PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

SISTEMAS INTELIGENTES Y SU APLICACIÓN A CENTRALES DE CICLO COMBINADO

RODRIGO ANDRÉS BERRÍOS ALVAREZ

Tesis presentada a la Dirección de Investigación y Postgrado
como parte de los requisitos para optar al grado de
Magister en Ciencias de la Ingeniería

Profesor Supervisor:
ALDO CIPRIANO ZAMORANO

Santiago de Chile, Noviembre 2008

© MMVIII, RODRIGO BERRÍOS A.

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

SISTEMAS INTELIGENTES Y SU APLICACIÓN A CENTRALES DE CICLO COMBINADO

RODRIGO ANDRÉS BERRÍOS ALVAREZ

Miembros del Comité:
ALDO CIPRIANO ZAMORANO
DAVID WATTS CASIMIS
MARCOS ORCHARD CONCHA
DAVID FULLER PADILLA

Tesis presentada a la Dirección de Investigación y Postgrado
como parte de los requisitos para optar al grado de
Magister en Ciencias de la Ingeniería

Santiago de Chile, Noviembre 2008

© MMVIII, RODRIGO BERRÍOS A.

-

AGRADECIMIENTOS

En primero lugar, quiero agradecer a mi profesor guía Don Aldo Cipriano por su paciencia y el apoyo que me entregó siempre y durante el desarrollo de este trabajo. Su guía fue fundamental para darle buen término a esta tesis.

Quiero darle las gracias a Pamela por su apoyo incondicional durante todo el tiempo en que estuve realizando esta tesis, y a Felipe por ser un gran compañero de trabajo y de oficina.

Agradezco a Alem Ogueda de Endesa Chile y a Rodrigo Paredes del Centro de Monitoreo y Diagnóstico de Endesa Chile, por su colaboración y ayuda en la medida que les fue posible. A Sergio Ávila por revisar y corregir detalladamente este documento, entregándome comentarios valiosos que pude incorporar.

De forma muy especial, agradezco también a Dow Química Chilena S.A. por el premio que me otorgó, que me facilitó en gran forma el trabajo involucrado en el desarrollo de esta tesis.

INDICE GENERAL

AGRADECIMIENTOS iv

INDICE DE FIGURAS ix

INDICE DE TABLAS xii

RESUMEN xiii

ABSTRACT xiv

1. INTRODUCCION 1
   1.1. Motivación 1
   1.2. Objetivos 3
   1.3. Marco de desarrollo de la Tesis 4
   1.4. Contribuciones de esta Tesis 4
   1.5. Organización de la Tesis 5

2. ANTECEDENTES 8
   2.1. Descripción del proceso 8
      2.1.1. Turbina de gas 9
      2.1.2. Caldera recuperadora 10
      2.1.3. Turbina de vapor 11
      2.1.4. Sistema de alimentación de agua/refrigeración 12
   2.2. Problemas de operación en centrales de CC 13
      2.2.1. Problemas de control 13
      2.2.2. Anomalías en el proceso 15
   2.3. Casos de estudio 18
      2.3.1. Control de partida/parada 18
      2.3.2. Control en régimen estable 19
      2.3.3. Diagnóstico para turbina de gas 21

2.3.4. Diagnóstico para caldera recuperadora 23
2.3.5. Diagnóstico para turbina de vapor 25
2.3.6. Diagnóstico para sistemas de refrigeración y ciclo de agua 25

3. MODELACIÓN DE UNA TURBINA DE GAS DE UNA CENTRAL DE CICLO COMBINADO 27
3.1. Introducción 27
3.2. Modelos fenomenológicos 28
3.3. Modelos simplificados 35
3.4. Modelos empíricos 37
3.4.1. Modelos ARX 38
3.4.2. Modelos de espacio de estados 40
3.4.3. Modelos de redes neuronales 46
3.4.4. Modelos difusos 51
3.4.5. Modelos difusos de Takagi-Sugeno 55
3.4.6. Otras técnicas 58
3.5. Modelos comerciales 58

4. DESARROLLO Y SIMULACIÓN DE UN MODELO PARA LA CENTRAL SAN ISIDRO 61
4.1. Estructura del modelo 61
4.2. Estimación y selección de parámetros 64
4.3. Modelos obtenidos con enfoque en simulación 72
4.4. Modelos obtenidos para predicción 80
4.5. Conclusiones 82

5. DETECCIÓN Y DIAGNÓSTICO DE FALLAS EN SISTEMAS DINÁMICOS 84
5.1. Introducción 84
5.2. Metodologías de detección y diagnóstico 85
5.3. Detección y diagnóstico de fallas basado en modelos 87
5.3.1. Esquemas de generación de residuos 90

5.3.2. Generación de residuos con estimación de parámetros 91
5.3.3. Generación de residuos con estimación de estado 93
5.3.4. Generación de residuos con ecuaciones de paridad 95
5.3.5. Generación de residuos con otras técnicas 97
5.4. Detección y diagnóstico de fallas basado en análisis estadístico 98
5.4.1. Detección de cambios en índices estadísticos 100
5.4.2. Análisis de componentes principales 101
5.4.3. Test T^2 de Hotelling 103
5.4.4. Test Q 104
5.4.5. Test de Scheffé 106
5.5. Detección y diagnóstico de fallas basados en otras técnicas 107

6. DESARROLLO DE UN SISTEMA DE DETECCIÓN Y DIAGNÓSTICO DE FALLAS PARA SENSORES DE UNA TURBINA DE GAS 108
6.1. Estructura del sistema de detección y diagnóstico 109
6.1.1. Análisis de los enfoques clásicos 109
6.1.2. Esquema propuesto de estimación de señales de falla 111
6.1.3. Esquema propuesto de diagnóstico de las señales de falla 111
6.2. Reconstrucción de la señal dañada 114
6.3. Entrenamiento del sistema 116
6.4. Resultados 119
6.4.1. Caso 1: Sensores sin falla 120
6.4.2. Caso 2: Sensores bajo falla abrupta 123
6.4.3. Caso 3: Sensores bajo falla no abrupta y aleatoria 128
6.4.4. Caso 4: Sensores bajo falla real 131
6.4.5. Caso 5: Falla de varianza del sensor 135
6.5. Esquema de aplicación del sistema 140
6.6. Conclusiones 141

7. CONCLUSIONES Y TRABAJO FUTURO 143

7.1. Conclusiones y comentarios generales 143
7.2. Temas de investigación futura 146

BIBLIOGRAFIA 148

ANEXO A. ESTIMACION DE RUIDO EN MODELO SS 157

INDICE DE FIGURAS

2.1 Diagrama general de una central de ciclo combinado. 9

2.2 Diagrama esquemático de una turbina de gas. 10

2.3 Diagrama esquemático de una caldera de recuperación. 11

2.4 Diagrama esquemático de una turbina de vapor. 12

2.5 Diagrama esquemático del sistema de condensación y refrigeración. 13

3.1 Modelo simplificado de una central de ciclo combinado. 36

3.2 Modelo simplificado de la TG. 37

3.3 Estructura de una red neuronal tipo perceptron multicapa. 46

3.4 Ejemplo de modelo de proceso de una central de CC en IPSEpro. 59

4.1 Esquema del modelo de la TG de Central San Isidro. 63

4.2 Evolución del error cuadrático medio para algoritmos de BP y de LM. 65

4.3 Ajuste en base a número de parámetros en modelos de redes MLP. 66

4.4 Ajuste en base a número de parámetros en modelos de redes MLP, con 2 capas. 67

4.5 Ajuste en base al número de parámetros en modelos RBFN. 68

4.6 Ajuste en base al número de parámetros en modelos difusos de TS. 69

4.7 Ajuste en base a los parámetros en modelos de TS con identificación conjunta 69

4.8 Ajuste en base al número de parámetros en modelos difusos de TS con clustering. 70

4.9 Comparación de los resultados para cada técnica. 71

4.10 Mejor ajuste a la salida del IGV. 73

4.11 Mejor ajuste a la temperatura de salida del compresor. 74

4.12 Mejor ajuste para el flujo de combustible. 75

4.13 Mejor ajuste para la potencia generada. 76

4.14 Mejor ajuste para la velocidad del eje. 77

4.15 Mejor ajuste para la temperatura a la salida de la turbina. 77
4.16 Mejor ajuste para la temperatura a la entrada del HRSG. 78
4.17 Mejor ajuste para presión en la cámara de combustión y temperatura del rotor. 79

5.1 Esquema tradicional de control de un proceso, afectado por fallas. 85
5.2 Estructura de un sistema FDD basado en modelos. 88
5.3 Esquema del proceso sometido a fallas con un sistema FDD. 89
5.4 Estructura general de un generador de residuos. 90
5.5 Esquema típico de generación de residuos. 91
5.6 Estimación de parámetros. a) Ecuación de error. b) Error de salida. 93
5.7 Esquema de ecuaciones de paridad. a) Error de salida. b) Ecuación de error. . 96

6.1 Esquema de generación de residuos basado en múltiples observadores. 110
6.2 Esquema de estimación directa de las señales de falla. 111
6.3 Tests Q, T^2 de Hotelling y de Scheffé para señal en operación normal. 112
6.4 Tests Q, T^2 de Hotelling y de Scheffé para señal en falla. 113
6.5 Señales de alarma filtradas para los 3 tests, en falla. 114
6.6 Estructura completa del sistema de mediciones tolerantes a fallas. 115
6.7 Estructura del estimador de fallas para el subsistema de cámara de combustión. 116
6.8 Patrón de entrenamiento para el estimador de fallas. 117
6.9 Ejemplo de falla generada a través del proceso estocástico planteado. 118
6.10 Estimación de las señales de falla. 120
6.11 Estimación mejorada de las señales de falla. 121
6.12 Aplicación del test Q y detección de la falla. 122
6.13 Señales de corrección de variables medidas. 123
6.14 Estimación de las señales de falla para falla abrupta. 124
6.15 Estimación mejorada de las señales de falla para falla abrupta y test Q. 125

6.16 Señales de corrección de variables medidas, para falla abrupta. 126
6.17 Señales corregidas, para falla abrupta. 127
6.18 Estimación de las señales de falla para falla no abrupta. 128
6.19 Estimación mejorada de las señales de falla para falla no abrupta y test Q. . . 129
6.20 Señales de corrección de variables medidas, para falla no abrupta. 130
6.21 Señales corregidas, para falla no abrupta. 131
6.22 Estimación de las señales de falla para falla real. 132
6.23 Estimación mejorada de las señales de falla para falla real y test Q. 133
6.24 Señales de corrección de variables medidas, para falla real. 134
6.25 Señales corregidas, para falla real. 135
6.26 Estimación de las señales de falla para falla de varianza. 136
6.27 Test Q para falla de varianza. 137
6.28 Análisis para falla de varianza. 138
6.29 Señales de corrección de variables medidas, para falla de varianza. 139
6.30 Señales corregidas, para falla de varianza. 140
6.31 Esquema de operación ideado para el módulo de mediciones tolerantes a fallas. 141

INDICE DE TABLAS

4.1 Resultados del ajuste de modelos para el sistema IGV. 73

4.2 Resultados del ajuste de modelos para el sistema del Compresor. 74

4.3 Resultados del ajuste de modelos para el sistema de combustible. 75

4.4 Ajuste de modelos para Cámara de combustión - Turbina (primeras 3 salidas). 76

4.5 Ajuste de modelos para Cámara de combustión - Turbina (últimas 3 salidas). . 78

4.6 Resultados del ajuste de modelos para la predicción en el sistema IGV. 80

4.7 Resultados del ajuste de modelos para la predicción en el sistema del compresor. 81

4.8 Resultados del ajuste de modelos para la predicción en el sistema de combustible. 81

4.9 Ajuste de modelos para predicción en Cámara de combustión - Turbina (1). . 81

4.10 Ajuste de modelos para predicción en Cámara de combustión - Turbina (2). . 82

6.1 Resultados para los casos en estudio. 138

RESUMEN

En esta tesis se estudia y analiza el uso de sistemas inteligentes (redes neuronales, lógica difusa, etc.) en aplicaciones para una central de generación eléctrica de ciclo combinado (central de CC). En una primera parte, se describe el proceso y se comentan los principales problemas asociados a la operación documentados en la literatura. Se realiza un estudio del estado del arte de las aplicaciones disponibles para centrales de CC, que atacan los problemas de operación antes mencionados. En una segunda parte, se estudia el problema de la modelación matemática de una central de CC. Se analizan modelos fenomenológicos, simplificados, empíricos, y modelos comerciales. Se estudian las ventajas que ofrecen los sistemas inteligentes sobre las técnicas convencionales, y luego se desarrolla un modelo para la simulación de la turbina de gas de la central de CC San Isidro, propiedad de Endesa Chile. Finalmente, en una tercera parte se estudia el problema de la detección y el diagnóstico de fallas en centrales de CC. Se analizan los enfoques clásicos, para luego proponer un nuevo esquema de detección, diagnóstico y corrección de fallas para sensores de la turbina de gas de la central San Isidro. Todas las conclusiones se realizan a partir de los resultados obtenidos bajo datos de operación reales de la planta.

Palabras Claves: sistemas inteligentes, centrales de ciclo combinado, modelos difusos, detección y diagnóstico de fallas, mediciones tolerantes a fallas.

ABSTRACT

In this thesis we study and analyze the use of intelligent systems (neural networks, fuzzy logic) in applications to a combined cycle power plant (CC plant). In the first part, the process is described and the major problems associated with the plant operation documented in the literature are discussed. An study of the state-of-the-art for CC power plants available applications is made, in line of the plant operation problems previously discussed. In a second part, we study the problem of the mathematical modeling of a CC power plant. Phenomenological, simplified, empirical, and business models are analyzed. We study the advantages of intelligent systems in comparison with the conventional techniques, and then develops a model for the simulation of the San Isidro CC power plant gas turbine, owned by Endesa Chile. Finally, in a third part, the problem of fault detection and diagnosis in CC power plants is studied. It discusses the classic approaches, and then a new detection, diagnosis and correction system for sensor faults in the San Isidro CC power plant gas turbine is proposed. All the conclusions are made from the results obtained under real operating data from the plant.

Keywords: intelligent systems, combined cycle power plant, fuzzy models, fault detection and diagnosis, fault tolerant measurement.

1. INTRODUCCION

1.1. Motivación

Las centrales de ciclo combinado juegan hoy un papel fundamental en la generación de electricidad a nivel nacional y mundial, debido a su alta eficiencia. Esto explica que se hayan desarrollado numerosas aplicaciones en la línea de control y monitorización para este tipo de centrales.

Una aplicación de control o monitorización consiste en el desarrollo de un algoritmo que resuelve un problema particular dentro del área estudiada, y que está diseñado de tal forma que es posible implementarlo de manera genérica sin importar el lenguaje particular tanto en software como posiblemente en hardware. En general, en el desarrollo de aplicaciones para procesos industriales se requiere de un conocimiento experto bastante acabado del sistema en estudio, muchas veces materializado a través de modelos. Sin embargo, en algunos casos no ha sido posible desarrollar modelos matemáticos validados para un rango de operación completo o suficientemente amplio de los procesos industiales más comunes. A esto se suma el problema de que el enfoque clásico de la teoría de control se basa en su mayor parte en modelos lineales de los procesos, lo que implica una linealización de los posibles modelos matemáticos dando lugar a errores que en la mayoría de los casos no son despreciables.

En lo que respecta a centrales de ciclo combinado, los modelos fenomenológicos existentes del proceso y sus correspondientes simplificaciones no son capaces de balancear una buena aproximación del comportamiento real con complejidad y costo computacional, lo cual representa una limitante para el desarrollo de aplicaciones para este tipo de proceso.

Los sistemas inteligentes corresponden a una línea de técnicas también denominadas de inteligencia artificial o inteligencia computacional, que están basadas en la inclusión de información heurística de los procesos a través de algoritmos inspirados en algunos procesos biológicos de razonamiento y evolución presentes en la naturaleza, cuya principal característica es el aprendizaje y la adaptabilidad. Entre éstas técnicas pueden mencionarse

los sistemas expertos, que incorporan el conocimiento de expertos operadores para la toma de decisiones; los sistemas basados en lógica difusa que integran rangos lingüísticos para la modelación, supervisión y control de sistemas; las redes neuronales que emulan el comportamiento biológico de las neuronas y sus conexiones e integran capacidades de aprendizaje, los algoritmos genéticos que se basan en la evolución biológica para realizar tareas de optimización, etc.

Dadas las características de adaptabilidad y aprendizaje, los sistemas inteligentes se han empleado para modelar procesos no lineales mediante la identificación de sus parámetros en base a los datos de entrada y salida del proceso, y el conocimiento experto que se tiene de él. El desarrollo de modelos con estas técnicas ha permitido mejores ajustes respecto a los enfoque clásicos en procesos industriales complejos, lo que abre las puertas para el desarrollo de aplicaciones, muchas veces también basadas en sistemas inteligentes. En el desarrollo de esta tesis se identifica un modelo de una central de ciclo combinado basado en técnicas clásicas y técnicas inteligentes, quedando de manifiesto el mejor desempeño de éstas últimas. También se considera el esfuerzo computacional que exigen estas técnicas para identificar sus parámetros.

Una de las principales problemáticas a la cual está sometido un proceso industrial es la presencia de fallas en sus componentes. El diseño de procesos en ingeniería es cada vez más complejo, integrando miles de partes cuyo funcionamiento conjunto permite realizar grandes tareas como por ejemplo generar energía eléctrica. El incremento de la complejidad de los procesos y con ello el número de componentes involucrados, aumentan naturalmente la posibilidad de que ocurran fallas en ellos. Esta problemática ha sido enfrentada a través de los denominados sistemas de detección y diagnóstico de fallas, que son aplicaciones enfocadas a dar aviso oportuno de la ocurrencia de fallas y su posible causa a los operadores de la planta, a través de alarmas físicas individuales o por medio de un sistema SCADA. En general, éstas aplicaciones se basan en el denominado enfoque basado en modelos, que como su nombre lo indica, necesita un modelo del proceso para luego realizar comparaciones y determinar si existe falla. Nuevamente, las aplicaciones de detección y diagnóstico basadas en modelos clásicos requieren la linealización de un modelo

matemático detallado de la planta, por lo que pueden incurrir en errores no despreciables. El enfoque sobre técnicas inteligentes permite el desarrollo con datos reales de la planta sin realizar mayores suposiciones.

Una de las fallas más recurrentes en los procesos industriales son las fallas de los sensores. Los enfoques existentes de aplicaciones de detección y diagnóstico de fallas para sensores se limitan a realizar una detección de la falla, mas no una reparación de la señal dañada. Además, la estructura clásica de la detección basada en modelos permite una detección intuitiva de fallas en sensores de variables de salida del proceso, pero muy oscura en sensores de variables de entrada. Esta tesis propone un nuevo esquema que transparenta esta detección, la que se combina con técnicas estadísticas que permiten realiza no sólo una detección y diagnóstico si no que también una reparación de las señales dañadas permitiendo la creación de un sistema de mediciones tolerantes a fallas para la planta, el cual se aplica a una turbina de gas.

Los desarrollos planteados en este trabajo han sido propuestos conjuntamente con ENDESA Chile, siendo la central de ciclo combinado San Isidro el foco de atención para su aplicación. Se trabajó en base a datos obtenidos de esta central mediante la data histórica y presente radicada en el Centro de Monitoreo y Diagnóstico (CMD) de ENDESA.

1.2. Objetivos

El objetivo general de este es:

- Desarrollar de forma teórico-práctica aplicaciones de sistemas inteligentes para el diagnóstico y monitorización de una central de ciclo combinado, que puedan aplicarse en los actuales sistemas de control y supervisión de las plantas generadoras térmicas nacionales.

Los objetivos específicos son:

- Realizar un estudio del estado del arte de aplicaciones existentes para centrales de ciclo combinado, y analizar las principales problemáticas analizadas, de tal

forma que pueda servir de punto de partida para el desarrollo de nuevas aplicaciones en la línea de centrales térmicas de generación.

- Desarrollar un modelo y un simulador dinámico para la turbina de gas de central San Isidro, que permita una simulación tanto en línea como fuera de línea, identificado a partir de datos reales y que considere las menos simplificaciones y suposiciones posibles sobre el proceso.
- Comparar el desempeño de las técnicas clásicas lineales y las técnicas enmarcadas en los sistemas inteligentes en la modelación de procesos industriales complejos.
- Desarrollar una aplicación de detección y diagnóstico de fallas para la turbina de gas de central San Isidro, realizando las menores simplificaciones y supuestos posibles sobre el proceso.
- Complementar la detección y el diagnóstico de fallas sobre sensores del proceso con una aplicación que permita reparar señales dañadas y entregarlas de manera auxiliar al sistema de control de la planta.

1.3. Marco de desarrollo de la Tesis

El desarrollo de este trabajo se enmarca en el Convenio ENDESA-Chilectra-Pontificia Universidad Católica de Chile para la Investigación en Sistemas Inteligentes para Innovación Energética, firmado en Julio de 2006 y con duración de 2 años.

Tanto la definición de los objetivos y problemáticas a tratar, como el enfoque en la línea de centrales térmicas de generación de ciclo combinado, fueron propuestos y acordados de forma conjunta entre ENDESA, el autor de esta tesis y el profesor supervisor. Esta alianza ha permitido el desarrollo de aplicaciones de real interés por la industria, y que también tienen importancia desde el punto de vista académico.

1.4. Contribuciones de esta Tesis

Las contribuciones de esta tesis pueden resumirse en 3 puntos principales:

(i) Recopilación de problemáticas y estado del arte de desarrollos para centrales de ciclo combinado, que puede servir como referencia y punto de partida para futuras investigaciones en la línea de centrales térmicas de generación eléctrica.

(ii) Análisis de las ventajas de la modelación inteligente sobre las técnicas tradicionales, realizando una comparación basada en el número de parámetros ocupado en cada modelo. Se considera y compara también el esfuerzo computacional requerido para la determinación de la mejor estructura y la identificación de parámetros en cada caso, concluyendo que los modelos difusos de Takagi Sugeno identificados de forma separada y subóptima representan la mejor alternativa para este tipo de aplicaciones.

(iii) Un nuevo esquema para un sistema de mediciones tolerantes a fallas que combina modelos y el análisis estadístico basado en descomposición de componentes principales. A diferencia de la gran mayoría de estudios referentes a la detección y diagnóstico de fallas en la literatura, este desarrollo realiza una detección de fallas de forma intuititva, y permite la reconstrucción de las señales dañadas de forma fácil, sin realizar mayores supuestos sobre el proceso en estudio.

1.5. Organización de la Tesis

La tesis se organiza en tres partes principales.

La primera parte consiste en una recopilación de antecedentes sobre el proceso. En el Capítulo 2 de este documento se realiza la descripción detallada de una central de ciclo combinado. Se describe el funcionamiento de cada uno de sus componentes, divididos en 4 sistemas principales: turbina de fas, caldera recuperadora, turbina de vapor y sistema de alimentación de agua/refrigeración. Posteriormente, se realiza un estudio acerca de los problemas de operación en las centrales más comunes documentados en la literatura, abarcando problemas de control y problemas de anomalías en el funcionamiento. El capítulo concluye con una recopilación de casos de estudio recientes sobre cada uno de los problemas de operación típicos, mostrando el estado del arte en el desarrollo de aplicaciones para centrales de ciclo combinado (centrales de CC).

En una segunda parte, se trata el problema de modelar y simular una turbina de gas de una central de CC. En el Capítulo 3 se describen modelos fenomenológicos, modelos fenomenológicos simplificados, modelos de entrada-salida y modelos comerciales de una turbina de gas. Para los modelos de entrada-salida, se describen distintas técnicas correspondientes a los enfoques clásicos (sistemas lineales) y a los modelos basados en técnicas inteligentes, como redes neuronales y modelos difusos.

En el Capítulo 4 se desarrolla un modelo para la turbina de gas de central San Isidro. Primero, se realiza un análisis de los modelos descritos en el Capítulo 3, seleccionando como mejor alternativa la modelación en base a datos de entrada-salida. Se determina la estructura del modelo en base a subsistemas considerados a partir de la descripción de la turbina de gas hecha en el Capítulo 1. Para estimar y seleccionar los parámetros y técnicas de modelación, se realiza una comparación del desempeño de cada técnica descrita en el Capítulo 3 en función del número de parámetros requeridos. Como resultado, se obtiene que los modelos difusos de Takagi-Sugeno identificados de forma subóptima muestran excelentes resultados conjugando un número de parámetros aceptable y un esfuerzo computacional accesible para aplicaciones en tiempo real. Posteriormente, se analizan los resultados obtenidos bajo los enfoques de simulación y predicción a N pasos de los modelos identificados para cada variable del proceso.

Finalmente, la tercera parte y final de la tesis trata el problema de detectar y diagnosticar fallas. En el Capítulo 5 se describen las metodologías clásicas de detección y diagnóstico de fallas. Primero, se describe la metodología de detección y diagnóstico basado en modelos. Se comentan las estructuras más utilizadas en la literatura, y se describen las principales técnicas de generación de residuos: estimación de paráemtros, estimación de estado, y ecuaciones de paridad. Se mencionan también otras técnicas de generación de residuos. En segundo lugar, se describen algunas de las técnicas más utilizadas en detección y diagnóstico de fallas basado en análisis estadístico, entre ellas la descomposición en componentes principales, test de Hotelling y test Q. Finalmente se mencionan otras técnicas aplicadas para algunos casos particulares.

El Capítulo 6 muestra el desarrollo de un sistema de detección y diagnóstico de fallas para sensores de una turbina de gas. En este capítulo, se construye un sistema de medición tolerante a fallas para la turbina de gas de central San Isidro. En primer lugar, se comenta la estructura seleccionada para el sistema bajo el análisis de las estructuras clásicas descritas en el capítulo 5. Se propone un esquema nuevo para detectar e identificar la señal de falla. Posteriormente, se analizan las ventajas de combinar los enfoques de detección basada en modelos y diagnóstico basado en análisis estadístico. Se realiza el diseño de un esquema de reconstrucción de la señal, y posteriormente se comenta la metodología de entrenamiento. Finalmente, se exponen 4 casos diferentes de fallas en los sensores, mediante los cuales se analiza el desempeño del sistema.

En el Capítulo 7 se presentan las conclusiones del trabajo y se proponen líneas de investigación futura basadas en esta tesis.

2. ANTECEDENTES

2.1. Descripción del proceso

Una central de ciclo combinado (CC) es una central eléctrica de alta eficiencia y que requiere costos de inversión más bajos que otras tecnologías de generación como la hidráulica. Estas centrales poseen dos tipos de turbinas para generar electricidad: la turbina de gas y turbina de vapor, llamadas así por la forma en que obtienen la energía para generar. Ambas turbinas se mezclan en un ciclo, de forma que la energía almacenada en el gas o vapor es transferida desde una turbina a otra.

El objetivo de combinar en un solo ciclo una turbina de gas con una turbina de vapor es reducir las pérdidas de calor y aumentar la eficiencia de la planta. El resultado más inmediato de combinar ambos ciclos es aprovechar la energía desperdiciada en los gases de escape de la turbina de gas, para producir energía eléctrica adicional. La eficiencia tradicional de una planta basada en una turbina de vapor es de alrededor de un 40%, mientras que para una basada en turbinas de gas es de 35% aproximadamente (Ordys, Pike, Johnson, Katebi, & Grimble, 1994), (Boissenin & Castanier, 1988). La central de CC logra eficiencias hasta de un 60% actualmente (Matta, Mercer, & Tuthill, 2000).

Existen diversos tipos de configuraciones de centrales de CC, en el sentido de como se interconectan los sistemas de vapor y gas, y cómo están dispuestas las turbinas y generadores en forma mecánica. Éstas configuraciones se revisan en (Ordys et al., 1994), como también las configuraciones de plantas de cogeneración de electricidad y calor. Una de estas configuraciones se muestra en la figura 2.1

Las centrales de CC se dividen en tres componentes principales: la turbina de gas, la caldera recuperadora y la turbina de vapor. Cada una de estas etapas posee a su vez subcomponentes, los que se detallan a continuación.

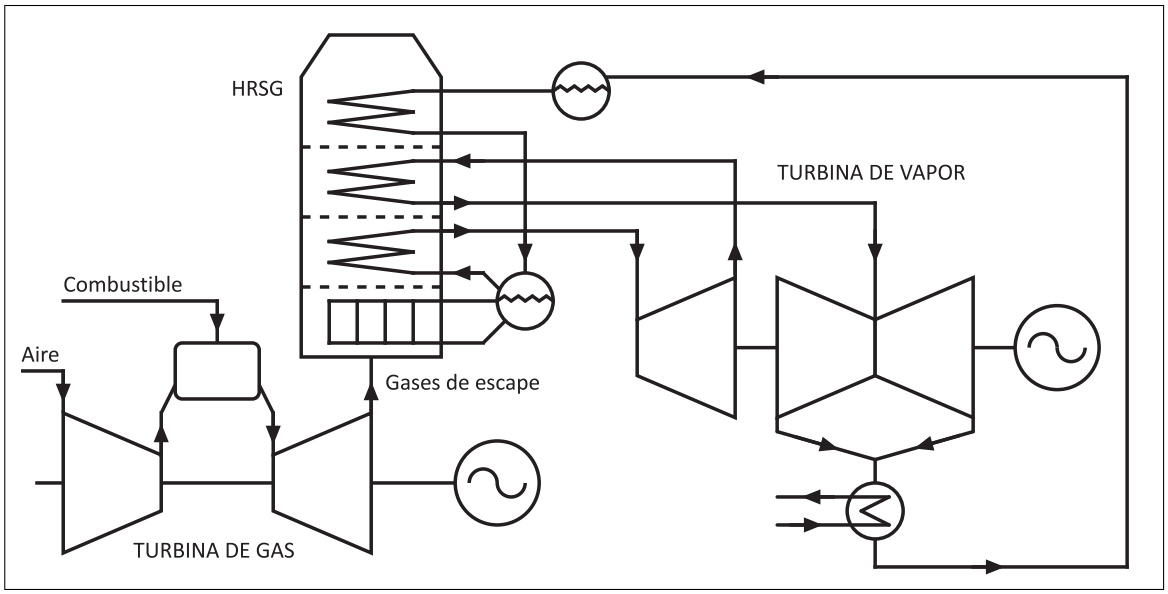

FIGURA 2.1. Diagrama general de una central de ciclo combinado.

2.1.1. Turbina de gas

La turbina de gas usa los gases a alta presión y temperatura generados por una combustión de gas o petróleo para generar electricidad. Para la operación, se requiere aire limpio, que se toma del exterior y se filtra, para luego inyectarse en un compresor (cuya energía es provista a través un mismo eje por la turbina de gas). El aire, cuya presión y temperatura fueron modificadas por el compresor, se inyecta a una cámara de combustión donde se mezcla con el combustible, que puede ser gas o petróleo (previamente temperados). Los gases de la combustión, a alta temperatura y presión, se inyectan dentro de la turbina de gas, que a través de varias etapas de presión convierte la energía del gas en torque mecánico.

Como muestra la figura 2.2, la turbina de gas consta de los siguientes componentes (Ordys et al., 1994): un filtro de aire, el compresor de aire, la caldera de combustión, la turbina propiamente tal, el ducto de escape, y el sistema de refrigeración (en circuitos de aire y agua, junto con aceite lubricante).

Modelos y parámetros de turbinas de gas pueden encontrarse en (Ordys et al., 1994), (Shields, 1989), (Rowen, 1983). Las turbinas pueden operar en ciclos termodinámicos

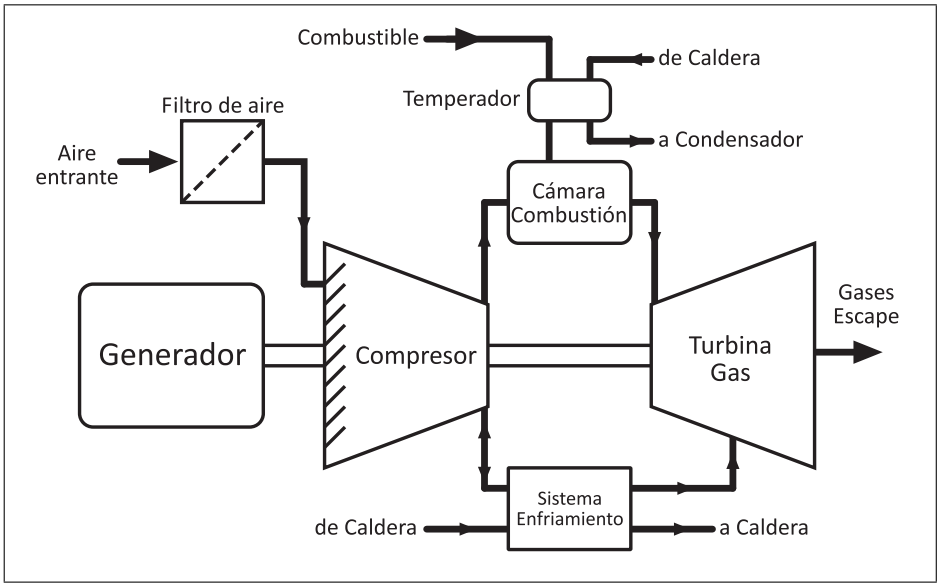

FIGURA 2.2. Diagrama esquemático de una turbina de gas.

tanto abiertos como cerrados, o divididas en etapas según las distintas presiones de operación.

2.1.2. Caldera recuperadora

La caldera recuperadora transforma la energía proveniente del escape de la turbina de gas en un incremento de temperatura en el sistema agua-vapor de la turbina de vapor. La figura 2.3 muestra un esquema de operación de una caldera recuperadora estándar. El circuito consta de un economizador, donde se incrementa la temperatura del agua proveniente de la alimentación, la que luego es depositada en el tambor. El tambor está conectado a elevadores, que causan la evaporación del agua. El agua, ahora convertida en vapor, abandona el tambor para ingresar a diversas etapas de calentamiento, hasta llegar a la condición de vapor sobrecalentado, el que se inyecta a la turbina de vapor (alta presión). La salida de vapor de la turbina es eventualmente atemperada por agua, y luego recalentada para ser reinyectada a la turbina de media presión. El escape de esta turbina alimenta nuevamente una turbina de baja presión.

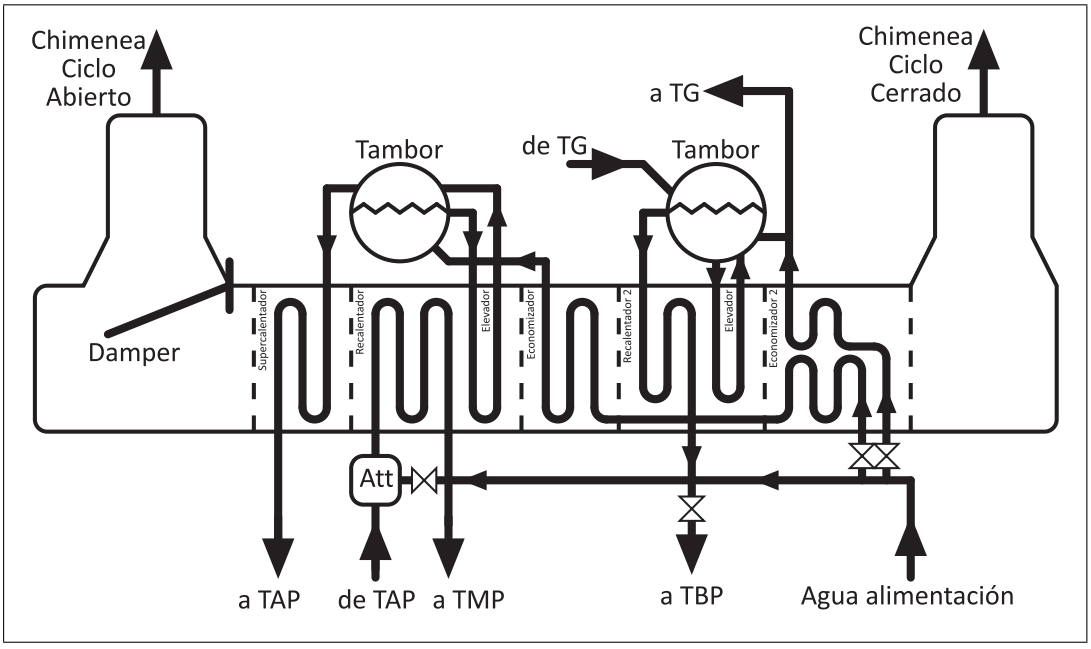

FIGURA 2.3. Diagrama esquemático de una caldera de recuperación.

La caldera cuenta también con una valvula de paso o damper, la que desvía los gases entrantes directamente a una chimenea de escape, y permite el incremento gradual de temperaturas dentro de la caldera.

Los componentes principales de la caldera recuperadora son el tambor, economizador, los elevadores, el recalentador y el sobrecalentador. En algunos casos se incluye una cámara de combustión adicional dentro de la caldera.

Modelos matemáticos y parámetros de calderas pueden encontrarse en (Ordys et al., 1994), (Rovnak & Corlis, 1991).

2.1.3. Turbina de vapor

Como se muestra en la figura 2.4, una turbina de vapor en general está compuesta de 3 etapas, según la presión del vapor: turbinas de alta, media y baja presión. La turbina convierte la energía almacenada en el vapor en la rotación de un eje mecánico, permitiendo su expansión y pérdida de presión. El vapor sobrecalentado ingresa desde la caldera a la turbina de alta presión, para luego ser reingresado a la caldera y recalentado, a presión

constante. Posteriormente, el vapor se inyecta a la turbina de media presión, para luego ingresar a la turbina de baja presión. En esta etapa el vapor puede mezclarse con vapor proveniente de la caldera, de forma de controlar presión y temperatura.

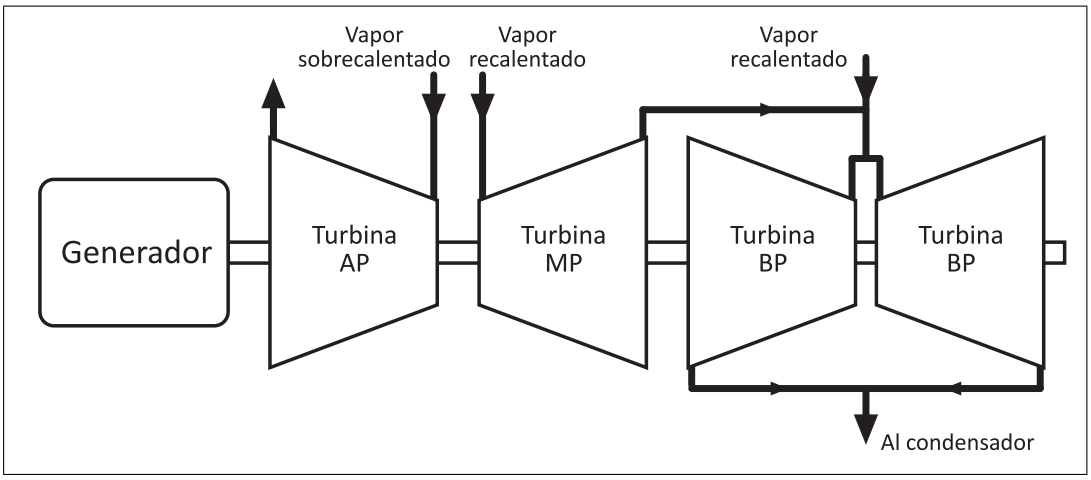

FIGURA 2.4. Diagrama esquemático de una turbina de vapor.

Los principales componentes de la turbina de vapor son las turbinas de alta, media y baja presión. Distintos tipos de turbinas pueden revisarse en (Bolland, 1991).

2.1.4. Sistema de alimentación de agua/refrigeración

El vapor saliente de la turbina de baja presión se transfiere al sistema de condensado, donde es enfriado por otro circuito de refrigeración y convertido en agua. Se extrae también el aire que contiene el flujo, en el desaireador. El agua se bombea a filtros para eliminar posibles impurezas, para luego ser bombeada de regreso al sistema de caldera recuperadora, cerrando así el ciclo agua-vapor. El circuito de refrigeración que forma parte del condensador es un circuito de agua separado, que puede abrirse o cerrarse, y es enfriado en las torres de enfriamiento.

La figura 2.5 muestra los principales componentes del sistema de refrigeración: el desaireador, el condensador, las bombas de condensado y alimentación, los filtros, y la torres de enfriamiento.

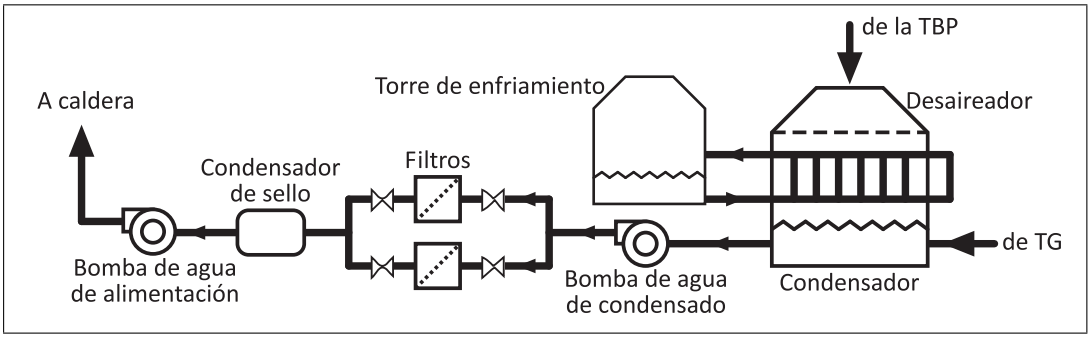

FIGURA 2.5. Diagrama esquemático del sistema de condensación y refrigeración.

2.2. Problemas de operación en centrales de CC

El proceso de generación de electricidad en una central de CC es bastante complejo. Involucra una gran cantidad de componentes, piezas y sensores que están sometidos a estrés térmico y mecánico, y considera variables tan diversas como presiones, temperaturas, velocidades, emisiones de contaminantes, etc. Debido a esta complejidad, surgen una serie de problemas asociados en dos categorías: problemas de control y problemas de anomalías sobre el funcionamiento normal.

2.2.1. Problemas de control

La central de CC es un proceso que involucra el control de muchas variables. El objetivo final de la planta es generar una cierta cantidad de potencia asignada por el despachador de carga. Sin embargo, sobre el control influye una serie de restricciones al proceso como nivel de emisiones de NOx, estrés térmico de los componentes, nivel máximo de oscilaciones mecánicas, térmicas y de presión, etc.

Los principales problemas de control se originan en dos estados de operación de la planta:

- Control del proceso de partida/parada de planta.
- Control de la planta en régimen permanente.

2.2.1.1. Control de partida/parada

La puesta en marcha de la planta es un proceso complicado, con muchas restricciones de diverso tipo, y que puede tomar un largo tiempo. No realizar un arranque exitoso trae como consecuencia directa una pérdida económica, ya que la unidad generadora no puede inyectar potencia hasta que ha sido estabilizada (fin de la fase de arranque).

En respuesta a esta problemática, se puede encontrar en la literatura numerosas aplicaciones que intentan abarcar los siguientes puntos:

- Reducción del tiempo empleado en la puesta en marcha.
- Reducción de las pérdidas de energía en el proceso de partida.
- Auto-adaptación de la partida a variables externas influyentes en el proceso (temperatura ambiente, humedad, etc).
- Control del nivel de emisiones de NOx.
- Control de estrés térmico de componentes claves del sistema.

2.2.1.2. Control en régimen permanente

Una vez que la planta entra en régimen estable y es conectada a la red, surgen nuevos desafíos de control que deben ser estudiados con el fin de minimizar pérdidas asociadas a la disponibilidad y estabilidad de la planta. Las principales aplicaciones y trabajos disponibles apuntan a los siguientes tópicos:

- Aumento de eficiencia de la planta.
- Inclusión de restricciones (emisiones de NOx, estrés térmico de componentes) en algoritmos de control óptimo.
- Reducción de vibraciones y oscilaciones del tipo mecánicas, térmicas y de presión en principales componentes, estabilizando la operación de la planta.
- Reducción de costos operacionales de la planta.
- Mantener estabilidad (principalmente de frecuencia) frente a cambios bruscos en la carga recibida de la red.

2.2.2. Anomalías en el proceso

Naturalmente, surgen anomalías en la operación de este tipo de centrales. Existe una gran cantidad de anomalías documentadas en la literatura. A grandes rasgos, éstas pueden descomponerse en las siguientes categorías:

- Fallas en turbina de gas.
- Fallas en caldera recuperadora.
- Fallas en turbina de vapor.
- Fallas en sistema de refrigeración.

2.2.2.1. Fallas en turbina de gas

Las anomalías más importantes en términos de consecuencias son aquellas que se producen en los componentes principales de la central de CC. Respecto a fallas en turbinas de gas, algunas de las más comunes son las siguientes (Miguel & Delgadillo, 2004):

- Disparo por alta o baja presión en el suministro de gas combustible: Puede alterar la aceleración de la máquina, sacándola del rango crítico de velocidad; si no enciendan los quemadores, se afecta la presión en la cámara de combustión y la temperatura en la TG. Puede verse agravada por una válvula de combustible con problemas, o un controlador desintonizado.
- Mal funcionamiento de la válvula de control de combustible: Puede deberse a un exceso de fricción, una característica de flujo inapropiada o una descalibración.
- Inestabilidades de la velocidad, ocasionadas por el cierre de la válvula de extracción del compresor (bleeding valve): Al cerrarse las válvulas puede producirse una inestabilidad en la velocidad, llevando el control a reducir la magnitud de la misma. Puede provocarse por una mala sintonía del control de velocidad o falta de prealimentación con predicciones de las variables de velocidad y presión.

- Taponamiento o problemas mecánicos con las toberas de combustión: Ocasiona que el gas combustible fluya de forma dispareja, ocasionando una mala distribución de temperaturas en la turbina, lo que puede provocar la detención de la máquina

El modo de operación más complejo de una central de CC es el proceso de partida o parada de la planta. Se han documentado diversos problemas que ocurren durante este proceso, de los cuales el más comun y relevante es el siguiente (Miguel & Delgadillo, 2004):

- Falla en el arranque de la TG por inestabilidad en la dinámica del proceso - sistema de control: Es un distanciamiento de la velocidad de la turbina con respecto a la banda permitida en torno a la referencia. Puede deberse a un ajuste incorrecto de los parámetros del controlador, una presión de combustible fuera de rango, un malfuncionamiento de la válvula de combustible, una característica de flujo de la válvula incorrecta, o un mal funcionamiento del motor de arranque.

2.2.2.2. Fallas en caldera recuperadora

En cuanto al sistema de vapor y caldera, algunas de las fallas más comunes son las siguientes:

- Bloqueo del filtro de agua: Se debe a la acumulación de impurezas dentro del filtro. Puede ocasionar una pérdida de presión en el circuito de agua, que se traduce en una baja de vapor disponible ocasionando pérdida de energía.
- Filtraciones en las tuberías del sistema agua-vapor: Pueden provocar una pérdida de agua y vapor generado, disminuyendo la presión y provocando una caída en la potencia de la turbina
- Falla en el cierre del damper: Genera inestabilidad y pérdida de control sobre las temperaturas dentro de la caldera recuperadora, pudiendo provocar daños por sobrecalentamiento de los componentes.

2.2.2.3. Fallas en turbina de vapor

En general el sistema de turbina de vapor, debido a su simpleza y robustez en comparación con el sistema de turbina de gas y caldera recuperadora, no presenta grandes fallas recurrentes que se repitan en la literatura. De todas formas existen estudios con respecto a las siguientes fallas:

- Vibraciones mecánicas, y de variables de temperatura y presión: Están principalmente relacionadas a desbalances en los álabes producto de corrosión provocada por un vapor no 100% puro.
- Aire de escape con agua: La presencia de agua en el escape de la turbina se considera como una falla en los sistemas de control, debido al peligro que este componente representa en términos de corrosión dentro de la turbina.
- Fallas generales de bloqueos de tubos y válvulas, ocurridas a lo largo de todo el sistema: Estas fallas tienen que ver con bloqueos y corrosión de los elementos mecánicos del sistema, y pueden provocar aumentos y bajas de presión, temperatura, y filtraciones dentro del sistema.

2.2.2.4. Fallas en el sistema de refrigeración

El sistema de refrigeración de la planta es un proceso fundamental que permite el funcionamiento correcto del ciclo principal de turbinas y calderas. Permite que los rangos de operación en términos de temperatura sean aceptables para las capacidades de los materiales, por lo que una falla en éste sistema podría provocar severos daños a los componentes principales de la planta.

Existen pocos trabajos concentrados en el sistema de refrigeración. Las falla documentada más recurrente en este sistema es la siguiente (Xue, Xiao, Bie, & Lu, 2005):

- Bomba de aceite fuera de funcionamiento: Puede provocar una paralización del suministro de aceite a soportes de piezas rotatorias, y con ello su destrucción.

2.3. Casos de estudio

Las aplicaciones de sistemas inteligentes desarrolladas para enfrentar las problemáticas descritas en el punto anterior son numerosas, para cada una de las partes que componen la central. A continuación se listan algunas de éstas aplicaciones, clasificadas según a que parte del proceso están destinadas.

