REAL TIME STATISTICAL PROCESS CONTROL FOR PLASMA ETCHING

by

Hai-Fang Guo

Memorandum No. UCB/ERL M91/61

2 July 1991

REAL TIME STATISTICAL PROCESS CONTROL FOR PLASMA ETCHING

by

Hai-Fang Guo

Memorandum No. UCB/ERL M91/61

2 July 1991

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

TITLE PAGE

REAL TIME STATISTICAL PROCESS CONTROL FOR PLASMA ETCHING

by

Hai-Fang Guo

Memorandum No. UCB/ERL M91/61

2 July 1991

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

REAL TIME STATISTICAL PROCESS CONTROL FOR PLASMA ETCHING

by

Hai-Fang Guo

ABSTRACT

Process variations in semiconductor manufacturing are responsible for degrading product quality and for limiting the competitiveness of the enterprise. In order to achieve tight control of the manufacturing process, a reliable and sensitive control procedure must be applied. One such approach-- Statistical Process Control (SPC)--is playing an important role in this direction.

Statistical Process Control was introduced over 60 years ago by Walter A. Shewhart. Although it had a tremendous impact on manufacturing, the data available for SPC six decades ago are quite different from the data available now. With the help of automated sensors and computers, real time data can be collected from the process. This change along with others makes the traditional SPC approach unsuitable and necessitates several modifications. Thus even as SPC is transforming IC production, the special needs of semiconductor production are transforming SPC [1].

This report presents a real time multivariate Statistical Process Control scheme that takes full advantage of real-time sensor data, which are collected from a single-wafer plasma etcher via the SECS II communication protocol. The scheme combines a time series model with Hotelling's T^2 statistic and produces a single parameter which carries information about the stability of the process. This parameter can then be applied to a standard process control chart. The method has been applied to a Lam Research Rainbow

Etcher and has successfully detected several introduced faults.

Signature: ______________________

Committee Chairman

Professor Costas J. Spanos

To my parents, my brother and Weijie
for their love, caring, sharing and supporting.

ACKNOWLEDGEMENT

I would like to express my deepest appreciation to my research advisor Professor Costas J. Spanos for the guidance and continuous support during the course of my graduate studies. I also thank Dean David A. Hodges for his valuable suggestions and help.

I would also like to thank Alan Miller of Lam Research for conducting the experiment; Bruce Sindahl of IBM for his insight on time series modeling; Jiahua Huang, Gary May, and the staff of the Berkeley Microfabrication Laboratory for their help with the etching experiments, and Adhi Gaduh for his on-line monitoring program.

It has been a great pleasure to work with my colleagues in BCAM group. Their help and friendship made my research both rewarding and enjoyable. I would like to thank Tom Luan, Sherry Lee, Eric Boskin, Raymond Chen, Gary May, Chris Hegarty, Hao-Cheng Liu, Bart Bombay, Chenny Wang, Sovarong Leang for many insightful discussions. Thanks to the early BCAM members--Dr. Norman Chang, Dr. KK Lin, Dr. Zhiming Lin--for their help and guidance. Special thanks to Hao-Cheng Liu, Sherry Lee and Bart Bombay for proofreading portions of this report.

My greatest gratitude is to my parents for their love, encouragement and support. I am also indebted to my boyfriend Weijie Yun for his understanding, patience and endless support.

This research was sponsored by SRC and IBM.

Table of Contents

List of Figures

Chapter 1 Introduction

1.1 Background and Motivation

To be successful in today's marketplace, a manufacturer must be able to consistently produce high quality, low cost products. Consequently, the Integrated Circuit (IC) manufacturing community is focusing its resources on achieving tight process control over critical manufacturing steps. Many tools and techniques are being used toward this end. Statistical Process Control (SPC) is prominent among them, as it can help in the timely detection of costly process shifts.

Historically, SPC has been used with process measurements in order to uncover equipment and process problems. Such problems are manifested by significant degradation in equipment operation and product quality. To discover this degradation, critical process parameters are monitored using various types of control charts. Traditionally, the measurements consist mainly of in-line readings collected from wafers after the completion of the process step in question. This method is helpful in detecting process drifts; however, it is not timely enough since there is significant delay between the occurrence of a drift and the resulting alarm. As production volume increases, faster response to process drifts becomes necessary in order to assure high product quality and low cost. In addition, obtaining in-line measurements for multi-chamber (cluster) equipment is even more complicated and further reduces equipment reliability and efficiency. Other kinds of data, such as real-time sensor data, are more desirable for quality control purposes [1].

Modern semiconductor manufacturing equipment can communicate internal sensor readings over standard RS232 ports using the SECSII protocol. This capability has been recognized as crucial for the diagnosis of equipment failures, and for the improvement of the overall product quality [2]. Unfortunately, in a high volume production facility the monitoring of multiple sensors results in an overload of information. Further, most of the popular SPC strategies cannot be applied to the real-time sensor readings, since these readings usually show non-stationary, auto-correlated and cross-correlated variations. A special type of SPC procedure is therefore needed to automate the processing of multiple, real time sensor data.

This report describes the development and the application of a novel SPC method used for the control of sophisticated semiconductor manufacturing equipment. This method uses time series filters [3] and multivariate statistics [4] to process internal machine parameters such as the position of the load coil, the tune vanes, the throttle valves, etc. These parameters are sampled several times per second, and the readings are filtered using a time series model. The filtered readings are then combined into a single variable with well defined statistical properties [5]. The value of this variable is calculated every few seconds, and is plotted against formally defined control limits. Real-time misprocessing alarms generated in this manner allow an operator to interrupt faulty runs and prevent any adverse effects on the equipment or the product. This alarm may be then used to initiate automated diagnosis [6].

This method has been applied on a Lam Research Rainbow single wafer plasma etcher, and it has successfully responded to several types of nonstandard process conditions that were introduced in a controlled fashion. These include mismatched RF components, different loading factors, gas leaks, and several miscalibrations of the equipment controls.

1.2 Thesis Organization

The report is structured as follows: Chapter 2 presents an introduction to SPC and some related concepts. Chapter 3 describes the real-time, multivariate SPC approach, which includes the time series model and Hotelling's T^2 statistic. The real-time multivariate SPC procedure has been applied on a Rainbow 4400 etcher and the results are presented in Chapter 5, along with a brief description of the Rainbow Etcher from a process engineer's perspective. Chapter 6 contains a summary and some suggestions for future extensions of this work.

Chapter 2 Traditional Statistical Process Control

2.1 Introduction

The concept of statistical control of a production sequence was introduced in 1924 by Walter A. Shewhart of the Bell Telephone Laboratories [7]. Statistical process control is a collection of methods whose objective is to improve the quality of a process by reducing the variability of its critical parameters. A control chart is among the tools whose function is "*to supply a continuous screening mechanism for detecting assignable causes of variation*" [3].

The basic concepts of SPC are presented in Section 2.2. The traditional SPC tool, the Shewhart control chart, is introduced in Section 2.3, along with the assumptions associated therewith. The applicability of this approach is further discussed in Section 2.4. The concepts introduced in this chapter are used later as the basis of the real-time SPC procedure.

2.2 Basic Concepts

2.2.1 State of Statistical Control

The state of statistical control is defined by Dr. Shewhart's statement that "*a phenomenon will be said to be controlled when, through the use of past experience, we can predict, at least within limits, how the phenomenon may be expected to vary in the future*"[8]. Any process, no matter how well it is designed and how carefully it is maintained, there will be subject to some natural variations due to secondary, but unavoidable causes. But as long as these variations are "nothing but the routine run by run variation," [1] the process is considered to be in the state of statistical control. This means that the future behavior of the process can be predicted within acceptable limits based on the knowledge acquired from past observations from the same process.

2.2.2 Assignable Cause and Chance Cause

When a process is in statistical control, there is only natural variation or "background noise" because of mechanisms known as *chance causes*. Ideally, a process will only be affected by chance causes; sometimes, however, a process can drift away from its normal level of operation, due to *assignable causes* such as significant environmental changes, miscalibrations of a machine, variations of the raw material, or human error. Assignable causes make a process unpredictable or cause it to loose the state of control as defined in Section 2.21. The main purpose of SPC is to detect an assignable cause and remove it as soon as possible in order to prevent quality degradation.

2.2.3 Cause-and-Effect Relationship

There is a cause and effect relationship between the existence of an Assignable Cause and the quality of a manufacturing process. Although critical parameters and process conditions vary for different processes, the output of a process will have its minimum variation when the critical parameters of the process remain constant. Consequently, if a process drift is discovered, then there must be some cause behind it, which can be directly attributed to the process parameters, in other words, "process variables can be utilized in a statistical fashion to estimate the quality of a process" [2]

This cause-and-effect relationship between assignable causes and process quality forms the basis of the traditional statistical process control as well as the real-time multivariate SPC.

2.3 Traditional Control Charts

There are many types of control charts such as the Shewhart control chart, the CUSUM chart, the Geometric Moving-Average (GMA) chart, etc. Below we discuss the Shewhart control chart.

2.3.1 Shewhart Control Chart

A control chart is a graphical representation of a statistical "hypothesis test" that is based on a set of data. The chart requires for its operation information about the routine process conditions, as summarized by the mean and the standard deviation for each of the parameters of interest. These values are estimated from data collected when the process is in statistical control. Figure 1 shows an $\overline{X}$ chart, which is a chart for the arithmetic average and it is usually used in conjunction with a range control chart (R chart) in order to control the location as well as the spread of a process parameter.

