JOURNAL OF CHEMOMETRICS
J. Chemometrics 2001; **15**: 285–298
DOI: 10.1002/cem.689

Application of PARAFAC2 to fault detection and diagnosis in semiconductor etch

Barry M. Wise[1]*, Neal B. Gallagher[1] and Elaine B. Martin[2]

[1]*Eigenvector Research, Inc., 830 Wapato Lake Road, Manson, WA 98831, USA*
[2]*Center for Process Analytics and Control Technology, University of Newcastle, Newcastle upon Tyne, UK*

SUMMARY

Monitoring and fault detection of batch chemical processes are complicated by stretching of the time axis, resulting in batches of different length. This paper offers an approach to the unequal time axis problem using the parallel factor analysis 2 (PARAFAC2) model. Unlike PARAFAC, the PARAFAC2 model does not assume parallel proportional profiles, but only that the matrix of profiles preserves its 'inner product structure' from sample to sample. PARAFAC2 also allows each matrix in the multiway array to have a different number of rows. It has previously been demonstrated how the PARAFAC2 model can be used to model chromatographic data with retention time shifts. Fault detection and, to a lesser extent, diagnosis in a semiconductor etch process are considered in this paper. It is demonstrated that PARAFAC2 can effectively model batch process data from semiconductor manufacture with unequal dimension in one of the orders, such as the unequal batch length problem. It is shown that the PARAFAC2 model has approximately the same sensitivity to faults as other competing methods, including principal component analysis (PCA), unfold PCA (often referred to as multiway PCA), trilinear decomposition (TLD) and conventional PARAFAC. The advantage of PARAFAC2 is that it is easier to apply than MPCA, TLD and PARAFAC, because unequal batch lengths can be handled directly rather than through preprocessing methods. It also provides additional diagnostic information: the recovered batch profiles. It is likely, however, that it is less sensitive to faults than conventional PARAFAC. Copyright © 2001 John Wiley & Sons, Ltd.

KEY WORDS: multiway; curve resolution; multivariate statistical process control; batch process monitoring

1. INTRODUCTION

Many important pharmaceuticals, specialty chemicals and polymers are made in batch chemical processes. Many steps in the manufacture of semiconductor devices, such as metal etch, can also be considered batch processes. Fault detection and diagnosis are important in these systems in order to assure consistent quality products and efficient process utilization. In each of these applications the time required to process a batch may vary. This results in data records of unequal length for each batch. In addition, even batches of equal length may not follow exactly the same time trajectory,

* Correspondence to: B. M. Wise, Eigenvector Research Inc., 830 Wapato Lake Road, Manson, WA 98831, USA.
E-mail: bmw@eigenvector.com
Contract/grant sponsor: United Kingdom Engineering and Physical Sciences Research Council (EPSRC)
Contract/grant sponsor: University of Newcastle

Received 29 February 2000
Accepted 23 February

taking longer in some steps of the processing and less time in others. This makes it less than straightforward to apply techniques such as unfold PCA and PARAFAC to the data, as these models assume that the range of normal process trajectories can be modeled as the sum over fixed time profiles. Several approaches have been developed to deal with this problem, such as dynamic time warping [1–3], truncation of the records, including centering of the record about a specific process event [4], and the introduction of a surrogate variable [5]. A number of these approaches are compared in Reference [6]. Each of these approaches alters the original data record in order to make it fit the model. It will be shown that PARAFAC2 [7,8] can be used on the original data directly, which is potentially advantageous.

2. DESCRIPTION OF THE DATA

The data used in this study come from the metal etch step in semiconductor processing, specifically the Al stack etch process. Data were collected on the commercially available Lam 9600 plasma etch tool [2]. In this process the TiN/Al–0·5% Cu/TiN/oxide stack is etched in an inductively coupled BCl_3/Cl_2 plasma. The key parameters of interest are the linewidth of the etched Al line, uniformity across the wafer, and the oxide loss. Process conditions must be kept constant in order assure consistent results.

The standard recipe for the process consists of a series of six steps. The first two steps are for gas flow and pressure stabilization. Step 3 is a brief plasma ignition step. Step 4 is the main etch of the Al layer terminating at the Al endpoint, with step 5 acting as the over-etch for the underlying TiN and oxide layers. Note that this is a single chemistry etch process, i.e. the process chemistry is identical during steps 3–5. Step 6 vents the chamber.

The metal etcher used for this study was equipped with three sensor systems: machine state, radio-frequency monitor (RFM) and optical emission spectroscopy (OES). The machine state sensors, built into the processing tool, collect machine data during wafer processing. The machine data consist of 40 process setpoints and measured and controlled variables sampled at 1 s intervals during the etch. These are engineering variables such as gas flow rates, chamber pressure and RF power. In this work, non-setpoint process variables with some normal variation were used for monitoring. Engineering judgment was also used to select variables that should impact product quality. These variables are listed in Table I.

The RFM sensors measure the voltage, current and phase relationships at the fundamental frequency of 13·56 MHz and the next four harmonics at four locations in the RF control system. The resulting 70 values are sampled every 3 s. The presence of each chemical species affects the plasma power and phase relationships in unique ways; thus the RFM sensors provide a surprising amount of chemical information.

Table I. Machine state variables used for process monitoring

1	BCl_3 flow	11	RF power
2	Cl_2 flow	12	RF impedance
3	RF bottom power	13	TCP tuner
4	RFB reflected power	14	TCP phase error
5	Endpoint A detector	15	TCP impedance
6	Helium pressure	16	TCP top power
7	Chamber pressure	17	TCP reflected power
8	RF tuner	18	TCP load
9	RF load	19	Vat valve
10	Phase error		

The OES is used to monitor the plasma in the range from 245 to 800 nm in three locations above the wafer using fiber optics. The original data consist of 2042 channels per location; however, in this work the data were preprocessed by integrating a much smaller number of peaks (40) in each of the three spectra which correspond to process gases and species evolving from the wafer due to the etch [2].

A major objective of this work was to determine which sensors, or combinations of sensors, are most useful for detecting process faults. Data from the three sensor systems were used to develop models of the process in a variety of ways, and the ability of the models to detect faults was tested.