2.3.1. Control de partida/parada

En (Matsumoto, Ohsawa, Takahasi, Akiyama, & Ishiguro, 1996), los autores proponen el uso de un sistema experto para optimizar la partida de la planta. Los objetivos de control son la regulación efectiva de las emisiones de NOx, y la operación de la máquina en rangos seguros que prolonguen su vida útil (disminuyendo estrés térmico). El sistema experto optimiza la agenda tanto de la turbina de gas como de la turbina de vapor. Como forma de representar la planta se usan reglas difusas, que permiten la modificación de los parámetros de la agenda. Un objetivo secundario es la minimización del tiempo de partida, y con ello la pérdida de energía.

En (Sanchez, Moheno, Suarez, & Quintero, 2002), se combinan técnicas de control predictivo con un modelo difuso de la planta del tipo Takagi-Sugeno, con el cual se realiza la predicción. Nuevamente, los objetivos de control tienen que ver con mantener la planta dentro de los límites de estrés térmico permitidos para sus componentes.

De la misma forma, en (Franke & Vogelbacher, 2006), se usa un control predictivo no lineal (NMPC) para minimizar los costos asociados a pérdidas de eficiencia dentro de la partida de la planta, respetando las restricciones de estrés térmico de los componentes. Los autores integran el desarrollo dentro de la plataforma Industrial^{IT} del SCADA 800xA de la firma ABB.

Finalmente, en (Albanesi et al., 2006), nuevamente se tiene por objetivo minimizar el tiempo de partida de la planta, usando un control multivariable tomando en consideración

las restricciones térmicas y de presión, pero en un contexto de dependencia con la red eléctrica. Se evalúan también factores como la regulación de carga y los cambios de frecuencia en la red en la partida y parada de la máquina.

2.3.2. Control en régimen estable

(Prasad, Swidenbank, & Hogg, 1998) se centran en el control en régimen estable de la planta, esto es, con la central conectada al sistema eléctrico, con la fase de partida concluida. En este trabajo se aplica una estrategia de control predictivo basada en una predicción de largo rango a través de un modelo neuronal. La aplicación está dirigida al control de la presión y temperatura del vapor sobrecalentado y recalentado inyectado a la turbina de vapor.

En (Kallappa & Ray, 1999), se desarrolla un control difuso de amplio rango de operación, diseñado para seguir cambios fuertes en la carga despachada a la central, de forma de conseguir una operación robusta frente a perturbaciones y extender la vida útil de las máquinas reduciendo el estrés. Los parámetros del controlados pueden ser sintonizados de forma de darle más peso a la eficiencia y desempeño de la planta o a la protección de los componentes de daños estructurales. El control se aplica sobre toda la sección de vapor de la planta.

En (Saez & Cipriano, 1999b), los autores realizan un control predictivo basado en modelos difusos, aplicado a la caldera recuperadora de la planta. El objetivo de control es minimizar los costos operacionales para un determinado nivel de carga en la central.

En (Poncia & Bittanti, 20 July 2001), se desarrolla también un controlador predictivo multivariable que permite la rápida respuesta frente a cambios de despacho de carga, y maximiza la eficiencia para las condiciones de operación impuestas.

En (Peng, Ozaki, Haggan-Ozaki, & Toyoda, 2002), los autores proponen también un control predictivo, pero basado en una arquitectura de modelos ARX no lineales exponenciales. El control se aplica a toda la sección de vapor de la planta, para lograr satisfacer la

demanda requerida manteniendo dentro de rango los valores de temperatura y presión de los componentes críticos.

En (H. Zhang, Li, & Cai, 2002), se propone un control adaptivo basado en dos partes. Primero, un ciclo de control difuso realiza el control principal sobre la planta, mientras que en segundo lugar existe un ciclo de compensación que sintoniza las reglas difusas en forma automática. El desarrollo se aplica al control del vapor sobrecalentado de la caldera.

(Jianhong, Laijiu, Sujin, & Yiqian, 2002) desarrollan un control PID difuso multivariable para el control de niveles en el desaireador y condensador, en el sistema de alimentación de agua de la caldera y sistemas de vapor.

Por otro lado, en (Jurado & Cano, 2003), se aplica un control neuro-difuso (tipo ANFIS) de 5 capas. El control se aplica a la planta completa, considerando como componentes principales la turbina de gas, caldera recuperadora y turbina de vapor. El objetivo es maximizar la potencia generada con el mínimo combustible posible, para las distintas condiciones de operación.

En (Ferrari-Trecate et al., 2004), los autores proponen un sistema híbrido del tipo MLD para modelar y controlar la planta. Se aplica un control predictivo con el objetivo de minimizar los costos de operación de la planta, considerando precios de combustibles y demandas de electricidad. La integración de dinámicas discontinuas permite realizar cambios directos entre los distintos estados de operación. El control está dirigido a la planta completa.

(Heo, Lee, & Garduno-Ramirez, 2006) proponen la utilización de PSO (Particle Swarm Optimization) para resolver el problema de optimización asociado al control multiobjetivo de la planta considerando restricciones del tipo térmico sobre los componentes, y la carga despachada a la planta. El control se aplica a toda la central.

(Jino, Matsumoto, & Kamito, 2006) proponen el uso de un controlador predictivo híbrido, basado en modelos de riesgo para las restricciones de operación segura en cada componente, y aplica una optimización estocástica de la función objetivo. El control está orientado hacia toda la sección de vapor de la planta.

Finalmente, en (Saez, Milla, & Vargas, 2007), se desarrolla un control supervisor predictivo difuso de turbinas de gas para la central de CC, basado en la aplicación de algoritmos genéticos. El objetivo de control es el desempeño económico y regulatorio de la turbina, variando dinámicamente los set points para el nivel regulatorio del control. Se realizan predicciones con un modelo difuso de la planta, y el control final se basa en una optimización que emplea algoritmos genéticos.

2.3.3. Diagnóstico para turbina de gas

Una de las principales líneas de estudio en cuanto a la resolución de anomalías en turbinas de gas, es aquella referente al estudio de fallas en las lecturas de los sensores asociados al proceso. En caso de que la falla no pueda detectarse, se podrían tomar acciones erróneas de control o ignorar condiciones de peligro para los componentes del sistema. En (Simani & Fantuzzi, 2000), los autores atacan el problema de fallas en sensores generando residuos con un banco de filtros de Kalman, uno para cada variable de salida del proceso. Los sensores estudiados son tanto de variables controladas como variables de salida. Los residuos son clasificados por una red neuronal, cuyo desempeño se evalúa para una topología RBFN y una MLP. Se concluye que para el caso MLP pueden detectarse desvíos de hasta un 1% del sensor respecto al valor real.

También es posible encontrar aplicaciones que atacan la ocurrencia de fallas en el proceso mismo, como puede verse en (Ibargüengoytia, Sucar, & Morales, 2000), en que los autores aplican la misma estructura de dos redes Bayesianas de (Ibargüengoytia, Sucar, & Vadera, 2001), pero directamente en el proceso, considerando variables como el flujo de combustible, de aire, la presión del combustible, las posiciones de las válvulas, de las IGVs, y las referencias de control de éstas últimas variables.

Otro trabajo relacionado con el tema de sensores en la turbina de gas es (Ibargüengoytia et al., 2001). Este artículo se centra en la validación de los datos obtenidos desde los sensores. El enfoque de los autores es realizar la detección en tiempo real, para lo cual usan una estructura basada en dos redes Bayesianas. Una primera red se encarga de representar

las dependencias e independencias de las lecturas de todos los sensores, mientras que la segunda red aísla el posible sensor fallado. La aislación de la falla es incremental, a través de un vector de probabilidades que se actualiza en todo momento. Este trabajo se evaluó en la central Gómez Palacio, en México, en 21 sensores de la turbina de gas.

(Palade, Patton, Uppal, Quevedo, & Daley, 2002) introducen lógica difusa para el diagnóstico de fallas en la turbina de gas, con énfasis en los actuadores. Combinan el conocimiento plasmado en el sistema fuzzy con la capacidad de aprendizaje de la red neuronal. Generan residuos mediante una red neuro-fuzzy del tipo Takagi-Sugeno que detecta la falla, para luego clasificar estos residuos con una red neuro-fuzzy del tipo Mamdani, para realizar la aislación.

En (Sampath, Gulati, & Singh, 2002) se realiza un análisis de desempeño de la turbina. Los autores simulan todas las posibles combinaciones de potenciales problemas con varios niveles de gravedad de las fallas, para luego compararlos con los actuales parámetros obtenidos desde los sensores. La diferencia entre los parámetros se convierte en una función objetivo, que es minimizada usando la técnica de Algoritmos Genéticos. El modelo que minimiza ésta diferencia entrega la naturaleza del problema.

(Arriagada, Genrup, Loberg, & Assadi, 2003), usan también redes neuronales para el diagnóstico de fallas en la turbina. Se basan en el entrenamiento y validación de la red con datos de operación sin fallas, de forma que la red es capaz de detectar la falla a partir de un conjunto de datos que describe el malfuncionamiento del sistema. Las fallas tratadas son falla en el compresor, fricción en álabes del compresor, bloqueo del flujo de aire hacia la caldera de combustión, filtración en el aire de enfriamiento del compresor a la turbina, sello de cubierta central gastado, erosión en los álabes de turbina, y bloqueo del aire de refrigeración en turbina.

En (Simani, 2005), el autor extiende el trabajo realizado en (Simani & Fantuzzi, 2000), para detectar fallas en el proceso utilizando una estructura similar a la antes propuesta. En el caso sin ruido, los residuos se generan con modelos ARX y modelos EIV MISO. El caso con ruido se trata con Filtros de Kalman. Las fallas consideradas son contaminación

del compresor, fallas en termocuplas, daño en el sello de la turbina por alta presión, y malfuncionamiento de la válvula de control de combustible.

En (Flores-Loredo, Ibargüengoytia, & Morales, 2005) se estudia la detección del desvío del comportamiento normal de la turbina en tiempo real, a través de una red Bayesiana. Las principales fallas atacadas son una baja en la presión del suministro de combustible, una falla en las válvulas de extracción del compresor (bleeding valves), y mal funcionamiento de la válvula de control de combustible.

Finalmente, en (Sánchez & Verde, 2006), se revisan las propiedades estructurales relacionadas con la supervisión de una turbina de gas con fallas en sus sensores. Los autores estudian la monitoreabilidad y observabilidad de la planta para determinar un set de restricciones (basadas en la redundancia analítica de las ecuaciones que describen la planta) de donde se puede obtener una estructura de residuos (los cuales permiten la detección de la falla).

2.3.4. Diagnóstico para caldera recuperadora

Entre las aplicaciones para caldera recuperadora, se encuentra la de (Sun, Chen, & Marquez, 2001) en la que se estudia la detección de filtraciones en tuberías y fallas en los sensores (desvíos de los valores reales), relacionados con la operación de la caldera recuperadora. Se aplica un método de mínimos cuadrados con factores de olvido variantes en el tiempo, para estimar los parámetros de un modelo que incluye factores para las filtraciones y fallas en los sensores.

(Liang-yu, Jian-qiang, & Bing-zhu, 2002) realizan un diagnóstico inteligente de fallas usando redes neuronales multicapa, en donde se usa un algoritmo de entrenamiento tipo Back-Propagation mejorado, en el que se cambia la tasa de corrección de error de forma dinámica, autoadaptiva. Los síntomas se tratan con conjuntos difusos. Las fallas analizadas son filtraciones en los tubos, bloqueos parciales en las tuberías, tuberías atestadas, acumulación de aire en los calentadores, y falla en las válvulas de extracción de calor.

Por otro lado, en (Yang & Liu, 2004) se aplican técnicas de Data Mining para el diagnóstico de fallas, basadas en la teoría de *fuzzy rough sets*. El objetivo es obtener la información necesaria desde los datos que maneja el sistema SCADA de la planta, para no realizar mayores intervenciones en la caldera. Las fallas tratadas son la pérdida de calor provocada por la incapacidad de la bomba de manejar todo el flujo drenado desde el calentador en altas condiciones de carga.

En (Xue et al., 2005), los autores pretenden predecir la ocurrencia de fallas en la caldera, para lo cual combinan las matemáticas difusas con una red neuronal tipo RBF. Se estima una tendencia de operación en base a las mediciones actuales y pasadas de los sensores. Las fallas son consideradas como cambios en los parámetros del sistema, que son estimados a través de la red. Las fallas tratadas son la falla en la bomba de combustible, el bloqueo del filtro de agua, y una falla en la máquina preparadora de combustible (para calderas con sistema de calentamiento adicional o auxiliar).

En (Ma, Ma, & Ma, 2005) se usa una red neuronal tipo RBF para detectar y diagnosticar fallas en el proceso termodinámico de la caldera. Se discuten distintos tipos de algoritmos de aprendizaje como Back-Propagation y Orthogonal Least Squares. También se discute el tratamiento de los síntomas, como variables continuas, o pertenecientes a conjuntos difusos discretos. Los tipos de fallas tratados son filtraciones en tubos del calentador, obstrucciones totales y parciales en tubos, acomulación de aire en el calentador, y valvula de extracción del calentadoren malfuncionamiento.

(Huang, Qi, & Liu, 2006) atacan las fallas ocurridas en el sistema de control de la caldera. Esto incluye mala sintonía o mal diseño de controladores, y fallas en actuadores y sensores. La detección y el diagnóstico se divide en tres niveles: de componente individual, ciclos de control locales, y el sistema en general. Se sugieren y comparan distintos tipos de técnicas, entre ellas observadores difusos, estimación de parámetros y redes neuronales.

Por último, en (Binsompasong, 2006) se estudian los efectos de una falla en la operación del damper, con el objeto de elaborar un mantenimiento predictivo para proteger tanto

a la turbina de gas como a la caldera recuperadora.

2.3.5. Diagnóstico para turbina de vapor

En cuanto a trabajos realizados referentes a la turbina de vapor, se puede revisar (Jian, Cai-xin, Weimin, Bide, & Ruijin, 2002), en que se diagnostican fallas ocasionadas por vibraciones en la turbina. Se utiliza la teoría de *rough sets* aplicada a una red neuronal, entrenada con Back-Propagation. A través de la correcta clasificación de las vibraciones se detectan anomalías en la operación e inestabilidades en el funcionamiento de la turbina.

En (Wu, Hu, Hu, & Zhao, 2002) se aplica una red neuronal con estructura SOM (*Self-Organizing Map*) al diagnóstico de fallas en la turbina. Se estudian fallas como filtraciones en los tubos de vapor, bloqueo en los tubos de escape, exceso de aire en el escape, polución de los tubos, fallas en válvulas de drenaje, y presencia de agua en el escape.

Por otro lado, en (Tian, Zhang, & Wang, 2006) se diagnostican fallas en la turbina aplicando Algoritmos Genéticos en el proceso de aprendizaje de una red neuronal.

Finalmente, en (Xiang, Huang, Huang, & Hu, 2006) se emplea una red neuronal de Wavelets, basada en PSO (*Particle Swarm Optimization*) para la detección de fallas en la turbina de vapor. La red clasifica patrones de datos. El entrenamiento y actualización de parámetros dentro de la red se realiza con el algoritmo PSO gracias a la posibilidad de prescindir del gradiente de la función. Para finalmente clasificar la falla, se usa un método híbrido.

2.3.6. Diagnóstico para sistemas de refrigeración y ciclo de agua

Entre estas aplicaciones destaca lo realizado en (Hu, He, & Hu, 2003), en donde los autores aplican un método similar al empleado en (Wu et al., 2002), en que se utiliza la estructura SOM para la detección de fallas en el sistema de condensado.

En (Tan & Lim, 2004) se aplica una red neuronal tipo fuzzy ARTMAP con extracción de reglas para la detección de fallas en el circuito de agua que refrigera el condensador

del ciclo de vapor. Se monitorea la correcta transferencia de calor en el condensador, y un bloqueo en las tuberías del sistema producto de agentes externos.

3. MODELACIÓN DE UNA TURBINA DE GAS DE CENTRAL DE CC

3.1. Introducción

En una central de ciclo combinado, sin duda la parte más importante de ésta es la turbina a gas. El funcionamiento de toda la central está ligado a este componente, que provee de energía tanto a su generador como a la sección de vapor a través de la caldera recuperadora. A diferencia de la turbina de vapor, la central no puede operar sin la turbina a gas.

Debido a la importancia de la turbina de gas (TG) en una central de ciclo combinado, y a la importancia de las centrales de ciclo combinado en de la generación de electricidad a nivel mundial, se han desarrollado y se siguen desarrollando múltiples aplicaciones de monitorización, control y diagnóstico de problemas en la TG. En la mayoría de estas aplicaciones, se hace uso de un modelo del proceso. Existen múltiples estudios, incluso tesis e investigaciones completas, orientadas a formular modelos de TG tanto en centrales de CC como en otras aplicaciones industriales. Como resultados de estos estudios, existen una gran cantidad de modelos con distinto grado de complejidad, y orientados a distintas aplicaciones.

Como parte del desarrollo de esta tesis, se pretende adaptar un modelo para la turbina de gas de Central termoeléctrica de ciclo combinado San Isidro. Esta central está ubicada en el sector Lo Venecia, a 8 kilómetros al oeste de la ciudad de Quillota, en la V Región, Chile. Tiene una capacidad instalada de 379 MW, y es propiedad de Compañía Eléctrica San Isidro S.A., filial de ENDESA Chile. El ciclo de gas aporta 238 MW y el ciclo de vapor aporta 132 MW. La generación media anual de esta central alcanza los 1646 GWh (período 2001-2002), con un factor de carga de 49,58%. La potencia generada por esta central se suministra al Sistema Interconectado Central (SIC).

3.2. Modelos fenomenológicos

Los modelos fenomenológicos de la TG responden a sus ecuaciones de balance termodinámico. Las TG fueron diseñadas para trabajar según el Ciclo de Brayton, que contempla expansión y compresión a entropía constante, y transferencia de calor en ambos sentidos a presión constante (Boyce, 1982), (Harman, 1981).

El compresor de la turbina obtiene aire del ambiente y lo comprime elevando su presión y con esto su temperatura. El aire comprimido entra a la cámara de combustión mezclándose con el combustible. Esta mezcla se quema a altas presiones, elevando notablemente su temperatura. Los gases de salida calientes entran a la turbina, expandiéndose. Acá, la presión de los gases ejerce fuerza sobre los álabes, haciendo girar el eje. La expansión de los gases continúa hasta igualar su presión con la del ambiente.

La variable más crítica del proceso es la temperatura a la entrada de la turbina, ya que es uno de los factores más influyentes en la potencia generada por la unidad. Además, en ciclo simple, entre mayor es la temperatura, mayor es la eficiencia (Westinghouse, 1975).

Para obtener un modelo dinámico relativamente manejable de forma matemática, es necesario realizar algunas suposiciones (Boyce, 1982), (Ailer, Santa, Szederkényi, & Hangos, 2001):

(i) Las propiedades físico-químicas se asumen constantes en cada parte de la turbina, como el calor específico a presión constante y volumen constante, constante especifica del gas y el exponente adiabático.

(ii) Las pérdidas de calor (por transmisión, convección y radiación) son despreciadas.

(iii) Se asume un coeficiente de pérdida de presión constante σ_I en el ducto de entrada.

(iv) En el compresor, la tasa de flujo másico es constante ($\dot{m}_{Cin} = \dot{m}_{Cout} = \dot{m}_C$), y no hay efecto de almacenaje de energía ($U_2 = cte$).

(v) En la cámara de combustión, se asume un coeficiente de pérdida de presión σ_{Comb}, y una eficiencia de combustión η_{Comb} constantes, se desprecia la entalpía del combustible, y se asume que la cámara de combustión es una región perfectamente mezclada (volumen de control). Esto significa que se desarrolla un modelo concentrado de parámetros y dimensión, y el valor de las variables dentro del volumen de control es igual que a la salida.

(vi) En la turbina, la tasa de flujo másico es constante ($\dot{m}_{Tin} = \dot{m}_{Tout} = \dot{m}_T$), y no hay efecto de almacenaje de energía ($U_4 = cte$).

(vii) En el deflector de gas, se asume un coeficiente de pérdida de presión constante σ_N.

La eficiencia de la unidad de turbina de gas (considerando compresor, cámara de combustión y turbina), se define como (Boyce, 1982), (Harman, 1981):

$$\eta = \frac{P_{turb} - P_{comp}}{Q_{in}} \tag{3.1}$$

En donde P_{turb} es la potencia de la turbina, P_{comp} la potencia del compresor, y Q_{in} es el calor inyectado como energía de entrada. Considerando como punto 1 la entrada al compresor, el punto 2 la salida del compresor y entrada a cámara de combustión, el punto 3 la salida de la cámara de combustión y entrada a turbina, y el punto 4 la salida de la turbina, el trabajo en el compresor W_C está dado por:

$$W_C = \dot{m}_a(h_2 - h_1) \tag{3.2}$$

En donde h_i corresponde a la entalpía a la entrada y salida del compresor, y $\dot{m}_a$ es la razón de flujo másico del aire. El trabajo en la turbina W_T se obtiene como:

$$W_T = (\dot{m}_a - \dot{m}_f)(h_3 - h_4) \tag{3.3}$$

En donde $\dot{m}_f$ es la razón de flujo másico de combustible, y h_i son las entalpías en la entrada y salida. El calor que entra al sistema está dado por la expresión:

$$Q_{in} = \dot{m}_f \cdot \text{LHV} = (\dot{m}_a + \dot{m}_f)h_3 - \dot{m}_a h_2 \tag{3.4}$$

En donde LHV es el poder calorífico inferior del combustible usado. Usando estas expresiones, la eficiencia del sistema es:

$$\eta = \frac{W_T - W_C}{Q_{in}} = \frac{(\dot{m}_a - \dot{m}_f)(h_3 - h_4) - \dot{m}_a(h_2 - h_1)}{(\dot{m}_a + \dot{m}_f)h_3 - \dot{m}_a h_2} \tag{3.5}$$

Bajo las suposiciones y algunas simplificaciones prácticas (Boyce, 1982), se puede calcular la eficiencia del ciclo como

$$\eta = 1 - \frac{T_4}{T_3} = 1 - \frac{T_1}{T_2} \tag{3.6}$$

Aplicando las leyes de conservación de energía y balance de masas (Ailer et al., 2001), se pueden obtener las ecuaciones para la turbina que se presentan a continuación.

Ecuación de balance de masa:

$$\frac{dm}{dt} = \dot{m}_{in} - \dot{m}_{out} \tag{3.7}$$

Ecuación de balance de energía interna U, donde se incluyen los flujos de calor y trabajo.

$$\frac{dU}{dt} = \dot{m}_{in} i_{in} - \dot{m}_{out} i_{out} + Q + W \tag{3.8}$$

Se puede transformar la ecuación de conservación de energía anterior considerando la relación con la temperatura:

$$\frac{dU}{dt} = c_v \frac{d}{dt}(T \cdot m) = c_v T \frac{dm}{dt} + c_v m \frac{dT}{dt} \tag{3.9}$$

Uniendo las ecuaciones anteriores, se puede obtener una ecuación para la temperatura

$$\frac{dT}{dt} = \frac{\dot{m}_{in} i_{in} - \dot{m}_{out} i_{out} + Q + W - c_v T(\dot{m}_{in} - \dot{m}_{out})}{c_v m} \tag{3.10}$$

Aplicando la ecuación del gas ideal ($pV = mRT$), se puede encontrar una ecuación para la presión:

$$\frac{dp}{dt} = \frac{p}{m}(\dot{m}_{in} - \dot{m}_{out}) + \frac{p}{T}\left(\frac{\dot{m}_{in} i_{in} - \dot{m}_{out} i_{out} + Q + W - c_v T(\dot{m}_{in} - \dot{m}_{out})}{c_v m}\right) \tag{3.11}$$

$$\frac{dp}{dt} = \frac{R}{c_v V}(\dot{m}_{in} i_{in} - \dot{m}_{out} i_{out} + Q + W) \tag{3.12}$$

Ecuación de balance de energía mecánica del eje compresor-turbina:

$$\frac{dE_{eje}}{dt} = \dot{m}_T c_{pgas}(T_3 - T_4)\eta_{mec} - \dot{m}_C c_{paire}(T_2 - T_1) - 2\pi\frac{3}{50} n M_{carga} \tag{3.13}$$

En donde n es la velocidad de rotación del eje. Estas ecuaciones pueden transformarse en otras que incluyan variables que puedan ser medidas en el proceso. Las ecuaciones transformadas son las siguientes:

$$\frac{dm_{Comb}}{dt} = \dot{m}_C + \dot{m}_{fuel} - \dot{m}_T \tag{3.14}$$

$$\frac{p_3}{dt} = \frac{p_3}{m_{Comb}}\frac{dm_{Comb}}{dt} + \frac{dp_3}{T_3 c_{vmed} m_{Comb}}(\dot{m}_C c_{paire} T_2 - \dot{m}_T c_{pgas} T_3) + Q_f \eta_{Comb} \dot{m}_{fuel} - c_{vmed} T_3 \frac{dm_{Comb}}{dt} \tag{3.15}$$

$$\frac{dp_3}{dt} = \frac{R_{med}}{c_{vmed} V_{Comb}}(\dot{m}_C c_{paire} T_2 - \dot{m}_T c_{pgas} T_3 + Q_f \eta_{Comb} \dot{m}_{fuel}) \tag{3.16}$$

$$\frac{dn}{dt} = \frac{1}{4\pi^2 \Theta n}\left(\dot{m}_T c_{pgas}(T_3 - T_4)\eta_{mec} - \dot{m}_C c_{paire}(T_2 - T_1) - 2\pi\frac{3}{50} n M_{carga}\right) \tag{3.17}$$

Para completar el modelo, se requieren algunas relaciones algebraicas entre las variables. Dos ecuaciones aparecen directamente de las suposiciones 3, 5 y 7, para las presiones después de compresor y turbina:

$$p_2 = \frac{p_3}{\sigma_{Comb}} \tag{3.18}$$

$$p_4 = \frac{p_1}{\sigma_I \sigma_N} \tag{3.19}$$

La ecuación de gas ideal entrega una ecuación para la cámara de combustión:

$$T_3 = \frac{p_3 V_{Comb}}{m_{Comb} R_{med}} \tag{3.20}$$

Ahora, la temperatura total después del compresor se encuentra usando la eficiencia isentrópica η_C:

$$T_2 = T_1 \left(1 + \frac{1}{\eta_C} \left(\left(\frac{p_2}{p_1} \right)^{\frac{\kappa_{aire}-1}{\kappa_{aire}}} - 1 \right) \right) \tag{3.21}$$

De la misma forma, se puede encontrar la temperatura después de la turbina usando la eficiencia isentrópica η_T:

$$T_4 = T_3 \left(1 - \eta_T \left(1 - \left(\frac{p_4}{p_3} \right)^{\frac{\kappa_{gas}-1}{\kappa_{gas}}} \right) \right) \tag{3.22}$$

Las siguientes ecuaciones describen el flujo másico en compresor y turbina:

$$\dot{m}_C = const(1) q(\lambda_1) \frac{p_1}{\sqrt{T_1}} \tag{3.23}$$

$$\dot{m}_T = const(2) q(\lambda_3) \frac{p_3}{\sqrt{T_3}} \tag{3.24}$$

En estas ecuaciones, $q(\lambda_1)$ y $q(\lambda_3)$ (tasas de flujos másicos sin dimensión) se calculan de la siguiente forma:

$$q(\lambda_1) = f_1\left(\frac{n}{\sqrt{\frac{T_1}{288.15}}}, \frac{p_2}{p_1}\right) \tag{3.25}$$

$$q(\lambda_3) = f_2\left(const(3)\frac{n}{\sqrt{T_3}}, \frac{p_3}{p_4}\right) \tag{3.26}$$

Por otro lado, las eficiencias isentrópicas de compresor y de la turbina se calculan como sigue:

$$\eta_C = g_1\left(\frac{n}{\sqrt{\frac{T_1}{288.15}}}, q(\lambda_1)\right) \tag{3.27}$$

$$\eta_T = g_2\left(const(3)\frac{n}{\sqrt{T_3}}, \frac{p_3}{p_4}\right) \tag{3.28}$$

Los parámetros, constantes y funciones que aparecen en estas ecuaciones, pueden determinarse con ayuda de medidas y las características técnicas de la turbina y el compresor. (Ailer et al., 2001) proponen 2 tipos de aproximaciones, las que se describen a continuación.

Primera aproximación

Las tasas de flujo másico adimensionales se aproximan como función lineal de la velocidad del eje (y número de revoluciones corregido) y como función lineal de la relación de presiones:

$$q(\lambda_1) = a_1\frac{n}{\sqrt{\frac{T_1}{288.15}}}\frac{p_2}{p_1} + a_2\frac{n}{\sqrt{\frac{T_1}{288.15}}} + a_3\frac{p_2}{p_1} + a_4 \tag{3.29}$$

$$q(\lambda_3) = c_1(const(3)\frac{n}{\sqrt{T_3}})\frac{p_3}{p_4} + c_2(const(3)\frac{n}{\sqrt{T_3}}) + c_3\frac{p_3}{p_4} + c_4 \tag{3.30}$$

La eficiencia isentrópica del compresor se calcula como función lineal de la velocidad del eje, de la relación de presiones y de la tasa de flujo másico sin dimensión:

$$\eta_C = b_1 \frac{n}{\sqrt{\frac{T_1}{288.15}}} q(\lambda_1) + b_2 \frac{n}{\sqrt{\frac{T_1}{288.15}}} + b_3 q(\lambda_1) + b_4 \tag{3.31}$$

$$\eta_T = d_1 (const(3) \frac{n}{\sqrt{T_3}}) \frac{p_3}{p_4} + d_2 (const(3) \frac{n}{\sqrt{T_3}}) + d_3 \frac{p_3}{p_4} + d_4 \tag{3.32}$$

Segunda aproximación

Las tasas de flujo másico adimensionales se aproximan como función lineal de la velocidad del eje (y número de revoluciones corregido) y como función cuadrática de la relación de presiones:

$$q(\lambda_1) = a_1 \frac{n}{\sqrt{\frac{T_1}{288.15}}} \left(\frac{p_2}{p_1}\right)^2 + a_2 \left(\frac{p_2}{p_1}\right)^2 + a_3 \frac{n}{\sqrt{\frac{T_1}{288.15}}} \frac{p_2}{p_1} + a_4 \frac{p_2}{p_1} + a_5 \frac{n}{\sqrt{\frac{T_1}{288.15}}} + a_6 \tag{3.33}$$

$$\begin{aligned} q(\lambda_3) = c_1 (const(3) \frac{n}{\sqrt{T_3}}) \left(\frac{p_3}{p_4}\right)^2 + c_2 \left(\frac{p_3}{p_4}\right)^2 + c_3 (const(3) \frac{n}{\sqrt{T_3}}) \frac{p_3}{p_4} \\ + c_4 \frac{p_3}{p_4} + c_5 (const(3) \frac{n}{\sqrt{T_3}}) + c_6 \end{aligned} \tag{3.34}$$

La eficiencia isentrópica del compresor se calcula como función lineal de la velocidad del eje (y número de revoluciones corregido), y como función cuadrática de la relación de presiones y de la tasa de flujo másico adimensional:

$$\eta_C = b_1 \frac{n}{\sqrt{\frac{T_1}{288.15}}} q(\lambda_1)^2 + b_2 q(\lambda_1)^2 + b_3 \frac{n}{\sqrt{\frac{T_1}{288.15}}} q(\lambda_1) + b_4 q(\lambda_1) + b_5 \frac{n}{\sqrt{\frac{T_1}{288.15}}} + b_6 \tag{3.35}$$

$$\begin{aligned} \eta_T = d_1 (const(3) \frac{n}{\sqrt{T_3}}) \left(\frac{p_3}{p_4}\right)^2 + d_2 \left(\frac{p_3}{p_4}\right)^2 + d_3 (const(3) \frac{n}{\sqrt{T_3}}) \frac{p_3}{p_4} \\ + d_4 \frac{p_3}{p_4} + d_5 (const(3) \frac{n}{\sqrt{T_3}}) + d_6 \end{aligned} \tag{3.36}$$

Los parámetros a_i, b_i, c_i y d_i deben identificarse a partir de datos de operación del proceso.

Un modelo alternativo puede encontrarse en (Ordys et al., 1994). Este modelo considera otras variables y las relaciona exclusivamente a través de ecuaciones algebraicas.

3.3. Modelos simplificados

Otro tipo de modelos ampliamente usados y comentados en la literatura son los modelos simplificados. A pesar que el modelo descrito en la sección anterior también se ha simplificado a través de las suposiciones realizadas, es suficientemente complejo como para no obviar ninguna de las características importantes de la dinámica de la turbina.

Estos modelos simplificados están en general orientados a estudios de sistemas de potencia, en los que se debe lidiar con el desarrollo de aplicaciones y simulaciones en las que aparecen involucradas una gran cantidad de generadoras.

Como ejemplo, en (Q. Zhang & So, 2000) se desarrolla un modelo dinámico de una central de CC para estudios de estabilidad de sistemas de potencia. El modelo de turbina de gas es un modelo dinámico simplificado lineal, basado en plantas de primer orden para cada componente de la turbina. La estructura de este modelo para turbina de gas (junto con el resto de la central) se muestra en la figura 3.1.

Se aprecia que los principales componentes de la turbina como el compresor, se modelan como funciones de transferencia de primer orden, con 1 sola entrada y 1 sola salida. Es evidente acá la simplificación, considerando la cantidad de variables que influyen y que se ven afectadas por la dinámica de cada componente. En este caso, las funciones F_i representan funciones no lineales. La salida de torque de la TG está representada por F_1, dada por la siguiente relación:

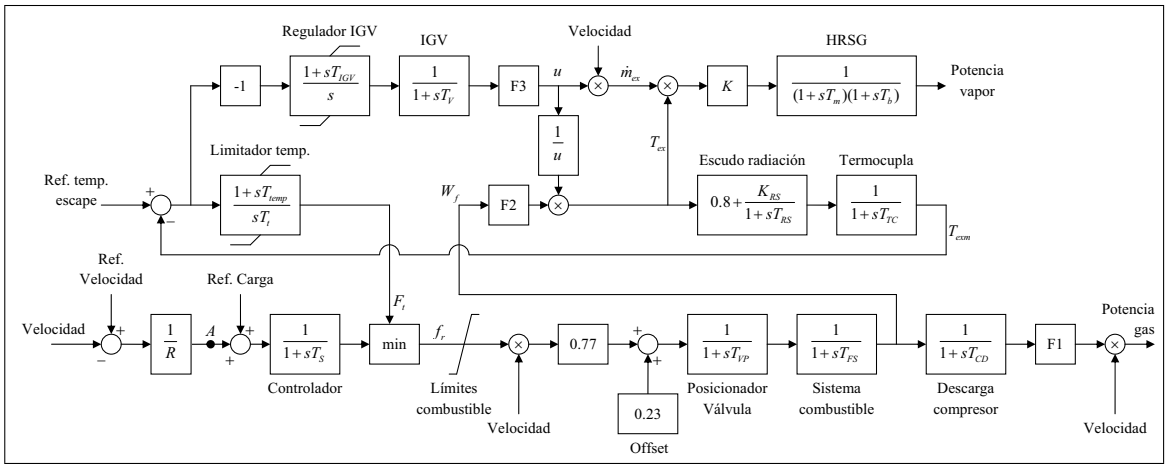

FIGURA 3.1. Modelo simplificado de una central de ciclo combinado.

$$F_1 : \text{Torque GT} = 1.3(W_f - 0.23) + 0.5(1 - \Omega)$$

Siendo W_f el flujo desde el compresor, y Ω la velocidad de la turbina. Las otras dos funciones no lineales son

$$F_2 : T_{ex} = T_r - 453(\Omega^2 - 4.21\Omega + 4.42)0.82(1 - W_f) + 722(1 - \Omega) + 1.94(1 - L_{igv})$$

$$F_3 : F_3 = (L_{igv})^{0.257}$$

Al ser éste un modelo orientado a estudios de sistemas de potencia, todas las variables se trabajan en unidades p.u. Otros modelos de este tipo se pueden encontrar en (Hung, Jul 1991); aquí la modelación de los componentes es de primer orden, pero se agrega más detalle en el sistema de control de la planta. En (Hannett & Khan, 1993) la turbina misma se modela como una planta de primer orden, la cámara de combustión como un retardo, y solo posee una función no lineal. En (Maslo & Andel, 2001) se usa un modelo muy simple de TG considerando plantas de primer orden para el compresor y sistema de combustible y una curva ajustada en base a datos para la TG; además se incluyen modelos simples de otros tipos de plantas. Otro enfoque que incluye un poco más de detalle se puede revisar en (Working Group, Aug 1994) y (Suzaki, Kawata, Sekoguchi, & Goto, 2000), en donde se

incluyen algunas ecuaciones no lineales dinámicas dentro de los principales componentes de la TG. El esquema usado para la TG en estos dos casos se resume en la figura 3.2.

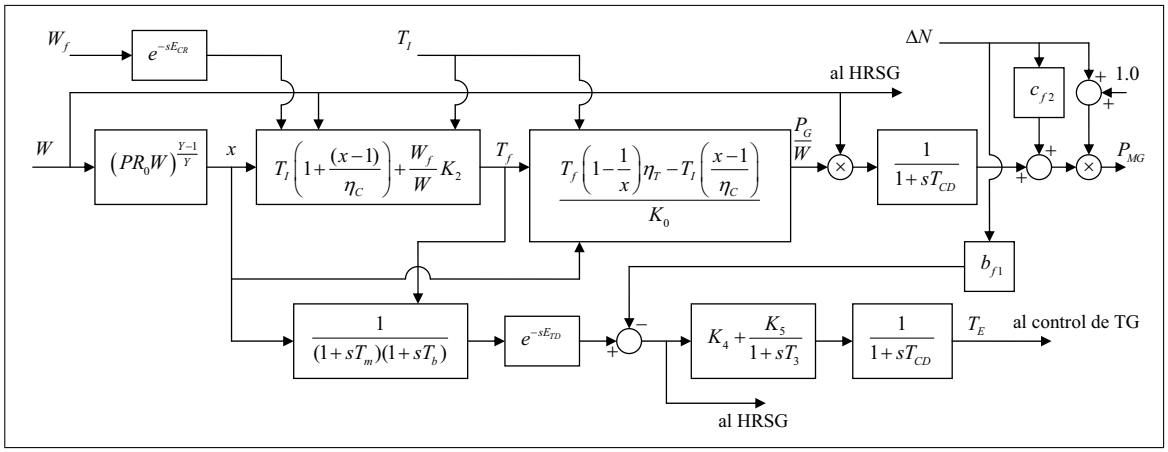

FIGURA 3.2. Modelo simplificado de la TG.

Como puede apreciarse en dicha figura, se incluyen algunas ecuaciones algebraicas para relacionar un mayor número de variables en este modelo, pero la dinámica del sistema sigue dada por una planta de primer orden. A esto se le agrega la dinámica del escudo de radiación y de la termocupla para enviar la señal medida al controlador de la TG.

3.4. Modelos empíricos

Otro enfoque en la modelación de un sistema dinámico corresponde a los modelos empíricos. En estos modelos se busca ajustar los datos a un tipo de estructura determinado, lineal o no lineal. Dicha estructura no tiene relación directa con el comportamiento físico del sistema a modelar, ni tampoco requiere conocimiento previo sobre relaciones matemáticas entre variables.

Estos modelos se utilizan ampliamente cuando no existe un conocimiento suficiente acerca del sistema en estudio, o cuando el proceso resulta muy difícil de modelar mediante relaciones físicas, químicas, termodinámicas, mecánicas o de otra índole fenomenológica.

En las aplicaciones de diagnóstico que utilizan modelos, debido a la precisión que deben tener los modelos utilizados, en general se emplean modelos empíricos ajustados a

partir de los datos disponibles, debido a la gran dificultad de encontrar modelos fenomenológicos con el nivel de detalle suficiente, y que sean manejables computacionalmente.

En (Saez & Cipriano, 1999a) se revisan algunas aplicaciones de estos modelos. En particular, se presenta una modelación difusa en base a modelos NARX de una central de CC; en (Roizman & Davydov, Jul 1999) se usan modelos difusos para realizar predicciones en variables de interés para realizar diagnóstico de fallas; y en (Palade, Patton, Uppal, Quevedo, & Daley, 2002) se usa un modelo neuro-difuso de la planta para generar residuos y realizar diagnóstico de fallas.

Se han formulado diversas técnicas para identificar modelos a partir de los datos de entrada y salida de la central. En esta sección se mencionan algunos de estos métodos, los que serán comparados más adelante. Básicamente, si y es la salida y x es la entrada del modelo, la identificación consiste en determinar una función f y un conjunto de parámetros θ tal que

$$y = f(x, \theta) + e \tag{3.37}$$

En donde e representa un error de ajuste entre la función f y la salida y.

3.4.1. Modelos ARX

El modelo ARX es uno de los modelos mas sencillos que existen. Considerando y como la salida y u la entrada del proceso a modelar, y e ruido blanco gaussiano, el modelo ARX plantea:

$$A(q)y(t) = B(q)u(t) + e(t) \tag{3.38}$$

Equivalentemente:

$$y(k) = \sum_{i=1}^{N} a_i y(t-i) + \sum_{i=0}^{M} b_i u(t-i) + e(t) \tag{3.39}$$

Formando el vector $x(t)$ según la siguiente estructura

$$x(t) = [y(t-1), \ldots, y(t-N), u(t), \ldots, u(t-M)]^T \tag{3.40}$$

y el vector θ con los parámetros del modelo según la siguiente estructura

$$\theta = [a_1, \ldots, a_N, b_0, \ldots, b_N]^T \tag{3.41}$$

el modelo se plantea como

$$y(t) = \theta^T x(t) + e(t) \tag{3.42}$$

Considerando el vector de parámetros θ como un vector constante en el tiempo, se obtiene un modelo lineal e invariante. Dado un conjunto de mediciones, se puede construir un set de r datos de la siguiente forma:

$$Y(r) = [y^1, y^2, \ldots, y^r]^T \tag{3.43}$$

$$X(r) = [x^1, x^2, \ldots, x^r]^T \tag{3.44}$$

El ajuste del modelo consiste en seleccionar un vector de parámetros θ que minimice la suma de los errores cuadráticos cometidos en la aproximación de cada pareja de datos

$$J = \sum_{i=1}^{r} (y^i - \theta^T(x^i))^2 \tag{3.45}$$

Si consideramos

$$\epsilon_i = y^i - \theta^T(x^i) \tag{3.46}$$

y con ello formamos el vector

$$E(r) = [\epsilon_1, \epsilon_2, \ldots, \epsilon_r]^T \tag{3.47}$$

se puede ver que

$$E = Y - \theta^T X \tag{3.48}$$

y con esto, la función objetivo de minimización es

$$J(\theta) = \sum_{i=1}^{r} (y^i - \theta^T(x^i))^2 = E^T E \tag{3.49}$$

en donde θ es la variable a optimizar. Es posible demostrar que la solución que minimiza esta cantidad es

$$\hat{\theta} = (X^T X)^{-1} X^T Y \tag{3.50}$$

Si se desea dar mayor importancia a ciertos datos durante el ajuste (por ejemplo, incrementar el peso de los datos más recientes contra los más antiguos), puede introducirse una matriz diagonal de pesos Λ, de forma que la función a minimizar es

$$J(\theta) = \sum_{i=1}^{r} \lambda_{ii} (y^i - \theta^T(x^i))^2 = E^T \Lambda E \tag{3.51}$$

En este caso, la solución al problema de ajuste es

$$\hat{\theta} = (X^T \Lambda X)^{-1} X^T \Lambda Y \tag{3.52}$$

3.4.2. Modelos de espacio de estados

El modelo en Espacio de Estados considera un sistema de la forma:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + w(t) \\ y(t) &= Cx(t) + Du(t) + v(t) \end{aligned} \tag{3.53}$$

Las matrices A, B, C y D son parametros invariantes en el tiempo, y $w(t)$ y $v(t)$ representan perturbaciones no medidas, asumidas como ruido blanco. En este caso, el problema consiste en determinar el orden del sistema n, y encontrar las matrices A, B, C y D, que

son los parámetros del modelo. El vector columna y tiene dimensión p, y el vector de entradas u tiene dimensión m.

Considerando la ecuación de salida del modelo, se puede encontrar que (Ljung, 1999):

$$\begin{aligned}
y(t+k) =& Cx(t+k) + Du(t+k) + v(t+k) \\
=& C(Ax(t+k-1) + Bu(t+k-1) + w(t+k-1)) + Du(t+k) + v(t+k) \\
=& CAx(t+k-1) + CBu(t+k-1) + Cw(t+k-1) + Du(t+k) + v(t+k) \\
=& \dots \\
=& CA^k x(t) + CA^{k-1}Bu(t) + CA^{k-2}Bu(t+1) + \dots + CBu(t+k-1) \\
& + CA^{k-1}w(t) + CA^{k-2}w(t+1) + \dots + Cw(t+k-1) \\
& + Du(t+k) + v(t+k)
\end{aligned} \tag{3.54}$$

Ahora, definiendo los vectores

$$Y_r(t) = \begin{bmatrix} y(t) \\ y(t+1) \\ \vdots \\ y(t+r-1) \end{bmatrix}, \qquad U_r(t) = \begin{bmatrix} u(t) \\ u(t+1) \\ \vdots \\ u(t+r-1) \end{bmatrix} \tag{3.55}$$

la ecuación (3.54) puede ser agrupada como

$$Y_r(t) = O_r x(t) + S_r U_r(t) + V(t) \tag{3.56}$$

en donde

$$O_r = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{r-1} \end{bmatrix}, \qquad S_r(t) = \begin{bmatrix} D & 0 & \cdots & 0 & 0 \\ CB & D & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ CA^{r-2}B & CA^{r-3}B & \cdots & CB & D \end{bmatrix} \tag{3.57}$$

y el bloque k-ésimo que compone $V(t)$ está dado por

$$V_k(t) = CA^{k-2}w(t) + CA^{k-3}w(t+1) + \ldots + Cw(t+k-2) + v(t+k-1) \quad (3.58)$$

La idea principal es estimar la matriz O_r, para obtener directamente las matrices A y C. Consideremos que existen mediciones para $y(t)$ y $u(t)$ para $t = 1, \ldots, N + r - 1$. Agrupando las N ecuaciones formadas según la ecuación (3.56), se pueden formar los siguientes vectores:

$$\begin{aligned} Y &= \begin{bmatrix} Y_r(1) & Y_r(2) & \ldots & Y_r(N) \end{bmatrix} \\ X &= \begin{bmatrix} x(1) & x(2) & \ldots & x(N) \end{bmatrix} \\ U &= \begin{bmatrix} U_r(1) & U_r(2) & \ldots & U_r(N) \end{bmatrix} \\ V &= \begin{bmatrix} V(1) & V(2) & \ldots & V(N) \end{bmatrix} \end{aligned} \quad (3.59)$$

La ecuación (3.56) puede reescribirse como

$$Y = O_r X + S_r U + V \quad (3.60)$$

Definiendo la matriz

$$U^{\perp} = I - U^T(UU^T)^{-1}U \quad (3.61)$$

se puede comprobar que $UU^{\perp} = 0$. Multiplicando por la derecha la ecuación (3.60) por $U^{\perp}$, se elimina el término U y se obtiene:

$$YU^{\perp} = O_r X U^{\perp} + VU^{\perp} \quad (3.62)$$

Consideremos ahora una matriz Φ de dimensiones $s \times N$ formada como

$$\Phi = \begin{bmatrix} \varphi_s(1) & \varphi_s(2) & \ldots & \varphi_s(N) \end{bmatrix} \quad (3.63)$$

Multiplicando la ecuación (3.62) por Φ^T y normalizando por N se obtiene

$$\frac{1}{N}YU^{\perp}\Phi^T = O_r\frac{1}{N}XU^{\perp}\Phi^T + \frac{1}{N}VU^{\perp}\Phi^T$$
$$G = O_rT_N + V_N \tag{3.64}$$

en donde se han definido las matrices (Ljung, 1999):

$$G = \frac{1}{N}YU^{\perp}\Phi^T, \qquad T_N = \frac{1}{N}XU^{\perp}\Phi^T, \qquad V_N = \frac{1}{N}VU^{\perp}\Phi^T \tag{3.65}$$

Si $V(t)$ y $\varphi_s(t)$ son no correlacionados, se cumple que

$$\lim_{N\to\infty} V_N = \lim_{N\to\infty}\frac{1}{N}VU^{\perp}\Phi^T = 0$$
$$\lim_{N\to\infty} T_N = \lim_{N\to\infty}\frac{1}{N}XU^{\perp}\Phi^T = T \tag{3.66}$$

y T es de rango completo n. Entonces:

$$G = O_rT + E_N \tag{3.67}$$

en donde $E_N \to 0$ cuando $N \to \infty$. Como $V(t)$ está formado por componentes de ruido blanco, $\varphi_s(t)$ puede definirse como

$$\varphi_s(t) = \begin{bmatrix} y(t-1) \\ \vdots \\ y(t-s_1) \\ u(t-1) \\ \vdots \\ u(t-s_2) \end{bmatrix} \tag{3.68}$$

Si se realiza la descomposición en valores singulares de G, se obtiene:

$$G = USV^T \tag{3.69}$$

Con $U^TU = I$, $V^TV = I$ y S matriz diagonal. Si G es de rango n, solo las primeras n columnas de la matriz diagonal S serán distinas de 0. Despreciando las columnas con

valores singulares 0, queda

$$G = USV^T = U_1 S_1 V_1^T \tag{3.70}$$

donde U_1, S_1 y V_1 corresponden a las matrices considerando la eliminación de las columnas sobrantes de S. Usando una aproximación para G según la ecuación (3.67), se obtiene

$$O_r T \approx U_1 S_1 V_1 \tag{3.71}$$

Multiplicando por V_1^T, se obtiene

$$O_r T V_1^T = O_r \tilde{T} \approx U_1 S_1 \tag{3.72}$$

Como T y V_1 son de rango completo, entonces $\tilde{T}$ también lo es. Debido a que $\tilde{T}$ es una matriz de rango completo, puede interpretarse como una matriz de cambio de base del espacio columna de O_r. Con esto se obtiene una aproximación

$$\hat{O}_r \approx U_1 S_1 \approx U_1 \approx U_1 R \tag{3.73}$$

con R cualquier matriz de rango completo n. Si antes de calcular la descomposición de valores singulares de G se premultiplica esta matriz por matrices W_1 y W_2, por la izquierda y derecha respectivamente para aumentar la estabilidad (formando $\hat{G} = W_1 G W_2$), se obtiene

$$\hat{O}_r \approx W_1^{-1} U_1 R \tag{3.74}$$

Es decir, se puede obtener la estimación de la matriz O_r a través de la descomposición en valores singulares de G, formada según la ecuación (3.65). A través de $\hat{O}_r$ se obtienen estimaciones de A y C de forma directa según su definición en (3.57).

