CONTROL TOLERANTE A FALLOS (PARTE I): FUNDAMENTOS Y DIAGNÓSTICO DE FALLOS [1]

Vicenç Puig, Joseba Quevedo, Teresa Escobet, Bernardo Morcego, Carlos Ocampo

Departament Enginyeria de Sistemes, Automàtica i Informàtica Industrial (ESAII) - Campus de Terrassa
Universitat Politècnica de Catalunya (UPC)
Rambla Sant Nebridi, 10. 08222 Terrassa (Spain)
e-mail: *{vicenc.puig, joseba.quevedo, teresa.escobet, bernardo.morcego, carlos.ocampo}@.upc.es*

Resumen: En este artículo se presentan los fundamentos del control tolerante a fallos, se introduce el análisis estructural como una herramienta útil para el análisis y el diseño tanto del sistema de diagnostico como de los mecanismos de tolerancia a fallos, finalizando con una revisión de los métodos de diagnóstico existentes. Algunas de las técnicas de diagnostico presentadas se aplican en un proceso real, basado en el sistema de control de la red de alcantarillado de Barcelona. En un segundo artículo se presentarán los mecanismos de tolerancia que se pueden activar una vez se ha diagnosticado el fallo y serán aplicados sobre el mismo ejemplo.

Keywords: Control tolerante, diagnóstico de fallos, detección de fallos, acomodación al fallo, reconfiguración del controlador.

1. INTRODUCCIÓN

Con el incremento del grado de dependencia de la sociedad moderna de los sistemas (automóviles, aviones, trenes, etc.) y procesos tecnológicos complejos (redes de distribución y producción de energía, agua, etc.), su disponibilidad y correcto funcionamiento se han convertido en una cuestión estratégica. Su incorrecto funcionamiento puede provocar pérdidas económicas, peligro para los operadores, inconvenientes para los usuarios, etc. Además, la automatización de los mismos mediante lazos de control automático, si bien ha permitido liberar a los operadores humanos de su control y operación manual, no los ha inmunizado frente a los ***fallos***. Se entiende por fallo todo cambio en el comportamiento de alguno de los componentes del sistema (desviación no permitida de alguna de sus propiedades o parámetros característicos) de manera que éste ya no puede satisfacer la función para la cual ha sido diseñado (Blanke, 2000).

Los sistemas de control automático son, pues, susceptibles a los fallos pudiendo verse amplificados por el lazo de control llegando a provocar su mal funcionamiento. Además, los lazos de control pueden ocultar los fallos evitando ser observados hasta alcanzar un grado tal que produzcan una ***avería*** irreparable que obligue a detener del sistema o proceso. Por ello existe una creciente necesidad e interés en desarrollar sistemas de control que

[1] Trabajo subvencionado por la CICYT del Ministerio de Ciencia y Tecnología Español (DPI2002-0350) y por la DGR de la Generalitat de Catalunya (grupo SAC 2001/SGR/00236).

puedan operar de forma aceptable incluso después de la aparición de un fallo y que sean capaces de parar el proceso antes de que se originen daños irreparables en el mismo. A este tipo de sistemas de control se les denomina ***tolerantes a fallos***. La tolerancia a fallos se entiende pues como la capacidad de un sistema de control para mantener los ***objetivos de control*** a pesar de la aparición de un fallo, admitiéndose una cierta ***degradación*** de sus ***prestaciones***.

En la bibliografía se consideran dos tipos de control tolerante a fallos: el ***pasivo*** y el ***activo***. El primero de ellos, utiliza la propiedad que tienen los sistemas realimentados de hacer frente a perturbaciones, cambios en la dinámica del sistema e incluso fallos en el mismo. Un cambio inesperado en el sistema crea un efecto sobre el mismo que se transmite al sistema de control que a su vez trata de compensarlo de forma más o menos rápida. En este sentido, el control tolerante pasivo consiste en un diseño robusto del sistema de control realimentado para hacerlo inmune a determinados fallos (Patton, 1997). Sin embargo, la teoría de control robusto muestra que sólo existen controladores robustos para una clase reducida de cambios en la dinámica del sistema provocados por los fallos. Además, un controlador robusto funciona de forma subóptima para la planta nominal puesto que sus parámetros se han obtenido mediante un compromiso entre prestaciones y robustez para toda la familia de plantas considerada, incluyendo los posibles fallos. Por otro lado, el ***control tolerante activo*** consiste en el ***diagnóstico*** en línea del fallo, es decir, en determinar el componente averiado, el tipo de avería, su tamaño e instante de aparición y, a partir de dicha información, activar algún mecanismo de acomodación del mismo o de reconfiguración del control o incluso dependiendo de la gravedad la parada el sistema.

Este enfoque, que será el utilizado en un artículo posterior para el diseño del sistema de control tolerante a fallos, exige disponer de un sistema de diagnóstico de fallos que, en tiempo real, pueda dar información a un sistema supervisor para que active algún mecanismo de acción correctora (Blanke, 2003).

La estructura del artículo es la siguiente: en la *Sección 2* se introducirá el problema de control tolerante. En la *Sección 3* se introducirá el problema de diagnóstico de fallos asociado al problema de control tolerante. En la *Sección 4* se presentará una revisión del diagnóstico de fallos mediante modelos. En la *Sección 5*, se presentará una aplicación de algunas de las técnicas de diagnóstico presentadas en este artículo aplicadas a un sistema de control real: el sistema de control de la red de alcantarillado de Barcelona. En la *Sección 6* se presentarán las conclusiones más relevantes sobre este primer artículo de control tolerante a fallos y finalmente, en un apéndice anexo se presentará un glosario de la terminología más utilizada habitualmente en este campo.

2. EL PROBLEMA DE CONTROL TOLERANTE

2.1 Definición del problema

Desde el punto de vista de teoría de sistemas, el control tolerante a fallos trata de la interacción entre un sistema dado (proceso) y su control. El término ***control*** debe considerarse en este caso en un sentido global, involucrando tanto la típica ley de control realimentado como aspectos de toma de decisiones que determinan la configuración del control. De cara a ilustrar mejor el alcance del problema de control tolerante, se introducirán algunas de las técnicas de control bien conocidas en el área.

Un problema de control estándar tiene como objetivo diseñar una ley de control ϑ, partiendo de un conjunto de objetivos O y un conjunto de restricciones, C, que describen el comportamiento dinámico del sistema (modelo matemático). El conjunto de restricciones, C, están determinadas por la estructura del modelo matemático S y sus parámetros, θ.

Definición 1 (Problema de control estándar)
ϑ = *solucionar* $\{O\ ,\ C(\theta)\}$.

En realidad, difícilmente un modelo matemático representa adecuadamente el comportamiento del sistema. Debido a problemas de perturbaciones, ruido en las medidas, dinámicas no modeladas, parámetros variantes en el tiempo e inciertos, etc., la solución al problema de control para conseguir los objetivos O, cuando se considera la incertidumbre en el modelo, se puede plantear suponiendo una estructura fija S para el mismo, pero con parámetros θ desconocidos y pertenecientes a un conjunto de parámetros Θ, aplicando técnicas de control robusto o adaptativo.

El control robusto trata de diseñar una ley de control ϑ que cumpla todo el conjunto de objetivos O teniendo cuenta todo el conjunto de parámetros Θ.

Definición 2 (Problema de control robusto)
ϑ = *solucionar* $\{O\ ,\ C(\Theta)\}$, *donde* Θ *es todo el conjunto posible de parámetros.*

El control adaptativo soluciona el problema estimando a cada iteración el valor de los parámetros θ del conjunto de posibles parámetros del sistema Θ.

Definición 3 (Problema de control adaptativo) *ϑ = solucionar $\{O\ ,\ C(\hat{\theta})\}$, donde $\hat{\theta}\ \epsilon\ \Theta$ se estima a cada iteración.*

La aparición de fallos en el sistema puede ocasionar tanto modificaciones en las restricciones C como cambios en los parámetros θ haciendo que el problema de calcular la ley de control ϑ no tenga solución a no ser que se modifique el conjunto de objetivos O.

Definición 4 (Problema de control tolerante a fallos) *ϑ = solucionar $\{O_f\ ,\hat{C}_f(\hat{\theta}_f)\}$, donde O_f representa el conjunto de objetivos y $\hat{C}_f(\hat{\theta}_f)$ representa la estimación de las restricciones y parámetros del sistema según su estado de funcionamiento (normal o en fallo).*

Existen, tal como se ha comentado en la introducción, dos enfoques para resolver el problema de control tolerante:

- ***Pasivo:*** basado en diseñar una única ley de control que sea capaz de alcanzar sus objetivos tanto en situación de funcionamiento normal como en fallo.
- ***Activo:*** basado en diseñar una ley de control diferente en función del estado del sistema (normal o en fallo) a partir de la estimación de las restricciones y parámetros del sistema $\hat{C}_f(\hat{\theta}_f)$ proporcionados por el diagnosticador.

El problema de control tolerante mediante el enfoque activo puede resolverse de dos formas: ya sea mediante la ***acomodación al fallo***, o bien, mediante la ***reconfiguración***. La acomodación al fallo consiste en resolver el problema manteniendo la estructura del controlador y modificando solamente los parámetros. Por otro lado, la reconfiguración consiste en cambiar las entradas y salidas del controlador así como reajustar la ley de control. La utilización de una u otra dependerá de los objetivos de control planteados y del fallo presente en el sistema.

Cabe mencionar que en el caso de control tolerante, la función objetivo O queda modificada, formulándose unas condiciones de operación óptimas (planta sin fallo) más unas condiciones de operación degradadas (planta con fallo).

