



UNIVERSIDAD SIMÓN BOLÍVAR

Decanato de Estudios de Postgrado

Doctorado en Ingeniería

**ALGORITMOS GENÉTICOS EN EL DISEÑO
INTEGRADO DE PROCESOS QUÍMICOS**

**Tesis Doctoral presentada a la Universidad Simón Bolívar por
Silvana R. Revollar C.**

**Como requisito parcial para optar al título de
Doctora en Ingeniería**

**Realizado con la tutoría de las Profesoras
Rosalba Lamanna (Tutora)
Pastora Vega (Co-tutora)**

Noviembre de 2010



UNIVERSIDAD SIMÓN BOLÍVAR
Decanato de Estudios de Postgrado
Doctorado en Ingeniería

ALGORITMOS GENÉTICOS EN EL DISEÑO INTEGRADO
DE PROCESOS QUÍMICOS

Esta Tesis Doctoral ha sido aprobada en nombre de la Universidad Simón Bolívar
por el siguiente jurado examinador:

Presidente

Miembro Externo

Miembro Principal

Miembro Principal

Miembro Principal-Tutor

Noviembre de 2010



UNIVERSIDAD SIMÓN BOLÍVAR

Decanato de Estudios de Postgrado

Doctorado en Ingeniería

VEREDICTO

VEREDICTO SOBRE EL TRABAJO DE TESIS DOCTORAL PRESENTADO POR
SILVANA REVOLLAR PARA OPTAR AL TÍTULO DE DOCTORA EN INGENIERÍA.

Quienes suscribimos, profesores...miembros del jurado designado por la Coordinación del Doctorado en Ingeniería de la Universidad Simón Bolívar para considerar y evaluar el trabajo de tesis doctoral, titulado: ALGORITMOS GENÉTICOS EN EL DISEÑO INTEGRADO DE PROCESOS QUÍMICOS presentado por Silvana R. Revollar C. dejamos constancia de lo siguiente :

En fe de todo lo cual, levantamos y firmamos el presente veredicto en el Valle de Sartenejas, Caracas, a los ... días del mes del año de dos mil ...

Presidente

Miembro Externo

Miembro Principal

Miembro Principal

Miembro Principal-Tutor

A mis padres



UNIVERSIDAD SIMÓN BOLÍVAR

Decanato de Estudios de Postgrado
Doctorado en Ingeniería

ALGORITMOS GENÉTICOS EN EL DISEÑO INTEGRADO DE PROCESOS QUÍMICOS

Presentado por: Silvana R. Revollar C.

Noviembre de 2010

RESUMEN

El *Diseño Integrado* es una metodología para llevar a cabo el diseño del proceso simultáneamente con el análisis de la dinámica de la planta y el diseño del sistema de control, en el marco de un problema de optimización que permite obtener una solución compromiso entre objetivos económicos y objetivos de *controlabilidad*. La complejidad de estos problemas hace necesario el uso de métodos de optimización poderosos, por lo que el objetivo principal de este trabajo es la aplicación de algoritmos genéticos en los problemas de diseño integrado, como una alternativa a los métodos tradicionales de optimización, ahondando al mismo tiempo en diversas extensiones o aspectos novedosos en los planteamientos de estos problemas. Los algoritmos genéticos son métodos de optimización capaces de enfrentar casos con discontinuidades, ruido y dinámicas complicadas. La formulación matemática a partir de una *superestructura* de la planta y contemplando la síntesis del proceso, selección de la estructura de control, dimensionamiento y cálculo de los controladores, se traduce en un problema de *optimización dinámica no lineal mixta entera*. En el algoritmo genético desarrollado se usa codificación mixta, con genes reales y binarios para tratar las variables continuas y discretas en el *problema de diseño integrado*, y además se incorporan técnicas heurísticas basadas en conocimiento para el ajuste del algoritmo. Se desarrollan diferentes aplicaciones de la metodología a tres procesos diferentes: un sistema de reactores, la torre de sulfatación en la operación de clarificación del azúcar, y el proceso de fangos activados en el tratamiento de aguas residuales. Se consideran especialmente casos de estudio que involucran decisiones estructurales, se evalúan índices de *controlabilidad* basados en el modelo no lineal y métricas de *controlabilidad* de estado y se aplican esquemas avanzados de control. En todos los casos, los resultados ponen en evidencia la ventaja de aplicar esta metodología, que permite obtener las dimensiones óptimas de la planta, junto con el punto de trabajo y un controlador sintonizado para garantizar los requerimientos de *controlabilidad* impuestos, así como también se constata la bondad de los algoritmos genéticos como método de resolución.

Palabras clave: Diseño integrado, síntesis de procesos, algoritmos genéticos, *controlabilidad*

ÍNDICE GENERAL

INDICE.....	vi
CAPITULO I. INTRODUCCIÓN.....	1
1. ESTADO DEL ARTE EN EL DISEÑO INTEGRADO DE PROCESOS.....	3
1.1. Metodologías Anticipativas.....	3
1.2. Diseño Simultáneo del Proceso y Control.....	7
2. PLATAFORMA DE OPTIMIZACION PARA EL DISEÑO INTEGRADO DE PROCESOS.....	19
3. OBJETIVOS.....	24
CAPITULO II. LA CONTROLABILIDAD DE LOS PROCESOS.....	28
1. INTRODUCCION.....	28
2. ANALISIS DE CONTROLABILIDAD BASADO EN EL MODELO LINEAL.....	31
2.1. Controlabilidad de Estado.....	34
2.2. Controlabilidad Entrada-Salida.....	35
3. CONTROLABILIDAD DE ESTADO PARA EL MODELO NO LINEAL.....	47
3.1. Controlabilidad basada en Métodos Numéricos.....	49
3.2. Condiciones de Controlabilidad de Estado No Lineal.....	52
4. CONCLUSION.....	60
CAPITULO III. DISEÑO INTEGRADO DE PROCESOS.....	62
1. INTRODUCCION.....	62
2. INTERACCION ENTRE EL DISEÑO Y EL CONTROL DEL PROCESO	64
2.1. Conflicto Economía-Controlabilidad.....	64
2.2. Controlabilidad y Estructura de Control.....	68
3. FORMULACION DEL PROBLEMA DE DISEÑO INTEGRADO DE PROCESOS..	70
4. FORMULACION DEL PROBLEMA DE DISEÑO INTEGRADO DE UN SISTEMA DE REACTORES.....	77
5. CONCLUSION.....	83
CAPITULO IV. ALGORITMOS GENÉTICOS.....	84
1. INTRODUCCION.....	84
2. LOS ALGORITMOS GENETICOS.....	85
2.1. Generalidades.....	85
2.2. Algoritmos Genéticos de Codificación Real.....	90
2.3. Operadores Genéticos.....	92
2.3.1. Selección.....	94
2.3.2. Cruce.....	98
2.3.3. Mutación.....	103
2.4. Tratamiento de las restricciones.....	105
3. EJEMPLO DE APLICACIÓN DEL ALGORITMO GENÉTICO AL DISEÑO INTEGRADO DE UN SISTEMA DE REACTORES.....	115
4. CONCLUSION.....	120
CAPITULO V. DISEÑO INTEGRADO DE UNA TORRE SULFITADORA DE CAÑA	

DE AZÚCAR.....	121
1. INTRODUCCION.....	121
2. DESCRIPCIÓN DE LA TORRE SULFITADORA.....	120
2.1.Modelo Matemático de la Torre Sulfidadora.....	123
2.2. Estudio de Controlabilidad de la Planta.....	128
3. FORMULACION DEL PROBLEMA DE OPTIMIZACION.....	131
4. DISEÑO INTEGRADO DE LA TORRE SULFITADORA CON UN CONTROLADOR PID.....	137
5. DISEÑO INTEGRADO DE LA TORRE SULFITADORA CON UN CONTROLADOR MPC.....	144
6. CONCLUSION.....	152
CAPITULO VI. DISEÑO INTEGRADO DEL PROCESO DE FANGOS ACTIVADOS.....	153
1. INTRODUCCION.....	153
2. EL PROCESO DE FANGOS ACTIVADOS.....	154
2.1. Modelo del Proceso de Fangos Activados	156
2.2. Operación del Proceso de Fangos Activados.....	164
3. DISEÑO INTEGRADO DEL PROCESO DE FANGOS ACTIVADOS. MODELO 1.....	170
3.1.Formulación General del Problema de Optimización.....	173
3.2. Casos de Estudio.....	178
3.2.1. Síntesis económica del proceso.....	178
3.2.2.Diseño integrado con un controlador PID.....	181
3.2.3.Diseño integrado con un controlador predictivo	187
4. DISEÑO INTEGRADO DEL PROCESO DE FANGOS ACTIVADOS PARA LA ELIMINACION DE MATERIA ORGÁNICA Y NUTRIENTES. MODELO 2.....	197
4.1.Formulación General del Problema de Optimización.....	199
4.2. Casos de Estudio.....	202
5. CONCLUSION.....	205
CAPITULO VII. CONCLUSION.....	207
REFERENCIAS.....	211
APENDICE I. ANALISIS DE CONTROLABILIDAD EN DISEÑOS PRELIMINARES...	222

CAPÍTULO I

INTRODUCCIÓN GENERAL

La necesidad de aplicar políticas de ahorro energético y protección al medio ambiente junto a la gran competitividad a nivel industrial, se traducen constantemente en mayores exigencias en cuanto a la optimización de los procesos químicos. Esto propicia una mayor automatización de los procesos, aumenta el interés en la aplicación de estrategias de control más eficientes e impulsa el diseño de plantas con unidades interconectadas que permitan ahorrar materia y energía, pero que añaden mayor dificultad a la operación. Una respuesta a estos desafíos es la metodología de *diseño integrado* de procesos, que permite incorporar el análisis del desempeño dinámico del sistema en el procedimiento de diseño, lo cual lleva a obtener diseños económicamente óptimos que garantizan inherentemente un desempeño dinámico satisfactorio, lo cual representa grandes ventajas desde los puntos de vista económico y operacional.

El diseño de un proceso químico, es una tarea extensa que empieza por la necesidad de obtener un producto que cumpla ciertas especificaciones, ya sea en la industria del petróleo, petroquímica, de alimentos o de manufactura entre muchas otras, y termina con la definición cuantitativa de todas las variables estructurales y de operación de una planta de producción.

La concepción y el diseño de un proceso químico se basan en la optimización de factores como los costos de construcción, los costos de operación y el consumo de energía, sujetos a restricciones de operabilidad del proceso y regulaciones ambientales. De acuerdo a estos requerimientos, se determina el modo de operación, la estructura física de la planta y las dimensiones de las unidades que la componen. Esto se lleva a cabo en distintas etapas del

procedimiento de diseño que comprende la síntesis estructural o síntesis del proceso, el diseño de los equipos y el cálculo del punto de operación.

En la *síntesis del proceso* se realiza el estudio preliminar de las operaciones básicas requeridas para transformar la materia prima en el producto final, y se analizan distintas alternativas estructurales de la planta física. En efecto, se trata de encontrar la mejor configuración desde el punto de vista de los beneficios económicos y minimización de los costos de inversión y operación. Es después de establecer la estructura de la planta que se estiman las dimensiones y condiciones óptimas de operación y, finalmente, se lleva a cabo la *ingeniería de detalle*, que abarca el diseño mecánico de los equipos y el desarrollo de los planos de instrumentación y de servicios auxiliares (Puigjaner et al., 2006).

Finalmente, la definición de los sistemas de control se realiza como una tarea posterior al diseño. Los sistemas de control son los que se encargan de adaptar la planta a los cambios a que está sujeta durante su operación, esto es, a variaciones en la demanda, en las especificaciones del producto, o en las características de las corrientes de alimentación. Sin embargo, las posibilidades del sistema de control no dependen únicamente de la estrategia escogida ni del correcto ajuste del controlador, sino de la dinámica de la propia planta, que queda ya definida en la etapa de diseño (Ziegler y Nichols, 1943; Skogestad y Wolff, 1992; Luyben, 1993; Luyben y Floudas, 1994). Es posible entonces que la planta presente una dinámica desfavorable que dificulte el funcionamiento de cualquier sistema de control (Bahri et al., 1996). En efecto, como las decisiones de diseño se fundamentan solo en la optimización de criterios económicos y de operación en estado estacionario, el proceso diseñado puede resultar muy difícil de controlar.

Se pone entonces en evidencia la necesidad de realizar un análisis riguroso de la controlabilidad del proceso como parte del problema de diseño, para evaluar cuál es el diseño de la planta y la estrategia de control que permiten satisfacer de manera conjunta los objetivos económicos y de operación dinámica (Ekawati, 2003). En este marco se debe tomar en cuenta que el beneficio económico, generalmente, tiene que ver con la minimización de los costos de inversión y operación en estado estacionario de la planta y con políticas de ahorro de energía. Mientras que la operación eficiente tiene que ver con el desempeño dinámico y con la

posibilidad de obtener los productos de acuerdo a las especificaciones y con poca variabilidad en su calidad: la controlabilidad. Esto lleva a definir la controlabilidad como la habilidad de la planta para lograr el desempeño dinámico deseado usando las variables manipuladas disponibles, dentro de los límites impuestos por las condiciones de operación (Skogestad y Wolff, 1992). Cualitativamente, la controlabilidad es una propiedad que indica qué tan bien rechaza perturbaciones una planta, qué tan severamente interactúan las múltiples variables y qué tan fácil se mueve de una condición de operación a otra (Luyben y Floudas, 1994).

En los últimos años, en ingeniería de procesos se ha impulsado el desarrollo de técnicas para incorporar criterios de controlabilidad en las etapas preliminares de diseño, estableciendo como filosofía la *Integración Diseño-Control*. El *Diseño Integrado* es una metodología donde se llevan a cabo el diseño del proceso simultáneamente con el análisis de la dinámica de la planta o incluso el diseño del sistema de control. La formulación matemática del *Diseño Integrado* se traduce en un problema de optimización con múltiples objetivos, donde se busca obtener una solución compromiso entre criterios económicos y criterios de controlabilidad. Existe un número importante de trabajos donde se proponen distintas alternativas para enfrentarse al *Diseño Integrado*. A continuación se presenta una revisión del estado del arte en esta área.

1. ESTADO DEL ARTE EN EL DISEÑO INTEGRADO DE PROCESOS

En la literatura sobre el estudio de la integración *Diseño-Control*, se tratan distintos aspectos del problema, como son la forma de introducir los criterios de controlabilidad en el procedimiento de diseño o la manera de formular y resolver el problema. Sin embargo, una fundamental clasificación se puede hacer distinguiendo las *Metodologías Anticipativas* de las metodologías de *Diseño y Control Simultáneo*. En las primeras, el diseño y control se llevan a cabo de forma secuencial, pero se examina la controlabilidad de las diversas alternativas de diseño de la planta para la selección del diseño más conveniente. En el *Diseño y Control Simultáneo* los criterios de control se introducen en el mismo procedimiento de diseño dando origen a un problema de optimización que tiene como resultado la planta que mejor satisface el compromiso entre economía y controlabilidad y, generalmente, permite obtener el controlador ya sintonizado.

1.1 METODOLOGÍAS ANTICIPATIVAS

Las *metodologías anticipativas* introducen técnicas de análisis de controlabilidad en los pasos de diseño del proceso de manera de obtener plantas más fáciles de controlar, pero se mantiene la separación entre las tareas de diseño y las de control. Generalmente, se enfocan en la optimización del modelo en estado estacionario y la evaluación de índices de controlabilidad basados en el modelo dinámico lineal, de manera de comparar distintas posibilidades de diseño en base a sus características operacionales para escoger un diseño económicamente óptimo y controlable. Dentro de esta clasificación se incluyen también algunos trabajos orientados a la síntesis de procesos, específicamente, al análisis del efecto de la estructura del proceso sobre la controlabilidad y otros donde se estudia el efecto de la estructura de control sobre el desempeño dinámico de diseños ya establecidos.

La *síntesis de procesos* se trata normalmente mediante métodos sistemáticos como la *descomposición jerárquica* y la *optimización de superestructuras*. El más utilizado cuando se trata del diseño de plantas completas es el *método jerárquico* que permite abordar estos problemas mediante su descomposición en una secuencia de subproblemas ordenados o jerarquizados de acuerdo al nivel de detalle de las decisiones de diseño involucradas. El método jerárquico comprende cinco niveles donde se decide el tipo de operación (continuo o por carga), la estructura entradas-salidas en el diagrama de flujo del proceso, los reciclos necesarios, el diseño de los sistemas de separación y el diseño de los sistemas de intercambio de calor (Douglas, 1988).

Entre los primeros en estudiar sistemáticamente la controlabilidad durante la *síntesis de procesos* se encuentran Fisher et al. (1988). Estos introducen, en el procedimiento de *síntesis jerárquica*, objetivos de control relacionados con el efecto de las perturbaciones típicas del proceso sobre los costos de operación y las restricciones. Por lo tanto, para cada posible diseño y en cada nivel de la síntesis, se verifican los grados de libertad entre las variables de control y las variables manipuladas y se examina el efecto de las perturbaciones, con el objeto de establecer si es necesario realizar modificaciones en el diseño para mejorar la controlabilidad. Más adelante, en Barton et al. (1992), se propone un análisis de controlabilidad más formal, donde se examina la estabilidad del proceso a lazo abierto, se selecciona la estructura de

control en función del beneficio económico, se aplica el arreglo de ganancias relativas en estado estacionario para determinar los posibles emparejamientos, y se estudian las limitaciones a la *controlabilidad funcional* evaluando la posibilidad de polos en el semiplano derecho y la norma del número de condición. En ambos trabajos, se evalúa la controlabilidad de diseños predeterminados obtenidos en base a la optimización económica del modelo estacionario y éstos se modifican en función de las deficiencias en la capacidad del proceso a ser controlado. Esto implica, generalmente, incluir más variables manipuladas en el diagrama de flujo (corrientes de purga, reciclos y equipos auxiliares), ignorar la optimización de las variables menos significativas, y sobredimensionar el diseño de algunos equipos para evitar que se violen restricciones. Este último es un aspecto criticable de este procedimiento.

Un trabajo que destaca dentro de las metodologías anticipativas es el de Luyben (1993) para escoger estructuras alternativas de un sistema de reactores tipo tanque agitado conectados en serie. La controlabilidad se evalúa utilizando un índice directamente relacionado con parámetros operacionales de la planta, como la relación entre el máximo calor retirado posible para el sistema de enfriamiento y el calor retirado en condiciones normales (Q_{max}/Q).

También existen procedimientos para analizar sistemáticamente la controlabilidad, que comprenden la evaluación de un conjunto de índices basados en el modelo lineal del proceso, y que cubren áreas como la interacción, el rechazo a perturbaciones y la factibilidad de aplicar control descentralizado. Entre éstos se encuentran los trabajos de Wolff et al. (1992) y Wolff (1994), para el diseño de plantas integradas, cuyos diagramas de flujo se caracterizan por ser complejos, y el de Wolff et al. (1994) para estudiar la *operabilidad* de plantas completas. En Skogestad (1994) también se estudia la controlabilidad para sistemas de una entrada-una salida con métodos de análisis frecuencial.

El campo de aplicación de estas metodologías ha sido extenso, se tratan tanto procesos complicados como la hidrodealquilación y el craqueo catalítico, como procesos más sencillos como los sistemas de reactores (TAC) y sistemas de neutralización, que son de gran interés por su carácter interactivo y no-lineal,

Otra visión del análisis de controlabilidad dentro de la *síntesis del proceso* lleva a la evaluación de la *controlabilidad estructural*. En el trabajo de Georgiou y Floudas (1990) se

propone un índice de *controlabilidad estructural* para determinar el diagrama de flujo óptimo de una planta en base a la minimización de los costos y a la vez, lograr rechazo total o disminución de la propagación de perturbaciones. Lee et al. (2001), relacionan el concepto de *controlabilidad estructural* con el recorrido de las perturbaciones a lo largo del proceso, siendo esta propiedad adecuada si una perturbación indeseable no se propaga. Esto se puede determinar a partir de la información relativa a la estructura, aunque se desconozcan mayores detalles del proceso. Representando el modelo del proceso como un “dígrafo”, se utilizan matrices de orden relativo para determinar el desacoplamiento del sistema, lo que permite seleccionar los mejores diagramas de flujo y descartar las alternativas no controlables.

La posibilidad de incluir el análisis de controlabilidad en la *síntesis del proceso*, en términos de la sensibilidad a perturbaciones, utilizando una descripción de modelos termodinámicos ha sido estudiada por Meeuse et al. (2000, 2001), Meusse (2002) y Meeuse y Grievink (2002). En estos trabajos, se relaciona la sensibilidad a la producción de entropía del sistema con su sensibilidad ante perturbaciones.

Finalmente, una metodología interesante para la integración Diseño-Control, basada en el análisis de *controlabilidad de estado* para sistemas no lineales la propone Ochoa (2005) mediante la evaluación de ciertas métricas de *controlabilidad práctica*. Esto contempla el modelado fenomenológico del proceso, la selección de variables manipuladas y *selección de la estructura (pareamiento) de control* (mediante el método de Lee et al. (2001)). También contempla la determinación del intervalo de acción disponible para las variables de entrada y la selección de escenarios y tolerancias a las perturbaciones. Se trata de la optimización de la planta en función de los costos de inversión y operación, incluyendo la evaluación de las métricas de controlabilidad y tomando en cuenta las restricciones sobre éstas. Finalmente, se lleva a cabo el diseño del sistema de control para la planta óptima a la que se le ha garantizado la controlabilidad en el punto de operación deseado. Una extensión de este trabajo para realizar el diseño integrado de sistemas acoplados se encuentra en Muñoz et al. (2007) quienes proponen una metodología que permite verificar la controlabilidad de los sistemas acoplados basada en el cálculo de la distribución de accesibilidad y la descomposición en estados controlables/no controlables. Se utiliza como ejemplo una planta intercambiador-tanque flash-reactor químico donde ocurre un proceso de separación de dos sustancias (Amoniaco-Agua).

En resumen, la literatura de metodologías anticipativas en diseño integrado, introduce el análisis de controlabilidad en las etapas iniciales del diseño para garantizar el buen desempeño dinámico de la planta. Este análisis se lleva a cabo al evaluar las posibles alternativas de la planta obtenidas en base a la optimización económica del modelo en estado estacionario donde se incluyen índices de controlabilidad a lazo abierto. Los criterios de controlabilidad, en los trabajos descritos anteriormente se centran en el estudio de efecto de las perturbaciones sobre las restricciones de operación y su propagación a lo largo del proceso, en el análisis de la información que aportan índices basados en el modelo lineal del proceso o en el modelo termodinámico y en el análisis de la estructura del proceso y del esquema de control. Entre los trabajos propuestos se destaca el de Ochoa et al., (2005) porque permite acotar a priori la controlabilidad basada en modelos no lineales.

1.2 DISEÑO INTEGRADO

Se denomina *Diseño Integrado* a la metodología que trata de manera simultánea el *diseño del proceso* y el *diseño del sistema de control* en el marco de un problema de optimización. La formulación de este problema, comprende variables de decisión, funciones objetivo y restricciones relacionadas con aspectos económicos y de desempeño del sistema de control. Este planteamiento permite obtener un diseño satisfactorio económico, con una estructura de control definida y, en muchos casos, con un controlador sintonizado, para cumplir de forma óptima los objetivos de control, lo cual es muy conveniente desde el punto de vista práctico y es una ventaja evidente de este método respecto al diseño tradicional.

La investigación en *Diseño Integrado* se ha visto impulsada por el gran avance en el campo de la informática, que ha hecho posible el desarrollo de métodos y herramientas de optimización cada vez más poderosas. De igual manera, el énfasis en el estudio de la *interacción entre el diseño y la controlabilidad* en el área de ingeniería de procesos, ha provisto de un buen número de técnicas para evaluar esta propiedad en el marco del problema de optimización. Por esta razón, en la literatura reciente se observa un incremento en las aplicaciones de esta metodología, así como distintas propuestas para su perfeccionamiento.

En las aplicaciones más completas de *diseño integrado* se contempla inclusive la selección de la *estructura del proceso* (*síntesis*) y la selección de la *estructura de control*

(pareamientos entre variables controladas y manipuladas y esquema de control). Cuando se considera la *síntesis* se plantea un problema de optimización basado en el análisis de una *superestructura* que contiene todas las posibles alternativas estructurales del proceso (*Síntesis algorítmica* o *síntesis automática*), cuyo objetivo es encontrar la estructura más económica que satisface los criterios de controlabilidad. La selección de la estructura de control se puede incluir en la misma *superestructura*. Esta formulación involucra variables continuas, que representan dimensiones y condiciones de operación, y variables discretas, que representan decisiones relacionadas con la estructura del proceso. La desventaja de la *síntesis algorítmica* es la dificultad para resolver rigurosa y eficientemente los problemas de optimización que se originan, por lo que su aplicación se limita a problemas de dimensión reducida o a etapas particulares del proceso, pues resultaría casi imposible generar y evaluar todas las alternativas existentes para el diseño de una planta completa (Puigjaner et al., 2006).

Diseñar simultáneamente el proceso y el control introduce la dificultad de sintonizar los controladores y evaluar su desempeño en el marco del problema de optimización. Debido a la complejidad del problema, en la mayoría de los trabajos se han aplicado esquemas tradicionales de control Proporcional Integral (PI), pero recientemente se ha propuesto el uso de técnicas avanzadas como el control predictivo (Sakizlis et al., 1994). En cuanto al cálculo de los parámetros del controlador, en algunos casos se introducen como variables de decisión en la optimización, mientras que en otros se sintonizan con métodos empíricos.

Un aspecto clave en el *diseño integrado* es la manera de cuantificar el comportamiento dinámico y su impacto sobre los costos. En la literatura se distinguen métodos que emplean *índices lineales* de controlabilidad y métodos basados en optimización del modelo dinámico con *índices numéricos*. En Ricardez-Sandoval et al. (2009) se presentan además los métodos basados en la evaluación de modelos lineales con incertidumbre usando como índices *normas de control robusto*.

De aquí, se entiende que dependiendo del alcance del problema, la estrategia de control y la manera de evaluar la controlabilidad se puede llegar a formulaciones matemáticas distintas del problema. En Sakizlis et al. (2004) también se realiza una revisión de las técnicas

dedicadas al *diseño simultáneo del proceso y el control* clasificando los métodos para formular el problema de diseño integrado en dos categorías:

1. El diseño de procesos económicamente óptimos capaces de operar eficientemente alrededor de un punto de operación estacionario, considerando dos objetivos: un índice de desempeño económico y una medida numérica del comportamiento dinámico como la integración del error cuadrático de control: ISE (“Integrated Square Error”). Estos métodos presentan el inconveniente de no poder definir la importancia relativa de los dos objetivos que compiten ni estudiar sistemáticamente el comportamiento dinámico de la planta.
2. El diseño considerando un único índice de desempeño económico, mientras que las especificaciones del sistema y su operación se representan usando modelos dinámicos en lugar de modelos en estado estacionario. Esto lleva a la aplicación de métodos de optimización dinámica para determinar la solución más económica que satisface las restricciones de *operabilidad*.

Introducir la optimización del modelo del proceso con todas sus particularidades en la formulación matemática del problema, permite un análisis más completo y realista del desempeño dinámico de la planta, que va más allá de la evaluación los índices de controlabilidad basados en un modelo linealizado. Sin embargo, esto acarrea mayor esfuerzo en el desarrollo de modelos detallados del proceso, que luego deben ser simplificados para poder resolver el problema de optimización, que finalmente limitan la aplicación de la metodología a casos de pequeña escala, debido a la complejidad del problema resultante (Lewin, 1999).

En las revisiones presentadas en la literatura (Lewin, 1999; Sakizlis et al., 2004; Seferlis y Georgiadis, 2004; Ricardez-Sandoval, 2009) se han establecido algunos criterios para organizar las distintas contribuciones en el área de diseño integrado. Sin embargo, aquí se propone una clasificación más amplia en función de todos los variados aspectos que contempla el problema, a saber: la definición del alcance del diseño (diseño solo o síntesis y diseño), estrategias de control a utilizar (estructura y controladores), las técnicas para evaluar la controlabilidad y cuantificar el efecto de perturbaciones e incertidumbre, la formulación del problema de optimización en sí, y los métodos para su resolución.

Es difícil abordar en profundidad todos los aspectos del problema a la vez, por esto, en la literatura se distinguen distintas líneas de investigación enfocadas en el estudio de algún aspecto particular, a saber:

1. *El alcance del problema de diseño.*

La formulación general del problema de diseño integrado incluye, además del dimensionamiento de la planta y el cálculo del punto de operación, decisiones estructurales relacionadas con la síntesis del proceso y los esquemas de control.

En Schweiger y Floudas (1994) por ejemplo, se presenta *la síntesis y diseño integrado* de redes de reactores químicos y sistemas compuestos reactor - columna de destilación. En otros trabajos aún más completos, como Luyben y Floudas (1994a,b), Walsh y Perkins (1994, 1996), Mohideen et al. (1996), Bahri et al. (1996, 1997), Bansal et al. (1998), y Ekawati (2003), se plantea y resuelve el problema de diseño incluyendo la *síntesis del proceso* y la *selección de la estructura de control*.

Muchos otros trabajos que consideran problemas de síntesis, se enfocan solo en la *selección de la estructura de control*: Narraway et al. (1991), Narraway y Perkins (1993, 1994), Mohideen et al. (1996), Heath et al. (2000), Kookos y Perkins (2001), Bansal et al. (2002), Sakizlis et al. (2003), Ekawati y Bahri (2003), Asteasuain et al. (2005, 2006, 2007), Flores-Tlacuahuac y Biegler / (2007), Patel et al. (2007), Exler et al. (2008), Hamid et al. (2010).

Entre los casos de estudio encontrado se tienen columnas de destilación binaria con distintos grados de complejidad y con integración de calor, sistema de reactor-separador-reciclo, evaporadores de múltiples efectos, reactores de polimerización, plantas de tratamiento de aguas, procesos de neutralización, el proceso de fangos activados y el proceso de Tennessee Eastman.

Por otra parte, muchas aplicaciones consideran ya definido el diagrama de flujo de la planta. Cabe mencionar algunas muy interesantes como Brengel y Seider (1992), que realizan el diseño integrado de un proceso de fermentación con una marcada no linealidad y tendencia a la inestabilidad; Ricardez-Sandoval et al. (2008) diseñan un proceso de mezcla con gran

incertidumbre en los parámetros; Miranda et al. (2008) aplican el procedimiento a un intercambiador de calor coaxial y una columna de destilación reactiva; Flores-Tlacuahuac y Biegler (2008) a un reactor de polimerización; Seon y Linninger (2010) a un proceso de isomerización. Destaca el trabajo de Ricardez-Sandoval et al. (2010) para sistemas de gran escala como el proceso de Tennessee Eastman. En esta categoría también entran los trabajos de Lenhoff y Morari (1982), Palazoglu y Arkun (1986), Blanco y Bandoni (2003) y Chawankul et al. (2005).

2. Las estrategias de control.

La idea del diseño integrado contempla que el sistema de control se diseña de manera simultánea con el proceso, aunque esto último se evita en algunos trabajos por simplicidad. En estos casos, se evalúa la controlabilidad a lazo abierto como en Luyben y Floudas (1994), o se supone control perfecto como en Narraway y Perkins (1993, 1994), y en Blanco y Bandoni (2003).

Los esquemas de control más utilizados en la mayoría de los casos se basan en controladores PI descentralizados (Narraway et al., 1991; Walsh y Perkins, 1994; Bahri, 1996; Sweiger y Floudas, 1997; Bansal et al., 2002; Exler et al., 2008, entre otros). Un paso hacia la aplicación de controladores y esquemas avanzados se observa en Kookos y Perkins (2001) quienes implementan un controlador PI multivariable. En Asteasuain et al. (2006) se utiliza una combinación de esquemas PI-retroalimentado y de alimentación adelantada (“*feed forward*”) multivariados y en Asteasuain et al. (2007) se utiliza un controlador PI multivariable y un esquema de control de relación.

Generalmente, los parámetros del controlador PI son variables de decisión en el problema de optimización, sin embargo, en Bahri (1996) y Bahri et al. (1997) se usan controladores PI prediseñados que son afinados luego al aplicar el procedimiento.

Brengel y Seider (1992) son de los primeros en proponer estrategias avanzadas de control, introduciendo un controlador predictivo (MPC) no lineal en el marco del problema de diseño integrado. En Loeblein et al. (1999) se utiliza un MPC sin restricciones, luego Sakizlis et al. (2003) y Sakizlis et al. (2004) implementan un controlador predictivo (MPC)

paramétrico, que permite calcular directamente las acciones de control evitando la optimización del controlador en línea. En Baker y Swartz (2006) se considera un MPC lineal con restricciones. Francisco y Vega (2006) y Francisco et al. (2010) también desarrollan aplicaciones de estrategias avanzadas de control basadas en MPC. En trabajos preliminares relacionados con esta tesis doctoral como Revollar et al. (2008) y Francisco et al. (2009), se encuentran aplicaciones de diseño integrado con controladores MPC basados en el modelo lineal, Revollar et al. (2010) se introduce un MPC basado en el modelo no lineal.

Otras estrategias se proponen en Chawankul et al. (2005), que utilizan un controlador por modelo interno (*IMC*), y también en Patel et al. (2007), Terrazas-Moreno (2008) y Miranda et al. (2008), quienes aplican esquemas de control óptimo.

3. Los métodos para evaluar la controlabilidad.

Las técnicas utilizadas para cuantificar la controlabilidad o en general el desempeño dinámico del proceso se separan en tres grupos:

a. Métodos basados en el uso de índices lineales de controlabilidad

Se puede citar aquí a Luyben y Floudas (1994), que utilizan índices de controlabilidad basados en el modelo lineal en estado estacionario expresados en función de los parámetros del proceso, como la matriz de ganancias relativas (RGA), el mínimo valor singular y el número de condición. Ya antes Palazoglu y Arkun (1986), aplican el análisis de valores singulares. También Blanco y Bandoni (2003) consideran el mínimo valor singular de la matriz de transferencia en estado estacionario como medida de controlabilidad. Finalmente, Narraway y Perkins (1993) utilizan modelos linealizados para estimar el margen entre el punto de operación de la planta y el óptimo económico que garantiza que ninguna de las restricciones del proceso sea violada debido al efecto de perturbaciones (*margen de retroceso*), siendo ésta una medida del potencial económico de las posibles estructuras de control.

b. Métodos basados en optimización del modelo dinámico no lineal con índices numéricos

El análisis de controlabilidad basado en el modelo dinámico se lleva a cabo calculando algún índice de comportamiento del modelo para un horizonte de tiempo predefinido. Una de las técnicas más utilizadas es el cálculo de la integral del cuadrado del error de control (ISE) a partir de la resolución del modelo dinámico no lineal. Ejemplos del uso de este índice se encuentran en Schweiger y Floudas (1997), Bansal et al. (1998), Asteasuain et al. (2006, 2007), Exler et al. (2008) y en Flores-Tlacuahuac y Biegler (2007), donde además del ISE, se usa el tiempo de establecimiento como índice adicional.

En Mohideen et al. (2006) se incorpora el análisis de *factibilidad dinámica* al problema de diseño integrado, que consiste a grandes rasgos en verificar que todas las restricciones de operación y restricciones sobre las variables de control se cumplan para todo el rango de parámetros con incertidumbre en el horizonte de tiempo establecido, lo que usualmente se logra mediante la imposición de restricciones de trayectoria en el problema.

Bahri et al. (1996, 1997) proponen el análisis de *operabilidad dinámica* en el marco de la *síntesis del proceso y selección de la estructura de control*. Este análisis comprende aspectos como la *estabilidad, controlabilidad y flexibilidad*, el objetivo es maximizar la economía del proceso sujeta a una dinámica regulatoria factible. Destaca el uso del *margen de retroceso* en base al modelo dinámico no lineal para relacionar los aspectos económicos con la *operabilidad*, al determinar la distancia entre el punto de operación óptimo en estado estacionario y el punto de operación dinámico de la planta. También se consideran la *factibilidad dinámica* e índices como el ISE y el tiempo de establecimiento. En Ekawati y Bahri (2003) se complementa este análisis introduciendo el *índice de controlabilidad de salida* (OCI) (Vinson y Georgakis, 2000).

En Exler et al. (2008) se evalúa un conjunto de índices de desempeño entre los que se incluye el ISE, junto con otras métricas en lazo abierto directamente relacionadas con el proceso de fangos activados tratado, entre las que se encuentran la energía de bombeo y la energía de aireación en el sistema. La sensibilidad de un proceso a perturbaciones y a cambios en las variables manipuladas son también índices interesantes, usados en el trabajo de Hamid et al. (2010), donde además se estudian aspectos relacionados con la termodinámica del proceso para determinar los parámetros óptimos de diseño.

c. Métodos robustos

En Chawankul et al. (2005) se introduce como medida del desempeño a lazo cerrado la peor variabilidad en la salida controlada debida a incertidumbre en el modelo. Ésta además, se integra en la función objetivo como el costo asociado a la variabilidad. También se imponen restricciones sobre la estabilidad robusta de la planta.

En Ricardez-Sandoval (2008, 2009) se propone una nueva técnica basada en el cálculo de límites estimados sobre las variables del proceso que determinan su flexibilidad, estabilidad y controlabilidad. El comportamiento no lineal del proceso a lazo cerrado alrededor del punto de operación nominal se describe mediante un modelo en espacio de estado con incertidumbre paramétrica. El modelo es utilizado para probar la estabilidad robusta y medir la variabilidad del proceso frente al peor escenario de perturbaciones e incertidumbre mediante la función de Lyapunov y el análisis de valores singulares estructurados.

4. El tratamiento de las perturbaciones e incertidumbre

En muchos trabajos se ignora el efecto de las perturbaciones e incertidumbre en los parámetros o se consideran perfiles de perturbaciones muy simples (Narraway et al., 1991; Narraway y Perkins, 1994; Schweiger y Floudas, 1994; Bahri, 1996; Kookos y Perkins, 2001).

Sin embargo se distingue un grupo de publicaciones enfocadas en el estudio del efecto que tienen sobre el desempeño dinámico y la economía del proceso, diferentes escenarios de perturbaciones e incertidumbre en los parámetros del modelo (Mohideen et al., 1996; Bahri et al., 1996, 1997; Bansal et al., 2000; Asteasuain et al., 2007; Ricardez-Sandoval et al. 2008, 2009, 2010). En Bandoni et al. (1994) se propone el “*algoritmo del peor caso*” donde se determina la máxima variación que pueden experimentar los parámetros con incertidumbre sin afectar la factibilidad del proceso. En Chawankul et al. (2005) se considera explícitamente la incertidumbre del modelo al aplicar técnicas de control robusto.

5. La formulación del problema de optimización y los métodos para su resolución.

La formulación matemática del diseño integrado depende del alcance del problema, de la manera de introducir los criterios de controlabilidad y del modelo utilizado para representar el

proceso. El planteamiento más sencillo se reduce a un problema de optimización no lineal (NP). Si se considera la síntesis del proceso o selección de la estructura de control, el problema se transforma en mixto-entero (MINLP). Se tiene finalmente un problema de optimización dinámica (MIDO) si se representa el comportamiento del proceso utilizando el modelo dinámico en lugar del modelo estacionario.

La naturaleza multi-objetivo del problema se maneja de diferentes maneras, ya que pueden considerarse los distintos objetivos como distintas funciones a optimizar, o puede optimizarse una función objetivo basada en aspectos económicos mientras los aspectos relacionados con la controlabilidad se introducen como restricciones.

a. Formulaciones multi-objetivo

En Luyben y Floudas (1994) se plantea un problema de programación no lineal mixta-entera (MINLP) multi-objetivo, donde se optimizan objetivos económicos y algunos de los índices lineales de controlabilidad. Desarrollos similares, pero que resultan solo en un problema de programación no lineal (NLP) se encuentran ya en Lenhoff y Morari (1982), Palazoglu y Arkun (1986). Blanco y Bandoni (2003) introducen las medidas de controlabilidad en este tipo de formulación haciendo uso de la teoría de optimización de autovalores.