Ahora, consideremos el modelo definido en (3.53). De la primera ecuación, y reemplazando la estimaciones obtenidas $\hat{A}$ y $\hat{C}$, se puede despejar:

$$\begin{aligned} \hat{x}(t+1) &= \hat{A}x(t) + Bu(t) \\ qx(t) - \hat{A}x(t) &= Bu(t) \\ x(t) &= (qI - \hat{A})^{-1}Bu(t) \end{aligned} \tag{3.75}$$

en donde q representa el operador de adelanto en el tiempo. Reemplazando esto en la ecuación para la salida, se obtiene una ecuación para la predicción de la salida

$$\hat{y}(t) = \hat{C}(qI - \hat{A})^{-1}Bu(t) + Du(t) \tag{3.76}$$

Luego, se tiene un problema similar al expuesto en la sección 3.4.1, el que puede ser resuelto en base a la minimización del error cuadrático de la predicción. Considerando la ecuación en $t = 0$, se puede encontrar la estimación para el estado inicial x_0, de la forma

$$\hat{y}(0) = \hat{C}(qI - \hat{A})^{-1}Bu(0) + Du(0) + \hat{C}(qI - \hat{A})^{-1}x_0 \tag{3.77}$$

De acá se obtienen finalmente la estimación para B y D (y el estado inicial x_0), que completa la estimación de los parámetros para el modelo.

Una elección tradicional de las matrices R, W_1 y W_2 es (Overschee & Moor, 1994):

$$R = I, \qquad W_1 = I, \qquad W_2 = \left(\frac{1}{N}\Phi U^{\perp}\Phi^T\right)^{-1}\Phi \tag{3.78}$$

Esta elección se hace evidente al estimar los estados X y los términos de ruido, como se muestra en el Anexo A.

3.4.3. Modelos de redes neuronales

Una red neuronal es un modelo matemático que se basa en la estructura de las neuronas del cerebro humano, integrando sus capacidades naturales de aprendizaje y almacenamiento de relaciones (Omidvar & Elliott, 1997). Existe un gran número de tipos de redes neuronales, que varían en el mecanismo de operación y grados de sofisticación.

Las redes neuronales consisten en general en un número de elementos de proceso (neuronas) interconectadas. La forma en que se conectan las neuronas y la naturaleza de la conexión, determinan la estructura de la red. Por otro lado, la forma en que se determinan los parámetros de las conexiones y de las propias neuronas, es gobernada por un algoritmo de entrenamiento. De esta forma las redes neuronales se pueden clasificar según sus estructuras y algoritmos de aprendizaje.

3.4.3.1. Red de perceptrones multicapa

Una de las redes neuronales más simples y más difundidas corresponde a la red de perceptrones multi-capa (*multi-layer perceptron*). La estructura de esta red se muestra en la figura 3.3.

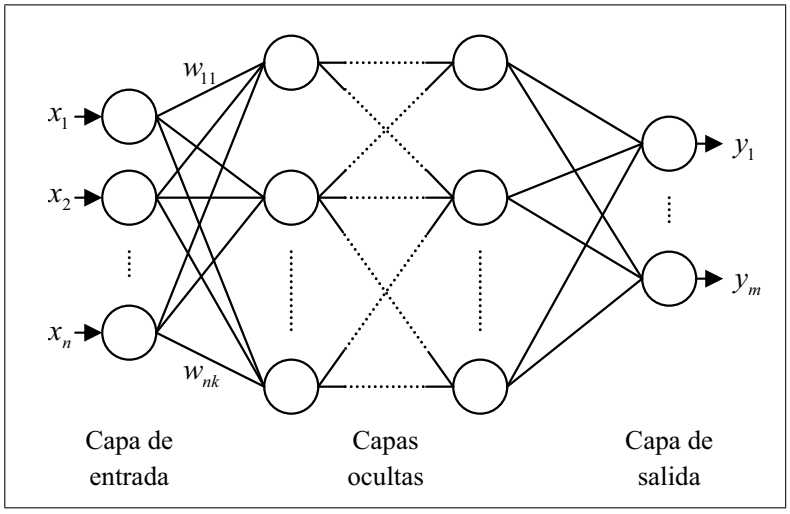

FIGURA 3.3. Estructura de una red neuronal tipo perceptron multicapa.

La salida de cada neurona se obtiene a partir de una función que recibe como argumento la suma ponderada de sus entradas más una constante de desvío (llamada función de

activación). Es decir, si se denomina h_j la salida de la neurona j, ésta se calcula como:

$$h_j = f\left(\sum_{i=1}^{n} w_{ij}x_i + \theta_j\right) \tag{3.79}$$

La función f puede ser lineal o no lineal. Funciones comúnmente usadas son las llamadas sigmoidales, por su forma en "S", como por ejemplo:

$$f_1(u) = \frac{1}{1+e^{-u}}, \qquad f_2(u) = \tanh(u), \qquad f_3(u) = \frac{u}{1+|u|} \tag{3.80}$$

Para determinar los parámetros de la red (los pesos de cada conexión, y el valor de los desvíos), el algoritmo más usado corresponde al denominado *backpropagation*. Este algoritmo es un proceso iterativo de entrenamiento basado en un descenso según el gradiente.

Consideremos una red de n entradas y 1 salida, con sólo una capa oculta con $j = 1, \ldots, p$ funciones de activación sigmoidales de la forma de $f_1(u)$ en la ecuación (3.80), y funcion de activación lineal en la neurona de salida. Definiendo

$$S_j = \sum_{i=1}^{n} w_{ij}x_i + \theta_j \tag{3.81}$$

la salida para cada neurona de la capa oculta queda determinada por

$$h_j = \frac{1}{1+e^{-S_j}} \tag{3.82}$$

La salida de la neurona de la capa de salida, y salida de la red, queda definida por

$$\hat{y} = \sum_{j=1}^{p} v_j h_j \tag{3.83}$$

Considerando un conjunto de m datos de entrenamiento (x, y), se busca minimizar el error cuadrático medio de aproximación (MSE) definido por

$$E = \frac{1}{2m}\sum_{k=1}^{m}(y_k - \hat{y_k})^2 \tag{3.84}$$

Los nuevos parámetros en cada iteración quedan definidos como

$$w_{ij}(t+1) = w_{ij}(t) - \eta \frac{\partial E}{\partial w_{ij}} \tag{3.85}$$

$$v_j(t+1) = v_j(t) - \eta \frac{\partial E}{\partial v_j} \tag{3.86}$$

$$\theta_j(t+1) = \theta_j(k) - \eta \frac{\partial E}{\partial \theta_j} \tag{3.87}$$

en donde η es un parámetro llamado "de aprendizaje", $0 < \eta < 1$, que se sintoniza de forma que el algoritmo converja. Para este caso, se tiene que

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial \hat{y}_k}\frac{\partial \hat{y}_k}{\partial h_j}\frac{\partial h_j}{\partial S_j}\frac{\partial S_j}{\partial w_{ij}} = -\frac{1}{m}\sum_{k=1}^{m}(y_k - \hat{y}_k)v_j h_j(1-h_j)x_i \tag{3.88}$$

$$\frac{\partial E}{\partial v_j} = \frac{\partial E}{\partial \hat{y}_k}\frac{\partial \hat{y}_k}{\partial v_j} = -\frac{1}{m}\sum_{k=1}^{m}(y_k - \hat{y}_k)h_j \tag{3.89}$$

$$\frac{\partial E}{\partial \theta_j} = \frac{\partial E}{\partial \hat{y}_k}\frac{\partial \hat{y}_k}{\partial h_j}\frac{\partial h_j}{\partial S_j}\frac{\partial S_j}{\partial \theta_j} = -\frac{1}{m}\sum_{k=1}^{m}(y_k - \hat{y}_k)v_j h_j(1-h_j) \tag{3.90}$$

En general, para la mayoria de los problemas no triviales, la velocidad de convergencia del algoritmo es muy lenta, requiriendo muchas iteraciones o "épocas". Para incrementar la velocidad del algoritmo, se introduce un término adicional denominado "momentum", quedando la ecuación de actualización de la siguiente forma:

$$w_{ij}(t+1) = w_{ij}(t) - \eta \frac{\partial E}{\partial w_{ij}} + \mu\left(w_{ij}(t) - w_{ij}(t-1)\right) \tag{3.91}$$

Una mejora sustancial en la velocidad de convergencia representa el algoritmo de Levenberg-Marquardt (Marquardt, 1963). En este algoritmo, se considera un vector de parámetros β que agrupa todos los parámetros de la red. En este caso:

$$\beta = [w_{11}, \ldots, w_{np}, v_1, \ldots, v_p, \theta_1, \ldots, \theta_p] = [\beta_1, \ldots, \beta_r] \tag{3.92}$$

Se considera un vector δ que representa la actualización de los parámetros. Se usa también la aproximación de primer orden

$$\hat{y}_k(x, \beta + \delta) \approx \hat{y}_k(x, \beta) + \nabla \hat{y}_k \delta \tag{3.93}$$

En donde $\nabla \hat{y}_k$ representa el gradiente de $\hat{y}_k$ respecto al vector de parámetros. El error cuadrático medio queda en función de δ de la forma

$$\begin{aligned} E(\delta) &= \frac{1}{2m} \sum_{k=1}^{m} \left(y_k - \hat{y}_k(x, \beta + \delta)\right)^2 \\ &= \frac{1}{2m} \sum_{k=1}^{m} \left(y_k - \hat{y}_k(x, \beta) - \nabla \hat{y}_k \delta\right)^2 \\ &= \frac{1}{2m} \left[y - \hat{y} - \frac{\partial(y_1, \ldots, y_m)}{\partial(\beta_1, \ldots, \beta_r)} \delta\right]^T \left[y - \hat{y} - \frac{\partial(y_1, \ldots, y_m)}{\partial(\beta_1, \ldots, \beta_r)} \delta\right] \end{aligned} \tag{3.94}$$

en donde

$$J = \frac{\partial(y_1, \ldots, y_m)}{\partial(\beta_1, \ldots, \beta_r)} \tag{3.95}$$

representa la matriz jacobiana de las m funciones y_k con respecto a los parámetros. En el óptimo, se debería cumplir que

$$\nabla E(\delta) = -\frac{1}{m} \left[y - \hat{y} - J\delta\right]^T J = 0 \tag{3.96}$$

De esta ecuación se puede despejar δ de la forma

$$\delta = (J^T J)^{-1} J^T (y - \hat{y}) \tag{3.97}$$

El nuevo vector de parámetros en cada iteración $\beta(t+1)$ se obtiene de $\beta(t+1) = \beta(t) + \delta$. Debido a que la matriz $J^T J$ puede ser no invertible, el algoritmo de Levenberg-Marquardt contempla un parámetro λ tal que

$$\delta = (J^T J + \lambda I)^{-1} J^T (y - \hat{y}) \tag{3.98}$$

Donde I es la matriz identidad. De esta forma, la matriz $J^T J + \lambda I$ será siempre invertible para algún λ. Cuando λ tiende a infinito, el algoritmo tiende a un algoritmo de descenso

en la dirección del gradiente. Los autores recomiendan usar un parámetro inicial $\lambda = \lambda_0$ de manera que el algoritmo converja, y un factor v de tal manera de que en cada iteración el parámetro λ se actualiza como $\lambda = \lambda/v$ si el error disminuye, y $\lambda = \lambda v$ en caso de que el error aumente.

3.4.3.2. Red de neuronas de base radial

Otro tipo de estructura de red neuronal muy utilizada para la formulación de modelos es la red de neuronas de base radial o *radial basis function networks* (RBFN). La estructura típica de esta red consta de una capa de entrada, una capa oculta con funciones de activación radiales en las neuronas, y una capa de salida con función de activación lineal. La función de activación se denomina "radial" ya que tiene como argumento la distancia de la entrada a un centro, bajo una norma determinada. Si h_j es la salida de la neurona $j = 1, \ldots, p$ de la capa oculta, ésta queda definida por

$$h_j = f\left(||x - c_j||\right) \tag{3.99}$$

En donde $x = [x_1, \ldots, x_n]$ y c_j es el centro de la neurona j. Usualmente se utiliza la norma euclideana como medida, y la función gaussiana como función de activación. En este caso

$$h_j = e^{-||x-c_j||^2/\beta^2} \tag{3.100}$$

y la salida será

$$\hat{y} = \sum_{j=1}^{p} v_j h_j = \sum_{j=1}^{p} v_j e^{-||x-c_j||^2/\beta^2} \tag{3.101}$$

Esta estructura se denomina "no normalizada". Una estructura normalizada utiliza una función de activación en la capa oculta de la forma

$$h_j = \frac{f\left(||x - c_j||\right)}{\sum_{j=1}^{p} f\left(||x - c_j||\right)} \tag{3.102}$$

Existen ventajas de la estructura normalizada en casos de aproximación de datos con componentes estocásticos.

Una red neuronal RBF es un aproximador universal, es decir, que dado un número suficiente de neuronas en la capa oculta, puede aproximar cualquier función contínua con un grado arbitrario de precisión (Wasserman, 1993).

Si se fijan los centros c_j y el parámetro β por cada neurona, se puede formar un vector $\theta = [v_1, \ldots, v_p]^T$, y un vector $H(x) = [h_1(x), \ldots, h_p(x)]^T$, de forma que

$$\hat{y} = \theta^T H(x) \tag{3.103}$$

La estimación de parámetros en este caso se realiza según el procedimiento de la sección 3.4.1. Para estimar también los parámetros c_j y β en base a la información disponible, puede emplearse el algoritmo de back-propagation o el algoritmo de Levemberg-Marquardt descrito en la sección 3.4.3.1.

3.4.4. Modelos difusos

Un modelo difuso es un tipo de modelo que se basa en descripciones lingüísticas de las relaciones existentes entre variables. Por ejemplo, cuando u *sube mucho*, entonces y *baja poco*. Un modelo o sistema difuso contiene los siguientes elementos:

(i) Una *base de reglas*, que contiene la cuantificación del conocimiento sobre el sistema.

(ii) Un *mecanismo de inferencias*, que intenta emular las decisiones que tomaría un experto humano con ciertas condiciones dadas.

(iii) Una *interface de fuzzificación*, que convierte las entradas en variables que el mecanismo de inferencias pueda manejar

(iv) Una *interfade de defuzzificación*, que convierte las conclusiones del mecanismo de inferencias en valores equivalentes al de las variables de salida.

Las reglas en un modelo difuso poseen la siguiente estructura

$$\text{IF } u_1 \text{ is } A_1 \text{ and}, \ldots, \text{and } u_n \text{ is } A_n \text{ THEN } y_q \text{ is } B_q \tag{3.104}$$

en donde la expresión lingüística u_j is A_j es cuantificada a través de un conjunto difuso $\tilde{A}_j$.

La fuzzificación de valores se hace a través de las denominadas funciones de pertenencia, que asocian valores numéricos de las variables a conjuntos lingüisticos según el grado de pertenencia que estos valores tengan. Dicho grado de pertenencia se evalúa mediante operaciones matemáticas sobre las funciones de pertenencia. Éstas funciones tienen diversas formas, y se denotarán por μ_i para la función de pertenencia i.

Para determinar las conclusiones, se usan diversos métodos. Uno de ellos es el Centro de Gravedad, que pretende calcular el centro de gravedad del área bajo las funciones de pertenencia μ_i, considerando como área para una entrada u el valor entre 0 y $\mu_i(u)$. De esta forma

$$y = \frac{\sum_i b_i \int \mu_i}{\sum_i \int \mu_i} \tag{3.105}$$

en donde b_i representa la distancia del centro de gravedad de cada $\mu_i(u)$ al eje vertical. Este método tiene la desventaja que en caso de que la función de pertenencia tenga área infinita, no se puede realizar el cálculo. Además, en general es complejo realizar el cálculo de $\int \mu_i$.

Otra forma alternativa de realizar la defuzificación es a través del centro promedio, en donde

$$y = \frac{\sum_i b_i \mu_i}{\sum_i \mu_i} \tag{3.106}$$

Las funciones más ampliamente usadas por su simplicidad en el primer caso, y en el segundo caso por ser funciones contínuas y diferenciables, son las funciones triángulo y funciones gausianas. En el caso de la función triangular, se tiene

$$\mu_i(u) = \begin{cases} \max\{0, 1 + \frac{u-c_i}{0.5w_i}\} & \text{si } u < c \\ \max\{0, 1 + \frac{c_i-u}{0.5w_i}\} & \text{de otra forma} \end{cases} \tag{3.107}$$

mientras que para la función gausiana, se tiene

$$\mu_i(u) = \exp\left(-\frac{1}{2}\left(\frac{u - c_i}{\sigma_i}\right)^2\right) \tag{3.108}$$

En el caso de querer cubrir extremos derechos o izquierdos de los intervalos de las variables u hasta ∞ y $-\infty$, se puede saturar la función de pertenencia en 1 a partir del centro c_i en cada función. En caso de sistemas multivariables, en donde la entrada u es un vector de dimensión m, se usa comunmente:

$$\mu_i(u) = \prod_{j=1}^{m} \mu_{ij}(u_j) \tag{3.109}$$

Consideremos ahora un sistema cuya entrada u es de dimensión m, y la salida y es de dimensión p. De forma general, sea x la entrada de un modelo difuso compuesta tanto por valores de u como posiblemente de y. Consideremos el modelo para cada una de los componentes de y de la forma

$$y = f(x, \theta) = \frac{\sum_{i=1}^{R} b_i \mu_i(x)}{\sum_{i=1}^{R} \mu_i(x)} \tag{3.110}$$

Los parámetros b_i, $i = 1, \ldots, R$ son los centros de las R funciones de pertenencia de salida. Definiendo

$$\xi_i(x) = \frac{\mu_i(x)}{\sum_{i=1}^{R} \mu_i(x)} \tag{3.111}$$

la ecuación (3.110) queda

$$f(x, \theta) = \sum_{i=1}^{R} b_i \xi_i(x) \tag{3.112}$$

Ahora, si se define

$$\xi(x) = [\xi_1, \xi_2, \ldots, \xi_R]^T, \qquad \theta = [b_1, b_2, \ldots, b_R]^T \tag{3.113}$$

se obtiene finalmente un modelo del tipo

$$y = f(x, \theta) = \theta^T \xi(x) \tag{3.114}$$

La estimación de parámetros para este modelo se realiza según los resultados de la sección 3.4.1. Esto implica que los parámetros propios de cada función de pertenencia μ_i están determinados de forma previa. Otra opción es realizar la estimación aplicando los algoritmos descritos en la sección 3.4.3.1.

Una forma alternativa de identificar un modelo difuso es mediante la agrupación de datos o *clustering*. El clustering particiona los datos en subconjuntos o grupos basados en similitudes entre los datos. Consideremos un sistema difuso con una entrada x de dimensión n, con funciones de pertenencia Gaussianas, de la forma (Passino & Yurkovich, 1997):

$$f(x,\theta) = \frac{\sum_{i=1}^{R} A_i \prod_{j=1}^{n} \exp\left(-\left(\frac{x_j - v_j^i}{2\sigma}\right)^2\right)}{\sum_{i=1}^{R} B_i \prod_{j=1}^{n} \exp\left(-\left(\frac{x_j - v_j^i}{2\sigma}\right)^2\right)} \tag{3.115}$$

en donde R es el número de clústeres o grupos (reglas). En este caso el vector de parámetros está dado por

$$\theta = [A_1, \ldots, A_R, B_1, \ldots, B_R, v_1^1, \ldots, v_n^1, \ldots, v_1^R, \ldots, v_n^R, \sigma] \tag{3.116}$$

Los parámetros se obtienen mediante un proceso iterativo sobre los datos de entrada. Supongamos que tenemos M datos de entrada-salida (y^i, x^i). Los primeros parámetros serán

$$v^1 = x^1, \qquad A_1 = y^1, \qquad B_1 = 1, \qquad R = 1 \tag{3.117}$$

A continuación, se procesan los datos de entrada salida de $i = 2, \ldots, M$. Si $|v^j - x^i| < \epsilon_f$ para algún $j = 1, \ldots, R$, entonces se actualizan los parámetros como

$$A_j = A_j + y^i, \qquad B_j = B_j + 1 \tag{3.118}$$

En caso en que $|v^j - x^i| > \epsilon_f$, $\forall j$, se agrega un nuevo cluster cuyo centro es el valor de x^i, y los parámetros se determinan como se hace en (3.117). El parámetro ϵ_f es un parámetro

de diseño del modelo, que indica que tan precisos son los clústers. Entre más pequeño es ϵ_f, mejor es la aproximación, pero a su vez mayor es el número de reglas R a usarse.

3.4.5. Modelos difusos de Takagi-Sugeno

En el caso de los modelos de Takagi-Sugeno, la estructura de las reglas tiene la siguiente forma:

$$\text{IF } u_1 \text{ is } A_1 \text{ and}, \ldots, \text{and } u_n \text{ is } A_n \text{ THEN } b_i = g_i(\cdot) \tag{3.119}$$

En donde b_i ya no es el centro de área de la función de pertenencia de salida, si no que es un valor calculado a través de la evaluación de la función g_i en sus argumentos, representados por "$\cdot$". No hay restricciones sobre la función g_i en un principio. La salida de un modelo de Takagi-Sugeno estará dada por

$$y = f(x, \theta) = \frac{\sum_{i=1}^{R} g_i(x)\mu_i(x)}{\sum_{i=1}^{R} \mu_i(x)} \tag{3.120}$$

Si consideramos una función $g_i(x)$ lineal en parámetros, como por ejemplo

$$g_i(x) = a_{i,0} + a_{i,1}x_1 + \cdots + a_{i,n}x_n \tag{3.121}$$

el modelo descrito en (3.120) puede formularse como

$$y = \frac{\sum_{i=1}^{R} a_{i,0}\mu_i(x)}{\sum_{i=1}^{R} \mu_i(x)} + \frac{\sum_{i=1}^{R} a_{i,1}x_1\mu_i(x)}{\sum_{i=1}^{R} \mu_i(x)} + \cdots + \frac{\sum_{i=1}^{R} a_{i,n}x_n\mu_i(x)}{\sum_{i=1}^{R} \mu_i(x)} \tag{3.122}$$

Usando la definición de $\xi_i(x)$ realizada en (3.111), se pueden construir los vectores:

$$\xi(x) = [\xi_1(x), \ldots, \xi_R(x), x_1\xi_1(x), \ldots, x_1\xi_R(x), \ldots, \ldots, x_n\xi_1(x), \ldots, x_n\xi_R(x)]^T \tag{3.123}$$

$$\theta = [a_{1,0}, \ldots, a_{R,0}, a_{1,1}, \ldots, a_{R,1}, \ldots, \ldots, a_{1,n}, \ldots, a_{R,n}]^T \tag{3.124}$$

Luego, se puede reescribir el sistema definido en (3.122) como

$$y = f(x, \theta) = \theta^T \xi(x) \tag{3.125}$$

en donde los parámetros θ se estiman según lo visto en la sección 3.4.1, considerando que los parámetros de cada función de pertenencia se encuentran fijados de forma previa. En

caso de que se requiera estimar todos los parámetros en base a la información disponible, se pueden aplicar los algoritmos de la sección 3.4.3.1

También es posible estimar un modelo difuso de Takagi-Sugeno mediante *clustering* (algoritmo *c-means*) (Passino & Yurkovich, 1997). De la misma forma que en la sección previa, el *clustering* busca agrupar los datos bajo similitudes, en este caso, la distancia geométrica al centro de cada grupo. Definamos μ_{ij} como los grados de pertenencia de cada dato de entrada x^i a cada *cluster* j, de forma escalar, y v^j el centro del *cluster*, un vector de $n \times 1$. Se pretende minimizar la siguiente función objetivo:

$$J = \sum_{i=1}^{M} \sum_{j=1}^{R} (\mu_{ij})^m |x^i - v^j|^2 \tag{3.126}$$

en que M es el número de datos disponibles para el entrenamiento, R es el número de clústeres (que equivalen al número de reglas), x^i, $i = 1, \ldots, M$ son los datos de entrada, v^j, $j = 1, \ldots, R$ son los centros de los clústeres, y μ_{ij} es el grado de pertenencia de x^i en el j-ésimo cluster. Además, $m > 1$ es un parámetro de diseño que indica que tan suaves y cuanto se traslapan las funciones de pertenencia, y $|x|^2 = x^T x$.

Consideremos un modelo difuso de Takagi-Sugeno tal que está definido por reglas de la forma

$$\text{IF } x \text{ is } H^j \text{ THEN } g_j(x) = a_{j,0} + a_{j,1}x_1 + \cdots + a_{j,n}x_n \tag{3.127}$$

en donde H^j es el conjunto difuso definido por la función de pertenencia $\mu_{H^j}(x)$. La salida del modelo está dada por la ecuación

$$y = f(x, \theta) = \frac{\sum_{j=1}^{R} g_j(x)\mu_{H^j}(x)}{\sum_{j=1}^{R} \mu_{H^j}(x)} \tag{3.128}$$

Resolviendo la minimización de J en forma iterativa, se puede encontrar una ley para la actualiación de los centros dada por:

$$v^j(k+1) = \frac{\sum_{i=1}^{M} x^i(\mu_{ij}(k+1))^m}{\sum_{i=1}^{M}(\mu_{ij}(k+1))^m} \tag{3.129}$$

en donde

$$\mu_{ij}(k+1) = \left[\sum_{r=1}^{R}\left(\frac{|x^i - v^j(k)|^2}{|x^i - v^r(k)|^2}\right)^{\frac{1}{m-1}}\right]^{-1} \tag{3.130}$$

y $k = 1, \ldots$ representa el número de iteración. La resolución recursiva termina cuando $|v^j(k+1) - v^j(k)| < \epsilon$, $\forall j$. En la ecuación (3.130) puede ocurrir que $|x^i - v^j| = 0$. En ese caso, se define $\mu_{ij} = 0$.

Finalmente, la función de pertenencia $\mu_{H^j}(x)$ estará dada por la cercanía de x a cada uno de los centros de cada cluster:

$$\mu_{H^j}(x) = \left[\sum_{r=1}^{R}\left(\frac{|x - v^j|^2}{|x - v^r|^2}\right)^{\frac{1}{m-1}}\right]^{-1} \tag{3.131}$$

Ahora, para obtener los parámetros $a_j = [a_{j,0}, \ldots, a_{j,n}]^T$ de cada una de las j reglas, se debe minimizar el error de aproximación de la función dada por

$$J_j = \sum_{i=1}^{M}(\mu_{ij})^2(y^i - a_j^T\hat{x}^i)^2 \tag{3.132}$$

para cada una de las $j = 1, \ldots, R$ reglas, en donde μ_{ij} es el grado de pertenencia de x^i al j-ésimo cluster dado por $\mu_{H^j}(x^i)$ definido en (3.131), y $\hat{x}^i = [1(x^i)^T]^T$. La resolución de este problema es equivalente al expuesto en la sección 3.4.1, en la ecuación (3.51). Los parámetros óptimos en este caso son

$$\hat{a}_j = (\hat{X}^T D_j^2 \hat{X})^{-1}\hat{X}^T D_j^2 Y \tag{3.133}$$

en donde $\hat{X} = [\hat{x}^1 \cdots \hat{x}^M]^T$, $Y = [y^1, \ldots, y^M]$, y $D_j = \text{diag}(\mu_{1j}, \ldots, \mu_{Mj})$ es una matriz diagonal con μ_{ij} en su elemento de la diagonal d_{ii}.

Otros métodos de identificación pueden revisarse en (Babuska & Verbruggen, 1995), (Babuska, Roubos, & Verbruggen, 1998).

3.4.6. Otras técnicas

Adicionalmente, existen otras técnicas o estructuras para modelación no lineal a través de sistemas inteligentes. Entre ellas se pueden mencionar las redes neuronales auto organizativas, mapas de Kohonen, redes tipo ART, redes LVQ (*Learning Vector Quantization*), redes CMAC (*Cerebellar Model Articulation Control*), redes GMDH (*Group Method of Data Handling*), redes neuronales recurrentes como las redes de Hopfield, redes de Jordan y Redes de Elman, etc.

También existen combinaciones de estructuras como las redes neuro-difusas (*Neuro-Fuzzy Networks*), mapas ART difusos, etc. Sin embargo, la aplicación de estas técnicas es poco generalizada en lo que respecta a modelación, encontrándose desarrollos en la literatura para aplicaciones muy específicas. Además, varias de estas técnicas se emplean más para la clasificación de patrones (mapas de Kohonen y mapas ART) que para la modelación.

3.5. Modelos comerciales

En el complejo tema del desarrollo de modelos para la creación de aplicaciones para centrales de ciclo combinado, existen modelos comerciales de centrales que pueden adquirirse. Estos modelos, formulados mediante software especializados, incluyen aplicaciones de ajuste del modelo, y de monitorización y diagnóstico del sistema dinámico.

Uno de los modelos más usados está implementado a través del software *IPSEpro* (www.ipsepro.com). IPSEpro es un conjunto de módulos de software para crear modelos de procesos y utilizarlos en todos los modos de operación de la planta. Estos módulos son capaces de:

- Calcular balances de calor y predecir el rendimiento de la planta en función de los parámetros de diseño
- Estimar costos durante el diseño conceptual de la planta

- Verificar y validar mediciones durante pruebas de aceptación
- Monitorizar y optimizar el rendimiento de la planta en línea y fuera de línea
- Planificar modificaciones y potenciaciones de plantas existentes

El corazón del programa es un sistema de modelamiento basado en balances térmicos y simulación de procesos. Además, posee un kit de desarrollo de modelos que permite agregar componentes detallados de forma personalizada a los procesos existentes. Este software ha sido especialmente diseñado para ser usado en plantas generadoras térmicas, plantas de ciclo combinado, recuperadoras de calor, y en general industrias que trabajen con procesos térmicos que involucren turbinas de gas, turbinas de vapor y calderas. En la figura 3.4 se muestra una interfaz gráfica del programa simulando un sistema de turbina de gas en una central de ciclo combinado.

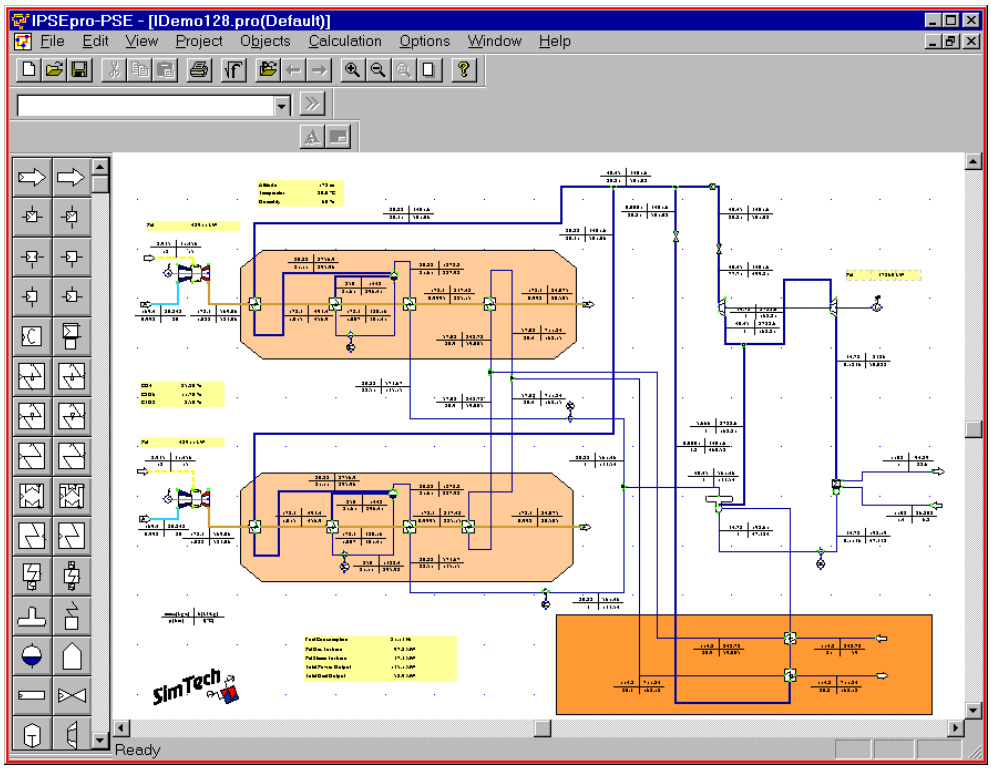

FIGURA 3.4. Ejemplo de modelo de proceso de una central de CC en IPSEpro.

En cuanto a la turbina de gas, IPSEpro posee una librería especial denominada *Gas Turbine Library*, que cuenta con modelos predefinidos de las turbinas comerciales más usadas dentro de centrales de CC. La librería entrega herramientas para evaluar los rendimientos específicos de cada parte de la turbina dentro del funcionamiento de la central de CC. Cada modelo cuenta con tablas de datos de comportamiento y parámetros de diseño reales. El modelo también incluye condiciones como temperatura, presión ambiental, altitud, etc., de forma de que el modelo se ajuste de la mejor forma al proceso real. Esta librería es capaz de trabajar en conjunto con el resto de las librerías del programa, como la *Advanced Power Plant Library*.

Otra alternativa disponible es *ProSim* (www.prosim.net), programa que permite realizar balances termodinámicos en procesos, pero no está orientado 100% al análisis y desarrollo de aplicaciones en centrales de CC.

Existen otros softwares como *Thermoflow GT Master* (www.thermoflow.com), pero tienen más bien fines académicos que de aplicación real sobre los procesos. También se encuentra el *Continuous System Simulation Language* (CSSL), en sus diversas versiones, que fue desarrollado por Jet Propulsion Labs y Simulation Council Inc. para la simulación de turbinas y sistemas en general dependiente del tiempo y bajo ecuaciones diferenciales no lineales. Una versión mas avanzada es la *Advanced Continuous Simulation Language* ACSL.

Puede revisarse también la revista especializada *Multiscale Modeling & Simulation* MMS, de la *Society for Industrial and Applied Mathematics*.

4. DESARROLLO Y SIMULACIÓN DE UN MODELO PARA LA CENTRAL SAN ISIDRO

En este capítulo se presenta la modelación realizada para la turbina de gas de la central de ciclo combinado San Isidro. El objetivo de esta modelación es la construcción de un simulador de la planta para el aprovechamiento en el desarrollo de aplicaciones de diagnóstico, monitorización y control, por lo cual el modelo se debe ajustar de la forma más precisa posible a los datos disponibles del proceso.

Uno de los objetivos del modelo, y de las aplicaciones comentadas, es su aplicación para la monitorización en tiempo real de la planta. Por este motivo, el modelo debe considerar una carga computacional lo más pequeña posible para generar sus resultados. Los modelos fenomenológicos discutidos en la sección 3.2 requieren una gran capacidad computacional para la solución de ecuaciones diferenciales no lineales y otras técnicas matemáticas que son requeridas en la resolución del modelo. Esto considerando modelos con un nivel de detalle suficiente para realizar aplicaciones de diagnóstico.

Por otro lado, si bien los modelos fenomenológicos simplificados presentados en la sección 3.3 presentan una carga computacional bastante manejable, no son capaces de entregar el nivel de detalle que requieren las aplicaciones de monitoreo y diagnóstico. Lo anterior sugiere construir un modelo de entrada-salida, y ajustarlo en base a la información disponible del proceso.

A continuación se presenta el proceso de construcción y selección de los modelos en base a la información disponible.

4.1. Estructura del modelo

Para la construcción del modelo, se seleccionaron las variables más relevantes para la operación de la planta. La selección está basada en los antecedentes proporcionados por los

operadores de la planta y del Centro de Monitoreo y Diagnóstico, respecto a las variables cuyo valor se monitorea constantemente y que determinan el estado de operación. Cabe destacar que para cada variable se cuenta con resgistros históricos de operación, provistos por el Centro de Monitoreo y Diagnóstico (CMD) de Endesa.

Las variables consideradas son las siguientes:

(i) Referencia y valor actual de la posición de los álabes de guía en el compresor (*Inlet Guide Vanes* o IGV).
(ii) Temperatura y presión ambientales a la entrada del compresor.
(iii) Referencia y valor actual de la velocidad de rotación del eje generador-turbina-compresor.
(iv) Temperatura de salida del compresor.
(v) Posición de la válvula de combustible, y flujo de combustible hacia la cámara de combustión.
(vi) Referencia y valor actual de la potencia generada.
(vii) Temperatura de salida de la turbina de gas y a la entrada de la caldera recuperadora o HRSG.
(viii) Temperatura aire de enfriamiento del rotor.

Según lo visto en el capítulo 2, los componentes principales de la turbina de gas son compresor, cámara de combustión y turbina. Acorde con esto, para el desarrollo del modelo de entrada-salida se consideraron 4 subsistemas de la turbina de gas, enumerados a continuación:

(i) Sistema de álabes en el compresor: relaciona la referencia del valor de las IGV con su valor real.
(ii) Compresor: relaciona el valor de las IGV, Velocidad del eje, temperatura y presión ambientales, con el valor de temperatura a la salida del compresor.

(iii) Sistema de combustible: relaciona la posición de la válvula con el flujo de combustible.

(iv) Cámara de combustión - turbina: relaciona la temperatura y presión de salida del compresor, el flujo de combustible, las referencias de potencia y velocidad, con los valores de velocidad del eje, potencia generada, temperatura de salida de la turbina y a la entrada del HRSG, y temperatura del aire de enfriamiento del rotor.

La conexión entre estos sistemas se ilustra en la figura 4.1. Debido a las altas temperaturas y presiones en la cámara de combustión, no se cuenta con medidas de esas variables en dicho punto que permitan diferenciar la cámara de combustión de la turbina de gas como subsistemas separados. Se incluye también el sistema de control de álabes (IGV) y de combustible, como principales elementos de control sobre la turbina.

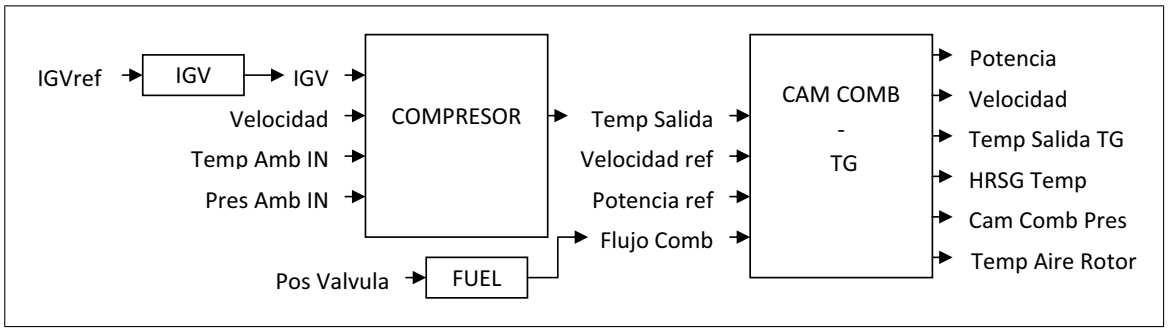

FIGURA 4.1. Esquema del modelo de la TG de Central San Isidro.

Se desarrollaron modelos según las técnicas vistas en la sección 3.4. Respecto a las figuras 3.1 y 3.2, se aprecia una diferencia en cuanto a que en dichos modelos se componen de plantas de primer orden que transforman las variables de forma lineal, siempre de una entrada y una salida, a través de una secuencia. En el esquema propuesto, no se realizan supuestos de interdependencia sobre las variables de entrada para cada subsistema. Existe también una diferencia en cuanto al tratamiento de variables. Por ejemplo, en el modelo propuesto no se consideran flujos másicos, y si se consideran variables ambientales. Esto se debe a los comentarios de los operadores respecto a las variables de mayor interés. A su vez, algunas variables como velocidad son calculadas a partir del movimiento de la turbina y no supuestas como externas en todo momento.

4.2. Estimación y selección de parámetros

Para cada uno de los sistemas descritos en la sección anterior, se ajustaron modelos según las técnicas de modelación lineal y difusa de Takagi-Sugeno revisadas en el capítulo 3. Cada sistema (IGV, Compresor, Combustible y Cámara de combustión - TG) se compone de un modelo MISO (*Multiple Input - Single Output*) por cada una de sus salidas. Las entradas pueden incluír retroalimentaciones de las salidas. La selección de los parámetros para cada modelo, y del mejor modelo para cada salida, se realizó en base a un índice de porcentaje de ajuste, definido de la siguiente forma:

$$FIT = 1 - \frac{|y - \hat{y}|}{|y - \bar{y}|} \tag{4.1}$$

en donde y representa un vector que contiene las mediciones de m salidas del sistema, $\hat{y}$ es el vector con m salidas estimadas por el modelo, e $\bar{y}$ es un vector de dimensión m en donde todos sus componentes son iguales al promedio de los valores del vector y. La operación $|\cdot|$ representa la norma euclideana entre vectores. Se seleccionó en cada caso el tipo de modelo y el conjunto de parámetros que maximizan este índice.

Para el entrenamiento de los modelos, se utilizó un conjunto de datos correspondiente a una puesta en marcha del sistema. Los datos están muestreados con un período de 1 minuto, durante un tiempo de 88 horas y 27 minutos (5307 muestras en total). Del total de muestras, se utilizó los primeros 2500 datos para entrenamiento, y el resto como datos de validación de los modelos. Estos datos, al pertenecer a una partida de planta, cumplen con entregar una exitación persistente sobre el sistema, y ser adecuados por tanto para la identificación de modelos.

Para los modelos tipo ARX, se evaluaron los resultados para distinta cantidad de pasos de regresión en cada variable, seleccionando el modelo con mejores resultados mediante una búsqueda simple. Se acotó el número de parámetros totales del modelo.

Para el modelo en espacio de estados, se seleccionó el orden del modelo en base a aquel que entregó una relación entre mejores resultados y número de parámetros, para el

mismo límite de parámetros que los modelos ARX. Los resultados de los modelos ARX y en espacio de estados se muestran junto a las otras técnicas en los graficos que siguen a continuación.

Para el modelo de redes neuronales con perceptrones multicapa, se realizó una estimación de parámetros aplicando el algoritmo back-propagation y el algoritmo de Levenberg-Marquardt, con un criterio de detención dado por una diferencia de la mejoría del índice MSE definido en (3.84) de $\Delta e = 10^{-8}$. Se consideran redes de 1 capa oculta de entre 1 y 15 neuronas, y de 2 capas ocultas de entre 1 y 10 neuronas por capa. La diferencia en la velocidad de convergencia de cada algoritmo se muestra en la figura 4.2, en donde se han ajustado los datos del sistema del compresor para una red de 3 neuronas en la capa oculta y 3 regresores por variable (42 parámetros), y con un criterio de detención de diferencia del error medio cuadrático de $\Delta e = 10^{-8}$. Se puede apreciar que el algoritmo de Levenberg-Marquartd no sólo converge en un número de iteraciones de entre 100 y 150 veces menor que el algoritmo de back-propagation, si que también logra un ajuste mejor para igual criterio de detención (unas 50 veces mejor).

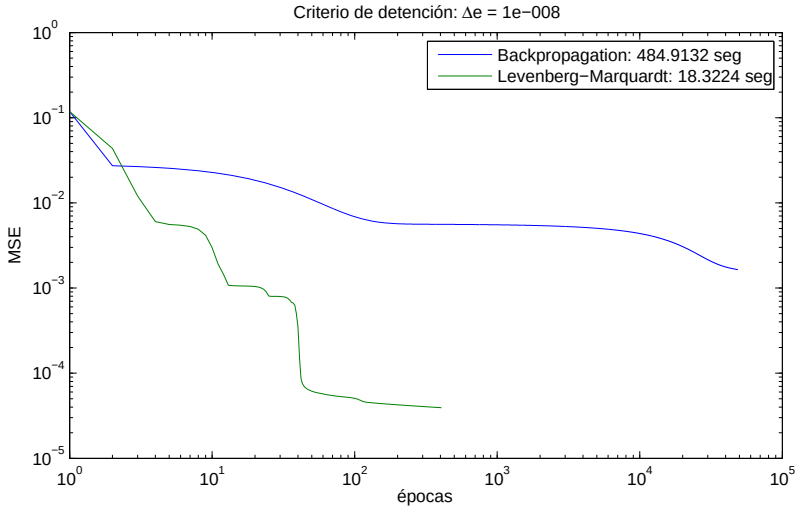

FIGURA 4.2. Evolución del error cuadrático medio para los algoritmos de back-propagation y de Levenberg-Marquardt.

La figura 4.3 muestra la evolución del valor del índice definido en (4.1) en el subsistema Compresor según el número de neuronas en la capa oculta, en función del número de

parámetros del modelo (debido a la cantidad de regresores), para una red de 1 capa oculta. En este caso, considerando R neuronas en la capa oculta, n variables de entrada al modelo y p regresores por cada variable, se obtienen npR parámetros debidos los pesos w_{ij}, R parámetros θ_j y R parámetros v_j, dando un total de $npR + 2R$ parámetros para el modelo.

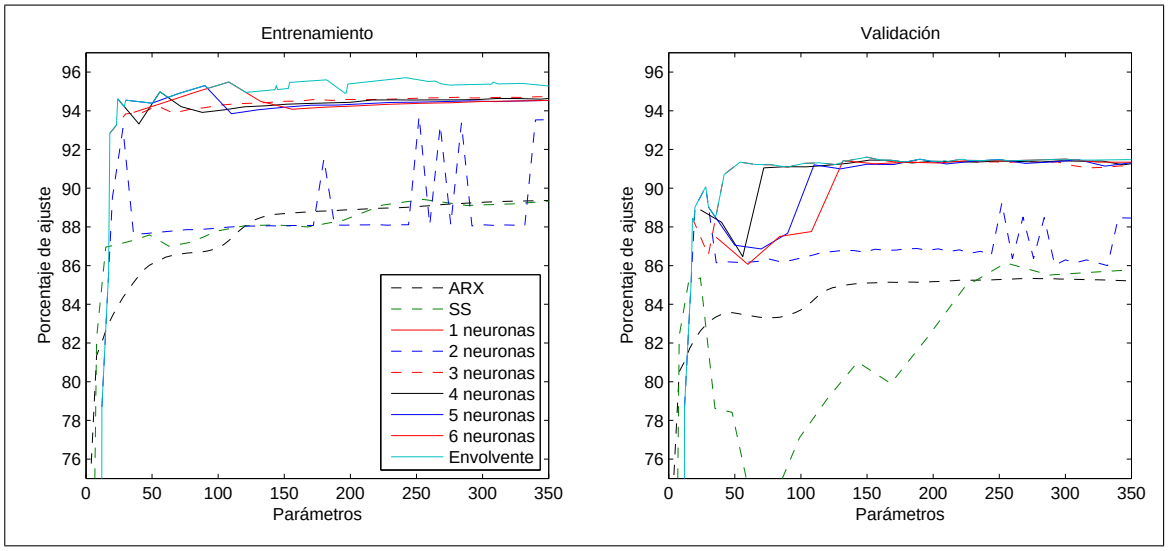

FIGURA 4.3. Ajuste en base al número de parámetros en modelos de redes perceptrón multicapa.

En esta figura, la envolvente representa los máximos valores considerando de 1 a 15 neuronas. Puede apreciarse que al incrementar las neuronas de la capa oculta, el ajuste en los datos de entrenamiento mejora levemente; sin embargo en la validación se llega a un límite que no puede sobrepasar.

La figura 4.4 muestra el índice de ajuste en función de los parámetros para el caso de 2 capas ocultas. En este caso, habiendo R_1 neuronas en la primera capa oculta y R_2 neuronas en la segunda, se tiene un total de $R_1(np + R_2 + 1) + 2R_2$ parámetros. La envolvente representa el máximo valor para determinado número de parámetros, de todas las combinaciones entre 1 a 10 neuronas en cada una de las dos capas ocultas. Se aprecia, en comparación con la figura 4.3, que la inclusión de una capa oculta adicional no significa una mejoría del ajuste significativa.

