



UNIVERSIDAD NACIONAL DEL SUR

TESIS DE DOCTOR EN INGENIERIA QUIMICA

**Monitoreo Estadístico de Procesos Batch.
Aplicaciones a Reactores de Polimerización**

Carlos Rodrigo Alvarez Medina

BAHIA BLANCA

ARGENTINA

2009

Página en Blanco

Página en Blanco

Página en Blanco



UNIVERSIDAD NACIONAL DEL SUR

TESIS DE DOCTOR EN INGENIERIA QUIMICA

**Monitoreo Estadístico de Procesos Batch.
Aplicaciones a Reactores de Polimerización**

Carlos Rodrigo Alvarez Medina

BAHIA BLANCA

ARGENTINA

2009

Página en Blanco

PREFACIO

Esta Tesis se presenta como parte de los requisitos para optar al grado Académico de Doctor en Ingeniería Química, de la Universidad Nacional del Sur y no ha sido presentada previamente para la obtención de otro título en esta Universidad u otras. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en la Planta Piloto de Ingeniería Química, dependiente del Departamento de Ingeniería Química de la Universidad Nacional del Sur y del CONICET, durante el período comprendido entre el 29 de Junio de 2004 y el 20 de Febrero de 2009, bajo la dirección de las Dras. Mabel Cristina Sánchez y Adriana Brandolin, ambas profesoras del Departamento de Ingeniería Química de la Universidad Nacional del Sur.

Carlos Rodrigo Alvarez Medina



UNIVERSIDAD NACIONAL DEL SUR
Secretaría General de Posgrado y Educación Continua

La presente tesis ha sido aprobada el / /, mereciendo
la calificación de(.....)

Página en Blanco

Agradecimientos

Muchas son las personas que tuvieron que ver, de una u otra forma, con que este trabajo llegara finalmente a buen puerto. No quisiera olvidarme de nadie, y un “*Muchas Gracias a Todos*”, aunque muy vago, es lo más justo que se me ocurre.

Quiero, sin embargo, mencionar especialmente a los principales artífices de este logro. En primer lugar agradezco enormemente a mis directoras, Mabel Sánchez y Adriana Brandolin por la guía, el apoyo y la colaboración que me dieron en todo momento durante estos cinco años en los que hemos estado trabajando juntos.

A Carolina, por todos los años de amor y compañía, por su paciencia e infinita generosidad para conmigo y por estar siempre a mi lado, tanto en los buenos momentos como en los otros.

A toda mi familia, que más allá de la distancia supo siempre hacerse sentir cerca. En especial a mis viejos y hermanos.

A mis compañeros de Plapiqui por todos y cada uno de los momentos vividos, desde los primeros campamentos hasta los partidos de basket y los “mortales” duelos de paddle. A toda la gente de Polímeros y Procesos por haber compartido conmigo gran parte del camino.

A los amigos que fui haciendo por la vida y no quiero dejar de recordar, Aaron, Adrián, Aníbal, Axel, Ceci P., Coti, Diego, Eli F., Eli V., Erica B., Erica S., Fede, Fernando, Guille, Gri, Hernán, José Miguel, Leo, Luis, Lore L., Lore M., Lucas, Mana, María, Martín, Marian, Maxi, Milton, Pablo H., Pablo M., Paola y Vani (cualquier semejanza con el orden alfabético es pura automatización).

Finalmente, quiero agradecer a CONICET por el apoyo económico que hizo posible este trabajo.

Página en Blanco

A mis Viejos

*Que primero me dieron la vida
y luego me dedicaron la suya.*

Página en Blanco

Resumen

Esta tesis comprende el desarrollo e implementación de nuevas metodologías de control estadístico multivariable (*MSPC*), su aplicación para monitorear la producción de procesos de polimerización en emulsión discontinuos (*batch*) y la comparación del desempeño de diferentes estrategias dedicadas al monitoreo estadístico de procesos *batch*.

Los procesos de polimerización presentan características que los hacen atractivos como casos de estudio para evaluar las técnicas de *MSPC*. Dada la complejidad de estos sistemas, y la gran cantidad de fuentes de variación que los afectan, la implementación de técnicas de monitoreo y control basadas en datos resulta una alternativa atractiva y realizable.

Los procedimientos de *MSPC* se componen de tres etapas fundamentales: detección, identificación y diagnóstico. El estado del sistema se monitorea en forma continua para evaluar si el mismo se encuentra operando en condiciones normales. Si se detecta un evento anormal, resulta necesario identificar las variables que señalan esta condición y luego diagnosticar la causa primera de la anomalía. Los estudios y desarrollos de esta tesis comprenden a las etapas de detección e identificación de fallas.

Se desarrolló una nueva metodología, denominada *OSS (Original Space Strategy)*, para la descomposición del estadístico de Hotelling en el espacio de las mediciones, que permite evaluar la influencia que tiene cada variable en el valor de dicho estadístico. Las principales ventajas de la estrategia de identificación propuesta son las siguientes: permite monitorear el proceso usando un sólo estadístico, reduce significativamente las ambigüedades en la identificación de fallas propias de otras técnicas existentes, evita la posible pérdida de información originada por la proyección de los datos en un espacio de variables latentes de dimensión incorrecta, proporciona una clara comprensión del significado físico de las

contribuciones negativas al estadístico y determina un valor límite para las mismas. Se propone la incorporación de la metodología *OSS* como herramienta de identificación en el nuevo procedimiento de monitoreo de procesos *batch* propuesto.

En esta tesis se analizaron en detalle los métodos más utilizados para el monitoreo de procesos *batch*. Los mismos se basan en técnicas de proyección tales como el Análisis de Componentes Principales (*PCA*) y el Análisis de Componentes Independientes (*ICA*). Se comparó el desempeño de estas estrategias con la propuesta en esta tesis, mediante su aplicación a un reactor de polimerización. El conjunto de datos analizados se obtuvo por simulación usando un modelo riguroso de un reactor de polimerización en emulsión de metacrilato de metilo. La comparación del desempeño se basó en los resultados obtenidos mediante la aplicación de los procedimientos a una serie de fallas conocidas. Los mismos se compararon en términos de: la capacidad de cada técnica para detectar la existencia de la falla y la exactitud con la que la metodología indica las variables sospechosas durante la etapa de identificación.

Los resultados de la aplicación de la estrategia propuesta, caracterizada por utilizar un solo estadístico, a este complejo caso de estudio muestran un muy buen desempeño de la misma, en lo referente a su velocidad de detección y exactitud en la identificación, cuando se lo compara con el obtenido usando *PCA* o *ICA*.

Abstract

This thesis involves the development and implementation of new methodologies for Multivariate Statistical Process Control (*MSPC*), their application to monitor the production of batch emulsion polymerization processes and the performance comparison among different techniques devoted to the statistical monitoring of batch processes.

Polymerization processes have distinctive features that make them attractive as case studies to evaluate *MSPC* techniques. Due to the inherent complexity of these systems and the huge amount of variation sources that affect them, the implementation of data-driven monitoring and control techniques arises as an attractive and feasible alternative.

The *MSPC* procedures are made up of three essential stages: detection, identification and diagnosis. The process state is evaluated continuously to determine if its operation is normal. If an abnormal event is detected, it is necessary to identify the variables that signal this condition, and then to diagnosis the first source of the anomaly. The studies and developments of this thesis involve the fault detection and identification stages.

A new methodology, called *OSS* (Original Space Strategy), is developed to decompose the Hotelling's statistic that allows evaluating the influence of each variable on the statistic value. The main advantages of the proposed identification strategy are: process monitoring is accomplished using only one statistic, it significantly reduces the ambiguities in fault identification inherent to other existing techniques, it avoids the possible loss of information that may arise when data are projected into a latent variable space of inappropriate dimension, it provides a clear understanding of the physical meaning of negative contribution to the statistic value and determines a limit for them. It should be highlighted the incorporation of the *OSS* methodology as identification tool for the new proposed monitoring procedure.

In this thesis the most used methods for batch process monitoring are analyzed in detail. They are based on projection techniques, such as, Principal Component Analysis (*PCA*) and Independent Component Analysis (*ICA*). A performance comparison is conducted among these strategies and the proposed one for their application to a polymerization reactor. The data set was obtained by simulation using a rigorous model of an emulsion polymerization reactor for the production of methylmetacrylate. The performance comparison is based on the results obtained from the application of the procedures to a set of known faults. They are compared in terms of the capacity of each technique to detect the existing faults and the accuracy of the methodology to indicate the suspicious variables during the identification stage.

The application results of the proposed strategy, which is characterized by the use of only one statistic, to this complex case study show that it has a good performance regarding its detection velocity and identification accuracy, in comparison with methods base on *PCA* o *ICA*.

Índice General

1. Introducción	3
1.1 Objetivos y Motivación	3
1.2 Revisión Bibliográfica – Estado del Arte	5
1.2.1 Modelado	9
1.2.2 Detección	15
1.2.3 Identificación	16
1.3 Organización de la Tesis	18
1.4 Referencias.....	20
2. Monitoreo Estadístico de Procesos utilizando Estrategias de Proyección	27
2.1 Introducción	27
2.2 Monitoreo Basado en el Análisis de Componentes Principales	33
2.2.1 Análisis de Componentes Principales (PCA)	33
2.2.2 Implementación de la Estrategia de Monitoreo	41
2.2.2.1 Etapa I.....	41
2.2.2.2 Etapa II.....	49
2.3 Monitoreo Basado en el Análisis de Componentes Independientes	55
2.3.1 Análisis de Componentes Independientes (ICA).....	55
2.3.2 Implementación de la Estrategia de Monitoreo	62
2.3.2.1 Etapa I.....	63
2.3.2.2 Etapa II.....	64
2.4 Monitoreo Basado en Cuadrados Mínimos Parciales	65
2.4.1 Cuadrados Mínimos Parciales (PLS).....	65
2.4.2 Implementación de la Estrategia de Monitoreo	67
2.5 Referencias.....	68
3. Monitoreo de Procesos Batch.....	73
3.1 Procesos <i>Batch</i> y Estructura de los Datos.....	74
3.2 Análisis Paralelo de Factores, Modelos <i>Tucker</i> y <i>MPCA</i>	77
3.3 Análisis Multimodo: <i>MPCA</i> , <i>MICA</i> , <i>MPLS</i>	79
3.3.1 Método de <i>Nomikos</i> y <i>MacGregor (NM)</i>	80
3.3.2 Método de <i>Wold</i> , <i>Kettaneh</i> , <i>Fridén</i> , y <i>Holmberg Original (WKFH)</i> y <i>Modificado (AT)</i>	87
3.3.3 Modelos locales (<i>ML</i>)	92
3.3.4 Modelos Evolutivos (<i>ME</i>).....	95

3.3.5	<i>Modelos de Ventana Móvil (VM)</i>	99
3.3.6	<i>Análisis Dinámicos (DPCA y DPLS)</i>	104
3.4	Batches con Diferente Número de Observaciones	105
3.5	Referencias	107
4.	Identificación de Situaciones Anómalas en el Espacio Original de las Mediciones..	113
4.1	Método de Mason, Young y Tracy (<i>MYT</i>)	114
4.2	Nuevo Método de Descomposición del Estadístico de Hotelling: <i>OSS (Original Space Strategy)</i>	116
4.3	Equivalencia entre <i>OSS</i> y los Gráficos de Contribución Generalizados	122
4.4	Casos de Estudio	124
4.4.1	<i>Ejemplo Numérico</i>	124
4.4.2	<i>Reactor Tubular</i>	129
4.5	Nueva Estrategia de Monitoreo de Procesos <i>Batch</i> Basada en <i>OSS</i>	133
4.5.1	<i>Etapa I</i>	133
4.5.2	<i>Etapa II</i>	134
4.6	Consideraciones generales	135
4.7	Referencias	137
5.	Descripción del Caso de Estudio	141
5.1	Reactor de Polimerización de Metacrilato de Metilo (<i>MMA</i>) en Emulsión	141
5.1.1	<i>Componentes de los Sistemas de Polimerización en Emulsión</i>	142
5.1.2	<i>Modelo Matemático</i>	145
5.1.2.1	Mecanismo de Reacción	145
5.1.2.2	Descripción del Reactor	149
5.1.2.3	Balances de Masa y Energía	150
5.1.3	<i>Implementación y Validación del Modelo</i>	163
5.1.4	<i>Generación de la Población de Referencia</i>	163
5.2	Referencias	165
6	Resultados y Discusión.....	169
6.1	Generación de Datos	169
6.1.1	<i>Cambios en el caudal de alimentación de monómero</i>	171
6.1.2	<i>Cambios en la temperatura de entrada del refrigerante</i>	172
6.1.3	<i>Variaciones en las condiciones iniciales</i>	174
6.2	Base de Comparación entre las Distintas Técnicas	174
6.3	Comparaciones entre <i>NM</i> , <i>WKFH-AT</i> , <i>OSS</i> , <i>VM</i> y <i>ME</i>	177
6.4	Resultados del Análisis de Componentes Independientes (<i>MICA</i>)	199
6.5	Referencias	208

7	Conclusiones y Trabajos Futuros.....	213
7.1	Conclusiones	213
7.2	Trabajos Futuros.....	215
	Apéndice 1	219
A1.1	Determinación del Número de Componentes a Retener	219
A1.2	Referencias.....	226
	Apéndice 2	227
A2.1	Test de Hipótesis	227
A2.2	Errores <i>Tipo I</i> y <i>Tipo II</i>	228
	Apéndice 3	231
A3.1	Estimación de Funciones de Densidad de Probabilidad. Método del Kernel (<i>KDE</i>).....	231
A3.2	Referencias.....	235
	Apéndice 4	237
A4.1	Modelo Matemático del Reactor Tubular Empleado en el Capítulo 4	237
	Apéndice 5	241
A5.1	Nomenclatura del Caso de Estudio.....	241
A5.2	Parámetros del Modelo.....	246
A5.3	Condiciones iniciales de los Batches de la Población de Referencia (<i>PR</i>).....	247
A5.4	Ruido de Medición	252
A5.5	Referencias.....	254
	Nomenclatura	255

Página en Blanco

Índice de Tablas

Tabla 3.1: Diferentes desdoblamientos posibles para el arreglo tridimensional $\underline{\mathbf{X}}$	80
Tabla 3.2: Pasos requeridos en la etapa de modelado fuera de línea (<i>Etapa I</i>)	81
Tabla 3.3: Modelos <i>MPCA</i> y <i>MICA</i> (desdoblamiento <i>batch-wise</i>).....	85
Tabla 3.4: Estadísticos y límites de control para los modelos <i>MPCA</i> y <i>MICA</i> (desdoblamiento <i>batch-wise</i>).....	86
Tabla 3.5: Cálculo de las <i>CVs</i> para <i>MPCA</i> y <i>MICA</i> (desdoblamiento <i>batch-wise</i>).....	87
Tabla 3.6: Modelos <i>MPCA</i> y <i>MICA</i> (desdoblamiento <i>variable-wise</i>)	93
Tabla 3.7: Estadísticos y límites de control para los modelos <i>MPCA</i> y <i>MICA</i> (desdoblamiento <i>variable-wise</i>)	94
Tabla 3.8: Cálculo de las <i>CVs</i> para <i>MPCA</i> y <i>MICA</i> (desdoblamiento <i>variable-wise</i>).....	95
Tabla 3.9: Modelos <i>MPCA</i> y <i>MICA</i> - Modelos Evolutivos	97
Tabla 3.10: Estadísticos y límites de control para los modelos <i>MPCA</i> y <i>MICA</i> - Modelos Evolutivos	98
Tabla 3.11: Cálculo de las <i>CVs</i> para <i>MPCA</i> y <i>MICA</i> - Modelos Evolutivos.....	99
Tabla 3.12: Modelos <i>MPCA</i> y <i>MICA</i> - Modelos de <i>VM</i>	102
Tabla 3.13: Estadísticos y límites de control para los modelos <i>MPCA</i> y <i>MICA</i> - Modelos de <i>VM</i>	103
Tabla 3.14: Cálculo de las <i>CVs</i> para <i>MPCA</i> y <i>MICA</i> - Modelos de <i>VM</i>	104
Tabla 4.1: Población de referencia (De Maesschalck y col., 2000)	125
Tabla 4.2: Observaciones defectuosas. Tests de prueba	125
Tabla 4.3: Valores de D para las observaciones de prueba	127
Tabla 4.4: Valores de c_i^D y $c_i^{T^2}$ para las observaciones de prueba.....	130
Tabla 4.5: Valores de $c_i^{T^2}$ para las observaciones A_1 a A_4	131
Tabla 4.6: Valores de c_i^D ($R = 3$) para las observaciones A_1 a A_4	132

Tabla 4.7: Valores de c_i^D ($R = 2$) para las observaciones A_1 a A_4	132
Tabla 6.1: Valores de Q_0 para las fallas 1 a 5.....	171
Tabla 6.2: Valores de los escalones aplicados a Q_0	172
Tabla 6.3: Valores de T_{C0} para las fallas 24 a 28.....	173
Tabla 6.4: Valores y posiciones de las pendientes en las rampas de T_{C0} para las fallas 29 a 46	173
Tabla 6.5: Condiciones iniciales para las fallas 47 a 82	174
Tabla 6.6: Resultados obtenidos con los métodos: <i>NM</i> , <i>WKFH</i> , <i>OSS</i> , <i>VM</i> y <i>EM</i>	184
Tabla 6.7: Resultados de <i>MICA</i>	206
Tabla A1.1: Datos de un estudio de audiometría.....	225
Tabla A1.2: Valores de R sugeridos por los diferentes criterios para los datos de la Tabla A1.1	225
Tabla A3.1: Distribuciones usadas como kernel en <i>KDE</i>	232
Tabla A3.2: Algoritmo de cálculo de $\hat{f}(x)$ usando <i>KDE</i>	235
Tabla A4.1: Parámetros del caso de estudio.....	239
Tabla A4.2: Variables de entrada para la simulación de las observaciones que constituyen la <i>NOC</i>	240
Tabla A5.1: Nomenclatura. Caso de estudio presentado en el Capítulo 5	241
Tabla A5.2: Valores de los parámetros del modelo.....	246
Tabla A5.3: Condiciones iniciales para los 146 <i>batches</i> utilizados para construir la <i>PR</i>	247
Tabla A5.4: Trayectorias de q_m^0 , T_{C0} y T^0 consideradas en la <i>PR</i>	251

Índice de Figuras

Figura 2.1: Esquema de la relación entre la calidad final del producto y las posibles fuentes de variación	30
Figura 2.2: Realizaciones de un experimento con dos variables medidas.....	35
Figura 2.3: Dependencia de la eliminación de los individuos con el orden de selección	43
Figura 2.4: Esquema del tipo de fallas detectadas por D y SPE	49
Figura 2.5: Representación esquemática de la metodología de monitoreo en línea	50
Figura 2.6: Comparación ICA vs. PCA para el caso de dos señales uniformemente distribuidas	56
Figura 3.1: Trayectoria de las J variables monitoreadas en el proceso desde $k=1$ a $k=K$	75
Figura 3.2: Estructura de las matrices de datos $\underline{\mathbf{X}} \in \mathbb{R}^{I \times J \times K}$	76
Figura 3.3: Disposición de los datos en los arreglos tridimensionales	77
Figura 3.4: Representación gráfica del modelo $PARAFAC$	78
Figura 3.5: Estrategia de desdoblamiento usada en $MPCA$ (desdoblamiento <i>batch-wise</i>)	79
Figura 3.6: Desdoblamiento BW y estandarización de los datos para el enfoque NM	82
Figura 3.7: Esquema de desdoblamiento y escalado de $\underline{\mathbf{X}}$ según $WKFH$	89
Figura 3.8: Dependencia de las tendencias remanentes en $\underline{\mathbf{X}}$ de acuerdo al tipo de escalado	90
Figura 3.9: Esquema de desdoblamiento y escalado de $\underline{\mathbf{X}}$ según AT	91
Figura 3.10: Esquema de desdoblamiento de $\underline{\mathbf{X}}$ para los ME	96
Figura 3.11: Esquema de desdoblamiento y escalado de $\underline{\mathbf{X}}$ para el enfoque de VM	101
Figura 4.1: Límites para x_1 y x_2 considerando sus distribuciones condicionales.....	114
Figura 4.2: Superficies de $T^2(x_1, x_2)$ y $T^2(x_2) \Big _{x_1=cte}$	118
Figura 4.3: $T^2 = T^2(x_k)$ dados $x_j, \forall j \neq k$	122
Figura 4.4: Reconstrucción de la varianza total:.....	126

Figura 4.5: Reconstrucción de la varianza de cada variable:.....	127
Figura 4.6: Reconstrucción de la varianza total (matriz de correlación)	132
Figura 5.1: Esquema de un sistema de polimerización en emulsión.....	146
Figura 5.2: Esquema del reactor	150
Figura 5.3: Conversión vs. tiempo.(Sajjadi y Yianneskis, 2003).	164
Figura 5.4: Peso molecular promedio en peso vs. tiempo (Sajjadi y Yianneskis, 2003).....	164
Figura 5.5: Conversión vs. tiempo. (Gao y Penlidis 2002).....	164
Figura 6.1: Trayectoria promedio de las J variables monitoreadas	170
Figura 6.2: Trayectoria temporal de las J variables monitoreadas para los I batches de la NOC	171
Figura 6.3: Perfil de D y valores de $c_{j,k}^D$, falla: 10, método: NM , $k=19$	179
Figura 6.4: Perfil de SPE y valores de $c_{j,k}^{SPE}$, falla: 10, método: NM , $k=19$	180
Figura 6.5: Perfil de D y valores de $c_{j,k}^D$, falla: 10, método: $WKFH-AT$, $k=19$	181
Figura 6.6: Perfil de SPE y valores de $c_{j,k}^{SPE}$ falla: 10, método: $WKFH-AT$, $k=19$	182
Figura 6.7: Perfil de T^2 y valores de $c_{j,k}^{T^2}$, falla: 10, método: OSS , $k=19$	183
Figura 6.8: Perfil de SPE para la falla 13 obtenido con NM	195
Figura 6.9: Perfil de SPE para la falla 13 (sin ruido) obtenido con NM	195
Figura 6.10: Gráfico cuantil-cuantil \mathcal{F}_{v_1, v_2} vs D	201
Figura 6.11: Gráfico cuantil-cuantil χ^2 vs SPE	201
Figura 6.12: Perfil de I_d y valores de $c_{j,k}^{I_d}$, falla: 10, método: $MICA$, $k=19$	203
Figura 6.13: Perfil de I_e y valores de $c_{j,k}^{I_e}$, falla: 10, método: $MICA$, $k=19$	204
Figura 6.14: Perfil de SPE y valores de $c_{j,k}^{SPE}$, falla: 10, método: $MICA$, $k=19$	205
Figura A1.1: Gráfico típico de autovalores de una matriz de covarianza	220
Figura A2.1: Interpretación gráfica de un test de hipótesis	229
Figura A3.1: Estimación de $f(x)$ basada en un conjunto de datos $\mathbf{x} \approx \mathcal{N}(0;1)$	233

Figura A3.2: Estimaciones de $f(x)$ usando distintos valores de h	234
Figura A4.1: Esquema del reactor considerado como caso de estudio	237

Página en Blanco

Capítulo 1

Introducción

1.1	Objetivos y Motivación	3
1.2	Revisión Bibliográfica – Estado del Arte.....	5
1.2.1	<i>Modelado</i>	9
1.2.2	<i>Detección</i>	15
1.2.3	<i>Identificación</i>	16
1.3	Organización de la Tesis.....	18
1.4	Referencias	20

Página en Blanco

1 Introducción

1.1 Objetivos y Motivación

El objetivo de este trabajo de tesis es estudiar, desarrollar e implementar metodologías de control estadístico multivariable para procesos de polimerización en emulsión discontinuos y semi-continuos.

Las estrategias de “control estadístico de procesos” (*SPC*) también se conocen con el nombre de técnicas de “monitoreo estadístico”, debido a que se utilizan para determinar y monitorear el nivel de control estadístico de un proceso productivo. Se considera que éste no se encuentra en control estadístico si se detectan desviaciones respecto de la condición normal o habitual, originadas por causas especiales. La identificación de las variables mediante las cuales se manifiesta el cambio permite tomar decisiones conducentes a la resolución del problema.

El proceso de polimerización en emulsión se utiliza en la fabricación de distintos tipos de productos, por ejemplo: adhesivos, látex para pinturas, recubrimientos para papeles y plásticos, pigmentos para tintas y recubrimientos para pisos, entre otros. Por su naturaleza heterogénea, este tipo de proceso presenta ciertas ventajas, tales como: ser relativamente fácil de manipular, permitir operar a bajas viscosidades y, por ende, ofrecer excelentes propiedades para la transferencia de calor. Dadas estas características es posible conseguir velocidades de reacción elevadas, obteniendo a la vez productos de alto peso molecular y con propiedades que no pueden lograrse con otros tipos de procesos de polimerización (Asua, 2001).

La polimerización en emulsión se puede llevar a cabo tanto en reactores discontinuos (*batch*) y semi-continuos (*semibatch*) como en equipos de operación continua. Los dos primeros se utilizan principalmente en producciones de poco volumen y se emplean para la

síntesis de polímeros de alto valor agregado, cuyas propiedades se diseñan en función de los requerimientos de cada cliente. En estos casos es crucial mantener dichas propiedades dentro de límites estrictos de calidad.

Dada la complejidad de este tipo de proceso, mantener en el tiempo las propiedades del producto en los valores deseados constituye un importante desafío. El control del proceso se dificulta debido a la elevada correlación que existe entre las variables que influyen sobre la velocidad de polimerización y las propiedades del producto, tales como las distribuciones de pesos moleculares y tamaños de partículas (Hamielec y Penlidis, 1995). Por ejemplo, se ha observado que la velocidad de polimerización depende fuertemente, entre otros factores, del número de partículas y concentración de monómero en las mismas, de la concentración de surfactante, e indirectamente de la concentración de iniciador. La presencia de impurezas también afecta el curso de la reacción. En relación a la calidad del producto, se sabe que la temperatura tiene gran influencia sobre los pesos moleculares. Por otra parte, las concentraciones de iniciador, surfactante e impurezas afectan la distribución de tamaños de partículas. Si las impurezas actúan como agentes de transferencia de cadena también pueden llevar a una disminución del peso molecular.

El control de calidad en los procesos de polimerización es un aspecto que aún no se ha resuelto en forma absoluta. En el caso de los reactores *batch* y *semibatch*, el control de la calidad final del producto resulta complejo ya que su comportamiento dinámico es altamente no lineal al igual que su cinética (Yabuki y col., 2002). Además, rara vez se dispone de sensores en línea para monitorear la evolución de la calidad durante el *batch*. En general, estos reactores se operan a lazo abierto respecto de las variables que definen la calidad y se automatizan para lograr una gran reproducibilidad en las condiciones de reacción. A pesar de dicha automatización, pueden producirse perturbaciones inesperadas debido, por ejemplo, a la

presencia imprevisible de impurezas en las materias primas (Zhang *y col.*, 1999). Por otra parte, la presencia de errores en las mediciones dificulta aún más el control del proceso.

A fin de monitorear la operación en línea de reactores de polimerización se han utilizado técnicas de control estadístico multivariable de procesos (*MSPC*). Uno de los principales atractivos de estas metodologías es que el monitoreo estadístico se efectúa utilizando mediciones en línea del proceso, en lugar de emplear mediciones de calidad fuera de línea. Esto brinda la posibilidad de realizar alguna acción correctiva durante el transcurso del batch.

La mayoría de las aplicaciones de *MSPC* a procesos *batch* se basan en el Análisis de Componentes Principales y en otras técnicas de Proyección en Estructuras Latentes (Nomikos y MacGregor, 1994). El desarrollo de estrategias eficientes de *MSPC* para un sistema de reacción de la complejidad descrita previamente resulta un desafío tanto académico como tecnológico.

La siguiente sección presenta un análisis detallado del estado del arte respecto del *MSPC*, que en esta tesis constituye el fundamento para el desarrollo de nuevas metodologías de monitoreo aplicables a procesos de polimerización en emulsión.

1.2 Revisión Bibliográfica – Estado del Arte

En esta sección se describe el estado del arte en lo referente al *SPC* resaltando las principales ideas y metodologías que han sido propuestas en esta área o incorporadas desde otras. Dado el alcance de esta tesis, el interés está puesto especialmente en los procesos *batch* y *semibatch*. Sin embargo, se hace también referencia a algunas de las técnicas aplicadas a los procesos continuos, dado que la mayoría de los conceptos básicos de *SPC* se han desarrollado y probado principalmente para los mismos.

En el área de la Quimiometría, de la cual el *SPC* puede considerarse parte, el concepto de *Datos* y el de *Información*, que en el lenguaje coloquial resultan equivalentes, tienen una gran diferencia. En esta tesis, se llama *Datos* a los signos usados (números, palabras, etc.) para describir sucesos y/o entidades. Sin importar la cantidad de datos con la que se cuente, por sí mismos éstos no pueden comunicar significado alguno. Su utilidad, y por lo tanto su importancia, está íntimamente ligada a la capacidad del analista de asociarlos en un contexto que le permita convertirlos en *Información*. Esta última, debe ser entendida como un conjunto de inferencias o conclusiones que se obtienen del análisis de los datos y que resultan significativas y pertinentes para el individuo que los analiza.

En las últimas décadas se ha observado un continuo avance en la tecnología relacionada con las mediciones, como así también una marcada disminución en los costos de la instrumentación de procesos. En consecuencia, la cantidad disponible de datos de operación de los procesos industriales se ha incrementado a tal punto que, en la mayoría de los casos, los datos se tornan imposibles de procesar y administrar de manera eficiente.

Diversos autores reconocen el potencial de conocimiento que contienen los datos de operación de un proceso. Disponer del mismo, permitiría intervenir en forma eficaz en el proceso para lograr su operación óptima durante la elaboración de un dado producto (Jaekle y MacGregor, 1998; Sebzalli y Wang, 2001). Esos mismos autores también resaltan las limitaciones y complicaciones asociadas a la obtención de dicha información.

La gran cantidad de datos disponibles, referente a casi cualquier actividad humana, presenta, hoy en día, un reto muy importante. La necesidad de extraer conocimiento e información útil a partir de estos datos no es exclusiva de la ingeniería de procesos. Actualmente, los problemas asociados a la obtención de información válida y confiable a partir de enormes bases de datos son abordados desde distintas áreas y puntos de vista. La

sociología, la administración de negocios y la industria son sólo algunas de las áreas que podrían mencionarse (Fayyad *y col.*, 1996; Goebel y Gruenwald, 1999; Apte *y col.*, 2002).

En lo que respecta a la ingeniería de procesos, se adoptaron y/o extendieron diversas técnicas provenientes de otros campos de la ciencia. Entre las más importantes pueden destacarse las que utilizan enfoques basados en principios fundamentales (Willisky, 1976; Basseville, 1988; Gertler, 1993) y las llamadas técnicas basadas en datos (*Data Driven Techniques*). Estas últimas se sustentan en la información del proceso contenida en las bases de datos de operación (Jackson, 1991; Dunia *y col.*, 1995; Dunia *y col.*, 1996; Kourti *y col.*, 1996; Kourti y MacGregor, 1996; Alcorta García y Frank, 1997; Jaeckle y MacGregor, 1998; 2000; Hyvärinen, 2000; Hyvärinen *y col.*, 2001; etc.).

Los enfoques que emplean modelos de principios fundamentales realizan una descripción detallada del proceso en cuestión y de la posible falla (o conjunto de fallas) para poder identificar la naturaleza y magnitud de la misma. Si bien estos resultados son de gran utilidad, dado que permiten diagnosticar completamente la falla, el esfuerzo requerido para modelar apropiadamente el sistema limita su aplicación práctica a problemas de pequeña y/o mediana escala.

En el caso de sistemas complejos o de gran escala, el problema de identificación de fallas mediante modelos basados en principios fundamentales resulta una opción poco viable puesto que, además de lograr una buena representación del sistema en estudio, dichos modelos deben estar adecuadamente implementados y optimizados desde el punto de vista numérico. Esta actividad no resulta trivial y demanda gran cantidad de tiempo y esfuerzo. Teniendo en cuenta estas limitaciones, numerosos autores atacaron el problema desde el punto de vista estadístico. El enfoque estadístico presenta una alternativa factible con un respaldo teórico robusto y

confiable dada la gran cantidad de técnicas propuestas, mejoradas y comprobadas en otras áreas de la ciencia para la resolución de problemas similares.

Las primeras aplicaciones de control estadístico en la industria de procesos hacían uso de las tradicionales cartas de Shewhart, *CUSUM* ó *EWMA* para monitorear, en forma individual, un conjunto de variables claves relacionadas con la calidad del producto y con la seguridad del proceso. El objetivo era detectar la ocurrencia de algún evento que pudiese estar alterando dichas variables, permitiendo así realizar acciones correctivas y minimizar su efecto. Sin embargo, estas estrategias no resultan adecuadas para la industria de procesos moderna donde se miden cientos o miles de variables, altamente correlacionadas, con una frecuencia del orden de los minutos o incluso segundos. Con el fin de considerar la naturaleza multivariada de los procesos industriales y comprender su efecto sobre el comportamiento de los mismos, se propusieron diferentes metodologías de *MSPC* durante las últimas dos décadas.

La implementación de una estrategia de *MSPC* se compone de dos etapas. En la primera se utilizan los datos de la población de referencia (*PR*) para formular el modelo que la representa y estimar los valores críticos de los estadísticos empleados para la detección de eventos anormales. La segunda etapa comprende las tareas de detección e identificación y se ejecuta en línea. Cada vez que se obtiene una nueva medición del proceso, se calculan los valores de los estadísticos. Si éstos resultan mayores que los valores límites, se detecta una situación anormal del proceso y luego se identifica(n) la(s) variable(s) mediante la(s) cual(es) se manifiesta la misma.

A continuación se analizan las contribuciones más importantes reportadas en la literatura sobre *MSPC* en relación con los aspectos de modelado, detección e identificación.

1.2.1 Modelado

Jackson (1980) fue uno de los primeros en proponer el uso de técnicas de proyección en estructuras latentes, tales como el Análisis de Componentes Principales (*PCA*) o Cuadrados Mínimos Parciales (*PLS*), con el fin de obtener un modelo del proceso a partir de los datos de la *PR*. La posibilidad de lograr una importante reducción en la dimensionalidad del espacio de descriptores, manteniendo la mayor parte de la información significativa sobre el sistema, es la característica fundamental que motivó a los investigadores de diversas áreas a utilizar estas herramientas en una gran variedad de aplicaciones. Kourti y MacGregor (1995) publicaron una revisión sobre los conceptos básicos de las técnicas *PCA* y *PLS* desde el punto de vista del monitoreo de procesos, tanto para el caso de procesos continuos como *semibatch* o *batch*.

Varios investigadores extendieron la metodología clásica del *PCA* a fin de considerar otros aspectos en el modelado. En tal sentido, Wangen y Kowalski (1988) introdujeron el concepto de *Multiblock PCA* que permite analizar diferentes bloques de variables que son medidas en el mismo proceso. De esta manera, las variables medidas en diferentes equipos o subsistemas pueden ser clasificadas como un bloque separado. En la misma línea Westerhuis y col. (1998) propusieron otro algoritmo denominado *PCA* jerárquico (*HPCA*). Por su parte, Qin y col. (2001) desarrollaron los métodos *Consensus PCA* y *Multiblock PLS*, en los cuales se presentan metodologías para el cálculo de las contribuciones de las variables y los bloques a las métricas utilizadas en el monitoreo de sistemas descentralizados.

El cambio paulatino del comportamiento del proceso durante la operación (debido a ensuciamiento de intercambiadores, agotamiento de catalizadores, etc.) fue estudiado por varios autores. Wold (1994) implementó una versión adaptiva de *PCA* combinando filtrado exponencial (*MEWMA*) con *PCA* o *PLS*. Por su parte, Li y col. (2000) utilizaron un enfoque alternativo que consiste en actualizar las matrices de correlación que definen los modelos del

PCA a medida que se van obteniendo nuevas observaciones. Además, señalaron que las metodologías propuestas años antes por Wold (1994) y Rännar y col. (1991) no constituían un enfoque estrictamente recursivo.

En los últimos años se han presentado diversos métodos tendientes a incorporar un nivel de detalle cada vez mayor en los modelos estadísticos utilizados. Además, se están realizando muchos esfuerzos orientados hacia la solución de los problemas que surgen durante la implementación de estas técnicas en procesos industriales reales. Por ejemplo, AlGhazzawi y Lennox (2008) presentaron un caso de aplicación industrial usando *Recursive PCA* y *Multiblock PCA*, en el cual se analizan las dificultades que surgen en la interpretación de los resultados y las cartas de control debido a la falta de entrenamiento de los operadores (no expertos en *SPC*). Además, discutieron los problemas asociados a la determinación de las regiones en las cuales el sistema de monitoreo no se comporta de la forma esperada. Una conclusión importante es que las técnicas de *SPC* se encuentran todavía en etapa de desarrollo y no son aún lo suficientemente robustas para su aplicación masiva en la industria de procesos.

En el caso particular de los procesos *semibatch* o *batch*, surgen complicaciones adicionales y bien diferenciadas de las inherentes a los sistemas continuos en la etapa de modelado. A continuación se presenta una revisión de las publicaciones más significativas que proponen soluciones a esta problemática.

El comportamiento multivariado y dependiente del tiempo de los sistemas *batch* hace que la estructura natural de los arreglos de datos sea tridimensional. Esto es así, ya que para realizar el análisis estadístico se debe disponer de datos correspondientes a I *batches*, que constituyen la denominada *PR* a partir de la cual se construirá el modelo estadístico. Estos

datos se obtienen realizando un conjunto de K observaciones de J variables para cada *batch*. Este conjunto de datos se ordena en un arreglo tridimensional $\underline{\mathbf{X}} \in \mathbb{R}^{I \times J \times K}$.

Arreglos como el descrito, se conocen como estructuras *Multiway* o multimodo, y fueron estudiados en otras áreas del conocimiento mucho tiempo antes que los ingenieros de procesos pusieran su interés en los mismos (Kroonenberg y de Leeuw, 1980; Kroonenberg, 1983, 1989).

Wold y col. (1987) demostraron la equivalencia entre las metodologías *Multiway* para *PCA* o *PLS* y el análisis *PCA* clásico, cuando este último se lleva a cabo sobre una matriz de datos que resulta del desdoblamiento del arreglo tridimensional en uno de dos dimensiones. Uno de los primeros grupos en aplicar estos resultados al problema de monitoreo de procesos *batch* fue el liderado por John F. MacGregor en *McMaster University*, Ontario, Canadá. Específicamente, Nomikos y MacGregor (1994) presentaron un trabajo donde se propone el uso de esta estrategia para el monitoreo de procesos discontinuos usando *PCA*. Este trabajo pionero es todavía utilizado como referencia por muchos autores para evaluar el desempeño de sus propias técnicas. En él se propone el desdoblamiento *batch-wise* para reordenar los arreglos tridimensionales. Las matrices correspondientes a cada una de las K observaciones son reubicadas una a la derecha de la otra en forma sucesiva, formando de esta manera una matriz bidimensional que conserva la dimensión del primer modo (filas) mientras que la del segundo (columnas) resulta del producto del número de variables (J) por el número de observaciones realizadas (K).

La facilidad de implementación del desdoblamiento *batch-wise* y de la interpretación de las matrices resultantes ha hecho que muchos autores eligieran este modo de desdoblamiento como punto de partida en sus trabajos. Sin embargo, esta metodología presenta algunos inconvenientes que deben mencionarse. En primer lugar, dado que generalmente el número de

batches de los que se dispone es limitado, la matriz bidimensional producida por el desdoblamiento *batch-wise* tiene un número mucho más elevado de columnas que de filas. De esta manera, la matriz de correlación, que se estima tomando el número de columnas como variables y el número de filas como individuos o puntos de observación, no resulta del todo confiable dada la gran cantidad de parámetros que deben ser estimados con relativamente poca información.

Otro inconveniente del desdoblamiento *batch-wise* está relacionado con la implementación en línea de la estrategia de monitoreo. A fin de determinar si el proceso se encuentra en estado de control, se requiere proyectar cada individuo obtenido en tiempo real en el espacio de las variables latentes. Para ello, dicho individuo debe tener la misma dimensión que los *batches* incluidos en la base de datos a partir de la cual se construyó el modelo. Dado que durante el monitoreo en línea los *batches* todavía no están completos, es necesario hacer estimaciones de los valores futuros de las variables medidas, desde la observación siguiente a la actual ($k+1$) hasta la última (K).

Se han propuesto diferentes metodologías con el objetivo fundamental de evitar la necesidad de predecir valores futuros para las observaciones, las cuales serán analizadas a continuación.

Una alternativa consiste en formular K modelos locales, uno para cada observación k , a fin de describir adecuadamente la dependencia de las variables de proceso entre ellas y con el tiempo, a la vez que se continúa empleando el método de desdoblamiento *batch-wise*.

El modelo local más simple contiene información exclusivamente de la observación k . A fin de incorporar la historia previa de la operación, Rännar y col. (1991) presentaron una adaptación del enfoque jerárquico propuesto para procesos continuos. Esta metodología utiliza modelos locales que incluyen el conocimiento previo del proceso en una fracción de la

matriz de proyección. En cada modelo se asigna mayor influencia a las últimas observaciones, mientras que las mediciones previas van perdiendo importancia en forma paulatina mediante el empleo de un factor de suavizado similar al usado en las técnicas *EWMA*.

La idea de la utilización de ventanas móviles (ampliamente difundida en el área de control de procesos) fue propuesta por Lennox *y col.* (2001). Los autores definieron una zona de influencia para cada modelo local, la cual determina la porción de la historia previa de la operación asociada con dicho modelo.

Ramaker *y col.* (2005) introdujeron los llamados modelos evolutivos. La metodología consiste en definir un modelo diferente para cada instante k que comprende todas las observaciones desde el inicio de la operación hasta k . De esta manera, toda la historia previa de la operación se incluye en el modelo. Las coordenadas del espacio obtenido para la primera observación resultan iguales tanto al utilizar un modelo evolutivo como uno local. Por otra parte, cuando se incluye finalmente la última observación, el modelo evolutivo es equivalente al propuesto por Nomikos y MacGregor (1994).

Otra alternativa destinada a evitar la estimación de valores futuros para las variables medidas se basa en modificar el tipo de desdoblamiento. Wold *y col.* (1998) presentaron una estrategia, referida como *BOL (Batch Observation Level)* por varios autores, en la cual se emplea un desdoblamiento vertical de la matriz de datos. Éste consiste en ubicar la matriz correspondiente a las observaciones del tiempo k debajo de la anterior, manteniendo así el número de variables y permitiendo la proyección de las observaciones a medida que las mismas están disponibles. En este trabajo, las matrices de proyección se obtienen de manera global, es decir, considerando las operaciones completas del principio al fin.

Westerhuis *y col.* (1999) remarcaron que el desdoblamiento propuesto por Wold *y col.* (1998) mantiene las componentes no lineales del proceso que generan la mayor variabilidad

en el modelo. De esta manera, las primeras componentes principales se usan para explicar dicha variabilidad, enmascarando así efectos importantes en una escala de variación más pequeña. Una solución a este problema fue introducida por Yoo *y col.* (2004a). En su estrategia, el desdoblamiento del arreglo tridimensional se hace en dos etapas. En la primera, el desdoblamiento *batch-wise* se emplea para realizar el centrado y escalado de cada matriz temporal. En la segunda, estas matrices se reordenan nuevamente una debajo de la otra, según lo propuesto por Wold *y col.* (1998). Cada una de las matrices temporales queda entonces estandarizada usando su propia media y desvío estándar en lugar de los parámetros globales.

Debe notarse que el método propuesto por Wold *y col.* (1998) resulta global desde el punto de vista del espacio obtenido para la proyección. Esto significa que el modelo resultante no considera los posibles cambios en la correlación de las variables en las distintas etapas del proceso. A fin de tener en cuenta esta variación, se incorporó el uso de matrices de correlación locales, ya en el espacio de las variables latentes (Van Sprang *y col.*, 2002). En consecuencia, estas estrategias resultan ser globales para la reducción de la dimensionalidad y locales para el testeado de hipótesis.

Los dos tipos de desdoblamientos explicados anteriormente se han usado ampliamente en el campo del monitoreo de procesos *batch*, y se aplicaron a muy diversos casos de estudio (Nomikos y MacGregor, 1994; Wold, 1994; Nomikos y MacGregor, 1995; Yoo *y col.*, 2004b; Ramaker *y col.*, 2005). Un enfoque completamente diferente fue presentado por Chen y Liu (2002) como una extensión de las técnicas propuestas años atrás por Ku *y col.* (1995) al caso de procesos *batch*. Los primeros proponen incorporar las características dinámicas del sistema en forma similar a como lo hacen los modelos autoregresivos (*ARMAX*), ampliamente usados en el campo de identificación de sistemas. La aplicación de esta técnica requiere un

reordenamiento en la estructura de los datos tal que la estrategia no puede agruparse con ninguna de las descriptas anteriormente.

Los enfoques y procedimientos para el tratamiento del arreglo tridimensional de datos citados previamente se utilizan como punto de partida común para la aplicación de diversas técnicas estadísticas multivariadas, tales como: *PCA*, *PLS*, Análisis de Componentes Independientes (*ICA*) y Análisis de Correspondencia (*CA*). La elección del tipo de desdoblamiento y de la técnica estadística tiene gran influencia en la capacidad de detección de la estrategia de monitoreo.

1.2.2 *Detección*

Las estrategias de *SPC* se basan en la realización de tests de hipótesis cada vez que se obtiene una nueva observación que describe el estado del proceso. La cantidad de tests a realizar depende de la metodología empleada. Los estadísticos más utilizados en el área de monitoreo de procesos son: el Estadístico de Hotelling (T^2) (Hotelling, 1931) y el Estadístico del Error Cuadrático de Representación (*SPE*). Este último sólo tiene sentido cuando se usan técnicas de proyección.

Algunos autores (Dunia y Qin, 1998; Qin, 2003) realizaron un análisis muy completo sobre la capacidad de los estadísticos para la detección de diferentes tipos de fallas. Si bien su interés estaba centrado en la detección de errores sistemáticos en los instrumentos en lugar de situaciones anómalas en el comportamiento del proceso, existe una gran analogía en el tratamiento del problema en los dos casos.

Qin (2003) destaca los roles asimétricos del T^2 y del *SPE* en la detección de fallas y analiza cualitativamente el tipo de eventos que podrían ser detectados por cada uno de éstos, dependiendo de las características de los datos que se consideraron durante la generación de

los modelos. El mismo autor propone además el uso de un índice combinado que resulta de la unificación de estas dos métricas.

1.2.3 Identificación

Luego que una falla o evento anormal es detectado, es necesario identificar la variable o el conjunto de variables que indica esa condición. La etapa de identificación o aislamiento es el paso en el cual se genera la información requerida para realizar el diagnóstico final del evento.

Las tareas de identificación no resultan triviales. Durante los últimos años se presentaron una gran cantidad de trabajos que proponen distintos enfoques para la determinación de la(s) variable(s) que señala(n) la anomalía en la operación.

La complejidad de esta etapa depende fuertemente de la fórmula matemática de la métrica que indica la existencia de un evento anormal. Por ejemplo, las fórmulas propuestas en la literatura para el cálculo de las contribuciones de cada variable al *SPE* son simples y directas. Las mismas evalúan la discrepancia que existe entre los valores medidos de cada variable y los predichos por el modelo de variables latentes que esté siendo utilizado (Nomikos y MacGregor, 1994; Kourti y MacGregor, 1995; Nomikos y MacGregor, 1995; Westerhuis *y col.*, 2000; Qin, 2003). Las variaciones más importantes en cómo se calculan las contribuciones a *SPE* consisten en considerar los desvíos en cada variable o los cuadrados de los mismos, haciendo que las contribuciones tengan o no la posibilidad de tomar valores negativos.

El estadístico T^2 es ampliamente utilizado en el análisis de sistemas multivariados en general y en el campo del monitoreo de procesos en particular. Si bien se presenta como una herramienta muy útil para la detección de eventos anormales, por sí mismo no brinda ninguna

información respecto de las posibles causas del desvío. Su fórmula matemática está compuesta, en su mayor parte, por productos cruzados de variables. Ésto dificulta tanto su interpretación como el cálculo de la influencia de cada variable sobre el valor final del estadístico.

Existen numerosos trabajos donde se proponen enfoques diferentes con el fin de dar una solución confiable al problema de identificación cuando se utiliza el estadístico T^2 . Jackson (1991) propuso la descomposición de T^2 en la suma de las contribuciones de cada componente principal (CP) al valor global de la métrica. Luego, la identificación de la(s) variable(s) sospechosa(s) se hace en términos del peso relativo de cada uno de los descriptores en la CP más influyente. El inconveniente de este enfoque es que, a diferencia de lo que ocurre en otras áreas de la ciencia, en los procesos químicos resulta casi siempre imposible asociar una variable (o un grupo de variables bien definido) a una CP en particular.

Posteriormente, MacGregor *y col.* (1994) propusieron evaluar las contribuciones de cada una de las variables a las CPs que están fuera de los límites de control correspondientes. Esta metodología es análoga a considerar cada CP como una carta de control independiente, tomando ventaja del hecho de que las CPs no están correlacionadas entre sí. Si bien este enfoque podría implementarse sin el perjuicio de no considerar la naturaleza multivariada del sistema para el caso de procesos continuos, ésta no es una alternativa conveniente para los procesos *batch* dado que presentan distintos niveles de correlación (globales y locales).

Nomikos (1996) y Miller *y col.* (1998) formularon el cálculo directo de las contribuciones de cada variable al valor final del T^2 , en lugar de pesarlas por su participación en las CP dominantes, presentando los correspondientes gráficos de contribución. Además, Westerhuis *y col.* (2000) extendieron esta teoría al caso de espacios latentes donde las direcciones podrían

eventualmente no ser ortogonales e introdujeron algunas nociones de cómo calcular los límites para cada contribución.

Por otra parte, Mason *y col.* (1995, 1997) propusieron la descomposición del T^2 en el espacio original de las mediciones, como una suma de J partes independientes. El primer término de esta suma se calcula como el cuadrado del valor del estadístico t de Student para la primera variable. El j -ésimo término ($j=1,2,\dots,J$) corresponde al estadístico t ajustado con las medias y desvíos estándar de las correspondientes funciones de probabilidad condicional dadas las $j-1$ variables anteriores. Puesto que no existe un ordenamiento fijo para las variables, pueden obtenerse hasta $J!$ particiones diferentes, aunque no independientes.

Dada la gran cantidad de términos que es necesario considerar en la descomposición de Mason *y col.* (1995, 1997), la interpretación de los resultados obtenidos no resulta fácil de llevar a cabo. Los mismos autores propusieron tener en cuenta sólo dos términos para cada variable. El primero la considera en forma independiente del resto y el segundo corresponde a su distribución condicionada a los valores de los restantes $J-1$ descriptores. Otros grupos de trabajo (Hawkins, 1991; Montgomery, 1996; Timm, 1996) formularon metodologías alternativas, que constituyen casos particulares de la técnica propuesta por Mason (Maravelakis, 2003).

1.3 Organización de la Tesis

El trabajo de tesis está organizado de la siguiente manera:

El presente capítulo contiene los objetivos y motivaciones de las tareas de investigación desarrolladas y proporciona una visión general sobre el estado del arte en el área del *MSPC*.

El Capítulo 2 presenta las técnicas de proyección en estructuras latentes que son, hasta el momento, las más utilizadas en el ámbito del *SPC*.

El Capítulo 3 discute en detalle las particularidades que presentan los procesos *batch* desde el punto de vista del *SPC* y analiza en profundidad las diferentes técnicas utilizadas para monitorear estos procesos.

El Capítulo 4 propone una nueva metodología, desarrollada durante este trabajo de tesis, para el monitoreo de procesos en el espacio de las variables originales.

El Capítulo 5 describe el proceso de polimerización en emulsión de metacrilato metilo (*MMA*), elegido como caso de estudio, y presenta también el desarrollo del modelo matemático del proceso y su implementación.

El Capítulo 6 muestra los resultados obtenidos al aplicar las diferentes técnicas de *SPC* analizadas en esta tesis al caso de estudio, la comparación de sus desempeños y algunas conclusiones parciales que de ésta se desprenden.

El Capítulo 7 contiene las conclusiones generales, además de las recomendaciones acerca del trabajo futuro a realizar en esta temática.

1.4 Referencias

- Alcorta García, E. y P. M. Frank, "Deterministic Nonlinear Observer-Based Term Approaches to Fault Diagnosis: A Survey". *Control Engineering Practice* **1997**, 5, (5), 663-670.
- AlGhazzawi, A. y B. Lennox, "Monitoring a Complex Refining Process Using Multivariate Statistics". *Control Engineering Practice* **2008**, 16, (3), 294-307.
- Apte, C., B. Liu, E. P. D. Pednault y P. Smyth, "Business Applications of Data Mining". *Communications of the ACM* **2002**, 45, (8), 49-53.
- Asua, J. M., "Control of Emulsion Polymerization Reactors Using a Hierarchical Fuzzy Logic/Model-Based Controller". *Polymer Reaction Engineering* **2001**, 9, (1), 37-67.
- Basseville, M., "Detecting Changes in Signals and Systems-A Survey". *Automatica* **1988**, 24, (3), 309-326.
- Chen, J. y K. C. Liu, "On-Line Batch Process Monitoring Using Dynamic PCA and Dynamic PLS Models". *Chemical Engineering Science* **2002**, 57, (1), 63-75.
- Dunia, R., S. J. Qin, T. F. Edgar y T. J. McAvoy, "Identification of Faulty Sensors Using Principal Component Analysis". *AIChE Journal* **1996**, 42, (10), 2797-2811.
- Dunia, R., S. J. Qin y T. F. Edgar (1995). Multivariable Process Monitoring Using Nonlinear Approaches. Proceedings of the American Control Conference, Seattle, WA, USA.
- Dunia, R. y S. J. Qin, "Subspace Approach to Multidimensional Fault Identification and Reconstruction". *AIChE Journal* **1998**, 44, (8), 1813-1831.
- Fayyad, U. M., G. Piatetsky-Shapiro, P. Smyth y R. Uthurusamy (1996). Advances in Knowledge Discovery and Data Mining, Menlo Park, California, The MIT Press.
- Gertler, J., "Residual Generation in Model-Based Fault Diagnosis". *Control, Theory and Advanced Technology* **1993**, 9, (1), 259-285.
- Goebel, M. y L. Gruenwald, "A Survey of Data Mining and Knowledge Discovery Software". *SIGKDD Explorations* **1999**, 2, (1), 20-33.
- Hamielc, A. E. y A. Penlidis (1995). Notes of 10th Annual Intensive Course on Polymerization Processes and Reaction Engineering, Porto Carras, Grecia.
- Hawkins, D. M., "Multivariate Quality Control Based on Regression-Adjusted Variables". *Technometrics* **1991**, 33, (1), 61-75.
- Hotelling, H., "The Generalization of Student's Ratio". *The Annals of Mathematical Statistics*. **1931**, 2, (3), 360-378.
- Hyvärinen, A., "Algorithms and Applications". *Neural Networks* **2000**, 13, (4-5), 411-430.
- Hyvärinen, A., J. Karhunen y E. Oja (2001). Independent Component Analysis, New York, Wiley-Interscience.

- Jackson, J. E., "Principal Components and Factor Analysis: Part I - Principal Components". *Journal of Quality Technology* **1980**, 12, (4), 201-213.
- Jackson, J. E. (1991). A User's Guide to Principal Components. Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics, New York, John Wiley & Sons, Inc.
- Jaekle, C. M. y J. F. MacGregor, "Industrial Applications of Product Design through the Inversion of Latent Variable Models". *Chemometrics and Intelligent Laboratory Systems* **2000**, 50, (2), 199-210.
- Jaekle, C. M. y J. F. MacGregor, "Product Design through Multivariate Statistical Analysis of Process Data". *AIChE Journal* **1998**, 44, (5), 1105-1118.
- Kourti, T., J. Lee y J. F. MacGregor, "Experiences with Industrial Applications of Projection Methods for Multivariate Statistical Process Control". *Computers and Chemical Engineering* **1996**, 20, (SUPPL.1), S745-S750
- Kourti, T. y J. F. MacGregor, "Process Analysis, Monitoring and Diagnosis, Using Multivariate Projection Methods". *Chemometrics and Intelligent Laboratory Systems* **1995**, 28, (1), 3-21.
- Kourti, T. y J. F. MacGregor, "Multivariate SPC Methods for Process and Product Monitoring". *Journal of Quality Technology* **1996**, 28, (4), 409-428.
- Kroonenberg, P. M., "Annotated Bibliography of Three-Mode Factor Analysis". *British Journal of Mathematical and Statistical Psychology* **1983**, 36, (1), 81-113.
- Kroonenberg, P. M., "The Analysis of Multiple Tables in Factorial Ecology. III. Three- Mode Principal Component Analysis: "Analyse Triadique Complete"". *Acta Oecologica/Oecologia Generalis* **1989**, 10, (3), 245-256.
- Kroonenberg, P. M. y J. de Leeuw, "Principal Component Analysis of Three-Mode Data by Means of Alternating Least Squares Algorithms". *Psychometrika* **1980**, 45, (1), 69-97.
- Ku, W., R. H. Storer y C. Georgakis, "Disturbance Detection and Isolation by Dynamic Principal Component Analysis". *Chemometrics and Intelligent Laboratory Systems* **1995**, 30, (1), 179-196.
- Lennox, B., G. A. Montague, H. G. Hiden, G. Kornfeld y P. R. Goulding, "Process Monitoring of an Industrial Fed-Batch Fermentation". *Biotechnology and Bioengineering* **2001**, 74, (2), 125-135.
- Li, W., H. H. Yue, S. Valle-Cervantes y S. J. Qin, "Recursive PCA for Adaptive Process Monitoring". *Journal of Process Control* **2000**, 10, (5), 471-486.
- MacGregor, J. F., C. Jaekle, C. Kiparissides y M. Koutoudi, "Process Monitoring and Diagnosis by Multiblock PLS Methods". *AIChE Journal* **1994**, 40, (5), 826-838.
- Marvelakis, P. (2003) An Investigation of Some Characteristics of Univariate and Multivariate Control Charts. Ph.D. Thesis, Department of Statistics, Athens University of Economics and Business.

- Mason, R. L., N. D. Tracy y J. C. Young, "A Practical Approach for Interpreting Multivariate T² Control Chart Signals". *Journal of Quality Technology* **1997**, 29, (4), 396-406.
- Mason, R. L., N. D. Tracy y J. C. Young, "Decomposition of T² for Multivariate Control Chart Interpretation". *Journal of Quality Technology* **1995**, 27, (2), 99-108.
- Miller, P., R. E. Swanson y C. F. Heckler, "Contribution Plots: The Missing Link in Multivariate Quality Control". *International Journal of Applied Mathematics and Computer Science* **1998**, 8, (4), 775-792.
- Montgomery, D. C., "Contributors to a Multivariate Statistical Process Control Chart Signal". *Communications in Statistics - Theory and Methods* **1996**, 25, (10), 2203-2213.
- Nomikos, P., "Detection and Diagnosis of Abnormal Batch Operations Based on Multi-Way Principal Component Analysis". *ISA Transactions* **1996**, 35, (3), 259-266.
- Nomikos, P. y J. F. MacGregor, "Monitoring of Batch Process Using Multiway Principal Component Analysis". *AIChE Journal* **1994**, 40, (8), 1361-1375.
- Nomikos, P. y J. F. MacGregor, "Multivariate SPC Charts for Monitoring Batch Processes". *Technometrics* **1995**, 37, (1), 41-59.
- Qin, S. J., "Statistical Process Monitoring: Basics and Beyond". *Journal of Chemometrics* **2003**, 17, (8-9), 480-502.
- Qin, S. J., S. Valle y M. J. Piovoso, "On Unifying Multiblock Analysis with Application to Decentralized Process Monitoring". *Journal of Chemometrics* **2001**, 15, (9), 715-742.
- Ramaker, H. J., E. N. M. Van Sprang, J. A. Westerhuis y A. K. Smilde, "Fault Detection Properties of Global, Local and Time Evolving Models for Batch Process Monitoring". *Journal of Process Control* **2005**, 15, (7), 799-805.
- Rännar, S., J. F. MacGregor y S. Wold, "Adaptive Batch Monitoring Using Hierarchical PCA". *Chemometrics and Intelligent Laboratory Systems* **1991**, 41, (1), 73-81.
- Sebzalli, Y. M. y X. Z. Wang, "Knowledge Discovery from Process Operational Data Using PCA and Fuzzy Clustering". *Engineering Applications of Artificial Intelligence* **2001**, 14, (5), 607-616.
- Timm, N. H., "Multivariate Quality Control Using Finite Intersection Tests". *Journal of Quality Technology* **1996**, 28, (2), 233-243.
- Van Sprang, E. N. M., H.-J. Ramaker, J. A. Westerhuis, S. P. Gurden y A. K. Smilde, "Critical Evaluation of Approaches for On-line Batch Process Monitoring". *Chemical Engineering Science* **2002**, 57, (18), 3979-3991.
- Wangen, L. E. y B. R. Kowalski, "A Multiblock Partial Least Squares Algorithm for Investigating Complex Chemical Systems". *Journal of Chemometrics* **1988**, 3, (1), 3-20.
- Westerhuis, J. A., S. P. Gurden y A. K. Smilde, "Generalized Contribution Plots in Multivariate Statistical Process Monitoring". *Chemometrics and Intelligent Laboratory Systems* **2000**, 51, (1), 95-114.

- Westerhuis, J. A., T. Kourti y J. F. MacGregor, "Analysis of Multiblock and Hierarchical PCA and PLS Models". *Journal of Chemometrics* **1998**, 12, (5), 301-321.
- Westerhuis, J. A., T. Kourti y J. F. MacGregor, "Comparing Alternative Approaches for Multivariate Statistical Analysis of Batch Process Data". *Journal of Chemometrics* **1999**, 13, (3-4), 397-413.
- Willsky, A. S., "A Survey of Design Methods for Failure Detection in Dynamic Systems". *Automatica* **1976**, 12, (6), 601-611.
- Wold, S., "Exponentially Weighted Moving Principal Components Analysis and Projections to Latent Structures". *Chemometrics and Intelligent Laboratory Systems* **1994**, 23, (1), 149-161.
- Wold, S., P. Geladi, K. Esbensen y J. Öhman, "Multi-Way Principal Components and PLS Analysis". *Journal of Chemometrics* **1987**, 1, (1), 41-56.
- Wold, S., N. Kettaneh, H. Friden y A. Holmberg, "Modelling and Diagnostics of Batch Processes and Analogous Kinetic Experiments". *Chemometrics and Intelligent Laboratory Systems* **1998**, 44, (1-2), 331-340.
- Yabuki, Y., T. Nagasawa y J. F. MacGregor, "Industrial Experiences with Product Quality Control in Semi-Batch Processes". *Computers and Chemical Engineering* **2002**, 26, (2), 205-212.
- Yoo, C. K., D. S. Lee y P. A. Vanrolleghem, "Application of Multiway ICA for On-line Process Monitoring of a Sequencing Batch Reactor". *Water Research* **2004a**, 38, (7), 1715-1732.
- Yoo, C. K., J. M. Lee, P. A. Vanrolleghem y I. B. Lee, "On-line Monitoring of Batch Processes Using Multiway Independent Component Analysis". *Chemometrics and Intelligent Laboratory Systems* **2004b**, 71, (2), 151-163.
- Zhang, J., A. J. Morris, E. B. Martin y C. Kiparissides, "Estimation of Impurity and Fouling in Batch Polymerisation Reactors through the Application of Neural Networks". *Computers and Chemical Engineering* **1999**, 23, (3), 301-314.

Página en Blanco

Capítulo 2

Monitoreo Estadístico de Procesos Utilizando Estrategias de Proyección

2.1	Introducción.....	27
2.2	Monitoreo Basado en el Análisis de Componentes Principales.....	33
2.2.1	<i>Análisis de Componentes Principales (PCA)</i>	33
2.2.2	<i>Implementación de la Estrategia de Monitoreo</i>	41
2.2.2.1	Etapa I.....	41
2.2.2.2	Etapa II	49
2.3	Monitoreo Basado en el Análisis de Componentes Independientes	55
2.3.1	<i>Análisis de Componentes Independientes (ICA)</i>	55
2.3.2	<i>Implementación de la Estrategia de Monitoreo</i>	62
2.3.2.1	Etapa I.....	63
2.3.2.2	Etapa II	64
2.4	Monitoreo Basado en Cuadrados Mínimos Parciales.....	65
2.4.1	<i>Cuadrados Mínimos Parciales (PLS)</i>	65
2.4.2	<i>Implementación de la Estrategia de Monitoreo</i>	67
2.5	Referencias	68

Página en Blanco

2 Monitoreo Estadístico de Procesos Utilizando Estrategias de Proyección

2.1 Introducción

La descripción de procesos químicos mediante modelos matemáticos es una práctica común en el área de la ingeniería de sistemas de procesos. En las últimas décadas, numerosos autores han propuesto estrategias de monitoreo y control de procesos basadas en modelos de principios fundamentales (Gertler, 1993; Venkatasubramanian y *col.*, 2003b). Dichos autores demuestran la potencialidad de estas estrategias para mejorar la operación de los procesos analizados.

Sin embargo, la aplicación de esas técnicas requiere, por lo general, implementar modelos detallados del proceso y de las posibles fallas (o conjunto de fallas) que se consideran. El esfuerzo requerido para modelar apropiadamente el sistema reduce su aplicación práctica sólo a problemas de pequeña y/o mediana escala. Además, cuando se utilizan modelos de principios fundamentales no es posible incluir directamente la descripción de la incertidumbre asociada a los valores de los parámetros. Tampoco resulta sencillo incorporarles las fluctuaciones en los valores de las variables no controladas y los posibles errores en las mediciones. Otro punto importante a considerar en este tipo de enfoque es que se podría llegar a un diagnóstico incorrecto si se realizaron simplificaciones erróneas en el modelo o hubo cambios en el comportamiento del proceso no previstos en el mismo (Venkatasubramanian y *col.*, 2003b).

