UNCLASSIFIE
SECURITY CLASSIFICATIC

(4)

REPOI		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TR41	**AD-A197 533**	**3. RECIPIENT'S CATALOG NUMBER**
4. TITLE (and Subtitle) Non-Linear Calibration Using Projection Pursuit Regression: Application to an Array of Ion Selective Electrodes		**5. TYPE OF REPORT & PERIOD COVERED** Technical Report - Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Kenneth R. Beebe and Bruce R. Kowalski		**8. CONTRACT OR GRANT NUMBER(s)** N00014-75-C-0536
9. PERFORMING ORGANIZATION NAME AND ADDRESS Laboratory for Chemometrics Department of Chemistry BG-10 University of Washington		**10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS** NR 051-565
11. CONTROLLING OFFICE NAME AND ADDRESS Materials Sciences Division Office of Naval Research		**12. REPORT DATE** July 1, 1988
		13. NUMBER OF PAGES 32
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)		**15. SECURITY CLASS. (of this report)**
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited.

DTIC
ELECTE
JUL 2 0 1988
S D

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Submitted and accepted for publication in ANALYTICAL CHEMISTRY.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Projection Pursuit
Non-Linear Regression
Sensor Array
Ion Selective Electrodes

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Please see reverse side.

A relatively new method of regression analysis called projection pursuit regression is introduced. This method performs linear and non-linear regression and differs from the more classical methods in that it is able to determine the form of the model as well as the model parameters. The method is applied to data obtained using an array of ion selective electrodes in solutions containing mixtures of Na^+ and K^+. Average relative errors of prediction of 0.4 and 5.3 for Na^+ and K^+, respectively were obtained.

OFFICE OF NAVAL RESEARCH

Contract N00014-75-C-0536

Task No. NR 051-565

TECHNICAL REPORT NO. 41

Non-Linear Calibration Using Projection Pursuit Regression:
Application to an Array of Ion Selective Electrodes

by

Kenneth R. Beebe and Bruce R. Kowalski

Prepared for Publication
in
Analytical Chemistry

Department of Chemistry. BG-10
University of Washington
Seattle. Washington   98195

A-1

July 1. 1988

Much effort has been expended in the field of analytical chemistry toward the development of selective sensors. The ultimate goal of this area of research is to build sensors that respond to only one analyte while ignoring all other analytes (interferents) that are present in the samples. Perhaps the most common example of the result of this effort is the development of ion selective electrodes (ISE) for the determination of ion concentrations in solutions. While some ISEs are relatively selective for the desired ions, all suffer from some degree of non-specificity. Unfortunately, in the field of sensor development this is a common occurrence.

Another approach to solving the problem of interferents is to use multiple non-selective sensors and employ multivariate mathematics (1) to perform the calibration and prediction. This was the approach taken in two recent papers (2,3) where arrays of ISEs were used to quantify mixtures of analytes. Analyte quantitation was achieved using either linear (2) or non-linear (3) regression techniques to model the response of the electrodes to the concentration of analytes in mixture samples. In both papers, the sensor responses were assumed to obey the relationship found in the set of extended Nernst equations,

$$E_{ij} = E_j^o + S_j \log\left(a_i + \sum_l K_{jl} \, a_{il} \right) \qquad 1)$$

where E_{ij} is the potential of the jth electrode in the array measured for the ith sample with respect to a suitable reference electrode; E_j^o is the intercept potential of the jth electrode; S_j is the slope of the response of the electrode in the absence of any interferents (analytes to which the sensor responds for which it was not designed); a_i is the activity of the analyte for which the electrode was made; a_{il} is the activity of the lth interfering ion; and K_{jl} is the selectivity coefficient of the jth electrode with respect to the lth interfering ion. In this study, only two analytes were present (sodium and potassium) and therefore equation 1 can be rewritten in the following form,

$$E_{ij} = E_j^o + S_j \log(c_{iNa^+} + K_j \, c_{iK^+}) \qquad 2)$$

where the order of the analytes has been arbitrarily chosen and the activities have been replaced by concentrations (c_{iNa^+}, c_{iK^+}).

Otto and Thomas (2) used calibration samples containing only one analyte to determine the slopes (S_j) for each electrode, rearranged equation 2, and used multiple linear regression (4) and partial least squares (5) to determine the remaining two parameters for the model (E_j^o and K_j). Beebe et al. (3) used non-linear regression based on a simplex algorithm and multiple linear regression (MLR) to determine the model parameters with no a priori information concerning the slope of the electrode responses.

This study will use a relatively new method of analysis called projection pursuit (6) to determine the model parameters. Projection pursuit is a nonparametric multivariate technique that allows the analyst to calibrate a system with no a priori information about the functional form of the calibration model. In other words, given the responses of J sensors to I samples containing mixtures of K analytes, projection pursuit can find an appropriate form for the model (log, parabolic, linear, etc.) as well as the model parameters. In the more common calibration procedures, knowledge of the functional form is an essential component. For example, in building the calibration model for an experiment involving absorption measurements, it is common practice to assume the instrument response follows Beer's Law and use a regression procedure to build the model. If the linearity criterion is not obeyed, the derived models are not valid, and the true models cannot be obtained using the normal linear regression techniques. Similarly, previous papers treating the calibration of arrays of ISEs (2,3) based the calibration models on the assumption that the electrodes obey a known response equation. Not knowing this "functional form" can make the calibration quite inaccurate. Furthermore, unexpected departure from the assumed functional form can yield erroneous results. For these reasons, non-parametric methods in general and projection pursuit in particular are powerful and versatile tools with a wide variety of possible applications.

