ELSEVIER

Bioelectrochemistry and Bioenergetics 40 (1996) 99–132

BIOELECTROCHEMISTRY AND BIOENERGETICS

Rapid and non-invasive quantification of metabolic substrates in biological cell suspensions using non-linear dielectric spectroscopy with multivariate calibration and artificial neural networks. Principles and applications

Andrew M. Woodward [a], Alun Jones [a,b], Xin-zhu Zhang [a], Jem Rowland [b], Douglas B. Kell [a,*]

[a] *Institute of Biological Sciences, Edward Llwyd Building, University of Wales, Aberystwyth, Dyfed SY23 3DA, UK*
[b] *Department of Computer Sciences, University of Wales, Aberystwyth, Dyfed SY23 3DD, UK*

Received 10 November 1995; revised 2 February 1996

Abstract

By studying the non-linear effects of membranous enzymes on an applied oscillating electromagnetic field, non-linear dielectric spectroscopy has previously been shown to produce qualitative information which is indicative of the metabolic state of a variety of organisms. In this study, we extend the method of non-linear dielectric spectroscopy to the production of data sets suitable for use with supervised multivariate analysis methods, in order to allow quantitative prediction of analyte concentrations in unknown samples, again using the alteration in the non-linear dielectric profile produced by these analytes via the metabolism of the cell (as effected via the operation of their membranous enzymes).

Non-stationarity in the extent of non-linear electrode polarization can interfere with the measurement of non-linear dielectric spectra; various electrode materials and configurations have been tested for their suitability for use in non-linear dielectric spectroscopy.

We exploit partial least-squares regression and artificial neural networks for the multivariate analysis of non-linear dielectric data recorded from yeast cell suspensions, and schemes for preprocessing these data to improve the precision of the prediction of analyte levels are developed and optimized. The resulting analytical methods are applied to the prediction of glucose levels in sheep and human blood, by both invasive and non-invasive measurements, and to the non-invasive measurement of process variables during a microbial fermentation.

Keywords: Biotechnology; Dielectric spectroscopy; Multivariate calibration; Neural networks; Non-invasive; Non-linear

1. Introduction to non-linear dielectric spectroscopy

When a suspension of cells is exposed to a static electric field, or to an alternating electric field whose frequency is low relative to that of the classical β-dielectric dispersion, it does not penetrate to the interior of the cell, and is dropped almost entirely across the plasma membrane of the cell, which is predominantly capacitive at these frequencies, and, due to its thinness, causes a substantial amplification of the field across the membrane [1,2]. In consequence, anything internal to the cell is essentially electrically invisible to a low-frequency electric field, but anything dielectrically active in the membrane may be expected to display properties associated with fields far stronger than that applied externally.

The dielectric response of biological tissue has long been assumed to be linear when the macroscopic exciting field is low, e.g. less than 0.1 V cm^{-1} as used typically, and such linear dielectric properties (see Fig. 1) have been reviewed by several workers [4–16]. Under these circumstances, an enzyme is typically treated as a hard sphere which relaxes linearly in an a.c. field at all but high field strengths [17]. However, substantial non-linear phenomena in the form of harmonics of the fundamental as measured across the membrane of nerve cells [18,19] and across black lipid membranes [20,21] have in fact been reported.

An enzyme which has different dipole moments in different conformations during its operation may be af-

* Corresponding author. Tel.: +44 1970 622334; fax: +44 1970 622354; e-mail: azw, auj, xxz, jjr, dbk@aber.ac.uk.

PII S0302-4598(96)05065-9

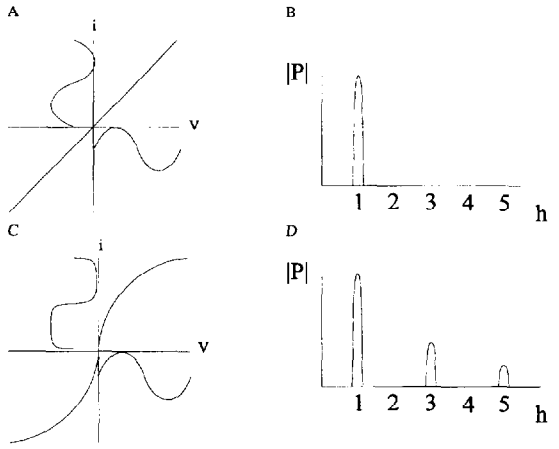

Fig. 1. Relationship between i/V curves and power spectra. (A) A linear i/V curve: a sinusoidal voltage produces a sinusoidal current at the same frequency. (B) Its power spectrum which contains only the fundamental (first harmonic). (C) A non-linear i/V curve: produces a distorted, imperfectly sinusoidal current. (D) Its power spectrum which also contains harmonics. These diagrams are intended to be merely illustrative; factors causing the production of odd-numbered vs. even-numbered harmonics may be found in a textbook on non-linear dynamics, e.g. [3].

fected by electromagnetic fields [22–24]. The change between states is unlikely to be smoothly or linearly related to the field due to the constraints imposed on the enzyme by its environment in the membrane, and it can be shown theoretically that the dielectric response of the material may be expected to be non-linear even at low applied fields [25–38]. Thus any perturbation of an applied field would be expected to be non-linear and, as with any (in this case, weakly) non-linear system, this behaviour would be expected to be manifested as the generation of harmonics of the applied frequency by the enzyme when the excitation is a single sinusoid [14,16,25,26]. More importantly, any asymmetry in the dielectric potential of the enzyme will lead to a rectification of field effects and the ability of the enzyme to harvest energy from the applied field, leading to the possibility of using the dielectric properties of the membranous enzymes to indicate and/or even to influence the metabolic state of cell suspensions.

In earlier work, we have demonstrated the presence, and investigated the properties, of these predicted harmonics as generated by a variety of cell suspensions, using a non-linear dielectric spectrometer designed around a standard IBM-compatible PC and realized almost completely in software with a minimum of extraneous hardware [39–44]. In the work described, we noted substantial changes, both qualitative and quantitative, in the non-linear dielectric spectra observed when resting cells were allowed to metabolize actively following the addition of appropriate substrates. Inhibitor and other studies indicated that, in yeast, the signal was due mainly to the H^+-ATPase located in the cells' plasma membrane [39,41], whereas, in a variety of prokaryotes, the main contributors to the generation of non-linear dielectricity were the electron transport chains located in the cytoplasmic membranes of these organisms [42,44].

Since this work was published, we have developed our system significantly, and have devised methods which allow the rapid, essentially non-invasive, quantitative analysis of metabolite levels by combining non-linear dielectric spectroscopy with modern chemometric methods of multivariate calibration and artificial neural networks (ANNs). Our purpose herein is to review this progress.

This paper is organized as follows. We first describe the modifications to our own implementation of a non-linear dielectric spectrometer as described previously [39–44], and rehearse the standard types of data which may be acquired. We then summarize our approaches to minimizing electrode polarization artefacts. We then compare the use of multivariate statistics with ANNs for the prediction of glucose concentrations in yeast cell suspensions metabolizing glucose in vitro and in fermenters in situ. Lastly, the approach is exploited for cognate measurements in blood cell suspensions from sheep and humans, and for non-invasive glucose measurements in vivo.

1.1. The non-linear dielectric spectrometer

The experimental hardware used for the following development is identical to that described previously [39–44].

Early on in the development work, but previously unpublished, we also carried out two essential control experiments [45].

The first of these (in 1990) was to investigate whether the harmonics were produced in the bulk of the suspension or by an interaction with the polarization layer. The suspension was isolated in a volume around the sensor electrodes with a dialysis membrane, the volume between this and the driver electrodes being filled with conductivity-matched supernatant as indicated in Fig. 2. Taking spectra with this arrangement, then substituting the supernatant for the suspension, produced essentially identical third harmonic signatures from a yeast suspension to those obtained when the dialysis membranes were removed, proving a bulk effect. An electrode surface effect would have given a null result.

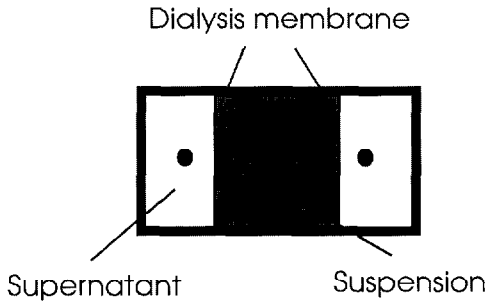

Fig. 2. Electrode chamber for control experiment to exclude direct electrode interactions with suspension.

The second (in 1991) was to investigate the use of a current source (as later used by others [46,47]) to provide the signal, instead of a voltage source as usually used. The rationale behind this was that the polarization harmonics might be removed by the feedback on the output of the current source, leaving those due to the biology, eliminating the need for a reference reading in the supernatant. However, we found that the biological harmonics were also sensed across the driver electrodes in a manner indistinguishable from the polarization harmonics and were also removed by the output feedback of the current source, leaving only a pure sinusoidal signal through the electrode chamber. The effect of the current source was to force an "apparent" linearity on the system. Therefore all subsequent work used a voltage source and alternative methods of dealing with the polarization problem.

Blake-Colman and coworkers [46,47] published a critique of some of our earlier work. However, this critique was based on a variety of misconceptions, and omitted to discuss our control experiments. Readers are therefore referred to our rebuttal [45] and to our original papers [39–44].

1.1.1. Recent developments to the spectrometer

In our published work to date, the data logging and spectral analysis programs were based on 256-point Fourier transforms, giving 128-point power spectra. A schematic diagram of the software used to record and analyse these standard spectrum sweeps is shown in Fig. 3.

Since we noticed over this period that the only phenomena that were reliably present in the spectra produced by the organisms investigated were harmonics, the calculation of interharmonic bins in the spectra merely wasted time and required the recording of unnecessarily long data series, resulting in unnecessarily long experimental times. Recording shorter data series allows wider frequency/voltage scans to be carried out in less time, and reduces the problems of time variation in electrode polarization (see later) during an experiment or series of experiments. Since only the harmonics appear to contain interesting information in the systems studied, the ideal would be to record only sufficient data to represent these, and to calculate a shorter Fourier transform in which each successive bin represents each successive harmonic.

If each block of data in a series is sampled at such a rate that it contains a whole number of cycles of the fundamental, the fundamental and its harmonics will fall exactly into spectral bins (e.g. 15 cycles per block places harmonics in every 15th bin as in Fig. 4(A)). In this case, the need for continuity at the block boundaries can be met with no windowing. Also the boxcar windowing implicit in the finite length of the time series and the infinite repeat of

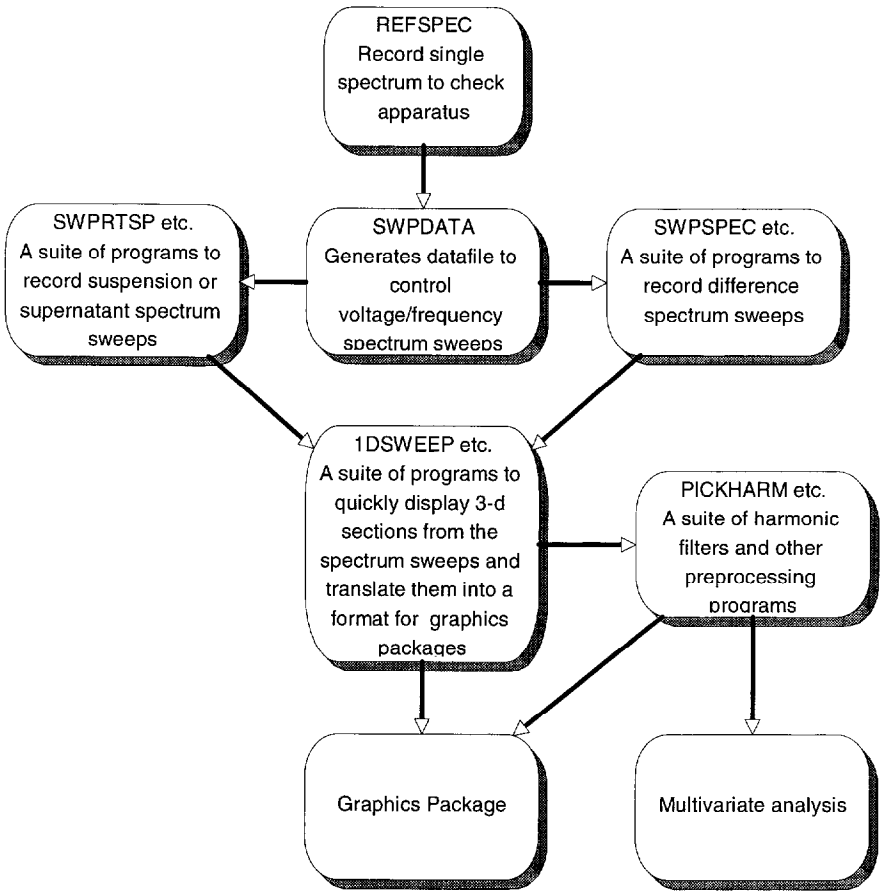

Fig. 3. Schematic diagram of spectrum sweep software.

this implicit in the Fourier transform are exactly representative of the true signal extended to infinity with no distortion, the only assumption being stationarity. Hence spectral leakage of any bin into adjacent bins is essentially eliminated (Fig. 4(B)).

If the sampling is chosen to include a single cycle in each block, each successive bin in the transform will contain each successive harmonic of the signal, uncontaminated by leakage from adjacent harmonics. So, if the first eight harmonics are of interest, then a 16-point fast Fourier transform can be applied to data containing one cycle per block, sampled at 16 samples per cycle (Fig. 4(D)). In reality, we would not wish to approach the Nyquist criterion this closely and would discard harmonics above, e.g. the fifth. It should be noted that the harmonics are beginning to be averaged out to slightly lower values, probably due to the increased noise bandwidth per bin, but that their relative sizes are preserved.

The importance of minimizing spectral leakage in single bin/harmonic spectra can be seen from Fig. 4(C), which is the same as Fig. 4(D) except that a minimum three term Blackman–Harris window [48] is applied to the data, smearing adjacent harmonics into each other. This spectrum can be seen to bear only a passing resemblance to the true windowed spectrum of Fig. 4(A). Overall, the above approach allows a speed up in data recording of an order of magnitude, although the great speed up in processing time resulting from the shorter Fourier transforms that now need to be calculated is essentially negligible on modern computers in comparison with the (physically determined) time taken to record the data.

The slightly modified schematic diagram, equivalent to that of Fig. 3, of the software used to perform these harmonic-per-bin sweeps is shown in Fig. 5.

Two-terminal networks have also been studied, and found to give essentially similar results, although they are much more sensitive to changes in the electrode polarization layer, and so tend to give noisier spectra.

1.2. Non-linear dielectric properties of yeast cell suspensions

The standard suspension of *Saccharomyces cerevisiae* was prepared as follows. Freeze-dried yeast (Allinson's baking yeast, obtained locally) was rehydrated to 50 mg dry wt. ml^{-1} in a solution of 1% yeast extract (w/v) in distilled water. This was allowed to stabilize for 2 h before experiments were carried out. This method was used to prepare all yeast suspensions used in this paper unless explicitly stated otherwise. No cell growth occurred under the conditions used.

In this suspension of *Saccharomyces cerevisiae*, an inhibitor study, together with the use of mutant strains, showed that the predominant source of the non-linear

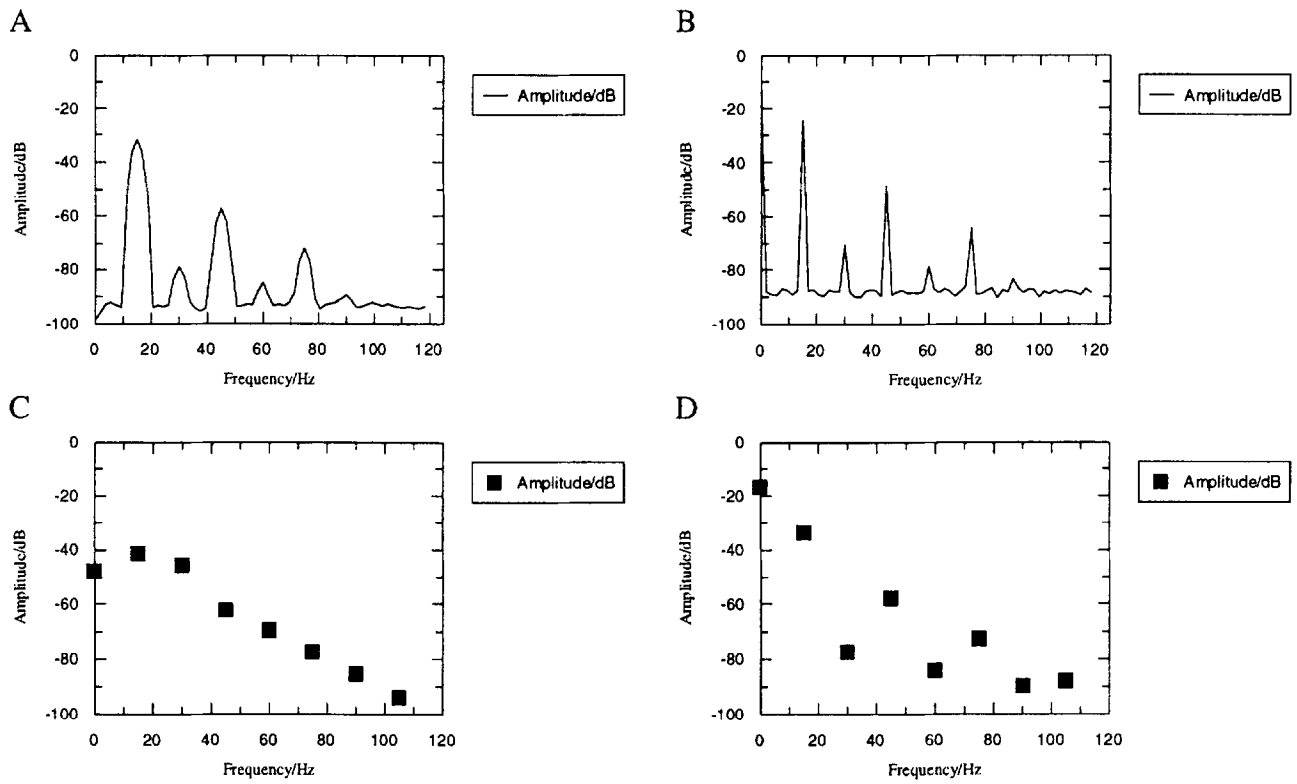

Fig. 4. Development of single-harmonic-per-bin data logging and processing: (A) 256-point Fourier transform with 15 cycles per block; (B) as (A) with no windowing to show the elimination of leakage; (C) 16-point Fourier transform with one cycle per block; leakage between harmonics renders the spectrum useless; (D) as (C) with no windowing restoring the integrity of the spectrum.

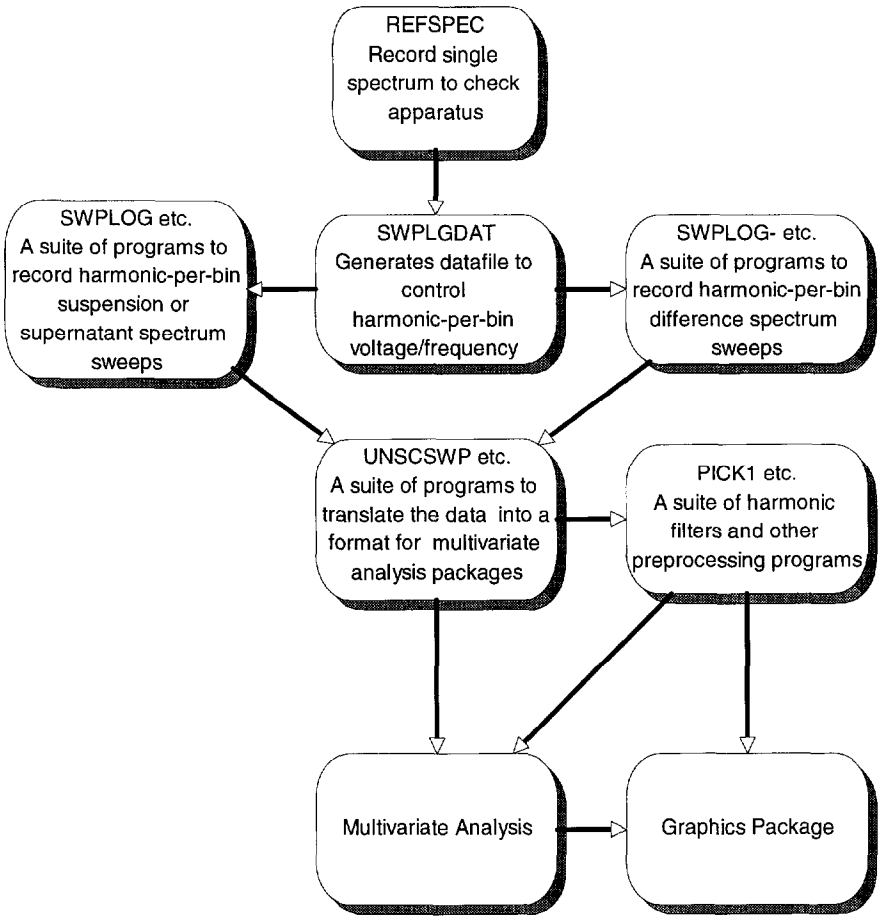

Fig. 5. Schematic diagram of modified software to allow the recording of rapid harmonic-per-bin spectrum sweeps.

signature in this organism is the membrane-located H^+-ATPase [39,41]. The harmonics are highly voltage and frequency windowed, with the peak of the frequency window for the resting enzyme coinciding rather neatly with its k_{cat} value [40]. On cooling the yeast suspension to 2 °C, the peak of the frequency window moved to 3 Hz, the voltage window not being significantly affected (data not shown).

From an energetic point of view, a generalized enzyme may be constrained to a potential surface which may typically look something like that in Fig. 6. In a resting state, at equilibrium, the yeast suspensions studied generated almost entirely odd-numbered harmonics [39], suggesting symmetry about the equilibrium of the ATPase, as discussed in Fig. 7(A).