Asumiendo que el funcionamiento del sistema realimentado puede ser descrito mediante las variables x_1 y x_2, la Figura 1 ilustra las diferentes regiones que deben ser consideradas en el diseño de un control tolerante a fallos. La región de comportamiento deseado es aquella en la que el sistema debe operar normalmente cumpliendo su función. El controlador se encarga de mantener sistema en dicha región a pesar de las perturbaciones e incertidumbre en el modelo utilizado para el diseño del lazo de control, e incluso en caso de pequeños fallos, aunque esa no es su función principal. La región de comportamiento degradado se corresponde con la región de funcionamiento a la que el sistema se desplaza después de la aparición de un fallo. En esta situación el controlador tolerante diseñado para reaccionar frente al mismo deberá de activar las acciones de recuperación preestablecidas de cara a evitar una mayor degradación que desplace al sistema hacia la región de comportamiento inadmisible e incluso de peligro.

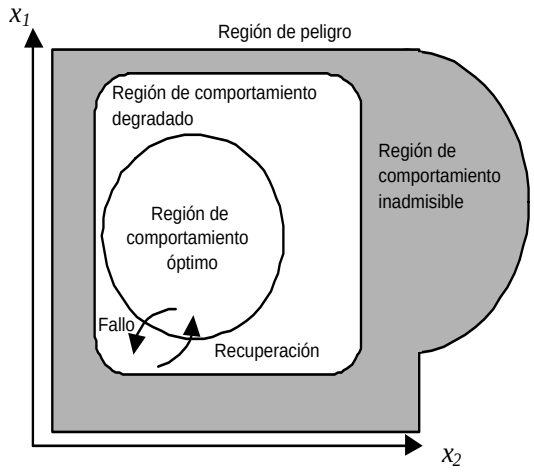

Figura 1. Regiones de Comportamiento.

2.2 Metodología de diseño

En la Figura 2 se presentan las etapas de una metodología sistemática para el diseño de sistemas de control tolerante siguiendo la propuesta por (Blanke, 2000). Las etapas de esta metodología se enumeran a continuación:

1. ***Análisis del Sistema*** a dos niveles: a nivel de componentes mediante un análisis de propagación de fallos a través de todos los subsistemas más relevantes, así como una evaluación de la severidad de los mismos y a nivel de estructura de cara a analizar la redundancia presente en el sistema que ayudará en el diseño del sistema de diagnóstico y reposición.
2. ***Diseño del Sistema de Diagnóstico*** a partir del análisis estructural y teniendo en cuenta las medidas disponibles y los fallos que se desean diagnosticar. En el caso de que no se puedan diagnosticar todos los fallos que se deseen, se deberá modificar la instrumentación disponible hasta conseguirlo. El sistema de diagnóstico de fallos deberá no sólo detectar y aislar los fallos sino también estimar su tamaño (cuantificación).
3. ***Diseño de los Mecanismos de Tolerancia*** para cada uno de los fallos considerados

según se trate de fallos en sensores, actuadores y/o planta.

4. ***Diseño del Supervisor*** a partir de la información acerca de los fallos proporcionada por el sistema de diagnóstico, el supervisor deberá activar los mecanismos de tolerancia que se han diseñado para cada uno de ellos.
5. ***Aplicación y test*** en simulación y sobre el sistema real.

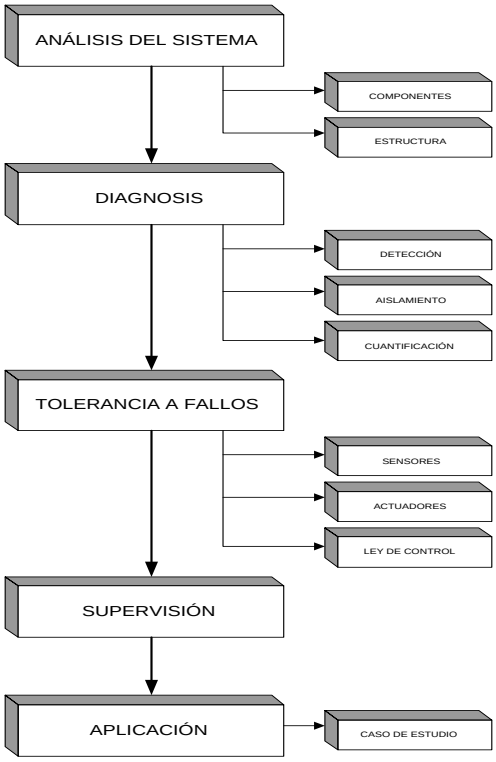

Figura 2. Metodología para el diseño de un sistema de control tolerante a fallos.

En los apartados siguientes se detallarán las dos primeras etapas de la metodología propuesta: análisis estructural y diagnosis, mientras que en un segundo artículo se tratan el resto de etapas.

3. ANÁLISIS DEL SISTEMA

Las etapas de diseño del sistema de diagnóstico y de los mecanismos de tolerancia a fallo precisan realizar un análisis del sistema:

- A nivel de componentes y sus interconexiones para realizar un análisis de propagación de fallos. El resultado de dicho análisis permitirá identificar una lista de componentes junto con sus modos de fallo que deberán de ser detectados y tratados mediante mecanismos de tolerancia a fallos.
- A nivel de la estructura del sistema con el objetivo de determinar el nivel de redundancia existente en el sistema de cara al aislamiento de los fallos críticos detectados en el análisis a nivel de componentes, así como de cara a la aplicación de mecanismos de tolerancia a fallos.

En ambos casos se analiza el sistema partiendo del nivel más bajo, utilizando el concepto de ***componente***, entendiendo como componente la unidad física considerada como indivisible, por ejemplo, un sensor, una bomba, una tubería, el controlador, etc.

3.1 Técnicas basadas en el análisis de propagación de fallos

El ***Análisis de Propagación de Fallos*** (FPA), propuesto por Blanke (1996), tiene como objetivo estudiar cuál es el efecto final de un fallo en cada uno de los componentes del sistema de control para poder posteriormente diseñar mecanismos de tolerancia que eviten su efecto. Si como resultado del FPA se encuentran fallos en determinados componentes que pueden ser críticos, éstos pasarán a constituir la lista de fallos a detectar, aislar y tratar mediante mecanismos de tolerancia.

El método de FPA tiene como punto de partida la técnica de ***modos de fallos y análisis de efectos*** (failure mode and effect analysis, FMEA) de componentes (Herrin, 1981). Se trata de un estándar comúnmente aceptado en la industria que permite analizar la propagación de los efectos de los fallos desde los componentes hacia el sistema. El método FMEA parte de un listado para cada componente de los diferentes ***modos de fallo*** y sus efectos. El resultado de dicho método es analizar qué efecto produciría sobre el sistema y su entorno la aparición de un modo de fallo en un determinado componente.

El análisis de la propagación de fallos se puede formalizar de forma adecuada utilizando métodos matriciales. Por lo tanto, la propagación de fallos entonces se puede ver como una relación binaria entre los fallos y sus efectos para cada componente o conjunto de componentes (Blanke, 2001).

Definición 5 (Matriz de propagación de fallos) Dada una relación binaria M, M: $F \times \varepsilon \rightarrow \{0,1\}$ del conjunto de fallos presentes en un componente $f_c \in F$ y el conjunto de sus efectos $e_c \in \varepsilon$, la matriz de propagación de fallos se define de la siguiente manera:

$$m_{ij} = \begin{cases} 1 & si\ f_{cj} = 1 \Rightarrow e_{ci} = 1 \\ 0 & en\ caso\ contrario \end{cases}$$

Utilizando la matriz de propagación de fallos M, la propagación de los efectos de los fallos en el i-ésimo componente se pueden expresar como:

$$e_{ci} \leftarrow M_i^f \otimes f_{ci} \qquad (1)$$

donde el operador booleano $\otimes$ realiza la siguiente operación:

$$e_{ci} \leftarrow (m_{i1} \wedge f_{c1}) \vee \cdots \vee (m_{in} \wedge f_{cn}) \qquad (2)$$

Cuando los efectos se propagan desde otros componentes se obtiene en el nivel i-ésimo:

$$e_{ci} \leftarrow M_i^f \otimes \begin{bmatrix} f_{ci} \\ e_{c(i-1)} \end{bmatrix} \qquad (3)$$

La descripción del efecto final de un determinado fallos en el sistema se obtiene mediante la interconexión de la descripción de los efectos en los componentes a través de la composición del operador $\otimes$. La evaluación de la severidad del efecto final de cada uno de los fallos permitirá identificar aquellos que deberán de ser detectados, aislados y tratados mediante los mecanismos de tolerancia.

3.2 Análisis estructural

El ***análisis estructural*** es el análisis de las características estructurales S del modelo, es decir, características que son independientes del valor de los parámetros. Es una forma de representar las relaciones existentes entre las variables y los parámetros resultantes del modelo y son independientes de la forma exacta bajo la cual se va a expresar dicho modelo (cuantitativo o cualitativo). A pesar de su simplicidad, dicho análisis permitirá obtener información muy útil de cara al diseño del sistema de diagnóstico así como de los mecanismos de tolerancia, puesto que permite, entre otros:

- identificar aquellos componentes que se pueden monitorizar,
- obtener relaciones de redundancia analítica que permitirán la detección y aislamiento de fallos, e
- identificar posibles mecanismos de tolerancia.