En Asteasuain et al. (2006) el *diseño y control simultáneo* de un reactor de polimerización se traduce en un problema de optimización multi-objetivo mixto-entero dinámico (MIDO). Los objetivos son una función de costos de inversión y operación y un índice dinámico similar al *ISE* que se relaciona con calidad del producto. El problema se resuelve aplicando un algoritmo de descomposición donde se tiene un *MINLP* maestro y un problema de optimización dinámica asociado.

b. Formulaciones de un objetivo económico con restricciones de controlabilidad.

En Narraway y Perkins (1993) se formula un problema de programación mixta entera (*MINLP*), en el cual se determina el punto de operación óptimo en estado estacionario y la penalización económica asociada con el “*margen de retroceso*” propuesto en Narraway et al. (1991). Esta metodología se extiende al caso no lineal en Narraway y Perkins (1994).

En Bahri (1996) se optimiza la economía del proceso y se garantiza una dinámica regulatoria factible mediante restricciones sobre las condiciones de *operabilidad dinámica*. El problema se resuelve aplicando un algoritmo iterativo de dos niveles, en el primero se obtiene la estructura, dimensiones y condiciones de operación mediante un *MINLP* y en el segundo nivel se examina la factibilidad de la solución mediante la resolución de problemas *NLP*. En Bahri et al. (1996), Bahri et al. (1997) se aplica esta metodología, en Ekawati y Bahri (2003) se amplía integrando un nuevo índice de controlabilidad al análisis de *operabilidad dinámica*.

Mohideen et al. (1996) proponen el siguiente planteamiento general del problema de optimización a formular para abordar el diseño integrado:

Minimizar: Costo total anualizado

Sujeto a:

Restricciones del modelo en ecuaciones diferenciales y algebraicas

Restricciones de trayectoria

Factibilidad en la operación

Variabilidad del proceso debido a perturbaciones e incertidumbre

La formulación matemática resulta en un problema de optimización dinámica mixta-entera (*MIDO*). El algoritmo de resolución propuesto requiere la descomposición en dos subproblemas y la aplicación de un procedimiento iterativo, primero se determina el diseño óptimo del proceso y la selección de la estructura de control y, definidos éstos, en el siguiente subproblema se evalúa la factibilidad de operación del proceso para todo el rango de perturbaciones e incertidumbre. Esta plataforma ha sido adoptada en los trabajos de Bansal et al. (2002), Sakizlis et al. (2003), Sakizlis et al. (2004). En Mohideen et al. (2006) se considera el costo de la estructura de control en el costo total anualizado.

También en Schweiger y Floudas (1997) se considera el objetivo económico, mientras que los objetivos de control, basados en el modelo dinámico no lineal, se tratan como restricciones. Se llega a un problema de control óptimo mixto entero (*MIOCP*) que luego se simplifica en un *MINLP/DAE*, es decir, en un problema de programación mixta entera con

ecuaciones diferenciales. Imponiendo distintos límites a las restricciones se desarrollan las curvas de Pareto que reflejan las soluciones compromiso.

Kookos y Perkins (2001) proponen otro algoritmo de descomposición que se basa en la generación de cotas superiores e inferiores sobre el desempeño económico de la planta. El algoritmo se lleva a cabo en dos pasos, primero se optimizan la topología de la planta y la estructura del sistema de control, en una segunda etapa, los parámetros continuos e invariantes en el tiempo se determinan por optimización dinámica.

En Chawankul et al. (2005) se incluye la variabilidad de la variable controlada en la función objetivo, imponiendo restricciones sobre las variables manipuladas para mejorar el rechazo a perturbaciones y restricciones que garantizan la estabilidad robusta. En este método la planta no lineal se representa con una familia de modelos lineales.

El trabajo de Asteasuain et al. (2007) es una extensión de Asteasuain et al. (2006) para considerar incertidumbre y perturbaciones, aunque se usa una función objetivo única relacionada con la calidad del producto. Se aplica el algoritmo “*del peor caso*” para resolver el problema, el cual consiste en una optimización de dos niveles. Primero se considera un conjunto de parámetros inciertos y luego se extiende para encontrar la mayor violación de las restricciones de operación sobre el dominio de parámetros con incertidumbre.

c. *Otras formulaciones*

En Brengel y Seider (1992) se propone una estrategia coordinada de optimización para enfrentar el problema de diseño simultáneo del proceso y del *MPC*, en donde la función basada en aspectos económicos es penalizada cuando la controlabilidad es deficiente. Esto se traduce en un problema de programación de dos niveles (“*Bilevel programming problem-BPP*”) el cual luego es simplificado para su resolución. Un procedimiento similar, también relacionado con la implementación de un *MPC* se propone en Baker y Swartz (2006) quienes introducen el problema cuadrático (*QP*) del controlador en la plataforma de diseño integrado sustituyéndolo por las condiciones de optimalidad de Karush-Kurn-Tucker, las cuales se introducen como restricciones en el problema de diseño integrado.

En Hamid et al. (2010) se propone una metodología alternativa de diseño integrado basada en modelos. Consiste en que el problema de optimización se descompone en cuatro etapas que se resuelven de manera jerárquica para reducir el espacio de búsqueda progresivamente: pre-análisis, análisis del diseño, análisis del controlador y finalmente la selección y verificación. En esta última etapa se calcula la función objetivo para el conjunto reducido de soluciones factibles halladas en las etapas previas.

En resumen, se han logrado grandes avances en el *área de diseño integrado*, aunque aún queda mucho por desarrollar. En cuanto al alcance del problema de diseño, se han descrito algunas formulaciones completas, que involucran decisiones discretas y evaluación de la dinámica a lazo cerrado que se han llevado a cabo con éxito, como Bansal et al. (2002) que tratan una columna de destilación binaria representada por un modelo riguroso, lo que resulta en un problema *MIDO* de gran dimensión (3000 ecs. diferenciales, 600 estados diferenciales, 85 variables binarias), y Asteasuain et al. (2007) que diseñan complejos reactores de polimerización implementando de manera simultánea esquemas de control retroalimentado y de alimentación adelantada. Sin embargo, en la mayoría de los casos, se trata sólo del diseño de un equipo o de unidades de procesos sencillas.

En cuanto a las técnicas para evaluar la controlabilidad, es notoria la falta de conciliación entre la visión de *controlabilidad de estado* y la de *controlabilidad entrada-salida*, donde se encuentran además muy pocas aplicaciones basadas en la *controlabilidad de estado*. En un trabajo desarrollado en el marco de esta tesis (Lamanna et al., 2009) se incorpora este análisis al *diseño y control simultáneo* de una torre de sulfitación, que aunque un ejemplo sencillo, pone de manifiesto el potencial de este método.

Aún en cuanto a la evaluación de la controlabilidad, cuando ésta se hace en base al modelo dinámico no lineal en presencia de perturbaciones, se obtienen problemas de optimización *dinámica mixta-entera (MIDO)*. Este análisis de controlabilidad ofrece numerosas ventajas, como la facilidad de interpretar los resultados, de relacionar directamente índices económicos con el desempeño dinámico y de analizar la flexibilidad del proceso en presencia de perturbaciones, pero es el elemento dentro del problema de optimización que requiere de mayor esfuerzo computacional. Esto hace evidente la necesidad de investigar en

dos vertientes: métodos de optimización lo suficientemente eficientes y robustos para tratar estos problemas, pero también métodos alternativos para evaluar la controlabilidad como los basados en la evaluación de índices robustos, para aliviar la carga computacional impuesta por la resolución de los modelos dinámicos.

Otro aporte de gran importancia que todavía asoma poco en la literatura es la incorporación de técnicas avanzadas de control en los diseños, los cuales introducen mejoras significativas en el desempeño dinámico del proceso, especialmente en los casos multivariados. Este es un aspecto que debe explotarse enfrentando a la vez el consecuente incremento de la dificultad del problema de optimización.

Como se puede ver, los avances en el análisis del desempeño dinámico y en la aplicación de técnicas avanzadas de control en el marco del problema de diseño integrado, están limitados por el nivel de complejidad que se puede llegar a tratar en los problemas de optimización resultantes. De aquí la fundamental importancia de los métodos que permiten resolver eficientemente los problemas de optimización no lineal mixta entera (MINLP y MIDO) que surgen de las diversas formulaciones, y que por ello requieren consideración por separado.

2. PLATAFORMA DE OPTIMIZACION PARA EL DISEÑO INTEGRADO

La formulación matemática del diseño integrado *de procesos*, generalmente se traduce en un problema de *Optimización No-Lineal Mixto-Entero (MINLP)* o en un problema de *Optimización Dinámica No-Lineal Mixta-Entera (MIDO)* dependiendo de la complejidad del planteamiento. La formulación del problema no lineal-mixto entero es de la forma:

$$\min_{\mathbf{x}, \mathbf{y}} f(\mathbf{x}, \mathbf{y})$$

Sujeto a las restricciones:

$$h(\mathbf{x}, \mathbf{y}) = 0 \quad (1.1)$$

$$g(\mathbf{x}, \mathbf{y}) \leq 0$$

$$\mathbf{x} \in \mathbb{R}, \mathbf{y} \in \{0,1\}^l$$

La función $f(\mathbf{x}, \mathbf{y})$ es una función de costos. Por lo general, el vector \mathbf{x} contiene variables del proceso como flujos, composiciones y temperaturas. El vector \mathbf{y} , de variables enteras, representa variables estructurales como la presencia o ausencia de conexiones y equipos, también representa elementos discretos en la entonación o estructura de los controladores (horizontes de predicción o selección de esquemas de control). Las restricciones de igualdad y desigualdad, $h(\mathbf{x}, \mathbf{y})$ y $g(\mathbf{x}, \mathbf{y})$, están relacionadas con los balances de masa, energía, relaciones termodinámicas, restricciones en las variables del proceso y restricciones lógicas asociadas a las conexiones entre las diferentes unidades de la superestructura.

El problema es de *Optimización Dinámica No-Lineal Mixta-Entera (MIDO)* cuando los parámetros presentan variaciones importantes en el tiempo, por lo que se consideran índices de desempeño dinámico y restricciones de trayectoria o punto final en el marco de la optimización. En este caso, tanto la función de costos como las restricciones pueden ser dependientes del tiempo: $f(\mathbf{x}, \mathbf{y}, t)$, $h(\mathbf{x}, \mathbf{y}, t)$ y $g(\mathbf{x}, \mathbf{y}, t)$.

Los problemas de optimización que envuelven variables continuas y discretas, se encuentran en numerosas aplicaciones en el campo de la ingeniería de procesos por lo que han sido ampliamente estudiados (Floudas, 1995; Grossmann, 1996). Son problemas difíciles de resolver debido a su naturaleza combinatoria y por la presencia de no convexidades y no linealidades que se identifican tanto en el dominio de las variables continuas como en el de las enteras (Floudas y Grossmann, 1995; Adjiman et al., 1997).

Se han desarrollado un buen número de técnicas de *programación no lineal mixta entera* (MINLP) para resolver estos problemas de manera global. Los más importantes son por ejemplo la Descomposición Generalizada de Benders (“Generalized Benders Decomposition” GBD) (Geoffrion, 1972), el método de Ramificación y Acotamiento (“Branch and Bound” B&B) (Gupta and Ravindran, 1985), el algoritmo de Aproximación Externa (“Outer Approximation” OA) y sus variantes (Duran y Grossmann, 1986; Kocis y Grossmann, 1987; Viswanathan y Grossmann, 1990), y el algoritmo del Plano Cortante Extendido (“Extended Cutting Plane” ECP) (Westerlund y Petterson, 1995). Estos métodos, gozan de una amplia

base teórica (Grossmann, 2002) y han sido estudiados y evaluados en distintos casos de estudio relacionados con la síntesis y diseño de procesos (Grossmann, 1990; Adjiman et al., 1997, Grossmann et al., 1999). Sin embargo, debido a la complejidad de los problemas de optimización mixta entera relacionados con procesos químicos y al continuo avance de los sistemas de cómputo, ésta sigue siendo un área de investigación activa.

En otras contribuciones se propone también la *programación disyuntiva* (Balas, 1985) para problemas de optimización mixta entera, en lugar de la tradicional *MINLP*. Esta formulación conocida como Programación Disyuntiva Generalizada (*GDP*) (Raman y Grossmann, 1994) hace uso de restricciones algebraicas, disyunciones y proposiciones lógicas para mantener la estructura lógica del problema y explotarla para resolverlo eficientemente. Para aprovechar los algoritmos de optimización existentes se han propuesto técnicas que requieren la reformulación del problema en *MINLP* relajados, como la caja conexa (“*Convex Hull Reformulation*”) (Grossmann y Ruiz, 2009). También se han planteado variantes del método de ramificación y acotamiento (Lee y Grossmann, 2000) y del de aproximación externa (Turkay y Grossmann, 1996) que aprovechan la estructura lógica que se logra con la *programación disyuntiva*.

Cuando los problemas de diseño incluyen consideraciones dinámicas, se trata de problemas de *Optimización Dinámica No-Lineal Mixta-Entera (MIDO)*. La solución de los *MIDO* ha sido abordada en la literatura con propuestas que destacan por sus elegantes formulaciones matemáticas y por sus propiedades teóricas. Desde el punto de vista de optimización, se reconocen dos métodos distintos para abordar la solución de los problemas de *Optimización Dinámica No-Lineal Mixta-Entera (MIDO)*. Una corriente (Chatzidoukas et al., 2003) propone la descomposición en una secuencia de problemas primal y maestro utilizando los métodos de Descomposición Generalizada de Benders (*GBD*) o de Aproximación Exterior (*OA*). La otra corriente (Biegler et al., 2002) propone la discretización completa de los estados y variables de control de manera de resolver simultáneamente el modelo en ecuaciones diferenciales y el problema de optimización como un problema de *programación no lineal mixta entera (MINLP)* de mayor dimensión cuya solución se obtiene usando cualquiera de los métodos antes mencionados (*B&B*, *GBD*, *OA*).

En el área de diseño integrado se han desarrollado algoritmos para abordar los *MIDO*, aunque envuelven procedimientos de resolución bastante complicados. En Mohideen et al. (1997), Kookos y Perkins (2001), Bansal et al. (2002) y Sakizlis et al. (2004) se proponen *métodos secuenciales o de descomposición* que requieren relajaciones del problema y/o resoluciones de secuencias de problemas de optimización no lineal (*NLP*) para llegar a una solución. Por ejemplo, en el método propuesto por Sakizlis et al. (2004) el *MIDO* se descompone en un problema *primal* de control óptimo donde se optimizan las variables continuas para generar una cota superior y un problema *maestro*, mixto-entero lineal (*MINL*) calculado a partir del resultado del problema *primal*, donde se optimizan las variables discretas para generar una cota inferior. Para la solución se sigue un procedimiento iterativo resolviendo el *MIDO* repetidas veces hasta que ambas cotas se cruzan. En Bahri et al. (2007) se propone la *discretización completa* para la solución del problema obteniéndose un problema de dos etapas, en la primera se resuelve el diseño estructural y en la segunda, se evalúa el desempeño dinámico en función de la máxima violación de las restricciones para el escenario de perturbaciones dado. Esta estrategia es complementada en Flores-Tlacuahuac y Biegler (2007) y Flores-Tlacuahuac y Biegler (2008) quienes introducen la técnica de programación disyuntiva para mejorar el desempeño del algoritmo.

Ambas metodologías exhiben un buen desempeño en la resolución de los *problemas de diseño y control simultáneo* (Kookos y Perkins, 2001; Bansal et al., 2002; Sakizlis et al., 2004; Bahri et al., 1997; Flores-Tlacuahuac y Biegler, 2007; y Flores-Tlacuahuac y Biegler, 2008). Sin embargo, es evidente la complejidad de las formulaciones teóricas, además, la descomposición o relajación en problemas de optimización no lineal *NLP* ocasiona inconvenientes relacionados con la dimensión de estos problemas (Flores-Tlacuahuac y Biegler, 2007) y representa un gran esfuerzo computacional cuando las variables enteras se encuentran lejos de su valor óptimo (Exler et al., 2008). Por otra parte, al utilizar métodos de optimización basados en gradiente para resolver los problemas *NLP*, pueden ser necesarias condiciones de convexidad y continuidad y puede haber dependencia de los puntos iniciales. En efecto, en algunos trabajos se han reportado dificultades y convergencia a mínimos locales debido a la naturaleza combinatoria y las características no lineales y no convexas de los problemas de optimización mixta-entera resultantes (Narraway y Perkins, 1994; Costa y Oliveira; 2001)

En general, como señalan Caballero et al. (2007) una de las grandes limitaciones cuando se resuelven problemas no lineales (*MINLP*, *MIDO*, *NLP*) es que en la mayor parte de las aplicaciones prácticas los problemas resultantes son no-convexos. En muchos de los casos los algoritmos propuestos sólo pueden garantizar un mínimo local y en algunos casos no son capaces de encontrar un punto factible.

Los *métodos globales de optimización* son una opción atractiva para este tipo de problemas. Entre ellos, los *métodos deterministas*, que proponen técnicas como los algoritmos de ramificación y acotamiento (Horst y Tuy, 1987), aproximaciones exteriores (Horst et al., 1992), métodos primal-dual (Floudas y Viswewaran, 1990) y métodos de plano cortante (Tuy et al., 1985), constituyen herramientas poderosas. En la literatura se encuentran distintas propuestas para la resolución de problemas *MINLP* y optimización dinámica de manera global (Ryoo y Sahandis, 1995; Esposito y Floudas, 2000; Tawarmelani y Sahandis, 2002; Papamichail y Adjiman, 2002; Banga et al., 2003) pero la rigurosidad en el tratamiento de las no convexidades involucra un gran esfuerzo computacional.

Por otra parte existen los *métodos estocásticos de optimización* (como algoritmos genéticos, recocido simulado o “simulated annealing”, búsqueda tabú, quimiotactismo bacteriano) que, dada su facilidad de implementación y versatilidad, han ido ganando popularidad en la medida en que los investigadores han reconocido el costo y las debilidades de los métodos de optimización y búsquedas tradicionales. Se basan en la imitación del comportamiento de algunos fenómenos naturales y son métodos libres de derivadas lo que les permite enfrentar problemas con discontinuidades, ruido y dinámicas complicadas. Tienen la ventaja de que su implementación es muy sencilla y es necesario poco conocimiento previo del problema a resolver (Caballero y Grossman, 2007), además, al basarse en reglas heurísticas es relativamente fácil adaptarlos a aplicaciones particulares.

Estos métodos han sido utilizados con buenos resultados en la optimización con variables discontinuas y discretas, aun cuando su capacidad para resolver problemas con restricciones ha sido ampliamente discutida (Summanwar et al., 2002). En el área de ingeniería de procesos se encuentran trabajos donde se demuestra el buen desempeño de los métodos estocásticos en problemas de optimización no lineal mixta entera como Androulakis y Venkatasubramanian,

(1991), Wang et al. (1998), Lewin et al. (1998), Costa y Oliveira (2001), Tsai and Chang (2001), Lin y Miller (2004), Lewin y Parag (2003) y Exler et al. (2008), entre muchos otros.

Los *algoritmos genéticos*, o denominados de manera más general *algoritmos evolutivos*, parecen ser los métodos estocásticos preferidos por la mayoría de los autores en el área de ingeniería química y control de procesos. Androulakis and Venkatasubramanian (1991), Wang et al. (1998), Tsai and Chang (2001) utilizan algoritmos genéticos para resolver el problema de la síntesis estructural, Lewin y Parag (2003) para la selección de la estructura y síntesis automática de esquemas de control descentralizado, Beraud et al. (2008) en la optimización de la estrategia de control de una planta de tratamiento de aguas. En Summanwar et al. (2002) se propone un *algoritmo genético multi-objetivo* y se evalúa su desempeño en la resolución de algunos problemas de *optimización mixta entera* y de *optimización dinámica* en ingeniería química. Sin embargo, no se han encontrado aplicaciones de algoritmos genéticos para enfrentar los problemas completos de diseño integrado, incluyendo la síntesis y/o selección de la estructura de control. Se considera que estos algoritmos son una buena alternativa a las técnicas de optimización tradicionales, principalmente porque son fáciles de realizar, y aun cuando requieren de un gran número de evaluaciones de la función objetivo para llegar a un resultado satisfactorio, el gasto computacional es menor que en los métodos deterministas.

En resumen, con la revisión de los estudios relacionados con el *diseño integrado de procesos* distinguiendo las distintas vertientes en el tratamiento del problema, se tienen criterios para establecer cuáles son las áreas que requieren mayor atención, a saber, la conciliación de la información obtenida a partir de los distintos métodos para evaluar la *controlabilidad*, la implantación de técnicas avanzadas de control, el desarrollo de métodos numéricos eficientes para la resolución de los problemas de diseño integrado y otras propuestas alternativas para tratar los aspectos más conflictivos . En base a esto se plantean los objetivos de este trabajo.

3. OBJETIVOS Y CONTENIDOS DEL TRABAJO

El objetivo general de esta tesis es el desarrollo de algoritmos genéticos como alternativa a los métodos tradicionales de optimización, para la solución de problemas *de síntesis y diseño*

integrado de procesos químicos, ahondando al mismo tiempo en diversas extensiones o aspectos novedosos de estos problemas.

En un marco muy general, se contempla la formulación del problema incluyendo la síntesis del proceso, selección de la estructura de control, su dimensionamiento y diseño del sistema de control, en donde una *superestructura* de la planta que representa todas las alternativas estructurales, se traduce en un problema de *optimización dinámica no lineal mixta entera*. A partir de este planteamiento general, se presentan casos particulares del problema en estudio para abordar distintas facetas del diseño integrado.

El algoritmo genético especialmente desarrollado usa una codificación mixta, con genes reales y binarios para tratar las variables continuas y discretas del *problema*. Además, presenta características particulares para mejorar su eficiencia, por la incorporación de técnicas heurísticas basadas en conocimiento.

Entonces, se plantean como objetivos específicos de esta investigación los siguientes:

- a) Resolución del problema de *diseño simultáneo del proceso y control* utilizando el algoritmo genético, considerando un objetivo económico y restricciones sobre índices de desempeño dinámico del proceso frente a un escenario realista de perturbaciones e índices basados en el modelo lineal.
- b) Un análisis sistemático de la controlabilidad en el marco del problema, donde se complementa el análisis de desempeño dinámico basado en el modelo no lineal, con la evaluación de métricas de *controlabilidad de estado* y medidas de controlabilidad basada en índices lineales.
- c) La aplicación de esquemas avanzados como los controladores predictivos basados en modelos (MPC).
- d) La modificación del algoritmo genético para enfrentar el planteamiento del problema con múltiples objetivos.

Para la evaluación de la metodología de diseño integrado, en este trabajo se utilizan como casos de estudio un sistema de reactores tipo tanque agitado (Luyben, 1993), una torre

de sulfitación perteneciente al proceso de clarificación de caña de azúcar (modelada en Alvarez et al. 2008) y el proceso de fangos activados de una planta de tratamiento de aguas residuales. Este último caso de estudio es el de mayor interés, debido a la complejidad de los procesos biológicos involucrados y a la necesidad de herramientas que permitan optimizar y mejorar la operabilidad de las plantas de tratamiento de aguas en el general de los procesos relacionados con la protección del ambiente. Se han utilizado dos modelos para representar el proceso de fangos activados, uno muy simple enfocado en una planta particular (Moreno et al., 1992) y uno generalizado basado en el Modelo del Proceso de Fangos Activados nro. 1 (ASM1) propuesto por la *International Water Association*.

Resultados relacionados con distintos aspectos de este trabajo han sido publicados. Los algoritmos genéticos se implementan con éxito en la resolución de problemas de diseño y control simultáneo con esquemas de control PI para el proceso de fangos activados simplificado en Revollar et al. (2004), Revollar et al. (2005), Francisco et al. (2005) y para el sistema de reactores tipo tanque agitado en Revollar et al. (2006). En Lamanna et al. (2009) se utiliza como caso de estudio la torre de sulfitación perteneciente al proceso de clarificación de caña de azúcar.

En Francisco et al. (2005) y Revollar et al. (2010) se incluyen comparaciones del comportamiento de los algoritmos genéticos con métodos numéricos clásicos y con otros métodos estocásticos.

En Revollar et al. (2008) y Francisco et al. (2009) se introducen técnicas avanzadas de control (MPC) en el marco del problema de diseño integrado del proceso de fangos activados simplificado.

En Lamanna et al. (2009) y en Revollar et al. (2010) se introduce el análisis de *controlabilidad de estado* en el marco del diseño integrado de una torre de sulfitación, considerando una estructura fija del proceso, con una estrategia de control PI y con un controlador predictivo no lineal, respectivamente. En este último se propone una formulación multiobjetivo.

En Revollar et al. (2010) se aplica la metodología, con una formulación multiobjetivo, a un proceso más complejo de fangos activados con un controlador MPC, considerando una estructura y dimensiones fijas del proceso.

La organización de este documento de tesis comprende seis capítulos. Después de la introducción general al tema y una revisión del estado de arte en el Diseño Integrado de procesos, presentada en el capítulo I, el capítulo II describe las distintas técnicas utilizadas para analizar la controlabilidad. En el capítulo III se discute respecto a la interacción entre el diseño y el control de procesos y se presenta la formulación general del problema de diseño integrado. El capítulo IV está dedicado a los algoritmos genéticos y su aplicación en la resolución de problemas de optimización mixta-entera. En el capítulo V se presentan los resultados de la aplicación de la metodología de diseño integrado a una torre de sulfitación perteneciente al proceso de clarificación de caña para la producción de azúcar. El capítulo VI corresponde finalmente a los resultados de diseño integrado para el proceso industrial de fangos activados en diferentes escenarios. Para terminar se presentan conclusiones y propuestas para futuros trabajos.

CAPÍTULO II

LA CONTROLABILIDAD DE LOS PROCESOS

1. INTRODUCCIÓN

Los procesos son diseñados para operar en estado estacionario en un punto de trabajo económico óptimo. Existen factores externos como perturbaciones, cambios en la demanda y en las condiciones ambientales que tienden a desviar el sistema de este punto de trabajo óptimo, afectando la operación económica de la planta. La operabilidad es la habilidad de un proceso para adaptarse a estos cambios. Esta viene dada por aspectos como la flexibilidad de la planta, que es la posibilidad de alcanzar el estado estacionario para un conjunto de distintos puntos de operación, el diseño del sistema de control y su *controlabilidad* (Schweiger y Floudas, 1997).

El control puede ser una tarea fácil en algunos casos, pero en otros son necesarias técnicas avanzadas para lograr un desempeño aceptable, incluso, existen procesos que no pueden ser controlados de ninguna manera. El diseño del proceso y las condiciones de operación seleccionadas tienen una influencia importante sobre su comportamiento dinámico. Si una planta se diseña basándose únicamente en criterios económicos y de operación en estado estacionario es posible que presente una dinámica desfavorable que dificulte el funcionamiento de cualquier sistema de control (Bahri et al, 1996). Por esta razón, la estrategia más conveniente para garantizar un desempeño dinámico satisfactorio, es incluir la *controlabilidad* como uno de los objetivos a optimizar en el planteamiento del problema de diseño de la planta.

La *controlabilidad* es una propiedad que mide la capacidad de un proceso para cumplir los objetivos de control a pesar de las perturbaciones externas y la incertidumbre en los parámetros de diseño; es una característica inherente al proceso y es independiente de la

selección de un controlador en particular (Luyben y Floudas, 1994). También se asocia la idea de *controlabilidad* con la habilidad de una planta para alcanzar y mantener la condición de operación en estado estacionario deseada.

En la formulación del problema de diseño de plantas, la definición de los objetivos económicos es directa, sin embargo el análisis de *controlabilidad* puede interpretarse de distintas maneras y su evaluación es una tarea difícil cuando se trata de procesos químicos que suelen ser sistemas complejos, no lineales y multivariados. Se dispone de numerosos métodos de evaluación de la *controlabilidad* con distintos requerimientos de información y de esfuerzo computacional, sin embargo, existe cierto desconocimiento respecto a la manera de utilizarlos y entenderlos (Morari, 1992). Además, la evaluación de los índices de *controlabilidad* en la fase inicial del diseño depende de información del proceso que aún se desconoce y muchos de estos índices están formulados en dominio frecuencial y no en el temporal en el que se enmarca el problema de diseño.

En la literatura relacionada con el diseño de sistemas de control el término *controlabilidad* es muy utilizado, aunque puede referirse a conceptos distintos. En la teoría de sistemas lineales se maneja el concepto de *controlabilidad de estado*, basado en el modelo matemático en espacio de estado. Mientras que en el área de diseño y control de procesos químicos, se ha desarrollado y utilizando el concepto de *controlabilidad* como medida del desempeño dinámico entrada-salida del sistema.

Los conocidos conceptos de *controlabilidad* y *observabilidad* de estado, introducidos por Kalman (1960), se basan en definiciones numéricas que a partir del modelo de estados del proceso, determinan la existencia de una solución completa para el problema de diseño del sistema de control. La *controlabilidad de estado* determina si es posible construir una señal de control sin restricciones, capaz de transferir el vector de estados de un estado inicial a un estado final en un tiempo finito. La *observabilidad* del sistema indica si todas las transiciones del estado afectan eventualmente a todos los elementos del vector de salida. Se dice que un sistema es completamente controlable si todos los estados son controlables y análogamente, un sistema es observable si es posible la observación de todos los estados (Ogata, 2003).

En los primeros trabajos orientados al estudio del compromiso entre el diseño del proceso y su capacidad de ser controlado se critica la aplicabilidad del concepto de *controlabilidad* de Kalman. Por ejemplo, Morari y Stephanopoulos (1980) indican que a partir del estudio de la *controlabilidad de estado* no es posible imponer acotamiento de las variables manipuladas, y tampoco, deducir cómo debe hacerse el control del proceso. Chenery (1997) completa esta crítica señalando que una vez definidos el estado inicial y final del sistema, la *controlabilidad de estado* no aporta información que permita imponer ninguna condición sobre la trayectoria entre ambos puntos. Sin embargo, está claro que la eficiencia del control es total cuando todos los estados del proceso son controlables (aun cuando esto sea difícil de lograr en la práctica), y desde esta perspectiva, el análisis de *controlabilidad y observabilidad de estado* es importante para incluir el control total de estados como un objetivo en el procedimiento de diseño.

El uso del término *controlabilidad* como una propiedad que mide la capacidad de la planta para ser controlada fue introducido inicialmente por Ziegler y Nichols (1943). Ellos la definen como “*la habilidad de un proceso de alcanzar y mantener el valor de equilibrio deseado*”, entendiéndose esta propiedad como una medida del buen desempeño de la planta. Estos autores presentan, además, el “*factor de recuperación*” para clasificar procesos en términos de *controlabilidad*, independientemente del controlador utilizado.

Más adelante, Rosembrock (1970) presenta el concepto de *controlabilidad funcional*, que establece si, para un sistema, existe un vector de trayectorias de entrada capaz de generar, desde la condición inicial, una trayectoria de salida que satisfaga ciertas condiciones de continuidad. Se dice que un proceso es *funcionalmente controlable* si el rango de la matriz de transferencia de la planta es igual al número de entradas. En el caso particular de plantas representadas por una matriz de transferencia cuadrada, la *controlabilidad funcional* se cumple si ésta es invertible. El análisis propuesto, indica la posibilidad o no de alcanzar un desempeño de control perfecto.

Morari (1983) propone estudiar del desempeño dinámico del sistema independientemente del controlador, ya que esta característica tiene que ver con el efecto de las limitaciones propias del sistema sobre la calidad del control. En este análisis, además de los aspectos antes

mencionados, se contemplan los cambios que se podrían efectuar sobre el proceso para mejorar la eficiencia del control. Para diferenciar esta idea de la *controlabilidad de estado*, se introduce el término “*resiliencia*” dinámica (“*dynamic resilience*”). Finalmente, se adopta la denominación de *controlabilidad entrada-salida* para distinguir este concepto.

La *controlabilidad entrada-salida* se entiende entonces como la habilidad de una planta de cumplir los objetivos de control, lo cual significa mantener las salidas dentro de ciertos límites alrededor de la referencia a pesar de las perturbaciones, ruido e incertidumbre del modelo, usando las entradas y mediciones disponibles (Skogestad, 1994). Desde este punto de vista, la *controlabilidad* es una propiedad inherente al proceso e independiente del controlador seleccionado, que abarca aspectos como la severidad de la interacción entre las variables, la capacidad de rechazar perturbaciones y la capacidad de moverse de un punto de operación a otro (Skogestad, 1994; Luyben y Floudas, 1994).

En este trabajo, es interesante el análisis de *controlabilidad entrada-salida* dado que establece una fuerte relación entre el diseño del proceso y su desempeño de control. Aunque se trata de un concepto cualitativo, se dispone de un buen número de indicadores para estimar la *controlabilidad entrada-salida*, los cuales típicamente se utilizan en el diseño integrado. En este capítulo se realiza, en la sección 2, una revisión acerca del concepto de *controlabilidad*. En las secciones siguientes se tratan detalle las técnicas para evaluar la *controlabilidad* en base a los modelos lineales y no lineales de la planta (secciones 3 y 4) y tanto desde el punto de vista de *controlabilidad entrada-salida* como de *controlabilidad de estado*, ya que ambos enfoques se consideran complementarios y se incluirán en la propuesta de diseño.

2. ANALISIS DE CONTROLABILIDAD BASADO EN EL MODELO LINEAL

Entre las técnicas para evaluar la *controlabilidad* se distinguen principalmente los métodos basados en el modelo dinámico lineal del proceso y los que se basan en el modelo dinámico no lineal. El análisis de *controlabilidad* será acertado en la medida que el modelo utilizado sea una buena representación del proceso. Si el proceso presenta características lineales alrededor del punto de operación, es suficiente la evaluación de la *controlabilidad* mediante índices basados en el modelo lineal, obteniéndose resultados significativos con poco esfuerzo computacional. Sin embargo, en algunos casos, especialmente cuando se trata de

procesos químicos complejos, puede ser necesaria la aplicación de técnicas basadas en el modelo no lineal, el cual resulta mucho más costoso computacionalmente (Morari, 1992; Chenery, 1997).

Para el caso lineal, considérese un sistema en tiempo continuo con múltiples entradas y múltiples salidas, dado por el siguiente modelo de estado:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A} \cdot \mathbf{x} + \mathbf{B} \cdot \mathbf{u} \\ \mathbf{y} &= \mathbf{C} \cdot \mathbf{x} + \mathbf{D} \cdot \mathbf{u}\end{aligned}\tag{2.1}$$

donde:

$\mathbf{u} = [\mathbf{u}_c \quad \mathbf{d}]$ es un vector de dimensión m que contiene todas las entradas,

\mathbf{d} es el vector de las perturbaciones de orden p ,

\mathbf{u}_c es el vector de las entradas manipuladas de orden $m-p$

\mathbf{x} es el vector de estados de dimensión n

\mathbf{y} es el vector de salida de dimensión r

\mathbf{A} ($n \times n$) es la matriz de estados

\mathbf{B} ($n \times m$) es la matriz de entrada

\mathbf{C} ($r \times n$) es la matriz de salida

\mathbf{D} ($r \times m$) es la matriz de salida directa.

Esta realización en espacio de estado (\mathbf{A} , \mathbf{B} , \mathbf{C} , \mathbf{D}) se expresa en el dominio de Laplace como:

$$\mathbf{y}(s) = \mathbf{G}(s)\mathbf{u}_c(s) + \mathbf{G}_d(s)\mathbf{d}(s)\tag{2.2}$$

donde:

$\mathbf{G}(s)$ ($r \times m$) es la matriz de transferencia del proceso.

$\mathbf{G}_d(s)$ ($r \times p$) es la matriz de transferencia de perturbaciones,

Ambas representaciones se relacionan mediante la siguiente expresión:

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{B} \cdot \mathbf{u}(s)\tag{2.3}$$

El diagrama de bloques del sistema a lazo abierto se muestra en la figura 2.1.a. en la figura 2.1.b. se ha agregado una matriz $G_c(s)$ y lazos de retroalimentación simple para el control.

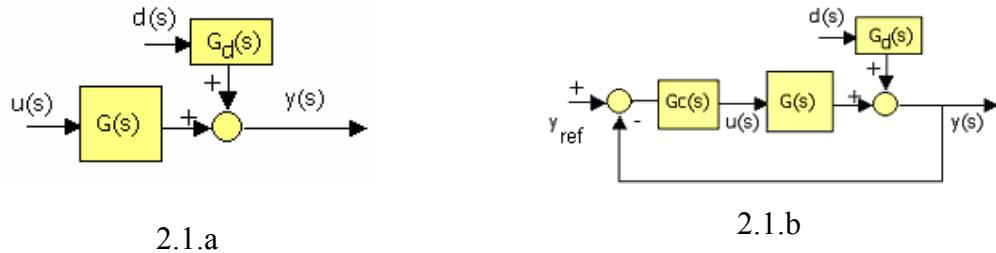


Figura 2.1. Diagrama de bloques del sistema en estudio. a. Lazo abierto b. Lazo cerrado

En este caso, el error de control es $e = [y_{ref} - y]^T$ y el objetivo de control es ajustar las variables de salida a los valores de referencia (y_{ref}) y rechazar las perturbaciones (d). Entonces, si no se considera ruido de medición, la ecuación entrada-salida para el sistema en lazo cerrado es:

$$y = (I + G \cdot G_c)^{-1} G \cdot G_c \cdot y_{ref} + (I + G \cdot G_c)^{-1} G_d \cdot d \quad (2.4)$$

Para este sistema, entonces, se pueden definir varias matrices de transferencia importantes:

- La *función de sensibilidad de la salida respecto de las perturbaciones* (Función de Sensibilidad) es:

$$S(s) = \frac{y(s)}{\tilde{d}(s)} = (I + G \cdot G_c)^{-1} \quad (2.5)$$

donde:

$$\tilde{d} = G_d \cdot d .$$

- La *función de sensibilidad de la salida respecto de la referencia* (Función de Sensibilidad Complementaria):

$$T(s) = \frac{y(s)}{y_{sp}(s)} = G \cdot G_c (I + G \cdot G_c)^{-1} \quad (2.6)$$

obsérvese que $\mathbf{S} + \mathbf{T} = \mathbf{I}$ y $\mathbf{T} = \mathbf{G} \cdot \mathbf{Gc} \cdot \mathbf{S}$

- La función de sensibilidad del control respecto a las perturbaciones:

$$\mathbf{M}(s) = \frac{\mathbf{u}_c(s)}{\tilde{\mathbf{d}}(s)} = (\mathbf{I} + \mathbf{G} \cdot \mathbf{Gc})^{-1} \mathbf{Gc} \quad (2.7)$$

La salida del sistema queda expresada en términos de estas funciones como:

$$\mathbf{y} = \mathbf{T} \cdot \mathbf{y}_{\text{ref}} + \mathbf{S} \cdot \mathbf{G}_d \cdot \mathbf{d} \quad (2.8)$$

Entonces, cuando aplica el control a lazo cerrado, la señal de control es:

$$\mathbf{u}_c = \mathbf{Gc} \cdot (\mathbf{y}_{\text{sp}} - \mathbf{y}) = \mathbf{Gc} \cdot (\mathbf{S} \cdot \mathbf{y}_{\text{ref}} - \mathbf{S} \cdot \mathbf{G}_d \cdot \mathbf{d}) \quad (2.9)$$

o bien, respecto a la Función de Sensibilidad Complementaria (\mathbf{T}):

$$\mathbf{u}_c = \mathbf{G}^{-1}(\mathbf{T} \cdot \mathbf{y}_{\text{sp}} - \mathbf{T} \cdot \mathbf{G}_d \cdot \mathbf{d}) \quad (2.10)$$

El análisis de *controlabilidad* basado en el modelo lineal, ya sea en espacio de estados o en matriz de transferencia, es útil para identificar las características del proceso que limitan su desempeño dinámico. En las secciones siguientes se resumen un conjunto de índices que permiten evaluar esta propiedad con una visión analítica y orientada al conocimiento del proceso.