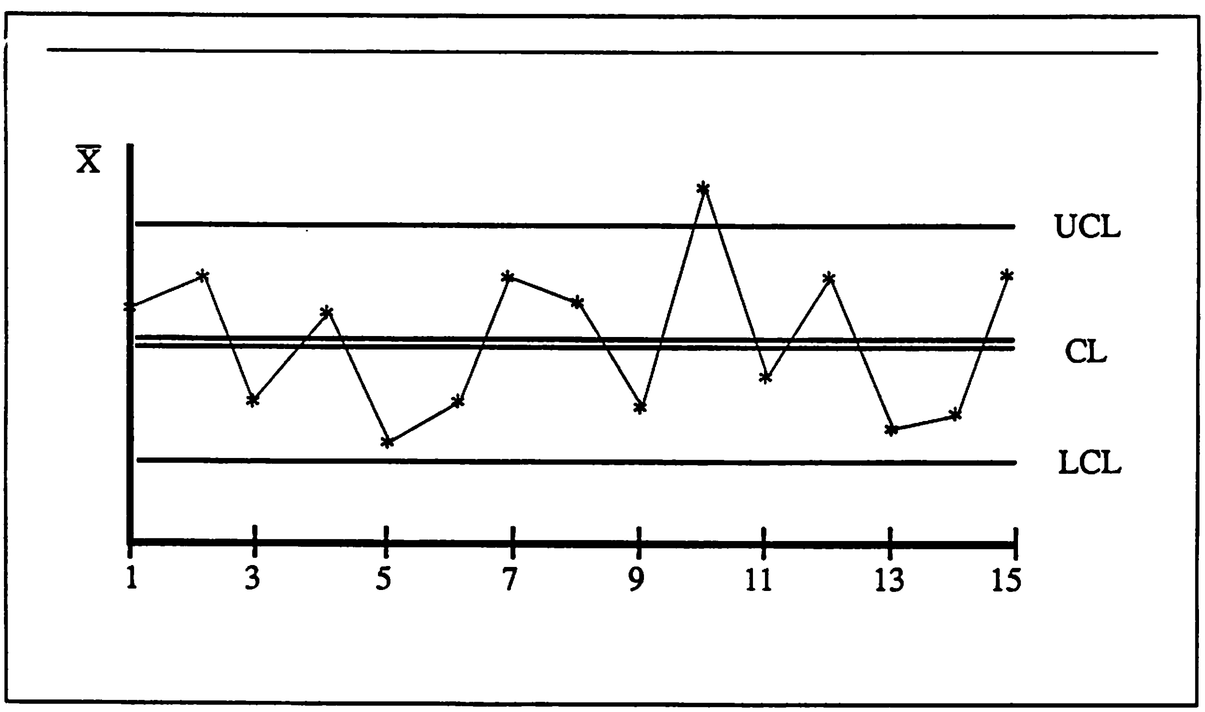


Figure 1. An $\overline{X}$ Chart

2.3.2 Statistical Basis of the Control Chart

The introduction of the Shewhart control chart played an important role in the history of quality control and reliability. It introduced an objective method to detect the significant departure of a product parameter from past observations. From a statistical point of view, SPC casts the decision-making process as a formal hypothesis test. A statistical hypothesis is a statement about the values the parameters generated by a probability distribution [1]. A formal hypothesis test is a very useful tool for many process control problems and also forms the basis for many statistical process control schemes. The *null hypothesis* (H_0) in this context states that the process under consideration is under statistical process control, while the *alternative hypothesis* (H_a) states that the process is out of statistical process control. To test these hypotheses, a random sample x is selected from the population of interest, and the suitable test statistic is calculated. The resulting score is tested against statistically defined limits, which are listed in Equation 1. The range of values that lead to the rejection of a hypothesis is called the *critical region* or the *rejection region*. For the simple Shewhart chart, the limits used to validate H_0 are given below:

$$\begin{aligned} \text{UCL} &= \mu + Z_{\frac{\alpha}{2}}\sigma_{\bar{x}} \\ \text{center line} &= \mu \\ \text{LCL} &= \mu - Z_{\frac{\alpha}{2}}\sigma_{\bar{x}} \\ &\text{where } \alpha \text{ is the type I error} \end{aligned} \tag{1}$$

Where α is the probability of rejecting H_0 by mistake, an occurrence known as the type I error. The probability of accepting H_0 by mistake is know as the type II error. A set of additional rules developed by Western Electric, called the *Western Electric Rules*, helps engineers to effectively interpret control charts [9].

An $\overline{X}$ chart is shown in Figure 2 with the distribution that illustrates its hypothesis testing nature.

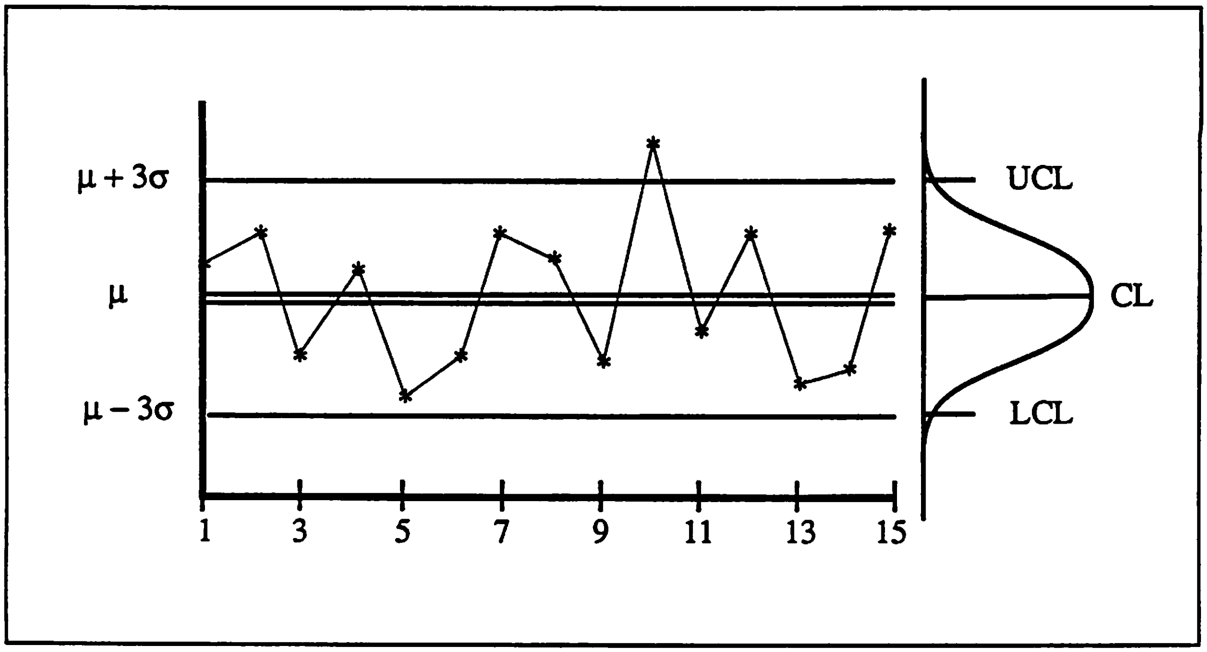


Figure 2. Hypothesis Test

A simple control chart is a device that uses data to test a hypothesis about a population. This discussion leads to next topic of this chapter, which introduces the assumptions needed for the application of a simple control chart.

2.3.3 Assumptions

A simple control chart tests the hypothesis for each of the data points using the same control limits. This means that all the data points under the test are being treated in the same way. Thus all data points must come from the same population. One of the assumptions associated with a simple control chart is that the data must be *Identically, Independently and Normally Distributed.* This is known as the *IIND* assumption. In order for the data to satisfy the *IIND* assumption, they must be consistent with Equation 2. This means that they must be normally distributed around the mean value μ with a standard deviation σ.

$$
\begin{gathered}
X_t = \mu + e_t \\
\text{where } t = 1, 2, \ldots \\
e_t \sim N(0, \sigma^2)
\end{gathered}
\tag{2}
$$

The *IIND* assumption is essential for the simple control chart, and without it the results would not truly reflect the process condition. It is this assumption that prevents us from applying the simple control chart to real-time sensor data, since such data are not *IIND*.

2.4 Problems

The *IIND* assumption covered in last section, restricts the application of the simple control chart. These control charts can only be applied to data points which come from the same population and are independently selected. Unfortunately, this assumption is not valid for real-time sensor data which are most likely to be both auto-correlated and cross-correlated.

To solve these problems, new statistical process control schemes are needed. A real-time multivariate SPC scheme will be presented in the next chapter.

Chapter 3 Real-Time Statistical Process Control

3.1 Introduction

The most popular SPC methods are based on the conventional Shewhart and CUSUM (Cumulative sum) control charts. However, one of the most important assumptions in using these schemes is that the data generated by the in-control process are *Identically, Independently* and *Normally Distributed*. This assumption is known as the *IIND* assumption and it plays a central role in SPC. The *IIND* assumption makes it improper to apply the traditional control charts to the real-time data directly, since real-time data are usually non-stationary, auto-correlated and cross-correlated, in the case of multi-parameter control, even when they originate from a process that is under control.

Modifications are therefore necessary in the traditional SPC scheme. Most modern equipment has an automated data acquisition capability that can generate an avalanche of real-time information. This information must be processed properly and efficiently. As process volume increases, instantaneous detection of process drifts becomes necessary as well.

To accommodate this situation, a SPC scheme is developed and applied to a test process. This scheme employs Econometric Time Series models and Hotelling's T^2 statistic [6]. Time series models transform data into *IIND* signals; Hotelling's T^2 statistic combines the multiple *IIND* signals into a single, well behaved statistical score. This scheme is capable of generating alarms on a real-time basis, and thus is able to prevent misprocessing.

The time series model and Hotelling's T^2 statistic are discussed in Section 3.2.2 and Section 3.2.3, respectively. Section 3.2.4 summaries the scheme and describes its implementation.

3.2 Time Series Modeling

A *model* is an algebraic statement that describes the statistical relationship between a variable of interest (such as an output) and other variables (such as inputs). A *time series model* describes the statistical dependency of the current observation on several previous observations [10].

3.2.1 Time Series Data

A Time series consists of observations generated at regular time intervals. Data collected from most modern semiconductor manufacturing equipment can be classified as a time series. These data can be auto-correlated. Such auto correlation might result by synchronizing readings with control actions. Modern equipment usually uses feedback control mechanisms to control critical parameters, such as the gas flow through the chamber of a plasma etcher. The sensors in the control loop sense the deviation of the parameter from the standard or target value. The controller tends to compensate (offset) the deviation by opening or closing a valve within the mass flow controller. Thus, if sampled on proper intervals, a high reading is likely to be followed by a low value and vice versa. Auto correlation might also result from a high sampling rate. With the help of the SECS II protocol, equipment can collect real-time data up to a rate of several Hz. At this rate, the monitored parameters are subject to "inertia" [11], so a high reading will be followed by another high reading. Figure 3 shows a sample time-series with a periodic pattern. In these two cases, the data cannot be described by the following Equation 2, unless we generalize it as follows:

$$\begin{gathered} X_t = f(X_{t-1}, X_{t-2}, \ldots, e_{t-1}, e_{t-2}, \ldots) + e_t \\ \text{where } t = 1, 2, \ldots \\ e_t \sim N(0, \sigma^2) \end{gathered} \tag{3}$$

The auto-correlation coefficient describes the statistical dependence of one observation on previous observations and takes values between 0 and 1. A zero value will be obtained when the observation of interest is independent from other observations, while a value of one indicates complete dependence. The following equation defines the auto-correlation coefficient between two readings that are k sampling intervals apart. This coefficient has been calculated using N consecutive observations.

$$r_k = \frac{\sum_{r=1}^{N-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^{N} (Z_t - \bar{Z})^2} \tag{4}$$