If glucose is added to the suspension to fuel proton transport by the ATPase, the shift away from equilibrium breaks the symmetry and even-numbered harmonics appear [39], as predicted by the analysis of Fig. 7(B), giving a measure of the activity or inactivity of this enzyme and the consequent metabolic state of the yeast cells [39].

The general situation, in various strains of yeast and in other organisms, is that the resting suspension will generate predominantly odd harmonics with some evens present, dependent on the exact shape of the potential well in which the enzyme sits. In the presence of metabolizable substrate (e.g. glucose for yeast), there is a less extreme shift in the balance of odd and even harmonics, favouring

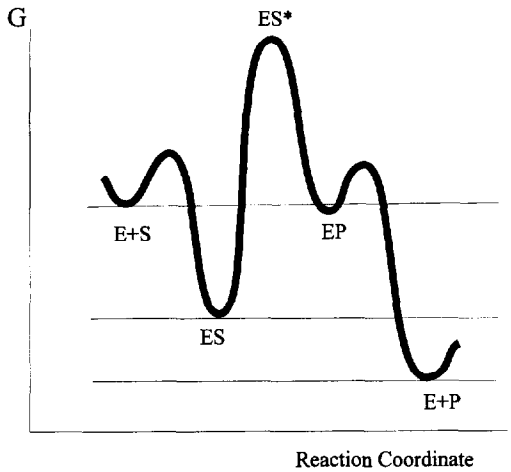

Fig. 6. A typical energy level diagram of an enzyme. The enzyme E binds to the reactant S. An "excited" intermediate ES* is formed which produces the bound product EP. This then dissociates to release the product P.

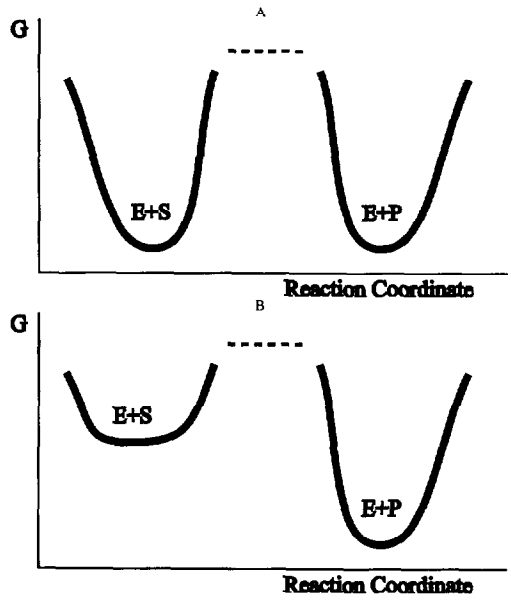

Fig. 7. (A) Enzyme at equilibrium. The potential energy function describing the potential energy well of E + S and E + P is symmetrical, containing mainly even terms. Therefore the equation of motion formed by differentiating this potential and describing the reaction with respect to the reaction coordinate produces mainly odd harmonics in response to sinusoidal excitation. (B) Enzyme not at equilibrium. The potential function is skewed and now requires substantial odd terms to describe the asymmetry. Hence the equation of motion now also produces substantial even harmonics.

the evens and tending to reduce the odds as the potential becomes more asymmetric during the operation of the enzyme. Also there is no fundamental reason to believe that the voltage/frequency windows of harmonics due to any metabolism-related effects will be the same as those for quiescent enzymes, particularly if the enzyme is chemically altered, with a consequent alteration of the rate constants of the equilibria between the conformational states during metabolism. The exact form of these responses is also found to be highly strain dependent within a species, frequency and voltage windows varying between strains.

Analysing the behaviour of the harmonics when the fundamental is varied over a range of frequencies/voltages allows the rapid collection of a very large amount of metabolism-dependent information, since in the present case we obtain data for the magnitude of each of a number of harmonics (usually five) generated when the system is excited at a number of voltages and frequencies. Since such data are extensive, and at least four-dimensional, it is necessary to exploit analytical tools capable of dealing with and visualizing such a volume of data.

Before we turn our attention to this, we shall address the practical experimental problem of electrode polarization instability.

2. Non-linear electrode polarization

The correct registration of the dielectric properties of biological systems would be easy were it not for electrode polarization, and a number of reviews of the unwanted contribution of electrode polarization to conventional (linear) dielectric measurements are available [4,9,49,50]. Polarization describes the fact that, when a current is forced to cross an electrode–electrolyte interface, there will be a resistance to such current flow and thus a tendency for a potential drop to occur across that interface. If the current is alternating, there can also be a phase shift between the current and the voltage, which therefore serves to contribute to the capacitance measured.

Apart from special techniques such as second harmonic a.c. voltammetry [51], a.c. measurements are usually taken assuming a linear response, with the current and voltage of the received signal measured only at the generator frequency. This is a zeroth order approximation to the reality that the current–voltage relationship of the electrode–electrolyte interface is non-linear, and an applied a.c. signal will necessarily generate harmonics of its frequency content [52,53], together with other phenomena, such as frequency mixing, pseudoperiodicity and chaos, classically associated with non-linear systems.

Since, in non-linear dielectric measurements, we are particularly looking for biologically generated harmonics, those produced simply by current flow across the electrode interface are a severe embarrassment. In order to obtain sensible results, the polarization harmonics must be preferably negligibly small or, failing that, stable. Unfortunately, neither of these is normally the case.

Electrode polarization is predominantly due to the requirement of a change of charge carriers from electrons in the electrode to ions in the solution and, consequently, is a fundamental property of metal–solution interfaces [54]. However, because of the non-linearities, the processes involved in polarization are much more complex than can be dealt with by simple equivalent circuit models, such as the standard Randles circuit used for linear models [55] and, anyway, it is not possible to deduce unambiguous mechanistic models from impedance measurements, only vice versa. The assumptions of most models to date, of reversibility at the electrode and reversion to its original state after the signal is removed, are in fact false [53].

The resistive Faradaic charge transfer contribution to the polarization impedance begins to become non-linear at small voltages and occurs predominantly at low frequency. The reactive part is linear to higher signal strength, happens first at high frequency and consequently can be neglected for present purposes [56].

Interfacial roughness and contamination can affect polarization, as can monocrystallinity or polycrystallinity [54]. Electrode polarization is strongly and mysteriously dependent on the exact surface preparation, and may de-

pend on the exact fractal nature of the electrode surface [52].

The polarization layer is also far from stable [57]. Spontaneous noise of up to 1 mV occurs electrochemically at electrodes. Its frequency ranges from d.c. to over 100 Hz, i.e. exactly the range of interest in non-linear dielectric spectroscopy. Spikes and semiperiodic activity are also generated and these can be larger still. Electrodes can display any of these modes, or jump from one to another. These fluctuations appear to be fundamental in electrochemical systems, and have been observed in stainless steel, copper, brass, aluminium, monel, solder, carbon, mercury, degenerating Ag/AgCl, silver and platinum. Stainless steel is mentioned as being particularly quiet [57]. The noise is reduced by a larger surface area of electrode, and also by a more reversible charge transfer process.

Fluctuations can also occur due to electromechanical activity if the apparatus is disturbed [57].

2.1. Effect of polarization on non-linear dielectric spectroscopy

In non-linear biological work, electrode polarization is a serious problem. It occurs most strongly at low frequencies (up to a few tens of Hertz) where the biology typically reacts most strongly to the electric field, and its fluctuations can be similar in size to or larger than the small changes due to biological activity (e.g. on glucose metabolism). It is therefore vital to control electrode polarization as far as possible. Since harmonics fluctuate proportionally more than the fundamental, non-linear measurements, which concentrate on these harmonics, are much more sensitive to polarization phenomena than are linear measurements, which concentrate exclusively on the relatively stable fundamental.

To obtain non-linear electrochemical reproducibility, we have found that electrode surfaces must be scrupu-

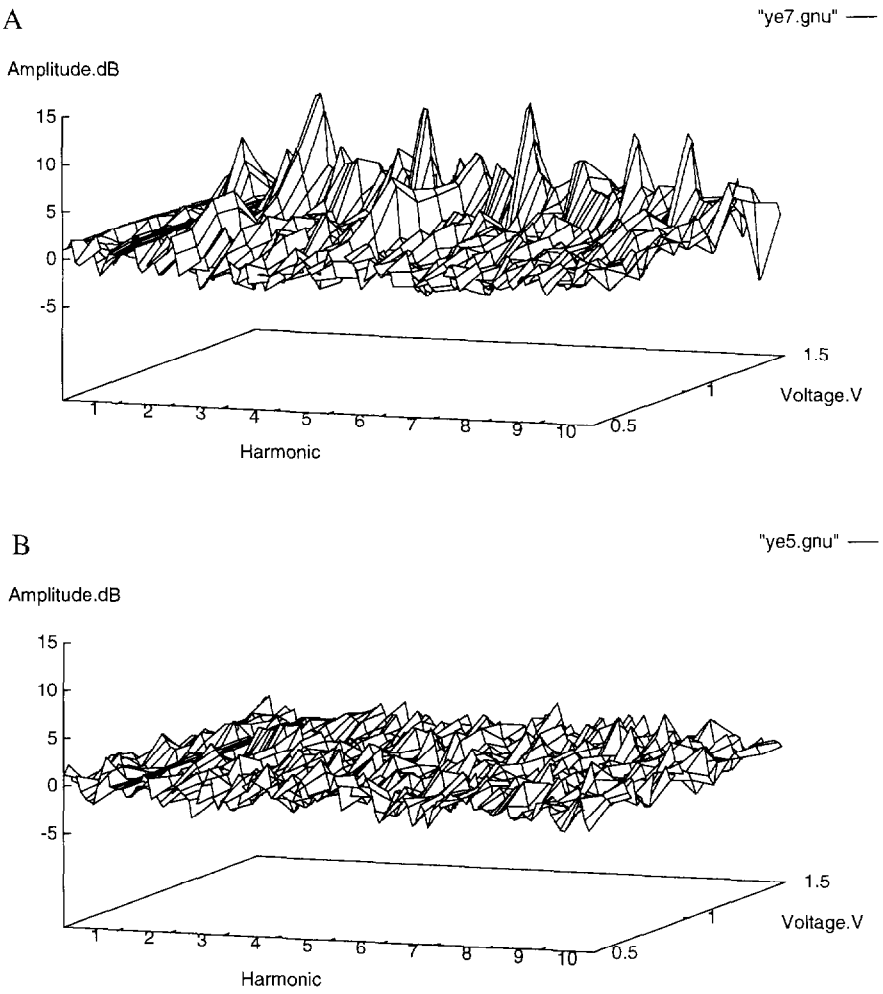

Fig. 8. (A) Supernatant minus supernatant for "dirty" gold electrodes in 100 mM KCl. This is a subtraction of two successive spectral sweeps, with an input sinusoid applied to the outer pins of a four-terminal electrode at 15 Hz with the voltage amplitude (zero to peak) swept between 0.5 and 1.5 V, and should be null. A data spike leading to a raised background in a spectral slice near the rear of the plot can be seen. Also spurious harmonics are common, showing that the polarization is changing between the two spectral sweeps. The time between sweeps was 3 min. (B) As (A), but for "clean" electrodes. The spurious harmonics and data spikes have disappeared, leaving only the innate electrochemical noise at the electrodes. The electrodes are now stable to the limit of this noise.

lously clean, and this is very difficult to achieve. If any contamination is present, the biologically relevant signal may be unstable, distorted or destroyed completely. Any results obtained will also be very dependent on the signal history! Only when electrodes are spotless will reasonably low-noise, repeatable results be obtained, although fluctuations will still occur. We find that these fluctuations occur on all timescales from slow drift over days, altering the effective characteristics of the sensing system between experiments, to second-to-second variations in polarization harmonics, which affect the stability of the sensing system during a single reading.

Fig. 8 shows a typical difference in stability between contaminated and clean electrodes.

Surfaces are very liable to contamination by many substances (proteins, sugars, lipids, etc.) which can bind to the electrode surface and distort results [58,59], proteins being particularly likely to cause trouble. Therefore the electrodes must be very carefully checked for surface interactions with any new chemical introduced into the test chamber before the results obtained can be trusted.

After cleaning, the stabilization time during which the electrode must be left (preferably under the working signal) in a working electrolyte can be up to days, but is

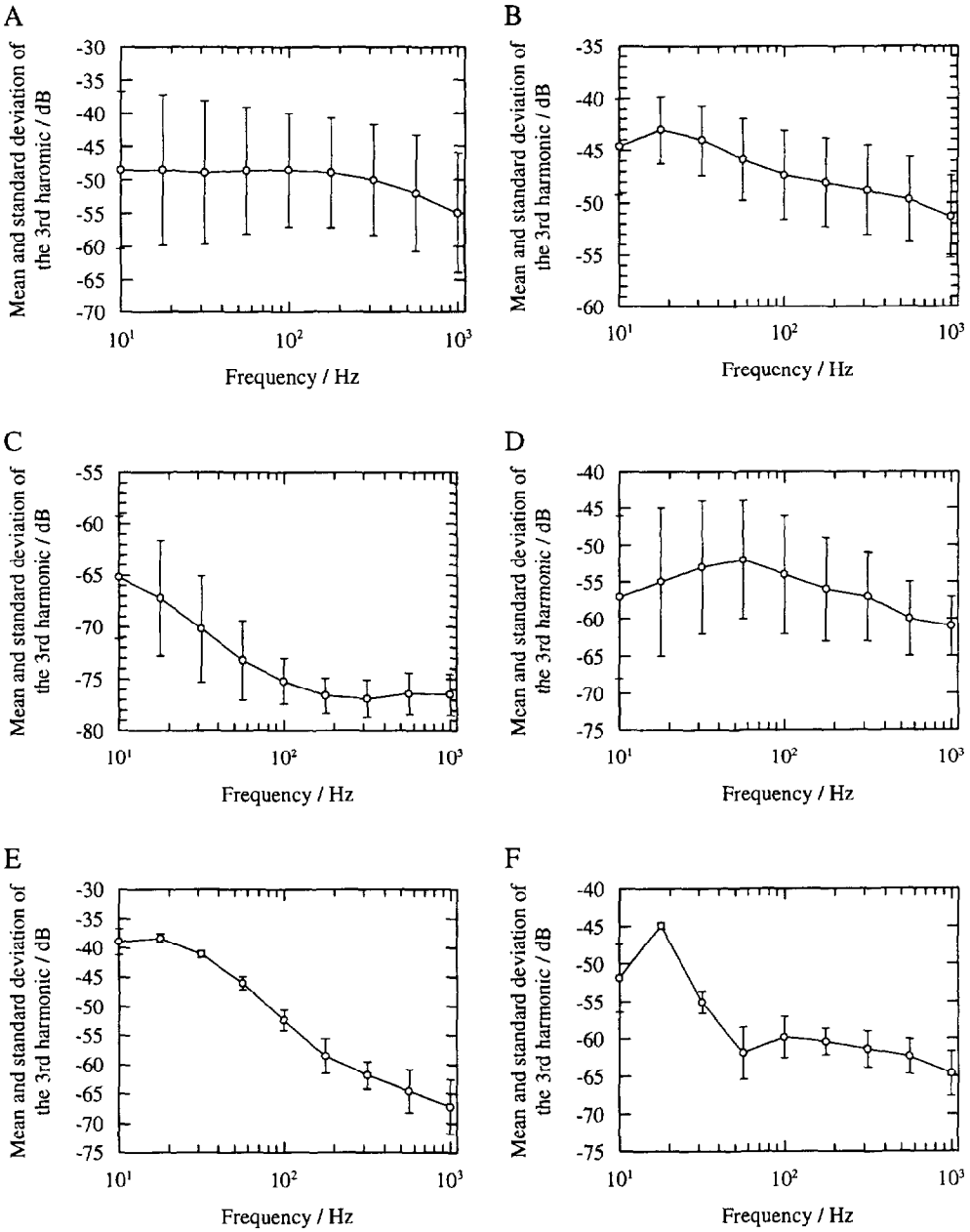

Fig. 9. Fluctuation of the third harmonic due to electrode polarization over 5 days: (A) tin electrodes; (B) platinum electrodes; (C) stainless steel electrodes; (D) copper electrodes; (E) gold electrodes; (F) Ag/AgCl disposable electrodes (with one set used for each reading). The applied signal was a swept sinusoid with a voltage of 1 V (zero to peak) and a frequency range from 10 to 1000 Hz. The mean and standard deviation were obtained from 24 samples. The electrodes were precleaned (except the Ag/AgCl disposables) by gentle abrasion and the electrolyte was 100 mM KCl.

typically minutes to hours [49,58–60]. Clean, stabilized electrodes need to be exposed to the working signal for a few cycles for transients to settle before a reading is taken, and there is a need for a few cycles' grace between readings for the electrodes to return to their original state [53]. For this reason, in all spectral sweeps performed in this work, a settling time of five blocks of data is used at each individual frequency/voltage combination to allow the electrodes to precondition under this exact signal before recording the corresponding spectrum.

Electrode cleaning to ensure repeatable non-linear dielectric spectra is a complex and often empirical task, due to the lack of knowledge of the exact form of the causative mechanisms in the non-linearity of the electrode–electrolyte interface, and to a similar lack of knowledge of the mechanisms of binding of substances to the electrode surface and the subsequent electrical activity of these substances. Ensuring an adequate degree of stability in the harmonics of this interface for non-linear work is much more difficult than ensuring the degree of stability in the fundamental necessary for linear work. No repeatable and certain ways of obtaining a quiet and repeatable reference signal from an individual electrode surface have been found. Abrasion works best, and most reliably, but generally requires a settling time under the working signal in the working medium, as suggested above, before reliable, repeatable control results can be obtained. If protein coating is suspected, biological washing liquids with proteases may sometimes work, as may applying a slowly alternating voltage to oxidize the surface contaminants and subsequently removing the products by the reduction half-cycle [60–62]. Platinization of the surface may work, and if it does not, it then makes it easy to abrade off the newly platinized layer to reach clean metal. In practice, the most reliable sequence of operations for gold electrodes has been found to be platinization, followed by abrasion, and then by a settling period as above, in which the electrodes are driven by a low-frequency signal with an amplitude of the order of 1 V.

The electrode cleaning process is deemed successful when a repeatable, artefact-free signature can be obtained from a well-known and reliable reference system. In our case, a resting yeast suspension, produced as above, is used to provide reference signatures, since its reliability and stability have been proven over time. When a clean, recognizable, repeatable, artefact-free third harmonic frequency window can be obtained from a frequency sweep covering its frequency range, the electrodes are ready for use.

The maximum amplitude of the signal which may be applied to the electrodes is also limited, since visible electrolysis occurs whenever the voltage exceeds ± 1.5 V zero-to-peak, and at these voltages the electrode surfaces are much more susceptible to contamination, and results become unstable.

Once clean, electrodes may stay stable for weeks, or become unstable in minutes. Continual control experiments, performed as indicated above, are vital during any series of experiments to make sure that the electrode surface behaviour has not substantially altered during the experiments, in which case the results must be abandoned and the experiments repeated from the last successful control.

2.2. Electrode materials

We have investigated the linearity and stability of a substantial variety of electrode materials. One important feature to mention at the outset is that the lowest linear impedance does not necessarily imply the lowest non-linearity. Platinum (including platinized Pt) has a lower polarization impedance than gold and is a favourite in linear impedance spectroscopy [15], but it still has substantial polarization harmonics and these fluctuate with time more than those of gold. This stability is more crucial in forming reliable multivariate models of non-linear dielectric processes than the absolute harmonic level.

A system to measure the stability of the harmonic levels is needed. The measure chosen to give an indication of the variability of a harmonic with respect to its level is the coefficient of variation, i.e. the standard deviation (SD) divided by the mean, of the power spectrum harmonics. With repeated measurements (24 in these studies), this value can be calculated for each voltage/frequency combination and an SD/mean vs. f vs. V surface plotted for a particular harmonic.

Values of the mean and SD from which this measure can be calculated are shown, for the third harmonic, in Fig. 9, for a single voltage in the centre of the range usually used in non-linear dielectric spectroscopy, and a range of frequencies covering those of most interest in this application.

These results show that the gold electrodes fluctuate least, relative to the size of the harmonic, followed very closely by Ag/AgCl disposables, photolithographically screen printed with the intention of a use-once-and-throw-away electrode. As can be seen, the repeatability of these disposables was no better than the fluctuations of the best metal electrodes.

A measure using data from the first five harmonics, based on the calculation of the standard deviation of the scores on the significant factors of a principal component analysis model (see later) formed on the electrode fluctuations, was also tried. This gave similar, but not identical, results to the SD/mean of the third harmonic data shown above in ranking the stability of the electrode materials. This is probably a more representative method for assessing the stability of electrodes destined to be used in a multivariate modelling environment, but we consider that more work needs to be done on the development of such measures of electrode reproducibility.

2.2.1. Performance limitations of Ag/AgCl disposable electrodes

The Ag/AgCl disposable electrodes need to be used in a solution containing at least 5 mM Cl^- to stabilize the surfaces, otherwise the signal applied to them will cause alteration of the electrode surface due to irreversible solution of Cl^- from the surface. Any alteration of the surface decreases with continued exposure to the signal, so that its effects can be reduced by feeding the electrode with a dummy signal sweep before actually recording any data.

If one set of "disposable" electrodes is used repeatedly, the SD/mean value will represent the overall variation including the electrode drift. This is typically of a similar order of magnitude over a period of 5 h as the random fluctuation component (i.e. the fluctuations are wideband). Also, with the abovementioned Ag/AgCl disposable electrodes, this drift is of a similar size to the differences between different individual electrode sets so that, at the present stage of electrode development, there is little benefit in repeatability from one experiment to the next in using a different set of electrodes for each reading. In fact, the tradeoff is that a single set will drift, but give slightly smaller random fluctuations and less noisy results if the modelling process used on the results can sift off the drift, whereas different electrodes eliminate the drift at the cost of increased noise in the data which the models must selectively reject. At present the benefits of this tradeoff are fairly even.

Since the electrodes alter when exposed to the a.c. voltage signals typically used in data sweeps, it is of interest to see how long it takes for them to stabilize in non-ideal conditions, with no chloride in the medium (Fig. 10). A rapid initial transient was observed in the harmonics, with detectable settling for up to 1.5 h and the fundamental still drifting up to 5 h. It is probably valid to regard the initial less than 10 s transient as a chemical equilibration of the electrode surface under the working signal and the longer term variation as the usual electrode polarization drift.