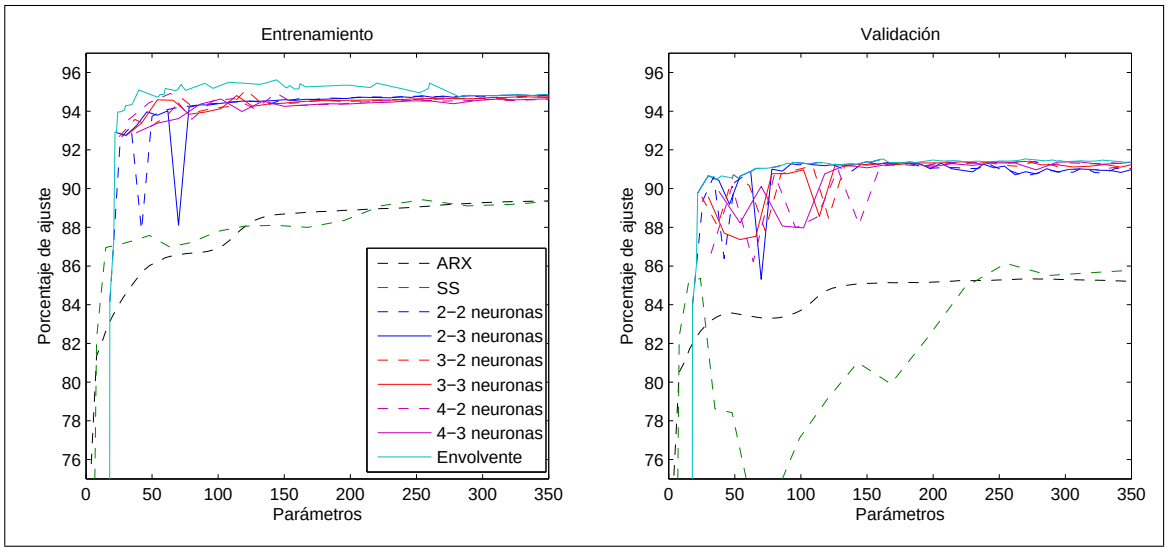

FIGURA 4.4. Ajuste en base al número de parámetros en modelos de redes perceptrón multicapa, con 2 capas ocultas.

Para el modelo neuronal de base radial (RBF), se consideraron de 2 a 25 neuronas en la capa oculta. Se realizó una primera estimación de parámetros fijando los centros c_j y el parámetros β por cada neurona. Los centros se fijaron de manera que quedaran uniformemente distribuídos entre los valores mínimos y máximos de cada variable. El parámetro β se fijó como la norma de la distancia entre centros adyacentes.

Considerando n variables de entrada al modelo, con p regresores por cada variable, y R neuronas en la capa oculta, al fijar los centros se obtienen n parámetros por centro por cada neurona R, y 1 parámetro en común β. Para la salida, se tienen R parametros debidos a las R neuronas. En total, se tienen $nR + R + 1$ parámetros para este modelo.

Al esttimar todos los parámetros usando algoritmos de descenso, se tienen np parámetros por centro y 1 parámetro β para cada neurona, junto con R parámetros de salida. Se obtienen en total $npR + 2R$ parámetros para este modelo. La figura 4.5 muestra la evolución del valor del índice de ajuste en función de los parámetros del modelo, según el número de neuronas en la capa oculta.

En la figura 4.5 se muestran los ajustes obtenidos para 2 a 6 neurones, y la envolvente correspondiente a los máximos valores considerando de 2 a 25 neuronas. Puede apreciarse

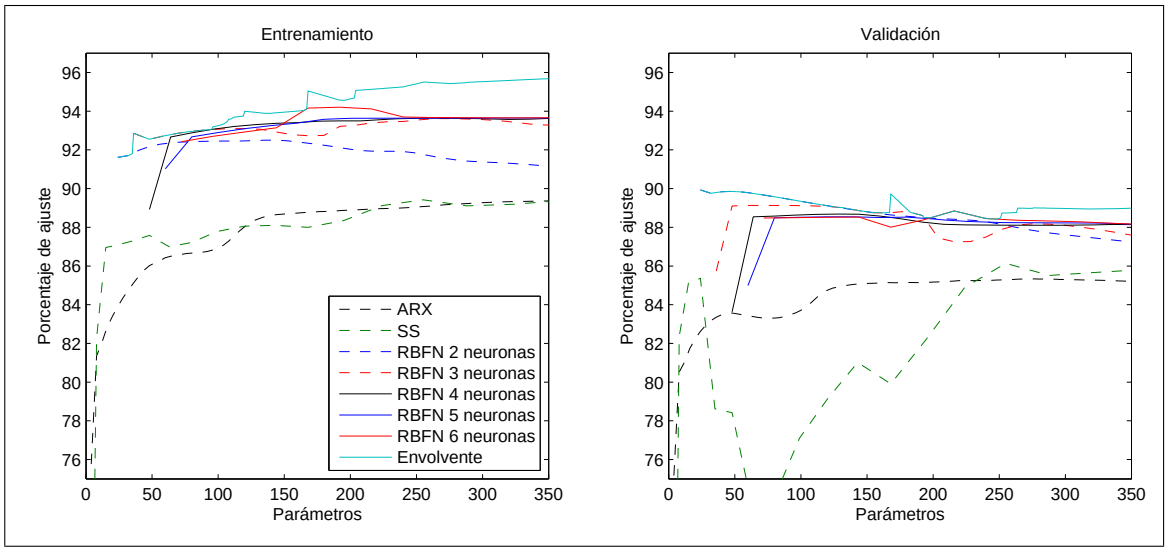

FIGURA 4.5. Ajuste en base al número de parámetros en modelos RBFN.

que para este tipo de modelo, si bien se logran buenos porcentajes de ajuste para los datos de entrenamiento, en la validación el índice cae y la diferencia con los modelos lineales no es tan sustancial.

Para los modelos difusos, en los del tipo Takagi-Sugeno, se consideraron de 2 a 15 reglas. Se realiza una primera estimación en donde los centros de las funciones de pertenencia se ubicaron arbitrariamente distribuidos de forma uniforme entre los valores mínimos y máximos de cada variable. El valor de *spread* de las funciones de pertenencia σ, se seleccionó como la distancia entre los centros. En un modelo difuso de Takagi-Sugeno con R reglas, se identifican np parámetros para cada una de las funciones de salida, y $2n$ parámetros por cada función de pertenencia (n por cada centro y n por cada *spread*), dando un total de $nR(p+2)$ parámetros a identificar en el modelo. La evolución del índice de ajuste en función de los parámetros se muestra en la figura 4.6. Al realizar una identificación de todos los parámetros del modelo, usando los algoritmos de descenso descritos en la sección 3.4.3.1, se deben considerar np parámetros de salida por regla, np parámetros por cada centro y np parámetros por cada *spread*, dando un total de $3npR$ parámetros para R reglas. La figura 4.7 muestra el indice de ajuste en función de los parámetros en este caso.

En este caso, el ajuste del modelo mejora respecto al caso subóptimo en el que se prefijan los parámetros; sin embargo el número de parámetros aumenta al no producirse repeticiones (en los centros y *spread* para cada regresor), por lo que finalmente para un mismo número de parámetros los resultados son prácticamente idénticos.

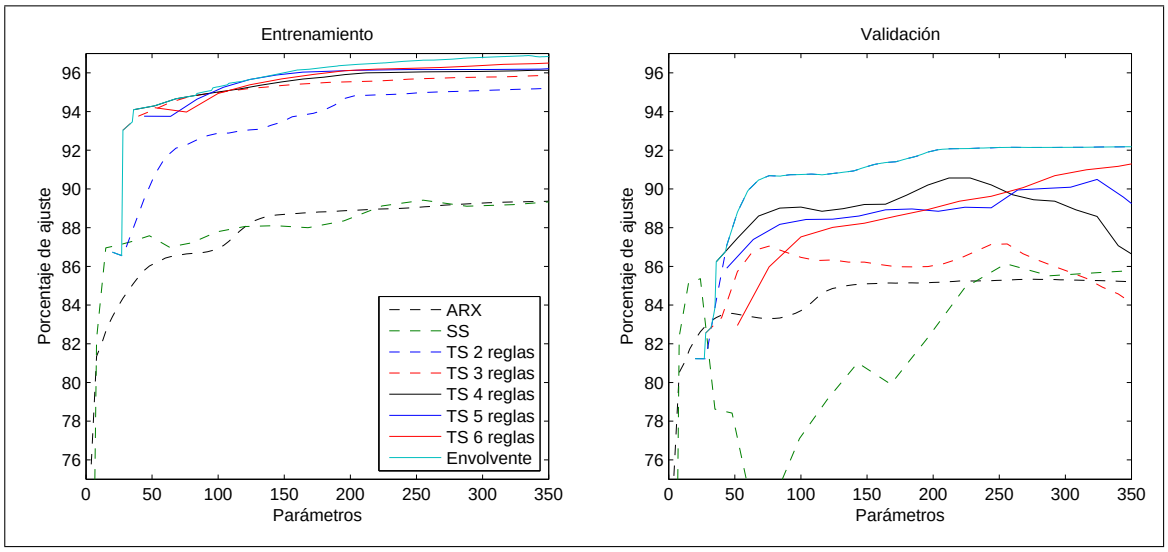

FIGURA 4.6. Ajuste en base al número de parámetros en modelos difusos de TS.

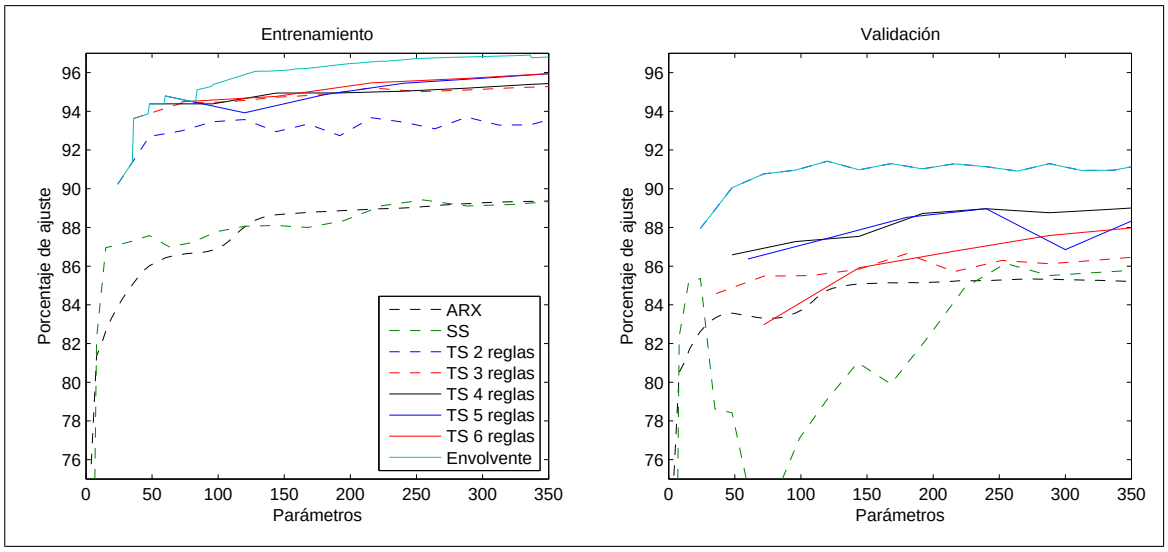

FIGURA 4.7. Ajuste en base al número de parámetros en modelos difusos de TS, identificando todos los parámetros.

Finalmente, para los modelos de Takagi-Sugeno con clustering, se consideraron de 2 a 15 reglas. Se seleccionaron como valores iniciales para los centros los usados en los modelos de Takagi-Sugeno convencionales. El número de centros y pasos de regresión se seleccionaron en base al mismo criterio. La figura 4.8 muestra la evolución del índice de ajuste en función del número de parámetros en el modelo. En este caso, se tienen np parámetros por cada una de las R reglas, n por cada centro de cada regla, y 1 parámetro m de ajuste de suavidad, sumando un total de $nR(p+1)+1$ parámetros a estimar.

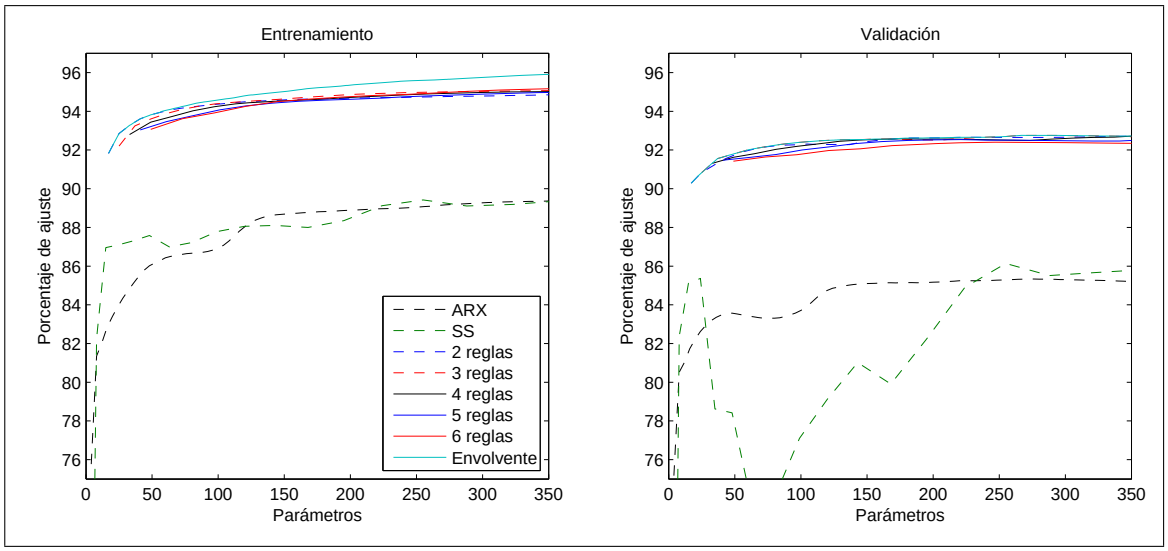

FIGURA 4.8. Ajuste en base al número de parámetros en modelos difusos de TS con clustering.

La envolvente representa el máximo ajuste considerando de 2 a 15 reglas. Como puede apreciarse, en este caso los ajustes son mucho mas estables en función del número de parámetros.

La figura 4.9 muestra una comparación entre las envolventes obtenidas en cada modelo. Se aprecia que el desempeño de los modelos inteligentes es bastante superior al de los modelos lineales tanto sobre los datos de entrenamiento como de validación. Entre los modelos "inteligentes", puede apreciarse un cierto mejor desempeño de los modelos difusos de Takagi-Sugeno y Takagi-Sugeno con *clustering*. La selección de los parámetros de las funciones de pertenencia de forma fija, permite realizar un entrenamiento muy simple

(lineal) para los parámetros de las funciones de salida. Comparado con el modelo en que todos los parámetros fueron identificados, el resultado es muy similar.

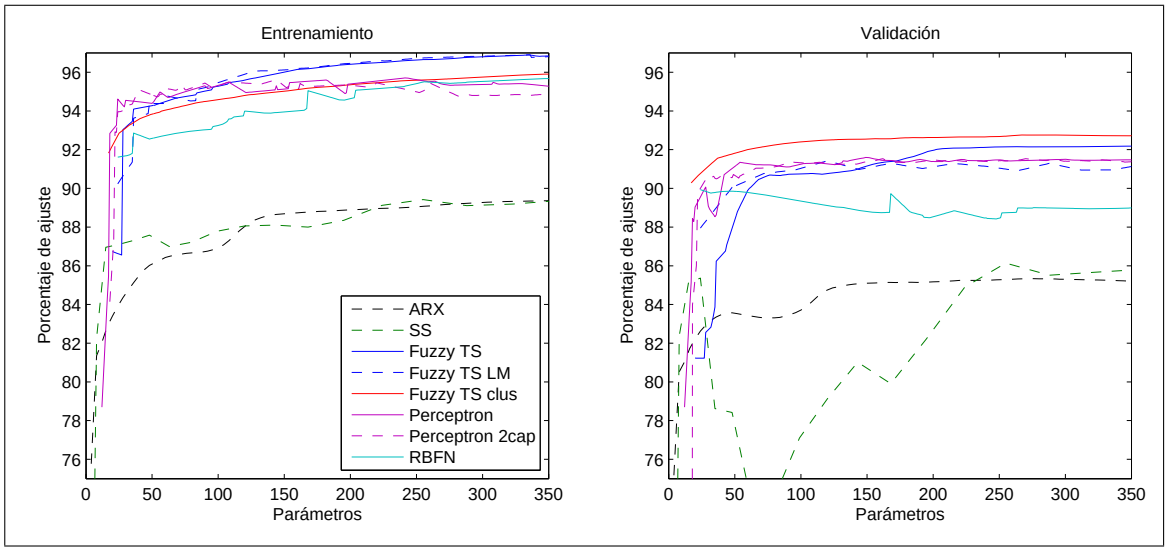

FIGURA 4.9. Comparación de los resultados para cada técnica.

Lo anterior no ocurre por ejemplo con el modelo neuronal de base radial. Al fijar los centros y el parámetro β con el mismo criterio que en los modelos difusos, se obtienen resultados muy pobres. En ese caso, la identificación de todos los parámetros con los algoritmos de descenso produce una diferencia significativa en el índice de ajuste.

Otro punto a favor de los modelos difusos, es el esfuerzo computacional requerido para su entrenamiento en comparación con las redes neuronales. El tiempo de entrenamiento de los parámetros la red de dos capas cuyos resultados se muestran en la figura 4.4 fue superior a 100 horas de cómputo (entrenamiento de todas las combinaciones de estructura), mientras que el tiempo de entrenamiento para el modelo difuso de Takagi-Sugeno no supera los 5 minutos (en el mismo PC).

Por lo anterior, se ha descartado el uso de redes neuronales y solo se trabajará con modelos difusos. En general, en todos los casos con mayor cantidad de parámetros se obtienen mejores resultados, por lo que se ha restringido el número de parámetros a $50n$, con n el número de entradas. Para todos los subsistemas, se cumple que el número de

datos de entrenamiento es mayor a 10 veces el número de parámetros a identificar, lo cual asegura empíricamente la invertibilidad de la matriz en (3.50).

4.3. Modelos obtenidos con enfoque en simulación

En el enfoque de simulación, se pretende emular el comportamiento del proceso para un período de tiempo indefinido, según determinados valores para las variables de entrada. El modelo recibe las lecturas de los valores de entrada para generar las salidas en todo instante de tiempo t. En caso de que el modelo considere entradas autoregresivas, se usan los valores anteriores de las salidas entregadas por el mismo modelo.

Al realizar la aproximación de modelos en el enfoque simulación, los modelos presentaron inestabilidades en simulaciones a largo plazo si se consideraban retroalimentaciones autoregresivas. Esto ocurre debido a que la estimación por mínimos cuadrados entrega una identificación de parámetros ajustando el error cuadrático medio de aproximación, pero no considera que los parámetros obtenidos aseguren estabilidad para el modelo. Al incluír entradas autoregresivas, a mayor cantidad de pasos de predicción, el error acumulado crece debido a que los parámetros asociados a estas entradas autoregresivas resultaron ser inestables. Es por esto que todos los modelos ajustados solo consideran entradas exógenas y sus regresores. Además, como se mencionó en la sección anterior, el mejor modelo en cada caso se basa en una selección entre los modelos con una limitación en la cantidad máxima de parámetros fijada en $50n$, con n el número de entradas al modelo.

Las siguientes tablas resumen la mejor estimación de las técnicas seleccionadas para cada uno de los sistemas descritos en la sección anterior.

Para el sistema de álabes IGV, la mejor aproximación a la salida, considerando el conjunto de datos de validación, está dado por la aproximación del modelo difuso de Takagi Sugeno, que se muestra en la figura 4.10. La línea punteada señala el límite de los datos

TABLA 4.1. Resultados del ajuste de modelos para el sistema IGV.

Tipo modelo	Porcentaje de ajuste	
	Entrenamiento	Validación
ARX	94.98	95.60
Espacio de Estados	94.72	95.41
Fuzzy Takagi-Sugeno	95.79	96.17
Fuzzy TS con Clustering	95.49	95.96

de entrenamiento y el comienzo de los datos de validación. En este caso, el modelo cuenta con 12 reglas y 3 regresores por cada variable.

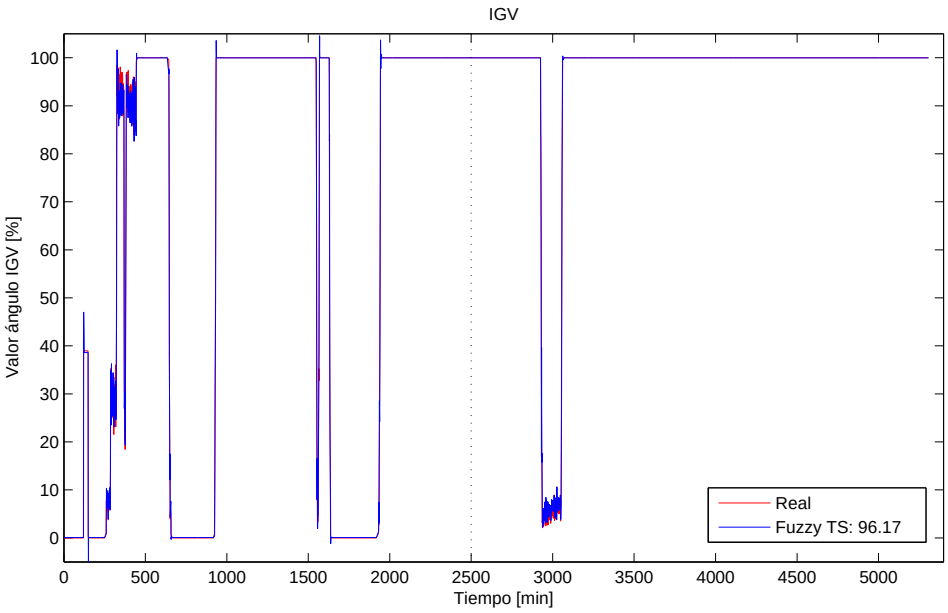

FIGURA 4.10. Mejor ajuste a la salida del IGV.

En el caso del compresor, el mejor ajuste para la salida (considerando datos de validación) está dado por una modelación difusa tipo Takagi-Sugeno con clustering, con 2 centros (reglas) y 23 regresores. El resumen de resultados obtenidos se muestra en la tabla 4.2. La aproximación se muestra en la figura 4.11.

TABLA 4.2. Resultados del ajuste de modelos para el sistema del Compresor.

Tipo modelo	Porcentaje de ajuste	
	Entrenamiento	Validación
ARX	88.89	85.17
Espacio de Estados	89.52	86.34
Fuzzy Takagi-Sugeno	96.52	91.91
Fuzzy TS con Clustering	95.27	92.62

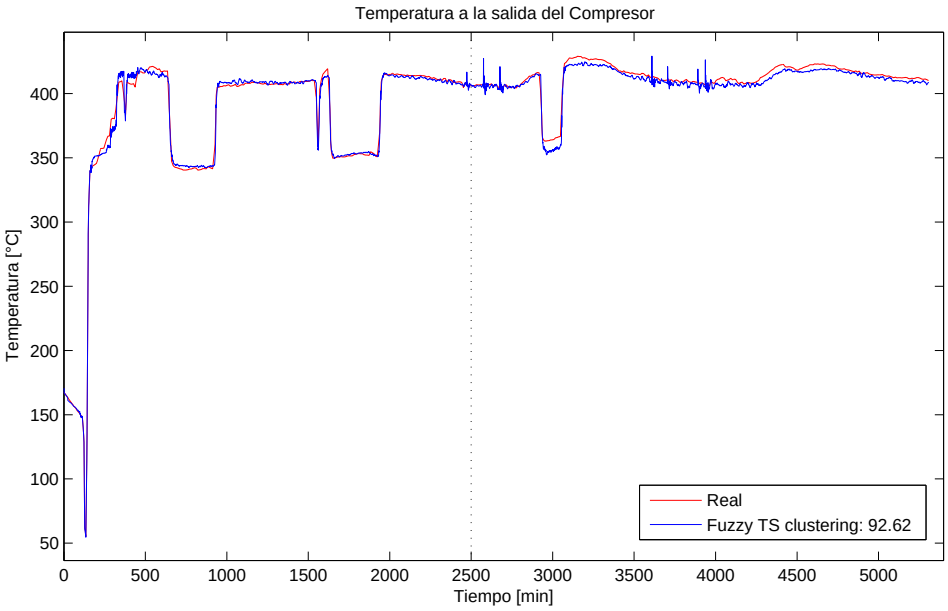

FIGURA 4.11. Mejor ajuste a la temperatura de salida del compresor.

En el caso del sistema de combustible, tal cual como lo muestra la tabla 4.3, la mejor aproximación está dada por un modelo difuso de Takagi-Sugeno con clustering, con 2 regresores por variable y 14 centros (reglas). Se puede apreciar gráficamente el ajuste obtenido para este sistema en la figura 4.12.

Para el sistema de cámara de combustión - turbina, en la variable de potencia generada, el mejor ajuste está dado por un modelo difuso de Takagi-Sugeno, con 3 regresores por

TABLA 4.3. Resultados del ajuste de modelos para el sistema de combustible.

Tipo modelo	Porcentaje de ajuste	
	Entrenamiento	Validación
ARX	89.85	89.08
Espacio de Estados	90.81	89.16
Fuzzy Takagi-Sugeno	95.28	93.87
Fuzzy TS con Clustering	95.42	94.13

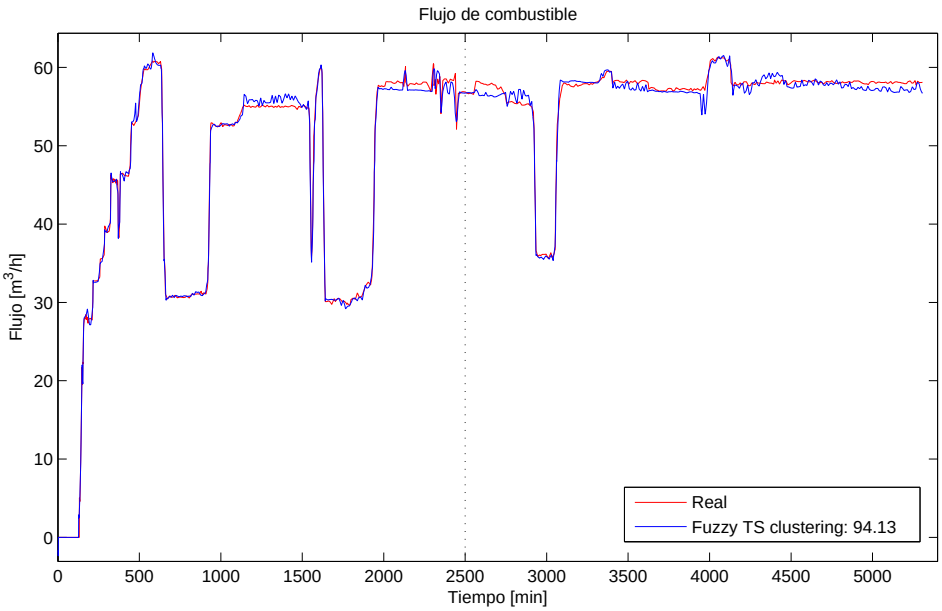

FIGURA 4.12. Mejor ajuste para el flujo de combustible.

variable y 9 centros (reglas), como se muestra en la tabla 4.4. La simulación de esta variable se muestra gráficamente en la figura 4.13.

En cuanto a la variable de velocidad, el mejor ajuste está dado también por el modelo difuso de Takagi-Sugeno. En este caso, se han empleado 15 regresores por variable, y 3 centros (reglas). El ajuste logrado para esta variable se muestra en la figura 4.14.

Para la temperatura a la salida de la turbina, el mejor ajuste está dado por un modelo difuso de Takagi-Sugeno con 1 regresor por variable de entrada y 15 centros (reglas). El ajuste de datos para esta variable se muestra en la figura 4.15.

TABLA 4.4. Resultados del ajuste de modelos para el sistema Cámara de combustión - Turbina (primeras 3 salidas).

Tipo modelo	Porcentaje de ajuste					
	POWER		SPEED		TGTEMP	
	Entr.	Valid.	Entr.	Valid.	Entr.	Valid.
ARX	95.30	94.85	81.93	76.17	72.70	65.77
Espacio Estados	92.67	92.62	89.46	86.33	75.83	66.16
Fuzzy TS	97.45	96.68	99.06	98.97	96.06	95.31
Fuzzy TS Cluster	97.22	96.49	97.76	97.75	95.04	93.75

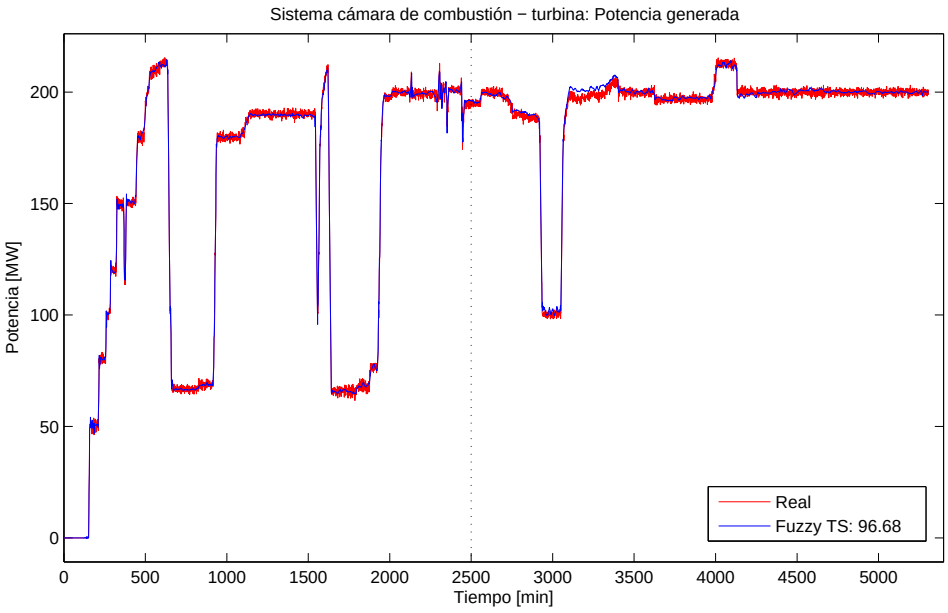

FIGURA 4.13. Mejor ajuste para la potencia generada.

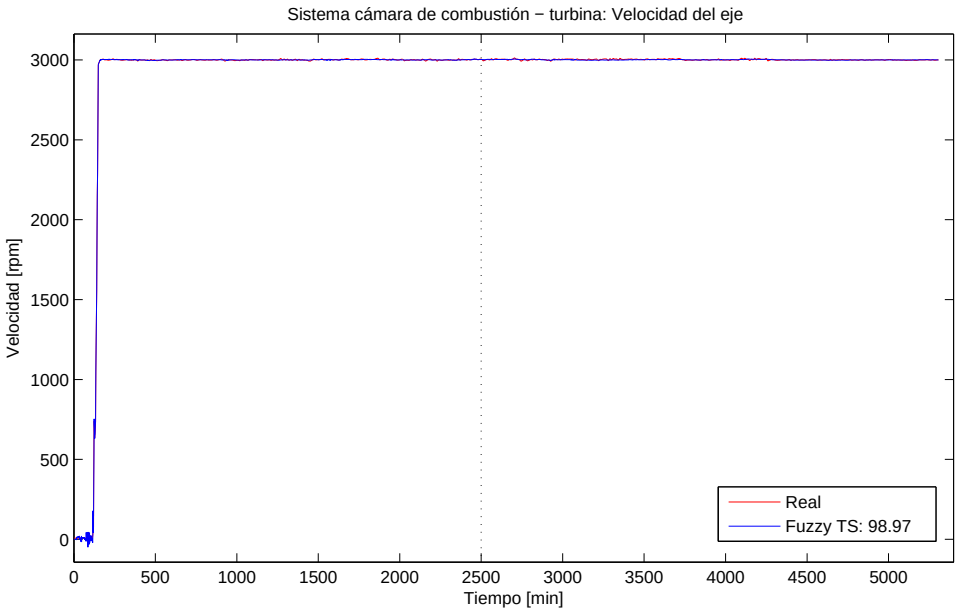

FIGURA 4.14. Mejor ajuste para la velocidad del eje.

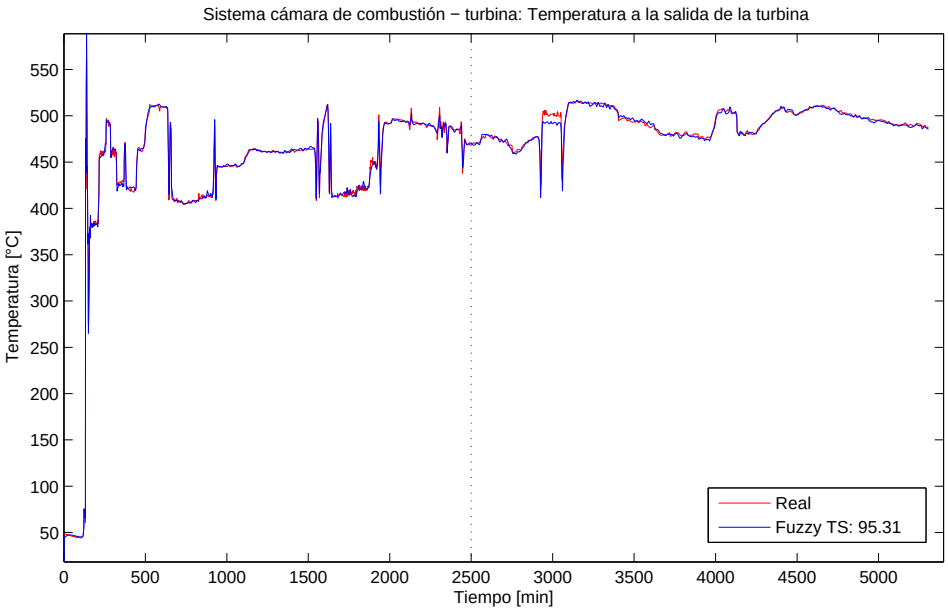

FIGURA 4.15. Mejor ajuste para la temperatura a la salida de la turbina.

Las variables de los ajustes de modelos de temperatura de entrada al HRSG, presión en la cámara de combustión y temperatura del aire de enfriamiento del rotor, se encuentran en la tabla 4.5.

TABLA 4.5. Resultados del ajuste de modelos para el sistema Cámara de combustión - Turbina (últimas 3 salidas).

Tipo modelo	Porcentaje de ajuste					
	HRTEMP		COMBPRES		RTRTEMP	
	Entr.	Valid.	Entr.	Valid.	Entr.	Valid.
ARX	84.39	78.12	91.82	88.10	81.38	77.15
Espacio Estados	84.48	76.75	82.50	81.18	79.48	80.13
Fuzzy TS	97.21	95.43	98.64	98.30	89.02	89.73
Fuzzy TS Cluster	96.91	94.08	98.03	97.90	95.23	94.57

Para la temperatura de entrada al HRSG, el mejor modelo es del tipo difuso de Takagi-Sugeno. En este modelo se usan hasta 6 pasos de regresión y 6 centros (reglas). Los resultados se muestran en la figura 4.16.

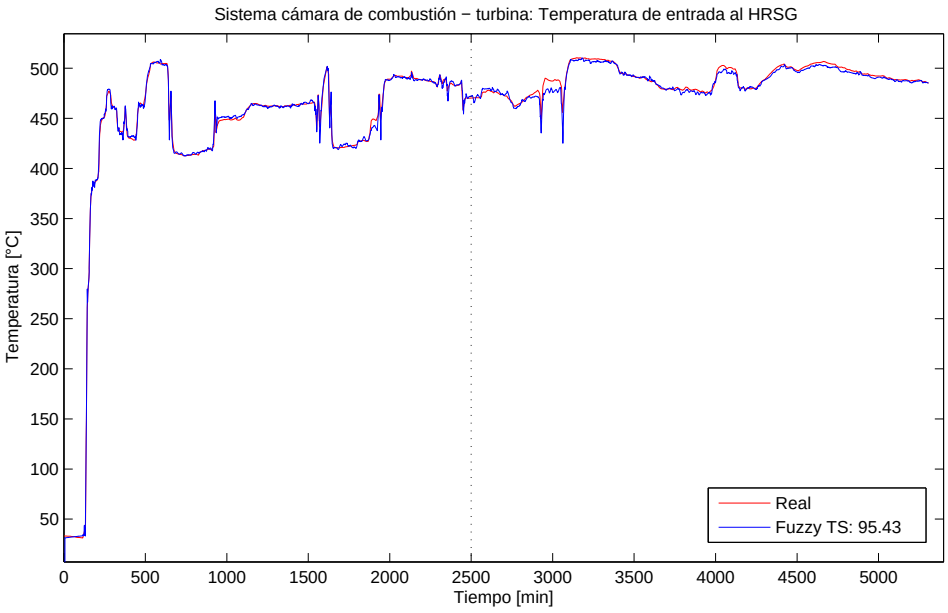

FIGURA 4.16. Mejor ajuste para la temperatura a la entrada del HRSG.

Para la presión en el compresor, se ajusta un modelo de Takagi-Sugeno con 2 regresores y 10 centros. Para el aire de enfriamiento del rotor, se usa un modelo Takagi-Sugeno con clustering, con 5 regresores y 8 centros. Estos ajustes se muestran en la figura 4.17.

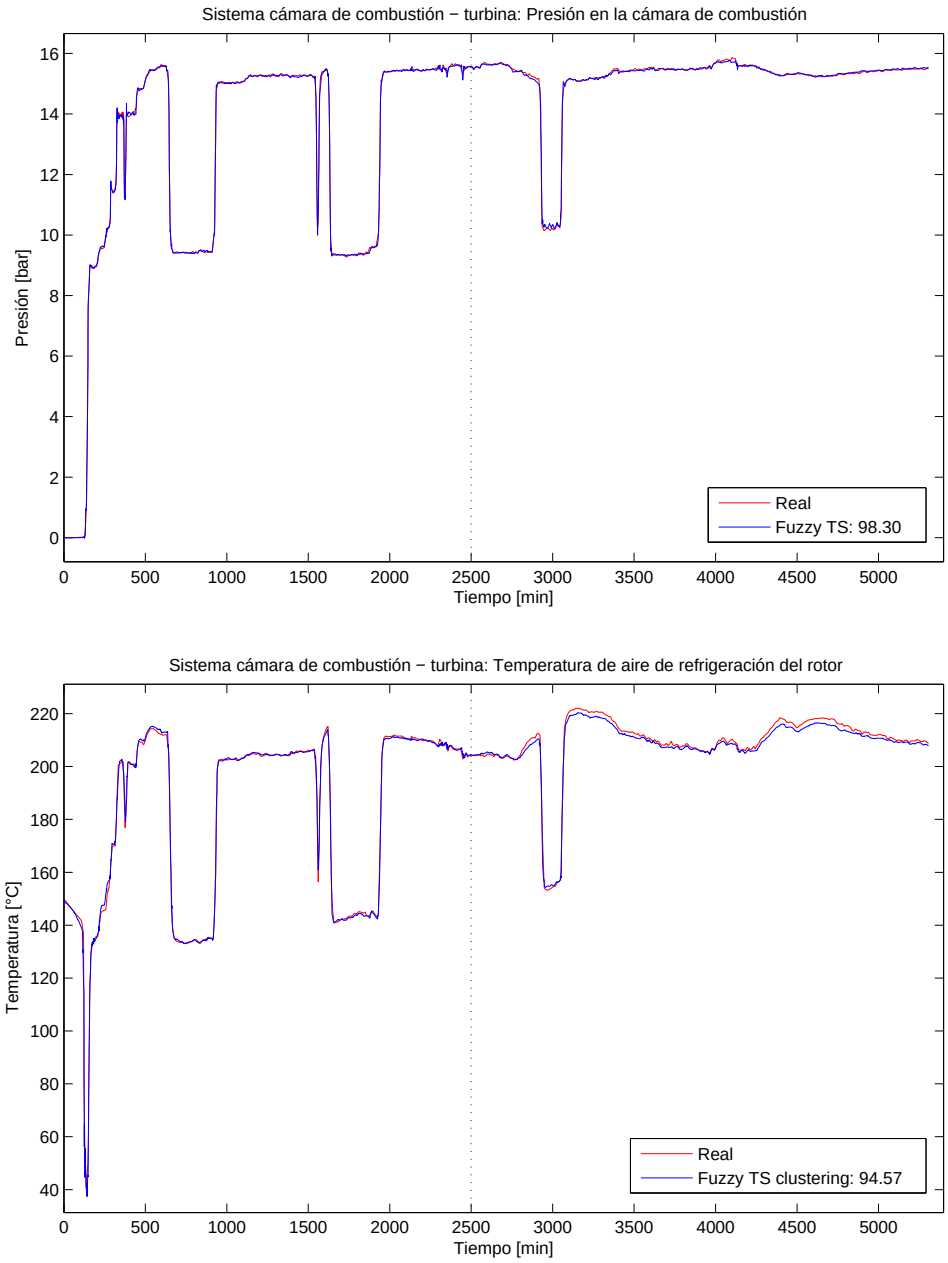

FIGURA 4.17. Mejor ajuste para la presión en la cámara de combustión y temperatura de aire de refrigeración del rotor.

Como puede verse, el ajuste de los modelos a los datos es bastante bueno, con errores cuadráticos medios normalizados que varían entre $6.9 \cdot 10^{-5}$ (caso compresor) y $2.5 \cdot 10^{-6}$ (caso velocidad del eje), lo que corresponde a errores en el índice de ajuste de 7.4% y 1.0% respectivamente. El resultado obtenido permite incluso usar estos modelos para realizar predicciones a corto plazo en aplicaciones de control predictivo.

4.4. Modelos obtenidos para predicción

Según el enfoque de predicción, se predice el valor de la salida para 1 o más instantes de tiempo, considerando toda la información disponible hasta el momento. El modelo recibe las lecturas de los valores tanto de entrada como de las salidas para generar las predicciones de las variables de salida en los siguientes instantes de tiempo. En cada modelo, se usan tanto las entradas exógenas como las retroalimentaciones autoregresivas para realizar la predicción de los siguientes instantes de tiempo.

Las siguientes tablas resumen los resultados de ajuste obtenidos para los enfoques de predicción.

TABLA 4.6. Resultados del ajuste de modelos para la predicción en el sistema IGV.

Tipo modelo	Ajuste (predicción 1 paso)		Ajuste (predicción 10 pasos)	
	Entrenamiento	Validación	Entrenamiento	Validación
ARX	95.11	95.70	69.49*	71.18*
Espacio de Estados	94.87	95.52	93.31	94.02
Fuzzy Takagi-Sugeno	96.13	92.47*	45.66*	53.19*
Fuzzy TS con Clustering	95.69	96.16	49.33*	52.18*

En estas tablas, los resultados marcados con * implican que fue necesario un filtraje para lograr el ajuste final. Esto se debe a las inestabilidades producidas por la identificación de parámetros asociados a las variables autoregresivas, comentadas anteriormente. Como puede desprenderse de las tablas 4.6, 4.7 y 4.8, el enfoque de predicción es más

TABLA 4.7. Resultados del ajuste de modelos para la predicción en el sistema del compresor.

Tipo modelo	Ajuste (predicción 1 paso)		Ajuste (predicción 10 pasos)	
	Entrenamiento	Validación	Entrenamiento	Validación
ARX	95.24	95.10	74.24	74.31
Espacio de Estados	99.21	99.12	90.43	88.31
Fuzzy Takagi-Sugeno	99.74	99.60	52.57*	53.65*
Fuzzy TS con Clustering	99.66	99.64	31.96	36.33

TABLA 4.8. Resultados del ajuste de modelos para la predicción en el sistema de combustible.

Tipo modelo	Ajuste (predicción 1 paso)		Ajuste (predicción 10 pasos)	
	Entrenamiento	Validación	Entrenamiento	Validación
ARX	98.69	98.88	82.14	84.34
Espacio de Estados	98.66	98.81	92.34	92.90
Fuzzy Takagi-Sugeno	99.04	99.12	78.30*	81.09*
Fuzzy TS con Clustering	98.99	99.08	79.78	82.21

TABLA 4.9. Resultados del ajuste de modelos para la predicción en el sistema Cámara de combustión - Turbina (primeras 3 salidas).

Tipo modelo	Porcentaje de ajuste (prediccion 1 paso)					
	POWER		SPEED		TGTEMP	
	Entr.	Valid.	Entr.	Valid.	Entr.	Valid.
ARX	97.31	97.07	99.09	99.07	95.03	95.07
Espacio Estados	96.97	96.86	99.00	99.00	96.68	96.67
Fuzzy TS	97.74	97.33	99.32	99.33	97.10	97.26
Fuzzy TS Cluster	97.72	97.37	99.88	99.88	98.34	98.25

efectivo que la simulación descrita en la sección anterior sólo hasta cierto número de pasos de predicción. En aplicaciones en que se requiere predecir por largos intervalos de tiempo, se hace más efectivo usar los modelos con el enfoque de simulación.

TABLA 4.10. Resultados del ajuste de modelos para la predicción en el sistema Cámara de combustión - Turbina (últimas 3 salidas).

Tipo modelo	Porcentaje de ajuste (predicción 1 paso)					
	HRTEMP		COMBPRES		RTRTEMP	
	Entr.	Valid.	Entr.	Valid.	Entr.	Valid.
ARX	99.60	99.59	98.89	98.93	98.54	98.66
Espacio Estados	99.40	99.40	97.23	97.05	98.03	98.26
Fuzzy TS	99.15	99.13	99.14	99.12	98.87	99.05
Fuzzy TS Cluster	99.69	99.67	99.24	99.21	99.50	99.35

No se han incluido los gráficos de ajuste para las predicciones de 1 paso ya que para todas las salidas, y con más razón para aquellas con mejor ajuste, es prácticamente indistinguible la salida real con la salida predicha.

4.5. Conclusiones

Como se aprecia en la sección 4.2, los modelos difusos de Takagi-Sugeno, en donde sus parámetros de salida son identificados linealmente mediante la fijación manual de los parámetros en las funciones de transferencia, son el balance más adecuado entre esfuerzo computacional y calidad en el ajuste de datos. A diferencia de los modelos neuronales tipo RBF, la fijación a priori de los parámetros correspondientes a las reglas (en el caso de los modelos neuronales, correspondientes a las neuronas de la capa oculta) entrega resultados muy parecidos a la identificación conjunta y óptima de los parámetros. Este resultado se debe al significado que tienen las reglas en los modelos difusos, que representan puntos de operación del sistema.

Un problema que se genera en la identificación es el hecho de que no se garantiza la estabilidad del modelo encontrado. La inclusión de salidas autoregresivas produjo inestabilidades en los modelos, que si bien son manejables en un corto plazo de predicción (de 1 a 5 pasos), se hacen muy significativas a largo plazo y en enfoque de simulación. Por esto que en los modelos con enfoque en simulación sólo se incluyen entradas exógenas y

sus regresores. Debido a este fenómento también en los modelos con enfoque de predicción, el desempeño de la identificación se reduce con el número de pasos de predicción considerados.

A pesar de estos problemas, los modelos obtenidos son lo suficientemente precisos para ser utilizados en simuladores o como base para el desarrollo de aplicaciones de control y monitorización.

5. DETECCIÓN Y DIAGNÓSTICO DE FALLAS EN SISTEMAS DINÁMICOS

5.1. Introducción

Todos los sistemas de la naturaleza (físicos, biológicos, mecánicos, etc.) pueden verse afectados por la ocurrencia de fallas en sus componentes, afectando con distinto grado de gravedad el funcionamiento general del proceso. Evidentemente, las posibilidades de que ocurran fallas en un sistema aumentan junto con la complejidad de éste. La complejidad de los procesos industriales aumenta día a día, acorde con su tamaño y grado de automatización. Así mismo, aumentan las posibilidades de ocurrencia de falla en algún componente. Debido a esto, en la actualidad ha crecido el grado de interés en la confiabilidad, seguridad y tolerancia a fallas en el diseño de procesos industriales y sistemas de control (Korbicz, Kościelny, Kowalczuk, & Cholewa, 2004).

En un sistema de control, las fallas pueden ocurrir en sensores, actuadores, y componentes del sistema a controlar, o en el hardware o software del controlador, como muestra la figura 5.1. Algunas de las fallas que afectan estos elementos pueden verse reflejadas en el sistema completo de forma inmediata, o ser anuladas por el lazo de control. Sin embargo, puede llegar un punto en que el lazo de control se vea sobrepasado y la falla haga sentir su efecto abrupta y poderosamente sobre el proceso. Una alternativa efectiva para tratar estos problemas es el control tolerante a fallas o FTC por sus siglas en inglés. En términos generales, un sistema de control tolerante a fallas es capaz de tomar acciones que permitan sobrellevar una falla menor del sistema y prevenir que sus efectos afecten al proceso completo.

