CHAPTER 5

Parametric Nonlinear Profiles

James D. Williams

INTRODUCTION

Profiles can take on several different functional forms, depending on the specific application. In some applications, the functional form of the profile cannot be adequately characterized using linear regression methods. Such profiles can be modeled either using parametric nonlinear regression or by nonparametric smoothing methods. Nonparametric regression techniques provide great flexibility in modeling the response. However, one disadvantage of nonparametric smoothing methods is that the subject-specific interpretation of the estimated nonparametric curve may be more difficult, and may not lead the user to discover as easily assignable causes that lead to an out-of-control signal.

Often, however, scientific theory or subject-matter knowledge leads to a natural parametric nonlinear function that describes the profiles well. Hence, one method is to model each profile by a parametric nonlinear regression function. The mathematical formulation for a *nonlinear profile* of an item is given by

$$y_{ij} = f(\mathbf{x}_{ij}, \boldsymbol{\beta}_i) + \epsilon_{ij}, \tag{5.1}$$

where $\mathbf{x}_{ij}$ is a $k \times 1$ vector of regressors for the jth observation of the ith profile, ϵ_{ij} is the random error, $\boldsymbol{\beta}_i$ is a $p \times 1$ vector of parameters for profile i, and f is a nonlinear function in the parameters $\boldsymbol{\beta}_i$. The random errors ϵ_{ij} are assumed to be i.i.d. normal random variates with mean zero and variance σ^2. In many applications, there is only one regressor ($k = 1$), but there are multiple parameters to monitor ($p > 1$). Williams et al. (2007), henceforth denoted WWB, give an example of a the

Statistical Analysis of Profile Monitoring, First Edition. Edited by Rassoul Noorossana, Abbas Saghaei and Amirhossein Amiri.

four-parameter logistic model often used to model dose–response profiles of a drug. The model is given by

$$y_{ij} = A_i + \frac{D_i - A_i}{1 + \left(\frac{x_{ij}}{C_i}\right)^{B_i}} + \epsilon_{ij}, \tag{5.2}$$

where y_{ij} is the measured response of the subject exposed to dose x_{ij} for batch i, $i = 1, \ldots, m$, $j = 1, \ldots, n$. In Equation (5.2), $k = 1$ and $p = 4$, giving four parameters to monitor, each parameter having a specific interpretation: A_i is the upper asymptote parameter, D_i is the lower asymptote parameter, B_i is the rate parameter for the ith batch, and C_i is the ED_{50} parameter. Another example is the "bathtub" function described in WWB where the density of particleboard is measured across the vertical profile. Note that for any given application, the specific form of the nonlinear function, f, in Equation (5.1) must be specified by the user.

The basic idea behind parametric nonlinear profile monitoring is to reduce the complex nonlinear profile into a few parameters through the nonlinear model estimation, then form a control chart scheme on the estimated parameters of each individual profile. The methods are usually multivariate in nature, since in most cases $p > 1$.

In a Phase I analysis, the concern is distinguishing between in-control conditions and the presence of assignable causes so that in-control parameters may be estimated for further product or process monitoring in Phase II analysis. If out-of-control observations are included in the estimation of in-control parameters, then the subsequent monitoring procedure will be less effective. Therefore, it is imperative in Phase I that outliers be identified and excluded from further analysis. Further, step or ramp shifts (if any) must be detected in the mean profile, so that in-control parameters may be estimated to reflect what would be expected from a stable process. In order to accomplish this a historical data set (HDS) consisting of m items sampled over time must be obtained. For each item i a response y_{ij} is observed and a set of predictor variables $\mathbf{x}_{ij}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, resulting in the quality profile for item i, i.e., $(y_{i1}, \mathbf{x}_{i1}), (y_{i2}, \mathbf{x}_{i2}), \ldots, (y_{in}, \mathbf{x}_{in})$.

5.1 NONLINEAR MODEL ESTIMATION

The n observations within each profile of the scalar model in Equation (5.1) are stacked into matrix form as $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in})'$, $\mathbf{f}(\mathbf{X}_i, \boldsymbol{\beta}_i) = (f(\mathbf{x}_{i1}, \boldsymbol{\beta}_i), f(\mathbf{x}_{i2}, \boldsymbol{\beta}_i), \ldots, f(\mathbf{x}_{in}, \boldsymbol{\beta}_i))'$, and $\epsilon_i = (\epsilon_{i1}, \epsilon_{i2}, \ldots, \epsilon_{in})'$. The vector form of the nonlinear model is then given by

$$\mathbf{y}_i = \mathbf{f}(\mathbf{X}_i, \boldsymbol{\beta}_i) + \epsilon_i, \quad i = 1, \ldots, m. \tag{5.3}$$

Estimates of $\boldsymbol{\beta}_i$ must be obtained for each sample of the HDS. This can be accomplished by employing the iterative Gauss–Newton procedure to obtain the maximum likelihood estimates. The Gauss–Newton algorithm requires the $m \times p$

matrix of derivatives of $f(\mathbf{X}_i, \boldsymbol{\beta}_i)$ with respect to $\boldsymbol{\beta}_i$, which is given by

$$\mathrm{D}_i = \frac{\partial \mathbf{f}(\boldsymbol{X}_i, \boldsymbol{\beta}_i)}{\partial \boldsymbol{\beta}_i} = \begin{bmatrix} \frac{\partial f(\mathrm{x}_{i1}, \beta_i)}{\partial \beta_{i1}} & \frac{\partial f(\mathrm{x}_{i1}, \beta_i)}{\partial \beta_{i2}} & \cdots & \frac{\partial f(\mathrm{x}_{i1}, \beta_i)}{\partial \beta_{ip}} \\ \frac{\partial f(\mathrm{x}_{i2}, \beta_i)}{\partial \beta_{i1}} & \frac{\partial f(\mathrm{x}_{i2}, \beta_i)}{\partial \beta_{i2}} & \cdots & \frac{\partial f(\mathrm{x}_{i2}, \beta_i)}{\partial \beta_{ip}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f(\mathrm{x}_{in}, \beta_i)}{\partial \beta_{i1}} & \frac{\partial f(\mathrm{x}_{in}, \beta_i)}{\partial \beta_{i2}} & \cdots & \frac{\partial f(\mathrm{x}_{in}, \beta_i)}{\partial \beta_{ip}} \end{bmatrix}. \tag{5.4}$$

Define $\mathbf{f}\left(\mathbf{X}_i, \hat{\boldsymbol{\beta}}_i^{(h)}\right) = f\left((\mathbf{x}_{i1}, \hat{\boldsymbol{\beta}}_i^{(h)}), f(\mathbf{x}_{i2}, \hat{\boldsymbol{\beta}}_i^{(h)}), \ldots, f(\mathrm{x}_{in}, \hat{\boldsymbol{\beta}}_i^{(h)})\right)'$, where $\hat{\boldsymbol{\beta}}_i^{(h)}$ is the estimate of $\boldsymbol{\beta}$ at iteration h, and define $\hat{\mathbf{D}}_i^{(h)}$ as the matrix of derivatives of f given in Equation (5.4) evaluated at $\hat{\boldsymbol{\beta}}_i^{(h)}$. Then the Gauss–Newton iterative solution for $\hat{\boldsymbol{\beta}}_i$ is given by

$$\hat{\boldsymbol{\beta}}_i^{(h+1)} = \hat{\boldsymbol{\beta}}_i^{(h)} + \left(\hat{\mathbf{D}}_i'^{(h)} \hat{\mathbf{D}}_i^{(h)}\right)^{-1} \hat{\mathbf{D}}_i'^{(h)} \left(\mathbf{y}_i - \mathbf{f}(\mathbf{X}_i, \hat{\boldsymbol{\beta}}_i^{(h)})\right).$$

Upon convergence of the algorithm, the estimated covariance matrix of $\hat{\boldsymbol{\beta}}_i$ is the estimated Fisher information matrix given by

$$\hat{Var}\left(\hat{\boldsymbol{\beta}}_i\right) = \hat{\sigma}_i^2 (\hat{\mathbf{D}}_i' \hat{\mathbf{D}}_i)^{-1},$$

where $\hat{\sigma}_i^2 = \sum_{j=1}^{m} (y_{ij} - f(\mathbf{x}_{ij}, \hat{\boldsymbol{\beta}}_i))^2/(n-p)$ and $\hat{\mathbf{D}}_i$ is the derivative matrix in Equation (5.4) evaluated at the converged parameter vector estimate $\hat{\boldsymbol{\beta}}_i$. Myers (1990, Chapter 9) and Schabenberger and Pierce (2002, Chapter 5) give a concise discussion of nonlinear regression model estimation. A more detailed treatment can be found in Gallant (1987) and Seber and Wild (1989).

Unlike linear regression, the exact distribution of parameter estimators in nonlinear regression is unobtainable, even if the errors ϵ_{ij} are assumed to be i.i.d. normal random variables. Instead, asymptotic results must be applied. Seber and Wild (1989, Chapter 12) give the asymptotic distribution of $\hat{\boldsymbol{\beta}}_i$ and the necessary assumptions and regularity conditions for the asymptotic distribution to be obtained. Given their regularity conditions hold and assuming that $n^{-1}\hat{\mathbf{D}}'_i\hat{\mathbf{D}}_i$ converges to some nonsingular matrix Ω_i as $n \to \infty$, then the asymptotic distribution of $\hat{\boldsymbol{\beta}}_i$ is given by

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_i - \boldsymbol{\beta}_i) \xrightarrow{D} N_p(0, \sigma^2 \Omega_i^{-1}), \tag{5.5}$$

where $\boldsymbol{\beta}_i$ is the asymptotic expected value of $\hat{\boldsymbol{\beta}}_i$ and N_p indicates a p-dimensional multivariate normal distribution. For practical purposes, the distribution given by Equation (5.5) is incalculable since the matrix Ω_i is unknown. Instead the approximate asymptotic distribution of $\hat{\boldsymbol{\beta}}_i$ is commonly used, given by

$$\hat{\boldsymbol{\beta}}_i \dot{\sim} N_p(\boldsymbol{\beta}_i, \sigma_i^2 (\mathbf{D}_i' \mathbf{D}_i)^{-1}). \tag{5.6}$$

This asymptotic distribution result will be exploited to determine approximate control limits in the multivariate T^2 control chart.

The standard estimator of $Var(\hat{\boldsymbol{\beta}}_i)$ is $\hat{\sigma}_i^2 \hat{\mathbf{D}}_i' \hat{\mathbf{D}}_i$. For the most traditional "in-control" case, $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ for all profiles $i = 1, \ldots, m$, where $\boldsymbol{\beta}$ is the in-control parameter vector. Consequently, the $\boldsymbol{\Omega}_i$ and $\mathbf{D}_i$ matrices are the same across all m items since all items have the same underlying model, f, the same x-values are observed, and the same values of $\boldsymbol{\beta}_i$. However, the $\hat{\mathbf{D}}_i$ matrices are not equal since the $\hat{\boldsymbol{\beta}}_i$ values vary from profile to profile.