En las últimas décadas, la *Simulación Estocástica* surgió como una alternativa para manejar cierto grado de incertidumbre cuando se emplean modelos de principios fundamentales. De esta forma los modelos pueden mejorarse incluyendo factores tales como

la incertidumbre asociada a los parámetros y variaciones en las condiciones de operación, entre otros (Näf, 1994). Las herramientas disponibles en el mercado para efectuar este tipo de tratamiento son aún bastante limitadas desde el punto de vista de su capacidad de análisis de resultados. En cualquier caso, dicho análisis necesita de la aplicación de técnicas estadísticas avanzadas, haciendo que esta metodología constituya un enfoque híbrido entre los modelos de principios fundamentales y las técnicas basadas en datos históricos.

La necesidad de una solución práctica a los problemas mencionados anteriormente alentó el uso de las técnicas basadas en datos (*Data Driven Techniques*) para el modelado de sistemas de proceso. Estas metodologías ya habían sido utilizadas con diversos propósitos en otras áreas del conocimiento, tales como las ciencias naturales, la economía y la agronomía (Kloek y Mennes, 1960; Jeffers, 1962; Vavra, 1972; Van Gerven y Oakland, 1973; Jackson, 1991). Dada la gran cantidad de técnicas propuestas, mejoradas y comprobadas para la resolución de problemas similares en éstas y otras áreas, los enfoques estadísticos proporcionaron una alternativa viable con un respaldo teórico robusto y confiable.

En el área particular del control de procesos, el *SPC* se ha aplicado exitosamente para incorporar las variaciones y comportamientos del proceso que no son tenidos en cuenta en el diseño de los lazos de control específicos (Venkatasubramanian *y col.*, 2003b).

El objetivo principal de las estrategias de *SPC* es comparar los diferentes estados del proceso durante su operación con aquellos que produjeron un producto final conforme con las características requeridas para éste. Estas técnicas son de sumo interés para el control de procesos ya que, en general, éstos contienen un gran número de variables que resultan muy difíciles de controlar.

Las estrategias de *SPC* poseen un inmenso potencial en lo relacionado a las actividades de control y operación en la industria de procesos moderna. La misma está sujeta a una gran

cantidad de requisitos medioambientales que obligan a adecuar los métodos de producción. Las disposiciones legales, juntamente con los requerimientos de calidad impuestos por la competitividad de los mercados, generan la necesidad de acotar tanto como sea posible el rango de variabilidad e incertidumbre al que están sujetos los procesos de producción.

Dada la alta especificidad y automatización de la industria moderna, las variaciones en los procedimientos de producción durante la fabricación de un determinado producto son cada vez menos significativas. Sin embargo, existen aún muchas fuentes de variabilidad asociadas de forma directa o indirecta a dichos procedimientos. Una de las más importantes es la intervención humana en estos procesos altamente automatizados. La forma en la que distintos operadores abren y cierran válvulas manuales o cómo intervienen manualmente los *setpoints* de los controladores ante situaciones que les resultan anormales, son ejemplos de estas fuentes de variación.

La Figura 2.1 muestra un esquema de las diferentes fuentes de variación que pueden afectar la calidad final del producto. Además de la influencia de la mano de obra pueden mencionarse, por ejemplo, las variaciones en la calidad de las materias primas y en el medio ambiente circundante.

Los cambios en las variables de proceso pueden clasificarse en dos grandes grupos: los que se hacen a propósito en las variables manipuladas y los cambios fortuitos que sufren las demás. Las perturbaciones que se inducen sobre las variables manipuladas del proceso son conocidas como causas asignables de variación, puesto que, dada una acción de cambio en alguna de ellas puede esperarse una determinada reacción de parte del proceso. Las fluctuaciones sobre las variables no manipuladas (originadas por fuentes aleatorias) son conocidas como causas no asignables, debido a la incapacidad de determinar fehacientemente la causa primera de las variaciones que se observan.



Figura 2.1: Esquema de la relación entre la calidad final del producto y las posibles fuentes de variación

Como se mencionó anteriormente, el objetivo del monitoreo de procesos es la detección e identificación de condiciones de operación anormales a fin de poder implementar alguna acción de control sobre las causas del problema (Kourti y MacGregor, 1995; Venkatasubramanian y *col.*, 2003a; Venkatasubramanian y *col.*, 2003b). Una vez que el sistema de monitoreo detecta una situación anómala, el personal responsable del proceso es el encargado de producir cambios en el mismo de modo de hacer que éste regrese al estado de control. Las acciones necesarias pueden ser llevarse a cabo por medio de:

- a. “Planes de acción para situaciones fuera de control” (OCAPS), que detallan las acciones a seguir una vez que se declara un estado fuera de control. Estas acciones están supeditadas al diagnóstico del problema.

- b. Lazos de control avanzado, que efectúan cambios programados para corregir el estado del proceso.

El monitoreo estadístico de procesos se compone de tres actividades fundamentales: detección del estado anómalo, identificación de la(s) variable(s) que señala(n) dicho estado y finalmente el diagnóstico, donde se indican las causas aparentes del problema. Estas estrategias deben, como mínimo, ser capaces de (Jackson, 1991):

- a. Responder si el proceso está o no bajo control.
- b. Fijar una cota para el *Error global de Tipo I*.
- c. Considerar las relaciones existentes entre las variables.
- d. Diagnosticar la causa del problema si el proceso se declara fuera de control.

El cumplimiento de los tres primeros puntos no presenta mayores complicaciones una vez que se elige la estrategia de monitoreo. Los niveles de confianza y los criterios para la declaración de los estados fuera de control son definidos por el usuario de acuerdo a las necesidades del caso particular y su conocimiento del proceso. Sin embargo, el último de los puntos de esta lista resulta particularmente difícil de satisfacer, especialmente cuando el número de variables consideradas es grande.

El monitoreo del proceso se lleva a cabo en forma continua a medida que las nuevas observaciones del sistema van siendo obtenidas. Cada una de estas observaciones se somete a una serie de transformaciones permitiendo así que la información que contiene sea usada para determinar si el proceso se encuentra o no bajo control. El conjunto de estas actividades componen la etapa de detección, y están directamente relacionadas con los primeros tres puntos mencionados anteriormente.

Cuando se detecta una situación anormal, el proceso se declara fuera de control y se inicia la etapa de identificación. El objetivo de ésta, es encontrar el conjunto de mediciones que indica aquella condición. Una vez que estas variables han sido identificadas, comienza el proceso de diagnóstico cuyo objetivo es encontrar la causa original de la desviación observada (que puede ser o no una alteración en alguna de las variables medidas).

Las etapas de identificación y diagnóstico se tratan generalmente en forma conjunta, refiriéndose a ellas como la etapa de diagnóstico del sistema. Sin embargo, la etapa de diagnóstico propiamente dicha tiene características muy particulares dado que requiere frecuentemente el uso adicional de sistemas expertos para encontrar el origen de la anomalía. Estos sistemas se basan generalmente en el empleo de técnicas tales como lógica difusa y redes neuronales (Ruiz, 2001; Yélamos Ruiz, 2008). En el tratamiento de procesos continuos esta diferenciación se hace más evidente que en el caso de procesos *batch*.

Si se consideran ahora las actividades relacionadas específicamente con la *implementación* de una estrategia de monitoreo de procesos, éstas se separan en dos etapas. Las *Etapas I y II* son también llamadas *modelado fuera de línea (Etapa I)* y *monitoreo en línea (Etapa II)* (Nomikos y MacGregor, 1994; NIST/Sematech, 2008).

El objetivo principal de la *Etapa I* es la obtención de un modelo que represente adecuadamente el proceso. Las tareas involucradas en esta fase son: el pretratamiento de los datos, la formulación del modelo y el cálculo de los valores críticos de los estadísticos usados para la detección de eventos anormales. En la *Etapa II* el proceso se monitorea en línea a fin de determinar si éste se encuentra en estado de control. Los modelos estadísticos desarrollados en la etapa anterior se utilizan para la tarea de detección de eventos anormales, y en caso de ser necesario, para la identificación de las variables que revelan el estado fuera de control.

Diversas técnicas basadas en métodos de proyección en estructuras latentes se han utilizado para el monitoreo de procesos durante las últimas décadas. Entre las más populares se encuentran: Análisis de Componentes Principales (*PCA*), Análisis de Componentes Independientes (*ICA*) y Cuadrados Mínimos Parciales (*PLS*) (Anderson, 1958; Jackson, 1991; Kourti y MacGregor, 1995; Lee y col., 2004). Los conceptos básicos de las mismas así como otros relacionados con su implementación para el monitoreo de procesos se discutirán en el resto de este capítulo. No se incluye la etapa de diagnóstico por estar fuera del alcance de esta tesis. Las estrategias *PCA* e *ICA* se describen con mayor detalle dado que la aplicación y el análisis del desempeño de las mismas son parte central de este trabajo de tesis.

2.2 Monitoreo Basado en el Análisis de Componentes Principales

2.2.1 Análisis de Componentes Principales (*PCA*)

Esta técnica ofrece la posibilidad de lograr reducciones importantes en la dimensionalidad del sistema, manteniendo la mayor parte de la información significativa. La idea fundamental del *PCA* es encontrar un nuevo espacio de variables en el cual la “cantidad de información” contenida en los datos de entrada se encuentre “ordenada” de forma tal que, sea posible excluir del análisis aquellas direcciones que no contienen una cantidad de información relevante.

La información contenida en los datos de entrada está dada por distintas cantidades o índices que pueden calcularse a partir de los mismos. Estos índices permiten determinar el estado de los sistemas descritos por esos datos, ya sea que correspondan a un conjunto o población, o a individuos aislados que serán comparados contra una población de referencia (*PR*).

En particular, *PCA* ofrece la posibilidad de mantener simultáneamente dos de estas cantidades:

- a. La distancia Euclídea entre las observaciones.
- b. La varianza total del sistema.

Cuando todas las Componentes Principales (*CPs*) se conservan para el análisis, las dos cantidades mencionadas se mantienen inalteradas. La principal característica del *PCA* es la de concentrar estas magnitudes en las primeras dimensiones, de manera tal que la eliminación de las últimas componentes no represente una pérdida significativa de información.

La reducción de la dimensionalidad del problema se realiza de manera tal que, las direcciones elegidas para la descripción del sistema en el espacio reducido maximizan la cantidad de información asociada con las dos cantidades mencionadas.

El cumplimiento de estos dos criterios puede verificarse analizando la Figura 2.2, donde se muestra el caso de un experimento en el cual sólo se miden dos variables, x_1 y x_2 . Consideremos una observación cualquiera p y su proyección p^* sobre la dirección definida por la recta a en el plano x_1 y x_2 . La distancia \overline{op} en el triángulo rectángulo $\left(\overset{\Delta}{opp^*} \right)$ puede calcularse aplicando el teorema de Pitágoras:

$$\overline{op}^2 = \overline{pp^*}^2 + \overline{op^*}^2. \quad (2.1)$$

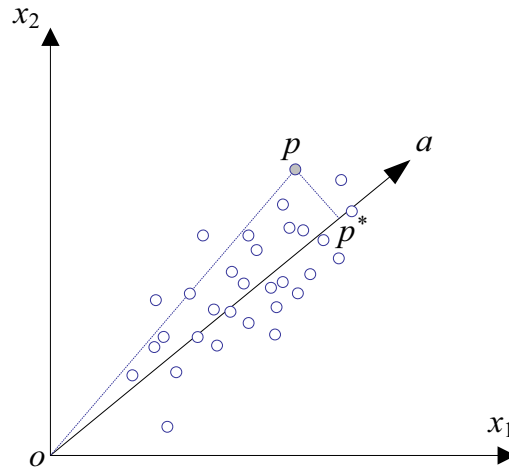


Figura 2.2: Realizaciones de un experimento con dos variables medidas

Para un dado conjunto de puntos, la distancia \overline{op} es invariante para cada uno de ellos. De esta manera, y según la ecuación (2.1), encontrar la dirección a que minimiza la distancia $\overline{pp^*}$ para el conjunto de todos los puntos p en la PR será equivalente a maximizar sus proyecciones en la dirección de a . De esta manera se maximiza también la varianza de las proyecciones de los datos en dicha dirección.

Si el procedimiento descrito se realiza en forma consecutiva para distintas direcciones a , sin repetir las ya encontradas e imponiendo condiciones de ortogonalidad entre las mismas, se maximizará la varianza retenida por las primeras direcciones manteniendo las distancias entre estos puntos lo más cercanas posible a los valores originales. De esta forma las CPs quedan ordenadas de manera tal que la primera es la que capta (describe y/o explica) la mayor variabilidad de los datos en términos de la varianza total. La segunda, hace lo propio con la varianza que no es explicada por la primera y así sucesivamente. Por lo tanto, cuando todas las CPs son retenidas en un análisis, tanto la varianza total del sistema como las distancias Euclídeas entre las observaciones se reconstruyen perfectamente en el espacio latente.

Las direcciones que definen el espacio de CPs para un dado conjunto de datos se obtienen como solución del siguiente problema de optimización (\mathcal{S}_j):

$$\mathcal{S}_j: \max_{\mathbf{p}} \mathcal{E}\{\mathbf{p}^T \mathbf{x}\}^2 = \mathcal{E}\{\mathbf{p}^T \mathbf{x} \mathbf{x}^T \mathbf{p}\} = \mathbf{p}^T \mathbf{\Sigma} \mathbf{p}, \quad (2.2)$$

s.a.

$$\mathbf{p}_j^T \mathbf{p}_j = 1, \quad \forall j = 1, 2, \dots, J, \quad (2.3)$$

$$\mathcal{E}\{\mathbf{p}_{j+1}^T \mathbf{x}\}^2 \geq \mathcal{E}\{\mathbf{p}_j^T \mathbf{x}\}^2, \quad \forall j = 1, 2, \dots, J, \quad (2.4)$$

donde \mathbf{p}_j son los vectores columna de la matriz de cambio de coordenadas $\bar{\mathbf{P}}$ que define el espacio de variables latentes, \mathbf{X} es la matriz de datos previamente centrada de dimensiones $(I \times J)$, $\mathcal{E}\{\cdot\}$ es el valor esperado de su argumento y J es el número de variables consideradas. Cada uno de los \mathbf{p}_j define la j -ésima dirección del espacio latente.

La solución del problema \mathcal{S}_j puede obtenerse en forma analítica usando los resultados del Teorema 2.1 (Anderson, 1958):

Teorema 2.1

Sea \mathbf{x} un vector aleatorio tal que $\mathcal{E}\{\mathbf{x}\} = \mathbf{0}$ y $\mathcal{E}\{\mathbf{x} \mathbf{x}^T\} = \mathbf{\Sigma}$. Existe entonces una transformación lineal ortogonal $\mathbf{u} = \mathbf{P}^T \mathbf{x}$ tal que la matriz de covarianza de \mathbf{u} es $\mathcal{E}\{\mathbf{u} \mathbf{u}^T\} = \bar{\mathbf{\Lambda}}$, con $\bar{\mathbf{\Lambda}} = \text{diag}\left([\lambda_1, \lambda_2, \dots, \lambda_J]^T\right)$ y donde $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_J \geq 0$ son las raíces de la ecuación característica $|\mathbf{\Sigma} - \lambda \mathbf{I}| = 0$.

Del Teorema 2.1 surge que λ_j es el j -ésimo autovalor de la matriz de covarianza $\mathbf{\Sigma}$, dado que es una de las raíces de la ecuación característica $|\mathbf{\Sigma} - \lambda \mathbf{I}| = 0$, siendo \mathbf{I} una matriz identidad de dimensiones $(J \times J)$. Así, el espacio de CPs correspondiente a un dado conjunto

de datos está definido por la matriz de autovectores asociada a la matriz de covarianza (o correlación según el caso) de dichos datos. Este resultado es utilizado frecuentemente en la literatura para proponer la obtención de una base del espacio de *CPs* mediante la descomposición en valores singulares de la matriz Σ , esto es:

$$\Sigma = \bar{\mathbf{P}}\bar{\Lambda}\bar{\mathbf{P}}^T, \quad (2.5)$$

donde $\bar{\mathbf{P}}$ es la matriz de autovectores asociada a Σ y es la misma que resulta de la solución del problema \mathcal{P} .

Dadas las propiedades de la traza de una matriz en su forma canónica (ecuación (2.5)) puede escribirse que:

$$\sigma_T = tr(\Sigma) = \sum_{j=1}^J \sigma_{jj} = \sum_{j=1}^J \lambda_j = tr(\bar{\Lambda}), \quad (2.6)$$

siendo σ_T la “varianza total” del sistema ya mencionada anteriormente en este capítulo.

Si en lugar de usar la matriz de covarianza se considera la matriz de correlación, la ecuación (2.6) resulta:

$$\sigma_T = tr(\Sigma) = \sum_{j=1}^J \sigma_{jj} = \sum_{j=1}^J \lambda_j = tr(\bar{\Lambda}) = J. \quad (2.7)$$

Como se indica en el Teorema 2.1: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_J \geq 0$. Esto implica que los términos de la sumatoria $\sum_{j=1}^J \lambda_j$ tienen menor peso conforme j aumenta. Por lo tanto, los últimos autovalores aportan menos que los primeros en la reconstrucción de σ_T . Así, dado que las primeras *CPs* capturan la mayor parte de σ_T , la dimensionalidad del mismo podría reducirse

eliminando las direcciones de orden superior, sin el perjuicio de una pérdida de información significativa en lo que a σ_T respecta.

En la práctica, los parámetros poblacionales $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$, media y matriz de covarianza respectivamente, son desconocidos. Dependiendo del caso, pueden usarse diferentes estimaciones para los mismos, siendo las más comunes la media muestral $\bar{\mathbf{x}}$ como estimador de $\boldsymbol{\mu}$ y la matriz de covarianza muestral \mathbf{S} ó de correlación \mathbf{Z} como estimador de $\boldsymbol{\Sigma}$.

Una vez obtenida la matriz de rotación $\bar{\mathbf{P}}$, la matriz de datos \mathbf{X} puede escribirse como:

$$\mathbf{X} = \bar{\mathbf{T}}\bar{\mathbf{P}}^T, \quad (2.8)$$

donde $\bar{\mathbf{T}}$ es la matriz de coordenadas de \mathbf{X} en el espacio de las *CPs* y se denominará matriz de *SCORES* en lo sucesivo.

De lo anterior se desprenden las siguientes relaciones:

$$\mathbf{S} = \frac{1}{(I-1)} \mathbf{X}^T \mathbf{X} = \bar{\mathbf{P}} \bar{\boldsymbol{\Lambda}} \bar{\mathbf{P}}^T, \quad (2.9)$$

$$\bar{\boldsymbol{\Lambda}} = \frac{1}{(I-1)} \bar{\mathbf{T}}^T \bar{\mathbf{T}} = \text{diag}([\lambda_1, \lambda_2, \dots, \lambda_J]^T), \quad \forall j = 1, 2, \dots, J. \quad (2.10)$$

Si se considera el caso en que el espacio de *CPs* se reduce de \mathbb{R}^J a \mathbb{R}^R , es decir que las últimas $J - R$ componentes se eliminan del análisis, las matrices de rotación y de *SCORES* definidas arriba pueden reescribirse como:

$$\bar{\mathbf{P}} = [\mathbf{P} \quad \tilde{\mathbf{P}}], \quad (2.11)$$

$$\bar{\mathbf{T}} = [\mathbf{T} \quad \tilde{\mathbf{T}}], \quad (2.12)$$

donde $\bar{\mathbf{P}} \in \mathbb{R}^{J \times J}$ es la matriz de cambio de coordenadas cuando se considera el espacio completo de *CPs*, $\mathbf{P} \in \mathbb{R}^{J \times R}$ es la matriz del cambio de coordenadas cuando sólo R *CPs* son retenidas y define la transformación lineal $\tau: \mathbb{R}^J \rightarrow \mathbb{R}^R$, y $\tilde{\mathbf{P}} \in \mathbb{R}^{J \times (J-R)}$ es el complemento de \mathbf{P} para dar $\bar{\mathbf{P}}$. Asimismo, $\bar{\mathbf{T}} \in \mathbb{R}^{I \times J}$ define las coordenadas de \mathbf{X} en el espacio de *CPs* completo, $\mathbf{T} \in \mathbb{R}^{I \times R}$ representa las coordenadas de \mathbf{X} en el espacio reducido y finalmente $\tilde{\mathbf{T}} \in \mathbb{R}^{I \times (J-R)}$ representan las coordenadas de los datos en el espacio excluido.

La ecuación (2.8) puede reescribirse como:

$$\mathbf{X} = \bar{\mathbf{T}}\bar{\mathbf{P}}^T = \begin{bmatrix} \mathbf{T} & \tilde{\mathbf{T}} \end{bmatrix} \begin{bmatrix} \mathbf{P}^T \\ \tilde{\mathbf{P}}^T \end{bmatrix} = \mathbf{T}\mathbf{P}^T + \tilde{\mathbf{T}}\tilde{\mathbf{P}}^T = \mathbf{T}\mathbf{P}^T + \mathbf{E} = \hat{\mathbf{X}} + \mathbf{E}. \quad (2.13)$$

Teniendo en cuenta que $\mathbf{T} \triangleq \mathbf{X}\mathbf{P}$, $\hat{\mathbf{X}}$ y \mathbf{E} , resultan:

$$\hat{\mathbf{X}} = \mathbf{X}\mathbf{P}\mathbf{P}^T = \mathbf{X}\mathbf{C}, \quad (2.14)$$

$$\mathbf{E} = \mathbf{X}\tilde{\mathbf{P}}\tilde{\mathbf{P}}^T = \mathbf{X}(\mathbf{I} - \mathbf{P}\mathbf{P}^T) = \mathbf{X}(\mathbf{I} - \mathbf{C}), \quad (2.15)$$

donde \mathbf{E} es la matriz de residuos, de las mismas dimensiones que \mathbf{X} , que contiene la parte de \mathbf{X} que no está representada en el espacio reducido de *CPs*.

La bondad de la representación del modelo de *CPs* puede medirse de distintas maneras. Una de las más utilizadas, debido a la facilidad de interpretación, se basa en la reconstrucción de la varianza total que se logra con cada componente de forma individual (ecuación (2.16)) o la alcanzada en forma acumulativa considerando las primeras R *CPs* (ecuación (2.17)):

$$\% \sigma_T^{ind} = \frac{\lambda_r}{J} \times 100, \quad \forall r = 1, 2, \dots, J, \quad (2.16)$$

$$\% \sigma_T^{acum} = \sum_{r=1}^R \lambda_r \times \frac{100}{J}. \quad (2.17)$$

La varianza de cada variable explicada por las primeras R CPs es también una medida muy usada para cuantificar la calidad del modelo obtenido. Se calcula como:

$$\frac{\|\hat{\mathbf{h}}_j\|^2}{\|\mathbf{h}_j\|^2} = \sum_{r=1}^R \frac{\lambda_r p_{j,r}^2}{s_{j,j}}, \quad (2.18)$$

siendo \mathbf{h}_j el vector que representa la variable j en el espacio definido por $\bar{\mathbf{P}}$.

El grado de correlación entre la j -ésima variable y la r -ésima CP puede obtenerse mediante la ecuación (2.19):

$$corr_{j,r} = \left(\frac{\lambda_r}{s_{j,j}} \right)^{\frac{1}{2}} p_{j,r}, \quad (2.19)$$

donde $p_{j,r}$ es la j -ésima componente del r -ésimo autovector de \mathbf{S} . Si el modelo se construye en base a la matriz de correlación $s_{j,j} = 1$.

Un punto central que debe considerarse cuando se utilizan técnicas de proyección en espacios latentes es la determinación del número de componentes que deberán ser retenidas para lograr una representación adecuada del sistema en el nuevo espacio de dimensión R . Existe una gran variedad de métodos para la determinación de R . Los criterios en los que se basan los mismos son a veces muy diferentes por lo que generalmente, dichos criterios no se satisfacen de manera simultánea. En el Apéndice 1 se explican en detalle algunos de los métodos más utilizados en el área de SPC para fijar R .

2.2.2 Implementación de la Estrategia de Monitoreo

2.2.2.1 Etapa I

a. Pretratamiento de los Datos

Dada una base de datos históricos del proceso que contiene I_0 mediciones de su operación en condiciones normales, el objetivo del pretratamiento es detectar la presencia de observaciones cuyo comportamiento no sea acorde al del resto de los datos del conjunto. Esta tarea permite refinar aún más la población que se tomará como referencia. Dado que las I observaciones remanentes resultan mucho más consistente entre sí, éstas definen el conjunto de Condiciones de Operación Normal (*NOC*) utilizado para formular el modelo del proceso.

En principio, las columnas de la matriz de datos crudos \mathbf{X}^0 , de dimensión $(I_0 \times J)$, se centran alrededor de las correspondientes medias muestrales. Dado que las mediciones son, por lo general, de naturaleza diferente (temperaturas, presiones, composiciones, caudales, etc.), éstas deben ser luego adimensionalizadas con el objeto de realizar su combinación lineal. El procedimiento de adimensionalización más frecuente consiste en dividir los valores de cada columna de \mathbf{X}^0 por su correspondiente desvío estándar (s_j). Otra alternativa es estandarizar cada variable según el rango de la misma (r_j) en la *PR*:

$$r_j = \max(x_j) - \min(x_j). \quad (2.20)$$

El próximo paso consiste en seleccionar los individuos de la base de datos que formarán la *PR*. Con tal propósito se han desarrollado diferentes metodologías, siendo los métodos de validación cruzada los más utilizados. Estos consisten en realizar sucesivamente los siguientes pasos: excluir un conjunto de individuos de la base de datos, modelar el proceso sin ellos y determinar la pertenencia de los individuos excluidos a la población remanente mediante la

aplicación de tests de hipótesis estadísticos. En función de los resultados de los sucesivos pasos se determina el conjunto de observaciones que representan la *NOC*. Un caso extremo de los métodos de validación cruzada es el método *leave one out*; en éste se excluye una sola observación por vez de la base de datos.

Una desventaja de los métodos de validación cruzada es que el orden en el cual las observaciones son eliminadas es de naturaleza combinatoria. En consecuencia, el hecho de que un individuo de la base de datos pertenezca o no a la *PR* depende del orden en que las observaciones son seleccionadas para el testeo. La Figura 2.3 (a) muestra esta situación. Se consideró una base de datos compuesta por 50 individuos, se aplicó el método *leave one out* y se representaron los valores del estadístico utilizado en el test de hipótesis (Q) y su valor límite o crítico. En la figura puede notarse que las observaciones indicadas con los números 12, 14, 16, 27 y 29 se encuentran por sobre los límites de confianza establecidos para la población. Sin embargo, si sólo se eliminan las observaciones 12, 14, 27 y 29, la observación 16 (ahora en la posición 14) no se señala como un individuo atípico o *outlier*, tal como se observa en la Figura 2.3 (b). Así, la pertenencia de dicha observación a la *PR* dependerá de si la misma fue eliminada o no antes que las demás observaciones.

A pesar de las limitaciones mencionadas anteriormente, los métodos de validación cruzada son ampliamente utilizados durante la *Etapa I* debido a la simplicidad del concepto y la facilidad de implementación.

Una alternativa para disminuir el efecto combinatorio en el orden de la selección de los individuos es permutar aleatoriamente la ubicación de los mismos en \mathbf{X}^0 , y ejecutar varias veces el algoritmo de eliminación hasta que los individuos remanentes en la población se mantengan sin variación por un determinado número de iteraciones.

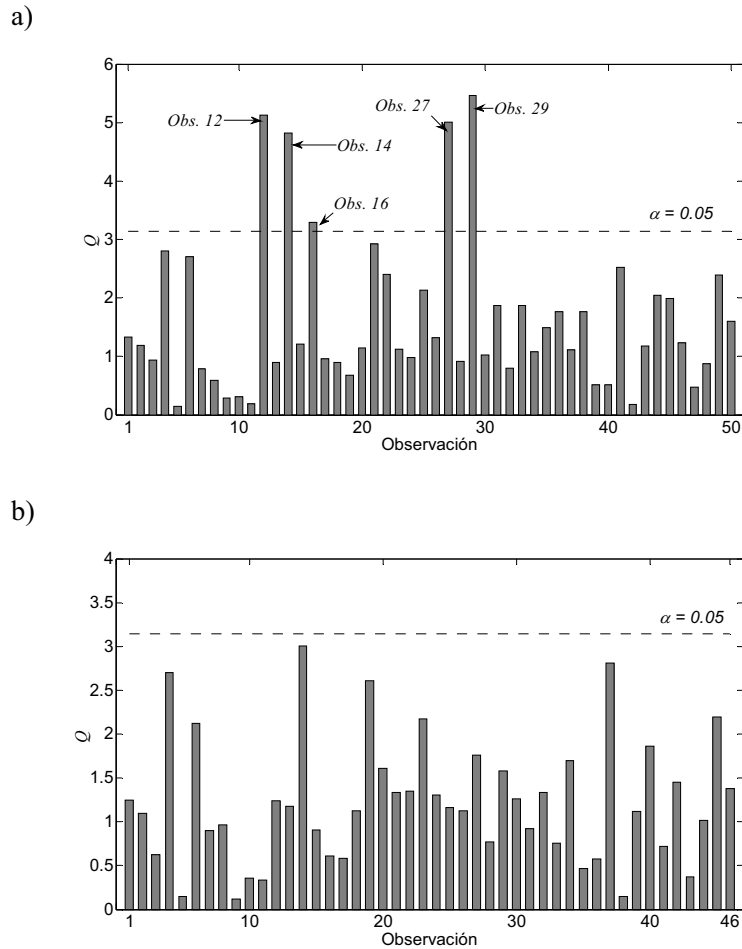


Figura 2.3: Dependencia de la eliminación de los individuos con el orden de selección

Las estrategias de monitoreo estadístico (*ICA*, *PLS*, etc) utilizan la metodología general, previamente descrita, para la constitución de la *PR*. En cada estrategia de proyección se introducen las variaciones asociadas al modelo formulado y sus correspondientes estadísticos.

b. Modelado

La *PR* representada por la matriz \mathbf{X} de dimensión $(I \times J)$ se emplea para la generación del modelo del proceso en el espacio de las variables latentes. En primera instancia se determina el número de *CPs* retenidas y luego se calculan las matrices \mathbf{P} , \mathbf{T} y \mathbf{E} que definen el modelo, según lo explicado en la Sección 2.2.1.

c. *Cálculo de los valores límites para los estadísticos*

Como se mencionó previamente, la detección de eventos anormales durante la *Etapa II* se efectúa mediante la aplicación de tests de hipótesis estadísticos. Los conceptos fundamentales de esta herramienta estadística se revisan en el Apéndice 2. El contraste entre la hipótesis nula (H_0 : no existen eventos anormales) y la alternativa (H_1 : existen eventos anormales) se realiza mediante la comparación de los valores de un conjunto de métricas con sus respectivos valores límites. Estos se calculan teniendo en cuenta: la naturaleza de la *PR*, el número de individuos considerados en ella (I), el número de variables o descriptores analizados (J) y el nivel de significación elegido para el test (α). Los estadísticos empleados en los tests de hipótesis dependen del tipo de análisis que se desea realizar (tipo de hipótesis que se desea testear) y del sistema particular considerado.

Las estrategias de monitoreo de procesos basadas en técnicas de proyección en variables latentes (*VLs*) hacen uso de un conjunto de estadísticos que funcionan en forma complementaria.

El estadístico de Hotelling (T^2) es el más empleado en el estudio de sistemas multivariados. Hotelling (1931) lo propuso como una generalización de la distribución t de Student al caso multivariado. Este estadístico es una forma cuadrática, y se representa frecuentemente como T^2 haciendo referencia a dicha condición y a su relación con la métrica de Student.

El estadístico T^2 se basa en la distancia de Mahalanobis, lo que le confiere la capacidad de evaluar cambios en la estructura de correlación del sistema. Permite, por ejemplo, comparar dos poblaciones entre sí o decidir si una observación particular puede o no considerarse como parte de una *PR* determinada. En relación con el monitoreo de procesos, el valor de T^2 es

una medida de la forma en la que una nueva observación viola la estructura de correlación observada en la *PR*. Por lo tanto, resulta útil para evaluar si las desviaciones que se observan en un proceso pueden o no considerarse dentro de la variabilidad normal del mismo.

Sea el vector de observaciones $\mathbf{x}_k^0 \in \mathbb{R}^J$, el correspondiente valor del estadístico T_k^2 se calcula como:

$$T_k^2 = \left(\mathbf{x}_k^{0T} - \bar{\mathbf{x}}^T \right) \mathbf{S}^{-1} \left(\mathbf{x}_k^0 - \bar{\mathbf{x}} \right), \quad (2.21)$$

donde $\bar{\mathbf{x}}$ es la media poblacional de \mathbf{X} y \mathbf{S} es la matriz de covarianza. Si T^2 se plantea en términos de la matriz de correlación \mathbf{Z} la ecuación (2.21) se debe escribir como:

$$T_k^2 = \mathbf{x}_k^T \mathbf{Z}^{-1} \mathbf{x}_k, \quad (2.22)$$

donde \mathbf{x}_k es la observación previamente estandarizada (centrada y escalada por su desvío estándar).

Considerando las ecuaciones (2.9), (2.10) y (2.13), y luego de cierto trabajo algebraico, puede obtenerse la ecuación (2.23):

$$T_k^2 = \mathbf{x}_k^T \mathbf{Z}^{-1} \mathbf{x}_k = \mathbf{x}_k^T \left(\bar{\mathbf{P}} \bar{\Lambda} \bar{\mathbf{P}}^T \right)^{-1} \mathbf{x}_k = \left(\mathbf{x}_k^T \bar{\mathbf{P}} \right) \bar{\Lambda}^{-1} \left(\bar{\mathbf{P}}^T \mathbf{x}_k \right) = \bar{\mathbf{t}}_k^T \bar{\Lambda}^{-1} \bar{\mathbf{t}}_k. \quad (2.23)$$

y una vez establecido el número de componentes a retener (R), la ecuación (2.23) se reordena como:

$$T_k^2 = \mathbf{x}_k^T \mathbf{Z}^{-1} \mathbf{x}_k = \begin{bmatrix} \mathbf{t}_k \\ \tilde{\mathbf{t}}_k \end{bmatrix}^T \begin{bmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \tilde{\Lambda} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{t}_k \\ \tilde{\mathbf{t}}_k \end{bmatrix}, \quad (2.24)$$

$$T_k^2 = \mathbf{t}_k^T \Lambda^{-1} \mathbf{t}_k + \tilde{\mathbf{t}}_k^T \tilde{\Lambda}^{-1} \tilde{\mathbf{t}}_k = D_k + T_{H_k}^2, \quad (2.25)$$

donde D_k es el valor del estadístico de Hotelling para la observación \mathbf{x}_k en el espacio de las VLs retenidas (\mathbb{L}^R) mientras que $T_{H_k}^2$ es el definido en el espacio de las VLs excluidas (\mathbb{L}^{J-R}). Este último es conocido como el estadístico de Hawkins (Jackson, 1991; Qin, 2003) y puede interpretarse como una suma ponderada de los errores de representación (Jackson, 1991).

El problema más importante asociado al cálculo de T_H^2 está relacionado al hecho que, en algunos casos, los autovalores contenidos en $\tilde{\Lambda}$ pueden ser muy cercanos a cero, causando problemas de mal condicionamiento.

Así, mientras el estadístico D detectará cambios en la estructura de correlación de los datos, T_H^2 analizará la magnitud de los errores de representación. Es de esperar, entonces, que violaciones a la estructura de correlación de \mathbf{Z} sean detectadas por D mientras que observaciones con un comportamiento muy diferente al usual sean detectadas por T_H^2 , dado que las mismas no están bien representadas en \mathbb{L}^R .

Puesto que \mathbf{x} , $\bar{\mathbf{t}}$, \mathbf{t} y $\tilde{\mathbf{t}}$ tienen todos media $\mathbf{0}$, tanto D como T_H^2 son siempre positivos excepto cuando $\mathbf{x} = \mathbf{0}$. Debido a ello, para estas métricas sólo se utiliza el límite de control superior (UCL).

Dado que las CPs resultan de una combinación lineal de las variables usadas como descriptores en \mathbf{X} y considerando el Teorema Central del Limite, se puede asumir que las CPs tienen una distribución más parecida a la normal que las variables originales en \mathbf{X} (Nomikos y MacGregor, 1994; Kourti y MacGregor, 1995; Nomikos y MacGregor, 1995). Si \mathbf{t}_k y $\tilde{\mathbf{t}}_k$ presentaran dicho comportamiento, las distribuciones para D y T_H^2 resultan (Qin, 2003):

$$D_k = \mathbf{t}_k^T \mathbf{\Lambda}^{-1} \mathbf{t}_k \approx \left(\frac{R(I^2 - 1)}{I(I - R)} \right) \mathcal{F}_{R, I-R, \alpha}, \quad (2.26)$$

$$T_{H_k}^2 = \tilde{\mathbf{t}}_k^T \tilde{\mathbf{\Lambda}}^{-1} \tilde{\mathbf{t}}_k \approx \left(\frac{(J - R)(I^2 - 1)}{I(I - (J - R))} \right) \mathcal{F}_{(J-R), I-(J-R), \alpha}, \quad (2.27)$$

donde \mathcal{F}_{g_1, g_2} es la función de distribución \mathcal{F} con g_1 grados de libertad (*g.l.*) en el numerador y g_2 *g.l.* en el denominador. El subíndice α indica el nivel de significación que fijará los valores críticos D_C y $T_{H_C}^2$.

El estadístico de Hawkins es poco utilizado cuando se emplean técnicas de monitoreo basadas en *PCA*. Esto se debe a los problemas de condicionamiento mencionados anteriormente. En general, los errores asociados a las predicciones se miden mediante el error cuadrático de representación (*SPE*):

$$SPE_k = \|\mathbf{e}_k\|^2 = \|(\mathbf{I} - \mathbf{P}\mathbf{P}^T) \mathbf{x}_k\|^2. \quad (2.28)$$

A diferencia de T_H^2 , *SPE* da la misma importancia a los errores de representación en todas las variables y no sufre los problemas de mal condicionamiento de T_H^2 .

Existen dos formas de calcular los valores críticos para el estadístico *SPE*. La primera fue propuesta por Box (1954), y se muestra en la ecuación (2.29).

$$SPE_C = g \chi_{h, \alpha}^2 \approx \frac{v}{2m} \chi_{2m^2/v, \alpha}^2, \quad (2.29)$$

donde m y v son la media y varianza, respectivamente, de los valores de SPE obtenidos para cada uno de los individuos de la PR . La segunda, introducida posteriormente por Jackson y Mudholkar (1979), es:

$$SPE_C = \theta_1 \left[1 - \frac{\theta_2 h_0 (1 - h_0)}{\theta_1^2} + \frac{z_\alpha (2\theta_2 h_0^2)^{1/2}}{\theta_1} \right]^{1/h_0}, \quad (2.30)$$

$$\theta_m = \sum_{r=R+1}^J \lambda_r^m \quad m = 1, 2, 3, \quad (2.31)$$

$$h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}, \quad (2.32)$$

donde z_α es la coordenada positiva para una probabilidad acumulada $(1 - \alpha)$ en la distribución $\mathcal{N}(0,1)$.

Los límites propuestos por Jackson y Mudholkar (1979) se utilizan generalmente en la *Etapa I*, con el objeto de estimar los errores de representación de las observaciones de la base de datos empleada para generar la PR . La fórmula propuesta por Box (1954) se aplica usualmente en la etapa de monitoreo en línea (*Etapa II*), dada su mayor simplicidad de cálculo (Nomikos y MacGregor, 1995).

Qin (2003) resalta los roles asimétricos de los estadísticos D y SPE en la detección de fallas, y analiza en forma cualitativa el tipo de eventos que podría ser indicado por cada uno de éstos. La Figura 2.4 muestra en forma esquemática el tipo de fallas que cada métrica debería detectar. En la misma se considera el espacio definido por las variables x_1 y x_2 y se indica la dirección de la primera CP . El rol del estadístico D es detectar fallas que rompan la estructura de correlación calculada a partir de los datos de referencia (**A**), mientras que SPE

sirve para controlar la distancia perpendicular entre una dada observación y el hiperespacio definido por las CPs retenidas (**B**).

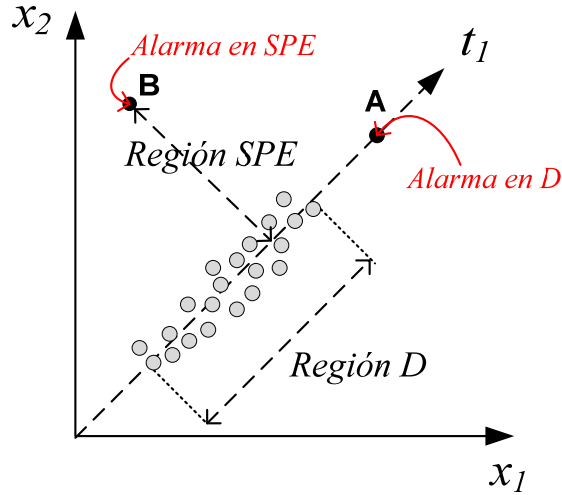


Figura 2.4: Esquema del tipo de fallas detectadas por *D* y *SPE*

2.2.2.2 Etapa II

La *Etapa II* es la fase de monitoreo en línea del proceso. Cada nueva observación \mathbf{x}_k^0 se centra, respecto de la media muestral $\bar{\mathbf{x}}$, y se estandariza utilizando la misma técnica aplicada a los datos de la *PR*.

Una vez obtenida \mathbf{x} , se calculan sus coordenadas en el espacio de las CPs (\mathbf{t}), los valores reconstruidos ($\hat{\mathbf{x}}$) y el error de representación (\mathbf{e}), como se muestra en las ecuaciones (2.33) a (2.35).

$$\mathbf{t}_k = \mathbf{P}^T \mathbf{x}_k, \tag{2.33}$$

$$\hat{\mathbf{x}}_k = \mathbf{P} \mathbf{P}^T \mathbf{x}_k = \mathbf{C} \mathbf{x}_k, \tag{2.34}$$

$$\mathbf{e}_k = \tilde{\mathbf{P}} \tilde{\mathbf{P}}^T \mathbf{x}_k = (\mathbf{I} - \mathbf{P} \mathbf{P}^T) \mathbf{x}_k = (\mathbf{I} - \mathbf{C}) \mathbf{x}_k. \tag{2.35}$$

Estas cantidades se emplean para el cálculo de los estadísticos usados para testear la hipótesis de que el proceso se encuentra operando en condiciones normales.

Si alguno de los estadísticos cae fuera de los límites correspondientes para tres observaciones consecutivas, el proceso se declara como fuera de control. Una vez detectado el estado anormal se hace necesario identificar cuál o cuáles son las variables medidas que indicaron tal condición. La Figura 2.5 muestra la estructura de las estrategia de monitoreo durante la *Etapa II*.

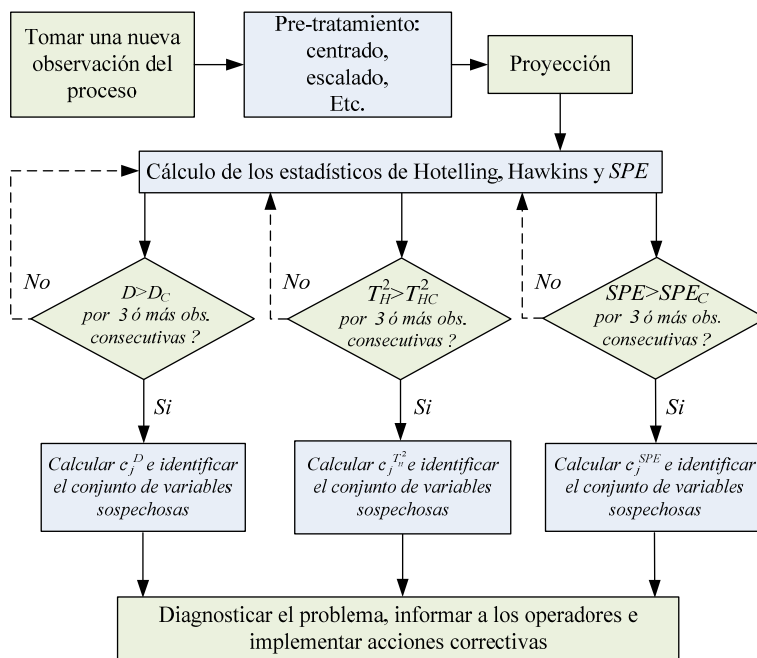


Figura 2.5: Representación esquemática de la metodología de monitoreo en línea

Se han propuesto diferentes metodologías para estimar la influencia de cada una de las variables sobre el valor final de los estadísticos considerados. A continuación se describen algunas de las más importantes para cada uno de ellos.

a. Estadístico D

Jackson (1991) propuso descomponer al estadístico D como la suma de las CPs retenidas en el análisis y realizar la identificación de las variables en términos de la importancia relativa de cada una (o de un conjunto de ellas) en las CPs dominantes.

La desventaja principal de este enfoque es que, en la gran mayoría de las aplicaciones industriales, no puede asociarse un significado físico claro a las CPs . Por lo tanto, resulta difícil identificar la variable que indica la falla.

MacGregor y col. (1994) plantearon evaluar la contribución de cada variable a las CPs que se encuentran fuera de la región de confianza. Los límites para cada CP se obtienen mediante las ecuaciones (2.36) y (2.37), bajo la suposición de que las CPs son aproximadamente normales.

$$UCL_r = \bar{t}_r + t_{I-1, \frac{\alpha}{2}} \sigma_{t_r} \left(I + \frac{I}{I} \right)^{\frac{1}{2}}, \quad (2.36)$$

$$LCL_r = \bar{t}_r - t_{I-1, \frac{\alpha}{2}} \sigma_{t_r} \left(I + \frac{I}{I} \right)^{\frac{1}{2}}, \quad (2.37)$$

donde UCL_r y LCL_r son los límites de control superior e inferior para la r -ésima CP (notar que el valor medio de todas las CPs es $\bar{t}_r = 0$), $t_{I-1, \alpha/2}$ es el valor de corte para la distribución de Student con $(I-1)$ g.l. para un nivel de significación α y σ_{t_r} es el desvío estándar de t_r en la PR .

El siguiente paso es el cálculo de la contribución de la j -ésima variable a dicha CP $(c_{j,k}^r)$, como muestra la ecuación (2.38):

$$c_{j,k}^r = x_{j,k} \mathbf{P}_{(j,r)}. \quad (2.38)$$

La variable, o grupo de variables, cuya contribución es más importante se indica como sospechosa de causar el elevado valor de D .

Si el conjunto de vectores que define la matriz de cambio de coordenadas no fuese ortonormal (como ocurre en *ICA*) o si los *SCORES* estuvieran correlacionados, es posible que el estadístico D detecte una falla pero ninguna de las *CPs* esté fuera de los límites de control o viceversa. En este último caso el operador no notará la ocurrencia de una situación anómala (Westerhuis y col., 2000).

Nomikos (1996) y Miller y col. (1998) propusieron el cálculo de las contribuciones de cada variable (*CVs*) al valor final del estadístico D , en lugar de ponderarlas por su participación en las *CPs* dominantes. En este caso, las posibles correlaciones existentes entre los *SCORES* tampoco se toman en cuenta en la estimación de las *CVs*. Además, los \mathbf{p}_r se asumen ortonormales ($\mathbf{P}^T \mathbf{P} = \mathbf{I}$). La expresión propuesta para el cálculo de las *CVs* es la siguiente:

$$c_{j,k}^D = \sum_{r=1}^R a_{r,r} t_{r,k} x_{j,k} p_{j,r}, \quad (2.39)$$

donde $a_{r,r}$ es el r -ésimo elemento de la diagonal de \mathbf{S}^{-1} .

Westerhuis y col. (2000) extendieron esta metodología al caso de espacios latentes donde las direcciones no son necesariamente ortonormales e introdujeron algunas nociones para el cálculo de los límites para cada contribución. Esta alternativa permite también tener en cuenta las posibles correlaciones entre las *CPs*:

$$D_k = \mathbf{t}_k^T \mathbf{S}^{-1} \mathbf{t}_k = \mathbf{t}_k^T \mathbf{S}^{-1} \left[\mathbf{x}_k^T \mathbf{P}^T (\mathbf{P}^T \mathbf{P})^{-1} \right]^T, \quad (2.40)$$

$$D_k = \mathbf{t}_k^T \mathbf{S}^{-1} \sum_{j=1}^J \left[x_{j,k} \mathbf{P}_{(j,:)} (\mathbf{P}^T \mathbf{P})^{-1} \right]^T, \quad (2.41)$$

$$D_k = \sum_{j=1}^J \mathbf{t}_k^T \mathbf{S}^{-1} \left[x_{j,k} \mathbf{P}_{(j,:)} (\mathbf{P}^T \mathbf{P})^{-1} \right]^T = \sum_{j=1}^J c_{j,k}^D, \quad (2.42)$$

donde $\mathbf{P}_{(j,:)}$ es la j -ésima fila de \mathbf{P} , \mathbf{S} es la matriz de covarianza de \mathbf{T} y $c_{j,k}^D$ es la contribución de la j -ésima variable de la observación k al estadístico D . Dada su definición, $c_{j,k}^D$ puede tomar valores negativos. Este aspecto merece especial consideración y se discute en detalle en el Capítulo 4. En esta tesis, las contribuciones definidas en la ecuación (2.42) se emplearán para calcular las *CVs*, a no ser que se indique expresamente lo contrario.

b. Estadístico T_H^2

Dadas las similitudes entre D y T_H^2 resulta lógico evaluar las *CVs* a este último de manera análoga a como que se explicó en el punto anterior.

$$T_{H_k}^2 = \sum_{j=1}^J \tilde{\mathbf{t}}_k^T \tilde{\mathbf{S}}^{-1} \left[x_{j,k} \tilde{\mathbf{P}}_{(j,:)} (\tilde{\mathbf{P}}^T \tilde{\mathbf{P}})^{-1} \right]^T = \sum_{j=1}^J c_{j,k}^{T_H^2}. \quad (2.43)$$

c. Estadístico SPE

Las *CVs* al estadístico SPE resultan más sencillas de calcular e interpretar que las anteriores, puesto que esta métrica no considera la correlación entre los errores y sólo evalúa la magnitud de los residuos en cada variable.

$$SPE_k = \|\mathbf{e}_k\|^2 = \|(\mathbf{I} - \mathbf{P}\mathbf{P}^T) \mathbf{x}_k\|^2 = \sum_{j=1}^J e_{j,k}^2 = \sum_{j=1}^J c_{j,k}^{SPE}. \quad (2.44)$$

Así, las variables cuyas c_j^{SPE} son más importantes revelan una situación anómala.

Westerhuis *y col.* (2000) propusieron usar límites para las *CVs*, los cuales se calculan a partir de las *CVs* de los datos considerados en la *PR*. De esta forma se tiene en cuenta que existen, de manera natural, diferencias considerables en las magnitudes de las *CVs* al valor del estadístico. La estimación de los valores críticos para las *CVs* se hace de la siguiente manera:

$$c_j^D \Big|_C = \overline{c_j^D} + \kappa \sigma_{c_j^D}, \quad (2.45)$$

donde $c_j^D \Big|_C$ es el valor crítico para las contribuciones de la j -ésima variable, $\overline{c_j^D}$ es el valor promedio de las contribuciones de la misma variable en la *PR* y $\sigma_{c_j^D}$ es el desvío estándar correspondiente. El valor de la constante κ puede ser definido por el usuario de acuerdo a las necesidades del caso. Si las *CVs* estuvieran normalmente distribuidas, un valor adecuado puede ser $\kappa = 3$ (Westerhuis *y col.*, 2000). Para las c_j^{SPE} los límites se establecen en términos de la distribución χ^2 utilizada para calcular los valores críticos de *SPE*

$$c_j^{SPE} \Big|_C = \frac{\sigma_{c_j^{SPE}}^2}{2\overline{c_j^{SPE}}} \chi^2 \left(2\left(\overline{c_j^{SPE}}\right)^2 / \sigma_{c_j^{SPE}}^2, \alpha \right), \quad (2.46)$$

donde $\overline{c_j^{SPE}}$ es el valor medio de las *CVs* en la *PR* y $\sigma_{c_j^{SPE}}^2$ es la varianza asociada a la misma.

Los valores límites deben tomarse sólo como orientativos, puesto que puede ocurrir que las suposiciones relacionadas a sus distribuciones no sean adecuadas. Una alternativa para solucionar este problema es la estimación de dichas distribuciones mediante métodos tales como el *KDE* (*Kernel Density Estimation*), detallado en el Apéndice 3.

2.3 Monitoreo Basado en el Análisis de Componentes Independientes

2.3.1 Análisis de Componentes Independientes (ICA)

Esta metodología, proveniente del área de reconocimiento de imágenes, se ha utilizado para el monitoreo de procesos en casos donde los datos de entrada \mathbf{X}^0 no responden a una distribución normal. Si bien *PCA* no presenta ninguna limitación para el tratamiento de datos no-normales, muchos autores señalaron que *ICA* es más adecuado para el tratamiento de los mismos y que muestra un mejor desempeño en la etapa de identificación de fallas (Hyvärinen y col., 2001; Lee y col., 2004; Yoo y col., 2004a).

Como su nombre lo indica, *ICA* está basado en la obtención de componentes estadísticamente independientes (*CI*s), lo que resulta ser una imposición mucho más severa que la simple correlación nula requerida a las *CP*s. A diferencia del *PCA*, *ICA* sí tiene restricciones relacionadas a la distribución de los datos de entrada. En este caso es necesario que las distribuciones de por lo menos $J - 1$ variables sean no-gaussianas. De no ser así, *ICA* no resulta una herramienta aplicable al conjunto de datos en cuestión. Otra diferencia importante es que las *CI*s no están sujetas a las restricciones de ortogonalidad requeridas a las *CP*s (ecuación (2.3)). Debido a esta última diferencia, algunos autores aseguran que las *CI*s son más apropiadas para representar variables con distribuciones no-normales (Hyvärinen y col., 2001; Lee y col., 2004; Yoo y col., 2004a).

Cuando las matrices de datos \mathbf{X} están formadas por datos normales, la ausencia de correlación entre los datos lograda mediante la aplicación de *PCA* es equivalente a la independencia. Por lo tanto, un modelo basado en *PCA* para estos datos dará como resultado un espacio de *CP*s estadísticamente independientes.

Consideremos el caso de dos señales independientes S_1 y S_2 uniformemente distribuidas en el intervalo $[-0,5, 0,5]$ (Figura 2.6 (a)), a las cuales se las combina entre sí de acuerdo a una matriz de mezcla \mathbf{M} . Las señales observadas (x_1 y x_2) serían similares a las mostradas en la Figura 2.6 (b). Los resultados que se obtienen al aplicar *PCA* e *ICA* se muestran en las Figura 2.6 (c) y (d), respectivamente.

En este ejemplo puede notarse que los resultados obtenidos con *ICA* reconstruyen las características subyacentes contenidas en las señales observadas (x_1 y x_2), en relación a las señales originales S_1 y S_2 , mucho mejor que los del *PCA*.

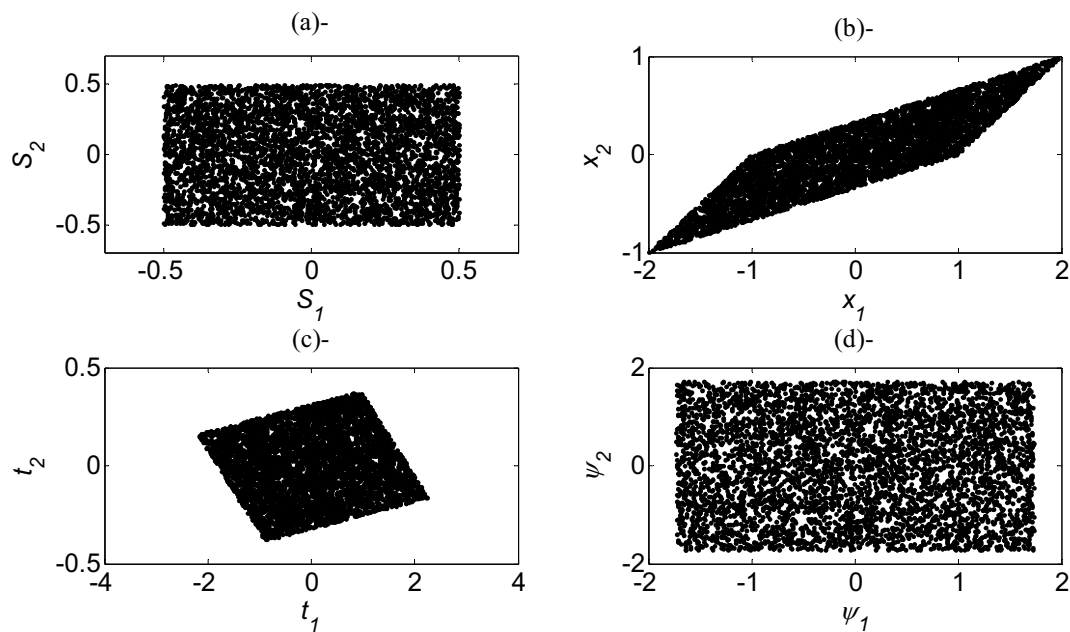


Figura 2.6: Comparación *ICA* vs. *PCA* para el caso de dos señales uniformemente distribuidas

ICA también se basa en un cambio de coordenadas que sirve para proyectar la matriz de datos \mathbf{X} en un espacio latente de dimensionalidad reducida. Se utiliza la transpuesta de \mathbf{X} como matriz de datos para el análisis. Siguiendo un procedimiento análogo al que se describió

para la obtención de las ecuaciones (2.13) a (2.15) del *PCA*, el modelo de *CI*s puede escribirse como:

$$\mathbf{X}^T = \bar{\mathbf{M}}\bar{\Psi} = \begin{bmatrix} \mathbf{M} & \tilde{\mathbf{M}} \end{bmatrix} \begin{bmatrix} \Psi \\ \tilde{\Psi} \end{bmatrix} = \mathbf{M}\Psi + \tilde{\mathbf{M}}\tilde{\Psi} = \mathbf{M}\Psi + \mathbf{E} = \hat{\mathbf{X}}^T + \mathbf{E}, \quad (2.47)$$

$$\bar{\Psi} = \bar{\mathbf{W}}\mathbf{X}^T, \quad (2.48)$$

$$\Psi = \mathbf{W}\mathbf{X}^T, \quad (2.49)$$

$$\hat{\mathbf{X}}^T = \mathbf{M}\mathbf{W}\mathbf{X}^T. \quad (2.50)$$

donde $\bar{\mathbf{M}} \in \mathbb{R}^{J \times J}$ es la matriz de mezcla de las señales independientes $\bar{\Psi}$, $\bar{\mathbf{W}} \in \mathbb{R}^{J \times J}$ es la matriz de rotación considerando todas las *CI*s, $\mathbf{W} \in \mathbb{R}^{A \times J}$ es la matriz del cambio de coordenadas cuando sólo se consideran A *CI*s y $\tilde{\mathbf{W}} \in \mathbb{R}^{(J-A) \times J}$ es el complemento de \mathbf{W} para dar $\bar{\mathbf{W}}$. Asimismo, $\bar{\Psi} \in \mathbb{R}^{J \times I}$ define las coordenadas de \mathbf{X}^T en el espacio completo de las *CI*s., $\Psi \in \mathbb{R}^{A \times I}$ representa las coordenadas de \mathbf{X}^T en el espacio reducido y $\tilde{\Psi} \in \mathbb{R}^{(J-A) \times I}$ contiene las del espacio excluido.

Las direcciones de la matriz de rotación $\bar{\mathbf{W}}$ pueden obtenerse mediante la maximización de la no-gaussianidad de los datos proyectados en el espacio que ésta define. Ésto se basa en el hecho de que la no-gaussianidad de un sistema da una medida de la independencia estadística de sus descriptores (Hyvärinen *y col.*, 2001).

Existen dos formas comunes de medir el grado de no-gaussianidad de un conjunto de datos: la Curtosis y la Negentropía. La *Curtosis* es el momento acumulado de cuarto orden de una distribución y se define como:

$$curt(x) = \mathcal{E}\{x^4\} - 3\mathcal{E}\{x^2\}^2. \quad (2.51)$$

Si se asume que la varianza de x es unitaria, la ecuación anterior se convierte en:

$$\text{curt}(x) = \mathcal{E}\{x^4\} - 3. \quad (2.52)$$

La Curtosis de una población normal es igual a cero, pero no lo es para casi todas las distribuciones no-gaussianas conocidas y puede tomar valores positivos o negativos.

Una variable aleatoria cuya Curtosis es menor que cero se conoce como variable sub-gaussiana ó platicúrtica, mientras que si la Curtosis toma valores positivos la variable se denomina supergaussiana ó leptocúrtica. Debido a que la Curtosis puede tomar valores positivos o negativos, generalmente se usa su valor absoluto o su cuadrado como medida de no-gaussianidad. La principal desventaja de la Curtosis como medida de no-gaussianidad es que la misma es muy sensible a la presencia de *outliers*.

En cuanto al concepto de Negentropía, éste se desprende del de entropía (concepto básico en la teoría de información), que se interpreta como la cantidad de información que puede obtenerse mediante la observación de las variables (Hyvärinen y col., 2001). Mientras más aleatorio e impredecible es el comportamiento de una variable, mayor es su entropía.

La entropía diferencial de un vector aleatorio \mathbf{x} con una función de densidad de probabilidad $f(\mathbf{x})$ se calcula como:

$$H(\mathbf{x}) = \int f(\mathbf{x}) \log(f(\mathbf{x})) d\mathbf{x}. \quad (2.53)$$

Un resultado importante es que una variable gaussiana tendrá la mayor entropía entre todas las distribuciones posibles con la misma varianza. Esto significa que la distribución normal es, en términos de información, la menos estructurada de todas las distribuciones posibles.

La Negentropía se define con el objetivo de conseguir una medida de no-gaussianidad que sea cero para la distribución normal y tenga valores positivos para todas las otras distribuciones posibles:

$$L(\mathbf{x}) = H(\mathbf{x}_{Gauss}) - H(\mathbf{x}), \quad (2.54)$$

donde \mathbf{x}_{Gauss} es un vector de variables normalmente distribuidas con la misma matriz de covarianza que \mathbf{x} .

La mayor desventaja de la Negentropía es que el cálculo a partir de su definición resulta muy engorroso puesto que es necesario estimar las funciones de densidad de probabilidad. Debido a esto, en la práctica se usan diferentes aproximaciones de la función de Negentropía como estimadores de no-gaussianidad. Por ejemplo, una de las formas clásicas para estimar la Negentropía, propuesta por Jones y Sibson (1987), es:

$$L(x) \approx \frac{1}{12} \mathcal{E}\{x^3\}^2 - \frac{1}{48} \text{curt}(x)^2. \quad (2.55)$$

En este caso la variable x tiene media cero y varianza unitaria. La validez de esta aproximación es bastante limitada puesto que la misma hereda la falta de robustez de la Curtosis respecto de los *outliers*. Una aproximación más robusta basada en el principio de máxima entropía fue presentada por Hyvärinen y col. (2001):

$$L(x) \approx \sum_{m=1}^N \kappa_m \left[\mathcal{E}\{G_m(x)\} - \mathcal{E}\{G_m(\nu)\} \right]^2, \quad (2.56)$$

donde κ_m es una constante positiva (que puede ser la misma para todos los términos), ν es una variable gaussiana con media cero y desvío unitario y G_m es alguna función cuadrática.

También es posible utilizar otras funcionalidades para G_m con el objetivo de hacer la estimación mucho más robusta, por ejemplo:

$$G_1(x) = \frac{1}{a_1} \log \{ \cosh(a_1 x) \}, \quad (2.57)$$

$$G_2(x) = -\exp \left\{ -\frac{a_2 x^2}{2} \right\}, \quad (2.58)$$

$$G_3(x) = x^4, \quad (2.59)$$

donde $1 \leq a_1 \leq 2$ y $a_2 \approx 1$.

De las tres funciones anteriores, la primera resulta ser la más adecuada para propósitos generales (Lee *et al.*, 2004). Aunque estas aproximaciones podrían no ajustar bien los perfiles de Negentropía teóricos, los resultados obtenidos mediante las mismas dan una idea comparativa lo suficientemente buena de la no-gaussianidad del conjunto de datos, cuando se los compara con ν (Hyvärinen *et al.*, 2001).

El primer paso del *ICA* (luego de la estandarización correspondiente) es el “blanqueado” de las señales. Esta operación consiste en una primera rotación del conjunto de datos en el espacio definido por una matriz $\bar{\mathbf{Q}}$, la cual se obtiene de la descomposición en valores singulares de la matriz \mathbf{S} .

$$\bar{\mathbf{Q}} = \bar{\mathbf{\Lambda}}^{-1/2} \bar{\mathbf{P}}^T, \quad (2.60)$$

donde $\bar{\mathbf{\Lambda}}$ y $\bar{\mathbf{P}}$ fueron definidas en la ecuación (2.9).