Where $k = 1, 2, \ldots$

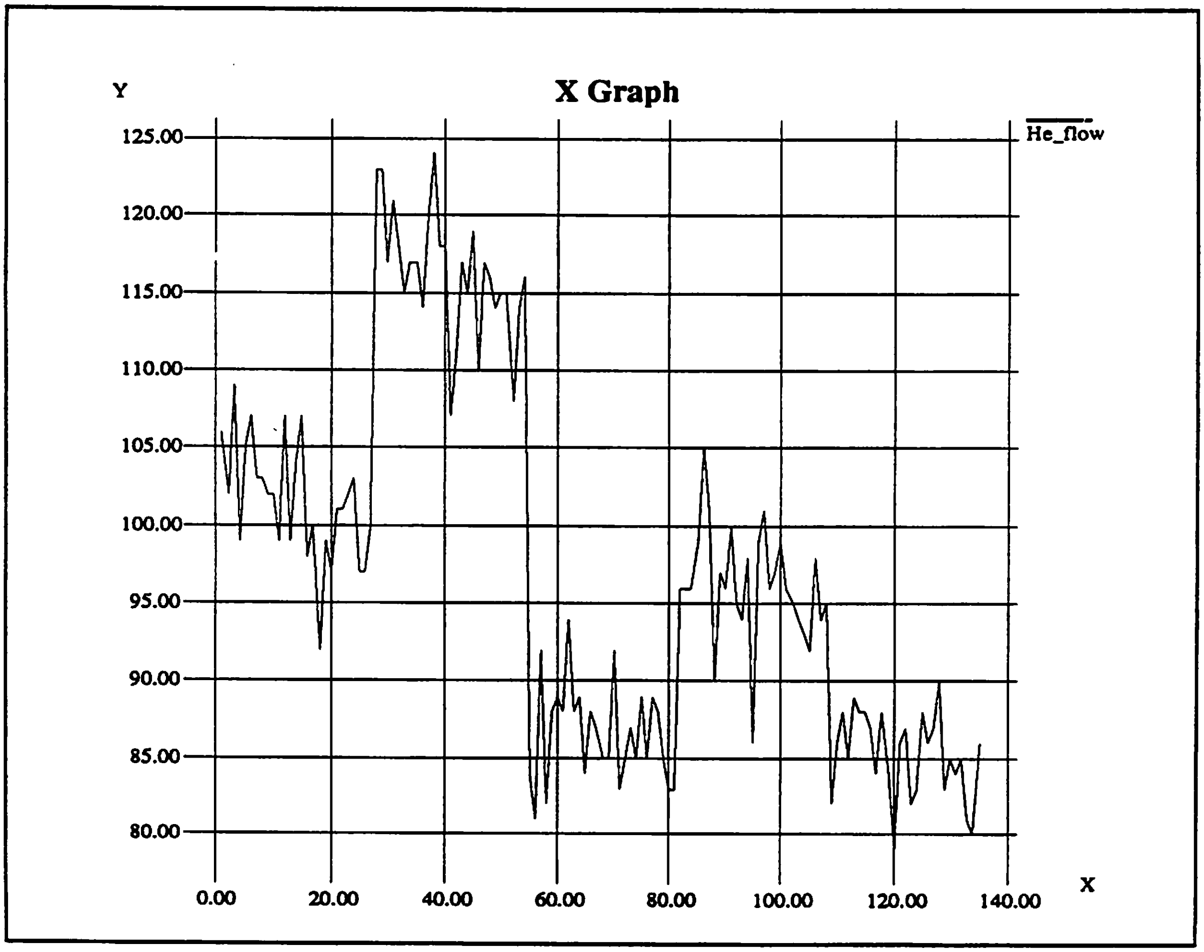


Figure 3. Sample Time-Series Data With Periodic Pattern

3.2.2 Objectives of Time Series Modeling

For the two cases described in the last section, the data are auto-correlated; thus the observations in the time series are not completely random. In fact, an observation is statistically related to the previous observation and can be predicted with certain accuracy, if this statistical relationship is known. The objective of time series modeling is to find this statistical relationship and use it to achieve [12]:

- Description - describe the features of a time series process.
- Explanation - infer the structural rules of behavior.
- Prediction - forecast future readings.
- Control - investigate the effects of changing model parameters.

In this project, the main purpose of the time series modeling is to find suitable models to filter real-time data used for statistical process control. The methods used to obtain the models are discussed in the next section.

3.2.3 Approach - Univariate Box-Jenkins Analysis

In this application, we employ univariate Box-Jenkins time series analysis (named after George E. P. Box and Gweilym M. Jenkins who introduced this procedure). The reason for using only univariate analysis is that the time-series behavior of one parameter can be explained by using only the past observations of this parameter. A Box-Jenkins time series model is also called an ARIMA(p, d, q) model, since it consists of three components (or filters) as illustrated in Figure 4. These components are the *auto-regressive* part (of order p), the *integration* part (of order d), and the *moving-average* part (of order q) [5].

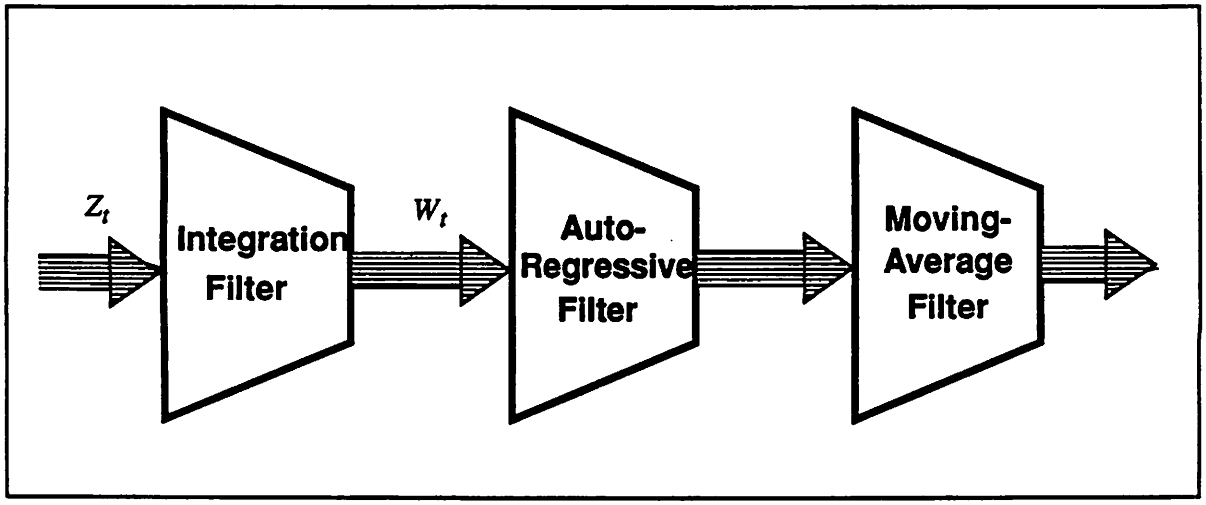


Figure 4. The Three Components of the ARIMA Model

The general form of the ARIMA(p, d, q) model is given below:

$$\begin{aligned} \phi(B)\,w_t &= \theta(B)\,a \\ \phi(B) &= 1-\phi_1 B-\phi_2 B^2-\dots-\phi_p B^p \\ \theta(B) &= 1-\theta_1 B-\theta_2 B^2-\dots\ \theta_q B^q \\ w_t &= \nabla^d z_t \qquad \text{where } (d \ge 0) \end{aligned} \tag{5}$$

Difference Operator: $\nabla z_t = z_t - z_{t-1}$

Backward Shift Operator: $B z_t = z_{t-1}$

where z_t is the original data, w_t is the differentiated signal and a is the *IIND* residual.

A condition for applying a time series model is that the signal must be stationary. This means that the mean, variance and autocorrelation function of the time series must be constant through time. The integration component of the ARIMA model as described above is used to convert a non-stationary data series to a stationary one, as the auto regressive and moving average models can only be applied to a stationary series. The differentiation can be used to achieve constant mean; taking the log or square root of the data might also be necessary in order to produce a constant variance.

The second part of the ARIMA model is the *auto-regressive* (AR) part, which describes the dependency of the current observation on previous observations, through the parameters ϕ_i, i = 1,..., p.

The third part of the ARIMA scheme is the *moving-average* (MA) part, which describes the dependency of the current observation on previous random shocks, by means of the parameters θ_j, j = 1,... q.

Sometimes the original data show seasonal periodic patterns - similar to the data shown in Figure 3. These patterns can be modeled by creating ARIMA models for the seasonal variation as well as for the individual samples. The composite model is known as a *Seasonal* ARIMA model or SARIMA(p, d, q)x(P, D, Q)$_S$, where p is the number of significant autocorrelations, d is the number of differentiations, q is the number of significant random shocks within each season, and P, D, Q are the autocorrelations, differentiations and moving average terms, taken across seasons of duration s [14]. The complete SARIMA(p, d, q)x(P, D, Q)$_S$ is shown in Equation 6.

$$\phi(B)\Phi(B^S)w_t = \theta(B)\Theta(B^S)a_t \tag{6}$$

A model must be obtained from the collected data when the process is under control; in this way the model describes the "good" process. Once a model has been developed, it can be used to *forecast* (or predict) each new value. The difference between the forecast value and the actual value is the forecasting error, which is, by definition, an *IIND* (*Identically, Independently,* and *Normally Distributed*) variable. The *residual* or *random shock* is the difference between the real-time data and the value predicted by the model:

$$\text{residual: } a_t \equiv Y_t - \hat{Y}_t \tag{7}$$

3.2.4 Modeling Procedure

In order to assist the user in obtaining a "good" ARIMA model, Box and Jenkins proposed a 3-step procedure. The procedure is outlined in the following figure [10].