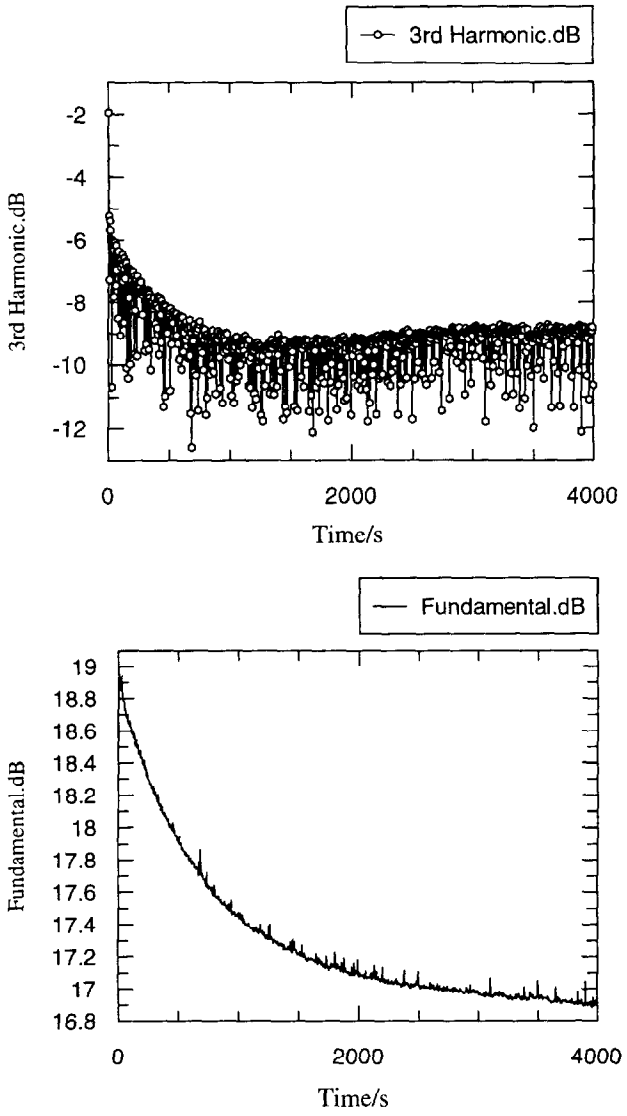

Fig. 10. Settling of third harmonic and fundamental in a single set of Ag/AgCl disposable electrodes. The medium used was yeast supernatant with no added chloride, and the fundamental had an amplitude of ± 1.2 V at a frequency of 15 Hz.

2.3. Alternative Ag / AgCl electrode configurations

Although disposable Ag/AgCl electrodes have clear practical advantages for commercial applications if their repeatability of manufacture can be improved, better results can be obtained in the laboratory with self-deposited layers of AgCl.

A standard four-terminal electrode system was used.

The outer silver pins have a layer of AgCl predeposited on their surface with a Princeton Applied Research model 174A polarographic analyser. The counter electrode is graphite, the reference electrode is Ag/AgCl and the silver electrode to be coated is the working electrode. The inner pins are gold. The effects of increasing the charge passed so as to deposit an increasingly thick layer of AgCl can be seen in Fig. 11. The polarization harmonics decrease until $600\,C\,m^{-2}$ of charge (10 mA for 60 s) have been passed; then there is no further benefit from increasing the layer thickness. The harmonics were measured using a solution of 100 mM KCl in the electrode chamber.

This configuration produces the frequency sweep spectrograph given in Fig. 12(A), which can be compared with the similar spectrograph in Fig. 12(B) produced under identical field conditions by an identical chamber with all-gold electrodes.

The Ag/AgCl electrodes show a considerable reduction in harmonics compared with gold electrodes. These also fall off more rapidly with increasing frequency. The linear impedance is reduced by a factor of two, but the linearity increases (harmonics decrease) with the chloride concentration in the medium, and high chloride concentrations can be inimical to biological cell suspensions.

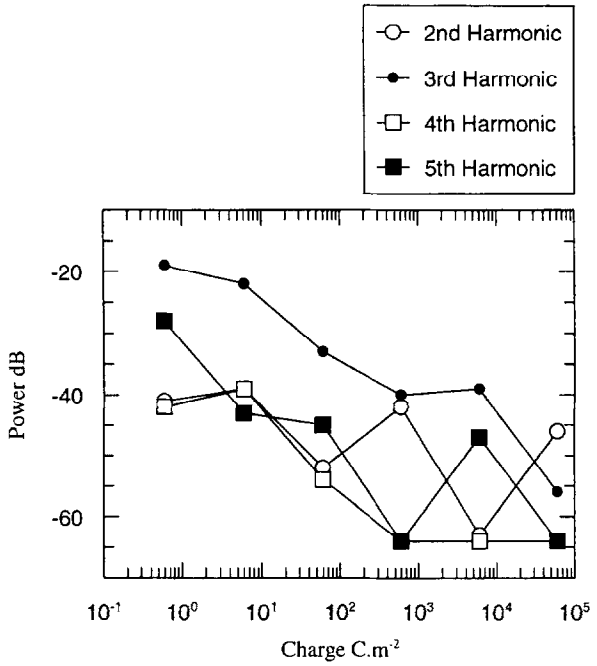

Fig. 11. The effect of the AgCl layer thickness (reflected in the amount of charge passed to produce it) on the polarization of Ag/AgCl electrodes.

A further substantial improvement in linearity without sacrificing biocompatibility can be achieved with the more complicated electrode chamber given in Fig. 13. In this configuration, no predeposition of AgCl is required since a thin, high integrity layer is automatically deposited by the working voltage signal from the very concentrated (typically greater than 3 M) KCl surrounding the driver electrodes.

The best results (in terms of lowered harmonics) can be obtained by cleaning the outer electrodes by mild abrasion and using them with no predeposition of AgCl. A sufficient layer of AgCl evidently builds up automatically under the a.c. signal, presumably due to rectification at the interfaces, and is in fact just visible as a light grey patina. The conductivity of the sample in the electrode chamber (in the range 3–15 mS cm^{-1}) has no noticeable effect on the levels of the harmonics (not shown).

We also studied the effect of the conductivity of KCl in the outer electrode channels of configuration 2 on the generation of harmonics. This shows a small, but noticeable, improvement to better linearity as the concentration of the KCl around the outer electrodes is increased, presumably due to a continuing thickening of the AgCl layer.

The harmonics are now down to a negligible level except at frequencies below a few Hertz, as shown in Fig. 14(A). An ultralow-frequency sweep from 1 to 10 Hz (linear) at ±8 V (giving approximately ±80 mV across the inner electrodes) with an electrolyte of 100 mM KCl shows that the harmonics drop below background (at −55 dB) at frequencies typically above 4 Hz.

The linear impedance of this arrangement is high (typically 20 kΩ at the working frequencies) because of the impedance of the frits, so a higher driver voltage must be used to obtain a working current through the electrolyte in the central portion of the electrode chamber between the inner electrodes. It is also almost invariant with frequency, another indication of the approximate linearity of the electrode interface.

However, the relative complexity of this configuration somewhat militates against its use in routine applications, and the slow leakage of concentrated KCl through the frits, even if these have a pore size of the order of a few nanometres, means that it is only suitable for short-term experiments in the laboratory over a period of hours.

To show the advantage of the AgCl coatings over raw silver driver electrodes, however, Fig. 14(B) shows a spectrograph with raw silver electrodes.

2.4. The practical effect of polarization variation on the experimental procedure

The variability of electrode polarization and electrode surface fouling are currently the most debilitating problems afflicting non-linear dielectric work at low frequencies. There is currently no satisfactory solution that can ensure repeatable electrode surfaces for each experiment or series of experiments, with the possible exception of configuration 2 above. However, more work is necessary to confirm the practicality of this configuration in working experiments on general suspensions.

The only workable approach with conventional electrodes involves continual controls vs. a reference to check the stability of the surface, followed by rejection of the series of experiments from the point at which these controls fail. This can make a series of comparative experiments a prolonged and cumbersome process which is currently unsuited to general use outside a laboratory.

All data sets presented in this paper have been subjected to this system of controls.

3. Multivariate data analysis

Let us consider an experiment to study the utilization of exogenous glucose by an organism, such as baker's yeast, as a function of time. A spectral data sweep is taken at intervals, and the corresponding "true" glucose level is measured by a reference (wet chemical) method. If, for example, a data sweep is performed with the fundamental set at each combination of ten frequencies and three voltages and, at each voltage and frequency, five harmonics are stored from the spectrum, then each sweep will have resulted in the acquisition of a data set consisting of 150 measured variables which correspond to, and are hopefully representative of, the single glucose concentration of interest. Clearly, forming a mathematical model

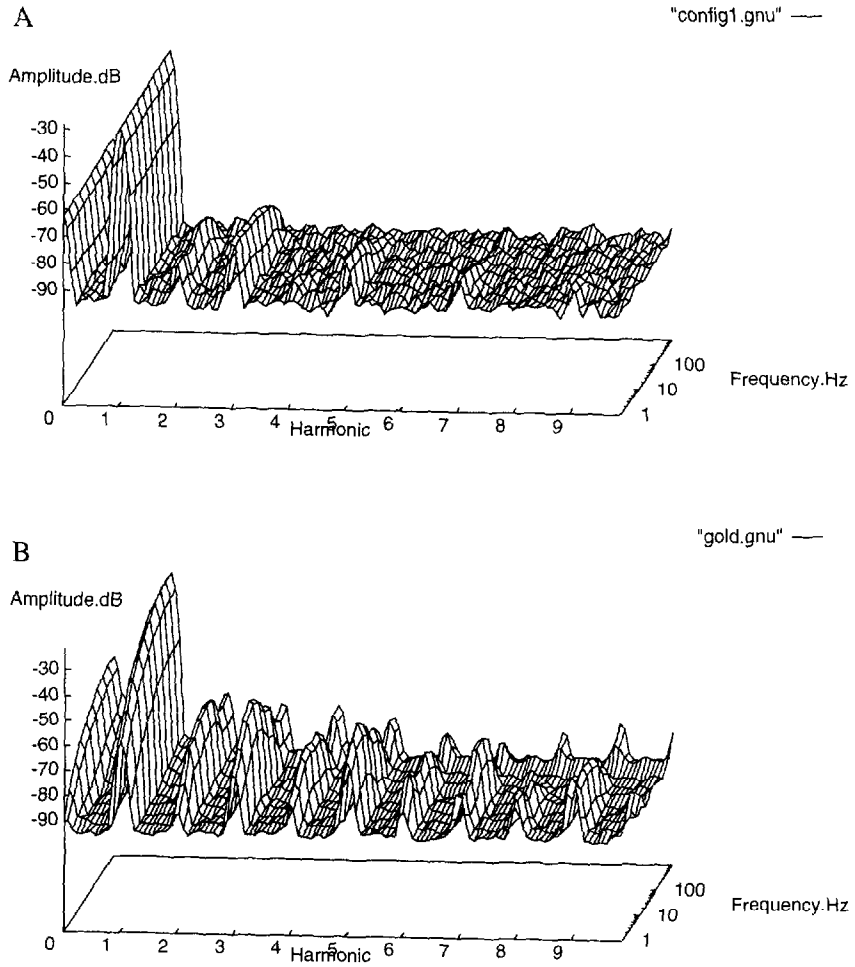

Fig. 12. (A) Spectrograph of Ag/AgCl electrodes in configuration 1. The applied field through the electrode chamber was $\pm 2\,V\,cm^{-1}$ and the frequency was scanned logarithmically from 1 to 100 Hz. The working medium was 100 mM KCl. (B) Spectrograph of gold electrodes taken under identical experimental conditions to (A).

from these variables to represent the glucose value "by eye" would be prohibitively difficult!

Multivariate modelling methods [63] provide a means of reducing data in multiple dimensions to their equivalent in a manageable number of dimensions, without rejecting significant information. Partial least-squares regression (PLSR) is an increasingly common tool in biochemical calibration problems.

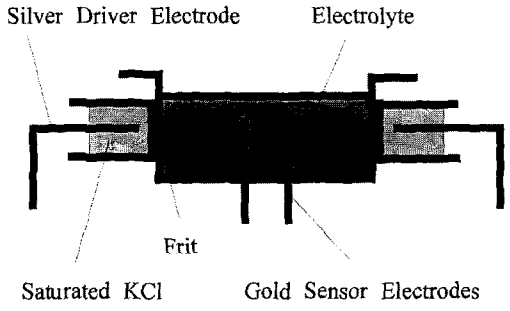

Fig. 13. Ag/AgCl electrodes: configuration 2. The very concentrated KCl adjacent to the driver electrodes is isolated from the sample in the chamber by sintered glass frits with a pore size of the order of nanometres.

3.1. Chemometrics

The use of PLSR and related methods in a wide range of sciences has led to the emergence of a new discipline, that of chemometrics. This is devoted to the study of more or less pragmatic multivariate methods in sciences, and the literature has grown rapidly over the last few years [64]. With the increasing use of PLS has come the realization that it is not well understood in terms of its statistical properties. The methods were designed to avoid making assumptions about structure within the data except those that are built into the path model. No assumption of normality is placed on the variables, for instance.

A simplified "crib sheet" of some practical considerations for new users of PLS modelling is included in Table 1.

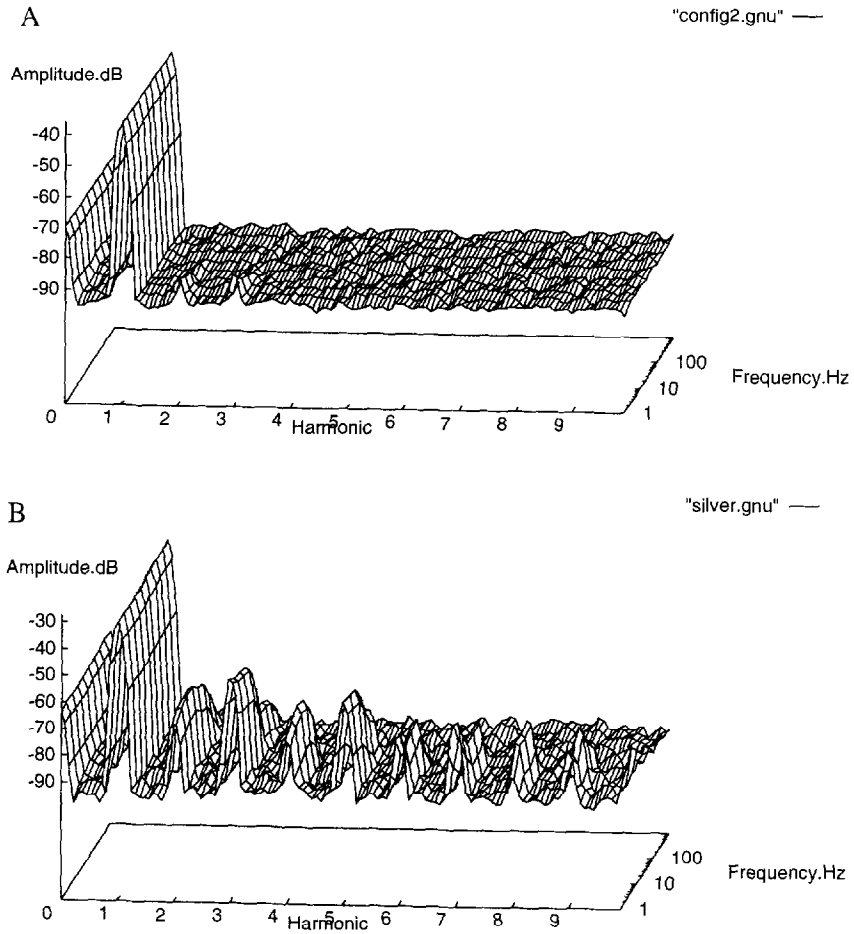

Fig. 14. (A) Spectrograph of Ag/AgCl electrodes in configuration 2. (B) Spectrograph of pure silver electrodes. The experimental conditions in both cases are the same as for Fig. 12.

3.2. Artificial neural networks

The multivariate modelling methods outlined above are very efficient at modelling predominantly linear and, to a degree, weakly non-linear phenomena. However, when the relationship between the x variables and the y data contains any degree of non-linearity, the models can begin to lose accuracy, and a fundamentally non-linear modelling method is required. Artificial neural networks (ANNs) fill this requirement.

ANNs are an increasingly well-known means of uncovering complex, non-linear relationships in multivariate data, whilst still being able to map the linearities. ANNs can be considered as collections of very simple "computational units" which can take a numerical input and transform it (usually via summation) into an output (see Refs. [65–82] for excellent introductions, and Refs. [83–106] for applications in analytical chemistry and microbiology). Also it has been proven that simple neural net architectures containing one arbitrarily large hidden layer using a non-linear

Table 1
Some hints in choosing parameters when setting up PLS models

1	Scaling of x variables: input x variables should be mean centred and scaled to $1/s$ where s is the standard deviation of the variable. If variables are small, significant variables will be underweighted in the model in favour of large insignificant variables
2	Number of exemplars in training set: need enough to fill parameter space evenly and to allow generalization
3	Number of input variables: those that do not contribute positively to discrimination may impair generalization and are best removed by pruning the input data
4	Validation: best to reserve some of the training data for validation on an independent, representative test set. If there are too few samples to allow this, leverage correction or leave-one-out cross-validation may be used. Leverage correction will typically overestimate the optimum number of factors in the model, whereas leave-one-out cross-validation will typically underestimate
5	Residuals: plotting the residuals Fiy against y indicates regions of unmodelled non-linearity. The residuals should be randomly distributed around zero. Any deterministic structure denotes non-linearity outside the modelling capabilities of PLS

squashing function can approximate any continuous mapping to arbitrary precision [107–111].

A simplified overview of some practical considerations for new neural net users is given in Table 2, and shows some parallels with PLS modelling [112].

Outside the rather inaccessible mathematical literature, there has been relatively little work on the statistical validation of neural network predictions. Thus although they can be trained to the optimal point, when challenged with a new stimulus the network will give its answer but, as yet, it is not possible to put accurate confidence limits on the prediction [72]. However, the link between statistics and neural networks is now becoming increasingly realized [103,113–115], and it is arguably only a matter of time before true statistical confidence limits (beyond simple mean ± SD on replicates) will be applied to neural network outputs.

4. Method comparison

Before considering the experimental aspects of this paper, a few words about method comparison are appropriate, since our experience is that the pitfalls in the most commonly encountered approaches are not widely recognized.

The most commonly used technique to compare how similar the estimates of an unknown measurement method (y) are to those obtained using a known or "gold standard" method (x) is linear regression analysis. This is used to reveal the numerical relationship between the two estimates for each sample in terms of a slope and intercept of the regression line, and a correlation coefficient to indicate the precision of this relation. Notwithstanding that its very ubiquity ensures its continued use, this linear regression approach is often used incorrectly, especially in situations where its fundamental assumptions are invalid [116–119].

The least-squares method assumes that there is no error on the values given by the reference method, and also that the errors in the equivalent values predicted by the test method are randomly distributed with a gaussian distribution [116,120]. The former assumption is only correct if one is able to use a gold standard method with negligible error. The latter is more insidious. Most practical measurement instruments include an error which is related to the absolute value of the reading, i.e. the standard error of the estimate at y is proportional to y. This violates the gaussian assumption which gives an identical standard error of the estimate at all y values. In other words, the measuring instrument has a constant coefficient of variation, not a constant measurement error. However, linear regression can be shown to be an adequate approximation for data with a coefficient of variation of less than 20% [116]. The practical effect of errors in the reference readings is to produce regression lines with artefactually underestimated gradient and similarly overestimated intercept, making systematic differences hard to characterize [117]; high measurement errors on data that are otherwise well correlated will give a slope of almost zero!

The effect of errors in the reference data means that, mathematically, there is no difference between the refer-

Table 2
Some hints in choosing parameters during the production of a feedforward backpropagation neural network calibration model to improve learning/convergence and generalization

1	Number of hidden layers: one is thought to be sufficient for most problems. More give a big increase in computational load with little, if any, benefit to the model
2	Number of nodes in hidden layer: PLS model gives a reasonable upper bound. Number of hidden nodes ⇐ optimum number of PLS factors. If this is not possible or desirable, then a suitable (but often pessimistic) rule of thumb says the number of hidden nodes = ln(number of inputs)
3	Architecture: fully interconnected feedforward net is most common. Many others exist, such as adaptive resonance theory, Boltzmann machine, direct linear feedthrough, Hopfield networks, Kohonen networks
4	Number of exemplars in training set: need enough to fill parameter space evenly and to allow generalization. When fewer are used, the network can "store" all the knowledge, predicting the training set superbly, but unseen data badly
5	Number of input variables: those that are weakly related or unrelated to the output data may impair generalization and are best removed from the input data set before training commences
6	Scaling of input and output variables: individual scaling on inputs improves learning speed dramatically and often the accuracy and precision of the predictions. There is a need to leave headroom, especially on the output layer
7	Updating algorithm: there are many variants on the original backpropagation (BP), some of which give small but worthwhile improvements. Others include radial basis functions, quick-prop, stochastic BP, Weigend weight eliminator. Standard BP [94] is still the most popular. For the non-specialist, there is little point in using the others unless satisfactory results cannot be obtained with this
8	Learning rate and momentum: need to be carefully chosen so that the net does not get stuck in local minima, nor "shoot" off in the wrong direction when encountering small bumps on the error surface. For standard BP, a learning rate of 0.1 and a momentum of 0.9 are a good starting point. These parameters may need adjusting before or during training if the net gets stuck in a suspected local minimum
9	Stability: best to remove a representative sample of the training data for validation. If there are insufficient samples to allow this, an approximation to the correct place to stop training in order to prevent overfitting can be made at the point at which the training error bottoms out after the first major drop (e.g. after about 30 epochs in Fig. 25)

ence and test data, so it is then invalid to choose either the x axis or the y axis as the reference for a regression analysis [116]. A method of correcting this problem has been proposed [121] by simultaneously minimizing the sum of the squares of the residuals on both axes, but this still assumes a constant error of measurement.