La base del espacio definida por la matriz $\bar{\mathbf{Q}}$ puede calcularse como combinación lineal de la base del espacio de *CPs*, por lo que podría decirse que el *PCA* es una etapa intermedia

del *ICA*. Al premultiplicar la matriz de datos transpuesta \mathbf{X}^T por la matriz de rotación $\bar{\mathbf{Q}}$, se obtiene:

$$\bar{\mathbf{Z}} = \bar{\mathbf{Q}}\mathbf{X}^T. \quad (2.61)$$

Luego, es posible escribir la ecuación (2.61) como:

$$\bar{\mathbf{Z}} = \bar{\mathbf{Q}}\mathbf{X}^T = \bar{\mathbf{Q}}\bar{\mathbf{M}}\bar{\mathbf{\Psi}} = \mathbf{B}\bar{\mathbf{\Psi}}. \quad (2.62)$$

La matriz de covarianza de $\bar{\mathbf{Z}}$ está dada por:

$$\mathcal{E}\{\mathbf{z}_i\mathbf{z}_i^T\} = \mathbf{B}\mathcal{E}\{\boldsymbol{\psi}_i\boldsymbol{\psi}_i^T\}\mathbf{B}^T = \mathbf{B}\mathbf{B}^T. \quad (2.63)$$

y finalmente, la matriz $\bar{\mathbf{W}}$ se calcula como:

$$\bar{\mathbf{W}} = \bar{\mathbf{B}}^T\bar{\mathbf{Q}}. \quad (2.64)$$

En esta tesis, el algoritmo *FastICA* propuesto por Hyvärinen y col. (2001) se utiliza para la determinación de la matriz \mathbf{B} . A partir de \mathbf{B} , dicho algoritmo estima tanto la matriz de rotación $\bar{\mathbf{W}}$ como la de los *SCORES* en el espacio de las *CI*s.

Al igual que *PCA*, la técnica *ICA* presenta la posibilidad de reducir el espacio de descriptores. En relación con este tema, Hyvärinen y col. (2001), en un contexto más general que el del monitoreo de procesos, indicaron una serie de ventajas del *ICA* en comparación con el *PCA*. Entre éstas destacan: la mayor facilidad de la interpretación de los resultados en los espacios reducidos, la minimización del efecto del ruido aleatorio y finalmente el incremento en la robustez del procedimiento, dada la posibilidad de que algunas *CI*s se superpongan (por ser una combinación lineal o estar muy cerca de serlo).

La determinación del número de *CI*s a retener (A) puede realizarse durante la etapa de blanqueado, o bien, luego de haber calculado las todas *CI*s. Hyvärinen y col. (2001), propusieron realizar la reducción en la etapa de blanqueado, y continuar el análisis con el conjunto reducido de variables. Dado que la etapa de blanqueado del *ICA* puede asimilarse a realizar un *PCA* sobre los datos de entrada, cualquiera de los métodos presentados en el Apéndice 1 puede utilizarse para seleccionar A .

Por otra parte, si todas las coordenadas son retenidas durante el blanqueado (y se considera el espacio completo expandido por $\bar{\mathbf{Q}}$), la reducción de las dimensiones del sistema se realiza después de obtener las direcciones para las J *CI*s. En este caso, deben tomarse en cuenta ciertas particularidades del método de obtención de $\bar{\mathbf{W}}$. A diferencia de lo que sucede en el *PCA*, las direcciones del espacio de *CI*s no tienen ningún criterio de ordenamiento particular. Por lo tanto, el primer paso para la reducción de la dimensionalidad en esta etapa del análisis, es el ordenamiento de las *CI*s según algún criterio. Entre los posibles criterios pueden mencionarse los que tienen en cuenta la no-gaussianidad de cada *CI*, la norma infinito de cada componente ($\|\boldsymbol{\psi}_i\|_\infty$) o su norma dos ($\|\boldsymbol{\psi}_i\|$) (Lee y col., 2004).

Una vez que se ordena el conjunto de *CI*s, la determinación de A debe realizarse aplicando algún criterio congruente con la metodología utilizada para el ordenamiento. Lee y col. (2004) señalaron que el uso combinado de $\|\boldsymbol{\psi}_i\|$ como criterio de ordenamiento y el *SCREE plot* para la determinación de A produce buenos resultados en el monitoreo de procesos.

2.3.2 Implementación de la Estrategia de Monitoreo

La metodología de monitoreo de procesos basada en *ICA* es similar a la que se explicó en la Sección 2.2.2. Si bien la aplicación de esta técnica en *SPC* es mucho más reciente que la de

PCA, los avances han sido mucho más rápidos ya que fue posible extrapolar fácilmente muchas mejoras y enfoques alternativos utilizados en *PCA*. A continuación se presentan las características distintivas de esta técnica para las *Etapas I y II*.

2.3.2.1 Etapa I

Los miembros de la *PR* se seleccionan siguiendo la metodología descrita en la Sección 2.2.2. Luego estos datos son utilizados para formular el modelo del proceso según se explicó en la Sección 2.3.1.

A fin de detectar la presencia de eventos anormales en la *Etapa II*, varios autores (Lee y col., 2004; Yoo y col., 2004b) propusieron el uso de tres estadísticos, cuyos límites o valores críticos se evalúan en la *Etapa I*.

El primer estadístico (I_d^2) es el homólogo del estadístico D utilizado en *PCA*. El segundo (I_e^2) es el correspondiente al estadístico de Hawkins. Finalmente, el tercero es el estadístico *SPE*, ya definido en la ecuación (2.28).

En el caso de *ICA* resulta interesante notar que $\mathcal{S}\{\Psi_i\Psi_i^T\} = \mathbf{I}$, debido al escalamiento usado en el blanqueado. A partir de este resultado, el estadístico I_d^2 se calcula directamente como (Lee y col., 2004):

$$I_{d_k}^2 = \Psi_k^T \Psi_k. \quad (2.65)$$

Yoo y col. (2004a) introdujeron el uso del estadístico (I_e^2) que se calcula mediante la siguiente expresión:

$$I_{e_k}^2 = \tilde{\Psi}_k^T \tilde{\Psi}_k. \quad (2.66)$$

Los subíndices d y e indican las partes “determinística” y “excluida” del modelo, cuyos espacios son expandidos por \mathbf{W} y $\tilde{\mathbf{W}}$, respectivamente.

Finalmente, el estadístico SPE se utiliza para monitorear la magnitud de los errores de representación. En este caso, SPE puede escribirse como:

$$SPE_k = \|\mathbf{e}_k\|^2 = \|(\mathbf{I} - \mathbf{Q}^{-1}\mathbf{B}\mathbf{W}^T)\mathbf{x}_k\|^2. \quad (2.67)$$

Tanto las $\boldsymbol{\psi}$ como las $\tilde{\boldsymbol{\psi}}$ son variables no-gaussianas y en general, las distribuciones asociadas son desconocidas. Por esto, los límites de confianza para cada uno de estos estadísticos deberán determinarse para cada aplicación particular mediante técnicas de estimación de funciones de densidad de probabilidad (*pdfs*). El método *KDE* es la técnica empleada en este trabajo de tesis para el cálculo de las *pdfs* de los estadísticos mencionados. En este método, la *pdf* de una única variable x , de la cual se cuenta con I repeticiones, se estima mediante la ecuación (2.68):

$$\hat{f}(x) = \frac{1}{Ih} \sum_{i=1}^I \mathcal{K} \left\{ \frac{x - x_i}{h} \right\}, \quad (2.68)$$

donde \mathcal{K} es una distribución conocida y h es un parámetro de suavizado. Una discusión más detallada sobre esta metodología puede consultarse en el Apéndice 3.

2.3.2.2 Etapa II

En la etapa de monitoreo en línea, cada nueva observación \mathbf{x}_k^0 es centrada y estandarizada tal como se explicó en la Sección 2.2.2.1 a fin de obtener \mathbf{x}_k .

Luego se obtienen las proyecciones para el caso de las observaciones individuales:

$$\boldsymbol{\psi}_k = \mathbf{W}\mathbf{x}_k, \quad (2.69)$$

$$\bar{\boldsymbol{\Psi}}_k = \bar{\mathbf{W}}\mathbf{x}_k \quad (2.70)$$

$$\hat{\mathbf{x}}_k = \mathbf{Q}^{-1}\mathbf{B}\bar{\boldsymbol{\Psi}}_k = \mathbf{Q}^{-1}\mathbf{B}\bar{\mathbf{W}}\mathbf{x}_k, \quad (2.71)$$

$$\mathbf{e}_k = (\mathbf{I} - \mathbf{Q}^{-1}\mathbf{B}\bar{\mathbf{W}})\mathbf{x}_k, \quad (2.72)$$

y con ellas se calculan los estadísticos utilizados en el testeo de hipótesis. Si el valor de alguna de estas métricas supera su valor crítico para tres observaciones consecutivas, se declara el estado fuera de control. Surge entonces la necesidad de identificar las variables sospechosas, para lo cual se evalúan las *CVs* a los valores de los estadísticos.

Las contribuciones para las dos primeras métricas se calculan como (Lee y col., 2004):

$$\mathbf{c}_k^{I_d^2} = \frac{\mathbf{Q}^{-1}\mathbf{B}\boldsymbol{\Psi}_k}{\|\mathbf{Q}^{-1}\mathbf{B}\boldsymbol{\Psi}\|} \|\boldsymbol{\Psi}_k\|, \quad (2.73)$$

$$\mathbf{c}_k^{I_e^2} = \frac{\mathbf{Q}^{-1}\tilde{\mathbf{B}}\tilde{\boldsymbol{\Psi}}_k}{\|\mathbf{Q}^{-1}\tilde{\mathbf{B}}\tilde{\boldsymbol{\Psi}}\|} \|\tilde{\boldsymbol{\Psi}}_k\|. \quad (2.74)$$

Las *CVs* al estadístico *SPE* se obtienen con una expresión equivalente a la de la ecuación (2.44):

$$SPE_k = \|\tilde{\mathbf{x}}_k\|^2 = \left\| (\mathbf{I} - \mathbf{Q}^{-1}\mathbf{B}\bar{\mathbf{W}}^T)\mathbf{x}_k \right\|^2 = \sum_{j=1}^J e_{j,k}^2 = \sum_{j=1}^J c_{j,k}^{SPE}. \quad (2.75)$$

Al igual que en el caso del *PCA* las *CVs* a I_d^2 y I_e^2 pueden tomar valores negativos.

2.4 Monitoreo Basado en Cuadrados Mínimos Parciales

2.4.1 Cuadrados Mínimos Parciales (PLS)

Consideremos ahora el caso en el que se dispone de dos matrices de datos diferentes. La primera de ellas, $\mathbf{X} \in \mathbb{R}^{I \times J}$ es la matriz de datos de operación descrita en las secciones anteriores. La segunda, $\mathbf{Y} \in \mathbb{R}^{I \times N}$, está compuesta por un grupo de N variables asociadas a la

calidad final del producto. Las técnicas analizadas en las secciones anteriores centran la atención del análisis en la obtención de las variables latentes que expliquen la variabilidad presente en \mathbf{X} . Sin embargo, la posibilidad de usar la información contenida en \mathbf{X} para predecir las variables de calidad en \mathbf{Y} resulta de gran interés desde el punto de vista del control de procesos (Kourti y MacGregor, 1995). El método *PLS* es, sin lugar a dudas, una de las metodologías más ampliamente utilizadas con este propósito.

Si bien existen muchas variantes del método *PLS*, la más utilizada consiste en realizar el análisis de la matriz de covarianza ($\mathbf{S}_{\mathbf{XY}}$):

$$\mathbf{S}_{\mathbf{XY}} = \mathbf{X}^T \mathbf{Y} \mathbf{Y}^T \mathbf{X}. \quad (2.76)$$

El propósito fundamental es encontrar un nuevo espacio de variables latentes en el cual la correlación entre las direcciones que definen al mismo y las del espacio definido por \mathbf{Y} sea la máxima posible. En esta estrategia, al igual que para *PCA*, las coordenadas del espacio latente están ordenadas en términos de dicha correlación (Kourti y MacGregor, 1995), y las coordenadas del nuevo espacio son también conocidas como *SCORES*.

Los valores de las proyecciones sobre cada una de las direcciones del espacio latente se obtienen como combinación lineal de las coordenadas en el espacio original:

$$\mathbf{t}_1 = \mathbf{X}_1 \mathbf{w}_1, \quad (2.77)$$

donde \mathbf{w}_1 es el autovector asociado al primer autovalor de $\mathbf{S}_{\mathbf{XY}}$ y $\mathbf{X}_1 = \mathbf{X}$. Una vez obtenido \mathbf{t}_1 se calculan los llamados vectores de regresión:

$$\mathbf{p}_1 = \mathbf{X} \mathbf{t}_1 / \mathbf{t}_1^T \mathbf{t}_1. \quad (2.78)$$

Luego se calcula la matriz de residuos, $\mathbf{X}_2 = \mathbf{X}_1 - \mathbf{t}_1 \mathbf{p}_1^T$, a partir de la cual se repite el procedimiento. Este algoritmo se conoce como *NIPALS* (*Non Linear Iterative Partial Least*

Squares) y se utiliza también para el cálculo de *CPs*. El número de componentes a retener se determina generalmente por el método de validación cruzada explicado en el Apéndice 1.

2.4.2 Implementación de la Estrategia de Monitoreo

La adaptación de las estrategias expuestas en las Secciones 2.2.2 y 2.3.2 puede hacerse fácilmente. Cuando se obtiene una nueva observación del proceso, se calculan las proyecciones de las observaciones, los vectores de regresión y los residuos del modelo como se indica a continuación:

$$t_r = x^T \mathbf{w}_r \quad \forall r = 1, \dots, R, \quad (2.79)$$

$$\hat{\mathbf{y}} = \mathbf{P}\mathbf{t}, \quad (2.80)$$

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}. \quad (2.81)$$

2.5 Referencias

- Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis Wiley Publications in Statistics, New York, John Wiley & Sons, Inc.
- Box, G. E. P., "Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, I. Effect of Inequality of Variance in the One-Way Classification". *The Annals of Mathematical Statistics* **1954**, 25, (2), 290-302.
- Gertler, J., "Residual Generation in Model-Based Fault Diagnosis". *Control Theory and Advanced Technology* **1993**, 9, (1), 259-285.
- Hotelling, H., "The Generalization of Student's Ratio". *The Annals of Mathematical Statistics* **1931**, 2, (3), 360-378.
- Hyvärinen, A., J. Karhunen y E. Oja (2001). Independent Component Analysis, New York, Wiley-Interscience.
- Jackson, J. E. y G. S. Mudholkar, "Control Procedures for Residuals Associated with Principal Component Analysis". *Technometrics* **1979**, 21, (3), 341-349.
- Jackson, J. E. (1991). A User's Guide to Principal Components. Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics, New York, John Wiley & Sons, Inc.
- Jeffers, J. N. R., "Principal Component Analysis of Designed Experiment". *Statistician* **1962**, 12, (2), 230-242.
- Jones, M. C. y R. Sibson, "What Is Projection Pursuit". *Journal of The Royal Statistics Society, Series A* **1987**, 150, (1), 1-36.
- Kloek, T. y L. B. M. Mennes, "Simultaneous Equations Estimation Based on Principal Components of Predetermined Variables". *Econometrica* **1960**, 28, (1), 45-61.
- Kourti, T. y J. F. MacGregor, "Process Analysis, Monitoring and Diagnosis, Using Multivariate Projection Methods". *Chemometrics and Intelligent Laboratory Systems* **1995**, 28, (1), 3-21.
- Lee, J. M., C. Yoo y I. B. Lee, "Statistical Process Monitoring with Independent Component Analysis". *Journal of Process Control* **2004**, 14, (5), 467-485.
- MacGregor, J. F., C. Jaeckle, C. Kiparissides y M. Koutoudi, "Process Monitoring and Diagnosis by Multiblock PLS Methods". *AIChE Journal* **1994**, 40, (5), 826-838.
- Miller, P., R. E. Swanson y C. F. Heckler, "Contribution Plots: The Missing Link in Multivariate Quality Control". *Applied Mathematics Computer Science*. **1998**, 8, (4), 775-792
- Näf, U. G., "Stochastic Simulation Using Gproms". *Computers and Chemical Engineering* **1994**, 18, (SUPPL.), S743-S747.

- NIST/Sematech (2008). e-Handbook of Statistical Methods, URL: <http://www.itl.nist.gov/div898/hadbook/>
- Nomikos, P. y J. F. MacGregor, "Monitoring of Batch Process Using Multiway Principal Component Analysis". *AIChE Journal* **1994**, 40, (8), 1361-1375.
- Nomikos, P. y J. F. MacGregor, "Multivariate SPC Charts for Monitoring Batch Processes". *Technometrics* **1995**, 37, (1), 41-59.
- Nomikos, P., "Detection and Diagnosis of Abnormal Batch Operations Based on Multi-Way Principal Component Analysis". *ISA Transactions* **1996**, 35, (3), 259-266.
- Qin, S. J., "Statistical Process Monitoring: Basics and Beyond". *Journal of Chemometrics* **2003**, 17, (8-9), 480-502.
- Ruiz, D. (2001). Fault Diagnosis in Chemical Plants Integrated to the Information System. Tesis Doctoral, Ingeniería Química, Barcelona, Universitat Politècnica de Catalunya.
- Van Gerven, D. P. y G. B. Oakland, "Univariate and Multivariate Statistical Model in the Analysis of Human Sexual Dimorphism". *Statistician* **1973**, 22, (2), 256-268.
- Vavra, T. G., "Factor Analysis of Perceptual Change". *Journal of Marketing Research* **1972**, 9, (2), 193-199.
- Venkatasubramanian, V., R. Rengaswamy y S. N. Kavuri, "A Review of Process Fault Detection and Diagnosis Part II: Qualitative Models and Search Strategies". *Computers and Chemical Engineering* **2003a**, 27, (3), 313-326.
- Venkatasubramanian, V., R. Rengaswamy, K. Yin y S. N. Kavuri, "A Review of Process Fault Detection and Diagnosis Part I: Quantitative Model-Based Methods". *Computers and Chemical Engineering* **2003b**, 27, (3), 293-311.
- Westerhuis, J. A., S. P. Gurden y A. K. Smilde, "Generalized Contribution Plots in Multivariate Statistical Process Monitoring". *Chemometrics and Intelligent Laboratory Systems* **2000**, 51, (1), 95-114.
- Yélamos Ruiz, I. (2008). A Global Approach for Supporting Operators' Decision-Making Dealing with Plant Abnormal Events. Tesis Doctoral, Ingeniería Química, Barcelona, Universitat Politècnica de Catalunya.
- Yoo, C. K., J. M. Lee, P. A. Vanrolleghem y I. B. Lee, "On-line Monitoring of Batch Processes Using Multiway Independent Component Analysis". *Chemometrics and Intelligent Laboratory Systems* **2004a**, 71, (2), 151-163.
- Yoo, C. K., D. S. Lee y P. A. Vanrolleghem, "Application of Multiway ICA for On-line Process Monitoring of a Sequencing Batch Reactor". *Water Research* **2004b**, 38, (7), 1715-1732.

Página en Blanco

Capítulo 3

Monitoreo de Procesos Batch

3.1	Procesos Batch y Estructura de los Datos.....	74
3.2	Análisis Paralelo de Factores, Modelos Tucker y MPCA.....	77
3.3	Análisis Multimodo: MPCA, MICA, MPLS	79
3.3.1	<i>Método de Nomikos y MacGregor (NM)</i>	80
3.3.2	<i>Método de Wold, Kettaneh, Fridén, y Holmberg Original (WKFH) y Modificado (AT)</i>	87
3.3.3	<i>Modelos locales (ML)</i>	92
3.3.4	<i>Modelos Evolutivos (ME)</i>	95
3.3.5	<i>Modelos de Ventana Móvil (VM)</i>	99
3.3.6	<i>Análisis Dinámicos (DPCA y DPLS)</i>	104
3.4	Batches con Diferente Número de Observaciones.....	105
3.5	Referencias.....	107

Página en Blanco

3 Monitoreo de Procesos *Batch*

Los procesos por lotes o *batch* se utilizan principalmente para la fabricación de productos de alto valor agregado en pequeña o mediana escala, tales como: polímeros, productos farmacéuticos y biológicos, entre otros. El monitoreo en línea del progreso del *batch* y la predicción en tiempo real de las variables de calidad son herramientas que permiten llevar a cabo la operación del mismo de manera satisfactoria y segura (Kourti y MacGregor, 1995; Kourti *y col.*, 1995). La aplicación de estas técnicas puede producir información relevante relacionada al estado del proceso, permitiendo así, tomar acciones correctivas antes que el efecto de alguna desviación de la operación normal arruine el lote.

Diversos investigadores han realizado esfuerzos tendientes a adaptar las técnicas descritas en el Capítulo 2 para su aplicación a procesos *batch* (Wold *y col.*, 1987; Rännar *y col.*, 1998; Nomikos y MacGregor, 1994; Kourti *y col.*, 1995; Nomikos y MacGregor, 1995; Wold *y col.*, 1998; Chen y Liu, 2002; Lee *y col.*, 2003b; Yoo *y col.*, 2004a; Yoo *y col.*, 2004b; Ramaker *y col.*, 2005, etc.). Sin embargo, el monitoreo de estos sistemas todavía representa un gran desafío, tal como se mencionó en el Capítulo 1.

Existen varias estrategias disponibles para el análisis de las estructuras *multi-modo* (*MM*) que resultan de este tipo de procesos. Los enfoques clásicos como *PARAFAC* y *Tucker 1, 2 y 3*, fueron adaptados dando origen a las versiones *MM* de las técnicas descritas en el capítulo previo. Estas se conocen como: *Multiway Principal Component Analysis (MPCA)* (Wold *y col.*, 1987; Basseville, 1988; Nomikos y MacGregor, 1994, 1995; Wold *y col.*, 1998; Smilde *y col.*, 2004), *Multiway Independent Component Analysis (MICA)* (Lee *y col.*, 2003a,b; Yoo *y col.*, 2004a;b) y *Multiway Partial Least Squares (MPLS)* (Kourti *y col.*, 1995; Kourti y MacGregor, 1996). Además, Dong y McAvoy (1996) y Rännar *y col.* (1998) presentaron

adaptaciones de otros procedimientos a procesos *batch*, tales como el *PCA* no lineal (*NLPCA*), basado en redes neuronales, y el *PCA* jerárquico (*HPCA*), respectivamente. La aplicación de cualquiera de las estrategias mencionadas permite determinar el estado del proceso analizando la progresión temporal de cada nuevo *batch* mediante alguno de los modelos en *VLs*.

En las siguientes secciones se discutirán las características más importantes de las técnicas de monitoreo de procesos *batch*. En particular, se examinarán en profundidad los métodos más difundidos para el análisis de los arreglos tridimensionales de datos, señalando las ventajas y desventajas de cada uno de ellos. Se tendrán en cuenta: la facilidad de la elaboración de las cartas de control, el pre-procesamiento de las nuevas observaciones, las capacidades de detección e identificación de las fallas y la interpretación de los resultados.

3.1 Procesos *Batch* y Estructura de los Datos

Los procesos *batch* o *semi-batch* son procesos cuya operación tiene un comienzo y un final determinados. Las materias primas se colocan en el interior de los equipos, dentro de los cuales se someten a una serie de operaciones que permiten obtener un producto de características diferentes a las de las materias primas utilizadas.

Los pasos y procedimientos que deben seguirse durante la operación de cada *batch* (calentamientos, niveles de agitación, adición de reactivos, etc.) se encuentran definidos en la “*receta*” de la operación. En ella se detallan la duración de cada tarea, los niveles deseados de las variables de proceso y toda la información requerida para operar el proceso de manera segura, satisfactoria y reproducible. El final de la operación puede ser un parámetro fijo de la receta en cuestión o, depender del valor de algunas de las variables, que pueden o no ser medidas en línea.

La Figura 3.1 muestra los datos obtenidos al monitorear un proceso *batch* genérico de duración t_{op} . Se dispone de K observaciones de J variables cada una. Estas observaciones son tomadas desde el inicio de la operación a $t = 0$ ($k = 1$), hasta el final de la misma a $t = t_{op}$ ($k = K$), con un período de muestreo que, por simplicidad, asumiremos constante ($\tau = t_{op} / (K - 1)$). Las mediciones del proceso se almacenan frecuentemente en una base de datos que permite disponer de las mismas cuando sea necesario.

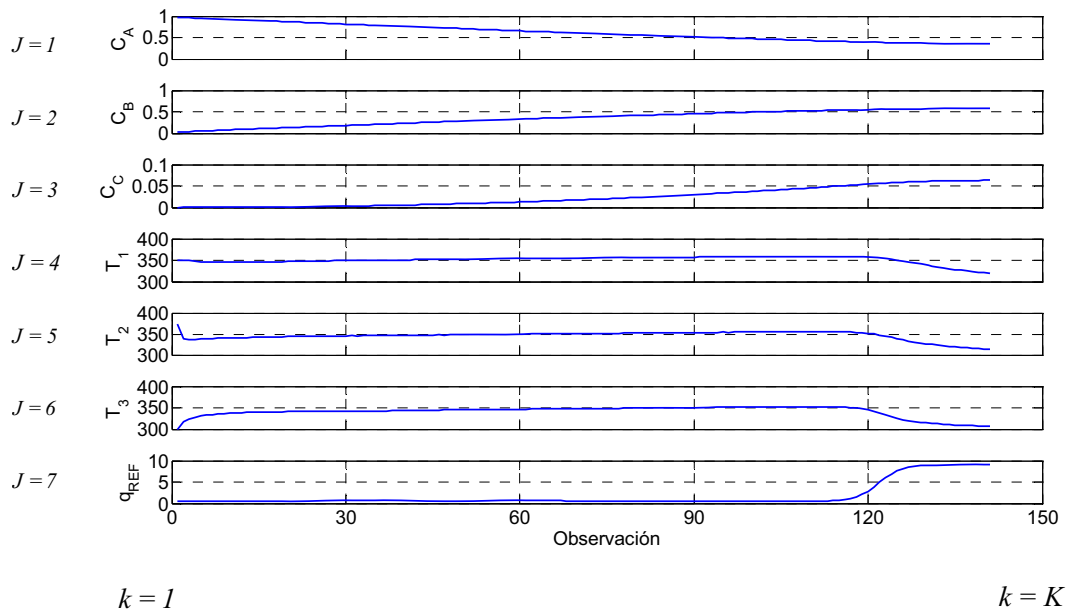


Figura 3.1: Trayectoria de las J variables monitoreadas en el proceso desde $k=1$ a $k=K$

Sea P el número total de realizaciones de un proceso que se encuentran almacenadas en una base de datos. Este conjunto de *batches* puede dividirse en dos grupos. El primero de ellos se forma con I_0 de los P *batches*, los cuales transcurrieron según se espera de la operación normal del proceso. El segundo se forma con los *batches* restantes, en los que tuvo lugar algún evento anormal. En cualquiera de estos dos conjuntos, cada uno de los *batches* puede ordenarse en una matriz \mathbf{X}_i^0 de K filas y J columnas

El primero de los conjuntos mencionados se usa en la *Etapa I* para formar la población de referencia (*PR*) que define las condiciones de operación normal (*NOC*). Los datos que la componen pueden reordenarse en un arreglo tridimensional (o tri-modo) $\underline{\mathbf{X}} \in \mathbb{R}^{I \times J \times K}$. La Figura 3.2 muestra la estructura de este arreglo. El primer modo representa los I *batches* en la *PR*, el segundo contiene las J variables (en cada una de las K observaciones para todos los *batches*) y el tercero contiene las K observaciones realizadas durante la operación de cada *batch*.

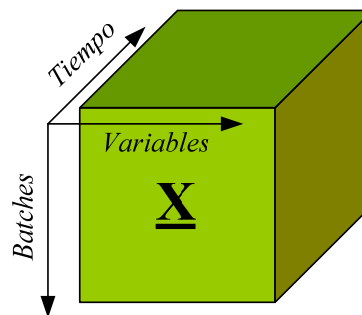


Figura 3.2: Estructura de las matrices de datos $\underline{\mathbf{X}} \in \mathbb{R}^{I \times J \times K}$

La Figura 3.3(a) muestra la ubicación del i -ésimo *batch* de la *PR* en el arreglo $\underline{\mathbf{X}}$. Esta capa horizontal del arreglo ($\underline{\mathbf{X}}_{(i,:,:)}, \forall i = 1, 2, \dots, I$) contiene las K observaciones correspondientes al *batch*. La Figura 3.3(b) representa las capas verticales que contienen la evolución temporal de cada una de las J variables para todos los *batches* ($\underline{\mathbf{X}}_{(:,j,:)}, \forall j = 1, 2, \dots, J$). Por último, la Figura 3.3(c) indica la disposición de las capas verticales que contienen las k -ésimas observaciones para todos los *batches* de la *PR* ($\underline{\mathbf{X}}_{(:, :, k)}, \forall k = 1, 2, \dots, K$).

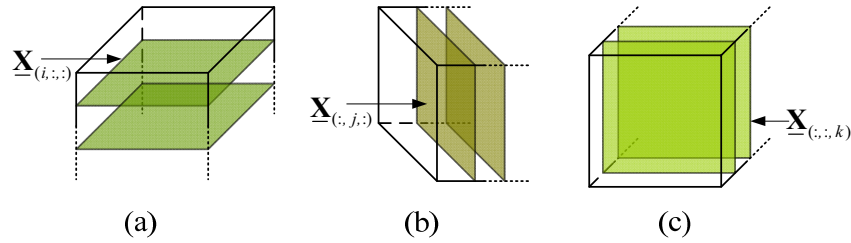


Figura 3.3: Disposición de los datos en los arreglos tridimensionales

La naturaleza tridimensional del arreglo $\underline{\mathbf{X}}$ impide la aplicación directa de las metodologías de análisis en estructuras latentes, tal como éstas se presentaron en el Capítulo 2. Por esa razón, y tal como se señaló en las secciones anteriores, durante los últimos años se dedicó un gran esfuerzo a la adaptación de las mismas para el análisis de procesos *batch* o *semi-batch*. Las principales estrategias propuestas con ese propósito se explican en las siguientes secciones.

3.2 Análisis Paralelo de Factores, Modelos *Tucker* y *MPCA*

El problema del análisis de estructuras *MM* se ha abordado desde distintas áreas de la ciencia. A lo largo de los años, se han propuestos numerosos métodos de análisis; cada uno de los cuales presenta ciertas ventajas y desventajas particulares. El Análisis Paralelo de Factores (*PARAFAC*) es una de las metodologías de análisis para arreglos de tres dimensiones y es originario de la psicometría. *PARAFAC* consiste en una descomposición multi-lineal del arreglo de datos $\underline{\mathbf{X}}$ en términos de las matrices \mathbf{A} , \mathbf{B} , y \mathbf{C} , según:

$$\underline{\mathbf{X}} = \sum_{f=1}^F \mathbf{a}_f \otimes \mathbf{b}_f \otimes \mathbf{c}_f + \underline{\mathbf{E}}, \quad (3.1)$$

$$x_{ijk} = \sum_{f=1}^F a_{if} b_{jf} c_{kf} + e_{ijk}, \quad (3.2)$$

donde F es el número de componentes o factores considerados. La Figura 3.4 muestra en forma gráfica las relaciones de las ecuaciones (3.1) y (3.2).

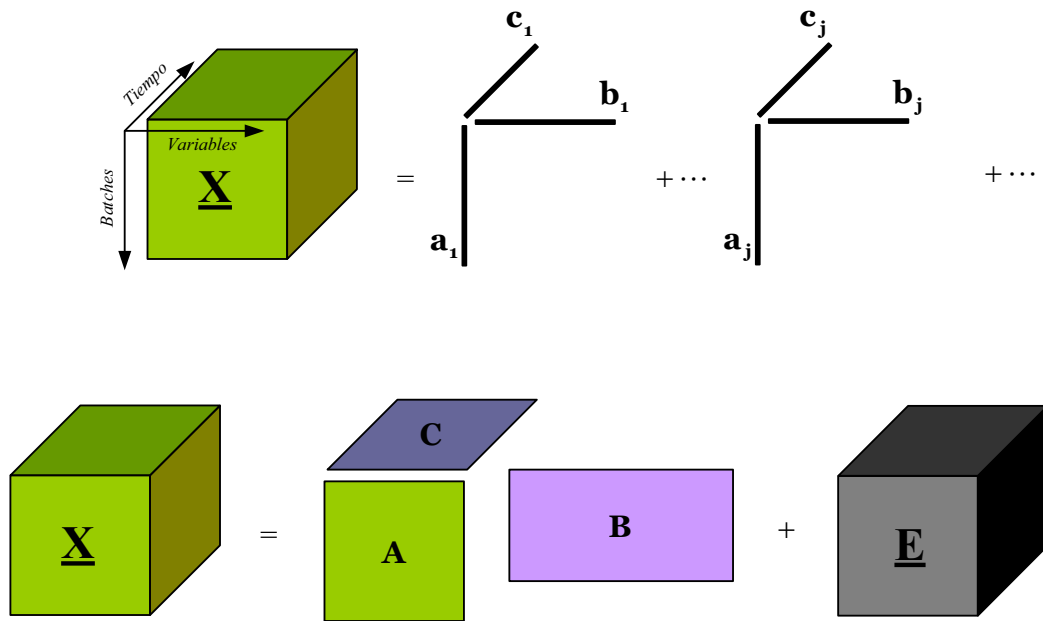


Figura 3.4: Representación gráfica del modelo *PARAFAC*

Al igual que *PARAFAC*, existen otros enfoques aplicables a este tipo de arreglos, por ejemplo los conocidos como *Tucker 1*, *2* y *3*. Muchos autores han evaluado el desempeño de estos modelos para el análisis de arreglos *MM*. Kiers (1991) demostró la existencia de un ordenamiento jerárquico entre dichos enfoques. Este autor indicó que *PARAFAC* puede considerarse como un caso particular del enfoque *Tucker 3*, el que a su vez, resulta ser una situación especial del *Tucker 1*. Este último es conocido también como enfoque *MPCA* y guarda una estrecha relación con el *PCA* descrito en el Capítulo 2 (Westerhuis y col., 1999). Algunos autores señalaron que el desempeño de los modelos *MPCA* es notoriamente superior a los de *Tucker 3* y *PARAFAC* (Anderson, 1958; Bro, 1997; Westerhuis y col., 1999). Por este motivo, se centrará la atención en las distintas modificaciones y mejoras propuestas para los enfoques del tipo *MPCA*.

3.3 Análisis Multimodo: *MPCA*, *MICA*, *MPLS*

Lohmöller y Wold (1980) demostraron que los arreglos tri-modo pueden descomponerse como el producto de Kronecker entre un vector y una matriz, y notaron que dicha descomposición es equivalente a “desdoblar” el arreglo tridimensional de la forma en que se indica en la Figura 3.5 (Wold y col., 1987).

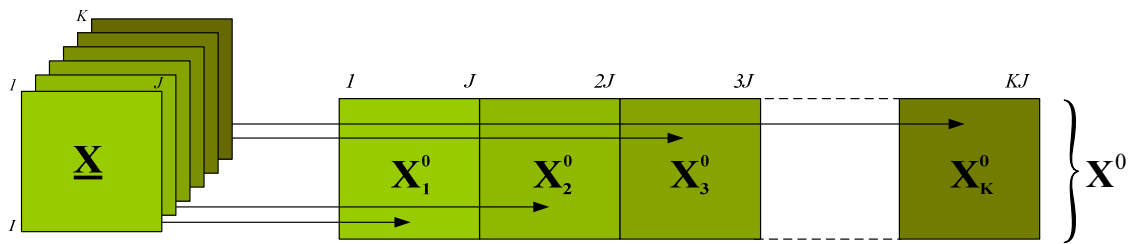


Figura 3.5: Estrategia de desdoblamiento usada en *MPCA* (desdoblamiento *batch-wise*)

Así, las columnas de la matriz \mathbf{X}^0 desde $1+(k-1)J$ hasta kJ corresponden a la k -ésima capa vertical mostrada en la Figura 3.3(c) ($\mathbf{X}_{(:, 1+(k-1)J:kJ)}^0 = \underline{\mathbf{X}}_{(:, :, k)}, \forall k = 1, 2, \dots, K$). Luego, es posible analizar a la matriz de datos crudos resultante, \mathbf{X}^0 , aplicando las técnicas presentadas en el Capítulo 2.

Según sea el modo que se conserva en la matriz \mathbf{X}^0 el desdoblamiento recibe diferentes denominaciones. El propuesto por Lohmöller y Wold (1980) se conoce como *batch-wise* (*BW*) en el ámbito del *SPC*, debido a que el primer modo (que define cada *batch*) es el que se mantiene en \mathbf{X}^0 . Posteriormente, diversos autores abordaron el análisis de los arreglos *MM* usando diferentes tipos de desdoblamiento de los arreglos tridimensionales (Wold y col., 1987; Nomikos y MacGregor, 1994; Wold, 1994; Nomikos y MacGregor, 1995; Lee y col., 2003b; Yoo y col., 2004b; etc.). La Tabla 3.1 muestra las dimensiones de las matrices obtenidas mediante diferentes tipos de desdoblamiento junto con el nombre de cada uno de ellos y la asignación de cada una de las capas de $\underline{\mathbf{X}}$ en \mathbf{X}^0 .

Tabla 3.1: Diferentes desdoblamientos posibles para el arreglo tridimensional $\underline{\mathbf{X}}$

	Reordenamiento		Nombre	Abreviatura
$\mathbf{X}^0_{(I \times KJ)}$	$\left(\mathbf{X}^0_{(:,1+(k-1)J:kJ)} = \underline{\mathbf{X}}_{(:,:,k)}^{\square \rightarrow}, \forall k = 1, 2, \dots, K \right)$	$\in \mathbb{R}^{I \times KJ}$	<i>batch-wise</i>	<i>BW</i>
$\mathbf{X}^0_{(KI \times J)}$	$\left(\mathbf{X}^0_{(1+(k-1)I:kl,:)} = \underline{\mathbf{X}}_{(:,:,k)}^{\square \downarrow}, \forall k = 1, 2, \dots, K \right)$ \equiv $\left(\mathbf{X}^0_{(1+(i-1)K:iK,:)} = \underline{\mathbf{X}}_{(i,,:)}^{\square \downarrow}, \forall i = 1, 2, \dots, I \right)$	$\in \mathbb{R}^{KI \times J}$	<i>variable-wise</i>	<i>VW</i>
$\mathbf{X}^0_{(K \times IJ)}$	$\left(\mathbf{X}^0_{(:,1+(i-1)J:iJ)} = \underline{\mathbf{X}}_{(i,,:)}, \forall i = 1, 2, \dots, I \right)$	$\in \mathbb{R}^{K \times IJ}$	<i>time-wise</i>	<i>TW</i>

Todos estos desdoblamientos presentan particularidades que fueron explotadas por varios autores con el objetivo de adaptar las técnicas clásicas al caso de procesos *batch*. A continuación se analizan las principales metodologías propuestas para el monitoreo de procesos *batch*, en las cuales se hace uso de alguno de estos desdoblamientos.

3.3.1 Método de Nomikos y MacGregor (NM)

Nomikos y MacGregor (1994, 1995) utilizaron el desdoblamiento propuesto por Lohmöller y Wold (1980) para el monitoreo de procesos *batch* en conjunto con las técnicas de *PCA* descritas en el Capítulo 2. Estos trabajos se consideran pioneros, y son tomados como referencia por muchos autores quienes comparan el desempeño de sus propias técnicas con el obtenido usando las propuestas descritas en los mismos.

Dada la *PR*, la *Etapa I* o de modelamiento fuera de línea consiste en tres pasos que se indican en la Tabla 3.2. En primer término se realiza el desdoblamiento de $\underline{\mathbf{X}}$ tal como se muestra en la Figura 3.6 para obtener la matriz \mathbf{X}^0 . Esta última es posteriormente escalada con el objetivo de: a) eliminar las variaciones temporales, que son la mayor fuente de no-

linealidad en los datos, b) eliminar las unidades de medición asociadas a cada variable y, c) establecer un rango de variación prácticamente homogéneo para todas las variables. Finalmente, los datos escalados \mathbf{X} se usan para obtener los modelos descritos en el Capítulo 2 (en el caso de los modelos *PLS*, la matriz \mathbf{Y} debe desdoblarse también en forma consistente).

Tabla 3.2: Pasos requeridos en la etapa de modelado fuera de línea (*Etapa I*)

<i>Etapa I, (NM)</i>	
1.	Desdoblamiento $\mathbf{X}_{(I \times J \times K)} \mapsto \mathbf{X}^0_{(I \times KJ)}$
2.	Escalado $\mathbf{X}^0_{(I \times KJ)} \mapsto \mathbf{X}_{(I \times KJ)}$
3.	Modelamiento <i>PCA, ICA, PLS</i>

El método *NM* usa las K observaciones realizadas a lo largo de toda la operación en la construcción de los modelos en *VLs*. Este tipo de modelo se conoce como modelo global (*MG*) debido a que se obtiene usando la información que describe al proceso en forma general durante toda su duración, sin resaltar las características especiales que algunas de sus etapas pudieran tener.

Entre las ventajas del enfoque de *NM* pueden señalarse la facilidad de interpretación de los resultados, dado que, cada uno de los *batches* en \mathbf{X} mantiene su identidad durante el análisis. Este hecho facilita el pre-tratamiento de los datos en la *Etapa I* debido a que cada uno de los *batches* se trata como un individuo. Así, los métodos usados en las etapas de refinamiento de las poblaciones se corresponden casi completamente con las actividades que se realizarán durante el monitoreo en línea de los nuevos *batches*.

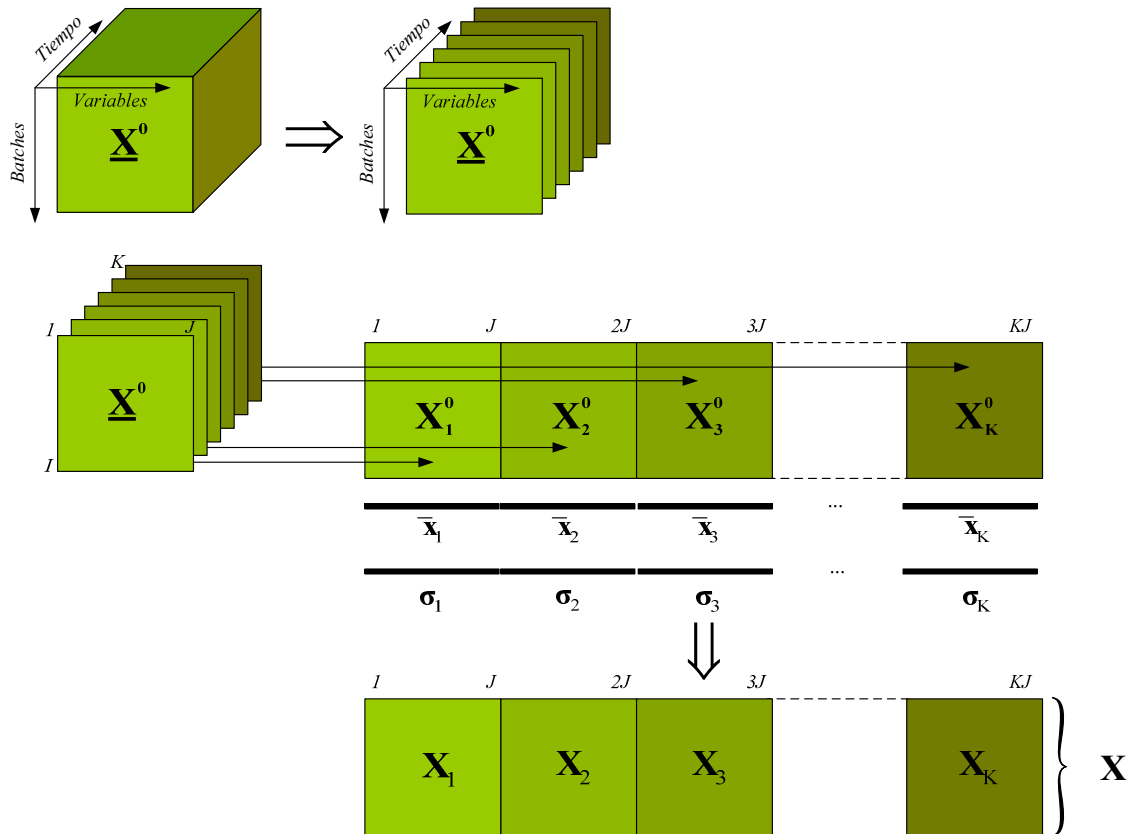


Figura 3.6: Desdoblamiento BW y estandarización de los datos para el enfoque NM

Una desventaja importante de este enfoque es que el número de variables crece proporcionalmente con el número de observaciones que se efectúan durante la operación del *batch*. Así, el proceso que era originalmente descrito por J variables medidas en K instantes de tiempo es ahora representado por KJ “*pseudo-variables*” que se usan simultáneamente como descriptores del proceso. En consecuencia, la matriz de covarianza \mathbf{S} contiene $KJ(KJ + 1)/2$ parámetros diferentes que deben estimarse a partir de los I *batches* de la *PR*. Si $KJ > I$, la información disponible no será suficiente para estimar en forma adecuada todos los parámetros de la matriz de correlación \mathbf{Z} . Así, dado que \mathbf{Z} define las direcciones de los espacios de *VLs*, una estimación pobre de la misma resulta en un cálculo inadecuado de los espacios latentes. De hecho, en estas situaciones la matriz \mathbf{Z} resulta mal condicionada (Jackson, 1991).

El método *NM* presenta otro inconveniente importante que está relacionado con su implementación para el monitoreo en línea. Cuando un nuevo *batch* se monitorea en tiempo real, las observaciones se van obteniendo en forma paulatina durante el transcurso de la operación. Sin embargo, como puede observarse en la Tabla 3.3, es condición necesaria para poder proyectar cada nueva observación que $\mathbf{x}_k \in \mathbb{R}^{K \times I}$. De esta manera, cuando sólo se han tomado k observaciones del proceso, las restantes $K - k$ deben ser estimadas de alguna forma. Nomikos y MacGregor (1994) propusieron tres métodos diferentes para la estimación de los estados futuros del proceso:

- a. *Método de la desviación actual.* En este caso, la desviación de la observación actual de \mathbf{x}_k^0 respecto del promedio correspondiente $(\bar{\mathbf{x}}_k)$ se propaga en las $(K - k)$ observaciones futuras, con respecto a sus correspondientes promedios.
- b. *Método de la desviación nula.* En este método los valores de los $(K - k)$ estados futuros se asumen iguales al promedio de las observaciones correspondientes en \mathbf{X} .
- c. *Método de estimación de observaciones faltantes.* En este método se aprovecha la capacidad del *PCA* para estimar datos faltantes en las observaciones.

Cualquiera de estos métodos permitirá estimar los valores de las observaciones futuras del proceso, haciendo posible la proyección de las observaciones actuales en los espacios latentes. Sin embargo, debe tenerse en cuenta que estas estimaciones podrían afectar el desempeño de las estrategias de monitoreo en lo que hace a su capacidad de detección, especialmente en las etapas tempranas del proceso (Nomikos y MacGregor, 1994).

Otro aspecto a tener en cuenta es que, Nomikos y MacGregor (1994, 1995) consideraron que la matriz de covarianza se mantiene constante durante todo el transcurso de la operación.

Ramaker y Van Sprang (2004) señalaron la importancia de considerar la variación de las matrices de covarianza de los *SCORES*, \mathbf{S}_k , en función del tiempo. Los cambios en la estructura de correlación a lo largo de la operación del proceso pueden ser incorporados en el análisis mediante un paso adicional durante la etapa de modelado. Luego de realizar el modelo en *VLs*, cada uno de los *batches* de la *PR* se somete al mismo método de estimación para las observaciones futuras que se aplicará a nuevos *batches*. Es decir, cada uno de los *batches* en la *PR* se monitorea como si fuese uno nuevo. Como resultado se obtendrán K conjuntos de trayectorias para cada uno de los *batches* de la *PR*. Cada una de estas trayectorias estará formada por dos partes: la primera con las mediciones realizadas hasta la k -ésima observación (\mathbf{x}_{med}) y la segunda compuesta por las estimaciones de los estados futuros del proceso, obtenidas según el método elegido (\mathbf{x}_{est}):

$$\mathbf{x}_k^T = \begin{bmatrix} \mathbf{x}_{med}^T & \mathbf{x}_{est}^T \end{bmatrix}. \quad (3.3)$$

$(1 \times K)J$ $1 \times kJ$ $1 \times (K-k)J$

En el paso siguiente, cada una de estas trayectorias se proyecta en el espacio de *VLs*. Los *SCORES* así obtenidos para todos los *batches* en la *PR* se agrupan en un arreglo \mathbf{T}_k , de acuerdo a la observación a la que corresponden (índice k), y se usan para la estimación de las matrices de covarianza \mathbf{S}_k .

$$\mathbf{S}_k = \frac{1}{(I-1)} \mathbf{T}_k^T \mathbf{T}_k. \quad (3.4)$$

La Tabla 3.3 contiene las ecuaciones necesarias para la aplicación de *MPCA* y *MICA* cuando se usa el desdoblamiento *batch-wise*. En la misma se indican también las dimensiones de cada una de las matrices y de los vectores involucrados en las Etapas *I* y *II*.

Tabla 3.3: Modelos *MPCA* y *MICA* (desdoblamiento *batch-wise*)

Desdoblamiento <i>batch-wise</i> (Nomikos y MaGregor, 1994)		
$\underline{\mathbf{X}}_{(I \times J \times K)} \Rightarrow \mathbf{X}^0_{(I \times KJ)} \quad \mathbf{X}^0_{(:, J+(k-1)J:kJ)} = \underline{\mathbf{X}}_{(:, :, k)} \Rightarrow \mathbf{X}^0 = \left[\underline{\mathbf{X}}_{(:, :, 1)} \quad \underline{\mathbf{X}}_{(:, :, 2)} \quad \cdots \quad \underline{\mathbf{X}}_{(:, :, K)} \right]$		
<i>MPCA</i>		
Etapa I	Proyecciones	Reconstrucciones
	$\begin{aligned} \mathbf{T} &= \mathbf{X} \mathbf{P} \\ (\mathbb{R}^{J \times R}) & (\mathbb{R}^{I \times KJ}) (\mathbb{R}^{KJ \times R}) \\ \tilde{\mathbf{T}} &= \mathbf{X} \tilde{\mathbf{P}} \\ (\mathbb{R}^{I \times KJ-R}) & (\mathbb{R}^{I \times KJ}) (\mathbb{R}^{KJ \times KJ-R}) \end{aligned}$	$\hat{\mathbf{X}} = \mathbf{X} \mathbf{P} \mathbf{P}^T$ $(\mathbb{R}^{I \times KJ}) \quad (\mathbb{R}^{I \times KJ}) (\mathbb{R}^{KJ \times R}) (\mathbb{R}^{R \times KJ})$
Residuos	$\mathbf{E} = \mathbf{X} \left(\mathbf{I} - \mathbf{P} \mathbf{P}^T \right)$ $(\mathbb{R}^{I \times KJ}) \quad (\mathbb{R}^{I \times KJ}) \left((\mathbb{R}^{KJ \times KJ}) - (\mathbb{R}^{KJ \times R}) (\mathbb{R}^{R \times KJ}) \right)$	
Etapa II	Proyecciones	Reconstrucciones
	$\begin{aligned} \mathbf{t}_k &= \mathbf{P}^T \mathbf{x}_k \\ (\mathbb{R}^{R \times 1}) & (\mathbb{R}^{R \times KJ}) (\mathbb{R}^{KJ \times 1}) \\ \tilde{\mathbf{t}}_k &= \tilde{\mathbf{P}}^T \mathbf{x}_k \\ (\mathbb{R}^{KJ-R \times 1}) & (\mathbb{R}^{KJ-R \times KJ}) (\mathbb{R}^{KJ \times 1}) \end{aligned}$	$\hat{\mathbf{x}}_k = \mathbf{P} \mathbf{P}^T \mathbf{x}_k$ $(\mathbb{R}^{KJ \times 1}) \quad (\mathbb{R}^{KJ \times R}) (\mathbb{R}^{R \times KJ}) (\mathbb{R}^{KJ \times 1})$
Residuos	$\mathbf{e}_k = \left(\mathbf{I} - \mathbf{P} \mathbf{P}^T \right) \mathbf{x}_k$ $(\mathbb{R}^{KJ \times 1}) \quad \left((\mathbb{R}^{KJ \times KJ}) - (\mathbb{R}^{KJ \times R}) (\mathbb{R}^{R \times KJ}) \right) (\mathbb{R}^{KJ \times 1})$	
<i>MICA</i>		
Etapa I	Proyecciones	Reconstrucciones
	$\begin{aligned} \Psi &= \mathbf{B}^T \mathbf{Q} \mathbf{X}^T \\ (\mathbb{R}^{A \times I}) & (\mathbb{R}^{A \times KJ}) (\mathbb{R}^{KJ \times KJ}) (\mathbb{R}^{KJ \times I}) \\ \tilde{\Psi} &= \tilde{\mathbf{B}}^T \mathbf{Q} \mathbf{X}^T \\ (\mathbb{R}^{\delta \times I}) & (\mathbb{R}^{\delta \times KJ}) (\mathbb{R}^{KJ \times KJ}) (\mathbb{R}^{KJ \times I}) \\ & \delta = KJ - A \end{aligned}$	$\hat{\mathbf{X}}^T = \mathbf{Q}^{-1} \mathbf{B} \mathbf{W} \mathbf{X}^T$ $(\mathbb{R}^{KJ \times I}) \quad (\mathbb{R}^{KJ \times KJ}) (\mathbb{R}^{KJ \times A}) (\mathbb{R}^{A \times KJ}) (\mathbb{R}^{KJ \times I})$
Residuos	$\mathbf{E}^T = \left(\mathbf{I} - \mathbf{Q}^{-1} \mathbf{B} \mathbf{W} \right) \mathbf{X}^T$ $(\mathbb{R}^{KJ \times I}) \quad \left((\mathbb{R}^{KJ \times KJ}) - (\mathbb{R}^{KJ \times KJ}) (\mathbb{R}^{KJ \times A}) (\mathbb{R}^{A \times KJ}) \right) (\mathbb{R}^{KJ \times I})$	
Etapa II	Proyecciones	Reconstrucciones
	$\begin{aligned} \psi_k &= \mathbf{B}^T \mathbf{Q} \mathbf{x}_k \\ (\mathbb{R}^{A \times 1}) & (\mathbb{R}^{A \times KJ}) (\mathbb{R}^{KJ \times KJ}) (\mathbb{R}^{KJ \times 1}) \\ \tilde{\psi}_k &= \tilde{\mathbf{B}}^T \mathbf{Q} \mathbf{x}_k \\ (\mathbb{R}^{\delta \times 1}) & (\mathbb{R}^{\delta \times KJ}) (\mathbb{R}^{KJ \times KJ}) (\mathbb{R}^{KJ \times 1}) \\ & \delta = KJ - A \end{aligned}$	$\hat{\mathbf{x}}_k = \mathbf{Q}^{-1} \mathbf{B} \mathbf{W} \mathbf{x}_k$ $(\mathbb{R}^{KJ \times 1}) \quad (\mathbb{R}^{KJ \times KJ}) (\mathbb{R}^{KJ \times A}) (\mathbb{R}^{A \times KJ}) (\mathbb{R}^{KJ \times 1})$
Residuos	$\mathbf{e}_k = \left(\mathbf{I} - \mathbf{Q}^{-1} \mathbf{B} \mathbf{W} \right) \mathbf{x}_k$ $(\mathbb{R}^{KJ \times 1}) \quad \left((\mathbb{R}^{KJ \times KJ}) - (\mathbb{R}^{KJ \times KJ}) (\mathbb{R}^{KJ \times A}) (\mathbb{R}^{A \times KJ}) \right) (\mathbb{R}^{KJ \times 1})$	

La forma de cálculo de los estadísticos cuando se usa *NM* se detalla en la Tabla 3.4. Las ecuaciones mostradas son las correspondientes al método *NM*, sin considerar las variaciones de *S* en el tiempo. Para incorporar esta característica en los análisis, las matrices Λ y $\tilde{\Lambda}$ deben reemplazarse por las matrices correspondientes \mathbf{S}_k y $\tilde{\mathbf{S}}_k$.

Los estadísticos D , T_H^2 , I_d , I_e y Q consideran todas las observaciones (tanto las realizadas como las estimadas) a diferencia del estadístico *SPE* que sólo considera los errores de representación de cada nueva observación \mathbf{x}_k . Las contribuciones de las variables a los

valores finales de los estadísticos (*CVs*) se calculan utilizando ecuaciones análogas a las presentadas en el Capítulo 2 (Westerhuis y col., 2000). La adaptación de dichas ecuaciones para el cálculo de las *CVs*, para ésta y las otras metodologías analizadas en este capítulo, puede llevarse a cabo con relativa facilidad.

Tabla 3.4: Estadísticos y límites de control para los modelos *MPCA* y *MICA* (desdoblamiento *batch-wise*)

Desdoblamiento <i>batch-wise</i> (Nomikos y MacGregor, 1994)			
<i>MPCA</i>			
	\mathbb{L}^R	\mathbb{L}^{JK-R}	Residuos
	$D_i = \mathbf{t}_i^T \Lambda^{-1} \mathbf{t}_i$	$T_{H_i}^2 = \tilde{\mathbf{t}}_i^T \tilde{\Lambda}^{-1} \tilde{\mathbf{t}}_i$	$Q_i = \sum_{j=1}^{KJ} e_{i,j}^2$
Distribuciones y valores críticos (asumiendo multinormalidad en las <i>CPs</i>)			
Etapa I	$D \approx \left(\frac{R(I^2 - I)}{I(I - R)} \right) \mathcal{F}_{R, I-R, \alpha}$	$T_H^2 \approx \kappa \mathcal{F}_{(KJ-R), I-(KJ-R), \alpha}$ $\kappa = \left(\frac{(KJ - R)(I^2 - I)}{I(I - (KJ - R))} \right)$	$Q_c = \theta_1 \left[1 - \frac{\theta_2 h_0 (1 - h_0)}{\theta_1^2} + \frac{z_\alpha (2\theta_2 h_0^2)^{1/2}}{\theta_1} \right]^{1/h_0}$ $h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2}$ $\theta_m = \sum_{r=R+1}^{KJ} \lambda_r^m \quad m = 1, 2, 3$
Etapa II	$D_k = \mathbf{t}_k^T \tilde{\Lambda}^{-1} \mathbf{t}_k$	$T_{H_k}^2 = \tilde{\mathbf{t}}_k^T \tilde{\Lambda}^{-1} \tilde{\mathbf{t}}_k$	$SPE_k = \sum_{j=1+(k-1)J}^{kJ} e_j^2$
	$D \approx \left(\frac{R(I^2 - I)}{I(I - R)} \right) \mathcal{F}_{R, I-R, \alpha}$	$T_H^2 \approx \kappa \mathcal{F}_{(KJ-R), I-(KJ-R), \alpha}$ $\kappa = \left(\frac{(KJ - R)(I^2 - I)}{I(I - (KJ - R))} \right)$	$SPE_{C,k} = g_k \chi_{h_k, \alpha}^2$ $g_k = \frac{v_k}{2m_k}$ $h_k = 2m_k^2 / v_k$ <i>m</i> y <i>v</i> definidas en ecuación (2.29)
<i>MICA</i>			
Etapa I	\mathbb{L}^R $I_{d_i}^2 = \Psi_i^T \Psi_i$	\mathbb{L}^{KJ-R} $I_{e_i}^2 = \tilde{\Psi}_i^T \tilde{\Psi}_i$	Residuos $SPE_i = \sum_{j=1}^{KJ} e_{i,j}^2$
Etapa II	$I_{d_k}^2 = \Psi_k^T \Psi_k$	$I_{e_i}^2 = \tilde{\Psi}_k^T \tilde{\Psi}_k$	$SPE_k = \sum_{j=1+(k-1)J}^{kJ} e_j^2$
<ul style="list-style-type: none"> Los límites de control para los estadísticos I_d, I_e y SPE en <i>MICA</i> se obtienen mediante el método <i>KDE</i> (Apéndice 3). 			

En particular, la Tabla 3.5 contiene las ecuaciones resultantes para el cálculo de las *CVs* cuando se usa el método *NM*.

En el método *NM* existe una gran probabilidad de que el estadístico T_H^2 presente los problemas de mal condicionamiento que se señalaron en el Capítulo 2. Esto se debe a que el número de *pseudo-variables* es por lo general mucho mayor que el número de componentes retenidos.

Tabla 3.5: Cálculo de las *CVs* para *MPCA* y *MICA* (desdoblamiento *batch-wise*)

Desdoblamiento <i>batch-wise</i> (Nomikos y MacGregor, 1994)		
<i>MPCA</i>		
$c_{j,k}^D = \mathbf{t}_k^T \mathbf{S}_k^{-1} \left[x_{j,k} \mathbf{P}_{(j,:)} (\mathbf{P}^T \mathbf{P})^{-1} \right]^T$	$c_{j,k}^{T_H^2} = \tilde{\mathbf{t}}_k^T \tilde{\mathbf{S}}_k^{-1} \left[x_{j,k} \tilde{\mathbf{P}}_{(j,:)} (\tilde{\mathbf{P}}^T \tilde{\mathbf{P}})^{-1} \right]^T$	$c_{j,k}^{SPE} = e_{j,k}^2$
<i>MICA</i>		
$\mathbf{c}_k^{I_d^2} = \frac{\ \Psi_k\ }{\ \mathbf{Q}^{-1} \mathbf{B} \Psi_k\ } \mathbf{Q}^{-1} \mathbf{B} \mathbf{W} \mathbf{x}_k; \quad c_{j,k}^{I_d^2} = \mathbf{c}_k^{I_d^2}(j)$	$\mathbf{c}_k^{I_e^2} = \frac{\ \tilde{\Psi}_k\ }{\ \mathbf{Q}^{-1} \tilde{\mathbf{B}} \tilde{\Psi}_k\ } \mathbf{Q}^{-1} \tilde{\mathbf{B}} \tilde{\mathbf{W}} \mathbf{x}_k; \quad c_{j,k}^{I_e^2} = \mathbf{c}_k^{I_e^2}(j)$	$c_{j,k}^{SPE} = e_{j,k}^2$

3.3.2 Método de Wold, Kettaneh, Fridén, y Holmberg Original (*WKFH*) y Modificado (*AT*)

Wold y *col.* (1998) propusieron el uso del desdoblamiento *VW* como una alternativa para solucionar algunas de las desventajas del método presentado por Nomikos y MacGregor, (1994). En este caso, cada una de las capas verticales ($\underline{\mathbf{X}}_{(:,k)}, \forall k = 1, 2, \dots, K$) se ubica debajo de la anterior formando así el arreglo $\mathbf{X}^0 = \left[\underline{\mathbf{X}}_{(:,1)}; \underline{\mathbf{X}}_{(:,2)}; \dots; \underline{\mathbf{X}}_{(:,K)} \right] \in \mathbb{R}^{KI \times J}$. Posteriormente, se escala la matriz \mathbf{X}^0 para obtener \mathbf{X} , la que finalmente se utiliza para la construcción de los modelos en *VLs*. La Figura 3.7 muestra un esquema detallado de este método (*WKFH*).

Como en el método *NM*, el método *WKFH* usa las K observaciones de los I *batches* de la *PR* para la construcción de los modelos, por lo que constituye también un modelo global del proceso.

Las principales ventajas de este método se resumen a continuación:

- a. No se necesitan estimaciones de los estados futuros del proceso.
- b. Cada una de las nuevas observaciones $\mathbf{x}^0 \in \mathbb{R}^{J \times I}$ se proyecta directamente en los espacios latentes luego de haber sido escalada usando la media y desvío estándar poblacionales ($\bar{\mathbf{x}}$ y $\bar{\boldsymbol{\sigma}}$).
- c. Debido al tipo de desdoblamiento utilizado (*VW*), el método *WKFH* permite el tratamiento de *batches* de distintas duraciones ($K_i \neq K_m$) en forma natural. La Sección 3.4 presenta una discusión detallada sobre el tratamiento de *batches* de longitudes diferentes (Lee y col., 2003b).

Una de las desventajas importantes asociadas a esta metodología proviene del tipo de escalado utilizado. Como puede verse en la Figura 3.7, los valores utilizados en el escalado de la matriz \mathbf{X}^0 son la media y varianza de las K observaciones realizadas sobre los I *batches* ($\bar{\mathbf{x}}$ y $\bar{\boldsymbol{\sigma}}$). Al centrar todos los datos en \mathbf{X}^0 , respecto de un único valor ($\bar{\mathbf{x}}$), la mayor fuente de no-linealidad (que proviene de la variación “determinística” del proceso) seguirá presente en \mathbf{X} luego del escalado.

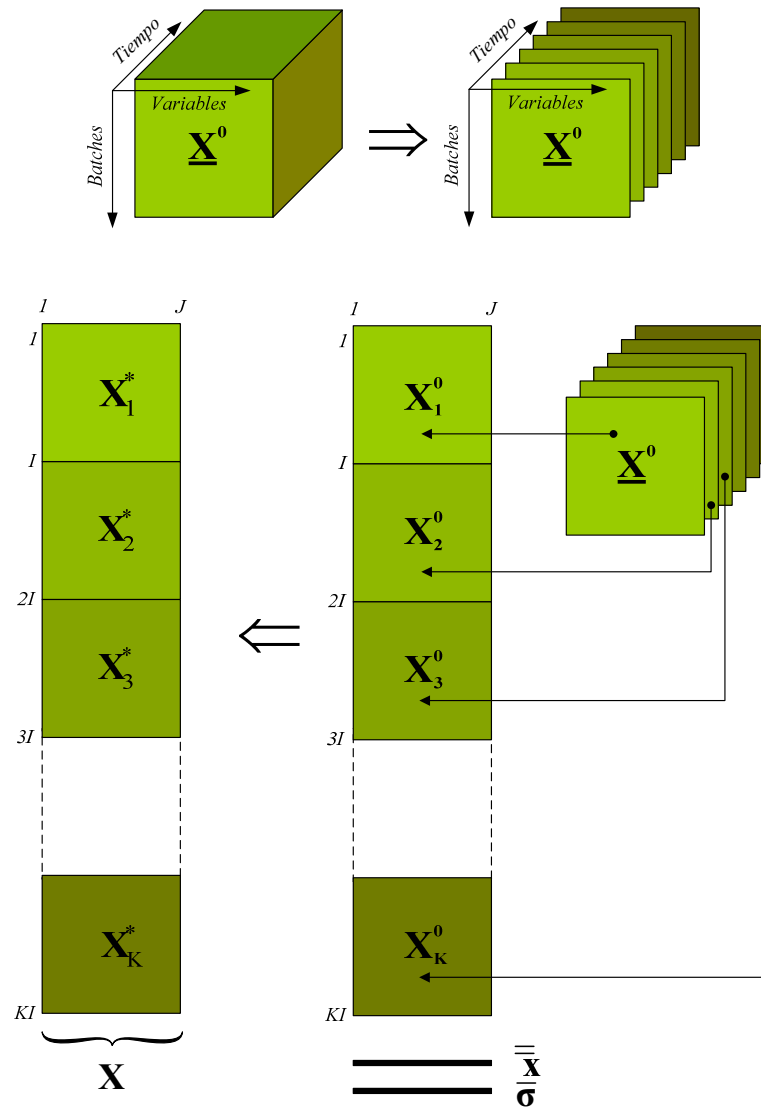


Figura 3.7: Esquema de desdoblamiento y escalado de \underline{X} según WKFH

La Figura 3.8 muestra los perfiles de tres variables medidas durante la operación de un reactor *batch*. En la primera columna se presentan las mediciones tal como se obtienen del proceso mientras que en la segunda se grafican los perfiles para las mismas variables luego de haber sido centradas según $\bar{\bar{x}}$. Se puede observar que la mayor fuente de variabilidad proviene de la dependencia temporal “determinística” de las variables originales. Finalmente, la tercera columna muestra los valores centrados usando las medias para cada uno de los K subgrupos de observaciones (\bar{x}_k). Puede notarse que de esta forma se eliminó la no-linealidad

propia del proceso, dejando sólo las fluctuaciones asociadas a la variabilidad intrínseca del mismo.