La estructura del sistema físico puede obtenerse a partir del modelo de comportamiento (normal) $S = (R, V)$, definido mediante un conjunto de m relaciones R en la que intervienen un conjunto de n variables V. En un modelo orientado a componentes, estas relaciones, denominadas ***relaciones primarias***, representan el comportamiento de cada uno de los componentes del sistema, mientras que el sistema físico completo constará de todo el conjunto de componentes acoplados entre sí.

El conjunto de variables V puede descomponerse en $V = X \cup K$, donde K es el subconjunto de variables observadas (medidas) y X es el subconjunto de variables no conocidas. El subconjunto de variables observadas contiene al subconjunto de variables de control U y al subconjunto de variables medidas Y, $K = Y \cup U$.

La estructura del modelo se puede representar mediante una matriz denominada ***matriz estructural*** (Staroswiecki, 1989).

Definición 6 (Matriz Estructural)
Se define como matriz estructural (SM) a la matriz cuyas filas se corresponden con las relaciones del modelo y las columnas se corresponden con las variables del modelo. Cada uno de los elementos de dicha matriz m_{ij} es '1' si y sólo si la variable de la columna j está contenida en la relación de la fila i, y '0' en caso contrario.

El análisis estructural propuesto por Cassar y Staroswiecki (1997) se basa en determinar el ***perfect matching*** (PM) entre variables X y ecuaciones R haciendo uso de un grafo bipartito $G = (R \cup X, A)$, siendo A el conjunto de arcos orientados entre variables y ecuaciones. De forma intuitiva, este emparejamiento es una asignación causal que asocia algunas variables del sistema con sus ecuaciones a partir de las cuales pueden ser calculadas. Las variables que no forman parte de este emparejamiento no pueden ser calculadas. Por otro lado, aquellas que pueden ser emparejadas de diferentes maneras presentan redundancia. Dicha redundancia será de utilidad para la detección de fallos, así como para el diseño de mecanismos de tolerancia. La determinación del PM se realiza utilizando el método de resolución gráfico denominado ***resolution process graph*** (RPG) (Cassar, 1997) teniendo en cuenta las restricciones causales asociadas a cada relación o ecuación, p.e., algunas relaciones puede ser que no sean invertibles o son invertibles en determinadas condiciones. En general no hay una solución única para PM. Cuando el número de ecuaciones R es mayor que el número de variables no conocidas X, algunas de las relaciones no se ven implicadas en el PM. Estas relaciones aparecen como un nodo sin utilizar en el RPG.

Definición 7 (Relación redundante)
Una relación redundante (RR) es una relación que no se ve involucrada en el PM en el grafo bipartito G=$(R \cup X, A)$.

Las RRs no son necesarias para determinar las variables no conocidas. Cassar y Staroswiecki (1997) muestran como este grafo puede ser utilizado para deducir las RRs. Cada RR origina una ***relación de redundancia analítica*** (ARR) cuando las variables no conocidas involucradas en RR son reemplazadas por su expresión formal siguiendo las trayectorias definidas en el RPG. Estas trayectorias remontan hacia atrás hasta conseguir variables observadas. Las $ARRs$ sólo contienen variables observadas y pueden ser evaluada a partir de las observaciones (Cordier, 2002).

Ejemplo 1 (Obtención de ARR)
Consideremos un ejemplo concreto en que RR_1 y RR_2 son expresiones del tipo:

$$\begin{aligned} RR_1(y_1, x) &= 0 \\ RR_2(y_2, x) &= 0 \end{aligned} \tag{4}$$

donde tenemos como variables medidas (y_1, y_2) $\in$ K y no medida $x \in X$.

Es posible obtener una nueva relación de redundancia, ***relación redundante deducida****, entre y_1 y y_2 ya que la variable no conocida x aparece en ambas relaciones. Por ejemplo, considerando que RR es invertible y combinando RR_1 con RR_2 se tiene:*

$$RR_2(y_2, RR_1^{-1}(y_1)) = 0 \tag{5}$$

En este caso la nueva relación es una ARR ya que en ella sólo aparecen variables medidas.

3.3 Análisis estructural extendido a componentes

Se puede generalizar el modelo de comportamiento de un sistema físico presentado en la *Sección 3.2* como *S=(R, V, SUPP)*, donde *SUPP* es conjunto de n_S ***soportes*** del sistema, entendiéndose por soporte de una relación primaria al componente asignado (Cordier, 2002). Cada *ARR* deducida del análisis estructural tendrá asignada un conjunto de componentes soporte, *supp(ARR)*. El soporte está constituido por un conjunto de componentes y un conjunto de sensores.

Definición 8 (Matriz estructural extendida)
La matriz estructural extendida (SM_e) se construye a partir de la matriz estructural SM asociando a cada relación del modelo (fila) el componente asociado (o soporte).

La matriz estructural extendida tiene por objetivo hacer un diagnóstico del sistema basado en componentes como se verá en la *Sección 4.2*. En la bibliografía se proponen algoritmos que permiten determinar las *RR*, junto con su soporte y las condiciones de activación, ya que un componente puede tener asignadas varias relaciones primarias en función de unas condiciones externas (Travé-Massuyes, 2003), (Pulido, 2002).

3.4 Tolerancia a fallos con respecto el análisis estructural

Frente a un fallo en un componente, el análisis estructural se ve modificado ya que una variable puede pasar de conocida a no conocida o algunas de las relaciones primarias pueden dejar de ser válidas suprimiendo o no las variables asociadas.

Definición 9 (Grado de redundancia)
El grado de redundancia de una variable medida y_i con respecto a un fallo en el componente sensor, f_i, vendrá dado por el número de RR que puedan estimar su comportamiento y que no estén afectadas por dicho fallo. Similarmente, el grado de redundancia de control de una variable controlada u_i con respecto a un fallo dado es el número de RR que permiten calcular u_i y no estén afectadas por dicho fallo.

Obviamente, para que el sistema sea tolerante y detectable al mismo tiempo una vez detectado un fallo, el grado de redundancia debe ser superior o igual a 2. La Definición 9 se puede extender de forma directa al caso de fallos múltiples.

4. EL PROBLEMA DE DIAGNÓSTICO

La primera tarea a realizar en un sistema de control tolerante activo consiste en el diagnóstico del fallo en tiempo real, llegando no sólo a su detección y aislamiento sino también a la estimación de su magnitud. Por lo tanto, el diagnóstico de fallo se puede a su vez dividir en tres etapas según su profundidad:

- **detección del fallo**: decisión de si existe o no un fallo así como la determinación de su instante de aparición.
- **aislamiento del fallo**: localización del componente en el cual se ha producido el fallo.
- **identificación y estimación del fallo**: identificación del modo de fallo y estimación de su magnitud.

Definición 10 (Problema de diagnóstico)
A partir de una secuencia de entradas U y salidas Y obtenidas a partir de los sensores instalados en el proceso, determinar la presencia de un fallo ***f*** *del conjunto de fallos posibles* ***F****.*

En la literatura existen tres grandes grupos de técnicas de diagnóstico de fallos según estén basadas en:

- el análisis de señales,
- modelos
- técnicas basadas en el conocimiento.

Puesto que en control tolerante se precisa llegar al nivel más profundo de las etapas de diagnóstico que se corresponde con la estimación de la magnitud del fallo, las técnicas de ***diagnóstico basadas en modelos*** (MBD) son las preferidas puesto que proporcionan de forma natural herramientas para realizar dicha estimación.

4.1 Detección de fallos basada en modelos

En los años 70, tal como aparece en Frank (2000), se inicia el diagnóstico basado en modelos como una aplicación de la teoría de observadores utilizados en el área de control automático. Durante los siguientes 25 años la comunidad de control automático, conocida como FDI, ha realizado numerosas contribuciones en este campo. Paralelamente, en los últimos 15 años ha habido un crecimiento importante de aportaciones procedentes de ciencias de la computación y de la comunidad de Inteligencia Artificial (esta comunidad utiliza las siglas DX para referirse al diagnóstico basado en modelos). La Figura 3 muestra resumidamente la evolución temporal de algunas de las técnicas utilizadas por las tres comunidades.

Definición 11 (Problema de detección)
A partir de una secuencia de entradas U y salidas Y obtenidas de los sensores instalados en el proceso a monitorizar, se trata de verificar la ***consistencia*** *con el comportamiento modelado:*

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t), \theta) \\ \mathbf{y}(t) &= \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t), \theta) \end{aligned} \tag{6}$$

donde:
$\boldsymbol{x} \in \Re^{nx}$,$\boldsymbol{u} \in \Re^{nu}$ *y* $\boldsymbol{y} \in \Re^{ny}$ *son los vectores de estado, entrada y salida de dimensión nx, nu and ny, respectivamente;*

$\boldsymbol{g}$ *y* $\boldsymbol{h}$ *son las funciones de espacio de estado y medida respectivamente;*

θ *es el vector de parámetros de dimensión p.*

La detección de una inconsistencia es indicativa de la presencia de un fallo.

