

UNIVERSITA' DEGLI STUDI DI PADOVA
FACOLTA' DI SCIENZE STATISTICHE
TESI DI LAUREA SPECIALISTICA IN STATISTICA E INFORMATICA



APPROCCI PARAMETRICI E NON PARAMETRICI PER LO STUDIO DI
PROFILI LINEARI E DI PROFILI LINEARI GENERALIZZATI

Relatrice: Prof.ssa Giovanna Capizzi

Laureando: Candeo Enrico

ANNO ACCADEMICO 2009-2010

Sommario

APPROCCI PARAMETRICI E NON PARAMETRICI PER LO STUDIO DI PROFILI LINEARI E DI PROFILI LINEARI GENERALIZZATI	1
Capitolo 1 Introduzione	7
Capitolo 2: Modelli e approcci di sorveglianza dei profili lineari.....	9
Capitolo 3: La Fase 2.....	17
3.1 Il caso Multivariato: le carte T^2 e la combinazione EWMA e R	17
3.2 Il caso univariato: le tre carte EWMA	19
3.3 Comparazione in termini di ARL	21
Capitolo 4 Approccio “Change Point”	25
4.1 Test basato sul rapporto di verosimiglianza (LR).....	25
4.2 Test basato sul rapporto di verosimiglianza generalizzato (GLR).....	28
Capitolo 5: Sorveglianza di profili lineari generali	33
5.1 Profilo lineare generale	33
5.2 Le carte MEWMA e EWMA₃ a confronto	35
5.3 Carta di controllo MEWMA con intervalli di campionamento variabili.....	37
5.4 Diagnostica dei profili lineari generali.	39
Capitolo 6: Approcci non parametrici alla sorveglianza dei profili lineari generali	43
6.1 Modello non parametrico	44
6.2 Lo schema di controllo NEWMA.....	44
6.3 Diagnostica nei profili.....	46
6.4 Comparazione della performance	49
6.5 Lo schema di controllo ACP	53
7. Illustrazione di un’applicazione: Il caso DRIE	57
7.1 L’applicazione del metodo MEWMA.....	59
7.2 L’applicazione del metodo NEWMA	61

7.3. L'applicazione del metodo ACP.	65
8. Conclusioni.....	67
Bibliografia	69

Capitolo 1 Introduzione

Molte applicazioni del controllo statistico di processo (SPC) assumono che questo sia adeguatamente rappresentato dalla distribuzione univariata di una specifica caratteristica di qualità o più in generale dalla distribuzione multivariata di un vettore costituito da più caratteristiche. In molte applicazioni aziendali si è visto che la qualità del processo è descritta in maniera più appropriata dalla relazione, nel seguito denominata “profilo”, che c'è tra una variabile risposta e una o più variabili esplicative. La sorveglianza dei profili è un'applicazione del controllo di processo ampiamente discussa nella recente letteratura sul controllo statistico di processo. L'analisi dei profili si suddivide in 2 momenti, detti Fase 1 e Fase 2. Raccolto un fissato numero di osservazioni del processo in esame, l'obiettivo della Fase 1 è quello di studiarne la struttura e determinarne la stabilità, eventualmente rimuovendo osservazioni anomale di cui si conoscono le cause. Dunque si stimano i valori dei parametri del processo In-Controllo che vengono usati nel disegno delle carte di controllo applicate durante la Fase 2. La performance di Fase 1 è spesso misurata in termini di probabilità di ottenere almeno un segnale di fuori controllo. Dal momento che l'interesse primario per la Fase 2 è quello di riconoscere il prima possibile il cambiamento di un parametro della regressione di un dato profilo, una misura di performance spesso utilizzata è l'*average run length* (ARL); dove per *run length* si indica il numero di osservazioni necessarie perché la carta dia un segnale di fuori controllo dopo che questo è realmente avvenuto.

L'obiettivo della tesi è quello di illustrare i vari approcci di tipo parametrico e non parametrico con la costruzione delle carte di controllo per la sorveglianza di profili lineari e di concludere con la presentazione di un esempio finalizzato a comprendere al meglio le metodologie descritte. La tesi si sviluppa a partire dal capitolo II dove prendiamo in esame le metodologie di analisi per la Fase 1: il caso multivariato con le carte T^2 e una combinazione di carte EWMA e R sviluppate di Kang e Albin (2000) e il caso univariato con le tre carte EWMA di Kim et al. (2003). Il capitolo III descrive e discute l'adattamento di tali metodologie alla sorveglianza di Fase 2. Il IV capitolo prende in esame due approcci utili alla costruzione delle carte per la sorveglianza di profili: l'approccio “*Change-Point*” e quello basato sul test rapporto di verosimiglianza

generalizzato, GLR. Nei due capitoli successivi vengono discussi gli approcci parametrici e non parametrici per i profili lineari generalizzati. In particolare, nel capitolo V presentiamo la carta con approccio parametrico MEWMA di Zou et al. (2007) mentre nel capitolo VI la carta non parametrica NEWMA di Zou et al.(2008) e quella ACP di Zou et al. (2008). Nel capitolo VII viene proposto un esempio che applica le ultime tre metodologie allo studio di un caso reale, il processo di incisioni di microcircuiti denominato DRIE, *“deep reactive ion etching”*. Infine nell’ultimo capitolo presentiamo le conclusioni finali.

Capitolo 2: Modelli e approcci di sorveglianza dei profili lineari

In questo capitolo presentiamo i modelli e gli approcci per l'analisi di Fase 1 dei profili lineari. Il disegno di uno schema di sorveglianza consiste nell'individuazione degli intervalli decisione (nel seguito limiti di controllo) a partire dai quali sia possibile concludere che il processo sia in controllo o fuori controllo. Il calcolo di tali limiti si effettua sotto l'ipotesi che il processo d'interesse sia stabile e in controllo. Quando i parametri del processo sono incogniti, gli intervalli di decisione degli schemi di controllo vengono stimati a partire da osservazioni campionarie. In particolare si estraggono dal processo campioni, detti di Fase 1, per verificare la stabilità, rilevare la presenza di osservazioni fuori controllo, stimare i parametri del processo incogniti e calcolare i limiti di controllo cosiddetti di prova.

Qualora una o più osservazioni di Fase 1 cadano fuori tali limiti, identificate le cause, queste osservazioni si eliminano dal data set preliminare e si ricalcolano stime e limiti di controllo. Si ripetono questi passaggi fino a quando il data set non viene epurato da tutti gli eventuali fuori-controllo.

Nel caso della sorveglianza di profili lineari si ipotizza che l'output del processo produttivo, oggetto di sorveglianza, sia una variabile casuale Y funzione lineare semplice di una variabile indipendente X , ovvero:

$$Y = A_{0j} + A_{1j}X_{ij} + \epsilon_{ij}, \quad i = 1, 2, \dots, n \text{ e } j = 1, 2, \dots, m, \quad [1]$$

dove A_{0j} e A_{1j} sono parametri della funzione e ϵ_{ij} sono errori casuali assunti indipendenti e identicamente distribuiti con distribuzione normale di media 0 e varianza σ_j^2 . I parametri in controllo per l'intercetta A_{0j} , il coefficiente angolare A_{1j} , e termine di errore ϵ dell'equazione [1] sono ignoti. Assumiamo che i valori di X siano fissati e identici per ogni singolo campione. I valori in controllo del processo sono ignoti. Uno degli obiettivi iniziali in Fase 1 è quello di identificare e rimuovere dal data set tutti i profili fuori controllo dopo che sono state rilevate possibili cause dovute al malf funzionamento del processo.

Le stime dei minimi quadrati per A_0 e A_1 per il campione j sono calcolate usando le formule:

$$a_{0j} = \bar{Y}_j - a_{1j}\bar{X}_j, \quad a_{1j} = S_{XY(j)}/S_{XX(j)},$$

dove $\bar{Y}_j = \sum_{i=1}^{n_j} Y_{ij}/n_j$, $\bar{X}_j = \sum_{i=1}^{n_j} X_{ij}/n_j$, $S_{XY(j)} = \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)Y_{ij}$, e $S_{XX(j)} = \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2$. La varianza σ_j^2 viene stimata dalla j^{th} media degli errori al quadrato MSE_j , dove $MSE_j = SSE_j/(n_j - 2)$. Qui $SSE_j = \sum_{i=1}^{n_j} e_{ij}^2$ è la somma dei quadrati dei residui, dove $e_{ij} = Y_{ij} - a_{0j} - a_{1j}X_{ij}$, $i = 1, 2, \dots, n_j$. La quantità $(n - 2)MSE_j/\sigma^2$ è distribuita come un chi-quadro con $(n - 2)$ gradi di libertà indipendentemente da a_{0j} e a_{1j} .

Gli stimatori dei minimi quadrati a_{0j} e a_{1j} sono distribuiti come una normale bivariata con media e matrice di varianza e covarianza date da

$$\mu = (A_0, A_1)^T$$

e:

$$\Sigma = \begin{bmatrix} \sigma_0^2 & \sigma_{01}^2 \\ \sigma_{01}^2 & \sigma_1^2 \end{bmatrix}, \quad [2]$$

rispettivamente.

Nella matrice [2] $\sigma_0^2 = \sigma^2(1/n_j + \bar{X}_j^2/S_{XX(j)})$, $\sigma_1^2 = \sigma^2/S_{XX(j)}$ e $\sigma_{01}^2 = -\sigma^2\bar{X}_j/S_{XX(j)}$ sono le varianze di a_{0j} , a_{1j} e la covarianza tra a_{0j} e a_{1j} .

Kang e Albin (2000) sviluppano un approccio multivariato per la sorveglianza dei profili lineari sotto l'assunzione che μ e Σ siano ignoti. Il loro approccio T^2 è basato su un vettore di stime di minimi quadrati per l'intercetta di Y e il suo coefficiente angolare. La statistica T^2 di riferimento per la sorveglianza dei parametri dell'intercetta e del coefficiente angolare è:

$$T_j^2 = \frac{m}{m-1} (z_j - \bar{z})^T S_2^{-1} (z_j - \bar{z}), \quad j = 1, 2, \dots, m, \quad [3]$$

dove $z_j = (a_{0j}, a_{1j})^T$, $\bar{z} = (\bar{a}_0, \bar{a}_1)^T$, $\bar{a}_0 = \sum_{j=1}^k a_{0j}/k$, $\bar{a}_1 = \sum_{j=1}^k a_{1j}/k$ e $S_2 = \begin{pmatrix} \hat{\sigma}_0^2 & \hat{\sigma}_{01} \\ \hat{\sigma}_{01} & \hat{\sigma}_1^2 \end{pmatrix}$, con $\hat{\sigma}_0^2, \hat{\sigma}_{01}, \hat{\sigma}_1^2$ indicanti le stime della varianza di a_{0j} , della covarianza tra a_{0j} e a_{1j} , e della varianza di a_{1j} . Questi stimatori si ottengono sostituendo a σ^2 la sua stima MSE. Il limite superiore della carta usata da Kang e Albin (2000) è:

$$UCL = 2F_{2,m(n-2),a}$$

Dove $F_{2,m(n-2),a}$ è il percentile $100(1 - a)$ di una distribuzione F con **2** e $m(n - 2)$ gradi di libertà. Quando la statistica [3] non è più grande del limiti UCL si conclude che i parametri della funzione di regressione sono in controllo. Si noti che i valori della statistica T^2 sono tra loro dipendenti poiché basati sulle stesse stime trovate dallo stesso campione di m profili, ma studi di simulazione mostrano che la marginale di una distribuzione F del test può approssimare al meglio la probabilità di un falso allarme.

Il secondo metodo di controllo proposto da Kang e Albin (2000) è di applicare alcuni schemi di sorveglianza standard ai residui di regressione ottenuti dal campione j :

$$e_{ij} = y_{ij} - A_0 - A_1 x_i, \quad i = 1, 2, \dots, n.$$

In particolare si suggerisce di combinare di una carta di tipo *EWMA* (*Exponentially Weighted Moving Average*, Roberts, 1959; Lucas et Saccucci, 1990) con una carta R per sorvegliare le variazioni dei parametri del processo. La combinazione di queste due carte è proposta per due motivi: il primo è che per identificare cambiamenti del processo di varianza la sola carta *EWMA* non è adeguata. L'altra ragione è che una carta *EWMA* basata sulla media dei residui non è sensibile ai cambiamenti in A_0 e A_1 dove la grandezza dei residui tende a essere grande, ma la loro media tende a essere piccola. Questo può succedere, per esempio, quando l'inclinazione della retta cambia, ma non la media dei valori di Y .

La media dei residui per il j^{th} campione è:

$$\bar{e}_j = \sum_{i=1}^n e_{ij} / n.$$

La statistica di controllo per la carta EWMA è la media pesata delle medie dei residui del campione j e le medie dei residui precedenti ed è data da:

$$z_j = \theta \bar{e}_j + (1 - \theta)z_{j-1},$$

dove θ ($0 < \theta \leq 1$) è una costante di lisciamento e $z_0 = 0$. Un segnale di fuori controllo è dato se il valore della statistica z_j cade al di fuori dei limiti di controllo superiore o inferiore così costruiti:

$$LCL = -L\sigma \sqrt{\frac{\theta}{(2-\theta)n}}, \quad UCL = L\sigma \sqrt{\frac{\theta}{(2-\theta)n}}, \quad [4]$$

dove $L > 0$ è una costante scelta per dare uno specifico valore dell'ARL in controllo. Quando la statistica cade fuori dai limiti di controllo si conclude che vi è stata una variazione nella media dei residui.

Per la carta R, Kang e Albin (2000) calcolano il range campionario $R_j = \max(e_{ij}) - \min(e_{ij})$, $j = 1, 2, \dots$, e i seguenti limiti inferiore e superiore:

$$LCL = \sigma(d_2 - Ld_3) \quad e \quad UCL = \sigma(d_2 + Ld_3), \quad [5].$$

I valori di d_2 e d_3 sono costanti e dipendono dal numerosità del campione. Questi valori possono essere trovati, tra tanti, anche nelle tabelle in Ryan (2000) o Montgomery (2001). Uno svantaggio di questo approccio è che se $n < 7$ non c'è limite di controllo inferiore e quindi non possono essere rilevate variazioni negative della varianza.

Per applicare questi due metodi al controllo della Fase 1, Kang e Albin modificano i limiti di controllo [4] e [5] sia per la carta Ewma che per la carta R sostituendo σ con \sqrt{MSE} .

$$LCL = -L\sqrt{MSE} \sqrt{\frac{\theta}{(2-\theta)n}}, \quad UCL = L\sqrt{MSE} \sqrt{\frac{\theta}{(2-\theta)n}},$$

$$LCL = \sqrt{MSE}(d_2 - Ld_3) \quad e \quad UCL = \sqrt{MSE}(d_2 + Ld_3).$$

Una difficoltà nell'applicare una carta EWMA in Fase 1 è che più gruppi di osservazioni potrebbero contribuire alla segnalazione di un unico fuori controllo e la possibilità di una loro eliminazione non è chiaramente definita.

Stover e Brill (1998) propongono a loro volta due metodi per la Fase 1. Il primo è un approccio simile al metodo T^2 di Kang e Albin (2000).

La statistica test di riferimento per il controllo dei parametri della funzione è:

$$T_j^2 = (z_j - \bar{z})^T S_1^{-1} (z_j - \bar{z}), \quad j = 1, 2, \dots, m \quad [6]$$

Il cambiamento sta nel fatto che gli stimatori della matrice di varianza e covarianza sono differenti. In particolare

$$S_1 = \begin{pmatrix} S_0^2 & S_{01} \\ S_{01} & S_1^2 \end{pmatrix},$$

dove $S_0^2 = \sum_{j=1}^m (a_{0j} - \bar{a}_0)^2 / (m - 1)$ è la varianza di a_{0j} , $S_1^2 = \sum_{j=1}^m (a_{1j} - \bar{a}_1)^2 / (m - 1)$ è la varianza di a_{1j} e $S_{01} = \sum_{j=1}^m (a_{0j} - \bar{a}_0)(a_{1j} - \bar{a}_1) / (m - 1)$ è la covarianza tra a_{0j} e a_{1j} . Il limite di controllo superiore, ottenuto da una distribuzione T^2 e dalla relazione con la distribuzione F , è:

$$UCL = 2(m + 1)(m - 1)F_{2, m-2, \alpha} / m(m - 2).$$

Resta il fatto che la distribuzione T^2 non è appropriata marginalmente per la statistica [6], dato che si stima la matrice di covarianza della popolazione usando la matrice di Fase 1 sugli stessi dati del campione, quindi tramite simulazioni si è visto che una più adeguata stima dell'errore di primo tipo del modello è data dalla distribuzione *Beta*. Il limite di controllo più adeguato è dato quindi da:

$$UCL = (m - 1)^2 B_{1, (m-3)/2, \alpha} / m.$$

Il secondo metodo proposto da Stover e Brill (1998) è una carta univariata per il controllo delle componenti principali, corrispondenti a vettori contenenti le stime dei parametri di regressione per ogni campione. Questo schema non viene sviluppato perché non se ne consiglia l'utilizzo in quanto non è capace di individuare cambiamenti dei parametri dell'intercetta e del coefficiente angolare.

Kim, Mahmoud and Woodall (2003) hanno proposto un altro approccio per l'analisi di Fase 1 dei profili lineari. Gli assunti di base per applicare questo tipo di metodo sono quelli di codificare i valori delle X in modo che la media dei valori codificati dia 0. Così

facendo gli stimatori per l'intercetta e per il coefficiente angolare sono indipendenti. In questo modo si ha la possibilità di sorvegliare ognuno dei tre parametri di regressione separatamente con una carta Shewhart. Il risultato del modello per la regressione lineare diventa:

$$Y_{ij} = B_{0j} + B_{1j}X'_{ij} + \epsilon_{ij}, \quad i = 1, 2, \dots, n_j, \quad j = 1, 2, \dots, m$$

dove $B_{0j} = A_{0j} + A_{1j}\bar{X}_j$, $B_{1j} = A_{1j}$, e $X'_{ij} = (X_{ij} - \bar{X}_j)$. In questo caso gli stimatori dei minimi quadrati per i parametri della regressione per il campione j sono $b_{0j} = \bar{y}_j$ e $b_{1j} = a_{1j} = S_{XY(j)}/S_{XX(j)}$. Per un processo in controllo b_{0j} e b_{1j} sono reciprocamente indipendenti e normalmente distribuiti di media B_0 e B_1 e varianza σ^2/n e $\sigma^2/S_{XX(j)}$ rispettivamente.