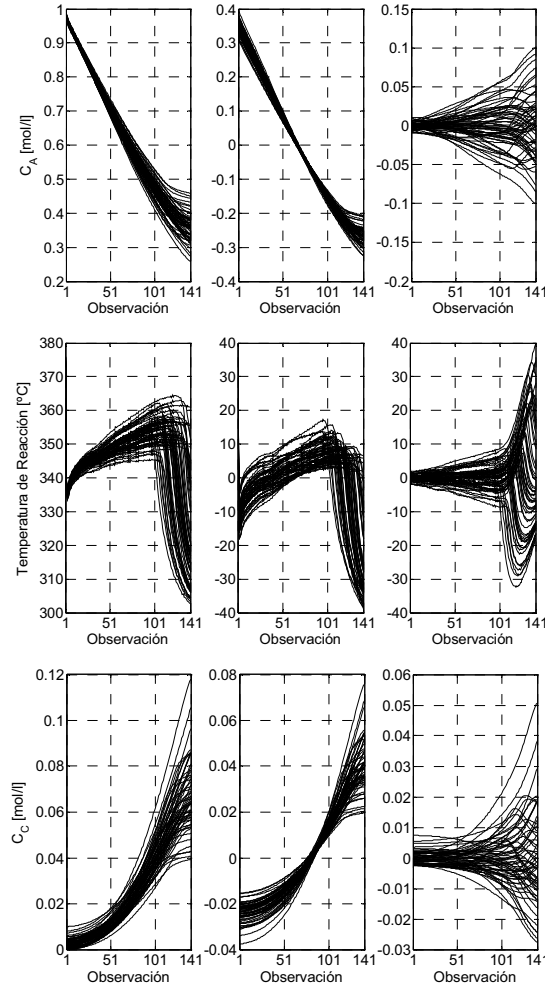


Figura 3.8: Dependencia de las tendencias remanentes en \mathbf{X} de acuerdo al tipo de escalado

Al realizar el análisis de VLs con el método $WKFH$, las primeras componentes estarán asociadas a la variación “determinística” remanente en la matriz de datos luego del escalado. Más aún, las variaciones aleatorias que intentan modelarse podrían quedar en la parte excluida del modelo (Westerhuis y *col.*, 1999). Una alternativa para solucionar este inconveniente es realizar el escalado de cada una de las capas $(\underline{\mathbf{X}}_{(:,k)}, \forall k = 1, 2, \dots, K)$ antes de ubicarlas debajo de la anterior, tal como se indica en el esquema de la Figura 3.9. Este procedimiento es equivalente a realizar el desdoblamiento y escalado de $\underline{\mathbf{X}}$ tal como lo sugiere NM y,

reacomodar posteriormente cada una de las particiones $\mathbf{X}_{(:,1+(k-1)J:kJ)}$ (con media nula y desvíos estándares unitarios) para formar $\mathbf{X}_{(I+(k-1)I:kl,:)}$. Este procedimiento, conocido como *AT*, ha sido utilizado por varios autores (Lee y col., 2003b; Yoo y col., 2004b; Aguado y col., 2007). El mismo conserva las ventajas del *WKFH*, evitando la presencia de la no-linealidad “determinística” del proceso mediante un centrado adecuado.

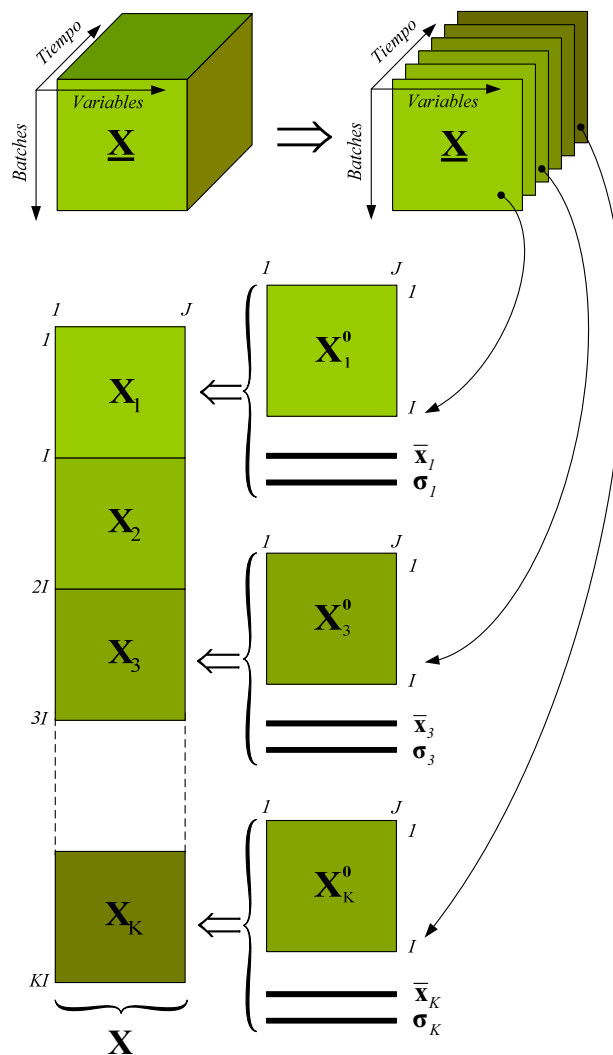


Figura 3.9: Esquema de desdoblamiento y escalado de \mathbf{X} según *AT*

Una práctica común (Van Sprang y col., 2002) es considerar la matriz de covarianza de \mathbf{X} (\mathbf{S}) para la obtención del modelo de *VLs* y la matriz de covarianza de cada capa de la matriz de *SCORES*, $\mathbf{T}_{(I+(k-1)I:kl.)}$, ($\mathbf{S}_{\mathbf{T}_k}$) para el cálculo del estadístico de Hotelling. De esta manera, el modelo resulta global desde el punto de vista de la proyección pero local para los tests de hipótesis.

Las Tablas 3.6 a 3.8 muestran las ecuaciones que definen las distintas proyecciones en *VLs*, los estadísticos utilizados y las *CVs* cuando se aplica el método *AT*.

Una desventaja asociada al desdoblamiento *VW* (independientemente del tipo de escalado) es que no se considera la naturaleza evolutiva de los procesos *batch*. El transcurso previo del *batch* desde el inicio de la operación hasta la *k*-ésima observación no se tiene en cuenta cuando se analiza una observación particular del proceso. En la literatura se han presentado técnicas que tratan, de alguna manera, de incorporar estas características en los modelos (Ramaker y Van Sprang, 2004; Lennox y col., 2001).

3.3.3 Modelos locales (*ML*)

Una alternativa a los métodos *WKFH* y *AT* es modelar cada una de las capas ($\underline{\mathbf{X}}_{(:, :, k)}, \forall k = 1, 2, \dots, K$) como un conjunto de datos por separado, ésto es, generar un modelo independiente para cada grupo de observaciones. En este caso, las ecuaciones que se introdujeron en el Capítulo 2 se aplican en forma directa a cada uno de estos conjuntos de datos, puesto que la variación temporal de las variables monitoreadas no se considera en absoluto.

Tabla 3.6: Modelos *MPCA* y *MICA* (desdoblamiento *variable-wise*)

Desdoblamiento <i>variable-wise</i> (Wold y col., 1998)			
$\mathbf{X}_{(I \times J \times K)} \Rightarrow \mathbf{X}^0_{(KI \times J)} \quad \mathbf{X}^0_{(1+(k-1)J:kJ,:)} = \underline{\mathbf{X}}_{(:, :, k)} \Rightarrow \mathbf{X}^0 = \left[\underline{\mathbf{X}}_{(:, :, 1)}; \underline{\mathbf{X}}_{(:, :, 2)}; \dots; \underline{\mathbf{X}}_{(:, :, K)} \right]$			
MPCA			
Etapas	Proyecciones	Reconstrucciones	Residuos
	$\mathbf{T} = \mathbf{X} \mathbf{P}$ $\tilde{\mathbf{T}} = \mathbf{X} \tilde{\mathbf{P}}$	$\hat{\mathbf{X}} = \mathbf{X} \mathbf{P} \mathbf{P}^T$	$\mathbf{E} = \mathbf{X} \left(\mathbf{I} - \mathbf{P} \mathbf{P}^T \right)$
Etapas	Proyecciones	Reconstrucciones	Residuos
	$\mathbf{t}_k = \mathbf{P}^T \mathbf{x}_k$ $\tilde{\mathbf{t}}_k = \tilde{\mathbf{P}}^T \mathbf{x}_k$	$\hat{\mathbf{x}}_k = \mathbf{P} \mathbf{P}^T \mathbf{x}_k$	$\mathbf{e}_k = \left(\mathbf{I} - \mathbf{P} \mathbf{P}^T \right) \mathbf{x}_k$
MICA			
Etapas	Proyecciones	Reconstrucciones	Residuos
	$\mathbf{\Psi} = \mathbf{B}^T \mathbf{Q} \mathbf{X}^T$ $\tilde{\mathbf{\Psi}} = \tilde{\mathbf{B}}^T \mathbf{Q} \mathbf{X}^T$ $\delta = J - A$	$\hat{\mathbf{X}}^T = \mathbf{Q}^{-1} \mathbf{B} \mathbf{W} \mathbf{X}^T$	$\mathbf{E}^T = \left(\mathbf{I} - \mathbf{Q}^{-1} \mathbf{B} \mathbf{W} \right) \mathbf{X}^T$
Etapas	Proyecciones	Reconstrucciones	Residuos
	$\boldsymbol{\psi}_k = \mathbf{B}^T \mathbf{Q} \mathbf{x}_k$ $\tilde{\boldsymbol{\psi}}_k = \tilde{\mathbf{B}}^T \mathbf{Q} \mathbf{x}_k$ $\delta = J - A$	$\hat{\mathbf{x}}_k = \mathbf{Q}^{-1} \mathbf{B} \mathbf{W} \mathbf{x}_k$	$\mathbf{e}_k = \left(\mathbf{I} - \mathbf{Q}^{-1} \mathbf{B} \mathbf{W} \right) \mathbf{x}_k$

Tabla 3.7: Estadísticos y límites de control para los modelos *MPCA* y *MICA* (desdoblamiento *variable-wise*)

Desdoblamiento <i>variable-wise</i> (Wold y col., 1998)			
MPCA			
\mathbb{L}^R	\mathbb{L}^{J-R}	Residuos	
$D_i = \mathbf{t}_i^T \mathbf{\Lambda}^{-1} \mathbf{t}_i$	$T_{H_i}^2 = \tilde{\mathbf{t}}_i^T \tilde{\mathbf{\Lambda}}^{-1} \tilde{\mathbf{t}}_i$	$Q_i = \sum_{j=1}^J e_{i,j}^2$	
<i>Distribuciones y valores críticos (asumiendo multinormalidad en las CPs)</i>			
Etapa I	$D \approx \left(\frac{R(I^2 - I)}{I(I - R)} \right) \mathcal{F}_{R, I-R, \alpha}$	$T_H^2 \approx \kappa \mathcal{F}_{(J-R), I-(J-R), \alpha}$ $\kappa = \left(\frac{(J - R)(I^2 - I)}{I(I - (J - R))} \right)$	$SPE_C = \theta_1 \left[1 - \frac{\theta_2 h_0 (1 - h_0)}{\theta_1^2} + \frac{z_\alpha (2\theta_2 h_0^2)^{1/2}}{\theta_1} \right]^{1/h_0}$ $h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2}$ $\theta_m = \sum_{r=R+1}^J \lambda_r^m \quad m = 1, 2, 3$
Etapa II	$D_k = \mathbf{t}_k^T \mathbf{S}_k^{-1} \mathbf{t}_k$	$T_{H_k}^2 = \tilde{\mathbf{t}}_k^T \tilde{\mathbf{S}}_k^{-1} \tilde{\mathbf{t}}_k$	$SPE_k = \sum_{j=1}^J e_j^2$
Etapa II	$D \approx \left(\frac{R(I^2 - I)}{I(I - R)} \right) \mathcal{F}_{R, I-R, \alpha}$	$T_H^2 \approx \kappa \mathcal{F}_{(KJ-R), I-(KJ-R), \alpha}$ $\kappa = \left(\frac{(KJ - R)(I^2 - I)}{I(I - (KJ - R))} \right)$	$SPE_{C,k} = g_k \chi_{h_k, \alpha}^2$ $g_k = \frac{v_k}{2m_k}$ $h_k = 2m_k^2 / v_k$ <i>m</i> y <i>v</i> definidas en ecuación (2.29)
MICA			
\mathbb{L}^R	\mathbb{L}^{J-R}	Residuos	
Etapa I	$I_{d_i}^2 = \mathbf{\Psi}_i^T \mathbf{\Psi}_i$	$I_{e_i}^2 = \tilde{\mathbf{\Psi}}_i^T \tilde{\mathbf{\Psi}}_i$	$Q_i = \sum_{j=1}^J e_{i,j}^2$
Etapa II	$I_{d_k}^2 = \mathbf{\Psi}_k^T \mathbf{\Psi}_k$	$I_{e_i}^2 = \tilde{\mathbf{\Psi}}_k^T \tilde{\mathbf{\Psi}}_k$	$SPE_k = \sum_{j=1+(k-1)J}^{kJ} e_j^2$
<ul style="list-style-type: none"> Los límites de control para los estadísticos I_d, I_e y SPE en <i>MICA</i> se obtienen mediante el método <i>KDE</i> (Apéndice 3). 			

Tabla 3.8: Cálculo de las *CVs* para *MPCA* y *MICA* (desdoblamiento *variable-wise*)

Desdoblamiento <i>variable-wise</i> (Wold y col., 1998)		
<i>MPCA</i>		
$c_{j,k}^D = \mathbf{t}_k^T \mathbf{S}^{-1} \left[x_{j,k} \mathbf{P}_{(j,\cdot)} (\mathbf{P}^T \mathbf{P})^{-1} \right]^T$	$c_{j,k}^{I_d^2} = \tilde{\mathbf{t}}_k^T \tilde{\mathbf{S}}^{-1} \left[x_{j,k} \tilde{\mathbf{P}}_{(j,\cdot)} (\tilde{\mathbf{P}}^T \tilde{\mathbf{P}})^{-1} \right]^T$	$c_{j,k}^{SPE} = e_{j,k}^2$
<i>MICA</i>		
$\mathbf{c}_k^{I_d^2} = \frac{\ \Psi_k\ }{\ \mathbf{Q}^{-1} \mathbf{B} \Psi_k\ } \mathbf{Q}^{-1} \mathbf{B} \mathbf{W} \mathbf{x}_k; \quad c_{j,k}^{I_d^2} = \mathbf{c}_k^{I_d^2}(j)$	$\mathbf{c}_k^{I_e^2} = \frac{\ \tilde{\Psi}_k\ }{\ \mathbf{Q}^{-1} \tilde{\mathbf{B}} \tilde{\Psi}_k\ } \mathbf{Q}^{-1} \tilde{\mathbf{B}} \tilde{\mathbf{W}} \mathbf{x}_k; \quad c_{j,k}^{I_e^2} = \mathbf{c}_k^{I_e^2}(j)$	$c_{j,k}^{SPE} = e_{j,k}^2$

3.3.4 Modelos Evolutivos (ME)

Ramaker y Van Sprang (2004) y Ramaker y col. (2005) introdujeron el concepto de “modelos evolutivos” como una herramienta para el monitoreo de procesos *batch*. El objetivo principal de esta estrategia es la obtención de un conjunto de modelos que consideren la evolución del proceso durante su operación. Para ello, el conjunto de datos contenidos en \mathbf{X} se utiliza para construir K matrices $\tilde{\mathbf{X}}_k^0$ que contienen las observaciones de los I *batches* de la *PR* desde la primera ($k = 1$) hasta la k -ésima, es decir:

$$\tilde{\mathbf{X}}_k^0 = \mathbf{X}_{(\cdot, :1:k)}^{\square \rightarrow}, \forall k = 1, 2, \dots, K \tag{3.5}$$

$(I \times kJ)$

La Figura 3.10 muestra el procedimiento de construcción de las matrices $\tilde{\mathbf{X}}_k^0$ a partir de \mathbf{X} . Cada una de estas matrices $\tilde{\mathbf{X}}_k^0$ se utiliza para generar un modelo en *VLs*, el cual se emplea para describir el transcurso de la operación desde su inicio hasta el momento en que se obtiene la k -ésima observación.

Las Tablas 3.9 a 3.11 detallan el conjunto de ecuaciones necesario para: la obtención de los modelos evolutivos en *VLs*, el cálculo de los estadísticos y sus valores límites así como la estimación de las *CVs*, en ese orden.

Debe notarse que el primer modelo de la serie de *MEs* será el modelo local correspondiente a $k = 1$ mientras que el último ($k = K$) es el modelo que se obtendría usando *NM*.

Ciertas ventajas y desventajas de los *ME* son comunes a algunos de los métodos explicados en las secciones precedentes. Entre sus principales ventajas se pueden mencionar:

- a. No se necesitan predicciones de los estados futuros del proceso.
- b. Se incorpora la información acerca de los estados pasados del proceso.
- c. El desdoblamiento utilizado para la obtención de cada $\tilde{\mathbf{X}}_k^0$ es *BW* por lo que la identidad de los *batches* se mantiene, facilitando el análisis en la etapa de modelado fuera de línea.

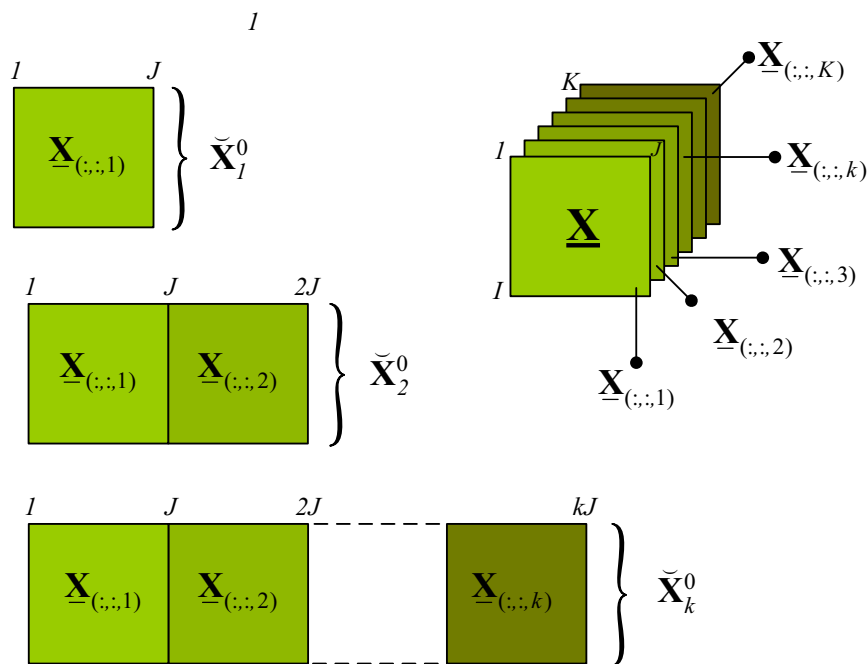


Figura 3.10: Esquema de desdoblamiento de $\underline{\mathbf{X}}$ para los *ME*

Por otra parte, las desventajas más importantes asociada a los *ME* son:

- a. Para valores grandes de K , cuando $kJ \geq I$ las matrices de covarianza $\tilde{\mathbf{S}}_k$ a partir de las cuales se calculan las matrices de rotación ($\tilde{\mathbf{P}}_k$ o $\tilde{\mathbf{W}}_k$ según sea el caso) son, por lo general, pobremente estimadas.

Tabla 3.9: Modelos *MPCA* y *MICA* - Modelos Evolutivos

Modelos Evolutivos (Ramaker y Van Sprang, 2004)			
$\underline{\mathbf{X}}_{(I \times J \times K)} \Rightarrow \mathbf{X}_k^0_{(I \times k)}$ $\tilde{\mathbf{X}}_k^0 = \underline{\mathbf{X}}_{(I \times kJ)}^{\rightarrow}, \quad \underline{\mathbf{X}}_{(:, :, k)} \Rightarrow \tilde{\mathbf{X}}_k^0 = \left[\underline{\mathbf{X}}_{(:, :, 1)} \quad \underline{\mathbf{X}}_{(:, :, 2)} \quad \cdots \quad \underline{\mathbf{X}}_{(:, :, k)} \right]$			
EMPCA			
Etapa I	Proyecciones $\tilde{\mathbf{T}}_k = \tilde{\mathbf{X}}_k \tilde{\mathbf{P}}_k$ $\tilde{\mathbf{T}}_k = \tilde{\mathbf{X}}_k \tilde{\mathbf{P}}_k$	Reconstrucciones $\hat{\tilde{\mathbf{X}}}_k = \tilde{\mathbf{X}}_k \tilde{\mathbf{P}}_k \tilde{\mathbf{P}}_k^T$	Residuos $\tilde{\mathbf{E}}_k = \tilde{\mathbf{X}}_k \left(\mathbf{I} - \tilde{\mathbf{P}}_k \tilde{\mathbf{P}}_k^T \right)$
Etapa II	$\tilde{\mathbf{t}}_k = \tilde{\mathbf{P}}_k^T \tilde{\mathbf{x}}_k$ $\tilde{\mathbf{t}}_k = \tilde{\mathbf{P}}_k^T \tilde{\mathbf{x}}_k$	$\hat{\tilde{\mathbf{x}}}_k = \tilde{\mathbf{P}}_k \tilde{\mathbf{P}}_k^T \tilde{\mathbf{x}}_k$	$\tilde{\mathbf{e}}_k = \left(\mathbf{I} - \tilde{\mathbf{P}}_k \tilde{\mathbf{P}}_k^T \right) \tilde{\mathbf{x}}_k$
EMICA			
Etapa I	Proyecciones $\tilde{\Psi} = \tilde{\mathbf{B}}_k^T \tilde{\mathbf{Q}}_k \tilde{\mathbf{X}}_k^T$ $\tilde{\Psi} = \tilde{\mathbf{B}}_k^T \tilde{\mathbf{Q}}_k \tilde{\mathbf{X}}_k^T$ <p style="text-align: center;">$\delta = kJ - A_k$</p>	Reconstrucciones $\tilde{\mathbf{X}}_k^T = \tilde{\mathbf{Q}}_k^{-1} \tilde{\mathbf{B}}_k \tilde{\mathbf{W}} \tilde{\mathbf{X}}_k^T$	Residuos $\tilde{\mathbf{E}}^T = \left(\mathbf{I} - \tilde{\mathbf{Q}}_k^{-1} \tilde{\mathbf{B}}_k \tilde{\mathbf{W}} \right) \tilde{\mathbf{X}}_k^T$
Etapa II	$\tilde{\psi}_k = \tilde{\mathbf{B}}_k^T \tilde{\mathbf{Q}}_k \mathbf{x}_k^T$ $\tilde{\psi}_k = \tilde{\mathbf{B}}_k^T \tilde{\mathbf{Q}}_k \mathbf{x}_k^T$ <p style="text-align: center;">$\delta = kJ - A_k$</p>	$\hat{\tilde{\mathbf{x}}}_k = \tilde{\mathbf{Q}}_k^{-1} \tilde{\mathbf{B}}_k \tilde{\mathbf{W}} \tilde{\mathbf{x}}_k$	$\tilde{\mathbf{e}}_k = \left(\mathbf{I} - \tilde{\mathbf{Q}}_k^{-1} \tilde{\mathbf{B}}_k \tilde{\mathbf{W}} \right) \tilde{\mathbf{x}}_k$

b. Al igual que en los modelos locales, cada uno de los K modelos obtenidos requiere la determinación del número de componentes (R_k o A_k) apropiado que permita representar a $\tilde{\mathbf{X}}_k^0$ adecuadamente.

Tabla 3.10: Estadísticos y límites de control para los modelos *MPCA* y *MICA* - Modelos Evolutivos

Modelos Evolutivos (Ramaker y Van Sprang, 2004)			
<i>EMPCA</i>			
\mathbb{L}^{R_k}	\mathbb{L}^{kJ-R_k}	Residuos	
$D_{i,k} = \tilde{\mathbf{t}}_{i,k}^T \tilde{\Lambda}_k^{-1} \tilde{\mathbf{t}}_{i,k}$	$T_{H_{i,k}}^2 = \tilde{\mathbf{t}}_{i,k}^T \tilde{\Lambda}_k^{-1} \tilde{\mathbf{t}}_{i,k}$	$Q_{i,k} = \sum_{j=1}^{kJ} \tilde{e}_{i,j,k}^2$	
<i>Distribuciones y valores críticos (asumiendo multinormalidad en las CPs)</i>			
<i>Etapa I</i>	$D_k \approx \left(\frac{R_k(I^2 - I)}{I(I - R_k)} \right) \mathcal{F}_{R_k, I - R_k}$	$T_{H,k}^2 \approx \kappa \mathcal{F}_{(kJ - R_k), I - (kJ - R_k), \alpha}$ $\kappa = \left(\frac{(kJ - R_k)(I^2 - I)}{I(I - (kJ - R_k))} \right)$	$SPE_{C,k} = \theta_1 \left[1 - \frac{\theta_2 h_0 (1 - h_0)}{\theta_1^2} + \frac{z_\alpha (2\theta_2 h_0^2)^{1/2}}{\theta_1} \right]^{1/h_0}$ $h_{0,k} = 1 - \frac{2\theta_{1,k} \theta_{3,k}}{3\theta_{2,k}^2}$ $\theta_{m,r} = \sum_{r=R_k+1}^{kJ} \lambda_r^m \quad m = 1, 2, 3$
<i>Etapa II</i>	$D_k = \tilde{\mathbf{t}}_k^T \tilde{\mathbf{S}}_k^{-1} \tilde{\mathbf{t}}_k$	$T_{H_k}^2 = \tilde{\mathbf{t}}_k^T \tilde{\mathbf{S}}_k^{-1} \tilde{\mathbf{t}}_k$	$SPE_k = \sum_{j=1+(k-1)J}^{kJ} \tilde{e}_j^2$
<i>Etapa I</i>	$D_k \approx \left(\frac{R_k(I^2 - I)}{I(I - R_k)} \right) \mathcal{F}_{R_k, I - R_k, \alpha}$	$T_{H,k}^2 \approx \kappa \mathcal{F}_{(kJ - R_k), I - (kJ - R_k), \alpha}$ $\kappa = \left(\frac{(kJ - R_k)(I^2 - I)}{I(I - (kJ - R_k))} \right)$	$SPE_{C,k} = g_k \chi_{h_k, \alpha}^2$ $g_k = \frac{v_k}{2m_k}$ $h_k = 2m_k^2 / v_k$ <i>m y v definidas en ecuación (2.29)</i>
<i>EMICA</i>			
\mathbb{L}^{R_k}	\mathbb{L}^{kJ-R_k}	Residuos	
$I_{d_{i,k}}^2 = \tilde{\Psi}_{i,k}^T \tilde{\Psi}_{i,k}$	$I_{e_{i,k}}^2 = \tilde{\Psi}_{i,k}^T \tilde{\Psi}_{i,k}$	$Q_{i,k} = \sum_{j=1}^{kJ} \tilde{e}_{i,j,k}^2$	
<i>Etapa II</i>	$I_{d_k}^2 = \tilde{\Psi}_k^T \tilde{\Psi}_k$	$I_{e_k}^2 = \tilde{\Psi}_k^T \tilde{\Psi}_k$	$SPE_k = \sum_{j=1+(k-1)J}^{kJ} \tilde{e}_j^2$
<ul style="list-style-type: none"> Los límites de control para los estadísticos I_d, I_e y SPE en <i>MICA</i> se obtienen mediante el método <i>KDE</i> (Apéndice 3). 			

- c. La aplicación del estadístico de Hawkins demandaría un gran esfuerzo de cálculo para las componentes excluidas en cada uno de los K modelos durante la *Etapa I*.

Tabla 3.11: Cálculo de las *CVs* para *MPCA* y *MICA* - Modelos Evolutivos

Modelos Evolutivos (Ramaker y Van Sprang, 2004)		
<i>EMPCA</i>		
$c_{j,k}^D = \tilde{\mathbf{t}}_k^T \tilde{\mathbf{S}}_k^{-1} \left[\tilde{\mathbf{x}}_{j,k} \tilde{\mathbf{P}}_{k,(j,:)} \left(\tilde{\mathbf{P}}_k^T \tilde{\mathbf{P}}_k \right)^{-1} \right]^T$	$c_{j,k}^{T^2} = \tilde{\mathbf{t}}_k^T \tilde{\mathbf{S}}_k^{-1} \left[\tilde{\mathbf{x}}_{j,k} \tilde{\mathbf{P}}_{k,(j,:)} \left(\tilde{\mathbf{P}}_k^T \tilde{\mathbf{P}}_k \right)^{-1} \right]^T$	$c_{j,k}^{SPE} = e_{j,k}^2$
<i>EMICA</i>		
$c_k^{I_d^2} = \frac{\ \tilde{\Psi}_k\ }{\ \tilde{\mathbf{Q}}^{-1} \tilde{\mathbf{B}} \tilde{\Psi}_k\ } \tilde{\mathbf{Q}}^{-1} \tilde{\mathbf{B}} \tilde{\mathbf{W}} \tilde{\mathbf{x}}_k; \quad c_{j,k}^{I_d^2} = c_k^{I_d^2}(j)$	$c_k^{I_e^2} = \frac{\ \tilde{\Psi}_k\ }{\ \tilde{\mathbf{Q}}_k^{-1} \tilde{\mathbf{B}} \tilde{\Psi}_k\ } \tilde{\mathbf{Q}}_k^{-1} \tilde{\mathbf{B}} \tilde{\mathbf{W}} \tilde{\mathbf{x}}_k; \quad c_{j,k}^{I_e^2} = c_k^{I_e^2}(j)$	$c_{j,k}^{SPE} = e_{j,k}^2$

3.3.5 Modelos de Ventana Móvil (VM)

El concepto de ventana móvil se utiliza ampliamente en el área del control de procesos como una herramienta que permite enfocar el análisis sobre un período de tiempo determinado. Lennox y col. (2001) fueron los primeros en proponer el uso de ventanas móviles como metodología en el caso del monitoreo de procesos *batch*.

El primer paso de esta estrategia consiste en desdoblarse \mathbf{X} mediante un desdoblamiento *BW* y luego realizar el escalado de la matriz resultante. A continuación la matriz \mathbf{X} se descompone en K arreglos $\tilde{\mathbf{X}}_k$ a partir de los cuales se genera un modelo independiente en *VLs* para cada uno de ellos. Las matrices $\tilde{\mathbf{X}}_k$ se obtienen como:

$$\begin{cases} \tilde{\mathbf{X}}_k = \mathbf{X}_{(:,1:kJ)} \Leftrightarrow k \leq \Delta k \\ \tilde{\mathbf{X}}_k = \mathbf{X}_{(:,(k-\Delta k+1)J:kJ)} \Leftrightarrow k > \Delta k \end{cases} \quad (3.6)$$

donde Δk es un número natural, conocido como “ancho de la ventana” ($\Delta k \in \mathbb{N}$, $1 \leq \Delta k \leq K$), que define el horizonte de tiempo que se considera en los modelos.

La Figura 3.11 muestra un esquema del procedimiento requerido para obtener $\tilde{\mathbf{X}}_k$ para cada uno de los períodos de muestreo (k).

En la ecuación (3.6), si $\Delta k = 1$ los K modelos de VM serán en realidad modelos locales como los discutidos en la Sección 3.3.3 mientras que si $\Delta k = K$ los modelos de VM resultarán equivalentes a los ME de la Sección 3.3.4.

Las principales ventajas de los modelos de VM son:

- a. No requieren predicciones de estados futuros del proceso.
- b. La calidad de la estimación de las matrices de covarianza para cada ventana de tiempo puede controlarse ajustando el ancho de la ventana (Δk).
- c. La cantidad de información sobre la historia previa del *batch* incorporada en el análisis puede manipularse mediante Δk .

Por otra parte, las desventajas más importantes asociadas a los modelos de VM son:

- a. Para valores grandes de Δk , (cuando $(k - \Delta k + 1)J > I$) las matrices de covarianza $\tilde{\mathbf{Z}}_k$ a partir de las cuales se calculan las matrices de rotación ($\tilde{\mathbf{P}}_k$ o $\tilde{\mathbf{W}}_k$ según sea el caso) son, por lo general, pobremente estimadas.
- b. Al igual que en los modelos locales y ME , cada uno de los K modelos obtenidos requerirá la determinación del número de componentes (R_k o A_k) apropiado para representar a $\tilde{\mathbf{X}}_k^0$ adecuadamente.

c. La calidad de la estimación de las matrices de correlación $\tilde{\mathbf{Z}}_k$ y la longitud del horizonte de tiempo considerado plantean una situación de compromiso para los valores de Δk .

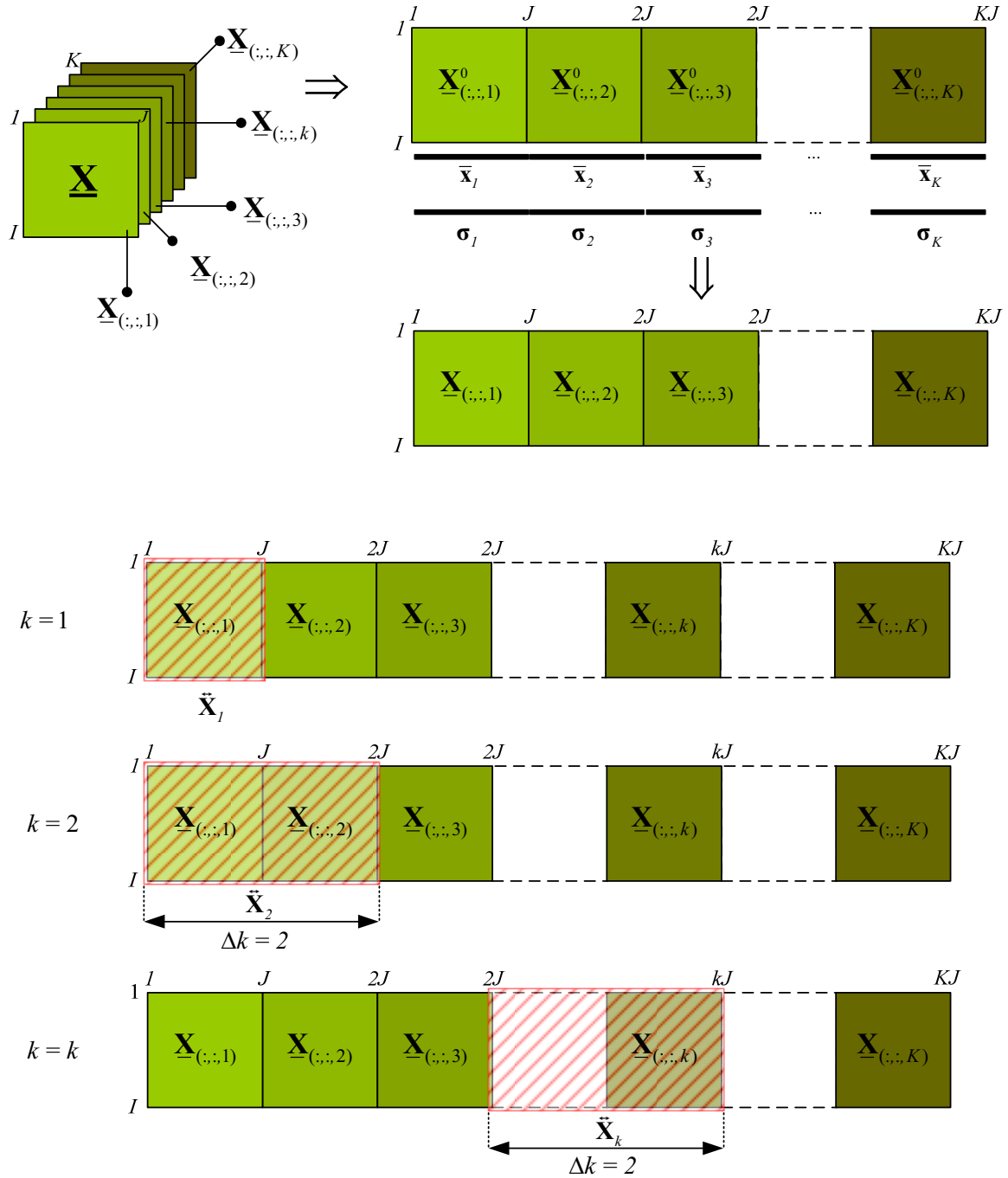


Figura 3.11: Esquema de desdoblamiento y escalado de \mathbf{X} para el enfoque de VM

Las ecuaciones para los modelos, métricas y *CVs* cuando se aplica la estrategia de *VM* se detallan en las Tablas 3.12 a 3.14.

Tabla 3.12: Modelos *MPCA* y *MICA* - Modelos de *VM*

Ventana Móvil (Lennox y col., 2001)			
$\underline{\mathbf{X}} \Rightarrow \tilde{\mathbf{X}}_k^0$ $\left(\begin{smallmatrix} I \times J \times K \\ I \times \delta \end{smallmatrix} \right)$ $\tilde{\mathbf{X}}_k^0 = \underline{\mathbf{X}}_{(:, :, (k-\Delta k+1)J : kJ)} \Rightarrow \tilde{\mathbf{X}}_k^0 = \left[\underline{\mathbf{X}}_{(:, :, k-\Delta k+1)}^0 \cdots \underline{\mathbf{X}}_{(:, :, k)}^0 \right]$ $\delta = \begin{cases} kJ \Leftrightarrow k \leq \Delta k \\ \Delta kJ \Leftrightarrow k > \Delta k \end{cases}$			
<i>MWMPCA</i>			
<i>Etapa I</i>	Proyecciones $\tilde{\mathbf{T}}_k = \tilde{\mathbf{X}}_k \tilde{\mathbf{P}}_k$ $\left(\begin{smallmatrix} I \times R_k \\ I \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} I \times \delta \\ \delta \times R_k \end{smallmatrix} \right)$ $\tilde{\tilde{\mathbf{T}}}_k = \tilde{\tilde{\mathbf{X}}}_k \tilde{\tilde{\mathbf{P}}}_k$ $\left(\begin{smallmatrix} I \times \delta - R_k \\ I \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} I \times \delta \\ \delta \times \delta - R_k \end{smallmatrix} \right)$	Reconstrucciones $\hat{\tilde{\mathbf{X}}}_k = \tilde{\mathbf{X}}_k \tilde{\mathbf{P}}_k \tilde{\mathbf{P}}_k^T$ $\left(\begin{smallmatrix} I \times \delta \\ I \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times R_k \\ R_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} R_k \times \delta \\ R_k \times \delta \end{smallmatrix} \right)$	Residuos $\tilde{\mathbf{E}}_k = \tilde{\mathbf{X}}_k \left(\mathbf{I} - \tilde{\mathbf{P}}_k \tilde{\mathbf{P}}_k^T \right)$ $\left(\begin{smallmatrix} I \times \delta \\ I \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} I \times \delta \\ \delta \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times R_k \\ R_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} R_k \times \delta \\ R_k \times \delta \end{smallmatrix} \right)$
<i>Etapa II</i>	$\tilde{\mathbf{t}}_k = \tilde{\mathbf{P}}_k^T \tilde{\mathbf{x}}_k$ $\left(\begin{smallmatrix} R_k \times I \\ R_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} R_k \times \delta \\ \delta \times I \end{smallmatrix} \right)$ $\tilde{\tilde{\mathbf{t}}}_k = \tilde{\tilde{\mathbf{P}}}_k^T \tilde{\tilde{\mathbf{x}}}_k$ $\left(\begin{smallmatrix} \delta - R_k \times I \\ \delta - R_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times I \\ \delta \times \delta - R_k \times \delta \end{smallmatrix} \right)$	$\tilde{\mathbf{x}}_k = \tilde{\mathbf{P}}_k \tilde{\mathbf{P}}_k^T \tilde{\mathbf{x}}_k$ $\left(\begin{smallmatrix} \delta \times I \\ \delta \times I \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times R_k \\ R_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} R_k \times \delta \\ R_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times I \\ \delta \times I \end{smallmatrix} \right)$	$\tilde{\mathbf{e}}_k = \left(\mathbf{I} - \tilde{\mathbf{P}}_k \tilde{\mathbf{P}}_k^T \right) \tilde{\mathbf{x}}_k$ $\left(\begin{smallmatrix} \delta \times I \\ \delta \times I \end{smallmatrix} \right) \left(\begin{smallmatrix} I \times \delta \\ \delta \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times R_k \\ R_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} R_k \times \delta \\ R_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times I \\ \delta \times I \end{smallmatrix} \right)$
<i>MWMICA</i>			
<i>Etapa I</i>	Proyecciones $\tilde{\Psi} = \tilde{\mathbf{B}}_k^T \tilde{\mathbf{Q}}_k \tilde{\mathbf{X}}_k^T$ $\left(\begin{smallmatrix} A_k \times I \\ A_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} A_k \times \delta \\ \delta \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times I \\ \delta \times I \end{smallmatrix} \right)$ $\tilde{\tilde{\Psi}} = \tilde{\tilde{\mathbf{B}}}_k^T \tilde{\tilde{\mathbf{Q}}}_k \tilde{\tilde{\mathbf{X}}}_k^T$ $\left(\begin{smallmatrix} \delta - A_k \times I \\ \delta - A_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times \delta \\ \delta \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times I \\ \delta \times I \end{smallmatrix} \right)$	Reconstrucciones $\hat{\tilde{\mathbf{X}}}_k^T = \tilde{\mathbf{Q}}_k^{-1} \tilde{\mathbf{B}}_k \tilde{\mathbf{W}} \tilde{\mathbf{X}}_k^T$ $\left(\begin{smallmatrix} \delta \times I \\ \delta \times I \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times \delta \\ \delta \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times A_k \\ A_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} A_k \times \delta \\ A_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times I \\ \delta \times I \end{smallmatrix} \right)$	Residuos $\tilde{\mathbf{E}}_k^T = \left(\mathbf{I} - \tilde{\mathbf{Q}}_k^{-1} \tilde{\mathbf{B}}_k \tilde{\mathbf{W}} \right) \tilde{\mathbf{X}}_k^T$ $\left(\begin{smallmatrix} \delta \times I \\ \delta \times I \end{smallmatrix} \right) \left(\begin{smallmatrix} I \times \delta \\ \delta \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times \delta \\ \delta \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times A_k \\ A_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} A_k \times \delta \\ A_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times I \\ \delta \times I \end{smallmatrix} \right)$
<i>Etapa II</i>	$\tilde{\Psi}_k = \tilde{\mathbf{B}}_k^T \tilde{\mathbf{Q}}_k \tilde{\mathbf{x}}_k$ $\left(\begin{smallmatrix} A_k \times I \\ A_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} A_k \times \delta \\ \delta \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times \delta \\ \delta \times I \end{smallmatrix} \right)$ $\tilde{\tilde{\Psi}}_k = \tilde{\tilde{\mathbf{B}}}_k^T \tilde{\tilde{\mathbf{Q}}}_k \tilde{\tilde{\mathbf{x}}}_k$ $\left(\begin{smallmatrix} \delta - A_k \times I \\ \delta - A_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times \delta \\ \delta \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times \delta \\ \delta \times I \end{smallmatrix} \right)$	$\tilde{\tilde{\mathbf{x}}}_k = \tilde{\mathbf{Q}}_k^{-1} \tilde{\mathbf{B}}_k \tilde{\mathbf{W}} \tilde{\tilde{\mathbf{x}}}_k$ $\left(\begin{smallmatrix} \delta \times I \\ \delta \times I \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times \delta \\ \delta \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times A_k \\ A_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} A_k \times \delta \\ A_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times I \\ \delta \times I \end{smallmatrix} \right)$	$\tilde{\mathbf{e}}_k = \left(\mathbf{I} - \tilde{\mathbf{Q}}_k^{-1} \tilde{\mathbf{B}}_k \tilde{\mathbf{W}} \right) \tilde{\tilde{\mathbf{x}}}_k$ $\left(\begin{smallmatrix} \delta \times I \\ \delta \times I \end{smallmatrix} \right) \left(\begin{smallmatrix} I \times \delta \\ \delta \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times \delta \\ \delta \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times A_k \\ A_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} A_k \times \delta \\ A_k \times \delta \end{smallmatrix} \right) \left(\begin{smallmatrix} \delta \times I \\ \delta \times I \end{smallmatrix} \right)$

Tabla 3.13: Estadísticos y límites de control para los modelos *MPCA* y *MICA* - Modelos de *VM*

Ventana Móvil (Lennox y col., 2001)		
EMPCA		
\mathbb{I}_{R_k}	$\mathbb{I}_{(k-\Delta k+I)J-R_k}$	Residuos
$D_{i,k} = \tilde{\mathbf{t}}_{i,k}^T \tilde{\Lambda}_k^{-1} \tilde{\mathbf{t}}_{i,k}$	$T_{H_{i,k}}^2 = \tilde{\mathbf{t}}_{i,k}^T \tilde{\Lambda}_k^{-1} \tilde{\mathbf{t}}_{i,k}$	$Q_{i,k} = \sum_{j=1}^{\delta J} e_{i,j,k}^2$
<i>Distribuciones y valores críticos (asumiendo multinormalidad en las CPs)</i>		
Etapla I $D_k \approx \left(\frac{R_k(I^2 - I)}{I(I - R_k)} \right) \mathcal{F}_{R_k, I - R_k}$	$T_{H,k}^2 \approx \kappa \mathcal{F}_{(\delta - R_k), I - (\delta - R_k), \alpha}$ $\kappa = \left(\frac{(\delta - R_k)(I^2 - I)}{I(I - (\delta - R_k))} \right)$ $\delta = (k - \Delta k + I)J$	$Q_{C,k} = \theta_1 \left[1 - \frac{\theta_2 h_0 (1 - h_0)}{\theta_1^2} + \frac{z_\alpha (2\theta_2 h_0^2)^{1/2}}{\theta_1} \right]^{1/h_0}$ $h_{0,k} = 1 - \frac{2\theta_{1,k} \theta_{3,k}}{3\theta_{2,k}^2}$ $\theta_{m,r} = \sum_{r=R_k+1}^{kj} \lambda_r^m \quad m = 1, 2, 3$
Etapla II $D_k = \tilde{\mathbf{t}}_k^T \tilde{\Lambda}_k^{-1} \tilde{\mathbf{t}}_k$	$T_{H_k}^2 = \tilde{\mathbf{t}}_k^T \tilde{\Lambda}_k^{-1} \tilde{\mathbf{t}}_k$	$SPE_k = \sum_{j=1}^{\delta J} e_{j,k}^2$
$D_k \approx \left(\frac{R_k(I^2 - I)}{I(I - R_k)} \right) \mathcal{F}_{R_k, I - R_k, \alpha}$	$T_{H,k}^2 \approx \kappa \mathcal{F}_{(\delta - R_k), I - (\delta - R_k), \alpha}$ $\kappa = \left(\frac{(\delta - R_k)(I^2 - I)}{I(I - (\delta - R_k))} \right)$	$SPE_{C,k} = g_k \chi_{h_k, \alpha}^2$ $g_k = \frac{v_k}{2m_k}$ $h_k = 2m_k^2 / v_k$ <i>m y v definidas en ecuación (2.29)</i>
EMICA		
Etapla I \mathbb{I}_{A_k}	$\mathbb{I}_{(k-\Delta k+I)J-A_k}$	Residuos
$I_{d_{i,k}}^2 = \tilde{\Psi}_{i,k}^T \tilde{\Psi}_{i,k}$	$I_{e_{i,k}}^2 = \tilde{\Psi}_{i,k}^T \tilde{\Psi}_{i,k}$	$Q_{i,k} = \sum_{j=1}^{\delta J} e_{i,j,k}^2$
Etapla II $I_{d_k}^2 = \tilde{\Psi}_k^T \tilde{\Psi}_k$	$I_{e_k}^2 = \tilde{\Psi}_k^T \tilde{\Psi}_k$	$SPE_k = \sum_{j=1}^{\delta J} e_{j,k}^2$
<ul style="list-style-type: none"> Los límites de control para los estadísticos I_d, I_e y SPE en <i>MICA</i> se obtienen mediante el método <i>KDE</i> (Apéndice 3). 		

Tabla 3.14: Cálculo de las CVs para MPCA y MICA - Modelos de VM

Ventana Móvil (Lennox y col., 2001)		
<i>EMPCA</i>		
$c_{j,k}^D = \tilde{\mathbf{t}}_k^T \tilde{\mathbf{\Lambda}}_k^{-1} \left[\tilde{\mathbf{x}}_{j,k} \tilde{\mathbf{P}}_{k,(j,:)} \left(\tilde{\mathbf{P}}_k^T \tilde{\mathbf{P}}_k \right)^{-1} \right]^T$	$c_{j,k}^{I_d^2} = \tilde{\mathbf{t}}_k^T \tilde{\mathbf{\Lambda}}_k^{-1} \left[\tilde{\mathbf{x}}_{j,k} \tilde{\mathbf{P}}_{k,(j,:)} \left(\tilde{\mathbf{P}}_k^T \tilde{\mathbf{P}}_k \right)^{-1} \right]^T$	$c_{j,k}^{SPE} = e_{j,k}^2$
<i>EMICA</i>		
$\mathbf{c}_{j,k}^{I_d^2} = \frac{\ \tilde{\Psi}_k\ }{\ \tilde{\mathbf{Q}}^{-1} \tilde{\mathbf{B}} \tilde{\Psi}_k\ } \tilde{\mathbf{Q}}^{-1} \tilde{\mathbf{B}} \tilde{\mathbf{W}} \tilde{\mathbf{x}}_k; \quad c_{j,k}^{I_d^2} = \mathbf{c}_k^{I_d^2}(j)$	$\mathbf{c}_{j,k}^{I_e^2} = \frac{\ \tilde{\Psi}_k\ }{\ \tilde{\mathbf{Q}}^{-1} \tilde{\mathbf{B}} \tilde{\Psi}_k\ } \tilde{\mathbf{Q}}^{-1} \tilde{\mathbf{B}} \tilde{\mathbf{W}} \tilde{\mathbf{x}}_k; \quad c_{j,k}^{I_e^2} = \mathbf{c}_k^{I_e^2}(j)$	$c_{j,k}^{SPE} = e_{j,k}^2$

3.3.6 Análisis Dinámicos (DPCA y DPLS)

Chen y Liu (2002) presentaron una adaptación de los enfoques dinámicos *DPCA* y *DPLS* propuestos por Ku y col. (1995) y Chen y McAvoy (1998) para procesos continuos. Las metodologías propuestas por Chen y Liu (2002), conocidas como *BDPCA* (*Batch Dynamic PCA*) y *BDPLS* (*Batch Dynamic PLS*), permiten incorporar las características dinámicas del proceso en el análisis del mismo. La auto-correlación de las variables del proceso se incorpora mediante el uso de modelos *ARMAX* (*Auto-Regresive Moving Average Exogenous*). Estos modelos hacen uso de ventanas de tiempo para cada una de las variables. De esta manera, no sólo se tiene en cuenta la correlación que existe entre las variables medidas para todo el conjunto de *batches*, sino que también se modela la correlación de las variables dentro de un mismo *batch* (Chen y Liu, 2002). Para cada uno de los *batches* en la *PR* se obtiene una matriz de correlación que sirve como estimador de la relación entre las mediciones durante la operación de un mismo *batch*. Posteriormente estas matrices se promedian para obtener la matriz de correlación que será utilizada para el cálculo de las matrices de proyección del modelo. En la literatura pueden encontrarse descripciones más detalladas de esta metodología y de sus versiones para procesos continuos (Ku y col., 1995; Chen y McAvoy, 1998; Wachs y Lewin, 1999; Chen y Liu, 2002).

3.4 Batches con Diferente Número de Observaciones

En las secciones precedentes se analizaron las estrategias más populares para el tratamiento de arreglos tri-modo en el ámbito del *SPC*. En todos los casos se considera que $\underline{\mathbf{X}} \in \mathbb{R}^{I \times J \times K}$, sin embargo, en la práctica existen situaciones que hacen que esta condición no sea satisfecha. Como se mencionó anteriormente, la operación y control de los procesos *batch* es una tarea compleja. La gran cantidad de fuentes de variación que existen, obliga a los operadores a solucionar problemas y tomar decisiones de manera constante durante el transcurso de la operación. Quizá la consecuencia más común de estos inconvenientes es que la duración del *batch* no sea uniforme. Con el objeto de cumplir con los parámetros de calidad establecidos, los operadores, basándose en su criterio y experiencia, acortan o prolongan la duración del lote según sea necesario. Así, el número total de observaciones obtenidas en cada *batch* no es siempre el mismo.

Esta situación causa muchos problemas a la hora de ordenar los *batches* a fin de obtener cualquiera de los modelos explicados en las secciones anteriores. En la literatura pueden encontrarse algunas propuestas para atacar este problema (Kassidas *y col.*, 1998; Lee *y col.*, 2003b; Ramaker *y col.*, 2003). Ramaker *y col.* (2005) aplicaron una técnica conocida como *Dynamic Time Warping (DTW)*, muy empleada en el área de reconocimiento de voz, a un conjunto de mediciones espectroscópicas. En dicho trabajo se presenta un análisis detallado de la metodología y su implementación.

La idea fundamental en el *DTW* es encontrar una función que permita minimizar las distancias entre un *batch* particular de la *PR* que se toma como base y cada uno de los otros *batches* del conjunto. La determinación de estas “*warping functions*” se hace mediante la técnica de programación dinámica, que permite resolver el problema de optimización resultante de una manera rápida y eficiente. El algoritmo de *DTW* posee algunos parámetros

asociados a las llamadas “restricciones globales” del problema de optimización que deben ser definidos por el usuario. Los valores de estos parámetros influirán en los resultados si dichas restricciones son activas. Como resultado de la aplicación de *DTW* sobre las trayectorias originales de los *batches*, se obtiene un nuevo conjunto de trayectorias de duración uniforme K_{Ref} . Debe notarse que la “deformación” a la que se somete cada uno de los *batches* de la *PR* introduce en el análisis una fuente más de incertidumbre.

3.5 Referencias

- Aguado, D., A. Ferrer, J. Ferrer y A. Seco, "Multivariate SPC of a Sequencing Batch Reactor for Wastewater Treatment". *Chemometrics and Intelligent Laboratory Systems* **2007**, 85, (1), 82-93.
- Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis Wiley Publications in Statistics, New York, John Wiley & Sons, Inc.
- Basseville, M., "Detecting Changes in Signals and Systems-A Survey". *Automatica* **1988**, 24, (3), 309-326.
- Bro, R., "PARAFAC. Tutorial and Applications". *Chemometrics and Intelligent Laboratory Systems* **1997**, 38, (2), 149-171.
- Chen, G. y T. J. McAvoy, "Predictive On-Line Monitoring of Continuous Processes". *Journal of Process Control* **1998**, 8, (5-6), 409-420.
- Chen, J. y K. C. Liu, "On-Line Batch Process Monitoring Using Dynamic PCA and Dynamic PLS Models". *Chemical Engineering Science* **2002**, 57, (1), 63-75.
- Dong, D. y T. J. McAvoy, "Batch Tracking Via Nonlinear Principal Component Analysis". *AIChE Journal* **1996**, 42, (8), 2199-2208.
- Jackson, J. E. (1991). A User's Guide to Principal Components. Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics, New York, John Wiley & Sons, Inc.
- Kassidas, A., P. A. Taylor y J. F. MacGregor, "Off-Line Diagnosis of Deterministic Faults in Continuous Dynamic Multivariable Processes Using Speech Recognition Methods". *Journal of Process Control* **1998**, 8, (5-6), 381-393.
- Kiers, H. A. L., "Hierarchical Relations among Three-Way Methods". *Psychometrika* **1991**, 56, (3), 449-470.
- Kourti, T. y J. F. MacGregor, "Process Analysis, Monitoring and Diagnosis, Using Multivariate Projection Methods". *Chemometrics and Intelligent Laboratory Systems* **1995**, 28, 3-21.
- Kourti, T. y J. F. MacGregor, "Multivariate SPC Methods for Process and Product Monitoring". *Journal of Quality Technology* **1996**, 28, (4), 409-428.
- Kourti, T., P. Nomikos y J. F. MacGregor, "Analysis, Monitoring and Fault Diagnosis of Batch Processes Using Multiblock and Multiway PLS". *Journal of Process Control* **1995**, 5, (4), 277-284.
- Ku, W., R. H. Storer y C. Georgakis, "Disturbance Detection and Isolation by Dynamic Principal Component Analysis". *Chemometrics and Intelligent Laboratory Systems* **1995**, 30, (1), 179-196.

- Lee, J. M., C. Yoo y I. B. Lee, "On-Line Batch Process Monitoring Using a Consecutively Updated Multiway Principal Component Analysis Model". *Computers and Chemical Engineering* **2003a**, 27, (12), 1903-1912.
- Lee, J. M., C. Yoo y I. B. Lee, "On-Line Batch Process Monitoring Using Different Unfolding Method and Independent Component Analysis". *Journal of Chemical Engineering of Japan* **2003b**, 36, (11), 1384-1396.
- Lennox, B., G. A. Montague, H. G. Hiden, G. Kornfeld y P. R. Goulding, "Process Monitoring of an Industrial Fed-Batch Fermentation". *Biotechnology and Bioengineering* **2001**, 74, (2), 125-135.
- Lohmöller, J. y H. Wold (1980). Three-Mode Path Models with Latent Variables and Partial Least Squares (PLS) Parameter Estimation. European Meeting of the Psychometrics Society, Groningen, Netherlands.
- Nomikos, P. y J. F. MacGregor, "Monitoring of Batch Process Using Multiway Principal Component Analysis". *AIChE Journal* **1994**, 40, (8), 1361-1375.
- Nomikos, P. y J. F. MacGregor, "Multivariate SPC Charts for Monitoring Batch Processes". *Technometrics* **1995**, 37, (1), 41-59.
- Ramaker, H.-J. y E. N. M. Van Sprang (2004). Statistical Batch Process Monitoring. Ph.D. Thesis Department of Chemical Engineering, Amsterdam, Universiteit van Amsterdam.
- Ramaker, H. J., E. N. M. Van Sprang, J. A. Westerhuis y A. K. Smilde, "Dynamic Time Warping of Spectroscopic Batch Data". *Analytica Chimica Acta* **2003**, 498, (1-2), 133-153.
- Ramaker, H. J., E. N. M. Van Sprang, J. A. Westerhuis y A. K. Smilde, "Fault Detection Properties of Global, Local and Time Evolving Models for Batch Process Monitoring". *Journal of Process Control* **2005**, 15, (7), 799-805.
- Rännar, S., J. F. MacGregor y S. Wold, "Adaptive Batch Monitoring Using Hierarchical PCA". *Chemometrics and Intelligent Laboratory Systems* **1998**, 41, (1), 73-81.
- Smilde, A. K., R. Bro y P. Geladi (2004). Multi-Way Analysis with Applications in the Chemical Sciences. John Wiley & Sons, Inc.
- Van Sprang, E. N. M., H.-J. Ramaker, J. A. Westerhuis, S. P. Gurden y A. K. Smilde, "Critical Evaluation of Approaches for On-Line Batch Process Monitoring". *Chemical Engineering Science* **2002**, 57, (18), 3979-3991.
- Wachs, A. y D. R. Lewin, "Improved PCA Methods for Process Disturbance and Failure Identification". *AIChE Journal* **1999**, 45, (8), 1688-1700.
- Westerhuis, J. A., S. P. Gurden y A. K. Smilde, "Generalized Contribution Plots in Multivariate Statistical Process Monitoring". *Chemometrics and Intelligent Laboratory Systems* **2000**, 51, (1), 95-114.

- Westerhuis, J. A., T. Kourti y J. F. MacGregor, "Comparing Alternative Approaches for Multivariate Statistical Analysis of Batch Process Data". *Journal of Chemometrics* **1999**, 13, (3-4), 397-413.
- Wold, S., "Exponentially Weighted Moving Principal Components Analysis and Projections to Latent Structures". *Chemometrics and Intelligent Laboratory Systems* **1994**, 23, (1), 149-161.
- Wold, S., P. Geladi, K. Esbensen y J. Öhman, "Multi-Way Principal Components-and PLS-Analysis". *Journal of Chemometrics* **1987**, 1, (1), 41-56.
- Wold, S., N. Kettaneh, H. Friden y A. Holmberg, "Modelling and Diagnostics of Batch Processes and Analogous Kinetic Experiments". *Chemometrics and Intelligent Laboratory Systems* **1998**, 44, (1-2), 331-340.
- Yoo, C. K., D. S. Lee y P. A. Vanrolleghem, "Application of Multiway ICA for On-Line Process Monitoring of a Sequencing Batch Reactor". *Water Research* **2004a**, 38, (7), 1715-1732.
- Yoo, C. K., J. M. Lee, P. A. Vanrolleghem y I. B. Lee, "On-Line Monitoring of Batch Processes Using Multiway Independent Component Analysis". *Chemometrics and Intelligent Laboratory Systems* **2004b**, 71, (2), 151-163.

Página en Blanco

Capítulo 4

Identificación de Situaciones Anómalas en el Espacio Original de las Mediciones

4.1	Método de Mason, Young y Tracy (MYT).....	114
4.2	Nuevo Método de Descomposición del Estadístico de Hotelling: OSS (Original Space Strategy).....	116
4.3	Equivalencia entre OSS y los Gráficos de Contribución Generalizados	122
4.4	Casos de Estudio	124
4.4.1	<i>Ejemplo Numérico</i>	124
4.4.2	<i>Reactor Tubular</i>	129
4.5	Nueva Estrategia de Monitoreo de Procesos Batch Basada en OSS	133
4.5.1	<i>Etapa I</i>	133
4.5.2	<i>Etapa II</i>	134
4.6	Consideraciones generales.....	135
4.7	Referencias.....	137

Página en Blanco

4 Identificación de Situaciones Anómalas en el Espacio Original de las Mediciones

Como se explicó en los capítulos precedentes, el monitoreo de procesos consiste en tres etapas fundamentales: detección, identificación y diagnóstico. La existencia de una falla o estado anómalo del proceso se pone de manifiesto cuando al menos uno de los estadísticos utilizados en la etapa de detección se encuentra por encima de su valor crítico en un número determinado de observaciones consecutivas. La posterior etapa de identificación consiste en determinar cuáles de las variables monitoreadas señalan la existencia de la misma. Ésto es, establecer el subconjunto de descriptores que puede indicarse como causante de una señal de alarma.

Se han propuesto varias metodologías aplicables a la etapa de identificación y diagnóstico. Algunas de ellas usan técnicas tales como lógica difusa y redes neuronales, especialmente en la etapa de diagnóstico, (Ruiz, 2001; Yélamos Ruiz, 2008), mientras que otras se basan en la descomposición de las métricas que detectan las fallas como una suma de las “contribuciones” de las variables (*CVs*) al valor final del estadístico correspondiente (MacGregor *y col.*, 1994; Nomikos, 1996). En este contexto, la técnica más difundida y utilizada es el método de los gráficos de contribución (Miller *y col.*, 1999; Westerhuis *y col.*, 2000).

En este capítulo se presenta una nueva metodología, desarrollada en el transcurso del trabajo de tesis (Alvarez *y col.*, 2007), para el cálculo de las *CVs* al estadístico de Hotelling en el espacio original definido por las mediciones. Esta estrategia presenta una ventaja comparativa frente a otros métodos aplicables en este espacio (Mason *y col.*, 1995, 1997), dado que se propone una descomposición única del estadístico T^2 , como suma de las

contribuciones de cada variable. Las bases geométricas que sustentan esta metodología permiten dar un significado físico a los valores de las contribuciones negativas que pueden resultar de la aplicación de este tipo de técnica, como así también, establecer una cota superior para la magnitud de las mismas.

Además, se ha desarrollado un nuevo método de monitoreo de procesos *batch* que emplea la descomposición del estadístico T^2 propuesta para la identificación de las variables sospechosas.

4.1 Método de Mason, Young y Tracy (MYT)

Mason *y col.* (1995, 1997) propusieron una metodología para la descomposición del estadístico T^2 , la cual permite estimar el efecto de cada una de las variables medidas en el valor final de esta métrica. Dichos autores consideraron a T^2 como la suma de varios términos. Cada sumando representa la distancia entre el valor observado de cada variable y su valor medio, considerando las distribuciones condicionales que resultan de asumir como constante el valor de las $J-1$ variables restantes. Al considerar los valores de dichas variables como fijos en los valores medidos, el entorno que define la región normal donde la j -ésima variable puede moverse se reduce considerablemente como muestra la Figura 4.1.