Theory

Notation Throughout this paper matrices will be represented with bold uppercase letters (**R**); vectors with bold lower case letters such as r_j to signify the jth column of **R**; and scalar quantities with plain upper and lower case letters (I, i).

Projection Pursuit In general, the goal of all multivariate calibration procedures is to estimate model parameters relating an IxJ matrix **R** containing the responses of J sensors or wavelengths to I calibration samples to an IxK matrix **C** containing the concentrations or characteristics of K analytes in the same I samples. Once an estimate of the model parameters are obtained, it is possible to predict the concentration of analytes in a new sample of unknown concentrations.

If the form of the model relating **R** to **C** is unknown, often the analyst "assumes" linearity and hopes for success. Another approach is to guess at a functional form and test its reasonableness with the calibration data. The problem with the latter approach is that there are too many functional forms from which to choose. There is literally an infinite number of models that can be constructed even for the simplest case of one response vector and one concentration vector.

Projection pursuit limits the number of choices by allowing the calibration data to determine an appropriate model. The procedure projects the K dimensional calibration data (the K columns of **C**) into a smaller space while retaining the multivariate structure. In other words, it determines the linear combination of predictor variables that is "best" related to the columns of **R**. As in any calibration procedure, the analyst must decide on which of the calibration matrices (concentration versus response) to use as the predictor variables. In this study, the concentration vectors were used in this role because the errors in the responses are presumed larger than those in the concentrations (see equation 3 where the projection pursuit model assumes the errors are primarily in r_j) and because of an interest in estimating the coefficients for the columns of **C** to compare to the earlier study.

When using this approach, projection pursuit can be used to estimate the model for one ion selective electrode at a time. The resulting projection pursuit model for the jth sensor is as follows,

$$r_j = G_j \left(\sum_{k=1}^{K} \alpha_{kj} \, c_k \right) + \varepsilon_j \qquad\qquad 3)$$

where r_j is an Ix1 vector corresponding to the response of the jth sensor to I calibration samples; α_{kj} is the coefficient for the kth predictor variable (c_k) where $\sum \alpha_k = 1$; G_j represents a smooth of the linear combination described by the α quantities (an explanation of this follows); and ε_j is the error associated with fitting the data for the jth sensor.

For each of the j sensors, projection pursuit's goal is to determine the α quantities and a smooth that minimizes the associated error (ε_j). It achieves this by first searching for appropriate α quantities, calculating a smooth given that linear combination of predictor variables, and calculating the error. It then iteratively searches for the linear combination that minimizes the error. The optimal linear combination is that which yields the most narrow band of points when plotted against r_j. A simple hypothetical problem using simple graphics will make this more clear.

For illustration purposes, it will be assumed that an analyst has obtained a calibration set of the response of one ISE to I mixtures of two analytes. Let r be the Ix1 vector of responses of the ISE to the I samples and C be the Ix2 matrix of concentrations. Again, for illustration purposes, assume that the first projection pursuit iteration yielded $0.37c_1 + 0.93c_2$ as the initial linear combination of predictor variables to describe r. From equation 3 it is clear that $\alpha_1 = 0.37$ and $\alpha_2 = 0.93$. Projection pursuit next determines whether or not this choice of α quantities is reasonable by "viewing" the relationship between r and the chosen linear combination of predictor variables (fig. 1). (In practice, the method "views" the result mathematically, and plots are included here only for instructional purposes.)

To determine the acceptability of the chosen α quantities, projection pursuit determines a *smooth* (7) going through the points in figure 1. A smooth is a continuous line that describes the

general behavior of all of the points. The value of the smooth at any point is based on the local average of q points on either side where q defines the bandwidth of the smooth. The smooth in figure 2 was obtained using running averages where q=5 followed by a polynomial smooth of degree 3. The important point of this illustration is to demonstrate how a smooth describes the general trend of of the data. The degree to which the smooth describes each point individually (as opposed to the overall trend) can be determined by adjusting the value of q. Although smooths are not common in calibration procedures, some common methods of analysis are very similar in their results. For example, the ordinary regression line can be thought of as a type of smooth describing the data with the restriction of linearity. Another commonly employed procedure is the use of spline functions (8) to approximate data when derivative spectra are desired. The spline function breaks the data into sections and fits polynomials to each section to form a continuous curve.

The smooth used by the projection pursuit is more like the spline function than the regression line in that it is not based on any model criterion such as linearity. The smooth is based solely on the behavior of the data and does not have any functional form. However, in situations where there is a real functional relationship (i.e. log, exponential) between the variables being smoothed, the data should reflect the relationship and the smooth should closely approximate the true model. For more details concerning the smooth employed by projection pursuit see reference 6.