Faults were induced in the data by changing the setpoints for controlled variables, such as chamber pressure and plasma power, from the normal recipe. The mean values of the controlled variables were then set back to those of the normal recipe. The effect is a data record where it appears as if a bias had developed in the sensor for the controlled variable.

Three experiments, numbered 29, 31 and 33, were performed to produce the data used here. The experiments were run several weeks apart. Forty-three wafers were processed in each experiment. The three experiments together generated data for 108 normal wafers and 21 wafers with induced faults, though there are several instances where the data records for the RFM and OES are not complete. These data are available in the Eigenvector Research Data Archive at http://www.eigenvector.com/Data/Data_sets.html.

Fault detection results for PCA, unfold PCA (MPCA), TLD and PARAFAC were reported in Reference [4]. In this work, PARAFAC2 is considered.

3. THE PARAFAC AND PARAFAC2 MODELS

The PARAFAC model was originally proposed by Harshman [9] and also by Carroll and Chang [10] under the name CANDECOMP. PARAFAC [11] models a three-dimensional array $\underline{\mathbf{X}}$ ($I \times J \times K$) as a summation over R outer products of triads of vectors. In order to allow the use of conventional linear algebra notation, it is convenient to describe the PARAFAC model in terms of each $I \times J$ slab of $\underline{\mathbf{X}}$, $\mathbf{X}_k$. Thus

$$\mathbf{X}_k = \mathbf{F}\mathbf{D}_k\mathbf{A}^{\mathrm{T}} + \mathbf{R}_k \tag{1}$$

where $\mathbf{F}$ is an $I \times R$ matrix of factor scores (for the row units), $\mathbf{A}$ is a $J \times R$ matrix of weights for the column units, $\mathbf{D}_k$ is a diagonal ($R \times R$) matrix containing the weights for the kth slab of $\underline{\mathbf{X}}$, and $\mathbf{R}_k$ denotes an $I \times J$ matrix of residuals. The PARAFAC model is typically fitted to the array $\underline{\mathbf{X}}$ through a procedure of alternating least squares in an attempt to minimize the sum of squared residuals in $\underline{\mathbf{R}}$.

In the PARAFAC model, all the $\mathbf{X}_k$ slabs are of fixed size, $I \times J$. As noted above, this might not always be the case. The PARAFAC2 model [10] was developed to handle the situation where the number of observations (row dimension) in each $\mathbf{X}_k$ may vary. The PARAFAC2 model is given by

$$\mathbf{X}_k = \mathbf{F}_k\mathbf{D}_k\mathbf{A}^{\mathrm{T}} + \mathbf{R}_k \tag{2}$$

which is identical to the PARAFAC model except that $\mathbf{F}$ is not fixed for all the slabs $\mathbf{X}_k$. Instead, there is a unique $n_k \times R$ matrix $\mathbf{F}_k$ for each $n_k \times J$ slab $\mathbf{X}_k$. It is clear that without some restrictions on the model this form would fit a principal component analysis (PCA) model of the data of the unfolded $\underline{\mathbf{X}}$. In PARAFAC2, however, the loading matrices $\mathbf{AD}_1$, …, $\mathbf{AD}_K$ must be proportional, which alone would not make the model unique. In addition, the constraint that the cross-product matrix $\mathbf{F}_k^{\mathrm{T}}\mathbf{F}_k$ is constant over k is imposed, which under fairly general conditions is enough to make the model unique (see discussion in Reference [7]). However, the appropriateness of this constraint deserves some consideration. This constraint is equivalent to requiring that the covariance matrix of each batch span

the same subspace. While this is clearly an approximation, it is reasonable to expect that the major components of the system covariance would be similar in practice. In the event that they are not, identification of this condition would be advantageous. It has been shown that systems that violate this constraint can be modeled with success [8].

It can be shown that the constraint $\mathbf{F}_k^{\mathrm{T}}\mathbf{F}_k = \mathbf{F}_j^{\mathrm{T}}\mathbf{F}_j$ for all $j, k = 1, \ldots, K$ is equivalent to requiring that $\mathbf{F}_k = \mathbf{P}_k\mathbf{F}$ for some columnwise orthonormal $(n_k \times R)$ matrix $\mathbf{P}_k$ and a fixed $(R \times R)$ matrix $\mathbf{F}$. Given this, the PARAFAC2 model can be identified by minimizing

$$\sigma(\mathbf{P}_1, \ldots, \mathbf{P}_k, \mathbf{F}, \mathbf{A}, \mathbf{D}_1, \ldots, \mathbf{D}_k) = \sum \|\mathbf{X}_k - \mathbf{P}_k\mathbf{F}\mathbf{D}_k\mathbf{A}^{\mathrm{T}}\|^2 \tag{3}$$

over all its arguments subject to the constraints that $\mathbf{P}_k^{\mathrm{T}}\mathbf{P}_k = \mathbf{I}_r$ and all the $\mathbf{D}_k$ are diagonal. Algorithms for identifying PARAFAC2 models are described in detail in Reference [7].

An existing PARAFAC2 model can be fit to new data. This is done by minimizing the function given in Equation (3) subject to the additional constraint that $\mathbf{F}$ and $\mathbf{A}$, determined from the calibration set, are fixed. The algorithm proceeds as shown in Reference [7]; however, all steps which update $\mathbf{F}$ and $\mathbf{A}$ are omitted.

All computations were performed in MATLAB 5·2 [12]. Code for fitting the 'calibration' PARAFAC2 models was developed by Rasmus Bro and used in Reference [8] and is freely available at http://www.models.kvl.dk/source/. Code for refitting the models to test data was developed in MATLAB by the authors.