5.2 PHASE I METHODS

The methods presented in this section are designed to detect out-of-control conditions in the HDS. Out-of-control cases are subsequently removed from the HDS, and the final in-control model parameters are estimated from the in-control HDS. The in-control estimates are then used for ongoing process monitoring in a Phase II analysis.

5.2.1 Multivariate T^2 Control Chart

In this section, the general framework of the multivariate T^2 statistic as well as its usefulness in a Phase I analysis is discussed. From the HDS there are m independent sample observation vectors, denoted, $\mathbf{w}_i$ $(i = 1, \ldots, m)$, each of dimension p. The general form of the T^2 statistic for sample i is

$$T_i^2 = (\mathbf{w}_i - \bar{\mathbf{w}})' \mathbf{S}^{-1} (\mathbf{w}_i - \bar{\mathbf{w}}), \tag{5.7}$$

where $\bar{\mathbf{w}} = \frac{1}{m} \sum_{j=1}^{m} \mathbf{w}_i$ and $\mathbf{S}$ is some estimator of the variance–covariance matrix of $\mathbf{w}_i$ (Mason and Young 2002). The T_i^2 statistics, $i = 1, \ldots, m$, are plotted against i, and out-of-control signals will be given for any T_i^2 value exceeding an upper control limit (UCL). It is usually assumed that each of the $\mathbf{w}_i$ vectors follows a multivariate normal distribution with common mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. This assumption is critical to finding the marginal distribution of T_i^2, as discussed in Section 5.2.1.1.

The vector of parameters $\boldsymbol{\beta}_i$ from the nonlinear regression model given in Equation (5.1) determines the shape of the curve $f(\mathbf{X}_i, \boldsymbol{\beta}_i)$. The multivariate T^2 statistic is employed to assess stability of the p parameters simultaneously. Individual control charts are not used for each of the p nonlinear regression parameters because of the correlation structure of the parameter estimators in nonlinear regression. The correlation structure can be accounted for in the multivariate T^2 statistic.

After calculating $\hat{\boldsymbol{\beta}}_i$ from each sample in the HDS, the average of the $\hat{\boldsymbol{\beta}}_i$ is calculated, which is denoted as $\bar{\hat{\boldsymbol{\beta}}}$. Further, some estimate of the covariance matrix is also calculated. Then $\mathbf{w}_i$ is replaced with $\hat{\boldsymbol{\beta}}_i$ and $\bar{\mathbf{w}}$ is replaced with $\bar{\hat{\boldsymbol{\beta}}}$ in Equation (5.7) to obtain the T^2 statistic

$$T_i^2 = \left(\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}}\right)' \mathbf{S}^{-1} \left(\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}}\right).$$

Large values of T_i^2 indicate abnormal $\hat{\boldsymbol{\beta}}_i$, suggesting that the profile for item i might be out-of-control.

There are several choices for the estimator $\mathbf{S}$. In their paper, WWB discuss three choices, which are given here.

The first choice for $\mathbf{S}$ is the *sample covariance* matrix, which is calculated by

$$\mathbf{S}_C = \frac{1}{m-1}\sum_{i=1}^{m}\left(\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}}\right)\left(\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}}\right)'. \tag{5.8}$$

Subsequently, the T_i^2 statistics are calculated as

$$T_{C,i}^2 = \left(\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}}\right)' \mathbf{S}_C^{-1}\left(\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}}\right). \tag{5.9}$$

Brill (2001) suggested use of the T_C^2 statistics in the context of monitoring nonlinear profiles of a chemical product. The advantage of this statistic is that it is very well understood and widely used. However, Sullivan and Woodall (1996) and Vargas (2003) demonstrated that T_C^2 is ineffective in detecting sustained shifts in the mean vector in a Phase I analysis. In fact, it was shown that as the step shift size increased, the power to detect the shift actually decreased.

An second choice of $\mathbf{S}$ is one based on *successive differences*, proposed originally by Hawkins and Merriam (1974) and again later by Holmes and Mergen (1993). To obtain the estimator, one first calculates the difference vector $\hat{\mathbf{v}}_i = \hat{\boldsymbol{\beta}}_{i+1} - \hat{\boldsymbol{\beta}}_i$ for $i = 1, \ldots, m-1$ and then stacks the transpose of these $m-1$ difference vectors into the matrix $\hat{\mathbf{V}}$ as

$$\hat{\mathbf{V}} = \begin{bmatrix} \hat{\mathbf{v}}_1' \\ \hat{\mathbf{v}}_2' \\ \vdots \\ \hat{\mathbf{v}}_{m-1}' \end{bmatrix}.$$

The estimator of the variance–covariance matrix is then calculated as

$$\mathbf{S}_D = \frac{\hat{\mathbf{V}}'\hat{\mathbf{V}}}{2(m-1)} \tag{5.10}$$

Sullivan and Woodall (1996) showed that $\mathbf{S}_D$ is an unbiased estimator of the true covariance matrix if the process is stable in Phase I. The resulting T_i^2 statistics are given by

$$T_{D,i}^2 = \left(\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}}\right)' \mathbf{S}_D^{-1}\left(\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}}\right). \tag{5.11}$$

Sullivan and Woodall (1996) and Vargas (2003) showed that a T^2 chart based on values of $T_{D,i}^2$ was effective in detecting both a step and ramp shift in the mean

vector during Phase I. Sullivan and Woodall (1996) also showed that the T_D^2 values are invariant to a full-rank linear transformation on the observations.

The third choice for $\mathbf{S}$ is a robust estimator of the variance–covariance matrix known as the minimum volume ellipsoid (MVE) estimator, first proposed by Rousseeuw (1984) and studied in profile monitoring for Phase I analysis by Jensen et al. (2007). To apply the MVE method, one finds outlier-robust estimates for both the in-control parameter vector and the variance–covariance matrix. The MVE estimator is based on finding the ellipsoid with the smallest volume that contains at least half of the $\hat{\boldsymbol{\beta}}_i$ vectors, $i = 1, \ldots, m$. The MVE estimator of $\boldsymbol{\beta}$ is the mean vector of the smallest ellipsoid, and the estimator of the variance–covariance matrix is the sample covariance matrix of the observations within the smallest ellipsoid multiplied by a constant to make the estimator unbiased for multivariate normal data. In a simulation study, Vargas (2003) studied the power properties of several different choices of $\mathbf{S}$ in the context of the T^2 statistic given in Equation (5.7) and found that the T^2 statistic based on the MVE estimators of $\boldsymbol{\beta}$ and the variance–covariance matrix was very powerful in detecting multivariate outliers. The MVE estimators of $\boldsymbol{\beta}$ and the covariance matrix are denoted by $\hat{\boldsymbol{\beta}}_{MVE}$ and $\mathbf{S}_{MVE}$, respectively. Hence, the third choice of T^2 is

$$T_{MVE,i}^2 = \left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\beta}}_{MVE}\right)' \mathbf{S}_{MVE}^{-1} \left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\beta}}_{MVE}\right), \quad i = 1, \ldots, m. \tag{5.12}$$

5.2.1.1 Control Limits

The distribution of the T_i^2 statistics for monitoring nonlinear profiles is more complex than in the case of linear profiles. The asymptotic distribution of $\hat{\boldsymbol{\beta}}_i$, $i = 1, \ldots, m$, must be employed instead. In order to determine the marginal distribution of T_i^2 in this case, it is assumed that the sample size, n, from each item in the baseline data set is of sufficient size such that the distributions of $\hat{\boldsymbol{\beta}}_i$, $i = 1, \ldots, m$, are approximately multivariate normal. The subsequent UCLs for the multivariate T^2 control charts are determined on the basis of this normality assumption.

In order to control the overall probability of a false alarm based on some appropriate UCL, the joint distribution of the T_i^2 values is required. However, these statistics are correlated since $\bar{\hat{\boldsymbol{\beta}}}$ and $\mathbf{S}$ are used in all T_i^2 statistics ($i = 1, \ldots, m$), thus making the joint distribution of the T_i^2 values difficult to obtain. As an alternative, Mahmoud and Woodall (2004) suggested using an approximate joint distribution assuming that the T_i^2 statistics are independent. Let α be the probability of a false alarm for any individual T_i^2 statistic. Then the approximate overall probability of a false alarm for a sample of m items is given by $\alpha_{overall} = 1 - (1 - \alpha)^m$. Thus, for a given overall probability of a false alarm, $\alpha = 1 - (1 - \alpha_{overall})^{1/m}$ is used in the calculation of UCLs. In their simulation study, Mahmoud and Woodall (2004) found that this approximation used to determine the UCLs performed well.

Tracy et al. (1992) noted in their paper that Gnanadesikan and Kettenring (1972) proved that for a stable process the marginal distribution of $T_{C,i}^2$ is proportional to a

β distribution, i.e.,

$$T_{C,i}^2 \frac{m}{(m-1)^2} \sim B\left(\frac{p}{2}, \frac{m-p-1}{2}\right).$$

Chou et al. (1999) give a formal proof. One key assumption is that the distribution of $\hat{\boldsymbol{\beta}}_i$ is approximately normal. Hence, an approximate UCL is

$$UCL_C = \frac{(m-1)^2}{m} B_{1-\alpha, p/2, (m-p-1)/2}, \tag{5.13}$$

where $B_{1-\alpha, p/2, (m-p-1)/2}$ is the $1-\alpha$ quantile of a β distribution with shape parameters $p/2$ and $(m-p-1)/2$.

The marginal distribution of the $T_{D,i}^2$ statistic is unknown. However, Williams et al. (2006) and Williams et al. (2009) gave an approximate distribution based on the chi-squared distribution for large sample sizes and an approximate distribution based on the $\boldsymbol{\beta}$ distribution for small sample sizes. For large sample sizes where $m > p^2 + 3p$, the UCL is

$$UCL_D = \chi^2(1-\alpha, p).$$

For small sample sizes where $m \leq p^2 + 3p$ and $p < 10$, the UCL is a vector given by

$$\mathbf{UCL}_D = (UCL_1, UCL_2, \ldots, UCL_m), \tag{5.14}$$

where

$$UCL_i = MV(m, i)\, BETA_{1-\alpha, \beta(m,p,i), \gamma(m,p,i)}, \quad i = 1, \ldots, m,$$

and $\beta(m, p, i)$ and $\gamma(m, p, i)$ are functions of m, p, and i that define the two shape parameters for the β distribution. The equations for $\beta(m, p, i)$ and $\gamma(m, p, i)$ given by Williams et al. (2006) are included in the Appendix.

The exact marginal distribution of $T_{MVE,i}^2$ is also unknown and intractable. Hence, in order to find the UCL for $T_{MVE,i}^2$ simulation must be employed.