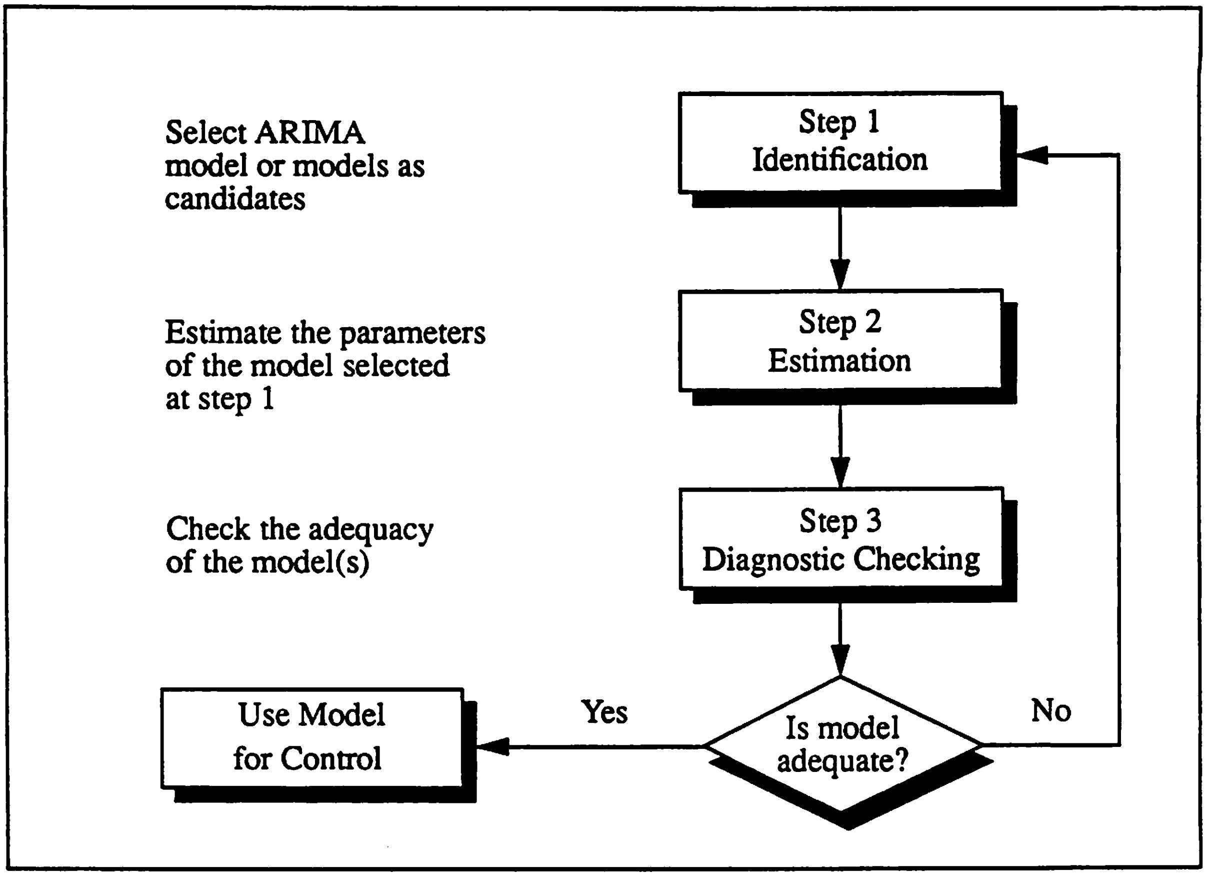


Figure 5. The 3-Step Procedure for ARIMA Modeling

Two devices are used to select the ARIMA models -- the autocorrelation function (acf) and the partial auto-correlation function (pacf). Both are calculated from the data and are compared with the theoretical acf and pacf patterns from known model structures. Once the structure of the model is established, the parameters of the selected models are estimated at the estimation step. At the diagnostic and checking steps, similar devices are used to check the adequacy of the selected model. If the models are not statistically adequate, the procedure is repeated. The procedure terminate when a satisfactory model is obtained.

3.3 Hotelling's T^2 Statistic

3.3.1 Correlated Data

While auto-correlation was used as a measure of the dependence of a variable on its past, *correlation* measures the dependency of one variable on the *other* variables. If several sensors are employed at the same time, their readings are very likely to be cross-correlated. For example, pressure readings of a plasma etcher are likely to be related to gas flow readings in the same chamber. Such a chamber pressure is plotted against the helium flow in Figure 6. The plot shows the cross-correlation between the two parameters in one particular experiment.

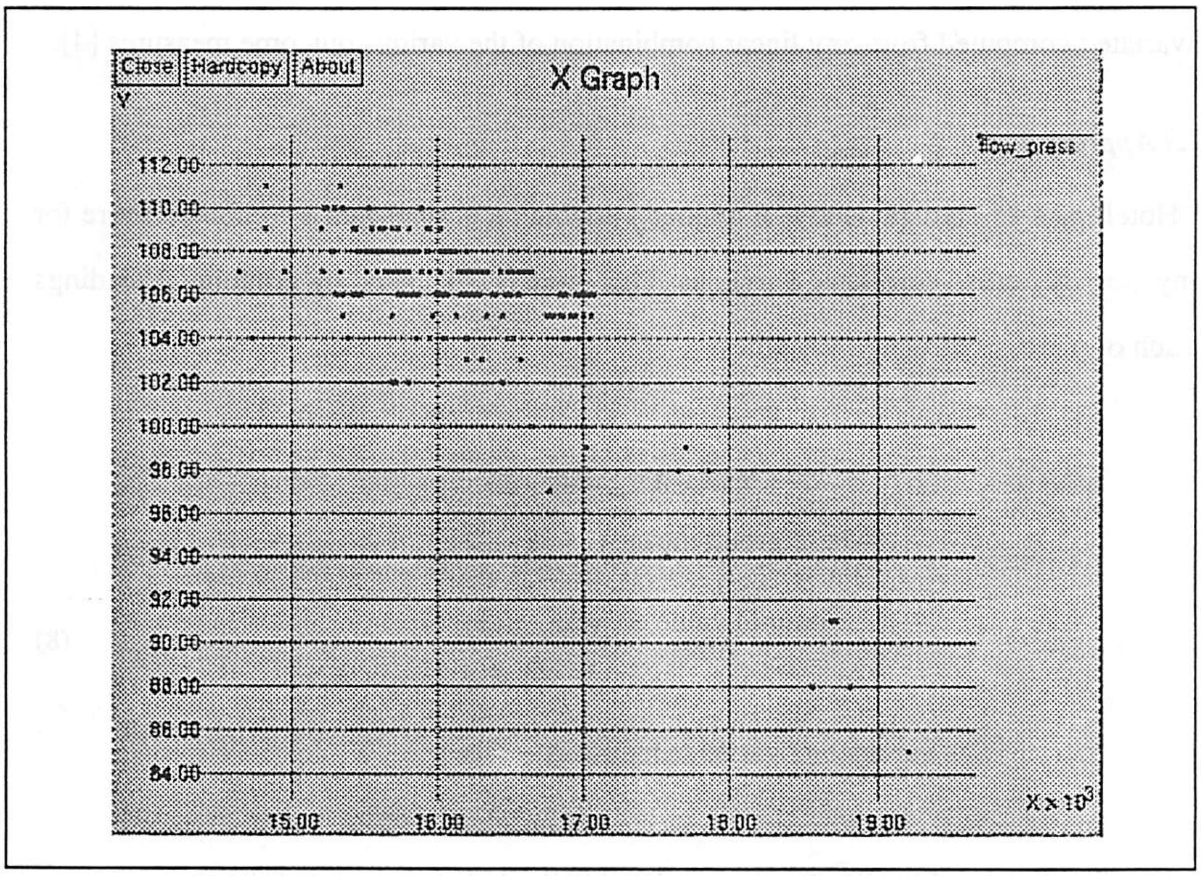


Figure 6. Sample cross-correlated data: chamber pressure plotted against He flow

3.3.2 Objective

Because multiple readings tend to be correlated, using a number of separate control charts can be misleading. It has been shown that as the number of correlated variables increases, the probability of generating false alarms from a control procedure that uses a large number of separate charts grows significantly [7]. Also from a process engineer's point of view, parameters that are tightly related should not be interpreted separately. In the case of plasma etching, for example, forward and reflected RF power readings are very closely related and treating them separately can lead to underestimation of type I and type II error rates.

The objective of Hotelling's T^2 statistic is to combine several cross-correlated variables into a single statistical score, which is simply the square of the maximum possible univariate t computed from any linear combination of the various outcome measures [4].

3.3.3 Approach

Hotelling's T^2 statistic is a well defined variable that represents a combined score for many, possibly cross-correlated variables. This score is calculated by grouping n readings of each of p cross-correlated variables:

$$\begin{gathered} T^2 = n(\bar{X} - \tilde{X})^T S^{-1} (\bar{X} - \tilde{X}) \\ \text{where group mean } \bar{X}^T = \begin{bmatrix} \bar{x}_1 & \dots & \bar{x}_p \end{bmatrix} \\ \text{nominal value } \tilde{X}^T = \begin{bmatrix} \tilde{x}_1 & \dots & \tilde{x}_p \end{bmatrix} \\ \text{variance-covariance matrix } S = \begin{bmatrix} s_1^2 & \dots & s_{1p} \\ \dots & \dots & \dots \\ s_{p1} & \dots & s_p^2 \end{bmatrix} \end{gathered} \tag{8}$$

The distribution of the T^2 statistic is related to the F distribution as follows:

$$T^2_{\alpha,p,n-1} = \frac{p(n-1)}{n-p} F_{\alpha,p,n-p} \tag{9}$$

This statistic takes a low value when the cross-correlated structure of the underlying variables remains constant. The T^2 score is very sensitive to any change in one or more of the variables. This score can be used in conjunction with a one-sided control chart, whose limit is set according to the number of variables, the sample size and the acceptable false alarm rate.

3.4 Summary and Implementation

The real-time SPC scheme takes multiple sensor data that are auto-correlated and cross-correlated, and then feeds them into individual time series filters that result in multiple, cross-correlated IIND residuals. The Hotelling's T^2 filter then combines the residuals into a single real-time alarm signal. This alarm signal can be used to initiate the BCAM diagnostic module [5]. This scheme is shown in Figure 7 [15].

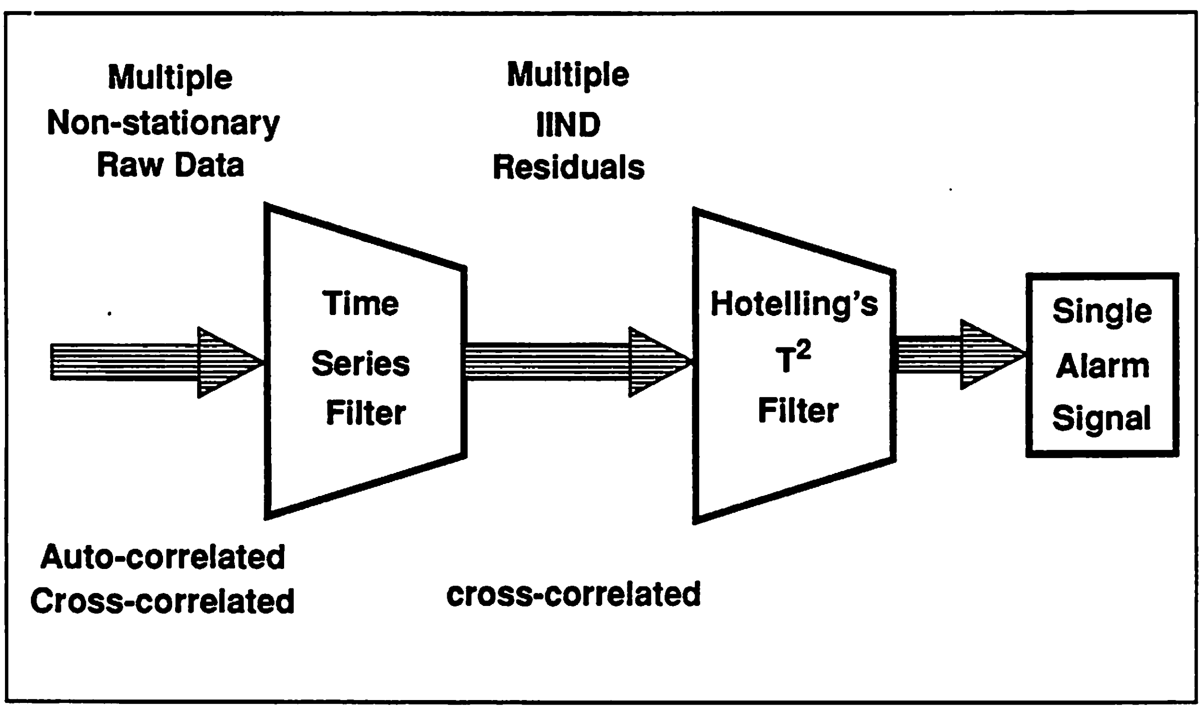


Figure 7. Summary of the real-time SPC scheme

A software package has been developed to implement this real-time SPC scheme. It includes three modules -- data manipulation, ARIMA modeling, and Hotelling's T^2 calculation and alarm generation. It is implemented both in SAS [16] and RS/1 [17]. Portions of the SAS code are attached in Appendix 3.1 - 3.3.