4. CONTROL LIMITS FOR MSPC

The central idea of multivariate or, in this case, multiway statistical process control is to first develop a model that describes the calibration data. New data are then compared to the calibration model to see how well it fits. Any assessment of the fit, however, must include some statistical measures of the similarity of the new data to the old. In the PARAFAC2 model there are several parameters which can be tracked. The first and perhaps most important is the lack of fit, measured by the sum of squared residuals. One problem with this measure in PARAFAC2 is that the size of the data record for each batch can be different. Thus we have chosen to normalize the sum of squared residuals by the number of time samples in order to obtain an average residual over the batch:

$$\mathrm{SSE}_n(k) = \mathrm{SSE}/n_k = \sum \|\mathbf{R}_k\|^2/n_k \tag{4}$$

Control limits can then be placed on the SSE_n through use of the χ^2 distribution. The SSE_n values for the calibration data are fit to the $g\chi^2 h$ distribution using the method of moments. This fit can be done easily by taking advantage of the fact that the mean of a χ^2 distribution is equal to the degrees of freedom and the mean is equal to twice the variance. The scale factor g is thus determined from

$$g = \mathrm{var}(\mathrm{SSE}_n)/(2 \times \mathrm{mean}(\mathrm{SSE}_n)) \tag{5}$$

while the degrees of freedom parameter h is

$$h = \mathrm{floor}(\mathrm{mean}(\mathrm{SSE}_n)/g) \tag{6}$$

where 'floor' indicates that the degrees of freedom h is rounded down to the next lowest integer. Limits can then be calculated for the desired confidence interval from the χ^2 distribution. In this work we choose 99% as the action limit.

As in PCA, limits can also be placed on the 'scores' for each of the batches, in our case the elements of **D**. In the development of limits we will assume that these are normally distributed, though later it will be shown that this is often not the case. Thus each of the R columns of **D** is characterized by a mean and a variance. Limits can be obtained using the standard normal deviate where the number of calibration samples exceeds 30, which is the case in this work. For smaller calibration sets the Student t statistic may be employed. The 99% confidence limits on the scores are the mean $\pm$ 2·58 times their standard deviation (leaving 0·5% in each tail of the distribution).

5. THE PROBLEM OF CORRELATED FACTOR SCORES

In PCA the scores are necessarily orthogonal, but in PARAFAC and PARAFAC2 this need not be the case. In PCA the factors are abstract, while PARAFAC and PARAFAC2 at least attempt to estimate intrinsic factors. There is no reason to believe that underlying factors in the data would be independent. An example of this is given in Figure 1, which shows the scores for the first two components in a PARAFAC2 model of the OES data for all the experiments. The calibration data are shown as pluses, while the test data are shown as asterisks with a label indicating the type of fault. Besides the clustering of the data, it is apparent that the scores of the calibration data are very correlated. This is due to the overall decline in signal strength of OES as the window into the etch chamber clouds over. Note that many of the fault wafers are obvious because they do not follow the trend. However, if limits are based simply on the individual calibration scores, none of the fault

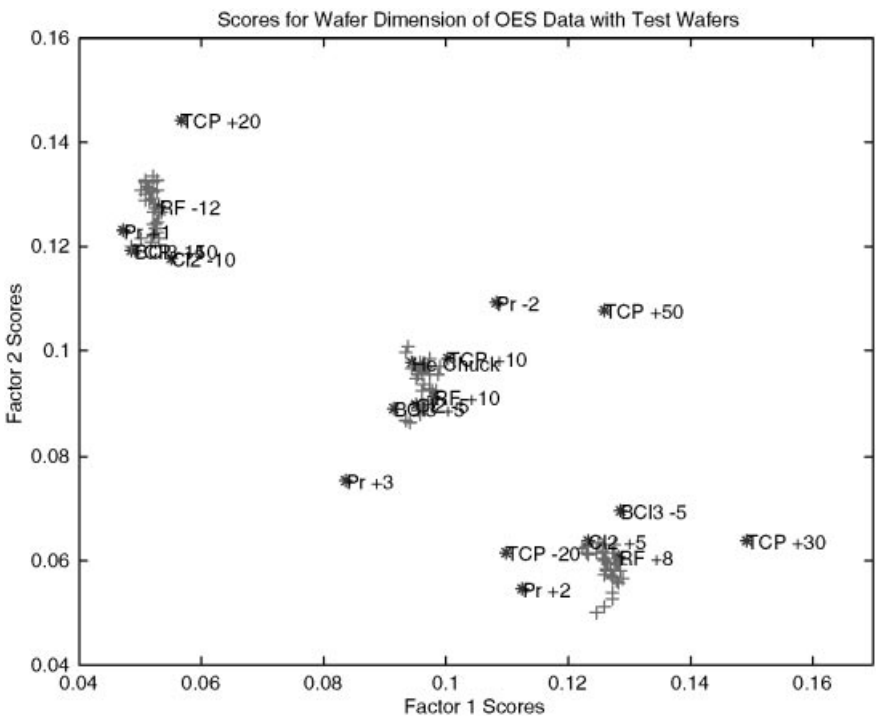

Figure 1. PARAFAC2 scores ($\mathbf{D}_k$) for all wafers based on model of normal wafers based on OES data.

