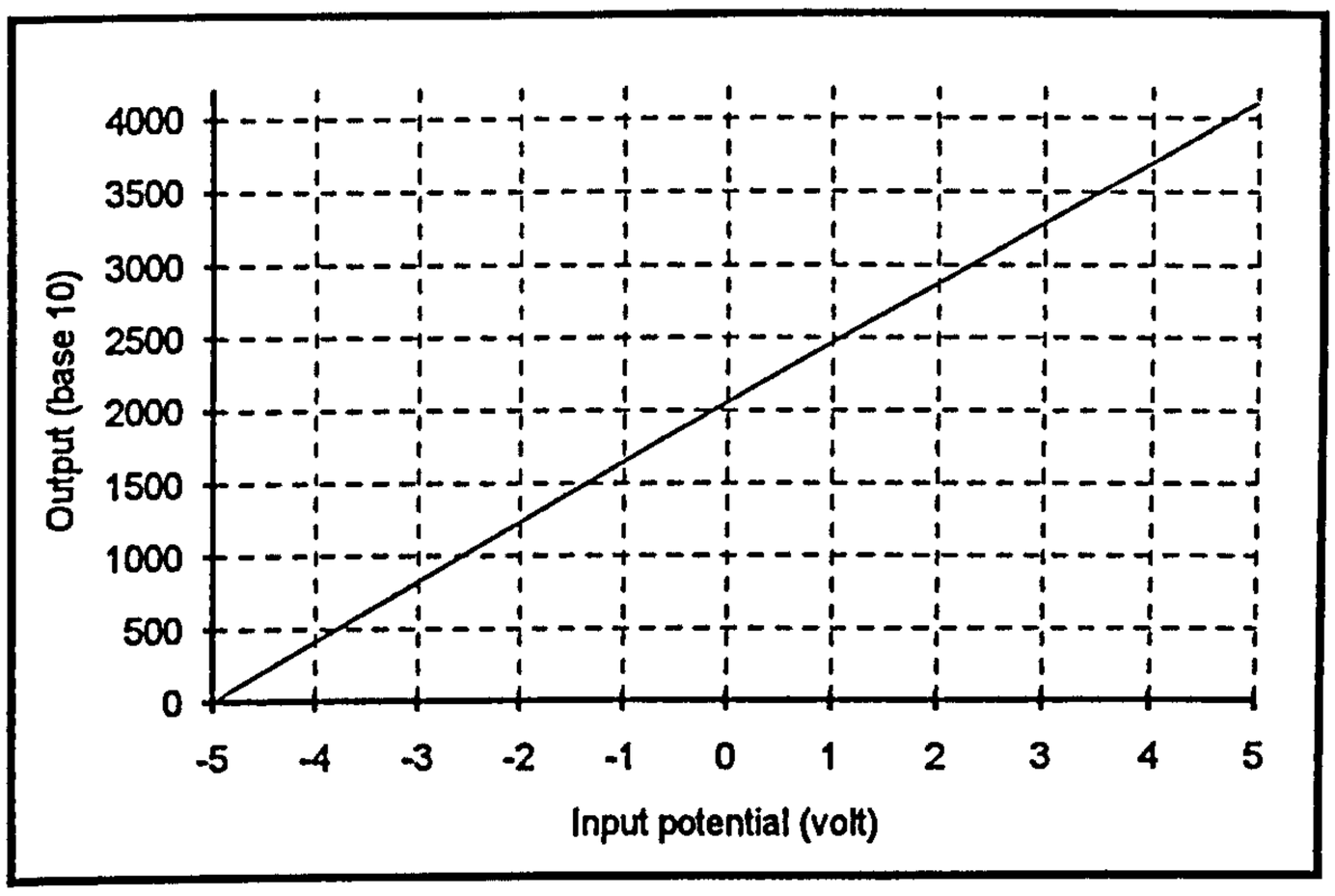



Figure 30. Calibration graph for A/D converter

## 4.3. Construction of the electroanalytical software

### 4.3.1. Overview

All the electroanalytical software that controls the potentiostat, applies the voltammetric wave form, reads the resultant current and processes the received data has been written in 80386 assembly language and Turbo Pascal® version 7. It is organised as a series of software layers. Each of these layers provides functions which, although complex within the layer, are regarded as being simple atomic functions by the layer above. The lowest level, called here the Hardware Interface Layer, controls the operation of the 48 way programmable Input / Output Card and thus the operation of the potentiostat and associated analytical hardware. The next layer, called the Logical Analysis Layer, is an intermediate 'system' level layer and executes the algorithms necessary to carry out a particular electrochemical analysis or Thin Film Mercury Electrode (TFME) production. The highest, the User Interface Layer presents to the user a mouse driven interface providing those operations that would be familiar to a electroanalytical chemist.

A similar set of layers performs the graphical functions responsible for plotting and manipulating a voltammogram. Once again these consists of a Low Level Graphics layer providing primitive graphics screen functions, a logical layer which provides as atomic those higher level functions required to graph data on an XY plot and a User Interface Layer which is accessed when the user wishes to display or manipulate a voltammogram.

The logical relationships between the layers is shown in Table 6, Page 102.

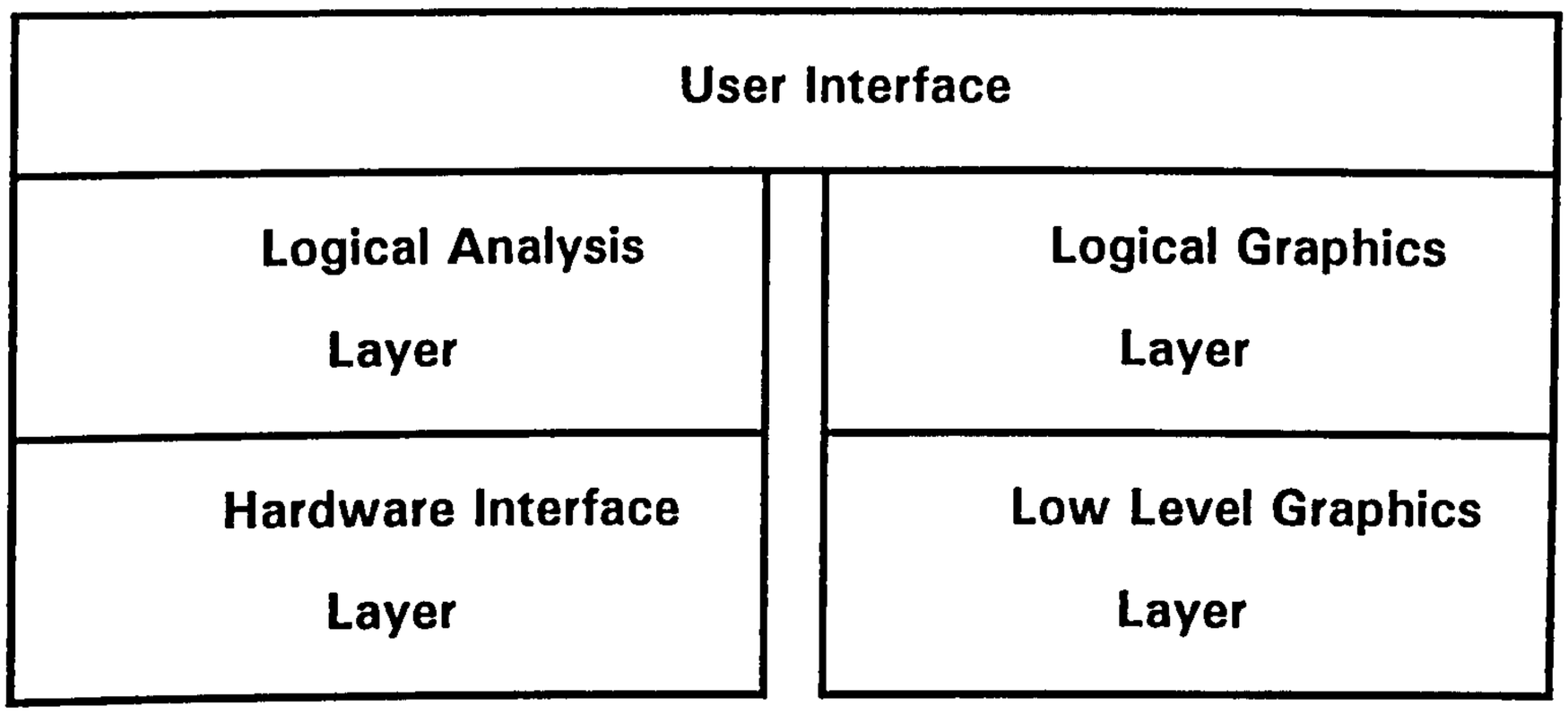



Table 6. Layered structure of the electroanalytical software

Each of the layers described above was designed, written and compiled as a number of separate units which provide various abstract views of data via member functions accessed through the interface section.

A useful feature of Turbo Pascal is the ability to use the keyword ASM. Following this keyword the programmer can program directly in 80x86 assembly language and yet remain within the Pascal source code file. During the compilation process the compiler switches from compile mode to assemble mode as required. Procedures and functions may be written that pass parameters and return values in a Pascal style yet internally contain nothing but assembly language. This feature was used to code the lowest level where I/O port access and bit manipulations were required.

Subsequent sections will describe the software layers in a top down fashion from the User Interface layer down to the lower levels.

### 4.3.2. User Interface Layer

The features in this layer are best described by figure 31 which illustrates the structure of the drop down menus.

Option choices are indicated using 'radio button' dialogue boxes and values are entered using 'text box' dialogue boxes. A sample of each type of dialogue box is given here, figure 32 showing radio buttons and figure 33 showing a text box. Screen dumps of all the dialogues available are given in appendix 4.

The Run Analysis option on the main menu starts the analytical run by applying the specified plate potential and then rotating the TFME for the specified time. Following this the electrode is rested without rotation and then the scan itself is commenced with data being recorded. Suitable messages are displayed to the user to confirm which phase of the analysis is currently in progress.

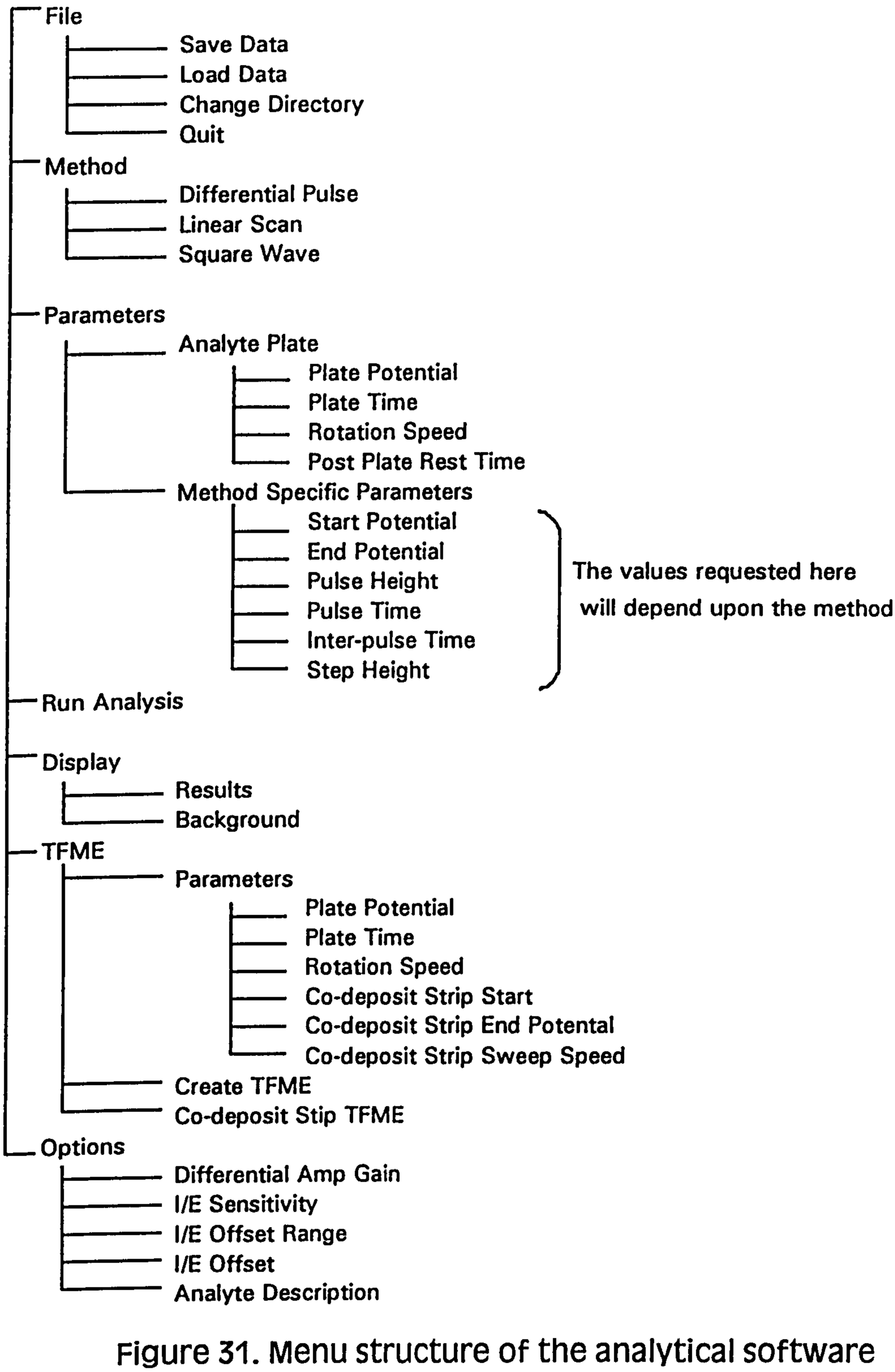



Figure 31. Menu structure of the analytical software

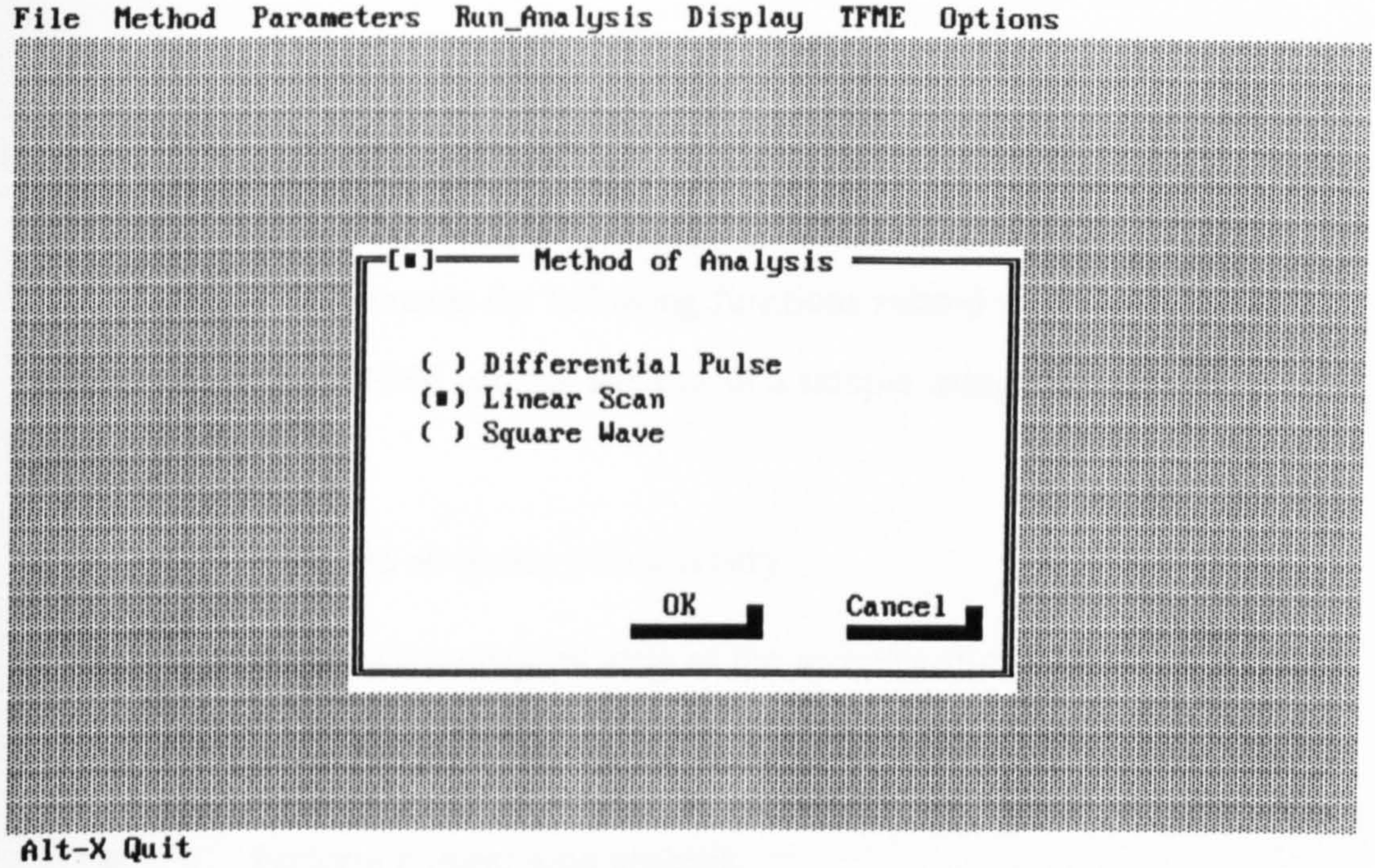



Figure 32. Typical 'radio button' dialogue box from the analytical software

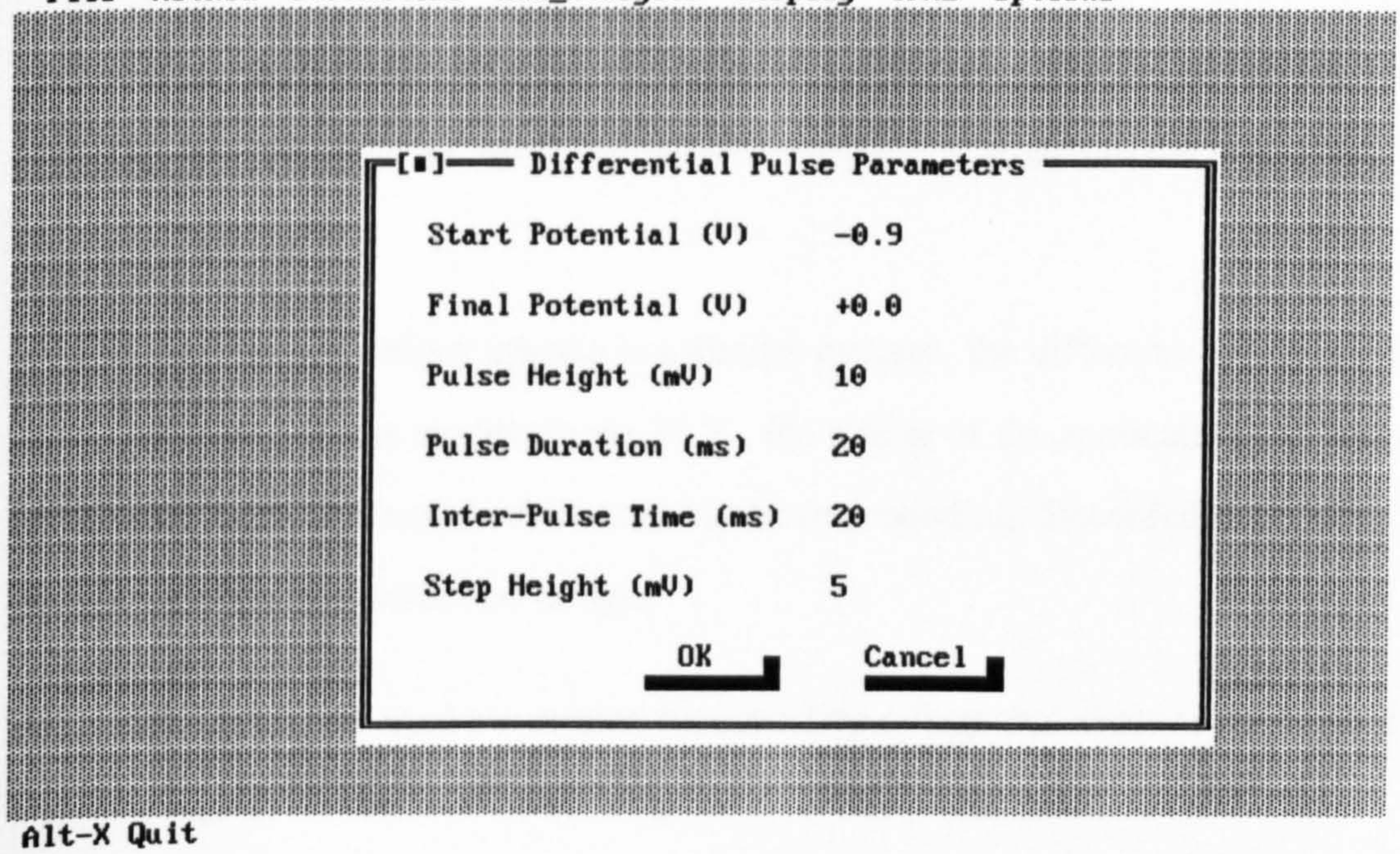



Figure 33. Typical 'text box' dialogue box from the analytical software

### 4.3.3. Logical Analysis Layer

This layer implements the following functions related to the creation and conditioning of the TFME and the analysis of a sample using a variety of waveforms.

1. Plate the electrode with mercury
2. Perform a co-deposit strip of the mercury film
3. Plate the mercury with the analyte
4. Perform a linear scan analysis
5. Perform a differential pulse analysis
6. Perform a square wave analysis

These functions are specified in appendix 6, table 45, page 287. They are called by the user interface layer and themselves call procedures in the hardware interface layer.

All of these functions operate in a similar manner, the difference being mainly in the waveform that is applied to the W.E., the timing of the application and setting of switches such as electrode rotate and electrode grounding. Recorded data are saved to dynamically allocated arrays.

The waveforms used are created by executing a loop that starts at the desired W.E. start potential and continues while the final W.E. potential has not been reached. By using an appropriate combination of changes in W.E. potential, delays, current recording times and sample and hold states the correct waveform can be

generated. As an example, a generalised form of the algorithm used for differential pulse analysis, omitting details, is given in Figure 34.

```
V_w.e. := start pd
While V_w.e. <= End pd do
        begin
Set Sample/Hold No. 2 to Sample          {start to sample i before the pulse}
apply V_w.e. to W.E.                     {send ramp potential to W.E.}
delay                                    {for time between pulses}
V_w.e. := V_w.e. + Pulse Height;         {calculate potential of pulse leading edge}
Store V_w.e. for bottom of pulse         {potential (x axis) data for voltammogram}
Set Sample/Hold No.2 to Hold             {capture i before the pulse}
apply V_w.e to W.E.                      {apply the pulse}
Set Sample/Hold  No. 1 to Sample         {start to sample i at the end of pulse}
delay                                    {for duration of the pulse}
Set Sample/Hold No.1 to Hold             {capture i at the end of the pulse}
Read & store A/D output                  {representation of Δ_i (y axis) after pulse}
V_w.e. := V_w.e. - pulse drop            {calculate height of pulse trailing edge}
end;
```

Figure 34. General algorithm for differential pulse analysis

For all applications the loop is started five iterations before the desired starting potential, in practice a few millivolts, in order to settle the instrument before data recording is started. This was found to be necessary to avoid abrupt changes in the recorded current at the start of the scan. The loop is stopped at the point when the potential, peak pulse potential in the case of differential pulse and square wave, has reached the desired value.

A particular feature of this software is the method by which delays are created. These need to have a resolution 1 ms or less to generate accurately the analytical waveforms but will need to be much longer when creating the TFME.

The built in Turbo Pascal procedure Delay(n), intended to pause program execution for n ms was found to be too unreliable for this application. Experiments were performed using clock interrupts after first speeding up the basic clock interrupt rate. Whilst this performed as required and permitted the coding of quite accurate delays from $10^{-3}$ to $10^{-5}$ s, the subsequent impact on the system's record of the real time and date was undesirable.

The method finally used relied upon a calibrated software loop written in assembly language. Software timing loops are generally to be avoided because of the unpredictability of interrupts. In this work however a calibration phase is run as soon as an analysis procedure is required. This waits for the low order byte of the system clock to change and then enters a loop that counts how many iterations can be performed before the byte changes again due to a real time clock interrupt 55 ms later. The number of loops performed in 1ms is calculated. Delays can then be achieved by performing the appropriate number of iterations of a second loop. The two loops are coded so that in all other respects they are identical to each other, including the use of dummy register increments simulating the counting of loops during the timing loop.

This approach was tested by;

1. Using the software to generate a square wave at various frequencies in the KHz region and examining the potential of the Working Electrode using an oscilloscope.

2. Creating long delays of 30 minutes to 1 hour and timing them with a stopwatch.

In both cases the outcome was satisfactory in spite of the interrupts that must have been occurring during the delays. The only concession made during these tests was that the disc cache was disabled during the test to prevent timings being altered by whether or not a cache flush and disc write occurred during the timing interval. In practice this is quite acceptable as there is no need for a disc cache during the analysis. Data can be stored in memory during the analytical run and dumped to disc afterwards. All subsequent work was done with the disc cache disabled.

### 4.3.4. Hardware Interface Layer

This layer of software was mainly written in assembly language because of the need to perform manipulation of data at the bit level.

The operation of this layer falls into two parts. Firstly, that concerned with the functional control of the potentiostat and secondly that concerned with the output of digitised W.E. potentials and the input of digitised W.E. currents.

A common feature of software in this layer is that there are three byte sized 'state variables' which define the current state of each one of the three I/O ports, A', B' or C' on the 48 way PIO card. The bit patterns in these state variables is reflected in the current state of operation of the potentiostat and current measuring hardware.

All requests to change the state of the hardware are first translated into setting the appropriate bit(s) in a data byte which has the initial value of 00(hex). This data byte is then ORed with the appropriate state variable to switch on that particular function, or NOTed and then ANDed with the state variable to switch off that function. The appropriate state variable is then stored in its allocated port and hence sent to the hardware.

In the case of multiple bit settings, for example setting the I/E offset which uses four bit binary encoding, the state variable is first ANDed with a mask to reset all the appropriate bits so that all bits in the state variable take on the correct value when the OR function is applied

The hardware control functions themselves fall into four groups; PIO set up, operational parameters set up, relay switching and measurement functions. These are listed in tables 7 and 8 and further defined in appendix 6, tables 46 to 49 on pages 288 to 291. All operate in the manner described above.

| PIO Board configuration Procedure |
|---|
| **Set up I/O Ports** *Sets up the operation mode of the ports and configures ports for input or output* |
| **Potentiostat Configuration Procedures** |
| **Set I to E Offset Range** *Sets the I to E converter offset range to either 5μA or 50μA* |
| **Set I to E Offset** *Sets the I to E converter offset to one of sixteen values spanning zero* |
| **Set Differential Amplifier Gain** *Sets the differential amplifier gain to one of four gains* |
| **Set I to E Sensitivity** *Sets the I to E converter sensitivity to one of six sensitivities* |
| **Ancillary switches** |
| **Turn On Motor** *Turns on the working electrode rotating motor* |
| **Turn Off Motor** *Turns on the working electrode rotating motor* |
| **Turn On Gas Purge** *Opens the Nitrogen gas purge valve* |
| **Turn Off Gas Purge** *Closes the Nitrogen gas purge valve* |
| **Connect Differential Amplifier Input 2** *Connects the inverting input of the differential amplifier to the working electrode via Sample/Hold 2* |
| **Ground Differential Amplifier Input 2** *Grounds the inverting input of the differential amplifier (for use in linear scan)* |
| **Connect Working Electrode** *Connects the working electrode to the I to E converter* |
| **Ground Working Electrode** *Grounds the working electrode (normally done whilst analyte plating)* |

Table 7. General purpose functions available in the hardware interface software layer

| Working Electrode Potential and Current Reading procedures |
|---|
| **Set Sample and Hold** *Sets sample/hold on either input one or two of the differential amplifier to either sample or hold* |
| **Send Voltage** *Converts a bipolar PD in the range ±2.5 V to the corresponding 12 bit D/A code and sends it out to the potentiostat* |
| **Voltage In** *Reads a 12 bit code from the A/D in the potentiostat and returns the value converted into a bipolar PD in the range ±2.5 V* |

Table 8. Measurement functions available in the hardware interface software layer

Six bits are used to encode the I to E sensitivity using a 1 out of n encoding. This apparently inefficient encoding system was used in order to simplify the hardware design by using six CMOS switches, in two packages, with each switch switching in one range. Again for simplicity, two bits are used to encode the offset range using 1 out of n. However four bits were used to binary encode the 16 values of offset, this being implemented using a single chip 1 pole, 16 way CMOS switch. The differential amplifier gain also uses binary encoding using two bits to program a two pole, four way CMOS switch which switches in pairs of resistors controlling the amplifier gain.

In order to apply a particular potential to the working electrode the procedure Send Voltage is called passing the value of the required potential in mV. A lookup table is then used to convert this potential into a 16 bit word which, when applied to the input of the D/A will cause it to output the correct potential.

Reading of an input value, which represented a current, is done by the Voltage In function which first sends a convert signal to the A/D converter and then enters a short software delay loop to ensure that 35ns elapses whist the conversion is carried out. After this time the A/D is put back into read mode and the 12 bit output built up in the BX register and returned by the function into a Pascal variable.

### 4.3.5. Logical Graphics Layer

This layer of software provides an intermediate layer between the user interface layer and the low level graphics layer. The functions available are listed in table 9 and specified in appendix 6, table 50, page 292

| |
|---|
| **Display Data** *Plots the analytical response data in as an XY graph on automatically scaled axes.* |
| **Set Title One** *Creates a single string suitable for the first title line of an XY plot and initialises conversion factors used by DisplayData* |
| **Set Title Two** *Initialises the second line of the XY plot title.* |
| **Set Title Three** *Initialises the third line of the XY plot title.* |

Table 9. List of functions available in the logical graphics layer

After carrying out an analysis the logical analysis layer presents the user interface layer with three pieces of data. These are the number of data points taken and two pointers to dynamically allocated arrays of integer containing the potentials, in mV, applied to the W.E. at each step and the corresponding potential received from the A/D converter and representing the (possibly differential) measured current in μA.

These data need preprocessing before they can be plotted in order to convert the potential from millivolt to volt and also to convert the received potential into a current using the I/E converter sensitivity and differential gain settings in force at the time of the analysis. The procedure Set Title One in this layer converts the enumerated types of I/E sensitivity, differential amplifier gain, I/E offset range, I/E offset and analytical method into suitable strings for the first title line of the graph.

It also uses the first two of these to calculate conversion factors which are subsequently used by Display Data to convert the received potential into a current.

The Display Data procedure accepts the two variables pointing to the analytical data together with the number of data points. It saves the three title strings, two generated from the operational parameters and the third, if any, from the analytical description that the user entered. Then, using the conversion factors calculated by Set Title One it converts the potential and received current data into units of volt and μA respectively and appends this processed data to the file. Finally it calls the low level graphics layer to plot the data in that file on auto scaled axes.

### 4.3.6. Low Level Graphics Level

The Low Level graph plotting module has been written in such a way that it can readily be used by other programs in this work. For example it has been used to display the sum squared prediction error of a neural network plotted against the number of training epochs in real time, whilst a neural network trains. This module provides access to several procedures and functions, listed in table 10 and specified in appendix 6, table 51, page 294.

Once the graph has been plotted several manipulative functions are available using one key press, see Table 11, Page 116. As a reminder to the user a list of these can be obtained by pressing '?' for help at any time a plot is displayed. This will show the available commands in a box overlaying the plot. The user may enter a command letter directly from this screen at which point the help screen is removed and the command executed.

The action of these commands will be obvious from the description but by way of illustration the operation of the zoom command will be detailed here.

| Graphics Functions |
|---|
| **Init Graphics** *Starts the graphics system in the best available mode* |
| **Shut Down Graphics** *Shuts down the graphics system and returns to the text screen.* |
| **Graphics Intiialised** *Determines if the graphics system is already started* |
| **Plot Data** *Plots the data held in the file created by Display Data* |
| **Draw Axis** *Plots the axis and legends and prints a three line graph title* |
| **Do Tick Marks On Y Axis** *Draws tick marks and prints the scale. Also calculates the plot to real world Y scaling value* |
| **Do Tick Marks On X Axis** *Draws tick marks and prints the scale. Also calculates the plot to real world X scaling value* |
| **Draw Curve Value** *Plots a coloured line to real world co-ordinates X, Y from last real world co-ordinates plotted or plots a point at X,Y if this is the first call to this procedure.* |
| **PlotValue** *Plots a pixel at real world co-ordinates X, Y.* |
| **Convert Plot XY To Value XY** *Converts a point given by the screen pixel co-ordinates to real world co-ordinates.* |
| **Convert Value XY To Plot XY** *Converts a point given by real world co-ordinates to screen pixel co-ordinates.* |
| **Find Max File Values** *Finds the maximum and minimum values of X and Y in the plot file.* |
| **Screen Dump** *Prints the screen on the printer* |
| **Make Display Colour** *Restores default graph colouring* |
| **Make Display Black And White** *Changes display to black on white to facilitate the transfer of the graph into another application via Windows Clipboard®* |

Table 10. List of the low level graphic functions provided

| Key | Operation |
|---|---|
| z | Zoom into a rectangular area of the plot defined by crosshairs using the mouse |
| o | Re-plot original curve |
| p | Dump plot to printer |
| s | Smooth the plot using a BSpline algorithm |
| a | Smooth the plot using a 5 point moving average |
| f | Smooth the plot using Fourier smoothing |
| S | Smooth the plot using a 13 point Savitzky-Golay polynomial |
| D | Plot the 2nd differential using a 13 point Savitzky-Golay polynomial |
| d | Plot a 1st differential using adjacent points |
| b | Fit a quadratic baseline using two portions of the curve defined using the mouse |
| h | Fit quadratics to the baseline and to the peak, plot and display the peak height |
| x | Display cross hairs and the corresponding potential - current coordinates |
| n | Normalise plot to span a height of 1.0 |
| m | Make display black on white for copy to Windows® Clipboard |
| c | Make display colour |
| Q | Exit graphical display |
| ? | Display command help screen |

Table 11. Commands available when graphical display of voltammetric plot is displayed

Pressing the 'z' key produces cross hairs on the screen which can be manipulated using the mouse. The potential of the W.E. and resultant current at the position of the cross hairs is displayed in real time, expressed in the same units as the axes. Pressing the left mouse button defines the start of a rectangular portion of

the display. The area so defined is outlined by a red rectangle, the bottom right hand corner of which moves in response to mouse movements, thus changing the area enclosed by the rectangle. A second press of the left mouse button produces a new display containing the area previously defined by the rectangle, enlarged to fill the available screen area. The graph axis are re-scaled as appropriate. This operation may be repeated indefinitely to any depth of zoom.

A useful feature of the program is that all commands operate on the data used to produce the plot currently being displayed. This data may be a transformed version of the original data due to previous user manipulations, for example via a smoothing or zoom operation.

Because of this it is possible to perform compound operations such as; plotting the n'th order differential of the plot by pressing the 'd' key 'n' times or smoothing the curve by a five point moving average to an arbitrary level of smoothing by pressing the key 'a' as many times as required.

The command to calculate the peak height is central to this thesis and is described in detail in the section of data pre processing on page154.

# 5. Methods of Investigation

## 5.1. Overview

It was initially proposed to investigate the application of neural networks to analytical chemistry by stripping voltammetry in two areas; (a) Calibration Mode, the application of neural networks after a chemical analysis as a predictive model from which the concentration of analytes can be predicted and (b) Control Mode, the use of a neural network in a control configuration to set optimum analytical parameters for the control software prior to the analysis of the sample. These were each intended to be studied separately.

A preliminary study of the behaviour of neural networks in this field was made using both computer simulated and real voltammetric data in order to ascertain the extent to which neural networks could map voltammograms to peak height. The synthetic data was generated by combining many of the features seen in a typical voltammogram. Details of this preliminary work, which included synthetic voltammograms of multiple species, may be found in Chapter 9 on page 210 and appendix 3 on page 253. The work showed clearly that the backpropagation network could be applied with some success to synthetic data.

Following this, a more detailed study using data obtained by analysing locally obtained curry powder for the presence of lead at trace levels was used in order to investigate the robustness of neural networks in the presence of real analytical variations. This work, described here in sec 5.2, page 121, extended the initial

study and showed that a neural network could indeed be used to model the relationship between voltammetric data and a corresponding analyte concentration. However, inherent problems in the analyte preparation and storage methods used by technical staff were detected. Statistically valid conclusions could not therefore be made from this curry powder data.

Finally, voltammetric data from a series of standards containing lead in the parts per billion range were produced by this author. Using this data, from which main conclusions of this thesis are drawn, calibrations of the instrument and predictions of unknown concentrations were performed using a variety of techniques, both traditional and using neural networks. The precision and accuracy of the predictions obtained in each case were compared. A description of the methodology used with this data begins on page 145.

In this latter set of experiments, three classes of data were investigated. The first of these contained good voltammograms with peaks whose height could easily be measured. The second group contained also those voltammograms that, due to the high lead concentration and the particular instrumental parameters used, saturated the instrument and produced flat topped peaks. The third group contained just the saturating voltammograms. The results of processing within these classes are given in the sections starting pages 163, 185 and 191 respectively.

Using the data in the first class it can be shown that using a neural network to post process voltammograms is superior to the traditional methods tested. The results obtained using the second and third classes are particularly significant as the data from which they originate cannot be processed using conventional, peak height means. Processing such data by neural network thus represents a novel method of producing quantitative analytical data from a voltammetric scan that might otherwise have to be repeated using more appropriate parameters.

This renders the control mode of operation, which was intended to adjust the parameters to obtain good peaks, less important than had been first thought. Work on the control mode is therefore postponed to a further period of research and not reported here.

For each of the investigations reported here, sections are included on the analyte preparation method, the production of the thin film mercury electrode, the data preparation method and the calibration / prediction results

## 5.2. Curry Powder data

### 5.2.1. Analyte preparation method for curry powder data

Four different commercially available curry powders were obtained. This is a product that contains organic compounds and also lead which has been picked up from ground water. 0.150 mg of each curry powder was digested in 5.0 ml of concentrated nitric acid (Aristar, BDH Laboratories) at high temperature and pressure inside a PTFE microwave bomb for 2 minutes. The resulting solution was placed in a 25.0 ml volumetric flask and made up to volume with deionised water. At this point the solution was expected to contain between 5 and 20 ppb lead. Four replicate samples of each of the four curry powders were prepared.

These sixteen solutions were then each diluted to 25% of their original concentration with deionised water and divided into four parts. One part was left uncontaminated. To the remaining three parts (sixteen samples in each part) technical staff added a standard solution of Lead Nitrate (BDH product No. 14036) in appropriate amounts so that the concentration of lead in these three parts was increased to cover a range of up to 200 ppb.

Each sample was then analysed for lead using a graphite furnace Atomic Absorption Spectrometer (AAS) operating at emission line 283.3 nm, in order to provide a known value for lead concentration.

The solutions were then re-analysed by differential pulse anodic stripping at a thin film mercury electrode using the author's own apparatus to provide

corresponding voltammetric data. Only the first 28 concentrations, which ranged in lead concentration, as given by AAS, from 1 ppb to 100 ppb were used for the neural network study.

### 5.2.2. Voltammogram production method for curry powder data

The working electrode used was pre plated with mercury. This was done by first polishing the working electrode for 30s using a slurry of aluminia in deionised water after which it was washed in deionised water, dried on a tissue and rinsed again in deionised water. The electrode was then rotated at 300 rpm in a 100 ppm mercury solution ($Hg(NO_3)_2$ in nitric acid) for 10 minutes. During this time the electrode was maintained at a potential of -0.70 V with reference to a saturated calomel electrode. An anodic linear scan was then performed from -0.70 V to 0.00 V at a scan rate of 0.33 $Vs^{-1}$ in order to strip off any material co-deposited with the mercury. Finally the electrode was rinsed using de ionised water.

Each of the 28 samples was analysed by first placing 15.0 ml of the solution in the cup and applying a potential of -0.60V vs. the Saturated Calomel Electrode (SCE) to the working electrode for 60s whilst the electrode was rotated at 300 rpm. This plated the mercury electrode with the material from the analyte. After this time rotation was stopped and the potential maintained at -0.60V for a further 15s.

Before the analysis proper was carried out a number of qualitative investigations were done to ascertain suitable values for the pulse width ($t_d$) and time between pulses ($t_r$). Rifkin[240] showed that for reversible reactions the peak current was independent of these values if $\frac{t_r}{t_d} > 2$ but that it increased substantially as $\frac{t_r}{t_d}$ became less than 1. However he also showed that peak current $\propto t_d^{-0.5}$ and suggested that $t_d$ should not be less than about 10 ms as the Faradic current needs

time to decay. A good heuristic suggested by Rifkin was to use values of $t_d$ of between 20 to 50 ms and to make $t_d = t_r$.

The dependence of peak current on pulse height ($E_a$) was shown by Rifkin to be almost proportional for values of $E_a$ less than $\frac{150}{n}$ where n = number of electrons transferred in the reaction. Large values $< E_a$ gave high sensitivity and low resolution whilst lower values gave low sensitivity but a higher resolution. A good heuristic suggested by Rifkin is to use a value of $E_a = \frac{100}{n}$, or 50 mV in the case of lead.

Accordingly, for the analysis of the curry powder analyte, the differential pulse scan was from -0.6 to -0.2V vs. SCE at a pulse height 50 mV. The pulse width used was 20 ms with a rest time of 20 ms between pulses. The step height was 1 mV. The current recording hardware was set to have a sensitivity of 0.125 $V\mu A^{-1}$, zero offset at an offset range of 5 μA and a gain of x1.

The 28 samples were analysed in a random order to prevent residual effects from one scan interfering with the next. Ten analytical runs were made on each sample with a new mercury film being created for each run.

### 5.2.3. Data set extraction and preprocessing for curry powder data

The particular selection of scan parameters resulted in 400 $\Delta i_{w.e.}$ data points for each scan. This constitutes one data vector. The entire data set therefore consisted of 280 data vectors, each containing 400 data points, representing ten replicate scans of 28 values of concentration as determined by AAS.

The target used for calibration and prediction was a single number in the range 1 to 100 representing the lead concentration in ppb as measured by AAS.

Each value of $\Delta i_{w.e.}$ was normalised to a value between 0 and 1 across the entire pattern set by first calculating the maximum and minimum value for each data point and then adjusting the value of that point according to the equation;

$$x_{i(new)} = \frac{\left(x_{i(old)} - x_{i(min)}\right)}{\left(x_{i(max)} - x_{i(min)}\right)} \qquad \{14\}$$

The AAS concentration value was similarly normalised.

After processing by neural network the output of the network was converted back to a concentration in ppb by using the equation

$$C' = \left[N \times \left(C_{max} - C_{min}\right)\right] + C_{min} \qquad \{15\}$$

where $N$ is the normalised output from the network, $C'$ is the network's prediction of the concentration and $C_{max}$ ,$C_{min}$ are the maximum and minimum values of concentration in the data set as given by AAS.

The entire set of vectors was then reorganised to create two data sets, each containing a training set and a test set.

The first data set used for the networks, called here the Classification Data Set, was obtained by sorting the entire set of data vectors in order of stated AAS concentration and then extracting every fifth vector and placing it in the test set. The remaining vectors constituted the training set. In this way the training set contained vectors from eight replicates of each of 28 concentrations and the test set contained two replicates (i.e. the 5th & 10th) from the same set of 28 concentrations.

This selection method caused the network to act in a classification mode where the test vector was a noisy representation of one of the (equally noisy) training

vectors. There were 224 training vectors and 56 test vectors with 28 possible outcomes (concentrations).

The second data set, called the Interpolation Data Set, was obtained by sorting the original vectors in order of concentration as before but extracting to the test set all ten replicate vectors for every third concentration value. This was intended to force the network to interpolate between the vectors of the training set and produce a concentration prediction that did not map directly onto one of its learnt concentrations.

This extraction method gave a training set of 200 vectors and a test set of 80 vectors with 8 possible outcomes.

The number of elements in the raw data vector was considered to be rather high. Input dimension reduction was therefore investigated using several methods of parameterisation, as shown in table 12.

| Parameters Extracted | Number of Network Inputs |
|---|---|
| Left hand valley minimum value, Right hand valley minimum value, Peak maximum value, Fitted left hand slope, Fitted right hand slope, Fitted background slope on left, Fitted background slope on right | 7 |
| Coefficients of a quadratic fitted to the peak | 3 |
| Coefficients of a quadratic fitted to the background, Coefficients of a quadratic fitted to the peak | 6 |
| As above plus the height calculated from fitted peak to fitted background | 7 |
| Coefficients of a 5th order polynomial fitted to the entire curve | 5 |

Table 12. Different methods employed in an attempt to reduce input dimensionality

These experiments showed that calibrations and predictions from parameterised input vectors appeared to be possible where the original lead peak was well defined. For many of the curves however, using extracted features was too unreliable due to poor peak definition and so, for the work reported here, the dimension of the normalised input vector was reduced to 15 data points simply by starting at data point 115 and taking every tenth data point up to data point 255. This corresponds to the values of $\Delta i_{w.e.}$ for scan potentials -0.485V to -0.345V in steps of 10 mV.

### 5.2.4. Calibration models investigated for curry powder data

The data obtained from the analysis of the curry powder was processed by two different neural network architectures, the general regression neural network and the multi layer perceptron network trained by back propagation. Each architecture was used with both types of training / test sets. The investigation of the multi layer perceptron network was further subdivided by using two different numbers of neurons in the hidden layer. A neural network written by the author and a commercial shell was used for the calibrations. To be consistent with the results presented for the lead standards however, all the results presented here were obtained using the commercial shell, NeuroShell 2 [1] . The preprocessing software was written by the author in Pascal.

The multi layer perceptron consisted of one input layer, one hidden layer and one output layer, fully connected between layers. The transfer function used for the hidden layer neurons was the sigmoid,

$$f(x) = \frac{1}{(1 + \exp(-x))} \qquad \{16\}$$

whilst the output layer used the linear function

$$f(x) = x. \qquad \{17\}$$

There were 15 inputs and 1 output. For the first architecture investigated 20 hidden layer neurons were used and for the second only 3 hidden neurons. No attempt was made to further optimise the number of hidden neurons.

The connection weights were initially set to random values in the range ±0.3. Throughout all the tests a value of 0.1 was used for the learning rate $\beta$, a value of 0.9 was used for the momentum $\alpha$ and weight updates were performed after each pattern presentation.

After every 200 presentations of a training pattern the entire test set was presented and the mean squared error in the output calculated. No weight updates were performed at this time. Each time the test set produced a new minimum value for the mean squared error the weights were saved to disc. Training was stopped when a considerable time had elapsed since the last minimum value and the last saved set of weights was used as the trained network.

The general regression neural network used throughout this work also used a sigmoid transfer for the hidden layer and a linear one for the output. The connection weights were again initially set to random values in the range ±0.3. For each data set used there were as many hidden units as there were patterns in the training set. Calibrations were made using both the Euclidean and the City Block distance metrics.

Once the network had been trained, a search procedure was employed in order to find the optimum value for the smoothing parameter $\sigma$. This was done by repeatedly presenting the test set to the network using different values of $\sigma$ and

measuring the error in the network's output. This process was halted when the value for the mean squared error could be improved no further. A typical plot showing this process for the interpolation data set is shown in Figure 35.

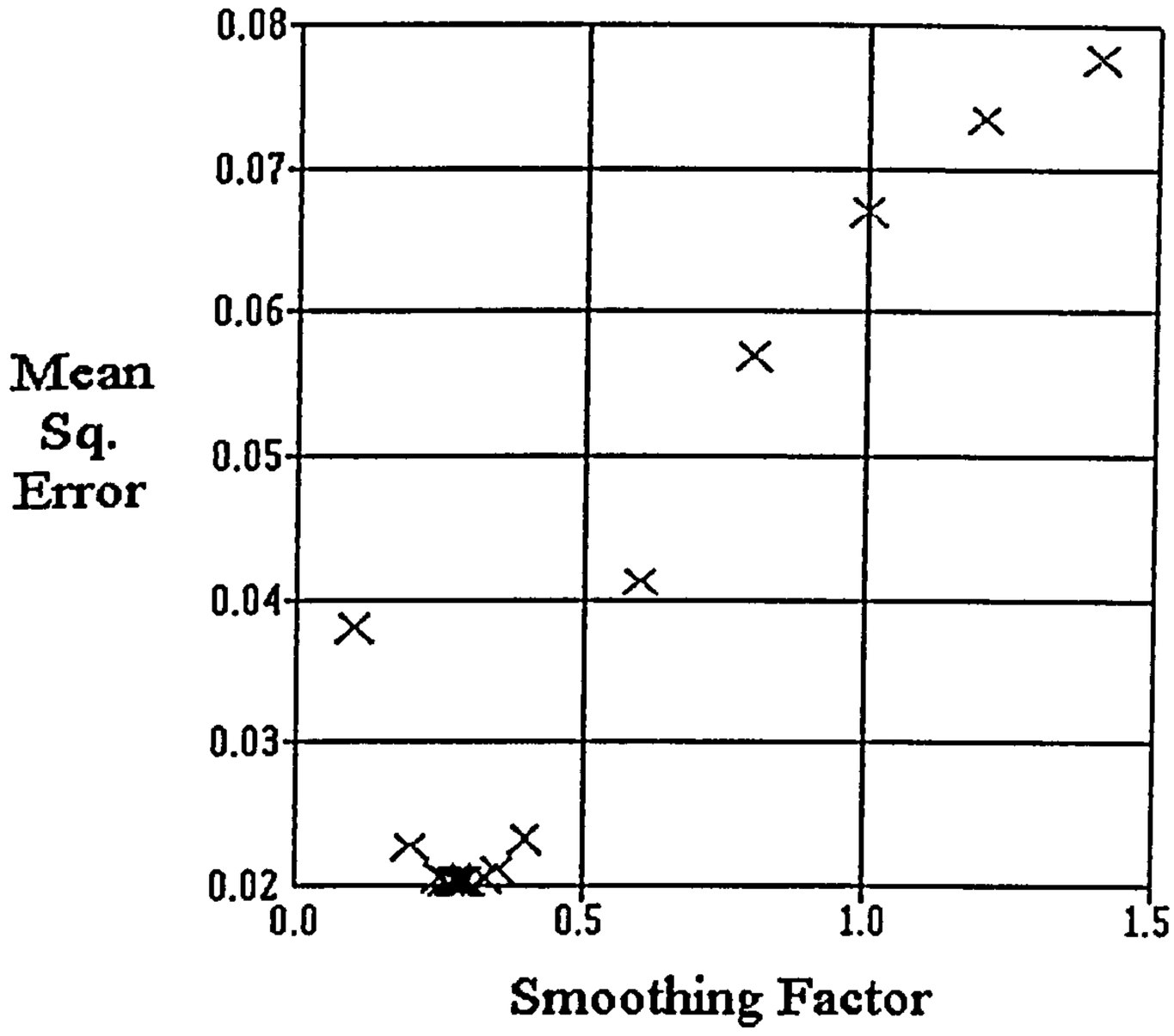



Figure 35. The progress of a search for an optimum smoothing factor in a GRNN network

### 5.2.5. Calibration quality metrics used for curry powder data

The following statistics were gathered for both the training and test sets.

a. The correlation coefficient between the predicted and target concentration values

b. $R^2$, the coefficient of multiple correlation between the predicted and target values.

c. Mean squared error in predicted values

d. Mean absolute error in predicted values

e. Min absolute error in predicted values

f. Max absolute error in predicted values

In addition a scatter graph was made of the predicted lead concentration plotted against the target concentration for the training set and test sets.

The premise made was that networks producing an $R^2$ closer to 1 and a smaller value for mean absolute error were better predictors of actual lead concentration than those with an $R^2$ not close to 1 and a larger value for the mean absolute error

## 5.3. Curry Powder Data Results

The results of processing the curry powder data have been previously published by the author (Manwaring[241]) and the reader is referred to this paper for the further details. A copy of the paper is given in appendix 1.

The voltammetric curves obtained from the curry powder using the pre plated electrode showed a considerable amount of noise. One constituent of this noise appeared to be regular and occurring at a frequency of about 0.3 Hz. This can clearly be seen in Figures 36 to 38. The author can give no satisfactory explanation for this but a similar phenomenon was noticed by De Vries[242] in 1967. Initial investigations indicated that earthing the metal box in which the potentiostatic electronics was housed slightly reduced the amplitude of the spikes as did disconnecting the external hold capacitors connected to pin 7 of the AD585 sample and hold chips. The problem was absent when the lead standards were used in the final part of this work (page 145).

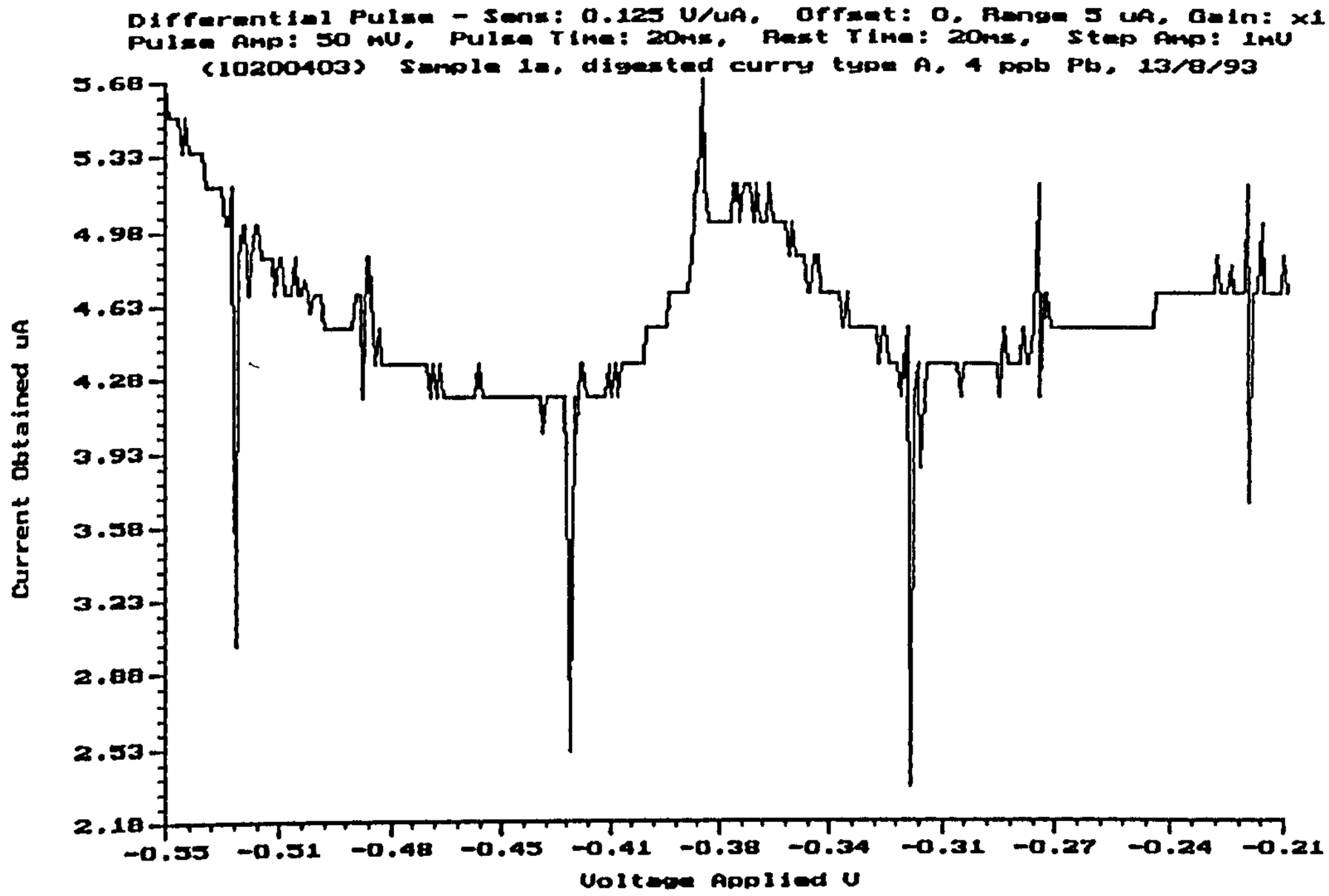



Figure 36. Noise spikes in a voltammogram of digested curry powder containing 4 ppb lead

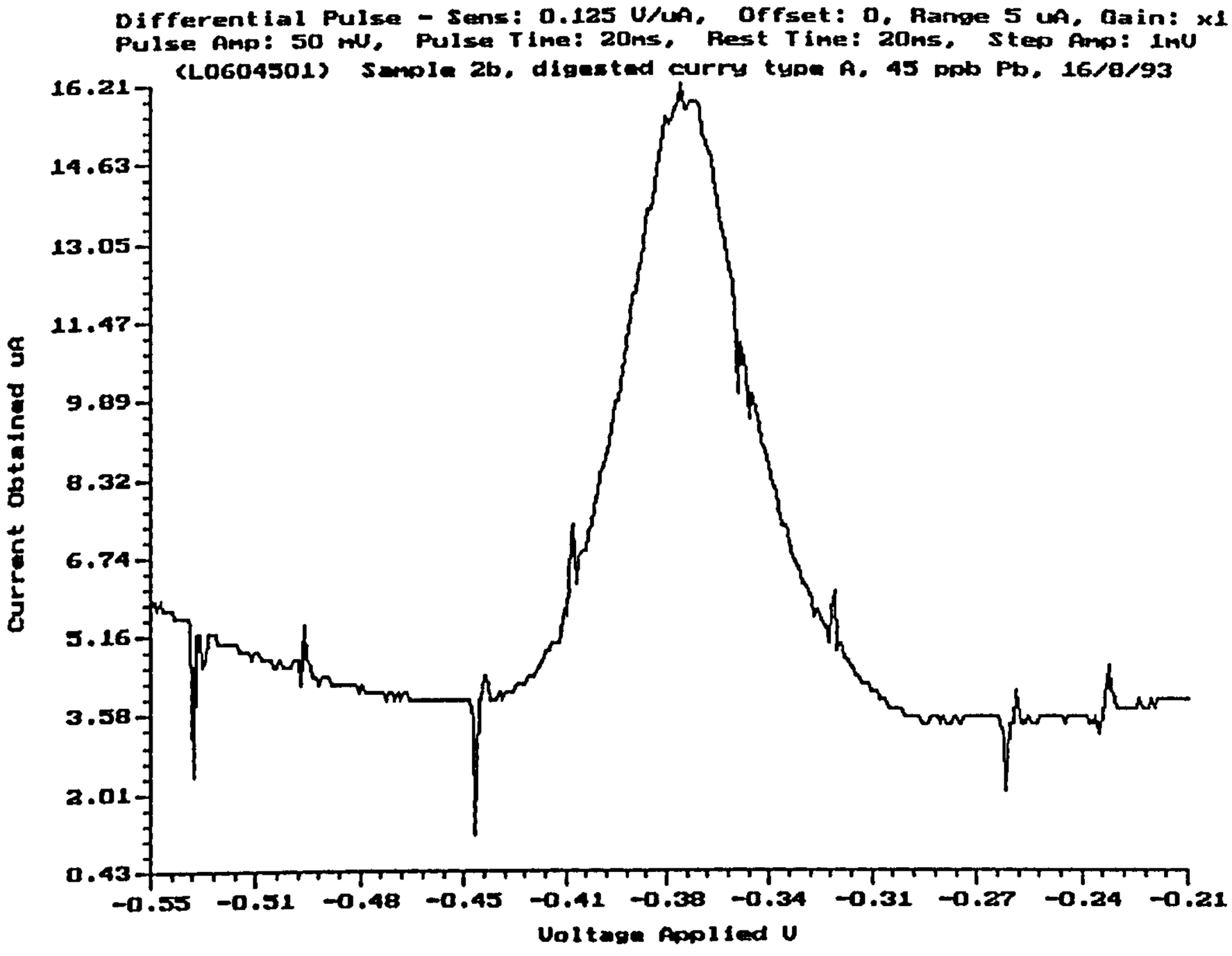



Figure 37. Noise spikes in a voltammogram of digested curry powder containing 45 ppb lead

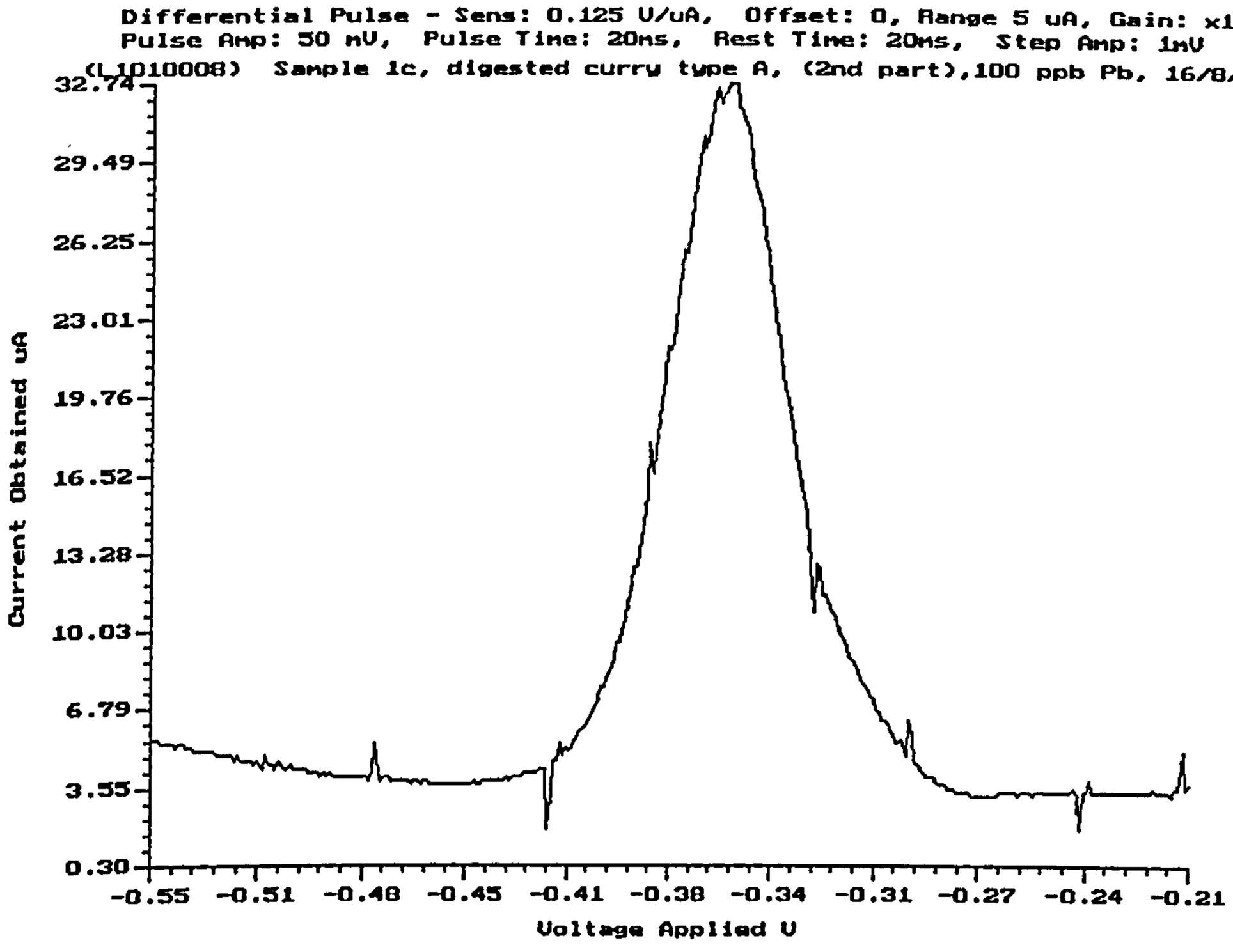



Figure 38. Noise spikes in a voltammogram of digested curry powder containing 100 ppb lead

The numerical results of these tests are shown in table 13 on page 134. Only the results for the test data are reproduced here as these are the ones with the greater significance, the training data results exhibited similar relationships but, as might be expected, they had a higher correlation between predicted and target values.

It would appear, therefore, that both network architectures were suitable for predicting lead concentration using data from $\Delta i_{w.e.}$ against $V_{w.e.}$ curves obtained from Differential Pulse Anodic Stripping, particularly when the number of data points in the curve is severely reduced.

| | $R^2$ | MSE | Mean Absolute Error (ppb) | Min Absolute Error (ppb)r | Max Absolute Error (ppb) | Correlation Coefficient |
|---|---|---|---|---|---|---|
| BP15:20:1 classification | 0.89 | 107.0 | 7.5 | 0.2 | 39.4 | 0.95 |
| BP 15:3:1 classification | 0.87 | 126.6 | 8.5 | 0.0 | 36.2 | 0.93 |
| GRNN city classification | 0.90 | 92.9 | 5.8 | 0.0 | 29.0 | 0.95 |
| GRNN Euclid classification | 0.97 | 30.8 | 1.8 | 0.0 | 26.5 | 0.98 |
| BP15:3:1 Interpolation | 0.75 | 204.6 | 11.4 | 0.0 | 33.7 | 0.89 |
| GRNN Euclid Interpolation | 0.76 | 119.1 | 12.4 | 1.1 | 25.7 | 0.87 |

Table 13. Tabulated results of processing the curry data with various network architectures

Both of the network architectures appeared to be able to model adequately the relationship between voltammogram curve shape and lead concentration. In general the predictive quality is better for classification than for interpolation.

The results of processing using the Back Propagation networks will be described first, concentrating upon predictions made from the test set rather than the training data set. Following this the results of the GRNN network will be given, also for the test set.

### 5.3.1.Back propagation

The results show considerable scatter on the test set predictions with the 15:20:1 network being marginally better than the 15:3:1 Both networks incorrectly allocated a large number of patterns to the same concentration, particularly at low values (< 20 ppb), see figures 40 and 41.

When required to interpolate between the learnt concentrations the networks appeared over sensitive to variations in the replicate Δi vs. $V_{w.e.}$ curve shapes and produced a different prediction for each one.

The time taken to train these networks for classification varied between 3 Hr 40 min for the 15:20:1 network and 42 min for the 15:3:1 network. 5366 and 2548 epochs respectively were required for training. Training for interpolation was faster, partly due to the reduction in the number of vectors in the training set. The 15:3:1 architecture, for example, reached its minimum error in only 8 minutes.

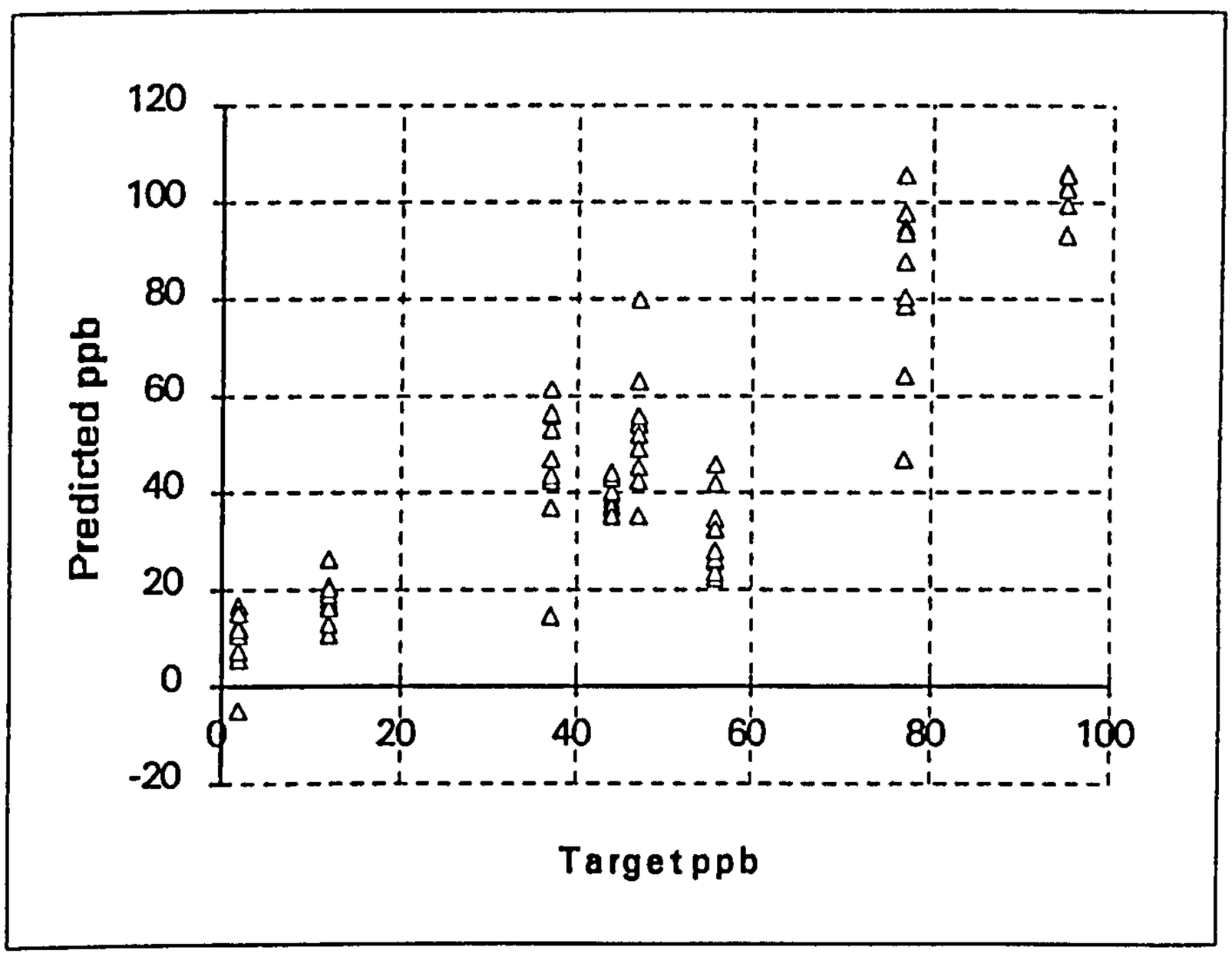



Figure 39. Predicted ppb against actual ppb for BP 15:3:1 using the curry data test set for interpolation

The predictive ability of the networks investigated can be seen from the plots which show the actual concentration plotted against the network's prediction of the concentration, both axes in parts per billion.

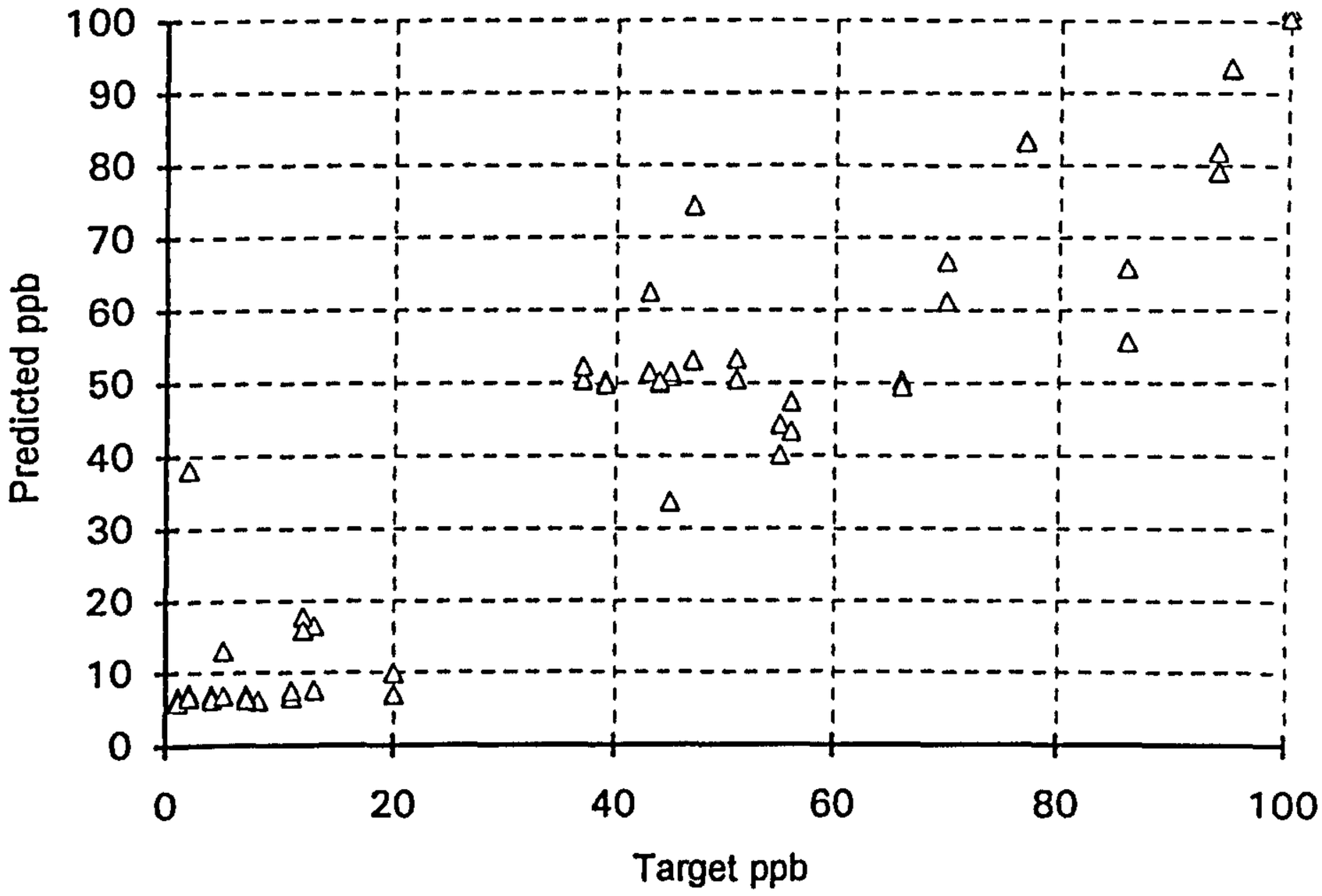



Figure 40. Predicted ppb against actual ppb for BP 15:3:1 using the curry data test set for classification

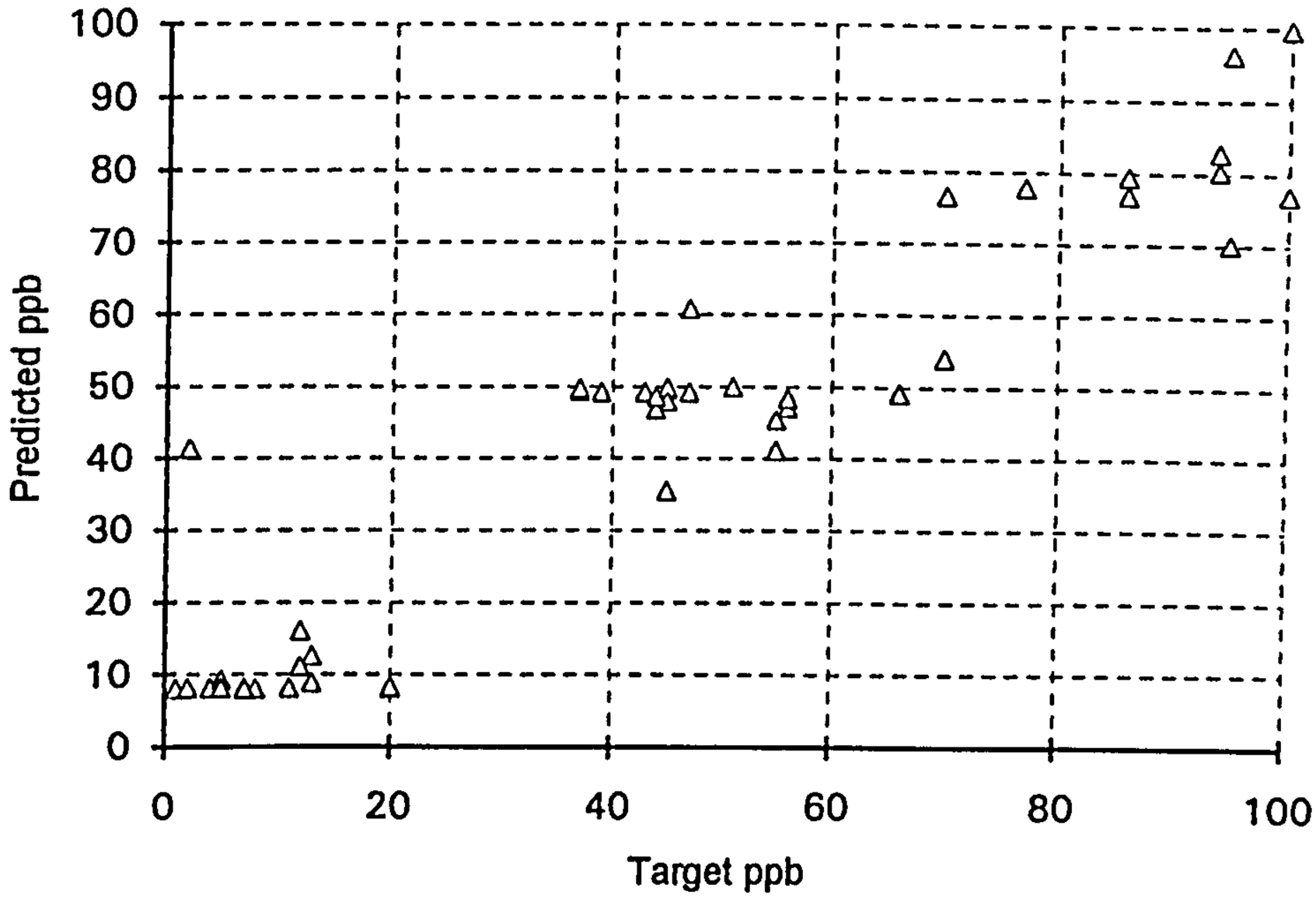



Figure 41. Predicted ppb against actual ppb for BP 15:20:1 using the curry data test set for classification

## 5.3.2. GRNN

When required to interpolate the GRNN showed less scatter than the Back Propagation network and appeared less sensitive to variations in curve shape between replicates.