2.1. CONTROLABILIDAD DE ESTADO

El análisis de *controlabilidad de estado* permite determinar si los estados del sistema son controlables a partir de las variables manipuladas elegidas, independientemente de la técnica de control a utilizar; además, permite obtener parámetros de diseño del equipo garantizando que el sistema específico es controlable.

Para un sistema lineal y multivariable representado por la ecuación (2.1) la condición necesaria y suficiente para la *controlabilidad completa del estado* es que los vectores \mathbf{B} , \mathbf{AB} , $\mathbf{A}^2\mathbf{B}, \dots, \mathbf{A}^{n-1}\mathbf{B}$ sean linealmente independientes, es decir, la *matriz de controlabilidad* de dimensión $n \times n$:

$$\mathbf{Co} = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} & \cdots & \mathbf{A}^n\mathbf{B} \end{bmatrix} \quad (2.11)$$

debe ser de rango completo (rango n) (Ogata, 2003). Si la *matriz de controlabilidad* (\mathbf{Co}) es cuadrada ($n \times n$), la condición para que el sistema sea controlable es que $\det(\mathbf{Co}) > 0$.

La *observabilidad completa* asociada al sistema (2.1) requiere que la *matriz de observabilidad* (2.12) de dimensión $n \times mxn$:

$$\mathbf{O} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \quad (2.12)$$

sea de rango n (Ogata, 2003).

La evaluación de la *controlabilidad de estado* tomando como base el modelo linealizado produce formulaciones matemáticas muy elegantes para el análisis del proceso. Sin embargo, algunos autores como Chenery (1997) indican que este análisis es relevante principalmente en el caso de plantas inestables, donde es recomendable verificar que los estados inestables sean controlables y observables. En otras condiciones, no aporta mayor información sobre la facilidad de controlar el proceso, ni establece restricciones sobre la trayectoria entre los estados iniciales y finales.

2.2. CONTROLABILIDAD ENTRADA-SALIDA

El desarrollo de técnicas para evaluar controlabilidad entrada-salida ha sido de gran interés en la ingeniería y control de procesos. Sobre este tema, se encuentra un buen número de trabajos (Morari, 1983; Holt y Morari, 1985; Skogestad y Morari, 1987; Skogestad et al, 1991; Wolff y Skogestad, 1992; Skogestad, 1995) dedicados al estudio de los factores que afectan el comportamiento dinámico de los procesos para establecer relaciones que permitan cuantificar su controlabilidad. Esta no es una tarea sencilla, dado que el concepto de controlabilidad es muy amplio y contempla distintas características de comportamiento del proceso.

Una idea bastante aceptada es asociar la *controlabilidad funcional* y la *resiliencia dinámica* con la posibilidad de lograr el *control perfecto* del proceso (Ekawati, 2003), siendo este a su vez posible si la planta es invertible. En efecto, un controlador ideal permite obtener

un seguimiento a la referencia y rechazo a perturbaciones perfecto si se expresa la señal de control en función de la inversa de la planta (Morari, 1983):

$$\mathbf{u} = \mathbf{G}^{-1} \cdot \mathbf{y}_{\text{ref}} + \mathbf{G}^{-1} \mathbf{G}_d \cdot \mathbf{d} \quad (2.13)$$

En la práctica, para un sistema retroalimentado se puede lograr algo similar si la función de transferencia del controlador (\mathbf{Gc}), se aproxima a la inversa de la planta, lo cual es posible si $\mathbf{S} \approx 0$ y $\mathbf{T} \approx I$ (Ver ec. 2.7).

En términos de la respuesta frecuencial, es posible aproximarse al control perfecto ($\mathbf{S} \approx 0$ y $\mathbf{T} \approx I$) cuando la magnitud de la ganancia de lazo es mayor que uno ($\|\mathbf{G} \cdot \mathbf{Gc}(jw)\| > 1$) lo cual ocurre para frecuencias inferiores a la del ancho de banda ($w < w_B$). Por el contrario, para frecuencias superiores a w_B se tiene que $\|\mathbf{G} \cdot \mathbf{Gc}(jw)\| \ll 1$, entonces $\mathbf{S} \approx I$ y $\mathbf{T} \approx 0$ y la respuesta no se verá afectada por la retroalimentación. En conclusión, para el control ideal de la planta (inversión) se requiere un ancho de banda alto, que se traduce en una alta ganancia a lazo cerrado, y rapidez de control (Wolf et al, 1992a; Skogestad y Wolff, 1992).

Existen algunas características estructurales de la planta a lazo abierto que reducen el ancho de banda y afectan la inversión son los *polos en el semiplano derecho* de la matriz $\mathbf{G}(s)$, como los *retardos de transporte* y los *ceros de fase no mínima* (Pedersen et al, 1998).

Los *retardos de transporte* y los *ceros en el semiplano derecho*, son elementos de fase no mínima que producen características indeseadas en la inversa de la planta haciéndolas inherentemente difíciles de controlar, al presentar elementos que introducen limitaciones sobre el ancho de banda alcanzable a lazo cerrado y disminuyen los márgenes de estabilidad. Los *retardos de transporte* se convierten en elementos predictivos no realizables en $\mathbf{G}^I(s)$. De la misma manera, los *ceros en el semiplano derecho* de $\mathbf{G}(s)$ se convierten en polos inestables de la inversa de la planta (Wolff, 1994; Chenery, 1997)

El desempeño dinámico del sistema también se ve afectado por otros factores asociados con la implementación del lazo de control como la *incertidumbre del modelo*, que puede producir dificultades e inexactitud en la obtención de la inversa de la planta, y las

restricciones sobre las variables manipuladas, ya que la planta deja de ser invertible si estas son violadas debido a la magnitud de las perturbaciones y/o cambios de referencia.

Por otra parte, las decisiones estructurales del proceso, como la selección de los posibles pareamientos entre las variables manipuladas y las variables controladas, son de gran importancia dentro del análisis de *controlabilidad*, ya que la compatibilidad entre éstas y la posible interacción con otras variables y lazos de control, influye significativamente sobre la posibilidad de lograr los objetivos de control.

El estudio de las características del proceso que limitan el *control perfecto* y el análisis de la interacción y direccionalidad entre las variables de entrada, salida y perturbaciones, han permitido el desarrollo de un buen número de índices para cuantificar la *controlabilidad*. Los procedimientos de análisis propuestos en la literatura, contemplan la evaluación de un conjunto de medidas para examinar los distintos aspectos de esta propiedad (Hovd y Skogestad, 1992; Wolff et al, 1992a; Wolff et al, 1992b; Wolff, 1994). A continuación se presentan las más utilizadas en el *diseño integrado de procesos*. Aunque en primer lugar, es necesario mencionar que para una correcta aplicación de los índices de controlabilidad que se describen, es crucial un escalamiento adecuado de las variables. Las variables deben ser escaladas para tomar valores en el intervalo [-1,1], por lo que las distintas magnitudes deben ser normalizadas para que sean inferiores a 1 para cualquier frecuencia. Wolff et al (1992) recomiendan normalizar tanto las entradas y salidas como las perturbaciones respecto a su rango permitido, lo cual implica escalar las matrices \mathbf{G} y \mathbf{G}_d .

a) *Ubicación de los polos de $\mathbf{G}(s)$ y estabilidad.*

La ubicación de los polos de $\mathbf{G}(s)$ aporta información general respecto a características de respuesta transitoria y la estabilidad del sistema en lazo abierto como es ampliamente conocido en la teoría de control clásico.

Un sistema se considera estable si, al ser excitado con una entrada acotada, su salida es también acotada. La estabilidad de un sistema se puede evaluar mediante el estudio de la ubicación de los autovalores (λ) de la matriz de estados \mathbf{A} o los polos (p) de la función de transferencia $\mathbf{G}(s)$. Se dice que el sistema es estable si $\lambda_i \leq 0$, $i=1, 2, 3$.

Si $\mathbf{G}(s)$ presenta un polo p inestable se impone un límite inferior al ancho de banda, ya que la frecuencia de cruce debe ser menor a la frecuencia del polo para asegurar una respuesta estable.

b) *Ganancias y análisis de los valores singulares.*

La *ganancia* se define como el cambio global en las salidas por unidad de cambio en las entradas, esta puede relacionarse con el comportamiento temporal como ganancia estática o instantánea o en función de la respuesta frecuencial. Además, en el caso de sistemas multivariados la ganancia depende de la combinación de los componentes de la entrada, es decir, de la dirección de la misma. En este caso, la ganancia global del sistema se calcula mediante la norma inducida de la matriz $\mathbf{G}(jw)$:

$$\|\mathbf{G}(jw)\| = \max_{\|\mathbf{u}\| \neq 0} \frac{\|\mathbf{G}(jw) \cdot \mathbf{u}(jw)\|}{\|\mathbf{u}(jw)\|} = \max_{\|\mathbf{u}\|=1} \|\mathbf{G}(jw) \cdot \mathbf{u}(jw)\| \quad (2.14)$$

Por otra parte, los *valores singulares* de una matriz \mathbf{G} de orden $n \times n$ son una medida de lo cerca que está la matriz de ser singular, es decir, de tener un determinante igual a cero. Se definen como:

$$\sigma_i = \sqrt{\lambda_i(\mathbf{G}^T \cdot \mathbf{G})} \quad i=1,2,\dots,n \quad (2.15)$$

donde:

λ : autovalores de la matriz (\mathbf{G})

Cualquier matriz $\mathbf{G}(s)$, puede descomponerse en sus valores singulares de la siguiente manera:

$$\mathbf{G} = \mathbf{U} \Sigma \mathbf{V}^H \quad (2.16)$$

donde:

Σ es una matriz diagonal con los valores singulares $\sigma_i(\mathbf{G})$ colocados de forma descendente a lo largo de la diagonal principal,

\mathbf{U} es una matriz unitaria de vectores singulares de salida u_i . Los vectores columna de \mathbf{U} representan las direcciones de salida de la planta, son ortogonales y de longitud unitaria (Ortonormales).

V es una matriz unitaria de vectores singulares de entrada v_i . Sus vectores columna son ortonormales y representan las direcciones de las entradas.

V^H es la matriz compleja conjugada transpuesta de V .

Los valores singulares, también llamados ganancias principales, indican la magnitud de la ganancia inducida por la matriz de transferencia en la transformación lineal $\mathbf{y} = \mathbf{G} \cdot \mathbf{u}$. Dado que V es una matriz unitaria ($V^H V = I$), la ecuación anterior se puede expresar como $\mathbf{G}V = \mathbf{U}\Sigma$ cuya columna i -ésima es $\mathbf{G}v_i = \sigma_i u_i$. Se puede observar que al considerar una entrada en la dirección v_i , la salida es en la dirección u_i y el valor singular σ_i es directamente la ganancia del sistema en esa dirección. Así, la descomposición en valores singulares en \mathbf{G} y \mathbf{Gd} es útil para establecer cuáles combinaciones de variables manipuladas tienen mayor efecto sobre la salida y qué perturbaciones producen las mayores variaciones (Wolf et al, 1992; Skogestad y Postlethwaite, 1996). Este análisis aplica también en el caso de plantas cuya matriz \mathbf{G} no es cuadrada.

c) *Máximo y mínimo valor singular.*

El *mínimo valor singular* $\sigma_{\min}(\mathbf{G})$ indica qué tan cerca está la matriz de ser singular y es la menor ganancia del proceso para todas las direcciones posibles de la entrada a una frecuencia dada. Se calcula como:

$$\sigma_{\min}(\mathbf{G}) = \|\mathbf{G}\|_2 = \min_{\|\mathbf{u}\|_2 \neq 0} \frac{\|\mathbf{G} \cdot \mathbf{u}\|_2}{\|\mathbf{u}\|_2} \quad (2.17)$$

El mínimo valor singular fue introducido como índice de *controlabilidad* por Morari (1983). Suponiendo que las variables están escaladas uniformemente, se recomienda que $\sigma_{\min}(\mathbf{G}) \geq 1$. Un valor grande de $\sigma_{\min}(\mathbf{G})$ significa que $\mathbf{G}(s)$ no está cercana a la singularidad y que el proceso es más resistente a perturbaciones. En cambio, valores pequeños de este parámetro indican que la planta requiere de variaciones de gran magnitud en las entradas manipuladas para un buen seguimiento a la referencia, lo que puede ocasionar saturación de los elementos de control (Skogestad y Postlethwaite, 1996).

El *máximo valor singular* $\sigma_{\max}(\mathbf{G})$ indica la mayor ganancia del proceso para todas las posibles direcciones de la entrada (\mathbf{u}) a una frecuencia dada. Se calcula a partir de la norma Euclidiana:

$$\sigma_{\max}(\mathbf{G}) = \|\mathbf{G}\|_2 = \max_{\|\mathbf{u}\|_2 \neq 0} \frac{\|\mathbf{G} \cdot \mathbf{u}\|_2}{\|\mathbf{u}\|_2} \quad (2.18)$$

Skogestad y Postlethwaite (1996) señalan que un *máximo valor singular* muy alto no representa un problema para el desempeño dinámico, aunque pueden producir mal funcionamiento de los actuadores y válvulas, por lo que se recomiendan sistemas con ganancias en estado estacionario inferiores a 50. Este fenómeno fue estudiado por Araujo y Skogestad (2006), quienes encontraron que una ganancia muy grande para frecuencias cercanas al ancho de banda a lazo cerrado puede producir problemas sobre desempeño a lazo cerrado.

d) *Número de condición.*

El *número de condición* de una matriz $\gamma(\mathbf{G})$ se define como la relación entre el máximo y el mínimo valor singular:

$$\gamma(\mathbf{G}) = \frac{\sigma_{\max}(\mathbf{G})}{\sigma_{\min}(\mathbf{G})} \quad (2.19)$$

y su valor deseable es cercano a la unidad.

Si $\gamma(\mathbf{G})$ es muy alto se considera que un sistema está mal condicionado, y ocurre que mientras el efecto de una combinación de variaciones en las entradas sobre la salida es mínimo, para otra combinación de variaciones en las entradas se produce un efecto significativo sobre la salida. Esto indica que la ganancia del sistema depende fuertemente de la dirección de la entrada, lo que dificulta el control (Wolff et al, 1992a).

En el caso de plantas con restricciones sobre las variables manipuladas, se requiere un número de condición pequeño y un $\sigma_{\min}(\mathbf{G})$ alto para evitar saturación de las entradas y lograr un buen desempeño dinámico.

El número de condición también indica la sensibilidad a la incertidumbre en el modelo o en las mediciones. Si el número de condición es pequeño ($\gamma(\mathbf{G}) \leq 2$) el sistema será poco sensible a la incertidumbre (Skogestad y Postlethwaite, 1996).

e) *Norma H_∞ del sistema.*

La norma H_∞ es otro índice de *controlabilidad* relacionado con el análisis de los valores singulares. Para cada frecuencia la matriz $G(jw)$ es distinta y por lo tanto los valores singulares $\sigma_i(w)$ son también distintos. La norma H_∞ del sistema es el máximo valor singular sobre todo el rango de frecuencias, es decir, es el máximo valor de la magnitud en la traza de Bode del máximo valor singular, y se define como:

$$\|\mathbf{G}(jw)\|_\infty = \max_w \sigma_{\max}(\mathbf{G}(jw)) \quad (2.20)$$

Físicamente, el valor de esta norma representa la máxima amplificación de la respuesta en la peor dirección de las entradas. Por esto, en el diseño de sistemas de control basado en la norma H_∞ se formula un problema de optimización donde se minimiza este valor.

f) *Sensibilidad a las perturbaciones y número de condición de la perturbación*

Las perturbaciones pueden ser rechazadas por la planta mientras su magnitud esté dentro de lo permitido por las restricciones de la variable manipulada. Esta limitación se asocia con la flexibilidad de la planta. La descomposición en valores singulares es útil para estudiar el efecto de las perturbaciones cuando existen restricciones sobre las entradas.

En sistemas multivariados una condición necesaria para lograr un control aceptable cuando las entradas están acotadas ($\|\mathbf{u}_{\max}\| \leq 1$), es que:

$$\sigma_i(\mathbf{G}) \geq |\mathbf{u}_i^H \cdot \mathbf{g}_d| - 1 \quad (2.21)$$

en las frecuencias en las que: $|\mathbf{U}_i^H \cdot \mathbf{g}_d| > 1$, y

\mathbf{u}_i^H es el i-ésimo vector singular de salidas de (Skogestad y Postlethwaite, 1996). La generalización $|\mathbf{u}_i^H \cdot \mathbf{g}_d| > 1$ puede interpretarse como la proyección de \mathbf{g}_d sobre el i-ésimo vector singular de salidas de la planta.

Otro criterio para evaluar el efecto de las perturbaciones tiene que ver con el análisis de los elementos de la matriz \mathbf{Gd} . Si se ha aplicado un escalamiento adecuado, la presencia de elementos mayores a uno en la matriz \mathbf{Gd} indica que es necesario aplicar alguna estrategia de control para reducir el efecto de las perturbaciones.

El número de condición de la perturbación indica qué tan grande debe ser la magnitud de la entrada (variable manipulada) para rechazar una perturbación unitaria, comparada con el efecto de la misma perturbación en la dirección que menos afecta la planta (Wolff et al, 1992b). El *número de condición de la perturbación* $\gamma_p(\mathbf{G})$ es presentado por Morari (1987):

$$\gamma_p(\mathbf{G}) = \frac{\|\mathbf{G}^{-1} \cdot \mathbf{g}_d\|_2}{\|\mathbf{g}_d\|_2} \sigma_{\max}(\mathbf{G}) \quad (2.22)$$

donde:

\mathbf{g}_d son las columnas de \mathbf{G}_d cuando se considera una perturbación única.

El *número de condición de la perturbación* también aporta información respecto al efecto de las restricciones sobre las entradas. Si este índice toma un valor muy grande sugiere que la entrada u puede sufrir un incremento muy grande para rechazar alguna perturbación y esto puede producir una violación en los límites de las variables manipuladas (Cheney, 1997). Un índice similar, denominado *Costo de la perturbación*, fue introducido por Weitz y Lewin (1996).

g) Controlabilidad estructural y matriz de ganancias relativas (RGA).

La *matriz de ganancias relativas* (RGA) (Bristol, 1966) mide la interacción entre las variables de entrada y salida en sistemas multivariados. Los elementos de la *matriz de ganancias relativas* (RGA) vienen dados por la relación entre la ganancia de un par entrada/salida, considerando todo el sistema a lazo abierto, y la ganancia del mismo par con todos los demás lazos funcionando perfectamente en automático.

La *matriz de ganancias relativas*, de una matriz \mathbf{G} cuadrada, es una matriz no singular, cuadrada, de dimensión $r \times r$. definida en estado estacionario como:

$$\mathbf{RGA}[\mathbf{G}(s)] = \mathbf{G}(s) \mathbf{x} [\mathbf{G}^{-1}(s)]^T \Big|_{s=0} \quad (2.23)$$

donde el operador \mathbf{x} indica la multiplicación elemento a elemento. La *RGA* requiere que la suma de los elementos de cada columna sea igual a 1.

El análisis de los elementos de la *RGA* permite identificar los mejores pareamientos entrada-salida: los elementos cercanos a uno en la RGA, indican que se puede escoger ese pareamiento sin que haya interacción con los otros posibles lazos de control. Si los elementos toman valores entre 0.5 y 1, se podría seleccionar ese pareamiento considerando que hay cierta interacción con los otros lazos, cuando tienen un valor de 0.5 significa que existe otra variable manipulada que influye de la misma manera sobre la variable controlada. Los elementos iguales a cero indican que, para esa combinación, la variable manipulada no afecta la variable controlada.

Las variables controladas y manipuladas que generen elementos negativos en la matriz de ganancias relativas, no deben ser emparejadas porque generan inestabilidades (Luyben y Floudas, 1994). Las plantas para las cuales se obtienen elementos de gran magnitud en la matriz de ganancias relativas (*RGA*) son difíciles de controlar porque presentan una alta sensibilidad a la *incertidumbre del modelo*.

La *RGA* es uno de los índices más utilizados para el análisis estructural del proceso, sin embargo, existen dudas sobre su eficiencia cuando se tienen más de dos pares de variables (Ekawati, 2003). Para enfrentar este problema se han propuesto medidas complementarias como el índice de Niederlinski (Niederlinski, 1971) y la matriz de ganancias relativas parciales (PRGA).

También se encuentran índices enfocados en interpretaciones distintas de la *controlabilidad estructural*. Para Lee et al. (2001), el concepto de *controlabilidad estructural* se relaciona con el recorrido de las perturbaciones a lo largo del proceso, lo que se puede determinar a partir de la información relativa a su estructura aunque se desconozcan mayores

detalles. Se dice que la *controlabilidad estructural* es buena si una perturbación indeseable no se propaga. En el trabajo citado, el modelo del proceso se describe como un “*dígrafo*” y se utilizan matrices de orden relativo para determinar el desacoplamiento del sistema. Mediante este criterio se pueden seleccionar los mejores diagramas de flujo y descartar las alternativas no controlables.

Las técnicas para evaluar la *controlabilidad* presentadas hasta aquí, se basan en el análisis de características particulares de la planta a lazo abierto, por lo cual, los índices de *controlabilidad* presentados son independientes del controlador. A continuación se presentan algunas técnicas para evaluar la *controlabilidad* del sistema a lazo cerrado, para las que se requiere del conocimiento del controlador. Este análisis se realiza para una estructura de control determinada, como la que se presenta en figura 2.1, donde \mathbf{G}_c es el controlador.

h) Polos y ceros del sistema a lazo cerrado.

Tradicionalmente se han utilizado los ceros y especialmente los polos del sistema a lazo cerrado para estudiar el comportamiento del sistema y entonces lograr con los controladores ubicación de los polos y ceros para establecer las características deseadas de respuesta a lazo cerrado. En esto se basan muchas de las técnicas de diseño de controladores, como por ejemplo el diseño por retroalimentación de estados o el diseño por modificación del lugar geométrico de las raíces.

i) Normas de las funciones de sensibilidad del sistema.

Las funciones de sensibilidad del sistema a lazo cerrado definidas en las ecuaciones (2.5) y (2.6), donde $\mathbf{S}(s) = (\mathbf{I} + \mathbf{G} \cdot \mathbf{G}_c)^{-1}$ y $\mathbf{T}(s) = \mathbf{G} \cdot \mathbf{G}_c (\mathbf{I} + \mathbf{G} \cdot \mathbf{G}_c)^{-1}$ respectivamente, proporcionan, mediante la aplicación de ciertas normas, valores numéricos objetivos acerca del comportamiento del sistema controlado en cuanto a estabilidad, seguimiento a la referencia y sensibilidad a cambios en el proceso, es decir, robustez (Skogestad y Postlewaite, 1996).

Retomando, la ecuación (2.8):

$$\mathbf{y} = \mathbf{T} \cdot \mathbf{y}_{\text{ref}} + \mathbf{S} \cdot \mathbf{G}_d \cdot \mathbf{d} .$$

se puede calcular la función del error del controlador:

$$\mathbf{e} = -\mathbf{S} \cdot \mathbf{y}_{\text{ref}} + \mathbf{S} \cdot \mathbf{G}_d \mathbf{d} \quad (2.24)$$

El cálculo del módulo de la función de sensibilidad permite conocer entonces la capacidad del sistema a lazo cerrado de eliminar errores, además de otros aspectos como robustez y sensibilidad al ruido.

Por ello se busca escoger el regulador (o el proceso, si se trata de un caso de diseño) que minimice el máximo valor de $|\mathbf{S}(jw)|$ para cualquier frecuencia. En el caso multivariable esto significa minimizar el máximo valor singular de la matriz de sensibilidad, es decir, la norma H_∞ de la función de sensibilidad:

$$\max_w \sigma_{\max} [\mathbf{S}(jw)] \quad (2.25)$$

lo que es equivalente a maximizar el margen de ganancia en el Diagrama de Nyquist (Puigjaner et al, 2006)

Para capturar las distintas condiciones que se exigen para la matriz \mathbf{S} , se puede definir un peso frecuencial Wp , e imponer la siguiente condición sobre el siguiente índice (basado en la norma H_∞)

$$\|Wp \cdot \mathbf{S}\|_\infty < 1 \quad (2.26)$$

Este índice se suele completar incluyendo también un límite superior sobre las magnitudes del control modelado por un peso Wu . Se define entonces el siguiente índice (de sensibilidad mixta) que considera ambos aspectos de forma conjunta, muy utilizado en diseño robusto de sistemas de control (Skogestad y Postlewaite, 1996):

$$\|N\|_\infty = \left\| \begin{pmatrix} Wp \cdot \mathbf{S} \\ Wu \cdot \mathbf{K} \cdot \mathbf{S} \end{pmatrix} \right\|_\infty \quad (2.27)$$

Los índices basados en la norma H_∞ son muy utilizados en control. Una aplicación interesante en el análisis de *controlabilidad* encuentra en Cao y Yang (2004), quienes proponen una metodología donde se plantea la reubicación de los polos de lazo cerrado para cumplir ciertos requerimientos de desempeño, como minimizar el error y los esfuerzos de

control. Esto se representa mediante un conjunto de desigualdades lineales matriciales (LMI), que se formula como un problema de optimización multiobjetivo. Si la solución de las LMI es factible existe al menos un controlador capaz de satisfacer las especificaciones de desempeño.

Otro bloque de índices está basado en la norma $L1$, tomada del análisis y diseño de sistemas de control robusto, en particular cuando existen restricciones. En concreto, es de interés la norma $\|M\|_1$, que es el valor máximo de la señal de control cuando se tiene la peor combinación de las perturbaciones afectando al sistema.

En resumen, como puede verse, se dispone de un buen número de indicadores de *controlabilidad* para evaluar distintas características de los sistemas lineales. Estos índices permiten obtener información valiosa respecto al comportamiento dinámico del proceso mediante cálculos sencillos, la cual es suficiente en la mayoría de los casos, donde se opera en un rango alrededor del punto de operación, en el cual el modelo lineal del sistema tiene validez (Morari, 1992). Sin embargo, existen algunos aspectos que han sido criticados en relación a la aplicación de los índices lineales de *controlabilidad*, incluso dentro de su rango de validez.

Chenery (1997) señala que se hacen suposiciones poco realistas en el cálculo de muchos de estos índices. Un ejemplo ocurre con la RGA que solo se aplica a plantas con matriz de transferencia cuadrada, es decir, con el mismo número de entradas y salidas, lo cual no siempre ocurre. O también, los índices que miden el efecto de las perturbaciones, porque éstas se consideran entradas acotadas.

Por otra parte, se ha criticado que muchos de los índices se manejan en el dominio frecuencial mientras que los requerimientos de desempeño del sistema se establecen en el dominio temporal, lo que ocasiona ciertas dificultades al interpretar los resultados de la evaluación. Por la misma razón, se hace difícil relacionar la información que aportan los índices con el sistema físico, por lo que no siempre se puede saber qué parámetros del diseño influyen sobre alguna característica desfavorable en la respuesta dinámica revelada por el análisis de *controlabilidad* (Chenery, 1997; Meeuse, 2000)

3. CONTROLABILIDAD DE ESTADO PARA EL MODELO NO LINEAL

Las técnicas para evaluar la *controlabilidad* descritas anteriormente son efectivas si las plantas pueden ser representadas por modelos lineales en un rango de operación lo suficientemente amplio. En algunos casos, la no linealidad del proceso es significativa, por lo que la planta exhibe un comportamiento dinámico mucho más complejo y se requieren de métodos más completos para evaluar la *controlabilidad*. Una característica importante de las plantas no lineales es que en los múltiples puntos en estado estacionario existentes dentro del rango de operación se pueden presentar comportamientos dinámicos muy distintos con diferentes características de estabilidad lo que produce efectos adversos para el control del proceso. Los valores de las variables de operación en los que se producen estos cambios se denominan *puntos de bifurcación*. Generalmente se evita operar la planta alrededor de éstos, e incluso, se sobredimensionan los diseños para evitar estos comportamientos (Pedersen et al, 1998). Sin embargo, el incremento en los estándares de calidad, las regulaciones ambientales y las exigencias económicas obligan a veces a crear diseños que operan en condiciones donde se presentan comportamientos *no lineales* que normalmente se evitarían (Morari, 1992). Por esta razón se ha despertado el interés en estudio de herramientas para evaluar la *controlabilidad* en las que se utilice directamente el modelo no lineal de la planta.

Los métodos para analizar la *controlabilidad* de una planta basándose en el modelo no lineal no están tan difundidos en la literatura como los índices lineales. Entre ellos se distinguen por una parte las *técnicas analíticas* y los *métodos numéricos basados en optimización*, para la *controlabilidad entrada-salida*, y por otra parte el análisis no lineal de *controlabilidad de estado*.

Las *técnicas analíticas* se enfocan en estudiar la posibilidad de obtener la inversa de la planta no lineal. La respuesta dinámica no forzada de un proceso no lineal está dada por la dinámica de la realización de mínimo orden de la inversa de la planta. Este análisis, denominado *dinámica de cero* (“*zero dynamics*”), revela si el sistema tiene una inversa estable o no (Daoutidis y Kravaris, 1991).

Li y Feng (1987) y Tsinias y Kalouptsidis (1991) estudiaron las condiciones necesarias y suficientes para la *controlabilidad funcional* de sistemas no lineales. Soloyev y Lewin (2003)

extienden la aplicación del *costo de la perturbación* (Índice similar al número de condición de la perturbación) a sistemas no lineales. Mijares et al (1985) y Maniousiouthakis y Nikolau (1989) proponen indicadores análogos al arreglo de ganancias relativas (RGA), en estado estacionario.

En la literatura se encuentra una corriente que aborda el análisis de *controlabilidad* estableciendo una relación formal entre la termodinámica del proceso y aspectos de la teoría de control no lineal basado en pasividad. Este análisis permite seleccionar los pareamientos entre las variables manipuladas y controladas que garantizan la estabilidad del sistema. Un sistema se hace *pasivo* cuando la tasa de cambio de la energía almacenada en el sistema está acotada por el producto de las variables manipuladas y las variables medidas (Ekawati, 2003). Ydschte y Alonso (1997) proponen el uso de este concepto para seleccionar los pareamientos entre variables de manera que la interconexión entre el proceso y el controlador sea pasiva.

En los trabajos de Meeuse (2000) y Meeuse y Grievink (2002) también se propone estudiar la *controlabilidad* en base al estudio de la termodinámica del proceso, estableciendo una relación entre la sensibilidad a la producción de entropía del sistema con su sensibilidad ante perturbaciones. Considerando modelos donde la producción de entropía depende de los flujos y fuerzas impulsoras que describen los fenómenos termodinámicos del proceso, se establecen relaciones entre estas variables que permiten minimizar la sensibilidad a perturbaciones. Este método es muy interesante aunque ha sido poco investigado hasta el momento y solo se ha aplicado a unidades de proceso sencillas (Intercambiador de calor, columna de destilación).

El análisis de *controlabilidad* no lineal mediante técnicas analíticas, al igual que en el caso lineal, requiere de gran experiencia para la interpretación la información que se obtiene de los distintos índices. Tienen la desventaja de que ninguna de estas técnicas permite estudiar más de uno de los aspectos limitantes de la *controlabilidad* a la vez, ni indica cómo producir un mejor diseño (Ekawati, 2003). En cambio, los *métodos numéricos* proveen resultados que son fáciles de interpretar y se relacionan directamente con los requerimientos de control especificados para la planta, ya que en general se optimizan el desempeño dinámico de la planta en función de estos requerimientos (Chenery, 1997). Los más usados entre estos

métodos se describen a continuación. Después, se discuten una serie de métricas para el análisis de la *controlabilidad de estado* basada en el modelo no lineal que puede complementar o, incluso ser una alternativa, para el estudio del comportamiento dinámico de estos sistemas.

3.1 CONTROLABILIDAD BASADA EN MÉTODOS NUMÉRICOS

En general, entre los índices numéricos de desempeño más utilizados para evaluar *controlabilidad* de sistemas no lineales se encuentran las integrales de las funciones de error como la *Integral del Cuadrado del Error* (ISE) y la *Integral del Valor Absoluto del Error* (IAE). A saber:

$$ISE = \int_{t=0}^{T_{\max}} (\mathbf{y}_{\text{ref}} - \mathbf{y})^2 \cdot dt \quad (2.25)$$

$$IAE = \int_{t=0}^{T_{\max}} |(\mathbf{y}_{\text{ref}} - \mathbf{y})| \cdot dt \quad (2.26)$$

donde, el tiempo inicial puede considerarse $t_i=0$ y T_{\max} es el tiempo máximo de simulación.

Se ha discutido respecto al uso del *ISE*, criticando aspectos como el hecho de que sólo refleja la dinámica de las variables medidas y desprecia la dinámica de las variables de estado no medibles. Además, la optimización de los parámetros de sintonización del controlador utilizando el *ISE* como índice puede presentar varios mínimos locales (Schweiger y Floudas, 1997). Sin embargo, a pesar de estos inconvenientes, su valor es proporcional a la calidad de las características dinámicas del proceso y se debe tomar en cuenta que es una medida fácil de calcular que, por ser una función diferenciable, permite el uso de métodos de gradiente para su optimización.

El *ISE* y el *IAE* son los índices más conocidos, pero existen también otras medidas dinámicas cuantitativas para evaluar el comportamiento del sistema, como las propuestas en los trabajos antes mencionados. Otros índices que permiten evaluar algunas características dinámicas del proceso son:

- Máxima desviación de la referencia:

$$Dev^{\max} = \max \{y_{ref} - y(t)\} = \|y_{ref} - y(t)\|_{\infty} \quad (2.27)$$

- Varianza del Error:

$$Var(e(t)) = \overline{e^2(t)} - (\overline{e(t)})^2 \quad (2.28)$$

donde:

$$\overline{e(t)} = \frac{\int_{t_i}^{t_f} e(t) dt}{T} \quad \text{y} \quad \overline{e^2(t)} = \frac{\int_{t_i}^{t_f} e^2(t) dt}{T}$$

- Varianza de las Variaciones de la Variable Manipulada:

$$Var(\Delta u) = \overline{\Delta u^2} - (\overline{\Delta u})^2 \quad (2.29)$$

donde:

$$\overline{\Delta u} = \frac{\int_{t_i}^{t_f} \Delta u dt}{T} \quad \text{y} \quad \overline{\Delta u^2} = \frac{\int_{t_i}^{t_f} \Delta u^2 dt}{T}$$

En la literatura se proponen distintas técnicas, algunas basados en los índices recién mencionados, para analizar la *controlabilidad* de los procesos en base a métodos numéricos de optimización:

Narraway y Perkins (1993) estiman un margen, entre el punto de operación de la planta y el óptimo económico, que garantiza que ninguna de las restricciones del proceso sea violada debido al efecto de perturbaciones (*margen de retroceso*), siendo ésta una medida del potencial económico de las posibles estructuras de control.

White (1991) estudia el efecto de las características fundamentales de la planta sobre su habilidad para desplazarse eficientemente de un punto de operación a otro, resolviendo un problema de control óptimo, en el cual se seleccionan las mejores trayectorias junto con los parámetros de la planta que afectan esta propiedad.

Georgiou y Floudas (1990) plantean la optimización de un índice de *controlabilidad estructural*. Esta es una matriz obtenida de la transformación del modelo dinámico,

independiente de las condiciones de operación y las no linealidades, que representa las relaciones entre las variables del proceso para una estructura particular. El rango de la matriz se calcula mediante la solución de un problema de programación entera lineal (ILP), si el rango es completo indica que la estructura es factible. Así, se determina el diagrama de flujo óptimo para una planta en base a la minimización de los costos. Un procedimiento similar se propone en Hopkins et al. (1998), quienes presentan la definición de *la Controlabilidad Estructural de Salida* (“*Output Structural Controllability*”, OSC) basado en una matriz de relaciones causa-efecto, y la proponen como herramienta cualitativa que ayuda a descartar, desde las etapas de síntesis, los diagramas de flujo que no sean estructuralmente controlables. Sin embargo, los autores consideran a la OSC como una condición necesaria pero no suficiente para garantizar la *controlabilidad* de un diseño.

Dimitriadis y Pistikopoulos (1995) introducen el *índice de flexibilidad dinámica* en el marco de un problema de optimización para una estructura fija del proceso. Éste representa la máxima desviación en el perfil de parámetros con incertidumbre que el diseño puede tolerar manteniéndose en la región de operación factible dentro de un horizonte finito de tiempo.

Vinson y Georgakis (1998) proponen el *índice de controlabilidad de la salida* (OCI), el cual cuantifica la habilidad del proceso de alcanzar el rango completo de valores deseados de salida con acciones de control en el rango disponible de las variables de entrada, en presencia de incertidumbre.

Estos métodos numéricos para evaluar la *controlabilidad* superan a los métodos basados en el modelo lineal en el sentido de que se pueden relacionar directamente con las especificaciones de desempeño. Permiten realizar el análisis en el dominio del tiempo y la optimización en el problema de diseño puede plantearse de manera de tomar en cuenta el efecto de las distintas características limitantes del desempeño de planta al mismo tiempo. Sin embargo, tienen como desventaja que requieren simulaciones del modelo no lineal, incrementando considerablemente el costo computacional del procedimiento.

3.2 CONDICIONES DE CONTROLABILIDAD DE ESTADO NO LINEAL

El modelo de estado puede servir de base también para un análisis de *controlabilidad* para el caso no lineal. El análisis sistemático de la *controlabilidad de estado* para sistemas no lineales se inició en la década de los 70 (Sussmann y Jurdjevic, 1972; Hermann y Krener, 1977), planteado un análogo no lineal del concepto de *controlabilidad de estado* en base al modelo lineal y revisando diversos aspectos de este concepto.

En primer lugar, en la definición de *controlabilidad de estado* para sistemas lineales se supone una entrada $u(t)$ sin restricciones por lo que la acción de control no es acotada, mientras que en la práctica la señal de control o la variable manipulada presentan claras limitaciones con importantes efectos sobre el desempeño del sistema. Por otra parte, en el caso de sistemas lineales, se puede definir la *controlabilidad completa de estado* gracias al principio de superposición, pero en sistemas no lineales esta globalidad se pierde, por lo que se hace necesario trabajar con el concepto de *controlabilidad local*, es decir, para un entorno del punto controlable (Ochoa, 2005).

Así, para los sistemas no lineales se define la *controlabilidad local práctica*, como una manera de abordar el problema desde un punto de vista local y tomando en cuenta el acotamiento de las acciones de control. Se dice que un sistema es localmente controlable en la práctica en un estado de equilibrio $x=x^*$, si el estado x^* es alcanzable en un tiempo finito, desde un punto inicial x_0 que pertenece a un vecindario de x^* , empleando un conjunto de acciones de control $u(t)$ perteneciente al espacio disponible U de las acciones de control o entradas manipuladas (Hermann y Krener, 1977).

Es importante introducir además el concepto de *alcanzabilidad* que es la habilidad para llegar a cualquier estado final deseado desde cualquier estado inicial dado en un tiempo finito, mediante un conjunto de acciones de control. En base a esto, se puede decir que la *controlabilidad* está relacionada con la posibilidad de alcanzar el estado de equilibrio deseado, independientemente del controlador y la estrategia de control a utilizar.

En este contexto, para el análisis de *controlabilidad local práctica* es necesario considerar una representación manejable del sistema, por lo que se consideran los sistemas dinámicos no lineales afines con sus entradas:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) + \sum_{i=1}^m \mathbf{g}_i(\mathbf{x}) \cdot u_i \\ \mathbf{y} &= \mathbf{h}(\mathbf{x})\end{aligned}\tag{2.30}$$

donde:

\mathbf{x} es el vector de estados,

\mathbf{y} es el vector de salidas,

u_i es un elemento del vector de entradas manipuladas,

$\mathbf{f}(\mathbf{x})$ es una función no lineal de los estados asociada con la respuesta natural del sistema (respuesta no forzada) la cual incluye, además, el efecto de las perturbaciones,

$\mathbf{g}(\mathbf{x})$ es una función no lineal de los estados, pero lineal con respecto a la entrada manipulada (sistema afín con la entrada manipulada) y asociada con la respuesta forzada, y

$\mathbf{h}(\mathbf{x})$ es una función que representa el mapeo desde el espacio de los estados hacia el espacio de las salidas.