Assumiamo ancora che i valori di X siano gli stessi per tutti i campioni e il processo sia in controllo; si può vedere che la quantità $b_{0j} - \bar{b}_0$, dove $\bar{b}_0 = \sum_{j=1}^m b_{0j}/m$, ha una distribuzione normale di media 0 e varianza $((m-1)/nm)\sigma^2$. Inoltre $m(n-2)MSE/\sigma^2$ si distribuisce come una chi-quadro con $m(n-2)$ gradi di libertà. Visto quindi che queste due quantità sono indipendenti, $(b_{0j} - \bar{b}_0)/\sqrt{MSE \frac{m-1}{nm}}$ segue una distribuzione t con $m(n-2)$ gradi di libertà. Per questo sembra ragionevole usare una carta di controllo Shewhart per sorvegliare l'intercetta B_0 con i seguenti limiti di controllo:

$$LCL = \bar{b}_0 - t_{m(n-2), \alpha/2} \sqrt{\frac{(m-1)MSE}{nm}} \quad e \quad UCL = \bar{b}_0 + t_{m(n-2), \alpha/2} \sqrt{\frac{(m-1)MSE}{nm}}$$

dove $t_{m(n-2), \alpha/2}$ è il percentile $100(1 - \alpha/2)$ di una distribuzione t con $m(n-2)$ gradi di libertà.

Assumendo sempre che il processo sia in controllo si può vedere inoltre come la quantità $b_{1j} - \bar{b}_1$, dove $\bar{b}_1 = \sum_{j=1}^m b_{1j}/m$, ha una distribuzione normale di media 0 e varianza $((m-1)/mS_{XX})\sigma^2$. Si constata quindi che $(b_{1j} - \bar{b}_1)/\sqrt{MSE \frac{m-1}{mS_{XX}}}$ ha una distribuzione t con $m(n-2)$ gradi di libertà.

I limiti di controllo di una carta Shewhart per sorvegliare B_1 sono posti uguali a:

$$LCL = \bar{b}_1 - t_{m(n-2),\alpha/2} \sqrt{\frac{(m-1)MSE}{mS_{XX}}} \text{ e } UCL = \bar{b}_1 + t_{m(n-2),\alpha/2} \sqrt{\frac{(m-1)MSE}{mS_{XX}}}.$$

Con il processo sempre in controllo si può vedere che

$$F_j = MSE_j / MSE_{-j}$$

ha una distribuzione F con $n-2$ e $(m-1)(n-2)$ gradi di libertà dove $MSE_{-j} = \sum_{i \neq j}^m MSE_i (m-1)$. Quindi una carta Shewhart per sorvegliare la varianza del processo σ^2 consiste nel riportare sulla carta a quantità F_j con i seguenti limiti di controllo:

$$LCL = F_{(n-2),(m-1)(n-2),\alpha/2} \text{ e } UCL = F_{(n-2),(m-1)(n-2),1-\alpha/2}$$

Una carta di controllo matematicamente equivalente per sorvegliare la varianza del processo si può ottenere disegnando i valori di MSE_j in una carta con i seguenti limiti di controllo:

$$LCL = \frac{mF_{(n-2),(m-1)(n-2),\frac{\alpha}{2}}}{m-1 + F_{(n-2),(m-1)(n-2),\frac{\alpha}{2}}} MSE \text{ e } UCL = \frac{mF_{(n-2),(m-1)(n-2),1-\frac{\alpha}{2}}}{m-1 + F_{(n-2),(m-1)(n-2),1-\frac{\alpha}{2}}} MSE$$

Quando i valori di X variano da campione a campione, si possono ottenere i limiti di controllo dei parametri del processo in corrispondenza di una specifica probabilità di falso allarme usando algoritmi di simulazione. Prima di passare allo studio dei cambiamenti nell'intercetta o nel coefficiente angolare si raccomanda sempre di controllare con carte di controllo per valutare la stabilità della varianza del processo in controllo. Una volta ottenuta questa si passa a controllare i parametri della funzione di regressione. Il metodo di Kim et al. (2003) si basa sul riportare le statistiche dipendenti in ognuna delle tre carte di controllo disegnate per la sorveglianza dei tre parametri del processo. Per esempio, se E_i rappresenta l'evento che l' i^{th} errore quadratico medio ecceda i limiti di controllo visti sopra, allora E_i e E_j ($i \neq j$) non sono indipendenti. Questo è il caso anche per l'intercetta che per l'inclinazione. Come visto prima però sappiamo che studi di simulazione mostrano che queste dipendenze non rappresentano un problema per la buona approssimazione della probabilità di falso allarme.

Capitolo 3: La Fase 2

Fino ad ora abbiamo visto algoritmi per lo studio dei profili lineari in Fase 1, divisi per autore e per metodologia. Presenteremo in questo capitolo e in quelli successivi strumenti di indagine di Fase 2, riprendendo gli approcci parametrici multivariati T^2 , la combinazione *EWMA-R* di Kang e Albin (2000), il caso univariato e le tre carte *EWMA* di Kim et al. (2003), lo schema *MEWMA* di Zou, Wang, Tsung (2007) per profili generalizzati, concludendo con due approcci non parametrici *NEWMA* di Zou, Wang e Tsung (2008) e *ACP* di Zou, Qiu e Hawkins (2008).

Il modello di Fase 2 è presentato sotto questa struttura:

$$Y_{ij} = A_0 + A_1 X_{ij} + \varepsilon_{ij}, \quad i = 1, 2, \dots, n \text{ e } j = 1, 2, \dots,$$

Dove le ε_{ij} sono realizzazioni i.i.d. da una normale standard $N(0,1)$. I valori delle X in ognuno dei campioni sono assunti costanti note. I valori dei parametri in controllo A_0 , A_1 , σ^2 sono assunti noti o stimati durante la Fase 1 da una data set che ci si aspetta rifletta al meglio la performance in controllo.

La performance delle carte di controllo di Fase 2 è solitamente misurata da alcuni parametri della distribuzione della *run length*, il numero osservazioni che vengono osservate prima che la carta produca un fuori controllo. Nella letteratura è utilizzata solitamente la sua versione media, Average Run Length *ARL*, per il confronto tra metodi di Fase 2.

3.1 Il caso Multivariato: le carte T^2 e la combinazione *EWMA* e *R*

Kang e Albin (2000) riprendono le stesse metodologie di Fase 1 per sorvegliare i profili lineari in Fase 2. Il loro primo approccio è una carta bivariata T^2 basata sulle stime dei minimi quadrati per l'intercetta e il coefficiente angolare assumendo noti i valori dei parametri.

La statistica T^2 è quindi:

$$T_j^2 = (z_j - \mu)^T \Sigma^{-1} (z_j - \mu), \quad j = 1, 2, \dots,$$

dove $z_j = (a_{0j}, a_{1j})^T$ è il vettore delle stime, e μ e Σ sono definite come nel capitolo precedente. Quando il processo è in controllo, la statistica T^2 segue una distribuzione chi quadro con 2 gradi di libertà. Il limite superiore di controllo della carta che sorveglia i parametri della funzione di regressione, fissato da Kang e Albin (2000) è:

$$UCL = \chi_{2,\alpha}^2,$$

dove $\chi_{2,\alpha}^2$ è il percentile $100(1 - \alpha)$ di una distribuzione chi quadro con due gradi di libertà. In caso di cambiamenti dei parametri del processo, assumendo che le X siano le stesse per tutti i profili, la statistica T^2 segue una distribuzione chi quadro non centrata con 2 gradi di libertà con parametro di non centralità:

$$\tau = n(\lambda + \beta\bar{X})^2 + \beta^2 S_{XX},$$

dove λ e β sono i cambiamenti in intercetta e coefficiente angolare rispettivamente. Il valore esatto dell'ARL di questa carta di controllo T^2 si può calcolare usando la formula:

$$ARL = \frac{1}{\Pr(T_j^2 > \chi_{2,\alpha}^2)}.$$

Se i valori di X invece non sono gli stessi per tutti i campioni, l'ARL corrispondente a una specifica probabilità di falso allarme può essere stimato tramite simulazione.

Il secondo metodo di Fase 2 di Kang e Albin (2000) è una carta *EWMA* per sorvegliare la deviazione media da un regressione lineare in controllo. Le deviazioni dalla regressione in controllo ottenute dal campione j^{th} sono calcolate usando la formula:

$$e_{ij} = Y_{ij} - A_0 - A_1 X_{ij}, \quad i = 1, 2, \dots, n_j,$$

la cui deviazione media per il campione j^{th} è $\bar{e}_j = \sum_{i=1}^{n_j} e_{ij} / n_j$. La statistica di controllo per la carta *Ewma* è data da:

$$EWMA_j = \theta \bar{e}_j + (1 - \theta) EWMA_{j-1}, \quad [7]$$

dove $0 < \theta \leq 1$ è il parametro di lisciamento e $EWMA_0 = 0$. Assumendo che $n_j = n$, $j = 1, 2, \dots$, i limiti di controllo per questa carta *EWMA* sono:

$$LCL = -L_1 \sigma \sqrt{\theta/n(2 - \theta)} \quad e \quad UCL = L_1 \sigma \sqrt{\theta/n(2 - \theta)},$$

dove $L_1 > 0$ è una costante scelta per dare uno specifico ARL in controllo.

Kang e Albin (2000), come visto per la metodologia di Fase 1, suggeriscono di usare una combinazione della carta EWMA [7] con una carta R per valutare al meglio la variazione della retta di regressione. Per la carta R disegnano il range dei campioni $R_j = \max(e_{ij}) - \min(e_{ij})$, $j = 1, 2, \dots$, in una carta con i limiti fissati:

$$LCL = \sigma(d_2 - L_2 d_3) \quad e \quad UCL = \sigma(d_2 + L_2 d_3),$$

dove $L_2 > 0$ è una costante scelta per produrre uno specifico ARL in controllo, e d_2 e d_3 sono costanti che dipendono dalla grandezza n del campione (Ryan, 2000 e Montgomery, 2001). Un segnale è dato ogni volta che una delle due carte produce un fuori controllo. Uno svantaggio per l'approccio con la carta R è che se $n < 7$ non c'è un limite inferiore e quindi di conseguenza non si può cogliere una diminuzione in varianza.

Se la numerosità del campione non è la stessa per tutti i campioni, si possono determinare i limiti di controllo di una carta EWMA e R corrispondenti a un determinato ARL in controllo tramite simulazione.

3.2 Il caso univariato: le tre carte EWMA

Kim et al. (2003) propongono un altro metodo per sorvegliare la Fase 2 del processo assumendo noti i valori dei parametri. La loro idea è di codificare i valori delle X all'interno del campione come descritto da Kim, Mahmoud e Woodall (2003) nell'approccio Shewhart di Fase 1. Codificando in questo modo i valori delle X , si ottengono così stimatori indipendenti per la regressione. Kim et al. (2003) suggeriscono dunque di sorvegliare i coefficienti di regressione, intercetta e coefficiente angolare, e la varianza usando tre carte *EWMA* separate. Si ha un segnale di allarme dopo che una delle tre carte *EWMA*, per l'intercetta, per il coefficiente angolare e per la varianza, produce un segnale di fuori controllo.

I valori per la statistica per la carta *EWMA* per sorvegliare l'intercetta Y , B_0 , sono:

$$EWMA_{I(j)} = \theta b_{0j} + (1 - \theta)EWMA_{I(j-1)},$$

dove $0 < \theta < 1$ è una costante di lisciamento e $EWMA_{I(0)} = B_0$. Un segnale di fuori controllo è dato dalla carta non appena la statistica $EWMA_I$ viola i limiti di controllo:

$$LCL = B_0 - L_I \sigma \sqrt{\theta/n(2-\theta)} \text{ e } UCL = B_0 + L_I \sigma \sqrt{\theta/n(2-\theta)},$$

dove $L_I > 0$ è una costante selezionata per produrre una specifica ARL in controllo.

La statistica EWMA per sorvegliare invece il coefficiente angolare B_1 , è data da:

$$EWMA_{S(j)} = \theta b_{1j} + (1 - \theta)EWMA_{S(j-1)},$$

dove $0 < \theta < 1$ è una costante di lisciamento e $EWMA_{S(0)} = B_1$. I limiti di controllo per la carta sono dati da:

$$LCL = B_1 - L_S \sigma \sqrt{\theta/n(2-\theta)} \text{ e } UCL = B_1 + L_S \sigma \sqrt{\theta/n(2-\theta)},$$

dove $L_S > 0$ anche qui è una costante selezionata per produrre una specifica ARL in controllo.

In fine per sorvegliare la varianza del processo, Kim et al. (2003) propongono uno schema EWMA basato sull'approccio di Crowder e Hamilton (1992), ma che usa i valori di MSE_j al posto della varianza del campione per trovare il valore della statistica EWMA. In particolare, la statistica EWMA per sorvegliare σ^2 è:

$$EWMA_{E(j)} = \max\{\theta(\ln MSE_j) + (1 - \theta)EWMA_{E(j-1)}, \ln \sigma_0^2\},$$

dove $0 < \theta < 1$ è la costante di lisciamento e $EWMA_{E(0)} = \ln \sigma_0^2$. Qui σ_0^2 è il valore in controllo della varianza del processo σ^2 . Il limite di controllo per individuare un eventuale incremento della varianza del processo è:

$$UCL = \ln \sigma_0^2 + L_E [\theta \text{Var}(\ln MSE_j) / 2 - \theta]^{1/2},$$

dove $L_E > 0$ è scelto come visto prima per ottenere uno specifico valore dell'ARL in controllo e

$$\text{VAR}(\ln MSE_j) \approx \frac{2}{n-2} + \frac{2}{(n-2)^2} + \frac{4}{3(n-2)^3} - \frac{16}{15(n-2)^5}.$$

Si potrebbe sviluppare una procedura con doppio limite di controllo, se fosse considerato importante rilevare una diminuzione nell'errore della varianza come di solito succede in alcune applicazioni pratiche.

3.3 Comparazione in termini di ARL

In questa sezione presentiamo un confronto in termini di ARL tra i metodi presentati prima, quindi tra EWMA-R, T^2 , e le tre carte EWMA. Le carte presentate sono disegnate per avere lo stesso ARL in controllo pari a 200. La costante di liscio θ nelle tre carte $EWMA_{I(j)}$, $EWMA_{S(j)}$ e $EWMA_{E(j)}$ è posta uguale a 0.2 come per la carta EWMA usata da Kang e Albin (2000). Si osservi che, in generale, a piccole costanti di liscio corrisponde una più veloce rilevazione di piccoli cambiamenti del processo. Nei loro studi di simulazione sono state utilizzate un totale di 10000 repliche per stimare i valori dell'ARL.

Il modello usato da Kang e Albin (2000) per il profilo lineare è $y_{ij} = 3 + 2x_{ij} + \varepsilon_{ij}$, dove ε_{ij} sono realizzazioni i.i.d. di una distribuzione normale di media zero e varianza unitaria. I valori fissati delle X sono 2, 4, 6 e 8, con media \bar{x} pari a 5. Nel metodo proposto da Kim et al. (2003) invece, trasformano i valori delle X usando la forma alternativa del modello e ottengono la nuova funzione lineare $y_{ij} = 13 + 2x_i^* + \varepsilon_{ij}$, dove ε_{ij} sono sempre realizzazioni di una $N(0,1)$. I valori di x_i^* sono -3, -1, 1 e 3 con media $\bar{x}_i^* = 0$.

Kim et al (2003) indicano con $EWMA_3$ la combinazione delle tre carte EWMA. Per l'implementazione dell' $EWMA_3$ ognuna delle tre EWMA viene disegnata in modo da ottenere la stessa ARL in controllo quando applicate individualmente. In particolare, al fine di avere separatamente un'ARL in controllo approssimativamente pari a 585, per la carta $EWMA_I$ per sorvegliare l'intercetta di Y , L_I è scelto pari a 3.0156; per l' $EWMA_S$ per il coefficiente angolare, L_S è posto pari a 3.0109. In fine L_E viene scelto pari a 1.3723. La combinazione delle tre carte di controllo ha un ARL complessivo pari a 200.

Le carte EWMA e R usata da Kang e Albin (2000) nei loro studi di simulazione hanno lo stesso $L = 3.1151$ che porta a una ARL in controllo approssimativamente di 802 per la EWMA, che monitora la deviazione media, e di 261 per la R che non ha limite inferiore quando $n = 4$. Anche in questo caso la combinazione delle due carte risulta avere un ARL totale di 200.

Per il confronto sono stati studiati quattro tipi diversi di cambiamenti, rispettivamente cambiamenti nell'intercetta e nel coefficiente angolare per il modello $Y = A_{0j} +$

$A_{1j}X_{ij} + \epsilon_{ij}$, diminuzione della varianza del processo e deriva del coefficiente angolare per il modello $Y_{ij} = B_{0j} + B_{1j}X'_{ij} + \epsilon_{ij}$ per il quale i valori delle variabili indipendenti sono codificati per essere 0.

La Tabella 1 dà i valori dell'ARL per un cambiamento in A_0 , che equivale anche a uno stesso cambiamento in B_0 in termini di σ , ovvero $A_0 + \lambda\sigma$. La proposta di Kim et al. (2003) è di poco più performante dello schema $EWMA/R$ e di gran lunga migliore della carta T^2 nell'individuare piccoli e moderati cambiamenti nell'intercetta, al contrario per quando si realizzano cambiamenti di grande entità. Per migliorare la performance della carta $EWMA_3$ nel caso di grandi cambiamenti, Kim et al. (2003) suggeriscono di aggiungere dei limiti di controllo di tipo Shewhart ad ognuna delle tre carte $EWMA$ come presentato da Lucas e Saccucci (1990).

Tabella 1: Comparazione dell'ARL con shift nell'intercetta $A_0 + \lambda\sigma$

Chart	λ									
	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
EWMA/R	66.5	17.7	8.4	5.4	3.9	3.2	2.7	2.3	2.1	1.9
T^2	137.7	63.5	28.0	13.2	6.9	4.0	2.6	1.8	1.5	1.2
EWMA ₃	59.1	16.2	7.9	5.1	3.8	3.1	2.6	2.3	2.1	1.9

Nella Tabella 2 si considera il caso di un cambiamento nel coefficiente angolare, A_1 , sempre in termini di σ , $A_1 + \beta\sigma$. Il metodo $EWMA_3$ si comporta leggermente meglio del metodo $EWMA/R$ su tutto il range di shift considerati. Anche in questo caso la carta T^2 si comporta meglio solo per grandi cambiamenti. La differenza comunque è minima. L'uso di una carta $EWMA$ multivariata come proposto da Lowry (1992) produrrebbe risultati ancora più performanti, ma l'interpretazione di un segnale di fuori controllo non risulterebbe per nulla semplice.

Tabella 2: Comparazione in termini di ARL per uno shift del coefficiente angolare

Chart	β									
	0.025	0.050	0.075	0.100	0.125	0.150	0.175	0.200	0.225	0.250
EWMA/R	119.0	43.9	19.8	11.3	7.7	5.8	4.7	3.9	3.4	3.0
T^2	166.0	105.6	60.7	34.5	20.1	12.2	7.8	5.2	3.7	2.7
EWMA ₃	101.6	36.5	17.0	10.3	7.2	5.5	4.5	3.8	3.3	2.9

La Tabella 3 mostra l'ARL per un incremento di σ , da σ a $\gamma\sigma$. In questo caso la carta $EWMA/R$ di Kang e Albin (2000) dà risultati migliori della carta $EWMA_3$ perché il cambiamento della deviazione standard è individuato più efficientemente da una carta R .