The Euclidean distance metric was found to be superior to the City Block, producing an almost perfect match between predicted and target values for classification, see figure 44. Very few patterns were incorrectly classified and the maximum absolute error of those that were was only 26.5 ppb.

Training times for the GRNN were particularly fast. The classification task using the City Block metric requiring only 2 min including a search of 25 candidate values for the smoothing parameter and the Euclidean metric needing only 1 min

50s including a search of 22 smoothing parameter values. Times for the interpolation task were similar.

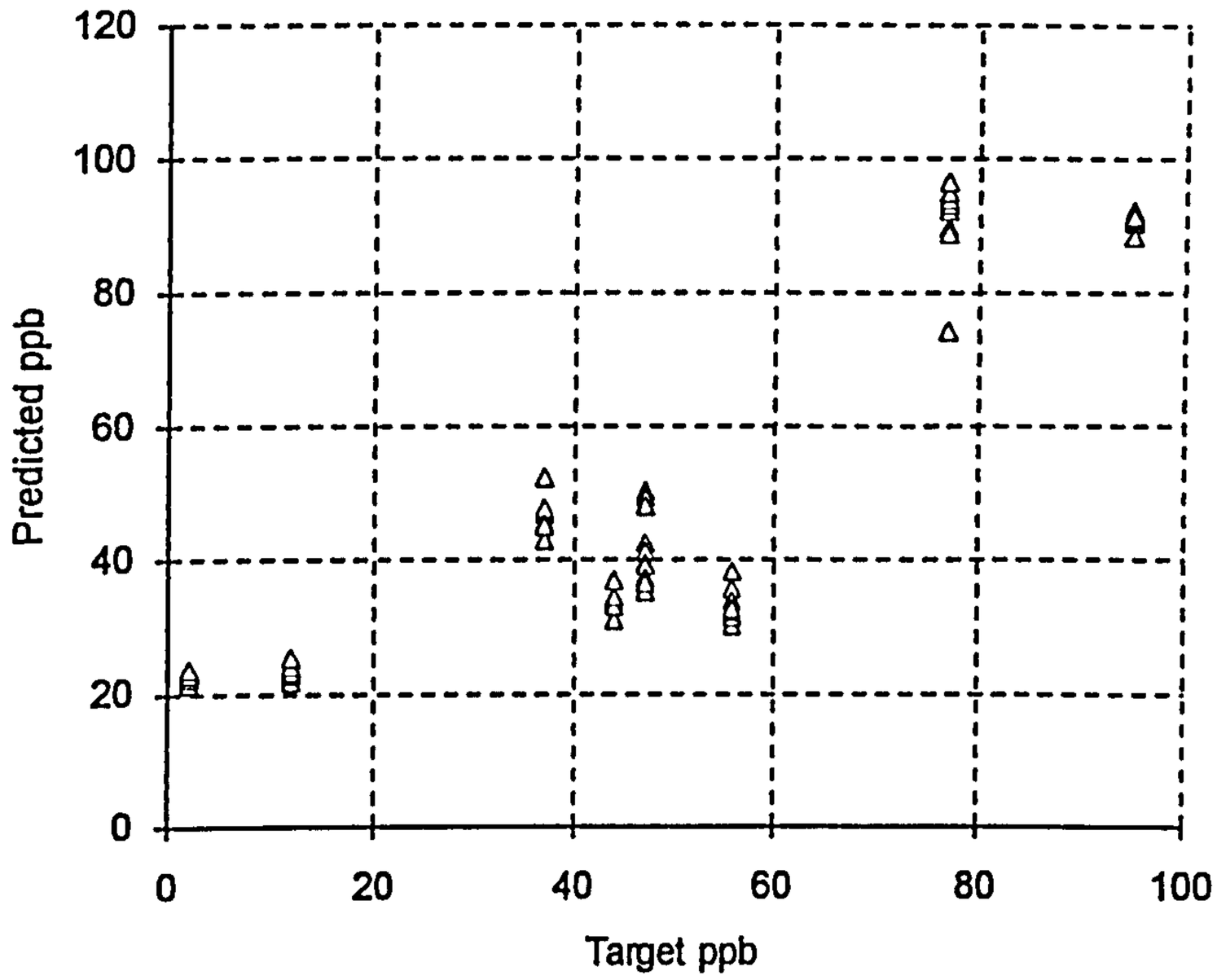



Figure 42. Predicted ppb against actual ppb for GRNN (Euclidean metric) using the curry data test set for interpolation

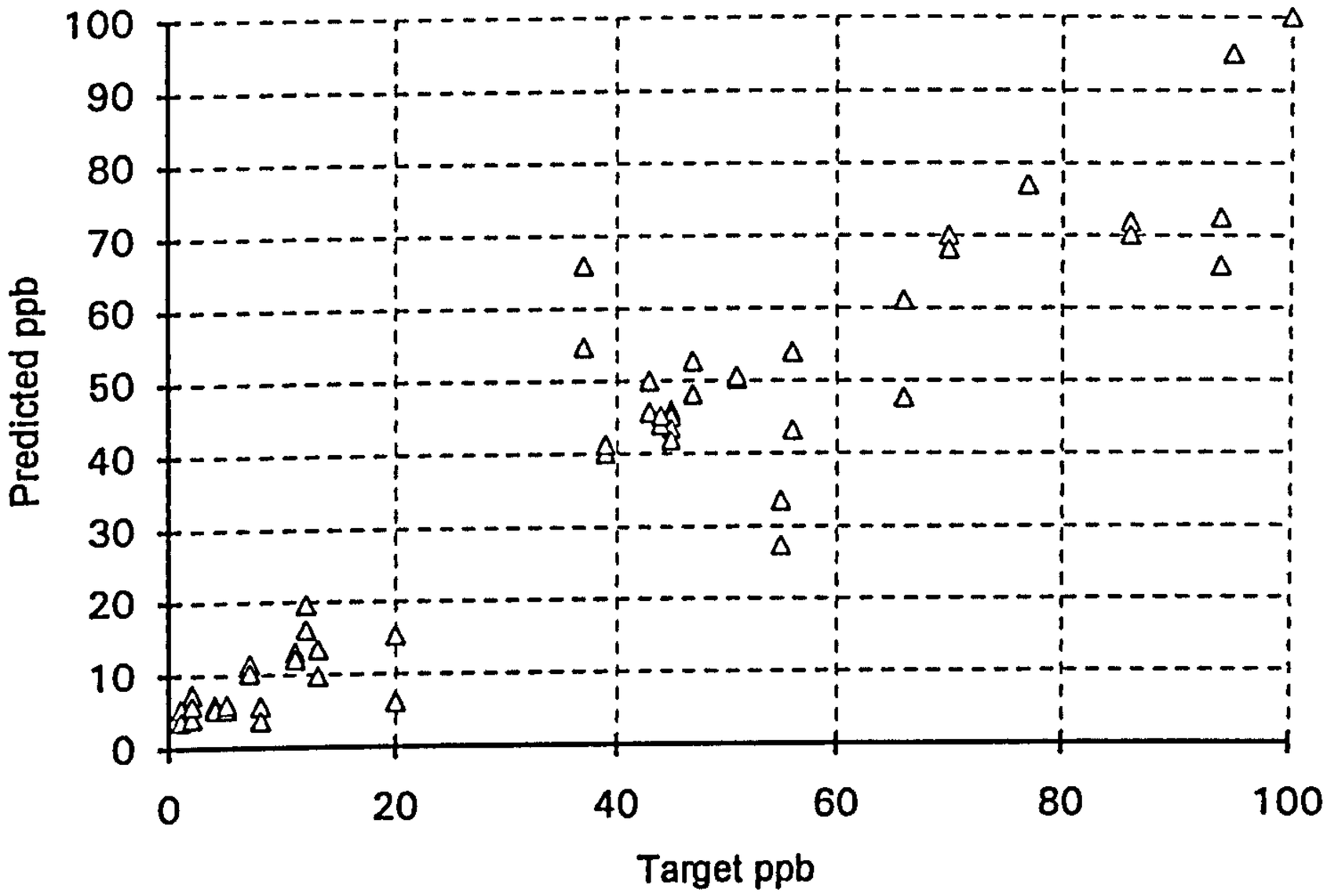



Figure 43. Predicted ppb against actual ppb for GRNN (city block metric) using the curry data test set for classification

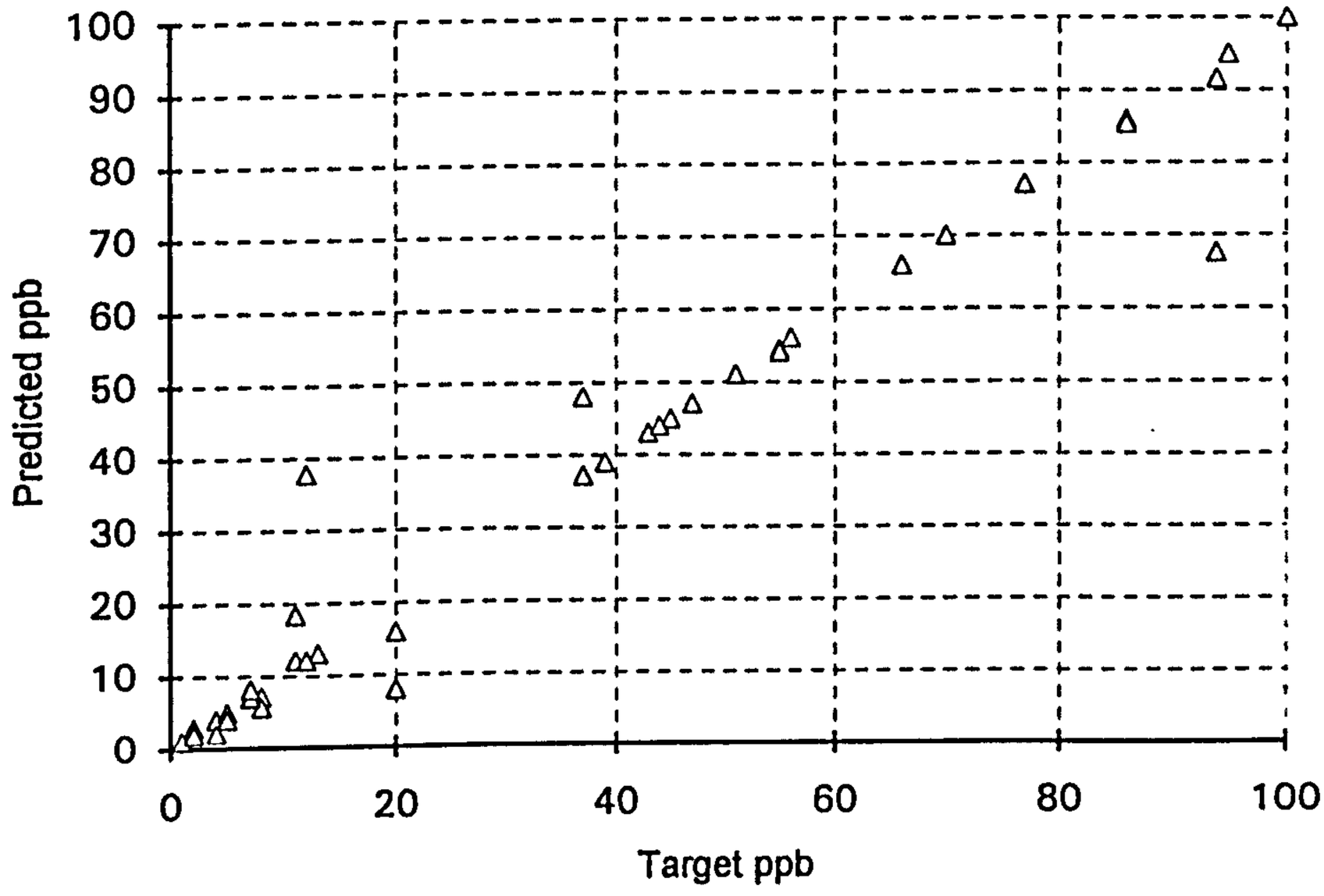



Figure 44. Predicted ppb against actual ppb for GRNN (Euclidean metric) using the curry data test set for classification

### 5.3.3. Observations on Results from Curry Powder Data

The previous results show that a neural network can be used to predict the concentration of lead in a real sample with a reasonable degree of precision when presented with a reduced form of stripping voltammogram. Classification appears to produce better predictions than interpolation, as might be expected. In this mode the GRNN performs better than the Back Propagation network.

However there are two disruptive factors that were observed through closer examination of the data used in this case that influenced the result obtained by the neural networks. These factors contributed to the outcome of the instrument calibration by neural network and resulted in the decision not to use this data for further work. The implications of these factors will now be discussed.

### 5.3.4. Reproducibility of electrode surface

Through investigating the results it was found that the mis-classified lead concentrations had large variations in the shape of the voltammogram within the ten replicate scans. The reason for this is probably due to the extent to which the mercury film could be reproduced when pre plating the electrode with mercury prior to the analyte deposition.

Wong[243] and Hume[244] both noted this reproducibility effect which was explained in detail by Štulíková[245] as being due to the geometric electrode surface being not uniform in its properties but consisting of sites of varying electrical activity. Only about one quarter to one half of the electrode surface consists of active sites with the rest being of low or very low activity.

Hence when deposited at a low potential, such as the -0.4v used when producing the electrode for the curry powder analysis , mercury first forms as

microscopic droplets only at the sites of highest activity. Subsequently, as it is easier to form mercury on mercury than on carbon, continued deposition is more likely on these droplets than elsewhere on the electrode. The result is that with this potential only a few, microscopically large droplets are formed. Depositing at a higher potential, such as the -0.9 V used later in this work with the in situ plated Florence electrode and the lead standards, causes mercury to be initially deposited at the less active as well as the active sites. This leads to the deposition of more initial droplets and hence to the production of a large number of fine droplets, giving a more uniform and hence more reproducible plating.

This effect is quite pronounced: Štulíková's work indicates that the droplets formed at -0.9 V are on average half the size of those formed at -0.4 V.

The implication of this in explaining the large variations in curve shape found within the ten replicates is that each time the electrode was re-polished, different numbers and positions of active sites were formed. Subsequent deposition at a low potential amplified these surface changes through the formation of a few, large mercury droplets. It is believed that this caused poor reproducibly of electrode surface which was reflected in the poor reproducibility of the stripping curves.

### 5.3.5. Reliability of target concentrations

A second factor that became apparent when investigating the data was the reliability of the target concentrations used in training.

As previously described, the samples used for this work were first analysed by technical staff using atomic absorption spectroscopy and then supplied to the author for repeat analysis. The samples were supplied in a number of different plastic bottles. Performing a check on the linearity of peak height with concentration revealed an underlying pattern shown graphically below in Figure 45. This shows

the average peak height over the 10 replicates plotted against the stated concentration for all the data. Adjacent data points have been joined for clarity, this is not meant to imply that there is necessarily a linear relationship when interpolating between points. A linear least squares fit of all the data is included on the plot.

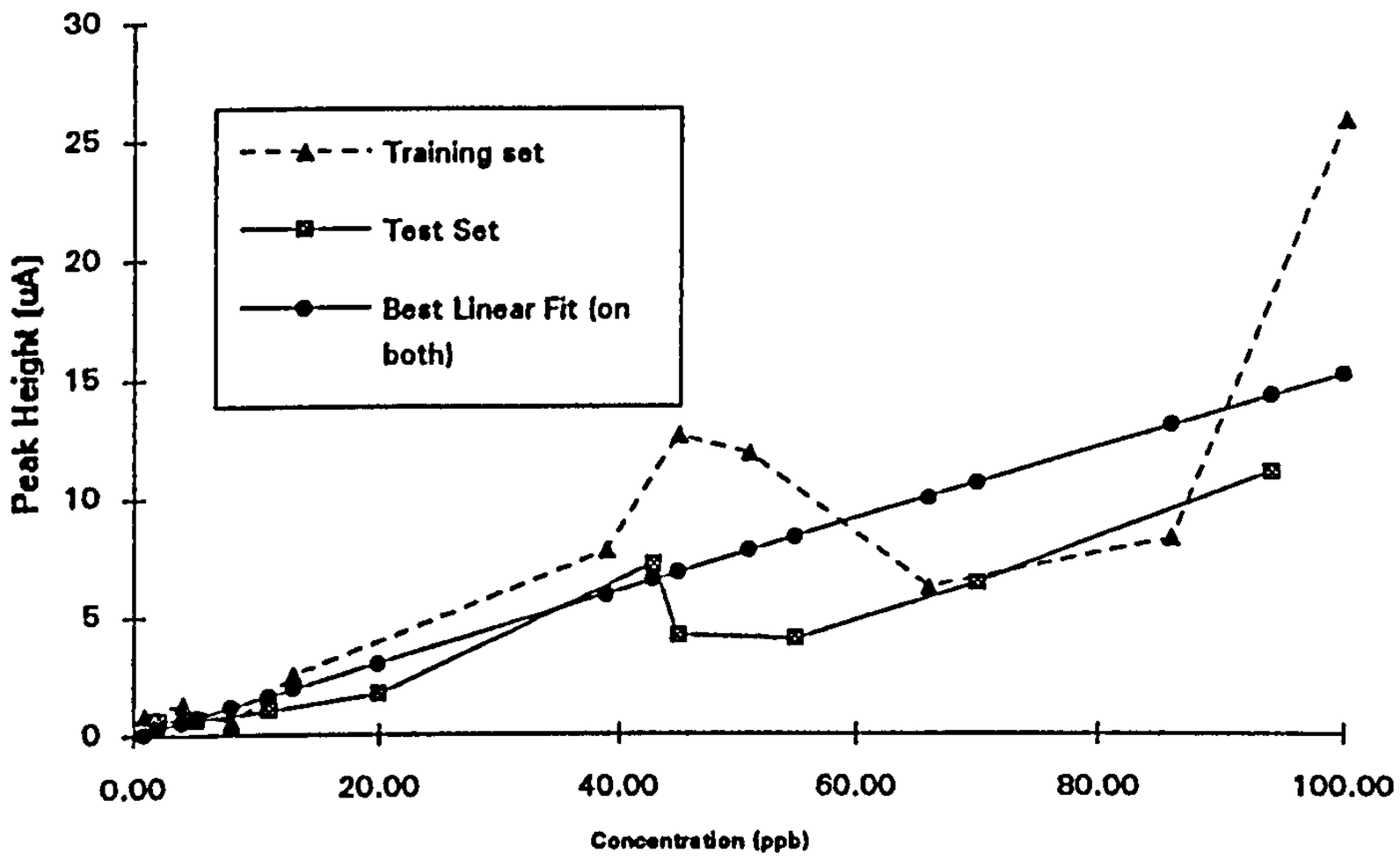



Figure 45. Voltammetric peak heights for the curry data plotted against AAS stated concentration

It was found that not only was the stated concentration sometimes much higher than expected (e.g. 77 ppb in unadulterated digestion that should have been about 2 ppb) but, as can be seen from the plot, the peak heights for some samples indicated that the actual lead concentration was much higher than that given by AAS.

This situation was investigated further and the type of bottle in which the sample was supplied was considered.

Four different types of plastic bottle were used by the technical staff to supply the samples. Some bottles were new ones but others had been used previously for

lead solutions and thoroughly washed. When the plot in figure 45 is reproduced including an indication of the type of bottle used for each sample, figure 46, it clearly shows that contamination of the samples by some of the bottles has taken place. Contamination by bottle type 3 appears to have taken place before analysis by AAS, producing a much higher than expected concentration for that sample.

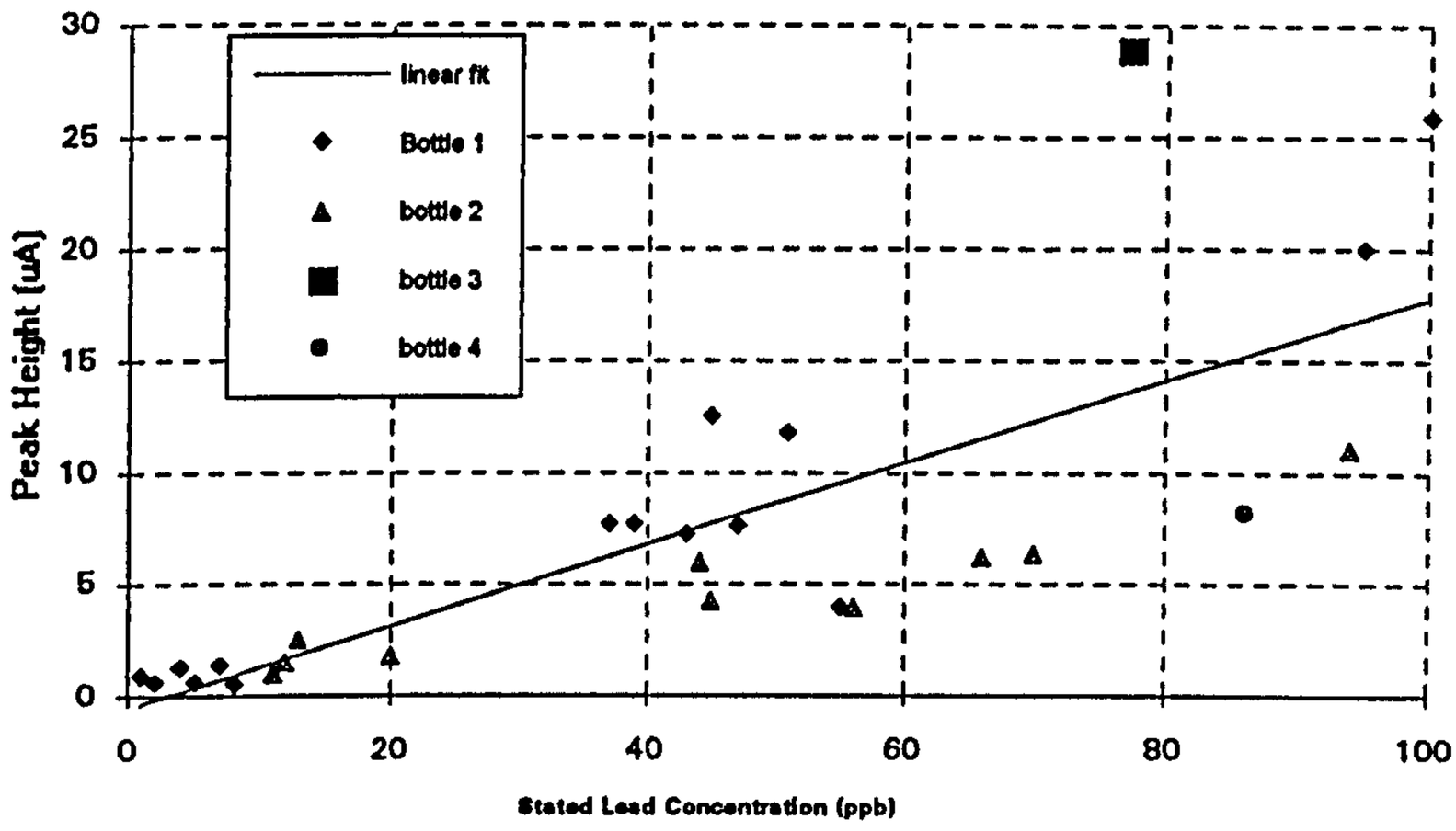



Figure 46. Plot of ASV peak heights showing contamination by certain bottle types

Bottles of type 1 however appear to have continued to contaminate the samples after AAS, generally producing a peak height that is higher than could be expected from the stated AAS concentration when the repeat analysis by ASV was carried out a few days later. This situation means that the targets given to the networks for similar voltammograms could differ over a wide range, by 45 and 70 ppb in one case for a 5μA peak height.

Given this condition it is surprising that the GRNN managed to perform so well on the test data. The explanation for this could be that the test set exhibited the

same anomalies relating to voltammogram shape and concentration as did the training set. However the lowest mean absolute error obtained reflects the poor quality of the original data.

The curry powder data produced some interesting and encouraging results as far as neural network prediction from voltammetric curves is concerned. However, because of the unreliability identified in this data it was decided not to use it in further research towards instrument calibration and to produce new lead standards instead.

## 5.4. Lead Standards Data

This section describes the work leading to the major conclusions of this thesis. Lead standards having a range of concentrations were prepared and analysed. The data from some of the analytical scans was then processed by a variety of methods in order to obtain predictions of the metal concentration in the remaining scans. The precision and accuracy of these methods were then compared using the hypothesis that a neural network model would perform better than the traditional method of using just the stripping peak heights.

### 5.4.1. Analyte preparation method

All the glassware used in the preparation of these standards had been previously washed and soaked in 10 % nitric acid for two weeks. Immediately before making up the standards it was rinsed first in deionised water and then in 10% nitric acid. After it had been used for making up a standard it was washed, rinsed and again stored in 10% nitric acid.

The lead standards were prepared from atomic spectroscopy grade lead nitrate, (Spectrosol, BDH Chemicals Ltd, product 14036) at a lead concentration of 1000 ppm (4.83 mmol $l^{-1}$ ). 1ml of this was first made up to 100 ml with 1 % nitric acid. This provided sub-standard A at a lead concentration of 10 ppm. 1ml of sub-standard A was then made up to 100 ml, again using 1% nitric acid, to produce substandard B at a concentration of 100 ppb.

A 2000 ppm solution of mercury was prepared by dissolving 1.3554g of analar mercuric nitrate ($Hg(NO_3)_2.\frac{1}{2}H_2O$) in 500 ml of 1% nitric acid. 5 ml of this was then made up to 100 ml with 1% nitric acid in order to produce a $5\times10^{-4}$M solution of mercury.

Immediately before the analytical run, quantities of sub-standard B ranging in 0.5 ml steps from 0.5 ml to 10.0 ml were put into a 25.0 ml volumetric flask together with 2.0 ml of the $5\times10^{-4}$M mercury solution and made up to volume using 1% nitric acid. The reason for adding the mercury to the lead solution is explained in the section on page 146.

This procedure produced solutions containing lead at concentrations from 10.0 to 200.0 ppb in steps of 10.0 ppb with each one also containing mercury at a concentration of $4\times10^{-5}$M.

Each standard was made up using the same pipettes in order to reduce random errors. These were rinsed in de ionised water, 10% nitric acid and 1% nitric acid between making up each solution. 15 ml of the final solution was placed in the analytical cup. The temperature of the solutions was 21°C

### 5.3.2. Working electrode preparation method for lead standards data

The electrode used for the processing of the lead standards was of the in situ plated electrode type developed by Florence[246], where the mercury and the species under examination are simultaneously plated onto the electrode substrate from the solution. Because of the shorter overall plating times and low concentration of mercury this results in a mercury film that is between 0.001 and 0.010 μm thick, orders of magnitude thinner than those used by previous workers who pre-plated the electrode with mercury.

De Vries[247] had showed previously that the resolution between neighbouring peaks improves as the mercury film thickness decreases and both he and Lund[248] confirmed that the peak current is independent of the thickness. The above technique is therefore to be preferred as it gives sharper and more reproducible peaks than electrodes formed by pre plating. This observation was confirmed by Goto[249] who also used the Florence electrode to investigate the mathematical transformations of the voltammetric signal.

There is some disagreement regarding the optimum concentration of mercury to be used. Lund[248] found that the most reproducible results were produced when a mercury concentration of $4 \times 10^{-5}$M was used whereas most of the work of Florence was carried out with $2 \times 10^{-5}$M solution. In this work the electrode was prepared using a solution containing the substance under test and mercury at a concentration of $4 \times 10^{-5}$M as previously described.

In general, because of the very thin layer of mercury accumulating at each deposition, the electrode is not cleaned between voltammetric scans but allowed to accumulate mercury while pre concentrating the test species during each of the plating phases. Florence showed that if this is done, the peak potential, $E_p$, changes very slightly on each scan due to the increased mercury film thickness. The peak height however was shown by Lund to be independent of scan number once one or two scans had been made . This property allowed the replicate scans performed in this work to be done without cleaning the electrode each time.

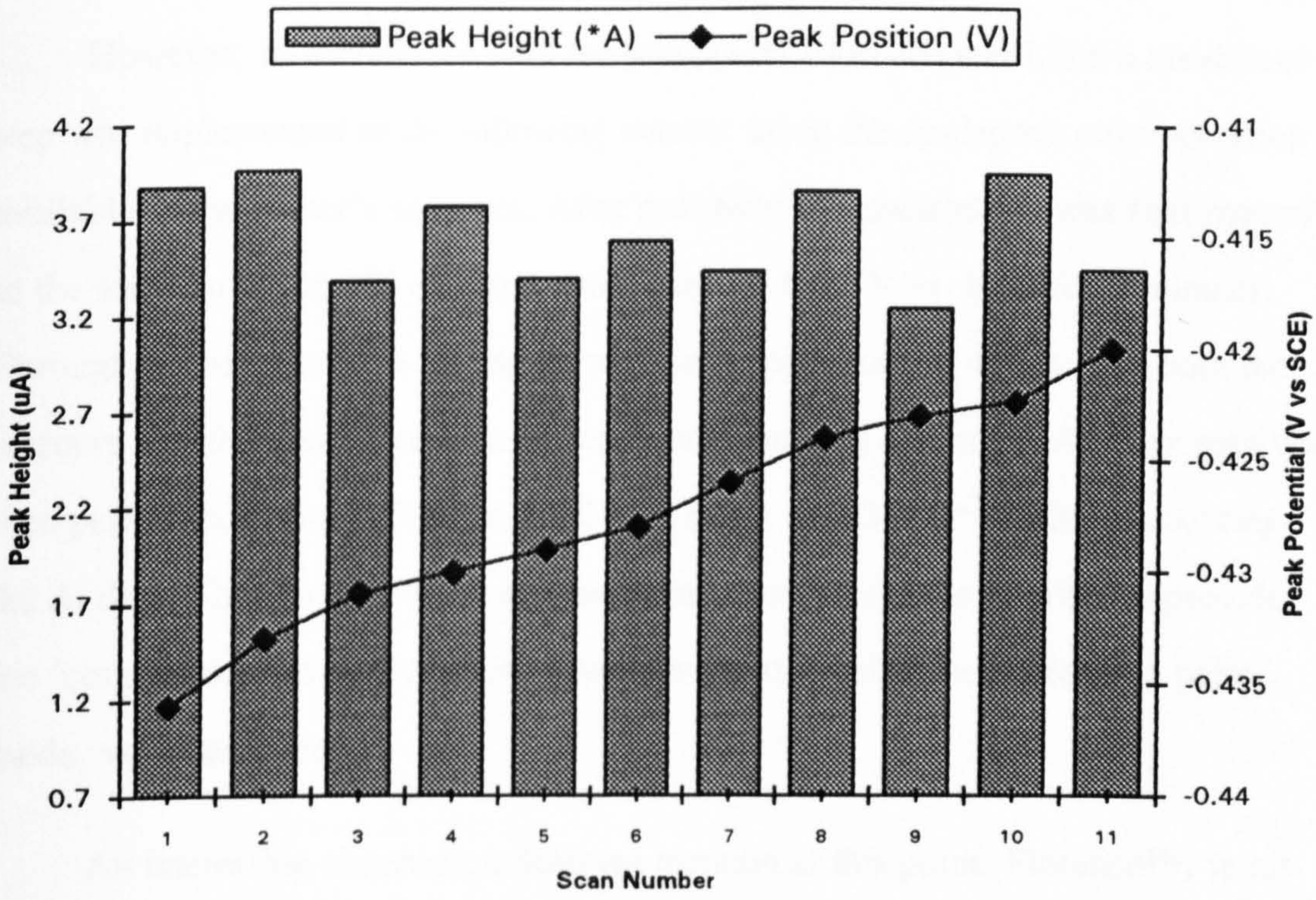



Figure 47. Variation of peak potential and height with scan number

As a confirmation of these observations, figure 47 shows the variation in peak height and peak potential obtained by the author for 11 replicate scans of a 100 ppb lead solution. The slight drift of the peak potential and the generally constant value of peak height can be seen. The rate of drift is in good agreement with that obtained by Florence.

Both Florence and Lund noted the need for what they termed a 'conditioning' stripping scan that was made before the actual measurements were taken on the second and subsequent scans. In their work the first scan appeared to consistently produce a lower stripping peak than subsequent scans. Neither author satisfactorily explained the reason for this, however, and inspection of the table and graph above indicates that the author did not experience this problem.

However, in accordance with the practice of Florence and Lund a conditioning step was implemented in the following manner using the co-deposit strip operation available in the author's software. After polishing the electrode it was first rotated in the test solution at 300 rpm and maintained at -0.70 V vs. SCE for 9 minutes. Throughout this phase the plating current was approximately 40 μA and both the mercury and the species under test were plated onto the electrode. A linear scan was then performed from -0.70 V to -0.10 V at a rate of 0.33 $Vs^{-1}$ without recording the Δi data. The 9 minute plate and the linear scan were made in order to provide the 'conditioning' step. Data from subsequent scans, using the differential pulse mode, were recorded.

An interesting observation deserves mention at this point. Florence[246], in his original paper of 1970, claimed that merely wiping the electrode with first a wet and then a dry tissue was sufficient for complete removal of the mercury film and rendered the electrode ready for another plating. He even reported that he had not polished his electrode for five years![250].

The author's experience with various tissues appears not to support this claim. Indeed, after wiping the electrode with a wide range of tissues it was found that subsequent plating was very poor, resulting in very low or non existent peaks. The reason for this is assumed to be that commercial tissues today have a lubricant or softener such as lanolin added, which leaves a film on the electrode preventing sufficient mercury formation. After experimenting with various polishing powders, the best, and most reproducible results in this work were achieved when, prior to forming the mercury film, the electrode was re-polished for 30s using water lubrication on a polishing stick obtained from Body Shop® and intended for polishing fingernails.

### 5.3.3. Differential pulse analysis method for lead standards data

The standards were prepared and analysed in a random order to prevent residual effects from one scan interfering with the next. Each standard was placed in turn in a new analytical cell and the electrode prepared and conditioned as previously described.

A single analytical scan consisted of rotating the electrode at 300 rpm for a plate time of 1 minute whilst holding it at a potential of -0.70 V vs. SCE. Unlike the technique of Florence however, rotation was stopped at the end of this period and a rest period of 30 s entered with the plate potential being maintained at -0.70 V. Finally the differential pulse scan from -0.70 V to -0.10 V using a step height of 2.0 mV was performed, also without rotation.

Four such scans were made of each solution using differential pulse mode at pulse heights of 5, 10, 20 and 50 mV. This was repeated to provide two voltammograms for each pulse height / concentration combination. The entire process was then repeated using newly made up solutions and with a deposition time of 3 minutes instead of 1 minute.

This provided 320 voltammograms i.e.16 analytical runs on each of 20 concentrations. In addition two replicate scans of 0 ppb lead concentration was made at each of the four pulse heights using a 3 minute plate time. The resulting database therefore consisted of 328 voltammograms each containing 300 $\Delta i$ data points.

The parameters of the system were: Offset value 0, Offset range 5μA, Sensitivity 125 $V\mu A^{-1}$ (i.e. the lowest sensitivity), pulse duration 20 ms, time between pulses 20 ms, gain x1.

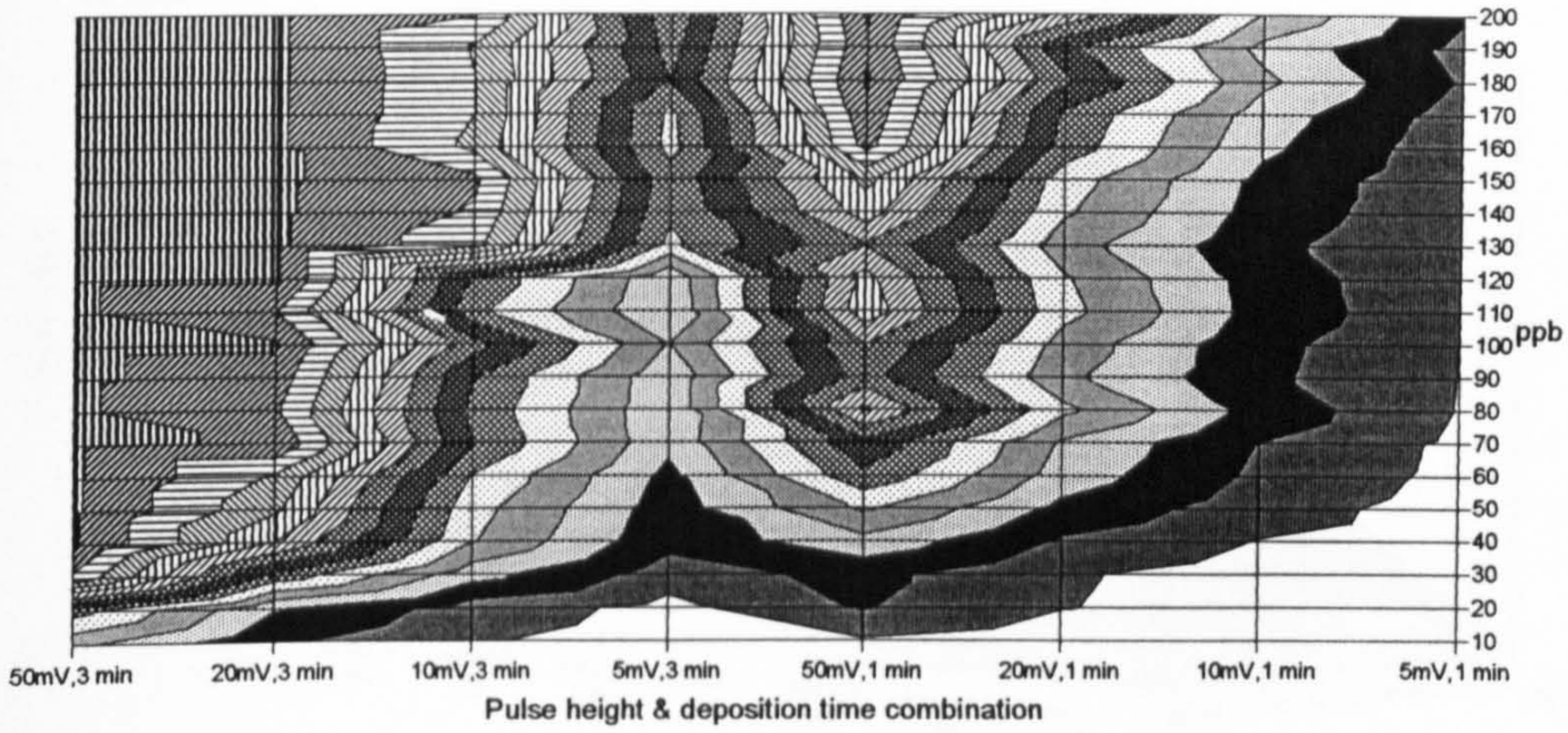



Figure 48. Contour plot of peak heights for all the scan parameter and lead concentration combinations

In order to gain some appreciation of the problem space in which these calibration models are operating, figure 48 shows a contour plot of peak height for each of the eight pulse height and deposition time combinations plotted against the lead concentration. It can be seen that a given peak height could correspond to several values of concentration depending upon the values of the other parameters. Peaks corresponding to high concentrations and high pulse heights that are saturating the I/E converter at 42.8 μA can be seen more clearly in figure 49. This illustrates the same data as figure 48 but plotted instead as a 3D graph.

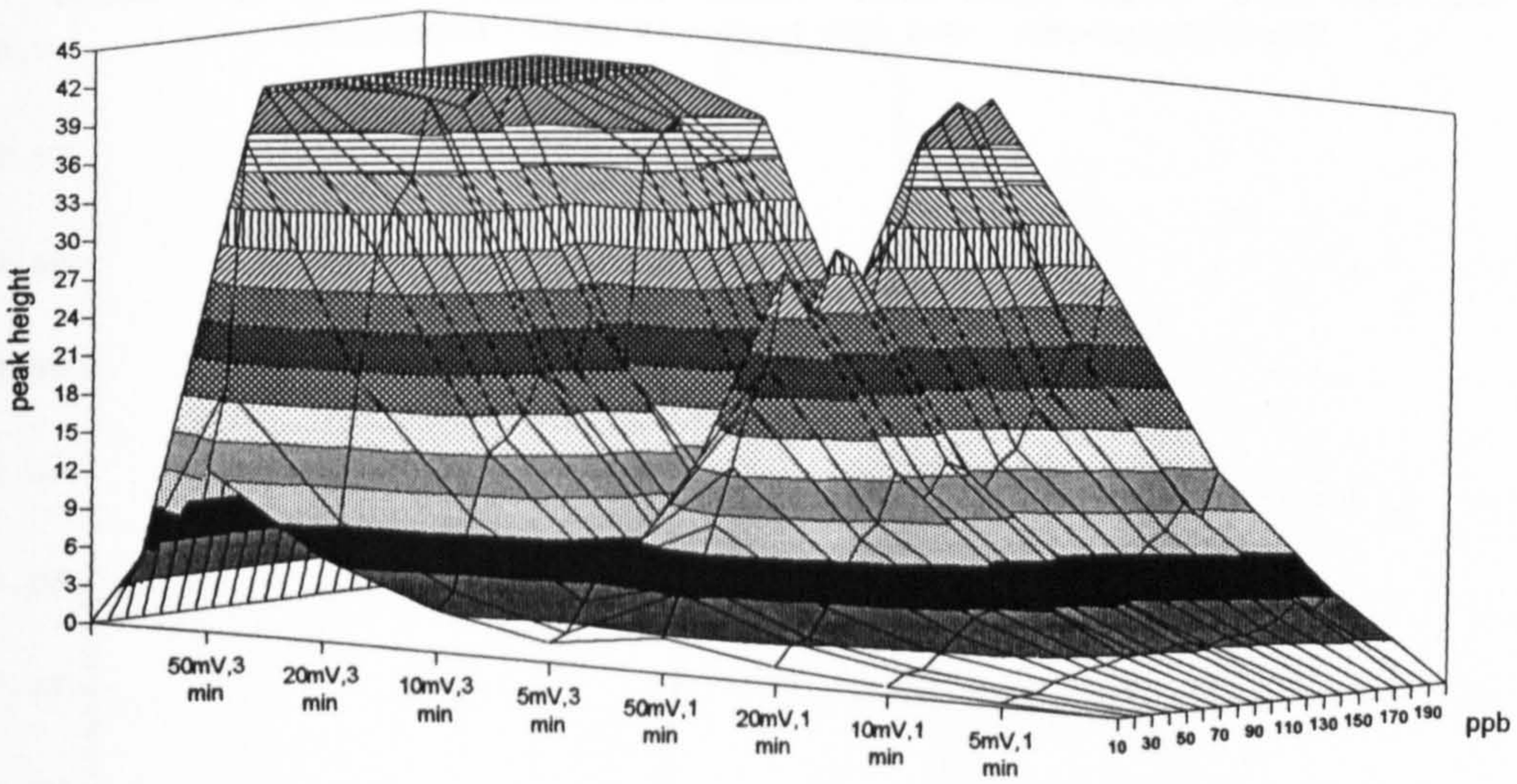



Figure 49. 3D plot of peak heights for various parameter and lead concentration combinations

### 5.3.4. Data set extraction for lead standards data

The data set of 328 voltammograms was split into three groups. The first group contained the 258 non saturating curves that had a peak current less than 42.74 μA and the second was obtained by using all of the original data vectors regardless of whether they saturated or not. The third group contained just the 70 curves that exhibited saturation typical of the type shown in figure 50.

These groups were then each split into a training (calibration) set and a test set by extracting approximately 25% of the vectors at random and placing them into the test set, with the remainder going into the training set.

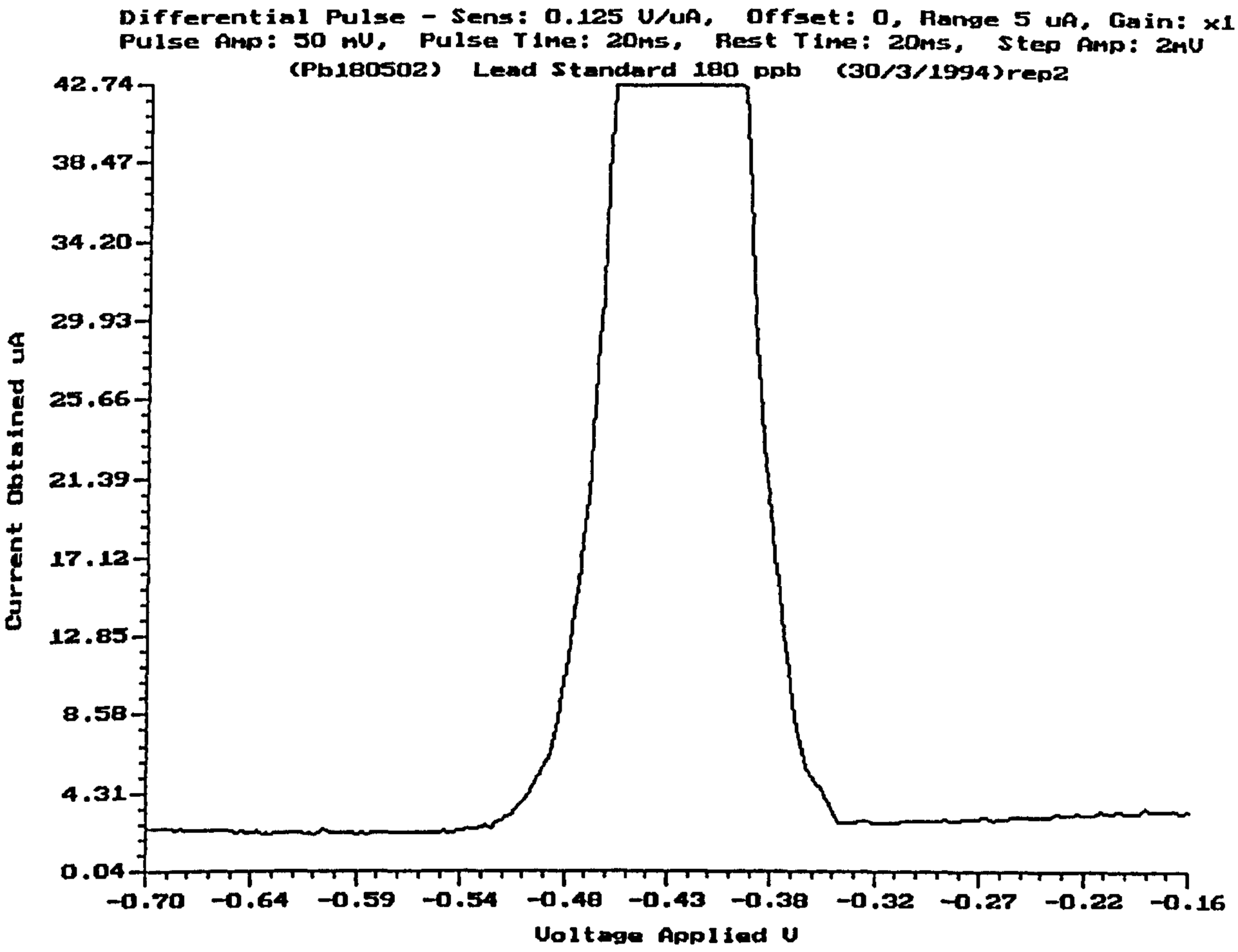



Figure 50. Flat topped peak produced by saturating the I/E converter

The breakdown of the vectors in each set is shown in table 14.

| Group | Vectors in training set | Vectors in test set |
|---|---|---|
| Complete | 229 (inc. 54 saturating ones) | 99 (inc. 16 saturated ones) |
| Non Saturating | 194 | 64 |
| Saturating | 52 | 18 |

Table 14. Breakdown of vectors in the training and test sets

### 5.3.5.Pre processing performed

Two kinds of preprocessing were performed, one to provide peak height data for the 'traditional' calibration method and one to provide a reduced dimensional data vector for use with the neural network models.

Using the raw data from the non saturating curves only, the peak height above a quadratic fitted to the baseline was calculated. This was done using the graphical tool developed as part of the analytical software and used a modified form of the technique used by Bond and Babaric[251].

To find a peak height, two rectangular areas judged to be on on the baseline before and after the peak are indicated using the mouse. The system then plots a quadratic fit to all the points indicated in the two areas in a similar manner to that previously carried out by Bond. However, in this current system, a third rectangular area may now defined, again using the mouse, which includes the likely position of the peak in question. Once this has been defined a second quadratic is fitted to the points in this area and the vertical distance, as measured on the Y axis, between the point of inflection of the peak quadratic and the fitted baseline quadratic is calculated. A vertical line is drawn indicating the position of the height measurement and both the potential of the W.E. and the peak height are displayed on the screen. The units used are the same as those on the axes (μA and Volt). Figure 51 shows a screen dump of a plot for which peak height calculation has been performed. The peak quadratic may not be very visible due to the quality of fit but it fits the points lying above approximately 22μA.

Two independent measurements of peak height were performed for each plot and the average of the two peak heights recorded.

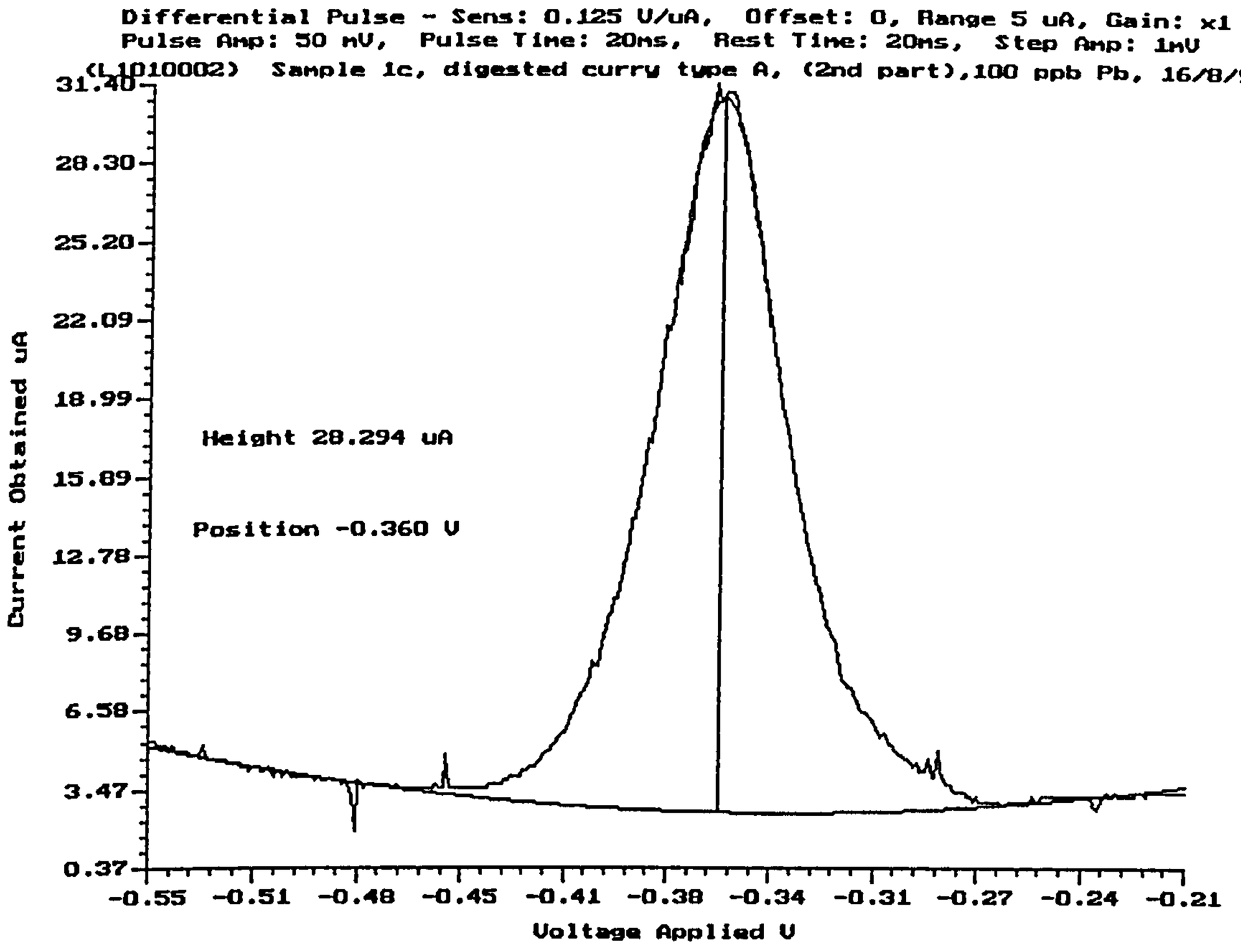



Figure 51. Screen dump of fitted quadratics and peak height display

The peak heights were then normalised to a 1 minute plate time at a 5 μA pulse height using the following equation which assumes a linear relationship between peak height and both deposit time and pulse height. This relationship is independently confirmed by Rifkin[81] and Dillard[84].

$$\text{Height}_{(\text{norm})} = \frac{\text{Height}_{(\text{original})}}{\text{Deposit Time}} \times \frac{5}{\text{Pulse Height}} \qquad \{18\}$$

For the neural network models the raw Δi data, with each data vector consisting of 300 data points was first processed in a similar manner as for the curry powder data by extracting the measured Δi at PD's ranging from -0.60 V to -0.26V in steps of 20.0 mV. This gave 17 current measurements for each analysis. In addition, the

corresponding plate time, pulse height, maximum and minimum values of Δi in the original vector and the standard deviation of the Δi points in the original vector were recorded. This was done for all voltammograms, whether saturating or not.

The 17 values of Δi were then scaled to values approximately in the range 0 to 1 using the formula

$$\Delta i_{(scaled)} = \frac{\Delta i_{(original)}}{42.8} \qquad \{19\}$$

Here 42.8 is the maximum possible value of Δi (in μA) that can be recorded with the sensitivity setting in force rounded up to the nearest 0.1 μA. The other components of the vector were normalised to the range 0-1 individually using the same formula as for the curry powder data, namely

$$x_{i(\text{scaled})} = \frac{\left(x_{i(\text{original})} - x_{i(\text{min})}\right)}{\left(x_{i(\text{max})} - x_{i(\text{min})}\right)} \qquad \{20\}$$

The target concentration in ppb, which ranged from 0 to 200 ppb, was similarly scaled to lie in the range 0 to 1.

Finally, two different sets of input vectors were created for training and testing the networks;

a) A 19 dimensional vector comprising the 17 scaled current measurements plus the scaled values of the deposit time and pulse height.

b) A 22 dimensional vector comprising the 17 scaled current measurements plus the scaled values of the deposit time, pulse height, maximum Δi, minimum Δi and standard deviation of the Δi.

The final data sets used are summarised in table 15, page 157. It must be stressed that these data sets form the starting point for the comparative research that

follows. Any errors in making up the solutions or performing the scans will be present in all data sets and, providing these errors are not too large, will not detract from the validity of the comparative study.

| Voltammograms used | Inputs used | Vectors in training set | Vectors in test set |
|---|---|---|---|
| All | 17 Δi + Td + Ea (total 19) | 229 | 99 |
| All | 17 Δi + Td + Ea + max + min + SD (total 22) | 229 | 99 |
| Non Saturating | Normalised average peak height (total 1) | 194 | 64 |
| Non Saturating | 17 Δi + Td + Ea (total 19) | 194 | 64 |
| Non Saturating | 17 Δi + Td + Ea + max + min + SD (total 22) | 194 | 64 |
| Saturating | 17 Δi + Td + Ea (total 19) | 52 | 18 |
| Saturating | 17 Δi + Td + Ea + max + min + SD (total 22) | 52 | 18 |

Table 15. Summary of data sets used for calibration of lead standards data

### 5.3.6. Calibration models investigated for lead standards data

For all the models used, one or more of the training sets was used to create, or calibrate, the model and the corresponding test set was used to obtain predictions of the 'unknown' lead concentrations.

The first calibration, which is similar to that performed routinely in many laboratories, acted as the base with which all other techniques could be compared.

Using the training set only, the normalised average peak height was plotted against the known lead concentration and a linear least squares regression performed in order to produce a calibration curve. This calibration curve was then used to obtain a prediction of the lead concentration for each of the averaged and normalised test set heights.

The second technique used was a partial least squares regression (PLSR), carried out in order to perform a multivariate calibration in a similar way to that used by Blank and Brown[208]. The calculations were performed by software written in Pascal by the author following an algorithm given in Martens[252]. The PLSR calibration was performed on the same non saturating data as that used for the peak height regression except that it used for its input the 22 dimensional data vector described previously, see page 156.

The calibration was repeated for various numbers of retained factors from 1 to 22 and a prediction made from the test set for each one. By plotting the error in test set prediction against the number of factors retained, an indication of the optimum number of factors to retain was ascertained. This process is detailed in the results on page 170. The PLSR possessing the optimum number of retained factors then provided the second method of calibration

Finally neural networks of a number of different designs were applied to the same, unsaturated data using both the 19 and 22 dimensional input vectors.

The neural network architectures for which results are reported here are:

a. A 19 input, 1 output (19:1) and a 22 input, 1 output (22:1), two layer, linear perceptron having one weight matrix. This was included to investigate whether the problem was linearly separable since, as it had no hidden layer, this model is only able to model linearly

separable problems. The single output unit had a linear transfer function.

b. A 138 input 1 output (138:1), two layer High Order Perceptron having one weight matrix and with the single output unit having a linear transfer function. The inputs comprised all possible multiples of pairs of elements from the 22 dimensional normalised vector described earlier. i.e.

$$(x_1),(x_1 \times x_2),(x_1 \times x_3),\ldots(x_1 \times x_n),(x_2),(x_2 \times x_3),(x_2 \times x_4),\ldots(x_2 \times x_n),\ldots(x_n) \quad \{21\}$$

c. A 19:30:1 and a 22:30:1 Back propagation, three layer MLP having either 19 or 22 input units, 30 hidden units and 1 output unit. The hidden units for both used the sigmoid transfer function and the output unit used the linear function.

d. A 19:194:1 and a 22:194:1General Regression Neural Network using a sigmoidal transfer function for the hidden layer units and a linear transfer function for the output unit. Only the Euclidean distance metric was used.

All the back propagation neural network models used batch update of weights. The learning rate and momentum were both set to 0.1 and the initial weights were random in the range ± 0.3. The GRNN also had its initial weights in the range ± 0.3 and was tested using an optimum smoothing factor that was found by the search procedure mentioned earlier. The commercial neural network tool, NeuroShell 2 was also used for this part of the study.

Except for the peak height calibration, this entire procedure was then repeated using first the data set containing all the vectors, saturating and non saturating and then the data set containing just the saturating curves.

### 5.3.7. Calibration quality metrics used for the comparative study of the techniques

For most of the models the predicted concentration of the test set could be obtained directly from the output of the model when presented with the test data as input. For the peak height regression however, the slope and intercept of the least squares fit on the training set heights and their corresponding lead concentrations was first calculated. These values were then used to calculate the predicted concentrations in the test set given the peak heights in that set. For each of the calibration models investigated, a plot was then made of predicted lead concentration against actual lead concentration for the test set.

The quality of the calibration made by each method was obtained by the following methods corresponding to those commonly found in the literature.

The first metric used was the average value of the 95% confidence interval for each data point over the range of actual concentrations used for the test. Jepson[253] , from whom the calculations for the confidence interval are taken, suggests that using confidence intervals is particularly suited to neural network models. Assuming no systematic error, this provides a measure of meaningful minimum reportable concentrations as well as giving the prediction precision in the same units as the prediction, namely ppb.

The average 95% prediction confidence in predicting a lead concentration from the data used as input to the model was obtained using the equation

$$\text{Average Confidence} = \frac{\sum_{i=1}^{n} C_i}{n} \qquad \{22\}$$

where n is the number of test data points and $C_i$ is the 95% confidence interval for test data point $i$.

The confidence interval $C_i$ is given by Jepson as

$$C_i = t_\alpha \times \sqrt{\frac{n-1}{n-2} s_y^2 \left(1-r^2\right)\left[1+\frac{1}{n}+\frac{\left(x-\bar{x}\right)^2}{\left(n-1\right)s_x^2}\right]} \qquad \{23\}$$

Here $n$ is the number of pairs of actual and predicted values, $t_a$ is the Student's t statistic with $n$-2 degrees of freedom and 95% confidence, $\bar{x}$ and $s_x^2$ are the mean and variance respectively of the actual values, $s_y^2$ is the variance of the predicted values and $r$ is Pearson's correlation coefficient given by

$$r = \frac{\sum_{i=1}^{n}\left(x_i y_i\right) - n\,\bar{x}\,\bar{y}}{n s_x s_y} \qquad \{24\}$$

The second metric reported for comparison of the different calibration methods is the percent standard error (%SEP). This metric is used by Blank[208] when reporting on the predictive ability of the partial least squares regression and enables the prediction precision to be expressed in percentage terms. This measure therefore reflects prediction errors calculated over all analyte concentrations using the concentration mean of the entire data set.

The %SEP is given by Blank as

$$\%\text{SEP} = \frac{100}{\bar{c}}\left[\frac{\sum_{i=1}^{n}\left(\hat{c}_i - c_i\right)^2}{p}\right]^{\frac{1}{2}} \qquad \{25\}$$

where $p$ is the number of test concentrations, $\hat{c}_i$ and $c_i$ are the predicted and actual concentration of the component in sample $i$ and $\bar{c}$ is the concentration mean of that component in the test set. The %SEP is included so that comparisons may be made with other researchers who commonly use this measure

$R^2$, the coefficient of multiple determination, was calculated as a means of obtaining an easily understandable measure of the prediction quality and linearity. Unlike the correlation coefficient, $R^2$ reveals any systematic error that may be present and includes a measure of the scatter present in the data points. The closer $R^2$ is to 1.0 the better the fit of the data points to the regression line. The value for $R^2$ was calculated by NeuroShell 2.

Finally the mean, minimum, maximum and standard deviation of the absolute errors for each calibration method are reported.

To summarise, eight metrics were used to assess the precision of prediction.

1. A graphical plot of predicted against actual concentration
2. The average 95% confidence interval
3. The %SEP
4. The value of $R^2$
5. Mean absolute error
6. Minimum absolute error
7. Maximum absolute error
8. Standard deviation of the absolute error

# 6. Lead Standards Results

## 6.1. Non Saturating Voltammograms

### 6.1.1. Introduction

The results obtained by processing the lead standards have also been previously published (Manwaring[254]). A copy of this paper is given in appendix 1.

The voltammograms from the lead standards produced using the in situ plated Florence electrode were considerably better than those produced for the curry powder using the pre plated electrode. They had a flatter baseline, probably due to less dissolved oxygen and organics, and exhibited much less noise. In addition there was less variation between voltammograms produced from the same concentration of lead. Figure 52 shows a typical voltammogram produced for a 10 ppb lead standard using a 3 minute plate time. The other analytical parameters are the same as for the curry powder scans. For comparison, figure 53 shows the voltammogram produced for the same concentration but with only 1 minute plate time. The quality of the voltammograms is clearly better than those produced earlier even though these are showing the raw data received and are not yet smoothed in any way. Note also that as expected, the peak height for 3 minutes plate time is approximately three times that for one minute.

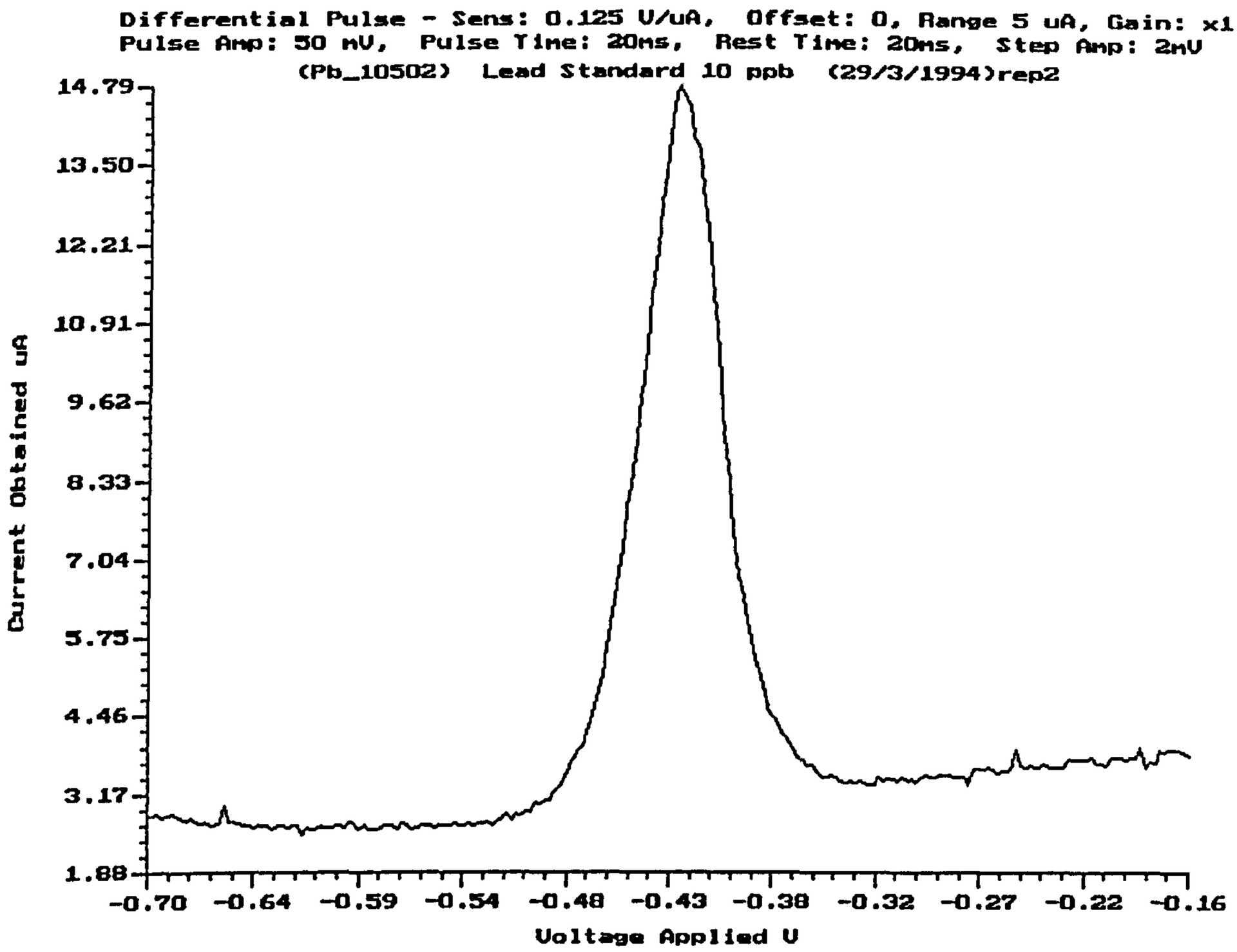



Figure 52. Voltammogram produced from a 10 ppb lead standard using the Florence electrode with a 3 minute plate time

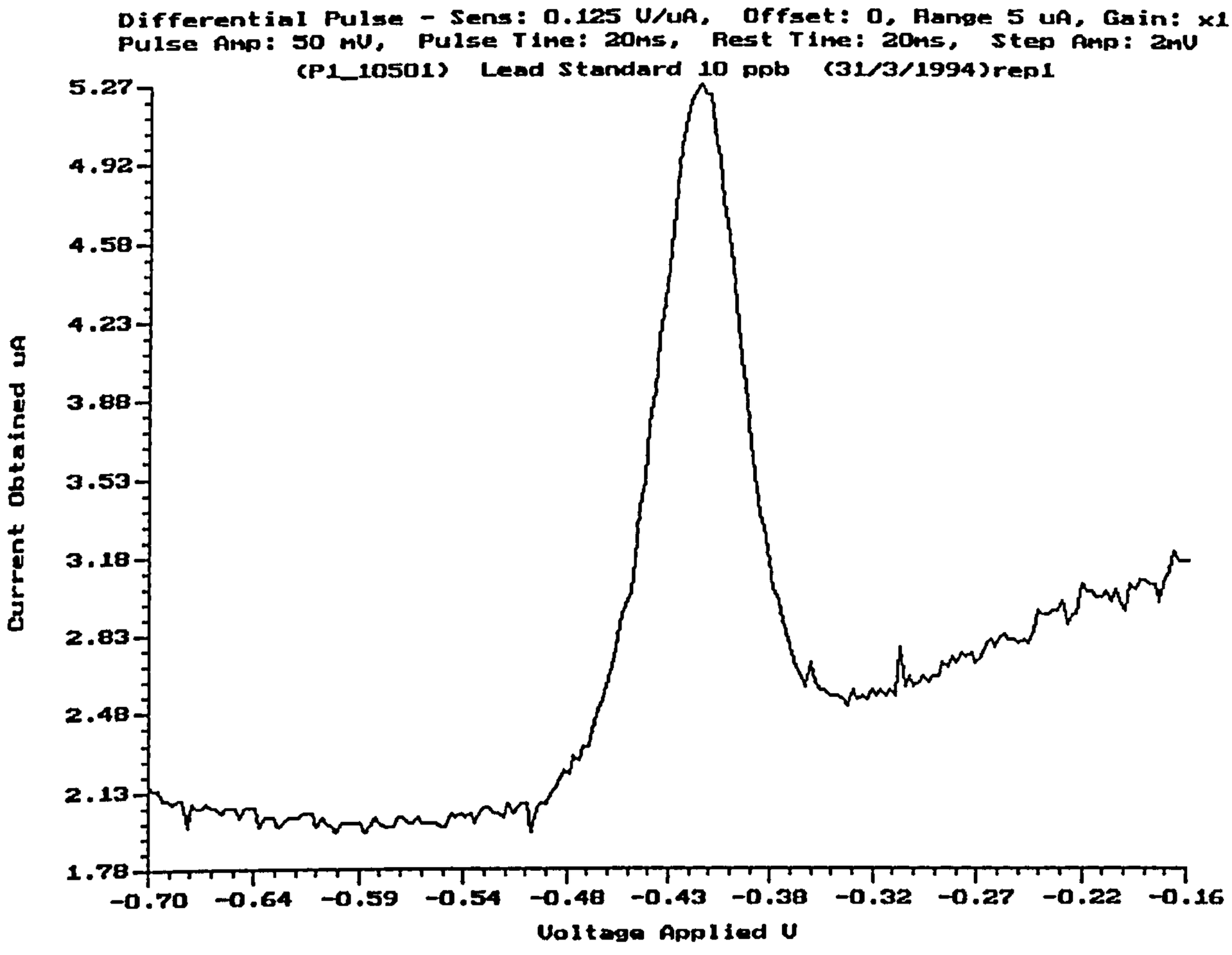



Figure 53. Voltammogram produced from a 10 ppb lead standard using the Florence electrode with a 1 minute plate time

Even at these low concentrations there is no evidence of the regular noise spikes apparent in the earlier data. The voltammogram for 100 ppb, a concentration for which these spikes were still clearly visible in the curry data, shows an extremely good signal to noise ratio even in the raw data. This voltammogram is reproduced in figure 54. Note also that the height is approximately ten times that for 10 ppb.

