FAULT DIAGNOSIS IN CHEMICAL PLANTS INTEGRATED TO THE INFORMATION SYSTEM

Thesis

presented by

Diego Ruiz

(Chemical Engineer)

for the degree of Doctor

of the Universitat Politècnica de Catalunya

Barcelona, March 2001

Departament d'Enginyeria Química

Escola Tècnica Superior d'Enginyers Industrials de Barcelona

Universitat Politècnica de Catalunya

A Silvia
To Silvia

A mis padres
To my parents

A todos los que hacen las cosas con amor
A tots els que fan les coses amb amor
To the people who make the things with love

Acknowledgements

I thank Professor Luis Puigjaner Corbella for the direction of this thesis. I would like to highlight his permanent drive and support. His example of work and dedication as well as his warnings have enriched me.

I thank José María Nougués for his guide and patient attention during the development of the thesis. I would like to note his honesty and knowledge in the area process control. His critic and advice have been determinant for the quality of this work.

I would like to express my gratitude to all the doctorate candidates (Jordi C., Frank, Arndt, María A., Zuly, María S., Eduard, Raimon, Quique, Edwin, Saúl, Suly, Joaquín, Francisco, Joan, Rosario, Chouaib, Eloy, Vicky, Danelia, Dridi, Kamal, Meng, Javier and Rodolfo), to the students that have realised the career's final project (Ana, Jordi V., Marta, Carles V., Sandra, Hans, Sebastián, Carles L., Espen, Alex, Aprilla, Dirk and Joel), to the informatics (David, Iván, Guillermo, José, Quim, Miguel, Montse, Adriá and Ricard), and to the professors from UPC (Moisès, Mariana, Antonio, Pan, Enric, Fakher, Recasens, Arnaldos, Joan DePablo, Joseba Quevedo and José Luis Cortina), for sharing with me different stages during these four years of intensive work. I also thank UPC non-academic personnel for their attention, specially Irene, Francina, Cristina, Castells and Francesc, from the Secretariat, and Maribel, from the cleaning personnel.

I thank all Spanish people and specially catalans for welcome me with open arms. I have felt as "at home" since the first day. I also would like to mention the friendship with many people that I have the chance to meet, specially who have been or are my flatmates (Pablo, Juan, Saúl, Marcus and Rachid).

I thank professors and researchers from Universidad Tecnológica Nacional (Facultad Regional Rosario, Argentina, where I have obtained the B.Sc. degree in Chemical Engineering), for their collaboration. I would like to mention specially Ph.D. Marta Basualdo who animated me to initiate the doctorate at UPC and to all the components of the Group of Research Applied to Chemical Engineering -GIAIQ- (Carlos Ruiz, Nicolás Scenna, Sonia Benz, Alejandro Santacruz and Miguel Muñoz, among others).

I thank Fundación YPF (Argentina) for the financial support during my two first years of doctorate. Through the J. Estenssoro Grant I had the possibility of assisting to courses, participating in international conferences related to my thesis and exposing my works to the scientific and technologic community.

I thank UPC and Generalitat de Catalunya (TDOC Grant) for the financial support during my two last years of doctorate.

I thank the following companies for their collaboration in the development of the industrial applications introduced in this thesis: Compañía Azucarera Concepción S.A. (CACSA), Tucumán, Argentina; Complejo Agroindustrial Camilo Cienfuegos (CAICC), La Habana, Cuba; Instituto Cubano de Investigaciones de los Derivados de la Caña de Azúcar (ICIDCA), La Habana, Cuba; Estación Experimental Agroindustrial Obispo Colombres (EEAOC), Tucumán, Argentina; SOTEICA S.R.L. (Tecnología e Ideas), Buenos Aires, Argentina; Petroquímica La Plata - Repsol-YPF, La Plata, Argentina; AEA Technology- Hyprotech, Barcelona, Spain.