Tabella 3: Comparazione di ARL per uno shift nella deviazione standard del processo

Chart	γ									
	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8	3.0
$EWMA/R$	34.3	12.0	6.1	3.9	2.9	2.3	1.9	1.7	1.5	1.4
T^2	39.6	14.9	7.9	5.1	3.8	3.0	2.5	2.2	2.0	1.8
$EWMA_3$	33.5	12.7	7.2	5.1	3.9	3.2	2.8	2.5	2.3	2.1

Da segnalare anche come, in questo caso la carta T^2 ha un performance soddisfacente anche se è una carta priva di memoria e non utilizza dunque le informazioni raccolte nel tempo. Concludendo, quando un incremento della deviazione standard porta a una incremento della varianza delle stime di intercetta e coefficiente angolare, utilizzare una $EWMA_5$ per lisciare gli stimatori non sembra essere di grande aiuto.

Un quarto caso è dato dal cambiamento del parametro di coefficiente angolare, δ , per l'equazione trasformata, da B_1 a $B_1 + \delta\sigma$

Tabella 4: Comparazione dell'ARL per uno shift per il coefficiente angolare B_1

Chart	δ									
	-0.2	-0.3	-0.4	-0.5	-0.6	-0.7	-0.8	-0.9	-1.0	
$EWMA/R$	76.7	33.7	15.3	7.5	4.2	2.6	1.8	1.4	1.2	
T^2	52.2	21.2	9.6	4.9	2.9	1.9	1.5	1.2	1.1	
$EWMA_3$	13.1	6.6	4.4	3.3	2.7	2.3	2.1	1.9	1.7	

Visto che è dimostrato che i valori dell'ARL per δ sono simmetrici intorno a $\delta = 0$, Kim et al (2003) considerano solo shift positivi. Quando il cambiamento nel coefficiente B_1 è di entità moderata, la carta $EWMA_3$ ha una performance migliore. Appena lo shift inizia a crescere ci ritroviamo nel caso in cui la carta T^2 migliora la sua performance. Anche in questo caso una correzione dei limiti con quelli tipo Shewhart o quelli con esatta varianza per la statistica $EWMA$ potrebbe portare dei miglioramenti.

Ultimo caso che però lascia pochi dubbi mostra che per cambiamenti simultanei per l'intercetta e slope il metodo $EWMA_3$ risponde meglio uniformemente rispetto a una carta di tipo $EWMA/R$.

Tabella 5: Confronto di ARL tra una $EWMA/R$ e una $EWMA$ in caso di shift combinato di Intercetta e Coefficiente angolare del modello

EWMA/R ARL EWMA ₃ ARL		δ									
		0.025	0.050	0.075	0.100	0.125	0.150	0.175	0.200	0.225	0.250
λ	0.05	179.1	169.9	156.6	140.8	123.2	105.1	88.8	73.4	60.2	49.6
		157.6	114.7	74.8	48.3	32.2	22.5	16.9	13.2	10.7	8.9
	0.10	139.5	133.6	125.4	115.5	103.5	90.4	78.3	65.7	55.6	46.3
		122.1	94.6	66.4	44.9	30.7	21.9	16.6	13.1	10.6	8.9
	0.15	96.8	94.2	90.3	85.1	78.5	70.9	63.0	55.3	47.7	40.9
		84.6	70.8	54.5	39.6	28.5	20.9	16.1	12.8	10.4	8.8
	0.20	64.8	63.8	62.1	59.7	56.6	52.9	48.5	44.0	39.2	34.6
		57.1	51.1	42.4	33.3	25.4	19.5	15.4	12.4	10.2	8.7
	0.25	44.3	43.8	42.9	41.8	40.3	38.4	36.1	33.6	30.8	28.1
		39.5	36.5	32.3	27.1	22.0	17.8	14.4	11.9	10.0	8.5
	0.30	31.0	30.8	30.5	29.9	29.2	28.3	27.1	25.7	24.2	22.5
		28.2	26.9	24.7	22.0	18.8	15.7	13.2	11.2	9.6	8.3
	0.35	22.9	22.8	22.6	22.2	21.9	21.4	20.7	19.9	19.1	18.0
		20.9	20.2	19.1	17.6	15.8	13.9	12.1	10.5	9.1	8.0
	0.40	17.4	17.3	17.2	17.1	16.8	16.6	16.2	15.8	15.3	14.7
		16.2	15.9	15.3	14.5	13.5	12.1	10.9	9.7	8.6	7.6
0.45	13.9	13.9	13.9	13.8	13.6	13.5	13.3	13.0	12.6	12.2	
	13.1	12.9	12.6	12.1	11.4	10.6	9.8	8.9	8.0	7.3	
0.50	11.5	11.5	11.4	11.3	11.3	11.1	11.0	10.8	10.6	10.3	
	10.8	10.8	10.6	10.3	9.9	9.3	8.7	8.1	7.5	6.9	

Capitolo 4 Approccio “*Change Point*”

Prima di proseguire con le carte di controllo per profili lineari generali e con metodi basati su metodi non parametrici, presentiamo un approccio utile per una loro migliore comprensione. Si tratta dell’approccio “*Change Point*” (Mahmoud et al. 2007) basato sul test rapporto di verosimiglianza per i profili lineari (Zou et al. 2007) e sul test rapporto di verosimiglianza generalizzato per profili lineari generali (Zou et al. 2008, Fan et al. 2001). I vantaggi di questi metodi sono un miglioramento della rilevazione di cambiamenti nei parametri del profilo e un miglioramento degli algoritmi di ricerca per determinare le fonti di variazione del processo e la localizzazione del punto di cambiamento.

4.1 Test basato sul rapporto di verosimiglianza (LR)

Presentiamo in questa sezione l’approccio *Change-Point* utile a testare la consistenza dei parametri di regressione di un data set composto da profili lineari.

In questa sezione prendiamo in esame l’ipotesi nulla che tutti i profili seguano un modello lineare semplice contro l’ipotesi alternativa che vi sia un cambiamento in uno o più parametri del modello dopo m_1 campioni, $m_1 = 1, 2, \dots, m - 1$. Per questo si propone un test basato sul test rapporto di verosimiglianza in un modello segmentato di regressione lineare semplice.

Il data set è composto come detto da m campioni casuali, dove ogni campione è composto da n_j coppie di osservazioni $(X_{ij}, Y_{ij}), i = 1, 2, \dots, n_j, n_j > 2, j = 1, \dots, m$. Il modello del profilo lineare che lega la variabile esplicativa X alla variabile risposta Y è:

Errore. L'origine riferimento non è stata trovata.

Il primo passo consiste nel combinare gli m profili in un unico campione di grandezza N . Quindi si applica il seguente modello di regressione lineare segmentato al campione combinato:

$$\begin{aligned}
Y_i &= A_{01} + A_{11}X_i + \varepsilon_i, \quad \theta_0 < i \leq \theta_1 \\
Y_i &= A_{02} + A_{12}X_i + \varepsilon_i, \quad \theta_1 < i \leq \theta_2 \\
&\vdots \\
Y_i &= A_{0s} + A_{1s}X_i + \varepsilon_i, \quad \theta_{s-1} < i \leq s,
\end{aligned}$$

dove $i = 1, 2, \dots, N$, θ_j è il punto di cambiamento tra i segmenti (di solito $\theta_0 = 0$ e $\theta_s = N$) e gli ε_i sono i termini dell'errore. La tecnica di regressione segmentata è usata per stimare il numero di segmenti s e la posizione del cambiamento θ_j . Fino a quando noi assumiamo che non vi siano stati cambiamenti di parametri all'interno di un campione, allora il punto di change point è riferito all'indice i , corrispondente alla fine del campione del profilo, i.e., $\theta_0 = 0$, $\theta_1 = n_1$, $\theta_2 = n_1 + n_2, \dots, \theta_m = N$. La statistica LRT usata per verificare l'ipotesi nulla di un singolo segmento contro l'alternativa di $s = 2$ segmenti è:

$$lrt_{m_1} = N \log \hat{\sigma}^2 - N_1 \log \hat{\sigma}_1^2 - N_2 \log \hat{\sigma}_2^2, \quad m_1 = 1, 2, \dots, m - 1, \quad [8]$$

dove $\hat{\sigma}^2$ è la stima di massima verosimiglianza (*MLE*) della varianza dell'errore per il modello di regressione stimato con tutti gli m campioni combinati in un unico campione N , $\hat{\sigma}_1^2$ è l'*MLE* della varianza dell'errore per il modello di regressione stimato con tutti i campioni precedenti a m_1 , combinati in un unico campione $N_1 = \sum_{j=1}^{m_1} n_j$ e $\hat{\sigma}_2^2$ invece si riferisce ai campioni successivi a m_1 messi assieme in un campione $N_2 = N - N_1$. Questo metodo può essere applicato ricorsivamente per identificare più punti di cambiamento nell'intero data set. Molto importante è notare che la statistica LRT [8], come abbiamo visto, è partizionata in tre termini i quali determinano il contributo relativo di un cambiamento nell'intercetta, nel coefficiente angolare o nella varianza, in caso di segnale di fuori controllo.

In questo approccio "change point" si applica la procedura della segmentazione binaria per stimare il punto di cambiamento e per determinare se ne abbiamo più di uno. Usando questa procedura, si ottengono i valori di lrt_{m_1} per tutti i possibili valori $m_1 = 1, 2, \dots, m - 1$ e si divide ognuno di questi per il valore atteso solo l'ipotesi nulla. Il metodo ci segnala la presenza di un change point se il massimo di queste statistiche eccede una data soglia. Il valore di m_1 che massimizza lrt_{m_1} è l'*MLE* del punto di cambiamento. Poi si divide il data set in due insieme distinti, prima e dopo di m_1 e si

applica la stessa procedura binaria ai due sub set. Questa procedura viene ripetuta fino a quando non si ha più evidenza di un cambiamento. L'approccio proposto permette non solo di segnalare che il processo non è più in controllo ma permette anche di stimare quando questo inizia il fuori controllo.

La statistica lrt_{m_1} può essere scritta anche sottoforma:

$$lrt_{m_1} = VAR_{\sigma^2} + VAR_{B_1} + VAR_{B_0}, \quad [9]$$

dove

$$VAR_{\sigma^2} = N \log\{(N_1 r^{2N_2/N} + N_2 r^{-2N_1/N})/N\},$$

$$VAR_{B_1} = N \log\{1 + (c_2 d_{B_1}^2 / c_1)\},$$

$$VAR_{B_0} = N \log\{1 + [(c_4 + c_3 d_{B_0}^2) / (c_1 + c_2 d_{B_1}^2)]\}.$$

Inoltre abbiamo che:

$r = \hat{\sigma}_1 / \hat{\sigma}_2$, $d_{B_0} = \bar{y}_2 - \bar{y}_1$, $d_{B_1} = (S_{xy2} / S_{xx2}) - (S_{xy1} / S_{xx1})$, $c_1 = N_1 \hat{\sigma}_1^2 + N_2 \hat{\sigma}_2^2$, $c_2 = S_{xx1} S_{xx2} / S_{xx}$, $c_3 = N_1 N_2 / N$, $c_4 = \{[N_1 N_2 (\bar{x}_2 - \bar{x}_1)^2 (S_{xx2} S_{xy1}^2 + S_{xx1} S_{xy2}^2) / N S_{xx} S_{xx1} S_{xx2}] - [N_1 N_2 (\bar{x}_2 - \bar{x}_1) (S_{xy1} + S_{xy2}) d_{B_0} / N S_{xx}] - [N_1^2 N_2^2 (\bar{x}_2 - \bar{x}_1)^2 d_{B_0}^2 / N^2 S_{xx}]\}$. L'espressione si semplifica considerevolmente se i valori di X sono gli stessi per ogni profilo con $c_4 = 0$.

Come anticipato, i tre fattori nella statistica [9], usata per indicare la presenza di un change point, determinano il contributo relativo del cambiamento nell'intercetta, nel coefficiente angolare.

In particolare, se il valore di VAR_{σ^2} è elevato si mette in evidenza l'instabilità in varianza. Visto che l'accuratezza dello stimatore dei coefficienti della regressione in controllo dipende dalla stabilità del processo in varianza, non è consigliabile applicare un test per verificare la stabilità dei coefficienti di regressione fino a quanto non si è sicuri che la varianza sia stabile. Inoltre un valore grande di VAR_{B_1} indica che le linee di regressione non sono parallele. Essendo plausibile dunque un cambiamento nel coefficiente angolare, non si andrà a verificare l'ipotesi di uguaglianza del parametro dell'intercetta.

L'idea di decomporre la statistica LRT in tre componenti è stata presa in considerazione da più autori come ad esempio Guilliksen e Wilks (1950) e Fatti e Hawkins (1986). I valori attesi in controllo per la statistica lrt_{m_1} non sono gli stessi per tutti i valori di m_1 . Per affrontare questo problema si propone di usare la prima correzione di Bartlett (Bartlett, 1937) che divide il test per un fattore normalizzato C_{m_1} . Se $C_{m_1} = E[lrt_{m_1}]$ le statistiche diventano:

$$lrtc_{m_1} = lrt_{m_1}/C_{m_1}, \quad m = 1, 2, \dots, m - 1,$$

tutte aventi gli stessi valori attesi.

4.2 Test basato sul rapporto di verosimiglianza generalizzato (GLR)

Si assuma che la relazione tra l'output Y e le variabili esplicative sia descrivibile dal seguente modello:

$$y_{ij} = g(x_{ij}) + \varepsilon_{ij}, \quad i = 1, 2, \dots, n_j \text{ e } j = 1, 2, \dots,$$

Si ipotizza che tale profilo abbia una varianza dell'errore stabile ma incognita, gli n_j abbiano tutti lo stesso valore n e che i punti, $\{x_{1j}, \dots, x_{nj}\}$ siano fissati per differenti j e indicati come $X = \{x_1, \dots, x_n\}$. L'assunzione della stabilità in varianza è solitamente valida per le applicazioni di calibrazione di industrie manifatturiere ed è consistente con la letteratura esistente sui profili lineari. Il modello "change-point" può essere espresso in questo caso come:

$$y_{ij} = \begin{cases} g(x_i) + \varepsilon_{ij}, & i = 1, \dots, n, \text{ if } 1 \leq j \leq \tau \\ g_1(x_i) + \varepsilon_{ij}, & i = 1, \dots, n, \text{ if } \tau \leq j \leq t \end{cases}$$

dove τ è un punto di cambiamento non noto, $g \neq g_1$ sono funzioni di regressioni non note in-controllo e fuori controllo rispettivamente e ε_{ij} sono i.i.d. errori con una ignota distribuzione F con media zero e varianza non nota σ^2 .

Per verificare la presenza di un possibile cambiamento in $\tau = k$, si utilizza un test di verosimiglianza generalizzato a due campioni (GLR) proposto da (Fan, Zhang, & Zhang, 2001) e (Zou, Qiu, & Hawkins, 2008) per l'ipotesi nulla sotto la quale g e g_1 sono uguali. In accordo con l'assunzione della normalità dell'errore, l'idea di base per derivare il test

GLR è quella di sostituire le ignote funzioni g e g_1 con le loro stime non parametriche costruite dai dati del profilo.

In particolare, le funzioni del test GLR in controllo, IC, e fuori controllo, OC, sono rispettivamente:

$$l_0 = -nt \ln(\sqrt{2\pi}\sigma) - \sum_{j=1}^t \left[\frac{1}{2\sigma^2} \sum_{i=1}^n (y_{ij} - \hat{g}^{(0)}(x_{ij}))^2 \right]$$

$$l_1 = -nt \ln(\sqrt{2\pi}\sigma) - \sum_{j=1}^k \left[\frac{1}{2\sigma^2} \sum_{i=1}^n (y_{ij} - \hat{g}^{(1)}(x_{ij}))^2 \right]$$

$$- \sum_{j=k+1}^t \left[\frac{1}{2\sigma^2} \sum_{i=1}^n (y_{ij} - \hat{g}_1(x_{ij}))^2 \right],$$

dove $\hat{g}^{(0)}$ identifica lo stimatore locale lineare del nucleo (LLKE) di g basato su t profili o campioni, e dove $\hat{g}^{(1)}$ e \hat{g}_1 stanno per l'LLKEs di g e g_1 basato sui primi k e rimanenti $t - k$ profili, rispettivamente. La statistica GLR al tempo t è definita dalla formula:

$$T_{h,k,t} = -2(l_0 - l_1) = \frac{k(t-k)}{t\sigma^2} (\bar{Y}_{0,k} - \bar{Y}_{k,t})^T V_h (\bar{Y}_{0,k} - \bar{Y}_{k,t})^T, \quad [10]$$

dove:

$$\bar{Y}_{i,m} = \frac{1}{m-i} \sum_{j=i+1}^m Y_j, \quad Y_j = (y_{1j}, \dots, y_{nj})^T,$$

$$V_h = W_h^T + W_h - W_h^{\otimes}, \quad W_h = (W_n(x_1), \dots, W_n(x_n))^T,$$

$$W_n(x_i) = (W_{n1}(x_i), \dots, W_{nm}(x_i))^T, \quad W_{nj}(x) = U_{nj}(x) / \sum_{i=1}^n U_{ni}(x),$$

$$U_{nj}(x) = K_h(x_j - x)[m_{n2}(x) - (x_j - x)m_{n1}(x)],$$

$$m_{nl}(x) = \frac{1}{n} \sum_{j=1}^n (x_j - x)^l K_h(x_j - x), \quad l = 1, 2,$$

$K_h(\cdot) = K(\cdot/h)/h$, K è il nucleo di una funzione simmetrica, h è una banda di lisciamento, e $A^{\otimes} = A^T A$. Dal momento che σ^2 è assunta ignota, lo sostituiamo con la sua stima non parametrica suggerita da Hall e Marron (1990)

$$\hat{\sigma}_t^2 = \frac{1}{t(n - df)} \sum_{j=1}^t (Y_j - W_h Y_j)^\otimes,$$

dove $df = \text{tr}(V_h)$.

Da un punto di vista asintotico $T_{h,k,t}$, data dalla [10], è simile a un test statistico basato sulla differenza di due stimatori della varianza costruiti da due data set, uno raccolto prima, l'altro raccolto dopo il punto di cambiamento. Come gli altri test basati sul parametro di lisciamo, anche in questo caso la scelta di h è rilevante e non così immediata in quanto è noto che un valore di h può essere ottimale per il lisciamo della curva, ma non per la potenza del test. Un test uniformemente più potente non esiste dovuto al fatto che la regressione non parametrica ha infinite dimensioni. Per il problema della non adeguatezza del test, Horowitz e Spokoiny (2001), suggeriscono la scelta di un h basato sul massimo di un test statistico dei momenti condizionato a una sequenza di parametri di lisciamo e dimostrano che il test risultante ha una certa ottimalità. Basandosi su questi risultati, Guerre e Lavergne (2005), suggeriscono di scegliere un h da una serie data di valori e dimostrano che il loro metodo è facile da usare e porta una buona performance in un'ampia varietà di casi.