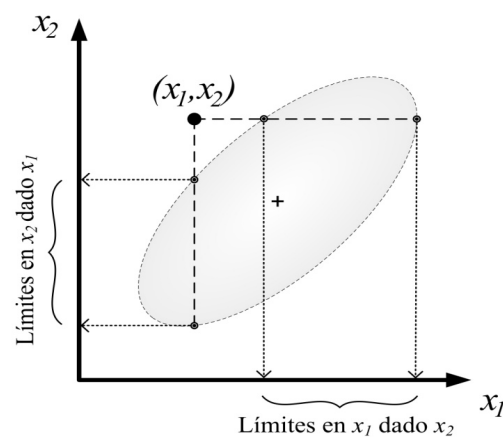


Figura 4.1: Límites para x_1 y x_2 considerando sus distribuciones condicionales

La descomposición propuesta por Mason y *col.* (1995, 1997) está formada por J términos ortogonales y se formula como:

$$T^2 = T_1^2 + T_{2 \cdot 1}^2 + \dots + T_{j \cdot 1, 2, \dots, j-1}^2 = T_1^2 + \sum_{j=1}^{J-1} T_{j+1 \cdot 1, \dots, j}^2, \quad (4.1)$$

$$T_1^2 = \frac{(x_1 - \bar{x}_1)^2}{s_1^2}, \quad (4.2)$$

$$T_{j \cdot 1, \dots, j-1}^2 = \frac{(x_j - \bar{x}_{j \cdot 1, \dots, j-1})^2}{s_{j \cdot 1, \dots, j-1}^2}, \quad (4.3)$$

$$\bar{x}_{j \cdot 1, \dots, j-1} = \bar{x}_j + \mathbf{b}_j^T (\mathbf{x}^{(j-1)} - \bar{\mathbf{x}}^{(j-1)}), \quad (4.4)$$

$$s_{j \cdot 1, \dots, j-1}^2 = s_x^2 - \mathbf{s}_{xx}^T \mathbf{S}_{xx}^{-1} \mathbf{s}_{xx}, \quad (4.5)$$

donde $\mathbf{b}_j^T = \mathbf{S}_{xx}^{-1} \mathbf{s}_{xx}$, $\mathbf{x}^{(j-1)}$ es el vector de observaciones excluida la j -ésima componente, y \mathbf{S}_{xx}^{-1} es la submatriz principal de \mathbf{S}^{-1} , de dimensiones $(J-1) \times (J-1)$, ésto es:

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{xx} & \mathbf{s}_{xx} \\ \mathbf{s}_{xx}^T & s_x^2 \end{bmatrix}. \quad (4.6)$$

Existen $J!$ ordenamientos posibles de los J componentes del vector de observación. Cada uno de estos ordenamientos producirá el mismo valor final para T^2 , pero los términos ortogonales que conforman la partición de T^2 serán diferentes en cada ordenamiento particular. Es posible corroborar que existen sólo $J \times 2^{(J-1)}$ términos independientes entre los $J \times J!$ términos que deberían ser evaluados como contribuciones potenciales a T^2 . De esta manera, la metodología *MYT* producirá resultados diferentes para las contribuciones condicionales $T_{j \cdot 1, \dots, j-1}^2$ según sea la ubicación de cada variable en el vector de observaciones.

Para eliminar estas ambigüedades, los mismos autores proponen considerar sólo dos términos para cada variable: su contribución individual (ecuación (4.2)) y el término condicional que resulta de considerar los restantes $J - 1$ términos como valores dados.

4.2 Nuevo Método de Descomposición del Estadístico de Hotelling: *OSS (Original Space Strategy)*

El estadístico de Hotelling es lo que se conoce como una forma cuadrática. Esto se debe a que la matriz de correlación o de covarianza es semidefinida positiva. Por lo tanto, los valores de T^2 , D ó I_d , según sea el caso, serán siempre mayores o iguales a cero. La ecuación (4.7) define al estadístico T^2 , y determina un paraboloides en el hiperespacio de dimensión \mathbb{R}^J , cuyo mínimo global ($T^2 = 0$) se encuentra en $\mathbf{x} = \bar{\mathbf{x}}$.

$$T^2 = (\mathbf{x} - \bar{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x} - \bar{\mathbf{x}}). \quad (4.7)$$

Consideremos el caso en el que sólo una variable, por ejemplo x_1 , se desvía de su valor medio, ésto es: $x_1 \neq \bar{x}_1$ y $x_j = \bar{x}_j \forall j \neq 1$. El valor del estadístico T^2 puede asociarse, a priori, a la desviación de x_1 , cuya contribución debería ser igual al valor final del estadístico. En el caso más simple, donde sólo se usan dos variables x_1 y x_2 como descriptores de la población, el vector de valores medios resulta $\bar{\mathbf{x}}^T = [\bar{x}_1 \ \bar{x}_2]^T$ y la matriz de covarianza \mathbf{S} se expresa como:

$$\mathbf{S} = \begin{bmatrix} s_1^2 & \rho_{12}s_1s_2 \\ \rho_{12}s_1s_2 & s_2^2 \end{bmatrix}, \quad (4.8)$$

donde s_1 y s_2 son los desvíos estándares muestrales de x_1 y de x_2 respectivamente, y ρ_{12} es el coeficiente de correlación entre las mismas. La ecuación (4.7) puede reescribirse entonces como sigue:

$$T^2 = \frac{1}{(1-\rho_{12})^2 s_1 s_2} \left(\frac{s_2}{s_1} (x_1 - \bar{x}_1)^2 + \frac{s_1}{s_2} (x_2 - \bar{x}_2)^2 - 2\rho_{12} (x_1 - \bar{x}_1)(x_2 - \bar{x}_2) \right). \quad (4.9)$$

El mínimo global de T^2 ocurrirá cuando ambas variables se encuentren en sus valores medios, esto es: $x_1 = \bar{x}_1$ y $x_2 = \bar{x}_2$. Sin embargo, cuando existe un desvío en cualquiera de los descriptores, digamos $x_1 \neq \bar{x}_1$, el mínimo para la métrica no se encontrará sobre la curva de nivel $x_2 = \bar{x}_2$ a menos que $\rho_{12} = 0$.

Sean δ_1 y δ_2 los desvíos de la observación \mathbf{x} en las direcciones x_1 y x_2 , respectivamente, los cuales constituyen el siguiente vector de desvíos:

$$\mathbf{x}^T - \bar{\mathbf{x}}^T = \boldsymbol{\delta}^T = [\delta_1 \ \delta_2]^T. \quad (4.10)$$

De acuerdo a los valores de los parámetros en \mathbf{S} , y para una dada magnitud de los desvíos ($\|\boldsymbol{\delta}\| = K$), si $\rho_{12} > 0$, el menor valor del estadístico T^2 ocurrirá si $sign(\delta_1) = sign(\delta_2)$. Por el contrario, si $\rho_{12} < 0$, T^2 será mínimo cuando $sign(\delta_1) \neq sign(\delta_2)$, según se desprende de la ecuación (4.9).

La Figura 4.2 muestra los posibles valores para el estadístico T^2 como función de x_2 mediante la intersección de la superficie $T^2 = T^2(x_1, x_2)$ con los planos definidos por $x_1 = x_1^a$ y $x_1 = x_1^b$. Como puede observarse, el mínimo valor posible para $T^2 = T^2(x_2) \Big|_{x_1=cte}$ y el valor de x_2 en el que se encuentra (x_2^{*,x_1}) dependen del valor de x_1 seleccionado.

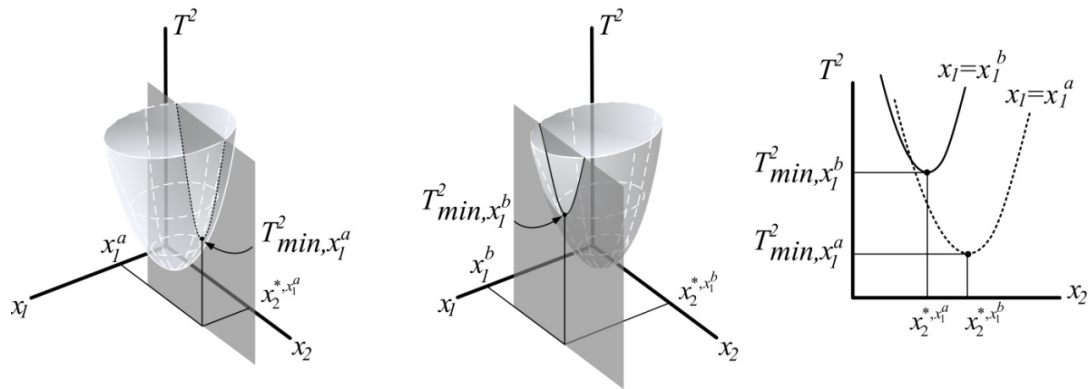


Figura 4.2: Superficies de $T^2(x_1, x_2)$ y $T^2(x_2)|_{x_1=cte}$

Para el caso general en el que se consideran J descriptores, la ecuación (4.7) se reescribe como:

$$T^2 = \sum_{i=1}^J \sum_{j=1}^J a_{i,j} x_i x_j, \tag{4.11}$$

donde $a_{i,j}$ es el i -ésimo elemento de la j -ésima columna de $\mathbf{A} = \mathbf{S}^{-1}$.

Con el fin de simplificar la nomenclatura para el tratamiento matemático, sin pérdida de generalidad, llamaremos x_k al valor de la desviación en la k -ésima variable respecto de su valor medio, ésto es:

$$x_k = x_k^{observado} - \bar{x}_k, \quad \forall j = 1..J. \tag{4.12}$$

Reformulando la ecuación (4.11), T^2 se escribe en términos de la k -ésima variable como:

$$T^2 = T_k^2 = a_{k,k} x_k^2 + \left(2 \sum_{\substack{j=1 \\ j \neq k}}^J a_{k,j} x_j \right) x_k + \sum_{\substack{i=1 \\ i \neq k}}^J \sum_{\substack{j=1 \\ j \neq k}}^J a_{i,j} x_i x_j. \tag{4.13}$$

La ecuación (4.13) es simplemente un ordenamiento particular de los sumandos, que permite poner de manifiesto el efecto de la k -ésima medición en el valor final del estadístico, para un valor dado en las restantes $J - 1$ mediciones. El valor x_k^* que produce el mínimo valor de T^2 , cuando los valores de los $J - 1$ descriptores restantes están fijos ($T_k^{2,Min}$) se calcula por medio de la derivada parcial de T^2 respecto de x_k .

$$\left. \frac{\partial T_k^2}{\partial x_k} \right|_{x_i=cte, i=1..J; i \neq k} = \left. \frac{\partial T^2}{\partial x_k} \right|_{x_i=cte, i=1..J; i \neq k} = 2a_{k,k}x_k + 2 \sum_{\substack{j=1 \\ j \neq k}}^J a_{k,j}x_j. \quad (4.14)$$

El mínimo local de $T^2 = T^2(x_k) \Big|_{x_i=cte, i=1..J; i \neq k}$ se obtiene igualando a cero la ecuación (4.14) de donde x_k^* y $T_k^{2,Min}$ resultan:

$$x_k^* = - \frac{\sum_{\substack{j=1 \\ j \neq k}}^J a_{k,j}x_j}{a_{k,k}}, \quad (4.15)$$

$$T_k^{2,Min} = - \frac{\left(\sum_{\substack{j=1 \\ j \neq k}}^J a_{k,j}x_j \right)^2}{a_{k,k}} + \sum_{\substack{i=1 \\ i \neq k}}^J \sum_{\substack{j=1 \\ j \neq k}}^J a_{i,j}x_i x_j = -a_{k,k}x_k^{*2} + \sum_{\substack{i=1 \\ i \neq k}}^J \sum_{\substack{j=1 \\ j \neq k}}^J a_{i,j}x_i x_j. \quad (4.16)$$

La diferencia entre el valor de T^2 correspondiente a la observación y $T_k^{2,Min}$ se muestra en la ecuación (4.17):

$$T^2 - T_k^{2,Min} = a_{k,k} (x_k - x_k^*)^2. \quad (4.17)$$

Esta diferencia puede usarse como una medida de la contribución de x_k al valor final de la métrica. Sin embargo, podrían aparecer algunos problemas en ciertos casos particulares.

Por ejemplo, consideremos el caso anterior, donde solo x_1 se desvía de \bar{x}_1 mientras que todas las demás variables permanecen en su valor medio. Si alguna de las variables que no presentan desviaciones tiene una alta correlación con x_1 , la contribución de la misma (calculada según la ecuación (4.17)) podría ser alta de cualquier manera. Para evitar este problema, la diferencia $T^2 - T_k^{2,Min}$ de la ecuación (4.17) se utiliza para obtener una nueva formulación. Sumando $T_k^{2,Min}$ para $k=1, \dots, J$ resulta:

$$\sum_{k=1}^J T_k^{2,Min} = \sum_{k=1}^J \left(-a_{k,k} x_k^{*2} + \sum_{\substack{i=1 \\ i \neq k}}^J \sum_{\substack{j=1 \\ j \neq k}}^J a_{i,j} x_i x_j \right) = \sum_{k=1}^J (-a_{k,k} x_k^{*2}) + \sum_{k=1}^J \left(\sum_{\substack{i=1 \\ i \neq k}}^J \sum_{\substack{j=1 \\ j \neq k}}^J a_{i,j} x_i x_j \right). \quad (4.18)$$

Teniendo en cuenta que la matriz de covarianza es una matriz simétrica y por lo tanto su inversa también lo es, la ecuación (4.18) se reduce a:

$$\sum_{k=1}^J T_k^{2,Min} = (J-2)T^2 + \sum_{k=1}^J a_{k,k} (x_k^2 - x_k^{*2}). \quad (4.19)$$

La ecuación (4.19) se reordena con el fin de obtener T^2 como una única suma de términos de la siguiente manera:

$$T^2 = \sum_{k=1}^J \frac{a_{k,k}}{2} \left[(x_k - x_k^*)^2 + (x_k^2 - x_k^{*2}) \right] = \sum_{k=1}^J a_{k,k} (x_k^2 - x_k^* x_k) = \sum_{k=1}^J c_k^{T^2}, \quad (4.20)$$

$$c_k^{T^2} = a_{k,k} (x_k^2 - x_k^* x_k) = a_{k,k} \left(x_k^2 + \frac{x_k}{a_{k,k}} \sum_{\substack{j=1 \\ j \neq k}}^J a_{k,j} x_j \right). \quad (4.21)$$

Cada término $c_k^{T^2}$ se considera como la contribución de cada variable al valor final del estadístico.

Las ecuaciones (4.20) y (4.21) proveen una metodología directa para la descomposición de T^2 como la suma de las contribuciones de todas las variables consideradas. Las mismas se obtienen en el espacio original de las mediciones usando la misma formulación para todas ellas. La descomposición del estadístico propuesta es de relevancia para la identificación de variables sospechosas, ya que evita las ambigüedades de los métodos previos.

Las bases geométricas que sustentan esta metodología permiten interpretar y dar un significado físico a los valores de $c_k^{T^2}$. La Figura 4.3 muestra los perfiles de $c_k^{T^2}$ como función de x_k . En la misma puede verse que $c_k^{T^2}$, al igual que T^2 , presenta una forma cuadrática con raíces en $x_k = 0$, $(x_k^{observado} - \bar{x}_k = 0)$ y $x_k = x_k^*$. Los valores de $c_k^{T^2}$ son negativos en el intervalo $0 < x_k < x_k^*$. El mínimo de $c_k^{T^2}$ ($c_{k,min}^{T^2}$) se encuentra en este intervalo y está ubicado en $x_k = x_k^* / 2$. Para valores de x_k fuera del intervalo $0 < x_k < x_k^*$, T^2 toma valores positivos que crecen con $|x_k|$. En el caso particular mostrado en la Figura 4.3, si $x_k < 0$ el sentido de la desviación contradice la estructura de correlación de \mathbf{S} y cualquiera sea la magnitud del apartamiento, éste resultará en valores positivos de $c_k^{T^2}$. Por el contrario, si $x_k > x_k^*$ la desviación ocurre en el sentido acorde con la estructura de correlación, pero su magnitud resulta exagerada.

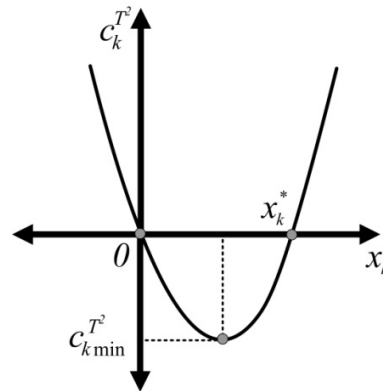


Figura 4.3: $T^2 = T^2(x_k)$ dados $x_j, \forall j \neq k$

4.3 Equivalencia entre OSS y los Gráficos de Contribución Generalizados

Los gráficos de contribución generalizados constituyen la metodología más ampliamente utilizada con propósitos de identificación en *SPC*. Westerhuis *y col.* (2000) propusieron una metodología para el cálculo de las *CVs* cuando se usan técnicas de proyección ampliando el alcance de algunas estrategias desarrolladas previamente por otros autores (Jackson, 1991; Miller *y col.*, 1999; Nomikos, 1996).

En esta sección se presenta un análisis que pone de manifiesto la equivalencia de los resultados de la estrategia presentada por Westerhuis *y col.* (2000) para el caso de *PCA* y los obtenidos mediante *OSS* cuando se retienen todas las *CPs* ($R = J$).

Las coordenadas de la matriz de datos \mathbf{X} en el espacio de *CPs* está dada por:

$$\bar{\mathbf{T}} = \mathbf{X}\bar{\mathbf{P}}, \quad (4.22)$$

donde, $\bar{\mathbf{T}}$ y $\bar{\mathbf{P}}$ representan las coordenadas de \mathbf{X} en el espacio completo de *CPs* y la matriz de cambio de coordenadas, respectivamente. El valor de la contribución de la k -ésima variable según Westerhuis *y col.* (2000) está dado por la ecuación (4.23):

$$c_k^D = \mathbf{t}_{new}^T \mathbf{B}^{-1} [x_{new,k} \mathbf{p}_k^T (\bar{\mathbf{P}}^T \bar{\mathbf{P}})^{-1}]^T. \quad (4.23)$$

La inversa de la matriz de covarianza \mathbf{B} se expresa como:

$$\mathbf{B}^{-1} = \left(\frac{\bar{\mathbf{T}}^T \bar{\mathbf{T}}}{I-1} \right)^{-1}, \quad (4.24)$$

y la proyección de cada nuevo vector de observaciones en el espacio de VLs es:

$$\mathbf{t}_{new}^T = \mathbf{x}_{new}^T \bar{\mathbf{P}}. \quad (4.25)$$

Si se considera el espacio completo de CPs , $\bar{\mathbf{P}}^T = \bar{\mathbf{P}}^{-1}$ dado que la matriz $\bar{\mathbf{P}}$ es ortonormal. En consecuencia, la ecuación (4.23) es equivalente a:

$$c_k^D = \mathbf{x}_{new}^T \bar{\mathbf{P}} \left(\frac{\bar{\mathbf{P}}^T \mathbf{X}^T \mathbf{X} \bar{\mathbf{P}}}{I-1} \right)^{-1} [x_{new,k} \mathbf{p}_k^T]^T = \mathbf{x}_{new}^T \bar{\mathbf{P}} (\bar{\mathbf{P}}^T \mathbf{S} \bar{\mathbf{P}})^{-1} \mathbf{p}_k x_{new,k}. \quad (4.26)$$

Dado que $\bar{\mathbf{P}}^T$, $\bar{\mathbf{P}}$ y \mathbf{S} son matrices cuadradas y no-singulares, puede escribirse:

$$c_k^D = \mathbf{x}_{new}^T \bar{\mathbf{P}} \bar{\mathbf{P}}^{-1} \mathbf{S}^{-1} (\bar{\mathbf{P}}^T)^{-1} \mathbf{p}_k x_{new,k}, \quad (4.27)$$

$$c_k^D = \mathbf{x}_{new}^T \mathbf{S}^{-1} \bar{\mathbf{P}} \mathbf{p}_k x_{new,k}. \quad (4.28)$$

El producto $\bar{\mathbf{P}} \mathbf{p}_k$ es un vector columna con un único elemento no nulo en la k -ésima componente, por lo que la ecuación (4.28) se reacomoda como:

$$c_k^D = \left(\sum_{j=1}^J x_{j,new} a_{jk} \right) x_{new,k} = a_{kk} x_k^2 + \left(\sum_{\substack{j=1 \\ j \neq k}}^J a_{kj} x_j \right) x_k = a_{kk} \left(x_k^2 + \frac{\sum_{\substack{j=1 \\ j \neq k}}^J a_{kj} x_j}{a_{kk}} x_k \right) = c_k^{T^2}. \quad (4.29)$$

De esta manera, queda demostrado que las contribuciones obtenidas por el método de Westerhuis *y col.* (2000) serán idénticas a las obtenidas mediante OSS cuando se retengan $R = J$ CPs.

4.4 Casos de Estudio

El desempeño de la estrategia OSS, para la identificación de las variables que señalan un evento anormal, se ilustra en esta sección mediante su aplicación a dos casos de estudio. Los valores de las contribuciones calculados mediante OSS se comparan con los obtenidos usando la metodología propuesta por Westerhuis *y col.* (2000) (ecuación (4.23)) en diversas condiciones.

4.4.1 Ejemplo Numérico

El primer caso de estudio no corresponde a un procesos químico en particular sino simplemente a un conjunto de datos propuesto por De Maesschalck *y col.* (2000) para el estudio de las propiedades del estadístico de Hotelling (ver Tabla 4.1). El conjunto de datos de referencia está formado por 20 observaciones de cuatro variables cada una ($\mathbf{X} \in \mathbb{R}^{20 \times 4}$). El vector de medias correspondiente es $\bar{\mathbf{x}} = [6 \quad 5,35 \quad 3,125 \quad 3,245]$, y las unidades de todas las variables son las mismas por lo que resulta posible el uso de matrices de covarianza.

Se consideraron además siete observaciones adicionales para evaluar el desempeño de las estrategias analizadas para la identificación de fallas (Tabla 4.2).

Puede notarse que los pares de observaciones TEST₁/ TEST₂ y TEST₃/TEST₄ tienen la misma distancia Euclídea a la media $\bar{\mathbf{x}}$, aunque algunas variables presentan desviaciones de diferente signo y magnitud.

Tabla 4.1: Población de referencia (De Maesschalck y col., 2000)

	x_1	x_2	x_3	x_4
1	4	3	1	2
2	5	4	2	3,5
3	8	7	3	4
4	8	6	5	4
5	9	7	2	3
6	6	3	5	3
7	6	5	3	2,5
8	10	8	2	3
9	2	3	1,5	3,4
10	4	4	3	3
11	6	6	6	4
12	6,5	4,5	0	2
13	9	8	5	5
14	4	5	1	1
15	4	6	3	5
16	6	7	2	4
17	2,5	4,5	6	4
18	5	5,5	8	3
19	7	5,5	1	2,5
20	8	5	3	3

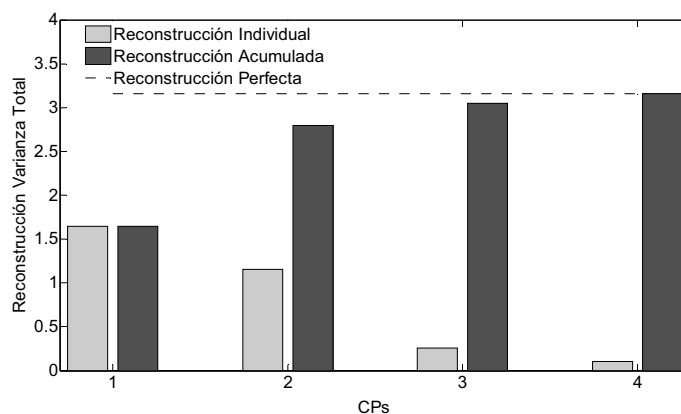
Tabla 4.2: Observaciones defectuosas. Tests de prueba

Nombre	Observación	Distancia Euclídea a \bar{x}
TEST ₁	[1,000 5,350 3,125 3,245]	5,000
TEST ₂	[11,00 5,350 3,125 3,245]	5,000
TEST ₃	[1,000 7,000 3,125 3,245]	5,265
TEST ₄	[11,00 7,000 3,125 3,245]	5,265
TEST ₅	[8,000 7,000 11,00 5,000]	8,475
TEST ₆	[2,000 8,000 8,000 7,000]	7,803
TEST ₇	[2,100 3,100 7,900 4,900]	6,769

La estrategia desarrollada por Westerhuis y col. (2000) se utiliza aquí en conjunto con *PCA* basado en la matriz de correlación. De esta manera se obtiene una mejor reconstrucción de las varianza de todas las variables que cuando se utiliza la matriz de covarianza. Las Figura 4.4 y 4.5 muestran los porcentajes de reconstrucción totales y de cada variable, obtenidos utilizando las matrices de covarianza y de correlación, respectivamente. Es evidente

que sólo las primeras dos *CPs* ($R=2$) serían suficientes para realizar el análisis. Sin embargo, con el fin de generalizar el ejemplo, se analizaron los modelos en *VLs* con $R=2, 3$ y 4 *CPs*.

(a)



(b)

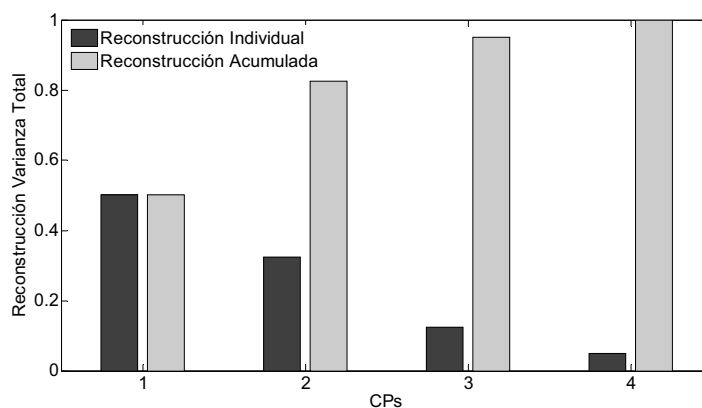


Figura 4.4: Reconstrucción de la varianza total:
(a)-Matriz de Covarianza, (b)-Matriz de Correlación

La Tabla 4.3 muestra los valores de D para distintos R y, si éstos exceden los límites correspondientes a los niveles de significación $\alpha=0,05$ o $\alpha=0,01$ son subrayados. Cabe notar que el estadístico T^2 tiene el mismo valor que D cuando todas las *CPs* son retenidas ($R=4$), por tal motivo no se lo incluyó.

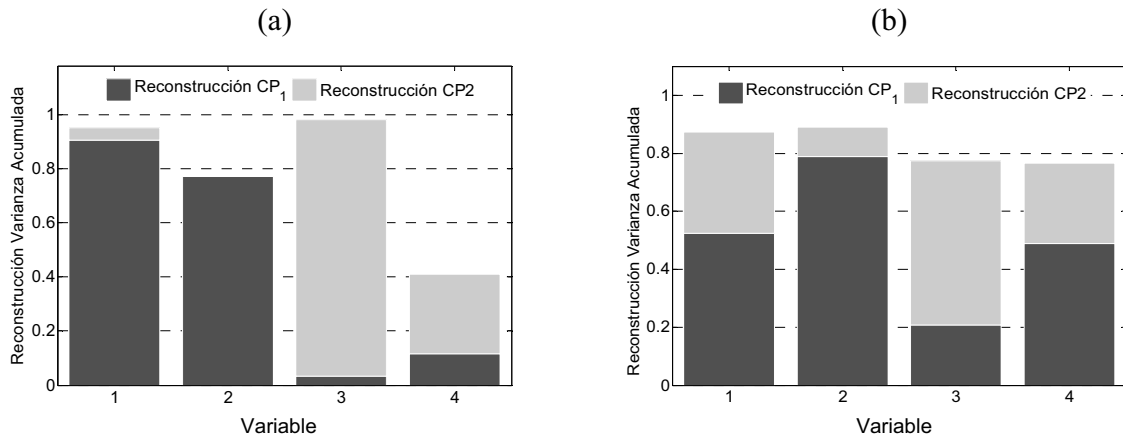


Figura 4.5: Reconstrucción de la varianza de cada variable:
 (a)-Matriz de covarianza, (b)-Matriz de correlación

Tabla 4.3: Valores de D para las observaciones de prueba

Observación	R	D	$D_{C,\alpha=0,05}$	$D_{C,\alpha=0,01}$
TEST ₁ , TEST ₂	4	11,92	14,99	23,80
TEST ₁ , TEST ₂	3	2,852	11,25	18,25
TEST ₁ , TEST ₂	2	1,718	7,88	13,33
TEST ₃	4	<u>24,49</u>	14,99	23,80
TEST ₃	3	2,198	11,25	18,25
TEST ₃	2	0,702	7,880	13,33
TEST ₄	4	5,832	14,99	23,80
TEST ₄	3	4,138	11,25	18,25
TEST ₄	2	3,315	7,880	13,33
TEST ₅	4	<u>15,36</u>	14,99	23,80
TEST ₅	3	<u>15,32</u>	11,25	18,25
TEST ₅	2	<u>10,22</u>	7,88	13,33
TEST ₆	4	<u>27,42</u>	14,99	23,80
TEST ₆	3	<u>20,34</u>	11,25	18,25
TEST ₆	2	<u>14,74</u>	7,880	13,33
TEST ₇	4	10,88	14,99	23,80
TEST ₇	3	10,12	11,25	18,25
TEST ₇	2	<u>10,12</u>	7,880	13,33

Los valores de los estadísticos para TEST₁ y TEST₂ son idénticos dado que x_1 presenta desvíos de la misma magnitud en los dos sentidos posibles, mientras que las tres variables restantes permanecen en sus valores promedio. Así, el signo de la desviación no tiene efecto en este caso particular puesto que, cualquiera sea el mismo, contradice de igual manera la estructura de correlación propuesta en S . En las observaciones TEST₃ y TEST₄ dos variables

se desvían simultáneamente de sus valores medios. La desviación en la variable x_2 es la misma en los dos casos. Por otro lado, los desvíos en la variable x_1 tienen la misma magnitud pero sentido contrario. En este caso puede notarse que $D_{\text{TEST}_3} > D_{\text{TEST}_4}$ y es mayor incluso que el valor límite correspondiente cuando todas las *CPs* son retenidas ($R = 4$). La diferencia en los valores de D para TEST_3 y TEST_4 se debe fundamentalmente a que las direcciones de las desviaciones en TEST_3 contradicen mucho más la estructura de correlación que las de TEST_4 . Además, del análisis de la Tabla 4.3 surge que los valores de D para TEST_3 son menores que sus correspondientes valores críticos (D_C) cuando se reduce la dimensión del espacio latente ($R < 4$).

Para la observación TEST_5 , D_{TEST_5} es mayor que el valor crítico para un nivel de significación $\alpha = 0,05$ pero no para $\alpha = 0,01$. Respecto de TEST_6 , D_{TEST_6} es mayor que D_C para ambos niveles de confianza. El valor del estadístico D para TEST_7 decrece lentamente cuando el número de *CPs* retenidas disminuye. Si se consideran sólo dos *CPs*, se detecta una falsa alarma para $\alpha = 0,05$.

La Tabla 4.4 muestra los valores de las *CVs* al estadístico D y T^2 para cada una de las observaciones consideradas. Las mismas se calcularon usando la ecuación (4.23) para el caso de las contribuciones generalizadas y la ecuación (4.21) para *OSS*. En el caso de las observaciones TEST_1 y TEST_2 , sólo la variable x_1 presenta un valor de contribución distinto de cero e igual, por lo tanto, al valor del estadístico correspondiente a dichas observaciones en cualquier espacio.

Respecto de TEST_3 , $c_1^D > c_2^D$ para todos los casos. Esto indica que los alejamientos en x_1 contradicen más la estructura de correlación que las desviaciones en la dirección de x_2 , en

todos los espacios analizados. Las *CVs* muestran también que, aunque $TEST_3$ y $TEST_4$ tienen la misma distancia a la media, las variaciones consideradas en $TEST_4$ contradicen menos la estructura de correlación de S , razón por la cual las contribuciones de x_1 y x_2 , son ambas menores a las de $TEST_3$ en el espacio original de las mediciones y cuando todas las *CPs* son retenidas.

Como puede observarse en la Tabla 4.4, los resultados de *OSS* son los mismos que los obtenidos aplicando la estrategia propuesta por Westerhuis *y col.* (2000) solamente cuando todas las *CPs* son consideradas. Otro punto interesante es que el signo de las contribuciones c_i^D cambia con el número de *CPs* consideradas, es decir que, una variable que sería indicada como sospechosa en un espacio parece comportarse de manera adecuada en el otro.

4.4.2 Reactor Tubular

El segundo caso de estudio consiste en un reactor tubular donde ocurre una única reacción irreversible cuya estequiometría es $A + B \rightarrow 3C$. El conjunto de variables medidas está conformado por diez mediciones: las composiciones de entrada de A, B y C, las temperaturas de alimentación al reactor y de entrada de refrigerante, el caudal de alimentación al reactor, la temperatura del reactor a los 10 y 20 metros y la temperatura y concentración de C a la salida de la unidad. Las mismas son identificadas como variables 1 a 10 en ese orden.

El modelo matemático del reactor, implementado en gPROMS[®], se utilizó para realizar 37 simulaciones, cuyos resultados se emplearon para generar la *PR*. El modelo y los datos de la *PR* se incluyen en el Apéndice 4. Además se simularon cuatro experiencias adicionales para comparar el desempeño de las diferentes estrategias, tal como se hizo en el ejemplo anterior. El test A_1 presenta un incremento en la composición del componente C en la corriente de alimentación. En el caso A_2 se simula un incremento en la temperatura de la corriente de

salida. La tercera experiencia, A_3 , tiene una reducción tanto en la temperatura del refrigerante como en la del reactor a los 10 m. El último caso considerado, A_4 , presenta un incremento en la temperatura de salida del reactor y en la fracción de C a la entrada del reactor.

El objetivo de este ejemplo es mostrar en forma comparativa los desempeños de las diferentes estrategias en la identificación de las fallas mencionadas más arriba. No se intenta proponer una estrategia para encontrar las causas originales de los desvíos.

Tabla 4.4: Valores de c_l^D y $c_l^{T^2}$ para las observaciones de prueba

Observación	c_1^D ó $c_1^{T^2}$	c_2^D ó $c_2^{T^2}$	c_3^D ó $c_3^{T^2}$	c_4^D ó $c_4^{T^2}$	D ó T^2	Estrategia
TEST ₁ , TEST ₂	11,92	0,000	0,000	0,000	11,92	4 CPs
TEST ₁ , TEST ₂	2,852	0,000	0,000	0,000	2,852	3 CPs
TEST ₁ , TEST ₂	1,718	0,000	0,000	0,000	1,718	2 CPs
TEST ₁ , TEST ₂	11,92	0,000	0,000	0,000	11,92	OSS
TEST ₃	16,59	7,906	0,000	0,000	24,49	4 CPs
TEST ₃	2,367	-0,169	0,000	0,000	2,198	3 CPs
TEST ₃	1,065	-0,362	0,000	0,000	0,702	2 CPs
TEST ₃	16,59	7,906	0,000	0,000	24,49	OSS
TEST ₄	7,256	-1,425	0,000	0,000	5,832	4 CPs
TEST ₄	3,337	0,801	0,000	0,000	4,138	3 CPs
TEST ₄	2,371	0,944	0,000	0,000	3,315	2 CPs
TEST ₄	7,256	-1,425	0,000	0,000	5,832	OSS
TEST ₅	1,024	-0,233	14,97	-0,402	15,36	4 CPs
TEST ₅	0,7743	0,121	15,10	-0,682	15,32	3 CPs
TEST ₅	-0,187	0,477	6,917	3,016	10,22	2 CPs
TEST ₅	1,024	-0,233	14,97	-0,402	15,36	OSS
TEST ₆	9,872	7,986	1,292	8,266	27,42	4 CPs
TEST ₆	3,465	0,681	0,239	15,96	20,34	3 CPs
TEST ₆	1,449	0,081	5,553	7,662	14,74	2 CPs
TEST ₆	9,872	7,986	1,292	8,266	27,42	OSS
TEST ₇	0,582	3,290	3,905	3,105	10,88	4 CPs
TEST ₇	2,626	1,261	4,242	1,996	10,13	3 CPs
TEST ₇	2,657	1,252	4,156	2,056	10,12	2 CPs
TEST ₇	0,582	3,290	3,905	3,105	10,88	OSS

De manera similar al caso anterior, se compararon los resultados obtenidos mediante OSS y los propuestos por Westerhuis *y col.* (2000). En este caso se consideraron modelos basados en PCA usando $R = 2$ y 3 CPs.

La Tabla 4.5 muestra los valores de las $c_j^{T^2}$ cuando se usa OSS en cada uno de los cuatro casos considerados. Las contribuciones más significativas en cada una de las fallas se resaltan en negrita. Para todas las corridas, los valores de T^2 exceden los valores críticos $T_{C,\alpha=0.05}^2 = 30,2$ y $T_{C,\alpha=0.01}^2 = 41,9$, y por lo tanto generan una alarma en el espacio original de mediciones.

Tabla 4.5: Valores de $c_j^{T^2}$ para las observaciones A_1 a A_4

	$c_1^{T^2}$	$c_2^{T^2}$	$c_3^{T^2}$	$c_4^{T^2}$	$c_5^{T^2}$	$c_6^{T^2}$	$c_7^{T^2}$	$c_8^{T^2}$	$c_9^{T^2}$	$c_{10}^{T^2}$
A_1	3,040	-0,020	336,8	3,770	39,85	0,540	21,59	2,770	-59,70	-3,290
A_2	-356,0	-33,10	-0,100	-207,0	-1486	5,300	74,70	4,100	1294	805,3
A_3	-523,0	35,70	1,000	-273,0	955,8	-0,900	504,0	16,20	-1287	680,0
A_4	11,75	-111,0	2,890	34,64	-912,0	1,320	-83,80	3,120	736,6	358,4

La reconstrucción de la varianza total de la PR mediante un modelo basado en PCA se muestra en la Figura 4.6. El criterio de Catell (Jackson, 1991) indica que se necesitan tres CPs para describir adecuadamente el proceso, logrando una reconstrucción de la varianza total del 75,4%.

La Tabla 4.6 contiene los valores de c_j^D calculados usando 3 CPs. Sólo el caso A_1 produce un valor del estadístico D que excede los valores críticos correspondientes ($D_{\alpha=0.05} = 9,404$, $D_{\alpha=0.01} = 14,405$). Cuando se consideran sólo dos CPs, ninguno de los casos simulados produce alarma en D (Tabla 4.7), teniendo en cuenta que los valores críticos correspondientes son $D_{\alpha=0.05} = 6,903$ y $D_{\alpha=0.01} = 11,130$.

Como puede observarse en la Tabla 4.5, para el caso A_1 , la máxima contribución en el espacio original corresponde a la tercera variable (concentración de C), lo que es congruente con la falla simulada. Los mismos resultados se obtienen con las otras estrategias (Tabla 4.6 y Tabla 4.7), aunque en el último caso no se detecta ninguna anomalía.

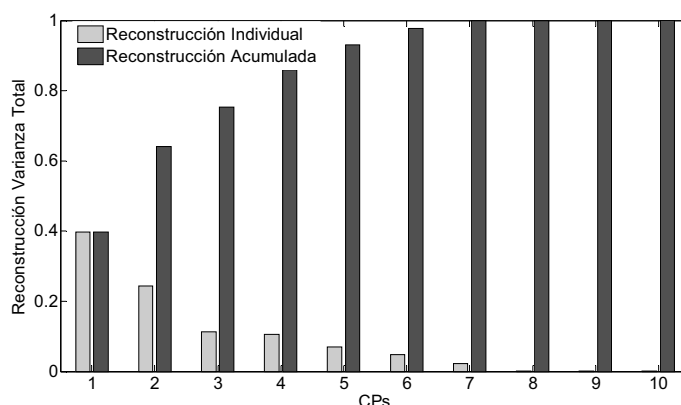


Figura 4.6: Reconstrucción de la varianza total (matriz de correlación)

Tabla 4.6: Valores de c_i^D ($R = 3$) para las observaciones A_1 a A_4

	c_1^D	c_2^D	c_3^D	c_4^D	c_5^D	c_6^D	c_7^D	c_8^D	c_9^D	c_{10}^D
A_1	0,436	-0,010	113,1	0,125	-0,036	0,084	-0,14	-0,180	0,019	0,204
A_2	0,035	0,007	-0,010	0,062	0,081	-0,030	0,245	-0,150	0,012	-0,06
A_3	0,133	0,098	-0,150	0,259	0,097	0,0435	0,3343	-0,189	0,116	0,011
A_4	-0,010	0,142	-0,060	0,081	0,157	0,083	0,250	-0,160	0,155	0,106

Tabla 4.7: Valores de c_i^D ($R = 2$) para las observaciones A_1 a A_4

	c_1^D	c_2^D	c_3^D	c_4^D	c_5^D	c_6^D	c_7^D	c_8^D	c_9^D	c_{10}^D
A_1	-0,045	-0,001	2,981	0,025	0,014	-0,010	-0,045	-0,034	0,010	-0,026
A_2	-0,060	-0,099	-0,007	0,275	0,868	-0,032	-0,062	0,412	0,702	0,001
A_3	-0,035	0,007	-0,010	0,062	0,081	-0,033	0,245	-0,146	0,119	-0,057
A_4	-0,003	0,132	-0,051	0,079	0,163	0,085	0,249	-0,159	0,155	0,109

Para la falla A_2 , la descomposición en el espacio original identifica un problema en la temperatura de salida del reactor (variable 9), situación que no ocurre en ninguno de los espacios latentes considerados.

En el caso A_3 , la Tabla 4.5 muestra que las contribuciones más importantes al estadístico T^2 son la temperatura del refrigerante (variable 5), la temperatura del reactor a los 10 m (variable 7) y la concentración de salida del componente C (variable 10). Por el contrario, si

se consideran dos o tres *CPs*, la única contribución importante se atribuye a la temperatura del reactor a los 10 m.

Los valores de las *CVs* para el caso A_4 indican desviaciones en la temperatura y concentración de *C* a la salida del reactor (variables 9 y 10). Esta situación no es detectada por los estadísticos en ninguno de los espacios latentes (Tabla 4.6 y Tabla 4.7). Debe mencionarse que algunas de las alarmas que no fueron detectadas por el estadístico *D* en los espacios de *VLs* podrían haber sido detectadas por el estadístico *SPE*, que se utiliza de manera complementaria al estadístico *D* para monitorear los residuos de las proyecciones.

4.5 Nueva Estrategia de Monitoreo de Procesos *Batch* Basada en *OSS*

La descomposición del estadístico T^2 propuesta en la Sección 4.2 se emplea para el desarrollo de una nueva estrategia de monitoreo de procesos *batch* aplicable en el espacio de las mediciones originales. La misma se basa en la formulación de modelos locales para cada observación.

La implementación de la técnica también se realiza en dos etapas: *Etapa I* (modelado fuera de línea) y *Etapa II* (monitoreo en línea). Las características distintivas de las mismas se presentan a continuación.

4.5.1 *Etapa I*

Inicialmente se obtiene información de los datos contenidos en las capas verticales de la matriz tridimensional $\underline{\mathbf{X}}$, que representa la *PR*, con el fin de utilizarla durante la etapa de monitoreo en línea.

Para cada capa vertical de $\underline{\mathbf{X}}$ ($\underline{\mathbf{X}}_{(:, :, k)}, \forall k = 1, 2, \dots, K$) se calcula el vector de medias muestrales $\bar{\mathbf{x}}_k$ y el vector de desvíos estándar $\boldsymbol{\sigma}_k$, con los cuales se estandariza $\underline{\mathbf{X}}_{(:, :, k)}$. Luego

se evalúa su matriz de correlación \mathbf{Z}_k y se la invierte. Los vectores $\bar{\mathbf{x}}_k$ y $\boldsymbol{\sigma}_k$ y la inversa de la matriz de correlación, \mathbf{Z}_k^{-1} , son almacenados.

Además la matriz $\underline{\mathbf{X}}_{(c, :, k)}$ estandarizada se utiliza para determinar el valor crítico del estadístico T_k^2 para un dado nivel de significación α . Para ello se calcula el estadístico T^2 para cada uno de los *batches* y se estima su función de densidad de probabilidad mediante el método *KDE* explicado en el Apéndice 3. Se selecciona como valor crítico del estadístico aquel para el cual la probabilidad acumulada de T_k^2 es 0,95.

También se calcula el valor medio y el desvío estándar de las contribuciones de la j -ésima variable en la k -ésima observación, $c_{j,k}^{T^2}$, representados como $\bar{c}_{j,k}$ y $\sigma_{c_{j,k}}$.

4.5.2 Etapa II

Durante la *Etapa II*, cada nueva observación \mathbf{x}_k^0 se centra, respecto de la media muestral $\bar{\mathbf{x}}_k$ y se estandariza utilizando el vector $\boldsymbol{\sigma}_k$. Luego se calcula el estadístico de Hotelling en función de la nueva medición estandarizada y \mathbf{Z}_k^{-1} . Si se excede el valor crítico, se realiza la identificación de las variables sospechosas mediante el cálculo de las *CVs* propuesto en la Sección 4.2. Una medición se identifica como sospechosa si excede alguno de los siguientes valores límites:

$$\tau_1 = \bar{c}_{j,k} + 2\sigma_{c_{j,k}} \quad (4.30)$$

$$\tau_2 = \bar{c}_{j,k} + 3\sigma_{c_{j,k}} \quad (4.31)$$

4.6 Consideraciones generales

La estrategia *OSS* permite descomponer el estadístico de Hotelling como una suma única de las *CVs* en el espacio original de las mediciones. Además proporciona una clara comprensión del significado físico de las contribuciones negativas y determina un valor límite para las mismas.

En comparación con otras técnicas aplicables en el mismo espacio, tiene la ventaja de evitar ambigüedades en la identificación de las variables que señalan el estado fuera de control y comparte con ellas el requerimiento de que la matriz de covarianza de los datos esté bien condicionada pues debe ser invertida.

Si se satisfacen las condiciones necesarias para la inversión de dicha matriz de covarianza, la estrategia *OSS* permite monitorear el proceso utilizando un sólo estadístico. Además evita la posible pérdida de información originada por la proyección de los datos en un espacio de variables latentes de dimensión incorrecta. Esto puede originar falsas detecciones, como en el caso del $TEST_7$ en el cual se produce un error *Tipo I*, o errores *Tipo II* como los que ocurren con $TEST_3$.

Los resultados de *OSS* son equivalentes a los obtenidos para las contribuciones generalizadas al estadístico D , definidas por Westerhuis *y col.* (2000), sólo cuando todas las *CPs* se retienen en un *PCA*. No obstante, cabe aclarar que la metodología también permite calcular las contribuciones en espacios latentes, para $R < J$, si la inversa de la matriz de covarianza se pre y post multiplica por las correspondientes matrices de cambios de coordenadas.

La descomposición del estadístico T^2 mediante la técnica *OSS* ha sido utilizada con diferentes propósitos. Entre ellos se destaca su incorporación como técnica de identificación

en la nueva estrategia de monitoreo de procesos *batch* presentada en la Sección 4.5 y su aplicación en un nuevo método de tratamiento de errores sistemáticos de valor constante en las mediciones de procesos continuos (Sánchez *y col.*, 2008).

4.7 Referencias

- Alvarez, C. R., A. Brandolin y M. C. Sanchez, "On the Variable Contributions to the D-Statistic". *Chemometrics and Intelligent Laboratory Systems* **2007**, 88, (2), 189-196.
- De Maesschalck, R., D. Jouan-Rimbaud y D. L. Massart, "The Mahalanobis Distance". *Chemometrics and Intelligent Laboratory Systems* **2000**, 50, (1), 1-18.
- Jackson, J. E. (1991). A User's Guide to Principal Components. Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics, New York, John Wiley & Sons, Inc.
- MacGregor, J. F., C. Jaeckle, C. Kiparissides y M. Koutoudi, "Process Monitoring and Diagnosis by Multiblock PLS Methods". *AIChE Journal* **1994**, 40, (5), 826-838.
- Mason, R. L., N. D. Tracy y J. C. Young, "Decomposition of T^2 for Multivariate Control Chart Interpretation". *Journal of Quality Technology* **1995**, 27, (2), 99-108.
- Mason, R. L., N. D. Tracy y J. C. Young, "A Practical Approach for Interpreting Multivariate T^2 Control Chart Signals". *Journal of Quality Technology* **1997**, 29, (4), 396-406.
- Miller, P., R. E. Swanson y C. F. Heckler, "Contribution Plots: The Missing Link in Multivariate Quality Control". *International Journal of Applied Mathematics and Computer Science*. **1999**, 8, (4), 775-792.
- Nomikos, P., "Detection and Diagnosis of Abnormal Batch Operations Based on Multi-Way Principal Component Analysis". *ISA Transactions* **1996**, 35, (3), 259-266.
- Ruiz, D. (2001). Fault Diagnosis in Chemical Plants Integrated to the Information System. Tesis Doctoral, Enginyria Química, Barcelona, Universitat Politècnica de Catalunya.
- Sánchez, M. C., C. R. Alvarez y A. Brandolin, "A Multivariate Statistical Process Control Procedure for Bias Identification in Steady-State Processes". *AIChE Journal* **2008**, 54, (8), 2082-2088.
- Westerhuis, J. A., S. P. Gurden y A. K. Smilde, "Generalized Contribution Plots in Multivariate Statistical Process Monitoring". *Chemometrics and Intelligent Laboratory Systems* **2000**, 51, (1), 95-114.
- Yélamos Ruiz, I. (2008). A Global Approach for Supporting Operators' Decision-Making Dealing with Plant Abnormal Events. Tesis Doctoral, Enginyria Química, Barcelona, Universitat Politècnica de Catalunya.

Página en Blanco

Capítulo 5

Descripción del Caso de Estudio

5.1	Reactor de Polimerización de Metacrilato de Metilo (MMA) en Emulsión.....	141
5.1.1	<i>Componentes de los Sistemas de Polimerización en Emulsión</i>	142
5.1.2	<i>Modelo Matemático</i>	145
5.1.2.1	Mecanismo de Reacción	145
5.1.2.2	Descripción del Reactor.....	149
5.1.2.3	Balances de Masa y Energía	150
5.1.3	<i>Implementación y Validación del Modelo</i>	163
5.1.4	<i>Generación de la Población de Referencia</i>	163
5.2	Referencias.....	165

Página en Blanco

5 Descripción del Caso de Estudio

En este capítulo se describe en detalle el sistema utilizado como caso de estudio y las ecuaciones del modelo matemático que lo representan. Se incluyen también los detalles de las simulaciones realizadas para definir la población usada como referencia (*PR*).

5.1 Reactor de Polimerización de Metacrilato de Metilo (*MMA*) en Emulsión

Los procesos de polimerización en emulsión tienen gran importancia a nivel industrial. Las características particulares de los mismos permiten la producción de algunos tipos de polímeros cuya obtención resulta muy complicada, cuando no imposible, por otros métodos de polimerización (Rawlings, 1985). Mediante este proceso es posible sintetizar polímeros aptos para un amplio rango de aplicaciones finales, por ejemplo: pinturas, gomas, adhesivos y hasta productos utilizados en química farmacéutica.

El proceso en emulsión es un sistema disperso donde coexisten varias fases. La naturaleza heterogénea de la mezcla reactiva tiene una influencia significativa sobre los mecanismos físicos y químicos de la reacción y sobre las propiedades del producto final (Tess y Poehlein, 1985).

Una de las ventajas más importantes del sistema en emulsión es que puede ser sometido a condiciones de operación más severas que un sistema en masa. Por ejemplo, es posible utilizar temperaturas de operación y velocidades de agitación elevadas, como así también concentraciones mayores de algunos de los reactivos. Otro aspecto a destacar es que en este proceso es factible alcanzar niveles elevados de conversión a altas velocidades de reacción. Las características mencionadas son las que permiten la obtención de un amplio rango de productos.

El comportamiento de éste o cualquier otro proceso de polimerización, desde el punto de vista del modelado matemático, es altamente no lineal (Ray y Villa, 1999) e involucra un gran número de variables. Lograr un control eficiente de la operación de estos reactores constituye un desafío importante (Bequette, 1991; Ogunnaike y Ray, 1994; Congalidis y Richards, 1998), más aún cuando la reacción se lleva a cabo en sistemas dispersos como el de emulsión.

5.1.1 Componentes de los Sistemas de Polimerización en Emulsión

En general, una emulsión o látex está formada por muchos compuestos de naturaleza química variada. Esto hace que los sistemas de polimerización en emulsión sean de carácter complejo ya que coexisten varias fases. La composición de un látex puede variar considerablemente de acuerdo a las características del polímero que se está produciendo y a los tratamientos posteriores a los que se someterán los productos de reacción. En algunos casos es necesario el agregado de *buffers* para evitar la aglomeración o de agentes de transferencia de cadena para controlar el peso molecular. Los componentes básicos de un látex, comunes a la mayoría de las recetas que se producen a nivel industrial son:

a. Agua

El agua es el componente mayoritario del látex. Su principal función es mantener la viscosidad de la mezcla reactiva en niveles relativamente bajos, además de servir como un excelente agente de transferencia de calor. Siendo la fase continua, el agua rodea al resto de las fases donde se acumulan los compuestos inmiscibles en ésta. Las moléculas de monómero son generalmente poco solubles en agua, razón por la cual, el agregado de cantidades importantes de esta materia prima llevará a que el monómero que no pueda disolverse tienda a constituir otra fase. Ésta, gracias a la agitación y el agregado de un agente dispersante está

conformada por gotas de monómero que servirán como reservorio de dicho material. La fase acuosa mantiene también dispersas a otras fases bien diferenciadas como: micelas formadas por una capa externa de moléculas de agente emulsionante en cuyo interior puede haber monómero y partículas de polímero.

b. Agentes emulsionantes

Los emulsionantes o surfactantes son compuestos tensoactivos, cuyo agregado tiene como objeto lograr una mayor estabilidad de las fases dispersas en el látex. En general, se trata de sales de sodio o potasio de ácidos orgánicos o sulfatos. Las moléculas de los surfactantes tienen dos partes bien diferenciadas: la cabeza, que contiene grupos químicos altamente hidrófilos y la cola formada por cadenas de grupos hidrófobos. La afinidad de cada uno de los extremos de las moléculas de los surfactantes hacia compuestos de polaridades diferentes hace que los mismos sean útiles para estabilizar las fases de la emulsión.

Si la concentración de surfactante en agua supera su valor de saturación (CMC), el exceso de emulsionante se separará de la fase acuosa agrupándose por el lado hidrófobo de las moléculas y dejando las cabezas hidrófilas hacia afuera. Esta conglomeración de moléculas de surfactante se conoce como micela. En su interior contiene eventualmente una cierta cantidad de monómero. Si un radical de la fase acuosa ingresa en una de estas micelas y da comienzo a las reacciones de polimerización, la micela se transforma en lo que se denomina una partícula de polímero. Esta partícula puede contener tanto radicales libres capaces de reaccionar, como polímero desactivado por la ocurrencia de las reacciones de terminación.

Si bien es posible que la polimerización ocurra en la fase acuosa, la cantidad de polímero generado en la misma es despreciable frente al formado en el interior de las partículas, que constituyen el lugar preferencial donde ocurren las reacciones de polimerización.

c. Monómero

La polimerización en emulsión se utiliza para obtener tanto homopolímeros como copolímeros. Los primeros están formados por un único monómero mientras que los segundos están constituidos por la unión química de dos o más monómeros diferentes. Las características de cada uno de los monómeros utilizados conferirán propiedades particulares tanto a la mezcla reaccionante como al producto final. Características tales como la capacidad de incorporar pigmentos y cargas, la facilidad para moldear el polímero y la estabilidad de las mezclas depende en gran medida de la naturaleza de los monómeros utilizados (Melchior y *col.*, 2000; Yang y *col.*, 2001).

Como se mencionó anteriormente, la solubilidad de los monómeros en agua es generalmente baja. Así, dado que la cantidad de monómero agregado al reactor generalmente supera la cantidad máxima que se disuelve en agua, el exceso de monómero se separa formando principalmente gotas de monómero que sirven como reservorio de material. El tamaño de estas gotas disminuye paulatinamente durante el transcurso de la reacción debido a la transferencia de masa entre las mismas y la fase acuosa. Además, parte del monómero se localiza dentro de las micelas, donde, a medida que se consume por reacción química, se repone como consecuencia de la transferencia de masa desde la fase acuosa.

d. Iniciador

Como su nombre lo indica, el iniciador es el compuesto responsable de comenzar las reacciones de polimerización. Sus moléculas se descomponen en la fase acuosa formando radicales muy reactivos. Estos radicales adicionan moléculas de monómero, transfiriéndole así su reactividad y permitiendo la formación de las cadenas de polímero, mediante reacciones con otras moléculas de monómero o polímero inactivo.

Se utilizan distintos tipos de iniciadores. Entre los más importantes se pueden mencionar a los que se descomponen por acción de la temperatura como el persulfato de potasio (*KPS*) o los del tipo redox, como el sistema peróxido de hidrógeno - ion ferroso, que se utilizan para llevar a cabo polimerizaciones a temperaturas relativamente bajas (O dian, 1991).

La Figura 5.1 muestra en forma esquemática la conformación de un sistema de polimerización en emulsión. En ésta pueden observarse todas las fases mencionadas, como así también los flujos de materia que pueden ocurrir desde y hacia cada una de ellas.

5.1.2 *Modelo Matemático*

Esta sección presenta una descripción, con mayor grado de detalle, de los fenómenos que tienen lugar en un reactor de polimerización en emulsión, así como las ecuaciones utilizadas para representarlos. El modelo presentado es una descripción rigurosa para el caso de un proceso de homopolimerización y está basado en hipótesis y ecuaciones propuestas en la literatura (Rawlings, 1985).

5.1.2.1 *Mecanismo de Reacción*

En un sistema de polimerización en emulsión en el que se utiliza un iniciador que se descompone por acción de la temperatura, se producen reacciones por radicales libres. Cuatro de ellas, consideradas fundamentales, se describen a continuación.

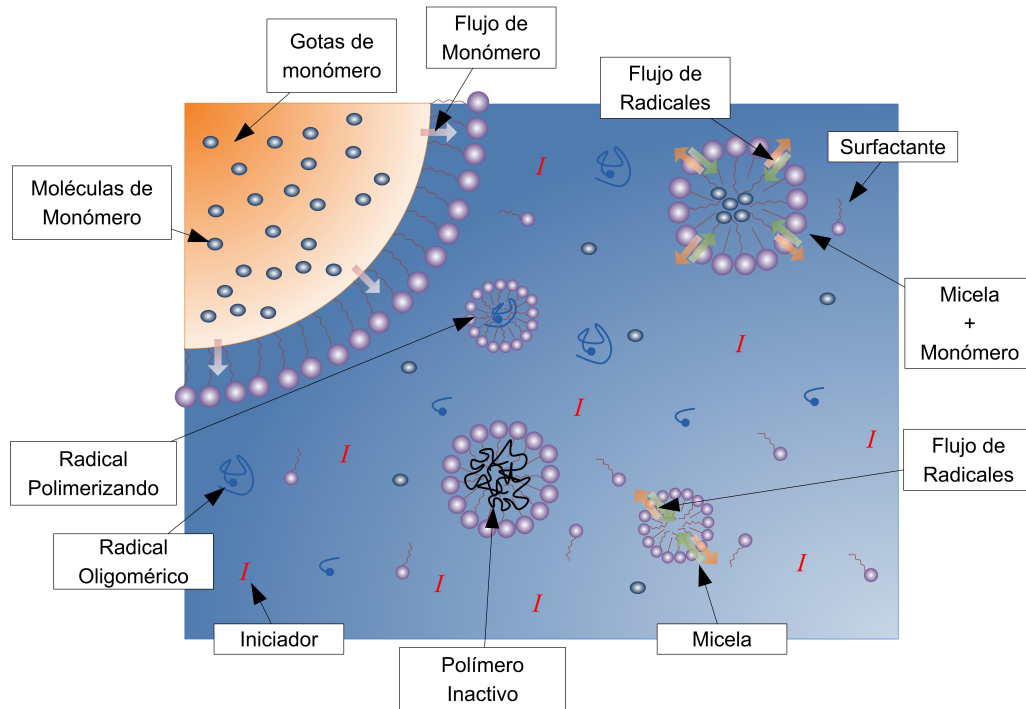
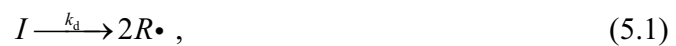


Figura 5.1: Esquema de un sistema de polimerización en emulsión

a. Iniciación

Para comenzar la reacción se utiliza un iniciador (I) soluble en agua, como el KPS . Éste se descompone por efecto de la temperatura dando dos moles de radicales ($R\bullet$) por mol de iniciador (ecuación (5.1)). Estos radicales pueden iniciar reacciones de propagación produciendo macromoléculas o, reaccionar con impurezas (J) del medio, dando como resultado especies X inactivas o inertes (ecuación (5.2)). Ambas reacciones se producen en fase acuosa.



b. Propagación

La propagación ocurre cuando los radicales generados en la etapa de iniciación reaccionan con el monómero circundante dando lugar a radicales de longitud l (R_l) o, cuando un radical de longitud cualquiera n reacciona con una molécula de monómero dando lugar a la creación de un nuevo radical de longitud $n+1$ (R_{n+1}). En ciertos sistemas de polimerización, las moléculas de polímero inactivo pueden presentar dobles enlaces terminales capaces de reaccionar y actuar como macromonómeros en una reacción de propagación. En este caso se puede producir la reacción de un radical de longitud n con un polímero inactivo de longitud m (P_m), produciendo un radical de longitud $m+n$ (R_{m+n}).

Frecuentemente se realiza la aproximación de cadena larga en este tipo de sistema, la cual permite considerar que la constante de propagación no depende de la longitud de las cadenas reaccionantes, y toma un único valor para cada temperatura. Además, es usual asumir el mismo valor en todas las fases reactivas (fase acuosa y fase polímero).



c. Transferencia de Cadena

Las reacciones de transferencia de cadena, son reacciones por las cuales un radical vivo de longitud n transfiere su actividad a una molécula de monómero (M), a algún agente de transferencia específico (Tr) o a una molécula de polímero inactivo (P_m).



d. Terminación

La terminación es el proceso mediante el cual dos radicales vivos reaccionan formando moléculas de polímero inactivo. Las reacciones de terminación pueden ser por combinación o desproporción. En el primer caso, los dos radicales se combinan para formar una molécula de polímero inactivo con una longitud de cadena igual a la suma de las de los radicales que reaccionan. En el caso de la terminación por desproporción, cada uno de éstos da lugar a una molécula de polímero inactivo cuyo peso molecular es el del radical respectivo. En una de ellas se forma un doble enlace terminal.



Estas reacciones pueden ocurrir tanto en la fase acuosa como en el interior de las partículas de polímero. En este caso también se considera que las constantes cinéticas correspondientes sólo dependen de la temperatura y su valor es el mismo para todas las fases.

Las ecuaciones anteriores incluyen a todas las reacciones consideradas en el modelo. Los balances de masa para cada una de las fases presentes en el reactor se presentan en la Sección 5.1.2.3.

5.1.2.2 Descripción del Reactor

La Figura 5.2 muestra un esquema del reactor y la instrumentación considerada para este caso de estudio. El sistema consiste en un reactor *semi-batch* no isotérmico refrigerado por agua que circula por una camisa. El modelo contempla posibles agregados de surfactante, iniciador y monómero durante el transcurso de la operación. Sin embargo, para el caso de estudio seleccionado sólo se considerará la alimentación continua de monómero fresco. El volumen de carga al reactor es de alrededor de 700 litros y se consideró que su diámetro es aproximadamente igual a la altura del líquido contenido durante la operación. El reactor cuenta con un lazo de control de temperatura (PI) cuyo objetivo es mantener a la misma aproximadamente constante, en general a 50 °C.

La operación del reactor puede dividirse en cinco etapas. En la primera, se introducen al reactor el agua, el surfactante y una dada cantidad de monómero. La segunda etapa, de calentamiento, tiene como objetivo llevar dicha mezcla hasta una temperatura adecuada para la reacción. Una vez que se alcanza la temperatura deseada, se agrega la totalidad del iniciador, se comienza la alimentación de monómero fresco y se conecta el lazo de control de temperatura. La cuarta etapa consiste en la reacción propiamente dicha, la que se efectúa por un lapso de 80 minutos. La última etapa de la operación consiste en la evacuación, limpieza y preparación del reactor para un nuevo ciclo. En este trabajo de tesis el foco de atención se centra sólo en la cuarta etapa, es decir, el paso de reacción en sí mismo.