5.2.2 Lack-of-Fit Control Chart

Another type of out-of-control profile is one where the functional form is different from the hypothesized model. Hence, in addition to monitoring the mean profile for abnormalities, it also may be of interest to check for changes in the profile function over time. This can be done by employing the lack-of-fit (LOF) control chart proposed by Williams et al. (2007), henceforth denoted WBWF. In order to compute the LOF control chart, the observations taken at each point must be replicated. The

replicated observations are denoted as $\mathbf{x}_{ijk}$, where $k = 1, \ldots, r$ and $r > 1$. Replication is common for dose–response monitoring schemes in the agricultural products and pharmaceutical industries, among others.

The LOF statistic proposed in WBWF compares the error sum of squares of a saturated model, sometimes called the means or full model, to the error sum of squares of the specified nonlinear regression model. The sum of squared errors for the saturated model is

$$SSE_i^{full} = \sum_{j=1}^{n} \sum_{k=1}^{r} (y_{ijk} - \bar{y}_{ij.})^2,$$

where $\bar{y}_{ij.} = r^{-1} \sum_{k=1}^{r} y_{ijk}$ is the mean of the r replications at the point j in profile i. Further, let $f(\mathbf{x}_{ijk}, \hat{\boldsymbol{\beta}}_i)$ represent the estimated profile replication k of profile i, then the sum of square errors for the nonlinear regression model is given by

$$SSE_i^{reg} = \sum_{j=1}^{n} \sum_{k=1}^{r} (y_{ijk} - f(\mathbf{x}_{ijk}, \hat{\boldsymbol{\beta}}_i))^2.$$

The LOF statistic compares the nonlinear model fit to the saturated model. The LOF statistic for profile i is

$$LOF_i = \frac{(SSE_i^{reg} - SSE_i^{full})/df^{LOF}}{SSE_i^{full}/df^{full}}, \tag{5.15}$$

where $df^{LOF} = n - p$ is the degrees of freedom for the LOF and $df^{full} = n(r - 1)$ is the degrees of freedom of the full model. Neill (1988) demonstrated that the LOF_i statistic for the nonlinear regression approximately follows an $\mathcal{F}$ distribution with df^{LOF} and df^{full} numerator and denominator degrees of freedom, respectively.

To assess LOF for profile i, one plots all LOF_i statistics versus i in an LOF chart. The UCL for the LOF chart is based on the $\mathcal{F}$ distribution, and is given by

$$UCL_{LOF} = \mathcal{F}(1 - \alpha, df^{LOF}, df^{full}), \tag{5.16}$$

where $\mathcal{F}(q, df_1, df_2)$ is the qth quantile from an $\mathcal{F}$ distribution with df_1 and df_2 numerator and denominator degrees of freedom, respectively. In those cases where LOF_i exceeds UCL_{LOF}, the proposed parametric nonlinear model is determined to be inadequate for profile i. Those profiles that exhibit LOF are subject to being removed from the HDS.

5.2.3 Monitoring the Variance σ^2

In addition to checking abnormal mean profiles in the HDS, it is also important to check for abnormal variance about each profile. This is analogous to monitoring

the process variance in the standard univariate case. In nonlinear profile monitoring, it is of interest to monitor the within-profile variability. The measure of within-profile variability is the mean square error (MSE) defined as $MSE_i = \sum_{j=1}^{n}(y_{ij} - \hat{y}_{ij})^2/(n-p)$, where $\hat{y}_{ij}$ is the predicted value of y_{ij} based on the nonlinear regression model in Equation (5.1). Wludyka and Nelson (1997) recommended a method to monitor variances based on an analysis-of-means-type test utilizing $S_i^2 = MSE_i$. In their paper, S_i^2 is plotted against i with associated lower- and upper-control limits equal to $(L_{\alpha,m,n-p})m\overline{S}^2$ and $(U_{\alpha,m,n-p})m\overline{S}^2$, respectively, where L and U are critical values given in their paper and $\overline{S^2}$ is the average of the S_i^2 values, $i = 1, \ldots, m$. For large n, their approximate upper and lower control limits are $\overline{S}^2 \pm h_{\alpha,m,\infty}\hat{\sigma}$, where h is a critical value given in Nelson (1983) and $\hat{\sigma} = \overline{S}^2\sqrt{2(m-1)/m(n-p)}$. The S_i^2 statistics are plotted on a separate control chart to monitor the variance of the error terms and LOF simultaneously with a T^2 control chart for the nonlinear regression parameters. However, use of this method is recommended only when within-profile error terms are independent.

5.2.4 In-Control Profile Estimate

The main purpose of a Phase I analysis is to estimate in-control parameters, including parameters for the mean profile and parameters for the variance. Then in ongoing process monitoring, new profiles are then monitored in a Phase II analysis against the baseline profile determined in Phase I. In Phase I, out-of-control profiles are removed from the HDS and in-control parameters are estimated from the remaining in-control profiles in the HDS. Let m^* be the number of in-control profiles remaining in the HDS. Then the in-control mean profile is estimated as $\mathbf{f}(\mathbf{X}, \boldsymbol{\beta})$ from Equation (5.3), where

$$\boldsymbol{\beta} = \frac{1}{m^*}\sum_{i=1}^{m^*}\hat{\boldsymbol{\beta}}_i. \tag{5.17}$$

The final estimate of the variance–covariance matrix of $\hat{\boldsymbol{\beta}}$, denoted by $\boldsymbol{\Sigma}_\beta$, is given by

$$\boldsymbol{\Sigma}_\beta = \frac{1}{m^*-1}\sum_{i=1}^{m^*}\left(\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}}\right)\left(\hat{\boldsymbol{\beta}}_i - \bar{\hat{\boldsymbol{\beta}}}\right)'. \tag{5.18}$$

Finally, the in-control estimate of σ^2 is given by

$$\sigma^2 = \frac{1}{m^*}\sum_{i=1}^{m^*}\sum_{j=1}^{n}\frac{(y_{ij} - f(\mathbf{x}_{ij}, \hat{\boldsymbol{\beta}}_i))^2}{n-p}. \tag{5.19}$$

One basic assumption is that the estimated parameters from Equations (5.17)–(5.19) are known and exact. This assumption affects the distributional

properties of the control chart statistics in Phase II. One should also check that $\mathbf{\Sigma}_\beta$ is positive definite, since this is a requirement in further stages of the analysis. One can check this assumption by calculating the eigenvalues of the $\mathbf{\Sigma}_\beta$ matrix and checking for any non-positive eigenvalues.

5.2.5 VDP Data Example

WWB gave the following example using the vertical density profile (VDP) data from Walker and Wright (2002) to illustrate the nonlinear profile monitoring methods of this section. The original data is available at the Web site http://filebox.vt.edu/users/bwoodall/VDP%20nonlinear%20profile%20data.txt. In the manufacture of particleboard, the quality characteristic monitored over time is the density properties of the finished boards. In manufactured particleboard, the density (in pounds per ft^3) near the core or center of a particleboard is much less than the density at the top and bottom faces of a board (see Young et al. 1999). The industry standard sampling procedure calls for a laser-aided density measuring device that scans fixed vertical depths of a board and records the density at each depth. Since the scan depths are fixed for each sample, the depth x_{ij} is denoted by simply x_j. Density measurements for this HDS were taken at depths of $x_j = (0.002)j$ inches, $j = 0, 1, 2, \ldots, 313$, giving a sequence of ordered pairs, (x_j, y_{ij}), $j = 1, \ldots, n$. The set of order pairs forms a *VDP* of the board. The HDS contains 24 particleboards measured in this way, and the 24 profiles are illustrated in Figure 5.1.

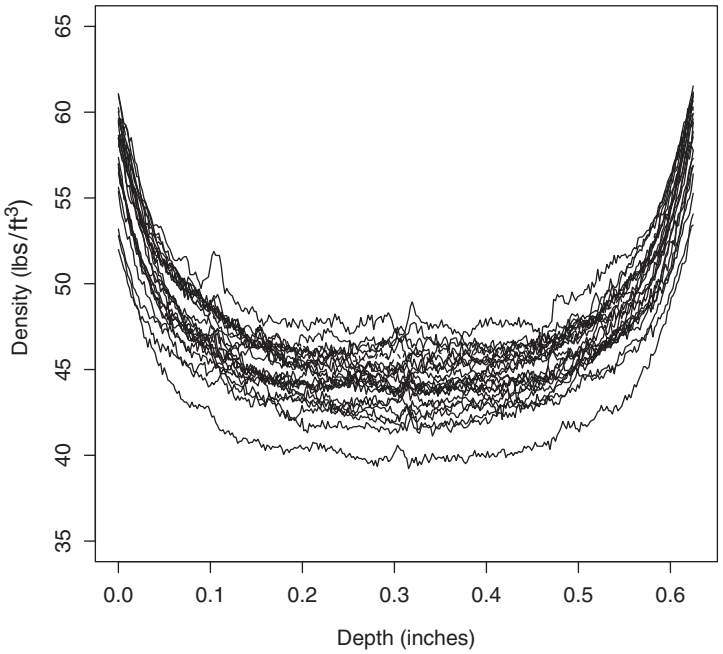

Figure 5.1 Vertical density profile (VDP) of 24 particleboards.

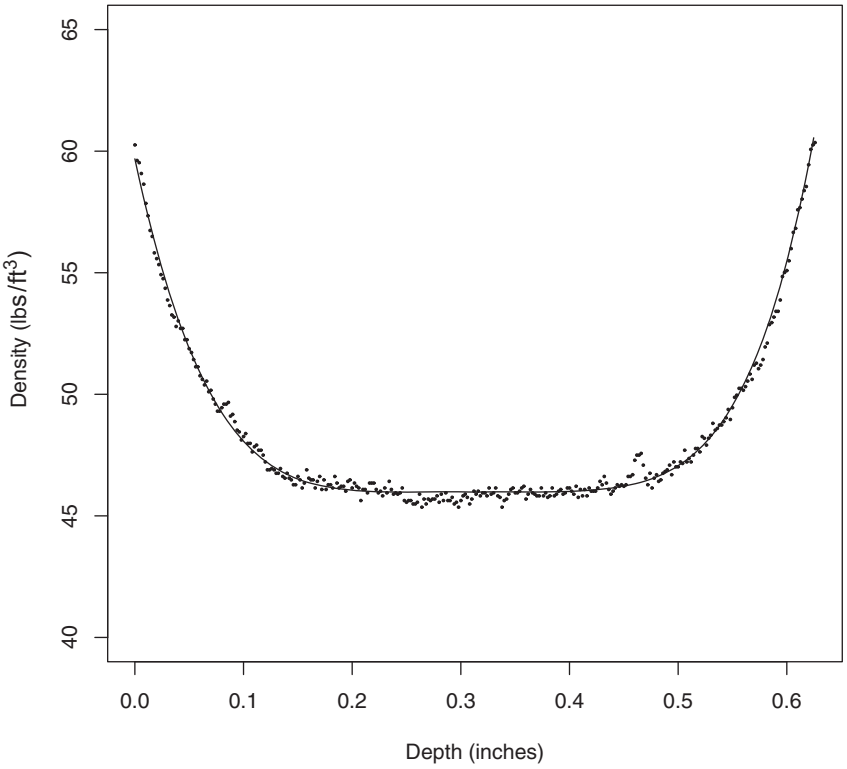

Figure 5.2 "Bathtub" function fit to board 1.