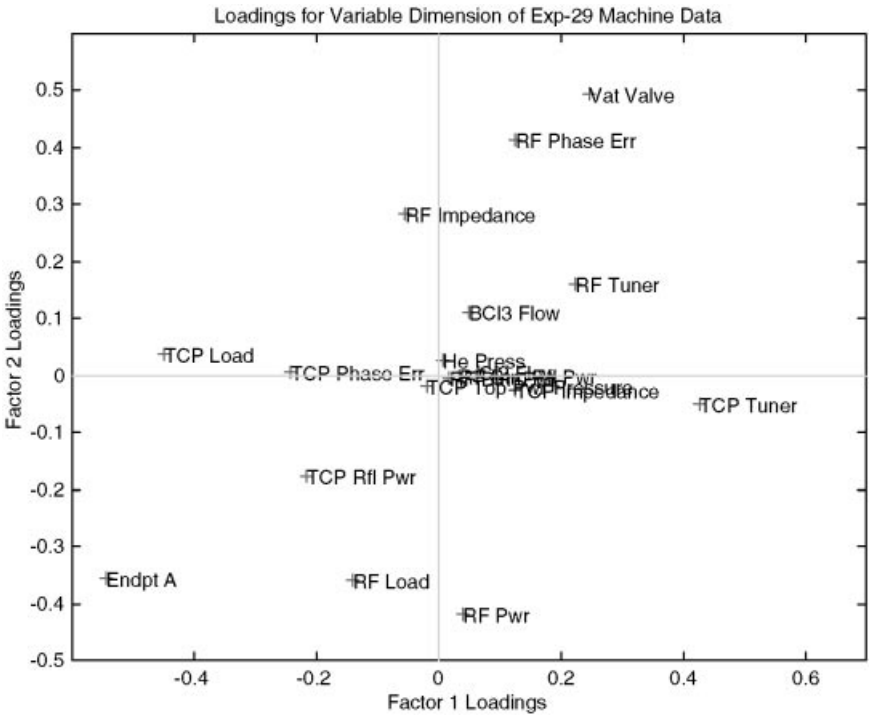

Figure 2. Variable loadings (**A**) for PARAFAC2 model of Exp-29 machine data.

wafers are detected as unusual, because none of the faults lie outside the range of the individual scores.

The problem of correlated batch scores can be solved by application of Hotelling's T^2 statistic. Once a PARAFAC2 model is identified from the calibration data, the elements of $\mathbf{D}_k$ are collected in a single matrix $\mathbf{C}$ ($K \times R$) such that the kth row of $\mathbf{C}$ is equal to the diagonal elements of $\mathbf{D}_k$. The matrix $\mathbf{C}$ is then autoscaled over all the batches to form the matrix $\mathbf{C}'$. T^2 values are then obtained for each of the K batches from

$$T_k^2 = \mathbf{C}'_k(\mathbf{C}'^{\mathrm{T}}\mathbf{C}')^{-1}\mathbf{C}'^{\mathrm{T}}_k \tag{7}$$

where $\mathbf{C}'_k$ is the kth row of $\mathbf{C}'$. Limits can be calculated for future batches using the F distribution from

$$T^2_{R,K,\alpha} = [R(K-1)/(K-R)]F_{R,K-R,\alpha} \tag{8}$$

T^2 values for new batches can be computed after the PARAFAC2 calibration model is fit to the new data. The new batch scores are scaled by the original mean and variance of $\mathbf{C}$ and then Equation (7) is applied with the new scores replacing $\mathbf{C}'$.

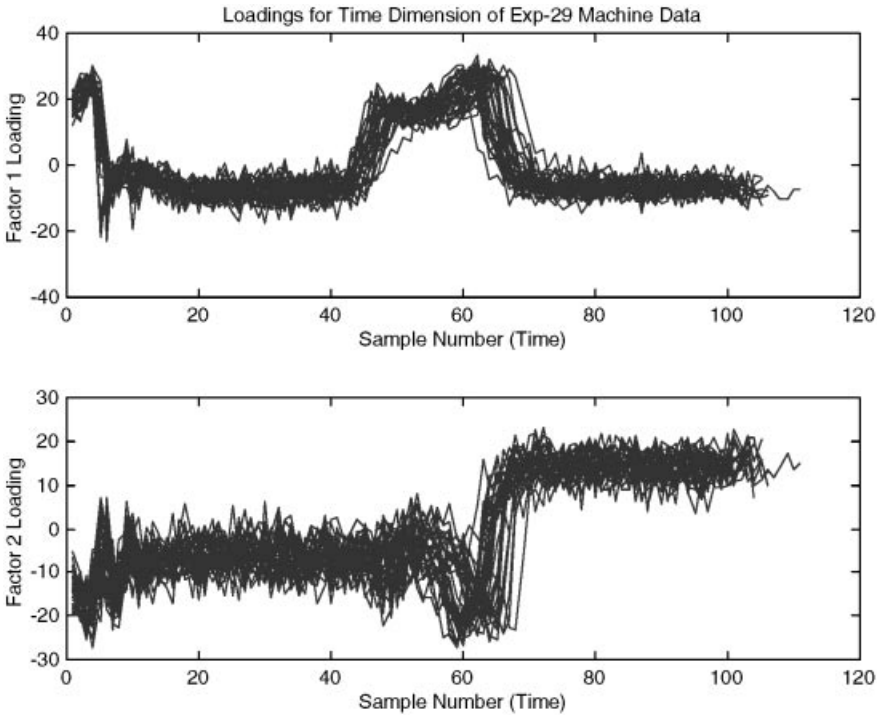

Figure 3. Time loadings ($\mathbf{F}_k$) for PARAFAC2 model of Exp-29 machine data.

6. SCALING

Scaling in multiway models can be a complex issue, as illustrated recently by Smilde [13]. A common method for scaling in MPCA models is to first remove the mean process trajectory by unfolding the $I \times J \times K$ array to $K \times IJ$ and mean centering the columns, followed by unfolding to $IK \times J$ and scaling the columns to unit variance [4]. Alternatively, the $I \times J \times K$ array can be unfolded to $K \times IJ$ and autoscaled, i.e. adjusted to mean zero and unit variance [14]. The choice depends upon how heavily one wants to weight periods of high process variance.

In PARAFAC2, however, the concept of a mean trajectory is meaningless, as the batches are potentially of different length and, even when they are not, are expected to have different time behavior. In this study the machine and RF data were unfolded to matrices of sum(n_k) $\times J$ (total number of rows over all the batches by number of variables) and then autoscaled. As pointed out in Reference [13], this is probably suboptimal. Instead, offsets should be determined for each variable such that the model best describes the data in their original form. This is a possible improvement to the method and should be considered in future work. Because the OES data were expected to have a natural zero, they was not centered or scaled at all.