Returning to the illustration, once a linear combination of predictor variables is chosen, projection pursuit calculates a smooth. The deviation of the points from the smooth is then used to determine whether the linear combination of predictor variables (c_k) is acceptable or if the procedure should continue iterating. If the fit is not acceptable, a non-linear regression technique (based on the method of Rosenbrock (9)) is used to choose a new set of α quantities to examine. A new smooth and corresponding fit are calculated, and the process is repeated until the deviation from the smooth converges. In the present hypothetical example it will be assumed that the α quantities determined in the first iteration ($\alpha = [0.37 \quad 0.93]$) were not the best possible and that the procedure continued searching for new α quantities until the following model was determined.

$$r = G(0.85\ c_1 + 0.53\ c_2) \tag{4}$$

Figure 3 is a plot of **r** versus this linear combination of **c** vectors where the desired fit of the data to the smooth is achieved. Projection pursuit has found a linear combination of predictor variables that results in a narrow band of points when **r** is plotted versus that linear combination.

Once the calibration model has been derived, prediction can be accomplished using two different schemes. The first is to examine the derived smooth and determine whether it corresponds to a known function. If an adequate function is available, it can be used in place of the smooth and the resulting model can be used for prediction. This was the approach taken by the present work to calibrate the array of ISEs because of the good correspondence between the smooth and the log function, the ease of implementation of this method, and the unavailability of the true smooth function as derived by the projection pursuit program.

The second approach to prediction is where projection pursuit is the most versatile. Instead of replacing the smooth with a known function, prediction can be performed using the smooth itself. This is possible because the smooth function is continuous within the range defined by the calibration samples and therefore interpolation can be performed. To illustrate how the smooth can be used for prediction, one can imagine estimating a calibration model where the instrument responses at three wavelengths are used as predictor variables and the concentration of an analyte is to be estimated. A possible model may be,

$$c = G\ (\ 0.21\ r_1 + 0.36\ r_2 + 0.91\ r_3\ ) \tag{5}$$

where G and $\alpha = [\ 0.21\ \ 0.36\ \ 0.91\ ]$ correspond to the final model. In this example note that the roles of the response and concentration vectors are not the same as in the calibration of ISEs. Here, the response vectors are treated as predictor variables while the concentration vector is treated as a dependent variable. These are not equivalent approaches, and the choice of which

method to use depends on the error structure and the goals of the experiment.

Once the calibration model has been determined (eq. 5), the analyst can predict the concentration of analytes in an unknown sample (ex. with response $r_{un} = [1 \ 2 \ 3]$) by evaluating the value of the estimated smooth at the point $(1)(0.21) + (2)(0.36) + (3)(0.91) = 3.66$.

The advantages of using the estimated smooth in this manner is that departures from ideality are modelled, and representative models for many systems that do not follow "common" functional forms can be obtained. This is because a smooth is not constrained to follow any predetermined functionality and is therefore more able to model the behavior of the data. This characteristic can make the method very valuable in an exploratory sense where unknown relationships between variables are sought. One caveat to the use of this method is that it is possible to overfit the data. Although the smooth used by the projection pursuit is robust and therefore not overly sensitive to outliers (non-representative calibration samples), it is still capable of fitting noise as well as data. Another limitation is that the method is not useful in situations where the unknown sample lies outside of the range of responses and concentrations defined by the calibration set. When a smooth is used as an integral part of the model, extrapolation is not possible. This is because smooths are a product of the data itself and therefore cannot be used to infer about behavior beyond the data. It should be noted, however, that both of these limitations are present in cases where the model is known and the analyst should be aware of the possible complications regardless of the method employed.

Calibration The functionality of the response of an ISE to mixtures of analytes (eq. 2) makes it very amenable to analysis using the projection pursuit algorithm. The data used in this study were taken from the third experiment in the study performed in reference 3. The data consisted of the responses of five ISEs to mixtures of sodium (0.1200 - 0.1650 M) and potassium (2.00 - 8.40 mM) ions in aqueous solutions (See table 1 for concentration levels of the 11 calibration and 9 prediction samples). An 11x5 matrix **R** was constructed with rows

corresponding to samples, and columns corresponding to sensors (see table 2), likewise, an 11x2 matrix C was formed with columns corresponding to sodium and potassium ion concentrations, respectively. To build the model for the jth electrode, projection pursuit was used to find α_{1j}, and α_{2j} as in equation 3 where r_j is the response of the jth sensor to the 11 calibration samples; and c_1 and c_2 are the concentrations of sodium and potassium in these samples, respectively.

For this study, the first approach to prediction was taken where the smooth was replaced with a functional form. As was expected, the log function corresponded well with the derived smooth. To compare the model results obtained using projection pursuit to those obtained in the earlier study using the same data, the following steps were followed to obtain projection pursuit estimates of E^o and S for each of the sensors.