En el caso de sistemas con única entrada, el término $\mathbf{g}(\mathbf{x})$ es un vector de acciones forzantes, mientras que para el caso de múltiples entradas donde \mathbf{u} es un vector de las m entradas manipuladas, $\mathbf{g}(\mathbf{x})$ es una matriz ($\mathbf{G}(\mathbf{x})$) de $n \times m$ componentes, donde n es el número de estados del sistema.

En el análisis de *controlabilidad* para sistemas lineales, se presenta la *matriz de controlabilidad* (\mathbf{Co} , ec. 2.11) como un indicador para determinar si el sistema es de estado controlable. Para sistemas no lineales se tiene la *matriz de controlabilidad local* (\mathbf{Wc}), la cual se define separadamente para sistemas de única entrada (Henson y Seborg, 1997) y para sistemas con múltiples entradas (Ochoa *et al.*, 2004).

Para un sistema de única entrada ($m=1$):

$$\mathbf{Wc} = \left[\mathbf{g}(\mathbf{x}) \quad ad_f \mathbf{g}(\mathbf{x}) \dots ad_f^{n-1} \mathbf{g}(\mathbf{x}) \right]_{n \times m}\tag{2.31}$$

donde, $ad_f g(x)$ es la operación de paréntesis de Lie entre los campos vectoriales $\mathbf{f}(x)$ y $\mathbf{g}(x)$ definida como:

$$[\mathbf{f}, \mathbf{g}] = \frac{\partial \mathbf{g}(x)}{\partial x} \mathbf{f} - \frac{\partial \mathbf{f}(x)}{\partial x} \mathbf{g}(x) \quad (2.32)$$

entonces,

$$(ad_f^2, \mathbf{g}) = [\mathbf{f}, [\mathbf{f}, \mathbf{g}]] \text{ y}$$

$$(ad_f^k, \mathbf{g}) = [\mathbf{f}, (ad_f^{k-1}, \mathbf{g})].$$

Para un sistema de múltiples entradas ($m=2,3,\dots,n$), el término $\mathbf{g}_i(x)$ no es un vector, sino una matriz que en adelante se denominará $\mathbf{G}(x)$, y que por simplicidad matemática se considerará conformada por m campos vectoriales:

$$\mathbf{G}(x) = [\mathbf{g}_1 \quad \mathbf{g}_2 \dots \mathbf{g}_m]_{n \times m} \quad (2.33)$$

En este caso, la matriz de *controlabilidad* está dada por:

$$\mathbf{Wc} = \begin{bmatrix} \mathbf{g}_1(x) \mathbf{g}_2(x) \dots \mathbf{g}_m(x) & ad_f \mathbf{g}_1(x) & ad_f \mathbf{g}_2(x) & ad_f \mathbf{g}_m(x) \\ ad_f^{n-1} \mathbf{g}_1(x) & ad_f^{n-1} \mathbf{g}_2(x) \dots ad_f^{n-1} \mathbf{g}_m(x) \end{bmatrix}_{nx(n \cdot m)} \quad (2.34)$$

donde los términos $ad_f^{n-1} \mathbf{g}_1(x)$ $ad_f^{n-1} \mathbf{g}_2(x)$ $ad_f^{n-1} \mathbf{g}_m(x)$, son las operaciones de paréntesis de Lie aplicadas recursivamente sobre cada uno de los campos vectoriales:
 $ad_f^{n-1} \mathbf{g}_i(x) = [\mathbf{f}, ad_f^{n-2} \mathbf{g}_i(x)]$.

Igual que para los sistemas lineales, para que el sistema sea de estado controlable se debe cumplir que la *matriz de controlabilidad local no lineal* sea de rango completo, es decir, si \mathbf{Wc} es cuadrada ($n \times n$) es necesario que $\det(\mathbf{Wc}) > 0$ y si \mathbf{Wc} no es cuadrada entonces $\text{rank}(\mathbf{Wc}) = n$.

En la literatura se encuentran algunas aplicaciones del análisis de *controlabilidad de estado* en sistemas no lineales. Por ejemplo, Ochoa (2005) identifica seis casos de *no controlabilidad práctica*, relacionados con aspectos tales como grados de libertad insatisfechos, estados no afectados por las acciones de control, acciones de control que no

afectan a ningún estado, saturación de las entradas, selección inadecuada de los elementos finales de control y restricciones físicas de los estados. Además, desarrolla unas *métricas de controlabilidad práctica*, para integrar la evaluación de la *controlabilidad de estado* de un sistema no lineal en el diseño de procesos. Con base en estas ideas, Muñoz (2007) propone una metodología para verificar la controlabilidad de sistemas acoplados no lineales. Finalmente, Alvarez (2008) utiliza una representación del sistema basada en una combinación de las matrices de controlabilidad y observabilidad (Matriz de Hankel), que permite calcular el impacto de cada variable de entrada (acción de control) sobre el comportamiento del sistema, con lo se establece un criterio para seleccionar las combinaciones de variables manipuladas y controladas en toda la planta.

El análisis de *controlabilidad local* de un sistema no lineal debe ser complementado con elementos que lo relacionen con las aplicaciones prácticas en el control de procesos. En la sección anterior fueron descritos algunos fenómenos que limitan el desempeño dinámico de una planta y que sirven de base para el desarrollo de algunos índices de *controlabilidad* basados en el modelo lineal. De la misma manera, Ochoa (2005) identifica algunas fallas comunes en el diseño de procesos, que le permiten formular cinco métricas de *controlabilidad* que se resumen a continuación: grados de libertad insatisfechos, estados no afectados por las acciones de control, acciones de control que no afectan a ningún estado, saturación de las entradas, selección inadecuada de los elementos finales de control y restricciones físicas de los estados. Si el sistema dinámico en estudio presenta por lo menos una de estas condiciones, dicho sistema será no controlable. Se trata entonces de condiciones necesarias pero no suficientes para determinar la *controlabilidad práctica* de un sistema dinámico, cuando son evaluadas por separado. Sin embargo, al evaluarlas a todas en conjunto sí podrá determinarse la *controlabilidad práctica* del sistema en cuestión.

j) Rango de la matriz de controlabilidad.

La primera métrica utilizada como condición necesaria en este análisis es el rango de la *matriz de controlabilidad*, la cual se encuentra asociada al caso de no controlabilidad por restricciones físicas de los estados. Mediante este índice se evalúa la posibilidad de alcanzar un estado deseado independientemente del tiempo y de la disponibilidad de las acciones de

control, por lo tanto, solamente tiene en cuenta la posibilidad física de alcanzar dichos estados deseados.

Este criterio consiste en determinar la independencia o dependencia lineal entre las filas o columnas de la *matriz de controlabilidad* (W_c), evaluando el determinante o el rango de la matriz. Si el sistema posee n estados y una sola variable manipulada, la *matriz de controlabilidad* es cuadrada y se debe cumplir que $\det(W_c) \neq 0$ para que el sistema sea controlable. Si el sistema analizado cuenta con más de una entrada manipulada ($m > 1$) la *matriz de controlabilidad* W_c no es cuadrada, y la condición para que el sistema sea de estado controlable es $\text{rank}(W_c) = n$.

k) Grados de libertad para el control

Los grados de libertad insatisfechos constituyen un caso de *no controlabilidad práctica*, que se presenta cuando el número de variables manipuladas es menor al número de variables que se desean controlar.

El cálculo de Grados de Libertad es directo:

$$G.D.L = N.V.M - N.V.C \quad (2.35)$$

donde:

$N.V.M$ es el Número de Variables Manipuladas (acciones de control)

$N.V.C$ el Número de Variables a Controlar (estados).

Si el sistema tiene $G.D.L < 0$, será no controlable en la práctica, y por lo tanto, será necesario modificar el diseño de manera que se puedan incluir nuevas variables manipuladas para así garantizar el cumplimiento de los grados de libertad para el control.

En algunos sistemas dinámicos puede darse el caso de que por más se realicen modificaciones en el diagrama de flujo no sea posible contar con un número de variables manipuladas igual al número de estados. Cuando esto ocurre, es necesario reducir el modelo de manera que las variables consideradas secundarias dejen de ser estados del sistema.

l) Determinante de la matriz asociada con la respuesta forzada

Un sistema dinámico no lineal compuesto por n estados e igual número de variables manipuladas, puede representarse por medio de la forma canónica afín con la entrada similar a la ecuación (2.35),

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x}) \cdot \mathbf{u} \quad (2.36)$$

En la definición *de controlabilidad local práctica* se recalca que el punto final que se desea alcanzar, y al cual se le evalúa la *controlabilidad*, es un punto de equilibrio forzado \mathbf{x}^* , donde en ausencia de perturbaciones ninguno de los estados del sistema cambia con el tiempo ($\dot{\mathbf{x}}=0$) cuando se aplica la acción de control forzante \mathbf{u}^* dada por la expresión obtenida a partir de la ecuación (2.41):

$$\mathbf{u}^* = -[\mathbf{G}(\mathbf{x}^*)]^{-1} \cdot \mathbf{f}(\mathbf{x}^*) \quad (2.37)$$

Se puede observar que si la inversa de la matriz \mathbf{G} no existe, entonces no será posible obtener una acción de control que permita garantizar la permanencia de los estados en el punto de equilibrio deseado, y por lo tanto, no se podrá garantizar la *controlabilidad práctica* del sistema.

Esta es la misma conclusión a la que se llega mediante el análisis de las limitaciones sobre el *control perfecto de la planta* para sistemas lineales, descrita en la sección anterior, sobre la que se fundamentan muchos de los índices de *controlabilidad* desde la perspectiva de *entrada-salida*. En este caso, la visión de control de los estados permite asociar esta condición con los modos controlables y no controlables del sistema.

La matriz \mathbf{G} no es invertible cuando se cumple: $\det(\mathbf{G})=0$, y por lo tanto, el sistema es no controlable. El determinante de una matriz puede ser cero cuando la matriz contiene una fila de ceros, cuando contiene una columna de ceros o cuando existe dependencia lineal entre sus filas o entre sus columnas.

Cuando la matriz presenta una fila de ceros, los elementos $g_{ij} = 0$ para todo $j = 1, 2, \dots, n$, en tal caso, la representación en la forma canónica (2.36) para la dinámica del estado x_i es

$\dot{x}_i = f_i(x)$, por lo que se puede apreciar que no existe dependencia de las entradas. Este comportamiento corresponde al segundo caso de *no controlabilidad práctica*, que se refiere a los estados no afectados por las acciones de control, específicamente, ninguna acción de control puede afectar al estado x_i , cuya dinámica solamente depende de su respuesta natural.

Para el caso en que el determinante es cero debido a que la matriz \mathbf{G} presenta una columna de ceros, los elementos $g_{ij} = 0$ para todo $i = 1, 2, \dots, n$, y por lo tanto, ninguno de los n estados presenta respuesta forzada g_{ij} . Esto se traduce en que la acción de control u_j no afecta a ninguno de los estados del sistema, por lo cual dicha acción de control debe ser descartada. Cuando esto ocurre, el sistema queda con un déficit de *grados de libertad para el control*, que es el primer caso de no controlabilidad. De esta manera, este análisis permite identificar si cada una de las variables manipuladas del sistema afecta por lo menos a uno de los estados, aún de manera preliminar se haya satisfecho la condición de los *grados de libertad para el control*.

Si el determinante de la matriz \mathbf{G} es cero debido a la dependencia lineal entre las filas o entre las columnas de la matriz, significa que existe dependencia entre dos o más estados, lo cual restringe enormemente el conjunto de estados de equilibrio forzado que pueden ser alcanzados, haciendo que el sistema sea no controlable en la práctica.

En resumen, el uso del determinante de la matriz \mathbf{G} como índice de *controlabilidad* permite detectar limitaciones para el control del proceso, ya sea debido a que uno de los estados presenta solo respuesta natural, a que una de las acciones de control deba ser descartada por no afectar a ninguno de los estados, o a que los estados no son independientes. Se podría decir que es una métrica que determina los grados de libertad efectivos para el control, ya que evalúa la existencia de relaciones causa–efecto entre las variables manipuladas y los estados a controlar.

m) Pertenencia de la acción de control forzante (u^) al intervalo disponible de acciones de control*

La acción de control forzante está dada por la ecuación (2.37): $\mathbf{u}^* = -[\mathbf{G}(\mathbf{x}^*)]^{-1} \cdot \mathbf{f}(\mathbf{x}^*)$. Por lo tanto, la existencia de la acción forzante o variable manipulada requerida para mantener al

sistema en el punto de equilibrio x^* se determina evaluando la inversa de la matriz $G(x^*)$. Por otra parte, el intervalo disponible de acciones de control (U) se denota por:

$$U = \left\{ \begin{array}{l} u_{1\min} \leq u_1 \leq u_{1\max} \\ u_{2\min} \leq u_2 \leq u_{2\max} \\ \vdots \\ u_{j\min} \leq u_j \leq u_{j\max} \\ \vdots \\ u_{n\min} \leq u_n \leq u_{n\max} \end{array} \right\} \quad (2.38)$$

donde:

$u_{j\min}$ y $u_{j\max}$ son respectivamente, los valores mínimo y máximo disponibles para la j-esima variable manipulada.

Si la acción de control u^* que fuerza al sistema a permanecer en el punto de equilibrio x^* no pertenece al intervalo disponible de acciones de control U , el sistema no será controlable en dicho punto. Aunque, con el conjunto de acciones de control disponibles, el punto deseado sea alcanzable, el sistema no logrará mantenerse allí porque es incapaz de ofrecer la acción de control requerida para tal fin. Por lo tanto, si $u^* \notin U$ el sistema no será controlable en la práctica.

Para determinar la *controlabilidad* de un sistema dinámico deben cumplirse todas las condiciones estudiadas, sólo entonces el sistema podrá declararse controlable en el punto de equilibrio evaluado.

Aunque en el campo de la ingeniería química, y especialmente en el diseño integrado de procesos, se ha dejado de lado el análisis de *controlabilidad de estado*, el conjunto de métricas presentadas permiten utilizar técnicas tradicionales de la teoría de control de estado en la evaluación del comportamiento dinámico del proceso en estudio mediante cálculos relativamente sencillos. La ventaja es que se trabaja el dominio del tiempo, pero el inconveniente de esta técnica es su naturaleza local, centrándose en la *controlabilidad* de los estados alrededor de un punto de equilibrio.

4. CONCLUSION

En este capítulo se describen las técnicas más utilizadas para analizar la *controlabilidad*, tanto desde la óptica del control entrada-salida como del de *controlabilidad* de estado.

Se presentan los distintos índices de controlabilidad (entrada-salida) basados en el modelo lineal entrada-salida, los cuales se asocian a ciertas características que limitan el control perfecto de la planta o a los problemas de direccionalidad de las entradas. Estos índices son ampliamente utilizados pero su aplicabilidad se limita al rango de validez del modelo en el caso de procesos con no linealidades significativas, y es difícil interpretar la información que aportan respecto al desempeño de la planta en términos de las características físicas del proceso, principalmente porque éste se maneja en el dominio frecuencial.

En cuanto al análisis de controlabilidad basado en el modelo no lineal, las simulaciones dinámicas del proceso a lazo cerrado frente a las distintas entradas y perturbaciones de interés son una herramienta valiosa para cuantificar su desempeño de control, sin embargo requieren de información que puede desconocerse en etapas preliminares del diseño. Las técnicas más utilizadas se basan en métodos de optimización de índices de desempeño dinámico, las cuales tienen numerosas ventajas porque se relacionan directamente con los requerimientos de respuesta y se trabaja en el dominio temporal, pero requieren de un esfuerzo computacional importante.

Finalmente, se discute el concepto de *controlabilidad de estado* de Kalman, en dos aspectos fundamentales: el acotamiento de las entradas o variables manipuladas y el carácter local que se debe tener en cuenta cuando se trabaja con sistemas no lineales. Se trata de la evaluación de un conjunto de *métricas* que mediante cálculos sencillos, exponen fallas típicas de diseño, como son: la evaluación de los *grados de libertad para el control*, que indica la relación entre el número de variables a controlar y el número de variables manipuladas, el *rango de la matriz de controlabilidad*, que indica si existe la posibilidad física de alcanzar los puntos de ajuste deseados para todos los estados del sistema, el *determinante de la matriz asociada con la respuesta forzada*, que, permite identificar los estados que no son afectados por las acciones de control o las acciones de control que no afectan a ningún estado, y la

pertenencia de la acción d forzante al intervalo disponible de acciones de control, que aporta información respecto a la saturación de las entradas.

La revisión del concepto de controlabilidad y la descripción de los distintos métodos para su evaluación en distintos escenarios, permite abordar con criterios claros el problema de interacción entre el diseño de los procesos y su controlabilidad y la aplicación de la metodología de diseño integrado que se abordan en el capítulo siguiente.

CAPÍTULO III

DISEÑO INTEGRADO DE PROCESOS

1. INTRODUCCIÓN

La *Integración Diseño-Control* busca conciliar aspectos frecuentemente conflictivos entre el diseño en estado estacionario, centrado en lograr el máximo beneficio económico, y los requerimientos de desempeño sobre el sistema de control para lograr una operación eficiente. El *diseño* tiene que ver con la síntesis del diagrama de flujo del proceso y la determinación de las condiciones nominales de operación, mientras que los aspectos de control se relacionan con la aplicación de estrategias para mantener las variables clave del proceso dentro de los límites especificados y minimizar los efectos de las perturbaciones que tienen impacto negativo sobre la calidad del producto y se resuelven tradicionalmente después del diseño.

Sin embargo de este modo, se corre el riesgo de diseñar un proceso que sea óptimo económicamente, pero que también sea muy difícil de controlar, exhibiendo un desempeño dinámico deficiente lo que, a la larga, hace más costosa su operación. Esto obliga a realizar modificaciones en la planta para mejorar su controlabilidad, evidenciando la fuerte interacción entre diseño y control. La repercusión económica de este fenómeno ha sido ampliamente reconocida en la ingeniería de procesos y de control. Esto ha llevado a desarrollar un buen número de trabajos enfocados en la integración de criterios de control en el procedimiento de diseño como se describe en la sección dedicada al estado del arte en el Capítulo I. En esta revisión se distinguen numerosas técnicas para incorporar criterios de *controlabilidad* en las etapas preliminares de diseño, estableciendo como filosofía la *Integración Diseño-Control*. Entre éstas se encuentran las *Metodologías Anticipativas*, enfocadas en la evaluación de la *controlabilidad* de posibles alternativas de diseño ya definidas, y los métodos de *Diseño*

Integrado de Procesos basados en la optimización simultánea del proceso y el sistema de control.

En las *metodologías anticipativas* se determinan, en primer lugar, los grados de libertad, los objetivos de control y las variables manipuladas para cada posible estructura. Luego, se analiza su comportamiento mediante la evaluación índices de controlabilidad estáticos. Por otra parte, el diseño y control simultáneo (*Diseño Integrado de Procesos*) presenta una gran ventaja conceptual, al enfocarse en la formulación de un problema de optimización cuya solución es el diseño que satisfaga, de la mejor manera posible el compromiso entre los objetivos económicos y los de control. Como se ha visto, se encuentran numerosas metodologías enfocadas en aspectos particulares del problema, pero no se ha establecido un criterio unificado para abordarlos en su conjunto. Los diversos planteamientos difieren en el alcance del problema de diseño, la estrategia de control utilizada, la forma de evaluar la controlabilidad y el tratamiento de las perturbaciones e incertidumbre del modelo. Estos llevan a formulaciones matemáticas de distinto nivel de complejidad cuya dificultad radica en la forma de introducir el análisis de controlabilidad y las estrategias de control. Además, para su resolución es necesario el desarrollo de algoritmos eficientes lo cual, como se ha observado, es un aspecto decisivo para el progreso en esta área.

La metodología de *diseño simultáneo del proceso y el sistema de control* tiene la ventaja de que el modelo del proceso con todas sus particularidades puede ser considerado en la formulación matemática del problema, lo que permite realizar un análisis más completo y realista del desempeño dinámico de la planta. Además, es posible obtener el diseño con el sistema de control ya sintonizado lo cual es otra gran ventaja desde el punto de vista práctico. Entre las desventajas, se encuentra la necesidad de un conocimiento detallado del modelo del sistema, que posiblemente no se tenga en las etapas preliminares del diseño y el gran esfuerzo computacional requerido para la resolución del problema de optimización que se presenta.

En la siguiente sección se estudia con más detenimiento la interacción entre el diseño y la *controlabilidad*. En la sección 3, se presenta la formulación del problema típico de síntesis y diseño simultáneo del proceso y del control, para concluir en la sección 4, con la elaboración

del ejemplo de aplicación de la metodología de *diseño integrado* de un sistema de reactores químicos.

2. INTERACCION ENTRE EL DISEÑO Y EL CONTROL DEL PROCESO

Entre los objetivos de diseño de un proceso, ya sea secuencial o integrado, se encuentran tanto lograr el beneficio económico esperado como la operación eficiente y segura de la planta. El beneficio económico generalmente se relaciona con la minimización de los costos de inversión y operación en estado estacionario de la planta y con políticas de ahorro de energía. La operación eficiente que se refiere al buen desempeño dinámico y la obtención de los productos de acuerdo a las especificaciones y con poca variabilidad en su calidad, depende de la controlabilidad.

La controlabilidad es inherente a las características de diseño de la planta (estructura y dimensiones). Las decisiones de diseño que representan una reducción en los costos de inversión y operación, suelen tener un impacto negativo sobre la respuesta dinámica. De aquí, el conflicto *economía-controlabilidad*. Por otra parte, la controlabilidad depende, obviamente, del diseño de los sistemas de control. Esto comprende, desde la selección de los puntos de operación (estados alcanzables) y variables manipuladas, hasta la selección y ajuste de los controladores de los que, generalmente, se espera rechazo a perturbaciones y seguimiento de la referencia con un tiempo de respuesta y estabilidad relativa aceptables.

2.1 CONFLICTO ECONOMÍA-CONTROLABILIDAD

En el diseño de procesos químicos son frecuentes los casos donde se evidencia el conflicto economía-controlabilidad. Para ilustrar esta situación con un ejemplo, se toma el sistema de reactores químicos estudiado por Luyben (1993).

Se trata de un proceso de reacción química de primer orden, en fase líquida, del tipo $A \rightarrow B$, exotérmica e irreversible, el cual se lleva a cabo en un conjunto de uno a tres reactores tipo tanque agitado continuo (TAC) conectados en serie. Analizando la respuesta estacionaria de este sistema se pueden establecer relaciones entre las variables de diseño.

Por ejemplo, se deduce que para una conversión (X_c) dada y una constante de reacción fija (k_r), los costos de inversión se reducen a medida que aumenta el número de reactores, ya que el volumen total de reacción es menor cuando se trata de equipos en serie. Sin embargo, la respuesta dinámica del sistema desmejora drásticamente, ya que los sistemas de dos o más reactores resultan inestables a lazo abierto en algunas condiciones de operación. Además, en la simulación de la respuesta a lazo cerrado se encuentra que un único reactor presenta mejor desempeño dinámico ante perturbaciones. Esto lleva a la conclusión de que los sistemas de dos o tres reactores son la mejor elección desde el punto de vista económico, pero presentan problemas importantes de control.

Enfocándose en los sistemas de múltiples reactores, que son más convenientes desde el punto de vista económico, se pueden estudiar posibles modificaciones estructurales para mejorar la *controlabilidad* (Luyben, 1993). Considerando que en la primera etapa se debería disipar más calor debido a una mayor concentración de reactivos, se propone aumentar su área de transferencia de calor, ya sea mediante la adición de un intercambiador de calor externo o utilizando dos reactores en paralelo que manejen el mismo volumen. Esto aumenta los costos pero mejora la eficiencia del control.

La *selección de la temperatura de reacción* para una conversión dada también tiene un efecto importante. Desde la perspectiva económica, la temperatura de reacción debe ser la máxima posible para minimizar el volumen del reactor y consecuentemente los costos. Sin embargo, al evaluar la dinámica, se observa en diseños con el mismo coeficiente de transferencia de calor pero distinto ΔT (diferencia entre la temperatura del reactor y la temperatura de la camisa), que a medida que aumenta ΔT es menor la capacidad del sistema controlado de restablecerse ante un cambio en el calor de reacción, observándose incluso saturaciones en la variable manipulada (F_j) cuando ocurren variaciones de gran magnitud.

Destaca en el trabajo de Luyben (1993) el uso de la relación Q_{max}/Q como un índice de *controlabilidad*. Q_{max} es la máxima tasa de transferencia de calor que se puede lograr con el flujo máximo de refrigerante. Luego, Q_{max}/Q es una medida de la habilidad del sistema de refrigeración para manejar requerimientos de transferencia de calor que superen lo establecido para las condiciones nominales de operación, y debe ser lo suficientemente grande en procesos

sometidos a perturbaciones e incertidumbre en sus propiedades cinéticas y termodinámicas. En la curva de Q_{max}/Q vs T_R existe un máximo, por lo tanto, puede tener dos valores de T_R para el mismo Q_{max}/Q . El criterio económico determina la selección del diseño con mayor temperatura (T_R) y consecuentemente menor volumen, ya que ambas posibilidades proporcionan la misma *controlabilidad*. Esta es la mejor solución compromiso entre economía y control para este caso de estudio.

Se estudia también el efecto de la conversión una temperatura de reacción dada, demostrando que es más fácil controlar los sistemas de reactores con una conversión más alta. Económicamente, conviene trabajar con una conversión baja porque se necesitan sistemas de menor tamaño, sin embargo, la relación Q_{max}/Q es mayor si la conversión es alta indicando una mejor *controlabilidad* debido al aumento del volumen de reacción y las áreas de transferencia de calor.

El mismo efecto sobre la *controlabilidad* se observa cuando aumenta el coeficiente de transferencia de calor U , ya que el aumento de U permite alcanzar mayores valores de Q_{max}/Q . Por otra parte, la operación a una alta velocidad de reacción afecta negativamente la *controlabilidad* del sistema porque el volumen del sistema y el área de transferencia de calor disminuyen y consecuentemente la relación Q_{max}/Q .

Finalmente, este ejemplo ilustra cómo parámetros del proceso como $\Delta T = (T_R - T_{j_i})$ y la relación Q_{max}/Q actúan como indicadores de desempeño dinámico que orientan en la toma de decisiones de diseño. Además, se observa que la mayoría de los casos, la relación Q_{max}/Q es inversamente proporcional a los costos de inversión, evidenciando claramente el compromiso entre los criterios económicos y de control.

Posteriormente, para el mismo caso de estudio, Subramanian et al (2001) realizan un análisis de *operabilidad* dinámica y en estado estacionario. Encuentran que los reactores operando a altas temperaturas (menos costosos) son inherentemente más rápidos en su respuesta y los sistemas de dos reactores en serie son más rápidos que uno simple cuando se opera a la misma temperatura. Esto muestra que las interacciones entre el diseño y control pueden ser analizadas desde diferentes puntos de vista llevando a resultados distintos dependiendo de la profundidad del estudio y de los objetivos de control.

El sistema de reactores es un ejemplo relativamente sencillo donde las interacciones entre diseño y control se evidencian claramente. Este fenómeno sin embargo, se presenta en la mayoría de los procesos químicos. A continuación se resumen otros casos analizados por Luyben (2004), en donde los requerimientos económicos conducen a decisiones que son contraproducentes desde el punto de vista de los criterios de control:

- *Determinación del área de transferencia de calor.* Un intercambiador de calor con una menor área de transferencia, implica una menor inversión de capital. Sin embargo, un área de intercambio pequeña requiere de un mayor gradiente de temperatura en condiciones normales de operación, lo cual limita el poder de reacción del sistema de control ante perturbaciones, como se observó en el sistema de reactores de Luyben (1993).
- *Diseño de columnas de destilación.* Desde el punto de vista económico es mejor un diseño donde la relación entre el reflujo actual y el reflujo mínimo sea pequeña, obteniéndose un menor número de platos y menores requerimientos de vapor. Sin embargo, en las columnas con una relación de reflujo alta se tiene un mejor desempeño ante perturbaciones, ya que hay una mayor capacidad de transferencia de calor.
- *Diseño de reactores con dos reactivos.* En este caso, se obtienen reactores más pequeños y es necesario un menor reflujo si se utiliza una concentración equimolar de ambos reactivos ya que la velocidad de reacción aumenta directamente con el producto de ambas concentraciones. Desde el punto de vista de control, al tener una gran cantidad de ambos reactivos disponible, la velocidad de reacción puede ser tan alta que se pueden producir picos de temperatura en algunos casos (sistema muy sensible). Para evitar este fenómeno se debe dimensionar el reactor en función de un reactivo limitante, de manera de reducir la velocidad de reacción, pero esto significa utilizar un reactor más grande y consecuentemente más costoso.
- *Dimensionamiento de las válvulas.* Para reducir los costos de bombeo es conveniente seleccionar válvulas que producen caídas de presión pequeñas, pero esto limita la relación entre los caudales máximo y mínimo que puede manejar la válvula (“rangeability”) haciendo más probable la saturación de este elemento final de control.

Los ejemplos descritos en esta sección ilustran el efecto que tienen sobre la dinámica del proceso controlado las decisiones de diseño basadas en la optimización de objetivos económicos en estado estacionario. Sin embargo, otras decisiones de diseño también pueden afectar negativamente la *controlabilidad* del proceso. En Ochoa (2005) se señala que un sistema puede llegar a ser no controlable en un punto de operación dado, simplemente, porque físicamente es imposible alcanzar dicho punto. Cita algunos ejemplos como:

- No tener en cuenta la posibilidad de que ocurran cambios de fase para el intervalo de condiciones de operación deseadas.
- En el caso de reacciones químicas reversibles, donde las propiedades termodinámicas del sistema limitan la conversión.
- Cuando se controlan propiedades fisicoquímicas del producto final, y el objetivo de control no es alcanzable debido a que la naturaleza química del producto no lo permite.
- La mala selección de los elementos finales de control, sin tomar en cuenta si su intervalo de acción es adecuado para la magnitud de las perturbaciones que afectan el sistema, o los efectos de tiempos muertos y zonas muertas que pueden presentar estos elementos.

En estos casos es posible que el sistema parezca ser controlable teóricamente, pero son las restricciones físicas de los estados y los elementos finales de control las que limitan la *controlabilidad*.

Por otra parte, la selección de la estructura de control, que viene dada por la definición de variables controladas y manipuladas, y el emparejamiento entre éstas, también es un aspecto determinante en la interacción entre el diseño del proceso y su *controlabilidad*, que se discute a continuación.

2.2 CONTROLABILIDAD Y ESQUEMAS DE CONTROL

La dinámica de los procesos químicos varía significativamente para distintas condiciones de operación debido a su comportamiento no lineal, además, generalmente se trata de sistemas multivariados. Por lo tanto, se necesitan distintas acciones de control e incluso distintas estructuras o esquemas de control para satisfacer el desempeño dinámico deseado.

El desarrollo de un sistema de control requiere de la definición de los objetivos de control, un conjunto de variables controladas, un conjunto de variables manipuladas, un conjunto de variables medidas y una estructura que interconecte las variables controladas y las variables manipuladas. De la selección de estos conjuntos e interconexiones depende la *controlabilidad* del sistema global.

En el diseño se seleccionan las variables manipuladas, entre las distintas entradas del proceso, en función de su impacto sobre las variables a controlar. La decisión es trivial en el caso de sistemas de una entrada-una salida, pero en el caso de sistemas multivariables para satisfacer los objetivos de control en un tiempo aceptable es necesario contar con un número de variables manipuladas igual o mayor al número de variables a ser controladas (Marlin, 1995). El cumplimiento de esta condición le proporciona flexibilidad suficiente al proceso de tal manera que puedan satisfacerse los objetivos de control.

En las etapas preliminares del procedimiento de diseño tradicional, la mayor parte de las plantas no son controlables. Esto muchas veces se debe a que no hay suficientes variables manipuladas en el diagrama de flujo del proceso para satisfacer todas las restricciones y mantener las variables de operación en su valor óptimo en presencia de perturbaciones. Dos fallas típicas en el diseño del proceso que influyen sobre la selección de una estructura de control adecuada y consecuentemente, sobre el desempeño dinámico del sistema controlado son:

- Alguno de los estados a controlar no es afectado por ninguna de las acciones de control posibles, por lo que al menos uno de los objetivos de control no puede ser satisfecho.
- Existe una acción de control que no afecta la dinámica de ningún estado, lo que equivale a tener una acción de control menos, por lo tanto los grados de libertad para el control realmente no se satisfacen. Es importante que todas las variables manipuladas seleccionadas afecten por lo menos a uno de los estados a controlar.

La selección de las variables manipuladas está sujeta a la validez del rango de acciones de control requerido. En la práctica se presentan restricciones relacionadas con la capacidad del elemento final de control o con la limitación de recursos (como en el caso de flujos de

servicio y de materias primas). En algunos casos puede ocurrir que para alcanzar un punto de operación especificado o en presencia de perturbaciones se requiera llevar por lo menos una de las entradas por fuera de su intervalo de acción disponible, lo cual ocasionará la saturación del elemento final de control, y a menos que se realicen modificaciones de diseño, el proceso será no controlable en dicho punto de operación.

La integración del diseño del proceso y su control se debe fundamentar en primer lugar en identificar decisiones de diseño que potencialmente puedan generar problemas en el desempeño del sistema de control, como algunos de los ejemplos descritos en esta sección, y luego, en resolver los conflictos *economía-controlabilidad* de la manera más conveniente. Este trabajo se centra en la metodología de *Diseño Integrado* para abordar de manera sistemática la *interacción diseño-control* desde las etapas preliminares de diseño.

En la sección siguiente, se propone una formulación general del problema de optimización resultante de este planteamiento.

3. FORMULACION DEL PROBLEMA DE DISEÑO INTEGRADO DE PROCESOS

El *Diseño Simultáneo del Procesos y el Sistema de Control*, concebido como un problema de optimización de objetivos económicos y objetivos de control parece ser la mejor manera de tratar sistemáticamente la interacción entre estos dos aspectos. Así, el planteamiento global del *Diseño Integrado de Procesos*, contempla la síntesis algorítmica del proceso para determinar su estructura y dimensiones, el análisis de la *controlabilidad* de la planta, la selección de la estructura de control, y la sintonización del o los controladores necesarios.

Para una mayor compresión, se presenta la *formulación del problema de diseño integrado*, como el nivel superior de una serie de problemas que van surgiendo del planteamiento del diseño del proceso y la incorporación de aspectos relacionados con el control. Se empieza por la formulación del problema básico de síntesis económica del proceso.

- n) *Síntesis algorítmica del proceso basada en criterios económicos*

La *síntesis del proceso* consiste en la selección de la estructura económicamente óptima de la planta, luego, se procede al dimensionamiento y cálculo del punto de operación. La síntesis de un proceso puede plantearse de forma algorítmica, como un problema de optimización basado en el análisis de una *superestructura* que contiene todas las alternativas estructurales posibles, cuyo objetivo es encontrar la mejor estructura y parámetros de la planta en función de los objetivos económicos. El modelo matemático contiene variables continuas, que representan dimensiones y condiciones de operación, y variables discretas, que representan decisiones relacionadas con la estructura del proceso. El problema de optimización formulado a partir de la *superestructura* es entonces de la forma:

$$\min_{\mathbf{x}, \mathbf{y}} f(\mathbf{x}, \mathbf{y})$$

Sujeto a las restricciones:

$$h(\mathbf{x}, \mathbf{y}) = 0 \quad (3.4)$$

$$g(\mathbf{x}, \mathbf{y}) \leq 0$$

$$\mathbf{x} \in \mathbb{R}, \mathbf{y} \in \{0,1\}^l$$

Esta formulación se distingue como un problema de *optimización no lineal mixto entero (MINLP)*, donde:

El vector \mathbf{x} contiene variables continuas que pueden ser caudales, presiones, temperaturas o dimensiones de unidades, entre otras variables de proceso, las cuales deben satisfacer los balances de materia, de energía y ecuaciones de diseño.

El vector \mathbf{y} contiene variables discretas y representan decisiones estructurales, como la presencia ($y=1$) o no ($y=0$) de un elemento en el diagrama de flujo del proceso.

La función objetivo $f(\mathbf{x}, \mathbf{y})$ generalmente representa los costos a minimizar o algún otro criterio de desempeño relacionado con la economía del proceso, como por ejemplo, el gasto de energía, el rendimiento de las reacciones o la penalización por violación a restricciones ambientales.

Las variables de decisión deben satisfacer ecuaciones $h(\mathbf{x},\mathbf{y})$ relacionadas con los balances de masa y energía en el proceso. Las inecuaciones $g(\mathbf{x},\mathbf{y})$ vienen dadas por restricciones de operación, restricciones asociadas a especificaciones de diseño, regulaciones ambientales o condiciones lógicas.

Al resolver el problema de síntesis algorítmica se debe obtener *la estructura, dimensiones y punto de operación* de la planta que mejor satisface los objetivos representados por la función $f(\mathbf{x},\mathbf{y})$.

o) Diseño integrado a lazo abierto

En el análisis de controlabilidad se puede hacer una separación entre los criterios de desempeño a lazo abierto en base a índices que cuantifican el efecto de los parámetros de diseño sobre la capacidad de la planta de ser controlada sin considerar el esquema de control y los índices de lazo cerrado, que requieren la evaluación del comportamiento del sistema con los controladores.

En el *Diseño Integrado de Procesos*, la evaluación de la controlabilidad a lazo abierto generalmente se introduce en el problema de *Síntesis Algorítmica* (3.4) añadiendo nuevos objetivos de optimización y restricciones, obteniéndose un problema optimización de la forma (Luyben y Floudas, 1994):

$$\min_{\mathbf{x},\mathbf{y}} [f_1(\mathbf{x},\mathbf{y}), f_2(\mathbf{x},\mathbf{y}), f_3(\mathbf{x},\mathbf{y}), \dots]$$

Sujeto a las restricciones:

$$\begin{aligned} h(\mathbf{x},\mathbf{y}) &= 0 \\ g(\mathbf{x},\mathbf{y}) &\leq 0 \end{aligned} \tag{3.5}$$

$$\eta^{Min} \leq \eta(\mathbf{x},\mathbf{y}) \leq \eta^{Max}$$

$$\mathbf{x} \in \mathbb{R}, \mathbf{y} \in \{0,1\}^l$$

Las restricciones $\eta(\mathbf{x},\mathbf{y})$ representan límites sobre las medidas de *controlabilidad* para lograr que el diseño cumpla ciertas condiciones que garantizan su buen desempeño dinámico,

haciéndolo más fácil de controlar. Estas pueden ser índices de *controlabilidad entrada-salida* y/o restricciones derivadas de la aplicación de las métricas de *controlabilidad práctica*. Es importante agregar, que para la aplicación de las métricas de *controlabilidad práctica* requiere de un análisis preliminar de *controlabilidad de estado* basado en el modelo no lineal del proceso.

Las funciones $f_i(\mathbf{x}, \mathbf{y})$ indican la posibilidad de incluir, además de los índices económicos, índices de controlabilidad u otro tipo de índice que relacione la economía y la operabilidad de la planta como objetivos de optimización. Así, dependiendo del criterio del diseñador, el problema (3.5) puede ser una *optimización multiobjetivo* que contemple objetivos económicos y de control, o simplemente la optimización de un objetivo económico sujeto a restricciones de controlabilidad.

p) Diseño integrado a lazo cerrado

La sintonización del controlador requiere del conocimiento del tipo de controlador a utilizar. En el marco del problema de Diseño Integrado, el controlador es típicamente PI o PID. Como ya se mencionó, se encuentran muy pocos trabajos en los que se hayan utilizado técnicas de control avanzado como el *Control Predictivo Basados en Modelos* (MPC), esto, debido a la dificultad que se añade al ya complejo problema de optimización.

Los controladores PID son los más utilizados a nivel industrial, son fáciles de utilizar y la requieren una mínima cantidad de recursos para su instalación. Existe un buen número de técnicas para determinar sus parámetros óptimos, como las reglas de sintonización basadas en un punto de la respuesta y las técnicas basadas en la integración de la respuesta completa donde se minimizan índices de desempeño como el ISE.