7. DIAGNOSTICS

Like PCA, PARAFAC and PARAFAC2 are rich in diagnostic information. Loadings in the variable dimension J give information about how the original variables relate to each other. As an example, consider a PARAFAC2 model of the experiment 29 (Exp-29) machine data for the normal wafers (34 wafers). A two-factor model captured 32·9% of the variation in the data (additional factors did not

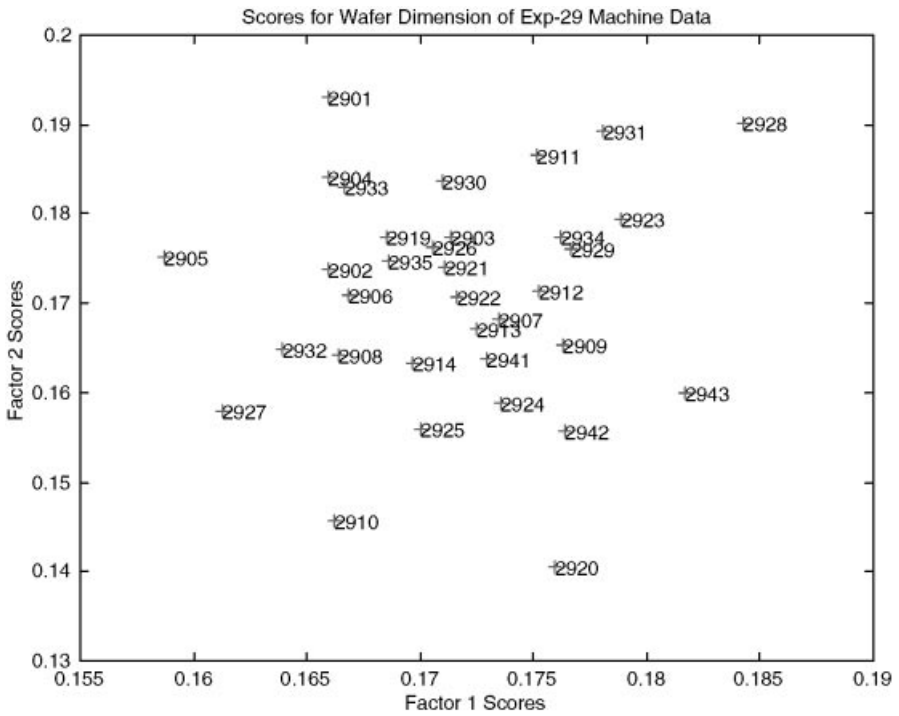

Figure 4. Batch scores ($\mathbf{D}_k$) of Exp-29 machine data.

appear to capture systematic variation). A bivariate loading plot for the variable dimension of these data is shown in Figure 2. Loadings in the time dimension are shown in Figure 3. The loadings for each batch are superimposed. Note that the stretching of the time axis is evident: different batches reach inflections and extremes at different time points, though the overall shapes are very similar. When PARAFAC models are fit to these data, the residuals (collected over the time points) are large in the transition areas (not shown). This is because the time stretching affects where the transitions occur, but the PARAFAC model assumes a fixed time trajectory for each of the factors. The batch scores are shown in Figure 4 and are keyed to the wafer number (two-digit experiment number followed by two-digit wafer number within the experiment). In this instance the batch scores appear quite random and show no obvious trend with wafer number.

It is also interesting to look at where the PARAFAC2 model does not fit. The percent variance explained in each variable is given in Figure 5. Here it is apparent that some of the variables, such as BCl_3 and Cl_2 flow, are not well explained by the model. In fact, 10 of the 19 variables have less than 20% of their variance explained. Inspection of the raw data (not shown) indicates that these variables are exceedingly noisy, with very little trend. Residuals can also be collected over the time dimension and, of course, the batch dimension.

The results of fitting the nine wafers with induced faults to the PARAFAC2 model of the normal Exp-29 machine data are shown in Figure 6. Here the fit of the new data to the model is summarized by charting the SSE, T^2 and scores on the individual factors. In all cases, limits are shown based on 99% confidence intervals. Wafers outside the 99% limits are marked with an asterisk. The batch scores for the wafers can be viewed as a bivariate plot as shown in Figure 7. Here the test wafers are marked with an asterisk. The pressure faults and TCP top power fault are very obvious. When the source of the fault

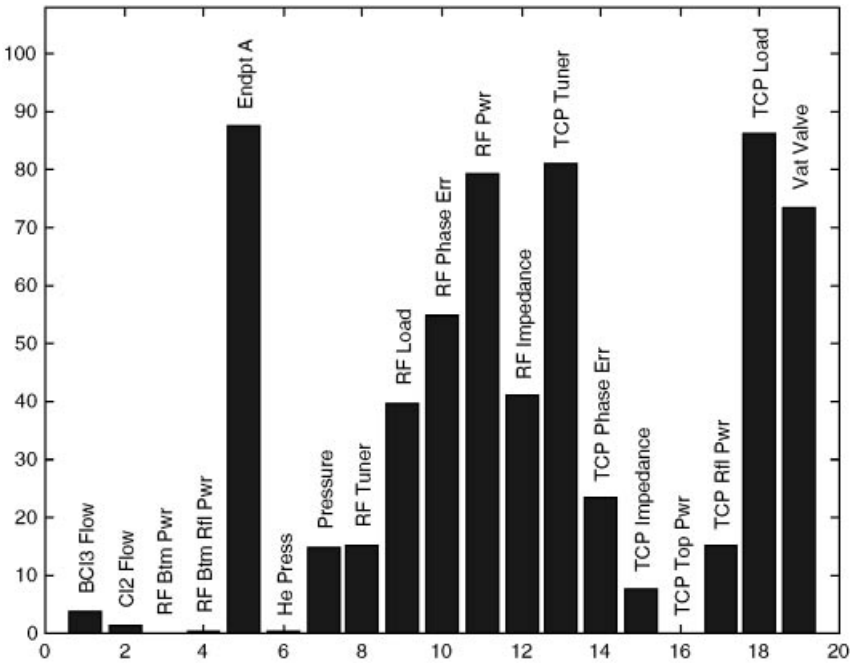

Figure 5. Percent variance explained by variable for PARAFAC2 model of Exp-29 machine data normal wafers.

is not known, it is possible to interrogate the model by plotting the contribution to the residual by variable over each of the wafers 15–17. Plots for the two pressure faults are shown in Figures 8 and 9. Residuals were collected over the variable dimension and are normalized by their standard deviation over the calibration wafers. This puts all the variables on a scale relative to their expected values. Limit lines are drawn for a 99% confidence assuming that they are normally distributed ($\pm 2{\cdot}58$ standard deviations). The importance of retaining the sign of the deviation is evident in the plots: the pattern in the fault where the pressure is too low is primarily the negative of the one where the pressure was too high. These contributions can be easily interpreted in these cases. If the overall pressure is too high but the pressure sensor reads normally (the Pr + 3 fault, Figure 7), the vat valve position has a large negative residual because it is not open nearly as far as it should be given the values of the other variables, including the apparent pressure. Other faults can be much more difficult to interpret, but with knowledge of the physics of the system they can often be understood.