The first step is to assume the final smooth corresponds to the log function for each sensor. This is suggested by the data as will be discussed later. The α quantities and the log function can then be used to transform the predictor variables (c_i) to yield a new vector l_j that is linearly related to r_j.

$$l_j = \log(\alpha_{1j} c_1 + \alpha_{2j} c_2) \qquad 6)$$

Regressing r_j onto l_j yields the following equation,

$$r_j = \beta_{0j} + \beta_{1j} \log(\alpha_{1j} c_1 + \alpha_{2j} c_2) \qquad 7)$$

To find the corresponding estimates of E_j^o and S_j, the following rearrangements of equation 7 can be made.

$$r_j = \beta_{0j} + \beta_{1j} \log\{\alpha_{1j} [c_1 + (\alpha_{2j}/\alpha_{1j}) c_2]\} \qquad 8)$$

$$r_j = [\beta_{0j} + \beta_{1j} \log(\alpha_{1j})] + \beta_{1j} \log[c_1 + (\alpha_{2j}/\alpha_{1j}) c_2] \qquad 9)$$

Comparing this equation to equation 2, yields the following equalities.

$$K_j = \alpha_{2j}/\alpha_{1j} \qquad\qquad 10)$$

$$S_j = \beta_{1j} \qquad\qquad 11)$$

$$E_j^0 = \beta_{0j} + \beta_{1j}\log(\alpha_{1j}) \qquad\qquad 12)$$

Note that equations 6-12, and the prediction equations that follow, are based on the assumption that the smooth determined by projection pursuit is equal to the log function. As stated earlier, in situations where no functional form can be found to represent the smooth, the calibration and prediction models can be constructed using the smooth function itself. This latter approach will not be employed in the present work but can prove to be a powerful alternative to the more traditional approach.

Prediction Once the calibration model has been estimated for each of the five electrodes, prediction of analyte concentrations for unknown samples can be accomplished by rearranging equation 7 and employing MLR. The first step is to write equation 7 in the following form,

$$10^{[(r_j - \beta_{0j})/\beta_{1j}]} = \alpha_{1j}c_1 + \alpha_{2j}c_2 \qquad\qquad 13)$$

In equation 13, r_j, c_1, and c_2 are in plain text because this equation is for one unknown sample as opposed to the general case of I samples in equation 7. For each unknown sample the value of r_j is measured, β_{0j} and β_{1j} are estimated, and therefore the left hand side of the equation is a known quantity. Therefore, for each sensor, one equation of the following form is found for the unknown sample,

$$x_j = \alpha_{1j}c_1 + \alpha_{2j}c_2 \qquad\qquad 14)$$

where x_j is the quantity on the left hand side of equation 13. This can be rewritten in vector notation as follows,

$$x_j = \alpha_j^T c \qquad \qquad 15)$$

If the x values calculated for each sensor (for the same unknown sample) are stacked to form a column vector (x), and the corresponding α_j for each sensor are also stacked to form a matrix A, the following equality will hold,

$$x = A\ c + \varepsilon \qquad \qquad 16) \quad .$$

where c is a vector of concentrations for the unknown sample. This equation can be solved using MLR to yield estimates for the elements of c in the following manner,

$$c = (A^T A)^{-1} A^T\ x \qquad \qquad 17)$$

By following these steps, the calibration model can be used to predict the concentrations of analytes in unknown samples.

Results and Discussion

Five electrodes were used to generate the data analyzed in this study; a Corning 476220 general purpose cation glass electrode, an Orion 94-11 sodium glass electrode with uncharacteristic non-selective behavior (3), and three plastic membrane based electrodes whose construction will not be discussed here. The sensors will be referred to as TNO, GP, EHPP, NA, and METH for sensors 1-5, respectively. For a more complete description of the electrodes, instrumentation, and

data acquisition the reader is referred to reference 3.

The first step was to use projection pursuit to determine the values of the α quantities in equation 3. For this study, 11 calibration samples were used to derive the parameters and the resulting models were used to predict the concentration of analytes contained in an additional 9 test samples. (It should be noted that projection pursuit is generally used in situations where the dimensionality of the problems and the number of samples used in the calibration step (I) are much larger than in this example. The ability of projection pursuit to successfully estimate the model parameters in the present study is due to the low level of noise present.) The resulting α quantities for the models of the five sensors are listed in table 3. Since the procedure for model estimation was identical for each of the five sensors, the details of only the first sensor will be discussed.

Assuming the functional form for sensor 1 is not known, the next step would be to plot r_1 versus the "best" linear combination of the c vectors ($.919c_1+.395 c_2$). This plot reveals that projection pursuit has done a good job at forming a tight band of points. A close examination of the plot also suggests that the true function describing the data has a slight curvature (negative second derivative). If a straight line is fit to these points and the residuals examined (fig. 4), it becomes clear that there is structure in the data that is not accounted for using a simple regression line on the untransformed data. The shape of the residual structure suggests a suitable transformation is to descend the ladder of powers for ($.919 c_1 + .395 c_2$) and the log function is a reasonable choice. In this study, there was obviously a strong bias toward using the log function as this is suggested by theory. In other situations the analyst may not have any a priori knowledge about the relationship between dependent and independent variables. In those situations, projection pursuit may provide clues concerning the functional relationships. From this information, hypotheses can be formulated and further investigation performed to verify or disprove the hypotheses.

For ISEs, the log function was suspected and the data substantiated the form of equation 2. A plot of log($.919 c_1 + .395 c_2$) versus r_1 revealed a more linear relationship. Figure 5 is the residual plot of the regression of r_1 on the transformed linear combination of c vectors where the

lack of structure in the new residuals verifies the log transformation as being acceptable. Additionally, the log function can be shown to be appropriate by comparing the magnitude of the residuals (taking into account that the log transformation has been performed). (Note: The original experiment was set up using a factorial design (10) to choose the levels for the concentration of analytes. For this reason, four sets of points in figures 4 and 5 are clustered together and are not ideal for determining the optimal transformation. If the experiment were to be repeated, it would be advisable to use a different scheme to select analyte concentration levels for the calibration samples.)