El control predictivo basado en modelos (MPC) es una técnica de control avanzado que cada vez es más popular tanto en el ámbito industrial como el académico. Su funcionamiento es sencillo e intuitivo ya que está basado en un modelo del proceso que permite calcular su comportamiento futuro y así determinar de forma óptima las señales de control más adecuadas. Es fundamental un buen modelo para predecir la salida, dependiendo de este se

tiene el control predictivo en el espacio de estados, control predictivo generalizado (GPC) y el control predictivo no lineal si el modelo del proceso presenta esta característica.

La *selección de la estructura de control* tiene que ver con las decisiones relacionadas con las configuraciones alternativas del sistema de control como la selección de las variables a controlar y manipular, la especificación de las variables medidas, el emparejamiento entre las variables controladas y manipuladas, el esquema de control y la selección del tipo de controlador. Al incorporar la selección de la estructura de control al problema de diseño se añade un conjunto de variables binarias y_c , que establecen la existencia o no de los posibles pareamientos variable controlada-variable manipulada, y restricciones relacionadas con las acciones de control (Skogestad y Larsson, 1998; Bansal et al, 2002). Esto incrementa el número de variables y restricciones en el problema (3.5). En la selección de los posibles emparejamientos se puede obtener información a partir de algunas métricas de controlabilidad práctica.

El MPC es un controlador discreto, es decir, proporciona señales de control en instantes discretos de tiempo. El algoritmo de control consiste en resolver un problema de optimización con restricciones donde se minimiza una función objetivo que contiene el error de la salida predicha respecto a la referencia y los esfuerzos de control, para obtener la señal de control óptima.

En el *diseño simultáneo del proceso y el sistema de control*, la evaluación de la *controlabilidad* puede introducirse mediante índices basados en el modelo linealizado, métricas de controlabilidad práctica, pero generalmente es necesario simular la respuesta dinámica del sistema en base al modelo no lineal para evaluar los índices de desempeño del control. Por lo tanto, se deben incluir las ecuaciones diferenciales que representan el modelo del proceso en el problema de optimización convirtiéndolo en un *Problema de Optimización Dinámica Mixto Entero*.

En base a todo lo anterior, el *Diseño Simultáneo del Proceso y el Sistema de Control* se plantea de la siguiente manera:

Dados:

- *El modelo del proceso,*
- *el conjunto de restricciones para la operación factible de la planta,*
- *una superestructura del proceso que describa las operaciones unitarias potenciales del proceso y sus interconexiones*
- *un conjunto de posibles configuraciones de control*
- *un conjunto de objetivos de control y tipos de controladores*
- *el escenario de perturbaciones,*
- *datos de costos de los equipos, costos de operación y costos de control*
- *un horizonte finito de tiempo*

Encontrar:

El diseño del proceso y el esquema de control con el menor costo anualizado que garantice la operación factible en todo el horizonte de tiempo.

Esto se traduce en la siguiente formulación:

Minimizar	<i>El costo total anualizado</i>
Sujeto a	<i>Ecuaciones de diseño,</i> <i>el modelo en ecuaciones diferenciales,</i> <i>restricciones sobre la variabilidad en las salidas</i> <i>factibilidad de la operación en el tiempo,</i> <i>ecuaciones del algoritmo de control.</i>

la cual da origen al siguiente problema de optimización dinámica no lineal, mixto-entero:

$$\min_{\mathbf{d}, \mathbf{x}, \mathbf{v}, \dots, v_v} J\left(\dot{\mathbf{x}}(t_f), \mathbf{x}(t_f), \mathbf{z}(t_f), \mathbf{v}(t_f), \boldsymbol{\theta}(t_f), \mathbf{v}_v, \mathbf{d}, \boldsymbol{\delta}, t_f\right)$$

Sujeto a las restricciones: (3.6)

$$fd\left(\dot{\mathbf{x}}(t_f), \mathbf{x}(t_f), \mathbf{z}(t_f), \mathbf{v}(t_f), \boldsymbol{\theta}(t_f), \mathbf{v}_v, \mathbf{d}, \boldsymbol{\delta}\right) = 0$$

$$fa\left(\dot{\mathbf{x}}(t_f), \mathbf{x}(t_f), \mathbf{z}(t_f), \mathbf{v}(t_f), \boldsymbol{\theta}(t_f), \mathbf{v}_v, \mathbf{d}, \boldsymbol{\delta}\right) \leq 0$$

$$fp\left(\dot{\mathbf{x}}(t_f), \mathbf{x}(t_f), \mathbf{z}(t_f), \mathbf{v}(t_f), \boldsymbol{\theta}(t_f), \mathbf{v}_v, \mathbf{d}, \boldsymbol{\delta}\right) \leq 0$$

$$f_o\left(\dot{\mathbf{x}}(t_f), \mathbf{x}(t_f), \mathbf{z}(t_f), \mathbf{v}(t_f), \boldsymbol{\theta}(t_f), \mathbf{v}_v, \mathbf{d}, \boldsymbol{\delta}\right) = 0$$

$$g(\mathbf{v}_v, \mathbf{d}, \boldsymbol{\delta}) \leq 0$$

$$h(\mathbf{v}_v, \mathbf{d}, \boldsymbol{\delta}) = 0$$

$$t_0 \leq t \leq t_f$$

$$\mathbf{x} \in \mathbb{R}, \boldsymbol{\delta} \in \{0,1\}^{l+m}$$

donde:

x es el vector de variables de estado del sistema,

z son las variables de salida,

v denota al vector de variables manipuladas,

v_v es el conjunto de variables de operación invariantes con el tiempo como los puntos de ajuste de los controladores,

d son variables de diseño independientes de la operación como las dimensiones de los equipos,

δ son las variables binarias que definen la estructura del proceso y la del sistema de control,

θ es el escenario de perturbaciones,

fd representa las ecuaciones diferenciales correspondientes a los balances de masa y energía,

fa representa las ecuaciones algebraicas como relaciones hidráulicas y termodinámicas,

f_o son las condiciones iniciales del sistema dinámico,

h y *g* son las restricciones de igualdad y desigualdad invariantes con el tiempo, y

t_f es el tiempo final.

Esta es la formulación general del problema de *Diseño Integrado de Procesos*. Dependiendo del alcance del problema, la forma de evaluar la *controlabilidad* y la estrategia de control utilizada, entre otros aspectos, se puede tratar como un problema *multi-objetivo* o

como uno de un objetivo con restricciones. Puede ser de optimización dinámica (MIDO) o no, y puede ser simplemente un problema de optimización no lineal (NLP) o un problema de optimización mixta-entera (MINLP).

El resultado de la optimización es una planta que satisface el mejor compromiso entre economía y controlabilidad, pero para lograr esto generalmente se requiere de algoritmos de optimización poderosos, dada la complejidad del problema de optimización.

A continuación se considera el diseño de un sistema de reactores basado en el caso de estudio de Luyben (1993), a modo de ejemplo, para ilustrar la aplicación de la metodología de diseño integrado y la formulación del problema de optimización.

4. DISEÑO INTEGRADO DE UN SISTEMA DE REACTORES

El problema en estudio se trata de un sistema con tres rectores del tipo tanque agitado continuo (CSTR) conectados en serie, analizados por Luyben (1993) para demostrar el fuerte compromiso existente entre el diseño de una planta y su controlabilidad. El sistema consiste en una reacción de primer orden del tipo $A \rightarrow B$, exotérmica e irreversible que se desarrolla en fase líquida. Se desea determinar la configuración de reactores con la cual se obtiene el menor de los costos, es decir, con uno, dos o tres reactores. Adicionalmente, se considera que para el sistema en estudio todos los reactores tienen el mismo volumen y la misma temperatura de reacción. La relación entre la altura y el diámetro de cada reactor es igual a 2. El resto de las variables y parámetros están resumidos en las tablas 3.1 y 3.2.

Tabla 3.1. Variables de entrada para el sistema de reactores químicos.

Entradas	Valores
Conversión (x)	0,95
Temperatura de entrada (To)	70 °F
Fracción molar de A en la alimentación (zo)	1
Flujo de alimentación (F)	100 lbmol/h
Temperatura de entrada de refrigerante (Tjo)	70 °F

La formulación del problema se realiza en diferentes etapas para ilustrar como va variando el modelo de optimización de acuerdo a los objetivos del diseño integrado. Primero se presenta el problema de síntesis solo con criterios económicos luego, se plantea el diseño integrado a lazo abierto y finalmente, se presenta la formulación del problema de *diseño y control simultáneo*.

Tabla 3.2. Parámetros físicos del sistema.

Parámetros físicos	Valores
Densidad del líquido (ρ)	50 lb./ ft ³
Peso molecular (M)	50 lb./ lbmol
Constante pre-exponencial de reacción (ko)	4,08.10 ¹⁰ 1/ h
Energía de activación (Ea)	30.000 BTU/lbmol
Constante R	1,99 BTU/lbmol.R
Capacidad calórica del líquido (Cp)	0,75 BTU/ lb.°F
Densidad del refrigerante (ρ_j)	62,3 lb./ ft ³
Capacidad calórica del refrigerante (Cpj)	1 BTU/ lb.°F
Coeficiente de transferencia de calor (U)	300 BTU/ft ² .h.°F

El problema de síntesis parte de una superestructura donde se representan las tres alternativas de diseño: un reactor, dos reactores conectados en serie, o tres reactores conectados en serie. Esta representación se muestra en la figura 3.1.

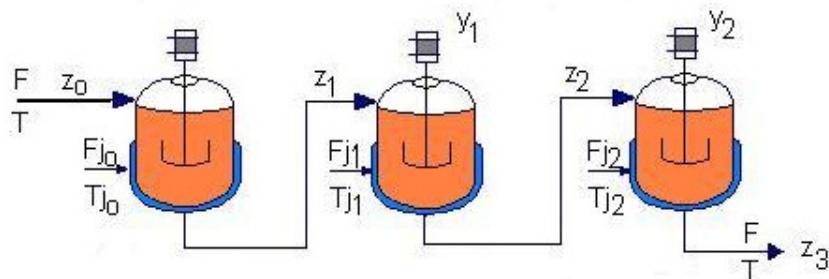


Figura 3.1. Superestructura que representa el sistema de uno a tres reactores químicos conectados en serie

A partir de la superestructura, se desarrolla el modelo matemático del sistema de reactores químicos en base a los balances de masa y energía en cada reactor.

$$\frac{dz_1}{dt} = \frac{F \cdot M}{V \cdot \rho} \cdot (z_o - z_1) - k \cdot z_1 \quad (3.7)$$

$$\frac{dT_1}{dt} = \frac{F \cdot M}{V \cdot \rho} \cdot (T_o - T_1) - \frac{k \cdot z_1}{M \cdot Cp} \cdot \Delta hr - \frac{Q_1}{V \cdot \rho \cdot Cp} \quad (3.8)$$

$$\frac{dz_2}{dt} = \frac{F \cdot M}{V \cdot \rho} \cdot (z_1 - z_2) - k \cdot z_2 + W1 \quad (3.9)$$

$$\frac{dT_2}{dt} = -\frac{k \cdot z_2 \cdot \Delta hr}{M \cdot Cp} - \frac{Q_2}{V \cdot \rho \cdot Cp} - W2 \quad (3.10)$$

$$\frac{dz_3}{dt} = \frac{F \cdot M}{V \cdot \rho} \cdot (z_2 - z_3) - k \cdot z_3 + W3 \quad (3.11)$$

$$\frac{dT_3}{dt} = \frac{k \cdot z_3 \cdot -\Delta hr}{M \cdot Cp} - \frac{Q_3}{V \cdot \rho \cdot Cp} - W4 \quad (3.12)$$

donde:

T_i : Temperatura de reacción del reactor i ,

z_i : Composición en el reactor i ,

Q_i : calor disipado en el reactor i , y

$$Q_i = U \cdot Ah \cdot (T_1 - T_{j_i}) \quad (3.13)$$

y los balances de energía en la camisa de cada reactor son:

$$\frac{dT_{j_1}}{dt} = \frac{F_{j_1}}{V_j} (T_{j_0} - T_{j_1}) + \frac{Q_1}{\rho_j \cdot Cp_j \cdot V_j} \quad (3.14)$$

$$\frac{dT_{j_2}}{dt} = \frac{F_{j_2}}{V_j} (T_{j_0} + T_{j_2}) + \frac{Q_2}{\rho_j \cdot Cp_j \cdot V_j} \quad (3.15)$$

$$\frac{dTj_3}{dt} = \frac{Fj_3}{Vj} (Tj_0 - Tj_3) + \frac{Q_3}{\rho_j \cdot Cp_j \cdot Vj} \quad (3.16)$$

Las variables $W1$ y $W2$ se utilizan para desactivar los balances asociados al segundo y tercer reactor en caso de que no existan. Su valor está definido por las restricciones lógicas que se presentan más adelante.

El objetivo del problema es determinar, simultáneamente, la estructura de la planta y las condiciones operacionales de temperatura, volumen de reactor y flujo de alimentación de refrigerante más económica. La función objetivo representa los costos de inversión, los cuales deben ser minimizados. Se deben satisfacer los balances de masa y energía de cada reactor y su camisa en estado estacionario. Además, para mantener la coherencia del modelo matemático y cumplir con los valores límites de las variables, se consideran restricciones lógicas y restricciones físicas sobre las variables. En resumen, la formulación completa del problema de optimización es la siguiente:

Función objetivo.

$$C = (1 + y_1 + y_2) \cdot 1916.9 \cdot (L)^{0.802} \cdot (D)^{1.066} \quad (3.17)$$

El diámetro, longitud característica y área de transferencia de calor, se obtienen de las relaciones geométricas asociadas a los tanques cilíndricos; dado que ambos reactores tienen el mismo volumen, estos parámetros son idénticos para cada reactor:

$$D = \left(\frac{2 \cdot V}{\pi} \right)^{1/3} \quad (3.18)$$

$$L = 2 \cdot D \quad (3.19)$$

$$Ah = 2 \cdot \pi \cdot (D)^2 \quad (3.20)$$

El conjunto de restricciones que componen el modelo viene dado por:

Restricciones operacionales sobre los balances de masa y energía

Se debe garantizar que la solución obtenida cumpla con el requerimiento de operación en estado estacionario:

$$r1 = \frac{F \cdot M}{V \cdot \rho} \cdot (z_o - z_1) - k \cdot z_1 \leq 0.00001 \quad (3.21)$$

$$r2 = \frac{F \cdot M}{V \cdot \rho} \cdot (T_o - T_1) - k \cdot z_1 \cdot \Delta hr - Q_1 \leq 0.05 \quad (3.22)$$

$$r3 = \frac{F \cdot M}{V \cdot \rho} \cdot (z_1 - z_2) - k \cdot z_2 + W1 \leq 0.00001 \quad (3.23)$$

Restricciones lógicas:

$$y_2 - y_1 \leq 0 \quad (3.24)$$

$$W1 \geq 0; \quad W1 = (1-x) \cdot k \cdot z_1 \quad (3.25)$$

$$W2 \geq 0; \quad W2 = (1-y_1) \cdot (k \cdot z_2 \cdot -\Delta hr) \quad (3.26)$$

$$W3 \geq 0; \quad W3 = (1-y_2) \cdot k \cdot z_2 \quad (3.27)$$

Restricciones físicas:

$$z_1 = (1-y_1) \cdot z_o \cdot (1-x) + y_1 \cdot \frac{F \cdot z_o}{F + V \cdot k} \quad (3.28)$$

$$z_2 = (1-y_2) \cdot z_o \cdot (1-x) + y_2 \cdot \frac{F \cdot z_1}{F + V \cdot k} \quad (3.29)$$

$$z_3 = (1-x) \cdot z_o \quad (3.30)$$

$$T_1 = T_2 = T_3 \quad (3.31)$$

$$120^{\circ}F \leq T_l \leq 200^{\circ}F \quad (3.32)$$

$$400 \text{ ft}^3 / \text{h} \leq Fj_1 \leq 1000 \text{ ft}^3 / \text{h} \quad (3.33)$$

Este modelo se traduce en un problema de *programación no lineal mixta-entera* (MINLP). Se tienen variables binarias de decisión, como y_1 y y_2 , para definir la existencia del segundo y del tercer reactor en las ecuaciones del modelo, y variables continuas como los volúmenes de los reactores (V), la temperatura de reacción T_i , y los flujos de refrigerante Fj_i .

Para tomar en cuenta el análisis de controlabilidad del sistema en etapas preliminares de diseño, se integran algunos criterios sencillos de controlabilidad a lazo abierto al planteamiento de problema de síntesis, convirtiéndolo en un problema de síntesis y diseño integrado de procesos. Estos criterios se introducen como restricciones del problema, los cuales están ligados a ciertas características que determinan una mayor estabilidad del sistema a lazo abierto y una mayor flexibilidad en la operación.

Para la estabilidad de un sistema, se imponen autovalores negativos a la matriz de estados (modelo linealizado). Por simplicidad, se considera suficiente evaluar la estabilidad del primer reactor para garantizar que el sistema total sea estable, pues en el caso de que existan dos o tres reactores en serie, el primero siempre está sometido a condiciones de operación más fuertes. Por otra parte, para este mismo sistema, Luyben (1993) propone también para el primer reactor la relación entre la tasa máxima de transferencia de calor y la tasa de transferencia de calor de diseño (Q_{1max}/Q_1) como una medida de la controlabilidad del sistema.

Restricciones de controlabilidad:

La restricción que garantiza la estabilidad del sistema a lazo abierto viene del análisis de los autovalores de la matriz de estados del primer reactor. Las variables de estado correspondientes son z_1 T_1 y Tj_1 y por lo tanto se obtienen tres autovalores (λ_i). Entonces:

$$\lambda_i \leq 0, i=1, 2, 3. \quad (3.34)$$

De manera de garantizar una mayor flexibilidad en la respuesta del sistema de transferencia de calor, se propone como restricción:

$$\frac{Q_{l,\max}}{Q_l} \geq U_q \quad (3.35)$$

donde U_q es el valor mínimo esperado para este parámetro.

El valor de Q_{\max} se obtiene de la ecuación (3.14) en estado estacionario, evaluada con el flujo máximo de refrigerante (cuatro veces el flujo de diseño) y la temperatura máxima de salida del refrigerante, que se calcula como:

$$Tj_{\max} = \frac{U \cdot Ah \cdot T_i + Fj_{\max} \cdot \rho j \cdot Cpj \cdot Tjo}{Fj_{\max} \cdot \rho j \cdot Cpj + U \cdot Ah} \quad (3.36)$$

5. CONCLUSION

Para resolver el conflicto *economía-controlabilidad* analizado en este capítulo se propone la formulación general del *problema de síntesis y diseño integrado*. Esta se basa en la minimización de costos económicos y de controlabilidad. Estos últimos pueden aparecer también como restricciones en el problema de optimización. Se distinguen restricciones de controlabilidad del proceso (lazo abierto) ó del sistema con sus controladores (lazo cerrado). El problema se completa con restricciones de operación estable (modelo estacionario) y restricciones físicas (límites) del proceso. También se definen las restricciones lógicas a imponer cuando se resuelve la selección de la estructura de la planta (síntesis).

En el tratamiento de los procesos no lineales se incluyen las condiciones de controlabilidad local de estado tanto en la definición de los esquemas de control (pareamientos) como en las restricciones físicas sobre las variables manipuladas (entradas).

El capítulo siguiente está dedicado a los algoritmos genéticos para la resolución de los problemas con restricciones como los planteados aquí.

CAPÍTULO IV

ALGORITMOS GENÉTICOS

1. INTRODUCCIÓN

Los métodos de optimización basados en cálculo o deterministas han sido ampliamente estudiados (Floudas, 1995; Grossmann, 1996; Caballero et al., 2007). En su formulación, un conjunto de condiciones necesarias y suficientes deben ser satisfechas para garantizar la existencia de las soluciones. En este tipo de cálculo existen a su vez métodos directos e indirectos. Los métodos indirectos obtienen los extremos locales a través de la resolución de un sistema de ecuaciones no lineales obtenidas de igualar el gradiente de la función objetivo a cero. Por otro lado, los métodos directos buscan los óptimos locales apoyados en la forma de la función y moviéndose en la dirección del gradiente local, escalando la función objetivo hasta llegar a un valor pico (extremo).

Los algoritmos de búsqueda basados en cálculo presentan algunas desventajas. En primer lugar, dependen de la existencia de derivadas por lo que pueden fallar si el espacio de búsqueda presenta discontinuidades. Por otra parte, el resultado obtenido puede ser un óptimo local porque no logran explorar la totalidad del espacio de posibles soluciones, ya que la búsqueda se basa en la información proporcionada por el gradiente de la función objetivo.

Por otra parte hay esquemas de optimización enumerativos que son algoritmos que buscan a través de un espacio de búsqueda finito, o de un espacio discretizado infinito, el valor de la función objetivo en cada punto del espacio de búsqueda. A pesar de que la simplicidad de este tipo de algoritmos es atractiva y de que la enumeración es un tipo de búsqueda “muy humano” (cuando el número de posibilidades es pequeña), tales esquemas pierden la carrera en busca de la robustez por una simple razón: falta de eficiencia.

Además, existen métodos de búsqueda aleatoria dirigida que se basan en las técnicas enumerativas pero usan información adicional para dirigir la búsqueda. Dos subclases importantes de estos algoritmos son el enfriamiento simulado (“*simulated annealing*”), basado en la transformación termodinámica para alcanzar estados de energía mínima, y los métodos basados en la selección natural comúnmente conocidos como computación evolutiva.

El desarrollo de las técnicas de computación evolutiva se remota a la década de los sesenta. Entre ellas se encuentran: las estrategias evolutivas (Rechemberg, 1973), la programación evolutiva (Fogel, 1994), la programación genética (Koza, 1992) y los algoritmos genéticos (Holland, 1975). Aunque algunos de estos métodos han sido diseñados para aplicaciones distintas y presentan diferencias en el tipo de datos utilizado y su posterior manejo, comparten el mismo principio de funcionamiento: una población de individuos está sometida a un proceso de transformación y durante su evolución los individuos luchan por sobrevivir, es decir, se “optimizan” (Michalewicz, 1996).

2. LOS ALGORITMOS GENETICOS

2.1 GENERALIDADES

Los algoritmos genéticos son métodos estocásticos que generan soluciones a problemas de búsqueda y optimización imitando, desde el punto de vista matemático, los mecanismos presentes en el proceso de evolución por selección natural. Esta técnica surge como producto de la investigación desarrollada por el Prof. Holland y sus estudiantes desde principios de los años 60. Se dio a conocer en *Adaptation in Natural and Artificial Systems* (Holland, 1975) dando origen a un nuevo campo de investigación en el área de la inteligencia artificial.

En la naturaleza, los organismos evolucionan de acuerdo a los principios de selección natural y supervivencia del más apto. Dentro de una población, los individuos con mayor capacidad de adaptación tienen mayor probabilidad de sobrevivir, generando mayor descendencia y transmitiendo su carga genética con atributos deseables a las generaciones futuras. Así, la evolución de las especies se puede interpretar como la búsqueda, entre un gran número de posibilidades, de la solución del problema de adaptación al medio, creando nuevas soluciones mediante la combinación de las características de los organismos existentes. La

formulación de los algoritmos genéticos surge de la analogía de este proceso con la solución de un problema de optimización.

Los seres vivos generalmente se organizan en poblaciones constituidas por individuos cuya información genética se encuentra almacenada en los cromosomas. Los cromosomas son cadenas ADN contenidas en las células del organismo que condensan toda la información relativa a las características del individuo. Cada característica está representada por genes, que son bloques funcionales de ADN localizados en una determinada posición (lugar) dentro del cromosoma. Las diferentes posibilidades de escoger una característica se denominan alelos. El genotipo es el conjunto de genes que constituye el material genético de un individuo, y da lugar al fenotipo que es en cambio el conjunto de las características físicas del individuo.

La descendencia se genera por el mecanismo de reproducción (cruce), mediante el cual se recombinan los genes de los padres para formar los cromosomas de los nuevos individuos. Eventualmente se puede producir una mutación en la descendencia, la cual es una modificación espontánea en el genotipo que se refleja como un cambio favorable o no en sus características y se puede transmitir a las generaciones sucesivas, introduciendo variabilidad genética en la población. El éxito de un individuo en el proceso de selección natural depende de su capacidad de reproducirse y sobrevivir, de manera de generar más descendientes y transmitir sus características a la siguiente generación generando un cambio en la población. Este cambio es también adaptativo, ya que los organismos adquirirán características favorables para su supervivencia y reproducción en el medio en que viven.

Trasladando estos conceptos de biología y genética al lenguaje matemático (Michalewicz, 1996), un individuo es una solución potencial del problema de optimización y está representado como un cromosoma dentro de una “población” de posibles soluciones (vector dentro de un conjunto). Los genes son las unidades que constituyen el cromosoma y codifican un elemento o característica de la posible solución ocupando un lugar particular en el mismo (elementos del vector). Los alelos son las distintas alternativas que puede presentar una misma característica, es decir, los diferentes estados que puede manifestar un gen (Valores posibles para cada elemento o variable). El genotipo viene a ser la información contenida en el

cromosoma codificado y el fenotipo es el significado real de la solución, es decir, el espacio de búsqueda.

La codificación de los cromosomas establece la relación entre el espacio del genotipo y el espacio del fenotipo. En el proceso de búsqueda llevado a cabo por el algoritmo genético, la transformación de las posibles soluciones se realiza manipulando representaciones codificadas, es decir, se trabaja en el espacio del genotipo.

La evolución de la población se traduce como la búsqueda del óptimo en el espacio de posibles soluciones, para lo cual se dispone de una función que permite evaluar la capacidad de cada individuo. La búsqueda se lleva a cabo mediante la aplicación de operadores genéticos, análogos a los procesos de selección natural, cruce, y mutación. El cruce permite intercambiar información entre dos posibles soluciones para crean una nueva, la mutación es un cambio al azar que ocurre con cierta probabilidad en uno o más genes del cromosoma para introducir cierta variabilidad en la población, mientras que la selección dirige la búsqueda hacia el área deseada en el espacio de posibles soluciones.

Este procedimiento se resume a continuación el esquema del *algoritmo genético básico* donde $P(t)$ es una población de individuos candidatos a solución de un problema determinado.

Inicio

$t \leftarrow 0$

Inicializar $P(t)$

Evaluar $P(t)$

Mientras (NO se cumpla la condición de parada) hacer

Inicio

$t \leftarrow t+1$

seleccionar $P'(t)$ de $P(t-1)$

$P''(t) \leftarrow$ cruce y mutación de $P'(t)$

evaluar $P''(t)$

Seleccionar $P(t)$ de $P''(t)$ y $P(t-1)$

Fin

Fin

Existen cinco componentes básicos que debe poseer una implementación del algoritmo genético (Michalewicz, 1996, Gen y Cheng, 2000): la representación de las posibles soluciones del problema (codificación), la manera de crear una población inicial de posibles soluciones, la función de evaluación que permite clasificar las soluciones en función de su calidad (*fitness* o *función de adaptación*), los operadores genéticos que alteran la composición genética de los descendientes durante la reproducción y los valores de los parámetros del algoritmo genético como la *tasa de mutación*, la *tasa de cruce* y el *tamaño de la población*.

Las características de los algoritmos genéticos los convierten en herramientas adecuadas para la resolución de los problemas de optimización complejos, no lineales y con restricciones que surgen en el área de investigación de operaciones e ingeniería. Una de sus principales ventajas, es que no exigen mayores requerimientos matemáticos sobre el problema de optimización, solo es necesaria la evaluación de la función objetivo para su implementación (Michelewicz et al, 1996). Pueden resolver problemas definidos en espacios de búsqueda continuos, discretos o mixtos y existen estrategias diseñadas para manejar eficientemente problemas de optimización con restricciones. Por otra parte, llevan a cabo una búsqueda multidireccional al mantener una población de soluciones potenciales y permitir el intercambio de información entre los elementos que la integran (Michalewics, 1996).

Estos métodos evolutivos combinan elementos de búsqueda dirigida y estocástica que les permite alcanzar un balance adecuado entre dos objetivos, aparentemente conflictivos, la explotación de la mejor solución y la exploración amplia del espacio de búsqueda (Gen y Cheng, 2000). Lo primero se refiere a la búsqueda local, que explota la mejor solución escalando hacia un óptimo cercano. Lo segundo, se refiere a la búsqueda aleatoria, que explora todo el espacio de posibles soluciones y permite escapar de óptimos locales. Sin embargo, para lograr el mencionado equilibrio todos los componentes del algoritmo genético deben ser examinados y ajustados cuidadosamente.

El algoritmo genético propuesto por Holland (1975), fue concebido originalmente con una representación de cadenas binarias de longitud fija (“*bit strings*”) para los cromosomas y con operadores de cruce y mutación binarios. Esta versión clásica se ha denominado *algoritmo genético básico, canónico* o simplemente se distingue como *algoritmo genético de*

codificación binaria. Este es el algoritmo genético más empleado, principalmente por la existencia de resultados teóricos que demuestran su efectividad (Holland, 1975; Goldberg, 1989; Goldberg, 1991).

A medida que se amplía su campo de aplicación es necesario desarrollar variantes del algoritmo genético clásico para algunos casos donde se requieren representaciones del cromosoma más complejas y operadores genéticos apropiados a estas estructuras o a las características del problema a resolver (Wright, 1991; Michalewicz et al, 1992; Herrera et al, 1994; Bui y Moon, 1995, Ronald, 1997). Entre otras ideas, se propone incorporar conocimiento del problema en la codificación del cromosoma y en los operadores genéticos (Davis, 1989). Así, se encuentran técnicas de codificación diseñadas para aplicaciones específicas como: *los vectores de números de punto flotante y vectores de números enteros* principalmente para problemas de optimización de funciones, *listas organizadas* para problemas de planificación y *matrices bidimensionales* para el problema de transporte lineal, entre otras (Herrera et al, 1998). Algunos autores, distinguen estas variantes denominándolas programas evolutivos “*evolution programs*” (Michalewicz; 1996) aunque otros autores simplemente generalizan la denominación de algoritmos genéticos y los clasifican de acuerdo a la técnica de codificación utilizada o la aplicación particular (Ronald, 1997; Herrera et al, 1998; Gen y Cheng, 2000).

Gran parte de los estudios teóricos que explican el funcionamiento de los algoritmos genéticos se han desarrollado para la representación binaria debido a su simplicidad.

Además, Gen y Cheng (2000) señalan la diferencia entre los métodos directos e indirectos para codificar las soluciones. En el primer caso se utiliza toda la información de la posible solución para formar el cromosoma, pero esto puede ser poco práctico cuando se trata de problemas complejos pues se pueden generar soluciones redundantes o ilegales. Los métodos indirectos utilizan solo una parte de la solución en el cromosoma y un decodificador para generar el resto de la información. Así, el algoritmo genético enfoca la búsqueda únicamente en la parte interesante de la solución. Sin embargo, para problemas de optimización de funciones con parámetros reales y problemas de optimización con restricciones en espacios continuos, típicos en aplicaciones de ingeniería, diversos autores

(Wright, 1991; Herrera et al, 1998; Gen y Cheng, 2000, Deb et al, 2001) recomiendan el uso de *algoritmos genéticos de codificación real*. En efecto la codificación real se presta especialmente para su aplicación en los problemas complejos de síntesis y diseño simultáneo de procesos y sistemas de control, que son el objeto de este trabajo.

2.2 ALGORITMOS GENÉTICOS DE CODIFICACIÓN REAL.

A continuación se presenta una descripción general del algoritmo genético con codificación real en la resolución de problemas de optimización con restricciones. Se explican con más detalle los operadores genéticos y los parámetros del algoritmo, y finalmente, se presentan las técnicas de tratamiento de restricciones más utilizadas.

Las limitaciones de la codificación binaria han despertado un gran interés en esta variante de los algoritmos genéticos, especialmente en el área de ingeniería, aunque existen paralelamente otros métodos, como las estrategias evolutivas y la programación evolutiva, formulados originalmente para trabajar con números reales. El empleo de algoritmos genéticos de codificación real no es aceptado por la mayoría de los investigadores teóricos por razones formales, sin embargo, en la práctica se encuentran aplicaciones donde han demostrado ser más efectivos (Davis, 1991; Wright, 1991; Michalewicz, 1992; Herrera et al, 1994, Herrera et al, 1995). Se han hecho esfuerzos para justificar de manera teórica la validez de la codificación real (Antonisse, 1989; Wright, 1991; Radcliffe, 1992; Surry and Radcliffe, 1997) los cuales se dirigen específicamente a la interpretación de la teoría de los esquemas en alfabetos de alta cardinalidad (Antonisse, 1989) y a la posibilidad de definir operadores genéticos que trabajen directamente en el espacio del fenotipo (Radcliffe, 1992).

La representación de las soluciones de problemas que involucran parámetros reales, en un algoritmo genético de codificación real, se formula de manera natural ya que la estructura topológica del espacio del *genotipo* es idéntica a la del espacio del *fenotipo* (Gen y Cheng, 2000). Esto es ventajoso porque evita los procesos de codificación y decodificación requeridos cuando se emplea la representación binaria, aumentando la velocidad de algoritmo genético (Herrera et al, 1998), y facilita el uso de operadores genéticos basados en técnicas aplicadas en métodos convencionales (Gen y Cheng, 2000).

El uso de códigos de parámetros reales mejora el desempeño del algoritmo genético en dominios de alta dimensionalidad y precisión, donde la codificación binaria con cadenas de longitud fija falla. Cabe destacar, además, que la representación con números reales permite explotar la gradualidad de las funciones de variables continuas, ya que pequeños cambios en las variables corresponden a pequeños cambios en la función objetivo (Herrera et al, 1998; Elliot et al, 2006). Además, esta técnica de codificación presenta la ventaja de lograr un mejor ajuste local de las soluciones utilizando operadores adecuados, lo que es difícil para los códigos binarios debido a la existencia de Riscos de Hamming (Herrera et al, 1998).

Otro aspecto importante, señalado en Davis (1989) y discutido por Michalewicz (1996) y Herrera et al (1998), es la necesidad de introducir conocimiento del dominio en el algoritmo de búsqueda. De acuerdo a Davis (1989) muchos de los problemas del mundo real no pueden ser manejados por el algoritmo genético de codificación binaria dado que, generalmente, existe conocimiento asociado al dominio real que debe ser considerado al transformar la solución por lo cual se debe incorporar este conocimiento al algoritmo genético, ya sea en el proceso de decodificación o expandiendo el conjunto de operadores genéticos. Este conocimiento es mucho más fácil de integrar al algoritmo utilizando codificación real, sobre todo en el caso de problemas con restricciones no triviales (Herrera et al, 1998).

Como ya se mencionó, en problemas de optimización con parámetros reales en dominios continuos, la codificación real es una manera sencilla de representar las posibles soluciones, ya que el espacio del genotipo y el fenotipo suelen coincidir. La técnica más utilizada en estos casos es la representación de *un gen-una variable* donde, el cromosoma que representa cada posible solución con N variables es un vector de números de punto flotante en R^N de la forma: $x = (x_1, x_2, \dots, x_N)$. La longitud de los cromosomas es la misma del vector constituido por las variables de la posible solución, cada gen es una variable x_i y los alelos son los posibles valores del parámetro x_i . Para aprovechar las ventajas que ofrece esta representación es necesario el uso de operadores genéticos diseñados para trabajar con números reales.

Existen otras estrategias para representar las soluciones con códigos reales, como el uso de enteros para representar cada dígito de la variable de manera de lograr un compromiso entre la codificación real y la binaria manteniendo los operadores tradicionales pero incrementando

la cardinalidad del alfabeto utilizado (Coello, 1999). En este caso, para una posible solución $x = (x_1, x_2, \dots, x_N)$, con valores numéricos establecidos a modo de ejemplo: $x = (10, 2.5, \dots, 5)$

estaría representada por una cadena de la forma: $\begin{bmatrix} 10 & 25 & \cdots & 5 \\ \underbrace{x_1}_{\cdot} & \underbrace{x_2}_{\cdot} & & \underbrace{x_N}_{\cdot} \end{bmatrix}$, suponiendo para cada gen una posición fija para el punto decimal. La precisión está limitada por la longitud de la cadena cromosómica, pero esta es modificable.

Una representación alternativa con números reales contempla el uso de enteros largos como genes, entonces la posible solución dada por el vector $x = (10.5, 2.517, \dots, 50.1)$ estaría representada por el cromosoma $x = (105, 2517, \dots, 501)$. Al usar esta técnica es necesario redefinir los operadores genéticos como cuando se trata con números reales pero mejora la eficiencia del procesador ya que el almacenamiento de enteros requiere menos memoria que el de números reales (Coello, 1999).

2.3 OPERADORES GENÉTICOS

Los algoritmos genéticos combinan estrategias de búsqueda aleatoria y búsqueda dirigida, lo que permite explorar en profundidad el espacio de posibles soluciones y explotar la mejor. Esto se logra mediante la aplicación de los operadores genéticos, de selección, cruce y mutación. El operador de selección dirige la búsqueda hacia zonas favorables, mientras que el cruce y la mutación permiten recorrer el espacio generando nuevas posibilidades en el conjunto de variables de decisión. En la figura 4.1 se muestra el procedimiento general de un algoritmo genético.

Los operadores genéticos deben seleccionarse o diseñarse de manera de mantener un balance adecuado entre la búsqueda aleatoria que llevan a cabo los operadores de recombinación (cruce y mutación) y la búsqueda dirigida por el operador de selección (Deb, 2000; Gen y Cheng, 2000). En el caso de los operadores de recombinación, señalan que el cruce explora toda el área alrededor de un óptimo local mientras que la mutación efectúa una búsqueda local para tratar de mejorar la solución (Gen y Cheng, 2000).

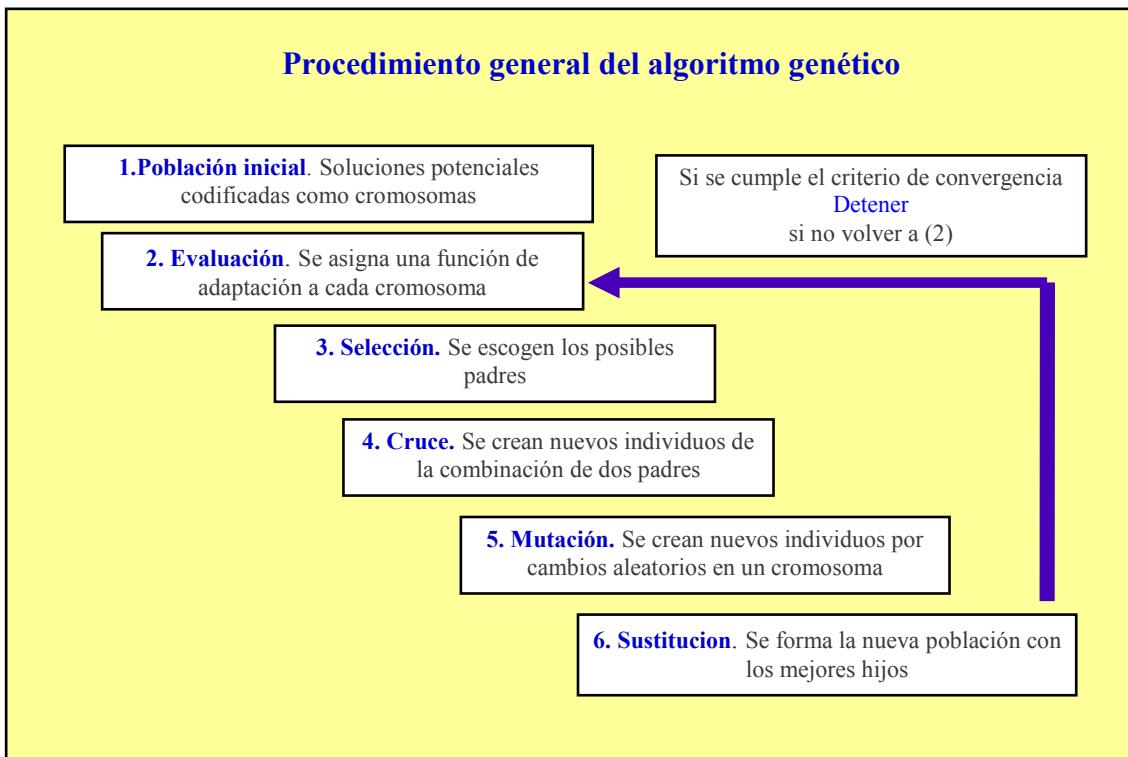


Figura 4.1. Evaluación de la función de adaptación para cada posible solución mientras se resuelve el problema con AG

En cuanto al operador de selección, se debe tomar en cuenta que si se aplica mucha presión selectiva favoreciendo la reproducción de las mejores soluciones de la población ésta perderá diversidad muy rápido, por lo cual se debe disponer de operadores de recombinación capaces de producir un conjunto de soluciones lo suficientemente distintas entre sí y de los padres. De lo contrario, el exceso de presión selectiva puede llevar a la convergencia prematura de la población a una solución sub-óptima. Por otra parte, si la presión selectiva es muy baja el algoritmo genético se comportará como un método de búsqueda aleatorio (Deb, 2000).