Use of the correlation coefficient alone for method comparison is dubious because it reflects the range of the data as much as the similarity between x and y data points [117,118]. If the data encompass only a small range, e.g. 100–120 arbitrary units, then despite an error of only 10% in the absolute values between the two methods, which agree with each other to satisfactory error bounds in individual (x,y) measurements, the data cluster will be almost circular, and the correlation coefficient will be low, suggesting no relation between the two methods. Conversely, as the data range increases, the correlation coefficient will increase, suggesting a better matching between the two methods despite the fact that they still agree to $\pm 10\%$. Westgard and Hunt [122] go so far as to say that, "The correlation coefficient is of no practical use in the statistical analysis of comparison data".

Regression also suffers from this phenomenon. Let us assume that, in the restricted range of data considered above, the regression line initially has an arbitrary slope. As the range of the data is extended, the equation of the regression line becomes closer to the true relation between the two methods, but will never actually equal it if there are errors in the x values, always overestimating the intercept and underestimating the slope. In a situation in which the range of the data is significantly curtailed, even a simple visual inspection of the data is more informative than a linear regression analysis [117].

It thus transpires that a simple line of identity between the estimates derived from the two methods (including percentage error boundaries, particularly in those common cases in which the data have a fixed coefficient of variation) is both more informative visually and more correct mathematically. An alternative to the percentage error criterion is the difference vs. mean plot, where each point on the y axis corresponds to the signed difference between the x and y values of a datum, and the x axis corresponds to their mean [117], a method closely allied to the display of residuals after a modelling process. This error measure makes no assumptions about the precision of the reference values, which is a failing with the percentage error, since the percentage is of the y data calculated against the x data [118]. A third approach to the situation is to plot the median (or mean) slope of all combinations of lines which can be drawn between any two data points [119,123,124]. This approach is also non-parametric, making no assumptions about the distribution of the data.

Other particular instances abound, particularly in medicine, where clinical criteria of efficacy are more important than statistical measures of any sort. The measurement of blood glucose in diabetics is one such example [125,126]. In this situation, a "region of acceptability" (which will be typically asymmetric about the ideal 1 : 1 line) as in the Clarke diagram [125] is the clinically correct criterion to use. In this, the errors in the predicted readings from the true readings are judged in such a way as to avoid terminal complications in the patient. This may require high precision in some regions of blood glucose concentration, while allowing large errors in other regions.

5. Experimental investigation of yeast metabolism using multivariate calibration of non-linear dielectric spectroscopic data

5.1. Data collection

The yeast *Saccharomyces cerevisiae*, prepared as a suspension as outlined above, was used as the test organism to determine the ability of multivariate modelling to predict levels of glucose in unseen suspensions, since it is readily available and its biological and dielectric properties are well researched. It was considered that the study of its metabolism in this way was adequately representative of the metabolism in batch culture. PLS was chosen as the most reliable of the easily available linear multivariate statistical methods, and the results from this were compared with those on the same data sets modelled with neural networks.

Five data sweeps were recorded with no glucose present; then glucose was added to a concentration of 200 mM and data sweeps were taken every 2 min until the glucose was used up. For this strain, this typically took 90 min from the addition of the glucose. (Cell counting procedures showed that cell growth did not occur in these experiments.) Five subsequent data sweeps were recorded after the glucose levels in the suspension had reached zero.

The glucose levels were measured every third data sweep with a Reflolux device designed for at-home testing by diabetics. We have measured the precision of this to be $\pm 10\%$. It has a detection range of 0.5 to 27 mM, so that the higher glucose concentration samples used in the yeast work were diluted before readings could be taken using this device.

Each data sweep scanned frequencies of 5 to 50 Hz in 5 Hz intervals at voltages of ± 1 V, 1.25 V and 1.5 V. At each voltage/frequency combination, a power spectrum was produced from 30 blocks of data averaged to reduce noise [127] and harmonics 1–5 of this spectrum were recorded to disk. The phase was not recorded, since previous work had shown that its presence did not usually improve the models formed, and its inclusion could in some cases degrade models due to variable selection considerations in accordance with the parsimony principle [128] if the phase was not well correlated with the glucose concentration. The suspension in the electrode chamber was then replaced with a conductivity-matched supernatant

and the sweep was repeated. The difference spectra due to the biology with the electrode polarization deconvolved can then be calculated. The suspension, supernatant and difference spectra were recorded to separate files. This leads to each sample (object) being composed of 150 x variables in each file.

For the following experiments, used as the basis for the present development of the multivariate calibration methods, three identical experimental runs (sweeps 1–3) were carried out on separate days. Predictions between these sweeps were all similar, so the model formed on sweep 1 (47 objects, 150 x variables) predicting sweep 2 (49 objects, 150 x variables) was chosen as the demonstration pair for all following development, unless otherwise stated.

Modelling and prediction were initially carried out on the difference data, since these are where the biological effects are greatest in proportion to the (ideally negligible) electrode polarization.

For the reasons discussed above, the method illustrated for comparison between the reference and predicted data was a simple superposition of the two data sets against time (presented as object number) to indicate the degree of fit of the prediction to the form of the reference plot (particularly the leading zeros; see later for a discussion of the importance of these in model assessment), together with a plot of the predicted/reference data compared with the line of identity to indicate the accuracy and precision of the prediction. However, in order to interface with traditional schemes of method comparison, and taking into account the discussion given above on its validity in such applications, a regression line was also derived for selected plots.

5.2. Multivariate analysis of yeast data using PLS

The PLS modelling processes were performed using THE UNSCRAMBLER 5 package (Camo A/S, Olav Tryggvasonsgt. 24, N-7011, Trondheim, Norway). This provides facilities for outlier detection and internal validation by either leverage correction or internal cross-validation.

In data as noisy as those produced by non-linear dielectric spectroscopy, due to the variations of the electrode–

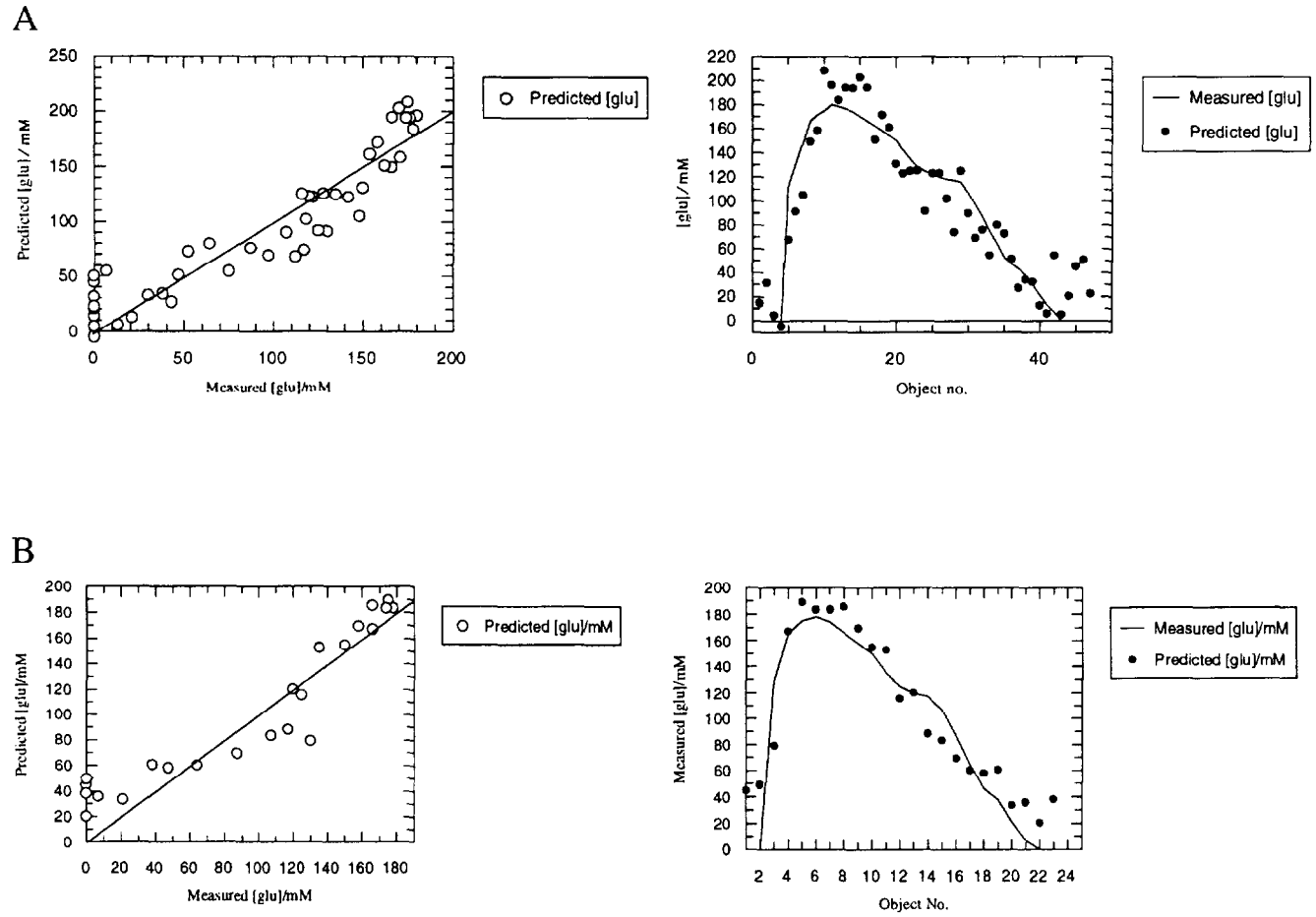

Fig. 15. (A) PLS prediction of the glucose concentration in a yeast cell suspension using difference spectra within one data run. The model was formed on every third sample and the remaining two-thirds of the data were predicted from this model. Each object corresponds to a time delay of approximately 2 min, the whole data set encompassing a period of approximately 90 min. The optimum model was formed with three PLS factors. (B) PLS prediction on suspension spectra within one data run. The model was formed on every odd-numbered sample and the remaining even-numbered data were predicted from this model. The optimum model was formed with two factors.

electrolyte interface, the outlier identification facilities are of limited use. If, during modelling, the residual variance does not decrease smoothly to the optimum number of factors, this is indicative of the presence of outliers, which must thus be identified individually by other means. In our experience, a comparison of scores plots and influence plots is the most helpful method for this. Scores are the projections of the data onto the individual factors, so that any point which is significantly outside the main cloud of the other data is suspicious. Influence is a plot of leverage vs. residual variance. A point with unexpectedly high leverage has a disproportionate effect on the model, and one with high residual variance is not well represented by the model. Thus a point with disproportionately high leverage and/or residual variance is likely to be an outlier and should be removed from the data set before modelling.

The models were always centred on the mean of the individual x variables, and these x variables were "scaled" to unit variance, i.e. to the reciprocal of the standard deviation of the individual variables. The optimal model was indicated by leave-one-out cross-validation [63].

5.2.1. Predictions inside a data set

If a model is formed on a representative selection of samples from one data sweep (e.g. sweep 1), and the remaining samples from that sweep are predicted, the training set has information on the same broad changes in electrode polarization as occur in the test set. Since the glucose results are taken every third sweep, these sweeps (non-interpolates) can be used to form a model to predict the remaining two-thirds of the data points (interpolates). We would expect the prediction to be better than that made between successive sweeps, due to the ability of the model to follow the changes in the electrode polarization between sweeps.

Cross-validation gives optimum modelling with three factors for a model formed on the non-interpolates of sweep 1. The predictions of interpolates by non-interpolates is good, as expected, but noisy (Fig. 15(A)).

Predictions made on the suspension data rather than the difference data are not substantially degraded, showing that the PLS can deconvolve the electrode polarization from the biology, given representative data on both the polarization and the biology (Fig. 15(B))

5.2.2. Predictions between data sets

A cross-validated model was formed on the difference spectra of one data run (sweep 1) and this model was used to predict a separate data run recorded on a separate day (sweep 2). As can be seen from Fig. 16, there is a significant modelling of glucose levels, but this model is very noisy. The prediction can be seen to be slightly worse than that of Fig. 15(A), suggesting that the polarization of the electrodes has changed slightly.

The prediction between these two difference spectrum sweeps is used as the basis for the following comparative development work on the modelling process.

5.2.3. Median vs. mean averaging

A major reason for the noise in the above prediction is that the data are polluted by "glitches". These are broadband, affecting all harmonics within a single spectrum as can be seen from Fig. 17, and are due to momentary hesitations of unknown origin in the A/D–D/A unit. The mean averaging normally used as the averaging method in computing the power spectrum is sensitive to these glitches, giving the raised harmonics noted above.

Forming a median-averaged power spectrum instead of the conventional mean-averaged spectrum removes the effect of these glitches at the cost of slightly increasing the

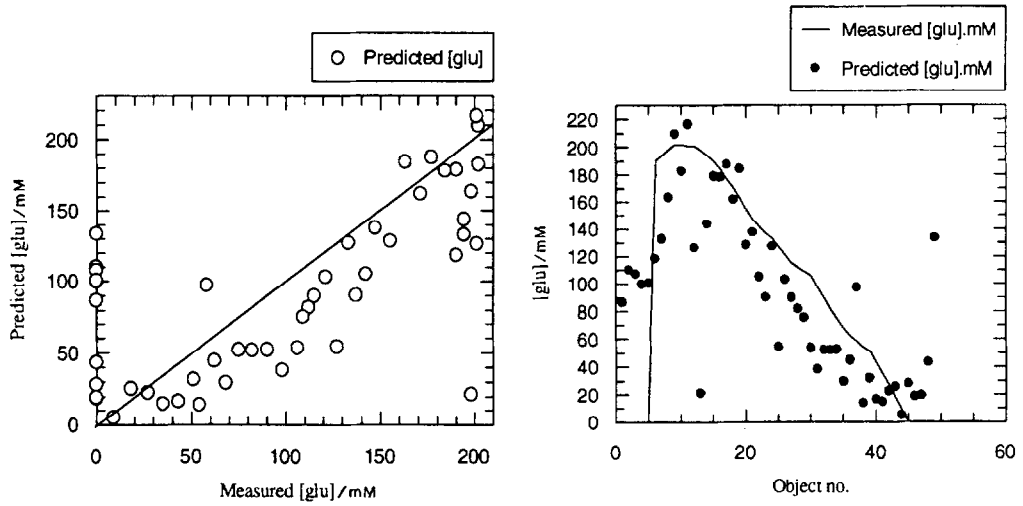

Fig. 16. PLS predictions between separate data runs. The model formed on sweep 1 is used to predict sweep 2. The optimum model, as judged by the minimum least-squares error of prediction of the known values, was formed with three factors.

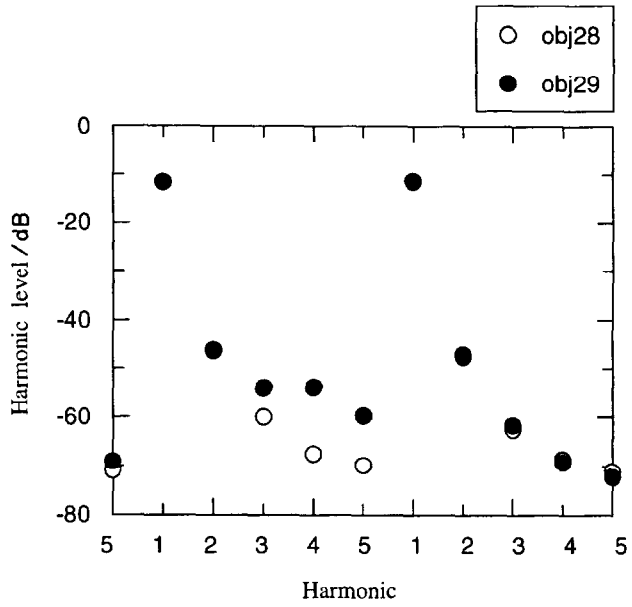

Fig. 17. A two-spectra-wide section of harmonic variables from two successive objects in a data run. As can be seen, a glitch has occurred in the left-hand spectrum in object 29, raising the harmonics above the true level of object 28 even after being mean averaged over 30 blocks.

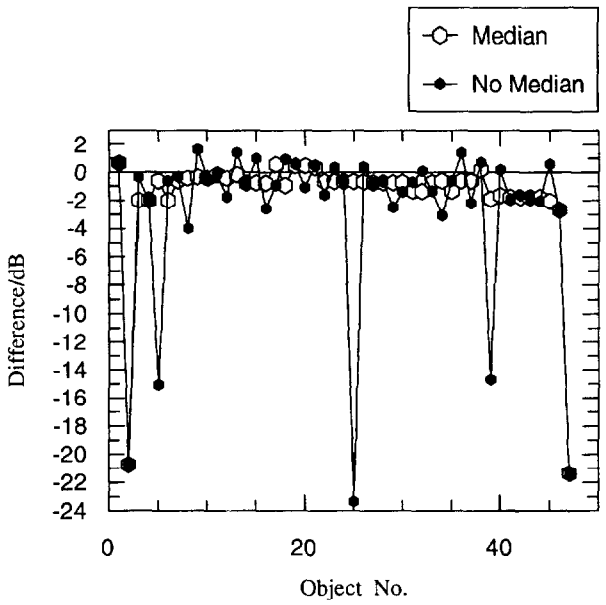

Fig. 19. The effect on a single variable of cleaning up noisy prerecorded mean-averaged power spectra with a five-point sliding median average filter. The difference spectrum shown is of a supernatant vs. supernatant sweep, and should ideally be null.

variance of the resulting averaged power spectrum as shown in Fig. 18.

With prerecorded files using a conventional power spectrum, the glitches can be removed by a sliding five-point median average down each individual variable, since the experimental sampling frequency is sufficiently high that it can be assumed that glucose-related effects will change only slowly between adjacent samples since [glu] does likewise (Fig. 19).

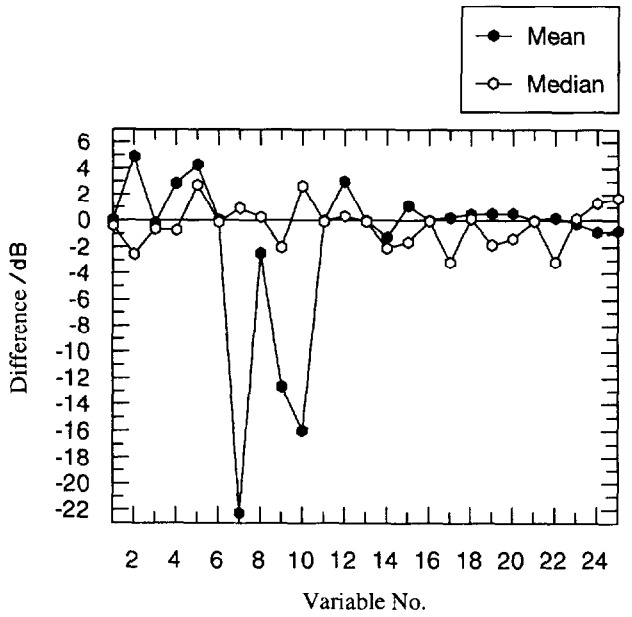

Fig. 18. The effect of median and mean averaging in the formation of the power spectrum. The first 25 variables of the difference spectrum of one object, for both a mean-averaged power spectrum and the equivalent median-averaged power spectrum, are superimposed to show the reduction of outliers which can be achieved by median averaging. The difference spectrum shown is of a supernatant vs. supernatant sweep, and should ideally be null.

It is clear that median averaging in this way removes noise from predictions very effectively, as can be seen in Fig. 20, both cleaning up the noise in the prediction and reducing the optimum number of factors to two, indicating a stronger modellable effect and allowing PLS to form a model using fewer latent variables.

5.2.4. *Variable selection*

As outlined above, our experiments and those of others [128–131] suggest that PLS modelling is actually somewhat limited in its ability to select the x variables that are most highly correlated with the desired y data and to ignore the relatively uncorrelated x variables; this suggests that PLS can become unreliable due to spurious modelling of coincident (chance) correlations in generally uncorrelated variables, especially if the number of variables is appreciably greater than the number of objects. This general problem is often referred to as the "curse of dimensionality" [132] and leads to greatly degraded generalization of the models formed. The variables fed to a PLS model should therefore preferably be the minimum number that contain all the necessary information about the process being modelled. Visual inspection of harmonic plots shows that the fourth and fifth harmonics contain little obvious variation with the glucose concentration.

Using the second harmonic only (30 variables) results in the much less noisy models of Fig. 21, but overesti-

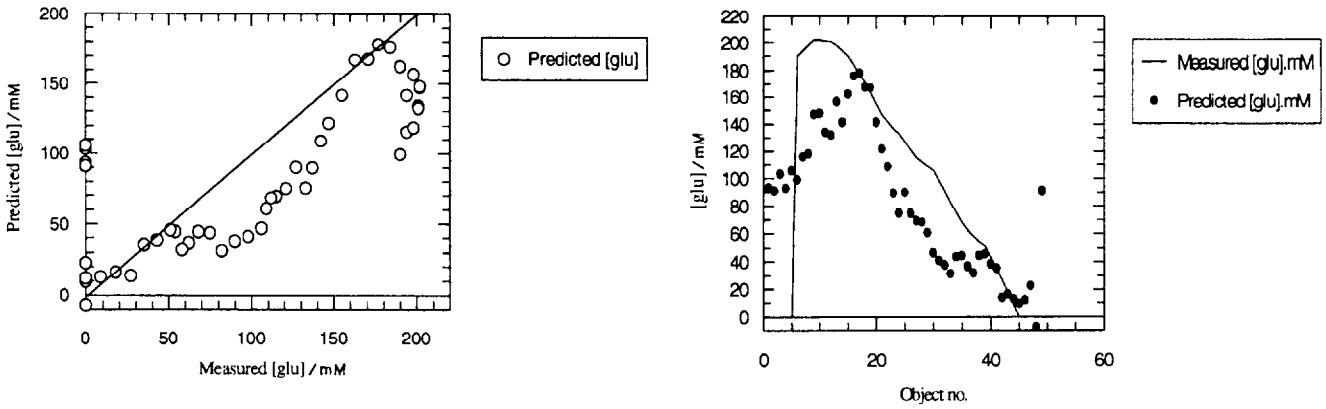

Fig. 20. The improvement in PLS predictions produced by median averaging the data after recording. The optimum model was formed with two factors instead of the three required for raw data.

mates low [glu] and slightly underestimates high [glu]. Using both second and third harmonics (60 variables) results in better low [glu] predictions, but a slightly noisier model than with the second harmonic alone. Therefore most information, but not all, is carried in the second harmonic. The third harmonic appears to add significant information about the low [glu] range, but its inclusion also begins to allow the modelling of noise. The addition of fourth and fifth harmonics merely increases the noise without improving the prediction. These findings are highly strain/organism dependent. The most significant harmonics/frequencies/voltages will differ greatly for different suspensions. They may differ in detail even for the same strain in different data runs in different conditions. The significant variables must be determined adaptively for all new batches of data in any automated analysis.