Chapter 4 Application Examples

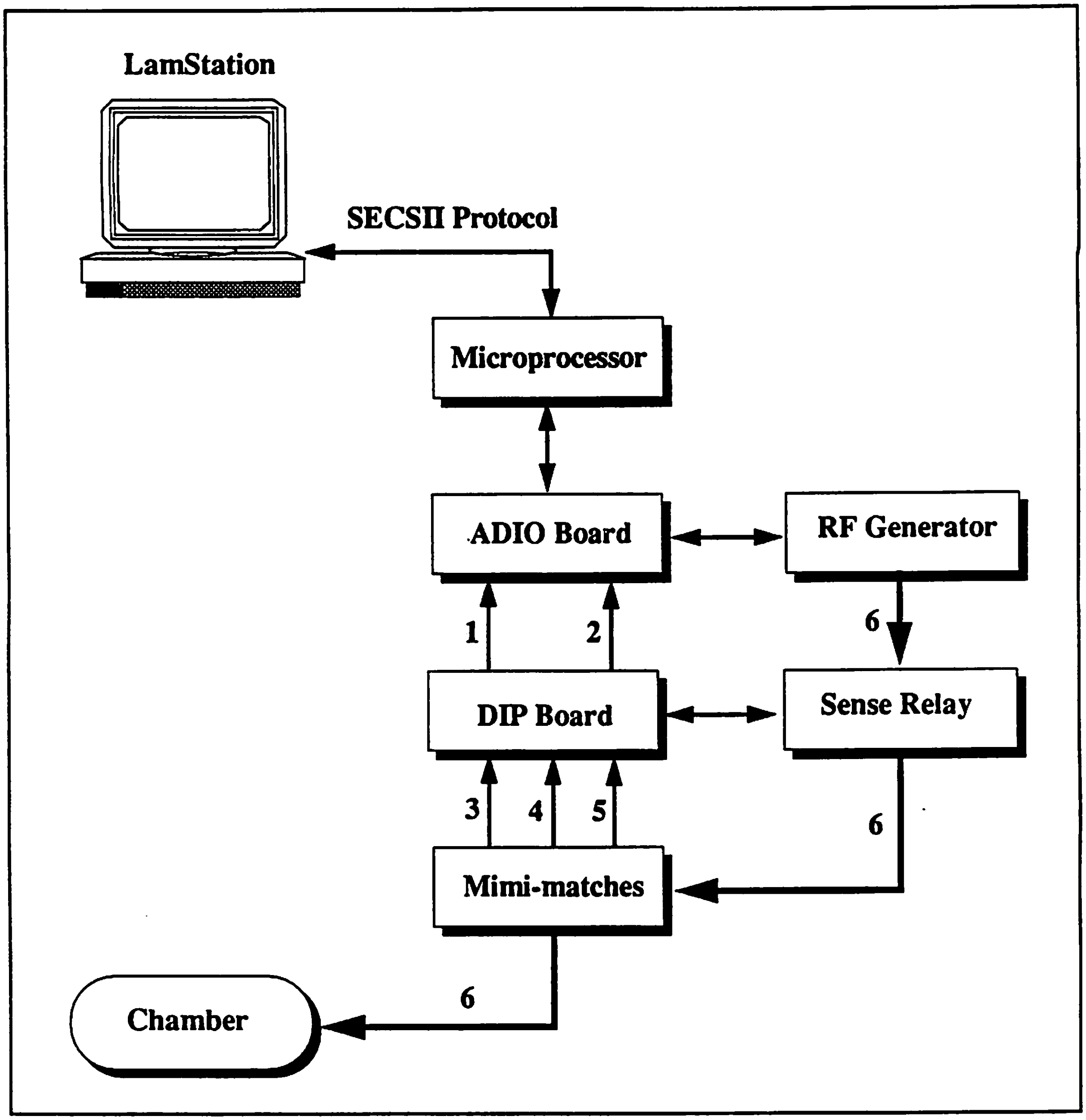
4.1 Introduction

Experiments are conducted to test the applicability of the real-time SPC scheme. To date it has been successfully applied on a Lam Rainbow plasma etcher, and it has been able to detect internal machine shifts that cannot be seen with the classical SPC procedures using wafer measurements.

4.2 The Test Vehicle - the Rainbow Etcher from Lam Research

The Lam Research Rainbow etcher is a state-of-the-art single wafer tool. There are three basic Rainbow machines, with applications for polysilicon, oxide, and metal films. The equipment provides automated processing of 4" through 8" wafers. For the experiment reported here, 6" inch patterned polysilicon wafers were etched using a Cl_2-based polysilicon etch recipe [18].

Through the SECS II protocol link a remote host can communicate directly to the Rainbow in order to acquire real-time analog data. The data conversion is done using the LamStation software package provided by Brookside Software Co. [19]. Up to 32 separate parameters can be sampled simultaneously with rates of up to 3 Hz. For this experiment, we monitored signals from the RF network because we found them to be very responsive to small process changes. The major functional components of the Rainbow RF network are shown in Figure 8 [20].



1	RF Phase Error
2	RF Impedance
3	RF Load Coil Position
4	RF Tune Vane Position
5	RF Peak to Peak Voltage
6	RF Coax Cable

Figure 8. Schematic of the RF Network

4.3 The Experiment

Two phases of an experiment designed to test the real-time SPC procedure have been completed to date.

The first phase of the experiment had two objectives. The first objective was to select the proper (most sensitive) parameters from all the available sensor readings and to find the proper time series models for the parameters. The second objective was to detect the introduced fault. This experiment was conducted at Lam Research's Qualifications Lab using a Rainbow etcher. First, the machine was calibrated to its "best" operating condition by running over one hundred wafers. Four polysilicon wafers are processed afterwards. These four wafers formed the *baseline* wafers, and as such they were used to select and characterize the appropriate time series models, and later to estimate the means and the variance-covariance matrix of the residual for the Hotelling's T^2 calculation. Following the baseline wafers, one wafer with a different loading factor - photoresist instead of polysilicon, was processed.

During the second phase of the experiment, the objective was to refine the time-series models, and to test the scheme by introducing, one by one, a number of "faults" into the process. Again, over one hundred wafers were processed to set up the proper process condition. Eight baseline wafers were etched before introducing faults. For each fault two wafers were processed, and the machine was then switched back to the baseline conditions. Two additional baseline wafers were processed before introducing the next fault. Several faults were introduced, including the replacement of several RF components with miscalibrated ones, miscalibrated recipe settings for the electrode gap, the pressure, etc. Chamber contamination was also introduced by not cleaning the chamber at the recommended intervals.

The main steps in establishing the control scheme during each of the two experimental phases are described in the following section.

4.4 Selecting the Relevant Sensor Readings

Originally, we collected readings from more than 25 parameters. Applying all of them into the scheme, however, proved to be too complicated and also unnecessary, since only some of them carried useful information. The criterion for selecting the relevant sensor readings was that the parameter must have some physical significance, and it also must be suitable for time series modeling. This meant that after applying a reasonably simple SARIMA model to the parameter readings, the resulting residuals should be *IIND*.

Five parameters were finally selected. They were the position of the RF tune vane, the position of RF load coil, the RF phase error, the plasma impedance, and the peak-to-peak voltage across the electrodes. As shown next, the statistical behavior of these readings conveys a comprehensive picture of the etching conditions.

Other readings of significance, such as the RF power, the chamber pressure, or the gas flows, were not used. We found that, since these parameters were actively controlled by the machine to fixed set points, their readings were very insensitive to internal machine changes.

4.5 Fitting the SARIMA Model

To build the SARIMA model we used the real-time data from the baseline wafers. The readings shown in Figure 9 were collected while processing the baseline wafers for the first experiment. It is obvious that the readings are not stationary, and that they have a seasonal pattern that repeats during the processing of each wafer. The seasonal component is evident from the patterns that can be seen repeating in each of the four baseline wafers.

Having decided to use a SARIMA model, it is important to select the appropriate model structure for each of the five parameters. Using the SAS statistical package, this task was accomplished by trial and error, until the reading residuals were IIND. The model structure selected was the SARIMA(0,1,1)x(1,1,0)$_{60}$, as listed in Equation 10. Although we used different coefficient values Φ_1 and θ_1 for each of the five parameters, the same structure was applicable to all five.

$$a_t = (Y_t - Y_{t-1}) - (Y_{t-60} - Y_{t-61}) - \Phi_1 x \\ [(Y_{t-60} - Y_{t-61}) - (Y_{t-120} - Y_{t-121})] + \theta_1 a_{t-1} \tag{10}$$