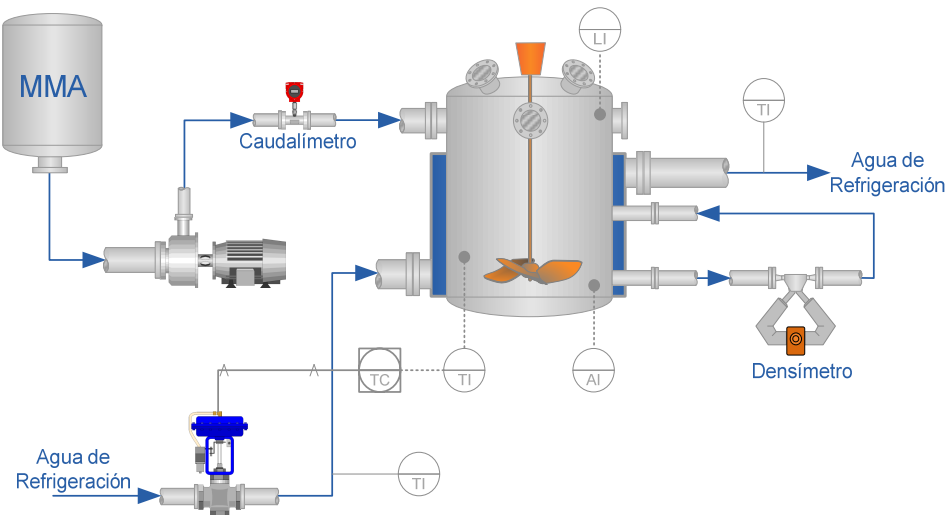


Figura 5.2: Esquema del reactor

5.1.2.3 Balances de Masa y Energía

A continuación se detallan los balances de masa para cada una de las especies contenidas en el reactor así como, las ecuaciones de temperatura resultantes de los balances de energía realizados en la camisa y en el reactor. La nomenclatura utilizada y los valores de los parámetros del modelo se encuentran detallados en el Apéndice 5.

a. Iniciador

El balance de masa del iniciador (I) para todo el volumen de control es:

$$\frac{d(I V_R V_w)}{dt} = -k_d I V_R V_w + q_I^0, \quad (5.11)$$

donde V_R es el volumen total de la mezcla reactiva, V_w es la fracción volumétrica de agua en el látex, k_d es la constante de descomposición térmica del iniciador y q_I^0 es el caudal molar de entrada de iniciador.

b. Radicales

El presente modelo considera la existencia de radicales tanto en la fase acuosa como en el interior de las partículas. Como se señaló anteriormente, en estas últimas es donde las reacciones de polimerización toman mayor relevancia. Las ecuaciones que describen la transferencia de radicales entre la fase acuosa (donde se generan los radicales de cadena corta) y las partículas de polímero resultan entonces de gran importancia para el modelo.

La velocidad de entrada de los radicales (e) en las partículas de polímero puede escribirse como:

$$e = k_{mp} 4\pi r^\eta R, \quad (5.12)$$

donde k_{mp} es la constante de velocidad de captura de radicales por parte de las partículas, r es el radio de la partícula considerada, R es la concentración total de radicales en la fase acuosa y finalmente η es un número que está relacionado a la naturaleza de la transferencia considerada. Si se asume que la transferencia se debe a un fenómeno de colisión $\eta = 2$, por el contrario, si se considera que el ingreso de los radicales es un fenómeno netamente difusivo entonces $\eta = 1$.

Para calcular la velocidad de desorción (d) de los radicales se asume que los de pequeña longitud, originados en algunas de las reacciones de transferencia de cadena, pueden migrar a la fase acuosa. La ecuación (5.13) presenta una de las expresiones más utilizadas para estimar dicha velocidad. Se tiene en cuenta: la transferencia de masa de moléculas de bajo peso molecular, la difusión a través de la partícula como paso controlante de este fenómeno, y la resistencia a la transferencia de masa en la fase acuosa (Rawlings, 1985).

$$d = \frac{3D_m \frac{k_{tr}}{k_p}}{3D_m \frac{MW_m}{\rho_m k_p \phi} + r^2}, \quad (5.13)$$

donde D_m es el coeficiente de difusividad efectiva, k_{tr} es la constante de transferencia efectiva, k_p es la constante de propagación, MW_m es el peso molecular del monómero, ρ_m es la densidad del monómero y ϕ es la fracción volumétrica de monómero en el interior de las partículas.

Las velocidades de absorción de los radicales por parte de las partículas de polímero y de desorción desde las mismas, entre otros factores, son función del tamaño de las mismas, como se observa en las ecuaciones (5.12) y (5.13). El tamaño de cada partícula de polímero depende fuertemente del tiempo transcurrido desde el nacimiento (nucleación) de la misma. Dado que este fenómeno de nucleación ocurre generalmente a lo largo de toda la operación, y que en las partículas puede o no haber un radical activo en un dado tiempo t , resulta difícil cuantificar la concentración de radicales en el interior de las mismas. Una solución posible para este problema fue propuesta por O'Toole (1965), quien utilizó el número promedio de radicales por partícula $\bar{i}(t',t)$ para cada instante de tiempo, teniendo en cuenta la historia de la misma, desde el momento $t = t'$ de su nacimiento.

$$\bar{i}(t',t) = \frac{\sum_{i=0}^{\infty} i F_i(t',t)}{F(t',t)} = \frac{a}{4} \frac{I_b(a)}{I_{b-1}(a)}, \quad (5.14)$$

$$a = \sqrt{V N_A \frac{e}{k_t}}, \quad (5.15)$$

$$b = 2V N_A \frac{d}{k_t}, \quad (5.16)$$

donde $I_b(a)$ es la función de Bessel modificada del primer tipo de orden b cuyo argumento es a , $F(t',t)$ son los moles de partículas nacidas en t' por unidad de tiempo y unidad de volumen del reactor debido a los fenómenos de nucleación micelar y homogénea, k_t es la constante de velocidad global de terminación, N_A es el número de Avogadro y V es el volumen total de las partículas por unidad de volumen del látex.

Las partículas de polímero pueden formarse de dos maneras diferentes. La primera, conocida como nucleación homogénea, ocurre cuando un oligómero en la fase acuosa reacciona con el monómero disuelto y alcanza una longitud de cadena determinada, conocida como longitud de cadena crítica (j_{cr}), a la cual el polímero se hace insoluble y se separa de la fase acuosa. Por otro lado, el mecanismo preferencial de formación de nuevas partículas se conoce como nucleación micelar o heterogénea y ocurre cuando un radical de la fase acuosa ingresa en una micela. La velocidad de entrada de los radicales en las micelas (P_m) puede calcularse como:

$$P_m = k_{mm} 4\pi r_{mic}^\eta R N_A m, \quad (5.17)$$

aquí, k_{mm} es la constante de velocidad de entrada de los radicales en las micelas, m es la concentración de micelas en el reactor y r_{mic} es el radio promedio de las micelas. El exponente η toma el valor de 2 o 1, según se adopte la teoría de la colisión o la difusiva.

El balance de radicales en la fase acuosa (R) puede escribirse entonces como:

$$\begin{aligned} \frac{d(R V_R V_W)}{dt} = & 2f k_d I V_R V_W - k_j J (R V_W V_R) \\ & - 4\pi k_{mp} N_A R \int_0^t F(t', t) V_R r^n dt' - P_m V_W V_R - k_{tw} R^2 V_W V_R, \quad (5.18) \\ & + \int_0^t \frac{3D_m \frac{k_{tr}}{k_p}}{3D_m \frac{MW_m}{\rho_m k_p \phi} + r^2} F(t', t) V_R \bar{i}(t', t) dt' \end{aligned}$$

donde f es un factor de eficiencia para la descomposición del iniciador, k_{tw} es la constante de velocidad de terminación en fase acuosa, J es la concentración de un agente inhibidor y k_j es la constante cinética correspondiente a la reacción entre éste y los radicales. Los términos del segundo miembro de la ecuación (5.18) representan, siguiendo el orden en que aparecen, velocidades de formación o desaparición de radicales por: descomposición del iniciador, reacción con inhibidores, absorción en las partículas y en las micelas, terminación en la fase acuosa y desorción desde las partículas.

c. Monómero

El balance global de monómero en el reactor puede escribirse como:

$$\frac{d(M V_R)}{dt} = -\frac{\rho_m}{MW_m} \int_0^t F(t', t) V_R k_p \bar{i}(t', t) \phi dt' + q_m^0, \quad (5.19)$$

donde M es la concentración de monómero en el reactor y q_m^0 es el caudal molar de alimentación de monómero. El monómero contenido en el reactor se encuentra distribuido en todas las fases, y su concentración en cada una de ellas dependerá de los fenómenos de transferencia de masa entre las mismas. La concentración de monómero en el interior de las partículas y su concentración en la fase acuosa se consideran generalmente en equilibrio

(Rawlings y Ray, 1988; Gao y Penlidis, 2002). La ecuación (5.20) expresa la relación entre las mismas:

$$1 - \varphi + \ln(\varphi) + \psi(1 - \varphi)^2 = \ln\left(\frac{M_w}{M_{\text{sat}}}\right), \quad (5.20)$$

donde φ es la fracción volumétrica de monómero en las partículas de polímero (que se considera igual para todas las partículas), ψ es el parámetro de interacción de Flory-Huggins, M_w es la concentración de monómero en la fase acuosa y M_{sat} el valor de saturación de M_w .

Si hay gotas de monómero presentes en el reactor ($M_w = M_{\text{sat}}$), φ se obtiene directamente de la ecuación (5.20). En caso contrario, la misma debe resolverse en forma simultánea con la ecuación (5.21).

$$M V_R = M_w V_w V_R + \frac{\rho_m}{MW_m} \int_0^t F(t', t) V_R V_w \varphi dt' + M_d V_R, \quad (5.21)$$

donde M_d es la concentración de monómero en las gotas.

El primer miembro de la ecuación (5.21) es la cantidad total de monómero en el interior del reactor en el tiempo t . El primer sumando del miembro derecho representa la cantidad de monómero disuelto en la fase acuosa y el segundo considera el monómero contenido en el interior de todas las partículas de polímero. Finalmente, el último término ($M_d V_R$) representa los moles de monómero en el interior de las gotas. Siempre que haya presente monómero en forma de gotas, la ecuación (5.21) puede escribirse como:

$$M V_R = M_{\text{sat}} V_w V_R + \frac{\rho_m \varphi}{MW_m} \int_0^t F(t', t) V_R V_w dt' + M_d V_R, \quad (5.22)$$

es decir, la fase acuosa puede considerarse saturada de monómero ($M_W=M_{sat}$) y φ (ecuación (5.20)) tendrá un valor constante. Las ecuaciones (5.21) y (5.22) no tienen en cuenta al monómero contenido en las micelas, cuya cantidad es despreciable frente al que se encuentra en las otras fases.

La conversión del sistema se determina generalmente fuera de línea utilizando métodos gravimétricos. La ecuación (5.23) propone una forma de cálculo para la conversión congruente con los resultados de dichas técnicas, mediante la evaluación de la proporción de monómero que forma parte del polímero, respecto del monómero total en el reactor.

$$x = \frac{N_A \rho_p \int_0^t F(t',t) V_W V_R (1-\varphi) dt'}{N_A \rho_p \int_0^t F(t',t) V_W V_R (1-\varphi) dt' + M V_R M W_m} \quad (5.23)$$

El numerador de la ecuación (5.23) representa la masa de polímero en el reactor mientras que el denominador es la suma de las masas de polímero y monómero presentes en mismo.

d. Velocidad de Crecimiento de las Partículas

La velocidad de crecimiento de las partículas de polímero (r_V) depende de la diferencia de densidades entre el monómero y el polímero y también de la velocidad de reacción en el interior de las partículas:

$$r_V = \frac{\rho_m}{\rho_p} \frac{k_p}{N_A} \frac{\varphi}{(1-\varphi)} \bar{i}(t',t) + \frac{V}{(1-\varphi)} \frac{d\varphi}{dt} \quad (5.24)$$

El primer término del segundo miembro de la ecuación (5.24) tiene en cuenta la variación del volumen debido al crecimiento de las macromoléculas por la reacción de propagación. El segundo contempla la variación del volumen debido al cambio de la fracción volumétrica del

monómero en la fase polímero, que sólo ocurre hacia el final de la reacción cuando no hay gotas de monómero presentes.

e. Balance de Masa del Surfactante

El balance global de surfactante en el reactor se muestra en la ecuación (5.25).

$$\frac{d(SV_w V_R)}{dt} = q_s^0. \quad (5.25)$$

En la expresión anterior se asume que los cambios en la masa de surfactante se deben sólo al flujo de entrada del mismo. La existencia de micelas dependerá de la cantidad de surfactante (S_w) que no esté siendo utilizada para estabilizar las gotas de monómero y/o las partículas de polímeros y puede estimarse a partir de la ecuación (5.26).

$$SV_w V_R = \frac{4\pi}{a_{ep}} \int_0^t F(t',t) V_R r^2 dt' + \frac{3MW_m V_R M_d}{a_{ed} \rho_m N_A r_d} + S_w V_w V_R. \quad (5.26)$$

El primer miembro de la expresión anterior corresponde a la cantidad total de surfactante presente en el reactor. El primer sumando del segundo miembro representa la cantidad de surfactante que cubre el área superficial de las partículas y el segundo a la que estabiliza a las gotas de monómero. Finalmente el tercer sumando representa el resto del surfactante, que se encontrará en la fase acuosa y también formando micelas en el caso que S_w supere la CMC del mismo (S_{WC}).

La concentración de las micelas presentes en el medio de reacción (m) se calcula por medio de las ecuaciones (5.27) a (5.29).

$$m' = (S_w - S_{WC})(a_{em} / 4\pi r_{mic}^2), \quad (5.27)$$

$$m = m' H(m), \quad (5.28)$$

$$H(m) = \begin{cases} 1, & m' \geq 0 \\ 0, & m' < 0, \end{cases} \quad (5.29)$$

$$a_{ep} = a_{ed} = a_{em}$$

donde a_{ep} , a_{ed} y a_{em} representan las áreas superficiales cubiertas por cada molécula de surfactante sobre las partículas de polímero, las gotas de monómero y las micelas, respectivamente, que por simplicidad se asumen iguales.

f. Volumen del Reactor

Las variaciones del volumen total del látex y de la fase acuosa se expresan así:

$$\frac{dV_R}{dt} = -\left(1 - \frac{\rho_m}{\rho_p}\right) \int_0^t F(t', t) V_R k_p \bar{i}(t', t) \phi dt' + q_w^0 \frac{MW_w}{\rho_w} + q_m^0 \frac{MW_m}{\rho_m} + q_s^0 \frac{MW_s}{\rho_s}, \quad (5.30)$$

$$\frac{d(V_R V_w)}{dt} = q_w^0 \frac{MW_w}{\rho_w}, \quad (5.31)$$

donde q_w^0 y q_s^0 son los caudales molares de alimentación, MW_w y MW_s son los pesos moleculares del agua y del surfactante, respectivamente, mientras que ρ_w y ρ_s son los valores correspondientes de densidad.

g. Efecto Gel

Contrariamente a lo que podría esperarse, es común que, a conversiones elevadas, se produzca un aumento relativo en la velocidad de polimerización respecto de sus valores iniciales. Este comportamiento puede explicarse mediante lo que se conoce como efecto gel. A niveles altos de conversión, la viscosidad en el interior de las partículas de polímero se

vuelve extremadamente alta lo que produce una disminución en la movilidad de las macromoléculas de alto peso molecular. Este fenómeno tiene un efecto directo en la velocidad de terminación puesto que disminuye la probabilidad de que dos radicales se encuentren y terminen. Esto último produce un aumento aparente en la velocidad de propagación. Desde el punto de vista del modelado, este efecto puede incluirse mediante el agregado de un factor de corrección (g_t) sobre la constante de velocidad de terminación k_t . Para el caso del *MMA*, el valor de g_t puede estimarse como (Rawlings, 1985):

$$g_t = \frac{\exp\{2(b_m x_{loc} + c_m x_{loc}^2)\}}{(1 - x_{loc})^2}, \quad (5.32)$$

$$x_{loc} = \frac{1 - \phi}{1 - \phi(1 - \frac{\rho_m}{\rho_p})}, \quad (5.33)$$

donde $b_m = -7.68$ y $c_m = -1.13$.

h. Balance Poblacional

El balance poblacional que permite calcular el número de partículas nacidas a un tiempo $t = t'$ por unidad de tiempo y unidad de volumen ($F(t', t)$) se muestra en la ecuación (5.34):

$$\frac{\partial F(t', t)}{\partial t} = \delta(t - t') P_m V_R V_W + \delta(t - t') k_p M R_{j_{cr}-1} V_R V_W, \quad (5.34)$$

donde $\delta(t - t')$ es la función delta de Dirac y $R_{j_{cr}-1}$ es la concentración en fase acuosa de los radicales de longitud $j_{cr} - 1$. El primer sumando del segundo miembro representa la aparición de partículas, ocurrida a $t = t'$, debido a la nucleación micelar. Por otra parte, en el segundo sumando se considera la formación de nuevas partículas debido a la nucleación homogénea por precipitación de los radicales oligoméricos que se tornan insolubles en la fase acuosa.

i. Pesos Moleculares

Los pesos moleculares promedios en peso Mw y en número Mn se calculan usando el método de los momentos (Katz y Saidel, 1967). Esta técnica permite simplificar la resolución del sistema de ecuaciones que describe la distribución de longitudes de los radicales y del polímero presentes en el reactor. Con ese propósito se multiplica por n^a a cada término del balance de masa de la macroespecie genérica, R_n o P_n , que contiene n unidades monoméricas. A continuación se suman las ecuaciones de balance para todos los tamaños posibles ($n=1\dots\infty$). Luego se reordenan los términos y se expresan en función de las definiciones de los momentos de los radicales, $\lambda_a = \sum_{n=1}^{\infty} n^a R_n$, y del polímero, $\mu_a = \sum_{n=1}^{\infty} n^a P_n$. De esta forma se obtienen las ecuaciones (5.35) a (5.42). Los términos del segundo miembro de cada una de estas ecuaciones tienen en cuenta las contribuciones correspondientes tanto a las reacciones químicas que ocurren en las partículas como a la transferencia de masa entre las partículas y la fase acuosa. Se desprecia la contribución de las macroespecies que se encuentran en la fase acuosa.

$$\frac{\partial \lambda_0}{\partial t} = -\frac{(k_{tc} + k_{td})\lambda_0^2}{VTP} + 4\pi k_{mp} RN_A \int_0^t \delta(t-t') F(t',t) V_R r_t'^2 dt' - \int_0^t \delta(t-t') dF(t',t) V_R \bar{i} dt' \quad , \quad (5.35)$$

$$\frac{\partial \lambda_1}{\partial t} = k_p \left(\frac{\varphi \rho_m}{MW_m} \right) - \frac{(k_{tc} + k_{td})\lambda_0 \lambda_1}{VTP} - k_{tr} \left(\frac{\varphi \rho_m}{MW_m} \right) (\lambda_1 - \lambda_0) + 4\pi k_{mp} RN_A \int_0^t \delta(t-t') F(t',t) V_R r_t'^2 dt' - \int_0^t \delta(t-t') dF(t',t) V_R \bar{i} dt' \quad , \quad (5.36)$$

$$\frac{\partial \lambda_2}{\partial t} = k_p \left(\frac{\varphi \rho_m}{MW_m} \right) (2\lambda_1 + \lambda_0) - \frac{(k_{tc} + k_{td})\lambda_0 \lambda_2}{VTP} - k_{tr} \left(\frac{\varphi \rho_m}{MW_m} \right) (\lambda_2 - \lambda_0) + 4\pi k_{mp} RN_A \int_0^t \delta(t-t') F(t',t) V_R r_t'^2 dt' - \int_0^t \delta(t-t') dF(t',t) V_R \bar{i} dt' \quad , \quad (5.37)$$

$$\frac{\partial \mu_0}{\partial t} = \frac{(k_{tc}/2 + k_{td})\lambda_0^2}{VTP} + k_{tr} \left(\frac{\varphi \rho_m}{MW_m} \right) \lambda_0, \quad (5.38)$$

$$\frac{\partial \mu_1}{\partial t} = \frac{(k_{tc} + k_{td})\lambda_0 \lambda_1}{VTP} + k_{tr} \left(\frac{\varphi \rho_m}{MW_m} \right) \lambda_1, \quad (5.39)$$

$$\frac{\partial \mu_2}{\partial t} = \frac{(k_{tc} + k_{td})\lambda_0 \lambda_2}{VTP} + \frac{k_{tc}\lambda_1^2}{VTP} + k_{tr} \left(\frac{\varphi \rho_m}{MW_m} \right) \lambda_2, \quad (5.40)$$

$$Mn = MW_m \frac{\lambda_1 + \mu_1}{\lambda_0 + \mu_0}, \quad (5.41)$$

$$Mw = MW_m \frac{\lambda_2 + \mu_2}{\lambda_1 + \mu_1}, \quad (5.42)$$

donde λ_0 , λ_1 y λ_2 son los momentos de orden cero, uno y dos de la distribución de longitudes de los radicales, respectivamente, μ_0 , μ_1 y μ_2 son los correspondientes momentos para el polímero, r_t es el radio de las partículas nacidas en t' y VTP es el volumen total de las partículas de polímero, que se estima mediante la ecuación (5.43):

$$VTP = \frac{4}{3} \pi N_A \int_0^t \delta(t-t') F(t', t) V_R r_t'^3 dt', \quad (5.43)$$

j. Balance de Energía

En los balances energía utilizados en este modelo no se detallan los fenómenos de transferencia de calor entre las diferentes fases del reactor ni la resistencia propia de cada una de ellas. Sólo se considera la transferencia entre el interior del reactor y el agua de refrigeración. En consecuencia, el balance global de energía se expresa como:

$$W^{\text{látex}} c_p^{\text{látex}} \frac{\partial(T - T_{\text{ref}})}{\partial t} = q_m^0 \rho_m c_p^m (T^0 - T_{\text{ref}}) - \Delta H_r r_{\text{pol}} - Q_C + \left(\left(\frac{\partial(MV_R)}{\partial t} c_p^m + \frac{\partial(V_R V_W)}{\partial t} c_p^w + \frac{\partial(W^p)}{\partial t} c_p^p \right) (T - T_{\text{ref}}) \right), \quad (5.44)$$

donde T es la temperatura de la mezcla reactiva, T_{ref} es una temperatura tomada como referencia, T^0 es la temperatura de la corriente de monómero fresco, $c_p^{\text{látex}}$, c_p^m , c_p^w y c_p^p son las capacidades caloríficas del látex, monómero, agua y polímero, respectivamente, ΔH_r es el calor de reacción y r_{pol} la velocidad de polimerización. Finalmente Q_C , W^p y $W^{\text{látex}}$ representan el calor transferido a la camisa, la masa total de polímero y la masa total en el interior del reactor, respectivamente, y se estiman así:

$$W^{\text{látex}} c_p^{\text{látex}} = MV_R MW_m c_p^M + V_R V_W c_p^w + c_p^p \rho_p \int_0^t F(t', t) V V_R (1 - \phi) dt' \quad (5.45)$$

$$W^p = \rho_p \int_0^t F(t', t) V V_R (1 - \phi) dt'. \quad (5.46)$$

$$Q_C = UA(T - T_C) \quad (5.47)$$

El balance de energía para la camisa es:

$$W^C c_p^w \frac{\partial T_C}{\partial t} = Q_C + f_C c_p^w (T_{C0} - T_C) \quad (5.48)$$

donde W^C es la masa de agua contenida en el interior de la camisa, T_{C0} es la temperatura de ingreso del agua de refrigeración y f_C es el caudal másico de agua a través de la camisa.

5.1.3 Implementación y Validación del Modelo

El modelo se implementó en gPROMS®. Además, se utilizó la interfase gOmatlab® para comunicar MatLab® y gPROMS® con el objeto de usar el modelo como una función en el ambiente de MatLab®. Algunos de los experimentos reportados por Gao y Penlidis (2002) y Sajjadi y Yianneskis (2003) se usaron para realizar el ajuste de las constantes pre-exponenciales de las reacciones de propagación, terminación y transferencia: A_{k_p} , A_{k_t} , y $A_{k_{tr}}$. Para ello se emplearon las rutinas de estimación de parámetros gEST, incluidas en gPROMS®.

El modelo describe en forma apropiada el comportamiento de las variables de interés en el transcurso del *batch*. A título de ejemplo, en las Figuras 5.3 a 5.5 se muestran algunos de los resultados calculados para las trayectorias de conversión y peso molecular promedio en peso, comparadas con las obtenidas experimentalmente.

5.1.4 Generación de la Población de Referencia

Se realizaron 146 simulaciones en las que se variaron las condiciones iniciales y los valores de las variables externas al reactor a fin de conseguir un conjunto de resultados apropiados para definir la población de referencia. Los datos de entrada del modelo, para cada una de dichas simulaciones, se muestran en el Apéndice 5. Por otra parte se utilizó el modelo para simular un conjunto de 82 fallas de proceso que fueron utilizadas para evaluar de manera comparativa el desempeño de las distintas estrategias de monitoreo analizadas. La descripción de dichas fallas se presenta en el Capítulo 6.

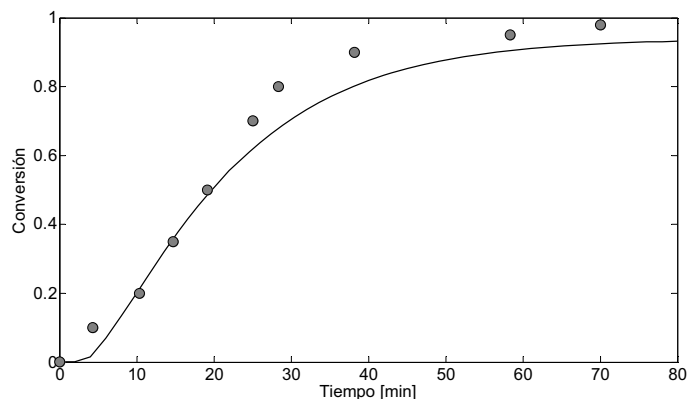


Figura 5.3: Conversión vs. tiempo. Modelo (—), datos experimentales (o) (Sajjadi y Yianneskis, 2003). Condiciones iniciales: $V_W=0,7$, $M = 1,96$ mol/l, $S = 0,0173$ mol /l de agua y $I = 0,0046$ mol/l

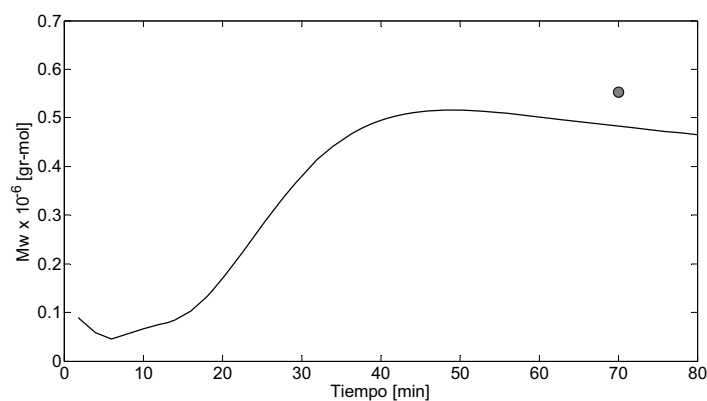


Figura 5.4: Peso molecular promedio en peso vs. . Modelo (—), datos experimentales (o) (Sajjadi y Yianneskis, 2003). Condiciones iniciales: $V_W=0,7$, $M = 1,96$ mol/l, $S = 0,0173$ mol /l de agua y $I = 0,0046$ mol/l

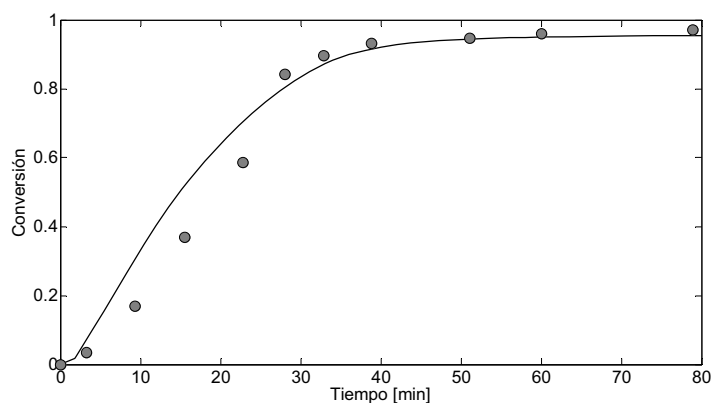


Figura 5.5: Conversión vs. tiempo. Modelo (—), datos experimentales (o) (Gao y Penlidis 2002). Condiciones iniciales: 30 g de MMA, 70 g de agua, $S = 0,02$ mol/l de agua y $I = 0,02$ mol/l

5.2 Referencias

- Bequette, B. W., "Nonlinear Control of Chemical Processes: A Review". *Industrial and Engineering Chemistry Research* **1991**, 30, (7), 1391-1413.
- Congalidis, J. P. y J. R. Richards, "Process Control of Polymerization Reactors: An Industrial Perspective". *Polymer Reaction Engineering* **1998**, 6, (2), 71-111.
- Gao, J. y A. Penlidis, "Mathematical Modeling and Computer Simulator/Database for Emulsion Polymerizations". *Progress in Polymer Science (Oxford)* **2002**, 27, (3), 403-535.
- Katz, S. y G. M. Saidel, "Moments of the Size Distribution in Radical Polymerization". *AIChE Journal* **1967**, 13, (2), 3129-326.
- Melchior, M., M. Sonntag, C. Kobusch y E. Jürgens, "Recent Developments in Aqueous Two-Component Polyurethane (2K-PUR) Coatings". *Progress in Organic Coatings* **2000**, 40, (1-4), 99-109.
- O'Toole, J. T., "Kinetics of Emulsion Polymerization". *Journal of Applied Polymer Science* **1965**, 9, (4), 1291- 1297.
- Odian, G. (1991). Principles of Polymerization, New York, John Wiley & Sons, Inc.
- Ogunnaike, B. A. y W. H. Ray (1994). Process Dynamics, Modeling, and Control, New York, Oxford University Press.
- Rawlings, J. B. (1985). Simulation and Stability of Continuous Emulsion Polymerization Reactors. Ph.D. Thesis, Chemical Engineering Department, Wisconsin, University of Wisconsin.
- Rawlings, J. B. y W. H. Ray, "Modeling of Batch and Continuous Emulsion Polymerization Reactors. Part I: Model Formulation and Sensitivity to Parameters". *Polymer Engineering and Science* **1988**, 28, (5), 237-256.
- Ray, W. H. y C. M. Villa, "Nonlinear Dynamics Found in Polymerization Processes: A Review". *Chemical Engineering and Science* **1999**, 55, (2), 275-290.
- Sajjadi, S. y M. Yianneskis, "Semibatch Emulsion Polymerization of Methyl Methacrylate with a Neat Monomer Feed". *Polymer Reaction Engineering* **2003**, 11, (4 SPEC. ISS.), 715-736.
- Tess, R. W. y G. W. Poehlein, Eds. (1985), Applied Polymer Science, ACS Symposium Series 285, Washington DC, American Chemical Society.
- Yang, C. H., H. J. Yang, T. C. Wen, M. S. Wu y J. S. Chang, "Mixture Design Approaches to IPDI-H6XDI-XDI Ternary Diisocyanate-Based Waterborne Polyurethanes". *Polymer* **1999**, 40, (4), 871-885.

Página en Blanco

Capítulo 6

Resultados y Discusión

6.1	Generación de Datos	169
6.1.1	<i>Cambios en el caudal de alimentación de monómero.....</i>	171
6.1.2	<i>Cambios en la temperatura de entrada del refrigerante.....</i>	172
6.1.3	<i>Variaciones en las condiciones iniciales</i>	174
6.2	Base de Comparación entre las Distintas Técnicas.....	174
6.3	Comparaciones entre <i>NM, WKFH-AT, OSS, VM y ME</i>	177
6.4	Resultados del Análisis de Componentes Independientes (<i>MICA</i>)	199
6.5	Referencias.....	208

Página en Blanco

6 Resultados y Discusión

Las técnicas de *MSPC* para procesos *batch* introducidas en los Capítulos 3 y 4 se aplicaron para monitorear la operación del reactor de polimerización de *MMA* descrito en el capítulo anterior. A continuación, se presentan y discuten los resultados del análisis del desempeño de estas estrategias para la detección e identificación de fallas.

6.1 Generación de Datos

El modelo del reactor de polimerización descrito en el Capítulo 5 se utilizó para realizar un conjunto de simulaciones en diferentes condiciones operativas, con el objetivo de emular un registro histórico de datos de operación. Este último fue empleado para definir las condiciones de operación normal (*NOC*) del proceso.

En una primera etapa se simularon 146 lotes. El pretratamiento de esta base de datos, mediante los procedimientos explicados en el Capítulo 2, permitió seleccionar 121 *batches* para conformar la *NOC* ($I = 121$). Se usaron ocho variables como descriptores del proceso ($J = 8$): caudal de alimentación de monómero (Q_0), temperatura de entrada del agua de camisa (T_{C0}), temperatura de salida del agua de camisa (T_C), temperatura de la mezcla reactiva (T), concentración de surfactante en la fase acuosa (S_W), nivel del reactor (h_R), densidad de la mezcla reactiva (ρ_R) y peso molecular promedio en peso (MW).

La duración de todos los *batches* es de 80 minutos. Las observaciones del proceso se toman cada dos minutos desde el inicio de la operación, dando un total de 41 observaciones por cada *batch* ($K = 41$). Las trayectorias promedio de los descriptores se muestran en la Figura 6.1.

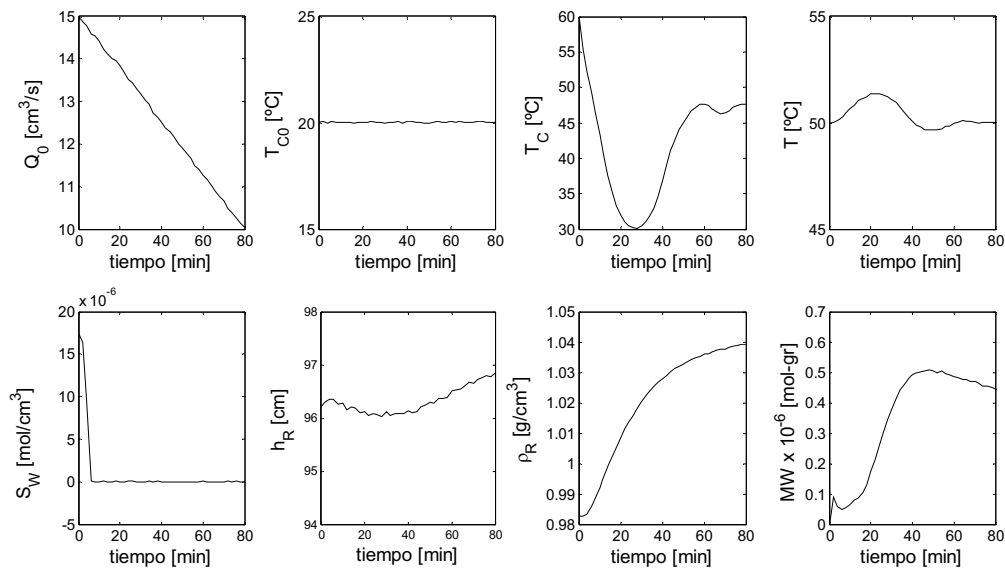


Figura 6.1: Trayectoria promedio de las J variables monitoreadas

Cada uno de los *batches* de la población de referencia (*PR*) se afectó artificialmente con ruido aleatorio con el objeto de emular el proveniente de un proceso real. La magnitud del ruido asociado a cada variable se determinó teniendo en cuenta las características de los instrumentos de medición disponibles en el mercado (Liptak, 2003). Los niveles de ruido en cada una de las variables se indican en el Apéndice 5. La Figura 6.2 muestra los valores de las J variables para los *batches* considerados en la *NOC*.

Se generó además un segundo conjunto de 82 *batches* en los cuales se simularon fallas conocidas. Estos últimos se analizaron mediante algunas de las técnicas descriptas en los Capítulos 3 y 4, con el objeto de evaluar el desempeño de las mismas en forma comparativa.

Se simularon fallas de distinta naturaleza. Para cada tipo de falla se contemplaron diferentes magnitudes de la desviación y diversos tiempos de ocurrencia. Las mismas están identificadas mediante una numeración creciente desde 1 hasta 82 y pueden agruparse de la siguiente manera.

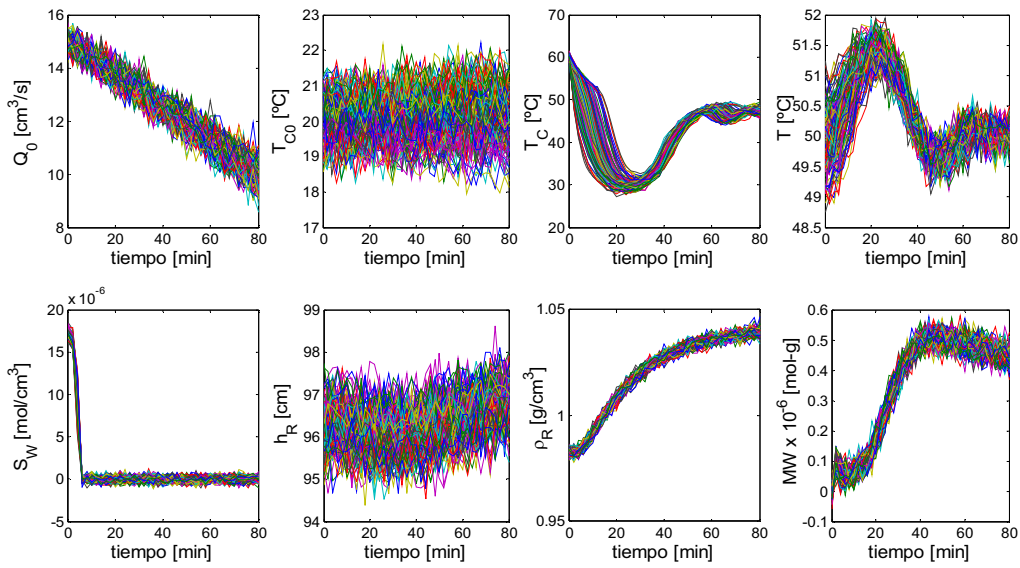


Figura 6.2: Trayectoria temporal de las J variables monitoreadas para los I batches de la NOC

6.1.1 Cambios en el caudal de alimentación de monómero

Se realizaron 23 simulaciones con variaciones en el caudal de alimentación de MMA fresco. Las primeras cinco fallas consideran una alimentación constante de monómero a diferentes caudales. La Tabla 6.1 muestra el valor del caudal para cada una de estas fallas.

Tabla 6.1: Valores de Q_0 para las fallas 1 a 5

Falla	Q_0 [cm^3/s]
1	15
2	12
3	10
4	17
5	20

El cambio en Q_0 respecto de su valor promedio se simula desde el principio de la operación y se extiende durante todo su transcurso. Se consideraron tanto valores más altos como más bajos que el valor inicial del caudal promedio.

Las fallas 6 a 23 consideran también variaciones en Q_0 . Sin embargo, en estos casos se simularon cambios en forma de escalón de diferente magnitud y que se ubican en distintos tiempos durante la operación. La Tabla 6.2 muestra la magnitud y duración de los distintos escalones considerados. La segunda columna indica el porcentaje de variación respecto del valor inicial del caudal promedio de la *NOC*. La tercera y cuarta señalan el tiempo y la observación donde la falla se pone de manifiesto por primera vez. Finalmente, la quinta y sexta columna indican el tiempo y la observación donde la perturbación finaliza.

Tabla 6.2: Valores de los escalones aplicados a Q_0

Falla	%	t_{inicio}	k_{inicio}	t_{final}	k_{final}
6	20	12	7	40	21
7	20	32	17	60	31
8	20	52	27	80	41
9	-20	12	7	40	21
10	-20	32	17	60	31
11	-20	52	27	80	41
12	10	12	7	40	21
13	10	32	17	60	31
14	10	52	27	80	41
15	-10	12	7	40	21
16	-10	32	17	60	31
17	-10	52	27	80	41
18	5	12	7	40	21
19	5	32	17	60	31
20	5	52	27	80	41
21	-5	12	7	40	21
22	-5	32	17	60	31
23	-5	52	27	80	41

6.1.2 *Cambios en la temperatura de entrada del refrigerante*

De manera similar al caso anterior, se realizaron 23 simulaciones en las que se emularon perturbaciones en la temperatura de alimentación del agua de camisa (T_{C0}). Las primeras

cinco fallas corresponden a valores fijos de T_{C0} durante todo el transcurso de la operación. La Tabla 6.3 presenta los valores de T_{C0} considerados en las fallas 24 a 28.

Tabla 6.3: Valores de T_{C0} para las fallas 24 a 28

Falla	T_{C0} [°C]
24	15
25	18
26	20
27	22
28	25

Las fallas 29 a 46 se simularon como cambios en forma de rampas. Cada uno de los perfiles de T_{C0} está representado en uno, dos o tres tramos en los que la misma varía linealmente en el tiempo. Los valores de las pendientes en cada tramo (m_i [°C/min]) y los intervalos de tiempo y observaciones que los definen (τ_i [min], k_i) se detallan en la Tabla 6.4.

Tabla 6.4: Valores y posiciones de las pendientes en las rampas de T_{C0} para las fallas 29 a 46

Falla	m_1	τ_1 [min]	k_1	m_2	τ_2 [min]	k_2	m_3	τ_3 [min]	k_3
29	0,04	0-10	1-6	0,14	10-40	6-21	0,02	40-80	21-41
30	0,04	0-30	1-16	0,14	30-60	16-31	0,02	60-80	31-41
31	0,04	0-50	1-26	0,14	50-80	26-41	-	-	-
32	-0,04	0-10	1-6	-0,14	10-40	6-21	-0,02	40-80	21-41
33	-0,04	0-30	1-16	-0,14	30-60	16-31	-0,02	60-80	31-41
34	-0,04	0-50	1-26	-0,14	50-80	26-41	-	-	-
35	0,14	0-40	1-21	0,2	40-60	21-31	0,02	60-80	31-41
36	0,14	0-60	1-31	0,2	60-80	31-41	0,02	-	-
37	0,14	0-80	1-41	-	-	-	-	-	-
38	-0,14	0-40	1-21	-0,2	40-60	21-31	-0,02	60-80	31-41
39	-0,14	0-60	1-31	-0,2	60-80	31-41	-	-	-
40	-0,14	0-80	1-41	-	-	-	-	-	-
41	0,2	0-20	1-11	0,04	20-30	11-16	0,14	30-80	16-41
42	0,2	0-20	1-11	0,04	20-40	11-21	0,14	40-80	21-41
43	0,2	0-20	1-11	0,04	20-50	11-26	0,14	50-80	26-41
44	-0,2	0-20	1-11	-0,04	20-30	11-16	-0,14	30-80	16-41
45	-0,2	0-20	1-11	-0,04	20-40	11-21	-0,14	40-80	21-41
46	-0,2	0-20	1-11	-0,04	20-50	11-26	-0,14	50-80	26-41

6.1.3 Variaciones en las condiciones iniciales

Los sistemas de polimerización en emulsión son muy sensibles a los cambios en las condiciones de operación, especialmente durante la primera etapa de la reacción. De este modo, pequeñas variaciones en las condiciones iniciales pueden tener un efecto muy marcado en el desarrollo del *batch*. Las fallas 47 a 82 consideran la variación de las condiciones iniciales de seis variables proceso, de a una por vez. La Tabla 6.5 contiene los valores de las perturbaciones simuladas en cada una de estas 36 fallas (6 fallas para cada variable).

Tabla 6.5: Condiciones iniciales para las fallas 47 a 82

Fallas 1° - 6°	Variable	1°	2°	3°	4°	5°	6°	Obs.
47 - 52	Volumen inicial de carga al reactor	-20%	-15%	-10%	10%	15%	20%	M
53 - 58	Concentración total de surfactante	-20%	-15%	-10%	10%	15%	20%	M
59 - 64	Fracción volumétrica de agua	-20%	-15%	-10%	10%	15%	20%	NM
65 - 70	Concentración de iniciador	-20%	-15%	-10%	10%	15%	20%	NM
71 - 76	Concentración de monómero	-20%	-15%	-10%	10%	15%	20%	NM
77 - 82	Temperatura del reactor	-10%	-7%	-5%	5%	7%	10%	M

M Variable Medida (directa o indirectamente)
 NM Variable No Medida

6.2 Base de Comparación entre las Distintas Técnicas

A continuación se mencionan algunos puntos importantes a tener en cuenta para llevar a cabo de manera adecuada la comparación de las diferentes técnicas que se analizan en este trabajo.

a. Alarmas y alertas

Cada una de las observaciones de los *batches* monitoreados se utiliza para realizar un test de hipótesis que permitirá decidir sobre el estado del proceso. Si el valor de alguno de los estadísticos utilizados cae fuera de sus límites para una dada observación, entonces se produce

una señal de *alerta*. Si ocurren tres alertas sucesivas se declara una *alarma*, indicando de esta manera que el *batch* está fuera de control.

b. Niveles de significación locales y globales

El nivel de significación utilizado en cada realización de un test de hipótesis se conoce como nivel de significación local (α_L) y es la probabilidad de cometer un error *Tipo I* en dicho contraste. El nivel de significación global (α_G) cuando se realizan K tests sucesivos es la probabilidad de cometer al menos un error de *Tipo I* en alguno de los K test realizados. Existen varias formas de estimar el valor del α_L necesario para conseguir el α_G deseado.

Si se considera que las mediciones son independientes entre sí, es decir, que el hecho de tomar una observación del proceso no afecta a los valores de las siguientes, el ajuste de Bonferroni (Ramaker y Van Sprang, 2004) permite despejar α_L de la siguiente expresión:

$$\alpha_G = 1 - (1 - \alpha_L)^{1/K}. \quad (6.1)$$

La aproximación más conocida para esta corrección es:

$$\alpha_G \approx \alpha_L K. \quad (6.2)$$

De la ecuación (6.2) se desprende que para disminuir la probabilidad global de error *Tipo I*, los α_L deberán tomar valores muy conservativos.

Las particularidades de las metodologías utilizadas en las técnicas que aquí se estudian hacen que la determinación del α_L no resulte trivial. Ramaker y Van Sprang (2004) abordaron el problema de las mediciones sucesivas en el contexto del monitoreo de procesos *batch*. Sin embargo, algunos puntos críticos que dificultan la obtención del α_L no fueron

contemplados en ese estudio. El hecho de que una *alarma* se declara cuando tres observaciones sucesivas caen fuera de los límites reduce drásticamente la probabilidad de error *Tipo I*. Además, una vez que se detecta una alarma, el *batch* es declarado fuera de control haciendo innecesaria la evaluación de las observaciones futuras. Estos factores complican el tratamiento matemático para la determinación del α_L . Por ese motivo, en este trabajo se estimó el α_L utilizando el enfoque empírico que se explica en la siguiente subsección.

Dado que el número de *batches* considerados no es muy grande ($I=121$), el ajuste empírico de los valores del α_L para un $\alpha_G=0,01$ puede arrojar valores mayores del 1%. Por esta razón, en este trabajo sólo se consideraron los límites correspondientes a $\alpha_G=0,05$.

c. Número Promedio de Errores Tipo I (AVTI)

Las diferencias intrínsecas que existen entre las metodologías analizadas podrían ocasionar que las mismas presenten, de manera sistemática, valores distintos en sus tasas de falsas alarmas. Estas pueden estimarse en base a los resultados obtenidos cuando los *batches* que forman la *NOC* se monitorean como si fuesen *batches* nuevos usando para ello el enfoque *leave one out*. El número promedio de errores *Tipo I* (*AVTI*) es una medida empírica de la tasa de falsas alarmas y puede estimarse como se muestra en la ecuación (6.3).

$$AVTI = \frac{\text{número de falsas alarmas detectadas}}{\text{número de tests realizados}} \quad (6.3)$$

Con el objeto de establecer una base apropiada para la comparación entre las diferentes técnicas, se modificaron sus correspondientes α_L con el fin de lograr un valor del *AVTI* de aproximadamente 0,05 para todas ellas. Una vez realizados estos cambios, los valores de





AVTI para todas las estrategias quedaron comprendidos entre 0,04 y 0,06 (entre 5 y 7 falsas alarmas).

6.3 Comparaciones entre *NM*, *WKFH-AT*, *OSS*, *VM* y *ME*

En esta sección se presenta una comparación entre las estrategias *NM* (Nomikos y MacGregor, 1994, 1995), *WKFH-AT* (Wold y col., 1998; Aguado y col., 2007), *OSS* (Alvarez y col., 2007), el enfoque de Ventana Móvil (*VM*) (Lennox y col., 2001) y los Modelos Evolutivos (*ME*) (Ramaker y Van Sprang, 2004).

Los valores de α_L de cada técnica fueron ajustados para conseguir un valor de *AVTI* de aproximadamente 0,05 tal como se indicó en la sección anterior. El número de componentes principales (*CPs*) retenidas se determinó teniendo en cuenta los resultados obtenidos por diferentes criterios (validación cruzada, *SCREE* plots y porcentaje de reconstrucción mínimo (Qin y Dunia, 2000)). Los valores de reconstrucción para todas las técnicas que usan proyección en *VLs* se encuentran alrededor del 60% de la varianza total. Si bien estos valores pueden parecer bajos si se los compara con algunos de los reportados en la literatura (Nomikos y MacGregor, 1994, 1995; Ramaker y Van Sprang, 2004), debe tenerse en cuenta que la relación entre las variables monitoreadas en este caso de estudio es altamente no lineal a diferencia de lo que ocurre con los empleados por algunos de estos autores.

Las cartas de control y los gráficos de contribución son la forma más común de presentar los resultados obtenidos en las etapas de detección e identificación, respectivamente. Sin embargo, dado el gran número de fallas simuladas y de técnicas analizadas, no resulta viable usar estas representaciones para todos los casos. A los efectos ilustrativos, las cartas de control y los gráficos de contribución para algunas de las metodologías consideradas se muestran sólo para la falla 10.

Las Figuras 6.3 a 6.7 muestran, para esta falla, las trayectorias de los estadísticos y las contribuciones de las variables obtenidas con los métodos *NM*, *WKFH-AT* y *OSS*, al tiempo en el que se detectan las fallas. En los gráficos de las trayectorias el área sombreada corresponde a la región de confianza para cada uno de los estadísticos. Los valores que generan una alerta se resaltan con , mientras que la primera señal de alarma está indicada con el símbolo  rodeando la observación correspondiente. En el caso de los gráficos de las contribuciones, las líneas indicadas con  y  representan los dos valores límites utilizados para las *CVs*. En particular, las Figuras 6.3 y 6.4 muestran las trayectorias temporales de los estadísticos *D* y *SPE*, así como los valores de las contribuciones $c_{j,k}^D$ y $c_{j,k}^{SPE}$ cuando se utiliza el método *NM*. Las Figuras 6.5 y 6.6 representan resultados análogos, obtenidos aplicando el enfoque *WKFH-AT*. Por último, la Figura 6.7 muestra el perfil de T^2 y las contribuciones $c_{j,k}^{T^2}$, que resultan de aplicar *OSS*. Esta falla fue indicada en la observación $k=19$ (36 min) por todos los estadísticos. Puede notarse también que en todos los casos, las descomposiciones de las métricas indican al caudal de alimentación Q_0 como variable sospechosa. Sin embargo, en el caso de las $c_{j,k}^{SPE}$ algunas otras variables exceden también los límites correspondientes.

La Tabla 6.6 contiene los resultados obtenidos con cada una de las técnicas consideradas para todas las fallas simuladas. En esta tabla, luego de la columna que identifica a las fallas (numeradas del 1 al 82), las dos siguientes contienen la identificación de la técnica y el número de observación en el que ocurrió la falla (k_{falla}), respectivamente.

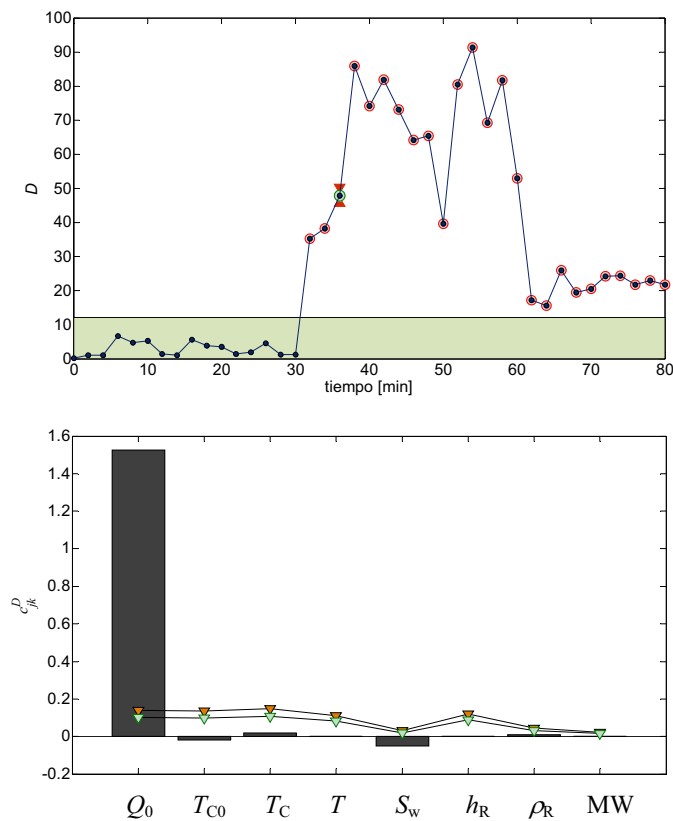


Figura 6.3: Perfil de D y valores de $c_{j,k}^D$, falla: 10, método: NM, $k=19$

El siguiente subgrupo de diez columnas presenta los resultados del monitoreo de cada uno de los *batches* defectuosos, mediante el uso del estadístico de Hotelling, ya sea en el espacio de las *CPs* o en el espacio original de las mediciones utilizado en *OSS*.

La primera de las columnas de dicho subgrupo indica la observación (k) donde la falla fue detectada por D o T^2 al 95% de confianza global (Alarma D). En todos los casos en los que se usaron proyecciones en *CPs*, sus distribuciones se consideraron como normales multivariadas a fin de obtener los valores críticos para D (Nomikos y MacGregor, 1994, 1995). En cambio las distribuciones de T^2 para *OSS* se obtuvieron mediante el método *KDE* (Apéndice 3).

En la segunda columna del mismo subgrupo se encuentra el tiempo a la señal de acción (*AST*) para D o T^2 , según sea el caso, expresado en términos de número de observaciones. El

AST es el tiempo transcurrido entre la introducción de una perturbación que altera la operación normal y el momento en el que los estadísticos indican alguna anomalía. Mientras menor sea AST mejor será el desempeño de la estrategia.

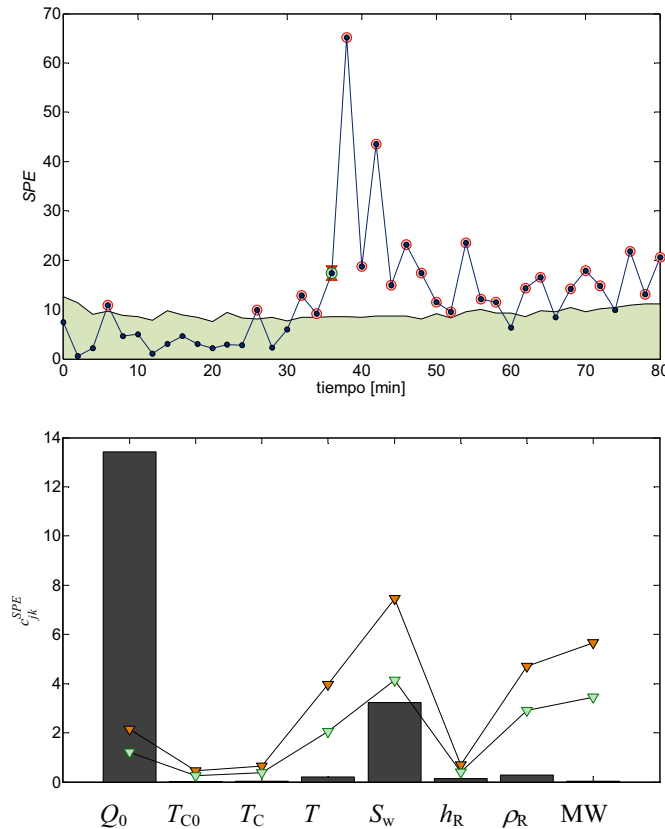


Figura 6.4: Perfil de SPE y valores de $c_{j,k}^{SPE}$, falla: 10, método: NM , $k=19$

Las siguientes ocho columnas contienen información sobre la contribución de cada una de las variables monitoreadas (Q_0 , T_{C0} , T_C , T , S_w , h_R , ρ_R , MW) al valor final del estadístico. A estas columnas se le asignan valores 2 ó 3 según sea que la CV en cuestión ($c_{j,k}^D$ ó $c_{j,k}^{T^2}$) exceda los valores críticos $\bar{c}_{j,k} + 2\sigma_{c_{j,k}}$ ó $\bar{c}_{j,k} + 3\sigma_{c_{j,k}}$, respectivamente. Aquí, $\bar{c}_{j,k}$ es el valor medio de las contribuciones de la j -ésima variable en la k -ésima observación, obtenido a partir de las $c_{j,k}^D$ y $c_{j,k}^{T^2}$ de la PR , y $\sigma_{c_{j,k}}$ es el desvío estándar correspondiente.

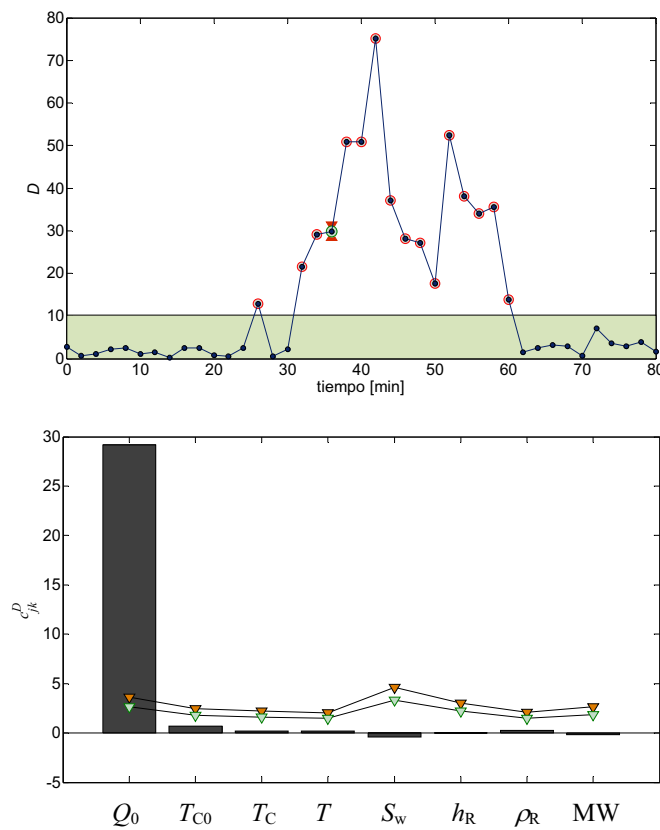


Figura 6.5: Perfil de D y valores de $c_{j,k}^D$, falla: 10, método: $WKFH-AT$, $k=19$

El segundo subgrupo de columnas corresponde al análisis del estadístico SPE siendo la estructura y ordenamiento equivalentes a lo explicado anteriormente. Los límites para las $c_{j,k}^{SPE}$ se obtienen como se explicó en el Capítulo 2 al 95% y 99% de confianza para las indicaciones 2 y 3, respectivamente.

Los espacios vacíos en la Tabla 6.6 indican que la falla no fue detectada en ninguna de las K observaciones o que ninguna de las CVs excede los límites correspondientes. En el caso de OSS los valores correspondientes al error de representación (SPE) estarán siempre vacíos dado que no se utiliza ninguna técnica de proyección en este caso.

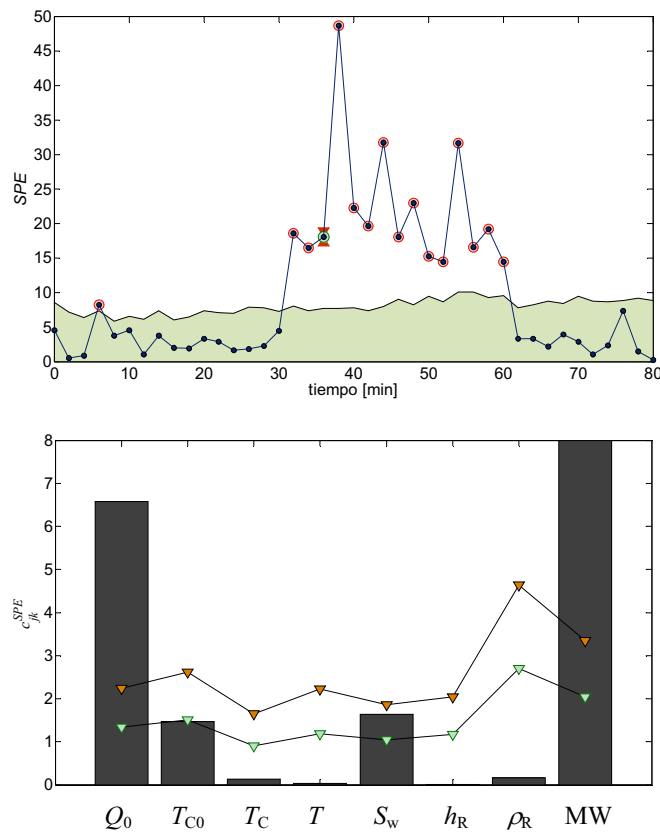


Figura 6.6: Perfil de SPE y valores de $c_{j,k}^{SPE}$ falla: 10, método: $WKFH-AT$, $k=19$

A continuación se analizan algunos aspectos relacionados con la capacidad de detección de las diferentes estrategias.

La influencia combinada del tiempo en que se produce la falla y su magnitud se observa en algunos ejemplos considerados. Las fallas 14 y 17 corresponden a un cambio escalón en el caudal de alimentación al reactor de +10% y -10%, respectivamente, a los 52 minutos de operación. Estas fallas no son detectadas por ninguna de las técnicas consideradas. Esto se atribuye al efecto combinado de la magnitud relativamente pequeña del valor de la falla y la ubicación temporal de la misma cerca del final de la operación. Los reactores de polimerización del tipo considerado como caso de estudio presentan la mayor sensibilidad a los cambios al inicio de la operación y las variaciones en el curso medio o final de la misma pueden no tener consecuencias mayores.

Las fallas 18 a 23 indican, de alguna manera, el límite de detección de las técnicas analizadas. Dichas fallas corresponden a escalones de +5% y -5% en Q_0 , efectuados a los 12, 32 y 52 minutos de operación. Entre estas fallas, las 18 y 21 fueron las únicas detectadas y sólo por los métodos de VM a través de SPE . Ninguno de los métodos fue capaz de producir una señal de alarma para el resto de estas perturbaciones. Pueden observarse efectos similares en los casos 25 a 27, debido a la pequeña magnitud de las fallas simuladas.