Uno de los componentes clave de un sistema FTC son las aplicaciones de detección y diagnóstico de fallas (FDD por sus siglas en inglés). Éstas permiten detectar componentes con fallas, y asociarlos con la generación de alarmas o reconfiguraciones apropiadas para enfrentar el problema. Las aplicaciones de FDD también son clave para el concepto de FTM o *Fault Tolerant Measurement*, cuyo objetivo es asegurar la confiabilidad de las

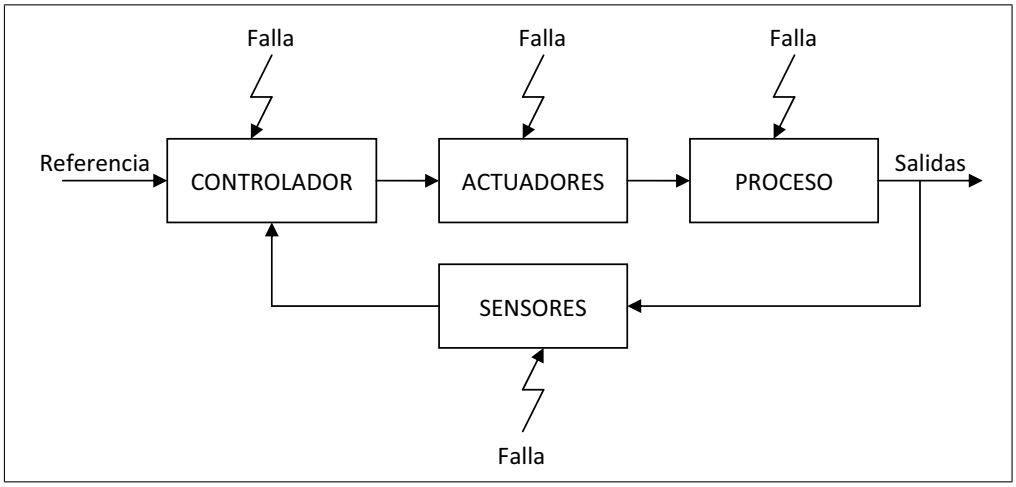

FIGURA 5.1. Esquema tradicional de control de un proceso, afectado por fallas.

mediciones de un proceso, reemplazando señales de sensores defectuosos por señales reconstruídas a partir de redundancias analíticas existentes o alguna otra fuente (Korbicz et al., 2004).

En este capítulo se comentan y analizan esquemas de detección y diagnóstico de fallas, con miras a la construcción de un sistema de medición tolerante a fallas (FTM) para una turbina de gas, presentado en el capítulo siguiente.

5.2. Metodologías de detección y diagnóstico

Existen diversas metodologías documentadas en la literatura para realizar detección y diagnóstico de fallas en un proceso. Una de las formas más básicas para realizar esta tarea es la denominada *redundancia física o redundancia de hardware*, que consiste en usar múltiples sensores y actuadores para realizar una misma tarea de medición y control de una determinada variable (Simani, Patton, & Fantuzzi, 2003). Bajo este esquema, se utiliza tradicionalmente el sistema de "votación" para decidir si ha ocurrido una falla entre los componentes redundantes. Este sistema consiste en discriminar el componente redundante que se comporta distinto de sus pares, siendo minoría.

Evidentemente, este enfoque presenta el problema del costo del equipamiento extra y de la mantención, además del espacio requerido para realizar la instalación de lo múltiples componentes.

Otro enfoque, en vista del costo en el que se incurre en la redundancia física, es realizar comparaciones cruzadas entre distintas mediciones del proceso. Estas comparaciones están basadas en relaciones analíticas entre variables de proceso a través de ecuaciones físicas, químicas, etc. A estas comparaciones se les denomina *redundancia analítica* (Chen & Patton, 1999).

Al usar la redundancia analítica, se realiza una comparación entre las diferenctes variables generando diferencias llamadas *síntomas o residuos*. Bajo condiciones normales de operación, los residuos generados deberian ser cero o valores muy cercanos a cero, mientras que al ocurrir una fallas el valor de los residuos es distinto de cero. De esta forma, los residuos permiten determinar cuando ocurre una falla en el proceso (Chen & Patton, 1999).

Usualmente la redundancia analítica se realiza comparando las señales medidas del proceso con valores estimados para dichas señales. La estimación se realiza a través de un modelo matemático del proceso, y los residuos se generan por la diferencia entre los valores medidos y estimados. También es posible incorporar síntomas heurísticos (como color, mal olor, presencia de humo, etc), integrando el conocimiento y la experiencia de operadores a través de sistemas expertos (Patton, Frank, & Clark, 2000). La forma más típica de realizar el diagnóstico es estableciendo cotas para los residuos, de tal manera que si dichas cotas son sobrepasadas, se determina un estado de falla (Chen & Patton, 1999). El problema de esta metodología es la variación de las condiciones del proceso según los distintos estados de operación, y la presencia de ruido en las mediciones. Por otro lado, la gran ventaja de este enfoque es que no requiere hardware adicional para realizar las tareas de diagnóstico. Los esquemas de detección y diagnóstico de fallas con redundancia analística que usan un modelo matemático del proceso para generar comparaciones, se denominan como *basados en modelos* (Simani et al., 2003).

Por otro lado, si sólo se cuenta con mediciones de las salidas del proceso, pueden aplicarse los llamados métodos basados en procesamiento de señales, en donde se trabaja en función de la detección de vibraciones, patrones específicos u otros comportamientos que se puedan identificar según cada proceso. Algunos métodos típicos basados en este enfoque son (Simani et al., 2003):

- Filtros pasabanda.
- Análisis espectral (basado en FFT).
- Transformadas Wavelet.
- Estimaciónd de entropías máximas.

También puede realizarse detección y diagnóstico de fallas basandose en las propiedades estadísticas de las variables, especialmente cuando existe influencia de componentes estocásticas. Las desviaciones con respecto a un comportamiento normal son detectadas con metodos como los siguientes:

- Estimacion de medias y varianzas.
- Árboles de decisión de Bayes.
- Test de maxima verosimilitud.
- Test basados en sumas acumulativas.

En cuanto a la identificación de fallas a partir de los síntomas generados, existen diversas técnicas documentadas en la literatura. Entre ellas, destacan los métodos de distancias geométricas o basados en cotas, los métodos probabilísticos y los que emplean sistemas inteligentes. Para cada uno de estos métodos, existe un sinnúmero de estructuras y combinaciones de técnicas documentadas (Chen & Patton, 1999).

5.3. Detección y diagnóstico de fallas basado en modelos

La detección y diagnóstico de fallas basada en modelos para sistemas dinámicos ha concentrado gran atención en las últimas dos décadas. Y no tan sólo como aplicaciones teóricas, también como aplicaciones industriales. Se puede encontrar una gran variedad de

métodos en la literatura, que usan modelos matemáticos del proceso, combinándolos con las teorías de control moderno (Simani et al., 2003).

El diagnóstico de fallas basado en modelos, como se dijo anteriormente, consiste en generar diferencias entre las variables medidas del proceso y sus estimaciones generadas por un modelo matemático. Estas diferencias, llamadas residuos, son cercanas a cero en condición normal y distintas de cero en presencia de falla. Se asume que ante una falla el valor del residuo debe variar significativamente respecto al estado de operación normal, para que la falla sea detectable. Se debe diseñar el sistema de detección con un número de residuos suficiente para tener sensibilidad sobre las fallas individuales que afectan a cada componente. El análisis de los residuos, por ejemplo cuando una cota es superada, conducen a la identificación de la falla.

La figura 5.2 muestra la estructura general de un sistema FDD basado en modelos.

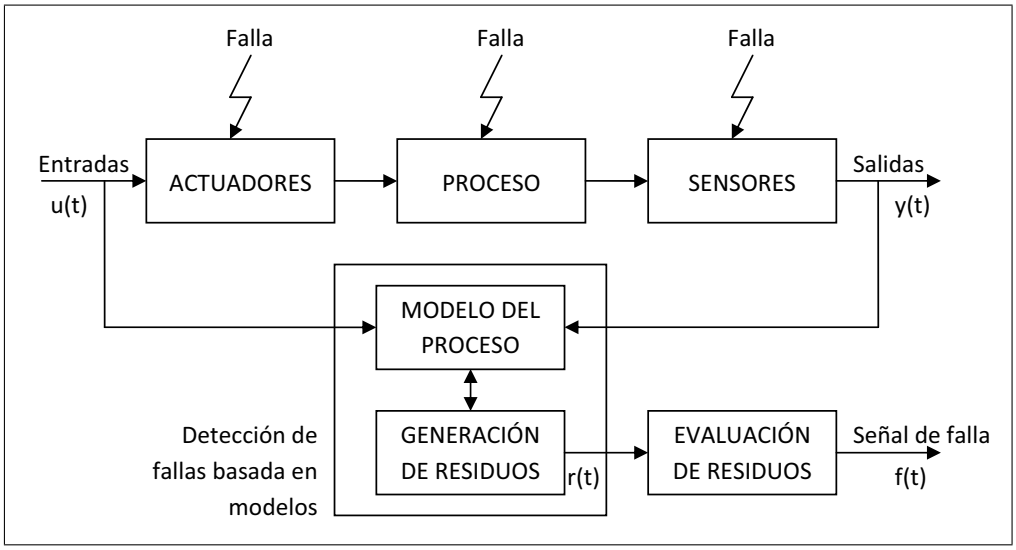

FIGURA 5.2. Estructura de un sistema FDD basado en modelos.

El bloque de generación de residuos, determina las señales de residuo en base a las mediciones disponbiles de variables de entrada y salida del proceso. Los residuos son cero o muy cercanos a cero en condición normal y significativamente distintos de cero en condición de falla, de forma que la presencia de una falla pueda detectarse. El bloque de evaluación de residuos examina los residuos generados en busca de fallas. La presencia de

falla se determina en base a reglas de decisión sobre el residuo. Esto puede significar una simple evaluación en base a cotas, distancias geométricas, o métodos estadísticos como el test de radio de verosimilitud (*generalized likelihood ratio test*), árboles de decisión, etc. (Patton et al., 2000).

Las fallas ocurridas en el proceso, desde el punto de vista matemático, pueden ser de carácter aditivo o multiplicativo. Considerando fallas aditivas, el esquema de la figura 5.1 puede reformularse como muestra la figura 5.3, incluyendo el posible sistema FDD.

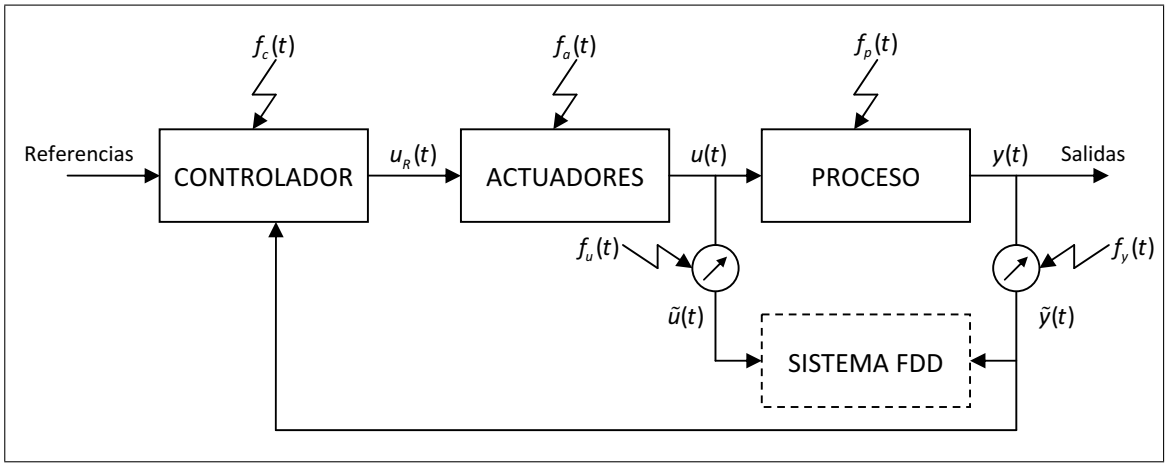

FIGURA 5.3. Esquema del proceso sometido a fallas con un sistema FDD.

En este caso, se considera que:

$$\begin{aligned} u_R(t) &= G_C(\tilde{y}(t), R(t)) + f_c(t) \\ u(t) &= u_R(t) + f_a(t) \\ y(t) &= G(u(t)) + f_p(t) \\ \tilde{u}(t) &= u(t) + f_u(t) \\ \tilde{y}(t) &= y(t) + f_y(t) \end{aligned} \tag{5.1}$$

En donde G_C representa la función de transferencia del controlador y G la del proceso. El sensor de variables de entrada u también podría estar ubicado a la entrada de los actuadores. En ese caso, se deben reemplazar las relaciones de (5.1) según corresponda.

5.3.1. Esquemas de generación de residuos

En base al esquema general de la detección y diagnóstico de fallas de la figura 5.2, la generación de residuos se representa de forma general como muestra la figura 5.4 (Basseville, 1988).

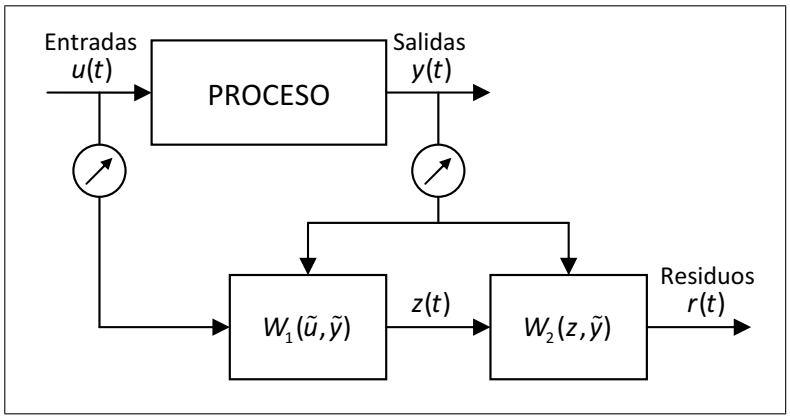

FIGURA 5.4. Estructura general de un generador de residuos.

En esta estructura, existe una señal auxiliar $z(t)$ que se genera por la función W_1, mientras que el residuo se calcula a partir de $z(t)$ y la medición de la salida $\tilde{y}(t)$ a través de W_2. Idealmente, en ausencia de falla se debe cumplir que

$$\begin{aligned} z(t) &= W_1(\tilde{u}, \tilde{y}) \\ r(t) &= W_2(z, \tilde{y}) = 0 \end{aligned} \tag{5.2}$$

El esquema de generación de residuos más usado se muestra en la figura 5.5 (Simani et al., 2003). En este caso, la función W_1 es un modelo del proceso implementado a través de un simulador o un observador. La función W_2 calcula el residuo como la diferencia entre la salida medida del proceso y la salida del simulador u observador.

Una vez generado el residuo, el método más simple y usado para la detección de la falla (evaluación del residuo) es comparar directamente la señal $r(t)$ o una función del residuo $J(r(t))$ con una cota fija ϵ o variable $\epsilon(t)$ de forma que

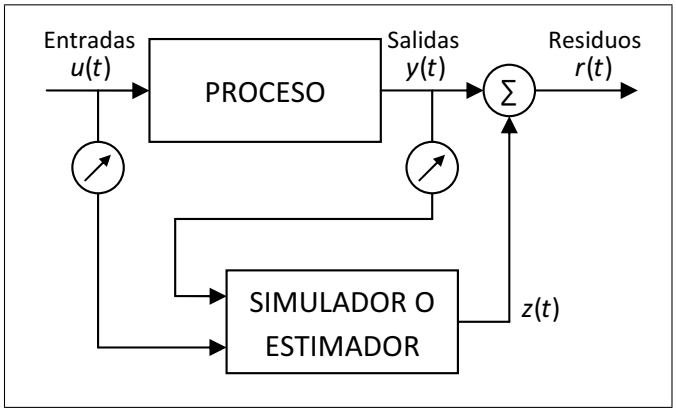

FIGURA 5.5. Esquema típico de generación de residuos.

$$\begin{aligned} J(r(t)) &\leq \epsilon(t) \quad \text{cuando} \quad f(t) = 0 \\ J(r(t)) &> \epsilon(t) \quad \text{cuando} \quad f(t) \neq 0 \end{aligned} \tag{5.3}$$

en donde $f(t)$ es un vector que contiene las señales de falla que se están detectando f_u, f_y, f_c, f_a y/o f_p señaladas en la figura 5.3. Cuando el proceso opera en régimen permanente en torno a un único estado de operación, las cotas fijas funcionan bastante bien para la detección (Simani et al., 2003). Cuando el proceso se mueve entre diversos estados durante la operación, cotas adaptivas son mucho más apropiadas (Chen & Patton, 1999).

Existen diversos métodos estudiados en la literatura, entre los cuales los más clásicos son las siguientes (Chen & Patton, 1999; Patton et al., 2000; Simani et al., 2003; Korbicz et al., 2004):

(i) Generación de residuos mediante estimación de parámetros.

(ii) Generación de residuos basado en observadores, o mediante estimación de estado.

(iii) Generación de residuos con ecuaciones de paridad.

5.3.2. Generación de residuos con estimación de parámetros

En la mayoría de los casos reales, los parámetros de un proceso no se conocen en su totalidad. Si la estructura del modelo del proceso es conocida, los parámetros pueden ser

estimados por ejemplo con los algoritmos estudiados en la sección 3.4, basados en los datos de entrada y salida del proceso.

La generación de residuos en mediante la estimación de parámetros se basa en la suposición de que las fallas del proceso se ven reflejadas en los parámetros del modelo. El procedimiento de detección consiste básicamente en realizar una estimación de parámetros en línea del proceso, y compararlos con los parámetros de un modelo de referencia, obtenido bajo una operación normal de la planta (Patton et al., 2000; Simani et al., 2003; Araya, 2004).

Existen básicamente dos métodos en esta línea: los métodos de ecuación de error, y los métodos de error de salida. Los métodos de ecuación de error consisten básicamente en considerar un modelo como el descrito en la ecuación (3.42), de la forma

$$\hat{y}(t) = \theta^T x(t) \tag{5.4}$$

La ecuación de error es

$$e(t) = y(t) - \hat{y}(t) = y(t) - \theta^T x(t) \tag{5.5}$$

El vector de parámetros θ se realiza minimzando la suma de los errores $e(t)$ al cuadrado, como se detalla en la seccion 3.4.1. La solución es

$$\hat{\theta} = (X^T X)^{-1} X^T Y \tag{5.6}$$

Esta solución puede aplicarse sobre una ventana de tiempo, o realizar una estimación de mínimos cuadrados recursivos, cuya solución viene dada por (Passino & Yurkovich, 1997):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t)x(t)\left(y(t) - \theta(t-1)^T x(t)\right) \tag{5.7}$$

donde

$$P(t) = P(t-1) - \frac{P(t-1)x(t)}{1 + x(t)^T P(t-1)x(t)} x(t)^T P(t-1) \tag{5.8}$$

Para mejorar la estimación, pueden aplicarse filtros. Por ejemplo en caso de que exista la presencia de ruido, pueden usarse filtros de Kalman.

En el caso de los metodos de error de salida, el error viene dado por

$$e(t) = y(t) - \hat{y}(\theta, t) \tag{5.9}$$

en donde

$$y(z) = \frac{\hat{B}(z)}{\hat{A}(z)} u(z) \tag{5.10}$$

En este caso, en que se utiliza directamente la función de transferencia del modelo, $e(t)$ es no lineal en los parámetros y no puede aplicarse el procedimiento de la sección 3.4.1. La estimación de parámetros se realiza por algoritmos numéricos, por ejemplo, algoritmos de descenso en la dirección del gradiente, lo que incrementa significativamente el esfuerzo computacional.

Los dos esquemas pueden verse gráficamente en la figura 5.6.

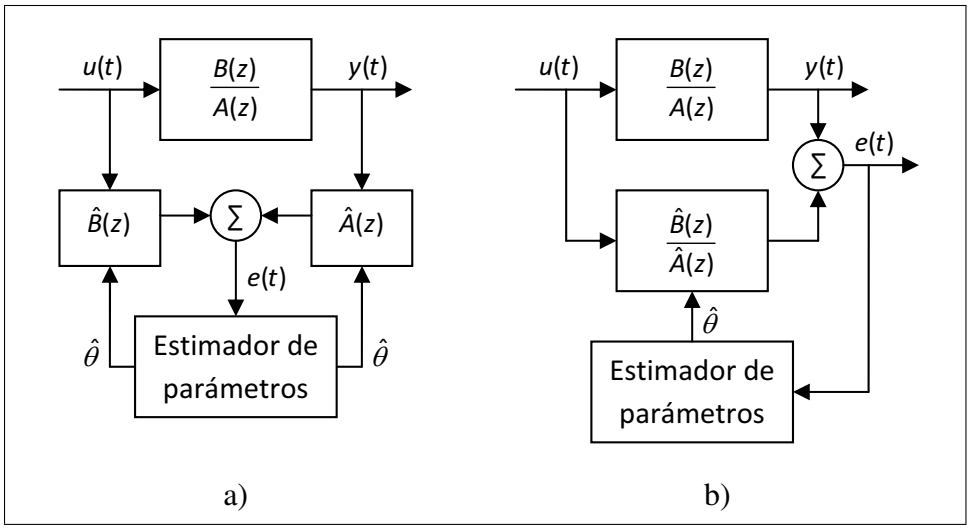

FIGURA 5.6. Estimación de parámetros. a) Ecuación de error. b) Error de salida.

5.3.3. Generación de residuos con estimación de estado

Básicamente, la generación de residuos con estimación de estado consiste en estimar las salidas del sistema a partir de las mediciones disponibles a través de observadores (observador de Luenberger en el caso determinístico, y filtro de Kalman en el caso con pertubaciones estocásticas). El error de estimación se usa como residuo (Simani et al., 2003).

En este caso, se considera un modelo dinámico lineal del proceso, en tiempo discreto y con parámetros invariantes en el tiempo, de la forma

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \tag{5.11}$$

Asumiendo que las matrices A, B y C son conocidas (lo que es bastante restrictivo), las variables no medidas se reconstruyen a través de un observador:

$$\hat{x}(t+1) = A\hat{x}(t) + Bu(t) + He(t) \tag{5.12}$$

$$e(t) = y(t) - C\hat{x}(t) \tag{5.13}$$

El error de estimación de estado es

$$e_x(t) = x(t) - \hat{x}(t) \tag{5.14}$$

$$e_x(t+1) = (A - HC)e_x(t) \tag{5.15}$$

Cuando el observador es estable, el error tiende a cero cuando $t \to \infty$. En caso de presencia de perturbaciones y fallas, el modelo de (5.11) toma la siguiente forma:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + Qv(t) + L_1 f(t) \\ y(t) &= Cx(t) + Rw(t) + L_2 f(t) \end{aligned} \tag{5.16}$$

en donde $v(t)$ y $w(t)$ corresponden a perturbaciones no medidas, $f(t)$ a las fallas que actuan sobre los estados y la salida a través de L_1 y L_2 (Patton et al., 2000). Acá, $f(t)$ tiene efectos aditivos ya que influyen sobre los estados y sobre el error en forma de suma. Si las fallas aparecen como cambios en los parámetros, el modelo (5.11) queda:

$$\begin{aligned} x(t+1) &= (A + \Delta A)x(t) + (B + \Delta B)u(t) \\ y(t) &= Cx(t) \end{aligned} \tag{5.17}$$

En este caso, los cambios ΔA y ΔB representan fallas con efecto multiplicativo (Patton et al., 2000)

Los métodos más comunes conocidos para detectar fallas bajo esta estructura son los siguientes (Chen & Patton, 1999; Patton et al., 2000; Simani et al., 2003):

(i) Observadores para procesos de múltiples salidas

- Un observador, excitado por una salida: un observador es activado por un sensor de variables de salida. Las otras salidas son reconstruídas y comparadas con sus mediciones.
- Banco de observadores, excitados por todas las salidas: varios observadores son diseñados para una señal de falla definida, y detectada a través de tests de hipótesis.
- Banco de observadores, excitados por una salida: múltiples observadores, uno para cada sensor. Las salidas estimadas se comparan con las salidas medidas.
- Banco de observadores, excitados por todas las salidas menos una, igual que el punto anterior pero los observadores se alimentan con todas las salidas menos la que se estudia.

(ii) Filtros para detección de falla en procesos de múltiples salidas

- La matriz de realimentación H del estimador de estado se elije de forma particular para las fallas $L_1 f(t)$ en una direccion determinada, y para fallas $L_2 f(t)$ en un plano determinado.

5.3.4. Generación de residuos con ecuaciones de paridad

La idea básica de las ecuaciones de paridad es realizar una comprobación de la consistencia de las medidas tomadas del proceso. Se considera una la planta de la forma

$$G_P(z) = \frac{B(z)}{A(z)} \tag{5.18}$$

y un modelo de la forma

$$G_M(z) = \frac{\hat{B}(z)}{\hat{A}(z)} \tag{5.19}$$

en donde tanto la estructura como los parámetros del estimador y de la planta se conocen *a priori* (Simani et al., 2003). Las fallas del sistema pueden reflejarse a través de una señal $f_u(t)$ en las variables de entrada y $f_y(t)$ en las variables de salida, luego

$$y(z) = G_P(z)u(z) + G_P(z)f_u(z) + f_y(z) \tag{5.20}$$

El vector de error o residuo de salida se forma como

$$r(z) = y(z) - \hat{y}(z) = (G_P(z) - G_M(z))u(z) + G_P(z)f_u(z) + f_y(z) \tag{5.21}$$

Asumiendo que $G_P(z) = G_M(z)$, se tiene que

$$r(z) = \frac{B(z)}{A(z)}f_u(z) + f_y(z) \tag{5.22}$$

También es posible formular el residuo de forma polinomial a través de una ecuación de error, como se muestra en la figura 5.7:

$$r(z) = \hat{A}(z)y(z) - \hat{B}(z)u(z) \tag{5.23}$$

$$= B(z)f_u(z) + A(z)f_y(z) \tag{5.24}$$

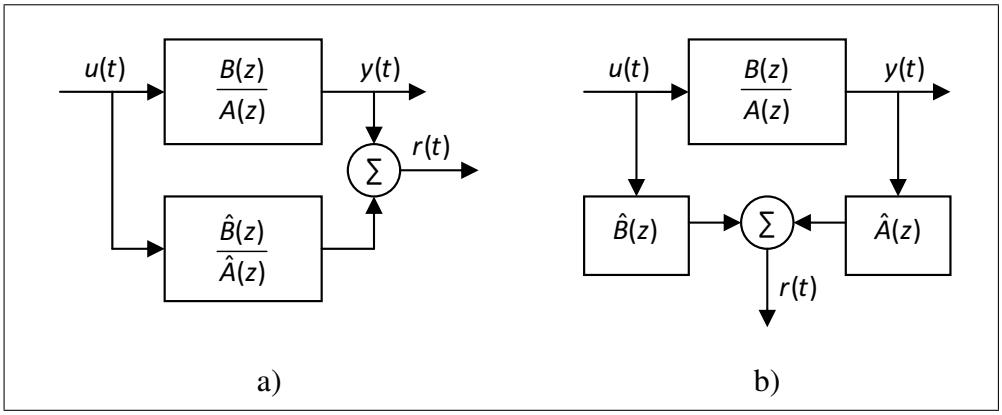

FIGURA 5.7. Esquema de ecuaciones de paridad. a) Error de salida. b) Ecuación de error.

Las ecuaciones (5.22) y (5.24) generan los residuos y son llamadas *ecuaciones de paridad*, bajo las suposiciones de ocurrencia de falla y de similitud total entre el modelo G_M y la planta G_P. La diferencia entre la estimación de parámetros y las ecuaciones de

paridad, es que en esta estructura se asumen los parámetros conocidos y constantes en todo momento, mientras que en la estimación de parámetros pueden variar todo el tiempo con el objetivo de minimizar el error. Para mejorar la estimación, el vector de errores puede filtrarse con una función $G_f(z)$ (Patton et al., 2000).

Las ecuaciones de paridad también pueden extenderse a sistemas en ecuaciones de estado. Considerando el modelo planteado en (5.11), y substituyendo la salida recursivamente, se puede formar la ecuación:

$$\begin{bmatrix} y(t) \\ y(t+1) \\ y(t+2) \\ \vdots \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} x(t) + \begin{bmatrix} 0 & 0 & 0 & \cdots \\ CB & 0 & 0 & \cdots \\ CAB & CB & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} u(t) \\ u(t+1) \\ u(t+2) \\ \vdots \end{bmatrix} \tag{5.25}$$

$$Y(t) = Tx(t) + QU(t) \tag{5.26}$$

Para formar la ecuación de paridad, es necesario eliminar la dependencia de las variables no medidas $x(t)$. Multiplicando la ecuación (5.26) por una matriz W tal que $WT = 0$, se obtiene un vector de residuos de la forma

$$r(t) = WY(t) - WQU(t) \tag{5.27}$$

5.3.5. Generación de residuos con otras técnicas

Aparte de las técnicas mencionadas anteriormente, existe una gran cantidad de otras técnicas para la generación de residuos. Muchas de ellas son modificaciones a partir de los tres enfoques clásicos tratados en las secciones previas.

Los métodos clásicos de diagnóstico de fallas basados en modelos suponen que los efectos de las fallas se ven reflejados en los estados o en las salidas, ya sean fallas en los componentes o en los sensores. El principal problema con estas técnicas es que la precisión del modelo del proceso afecta directamente la eficacia de la detección y la sensibilidad del diagnóstico (Patton et al., 2000). Más aún, la mayoría de los proceso industriales tienen

un comportamiento no lineal, y no pueden ser modelados a través de las técnicas mencionadas anteriormente. Muchas veces ni siquiera es posible formular modelos analíticos de procesos demasiado complejos, como se vio en el capítulo 3.

Por lo anterior, también son utilizados principalmente los modelos difusos y las redes neuronales, por su capacidad de modelar proceso no lineales, además de la capacidad de incorporar información heurística como la experiencia de los operadores, en el caso de los modelos difusos. La desventaja, como ya se vio anteriormente, es que se deben contar con datos de entrada y salida del proceso para realizar la estimación de sus parámetros. En el caso de obtener datos para el proceso en condiciones normales, en general no hay mayores problemas ya que todas las plantas industriales modernas cuentan con sistemas SCADA que reciben mediciones en línea de las variables importantes provenientes de los sensores. El problema surge al requerir datos de operación en estado de falla. Estos datos en general no se encuentran disponibles, y someter el proceso a la falla para obtener la información es impensado. Debido a que los parámetros y estructura en estos modelos no guardan relación con la estructura y parámetros físicos, es imposible generar artificialmente una falla sin un conocimiento muy acabado.

La forma más habitual en como se aplican estas técnicas es según el esquema clásico de la figura 5.5.

5.4. Detección y diagnóstico de fallas basado en análisis estadístico

Los métodos clásicos de detección de fallas basada en modelos descritos en la sección 5.3 plantean que una falla causa cambios en ciertos parámetros físicos y mediciones que conducen a un cambio en parámetros y estados del modelo. Éstos métodos requieren un conocimiento *a priori* muy acabado acerca de la relación de las fallas y los parámetros o estados del modelo. Esto es extremadamente difícil de realizar en procesos industriales complejos, y en muchos casos imposible, por lo que en general sólo pueden abordarse ciertas fallas bajo determinadas condiciones.

Por otro lado, las técnicas de modelación inteligente descritas en el capítulo 3 permiten modelar efectivamente procesos complejos y no lineales. Sin embargo, tal como se comenta en la sección 5.3.5, requieren de gran cantidad de información disponible.

Otro enfoque en la detección y el diagnóstico de fallas es el análisis estadístico de las señales del proceso. Existen técnicas estadísticas que pueden aplicarse a las señales disponibles para determinar la posible presencia de fallas o detectar anomalías en la operación, prescindiendo de modelos o de gran cantidad de información disponible del proceso.

El método más simple consiste en verificar límites de operación para las señales (Korbicz et al., 2004):

$$y_{MIN} \leq y(t) \leq y_{MAX} \tag{5.28}$$

Si la señal $y(t)$ sale fuera del rango establecido por las cotas, se establece la presencia de alguna falla o anomalía. Evidentemente, el diagnóstico de la falla es muy difícil en este caso, debido a la poca información con la que se cuenta. Los valores de las cotas para el rango aceptable se establecen según el conocimiento experto que se tiene sobre el proceso. Estas cotas también pueden aplicarse a las razones de cambio de la señal:

$$\Delta y_{MIN} \leq y(t) - y(t-1) \leq \Delta y_{MAX} \tag{5.29}$$

Esta forma de detección puede mejorarse usando zonas muertas, límites con histéresis, etc. Sin embargo, tanto la capacidad de detección de falla como el error cometido con falsos positivos es bastante limitado. Las cotas pueden ser adaptivas, en función del tiempo, pero establecer dichas funciones puede ser muy complejo.

A continuación se muestran algunos métodos estadísticos que se aplican a las señales cuando se trabaja en ausencia de modelos.

5.4.1. Detección de cambios en índices estadísticos

Cuando la variable del proceso que se está monitorizando tiene un comportamiento estocástico, se puede establecer su media y su varianza como:

$$\bar{y} = E\{y(t)\} \tag{5.30}$$

$$\sigma_y^2 = E\{[y(t) - \bar{y}]^2\} \tag{5.31}$$

En un caso de comportamiento normal del proceso, los valores de $\bar{y}$ y σ_y^2 son generalmente constantes. Se pueden obtener índices de cambio de estos indicadores para generar síntomas del proceso, de la forma (Patton et al., 2000):

$$\Delta y = E\{y(t) - \bar{y}\} \tag{5.32}$$

$$\Delta\sigma = E\{\sigma_y(t) - \sigma_y\} \tag{5.33}$$

Estos indicadores de cambio se realizan respecto a los estadísticos obtenidos en el caso de operación normal. Usualmente suele tomarse una cota de la forma (Simani et al., 2003)

$$|\Delta y| \leq \epsilon\sigma_y, \quad \epsilon \geq 2 \tag{5.34}$$

También se puede realizar la detección de un cambio en base a estos síntomas, aplicando tests de cambio de media o cambio de varianza, como el test t o el test de χ^2, test F para comparación de varianzas, etc (Basseville & Nikiforov, 1993). Para manejar múltiplies cambios sobre diversas señales, puede construirse un vector de cambio

$$\Delta Y = [\Delta y_1, \ldots, \Delta y_m] \tag{5.35}$$

En este caso, pueden aplicarse algoritmos de distancia geométrica o clasificaciones con redes neuronales (Simani et al., 2003).

Otra opción es aplicar algoritmos basados en sumas acumulativas o CUMSUM por sus iniciales en inglés, en donde las pruebas estadísticas se realizan sobre un índice generado por la suma acumulativa de valores en una ventana móvil, lo que permite filtrar los datos y reducir variabilidades estadísticamente no significativas (Patton et al., 2000)

En caso de aplicarse a parámetros, se utiliza también el test de verosimilitud *generalized likelihood ratio*. Este test también usa el indicador formado con sumas acumulativas (Basseville & Nikiforov, 1993).

5.4.2. Análisis de componentes principales

El análisis de componentes principales es una técnica que permite transformar un número determinado de muestras de variables, en un número (si se desea, menor) de variables a través de una proyección. Una de las características principales de esta transformación, es que las nuevas variables no están correlacionadas.

Supongamos que se tienen m mediciones de n variables estocásticas con distribución gaussiana, de media nula y varianza unitaria, agrupadas en los vectores $x_1, \ldots, x_n$. Se forma la matriz X de la forma

$$X = [x_1, \ldots, x_n]^T \tag{5.36}$$

Cuando los datos no poseen las características de media nula y varianza unitaria, se pueden normalizar a través de la relación (Xu & Wang, 2006)

$$\bar{X} = D_s^{-1}(X - E\{x\})) \tag{5.37}$$

en donde D_s corresponde a una matriz diagonal cuyos elementos son las varianzas de las variables, y $E\{x\} = [\mu_1, \ldots, \mu_n]$ son las medias de las variables x. La matriz de covarianza de los n vectores de variables es

$$S = \begin{bmatrix} s_1^2 & s_{12} & \cdots & s_{1n} \\ s_{12} & s_2^2 & \cdots & s_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{1n} & s_{2n} & \ldots & s_n^2 \end{bmatrix} \tag{5.38}$$

en donde s_i^2 es la varianza de la variable i, y s_{ij} es la covarianza entre las variables i y j. La descomposición en vectores y valores propios de la matriz S corresponde a

$$U^T S U = L \tag{5.39}$$

En donde U es una matriz ortonormal cuyas columnas son los vectores propios de S y L es una matriz diagonal con los valores propios de S. La transformación al espacio de componentes principales consiste en (Jackson, 2003)

$$z = U^T x \tag{5.40}$$

en donde x corresponde a un vector $n \times 1$ con observaciones de las n variables originales. Las variables transformadas se denominan los *componentes principales* de x, y las observaciones individuales transformadas se denominan puntajes-z (*z-scores*). El componente principal i-ésimo corresponde a

$$z_i = u_i^T x \tag{5.41}$$

Una de las características de estas nuevas variables o componentes principales es que son no correlacionadas entre si, y el sistema de coordenadas que ellas establecen maximiza la varianza de los datos en la dirección de cada uno de los componentes principales. Como U es ortonormal, se cumple que $U^T U = UU^T = I$, con I la matriz identidad. Luego, se puede establecer la transformación inversa desde el espacio de componentes principales al espacio de las variables originales x como

$$x = Uz \tag{5.42}$$

Los componentes principales se utilizan en detección y diagnóstico para establecer límites de funcionamiento como los descritos en la sección anterior. La ventaja en este caso es que al ser las variables z_i no correlacionadas entre si, el error tipo I (casos de falla no detectados) que se comete al establecer los límites de operación con un grado α de confianza en base a la distribución t puede calcularse con exactitud, siendo $1 - \alpha^n$ considerando un nivel de confianza α para cada variable. En el caso de las variables x, generalmente con un nivel de correlación distinto de cero, es prácticamente imposible determinar el error (Jackson, 2003).

Los valores propios l_i representan la varianza asociada al componentes principal i. Sin pérdida de generalidad, se puede asumir que la matriz L en (5.39) contiene los valores

propios de S, l_i, ordenados de mayor a menor. El valor

$$\hat{s} = \frac{\sum_{i=1}^{r} l_i}{\sum_{i=1}^{n} l_i} \cdot 100 \tag{5.43}$$

representa entonces el porcentaje de la varianza del proceso compuesto por las variables x, explicado por los $r < n$ primeros componentes principales z. Si sólo se seleccionan los vectores asociados a los r valores propios más significativos para formar una matriz $\hat{U}$, y según la ecuación (5.42), se obtiene una aproximación de los valores x de la forma

$$\hat{x} = \hat{U}z \tag{5.44}$$

Al espacio generado por $\hat{U}$ se le denomina espacio de componentes principales, y al espacio generado por $\tilde{U} = I - \hat{U}$ se le denomina espacio residual. En general, se asocia la varianza explicada por los mayores componentes principales a variabilidades propias del proceso, mientras que a los más pequeños se les asocian los efectos de variabilidades no explicados por el proceso como ruido (Jackson, 2003).

El uso de componentes principales suele pasar por aplicar las técnicas descritas en la sección anterior a los componentes principales z o a los valores filtrados $\hat{X}$ (Jackson, 2003).

5.4.3. Test T^2 de Hotelling

El test T^2 de Hotelling es una generalización multivariable del test de t de Student, que permite establecer rangos para la variabilidad general de las mediciones y ver si se encuentran dentro de límites de operación normal. Considerando los supuestos sobre los datos x realizados en la sección anterior, se forma el índice T^2 de la forma (Jackson, 2003)

$$T^2 = x^T S^{-1} x \tag{5.45}$$

Donde S es la matriz de covarianza definida en la sección anterior. Despejando S de la ecuación (5.39) se obtiene

$$S = ULU^T \tag{5.46}$$

Por lo tanto, se cumple que

$$S^{-1} = (ULU^T)^{-1} \tag{5.47}$$

$$S^{-1} = UL^{-1}U^T \tag{5.48}$$

Reemplazando esto en la ecuación (5.45), se obtiene

$$T^2 = x^T UL^{-1}U^T x \tag{5.49}$$

$$T^2 = z^T L^{-1} z \tag{5.50}$$

Luego, el índice T^2 puede obtenerse también a través de los componentes principales z. De la misma forma, para cancelar efectos del ruido y mantener solo la variabilidad atribuída al proceso, pueden seleccionarse $r < n$ componentes principales para generar el indice, o los valores filtrados $\hat{x}$. El nivel de confianza α para el índice T^2, considerando m mediciones de n variables, se asocia con la distribución F de Fisher de la siguiente forma (Jackson, 2003)

$$T^2_{n,m,\alpha} = \frac{n(m-1)}{m-n} F_{n,m-n,\alpha} \tag{5.51}$$

Cuando el índice T^2 calculado excede $T^2_{n,m,\alpha}$, se dice que el proceso esta "fuera de control" o ante la presencia de una falla. Este índice permite eliminar el problema del cálculo del error tipo I al existir varias variables involucradas.

5.4.4. Test Q

El estadístico Q está asociado directamente con la descomposición en componentes principales. Como se aprecia en la ecuación (5.44), los valores de X pueden ser aproximados o filtrados por los r componentes principales más relevantes, dando origen a una aproximación $\hat{X}$. De esta forma:

$$X = \hat{X} + (X - \hat{X}) \tag{5.52}$$

$$X = \hat{U}\hat{Z} + (UZ - \hat{U}\hat{Z}) \tag{5.53}$$

$$X = \hat{U}\hat{U}^T X + (UU^T - \hat{U}\hat{U}^T)X \tag{5.54}$$

$$X = \hat{C}X + \tilde{C}X \tag{5.55}$$

El espacio generado por la matriz $\hat{C}$ se denomina espacio de componentes principales (también llamado núcleo), mientras que el espacio generado por $\tilde{C}$ se denomina espacio residual. El índice Q o *square prediction error* SPE se calcula como

$$Q = (x - \hat{x})^T (x - \hat{x}) = ||\tilde{C}x||^2 \tag{5.56}$$

Considerando las siguientes definiciones

$$\theta_1 = \sum_{i=r+1}^{n} l_i \tag{5.57}$$

$$\theta_2 = \sum_{i=r+1}^{n} l_i^2 \tag{5.58}$$

$$\theta_3 = \sum_{i=r+1}^{n} l_i^3 \tag{5.59}$$

$$h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2} \tag{5.60}$$

Es posible demostrar que el valor c definido como

$$c = \theta_1 \frac{\left[\left(\frac{Q}{\theta_1}\right)^{h_0} - \frac{\theta_2 h_0(h_0-1)}{\theta_1^2} - 1\right]}{\sqrt{2\theta_2 h_0^2}} \tag{5.61}$$

distribuye aproximadamente normal con media cero y varianza unitaria (Jackson & Mudholkar, 1979). De esta forma, es posible obtener un valor para un nivel de confianza α sobre el índice Q de la forma

$$Q_\alpha = \theta_1 \left(\frac{c_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + \frac{\theta_2 h_0(h_0-1)}{\theta_1^2} + 1 \right)^{1/h_0} \tag{5.62}$$

en donde c_α es un valor donde la distribución normal acumulada abarca un área de α. Los valores de Q que excedan el nivel Q_α, significa que no estan adecuadamente representados

por el modelo filtrado bajo componentes principales. Esto se traduce en que la variabilidad debido a efectos no explicados como ruido o fallas se hace muy fuerte.

El índice Q se aplica de la misma forma que la prueba de T^2. En general, la prueba de Q es mas certera que el test T^2 para determinar la correcta correlación entre mediciones de un proceso (Li, Yue, Valle-Cervantes, & Qin, 2000; Jackson, 2003; Xu & Wang, 2006).

5.4.5. Test de Scheffé

El test de Scheffé se usa para identificar que variables son una posible fuente de perturbaciones en un sistema con n variables medidas (Gómez, Unbehauen, Kortmann, & Peters, 1996; González, Orchard, Cerda, Casali, & Vallebuona, 2003). Se basa en el análisis de las mediciones del proceso en una ventana móvil de N muestras, en donde se analizan las desviaciones de las variables respecto a una referencia (residuos), por ejemplo, un conjunto de datos tomados en estado de operación normal. Considerando

$$r_i(t) = y_i(t) - y_i^*(t) \tag{5.63}$$

en donde $y_i(t)$ es la medición de la variable i e $y_i^*(t)$ es la referencia, se generan n índices f_{oi} de la forma

$$f_{oi} = \left(\frac{\bar{r}_i}{\sigma_i}\right)^2 \frac{N(N-n)}{(N-1)n} \tag{5.64}$$

en donde $\bar{r}_i$ es la media del valor r_i y σ_i su desviación estándar. En ausencia de perturbaciones (o fallas), se cumple con un nivel de confianza α que

$$f_{oi} \leq F_{n,N-n,\alpha} \tag{5.65}$$

A diferencia de los test T^2 y Q, este índice se calcula de forma separada para cada variable.

5.5. Detección y diagnóstico de fallas basados en otras técnicas

Existen otras técnicas para el diseño de sistemas de detección y diagnóstico de fallas, como por ejemplo los sistemas basados en análisis de señales. Entre los métodos más utilizados se pueden encontrar:

- Análisis espectral (transformadas de Fourier, espectros de frecuencia, DFT, FFT)
- Funciones de correlación
- Transformada Wavelet
- Filtros en tiempo y frecuencia (pasabajos, pasabanda, etc)
- Filtros de partículas

El principal uso de estas técnicas es sobre procesos que presenten características determinadas que puedan ser explotadas por estos métodos, por ejemplo la detección de vibraciones sobre máquinas rotatorias, taladros, etc., y no cuentan con un esquema estructurado de aplicación.

6. DESARROLLO DE UN SISTEMA DE DETECCIÓN Y DIAGNÓSTICO DE FALLAS PARA SENSORES DE UNA TURBINA DE GAS

En general, los procesos industriales complejos como las plantas de ciclo combinado cuentan con sistemas de monitorización que ponen al tanto al operador de cualquier anomalía que ocurre en la planta. En caso de falla, es usual que la planta se detenga para evitar daños a los componentes. Las fallas graves son poco usuales en la operación de la planta.

Uno de los tipos de falla más comunes es la falla de instrumentación. Esta se manifiesta como una medición errónea de la variable involucrada. En caso de sensores que estan propensos a una falla, por ejemplo las termocuplas que monitorean la temperatura de las bobinas del estator, se realiza una redundancia física para que en caso de ocurrir una falla no se pierda dicha medicion. Sin embargo, los operadores de todas formas reciben las alarmas debido a las mediciones fuera de lugar que entregan los sensores en falla. En caso de que el sensor no posea la redundancia física, la medición se pierde y no se sabe de forma inmediata si la falla es del proceso o del sensor. Debido a los cortos tiempos de reacción disponibles para tomar decisiones, siempre se opta por la prevención del daño mayor, muchas veces deteniendo la planta de forma innecesaria.

En este capítulo se presenta el desarrollo de un sistema de detección y diagnóstico de fallas para sensores de una turbina de gas, integrado en un módulo de mediciones tolerantes a fallas. El objetivo del sistema es monitorizar en tiempo real las mediciones del proceso, y eventualmente determinar la presencia de una falla en estas mediciones. Ante la presencia de una falla, el sistema realiza una estimación de la medición correcta y reemplaza la medición fallada por la corregida.

El sistema está construído en base a la combinación de las técnicas de detección y diagnóstico de fallas basado en modelos y en análisis estadísticos descritos en la sección anterior, combinando las ventajas de cada enfoque. El desarrollo se realiza en base a la información obtenida desde sensores de la turbina de gas de Central San Isidro, propiedad de ENDESA, y provista desde el Centro de Monitoreo y Diagnóstico (CMD).

6.1. Estructura del sistema de detección y diagnóstico

6.1.1. Análisis de los enfoques clásicos

Los sistemas de detección y diagnóstico de fallas basados en modelos tienen la ventaja de adaptarse a la dinámica del proceso, y en muchos casos operar bajo distintas condiciones. Sin embargo, luego de la generación de los residuos no se cuenta con estudios que permitan determinar la presencia de perturbaciones o fallas bajo cierto nivel de confianza en procesos no lineales. Por ejemplo, en (Chen & Patton, 1999) se revisan distintas técnicas basadas en los enfoques clásicos que conlcuyen con dificultad índices estadísticos relacionados con los errores de modelación y la presencia de fallas, pero todos ellos se basan sólo en modelos lineales ya sea tipo ARX o en espacio de estados. Lo mismo ocurre en (Simani et al., 2003), en que se omite el análisis bajo ruido y perturbaciones ocasionadas por la linealización de modelos de procesos. Por otro lado, en (Korbicz et al., 2004) si se incorporan técnicas no lineales basadas en inteligencia artificial, pero más bien enfocadas a la generación de los residuos en distintos casos.

En el caso de la detección de fallas en sensores, las estructuras clásicas generan un desafío en la clasificación de los residuos. Es relativamente simple detectar las fallas en sensores de salida: se compara la salida entregada por el modelo contra la medición real, bajo un esquema simple como el de la figura 5.5. Sin embargo, ante la presencia de una falla en un sensor de variables de entrada, la detección ya no es trivial ya que el residuo arrastra la falla en el modelo. En este caso, la detección de fallas de sensores de entrada se vuelve oscura. En el caso de un esquema de múltiples modelos como el de la figura 6.1, si bien se cuenta con distintos modelos para las salidas, el caso es el mismo. Al haber una falla en un sensor de entrada, ésta se propaga hacia todas las salidas a través de todos los modelos, dificultando la detección.