where Φ is the seasonal AR coefficient

θ is the MA coefficient

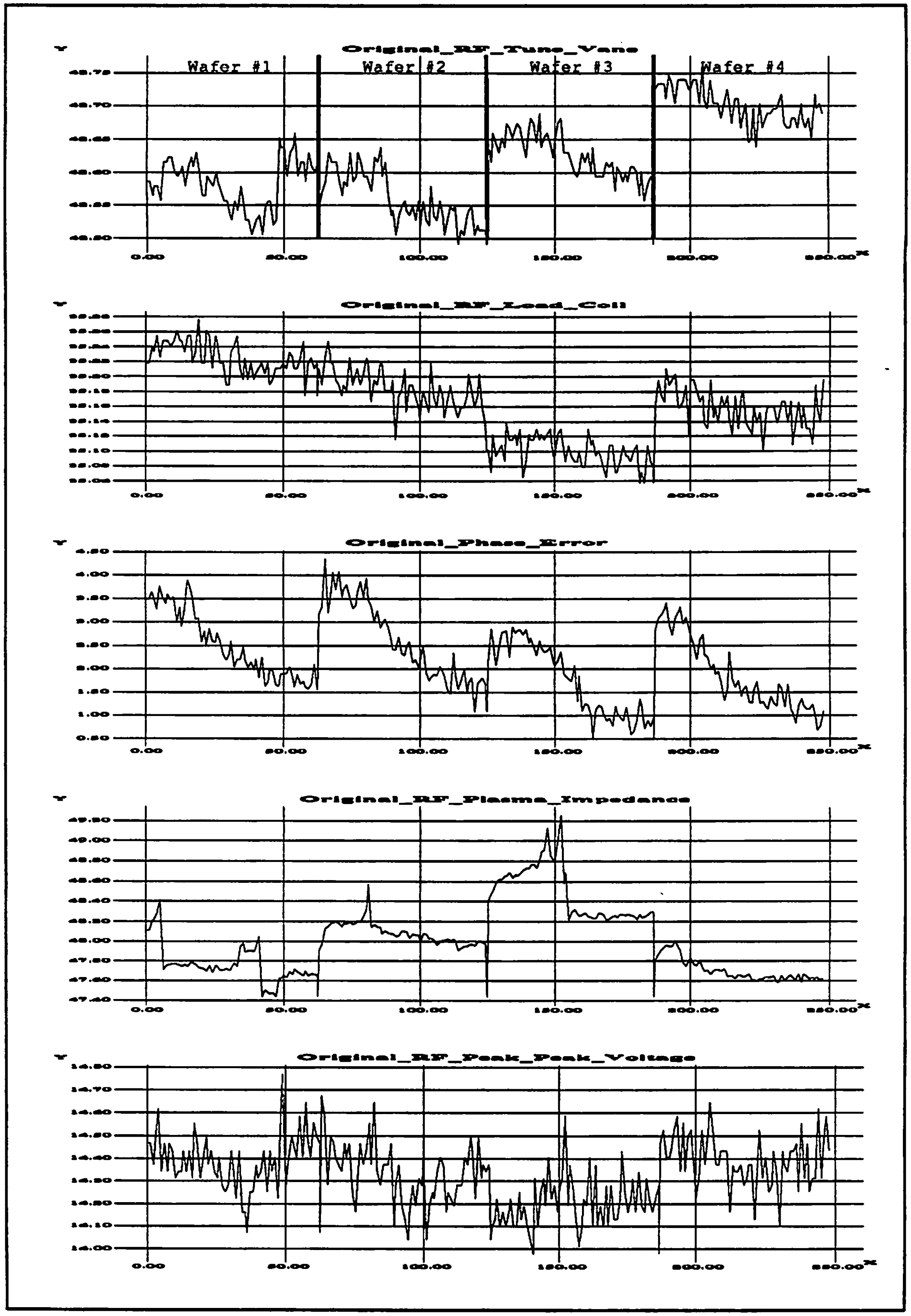


Figure 9. The Original Real-time Data for the Four Baseline Wafers

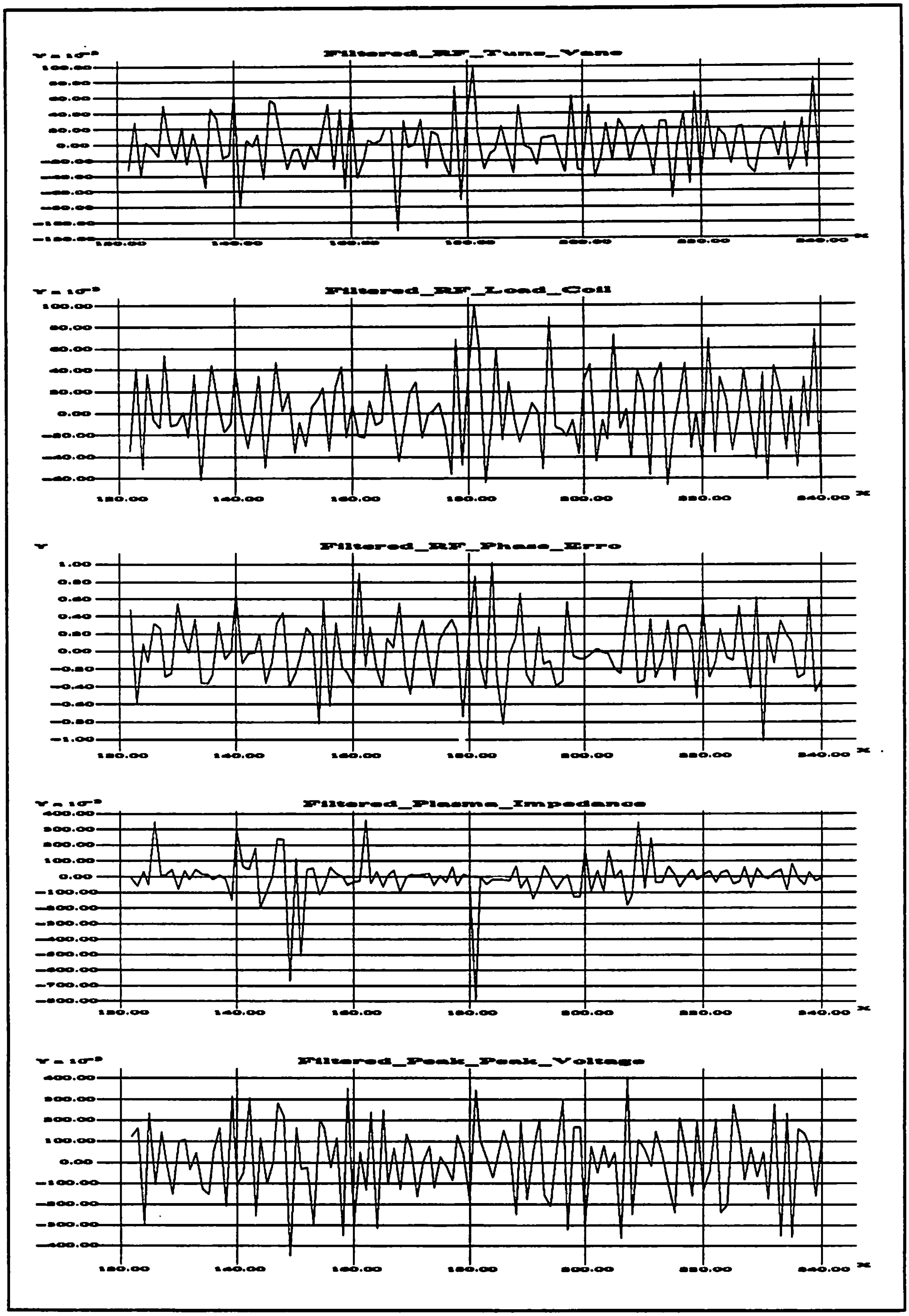


Figure 10. The *IIND* Residuals for the Four Baseline Wafers

The *residuals* a_t, defined as the real-time data minus the fitted SARIMA model as shown in Equation 7 of Chapter 3 are *IIND*; thus they can be used in traditional SPC procedures. The residuals for the five observed parameters are plotted in Figure 10.

4.6 Applying Hotelling's T^2 Statistic

In order to use Equation 8 of Chapter 3 to calculate the T^2 statistic, it is necessary to estimate the mean values and the variance-covariance matrix for the five residuals. These values have been obtained from the baseline wafer data using SAS. The T^2 statistic can be applied to a Hotelling T^2 control chart. This single-sided chart can then respond to process drifts by generating an alarm, whenever the T^2 score takes values beyond the upper control limit.

4.7 Experimental Results

4.7.1 First Phase of the Experiment

The T^2 statistic obtained from the first phase of the experiment is plotted in Figure 11.

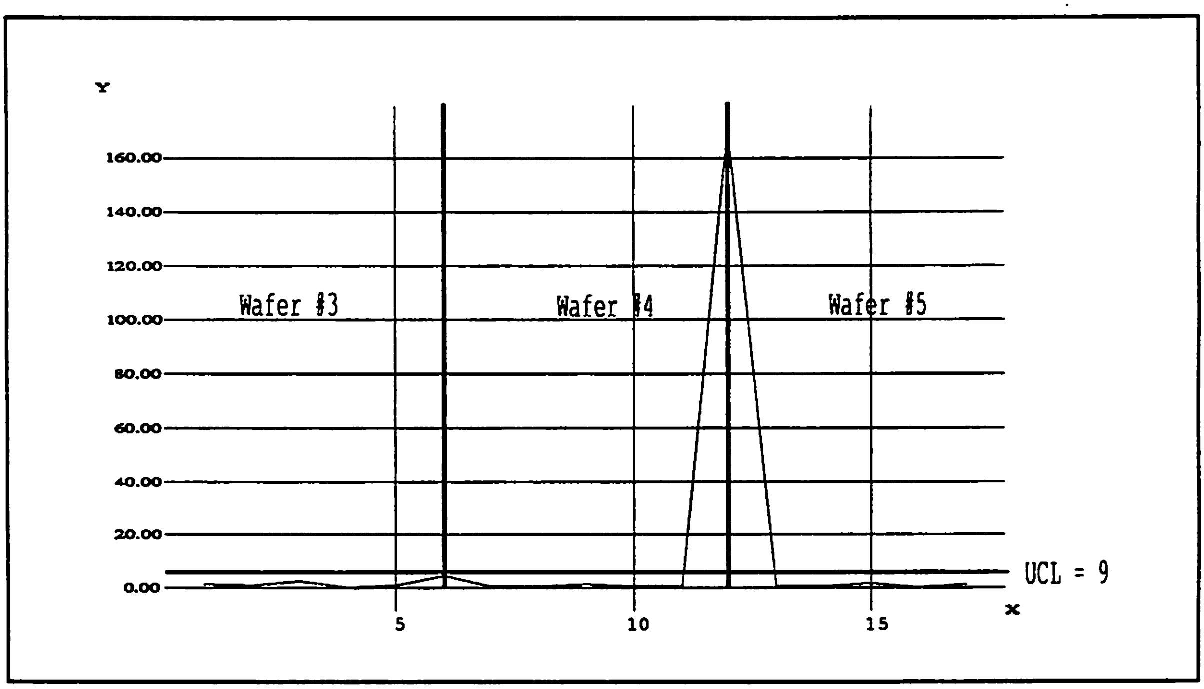
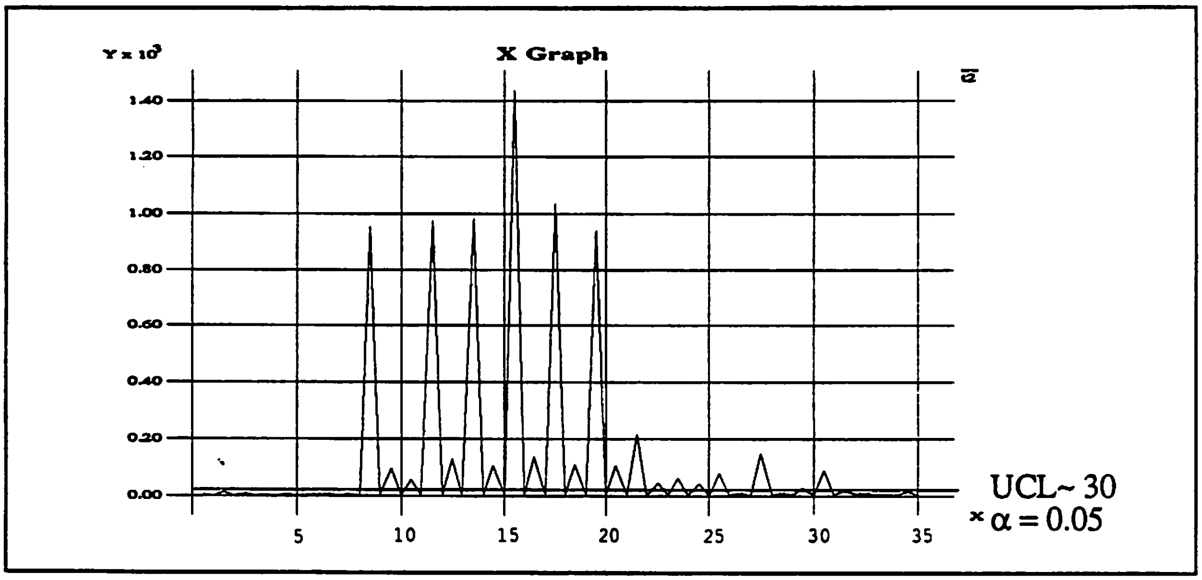


Figure 11. The T^2 Control Chart for the First Phase of the Experiment

Since the fifth wafer had poorly developed photoresist on top of the polysilicon, the loading factor is different from that of the baseline wafers; thus the plasma parameters are different. This process change was detected by the T^2 control chart by generating a score that easily exceeds the UCL of it.