4.1.1. Detección de fallos basada en modelos cuantitativos. En el caso de utilizar modelos cuantitativos, una forma de verificar la consistencia entre el modelo y las medidas de las entradas/salidas es generar, a partir de las mismas ($\boldsymbol{u}$,$\boldsymbol{y}$) y del modelo, una estimación de las salidas $\hat{\mathbf{y}}$. La consistencia entre el sistema real y el modelado se evalúa a cada instante de tiempo mediante la diferencia

$$\mathbf{r}(t) = \mathbf{y}(t) - \hat{\mathbf{y}}(t) \tag{7}$$

conocida como ***residuo***, o bien, a partir de relación más compleja que involucra a las entradas $\boldsymbol{u}(k)$ y salidas medidas $\boldsymbol{y}(k)$ así como los parámetros del sistema denominada ***relación de redundancia analítica***

$$\phi(t) = \mathbf{f}(\mathbf{u}(t), \mathbf{y}(t), \theta) \tag{8}$$

Las técnicas más utilizadas para generar residuos mediante modelos analíticos son:

- Ecuaciones de paridad (Gertler, 1991)
- Observadores (Chen, 1999).

mientras que las técnicas de generación de relaciones de redundancia analítica más conocidas son las basadas en:

- Espacio de paridad (Chow, 1984)
- Análisis estructural (Staroswiecki, 2000).

Últimamente se han llegado a obtener equivalencias entre las diferentes técnicas de generación de residuos y de relaciones de redundancia analítica (Gertler, 1991), (Ding, 1999).

Idealmente, en ausencia de fallos, el residuo debería ser siempre nulo, mientras que en presencia de un fallo, sea cual sea el tipo de fallo, el residuo debería ser diferente de cero. Por tanto, en condiciones ideales, el test de detección de fallo consistiría en comprobar si el residuo $\boldsymbol{r}(t)$ es nulo o no. Sin embargo, la presencia habitual de perturbaciones, ruido y errores de modelización provoca también que los residuos no sean nulos, interfiriendo en la detección de posibles fallos. Por ello, hace falta diseñar residuos que estén afectados lo menos posible por las perturbaciones, ruido y dinámica no modelada, es decir, ***robustos*** frente a dichos efectos. La robustez en detección de fallos se puede alcanzar en la generación de los residuos (***robustez activa***) o en la fase de toma de decisión (***robustez pasiva***).

El empuje de la teoría del control robusto en los años 80 y 90 impulsó la búsqueda de técnicas activas de diagnóstico robusto. Se han de destacar aportaciones como la utilización del conocimiento estadístico de las perturbaciones (Basseville, 1993); la utilización de funciones de transferencia para el cálculo de los residuos que permite un desacoplo respecto de las perturbaciones (Gertler, 1998), (Chen, 1999) o la utilización de unos índices de prestaciones para generar los residuos como propone Frank (1991), entre otros.

La robustez pasiva consiste en tener en cuenta la incertidumbre en la generación de los umbrales que han de superar los residuos para considerar la existencia de fallos. Estos umbrales por lo general dependen de las condiciones de operación del proceso controlado y si está en estado transitorio o permanente, por lo que es posible determinar empíricamente o analíticamente unos umbrales adaptativos que permitan realizar una detección robusta de fallos. Las técnicas de umbrales adaptativos fueron propuestas inicialmente por Clark (1989), que sugirió una relación empírica entre el punto de operación y el correspondiente umbral para la detección. Otra técnica es la propuesta por Emami-Naemi (1988) que desarrolló una relación teórica, basada en H_∞, entre el punto de operación, la incertidumbre del modelo y el umbral de detección. Frank (1991) también utiliza

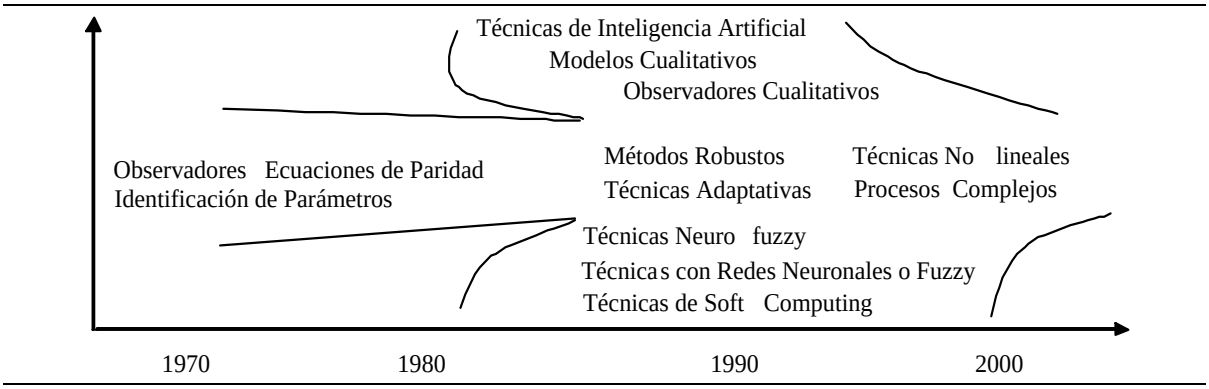

Figura 3. Evolución histórica del diagnóstico basado en modelos.

técnicas basadas en H_∞ para obtener umbrales adaptativos. Otra técnica basada en optimización dinámica asumiendo incertidumbre en los parámetros del modelo fue propuesta por Horak (1988) y varias propuestas utilizando modelos intervalares en simulación, predicción y observación han sido descritas en Puig (2002a).

Las técnicas pasivas de detección de fallos tienen la ventaja de que permiten abordar el problema aún cuando existan incertidumbres en los parámetros del modelo y sin tener que realizar ninguna simplificación como en el caso de las técnicas activas. Las técnicas pasivas no evitan el efecto de la incertidumbre del modelo sino que lo propagan hasta conseguir en todo momento unos valores que delimitan los residuos del proceso sin fallo. Por tanto, mientras los residuos satisfagan su pertenencia al conjunto de valores posibles no se indicará la existencia de fallo alguno, ya que a este nivel los valores de los residuos pueden ser debidos a errores de modelado, perturbaciones o ruido. Evidentemente, estas técnicas tienen como inconveniente que fallos pequeños pueden no ser detectados al no superar los umbrales de detección establecidos por la existencia de incertidumbre en el modelo.

Otra técnica muy utilizada para el diagnóstico de fallos basado en modelos analíticos es la identificación de parámetros (Isermann, 1993) y la utilización de observadores adaptativos. En comparación con las técnicas generadoras de residuos presentadas anteriormente, éstas últimas proporcionan un conocimiento más profundo de la magnitud del fallo que puede ser muy útil posteriormente para el análisis y el diseño de controladores tolerantes a fallos.

4.1.2. Detección de fallos basada en modelos cualitativos o semicualitativos. En ciertos casos es difícil disponer del conocimiento completo del proceso para construir un modelo analítico suficientemente representativo del mismo y, por lo tanto, se puede hacer inevitable grandes desviaciones entre la realidad y el modelo que hagan inservible este procedimiento para diagnosticar fallos.

De forma alternativa, un conocimiento incompleto puede ser tratado de forma abstracta mediante un modelo cualitativo que enfatice distinciones y relaciones primarias del proceso e ignore relaciones no importantes o desconocidas. Aunque los modelos cualitativos son por naturaleza imprecisos, pueden estar capacitados para representar bien el comportamiento del proceso complejo. En este caso, se utilizan conjuntos de valores catalogados mediante un atributo (positivo, negativo, disminuye, ...) en lugar de simples valores numéricos como elementos de base para la representación de modelos cualitativos. Otro tipo de modelos denominados semicualitativos utilizan conjuntos de valores caracterizados por intervalos o por conjuntos borrosos.

En los últimos años, el estudio de modelos cualitativos o semicualitativos para la monitorización y el diagnóstico de fallos está teniendo una gran respuesta (Kuipers, 1994), (Leitch R., 1994), (Travé-Massuyes, 2001). Los descriptores cualitativos de las variables pueden ser signos (De Kleer, 1984), intervalos (Puig, 2002a) o conjuntos borrosos (Shen, 1993). Inclusive, los conjuntos borrosos pueden ser una serie de intervalos, utilizando el principio de identidad "α-corte" (Nguyen, 1978) reduciendo el tratamiento borroso en cálculo intervalar (Puig, 2002b).

Entre las técnicas más relevantes que utilizan modelos cualitativos o semicualitativos para el diagnóstico de fallos se pueden citar:

- Observadores cualitativos utilizando ecuaciones diferenciales cualitativas (QDE). Las QDE son una extensión de las ecuaciones diferenciales ordinarias que utilizan variables y parámetros intervalares y las funciones imprecisas pueden ser representadas mediante reglas IF THEN. Un conocido algoritmo que simula las QDE es el QSIM de (De Kleer, 1984). Y los observadores cualitativos son una extensión de aquellos simuladores que

reducen el número de comportamientos irrelevantes (Zhuang, 1997).
- Detección de fallos utilizando envolventes que engloban la respuesta de la familia de sistemas representados por modelos cualitativos que utilizaban intervalos de valores en lugar de simples valores numéricos (también llamados modelos semi-cuantitativos). En (Bonarini, 1994), (Puig, 2003), los parámetros y las variables de estado intervalares son tratadas como un hipercubo que evoluciona a lo largo del tiempo y que proporciona unas envolventes adaptativas donde debe estar siempre la respuesta del sistema real sin fallo.