Por ejemplo, esta estructura se usa en (Simani, Fantuzzi, & Beghelli, 2000), y en (Simani, Fantuzzi, & Spina, 1998). Luego, para clasificar los residuos, se utilizan redes neuronales que permiten detectar las fallas. Si bien bajo este esquema es posible detectar la

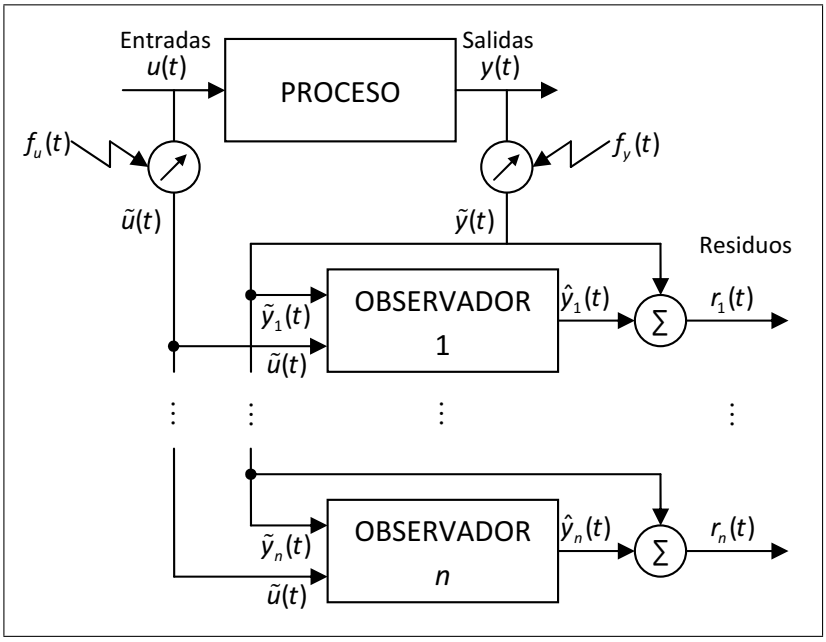

FIGURA 6.1. Esquema de generación de residuos basado en múltiples observadores.

falla (como se realiza en dichos desarrollos), no es posible reconstruír las señales dañadas (especialmente las señales de entrada).

Por otro lado, los desarrollos basados en análisis estadístico entregan los resultados de la detección de perturbaciones bajo un nivel de confianza determinado, pero requieren de fuertes supuestos sobre el proceso para realizar las conclusiones (por ejemplo, procesos estocásticos de media nula y varianza unitaria). Por tanto, el proceso en estudio queda muy restringido, en general, a operar en un régimen estable y que no presente cambios en el estado de operación abruptos. En las centrales de ciclo combinado esta situación es imposible de asegurar, ya que en el contexto en que están inmersos estos procesos los cambios abruptos son habituales (cambios en el despacho de carga, salida de otras unidades de la red fuera de operación, cambio del suministro de combustible entre petróleo diesel y gas, etc.), por lo que los algoritmos estadísticos interpretan como fallas los que en realidad son cambios en el estado de operación (ver gráficos de operación de la central en el capítulo 4). Otra desventaja es que estos algoritmos pueden detectar la falla, pero no determinar su procedencia, ni mucho menos reconstruir la señal.

6.1.2. Esquema propuesto de estimación de señales de falla

Para resolver estos problemas se propone primero una estructura que reemplaza la generación de residuos como tales por una estructura que estima directamente las señales de fallas ocurridas sobre los sensores, como se muestra en la figura 6.2.

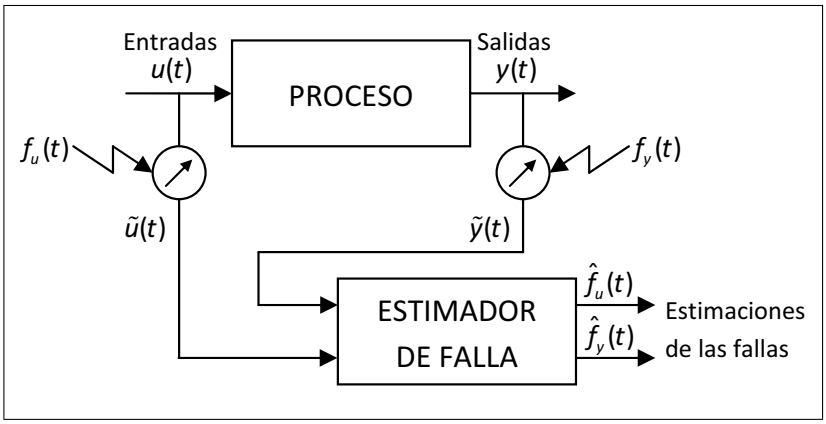

FIGURA 6.2. Esquema de estimación directa de las señales de falla.

Bajo este esquema, la detección de la falla se realiza de forma intuitiva cuando el valor estimado de la falla en sensores de entrada $\hat{f}_u$ o en sensores de salida $\hat{f}_y$ es distinto de cero. Además, la reconstrucción de la señal dañada se realiza simplemente sumando o restando la estimación de la falla $\hat{f}_u$ o $\hat{f}_y$ al sensor correspondiente. El estimador de falla debe ser capaz de "modelar" la ocurrencia de fallas en base a las entradas y salidas del sistema. Para esto, se construye usando las técnicas de modelación inteligente basadas en sistemas difusos de Takagi-Sugeno vistas en el capítulo 3.

6.1.3. Esquema propuesto de diagnóstico de las señales de falla

Como se comentó anteriormente, el análisis estadístico permite detectar perturbaciones o fallas en un proceso bajo suposiciones como media cero y varianza unitaria. Esto restringe su aplicación a procesos que mantengan un estricto régimen estable.

Sin embargo, bajo el esquema de estimación de señales de falla propuesto en la subsección anterior, en estado normal las señales entregadas por el estimador son cero o cercanas a cero, y pueden ser escaladas a través de la ecuación (5.37). Frente a cambios abruptos en

el proceso, las señales no deberían variar abruptamente por lo que los cambios significativos sólo están asociados a la presencia de fallas. Luego, es posible aplicar los enfoques basados en análisis estadístico descritos en él capítulo anterior. Para eliminar los efectos del error de modelación cometido al construír el estimador, se usa el análisis de componentes principales para eliminar los componentes menos significativos asociados al ruido y los errores de estimación.

Cabe destacar que el análisis de componentes principales requiere que las variables que se consideran en la descomposición tengan una distribución gaussiana. En este caso, dichas variables provienen del estimador de fallas. No existen en la literatura estudios relacionados con la distribución de probabilidad del error de modelación en sistemas difusos. Sin embargo, se asumirá que tienen una media cero, varianza constante, y sque en operación normal se comportan de una forma aproximadamente gaussiana.

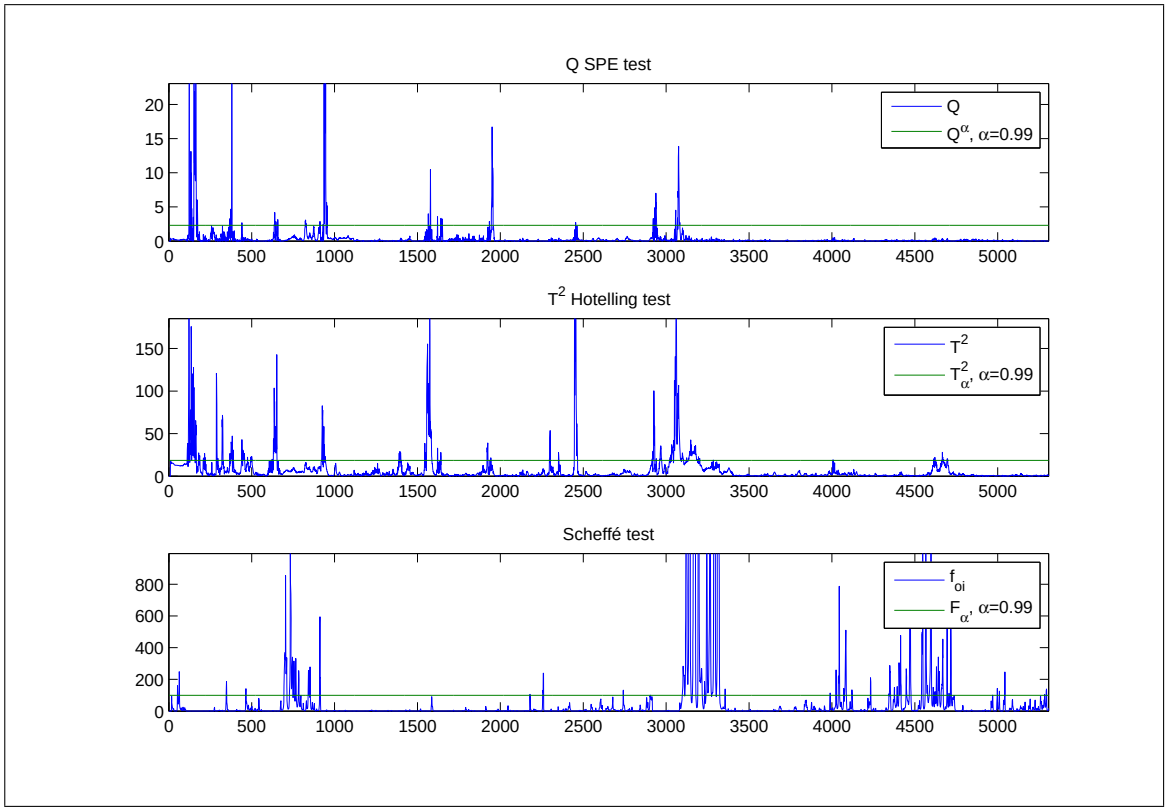

FIGURA 6.3. Tests Q, T^2 de Hotelling y de Scheffé para señal en operación normal.

Debido a que el test Q es más certero que el test T^2 y el test de Scheffé para clasificar anomalías (Li et al., 2000; Jackson, 2003; Xu & Wang, 2006), se utiliza dicho test para determinar la presencia de falla. Esta característica puede verse en la figura 6.3, que compara el desempeño de los tres indicadores considerando el residuo del estimador de falla ante operación normal. Se aprecia que el test Q genera menos falsas alarmas al producir cruces por el umbral de confianza en menos ocasiones y durante menos tiempo.

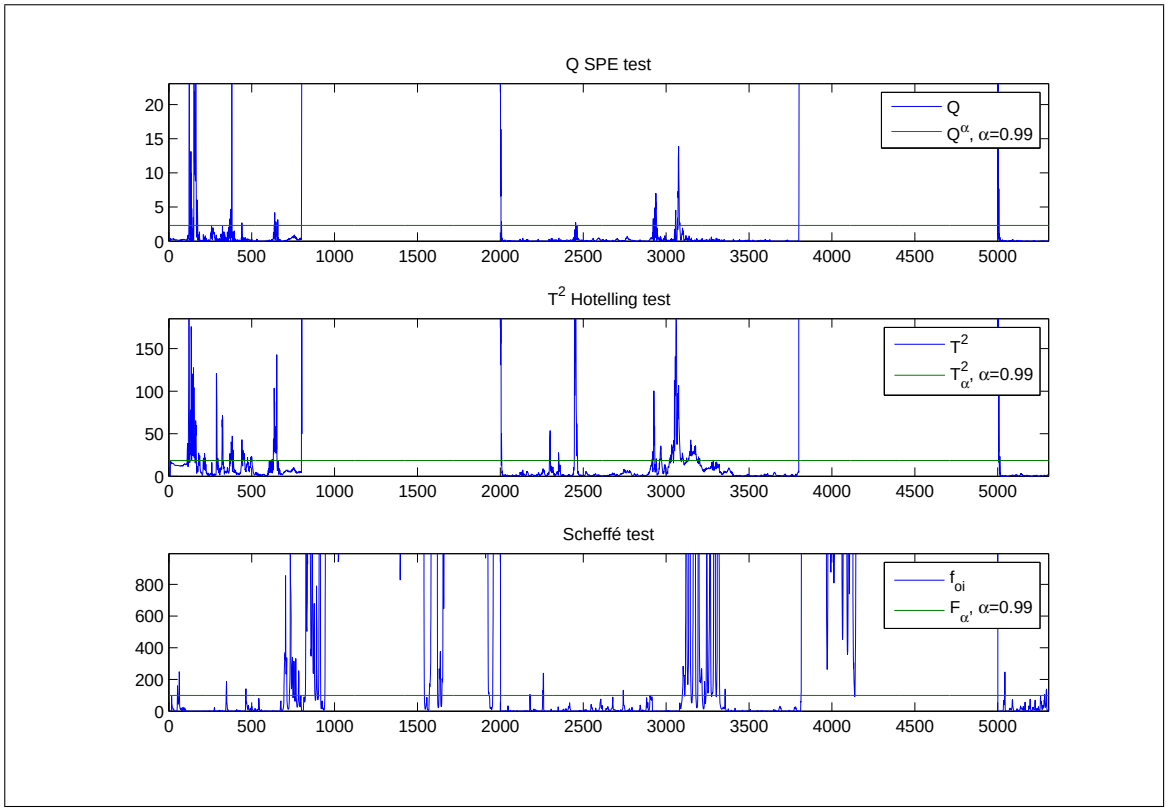

FIGURA 6.4. Tests Q, T^2 de Hotelling y de Scheffé para señal en falla.

La figura 6.4 muestra el resultado de los tres tests frente a una falla abrupta que se repite dos veces en el período de tiempo. En la figura 6.5 se muestran las alarmas generadas por sobrepasar los niveles de confianza. Puede verse que tanto el test T^2 de Hotelling como el test Q detectan correctamente la falla, pero el test Q tiene notoriamente menos falsas alarmas que los otros dos tests.

En estas pruebas se han establecido ventanas de 10 datos para el cálculo de los tests, y el filtraje de las alarmas.

Una vez detectada la presencia de falla en el sistema, se selecciona el sensor fallado a través de aquel sensor que aporta de mayor forma al índice Q, de la forma

$$q_i = |x_i - \hat{x}_i|^2 \tag{6.1}$$

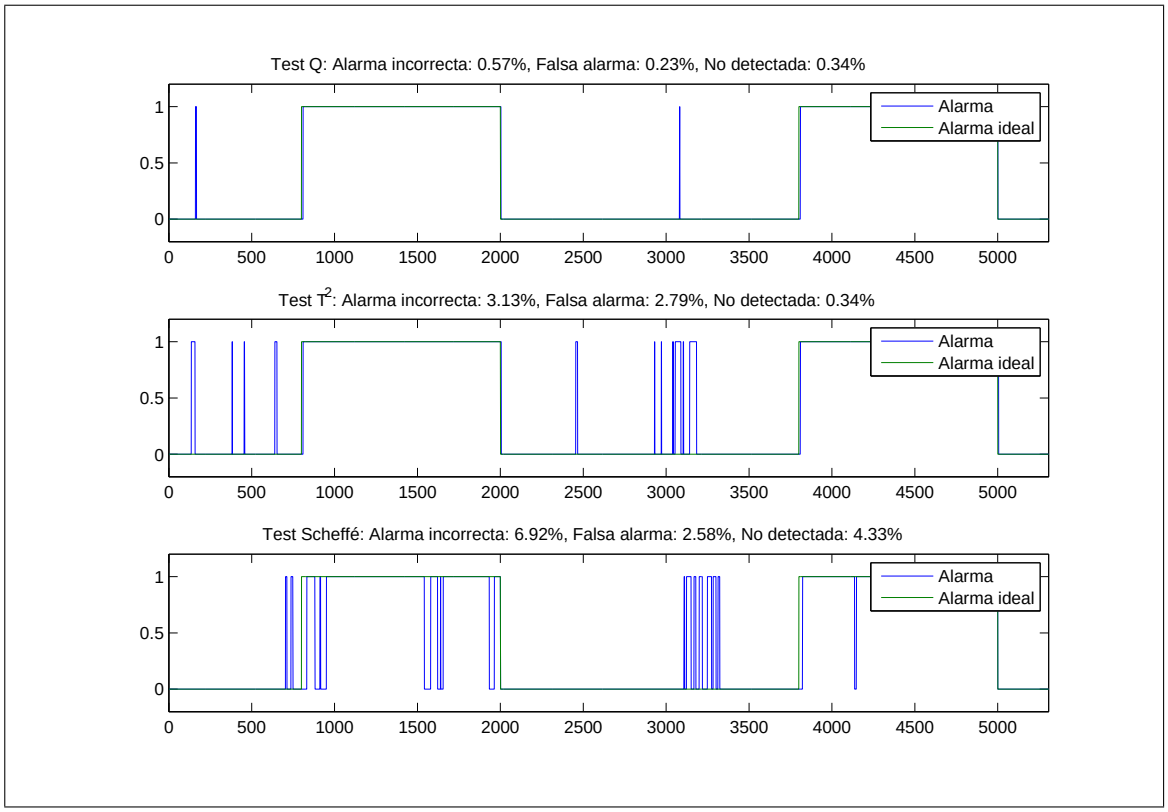

FIGURA 6.5. Señales de alarma filtradas para los 3 tests, en falla.

6.2. Reconstrucción de la señal dañada

La salida del estimador de falla es tratada con análisis de componentes principales y luego chequeada con el test Q. Una vez que se determina la presencia de falla global, se determina cuál es el sensor fallado.

Para realizar la reconstrucción de la señal deñada, se eliminan las estimaciones de falla de todos los sensores que no estén fallados. Es decir, si el sensor fallado corresponde a u_p, entonces se establece

$$\hat{f}^*_{u_i} = 0, \forall i \neq p, \quad \hat{f}^*_{u_p} = \hat{f}_{u_p}, \quad \hat{f}^*_y = 0 \tag{6.2}$$

De la misma forma, en caso de que el sensor fallado corresponda a y_q, se establece

$$\hat{f}_u = 0, \quad \hat{f}_{y_j} = 0, \forall j \neq q, \quad \hat{f}^*_{y_q} = \hat{f}_{y_q} \tag{6.3}$$

La reconstrucción de la señal se realiza de la siguiente forma:

$$\begin{bmatrix} \hat{u}(t) \\ \hat{y}(t) \end{bmatrix} = \begin{bmatrix} \tilde{u}(t) \\ \tilde{y}(t) \end{bmatrix} - \begin{bmatrix} \hat{f}^*_u(t) \\ \hat{f}^*_y(t) \end{bmatrix} \tag{6.4}$$

en donde $\hat{u}$ y $\hat{y}$ corresponden a las señales reparadas, $\tilde{u}$ y $\tilde{y}$ son las señales medidas y $\hat{f}_u$ y $\hat{f}_y$ son las estimaciones de las fallas sobre los sensores. La estructura completa del sistema se muestra en la figura 6.6.

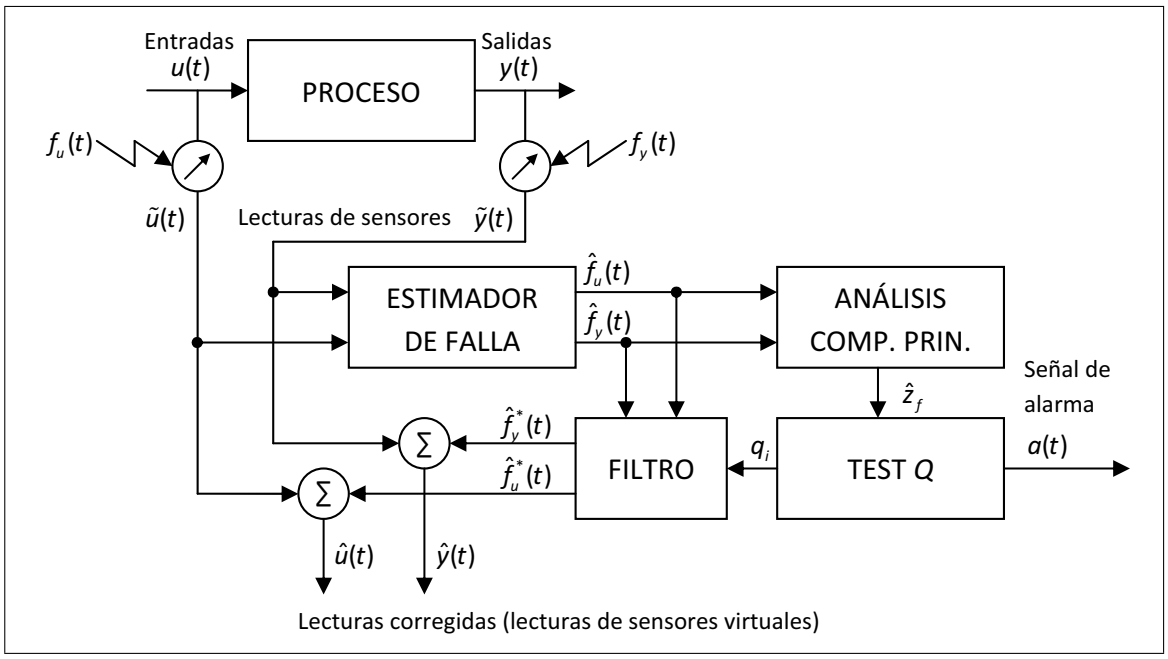

FIGURA 6.6. Estructura completa del sistema de mediciones tolerantes a fallas.

A la reconstrucción de las señales a partir de algoritmos de software se les denomina lecturas de *sensores virtuales* o *soft sensors*.

6.3. Entrenamiento del sistema

Como se ve en la figura 6.2, el sistema ajusta un modelo cuyas entradas corresponden a las variables de entrada y salida del proceso, que en este caso, son los subsistemas planteados en el capítulo 4. Las salidas del modelo son las señales de falla (supuestas como aditivas) sobre cada medición. Por ejemplo, para el sistema de turbina de gas el estimador de falla correspondería a la figura 6.7.

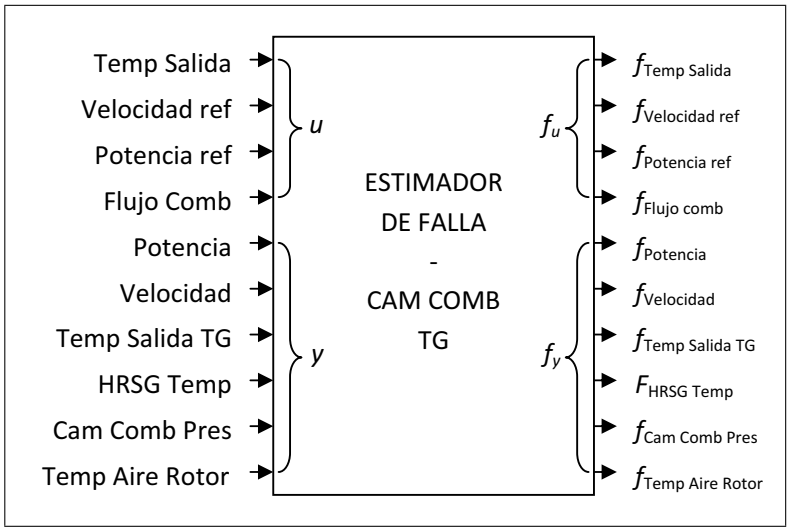

FIGURA 6.7. Estructura del estimador de fallas para el subsistema de cámara de combustión - turbina.

Tal cual como se hizo para el modelo del proceso en el capítulo 4, se ajusta un modelo para cada una de las salidas. Para identificar los parámetros, se utilizó el conjunto de datos de partida de planta disponible, usando los primeros 3000 datos como entrenamiento y el resto como validación. En este caso, no se cuenta con los datos suficientes ya que las salidas del estimador de fallas (las fallas en sí), son cero para los datos originales. Por esto se generó artificialmente un vector de entrenamiento que simula ser la falla para cada sensor. Este patrón de entrenamiento se muestra en la figura 6.8.

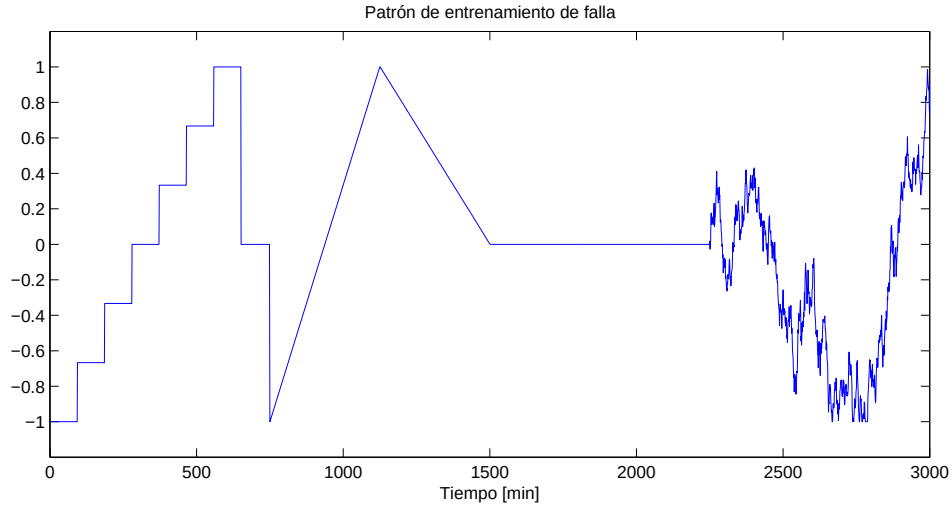

FIGURA 6.8. Patrón de entrenamiento para el estimador de fallas.

Como puede verse en la figura, el patrón de entrenamiento está dividido en 4 secciones. Primero, se incluyen cambios abruptos en forma de escalón. En segundo lugar, se incluye un cambio no abrupto en forma de rampa. En tercer lugar, se incluye un efecto nulo para asegurar una salida cercana a cero en ausencia de falla, y finalmente se incluye un movimiento aleatorio ruidoso. El patrón se muestra normalizado con amplitud unitaria, la que es reajustada según la magnitud de las mediciones para cada variable.

El entrenamiento se realiza sumando este patrón (debidamente escalado) al sensor correspondiente, y planteándolo como señal de falla de salida. Debido a como se ha planteado el comportamiento ideal del estimador, ante la presencia de una falla en un sensor, ésta se debe ver reflejada en la estimación de la falla para ese sensor, pero no debe afectar la estimación de las otras señales de falla (cuyos modelos también reciben la señal fallada como entrada). Por esto se realizan 3 iteraciones de entrenamiento a los datos, con las siguientes características:

(i) Se realiza el entrenamiento considerando como entradas las mediciones de las variables, sumando al sensor correspondiente el patrón de entrenamiento, y como salida también el patrón de entrenamiento.

(ii) Se realiza el entrenamiento considerando como entradas las mediciones de las variables, sumando al sensor correspondiente el patrón de entrenamiento y a los demás sensores fallas aleatorias, y como salida también el patrón de entrenamiento.

(iii) Se realiza el entrenamiento considerando como entradas las mediciones de las variables, dejando al sensor correspondiente igual y sumando a los demás sensores fallas aleatorias, y como salida un vector nulo.

La señal de falla aleatoria se genera a través de un proceso estocástico de la siguiente forma:

$$x(0) = 0 \tag{6.5}$$

$$x(t) = x(t-1) + z \tag{6.6}$$

en donde z sigue una distribución normal con media cero y varianza unitaria. Un ejemplo de este proceso se muestra en la figura 6.9.

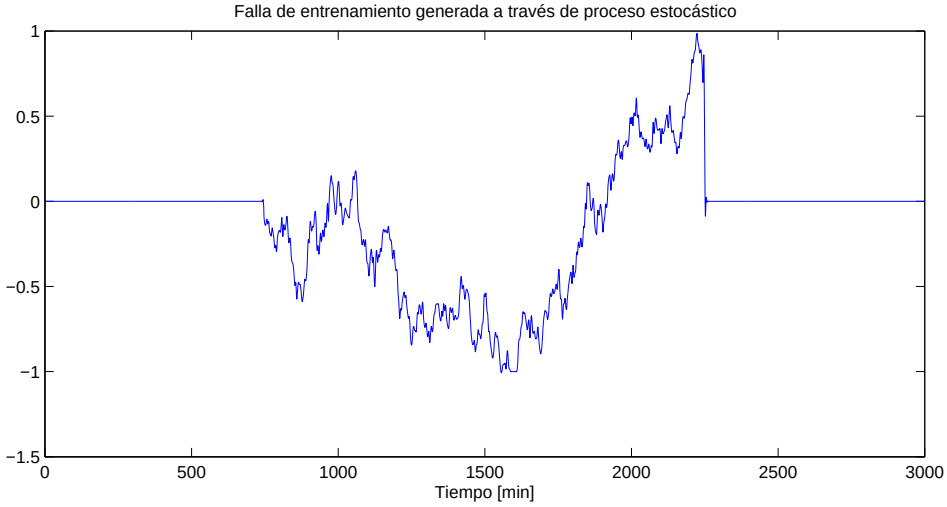

FIGURA 6.9. Ejemplo de falla generada a través del proceso estocástico planteado.

Para seleccionar el mejor modelo, se utiliza la misma metodología empleada para la modelación del proceso: se limitan los parámetros a $50n$ donde n es el número de entradas al modelo, y se usan de 2 a 15 reglas. Además, para mejorar el ajuste obtenido se entrena

un segundo modelo difuso de Takagi-Sugeno cuya entrada es la salida del estimador de falla.

6.4. Resultados

A continuación se muestran los resultados para el proceso de detección y diagnóstico de fallas, y reconstrucción de señales dañadas para el subsistema de la cámara de compresión - turbina de gas. La estructura de estimación de falla se muestra en la figura 6.7.

Para simplificar los resultados, no se han estimado modelos ni para la variable de referencia de velocidad ni para la referencia de potencia (se asume que no ocurrirán fallas en estas mediciones).

La evaluación de resultados se realizará de acuerdo a los siguientes cuatro indicadores de desempeño:

- El tiempo de falsas alarmas generadas: Se evaluará como el porcentaje de tiempo en que el sistema entrega una falsa alarma, respecto al tiempo total de simulación.
- Las fallas no detectadas: Se evaluará como el porcentaje de tiempo en que el sistema no detecta la presencia de una falla existente, respecto al tiempo total de simulación.
- El tiempo de detección: Medido a través del número de pasos de muestreo necesarios, una vez ocurrida la falla, para que el sistema logre detectar su presencia.
- La calidad de la reconstrucción: Medido a través del índice de ajuste FIT definido en (4.1) de la señal reconstruida respecto a la señal original de referencia sin falla.

Es necesario clarificar que usualmente el indicador de falsas alarmas del sistema está relacionado con una estimación estadística en base a la distribución de probabilidad de dicho error. Debido a la escasez de datos y a la imposibilidad de la determinación analítica de dicha distribución, se ha elegido el índice alternativo descrito anteriormente.

6.4.1. Caso 1: Sensores sin falla

En este caso, no se ha aplicado ninguna falla sobre las mediciones originales. La figura 6.10 muestra la estimación de las señales de falla realizada por el estimador.

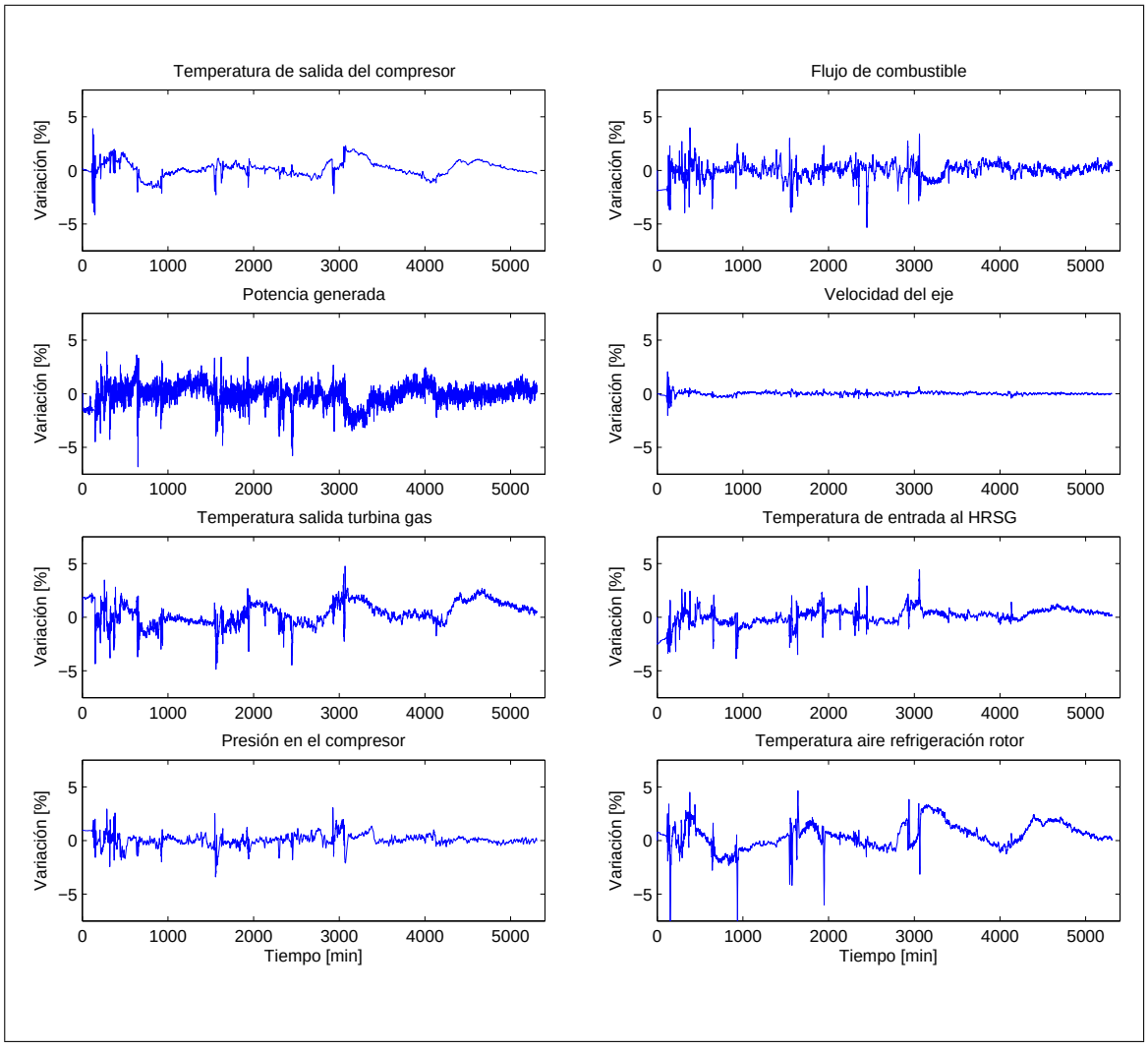

FIGURA 6.10. Estimación de las señales de falla.

En esta figura, se muestran los valores estimados de las fallas en forma de porcentaje respecto al valor de la variable a la que corresponde cada medición. Como puede verse, el error de modelación en este caso causa ruido en todas las estimaciones con una magnitud general de un 5%.

La figura 6.11 muestra la salida mejorada del estimador, a través de un modelo difuso en serie.

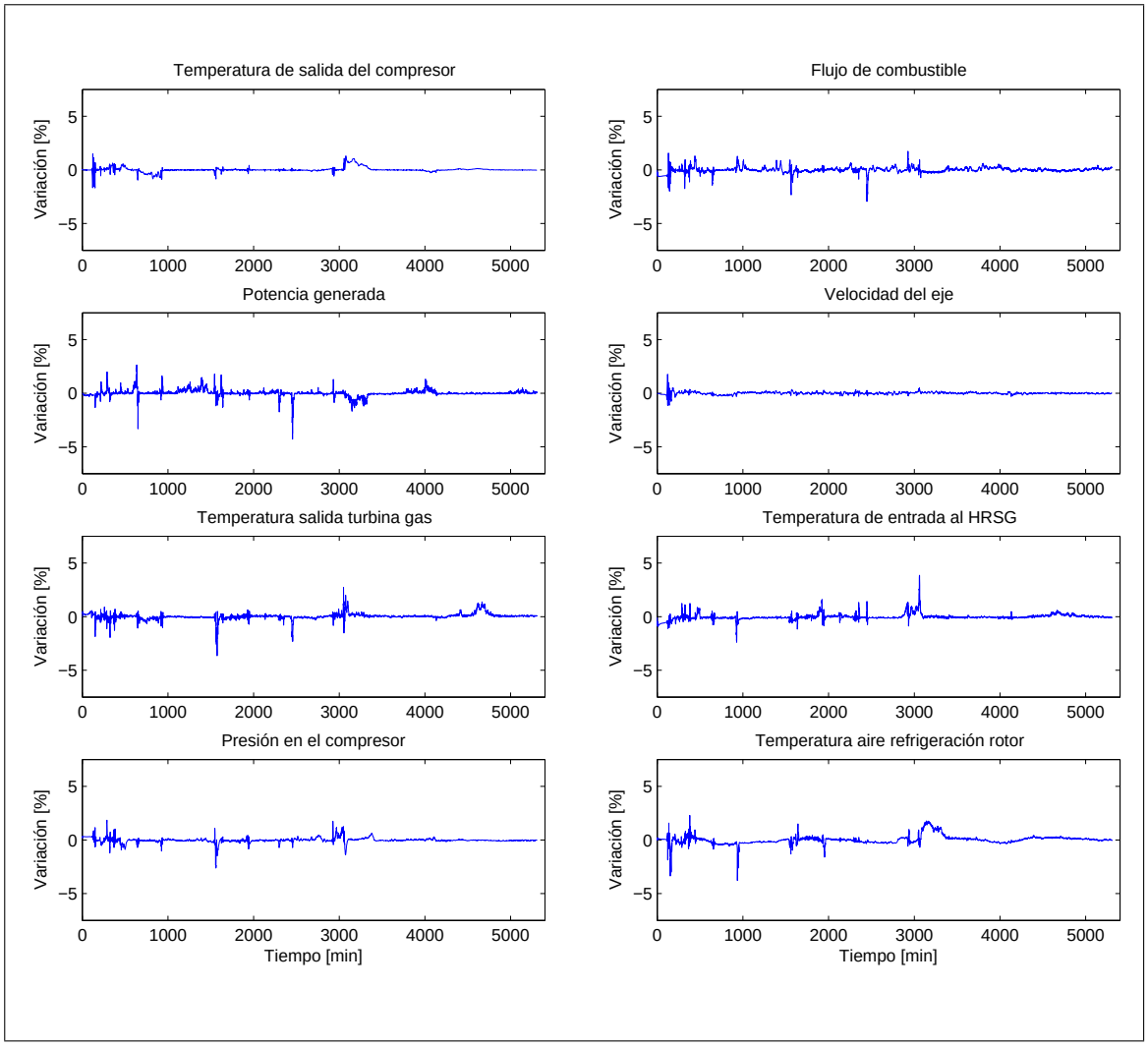

FIGURA 6.11. Estimación mejorada de las señales de falla.

Puede verse que la estimación de las fallas ahora es mucho mas cercana al valor nulo. La mejoría a través del modelo difuso de Takagi-Sugeno es bastante efectiva en este caso.

Ahora, a estas señales se les aplica la descomposición en componentes principales. Se seleccionan 7 de los 8 componentes principales, que para este caso (calculados en base al largo total de los datos) explican un 94.69% de la varianza de las señales de falla. Se aplica

el test Q para detectar las fallas, usando un nivel de confianza de $\alpha = 99\%$, dando como resultado el gráfico mostrado en la figura 6.12.

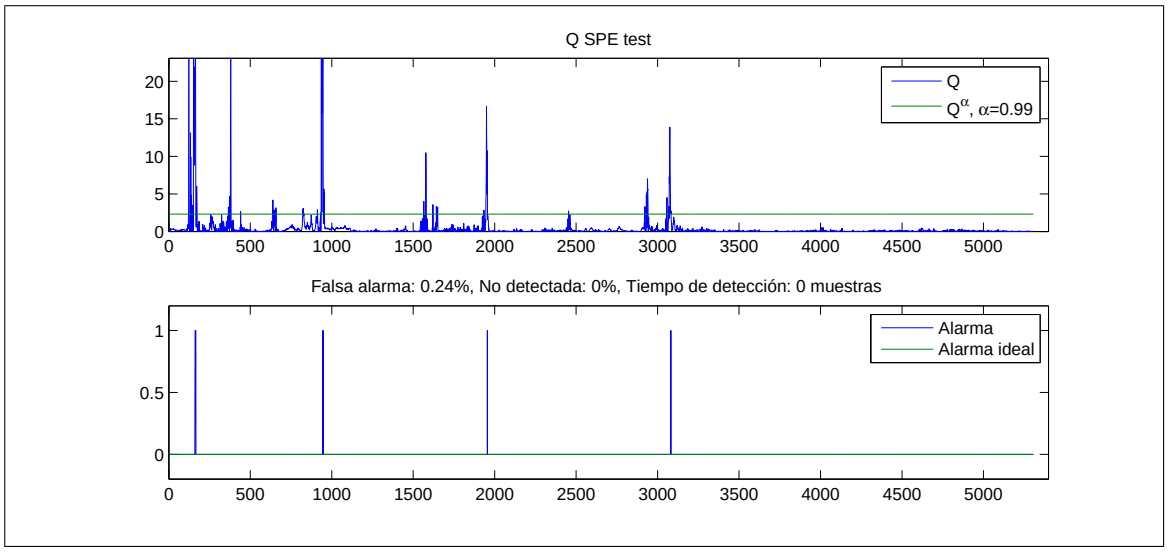

FIGURA 6.12. Aplicación del test Q y detección de la falla.

En este caso, la señal de alarma se ha filtrado como el promedio móvil de 3 instantes sucesivos. Como puede verse, aunque existen falsas alarmas (en los datos de validación), su presencia en el total de tiempo monitoreado es apenas de un 0.3%, esperable para el nivel de confianza planteado.

Con esta información, se filtran las estimaciones mejoradas de las fallas que se muestran en la figura 6.11, entregando las señales de corrección de sensores que se muestran en la figura 6.13.

Como puede apreciarse, la corrección en este caso en prácticamente nula, concordando con un estado de operación sin falla.

Para las futuras aplicaciones del test Q, se han considerado las matrices $\hat{C}$ y $\tilde{C}$ de la descomposición en componentes principales obtenidas en este caso, debido a que se pretenden detectar variaciones respecto al comportamiento de operación normal, y esta descomposición guarda las características del error de modelación del estimador bajo dichas condiciones.

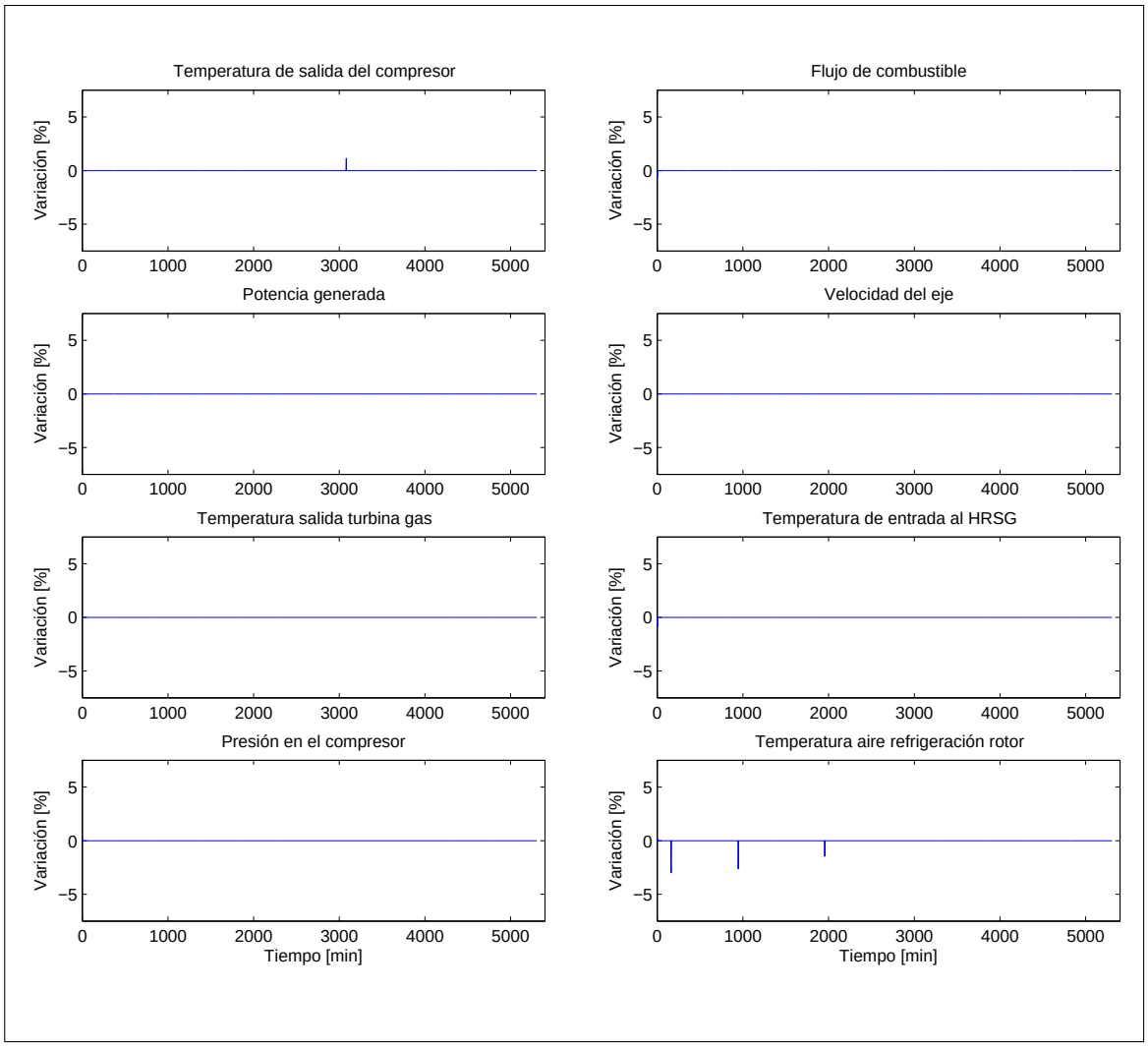

FIGURA 6.13. Señales de corrección de variables medidas.

6.4.2. Caso 2: Sensores bajo falla abrupta

En este caso, se genera artificialmente una falla abrupta en forma de escalón la que se aplica al sensor de temperatura de salida del compresor. La falla generada se muestra junto a las estimaciones generadas.

La figura 6.14 muestra las señales de falla estimadas. Como puede verse, a pesar de existir bastante ruido y cierta incidencia de la falla en las otras estimaciones, en el sensor fallado la estimación es bastante certera.

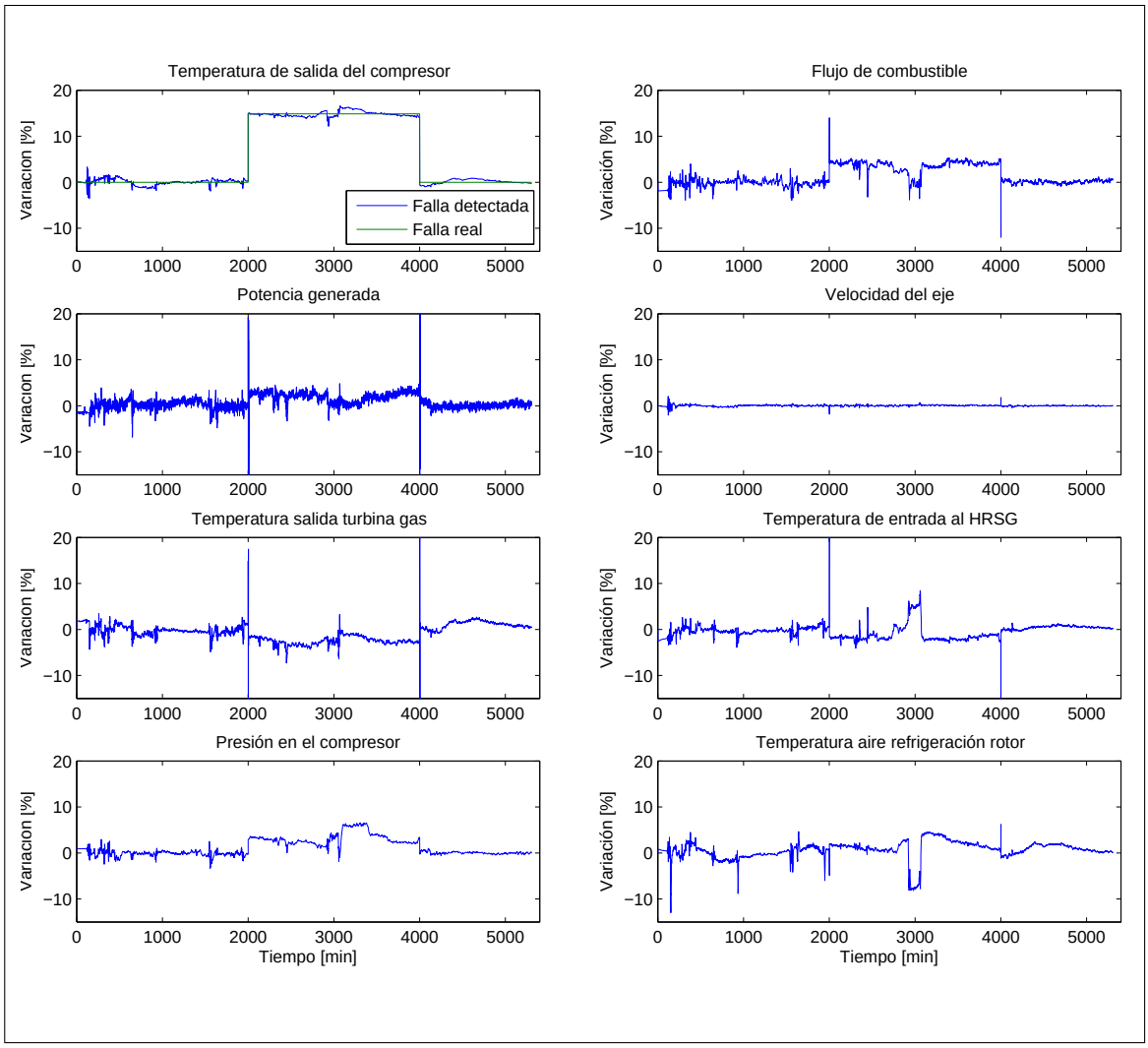

FIGURA 6.14. Estimación de las señales de falla para falla abrupta.