4.7.2 Second Phase of the Experiment

The results for the second experiment were also conclusive, as shown below in Figure 12.



Wafer 1- 7	Baseline wafers
Wafer 8 - 10	Mini-Match #1
Wafer 11-12	Mini-Match #3
Wafer 13-14	Mini-Match #2
Wafer 15-16	DIP Board #2
Wafer 17-18	DIP Board #1
Wafer 19-20	Mini-Match #4
Wafer 21-22	Miscalibrated Gap Distance
Wafer 23-24	Miscalibrated Pressure
Wafer 25-26	Miscalibrated Flow
Wafer 27-28	Sense Relay #2
Wafer 29-30	Sense Relay #3
Wafer 31-32	Chamber Contamination
Wafer 33-34	Out-Off Specification Cable

Figure 12. The T^2 Control Chart for the second phase of the Experiment

The faults associated with the alarms are shown on the plot. This scheme is able to pick up most of the process faults that we have introduced. The significance of this scheme is that it can detect very slight process changes which might affect the performance of the machine, but cannot be seen on traditional control charts of etch rate or uniformity. Such control charts are shown in Figures 13 and 14.

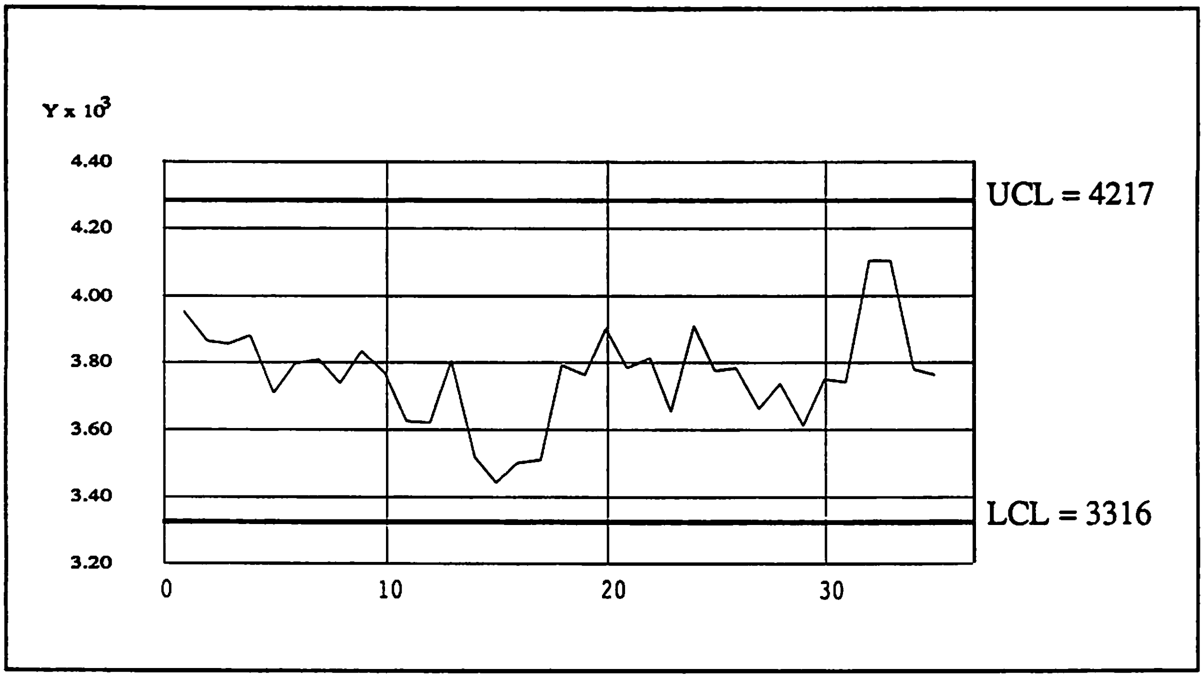


Figure 13. The Etch Rate Control Chart.

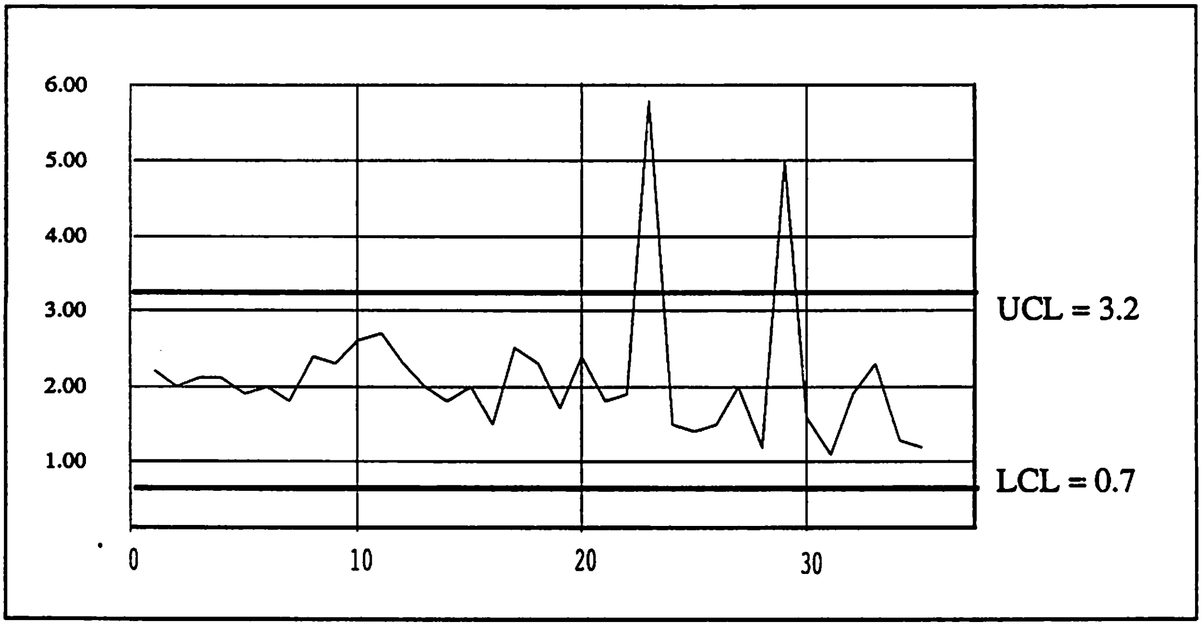


Figure 14. The Etch non-Uniformity Control Chart

The two alarms observed in the non-uniformity chart were traced to wafer problems that were unrelated to our experiment.

Chapter 5 Conclusions and Future Plans

We have presented a novel application of real-time statistical process control for monitoring state-of-the-art semiconductor manufacturing equipment. Through actual monitoring examples, we have shown that this technique can successfully flag internal machine variations long before their effects can be seen on the wafers.

Because of the ability of this technique to observe changes within the internal operation of the machine, we also expect it to be useful in driving an automated diagnostic package that, in real-time, would be able to supply more information to the operator. Adjustments, maintenance scheduling, and further investigation might be planned by the operator based on this information.

Although not reported here, a third experiment is currently under way. Its objective is to quantify the sensitivity of our procedure. To this end, we have introduced miscalibrations of varying sizes and are currently analyzing the results.

The T^2 charts that were presented here have been obtained off-line. Currently, we are working on the real-time implementation of this scheme. Since the only operations that are taking place in real-time involve the evaluation of the SARIMA forecasting error and the computation of the T^2 score, we feel that a workstation with moderate capabilities will be able to easily monitor up to 3 samples/second. Operation beyond this speed, will be limited by the bandwidth of the SECS-II serial port.

We are currently implementing this scheme as part of the Berkeley Computer-Aided Manufacturing system. BCAM is a Unix-based CAM framework that supports monitoring, modeling, real-time and in-line model-based SPC, diagnosis and feed-forward feedback control. This system is being implemented on Sun4 and IBM RS6000 workstations for the concurrent control of multiple equipment operating alone or in workcell configurations.

Appendix 3.1

SAS Implementation of the Data Manipulation Module

```
/**;
;;; SPC module of BCAM ;
;;;
;;; Copyright (c) 1989 Regents of the University of California ;
;;;
;;; Permission to use, copy, modify, and distribute this software and its;
;;; documentation for any purpose and without fee is hereby granted, ;
;;; provided that the above copyright notice appear in all copies and ;
;;; that both that copyright notice and this permission notice appear in ;
;;; supporting documentation, and that the name of the University of ;
;;; California not be used in advertising or publicity pertaining to ;
;;; distribution of the software without specific, written prior ;
;;; permission. The University of California makes no representations ;
;;; about the suitability of this software for any purpose. It is ;
;;; provided "as is" without express or implied warranty. ;
;;;
;;; Author: Hai-Fang Guo ;
;;; $Revision: 1.0 $ ;
;;; $Date: 91/05/17 17:04:01 $ ;
**/;


*** USED TO SELETE THE WINDOW FOR ANALYSIS -- THE ETCHING STEP ;
*** USE THE RF FORWARD POWER AS STANDARD TO SELECT THE WINDOW;
*** IN THIS CASE, C14;

 DATA PARTSUB ;
 SET SASDATA.TOTAL ;
 RETAIN IND 0;
 POWER=270;
 HIGH=1.05*POWER;
 LOW =0.95*POWER;
 IF TEST="MATCH2" OR TEST ="BASE4" OR TEST="BASE5" OR TEST="BASE6"
 OR TEST="CONTAM";
 IF TYPE ="POLY";
 IF C14<LOW OR C14>HIGH THEN DELETE;
 TIMEDIFF=C1-LAG1(C1);
 IF TIMEDIFF>50 THEN IND=0;
 ELSE IND=IND+1;
 IF WAFER^=LAG1(WAFER) THEN IND=0;
 REC1=LAG1(REC);
 DIFF=REC-REC1;
 IF DIFF^=1 THEN DELETE;
 DROP C6 C7 C9 C11 C12 C13 C25 C26 C29;
 RUN;


*** USED TO CHOOSE THE RIGHT PART OF THE WINDOW FROM 4 TO 63 ;

 DATA SUB;
```

```
SET PARTSUB;
RETAIN I 0;
RETAIN MINC2 2000 MINC3 2000 MINC4 2000;
RETAIN MINC8 2000 MINC10 2000 MINC16 2000;
RETAIN MINC18 2000 MINC19 2000 MINC20 2000;
RETAIN MINC20 2000 MINC21 2000 MINC22 2000;
RETAIN MINC23 2000 MINC24 2000;

IF MINC8>C8 THEN MINC8=C8;
IF MINC10>C10 THEN MINC10=C10;
IF MINC16>C16 THEN MINC16=C16;
IF MINC18>C18 THEN MINC18=C18;
IF MINC23>C23 THEN MINC23=C23;
IF MINC22>C22 THEN MINC22=C22;
IF MINC24>C24 THEN MINC24=C24;

LAG1C16 = LAG1(C16);
DIFF1 = C16 - LAG1C16;
LAG22C16 = LAG22(C16);
DIFF22 = C16 - LAG22C16;

IF IND<4 OR IND>23 THEN DELETE;
I = I + 1;
RUN;


*** USED TO CHECK THE WINDOWED DATA ;
PROC PRINT DATA= SUB;
VAR TEST WAFER IND C16 LAG1C16 DIFF1 DIFF22 C18 C20 C21 C22 ;
RUN;

*** USED TO OUTPUT THE SELECTED VARIABLES TO THE OUT.TEST FILE ;
*** FOR OUTPUT FILES;

CMS FILEDEF OUT16 DISK OUT16 TEST;
DATA _NULL_;
SET SUB;
FILE OUT16 ;
INDEX=_N_;
MARK=INDEX-0.5;
IF IND=1 THEN PUT MARK MINC16;
PUT INDEX C16;
RUN;
```