Because PARAFAC2 fits the time loadings to each batch, these traces can also be compared to the calibration data. This is shown in Figure 10. Here the traces from the test data are shown with the limits derived from the mean of the calibration traces $\pm 2{\cdot}58$ standard deviations. The traces of the test data fit largely within the calibration data, though some overall deviations in the shape are apparent for some of the wafers.

8. RESULTS AND DISCUSSION

The results of the fault detection tests are shown in Table II. The data sets considered are listed along the left side. The 'global' results refer to building models on data from all three experiments, while 'local' refers to building separate models for each experiment. Results reported in Reference [4] for TLD, PARAFAC, MPCA and PCA on the wafer means are included for comparison. (Note, however, that T^2

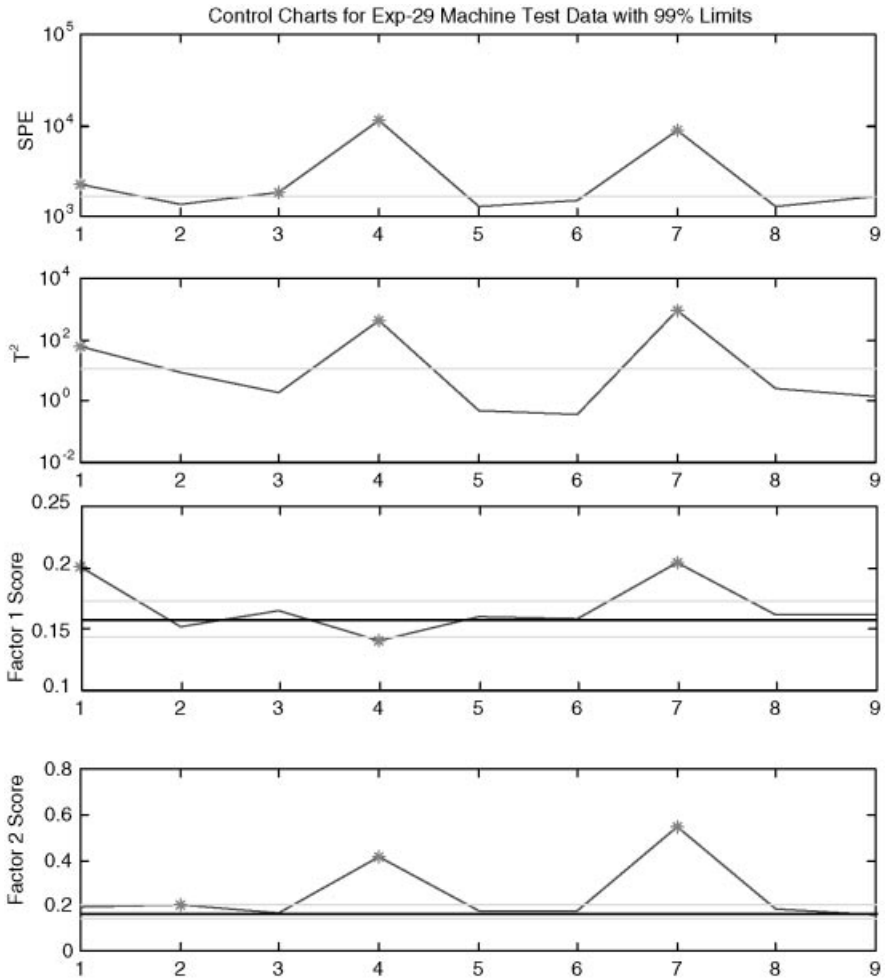

Figure 6. MSPC control charts for Exp-29 faults on PARAFAC2 model of Exp-29 normal machine data.

was not used on TLD and PARAFAC scores.) For a fault to be detected, the SPE, T^2 or one of the factor scores was outside the calculated 99% limits. Values in the table are the number of faults caught out of the 20 available (fault wafers with all three sensor sets available). In general, the sensitivity of the PARAFAC2 model was very similar to that of the methods discussed in Reference [4], somewhat more sensitive than MPCA, but not as sensitive as PARAFAC, though not significantly so.

It is apparent that PARAFAC2 is slightly less sensitive than PARAFAC. This is evidently due to the additional flexibility of the model, which allows for significant deviation in the characteristic time behavior of the process. In fact, if fitted time profiles are not considered, PARAFAC2 would allow for complete reordering of the data records without indicating a fault. This is a potential drawback of the method. However, the availability of the fitted time profiles is very useful as it provides much diagnostic information with regard to the trajectory of the particular batch. This is not available in MPCA, TLD and PARAFAC, though residuals collected across time points may provide similar information. The biggest single advantage of PARAFAC2 is its ability to deal with data records of different length in a 'natural' way without requiring preprocessing of the data.