La figura 6.15 muestra la mejora en las estimaciones de las fallas generadas por el modelo en serie. Se aprecia una mejoría notable con relación a la reducción del ruido de las estimaciones de otras fallas, y en el caso del sensor que sufre la falla abrupta, se mejora la estimación.

También se aprecia en esta figura el test Q realizado a los datos bajo la descomposición en componentes principales. Como puede verse, nuevamente existen algunas falsas alarmas pero que no superan el 0.3%. El test determina claramente la presencia de la falla.

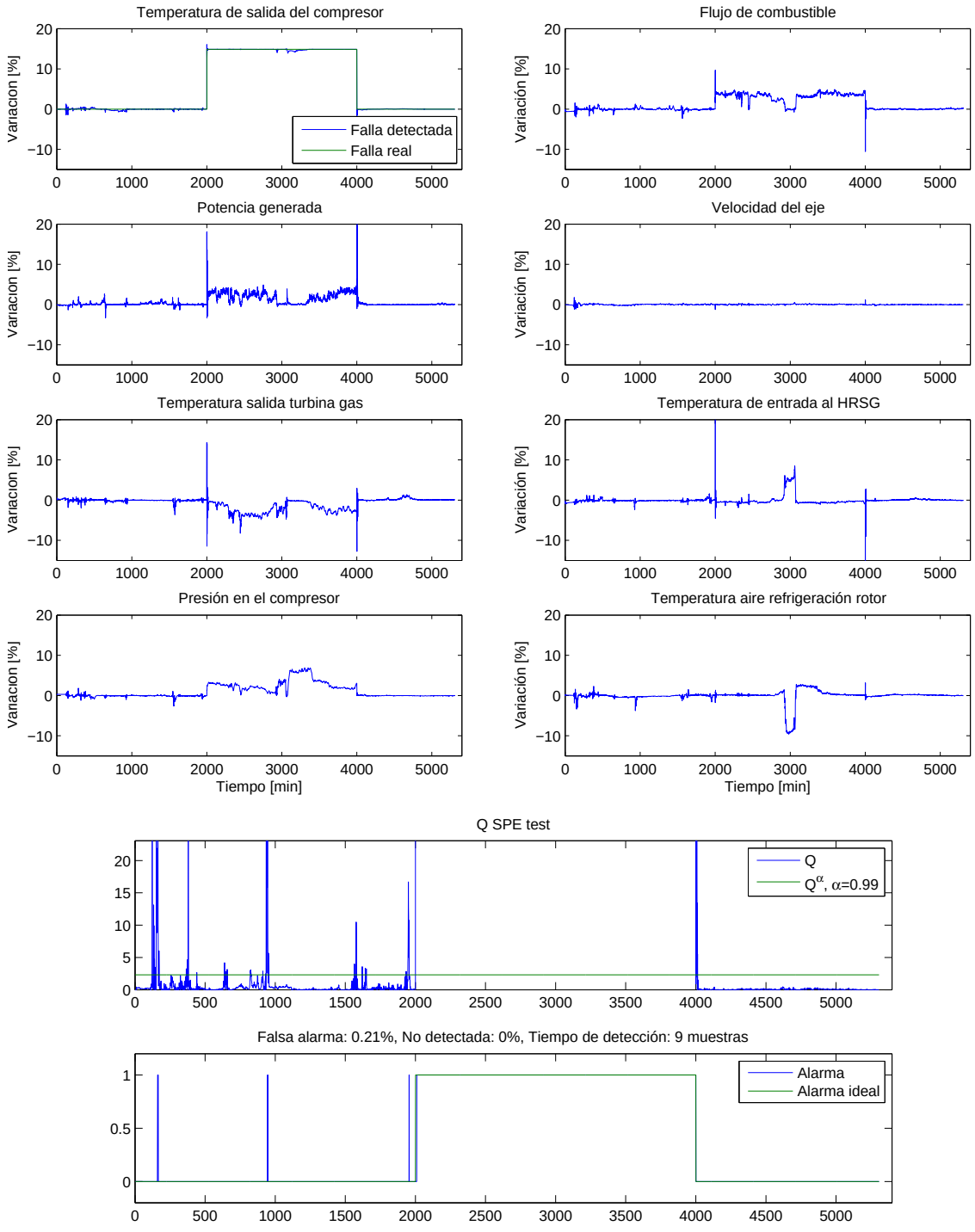

FIGURA 6.15. Estimación mejorada de las señales de falla para falla abrupta y test Q.

Una vez obtenidos los datos de la detección por parte del test Q, se realiza el filtraje de las estimaciones de falla generadas. Las señales de corrección del error se muestran en la figura 6.16.

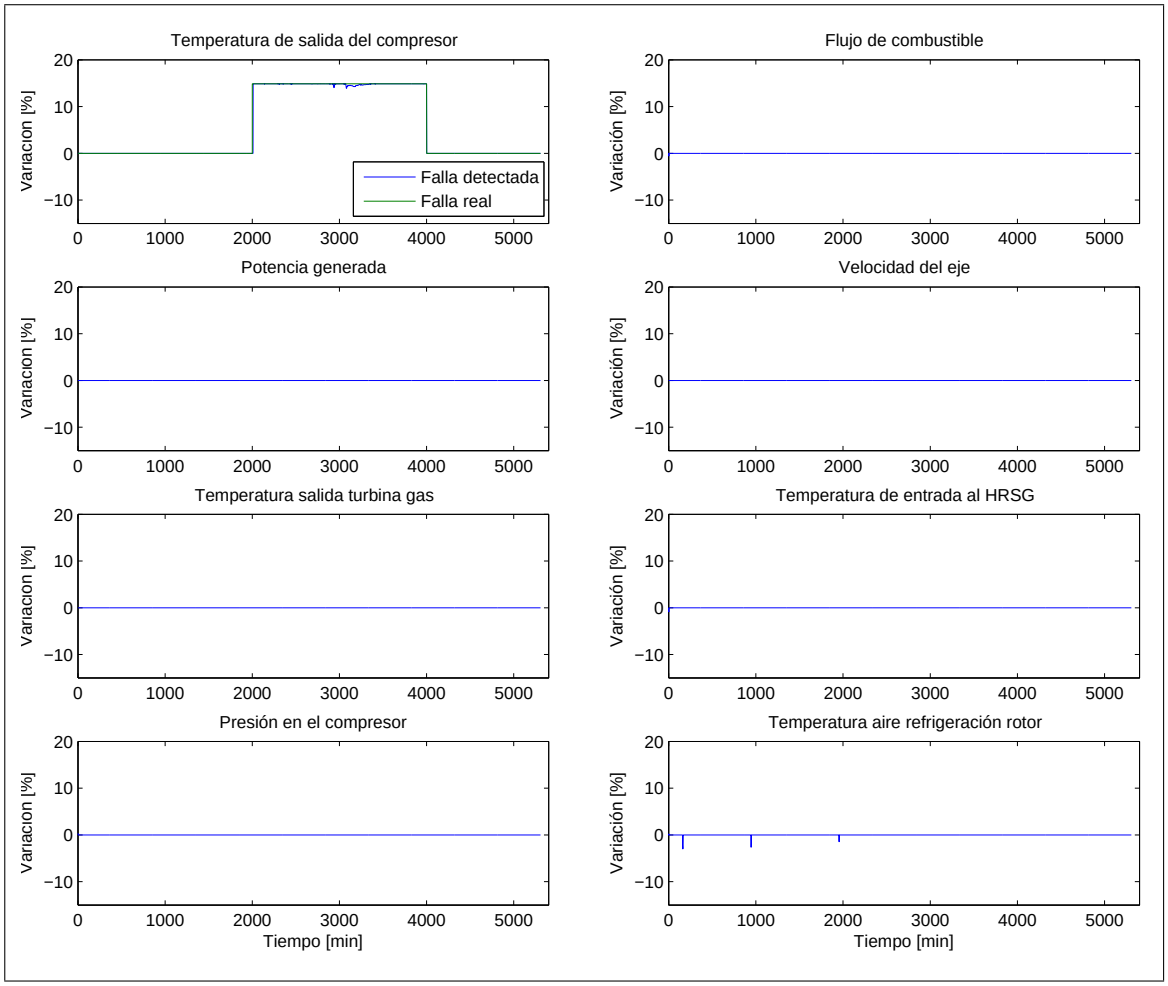

FIGURA 6.16. Señales de corrección de variables medidas, para falla abrupta.

Puede apreciarse claramente en esta figura la eficacia de la detección y selección a través del test Q y la descomposición en componentes principales. Todas las estimaciones de las fallas en los sensores en operación normal fueron eliminadas y definidas en cero, mientras que mantiene la estimación mejorada para la falla ocurrida en el sensor de temperatura de salida del compresor.

Las señales generadas por los sensores virtuales (señales corregidas), junto con la señal original y la medición con falla se muestran en la figura 6.17

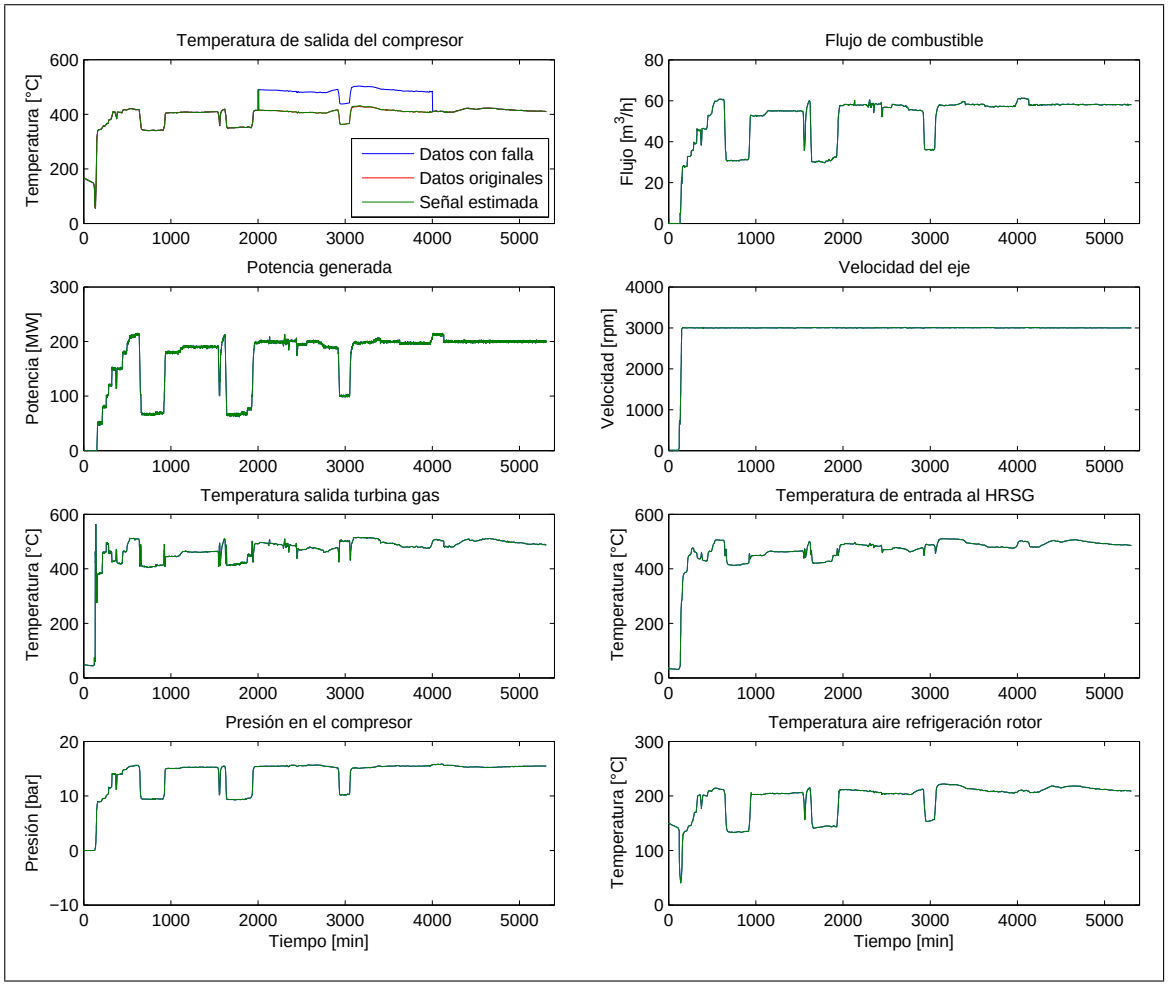

FIGURA 6.17. Señales corregidas, para falla abrupta.

Es claro en esta figura como la señal reparada se ajusta de manera prácticamente perfecta a la señal original antes de haber introducido la falla. En cuanto a los otros sensores, puede verse como se mantienen inalterados acorde con las correcciones mostradas en la figura 6.16.

6.4.3. Caso 3: Sensores bajo falla no abrupta y aleatoria

En este caso se simula artificialmente una falla de carácter no abrupto y con comportamiento aleatorio, sobre la variable de potencia generada. En la figura 6.18 se aprecia la estimación de las fallas.

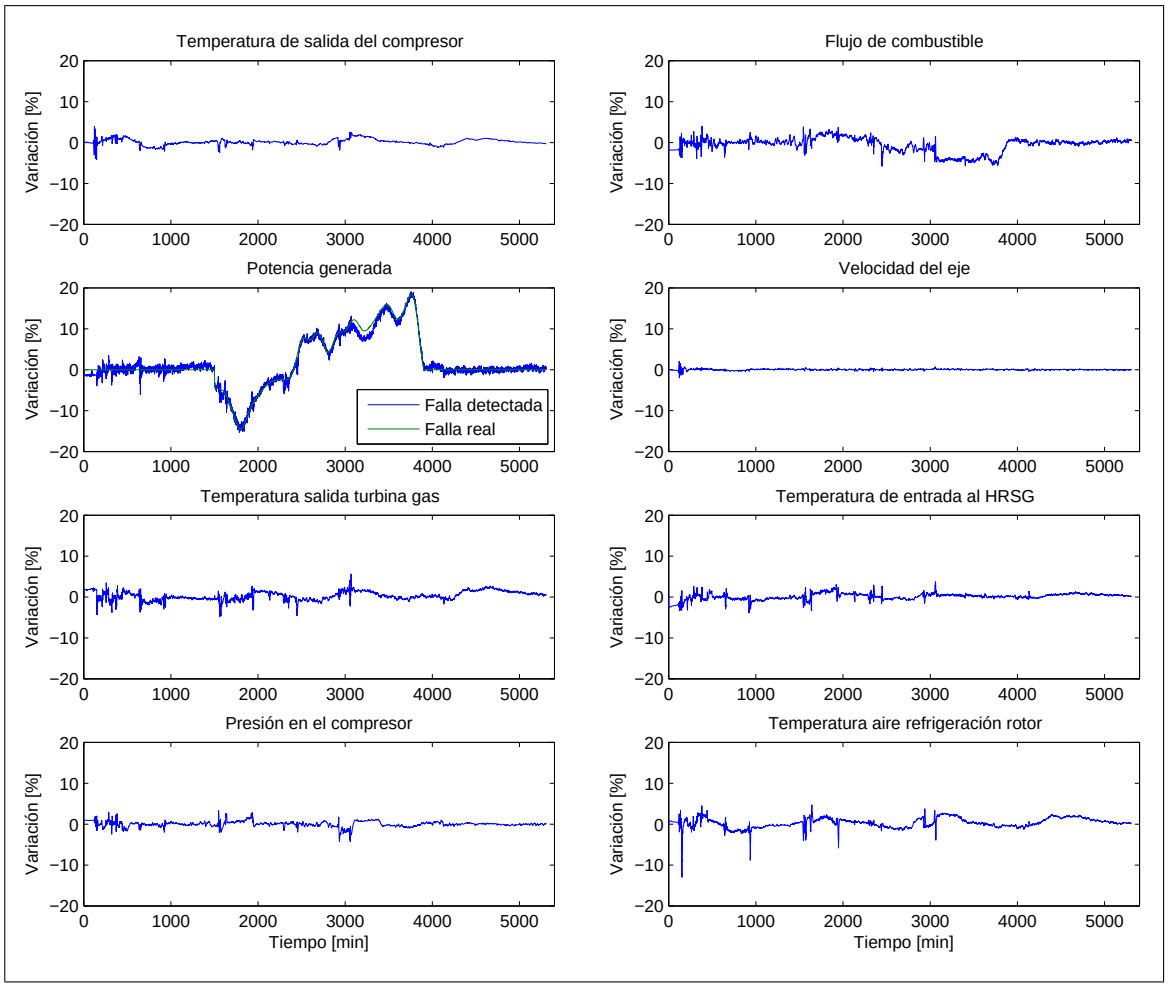

FIGURA 6.18. Estimación de las señales de falla para falla no abrupta.

En este caso puede verse con claridad como el estimador de la falla de la variable involucrada (potencia generada) actúa con mucha mayor claridad que los demás estimadores. En ellos permanece el ruído de modelación y los efectos de la falla del sensor.

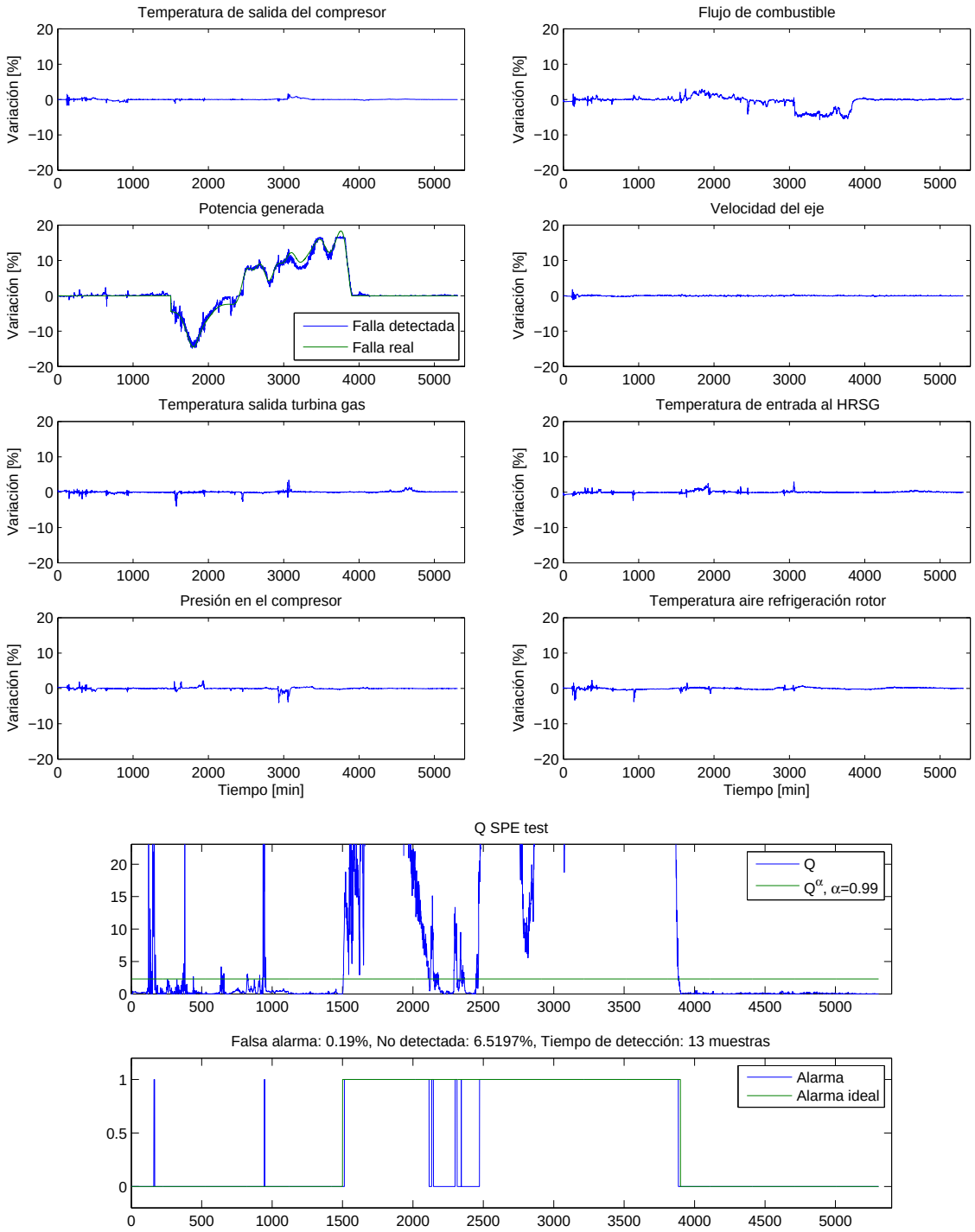

FIGURA 6.19. Estimación mejorada de las señales de falla para falla no abrupta y test Q.

La estimación mejorada se muestra en la figura 6.19, junto con el test Q para este caso. Puede apreciarse que en este caso se tienen más problemas para la detección, habiendo intervalos en que no se detecta la falla correctamente. Esto tiene que ver con el hecho de que la señal de falla realiza un cruce por cero, luego en los valores muy cercanos a ese punto de cruce se producen los problemas. La falla no detectada asciende a un total de un 6.7% en este caso. Las señales corregidas se muestran en la figura 6.20.

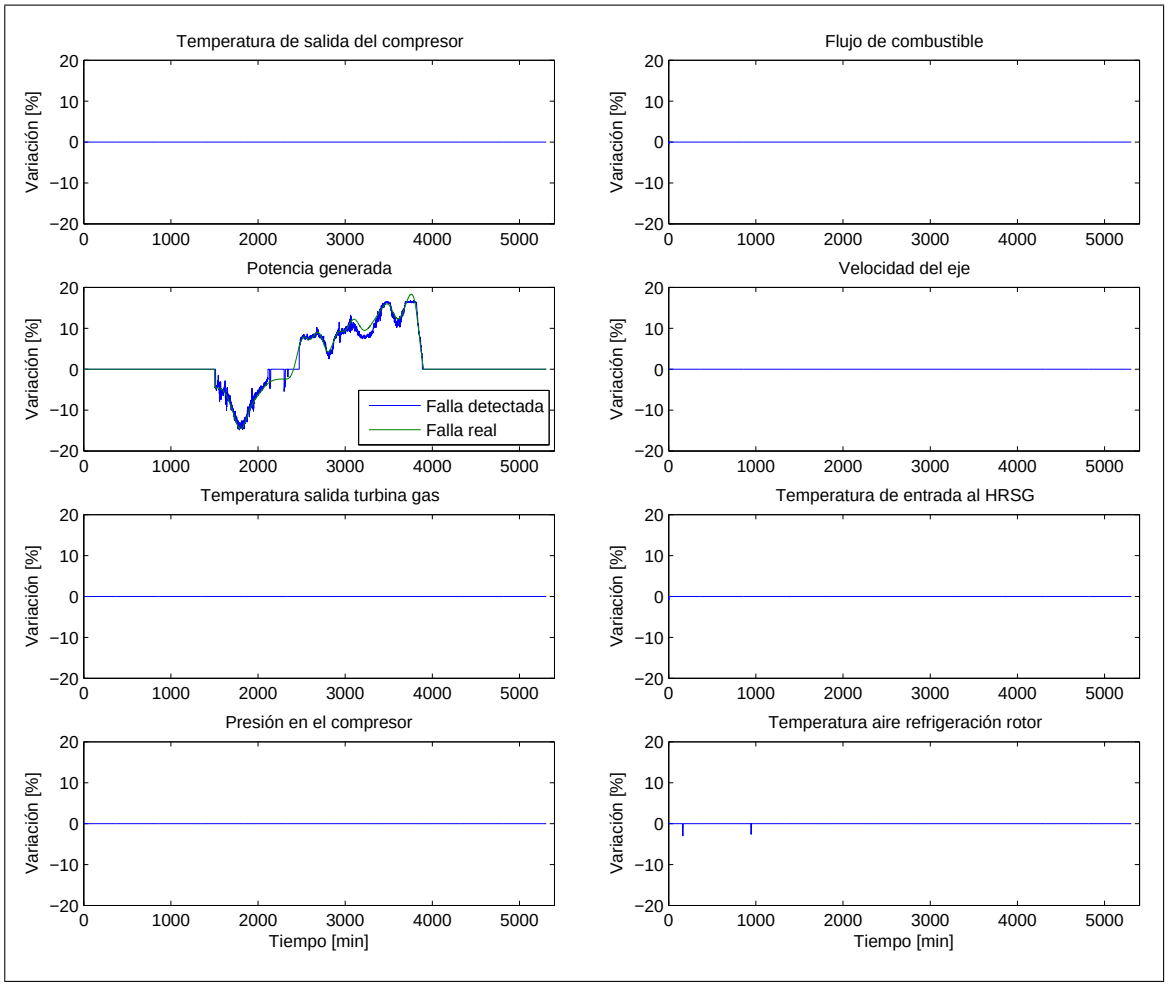

FIGURA 6.20. Señales de corrección de variables medidas, para falla no abrupta.

Puede verse que a pesar de los problemas de detección, las señales de corrección son bastante satisfactorias. El problema en el cruce por cero queda de manifiesto en el gráfico

de la señal de corrección de la variable con falla. Las señales corregidas en este caso se muestran en la figura 6.21.

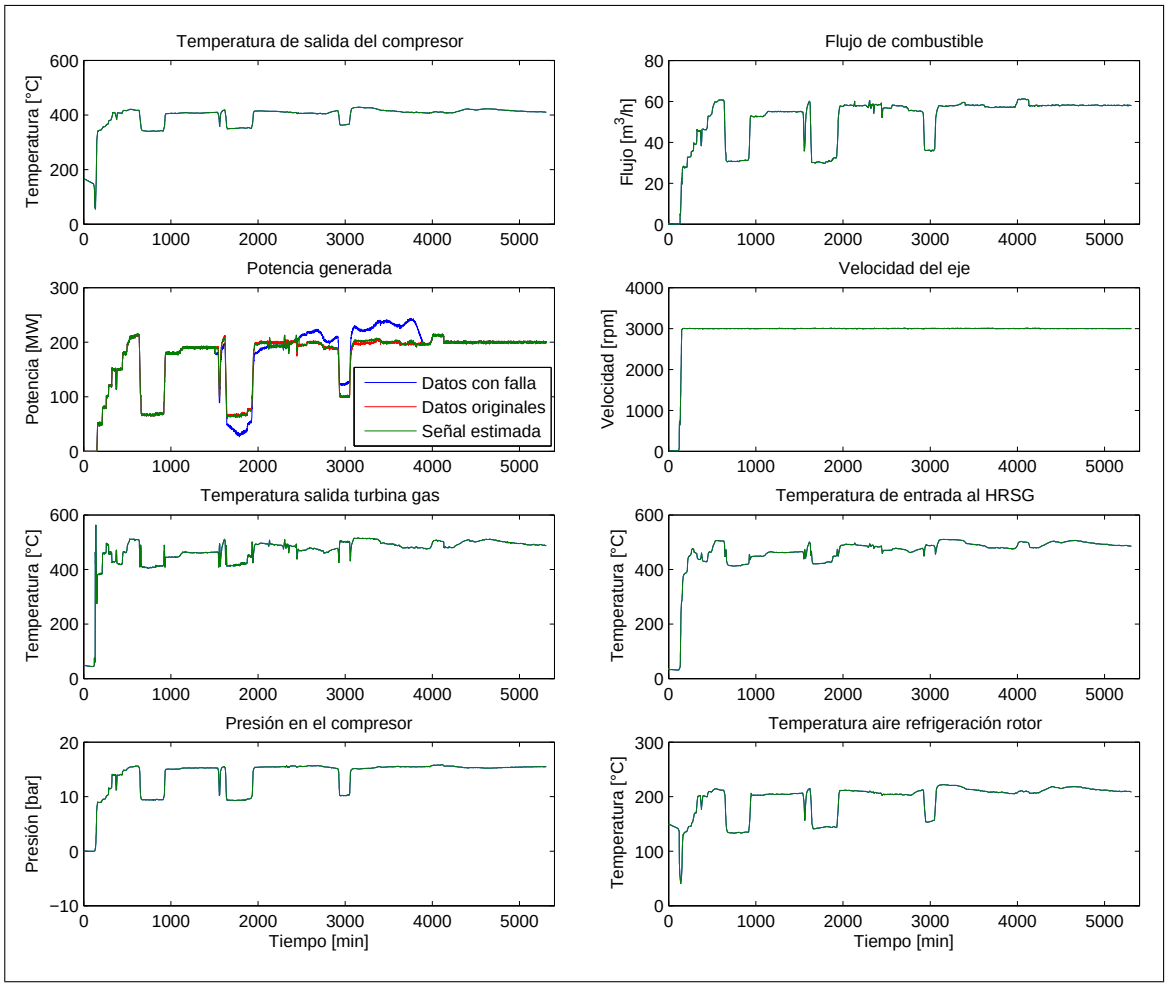

FIGURA 6.21. Señales corregidas, para falla no abrupta.

Se ve que a pesar del error cometido en la detección, la señal estimada sigue siendo prácticamente idéntica a los datos originales.

6.4.4. Caso 4: Sensores bajo falla real

En este caso se aplica el sistema diseñado a un conjunto de datos correspondiente a una falla real, ocurrida sobre el sensor de temperatura de aire de refrigeración. La falla ocurre el día 15 de marzo de 2008, a las 18:50, producto de la caída de la termocupla asociada a

esta medición. Al ser una falla real, no se cuenta con una referencia con la cual comparar los resultados. La figura 6.22 muestra las estimaciones de las señales de falla en este caso.

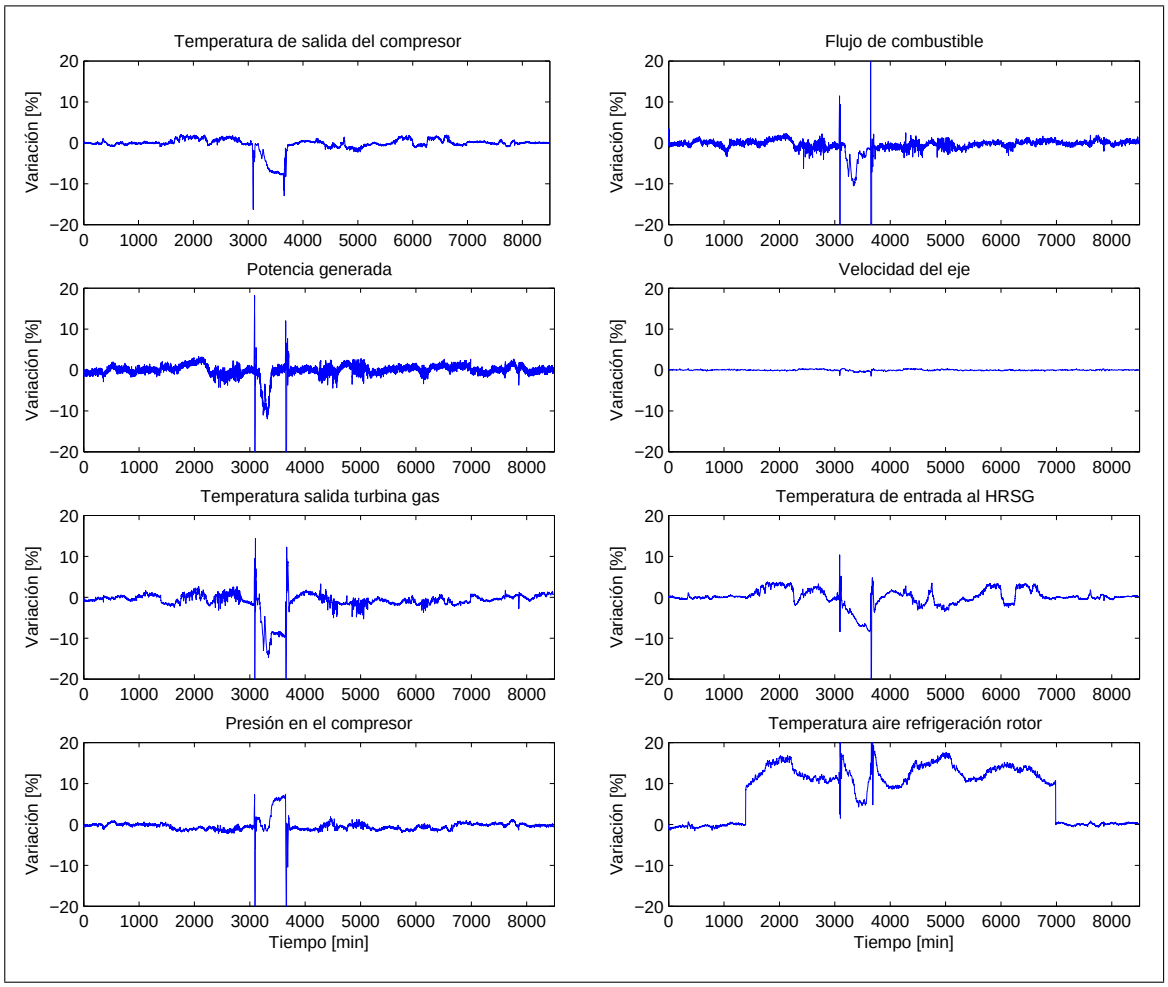

FIGURA 6.22. Estimación de las señales de falla para falla real.

Puede verse que en general en este caso el modelo funciona bastante bien ya que las estimaciones de los sensores que no tienen falla es siempre cercana a cero, salvo por una perturbación entre los 3000 y 4000 minutos. Esta perturbación pertenece al proceso y se debe al cambio entre gas y petróleo diesel en la alimentación de la planta. Las estimaciones mejoradas y el test Q se muestran en la figura 6.23.

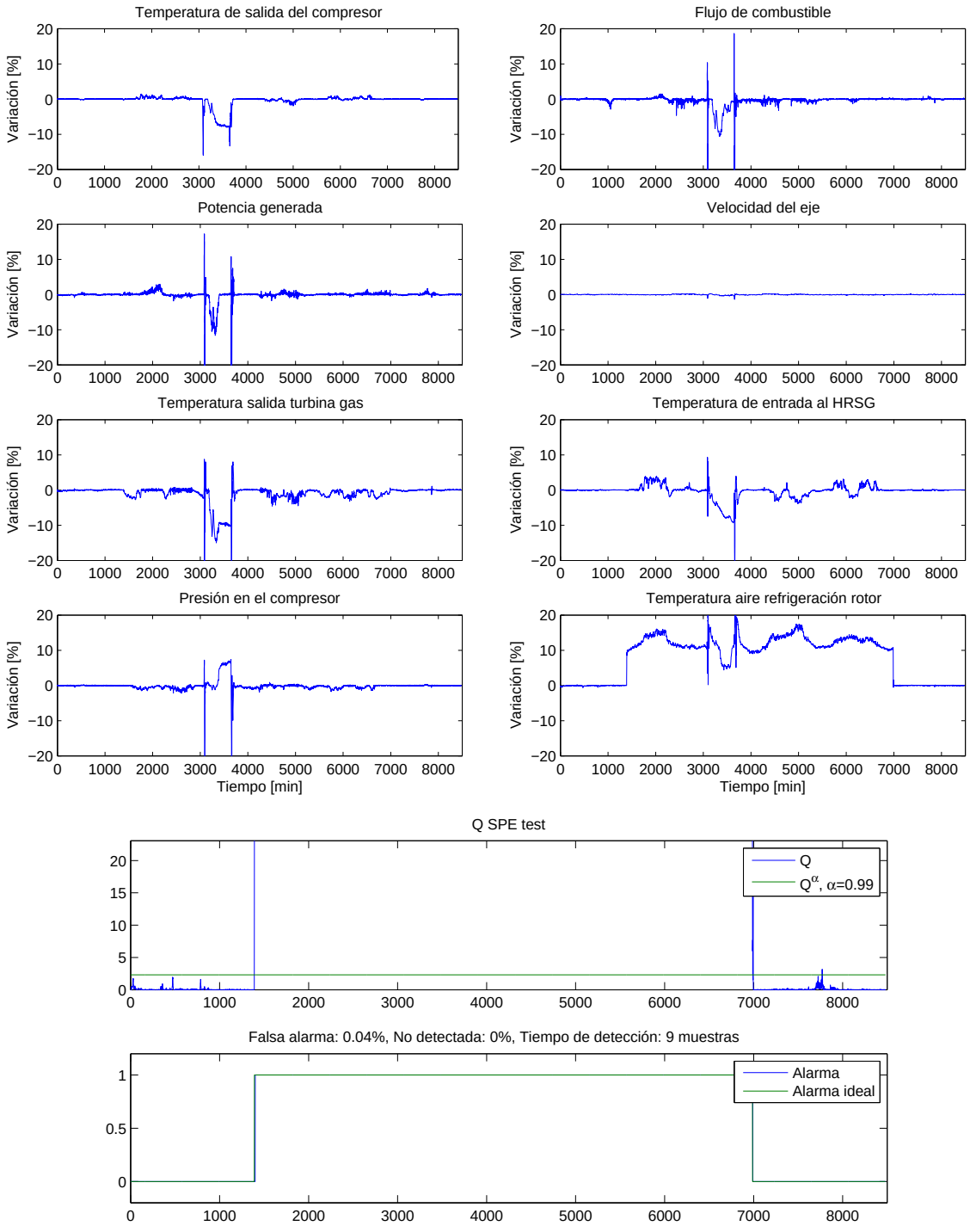

FIGURA 6.23. Estimación mejorada de las señales de falla para falla real y test Q.

En esta figura se aprecia como la detección en este caso ocurre de forma impecable. La alarma ideal del gráfico del test Q se genera artificialmente a través de la inspección manual de los datos en presencia de falla. El filtraje correspondiente a las señales de corrección luego del test se muestra en la figura 6.24.

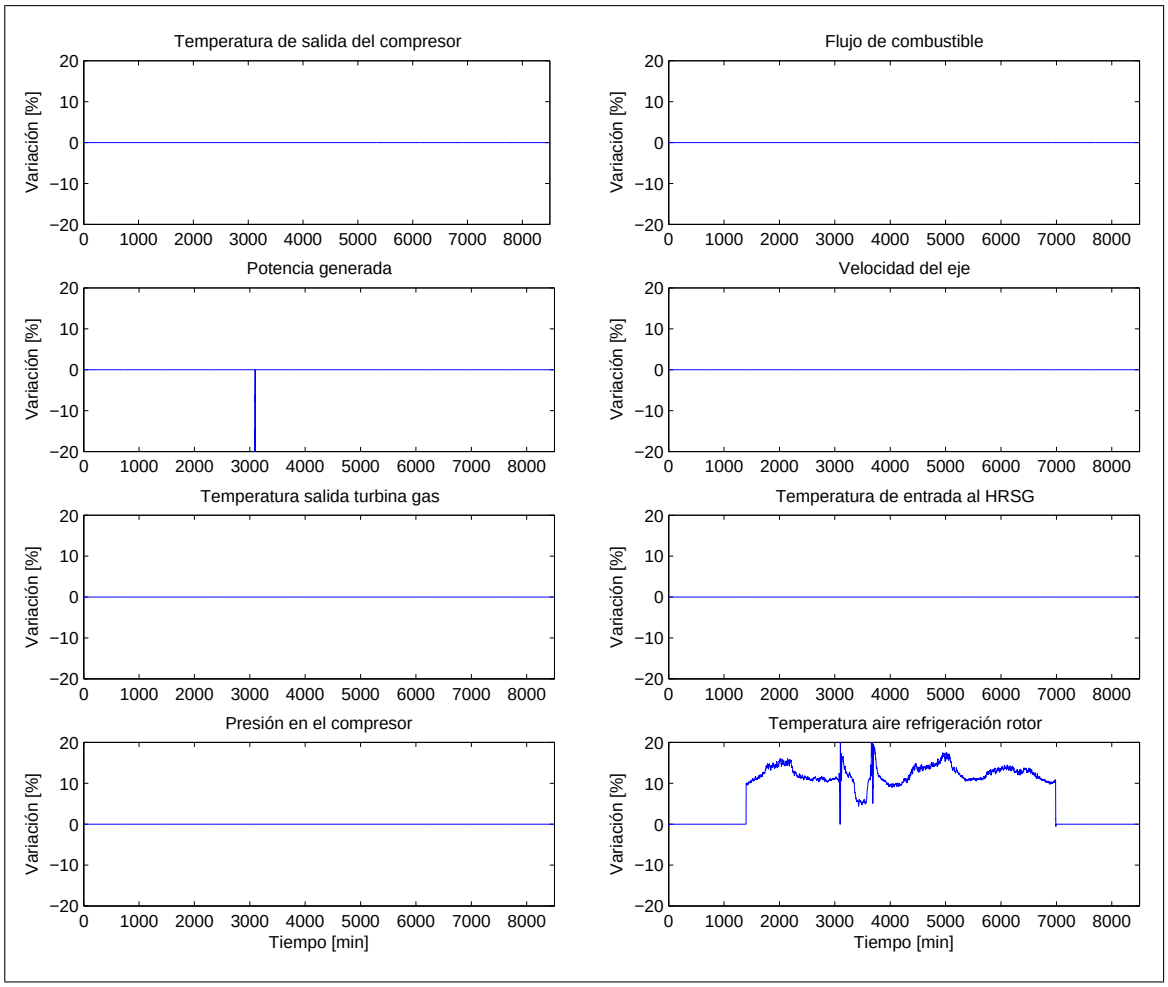

FIGURA 6.24. Señales de corrección de variables medidas, para falla real.

En esta figura se puede apreciar la potencialidad del filtraje a través de la descomposición en componentes principales. Se observa como una fuerte perturbación en los datos, debida al cambio del tipo de combustible (que no representa una falla), es eliminada completamente de las señales de corrección luego del filtraje. Sólo permanece la señal de corrección sobre el sensor efectivamente fallado.

Finalmente, las señales corregidas se muestran en la figura 6.25.

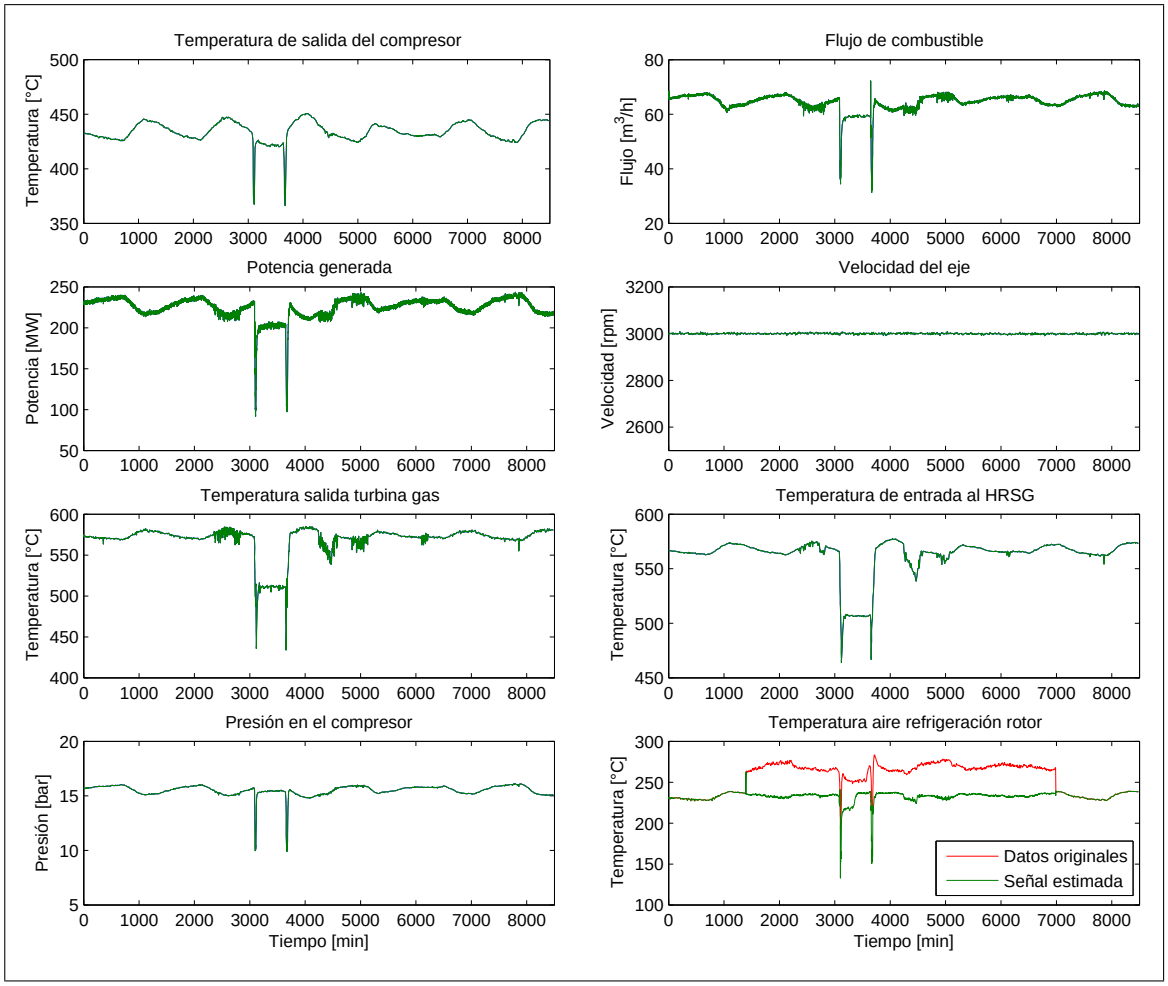

FIGURA 6.25. Señales corregidas, para falla real.

Se aprecia que a pesar de que no se cuenta con una referencia precisa de la falla, la estimación realizada para el sensor dañado concuerda a primera vista con el comportamiento de las demás señales asociadas a este proceso.

6.4.5. Caso 5: Falla de varianza del sensor

Si bien los casos previamente comentados consideran varios tipos de fallas, en general éstos tienen relación con fallas en el valor esperado de la medición del sensor para cada instante t, $E\{u(t)\}$ o $E\{y(t)\}$ según corresponda.

En este caso se analiza la presencia de una falla en la varianza del sensor, simulada como un ruido con distribución gaussiana de amplitud creciente aplicada al sensor de temperatura de salida del compresor.

La figura 6.26 muestra las estimaciones de las señales de falla en este caso.

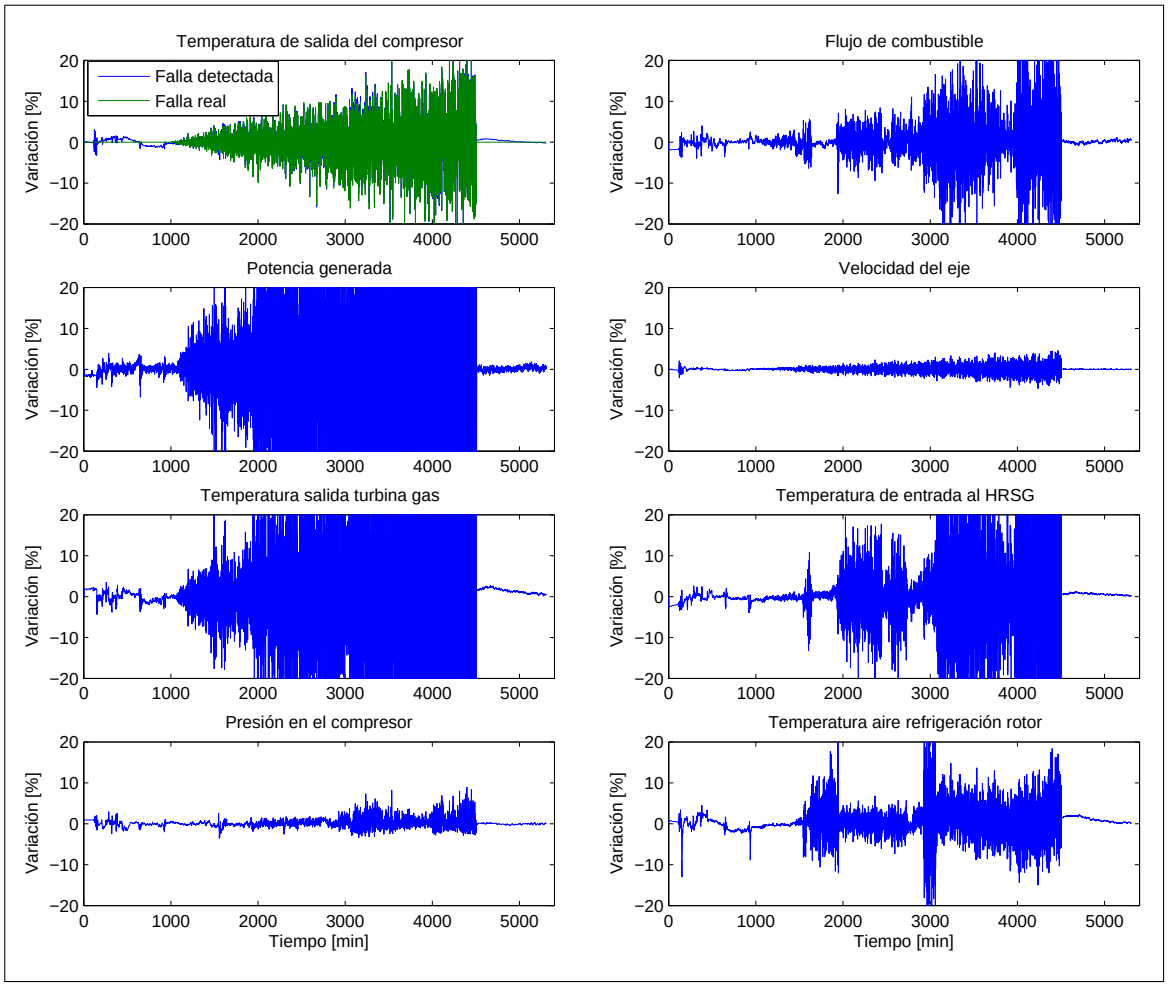

FIGURA 6.26. Estimación de las señales de falla para falla de varianza.