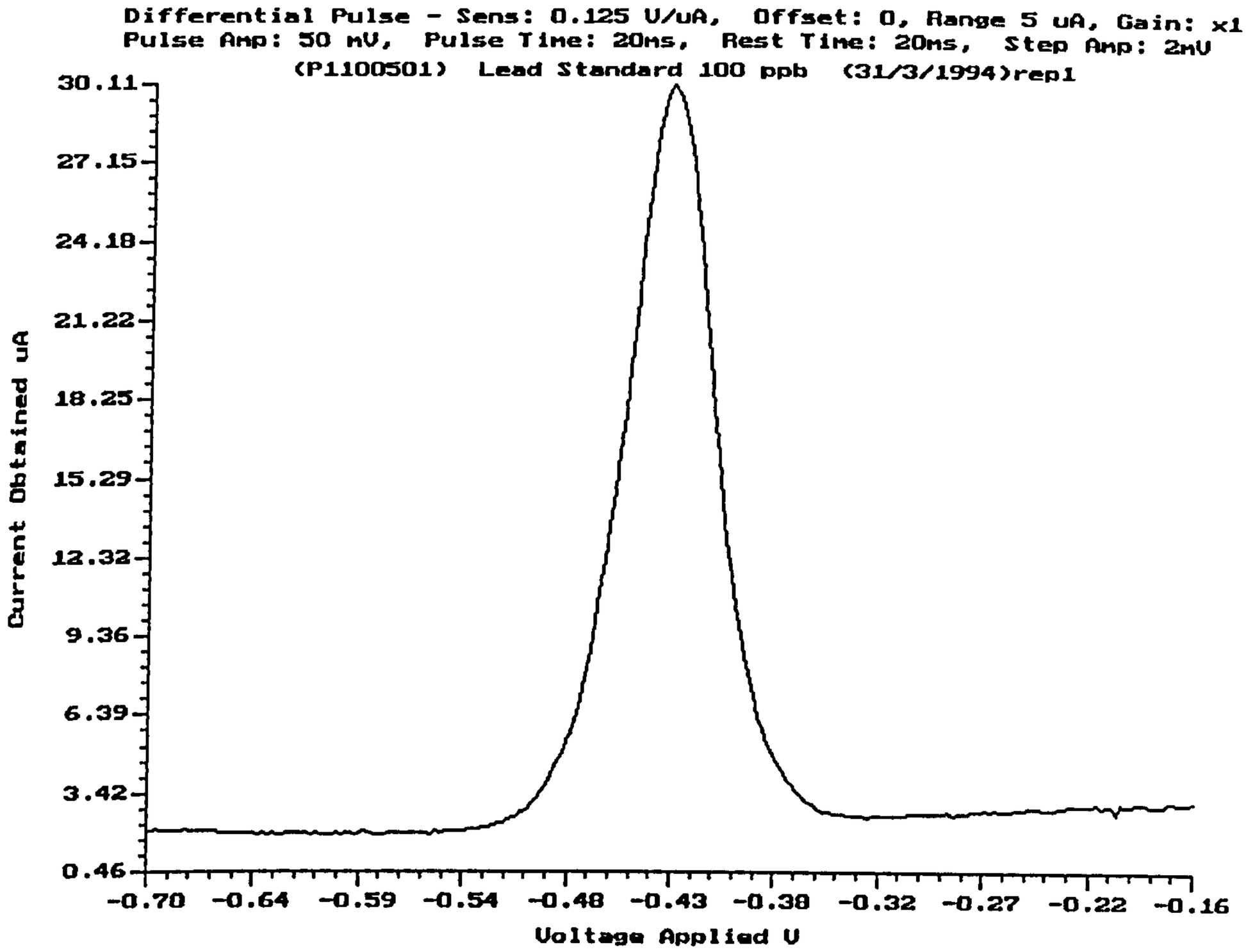



Figure 54. Voltammogram produced from a 100 ppb lead standard using the Florence electrode with a 1 minute plate time

These high quality voltammograms made calculating an accurate peak height a relatively simple matter, given the software already developed. Calibrations from these heights could therefore be made with equal ease. The reduced dimensionality data vector used for the other models however does not contain so much redundancy.

Figure 55 shows a plot of the first 17 elements from two such reduced dimensionality vectors. These elements correspond to the normalised values of $\Delta i$. They are shown plotted against $V_{w.e.}$ for voltammograms from 40 and 50 ppb lead scanned using a 5 mV pulse height and a 1 minute plate time. The anomalous data

point for 50 ppb between -0.5 and -0.45 is due to a point in a region of high noise being selected from the original voltammogram

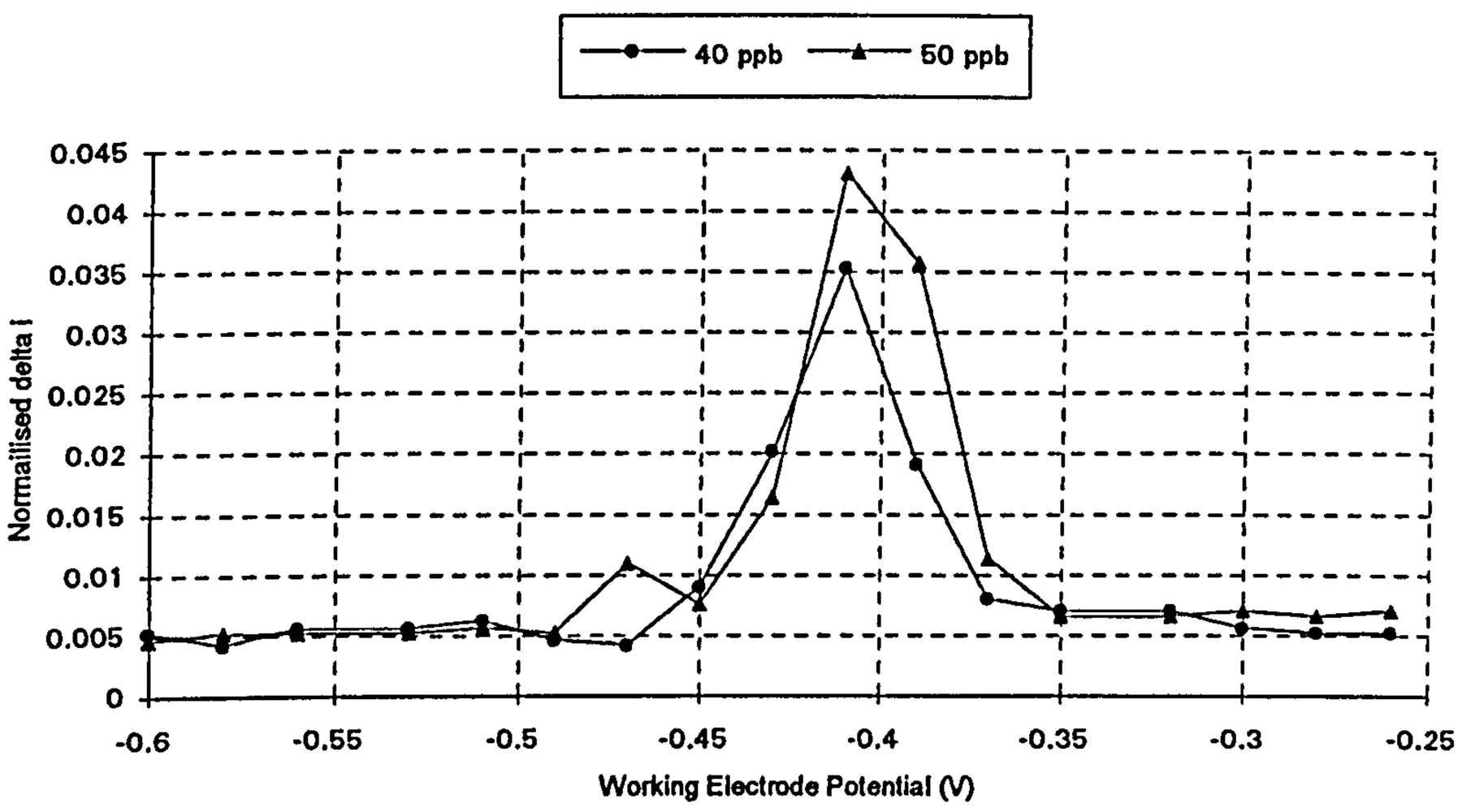



Figure 55. Part of two reduced dimesionality data vectors corresponding to delta i

A succinct tabulation of all the results obtained from the lead standards data is given in the summary in tables 32 to 34, pages 193 and 194. What follows here is a breakdown of the results under the three classes of non saturating voltammograms, all voltammograms and saturating voltammograms only, as previously described.

The results of processing the voltammograms by performing a least squares regression on the training set peak heights will be described first. Following this will be the results of using partial least squares regression, elaborating on the method used to optimise the number of factors extracted from the regression. Finally, the results of performing a similar exercise to optimise the number of hidden units in the MLP will be described, prior to giving the results of calibrations using neural networks.

## 6.1.2.Prediction from Least Squares Regression on peak heights

A plot of the normalised peak heights for all the non saturating voltammograms in both the training and test sets is shown below in figure 56. Each triangle represents the normalised peak height of one voltammogram. It can been seen that there is a reasonably linear correlation between the normalised heights and the concentration.

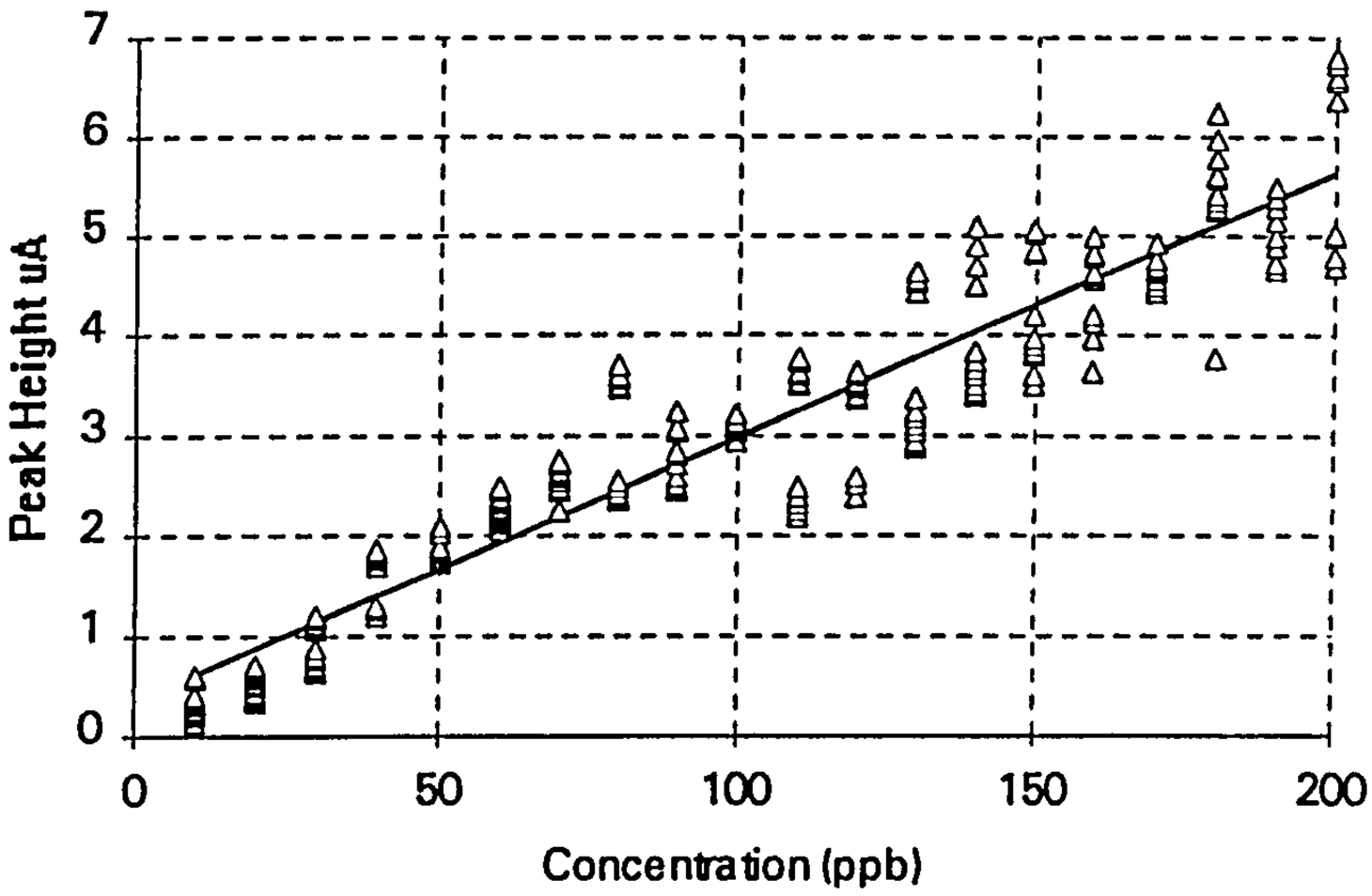



Figure 56. Normalised calibration curve for peak heights against concentration for all non saturating voltammograms

Using this regression line, the predicted concentrations of the test set samples were found from their normalised peak heights. The known concentration of each sample was then plotted against the predicted concentration for each sample.

The result of performing this operation is shown in figure 57. Here the 95% prediction confidence limits have been plotted (dotted lines) either side of a least squares regression line to indicated the quality of the calibration performed in this manner. This will be the format of all subsequent calibration plots. The indicative values that may be extracted for this plot are given below in table 16.

| Avg Conf. (± ppb) | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) |
|---|---|---|---|---|---|---|
| 30.0 | 16.8 | 0.89 | 13.8 | 37.9 | 0.1 | 10.4 |

Table 16. Tabulated results for peak height calibration

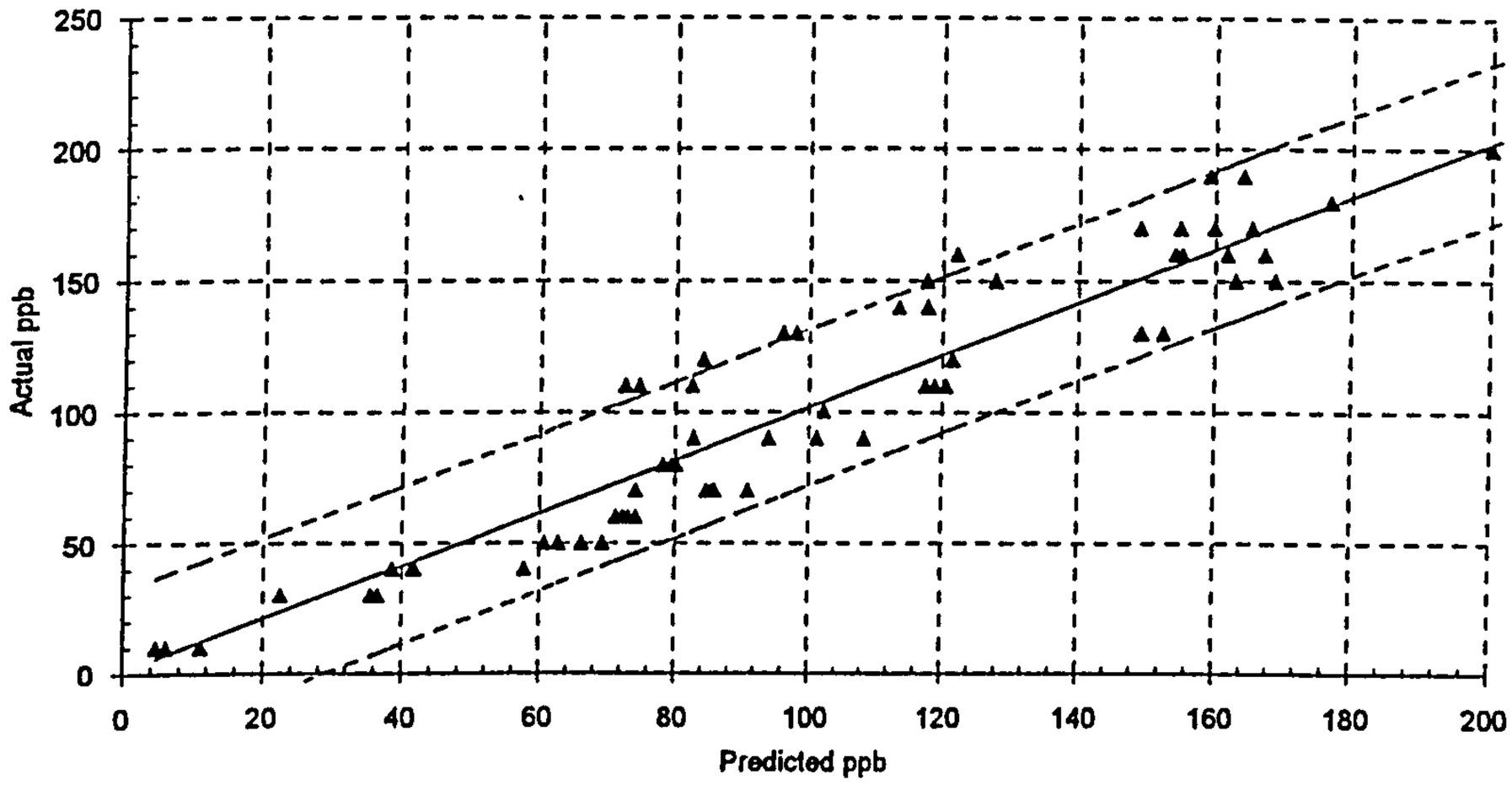



Figure 57. Calibration of lead standards from regression on training set peak heights

This plot, and the values in the above table, are used as a base with which to compare other calibration methods. Of these the reader's attention is particularly

drawn to the results for the 22:30:1 three layer neural network trained by back propagation, the calibration curve for which appears on page 184.

### 6.1.3.Prediction from partial least squares

#### 6.1.3.1. Optimisation of number of Factors in PLS

The data used for this calibration was only that which did not saturate the hardware and therefore comprised 194 calibration vectors and 64 test vectors. All the 22 elements described on page 156, including pulse height, plate time, $\Delta_{i(max)}$, $\Delta_{i(min)}$ and the standard deviation of $\Delta_i$ were used as the independent variables input to the PLS regression. The lead concentration was the single dependent variable. All the data was mean centred and each vector scaled to a length of 1 before calculating the regression. The training set was used to build the regression model and then the test set used to obtain predictions of the lead concentration which were finally plotted as before.

The quality of a PLS regression depends upon the number of 'factors' extracted from the data. In this work the regression and prediction was repeated 22 times, with each regression extracting a number of factors from 1 to 22. The percent standard error on the test set predictions calculated for each one. The results of this exercise are shown graphically in Figure 58 which is a plot of the percent standard error of prediction in the test set plotted against the number of factors retained.

This graph indicates that the optimum number of factors to retain was 9, this number producing the lowest percent standard error. Using fewer factors than this appeared to create a model with an inadequate fit and hence a high %SEP. If slightly more than 9 factors are used the model again becomes worse due to the over

fitting of the training data. Once about 15 factors have been extracted there is little further information in the training data set and there is no change in the %SEP.

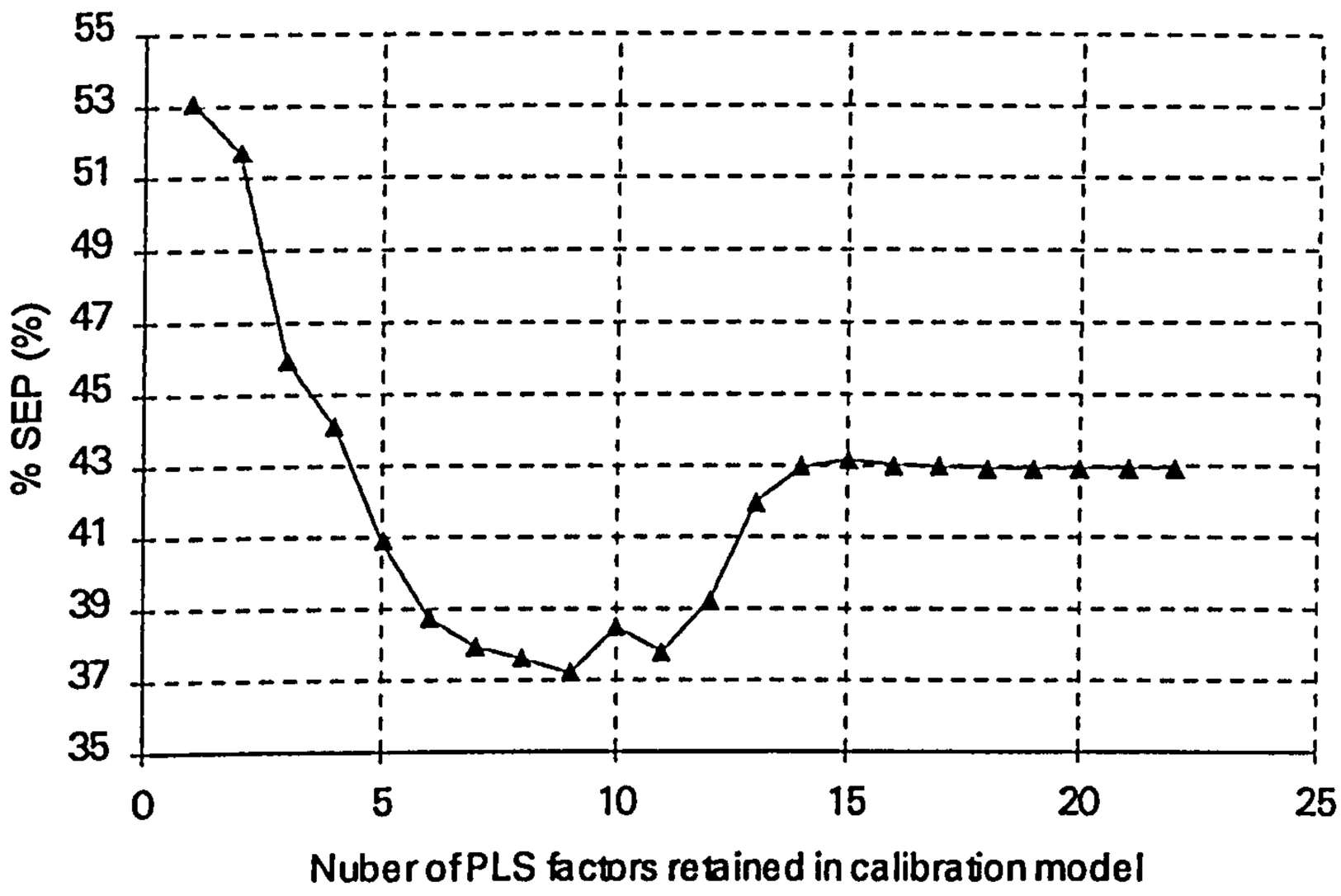



Figure 58. % SEP for non saturating test set against number of PLS factors retained

## 6.1.3.2. Prediction confidence Limits Obtained using PLS

The initial results of predicting the test set concentration using a nine factor linear PLS model of the calibration set were not very encouraging. Figure 59 shows the predictions of the test set plotted in the same format as that used for the peak heights. The numerical indicators are given below in table 17. It can be seen that these values indicate a worse fit using this model than peak heights regression.

| Avg Conf. (± ppb) | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) |
|---|---|---|---|---|---|---|
| 64.8 | 37.2 | 0.50 | 30.3 | 84.3 | 2.7 | 22.9 |

Table 17. Tabulated results for PLS regression

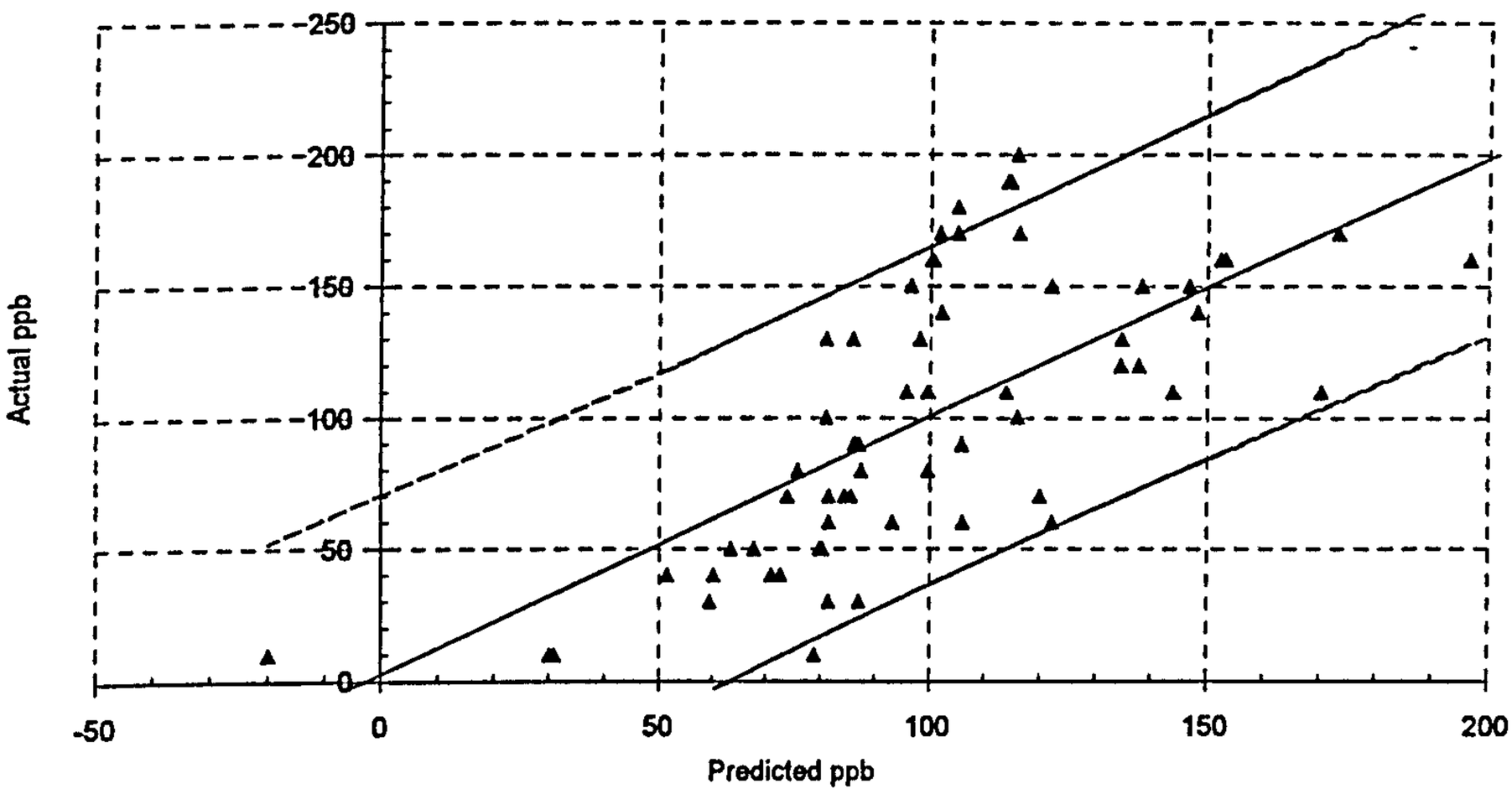



Figure 59. Actual vs. predicted concentration for non saturated standard curves using linear PLS

### 6.1.4.Prediction from 19 input, two layer perceptron

The results of these investigations indicated that the problem did not lend itself to easy modelling by such a linear system. Figure 60 and table 18 show the results of training with just 19 inputs.

| Avg Conf. (± ppb) | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) |
|---|---|---|---|---|---|---|
| 60.5 | 36.2 | 0.56 | 27.6 | 80.8 | 0.0 | 24.5 |

Table 18. Tabulated results for 19:0:1 neural network

The average 95% prediction confidence here is very close to that for PLS. Indeed closer examination indicates that similar test vectors in each calibration have been allocated similar concentration predictions.

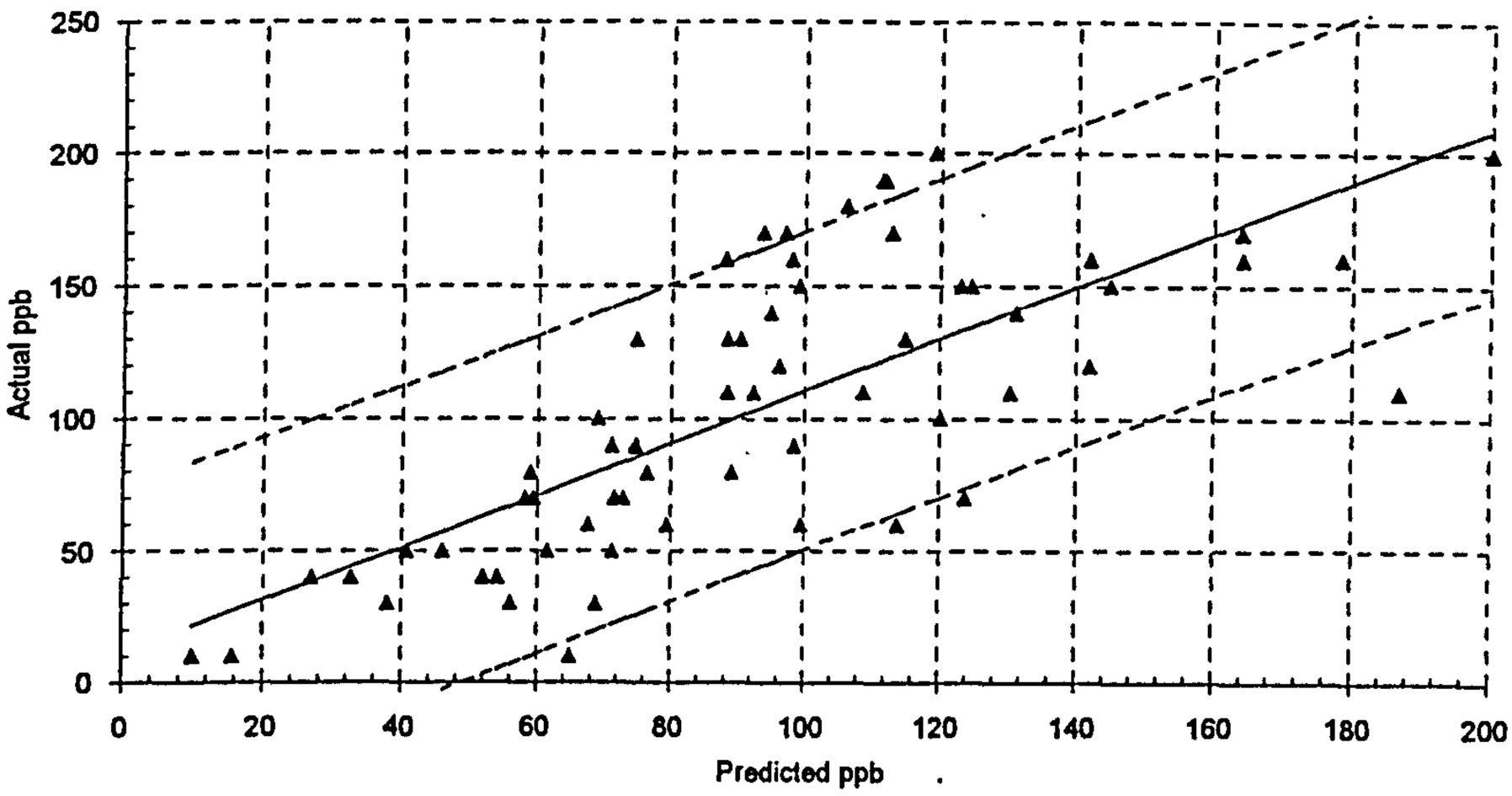



Figure 60. Test set calibration using a 19 input perceptron trained on non saturating data

### 6.1.5.Prediction from 22 input, two layer perceptron

Increasing the number of inputs to include the maximum, minimum and standard deviation of the points in the original voltammogram produced the plot shown in figure 61 and table 19.

| Avg Conf. (± ppb) | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) |
|---|---|---|---|---|---|---|
| 59.0 | 35.1 | 0.58 | 25.8 | 90.0 | 0.0 | 24.8 |

Table 19. Tabulated results of 22:0:1 neural network

The average 95% prediction confidence is slightly better than both the 19 input, two layer network and the PLS regression but still not as good as using the peak height regression.

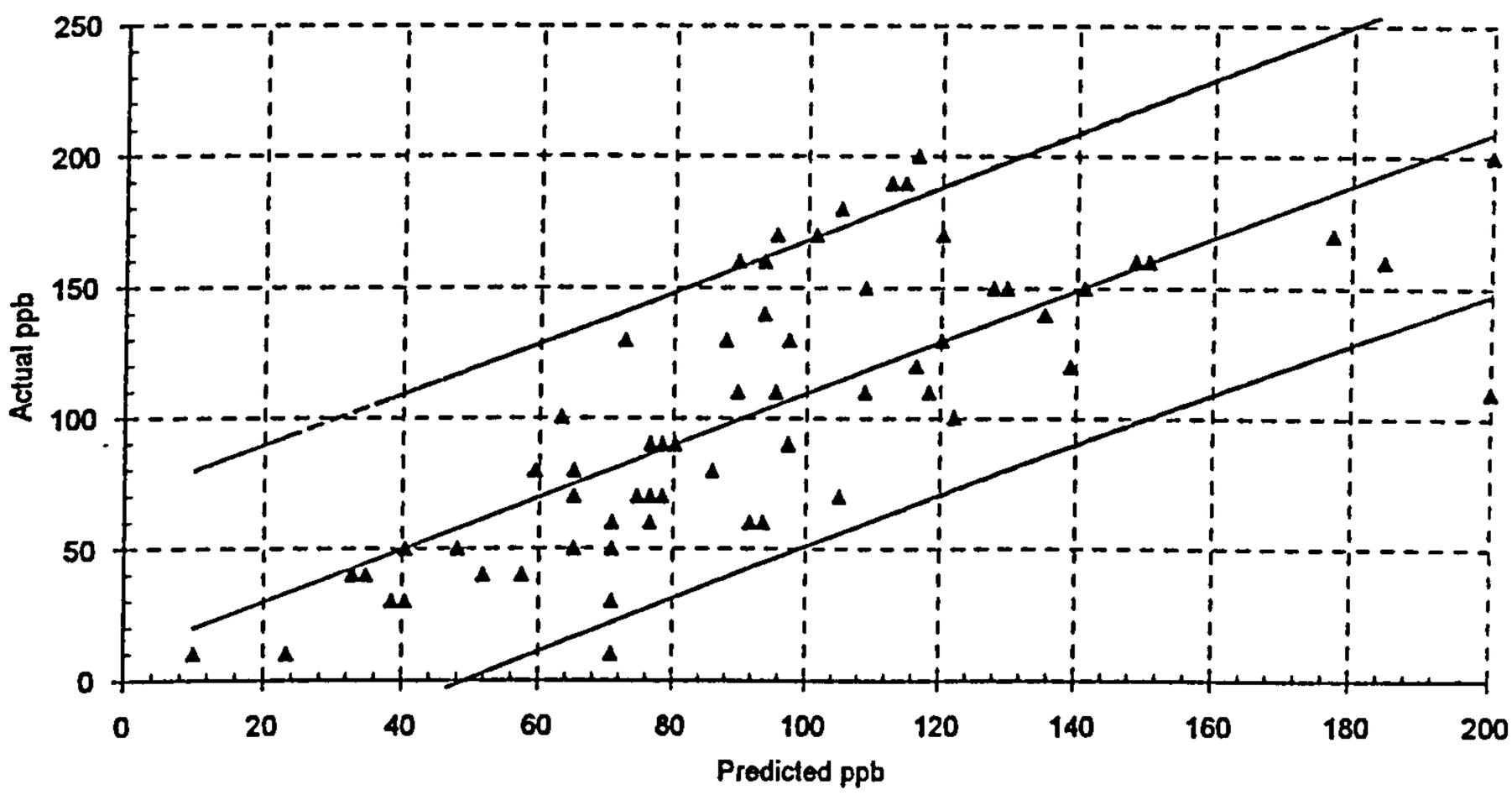



Figure 61. Test set calibration using a 22 input perceptron trained on non saturating data

### 6.1.6.Prediction from 138:0:1 two layer high order perceptron

An attempt to provide more information to the network by providing it with combinations of inputs instead of the raw inputs produced the 138:0:1 network for which the calibration plot is shown in figure 62. Unfortunately the average 95% prediction confidence achieved by this network is the worst of all the calibration methods attempted. Only 68.8% could be achieved with this architecture, see figure 62 and table 20.

| Avg Conf. (± ppb) | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) |
|---|---|---|---|---|---|---|
| 68.8 | 50.1 | 0.43 | 42.0 | 150.1 | 0.0 | 28.9 |

Table 20. Tabulated results of 138:0:1 high order neural network

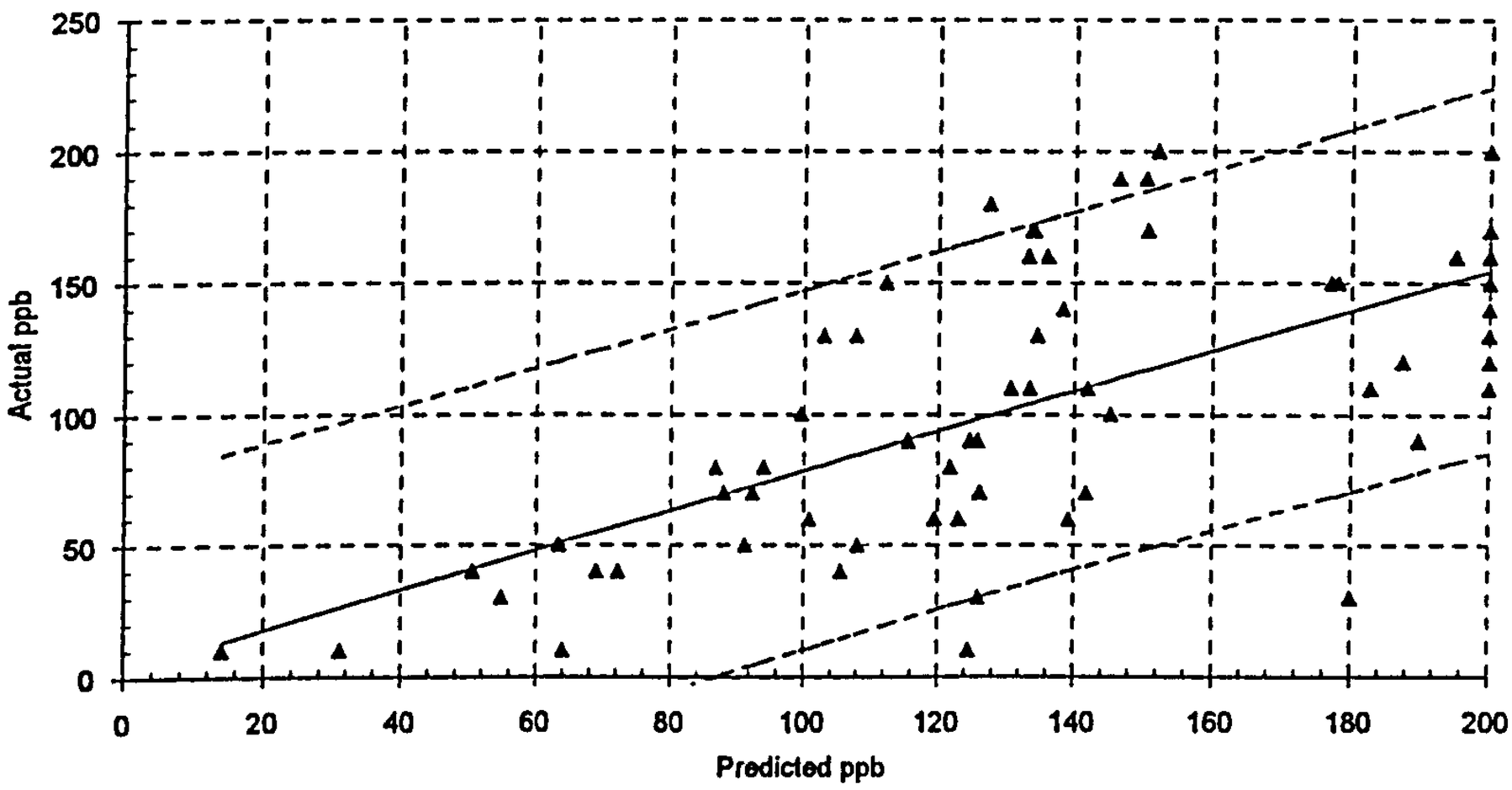



Figure 62. Test set calibration using a 138 input, two layer, high order perceptron trained on non saturating data

The following sections deal with three layer networks that exhibited considerably better performance than the models reported above.

### 6.1.7.Prediction using the 19:194:1 GRNN

After the success of the GRNN used to process the data from digested curry powder the first multilayer neural architecture studied, a 19 input GRNN trained using Euclidean distance metric proved somewhat disappointing. The calibration curve for this model is shown in figure 63 and table 21.

| Avg Conf. (± ppb) | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) |
|---|---|---|---|---|---|---|
| 31.4 | 18.6 | 0.88 | 12.7 | 85.3 | 0.0 | 14.1 |

Table 21. Tabulated results of 19:194:1 GRNN

The average 95% prediction confidence obtained using this data was approximately half that obtained by either PLS or either of the two layer neural networks, indicating the superior modelling ability of multi layer networks. However the results of this network are still only of the same order as predictions made using the base method of peak height regression.

Some encouragement must be drawn from this result however, as it indicates that the problem is not linearly separable in the original input space and that a multi layer network might therefore be able to extract more information from the curve than using models based upon peak height alone. It should be remembered at this point that the data vector being processed is a vastly reduced version of the vector from which the peak height was calculated (See Figure 55, page167)

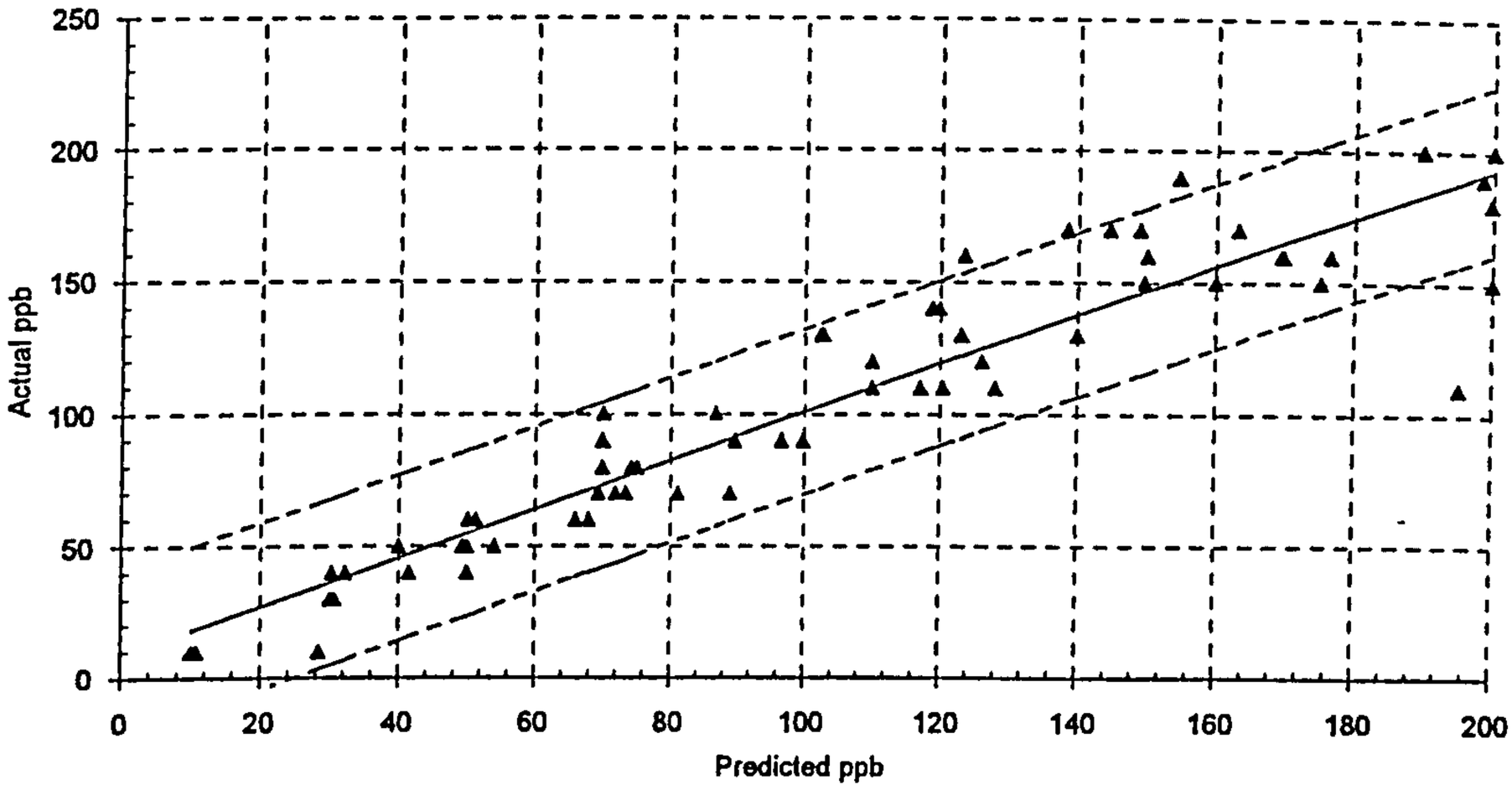



Figure 63. Test set calibration using a 19 input GRNN trained on non saturating data using the Euclidean distance metric

### 6.1.8.Prediction using the 22:194:1 GRNN

Increasing the number of inputs to the GRNN to 22 by including the additional information described earlier produced a calibration curve that was slightly better than that obtained by the peak heights. Figure 64 and table 22 illustrate the calibration.

| Avg Conf. (± ppb) | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) |
|---|---|---|---|---|---|---|
| 27.5 | 16.2 | 0.91 | 12.5 | 50.0 | 0.0 | 10.8 |

Table 22. Tabulated results for 22:194:1 GRNN

It is interesting to note however that whilst this model appears to perform slightly better than using peak heights, inspection shows that the points representing predicted concentration are more scattered at the extremes of the concentration range than they are for the latter method.

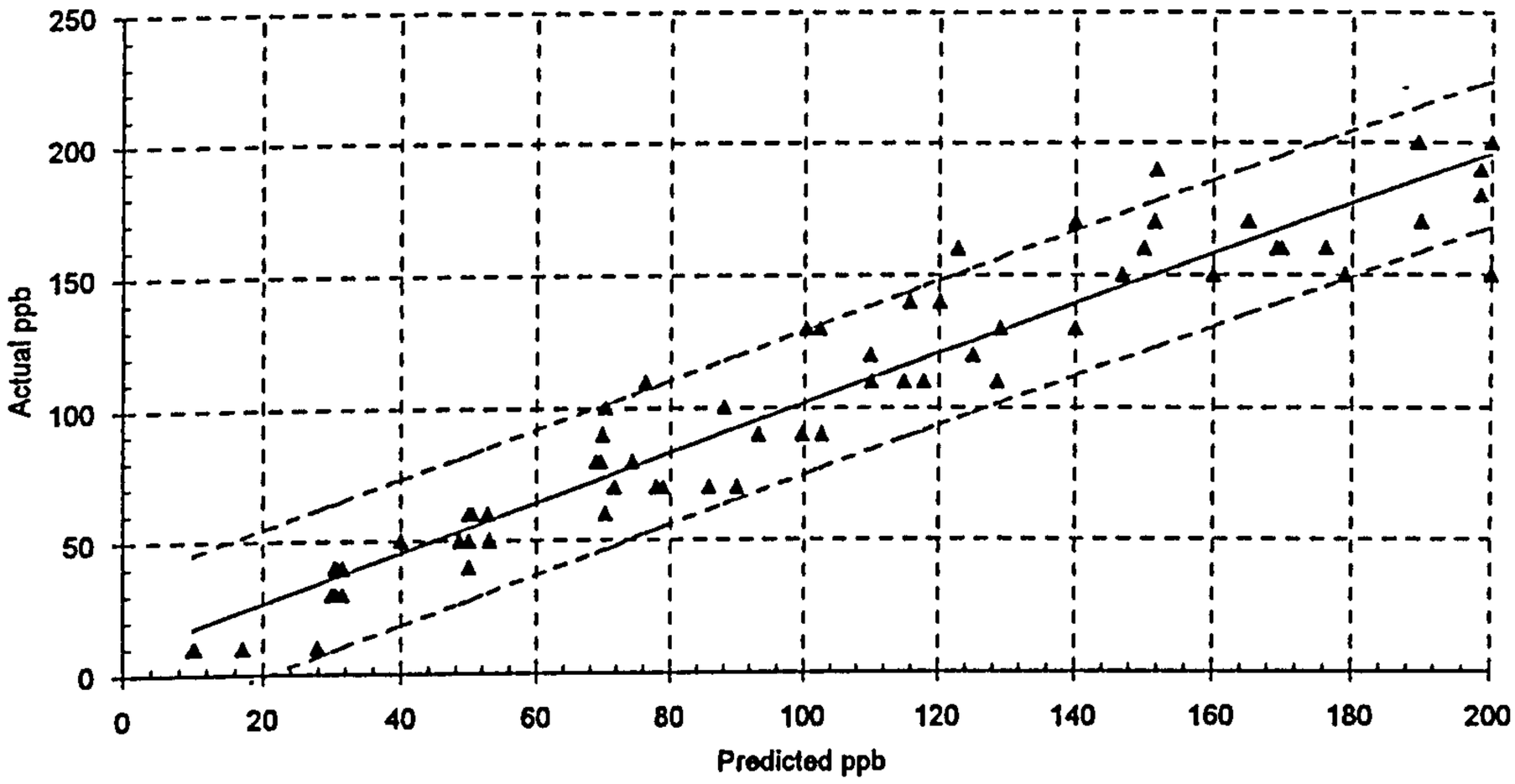



Figure 64. Test set calibration using a 22 input GRNN trained on non saturating data using the Euclidean distance metric

## 6.1.9.Prediction from Multi Layer Perceptrons

### 6.1.9.1. Optimisation of Number of Hidden Units

The same 22 input non saturating data used to find the optimum number of factors for PLS was used here to investigate the effect of varying the number of hidden units in a three layer MLP in an attempt to optimise that model. Using 22 inputs and the usual single output, 10 different MLPs having hidden layers containing from 1 to 45 sigmoidal processing units were trained on the non

saturating standard lead curves. The output unit had a linear transfer function. After each training epoch the test set was presented (without weight updates) and the mean squared error recorded. Training continued until the mean squared error on the test set showed an obvious rising trend from its minimum value. The minimum test set MSE reached for that architecture was recorded. The entire process was then repeated twice using a different set of random starting weights each time and the results for each architecture averaged.

| Number of Hidden Units | Average Minimum MSE for the test set |
|---|---|
| 1 | 0.032 |
| 5 | 0.028 |
| 10 | 0.014 |
| 15 | 0.011 |
| 20 | 0.012 |
| 25 | 0.011 |
| 30 | 0.008 |
| 35 | 0.012 |
| 40 | 0.012 |
| 45 | 0.015 |

Table 23. MSE reached by a back propagation MLP with different numbers of hidden units

The results, shown in table 23 and figure 65, indicate that 30 hidden units are optimal for this problem as that number produces the lowest MSE on the test set.

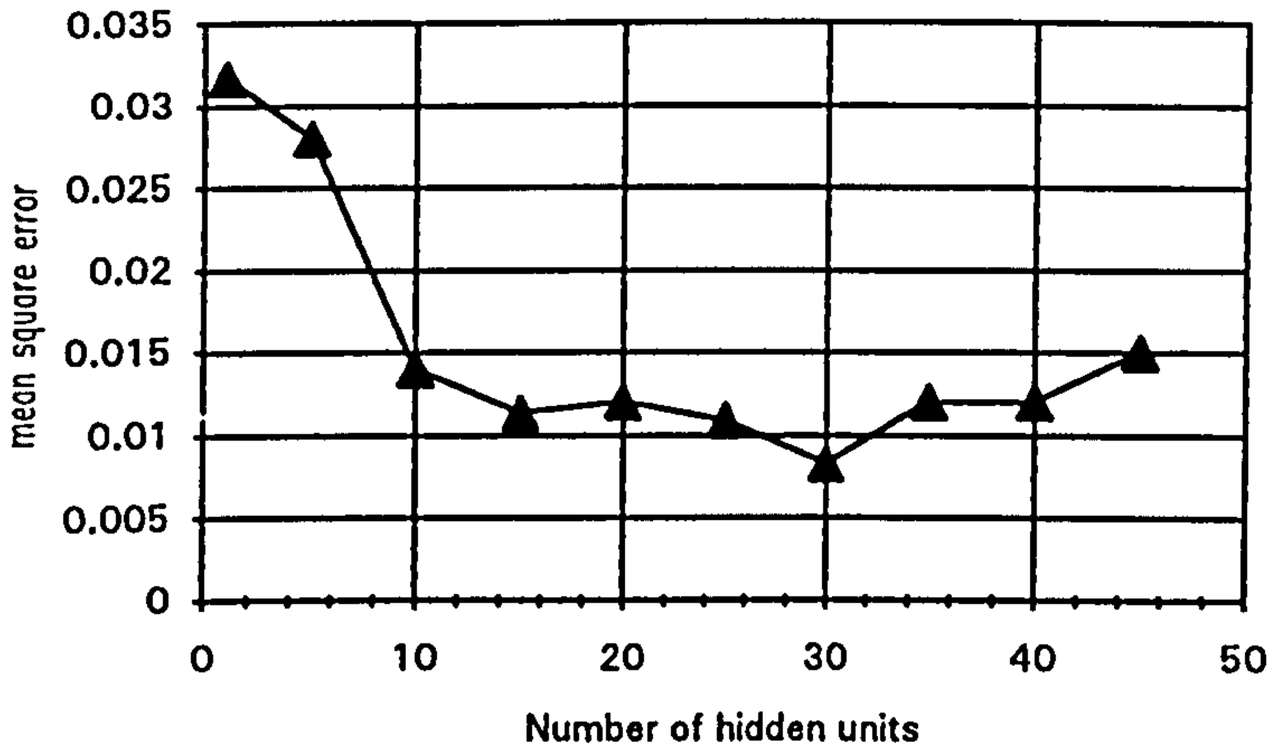



Figure 65. Plot of MSE reached against number of hidden units for a back propagation MLP

It can be concluded that a network with fewer than 10 hidden units is not able to model adequately the training set data and those networks with many more than 35 are over fitting the same data. The optimum number of hidden units, here taken as 30, does not appear to be quite so critical as the optimum number of factors to retain when using PLS. Adequate performance was obtained with any number between 15 and 40. For the remainder of this work however MLP networks having 30 hidden units were used.

### 6.1.9.2. Effect of initial starting weights

At this point the effect on the final mean squared error of changes in the initial network weights was investigated. The network used for this test was the 22:30:1 network training on the non saturating curves. Eight different networks were trained, each with a different starting seed for the random number generator used to initialise the weights. Each was trained until the MSE had reached its minimum. Sigmoidal hidden units were used with a linear output unit.

| Seed | MSE reached |
|---|---|
| 1 | 0.010 |
| 2 | 0.011 |
| 3 | 0.010 |
| 4 | 0.011 |
| 5 | 0.010 |
| 6 | 0.010 |
| 7 | 0.013 |
| 8 | 0.014 |

Table 24. MSE reached by a 22 Input MLP for different starting weights

The results, shown in Table 24 have a mean of 0.0112 and a standard deviation of 0.0016. This seems to indicate that the global minimum had been reached by these networks and that the particular set of starting weights used was not acting as a barrier to effective training.

### 6.1.10. Prediction using the 19:30:1 multi layer perceptron

The MLP's trained using back propagation performed much better than any of the other models, producing average 95% prediction confidence limits approximately half as wide as those of the next best model, the 22:194:1 GRNN.

The calibration plot for the 19:30:1 model is shown in figure 66 and the results in table 25. Note that in order to provide an easy basis for comparison, the abscissa in figure 66 has been retained at an upper limit of 200 ppb. The

predictions for the two actual concentrations of 200 ppb in the test set were in fact very slightly above this figure and hence do not appear on the plot. These predictions were 202 and 204 ppb respectively (with predictions rounded to the nearest ppb).

| Avg Conf. (± ppb) | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) |
|---|---|---|---|---|---|---|
| 16.7 | 10.0 | 0.97 | 8.4 | 26.9 | 0.4 | 5.7 |

Table 25. Tabulated results for 19:30:1 neural network

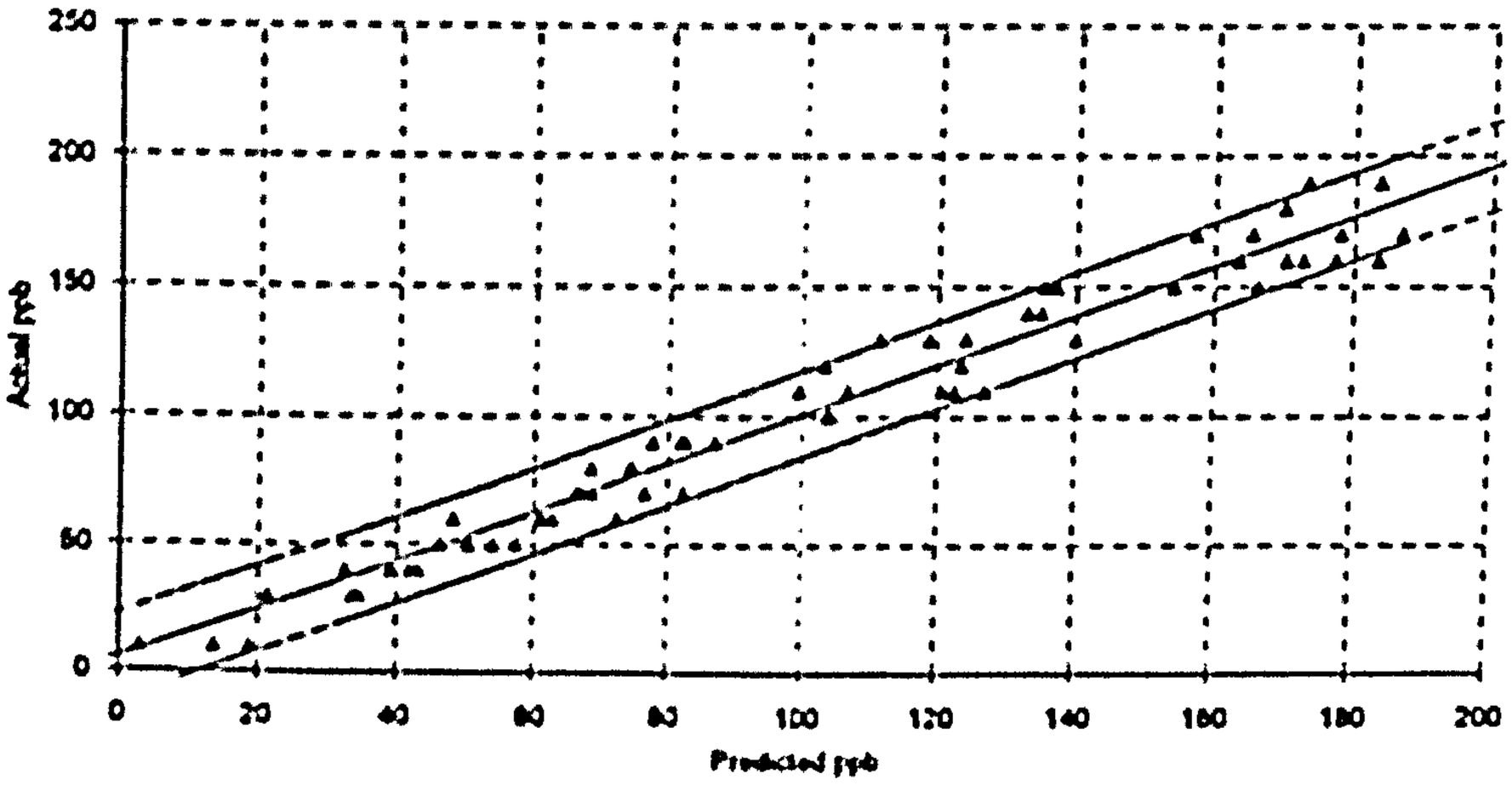



Figure 66. Test set calibration using a 19 input back propagation MLP trained on non saturating data

The 19:30:1 network gave nearly the best performance of all the models and is much better than using peak heights alone.

### 6.1.11. Prediction using the 22:30:1 multi layer perceptron

The best predictions of any model studied were achieved by the 22:30:1 multilayer perceptron. This network achieved ±16.1 ppb prediction confidence limits with a %SEP of 9.4 %, a considerable improvement over using peak heights, even when only 22 Δi data points are taken into consideration.

The results for the network are shown in table 26 and figure 67, where the abscissa has again been retained at an upper limit of 200. The two predictions for 200 ppb were 202 and 203 ppb.

| Avg Conf. (± ppb) | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) |
|---|---|---|---|---|---|---|
| 16.1 | 9.4 | 0.97 | 7.9 | 20.6 | 0.0 | 5.5 |

Table 26. Tabulated results for 22:30:1 neural network

If figure 67 is compared to figure 57, duplicated here, which showed the 'base' predictions obtained from a peak height calibration, the narrowness of the 95% confidence limits obtained by the neural network and the consequent improvement in concentration predictions are obvious.

The predictive data from the non saturating voltammograms represent the first important result of this work. It shows that neural networks can offer a considerable advantage over traditional methods of producing calibration curves from voltammograms.

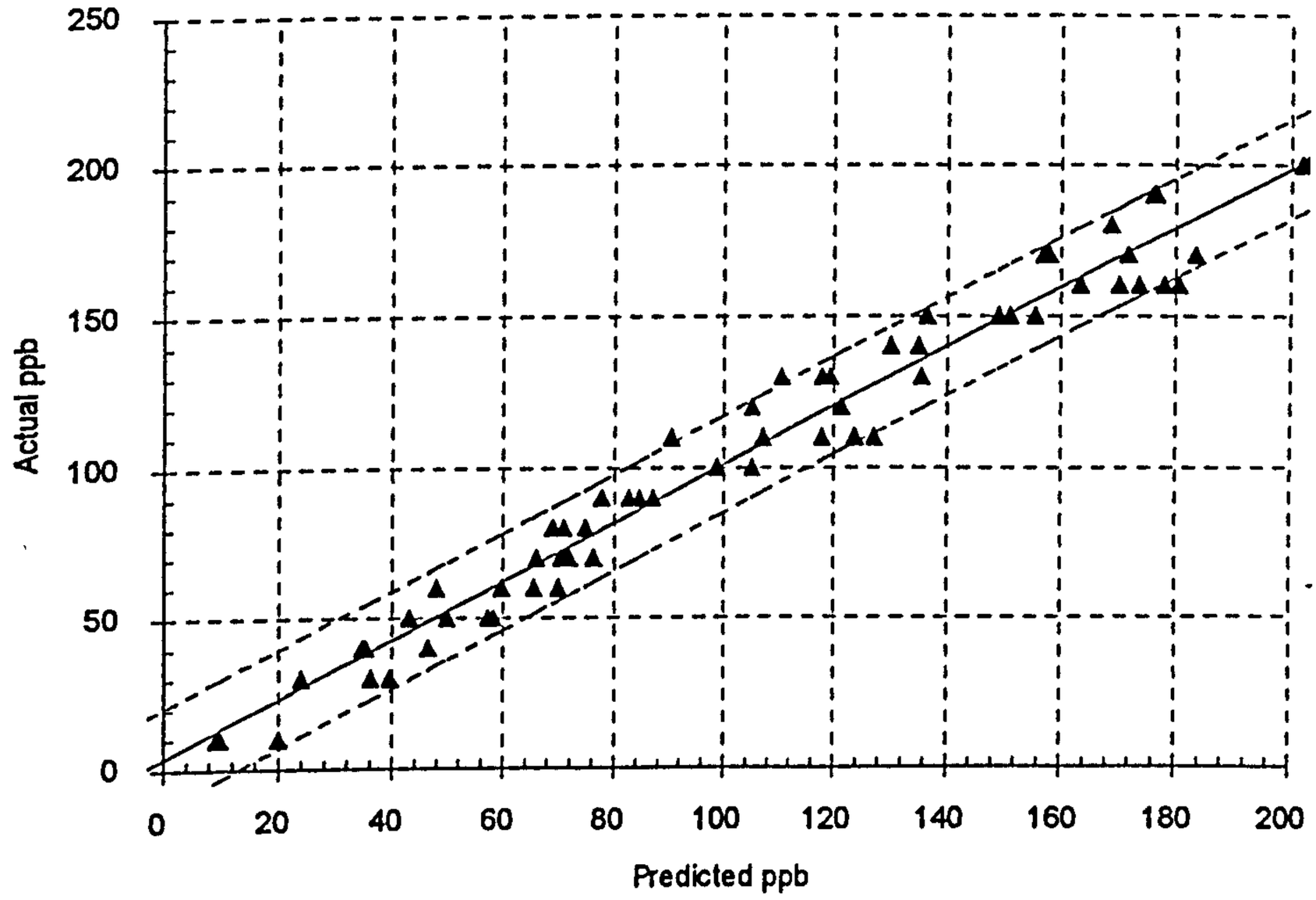



Figure 67. Test set calibration using a 22 input back propagation MLP trained on non saturating data

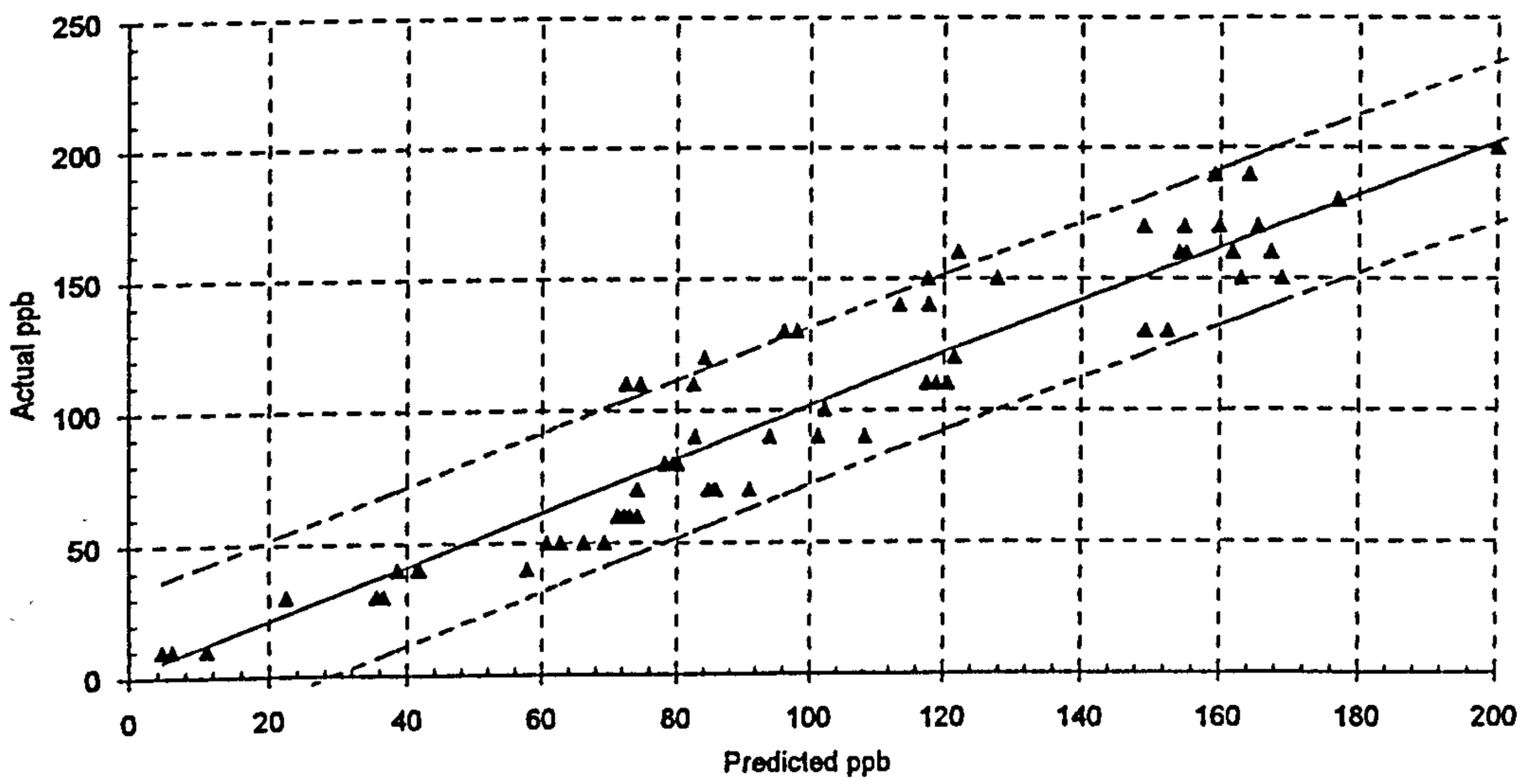



Figure 68. Copy of figure 57, Calibration of lead standards using peak heights

## 6.2. Saturating and Non Saturating Voltammograms

Saturating voltammograms were produced for a range of different concentrations as a low concentration could be made to produce a saturated voltammogram if the pulse height and deposition time were large whereas a high concentration sample would start to produce saturated voltammograms at low pulse heights and deposit times.

The 16 saturating vectors in the test set comprising all the data had target concentrations of between 50 and 200 ppb and were roughly evenly spread over this range, see table 27. The 54 saturating vectors in the training set showed a similar spread from 40 ppb to 200 ppb.

| target ppb | 50 | 60 | 70 | 90 | 100 | 110 | 120 | 130 | 150 | 160 | 160 | 160 | 170 | 180 | 190 | 200 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| pulse height | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 1 | 3 | 3 |
| deposition time | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 20 | 20 | 20 | 50 | 50 | 20 | 50 | 50 | 50 |

Table 27. Parameters of saturating voltammograms in the combined test set

Networks with no hidden layer were not used with this data, nor was it applicable to attempt a peak height calibration as the peak heights could not be ascertained for all the data. The only networks used therefore were the GRNN and the MLP, each with either 19 or 22 inputs.

### 6.2.0.1. Prediction using the 19:229:1 and 22:229:1 GRNNs

It was expected that all network models would perform poorly on this data due to wide variation in the vectors from saturating and non saturating voltammograms with the same target concentration.

However, both of the GRNN networks, with 19 or 22 inputs, performed almost as well with this data as they had with data containing no saturating vectors. The 19 input network produced an average 95% prediction confidence of ± 24.9 ppb and the 22 input produced ± 26.4 ppb. Both still out performed the peak height regression even when this was using only non saturating data.