Finally, I thank all my family, who live here and who live in my native land, for their inestimable support. In spite of the distances I always feel them near me. In special, I would like to highlight the example of work and love that I have received from my parents, and the admiration I feel for my brothers and my sister. I also would like to mention the love that every day I receive from my girlfriend and that has helped me very much in my last stage of the thesis.

Diego Ruiz Massa

Agradecimientos *(Acknowledgements, in Spanish)*

Agradezco al Profesor Luis Puigjaner Corbella la dirección de esta tesis. Quiero destacar su empuje y apoyo permanentes. Su ejemplo de trabajo y dedicación así como sus llamadas de atención han sido muy enriquecedoras para mí.

Agradezco a José María Nougués por su guía y paciente atención durante el desarrollo de la tesis. Quiero remarcar su honestidad y conocimientos sobre control de procesos. Sus críticas y consejos han sido determinantes para la calidad del presente trabajo.

Quiero expresar mi gratitud a todos los doctorandos (Jordi C., Frank, Arndt, María A., Zuly, María S., Eduard, Raimon, Quique, Edwin, Saúl, Suly, Joaquín, Francisco, Joan, Rosario, Chouaib, Eloy, Vicky, Danelia, Dridi, Kamal, Meng, Javier y Rodolfo), a los estudiantes que han realizado el proyecto de fin de carrera (Ana, Jordi V., Marta, Carles V., Sandra, Hans, Sebastián, Carles L., Espen, Alex, Aprilla, Dirk y Joel), a los informáticos (David, Iván, Guillermo, José, Quim, Miguel, Montse, Adriá y Ricard) y a los profesores de la UPC (Moisès, Mariana, Antonio, Pan, Enric, Fakher, Recasens, Arnaldos, Joan DePablo, Joseba Quevedo y José Luis Cortina), quienes han compartido conmigo distintas etapas durante estos cuatro años de intensa labor. Agradezco también la atención del personal no docente de la UPC, especialmente a Irene, Francina, Cristina, Castells y Francesc, de Secretaría; y a Maribel, del personal de limpieza.

Agradezco a todos los españoles y en especial a los catalanes, por recibirme con los brazos abiertos. Me he sentido "como en casa" desde el primer día. También quiero destacar la amistad de muchas personas que tuve la dicha de conocer, especialmente quienes han sido o son mis compañeros de piso (Pablo, Juan, Saúl, Marcus y Rachid).

Agradezco la colaboración de profesores e investigadores de la Universidad Tecnológica Nacional, Facultad Regional Rosario, Argentina, de la que soy egresado como ingeniero químico. Quiero mencionar especialmente a la Dra. Ing. Qca. Marta Basualdo quien me animó en la realización del doctorado en la UPC y a todos los integrantes del Grupo de Investigación Aplicada a la Ingeniería Química (Carlos Ruiz, Nicolás Scenna, Sonia Benz, Alejandro Santacruz y Miguel Muñoz, entre otros).

Agradezco el apoyo financiero de la Fundación YPF (Argentina) durante mis dos primeros años de doctorado. A través de la beca J. Estenssoro tuve la posibilidad de hacer cursos, participar en congresos internacionales relacionados con mi tesis y exponer mis trabajos a la comunidad científica y tecnológica.

Agradezco el apoyo financiero de la UPC y de la Generalitat de Catalunya (beca TDOC) durante mis dos últimos años de doctorado.

Agradezco la colaboración de las siguientes empresas en el desarrollo de las aplicaciones industriales presentadas en esta tesis: Compañía Azucarera Concepción S.A. (CACSA), Tucumán, Argentina; Complejo Agroindustrial Camilo Cienfuegos (CAICC), La Habana, Cuba; Instituto Cubano de Investigaciones de los Derivados de la Caña de Azúcar (ICIDCA), La Habana, Cuba; Estación Experimental Agroindustrial Obispo Colombres (EEAOC), Tucumán, Argentina; SOTEICA S.R.L. (Tecnología e Ideas), Buenos Aires, Argentina; Petroquímica La Plata - Repsol-YPF, La Plata, Argentina; AEA Technology- Hyprotech, Barcelona, España.