In particolare si definisce il seguente insieme di possibili valori del parametro di lisciamo

$$H_n = \{h_j = h_{max} a^{-j} : h_j \geq h_{min}, j = 0, \dots, J_n\},$$

dove $0 < h_{min} < h_{max}$ sono i limiti superiori e inferiori, e $a > 1$ è un parametro. Il numero dei valori di H_n è $J_n < \log_a(h_{max}/h_{min})$. Seguendo Guerre e Lavergne (2005), una conveniente scelta di h è data da:

$$\tilde{h} = \arg \max_{h \in H_n} \{(T_{h,k,t} - \mu_h) - (T_{h_0,k,t} - \mu_{h_0}) - \gamma_n v_{h,h_0}\},$$

dove $\gamma_n > 0$ è un parametro di penalità, μ_h è la media di $T_{h,k,t}$ e v_{h,h_0}^2 è la varianza di $T_{h,k,t} - T_{h_0,k,t}$. Dopo aver scelto il giusto valore della banda di lisciamo, il test statistico diventa:

$$\tilde{T}_{k,t} = (T_{\tilde{h},k,t} - \mu_{\tilde{h}}) / v_{h_0}, \quad [11]$$

dove $v_{h_0}^2$ è la varianza di $T_{h_0,k,t}$, $\mu_{\tilde{h}} = \sum_{i=1}^n V_{\tilde{h}}^{(ii)}$, e $V_{\tilde{h}}^{(ij)}$ denota (i,j) -elemento della matrice $V_{\tilde{h}}$. Per un dato valore di h e h_0 , si può vedere che stimatori consistenti per $v_{\tilde{h}}^2$ e v_{h,h_0}^2 sono rispettivamente:

$$\hat{v}_h^2 = 2 \sum_{i=1}^n \sum_{j=1}^n [V_h^{(ij)}]^2, \quad \hat{v}_{h,h_0}^2 = 2 \sum_{i=1}^n \sum_{j=1}^n [V_h^{(ij)} - V_{h_0}^{(ij)}]^2.$$

Il criterio di selezione di Guerre e Lavergne (2005), ha ottime proprietà; prima quella che favorisce una statistica sotto H_0 , che garantisce che la distribuzione asintotica di $\tilde{T}_{k,t}$ sia la stessa di $(T_{h_0,k,t} - \mu_{h_0})/v_{h_0}$. Inoltre, sempre asintoticamente, il parametro di liscio selezionato non porta ad aumenti dell'errore di primo tipo. Una seconda proprietà è quella che questa procedura permette l'uso di v_{h_0} al posto di $v_{\tilde{h}}$ in $\tilde{T}_{k,t}$ che incrementa asintoticamente la potenza del test.

Nei casi reali il punto esatto di cambiamento, τ , è sconosciuto. Nell'ipotesi di un cambiamento nella funzione di regressione g al tempo t , consideriamo il test statistico GLR adattivo:

$$\tilde{T}_{\hat{\tau}} = \max_{1 < k < t} \tilde{T}_{k,t},$$

che è il massimo tra tutti i punti di shift possibili. Lo stimatore $\hat{\tau}$ è utilizzato al posto di τ .

Capitolo 5: Sorveglianza di profili lineari generali

Le metodologie di controllo che abbiamo visto fino ad ora si basano sul fatto che un semplice profilo lineare possa essere adeguatamente rappresentato da una linea retta. D'altra parte però come un profilo lineare può caratterizzare svariate applicazioni, profili lineari generali, che includono sia regressioni polinomiali che regressioni lineari multiple, forse potrebbero essere più adeguati per rappresentare la maggior parte delle applicazioni industriali. Zou, Wang e Tsung (2008) propongono una carta EWMA multivariata (*MEWMA*) per sorvegliare per mezzo di un'unica carta i cambiamenti dei parametri di regressione e la varianza del processo. A differenza degli approcci basati sull'applicazione di più carte per la sorveglianza dei profili, la complessità di questa carta singola multivariata non cresce con il crescere del numero dei parametri del profilo. La carta MEWMA può essere disegnata facilmente e ha una performance soddisfacente. Inoltre questo schema può essere potenziato con l'utilizzo di una schema di campionamento a intervalli variabili (VSI) che ne aumenta l'efficienza nel cogliere cambiamenti del profilo. Questo schema di controllo collegato a un test rapporto di verosimiglianza consente l'identificazione del punto di cambiamento e del parametro del profilo che ha subito un cambiamento.

5.1 Profilo lineare generale

Si assume che per il j^{th} campione raccolto, abbiamo osservazioni (X_i, Y_i) , dove Y_i è un vettore n_i -variato e X_i è una matrice $n_i \times p$ ($n_i > p$). Inoltre si assume che quando il processo è in controllo il modello sottostante è:

$$Y_i = X_i\beta + \varepsilon_i, \quad [12]$$

dove $\beta = (\beta^{(1)}, \beta^{(2)}, \dots, \beta^{(p)})$ è il vettore p -dimensionale dei coefficienti e gli ε_i sono realizzazioni i.i.d. da una normale multivariata con media 0 e varianza $\sigma^2 I$. Un'ulteriore assunzione riguarda le X_i che supponiamo essere $(1, X_i^*)$, dove le X_i^* sono ortogonali a 1 e $\mathbf{1}$ è un vettore n -variato di tutti 1. Diversamente possiamo ottenere questa forma attraverso adeguate trasformazioni. La numerosità n_j è uguale per tutti i campioni e la variabile esplicativa X_i si assume essere fissata per differenti i . Questo è praticamente lo

stesso caso utilizzato da Kang e Albin (2000), Kim et al. (2003) e Woodall(2004) dove l'interesse dello studio erano semplici regressioni lineari.

Quando si studia un profilo come quello descritto dal modello [12], i $p + 1$ parametri, ovvero i p coefficienti e la deviazione standard σ , devono essere controllati simultaneamente. Si definisca:

$$Z_i(\beta) = (\hat{\beta}_i - \beta)/\sigma, [13]$$

e inoltre

$$Z_i(\sigma) = \Phi^{-1}\{\chi^2((n - p)\hat{\sigma}_i^2/\sigma^2; n - p)\}, [14]$$

dove $\hat{\beta}_i = (X^T X)^{-1} X^T Y_i$, $\hat{\sigma}_i^2 = \frac{1}{(n-p)} (Y_i - X\hat{\beta}_i)^T (Y_i - X\hat{\beta}_i)$, $\Phi^{-1}(\cdot)$ è la funzione di ripartizione di una normale standard, e $\chi^2(\cdot, \nu)$ è una distribuzione chi quadro con ν gradi di libertà. La trasformazione [14] ha la proprietà che la distribuzione di $Z_i(\sigma)$ non dipende dalla grandezza del campione n , quando il processo è in controllo. Questo ci permetterà successivamente di occuparci anche del caso di intervalli di campionamento variabili. Inoltre anche la scelte dei limiti di controllo non è influenzata da n così che la carta di controllo può essere disegnata più facilmente. Un altro vantaggio portato dalla trasformazione di $\hat{\sigma}_i^2$ in $Z_i(\sigma)$ è che la distribuzione di $Z_i(\sigma)$ è simmetrica così che la carta di controllo possa essere maggiormente sensibile alla decrescita della deviazione standard. Quando il processo è in controllo il vettore $(p + 1)$ variato $Z_i = (Z_i^T(\beta), Z_i(\sigma))^T$ segue una distribuzione multinormale con media 0 e matrice di varianza e covarianza $\Sigma = \begin{pmatrix} (X^T X)^{-1} & 0 \\ 0 & 1 \end{pmatrix}$.

Lo schema per sorvegliare il profilo lineari generale [12] viene denominato *MEWMA* ed è definito dalla seguente statistica *EWMA*:

$$W_j = \lambda Z_j + (1 - \lambda)W_{j-1}, \quad j = 1, 2, \dots,$$

dove W_0 è un vettore iniziale $(p + 1)$ dimensionale e λ è un parametro di lisciamento ($0 < \lambda < 1$).

I segnali della carta sono dati da:

$$U_j = W_i^T \Sigma^{-1} W_i > L \frac{\lambda}{2-\lambda}, [15]$$

dove $L > 0$ è scelto per ottenere una precisa ARL in controllo.

5.2 Le carte MEWMA e EWMA₃ a confronto

Lo schema MEWMA (Zou et al. 2007) introdotto per lo studio di profili lineari generali può anche essere utilizzato per la sorveglianza di profili lineari semplici. In questa sezione quindi si è scelto di confrontare la carta MEWMA con il modello EWMA₃ qui denominato KMW, proposto da Kim et al. (2003) per il monitoraggio dei profili lineari.

Per semplicità e coerenza con la letteratura, si assume un ARL in controllo pari a 200. Il profilo di riferimento con parametri $A_0 = 3$, $A_1 = 2$, $\sigma^2 = 1$, e $x_i = 2, 4, 6, 8$ è quello proposto da Kang e Albin (2000) e già descritto nel paragrafo 3.3. Il disegno della corrispondente carta KMW, suggerito da Kim et al. (2003), consiste nel fissare i limiti di controllo pari a 3.0156, 3.0109 e 1,3723 per le tre carte EWMA con costante di lisciamiento $\lambda = 0.2$. Nel caso di parametri noti, l'ARL totale in controllo è pari a 200, mentre per ogni singola carta è pari a 584. La performance dei due schemi è valutata su 50000 simulazioni e approssimazioni di catene di Markov. Inoltre i tipi di shift considerati sono consistenti con quelli di Kim et al. (2003), quindi con cambiamenti nei parametri A_0 , A_1 , σ e B_1 .

Tabella 6: Confronto dell'ARL tra MEWMA e KMW per shift in A_0 , A_1 , deviazione standard e B_1

δ_1	A_0		A_1			σ			B_1		
	MEWMA	KMW	δ_1	MEWMA	KMW	δ_1	MEWMA	KMW	δ_1	MEWMA	KMW
.1000	131.5	133.7	.0250	99.0	101.6	1.1000	76.2	72.8	.0500	120.5	120.8
.2000	59.9	59.1	.0375	57.4	61.0	1.1500	48.7	48.1	.0750	77.3	77.3
.3000	29.6	28.3	.0500	35.0	36.5	1.2000	33.2	33.5	.1000	50.0	49.1
.4000	17.2	16.2	.0625	23.1	24.6	1.2500	24.1	24.9	.1500	24.0	22.8
.5000	11.5	10.7	.0750	16.4	17.0	1.3000	18.4	19.4	.2000	14.0	13.1
.6000	8.5	7.9	.1000	9.8	10.3	1.4000	12.1	12.7	.2500	9.5	8.9
.8000	5.8	5.1	.1250	6.9	7.2	1.6000	7.0	7.2	.3000	7.1	6.6
1.0000	4.1	3.8	.1500	5.3	5.5	1.8000	4.9	5.1	.4000	4.7	4.4
1.5000	2.6	2.4	.2000	3.7	3.8	2.2000	3.1	3.2	.5000	3.6	3.3
2.0000	2.0	1.9	.2500	2.9	2.9	2.6000	2.3	2.5	.7000	2.5	2.3
						3.0000	1.9	2.1	.9000	2.0	1.9

La Tabella 6 mostra che in caso di cambiamenti elevati per il coefficiente angolare A_1 le carte di controllo a confronto hanno un comportamento simile, mentre per cambiamenti più piccoli la carta MEWMA ha una performance leggermente migliore. Per quanto riguarda i cambiamenti nella deviazione standard, σ , il metodo di Zou et al.

(2007) sembra dare migliori risultati ma solo in caso di piccoli cambiamenti. Lo schema *KMW* infine presenta una performance leggermente migliore per rilevare moderati e grandi cambiamenti per l'intercetta A_0 e B_1 , ma la differenza è veramente trascurabile.

Tabella 7: ARL di una carta MEWMA per cogliere un decremento in varianza

δ_1	.10	.15	.20	.25	.30	.35	.40
ARL	3.3	3.9	4.5	5.3	6.4	7.8	9.7
δ_1	.45	.50	.55	.60	.65	.70	.75
ARL	12.5	16.5	22.9	33.0	49.1	74.9	114.5

Viene considerato nel confronto anche il caso di un cambiamento simultaneo nell'intercetta e inclinazione. In questo caso lo schema MEWMA ha una performance migliore nella maggior parte dei casi tranne quando uno dei due cambiamenti è molto grande e l'altro invece molto piccolo.

Tabella 8: Confronto dell'ARL tra MEWMA e KMW per uno shift combinato nell'intercetta, δ_1 , e coefficiente angolare, δ_2 .

MEWMA	δ_2										
	.025	.050	.075	.100	.125	.150	.175	.200	.225	.250	
δ_1	.05	155.8	111.0	72.9	48.0	32.8	23.5	17.7	13.9	11.3	9.5
		157.6	114.7	74.8	48.3	32.2	22.5	16.9	13.2	10.7	8.9
	.10	118.0	89.2	62.1	42.8	30.2	22.2	16.9	13.5	11.0	9.3
		122.1	94.6	66.4	44.9	30.7	21.9	16.6	13.1	10.6	8.9
	.15	82.2	66.3	49.5	36.1	26.7	20.2	15.8	12.8	10.6	9.0
		84.6	70.8	54.5	39.6	28.5	20.9	16.1	12.8	10.4	8.8
	.20	56.4	48.0	38.2	29.6	22.9	18.1	14.5	12.0	10.1	8.7
		57.1	51.1	42.4	33.3	25.4	19.5	15.4	12.4	10.2	8.7
	.25	39.5	35.0	29.4	24.0	19.5	15.9	13.2	11.1	9.5	8.2
		39.5	36.5	32.3	27.1	22.0	17.8	14.4	11.9	10.0	8.5
	.30	28.7	26.2	22.9	19.6	16.5	13.9	11.8	10.2	8.8	7.8
		28.2	26.9	24.7	22.0	18.8	15.7	13.2	11.2	9.6	8.3
	.35	21.7	20.2	18.3	16.1	14.0	12.2	10.6	9.3	8.2	7.3
		20.9	20.2	19.1	17.6	15.8	13.9	12.1	10.5	9.1	8.0
	.40	17.0	16.1	14.9	13.5	12.0	10.7	9.5	8.5	7.6	6.9
		16.2	15.9	15.3	14.5	13.5	12.1	10.9	9.7	8.6	7.6
	.45	13.7	13.2	12.4	11.4	10.5	9.5	8.6	7.8	7.1	6.5
		13.1	12.9	12.6	12.1	11.4	10.6	9.8	8.9	8.0	7.3
	.50	11.4	11.1	10.5	9.9	9.2	8.5	7.8	7.2	6.6	6.1
		10.8	10.8	10.6	10.3	9.9	9.3	8.7	8.1	7.5	6.9

Possiamo concludere quindi che per cogliere cambiamenti nei profili lineari, lo schema *MEWMA* ha almeno la stessa performance dello schema combinato *KMW* di Kim et al. (2003).

5.3 Carta di controllo MEWMA con intervalli di campionamento variabili

Un miglioramento del modello *MEWMA* si può ottenere con l'applicazione di un intervallo di campionamento variabile (VSI). Lo schema VSI è un approccio atto a migliorare l'efficienza degli schemi di sorveglianza. Tuttavia l'applicazione di questo a carte di controllo multiple per i profili, come lo schema *KMW* non è così semplice. Questo perché i tre limiti di controllo e i tre limiti di preallarme devono essere scelti contemporaneamente così che tutte le carte distintamente diano lo stesso tasso medio di campionamento in controllo e lo stesso tasso medio di falsi allarmi.

I convenzionali schemi di campionamento per la sorveglianza dei processi adottano un tasso di campionamento fisso (FSR): campioni di numerosità fissa (FSS) vengano cioè estratti a intervalli di campionamento fissi (FSI). Negli ultimi tempi molte modifiche sono state proposte per ottenere un'individuazione più veloce di cambiamenti nei parametri del processo. Tra i metodi proposti proprio lo schema VSI è quello più popolare e efficiente per incrementare l'efficienza nella rilevazione. In un carta di controllo VSI, gli intervalli di campionamento variano in funzione della statistica di controllo. L'idea di base è di usare un intervallo di tempo più corto in presenza di un possibile cambiamento del processo, mentre di allungare i tempi tra campionamenti nel caso non si ipotizzino variazioni. Studi sul tema mostrano che usare solo due possibili valori per gli intervalli di campionamento è più che sufficiente, consideriamo quindi due possibili valori: $0 < d_1 \leq d_2$.

Nell'applicare lo schema VSI alla carta *MEWMA*, Zou, Wang et Tsung (2007), aggiungono un limite di allarme all'interno dei classici limiti di controllo $0 \leq L_1 < L$ che determina quale intervallo di campionamento bisogna utilizzare successivamente. In particolare un lasso di tempo maggiore viene usato se dopo l'osservazione, i , la statistica U_i , non supera il limite di allarme $L_1 \frac{\lambda}{2-\lambda}$. D'altra parte invece si fa uso di un intervallo più corto se U_i supera tale limite ma non supera $L \frac{\lambda}{2-\lambda}$. Se invece la statistica U_i cade completamente al di fuori di quest ultimo allora si ha un segnale di fuori controllo come

nel caso tradizionale di una carta *MEWMA*. Quando si valuta la performance di una carta di controllo con schema *VSI*, bisogna prendere in considerazione sia il tempo medio di segnale, *ATS*, sia il numero medio di campioni prima di avere il segnale, *ANSS*, perché l'*ATS* non è una semplice costante moltiplicativa di *ANSS* come nel caso *FSI*.

Uno studio comparativo con le carte *MEWMA* classiche dimostra l'aumento in termini di efficienza dello schema *VSI – MEWMA*. Il confronto richiede che tutte le carte comparate abbiano lo stesso tasso di campionamento in controllo e stesso tasso di falsi allarmi. Questo assicura che le carte abbiano identico *ATS* e *ANSS* quando il processo è in controllo. Il paragone viene fatto in termini di "steady-state" *ATS* (*SSATS*), definito come il tempo che dobbiamo aspettare dal momento in cui c'è stato un cambiamento al momento in cui la carta ci segnala il fuori controllo, sotto l'assunzione che la statistica di controllo, verificatosi il cambiamento, abbia raggiunto la stazionarietà in distribuzione.

Tabella 9: Confronto di *SSATS* per una *FSI MEWMA* e una *VSI MEWMA* per uno shift in intercetta e deviazione standard

A_0					σ				
δ_1	FSI	VSI			δ_1	FSI	VSI		
		$d_1 = .5$ $d_2 = 1.25$	$d_1 = .25$ $d_2 = 1.5$	$d_1 = .1$ $d_2 = 1.9$			$d_1 = .5$ $d_2 = 1.25$	$d_1 = .25$ $d_2 = 1.5$	$d_1 = .1$ $d_2 = 1.9$
.1	127.9	124.4	122.2	120.0	.1	2.7	2.1	1.9	1.8
.2	57.6	51.9	48.3	45.2	.3	5.8	4.3	3.6	3.3
.4	28.1	23.3	20.4	18.1	.5	15.9	11.2	8.7	7.1
.4	16.1	12.6	10.6	9.2	.7	73.8	63.9	57.6	51.1
.5	10.6	8.0	6.6	5.8	1.1	73.2	68.9	66.3	63.9
.6	7.6	5.7	4.8	4.2	1.2	31.2	27.4	25.2	23.4
.8	4.8	3.6	3.1	2.8	1.4	16.9	14.1	12.5	11.4
1.0	3.4	2.6	2.3	2.1	1.8	10.8	8.8	7.7	6.9
1.5	2.0	1.6	1.4	1.4	2.2	4.0	3.2	2.7	2.6
2.0	1.4	1.1	1.0	1.1	2.6	2.4	1.9	1.7	1.7
3.0	.8	.7	.8	.9	3.0	1.3	1.1	1.1	1.2

La Tabella 9 mostra i valori di *SSATS* per la carta *VSI* e *FSI-EWMA* per profili lineari per cui si studia un cambiamento nell'intercetta e nella deviazione standard. L'*ATS* e l'*ANSS* per entrambe le carte è posto pari a 200, così il tasso di campionamento medio in controllo per carte *VSI* è di 1 campione per unità di tempo. Possiamo concludere che aggiungere lo schema *VSI* porta a una sostanziale riduzione nel tempo richiesto per individuare piccoli o moderati cambiamenti. In generale l'intervallo d_1 dovrebbe essere il più piccolo possibile per una migliore performance, questo però dipende anche dai limiti fisici del processo, quindi dalla possibilità di raccogliere dati dopo un breve lasso di tempo. Per

quanto riguarda d_2 invece l'intervallo dovrebbe essere grande in modo che la carta di controllo abbia un'accettabile tasso medio di campionamento.