The results are grouped here in table 28 as the values were very similar.

| Avg Conf. (± ppb) | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) | |
|---|---|---|---|---|---|---|---|
| 24.9 | 14.2 | 0.93 | 11.3 | 52.9 | 0.0 | 10.0 | 19 input GRNN |
| 26.4 | 15.0 | 0.92 | 12.0 | 61.8 | 0.0 | 10.5 | 22 input GRNN |

Table 28. Tabulated results for 19 and 22 input GRNN on all saturating and non saturating data

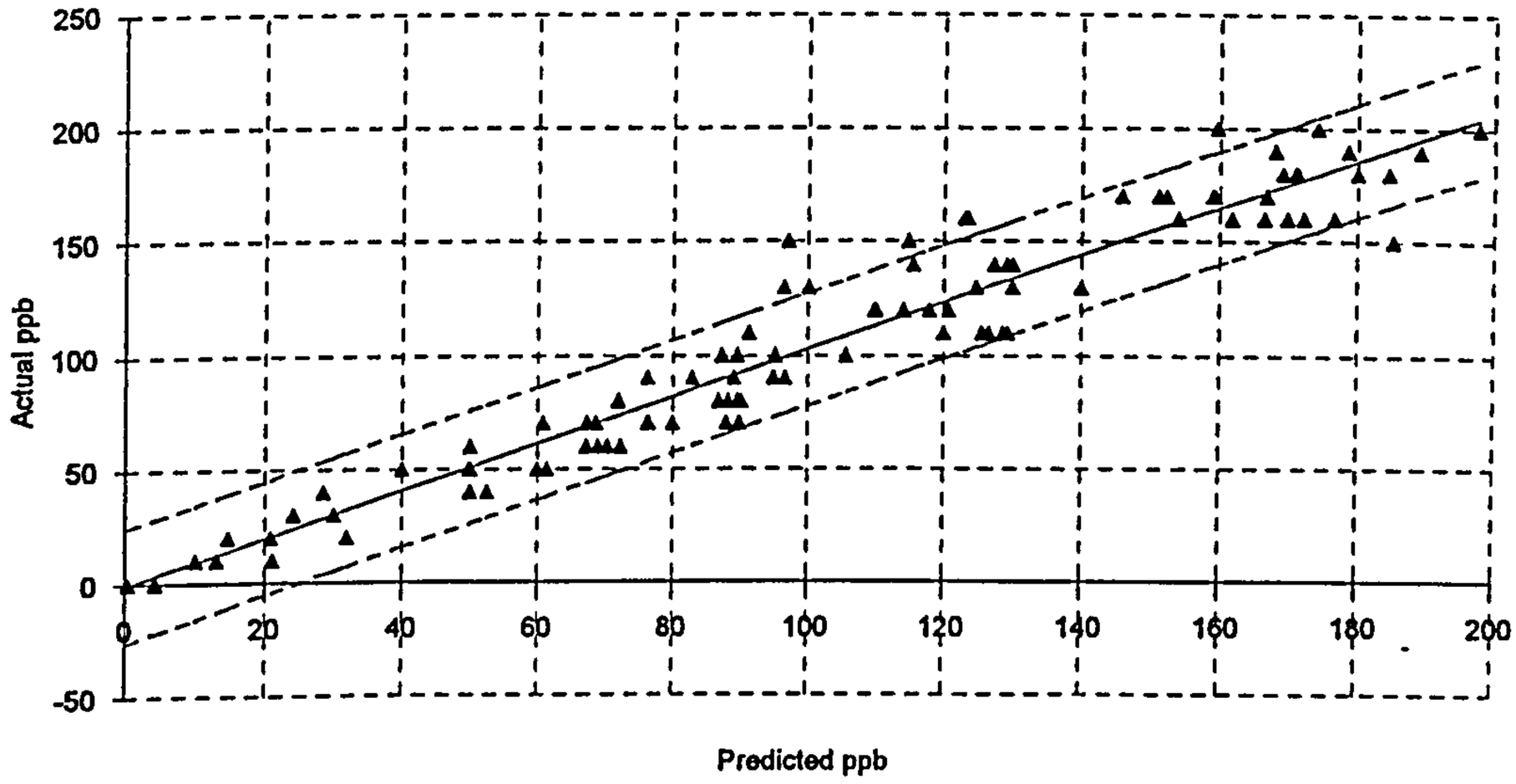



Figure 69. Test set calibration using a 19 input GRNN trained on both saturating and non saturating data

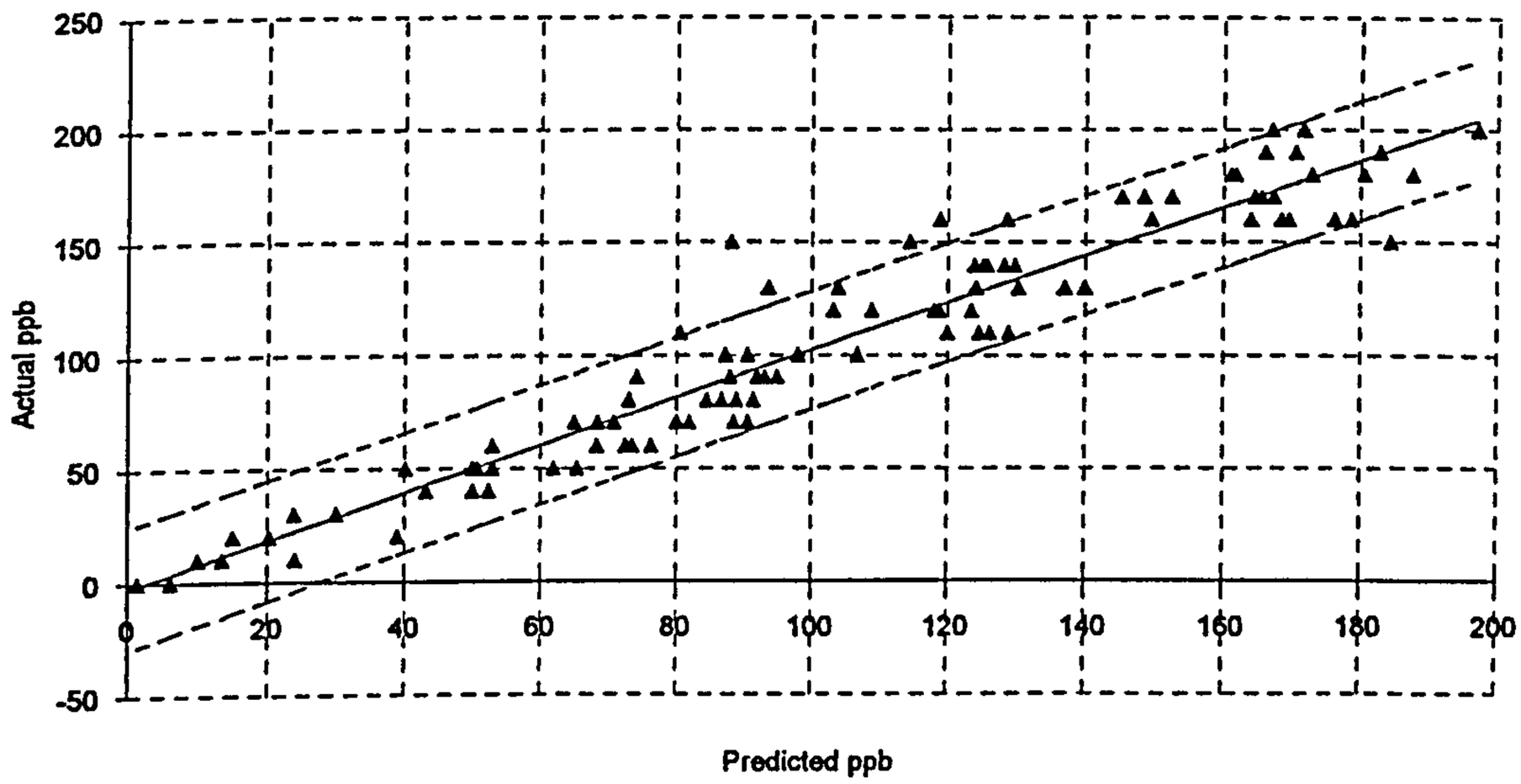



Figure 70. Test set calibration using a 22 input GRNN trained on both saturating and non saturating data

### 6.2.0.2. Prediction using the 22:30:1 MLP

The 22 input MLP network performed slightly worse on this data than it did on non saturating data. However with a mean 95% confidence of ± 23.9 ppb it is still better than that produced by the both the GRNN and the peak heigiit models.

| Avg Conf. (± ppb) | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) |
|---|---|---|---|---|---|---|
| 23.9 | 13.3 | 0.93 | 10.8 | 45.8 | 0.0 | 9.2 |

Table 29. Tabulated results of 22:30:1 neural network on saturating and non saturating data

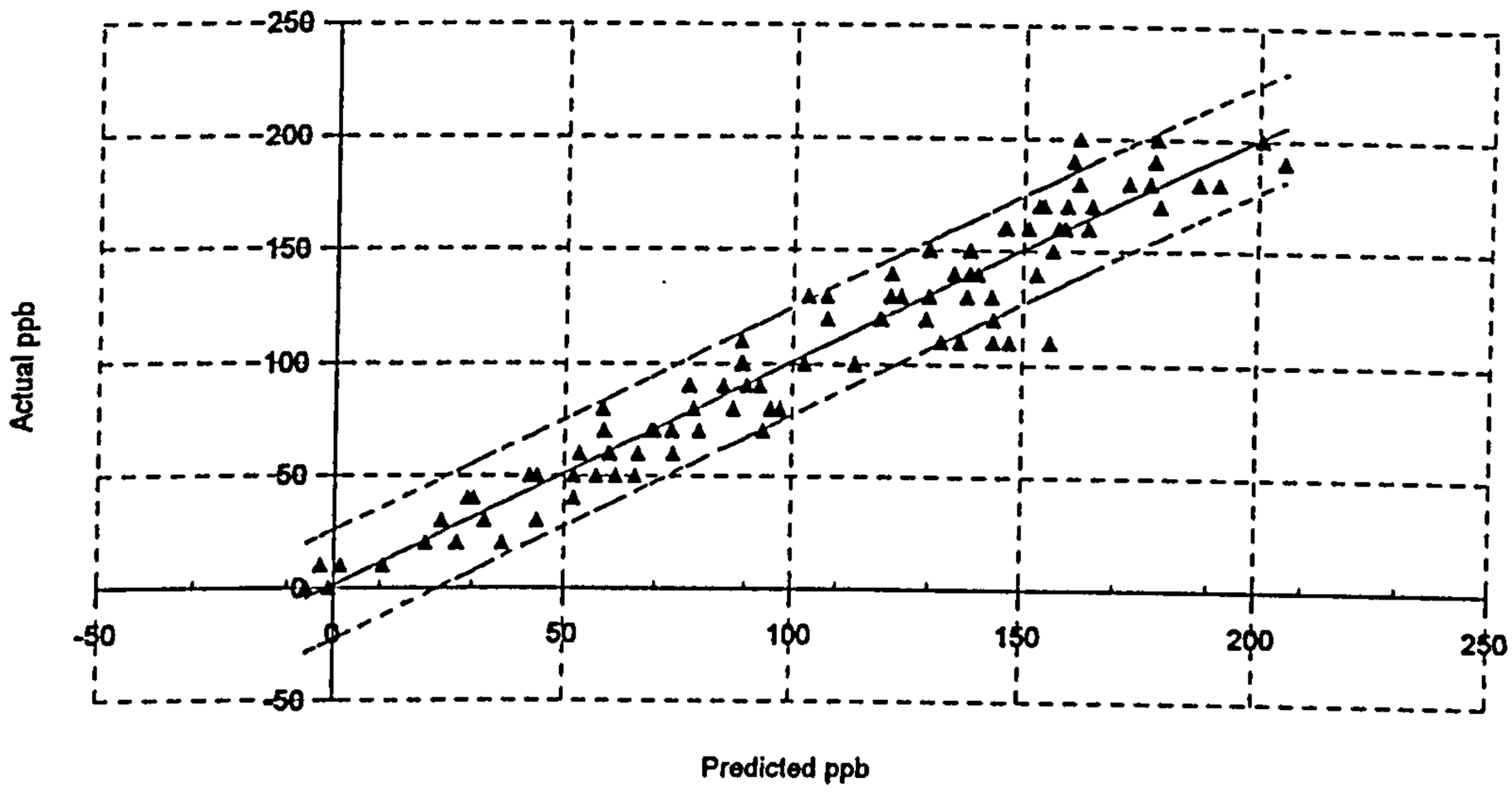



Figure 71. Test set calibration using a 22 input back propagation MLP trained on both saturating and non saturating data

### 6.2.0.3. Prediction using the 19:30:1 MLP

Although it was presented with less total information on its inputs, the best network model for this data was the 19 input MLP. This network performed almost as well as it had done with the non saturating data, producing an average 95% prediction confidence of ± 18.4 ppb, roughly half as wide as using peak heights on non saturated data.. Mean and standard deviation of the absolute errors also compare well with that model.

| Avg Conf. (± ppb) | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) |
|---|---|---|---|---|---|---|
| 18.4 | 10.3 | 0.96 | 8.3 | 29.8 | 0.0 | 7.2 |

Table 30. Tabulated results of 19:30:1 neural network on saturating and non saturating data

The calibration plot for this model is shown in figure 72. It can be seen that it is not possible to distinguish those points which were predicted by processing saturating voltammograms from those arising out of voltammograms in which the stripping peak was well defined.

A possible explanation for the 19 input MLP performing better than the 22 input MLP with this data is that one of the inputs in the 22 input model is a normalised version of the maximum value in the original voltammogram. The value of this maximum in saturating voltammograms remains constant at 42.8 μA regardless of the concentration in the original sample. Hence, by providing the

network with an input that remains as the same value for different target concentrations, this input is sometimes providing conflicting data.

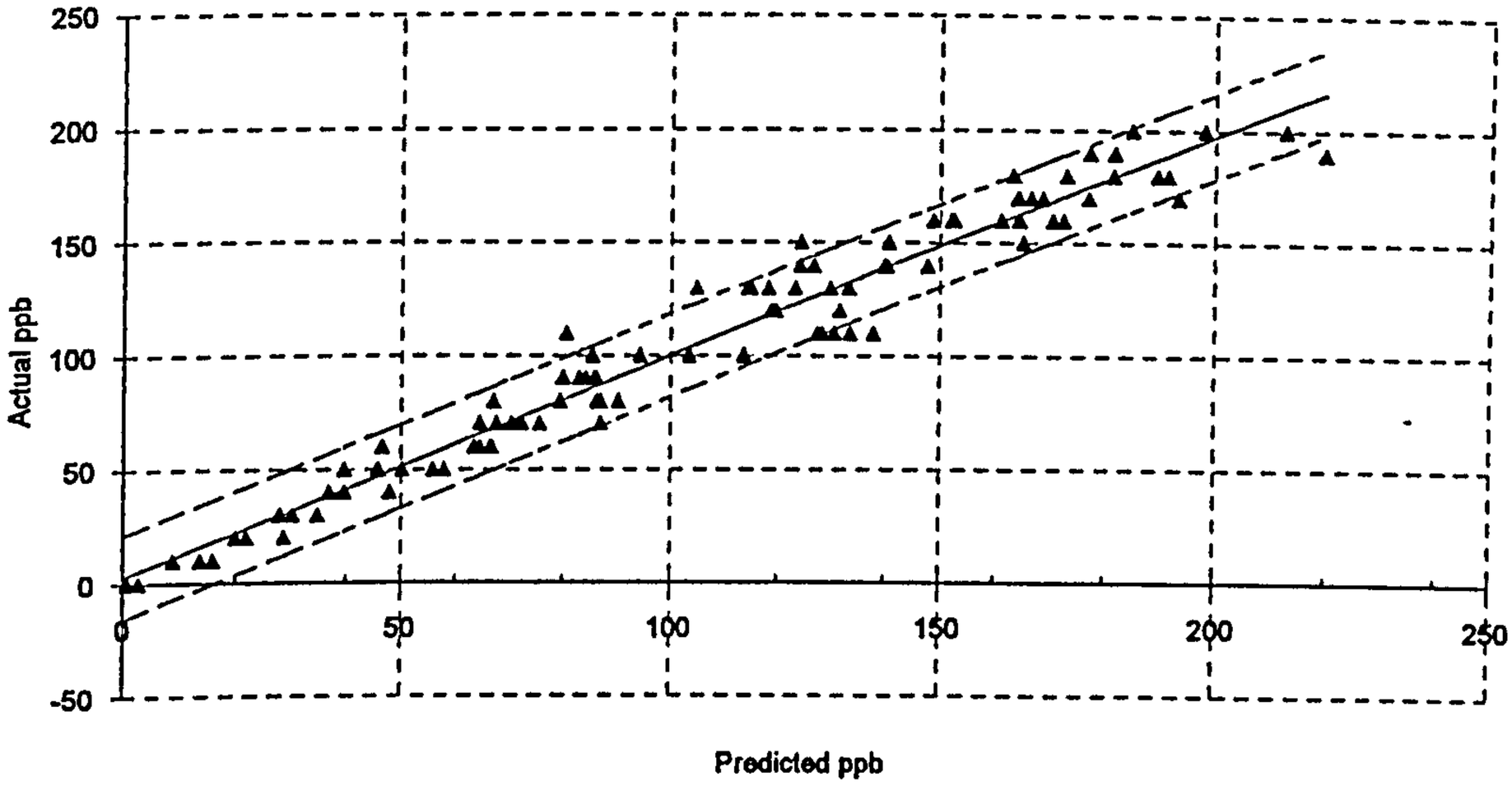



Figure 72. Test set calibration using a 19 input back propagation MLP trained and tested on both saturating and non saturating data

The input that provided a normalised standard deviation of the original data points would still be providing a range of different values for the different vectors. These values would however be distorted in saturating vectors, tending to be lower than they would be if the true maximum of the peak were available.

The fact that neural network models are able to make predictions from data that includes saturating voltammograms represents the second important result of this work. Providing a prediction for lead concentration from voltammograms that may or may not saturate the equipment releases the analyst from having to be too careful about the choice of operational parameters used during the scan.

## 6.3. Saturating Voltammograms Only

To complete this work on the lead standards an attempt was made to predict the lead concentration by training and testing networks on data that was entirely produced from saturating voltammograms. This has little practical application since one would not deliberately saturate the equipment but the results are included here as they represent a third significant aspect of this work.

The test set included a representative sample of all the saturating vectors. The parameters of the ones selected are given in table 31.

| Pulse Height (mV) | Deposit time (minute) | Concen-tration (ppb) |
|---|---|---|
| 20 | 3 | 70 |
| 50 | 3 | 70 |
| 20 | 3 | 120 |
| 50 | 3 | 120 |
| 20 | 3 | 140 |
| 20 | 3 | 150 |
| 50 | 3 | 150 |
| 20 | 3 | 170 |
| 20 | 3 | 170 |

| Pulse Height (mV) | Deposit time (minute) | Concen-tration (ppb) |
|---|---|---|
| 50 | 3 | 170 |
| 50 | 1 | 170 |
| 20 | 3 | 180 |
| 50 | 3 | 180 |
| 50 | 1 | 180 |
| 20 | 3 | 190 |
| 20 | 3 | 190 |
| 50 | 3 | 190 |
| 50 | 1 | 200 |

Table 31. Parameters of test set used for calibration with just saturating voltammograms

As with the combined test set, only the GRNN and three layer, back propagation neural network models were used with this data. The average 95%

prediction confidence limits in predicting a lead concentration from calibration curves produced in this way showed little difference between the different network models used. When placed in order of predictive ability they did nevertheless show the same pattern as that displayed when used on data containing all the vectors. The worst performer was the 19 input GRNN and the best was the 19 input back propagation MLP. The prediction confidence limits were also similar to those obtained using the combined data, ranging from ± 20.4 ppb for the 19 input GRNN to ± 16.5 ppb for the 19 input MLP.

The calibration plots are visually rather similar. Consequently for reasons of space only the plot for the 19 input MLP is reproduced here in figure 73. The tabulated results of all four network models are given in the summary in table 34.

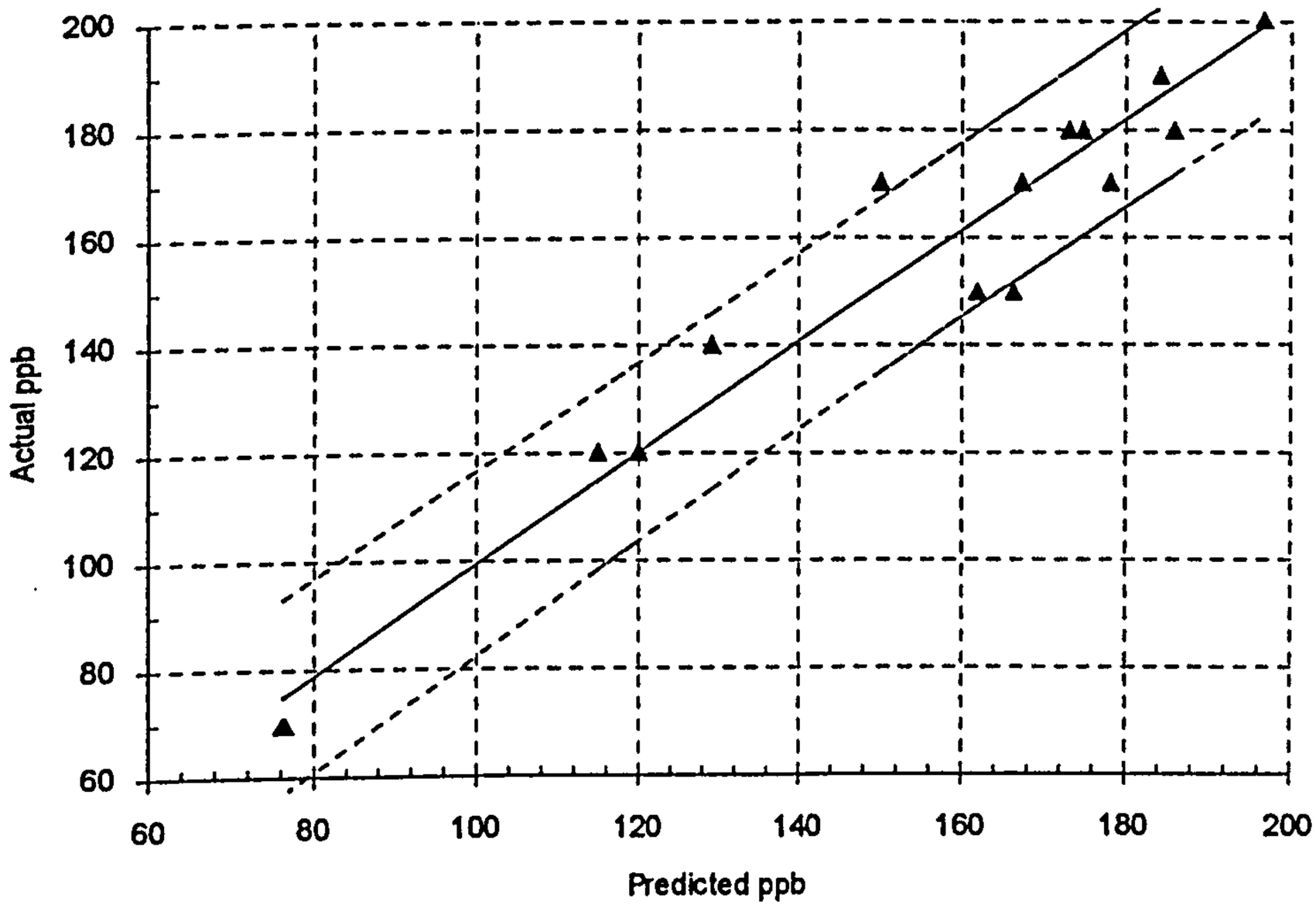



Figure 73. Test set calibration using a 19 input back propagation MLP trained and tested on just saturating data

## 6.4. Lead Standards Results Summary

All the results from work on the lead standards are summarised in tables 32 to 34. In these tables the rows are sorted in order of descending average 95% prediction confidence. It can be seen that every three layer neural network is a better predictor of the actual lead concentration than obtained using peak heights alone.

| Network / Calibration model | Avg Conf. (± ppb) | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) |
|---|---|---|---|---|---|---|---|
| BP 138 input, two layer, high order neural network | 68.8 | 50.1 | 0.43 | 42.0 | 150.1 | 0.0 | 28.9 |
| 22 input, 9 factor PLS | 64.8 | 37.2 | 0.50 | 30.3 | 84.3 | 2.7 | 22.9 |
| 19 input, two layer neural network | 60.5 | 36.2 | 0.56 | 27.6 | 80.8 | 0.0 | 24.5 |
| 22 input, two layer neural network | 59.0 | 35.1 | 0.58 | 25.8 | 90.0 | 0.0 | 24.8 |
| 19 input GRNN | 31.4 | 18.6 | 0.88 | 12.7 | 85.3 | 0.0 | 14.1 |
| Regression on training set heights | 30.0 | 16.8 | 0.89 | 13.8 | 37.9 | 0.1 | 10.4 |
| 22 input GRNN | 27.5 | 16.2 | 0.91 | 12.5 | 50.0 | 0.00 | 10.8 |
| 19 input, three layer neural network (Back Prop) | 16.7 | 10.0 | 0.97 | 8.4 | 26.9 | 0.4 | 5.7 |
| 22 input, three layer neural network (Back Prop) | 16.1 | 9.4 | 0.97 | 7.9 | 20.6 | 0.0 | 5.5 |

Table 32. Results summary - Non saturated data sets

| Network<br><br>(Peak height regression shown for comparison) | Avg Conf. (± ppb) | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) |
|---|---|---|---|---|---|---|---|
| **Regression on training set heights** (for non saturating) | 30.0 | 16.8 | 0.89 | 13.8 | 37.9 | 0.1 | 10.4 |
| **19 input GRNN** | 24.9 | 14.2 | 0.93 | 11.3 | 52.9 | 0.0 | 10.0 |
| **22 input GRNN** | 26.4 | 15.0 | 0.92 | 12.0 | 61.8 | 0.0 | 10.5 |
| **22 input, three layer neural network (Back Prop)** | 23.9 | 13.3 | 0.93 | 10.8 | 45.8 | 0.0 | 9.2 |
| **19 input, three layer neural network (Back Prop)** | 18.4 | 10.3 | 0.96 | 8.3 | 29.8 | 0.0 | 7.2 |

Table 33. Results summary - saturated and non saturated data sets

| Network<br><br>(Peak height regression shown for comparison) | Avg Conf. ± ppb | % SEP | R^2 | Mean Abs Err (ppb) | Max Abs Err (ppb) | Min Abs Err (ppb) | S. D. Abs Err (ppb) |
|---|---|---|---|---|---|---|---|
| **Regression on training set heights** (for non saturating) | 30.0 | 16.8 | 0.89 | 13.8 | 37.9 | 0.1 | 10.4 |
| **19 input GRNN** | 20.4 | 7.7 | 0.92 | 9.3 | 26.2 | 0.0 | 7.4 |
| **22 input GRNN** | 21.3 | 7.7 | 0.91 | 9.6 | 24.5 | 0.0 | 7.5 |
| **22 input, three layer neural network (Back Prop)** | 18.3 | 6.5 | 0.93 | 8.7 | 19.6 | 0.2 | 5.1 |
| **19 input, three layer neural network (Back Prop)** | 16.6 | 5.53 | 0.95 | 7.2 | 20.0 | 0.2 | 4.9 |

Table 34. Results summary - Only saturated data sets

# 7. Discussion of Results

## 7.1. Lead Standards Error Analysis

There are a number of areas in this work where errors were introduced. An indicative list is given below followed by a discussion on the effect that these errors may have had on the results previously given.

a. The target values for all the models in this work were the assumed concentrations of lead in the standard solutions made up. These values are subject to error arising from a number of factors to do with the weighing out of the chemicals, the use of the volumetric glassware and the experience (or lack thereof) of the operator.

b. The potentiostat and associated electronics that scanned the analytes and produced the raw data also introduced error into the resulting voltammograms.

c. The method of peak height estimation used on the voltammograms as the basis of the comparative calibration was susceptible to the points at which the quadratics were fitted.

d. All the calibration models, peak height and neural network, contained errors due to the need to generalise from the training set rather than to fit it exactly.

Errors in area (c), the measuring of the peak height via software, were investigated by performing ten replicate measurements of the peak height for a single voltammogram produced from a 100 ppb lead standard. This was produced using a plate time of 1 minute and a pulse height of 50 mV. From these ten measurements the mean peak height found was 27.83 μA and the standard deviation of the measurements was 0.05 μA.

Errors in area (b), the instrumental errors, are difficult to isolate from those in inherent the measurement of the peak height. The method used here is to investigate the variation in peak height recorded for 10 replicate scans on the same sample of digested curry powder containing a nominal 45 ppb lead, scanned using a deposit time of 1 minute and a pulse height of 50 mV. This process assumes that the peak height measurement process is considerably more precise than the process producing the peak. The measured mean peak height obtained over the ten replicate scans with this data was 13.17 μA with a standard deviation of 0.64 μA .

Experiments confirmed that whilst a blank solution produces a background signal above 0 μA, a quadratic fitted to two intervals either side of where the lead peak would normally appear and a second one fitted to the remaining central interval, will coincide and record a measured peak height of 0.0 μA.

Thus, assuming a linear relationship between peak height and concentration, a height of 13.17 μA for 45 ppb indicates that with these instrumental parameters, neighbouring samples differing by 10 ppb would be separated by a difference in peak height of 2.9 μA . Since the measured peak height corresponding to any one concentration has a standard deviation of 0.64 μA, peaks from neighbouring concentrations would be separated by 4.5 standard deviations. This indicates that the instrument should be capable of producing peaks that are resolved into reasonably separated bands corresponding to each 10 ppb division.

The standard deviation of the peak height measurement process (0.05 μA), is small compared to the 0.64 μA above and has been ignored in this work.

Errors in area (a), the making up of the solutions, can be estimated by making up replicate samples of lead solution and measuring the lead concentration in the final analyte using a standard technique such as carbon furnace atomic absorption spectrometry. Alternatively, given the data above, an estimation of the error may be made from the variation in peak height about the least squares fit when the peak heights are plotted against the nominal lead concentration. A possible metric would be the % standard error calculated using the equation

$$\% \text{ error} = \frac{\sqrt{\sum e_i^2}}{\bar{y}} \times 100 \qquad \{26\}$$

where the error, $e_i$, is the difference between the desired concentration and the concentration calculated from the regression line using the measured peak height and $\bar{y}$ is the average of the desired concentrations used in the regression.

Calculating this for the voltammograms produced by a 1 minute plate and 5mV pulse height showed that solutions were being made up to within a standard error of 10.9% of the desired value.

The above discussion however concerns itself with the mechanisms producing the voltammograms and the measurement of the peak height necessary for the calibration based on heights. However, the initial set of data supplied to each of the calibration models, whether based upon peak height or a neural network, is the same. This data being the voltammograms themselves. On this basis the only error introduced from this point on is that in area (d), the error inherent in the modelling method itself. Thus the comparison of different modelling methods may be carried

out from the starting point of the voltammograms themselves, however they are produced.

It is from the observation of this modelling error, measured using $R^2$, %SEP, 95% confidence limits and statistics drawn from the absolute errors, that the conclusions of this thesis are drawn.

## 7.2. Summary of Main Observations

The following summarises the main observations made when performing differential pulse, anodic stripping voltammetry of trace level lead solutions using computerised instrumentation and applying neural network and traditional processing to the data obtained.

1. Several neural network paradigms are able to predict the lead concentration from either a complete anodic stripping voltammogram or one that has had the number of data point severely reduced.

2. From good, non saturating curves a backpropagation neural network is able to predict species concentration with a lower error and a higher confidence than the equivalent prediction obtained by peak height calibration or partial least squares. This applies even when the number of inputs to the network is very much less than the number of data points in the original voltammogram.

3. The backpropagation neural network is able to estimate species concentration with an acceptable error even from curves exhibiting saturation. These curves may be obtained as a result of using analytical parameters which, in hindsight, prove to have been inappropriate for the actual, high, concentration level in the solution.

4. Including the analytical parameters used during analysis as inputs to the neural network enables it to model a much wider range of instrument response / analyte concentration than is possible using a peak height

calibration at a single instrument setting. Even if different instrument settings are allowed for in the calibration by scaling of the peak height, the neural network model still shows superior performance.

5. The GRNN network outperformed the backpropagation network on voltammetric curves containing considerable non linearity due to artifacts in the data but both networks successfully modelled the data.

6. A fully computerised instrument is feasible, has been shown to work and can contain elements that render it able to be controlled by an external, optimising system, possibly driven by a neural network

7. Gray code, binary and decimal outputs may all be used with advantage. The Gray performs marginally better than binary. When a linear transfer function is used on the output with a a single decimal value encoding the target, predictions can be made with sufficient resolution to predict a concentration level to within 20 ppb for lead.

8. Neural networks can be used to estimate the concentration of several species simultaneously from a single simulated voltammogram. (Details of this auxiliary work are given in appendix 2)

# 8. Conclusions and Recommendations for Further Work

## 8.1. Final Conclusions and Justification of Aims

It has already been said that this work does not claim to push back the boundaries of fundamental research into neural networks, nor does it propose a similar claim for analytical chemistry. Nevertheless, this work represents novel research and an advancement of knowledge as it brings together two unrelated fields and describes a new application of neural networks, previously unreported in the literature.

One of the original aims of this work was to ascertain whether neural networks could be used to good effect in producing quantitative predictions of trace metal concentrations from data obtained by differential pulse stripping analysis. The data was originally intended to be extracted from good quality voltammetric scans that produced a clearly defined stripping peak.

This has been achieved. The major conclusions of this work are that even relatively uncomplicated neural networks are able to produce the required quantitative predictions.

Indeed, the neural network model is shown to be superior to a traditional method and able to process scans performed at a range of pulse heights and deposition times without the need to re calibrate the instrument.

A second aim of this work was to investigate a control mechanism whereby a neural network could identify a poor quality peak, for example one that exhibits saturation caused by inappropriate stripping parameters, and initiate a repeat of the scan using different values. The reason for doing so being to obtain a non saturating peak from which a prediction of analyte concentration could be made.

The results of this work show that, at least within a limited range, such a mechanism is unnecessary. It is shown that good predictions can be made by neural networks from scans that would have been of no use if processed by traditional methods. If the instrumental parameters used in the analytical scan are included in the neural network's input, good calibrations may be made from saturating and non saturating scans alike and valid quantitative predictions of concentration obtain from unseen test data obtained from scans that are similarly variable in quality.

Nevertheless, although postponed to a further period of investigation, the goal of controlling the analysis by neural network remains one worthy of research. The groundwork for achieving this goal is in place. The software developed in the earlier part of the research, and used to operate the computerised instrument, has been written in a way that facilitates further levels of control, directly or indirectly by other pieces of software. The results so far obtained seem to suggest that it would not be a difficult matter to train a neural network to detect inappropriate instrument settings and to have this control provided by the neural network outputting the optimum settings.

## 8.2. Statistical Justification for the Conclusions

The main conclusions of this work rest upon the evidence that neural network processing, in particular by the 22:30:1 backpropagation network, provides more precise calibrations than a calibration curve produced from peak heights alone. Accordingly a one tailed Student's t test was devised to ascertain the significance of the apparently lower 95% prediction interval obtained by MLP over peak height calibration and also the perceived lower mean absolute error in prediction. The cross validation procedure adopted was as follows.

The file containing data from unsaturated peaks was first randomised by making a single pass through the file. As each row of the data matrix was reached during the pass it was swapped with a row chosen at random from the entire file. After randomising the data, four cross validation training and testing data sets in the ratio 75% training 25% testing were extracted.

This was done by taking the vectors 1 to 64 to be the test data for set 0 with the remaining vectors going into the training set for set 0. Then the next 64 vectors , numbers 65 to 129, were extracted to go into the test set for set 1 with vectors 1 to 64 and 130 onwards going into the training set for set 1. This process was repeated until four sets were obtained and each vector had been in both the test set and the training set once.

Each of these four sets was then processed to obtain a prediction of test set lead concentration using (a) a least squares regression on the peak heights found in the original voltammograms and (b) a 22 input MLP trained using the same methods and training parameters as previously described.

Each of these eight sets of 64 lead predictions was then processed to obtain the 95% prediction limits in predicting the lead concentration, the average 95% prediction limit for each set and each method and the mean absolute error for each set and each method. These figures, along with other indicative values are given below in tables 35 and 36.

| Calibration from heights of training set | | | | | | |
|---|---|---|---|---|---|---|
| Set | slope | intercept (ppb) | Avg 96% Conf (ppb) | min Abs error (ppb) | max Abs error (ppb) | mean Abs error (ppb) |
| 0 | 0.98 | 2.06 | 31.5 | 0.9 | 37.4 | 15.3 |
| 1 | 1.02 | -2.21 | 31.5 | 0.2 | 44.0 | 13.9 |
| 2 | 1.03 | -3.49 | 29.4 | 0.2 | 53.5 | 12.9 |
| 3 | 0.97 | 3.64 | 31.8 | 0.3 | 41.1 | 15.2 |

Table 35. Results of cross validation using peak height calibration

| Calibration by neural network | | | | | | |
|---|---|---|---|---|---|---|
| Set | slope | intercept (ppb) | Avg 96% Conf (ppb) | min Abs error (ppb) | max Abs error (ppb) | mean Abs error (ppb) |
| 0 | 1.01 | -1.42 | 14.7 | 0.0 | 28.2 | 6.1 |
| 1 | 0.95 | 4.17 | 22.2 | 0.2 | 32.7 | 10.0 |
| 2 | 1.00 | 1.21 | 18.1 | 0.0 | 31.3 | 7.8 |
| 3 | 0.99 | 1.12 | 19.7 | 0.5 | 33.9 | 8.9 |

Table 36. Results of cross validation using a neural network

A one tailed paired Student's t test was then performed on the average 95% confidence limits and the mean absolute error values for both calibration models in order to ascertain if the apparent improvement in performance offered by the neural network was significant at the 95% level

### 8.2.1. Significance of mean absolute error figure

Microsoft Excel© was used to perform a one tailed paired Student's t test on the mean absolute error of the network ($x_1$) and the peak heights calibration ($x_2$). The null hypothesis, $H_0$, was that the means of the two sets of errors, $\mu_1$ and $\mu_2$, from each calibration method, were equal. The alternative hypothesis, $H_1$, was that they were not equal. This gave the following results.

Null hypothesis $H_0$: $\mu_1 = \mu_2$ or $\mu_1 - \mu_2 = 0$.

Critical region with 3 degrees of freedom and $\alpha = 0.05$ is $t > 2.353$

$t(x_1, x_2)$, calculated by Excel, $= 5.387$, with a one-tail probability of $P(T<=t)$ of 0.006.

The decision is therefore to reject the null hypothesis that the means of the two sets of mean absolute error are equal.

Since the t value obtained falls in the right tail of the critical region it can be concluded that processing by neural network does indeed produce a lower mean absolute error in prediction than does using calibration curves obtained from peak heights.

### 8.2.2. Significance of mean 95% confidence error figure

Microsoft Excel© was again used to perform a one tailed paired students t test on the 95% confidence interval of the network ($y_1$) and the heights calibration ($y_2$) gave the following results.

The null hypothesis, $H_0$, was that the means of the two sets of errors, $\mu_1$ and $\mu_2$, from each calibration method, were equal. The alternative hypothesis, $H_1$, was that they were not equal. This gave the following results.

Null hypothesis $H_0$: $\mu_1 = \mu_2$ or $\mu_1 - \mu_2 = 0$.

Critical region with 3 degrees of freedom and $\alpha = 0.05 = t > 2.353$

t ($y_1$ , $y_2$ ), calculated by Excel, = 7.776 with a one-tail probability of P(T<=t) of 0.002.

The decision is therefore to reject the hypothesis that the means of the two 95% confidence intervals are equal.

Since the t value obtained for this data also falls in the right tail of the critical region it can be concluded that processing by neural network produces a more precise prediction. That is, the 95% confidence limits in predicting a concentration are smaller than those obtained using calibration curves obtained from peak heights

## 8.3. Recommendations for further work

A period of research should lead to as many questions as it does answers and this work is no exception. The conclusions here indicate that neural networks are better at obtaining predictions of lead concentration from voltammetric scans than using a calibration curve based upon peak heights. Calibrations using the technique of standard additions, where a known quantity of the metal under investigation is introduced into the test sample and the scan repeated, were deliberately not investigated in this study. This technique has the advantage that the chemical matrix used for the calibration is the same as that containing the original concentration of test metal. It has the disadvantage that several scans must be made on the same sample in order to obtain a reliable prediction. In addition, the method used to measure the response is still the peak height, which this work suggests could be improved upon.

One area that should be investigated therefore, is a comparison of the neural network calibration technique detailed in this work with the method of standard additions using real analytical samples. Although not particularly suited to a field instrument, it would be interesting to know if the network approach could still produce better results.

A second area of research is to investigate the question of why the neural network appeared to produce superior performance. Theory suggests a relatively simple linear relationship between peak height and concentration, hence the additional information in the voltammogram ought to have been redundant. At this stage no attempt has been made to explain the results reported here but it would be useful to understand what features of the data the neural network is utilising.

Preliminary investigations into the determination of more than one element at a time has been investigated by the author for simulated curves (see chapter 9) and this shows very promising results. Neural networks appear to be able to model very well the convoluted curves produced by simulated overlapping voltammograms and there is a wealth of literature on the subject, some of which is reported here in the background study. A further study needs to be made to ascertain whether, how, and to what extent, neural networks can be used to predict concentrations of all metals of interest in a single differential pulse scan.

In order to expand the range of the instrument it must operate over a wider range of sensitivities, pulse heights, current offsets and pulse mark space ratios. This work indicates that the network can use information relating to pulse height with advantage. Further work needs to be done to provide a control loop to extend the range of the instrument to cover all the concentrations that it is able to measure. Furthermore the operational parameter settings could possibly be set automatically, following a single initial scan. The use of a neural network as a control mechanism is facilitated by the 'hooks' built into the current software enabling instrumental parameters to be read from a file Further work could be done to recognise poor quality scans caused by a number of factors and to map these to a set of parameters which converge to an optimum scan which could be used for concentration estimation.

A worthy piece of future work would be the development of this work into a completely portable field instrument. The electronic hardware could be reduced in size and the neural network could be contained on a single chip. The chemical cell used for the analysis is less easily miniaturised but some advances in micro cells have been reported. There would remain serious chemical problems involving complexing of heavy metals with other agents but it is possible that a robust field

instrument capable of detecting toxic metals in the parts per billion range in a matrix such as sea water could could be developed.

Finally, the work of this thesis has used peak shaped curves from voltammetric analysis. An obvious question therefore is whether the technique employed here can be advantageously applied to other, unrelated, instrumental responses that have a similar peak shaped response.

The topic of further research would appear to be not so much a question of what to do but which, out of the numerous options, to pursue first.

# 9. Epilogue - Simultaneous Prediction of Multiple Species

As part of the preliminary work, synthetic mixtures of lead and thallium were generated using a computer program written by the author. This program generated many of the features seen in a typical voltammogram and then combined them all into one curve. A neural network was then trained to predict concentrations of both metals simultaneously. This work was originally carried out as an initial investigation into the feasibility of using neural networks to process voltammograms. Although not investigated further than detailed here, a summary of the work is included in this thesis as it represents an area warranting further research.

The synthetic voltammogram was constructed by combining the following features:

1. The two peak shaped responses themselves, generated using the equation given in Engblom[120,121], (repeated here on page 36).
2. A quadratic background (after Bond and Babaric[251])
3. A small random noise function on each of the data points
4. A small random noise function on each of the peak positions
5. A skew on each of the peaks

By inspecting and manually fitting real voltammetric curves generated using a Ag/AgCl electrode and a preplated electrode, the following parameters were used in the generation of the synthetic mixtures.

| | | |
|---|---|---|
| Peak Position for Lead | = | -0.396 V wrt SCE |
| Peak Position for Thallium | = | -0.460 V wrt SCE |
| Number of electrons transferred for lead | = | 2 |
| Number of electrons transferred for thallium | = | 1 |
| Half width for 1 electron transfer | = | 0.0754 V |
| Noise on data points | = | ± a random value between 0 and 0.2 μA |
| Noise on peak position | = | ± a random value between 0 and 2 mV |
| Base line equation ($ax^2+bx+c$) where | a = | 30 ± a random value between 0 and 3 |
| | b = | 20 ± a random value between 0 and 3 |
| | c = | 5.8 ± a random value between 0 and 1 |
| Skew = | | $\frac{1}{\sqrt{V}}$ where $V$ = scan potential |

A number of synthetic mixtures of lead and thallium were generated, each one containing a different combination of concentrations. Concentrations were generated from 0 to 100 ppb in steps of 10 ppb for both lead and thallium and 3 simulated repetitions were made. There were 60 data points in each curve.

All 60 data points were used as an input to the neural network, the target being two real numbers representing the concentration of lead and thallium respectively. Both a three layer backpropagation network with 42 hidden nodes and a GRNN network were trained with the same training set created by extracting 80% of the patterns with the remaining 20% going into the test set.

The results of training such networks are shown graphically below.

## 9.1 Backpropagation Neural Network

There were 181 training patterns and 61 test patterns. The input for backpropagation was scaled to ±1, the learning rate was 0.1, momentum 0.1 and initial weights set to ±0.3.

Training time was 1hr 20 min on a 40MHz 80386 with maths co-processor and took 375 epochs, a epoch being defined as one complete presentation of the training set.

At intervals of 200 pattern presentations the test set was presented without updating the weights. A graph of the error in the test set as the training proceeded is shown in figure 74.

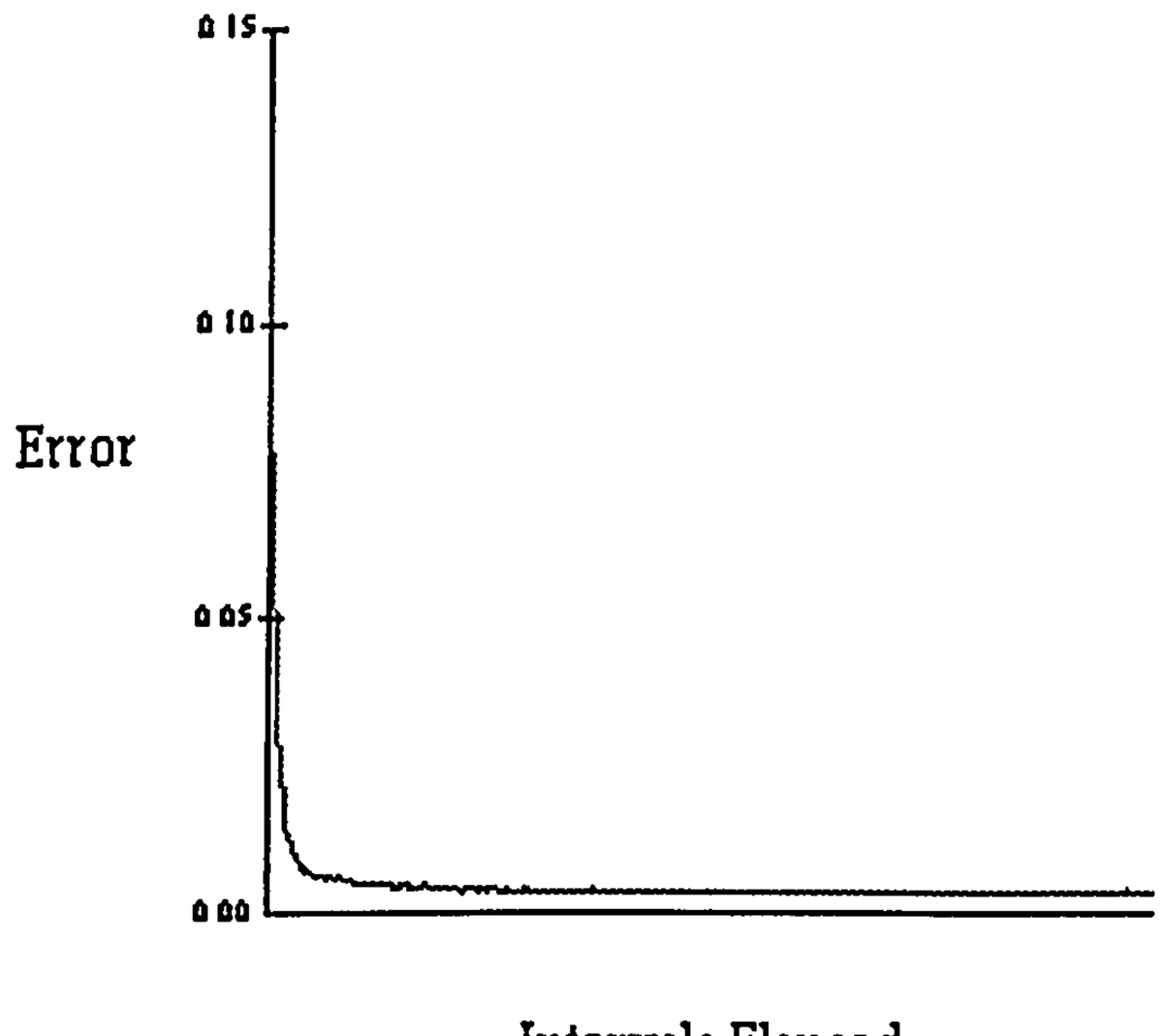



Figure 74. Graph of test set error against training epochs

Once the network had trained the training set and test set were both independently presented and the predicted concentrations of lead and thallium obtained. For the training set the curve of predicted concentration against target concentration showed a good correlation. Figures 75.and 76.below illustrate this.

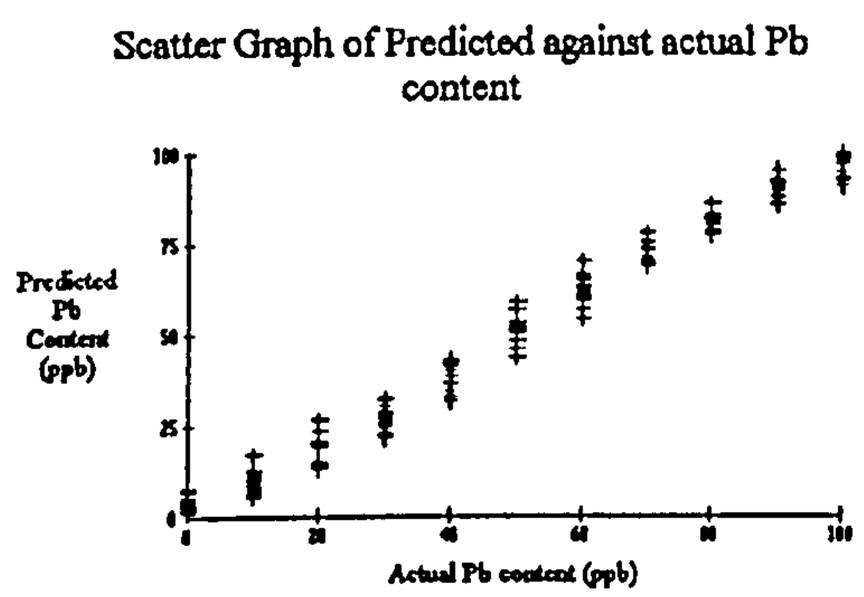



Figure 75. Errors in lead prediction for training set (BP)

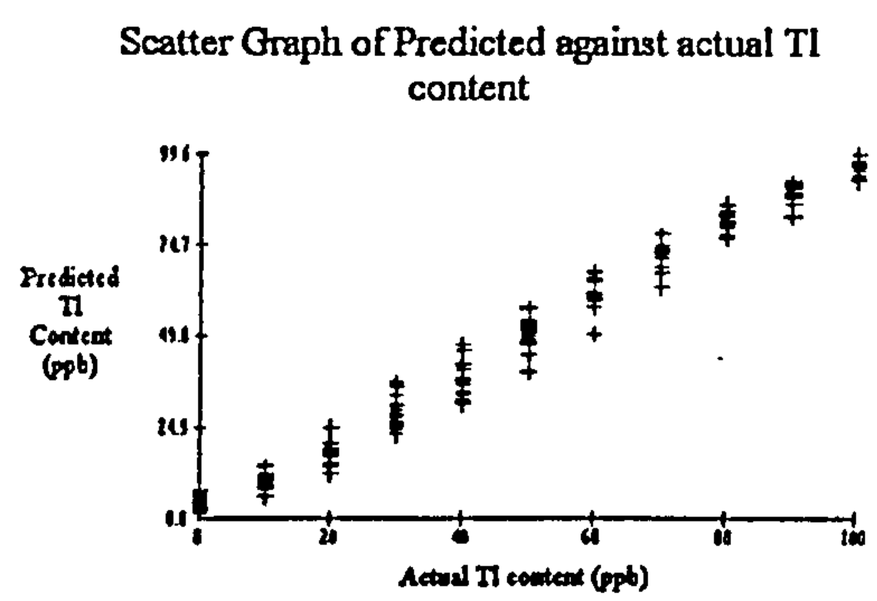



Figure 76. Errors in thallium prediction for training set (BP)

A scatter graph of the errors in lead and thallium prediction for the training set is shown in figure 77.

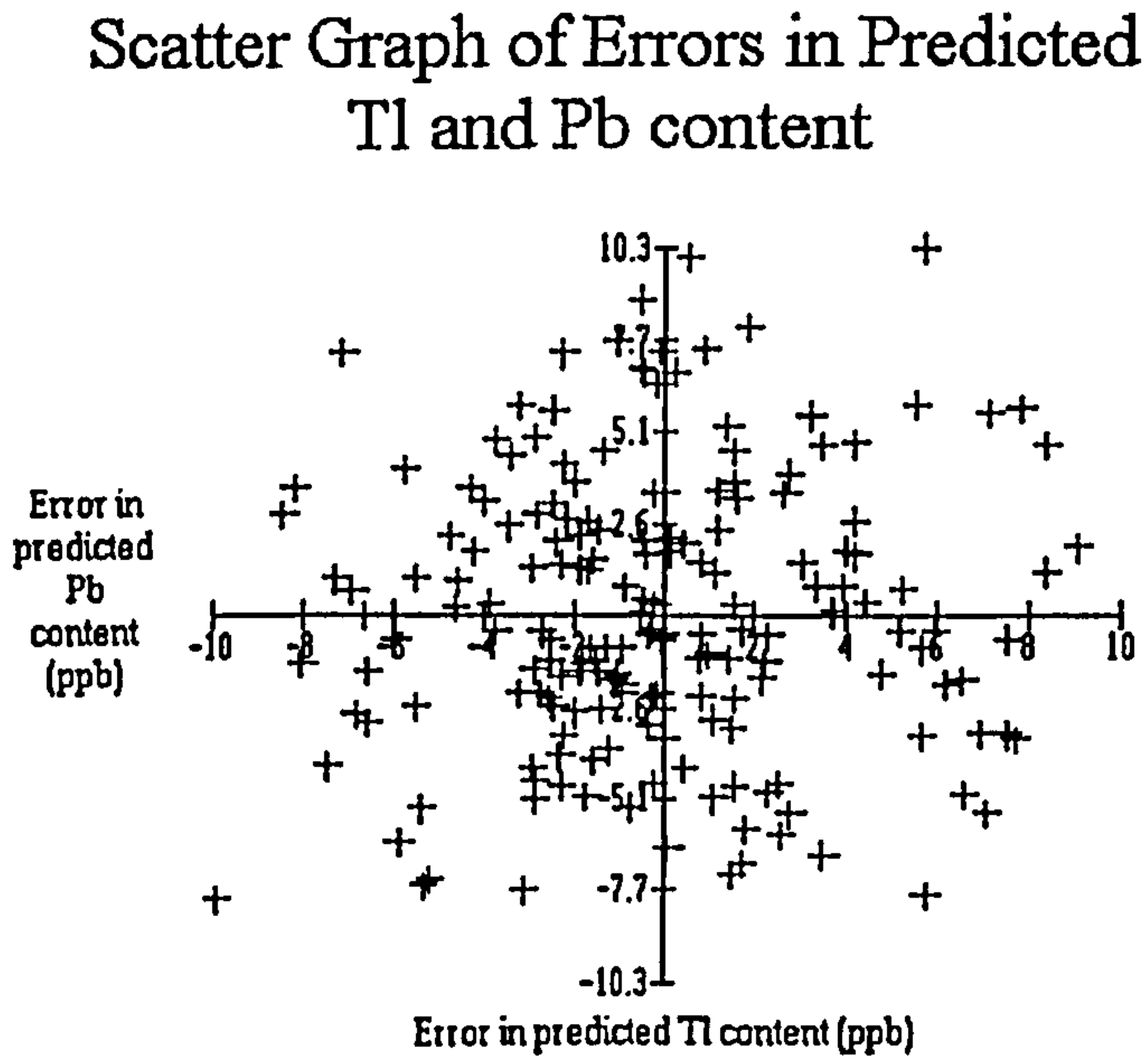



Figure 77. Errors in training set for back propagation

Equivalent graphs for the test sets are shown below

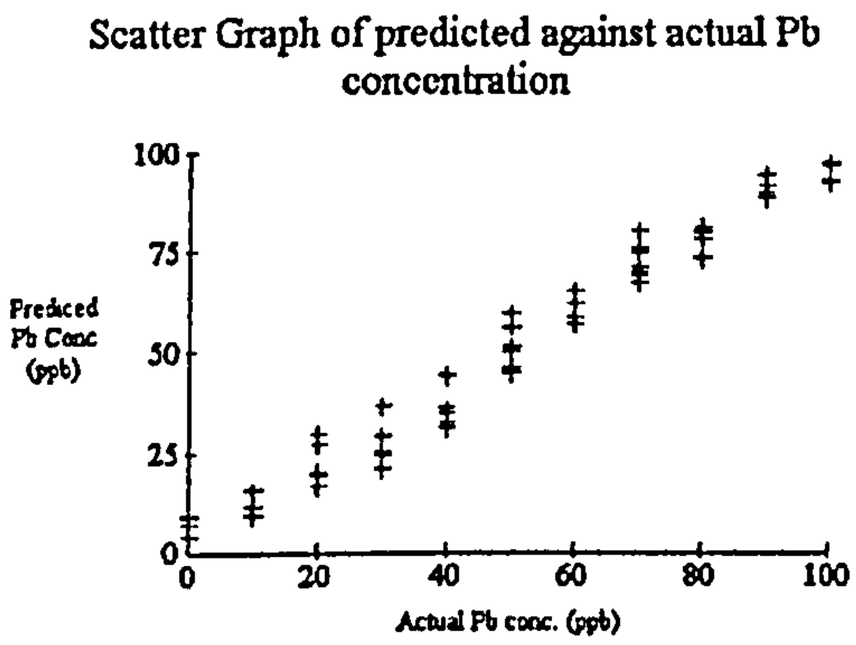



Figure 78. Errors in lead prediction for test set (BP)

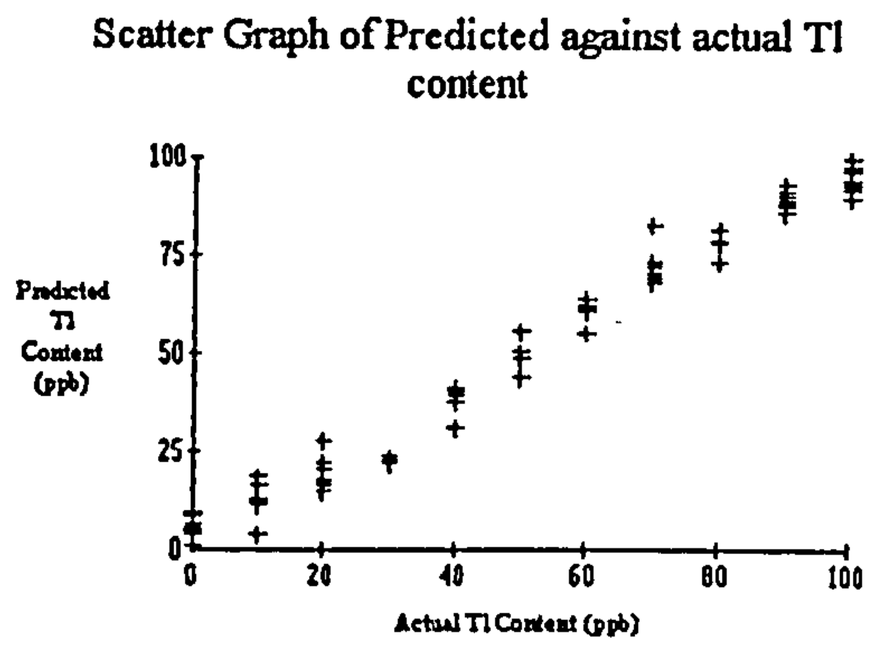



Figure 79. Errors in thallium prediction for test set (BP)

The corresponding scatter graph of errors for the test set is shown here. It can be seen that the errors in thallium are spread over a larger range than those for lead, but that both metals are predicted to within a simulated ±15 ppb.

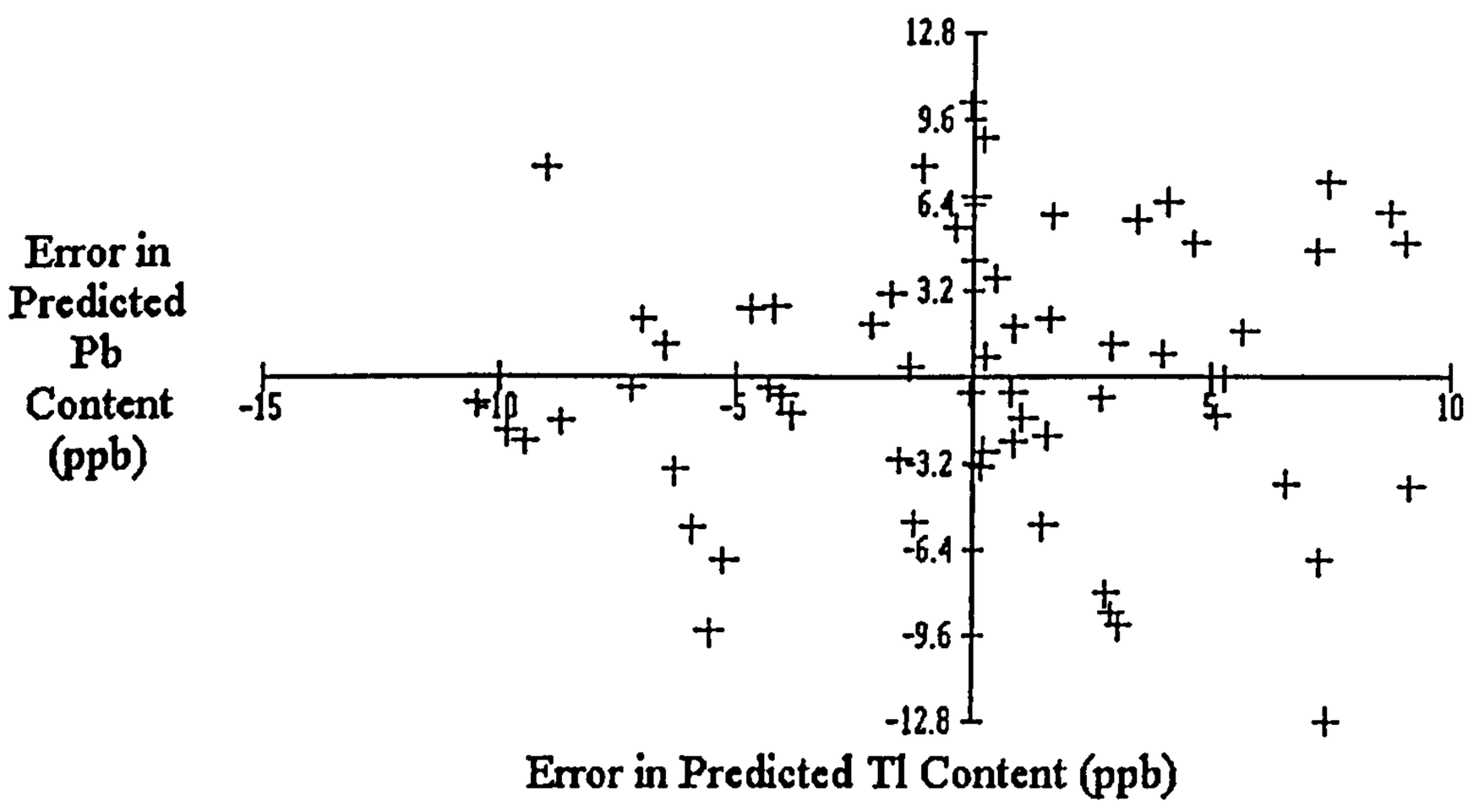



Figure 80. Errors in test set for backpropagation

## 9.2 General Regression Neural Network

This network had 60 inputs, 181 hidden units and 2 outputs. The smoothing factor was found as previously described, with the best value for the test set being 0.276. Training time was 3.5 minutes including the search for the best smoothing factor. The graphs for lead and thallium prediction for the training set are shown below.

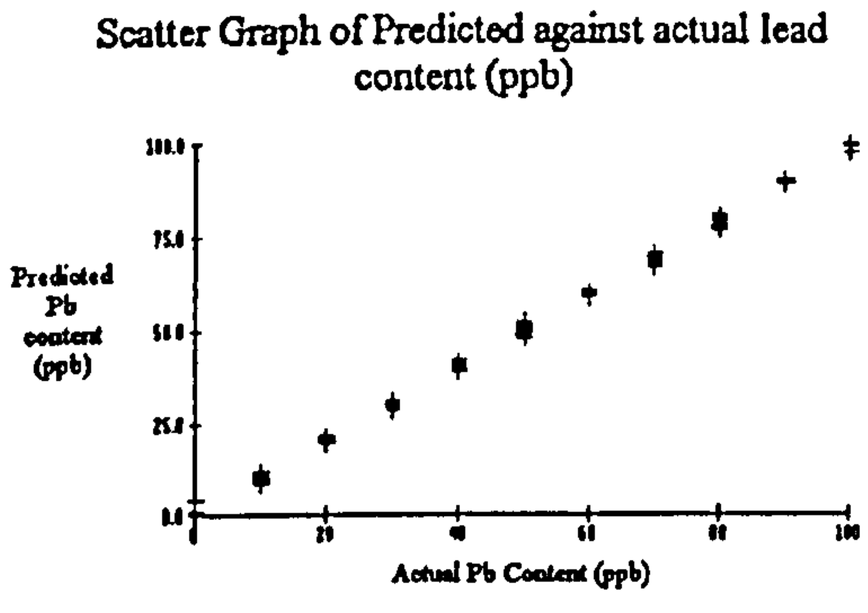



Figure 81. Errors in lead prediction for training set (GRNN)

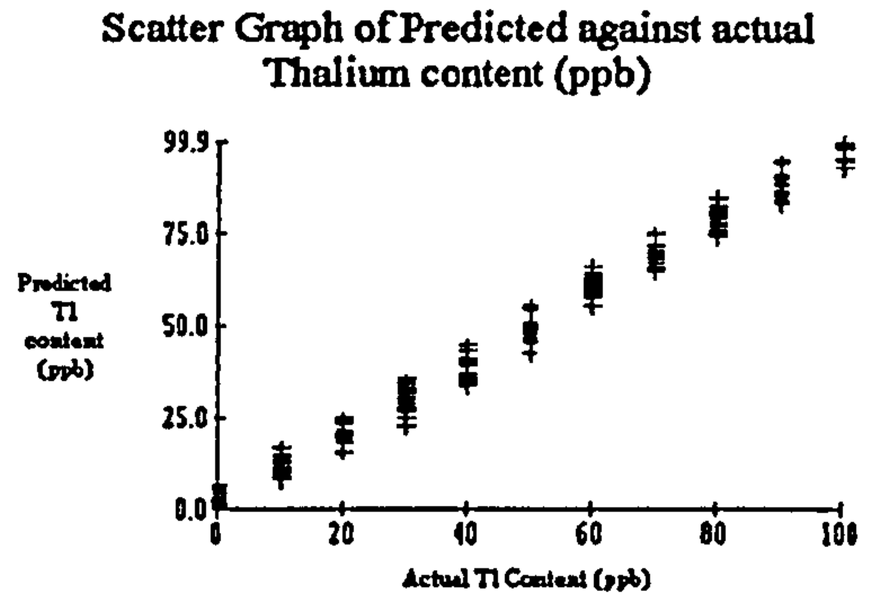



Figure 82. Errors in thallium prediction for training set (GRNN)

On the training set the GRNN appears to perform better than the back propagation and seems particularly good at predicting the lead content. A scatter graph of errors of given below which indicates a cluster of small errors in lead prediction combined with a few lead predictions having much larger errors.

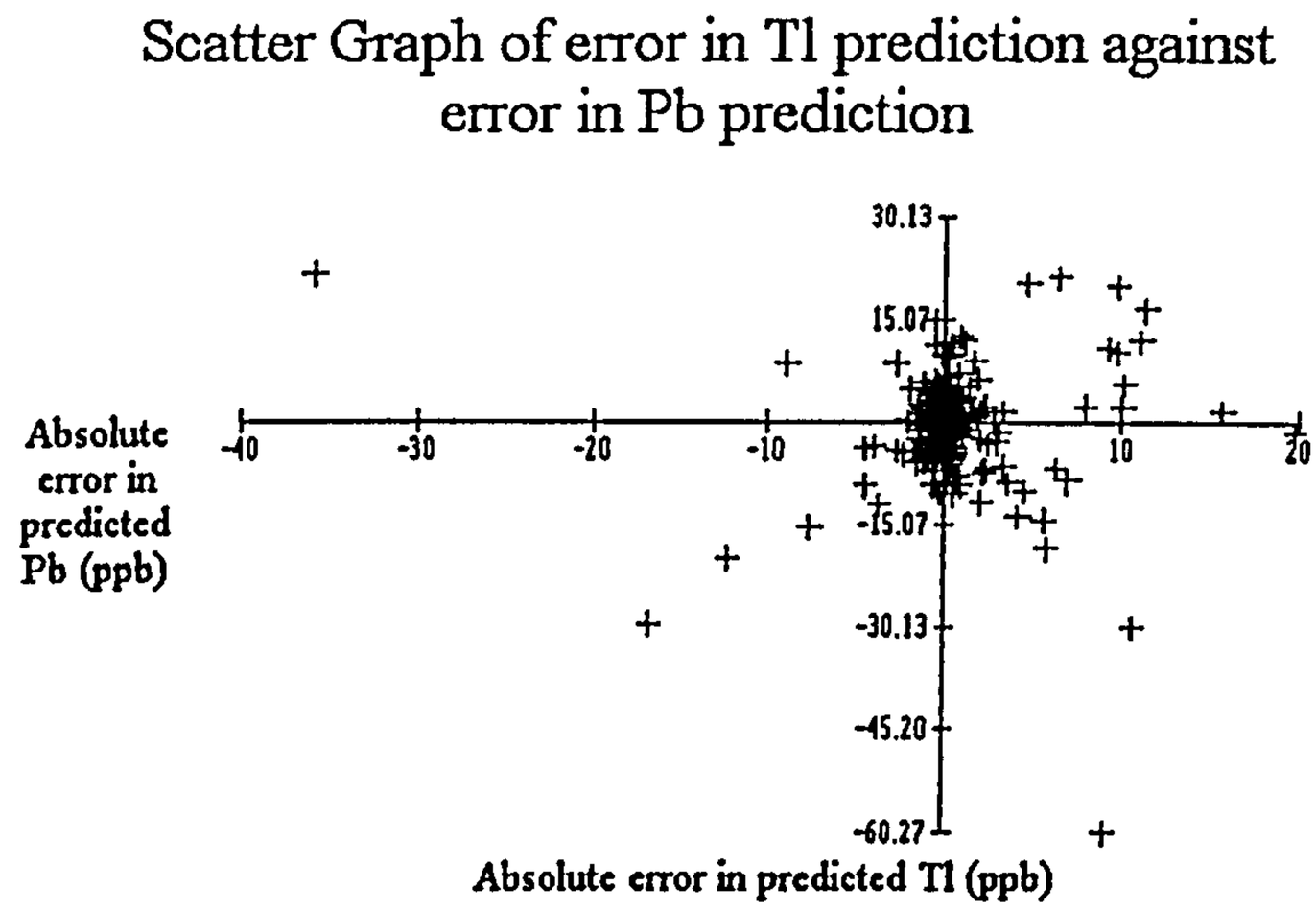



Figure 83. Errors in training set for GRNN

When the GRNN was used to process the test set the errors for both lead and thallium were larger than for the training set, showing a uniform scatter. These results are shown below. The scatter graph of errors for lead and thallium show the same pattern as that obtained with the training set with the thallium errors greater than those for lead and being greater when under predicting. (See figure 86).

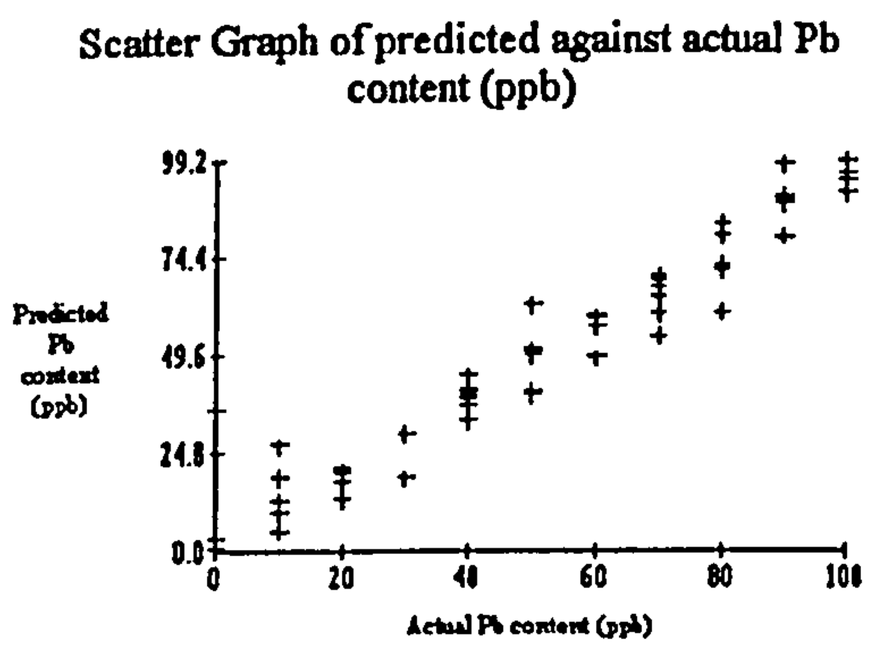



Figure 84. Errors in lead prediction for test set (GRNN)

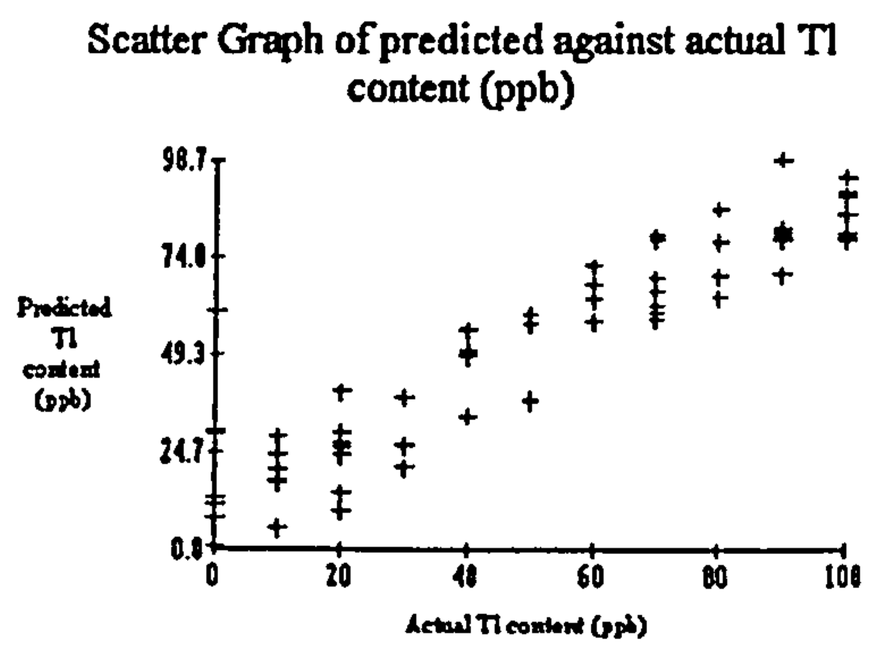



Figure 85. Errors in thallium prediction for test set (GRNN)

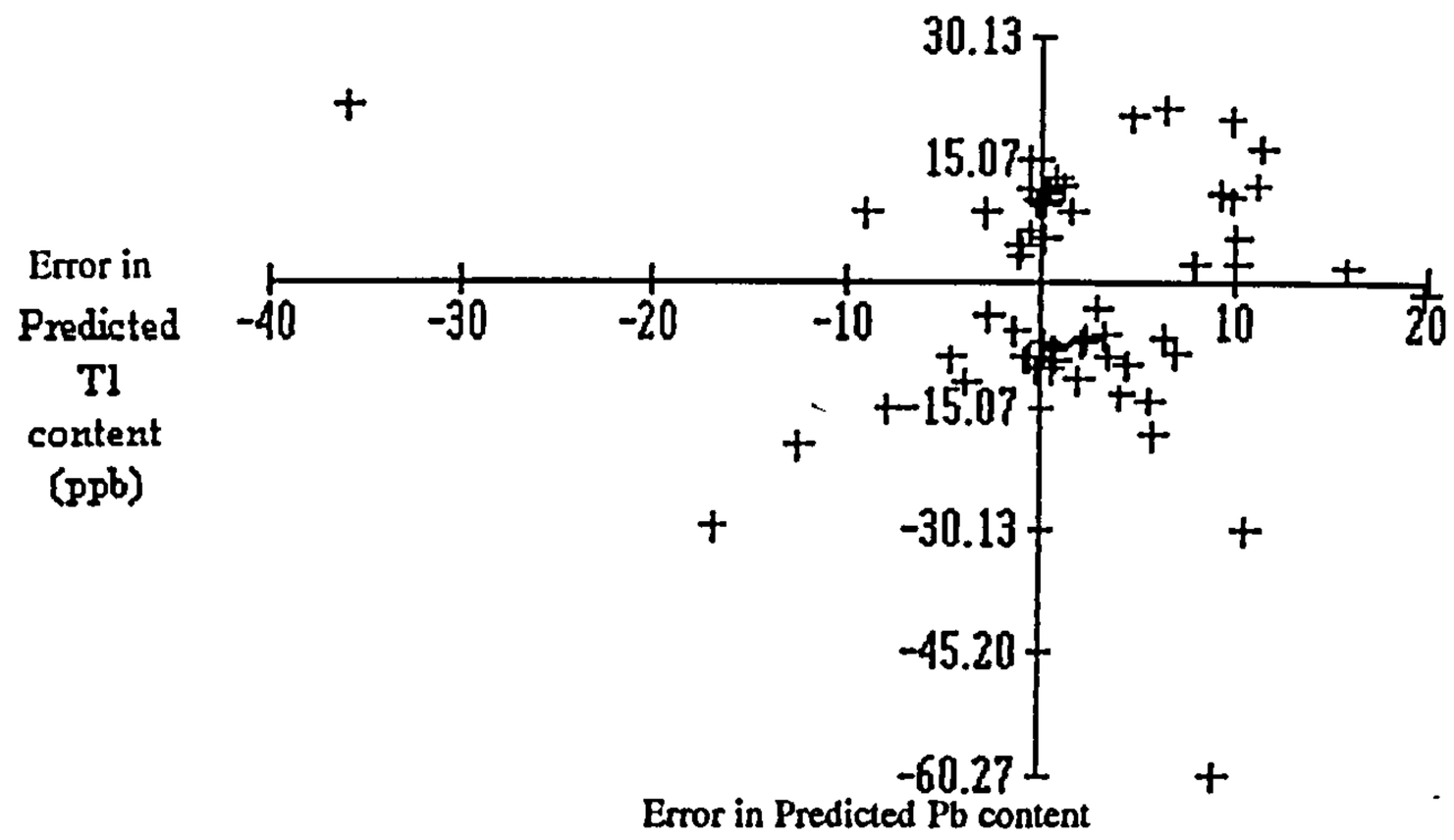



Figure 86. Errors in test set for GRNN

The results show that concentration estimations may be made from this data for two species simultaneously even when the peaks overlap, as they did for large simulated concentrations of the metals. In this simulated experiment typical absolute errors in estimation were of the order of 10 to 15 ppb for the both the training set and the test set. Further experiments were also briefly carried out with simulated three element mixtures containing lead, thallium and indium. First results are encouraging but are not reported here.