Finalmente, agradezco el inestimable apoyo de toda mi familia, los que viven aquí y los que viven en mi tierra. A pesar de las distancias siempre los siento muy cerca. En especial, quiero resaltar el ejemplo de trabajo y amor que recibí de mis padres, y la admiración que siento por mis hermanos y mi hermana. También quiero mencionar el amor que día a día recibo de mi novia y que me ha ayudado mucho en mi última etapa de la tesis.

Diego Ruiz Massa

Contents

Figures Index

Tables index

Summary

The growing necessities of improvements in product quality and process productivity, plus the environmental and safety requirements, have motivated that the chemical plant designs and their control systems are very complex. This high complexity makes necessary the assistance to plant operators for fault diagnosis and the following decision-making in order to manage abnormal situations (Chapter 1, Introduction).

The pretended contribution of this thesis deals with the implementation of a fault diagnosis system in chemical plants integrated to the monitoring, management and control system (Chapter 2, Objectives).

First, the state of the art is presented (Chapter 3), describing the existing methods for fault diagnosis. Fault diagnosis methods can be classified in three main groups: historical based, knowledge-based and combination of both. The use of a wise combination where methods are adapted to each other in order to improve the advantages and to reduce the drawbacks seems to be the best option. Furthermore, these methods overcome the so called hybrid systems which consist in the implementation of some different methods (and then deciding among the different obtained responses) due to the difficulties related to the implementation of each individual methodology. Also, a special overview of the fault diagnosis methods for batch processes is presented. There exist very few methods due to the high complexity involved. Moreover, it has to be taken into account that the information provided by the fault diagnosis system will be utilised by other levels in the information system as the scheduling system.

The fault diagnosis system that is introduced in this thesis consists in a combination of a pattern recognition approach based on artificial neural networks and an inference

system based on fuzzy logic. Therefore, a brief overview about the concepts of neural network and fuzzy logic technologies is introduced (Chapter 4, Tools and techniques). A summary about Hazard and Operability analysis (HAZOP) and the signal pre-processing using wavelets is also included because such techniques are utilised in the design and implementation of the proposed fault diagnosis system.

The general structure of the fault diagnosis system is shown in Figure 5.3 (Chapter 5, Methodology). The fuzzy logic system complements the artificial neural network in a block oriented configuration. The inputs to the system are the direct or indirect measurements from the plant and the output consists in a signal for each fault (0: no fault; 1: fault).

The information needed to develop the fault diagnosis system includes the historical data, the hazard and operability study and the model of the chemical plant. First, the possible faults with high importance are defined, according to the historical data and the hazard analysis, taking into account the economic impact and the frequency of them. The artificial neural network is trained with historical data of faults occurred in the past, with the aim of recognising the respective patterns. In the case that the corresponding historical data are not available, for example due to the no occurrence of the fault, the patterns are obtained through simulation, using the plant model. The fuzzy logic system contains a set of if-then rules that can be of two types: those based in the process knowledge, by the hazard analysis or by the experience with simulation, and those based on the experience with the use of an artificial neural network, previously trained. Other novel aspect is the possibility of artificial neural network training by using signals features that are extracted by its pre-processing using wavelets. This alternative allows a higher fault diagnosis system performance in batch and complex continuous chemical plants.

In order to optimise the parameters of the components of the fault diagnosis system, a performance index is proposed, Equation (5.1) and Figures 5.7 and 5.8, which takes into account the time required for a correct diagnosis. The performance index is also utilised to compare the proposed fault diagnosis system against other methods.

The selection of the type of artificial neural network is not a critical aspect for the success of the system implementation. Nevertheless, in the performed applications the following neural networks have been preferred: a probabilistic neural network to classify fault patterns given by features extracted by signal pre-processing using wavelets, a radial basis function network for the case of training with variables' profiles directly, except for large amount of training data where the backpropagation network is

chosen, and auto -associative neural networks for the special case of fault diagnosis in sensors.