Appendix 3.2

SAS Implementation of SARIMA Modeling

```
/**;
;;; SPC module of BCAM ;
;;;
;;; Copyright (c) 1989 Regents of the University of California ;
;;;
;;; Permission to use, copy, modify, and distribute this software and its;
;;; documentation for any purpose and without fee is hereby granted, ;
;;; provided that the above copyright notice appear in all copies and ;
;;; that both that copyright notice and this permission notice appear in ;
;;; supporting documentation, and that the name of the University of ;
;;; California not be used in advertising or publicity pertaining to ;
;;; distribution of the software without specific, written prior ;
;;; permission. The University of California makes no representations ;
;;; about the suitability of this software for any purpose. It is ;
;;; provided "as is" without express or implied warranty. ;
;;;
;;; Author: Hai-Fang Guo ;
;;; $Revision: 1.0 $ ;
;;; $Date: 91/05/17 17:04:01 $ ;
**/;


*****************************************************************;
*** ARIMA ANALYSIS -- TO FIND ONE MODEL FOR EACH PARAMETER ;
*****************************************************************;
PROC ARIMA DATA=SUB;
IDENTIFY VAR=C16(1,20) NLAG=70 CENTER CLEAR ;
ESTIMATE P=1 Q=(0)(20) PLOT GRID ;
FORECAST OUT=B1 BACK=152 LEAD=160 ID=I ;

IDENTIFY VAR=C18(1,20) NLAG= 70 CENTER CLEAR ;
ESTIMATE Q=(0)(20) PLOT GRID;


IDENTIFY VAR=C20(1,20) NLAG=170 CENTER CLEAR ;
ESTIMATE Q =(0)(20) PLOT;


IDENTIFY VAR=C21(1,20) NLAG=170 CENTER CLEAR ;
ESTIMATE Q=(0)(20) PLOT;


IDENTIFY VAR=C22(1,20) NLAG=170 CENTER CLEAR ;
ESTIMATE P= 1 Q=(0)(20) PLOT;


PROC PLOT DATA=B1;
PLOT FORECAST*I='F' C16*I='*' L95*I='L' U95*I='U'
/OVERLAY ;
```

```
RUN;

**********************************************************************;
*** USED TO FILTER THE DATA -- CREATE A NEW DATA SET CALLED FILTD ;
**********************************************************************;
DATA FILTD;
RETAIN X16 0 A16 0 X18 0 A18 0 X20 0 A20 0 X21 0 A21 0 X22 0 A22 0;
SET SUB;

PHI161 = -0.233387;
T1620 = 0.613464;
DIFF161=C16 - LAG1(C16);
DIFF1620=LAG20(C16)-LAG21(C16);
DIFF1640=LAG20(LAG20(C16))-LAG20(LAG21(C16));
IF I < 23 THEN X16 = 0;
A16 =DIFF161-DIFF1620-PHI161*(DIFF161-DIFF1620)+T1620*LAG19(X16);
X16 = A16 ;
TERM3 = T1620*LAG19(X16);
TERM1 = DIFF161-DIFF1620;
TERM2 = (-1)*PHI161*(DIFF161-DIFF1620);
TERM0 = LAG19(X16);


*****************************************************************;
*** FILTER OUT THE ... DATA IN FILE FILTD, PUT THE NEW DATA ;
*****************************************************************;
DATA SASDATA.FILTERED;
SET FILTD;
INDEX=_N_;
IF A16= 0 THEN DELETE;
IF INDEX<181 THEN DELETE;
KEEP A16 A18 A20 A21 A22;
RUN;

PROC PRINT DATA=FILTD;
VAR TEST WAFER I IND TERM1 TERM2 TERM3 TERM0 A16 X16;
RUN;
```

Appendix 3.3

SAS Implementation of Hotelling's T^2 Module

```
/**;
;;; SPC module of BCAM ;
;;;
;;; Copyright (c) 1989 Regents of the University of California ;
;;;
;;; Permission to use, copy, modify, and distribute this software and its;
;;; documentation for any purpose and without fee is hereby granted, ;
;;; provided that the above copyright notice appear in all copies and ;
;;; that both that copyright notice and this permission notice appear in ;
;;; supporting documentation, and that the name of the University of ;
;;; California not be used in advertising or publicity pertaining to ;
;;; distribution of the software without specific, written prior ;
;;; permission. The University of California makes no representations ;
;;; about the suitability of this software for any purpose. It is ;
;;; provided "as is" without express or implied warranty. ;
;;;
;;; Author: Hai-Fang Guo ;
;;; $Revision: 1.0 $ ;
;;; $Date: 91/05/17 17:04:01 $ ;
**/;


************************************;
*** USED TO PRODUCE THE COV MATRIX ;
************************************;
 PROC CORR DATA=SASDATA.BASELINE COV OUTP=COVM ;
 VAR A16 A18 A20 A21 A22;
 RUN;

**********************************************************************;
*** USED TO PRODUCE THE MEAN TABLE -- GROUP THE DATA SS(SAMPLE SIZE) ;
*** TOGETHER, TAKE THE MEAN VALUES WHICH WILL BE USED TO CALCULATE;
*** THE T^2 IN PROC MATRIX ;
**********************************************************************;

*** USED TO DO MATRIX OPERATION -- ;
 PROC MATRIX;
 FETCH X DATA=SASDATA.FILTERED /*INCLUDE RAIN5 WHOLE DATA FOR T^2*/;
 FETCH M DATA=SASDATA.COVMAT;
 FETCH Y DATA=SASDATA.BASELINE; /* WINDOW W/O RAIN5 FOR GRANDMEAN*/;

 SS=10 ; /* SAMPLE SIZE N */;
 TR=NROW(X);
 NP=TR#/SS + 1; /* NUMBER OF SUBGROUPS = N OF POINTS IN T^2 */;

 R=NROW(Y);
 C=NCOL(Y);
 COVMATI=INV(M);
 GMEAN=Y(+, )#/R; /* GRAND MEAN OF THE SELECTED WIDNOW */;
```

```
DO I=1 TO NP BY 1; /* CALCULATE GROUP MEAN WITH SIZE=SS */;
IF I=1 THEN DO;
S=1;
E=I#SS-1;
A1=X(S:E, );
MEANMAT=A1(+, )#/E;
END;
IF I^= 1 THEN DO;
START= (I-1)#SS; /* STORED IN MATRIX MEANMAT */;
END=I#SS - 1;
A=X(START:END, );
MEAN=A(+, )#/SS;
MEANMAT=MEANMAT//MEAN;
END;
END;

DO I=1 TO NP BY 1;
DIFF=MEANMAT(I, )-GMEAN;
DI=DIFF';
VAL=DIFF*COVMATI*DI#SS;
T2MAT=T2MAT//VAL;
END;

OUTPUT T2MAT OUT=T2;

PRINT SS NP TR;
RUN;


CMS FILEDEF OUT8 DISK OUT8 TEST;
DATA _NULL_;
SET T2;
FILE OUT8 ;
INDEX=_N_;
PUT INDEX COL1;
RUN;
```

REFERENCES

[1] Costas J. Spanos, "Statistical Process Control in Semiconductor Manufacturing", memo, February 1991.

[2] William E. Barkman, "In-Process Quality Control for Manufacturing", New York and Basel: Marcel Dekker, Inc., 1989.

[3] G.E.P.Box and G.M.Jenkins, "Time Series Analysis: Forecasting and Control", 2nd ed., San Francisco: Holden-Day, 1976.

[4] Richard Harris, "A Primer of Multivariate Statistics", Academic Press, 1975.

[5] Douglas C. Montgomery and David J. Friedman, "Statistical Process Control in a Computer-Integrated Manufacturing Environment", In *Statistical Process Control in Automated Manufacturing*, edited by J. Keats, P67-87. New York and Basel: Marcel Dekker, Inc.,1989.

[6] Gary S. May and Costas J. Spanos, "Automated Malfunction Diagnosis of a Plasma Etcher", *Proceedings of the 1991 International Semiconductor Manufacturing Science Symposium*, May 1991.

[7] Douglas C. Montgomery, "Introduction to Statistical Quality Control", 2nd ed., New York: John Wiley & Sons, 1990.

[8] Walter Andrew Shewhart, "Economic Control of Quality of Manufactured Product", New York: Van Nostrand. (Republished in 1981, with a dedication by W. Edwards Deming, by the American Society for Quality Control, Milwaukee, WI.)

[9] Western Electric Co., "Statistical Quality Control Handbook", 1956.

[10] Allan Pankratz, "Forecasting With Univariate Box-Jenkins Models -- Concepts and Cases", New York: John Wiley & Sons, 1983.

[11] Costas J. Spanos, "On-Line Statistical Process Control of VLSI Manufacturing Equipment", Research proposal to IBM, November 1989.

[1] Bruce Sindahl, "Estimated Time Series and CIM", personal communication, April 1990.

[2] Walter Vandaele, "Applied Time Series and Box-Jenkins Models", New York: Academic Press, 1983.

[3] Christopher Chatfield, "The Analysis of Time Series - An Introduction", 4th ed., New York: Chapman and Hall, 1989.

[4] Hai-Fang Guo, Costas J. Spanos and Alan J. Miller, "Real-Time Statistical Process Control for Plasma Etching", *Proceedings of the 1991 International Semiconductor Manufacturing Science Symposium,* May 1991.

[5] "SAS/ETS User's Guide", Version 5 Ed., SAS Institute Inc.

[6] "RS/1 QCA User's Guide: RS series quality control analysis", BBN Software Products Corporation.

[7] Lam Research, "Rainbow: Excellence in Etch", 1991.

[8] P. Byrne and K. Heiman, "Equipment Analysis with the SECS Protocol and Personal Computers", ISMSS 1989.

[9] Alan J. Miller, "LamStation Performance and Modeling Evaluation", Lam Research, 1990.