This work indicated that it may be possible to predict simultaneously the concentration of mixtures of metals if the individual voltammograms are purely additive. This simulation assumes no interaction between the two metals. However, as previously stated, this work formed part of the preliminary investigations into the feasibility of the process and was not carried out as rigorously as the latter parts of this work. Nevertheless it does have scope for further work.

# Bibliography

[1]Ward Systems Inc, Frederick, Maryland, USA.

[2]Heyrovský J. Chem. Listy. 1922, 16, 256.

[3]Heyrovský J. Phil. Mag. 1923, 45, 303.

[4]Heyrovský J. Shikata M., Rec. trav. chim. 1925, 44, 496.

[5]Skoog D. A., West D. M. and Holler F. J. Fundamentals of analytical chemistry. London: Saunders College Pub.

[6]Strobel H. and Heineman W. Voltammetry: Stationary solution techniques. *In*: Strobel H., ed. Chemical Instrumentation : A systematic approach.1989. London. John Wiley.

[7]Riley T. and Watson A. Polarography and other voltammetric methods. 1987. London. John Wiley.

[8]Riley T. and Tomlinson C. Principles of electroanalytical methods. 1987. London. John Wiley.

[99]Stock J. T. Electrochemistry in retrospect. *In*: Stock J. T. ed. Electrochemistry past and present. 1989, New York. ACS Symposium series.

[10]Wang J. Stripping Analysis. 1985. New York. VCH Publishers.

[11]Fried I. and Elving P. J. The rotating disk electrode. Anal. Chem. 1965, 37, 803-806.

[12]Anderson J. E. and Bond A. M. Digital Alternating Current Polarography with Microprocessor-based instrumentation. Anal. Chem. 1981, 53, 1394-1398.

[13]Wightman R. M. and Wipf D. O. High speed cyclic voltammetry. Acc. Chem. Res. 1990, 23(3), 64-70.

[14]Osteryoung J. G. and Osteryoung R. A. Pulse polarographic analysis of toxic heavy metals. American Laboratory. 1972, 4, 8-19.

[15]Osteryoung J. G. and Osteryoung R. A. Square wave voltammetry. Anal. Chem. 1985, 57(1), 101A.

[16]Murphy M. M., O'Dea J. J. and Osteryoung J. G. Pulse voltammetry at microcylinder electrodes. Anal Chem. 1991, 63, 2743-2750.

[17]Bond A. M. and Grabaric B. S. Fast sweep differential pulse voltammetry at a dropping mercury electrode with computerised instrumentation. Anal. Chem. 1979, 51, 126-128.

[18]Wang J. Stripping Voltammetry. 1985. New Mexico. VCH Pub.

[19]Kopanica M. and Opekar F. Electrochemical stripping analysis. *In*: Kalvoda R. ed. Electrochemical methods in chemistry. 1987, London, Plenum Press.

[20]Brainina Kh. Z. Stripping Voltammetry in Chemical Analysis. 1974. London. John Wiley & Sons.

[21]Graabæk A. M. and Jeberg B. Trace metal analysis by computerised stripping potentiometry. Int. Lab. 1992, September, 33-38.

[22]Pomeroy R. S., Denton M. B. and Armstrong N. R. Voltammetry at the thin film mercury electrode (TFME). Journal of Chemical Education, 1989, 66(10), 877-880.

[23]DeAngelis T. P., Bond R. E., Brooks E. E. and Heineman W. R. Thin layer differential pulse voltammetry. Anal. Chem. 1977, 49(12), 1792-1797.

[24]Copeland T. R. and Skogerboe R. K. Anodic Stripping Voltammetry. Anal. Chem. 1974, 46, 1257-1268.

[25]Copeland T. R., Christie J. H., Osteryoung R. A. and Skogerboe R. K. Analytical applications of pulsed voltammetric stripping at thin film mercury electrodes. Anal. Chem. 1973, 45, 2171-2174.

[26]Edwards C. A. Anodic stripping voltammetry - a review. 1976. Water Research Centre Technical Report.

[27]Svensmark B. Staircase voltammetry as a rapid detection method for anodic stripping determination of lead. Anal. Chim. Acta. 1987, 197, 239-248.

[28]Eisner U., Turner J. A. and Osteryoung R. A. Staircase voltammetric stripping analysis at thin film mercury electrodes. Anal. Chem. 1976, 48, 1608-1610.

[29]Anderson J. E. and Bond A. M. Microprocessor controlled instrument for the simultaneous generation of square wave, alternating current, direct current and pulse polarograms. Anal. Chem. 1983, 55, 1934-1939.

[30]Mlakar M. and Lovric M. Comparison of stripping methods at thin film mercury electrodes. Analyst. 1990, 115, 45-48.

[31]Feldman B. J., Osterloh J. D., Hata B. H. and D'Alessandro A. Determination of lead in blood by square wave anodic stripping at a carbon disc ultramicroelectrode. Anal. Chem. 1994, 66, 1983-1987.

[32]Christie J. H., Osteryoung, J. G. and Osteryoung R. A. Instrumental artifacts in differential pulse polarography. Anal. Chem. 1973, 45, 210-215.

[33]Beebe K. R., Blaser W. W., Bredeweg R. A. et. al. Process analytical chemistry. Anal. Chem. 1993, 65, 199R-216R.

[34]Bersier P. M. and Bersier J. Polarography, voltammetry and tensammetry: Tools for day to day analysis in the industrial laboratory. Analyst. 1989, 114, 1531-1535.

[35]Bersier P. M., Howell J. and Bruntlett C. Advanced electroanalytical techniques versus atomic absorption spectrometry, inductively coupled plasma atomic emission spectrometry and inductively coupled plasma mass spectrometry in environmental analysis. Analyst. 1994, 119, 219-232.

[36]Bond A. M. Comparison of semi-integral, semi-differential, direct current linear sweep, direct current derivative sweep, pulse and related voltammetric methods by computerised instrumentation. Anal. Chem. 1980, 52, 1318-1322.

[37]Anderson J. E. and Bond A. M. Comparison of differential pulse and variable amplitude pseudo derivative normal pulse polarography with microprocessor based instrumentation. Anal Chem. 1981, 53, 504-508.

[38]Willard H. H. and Merritt Jr L. Instrumental Methods of Analysis. 1988. California. Wadsworth Press.

[39]Parry E. P. and Osteryoung R. A. Evaluation of analytical pulse polarography. Anal. Chem. 1965, 37(13), 1634-1637.

[40]Rifkin S. C. and Evans D. E. Analytical evaluation of differential pulse voltammetry at stationary electrodes using computer based instrumentation. Anal. Chem. 1976, 48(14), 2174-2179.

[41]Gates S. C. Laboratory data collection with an IBM PC. Byte. 1984, May, 366-378.

[42]Okamura K. and Aghai-Tabriz K. A low cost data acquisition system. Byte. 1985, February, 199-202.

[43]He P., Avery J. P. and Faulkner L. R. Cybernetic control of electrochemical repertoire. Anal. Chem. 1982, 54(12), 1313A-1326A.

[44]Betteridge D. and Goad T. B. The impact of microprocessors on analytical instrumentation. Analyst. 1981, 106, 257-282.

[45]Tacussel J., Leclerc P. and Fombon J. J. Pulse voltammetry and polarography: recent progress in microprocessor instrumentation. J. Electroanal. Chem. 1986, 214, 79-94.

[46]Gunasingham H. Development of Electrochemical Instrumentation. *In*: Stock J. T. Electrochemistry Past and Present. 1989. New York. ACS Symposium Series. 236-253.

[47]Osteryoung R. A. and Wechter C. *In*: Stock J. T. Electrochemistry Past and Present. 1989. New York. ACS Symposium Series. 380-395.

[48]Keller H. E. and Osteryoung R. A. Application of a computerised electrochemical system to pulse polarography at a hanging mercury drop electrode. Anal. Chem. 1971, 43(3), 342-348.

[49]Landowne R. A., Morosani R. W., Herrmann R. A. et. al. Computer acquisition and analysis of gas chromatographic data. Anal. Chem. 1972, 44(12), 1961-1971.

[50]Granéli A., Jagner D. and Josefson M. Microcomputer system for potentiometric stripping analysis. Anal. Chem. 1980, 52, 2220-2223.

[51]Brown O. R. Control of electrochemical experiments by an inexpensive personal microcomputer. Electrochim. Acta.1982, 27(1), 33-46.

[52]Gustine D. L. and McCulloch J. Versatile microcomputer controlled, automated gradient analytical high performance liquid chromatography system. J. Chromatography. 1984, 316, 407-414.

[53]Paul D. W., Ridgway T. H. and Heinman W. R. Microprocessor controlled potentiostat for twin electrode voltammetry. Anal. Chim. Acta. 1983, 146, 125-134.

[54]Kalcher K. and Jorde C. Improved data processing for electroanalytical measurements. Computers Chem. 1986, 10(3), 201-218.

[55]Imaino W., Munce C., Yerry M. et. al. Data acquisition and signal processing using a laptop personal computer. Rev. Sci. Instrum. 1991, 62(2), 516-521.

[56]Lysaght M. J., van Zee J. A. and Callis J. B. Laptop chemistry: a fibre optic, field portable, near infrared spectrometer. Rev. Sci. Instrum. 1991, 62(2), 507-515.

[57]Ziegler E. and Schomburg G. The application of colachrom, a new command language for chromatographic data processing. J. Chromatography. 1984, 290, 339-350.

[58]Miller R. M. and Thomas K. E. Design of a computer controlled electroanalytical system. Analyst. 1989, 114, 777-783.

[59]Thomas K. E. A flexible approach to the design of voltammetric instrumentation. PhD Thesis. 1986, Dept. of Instrumentation and analytical science, University of Manchester.

[60]Gerth D. J., Howell T. and Shupack S. I. An object oriented data handling system for spectral or chromatographic data acquisition and analysis. Computers Chem. 1992, 16(1), 35-39.

[61]Carr J. J. Dealing with noise and interference in electronic instrumentation circuits. Elector Electronics. 1992, July, 20-23.

[62]Bromba M. U. A. and Ziegler H. Digital smoothing of noisy spectra. Anal. Chem. 1983, 55, 648-653.

[63]Deutscher R. L. and Fletcher S. Magic sampling - a digital sampling strategy that discriminates against interference (noise). Electrochim. Acta. 1990, 35, 1797-1804.

[64]DeNoyer L. K. and Dodd J. G. Maximum likelihood smoothing of noisy data. Int. Lab. 1990, June, 16-23.

[65]Lee T. A., Headley L. M. and Hardy J. K. Noise reduction of gas chromatography / mass spectrometry data using principal component analysis. Anal. Chem. 1991, 63, 357-360.

[66]Aubanel E. E. and Oldham K. B. Fourier smoothing. Byte. 1985, February, 207-218.

[67]Horlock G. Digital data handling of spectra utilising fourier transforms. Anal. Chem. 1972, 44, 943-947.

[68]Hayes J. W., Glover D. E., Smith D. E., et. al. Some observations on digital smoothing of electroanalytical data based upon the fourier transform. Anal. Chem. 1973, 45, 277-284.

[69]Felinger A., Pap T. L. and Inczédy J. Improvement in the signal to noise ratio of chromatographic peaks by fourier transform. Anal. Chim. Acta. 1991, 248, 441-446.

[70]Savitzky A. and Golay M. J. E. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 1964, 36, 1627-1639.

[71]Steiner J., Termonia Y. and Deltour J. Comments on smoothing and differentiation of data by simplified least square procedure. Anal. Chem. 1972, 44, 1906-1909.

[72]Bromba M. U. A. and Ziegler H. Application hints for Savitzky-Golay digital smoothing filters. Anal. Chem. 1981, 53, 1583-1586.

[73]Southampton Electrochemistry Group. Instrumental methods in electrochemistry. 1985. Chichester. Ellis Horwood Ltd.

[74]Pavididis T. Structured pattern recognition. 1977. New York. Springer-Verlag.

[75]Merritt W. and Settle D. Instrumental methods of analysis. 1981. London. John Wiley.

[76]Cacedi M. S. and Cacheris W. P. Fitting curves to data. Byte, 1984. May. 340-362.

[77]De Vries W. T. and Van Dalen E. J. Electroanal. Chem. 1964, 8, 366.

[78]De Vries W. T. Exact treatment of anodic stripping voltammetry with a plane mercury film electrode. J. Electroanal. Chem. 1965, 9, 448-456,

[79]De Vries W. T. and Van Dalen E. J. Electroanal. Chem. 1967, 14, 315-327.

[80]Osteryoung R. A. and Christie J. H. Theoretical treatment of pulsed voltammetric stripping at the thin file mercury electrode. Anal. Chem. 1974, 46, 351-355.

[81]Rifkin S. C. and Evans D. H. General equation for voltammetry with step functional potential changes applied to differential pulse voltammetry. Anal. Chem. 1976, 48, 1616-1618.

[82]Dillard J. W., Turner J. A. and Osteryoung, R. A. Digital simulation of differential pulse polarography with incremental time changes. Anal. Chem. 1977, 49, 1246-1250.

[83]Dillard J. W. and Hanck K. W. Digital simulation of differential pulse polarography. Anal. Chem. 1976, 48, 218-222.

[84]Dillard J. W., O'Dea J. J. and Osteryoung R. A. Analytical implications of differential pulse polarography of irreversible reactions from digital simulation. Anal. Chem. 1978, 51, 115-119.

[85]Ruzic I. and Sluyters-Rehbach M. The current-potential relationship for differential pulse polarography. Anal. Chim. Acta. 1978, 99, 177-182.

[86]Heijne G. J. M and Van der Lineden W. E., The current-potential relationship in differential pulse polarography. Anal. Chim. Acta. 1976, 82, 231.

[87]Kies H. L. private communication to Heijne G. J. M. September 1977.

[88]Heijne G. J. M and Van der Lineden W. E., The current-potential relationship in differential pulse polarography: a revision. Anal. Chim. Acta. 1978, 99, 183-187.

[89]Bond A. M. and Grabaric B. S. Correction for background current in differential pulse, alternating current and related polarographic techniques in the determination of low concentrations with computerised instrumentation. Anal. Chem. 1979, 51, 337-341.

[90]Mellado, J. M. R., Blazquez M. and Dominguez M. A curve fitting program set for handling of differential pulse polarograms. Comput. Chem. 1988, 12, 257-266.

[91]Kalman R. E. J. Basic Eng. 1960, 82, 34.

[92]Rutan S. C. Adaptive Kalman Filtering. Anal. Chem. 1991, 63, 1103A-1109A.

[93]Yongnian N. Curve resolution and quantification of pyrazines by differential pulse polarography using a Kalman filter approach. Analyst. 1993, 118, 1049-1053.

[94]Seelig P. F. and Blount H. N. Kalman filter applied to anodic stripping voltammetry : theory. Anal. Chem. 1976, 48, 252-258.

[95]Seelig P. F. and Blount H. N. Experimental evaluation of recursive estimation applied to linear sweep anodic stripping voltammetry for real time analysis. Anal. Chem. 1979, 51, 327-337.

[96]Poulisse H. N. J. Multicomponent analysis computations based on Kalman filtering. Anal. Chim. Acta. 1979, 112, 361-374.

[97]Wentzell P. D. Modelling chemical response surfaces with the Kalman filter. Anal. Chem. 1988, 60, 905-911.

[98]Perone S. P. Application of derivative techniques to stationary electrode polarography, Anal. Chem. 1965, 37, 2-9.

[99]Grushka E. and Israeli D. Characterization of Overlapped Chromatographic peaks by their second derivative. The limit of the method. Anal. Chem. 1990, 62, 717-721.

[100]Oppenheimer L., Capizzi T. P., Weppelman R. M. et. al. Determining the lowest limit of reliable assay measurement. Anal. Chem. 1983, 55, 638-643.

[101]Mitchell G. G., Mills W. N., Garden J. S. et. al. Multiple curve procedure for improving precision with calibration curve based analyses. Anal. Chem. 1977, 49, 1655-1660.

[102]Gutknecht W. F. and Perone S. P. Numerical deconvolution of overlapping stationary electrode polarographic curves with an online digital computer. Anal. Chem. 1970, 42, 906-917.

[103]Bond A. M. and Grabaric B. S. Simple approach to the problem of overlapping waves using a microprocessor controlled polarograph. Anal. Chem. 1976, 48, 1624-1228.

[104]Grotch S. L. Computer techniques for identifying low resolution mass spectra. Anal. Chem. 1971, 43, 1362-1370.

[105]Grotch S. L. Computer identification of mass spectra using highly compressed spectral codes. Anal. Chem. 1973, 45, 2-6.

[106]Caruana R. A., Searle R. B., Heller T. et. al. Fast algorithm for the resolution of spectra. Anal. Chem. 1986, 58, 1162-1167.

[107]Windig W., Haverkamp J. and Kistemaker P. G. Interpretation of sets of pyrolysis mass spectra by discriminant analysis and graphical rotation. Anal. Chem. 1983, 55, 81-88.

[108]Gillette P. C., Lando J. B. and Koenig J. K. Factor analysis for separation of pure component spectra from mixture spectra. Anal. Chem. 1983, 55, 630-633.

[109]Vogt N. B. Polynomial principal components regression: an approach to analysis and interpretation of complex mixture relationships in multivariate environmental data. Chemometrics and Intelligent Laboratory Systems. 1989, 7, 119-130.

[110]Tauler R., Casassas E. and Izquierdo-Ridorsa A. Self modelling curve resolution in studies of spectrometric titrations of multi-equilibria systems by factor analysis. Anal. Chim. Acta. 1991, 248, 447-458.

[111]Donahue S. M. and Brown C. W. Successive average orthogonalization of spectral data. Anal. Chem. 1991, 63, 980-985.

[112]Fay, M. J., Proctor A., Hoffmann D. P. et. al. Improved principal component analysis of noisy data. Anal. Chem. 1991, 63, 1058-1063.

[113]Vanslyke S. J. and Wentzell P. D. Real time principal component analysis using parallel Kalman filter networks for peak purity analysis. Anal. Chem. 1991, 63, 2512-2519.

[114]Glajch J. L., Kirkland J. J. and Squire K. M. Optimization of solvent strength and selectivity for reversed phase liquid chromatography using an interactive mixture design statistical technique. J. Chromatogr. 1980, 199, 57-79.

[115]Toft J. and Kvalheim O. M. Resolution parameter for multidetection chromatography. Anal. Chem. 1993, 65, 2270-2275.

[116]Elliot D. F. and Ramamohan R. Fast Transforms - Algorithms Analyses Applications. 1982. New York. Academic Press.

[117]Grabaric B. S. O'Halloran R. J. and Smith D. E. Resolution enhancement of AC polarographic peaks by deconvolution using the fast fourier transform. Anal. Chim. Acta. 1981, 133, 349-358.

[118]Skarjune R. Derivation and implementation of an efficient fast fourier transform algorithms (EFFT). Comput. Chem. 1986, 10, 241-251.

[119]Jackson R. J. and Griffiths P. R. Comparison of fourier self deconvolution and maximum likelihood restoration for curve fitting. Anal. Chem. 1991, 63, 2557-2563.

[120]Engblom S. O. The fourier transform of a voltammetric peak and its use in resolution enhancement. J. Electroanal. Chem. 1990, 296, 371-394.

[121]Engblom S. O. Fourier transform of a reversible linear sweep voltammogram. Anal. Chem. 1992, 64, 2530-2538.

[122]Davis L. Handbook of genetic algorithms. 1991. London. Von Nostrand Reinhold.

[123]Goldberg D. E. Genetic algorithms in search, optimisation and machine learning. 1989. New York. Addison Wesley Pub.

[124]Rosenblatt F. The perceptron: A perceiving and recognising automaton. 1957. Cornell Aeronautical Laboratory.

[125]Cybenko G. Approximation by superposition of a sigmoidal function. Math. Control Signals Systems. 1989, 2, 303-314.

[126]McCulloch W. and Pitts W. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics 1943. 7.

[127]Hebb D. Organisation of behaviour. 1949. New York. John Wiley.

[128]Widrow B. and Hoff M. E. Jr. Adaptive switching circuits. IRE WESCON Convention record. 1960, 4, 126-134.

[129]Minsky M. and Papert S. A. Perceptrons: An introduction to computational geometry. 1969. Cambridge MA. MIT Press.

[130]Giles C. L. and Maxwell T. Learning, invariance and generalisation in high order neural networks. Applied Optics. 1987, 26, 4272-4978.

[131]Kohonen T. Learning vector quantisation for pattern recognition. Technical Report TKK-F-A601. 1986. Helsinki University of Technology Finland.

[132]Anderson J. A., Silverstein S. A., Ritz S. A. and Jones R. S. Distinctive features, categorical perception and probability learning: Some applications of a neural model. Psychological Review. 1977, 84,413-451.

[133]Grossberg S. Learning by neural networks. *In*: Grossberg S. Ed. Studies of mind and brain. 1982, Boston, Reidel Pub. 1-52.

[134]Klimasauskas C. C. ed. The 1989 neurocomputing bibliography. 1989, Cambridge MA. MIT Press.

[135]Carpenter G. A. and Grossberg S. Category learning and adaptive pattern recognition, a neural network model. Proc of 3rd army conf. on applied mathematics and computation.1985, ARO Report 86-1, 37-56.

[136]Werbos P. Beyond regression: New tools for prediction and analysis in the behavioural sciences. PhD thesis. August 1974. Harvard, Cambridge, MA.

[137]Parker D. B, Lemma Logic. Technical Report TR-47, April 1985. Centre for Computational Research in Economics and Management Science, MIT, Cambridge MA.

[138] McClelland J, Rumelhart D. Explorations in parallel distributed processing, vol. 1 and 2. 1986. MIT press, Cambridge, MA.

[139]Lippmann R. P. An introduction to computing with neural nets. IEEE ASSP. April 1987, 5-22.

[140]Kohonen T. An introduction to neural computing. Neural Networks. 1988, 1, 3-16.

[141]Kosko B. Constructing an auto associative memory. Byte. September 1987, 137-144.

[142]Nelson M. M. A practical guide to neural nets. 1990. New York. Addison Wesley.

[143]Haykin S. Neural networks a comprehensive foundation. 1994. New York. Macmillian.

[144]Blum A. Neural networks in C++. 1992. New York. John Wiley & Sons.

[145]Freeman J. A. and Skapura D. M. Neural networks algorithms, applications and programming techniques. 1991. New York. Addison Wesley.

[146]Pao Y. Adaptive Pattern Recognition and Neural Networks. 1989, New York, Addison Wesley.

[147]Müller B. and Reinhardt J. Neural Networks 2nd (corrected) Ed. 1991. New York. Springer-Verlag.

[148]Fausett L. Fundamentals of Neural Networks. 1994. New Jersey. Prentice Hall.

[149]Nelson M. M. and Illingworth W. A Practical guide to Neural Networks. 1990, London, Addison Wesley.

[150]Zupan J. and Gasteiger J. Neural networks for chemists. 1993. New York. VCH.

[151]Fahlman S. E. An Empirical study of learning speed in back propagation networks. Technical report CMU-CS-88-162. 1988, Air force Wright Aeronautical laboratories. Ohio.

[152]Hirose Y., Yamashita K. and Hijiya S. Back propagation algorithms which varies the number of hidden units. Neural Networks. 1991, 4, 61-66.

[153]Weir M. K. A method for self determination of adaptive learning rates in back propagation. Neural Networks. 1991, 4, 371-379.

[154]Rigler A. K., Irvine J. M. and Vogl T. P. Rescaling of variables in back propagation learning. Neural Networks. 1991, 4, 225-229.

[155]Weymaere N. and Martens J. P. A fast and robust learning algorithm for feed forward neural networks. Neural Networks. 1991, 4, 361-369.

[156]Specht D. F. A General Regression Neural Network. IEEE Trans. Neural Networks. 1991, 2, 568-578.

[157]Lang K. J., Waibel A. H. and Hinton. G. E. A time delay neural network for isolated word recognition. Neural Networks. 1990, 3, 23-43.

[158]Kämmerer B. R. and Küpper W. A. Experiments for isolated word recognition with single and two layer perceptrons. Neural Networks. 1990, 3, 693-706.

[159]Jordan M. I. Serial order: A parallel distributed processing approach. Technical Report 8604. 1986, University of California. San Diego.

[160]Elman J. L. Finding structures in time. Cognitive Science. 1990, 14, 179-211.

[161]Elman J. L. Incremental learning or the importance of starting small. Technical Report 9101. 1991, University of California, San Diego.

[162]Elman J. L. Distributed representations, simple recurrent networks and grammatical structure. Machine Learning, 1991, 7, 195-225.

[163]Torreele J. Temporal processing with recurrent networks: An evolutionary approach. Proc. 6th Int. Conf. on Genetic Algorithms. 1991, San Diego, 555-561.

[164]Fu K. S. Syntactic methods in pattern matching. 1974, New York, Academic Press.

[165]Pavilidis T. Structural pattern recognition. 1977, New York, Springer-Verlag.

[166]Roberts S. and Tarassenko L. New method of automated sleep quantification. Med. & Biol. Eng. & Comp. 1992, September, 509-517.

[167]Roberts S. and Tarassenko L. Analysis of the sleep EEG using a multilayer network with spatial organisation. IEE Proc.-F, 1992, 139 (6), 420-425.

[168]Eberhart R. C. and Dobbins R. W. Detection of electroencephalogram spikes *In:* Eberhart R. C. and Dobbins R. W eds. Neural network PC tools. 1990, New York, Academic Press, 215-234.

[169]Levin E., Gewirtzman R. and Inbar G. Neural network architecture for adaptive system modelling and control. Neural Networks. 1991, 4, 185-191.

[170]Jang J. R. Self learning fuzzy controllers based on temporal back propagation. IEEE Trans. Neural Networks. 1992, 3(5), 714-723.

[171]Lister J. B., Schnurrenberger H. Staeheli N et. al. Neural networks in front end processing and control. IEEE Trans. on Nuclear Science. 1992, 39(2), 49-57.

[172]Jurs P.C., Kowalski B. R., Isenhour T. L., Reilly C. N., Computerized learning machines applied to chemical problems: Investigation of convergence rate and predictive ability of adaptive binary pattern classifiers. Anal Chem.1969, 41, 690-695.

[173]Kowalski B. R., Jurs P.C., Isenhour T. L., Reilly C. N., Computerized learning machines applied to chemical problems: Multicategory pattern classification by least squares. Anal Chem., 1969, 41,695-700.

[174]Kowalski B. R., Jurs P.C., Isenhour T. L., Reilly C. N., Computerized learning machines applied to chemical problems: Interpretation of infrared spectrometry data. Anal Chem., 1969, 41,1945-1949.

[175]Jurs P.C., Kowalski B. R., Isenhour T. L., Reilly C. N., An investigation of combined patterns from diverse analytical data using computerized learning machines. Anal Chem., 1969, 41,1949-1953.

[176]Wangen L. E., Isenhour T. L. Semiquantitative analysis of mixed gamma-ray spectra by computerized learning machines. Anal Chem, 1970, 42, 737-743.

[177]Jurs P. C., Machine intelligence applied to chemical systems. Anal Chem, 1971, 43, 22-26.

[178]Sybrandt L. B., Perone S. P. Computerized learning machine applied to qualitative analysis of mixtures by stationary electrode polarography. Anal Chem., 1971, 43, 382-388.

[179]Sybrandt L. B., Perone S. P. Computerized pattern classification of strongly overlapped peaks in stationary electrode polarography. Anal Chem., 1972, 44, 2331-2339.

[180]Tunnicliff D. D., Wadsworth P. A., Application of pattern Separation techniques to mass spectrometric data, Anal Chem., 1973, 45, 12-20.

[181]Thomas Q. V., Kruger L., Perone S. P, Computer-assisted optimization of anodic stripping voltammetry. Anal Chem, 1976, 48, 761-766.

[182]Byers W. A., Freiser B. S., Perone S. P. Structural and activity characterisation of organic compounds by electroanalysis and pattern recognition. Anal Chem,1983, 55, 620-625.

[183]Kaplan D., Raphaeli D., Ben-Yaakov S. Application of personal computers in the analytical laboratory - III. Talanta, 1987, 34, 709-714.

[184]Wold H. *In*: Krishnaiah P. R. ed. Multivariate Analysis.1966, New York, Academic Press.

[185]Wold H. Soft Modelling: The basic design and some extensions. In Systems under indirect observation, causality-structure-prediction (Eds Jöreskog K. G. and Wold H.) 1981, North Holland, Amsterdam. [186]Wold H. *In*: Martens H. and Russwurm H. eds. Food research and data analysis.1983. London. Applied Science Pub.

[187]Wold S., Martens H., Wold H., The multivariate calibration problem in chemistry solved by the PLS method. *In*: Ruhe A., Kàgström B. eds. Proc. Conf. Matrix pencils.1982, Lecture Notes in Mathematics. 1982. Heidelberg. Springer Verlag. 286-293.

[188]Lindberg W., Persson J., Wold S. Partial Least Squares Method for Spectrofluorimetric Analysis of Mixtures of Humic Acid and Ligninsulfonate, Anal. Chem. 55, 1983, 643-648.

[189]Geladi P., Kowalski B., Partial Least Squares Regression: A Tutorial, Anal. Chim. Acta. 1986, 185, 1-17.

[190]Wold S., Kettaneh-Wold N., Skagerberg B. Non Linear PLS Modeling, Chemom. Intell. Lab. Sys. 1989, 7, 53-65.

[191]Wold S., Nonlinear partial least squares modelling II. Spline inner relation. Chemom. Intell. Lab. Sys. 1992, 14, 71-84.

[192]Haarland D. M., Thomas E. V., Partial Least Squares Methods for Spectral Analyses. 1. Relation to other Quantitative Calibration Methods and the Extraction of Qualitative Information. Anal. Chem. 1988, 60, 1193-1202.

[193]Haarland D. M., Thomas E. V., Partial Least Squares Methods for Spectral Analyses. 2. Application to Simulated and Glass Spectral Data. Anal. Chem. 1988, 60, 1202-1208.

[194]Weijer A. P., Buydens L. Kateman G. et. al. Neural networks as a soft modelling technique for the quantitative description of the relation between physical strauctures and the mechanical properties of poly(ethylene terephthalate) yarns. Chemom. Intell. Lab. Syst. 1992, 16, 77-86.

[195]Long J. R., Gregoriou V. G., Gemperline P. J. Spectroscopic calibration and quantitation using artificial neural networks. Anal Chem, 1990, 62, 1791-1797.

[196]Long J. R., Mayfield H. T., Henley M. V., Kromann P. R. Pattern recognition of jet fuel chromatographic data by artifical neural networks with backpropagation of error. Anal Chem, 1991, 63, 1256-1261.

[197]Wythoff B. J., Levine S. P., Tomellini S. A. Specrtal peak verification and recognition using a multilayered neural network. Anal Chem 1990, 62, 2702-2709.

[198]Kalltenbach T. F., Small G. W., Development and optimisation of piecewise linear discriminants for the automated detection of chemical species, Anal. Chem. 1991, 63, 936-944.

[199]Perkins J. H., Hasenoehrl E. J. Griffiths P. R., Expert System Based on Principal Component Analysis for the Identification of Molecular Structures from Vapor-Phase Infrared Spectra. 1. Theory . Identification of Alcohols, Anal. Chem.1919, 63, 1738-1747.

[200]Peterson K. L., Counter-Propagation Neural Networks in the Modelling and Prediction of Kovats Indices for Substituted Phenols, Anal. Chem. 1992, 379-386.

[201]Hasenoehrl E. J., Perkins J. H., Griffiths P. R., Rapid Functional Group Characterisation of Gas Chromatography / Fourier Transform Infrared Spectra by a Principal Components Analysis Expert System, Anal. Chem. 1992, 64, 705-710.

[202]Bhandare P., Mendelson Y., Peura R. A., et. al. Multivariate Determination of Glucose in Whole Blood Using Partial Least-Squares and Artificial Neural Networks Based on Mid-Infrared Spectroscopy, Appl. Spectosc.1993, 47, 1214-1221.

[203]Gemperline P. J., Long J. R., Gregoriou V. G. Non linear multivariate calibration using principal components regression and artificial neural networks. Anal Chem. 1991, 63, 2313-2323.

[204]Allanic A. L., Jézéquel J. Y., André J. C. Application of neural networks theory to identify two dimensional fluorescence. Anal. Chem. 1992, 64, 2618-2622.

[205]Borggaard C., Thodberg H. H. Optimal minimal neural interpretation of spectra. Anal. Chem. 1992, 64, 545-551.

[206]McAvoy T. J., Wang N. S., Nadidu S. et. at. Interpreting biosenor data via backpropagation. Proc. Int. Joint Conf. on Neural Networks. 1989, Vol 1, 227-233.

[207]Blank T. B., Brown S. D. Non linear multivariate mapping of chemical data using feed-forward neural networks. Anal Chem 1993, 65, 3081-3089.

[208]Blank T. B., Brown S. D. Data processing using neural networks. Anal. Chim. Acta. 1993, 277, 273-287.

[209]Goodacre R., Edmonds A. N., Kell D. B. Quantitative analysis of the pyrolysis-mass spectra of complex mixtures using artificial neural networks: application to amino acids in glycogen. J. Anal. Appl. Pyrol. 1993, 26, 93-114.

[210]Goodacre R., Kell D. B., Rapid and quantitative analysis of bioprocess using pyrolysis mass spectrometry and neural networks: Application to indole production. Anal. Chim. Acta. 1993, 279, 17-26.

[211]Goodacre R., Kell D. B., Bianchi G. Rapid assessment using pyrolysis mass spectrometry and artificial neural networks of the adulteration of virgin olive oils by other seed oils. J. Sci. Food Agric. 1993, 63, 297-307.

[212]Gallant S. R., Franleigh S. P., Cramer S. M. Deconvolution of overlapping chromatographic peaks using a cerebellar model arithmetic computer neural network Chemom. Intell. Lab. Sys. 1993, 18, 41-57.

[213]Gasteiger J., Zupan J. Neural networks in chemistry. Angew. Chem. Int. Engl. 1993, 32, 503-527.

[214]Kateman G. Neural networks in analytical chemistry? Chemom. Intell. Lab. Sys. 1993, 19, 135-142.

[215]Næs T., Kvaal K., Isaksson T. Miller C., Artificial neural networks in multivariate calibration, J. Near Infrared Spectrosc. 1993, 1, 1-11.

[216]Bond A. M., Heritage I. D. and Thormann W. Strategy for trace metal determination in sea water by anodic stripping voltammetry using a computerized multi time domain measurement method. Anal. Chem. 1986, 58, 1063-1066.

[217]Williams G. and D'Silva C. Hand held instrument for environmental monitoring. Analyst, 1994, 119, 187-190.

[218]Palys M., Bos M. and van der Linden W. E. Automated polarographic elucidation of electrode mechanisms by means of a knowledge based system. Anal. Chim. Acta. 1991, 248, 429-439.

[219]Smits J. R., Breedveld L. W., Derksen M. W. et. al. Pattern classsification with artificial neural networks: Classification of algae based upon flow cytometer data. Anal. Chim. Acta. 1992, 258, 11-25.

[220]von Fraunhofer, J. A., Banks, C. H. Potentiostat and its applications, 1972, London. Butterworths press.

[221]Blanchet P. F. Simple computer-supported photoelectrochemical instrument for studies of ultrathin films on transparent electrodes. Rev. Sci. Instrum. 1989, 60, 2750-2755.

[222]Cath P. G. and Peabody A. M. High speed current measurements. Anal. Chem. 1971, 43(11), 91A-99A.

[223]Booman G. L. and Holbrook W. B. Optimum stabilisation networks for potentiostats with application to a polarographh using transistor operational amplifiers. Anal. Chem. 1965, 37(7), 795-802.

[224]Brown, E. R.,Smith, D. E., Booman, G. L. A study of operational amplifier potentiostats employing positive feedback for *iR* compensation. 1. Theoretical analysis of stability and bandpass characteristics. Anal. Chem. 1968, 40(10), 1411-1432.

[225]Kanazawa, K. K., Galwey, R. K. Current follower stabilisation in potentiostats. Anal. Chem. 1977, 49(4), 677-678.

[226]Meyer J. J. Potentiostat with a positive feedback *iR* compensation and a high sensitivity current follower indicator circuit for direct determination of high second order rate constants. Anal. Chem. 1982, 54, 207-212.

[227]Tallman D. E., Shepherd G. and MacKellar W. J. A wide bandwidth computer based potentiostat for fast voltammetry at microelectrodes. J. Electroanal. Chem. 1990, 280, 327-340.

[228]Harrington M. S., Anderson L. B., Robbins J. A. and Karweik D. H. Instruction manual and schematics. Physics auxiliary puplication service document number RSINA-60-3323-23, 1987.

[229]Harrington M. S., Anderson L. B., Robbins J. A. and Karweik D. H. Multiple electrode potentiostat. Rev. Sci. Instrum. 1989, 60(10), 3323-3329.

[230]Candela D. and McAllaster D. R. Synchronous pulse generator encompassing a wide range of time scales. Rev. Sci. Instrum. 1991, 62(2), 522-526.

[231]Bond A. M. and Norris A. Inexpensive microprocessor controlled programmable function generators for use in electrochemistry. Anal. Chem. 1980, 52, 367-371.

[232]Analog Devices. Data Conversion Products Databook. 1988, 2.105 - 2.112

[233]Jayaweera, P., Ramaley, L., Low Noise, Fast Response Potentiostat for Pulse electrode studies, Anal. Inst.1986, 15, 259-277.

[234]Švestka, M., Bond, A.M. Modified current to voltage converters for electrochemical instrumentation. J. Electroanal. Chem. 1986, 200, 35-46.

[235]Analog Devices. Data Conversion Products Databook. 1988, 7.13 - 7.16

[236]Davis, J. E., Clifford Toren Jr, E. On the instability of current followers in potentiostat circuits. Anal. Chem. 1974, 46(6) , 647-650.

[237]Bezman, R., McKinney, P. S. Anal Chem. 1969, 41, 1560.

[238]Analog Devices. Linear Products Databook. 1988, 4.61 - 4.72

[239]Analog Devices. Data Conversion Products Databook. 1988, 3.61 - 3.71

[240]Rifkin S. C., Evans D. H. Analytical Evaluation of differential Pulse Voltammetry at stationary electrodes using computer based instrumentation. Anal. Chem. 1976, 48, 2174-2179.

[241]Manwaring H. S. A comparison of backpropagation and general regression neural networks in quantative trace metal analysis using anodic stripping voltammetry and cybernetic instrumentation. Neural Compt & Applic. 1994, 2, 168-178.

[242]De Vries W. T. Linear potential-sweep voltammetry at a plane mercury-film electrode Electroanal. Chem & Interfacia. 1967, 14, 315-327.

[243]Wong D. K. Y., Ewing A. G. Anodic stripping voltammetry at mercury films deposited on ultrasmall carbon-ring electrodes. Anal. Chem. 1990, 62, 2697-2702.

[244]Hume D. N., Carter J. M., Chem. Anal. (Warsaw). 1972, 17, 747.

[245]Štulíková M. The deposition and stripping of mercury on a glassy carbon rotating disc electrode. Electroanalytical Chemistry and Interfacial Chemistry, 1973, 48, 33-45.

[246]Florence T. M., Anodic Stripping Voltammetry with a Glassy Carbon Electrode Mercury Plated in situ., Journal of Electroanalytical Chemistry, 1970,27, 273-281.

[247]De Vries W. T. J. Exact treatment of Anodic Stripping Voltammetry with a plane Mercury Film Electrode. Elecroanal. Chem. 1965, 9, 448-456.

[248]Lund W., Salberg M., Anodic stripping voltammetry with the Florence mercury film electrode. Determination of copper, lead and cadmium in sea water., Anal. Chim. Acta. 1975, 76, 131-141.

[249]Goto M., Ikenoya K., Ishii D., Anodic Stripping Semidifferential electroanalysis with thin film mercury electrode formed in situ. Anal. Chem. 1979, 51, 110-115.

[250]Florence T. M., Farrar Y. J., J. Electroanal. Chem. Interfacial Electrochem. 1974, 51, 191.

[251]Bond, A. M., Grabaric, B. S. Correction for Background Current in Differential Pulse, Alternating Current and related Polargraphic Techniques in the Determination of Low concentrations with Computerized Instrumentation, Anal. Chem.1979, 51, 337-341.

[252]Martens H., Næs T., Multivariate Calibration, 1989. New York. John Wiley.121-123.

[253]Jepson B, Collins A. Evans A. Post neural network procedure to determine expected prediction values and their confidence limits, Neural Comput & applic, 1993, 1, 224-228.

[254]Manwaring H. S. The use of an artificial neural network to improve precision in trace level, quantitative analysis of heavy metal pollutants. Proc. 4th Int. Conf. on Artificial Neural Networks. 1995, 375-380.

# Appendices

# Appendix 1. Published papers

(page numbering temporarily suspended)

Neural Comput & Applic (1994)2:168–178


Neural
Computing
& Applications

# A Comparison of Backpropagation and General Regression Neural Networks in Quantative Trace Metal Analysis using Anodic Stripping Voltammetry and Cybernetic Instrumentation

Howard S. Manwaring

Faculty of Technology and Computing, West Herts College, Watford, Hertfordshire, UK

*A comparison is made between backpropagation and general regression neural networks for the prediction of parts per billion lead concentration when used to process data obtained from digested curry powder by the electrochemical analysis method of differential pulse, anodic stripping at a thin film mercury electrode (TFME). Two data sets are used, one requiring the net to classify an unknown analytical data vector into one of a number of previously learnt concentrations, and one requiring the net to predict the probable concentration of an unknown sample by interpolation of the already learnt concentrations. For both of these data sets the general regression neural network is shown to train faster and to provide results superior to those obtained by backpropagation.*

**Keywords:** Neural network; General regression; Backpropagation; Anodic stripping voltammetry; Trace metal chemical analysis; Automated instrument

## 1. Background and Motivation

This work grew out of research into the development of a cybernetic instrument employing neural networks to automate the determination of trace levels (a few parts per billion) of heavy metals in solution by peak quantisation using anodic stripping. A computer controlled analytical instrument has been built and software written providing a user interface with access to a range of analytical procedures and user defined parameters. Current research is aimed at processing the resulting analytical data by a neural network to deliver direct concentration predictions for the various heavy metals detected. The network comparison given here forms part of that work. Some attempts have already been made at applying learning algorithms and neural networks to this and related fields. Wythoff [1] used a supervised network to recognise spectral peaks when they occurred against a noisy background. Sybrant [2] detected overlapping peaks in stationary electrode polarography using a pattern recognition method, and much work [3–6] has been done on the use of neural networks in interpreting IR and absorption spectra.

### 1.1. Electrochemical Analysis

The determination of trace quantities of heavy metals in such places as foodstuffs, ground water and soils is an increasingly important issue. Environmental pollution has far reaching consequences, and the trend is towards on-site determination using portable instrumentation. Many techniques are used to assess concentration levels of such substances as lead, cadmium and lithium. Differential pulse anodic stripping has the advantages of requiring minimal equipment and yet is capable of determining the concentration of various electroactive species in solution, even when the concentration of these species is very low, typically 1 part in $10^8$.

Anodic stripping has been well documented elsewhere. A very readable account is given by

*Original manuscript received 12 February 1994*

*Correspondence and offprint requests to*: H.S. Manwaring, 89 Kenilworth Road, Edgware, Middlesex HA8 8XA, UK.

Graabæk [7], and a more thorough overview of the technique is given by Copeland [8]. Pomeroy [9] extends this by describing the construction of the necessary electronics and hardware, and giving examples of typical results. An important paper by Rifkin [10] describes how a computer-based instrument was designed to provide flexibility in the analytical method, and shows the effects of changes in the various analytical parameters. Copeland [11] describes typical analyte preparation methods, and investigates optimum experimental conditions.

However, for the purposes of this paper a brief overview of the technique will be given here.

Every electroactive species undergoing an oxidation or reduction reaction, such as that occurring when a metal is plated or stripped off electrically, has a standard potential $E^{O}$ at which this occurs. For lead this is around $-0.12$ V with respect to the normal hydrogen electrode. If the potential between two electrodes in a solution containing the species is slowly changed from a value less than $E^{O}$ to a value greater than $E^{O}$, or *vice versa* in some cases, an increase in current will be observed at $E^{O}$ when the species oxidises into solution or reduces, i.e. plates, onto the electrode. The potential at which this increase occurs is characteristic of the metal being oxidised or reduced and the size of the increase, relative to the underlying background current, is indicative of the concentration of the metallic ions in the solution. For trace metal analysis these currents are typically in the region of $5 \times 10^{-9}$ A and change by a few nanoAmp during the analytical run.

In anodic stripping, the solution under test is placed in a specially constructed cell containing three electrodes. Two of these, the working electrode and the auxiliary electrode, act as cathode and anode as in a simple electroplating arrangement. The third is a specially constructed reference electrode providing a reference potential $V_{ref}$, with respect to which the potential of the working electrode can be measured. Feedback circuitry enables the potential of the working electrode to be accurately set and maintained at any value regardless of the current flowing through the solution. The potential of the working electrode, $V_{w.e.}$, and the current flowing into or out of it is measured during the analysis.

When using a TFME the working electrode is formed by electroplating a few microns of mercury onto a polished, vitreous carbon disc mounted in the end of a PTFE rod while it is rotating at about 2000 rpm. Figure 1 shows the arrangement of a TFME electrode built by the author for this work. Before the analytical run itself, the working electrode is held at a potential several tenths of a volt below $E^{O}$ for the metal being investigated. During this time the metal is plated onto the mercury, where it forms an amalgam. In the analysis phase the potential of the working electrode is gradually changed in an anodic direction using a variety of waveforms, until it is several tens of mV above $E^{O}$. Around $E^{O}$ the trace metal is stripped out of the mercury back into solution and the current increase occurs.

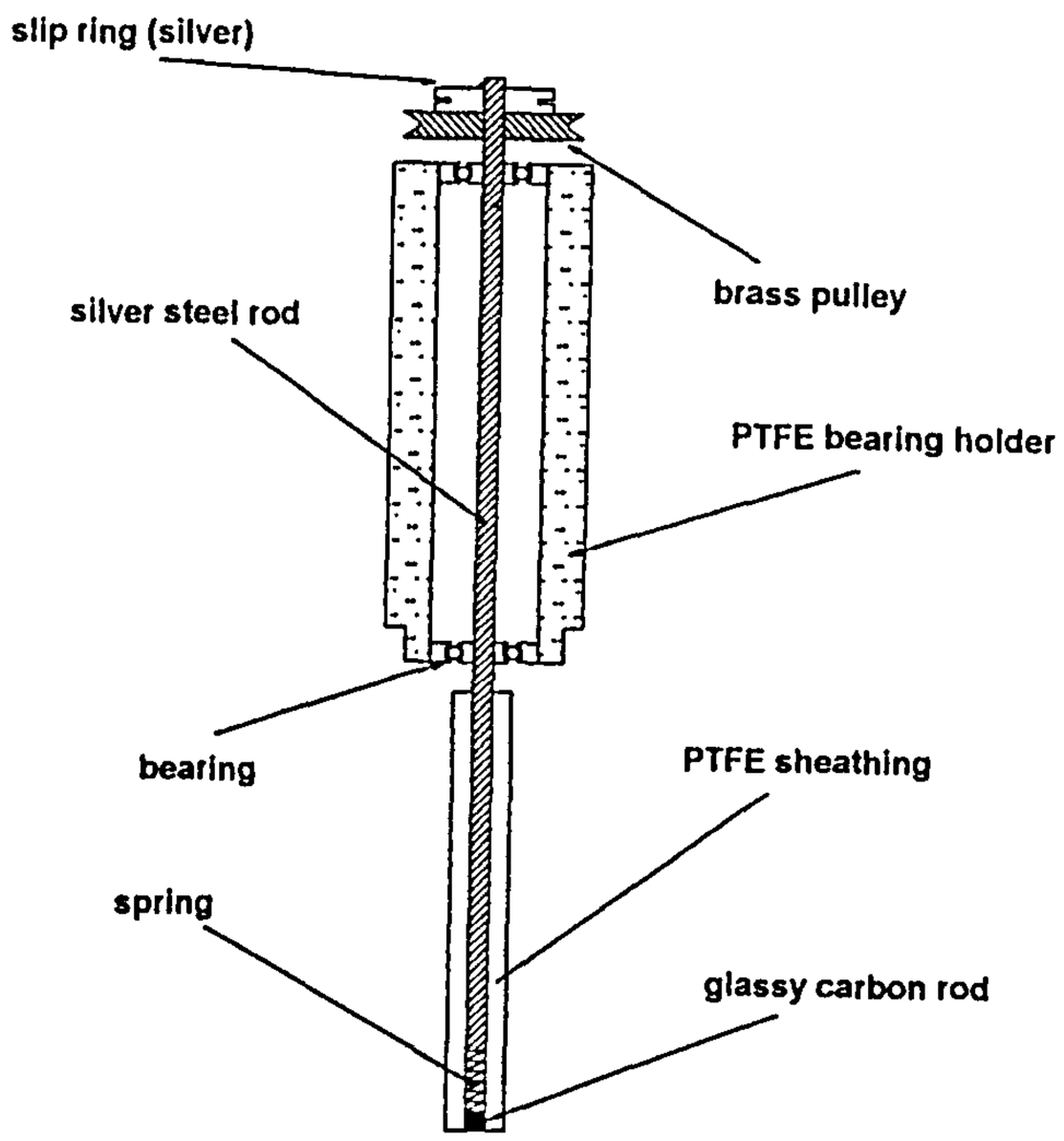



**Fig. 1.** Construction of a rotatable, glassy carbon, thin film mercury electrode built as part of the instrument developed by the author. The diameter of the carbon disk in the end of the PTFE rod is 2 mm and the overall length of the arrangement is 15 cm.

If differential pulse analysis is being performed, the waveform is a pulse of relatively large (50–100 mV) amplitude and short (10–20 ms) duration superimposed on a shallow potential ramp (~5 mV $s^{-1}$). The current flowing through the working electrode is measured just before the pulse is applied, and again just before the pulse ends. The difference in these two currents, $\Delta i_{w.e.}$, is recorded along with the corresponding ramp potential appearing at the working electrode, $V_{w.e.}$, at the time that the pulse was applied (see Fig. 2). The resulting curve of $\Delta i_{w.e.}$ against $V_{w.e.}$ is a stripping current peak superimposed on a roughly parabolic baseline due to the background current.

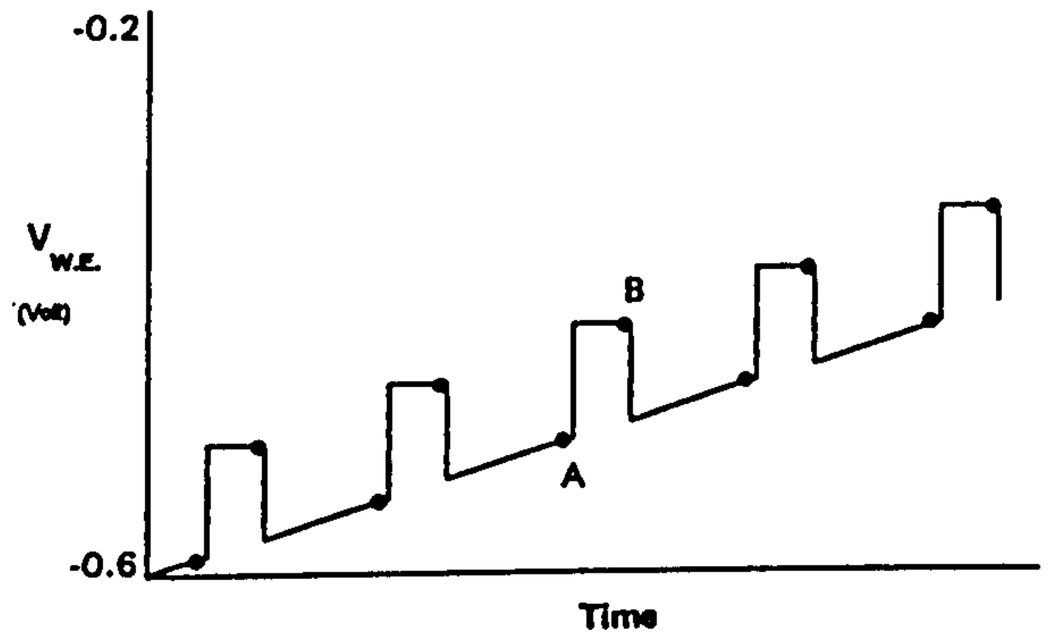



Fig. 2. Typical waveform applied to the working electrode during differential pulse anodic stripping. For each pulse the current is measured at points A and B, the difference between these values is recorded along with the potential at point A. Note that the scale of this waveform has been exaggerated for clarity.

## 1.2. The Problem

It can be shown [12] that the height of the $\Delta i_{w.e.}$ peak is given by the equation

$$(\Delta i_{w.e.})_{max} = \frac{nFAD_0^{1/2}C_0^*}{\pi^{1/2}(\tau - \tau')^{1/2}} \times \left(\frac{1-\sigma}{1+\sigma}\right) \qquad (1)$$

where

$$\sigma = \exp\left(\frac{nF}{RT}\frac{\Delta E}{2}\right) \qquad (2)$$

Here $n$ is the number of electrons transferred, $F$ is the Faraday constant, $A$ is the area of the mercury film, $D_0$ is the diffusion coefficient, $C_0^*$ is the bulk concentration of the species, $(\tau - \tau')$ is the pulse width, $R$ is the gas constant and $\Delta E$ is the pulse height.

The measured peak height above background ought, therefore, to be proportional to the bulk concentration for a given pulse height and width. Unfortunately, dissolved oxygen can give rise to a considerable and variable background current, organic compounds can produce a spurious signal, changes in sample pH can shift the peaks anodically or cathodically, and the signal to noise ratio is often very small. For low concentrations the height, or even existence, of any peak may be difficult to determine without several repetitions, each of which will be subject to the variations mentioned above.

Figures 3 and 4 are screen dumps from the instrument developed by the author. Figure 3 shows the result of analysing a blank solution of KCl and, for comparison, Fig. 4 shows the same solution with 10 parts per billion lead added. Smoother and flatter baselines can be obtained by careful sample preparation and deoxygenation.

The usual method of determining the concentration of a suspected species is by standard addition with repeated analytical runs. However, the purpose of this work was to determine if data from the raw solution could be used directly as input to a neural network and an estimate of the lead concentration made regardless of the above interferences and without time consuming sample preparation.

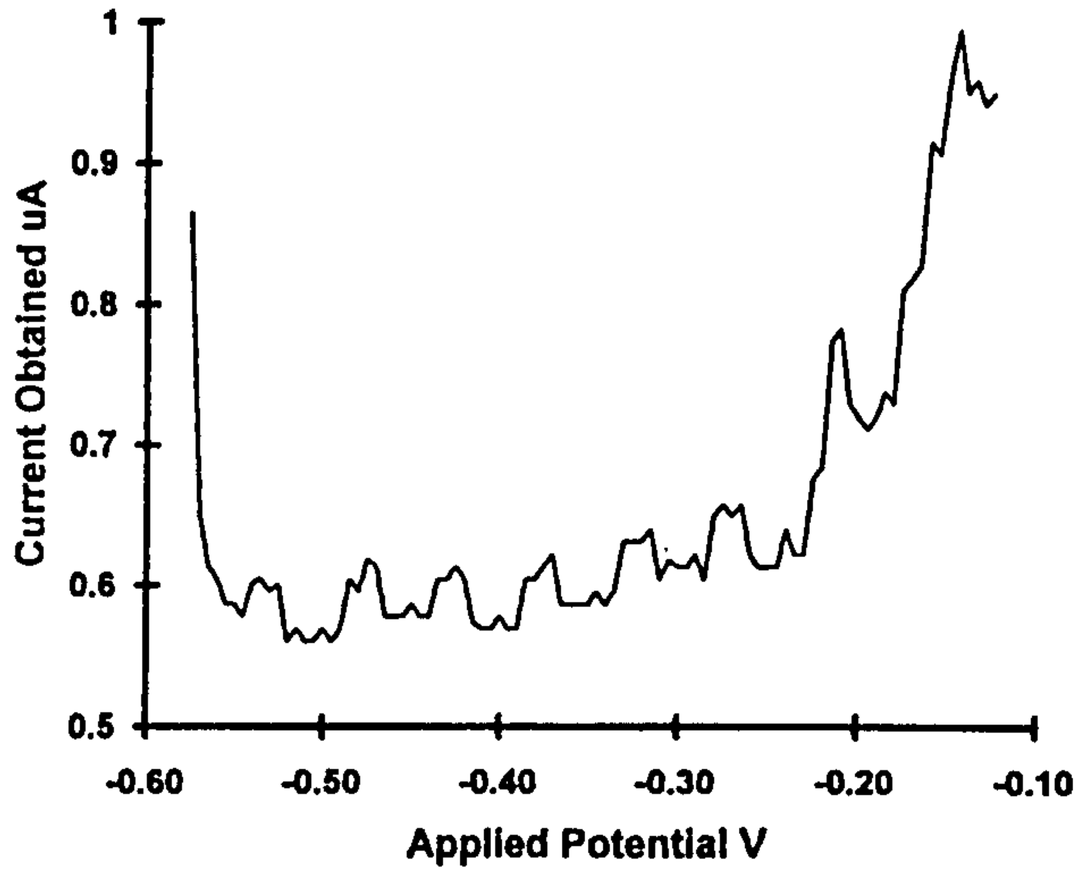



Fig. 3. Typical results obtained from performing differential pulse anodic stripping on a blank solution of KCl without deoxygenation.

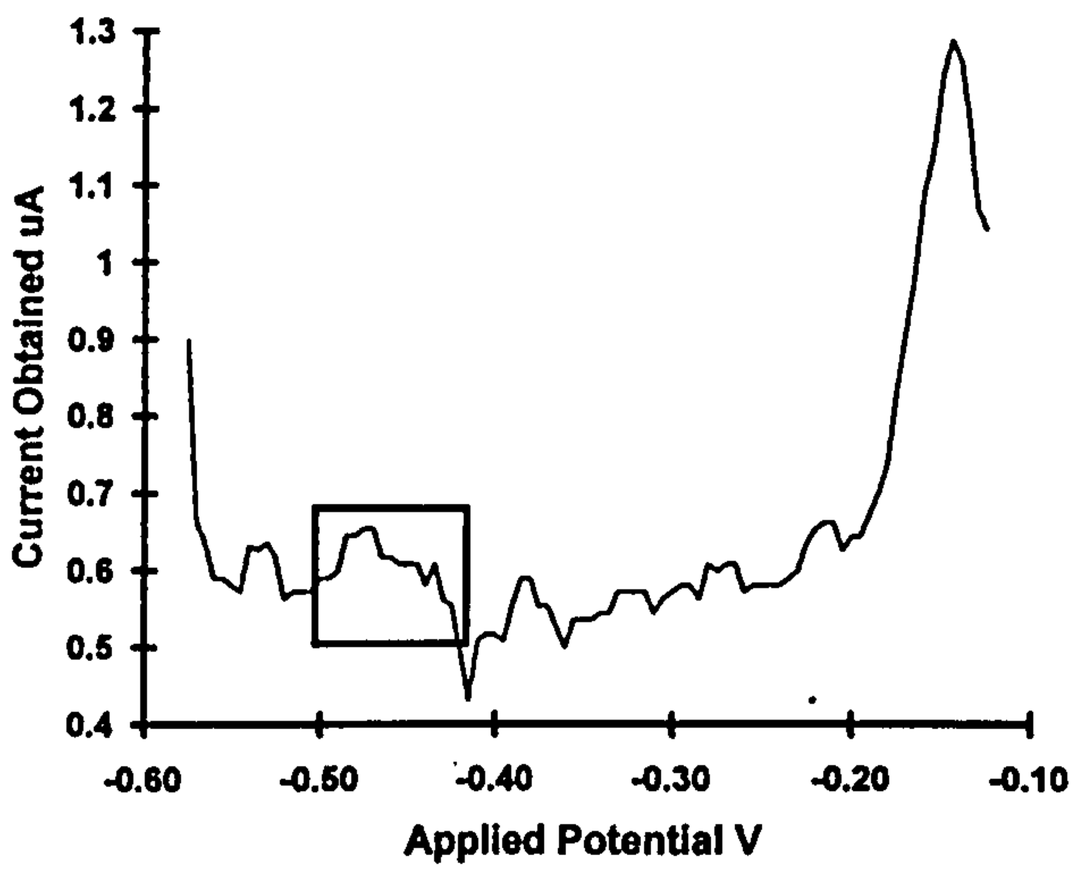



Fig. 4. Typical results obtained from performing differential pulse anodic stripping on the same oxygenated solution as in Fig. 2 but with 10 parts per billion lead added. The stripping peak caused by this lead is indicated by the dark square.

## 2. The Data

The system was used to determine the lead content of four different commercially obtained curry powders. This is a product that would naturally contain many organic compounds, but for which the maximum

allowed lead concentration is 20 ppm. The raw data from the instrument, consisting of up to 4000 pairs of pulse potential and resulting differential current measurements, is stored on disk by the analytical software.

### 2.1. Production

The analytical solutions used were obtained by digesting the curry powder in concentrated nitric acid at high temperature and pressure inside a PTFE container subjected to high energy microwave radiation. The resulting solutions were expected to have lead concentrations of about 5 parts per billion. The solutions were divided, and some had lead added in known amounts in order to give a range of nominal concentrations up to 500 ppb. Each sample was then analysed by atomic absorption spectroscopy (AAS) to provide a known concentration value. They were then re-analysed by differential pulse anodic stripping at a TFME using the system described above without deoxygenation or any other preparation. The particular selection of parameters used resulted in 400 values of $\Delta i_{w.e.}$ for each sample, corresponding to pulse potentials of between −0.6 V and −0.2 V in steps of 1 mV. Ten replicate runs were made for each sample, with a new mercury film being created for each run. Only the first 28 concentrations, ranging in lead concentration from 1–100 ppb were used in this study.

### 2.2. Preprocessing

The data set consisted of 280 data vectors each containing 400 $\Delta i_{w.e.}$ data points and representing 28 values of concentration, as determined by AAS. The target was a single number in the range 1–100 representing the concentration in ppb as measured by AAS. All these values were normalised across the entire pattern set by first calculating the maximum and minimum value for each data point and then adjusting the value of that point according to the equation

$$x_{i(new)} = [x_{i(old)} \times (x_{i(\max)} - x_{i(\min)})] + x_{i(\min)} \quad (3)$$

The AAS concentration value was similarly normalised.

### 2.3. Post Processing

The output of the network was converted to a concentration in ppb using the equation

$$C' = [N \times (C_{\max} - C_{\min})] + C_{\min} \quad (4)$$

where $N$ is the normalised output from the network, $C'$ is the network's prediction of the concentration, and $C_{\max}$, $C_{\min}$ are the maximum and minimum values of concentration in the data set as given by AAS.

### 2.4. The Two Data Sets Used

The entire data set was split into a training set and a test set. A third, validation set, was not employed at this stage as the intention was to produce a qualitative comparison of network architectures and not a finalised, working network. The methods employed in the extraction of the test set gave rise to the two data sets.

The first data set used for the networks, here called the *classification data set*, was obtained by sorting the entire set of data vectors in order of AAS stated concentration, and then extracting every fifth vector and placing it in the test set. The remaining vectors constituted the training set. In this way, the training set contained vectors from eight replicates of each of 28 concentrations and the test set contained two replicates (i.e. the 5th & 10th) from the same set of 28 concentrations. This selection method caused the network to act in a classification mode, where the test vector was a noisy representation of one of the (equally noisy) training vectors. This gave 224 training vectors and 56 test vectors with 28 possible outcomes.

The second data set used, called the *interpolation data set*, was obtained by sorting the data set in order of concentration as before, but instead extracting all ten replicate vectors for every third concentration value up to 100 ppb. This forced the network to act in a prediction mode and to interpolate between the input vectors for the learnt concentrations and produce a concentration prediction that did not map on to one of its learnt concentrations. This gave a training set of 200 vectors and a test set of 80 vectors with eight possible outcomes. It is recognised that these training set to test set ratios are rather high.

Input dimension reduction was investigated using several methods. Experiments in reducing inputs by extracting features, as shown in Table 1, showed some promise where the peak was well defined. In particular, considerable information appeared to be held in the left-hand slope of the peak and for higher concentrations reasonable network predictions could be made when just this one value was used as input.

For many of the curves, however, using extracted

Table 1. Examples of parameterisation methods used to reduce the dimension of the input vector.

| Parameters extracted | No. of network inputs |
|---|---|
| Left-hand valley minimum value, Right-hand valley minimum value, Peak maximum value, Fitted left-hand slope, Fitted right-hand slope, Fitted background slope on left, Fitted background slope on right | 7 |
| Coefficients of a quadratic fitted to the peak | 3 |
| Coefficients of a quadratic fitted to the background, Coefficients of a quadratic fitted to the peak | 6 |
| As above plus the height calculated from fitted peak to fitted background | 7 |
| Coefficients of a 5th order polynomial fitted to the entire curve | 5 |

features was too unreliable due to poor peak definition, and so as it was expected that the lead peak would occur at about −0.380 mV, the dimension of the input vector was reduced to 15 data points by simply starting at data point 115 and taking every tenth data point up to data point 255. This corresponds to the values of $\Delta i_{w.c.}$ for ramp potentials −0.485 to −0.345 in steps of 10 mV. The target remained as one continuously valued number.

## 3. Neural Network Architectures Investigated

Two network architectures were investigated: standard backpropagation and the general regression neural network. Investigation of the backpropagation network was subdivided by using two different numbers of neurons in the hidden layer. Each architecture was used with both training data sets.

### 3.1. Standard Backpropagation Network

This network is the familiar network formalised by Werbos [13] and elaborated first by Parker [14] and then McClelland [15]. It consisted of one input layer, one hidden layer and one output layer, fully connected between layers. The transfer function used for the hidden layer neurons was the sigmoid

$$f(x) = \frac{1}{(1 + \exp(-x))} \tag{5}$$

whilst the output layer used the linear function

$$f(x) = x \tag{6}$$

There were 15 inputs and 1 output. The number of hidden layer neurons for the first architecture investigated was 20, this being calculated from the heuristic

$$h = 0.5(i + o) + \sqrt{p} \tag{7}$$

where $h$ is the number of hidden neurons, $i$ and $o$ the number of inputs and outputs, respectively, and $p$ is the number of training patterns. This heuristic has been found to be better than the often used alternative one of

$$h = 2 \times \sqrt{i + o} \tag{8}$$

For the second architecture investigated, only three hidden neurons were used. No attempt was made to further optimise the number of hidden neurons.

The connection weights were initially set to random values in the range ±0.3. Throughout all the tests, a value of 0.1 was used for the learning rate β, a value of 0.9 was used for the momentum α and weight updates were performed after each pattern presentation. After every 200 presentations of a training pattern the entire test set was presented, and the mean squared error calculated. No weight updates were performed at this time. Each time the test set produced a new minimum value for the mean squared error the weights were saved to disk. Training was stopped when a considerable time had elapsed since the last minimum value, and the last saved set of weights was used as the trained network.

### 3.2. General Regression Neural Network

The general regression neural network belongs to the class of supervised networks, but has the advantage of being able to train quickly on sparse data sets and produce continuous valued outputs. It works by measuring how far in $N$ dimensional

space a given unknown input pattern is from the patterns in the training set, where $N$ is the number of inputs in the pattern. There are as many neurons in the hidden layer as there are training patterns. When a new pattern is presented to the network, the output that is predicted by the network is a proportional amount of all of the outputs in the training set. This proportion is based upon how far the new pattern is from the known patterns in the training set.

The regression of a dependent variable $Y$, on an independent variable $X$, is formed by computing the most likely value of $Y$ for each value of $X$. This computation is based upon a finite number of possible noisy measurements of $X$ and the corresponding value of $Y$.

Usually, it is necessary to know the underlying function, for example the linear function

$$Y = mX + c \tag{9}$$

and then estimate the coefficients $m$ and $c$, from which an estimate of $Y$ for a given $X$ can be calculated.

Specht [16] describes a method that expresses the functional form as a probability density function determined empirically from the observed data set, thus requiring no *a priori* knowledge of the underlying function. In essence, the network's estimate for $\hat{Y}(X)$ can be thought of as a weighted average of all observed values $Y_i$, where each observed value is weighted according to its distance from $X$.

Two ways of measuring the distance between two vectors can be used in the GRNN, the Euclidean distance and the City Block measure. These result in the following equations for $\hat{Y}(X)$:

Euclidean;

$$\hat{Y}(X) = \frac{\sum_{i=1}^{n} Y^i \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^{n} \exp\left(-\frac{D_i^2}{2\sigma^2}\right)} \tag{10}$$

where the scalar distance quantity $D_i^2$ is defined as

$$D_i^2 = (X - X^i)^T (X - X^i) \tag{11}$$

City Block;

$$\hat{Y}(X) = \frac{\sum_{i=1}^{n} Y^i \exp\left(-\frac{C_i}{\sigma}\right)}{\sum_{i=1}^{n} \exp\left(-\frac{C_i}{\sigma}\right)} \tag{12}$$

where the distance $C_i$ is defined as

$$C_i = \sum_{j=1}^{p} |X_j - X_j^i| \tag{13}$$

In both cases, $n$ is the number of sample observations and $p$ is the dimension of the vector variable $X$.

The parameter $\sigma$ is a smoothing parameter which affects the network's ability to generalise. A very large value tends to cause $\hat{Y}(X)$ to become the sample mean of the observed $Y^i$, as $\sigma$ approaches zero $\hat{Y}(X)$ assumes the value of $Y^i$ associated with the observation closest to $X$. At intermediate values, all values of $Y^i$ are taken into account but those corresponding to points closer to $X$ are given more weight.

Although the City Block distance can be computed faster than the Euclidean distance, it will be shown that this metric produces predictions that are not as accurate as using the Euclidean metric.

Throughout this work, the hidden layer neurons used the sigmoid transfer function given in Eq. (5) and the output neuron used a linear function (6). The connection weights were initially set to random values in the range ±0.3.

A search procedure was employed to find the optimum value for the smoothing parameter $\sigma$. This was done by repeatedly presenting the test set to the network using different values of $\sigma$ and measuring the error in the network's output. This process was halted when the value for the mean squared error could be improved no further. A typical plot showing this process for the interpolation data set is shown in Fig. 5.

## 4. Results and Discussion

The computer used for this work was a 40 MHz 80386-based machine fitted with 7 MBytes of RAM and a 80387 maths coprocessor. Both hand crafted and commercial neural network software was used, the commercial software being run under Windows v. 3.1 in enhanced mode.

After each network architecture had been trained on each data set the following statistics were gathered for both the training and test sets:

- the correlation coefficient between the predicted and target values,
- $R^2$, the coefficient of multiple determination between the predicted and target values. This is assumed to be a better measure of the closeness of actual and predicted values than the correlation coefficient,
- mean squared error in predicted values,
- mean absolute error in predicted values,

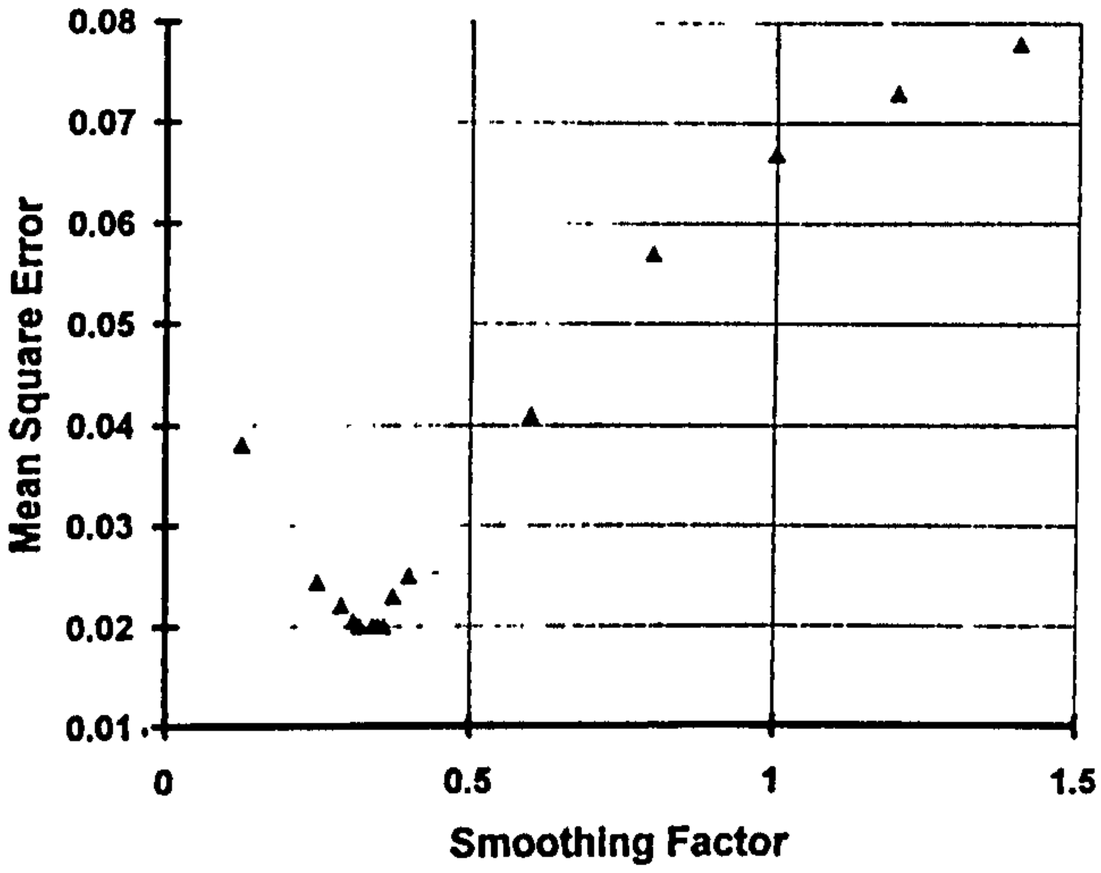



**Fig. 5.** Plot showing the progress of a search for an optimum GRNN smoothing parameter. This was using a Euclidean distance metric on the interpolation test data set.

- min absolute error in predicted values,
- max absolute error in predicted values.