5.2.5. *Linearizing x data*

The reference measurements are linear with glucose concentration, but the spectrum harmonics contain a logarithmic transformation, being formed from a classic power spectrum. It may therefore be wondered if linearizing these

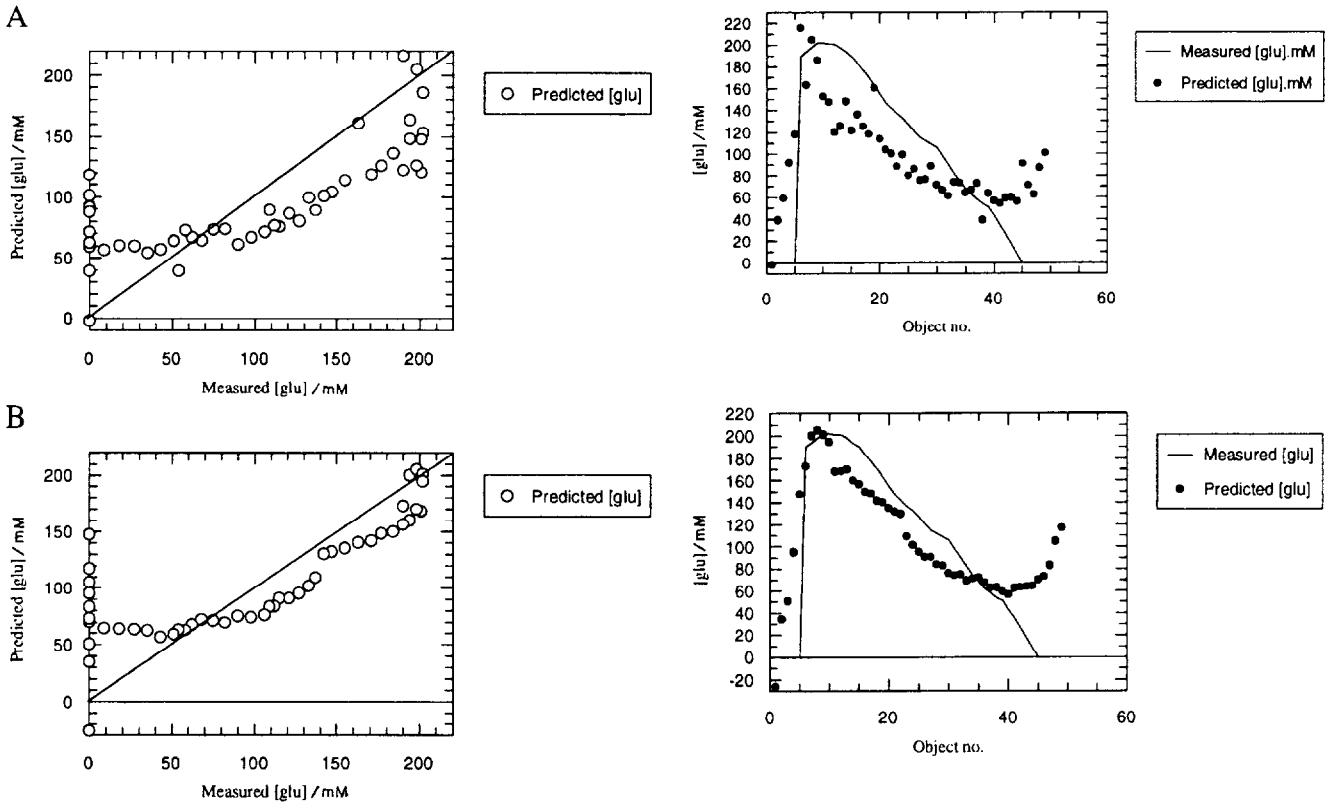

Fig. 21. The effect of variable selection (in this case, use of only the second harmonic) on the quality of the PLS model: (A) raw data; (B) median-averaged data. The predictions are obtained from a model with two factors optimal.

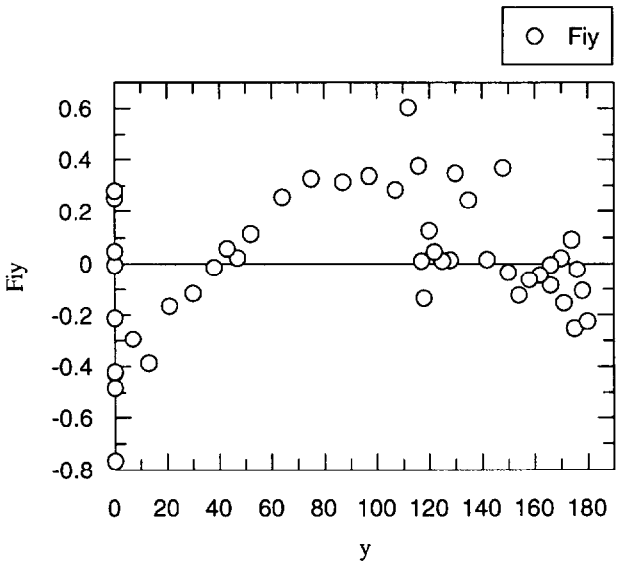

Fig. 22. Unmodelled non-linearity in the PLS model below a glucose concentration of 100 mM: the structure in the plot, indicating the unmodelled non-linearity in the data, can clearly be seen.

will make a worthwhile improvement to the predictions; however, doing so makes little alteration to the predictions, the plots being similar to the above; one interpretation of this is that the limiting factor in improving the model is the inherent underlying non-linearity in the response of the harmonics to the glucose concentration.

Referencing the harmonic levels to the fundamental also makes negligible improvement to the predictions, suggesting that the fundamental has little effect on the model, i.e. it is formed predominantly on the non-linear aspects of the dielectric response. Again the plots are much the same as above.

5.2.6. PLS residuals and unmodelled non-linearity

All permutations of PLS predictions between sweep 1, sweep 2 and the very similar but not shown sweep 3 (which was experimentally identical and produced data which look and behave very similarly) all converge to very similar clean, tightly modelled predictions of very similar curved lines, with small gain/d.c. anomalies, in which the predicted data either span a slightly different range of [glu] to the true values and/or are displaced by a small constant value due to slight alterations of the electrode surfaces between the sweeps. PLS models are observed to be extremely sensitive to baseline shifts in the variables. Consequently, even the slightest variation in electrode polarization can cause large gain/d.c errors in predictions. The so-called multiplicative scatter correction [63] has

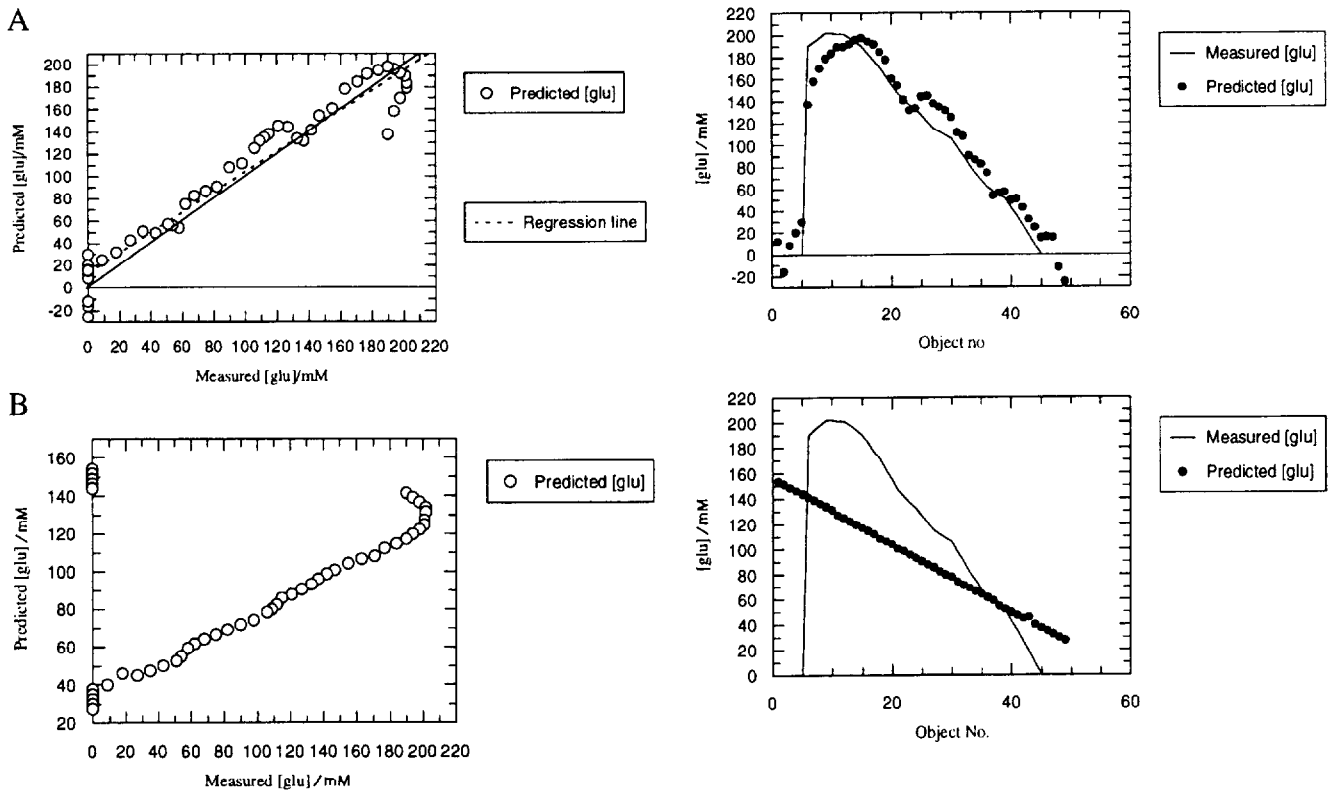

Fig. 23. The effect of adding time indices to the x data to assist the PLS regression to model out electrode drift. (A) Two-factor prediction obtained on median-averaged second harmonic data; this technique is shown to be accurate and precise to within the errors of the reference method. The regression line is also shown for compatibility with traditional methods of regression analysis, commonly but incorrectly applied to method comparison data. The slope of this line is 0.91 and its intercept is 13.4. The correlation coefficient is 0.97. (B) Equivalent control experiment demonstrating that this modelling is not occurring purely on the time variables.

been applied to near-infrared data in an attempt to remove/reduce this sensitivity to baseline, but the results when applied to our data were ambiguous, giving improvements in some predictions, but no improvement, or worsening, in others.

The consistently tightly modelled, but curved, predictions obtained between sweeps 1 and 3 suggest that although PLS is modelling what it sees very reliably, the underlying process being modelled is too non-linear for PLS to cope with. For a linear, fully modelled process, a plot of residuals *Fiy* vs. *y* (where *y* is the true glucose concentration) should be randomly distributed with no structure [63]. This plot (Fig. 22) for the model formed on sweep 1 shows a clear structure, falsifying the hypothesis that the underlying processes are adequately linear for modelling using a purely linear PLS model.

The residuals structure is most marked below 100 mM and explains why predictions are generally good above this value, but curve progressively away from the ideal 1 : 1 line below it. This suggests that non-linear modelling techniques, such as neural nets (see above) or non-linear PLS [133–139], may model the glucose concentration data more effectively over the whole range of concentration.

5.2.7. Time indices

Adding a time variable to the *x* data is a technique that has been used to assist multivariate methods to model time-dependent phenomena, such as microbial fermentations [140,141]. The slow drift in our electrodes is monotonic, whereas the response of the suspension to glucose is not. The inclusion of the time parameter can thus substantially assist the predictions to model the low glucose concentrations and leading zeros by helping the PLS regression model out the drift while ignoring the biology. Increasing the number of time variables further improves the predictions up to an optimum of four time indices per 30 harmonic variables as shown in Fig. 23(A). The control experiment of Fig. 23(B), modelling only the time variables against glucose concentrations, gives a much poorer prediction, showing that in this case there is no modelling value in time variables alone, but that they can be very valuable in modelling out slowly changing backgrounds, such as electrode drift.

At this point, the PLS prediction agrees with the reference values to within the measured error bounds of the latter for all points except the leading zeros (which are still quite well predicted). This again suggests that electrode variations are the major source of problems in predictions between (and within) runs.

5.2.8. Variable selection by PLS pruning

An alternative method for automatically pruning variables to produce an improved model was devised and applied to the above data sets. The methodology used here was termed "PLS pruning".

After each model is formed, the user has the opportunity to mark a certain number of the variables as irrelevant. These variables will no longer be used in subsequent modelling. The variables can be marked by hand, thereby allowing "what if we only measured....?" questions to be addressed. The second harmonic predictions above are produced by this method, using prior knowledge about the effects of glucose on the harmonic patterns of yeast.

The second, and more interesting, exclusion criterion involves low weighting in the model. The weightings matrix is examined and those variables which have the smallest values are marked as excluded. The effect of this is basically to adjust the weightings to zero for irrelevant variables. Using PLS pruning, we can measure over a wide parameter space and expect to remove the effects of irrelevant variables. This handles both the problems above and should allow models to be improved greatly. It also has the great value that it requires no prior knowledge of the system under study.

As an extension to this process, we may automate the exclusion process using validation. Rather than removing a set of variables on the basis of a single model, we may perform the modelling/exclusion process a number of times and choose the model which gives the best root-mean-square error (RMSE) by, for example, cross-validation. Using this arrangement, we are searching for a minimum in a rectangular "error surface", where the two dimensions are the number of factors and number of variables. We first use cross-validation to find the appropriate number of factors using the full variable set. Having done this, we make a model and remove the least relevant variable on the basis of its weighting. This process is then iterated until one variable remains. The point in the two-dimensional space for which the calculated RMSE is minimal indicates the best number of factors and variables which should be used.

In terms of measurement, the method is the same as that used previously, except that we can now choose to measure over a much wider range of frequencies and amplitudes. Increasing the number of harmonics is likely to be of little use because we have already seen that the higher harmonics are close to the noise baseline. The pruning process is used to reduce the area of the sampling space used for model formation. Of course, once a suitable minimum error has been found, we can then examine the list of remaining variables and prune the physical sampling range accordingly. This provides the additional advantage of reducing the acquisition time.

A more detailed investigation of the use of pruning for variable selection follows. We first need to show that the pruning process does indeed provide an improvement over the arbitrary selection of variables.

The following values were calculated.

1. Root-mean-square error of prediction (RMSEP) for the best PLS model based on the full variable set.
2. RMSEP for the best PLS model for each subset of variables representing a single harmonic.

3. Minimum, mean, maximum and standard error of the RMSEP for the best PLS models for 1000 randomly selected variable subsets containing 30 variables.
4. The complete RMSEP profile for the pruning process.

These values, displayed in Table 3, will now be discussed.

We first note that the fundamental response generates a poor model, having high RMSEP and requiring a larger number of factors to generate even this model. Once again, this highlights the fact that the biological dielectric effect in the suspension is predominantly non-linear. Secondly, by using the prior information of the harmonic content, we are able to generate significantly better models by selecting "good" harmonics. By using the second harmonic alone, we can generate an improved model over that generated from the full variable set. For comparison, we consider the random selections. The average RMSEP for 1000 models was 56.2 mM. The second harmonic model performs significantly better than most random models, with the fifth being the only other which does better than the mean random selection.

In the above, we considered only particular harmonics, and we must now generalize to remove the least important variables in turn. Fig. 24 demonstrates the effects of pruning on the RMSEP for the model formed on sweep 1 predicting the data of sweep 2. As can be seen, the action of pruning provides a significant improvement in the model. The best model occurs with only five variables, giving an RMSEP of 40.9 mM from a one-factor model. The low number of variables, and the sharpness of the minimum at five, suggests that this may be an artefact local to the data sets, rather than a true minimum. A more reliable dip lies between 11 and 16 variables, this range yielding RMSEP values between 44.6 and 45.1 mM for four-factor models. These pruned models provide an improvement over both the random variable set and the full variable set. They have comparable performance to the second harmonic selection, but for their generation we did not have to rely on a prior knowledge of harmonic responses. The pruning process therefore forms a method for improving models when little is known about the effects present within the data. Examination of the remaining variables in a pruned set gives a good insight into important areas of the excitation space.

Table 3
RMSEP values for assorted variable selection methods

Model	RMSEP	Factors
Full variable set	49.13	3
Fundamental frequency	78.64	5
Second harmonic	45.39	2
Third harmonic	57.09	3
Fourth harmonic	67.87	1
Fifth harmonic	52.70	1
Random selection (min)	35.22	–
Random selection (mean)	56.24	–
Random selection (max)	119.57	–
Random selection (SD)	10.52	–

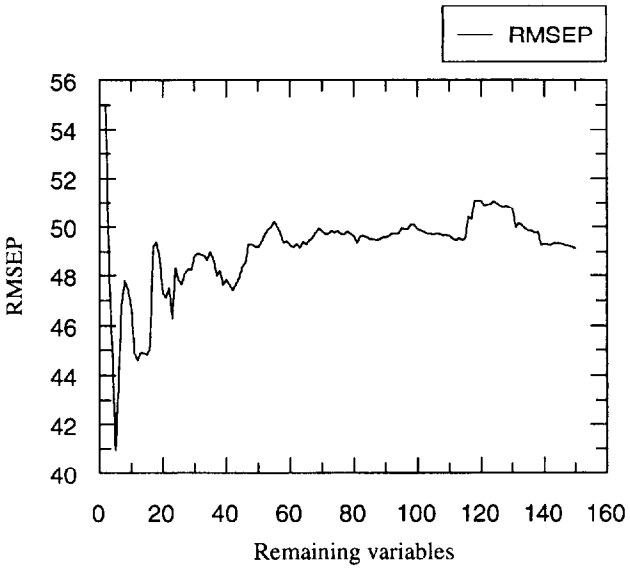

Fig. 24. Improvement in prediction error by progressive removal of least relevant x variables.

5.3. Multivariate analysis of yeast data using artificial neural networks

All neural network modelling was carried out on a commercial program called NEURALDESK (Neural Computer Sciences, Lulworth Business Centre, Nutwood Way, Totton, Southampton, Hampshire, UK), which runs under Microsoft Windows 3.1 (or Windows NT) on an IBM-compatible PC. To ensure maximum speed, an accelerator board for the PC (NeuSprint) based on the AT and T DSP32C chip, which effects a speed enhancement of some 100-fold over a 386 processor, permitting the analysis (and updating) of some 400 000 weights per second, was used.

In all the following predictions, a stochastic backpropagation algorithm was used with all NEURALDESK settings default (architecture x-8-1; learning rate, 0.1; momentum, 0.9) unless stated otherwise.

The ease of convergence of models formed on sweeps 1, 2 and 3 varies significantly for individual initial weight configurations. This suggests that the error surface is very convoluted with many local minima. Starting from a random weight initialization, the global minimum will be found in only a small proportion of model configurations or will only be found after prolonged training involving much interactive fine tuning of the net parameters. However, when the initial weights start the net training in the vicinity of the global minimum, training to the minimum RMSEP can be very rapid as shown in Fig. 25.

5.3.1. Headroom

A variety of preliminary runs showed that scaling training and test inputs and, particularly outputs, assists predictions by reducing the squashing effect of the sigmoid on the hidden nodes in order to prevent rapid saturation of

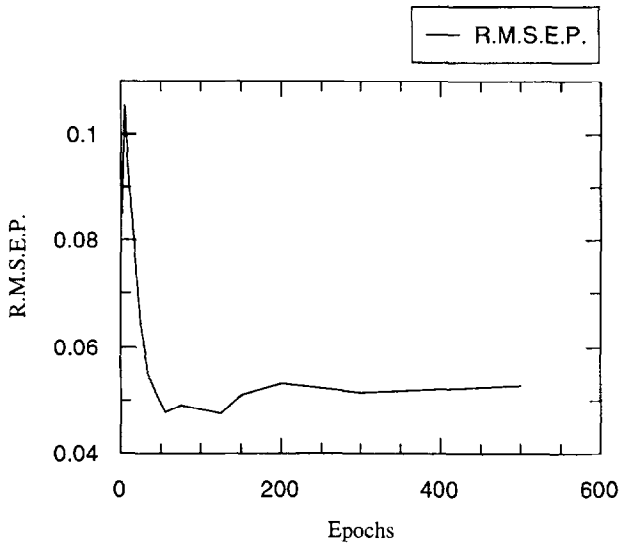

Fig. 25. Root-mean-square error of prediction of the test data as a function of the number of training epochs with the net updated at each epoch. The net parameters and data headroom are the defaults oulined in the text. Optimum training is achieved here at 50 epochs with overfitting slowly occurring above this. The minimum can occur at as few as 10 epochs if the initial weights are fortuitously close to the global minimum.

hidden nodes, and by simulating linear output nodes which are found to be beneficial over sigmoid output nodes, particularly in the case of extrapolate prediction or where the training set incompletely covers the y data interval. The entire x matrix may be scaled simultaneously, but we have found [142] that this often conceals small but significant variables in favour of large but less significant ones, and that better modelling is achieved with each x variable scaled individually.

Lowering the input scaling between 0.2–0.8 to 0.4–0.6 has no significant effect on the predictions, but scaling inputs outside the range 0.2–0.8 degrade the predictions. Scaling outputs outside 0.4–0.6 are similarly detrimental. At these headroom levels, the output nodes are approximately linear. The situation in which no headroom is included, such that both x and y variables are scaled between zero and unity, is shown in Fig. 26.