Como puede verse en la figura, en este caso la estimación de señales de falla realizada por el estimador es defectuosa en las variables que no estan sujetas a falla. Esto se debe a la rapidez del cambio en la señal introducido por el ruido. Como el ruido realiza cruces intermitentes por cero, y como se vio en el caso de la falla no abrupta, el test Q tendrá

problemas para realizar la detección de fallas en dichos puntos. El test Q en este caso se muestra en la figura 6.27.

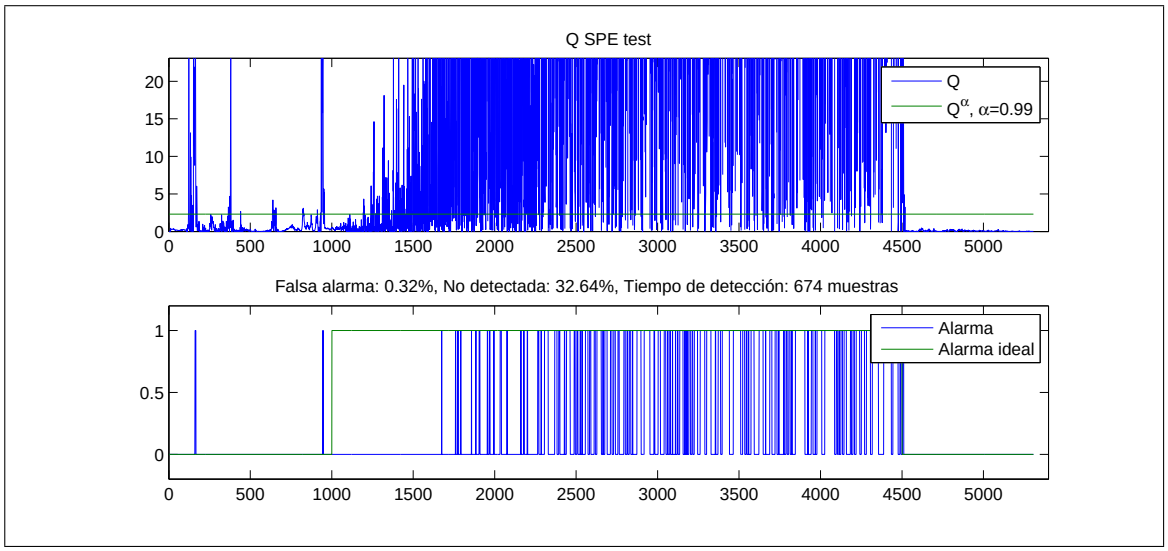

FIGURA 6.27. Test Q para falla de varianza.

Las señales de corrección en este caso son de baja calidad, con mucha interferencia por las malas estimaciones de los sensores sin falla.

Para corregir esta situación, se agrega un módulo adicional al sistema que calcula las varianzas de las estimaciones de falla, $V\{\hat{f}_u(t)\}$ y $V\{\hat{f}_y(t)\}$, a través de una ventana móvil de n muestras pasadas en cada instante t. Se genera un índice de varianzas conjuntas de la forma

$$V_f(t) = \max\left(V\{\hat{f}_{u_i}(t)\}, V\{\hat{f}_{y_j}(t)\}\right), \forall i, j \tag{6.7}$$

De esta forma, se detecta una *falla de varianza* en el caso de que este índice supere una cota establecida empíricamente en base al comportamiento del estimador ante ausencia o presencia de fallas normales (fallas de valor esperado). La figura 6.28 muestra la evolución de este índice y la alarma por falla de varianza en el caso de las señales de la figura 6.26.

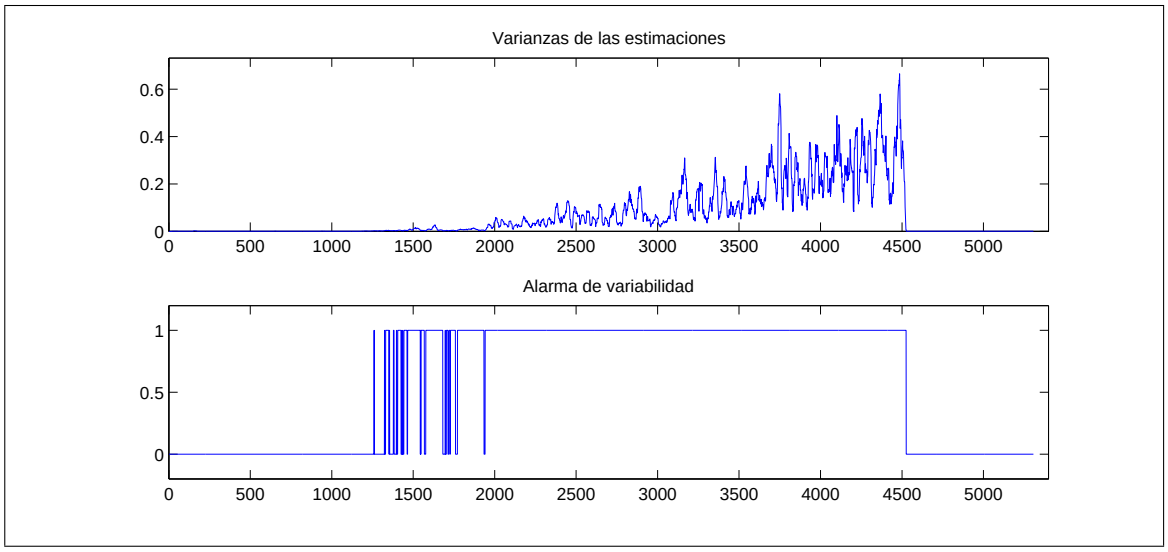

FIGURA 6.28. Análisis para falla de varianza.

En este caso, se determina el sensor en falla realizando la estimación de la varianza de las señales corregidas preliminarmente de la forma:

$$V\left\{\begin{bmatrix}\hat{u}(t)\\ \hat{y}(t)\end{bmatrix}\right\} = V\left\{\begin{bmatrix}\tilde{u}(t)\\ \tilde{y}(t)\end{bmatrix} - \begin{bmatrix}\hat{f}_u(t)\\ \hat{f}_y(t)\end{bmatrix}\right\} \tag{6.8}$$

El sensor cuya varianza es la menor (debido a la anulación realizada por una correcta estimación de la señal de falla), se declara como el sensor en falla. La figura 6.29 muestra las señales de corrección utilizando el indicador de varianza planteado. La figura 6.30 muestra las señales corregidas en este caso.

TABLA 6.1. Resultados para los casos en estudio (FA: Falsa alarma, ND: No detectada, TD: Tiempo de deteccion, FIT= Índice de ajuste).

Caso	Esquema original				Módulo varianza			
	FA	ND	TD	FIT	FA	ND	TD	FIT
Caso 2: Abrupto	0.21	0.00	9	93.55	0.66	0.00	0	98.94
Caso 3: No abrupto	0.19	6.52	13	95.66	0.19	6.52	13	95.66
Caso 4: Real	0.04	0.00	9	N/A	0.04	0.00	9	N/A
Caso 5: Varianza	0.32	32.64	674	53.67	0.66	3.58	259	94.11

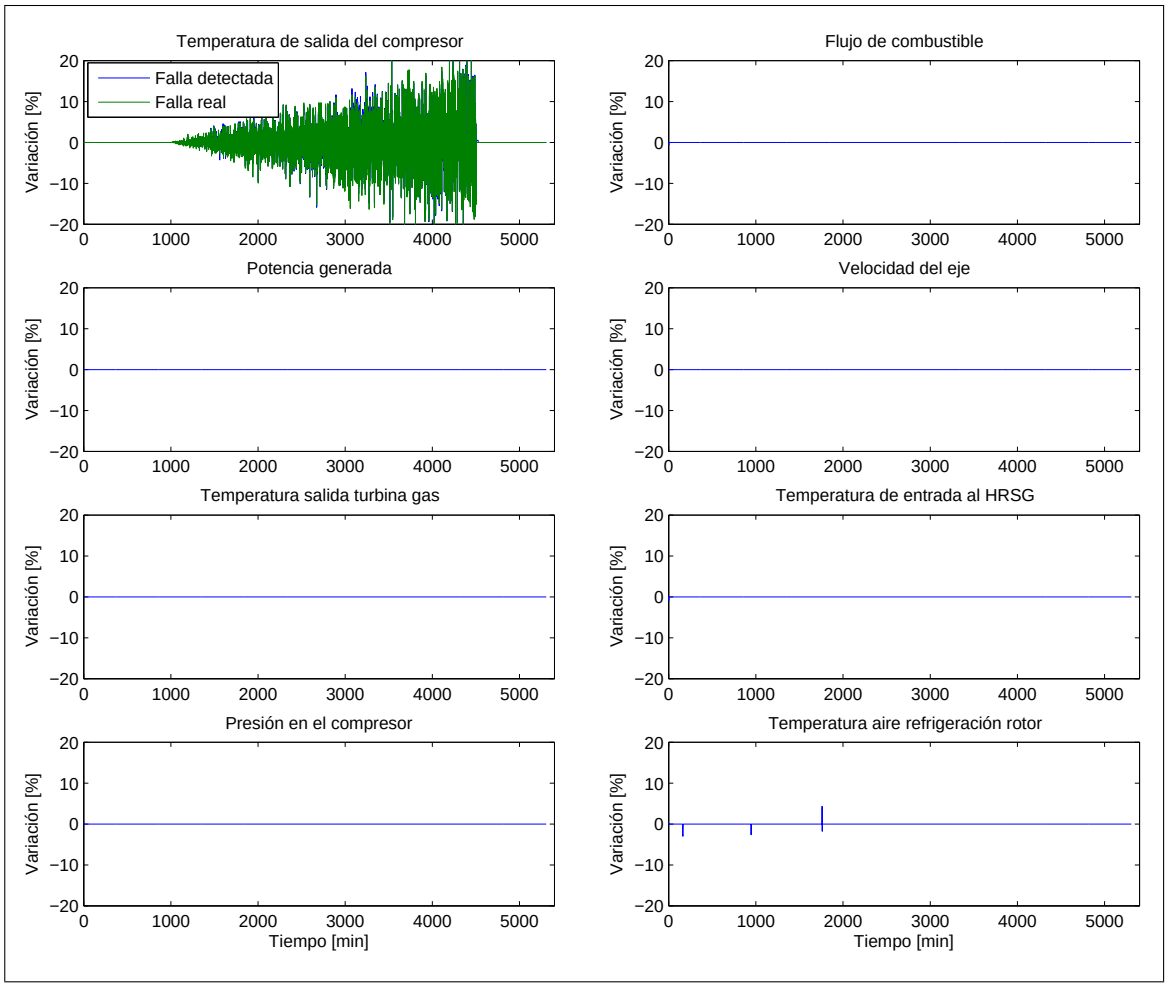

FIGURA 6.29. Señales de corrección de variables medidas, para falla de varianza.

Se aprecia que en este caso, la corrección realizada por el sistema con el módulo de detección de fallas de varianza es correcta. La tabla 6.1 resume los indicadores de desempeño para los distintos casos en análisis antes y después de la aplicación del módulo de detección de fallas de varianza. Se ve que en el caso de la falla de varianza la mejora de desempeño en todos los indicadores es dramática. Incluso se aprecia una mejora en el caso 2 de falla abrupta en el tiempo de detección, que pasa a ser instantáneo una vez ocurrida la falla. El indicador de falsa alarma se incrementa, pero tiene relación con la extensión de la alarma una vez concluídá la falla, y no con la aparición de falsas alarmas en otros puntos de la simulación.

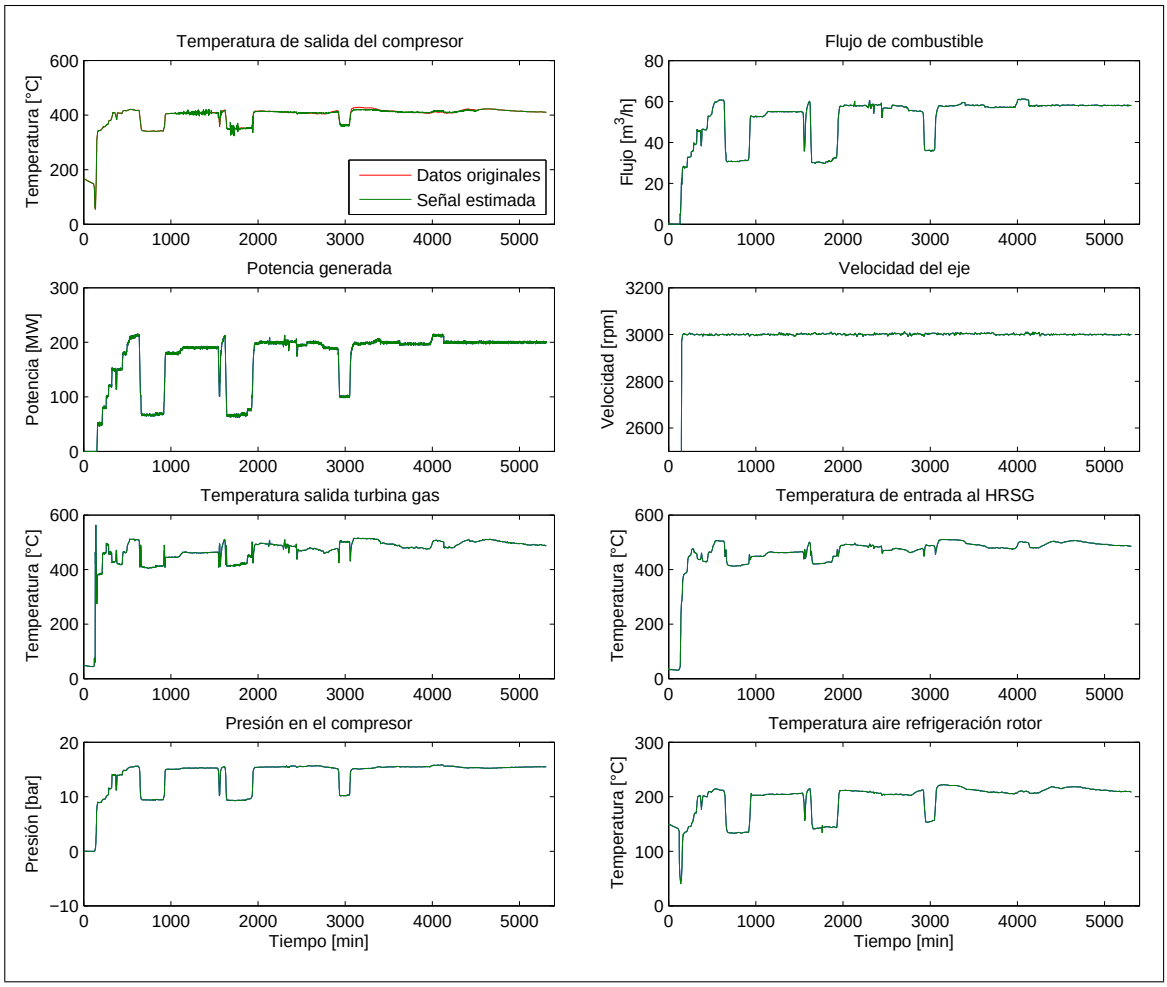

FIGURA 6.30. Señales corregidas, para falla de varianza.

6.5. Esquema de aplicación del sistema

El sistema de detección, diagnóstico y corrección de fallas planteado en las secciones anteriores está pensado para incorporarse como un módulo anexo al sistema de supervisión y control SCADA de la central de ciclo combinado. El esquema propuesto de funcionamiento se muestra en la figura 6.31.

Como se aprecia, se incluye un módulo de entrenamiento que realiza el ajuste de los parámetros del estimador de falla en caso de pequeños desajustes luego de largos períodos de tiempo en operación. Se incluye una interfaz de usuario en donde el sistema realiza una labor de apoyo al operador entregando las alarmas y las mediciones corregidas de las

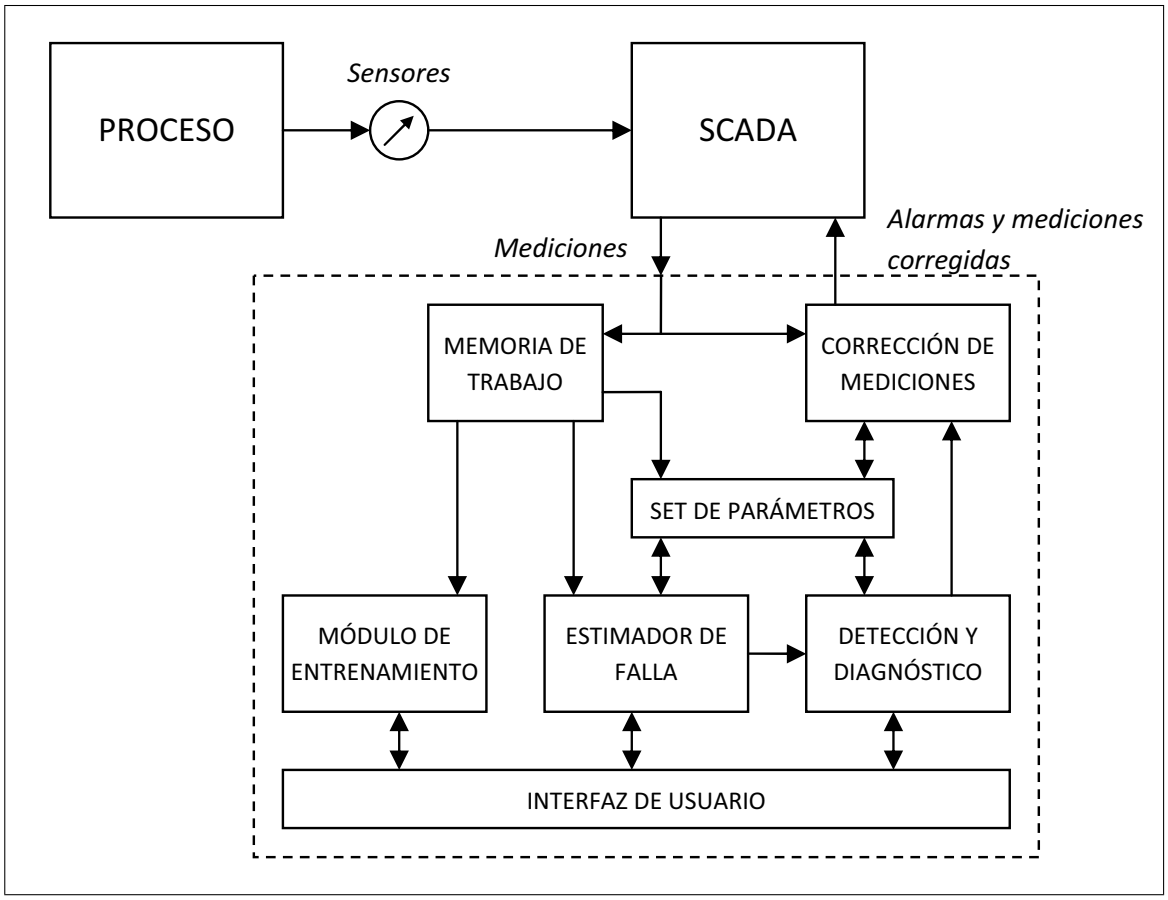

FIGURA 6.31. Esquema de operación ideado para el módulo de mediciones tolerantes a fallas.

señales. La conexión con el sistema SCADA pretende, una vez que el sistema se ha validado totalmente, realizar el reemplazo directo de las medidas provenientes de los sensores por las medidas del módulo de mediciones tolerantes a fallas.

6.6. Conclusiones

En el desarrollo de este capítulo se aprecia como el esquema propuesto de identificación de señales de falla es factible de implementar a través del uso de técnicas inteligentes, en particular, de los modelos difusos de Takagi-Sugeno. La combinación de los resultados de esta técnica con el análisis de componentes principales se convierten en una

potente herramienta para detectar, diagnosticar y corregir fallas en las mediciones de variables a través de sensores. Además, la inclusión de un módulo de monitoreo de variabilidad incrementa la robustez del sistema frente a fallas de varianza sobre los sensores.

Cabe destacar que el desarrollo realizado fue pensado en todo momento para su aplicación en línea y en tiempo real a la planta. El análisis y los graficos generados no se realizan en base a un análisis *a posteriori* de los datos, si no que se aplican en forma secuencial tomando en cuenta sólo la información pasada. También se consideró el esfuerzo computacional requerido. Los algoritmos se desarrollaron en el lenguaje de MATLAB, sin embargo están estructurados como funciones independientes que no dependen de los comandos propios del software utilizado, por lo que son fácilmente portables a lenguajes estándar como C++ o el empleado por los módulos de desarrollo de los sistemas SCADA de las plantas de ciclo combinado. Sin perjuicio de lo anterior, el desarrollo puede ser utilizado para un análisis fuera de línea basado en datos históricos.

7. CONCLUSIONES Y TRABAJO FUTURO

7.1. Conclusiones y comentarios generales

En el desarrollo de esta tesis se comprobó como los sistemas inteligentes se aplican al desarrollo de aplicaciones a una central de ciclo combinado, tanto en modelación y simulación como en la detección y diagnóstico de fallas.

La creciente complejidad de los procesos industriales dejan en clara desventaja los enfoques clásicos basados en modelación lineal para el desarrollo de aplicaciones de monitorización y diagnóstico. En el capítulo 4 quedó de manifiesto la importante diferencia de casi 6 puntos porcentuales en promedio entre los enfoques lineales y las técnicas inteligentes. Sin embargo, esta mejora tiene un costo que es el esfuerzo computacional que requiere el entrenamiento e identificación de estructura y parámetros de los modelos basados en inteligencia artificial. Las redes neuronales, cuya principal desventaja es que son modelos de caja negra en que sus parámetros no son interpretables de forma intuitiva, requieren un considerable esfuerzo de cómputo para su entrenamiento, el que puede realizarse sólo por algoritmos numéricos debido a la incapacidad de obtener una solución analítica cerrada para un conjunto óptimo de parámetros.

Acá surge la ventaja de los modelos difusos de Takagi-Sugeno, en los cuales se puede fijar de los parámetros de las funciones de pertenencia de forma manual, obteniendo una solución cerrada (en base a mínimos cuadrados) para los parámetros en las funciones de salida. Esta fijación de parámetros se realiza en base al sentido común y la incorporación de la experiencia de los operadores, debido a que representan puntos de operación del proceso general. El modelo obtenido por esta vía, a pesar de tener un carácter subóptimo, logró índices de ajuste incluso superiores a los conseguidos por el mismo tipo de modelo con estimación óptima de parámetros o por los modelos en redes neuronales, con un esfuerzo computacional notablemente inferior (por ejemplo, 5 minutos para estimar estructura y parámetros contra más de 100 horas en el caso de una red neuronal tipo perceptrón multicapa).

En este caso, el tiempo de cómputo no es una desventaja de la modelación inteligente en comparación a las técnicas clásicas.

La versatilidad de los modelos difusos de Takagi-Sugeno queda de manifiesto en los capítulos 5 y 6, en donde se aplican como estimadores de falla para el sistema de detección y diagnóstico de fallas en sensores. En este caso, el ajuste proporcionado por este tipo de modelo es suficientemente bueno como para realizar la reconstrucción de señales luego de ser detectada una falla. Acá es clave la aplicación de la modelación inteligente para lograr los resultados esperados.

El sistema de medición tolerante a fallas propuesto en el capítulo 6 completa una de las líneas propuestas por Silvio Simani en su tesis de doctorado, base para el libro (Simani et al., 2000). Tanto en este libro como en la mayoría de la literatura disponible sobre detección y diagnóstico de fallas en sensores se utilizan modelos linealizados del proceso, y se trabaja sobre los resultados obtenidos del modelo y no sobre los datos reales del proceso. El desarrollo realizado suple esta falencia proponiendo una estructura que se aplica directamente sobre el proceso sin realizar mayores suposiciones. Otro punto importante es que los desarrollos disponibles en la literatura sobre FDD en sensores (tanto de plantas de ciclo combinado como otros procesos) solo tratan en general la detección de la falla y no la reparación de ésta. El diseño realizado proporciona un esquema robusto de reparación de fallas, siendo una contribución real e importante al estado del arte de sistemas de detección y diagnóstico.

En el caso del sistema de detección y diagnóstico de fallas, los sistemas inteligentes introducidos en la forma de un sistema difuso de Takagi-Sugeno, aparecen complementados por la aplicación del análisis estadístico mediante descomposición en componentes principales. La combinación de las ventajas de cada enfoque dan origen a buenos resultados para la monitorización de las mediciones.

Tanto los algoritmos de modelación y simulación empleados como los del sistema de medición tolerante a fallas, fueron desarrollados sobre el sistema MATLAB. Sin embargo, fueron pensados y estructurados de forma estándar, sin recurrir a comandos propietarios del

software, de forma que pueden ser fácilmente adaptados a otros lenguajes de programación como C++ o los empleados por los módulos de desarrollo de los sistemas SCADA de las plantas. El esfuerzo computacional empleado siempre fue considerado, de forma que los desarrollos planteados en esta tesis están pensados para ser aplicados en línea y en tiempo real a los procesos.

Otro comentario importante es que el conjunto de datos con que se trabajó fue muy rico, al contener una partida del sistema y recorrer así diversos puntos de operación. De todas formas, el desempeño de los desarrollos realizados es mejorable considerando un conjunto de datos elegido más cuidadosamente y que abarque absolutamente todos los estados de operación del proceso. Esto no fue posible realizarlo en la práctica, pero es posible realizarlo en una futura implementación de estos desarrollos gracias a las capacidades de entrenamiento de los sistemas inteligentes empleados.

Una de las potencialidades mayores de este trabajo es que no se realizan supuestos mayores sobre el proceso en estudio, por lo que los resultados aquí obtenidos pueden extenderse a otros procesos industriales.

Respecto a la aplicabilidad de los desarrollos realizados en esta tesis, es importante mencionar que todas las técnicas y los algoritmos involucrados fueron desarrollados pensando en una fácil migración a un lenguaje adecuado y en una implementación en tiempo real. Actualmente, la central San Isidro es monitorizada desde el CMD de Endesa Chile, bajo el software PI de OSISOFT. Este software permite el desarrollo de aplicaciones a través de sus módulos de expansión. Tanto el modelo generado como el sistema de detección y diagnóstico de fallas en sensores de la turbina de gas pueden implementarse en módulos anexos para ser visualizados desde el CMD, y ser finalmente validados. Con respecto a eso último, y pese a los buenos resultados obtenidos, debe realizarse una extensa validación, utilizando nuevos datos disponibles, sobre los desarrollos realizados al momento de su implementación, de forma de generar la confianza necesaria en el sistema como para ser un apoyo en la toma de decisiones futuras.

Como conclusión general, la aplicación de los sistemas inteligentes resultó decisiva en el éxito de los desarrollos de esta tésis. Como puede verse, el grueso de los resultados dependen de las ventajas ofrecidas por los métodos englobados en esta denominación.

7.2. Temas de investigación futura

Como temas de investigación futura, y a la luz de los resultados obtenidos en esta tesis, se propone:

(i) Estudio detallado del desempeño de la identificación de modelos difusos de Takagi-Sugeno fijando los parámetros de funciones de pertenencia a priori, en comparación con la estimación conjunta mediante algoritmos de descenso: Como se muestra en el ajuste de parámetros realizado en el capítulo 4, el modelo difuso de Takagi-Sugeno identificado de forma subóptima al prefijar centros y *spreads* de las funciones de pertenencia no se diferencia significativamente del modelo cuyos parámetros se identifican en forma conjunta y óptima a través del algoritmo de Levenberg-Marquardt. Sin embargo, el uso del modelo subóptimo representa una gran ventaja desde el punto de vista del esfuerzo computacional invertido en su entrenamiento. Resulta interesante entonces estudiar si este comportamiento es recurrente sobre otros procesos industriales, y si es posible demostrarlo de forma teórica.

(ii) Estudio de la aplicación del esquema propuesto en la detección de fallas en sensores, a la detección y el diagnóstico de fallas propias del proceso, como las comentadas en el capítulo 2 como más recurrentes: Si bien las fallas en sensores son el tipo de falla más frecuente en una central de ciclo combinado, es interesante de todas formas generar un sistema de detección y diagnóstico de las fallas propias del proceso, que pueda operar en conjunto con el sistema de mediciones tolerante a fallas.

(iii) Estudio de modificaciones sobre el esquema propuesto en la detección de fallas sobre sensores, para proporcionar una capacidad de tolerancia a fallas simultáneas: Si bien el sistema de mediciones tolerantes a fallas diseñado en esta tesis es capaz de manejar de forma separada las fallas en los sensores, es interesante estudiar el efecto de diversas fallas ocurridas de forma simultánea y adaptar el esquema propuesto para hacerlo robusto frente a este escenario.

(iv) Estudio de la aplicabilidad del esquema propuesto para otro tipo de procesos industriales: El esquema propuesto de mediciones tolerantes a fallas considera un esquema general, pero sólo fue probado y adaptado considerando información de una central de ciclo combinado. El problema de la ocurrencia de fallas en sensores es relevante en todos los procesos industriales complejos, luego es interesante estudiar si el diseño propuesto es funcional en procesos en general, y determinar cuáles son sus limitaciones de aplicabilidad.

BIBLIOGRAFIA

Ailer, P., Santa, I., Szederkényi, G., & Hangos, K. M. (2001, January). Nonlinear model-building of a low-power gas turbine. *Periodica Polytechnica Ser. Transportation Engineering*, *2*, 117-135.

Albanesi, C., Bossi, M., Magni, L., Paderno, J., Pretolani, F., Kuehl, P., et al. (2006, December). Optimization of the start-up procedure of a combined cycle power plant. *45th IEEE Conference on Decision and Control*, 1840-1845.

Araya, J. (2004). Estimación de parámetros para su aplicación a detección y diagnóstico de fallas. *Tesis de Magíster, Pontificia Universidad Católica de Chile*.

Arriagada, J., Genrup, M., Loberg, A., & Assadi, M. (2003, November). Fault diagnosis system for an industrial gas turbine by means of neural networks. In *International gas turbine congress.*

Babuska, R., Roubos, J., & Verbruggen, H. (1998, May). Identification of mimo systems by input-output ts fuzzy models. *Fuzzy Systems Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on*, *1*, 657-662 vol.1.

Babuska, R., & Verbruggen, H. (1995). Identification of composite linear models via fuzzy clustering. *In Proceedings of European Control Conference*, 1207–1212.

Basseville, M. (1988). Detecting changes in signals and systems—a survey. *Automatica*, *24*(3), 309–326.

Basseville, M., & Nikiforov, I. V. (1993). *Detection of abrupt changes: theory and application*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc.

Binsompasong, S. L. P. J. J., P.; Leeragreephol. (2006, October). Case study of protecting gas turbine trip from damper closing failure and hrsg trip from solenoid coil of hydraulic damper burned. *SICE-ICASE, 2006. International Joint Conference*, 467-471.

Boissenin, Y., & Castanier, A. (1988). *Choosing the right combined cycle power plant*. Alsthom Gas Turbines Reference Library.

Bolland, O. (1991, April). A comparative evaluation of advanced combined cycle alternatives. *ASME Journal of Engineering for Gas Turbines and Power*, *113*, 190-197.

Boyce, M. P. (1982). *Gas turbine engineering handbook*. Gulf Publishing Company, Houston, Texas.

Chen, J., & Patton, R. J. (1999). *Robust model-based fault diagnosis for dynamic systems*. Norwell, MA, USA: Kluwer Academic Publishers.

Ferrari-Trecate, G., Gallestey, E., Letizia, P., Spedicato, M., Morari, M., & Antoine, M. (2004, September). Modeling and control of co-generation power plants: a hybrid system approach. *IEEE Transactions on Control Systems Technology*, *12*(5), 694-705.

Flores-Loredo, Z., Ibargüengoytia, P. H., & Morales, E. F. (2005, November). On line diagnosis of gas turbines using probabilistic and qualitative reasoning. *Intelligent Systems Application to Power Systems, 2005. Proceedings of the 13th International Conference on*, 5 pp.-.

Franke, R., & Vogelbacher, L. (2006, December). Nonlinear model predictive control for cost optimal startup of steam power plants. *Automatisierungstechnik*, *54*, 630-637.

Gómez, E., Unbehauen, H., Kortmann, P., & Peters, S. (1996). Fault detection and diagnosis with the help of fuzzy-logic and with application to a laboratory turbogenerator. *In Proceedings of 13th IFAC World Congress, San Francisco*, 235–240.

González, G., Orchard, M., Cerda, J., Casali, A., & Vallebuona, G. (2003). Local models for soft-sensores in a rougher flotation bank. *Minerals Engineering*, *16*, 441–453.

Hannett, L., & Khan, A. (1993). Combustion turbine dynamic model validation from tests. *IEEE Transactions on Power Systems*, *8*, 152-158.

Harman, R. T. (1981). *Gas turbine engineering*. John Wiley & Sons, New York.

Heo, J. S., Lee, K. Y., & Garduno-Ramirez, R. (2006, June). Multiobjective control of power plants using particle swarm optimization techniques. *IEEE Transaction on Energy Conversion*, *21*(2), 552-561.

Hu, N.-S., He, N.-N., & Hu, S. (2003, November). Fault diagnosis of the steam turbine condenser system based on som neural network. *Machine Learning and Cybernetics, 2003 International Conference on*, *2*, 1222-1225 Vol.2.

Huang, X., Qi, H., & Liu, X. (2006, June). Implementation of fault detection and diagnosis system for control systems in thermal power plants. *Intelligent Control and Automation, 2006. WCICA 2006. The Sixth World Congress on*, *2*, 5777-5781.

Hung, W. (Jul 1991). Dynamic simulation of gas-turbine generating unit. *Generation, Transmission and Distribution [see also IEE Proceedings-Generation, Transmission and Distribution], IEE Proceedings C*, *138*(4), 342-350.

Ibargüengoytia, P. H., Sucar, L. E., & Morales, E. (2000). A probabilistic model approach for fault diagnosis. In *MICAI* (p. 687-698).

Ibargüengoytia, P. H., Sucar, L. E., & Vadera, S. (2001, November). Real time intelligent sensor validation. *IEEE Transactions on Power Systems*, *16*(4), 770-775.

Jackson, J. (2003). *A user's guide to principal components (wiley series in probability and statistics)*. Wiley-Interscience.

Jackson, J., & Mudholkar, S. (1979). Control procedures for residuals associated with principal component analysis. *Technometrics*, *21*(3), 341–349.

Jian, O., Cai-xin, S., Weimin, B., Bide, Z., & Ruijin, L. (2002). A steam turbine-generator vibration fault diagnosis method based on rough set. *Power System Technology, 2002. Proceedings. PowerCon 2002. International Conference on*, *3*, 1532-1534 vol.3.

Jianhong, L., Laijiu, C., Sujin, Z., & Yiqian, W. (2002). Multivariable fuzzy pid control system for deaerator's and condenser's levels in a thermal power unit. *Cybernetics and Systems*, *33*, 483-506.

Jino, Y., Matsumoto, S., & Kamito, A. (2006, October). Risk based model predictive control with hybrid system structure and its application to thermal power plants. *SICE-ICASE, 2006. International Joint Conference*, 914-919.

Jurado, F., & Cano, A. (2003, February). *Fuzzy logic control of a combined cycle power plant using biomass* (No. 2).

Kallappa, P., & Ray, S. (1999). Fuzzy wide-range control of fossil power plants for life extension and robust performance. *Proceedings of the 1999 American Control Conference*, *5*, 3231-3235 vol.5.

Korbicz, J., Kościelny, J., Kowalczuk, Z., & Cholewa, W. (2004). *Fault diagnosis: Models, artificial intelligence, applications*. New York, NY, USA: Springer-Verlag New York, Inc.

Li, W., Yue, H. H., Valle-Cervantes, S., & Qin, S. (2000). Recursive PCA for adaptive process monitoring. *Journal of Process Control*, *10*, 471–486.

Liang-yu, M., Jian-qiang, G., & Bing-zhu, W. (2002). Fault intelligent diagnosis for high-pressure feedwater heater system of a 300 mw coal-fired power unit based on improved bp neural network. *Power System Technology, 2002. Proceedings. PowerCon 2002. International Conference on*, *3*, 1535-1539 vol.3.

Ljung, L. (Ed.). (1999). *System identification (2nd ed.): theory for the user*. Upper Saddle River, NJ, USA: Prentice Hall PTR.

Ma, Y.-G., Ma, L.-Y., & Ma, J. (2005, August). RBF neural network based fault diagnosis for the thermodynamic system of a thermal power generating unit. *Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on*, *8*, 4738-4743 Vol. 8.

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, *11*(2), 431–441.

Maslo, K., & Andel, J. (2001). Gas turbine model using in design of heat and power stations. *Power Tech Proceedings, 2001 IEEE Porto*, *4*, 6 pp. vol.4-.

Matsumoto, H., Ohsawa, Y., Takahasi, S., Akiyama, T., & Ishiguro, O. (1996, June). An expert system for startup optimization of combined cycle power plants under nox emission regulation and machine life management. *IEEE Transaction on Energy Conversion*, *11*(2), 414-422.

Matta, R. K., Mercer, G. D., & Tuthill, R. S. (2000). Power systems for the 21st century - "H" gas turbine combined-cycles. *GE Power Systems*.

Miguel, M. I., & Delgadillo, A. (2004, Febrero). *Reporte de estudio: Descripción de fallas en turbinas de gas* (Tech. Rep.). Cuernavaca, Morelos: Instituto de Investigaciones Eléctricas.

Omidvar, O. M., & Elliott, D. (Eds.). (1997). *Neural systems for control*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Ordys, A., Pike, A., Johnson, M., Katebi, R., & Grimble, M. (1994). *Modelling and simulation of power generation plants*. London: Springer Verlag.

Overschee, P. V., & Moor, B. D. (1994). N4sid: subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica*, *30*(1), 75–93.

Palade, V., Patton, R. J., Uppal, F., Quevedo, J., & Daley, S. (2002). Fault diagnosis of an industrial gas turbine using neuro-fuzzy methods. In *Proc. of the 15th IFAC World Congress* (p. 2477-2482).

Palade, V., Patton, R. J., Uppal, F. J., Quevedo, J., & Daley, S. (2002, July). Fault diagnosis of an industrial gas turbine using neuro-fuzzy methods. In *15th IFAC World Congress* (p. 2477-2482). Barcelona, Spain.

Passino, K. M., & Yurkovich, S. (1997). *Fuzzy control*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.

Patton, R., Frank, P., & Clark, R. (2000). *Issues of fault diagnosis for dynamic systems*. Springer-Verlag London, Limited.

Peng, H., Ozaki, T., Haggan-Ozaki, V., & Toyoda, Y. (2002, March). A nonlinear exponential arx model-based multivariable generalized predictive control strategy for thermal power plants. *IEEE Transactions on Control Systems Technology*, *10*(2), 256-262.

Poncia, G., & Bittanti, S. (20 July 2001). Multivariable model predictive control of a thermal power plant with built-in classical regulation. *International Journal of Control*, *74*, 1118-1130.

Prasad, G., Swidenbank, E., & Hogg, B. W. (1998, June). A neural net model-based multivariable long-range predictive control strategy applied in thermal power plant control. *IEEE Transaction on Energy Conversion*, *13*(2), 176-182.

Roizman, O., & Davydov, V. (Jul 1999). Neuro-fuzzy computing for large power transformers monitoring and diagnostics. *Fuzzy Information Processing Society, 1999. NAFIPS. 18th International Conference of the North American*, 248-252.

Rovnak, J. A., & Corlis, R. (1991, June). Dynamic matrix based control of fossil power plants. *IEEE Transactions on Energy Conversion*, *6*(2), 320-326.

Rowen, W. I. (1983, October). Simplified mathematical representations of heavy-duty gas turbines. *ASME Journal of Engineering for Power*, 865-871.

Saez, D., & Cipriano, A. (1999a). Fuzzy modelling for a combined cycle power plant. *Fuzzy Systems Conference Proceedings, 1999. FUZZ-IEEE '99. 1999 IEEE International*, *2*, 1186-1190 vol.2.

Saez, D., & Cipriano, A. (1999b). Fuzzy models based economic predictive control for a combined cycle power plant boiler. *Proceedings of the 1999 IEEE International Symposium on Intelligent Control/Intelligent Systems and Semiotics*, 417-422.

Saez, D., Milla, F., & Vargas, L. S. (2007, September). Fuzzy predictive supervisory control based on genetic algorithms for gas turbines of combined cycle power plants. *IEEE Transaction on Energy Conversion*, *22*(3), 689-696.

Sampath, S., Gulati, A., & Singh, R. (2002, July). Artificial intelligence techniques for gas turbine engine fault diagnosis. In *38th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit.* Indianapolis, Indiana, USA.

Sanchez, E., Moheno, M., Suarez, D. A., & Quintero, A. (2002). Electric power fossil plant start-up combining predictive control and takagi-sugeno fuzzy modelling. *Proceedings of the 2002 IEEE International Symposium on Intelligent Control*, 116-120.

Shields, C. (1989). Current developments in gas turbine combined cycle power plant. *Publication 453, Institution of Diesel and Gas Turbine Engineers*.

Simani, S. (2005, August). Identification and fault diagnosis of a simulated model of an industrial gas turbine. *Industrial Informatics, IEEE Transactions on*, *1*(3), 202-216.

Simani, S., & Fantuzzi, C. (2000). Fault diagnosis in power plant using neural networks. *Inf. Sci.*, *127*(3-4), 125–136.

Simani, S., Fantuzzi, C., & Beghelli, S. (2000, Sep). Diagnosis techniques for sensor faults of industrial processes. *Control Systems Technology, IEEE Transactions on*, *8*(5), 848-855.

Simani, S., Fantuzzi, C., & Spina, R. (1998, Sep). Application of a neural network in gas turbine control sensor fault detection. *Control Applications, 1998. Proceedings of the 1998 IEEE International Conference on*, *1*, 182-186 vol.1.

Simani, S., Patton, R., & Fantuzzi, C. (2003). *Model-based fault diagnosis in dynamic systems using identification techniques*. Secaucus, NJ, USA: Springer-Verlag New York, Inc.

Sánchez, M., & Verde, C. (2006, June). Analytical redundancy for a gas turbine of a combined cycle power plant. In *American control conference* (p. 4442-4447). Minneapolis, Minnesota, USA.

Sun, X., Chen, T., & Marquez, H. J. (2001). Detecting leaks and sensor biases by recursive identification with forgetting factors. *Decision and Control, 2001. Proceedings of the 40th IEEE Conference on*, *4*, 3716-3721 vol.4.

Suzaki, S., Kawata, K., Sekoguchi, M., & Goto, M. (2000). Mathematical model for a combined cycle plant and its implementation in an analogue power system simulator. *Power Engineering Society Winter Meeting, 2000. IEEE*, *1*, 416-421 vol.1.

Tan, S. C., & Lim, C. P. (2004, June). Application of an adaptive neural network with symbolic rule extraction to fault detection and diagnosis in a power generation plant. *Energy Conversion, IEEE Transaction on*.

Tian, Z., Zhang, H., & Wang, Y. (2006). Failure diagnosis of steam turbine-generator set based on genetic algorithm and bp network. In *Isda '06: Proceedings of the sixth international conference on intelligent systems design and applications (isda'06)* (pp. 930–935). Washington, DC, USA: IEEE Computer Society.

Wasserman, P. D. (1993). *Advanced methods in neural computing*. New York, NY, USA: John Wiley & Sons, Inc.

Westinghouse, E. C. (1975). *Gas turbine systems primer.*

Working Group, o. (Aug 1994). Dynamic models for combined cycle plants in power system studies. *Power Systems, IEEE Transactions on*, *9*(3), 1698-1708.

Wu, J.-F., Hu, N.-S., Hu, S., & Zhao, Y. (2002). Application of som neural network in fault diagnosis of the steam turbine regenerative system. *Machine Learning and Cybernetics, 2002. Proceedings. 2002 International Conference on*, *1*, 184-187 vol.1.

Xiang, C., Huang, X., Huang, D., & Hu, J. (2006). Wavelets neural network based on particle swarm optimization algorithm for fault diagnosis. In *Icicic '06: Proceedings of the first international conference on innovative computing, information and control* (pp. 320–323). Washington, DC, USA: IEEE Computer Society.

Xu, T., & Wang, Q. (2006). Application of MSPCA to sensor fault diagnosis. *Acta Automatica Sinica*, *32*, 417–421.

Xue, G., Xiao, L., Bie, M., & Lu, S. (2005, May). Fault prediction of boilers with fuzzy mathematics and RBF neural network. *Communications, Circuits and Systems, 2005. Proceedings. 2005 International Conference on*, *2*, -1016.

Yang, P., & Liu, S. S. (2004, December). Fault diagnosis for boilers in thermal power plant by data mining. *Control, Automation, Robotics and Vision Conference, 2004. ICARCV 2004 8th*, *3*, 2176-2180 Vol. 3.

Zhang, H., Li, M., & Cai, L. (2002). Fuzzy adaptive control of nonlinear processes with feedforward compensator and its application. *Cybernetics and Systems*, *33*, 171-187.

Zhang, Q., & So, P. (2000). Dynamic modelling of a combined cycle plant for power system stability studies. *Power Engineering Society Winter Meeting, 2000. IEEE*, *2*, 1538-1543 vol.2.

ANEXO A. ESTIMACIÓN DE LOS TÉRMINOS DE RUIDO PARA EL MODELO DE ECUACIÓN DE ESTADO

En este anexo se estiman los coeficientes asociados al ruido en el modelo de ecuación de estado definido en (3.53).

Consideremos el problema de predecir k pasos para la salida y, a partir de las salidas anteriores y las entradas u. Utilizando las definiciones realizadas, el problema es estimar θ_k y γ_k tal que

$$y(t+k-1) = \theta_k^T \varphi_s(t) + \gamma_k^T U_r(t) + \epsilon(t+k-1) \tag{A.1}$$

Considerando esta ecuación simultáneamente para r predictores, pueden agruparse las nuevas ecuaciones como

$$Y_r(t) = \Theta\varphi_s(t) + \Gamma U_r(t) + E(t) \tag{A.2}$$

Este problema corresponde a un problema de estimación equivalente al visto en la sección 3.4.1. Repitiendo la ecuación (A.2) para todos los tiempos t, se puede formular el conjunto de nuevas ecuaciones en la ecuación matricial:

$$Y = \Theta\Phi + \Gamma U + E \tag{A.3}$$

Extendiendo los resultados obtenidos en la sección 3.4.1, la estimación que minimiza el error cuadrático de los parámetros es

$$\begin{bmatrix} \hat{\Theta} & \hat{\Gamma} \end{bmatrix} = \begin{bmatrix} Y\Phi^T & YU^T \end{bmatrix} \begin{bmatrix} \Phi\Phi^T & \Phi U^T \\ U\Phi^T & UU^T \end{bmatrix}^{-1} \tag{A.4}$$

Realizando una inversión de matriz por bloques, se obtiene

$$\hat{\Theta} = YU^{\perp}\Phi^T \left(\Phi U^{\perp}\Phi^T\right)^{-1} \tag{A.5}$$

De esta forma, la predicción $\hat{Y} = \hat{\Theta}\Phi$ será

$$\begin{aligned}\hat{Y} &= YU^{\perp}\Phi^T\left(\Phi U^{\perp}\Phi^T\right)^{-1}\Phi \\ &= \frac{1}{N}YU^{\perp}\Phi^T\left(\frac{1}{N}\Phi U^{\perp}\Phi^T\right)^{-1}\Phi \\ &= W_1 G W_2 \end{aligned} \tag{A.6}$$

En donde

$$W_1 = I, \qquad W_2 = \left(\frac{1}{N}\Phi U^{\perp}\Phi^T\right)^{-1}\Phi \tag{A.7}$$

Se reconoce en esta expresión la matriz $\hat{G}$. Realizando la descomposición en valores singulares, y seleccionando los n valores más significativos, se obtiene

$$\hat{Y} \approx U_1 S_1 V_1^T \tag{A.8}$$

Despejando U_1 de la ecuación (3.73) y reemplazando, se obtiene

$$\hat{Y} \approx O_r R^{-1} S_1 V_1^T \tag{A.9}$$

Considerando la ecuación (3.60), se puede obtener una predicción de k pasos mediante la ecuación

$$\hat{Y} = O_r \hat{X} \tag{A.10}$$

Se desprende entonces de las ecuaciones (A.9) y (A.10) que

$$\hat{X} = R^{-1} S_1 V_1^T \tag{A.11}$$

Con esto, se pueden estimar los ruidos como

$$\begin{aligned} w(t) &= \hat{x}(t+1) - \hat{A}\hat{x}(t) - \hat{B}u(t) \\ v(t) &= y(t) - \hat{C}\hat{x} - \hat{D}u(t) \end{aligned} \tag{A.12}$$

A partir de esta información, se pueden estimar las matrices de covarianza de estos ruidos.