4.1.3. Detección de fallos mediante técnicas de soft-computing. En el caso de tener que detectar fallos en procesos complejos, puede ocurrir que no existen modelos matemáticos del mismo. Una forma de obtener dichos modelos a partir de los datos de entrada/salida en situación de buen funcionamiento es utilizando técnicas denominadas de ***soft-computing***:

- Redes Neuronales
- Lógica Borrosa
- Algoritmos Genéticos

o combinaciones de ellas para diagnosticar de forma robusta procesos no lineales de los que no se dispone de un conocimiento analítico. Entre otras técnicas destacan:

- Observador basado en redes neuronales (Marcu, 1999), cuyas redes neuronales representan modelos no lineales multi-entrada y multi-salida de tipo ARMA. El tipo de redes neuronales que se utilizan son una estructura mixta de perceptrón dinámico multi-capa (DMLP-MIX). El entrenamiento del modelo se realiza aplicando "backpropagation" dinámico.
- Observador basado en lógica borrosa (Chen, 1999) que representa un conjunto de observadores analíticos lineales cuyas salidas son fusionadas siguiendo los modelos borrosos propuestos por Tagaki-Sugeno. Utilizando esta técnica un sistema con dinámica no lineal se describe mediante un conjunto de modelos que linealizan el comportamiento en torno a unos puntos de funcionamiento.
- Diagnóstico basado en redes jerárquicas borroso neuronales que consiste en integrar una estructura con capacidad de razonamiento cualitativo (borroso) y la ventaja de poseer un mecanismo de aprendizaje (redes neuronales), tal como se propone en (Calado, 1999).

4.2 Aislamiento del fallo

Una vez realizada la detección, es necesario realizar un análisis de la causa que ha provocado la no consistencia, es decir diagnosticar el fallo. Este tipo de análisis puede realizarse siguiendo dos metodologías distintas, ya sea utilizando (Cordier, 2002):

- matriz de firmas de fallo (enfoque FDI o de Ingeniería de Control), o bien,
- el diagnóstico basado en consistencia (enfoque DX o de Inteligencia Artificial)

En ambos casos, se parte de un conjunto de indicadores de consistencia que se corresponden con las denominadas ***relaciones de redundancia analítica*** ARRs. Dichas relaciones se obtienen a partir de las relaciones elementales de los componentes del sistema a base de eliminar las variables intermedias no medidas, por lo tanto quedan vinculadas a los posibles fallos que pueden presentar los componentes utilizados para generarlas, tal como se ha descrito en la *Sección 3.2*.

4.2.1. Aislamiento del fallo según FDI. En el enfoque FDI, a partir del conjunto de residuos se define la ***matriz de firmas de fallo teórica***, Σ.

Definición 12 (Matriz firmas de fallo teórica) *La matriz de firmas de fallo teórica Σ contiene de forma codificada la dependencia de un determinado fallo (columna de la matriz) con cada residuo (fila de la matriz). Un elemento Σ_{ij} de esta matriz es igual a 1 si el fallo de la columna j influye en el residuo de la fila i, en caso contrario será 0, partiendo de la hipótesis de fallo simple. En el caso de que es consideren fallos múltiples, el número de columnas de la matriz de signatura de fallos teórica debería aumentarse hasta considerar todas las posibles combinaciones.*

En tiempo real, cada uno de los residuos $r_i(k)$ considerado será evaluado respecto a su umbral τ_i, siguiendo la metodología explicada en la sección anterior. El resultado proporciona un conjunto de ***firmas de fallo observadas*** del sistema $\mathbf{s}(k) = [s_1(k), \cdots, s_n(k)]$, donde

$$s_i(k) = \begin{cases} 0 & if \quad r_i(k) < \tau_i \\ \\ 1 & if \quad r_i(k) > \tau_i \end{cases} \tag{9}$$

Entonces, el aislamiento del fallo consistirá en encontrar cuál de las firmas de fallo de la matriz de signatura de fallos se aproxima más a la firma $\boldsymbol{s}(k)$ encontrada experimentalmente. El grado de aproximación se mide calculando alguna distancia entre las dos. En la práctica, las más utilizadas son la distancia Euclídea o la distancia de Hamming. Si, por ejemplo, utilizásemos la distancia de

Hamming, el procedimiento de aislamiento daría un vector de distancias de cada firma de fallo $\mathbf{d}(k) = [d_1(k), \cdots, d_n(k)]$, donde:

$$d_j(k) = \sum_{i=1}^{n} (\Sigma_{ij} \oplus s_i(k)) \tag{10}$$

y $\oplus$ es el operador lógico XOR. La firma de fallo teórico que presentara la distancia menor indicaría cuál es el posible fallo del sistema.

Ejemplo 2 (Firma de Fallo)
Dada una matriz de firma de fallos

	f_1	f_2	f_3	f_4	f_5
ARR_1	1	0	1	0	0
ARR_2	0	1	0	1	0
ARR_3	0	1	1	0	1

si la firma de fallos observada en un instante dado es $\mathbf{s} = [1, 0, 1]$*, el vector de distancias de Hamming es* $\mathbf{d} = [2, 1, 3, 0, 1]$*, para este caso. Esto permitiría deducir, a partir de escoger aquel fallo que presenta una firma teórica de distancia menor al fallo observado, que el fallo presente en el sistema es* f_3*.*

4.2.2. Aislamiento del fallo según DX. El diagnóstico en DX se fundamenta en la teoría de razonamiento lógico propuesta por Reiter (1987), la cual fue extendida y formalizada por De Kleer y Williams (1990). Para hacer el diagnóstico utiliza los términos ***sospechoso*** y ***candidato***. El primero se refiere a todo componente al que el sistema de diagnóstico identifique como posible responsable de la aparición de una discrepancia, mientras que el segundo se refiere a aquel componente o conjunto de componentes cuyo funcionamiento incorrecto explicaría todas las discrepancias observadas.

Cuando se detecta un síntoma, esto es, se produce una discrepancia, se deduce que alguno de los componentes involucrados en la predicción del valor discrepante de las observaciones ha de estar funcionando de forma anómala.

La característica fundamental de esta técnica es que realiza el diagnóstico de forma iterativa:

- Detección de conflictos. Consiste en enunciar el conjunto de componentes sospechosos en función de las discrepancias observadas.
- Diagnóstico. Se encargaría de encontrar los candidatos a partir de los sospechosos.
- Discriminación de hipótesis. Tendría que refinar el conjunto de candidatos si tras la fase anterior existiese más de uno.

Las relaciones de redundancia analítica no contienen información suficiente para que el diagnóstico DX sea posible. En el enfoque DX, a partir del conjunto de residuos, se define una ***matriz de firmas de fallo teórica extendida***, Σ_e que incorpora los componentes asociados a cada uno de ellos.

Definición 13 (Matriz de firmas teórica extendida) La matriz de firmas de fallo teórica Σ_e *contiene de forma codificada la dependencia de un determinado fallo que en este caso coincide con el componente o soporte de la ARR (columna de la matriz) con cada generador de síntomas o ARR (fila de la matriz). Un elemento* Σ_{eij} *de esta matriz es igual a 1 si el fallo de la columna j influye en el residuo de la fila i, en caso contrario será 0.*

Ejemplo 3 (Matriz Teórica Extendida)
Partiendo de la firma de fallos definida en el Ejemplo 2, el razonamiento DX sería:

- *Detección de conflictos: {*f_1*,* f_3*} y {*f_2*,*f_3*,* f_5*}.*
- *Diagnóstico: {*f_3*}, {*f_1*,* f_2*},{*f_1*,* f_5*}.*
- *Refinamiento del diagnóstico, en el caso de considerar fallo simple: {*f_3*}.*

Tal y como se comenta en Cordier (2002), la detección de conflictos a partir de la matriz de firma de fallos consistiría en evaluar de forma independiente cada una de las filas. Una vez detectados los posibles conflictos, debe utilizarse la teoría de razonamiento basado en consistencia para producir el diagnóstico.

Obsérvese que en las condiciones del ejemplo el resultado de diagnóstico FDI y DX es equivalente. Esto es generalizable bajo la hipótesis de fallos simples y en el caso en que los fallos sean observables por todas las *ARR* relacionadas con él.

4.3 Estimación el tamaño del fallo

La etapa de aislamiento permite determinar el fallo presente en el sistema pero no su magnitud. La estimación de dicha magnitud hace falta en determinados mecanismos de tolerancia a fallo.

Definición 14 (Estimación del tamaño del fallo) La estimación del tamaño de un fallo f_i *consiste en determinar su magnitud y evolución histórica.*

Existen diversas técnicas para la estimación del tamaño del fallo (Blanke, 2003). La mayoría se basan en que se ha llegado a modelar el modo en como actúa el fallo:

$$r_i(k) = G_{ij}(q) f_j(k) \tag{11}$$

donde $G_{ij}(q)$ es el modelo de cómo afecta el fallo f_j al residuo r_i. A partir de dicho modelo del

fallo se puede calcular la ***sensibilidad*** del residuo frente a dicho modo:

$$S_{ij} = \frac{dr_i}{df_j} = G_{ij}(q) \tag{12}$$

que permite calcular la magnitud del fallo a partir de la magnitud del residuo de acuerdo con:

$$f_j(k) = \frac{r_i(k)}{S_{ij}} = \frac{r_i(k)}{G_{ij}(q)} \tag{13}$$

5. CASO DE ESTUDIO: CONTROL TOLERANTE RED DE ALCANTARILLADO DE BARCELONA

La aplicación que se describe a continuación trata del control global de la red de alcantarillado de Barcelona que tiene como objetivo minimizar las inundaciones y vertidos al mar actuando sobre los elementos activos de la red de alcantarillado (depósitos de retención, compuertas de derivación, estaciones de bombeo) en caso de lluvia intensa. Un importante reto de esta aplicación es que su funcionamiento efectivo debe producirse en condiciones meteorológicas muy adversas, lo que se traduce en una gran probabilidad de que se produzcan errores en algunos de los más de 100 sensores (pluviómetros y limnímetros) que dispone la red de Barcelona y/o un mal funcionamiento en los actuadores de los depósitos de retención, compuertas de derivación o estaciones de bombeo, por lo que el control óptimo de la red de alcantarillado debe ser tolerante a fallos (Figueras, 2004).