The on-line implementation of the fault diagnosis system is performed with the algorithm shown in Figure 5.18 and the subprograms shown in Figures 5.19 and 5.20.

The signals provided by the fault diagnosis system can be used by the scheduling system to update the schedule in the most effective way, by the control system to take automated control actions and by plant's operators as support for decision-making. The basis of the translation of the system output for its utilisation at other levels in the information system have been settled. The proposed strategy is based on the hazard and operability analysis.

The proposed system is the result of successive improvements, by working with different case studies. The academic scenarios correspond to a continuous chemical plant with a recycle stream and a batch reactor. The pilot plant scale cases correspond to scenarios built at UPC: a reactor gasifier, a fed-batch reactor and a multipurpose batch chemical plant. The industrial scenarios correspond to two sugar refineries and a sector of a petrochemical plant. A description of the case studies is done in Chapter 6 (Case studies).

Different aspects of the proposed methodology are illustrated by the implementation of the proposed fault diagnosis system in different case studies (Chapter 7, Results and discussion).

In the case of the chemical plant with recycle, the advantages of the proposed system with respect to a neural network or an expert system working alone, are shown. Such advantages include the speed of diagnosis, the exactness of multiple fault diagnosis and the no existence of cases of false diagnosis.

In the case of the batch reactor, the use of fault patterns based on the extracted features by signal pre-processing using wavelets for artificial neural network training is shown. Furthermore, the generation of if-then rules for the fuzzy logic system, based on the hazard analysis is illustrated, too. A comparison against a statistical method is done and a better performance of the proposed system is observed in relation to its capacity to isolate faults.

Implementation in the case of the gasifier, using real data, allows to show that the system has better performance than statistical techniques. The system implementation is also illustrated in the case of semi-batch and batch plants. The translation of the system output for its utilisation at other levels in information system is treated in these scenarios.

Implementation of the proposed fault diagnosis system has been tested and validated in three industrial cases (Chapter 8, Industrial applications). In the first case, one of the sugar cane refineries, results are shown for the application of the fault diagnosis system in the sector of Boiling and Crystallisation, where problems derived from the fluctuations in the steam supply and vacuum generation affect directly the quality product. A method is proposed to update the system with the participation of plants' operators (Table 8.3). In the second case, a sector of other sugar refinery, an application of the fault diagnosis system is presented with the additional difficulty of the lack of on-line measurements from the plant. A system has been given to the company to assist plant's operators for decision-making. Such a tool allows them diagnosing possible faults, after performing some observations and measurements from the plant in "off line" mode. Finally, the system implementation is shown in a sector of a petrochemical plant. The proposed fault diagnosis system is able to adequately diagnose fluctuations in the flowrates of hot oil to the reboilers of the distillation columns. Furthermore, it is possible to take advantage of some commercial softwares that are familiar to plant engineers.

The simplicity of the development and the flexible strategy of implementation of the proposed fault diagnosis system give a promising future to the presented technology (Chapter 9, Conclusions).

Resumen *(Summary, in Spanish)*

Las necesidades crecientes de mejorar la calidad de los productos y la rentabilidad de los procesos, sumado a las exigencias de protección del medio ambiente y la seguridad de operación, han motivado que los diseños de plantas químicas y sus sistemas de control sean muy complejos. Esta elevada complejidad hace necesaria la asistencia a los operadores en la diagnosis de fallos y la consiguiente toma de decisiones para gestionar las situaciones anormales (Capítulo 1, Introducción).

La contribución que se pretende con esta tesis se refiere a la implantación de un sistema de diagnosis de fallos en plantas químicas completas integrado al sistema de supervisión, gestión y control de la producción (Capítulo 2, Objetivos).