5.4 Diagnostica dei profili lineari generali.

Nella pratica del controllo qualità, insieme con una veloce identificazione del cambiamento del processo, un altro aspetto critico è quello di diagnosticare quale/i parametro/i abbia prodotto un segnale di fuori controllo. Una metodologia di indagine volta a localizzare il punto di cambiamento nel processo e a isolare quale parametro sia cambiato nel profilo, aiuta gli operatori ad identificare ed eliminare le cause di un segnale.

Per stabilire in quale punto vi sia stato il cambiamento nel monitoraggio dei profili viene usato uno stimatore di massima verosimiglianza della statistica "change point". Si assume che un segnale di fuori controllo sia stato segnalato nel sottogruppo k di una carta di controllo *MEWMA*. Si propone quale stimatore del punto di cambiamento, change point, τ :

$$\tau = \arg \max_{0 \leq t < k} \{lr(tn, kn)\},$$

dove $lr(tn, kn)$ è una statistica GLR.

Una volta identificato il punto in cui si sia verificato il cambiamento, è ugualmente critico cogliere quale parametro del profilo sia cambiato. Le carte di controllo costruite da Kim et al. (2003), sono di più facile interpretazione rispetto alla carta di Kang e Albin (2000), in quanto la prima è costruita con la combinazione di tre carte *EWMA* e una volta ricevuto un segnale di fuori controllo si può andare a controllare quale sia il problema nella singola carta. Questo preclude però la verifica del caso vi sia un cambiamento in due parametri contemporaneamente. A primo impatto sembrerebbe che per la carta *MEWMA* sia impossibile identificare quale parametro sia cambiato. In realtà è possibile farlo grazie a un test parametrico. In particolare:

$$\tilde{\beta}_{t,k} = \frac{1}{k-t} (X^T X)^{-1} X^T \sum_{j=t+1}^k Y_j,$$

e inoltre

$$\tilde{\sigma}_{t,k}^2 = \frac{1}{(k-t)n-p} \sum_{j=t+1}^k (Y_j - X\tilde{\beta}_{t,k})^T (Y_j - X\tilde{\beta}_{t,k}).$$

In pratica si suppone che la carta *MEWMA* abbia segnalato un allarme al k –esimo campione e che si sia già ottenuto la stima del punto di cambiamento. Si è dunque interessati a capire quale o quali parametri abbiano subito una variazione. I test per l'intercetta, la deviazione standard e per i coefficienti di regressione sono i seguenti. Un test t per il cambiamento nell'intercetta con $(k - \hat{t})n - p$ gradi di libertà. La statistica test è:

$$T_{TEST} = \frac{\sqrt{(k - \hat{t})n}(\beta_{\hat{t},k}^{(1)} - \beta^{(1)})}{\tilde{\sigma}_{\hat{t},k}},$$

dove $\beta_{\hat{t},k}^{(1)}$ denota il primo componente di un vettore p dimensionale, $\tilde{\beta}_{\hat{t},k}$. Il test usato per la deviazione standard è un chi-quadrato con $(k - \hat{t})n - p$ gradi di libertà e la statistica test è:

$$\chi_{TEST}^2 = \frac{[(k - \hat{t})n - p]\tilde{\sigma}_{\hat{t},k}^2}{\sigma^2}.$$

Per il resto dei $p - 1$ parametri $(\beta^{(2)}, \dots, \beta^{(p)})$, considerando la correlazione tra gli stimatori, si propone un test con regione di rifiuto:

$$F_{TEST}^{(i)} = \frac{(k - \hat{t})(\beta_{\hat{t},k}^{(i)} - \beta^{(i)})^2}{m_{ii}\tilde{\sigma}_{\hat{t},k}^2} > F_{\alpha}(p - 1, (k - \hat{t})n - p, R),$$

per ogni $i = 2, \dots, p$, dove m_{ii} sono gli elementi della diagonale di $M = (X^T X)^{-1}$ e $R = \text{diag}\{m_{11}^{-1/2}, \dots, m_{pp}^{-1/2}\} M \text{diag}\{m_{11}^{-1/2}, \dots, m_{pp}^{-1/2}\}$ è la matrice di correlazione di $\hat{\beta}$ e $F_{\alpha}(p - 1, (k - \hat{t})n - p, R)$ è il percentile superiore, α , di una distribuzione F multivariata con parametri $(p - 1, (k - \hat{t})n - p, R)$. Come si può notare, i test statistici proposti non seguono realmente la distribuzione nulla data, cioè $(t, F, \text{chi. quadro})$, quando il processo è in controllo, poiché sono ottenuti sotto la condizione che la carta *MEWMA* abbia segnalato un fuori controllo e sia stata ottenuta una stima del punto di cambiamento. Questo fatto fondamentale non dovrebbe cambiare nulla nella diagnostica, ma potrebbe ripercuotersi sul errore di I tipo. In aggiunta, l'efficienza di

ognuno di questi test dipende dall'accuratezza della stima del punto di cambiamento. Infine, quando p è grande, questo $p + 1$ test non permette un'accurata decisione se si vuole tenere sotto controllo l'errore di I tipo.

D'altra parte questi metodi di diagnostica si comportano bene in molti casi e sono in conclusione un valido aiuto agli operatori per prendere decisioni sui profili in esame.

Capitolo 6: Approcci non parametrici alla sorveglianza dei profili lineari generali

Si assuma che per il j^{th} profilo, si disponga delle osservazioni (X_j, Y_j) , dove Y_j è un vettore n -variato di risposta e X_j è una matrice di regressori $n_j \times p$. Si assume inoltre che quando il processo è in controllo il modello sottostante è:

$$Y_{ij} = g(X_j^{(i)}, \beta) + \varepsilon_{ij}, \quad i = 1, 2, \dots, n_j \text{ e } j = 1, 2, \dots, [16]$$

dove $X_j^{(i)}$ denota l' i^{th} riga di X_j , β è un vettore q -dimensionale di parametri e gli ε_{ij} sono delle variabili casuali i.i.d. con distribuzione normale di media 0 e varianza σ^2 .

La funzione g può essere sia lineare sia non lineare. Nel caso di profili non lineari, gli schemi di controllo parametrici, quali la carta di controllo KMW (Kim et al., 2003) e la MEWMA (Zou et al., 2007), possono essere comunque applicati ottenendo gli stimatori ai minimi quadrati per i parametri di ogni singolo profilo. Si noti che, diversamente dalla regressione lineare, l'esatta distribuzione di tali stimatori non è ottenibile. Le carte parametriche possono essere utilizzate per la sorveglianza del modello [16] visto che i parametri stimati solitamente hanno una distribuzione asintoticamente normale. Gli studi di simulazione indicano però che schemi di controllo approssimati non soltanto peggiorano le proprietà dell'ARL in controllo, ma hanno inoltre un effetto rilevante sulla performance fuori controllo. Un altro problema inerente al monitoraggio del modello [15] con metodi parametrici è che questi si basano sull'assunto di base che i modelli IC e OC abbiano la stessa forma, ma che cambino solamente i parametri. Benché il modello IC possa essere determinato prima che il monitoraggio abbia inizio, quello OC non può essere specificato facilmente, specialmente quando la curva in controllo ha forme complicate come nel caso di molti modelli non lineari. I modelli parametrici di solito sono dunque più performanti quando vengono disegnati su un dato modello OC, ma presentano qualche problema quando è difficile specificazione il modello del profilo fuori controllo.

I modelli non parametrici quindi si pongono come obiettivo quello di risolvere il problema dello studio dei profili non lineari e di superare l'incapacità dei metodi parametrici nell'identificare certi tipi di cambiamento a causa di un'errata specificazione del profilo fuori controllo.

6.1 Modello non parametrico

Per facilità di presentazione si presenta il modello [16] nel caso unidimensionale:

$$y_{ij} = g(x_{ij}) + \varepsilon_{ij}, \quad i = 1, 2, \dots, n_j \text{ e } j = 1, 2, \dots, \quad [17]$$

dove x_{ij} denota il valore della regressione per l' i^{th} osservazione per il j^{th} profilo. Si noti che in questo modello abbiamo tolto il parametro β perché poniamo la nostra attenzione sul metodo non parametrico. La numerosità si assume la stessa per ogni campione e la variabile esplicativa $x_{1j}, x_{2j}, \dots, x_{n_j}$, fissata per differenti j , con $j=1, 2, \dots$. Nei contesti di regressione non parametrica questo contesto viene chiamato "*common fixed design*". In particolare possiamo inoltre assumere che $x_1 \leq x_2 \leq \dots \leq x_n$ e che le x_i variano in un intervallo tra $[0, 1]$.

Si consideri il caso di Fase 2 nel quale la funzione di regressione in controllo e la varianza, dette g_0 e σ_0 , sono assunte note o che il data set di Fase 1 sia sufficiente per stimare al meglio la funzione di regressione. L'obiettivo durante la Fase 2 è quello di identificare qualsiasi cambiamento nella funzione del modello e nella varianza del profilo. Un cambiamento nella funzione di regressione indica l'esistenza di una dilatazione o di una contrazione del profilo. Un incremento della varianza è dovuto a una poca accuratezza nella lavorazione del profilo invece un decremento ci indica un possibile miglioramento del processo.

6.2 Lo schema di controllo NEWMA

Dato il data set composto da j profili, le osservazioni sono $\{x_i, y_i\}_{i=1}^n$. Sia

$$Z_j = (Y_j - G_0)/\sigma_0$$

dove

$$Y_j = (y_{j1}, y_{j2}, \dots, y_{jn})^T.$$

Dato che sia la regressione che la varianza devono essere sorvegliate è necessario un test non parametrico per σ_j . Assumendo soddisfatte le condizioni seguenti:

- esista una densità positiva, $f(u)$, che sia Lipschitz continua, limitata inferiormente in 0 e i punti di controllo, x_i , soddisfino:

$$\int_0^{x_i} f(u) du = \frac{i}{n}, \quad i = 1, \dots, n.$$

- $g_0(u)$ abbia una derivata seconda continua;
- la funzione $K(t)$ sia simmetrica e limitata;
- La banda di lisciamento h soddisfi le condizioni $h \rightarrow 0$ e $nh^{3/2} \rightarrow \infty$ e $nh^8 \rightarrow 0$.

Sotto queste condizioni, senza tener conto se la funzione di regressione sia cambiata o meno, si ha:

$$\hat{\sigma}_j^2 = \frac{1}{n} (Z - WZ)^\otimes = \sigma_j^2 + O_p\left(n^{-\frac{1}{2}}\right) + O_p((nh)^{-1}),$$

Dal momento che Z_j è un vettore con distribuzione normale standard multivariata, $n\hat{\sigma}_j^2 = (Z - WZ)^\otimes = Z^T(I - V)Z$ è una forma quadratica di un vettore casuale normale. La sua distribuzione è equivalente alla distribuzione di una combinazione lineare di χ_1^2 variabili indipendenti con coefficienti dati dagli auto valori di $(I - V)$.

La statistica $(Z - WZ)^\otimes$ può utilizzarsi per condurre un test su un possibile cambiamento in varianza, dove un valore grande indicherebbe un incremento nella varianza del processo. A partire da questo test per la varianza si può implementare un carta di tipo EWMA, più sensibile ai cambiamenti di piccola e media entità rispetto a uno schema Shewhart.

Inizialmente si trasforma $\hat{\sigma}_j^2$ in una variabile normale:

$$\tilde{\sigma}_j = \Phi^{-1}\{\psi(n\hat{\sigma}_j^2; I - V)\}, \quad [18]$$

dove $\Phi^{-1}(\cdot)$ è la funzione di ripartizione di una normale standard e $\psi(\cdot; A)$ è la funzione di ripartizione di una variabile casuale $n\hat{\sigma}_j^2$ quando il processo è in controllo. Possiamo

notare che la trasformazione [16] della varianza, con una successiva applicazione di una carta *EWMA*, è stata proposta da Quesenberry (2005) e da Chen, Cheng e Xie (2001). Il vantaggio di questa trasformazione è che la distribuzione di $\tilde{\sigma}_j$ è simmetrica e quindi è possibile introdurre una carta di controllo che, pur non essendo di semplice costruzione, può rilevare cambiamenti della varianza sia positivi che negativi. A tal fine è necessario ricavare $\psi(\cdot; A)$. Qui si introduce il metodo che associa i primi tre momenti della distribuzione di $n\hat{\sigma}_j^2$ con quelli di una distribuzione chi quadrato che meglio approssima la forma quadratica. Successivamente indicando con $U_j = (Z_j^T, \tilde{\sigma}_j)^T$ un vettore casuale $(n + 1)$ -variato e $\Sigma = \begin{pmatrix} V & 0 \\ 0 & 1 \end{pmatrix}$, una matrice simmetrica $(n + 1)$ -dimensionale, la statistica di controllo *EWMA* può così essere definita come:

$$E_j = \lambda U_j + (1 - \lambda)E_{j-1}, \quad j = 1, 2, \dots,$$

dove E_0 è un vettore iniziale $(n + 1)$ -dimensionale e λ è costante di lisciamento definita nell'intervallo $[0,1]$. La carta lancia un segnale quando:

$$Q_j = E_j^T \Sigma E_j > L \frac{\lambda}{2-\lambda}, \quad [19]$$

dove $L > 0$ è scelto per dare una specifica ARL in controllo. Lo schema [19] così presentato prende il nome di *NEWMA* che è analogo allo schema *MEWMA* [15] presentato da Zou et al. (2007) per sorvegliare profili lineari. Un punto di discussione aperto è che Z_j e $\tilde{\sigma}_j$ solitamente sono dipendenti, ma di questo non si tiene conto nella matrice Σ . Ponendo Z_j e $\tilde{\sigma}_j$ insieme, il loro risultato U_j ha una distribuzione congiunta incognita con distribuzione marginale normale. Ci si aspetta dunque che lo schema proposto rimanga valido anche ignorando questa correlazione poiché questo non influenza in maniera sostanziale il test sulla regressione e sulla varianza.

6.3 Diagnostica nei profili

Come fatto in precedenza discutiamo l'analisi di un profilo lineare generale e i passi necessari per una diagnosi del processo.

Il primo passo è quello di identificare il punto di cambiamento. E' necessario infatti separare i profili fuori controllo dall'intero data set di profili osservati così da permettere un'accurata inferenza sui vari tipi di cambiamenti. Per fare questo usiamo uno stimatore

basato sul test rapporto di verosimiglianza generalizzato (GLR). Si assuma che la carta *NEWMA* segnali un fuori controllo per il profilo k . Lo stimatore del punto di cambiamento, τ , è dato da:

$$\hat{\tau} = \arg \max_{0 \leq t < k} \{lr(tn, kn)\},$$

dove $lr(tn, kn)$ è il test GLR. Questa formulazione è simile a quella di Zou et al (2007), vista nei capitoli precedenti, però ma qui viene usato un metodo di lisciamento non parametrico al posto di uno stimatore di massima verosimiglianza per massimizzare il test GLR.

Una volta stimato il punto di cambiamento il secondo passo consiste nello stabilire se la varianza è stabile questo perché, una volta che vi è stato un cambiamento, non si può ottenere una accurata inferenza sulla regressione. Per far questo si suggerisce quindi di verificare se il cambiamento si è verificato nei $(k - \hat{\tau})$ OC profili. Usando una estensione dello stimatore della varianza, si può applicare un test statistico con la seguente regione di accettazione:

$$\psi_1^{-1}\left(\frac{\alpha}{2}; I - V, k - \hat{\tau}\right) < \sum_{j=\hat{\tau}+1}^k Z_j^T (I - V) Z_j < \psi_1^{-1}\left(1 - \frac{\alpha}{2}; I - V, k - \hat{\tau}\right), [20]$$

dove $\psi_1(\cdot; A, l)$ è la funzione di ripartizione della somma di l variabili indipendenti $Z^T A Z$ e $\psi_1^{-1}(\alpha; A, l)$ è il percentile α di una distribuzione $\psi_1(\cdot; A, l)$ che può essere approssimata da $\psi_1(\cdot; A)$ usando le proprietà additive di un variabile casuale chi-quadro. Se l'ipotesi nulla della stabilità in varianza contro l'ipotesi alternativa di un possibile cambiamento nella varianza del processo è rifiutata, allora un incremento della varianza corrisponde a una maggiore inaccuratezza del processo o a un incremento dell'errore di misura, mentre un riduzione della varianza indica un miglioramento del processo produttivo.

Il terzo passo è dunque volto a determinare se la funzione di regressione è cambiata nei parametri. Se la stabilità della varianza è verificata, si deve determinare quali parametri della regressione siano cambiati. Possiamo adottare ancora il test GLR con un semplice trasformazione del test [18] dato da:

$$\frac{1}{(k - \hat{\tau})} \left(\sum_{j=\hat{\tau}+1}^k Z_j \right)^T V \left(\sum_{j=\hat{\tau}+1}^k Z_j \right) > \psi^{-1}(1 - \alpha; V),$$

dove ancora $\psi_1^{-1}(\alpha; A)$ è il percentile α di una distribuzione di forma quadratica $Z^T AZ$. Se si accetta l'ipotesi nulla, che nessun cambiamento si sia verificato nella funzione di regressione, possiamo concludere che la carta NEWMA ha dato un segnale di falso allarme in quanto tutti i parametri verificati sono in controllo mentre la nostra analisi era iniziata a seguito del segnale di fuori controllo della carta NEWMA. Se si accetta l'ipotesi alternativa riscontrando un effettivo cambiamento nei parametri del processo possiamo proseguire con il passaggio successivo.

Quest'ultimo passaggio è di aiuto per vedere quale parte della curva di regressione è cambiata, qual è la funzione di regressione OC e quanto grande è la differenza dei modelli in controllo e fuori controllo. Per ottenere risposte di facile interpretazione a questi problemi si utilizzano dei metodi grafici integrati dalle conoscenze a priori sul processo. In particolare, si suggerisce di disegnare la curva non parametrica lisciata della media dei $(k - \hat{\tau})$ profili, $\frac{1}{(k - \hat{\tau})} \sum_{j=\hat{\tau}+1}^k WY_j$, insieme ai profili in controllo.