La aplicación de control global (Cembrano, 2004) requiere el uso de un modelo operativo de la dinámica y de las restricciones de la red para calcular, en un horizonte de futuro, estrategias óptimas de control de los actuadores, teniendo en cuenta información sobre el estado actual de la red proporcionada por sensores conectados a un SCADA y a partir de una apropiada predicción de lluvia.

El caso, que se estudia en esta sección, corresponde a una importante zona del centro de la ciudad de Barcelona compuesta por varias subcuencas que tiene una superficie de 22,6 Km^2, con un depósito de retención (Escola Industrial) que dispone de dos compuertas de entrada y de salida y una compuerta de derivación (Tarragona), 11 depósitos virtuales que contienen todo el volumen que almacena la red de alcantarillado en esta zona, 12 pluviómetros que miden la intensidad de lluvia y 10 limnímetros que proporcionan los niveles de los colectores principales (Figura 4).

A continuación se presentará como se pueden aplicar las técnicas de diagnóstico de fallos en los sensores utilizados para el control global de la

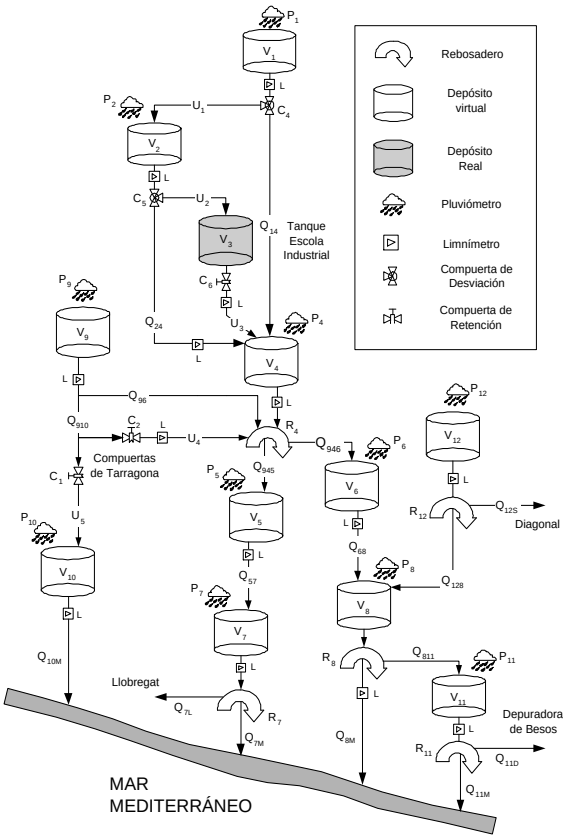

Figura 4. Modelo de la cuenca piloto de Barcelona.

red (pluviómetros y limnímetros). En un segundo artículo se presentarán como se pueden implementar mecanismos de tolerancia a fallos una vez han sido detectados de forma que el control global sea tolerante.

5.1 Diagnóstico de fallos en pluviómetros y limnímetros

La detección y aislamiento de los sensores en fallo se ha realizado mediante modelos analíticos (ver Sección 4).

El modelo utilizado para cada pluviómetro para diagnosticar fallos es de tipo estático:

$$\hat{P}_{ut}(k) = f(P_1(k), P_2(k), ..., P_m(k)) \tag{14}$$

siendo $\hat{P}_{ut}(k)$ la estimación al instante k de la medida proporcionada por el pluviómetro P_{ut} en función de otras mediciones reales en el mismo instante $P_1(k), P_2(k)$, .., $P_m(k)$. Otros modelos (series temporales) fueron descartados por la dificultad de prever adecuadamente la lluvia futura a partir de valores históricos. Los pluviómetros que interviene en cada una de estas relaciones estáticas (14) han sido derivados a partir de un análisis de correlación de 48 escenarios producidos durante un registro de 5 años. En la Figura 5 se muestra gráficamente los pluviómetros mas correlacionados con el pluviómetro P_1.

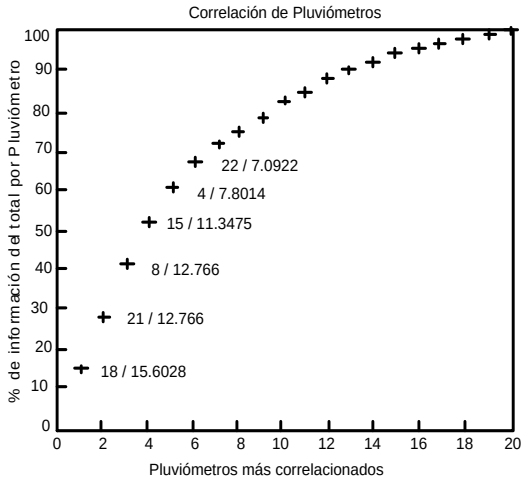

Figura 5. Pluviómetros más correlacionados con el pluviómetro P1.

Seleccionando los tres pluviómetros más correlacionados con cada uno de los pluviómetros de la red se han generado un conjunto de relaciones de redundancia analítica que permiten obtener la matriz de firmas de fallos teórica presentada en la Figura 6.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22
P1	X							X										X			X	
P2		X										X	X		X							
P3			X													X	X		X			
P4				X															X	X	X	
P5					X								X	X						X		
P6						X				X	X			X								
P7			X				X			X									X			
P8	X							X							X							X
P9									X								X	X	X			
P10						X				X	X			X								
P11						X				X	X			X								
P12		X										X	X		X							
P13		X			X							X	X									
P14						X					X			X		X						
P15		X													X						X	X
P16			X			X										X			X			
P17			X						X								X		X			
P18	X								X									X			X	
P19			X	X													X		X			
P20				X	X								X							X		
P21				X											X			X			X	
P22		X													X						X	X

Figura 6. Matriz de firmas de fallo teórica para la red de pluviómetros de Barcelona (en columnas se presentan los fallos y en filas los modelos, siguiéndose el criterio que una X representa incidencia y un cuadro en blanco no incidencia).

El modelo utilizado para el diagnóstico de fallos en los limnímetros se ha deducido a partir de la ecuación de balance de masas de los depósitos virtuales y reales en que se representado la red de alcantarillado (Figura 7):

$$L_{down}(k+1) = aL_{down}(k) + bL_{up}(k) + cI(k) \quad (15)$$

que establece el balance de masas entre el caudal de entrada de lluvia I al instante k y los niveles aguas arriba ($L_{up}(k)$) y aguas abajo ($L_{down}(k)$) de los colectores de una cuenca.

Utilizando el análisis estructural descrito en la *Sección 3.2* y a partir del modelo de la cuenca piloto (Figura 4) empleando la metodología de modelado basada en depósitos virtuales (Cembrano, 2004), se han deducido un conjunto de relaciones de redundancia analítica entre los diferentes limnímetros que permiten obtener la matriz de firmas fallos teórica (Figura 8).

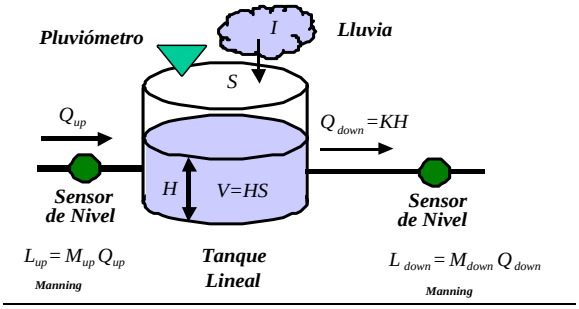

Figura 7. Modelo de una cuenca mediante un depósito virtual.

	L3	L7	L8	L9	L11	L16	L19	L20	L27	L39	L41	L47	L56	L80	P4	P6	P16	P20	
L3	X								X							X			L3
L7		X											X				X		L7
L8			X			X												X	L8
L9				X											X				L9
L11			X	X	X													X	L11
L16						X				X				X				X	L16
L19							X												L19
L20								X											L20
L27						X			X									X	L27
L39										X							X		L39
L41										X	X						X		L41
L47											X	X							L47
L56													X				X		L56
L80											X	X		X					L80

Figura 8. Matriz de firmas de fallo teórica para la red de limnímetros de Barcelona (se ha seguido el mismo convenio que en el caso de los pluviómetros).

5.2 Aplicación a escenarios reales

Una vez obtenidas las relaciones de redundancia analítica para pluviómetros y limnímetros se han calibrado utilizando técnicas de identificación intervalar (Ploix, 1999) que permiten obtener un modelo intervalar, él cual proporciona una predicción de tipo envolvente que engloba todos las medidas obtenidas en escenarios sin fallo. En la Figura 9 se presenta el resultado de la validación de la calibración del modelo intervalar para el limnímetro L16.

En las Figuras 10 y 11 se presentan los resultados de las pruebas de detección para los limnímetros L16 y L80, respectivamente, en un escenario en el que el limnímetro L80 está en fallo. Se puede observar que el fallo se detecta en el instante 1000. En la Figura 12 se presenta para este instante el aislamiento que realizaría un algoritmo basado en

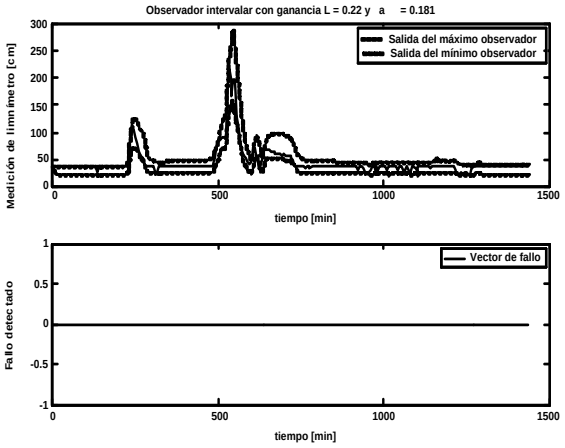

Figura 9. Validación de la calibración del modelo intervalar del limnímetro L16 para un escenario sin fallo.

el cálculo de la mínima distancia entre la firma de fallo observado y la teórica (Sección 4).