Table 4
Importance of specific harmonics to the neural network model

	First	Second	Third	Fourth	Fifth
Mean harmonic weighting	0.455	0.47	0.51	0.435	0.4

We therefore found it appropriate for subsequent predictions to use x variables scaled to the range 0.2–0.8 and y data scaled to the range 0.4–0.6 for training and predictions, then rescaled to the correct values for plotting the results.

A model formed on sweep 1, predicting sweep 2 using raw data with no median averaging or any other preprocessing, gives Fig. 27. This predicts the general features of the glucose concentration vs. time relation, including the leading zeros, but is very noisy.

5.3.2. Median averaging

As observed in PLS models, median averaging the variables cleans up and improves the finite-glucose readings drastically, but has a tendency to smear out the leading zeros so that they are less well predicted, as indicated in Fig. 28.

5.3.3. Variable selection

One way to determine the importance of individual inputs in affecting the outputs of a neural net model is to apply a unit diagonal test matrix to the trained net, with each column being applied seriatim. Such a procedure shows that the third harmonic variables contribute more to (give a higher output in) the model formed on this particular data set than second than first than fourth and fifth (Table 4), although there is not a large difference in significance.

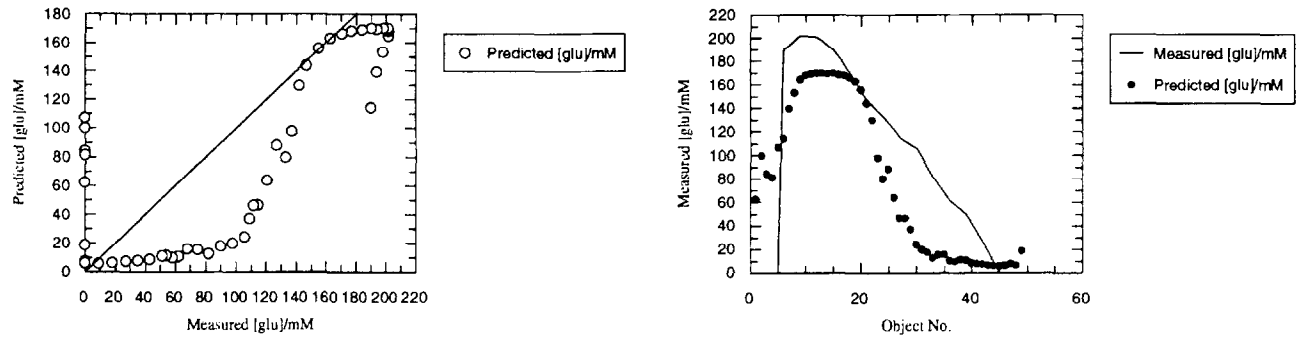

Fig. 26. Neural network prediction of glucose concentration with no headroom on either x or y variables, i.e. all variables were individually normalized in the range 0–1. The optimum training occurred at 70 epochs with a training error of 0.025. It can be seen that the predictions tend to reflect the sigmoidal nature of the nodes as much as the relation between the data. In order to show the effect more clearly, the data used here were median averaged and should be compared with those in Fig. 25.

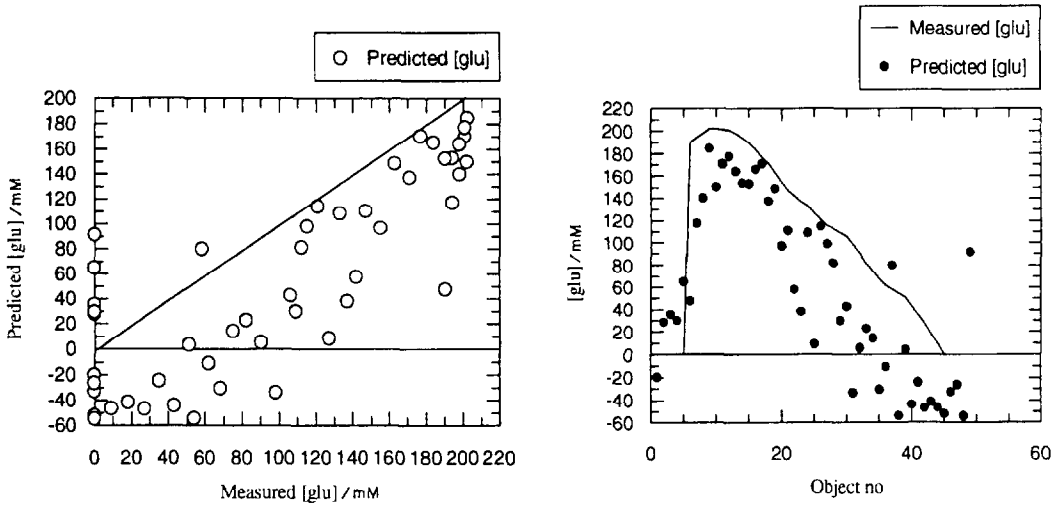

Fig. 27. Neural network prediction of glucose levels in yeast cell suspensions between separate data runs using the raw data with no preprocessing other than normalization and headroom scaling. The optimum training occurred at 36 epochs at a training error of 0.02.

Pruning the input variables to include only second and third harmonics smooths out the predictions a little further, but begins to make them more inaccurate, showing that most but not all the data relating to the neural net's ability to model [glu] are in the second and third harmonics. Training and predicting only on the second harmonic reduces the performance of the model as shown in Fig. 29, in contrast with the finding with PLS (see above), so that the full data set was used for subsequent predictions. Neural networks appear to be more robust to spurious modelling of uncorrelated data than PLS and so require less rigour in the choice of the most relevant variables.

5.3.4. *The leading zeros problem*

It is reasonable to assume that, biochemically, the yeast's resting state before the addition of glucose will not necessarily be identical to the resting state to which it returns after the glucose is used up, since storage polymers, such as glycogen and trehalose, will have been formed as a result of glucose metabolism [143–145]. On this basis, it may be expected that the leading zeros will be predicted less well if the model forms predominantly on the much larger section of finite- and post-glucose data. The multivariate method used has, in effect, to model two separate systems. However, the prediction to some degree of the leading zeros is a vital check that the model is not merely forming on drifts and trends in the data, since data reflecting glucose utilization are monotonic. If modelling were to occur merely on the basis of a trend, the leading zeros would be predicted to the same absolute levels as the initial glucose concentration, since the model would see them as identical to the high glucose readings if merely modelling a trend in the data as the x variables would be identical for samples taken both before and immediately after the addition of glucose. Therefore the prediction of the leading zeros acts as a marker that the model is actually forming on a glucose-related response, and not merely on unconnected coincidental trends. Accordingly, it is always imperative to display both the predicted vs. measured plots and the predicted vs. object number plots in order to determine the degree of fit of the leading zeros [130].

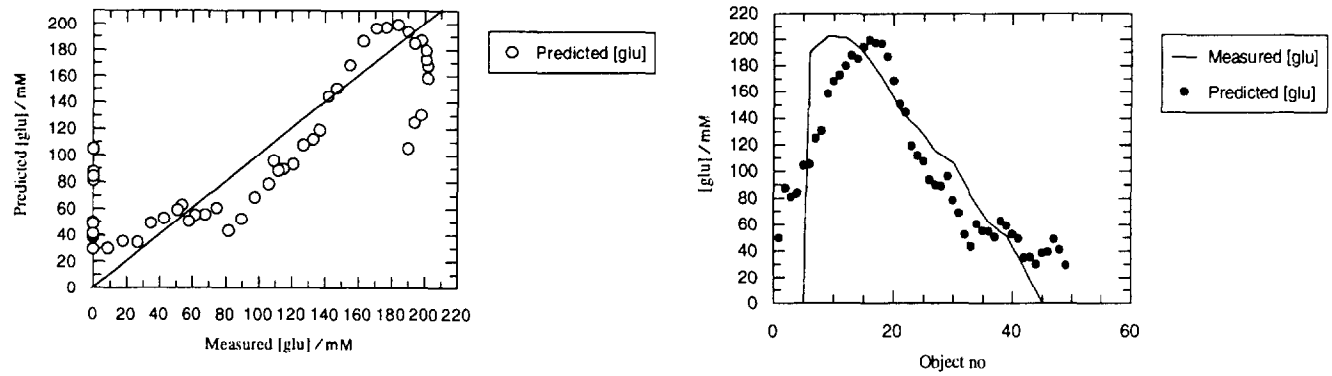

Fig. 28. Improvement in neural network prediction produced by median averaging the data of Fig. 27 after recording. The optimum training occurred at 25 epochs at a training error of 0.02.

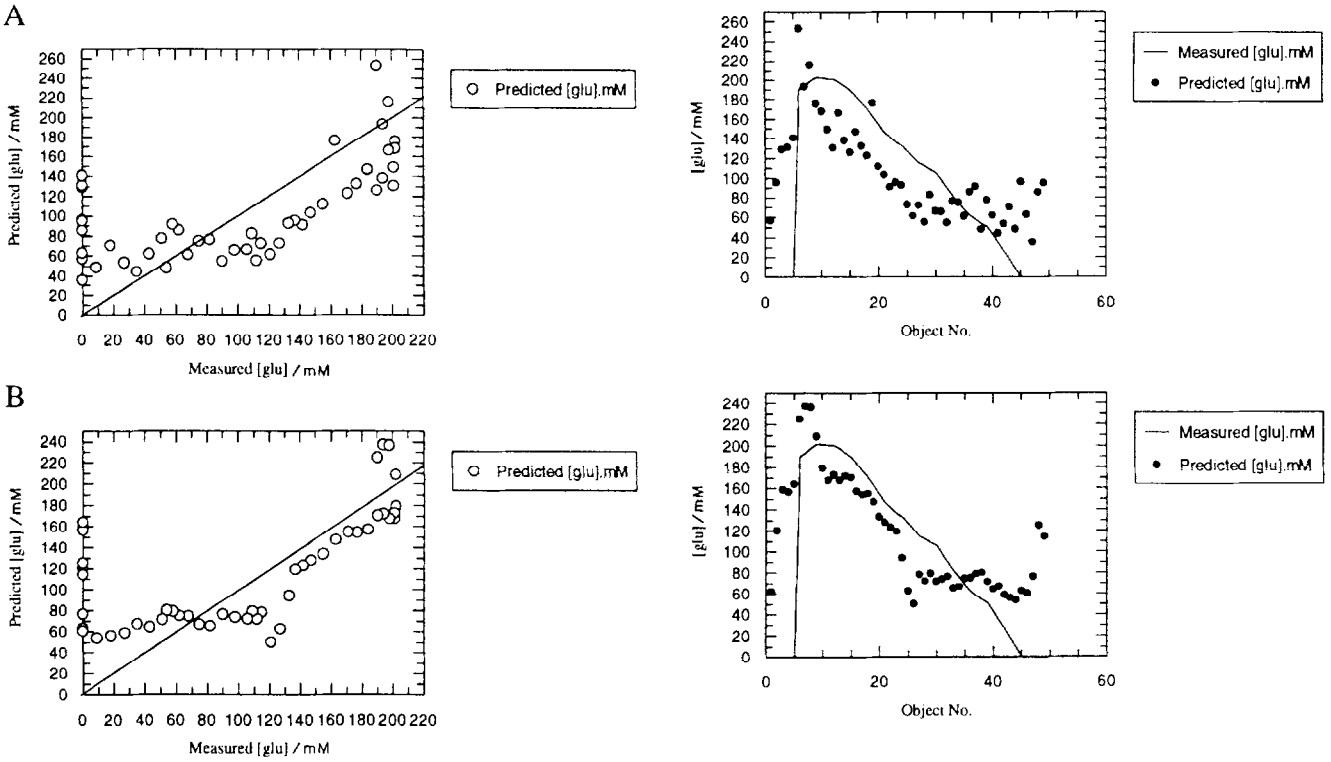

Fig. 29. The effect of variable selection (in this case, use of only the second harmonic) on the quality of the neural network model: (A) raw data; (B) median-averaged data. The optimum training occurred at 60 epochs at a training error of 0.02.

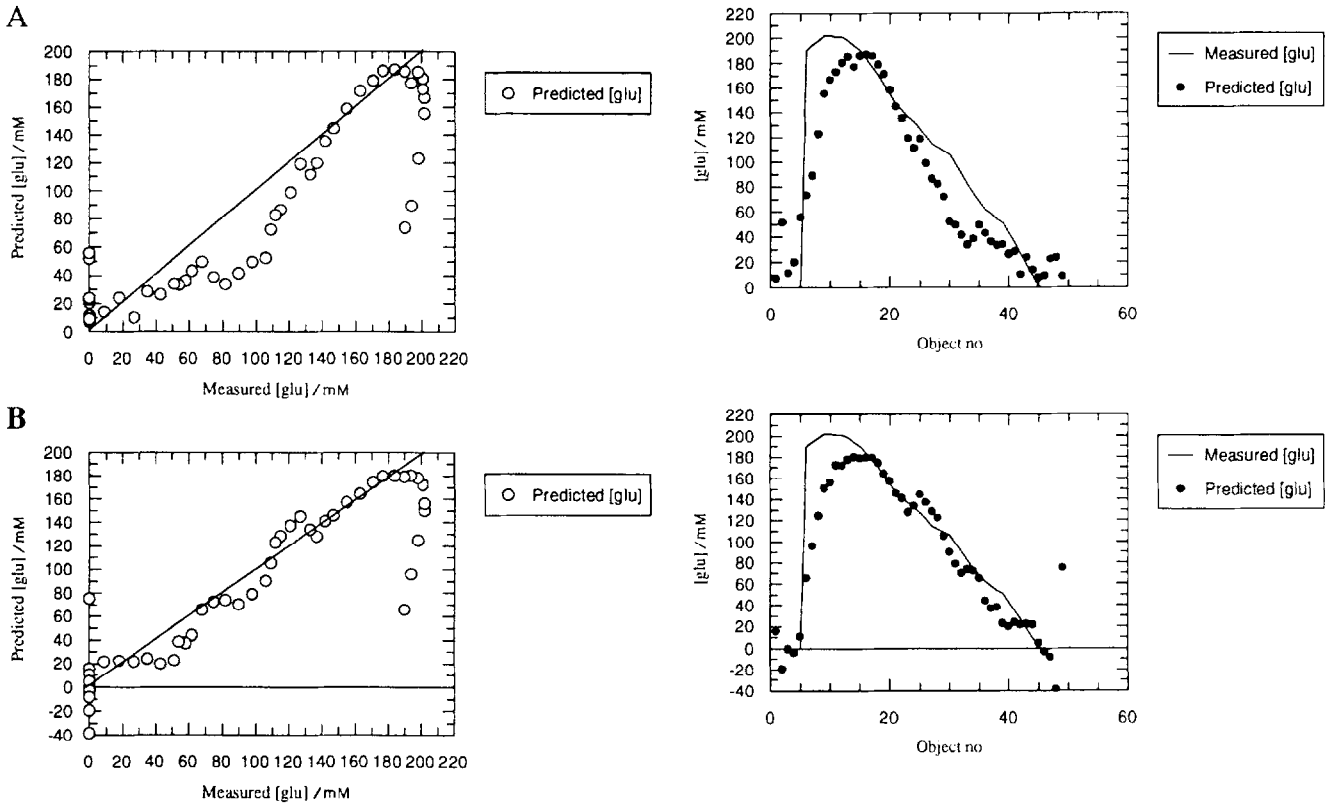

Fig. 30. Improvement in modelling of leading zeros. (A) Non-median-averaged leading zeros. The optimum training occurred at 55 epochs at a training error of 0.015. (B) Non-median-averaged leading zeros repeated for a similar number of instances per epoch as those from finite-glucose measurements. The optimum training occurred at nine epochs at a training error of 0.02.

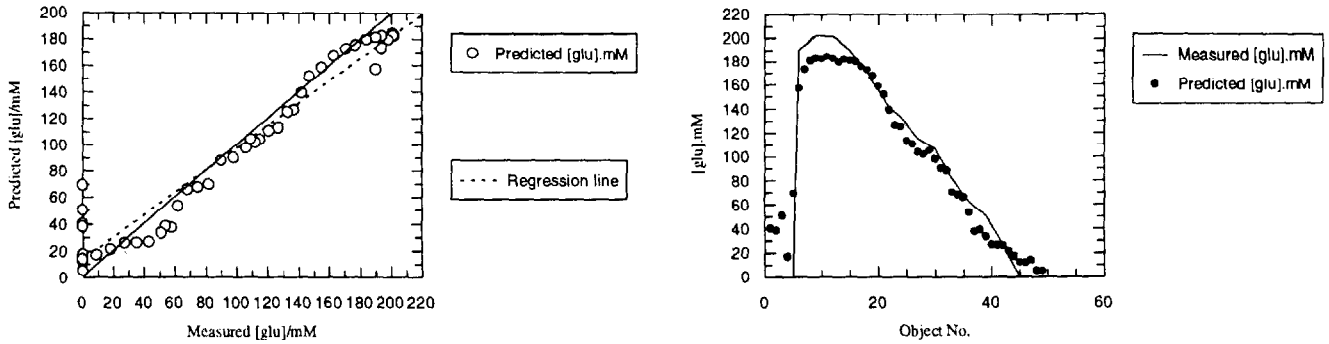

Fig. 31. Prediction of sweep 2 by sweep 1. Five time index variables were added to the x data. The leading zeros were non-median averaged as in Fig. 30(A), but not repeated to form a greater data length as in Fig. 30(B). The optimum training occurred at 30 epochs at a training error of 0.02. Again, the regression line is included for comparative purposes. This has a slope of 0.84 and an intercept of 13.5. The correlation coefficient is 0.97.

Exploiting this argument, it is indeed found that substituting the non-median-averaged leading zeros into the median-averaged data gives the advantage of both well-predicted (if noisier) leading zeros and a closely fitted finite-glucose curve as shown in Fig. 30(A).

Increasing the number of leading zeros, by direct copying, until there are roughly as many as non-zero readings, so that the net sees the zeros as often as the non-zeros during each training epoch, improves the predictions substantially as can be seen in Fig. 30(B). All the points are well modelled except for the few after the addition of glucose, when the yeast could not be expected to respond instantly anyway. It is known that there is a significant phase of activation of H^+-ATPase following the addition of glucose to a resting cell suspension [146,147].

5.3.5. PLS scores as inputs

In weakly non-linear problems, it is possible to speed up the learning of a net by forming a PLS model of the process under study, and using the PLS scores as the inputs to the net. This allows a substantial dimensional reduction in the data, so that the net has to train many fewer weights. On our data, we found that this procedure did not improve the precision of prediction, but did allow slightly faster training in models which converged slowly. Most of our models on sweeps 1–3, however, converged quickly with the raw data as inputs.

5.3.6. Time indices

Time indices added as extra x variables for the same reasons as outlined above (five added to end of 150 variables in the data of Fig. 30(A)), in order to allow the net to model a drift in baseline, made the further improvement to the predictions indicated in Fig. 31. With the assistance of this technique, the prediction of sweep 2 by sweep 1 is now within the error bounds of the reference readings except for the leading zeros (which can be improved by selectively increasing the number of training epochs for these zeros as outlined above).

5.3.7. Architecture optimization

Using different training algorithms on this final configuration of data had little effect on the effectiveness of modelling. Standard backpropagation is much the same as stochastic backpropagation. Weigend weight eliminator and skeletonizing backpropagation both make negligible improvement, as might be expected from the similarity of the significance values for individual harmonics shown in Table 4.

If the neural network has too many hidden nodes, it can overfit the training set, resulting in poor generalization [148]. Ideally, the net should use the minimum number of hidden nodes in the minimum number of hidden layers which allow modelling to the required precision. Building a net from scratch, modelling to the optimum RMSEP with no hidden nodes, then with one hidden node and then two

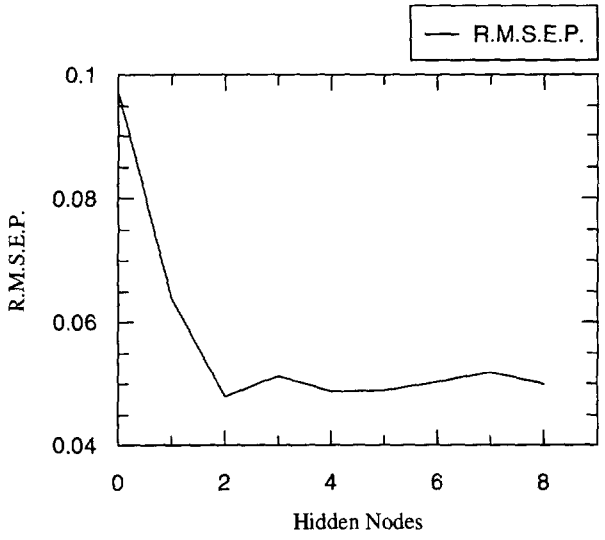

Fig. 32. The effect of building a neural net from scratch. The optimum prediction error is obtained with two hidden nodes. There is no significant degradation or improvement in exceeding this up to the NEURALDESK default architecture of eight hidden nodes.

nodes etc. gives no further improvement after two nodes as shown in Fig. 32.

The model formed with two nodes produces predictions essentially identical to those with the default eight nodes; therefore it appears that, for these predictions, overfitting is not a serious problem, but no benefit is obtained by using more than two hidden nodes. This intrinsic dimensionality indicated by the net reassuringly coincides well with that indicated by the optimum number of factors in the equivalent PLS predictions.

We hypothesize that the intrinsic dimensionality of this system is the number of macrostates (or electrically active transitions between them) that the enzyme adopts in its multistate operation as outlined above.

Finally, with the present data, the use of two or more hidden layers did not improve the efficacy of the model, consistent with the view that the intrinsic dimensionality of the nonlinear dielectric response is indeed small.

5.4. Predictions on suspension files alone

The real advantage to multivariate modelling is the possibility that the modelling procedure may be able to pick out changes in the suspension spectra that are related to glucose (or other metabolites or even inhibitors), without the need to record a supernatant spectrum, en route to forming difference spectra. Indeed, in many (most?) practical situations for measurements in situ, it is inconvenient or impossible to take a reference reading.