The next step of the calibration procedure is the regression of the vector of responses onto the vector l_j (eq. 7). This regression step yielded the coefficients found in table 4. These values can be used to estimate the parameters in the original Nernst equation (eq. 2) using the equalities found in equations 10-12. The estimated parameters using projection pursuit and those of the earlier study of Beebe et al. are listed in table 5. The agreement between the two methods is very good and one would expect the predictive abilities of the two methods to be comparable. Note that it is possible to improve the projection pursuit model once the log function was chosen to replace the smooth. One approach to calibration is to use projection pursuit to find the functional relationships and some other non-linear regression technique to determine the model parameters given the estimated functional form for the calibration model.

Table 6 lists the results of using the projection pursuit model to predict the concentrations of analytes in the nine test samples using equation 17. The results of the earlier study using the simplex model are also included for comparison. These results show that the model derived using the simplex method yielded slightly better prediction than the model determined using projection pursuit. This is a reasonable result because of the approach to prediction that was followed. The projection pursuit smooth was assumed to equal the log transformation and the α quantities calculated for the smooth were used with the log function. These α quantities were not exactly optimal for the log function and therefore the projection pursuit model did not perform as well as the model derived using the simplex procedure.

As stated earlier, a second approach would have been to use the smooth itself as a part of the calibration model (eq. 5). Both of these are reasonable options and either could be used depending on whether the analyst has more confidence in the theory (in which case the functional model would be desirable) or the calibration samples (where the smooth would be used).

Conclusion

Projection pursuit has been presented along with an example of its use. Although ion selective electrodes were used in this study, it is not to be inferred that this is the method of choice for the calibration of ISEs. ISE data were used to illustrate the effectiveness and capabilities of projection pursuit for calibration and model estimation in general. Many other possible applications can be imagined and a variety of approaches to the data analysis can be used depending on the particular type of data at hand. The most powerful aspect of the technique is that there is no need to assume any functional relationship between variables under investigation. The method can be used to verify assumed relationships, detect outliers, and determine functional relationships between variables that are unknown. Furthermore, the method is not tied to any fixed functional form. If the data do not follow some standard form, the more common modes of analysis cannot be used for model building. Using the smooths as the functional forms allow the analyst to build models that are uniquely characteristic of the system under investigation.

Acknowledgement

The projection pursuit program used was a written as a macro on the computer package S. The authors thank the University of Washington department of Statistics for the use of both their computer and software.

References:

1) Ramos, L.S.; Beebe, K.R.; Carey, W.P.; Sanchez, M.E.; Erickson, B.C.; Wilson, B.E.; Wangen, L.E.; Kowalski, B.R. *Anal. Chem.* **1986**, *58*, 294R.

2) Otto, M.; Thomas, J.D.R. *Anal. Chem.* **1985**, *57*, 2647.

3) Beebe, K.R.; Uerz D.; Sandifer, J.; Kowalski, B.R. *Anal. Chem.*, **1988**, *60*,66-71.

4) Draper, N.R.; Smith, H. *Applied Regression Analysis* **1981**, Wiley: New York.

5) Geladi, P.; Kowalski, B.R. *Anal. Chim. Act.* **1986**, *185*, 1-17.

6) Friedman, J.H.; Stuetzle, W. *J. Am. Stat. Assoc.* **1981**, *76*, 817-823.

7) Tukey, J.W. *EDA Exploratory Data Analysis* **1977**, Addison-Wesley: Mass.

8) Gans, P.; Gill, J.B. *Appl. Spectrosc.* **1984**, *38*, 370-376.

9) Rosenbrock, H.H. *Computer Journal* **1960**, *3*, 175-184.

10) Box, G.E.P.; Hunter, W.G.; Hunter, J. S. *Statistics for Experimenters* **1978**, Wiley: New York.

Figure Captions:

Figure 1. For the hypothetical example, plot of r versus $0.37c_1 + 0.93c_2$.

Figure 2. Plot of r versus $0.37c_1 + 0.93c_2$ with a smooth included. This plot shows the relatively poor fit of the data to the initial smooth estimate.

Figure 3. Plot of r versus $0.85c_1 + 0.53c_2$ for the hypothetical final model where the desired tight belt of points is achieved.

Figure 4. Plot of residuals versus fitted values resulting from the regression of r_1 onto $(.919c_1 + .395c_2)$.

Figure 5. Plot of residuals versus fitted values resulting from the regression of r_1 onto $\log(.919c_1 + .395c_2)$.

Figure 1.

$$0.37\,c_1 + 0.93\,c_2$$

Figure 2.