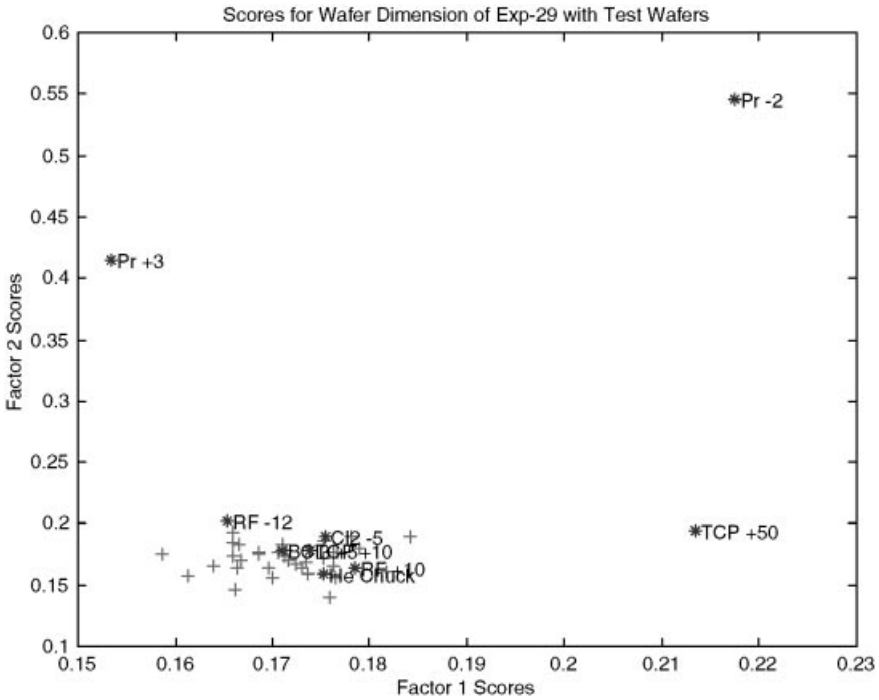

Figure 7. Batch scores ($\mathbf{D}_k$) of Exp-29 fault data on model of Exp-29 normal data.

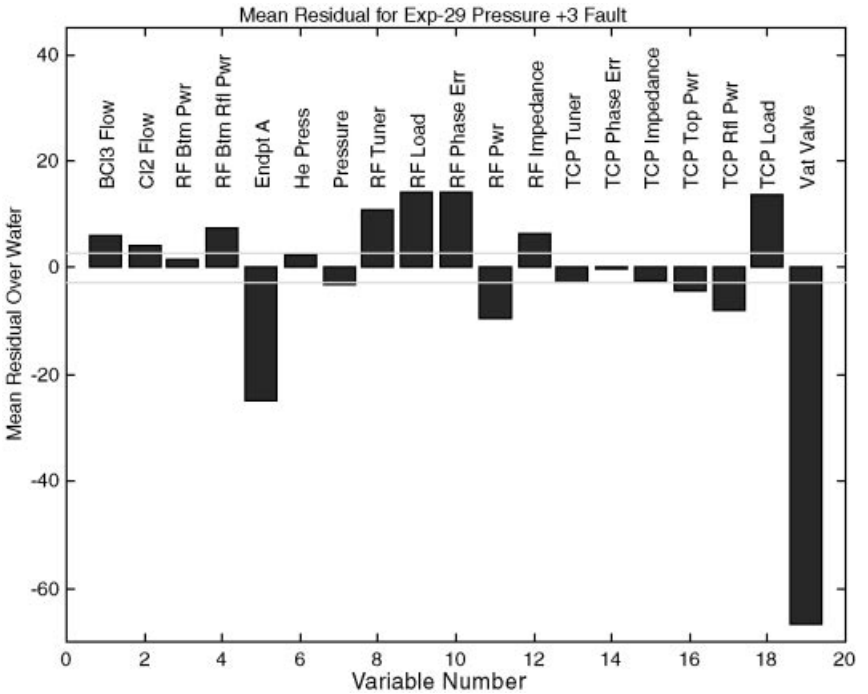

Figure 8. Contribution plot for wafer with 3 Torr positive pressure fault.

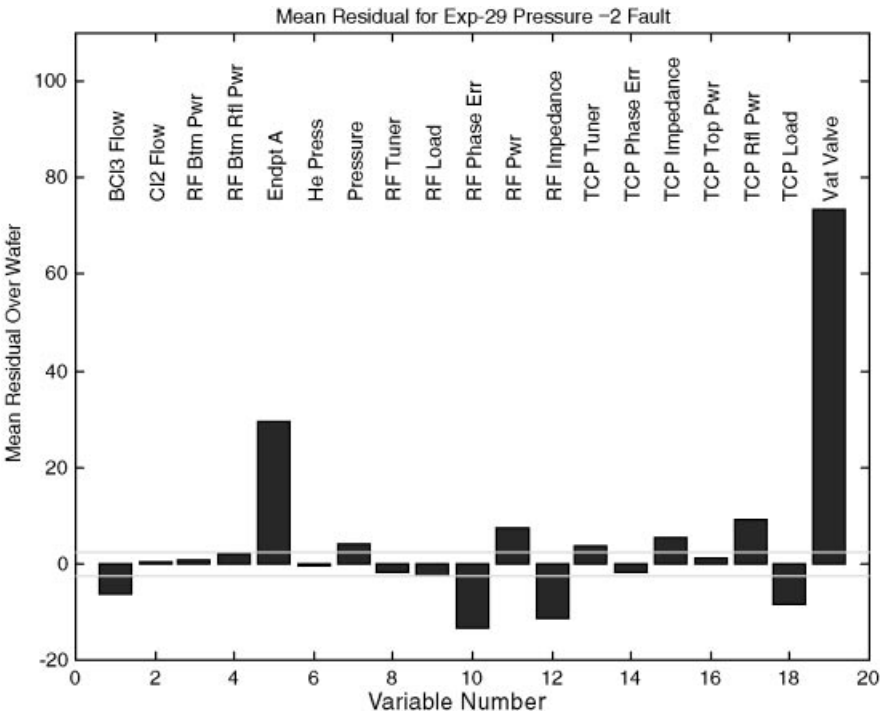

Figure 9. Contribution plot for wafer with 2 Torr negative pressure fault.