To this end, we studied sweeps 1–3 as described above, but using only the suspension data from the experiments, rather than deconvolving the supernatant data to produce difference spectra. Suspension files, when leading zeros are repeated, as in Fig. 29(B), give very similar predictions to those obtained from difference files (with a small gain and d.c. offset due to the slight change in the electrode polarization between the two data runs), removing the need for a supernatant reading where this is difficult to obtain (Fig. 33), suggesting that neural networks are perfectly capable of deconvolving what can be large electrode-generated harmonics from the small biological effects, as long as the former are reproducible. Addition of time indices did not make any significant improvement to this prediction.

5.4.1. Fermenter experiments

The simple batch experiments carried out above were performed on resting cell suspensions in vitro. We therefore performed a yeast fermentation, as described in batch form in Ref. [149], and the above methods were used to record and analyse the non-linear dielectric spectral sweeps in conjunction with a number of more conventional biological measurements of interest in fermentations generally.

A yeast culture was grown overnight and added to a nutrient solution containing nutrients and glucose at approximately 300 mM. The culture was agitated throughout the experiment. At intervals of just over 1 h, a sample of the culture was extracted from the fermenter and a number of biological variables were measured by conventional methods. Reflolux measurements of glucose concentration were recorded approximately every 20 min.

The acquisition software was configured to average 100 power spectra at each point in a logarithmic frequency range of 10–1000 Hz in 30 steps, and a linear voltage range of 0.5–1.5 V in six steps, recording the first five harmonics at each voltage/frequency combination. Sampling took place continuously, with each frequency/voltage sweep being immediately followed by the next. The overall time for each sweep was 157 s.

The experiment continued until all the glucose had been used and the biological variables indicated that the culture was beginning to change its state. This process took around 10 h.

The following biological variables were measured: glucose concentration, pH, percentage viability, percentage budding, low-frequency conductance of the culture (G), wet weight, dry weight, cell count and peak channel number (PCN) (a measurement of the modal cell size obtained by flow cytometry) [150].

The number of nonlinear biological dielectric spectroscopy samples far exceeded that of the biological variables, by virtue of the fast acquisition cycle for nonlinear

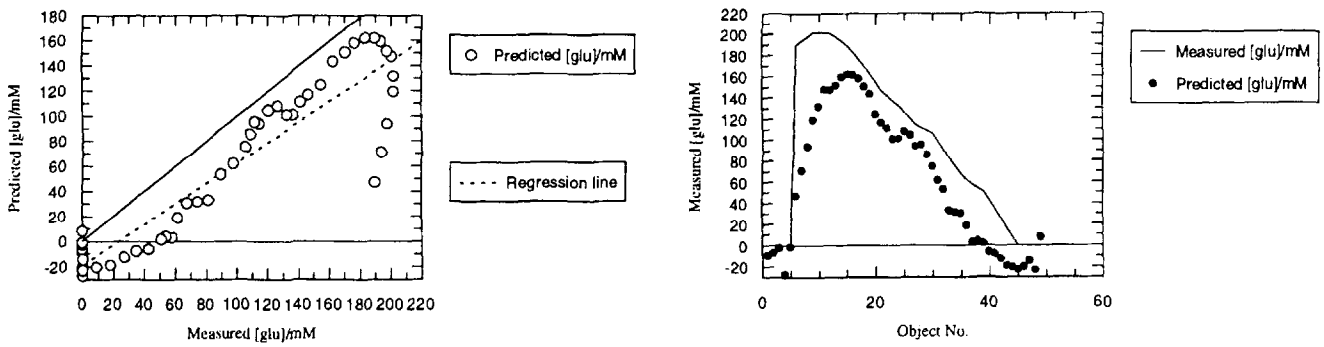

Fig. 33. Neural network prediction using suspension data. The data were median averaged, but the leading zeros were not, and these zeros were extended as in Fig. 30(B). The optimum training occurred at 59 epochs at a training error of 0.15. The regression line is included for comparative purposes. This has a slope of 0.81 and an intercept of −18.8. The correlation coefficient is 0.93.

biological dielectric spectroscopy. The great majority of nonlinear biological dielectric spectroscopy samples did not, therefore, have corresponding biological values and could not be used for modelling. These points were used to test the models. The predictions for these points were calculated and plotted against time on the same graph as the corresponding biological variables. The curves obtained gave an indication of the performance of the model for the unmeasured points, on the basis that the biological variations should occur smoothly.

For this experiment, it was appropriate to use PLS2 [63] to model a number of y variables simultaneously. The results of a two-factor model of all variables are now presented. Fig. 34 shows the predictions based on a two-factor PLS2 model formed on the nine sample points. The model thus generated was used to predict all 234 points.

The model appears to generate reasonable predictions for glucose concentration, pH and cell count (wet and dry weight being similar measures to the cell count and producing essentially identical predictions), in that these show good interpolations between the training points. The model performs less well in terms of precision for percentage budding and "G out", but follows the form of the measured curve quite closely.

However, this approach does not produce a usable model for PCN or percentage viability (not shown), since

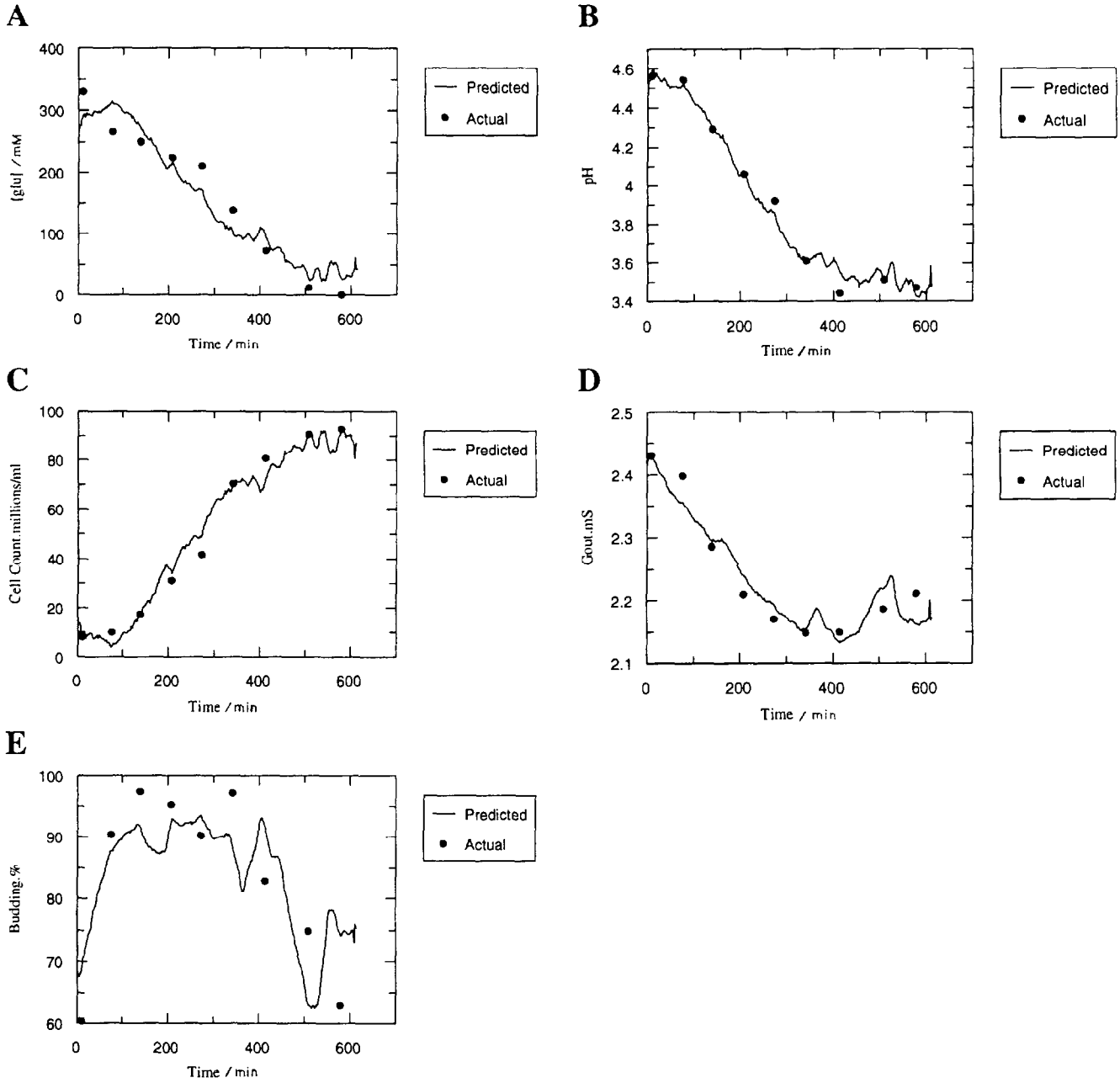

Fig. 34. Two-factor PLS2 predictions of various parameters measured during fermentation: (A) extracellular glucose concentration; (B) pH; (C) cell count; (D) extracellular conductivity; (E) percentage of cells budding.

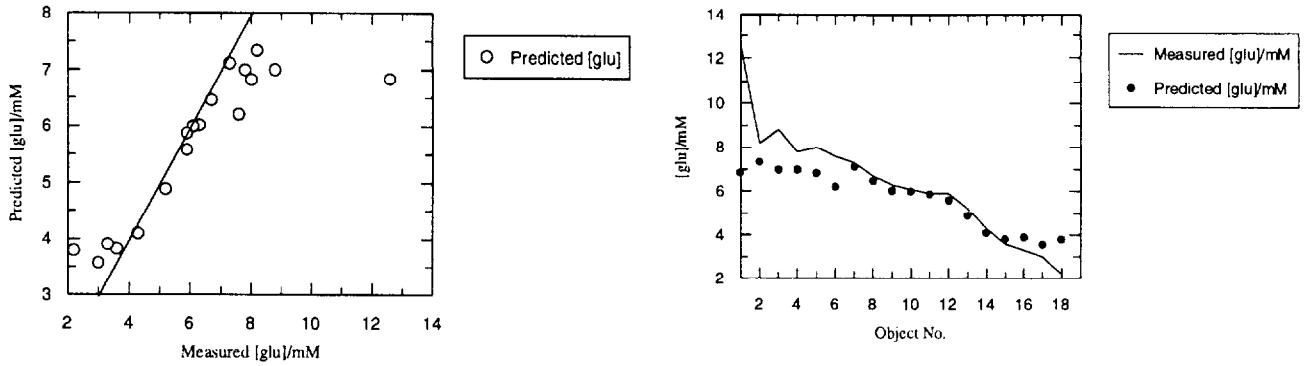

Fig. 35. Prediction, within a single data set, of even-numbered samples by odd-numbered samples for glucose concentration in sheep's blood. The data cover a time interval of 7 h. The optimum prediction is obtained with two PLS factors.

both are measured to be roughly constant throughout the experiment, with small variations not simply related to any of the other parameters measured and hence more difficult for PLS2 to model. Overall, however, it may be observed that the non-linear dielectric approach provides a novel, rapid and non-invasive method for the measurement of biological properties in fermentations in situ.

Fermentations involve measurements in difficult biological media containing cells and, having shown that these could be accomplished successfully, it was of interest to determine whether non-linear dielectric spectroscopy could be used to determine glucose levels in blood. To this end, a short series of experiments was carried out on sheep's blood.

5.4.2. Sheep's blood in vitro

The blood, drawn from a sheep which had not received food overnight, was placed in a heparinizing vial and carried back to the laboratory. Preliminary experiments showed that raw blood proteins contaminated the electrode surfaces immediately and prevented repeatable results. Hence the blood was spun down and resuspended three times, in a buffer consisting of 150 mM NaCl and 20 mM KH_2PO_4 adjusted to pH 7.3, in order to remove these plasma proteins. The blood was ready for experimentation within 1 h of being drawn from the sheep.

The resting concentration of glucose in sheep's blood is typically 3 mM or less, and this forms the baseline level for these experiments.

Glucose (10 mM) was added to the washed resting blood and data sweeps were taken at 15 min intervals over 7 h (typical time for the blood glucose concentration to return to baseline). Suspension spectra only were recorded with no supernatant references. The data sweeps covered five voltages equidistantly spaced throughout the voltage range from 0.5 to 1.5 V and nine frequencies logarithmically spaced across the frequency range from 10 Hz to 1 kHz. The spectra were calculated using mean averaging over 30 blocks of data per voltage/frequency combination.

PLS modelling on the odd samples of any of these data sweeps gives good predictions of the even samples with two factors, as shown in Fig. 35. This is a good test of whether there would be a modellable effect in the absence of electrode polarization variation, since both the modelled and predicted samples contain similar information on any variation which may have occurred during the experiment. The anomalous prediction of the even sample at 12.3 mM measured [glu] is due to this reading being outside the range of [glu] in the odd samples on which the model was made. PLS is poor at extrapolation.

Good straight line predictions between runs performed on different batches of blood taken from the same sheep

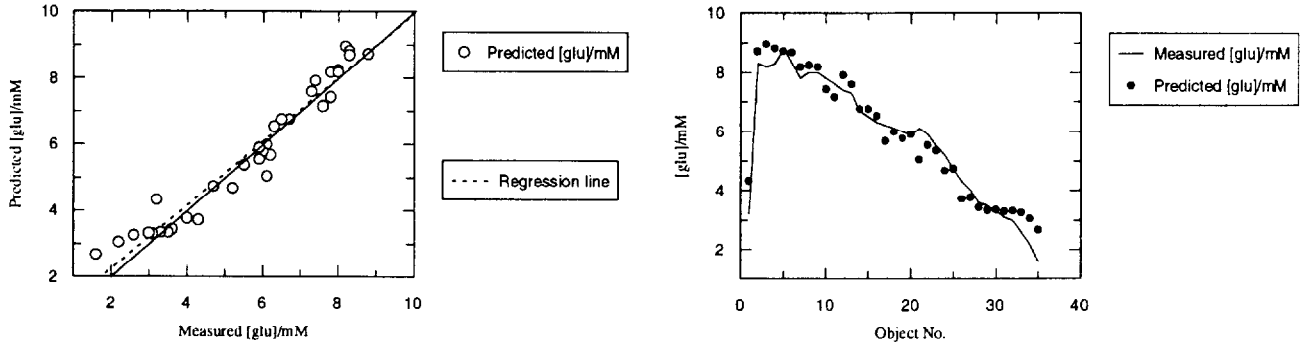

Fig. 36. Prediction between separate data sets. Data cover a time interval of 7 h, and the optimum prediction requires three PLS factors. Again, the regression line is included for comparative purposes. This has a slope of 0.96 and an intercept of 0.32. The correlation coefficient is 0.97.

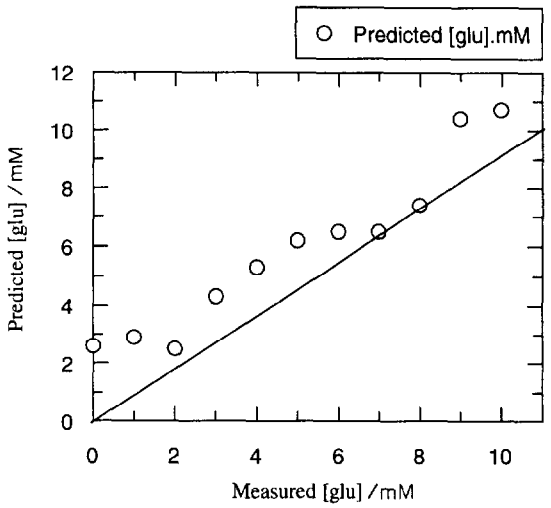

Fig. 37. PLS prediction between two titrations of glucose in the supernatant. The optimum model required eight PLS factors indicating a weak relationship.

on different days can be obtained with attention to outliers and removing from the data sets all points which are not within the ranges of both modelling and prediction data sets (Fig. 36), but they typically include gain/d.c. anomalies due to the sensitivity (caused by electrode variations) of PLS models to baseline differences between runs.

Predictions between different data sets are more variable than those using yeast. This is probably due to a greater tendency to electrode fouling even in washed erythrocytes, leading to a lesser degree of reproducibility of polarization conditions between otherwise identical experiments. Also, the inhibitor studies necessary to identify the active enzyme(s) have not yet been carried out, and the experimental conditions producing the optimum and most stable operating conditions for this enzyme(s) are indeterminate.

It is necessary to check that the models are forming on real effects rather than merely chance correlations in random noise, and that the predictions depend on the presence

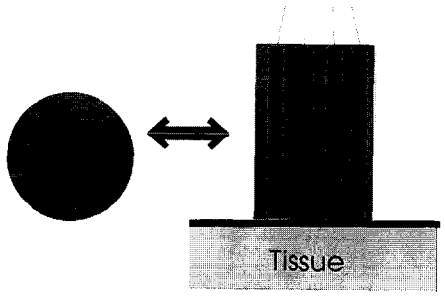

Fig. 38. A probe with flush gold electrodes used to record non-invasive non-linear dielectric spectra in vivo.

of erythrocytes. Data sets containing computer-generated random numbers, with the same number of objects and variables as the real data, produce random predictions, showing that modelling on chance correlations is not a problem for this size of data set.

Glucose control spectra taken by adding glucose to supernatants cannot be predicted from sheep's blood models, or vice versa; therefore the predicted relations shown above between sheep's blood data sets depend on the presence of erythrocytes. However, glucose controls can predict other glucose controls with typically eight PLS factors, showing that glucose can affect electrode surfaces in a modellable way. However, the predictions require many PLS factors, denoting only a weak effect, whereas the predictions in the presence of erythrocytes require only a few PLS factors, showing that these predictions are formed predominantly on the reaction of the erythrocytes to the presence of glucose. Fig. 37 shows the best prediction pair out of six combinations of three control data sets. Other control pairs show poorer modelling.

5.4.3. Human blood in vivo

An electrode probe with flush electrodes in the configuration of Fig. 38 was matched to the subject's forearm with saline making sure that the position and orientation were the same for each data sweep, these data sweeps being

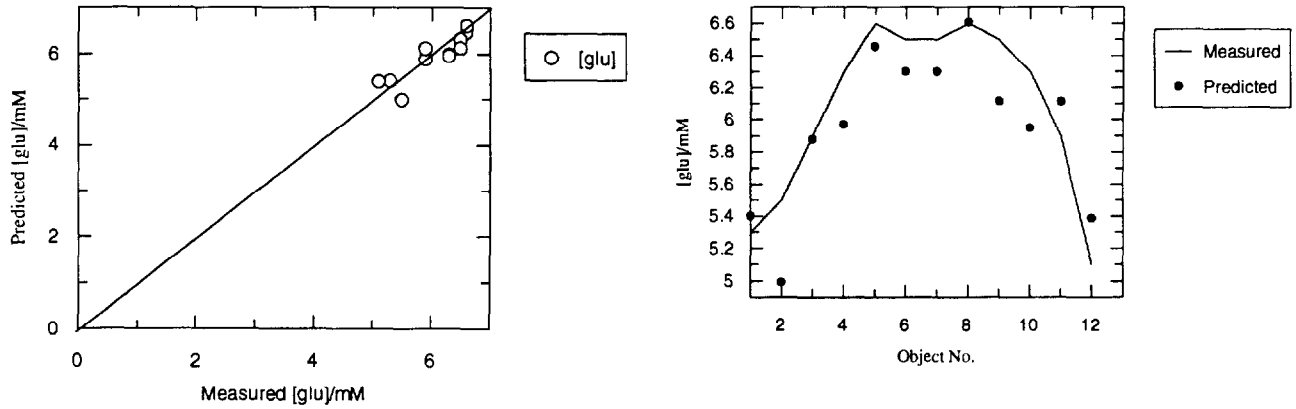

Fig. 39. Prediction, within a single data set, of even-numbered samples by odd-numbered samples. The data cover a time interval of 30 min. The optimum prediction is obtained with two PLS factors.

identical to those described above for sheep's blood in vitro.

The 64 kg subject, having starved for 16 h to allow blood glucose to reach a baseline level, ingested 43 g of glucose; the blood glucose level was measured as a function of time with the Reflolux simultaneously with dielectric sweeps. Reference sweeps were not possible, and so the spectra recorded were the single-sided spectra equivalent to the suspension sweeps in sheep's blood and yeast above. This subject metabolized glucose rather rapidly, so that the range of glucose concentrations is not as great as expected.

Modelling on the odd-numbered samples of any of these data sweeps gave good predictions of the even-numbered samples with two PLS factors, as in Fig. 39.

The absolute blood glucose level is harder to reproduce precisely with a live subject but, when two data sweeps are used, by modelling and predicting only on those points within the glucose concentration range of both sweeps, the predictions between the two sets can be good, although not as repeatable as the yeast predictions or sheep's blood predictions. Fig. 40 shows one of the more satisfactory results.

Models and predictions tend to be very noisy, and show a poor repeatability from run to run. Gain and d.c. offsets are common. However, the possibility of obtaining occasional good predictions suggests a viable method if the causes of both systematic and random errors can be removed. It is felt that much of this problem can be laid at the door of the changing electrode surface and polarization state, and the non-identical nature of matching between electrode and skin, from run to run and even from data point to data point. If this problem can be solved, predictions should improve markedly. This conclusion is strongly reinforced by the consistently good predictions of data points within a single run by models formed on other (independent) data points within that run (odds vs. evens), since PLS has the same electrode fluctuation data for both model and prediction, and seems to be able to cope well with separating this from the glucose concentration data.

6. Conclusions

Individual non-linear dielectric signatures and activity markers correlating with the known biochemistry of the organism have been discovered in yeast cells, plant cells, bacteria and erythrocytes, suggesting a wide range of applicability of the technique both for the identification of an organism and the determination of its metabolic state.

The ability to detect the action/inaction of a membrane-bound enzyme allows the use of whole-cell biosensors for quick and easy detection or assay of substances affecting the operation of these enzymes. For instance, toxic chemicals in drinking water could be detected by their inhibitory effect on the metabolism of the cells in the sensor.

The method lends itself to the analysis of any whole-cell system containing substrates, products or inhibitors of metabolism, and assays can be carried out non-invasively on the specimen in situ. The performance of fermentations and the measurement of glucose levels in cell substrates by its effect on cell metabolism are examples which have been illustrated. A novelty of the approach is that the method interrogates the biology directly to give an account of what it is seeing metabolically.