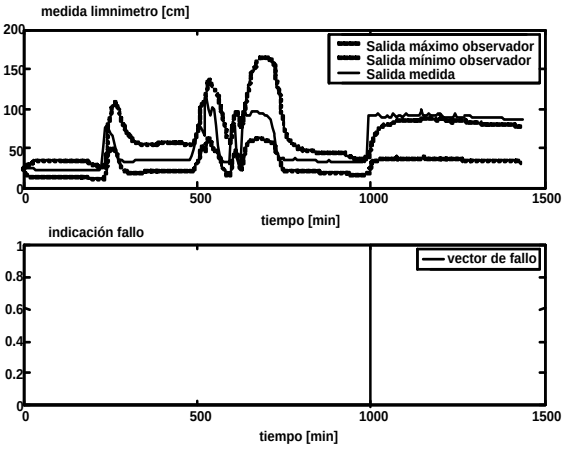

Figura 10. Test de detección en el limnímetro L16 en un escenario con fallo.

6. CONCLUSIONES

Como se ha visto en este trabajo, el control tolerante a fallos integra técnicas como detección y diagnóstico de fallos, análisis estructural, etc., que son motivo de investigación y desarrollo desde hace más tiempo que el mismo concepto de control tolerante. Incluye aspectos muy interesantes y nada obvios como son la capacidad de detectar y aislar fallos en procesos realimentados, la severidad de los fallos, la capacidad de recuperabilidad del proceso controlado en fallo,... tal como se ha visto en los apartados precedentes. En este artículo, se han presentado los fundamentos del control tolerante a fallos así como una metodología de diseño. Posteriormente, se ha pasado a describir las primeras etapas de dicha metodología, concretamente, el análisis estructural y el diagnóstico de

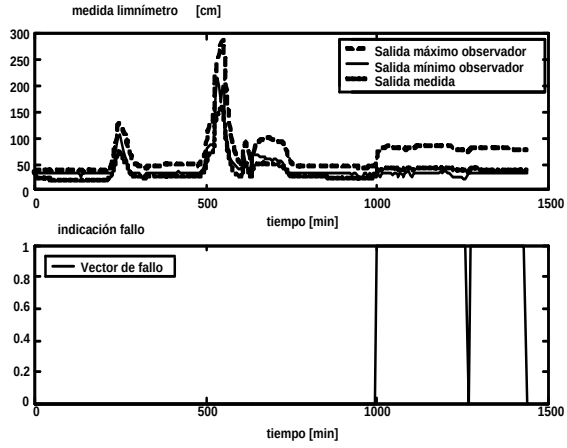

Figura 11. Test de detección en el limnímetro L80 en un escenario con fallo.

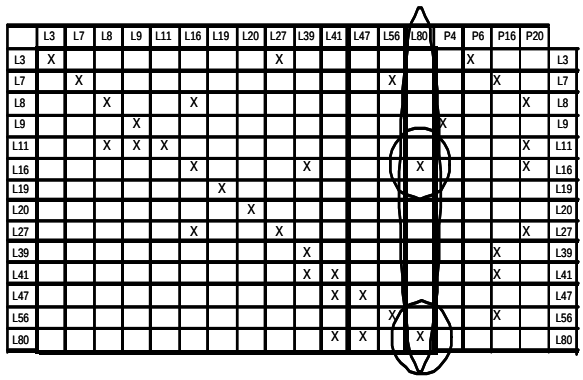

Figura 12. Aislamiento de un fallo el limnímetro L80 a partir de las pruebas de detección presentados en las Figuras 10 y 11.

fallos. Como ejemplo de aplicación de las técnicas de análisis estructural y de diagnóstico de fallos propuestas, se ha utilizado un caso real basado en la red de alcantarillado de Barcelona donde se desea diagnosticar fallos en los sensores y actuadores utilizados para el control global de la misma de cara a la activación de mecanismos de tolerancia que permitan un control tolerante.

En un segundo artículo se presentaran los mecanismos de tolerancia, junto con la descripción de un proceso real en el que ha empezado a incorporarse herramientas de control tolerante.

Por otra parte, dada la juventud y empuje del tema, existe un importante colectivo de grupos de investigación muy activos a nivel internacional que cooperan entre ellos (DAMADICS, BRIDGE, CHEM, IFATIS, FTCOSY, DEFTAC) para dar respuestas teóricas a muchos de los problemas presentados en este artículo y el siguiente, así como su aplicación ya sea por los mismos grupos de investigación o por empresas punteras de sectores aeronáuticos, automoción, energéticos (nuclear), químicos, petroquímicos y de componentes

de automatización (robots), principalmente. Por ello cabe esperar la aparición de soluciones reales fiables de control tolerante en los próximos años.

AGRADECIMIENTOS

Este trabajo ha sido subvencionado por la CICYT del Ministerio de Ciencia y Tecnología Español (DPI2002-0350) y por la DGR de la Generalitat de Catalunya (grupo SAC 2001/SGR/00236). Los autores también desean agradecer el apoyo recibido por la empresa CLABSA S.A. en la aplicación de este trabajo.

REFERENCIAS

Basseville, M. y I.V.Ñikiforov (1993). *Detection of abrupt changes: theory and applications*. Prentice Hall.

Blanke, M. (1996). Consistent design of dependable control systems. *Control Engineering Practice* **4**, 1305–1312.

Blanke, M. (2000). What is fault-tolerant control?. *Proceedings of IFAC SAFEPROCESS'00* **35**, 123–126.

Blanke, M. (2001). Concepts and methods in fault-tolerant control. *Proceedings of American Control Conference*.

Blanke, M., M. Kinnaert J. Lunze y M. Staroswiecki (2003). *Diagnosis and fault-tolerant control*. Springer-Verlag. Germany.

Bonarini, A. y G. Bontempi (1994). A qualitative simulation approach to fuzzy dynamical models. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* **4**, 258–313.

Calado, J.M.F. y Sa de Costa, J.M.G. (1999). Online fault detection and diagnosis based on a coupled system. *Proceedings of European Control Conference*.

Cassar, J. y Staroswiecki, M. (1997). A structural approach for the design or failure detection and identification systems. *IFAC/IFIP/IMACS Conference on Control of Industrial Processes*.

Cembrano, G., Quevedo J. Salamero M. Puig V. Figueras J. y Martí J. (2004). Optimal control of urban drainage systems: a case study. *Control Engineering Practice*.

Chen, J. y Patton, R.J. (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publishers.

Chow, E.Y. y Willsky, A.S. (1984). Analytical redundancy and the design of robust failure detection systems. *IEEE Transactions on Automatic Control* **AC-29**, 603–614.

Clark, R.N. (1989). State estimation for instrument fault detection. In: *Fault diagnosis in dynamic systems, theory and application* (Clark RN Patton RJ, Frank PM, Ed.). Prentice Hall. New York.

Cordier, M.O., Dague P. Dumas M. Lévy M. Montmain J. Staroswiecki M. y Travé-Massuyès L. (2002). Comparative analysis of ai and control theory approaches to model-based diagnosis. *European Conference on Artificial Intelligence (ECAI'2000)* pp. 274–279.

De Kleer, J. y Brown, J.S. (1984). Qualitaitve physics based on confluences. *Artificial Intelligence* **24**, 77–83.

De Kleer, J. y Williams, B. (1990). Diagnosing multiple faults. *Artificial Intelligence* **32**, 97–130.

Ding, X. y Jeinsch, T. (1999). An approach to analysis and design of observer and parity relation based fdi systems. *Proceedings of XIV IFAC World Congress*.

Emami-Naeini, A., Akhter M.M. y Rock S.M. (1988). Effect of model uncertainty on failure detection: the threshold selector. *IEEE Transactions on Automatic Control* **AC-33**, 1106–1115.

Figueras, J., Puig V. Quevedo J. (2004). Contribution to the optimal control of sewage networks including fault-tolerant capabilities. *1ª Jornada de Investigación en Automática, Visión y Robótica. Universidad Politécnica de Cataluña*.

Frank, P.M. (1991). Enhancement of robustness in observer-based fault detection. *IFAC Symposium SAFEPROCESS* **2**, 275–287.

Frank, P.M., Ding S.X. y Köppen-Seliger B. (2000). Current developments in the theory of fdi. *IFAC Symposium SAFEPROCESS* **1**, 16–27.

Gertler, J. (1998). *Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker. New York.

Gertler, J.J. (1991). Analytical redundancy methods in fault detection and isolation. *IFAC Symposium SAFEPROCESS* **1**, 9–21.

Herrin, S.A. (1981). Maintainability applications using the matrix fmea technique. *IEEE Transactions R-30* pp. 212–217.

Horak, D.T. y Guidance, J. (1988). Failure detection in dynamic systems with modelling errors. *Control and Dynamics* **11(6)**, 508–516.

Isermann, R. (1993). Fault diagnosis of machines via parameter estimation and knowledge processing. *Automatica* **29**, 815–836.