In addition, a scatter graph was obtained of the predicted lead concentration plotted against the target concentration for the training set and test sets. Only the test set plots are reproduced here; the training sets exhibiting similar relationships but, as might be expected, having a higher correlation between predicted and target values. The numeric results are presented in Tables 2–7, where the absolute errors are measured in parts per billion.

### 4.1. The First Data Set – Classification

**Backpropagation.** The first architecture with 20 hidden layer neurons reached a minimum mean squared error value for the test set of 0.0437. This was achieved after 5366 epochs and took 3 hrs 40 min (an epoch being one presentation of the entire training set).

The second architecture with three hidden neurons reached a minimum mean squared error for the test set of 0.0516 after 42 mins and needed 2548 epochs.

Both exhibit considerable scatter on the test set predictions with the 15:20:1 network (Fig. 6), being marginally better than the 15:3:1 (Fig. 7), although it had a larger maximum absolute error of 39.4 ppb as opposed to 36.1 ppb for the 15:3:1. Both networks incorrectly allocate a large number of patterns to the same concentration, particularly at low values (<20 ppb).

**GRNN.** Using the City Block metric (Fig. 8), the correlation between predicted and target values is much better than either of the backpropagation architectures, and shows less scatter. The maximum absolute error is lower (29.0 ppb), and the low end of the concentration range is better separated. This network took a total of 2 min to train including a search of 25 candidate values for the smoothing parameter.

Using Euclidean distances (Fig. 9), there is an almost perfect match between predicted and target values. Very few patterns are incorrectly classified, and the maximum absolute error of those that were is only 26.5 ppb. The time taken was 1 min 50 s, including a search of 22 smoothing parameter values.

Note: The curve of $\Delta i_{w.e.}$ against $V_{w.e.}$ corresponding to the misclassified concentrations were subsequently found to have large variations in curve shape within the ten replicates.

### 4.2. The Second Data Set – Interpolation

**Backpropagation.** After 8 min of training the minimum error in the test set was reached. The 15:3:1 architecture once again shows considerable scatter (see Fig. 10), even producing a prediction of a negative lead concentration. The maximum absolute error, at 33.7 ppb, is better than for the classification data set, but still poor. The network appears oversensitive to variations in the replicate $\Delta i_{w.e.}$ vs. $V_{w.e.}$ curve shapes, and produces a different prediction for each one.

**Table 2.** Results using backpropagation (15:20:1), with test set and classification data.

| File processed | *R* squared | Mean avg squared error | Mean absolute error | Min absolute error | Max absolute error | Correlation coefficient |
|---|---|---|---|---|---|---|
| Entire data | 0.9016 | 95.378 | 6.833 | 0.063 | 40.556 | 0.950 |
| Training | 0.9046 | 92.451 | 6.657 | 0.063 | 40.556 | 0.951 |
| Test | 0.8895 | 107.008 | 7.540 | 0.150 | 39.381 | 0.945 |

**Table 3.** Results using backpropagation (15:3:1), with test set and classification data.

| File processed | *R* squared | Mean avg squared error | Mean absolute error | Min absolute error | Max absolute error | Correlation coefficient |
|---|---|---|---|---|---|---|
| Entire data | 0.8832 | 113.131 | 7.873 | 0.002 | 43.994 | 0.941 |
| Training | 0.8867 | 109.760 | 7.724 | 0.040 | 43.994 | 0.943 |
| Test | 0.8693 | 126.615 | 8.470 | 0.002 | 36.175 | 0.932 |

**Table 4.** Results using GRNN (City Block metric), with test set and classification data.

| File processed | *R* squared | Mean avg squared error | Mean absolute error | Min absolute error | Max absolute error | Correlation coefficient |
|---|---|---|---|---|---|---|
| Entire data | 0.9830 | 16.514 | 1.398 | 0.000 | 41.506 | 0.992 |
| Training | 0.930 | 67.838 | 4.696 | 0.000 | 30.662 | 0.965 |
| Test | 0.9041 | 92.894 | 5.782 | 0.000 | 29.016 | 0.953 |

**Table 5.** Results using GRNN (Euclidean), with test set and classification data.

| File processed | *R* squared | Mean avg squared error | Mean absolute error | Min absolute error | Max absolute error | Correlation coefficient |
|---|---|---|---|---|---|---|
| Entire data | 0.9901 | 9.593 | 0.711 | 0.000 | 26.522 | 0.995 |
| Training | 0.9956 | 4.295 | 0.438 | 0.000 | 23.136 | 0.998 |
| Test | 0.9682 | 30.784 | 1.802 | 0.000 | 26.522 | 0.984 |

**Table 6.** Results using backpropagation (15:3:1), with test set and interpolation data.

| File processed | *R* squared | Mean avg squared error | Mean absolute error | Min absolute error | Max absolute error | Correlation coefficient |
|---|---|---|---|---|---|---|
| Entire data | 0.7813 | 211.859 | 11.147 | 0.017 | 37.778 | 0.900 |
| Training | 0.7865 | 214.773 | 11.042 | 0.017 | 37.778 | 0.921 |
| Test | 0.7535 | 204.575 | 11.409 | 0.026 | 33.747 | 0.889 |

**GRNN.** This again trained in 2 min after a search of 25 smoothing parameter values. As with classification, the GRNN shows much less scatter than the backpropagation network, and is less sensitive to variations in curve shape within replicates (see Fig. 11). The maximum absolute error (26.6 ppb) is lower than for backpropagation performing the same task.

## 5. Conclusions

It can be seen that the general regression neural network is superior in its performance to the backpropagation network when applied to the task of predicting lead concentration using data from $\Delta i_{w.e.}$ against $V_{w.e.}$ curves obtained from differential

**Table 7.** Results using GRNN (Euclidean), with test set and interpolation data.

| File processed | *R* squared | Mean avg squared error | Mean absolute error | Min absolute error | Max absolute error | Correlation coefficient |
|---|---|---|---|---|---|---|
| Entire data | 0.6627 | 326.820 | 14.814 | 0.000 | 47.156 | 0.820 |
| Training | 0.6244 | 377.904 | 15.765 | 0.000 | 47.156 | 0.812 |
| Test | 0.7601 | 119.109 | 12.435 | 1.051 | 25.687 | 0.874 |

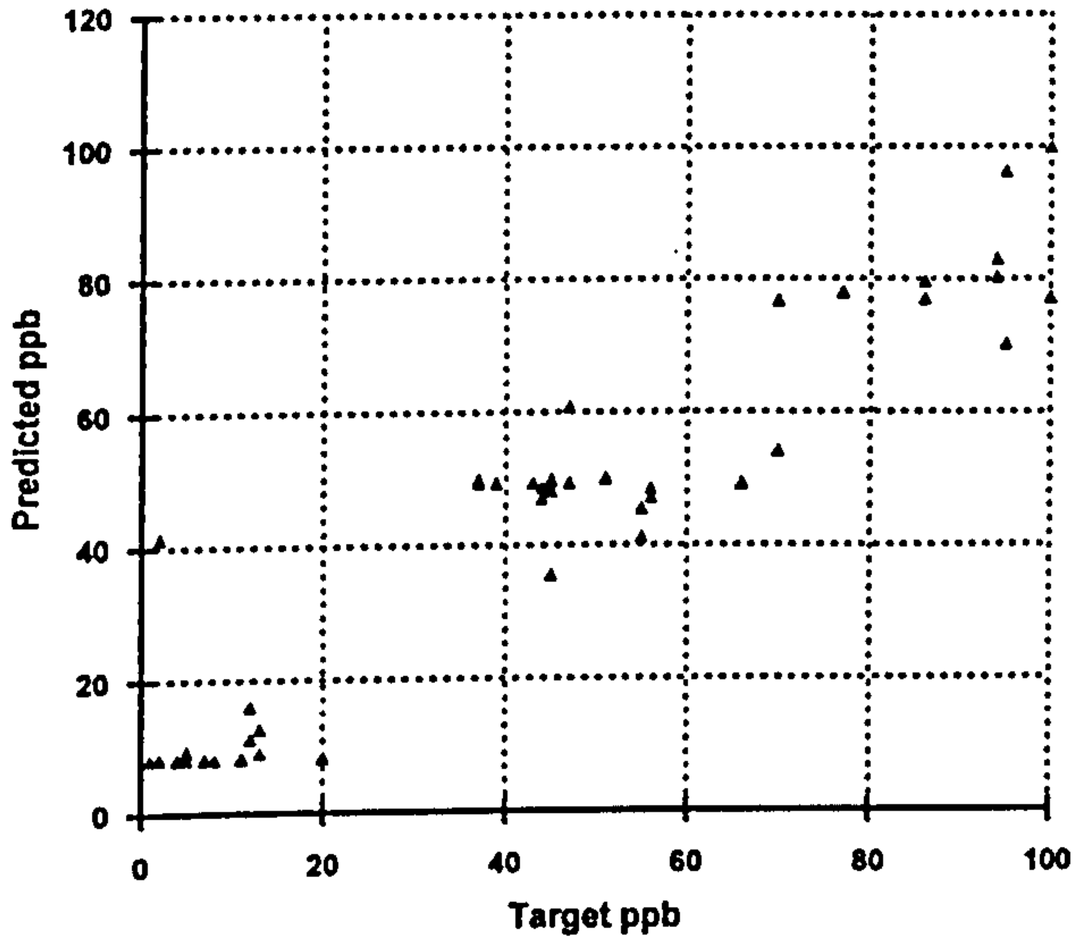



Fig. 6. Scatter plot for 15:20:1 backpropagation network, test set, classification data.

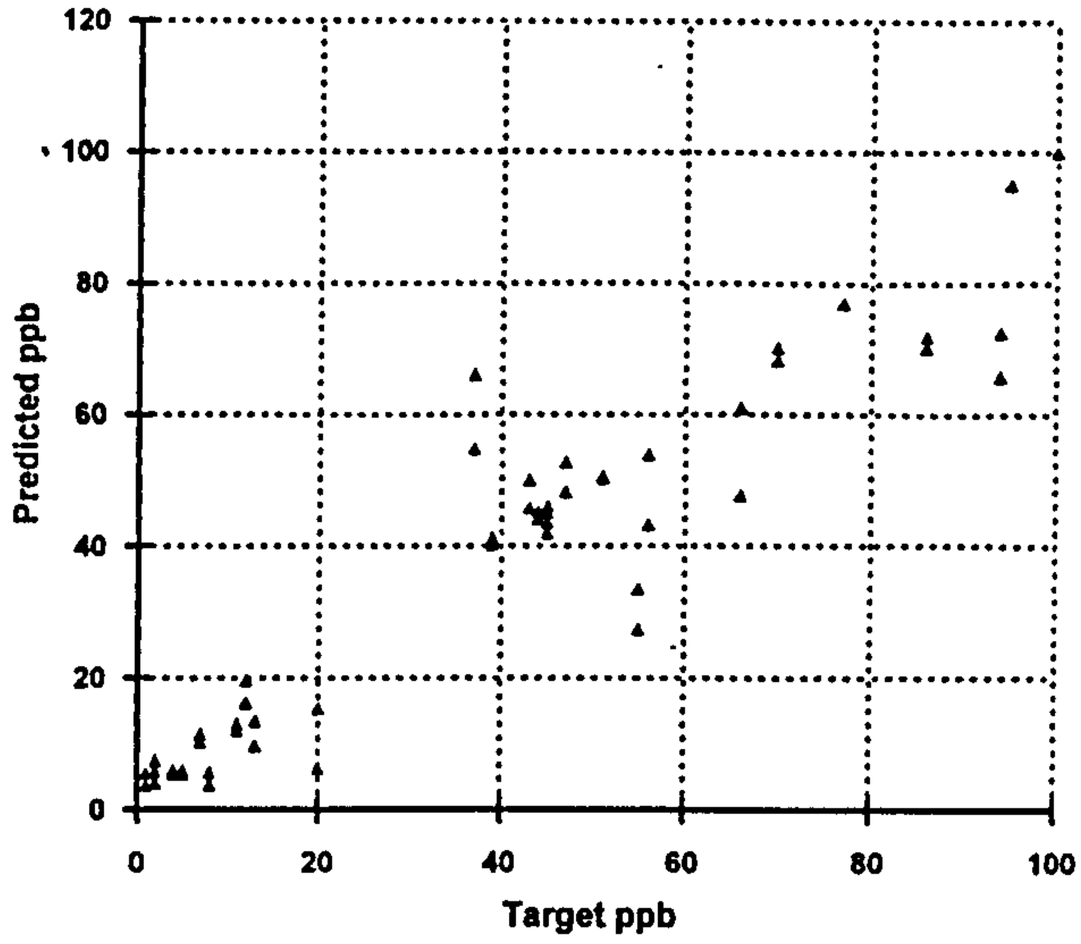



Fig. 8. Scatter plot for GRNN using City Block metric, test set, classification data.

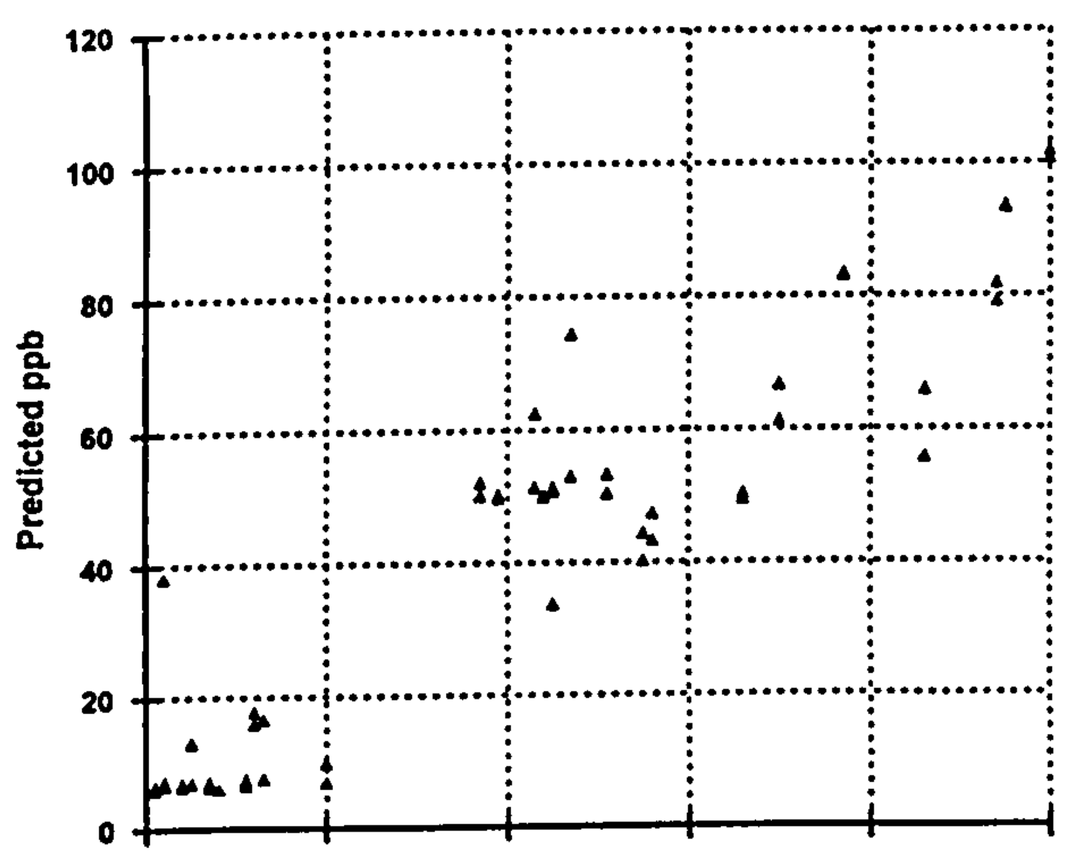



Fig. 7. Scatter plot for 15:3:1 backpropagation network, test set, classification data.

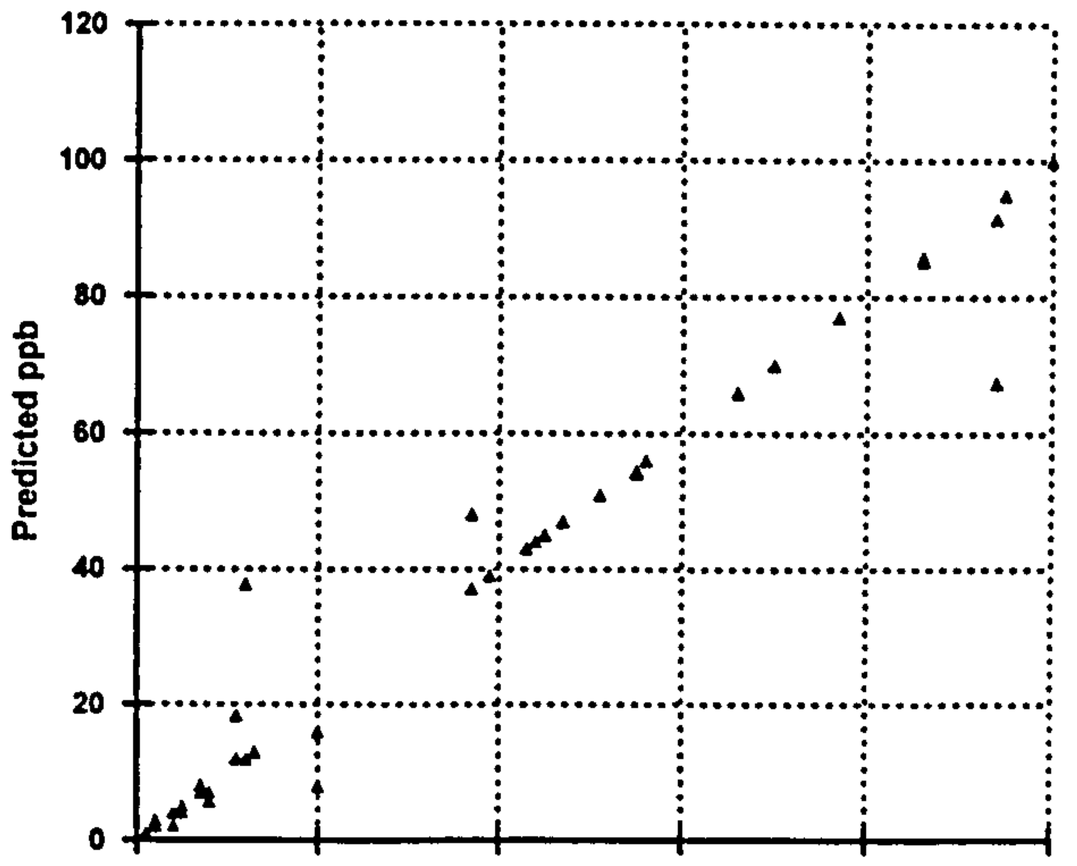



Fig. 9. Scatter plot for GRNN using Euclidean metric, test set, classification data.

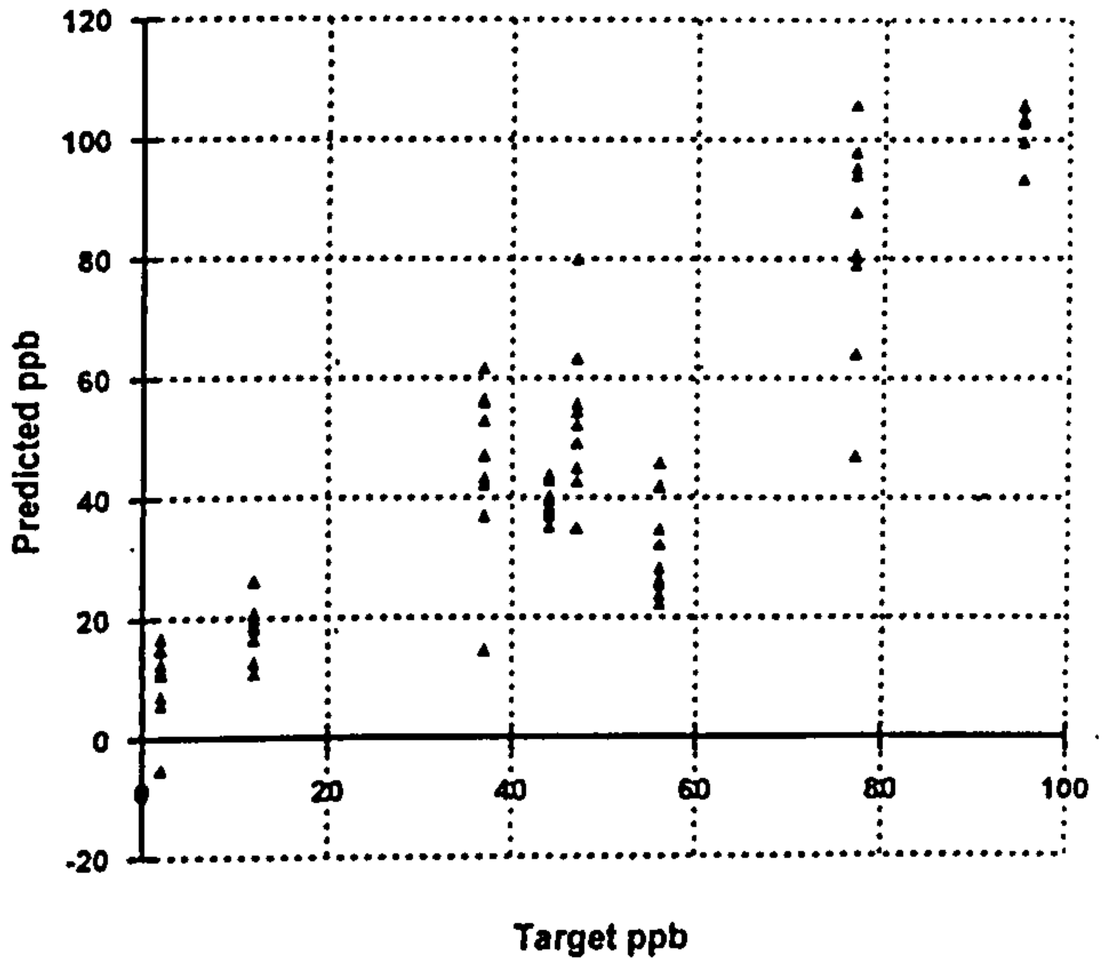



**Fig. 10.** Scatter plot for 15:3:1 backpropagation network, test set, interpolation data.

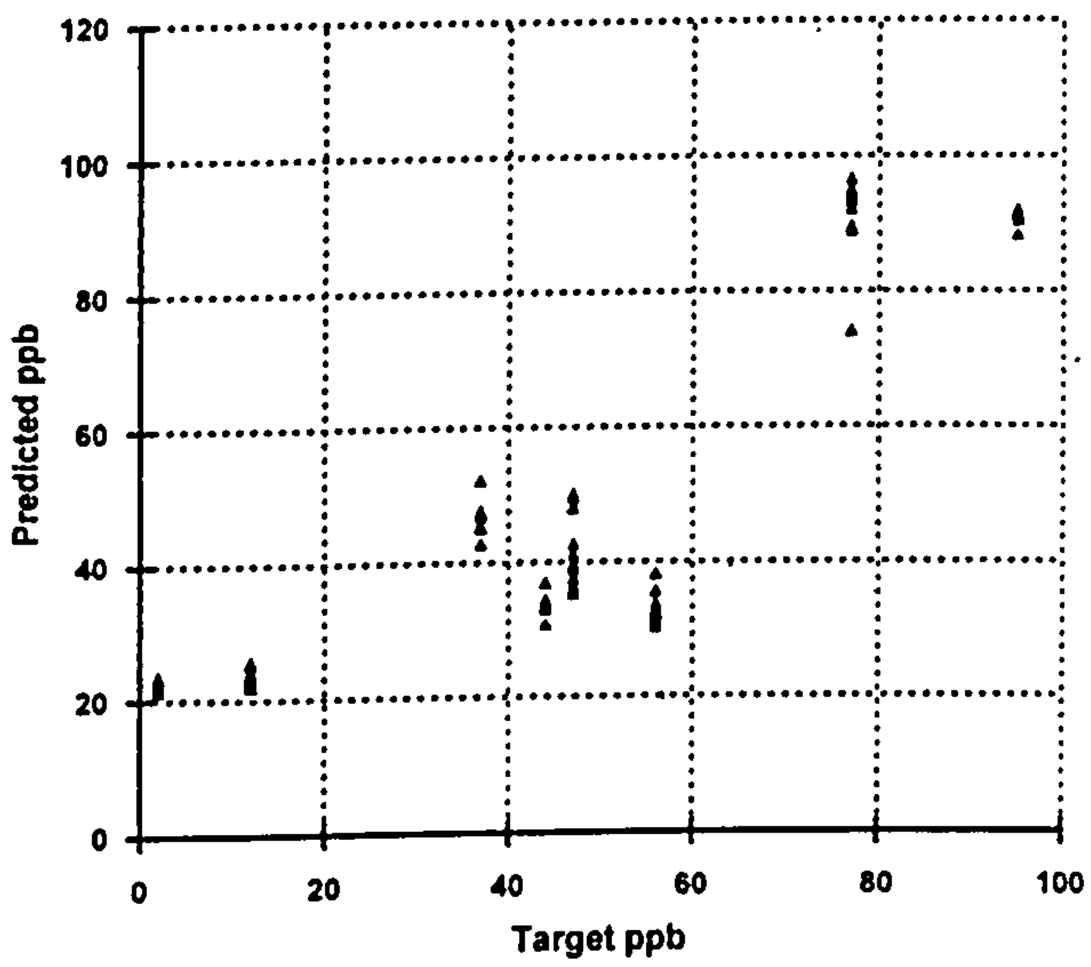



**Fig. 11.** Scatter plot for GRNN using Euclidean metric, test set, interpolation data.

pulse anodic stripping, even when the number of data points in the curve is severely reduced.

The GRNN is faster to train, exhibits less scatter, produces a lower maximum absolute error and is more tolerant of noise or random fluctuations on the input data.

The statistical *t*-test performed on the analytical data indicated that ±20 parts per billion is the minimum resolution that could reasonably be expected from this analytical data. The maximum absolute error value of 26.6 parts per billion, obtained by the GRNN, compares well with this figure.

## 6. Further Work

More work clearly needs to be done using larger data sets before a trained network can be expected to perform reliably in the field. Once this is done, the trained network will be incorporated into the analytical software already produced so that the entire process from analysis to concentration prediction is automated. The research will also to expand to consider the quantitative prediction of multiple species simultaneously using this same analytical technique.

**Acknowledgements.** I would like to thank West Herts College for the support they have given me by providing me with the resources and time to carry out this work.

The commercial neural network software used for this work was NeuroShell 2®, produced by Ward Systems Group inc. of Frederick, MD, USA.

Windows® is a registered trademark of the Microsoft Corporation.

## References

1. Wythoff BJ, Levine SP, Tomellini SA. Spectral peak verification and recognition using a multilayer neural network. Anal Chem 1990; 62: 2702–2709.
2. Sybrandt LB, Perone SP. Computerized pattern classification of strongly overlapped peaks in stationary electrode polagraphy. Anal Chem 1972; 44: 2331–2339.
3. Munk ME, Madison MS, Robb EW. Neural network models for infrared spectrum interpretation. Mikrochim Acta 1991; 11: 505–514
4. Borggaard C, Thodberg HH. Optimal minimal neural interpretation of spectra. Anal Chem 1992; 64: 545–551
5. Gemperline PJ, Long JR, Gregoriou VG. Nonlinear multivariate calibration using principal components regression and artificial neural networks. Anal Chem 1991; 63: 2313–2323
6. Blank TB, Brown SD. Nonlinear multivariate mapping of chemical data using feed-forward neural networks. Anal Chem 1993; 65: 3081–3089
7. Graabæk AM, Jeberg B. Trace element analysis by computerised stripping potentiometry. Int Lab 1992; 33–38
8. Copeland TR, Skogerboe RK. Anodic stripping voltammetry. Anal Chem 1974; 46: 1257A–1268A
9. Pomeroy RS, Denton MB, Armstrong NR. Voltammetry at the thin-film mercury electrode. J Chem Educ 1989; 66: 877–880
10. Rifkin SV, Evans DH. Analytical evaluation of differential pulse voltammetry at stationary electrodes using computer based instrumentation. Anal Chem 1976; 48: 2174–2179
11. Copeland TR, Christie JH, Osteryoung RA, Skogerboe RK. Analytical applications of pulsed stripping

and thin film mercury electrodes. Anal Chem 1973; 45: 2171–2174
12. Bard AJ. Controlled potential microelectrode techniques – potential step methods. In: Bard AJ. Electrochemical Methods. Wiley, 1980, pp. 136–212
13. Werbos P. Beyond regression: New tools for prediction and analysis in the behavioral sciences. PhD thesis, Harvard, Cambridge, MA, August 1974
14. Parker DB, Learning Logic. Technical Report TR-47, Centre for Computational Research in Econics and Management Science, MIT, Cambridge, MA, April 1985
15. McClelland J, Rumelhart D. Explorations in Parallel Distributed Processing; Vols 1 and 2. MIT Press, 1986
16. Specht DF. A general regression neural network. IEEE Trans Neural Net 1991; 2: 568–578

# THE USE OF AN ARTIFICIAL NEURAL NETWORK TO IMPROVE PRECISION IN TRACE LEVEL, QUANTITATIVE ANALYSIS OF HEAVY METAL POLLUTANTS

**H. S. Manwaring**

**West Herts College (Associate College of the University of Hertfordshire), England**

## ABSTRACT

The author has used various neural networks to process the response obtained from an electroanalytical technique used for the analysis of trace metal pollutants in liquids. A previous paper, Manwaring ([1]), compared the capabilities of the GRNN and MLP in this respect. In this paper it is shown that using the neural network to make predictions of unknown sample concentrations shows an improvement, by a factor of about two, on the mean absolute error and the prediction confidence when compared with a traditional, calibration curve technique. In addition the neural network method is shown to produce reliable predictions even with instrumental responses that are completely unsuitable for traditional processing.

## INTRODUCTION

The analysis of river water for environmental heavy metal pollutants such as lead may be carried out by an electrochemical method known as Differential Pulse Anodic Stripping Voltammetry (DPASV), to be described later. The response of an instrument performing this type of analysis, known as a voltammogram, may be viewed as an asymmetric peak of a few μA amplitude superimposed upon a roughly quadratic background. Figure 1 is a screen dump from an instrument built by the author and shows the response obtained when used to analyse a solution containing lead at a concentration of 10 parts per billion.

A common way of calibrating such an instrument is by the construction of a calibration curve based upon the relationship between the height of the stripping peak above the interpolated background and the known metal concentration in a set of standard samples. Unfortunately, the height of the peak, and to a lesser extent its position, is dependent not only on the concentration of pollutant by also upon several instrumental parameters which also affect the sensitivity and resolution of the instrument. In order to obtain meaningful results therefore, the scans used to construct the calibration curves must either be made using only one set of operational parameters or the peak height obtained must be scaled in some way to take into account the different parameters used.

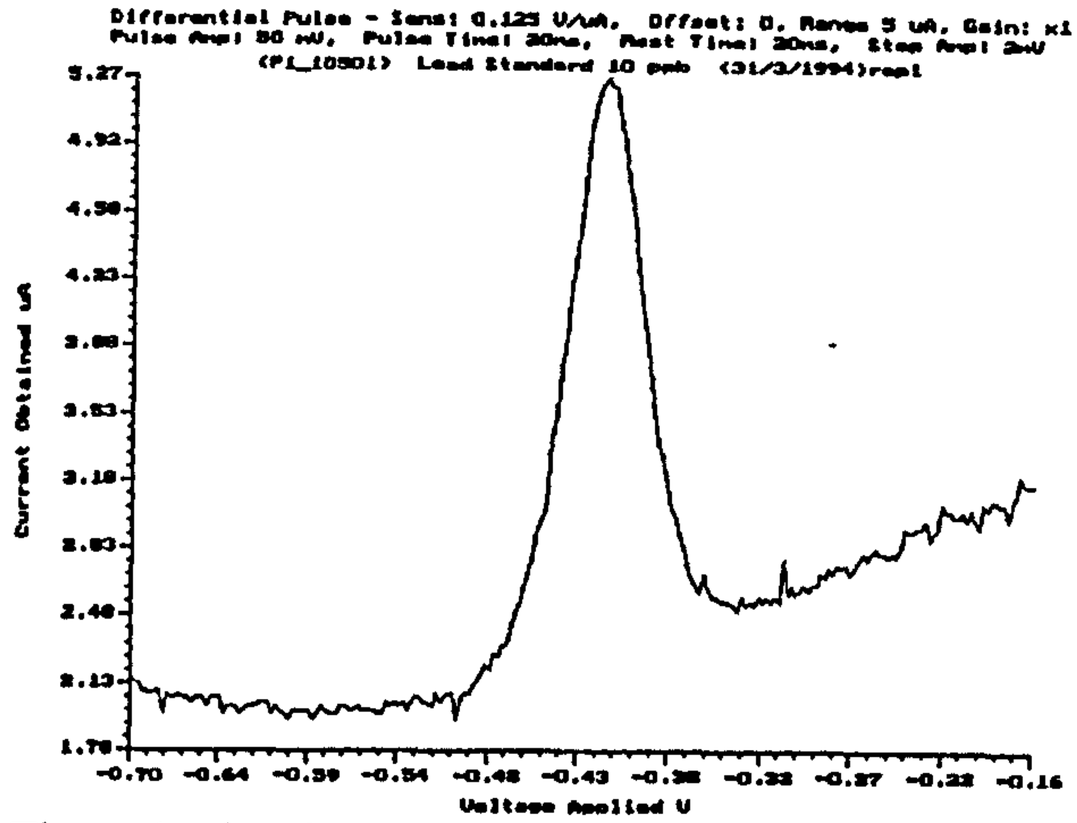



Figure 1: A voltammogram from a typical DPASV scan

If the optimum DPASV operational parameters for a given analyte sample are unknown, as is usually the case, it is possible inadvertently to saturate the instrument. In this case a flat topped peak is produced, figure 2, for which the true peak height cannot be found. These responses cannot be used in a calibration curve technique. Scans must be repeated several times using different parameters until a good peak is obtained.

The author has built a computer controlled instrument which performs an analysis by DPASV at the parts per billion level. A neural network is trained on responses obtained from calibration samples which have been analysed using a number of different operational parameters. The inputs used to the neural network are computed from both the instrumental response and the various parameters in effect at the time of the analysis. The single output is a direct quantitative measure, in ppb, of the concentration of metal in the original sample.

The neural network approach is able accurately to model the response / concentration relationship and also eliminates the need for repeat scans as it suffers only a very slight loss of precision and accuracy when estimating from a saturated, flat topped response. Indeed, the performance achieved with such responses is better than the calibration curve method achieves using good peaks.

*'Artificial Neural Networks', 26th-28th June 1995, Conference Publication No. 409, © IEE 1995*

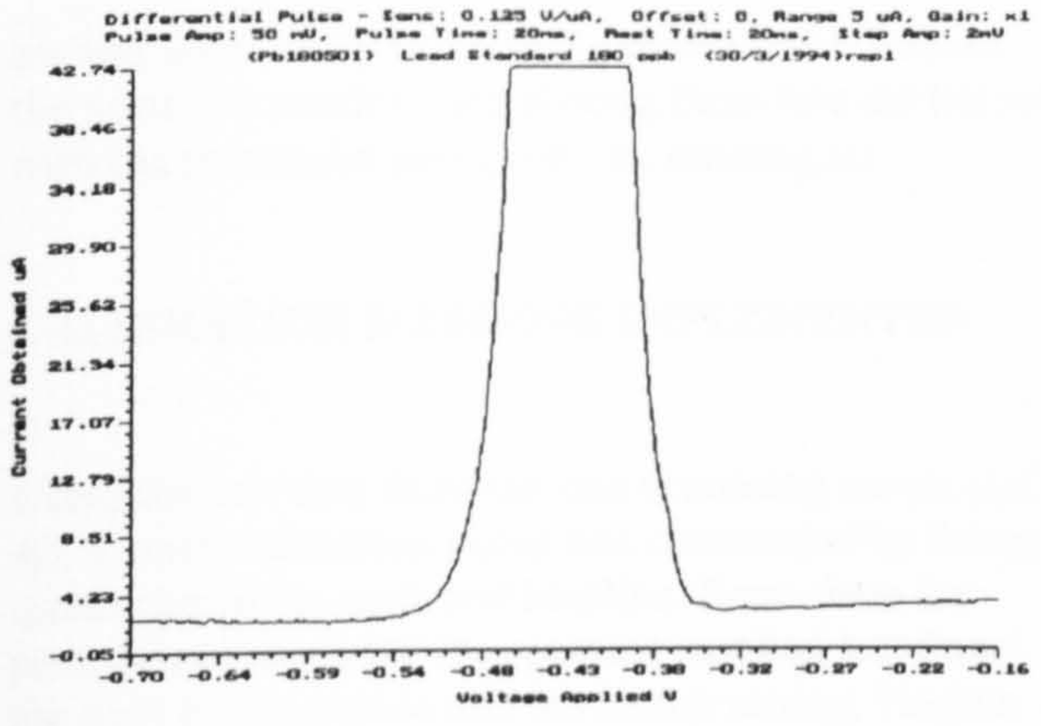



Figure 2: A saturating DPASV voltammogram

## ANODIC STRIPPING VOLTAMMETRY

Differential Pulse Stripping Voltammetry is very well described in Wang ([2]), Kopanica ([3]) and Brainina ([4]). The use of true neural networks in analytical chemistry appeared in 1990 when Long ([5]) trained a back propagation network to estimate the protein content of wheat from its near infrared spectrograph. Wythoff ([6]) used a simple back propagation network to identify signal peaks in infrared spectra whilst McAvoy ([7]) found that the backpropagation MLP out performed linear techniques in de convoluting fluorescence spectra. Blank ([8], [9]) used both real and synthetic IR data to compared non-linear forms of Partial Least Squares to a Back Propagation neural network. To date however little seems to have been written regarding the use of neural networks with DPASV.

In DPASV the solution under test is placed in a specially constructed cell containing three electrodes. Two of these, the Working Electrode, which is formed by electroplating a few microns of mercury onto a polished, vitreous carbon disc, and the Auxiliary Electrode, a fine platinum wire, act as cathode and anode as in a simple electroplating arrangement. The third is a specially constructed reference electrode providing a reference potential with respect to which the potential of the Working Electrode can be measured. Feedback circuitry enables the potential of the Working Electrode to be accurately set and maintained at any value regardless of the current flowing through the solution.

Before the analytical run the Working Electrode is rotated and held at a potential of about -0.6V for a short time $t_p$. During this time any metallic pollutants are plated onto the mercury where they form an amalgam. In the analysis phase the potential of the Working Electrode is gradually changed in an anodic direction, using a ramped, pulse shaped wave form of pulse height $E_p$ and variable mark - space ratio, until it is at about -0.2V. Around the standard potential for the trace metal in question, the metal is stripped out of the mercury, i.e. oxidised, back into solution and a current increase occurs due to the Faradic reaction. The potential of the Working Electrode ($E_{w.e.}$) at the time of each pulse and the difference in current flowing ($\Delta i$) just before the pulse and at the end of the pulse provides the instrumental response.

## DATA SET PRODUCTION

The longer $t_p$ is the more metal is plated and the higher the response peak obtained. Also, within certain limits, the higher the pulse height the greater the sensitivity i.e. the higher the peak, but the lower the resolution. For this study standard solutions containing lead at concentrations of 0 to 200 ppb in steps of 10 ppb were made up by dissolving lead nitrate in nitric acid. Four scans were made of each solution using DPASV at pulse heights ($E_p$) of 5, 10, 20 and 50 mV and a deposit time ($t_p$) of 60s. Two replicate scans were carried out. This entire process was then repeated using new solutions and with a deposition time of 180s. This provided 320 voltammograms i.e. 16 analytical runs on each of 20 concentrations. In addition two replicate scans of 0 ppb lead concentration was made at each of the four pulse heights using a 180s deposit time. The resulting database therefore consisted of 328 voltammograms each of which contained 300 $\Delta i$ data points. Figure 2 shows a 3D plot of the data set showing the peak height for all concentrations and parameter combinations. Saturating peaks are in the top left and top centre of the diagram.

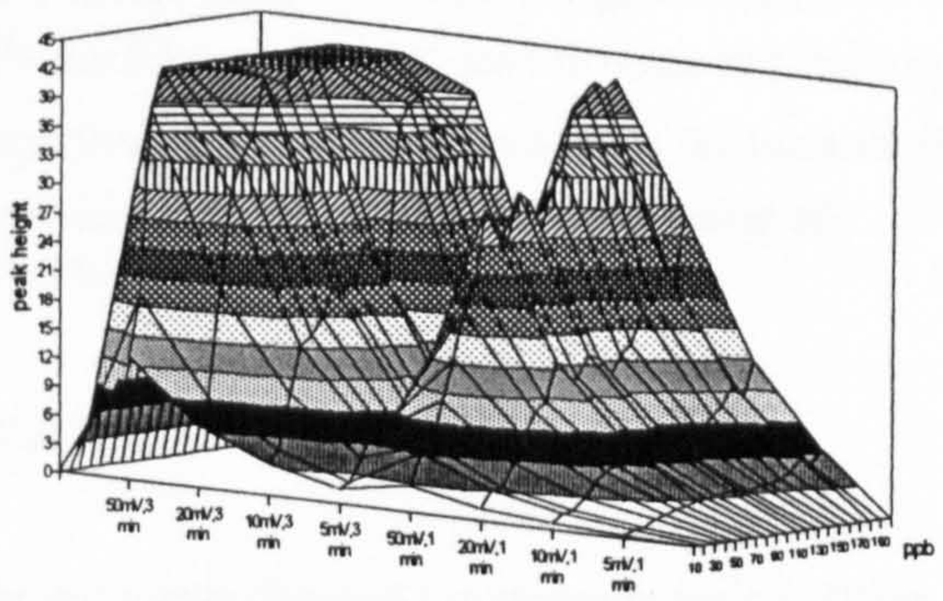



Figure 2: 3D plot showing the surface being modelled.

The data set was split into three groups. The first group (set A) contained curves that did not exhibit saturation, the second group (set B) contained the saturating curves with flat topped responses and the third group (set C) was obtained by using all of the original data vectors regardless of whether they saturated or not. These three groups were then each split into a training

set and a test set by extracting approximately 25% of the vectors at random and placing them into the test set with the remainder going into the training set.

## CALIBRATION METHODS IMPLEMENTED

Using the raw data from the non saturating curves (set A), a 'base' calibration curve was constructed by fitting quadratics to the peak and baseline. From these the peak height above the interpolated baseline was found for each concentration and parameter setting. Heights from replicate curves were averaged. These heights were then normalised to a 1 minute plate time ($t_p$) at a 5 mA pulse height ($E_p$). Using the training set only, the normalised mean peak height was plotted against the known lead concentration and a linear least squares regression performed in order to produce a calibration curve. This calibration curve was then used to obtain an estimate of the lead concentration for each of the averaged and normalised test set heights in the non saturating data set.

For the neural network models the raw $\Delta i$ data, with each response consisting of 300 data points, was reduced to a 17 element input vector by extracting the $\Delta i$ values corresponding to working electrode potentials of -0.6V to -0.26V in steps of 20 mV. The vector was then extended to either 19 components by appending the two operational parameters of plate time ($t_p$) and pulse height ($E_p$) or to 22 components by also appending the three calculated values of maximum $\Delta i$, minimum $\Delta i$ and the standard deviation of the $\Delta i$ points in the original instrument response. The 17 values of $\Delta i$ were then scaled to values approximately in the range 0 to 1 using the formula

$$x' = \frac{x}{42.8} \qquad (1)$$

where 42.8 is the maximum possible $\Delta i$ value, in μA, that can be recorded. The remaining components of the vector were each normalised to the range 0-1 individually using their particular maximum and minimum. The single target was the concentration in ppb, similarly scaled to lie in the range 0 to 1.

## NEURAL NETWORK ARCHITECTURES USED

Various neural network architectures having 22 or 19 inputs were investigated including GRNN, two layer MLPs, three layer MPLs and a two layer high order MLP using 138 inputs constructed from all possible multiples of input pairs. In addition a model was constructed using Partial Least Squares as this has been shown, Wold ([10],[11]), to be of some benefit with this type of chemical response.

All the three layer MLPs used a sigmoidal transfer function for the hidden layer and all models used a linear transfer function for the single output unit. The optimum number of hidden units for the three layer architectures was found empirically to be 30.

For each of the models investigated, including the peak height calibration curve, a plot was made of the actual lead concentration against the predicted lead concentration for the unseen test set.

The quality of prediction for each model was then obtained by calculating various metrics of which two, the mean absolute error in prediction and the mean value of the 95% prediction confidence interval for each data point, are reported here.

The mean 95% prediction confidence in predicting a lead concentration from the data used as input to the model was obtained using the equation

$$\frac{\sum_{i=1}^{n} C_i}{n} \qquad (2)$$

where n is the number of test data points and $C_i$ is the 95% confidence interval for test data point *i*. The use of this confidence interval when assessing neural network predictions is discussed in Jepson ([12]) where the equation

$$C_i = t_\alpha \times \sqrt{\frac{n-1}{n-2} s_y^2 \left(1-r^2\right)\left[1+\frac{1}{n}+\frac{\left(x-\bar{x}\right)^2}{\left(n-1\right)s_x^2}\right]} \qquad (3)$$

is used to calculate the 95% confidence limit. Here $n$ is the number of pairs of actual and predicted values, $t_a$ is the student's t statistic with $n$-1 degrees of freedom and 95% confidence, $\bar{x}$ and $s_x^2$ are the mean and variance respectively of the actual values, $s_y^2$ is the variance of the predicted values and $r^2$ is the coefficient of determination.

## RESULTS

The full results obtained are shown in table 1. Using the peak height calibration curve obtained from the set A training set, concentration estimates for each of the test set vectors were obtained. A plot of the actual concentration vs the predicted concentration for the test set is shown in figure 3 together with a linear regression and the 95% limits in prediction. The mean absolute error of the predictions shown here is 14.2 ppb and the mean value of the 95% confidence limits over the range shown for this plot is ± 30 ppb.

TABLE 1 - Summary of results

| | Absolute Error (ppb) | | | Average 95% Confidence (± ppb) | | |
|---|---|---|---|---|---|---|
| Net Type | Non saturating (set A) | Only Saturating (Set B) | All peaks (Set C) | Non saturating (set A) | Only Saturating (Set B) | All peaks (Set C) |
| 138 :0:1 high order net | 42.0 | | | 68.8 | | |
| 22 input, 9 factor PLSR | 30.3 | | | 64.7 | | |
| 19:0:1 MLP (BP) | 27.6 | | | 60.4 | | |
| 22:0:1 MLP (BP) | 25.8 | | | 58.9 | | |
| Peak Height Calibration | 14.2 | | | 30.0 | | |
| 19 input GRNN | 12.5 | 9.3 | 11.3 | 31.4 | 20.4 | 24.9 |
| 22 input GRNN | 12.5 | 8.7 | 12.0 | 27.5 | 21.3 | 26.4 |
| 19:30:1 MLP (BP) | 8.4 | 8.7 | 8.3 | 16.7 | 18.3 | 18.4 |
| 22:30:1 MLP (BP) | 6.1 | 9.6 | 10.3 | 16.1 | 18.3 | 23.9 |

The best neural network performance on this data set was obtained by a 22 input, three layer MLP having 30 hidden units and trained using a batch update backpropagation algorithm (figure 4). On the same unseen, non saturating curves as those used above this achieved a mean absolute error of 6.1 ppb, and a mean 95% confidence of ±16.1 ppb.

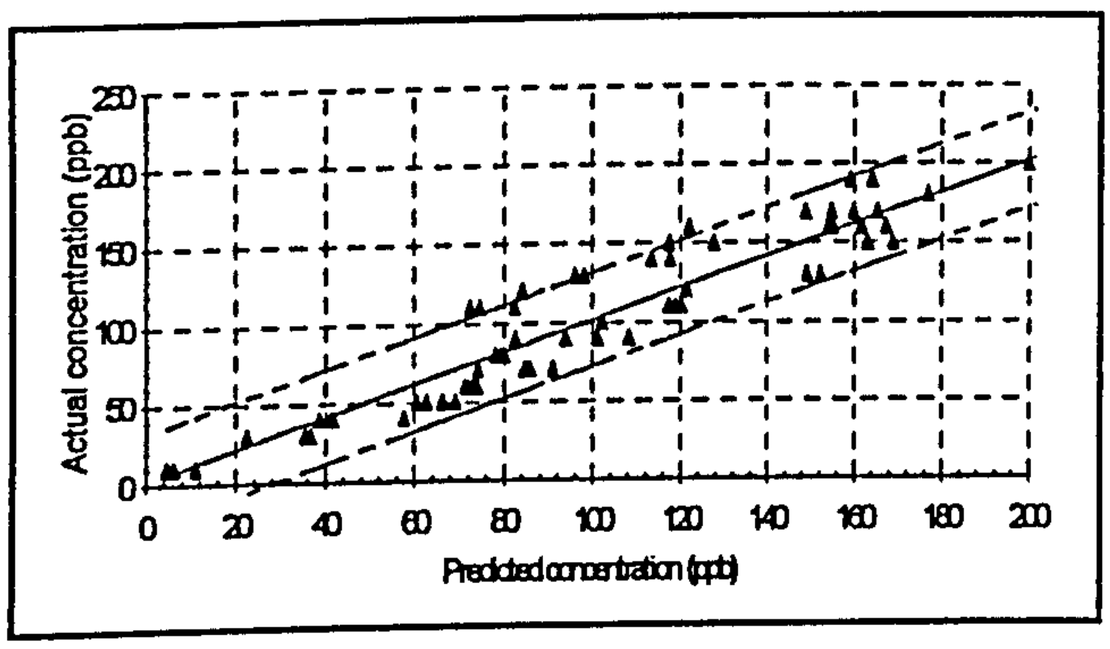



Figure 3: Test set concentration predictions for peak height calibration (set A)

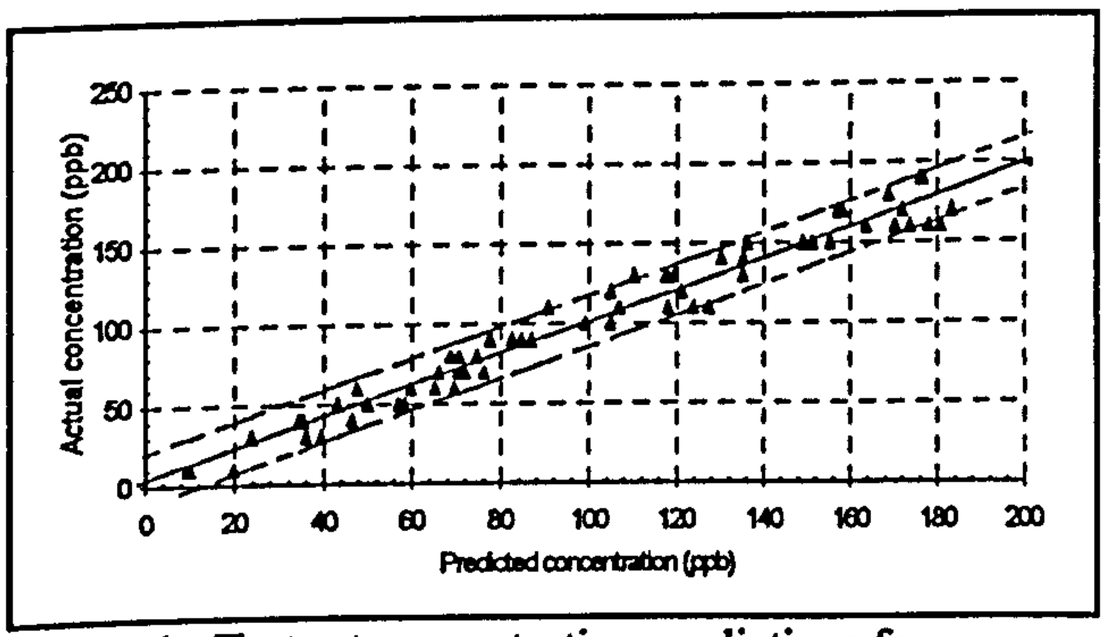



Figure 4: Test set concentration predictions for 22:30:1 neural network calibration (set A)

The voltammograms exhibiting saturation cannot be used in the peak height calibration since it is not possible to estimate the height of the peak. However a 19:30:1 backpropagation MLP was successfully trained using just the saturating curves and when tested on the saturating test set achieved a mean absolute error of 8.7 ppb with a mean 95% confidence of ±18.3 ppb.

The second data set that could not be used with a peak height calibration is set C which contained a mixture of saturating and unsaturating curves, as might be obtained in practice. With this set the mean absolute error and mean 95% confidence at 8.3 ppb and ±18.4 ppb respectively, are still better, than the values obtained for the peak height method.

**Calibration by right hand peak slope**

As part of the investigation into these results attempts were made to parameterise the original curve's data points. One such attempt was to use ten data points taken from the right hand slope of the peak as a vector, the operational parameters were appended and the vector normalised as described previously. When this right hand slope was used with set C to train and test a 12:20:1 backpropagation MLP it was found that a mean absolute error of 19.5 ppb and a mean 95% confidence of ±38.5 ppb could be obtained. The calibration graph for the test set is shown in figure 5.

These particular figures are similar to those obtained by a simple peak height method and do not represent the significant advantage of the neural network technique outlines above. However they do give an insight into alternative sources for information related to the concentration.

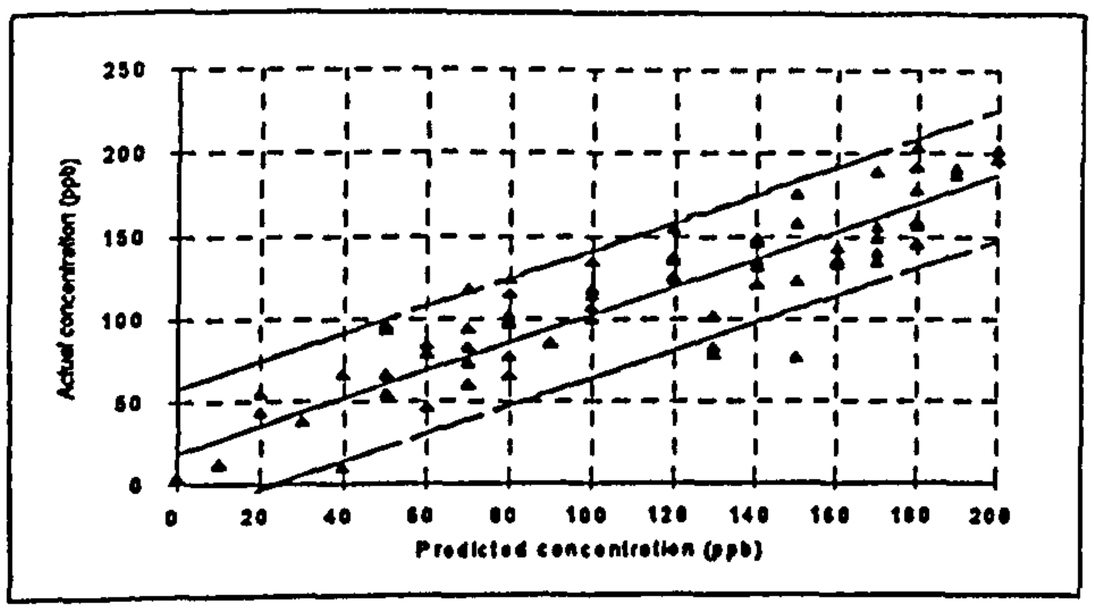



Figure 4: Test set concentration predictions for 12:20:1 neural network using RHS slope of the peak as input (set C)

## CROSS VALIDATION EXERCISE

The main conclusion of this work rest upon the hypothesis that neural network processing, in particular by backpropagation MLP, provides more precise calibrations than a calibration curve produced from peak heights alone. Accordingly a one tailed student's t test was devised to ascertain the significance of the apparently lower 95% prediction interval obtained by MLP over peak height calibration and also the perceived lower mean absolute error in prediction. The procedure was as follows.

The file containing data from unsaturated peaks was first randomised by making a single pass through the file. As each row of the data matrix was reached during the pass it was swapped with a row chosen at random from the entire file. After randomising the data, four cross validation training and testing data sets in the ratio 75% training 25% testing were extracted.

This was done by taking the vectors 1 to 64 to be the test data for set 0 with the remaining vectors going into the training set for set 0. Then the next 64 vectors, numbers 65 to 129, were extracted to go into the test set for set 1 with vectors 1 to 64 and 130 onwards going into the training set for set 1. This process was repeated until four sets were obtained and each vector had been in both the test set and the training set once.

Each of these four sets was then processed to obtain a prediction of test set lead concentration using (a) a least squares regression on the peak heights found in the original voltammograms and (b) a 22 input MLP trained using the same methods and training parameters as previously described.

Each of these eight sets of 64 lead predictions was then processed to obtain the 95% prediction limits in predicting the lead concentration, the average 95% prediction limit for each set and each method and the mean absolute error for each set and each method. This gave sets of paired data related to the two methods. The values obtained, along with other indicative values are given below in tables 3 and 4.

| Set | Avg 96% Conf (ppb) | min abs error (ppb) | max abs error (ppb) | mean abs error (ppb) |
|---|---|---|---|---|
| 0 | 31.46 | 0.89 | 37.40 | 15.31 |
| 1 | 31.50 | 0.17 | 43.90 | 13.88 |
| 2 | 29.36 | 0.20 | 53.55 | 12.87 |
| 3 | 31.81 | 0.31 | 41.08 | 15.17 |

TABLE 3 - Results of cross validation using peak height calibration

| Set | Avg 96% Conf (ppb) | min abs error (ppb) | max abs error (ppb) | mean abs error (ppb) |
|---|---|---|---|---|
| 0 | 14.67 | 0.02 | 28.22 | 6.11 |
| 1 | 22.23 | 0.19 | 32.67 | 9.99 |
| 2 | 18.07 | 0.04 | 31.28 | 7.76 |
| 3 | 19.68 | 0.46 | 33.85 | 8.91 |

TABLE 4 - Results of cross validation using a neural network

### Significance of cross validation results

Microsoft Excel© was used to perform a one tailed paired student's t test on the mean absolute error of the network and the peak heights calibration ($x_1$ & $x_2$) and the mean 95% confidence in the estimates ($y_1$ & $y_2$). The null hypothesis in both cases was that the means of the two sets of errors (or the two sets of confidence values) from each calibration method were equal. The alternative hypothesis was that they were not equal. This gave the following results where the critical region with 3 degrees of freedom and $\alpha = 0.05$ is was taken to be $t > 2.353$.

For the mean absolute error, $t(x_1, x_2)$, calculated by Excel, = 5.387, with a one-tail probability of P(T<=t) of 0.006.

For the average 95% confidence limits, $t(y_1, y_2)$, calculated by Excel, = 7.776 with a one-tail probability of P(T<=t) of 0.002.

The decision in both cases therefore to reject the null hypothesis that the means are equal. Since the t values obtained fall in the right tail of the critical region it can be concluded that processing by neural network does indeed produce both a lower mean absolute error in prediction than does using calibration curves obtained

from peak heights. Similarly it may be concludes that the 95% confidence limits in predicting a concentration are smaller than those obtained using calibration curves obtained from peak heights.

## CONCLUSION

The results of this work have shown that even relatively uncomplicated neural networks are able to produce the required quantitative predictions from ASV instrument responses. The data processed may be obtained from scans performed at a range of pulse heights and deposit times without re calibrating the instrument.

The neural network model appears to be superior to the peak height method in both accuracy and precision, even when using relatively crude pre-processing of the data. The operator is not limited to using only one set of operational parameters as the parameters in force at the time of the scan are taken into account by the neural network model.

Furthermore it is shown that good predictions can be made by neural networks from scans that would have been of no use if processed using traditional methods due to saturation. Calibrations may be made from saturating and non saturating scans alike and valid quantitative predictions of concentration obtain from test data that is similarly variable in quality.

## REFERENCES


1. Manwaring H. S. 1994, "A comparison of backpropagation and general regression neural networks in quantative trace metal analysis using anodic stripping voltammetry and cybernetic instrumentation.", Neural Compt & Applic, 2, 168-178.

2. Wang J. 1985, "Stripping Voltammetry", VCH Pub. New Mexico.

3. Kopanica M. and Opekar F. 1987, "Electrochemical stripping analysis." *In*: Kalvoda R. ed. "Electrochemical methods in chemistry", Plenum Press, London

4. Brainina Kh. Z. 1974, "Stripping Voltammetry in Chemical Analysis", John Wiley & Sons, London.

5. Long J. R. et al. 1990, "Spectroscopic calibration and quantitation using artificial neural networks", Anal Chem, 62, 1791-1797.

6. Wythoff B. J. et al. 1990, "A. Spectral peak verification and recognition using a multilayered neural network. Anal Chem , 62, 2702-2709.

7. McAvoy T. J. et. at. 1989, "Interpreting biosenor data via backpropagation", Proc. Int. Joint Conf. on Neural Networks, 1, 227-233.

8. Blank T. B., Brown S. D. 1993, "Non linear multivariate mapping of chemical data using feed-forward neural networks", Anal Chem , 65, 3081-3089.

9. Blank T. B., Brown S. D. 1993, "Data processing using neural networks", Anal. Chim. Acta. , 277, 273-287.

10. Wold S. et al. 1989, "Non Linear PLS Modelling", Chemom. Intell. Lab. Sys. , 7, 53-65.

11. Wold S. 1992, "Non-linear partial least squares modelling II. Spline inner relation.", Chemom. Intell. Lab. Sys. , 14, 71-84.

12. Jepson B. et al. 1993, "Post neural network procedure to determine expected prediction values and their confidence limits", Neural Comput & applic., 1, 224-228.


## ACKNOWLEDGEMENTS


The author wishes to thank the following institutions:

West Herts College for assisting with the funding of this research and for providing the time and facilities to carry it out.

University of Hertfordshire for providing laboratory assistance and support through their PhD programme.

# Appendix 2. Ancillary work carried out as part of this investigation

## A2.1. Synthetic Curve generation and fitting according to a Pearson VII Line

As part of an attempt to reduce the number of inputs to a neural network it was investigated whether an equation could be fitted to voltammetric curves and the coefficients of that equation used as inputs rather than representations of the actual data points.

De Weijer[194], in Curve fitting using natural computation, describes a method of using a Pearson VII line to fit peak shaped chromatographic instrument responses. During the work on this current thesis the author investigated the feasibility of doing the same for voltammograms of a sample of water taken from a local river. The river water was scanned by the apparatus previously described using differential pulse and the instrument response read into the Excel spreadsheet.

The equation given by De Weijer is;

$$A(\nu) = \sum_{i=1}^{n} \frac{A_{i,0}}{\left[1 + 4Z_i^2\left(2^{\frac{1}{m_i}} - 1\right)\right]^{m_i}} \qquad \{27\}$$

where;

$A_{i,0}$ = height in the centre of peak $i$

$$Z = \frac{\left(\nu - \nu_{i,0}\right)}{H_i}$$

$\nu_{i,0}$ = peak position of peak $i$

$H_i$ = the half width of peak $i$

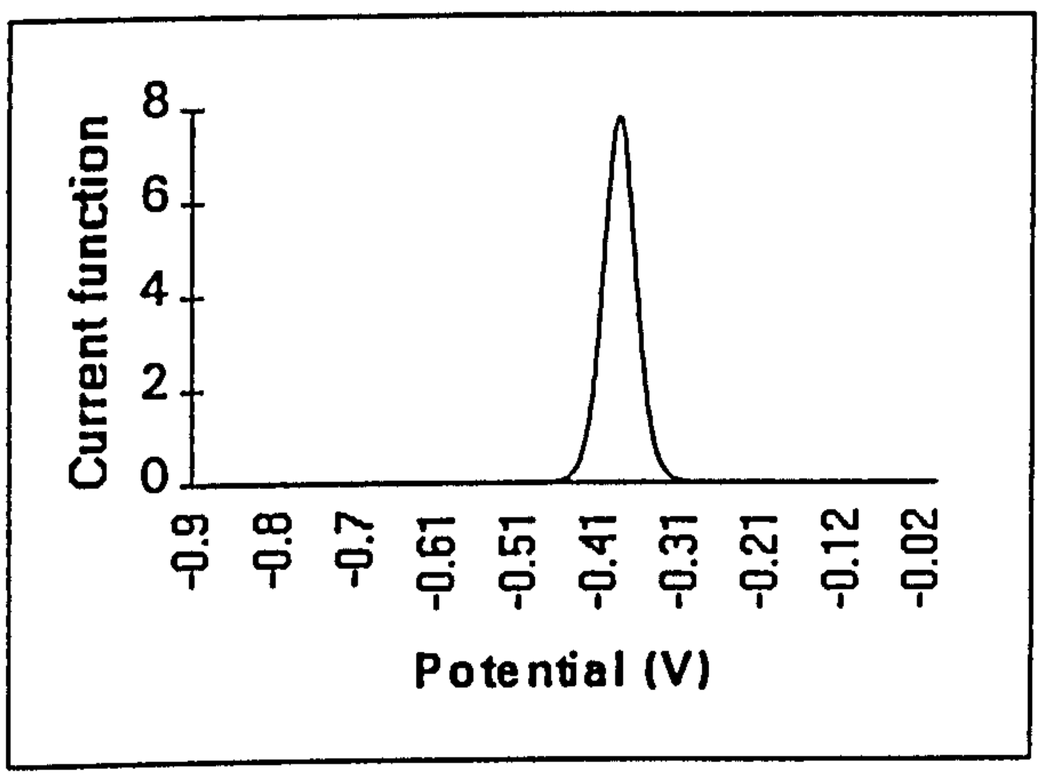
$m_i$ = the trailing factor of peak $i$

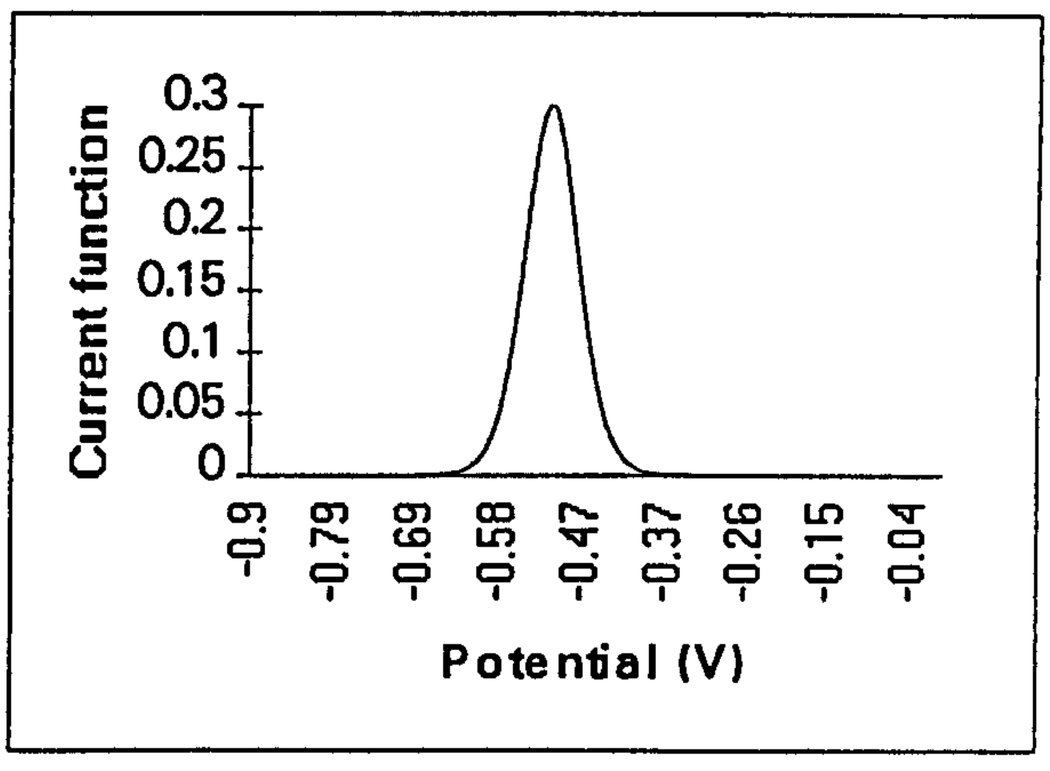
$n$ = the number of peaks

Using Excel, two such curves were generated. These were convoluted together and a small linear offset added to each point. A graph was then plotted containing

both the real voltammetric data and the combined synthetic curve. By manually adjusting the various parameters of the curves a reasonable fit was obtained on the real curve.

Below are shown the two individual peaks, generated by the above equation before they were convoluted.



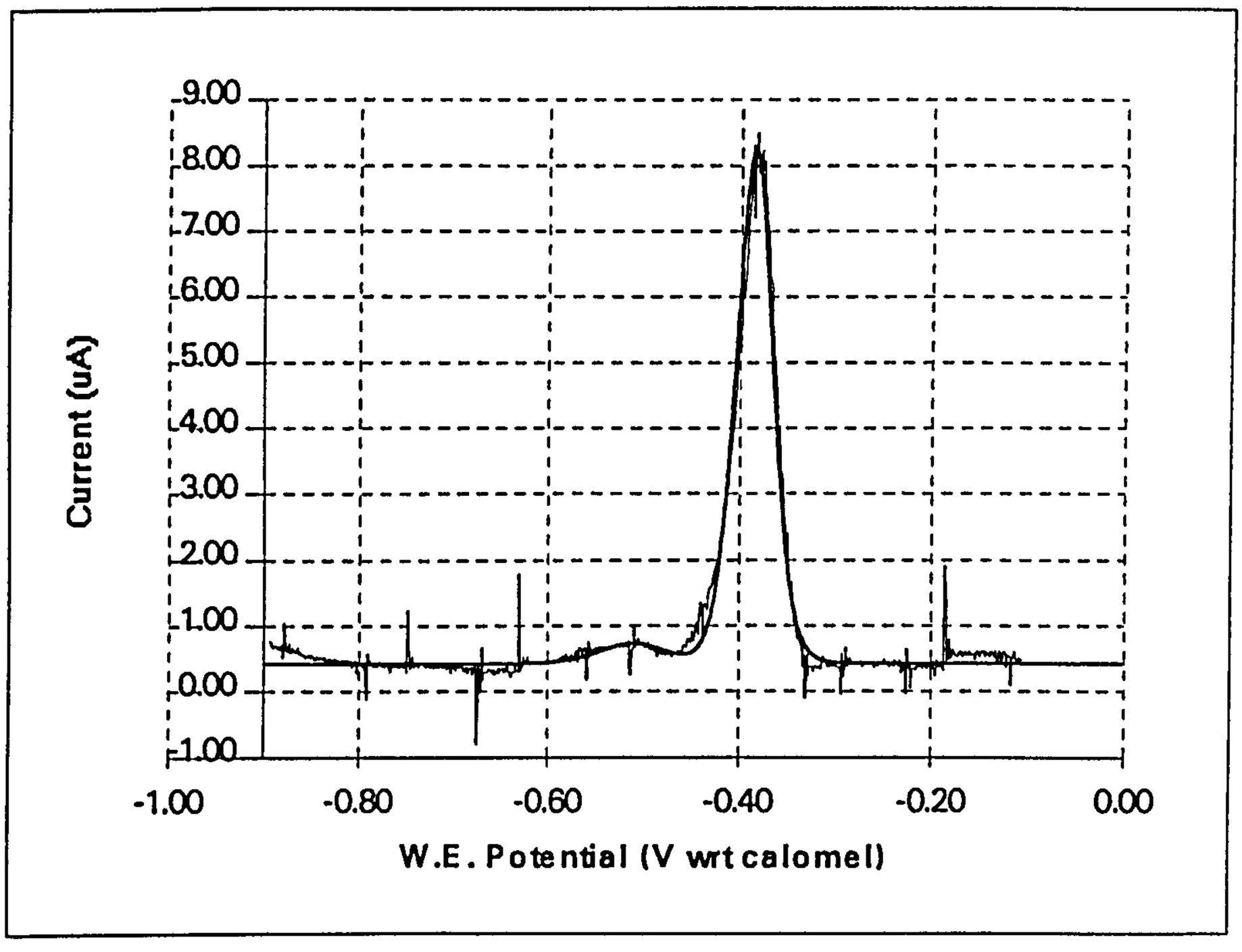

Curve 1 (lead) Curve 2 (thallium ?)

The combined curve is plotted on the following graph along with the real curve. The real curve may be readily identified as the one have noise spikes.

The final parameters used were

| peak 1 | | | peak 2 | |
|---|---|---|---|---|
| peak position | -0.385 | | peak position | -0.51 |
| peak height | 7.8 | | peak height | 0.3 |
| half width | 0.044444 | | half width | 0.08 |
| trailing edge curvature | 7 | | trailing edge curvature | 7 |

The constant base offset added to each data point was 0.43.




Combined curve and real curve plotted on the same axis

It seems that the half width value used in the equation is actually the reciprocal of the actual peak width as when the reciprocal of this value is taken the real half

widths of peak 1 and 2 are 45 mV and 25 mV respectively, close to the values seen in the real curve.

Judging by the peak positions, the metals in the river water seem to be lead and thallium. Indeed, a second test of the water using ASS indicated that the water contained lead at a concentration of 800 ppb.

This exercise indicated that it may be possible to use this equation to fit a Pearson VII curve or curves to a voltammogram and raises the additional possibility of using the parameters of these fitted curves obtained as inputs to a neural network trained to predict the concentration.

## A2.2. Investigation into the optimisation of instrument parameters using a genetic algorithm

Some time was spent during this work investigating techniques that could be utilised to optimise the operational parameters of the instrument so that the best voltammogram was produced. The intention was that after this had been done the neural network would be able more precisely to predict the concentration of the test substance.

As a first attempt at performing a similar operation, the author used the same program as that discussed in appendix 2 to generate synthetic mixtures containing two peaks.

The investigation was to optimise the parameters used by the curve generation program so that a form of resolution enhancement of overlapping peaks was performed. To this end, since the curves could be produced anywhere on the axis, an optimum combined curve was defined. This was one that had the two peaks as tall as possible, with the positions of the peaks as similar as possible and as far to the right as possible. However the valley between them was to be as deep as possible.

In this simulation the parameters being controlled were those generating the synthetic curves. In the real example it was intended that they would be the operational parameters of the analytical instrument which, indirectly, generate the real curve.

A genetic algorithm was written in Pascal to perform this optimisation. The general operation of the genetic algorithm followed traditional lines and was modelled after the work of Davis[122] and Goldberg[123].

The chromosome was coded as a string of 18 bits encoding the six curve parameters (3 for each curve, i.e. position, height and width). Both Gray and binary encoding were tried for this encoding. Initially a population of 50 such chromosomes were created at random. Each chromosome was passed in turn to the program that used them to generate the two $sech^2$ curves and convolute the result.

Following this the 'resolution parameter' was found for each convoluted curve by;

$$resolution = \frac{a(height_1 \times height_2) + b(depth_{valley} \times position_1 \times position_2)}{c}$$

where $a$, $b$, and $c$ are constants

Following the principles of genetic algorithms, pairs of mates were selected from this population with a probability proportional to their resolution. Crossover between these pairs was performed at a random position with a probability of 0.7 followed by the mutation of each bit with a probability of 0.001.