Non è così importante che i test statistici per il secondo e terzo step seguano la vera distribuzione nulla data quando il processo è in controllo. Tali test sono infatti ottenuti sotto la condizione che la carta *NEWMA* abbia segnalato un fuori controllo e che il punto di cambiamento sia stato individuato. Possiamo quindi affermare che questo aspetto può influire sull'errore di I tipo e che l'efficacia dipende dall'accuratezza della stima del punto di cambiamento. In alcuni casi, la differenza tra i modelli IC e OC può fornire informazioni fuorvianti sulla variazione dei valori dei parametri di regressione ma non sulla variazione della forma del modello in sé. Questi casi accadono in alcuni modelli semplici quali le regressioni lineari più semplici e le regressioni quadratiche. Per la diagnosi dei fuori controllo, si usa la funzione con i parametri stimati per sostituire $g_0(x)$ in Z e si applica ancora il test GLR. Se l'ipotesi nulla dell'assenza di qualsiasi variazione nella forma del modello è accettata, si possono usare i test parametrici di Zou et al (2007) per identificare quali parametri sono cambiati.

6.4 Comparazione della performance

Per valutare la performance della carta *NEWMA*, questa viene confrontata con due approcci differenti. Il primo che viene denominato come NM, tratta Z_j come un vettore multivariato e usa lo schema *MEWMA* per costruire la carta di controllo. Nonostante la sua semplicità e convenienza, questo approccio ha due seri inconvenienti: ignora completamente la continuità delle risposte per i vicini punti covariati, mentre per n grande accumula rumore stocastico con una conseguente perdita di efficienza. Il secondo approccio denominato PM, è lo schema parametrico proposto da Zou et al. (2007), che ottiene la stime dei parametri per ogni profilo e poi applica lo schema *MEWMA* alla trasformazione dei parametri stimati.

Di seguito illustriamo i confronti tra le tre carte, NM, PM e *NEWMA*. Per semplicità e consistenza con la letteratura poniamo $\tau = 0$ e prendiamo in esame solo il caso con una ARL in controllo pari a 200. Inoltre si restringe lo studio a un disegno di punti equidistanti:

$$x_i = i - \frac{0.5}{n}, \quad i = 1, \dots, n.$$

Per la carta *NEWMA* la banda, h , è data dalla formula empirica:

$$h_E = c \times \left(\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x})^2 \right)^{1/2} n^{-1/5}, [21]$$

dove $c = 1.0, 1.5, e 2.0$. Tutti i risultati in termini di ARL sono valutati su 10000 simulazioni. Per illustrare l'efficienza dello schema *NEWMA*, la simulazione si basa su due modelli IC: un modello quadratico e uno non lineare. Si noti che il modello quadratico di regressione cade all'interno della classe dei modelli lineari. Per ogni modello si considera $n = 20, e 40$. Il numero e la varietà di possibili modelli OC e la combinazione dei parametri sarebbe troppo grande per ottenere un confronto completo. Visto l'obiettivo di dimostrare la robustezza della carta *NEWMA* si è deciso quindi di mostrare solo alcuni modelli OC più rappresentativi.

Il modello IC di forma quadratica è il seguente:

$$y_{ij} = 1 + 2x_i + 3x_i^2 + \varepsilon_{ij}, \quad i = 1, \dots, n,$$

dove $\sigma_0^2 = 1$.

Tre modelli OC sono scelti come segue:

$$(I): y_{ij} = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_{ij},$$

$$(II): y_{ij} = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \varepsilon_{ij},$$

$$(III): y_{ij} = 1 + 2x_i + 3x_i^2 + \beta \sin(2\pi\beta x_i) + \varepsilon_{ij}.$$

Questi tre modelli OC corrispondono ai seguenti scenari: (I) la struttura della regressione non cambia e solamente i parametri subiscono una variazione; (II) la struttura viene trasformata in un nuovo modello lineare; (III) la regressione lineare cambia in un modello non lineare.

Alcuni parametri dei tre modelli descritti sono riportati nella Tabella 10.

Tabella 10: Parametri dei tre modelli OC per lo Scenario 1

	Model (I)				Model (II)					Model (III)		
	β_0	β_1	β_2	σ	β_0	β_1	β_2	β_3	σ	β_1	β_2	σ
(i)	1.0	2.0	3.1	1.0	.8	4.4	-3.0	4.0	1.0	.1	1.0	1.0
(ii)	1.0	2.1	3.1	1.0	.8	4.4	-3.0	4.1	1.0	.2	1.0	1.0
(iii)	1.1	2.1	3.1	1.0	.8	4.4	-3.2	4.1	1.0	.2	.8	1.0
(iv)	1.0	2.0	3.0	1.1	1.0	4.4	-3.2	4.1	1.0	.2	1.3	1.0
(v)	1.0	2.0	3.0	.7	.8	4.4	-3.0	4.0	1.1	.3	1.5	1.0
(vi)	1.1	2.1	3.1	1.1	.8	4.5	-3.0	4.0	1.1	.3	1.5	1.1

Tabella 11: Confronto di ARL tre NEWMA, NM e PM per lo Scenario 1

		$n = 20$					$n = 40$				
		NEWMA			NM	PM	NEWMA			NM	PM
		$c = 1.0$	$c = 1.5$	$c = 2.0$			$c = 1.0$	$c = 1.5$	$c = 2.0$		
(I)	(i)	150.1	144.7	140.0	171.3	137.1	120.6	115.3	110.1	162.3	103.1
	(ii)	66.6	60.8	56.7	98.2	54.2	37.9	34.0	31.8	79.4	29.2
	(iii)	20.7	18.3	17.5	33.1	16.5	11.2	10.3	9.8	22.4	9.1
	(iv)	30.5	27.4	27.0	43.7	26.5	17.5	16.2	15.5	30.5	14.8
	(v)	8.2	6.6	5.9	*	5.3	4.2	3.7	3.5	*	3.2
	(vi)	12.5	11.4	11.0	17.3	10.8	7.4	7.0	6.7	11.8	6.3
(II)	(i)	104.6	104.2	113.4	132.2	199.7	66.5	64.0	67.5	114.2	200.1
	(ii)	89.4	87.5	92.5	119.2	154.1	52.5	50.3	52.1	99.3	120.2
	(iii)	78.8	75.7	79.2	109.3	121.0	45.0	42.4	42.8	89.3	84.9
	(iv)	24.2	22.1	21.1	38.9	22.8	12.7	11.6	11.2	26.2	12.2
	(v)	24.5	23.7	22.6	35.8	25.7	13.9	13.0	12.8	24.4	14.0
	(vi)	22.4	20.6	20.3	32.3	22.5	12.6	12.0	11.6	21.7	12.6
(III)	(i)	106.8	103.1	102.3	137.2	119.6	69.9	66.7	66.2	122.0	81.3
	(ii)	37.8	35.6	34.8	60.1	47.1	20.0	18.4	18.0	43.1	25.0
	(iii)	35.8	32.2	31.3	57.7	31.5	19.0	16.9	16.5	40.6	16.3
	(iv)	37.3	36.2	38.2	58.1	51.8	19.5	18.3	18.6	41.7	27.2
	(v)	18.0	18.3	20.4	29.5	27.4	9.7	9.5	10.0	18.3	15.6
	(vi)	11.2	11.1	11.4	15.5	13.4	6.8	6.6	6.6	10.6	7.7
L		20.25	17.25	15.63	38.69	13.87	21.66	18.28	16.50	65.20	13.87

I limiti di controllo sono riportati nell'ultima riga della Tabella 11.

Sotto il modello OC (I) la carta PM ha una efficienza superiore, come ci si poteva aspettare, vista la coerenza tra il modello OC e IC entrambi di forma quadratica. La carta *NEWMA* offre comunque una performance soddisfacente e la differenza con la carta PM non è così rilevante, anche quando usiamo un valore relativamente piccolo per la banda ($c = 1.0$). I modelli basati sul metodo smooth, *NEWMA* e PM, hanno una migliore resa della carta NM che non si rifà a questa metodologia.

Nei casi (i)-(iii) dei modelli (II) la carta PM è di gran lunga peggiore della carta *NEWMA* e seppur di poco anche della carta NM. Nello specifico nel caso (i) del modello (II) l'ARL IC e OC della carta PM sono entrambe uguali circa 200, poiché quando una forma quadratica viene adattata ai dati usando i minimi quadrati, i coefficienti stimati saranno molto simili a quelli del modello IC. Questo fenomeno dimostra lo svantaggio delle carte di controllo parametriche e la loro inefficienza quando il modello OC è mal specificato. Infatti queste due curve hanno leggere differenze nella totalità del profilo, ma hanno alcuni cambiamenti più accentuati in prossimità dei punti di rilevazione della carta. La carta *NEWMA* quindi può catturare questi dettagli più accuratamente e segnalare un allarme più velocemente.

Il modello (III) è una forma quadratica del modello più un'oscillazione di una funzione seno. La carta PM coglie le differenze nella maggior parte dei casi, anche se il metodo parametrico non è il più corretto. In generale però è leggermente peggiore della carta *NEWMA* ad eccezione del caso (iii) dove le differenze tra IC e OC non sono rilevanti.

Per il secondo confronto consideriamo lo schema NM e *NEWMA* che sono metodologie non parametriche mentre non prendiamo in considerazione PM in quanto è un metodo parametrico e quindi poco appropriato per lo studio dei modelli non lineari. Il modello non lineare in controllo di riferimento è:

$$y_{ij} = 1 - \exp(-x_{ij}) + \varepsilon_{ij}, \quad i = 1, 2, \dots$$

e i tre modelli OC sono:

$$(I): y_{ij} = 1 - \beta_1 \exp(-x_i^{\beta_2}) + \varepsilon_{ij},$$

$$(II): y_{ij} = 1 - \exp(-x_i) + \beta_1 \cos(\beta_2 \pi(x_i - 0.5)) + \varepsilon_{ij},$$

$$(III): y_{ij} = \frac{1}{1 + \beta_1 x_i^{\beta_2}} + \varepsilon_{ij}.$$

I parametri scelti sono descritti in questo caso nella Tabella 12:

Tabella 12: Parametri dei tre modelli OC per lo Scenario 2

	Model (I)			Model (II)			Model (III)		
	β_1	β_2	σ	β_1	β_2	σ	β_1	β_2	σ
(i)	1.0	1.3	1.0	.2	3.0	1.0	10.0	.1	1.0
(ii)	1.0	1.5	1.0	.3	3.0	1.0	10.0	.2	1.0
(iii)	1.1	1.0	1.0	.2	2.0	1.0	40.0	1.5	1.0
(iv)	1.3	1.0	1.0	.3	2.0	1.0	60.0	1.5	1.0
(v)	1.2	1.0	1.1	.2	4.0	1.1	50.0	1.2	1.1
(vi)	1.0	1.2	.7	.2	4.0	1.3	-3.0	4.0	1.1

Tabella 13: Confronto di ARL tra NEWMA e NM per lo Scenario 2

		$n = 20$				$n = 40$			
		NEWMA			NM	NEWMA			NM
		$c = 1.0$	$c = 1.5$	$c = 2.0$		$c = 1.0$	$c = 1.5$	$c = 2.0$	
(I)	(i)	144.5	139.9	134.2	168.3	110.3	103.1	100.6	155.0
	(ii)	99.1	93.0	86.2	131.2	63.8	58.7	55.1	113.3
	(iii)	113.7	109.1	102.2	143.0	79.5	72.7	67.4	129.1
	(iv)	19.9	17.8	16.6	31.8	10.6	9.8	9.4	21.0
	(v)	17.9	16.6	15.7	25.4	10.3	9.7	9.2	17.1
	(vi)	8.2	6.5	5.8	*	4.2	3.7	3.5	*
(II)	(i)	39.4	40.9	45.9	59.3	20.3	20.1	21.4	43.3
	(ii)	17.9	18.1	20.6	27.1	9.7	9.4	9.9	18.4
	(iii)	37.9	35.2	34.4	60.2	20.0	18.4	17.8	43.4
	(iv)	17.3	16.0	15.5	27.0	9.4	8.8	8.7	18.3
	(v)	17.4	17.8	19.0	24.1	9.8	9.9	10.3	16.1
	(vi)	5.4	5.0	4.9	8.5	3.5	3.3	3.3	6.0
(III)	(i)	69.2	61.9	58.4	99.4	39.1	34.9	33.2	80.1
	(ii)	40.2	36.7	33.9	64.8	21.5	19.4	18.1	46.5
	(iii)	11.8	10.9	10.6	17.9	6.9	6.4	6.2	12.1
	(iv)	15.4	14.6	13.9	23.7	8.4	8.0	7.7	15.7
	(v)	14.0	13.3	12.9	19.8	7.9	7.5	7.4	12.8
	(vi)	7.8	7.3	7.0	11.8	4.8	4.5	4.4	8.0
L		20.25	17.25	15.63	38.69	21.66	18.28	16.50	65.20

Si noti che i limiti di controllo per la carte NEWMA sono uguali al caso precedente poiché non dipendono dal modello in controllo ma solo da n , h e X . Dalla Tabella 13 vediamo che la carta NEWMA si comporta sempre meglio della carta NM per tutti i tipi di modello OC e per tutte le variazioni dei parametri considerate

A conclusione di questo doppio confronto, prima sul modello di forma quadratica poi sul modello lineare, la carta *NEWMA* mostra una maggiore robustezza e sensibilità ai cambiamenti nella funzione di regressione e nella varianza rispetto alle due metodologie *NM* e *PM*. Usare questo tipo di carta quindi conviene sempre perché in termini di efficienza porta un risultato simile se non migliore ai metodi alternativi parametrici che sono influenzati dal problema dell'errata specificazione del modello fuori controllo. La scelta della banda porta degli effetti sull'ARL fuori controllo, ma usando la formula [19] si possono ottenere risultati soddisfacenti in un'ampia varietà di situazioni. I differenti valori di $c \in [1.0, 2.0]$ offrono risultati di ARL. Un valore piccolo di c di solito è più efficace nel cogliere un cambiamento marcato in un'area precisa, invece una valore grande si comporta meglio quando la differenza tra i modelli *IC* e *OC* non è marcata.

6.5 Lo schema di controllo ACP

Abbiamo visto nelle sezioni precedenti un approccio non parametrico presentato da Zou, Tsung e Wang (2008) per il modello [17]:

$$y_{ij} = g(x_{ij}) + \varepsilon_{ij}, \quad i = 1, 2, \dots, n_j \text{ e } j = 1, 2, \dots,$$

La carta vista prima quindi integrava un classico schema *EWMA* con una procedura basata su un test *GLR* per trattare il caso di una regressione non parametrica.

Uno dei svantaggi dello schema *NEWMA* è che in molte applicazioni pratiche le assunzioni di base del modello non sono soddisfatte e quindi la sua performance non è soddisfacente. Per prima cosa la carta si rifà alla vera funzione di regressione in controllo, g , e alla varianza dell'errore σ^2 del modello. Nella pratica sia g , che σ^2 sono solitamente incogniti e necessitano di un data set in controllo per essere stimati e se questo è di piccole dimensioni, vi sarà una certa incertezza nella stima dei parametri che porta a distribuzione distorta della run-length in controllo della carta. Anche se i limiti di controllo della carta vengono aggiustati adeguatamente per ottenere un desiderato comportamento dell'ARL in controllo, l'OC ARL potrebbe essere gravemente compromesso (cf. Jones (2002)). In secondo luogo la carta *NEWMA* assume che la distribuzione dell'errore sia normale, mentre nella pratica, anche in questo caso, la maggior parte delle volte questa è sconosciuta. In questi casi è molto difficile costruire una carta *NEWMA* adeguata. Un terzo problema riscontrato è che la carta dipende dalla

scelta della banda di lisciamento usata per il profilo. Per ultimo diciamo ancora che la migliore performance della carta si ottiene solo dopo che questa viene collegata alla grandezza dello shift che il più delle volte però è sconosciuto.

Zou, Qiu e Hawkins (2008), quindi propongono una nuova carta di controllo che ha come oggetto di risolvere queste problematiche. Questa carta adotta l'approccio *change point* suggerito da Hawkins, Qiu e Kang (2003) che simultaneamente stima i parametri e controlla le condizioni di fuori controllo. Una procedura adattiva per la selezioni di parametri di lisciamento è incorporata nella costruzione della carta di controllo così da adattarsi al meglio ai casi in cui vi siano differenze tra le funzioni di regressione in controllo e fuori controllo con una conseguente maggiore robustezza a varie condizioni dei profili OC. Successivamente una procedura bootstrap è usata per determinare i limiti di controllo senza conoscere la distribuzione dell'errore ma solamente usando m_0 profili in controllo per stimarne la distribuzione.

Assumiamo che siano disponibili m_0 profili in - controllo. In molte applicazioni pratiche, si tende a raccogliere un numero elevato di osservazioni per ogni profilo nell'analisi di Fase 1 a differenza della Fase 2, perché le analisi più approfondite sono svolte nella la prima fase, mentre le condizioni di fuori controllo possono essere comunque catturate anche con un numero minore di campioni. Distinguiamo con n_0 in numero di osservazioni in ogni profilo m_0 e n per indicare il numero di osservazioni in ogni profilo futuro con $n \leq n_0$. Inoltre sia $X_{IC} = \{x_1, \dots, x_n, x_{n+1}, \dots, x_{n_0}\}$ la matrice del disegno corrispondente agli m_0 profili IC con le corrispondenti variabili risposta $\{y_{1j}, \dots, y_{nj}, y_{(n+1)j}, \dots, y_{n_0j}\}$, per $j = 1, \dots, m_0$. La procedura di Fase 2 è così descritta:

- dopo aver ottenuto il $(t - m_0)$ profilo di Fase 2, con $t > m_0$, si calcola:

$$lr_{m_0,t} = \max_{m_0 \leq k < t} \tilde{T}_{k,t}.$$

- se $lr_{m_0,t} > h_{m_0,t,\alpha}$, dove $h_{m_0,t,\alpha}$ è un determinato limite di controllo, si segnala un fuori controllo. Dopo il segnale, la procedura di diagnostica descritta da Zou et al. (2008) può essere utilizzata per localizzare il significato del cambiamento del profilo
- se invece $lr_{m_0,t} < h_{m_0,t,\alpha}$ la sorveglianza del processo continua ottenendo il $(t + 1)$ profilo di Fase 2 e ripetendo i primi due passi.

Questo schema differisce dalla carta *NEWMA* per il fatto che la funzione IC e la varianza dell'errore in ogni profilo osservato possono essere entrambi non noti, mentre per la *NEWMA* i parametri vengono assunti noti. Si noti che calcolare la statistica di controllo $lr_{m_0,t}$ comporta la stima della varianza dell'errore σ^2 per ogni t . Quando t è dato, uno stimatore consistente di σ^2 è dato dalla formula:

$$\hat{\sigma}_t^2 = \frac{1}{t(n-df)} \sum_{j=1}^t (Y_j - W_h Y_j)^{\otimes}.$$

Questo stimatore può essere aggiornato facilmente al $t+1$ esimo profilo con la formula:

$$\hat{\sigma}_{t+1}^2 = [(t+m_0)(n-df)\hat{\sigma}_t^2 + (Y_t - W_{h_b} Y_t)^{\otimes}] / [(t+1+m_0)(n-df)],$$

dove h_b è una banda di liscio pre-specificata. Nelle applicazioni pratiche sarebbe più conveniente disegnare nelle carte di controllo la statistica normalizzata $lr_{m_0,t}/h_{m_0,t,\alpha}$ verso t . In alcuni casi, il limite di controllo normalizzato è costante e pari a 1.