En primer lugar se presenta un resumen del estado del arte (Capítulo 3), describiendo los métodos existentes para la diagnosis de fallos. Los mismos se pueden clasificar en tres grupos principales: los basados en datos históricos, los basados en el conocimiento del proceso y combinaciones de ambos. El uso de una combinación inteligente en la que los métodos se adaptan uno al otro para mejorar las ventajas individuales y reducir los puntos débiles parece ser la mejor opción. Además, estos sistemas superan a los llamados sistemas híbridos que consisten en implementar varios métodos diferentes y luego decidir entre los distintos resultados obtenidos, debido a las dificultades de implementación de cada técnica individual. También se presenta una revisión especial de los métodos actuales para la diagnosis de fallos en plantas discontinuas. Existen muy pocos métodos para estos casos dada la elevada complejidad de implantación. Además, se debe tener en cuenta que la información que suministre el sistema de diagnosis será utilizado a otros niveles del sistema informático de la planta como el de programación de la producción.

El sistema de diagnosis de fallos que se presenta en esta tesis consiste en una combinación de un sistema de reconocimiento de patrones basado en redes neuronales artificiales y un sistema de inferencia basado en la lógica difusa. Se incluye entonces una breve revisión de los conceptos de la tecnología de redes neuronales y la lógica difusa (Capítulo 4, Herramientas y técnicas). Además, se hace un resumen sobre el análisis de riesgo y operabilidad (HAZOP) y el pre-procesamiento de señales usando *wavelets*, ya que estas técnicas se usan en el diseño e implementación del sistema de diagnosis que se propone.

La estructura general del sistema de diagnosis de fallos se muestra en la Figura 5.3 (Capítulo 5, Metodología). El sistema de lógica difusa complementa a una red neuronal en una combinación orientada a bloques. La entrada al sistema son las mediciones directas o indirectas de la planta y la salida consiste en una señal para cada fallo (0: no fallo; 1: fallo).

La información necesaria para desarrollar el sistema de diagnosis incluye los datos históricos, un análisis de riesgo y operabilidad y un modelo de la planta química. Primero se definen los fallos posibles con alta importancia, de acuerdo con los datos históricos y el análisis de riesgo, teniendo en cuenta el impacto económico y la frecuencia de los mismos. La red neuronal se entrena con datos históricos de fallos ocurridos en el pasado, con el objeto de reconocer los patrones respectivos. En el caso de que no se posean los datos históricos de alguno de los fallos, por ejemplo porque nunca hayan ocurrido, se obtienen los patrones mediante la simulación, usando el modelo de la planta. El sistema de lógica difusa contiene un conjunto de reglas si-entonces que pueden ser de dos tipos: las basadas en el conocimiento de la planta, mediante el análisis de riesgo o la experiencia con la simulación, y las basadas en la experiencia con el uso de la red neuronal, previamente entrenada. Otro aspecto novedoso es la posibilidad de entrenar la red neuronal con "características" extraídas de las variables medidas mediante su pre-procesamiento con *wavelets*. Esta variante permite obtener un alto rendimiento del sistema de diagnosis en plantas químicas discontinuas y continuas complejas.

Para optimizar los parámetros de los componentes del sistema de diagnosis se propone un índice de rendimiento, Ecuación (5.1) y Figuras 5.7 y 5.8, que tiene en cuenta el tiempo requerido para una diagnosis correcta. Además, se utiliza el índice de rendimiento para comparar el sistema propuesto con otros métodos.

La selección del tipo de red neuronal no es un aspecto determinante del éxito del sistema. No obstante, en las aplicaciones realizadas se han preferido las siguientes redes neuronales: una red neuronal probabilística para clasificar patrones de fallos dados por "características" extraídas mediante el pre-procesamiento con *wavelets*, una red neuronal con funciones en base radial para el caso del entrenamiento con perfiles de las variables, salvo que los datos de entrenamineto sean cuantiosos, donde se ha elegido una red "backpropagation", y una red auto-asociativa para el caso especial de la diagnosis de fallos en sensores.

La implementación en línea del sistema de diagnosis se realiza con el algoritmo esquematizado en las Figura 5.18 y los subprogramas mostrados en las Figuras 5.19 y 5.20.