Algunos operadores genéticos empleados en el algoritmo genético de codificación real se derivan de los operadores básicos para la representación binaria. Pero, se han desarrollado algunas variantes de los operadores para aprovechar las ventajas de la representación con números reales. A continuación se presentan las técnicas para llevar a cabo los procesos de selección, cruce y mutación en el algoritmo genético desarrollado, y los operadores genéticos especializados para la representación real de un gen-una variable.

2.3.1 Selección

Este mecanismo imita el proceso de escogencia que determina la supervivencia del más apto en la naturaleza. Por lo general es el primer operador que se aplica a la población, y su función es seleccionar para su reproducción preferiblemente a los individuos o cromosomas que exhiben la mejor función de adaptación de manera que éstos puedan contribuir con la mayor cantidad de descendientes en la siguiente generación.

Se han propuesto distintas técnicas de selección. Entre estos se encuentran los métodos proporcionales, como la selección por ruleta, los métodos de jerarquización, escalamiento, la selección por torneo, las estrategias de selección ($\mu+\lambda$) y (μ,λ) y el criterio de elitismo. En Goldberg y Deb (1991) se analizan en profundidad los distintos métodos de selección.

A continuación se describen estas estrategias de selección, considerando que la población actual P está constituida por N cromosomas o individuos (p_i) cuya función de adaptación es $f(p_i)$. El conjunto de individuos seleccionados se denomina P' .

a) *Selección por ruleta:*

Al aplicar este método, la probabilidad de selección o supervivencia de cada individuo es proporcional a su función de adaptación (“fitness”). La probabilidad de selección se calcula como (Herrera et al, 1998):

$$p_{sel}(p_i) = \frac{f(p_i)}{\sum_{j=1}^N f(p_j)} \quad (4.1)$$

Siguiendo un procedimiento que imita una ruleta, se asigna a cada individuo de la población una porción de la ruleta que corresponde a su probabilidad de selección ($p_{sel}(p_i)$) y luego, se “gira” la ruleta un número de veces igual a la cantidad de individuos deseado para la reproducción. En cada “lanzamiento” de la ruleta se escoge un individuo que formará parte de P' . Esta estrategia de selección se denomina *muestreo estocástico con reemplazo*, fue propuesto por Holland (1975), y es el más usado.

Sin embargo, Herrera et al (1998) señala que el algoritmo más eficiente es el *muestreo estocástico universal*. Este método simula una ruleta con N punteros distribuidos uniformemente a lo largo de su circunferencia, por lo cual, en un solo giro, se determina el número de copias de cada individuo (p_i) que integrarán P' , el cual está dado por el número de punteros que señalan la porción de ruleta que corresponde a (p_i).

Los esquemas de selección proporcional como los recién mencionados, además de ser inherentemente lentos (Deb, 2000), presentan el inconveniente de que un cromosoma sobresaliente puede dominar el proceso de selección en las primeras generaciones creando peligro de convergencia prematura. Además, cuando la población presenta valores de la función de adaptación (FA) muy cercanos, la presión selectiva es muy baja y la búsqueda tiende a hacerse aleatoria (Gen y Cheng, 2000). Los mecanismos de escalamiento y jerarquización se han propuesto para reducir este problema.

b) Jerarquización (Ranking)

En el método de jerarquización se ordena a la población del mejor al peor individuo de acuerdo al valor de la función de adaptación (FA) y se asigna a cada uno una probabilidad de selección proporcional a esta posición. Esto permite una mayor diferenciación de la población, aún en generaciones posteriores cuando sus valores de FA son muy cercanos, regulando la presión selectiva sobre los individuos (Goldberg y Deb, 1991). Se encuentran estrategias de jerarquización lineal y de jerarquización exponencial.

A modo de ejemplo se presenta un método de jerarquización lineal en la cual la probabilidad de selección de cada individuo está dada por:

$$p_{sel}(p_i, t) = Min - (Max - Min) \frac{Rank(i, t) - 1}{N - 1} \quad (4.2)$$

donde t es la generación actual, Max es un valor escogido arbitrariamente tal que $0 \leq Min \leq 1$ y $Max = (2 - Min)$. Estos valores constituyen los nuevos espacios asignados a cada individuo en la ruleta.

c) Escalamiento

Los métodos de escalamiento cumplen con dos objetivos básicos, mantener una diferencia razonable entre los rangos relativos de la función adaptación de cada cromosoma y prevenir la supremacía de los individuos sobresalientes en las primeras generaciones para estimularla en las generaciones posteriores. La idea en este método es realizar una transformación para adecuar los valores reales de la función de adaptación a un rango de valores deseados de probabilidad de selección que luego se utilizarán para delimitar el espacio de cada individuo en la ruleta.

Existen distintas estrategias de escalamiento de la función de adaptación, que pueden ser lineales, logarítmicas, de truncación *sigma* y de potenciación, entre otras. Si la relación entre la función de adaptación original y la escalada es constante se habla de *escalamiento estático*, si la transformación es variable respecto a otros parámetros se denomina *escalamiento dinámico* (Gen y Cheng, 2000).

En el caso de la técnica de *escalamiento por truncación sigma*, la probabilidad de seleccionar un individuo depende de su “*fitness*”, la media y la desviación estándar de la población.

$$p_{sel}(p_i, t) = \begin{cases} 1 + \frac{f(p_i, t) - \bar{f}(t)}{2\sigma(t)} & \text{si } \sigma(t) \neq 0 \\ 1 & \text{si } \sigma(t) = 0 \end{cases} \quad (4.3)$$

donde:

\bar{f} es la media de la función de adaptación en la población actual y $\sigma(t)$ es la desviación estándar de la población (instante t).

De esta manera, en las primeras generaciones la desviación estándar tendrá un valor muy alto impidiendo que los mejores individuos tomen porciones más grandes de la ruleta. Al final, la desviación estándar será más baja lo cual favorecerá la probabilidad de selección de los individuos más aptos (Coello, 2009).

d) Selección por torneo

Este procedimiento de selección se lleva a cabo seleccionando aleatoriamente un conjunto de individuos y tomando los mejores para la reproducción. Esto se repite el número de veces que sea necesario hasta completar la cantidad de cromosomas requerido para la reproducción (P'). El número de cromosomas que integran el conjunto seleccionado se denomina *tamaño del torneo*. Comúnmente se lleva a cabo entre dos individuos tomando el nombre de *torneo binario* (Gen y Cheng, 2000).

Existe una variante probabilística del torneo, que se diferencia de la anterior en la selección del ganador del torneo. En lugar de seleccionar siempre el mejor como ganador del torneo, se introduce la probabilidad de seleccionar algunas veces al peor individuo como ganador de acuerdo a un parámetro p con $0.5 \leq p \leq 1$ (Coello, 2009).

Este método se ha hecho popular porque es similar a la jerarquización en términos de la presión de selección pero es computacionalmente más eficiente pues no requiere del cálculo de probabilidades de selección ni de otros parámetros aparte del “*fitness*” (Deb, 2000).

e) Selección $(\lambda+\mu)$ y selección (λ,μ)

Estas son técnicas deterministas tomadas de las Estrategias Evolutivas. Solo se escogen los mejores cromosomas de la población (P) y el método no permite repetición de cromosomas en P . En la estrategia $(\lambda+\mu)$, en una generación dada donde hay M padres y L descendientes, se ordenan los $M+L$ miembros de la población de acuerdo a su función de adaptación y se escogen los M mejores como padres de la siguiente generación. En cambio, al aplicar la estrategia (λ, μ) con $L > M$, solo los L descendientes se ordenan de acuerdo a su función de adaptación y los M mejores cromosomas de dicho conjunto se escogen como los padres de la siguiente generación (Gen y Cheng, 2000; Costa y Oliveira, 2000).

f) Elitismo

Este es un criterio que muchas veces se combina con otros métodos de selección, para mejorar el desempeño de los algoritmos genéticos.

Cuando se crea una nueva población a partir de una anterior a través de mutaciones y de cruces, existe una alta probabilidad de que se pierda el mejor cromosoma. El *elitismo* obliga a

mantener el mejor o los mejores cromosomas de la población actual en la nueva generación antes del apareamiento, previniendo la pérdida de la mejor solución encontrada.

2.3.2. Cruce

El proceso de cruce se lleva a cabo sobre los individuos seleccionados para la reproducción, permitiendo intercambiar información entre dos cromosomas seleccionados como *padres*, los cuales combinan sus genes para producir nuevos individuos, que probablemente presentarán mejores características. El cruce se usa como el operador principal del algoritmo genético y su función es la exploración del espacio de búsqueda mediante la generación de una amplia gama de posibles soluciones (Herrera et al, 1998).

Otro aspecto importante es que la búsqueda debe desarrollarse tratando de conservar, dentro de lo posible, la información almacenada en los individuos seleccionados como padres pues éstos son los que presentan las mejores características. La *tasa de cruce* es un parámetro que ayuda a mantener los mejores *padres*, definiendo una probabilidad alta de que un individuo seleccionado sea sometido al operador de cruce, de lo contrario, se copia directamente en la siguiente generación (Deb, 2000).

La definición de los operadores de cruce depende fuertemente de la estrategia de representación utilizada. Entre los distintos operadores de cruce que se encuentran en la literatura cabe mencionar: los clásicos para codificación binaria (Holland, 1975; Goldberg, 1989, Eshelman et al, 1989) y los operadores diseñados para codificación real (Herrera et al, 1998; Gen y Cheng, 2000; Deb, 2000).

A continuación se describen los operadores clásicos que pueden extenderse directamente a la codificación real y los operadores especializados tomados de la recopilación de Herrera et al, (1998) y Herrera et al, (2003) para este tipo de codificación.

El cruce clásico para codificación binaria (Holland, 1975; Goldberg, 1989) se lleva a cabo entre dos cadenas de bits que representan a los cromosomas X y Y con N genes y longitud L , escogidos al azar para cruzarse entre el conjunto de posibles padres:

$$X = \underbrace{110110}_{x_1} \underbrace{101011}_{x_2} \underbrace{110}_{x_3} \cdots \underbrace{0010}_{x_N} \quad Y = \underbrace{00110}_{y_1} \underbrace{111111}_{y_2} \underbrace{001}_{y_3} \cdots \underbrace{0100}_{y_N}$$

Se toma una posición R al azar en el intervalo $\{1, \dots, L-1\}$ que determina el lugar donde se dividen ambos cromosomas para intercambiar sus características y producir los descendientes:

$$H_1 = \underbrace{11010}_{h_1^1} \underbrace{101011}_R \underbrace{001 \dots 0100}_{x_N^1} \quad H_2 = \underbrace{00110}_{y_1} \underbrace{111111}_R \underbrace{110 \dots 0010}_{y_N}$$

En la figura 4.2 se ilustra cómo se lleva a cabo este tipo de cruce.

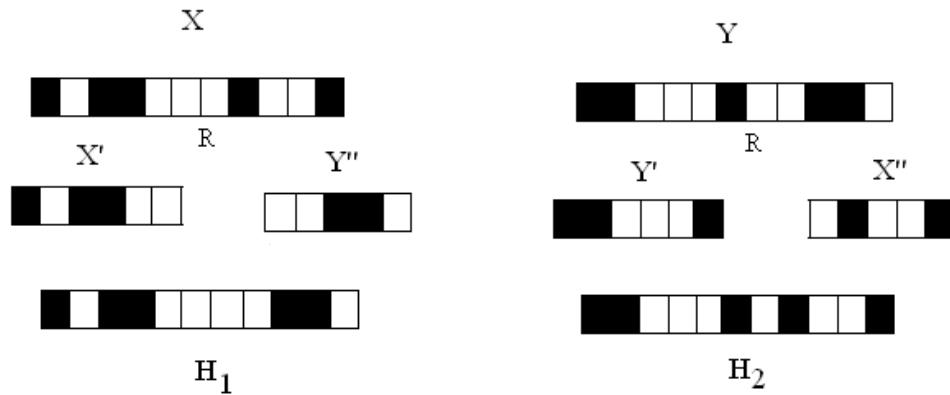


Figura 4.2. Cruce simple para codificación binaria

Entre las variantes más utilizadas del cruce simple binario se encuentra el “*n-point crossover*” (Eshelman et al, 1989) que es una generalización del cruce simple, donde se escoge un número de posiciones del cromosoma R : R_1, R_2, \dots, R_p para dividir en segmentos los padres los cuales se intercambiarán para formar los dos descendientes correspondientes. En Eshelman et al (1989) se pueden encontrar más operadores para la codificación binaria.

Para describir los operadores diseñados para codificación real se consideran dos padres $X = (x_1, x_2, \dots, x_N)$, $Y = (y_1, y_2, \dots, y_N)$, tomados al azar entre los individuos seleccionados para reproducción, donde N el número de genes que es equivalente al número de variables. El intervalo de acción del gen i se define como $[a_i, b_i]$.

g) *Cruce plano (“Flat crossover”)*

Esta técnica fue propuesta por Radcliffe (1991). Al combinar los padres se produce un descendiente $H = (h_1, h_2, \dots, h_N)$ generando aleatoriamente los valores de cada gen h_i en el intervalo comprendido entre los valores de los genes correspondientes de los padres $[x_i, y_i]$.

h) Cruce simple (Wright, 1991, Michalewicz, 1996)

Se escoge aleatoriamente un punto de ruptura $R \in \{1, \dots, N-1\}$, donde se dividen los cromosomas que representan a los padres para combinar ambos vectores y construir dos descendientes tales que:

$$H_1 = (x_1, \dots, x_R, y_{R+1}, \dots, y_N) \quad H_2 = (y_1, \dots, y_R, x_{R+1}, \dots, x_N)$$

i) Cruce discreto (Mühlenbein et al., 1993)

Esta técnica produce un descendiente $H = (h_1, h_2, \dots, h_N)$ donde cada gen h_i se obtiene tomando aleatoriamente uno de los dos valores de los genes de los padres $\{x_i, y_i\}$.

j) Cruce aritmético (Michalewicz, 1996)

Al aplicar este tipo de cruce se obtienen dos descendientes de la forma $H_K = (h_1^K, h_2^K, \dots, h_N^K)$ K:1,2, donde cada gen h_i^K se genera haciendo:

$$h_i^1 = \lambda \cdot x_i + (1 - \lambda) y_i \quad (4.4)$$

$$h_i^2 = \lambda \cdot y_i + (1 - \lambda) x_i \quad (4.5)$$

con $\lambda \in [0, 1]$. Si se toma λ constante se denomina *cruce aritmético uniforme*, cuando se escoge λ variable se usa el llamado *cruce aritmético no-uniforme*. El caso especial en el que se toma $\lambda=0.5$ se conoce como cruce promedio (“average crossover”).

k) Cruce extendido lineal (Mühlenbein et al., 1993)

Se produce un descendiente obteniendo cada gen mediante la relación:

$$h_i = x_i + \alpha(y_i - x_i) \quad (4.6)$$

con $\alpha \in [-0.25, 1.25]$.

l) Cruce lineal (Wright, 1991)

Este tipo de cruce genera tres descendientes $H_K = (h_1^K, h_2^K, \dots, h_N^K)$ K:1,2,3 cuyos genes se construyen haciendo:

$$h_i^1 = \frac{1}{2} \cdot x_i + \frac{1}{2} y_i \quad (4.7)$$

$$h_i^2 = \frac{3}{2} \cdot x_i - \frac{1}{2} y_i \quad (4.8)$$

$$h_i^3 = -\frac{1}{2} \cdot x_i + \frac{3}{2} y_i \quad (4.9)$$

Luego, se escogen los mejores dos descendientes de este conjunto, aplicando de cierta manera un mecanismo de selección.

m) Cruce heurístico (Wright, 1991)

Esta técnica de cruce toma en cuenta el valor de la función de adaptación de los padres. Suponiendo que X es el parente con el mejor “fitness”, se genera un descendiente obteniendo sus genes como:

$$h_i = r(x_i - y_i) + x_i \quad (4.10)$$

donde r es un número aleatorio en el intervalo [0,1].

Estos son algunos de los operadores de cruce encontrados en la literatura. Además, se tienen el cruce BLX- α (Eshelman et al, 1993), el cruce lineal BGA (Schlierkamp-Voosen, 1994) y el cruce basado en conectivos difusos (Herrera et al, 2004) entre otros.

El siguiente paso, después de la generación de los descendientes, tiene que ver con la formación de la nueva población. Los autores proponen generar cuatro descendientes por cada

par de individuos pertenecientes al conjunto P'' constituido por la mitad de los individuos del conjunto de cromosomas seleccionados como padres (P'). Cada descendiente se obtiene mediante la aplicación de uno de los operadores antes mencionados. Los cuatro formarán parte de la nueva población, dos sustituyen a sus respectivos padres y los otros dos sustituyen otros dos individuos pertenecientes a $(P' - P'')$. De esta manera, aplicando adecuadamente los operadores se logra incluir en la población: dos individuos con propiedades de exploración, un individuo de la zona de explotación y otro de la zona de explotación relajada.

En la figura 4.3 (Herrera et al, 1998) se ilustra cómo actúa cada uno de estos operadores, mostrando el rango de acción de uno de los genes resultantes después del proceso de cruce el intervalo $[a_i, b_i]$.

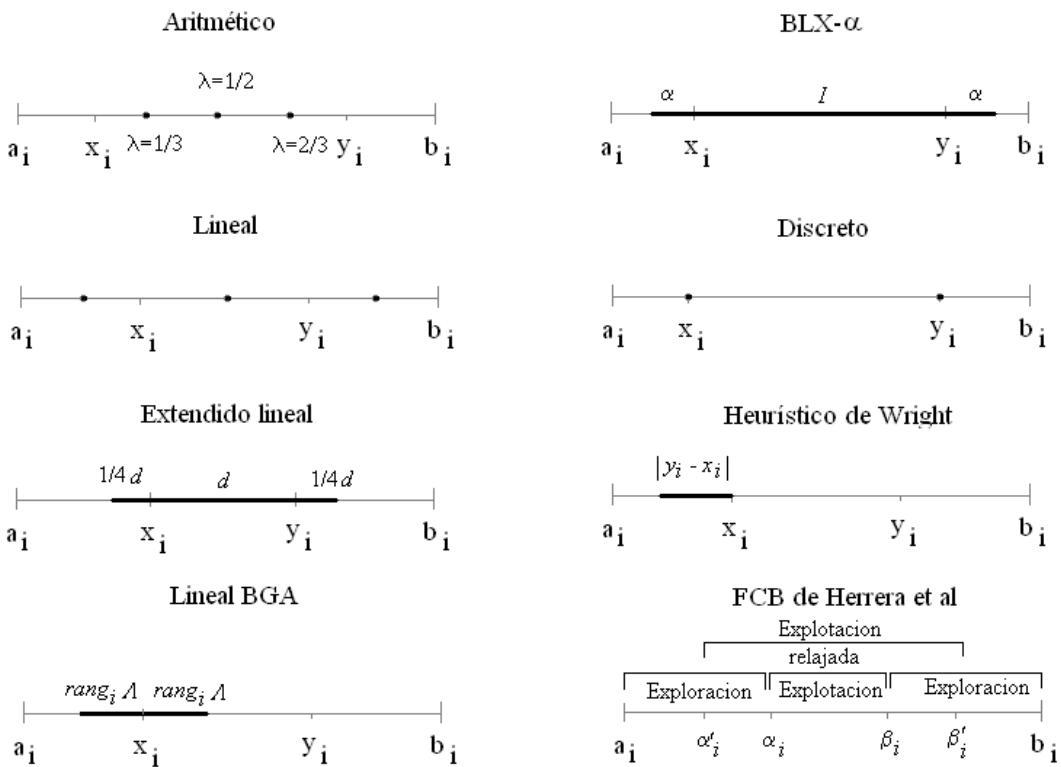


Figura 4.3. Rango de acción de los distintos operadores para un gen

La descripción que se acaba de presentar de los operadores de cruce es una breve reseña tomada de Herrera et al, (1998) quienes hacen una buena revisión de las técnicas para manejar algoritmos genéticos de codificación real. En Gen y Cheng (2000) se presenta una explicación

más detallada de algunas de estas técnicas, haciendo énfasis en los fundamentos matemáticos que las sustentan. En la literatura también se encuentra un buen número de trabajos que estudian el efecto del operador de cruce sobre el desempeño general del algoritmo genético, entre ellos se encuentran los de Deb (2002) y Herrera et al (2003).

2.3.3. Mutación

La mutación es un mecanismo necesario para mantener la diversidad de la población introduciendo la probabilidad de que ocurra un cambio inesperado y favorable en alguno de los genes. Este operador es útil para mejorar localmente las soluciones (Deb, 2000).

En general, el operador de cruce se considera el motor de la búsqueda que lleva a cabo el algoritmo genético y la mutación se toma como un operador de respaldo. Sin embargo, algunos trabajos donde se compara el efecto de ambos operadores sobre la búsqueda indican que en algunos casos el papel de la mutación puede ser tan importante como el del cruce (Gen y Cheng, 2000).

A continuación se describen algunos operadores de mutación comúnmente utilizados en la literatura (Herrera et al, 1998; Gen y Cheng, 2000). Suponiendo que H , $H = (h_1, h_2, \dots, h_N)$ es el individuo que va a ser sometido a mecanismo de mutación y h_i' es el gen resultante de la aplicación del operador.

a) *Mutación aleatoria (Michalewicz, 1996)*

En este caso el cromosoma H se convierte en $H = (h_1, \dots, h_i', \dots, h_N)$ y h_i' es un número aleatorio (uniforme) del intervalo $[a_i, b_i]$.

b) *Mutación no-uniforme (Michalewicz, 1996)*

Se muta un gen h_i del individuo de acuerdo a una de las siguientes opciones que dependen del parámetro aleatorio $\tau \in \{0,1\}$:

$$\begin{cases} h_i' = h_i + \Delta(t, b_i - h_i) & \text{si } \tau = 0 \\ h_i' = h_i - \Delta(t, h_i - a_i) & \text{si } \tau = 1 \end{cases} \quad (4.11)$$

La función $\Delta(t, y)$ se define como:

$$\Delta(t, y) = y \cdot r \left(1 - \frac{t}{g_{\max}}\right)^b \quad (4.12)$$

donde r es un número aleatorio en el intervalo $[0,1]$, g_{\max} es el número máximo de generaciones, b es un número escogido por el usuario que determina el grado de dependencia del número de iteraciones y t es la generación actual. Esta función permite explorar uniformemente en todo el espacio, inicialmente, y luego localmente, apuntando al logro de precisión y ajuste fino de las soluciones (Gen y Chen, 2000)

c) *Mutación direccional (Gen et al, 1996)*

Este método dirige la mutación en la dirección del gradiente, aunque también puede aplicarse de manera de evitar que los cromosomas se estanquen en una zona del espacio de búsqueda. El cromosoma mutado se calcula como:

$$H' = H + r \cdot \vec{d} \quad (4.13)$$

donde \vec{d} es la dirección aproximada del gradiente cuya componente i se determina haciendo:

$$d_i = \frac{f(h_1, \dots, h_i + \Delta h_i, \dots, h_N) - f(h_1, \dots, h_i, \dots, h_N)}{\Delta h_i} \quad (4.14)$$

tomando Δh_i como un valor real muy pequeño.

d) *Mutación gaussiana*

Esta técnica de mutación se deriva de las utilizadas en las estrategias evolutivas, para las cuales se definen individuos compuestos por las variables de decisión del problema a tratar y los parámetros del algoritmo de la forma $H = (H, \sigma)$ donde σ es el vector de las desviaciones estándar. Los descendientes mutados se generan como:

$$\sigma' = \sigma e^{N(0, \Delta \sigma)} \quad (4.15)$$

$$H' = H + N(0, \Delta\sigma') \quad (4.16)$$

donde $N(0, \Delta\sigma')$ es un vector de números aleatorios independientes (Gaussianos) con una media de cero y una desviación estándar σ .

2.4. TRATAMIENTO DE LAS RESTRICCIONES

Los algoritmos genéticos son poco hábiles en el manejo de restricciones no triviales. Durante el proceso de evolución en el espacio del genotipo se pueden generar soluciones ilegales o soluciones no factibles que afectan el desempeño del algoritmo genético, tal como se muestra en la figura 4.4 (Gen y Cheng, 2000). La *illegalidad* tiene que ver con la representación de los individuos, ya que se pueden producir cromosomas que al decodificarlos no representen una posible solución del problema (Gen y Cheng, 2000), lo cual se relaciona directamente con la técnica de codificación utilizada. En cambio, la no-factibilidad se refiere a soluciones que, decodificadas, no pertenecen a la región factible del espacio de búsqueda, es decir, soluciones que no cumplen todas las restricciones del problema de optimización.

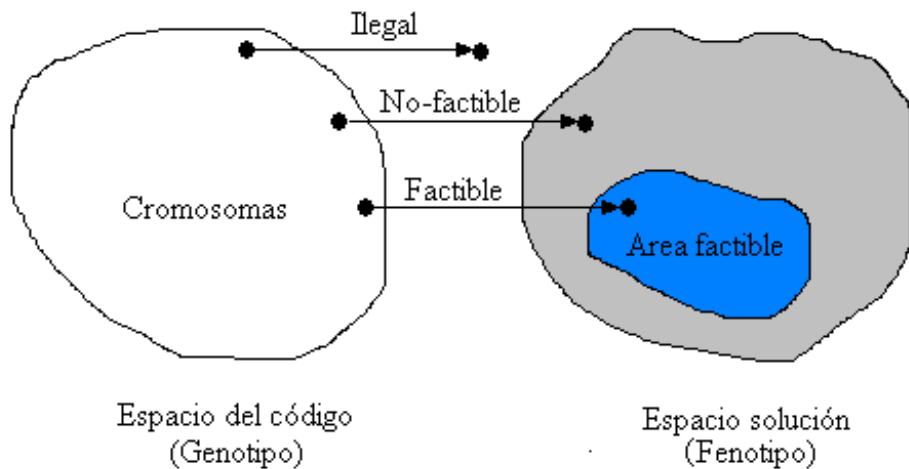


Figura 4.4 Espacio de soluciones factibles, no factibles e ilegales

En la literatura se presentan diversas estrategias para enfrentar los problemas de optimización con restricciones con un algoritmo genético de codificación real (Michalewicz, 1995; Summanwar et al, 2002; Coello, 2002 Michalewicz et al, 2006). Coello (2002) presenta

una revisión bastante completa de estas técnicas clasificándolas de acuerdo al siguiente esquema:

- a) Funciones de penalización*
- b) Técnicas de rechazo de los individuos no factibles (“Death penalty”)*
- c) Representaciones y operadores especiales*
- d) Algoritmos de reparación*
- e) Separación de objetivos y restricciones*
- f) Métodos híbridos*

A continuación se presenta una breve reseña de cada una de ellas, haciendo énfasis en los métodos de penalización que han sido los más utilizados para la aplicación específica que se trata en este trabajo.

a) Métodos de penalización

Las técnicas de penalización son probablemente las más utilizadas por los algoritmos genéticos en los problemas de optimización con restricciones (Deb, 2000). Esta técnica transforma el problema restringido en un problema sin restricciones penalizando las soluciones no factibles.

La idea básica de la técnica de penalización de restricciones proviene de los métodos de optimización convencionales. En los problemas de optimización clásicos se distinguen dos tipos de funciones de penalización: interior y exterior. Los métodos de penalización interior utilizan un factor de penalización que toma valores bajos en los puntos alejados de los límites establecidos por las restricciones y tiende a infinito a medida que los puntos se acercan a la frontera de la región factible. Si la búsqueda inicia con un punto factible, los candidatos a solución generados a partir de éste se mantendrán dentro de la región factible, ya que la penalización hace que los límites de las restricciones actúen como barreras durante el proceso de optimización (Coello, 2002). En el caso de los métodos de penalización exterior, la

búsqueda inicia con un punto no factible, desde el cual se avanza, a lo largo de proceso de optimización, hacia la región factible. Los métodos de penalización exterior son los preferidos para los algoritmos genéticos porque no se requiere una solución inicial factible (Coello, 2002).

La forma general de la función de penalización exterior es (Deb, 1995; Reklaitis et al., 1983; Summanwar et al, 2002)

$$F(x) = f(x) + \left[\sum_{i=1}^{pd} r_i \cdot G_i + \sum_{j=1}^{pi} c_j \cdot L_j \right] \quad (4.17)$$

con:

$$G_i = \max[0, g_i(x)]^\beta \quad (4.18)$$

$$L_j = |h_j(x)|^\gamma \quad (4.19)$$

r_i y c_j son constantes positivas escogidas como factores de penalización, pd y pi son el número de restricciones de desigualdad y de igualdad respectivamente, g_i y h_i corresponden a las restricciones de desigualdad y de igualdad respectivamente, y α y γ se toman normalmente como 1 o 2.

En el diseño de la función de penalización se debe tomar en cuenta la regla de la *mínima penalidad* (“minimum penalty rule”), que establece que la penalización debe mantenerse tan baja como sea posible, justo debajo del límite en el cual las soluciones óptimas son no factibles. Respecto a esto puede decirse lo siguiente:

- Imponiendo una penalización muy fuerte, se disminuye la capacidad de exploración, corriendo el riesgo de que la población converja prematuramente al primer individuo no óptimo que cumpla las restricciones desplazando a algunos individuos no factibles que pueden estar cerca del óptimo y que al reproducirse podrían acercarse mucho más

- Imponiendo penalización moderada se perderá mucho tiempo explorando la región no factible, ya que el término correspondiente a la penalización será despreciable frente a la función objetivo y, probablemente, la búsqueda convergerá en una solución no factible.

Existen diversas técnicas de penalización para trabajar con algoritmos genéticos, entre las que se encuentran las funciones de penalización estática, penalización dinámica, penalización basada en “*simulated annealing*”, penalización adaptativa y penalización co-evolucionaria. Estas se describen en detalle en Michalewics (1995), Michalewicz et al, (1996), Coello (2002), aquí se presentan únicamente las más utilizadas en la literatura.

En las técnicas de *penalización estática*, los coeficientes de penalización no dependen del número actual de generaciones, permaneciendo constantes durante todo el proceso de evolución. Al usar este método la población inicial puede generarse de manera aleatoria por lo cual estará constituida por individuos factibles y no factibles.

Una forma típica de esta función de penalización es:

$$F(x) = f(x) + \left[\sum_{i=1}^m r_{k,i} \cdot \max[0, g_i(x)]^2 \right] \quad (4.20)$$

donde se sigue la misma notación de la ecuación 4.21, con m el número total de restricciones. Las restricciones de igualdad se transforman en restricciones de desigualdad haciendo $|h_i(x)| \leq \varepsilon$, donde ε es un valor muy pequeño (Coello, 2002). Los factores de penalización se escogen de manera que la violación de la restricción sea del mismo orden de magnitud de la función objetivo (Summanwar et al, 2002).

También es posible normalizar las restricciones aplicando la siguiente transformación Deb (2001):

$$\underline{g}_i \leq b_i \rightarrow g_i \equiv \frac{g_i}{b_i} - 1 \geq 0 \quad (4.21)$$

Para este caso particular la notación \underline{g}_i para las restricciones sin normalizar y g_i para las restricciones normalizadas. Así, se puede usar un factor de penalización único (Deb, 1995) obteniendo una función de la forma:

$$F(x) = f(x) + R \cdot \sum_{i=1}^m \max[0, g_i(x)]^2 \quad (4.22)$$

Homaifar et al (1994) proponen una estrategia de penalización en la cual se determinan los valores adecuados de los factores de penalización (R_{ij}) de acuerdo a los niveles de violación definidos para cada restricción, de manera que a mayor nivel de violación se tienen valores más altos del coeficiente de penalización. Los individuos se evalúan utilizando la siguiente fórmula

$$F(x) = f(x) + \sum_{i=1}^m R_{ij} \cdot f_j^2(x) \quad (4.23)$$

donde $f(x)$ es la función objetivo, f_j es la desviación de la restricción j y R_{ij} es el factor de penalización para el i -esimo nivel de violación de la restricción j . La debilidad de este método es la cantidad de parámetros que requiere en función del número de restricciones y de los niveles de violación asignados.

Existen otros métodos para tratar restricciones basados en funciones de penalización estáticas, sin embargo, esta técnica presenta algunos inconvenientes principalmente debido a la sensibilidad del algoritmo a variaciones en el factor de penalización. Si se toma un valor muy pequeño puede que se obtengan soluciones no factibles, mientras que si se toma un valor muy alto puede ejercer demasiada presión selectiva. Se deben realizar varios ensayos hasta encontrar un valor adecuado para los pesos asociados a la penalización lo cual requiere de un gran esfuerzo experimental (Deb, 2001).

Como alternativa se han propuesto métodos de penalización dinámicos, en los cuales se toma información de la generación actual para el cálculo de los factores de penalización. Entre éstos cabe destacar el método de Michalewicz y Attia (1994) basado en el “*Simulated annealing*”.

Algunos autores consideran que los métodos de penalización dinámica funcionan mejor que la penalización estática, sin embargo, ambos presentan los mismos inconvenientes de sensibilidad a los factores de penalización. Además, si los parámetros ajustables requeridos por el método no tienen valores adecuados, el algoritmo convergerá a una solución sub-óptima (Coello, 2002).

b) Métodos de rechazo

El rechazo consiste en eliminar los individuos que no satisfacen las restricciones. Se utiliza frecuentemente en muchas de las técnicas evolutivas porque de cierto modo es computacionalmente eficiente al no ser necesario evaluar las soluciones no factibles ni determinar el grado de desviación de respecto a las restricciones. Sin embargo, cada vez que se rechaza un individuo, se inicia un proceso iterativo para generar uno nuevo que satisfaga las restricciones, lo cual puede ser un proceso arduo si se trata de un problema con restricciones complejas.

Por esto, la efectividad de esta estrategia se limita a aplicaciones donde la región factible es convexa y ocupa una parte razonable de la totalidad del espacio de búsqueda. En muchos problemas la población inicial cuenta con un número muy reducido de individuos que cumplen las restricciones, siendo más conveniente en estos casos mejorar las soluciones existentes en lugar de eliminarlas (Michalewicz et al, 1996).

c) Representaciones y operadores especiales

Diversos autores han ideado representaciones especiales de los cromosomas junto con operadores adecuados para enfrentar los problemas con restricciones. La idea es utilizar una representación particular que simplifique el espacio de búsqueda para luego aplicar operadores especializados que preserven la factibilidad de las posibles soluciones. Esta estrategia se utiliza comúnmente en problemas donde es muy difícil encontrar al menos una solución factible (Coello, 2002).

Entre estas técnicas es interesante mencionar el algoritmo GENOCOP desarrollado por Michalewicz y Xiao (1995) que ha sido evaluado en distintos problemas de prueba produciendo buenas soluciones. Este método parte de una población factible y supone

restricciones lineales y una región factible convexa, con lo que simplifica el espacio de búsqueda. Esta es su principal debilidad, por el esfuerzo de cómputo que involucra encontrar una población factible y la limitación impuesta por el requerimiento de linealidad en las restricciones.

También se encuentran métodos que se enfocan en ubicar la frontera del espacio de búsqueda factible, lo cual tiene bastante sentido si se toma en cuenta que en muchos de los problemas de optimización no lineal al menos una de las restricciones se encuentra activa en presencia del óptimo global. Esta técnica hace uso de penalizaciones adaptativas o mecanismos similares para cruzar el borde la región factible (Coello, 2002). Se requiere un procedimiento de inicialización para encontrar una población inicial factible y unos operadores especiales, tales que el cruce permita encontrar todos los puntos entre los padres y la mutación actúe de manera que pequeños cambios en los individuos se reflejen de manera proporcional en la función de costo (Coello, 2002).

Una estrategia interesante es el uso de decodificadores (“*Decoders*”), los cuales consisten en cromosomas que indican cómo construir una solución factible. Cada “*Decoder*” impone una relación de transformación T entre la solución factible y la solución decodificada. Para aplicar esta técnica se deben cumplir algunas condiciones: i) por cada solución factible “ s ” debe existir una solución decodificada “ d ” ii) cada solución decodificada “ d ” corresponde a una solución factible “ s ” iii) todas las soluciones factibles deben estar representadas por el mismo número de soluciones decodificadas “ d ” iv) la transformación T debe ser computacionalmente rápida v) los pequeños cambios en las soluciones decodificadas deben reflejarse como pequeños cambios en la solución como tal.

El uso de decodificadores, de acuerdo a los resultados presentados en Koziel y Michalewicz (1999), es una de las estrategias que aporta mejores resultados en el tratamiento de las restricciones. Sin embargo, presenta como principales desventajas el mayor esfuerzo computacional que requiere su aplicación y que en espacios de búsqueda no-convexos es muy difícil cumplir la condición (v) (Coello, 2002).

d) Métodos de reparación

Los algoritmos de reparación se han utilizado eficientemente en problemas de optimización combinatoria donde es relativamente fácil “reparar” un individuo que no es factible. Este método aproxima un cromosoma no-factible a la solución válida más cercana o lo reconstruye para obtener una solución válida que puede ser sustancialmente distinta. Las versiones reparadas del cromosoma se utilizan únicamente en la evaluación de la función de adaptación; solo en algunos casos, dependiendo de un parámetro de probabilidad, pueden ser introducidas en la población sustituyendo al individuo original (Coello, 2002; Summanwar et al, 2002).

Esta estrategia se ha incorporado en el software GENOCOP III para optimización de problemas no lineales (Michalewicz y Nazhiyath, 1995). Sin embargo, se debe tomar en cuenta que el proceso de transformación de una solución no-factible a una factible no siempre es viable. En algunos casos la reparación de un cromosoma puede ser tan compleja como la resolución del problema original. El algoritmo de reparación se debe diseñar para cada problema particular, lo cual es una debilidad de este método, pues no existe una metodología general para diseñar una estrategia de reparación (Michalewicz et al, 1996).

e) Separación de objetivos y restricciones

En la literatura se encuentran algunos métodos de tratamiento de las restricciones que plantean separar los objetivos de la optimización y las restricciones, entre estos cabe destacar las técnicas que consideran la *superioridad de los puntos factibles* y las técnicas de *optimización multiobjetivo*.

Entre las técnicas basadas en la suposición de superioridad de los puntos factibles se encuentran las propuestas de Deb (2000), Powel y Skolnick (1993) y el algoritmo CONGA (Constraint based Numeric Genetic Algorithm) de Hinterding y Michalewicz (1998).