The best 50 of these 100 curves were then saved as the 'new' population and the process repeated. Optionally, the best curve was plotted by passing it to the graph plotting module described earlier in Chapter 4.

This process ended when no further improvement in the resolution parameter could be made.

The results of this excursion into another area of machine learning showed that it was possible, even using the rather crude metrics described here, to optimise the parameters for curve generation.

However, a drawback using this approach in a real voltammetric experiment would be the time that it would take to optimise. Generating 30 or so generations, each one containing two sets of 50 curve parameters only took a few minutes using a 40 MHz computer. Generating the same curves where the resolution parameters are found by performing real voltammetric scans, with each one taking around 90s would take over three days!

For this reason, although the work was interesting and thought provoking, it was abandoned at this time. Elements of this work could, however, be employed. If a neural network was trained in a control mode, mapping existing operational parameters and existing voltammograms to an output of better operational parameters, the task of optimisation may only take a few cycles and hence be complete in a short time.

# Appendix 3. Initial investigations using Gray & Binary output encoding

In this initial study the aim was to investigate qualitatively if neural networks could perform at all well using voltammetric data. Hence the values and materials used during the preparation of the samples were not critical and were chosen from those used during the initial development of the instrument.

## A3.1. Methods used

The analytes for the initial study were prepared by taking 100 units of a 0.05M Potassium Nitrate buffer as a blank and performing 20 runs by differential pulse voltammetry on this alone. Then 1 part of 1 ppm Lead was added to give a nominal concentration of 10 ppb Lead. Again 20 runs were performed. Another 1 part of 1 ppm Lead was added and 20 runs performed. This was repeated until a nominal concentration of 50 ppb had been analysed giving 20 sets of data for each of 6 concentrations of lead from 0 to 50 ppb in nominal steps of 10 ppb.

### A3.1.1. Working electrode preparation method

The working electrode used was pre plated with mercury . This was achieved by first polishing the working electrode for 30s using a slurry of alumina in deionised water after which it was washed in de ionised water, dried on a tissue and rinsed again in deionised water. The electrode was then rotated at 300 rpm in a mercury solution ($Hg(NO_3)_2$ in 0.1M nitric acid) at a concentration of 2000 ppm for

60 s. During this time the electrode was maintained at a potential of -0.4v with reference to a home made silver - silver chloride electrode.

An anodic linear scan was then performed from -0.4 V to 0.0 V at a scan rate of 0.33 $Vs^{-1}$ in order to strip off any co-deposited materials. Finally the electrode was rinsed using de ionised water.

### A3.1.2. Differential pulse analysis method

Plating the electrode with the material from the analyte was done by applying a potential of -0.7V vs. Ag/AgCl to the working electrode and maintaining that potential for 20s whilst the electrode was rotated at 300 rpm. After this time rotation was stopped and the potential maintained at -0.7V for a further 20ms. This rest period allowed the analyte to become quiescent so that further electrochemical activity took place in a diffusion controlled environment.

Following this period a differential pulse scan was made from -0.7V to -0.1V using a pulse height of 20 mV, a pulse width of 20 ms and a rest period between pulses of 40 ms. The potential ramp was created by stepping the potential by 5 mV after each pulse. The sensitivity of the instrument was maintained at 250 $nAV^{-1}$, the gain set to x1 and zero offset was applied.

### A3.1.3. Data set extraction and processing

Each analytical run produced 94 digital values corresponding to the sampled differential current for each pulse. These values were recorded directly from the potentiostat's A/D output with no pre-processing (for example to convert them into actual values of current). The first and last 2 points of each data set were discarded to eliminated spurious end effects. The remaining 90 raw data points were taken as the value of the input vector for the network. The values could possibly have been chosen so that the effect of 50 Hz mains interference was partly eliminated. However this was deliberately not done.

The target network output values represented the quantised concentration of lead in steps of 10 parts per billion. These were assigned in one of two ways. Either using three real numbers representing the values 0 to 5 in binary or, again using three numbers, as the values 0 to 5 using the Gray encoding system. The target vector allocated for each concentration is given in the table below. In order to improve network performance and prevent saturation of the sigmoid transfer function, target values of 0 were actually encoded as 0.1 and those of 1 encoded as 0.9. The network was required to learn the mapping between the input vector, i.e. the analytical curve, and the output value given below in table 37.

The training and test data sets were prepared in the following way. First each 90 value data set was normalised to the range 0 to 1 and stored in a file with the corresponding three value target output vector appended.

| Concentration | Binary Encoding | | | | Gray Encoding | | |
|---|---|---|---|---|---|---|---|
| | Output 3 | Output 2 | Output 1 | | Output 3 | Output 2 | Output 1 |
| 0 | .1 | .1 | .1 | | .1 | .1 | .1 |
| 10ppb | .1 | .1 | .9 | | .1 | .1 | .9 |
| 20ppb | .1 | .9 | .1 | | .1 | .9 | .9 |
| 30ppb | .1 | .9 | .9 | | .1 | .9 | .1 |
| 40ppb | .9 | .1 | .1 | | .9 | .9 | .1 |
| 50ppb | .9 | .1 | .9 | | .9 | .9 | .9 |

Table 37. Target encoding system used for binary and Gray

The even numbered patterns from each of the 20 runs at each concentration were then concatenated into one file containing 60 input vectors and their corresponding target output vector. This was used as the training set. The odd numbered patterns were similarly concatenated and used as the test set. This gave 60 patterns in each set, each containing 93 values (90 input + 3 output). These two sets were replicated with one train / test set using binary and one using Gray encoding for the target vector.

The network used was a standard MLP trained using backpropagation and was written in Visual Basic using primitive neural network tools available in a DLL called NeuroWindows (Ward Systems). It had 90 inputs, 3 outputs and a variable number of hidden nodes. The learning rate for the hidden slab was 0.3 and that for the output slab was 0.2. The momentum for all nodes was 0.9.

### A3.1.4. Network training method

The network training and testing was divided into two sessions, one for each output encoding. Each training session consisted of training the network on 300 different combinations of threshold mean squared error (mse) i.e. the lowest error value accepted before the network is considered to be trained, and the number of hidden nodes in the network .

The number of hidden nodes was changed in integral steps from 2 to 15 and for each of these 14 network architectures the network was trained to one of 20 threshold mse ranging from 0.1 to 0.005 on steps of 0.005. The network was completely retrained for each architecture and threshold combination using random starting weights of between +2 and -2 for the hidden slab and between +3 and -3 for the output slab.

After the network had reached its threshold mse for each combination it was tested using both the training set and the previously unseen test set. The number of correct classifications for each combination was recorded. Correct in this context is defined as each output node being $>0.45$ for a target of 0.9 and $<= 0.45$ for a target of 0.1.

This data was then sorted using the percentage correct on the test set as the first key and the percentage correct on the training set as the second key. In this way it was possible to find the combination that produced the best generalisation for each of the two target encoding systems.

## A3.2. Results of Initial studies

Typical curves produced by the instrument using the initial parameters, analyte production and electrode preparation are shown in Appendix 3. Note the large peak caused by a copper impurity at the right hand end of the curve. It is interesting to compare these with the much better quality curves obtained using the in situ plated, Florence electrode used for the later part of this work (see for example Figure 53 on page 165).

The data obtained from training using different combinations of number of hidden units and final mean squared error was first sorted using the percentage correct on the test set as the first key and the percentage correct on the training set as the second key. In this way it was possible to find the combination that produced the best generalisation for each of the two target encoding systems.

The graph below, figure 87, shows the training curve for best Binary encoded combination, a 90-15-3 network training to a threshold of 0.035 mean squared error. The training took about 4 minutes. For comparison, figure 88 shows the training curve for best Gray encoding combination, a 90-4-3 network training to a threshold of 0.02 mean squared error. The training time for this arrangement was about 2 minutes.

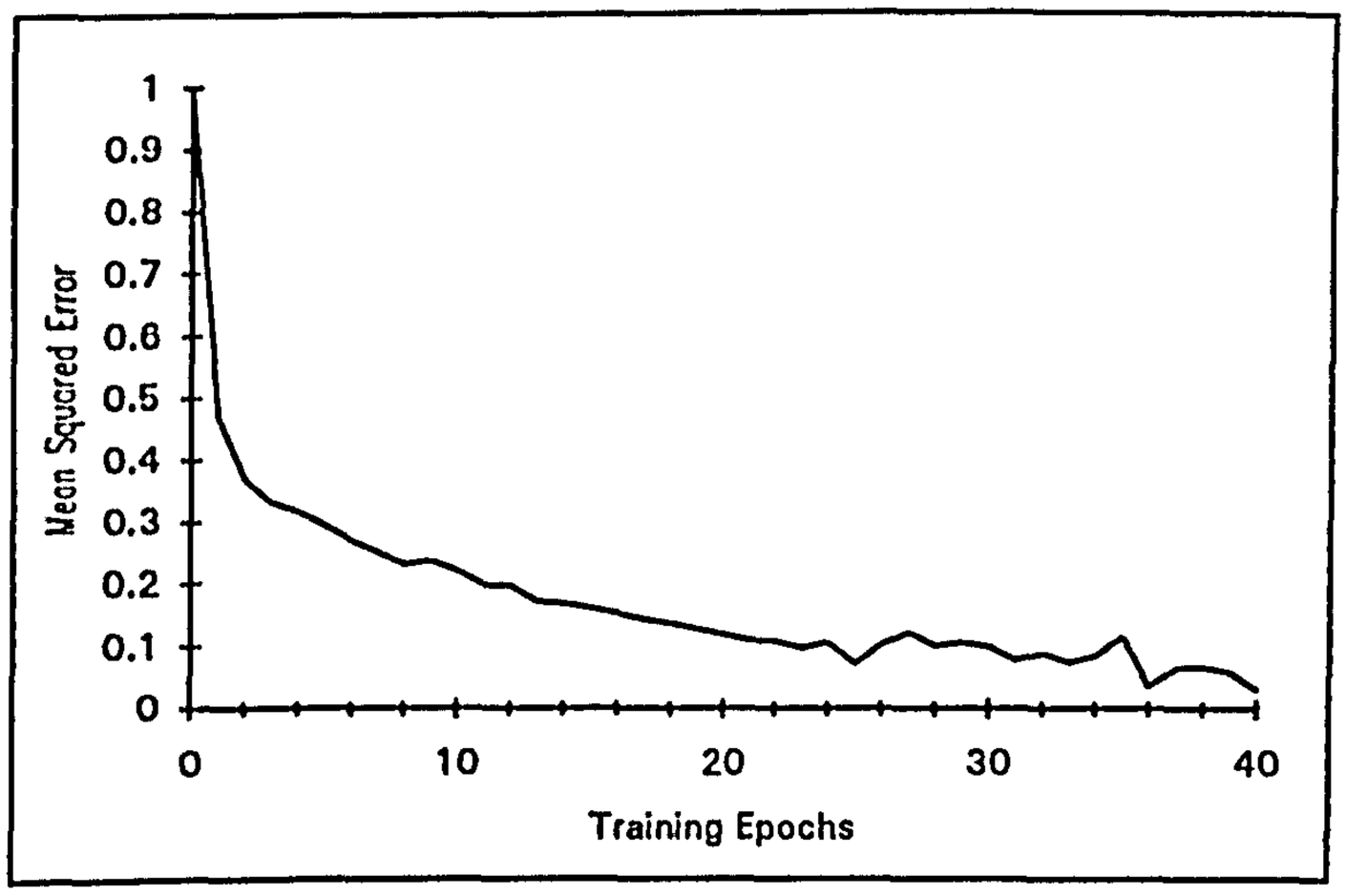



Figure 87. Training curve for binary encoding of targets

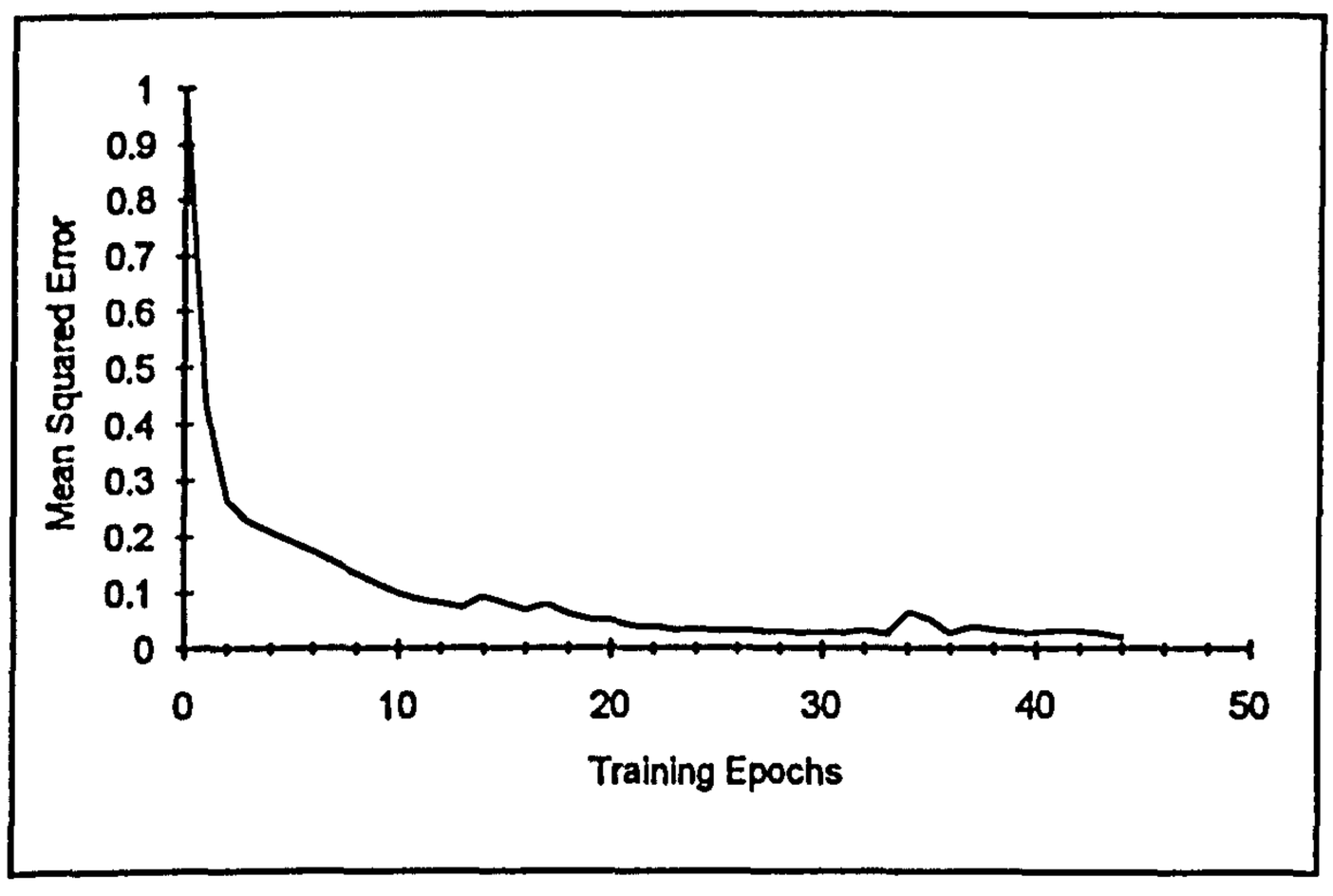



Figure 88. Training curve for Gray target encoding

The resulting hidden node / mse combination for the highest % correct on the test set in each case is tabulated below in table 38. The entire training and test procedure was repeated using different starting weights and almost identical results were obtained.

| | Binary | Gray |
|---|---|---|
| Number of Hidden Units | 15 | 4 |
| MSE Training Set | 0.0218 | 0.0297 |
| MSE Test Set | 0.1248 | 0.0318 |
| Highest % correct on Training Set | 100.0 | 98.3 |
| Highest % correct on Test Set | 85.0 | 95.0 |

Table 38. Initial results obtained using binary and Gray target encoding

The incorrect classifications for the Binary and Gray target encoding are tabulated below in table 39. Here Out3 - Out1 is the network output vector and Tar3 - Tar1 is the target vector. The mean squared error (mse) is the mean of the squares of the errors for each output node.

It is clear that the network is having difficulty distinguishing between the pattern corresponding to 000 and that corresponding to 001 when coded in Binary. These are the patterns for 10ppb and 20ppb Lead respectfully. An alternative explanation could be that the network has difficulty in classifying those outputs with a 1 in the final position. However, this would have caused it to mis-classify every odd output value, which it did not do.

| Testing using Binary Coded outputs on the test data set<br>Network size 90-15-3, trained to 0.035 mse | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Pat No. | Out 3 | Out 2 | Out 1 | Tar 3 | Tar 2 | Tar 1 | mse | correct |
| 37 | .0954 | .0484 | .4611 | .1 | .1 | .1 | .133 | N |
| 38 | .0799 | .7316 | .1205 | .1 | .1 | .9 | 1.007 | N |
| 44 | .0808 | .6237 | .1679 | .1 | .1 | .9 | .811 | N |
| 50 | .0787 | .7936 | .0668 | .1 | .1 | .9 | 1.176 | N |
| 51 | .0697 | .9337 | .5143 | .1 | .9 | .1 | .174 | N |
| 56 | .0889 | .6375 | .195 | .1 | .1 | .9 | .786 | N |
| 58 | .3455 | .6746 | .4294 | .1 | .9 | .9 | .333 | N |

Table 39. Incorrectly classified targets using binary encoding

| Testing using Gray Coded outputs on the training data set<br>Network size 90-4-3 trained to 0.02 mse | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Pat No. | Out 3 | Out 2 | Out 1 | Tar 3 | Tar 2 | Tar 1 | mse | correct |
| 58 | .3455 | .6746 | .4294 | .1 | .9 | .9 | .333 | N |
| 58 | .5527 | .9014 | .0329 | .1 | .9 | .1 | .209 | N |
| 59 | .9433 | .8751 | .6244 | .9 | .9 | .1 | .278 | N |

Table 40. Incorrect classifications using Gray encoding

Using Gray output encoding, see table 40, there does not appear to be an obvious pattern to the type of target that the network found difficulty in predicting. In addition there are fewer erroneously classified patterns than when using binary encoding.

## A3.3. Summary of initial studies

The standard Back Propagation neural network can be an effective tool in the classification of voltammetric curves in the determination of lead concentration. It appears able to generalise well in the presence of noise. Using an encoding scheme that changes in only one dimension for each valid code i.e. a Gray code produces a network that generalises better than one using pure binary encoding.

## A3.4. Stripping Curves obtained

There follows typical voltammograms obtained when using a pre-plated electrode during the study of Gray and Binary neural network output encoding

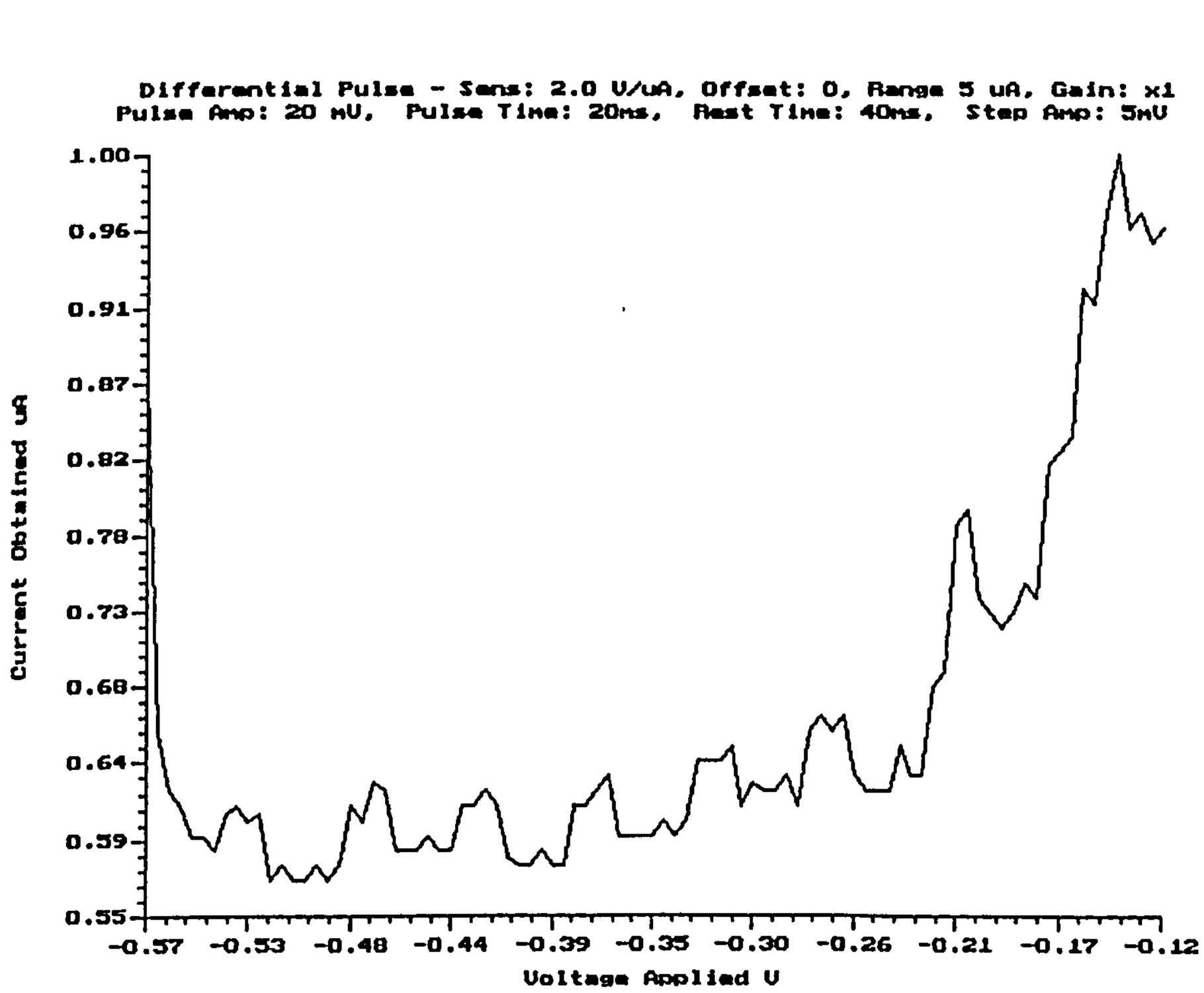



Figure 89. Typical curve for blank NaCl solution

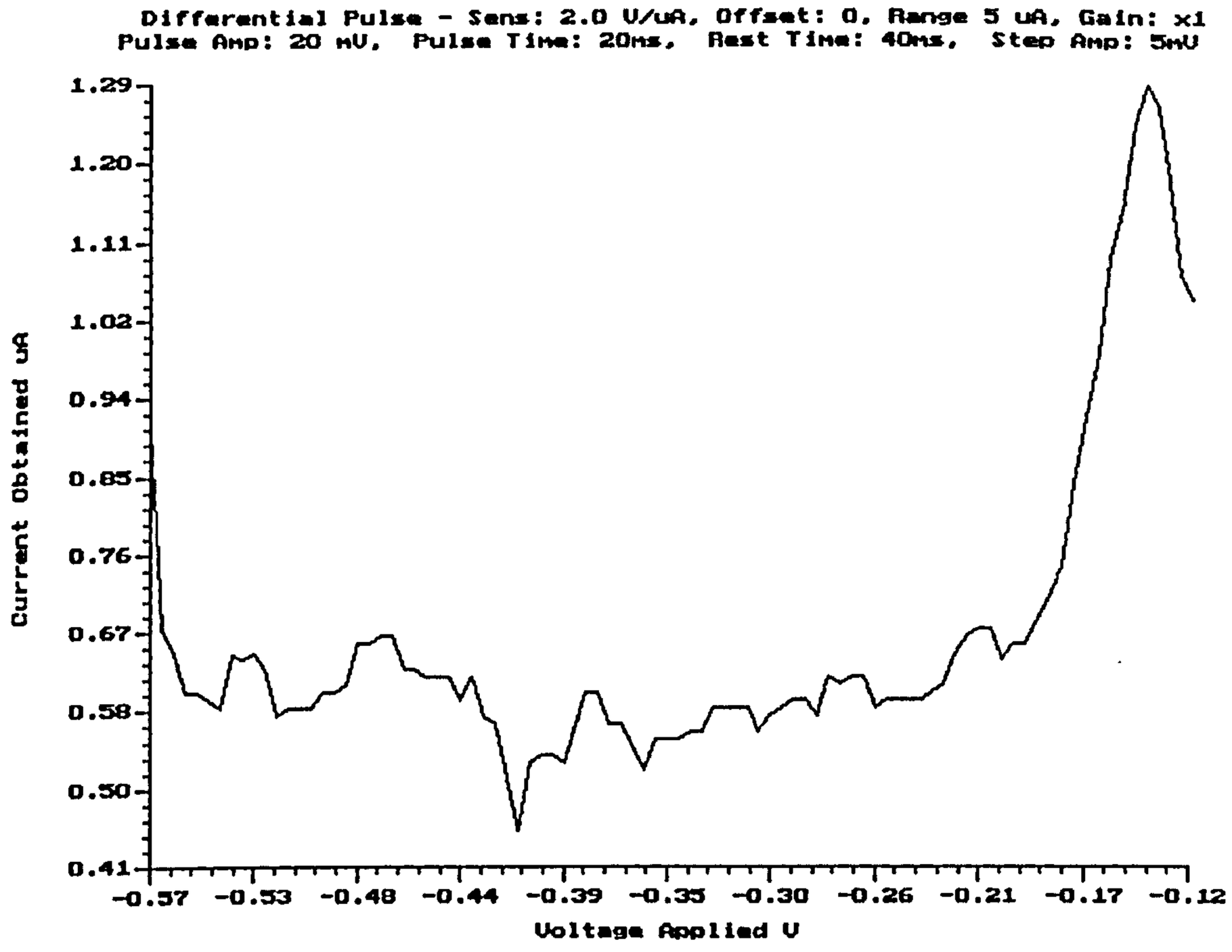



Figure 90. Typical curve for 10 ppb lead

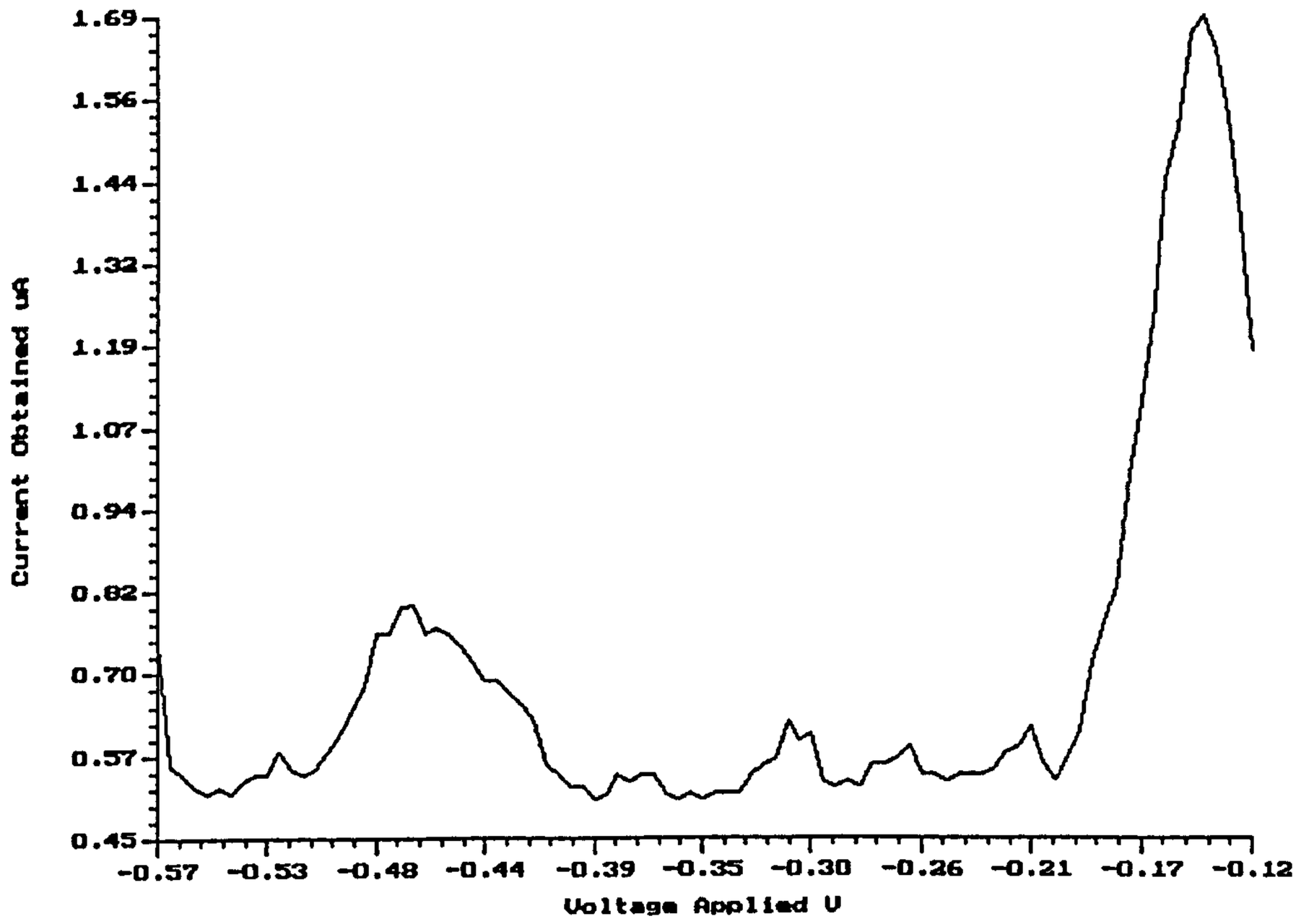



Figure 91 Typical curve for 20 ppb Pb

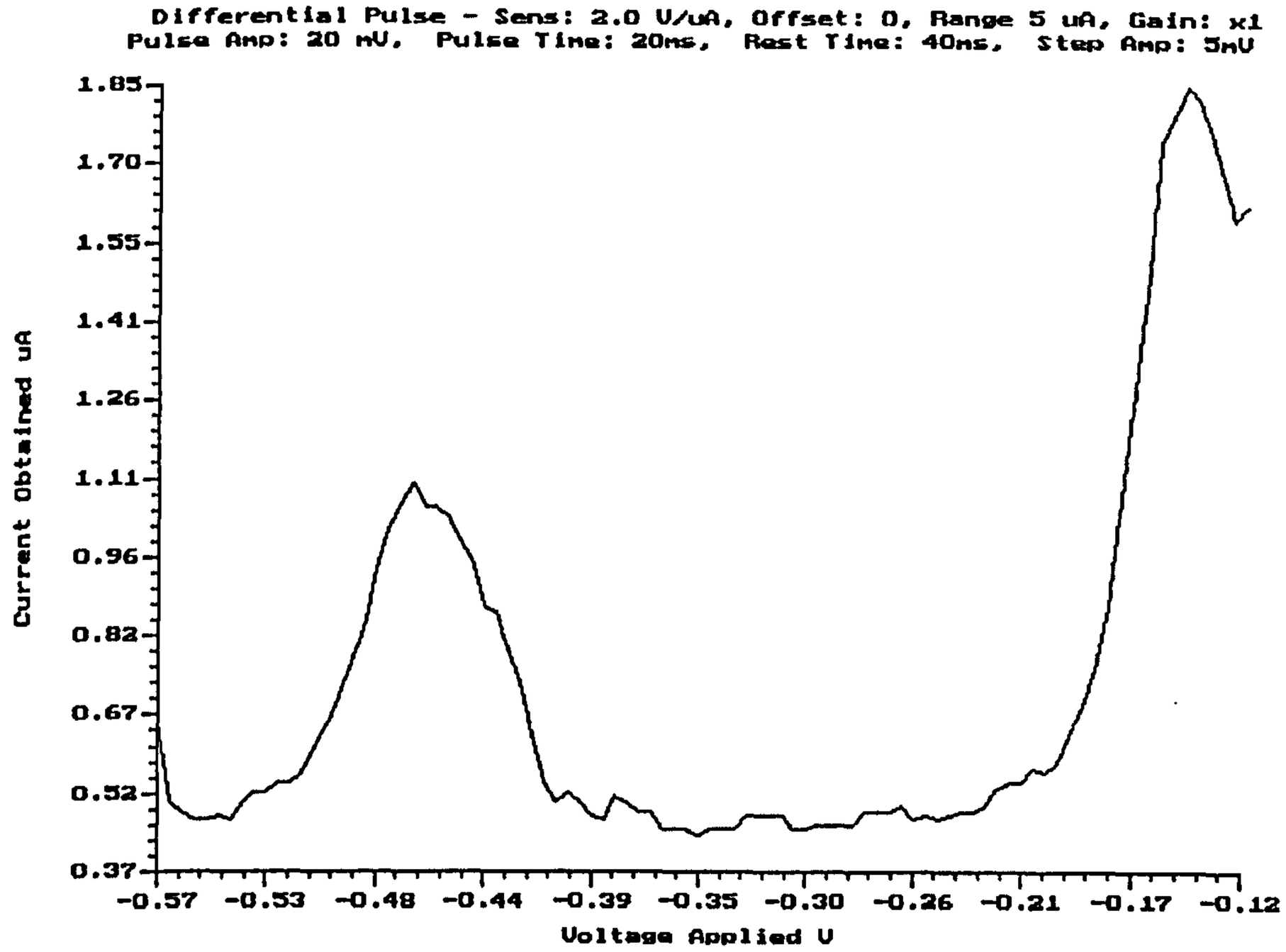



Figure 92 Typical curve for 30 ppb lead

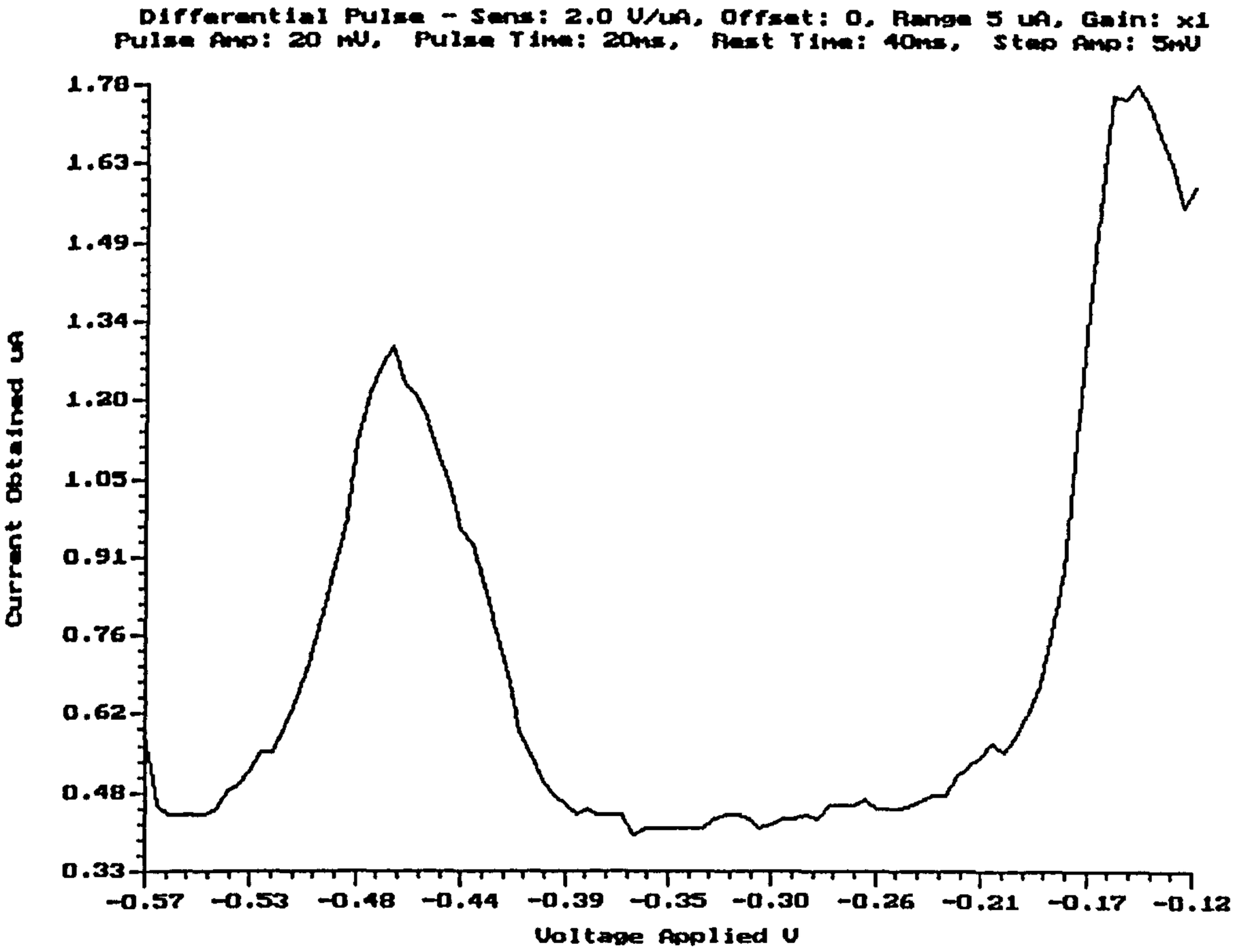



Figure 93. Typical curve for 40 ppb lead

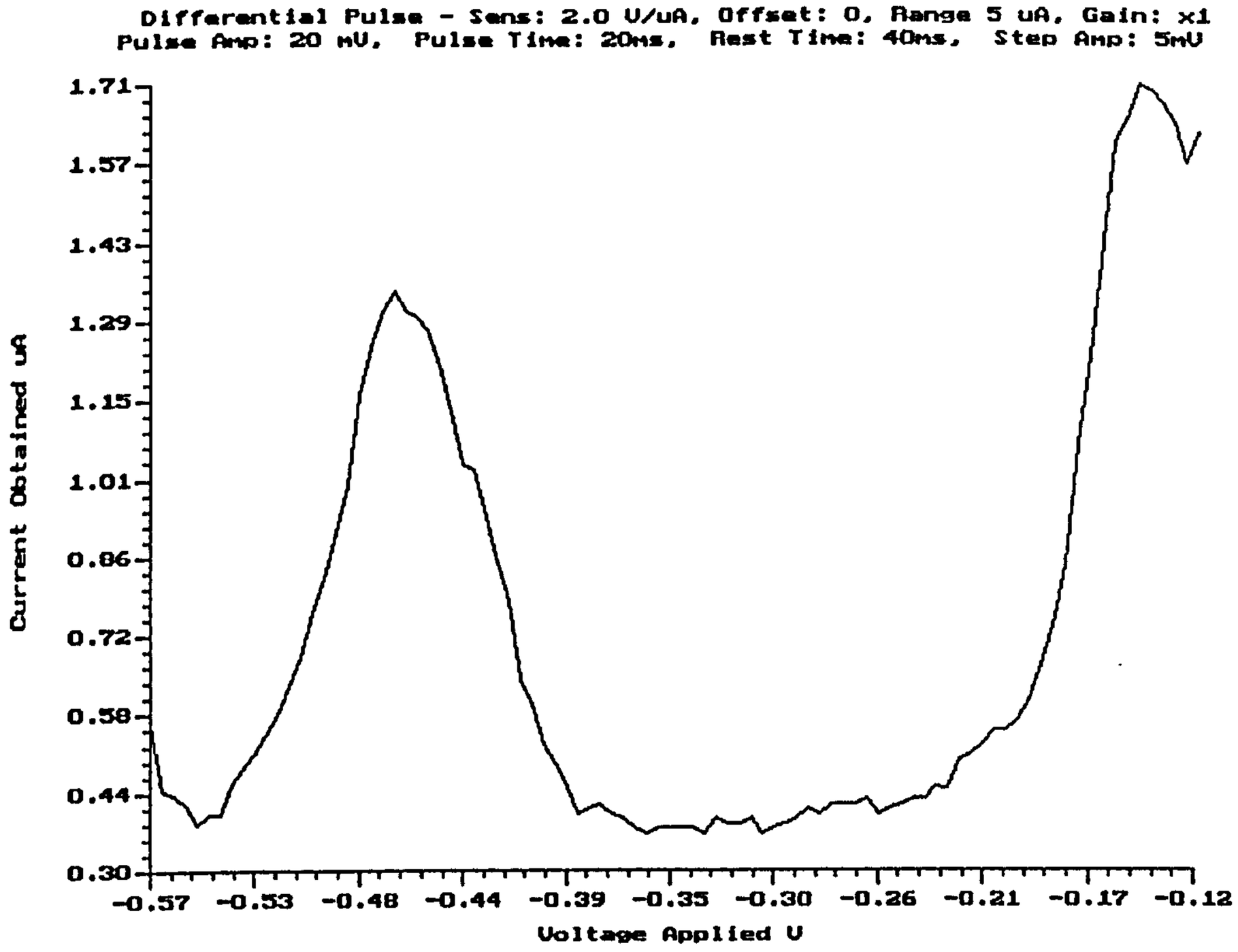



Figure 94. Typical curve for 50 ppb lead

## A3.5. Tabulated results

Testing using Binary Coded outputs on the training data set

Network size 90-15-3, trained to 0.035 mse

| Pat No. | Out 3 | Out 2 | Out 1 | Tar 3 | Tar 2 | Tar 1 | mse | correct |
|---|---|---|---|---|---|---|---|---|
| 1 | .1084 | .0475 | .0615 | .1 | .1 | .1 | .004 | Y |
| 2 | .0907 | .0677 | .7024 | .1 | .1 | .9 | .04 | Y |
| 3 | .0901 | .8397 | .0284 | .1 | .9 | .1 | .009 | Y |
| 4 | .0718 | .9342 | .6419 | .1 | .9 | .9 | .069 | Y |
| 5 | .5885 | .4463 | .1721 | .9 | .1 | .1 | .222 | Y |
| 6 | .9121 | .0761 | .7312 | .9 | .1 | .9 | .029 | Y |
| 7 | .1108 | .0484 | .0624 | .1 | .1 | .1 | .004 | Y |
| 8 | .0815 | .0841 | .726 | .1 | .1 | .9 | .031 | Y |
| 9 | .0814 | .8929 | .0531 | .1 | .9 | .1 | .003 | Y |
| 10 | .0697 | .9342 | .6948 | .1 | .9 | .9 | .044 | Y |
| 11 | .7366 | .2849 | .0627 | .9 | .1 | .1 | .062 | Y |
| 12 | .9225 | .0723 | .5924 | .9 | .1 | .9 | .096 | Y |
| 13 | .123 | .0443 | .0645 | .1 | .1 | .1 | .005 | Y |
| 14 | .1057 | .0783 | .8178 | .1 | .1 | .9 | .007 | Y |
| 15 | .0765 | .8988 | .0358 | .1 | .9 | .1 | .005 | Y |
| 16 | .0742 | .9285 | .6845 | .1 | .9 | .9 | .048 | Y |
| 17 | .6978 | .3119 | .23 | .9 | .1 | .1 | .103 | Y |
| 18 | .9235 | .0665 | .7749 | .9 | .1 | .9 | .017 | Y |
| 19 | .113 | .0557 | .0178 | .1 | .1 | .1 | .009 | Y |
| 20 | .1031 | .0492 | .7593 | .1 | .1 | .9 | .022 | Y |
| 21 | .0769 | .9211 | .0591 | .1 | .9 | .1 | .003 | Y |
| 22 | .0723 | .9313 | .7678 | .1 | .9 | .9 | .019 | Y |
| 23 | .7312 | .2828 | .1783 | .9 | .1 | .1 | .068 | Y |
| 24 | .9357 | .0564 | .8055 | .9 | .1 | .9 | .012 | Y |
| 25 | .1026 | .0722 | .0327 | .1 | .1 | .1 | .005 | Y |
| 26 | .1066 | .0506 | .8262 | .1 | .1 | .9 | .008 | Y |
| 27 | .0736 | .9245 | .0737 | .1 | .9 | .1 | .002 | Y |
| 28 | .0896 | .9099 | .6999 | .1 | .9 | .9 | .040 | Y |
| 29 | .8046 | .2057 | .1561 | .9 | .1 | .1 | .023 | Y |
| 30 | .9431 | .0491 | .8875 | .9 | .1 | .9 | .005 | Y |
| 31 | .0844 | .0917 | .0571 | .1 | .1 | .1 | .002 | Y |
| 32 | .1082 | .0594 | .7797 | .1 | .1 | .9 | .016 | Y |
| 33 | .0718 | .9289 | .1028 | .1 | .9 | .1 | .002 | Y |
| 34 | .0764 | .9232 | .8486 | .1 | .9 | .9 | .004 | Y |
| 35 | .8771 | .1329 | .0746 | .9 | .1 | .1 | .002 | Y |
| 36 | .9422 | .0483 | .9167 | .9 | .1 | .9 | .005 | Y |
| 37 | .0769 | .082 | .0455 | .1 | .1 | .1 | .004 | Y |
| 38 | .1081 | .0631 | .7592 | .1 | .1 | .9 | .021 | Y |
| 39 | .0821 | .9143 | .0479 | .1 | .9 | .1 | .003 | Y |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 40 | .0727 | .923 | .8914 | .1 | .9 | .9 | .001 | Y |
| 41 | .7724 | .2293 | .2744 | .9 | .1 | .1 | .063 | Y |
| 42 | .953 | .0413 | .905 | .9 | .1 | .9 | .006 | Y |
| 43 | .0759 | .0812 | .0739 | .1 | .1 | .1 | .002 | Y |
| 44 | .0959 | .0712 | .7872 | .1 | .1 | .9 | .014 | Y |
| 45 | .0754 | .9262 | .0733 | .1 | .9 | .1 | .002 | Y |
| 46 | .0763 | .9214 | .8829 | .1 | .9 | .9 | .001 | Y |
| 47 | .8554 | .1485 | .1831 | .9 | .1 | .1 | .011 | Y |
| 48 | .9438 | .0479 | .9076 | .9 | .1 | .9 | .005 | Y |
| 49 | .0823 | .064 | .0642 | .1 | .1 | .1 | .003 | Y |
| 50 | .0882 | .1519 | .7966 | .1 | .1 | .9 | .014 | Y |
| 51 | .0718 | .9317 | .1438 | .1 | .9 | .1 | .004 | Y |
| 52 | .1131 | .8779 | .752 | .1 | .9 | .9 | .023 | Y |
| 53 | .8736 | .1311 | .1746 | .9 | .1 | .1 | .007 | Y |
| 54 | .9442 | .047 | .9246 | .9 | .1 | .9 | .005 | Y |
| 55 | .0835 | .0607 | .0978 | .1 | .1 | .1 | .002 | Y |
| 56 | .098 | .1217 | .772 | .1 | .1 | .9 | .017 | Y |
| 57 | .0759 | .9271 | .0852 | .1 | .9 | .1 | .002 | Y |
| 58 | .0754 | .9253 | .8023 | .1 | .9 | .9 | .011 | Y |
| 59 | .8751 | .1317 | .1578 | .9 | .1 | .1 | .005 | Y |
| 60 | .9431 | .0492 | .8565 | .9 | .1 | .9 | .006 | Y |

mean error squared = 2.128272E-02

percent correct = 100.0%

Testing using Binary Coded outputs on the test data set

Network size 90-15-3 trained to 0.035 mse

| Pat No. | Out 3 | Out 2 | Out 1 | Tar 3 | Tar 2 | Tar 1 | mse | correct |
|---|---|---|---|---|---|---|---|---|
| 1 | .0866 | .0523 | .1097 | .1 | .1 | .1 | .003 | Y |
| 2 | .0929 | .1253 | .8039 | .1 | .1 | .9 | .010 | Y |
| 3 | .0729 | .9325 | .1405 | .1 | .9 | .1 | .003 | Y |
| 4 | .0755 | .921 | .8996 | .1 | .9 | .9 | .001 | Y |
| 5 | .8588 | .145 | .2064 | .9 | .1 | .1 | .015 | Y |
| 6 | .9482 | .0448 | .9164 | .9 | .1 | .9 | .006 | Y |
| 7 | .0738 | .1021 | .1168 | .1 | .1 | .1 | .001 | Y |
| 8 | .1048 | .1283 | .6995 | .1 | .1 | .9 | .041 | Y |
| 9 | .0766 | .9263 | .1341 | .1 | .9 | .1 | .002 | Y |
| 10 | .0983 | .8965 | .895 | .1 | .9 | .9 | .000 | Y |
| 11 | .8582 | .1378 | .3368 | .9 | .1 | .1 | .059 | Y |
| 12 | .9456 | .0441 | .953 | .9 | .1 | .9 | .008 | Y |
| 13 | .0901 | .0564 | .3929 | .1 | .1 | .1 | .088 | Y |
| 14 | .0819 | .3716 | .5466 | .1 | .1 | .9 | .199 | Y |
| 15 | .074 | .9332 | .1397 | .1 | .9 | .1 | .003 | Y |
| 16 | .1072 | .8867 | .8744 | .1 | .9 | .9 | .001 | Y |
| 17 | .8834 | .1213 | .1689 | .9 | .1 | .1 | .005 | Y |
| 18 | .9513 | .0413 | .9454 | .9 | .1 | .9 | .008 | Y |
| 19 | .0843 | .0588 | .2649 | .1 | .1 | .1 | .029 | Y |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 20 | .0954 | .2785 | .6071 | .1 | .1 | .9 | .118 | Y |
| 21 | .0694 | .9341 | .2985 | .1 | .9 | .1 | .042 | Y |
| 22 | .1436 | .8373 | .7425 | .1 | .9 | .9 | .031 | Y |
| 23 | .8949 | .1081 | .2015 | .9 | .1 | .1 | .01 | Y |
| 24 | .9539 | .0411 | .898 | .9 | .1 | .9 | .006 | Y |
| 25 | .0858 | .054 | .1462 | .1 | .1 | .1 | .004 | Y |
| 26 | .0819 | .4142 | .5233 | .1 | .1 | .9 | .241 | Y |
| 27 | .0722 | .93 | .2525 | .1 | .9 | .1 | .025 | Y |
| 28 | .0983 | .8995 | .8813 | .1 | .9 | .9 | .000 | Y |
| 29 | .8799 | .1147 | .3761 | .9 | .1 | .1 | .077 | Y |
| 30 | .9505 | .0423 | .9313 | .9 | .1 | .9 | .007 | Y |
| 31 | .099 | .0382 | .741 | .1 | .1 | .1 | .415 | N |
| 32 | .0949 | .5232 | .1343 | .1 | .1 | .9 | .765 | N |
| 33 | .0715 | .9333 | .3199 | .1 | .9 | .1 | .050 | Y |
| 34 | .1462 | .8477 | .8259 | .1 | .9 | .9 | .010 | Y |
| 35 | .9084 | .0872 | .3857 | .9 | .1 | .1 | .082 | Y |
| 36 | .9535 | .0389 | .9591 | .9 | .1 | .9 | .010 | Y |
| 37 | .0954 | .0484 | .4611 | .1 | .1 | .1 | .133 | N |
| 38 | .0799 | .7316 | .1205 | .1 | .1 | .9 | 1.007 | N |
| 39 | .0699 | .9317 | .4486 | .1 | .9 | .1 | .123 | Y |
| 40 | .2449 | .7327 | .5346 | .1 | .9 | .9 | .182 | Y |
| 41 | .9018 | .0997 | .2644 | .9 | .1 | .1 | .027 | Y |
| 42 | .9503 | .0422 | .9411 | .9 | .1 | .9 | .008 | Y |
| 43 | .0846 | .0679 | .2219 | .1 | .1 | .1 | .016 | Y |
| 44 | .0808 | .6237 | .1679 | .1 | .1 | .9 | .811 | N |
| 45 | .0707 | .9328 | .3898 | .1 | .9 | .1 | .086 | Y |
| 46 | .0958 | .8961 | .9201 | .1 | .9 | .9 | .000 | Y |
| 47 | .9094 | .0896 | .3126 | .9 | .1 | .1 | .045 | Y |
| 48 | .9535 | .0395 | .9477 | .9 | .1 | .9 | .009 | Y |
| 49 | .0917 | .0498 | .3779 | .1 | .1 | .1 | .080 | Y |
| 50 | .0787 | .7936 | .0668 | .1 | .1 | .9 | 1.176 | N |
| 51 | .0697 | .9337 | .5143 | .1 | .9 | .1 | .174 | N |
| 52 | .152 | .8512 | .7619 | .1 | .9 | .9 | .024 | Y |
| 53 | .861 | .1402 | .1817 | .9 | .1 | .1 | .010 | Y |
| 54 | .9523 | .0388 | .9642 | .9 | .1 | .9 | .011 | Y |
| 55 | .0886 | .0611 | .2813 | .1 | .1 | .1 | .035 | Y |
| 56 | .0889 | .6375 | .195 | .1 | .1 | .9 | .786 | N |
| 57 | .0742 | .9286 | .2008 | .1 | .9 | .1 | .012 | Y |
| 58 | .3455 | .6746 | .4294 | .1 | .9 | .9 | .333 | N |
| 59 | .9306 | .0727 | .2228 | .9 | .1 | .1 | .017 | Y |
| 60 | .9514 | .0374 | .9795 | .9 | .1 | .9 | .013 | Y |

mean error squared = 0.1248855

percent correct = 85.0%

Testing using Gray Coded outputs on the training data set

Network size 90-4-3 trained to 0.02 mse

| Pat No. | Out 3 | Out 2 | Out 1 | Tar 3 | Tar 2 | Tar 1 | mse | correct |
|---|---|---|---|---|---|---|---|---|
| 1 | .1104 | .0698 | .0913 | .1 | .1 | .1 | .001 | Y |
| 2 | .111 | .0848 | .702 | .1 | .1 | .9 | .040 | Y |
| 3 | .0513 | .5534 | .9554 | .1 | .9 | .9 | .126 | Y |
| 4 | .035 | .965 | .3914 | .1 | .9 | .1 | .093 | Y |
| 5 | .6434 | .8899 | .0273 | .9 | .9 | .1 | .071 | Y |
| 6 | .9517 | .8765 | .7084 | .9 | .9 | .9 | .040 | Y |
| 7 | .1104 | .0698 | .0913 | .1 | .1 | .1 | .001 | Y |
| 8 | .1113 | .0849 | .7272 | .1 | .1 | .9 | .030 | Y |
| 9 | .0519 | .5568 | .96 | .1 | .9 | .9 | .124 | Y |
| 10 | .0364 | .9654 | .3766 | .1 | .9 | .1 | .085 | Y |
| 11 | .6672 | .8866 | .0345 | .9 | .9 | .1 | .059 | Y |
| 12 | .9525 | .8763 | .714 | .9 | .9 | .9 | .038 | Y |
| 13 | .1112 | .0697 | .0914 | .1 | .1 | .1 | .001 | Y |
| 14 | .1058 | .1153 | .9653 | .1 | .1 | .9 | .005 | Y |
| 15 | .075 | .3099 | .9816 | .1 | .9 | .9 | .356 | N |
| 16 | .0508 | .9589 | .3271 | .1 | .9 | .1 | .057 | Y |
| 17 | .7079 | .8851 | .0358 | .9 | .9 | .1 | .041 | Y |
| 18 | .9596 | .8766 | .7913 | .9 | .9 | .9 | .016 | Y |
| 19 | .1104 | .0698 | .0913 | .1 | .1 | .1 | .001 | Y |
| 20 | .1101 | .0923 | .8521 | .1 | .1 | .9 | .002 | Y |
| 21 | .0424 | .7087 | .9282 | .1 | .9 | .9 | .041 | Y |
| 22 | .0618 | .9604 | .2637 | .1 | .9 | .1 | .032 | Y |
| 23 | .7177 | .8831 | .0341 | .9 | .9 | .1 | .038 | Y |
| 24 | .9647 | .877 | .8475 | .9 | .9 | .9 | .007 | Y |
| 25 | .1104 | .0698 | .0913 | .1 | .1 | .1 | .001 | Y |
| 26 | .1072 | .108 | .9513 | .1 | .1 | .9 | .003 | Y |
| 27 | .043 | .7032 | .9286 | .1 | .9 | .9 | .043 | Y |
| 28 | .1086 | .9492 | .1892 | .1 | .9 | .1 | .010 | Y |
| 29 | .7696 | .8781 | .0411 | .9 | .9 | .1 | .021 | Y |
| 30 | .9699 | .8764 | .8854 | .9 | .9 | .9 | .006 | Y |
| 31 | .1104 | .0699 | .092 | .1 | .1 | .1 | .001 | Y |
| 32 | .1054 | .1111 | .9334 | .1 | .1 | .9 | .001 | Y |
| 33 | .0511 | .6123 | .9524 | .1 | .9 | .9 | .088 | Y |
| 34 | .1078 | .9526 | .1759 | .1 | .9 | .1 | .009 | Y |
| 35 | .8114 | .8754 | .0638 | .9 | .9 | .1 | .010 | Y |
| 36 | .9703 | .8766 | .8897 | .9 | .9 | .9 | .006 | Y |
| 37 | .1104 | .0699 | .0917 | .1 | .1 | .1 | .001 | Y |
| 38 | .1059 | .1132 | .9568 | .1 | .1 | .9 | .003 | Y |
| 39 | .0453 | .6694 | .9445 | .1 | .9 | .9 | .058 | Y |
| 40 | .1111 | .9523 | .1745 | .1 | .9 | .1 | .008 | Y |
| 41 | .7746 | .8803 | .0517 | .9 | .9 | .1 | .018 | Y |
| 42 | .9718 | .8768 | .9055 | .9 | .9 | .9 | .006 | Y |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 43 | .1104 | .0699 | .0929 | .1 | .1 | .1 | .001 | Y |
| 44 | .1088 | .0978 | .8866 | .1 | .1 | .9 | .000 | Y |
| 45 | .0478 | .6424 | .9405 | .1 | .9 | .9 | .071 | Y |
| 46 | .1282 | .9501 | .1545 | .1 | .9 | .1 | .006 | Y |
| 47 | .8331 | .876 | .0856 | .9 | .9 | .1 | .005 | Y |
| 48 | .9703 | .8765 | .8895 | .9 | .9 | .9 | .006 | Y |
| 49 | .1104 | .07 | .0943 | .1 | .1 | .1 | .001 | Y |
| 50 | .1096 | .109 | .9804 | .1 | .1 | .9 | .007 | Y |
| 51 | .0427 | .7495 | .9098 | .1 | .9 | .9 | .026 | Y |
| 52 | .196 | .9372 | .1233 | .1 | .9 | .1 | .011 | Y |
| 53 | .8469 | .8738 | .09 | .9 | .9 | .1 | .004 | Y |
| 54 | .9711 | .8762 | .893 | .9 | .9 | .9 | .006 | Y |
| 55 | .1104 | .0702 | .0986 | .1 | .1 | .1 | .001 | Y |
| 56 | .1033 | .1264 | .9748 | .1 | .1 | .9 | .006 | Y |
| 57 | .0329 | .838 | .8542 | .1 | .9 | .9 | .010 | Y |
| 58 | .0884 | .9546 | .209 | .1 | .9 | .1 | .015 | Y |
| 59 | .8443 | .8727 | .079 | .9 | .9 | .1 | .004 | Y |
| 60 | .9674 | .8774 | .876 | .9 | .9 | .9 | .006 | Y |

mean error squared = 2.972622E-02

percent correct = 98.33%

Testing using Gray Coded outputs on the test data set

Network size90-4-3 trained to 0.02 mse

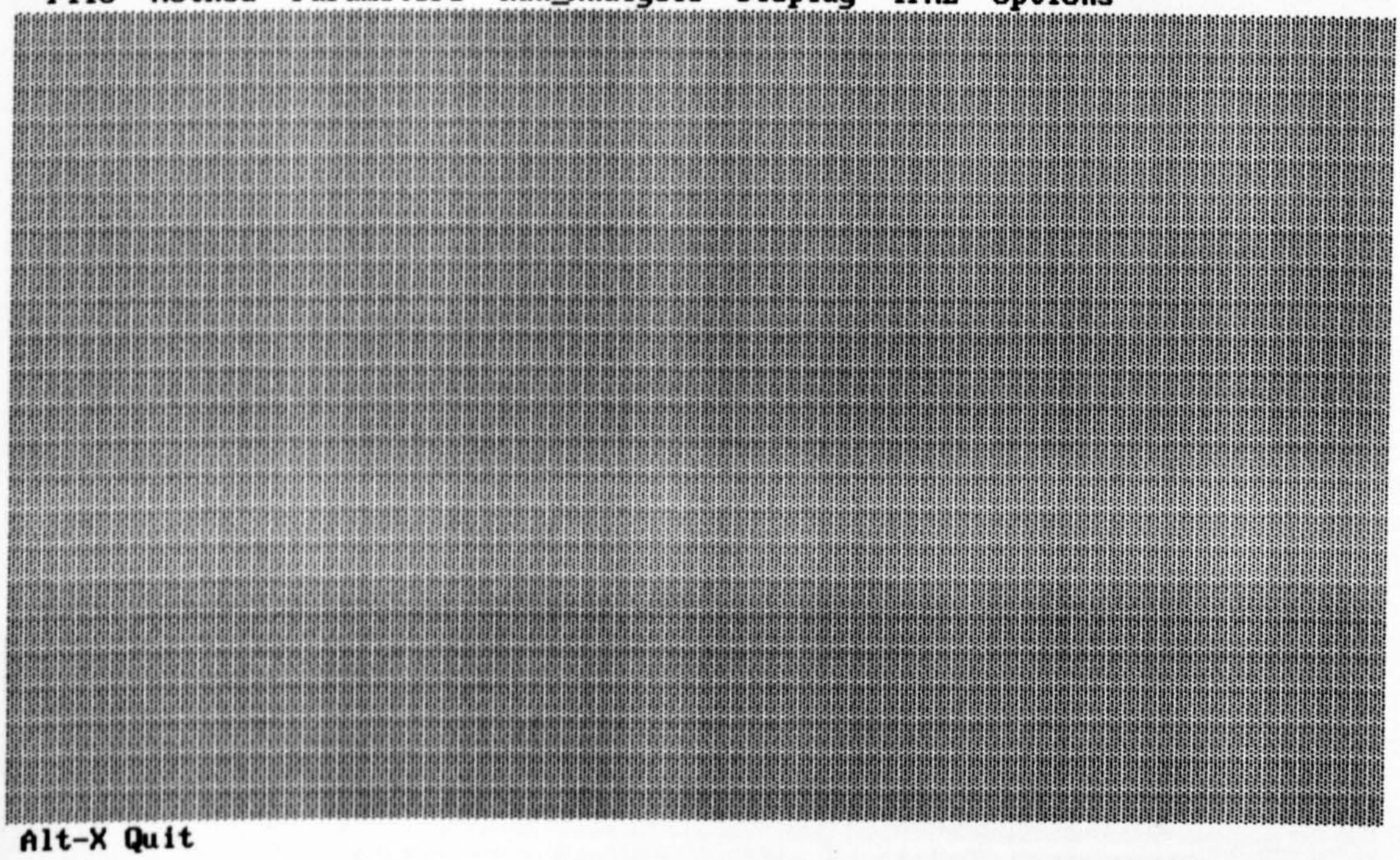
| Pat No. | Out 3 | Out 2 | Out 1 | Tar 3 | Tar 2 | Tar 1 | mse | correct |
|---|---|---|---|---|---|---|---|---|
| 1 | .1104 | .0704 | .1015 | .1 | .1 | .1 | .001 | Y |
| 2 | .108 | .113 | .98 | .1 | .1 | .9 | .007 | Y |
| 3 | .0303 | .8898 | .7908 | .1 | .9 | .9 | .017 | Y |
| 4 | .1481 | .9475 | .1339 | .1 | .9 | .1 | .006 | Y |
| 5 | .8373 | .874 | .0768 | .9 | .9 | .1 | .005 | Y |
| 6 | .9717 | .8761 | .8975 | .9 | .9 | .9 | .006 | Y |
| 7 | .1104 | .0704 | .1027 | .1 | .1 | .1 | .001 | Y |
| 8 | .0993 | .1404 | .9739 | .1 | .1 | .9 | .007 | Y |
| 9 | .0274 | .9027 | .7704 | .1 | .9 | .9 | .022 | Y |
| 10 | .269 | .9336 | .0816 | .1 | .9 | .1 | .030 | Y |
| 11 | .8656 | .875 | .1323 | .9 | .9 | .1 | .003 | Y |
| 12 | .9721 | .8764 | .9024 | .9 | .9 | .9 | .006 | Y |
| 13 | .1103 | .073 | .1572 | .1 | .1 | .1 | .004 | Y |
| 14 | .109 | .1162 | .9885 | .1 | .1 | .9 | .008 | Y |
| 15 | .0264 | .9184 | .7242 | .1 | .9 | .9 | .037 | Y |
| 16 | .293 | .9308 | .0758 | .1 | .9 | .1 | .039 | Y |
| 17 | .861 | .8727 | .1053 | .9 | .9 | .1 | .002 | Y |
| 18 | .973 | .8761 | .9067 | .9 | .9 | .9 | .006 | Y |
| 19 | .1105 | .072 | .1423 | .1 | .1 | .1 | .003 | Y |
| 20 | .1019 | .1394 | .9895 | .1 | .1 | .9 | .010 | Y |
| 21 | .0331 | .8716 | .827 | .1 | .9 | .9 | .011 | Y |
| 22 | .2444 | .9331 | .1177 | .1 | .9 | .1 | .022 | Y |

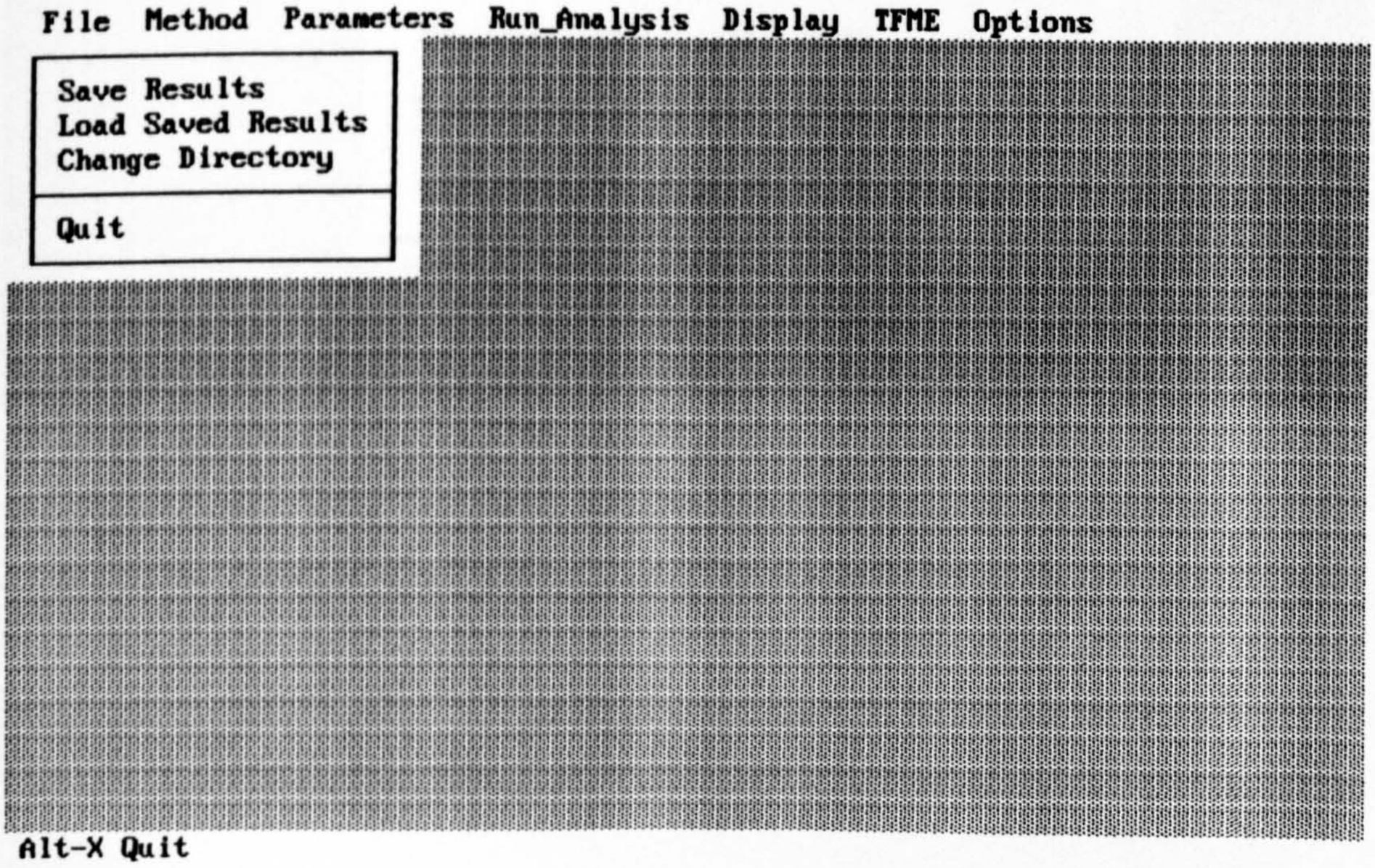
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 23 | .8887 | .873 | .1798 | .9 | .9 | .1 | .007 | Y |
| 24 | .9718 | .8766 | .9045 | .9 | .9 | .9 | .006 | Y |
| 25 | .1104 | .0705 | .1034 | .1 | .1 | .1 | .001 | Y |
| 26 | .1077 | .1207 | .9893 | .1 | .1 | .9 | .008 | Y |
| 27 | .0262 | .9161 | .7529 | .1 | .9 | .9 | .027 | Y |
| 28 | .2783 | .9327 | .0762 | .1 | .9 | .1 | .033 | Y |
| 29 | .892 | .8742 | .2042 | .9 | .9 | .1 | .012 | Y |
| 30 | .9724 | .8762 | .9037 | .9 | .9 | .9 | .006 | Y |
| 31 | .1104 | .079 | .3937 | .1 | .1 | .1 | .087 | Y |
| 32 | .0812 | .2422 | .9827 | .1 | .1 | .9 | .027 | Y |
| 33 | .0227 | .9438 | .6717 | .1 | .9 | .9 | .060 | Y |
| 34 | .3816 | .9216 | .0591 | .1 | .9 | .1 | .081 | Y |
| 35 | .9278 | .876 | .4669 | .9 | .9 | .1 | .136 | N |
| 36 | .9736 | .8761 | .9112 | .9 | .9 | .9 | .006 | Y |
| 37 | .1106 | .0763 | .2908 | .1 | .1 | .1 | .037 | Y |
| 38 | .0949 | .1708 | .9898 | .1 | .1 | .9 | .013 | Y |
| 39 | .0334 | .9047 | .7537 | .1 | .9 | .9 | .026 | Y |
| 40 | .3655 | .9219 | .0918 | .1 | .9 | .1 | .071 | Y |
| 41 | .9074 | .8727 | .2562 | .9 | .9 | .1 | .025 | Y |
| 42 | .9728 | .876 | .905 | .9 | .9 | .9 | .006 | Y |
| 43 | .1105 | .072 | .14 | .1 | .1 | .1 | .002 | Y |
| 44 | .0998 | .1492 | .9895 | .1 | .1 | .9 | .010 | Y |
| 45 | .0312 | .9216 | .7102 | .1 | .9 | .9 | .041 | Y |
| 46 | .2915 | .9316 | .0756 | .1 | .9 | .1 | .038 | Y |
| 47 | .924 | .8738 | .3857 | .9 | .9 | .1 | .083 | Y |
| 48 | .9732 | .8762 | .9099 | .9 | .9 | .9 | .006 | Y |
| 49 | .1105 | .0744 | .2124 | .1 | .1 | .1 | .013 | Y |
| 50 | .1008 | .152 | .9914 | .1 | .1 | .9 | .011 | Y |
| 51 | .0245 | .9592 | .5586 | .1 | .9 | .9 | .126 | Y |
| 52 | .3979 | .9194 | .0517 | .1 | .9 | .1 | .091 | Y |
| 53 | .843 | .879 | .129 | .9 | .9 | .1 | .005 | Y |
| 54 | .9735 | .8763 | .9121 | .9 | .9 | .9 | .006 | Y |
| 55 | .1104 | .0734 | .1764 | .1 | .1 | .1 | .007 | Y |
| 56 | .0798 | .2562 | .9847 | .1 | .1 | .9 | .032 | Y |
| 57 | .0365 | .8632 | .8019 | .1 | .9 | .9 | .015 | Y |
| 58 | .5527 | .9014 | .0329 | .1 | .9 | .1 | .209 | N |
| 59 | .9433 | .8751 | .6244 | .9 | .9 | .1 | .278 | N |
| 60 | .9739 | .8764 | .9154 | .9 | .9 | .9 | .006 | Y |