Figure 3.

r

$0.85\,c_1 + 0.53\,c_2$
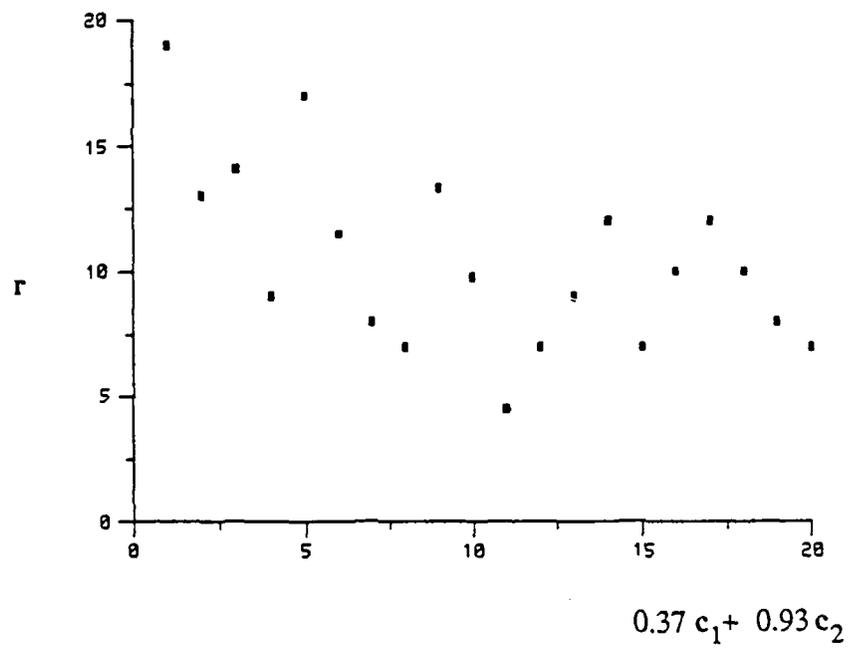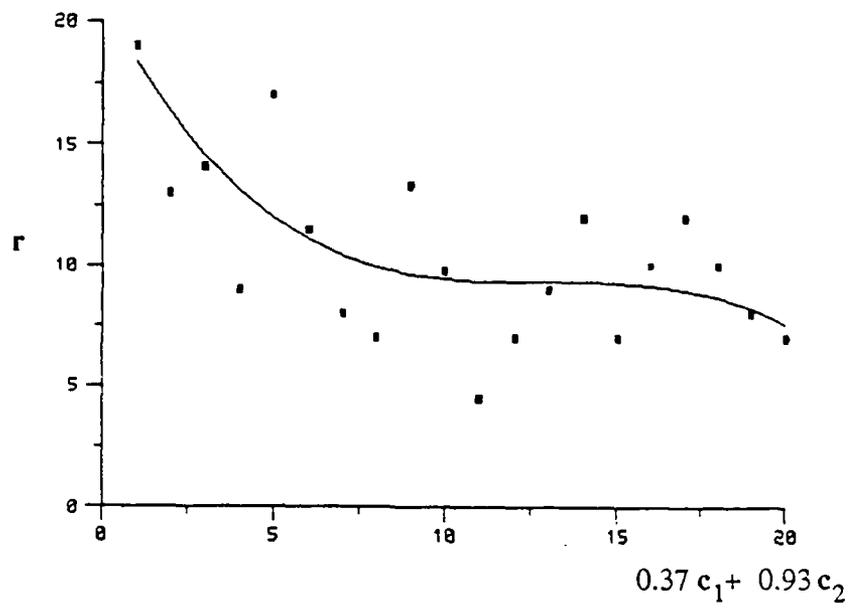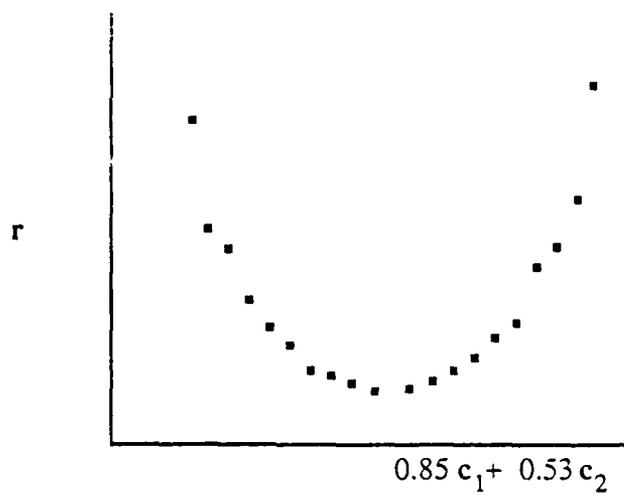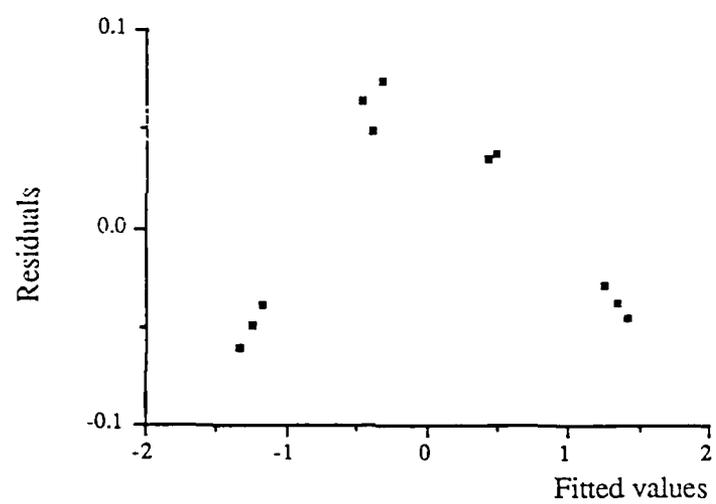
Figure 4.