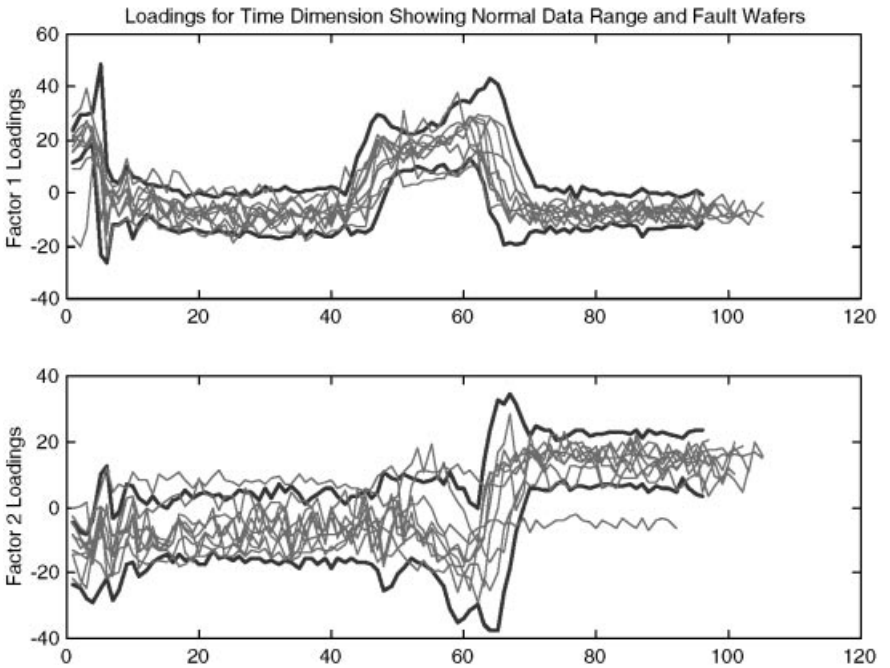

Figure 10. Fitted time profiles ($\mathbf{F}_k$) for test wafers with range of calibration wafers shown.

Table II. Faults caught by PARAFAC2 compared with TLD, PARAFAC, MPCA and PCA on the means

		TLD	PARAFAC	MPCA	PCA/means	PARAFAC2
Global	Machine	11	12	10	10	7
	RFM	7	6	11	9	12
	OES	9	6	6	5	8
Means		9·0	8·0	9·0	8·0	9·0
Local	Machine	14	17	11	16	13
	RFM	12	14	14	12	11
	OES	12	11	6	13	11
Means		12·7	14·0	10·3	13·6	11·7
Overall means		10·8	11·0	9·7	10·8	10·3

ACKNOWLEDGEMENTS

The data used in this study were collected by Texas Instruments as part of the SEMATECH J-88 project. This work was funded by a grant from the United Kingdom Engineering and Physical Sciences Research Council (EPSRC). Additional funding and support were provided by the University of Newcastle. The authors would like to thank Rasmus Bro for the original PARAFAC2 code used in this study, and for his comments on the manuscript.

REFERENCES

1. Gollmer K, Posten C. Detection of distorted pattern using dynamic time warping algorithm and application for supervision of bioprocesses. *Prepr. IFAC Workshop on On-line Fault Detection and Supervision in the Chemical Process Industries*, Newcastle, June 1995; 115–120.
2. Barna GG. Procedures for implementing sensor-based fault detection and classification (FDC) for advanced process control (APC). *SEMATECH Tech Transfer Doc 97013235A-XFR*, 1997.
3. White D, Barna GG, Butler SW, Wise B, Gallagher N. Methodology for robust and sensitive fault detection. *Electrochemical Society Meet.*, Montreal, May 1997; **97–9**: 55–79.
4. Wise BM, Gallagher NB, Butler SW, White Jr DD, Barna GG. A comparison of principal components analysis, multi-way principal components analysis, tri-linear decomposition and parallel factor analysis for fault detection in a semiconductor etch process. *J. Chemometrics* 1999; **13**: 379–396.
5. Wold S, Kettaneh N, Fridén H, Holmberg A. Modelling and diagnostics of batch processes and analogous kinetic experiments. *Chemometrics Intell. Lab. Syst.* 1998; **44**: 331–340.
6. Rothwell S, Martin EB, Morris AJ. Comparison of methods for dealing with uneven length batches with aplication to MPCA monitoring of a batch process. *Proceedings 7th International Conference on Computer Applications in Biotechnology (CAB7)*, Osaka, Japan, June 1998.
7. Kiers HAL, ten Berge JMF, Bro R. PARAFAC2—Part I. A direct fitting algorithm for the PARAFAC2 model. *J. Chemometrics* 1999; **13**: 275–294.
8. Bro R, Anderson CA, Kiers HAL. PARAFAC2—Part II. Modeling chromatographic data with retention time shifts. *J. Chemometrics* 1999; **13**: 295–310.
9. Harshman RA. Foundations of the PARAFAC procedure: model and conditions for an 'explanatory' multi-mode factor analysis. *UCLA Working Papers Phonet.* 1970; **16**: 1.
10. Carroll JD, Chang J. Analysis of individual differences in multidimensional scaling via an *N*-way generalization of 'Eckart–Young' decomposition. *Psychometrika* 1970; **35**: 283–319.
11. Bro R. PARAFAC. Tutorial and applications. *Chemometrics Intell. Lab. Syst.* 1997; **38**: 149–171.
12. *MATLAB 5 User's Guide*. The MathWorks: Natick, MA, 1998.
13. Smilde AK. Comments on three-way analyses used for batch process data. *J. Chemometrics* 2001; **15**: 19–27.
14. Westerhuis JA, Kourti T, MacGregor JF. Comparing alternative approaches for multivariate statistical analysis of batch process data. *J. Chemometrics* 1999; **13**: 397–414.

15. Wise BM, Ricker NL, Veltkamp DJ. Upset and sensor failure detection in multivariate processes. *AIChE Ann. Meet*., San Francisco, CA, November 1989.
16. Wise BM, Veltkamp DJ, Ricker NL, Kowalski BR, Barnes SM, Arakali V. Application of multivariate statistical process control (MSPC) to the West Valley slurry-fed ceramic melter process. *Waste Management '91 Proc*., Tucson, AZ, 1991; **2**: 169–176.
17. Miller P, Swanson RE, Heckler CE. Contribution plots: the missing link in multivariate quality control. *Appl. Math. and Comp. Sci.* 1998; **8**: 775–792.