The nonlinear function used to model particleboard profile i is a "bathtub" function given by

$$f(x_{ij}, \boldsymbol{\beta}) = \begin{cases} a_1(x_{ij} - c)^{b_1} + d & x_j > c \\ a_2(-x_{ij} + c)^{b_2} + d & x_j \leq c \end{cases}, \quad i = 1, \ldots, m; \quad j = 1, \ldots, n, \tag{5.20}$$

where $\boldsymbol{\beta} = (a_1, a_2, b_1, b_2, c, d)$. One advantage of this nonlinear model is the interpretability of the model parameters. For example, a_1, a_2, b_1, and b_2 determine the "flatness," c is the center, and d is the bottom, or the "level" of the curve. Differing values of a_1 and a_2 or different values of b_1 and b_2 allow for an asymmetric curve about the center c. Figure 5.2 contains the "bathtub" function fit to board 1 from the VDP data.

Although this parametric nonlinear regression model does not explain all of the variability in the VDP of board 1, the bathtub function model has $R^2 > 0.9999$, and is acceptable for the purposes of this analysis. The bathtub function of Equation (5.20) was fit to each of the 24 boards in the HDS, and the T_C^2, T_D^2, and T_{MVE}^2 statistics of Equations (5.9), (5.11), and (5.12), respectively, were calculated based on the $\hat{\boldsymbol{\beta}}_i$

Table 5.1 Estimated Parameter Values and T^2 Statistics for the VDP Data

Board	$\hat{a}_1$	$\hat{a}_2$	$\hat{b}_1$	$\hat{b}_2$	$\hat{c}$	$\hat{d}$	T_C^2	T_D^2	T_{MVE}^2	UCL_D
1	6560	3259	5.63	4.40	45.98	0.29	2.65	1.91	6.00	27.88
2	470	291	3.01	2.74	42.08	0.32	7.56	5.27	6.97	22.29
3	1812	2871	3.99	5.02	47.66	0.34	5.83	7.17	8.64	22.27
4	6171	15,009	4.25	7.39	46.63	0.39	12.21	17.28	1131.81	22.24
5	4963	2251	5.14	4.20	43.43	0.30	1.65	2.27	2.88	22.21
6	4556	3758	5.28	4.72	40.13	0.30	8.49	13.03	9.83	22.17
7	5542	3815	5.25	5.00	44.15	0.31	2.15	3.49	3.58	22.12
8	3664	2979	4.89	4.41	44.06	0.30	0.79	0.97	2.69	22.07
9	28,041	8872	7.58	4.95	43.22	0.26	4.62	7.10	385.03	22.01
10	1640	1207	4.17	3.39	41.84	0.28	4.30	5.05	4.61	21.95
11	3492	1031	5.82	3.17	46.06	0.25	8.66	8.95	10.00	21.91
12	915	750	3.45	3.52	44.37	0.32	1.80	1.99	2.22	21.88
13	989	1392	3.58	4.05	45.47	0.32	3.42	4.42	5.18	21.88
14	1474	620	4.82	3.29	42.52	0.27	3.28	4.50	7.04	21.91
15	129,068	5420	12.40	3.33	45.90	0.15	21.45	22.18	17018.91	21.95
16	10,166	3822	5.83	4.86	44.19	0.30	3.83	5.60	12.93	22.01
17	1483	603	4.07	3.26	44.83	0.30	2.30	2.53	2.36	22.07
18	31,156	31,069	7.70	5.94	46.46	0.27	14.55	19.75	8221.00	22.12
19	418	198	3.22	2.67	42.84	0.30	4.58	3.90	5.16	22.17
20	3207	4741	4.88	5.02	44.45	0.30	5.34	5.59	34.00	22.21
21	672	773	3.37	3.37	44.46	0.31	2.64	3.42	2.79	22.24
22	3520	1807	5.10	4.01	45.52	0.29	1.71	1.37	1.73	22.27
23	1979	845	4.24	3.66	45.53	0.32	4.45	4.85	7.38	22.29
24	6095	26,778	5.41	6.67	44.46	0.31	9.75	10.55	6676.21	27.88

values. Parameter estimates for each of the twenty-four boards and the corresponding T^2 statistics are given in Table 5.1.

The control limits for the T_C^2 and T_D^2 statistics were calculated from Equations (5.13) and (5.14), respectively. The UCL for the T_{MVE}^2 statistic was simulated to achieve an overall probability of a signal equal to 0.05 for $m = 24$ boards. In this simulation, samples were drawn from a multivariate normal distribution of dimension six, mean vector zero, and variance–covariance matrix I. The identity covariance matrix was used since the in-control performance of the methods does not depend on the assumed in-control parameter vector or the variance–covariance matrix. The simulation was repeated 200,000 times, which gives a standard error for the estimated control limits of less than 0.0005. The UCL values are 14.72 and 65.37, for the T_C^2 and T_{MVE}^2 control charts, respectively. The control limit vector for the T_D^2 control chart is given in Table 5.1.

The purpose of a Phase I analysis is to identify "outlying" or out-of-control boards or a shift in the process that might affect the estimation of in-control parameters. Subsequently, the three T^2 control charts were compared for assessing process stability and identifying outlying profiles. Figure 5.3 below is a plot of all three T^2 control charts for the VDP data.

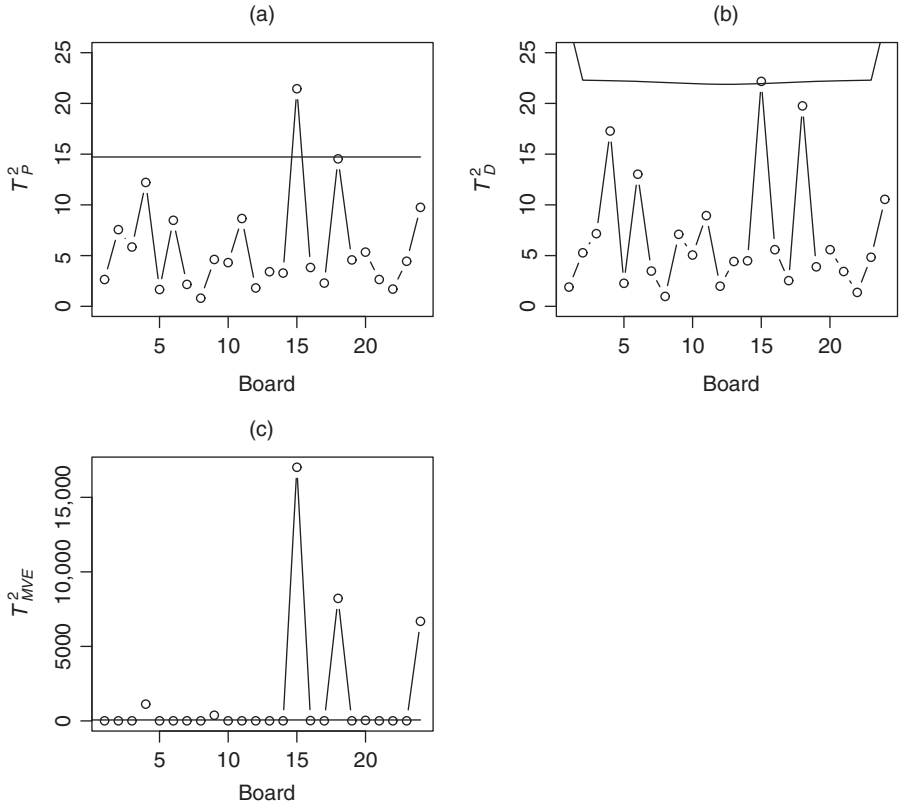

Figure 5.3 The T^2 control charts for the VDP data. (a) The T_C^2 control chart based on the sample covariance matrix, (b) T_D^2 control chart based on the successive differences estimator, and (c) T_{MVE}^2 control chart based on the minimum volume ellipsoid.

Both the T_C^2 and the T_D^2 control charts signal that board 15 is a potential out-of-control profile, and the profile for board 18 produces a signal in the T_C^2 control chart. Note that the T_D^2 statistic accentuates the same outlying observations of the T_C^2 chart, but has a larger UCL. Sullivan and Woodall (1996) found that the T_C^2 control chart has greater power to detect isolated outlying observations than the T_D^2 control chart based on the successive differences variance–covariance matrix estimator; however, the T_D^2 chart is better for detecting a sustained shift in the mean vector. For this HDS, there is no apparent sustained shift in the regression parameter vector.

The T_{MVE}^2 control chart based on the minimum variance ellipsoid estimator indicates that boards 4, 9, 15, 18, and 24 have outlying profiles. The most pronounced outlier is board 15, which both the T_C^2 and the T_D^2 charts also indicated as the most severe outlier. As shown in Vargas (2003), the T_{MVE}^2 control chart is very powerful in detecting multivariate outliers. Investigating the table of parameter estimates for these boards, given in Table 5.1, it seems reasonable that the boards 15 and 18 are outliers, with boards 4, 9, and 24 worthy of further investigation.

In addition to monitoring the regression parameter vectors of the profiles in a Phase I analysis, the variation about the profiles should be monitored to check for stability. In their paper, WWB recommended using the methods of Wludyka and Nelson (1997) to monitor the variance σ^2. Use of their method is appropriate when the error terms within a profile are independent. In this VDP example, however, the within-profile density measurements are spatially correlated so the variance chart is not used.

5.3 PHASE II METHODS

In-control model parameters are estimated in a Phase I analysis. A Phase II analysis consists of continuous monitoring of ongoing profiles from the product or process, denoted by $\mathbf{y}_i, i = 1, \ldots$. The methods presented in this section are designed to detect out-of-control conditions of new profiles compared to the baseline profile estimated from the Phase I analysis. For each new profile, $\mathbf{y}_i$, the nonlinear model estimation methods of Section 5.1 are applied to obtain the estimated parameter vector $\hat{\boldsymbol{\beta}}_i$, the fitted sample profile $\mathrm{f}(\mathbf{X}_i, \hat{\boldsymbol{\beta}}_i)$, and the estimate of the profile variance S_i^2. Phase II control charts are then applied on $\hat{\boldsymbol{\beta}}_i$ and S_i^2 to detect changes in the mean profile, changes in the variance of the profile, or both.