Figure 5.
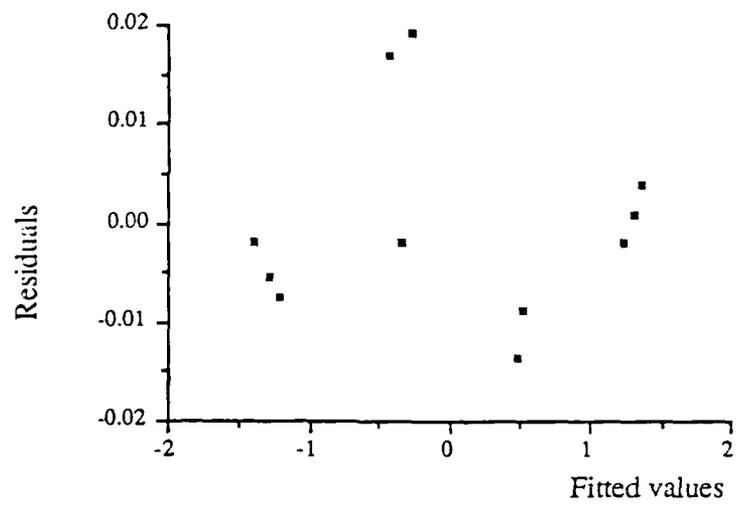
Table I.   Composition of samples used for calibration and prediction.

| | Calibration | | | Prediction | |
Sample	$[Na^+]$ M	$[K^+]$ mM	Sample	$[Na^+]$ M	$[K^+]$ mM
1	0.1200	2.00	1	0.1200	3.82
2	0.1200	5.48	2	0.1200	7.00
3	0.1200	8.40	3	0.1350	3.82
4	0.1350	2.00	4	0.1350	7.00
5	0.1350	5.48	5	0.1500	2.00
6	0.1350	8.40	6	0.1500	7.00
7	0.1500	3.82	7	0.1500	8.40
8	0.1500	5.48	8	0.1650	3.82
9	0.1650	2.00	9	0.1650	7.00
10	0.1650	5.48			
11	0.1650	8.40			

Table II. Responses of sensors to calibration and prediction samples.

Calibration sample			Sensor		
	TNO	GP	EHPP	NA	METH
1	-3.10	-49.00	4.80	13.20	-6.40
2	-2.50	-45.40	6.20	15.60	-5.70
3	-2.40	-43.10	6.90	17.10	-5.50
4	-0.20	-46.60	7.10	16.50	-4.20
5	0.10	-43.50	8.40	18.30	-3.60
6	0.30	-41.20	9.40	19.70	-3.40
7	2.10	-43.00	9.60	19.70	-1.20
8	2.30	-41.60	10.10	20.50	-0.90
9	4.40	-42.60	10.90	20.90	0.90
10	4.40	-40.00	11.80	22.30	1.20
11	4.40	-38.20	12.40	23.20	1.30

Prediction

	TNO	GP	EHPP	NA	METH
1	-2.80	-47.10	5.40	14.40	-6.10
2	-2.60	-44.30	6.50	16.40	-5.70
3	-0.10	-45.00	7.90	17.50	-3.90
4	0.10	-42.30	8.90	19.00	-3.50
5	2.00	-44.60	9.00	18.90	-1.50
6	2.30	-40.60	10.40	21.00	-0.80
7	2.20	-39.70	10.70	21.40	-0.80
8	4.50	-41.10	11.50	21.70	1.20
9	4.40	-39.00	12.10	22.70	1.20

Table III.   Results of projection pursuit of r_j and C.

Sensor	α_1	α_2
TNO	0.919	0.395
GP	0.152	0.988
EHPP	0.403	0.915
NA	0.310	0.951
METH	0.819	0.574

Table IV. Results of regression step of calibration phase.

Sensor	β_0	β_1
TNO	46.40	51.63
GP	35.32	49.70
EHPP	63.18	44.81
NA	91.34	55.18
METH	46.38	52.72

Table V.   Final models derived from projection pursuit and simplex (3).

	Model					
	Projection pursuit			Simplex		
Sensor	E^O	S	K	E^O	S	K
TNO	44.50	51.63	0.43	44.51	51.64	0.44
GP	-5.36	49.70	6.51	-5.36	49.69	6.51
EHPP	45.47	44.81	2.27	45.37	44.78	2.39
NA	63.25	55.18	3.07	63.07	55.07	3.18
METH	41.81	52.72	0.70	41.81	52.72	0.70

Table VI. Predicted concentrations and (% relative error) of prediction for test samples.