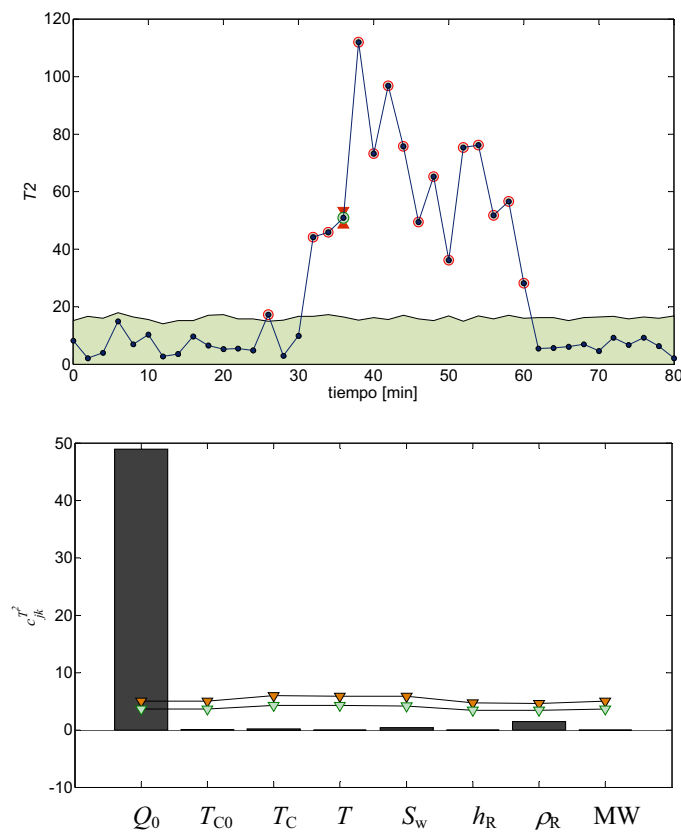


Figura 6.7: Perfil de T^2 y valores de $c_{j,k}^{T^2}$, falla: 10, método: OSS, $k=19$

Tabla 6.6: Resultados ... (Continuación)

K _{plata}	AST D	Q ₀	T _c	T _{co}	T	S _w	h _r	Pr	MW	Alarma SPE	Q ₀	T _c	T _{co}	T	S _w	h _r	Pr	MW	Alarma SPE	K _{plata}	AST D	Q ₀	T _c	T _{co}	T	S _w	h _r	Pr	MW	Alarma SPE				
																															Alarma D	Alarma SPE		
21	7									10	3	2			2				2	3	1													
																															Nomitos y MacGregor (NM)			
																															WKFFH-AT			
																															OSS			
																															OSS -I/M (Δk=2)			
22	17									10	3	2			2				2	3	1													
																															Nomitos y MacGregor (NM)			
																															WKFFH-AT			
																															OSS			
																															OSS -I/M (Δk=2)			
23	27									10	3	2			2				2	3	1													
																															Nomitos y MacGregor (NM)			
																															WKFFH-AT			
																															OSS			
																															OSS -I/M (Δk=2)			
24	1									26	25	3	2	3	3	3	2	3	2	3	3	2												
																															Nomitos y MacGregor (NM)			
																															WKFFH-AT			
																															OSS			
																															OSS -I/M (Δk=2)			
25	1									10	9	2	3	3	3	3	3	2	3	3	1													
																																Nomitos y MacGregor (NM)		
																																WKFFH-AT		
																																OSS		
																																OSS -I/M (Δk=2)		
26	1									10	3	2			2				2	3	1													
																																Nomitos y MacGregor (NM)		
																																WKFFH-AT		
																																OSS		
																																OSS -I/M (Δk=2)		
27	1									13	12	3	3	3	3	2			2	3	1													
																																Nomitos y MacGregor (NM)		
																																WKFFH-AT		
																																OSS		
																																OSS -I/M (Δk=2)		
28	1									15	14	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2	3	2	2	2			
																																Nomitos y MacGregor (NM)		
																																WKFFH-AT		
																																OSS		
																																OSS -I/M (Δk=2)		
29	1									15	14	3	3	3	3	3	3	3	3	3	15	14	3	3	3	3	3	3	3	3	3			
																																Nomitos y MacGregor (NM)		
																																WKFFH-AT		
																																OSS		
																																OSS -I/M (Δk=2)		
30	1									19	18	3	3	3	3	3	3	3	3	3	19	18	3	3	3	3	3	2	3	2	2	2		
																																	Nomitos y MacGregor (NM)	
																																	WKFFH-AT	
																																	OSS	
																																	OSS -I/M (Δk=2)	

Tabla 6.6: Resultados ... (Continuación)

ID	Alarma D										Alarma SPE									
	K _{salida}	T ₀	T _c	T _w	P _r	MW	Q ₀	T ₀	T _c	T _w	P _r	MW	Q ₀	T ₀	T _c	T _w	P _r	MW		
51	Nomitos y MacGregor (NM)																			
	WKFFH-AT																			
	OSS																			
	OSS - /M ($\Delta k=2$)																			
	Modelos Locales (ML)																			
	Ventana Móvil ($\Delta k=2$)																			
52	Nomitos y MacGregor (NM)																			
	WKFFH-AT																			
	OSS																			
	OSS - /M ($\Delta k=2$)																			
	Modelos Locales (ML)																			
	Ventana Móvil ($\Delta k=2$)																			
53	Nomitos y MacGregor (NM)																			
	WKFFH-AT																			
	OSS																			
	OSS - /M ($\Delta k=2$)																			
	Modelos Locales (ML)																			
	Ventana Móvil ($\Delta k=2$)																			
54	Nomitos y MacGregor (NM)																			
	WKFFH-AT																			
	OSS																			
	OSS - /M ($\Delta k=2$)																			
	Modelos Locales (ML)																			
	Ventana Móvil ($\Delta k=2$)																			
55	Nomitos y MacGregor (NM)																			
	WKFFH-AT																			
	OSS																			
	OSS - /M ($\Delta k=2$)																			
	Modelos Locales (ML)																			
	Ventana Móvil ($\Delta k=2$)																			
56	Nomitos y MacGregor (NM)																			
	WKFFH-AT																			
	OSS																			
	OSS - /M ($\Delta k=2$)																			
	Modelos Locales (ML)																			
	Ventana Móvil ($\Delta k=2$)																			
57	Nomitos y MacGregor (NM)																			
	WKFFH-AT																			
	OSS																			
	OSS - /M ($\Delta k=2$)																			
	Modelos Locales (ML)																			
	Ventana Móvil ($\Delta k=2$)																			
58	Nomitos y MacGregor (NM)																			
	WKFFH-AT																			
	OSS																			
	OSS - /M ($\Delta k=2$)																			
	Modelos Locales (ML)																			
	Ventana Móvil ($\Delta k=2$)																			
59	Nomitos y MacGregor (NM)																			
	WKFFH-AT																			
	OSS																			
	OSS - /M ($\Delta k=2$)																			
	Modelos Locales (ML)																			
	Ventana Móvil ($\Delta k=2$)																			
60	Nomitos y MacGregor (NM)																			
	WKFFH-AT																			
	OSS																			
	OSS - /M ($\Delta k=2$)																			
	Modelos Locales (ML)																			
	Ventana Móvil ($\Delta k=2$)																			

Tabla 6.6: Resultados ... (Continuación)

K _{salida}	Alarma D	AST D	Q̄ ₀	T ^{CO}	T ^C	T	S ^w	h _r	Pr	MW	Alarma SPE	Q̄ ₀	T ^{CO}	T ^C	T	S ^w	h _r	Pr	MW
76	1	28	27	Nomitos y MacGregor (NM)															
				WKFFH-AT															
				OSS															
				OSS -/M (Δk=2)															
				Modelos Locales (ML)															
77	1	3	2	Nomitos y MacGregor (NM)															
				WKFFH-AT															
				OSS															
				OSS -/M (Δk=2)															
				Modelos Locales (ML)															
78	1	3	2	Nomitos y MacGregor (NM)															
				WKFFH-AT															
				OSS															
				OSS -/M (Δk=2)															
				Modelos Locales (ML)															
79	1	3	2	Nomitos y MacGregor (NM)															
				WKFFH-AT															
				OSS															
				OSS -/M (Δk=2)															
				Modelos Locales (ML)															
80	1	3	2	Nomitos y MacGregor (NM)															
				WKFFH-AT															
				OSS															
				OSS -/M (Δk=2)															
				Modelos Locales (ML)															

Tabla 6.6: Resultados ... (Continuación)

	Falla	Alarma D										Alarma SPE									
		AST D	Q_0	T_{co}	T_c	T	S_w	I_k	D_k	MW	AST SPE	Q_0	T_{co}	T_c	T	S_w	I_k	D_k	MW		
81	Nomikos y MacGregor (NM)	4	3	3		3	3					3	2		2	3	3	3		3	
	WKFH-AT	3	2	2		3	3	3				3	2		2	3	2	3		2	
	OSS	3	2			3															
	OSS-VM ($\Delta k=2$)	3	2			3															
	Modelos Locales (ML)	3	2			3	3	3				3	2			3		2	2	2	
	Ventana Móvil ($\Delta k=2$)	3	2			3	3	3			3	2	3			3	2	3		2	
	Ventana Móvil ($\Delta k=3$)	3	2			3	3	3		3	2	3				3	2	3			
	Modelos Evolutivos	3	2			3	3	3			6	5		3	3	3					
82	Nomikos y MacGregor (NM)	3	2			3	3	3		3	3		3	3	3						
	WKFH-AT	3	2			3	3			3	2		3	3	3		3				
	OSS	3	2			3															
	OSS-VM ($\Delta k=2$)	3	2			3															
	Modelos Locales (ML)	3	2			3	3	3		2	3	2	3	3	3	3	3	2			
	Ventana Móvil ($\Delta k=2$)	3	2			3	3	3		2	3	2		3	3	3				2	
	Ventana Móvil ($\Delta k=3$)	3	2			3	3	3			3	2		3	3	3				2	
	Modelos Evolutivos	3	2			3	3	3			3	2		3	3	3				2	

Las variaciones en las condiciones iniciales deberían tener, en este tipo de sistema, una gran influencia en la trayectoria de las variables durante el transcurso de la operación. Algunas alteraciones en las condiciones iniciales fueron simuladas sobre variables no medidas como la concentración de monómero (M_0), la fracción volumétrica de agua (V_w) y la cantidad inicial de iniciador agregado (I_0). Las fallas 71 a 75, correspondientes a variaciones en M_0 no produjeron señal de alarma en ningún caso. Las variaciones en la cantidad de iniciador también presentaron dificultades para la detección.

Del análisis de los resultados obtenidos surgen las siguientes conclusiones sobre las capacidades de detección de las estrategias examinadas:

- a. La mayoría de las técnicas tiene buena capacidad de detección de fallas.
- b. Los enfoques de VM más PCA y ME presentan, por lo general, un retardo en la detección de la falla cuando se los compara con los desempeños de NM y WKFH-AT. Este retraso es generalmente mayor en el estadístico de Hotelling que en SPE y su tamaño depende del tipo y magnitud de la falla considerada. Mientras más importante es la magnitud de la falla, menor es el retardo en la detección.

- c. Algunas de las fallas simuladas en variables no medidas, como la concentración de iniciador (I_0) o la fracción volumétrica de agua (V_w), fueron mayoritariamente detectadas por los métodos de VM y ME , como es el caso de las fallas 61, 70 y 76.
- d. El 73% de las fallas simuladas fue detectado por alguna de las siguientes estrategias: OSS , NM o $WKFH-AT$ y el 65% fue indicado por las tres estrategias simultáneamente.
- e. Puede observarse en los resultados de la Tabla 6.6 que el desempeño de OSS es similar o superior al de las estrategias NM y $WKFH-AT$. En el 77% de los casos en los que alguna de las tres técnicas (OSS , NM o $WKFH-AT$) indicaron la existencia de la falla, OSS la detectó con el menor valor de AST en forma simultánea con NM ó $WKFH-AT$.
- f. En el 40% de los casos en los que alguna de las tres técnicas (OSS , NM o $WKFH-AT$) indicó la existencia de la falla mediante el estadístico de Hotelling, las tres estrategias tuvieron los mismos valores de AST mientras que en el 30% fueron detectadas por OSS en forma simultánea con sólo una de las otras técnicas. En el 20% de los casos los valores de AST para OSS fueron los menores.
- g. El 5% de las alarmas obtenidas con OSS tienen valores de AST mayores que los de NM o $WKFH-AT$, para T^2 y D respectivamente.
- h. Sólo el 5% de los casos en los que NM o $WKFH-AT$ detectaron alguna falla ésta no fue señalada por OSS .
- i. El uso del enfoque de VM combinado con OSS no ofrece ninguna mejora con respecto a los resultados obtenidos exclusivamente con OSS .

- j.* El enfoque de modelos *PCA* locales presenta resultados similares a las técnicas *NM*, *WKFH-AT* y *OSS* cuando el tamaño de la falla simulada es considerablemente alto. Cuando las fallas simuladas no son importantes los modelos locales muestran retardos en la detección en comparación con los métodos mencionados anteriormente.
- k.* En algunos casos los valores de *AST* para los modelos locales son menores que los correspondientes a *OSS* (fallas 34, 40, 41, 59). En éstos la detección de la falla viene dada por *SPE* y no por *D*.
- l.* La detección de la falla 13 mediante *NM* ocurrió en la observación 16 a través de *SPE*. Sin embargo, la falla fue simulada a partir de la observación 17, por lo que *AST* resulta negativo. Esto indica la existencia de una falsa alarma. Dado que las fallas se simularon sobre la base de la trayectoria promedio no sería posible, en principio, declarar una falla antes de $k=19$. No obstante, los valores de ruido que se agregan sobre la trayectoria de la falla serían los causantes de este comportamiento.

Las Figura 6.8 y Figura 6.9 muestran los perfiles de *SPE* obtenidos para *NM* con y sin ruido, respectivamente. En la Figura 6.9 se observa que el valor de *SPE* no alcanza a disparar la alarma pero tiene un incremento claro a partir de la observación $k=17$ (32 min).

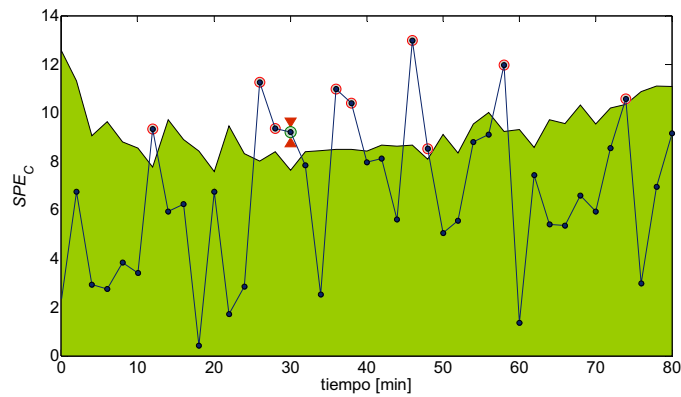


Figura 6.8: Perfil de SPE para la falla 13 obtenido con NM

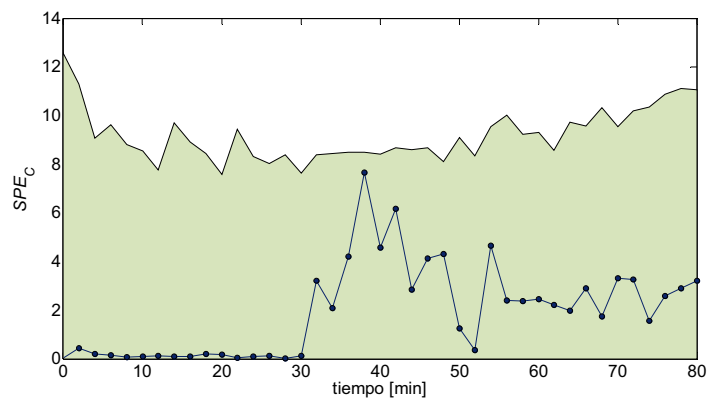


Figura 6.9: Perfil de SPE para la falla 13 (sin ruido) obtenido con NM

En cuanto al desempeño de las estrategias en la etapa de identificación, la Tabla 6.6 muestra la descomposición de los estadísticos en términos de las CVs para el tiempo donde se declara la falla.

El análisis de estos resultados indica que resulta más fácil la interpretación de las causas de cada falla utilizando las CVs al estadístico de Hotelling que al SPE , para casi todos los casos analizados. En la mayoría de ellos existen más variables fuera de control debidas a los residuos generados por las proyecciones que a las variaciones sistemáticas que se ponen de manifiesto en D o T^2 . Debe mencionarse que cuando los valores de SPE exceden los límites máximos, el valor de D y por lo tanto las CVs asociadas pueden resultar poco confiables (Jackson, 1991; Nomikos y MacGregor, 1994; Westerhuis y col., 2000). Sin

embargo, en la gran mayoría de los casos, las variables señaladas como sospechosas en D o T^2 guardan una estrecha relación con la causa original de la falla.

Resulta claro también que a mayor demora en la detección de una dada perturbación más difícil será la identificación de la misma. Esto se debe a los efectos de su propagación hacia el resto de las variables medidas. De manera similar, la existencia de alteraciones en los valores de las variables no medidas debe ponerse de manifiesto en términos de las variables monitoreadas. Por lo tanto, es de esperar que las tareas de identificación resulten mucho más complicadas que cuando la falla se produce sobre una variable medida directa o indirectamente.

En el primer grupo de fallas (fallas 1 a 23), el valor de Q_0 fue indicado como una de las principales desviaciones respecto del transcurso esperado en casi todos los casos donde la detección fue posible. Las indicaciones obtenidas mediante la descomposición del estadístico de Hotelling (T^2 y/o D) resultan más clara que las correspondientes al SPE , incluso cuando la alarma en SPE es simultánea o anterior a la de D . La interpretación de los resultados de OSS es clara y precisa en todos los casos en los que T^2 dio lugar a alarmas. El uso conjunto del enfoque de VM con OSS no goza sin embargo de las mismas capacidades. Una de las causas posibles de este comportamiento reside en el hecho de que las matrices de correlación obtenidas usando el enfoque de VM en el espacio original de las mediciones tienen números de condición demasiado altos y el mismo aumenta drásticamente con el valor de Δk . En consecuencia, las matrices calculadas por inversión de las de correlación resultan poco confiables, deteriorando la estimación de T^2 y las correspondientes $c_{j,k}^{T^2}$.

Respecto del resto de las técnicas, las contribuciones obtenidas para D en este primer grupo de fallas indican también a Q_0 como una de las variables responsables de la señal de alarma. En la mayor parte de los casos la indicación es precisa y única, sin embargo en

algunas situaciones aparecen también otras variables indicadas como sospechosas, siendo posible encontrar una relación lógica para las mismas. Las variaciones en Q_0 pueden producir cambios significativos en la densidad de la mezcla reactiva, en el peso molecular promedio y hasta en la temperatura del reactor si la perturbación en el caudal de alimentación es importante (fallas 3 y 5) o si la falla es detectada mucho después que la misma se inició (falla 1). Este razonamiento puede usarse para la interpretación de las $c_{j,k}^{SPE}$.

Cuando varias CVs exceden sus valores límites de manera simultánea, la interpretación de estos resultados se vuelve muy complicada dificultando más aún las tareas relacionadas con el diagnóstico de la falla. Teniendo en cuenta estos inconvenientes, algunos autores han propuesto el uso de técnicas tales como redes neuronales y/o lógica difusa para las etapas de identificación-diagnóstico. Sin embargo, la gran cantidad de fallas y situaciones anormales “rotuladas” con las que debe contarse hacen difícil su implementación en la práctica.

En el segundo grupo de fallas (24 a 46) las indicaciones de las variables sospechosas obtenidas mediante OSS son siempre consistentes con la causa real de la alarma. Para las fallas 29 y 31, la demora en la detección de las mismas podría ser la responsable de valores altos en la temperatura del reactor (T). La combinación con estrategias de VM tampoco resultó beneficiosa en estos casos. Las contribuciones obtenidas mediante la descomposición de D (Westerhuis y col., 2000) no resultan tan claras como para el grupo de fallas anterior. En muchos casos la temperatura del reactor (T) y la temperatura de salida del refrigerante (T_C) se indican como sospechosas (ej.: fallas 33, 34, 36, 43, 44, 46). Como se explicó anteriormente, ésto puede adjudicarse a la demora en la detección de la falla y a la consecuente propagación de sus efectos en T y T_C . Sin embargo, la concentración de surfactante (S_W), el nivel del reactor (h_R) y la densidad de la mezcla (ρ_R) también se indican como variables sospechosas en otros casos (fallas 30, 31, 32, 35, 38 y 39). Si bien dichas variables pueden ser afectadas por

los cambios de T , estas descomposiciones no señalan a ésta como variable sospechosa (en la mayor parte de los métodos). Una posible explicación para este comportamiento proviene de la naturaleza de las fallas simuladas. Debido a que las fallas en la temperatura de entrada del agua de refrigeración (T_{C0}) se simularon como una secuencia de rampas (Tabla 6.4), es posible que T experimente cambios tales que no sean suficientes para ponerse de manifiesto en forma directa, sino por cambios marcados en el resto de las variables de proceso. Al igual que en el grupo de fallas anterior, los valores de las contribuciones obtenidas para SPE resultan más complejas de interpretar que las correspondientes al estadístico de Hotelling. Los razonamientos anteriores que permiten entender la propagación de los efectos de la fallas son válidos también para explicar los resultados de esta métrica. En particular, en la mayor parte de los casos en los que S_W , h_R o ρ_R aparecen indicadas como sospechosas las CVs indican lo propio para T . En algunos casos, variables totalmente independientes como Q_0 también aparecen resaltadas.

Respecto del último grupo de fallas simuladas puede notarse un comportamiento similar en lo que hace a la identificación de las mismas. Las indicaciones obtenidas para D y T^2 para las fallas 47 a 52 indican acertadamente a h_R como la variable sospechosa. La identificación mediante las $c_{j,k}^{SPE}$ no resulta clara. Lo mismo ocurre con las fallas simuladas en la carga inicial de surfactante. Las fallas 59 a 64 son variaciones en la cantidad inicial de agua (V_W), que es una variable no medida. En la mayoría de los casos las variaciones iniciales de V_W fueron detectadas con valores altos de AST . Las variables indicadas como sospechosas por la descomposición de D y/o T^2 fueron las temperaturas del reactor (T) y de salida del agua de camisa (T_{C0}). Las indicaciones de las $c_{j,k}^{SPE}$ resultaron más dispares. Sin embargo, la densidad de la mezcla reactiva (ρ_R) fue resaltada en la mayor parte de los casos. Las alteraciones en la cantidad de iniciador y en la concentración inicial de monómero generaron

pocas alarmas. Las variaciones en T al inicio de la reacción se pusieron de manifiesto rápidamente. Dada la importancia de la temperatura en este tipo de procesos, y teniendo en cuenta que la primera etapa de la reacción es en la que ocurren la mayor cantidad de cambios en el sistema, es de esperar que las variaciones de la temperatura al inicio de la operación se propaguen rápidamente sobre las otras variables. Las contribuciones obtenidas para T^2 y D indican a T , S_w y T_C como las variables sospechosas en la mayoría de los casos. Las contribuciones a SPE resultan también mucho más difíciles interpretar.

6.4 Resultados del Análisis de Componentes Independientes (*MICA*)

Otra de las técnicas de proyección ampliamente utilizada en *SPC* es el Análisis de Componentes Independientes (*ICA*). Como se señaló en el Capítulo 2 las técnicas basadas en *ICA* se usan frecuentemente en el análisis de conjunto de datos no-gaussianos. Muchos autores señalaron que *ICA* es más adecuado que *PCA* para el tratamiento de este tipo de datos y que muestra un muy buen desempeño en las etapas de detección e identificación (Hyvärinen y col., 2001; Lee y col., 2004; Yoo y col., 2004).

En esta sección se presentan los resultados obtenidos mediante la aplicación de *MICA* tal como se propone en el trabajo de Yoo y col. (2004). El algoritmo *FastICA* propuesto por Hyvärinen y col. (2001) fue usado con el objetivo de encontrar la matrices \mathbf{W} , $\bar{\mathbf{W}}$. El conjunto de rutinas necesarias para el análisis de los datos puede obtenerse en el sitio web del Laboratorio de Ciencias de la Computación y Ciencias de la Información de la Universidad de Helsinki: <http://www.cis.hut.fi/projects/ica/fastica/>.

Una condición necesaria para la aplicabilidad del *ICA* es que los datos muestren un comportamiento no-gaussiano. Existen varias formas para determinar la no-gaussianidad de un conjunto de datos incluyendo algunos tests de hipótesis específicos. En este trabajo sin

embargo, se utiliza una herramienta mucho más simple para tener alguna idea de la no-gaussianidad presente en el conjunto de observaciones de la *NOC*. Los gráficos cuantil-cuantil permiten comparar datos provenientes de dos poblaciones cualesquiera. Si ambos conjuntos pueden ser descritos por la misma distribución, sus elementos, luego de cierto reordenamiento, se comportarían de manera tal de describir una recta en un plano de coordenadas. Una descripción más profunda sobre los gráficos cuantil-cuantil puede encontrarse en Martínez y Martínez (2002).

Las Figuras 6.10 y 6.11 muestran los gráficos cuantil-cuantil para los estadísticos D y SPE respectivamente cuando se utiliza *WKFH-AT*. Estos gráficos comparan los valores de dichas métricas con las correspondientes distribuciones de referencia que éstos seguirían si los datos en \mathbf{X} tuviesen una distribución normal multivariada. En ambos gráficos puede observarse un comportamiento similar. Las distribuciones de los dos estadísticos muestran un apartamiento de las referencias para valores altos de las métricas. Resulta entonces válido y de interés analizar las posibles ventajas asociadas al uso de *MICA* para el caso de estudio considerado.

Las Figuras 6.12 a 6.14 muestran los perfiles de I_d , I_e y SPE para la falla 10, juntamente con la correspondiente descomposición de los estadísticos en el tiempo en cual se declara la alarma. Al igual que en los casos anteriores, el área sombreada corresponde a la región de confianza para cada uno de los estadísticos. Los valores que generan una alerta se resaltan con \circ y la primera señal de alarma está indicada con el símbolo \blacktriangledown . En el caso de los gráficos de las contribuciones, las líneas indicadas con \blacktriangledown y \blacktriangledown muestran los valores límites para las *CVs* para los niveles 2 y 3, respectivamente

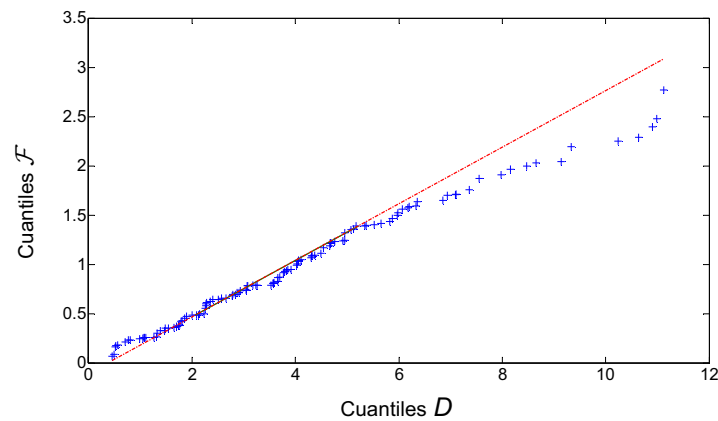


Figura 6.10: Gráfico cuantil-cuantil \mathcal{F}_{v_1, v_2} vs D

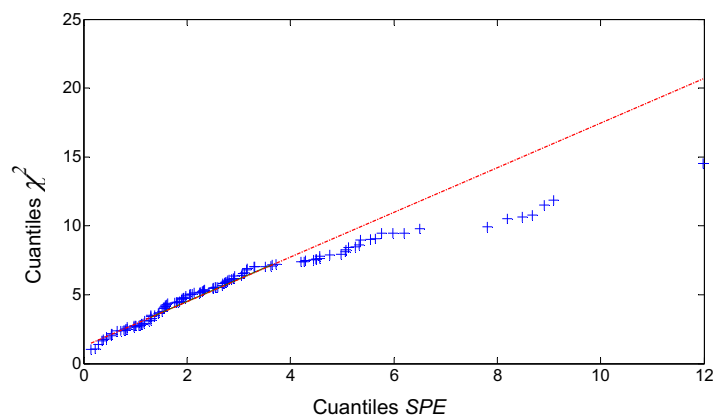


Figura 6.11: Gráfico cuantil-cuantil χ^2 vs SPE

En el caso de la falla 10, todos los estadísticos indican la existencia de la misma en la observación $k=19$. Sin embargo, y a diferencia de las técnicas consideradas en la sección anterior, es SPE el que muestra mejor desempeño en la etapa de identificación.

La Tabla 6.7 muestra, para todas las fallas consideradas, los valores de las tres métricas (I_d , I_e , SPE) y sus respectivos AST , como así también las contribuciones de los descriptores a cada estadístico ($c_{j,k}^{I_d}$, $c_{j,k}^{I_e}$, $c_{j,k}^{SPE}$). Éstas se calcularon mediante la descomposición de las métricas, tal como se describió en el Capítulo 2.

De la comparación entre los resultados mostrados en las Tablas 6.6 y 6.7 surge que el desempeño de $MICA$ para la detección de situaciones anormales es muy similar al obtenido

con las técnicas analizadas en la sección anterior. Sin embargo, se logró una disminución de los AST para SPE , cuando se los compara con los de la misma métrica en las técnicas clásicas.

La notable mejora en las capacidades de identificación del estadístico SPE es un punto que merece destacarse. Los valores de $c_{j,k}^{SPE}$ obtenidos para $MICA$ resultan mucho más claros que los calculados para cualquiera de las técnicas de la sección anterior. En todos los casos en los cuales SPE indicó la falla en el primer grupo de *batches* defectuosos (fallas 1 a 23), Q_0 fue correctamente indicada como la principal variable sospechosa. Los valores obtenidos para $c_{j,k}^{I_d}$ y $c_{j,k}^{I_e}$ para estos mismos *batches* permitieron corroborar este resultado, sin embargo sus indicaciones no fueron tan claras como las de SPE . Las $c_{j,k}^{SPE}$ asociadas a las fallas del segundo grupo (24 a 46), en las que se simularon cambios en la temperatura de ingreso de refrigerante, señalaron en forma acertada a T_{C0} como variable sospechosa. Los valores de $c_{j,k}^{I_e}$ destacan también a esta variable como principal responsable del desvío cuando resulta posible realizar la identificación. En algunos casos la temperatura de salida del refrigerante (T_C) tiene también valores altos de $c_{j,k}^{I_e}$, lo que podría deberse a los valores altos de AST . La identificación a partir de I_d resultó esencialmente pobre para este grupo de fallas.

Finalmente, para el tercer grupo de fallas (cambios en las condiciones iniciales) la descomposición de SPE también fue la que mejor indicó la causa de las mismas.

El subgrupo de fallas 47 a 52 generó rápidamente alarmas en I_d y SPE . No obstante, sólo las $c_{j,k}^{SPE}$ indicaron de forma adecuada la raíz del problema. Sólo dos de las fallas simuladas en S_w fueron detectadas por $MICA$ en el sub-grupo de fallas 53 a 58 y tan sólo una fue correctamente identificada por $c_{j,k}^{I_d}$. Ninguno de los estadísticos utilizados fue capaz de

detectar las fallas 65 a 70 o las 71 a 76, simuladas en las concentraciones iniciales de iniciador (I_0) y de monómero (M_0), respectivamente. Al igual que en los casos analizados en la sección precedente, la identificación de las fallas 77 a 82 resulta complicada debido a la rápida propagación de la falla en T en el resto de las variables medidas.

El análisis de los resultados obtenidos para *MICA* permite concluir que esta técnica no produce mejoras significativas para la detección de las fallas simuladas en comparación con *NM*, *WKFH-AT* y *OSS*.

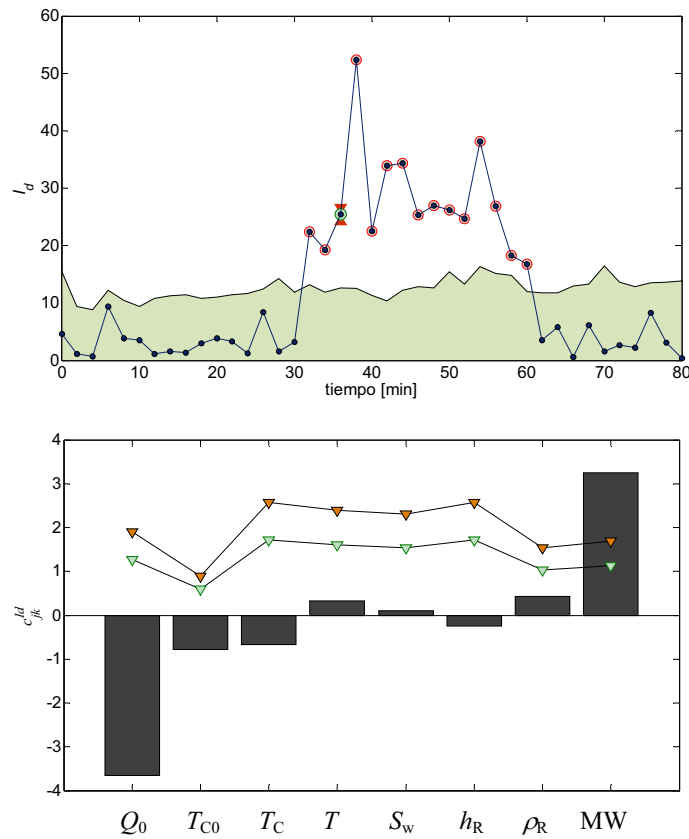


Figura 6.12: Perfil de I_d y valores de $c_{j,k}^{I_d}$, falla: 10, método: *MICA*, $k=19$

La notoria mejora en la capacidad de identificación de *SPE* se atribuye a que las técnicas de Componentes Independientes expanden espacios en cuya base no se imponen restricciones de ortogonalidad. De esta manera los versores base del espacio latente pueden copiar con mayor facilidad las relaciones impuestas por la correlación de los datos. La falta de precisión

en la identificación observada en I_d e I_e se adjudican al hecho de que si bien las CI_s no están correlacionadas desde el punto de vista global, existen correlaciones remanentes en cada uno de los subgrupos temporales que no se tienen en cuenta en este caso. La inclusión de matrices de correlación temporal (S_k) para las CI_s podría, en principio, mejorar notablemente la capacidad de identificación de dichos estadísticos.

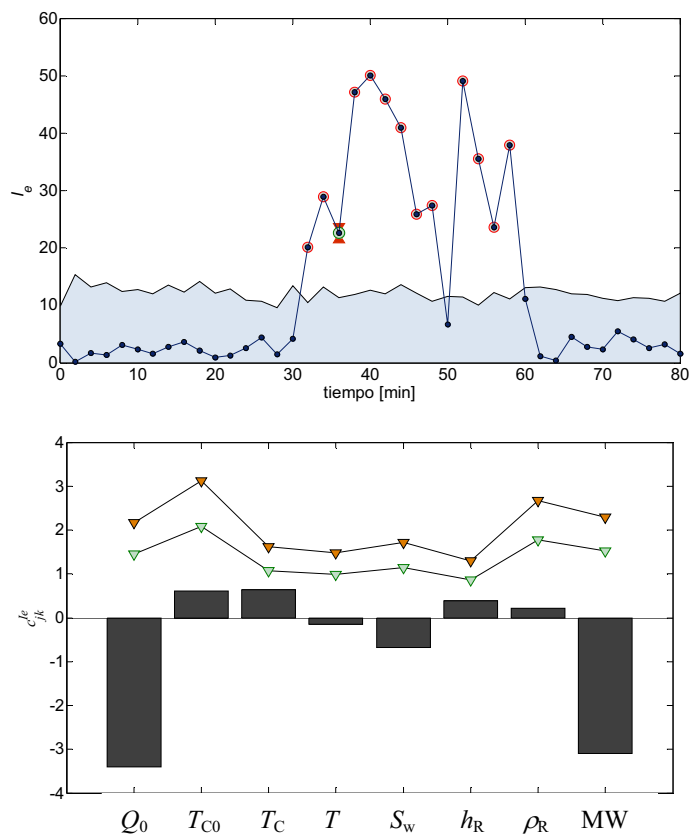


Figura 6.13: Perfil de I_e y valores de $c_{j,k}^{I_e}$, falla: 10, método: *MICA*, $k=19$

La comparación de los resultados de *MICA* con los de *OSS*, presentados en la Tabla 6.6, se realiza en base a los casos para los cuales alguna de las estrategias *OSS*, *NM*, *WKFH-AT* o *MICA* detectó la falla simulada. Es posible realizar las siguientes observaciones:

El 12% de las fallas que no fueron detectadas por *MICA* fueron señaladas correctamente por *OSS* (fallas 16, 54, 57, 62, 69). Sin embargo no se encontraron ejemplos de la situación inversa.

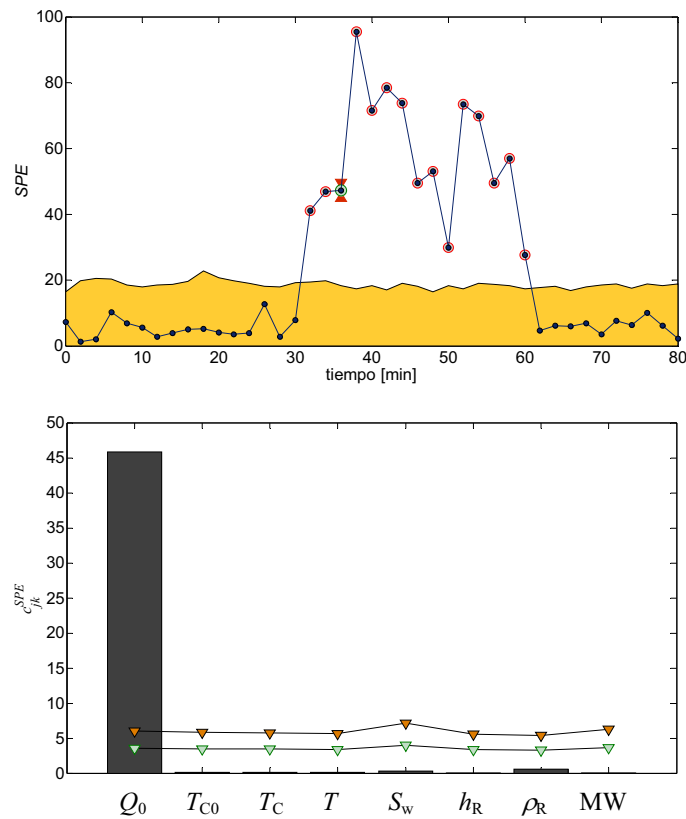


Figura 6.14: Perfil de *SPE* y valores de $c_{j,k}^{SPE}$, falla: 10, método: *MICA*, $k=19$

Si se considera el *AST*, *OSS* presentó valores inferiores a los de *MICA* en el 10% de los tests (fallas 12, 35, 37, 38, 39 y 59), la condición inversa se produjo sólo en el 5% de los casos analizados (fallas 30, 31, 64).

En algunos tests, la primera indicación del estado anómalo de *MICA* se debe a I_e pero no es posible llevar a cabo la identificación del mismo (40, 44).

6.5 Referencias

- Aguado, D., A. Ferrer, J. Ferrer y A. Seco, "Multivariate SPC of a Sequencing Batch Reactor for Wastewater Treatment". *Chemometrics and Intelligent Laboratory Systems* **2007**, 85, (1), 82-93.
- Alvarez, C. R., A. Brandolin y M. C. Sánchez, "On the Variable Contributions to the D-Statistic". *Chemometrics and Intelligent Laboratory Systems* **2007**, 88, (2), 189-196.
- Hyvärinen, A., J. Karhunen y E. Oja (2001). Independent Component Analysis, New York, John Wiley & Sons.
- Jackson, J. E. (1991). A User's Guide to Principal Components. Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics, New York, John Wiley & Sons, Inc.
- Lee, J. M., C. Yoo y I. B. Lee, "Statistical Process Monitoring with Independent Component Analysis". *Journal of Process Control* **2004**, 14, (5), 467-485.
- Lennox, B., G. A. Montague, H. G. Hiden, G. Kornfeld y P. R. Goulding, "Process Monitoring of an Industrial Fed-Batch Fermentation". *Biotechnology and Bioengineering* **2001**, 74, (2), 125-135.
- Liptak, B. G. (2003). Instrument Engineers' Handbook, Volume 1: Process Measurement and Analysis, London, CRC PRESS.
- Martinez, W. L. y A. R. Martinez (2002). Computational Statistics Handbook with Matlab, London, Chapman & Hall/CRC.
- Nomikos, P. y J. F. MacGregor, "Monitoring of Batch Process Using Multiway Principal Component Analysis". *AIChE Journal* **1994**, 40, (8), 1361-1375.
- Nomikos, P. y J. F. MacGregor, "Multivariate SPC Charts for Monitoring Batch Processes". *Technometrics* **1995**, 37, (1), 41-59.
- Qin, S. J. y R. Dunia, "Determining the Number of Principal Components for Best Reconstruction". *Journal of Process Control* **2000**, 10, (2), 245-250.
- Ramaker, H.-J. y E. N. M. Van Sprang (2004). Statistical Batch Process Monitoring. Ph.D. Thesis, Department of Chemical Engineering, Amsterdam, Universiteit van Amsterdam.
- Westerhuis, J. A., S. P. Gurden y A. K. Smilde, "Generalized Contribution Plots in Multivariate Statistical Process Monitoring". *Chemometrics and Intelligent Laboratory Systems* **2000**, 51, (1), 95-114.
- Wold, S., N. Kettaneh, H. Friden y A. Holmberg, "Modelling and Diagnostics of Batch Processes and Analogous Kinetic Experiments". *Chemometrics and Intelligent Laboratory Systems* **1998**, 44, (1-2), 331-340.

Yoo, C. K., J. M. Lee, P. A. Vanrolleghem y I. B. Lee, "On-Line Monitoring of Batch Processes Using Multiway Independent Component Analysis". *Chemometrics and Intelligent Laboratory Systems* **2004**, 71, (2), 151-163.

Página en Blanco

Capítulo 7

Conclusiones y Trabajos Futuros

7.1 Conclusiones	213
7.2 Trabajos Futuros.....	215

Página en Blanco

7 Conclusiones y Trabajos Futuros

7.1 Conclusiones

A lo largo de este trabajo de tesis se discutieron las bases teóricas que sustentan muchas de las metodologías existentes en el área del *SPC*. Los problemas asociados a la implementación de este tipo de estrategias son variados y dependen fuertemente de la complejidad inherente a cada aplicación. Los procesos *batch*, en particular, presentan una serie de complicaciones adicionales dada la naturaleza propia de los datos que se usan para describirlos y fueron el tema central de este estudio. Los análisis y comparaciones de las técnicas analizadas en esta tesis se realizaron en base a los resultados de su aplicación a un sistema de polimerización en emulsión de metacrilato de metilo. Se analizaron y discutieron las principales ventajas y desventajas de cada uno de los enfoques considerados en forma comparativa y se establecieron los puntos de partida para la interpretación y comparación de los resultados.

Se prestó particular atención a un conjunto de procedimientos basados en *PCA*, que es sin duda, la metodología más utilizada en el área de *SPC*, tanto en procesos continuos como *batch*. Los resultados obtenidos mediante dichas técnicas mostraron que, para este caso de estudio particular, las diferentes modificaciones propuestas al método original presentado por Nomikos y MacGregor no resultan en mejoras significativas en el desempeño de la estrategia de control en su conjunto. Sin embargo, la eliminación de la necesidad de estimar los valores futuros para las mediciones de proceso y la incorporación de las estructuras de varianza locales son puntos a resaltar en metodologías tales como *WKFH-AT*, *ML* y *VM*, por ejemplo. En general, los resultados obtenidos mostraron que el desempeño de la gran mayoría de estas

técnicas es comparable desde el punto de vista de la detección de fallas ya que ninguna de ellas probó ser claramente superior a las demás.

Las diferencias más marcadas estuvieron relacionadas a la etapa de identificación de las fallas. La identificación del conjunto de variables sospechosas se hizo mediante el cálculo de las CVs al valor de los estadísticos utilizados en el monitoreo. En general, las $c_{j,k}^{SPE}$ resultaron mucho más difíciles de interpretar que las $c_{j,k}^D$, incluso en los casos en los que la detección por parte de SPE fue simultánea o anterior a la correspondiente al estadístico D . Del análisis de los resultados obtenidos aplicando $MICA$ en base a un desdoblamiento VW se desprende que el desempeño de dicha metodología en lo que hace a la detección de situaciones anormales es comparable a los obtenidos usando técnicas de $MPCA$. Sin embargo, en la etapa de identificación de fallas, los valores de CVs a los estadísticos I_d y I_e no resultaron tan claros como las de $c_{j,k}^{SPE}$, siendo este último el más preciso de los tres cuando se usa esta metodología. Esto se atribuye al hecho de que las técnicas basadas en ICA no están sometidas a restricciones de ortogonalidad sobre los vectores que forman la base, por lo que resultan más versátiles para expandir cualquier espacio.

El aporte más importante de esta tesis está directamente ligado a la etapa de identificación. El mismo consiste en una propuesta para la descomposición del estadístico de Hotelling (T^2) en el espacio original de las mediciones como una suma única de J términos. Esta metodología, llamada OSS (*Original Space Strategy*), permite además una interpretación sencilla de los valores de las contribuciones $c_{j,k}^{T^2}$. El proceso se monitorea en el espacio original de las mediciones siendo T^2 la única métrica necesaria para el análisis. Los resultados de la aplicación de OSS mostraron un muy buen desempeño en lo referente a las velocidades de detección cuando se lo compara con el obtenido mediante NM , $WKFH-AT$ o

MICA. La interpretación de las $c_{j,k}^{T^2}$ resulta mucho más clara y precisa en la gran mayoría de los casos en los que T^2 detecta la existencia de la falla. Las principales ventajas de este enfoque son las siguientes: evita las ambigüedades que pueden resultar cuando se consideran las *CVs* de más de una métrica, es aplicable en el espacio original de mediciones y también en los espacios latentes, tiene asociada una interpretación geométrica sencilla y permite una interpretación física de las contribuciones negativas. Cabe aclarar que las ventajas asociadas a la interpretación geométrica de la metodología *OSS* se mantendrán incluso cuando esta descomposición se aplique en algún espacio de *VLs*.

La estrategia *OSS* mostró también un muy buen desempeño cuando se la aplicó para la detección de errores sistemáticos en las mediciones. Los resultados del empleo de *OSS* a sistemas continuos, en los que se simuló errores sistemáticos en las observaciones, mostraron las bondades de esta descomposición para su aplicación en ese contexto.

El principal inconveniente asociado a esta metodología está relacionado a la cantidad de información disponible para la elaboración de los modelos. Aquellos sistemas en los que se cuenta con un gran número de descriptores necesitarán también un número grande de observaciones, a fin de obtener estimaciones apropiadas de las matrices de covarianza o correlación y evitar así problemas de mal condicionamiento. La aplicación de *OSS*, en su estado actual, está restringida a problemas en los que se dispone de una cantidad suficiente de datos y de conocimiento del proceso tal que las matrices de correlación resulten bien condicionadas.

7.2 Trabajos Futuros

En cuanto a los futuros trabajos a realizarse en esta área pueden mencionarse los siguientes:

El análisis de las relaciones entre los estadísticos T^2 , D , T_H^2 y SPE es un punto que aún debe ser explorado de manera rigurosa. La determinación y delimitación de posibles zonas grises en la etapa de detección debidas a la reducción de las dimensiones del espacio original, resulta de sumo interés dado que brindará información relacionada a la existencia de falsas alarmas y fallas en las detecciones. Asimismo, es importante analizar formas alternativas para el cálculo de los límites de las CVs con el objeto de mejorar el desempeño de las diferentes técnicas en la etapa de identificación.

Respecto a OSS en particular, es necesario analizar algunos puntos clave con el objeto de poder extender esta metodología a problemas de mediana y gran escala. Los problemas asociados con el número de condición de las matrices de covarianza en procesos con muchas variables medidas, mencionados anteriormente, se abordarán mediante el empleo de técnicas de regularización de matrices. Estos procedimientos permitirían, en principio, mejorar los números de condición de estas matrices, haciendo posible el análisis en el espacio original de las observaciones sin necesidad de aplicar técnicas de proyección.

Se requiere un análisis exhaustivo de otros métodos de cálculo de las CVs desarrollados durante este trabajo de tesis. Estos se basan en el conocimiento del vecino en control más cercano ($NICN$). Esta información permite tener una idea de cuál es la desviación mínima que debe ser explicada para poder interpretar la falla. Los resultados preliminares obtenidos a partir de la aplicación de esta estrategia son muy prometedores.

La ampliación de las capacidades de éstas y otras técnicas de identificación mediante el uso de metodologías de análisis exploratorio tales como *biplots* pueden resultar en una mejora significativa relacionada con la facilidad de interpretación del estado del proceso, la cual será analizada en el futuro.

El análisis de las similitudes existentes entre las metodologías de identificación usadas en *SPC* y las empleadas en el área de tratamiento de errores gruesos en las observaciones está aún poco desarrollado. La vinculación de estas herramientas proveería un marco común para la determinación de situaciones anormales ocasionadas tanto por un cambio real en las variables de proceso como por la existencia de errores sistemáticos en las mediciones.

Página en Blanco

Apéndice 1

A1.1 Determinación del Número de Componentes a Retener

a. Criterio del Porcentaje de Varianza Acumulado

El criterio del porcentaje de varianza acumulado (Valle y col., 1999) se fundamenta en el índice de calidad de representación introducido en la ecuación (2.17). En éste se considera que mientras más alto sea el porcentaje de la varianza total reconstruido por el modelo, mejor será la calidad del mismo. Sin embargo, como se explicó en la Sección 2.2.1, la proporción de la varianza explicada por cada componente adicional decrece con el número de componentes considerado. Debido a esto, el intento de reconstruir porcentajes muy altos de varianza tales como el 90 ó 95 % llevará casi inevitablemente al uso de una gran cantidad de *CPs*. En consecuencia, la decisión final del valor adecuado para R resultará una solución de compromiso entre dicho porcentaje y la cantidad de *CPs* consideradas, dado que el valor mínimo aceptado para la reconstrucción de la varianza es en sí una decisión subjetiva.

b. SCREE Test. Porcentaje de Varianza Residual

El método *SCREE* fue propuesto por Cattell (1966). Este método se basa en el análisis de la diferencia existente entre las magnitudes de los autovalores de las matrices de covarianza o correlación. Dado que los últimos autovalores son por lo general muy pequeños comparados con los primeros, es posible determinar de manera gráfica cuáles de ellos resultan importantes y cuáles podrían ser descartados.

El criterio propuesto por Cattell sugiere conservar sólo las primeras R CPs , incluyendo hasta aquella en la cual puede detectarse un quiebre en la tendencia de los autovalores. Este quiebre (llamado codo o rodilla en buena parte de la literatura) indica cuándo las CPs dejan de contener información relevante y comienzan a incluir variaciones aleatorias (Jackson, 1991).

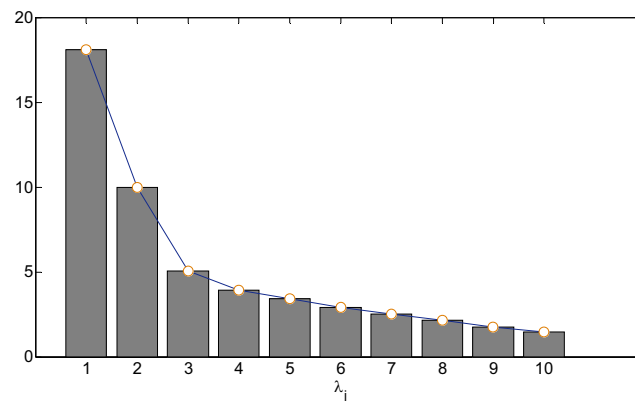


Figura A1.1: Gráfico típico de autovalores de una matriz de covarianza

La Figura A1.1 muestra un gráfico típico de autovalores donde existen grandes diferencias entre los primeros y los últimos valores de la serie. Según el criterio de Cattell, cuatro CPs ($R = 4$) serían suficientes para representar el sistema.

Entre las desventajas principales que muestra este método, pueden mencionarse las siguientes:

- a. No existen garantías de que se produzcan quiebres visibles en la tendencia de los autovalores.
- b. En ciertos casos puede existir más de un quiebre. En algunos sistemas es posible asociar este tipo de comportamiento a cambios sistemáticos en la fuente de variación de los datos (Jackson, 1991). El criterio y conocimiento del problema por parte del usuario es crucial en este tipo de situaciones.

c. Diferencias muy grandes entre los primeros autovalores (de órdenes de magnitud) pueden causar errores al momento de determinar la ubicación exacta del quiebre. En este caso es recomendable el uso de escalas logarítmicas en los gráficos de autovalores.

A pesar de los inconvenientes mencionados, el *SCREE* test se usa ampliamente debido a su simplicidad, habiéndose propuesto varias modificaciones para mejorar su desempeño.

c. *Autovalor Promedio*

Otra técnica frecuentemente utilizada es la comparación de cada λ_r ($r=1,2,\dots,J$) contra el promedio de los J autovalores. En este caso las *CPs* cuyos autovalores asociados sean mayores que el promedio serán retenidas para el análisis.

El valor promedio de los autovalores se calcula como:

$$\bar{\lambda} = \frac{tr(\bar{\Lambda})}{J}. \quad (\text{A1.1})$$

En el caso considerado en la Figura A1.1, $\bar{\lambda} = 5,13$ y sólo dos *CPs* ($R=2$) serían retenidas.

Si el modelo se construye a partir de la matriz de correlación, la $tr(\bar{\Lambda}) = J$ por lo que todas las componentes cuyos autovalores asociados cumplan que $\lambda_r \geq 1$ serán retenidas.

d. *Análisis Paralelo*

El análisis paralelo consiste básicamente en la aplicación del *SCREE* test a dos conjuntos de datos diferentes: la matriz de datos \mathbf{X} y otra especialmente formulada para este análisis. Esta última debe ser de las mismas dimensiones que \mathbf{X} , y contener los datos generados

aleatoriamente con una distribución normal estandarizada y cuyas variables no estén correlacionadas.

A continuación, se grafican los autovalores de las matrices de correlación para los dos conjuntos de datos. Se considera que los autovalores que quedan por encima del punto de cruce de ambas curvas son aquellos que contienen la información más significativa mientras que los restantes contienen mayoritariamente ruido aleatorio.

Valle *y col.* (1999) señalaron que cuando el número de observaciones es finito, los primeros autovalores de la matriz auxiliar resultan mayores que 1 mientras que los últimos están por debajo de la unidad. Sin embargo, cuando estas matrices son generadas para un número de observaciones tendiendo a infinito, todos los autovalores de las matrices de correlación se aproximan a la unidad, por lo que el método del análisis paralelo se vuelve equivalente al del promedio de los autovalores.

e. Criterio de Autocorrelación

Shrager y Hendler (1982) propusieron el uso del índice de autocorrelación para determinar cuáles son las *CPs* que contienen información relevante y cuáles están compuestas mayoritariamente por variaciones aleatorias. El índice de autocorrelación se define a partir de los componentes de la matriz de *SCORES*, de la siguiente manera:

$$AC_r = \sum_{i=1}^{I-1} t_{i,r} t_{i+1,r}. \quad (A1.2)$$

Si el índice de autocorrelación para la *r*-ésima *CP* (AC_r) es mayor que un valor predeterminado (usualmente 0.5), se considera que dicha componente posee información significativa para el análisis. Por el contrario, si AC_r es menor que dicho umbral, se asume que el comportamiento es suficientemente oscilatorio y se atribuye al ruido aleatorio.

Esta técnica presenta dos desventajas fundamentales (Valle y col., 1999). La primera es que el umbral para la correlación es un número impuesto en forma subjetiva por el usuario. La segunda es que las *CPs* con valores de varianza grandes (las que podrían contener información importante) pueden tener bajos niveles de autocorrelación y podrían ser eliminadas de forma errónea.

f. Criterio de Validación Cruzada

El método de validación cruzada es quizá el más utilizado para la selección del número de *CPs* a retener. En esta estrategia, el conjunto total de observaciones (I) se divide aleatoriamente en G subgrupos de I_g observaciones. Luego, se construyen G modelos diferentes, en los cuales sólo se considera la primera *CP*. Cada modelo deja fuera del estudio un subgrupo diferente de observaciones. Este procedimiento se repite tomando las primeras dos *CPs* y así sucesivamente hasta incluir las J componentes que expanden el espacio completo.

Los errores de representación para cada uno de los individuos en el grupo excluido se calculan en cada uno de los pasos. El promedio de los errores cuadráticos de representación ($PRESS_r$) para las I observaciones, cuando el modelo considera r componentes, se calcula como:

$$PRESS_r = \frac{I}{IJ} \sum_{g=1}^G \sum_{i=1}^{I_g} (\hat{\mathbf{x}}_i - \mathbf{x}_i)^T (\hat{\mathbf{x}}_i - \mathbf{x}_i). \quad (\text{A1.3})$$

La decisión de si debe o no adicionarse una nueva componente al modelo se toma en base al valor del estadístico \mathcal{W} (Jackson, 1991).

$$\mathcal{W} = \frac{(PRESS_{r-1} - PRESS_r) / D_M}{PRESS_r / D_R}, \quad (\text{A1.4})$$

$$D_M = I + J - 2r, \quad (A1.5)$$

$$D_R = J(I - I) - \sum_{m=1}^r (I + J - 2m). \quad (A1.6)$$

Si $\mathcal{W} > 1$ la *CP* correspondiente se debe retener para el análisis. Es posible que si se siguieran analizando los resultados luego de la primera ocurrencia de $\mathcal{W} < 1$, algunos de los valores posteriores de \mathcal{W} vuelvan a estar por encima de la unidad.

Otro criterio basado sólo en los valores de $PRESS_r$ fue propuesto por Wold (1978). En este caso el procedimiento es similar al descrito pero en lugar de usar el estadístico \mathcal{W} como criterio de corte, se busca el valor de r para el cual se observa un mínimo local en el valor de $PRESS_r$.

En la Tabla A1.1 se incluye un conjunto de datos reportados por Jackson (1991), que se utilizarán para ilustrar el desempeño de las técnicas descritas. Estos datos comprenden los resultados de estudios de audiometría realizados a una población de 100 individuos, y consisten en los valores de mediciones efectuadas en cada oído para cuatro frecuencias diferentes. La Tabla A1.2 muestra los valores de R que se obtienen al aplicar cada uno de los métodos descritos anteriormente al conjunto de datos mencionado.

Los resultados de la Tabla A1.2 revelan la existencia de discrepancias importantes en el número resultante de componentes a retener. Para algunos criterios como el del autovalor promedio, autocorrelación y el análisis paralelo solo dos *CPs* capturan la variabilidad más importante del sistema. El criterio del porcentaje de reconstrucción (para un % reconstrucción $> 90\%$) y el gráfico SCREE recomiendan el uso de 4 *CPs* mientras que el criterio de validación cruzada indica que el valor apropiado para R es cinco.

Tabla A1.1: Datos de un estudio de audiometría

	Oído Izquierdo				Oído Derecho					Oído Izquierdo				Oído Derecho			
	500	1000	2000	4000	500	1000	2000	4000		500	1000	2000	4000	500	1000	2000	4000
1	0	5	10	15	0	5	5	15	51	-10	-10	-10	5	-10	-10	-5	0
2	-5	0	-10	0	0	5	5	15	52	5	10	20	25	0	5	15	30
3	-5	0	15	15	0	0	10	25	53	-10	-10	5	25	-10	-10	-10	40
4	-5	0	-10	-10	-10	-5	-10	10	54	5	10	0	-10	0	5	-5	15
5	-5	-5	-10	10	0	-10	-10	50	55	15	20	10	60	20	20	0	25
6	5	5	5	-10	0	5	0	20	56	-10	-10	-10	5	-10	-10	-5	-10
7	0	0	0	20	5	5	5	10	57	-5	-5	-10	30	0	-5	-10	15
8	-10	-10	-10	-5	-10	-5	0	5	58	-5	-5	0	5	-5	-5	0	10
9	0	0	0	40	0	0	-10	10	59	-5	5	5	40	0	0	0	60
10	-5	-5	-10	20	-10	-5	-10	15	60	5	10	30	20	5	5	20	10
11	-10	-5	-5	5	5	0	-10	5	61	5	5	0	10	-5	5	0	10
12	5	5	10	25	-5	-5	5	15	62	0	5	10	35	0	0	5	20
13	0	0	-10	15	-10	-10	-10	10	63	-10	-10	-10	0	-5	0	-5	0
14	5	15	5	60	5	5	0	50	64	-10	-5	-5	20	-10	-10	-5	5
15	5	0	5	15	5	-5	0	25	65	5	10	0	25	5	5	0	15
16	-5	-5	5	30	5	5	5	25	66	-10	0	5	60	-10	-5	0	65
17	0	-10	0	20	0	-10	-10	25	67	5	10	40	55	0	5	30	40
18	5	0	0	50	10	10	5	65	68	-5	-10	-10	20	-5	-10	-10	15
19	-10	0	0	15	-10	-5	5	15	69	-5	-5	-5	20	-5	0	0	0
20	-10	-10	-5	0	-10	-5	-5	5	70	-5	-5	-5	5	-5	0	0	5
21	-5	-5	-5	35	-5	-5	-10	20	71	0	10	40	60	-5	0	25	50
22	5	15	5	20	5	5	5	25	72	-5	-5	-5	-5	-5	-5	-5	5
23	-10	-10	-10	25	-5	-10	-10	25	73	0	5	45	50	0	10	15	50
24	-10	0	5	15	-10	-5	5	20	74	-5	-5	10	25	-10	-5	25	60
25	0	0	0	20	-5	-5	10	30	75	0	-10	0	60	15	0	5	50
26	-10	-5	0	15	0	0	0	10	76	-5	0	10	35	-10	0	0	15
27	0	0	5	50	5	0	5	40	77	5	0	0	15	0	5	5	25
28	-5	-5	-5	55	-5	5	10	70	78	15	15	5	35	10	15	-5	0
29	0	15	0	20	10	-5	0	10	79	-10	-10	-10	5	-5	-5	-5	5
30	-10	-5	0	15	-5	0	10	20	80	-10	-10	-5	15	-10	-10	-5	5
31	-10	-10	5	10	0	0	20	10	81	0	-5	5	35	-5	-5	5	15
32	-5	5	10	25	-5	0	5	10	82	-5	-5	-5	10	-5	-5	-5	5
33	0	5	0	10	-10	0	0	0	83	-5	-5	-10	-10	0	-5	-10	0
34	-10	-10	-10	45	-10	-10	5	45	84	5	10	10	20	-5	0	0	10
35	-5	10	20	45	-5	10	35	60	85	-10	-10	-10	5	-10	-5	-10	20
36	-5	-5	-5	30	-5	0	10	40	86	5	5	10	0	0	5	5	5
37	-10	-5	-5	45	-10	-5	-5	50	87	-10	0	-5	-10	-10	0	0	-10
38	5	5	5	25	-5	-5	5	40	88	-10	-10	10	15	0	0	5	15
39	-10	-10	-10	0	-10	-10	-10	5	89	-5	0	10	25	-5	0	5	10
40	10	20	15	10	25	20	10	20	90	5	0	-10	-10	10	0	0	0
41	-10	-10	-10	20	-10	-10	0	5	91	0	0	5	15	5	0	0	5
42	5	5	-5	40	5	10	0	45	92	-5	0	-5	0	-5	-5	-10	0
43	-10	0	10	20	-10	0	15	10	93	-5	5	-10	45	-5	0	-5	25
44	-10	-10	10	10	-10	-10	5	0	94	-10	-5	0	10	-10	5	-10	10
45	-5	-5	-10	35	-5	0	-10	55	95	-10	-5	0	5	-10	-5	-5	5
46	5	5	10	25	10	5	5	20	96	5	0	5	0	5	0	5	15
47	5	0	10	70	-5	5	15	40	97	-10	-10	5	40	-10	-5	-10	5
48	5	10	0	15	5	10	0	30	98	10	10	15	55	0	0	5	75
49	-5	-5	5	-10	-10	-5	0	20	99	-5	5	5	20	-5	5	5	40
50	-5	0	10	55	-10	0	5	50	100	-5	-5	-10	-10	-5	0	15	10

Tabla A1.2: Valores de *R* sugeridos por los diferentes criterios para los datos de la Tabla A1.1

Método	Porcentaje de Reconstrucción	SCREE	Autovalor Promedio	Análisis Paralelo	Autocorrelación	Validación Cruzada
<i>R</i> =	4	4	2	2	2	5

A1.2 Referencias

- Cattell, R. B., "The Scree Test for the Number of Factors". *Multivariate Behavioural Research* **1966**, 1, (2), 245-276.
- Jackson, J. E. (1991). A User's Guide to Principal Components. Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics, New York, John Wiley & Sons, Inc.
- Shrager, R. I. y R. W. Hendler, "Titration of Individual Components in a Mixture with Resolution of Difference Spectra, PKS, and Redox Transitions". *Analytical Chemistry* **1982**, 54, (7), 1147-1152.
- Valle, S., W. Li y S. J. Qin, "Selection of the Number of Principal Components: The Variance of the Reconstruction Error Criterion with a Comparison to Other Methods". *Industrial and Engineering Chemistry Research* **1999**, 38, (11), 4389-4401.
- Wold, S., "Cross Validatory Estimation of the Number of Components in Factor and Principal Components Analysis". *Technometrics* **1978**, 20, (4), 397-406.

Apéndice 2

A2.1 Test de Hipótesis

El test de hipótesis es una herramienta fundamental de la estadística ya que permite al usuario sacar conclusiones tales como si un individuo puede o no considerarse como parte de una población, o si dos individuos pueden o no considerarse diferentes entre sí.

La hipótesis que se desea testear se denomina hipótesis nula H_0 . El término “nula” indica que ésta es la hipótesis que se mantendrá en el caso de no encontrar evidencia suficiente de que la misma no es cierta. Un punto importante a tener en cuenta es que la hipótesis nula nunca es tomada como cierta, sino que se concluye que existe o no evidencia de que no lo es (para un dado nivel de confianza).

En el enfoque de contraste de hipótesis que se plantea en los tests, se considera la hipótesis alternativa H_1 que puede o no estar enunciada en forma explícita. La hipótesis alternativa no siempre es la negativa de H_0 sino que puede, en algunos casos, definirse en forma más específica.

La estructura básica de un test de hipótesis se muestra en la ecuación (A2.1).

$$\begin{array}{cc} \text{(a)} \left\{ \begin{array}{l} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{array} \right. & \text{(b)} \left\{ \begin{array}{l} H_0 : \theta = \theta_0 \\ H_1 : \theta = \theta_1 \end{array} \right. \end{array} \quad (\text{A2.1})$$

En el contraste (a), planteado en la ecuación (A2.1), sólo se intenta testear si el valor de θ puede o no considerarse igual a θ_0 . En el (b) se desea determinar si θ es igual a θ_0 o a θ_1 .

Las suposiciones y simplificaciones a los que cada contraste está sujeto dependerán del tipo de comparación planteada y de las características de las poblaciones consideradas. Las decisiones se toman en términos del valor de algún estadístico que permite aceptar o no H_0 para un dado nivel de confianza.

A2.2 Errores *Tipo I* y *Tipo II*

Cada vez que se plantea un test de hipótesis se debe estar conciente de que la aceptación o no de H_0 tiene asociada en forma intrínseca una dada probabilidad de cometer un error en dicha decisión.

Cuando se realiza un test de hipótesis existen cuatro situaciones posibles:

- a. H_0 es verdadera y es aceptada.
- b. H_0 es verdadera pero es rechazada.
- c. H_0 es falsa y es rechazada.
- d. H_0 es falsa pero es aceptada.

En los casos *a* y *c* no se comete error en la aceptación o rechazo de H_0 . Por otra parte, el error cometido en la situación indicada en el punto *b* es conocido como *Error Tipo I* mientras que el descrito en el punto *d* se conoce como *Error Tipo II*.

La probabilidad de cometer un *Error Tipo I* se identifica con la letra griega α y se conoce como nivel de significación del test, mientras que la probabilidad de incurrir en un *Error Tipo II* se indica con la letra griega β . La cantidad $(1 - \beta)$ se conoce como potencia del test y es la probabilidad de escoger H_1 cuando ésta es cierta $P(\text{escoger } H_1 | H_1 \text{ es cierta})$.

La probabilidad de *Error Tipo I* (α), es un parámetro definido por el usuario y toma normalmente valores del 10%, 5%, 2% o 1%. La elección del valor de α está relacionada al riesgo que el analista está dispuesto a correr al rechazar H_0 .