5.3.1 Multivariate T^2 Control Chart

Analogous to the multivariate T^2 control chart given in Equation (5.7) for a Phase I application, the multivariate T^2 control chart to monitor $\hat{\boldsymbol{\beta}}_i$ for Phase II applications is given by

$$T_i^2 = (\hat{\boldsymbol{\beta}}_i - \boldsymbol{\beta})' \boldsymbol{\Sigma}_\beta^{-1} (\hat{\boldsymbol{\beta}}_i - \boldsymbol{\beta}), \tag{5.21}$$

where $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}_\beta$ are given in Equations (5.17) and (5.18), respectively. Since $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}_\beta$ are assumed to be known, then T_i^2 of Equation (5.21) has a chi-square distribution with p degrees of freedom. Hence, a UCL for this control chart is given by $UCL_{T^2} = \chi^2(1 - \alpha, p)$.

In their paper, WBWF give a more conservative UCL for T_i^2 by relaxing the assumption that the parameters in Equations (5.17) and (5.18) are known and exact. This UCL is based on the fact that these parameters are estimated rather than known. This UCL is based on a results in Mason et al. (2001), and is given by

$$UCL_{T^2} = \frac{p(m^* + 1)(m^* - 1)}{m^*(m^* - p)} \mathcal{F}(1 - \alpha, p, m^* - p),$$

where $\mathcal{F}(\alpha, df_1, df_2)$ is the α quantile of an F distribution with df_1 numerator and df_2 denominator degrees of freedom. Whenever T_i^2 exceeds UCL_{T^2} then profile i is declared to be out of control, and is subject to further investigation.

5.3.2 EWMA Control Chart

Vaghefi et al. (2009) extended the EWMA control chart methods of the linear profile given in Kang and Albin (2000) to the nonlinear profile case. This method is based on monitoring the residuals between the reference and sample profiles, and can be applied to either parametric nonlinear profiles $\mathbf{f}(\mathbf{X}_{ij}, \boldsymbol{\beta}_i)$ or nonparametric nonlinear profiles, that is, mean profile curves estimated using nonparametric regression techniques. The EWMA chart is based on monitoring the expected value of

$$e_{ij} = y_{ij} - f(\mathbf{x}_{ij}, \boldsymbol{\beta}), \tag{5.22}$$

which is the error between the reference profile and the sample profile. Recall that $\boldsymbol{\beta}$ was calculated from the in-control profiles of the HDS in a Phase I analysis, as described in Section 5.2.4. The EWMA statistics are calculated as $z_i, i = 1, 2, 3, \ldots,$ where $z_i = \lambda \bar{e}_i + (1 - \lambda) z_{i-1}$, $\bar{e}_i$ is the sample mean of the residuals from profile i, $0 < \lambda \leq 1$ is the smoothing parameter, and z_0 is assumed to equal zero. The upper and lower control limits of the EWMA control chart are symmetric around zero, and are given by

$$\pm c_1 \sigma \sqrt{\frac{\lambda}{(2 - \lambda)n}},$$

where c_1 is a constant chosen to produce the desired average run length (ARL) and $\sigma = \sqrt{\sigma^2}$ from Equation (5.19).

5.3.3 MCUSUM Control Chart

Vaghefi et al. (2009) proposed a multivariate cumulative sum (MCUSUM) control chart to monitor the parameter vector $\hat{\boldsymbol{\beta}}_i$. Their method assumes that $\hat{\boldsymbol{\beta}}_i$ comes from a multivariate normal distribution with either an in-control mean denoted by $\boldsymbol{\beta}_0$ or an out-of-control mean denoted by $\boldsymbol{\beta}_\delta$, where $\boldsymbol{\beta}_\delta = \boldsymbol{\beta} + \boldsymbol{\delta}$ and $\boldsymbol{\delta}$ is the magnitude of the shift one is interested in detecting. The MCUSUM statistic is given by

$$S_i = max(S_{i-1} + \mathbf{a}'(\hat{\boldsymbol{\beta}}_i - \boldsymbol{\beta}) - 0.5D, 0), \tag{5.23}$$

where $S_0 = 0$,

$$D = \sqrt{\boldsymbol{\delta}' \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\delta}}$$

and $\mathbf{a}$ is a $p \times 1$ vector given by

$$\mathbf{a}' = \frac{\boldsymbol{\delta}' \boldsymbol{\Sigma}_\beta^{-1}}{D}.$$

A signal is given whenever S_i exceeds the UCL, which can set to obtain an overall ARL.

5.3.4 Control Charts Based on Metrics

Other Phase II control charts that can be applied to either parametric or nonparametric linear profiles are those given in Vaghefi et al. (2009) based on metrics. These charts, like the EWMA chart in Section 5.3.2, are based on the residual error e_{ij} of Equation (5.22). Each metric is designed to measure a different aspect of the "distance" between the reference profile and the sample profile. The metrics are defined as

$$
\begin{aligned}
M_{i1} &= \max_{1\le j\le n} |e_{ij}| \\
M_{i2} &= \sum_{j=1}^{n} |e_{ij}| \\
M_{i3} &= \sum_{j=1}^{n} n^{-1}|e_{ij}| \\
M_{i4} &= \sum_{j=1}^{n} e_{ij}^2 \\
M_{i5} &= \left| S(\mathbf{y}_i) - \int_{x_1}^{x_n} f(\mathbf{x}_i, \boldsymbol{\beta}) dx \right|,
\end{aligned}
$$

where $S(\mathbf{y}_i)$ approximates the integral of y between the range of the x_{ij} by the trapezoidal method. The four metrics M_{i1} through M_{i4} represent the maximum deviation, sum of absolute deviations, the mean absolute deviation, and the sum of square deviations between the reference profile and the sample profile, respectively. The metric M_{i5} represents the absolute difference between the area under the reference profile and the area under the sample profile.

The UCL for the control charts based on metrics $M_{i1} - M_{i5}$ must be obtained via simulation in order to obtain the desired ARL.

5.3.5 Lack-of-Fit Control Chart

If there exists replication in the observed profile at each point $\mathbf{x}_{ij}$, then LOF can be monitored at time i using methods analogous to those in Section 5.2.2. From the sample profile obtained at time i, the LOF statistic is calculated as

$$
LOF_i = \frac{(SSE_i^{reg} - SSE_i^{full})/df_i^{LOF}}{SSE_i^{full}/df^{full}}.
$$

As mentioned in Section 5.2.2, the LOF_i statistic follows an approximate $\mathcal{F}$ distribution with df^{LOF} and df^{full} numerator and denominator degrees of freedom, respectively. The LOF_i statistics are plotted by i in a LOF chart. The associated UCL is given by

$$
UCL_{LOF} = \mathcal{F}(1-\alpha, df^{LOF}, df^{full}).
$$

Whenever LOF_i exceeds UCL_{LOF}, then it is determined that the proposed dose–response model has changed for profile i, and is subject to further investigation.

5.3.6 Monitoring the Variance

In addition to monitoring the mean of the profiles, it is also important to check for changes in the residual variance of the profiles. From a Phase I analysis, the in-control residual variance is estimated to be σ^2 of Equation (5.19). Similar to the Phase I control chart for the variance given in Section 5.2.3, the Phase II control statistic to monitor the variance is based on the estimated within-profile error variance, estimated by the MSE. The chart statistic is calculated as

$$\chi^2 = \frac{(n-p)S_i^2}{\sigma^2},$$

which has an asymptotic χ^2 distribution with $n - p$ degrees of freedom. Hence, an approximate UCL for the chart is given by $UCL = \chi^2(1-\alpha, n-p)$.

5.4 VARIANCE PROFILES

In some applications where $\mathbf{x}_{ij}$ is replicated r times, such as is common in dose–response profiles of a drug or chemical, one may find that the variance is not homoscedastic throughout the range of $\mathbf{x}_{ij}$. One would expect that the variability in the response of an organism to large doses of a chemical would be much smaller than the variability in the response to small doses. In this case, the variance can be described as some function of dose. In their paper, WBWF give an example of a heteroscedastic dose–response relationship of a chemical in the agricultural crop protection industry.

The body of literature on variance function modeling is extensive. A thorough treatment of variance function estimation is given in Davidian and Carroll (1987), Carroll and Ruppert (1988), and Arbogast and Bedrick (2004). Davidian and Carroll (1987) characterized the general variance function model as

$$\text{Var}(y_{ijk}) = \sigma_{ij}^2 = \sigma_i^2 g(\mathbf{z}_{ij}, \boldsymbol{\beta}_i, \boldsymbol{\theta}_i),$$

where in this case σ_i^2 is a scale parameter for profile i and $g(\cdot)$ is some function of regressor variables $\mathbf{z}_{ij}$, the parameter vector $\boldsymbol{\beta}_i$, and other parameters $\boldsymbol{\theta}_i$. The variance predictor variables $\mathbf{z}_{ij}$ can be $\mathbf{x}_{ij}$, but not necessarily so. The form of $g(\cdot)$ will depend on the specific application.

When there are replications at every $\mathbf{x}_{ij}$, then an unbiased estimator of σ_{ij}^2 can be obtained. The estimator is

$$\hat{\sigma}_{ij}^2 = S_{ij}^2 = \frac{1}{r-1}\sum_{k=1}^{r}(y_{ijk} - \bar{y}_{ij.})^2. \tag{5.24}$$

If it is assumed that y_{ijk}, $k = 1, \ldots, r$, are i.i.d. normal random variables (i.e., the r replications within a given dose are independent), then

$$\frac{(r-1)S_{ij}^2}{\sigma_{ij}^2} \sim \chi^2(r-1). \tag{5.25}$$

A useful variance model proposed by Bellio et al. (2000) is the so-called power of x (POX) model, given by

$$\sigma_{ij}^2 = \sigma_i^2 g(\mathbf{z}_{ij}, \boldsymbol{\beta}_i, \boldsymbol{\theta}_i) = a_{0,i} x_{ij}^{\theta_{1,i}}. \tag{5.26}$$

Then, as suggested by Aitkin (1987), using the distributional result of Equation (5.25) the S_{ij}^2 may be modeled using a generalized linear model (GLIM) framework with the natural logarithm link function and gamma errors. Specifically, the GLIM model used is S_{ij}^2 following a gamma distribution with scale parameter equal to $\frac{r-1}{2}$ and mean function equal to

$$\sigma_{ij}^2 = \exp\{\theta_{0,i} + \theta_{1,i} \log(x_{ij})\}. \tag{5.27}$$

Note that $\theta_{0,i} = \log(a_{0,i})$. This model is based on the assumption that $\log(S_{ij}^2)$ has a simple linear relationship with $\log(x_{ij})$. Subject-matter theory and experience often give rise to an appropriate variance model.

Once the model parameters in Equation (5.27) are estimated using GLIM techniques, then estimates of σ_{ij}^2 may be obtained. Let $\hat{\theta}_{0,i}$ and $\hat{\theta}_{1,i}$ be the GLIM estimators of $\theta_{0,i}$ and $\theta_{1,i}$, respectively. These estimators have an asymptotic normal distribution, and the estimator of σ_{ij}^2 is

$$(\hat{\sigma}_{ij}^2)^{GLIM} = \exp\{\hat{\theta}_{0,i} + \hat{\theta}_{1,i} \log(x_{ij})\}. \tag{5.28}$$

The expression in Equation (5.28) is referred to as the estimated *variance profile* for sampling period i.