Powel y Skolnick (1993) proponen la siguiente función para evaluar los individuos:

$$F(x) = \begin{cases} f(x) & \text{si } es \text{ factible} \\ 1 + R \left(\sum_{i=1}^{pd} \max[0, g_i(x)] + \sum_{i=1}^{pi} |h_i(x)| \right) & \text{Sino} \end{cases} \quad (4.24)$$

donde R es una constante definida por el usuario. La $f(x)$ correspondiente a las soluciones factibles está escalada en el intervalo $(-\infty, 1)$ mientras que la evaluación de las soluciones no factibles arroja valores en el intervalo $(1, \infty)$ en el cual están escaladas $g_i(x), h_i(x)$. Esta estrategia de evaluación se combina con un mecanismo de selección por jerarquización lineal, donde la presión selectiva es baja en las primeras generaciones y se hace más fuerte a medida que el algoritmo converge a una solución. Este mecanismo de selección hace que el método sea poco sensible a los valores de la constante R , evitando los problemas típicos asociados a los factores de penalización.

Deb (2000) propone un método similar, donde se utiliza una función de la forma:

$$F(x) = \begin{cases} f(x) & \text{if } g_i(x) \leq 0 \quad \forall i = 1, 2, \dots, pd, \text{ y } |h_i(x)| \leq \delta \\ f_{\max} + \sum_{i=1}^{pd} \max[0, g_i(x)] + \sum_{i=1}^{pi} |h_i(x)| \end{cases} \quad (4.25)$$

en la que f_{\max} es el valor de la función objetivo del peor individuo factible de la población y la principal diferencia respecto al planteamiento de Powel y Skolnick (1993) es que no es necesario el factor de penalización R . Esta función se combina con un esquema de selección por torneo binario que sigue los siguientes criterios:

- Si ambas soluciones son factibles, se escoge la que posea el mejor valor de la función objetivo.
- Entre una solución factible y una no-factible se escoge la solución factible.
- Si ambas soluciones son no-factibles se toma la que presente la menor desviación en las restricciones

En el torneo se aplica, además, un *operador de nicho*, que solo permite comparar soluciones que se encuentran dentro de un límite dentro de su distancia Euclídea. De esta manera, si se encuentran soluciones alejadas entre sí dentro del espacio factible se puede asegurar su permanencia en las siguientes generaciones, manteniendo así la diversidad de la población.

En el caso de los *métodos basados en optimización multiobjetivo* se plantea redefinir el problema de optimización de un objetivo con restricciones, convirtiéndolo en uno de optimización multiobjetivo con $m+1$ objetivos, siendo m el número total de restricciones (Coello, 2002). Más adelante, se tratará con más detalle el tema de los algoritmos genéticos multiobjetivo, en esta sección solo se describen las aplicaciones de estas técnicas para el manejo de las restricciones.

Entre estos métodos se encuentran el COMOGA propuesto por Surry and Radcliffe (1997) el cual utiliza un ordenamiento (“*ranking*”) de Pareto para tratar las restricciones. Este método resuelve el problema utilizando una aproximación dual, en la cual una parte de la población es seleccionada en base el ordenamiento de Pareto y otra parte se selecciona de acuerdo al valor de la función objetivo (Coello, 2002; Summanwar et al, 2002). Summanwar et al (2002) presentan una versión modificada del COMOGA, donde se utiliza codificación real, selección basada en el rango de los individuos, mutación no-uniforme y se mantienen dos listas de clasificación de los individuos basadas en el ordenamiento de Pareto y en la función objetivo, para evitar eliminar individuos de manera inapropiada. En su trabajo se compara el desempeño del COMOGA y el método de Deb (2000) en la resolución de problemas de optimización de distintos niveles de dificultad, encontrando resultados favorables aunque se evidencia el mayor esfuerzo computacional requerido por el COMOGA.

El VEGA (Vector Evaluated Genetic Algorithm) es una técnica propuesta por Schaffer (1984) donde se divide la población en m subpoblaciones que son evaluadas de acuerdo a uno de los objetivos de la optimización. El mecanismo de selección se lleva a cabo en función de cada objetivo pero el cruce se realiza en las fronteras de las subpoblaciones, además, se aplican heurísticas para evitar que la población converja prematuramente. Sin embargo, algunos estudios señalan que el efecto de aplicar este método es el mismo que el de realizar una suma ponderada de los objetivos (Michalewicz, 1995).

Otras técnicas interesantes, basadas en optimización multiobjetivo, han sido propuestas por Jimenez y Verdegay (1999) y Coello (2000).

Los *métodos híbridos* combinan técnicas comúnmente utilizadas para tratar restricciones en los algoritmos genéticos con otro tipo de técnica, que frecuentemente es un método de

optimización numérica. Entre ellos se encuentran: los métodos con multiplicadores de Lagrange (Adeli y Cheng, 1994; Kim y Myung, 1997) que integran la función típica de penalización de las restricciones con el método primal-dual de optimización, los algoritmos culturales de Reynolds (1994), los sistemas inmunes (Forrest y Perelson, 1991), los métodos de colonias de hormigas (Dorigo et al, 1991; Dorigo et al, 1996) y los métodos que introducen lógica difusa como el propuesto por Van Lee (1995).

3. EJEMPLO DE APLICACIÓN DEL ALGORITMO GENETICO AL DISEÑO INTEGRADO DE UN SISTEMA DE REACTORES

Se resolvió el problema de síntesis y diseño integrado del sistema de reactores químicos planteado en el capítulo III como ejemplo de aplicación de los algoritmos genéticos. Para analizar el efecto de las restricciones de controlabilidad sobre el diseño del proceso, se resolvió en primer lugar el problema de síntesis de procesos sin tomar en cuenta las restricciones de controlabilidad, luego, se añadió la restricción sobre los autovalores de la matriz de estados y finalmente, se agregó la restricción sobre el parámetro Q_{\max}/Q . A continuación se describe brevemente la implementación del algoritmo genético al resolver este problema, la evaluación de desempeño del algoritmo y finalmente, se presentan los resultados del problema de síntesis y diseño integrado.

a) Implementación del algoritmo genético

El algoritmo genético se desarrolla en Matlab ®. En una etapa de inicialización se genera una población aleatoria de individuos que representan posibles soluciones al problema, esto se hace cuidando que la población tenga la misma cantidad de individuos para cada alternativa estructural. Cada individuo es un vector que contiene las variables a optimizar, que en este caso son los valores normalizados del volumen de reacción (V), la temperatura de reacción (T), el flujo de refrigerante alimentado al primer reactor (F_{j1}) y las variables enteras y_1 y y_2 , que representan la existencia del segundo y tercer reactor. Los individuos son de la forma:

$$[V \ T \ F_{j1} \ y_1 \ y_2]$$

Para cada cromosoma se determina la función de costos (función objetivo) y se evalúan las restricciones, aplicando la estrategia de penalización dada por la ecuaciones 4.24 en el caso

que no se cumplan. La población evoluciona, a lo largo de varias generaciones, a través de los operadores de selección, cruce aritmético, mutación aleatoria y sustitución hasta cumplir con el número de iteraciones máximo establecido como criterio de parada. El mejor individuo de la población final se reporta como solución del problema.

Se establece un conjunto de valores base de los parámetros del algoritmo genético para la solución del problema: el tamaño de la población es de 30 individuos. La selección se realiza en un 35 % por elitismo y 65 % por torneo. La tasa de cruce es del 85 % y la tasa de mutación es del 10 %. El operador de cruce utilizado es el cruce aritmético y la renovación de la población esta dada por una tasa de sustituidos del 50 %.

El desempeño del algoritmo genético, en cuanto a calidad de la solución obtenida y eficiencia, está influenciado por los valores de los parámetros del algoritmo. Para evaluar su sensibilidad a la variación de los parámetros, se realizan modificaciones en algunos de los parámetros, como el número de individuos que componen la población, el porcentaje de individuos sustituidos y el número de iteraciones máximo.

La evaluación del desempeño del algoritmo genético al resolver este problema, se realiza para la síntesis económica del proceso, limitado a un sistema de uno o dos reactores. Así, se reduce el número de restricciones del problema para evitar otro tipo de efectos.

b) Evaluación del desempeño del algoritmo genético ante variaciones en sus parámetros

Se realizan variaciones en cada uno de los parámetros a evaluar mientras los demás parámetros mantienen en sus valores base. Se realizan cinco pruebas de 1000 iteraciones en cada caso y se calcula el promedio de la función objetivo para medir la calidad de los resultados.

En la tabla 4.1 se muestran los resultados obtenidos al variar el número máximo de iteraciones. Se observa como los resultados mejoran a medida que aumenta el número de iteraciones. Sin embargo, al superar las 1000 generaciones la variación en la solución no es significativa lo que indica que ese número de generaciones es suficiente para alcanzar una solución satisfactoria.

Tabla 4.1. Influencia del número de generaciones máximo sobre los resultados obtenidos al resolver el problema de síntesis de reactores químicos

Nro de Iteraciones	500	1000	1500	2000
F. Objetivo (\$)	129366.0	95762.4	95449.4	90586.4
Nro. de unidades	2	2	2	2

En la tabla 4.2 se presentan los resultados del problema de síntesis del sistema de reactores químicos realizando variaciones en el número de individuos de la población, el porcentaje de seleccionados por elitismo y el porcentaje de sustituidos manteniendo el resto de los parámetros en los valores base (*).

Tabla 4.2. Influencia de los parámetros del sobre el desempeño del algoritmo genético al resolver el problema de síntesis de reactores químicos.

Parámetro	Variación		
Numero de Individuos	20 individuos	30 individuos	40 individuos
F. Objetivo Promedio (\$)	102696,60	95762.40	95010.75
Nro. de unidades	2	2	2
% Elitismo en la selección	35% *	50%	65%
F. Objetivo Promedio (\$)	98390,00	92233,60	96748,67
Nro. de unidades	2	2	2
% Sustituidos	35%	50% *	65%
F. Objetivo Promedio (\$)	99704,80	98390,00	95449,40
Nro. de unidades	2	2	2

En los resultados presentados en la tabla 4.2 se observa que el porcentaje de seleccionados por elitismo óptimo para este problema es del 50 %. El porcentaje de sustituídos

produce una ligera mejora en los resultados obtenidos. En cuanto al número de individuos, se observa que reducir la población a una de 20 individuos empeora el desempeño del algoritmo, pero incrementar el número de individuos a 40 no representa una mejora importante en la calidad de la solución.

A continuación se presentan los resultados del problema de síntesis y diseño integrado del sistema compuesto por uno a tres reactores químicos.

c) Resultados del problema de síntesis y diseño integrado

En la tabla 4.3 se presentan los resultados de la síntesis del sistema de reactores con restricciones distintas de controlabilidad, entre las que se encuentran el caso económico, el diseño integrado a lazo abierto con una restricción sobre los autovalores de lazo abierto y luego, el diseño considerando además restricciones sobre la magnitud la relación Q_{\max}/Q .

Se puede observar que el resultado del problema de síntesis de reactores químicos enfocado únicamente en la minimización de los costos es el sistema compuesto por tres reactores operando a una temperatura cercana a la máxima. Esto quiere decir, que en un sistema de tres reactores se puede alcanzar la conversión especificada, con un volumen total mucho menor al requerido para realizar el mismo proceso en un reactor único. Sin embargo, al incluir las restricciones de controlabilidad, el resultado es un sistema de un reactor, cuyo volumen aumenta a medida que aumenta la severidad de las restricciones impuestas.

Para este sistema, Luyben (1993), menciona que la temperatura óptima de reacción es la más alta posible pues en estas condiciones se requiere de un menor volumen de reacción, lo cual se traduce en menores costos. Sin embargo, al considerar la controlabilidad del sistema, se encuentra que reducir el volumen del reactor involucra la reducción del área de transferencia de calor lo cual dificulta la operación del sistema de refrigeración.

En los resultados presentados se puede observar que la parte real de los autovalores es positiva en el sistema de tres reactores (el más económico) indicando inestabilidad a lazo abierto. Para este sistema, la relación Q_{\max}/Q es muy pequeña, lo cual indica poca flexibilidad en el proceso. Cuando se introduce la restricción que garantiza autovalores de lazo abierto negativos, el volumen de reacción aumenta al igual que la relación Q_{\max}/Q . Finalmente se

encuentra que a medida que se exige una mayor relación Q_{\max}/Q el volumen aumenta, afectando de la misma forma los costos de construcción. Todo esto refleja un fuerte compromiso entre los criterios económicos y los criterios de controlabilidad para un proceso de reacción química, el cual debe manejarse adecuadamente para obtener un diseño satisfactorio.

Tabla 4.3. Resultados del problema de síntesis y diseño integrado del sistema de uno a tres reactores químicos

	Síntesis económica	Autovalores Negativos	Autovalores Negativos $Q_{\max}/Q > 2,5$	Autovalores Negativos $Q_{\max}/Q > 3$
<i>F. objetivo (\$)</i>	70910	120000	145720	412370
<i>Nro. de unidades</i>	3	1	1	1
<i>Volumen</i>	36.35	494.25	674.75	3586.5
<i>Temperatura</i>	198.87	193.03	184.34	141.39
z_1	0.37	0.05	0.05	0.05
z_2	0.14	-	-	-
z_3	0.05	-	-	-
Q_{\max}/Q_1	1.27	2.4	2.51	3
Q_{\max}/Q_2	1.94	-	-	-
Q_{\max}/Q_3	0	-	-	-
<i>Parte real autoval</i>	29.69	-0.5071	-0.3303	-0.0379
	-28.2229	-13.7541	-13.60	-12.36
	-1.6869	-0.5071	-0.3303	-0.0379

4. CONCLUSION

En este capítulo se describen los algoritmos genéticos, que constituyen un método estocástico de optimización muy poderoso y que se aplican como alternativa a los métodos tradicionales de optimización, para la resolución de los problemas de *diseño integrado*. Se proponen algoritmos basados en codificación real que son los más adecuados cuando se trabaja con variables y restricciones de procesos físicos. Se revisa la literatura sobre aplicaciones similares haciendo hincapié en los distintos operadores genéticos y las alternativas de tratamiento de las restricciones.

Finalmente, se presenta el módulo informático de algoritmos genéticos desarrollado para este trabajo, con los alentadores resultados de una primera aplicación a la síntesis y diseño del sistema de reactores.

En los capítulos siguientes se exponen los resultados de otros dos casos de diseño integrado de procesos más complejos. Comparaciones de los algoritmos genéticos con los resultados obtenidos con métodos de optimización tradicionales y/o con otros métodos estocásticos se resumen también en el capítulo VI.

CAPÍTULO V

DISEÑO INTEGRADO DE UNA TORRE SULFITADORA DE CAÑA DE AZÚCAR

1. INTRODUCCIÓN

En este capítulo se lleva a cabo *el diseño integrado de una torre de sulfitación* perteneciente a la etapa de clarificación de jugo de caña en un ingenio azucarero. Después de una breve descripción del proceso, se retoma la deducción del modelo semifísico de base fenomenológica presentada en los trabajos de Prada et al. (1984) y Alvarez et al. (2009), para realizar el estudio del comportamiento dinámico y la controlabilidad de estado del sistema. Finalmente se presenta la formulación y los resultados del diseño integrado de la sulfitadora con controladores PI y MPC.

2. DESCRIPCIÓN DE LA TORRE SULFITADORA

La etapa de clarificación en un ingenio azucarero es la encargada de retirar la mayor cantidad de impurezas en el jugo de caña ya filtrado, previo a los procesos de evaporación y posterior cristalización del azúcar. Aunque existen varias configuraciones posibles de proceso para preparar el jugo entre la molienda y la entrada al clarificador, todas buscan la mayor reducción de impurezas minimizando la pérdida de materia prima (sacarosa). Los clarificadores separan mediante decantación sustancias indeseadas que floculan por la acción de tensoactivos, adicionados al jugo bajo condiciones de pH y temperatura adecuadas. Los sólidos no azúcares van al fondo del tanque clarificador en forma de lodo (cachaza), mientras que el jugo claro sale por la parte superior del tanque y va directamente al proceso de evaporación. En tal sentido, el control de pH resulta crítico en la clarificación para evitar la inversión de sacarosa, principal causante de la pérdida de azúcar al final del proceso. La sacarosa, un disacárido compuesto por una glucosa y una fructosa, se invierte para formar

(+)-D-glucosa y (-)-D-fructosa, ambas sustancias con menor tendencia a cristalizar y endurecer. Las condiciones que favorecen tal inversión son el calentamiento en medio ácido del jugo y la presencia de la enzima invertasa. La invertasa se activa entre 25 y 55°C con valores de pH entre 4.5 y 5.0.

Un proceso típico de clarificación es como sigue: al jugo que proviene del filtrado con un pH ácido entre 5.1 y 5.6 unidades, se adiciona lechada de cal en un primer tanque, con el fin de aumentar el pH hasta valores entre 6.5 y 7.0. Tal lechada puede ser solo cal en agua o cal con meladura (llamada sacarato). Se busca minimizar pérdidas de sacarosa, además de lograr que la cal adicionada ayude a precipitar impurezas y aumente la capacidad del coagulante. El paso siguiente ocurre en el tanque (o torre) sulfidador(a). Allí se adiciona, en forma de finas burbujas, una corriente de humos (SO_2 , N_2 , O_2) proveniente de la quema de azufre sólido en un horno. El jugo sufre un aumento de temperatura de hasta 8°C por el contacto con los humos. El dióxido de azufre, además de servir como agente antiséptico, transforma compuestos que dan coloración oscura al jugo en compuestos incoloros. Este efecto se mantiene incluso durante los calentamientos posteriores. La adición del dióxido de azufre entrega un jugo con pH entre 4.3 y 4.5, justo en el límite inferior del intervalo en el cual se produce la pérdida de sacarosa. Ocurre después una segunda alcalinización con la adición de cal, que se espera arrastre partículas pequeñas y torne insolubles algunos contaminantes que están originalmente solubilizados en el jugo. Luego de esta etapa, se acostumbra calentar el jugo hasta 105°C, para facilitar la formación de flóculos de alto peso molecular que puedan ser decantados por la gravedad. Finalmente, el jugo caliente llega al tanque clarificador donde se mezcla con un floculante aniónico, con dosificación cercana a las 5ppm. De este modo, se logra separar por decantación los sólidos o cachaza, que van al fondo, del jugo claro que sale por la parte superior. Un diagrama de flujo de proceso general se ve en la figura 5.1.

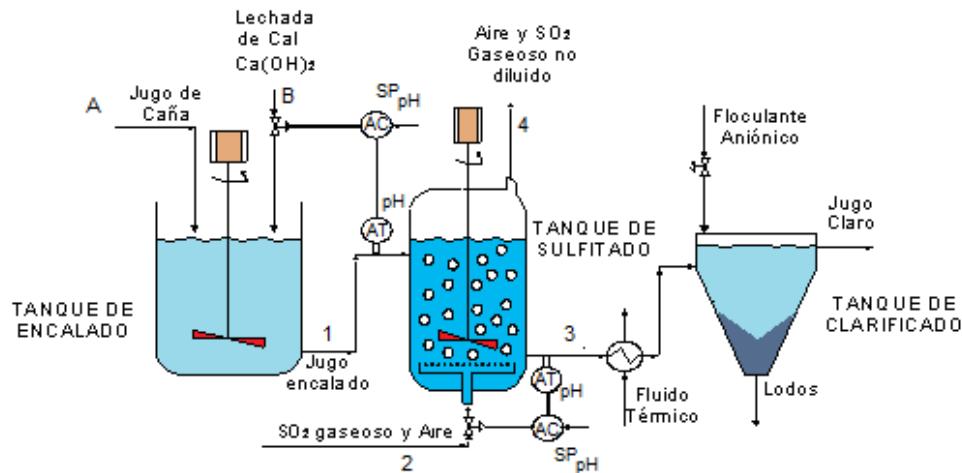


Figura 5.1. Diagrama de Flujo del Proceso general para el proceso de clarificación

Todo el proceso de clarificación se ve afectado por varias perturbaciones, entre las que cabe mencionar: cambios en la materia prima (tipo y calidad de la caña), preparación de la cal o sacarato, flujos de jugo cambiantes por alteraciones en la operación del molino, flujos de entrada de corrientes de recirculación del proceso (jugo filtrado, aguas dulces, etc.).

2.1 MODELO MATEMÁTICO DE LA TORRE SULFITADORA

Se trata de un modelo semi-físico de base fenomenológica, puesto que toma su estructura de los balances de materia de las sustancias presentes en el proceso, pero adiciona formulaciones empíricas para varios de sus parámetros. El modelo se basa en las especies químicas SO_3^{2-} y Ca^{2+} , se estudian por separado los fenómenos de transferencia del gas al líquido y las reacciones químicas que tienen lugar, procediendo de modo similar al modelo presentado en Vega (1987) y en Prada et al. (1984), pero ampliando la explicación al efecto total del encalado previo e incluyendo una justificación desde la teoría de invariantes de reacción (Asbjørnsen, 1972; Fjeld et al. 1974; Gustaffson, 1982).

De acuerdo con Alvarez et al. (2009), en la torre sulfitadora se pueden tomar al menos dos subsistemas o “Sistemas de Proceso”, uno asociado al jugo que desciende y otro asociado al gas que asciende (ver figura 5.2).

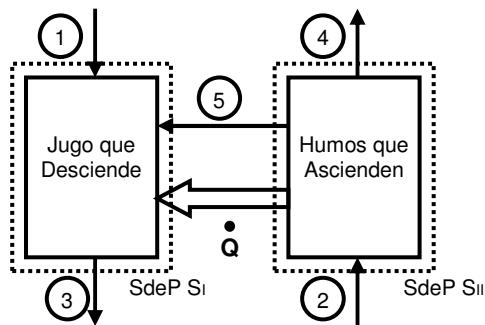


Figura 5.2. Diagrama de bloques mostrando los Sistemas de Proceso definidos y sus interacciones

Las reacciones químicas que se dan en el proceso de encalado y sulfitación son:



donde se pueden definir los siguientes invariantes de reacción, equivalentes a estados del proceso:

$$X = [\text{Ca(OH)}^+] + [\text{HSO}_3^-] = [\text{OH}^-] - [\text{H}^+] \quad (5.5)$$

$$C = [\text{H}_2\text{SO}_3] + [\text{HSO}_3^-] \quad (5.6)$$

siendo X el invariante correspondiente al balance de carga electrónica de la solución y C el balance del anión último, relacionado con la concentración de H_2SO_3 en la solución.

El fenómeno de transferencia de masa del gas al líquido se estudia en Vega (1987). En la superficie de contacto gas-agua tiene lugar la reacción química irreversible presentada en la ecuación (5.1). Su constante cinética es muy alta (Beg et al., 1991), por lo tanto, se puede considerar que todo el SO_2 que se disuelve en el jugo reacciona con el agua para formar ácido sulfuroso. En estas condiciones, es necesario formular la transferencia de masa entre el gas

(SO₂) que asciende en la torre y el jugo (solución que baja en la torre). Esta transferencia puede ser vista como un flujo de SO₂ que pasa del gas a la solución.

Los flujos se identifican de acuerdo a la numeración de corrientes de la figura 5.2, donde F_1 y F_2 son el flujo de jugo de caña y el flujo de gas (compuesto por aire y SO₂) alimentados a la torre, F_3 es el flujo de solución tratada, F_4 es el flujo de aire y SO₂ no disuelto que sale de la torre y F_5 es el flujo de SO₂ transferido del gas a la solución. Se realiza un balance de masa, primero total y luego por componente (para los dos invariantes) en el Sistema de Proceso seleccionado:

$$\frac{dM_{Jugo}}{dt} = F_1 \cdot \rho_1 + F_5 \cdot \rho_5 - F_3 \cdot \rho_3 \quad (5.7)$$

$$\frac{dX}{dt} = \frac{1}{V_{Jugo}} [F_1 \cdot [Ca(OH)^+]_l + \alpha \cdot F_5 - F_3 \cdot X] \quad (5.8)$$

$$\frac{dC}{dt} = \frac{1}{V_{Jugo}} [F_5 - F_3 \cdot C] \quad (5.9)$$

Al estudiar la reacción se supone que para cada burbuja:

- La velocidad de reacción es proporcional a la superficie de contacto y a la concentración de gas en su interior.
- El dióxido de azufre se comporta como un gas perfecto (ideal).

Algunos estudios realizados por Prada y Vega (1981), con datos de la fábrica Onésimo Redondo de Valladolid (España), justifican la aproximación del proceso a uno de primer orden:

$$F_5(t) + \tau(T) \cdot \frac{dF_5(t)}{dt} = K(T) \cdot F_2(t) \quad (5.10)$$

Aquí, la constante de tiempo (τ) y la ganancia (K) varian con la temperatura (T). Sin embargo, experimentalmente puede comprobarse que dichas variaciones no son significativas. Una

estimación adecuada de los valores que pueden usarse para este caso particular es $K=0.95$ y $\tau = 1.7\text{s}$ (Vega, 1987).

Alvarez et al. (2009) complementan esta relación recurriendo al formalismo típico para transferencias: fuerza impulsora, resistencia y flujo neto, común a las leyes de gradiente (Fradkov et al., 1999). Así, la transferencia de masa se formula como una función de un coeficiente global de transferencia de masa y una fuerza impulsora. Dicha fuerza impulsora se puede ver como la diferencia entre la concentración de saturación del SO₂ en la solución (C_{sat}) y la concentración actual de SO₂ en dicha solución (segundo invariante C):

$$F_5 = K_{TdeM} \cdot (C_{sat} - C) = k \cdot F_2 \cdot h \cdot (C_{sat} - C) \quad (5.11)$$

donde:

F_5 es, como ya se mencionó, el flujo de SO₂ que pasa del gas a la solución y, K_{TdeM} es el coeficiente de transferencia de masa.

En esta expresión se ha reemplazado el coeficiente global de transferencia de masa (que ya incluía el área de transferencia) por una función: $K_{TdeM} = k \cdot F_2 \cdot h$, que depende de una constante (k), del flujo de gas que entra a la torre (F_2) y la altura total de la torre (h), que aparece como nuevo parámetro del modelo.

En cuanto a los datos utilizados para validar esta relación, se conoce que $C_{sat}=0.016\text{mol/l}$ (Mondal, 2007) es la concentración de saturación del SO₂ disuelto en el jugo tomado como si fuera agua. Esto último no está alejado de la realidad, puesto que en el jugo no existen compuestos químicos que alteren significativamente dicha solubilidad (Hugot, 1986). Ahora, mediante el ajuste de datos experimentales que prescriben eficiencias de transferencia de SO₂ gaseoso al H₂SO₃ solubilizado, en torres entre 3.5 y 7.5 metros, de hasta el 95% , se llega a que $k = 3615\text{m}^2/\text{mol}$ (Alvarez et al., 2009). Finalmente, aunque el fenómeno es dinámico, es decir, por algunos instantes puede detectarse disuelto en el jugo algo de SO₂ sin reaccionar (reacción química (5.1)), el tiempo característico de dicha dinámica es de aproximadamente un segundo, por lo que puede asumirse instantánea.

Tomando como base los balances de masa y la relación constitutiva mencionada anteriormente, se obtienen las dos ecuaciones de estado:

$$\frac{dX}{dt} = -\frac{F_1}{V} X + \frac{F_1}{V} [Ca(OH)^+]_l - \frac{\alpha \cdot k \cdot h \cdot C_{sat}}{V} F_2 + \frac{\alpha \cdot k \cdot h}{V} F_2 \cdot C \quad (5.12)$$

$$\frac{dC}{dt} = -\frac{F_1}{V} C + \frac{k \cdot h \cdot C_{sat}}{V} F_2 - \frac{k \cdot h}{V} F_2 \cdot C \quad (5.13)$$

siendo el primer estado X balance de carga electrónica en la solución) y el segundo estado C (balance del anión último). Además, α es la fracción de ácido sulfuroso ionizado en el jugo:

$$\alpha = \frac{-K_{eq_2} + \sqrt{(K_{eq_2})^2 + 4 \cdot K_{eq_2} \cdot C}}{2 \cdot C} \quad (5.14)$$

que se calcula en función de la constante de equilibrio del ácido sulfuroso (K_{eq_2}):

$$K_{eq_2} = \frac{[HSO_3^-][H^+]}{[H_2SO_3]} = \frac{(\alpha \cdot [H_2SO_3]_{inicial})^2}{(1-\alpha) \cdot [H_2SO_3]_{inicial}} \quad (5.15)$$

y cuyo valor numérico está dado por:

$$K_{eq_2} = 10^{\left(\frac{853}{T_{abs} - 4.74}\right)} \quad (5.16)$$

donde:

T_{abs} es la temperatura de la solución.

Finalmente, el valor del pH a la salida de la torre de sulfitado se calcula luego de hallar la siguiente expresión polinómica para la concentración del ión H^+ :

$$[H^+]^2 + X \cdot [H^+] - K_w = 0 \quad (5.17)$$

donde:

K_w es la constante de disociación del agua, cuyo valor numérico se calcula como:

$$K_w = 0.45344659 \cdot \exp\left(\frac{-19010.7323}{R \cdot T_{abs}}\right) \quad (5.18)$$

La solución de la ecuación (5.17) para la concentración de iones H^+ , ya factorizada para detectar signo y producir siempre concentraciones mayores que cero, es:

$$[H^+] = \frac{-X \pm \sqrt{X^2 + 4 \cdot K_w}}{2} = \frac{X}{2} \cdot \left[-1 + \text{sign}(X) \sqrt{1 + \frac{4 \cdot K_w}{X^2}} \right] \quad (5.19)$$

El modelo del proceso queda entonces determinado, con dos estados (ecuaciones 5.12 y 5.13) y una salida (ecuación 5.19).

2.2 ESTUDIO DE CONTROLABILIDAD DE LA PLANTA

Pensando en el control de la etapa de clarificación como un todo, se propone el siguiente apareamiento de los estados con las acciones de control disponibles:

$$x_1 = X \text{ controlado con } u_1 = [Ca(OH)]^I$$

$$x_2 = C \text{ controlado con } u_2 = F_2.$$

y se reescribe el modelo en variables de estado como:

$$\frac{dx_1}{dt} = -\frac{F_1}{V} x_1 + \frac{F_1}{V} u_1 - \frac{\alpha \cdot k \cdot h \cdot C_{sat}}{V} u_2 + \frac{\alpha \cdot k \cdot h}{V} x_2 \cdot u_2 \quad (5.20)$$

$$\frac{dx_2}{dt} = -\frac{F_1}{V} x_2 + \frac{k \cdot h \cdot C_{sat}}{V} u_2 - \frac{k \cdot h}{V} x_2 \cdot u_2 \quad (5.21)$$

De este modo se logra la separabilidad de estados necesaria para aplicar el análisis de *controlabilidad de estado no lineal* descrito en el capítulo II, basado en el modelo del proceso definido en la forma afín con la entrada (ec. 2.30):

$$\frac{d\mathbf{x}}{dt} = \begin{bmatrix} -\frac{F_1}{V} & 0 \\ 0 & -\frac{F_1}{V} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \frac{F_1}{V} & -\frac{\alpha \cdot k \cdot h \cdot C_{sat}}{V} + \frac{\alpha \cdot k \cdot h}{V} x_2 \\ 0 & \frac{k \cdot h \cdot C_{sat}}{V} - \frac{k \cdot h}{V} x_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad (5.22)$$

Se evalúan entonces en su conjunto los siguientes índices de controlabilidad:

a) *Grados de libertad para el control.*

La selección de dos acciones de control (variables manipuladas) para dos variables controladas garantiza que no hay grados de libertad (*GL*) insatisfechos en el sistema de dos estados, por tanto, el índice de controlabilidad práctica asociado a los grados de libertad se cumple en este caso (*GL*=0).

b) *Rango de la matriz de controlabilidad.*

A partir de la formulación matricial previa (5.22) se construye la matriz de controlabilidad no lineal (ec. 2.34) para este sistema:

$$\mathbf{W}_c = \begin{bmatrix} \frac{F_1}{V} & \frac{-\alpha \cdot k \cdot h \cdot C_{sat}}{V} + \frac{\alpha \cdot k \cdot h}{V} x_2 & \frac{F_1^2}{V^2} & \frac{-\alpha \cdot k \cdot h \cdot C_{sat} \cdot F_1}{V^2} \\ 0 & \frac{k \cdot h \cdot C_{sat}}{V} - \frac{k \cdot h}{V} x_2 & 0 & \frac{k \cdot h \cdot C_{sat} \cdot F_1}{V^2} \end{bmatrix} \quad (5.23)$$

El índice básico de controlabilidad de estado se refiere al rango de la matriz \mathbf{W}_c , que debe ser igual al número “*n*” de estados en un sistema controlable (Hermann y Krener, 1977; Ochoa et al., 2004). En este caso el rango de $\mathbf{W}_c = 2 = n$ para todo posible diseño de la sulfitadora. En efecto, $\mathbf{W}_c < 2$ ocurriría solo si ambas filas de la matriz \mathbf{W}_c fuesen idénticas, es decir, si se cumplieran las tres condiciones siguientes:

$$\left\{ \begin{array}{l} \bullet \quad \frac{F_1}{V} = 0; \quad \frac{F_1^2}{V^2} = 0 \\ \bullet \quad \left(\frac{-\alpha \cdot k \cdot h \cdot C_{sat}}{V} + \frac{\alpha \cdot k \cdot h}{V} x_2 \right) = \left(\frac{k \cdot h \cdot C_{sat}}{V} - \frac{k \cdot h}{V} x_2 \right) \\ \bullet \quad \left(\frac{-\alpha \cdot k \cdot h \cdot C_{sat} \cdot F_1}{V^2} \right) = \left(\frac{k \cdot h \cdot C_{sat} \cdot F_1}{V^2} \right) \end{array} \right. \quad (5.24)$$

o finalmente, si $F_1=0$ ó $V=\infty$ ó $\alpha=-1$, lo cual es imposible en procesos reales. Esto indica que en el sentido estricto, el sistema es controlable en todo el espacio de estado.

c) *Determinante de la matriz asociada a la respuesta forzada.*

En el análisis de Ochoa (2005) se recalca que la evaluación de la controlabilidad del proceso no lineal está asociada al punto final de equilibrio forzado \mathbf{x}^* que se desea alcanzar. En ausencia de perturbaciones ninguno de los estados del sistema cambia con el tiempo cuando se aplica la acción de control forzante \mathbf{u}^* dada por:

$$\mathbf{u}^* = -[\mathbf{G}(\mathbf{x}^*)]^{-1} \cdot \mathbf{f}(\mathbf{x}^*) \quad (5.25)$$

Si la inversa de la matriz \mathbf{G} no existe, entonces no será posible obtener una acción de control que permita garantizar el punto de equilibrio deseado. La existencia de la inversa $[\mathbf{G}]^{-1}$ se garantiza si $\det(\mathbf{G}) \neq 0$.

Para la sulfitadora:

$$\det(\mathbf{G}) = \frac{F_1 \cdot k \cdot h \cdot C_{sat}}{V^2} - \frac{F_1 \cdot k \cdot h}{V^2} x_2 \neq 0 \quad (5.26)$$

La ecuación anterior se cumple si se puede establecer que C_{sat} es diferente del estado $x_2 = C$. Analizando el modelo, se evidencia que la concentración de saturación de SO₂ disuelto en soluciones acuosas difícilmente puede ser igual a la concentración de ácido en la torre sulfitadora (C) a las condiciones típicas de operación del proceso (Mondal, 2007).

d) *Pertenencia de la acción de control forzante (\mathbf{u}^*) al intervalo disponible de acciones de control.*

Finalmente, es necesario verificar la pertenencia de las acciones de control a los rangos disponibles. Utilizando la ecuación (5.26) es posible calcular las entradas asociadas al punto de equilibrio (X^*, C^*) de la planta:

$$\mathbf{G}^{-1} = \begin{pmatrix} \frac{V}{F_1} & \frac{\alpha \cdot V}{F_1} \\ 0 & \frac{V}{k \cdot h \cdot (x_2 - C_{sat})} \end{pmatrix} \quad (5.27)$$

y luego:

$$u_1^* = X^* + \alpha \cdot C^* \quad (5.28)$$

$$u_2^* = \frac{F_1 \cdot C^*}{k} \quad (5.29)$$

3. FORMULACION DEL PROBLEMA DE OPTIMIZACION

A continuación se presenta el planteamiento del problema de *Diseño Integrado de la Torre Sulfitadora*. Éste consiste en, dados el caudal y la temperatura del jugo de caña que llega a la etapa de sulfitación, determinar las dimensiones, el punto de operación y los parámetros del regulador a usar, que optimicen el funcionamiento de esta etapa. Esto se logra minimizando los costos de inversión y de operación de la planta y un índice de desempeño dinámico a lazo cerrado, mientras se imponen restricciones sobre las condiciones de controlabilidad de estado.

Un paso previo a la formulación matemática del problema de diseño integrado, es la definición de los objetivos económicos, de operación y de control, en función de los requerimientos del proceso. Para el caso particular de la torre sulfitadora seleccionada, se desea obtener una corriente de salida de jugo de caña con un pH alrededor de 4.5, con un flujo de jugo de caña a la entrada de 11.05 lit/s.

Los objetivos económicos se centran en minimizar los costos de inversión, relacionados directamente con la altura de la torre y los costos de operación, relacionados con el flujo de gas (SO_2). Entre los objetivos de operación se tiene que el flujo máximo de gas no supere los 0.013 mol/s y que el pH del jugo entrante (pHe), regulado previamente en el tanque de encalado, se encuentre entre 8 y 9 unidades.

En cuanto a los objetivos de control, en primer lugar, los diseños deben satisfacer las métricas de controlabilidad de estado calculadas en la sección anterior, para garantizar que la planta sea controlable. Dos de los índices calculados, se cumplen a priori, la matriz de controlabilidad tiene rango completo y no hay grados de libertad insatisfechos, pero se debe verificar que existe la inversa de la matriz asociada a la respuesta forzada y que las relaciones entre las variables de estado y los parámetros de diseño permitan obtener valores en los rangos aceptables para las entradas de control, según las ecuaciones (5.28) y (5.29). Por otra parte,

operando a lazo cerrado se debe minimizar el error en la respuesta cuando el sistema está sometido a perturbaciones típicas del proceso, lo cual se mide utilizando el ISE (*Integral del cuadrado del error*). Además, se desea que el elemento final de control, que en este caso es la válvula que regula el flujo de gas (F_2), se sature el menor número de veces posible. A continuación se presenta la descripción del sistema de control.

Para la medición de las salidas, se tiene un sensor de pH ubicado en la descarga de la torre sulfitadora, cuya señal permite obtener el primer estado X a través de la ecuación (5.19). La medición del segundo estado C es mucho más compleja puesto que se deberían detectar grupos sulfitos (grupos SO_3^{2-}) donde se encuentren (Light, 1997). Sin embargo, utilizando el sensor de pH de salida y un sensor de pH en la entrada, y mediante el balance iónico en la sulfitadora, es posible reconstruir el segundo estado: C (Alvarez et al., 2001).

Un apareamiento variables manipuladas-estados como el presentado arriba, se justifica en un montaje real si se tiene en cuenta que en el tanque de encalado (previo a la sulfitadora), se realiza la adición de lechada de cal de modo que se logre la concentración de $\text{Ca}(\text{OH})^+$ deseada en la línea de entrada a la sulfitadora ($C_{\text{Ca}(\text{OH})^+,1}$). Esta es la primera acción de control, que actúa sobre el primer estado del proceso (invariante asociado con la carga electrónica (X)). Por otra parte, la adición de SO_2 gaseoso (F_2), que conlleva una reacción instantánea $\text{SO}_2\text{-H}_2\text{O}$ pero con fracción α disociada del ácido sulfuroso que se produce, conforma la segunda acción de control, utilizada para regular la concentración del anión último (segundo estado C), de vital importancia para el blanqueo y desinfección.

A pesar de que, para instalar lazos de control para los dos estados se requieren dos sensores, lo que ocurre normalmente en los procesos industriales para producir azúcar, es que sólo se dispone de un sensor de pH en la línea de salida de jugo de la sulfitadora. Por lo tanto, el control de la sulfitadora utilizando la instrumentación típica de este proceso a nivel industrial (solo un medidor de pH), se limita a tomar como única acción de control el flujo gas (SO_2) desde el quemador de azufre (F_2) para regular el pH del jugo a la salida de la sulfitadora.

Por consiguiente, se debe resaltar el desplazamiento de los objetivos de control del espacio de estado (control total del estado $x_1=X$ y $x_2=C$) a un objetivo de control único en el

espacio entrada-salida (la salida única del proceso dada por el sensor de pH en la descarga de jugo desde la sulfitadora).