Dalla costruzione di $lr_{m_0,t}$, è facile da vedere che questa non dipende dalla regressione in controllo g come anche da $h_{m_0,t,\alpha}$.

Zou, Qiu e Hawkins (2008), propongono una procedura bootstrap per la determinazione di $h_{m_0,t,\alpha}$ basato su m_0 profili IC. Per prima cosa bisogna trasformare in media gli m_0 profili IC e poi questi dati vengono liscio secondo un nucleo lineare. Il risultato è una matrice $m_0 \times n_0$ di residui:

$$\hat{e}_{ij} := y_{ij} - W_{n_0}(x_i) \bar{Y}_{0,m_0}, \quad j = 1, \dots, m_0, \quad i = 1, \dots, n_0,$$

dove $W_{n_0}(x_i)$ è un operatore di liscio definito precedentemente, usando X_{IC} al posto di X e h_b . Successivamente, si genera un profilo bootstrap $\{(x_i, y_i^*), i = 1, \dots, n\}$ definendo $y_i^* = e_i^*$ per $i = 1, \dots, n$, dove e_i^* è campionato con reinserimento da $\{\hat{e}_{ij}, i = 1, \dots, n_0, j = 1, \dots, m_0\}$. Dunque, per ogni valore di t , con questa procedura di ricampionamento, si simula il processo di analisi, incluso la generazione dei primi m_0 profili in controllo, e i futuri profili e il valore $lr_{m_0,t}^*$ corrispondente. L'intera procedura viene ripetuto per B volte. Poi per un data probabilità di falso allarme, α , che

corrisponde a un ARL IC pari a $1/\alpha$, i limiti di controllo $h_{m_0,t,\alpha}$ possono essere approssimati da :

$$\Pr(lr_{m_0,t}^* > h_{m_0,t,\alpha} | lr_{m_0,i}^* \leq h_{m_0,i,\alpha} \leq i < t) = \alpha, \text{ for } t > 1,$$

$$\Pr(lr_{m_0,1}^* > h_{m_0,1,\alpha}) = \alpha.$$

Ovviamente le probabilità vanno interpretate come frequenze nelle B bootstrap replicazioni.

Basate su esperienze numeriche e risultati empirici di Hawkins, Qui e Kang (2003) che hanno discusso limiti di controllo simili basati su metodi bootstrap, si può osservare che $h_{m_0,t,\alpha}$ converge gradualmente a una costante, al crescere di t .

7. Illustrazione di un'applicazione: Il caso DRIE

In questo capitolo viene illustrata un'applicazione dei metodi presentati allo studio di un processo di incisione (DRIE). Tale processo, utilizzando un plasma con caratteristiche chimico reattive, rimuove del materiale dalle sottili lastre di materiale semiconduttore (wafer) come il silicio, da cui si ricavano microcircuiti.

Il processo DRIE si svolge all'interno della camera centrale di una macchina per induzione di due diversi tipi di plasma di silicene con al di sotto un piano dove appoggiare il wafer. Il sistema prima rilascia il plasma, che incide nella camera per ricavare un circuito desiderato, poi vengono introdotti differenti tipi di gas che generano un film di protezione delle pareti dei circuiti. I due step di incisione e deposizione sono ripetuti alternativamente fino a quando non viene raggiunto il limite di tempo stabilito dal processo oppure quando un modulo di rilevazione conferma la corretta profondità dell'incisione. La qualità del processo viene misurata in laboratorio con una scansione al microscopio elettronico. Una delle caratteristiche di qualità più importanti è il profilo della sezione che porta informazioni importanti sulle operazioni di incisione. Il profilo desiderato è quello con gli angoli smussati e pareti laterali perpendicolari come riportato al centro della Figura 1 che è chiamato profilo anisotropo. Idealmente le pareti laterali sono perpendicolari al fondo con un certo grado di lisciamento sugli angoli.

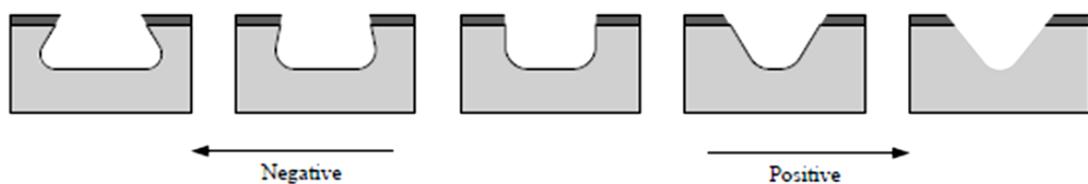


Figura 1: Possibili profili di incisione di un processo DRIE

Le applicazioni industriali correnti studiano l'angolo di incidenza delle pareti del profilo che, anche se di facile misurazione, contengono informazioni incomplete sulla sezione e in alcuni casi non differenziano abbastanza tra i profili fuori controllo.

Gli schemi proposti non vengono applicati all'intero profilo inciso nel wafer perché questo non può essere visto come un modello polinomiale o un modello lineare

generalizzato, ma vengono presi in esame solamente gli angoli che comunque contengono una sufficiente quantità di informazioni per riuscire a distinguere le condizioni di fuori controllo. I due lati della sezione presentano una forma simmetrica quindi viene preso in esame solo l'angolo di sinistra per implementare gli schemi di controllo.

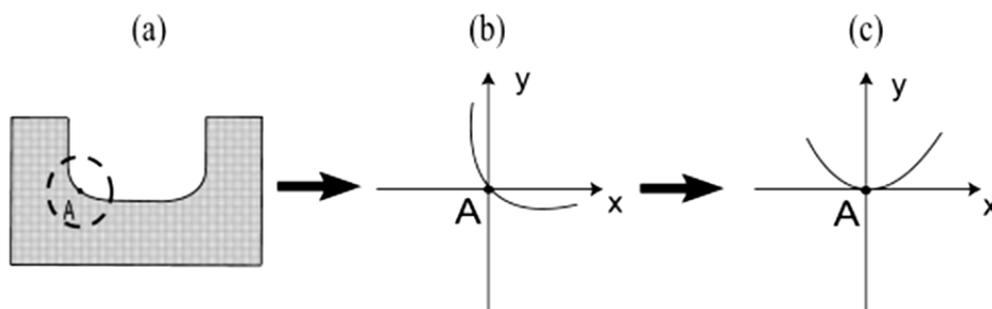


Figura 2: Modellazione del profilo DRIE

Come possiamo notare nella Figura 2 si ricava un punto di riferimento A secondo delle precise coordinate scelte precedentemente. Trovato il punto A, si ruota in senso antiorario l'angolo del profilo di 45 gradi come si vede nella Figura 2 (b) e (c). La curva dell'angolo ruotato quindi può essere ben descritta da un modello polinomiale quadratico.

Il data set di riferimento è costituito da 18 profili anisotropi in controllo. Per ogni profilo vengono fissati gli $x_i, i = 1, \dots, 11$, che sono valori equidistanti compresi tra -2.5 e 2.5 con distanza pari a 0.5 .

Il modello IC ottenuto è:

$$y_{ij} = ax_i^2 + \varepsilon_{ij}, \quad i = 1, \dots, n,$$

dove $a = 0.62$ e $\sigma = 0.4$. Il modello ha quindi i parametri: $\beta = (\beta^{(1)}, \beta^{(2)}, \beta^{(3)}) = (1.55, 0, 0.62)$. Il cambiamento dei coefficienti dei parametri di regressione β , indica un profilo negativo o positivo causato da una eccessiva incisione o al contrario da una non sufficiente pulizia. Un incremento in varianza invece corrisponde a un profilo ruvido o a una poca accuratezza nel processo di *etching*. Un decremento in varianza della curva invece indica un incremento della capacità del processo fermo restando il non

cambiamento dei parametri del modello. Con queste considerazioni di base passiamo ad applicare i modelli *MEWMA*, *NEWMA* e *ACP*.

7.1 L'applicazione del metodo MEWMA.

I passi per l'implementazione della carta di controllo MEWMA al processo DRIE sono i seguenti.

Per prima cosa si sceglie l'ARL in controllo desiderata e la costante di lisciamento λ e si determinano i limiti di controllo, L , basati su p , ARL_{IC} e λ . Nel caso preso in esame con

$$p=3, ARL = 370 \text{ e } \lambda = 0.2 \text{ si trova } L = 15.41 \text{ Di conseguenza } L \frac{\lambda}{2-\lambda} \text{ sar\`a } 1.71.$$

Il secondo passo consiste nel sorvegliare il processo e ottenere le osservazioni, y_{ij} , del prodotto a valori fissati di x_{ij} . Successivamente si riportano sul grafico i valori della statistica U_j e si confronta con i limiti di controllo (Figura 3).

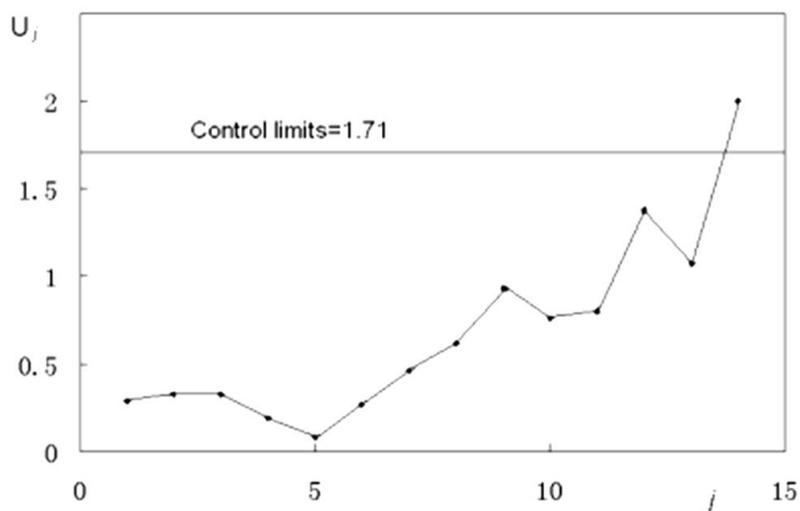


Figura 3: La carta MEWMA basata sui dati del processo DRIE

Poiché è molto dispendioso e difficile raccogliere campioni di profili fuori controllo in laboratorio, viene artificialmente aggiunto ai dati dell'esempio uno shift di $\beta^{(3)}$ per il modello IC da 0.62 a 0.67 dopo il quinto campione e vengono generati campioni fuori controllo tramite una simulazione Monte Carlo. Le osservazioni y_{ij} e la corrispondente statistica U_j sono riportati nella Tabella 14.

Tabella 14: Dati per l'esempio con uno shift in $\beta^{(3)}$ dopo il quinto campione

j	y_{ij}											U_j	lr
0													10.59
1	2.64	2.70	2.10	.35	.32	.35	.94	.31	1.15	2.70	4.09	.29	13.15
2	4.60	2.51	1.28	.94	-.09	-.29	.96	.82	1.38	1.66	3.57	.33	14.43
3	4.21	2.09	1.36	1.18	.34	-.67	.54	-.03	1.51	2.82	3.70	.33	14.92
4	3.11	2.46	1.59	.46	-.45	.44	.05	.49	1.71	2.79	4.06	.19	17.07
5	4.14	2.19	1.08	.47	.29	-.08	.00	.80	1.05	3.02	4.00	.08	17.78
6	3.64	2.87	.25	.51	.24	.37	-.53	.61	1.40	2.35	4.31	.27	17.65
7	4.40	2.58	1.25	.41	.13	.19	-.31	-.09	1.15	3.42	4.07	.46	14.09
8	3.92	2.78	1.74	.10	.61	-.99	-.02	.30	2.05	2.56	3.61	.62	13.03
9	4.66	2.97	1.50	1.06	.12	-.48	-.54	.48	1.21	3.01	3.73	.93	9.15
10	4.34	2.27	1.31	.52	.19	-.10	.14	1.07	1.04	3.02	3.86	.76	11.11
11	3.89	1.82	1.31	.08	-.11	.49	-.18	-.28	1.68	2.62	4.12	.80	11.12
12	4.52	2.43	1.50	.46	-.38	-.61	.01	1.32	1.20	2.71	4.42	1.38	9.67
13	3.53	2.59	2.57	.26	.01	.13	.12	.43	2.02	2.39	4.20	1.07	14.15
14	4.42	3.02	1.93	-.37	.30	-.95	-.65	1.14	.76	1.55	3.93	2.00	

La Figura 4 mostra i profili IC e OC in sequenza temporale. Dopo il quinto campione il processo inizia a generare profili negativi che sono considerati difettosi. Possiamo notare che la grandezza del cambiamento in questi profili è solitamente piccola ed è per questo che sono necessari schemi di controllo sensibili ai piccoli cambiamenti. Dalla Figura 3 e dalla Tabella 14 possiamo vedere come lo schema *MEWMA* segnali velocemente un cambiamento al campione numero 14.

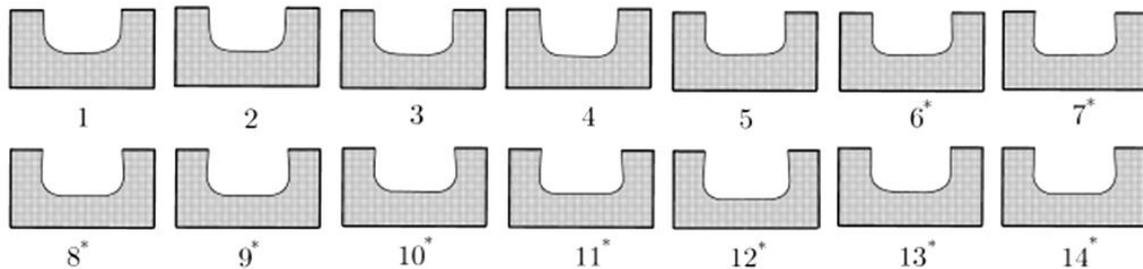


Figura 4: I 14 profili dove con il simbolo “*” vengono segnalati quelli OC.

Il terzo passo consiste nel valutare il valore del test rapporto di verosimiglianza generalizzato $lr(jn, 14n)$ per $j = 0, 1, \dots, 13$, che troviamo nell'ultima colonna della Tabella 14 e possiamo vedere come il massimo corrisponde a $j = 5$ con $lr(5n, 14n) = 17.78$ che indica precisamente l'effettivo punto di cambiamento, τ .

Il quarto passo serve per verificare quale parametro è effettivamente cambiato e quindi utilizziamo i test visti precedentemente nel capitolo dedicato. Per i tre test otteniamo i valori. $T_{TEST} = -0.427$, $\chi^2_{TEST} = 115.3$, $F_{TEST}^{(2)} = 0.019$ e $F_{TEST}^{(3)} = 13.4$. Considerando un livello di significatività, $\alpha = 0.05$ segue che $|T_{TEST}| < |t(0.025; 9n - 3)| = 1.985$, che $70.8 = \chi^2(0,025; 9n - 3) < \chi^2_{TEST} < \chi^2(0,975; 9n - 3) = 125$, che $F_{TEST}^{(2)} <$

$F_{0.95}(2,9n - 3, R) = 3.94$ e che $F_{TEST}^{(3)} > F_{0.95}(2,9n - 3, R) = 3.94$ dove $t(\alpha, \nu)$ e $\chi^2(\alpha; \nu)$ sono i percentili inferiori di una distribuzione t di Student e chi-quadro rispettivamente con ν gradi di libertà. Questa diagnosi quindi conclude positivamente che c'è stato uno shift positivo nel parametro $\beta^{(3)}$. Se nella realtà ci fossimo trovati davanti a questo tipo di problema con un cambiamento del parametro di secondo ordine sarebbe stato necessario cambiare i settaggi della macchina e riesaminare le condizioni del processo DRIE.

Il quinto e ultimo passo ci dice, una volta identificati gli eventuali parametri fuori controllo, di risolvere il problema e quindi di ripartire dal primo passo del processo di sorveglianza.

7.2 L'applicazione del metodo NEWMA

Le considerazioni di partenza sono le stesse fatte per la carta di tipo *MEWMA* quindi il modello di riferimento è lo stesso

$$y_{ij} = 0.62x_i^2 + \varepsilon_{ij}, \quad i = 1, \dots, n,$$

come lo è anche il data set di riferimento che contiene i cinque profili in controllo e i nove fuori controllo. I 5 passi dell'analisi sono anche qui gli stessi del caso *MEWMA* e andiamo a vederli più approfonditamente.

Passo 1. Viene adottato il nucleo di una funzione Epanechnikov per $K(\cdot)$ e viene calcolato h con $c = 1.5$. In accordo con i valori dei punti x_i si ottiene data un'ARL=370 e $\lambda = 0.2$. il limite di controllo $L = 18.09$. Di conseguenza, $L \frac{\lambda}{2-\lambda}$ sarà 2.01.

Passo 2. Si inizia la sorveglianza del processo ottenendo i vettori Z_j , E_j e quindi anche la statistica Q_j che viene confrontata con il limite di controllo trovato al passo precedente. Nella Figura 5 la carta segnala un all'arme al 14 campione come effettivamente fa anche la carta parametrica *MEWMA*.

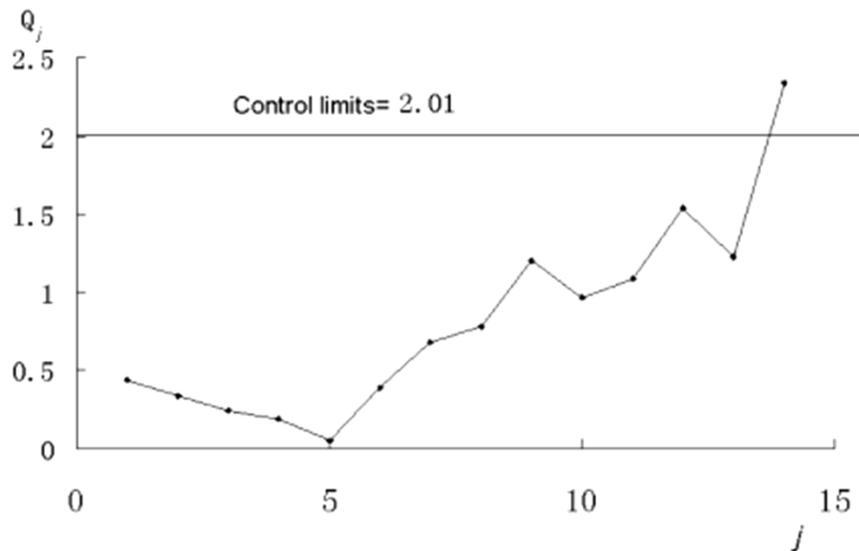


Figura 5: Carta NEWMA per il monitoraggio di un processo DRIE

Passo 3. Dai valori del test rapporto di verosimiglianza generalizzato $lr(jn, 14)$ si può vedere che il massimo lo otteniamo a $j = 6$ per il quale $lr(6n, 14n)$ è uguale a 18.60. Nella realtà il punto di change point è $\tau = 5$ che viene indicato precisamente dal metodo parametrico di Zou et al (2007). In questo caso lo schema non parametrico ottiene risultati meno accurati ma comunque vicini alla diagnosi reale.