In order to check that the estimated variance profiles are in-control, one looks for unusual values of

$$\hat{\boldsymbol{\theta}}_i = \begin{bmatrix} \hat{\theta}_{0,i} \\ \hat{\theta}_{1,i} \end{bmatrix}.$$

For example, an extremely large value of $\theta_{1,i}$ could indicate that the overall variability for variance profile i is too large. On the other hand, an out-of-control value of $\theta_{1,i}$ indicates that the rate of increase or decrease of heterogeneity of variance is different for variance profile i. For positive values of $\theta_{1,i}$, the variability increases with x_{ij}, whereas for negative values the variability decreases with x_{ij}.

Monitoring the variance profile in a Phase I application can be done using the multivariate T^2 control chart methods described in Section 5.2.1 applied

to $\hat{\boldsymbol{\theta}}_i, i = 1, \ldots, m$. The associated UCL values for the T^2 control charts are based on the assumption that the $\hat{\boldsymbol{\theta}}_i$ have an asymptotic multivariate normal distribution with mean vector $\boldsymbol{\theta}$ and covariance matrix $\boldsymbol{\Sigma}_\theta$. A common estimator of the mean is $\bar{\hat{\boldsymbol{\theta}}} = m^{-1} \sum_{i=1}^{m} \hat{\boldsymbol{\theta}}_i$. The estimator of the covariance matrix can be based on the sample covariance matrix estimator of Equation (5.8), the successive differences estimator of Equation (5.10), or the MVE estimator of Section 5.2.1. The associated UCL values of the T^2 charts are calculated in the same manner as described in Section 5.2.1.1.

It is important to note that in the case of heteroscedasticity, a more appropriate mean profile estimation method is to use a weighted Gauss–Newton estimation method, where the weights are given by the $n \times n$ diagonal weight matrix $\mathbf{W}_i = \text{diag}\{(\hat{\sigma}_{ij}^2)^{GLIM}\}$. This will affect both the estimate of $\boldsymbol{\beta}$ and the associated estimate of the variance–covariance matrix $\boldsymbol{\Sigma}_\beta$.

With the aid of the T^2 charts and the associated UCLs, the out-of-control variance profiles are investigated an possibly removed from the HDS. With the remaining m^* profiles, the in-control mean vector $\boldsymbol{\mu}_\theta$ and covariance matrix $\boldsymbol{\Sigma}_\theta$ for use in continuous process monitoring are calculated as

$$\boldsymbol{\mu}_\theta = \frac{1}{m^*} \sum_{i=1}^{m^*} \hat{\boldsymbol{\theta}}_i \tag{5.29}$$

and

$$\boldsymbol{\Sigma}_\theta = \frac{1}{m^* - 1} \sum_{i=1}^{m^*} \left(\hat{\boldsymbol{\theta}}_i - \boldsymbol{\mu}_\theta\right) \left(\hat{\boldsymbol{\theta}}_i - \boldsymbol{\mu}_\theta\right)'. \tag{5.30}$$

5.4.1 Phase II Variance Profile Monitoring

To monitor the variance profile collected at time t, the chosen variance profile model from the Phase I analysis is employed to estimate the value of $\boldsymbol{\theta}$ from the profile generated at time t. The estimated mean vector and variance–covariance matrix of $\hat{\boldsymbol{\theta}}$, given in Equations (5.29) and (5.30), are then used to construct a T^2 control chart. The $T_t^2,\ t = 1, 2, 3, \ldots,$ statistics are calculated as

$$T_t^2 = \left(\hat{\boldsymbol{\theta}}_t - \boldsymbol{\mu}_\theta\right)' \boldsymbol{\Sigma}_\theta^{-1} \left(\hat{\boldsymbol{\theta}}_t - \boldsymbol{\mu}_\theta\right).$$

The UCL for this chart is given by

$$UCL_{T_{\hat{\theta}}^2} = \frac{2(m^* + 1)(m^* - 1)}{m^*(m^* - 2)} \mathcal{F}(1 - \alpha_{II}, p, m^* - 2)$$

(Mason et al. 2001). Whenever $T_t^2 > UCL_{T_{\hat{\theta}}^2}$ then the chart signals, and hence profile t is subject to further investigation.

5.4.2 Dose–Response Data Example

To illustrate the profile monitoring methods for heteroscedastic data, the HDS from DuPont Crop Protection given in WBWF is analyzed. The data consist of 44 weeks ($m = 44$) of in vivo bioassay results for standards (commercial crop protection products) run alongside experimental compounds over a 1-year time period. The bioassay procedure was employed using a commercial crop protection product and a test organism. Because of the proprietary nature of the bioassay, the commercial compound and test organism used are undisclosed.

The commercial compound was diluted to eight doses (0.003, 0.009, 0.028, 0.084, 0.25, 0.76, 2.27, and 6.8) and replicated four times at each dose ($r = 4$) in 96-well microtiter plates for each sampling period i. A spectrophotometer measured the optical density (OD) of the plant organism after the inoculation period. Both treated and untreated wells were measured for growth inhibition. The PC_{ijk} values were calculated using the median OD (M_i) from 96 replications of untreated wells.

Based on subject-matter theory of the biological process, the proposed variance model is the POX model given in Equation (5.26). Equation (5.24) was employed to calculate S_{ij}^2 for $i = 1, \ldots, 44$ and $j = 1, \ldots, 4$, used the distributional properties of the S_{ij}^2 statistics given in Equation (5.25), and employed a GLIM model from Equation (5.27). Using this model values of the estimators, $\hat{\theta}_{0,i}$ and $\hat{\theta}_{1,i}$ were obtained for all 44 weeks of the DuPont data.

To graphically visualize the goodness-of-fit of this model, the $\log(S_{ij}^2)$ versus the logarithm of dose is plotted, superimpose with the logarithm of the predicted values of the variance from Equation (5.28), given by $\log((\hat{\sigma}_{ij}^2)^{GLIM})$. The plot is given in Figure 5.4.

To check for unusual variance profiles, $T_{MVE,i}^2$ and $T_{D,i}^2$ statistics were calculated based on the values of $\hat{\boldsymbol{\theta}}_i$, as given in Equations (5.12) and (5.11), respectively. The overall α was chosen to be $\alpha_{overall} = 0.05$. The UCL for the chart based on $T_{D,i}^2$ is 13.51 and the UCL for the chart based on the MVE (obtained via simulation) is approximately 23. The two charts are given in Figure 5.5.

The T^2 chart based on the $T_{D,i}^2$ statistic does not signal, whereas observations 6, 20, 22, 24, 26, and 45 produced a signal in the chart based on $T_{MVE,i}^2$. Recall that the T^2 chart based on $T_{D,i}^2$ statistics is not as powerful in detecting multivariate outliers as the chart based on the $T_{MVE,i}^2$ statistics. The observations that signal are extreme in both $\theta_{0,i}$ and $\theta_{1,i}$. In Figure 5.6, the 44 ordered pairs $\hat{\theta}_{0,i}$ and $\hat{\theta}_{1,i}$ are plotted. The 6 weeks that produced the signal are extreme in both $\theta_{0,i}$ and $\theta_{1,i}$. All are associated with variance profiles with a positive slope value $\hat{\theta}_{1,i}$ Hence, these observations are removed from the HDS, resulting in $m' = 38$.

WNLS was used to estimate the mean profiles, where the estimated weights are $(\hat{\sigma}_{ij}^{-2})^{GLIM}$ from Equation (5.28). The estimated mean profiles for all 44 weeks are given in Figure 5.7.

To check for appropriateness of the mean profile model, the LOF statistic of Equation (5.15) were calculated based on the weighted sums of squares. From Equation

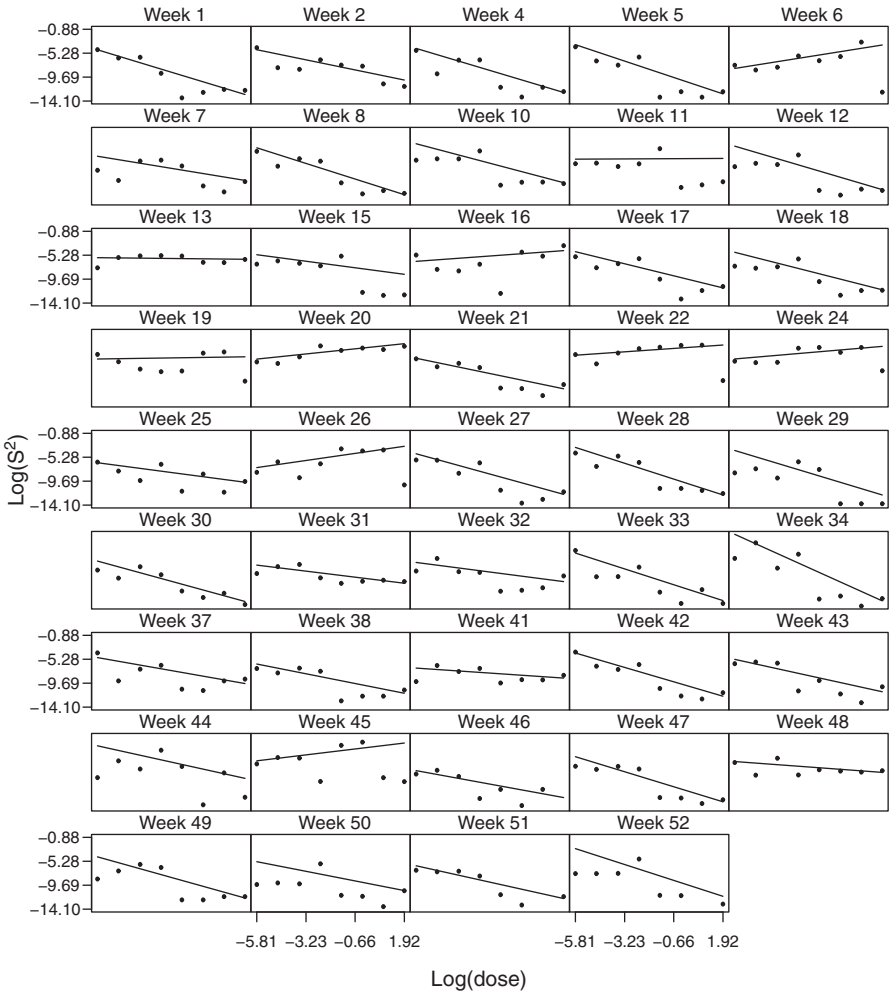

Figure 5.4 Estimated variance profiles for all 44 weeks.

(5.16) the approximate UCL associated with the chart is 5.26, and the chart is given in Figure 5.8.

Although weeks 21, 30, 32, and 33 produced a signal, only observations 21 and 32 were removed from the HDS. Weeks 30 and 33 exhibit only marginal LOF. This leaves $m'' = 36$ profiles remaining in the HDS.