The hardware required to achieve this is simple, cheap and robust, most of the complexity being in the controlling and analysing software, and the process can be sufficiently rapid to allow real-time control applications in, for example, fermentation processes. Improvements in electrode performance will continue to assist the further development of non-linear dielectric spectroscopy in bioelectrochemical systems.

Acknowledgements

We thank the Chemicals and Pharmaceuticals Directorate of the UK BBSRC for financial support of this work, the EPSRC for a studentship to A.J., the Overseas Research Studentship Scheme for support of X-z.Z., Andy

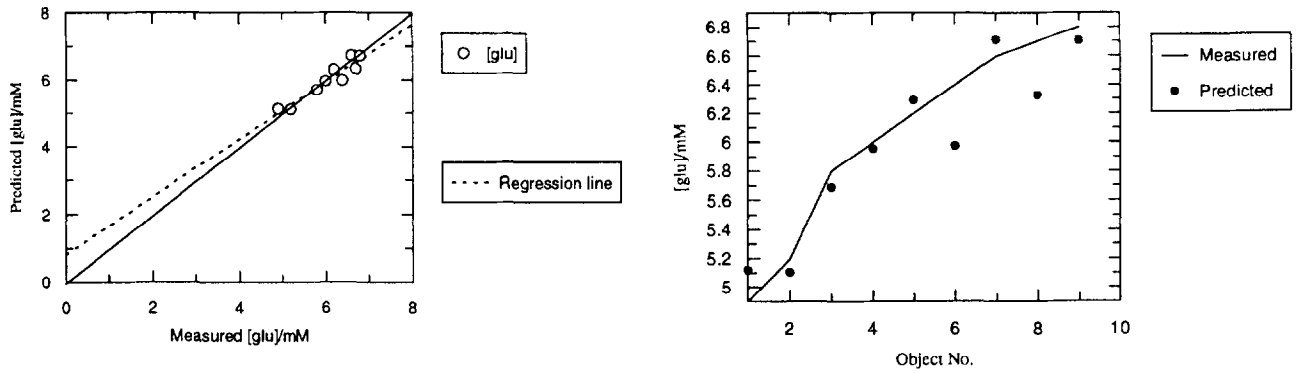

Fig. 40. Prediction between separate data sets. Data cover a time interval of 20 min, and the optimum prediction requires two PLS factors. Again, the regression line is included for comparative purposes. This has a slope of 0.85 and an intercept of 0.82. The correlation coefficient is 0.95.

McShea for assistance with electrode measurements, Brian Davies (Peithyll) for the provision of sheep's blood and Drs. Hazel and Chris Davey for assistance with the fermenter experiments.

References

[1] U. Zimmermann, Biochim. Biophys. Acta, 694 (1982) 227.
[2] T.Y. Tsong and R.D. Astumian, Bioelectrochem. Bioenerg., 15 (1986) 457.
[3] K. Hayashi and N. Sakamoto, Dynamic Analysis of Enzyme Systems, Springer-Verlag, New York, 1986.
[4] E.H. Grant, G.P. South and R.J. Sheppard, Dielectric Properties of Biological Molecules in Solution, Oxford University Press, Oxford, 1978.
[5] O. Schanne and E.R.P. Ceretti, Impedance Measurements in Biological Cells, Wiley, Chichester, 1978.
[6] R. Pethig, Dielectric and Electronic Properties of Biological Materials, Wiley, Chichester, 1979.
[7] D.B. Kell and C.M. Harris, J. Bioelectricity, 4 (1985) 317.
[8] K.R. Foster and H.P. Schwan, in C. Polk and E. Postow (Eds.), CRC Handbook of Biological Effects of Electromagnetic Fields, CRC Press, Boca Raton, 1986, p. 27.
[9] D.B. Kell, in A.P.F. Turner, I. Karube and G.S. Wilson (Eds.), Biosensors; Fundamentals and Applications, Oxford University Press, Oxford, 1987, p. 428.
[10] R. Pethig and D.B. Kell, Phys. Med. Biol., 32 (1987) 933.
[11] K.R. Foster and H.P. Schwan, CRC Crit. Rev. Biomed. Eng., 17 (1989) 25.
[12] C.L. Davey and D.B. Kell, in R. Paris (Ed.), Electric Field Phenomena in Biological Systems, IOP Short Meetings Series 21, Institute of Physics, London, 1989, p. 51.
[13] S. Takashima, Electrical Properties of Biopolymers and Membranes, Adam Hilger, Bristol, 1989.
[14] C.L. Davey and D.B. Kell, in M. O'Connor, R.H.C. Bentall and J.S. Monahan (Ed.), Emerging Electromagnetic Medicine, Springer, Berlin, 1990, p. 19.
[15] D.B. Kell and C.L. Davey, in A.E.G. Cass (Ed.), Biosensors; A Practical Approach, Oxford University Press, Oxford, 1990.
[16] C.L. Davey and D.B. Kell, The low-frequency dielectric properties of biological cells, in D. Walz, H. Berg and G. Milazzo (Eds.), Bioelectrochemistry of Cells and Tissues, Birkhäuser, Zürich, 1995, pp. 156–207.
[17] P. Debye, Polar Molecules, Dover Press, New York, 1929.
[18] L.J. De Felice, W.J. Adelman, Jr., D.E. Clapham and A. Mauro, in W.J. Adelman, Jr. and D.E. Goldman (Eds.), The Biophysical Approach to Excitable Membranes, Plenum, New York, 1981, p. 37.
[19] S. Miyamoto and H.M. Fishman, IEEE Trans. Biomed. Eng., BME-33 (1986) 644.
[20] W. Carius, J. Colloid Interface Sci., 57 (1976) 301.
[21] V.S. Sokolov and V.G. Kuz'min, Biofizika, 25 (1980) 170.
[22] E.H. Sepersu and T.Y. Tsong, J. Membr. Biol., 74 (1983) 191.
[23] E.H. Sepersu and T.Y. Tsong, J. Biol. Chem., 259 (1984) 7155.
[24] D.S. Liu, R.D. Astumian and T.Y. Tsong, J. Biol. Chem., 265 (1990) 7260.
[25] D.B. Kell, Protonmotive energy-transducing systems: some physical principles and experimental approaches, in C.J. Anthony (Ed.), Bacterial Energy Transduction, Academic Press, London, 1988, p. 429.
[26] H.V. Westerhoff, R.D. Astumian and D.B. Kell, Ferroelectrics, 86 (1988) 79.
[27] D.B. Kell, R.D. Astumian and H.V. Westerhoff, Ferroelectrics, 86 (1988) 59.
[28] D.B. Kell, ISI Atlas Sci., Biochemistry, 1 (1988) 25.
[29] D.B. Kell, Coherent properties of energy-coupling membrane systems, in H. Frohlich (Ed.), Biological Coherence and Response to External Stimuli, Springer, Heidelberg, 1988, p. 233.
[30] H.V. Westerhoff, T.Y. Tsong, P.B. Chock, Y. Chen and R.D. Astumian, Proc. Natl. Acad. Sci. USA, 83 (1986) 4734.
[31] R.D. Astumian, P.B. Chock, T.Y. Tsong, Y. Chen and H.V. Westerhoff, Proc. Natl. Acad. Sci. USA, 84 (1987) 434.
[32] T.Y. Tsong and R.D. Astumian, Progr. Biophys. Mol. Biol., 50 (1987) 1.
[33] T.Y. Tsong and R.D. Astumian, Annu. Rev. Physiol., 50 (1988) 273.
[34] R.D. Astumian and B. Robertson, J. Chem. Phys., 91 (1989) 4891.
[35] R.D. Astumian, P.B. Chock, T.Y. Tsong and H.V. Westerhoff, Phys. Rev. A, 39 (1989) 6416.
[36] T.Y. Tsong, D.-S. Liu, F. Chauvin, A. Gaigalas and R.D. Astumian, Bioelectrochem. Bioenerg., 21 (1989) 319.
[37] B. Robertson and R.D. Astumian, Biophys. J., 57 (1990) 689.
[38] G.R. Welch and D.B. Kell, Not just catalysts: the bioenergetics of molecular machines, in G.R. Welch (Ed.), The Fluctuating Enzyme, Wiley, New York, 1986, pp. 451–492.
[39] A.M. Woodward and D.B. Kell, Bioelectrochem. Bioenerg., 24 (1990) 83.
[40] A.M. Woodward and D.B. Kell, Bioelectrochem. Bioenerg., 25 (1991) 395.
[41] A.M. Woodward and D.B. Kell, FEMS Microbiol. Lett., 84 (1991) 91.
[42] A.M. Woodward and D.B. Kell, Bioelectrochem. Bioenerg., 26 (1991) 423.
[43] D.B. Kell and A.M. Woodward, Anal. Proc., 28 (1991) 378.
[44] A. McShea, A.M. Woodward and D.B. Kell, Bioelectrochem. Bioenerg., 29 (1992) 205.
[45] A.M. Woodward and D.B. Kell, Biosensors Bioelectron., 10 (1995) 639.
[46] B.C. Blake-Colman, M.J. Hutchings and P. Silley, Biosensors Bioelectron., 9 (1994) 231.
[47] M.J. Hutchings, B.C. Blake-Colman and P. Silley, Biosensors Bioelectron., 9 (1994) 91.
[48] F.J. Harris, Proc. IEEE, 66 (1978) 51.
[49] H.P. Schwan, in W.L. Nastuk (Ed.), Physical Techniques in Biological Research, Vol. VIB, Academic Press, New York, 1963, pp. 323–407.
[50] C.L. Davey, G.H. Markx and D.B. Kell, Eur. Biophys. J., 18 (1990) 255.
[51] A.J. Bard and L.R. Faulkner, Electrochemical Methods, Wiley, Chichester, 1980.
[52] H.P. Schwan, Ann. Biomed. Eng., 20 (1992) 269.
[53] M. Moussavi, H.H. Sun, H.P. Schwan and A. Richter, Ann. Biomed. Eng., 18 (1990) 505.
[54] R. De Levie, Ann. Biomed. Eng., 20 (1992) 337.
[55] J.E.B. Randles, Kinetics of rapid electrode reactions, Discuss. Faraday Soc., 1 (1947) 11.
[56] E.T. McAdams and J. Jossinet, Ann. Biomed. Eng., 20 (1992) 307.
[57] A.H. Flasterstein, Med. Biol. Eng., 4 (1966) 583.
[58] C.D. Ferris, Introduction to Bioelectrodes, Plenum, 1974.
[59] T.K. Chen, Anal. Chem., 64 (1992) 1264.
[60] L.A. Larew and D.C. Johnson, J. Electroanal. Chem., 262 (1989) 176.
[61] S. Hughes and D.C. Johnson, Anal. Chim. Acta, 149 (1983) 1.
[62] G.G. Neuburger and D.C. Johnson, Anal. Chim. Acta, 192 (1987) 205.
[63] H. Martens and T. Næs, Multivariate Calibration, Wiley, 1989.
[64] L.S. Ramos, K.R. Beebe, W.P. Carey, E.M. Sánchez, B.C. Erickson, B.E. Wilson, L.E. Wangen and B.R. Kowalski, Anal. Chem., 58 (1986) 294R.
[65] I. Aleksander and H. Morton, An Introduction to Neural Computing, Chapman and Hall, London, 1990.

[66] D.J. Amit, Modeling Brain Function; The World of Attractor Neural Networks, Cambridge University Press, Cambridge, 1989.
[67] R. Beale and T. Jackson, Neural Computing: An Introduction, Adam Hilger, Bristol, 1990.
[68] G.A. Carpenter and S. Grossberg, Pattern Recognition by Self-Organizing Neural Networks, MIT Press, Cambridge, MA, 1991.
[69] J.D. Cowan and D.H. Sharp, Q. Rev. Biophys., 21 (1988) 365.
[70] R.C. Eberhart and R.W. Dobbins, Neural Network PC Tools, Academic Press, London, 1990.
[71] S.I. Gallant, Neural Network Learning, MIT Press, Cambridge, MA, 1993.
[72] R. Hecht-Nielsen, Neurocomputing, Addison-Wesley, Massachusetts, 1990.
[73] J. Hertz, A. Krogh and R.G. Palmer, Introduction to the Theory of Neural Computation, Addison-Wesley, California, 1991.
[74] T. Kohonen, Self-Organization and Associative Memory, Springer-Verlag, Berlin, 1989.
[75] J.R. Long, H.T. Mayfield, M.V. Henley and P.R. Kromann, Anal. Chem., 63 (1991) 1256.
[76] J.L. McClelland and R.E. Rumelhart, Explorations in Parallel Distributed Processing; A Handbook of Models, Programs and Exercises, MIT Press, Cambridge, MA, 1988.
[77] Y.-H. Pao, Adaptive Pattern Recognition and Neural Networks, Addison-Wesley, Reading, MA, 1989.
[78] P. Peretto, An Introduction to the Modelling of Neural Networks, Cambridge University Press, Cambridge, 1992.
[79] D.E. Rumelhart, J.L. McClelland and The PDP Research Group, Parallel Distributed Processing. Experiments in the Microstructure of Cognition, MIT Press, Cambridge, Massachusetts, 1986.
[80] P.K. Simpson, Artificial Neural Systems, Pergamon, Oxford, 1990.
[81] P.D Wasserman, Neural Computing: Theory and Practice, Van Nostrand Reinhold, New York, 1989.
[82] P.D. Wasserman and R.M. Oetzel, Neural Source: the Bibliographic Guide to Artificial Neural Networks, Van Nostrand Reinhold, New York, 1989.
[83] J.W. Ball and P.C. Jurs, Anal. Chem., 65 (1993) 505.
[84] K.R. Beebe, W.W. Blaser, R.A. Bredeweg, J.P. Chauvel, R.S. Harner, M. LaPack, A. Leugers, D.P. Martin, L.G. Wright and E.D. Yalvac, Anal. Chem., 65 (1993) 199R.
[85] B.T. Blank and S.D. Brown, Anal. Chem., 65 (1993) 3081.
[86] L. Boddy and C.W. Morris, Neural network analysis of flow cytometry data, in D. Lloyd (Ed.), Flow Cytometry in Microbiology, Springer-Verlag, London, 1993, pp. 159–169.
[87] C. Borggaard and H.H. Thodberg, Anal. Chem., 64 (1992) 545.
[88] A. Bos, M. Bos and W.E. van der Linden, Anal. Chim. Acta, 256 (1992) 133.
[89] A. Bruchmann, P. Zinn and C.M. Haffer, Anal. Chim. Acta, 283 (1993) 869.
[90] J. Chun, E. Atalan, A.C. Ward and M. Goodfellow, FEMS Microbiol. Lett., 107 (1993) 321.
[91] B. Curry and D.E. Rumelhart, Tetra. Comput. Meth., 3 (1990) 213.
[92] E. Erwin, K. Obermayer and K. Schulten, Biol. Cybernet., 67 (1992) 47.
[93] R. Freeman, R. Goodacre, P.R. Sisson, J.G. Magee, A.C. Ward and N.F. Lightfoot, J. Med. Microbiol., 40 (1994) 170.
[94] P.J. Gemperline, J.R. Long and V.G. Gregoriou, Anal. Chem., 63 (1991) 2313.
[95] R. Goodacre and D.B. Kell, Anal. Chim. Acta, 279 (1993) 17.
[96] R. Goodacre, A.N. Edmonds and D.B. Kell, J. Anal. Appl. Pyrol., 26 (1993) 93.
[97] R. Goodacre, M.J. Neal and D.B. Kell, Anal. Chem., 66 (1994) 1070.
[98] D.B. Kell and C.L. Davey, Bioelectrochem. Bioenerg., 28 (1992) 425.
[99] T.J. McAvoy, H.T. Su, N.S. Wang, M. He, J. Horvath and H. Semerjian, Biotechnol. Bioeng., 40 (1992) 53.
[100] T. Rataj and J. Schindler, Binary, 3 (1991) 159.
[101] D. Richard, C. Cachet, D. Cabrol-Bass and T.P. Forrest, J. Chem. Inf. Comput. Sci., 33 (1993) 202.
[102] R. Shadmehr, D. Angell, P.B. Chou, G.S. Oehrlein and R.S. Jaffe, J. Electrochem. Soc., 139 (1992) 907.
[103] J.R.M. Smits, P. Shoenmaker, A. Stehmann, F. Sijstermans and G. Kateman, Chemom. Intell. Lab. Syst., 18 (1993) 27.
[104] S.H. Weiss and C.A. Kulikowski, Computer Systems That Learn: Classification and Prediction Methods from Statistics, Neural Networks, Machine Learning, and Expert Systems, Morgan Kaufmann Publishers, California, 1991.
[105] J. Zupan, M. Novi, X. Li and J. Gasteiger, Anal. Chim. Acta, 292 (1994) 219–234.
[106] R. Goodacre, S. Trew, C. Wrigley-Jones, G. Saunders, M.J. Neal, N. Porter and D.B. Kell, Anal. Chim. Acta, 313 (1995) 25.
[107] G. Cybenko, Math. Control Signals Syst., 2 (1989) 303.
[108] K. Funahashi, Neural Networks, 2 (1989) 183.
[109] K. Hornik, M. Stinchcombe and H. White, Neural Networks, 2 (1989) 359.
[110] K. Hornik, M. Stinchcombe and H. White, Neural Networks, 3 (1990) 551.
[111] H. White, Neural Networks, 3 (1990) 535.
[112] T. Næs, K. Kvaal, T. Isaksson and C. Miller, J. Near Infrared Spectrosc., 1 (1993) 1.
[113] H.A. Bourlard and N. Morgan, Connnectionist Speech Recognition: A Hybrid Approach, Kluwer Academic Publishers, Boston, MA, 1994.
[114] W.L. Buntine and A.S. Weigend, Complex Systems, 5 (1991) 603.
[115] W.S. Sarle, Neural networks and statistical models, Proc. 19th Ann. SAS Users Group Int. Conf., 1994, SAS, 1994, pp. 1–13.
[116] P.J. Cornbleet and N. Gochman, Clin. Chem., 25 (1979) 432.
[117] D.G. Altman and J.M. Bland, The Statistician, 32 (1983) 307.
[118] M.J. Bland and D.G. Altman, The Lancet, February 8 (1986) 307.
[119] G.C. Garber, R.P. Mallon and A.S. Swern, Anal. Chem., 65 (1993) 480.
[120] U. Feldmann, B. Schneider and H. Klinkers, J. Clin. Chem. Clin. Biochem., 19 (1981) 121.
[121] W.E. Deming, Statistical Adjustment of Data, Wiley, New York, 1943.
[122] J.O. Westgard and M.R. Hunt, Clin. Chem., 19 (1973) 49.
[123] H. Passing and W. Bablock, J. Clin. Chem. Clin. Biochem., 21 (1983) 709.
[124] H. Passing and W. Bablock, J. Clin. Chem. Clin. Biochem., 22 (1984) 431.
[125] W.L. Clarke, D. Cox, L.A. Gonder Frederick, W. Carter and S.L. Pohl, Diabetes Care, 10 (1987) 622.
[126] D. Cox, W.L. Clarke, L.A. Gonder Frederick, S.L. Pohl, C. Hoover, A. Snyder, L. Zimbelman, W. Carter, S. Bobbit and J. Pennebaker, Diabetes Care, 8 (1985) 529.
[127] P.D. Welch, IEEE Trans. AU, 15 (1967) 70.
[128] M.B. Seasholtz and B. Kowalski, Anal. Chim. Acta, 277 (1993) 165.
[129] A.J. Miller, Subset Selection in Regression, Chapman and Hall, London, 1990.
[130] D.B. Kell and B. Sonnleitner, Trends Biotechnol., 13 (1995) 481.
[131] M. Baroni and G. Constantino, Quant. Struct-Act. Relat., 12 (1993) 9.
[132] R.O. Duda and P.E. Hart, Pattern Classification and Scene Analysis, Wiley, New York, 1973.
[133] V.-M. Taavitsainen and P. Kohonen, Chemometr. Intell. Lab. Syst., 14 (1992) 185.
[134] A. Höskuldsson, J. Chemometr., 6 (1992) 307.
[135] I.E. Frank, Chemometr. Intell. Lab. Syst., 8 (1990) 109.
[136] O.M. Kvalheim, D.W. Aksnes, T. Brekke, M.O. Eide, E. Sletten and N. Telnæs, Anal. Chem., 57 (1985) 2858.
[137] S. Wold, Chemometr. Intell. Lab. Syst., 14 (1992) 71.
[138] S. Wold, N. Kettaneh-Wold and B. Skagerberg, Chemometr. Intell. Lab. Syst., 7 (1989) 53.

[139] B.J. Wythoff, Chemometr. Intell. Lab. Syst., 20 (1993) 129.
[140] G. Montague and A.J. Morris, Trends Biotechnol., 12 (1994) 312.
[141] J. Glassey, G.A. Montague, A.C. Ward and B.V. Kara, Process Biochem., 29 (1994) 387.
[142] M.J. Neal, R. Goodacre and D.B. Kell, On the analysis of pyrolysis mass spectra using artificial neural networks. Individual input scaling leads to rapid learning, Proc. World Congress on Neural Networks, 1994, San Diego, I-318–I-323.
[143] S.H. Lillie and J.R. Pringle, J. Bacteriol., 143 (1980) 1384.
[144] J.C. Slaughter and T. Nomura, Enz. Microb. Technol., 14 (1992) 64.
[145] J.M. Thevelein and S. Hohmanm, TIBS, 20 (1995) 3.
[146] W.L. Bryan and R.W. Silman, Enz Microb. Technol., 13 (1991) 2.
[147] R. Serrano, FEBS Lett., 156 (1983) 11.
[148] E.B. Baum and D. Hassler, Neural Computation, 1 (1989) 151.
[149] G.H. Markx, C.L. Davey and D.B. Kell, J. Gen. Microbiol., 137 (1991) 735.
[150] H.M. Davey, C.L. Davey and D.B. Kell, On the determination of the size of microbial cells using flow cytometry, in D. Lloyd (Ed.), Flow Cytometry in Microbiology, Springer-Verlag, Heidelberg, 1993, pp. 49–65.