En la Figura A2.1 se observan las funciones de densidad de probabilidad (*pdf*) correspondientes a ambas hipótesis. El valor p_α indica el límite de valores muestrales hasta el cual H_0 será aceptada para un dado nivel de significación α . El área α corresponde a la probabilidad acumulada desde $p_\alpha \rightarrow \infty$. Por el contrario el área β muestra la probabilidad de aceptar H_0 cuando H_1 es verdadera

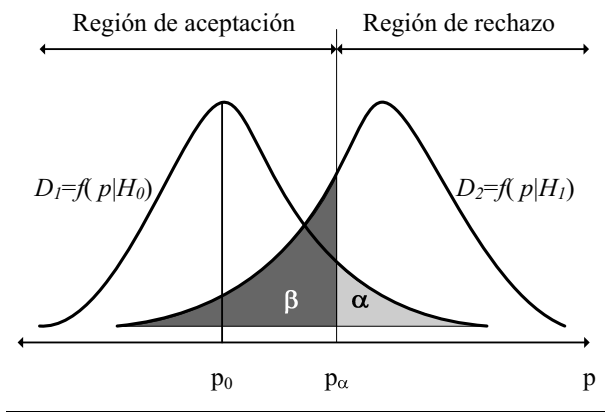


Figura A2.1: Interpretación gráfica de un test de hipótesis

Aunque H_1 estuviese definida en forma explícita y se estuviese evaluando una hipótesis alternativa particular, los valores de α y β no pueden estimarse de manera simultánea. No obstante, puede observarse que mientras menores son los valores de α , mayor serán los de la probabilidad β . Debido a la imposibilidad de conocer simultáneamente α y β , la potencia de los test de hipótesis se estima en la práctica en función de los datos contenidos en las poblaciones utilizadas como referencia, los cuales fueron previamente clasificados.

Página en Blanco

Apéndice 3

A3.1 Estimación de Funciones de Densidad de Probabilidad. Método del Kernel (KDE)

El método *Kernel Density Estimation (KDE)* es un procedimiento no paramétrico de estimación de las funciones de densidad de probabilidad. Dado un cierto conjunto de datos (\mathbf{X}), la función de Densidad de Probabilidad $f(x)$ que describe su estructura puede estimarse según:

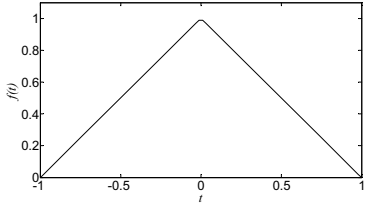
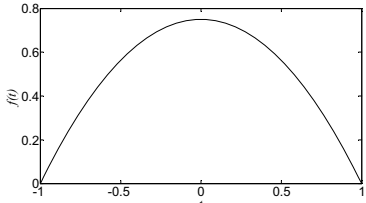
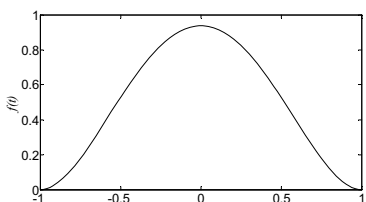
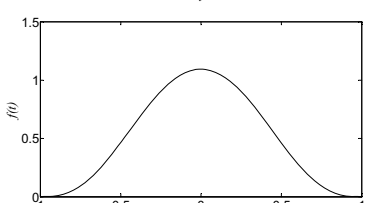
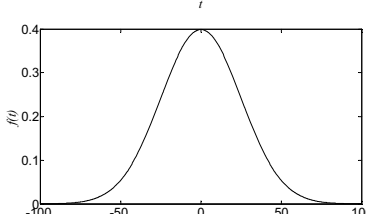
$$\hat{f}(x) = \frac{1}{Ih} \sum_{i=1}^I \mathcal{K} \left\{ \frac{x-x_i}{h} \right\}, \quad (\text{A3.1})$$

donde I es el número de elementos x_i en el conjunto de datos analizado (\mathbf{X}), x es el valor de la variable en el cual quiere estimarse $f(x)$, h es un parámetro llamado ancho de ventana el cual actúa como un factor de peso o suavizado y $\mathcal{K}\{t\}$ es lo que se conoce como *kernel*, debiendo cumplir con la condición:

$$\int_{-\infty}^{\infty} \mathcal{K}\{t\} dt = 1. \quad (\text{A3.2})$$

Frecuentemente la función *kernel* es una función simétrica, siendo la más utilizada la distribución de probabilidad normal estándar. Sin embargo, resulta válido utilizar otros tipos de funciones tales como triangulares o Epanechnikov, entre otras (Martinez y Martinez, 2002; Bowman y Azzalini, 1997). En la Tabla A3.1 se resumen las características de algunas funciones generalmente usadas como *kernel*.

Tabla A3.1: Distribuciones usadas como *kernel* en *KDE*

Distribución	
Triángular	$\mathcal{K}(t) = (1 - t) \quad -1 \leq t \leq 1$ 
Epanechnikov	$\mathcal{K}(t) = \frac{3}{4}(1 - t^2) \quad -1 \leq t \leq 1$ 
Biweight	$\mathcal{K}(t) = \frac{15}{16}(1 - t^2)^2 \quad -1 \leq t \leq 1$ 
Triweight	$\mathcal{K}(t) = \frac{35}{32}(1 - t^2)^3 \quad -1 \leq t \leq 1$ 
Normal	$\mathcal{K}(t) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{t^2}{2}\right\} \quad -\infty < t < \infty$ 

Según la ecuación (A3.1), la estimación de la *pdf* de \mathbf{X} se obtiene mediante la suma ponderada de las funciones simétricas elegidas como *kernel* centradas alrededor de cada uno de los I puntos considerados (x_i). Dado que resultan de una combinación lineal de las $\mathcal{K}\{t\}$, los estimadores $\hat{f}(t)$ heredan las propiedades de las mismas en lo que hace a continuidad y derivabilidad.

La Figura A3.1 muestra la estimación de la función $\hat{f}(x)$ para un conjunto de 15 puntos normalmente distribuidos con media nula y desvío estándar unitario usando la distribución normal estándar como *kernel*.

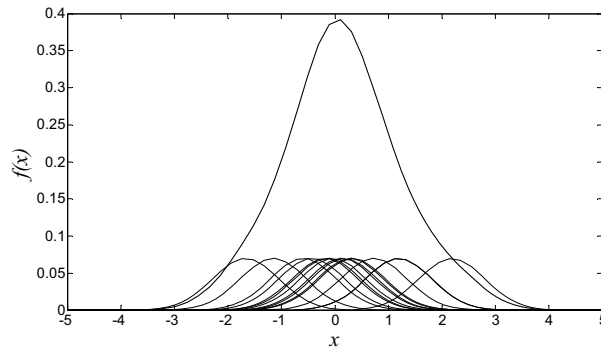


Figura A3.1: Estimación de $f(x)$ basada en un conjunto de datos $\mathbf{x} \approx \mathcal{N}(0;1)$

La selección del valor para el parámetro h es central en esta técnica. La Figura A3.2 muestra las estimaciones de $f(x)$ obtenidas para el mismo conjunto de datos usado en la Figura A3.1 para distintos valores de h . Valores bajos de h darán lugar a estimaciones poco definidas las cuales tendrán una gran cantidad de picos (Figura A3.2 (a)). Por otro lado, valores altos para h llevarán a la obtención de curvas suaves pero que pueden, en algunos casos, disimular la presencia de modos de menor importancia en la distribución.

La determinación de los valores apropiados para h puede hacerse mediante la minimización de la función *AMISE* (*Asymptotic Mean Integrated Square Error*), la cual dependerá del tipo de función elegida como *kernel* (Martinez y Martinez, 2002). Por ejemplo, el valor que minimiza la *AMISE*, cuando se elige la distribución de probabilidad normal estándar como *kernel*, es:

$$h = \left(\frac{4}{3}\right)^{1/5} \sigma I^{-1/5} \approx 1,06 \sigma I^{-1/5}, \quad (\text{A3.3})$$

siendo σ el desvío estándar del vector \mathbf{x} .

Los pasos requeridos para la implementación de *KDE* se muestran en la Tabla A3.2 (Martinez y Martinez, 2002). Para distribuciones que sólo tienen valores positivos, como las χ^2 ó \mathcal{F} por ejemplo, los valores de \mathbf{X} se transforman usando logaritmos.

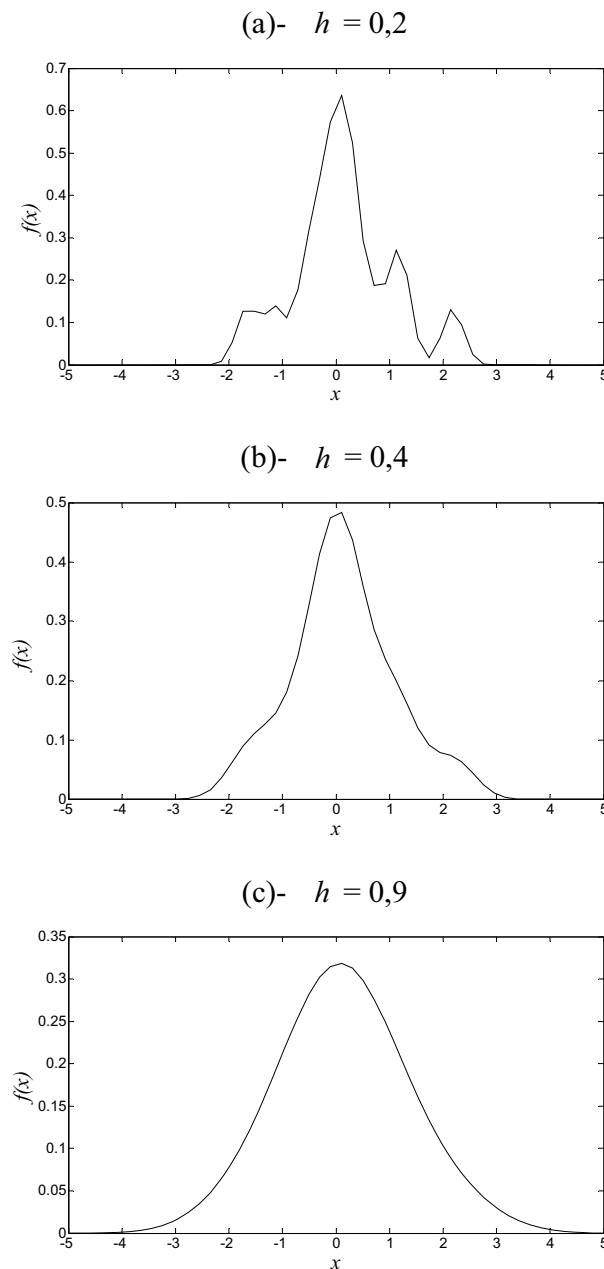


Figura A3.2: Estimaciones de $f(x)$ usando distintos valores de h

Tabla A3.2: Algoritmo de cálculo de $\hat{f}(x)$ usando *KDE*

-
1. Seleccionar la función que se utilizará como *kernel* y un valor apropiado para el ancho de ventana h . Definir el dominio sobre el cual se realizará la estimación de $f(x)$ y establecer un conjunto de puntos de estimación de dicho dominio (\mathbf{Y}).
 2. Para cada uno de los puntos x_i en \mathbf{X} evaluar el valor de la función $\mathcal{K}_i(y)$ en todos los puntos y_i de \mathbf{Y} .

$$\mathcal{K}_i(y) = \mathcal{K} \left\{ \frac{y - x_i}{h} \right\}; \quad i = 1, 2, \dots, I.$$

Como resultado de ésto se obtendrá un conjunto de I curvas centradas en cada uno de los puntos muestrales x_i .

3. Ponderar cada una de estas curvas por el factor $1/h$.
 4. Calcular el valor de $\hat{f}(y)$ para cada y_i como el promedio de las I curvas obtenidas en el paso 3.
-

A3.2 Referencias

- Bowman, A. W. y A. Azzalini (1997). Applied Smoothing Techniques for Data Analysis, Oxford, Oxford University Press.
- Martinez, W. L. y A. R. Martinez (2002). Computational Statistics Handbook with Matlab, London, Chapman & Hall/CRC

Página en Blanco

Apéndice 4

A4.1 Modelo Matemático del Reactor Tubular Empleado en el Capítulo 4

El modelo del reactor tubular considerado en el Capítulo 4, corresponde a un caso de estudio incluido en las librerías que acompañan a gPROMS®. El mismo consiste en un reactor tubular (de un solo tubo) en el cual ocurre difusión radial y también axial. En él se lleva a cabo una única reacción $A + B \rightarrow 3C$. El reactor se refrigera por medio de una camisa concéntrica tal como se muestra en la Figura A4.1.

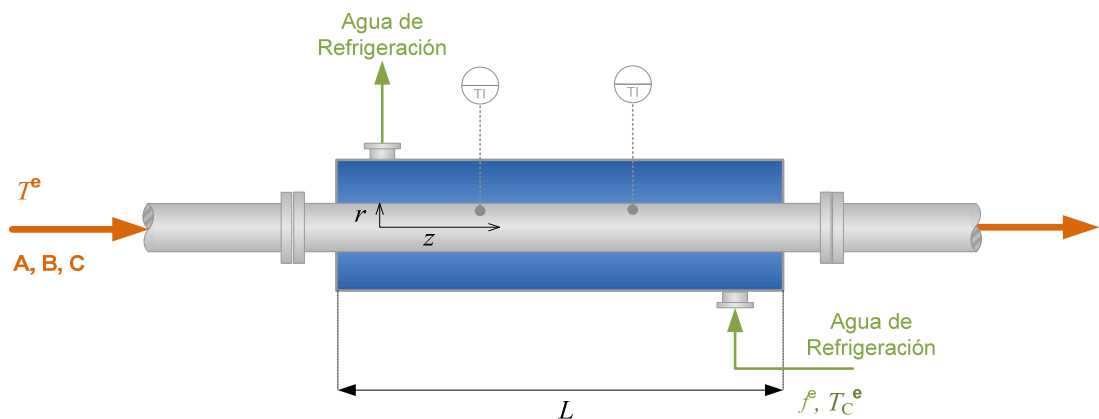


Figura A4.1: Esquema del reactor considerado como caso de estudio

Los balances de masa para cada uno de los 3 componentes del modelo se muestran en la siguiente expresión:

$$\frac{\partial C_i}{\partial t} = -v \frac{\partial C_i}{\partial z} + D_z \frac{\partial^2 C_i}{\partial z^2} + \frac{D_r}{r} \frac{\partial C_i}{\partial r} \left(r \frac{\partial C_i}{\partial r} \right) + \sum_{j=1}^J v_j r \quad i = A, B \text{ o } C; z \in (0, L); r \in (0, R), \quad (\text{A4.1})$$

donde z y r son los valores de las coordenadas axiales y radiales respectivamente, v es la velocidad lineal del flujo en la dirección z , C_i es la concentración de cada uno de los

componentes considerados, D_z y D_r son los coeficientes de dispersión axial y radial respectivamente, r es la velocidad de reacción y ν_i es el coeficiente estequiométrico de cada uno de los componentes.

El balance de energía en el cilindro central se formula como:

$$\rho C_p \frac{\partial T}{\partial t} = -v \frac{\partial T}{\partial z} + k_z \frac{\partial^2 C_i}{\partial z^2} + \frac{k_r}{r} \frac{\partial T}{\partial r} \left(r \frac{\partial T}{\partial r} \right) + \Delta H r, \quad (\text{A4.2})$$

donde ρ y C_p son la densidad y la capacidad calorífica de la mezcla reactiva, T es la temperatura de la mezcla reactiva, k_z y k_r son los coeficientes de conductividad térmica en el sentido axial y radial, respectivamente, y ΔH es la entalpía de reacción.

El balance de energía en la camisa se formula como:

$$\rho_c V C_{p,c} \frac{\partial T_c}{\partial t} = f_c C_{p,c} (T_c^e - T_c) + Q, \quad (\text{A4.3})$$

siendo:

$$Q = US \int_0^L (T(z, R) - T_c) dz, \quad (\text{A4.4})$$

donde ρ_c y $C_{p,c}$ son la densidad y la capacidad calorífica del fluido que circula por la camisa, f_c es el caudal de refrigerante, V es el volumen de la camisa, T_c es la temperatura en la camisa (considerada uniforme), T_c^e es la temperatura de la corriente de entrada de refrigerante, Q es el calor intercambiado entre el cilindro interior y la camisa, U es el coeficiente de transferencia de calor y S es la sección longitudinal de intercambio.

La velocidad de reacción se calcula con la siguiente expresión:

$$r = A \exp\left\{-\frac{E}{R_g T}\right\} C_A C_B, \quad (\text{A4.5})$$

donde A y E son el factor pre-exponencial y la energía de activación de la constante cinética de la reacción, respectivamente y R_g es la Constante Universal de los gases.

Finalmente, las condiciones de borde resultan:

$$-k_r \frac{\partial T}{\partial r} \Big|_{r=R} = U_h (T - T_C), \quad \forall z \in (0, L), \quad (\text{A4.6})$$

Los valores y unidades de los parámetros utilizados en este trabajo se detallan en la Tabla A4.1:

Tabla A4.1: Parámetros del caso de estudio

Parámetro	Valor	Unidades
v_1, v_2	-1	-
v_3	3	-
v	8,8	m/s
L	30	m
R	0,5	in
ρ	1,293	kg/m ³
ρ_C	2000	kg/m ³
C_p	992	J/kg K
$C_{p,C}$	123900	J/kg K
D_z	0,0001	m ² /s
D_r	0,01	m ² /s
D_z	0,0001	m ² /s
k_r	0,5	W/m K
k_z	0,05	W/m K
ΔH	1200000	J/mol
E	113370,0	
A	11,45E-3	
R_g	8,314	J/mol K
U	9,6	W/m ² K
U_h	96,0	W/m ² K

Los valores de las condiciones de entrada para las 37 observaciones usadas para construir la *NOC* se detallan en la Tabla A4.2.

Tabla A4.2: Variables de entrada para la simulación de las observaciones que constituyen la *NOC*

	C_1^e [mol/m ³]	C_2^e [mol/m ³]	$C_3^e \times 10^3$ [mol/m ³]	T^e [°C]	T_c^e [°C]	f_c^e [kg/s]
1	0,20	4,04	1,94	633,95	626,47	0,102
2	0,21	4,20	1,98	630,75	626,67	0,104
3	0,21	3,89	1,98	638,51	624,39	0,101
4	0,22	4,27	2,04	616,99	624,76	0,105
5	0,22	3,92	2,10	622,17	624,89	0,103
6	0,21	3,95	1,78	614,90	623,76	0,097
7	0,22	4,06	1,91	608,66	623,63	0,103
8	0,21	4,34	2,09	610,27	625,12	0,097
9	0,22	3,81	1,84	637,35	625,84	0,101
10	0,21	4,00	2,12	619,39	624,62	0,102
11	0,20	4,27	2,02	623,82	624,47	0,103
12	0,21	3,85	1,75	639,57	624,22	0,099
13	0,22	4,12	1,80	615,46	626,34	0,099
14	0,20	3,90	1,73	637,75	625,46	0,100
15	0,20	4,09	1,82	630,52	624,04	0,101
16	0,20	4,17	1,93	624,59	625,49	0,095
17	0,22	3,94	1,99	620,55	625,66	0,100
18	0,22	3,89	1,94	626,11	625,59	0,097
19	0,21	3,93	1,81	613,78	623,71	0,096
20	0,21	3,97	1,82	625,08	623,31	0,097
21	0,20	4,21	1,90	622,27	624,14	0,099
22	0,22	4,22	1,85	630,66	623,61	0,099
23	0,20	4,09	1,84	641,11	626,34	0,104
24	0,21	4,18	2,03	614,16	623,87	0,097
25	0,22	4,18	2,17	611,32	623,31	0,099
26	0,21	3,94	1,94	627,30	625,84	0,098
27	0,20	4,08	2,07	641,52	625,13	0,097
28	0,21	4,07	2,12	608,25	624,15	0,104
29	0,20	4,08	2,07	638,04	624,53	0,098
30	0,22	4,24	1,84	608,36	623,18	0,102
31	0,20	3,94	2,06	625,70	626,41	0,099
32	0,22	4,17	2,13	614,18	626,33	0,099
33	0,20	4,05	1,79	632,60	624,10	0,100
34	0,22	4,18	1,82	616,23	625,25	0,097
35	0,20	3,84	2,00	640,28	623,75	0,100
36	0,22	4,20	2,07	612,24	625,35	0,101
37	0,20	3,92	2,00	625,76	624,41	0,095

Apéndice 5

A5.1 Nomenclatura del Caso de Estudio

Tabla A5.1: Nomenclatura. Caso de estudio presentado en el Capítulo 5

Variable	Descripción
a	Argumento de la función de Bessel modificada del primer tipo
A	Área de transferencia de calor entre el reactor y la camisa
A_{k_p}	Factor pre-exponencial de la constante cinética de la reacción de propagación
A_{k_t}	Factor pre-exponencial de la constante cinética de la reacción de terminación global
$A_{k_{tr}}$	Factor pre-exponencial de la constante cinética de la reacción de transferencia efectiva
a_{ed}	Área superficial cubierta por una molécula de surfactante sobre las gotas de monómero
a_{em}	Área superficial cubierta por una molécula de surfactante sobre las micelas
a_{ep}	Área superficial cubierta por una molécula de surfactante sobre las partículas de polímero
b	Orden de la función de Bessel modificada del primer tipo
b_m	Constante de la correlación del efecto gel
c_m	Constante de la correlación del efecto gel
CMC	Abreviatura para concentración micelar crítica
$c_p^{\text{látex}}$	Capacidad calorífica del látex
c_p^m	Capacidad calorífica del monómero
c_p^w	Capacidad calorífica del agua
c_p^p	Capacidad calorífica del polímero
d	Velocidad de desorción de los radicales desde las partículas

Tabla A5.1: Nomenclatura... (continuación)

Variable	Descripción
D_m	Coeficiente de difusividad efectiva
e	Velocidad de entrada de los radicales en las partículas de polímero
E_{k_p}	Energía de activación de la constante cinética de la reacción de propagación
E_{k_t}	Energía de activación de la constante cinética de la reacción de terminación global
$E_{k_{tr}}$	Energía de activación de la constante cinética de la reacción de transferencia efectiva
f	Factor de eficiencia para la descomposición del iniciador
$F(t', t)$	Moles de partículas nacidas en t' por unidad de tiempo y unidad de volumen del reactor
f_c	Caudal másico de agua en la camisa
$F_i(t', t)$	Moles de partículas con i radicales nacidas en t' por unidad de tiempo y unidad de volumen del reactor
g_t	Factor de corrección por el efecto gel en k_t
I	Símbolo del iniciador o su concentración en la fase acuosa
$I_b(a)$	Función de Bessel modificada del primer tipo de orden b y argumento a
$\bar{i}(t', t)$	Número de radicales promedio en el tiempo t en el interior de cada partícula nacida en t'
j_{cr}	Longitud crítica (máxima) de cadena de los radicales en fase acuosa
J	Símbolo del agente inhibidor o su concentración
k_d	Constante de velocidad de descomposición del iniciador
k_j	Constante de velocidad de la reacción de inhibición
k_{mm}	Constante de velocidad de entrada de los radicales en las micelas
k_{mp}	Constante de velocidad de captura de radicales por parte de las partículas
k_p	Constante de velocidad de propagación
KPS	Abreviatura de persulfato de potasio
k_t	Constante de velocidad global de terminación
k_{tc}	Constante de velocidad de terminación por combinación

Tabla A5.1: Nomenclatura... (continuación)

Variable	Descripción
k_{td}	Constante de velocidad de terminación por desproporción
k_{tr}	Constante de velocidad de transferencia efectiva
k_{tm}	Constante de velocidad de transferencia al monómero
k_{tp}	Constante de velocidad de transferencia al polímero
k_{trt}	Constante de velocidad de transferencia a Tr
k_{tc}	Constante de velocidad de terminación por combinación
k_{td}	Constante de velocidad de terminación por desproporción
k_{tw}	Constante de terminación en fase acuosa
m	Concentración de micelas
M	Símbolo de monómero o su concentración en el reactor
M_d	Concentración de monómero en las gotas
Mn	Peso molecular promedio en número del polímero
Mw	Peso molecular promedio en peso del polímero
M_{sat}	Concentración de saturación del monómero en la fase acuosa
M_w	Concentración de monómero en la fase acuosa
MW_m	Peso molecular del monómero
MW_w	Peso molecular del agua
MW_s	Peso molecular del surfactante
N_A	Número de Avogadro
P_m	Velocidad de entrada de los radicales en las micelas
P_n	Polímero inactivo con n unidades monoméricas
Q_c	Calor transferido entre el reactor y la camisa
q_i^0	Caudal molar de entrada de iniciador
q_m^0	Caudal molar de alimentación de monómero
q_s^0	Caudal molar de alimentación de surfactante
q_w^0	Caudal molar de entrada de agua

Tabla A5.1: Nomenclatura... (continuación)

Variable	Descripción
r	Radio de las partículas
R	Concentración total de radicales en la fase acuosa
R'	Símbolo de radical de iniciación
r_d	Radio de las gotas de monómero
R_g	Constante Universal de los gases
R_n	Radical con n unidades monoméricas
r_{mic}	Radio de las micelas
r_{pol}	Velocidad de polimerización
$r_{t'}$	Radio de las partículas nacidas en t'
R_{Tr}	Radical en la reacción de transferencia a Tr
r_v	Velocidad de crecimiento de las partículas
S	Símbolo de surfactante o su concentración total
S_w	Concentración de surfactante en la fase acuosa y en las micelas cuando corresponda
S_{wc}	Concentración crítica para la formación de micelas
t	Tiempo
t'	Tiempo de nacimiento
T	Temperatura de la mezcla reactiva
T^0	Temperatura de la alimentación de monómero
T_C	Temperatura de salida del agua de camisa
T_{C0}	Temperatura de entrada del agua de camisa
Tr	Agente de transferencia de bajo peso molecular
T_{ref}	Temperatura de referencia
U	Coefficiente global de transferencia de calor
V	Volumen total de las partículas por unidad de volumen de látex
V_R	Volumen total del látex
VTP	Volumen total ocupado por las partículas de polímero
V_w	Fracción volumétrica de agua en el látex

Tabla A5.1: Nomenclatura... (continuación)

Variable	Descripción
W^C	Masa de agua contenida en la camisa
$W^{\text{látex}}$	Masa total de látex en el reactor
W^P	Masa total de polímero en el reactor
x	Conversión de monómero en el reactor
X	Especie inerte
x_{loc}	Conversión de equilibrio del monómero en las partículas
Letras griegas	
$\delta(t - t')$	Función delta de Dirac
ΔH_r	Calor de polimerización
φ	Fracción volumétrica de monómero en el interior de las partículas
φ_{sat}	Fracción volumétrica de equilibrio del monómero en el interior de las partículas
η	Orden de las velocidades de captura por parte de las partículas o de desorción desde las mismas respecto del radio de la partícula
λ_0	Momento de orden cero de la distribución de longitudes de los radicales
λ_1	Momento de primer orden de la distribución de longitudes de los radicales
λ_2	Momento de segundo orden de la distribución de longitudes de los radicales
μ_0	Momento de orden cero de la distribución de longitudes del polímero
μ_1	Momento de primer orden de la distribución de longitudes del polímero
μ_2	Momento de segundo orden de la distribución de longitudes del polímero
ρ_m	Densidad del monómero
ρ_p	Densidad del polímero
ρ_s	Densidad del surfactante
ρ_w	Densidad del agua
ψ	Parámetro de Flory-Huggins

A5.2 Parámetros del Modelo

Tabla A5.2: Valores de los parámetros del modelo

Variable	Valor	Unidades	Ref.
R_g	1,987	[cal mol ⁻¹ K ⁻¹]	†
N_A	$6,023 \times 10^{23}$	-	†
ρ_m	0,919	[g/cm ³]	†
ρ_p	1,19	[g/cm ³]	†
MW_m	100,13	[g/mol]	†
φ_{sat}	0,73	-	†
ψ	0,61332	-	†
η	2	-	†
k_{mm}	28	[cm/s]	†
k_{mp}	28	[cm/s]	†
D_m	$1,1 \times 10^{-7}$	[cm ² /s]	†
M_{sat}	$1,56 \times 10^4$	[mol/ cm ³]	†
c_p^m	0,430	[kcal kg ⁻¹ K ⁻¹]	§
c_p^p	1,6	[kJ kg ⁻¹ K ⁻¹]	‡
c_p^w	4,18	[kJ kg ⁻¹ K ⁻¹]	§
ΔH_r	-57,8	[kJ/mol]	†
S_{wC}	$1,73 \times 10^{-6}$	[mol/ cm ³]	†
a_{em}	50×10^{-16}	[cm ²]	†
r_{mic}	$2,5 \times 10^{-7}$	[cm]	†
k_d	$k_d = 1,8 \times 10^{17} \exp\{-34100/R_g T\}$	[s ⁻¹]	†
f	0,5	-	†
UA	2,5	[kJ s ⁻¹ K ⁻¹]	
E_{k_p}	4353	[cal mol ⁻¹]	†
A_{k_p}	$4,39 \times 10^8$	[l mol ⁻¹ s ⁻¹]	gEST
E_{k_t}	701	[cal mol ⁻¹]	†
A_{k_t}	$3,12 \times 10^{12}$	[l mol ⁻¹ s ⁻¹]	gEST
$E_{k_{tr}}$	4353	[cal mol ⁻¹]	†
$A_{k_{tr}}$	$2,45 \times 10^3$	[l mol ⁻¹ s ⁻¹]	gEST

‡ Brandrup y col. (1999)

§ Gao y Penlidis (2000)

† Rawlings (1985)

A5.3 Condiciones iniciales de los *Batches* de la Población de Referencia (*PR*)

En la Tabla A5.3 se muestran las condiciones iniciales utilizadas para simular los 146 batches con los cuales se construyó la *PR* en la *Etapa I* del monitoreo del proceso.

Tabla A5.3: Condiciones iniciales para los 146 *batches* utilizados para construir la *PR*

<i>Batch</i>	V_R [l]	V_W [-]	$S \times 10^5$ [mol/cm ³]	T [°C]	T_C [°C]	$I \times 10^6$ [mol/cm ³]	$M \times 10^3$ [mol/cm ³]	ρ_R [g/cm ³]
1	694,95	0,784	1,75	49,95	60,91	4,68	1,99	0,98
2	680,57	0,780	1,72	50,19	61,24	4,57	2,02	0,98
3	703,36	0,789	1,70	50,12	59,39	4,54	1,94	0,98
4	686,62	0,785	1,74	49,64	60,35	4,69	1,97	0,98
5	713,87	0,788	1,72	50,55	59,14	4,65	1,95	0,98
6	699,56	0,785	1,75	50,15	60,16	4,57	1,97	0,98
7	702,04	0,787	1,77	49,66	58,97	4,64	1,96	0,98
8	697,82	0,787	1,72	50,57	59,24	4,60	1,95	0,98
9	708,47	0,783	1,67	50,18	60,62	4,70	1,99	0,98
10	725,47	0,789	1,76	49,30	60,42	4,70	1,93	0,98
11	698,41	0,789	1,73	51,02	61,09	4,59	1,93	0,98
12	701,33	0,785	1,74	50,29	60,39	4,63	1,97	0,98
13	712,45	0,789	1,76	48,72	59,37	4,65	1,94	0,98
14	700,69	0,780	1,67	50,31	60,60	4,71	2,02	0,98
15	698,88	0,784	1,72	50,85	59,87	4,67	1,98	0,98
16	690,29	0,785	1,71	50,43	59,35	4,60	1,97	0,98
17	684,41	0,785	1,73	50,88	60,06	4,57	1,97	0,98
18	708,33	0,781	1,78	49,39	59,25	4,65	2,01	0,98
19	718,94	0,794	1,73	48,46	60,51	4,55	1,89	0,98
20	691,93	0,781	1,77	51,19	61,35	4,65	2,01	0,98
21	710,01	0,787	1,71	50,26	60,47	4,71	1,95	0,98
22	714,63	0,786	1,73	50,01	60,63	4,64	1,96	0,98
23	681,41	0,783	1,71	49,73	60,08	4,64	1,99	0,98
24	683,19	0,784	1,75	48,98	59,87	4,66	1,98	0,98
25	706,66	0,792	1,73	50,15	60,70	4,72	1,91	0,98
26	695,33	0,791	1,75	49,08	59,84	4,74	1,91	0,98
27	708,05	0,784	1,75	49,44	60,06	4,54	1,98	0,98
28	709,52	0,789	1,73	49,86	59,58	4,64	1,93	0,98
29	708,31	0,792	1,76	50,50	59,45	4,67	1,91	0,98
30	715,05	0,787	1,80	50,25	59,79	4,60	1,95	0,98
31	707,80	0,787	1,75	49,10	59,05	4,82	1,95	0,98
32	713,89	0,790	1,73	50,99	60,20	4,69	1,93	0,98
33	685,97	0,790	1,76	50,02	60,08	4,56	1,93	0,98

Tabla A5.3: Condiciones... (continuación)

<i>Batch</i>	V_R [l]	V_W [-]	$S \times 10^5$ [mol/cm ³]	T [°C]	T_C [°C]	$I \times 10^6$ [mol/cm ³]	$M \times 10^3$ [mol/cm ³]	ρ_R [g/cm ³]
34	699,77	0,785	1,74	51,25	59,97	4,69	1,97	0,98
35	698,17	0,793	1,70	49,19	58,20	4,68	1,90	0,98
36	681,29	0,781	1,75	49,48	59,59	4,68	2,01	0,98
37	703,00	0,784	1,73	49,49	60,20	4,49	1,99	0,98
38	687,67	0,787	1,73	49,93	60,30	4,74	1,95	0,98
39	716,51	0,786	1,77	49,35	60,14	4,50	1,96	0,98
40	690,61	0,789	1,79	49,36	59,97	4,71	1,93	0,98
41	706,17	0,790	1,72	48,41	58,72	4,54	1,93	0,98
42	702,56	0,785	1,74	49,44	60,27	4,65	1,98	0,98
43	689,24	0,783	1,71	50,17	59,90	4,70	1,99	0,98
44	674,68	0,782	1,75	49,88	60,39	4,69	2,01	0,98
45	699,31	0,787	1,67	49,89	60,48	4,63	1,96	0,98
46	688,21	0,784	1,76	49,92	59,93	4,55	1,98	0,98
47	705,92	0,778	1,74	49,67	59,73	4,56	2,04	0,98
48	719,75	0,787	1,77	49,53	59,74	4,50	1,95	0,98
49	692,49	0,787	1,73	49,72	59,24	4,67	1,96	0,98
50	704,44	0,789	1,70	50,15	60,19	4,67	1,94	0,98
51	688,23	0,784	1,69	49,36	60,99	4,59	1,98	0,98
52	699,77	0,786	1,71	49,90	59,54	4,63	1,96	0,98
53	699,44	0,782	1,74	50,50	59,60	4,66	2,00	0,98
54	700,00	0,784	1,72	49,41	60,58	4,64	1,98	0,98
55	712,78	0,797	1,76	49,84	61,23	4,78	1,86	0,98
56	678,14	0,784	1,71	49,95	58,64	4,45	1,98	0,98
57	705,00	0,790	1,73	49,76	59,17	4,65	1,93	0,98
58	710,45	0,783	1,73	48,61	59,39	4,65	1,99	0,98
59	706,74	0,790	1,75	50,93	60,35	4,58	1,93	0,98
60	700,47	0,782	1,73	50,43	59,64	4,69	2,00	0,98
61	707,90	0,786	1,69	49,75	60,02	4,64	1,97	0,98
62	706,64	0,788	1,70	49,56	59,75	4,69	1,94	0,98
63	697,02	0,780	1,74	50,17	60,78	4,70	2,01	0,98
64	696,55	0,782	1,72	49,65	60,65	4,62	2,00	0,98
65	682,79	0,780	1,71	50,04	60,56	4,73	2,02	0,98
66	703,67	0,781	1,72	50,39	60,10	4,61	2,01	0,98
67	716,84	0,782	1,74	49,33	59,36	4,59	2,00	0,98
68	709,32	0,791	1,73	49,85	59,18	4,58	1,92	0,98
69	710,98	0,791	1,69	50,04	60,75	4,68	1,92	0,98
70	702,78	0,780	1,74	48,99	59,06	4,65	2,02	0,98
71	691,34	0,788	1,77	49,46	59,46	4,62	1,95	0,98
72	712,63	0,794	1,72	50,24	58,62	4,65	1,89	0,98
73	698,47	0,785	1,74	49,61	60,65	4,71	1,97	0,98
74	704,55	0,789	1,73	51,02	61,15	4,69	1,94	0,98

Tabla A5.3: Condiciones.... (continuación)

<i>Batch</i>	V_R [l]	V_W [-]	$S \times 10^5$ [mol/cm ³]	T [°C]	T_C [°C]	$I \times 10^6$ [mol/cm ³]	$M \times 10^3$ [mol/cm ³]	ρ_R [g/cm ³]
75	701,03	0,792	1,72	50,09	60,54	4,52	1,91	0,98
76	692,59	0,784	1,80	48,76	60,11	4,63	1,98	0,98
77	693,47	0,784	1,73	49,70	59,31	4,61	1,98	0,98
78	705,18	0,783	1,69	49,57	59,40	4,67	1,99	0,98
79	688,92	0,787	1,78	50,07	60,98	4,60	1,95	0,98
80	709,11	0,788	1,74	49,85	60,92	4,71	1,95	0,98
81	706,64	0,787	1,70	49,81	59,51	4,61	1,96	0,98
82	690,41	0,787	1,75	49,51	60,64	4,51	1,96	0,98
83	693,95	0,779	1,70	49,56	58,54	4,69	2,03	0,98
84	703,82	0,787	1,77	49,90	58,74	4,45	1,95	0,98
85	702,73	0,783	1,73	50,17	61,61	4,71	1,99	0,98
86	700,25	0,785	1,75	49,95	59,94	4,62	1,97	0,98
87	688,29	0,789	1,71	51,16	62,10	4,50	1,94	0,98
88	695,63	0,795	1,74	50,42	60,14	4,64	1,89	0,98
89	686,16	0,787	1,74	50,06	60,74	4,64	1,95	0,98
90	687,68	0,788	1,79	49,46	59,86	4,55	1,94	0,98
91	717,18	0,788	1,73	49,69	59,58	4,59	1,95	0,98
92	700,65	0,791	1,80	49,06	61,15	4,73	1,92	0,98
93	685,80	0,785	1,78	49,75	59,30	4,59	1,97	0,98
94	686,84	0,789	1,69	51,17	59,59	4,69	1,94	0,98
95	684,26	0,783	1,72	50,50	60,59	4,69	1,99	0,98
96	696,95	0,792	1,73	50,04	60,68	4,63	1,91	0,98
97	701,50	0,787	1,76	50,06	60,43	4,50	1,96	0,98
98	707,66	0,791	1,71	50,51	60,83	4,66	1,92	0,98
99	686,38	0,782	1,71	51,49	59,16	4,57	2,00	0,98
100	696,94	0,785	1,73	50,58	59,97	4,52	1,97	0,98
101	685,85	0,791	1,74	50,45	59,02	4,66	1,92	0,98
102	684,61	0,784	1,76	50,37	59,78	4,67	1,98	0,98
103	710,86	0,795	1,75	50,67	59,87	4,65	1,88	0,98
104	700,13	0,788	1,68	50,84	59,13	4,55	1,95	0,98
105	692,47	0,775	1,75	50,03	58,44	4,60	2,07	0,98
106	709,40	0,785	1,76	49,79	59,27	4,66	1,98	0,98
107	702,70	0,790	1,75	50,15	60,00	4,50	1,93	0,98
108	688,45	0,783	1,77	50,66	58,62	4,57	1,99	0,98
109	715,63	0,787	1,74	50,81	61,01	4,54	1,95	0,98
110	717,25	0,787	1,73	50,61	58,35	4,76	1,95	0,98
111	713,28	0,781	1,66	49,89	60,09	4,68	2,01	0,98
112	692,02	0,782	1,75	49,78	61,19	4,57	2,00	0,98
113	684,93	0,785	1,74	50,36	60,04	4,64	1,97	0,98
114	699,15	0,787	1,72	50,62	60,13	4,61	1,95	0,98
115	696,14	0,785	1,71	49,62	59,98	4,56	1,97	0,98

Tabla A5.3: Condiciones... (continuación)

<i>Batch</i>	V_R [l]	V_W [-]	$S \times 10^5$ [mol/cm ³]	T [°C]	T_C [°C]	$I \times 10^6$ [mol/cm ³]	$M \times 10^3$ [mol/cm ³]	ρ_R [g/cm ³]
112	692,02	0,782	1,75	49,78	61,19	4,57E-06	2,00E-03	0,98
113	684,93	0,785	1,74	50,36	60,04	4,64E-06	1,97E-03	0,98
114	699,15	0,787	1,72	50,62	60,13	4,61E-06	1,95E-03	0,98
116	690,16	0,787	1,70	49,00	60,14	4,65E-06	1,96E-03	0,98
117	705,81	0,786	1,73	49,97	60,71	4,63E-06	1,96E-03	0,98
118	717,37	0,784	1,74	50,37	60,15	4,54E-06	1,99E-03	0,98
119	690,12	0,789	1,72	51,05	60,41	4,60E-06	1,94E-03	0,98
120	697,13	0,786	1,73	49,78	60,11	4,67E-06	1,96E-03	0,98
121	707,74	0,789	1,73	50,53	60,22	4,53E-06	1,94E-03	0,98
122	690,03	0,791	1,71	49,48	59,21	4,58E-06	1,91E-03	0,98
123	685,98	0,783	1,77	49,16	59,24	4,59E-06	1,99E-03	0,98
124	699,24	0,792	1,77	49,07	60,28	4,76E-06	1,91E-03	0,98
125	705,66	0,787	1,72	49,13	59,28	4,73E-06	1,95E-03	0,98
126	693,05	0,787	1,70	49,60	60,65	4,69E-06	1,96E-03	0,98
127	698,25	0,794	1,76	49,01	60,09	4,65E-06	1,89E-03	0,98
128	694,93	0,791	1,73	50,37	62,19	4,63E-06	1,92E-03	0,98
129	699,07	0,793	1,71	49,82	60,33	4,51E-06	1,90E-03	0,98
130	717,91	0,782	1,73	49,14	58,95	4,59E-06	2,00E-03	0,98
131	684,28	0,782	1,77	49,34	60,40	4,64E-06	2,00E-03	0,98
132	705,48	0,784	1,71	49,95	59,59	4,55E-06	1,98E-03	0,98
133	724,64	0,792	1,73	50,33	60,85	4,58E-06	1,91E-03	0,98
134	684,16	0,790	1,75	49,66	58,93	4,52E-06	1,93E-03	0,98
135	688,07	0,787	1,72	49,63	60,38	4,59E-06	1,96E-03	0,98
136	712,11	0,784	1,79	49,50	58,69	4,59E-06	1,98E-03	0,98
137	695,45	0,789	1,77	50,62	58,85	4,60E-06	1,93E-03	0,98
138	683,89	0,788	1,70	49,83	60,24	4,60E-06	1,94E-03	0,98
139	703,68	0,784	1,71	49,90	59,89	4,50E-06	1,98E-03	0,98
140	718,12	0,791	1,77	49,16	59,10	4,71E-06	1,92E-03	0,98
141	708,26	0,782	1,75	50,21	59,77	4,73E-06	2,00E-03	0,98
142	722,84	0,783	1,75	51,79	59,53	4,61E-06	1,99E-03	0,98
143	705,89	0,788	1,68	50,19	59,28	4,57E-06	1,94E-03	0,98
144	696,04	0,781	1,75	50,16	58,81	4,59E-06	2,01E-03	0,98
145	686,70	0,785	1,72	49,04	60,25	4,63E-06	1,97E-03	0,98
146	686,98	0,786	1,70	50,48	60,27	4,65E-06	1,96E-03	0,98

Las trayectorias correspondientes a q_m^0 , T_{C0} y T^0 se describieron mediante perfiles de referencia a los que se adicionaron pequeñas perturbaciones aleatorias, con el objeto de incorporar las posibles variaciones “reales” que pueden sufrir dichas variables durante la

operación del proceso. La Tabla A5.4 muestra los perfiles de referencia y valores máximos para cada una de las perturbaciones

Para cada uno de los *batches* considerados en la *PR*, los perfiles de estas tres variables se representan en MatLab® como vectores de dimensión 161×1 . Cada una de estas componentes es tomada por gPROMS® como el valor de la variable, para la simulación de un intervalo de 30 segundos de operación. Luego de transcurrido este período gPROMS® toma la componente siguiente y continúa la simulación por otros 30 segundos. Este procedimiento se sigue hasta completar los 80 minutos de operación. La comunicación ente MatLab® y gPROMS® se realiza mediante la interface gOmatlab.

Luego de los análisis correspondientes a la *Etapa I*, los *batches*: 10, 13, 19, 20, 32, 34, 41, 44, 55, 58, 67, 76, 87, 94, 95, 99, 101, 109, 119, 128, 130, 133, 136, 140 y 142 fueron eliminados de la *PR*. Así, las condiciones de operación normal (*NOC*) quedan definidas por el subconjunto de los 121 *batches* restantes.

Respecto de las fallas simuladas, las condiciones iniciales correspondientes a las fallas 1 a 46 se tomaron iguales al valor promedio de las consideradas para los 121 *batches* seleccionados para definir las *NOC*. Las perturbaciones simuladas en q_m^0 y T_{C0} en los grupos de fallas 1 a 23 y 24 a 46 se detallan en la Tablas 6.1 a 6.4. Las fallas 47 a 82 se simularon teniendo en cuenta cambios en los valores iniciales de V_R , S_W , V_W , I , M y T (Tablas 6.5).

Tabla A5.4: Trayectorias de q_m^0 , T_{C0} y T^0 consideradas en la *PR*

Variable	Perfil de Referencia	Perturbación Máxima (+/-)
q_m^0	$q_m^0(t) = 15 - \frac{5}{80}t$ [cm ³ /s]	5%
T_{C0}	$T_{C0}(t) = 20$ °C	2%
T^0	$T^0(t) = 50$ °C	2%

Estos cambios se realizaron de a uno por vez manteniendo los valores de las restantes condiciones iniciales iguales a los correspondientes a las fallas 1 a 46.

A5.4 Ruido de Medición

A continuación se describen los criterios utilizados en la selección de los niveles de ruido para cada una de las variables medidas. La magnitud del ruido asociado a cada variable se estableció de acuerdo a las características de los instrumentos de medición correspondientes (Liptak, 2003).

a. Medición de Densidad

Para la medición en línea de la densidad de la mezcla reactiva se considera el uso de medidores oscilatorios de Coriolis los cuales son aplicables a líquidos, gases y barros. El rango de aplicabilidad de los mismos es desde 0,00 a 5,0 g/cm³. La inexactitud típica de este tipo de instrumentos es de $\pm 0,002$ g/cm³ cuando se mide en el rango de 0,3 a 3,0 g/cm³ siendo los valores mínimos y máximos $\pm 0,002$ y $\pm 0,05$ g/cm³, respectivamente. El valor de repetitividad típica de estos instrumentos es de $\pm 0,001$ g/cm³ en el mismo rango.

Teniendo en cuenta estos valores, el nivel de máximo para el ruido en la medición de ρ_R se asumió como $\pm 0,005$ de forma de contemplar otras posibles fuentes de ruido asociadas a esta medición como ruido eléctrico, vibraciones en las líneas, etc.

b. Medición de Temperatura

Uno de los tipos de sensores más populares para la medición de temperatura se basan en termo-resistencias. El rango de aplicación de las mismas es bastante amplio (−183 a 631 °C). Los sensores Clase A tienen una inexactitud de aproximadamente $\pm 0,03$ °C ó 0,01% del

rango total. La inexactitud de los sensores Clase B es de $\pm 0,3\text{ }^{\circ}\text{C}$ ó $0,12\%$ del rango. Dichos valores se ven afectados por el resto de los componentes del sistema.

De forma similar al caso anterior, los niveles máximos de ruido considerados en las mediciones de temperatura se asumieron un poco mayores a la inexactitud del instrumento. Considerando el uso de sensores Clase A y B, los valores máximos establecidos para el ruido asociado a la medición de temperatura fueron $\pm 0,5\text{ }^{\circ}\text{C}$ en T y $\pm 1,0\text{ }^{\circ}\text{C}$ para T_{C0} y T_C , respectivamente.

c. Medición de Caudal

Para la medición del caudal de alimentación de monómero se considera un medidor de turbina. Este tipo de sensores alcanzan repetitividades desde $0,02\%$ a 5% aproximadamente, dependiendo del diseño. El nivel de ruido considerado para la medición de Q_0 es de $\pm 6,5\%$ del valor nominal inicial.

d. Medición de Nivel

Los medidores laser de nivel tienen una inexactitud típica de $\pm 5\text{ mm}$. En este trabajo se usó el doble de dicho valor como cota máxima para el error de medición en el nivel del reactor.

e. Medición de Tensión Superficial

La información sobre medición de tensión superficial en línea no es común en la literatura relacionada a la instrumentación de procesos. Por esta razón, y dada las dificultades encontradas para obtener una idea de la precisión y repetitividad de estos sistemas se adoptó arbitrariamente un nivel de ruido máximo para S_w del 5% del valor inicial promedio.

f. Medición de Peso Molecular

En la actualidad no existe un método homologado para la estimación en línea del peso molecular de los polímeros. Sin embargo, se han propuesto métodos para su estimación indirecta mediante la medición de viscosidad y densidad de la mezcla (Vega y col., 2001). En esta tesis se consideró el valor del ruido en MW como el 5% del valor máximo promedio alcanzado por los *batches* de la *NOC*.

A5.5 Referencias

- Brandrup, J., E. H. Immergut y E. A. Grulke 1999. Polymer Handbook, New York, John Wiley and Sons, Inc.
- Gao, J. y A. Penlidis, "Mathematical Modeling and Computer Simulator/Database for Emulsion Polymerizations". *Progress in Polymer Science* **2002**, 27, (3), 403-535.
- Liptak, B. G. (2003). Instrument Engineers' Handbook, Volume 1: Process Measurement and Analysis, London, CRC PRESS.
- Rawlings, J. B. 1985. Simulation and Stability of Continuous Emulsion Polymerization Reactors. Ph.D. Thesis, Chemical Engineering Department, Wisconsin, University of Wisconsin.
- Vega, M. P., E. L. Lima y J. C. Pinto, "In-line Monitoring of Weight Average Molecular Weight in Solution Polymerizations Using Intrinsic Viscosity Measurements". *Polymer* **2001**, 42, (8), 3909-3914.

Nomenclatura

Nomenclatura General

Variables	Descripción
$\mathbf{0}$	Matriz o vector nulo
A	Número de <i>CI</i> s retenidas
\mathbf{A}	Matriz en el método <i>PARAFAC</i>
a_{ij}	Elemento de la <i>i</i> -ésima fila, <i>j</i> -ésima columna de la matriz \mathbf{A}
AST	Tiempo a la señal de alarma
$AVTI$	Número promedio de errores <i>Tipo I</i>
\mathbf{B}	Matriz en el método <i>PARAFAC</i>
\mathbf{C}	Matriz en el método <i>PARAFAC</i>
$\bar{c}_{j,k}$	Contribución promedio al estadístico indicado en $\mathbb{C}_{j,k}$ de la <i>j</i> -ésima variable en la <i>k</i> -ésima observación
$c_{j,k}^D$	Contribución de la <i>j</i> -ésima variable al estadístico D en la <i>k</i> -ésima observación
$c_{j,k}^{I_d^2}$	Contribución de la <i>j</i> -ésima variable al estadístico I_d^2 en la <i>k</i> -ésima observación
$c_{j,k}^{I_e^2}$	Contribución de la <i>j</i> -ésima variable al estadístico I_e^2 en la <i>k</i> -ésima observación
$c_{j,k}^{SPE}$	Contribución de la <i>j</i> -ésima variable al estadístico SPE en la <i>k</i> -ésima observación
$c_{j,k}^{T^2}$	Contribución de la <i>j</i> -ésima variable al estadístico T^2 en la <i>k</i> -ésima observación
$c_{j,k}^{T_H^2}$	Contribución de la <i>j</i> -ésima variable al estadístico T_H^2 en la <i>k</i> -ésima observación
$curt(x)$	Curtosis de x

Variables	Descripción
D	Estadístico de Hotelling en el espacio de las <i>CPs</i> retenidas
D_c	Valor crítico del estadístico Hotelling en el espacio de las <i>CPs</i> retenidas
$\mathbf{e}, \mathbf{E}, \underline{\mathbf{E}}$	Vector, matriz bidimensional y arreglo tridimensional de residuos
\mathcal{F}_{g_1, g_2}	Distribución F con g_1 grados de libertad en el numerador y g_2 en el denominador
G	Número de grupos usado en el método de validación cruzada.
$H(x)$	Entropía de la variable x
H_0	Hipótesis nula
H_1	Hipótesis alternativa
h_R	Nivel de líquido en el reactor
i	Ordinales de conteo del primer modo (<i>batches</i>)
I'	Número total de <i>batches</i> en la base de datos históricos
I	Número de individuos considerados en <i>PR</i>
\mathbf{I}	Matriz identidad
I_0	Número de <i>batches</i> considerados en la etapa de pre-tratamiento de datos
I_d^2	Estadístico correspondiente a la parte determinística en <i>ICA</i>
I_e^2	Estadístico correspondiente a la parte excluida en <i>ICA</i>
j	Ordinales de conteo del segundo modo (variables)
J	Número de variables o descriptores considerados
k	Ordinales de conteo del tercer modo (observaciones)
K	Número total de observaciones

Variables	Descripción
k_{falla}	Observación en la que se produce la falla
k_{final}	Observación en la que finaliza la falla
k_{inicio}	Observación en la que se inicia la falla
$L(x)$	Negentropía de la variable x
M	Matriz de mezcla de señales independientes
M_0	Concentración de Monómero
$m_i, i=1,2,3$	Pendientes de las rampas de T_C en las fallas 29 a 46
m_k	Valor promedio de SPE de la población de referencia para la k -ésima observación
\mathbb{N}	Espacio de números naturales
$\mathcal{N}(\mu; \sigma)$	Función de probabilidad normal con media μ y varianza σ
$p_{j,r}$	j -ésimo componente del r -ésimo autovector
P_r	r -ésimo autovector de la matriz de correlación
P	Matriz de proyección en el espacio de las CPs retenidas
\bar{P}	Matriz de autovectores de la matriz de correlación
\tilde{P}	Matriz de proyección en el espacio de las CPs excluidas
Q	Error total de representación en la <i>Etapa I</i>
Q	Matriz de blanqueado en <i>ICA</i>
Q_0	Caudal de alimentación de monómero fresco
R	Número de CPs retenidas
\mathbb{R}	Espacio de números reales
S	Matriz de covarianza muestral
S_k	Matriz de covarianza muestral para la observación k

Variables	Descripción
S_{xx}	Matriz de covarianza del método <i>MYT</i>
s_{xx}	Vector de covarianzas del método <i>MYT</i>
SPE	Error cuadrático de predicción
SPE_C	Valor crítico del error cuadrático de predicción
S_w	Concentración de surfactante en la fase acuosa
t	Tiempo
\mathbf{t}	Vector de coordenadas de una observación individual en el espacio de las <i>CPs</i> retenidas
T	Temperatura de la masa reactiva
\mathbf{T}	Matriz de coordenadas en el espacio de <i>CPs</i> retenidas
$\bar{\mathbf{T}}$	Matriz de coordenadas en el espacio completo de <i>CPs</i>
$\tilde{\mathbf{T}}$	Matriz de coordenadas en el espacio completo de <i>CPs</i> excluidas
T^2	Estadístico de Hotelling en el espacio de las variables originales
T_C^2	Valor crítico del estadístico de Hotelling en el espacio de las variables originales
T_C	Temperatura de salida del agua de refrigeración
T_{C0}	Temperatura de entrada del agua de refrigeración
t_{final}	Tiempo final de la falla
t_{inicio}	Tiempo de inicio de la falla
T_H^2	Estadístico de Hawkins
$T_{H_C}^2$	Valor crítico del estadístico de Hawkins
\mathbf{t}_{new}	Vector de proyección en de una nueva observación en el espacio de las <i>CPs</i>

Variables	Descripción
t_{op}	Duración del <i>batch</i>
\mathbf{v}	Vector genérico (vector columna)
V_W	Fracción volumétrica de agua en el reactor
\mathbf{W}	Matriz de proyección en el espacio de las <i>ICs</i> retenidas
$\bar{\mathbf{W}}$	Matriz de proyección en el espacio completo de las <i>ICs</i>
$\tilde{\mathbf{W}}$	Matriz de proyección en el espacio de las <i>ICs</i> excluidas
x	Magnitud escalar representando un dato o medición
\mathbf{x}	Vector de mediciones ó vector de observación
x_j	Magnitud la <i>j</i> -ésima variable en el vector de observación \mathbf{x}
\mathbf{x}^0	Vector de mediciones obtenidas del proceso
$\bar{\mathbf{x}}$	Vector de medias muestrales
\mathbf{X}	Matriz de datos bidimensional
$\mathbf{X}_{(I \times KJ)}, \mathbf{X}_{(IK \times J)}$	Matrices de datos bidimensionales obtenidas mediante algún método de desdoblamiento
$\underline{\mathbf{X}}$	Arreglo de datos de tres o más dimensiones
$\bar{\bar{\mathbf{x}}}$	Media generalizada en el método <i>WKFH</i>
\mathbf{X}_k^0	Matriz de datos sin escalar para la observación <i>k</i>
$\hat{\mathbf{x}}, \hat{\mathbf{X}}, \underline{\hat{\mathbf{X}}}$	Vector, matriz bidimensional y arreglo tridimensional estimados por el modelo de variables latentes
\mathbf{x}_{new}	Vector correspondiente a una nueva observación en el espacio original de las mediciones
$\hat{\mathbf{x}}_{new}$	Vector correspondiente a la estimación de una nueva observación en el espacio original de las mediciones
\mathbf{Y}	Matriz de datos de propiedades finales del producto

Variables	Descripción
\mathbf{Z}	Matriz de correlación muestral
Operaciones Matemáticas	
$ x $	Valor absoluto de x
$ \mathbf{A} $	Determinante de la matriz \mathbf{A}
$\ \mathbf{v}\ $	Norma 2 del vector \mathbf{v}
$\ \mathbf{A}\ _{\infty}$	Norma infinito de la matriz \mathbf{A}
$diag(\mathbf{v})$	Matriz diagonal formada por los elementos del vector \mathbf{v}
$\mathcal{E}\{\mathbf{x}\}$	Valor esperado del vector \mathbf{x}
$P(a b)$	Probabilidad condicional de a dada la condición b
$r(\mathbf{A})$	Rango de la matriz \mathbf{A}
$sign(arg)$	Signo del argumento “ arg ”
$tr(\mathbf{A})$	Traza de la matriz \mathbf{A}
Convenciones	
\cong	Aproximadamente igual
\approx	Se distribuye como
\triangleq	Igual por definición a
$\mathbf{A}_{(a,b,c,d)}$	Partición de la matriz \mathbf{A} formada por las filas a a b y las columnas c a d
$\mathbf{A}_{(:,c,d)}$	Partición de la matriz \mathbf{A} formada por todas las filas y las columnas c a d

Convenciones

$\mathbf{A}_{(ab,:)}$	Partición de la matriz \mathbf{A} formada por las filas a a b y todas las columnas
$\underline{\mathbf{A}}_{(ab,c:d,ef)}$	Extensión de las convenciones anteriores para arreglos tridimensionales $\underline{\mathbf{A}}$ formadas por las filas a a b y las columnas c a d
$[\mathbf{X}_1 \ \mathbf{X}_2 \ \dots \ \mathbf{X}_K]$	Concatenación horizontal de las matrices \mathbf{X}_1 hasta \mathbf{X}_K
$[\mathbf{X}_1; \ \mathbf{X}_2; \ \dots \ \mathbf{X}_K]$	Concatenación vertical de las matrices \mathbf{X}_1 hasta \mathbf{X}_K

Letras Griegas

α	Nivel de significación para el testeo de hipótesis
α_G	Nivel de significación global para el testeo de hipótesis
α_L	Nivel de significación local para el testeo de hipótesis
β	Probabilidad de error <i>Tipo II</i> asociada a un test de hipótesis.
χ^2	Distribución chi-cuadrado
δ	Vector de desviaciones
Δk	Ancho de la ventana móvil
λ_r	r -ésimo autovalor de la matriz de covarianza
Λ	Matriz de los primeros R autovalores de la matriz de correlación
$\bar{\Lambda}$	Matriz de autovalores de la matriz de correlación
$\tilde{\Lambda}$	Matriz de los últimos $J-R$ autovalores de la matriz de correlación
μ_j	Valor medio poblacional de la variable j
$\boldsymbol{\mu}$	Vector de medias poblacionales

Letras Griegas

V_k	Varianza de <i>SPE</i> de la población de referencia para la <i>k</i> -ésima observación
Σ	Matriz de covarianza poblacional
ρ	Coeficiente de correlación
ρ_R	Densidad de la mezcla reactiva
σ_{ij}	Covarianza entre las variables <i>i</i> y <i>j</i>
σ_j^2	Varianza poblacional de la variable <i>j</i>
σ_T	Varianza total del sistema
$\sigma_{c_{j,k}}^{\#}$	Desvío estándar de las contribuciones al estadístico indicado en $\#$ de la <i>j</i> -ésima variable en la <i>k</i> -ésima observación
τ	Intervalo de muestreo
ψ	Valor de la coordenada individual en el espacio de <i>ICs</i>
Ψ	Vector de coordenadas en el espacio de <i>ICs</i>
Ψ	Matriz de coordenadas de X en el espacio de <i>ICs</i> retenidas
$\bar{\Psi}$	Matriz de coordenadas de X en el espacio completo de <i>ICs</i>
$\tilde{\Psi}$	Matriz de coordenadas de X en el espacio de <i>ICs</i> excluidas

Subíndices, superíndices y tildes

$\hat{\#}$	Referente a la reconstrucción obtenida a partir de un espacio latente
$\bar{\#}$	Variables en los espacios latentes cuando se consideran todas las direcciones
$\tilde{\#}$	Variables en los espacios latentes excluidos
$\sim\#$	Referente a los modelos evolutivos

Subíndices, superíndices y tildes

$\overset{\sim}{\cdot}$	Referente a los modelos de ventana móvil
\cdot_k	Variable evaluada en la k -ésima observación
\cdot_{new}	Correspondiente a una nueva observación durante la etapa de monitoreo en línea
\cdot_{Ref}	Relativo a la población de referencia
\cdot^T	Transpuesta de la variable
\cdot^{ind}	Propiedad individual asociada a una dada componente
\cdot^{accum}	Valor de una dada propiedad acumulado por un grupo de componentes

Abreviaturas

<i>ARMAX</i>	Auto-Regresive Moving Average Exogenous
<i>AT</i>	Método de Wold, Kettaneh, Fridén, y Holmberg Modificado
<i>BDPCA</i>	Batch Dynamic <i>PCA</i>
<i>BDPLS</i>	Batch Dynamic <i>PLS</i>
<i>BOL</i>	Batch observation level
<i>BW</i>	Desdoblamiento <i>batch-wise</i>
<i>CA</i>	Análisis de correspondencia
<i>CAM</i>	Análisis Basado en Cambios Acumulativos
<i>CI</i>	Componente Independiente
<i>CP</i>	Componente Principal
<i>CUSUM</i>	Método de sumas acumulativa
<i>CV</i>	Contribución de la variable a los valores finales de los estadísticos utilizados.

Abreviaturas	
<i>DPCA</i>	Enfoques Dinámicos <i>PCA</i>
<i>DPLS</i>	Enfoques Dinámicos <i>PLS</i>
<i>DTW</i>	Dynamic Time Warping
<i>EWMA</i>	Exponential weighted moving average
<i>g.l.</i>	Grados de libertad
<i>HPCA</i>	<i>PCA</i> jerárquico
<i>ICA</i>	Análisis de Componentes Independientes
<i>KDE</i>	Kernel Density Estimation
<i>LCL</i>	Límite de Control Inferior
<i>LCL_r</i>	Límite de Control Inferior para la <i>r</i> -ésima <i>CP</i>
<i>ME</i>	Modelos Evolutivos
<i>MEWMA</i>	Promedio móvil multivariado con suavizado exponencial
<i>MG</i>	Modelo Global
<i>MICA</i>	Multi-way <i>ICA</i>
<i>ML</i>	Modelos Locales
<i>MM</i>	Multi-Modo
<i>MMA</i>	Metacrilato de metilo
<i>MPCA</i>	Multi-way <i>PCA</i>
<i>MPLS</i>	Multi-way <i>PLS</i>
<i>MSPC</i>	Control Estadístico multivariable
<i>MYT</i>	Método de Mason, Young y Tracy
<i>MW</i>	Peso molecular promedio en peso
<i>NIPALS</i>	Método de cuadrados mínimos iterativos

Abreviaturas

<i>NLPCA</i>	<i>PCA</i> no lineal
<i>NM</i>	Método de Nomikos y MacGregor
<i>NOC</i>	Condiciones normales de operación
<i>OSS</i>	Original Space Strategy
<i>OSS-VM</i>	Original Space Strategy con ventana móvil
<i>PARAFAC</i>	Análisis paralelo de Factores
<i>PCA</i>	Análisis de Componentes Principales
<i>pdf</i>	Función de densidad de probabilidad
<i>PLS</i>	Cuadrados Mínimos Parciales
<i>PR</i>	Población de Referencia
<i>SCORES</i>	Coordenadas de las observaciones en los espacios de variables latentes
<i>SPC</i>	Control Estadístico de Procesos
<i>SCREE</i>	Método de selección del número de componentes principales a retener
<i>TW</i>	Desdoblamiento <i>time-wise</i>
<i>UCL</i>	Límite de Control Superior
<i>UCL_r</i>	Límite de Control Superior para la <i>r</i> -ésima <i>CP</i>
<i>VL</i>	Variable latente
<i>VM</i>	Modelos de Ventana Móvil
<i>VW</i>	Desdoblamiento variable-wise
<i>WKFH</i>	Método de Wold, Kettaneh, Fridén, y Holmberg Original
<i>WKFH-AT</i>	Método de Wold, Kettaneh, Fridén, y Holmberg Modificado