Next unusual values of estimated parameters were examined using a T^2 chart based on the $T^2_{MVE,i}$ and $T^2_{D,i}$ statistics, as given in Equations (5.12) and (5.11), respectively. The UCLs associated with the charts are 38 and 17.68, respectively. The charts are given in Figure 5.9.

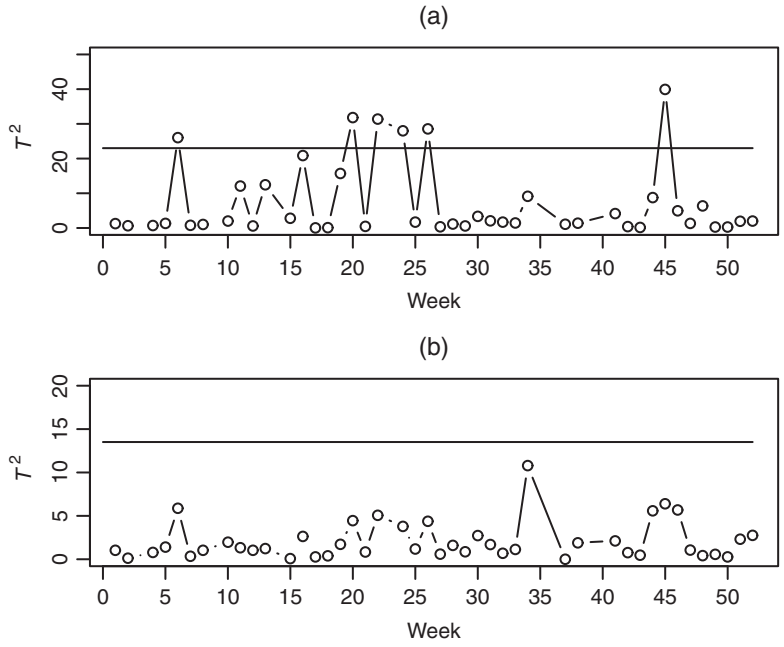

Figure 5.5 T^2 charts based on (a) $T^2_{MVE,i}$ and (b) $T^2_{D,i}$ to detect unusual variance profiles.

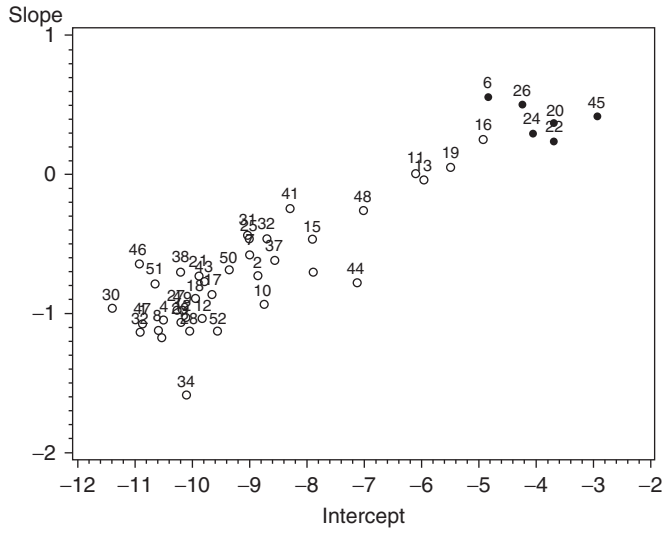

Figure 5.6 Scatterplot of values of $\hat{\theta}_{0,i}$ and $\hat{\theta}_{1,i}$ for all 44 weeks of the DuPont HDS. Those six observations that produced a signal in the T^2 chart are indicated with the solid black dot.

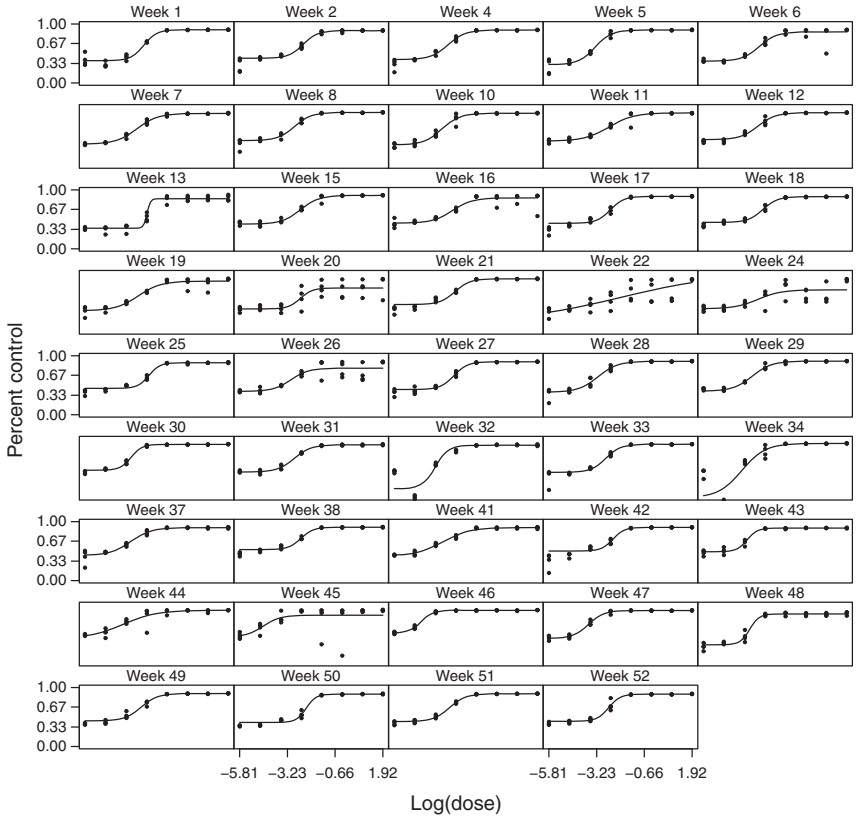

Figure 5.7 Estimated mean profiles based on estimated weights for all 44 weeks.

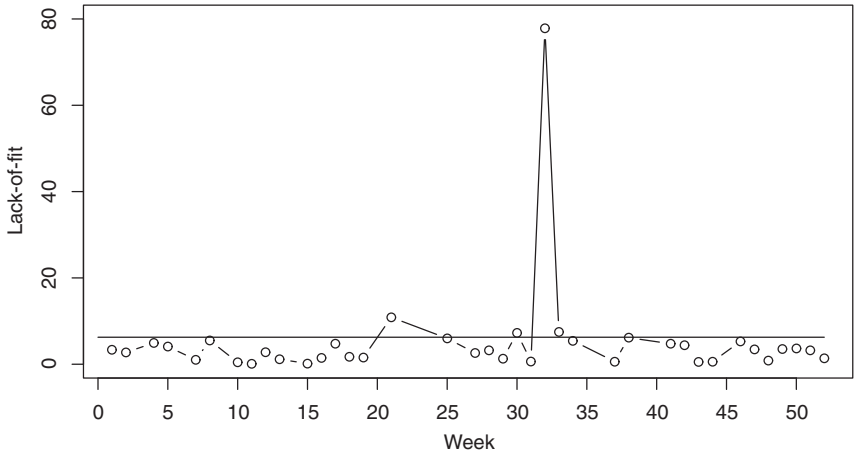

Figure 5.8 Lack-of-fit chart based on the weighted sums of squares.

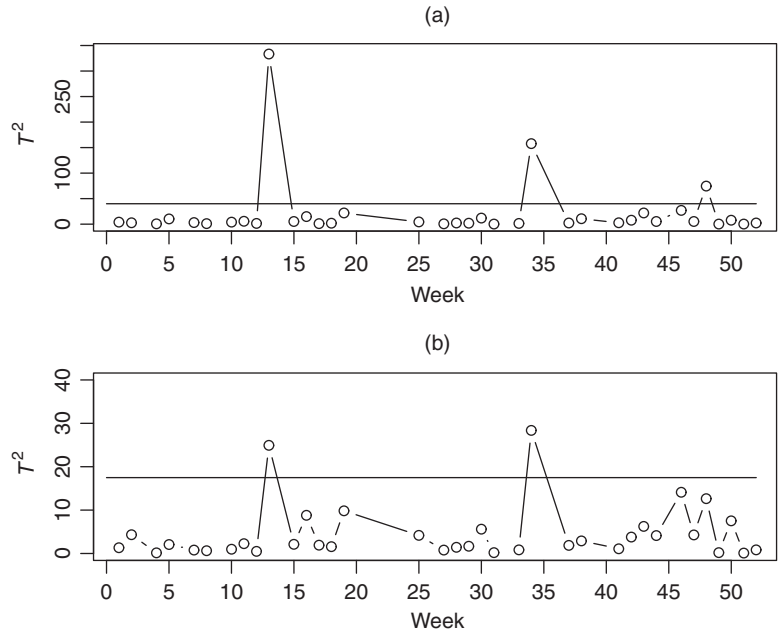

Figure 5.9 T^2 charts based on (a) $T^2_{MVE,i}$ and (b) $T^2_{D,i}$ to detect unusual values of the parameters estimates of the mean profiles.

Weeks 13, 34, and 48 produced a signal in the T^2 chart based on $T^2_{MVE,i}$, and observations 13 and 34 produced a signal in the chart based on $T^2_{D,i}$. All three of these weeks were removed from the HDS. Subsequently, a second T^2 chart was created to further check for out-of-control profiles after removing these weeks. The new UCLs for the two charts are 40 and 17.49, respectively. The charts are given in Figure 5.10.

Week 46 was identified as being out of control and was removed from the HDS. The remaining weeks were presumed to be in-control. A plot of the in-control estimated mean profiles is given in Figure 5.11.

The remaining $m^* = 32$ profiles are now used to estimate the in-control mean vector and variance–covariance matrix of $\hat{\boldsymbol{\theta}}_i$ and the in-control mean vector and variance–covariance matrix of $\hat{\boldsymbol{\beta}}_i$ for future process monitoring. From Equations (5.29) and (5.30), the estimates for the variance profile are calculated as

$$\boldsymbol{\mu}_\theta = \begin{bmatrix} -9.326028 \\ -0.765682 \end{bmatrix}$$

and

$$\boldsymbol{\Sigma}_\theta = \begin{bmatrix} 2.4730289 & 0.5147257 \\ 0.5147257 & 0.1396993 \end{bmatrix}.$$

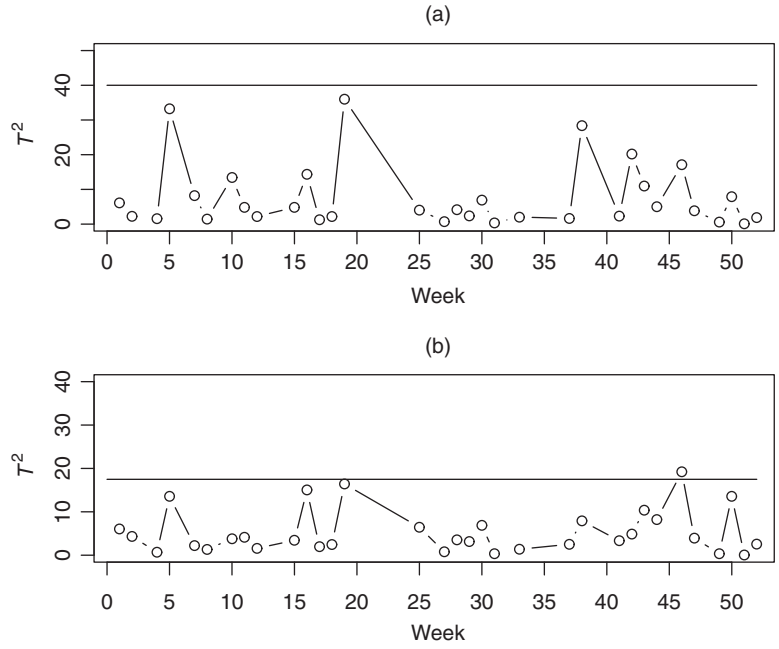

Figure 5.10 Second set of T^2 charts based on (a) $T^2_{MVE,i}$ and (b) $T^2_{D,i}$ to detect unusual values of the parameters estimates of the mean profiles.