Las señales dadas por el sistema de diagnosis pueden ser usadas por el sistema de programación de la producción para actualizar el plan de la manera más efectiva, por el sistema de control para actuar en forma automática y por los operadores de planta como soporte para la toma de decisiones. Se han sentado las bases para la traducción de la salida del sistema de diagnosis para su utilización por los demás niveles del soporte informático. Se usa una estrategia basada en el análisis de riesgo y operabilidad de la planta.

El sistema propuesto es consecuencia de sucesivas mejoras, al trabajar con diferentes casos de estudio. Los escenarios académicos corresponden a una planta química continua con una corriente de reciclo y un reactor discontinuo. Los casos a escala de planta piloto corresponden a escenarios construidos en la UPC: un reactor de gasificación, un reactor semicontinuo y una planta discontinua multipropósito. Los escenarios industriales corresponden a dos refinerías de azúcar y a un sector de una planta petroquímica. Una descripción de los casos de estudio se realiza en el Capítulo 6 (Casos de estudio).

Distintos aspectos de la metodología propuesta se ilustran mediante la implementación del sistema de diagnosis propuesto en los distintos casos de estudios (Capítulo 7, Resultados y discusión).

En el caso de la planta con reciclo se muestran las ventajas del sistema propuesto respecto a una red neuronal y a un sistema experto que trabajan por separado. Tales ventajas incluyen la velocidad de diagnosis, la exactitud en la diagnosis de fallos múltiples y la inexistencia de casos de diagnosis falsa. También se muestra una comparación de distintos tipos de redes neuronales y la optimización de sus parámetros, usando el índice de rendimiento propuesto. Además se muestra la aplicación de redes neuronales autoasociativas para diagnosicar fallos en sensores.

En el caso del reactor discontinuo, se muestra el uso de patrones de fallos basados en las "características", extraídas mediante el pre-procesamiento de las mediciones con *wavelets*, para entrenar la red neuronal. Además, se ilustra la generación de reglas si-entonces del sistema de lógica difusa, en base al análisis de riesgo. Se realiza una comparación con un método estadístico, observándose un mejor rendimiento del sistema propuesto en relación con su capacidad de aislar los fallos.

La implementación en el caso del reactor de gasificación, usando datos reales, permite mostrar que el sistema tiene mejor rendimiento que los métodos estadísticos. También se ilustra la implantación del sistema en casos de plantas semicontinuas y discontinuas. La traducción de la salida del sistema para su utilización en otro niveles del sistema informático de la planta se ilustra en estos escenarios.

La implementación del sistema de diagnosis propuesto se ha verificado y validado en tres casos industriales (Capítulo 8, Aplicaciones industriales). En el primer caso, una refinería de azúcar de caña, se muestra el resultado de la aplicación del sistema de diagnosis en el sector de "tachos de cocimiento", donde los problemas derivados de las fluctuaciones en el suministro de vapor y en la generación de vacío afectan directamente la calidad del producto final. Se propone un método para actualizar el sistema con la participación de los operadores de planta (Tabla 8.3). En el segundo caso, un sector de otra refinería, se presenta la aplicación del sistema con la dificultad añadida de no contar con mediciones en línea de la planta. Se ha provisto a la compañía de un sistema para la asistencia de los operadores en la toma de decisiones. Tal herramienta les permite diagnosticar posibles fallos, luego de hacer algunas observaciones y mediciones "fuera de línea" de la planta. Finalmente, se muestra la implementación del sistema en un sector de una planta petroquímica. El sistema de diagnosis propuesto es capaz de diagnosticar adecuadamente las fluctuaciones de caudal de fluido calefactor de los re-hervidores de las columnas de destilación. Además, se pueden aprovechar programas comerciales que resultan familiares a los ingenieros de planta.

La simplicidad del desarrollo y la flexible estrategia de implementación del sistema propuesto auguran un futuro promisorio a la tecnología presentada (Capítulo 9, Conclusiones).