	Model			
	Projection pursuit		Simplex	
Sample	$[Na^+]$ M	$[K^+]$ mM	$[Na^+]$ M	$[K^+]$ mM
1	0.1202 (0.2)	3.50 (8.4)	0.1198 (0.2)	3.73 (2.4)
2	0.1999 (0.1)	6.92 (1.1)	0.1199 (0.1)	6.86 (2.0)
3	0.1342 (0.6)	4.26 (11.5)	0.1351 (0.1)	3.89 (1.8)
4	0.1341 (0.7)	7.53 (7.6)	0.1349 (0.1)	7.15 (2.1)
5	0.1493 (0.5)	2.13 (6.5)	0.1498 (0.1)	2.03 (1.5)
6	0.1498 (0.1)	6.99 (0.1)	0.1497 (0.2)	7.00 (0.0)
7	0.1486 (0.3)	8.48 (1.0)	0.1486 (0.9)	8.45 (0.6)
8	0.1668 (1.1)	3.50 (8.4)	0.1665 (0.9)	3.69 (3.4)
9	0.1645 (0.3)	6.79 (3.0)	0.1641 (0.5)	7.03 (0.4)
Average % relative error	(0.4)	(5.3)	(0.3)	(1.6)

TECHNICAL REPORT DISTRIBUTION LIST, GEN

	No. Copies		No. Copies
Office of Naval Research Attn: Code 413 800 N. Quincy Street Arlington, Virginia 22217	2	Dr. David Young Code 334 NORDA NSTL, Mississippi 39529	1
Dr. Bernard Douda Naval Weapons Support Center Code 5042 Crane, Indiana 47522	1	Naval Weapons Center Attn: Dr. A. B. Amster Chemistry Division China Lake, California 93555	1
Commander, Naval Air Systems Command Attn: Code 310C (H. Rosenwasser) Washington, D.C. 20360	1	Scientific Advisor Commandant of the Marine Corps Code RD-1 Washington, D.C. 20380	1
Naval Civil Engineering Laboratory Attn: Dr. R. W. Drisko Port Hueneme, California 93401	1	U.S. Army Research Office Attn: CRD-AA-IP P.O. Box 12211 Research Triangle Park, NC 27709	1
Defense Technical Information Center Building 5, Cameron Station Alexandria, Virginia 22314	12	Mr. John Boyle Materials Branch Naval Ship Engineering Center Philadelphia, Pennsylvania 19112	1
DTNSRDC Attn: Dr. G. Bosmajian Applied Chemistry Division Annapolis, Maryland 21401	1	Naval Ocean Systems Center Attn: Dr. S. Yamamoto Marine Sciences Division San Diego, California 91232	1
Dr. William Tolles Superintendent Chemistry Division, Code 6100 Naval Research Laboratory Washington, D.C. 20375	1		

ABSTRACTS DISTRIBUTION LIST, 051B

ABSTRACTS DISTRIBUTION LIST, 051B

Professor J. Janata
Department of Bioengineering
University of Utah
Salt Lake City, Utah 84112

Dr. J. DeCorpo
NAVSEA
Code 05R14
Washington, D.C. 20362

Dr. Charles Anderson
Analytical Chemistry Division
Athens Environmental Laboratory
College Station Road
Athens, Georgia 30613

Dr. Ron Flemming
B 108 Reactor
National Bureau of Standards
Washington, D.C. 20234

Dr. Frank Herr
Office of Naval Research
Code 422CB
800 N. Quincy Street
Arlington, Virginia 22217

Professor E. Keating
Department of Mechanical Engineering
U.S. Naval Academy
Annapolis, Maryland 21401

Dr. M. H. Miller
1133 Hampton Road
Route 4
U.S. Naval Academy
Annapolis, Maryland 21401

Dr. Clifford Spiegelman
National Bureau of Standards
Room A337 Bldg. 101
Washington, D.C. 20234

Dr. Denton Elliott
AFOSR/NC
Bolling AFB
Washington, D.C. 20362

Dr. B. E. Spielvogel
Inorganic and Analytical Branch
P.O. Box 12211
Research Triangle Park, NC 27709

Ms. Ann De Witt
Material Science Department
160 Fieldcrest Avenue
Raritan Center
Edison, New Jersey 08818

Dr. A. Harvey
Code 6110
Naval Research Laboratory
Washington, D.C. 20375

Mr. S. M. Hurley
Naval Facilities Engineering Command
Code 032P
200 Stovall Street
Alexandria, Virginia 22331

Ms. W. Parkhurst
Naval Surface Weapons Center
Code R33
Silver Spring, Maryland 20910

Dr. M. Robertson
Electrochemical Power Sources Division
Code 305
Naval Weapons Support Center
Cra..., Indiana 47522

Dr. Andrew T. Zander PI204
Perkin-Elmer Corporation
901 Ethan Allen Highway/MS905
Ridgefield, Connecticut 06877

ABSTRACTS DISTRIBUTION LIST, 051B

Dr. Marvin Wilkerson
Naval Weapons Support Center
Code 30511
Crane, Indiana 47522

Dr. J. Wyatt
Naval Research Laboratory
Code 6110
Washington, D.C. 20375

Dr. J. MacDonald
Code 6110
Naval Research Laboratory
Washington, D.C. 20375

Dr. H. Wohltjen
Naval Research Laboratory
Code 6170
Washington, D.C. 20375

Dr. John Hoffsommer
Naval Surface Weapons Center
Building 30 Room 208
Silver Spring, Maryland 20910

Dr. Robert W. Shaw
U.S. Army Research Office
Box 12211
Research Triangle Park, NC 27709