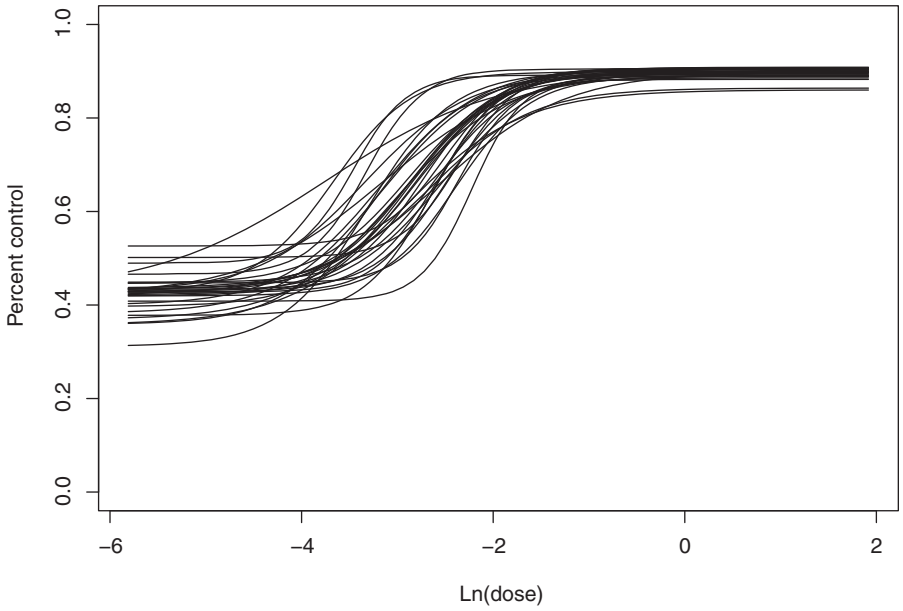

Figure 5.11 Estimated in-control mean profiles.

From Equations (5.17) and (5.19), the estimates for the mean profile are calculated as

$$\boldsymbol{\mu}_\beta = \begin{bmatrix} 0.8959855 \\ 2.3857821 \\ 0.0608633 \\ 0.4227484 \end{bmatrix}.$$

and

$$\boldsymbol{\Sigma}_\beta = \begin{bmatrix} 0.0001282 & -0.000134 & -0.000055 & 0.0000786 \\ -0.000134 & 0.4280911 & 0.0067914 & 0.0120498 \\ -0.000055 & 0.0067914 & 0.0004831 & 0.0002597 \\ 0.0000786 & 0.0120498 & 0.0002597 & 0.0017581 \end{bmatrix}.$$

APPENDIX

The function for $\beta(m, p, i)$ is

$$\beta(m, p, i) = I_{\{i=1,m\}} \left(\frac{p}{2} - \frac{1}{a_{11}(m - b_{11})} \right) + I_{\{i=2,\ldots,m-1\}} (a_{12} p + b_{12}),$$

and the function for $\gamma(m, p, i)$ is given by

$$\gamma(m, p, i) = I_{\{i=1,m\}} a_{21} + I_{\{i=2,\ldots,m-1\}} \left[a_{22} \left(i - \frac{m+1}{2} \right)^2 + b_{22} \right],$$

where

$$I_{\{i=1,m\}} = \begin{cases} 1 & \text{if } i = 1 \text{ or } i = m \\ 0 & \text{otherwise} \end{cases}$$

$$I_{\{i=2,\ldots,m-1\}} = \begin{cases} 1 & \text{if } 2 \le i \le m-1 \\ 0 & \text{otherwise} \end{cases}$$

$$a_{11} = 6.356 e^{-0.825p} + 0.06$$
$$b_{11} = 0.5564p + 0.9723$$
$$a_{12} = 0.54 - 0.25 e^{-0.25(m-15)}$$
$$b_{12} = -0.085 + 0.2 e^{-0.2(m-22)}$$
$$a_{21} = (-0.5m + 2)p + \frac{1}{3}(m+3)(m-5)$$
$$a_{22} = 0.99 + 0.38 e^{0.38(p-13.5)} - \frac{1}{0.25 e^{-0.25(p-10)} \left(m - 11 + \frac{(p-7)^2}{3} \right)}$$
$$b_{22} = (0.07 e^{-0.07(m-42)} - 1.95)p + 0.0833 m^2$$

REFERENCES

Aitkin, M. (1987) Modelling variance heterogeneity in normal regression using GLIM. *Applied Statistics*, **36**, 332–339.

Arbogast, P. G. and Bedrick, E. J. (2004) Model-checking techniques for linear models with parametric variance functions. *Technometrics*, **46**, 404–410.

Bellio, R., Jensen, J. E., and Seiden, P. (2000) Applications of likelihood asymptotics for nonlinear regression in herbicide bioassays. *Biometrics*, **56**, 1204–1212.

Brill, R. V. (2001) A case study for control charting a product quality measure that is a continuous function over time. Presentation at the 47th Annual Fall Technical Conference, Toronto, Ontario.

Carroll, R. J. and Ruppert, D. (1988) *Transformation and Weighting in Regression*. Chapman and Hall, New York.

Chou, Y. -M., Mason, R. L., and Young J. C. (1999) Power comparisons for a hotelling's T^2 statistic. *Communications in Statistics, Part B—Simulation and Computation*, **28**, 1031–1050.

Davidian, M. and Carroll, R. J. (1987) Variance function estimation. *Journal of the American Statistical Association*, **82**, 1079–1091.

Gallant, A. R. (1987) *Nonlinear Statistical Models*. John Wiley & Sons, Inc., New York.

Gnanadesikan, R. and Kettenring, J. R. (1972) Robust estimates, residuals, and outlier detection with multiresponse data. *Biometrics*, **28**, 81–124.

Hawkins, D. M. and Merriam, D. F. (1974) Zonation of multivariate sequences of digitized geologic data. *Mathematical Geology*, **6**, 263–269.

Holmes, D. S and Mergen, A. E. (1993) Improving the performance of the T^2 control chart. *Quality Engineering*, **5**, 619–625.

Jensen, W. A., Birch, J. B., and Woodall, W. H. (2007) High breakdown estimation methods for phase I multivariate control charts. *Quality and Reliability Engineering International*, **23** (5), 615–629.

Kang, L. and Albin, S. L. (2000) Online monitoring when the process yields a linear profile. *Journal of Quality Technology*, **32**, 418–426.

Mahmoud, M. A. and Woodall, W. H. (2004) Phase I analysis of Linear Profiles with Calibration Applications. *Technometrics*, **46**, 380–391.

Mason, R. L., Chou, Y. M., and Young, J. C. (2001), Applying Hotelling's T^2 Statistic to Batch Processes, *Journal of Quality Technology* **33**, pp. 466–479.

Mason, R. L. and Young, J. C. (2002) *Multivariate Statistical Process Control with Industrial Applications*. SIAM, Philadelphia.

Myers, R. H. (1990) *Classical and Modern Regression with Applications*. 2nd Edition. Duxbury Press, Belmont, California.

Neill, J. W. (1988) Testing for lack-of-fit in nonlinear regression. *The Annals of Statistics*, **16**, 733–740.

Nelson, L. S. (1983) Exact critical values for the analysis of means. *Journal of Quality Technology*, 15, 40–44.

Rousseeuw, P. J. (1984) Least median of squares regression. *Journal of the American Statistical Association*, **79**, 871–880.

Schabenberger, O. and Pierce F. J. (2002) *Contemporary Statistics for the Plant and Soil Sciences*. CRC Press, Boca Raton, Florida.

Seber, G. A. F. and Wild, C. J. (1989) *Nonlinear Regression*. John Wiley & Sons, Inc., New York.

Sullivan, J. H. and Woodall, W. H. (1996) A comparison of multivariate control charts for individual observations. *Journal of Quality Technology*, **28**, 398–408.

Tracy, N. D., Young, J. C., and Mason, R. L. (1992) Multivariate control charts for individual observations. *Journal of Quality Technology*, **24**, 88–95.

Vaghefi, A., Tajbakhsh, S. D., and Noorossana, R. (2009) Phase II monitoring of nonlinear profiles. *Communications in Statistics—Theory and Methods*, **38**, 1834–1851.

Vargas J. A. (2003) Robust estimation in multivariate control charts for individual observations, *Journal of Quality Technology*, **35**, 367–376.

Walker, E. and Wright, S. (2002) Comparing curves using additive models. *Journal of Quality Technology*, **34**, 118–129.

Williams, J. D., Woodall, W. H., Birch, J. B. and Sullivan, J. H. (2006) Distribution of hotelling's T^2 statistic based on successive differences. *Journal of Quality Technology*, **38**,217–229.

Williams, J. D., Birch, J. B., Woodall, W. H., and Ferry, N. M. (2007) Statistical monitoring of heteroscedastic dose-response profiles from high-throughput screening. *Journal of Agricultural Environmental, and Biological Statistics*, **12**, 216–235.

Williams, J. D., Sullivan, J. H., and Birch, J. B. (2009) Maximum value of hotelling's T^2 statistic based on the successive differences covariance matrix estimator. *Communications in Statistics: Theory and Methods*, **38** (4), 471–483.

Williams, J. D., Woodall, W. H., and Birch, J. B. (2007) Statistical monitoring of nonlinear product and process quality profiles. *Quality & Reliability Engineering International*, **23**, 925–941.

Wludyka, P. S. and Nelson, P. R. (1997) An analysis-of-means-type test for variances from normal populations. *Technometrics*, **39**, 274–285.

Young, T. M., Winistorfer, P. M., and Wang, S. (1999) Multivariate control charts of MDF and OSB vertical density profile attributes. *Forest Products Journal*, **49**(5), 79–86.