Kuipers, B. (1994). *Qualitative Reasoning - Modelling and Simulation with Incomplete Knowledge*. MIT Press.. Cambridge, MA.

Leitch R., Shen, Q. Conghil G.-Chantler M. y Slater A. (1994). Qualitative model-based diagnosis of dynamic systems. *Colloquim of the Institution of Measurement and Control*.

Marcu, T., Matcovschi M.H. y Frank-P.M. (1999). Neural observer-based approach to fault detection and isolation of a three-tank systems. *Proceedings of European Control Conference*.

Nguyen, H.T. (1978). A note on the extension principle for fuzzy sets. *Journal of Mathematical Analysis and Applications* **64**, 369–380.

Patton, R. J. (1997). Fault-tolerant control: the 1997 situation. *Proceedings of IFAC Symposium on SAFEPROCESS* pp. 1033–1055.

Ploix, S., Adrot O. y Ragot-J. (1999). Parameter uncertainty computation in static linear models. *IEEE Conference on Decision and Control*.

Puig, V., Quevedo J. Escobet T. y De las Heras S. (2002a). Robust fault detection approaches using interval models. *IFAC World Congress*.

Puig, V., Saludes J. y Quevedo J. (2003). Worst-case simulation of discrete linear time-invariant interval dynamic systems. *Reliable Computing* **9(4)**, 251–290.

Puig, V. y Quevedo, J. (2002b). Passive robust fault-detection using fuzzy parity equations. *Mathematics and computers in Simulation* **60**, 193–207.

Pulido, B. y Alonso, C. (2002). Possible conflicts, arrs, and conflicts. *13th International Workshop on Principles of Diagnosis (DX-02)* pp. 122–128.

Reiter, R. (1987). A theory of diagnosis from the first principles. *Artificial Intelligence* **32**, 57–95.

Shen, Q. y Leitch, R. (1993). Fuzzy qualitative simulation. *IEEE Transactions on SMC*.

Staroswiecki, M., Cassar J.P. y Declerk P. (2000). A structural framework for the design of FDI system in large scale industrial plants. In: *Issues of Fault Diagnosis for Dynamic Systems* (Clark RN Patton RJ, Frank PM, Ed.). Springer-Verlag.

Staroswiecki, M. y P. Declerck (1989). Analytical redundancy in non-linear interconnected systems by means of structural analysis. *IFAC/IMACS/IFORS Conference AIPAC'89*.

Travé-Massuyes, Escobet, T. Pons R. y Tornil S. (2001). The caen diagnosis system and its automatic modelling method. *Computación y Sistemas*.

Travé-Massuyes, Escobet, T. y Spanache S. (2003). Diagnosability analysis based on component supported analytical redundancy relations. *Proceedings IFAC SAFEPROCESS* pp. 887–890.

Zhuang, Z. y Frank, P.M. (1997). Qualitative observer and its application to fault detection and isolation systems. *Systems and Control Engineering* **211**, 253–262.

ANEXO: TERMINOLOGÍA

Conceptos Básicos

Fallo (*fault*): desviación no permitida de, al menos, una propiedad característica o parámetro de un sistema de su condición aceptable, usual o estándar. Un fallo es la aparición de un modo de fallo.

Fallo abrupto (*abrupt fault*): fallo cuyo efecto aparece repentinamente (por ejemplo, modelado mediante un escalón).

Fallo incipiente (*incipient fault*): Fallo cuyo efecto aparece progresivamente (por ejemplo, modelado mediante una rampa).

Avería (*failure*): interrupción permanente de la capacidad de un sistema para realizar una función requerida bajo las condiciones de operación especificadas.

Diagnóstico de fallos (*fault diagnosis*): determinación del tipo, tamaño, localización e instante de aparición de un fallo. Incluye la detección, el aislamiento y la estimación del fallo.

Detección de fallos (*fault detection*): determinación de la presencia de fallos en el sistema así como el instante de su aparición.

Aislamiento de fallos (*fault isolation*): determinación del tipo, localización e instante de detección de un fallo. Se realiza después de la etapa de detección.

Estimación del fallo (*fault estimation*): determinación del tamaño y comportamiento del fallo durante el tiempo.

Modelado del fallo (*fault modeling*): determinación de un modelo (matemático) que describe un determinado fallo.

Modos de fallo (fault modes): descripción de los tipos de fallo que puede presentar un componente (matemáticamente). Cada modo de fallo suele tener asociado unos efectos que se dan en mayor o menor medida según la magnitud del mismo.

Análisis de los efectos (effects analysis): consiste en el estudio de la aparición de un modo de fallo en un componente y de cómo se propaga por el sistema.

Sistema tolerante a fallos (*fault tolerant system*): sistema que, ante la aparición de un fallo, mantiene su función con o sin degradación de prestaciones, pero sin desembocar en una avería a nivel de subsistema o sistema.

Supervisión

Supervisión (*supervision*): es una actividad de alto nivel que engloba las actividades de monitoriza-

ción (o vigilancia), la detección y diagnóstico de fallos y el control supervisor o de planta.

Control supervisor (*supervisory control*): es la actividad que se lleva a cabo sobre un conjunto de controladores para asegurar que sus objetivos de control se cumplen. Esta definición también se utiliza para el termino supervisión local.

Supervisión de planta (*plant-wide supervision*): es la actividad cuyo objetivo es asegurar que las trayectorias de ciertas variables clave de un proceso siguen de forma adecuada unas trayectorias de referencia dadas. Esta definición también se utiliza para el termino supervisión global.

Supervisor (*supervisor*): entidad (humana o artificial) que realiza la supervisión de un proceso mediante el diagnóstico de fallos, determinación y ejecución de las acciones correctoras en presencia de fallos.

Monitorización (*monitoring*): determinación continua en tiempo real del estado de operación de un sistema mediante el registro y análisis de información significativa e indicación de sus anomalías de comportamiento. Otros términos que se utilizan para este concepto son el de vigilancia (surveillance) y observación del estado.

Sistema tolerante a fallos (*fault-tolerant system*): sistema que permite acomodar un fallo con o sin degradación de prestaciones, pero sin desembocar en una avería a nivel de subsistema o sistema.

Diagnóstico de fallos

Redundancia (*redundancy*): uso de más de un método de obtener el estado o características de un sistema.

Redundancia física o hardware (*hardware redundancy*): uso de más de un instrumento independiente para conseguir una determinada función.

Redundancia analítica (*analytical redundancy*): uso de dos o más formas, no necesariamente idénticas, de determinar una variable donde una forma utiliza un modelo matemático del sistema de forma analítica.

Discrepancia (*discrepancy*): comportamiento anómalo de un valor físico o inconsistencia entre varios valores físicos y la relación entre ellos.

Residuo (*residual*): señal que contiene información del fallo basada en la desviación entre las medidas de las entradas/salidas del sistema (comportamiento real) y estimaciones obtenidas mediante un modelo del mismo (comportamiento modelado). El residuo describe el grado de consistencia entre el comportamiento real y el modelado.

Generación del residuo (*residual generation*): determinación del residuo a partir del modelo y las entradas/salidas del sistema.

Evaluación del residuo (*residual evaluation*): análisis del residuo de con fin de detectar, aislar e identificar el fallo.

Umbral (*threshold*): valor del residuo a partir del cual se considera la existencia de un fallo.

Análisis estructural (*structural analysis*): análisis de las propiedades estructurales de los modelos, es decir, de las propiedades que son independientes de los valores concretos de los parámetros. Se utiliza para la obtención de las ecuaciones que permitirán generar los residuos.

Control tolerante

Tolerancia a fallos (*fault-tolerance*): capacidad de un sistema de control de mantener los objetivos de control a pesar de la aparición de un fallo, aceptando una posible degradación de las prestaciones. La tolerancia a fallos se puede obtener mediante la acomodación del fallo o mediante la reconfiguración del controlador o del sistema.

Tolerancia activa a fallos (*active fault-tolerance*): sistema tolerante a fallos con diagnosis y acomodación explícita de los mismos.

Tolerancia pasiva a fallos (*passive fault-tolerance*): sistema tolerante a fallos sin diagnosis y acomodación explícita de los mismos. La tolerancia, en este caso, se basa en el diseño del controlador para que sea insensible (robusto) a un conjunto restringido de fallos.

Leyes de control admisibles (*admissible control laws*): conjunto de algoritmos que se pueden implementar para resolver un determinado problema de control.

Objetivos de control (*control objectives*): conjunto de especificaciones que el sistema controlado debe alcanzar cuando se utiliza una determina ley de control.

Restricciones de control (*control restrictions*): conjunto de relaciones funcionales que el comportamiento del sistema controlado debe satisfacer durante el tiempo.

Problema de control estándar (*standard control problem*): consiste en seleccionar una ley de control dentro de un conjunto de leyes de control admisibles, de forma que el sistema controlado alcance los objetivos de control a la vez que su comportamiento satisface un conjunto de restricciones de control.

Acomodación de fallos (*fault accomodation*): 1. Acción correctora (basada en el cambio de operación del sistema) que evita que un cierto fallo

desemboque en un efecto final no deseado. 2. Cambio en los parámetros del controlador o en su estructura para evitar los efectos de un fallo. Las entradas y/o salidas del controlador continúan siendo las mismas. Los objetivos de control se alcanzan aunque las prestaciones se pueden degradar.

Reconfiguración (*reconfiguration*): cambio en las entradas y/o salidas del controlador a través de un cambio en la estructura del controlador y sus parámetros. Los objetivos de control se alcanzan aunque las prestaciones se pueden degradar.