Passo 4. Utilizzando i test visti nel capitolo di riferimento si ottiene un p-value pari a 0.209 e 0.002 che indicano che la varianza è stabile mentre sembra esserci uno shift nella funzione di regressione

Infine si disegna la curva locale di lisciamento della media degli ultimi otto campioni e la curva corrispondente al modello in controllo (Figura 6).

Dalla Figura 6 si può vedere che il profilo OC presenta un trend negativo dovuto forse a un incisione troppo accentuata del processo.

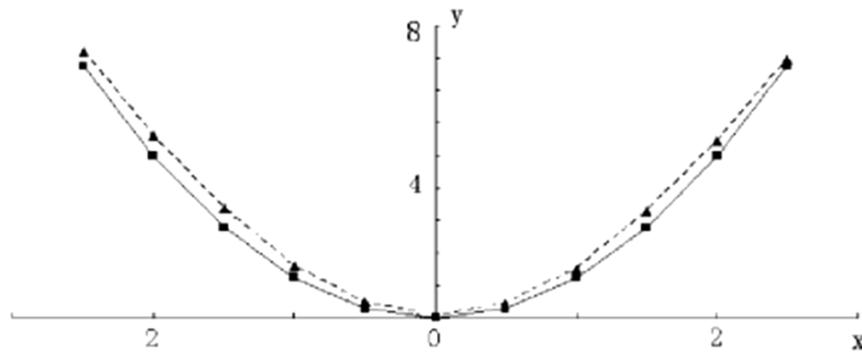


Figura 6: Diagnosi grafica per il processo DRIE. La linea continua rappresenta il modello in controllo, mentre quella tratteggiata il modello OC stimato

A questo punto per dimostrare l'efficacia del metodo non parametrico prendiamo in considerazione l'intero profilo trasformato come vediamo nella Figura 7.

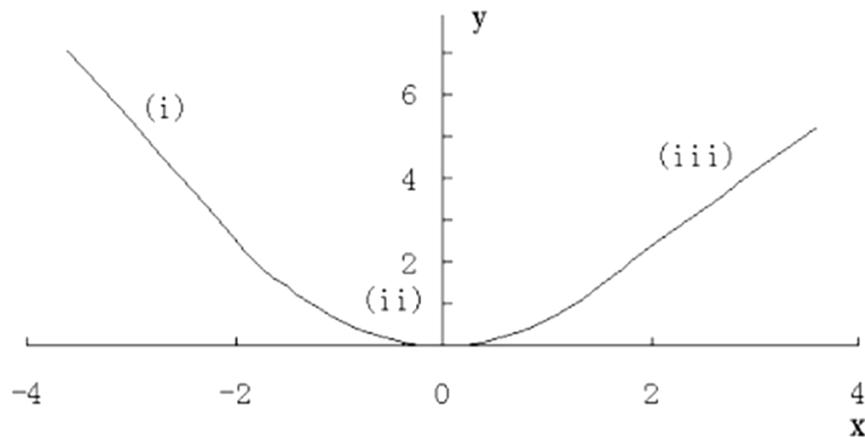


Figura 7: Il profilo del processo DRIE

Mentre prima prendevamo in considerazione solo la parte centrale (ii) del profilo, ora si considera l'intera curva. Questa non può essere vista come un modello lineare generale o polinomiale e i quindi metodi parametrici non possono essere applicati direttamente. Si applica quindi uno schema *NEWMA*. Con questo tipo di carte per avere un performance ottimale è necessario un numero di osservazioni elevato specialmente quando il profilo è complicato. Per questo vengono prese misurazioni su 35 punti scelti nel modo seguente: tra -3.6 e -1.8 sette punti a distanza di 0.3 come anche tra 1.8 e 3.6 , mentre tra -1.5 e 1.5 i punti sono ventuno a distanza di 0.15 . A questo punto vengono raccolti i campioni su 18 profili anisotropi in controllo. La linea di lisciamento della curva della media di questi 18 IC profili è riportata nella Figura 8.

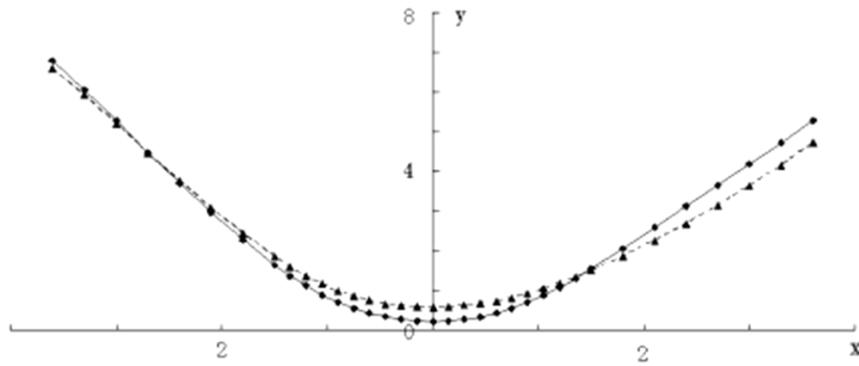


Figura 8: Grafico del nuovo profilo DRIE. La linea continua rappresenta il modello IC, mentre quella tratteggiata il modello OC stimato

Questa curva può essere considerata la stima (\hat{G}_0) del vero modello (G_0) in controllo e così non c'è necessità di modellare questo profilo con forme polinomiali. Inoltre la stima della varianza è di 0.409. Da qui possiamo partire con l'analisi e le procedure di diagnostica.

Similmente a quanto fatto prima, usiamo un ARL pari a 370, $\lambda = 0.2$ e $c = 1.5$ e un limite di controllo pari a 2.30. Vengono generati inoltre altri 6 profili in controllo aggiungendo errori casuali al modello G_0 stimato dai 18 profili anisotropi e altri 3 profili non in controllo che vengono considerati negativi. La statistica di riferimento U_j e la carta $NEWMA$ ottenuta la vediamo in Figura 9.

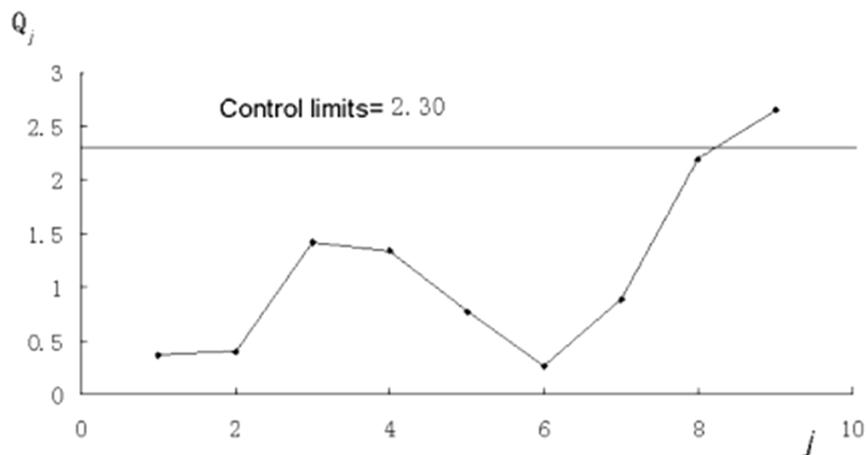


Figura 9: NEWMA carta per il nuovo esempio

Si può vedere che la carta segnala un allarme dopo il terzo. Osservando il test rapporto $lr(jn, 9n)$ si osserva che il massimo lo si ha a $j = 6$. Questo risultato indica accuratamente il momento in cui è avvenuto il cambiamento. Inoltre dai test di

diagnostica otteniamo i p-value 0.633 e 10^{-6} che indicano che la funzione di regressione è cambiata. Un'ulteriore conferma ci viene data dal grafico in Figura 8 dove vengono confrontate la curva del modello IC e quella della media dei 3 profili fuori controllo. Vediamo come ci sia un trend positivo dovuto verosimilmente a una non completa pulizia del fondo del minicircuito. Questo suggerisce una revisione dei settaggi della macchina e un riesame del processo DRIE.

7.3. L'applicazione del metodo ACP.

Lo schema di controllo *ACP* di Zou, Qiu e Hawkins (2008) viene applicato all'intero profilo del processo DRIE. Il data set di riferimento è quello dei 18 profili in controllo presi in considerazione dei quali vengono utilizzati i primi otto per implementare la procedura di bootstrap mentre i restanti dieci vengono scartati. Per ognuno di questi otto profili vengono misurati settanta punti come richiesto dalla procedura di bootstrap, questo permette di avere una buona approssimazione della performance IC desiderata anche se non viene fatta alcuna assunzione sulla distribuzione degli errori. Anche in questo caso l'*ARL* in controllo è 370 e tutti i parametri sono ricavati tramite simulazione. La carta *ACP* è usata per sorvegliare i profili in Fase 2 ognuno dei quali ha 35 osservazioni. Come visto prima ci sono nove profili di Fase 2 nel data set e gli ultimi tre sono negativi.

La Tabella 15 riporta vari risultati

Tabella 15: Risultati della carta ACP per il monitoraggio del data set DRIE

t	ACP				
	$lr_{m_0,t}$	$\hat{\sigma}_t$	$h_{m_0,t,\alpha}$	$lr_{m_0,t}/h_{m_0,t,\alpha}$	$\hat{\tau}$
9	0.740	0.431	6.526	0.113	8
10	0.924	0.420	6.816	0.136	8
11	2.493	0.423	7.112	0.351	8
12	1.957	0.420	7.194	0.272	8
13	1.126	0.418	7.437	0.151	8
14	1.573	0.416	7.741	0.203	8
15	2.423	0.411	8.136	0.298	14
16	6.771	0.411	8.487	0.798	14
17	9.053	0.408	8.702	1.040	14

La carta di controllo basata sulla statistica rapporto $lr_{m_0,t}/h_{m_0,t,\alpha}$ viene mostrata in Figura 10 dove il limite di controllo è 1. Sia da quest'ultima che dalla Tabella 15 possiamo vedere come il segnale di allarme sia dato al 17-esimo profilo in accordo con la carta *NEWMA* di Zou, Tsung e Wang (2008). Anche il punto di cambiamento viene stimato perfettamente in $\tau = 14$. Il vantaggio di questa carta rispetto a quella *NEWMA* è che questa non richiede un metodo di diagnostica separata.

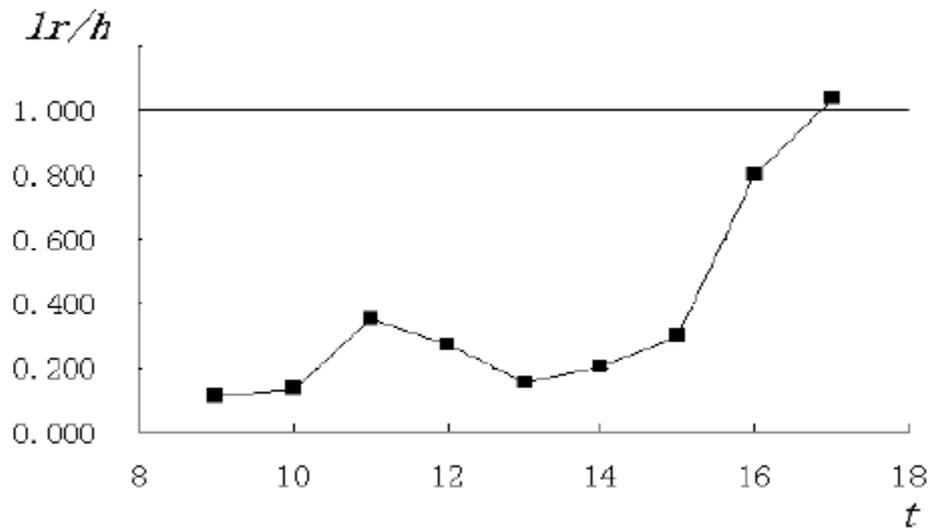


Figura 10: La carta di controllo ACP per il processo DRIE

Concludendo i metodi proposti rispondono perfettamente alle esigenze aziendali quali una veloce segnalazione in caso di cambiamento dei parametri di regressione della funzione sia in caso di cambiamento nella varianza del processo.

8. Conclusioni

L'obiettivo della tesi era quello di illustrare alcuni degli approcci parametrici e non parametrici per lo studio dei profili lineari e dei profili lineari generali. L'analisi dei profili lineari si divide in due momenti denominati Fase 1 e Fase 2. L'obiettivo principale per lo studio di Fase 1 è di ricavare, dai campioni raccolti, un modello di riferimento per il processo in controllo da adottare per lo studio di Fase 2. Le carte di controllo che abbiamo illustrato sono: il metodo univariato $EWMA_3$ o KMW di Kim et al. (2003), e il metodo multivariato EWMA/R e T^2 di Kang e Albin (2000). Queste carte vengono applicate per sorvegliare possibili cambiamenti nei parametri della regressione, intercetta e coefficiente angolare, e nella varianza del processo. Le due metodologie si differenziano dal fatto che, ognuna delle tre carte EWMA di Kim et al (2003) studiano un singolo parametro del modello di regressione e la varianza, mentre sia la carta EWMA/R sia la carta T^2 li studiano contemporaneamente. Gli stessi metodi di analisi sono utilizzati per la sorveglianza del profilo in Fase 2. I confronti tra i vari metodi mostrano che generalmente l'approccio $EWMA_3$ si comporta meglio della carta EWMA/R per rilevare cambiamenti moderati per l'intercetta mentre la carta T^2 è più adeguata per cambiamenti più grandi. Stessi risultati sono stati rilevati anche per cambiamenti del coefficiente angolare. Per quanto riguarda i cambiamenti della varianza il metodo EWMA/R si comporta decisamente meglio degli altri due in quanto la carta R è più adatta a rilevare tale tipo di cambiamento. Quando ci troviamo di fronte a cambiamenti simultanei di intercetta e coefficiente angolare la carta univariata EWMA ha un'efficienza maggiore rispetto alle altre due. Si può quindi affermare che studiare i singoli parametri della regressione con carte che studiano i singoli parametri porta a risultati più efficienti sul piano dell'ARL fuori controllo e a una più semplice interpretazione degli eventuali segnali di fuori controllo. Infatti diventa difficile rilevare quale parametro sia cambiato nelle carte multivariate.

Dopo è stato introdotto un altro schema di controllo basato sempre su carta EWMA ma per profili non lineari. La carta MEWMA proposta da Zou, Wang e Tsung (2007) si propone di studiare i profili lineari generali, cioè quando il modello di riferimento non è più lineare. Tramite simulazione abbiamo confrontato le carte EWMA3 con questo metodo per rilevare shift dei parametri della funzione di regressione in controllo sia per

identificare derive della varianza del processo. I risultati hanno dimostrato che i due schemi di controllo hanno una performance simile. La carta MEWMA invece si comporta meglio quando più parametri cambiano contemporaneamente. Abbiamo visto inoltre che è possibile migliorare la performance della carta applicando uno schema con un intervallo di campionamento variabile (VSI). I tempi di rilevazione di possibili cambiamenti decrescono in maniera significativa. Possiamo quindi concludere che a livello pratico questo schema di campionamento può essere presa in considerazione per migliorare la qualità del processo, anche se spesso bisogna scontrarsi con i limiti dettati dalla realtà produttiva dove tale raccolta dei dati può non essere così semplice. Un esempio lo abbiamo nel processo DRIE descritto nell'esempio dove la misurazione viene effettuata tramite microscopio elettronico. L'esempio di questo processo ci viene in aiuto per presentare meglio i risultati del confronto tra il metodo MEWMA e gli approcci non parametrici, NWEMA e ACP. Dall'analisi svolta abbiamo visto come le tre metodologie diano risultati soddisfacenti per la rilevazioni dei cambiamenti della funzione di regressione. La carta MEWMA per la parte centrale della funzione fornisce una performance migliore della carta NEWMA anche se la differenza può essere trascurabile. Per le carte non parametriche il confronto si è basato sulla totalità della curva dove un modello parametrico non era idoneo alla descrizione del profilo. Qui abbiamo visto che sia la carta NEWMA che la carta ACP riescono a individuare perfettamente il cambiamento della curva. Possiamo concludere quindi che tutte e tre le carte producono in modo soddisfacente una veloce risposta in caso di fuori controllo dei parametri della funzione di regressione o della varianza del processo oltre che una facile individuazione dell'eventuale punto di cambiamento.

Bibliografia

Colosimo, B. M., & Pacella, M. (2007). On the Use of Principle Component Analysis to Identify Systematic Pattern in Roundness Profiles. *Quality and Reliability Engineering International* , 23, p. 707-725.

Kang, L., & Albin, S. L. (2000). On-Line Monitoring When the Process Yields a Linear Profile. *Journal of Quality Technology* , 32, p. 418-426.

Kazemzadeh, R. B., Noorossana, R., & Amiri, A. (2008). Phase I Monitoring of Polinomial Profiles. *Communication in Statistics - Theory and Methods* (37), 1671-1686.

Kim, K., Mahmoud, M. A., & Woodhall, W. H. (2003). On the Monitoring of Linear Profiles. *Journal of Quality Technology* , 35, p. 317-328.

Mahmoud, M. A., & Woodall, W. (2004). Phase I Analysis of Linear Profiles With Calibration Applications. *Technometrics* , 46 (4), 380-391.

Mahmoud, M. A., Parker, P., Woodhall, W. H., & Hawkins, D. M. (2007). A Change Point Method for Linear Profile Data. *Quality and Reliability Engineering International* , 23, p. 247-268.

Williams, J. D., Woodall, W., & Birch, J. B. (2007). Statical Monitoring of Nonlinear Product and Process Quality Profiles. *Quality and Reliability Engineering International* (23), 925-941.

Woodall, W. H., Spitzner, D. J., Montgomery, D. C., & Gupta, S. (2004, Luglio). Using Control Chart to Monitor Process and Product Quality Profiles. *Journale of Quality Technology* , 36 (3), p. 309-319.

Zou, C., Qiu, P., & Hawkins, D. (2008). Nonparametric Control Chart for Monitoring Profiles Using Change Point Formulation ad Adaptive Smoothing. *Statistica Sinica* , 1-24.

Zou, C., Wang, Z., & Tsung, F. (2007, Novembre). Monitoring General Linear Profiles Using Multivariate Exponentially Weighted Moving Average Schemes. *Technometrics* , 49 (4), p. 395-408.

Zou, C., Wang, Z., & Tsung, F. (2008, Novembre). Monitoring Profiles Based on Nonparametric Regression Methods. *Technometrics* , 50 (4), p. 512-526.

Zou, C., Zhang, Y., & Wang, Z. (2006). A Control Chart Based on a Change-Point model for Monitoring Linear Profiles. *IIE Transaction* , 1093-1103.