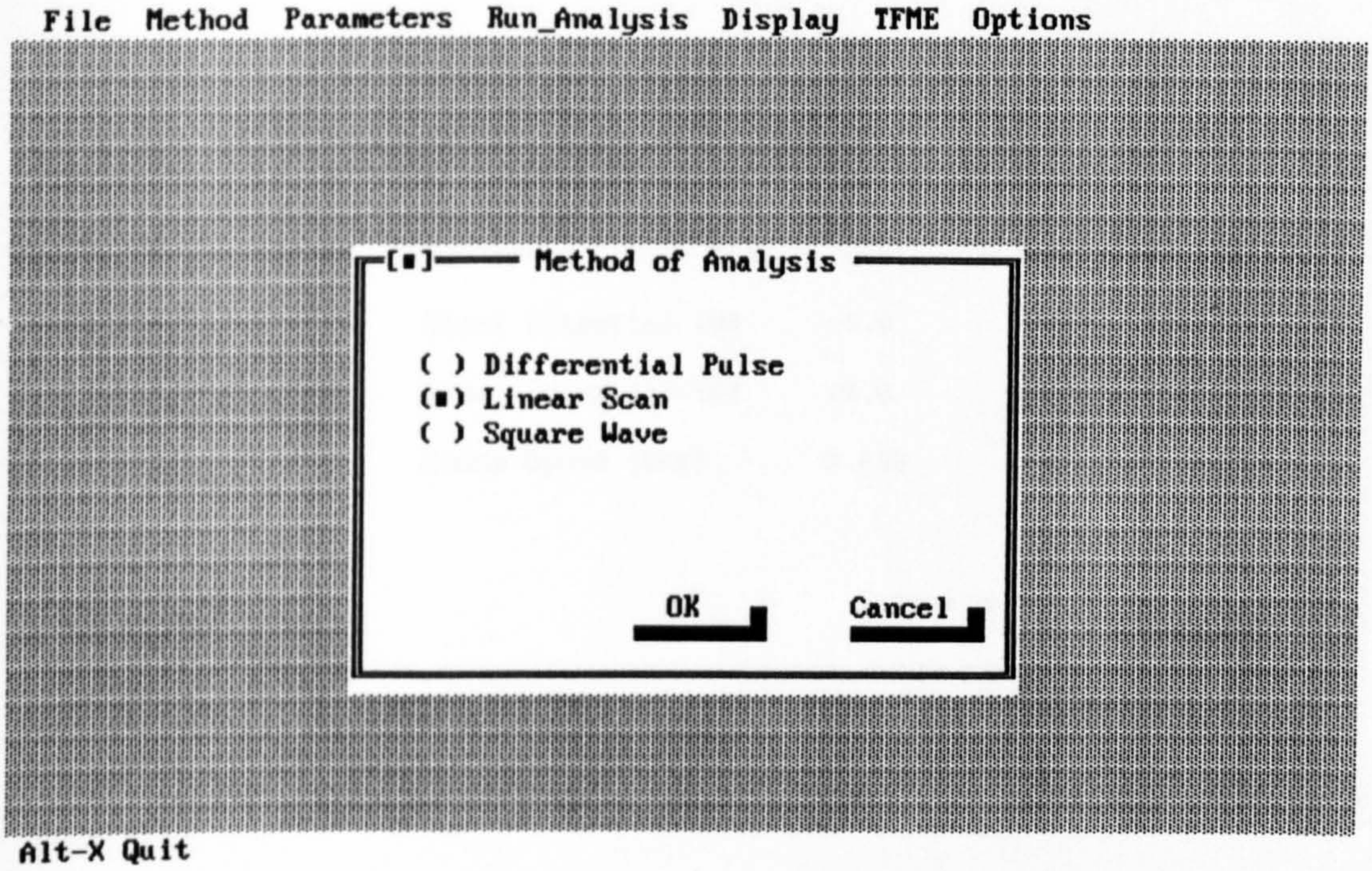
mean error squared = 3.180071E-02

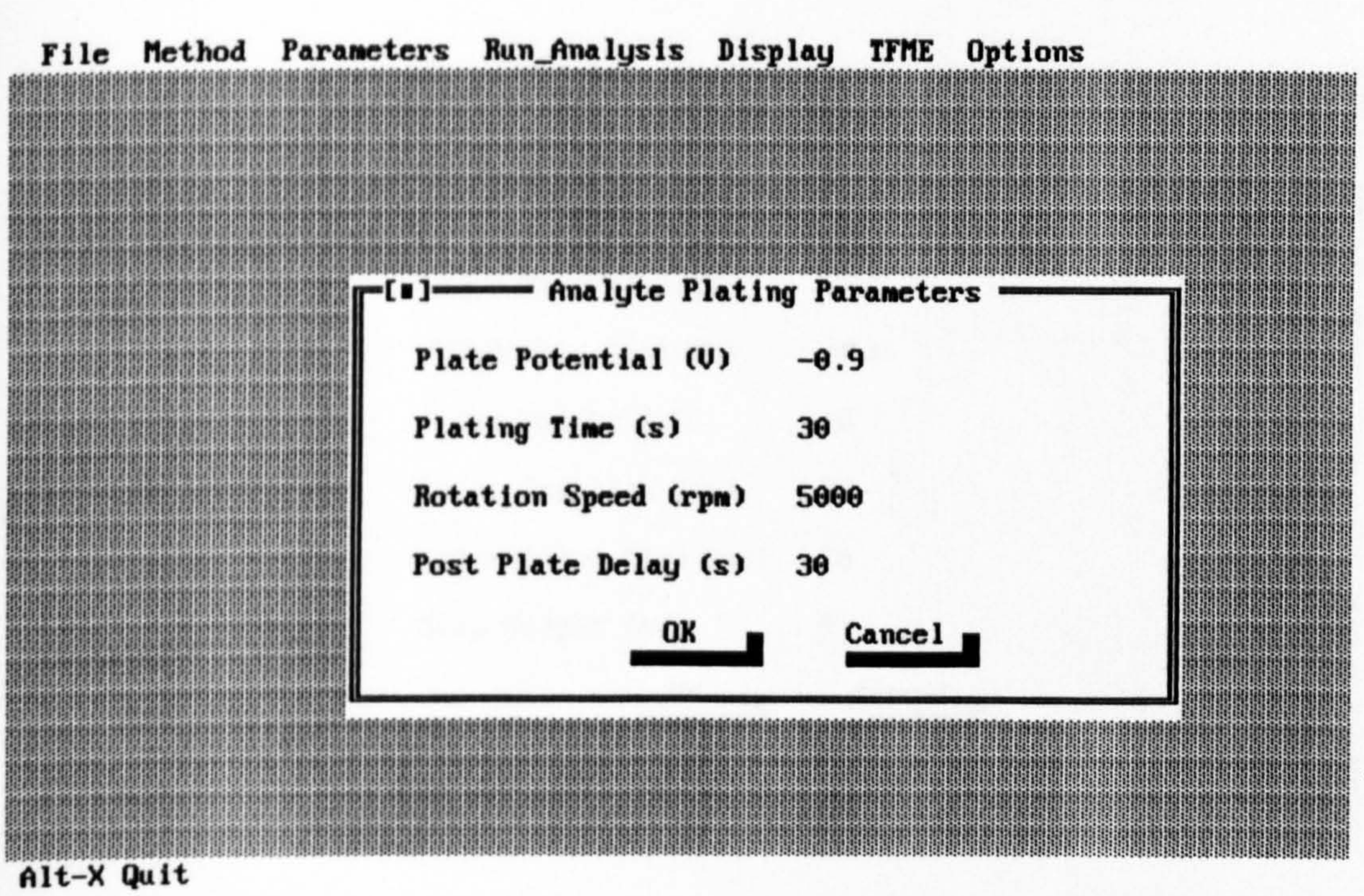
percent correct = 95.0%

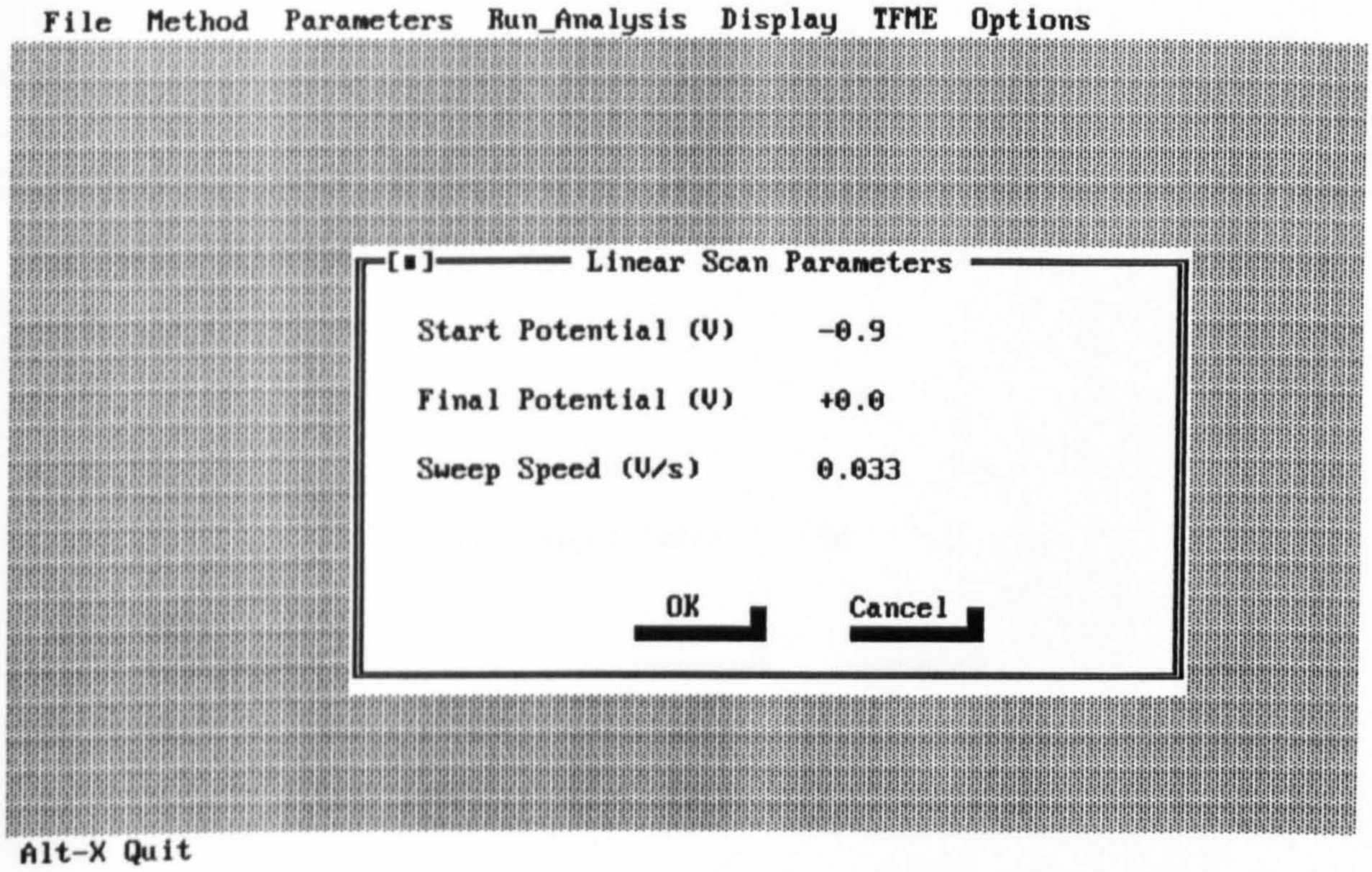
# Appendix 4 User interface screens from analytical software

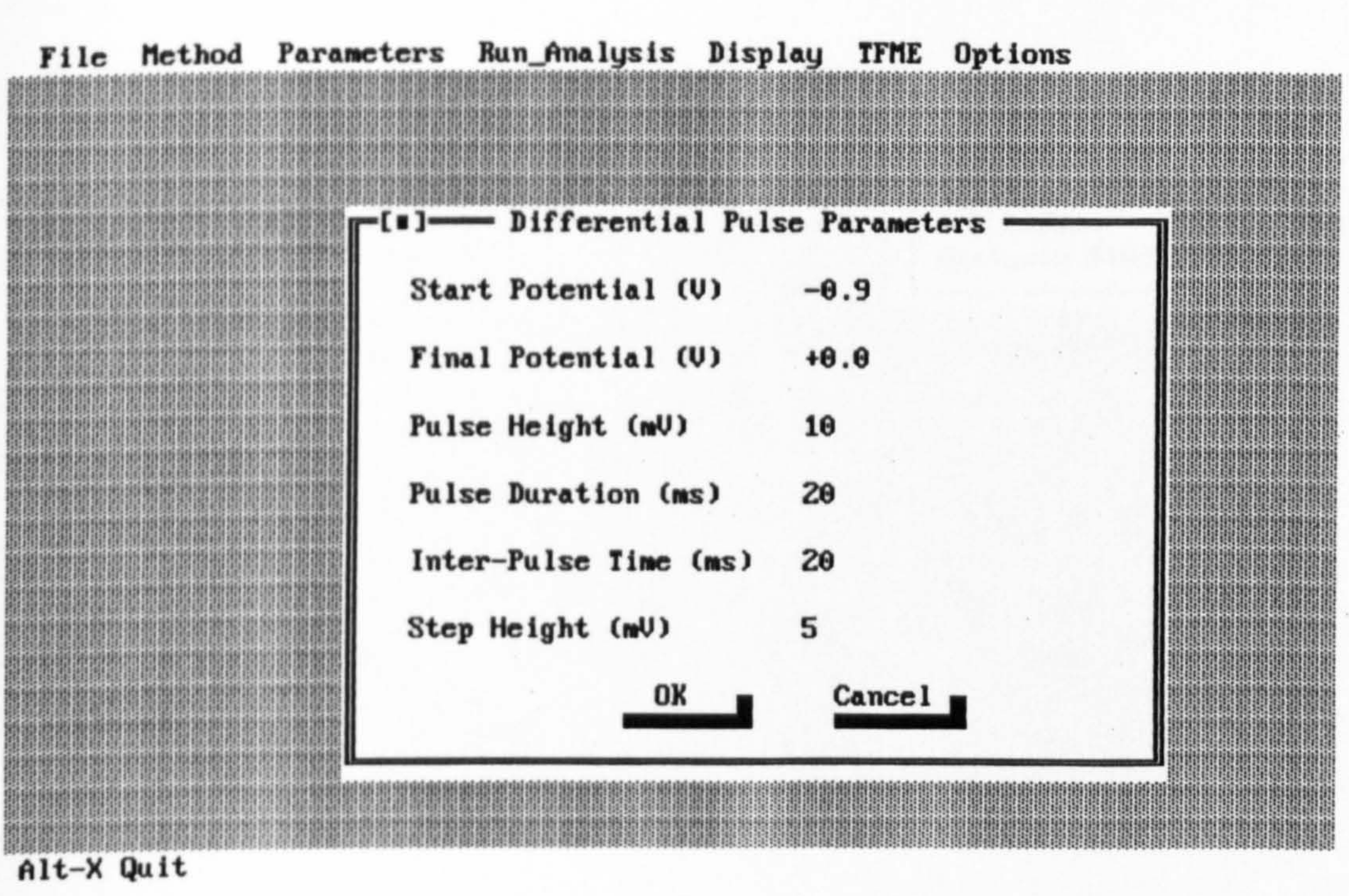



File Method Parameters Run_Analysis Display TFME Options
[■] Method of Analysis
( ) Differential Pulse
(■) Linear Scan
( ) Square Wave
OK
Cancel
Alt-X Quit


File Method Parameters Run_Analysis Display TFME Options
[■] Analyte Plating Parameters
Plate Potential (V) -0.9
Plating Time (s) 30
Rotation Speed (rpm) 5000
Post Plate Delay (s) 30
OK
Cancel
Alt-X Quit


File Method Parameters Run_Analysis Display TFME Options
[■] Linear Scan Parameters
Start Potential (V) -0.9
Final Potential (V) +0.0
Sweep Speed (V/s) 0.033
OK
Cancel
Alt-X Quit


File Method Parameters Run_Analysis Display TFME Options
[■] Differential Pulse Parameters
Start Potential (V) -0.9
Final Potential (V) +0.0
Pulse Height (mV) 10
Pulse Duration (ms) 20
Inter-Pulse Time (ms) 20
Step Height (mV) 5
OK
Cancel
Alt-X Quit

File Method Parameters Run_Analysis Display TFME Options

[■] Square Wave Parameters

| | |
|---|---|
| Start Potential (V) | -0.9 |
| Final Potential (V) | +0.0 |
| Frequency (Hz) | 200 |
| Scan Rate (V/s) | 0.1 |
| Pulse Height (mV) | 50 |

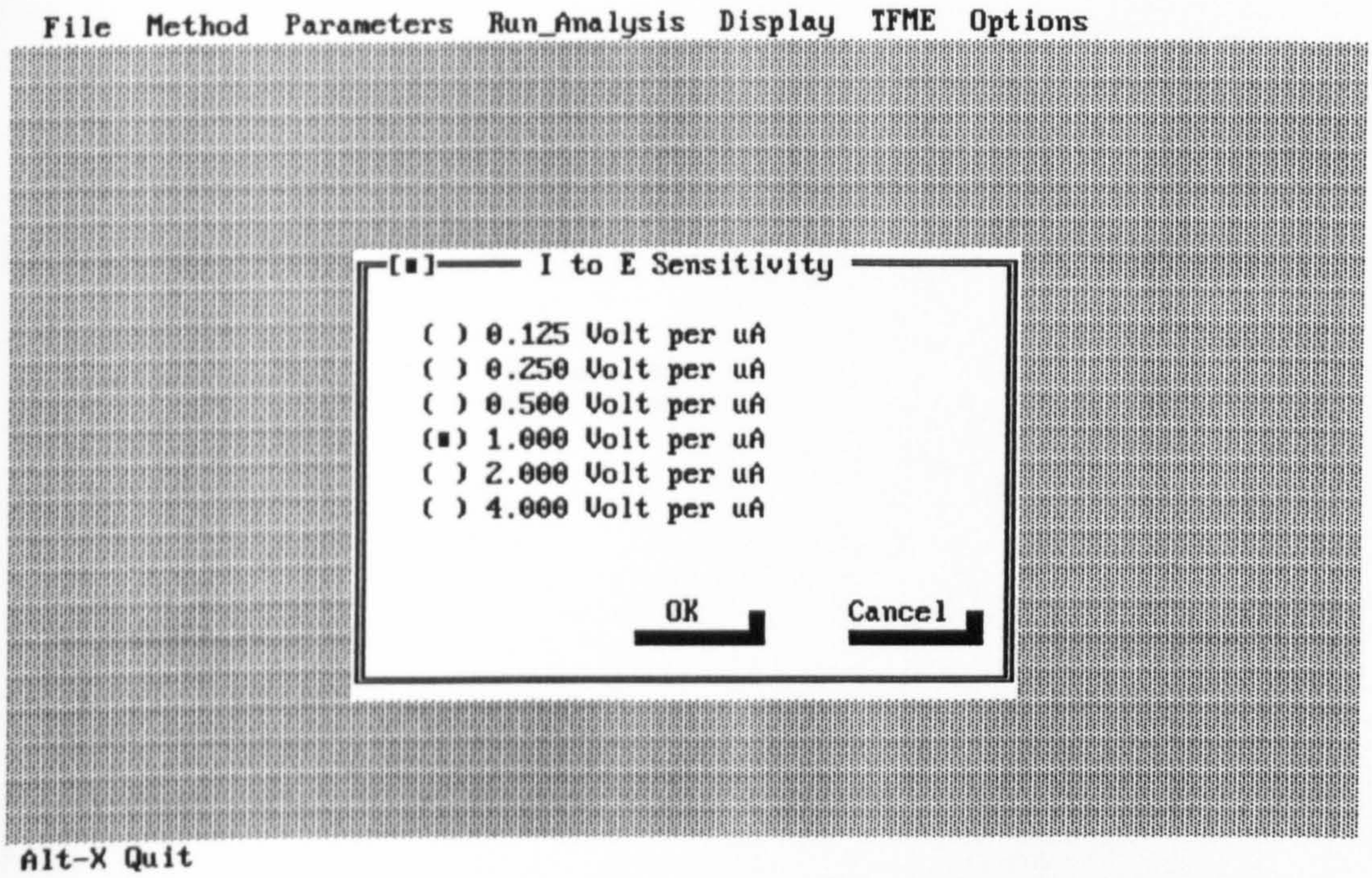
OK Cancel

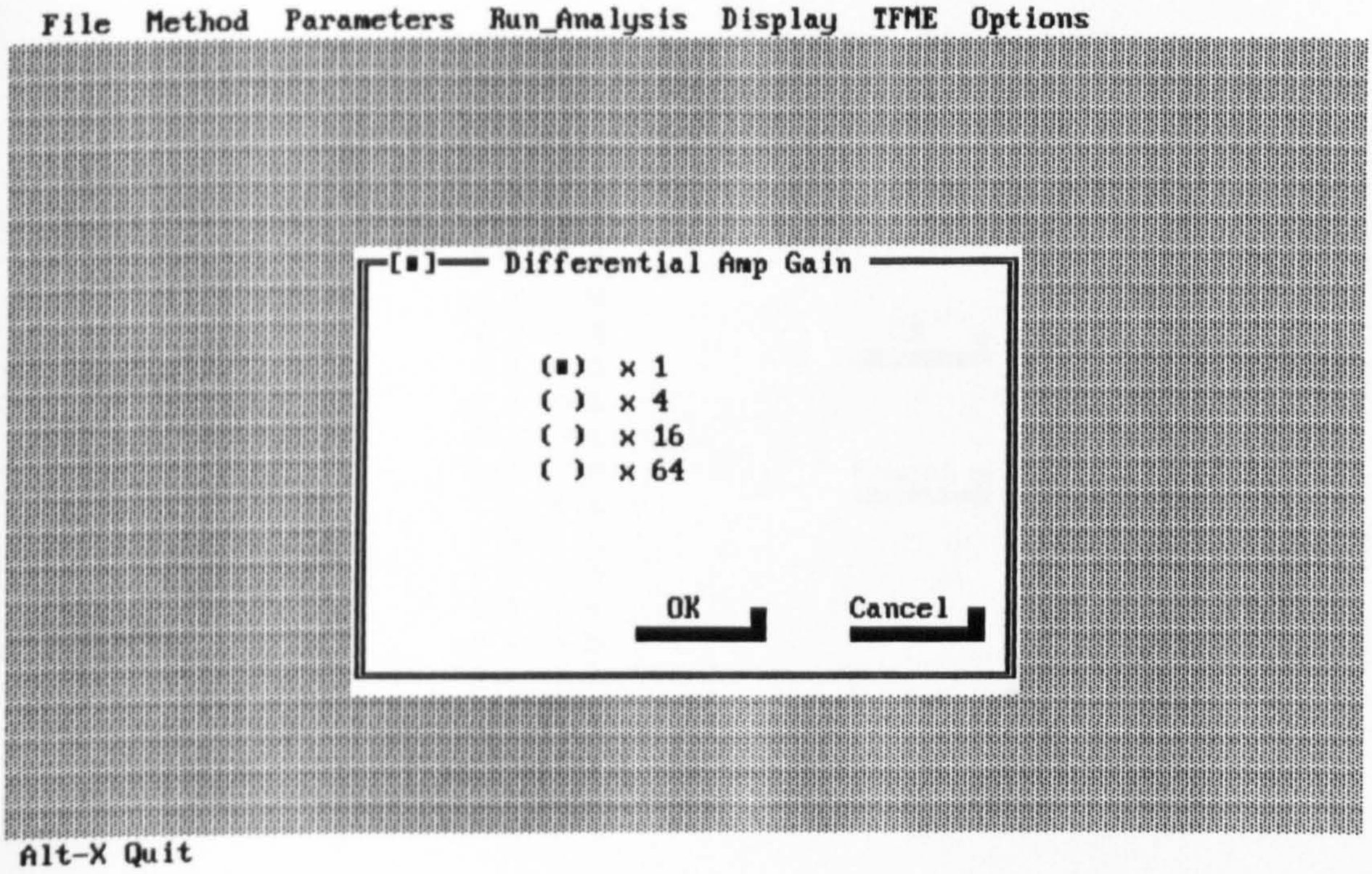
Alt-X Quit

File Method Parameters Run_Analysis Display TFME Options

I/E Sensitivity F8
Diff Amp Gain F9
I/E Offset Range F10
I/E Offset
Analysis Description

Alt-X Quit


File Method Parameters Run_Analysis Display TFME Options
[■] I to E Sensitivity
( ) 0.125 Volt per uA
( ) 0.250 Volt per uA
( ) 0.500 Volt per uA
(■) 1.000 Volt per uA
( ) 2.000 Volt per uA
( ) 4.000 Volt per uA
OK
Cancel
Alt-X Quit


File Method Parameters Run_Analysis Display TFME Options
[■] Differential Amp Gain
(■) x 1
( ) x 4
( ) x 16
( ) x 64
OK
Cancel
Alt-X Quit

File Method Parameters Run_Analysis Display TFME Options

[■] I to E OffsetRange

(■) 5 micro Amp
( ) 50 micro Amp

OK Cancel

Alt-X Quit

File Method Parameters Run_Analysis Display TFME Options

[■] I to E Offset

( ) - 8
( ) - 7
( ) - 6
( ) - 5
( ) - 4
( ) - 3
( ) - 2
( ) - 1
(■) 0
( ) + 1
( ) + 2
( ) + 3
( ) + 4
( ) + 5
( ) + 6
( ) + 7

OK
Cancel

Alt-X Quit

File Method Parameters Run_Analysis Display TFME Options

[■] Analyte Description

Description

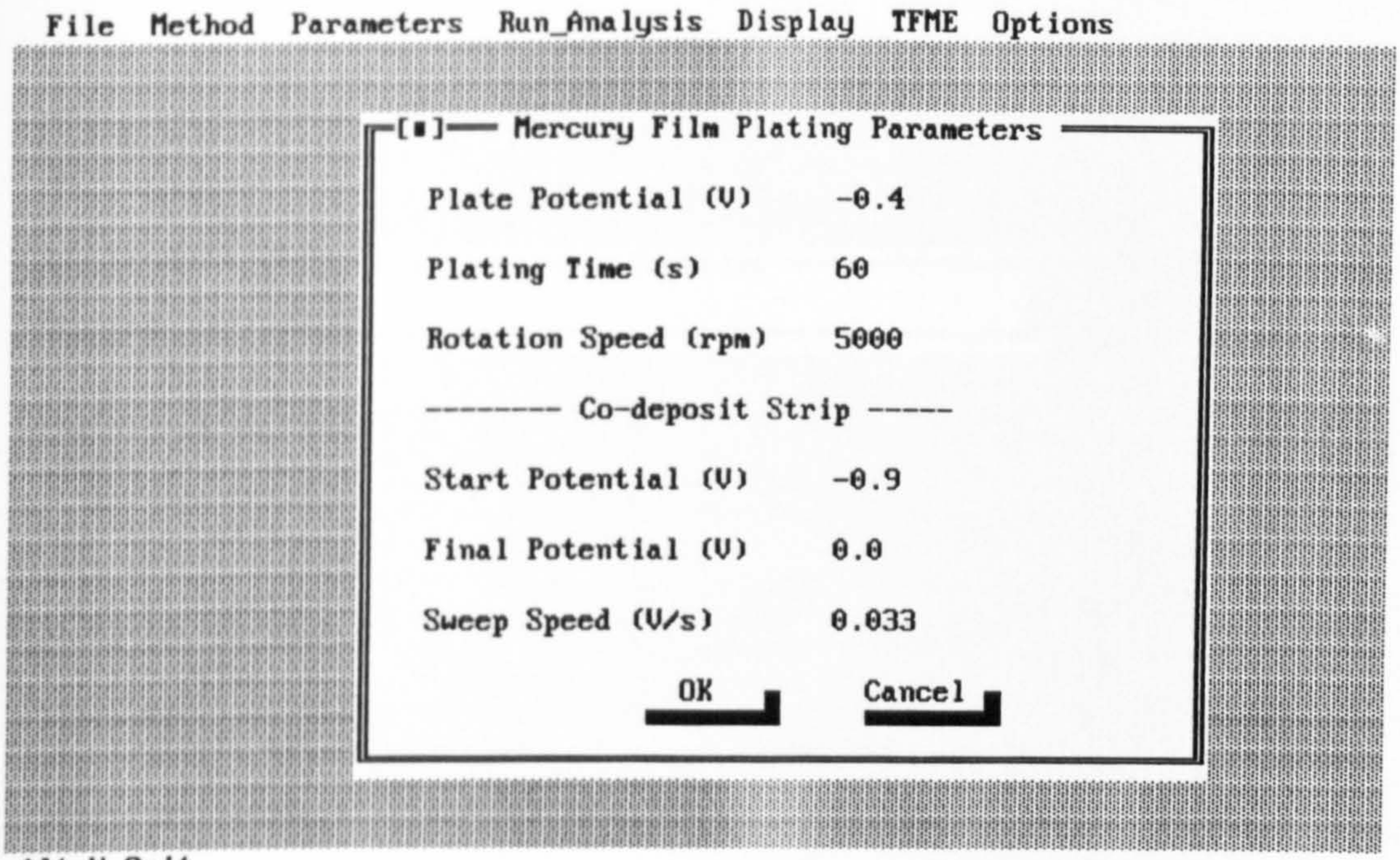
OK Cancel

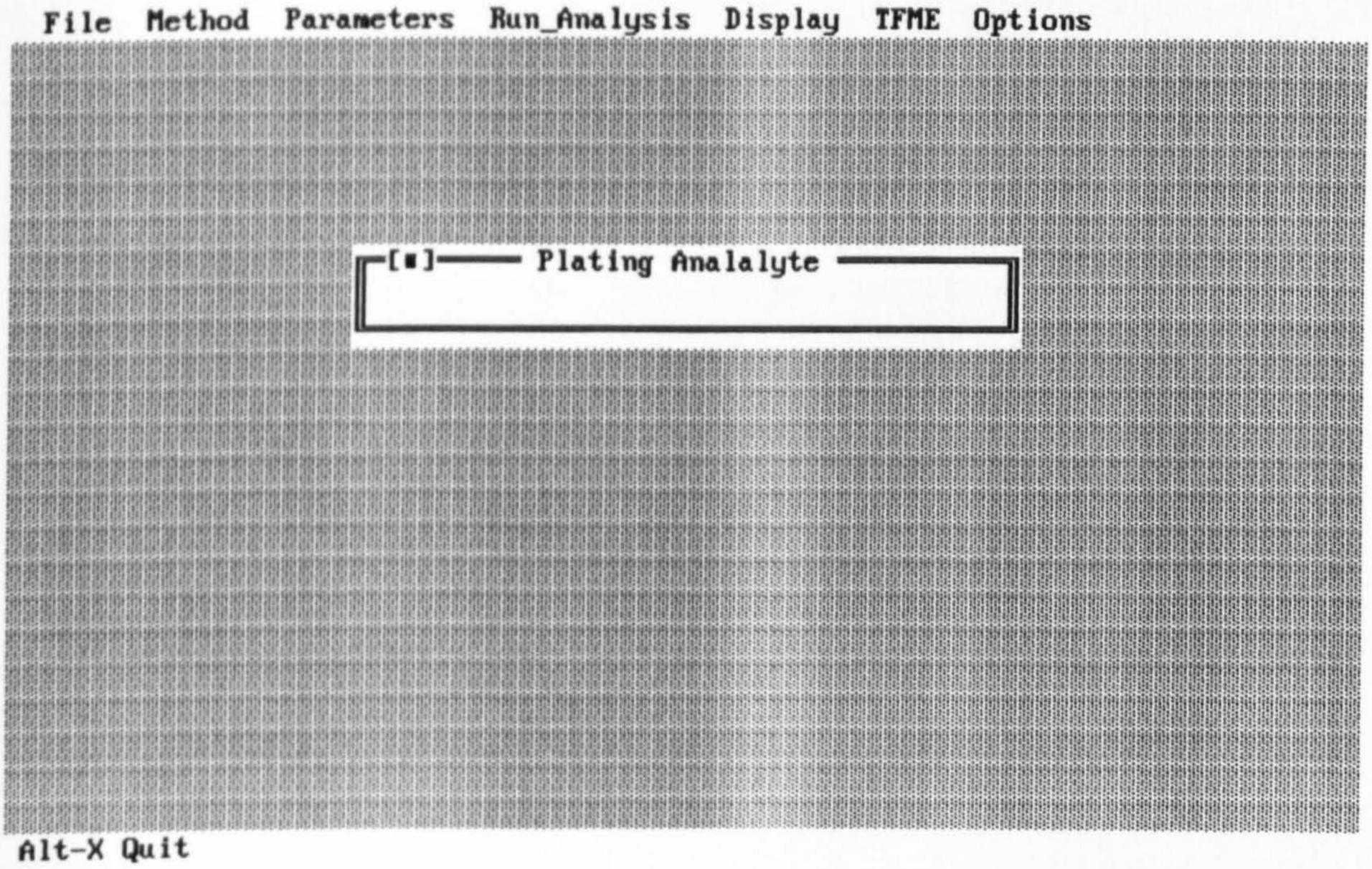
Alt-X Quit

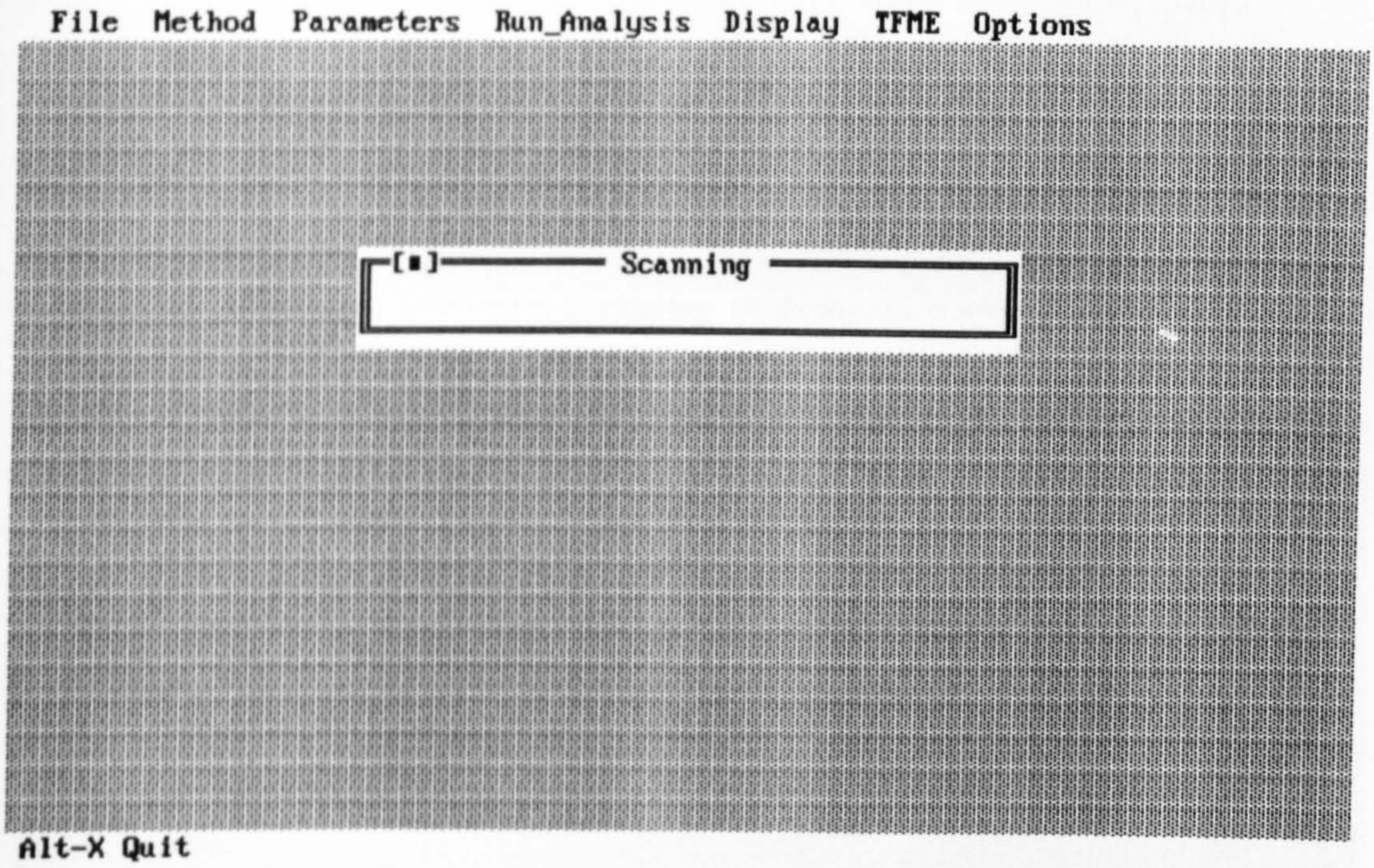
File Method Parameters Run_Analysis Display TFME Options

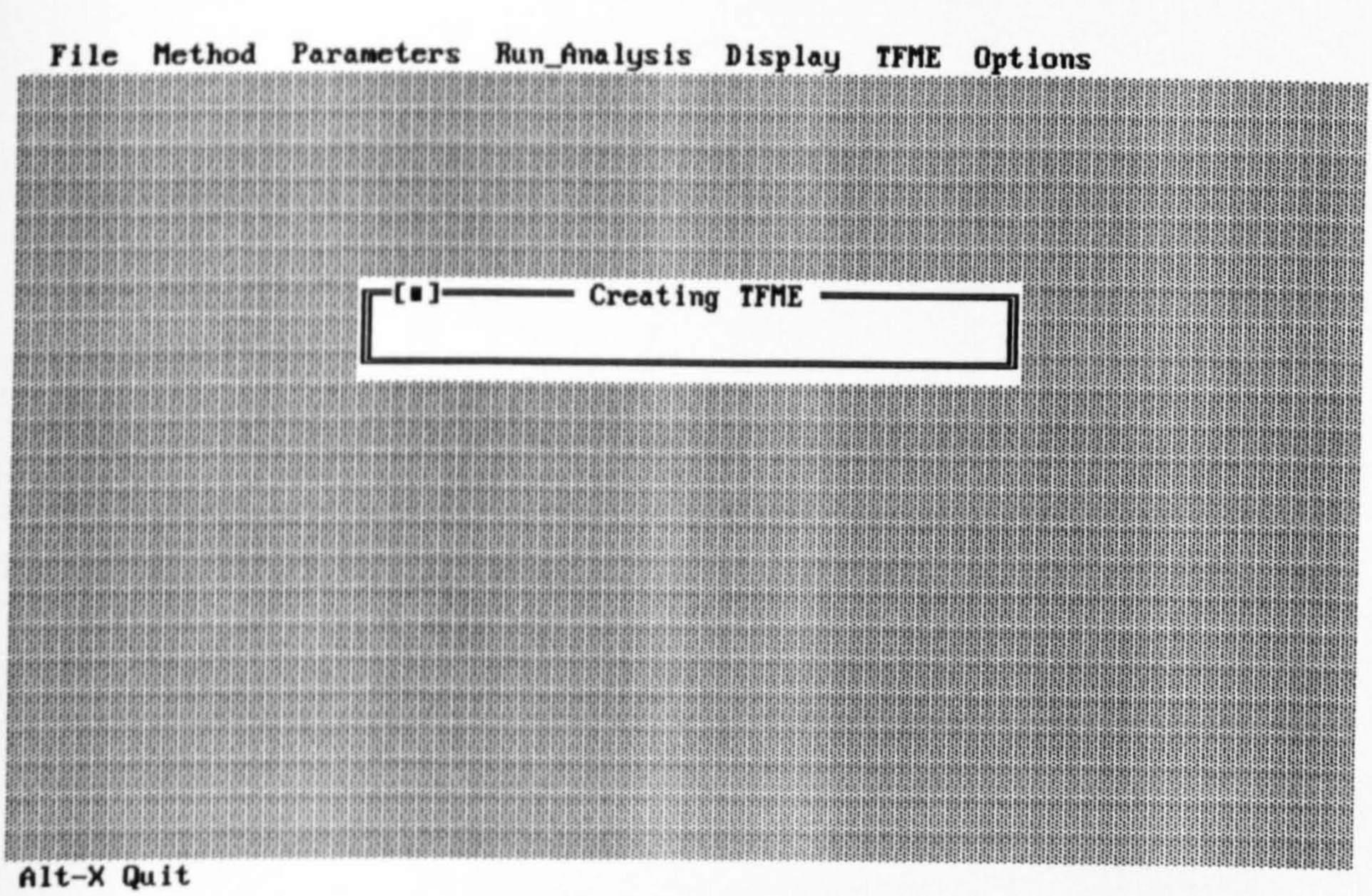
Parameters F6
Create Hg Electrode F7
Strip Co-Deposits

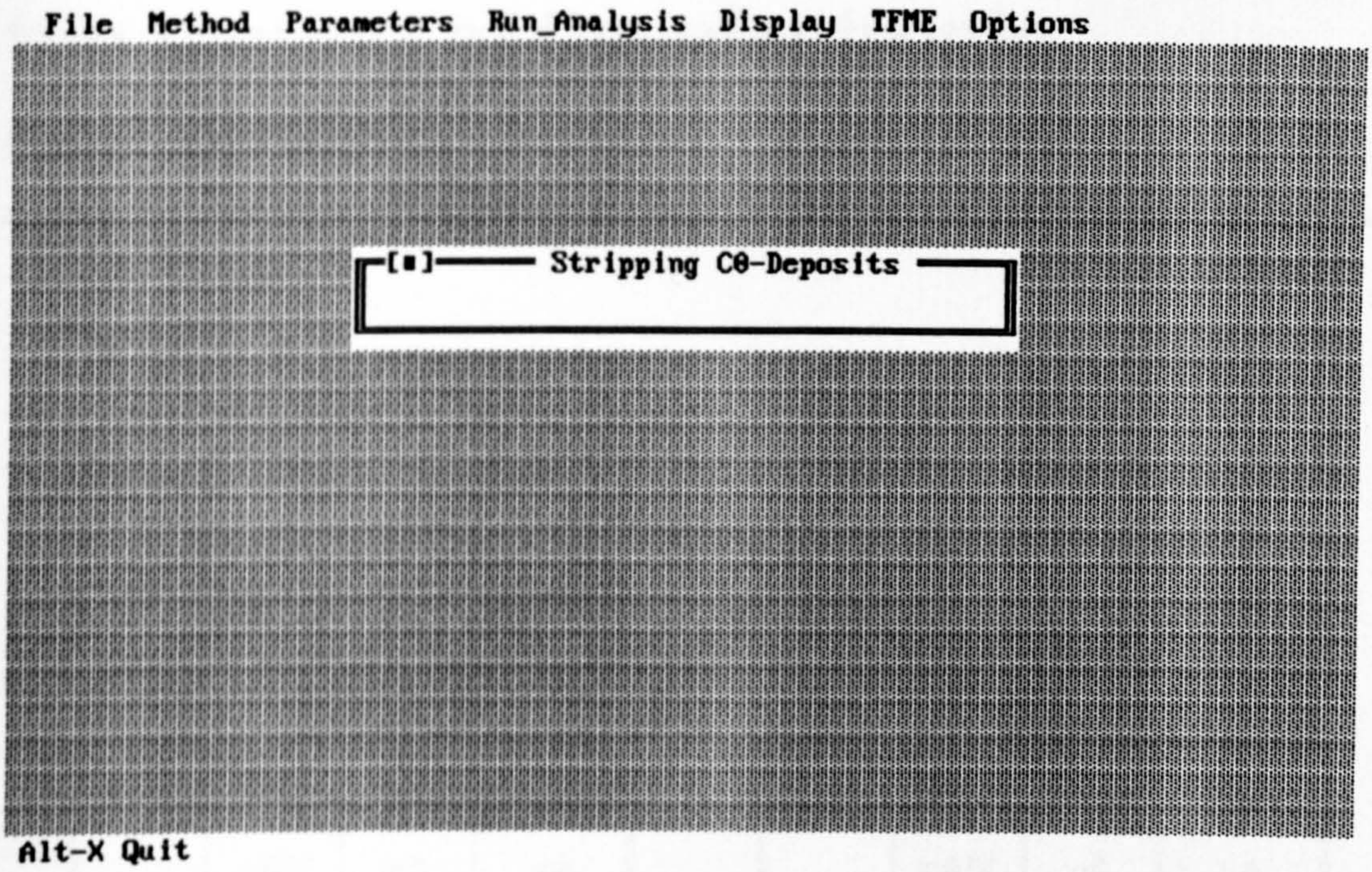
Alt-X Quit


File Method Parameters Run_Analysis Display TFME Options
[■] Mercury Film Plating Parameters
Plate Potential (V) -0.4
Plating Time (s) 60
Rotation Speed (rpm) 5000
-------- Co-deposit Strip -----
Start Potential (V) -0.9
Final Potential (V) 0.0
Sweep Speed (V/s) 0.033
OK
Cancel
Alt-X Quit


File Method Parameters Run_Analysis Display TFME Options
[■] Plating Analalyte
Alt-X Quit


File Method Parameters Run_Analysis Display TFME Options
[■]
Scanning
Alt-X Quit


File Method Parameters Run_Analysis Display TFME Options
[■]
Creating TFME
Alt-X Quit


File Method Parameters Run_Analysis Display TFME Options
[■] Stripping CΘ-Deposits
Alt-X Quit

# Appendix 5 Data obtained during the testing of the electronic equipment

## Testing of the I to E converter.

The following table shows the saturation currents for the I to E converter for all combinations of sensitivity and differential amplifier gain.

| Gain | Sens nA/V | $I_{max}$ μA | $I_{min}$ μA |
|---|---|---|---|
| 1 | 250 | 1.075 | -1.076 |
| 1 | 500 | 2.273 | -2.275 |
| 1 | 1000 | 4.950 | -4.955 |
| 1 | 2000 | 9.709 | -9.718 |
| 1 | 4000 | 20.161 | -20.076 |
| 1 | 8000 | 42.958 | -42.777 |
| 4 | 250 | 0.269 | -0.269 |
| 4 | 500 | 0.568 | -0.569 |
| 4 | 1000 | 1.244 | -1.239 |
| 4 | 2000 | 2.427 | -2.430 |
| 4 | 4000 | 5.040 | -5.072 |
| 4 | 8000 | 10.684 | -10.694 |
| 16 | 250 | 0.067 | -0.067 |
| 16 | 500 | 0.144 | -0.142 |
| 16 | 1000 | 0.309 | -0.310 |
| 16 | 2000 | 0.607 | -0.607 |
| 16 | 4000 | 1.267 | -1.261 |
| 16 | 8000 | 2.671 | -2.674 |
| 64 | 250 | 0.017 | -0.017 |
| 64 | 500 | 0.036 | -0.036 |
| 64 | 1000 | 0.077 | -0.077 |
| 64 | 2000 | 0.152 | -0.152 |
| 64 | 4000 | 0.317 | -0.315 |
| 64 | 8000 | 0.668 | -0.668 |

Table 41. Saturation current values for various gains and sensitivities

The positive values in table 41 are shown below, plotted on a log-log scale.

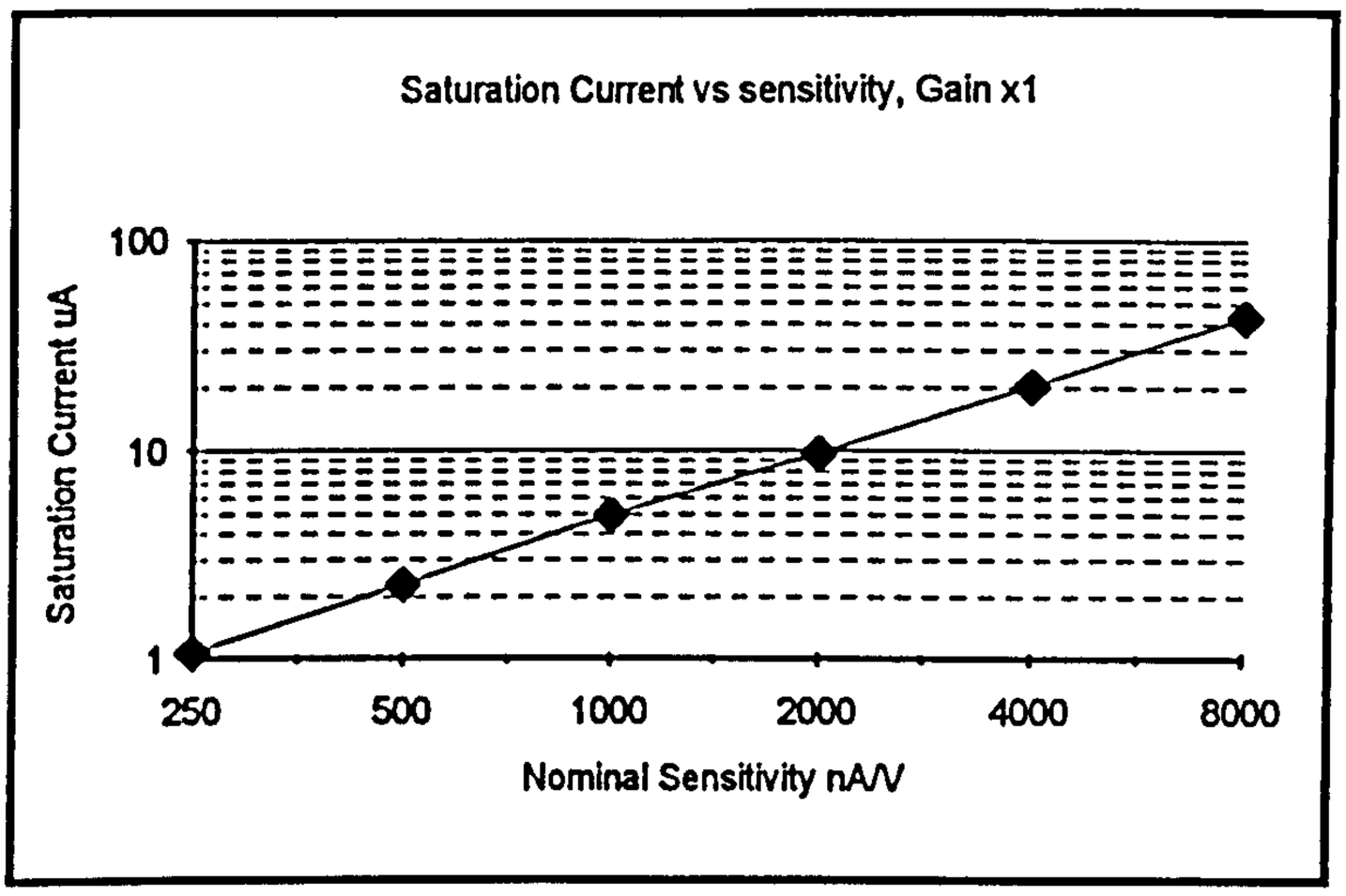



Figure 95. Saturation current - Gain x1

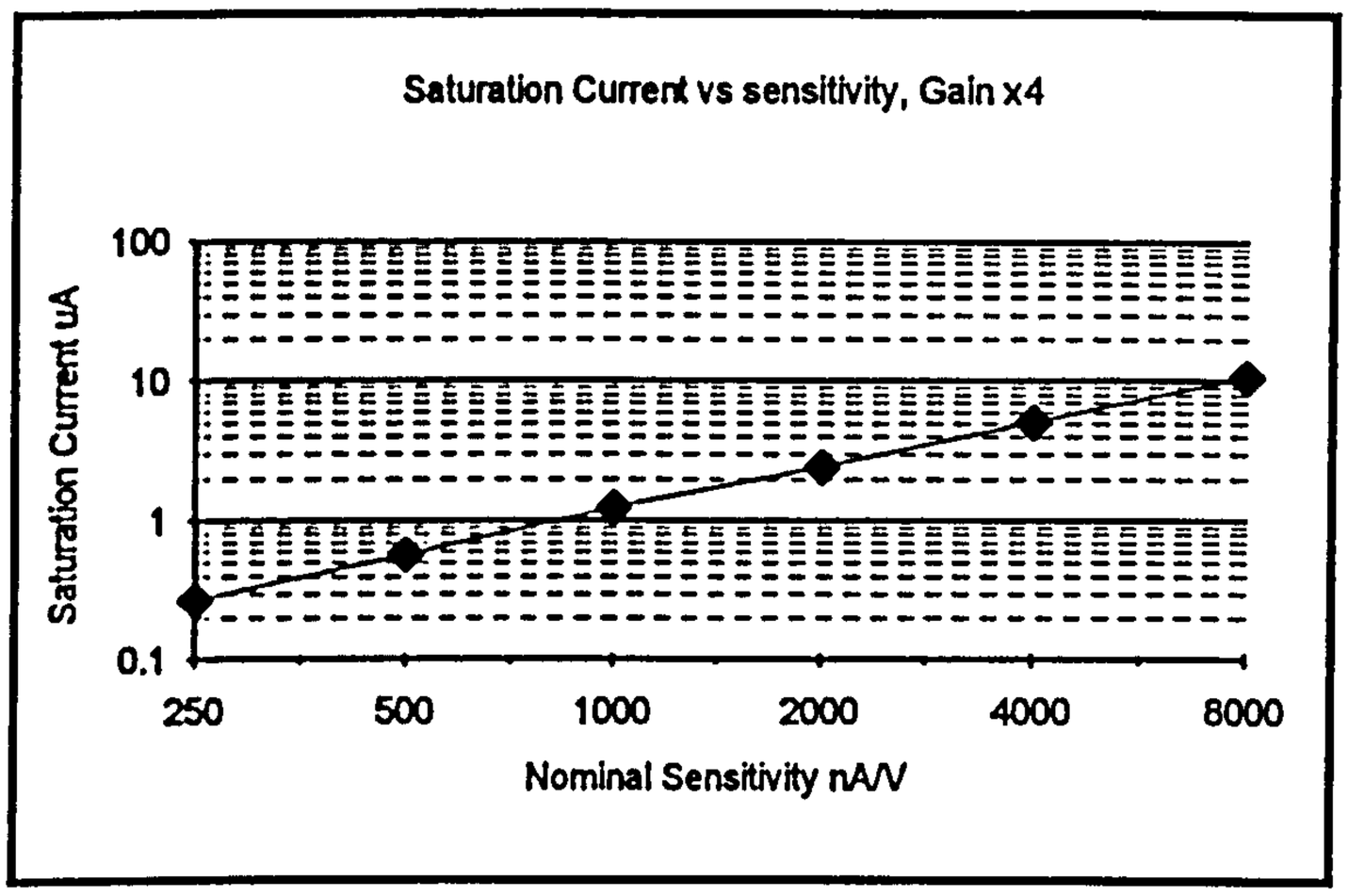



Figure 96. Saturation current - Gain x4

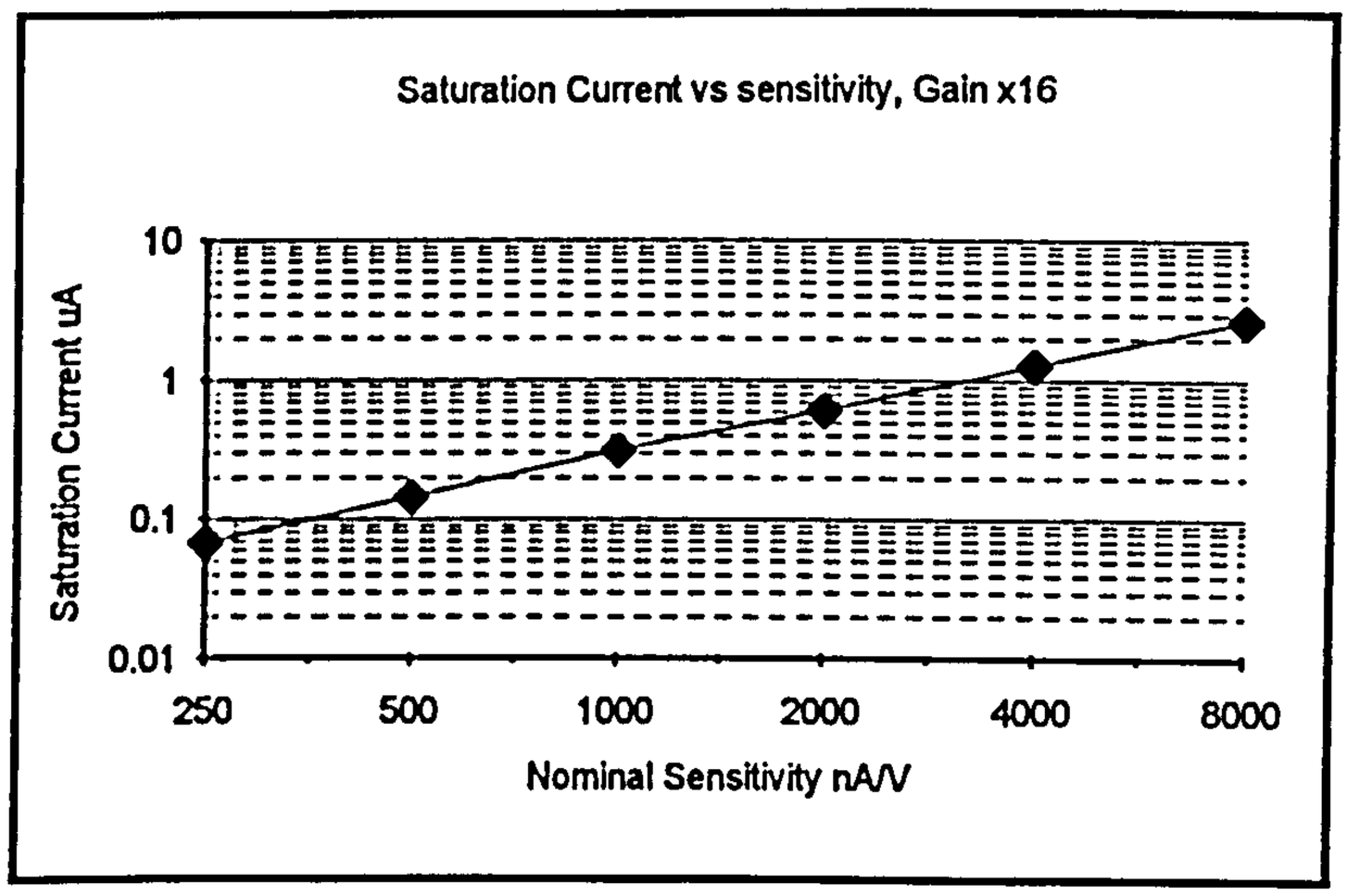



Figure 97. Saturation current - Gain x16

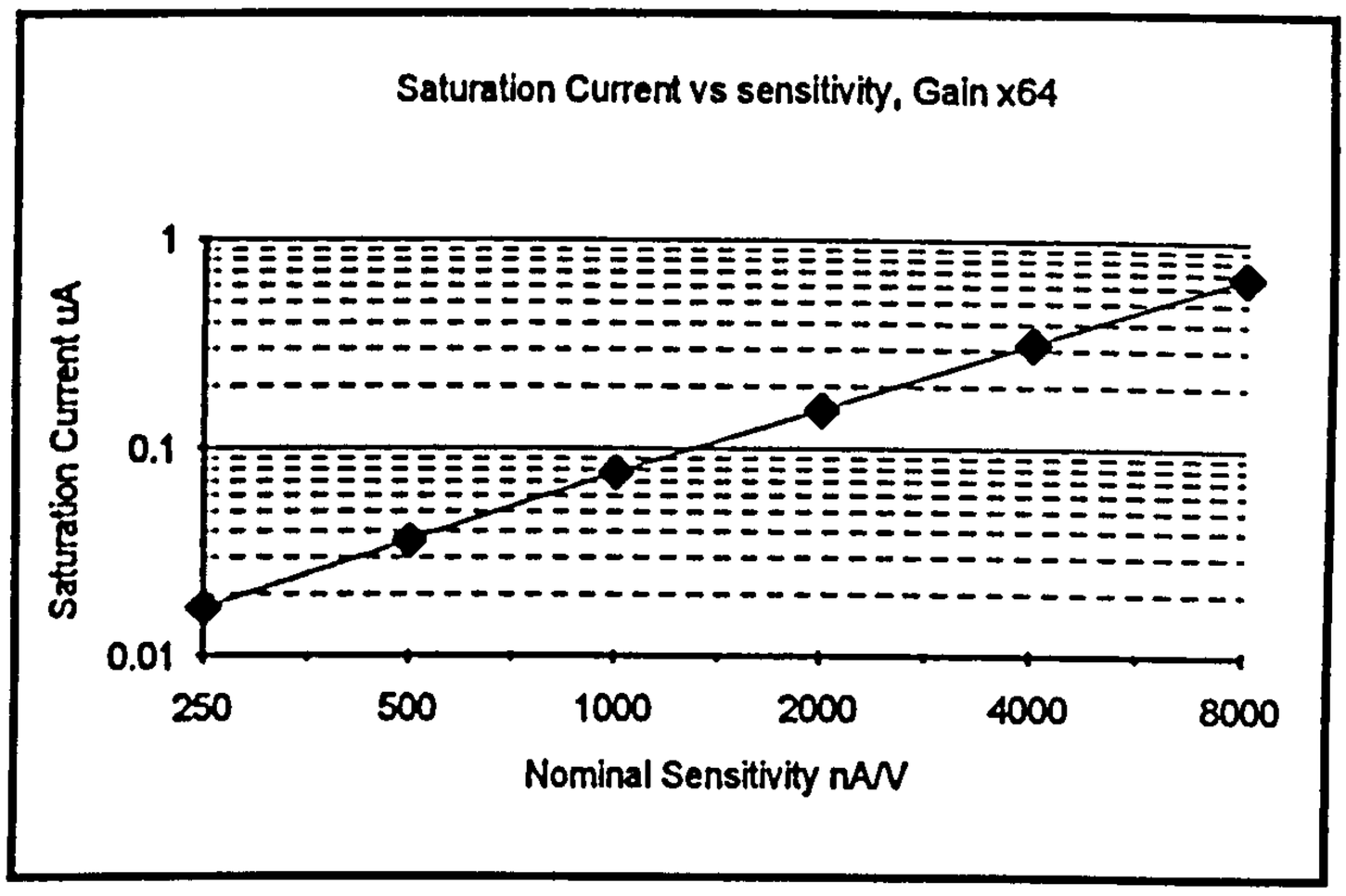



Figure 98. Saturation current - Gain x64

| Address Line | | Offset Range |
|---|---|---|
| 1 | 0 | |
| 0 | 1 | ±5μA |
| 1 | 0 | ±50A |

Table 42. I to E converter offset range settings

| Address Line | | | | Offset Number |
|---|---|---|---|---|
| 20 | 21 | 22 | 23 | |
| 0 | 0 | 0 | 0 | -8 |
| 0 | 0 | 0 | 1 | -7 |
| 0 | 0 | 1 | 0 | -6 |
| 0 | 0 | 1 | 1 | -5 |
| 0 | 1 | 0 | 0 | -4 |
| 0 | 1 | 0 | 1 | -3 |
| 0 | 1 | 1 | 0 | -2 |
| 0 | 1 | 1 | 1 | -1 |
| 1 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 1 | 1 |
| 1 | 0 | 1 | 0 | 2 |
| 1 | 0 | 1 | 1 | 3 |
| 1 | 1 | 0 | 0 | 4 |
| 1 | 1 | 0 | 1 | 5 |
| 1 | 1 | 1 | 0 | 6 |
| 1 | 1 | 1 | 1 | 7 |

Table 43. I to E converter offset settings

| | 5 μA Range (R = 1 MΩ) | | 50μA Range (R = 100 kΩ) | | Calculated Offset Current | |
|---|---|---|---|---|---|---|
| Offset Number | i/e offset sw pin 10 V | i/e Amp Pin 2 V | i/e offset sw pin 12 V | i/e Amp Pin 2 V | 5 μA Range μA | 50μA Range μA |
| -8 | -2.51 | -1.266 | -2.51 | -1.26 | -1.244 | -12.5 |
| -7 | -2.26 | -1.029 | -2.26 | -1.028 | -1.231 | -12.32 |
| -6 | -1.95 | -0.757 | -1.95 | -0.756 | -1.193 | -11.94 |
| -5 | -1.69 | -0.516 | -1.69 | -0.514 | -1.174 | -11.76 |
| -4 | -1.29 | -0.157 | -1.29 | -0.156 | -1.133 | -11.34 |
| -3 | -0.99 | -0.016 | -0.99 | -0.016 | -0.974 | -9.74 |
| -2 | -0.67 | -0.017 | -0.67 | -0.017 | -0.653 | -6.53 |
| -1 | -0.35 | -0.017 | -0.35 | -0.017 | -0.333 | -3.33 |
| 0 | -0.002 | -0.017 | -0.002 | -0.017 | 0.015 | 0.15 |
| 1 | 0.33 | -0.017 | 0.33 | -0.018 | 0.347 | 3.48 |
| 2 | 0.67 | -0.018 | 0.67 | -0.018 | 0.688 | 6.88 |
| 3 | 1.06 | 0.003 | 1.06 | 0.043 | 1.057 | 10.17 |
| 4 | 1.35 | 0.259 | 1.35 | 0.26 | 1.091 | 10.9 |
| 5 | 1.73 | 0.607 | 1.73 | 0.608 | 1.123 | 11.22 |
| 6 | 2.11 | 0.947 | 2.11 | 0.949 | 1.163 | 11.61 |
| 7 | 2.58 | 1.367 | 2.58 | 1.368 | 1.213 | 12.12 |

Table 44. Calibration of I/E current offset

# Appendix 6 Functions provided by the Analytical Software

| Analytical Procedures |
|---|
| procedure DoLinearScan(Vstart, Vend, ScanRate:real);<br><br>Performs a linear scan from *Vstart* (mV) to *VEnd* (mV) at a scan rate of *ScanRate* ($Vs^{-1}$)<br><br>The output PD (mV) and the input current (μA) are recorded at 1 mV intervals. |
| procedure DoSquareWave(VStart, VEnd, Scanrate:real; f, Ep : integer);<br><br>Performs a square wave scan from *Vstart* (mV) to *VEnd* (mV) at a scan rate of *ScanRate* ($Vs^{-1}$). The square wave frequency is *f* Hz and the pulse amplitude is *Ep* (mV).<br><br>The output PD (mV) and the input current (μA) are recorded for each pulse. |
| procedure DoDiffPulse(VStart,VEnd : real; PulseHeight, StepHeight : integer;<br>PulseDuration, InterPulseTime :word );<br><br>Performs a differential pulse scan from *Vstart* (mV) to *VEnd* (mV) rising by *Stepheight* (mV) each pulse. The pulse height is *PulseHeight* (mV). The duration of each pulse is *PulseDuration* (ms) and the rest time is *InterPulseTime* (ms).<br><br>The output PD (mV) and the difference between the input current before and at the end of the pulse (μA) are recorded for each pulse. |
| Procedure DoHgPlate(PD:real; time : word; Speed : integer);<br><br>Applies a potential of *PD* (mV) to the electrode and rotates it at *Speed* (rpm) for *time* (ms) |
| Procedure StripHg(Vstart, Vend, ScanRate:real);<br><br>Performs a linear scan from *Vstart* (mV) to *VEnd* (mV) at a scan rate of *ScanRate* ($Vs^{-1}$)<br><br>No data is recorded |
| Procedure DoAnalytePlate(PD:real; time : word; Speed : integer; PostDelay :<br>word;measuring : boolean);<br><br>Applies a potential of *PD* (mV) to the electrode and rotates it at *Speed* (rpm) for *time* (ms). Then turns off rotation and maintains the potential for a further *PostDelay* (ms).<br><br>If *measuring* is true, PD and current data are recorded every ms and may be displayed as a plot through the display menu. |

Table 45. The analytical procedures available in the logical analysis layer of the electrochemical analysis software

Table 46, Page 288 specifies the function required to initialise the PIO. This sends the byte $8A_{(hex)}$ to the control port of port 1 in order to set all 8 bits of port A plus the lower 4 bits of port C to outputs and all 8 bits of port B plus the upper 4 bits of port B to inputs. These two sets of 12 bits are used to output the digitised desired working electrode potential and to input the digitised potential representation of the working electrode current (differential or direct) respectfully.

| PIO Board configuration Procedure |
| --- |
| procedure SetUpIOPorts<br>Sets up the mode of operation of the ports and configures ports for input or output<br>Sets bits;<br>Control mask to set B & C Hi in and A & C lo Out = $8A<br>Control mask to set up ports A'B'C' all to output = $80 |

Table 46. Low level control functions (PIO board)

Table 47, Page 289 lists the various functions that control the operating parameters of the potentiostat. These are;

- The Current to Voltage Converter Sensitivity
- The Current To Voltage Offset Range
- The Current To Voltage Offset Value
- The Differential Amplifier Gain

The current to voltage converter can be set to one of six conversion sensitivities from 8000 $nAV^{-1}$ to 250 $nAV^{-1}$ in order to allow for high analyte

concentrations to produce large currents without the converter clipping the converted PD yet still allow for small currents to be converted to measurable PD's

| **Potentiostat Configuration Procedures** |
|---|
| procedure SetItoEOffsetRange(OffsetRange:ItoEOffsetRangeType)<br>Sets the I to E converter offset range to either 5μA or 50μA<br>mask used to clear offset range bits $FC<br>Sets bits; 5 μAmp = $01, 50μA = $02 |
| procedure SetItoEOffset(Offset:ItoEOffsetType)<br>Sets the I to E converter offset to one of sixteen values spanning zero<br>mask used to clear I/E offset bits $0F<br>Sets bits;<br>I/E Offsets (-8 to +7) -8 = $00, - 7 = $10, - 6 = $20, - 5 = $30 ... +7 = $F0 |
| procedure SetAmpGain(gain:AmpGainType)<br>Sets the differential amplifier gain to one of four gains<br>mask used to clear differential amplifier gain bits $FC<br>Sets bits;<br>Delta Amp Gain x1 = $00, x4 = $01, x16 = $02, x64=$03 |
| procedure SetItoESens(NanoAmpPerVolt:ItoESensType)<br>Sets the I to E converter sensitivity to one of six sensitivities<br>mask used to clear I/E sensitivity bits $03<br>Sets bits;<br>I/E conversion range (nA/v) 8000 = $04, 4000 = $08, 2000 = $10, 1000 = $20, 500 = $40, 250 = $80 |

Table 47. Low level control functions (potentiostat control)

| Ancillary switches |
| --- |
| procedure TurnOnMotor (Speed:integer)<br>Turns on the working electrode rotating motor<br>Sets bit $02 |
| procedure TurnOffMotor<br>Turns off the working electrode rotating motor<br>Resets bit $02 |
| procedure TurnOnGasPurge<br>Opens the Nitrogen gas purge valve<br>Sets bit $01 |
| procedure TurnOffGasPurge<br>Closes the Nitrogen gas purge valve<br>Resets bit $01 |
| procedure ConnectDeltaAmpInput2<br>Connects the inverting input (pin 19) of the differential amplifier to the working electrode via Sample/Hold 2<br>Resets bit $04 |
| procedure GroundDeltaAmpInput2<br>Grounds the inverting input (pin 19) of the differential amplifier (for use in linear scan)<br>Sets bit $04 |
| procedure ConnectWorkingElectrode<br>Connects the working electrode to the I to E converter (normally done before a scan)<br>Sets bit $08 |
| procedure GroundWorkingElectrode<br>Grounds the working electrode (normally done before analyte plating)<br>Resets bit $08 |

Table 48. Low level control functions (ancillary)

Table 49, Page 291 lists the functions that control the application of a particular potential to the working electrode, the subsequent reading of the

converted differential cell current and the control of the two sample and holds on the inputs of the differential amplifier.

| Working Electrode Potential and Current Reading procedures |
|---|
| procedure SetSampleHold(number : integer; mode : SampleHoldType)<br><br>Sets sample/hold on either input one or two of the differential amplifier to either sample or hold<br><br>Affects bits;<br><br>Sample/Hold 1 $10 set = sample, reset = hold<br><br>Sample/Hold 2 $20 set = sample, reset = hold |
| procedure SendVoltage (Vout : integer)<br><br>Converts a bipolar PD in the range ±2.5 V to the corresponding 12 bit D/Acode and sends it out to the potentiostat |
| function VoltageIn:integer<br><br>Reads a 12 bit code from the A/D in the potentiostat and returns the value converted into a bipolar PD in the range ±2.5 V |

Table 49. Low level control functions (W.E.)

| |
|---|
| Procedure DisplayData(var XDataArray, YDataArray : AnalysisArrayType; NumDataPoints : integer; StayOnScreen : boolean);<br><br>Plots the *NumDataPoints* of data in *XdataArray* and *YdataArray* as an XY graph on auto scaled axes. If *StayOnScreen* is true the user may manipulate the plot with various commands and must exit the graphical display by entering 'x' to exit graphics. If false this facility is disabled and graphics is terminated by a call to ShutDownGraphics |
| Procedure SetTitleOne(Sensitivity : ItoESensType; OffsetRange : ItoEOffsetRangeType; Offset : ItoEOffsetType; CurrentGain : AmpGainType; method : string);<br><br>Converts the enumerated types of *Sensitivity, OffsetRange, Offset, CurrentGain* and *Method* into a single string suitable for the first title line of an XY plot. Also initialises conversion factors used by DisplayData |
| Procedure SetTitleTwo(TitleTwo : string);<br><br>Initialises the second line of the XY plot title to the string *TitleTwo.* |
| Procedure SetTitleThree(TitleThree : string);<br><br>Initialises the third line of the XY plot title to the string *TitleThree.* |

Table 50. Functions available in the logical graphics software layer

| Graphics Functions |
|---|
| Procedure InitGraphics( var ok : integer);<br><br>Starts the graphics system in the best available mode, returns zero or an error code |
| Procedure ShutDownGraphics;<br><br>Shuts down the graphics system and returns to the text screen. |
| Function GraphicsIntiialised : boolean;<br><br>Returns true if the graphics system is already started, false otherwise |
| Procedure PlotData(FileToUse: string; DisplayMethod:DispMethodT; ColourToUse:word; StayOnScreen : boolean);<br><br>Plots the data on file *FileToUse*, in colour *ColourToUse* using the *DispMethodT* of display (see text). If *StayOnScreen* is true the user may manipulate the plot with various commands and must exit the graphical display by entering 'x' to exit graphics. If false this facility is disabled and graphics is terminated by a call to ShutDownGraphics |

| |
|---|
| Procedure DrawAxis(Xtitle, Ytitle, GraphTitle1, GraphTitle2, GraphTitle3 : string);<br><br>Plots the X and Y axis and gives them legends, also prints a three line graph title |
| Procedure DoTickMarksOnYAxis(YMin,Ymax: real; numticks, interval : integer);<br><br>Draws *numticks* tick marks spaced *interval* apart and prints the scale going from *Ymin* to *Ymax*. Also calculates and initialises the plot to real world Y scaling value |
| Procedure DoTickMarksOnXAxis(Xmin, Xmax : real; numticks, interval : integer);<br><br>Draws *numticks* tick marks spaced *interval* apart and prints the scale going from *Xmin* to *Xmax*. Also calculates and initialises the plot to real world X scaling value |
| Procedure DrawCurveValue(X, Y : real; ColourToUse : word);<br><br>Plots a *ColourToUse* coloured line to real world co-ordinates *X, Y* from last real world co-ordinates plotted or plots a point at *X,Y* if this is the first call to this procedure. |
| Procedure PlotValue(X, Y : real; colourtouse : word);<br><br>Plots a *ColourToUse* pixel at real world co-ordinates *X, Y* |
| Procedure ConvertPlotXYToValueXY(var ValX, ValY : real; PlotX, PlotY : integer);<br><br>Converts a point given by the screen pixel co-ordinates *PlotX, PlotY* to real world co-ordinates *ValX, ValY*. |
| Procedure ConvertValueXYToPlot(ValX,ValY : real; var PlotX,PlotY :integer;var OffScale:boolean);<br><br>Converts a point given by real world co-ordinates *ValX, ValY* to screen pixel co-ordinates *PlotX, PlotY*. Offscale is true if *PlotX* or *PlotY* is out of range |
| Procedure FindMaxFileValues(infile : string; var MinXValue, MaxXvalue, MinYValue, MaxYValue : real);<br><br>Finds the maximum and minimum values of X and Y in the plot file *infile* |
| Procedure ScreenDump;<br><br>Prints screen on the printer |
| Procedure MakeDisplayColour;<br><br>Restores default graph colouring |
| Procedure MakeDisplayBlackAndWhite;<br><br>Changes display to black on white to facilitate the transfer of the graph into another application via Windows® Clipboard |
| Procedure MovingAverage;<br><br>Replaces graph with a five point moving average |

| |
|---|
| Procedure ShowFinalPlotVals(ThisEpoch : integer;TotalSumSquares,CriticalError : real);<br><br>Displays values of *ThisEpoch*, *TotalSumSquares* and *CriticalError* and time if cessation of training (for use with a Neural Net learning curve plot) |
| Procedure ShowCurrentError(TSS:real);<br><br>Prints current value of *TSS* (total sum squared error) in the bottom right hand corner of the screen (for use with a Neural Net learning curve plot) |

Table 51. Table of available functions in the low level graphics module