Esta reducción del problema lo muestra como uno de una entrada – una salida, con la salida asociada directamente al primer estado ($x_1 = X$) y considerando el otro estado (C) como una perturbación para la sulfitadora. Ese supuesto no está lejos de la realidad del montaje mostrado en la figura 5.1, ya que C es consecuencia directa del encalado y a su vez afecta a X . Sin embargo, debe recordarse que la aproximación por invariantes de reacción establece dos estados desde los cuales se reconstruye el pH como salida del proceso. Por lo tanto, el problema es monovariable desde la visión entrada-salida, pero resulta ser de dos entradas (F_2 y $[Ca(OH)^+]$), dos estados (X y C) y una salida (pH), desde la visión de control de estado. Aunque se supone que en el encalado previo se controla la concentración de $Ca(OH)^+$, ésta es una de las perturbaciones si sólo se tiene una acción de control (F_2) junto con el caudal de jugo F_1 . En la figura 5.3 se muestra el Diagrama de Flujo de Información que representa al modelo dinámico de la sulfitadora dado por las ecuaciones (5.20) y (5.21). Lo recién discutido se aprecia mejor en este diagrama, donde se ve el efecto que tiene manipular el flujo de gas (F_2) para controlar el pH, como se hace en los ingenios. Sin embargo, otras estructuras de control podrían proponerse para lograr un mejor desempeño.

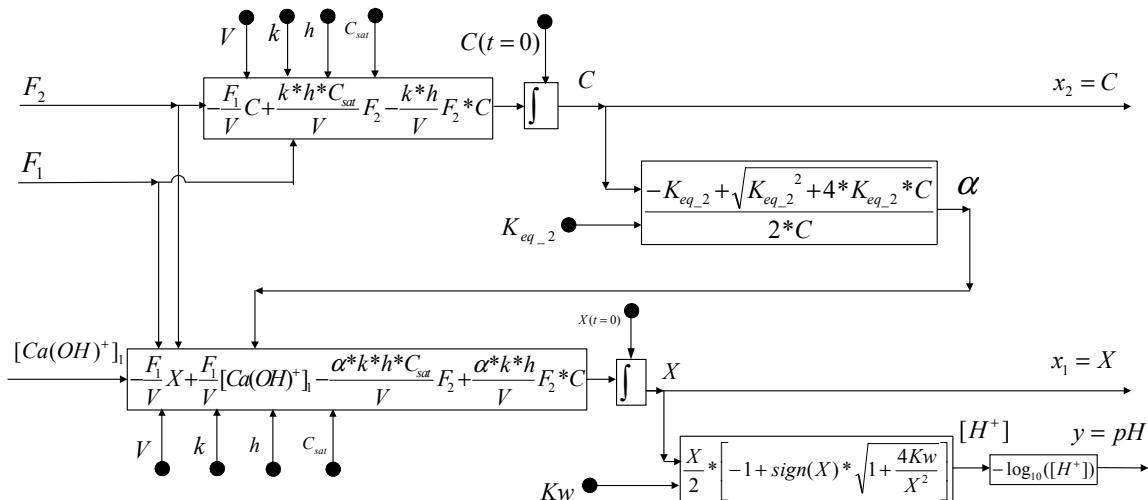


Figura 5.3. Representación del modelo de la sulfitadora en un diagrama de flujo de información.

Con la intención de mantener la automatización del proceso con el menor costo económico posible, es práctica común en los ingenios dotar al proceso de encalado y sulfatado de un solo lazo de control: el de pH. Buscando una réplica exacta de la situación más frecuente en la práctica industrial, en este trabajo se reduce entonces el proceso de un sistema multivariable a un proceso monovariable. Aunque existe interacción entre las dos dinámicas, se puede observar del modelo que tal interacción es en una sola dirección: sólo de una variable sobre la otra, lo que hace válida tal reducción.

Finalmente, la estrategia de control para este caso de estudio se resume en controlar el pH a la salida manipulando el flujo de gas (F_2), el estado C se constituye así en una variable intermedia (como se ve en la figura 5.3), mientras que la concentración de iones de calcio provenientes del encalado y el flujo de jugo (F_1) son las perturbaciones.

Formulación matemática del problema

Al formular el problema de optimización se pretende determinar la altura de la torre sulfitadora (h), el flujo de SO₂ (F_2) y el pH del jugo de entrada (pHe), en estado estacionario, que optimicen el funcionamiento de esta planta, minimizando los costos de inversión, de operación y control, cumpliendo las restricciones impuestas sobre los índices de controlabilidad de estado y de desempeño dinámico.

Función objetivo:

La función objetivo contiene los costos de construcción y operación de la planta (en unidades monetarias normalizadas, UMN), del siguiente modo:

$$f_{costo} = w_1 \cdot \bar{h}^2 + w_2 \cdot \bar{F}_2 + factorcontrol \cdot ISE \quad (5.30)$$

donde:

\bar{h} : altura normalizada de la torre de sulfitación,

\bar{F}_2 : el flujo normalizado de gas que entra a la torre,

$w_1, w_2, factorcontrol$: pesos de la función de costos.

El primer término representa los costos de construcción y el segundo es proporcional a los costos de operación. Se seleccionan arbitrariamente los valores de 0.3 y 0.7 para w_1 y w_2 respectivamente, considerando criterios de diseño ya establecidos (Walas, 1990). El tercer término, proporcional al ISE, es para representar los costos relacionados directamente con el funcionamiento del sistema de control y que son mayores a medida que hay mayor desviación respecto al punto de operación deseado.

La minimización de la función de costo combinada recién descrita debe cumplir con las restricciones que se presentan a continuación.

Restricciones de balances:

Se resuelven las ecuaciones de estado (5.20) y (5.21) para obtener un punto de operación estacionario.

Restricciones físicas:

De acuerdo con el modelo de referencia de la planta real utilizada, se estima que la relación altura y sección del depósito de la sulfitadora es 10 aproximadamente.

Se fijan también los límites físicos razonables de las variables tanto de pH a la salida de la torre como del flujo de gas SO₂ introducido en el líquido del sulfitador:

$$pH \cong 4.5 \Rightarrow |pH - 4.5| \leq \varepsilon \quad (5.31)$$

$$F_2 < 0.013 \text{ mol/s} \quad (5.32)$$

Otras restricciones, que deben introducirse para garantizar valores lógicos en algunos parámetros son:

$$0 < \alpha < 1 \quad (5.33)$$

$$0 < k < 1 \quad (5.34)$$

Restricciones de controlabilidad:

Las restricciones derivadas de la aplicación de las métricas de controlabilidad de estado, ecuaciones (5.26), (5.27) y (5.28) son:

$$\left(\frac{F_1 \cdot k \cdot h \cdot C_{sat}}{V^2} - \frac{F_1 \cdot k \cdot h}{V^2} x_2 \right) \neq 0 \Rightarrow \text{abs} \left(\frac{F_1 \cdot k \cdot h \cdot C_{sat}}{V^2} - \frac{F_1 \cdot k \cdot h}{V^2} x_2 \right) \leq \varepsilon \quad (5.35)$$

$$u_1^* \rightarrow u_{1min} \leq X^* + \alpha \cdot C^* \leq u_{1max} \quad (5.36)$$

$$u_2^* \rightarrow u_{2min} \leq \frac{F_1 \cdot C^*}{k} \leq u_{2max} \quad (5.37)$$

Se tienen dos restricciones adicionales que requieren la evaluación de la respuesta dinámica a partir de las ecuaciones 5.12 y 5.13 del modelo. Una restricción se establece sobre la señal de control para evitar que la válvula se sature durante la operación. Ésta se define como:

$$Sat \leq 0 \quad (5.38)$$

donde:

Sat: número de veces que la señal de control alcanza sus valores máximo y mínimo, saturando la válvula.

La otra se impone una sobre el índice ISE, definido como:

$$ISE = \int_{t=0}^{Tmax} \left(\frac{pH_{ref} - pH}{pH \text{ span}} \right)^2 \cdot dt \quad (5.39)$$

donde:

pH span es el máximo rango de variación previsto para el pH,

Tmax es el tiempo de operación considerado.

En el caso de que no se incluyan los costos asociados al control (*factorcontrol*=0) en la función objetivo, se impone la condición:

$$ISE \leq ISE_{max} \quad (5.40)$$

para garantizar los requerimientos mínimos de desempeño dinámico de la planta.

4. DISEÑO INTEGRADO DE LA TORRE SULFITADORA CON UN CONTROLADOR PID

Se propone un lazo de control retroalimentado con un controlador PID, para ilustrar el diseño integrado con una estrategia típica de control en los ingenios azucareros. No obstante, la metodología es directamente extensible a aplicaciones de control avanzado como los controladores predictivos multivariados (Francisco y Vega, 2006)

La ley de control PID está dada por:

$$u_2(t) = \bar{F}_2 + K_p \cdot (pH_{ref} - pH(t)) + \frac{K_p}{Ti} \int_{t_0}^t (pH_{ref} - pH(t)) dt + K_p \cdot Td \cdot \frac{d(pH_{ref} - pH(t))}{dt} \quad (5.41)$$

donde:

u_2 : acción de control sobre el flujo de gas,

\bar{F}_2 : valor del flujo de gas en estado estacionario, corresponde al calibrado del controlador.

La acción derivativa en la ley de control, siguiendo reglas clásicas de sintonización, se toma como $Td=0.16Ti$. Los parámetros Kp y Ti , en cambio, son ajustados en el procedimiento de diseño.

La respuesta dinámica de la planta se evalúa utilizando el modelo no lineal del proceso (5.20-5.21), por un tiempo de simulación de 3000seg, considerando perturbaciones sobre el flujo y pH de la corriente de entrada (F_e y pHe) como las que se muestran en la figura 5.4. En el caso de pHe la figura es referencial ya que el valor inicial de la variable es el pH de entrada en estado estacionario obtenido en la optimización (\bar{pHe}), sobre el cual ocurren variaciones entre $9.74 \cdot 10^{-5}$ y $9.74 \cdot 10^{-7}$.

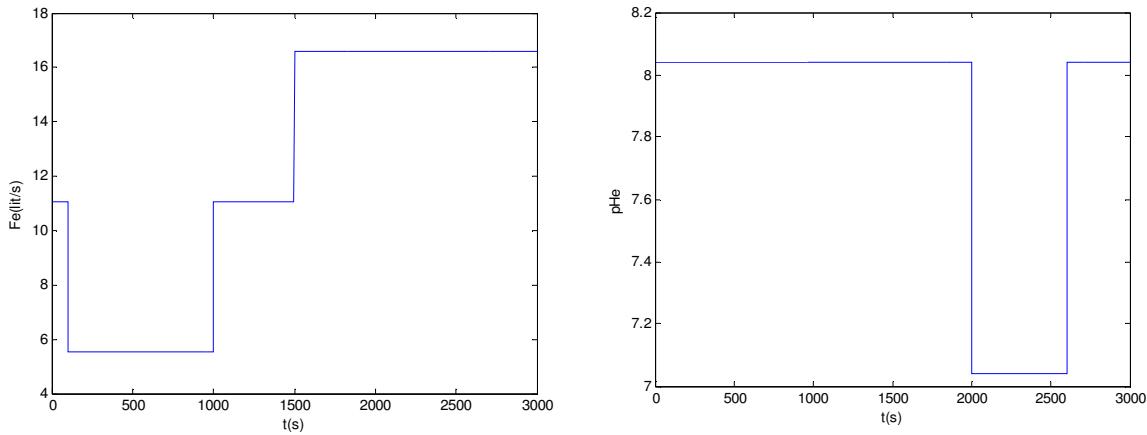


Figura 5.4. Régimen de perturbaciones en el flujo de entrada Fe y el pH de entrada pHe .

En el algoritmo genético desarrollado para resolver la optimización, la solución típica se codifica en un cromosoma que es un vector real de longitud fija que contiene las siguientes variables:

$$[h, \bar{F}_2, \overline{pHe}, K_p, Ti]$$

Se usan los operadores de selección por ruleta, el cruce aritmético y la mutación aleatoria, ajustados tomando como base los resultados de trabajos previos (Revollar et al., 2004; Revollar et al. 2005; Revollar et al., 2006). Las soluciones que no cumplen las restricciones se tratan con una función de penalización (ecuación 4.24). El problema se resuelve con una población de 100 individuos, una tasa de mutación decreciente que va de 0.1 a 0.02 y un número máximo de iteraciones de 150. La evaluación de la calidad (adaptación) de cada individuo, es decir, de la función objetivo (5.30) se realiza en el algoritmo genético según el esquema de la figura 5.5.

El diseño se centra en lograr una especificación de $pH \approx 4.5$ a la salida de la planta. Con la intención de comparar diferentes exigencias de control, se plantean dos escenarios para la resolución del problema:

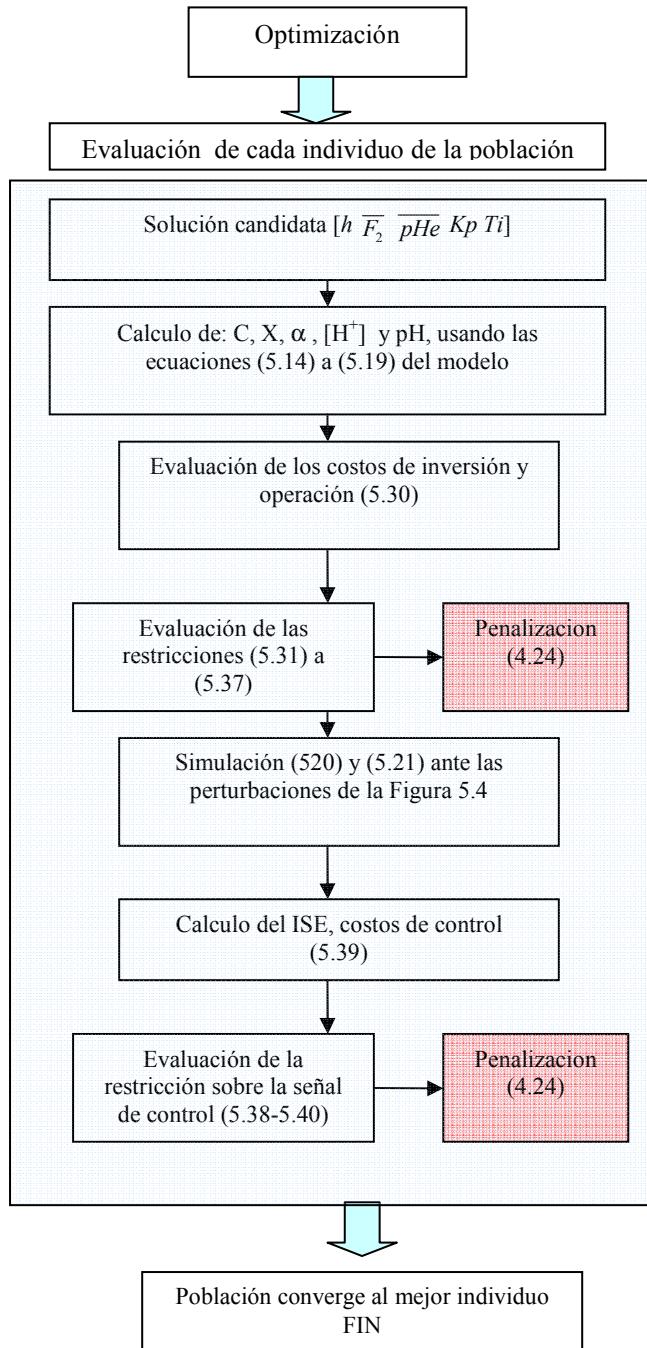


Figura 5.5. Proceso del algoritmo genético correspondiente a la evaluación de la función de adaptación para cada posible solución.

El diseño integrado considerando los costos de control en la función objetivo (5.30), con $factorcontrol=0.1$, para obtener la planta más económica posible con el mejor desempeño dinámico.

- El diseño integrado con el ISE sólo como restricción de controlabilidad (ec. 5.40), (*factorcontrol=0* en la ec. 5.30), para obtener el diseño más económico que satisface los requerimientos mínimos de desempeño dinámico.

Los valores de los parámetros de la planta, entradas y condiciones de diseño establecidas previamente se muestran en la tabla 5.1.

Tabla 5.1. Parámetros de la torre de sulfitación

Parámetros	
<i>pH salida</i>	4.5
<i>Temperatura de operación (Te)</i>	328° K
<i>Concentración de saturación (Csat)</i>	0.016 mol/lit
<i>Flujo de entrada del jugo (F₁)</i>	11.05 lit/s
<i>Coeficiente asociada a la transferencia de masa (k)</i>	7.45 lit/(m.mol)
<i>Relación altura/diámetro (h/d)</i>	10
<i>ISE_{max}</i>	8
Parámetros relacionados con la válvula reguladora de F₂:	
<i>Constante de tiempo de la válvula (τv)</i>	5
<i>Capacidad de la válvula</i>	50
<i>F₂ operación/F_{2max}</i>	0.5

En la tabla 5.2 se comparan los resultados obtenidos al diseñar la planta considerando los dos escenarios planteados anteriormente. En ambos casos se verifica la controlabilidad de estado obteniéndose un determinante de la matriz asociada a la respuesta forzada distinto de cero. Además, los valores de la acción de control forzante están dentro del intervalo disponible de acciones de control, como se observa en la tabla 5.2.

En las figuras 5.6 y 5.7 se muestran las respuestas en lazo cerrado de las plantas diseñadas y las respuestas de las mismas a lazo abierto, para poner en evidencia el efecto del controlador sobre el comportamiento general. En la figura 5.6 se puede apreciar que la planta obtenida en el escenario 1 muestra una respuesta menos sensible a perturbaciones, lo cual está

acorde con la optimización del ISE. En la figura 5.8 se muestran las señales de control para ambos casos, donde se puede observar que no hay saturación del elemento final de control.

Tabla 5.2. Resultados del diseño integrado de la planta

	Diseño Integrado con ISE en costos (Escenario 1)	Diseño Integrado con restricción sobre el ISE (Escenario 2)
Costo (UMN)	0.16	0.13
$h (m)$	4.99	3.38
$\overline{F_2} (mol/s)$	0.0013	0.0018
\overline{pHe}	8.57	8.53
K_p	-0.114	-0.0311
T_i	16.45	9.12
T_d	2.63	1.46
ISE	0.20	5.65
u_1	$9.74 \cdot 10^{-7} \leq 3.62 \cdot 10^{-5} \leq 9.74 \cdot 10^{-5}$	$9.74 \cdot 10^{-7} \leq 3.28 \cdot 10^{-5} \leq 9.74 \cdot 10^{-5}$
u_2	$5.20 \cdot 10^{-5} \leq 1.30 \cdot 10^{-3} \leq 2.0 \cdot 10^{-3}$	$7.20 \cdot 10^{-5} \leq 1.80 \cdot 10^{-3} \leq 3.6 \cdot 10^{-3}$
Sat	0	0

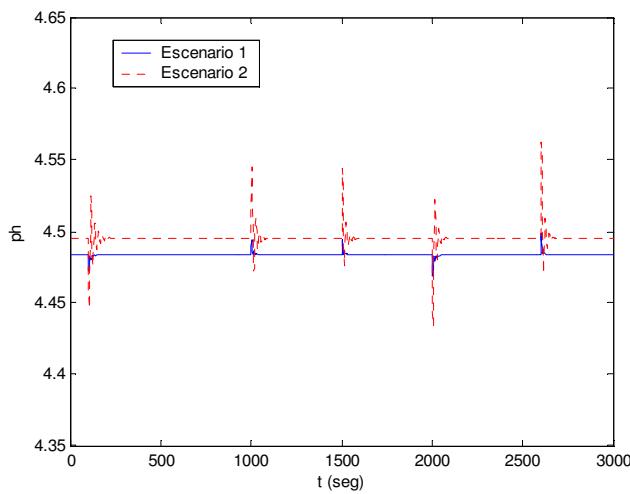


Figura 5.6. Respuesta a lazo cerrado (pH) de las plantas diseñadas ante las perturbaciones descritas en la figura 5.4

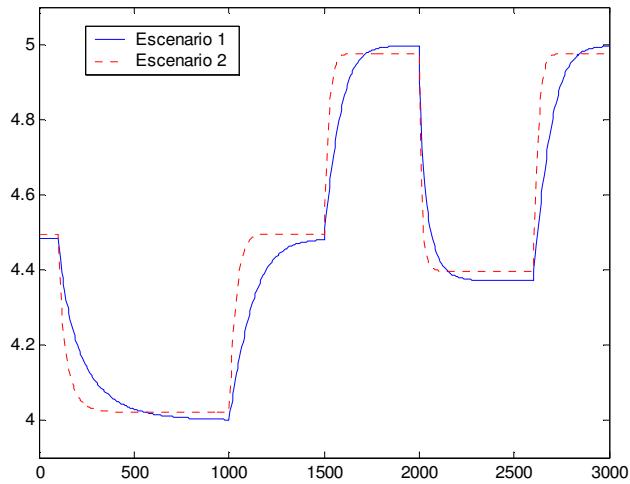


Figura 5.7. Respuesta a lazo abierto de las plantas diseñadas frente a las perturbaciones descritas en la figura 5.4

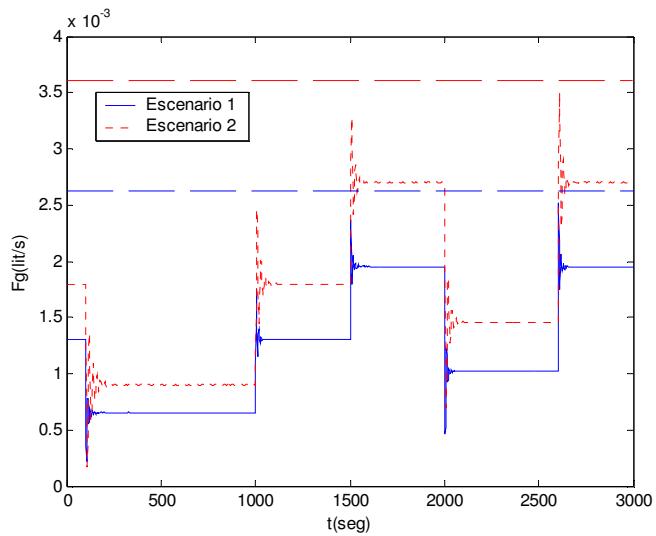


Figura 5.8. Señal de control de los diseños obtenidos en ambos escenarios.

Se resuelve también el problema de diseño siguiendo la metodología secuencial clásica, con la intención de compararla con el diseño integrado. Para ello se realizan, en primer lugar, los mapas de operación de la torre sulfitadora para el pH de salida deseado, evaluando los costos de inversión y operación en cada caso. Se escoge la planta que cumple las especificaciones con el mínimo costo y, para esta planta, se prueban varios controladores

minimizando el ISE como medida de desempeño. De esta manera se determina el mejor controlador para la planta más económica.

Tabla 5.3. Resultados del diseño de la planta aplicando la metodología clásica

Diseño secuencial	
<i>Costo (UMN)</i>	0.093
\bar{F}_2 (mol/s)	0.0013
\overline{pHe}	8
K_p	-0.087
T_i	9.786
T_d	1.566
<i>ISE</i>	0.625
u_1	$9.74 \cdot 10^{-7} \leq 9.74 \cdot 10^{-5} \leq 9.74 \cdot 10^{-5}$
u_2	$5.20 \cdot 10^{-5} \leq 1.3 \cdot 10^{-3} \leq 2.6 \cdot 10^{-3}$
<i>Sat</i>	2

En la tabla 5.3 se muestran los resultados obtenidos, y en la figuras 5.9 y 5.10 se encuentran la respuesta a lazo cerrado y la señal de control respectivamente. La planta diseñada es bastante pequeña y con costos menores a las diseñadas aplicando la metodología integrada. El ISE alcanza valores bastante bajos, pero la señal de control llega a saturarse, como se puede ver tanto en la tabla 5.3 como en la figura 5.10. Esta no es una condición deseable en la operación. Aun cuando resulta una planta bastante económica, su desempeño dinámico no logra superar el de las plantas anteriores, presentando una oscilación continuada tanto en la salida del sistema como en la señal de control, a pesar de las buenas condiciones de controlabilidad de estado que la planta tiene intrínsecamente.

Es importante destacar además, que aplicar la metodología de diseño integrado representa una gran ventaja, pues se obtienen simultáneamente las dimensiones de la planta, junto con el punto de trabajo, con un controlador sintonizado de manera de garantizar los requerimientos de controlabilidad impuestos.

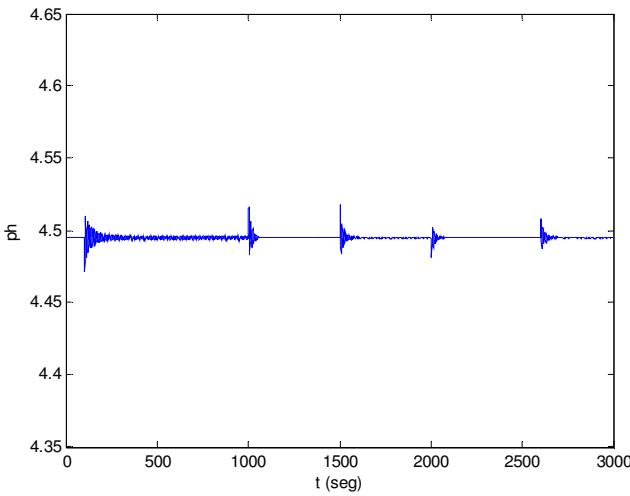


Figura 5.9. Respuesta a lazo cerrado (pH) de la planta obtenida aplicando la metodología de diseño clásico

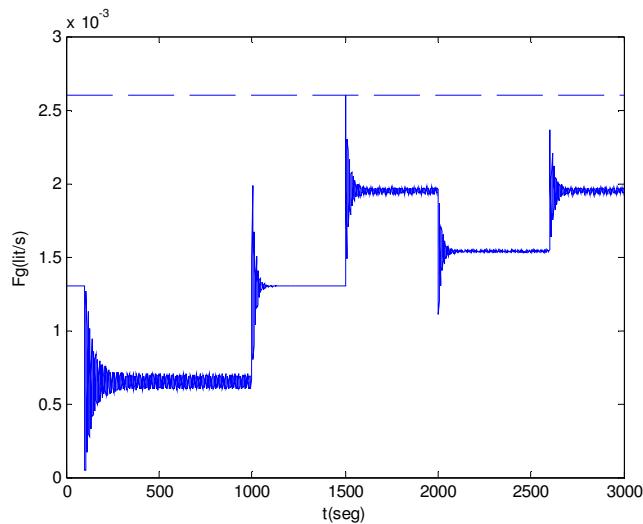


Figura 5.10. Señal de control para la planta obtenida aplicando la metodología de diseño clásico.

5. DISEÑO INTEGRADO DE LA TORRE SULFITADORA CON UN CONTROLADOR PREDICTIVO (MPC)

La aplicación de controladores avanzados en el marco del diseño integrado de procesos se ve limitada por la complejidad de los problemas de optimización resultantes. En los trabajos de Brengel y Seider (1992), Sakizlis et al. (2003), Sakizlis et al. (2004) y Francisco y

Vega (2006) se proponen distintas estrategias para introducir controladores predictivos basados en modelos (“*Model Based Predictive Controllers*”-MPC) en el diseño integrado. Esto tiene como inconveniente la necesidad de resolver un problema de optimización en línea para calcular la ley de control en cada paso de la optimización de la planta con su sistema de control (Sakizlis et al., 2004). Sin embargo, las cualidades de los MPC y las ventajas que ofrecen sobre los esquemas PID convencionales (Maciejowsky, 2002) justifican el interés de su incorporación en los esquemas considerados para el diseño integrado.

El cálculo de la ley de control del MPC se formula como un problema de optimización de la siguiente función de costos:

$$\min_{\Delta u} (V(k)) = \sum_{i=H_w}^{H_p} (\hat{y}(k+i|k) - r(k+i|k))^2 \cdot W_y(i) + \sum_{i=0}^{H_c-1} (\Delta \hat{u}(k+i|k))^2 \cdot W_u(i) \quad (5.42)$$

sujeta a restricciones sobre las entradas, la predicción de las salidas y las variaciones en las variables manipuladas:

$$\begin{aligned} u_{\min} < \hat{u}(k+i|k) &< u_{\max} \quad i = 0, \dots, H_c \\ y_{\min} < \hat{y}(k+i|k) &< y_{\max} \quad i = H_w, \dots, H_p \\ \Delta u_{\min} < \Delta \hat{u}(k+i|k) &< \Delta u_{\max} \quad i = 0, \dots, H_c \end{aligned}$$

donde:

\hat{y} : salida futura estimada,

r : trayectoria de referencia,

Δu : incrementos de la señal de control,

H_p : horizonte de predicción,

H_c : horizonte de control,

$W_y(i)$: matriz diagonal que representa los pesos de los esfuerzos de control a lo largo del horizonte de control,

$W_u(i)$: matriz diagonal que representa los pesos de los errores de las salidas respecto a la trayectoria de referencia marcada.

Como modelo de predicción para estimar las salidas futuras \hat{y} se utiliza generalmente el modelo discretizado lineal en espacio de estado del proceso. Las matrices W_y y W_u dependen del instante de muestreo i , posibilitando la penalización del error únicamente en ciertos puntos, denominados puntos de coincidencia.

Este es un problema de optimización cuadrático (QP), que se resuelve en cada instante de muestreo k y permite determinar la secuencia de acciones de control $\Delta\hat{u}(k+i | k)$ óptima que se debe aplicar sobre el sistema en el instante de tiempo $k+1$ hasta que se cumpla el horizonte de control H_c , y a partir de ese instante los Δu son nulos, es decir:

$$\hat{u}(k+i | k) = \hat{u}(k + H_c - 1 | k) \quad \forall i \geq H_c$$

Este problema se debe articular en el marco del problema de diseño integrado, dando lugar a un problema de optimización de dos niveles, en un nivel superior se calculan los parámetros de la planta y los parámetros del controlador predictivo y en un nivel inferior se calcula la ley de control del MPC resolviendo el problema (5.42). En el diseño integrado los parámetros del MPC (W_y , W_u , H_p , H_c) se determinan mediante *sintonía automática*, para lo cual se encuentran distintas técnicas en la literatura (Ali et al., 1993; Francisco et al., 2005; Li y Du, 2002; Vega et al., 2007).

En la aplicación presentada aquí, se pretende llevar a cabo el diseño integrado de la torre de sulfitación utilizando un esquema de control predictivo no lineal. Se trata de un ejemplo sencillo donde se considera solo una variable controlada (pH a la salida) y una variable manipulada (Flujo de SO₂-F₂) para ilustrar cómo se aplica la metodología de diseño integrado con un esquema de control predictivo basado en el modelo no lineal del proceso.

Generalmente se usan métodos deterministas de optimización para hallar la ley de control del MPC. El proceso de optimización debe ser bastante rápido para lograr un control eficiente, por lo que suelen emplearse modelos linealizados de la planta para la predicción, lo que produce problemas de estabilidad y robustez en el caso de sistemas con no linealidades significativas.

Por otra parte, algunas técnicas de sintonía automática de los parámetros del controlador requieren simulaciones de la dinámica del modelo para el cálculo de los índices a optimizar. Esto hace bastante lenta y costosa la optimización cuando se lleva a cabo con los métodos convencionales.

En este caso se propone una técnica de enumeración restringida para realizar la optimización que calcula la señal de control del MPC. Este método sin derivadas permite resolver fácil y rápidamente el problema de optimización del nivel inferior. Se aprovecha esta ventaja para incluir en el procedimiento el uso del modelo no lineal del proceso para la predicción de las salidas, y la evaluación de los índices de controlabilidad dinámicos para la sintonía automática del controlador.

La técnica de enumeración restringida se basa en relacionar los movimientos de la señal de control que se calcula con la posición actual del elemento final de control. Se utiliza esta información y el error de control para reducir el espacio de búsqueda. En cada paso de optimización, la búsqueda se realiza sobre una malla que define un conjunto arbitrario y equidistante de posiciones del elemento final de control (M_{FCE}) centrada en la posición actual del elemento. El tamaño de la malla es proporcional a la magnitud del error de control y el signo del error de control permite descartar una región de la malla en función del movimiento que debe realizar el elemento final de control. Para simplificar aún más el procedimiento se considera un horizonte de control de un solo paso.

Formulación matemática del problema

El problema de optimización es básicamente el mismo descrito por las ecuaciones 5.30 a 5.40, con el esquema de control avanzado MPC (ley de control dada en la ec. 5.42). Para la resolución se optimizan separadamente los costos de construcción y operación por una parte, y el desempeño dinámico del sistema por otra, de manera que la formulación matemática del diseño integrado resulta en un *problema de optimización mixta-entera multio-bjetivo*. Los objetivos en conflicto son: los costos de inversión y operación (ec. 5.30, con *factorcontrol*=0) y el desempeño dinámico cuantificado con el ISE (ec. 5.39).

Las variables de decisión relacionadas con el proceso siguen siendo la altura de la torre (h), y los valores estacionarios de F_2 y pHe . Los parámetros del controlador a optimizar son: Wu , Wy , Hp , M_{FCE} , donde Hp y M_{FCE} son variables enteras.

El problema se resuelve utilizando el algoritmo genético de codificación real, con una modificación para tratar problemas multi-objetivo. El cromosoma que representa las posibles soluciones es el vector:

$$[h, \bar{F}_2, \overline{pHe}, Wy, Wu, Hp, M_{FCE}].$$

La evaluación de cada posible diseño involucra el cálculo de los costos de inversión y operación, la optimización en línea del controlador predictivo y la simulación del modelo, como se representa en la figura 5.11.

En un problema de optimización multi-objetivo la meta es encontrar tantas soluciones óptimas de Pareto como sea posible en lugar de encontrar un óptimo absoluto. Al trabajar con poblaciones de posibles soluciones, los algoritmos genéticos se prestan de manera natural a la resolución de estos problemas (Deb, 2001). El conjunto de soluciones óptimas de Pareto consiste en los vectores de variables de decisión para los cuales uno de los objetivos de la optimización no puede ser mejorado sin empeorar algún otro (Summanwar et al, 2002).

El algoritmo genético se modifica para manejar dos objetivos de optimización: f_1 corresponde a los costos de inversión y operación y f_2 , es el ISE, y corresponde al desempeño dinámico. Igual que en el caso anterior, las violaciones de las restricciones se tratan con una función de penalización, designada con f_1 para las violaciones en las restricciones de proceso y con f_2 para las violaciones en las restricciones de control penalizan. En el proceso de selección de los padres, la población se ordena en función de cada objetivo por separado por lo que se tiene una lista ordenada para cada objetivo. Se descartan los peores individuos de cada lista (25% del total), luego, se identifican los cromosomas que representan las mejores soluciones compromiso mediante una función de adaptación basada en la posición de los individuos en cada lista ordenada.

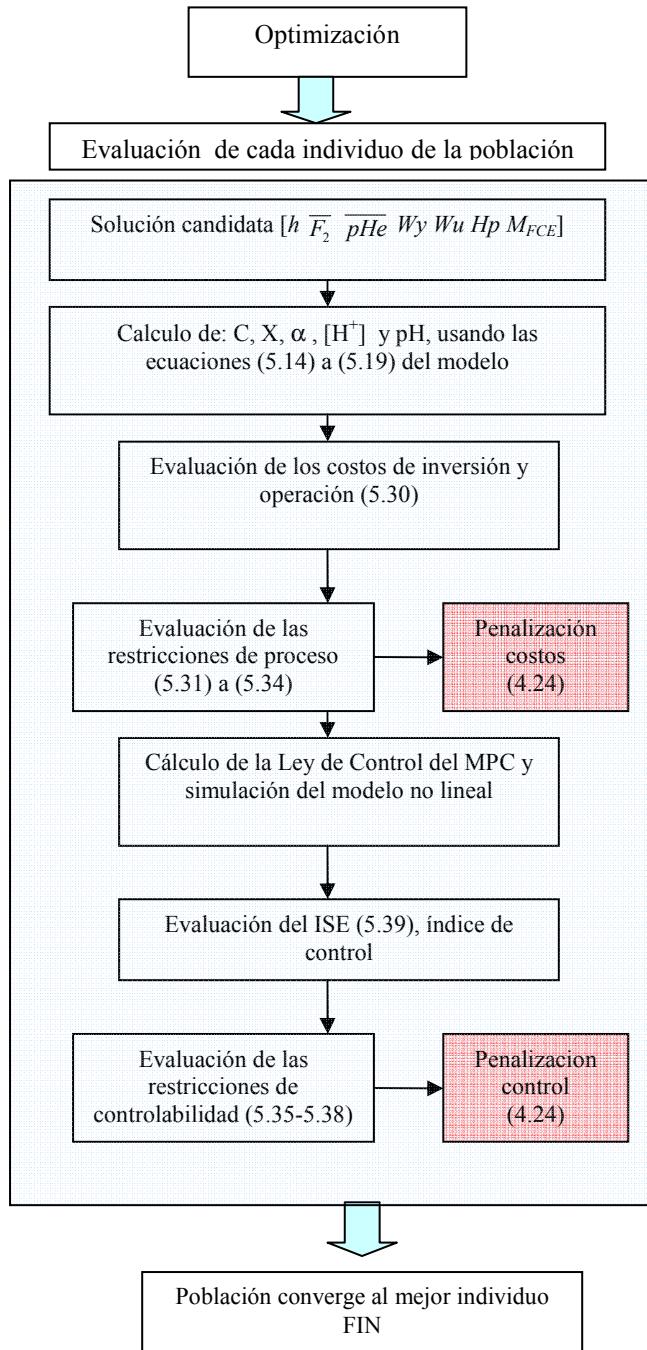


Figura 5.11. Paso de evaluación de la función de adaptación y el controlador MPC para cada posible solución, dentro del algoritmo genético.

En el algoritmo genético se utiliza el cruce aritmético y una tasa de mutación decreciente (0.1 a 0.02). El tamaño de la población es de 200 individuos, se establece un máximo de 50 generaciones y se realizan 10 pruebas para verificar la validez de los resultados.

En la tabla 5.4 se presentan los resultados de la optimización. Se plantean dos escenarios:

- El diseño clásico, donde se obtiene la planta más económica y se controla con el MPC no lineal sintonizado de forma automática optimizando el ISE.
- El diseño integrado para obtener la mejor solución compromiso entre costos de inversión y operación y desempeño dinámico.

Tabla 5.4. Resultados del diseño con un controlador predictivo (MPC) no lineal

	Diseño Clásico	Diseño Integrado
Costo (UMN)	009	0.10
ISE	31	10
h (m)	3	3.61
\overline{F}_2 (mol/s)	0.0013	0.0012
pHe	8	8
W_y	792	493
W_u	237	155
H_p	4	5
M_{FCE}	21	17
u_1	$9.74 \cdot 10^{-7} \leq 9.74 \cdot 10^{-6} \leq 9.74 \cdot 10^{-5}$	$9.74 \cdot 10^{-7} \leq 9.74 \cdot 10^{-6} \leq 9.74 \cdot 10^{-5}$
u_2	$5.20 \cdot 10^{-5} \leq 1.3 \cdot 10^{-3} \leq 2.6 \cdot 10^{-3}$	$7.20 \cdot 10^{-5} \leq 1.80 \cdot 10^{-3} \leq 3.6 \cdot 10^{-3}$
Sat	0	0

En las figuras 5.12 y 5.13 se muestran la respuesta dinámica y la señal de control para ambos diseños. Ambas plantas exhiben un buen desempeño dinámico, como es de esperarse por las buenas características de controlabilidad que presenta la planta y por la eficiencia del control predictivo. Aun así, la aplicación de la metodología de diseño integrado permite lograr

una mejora del 60% en el ISE y una reducción del 10% en los costos. Además, los resultados demuestran que haciendo uso de métodos de optimización alternativos se pueden superar los inconvenientes relacionados con la complejidad del problema para introducir estrategias de control avanzado en el marco del diseño integrado.

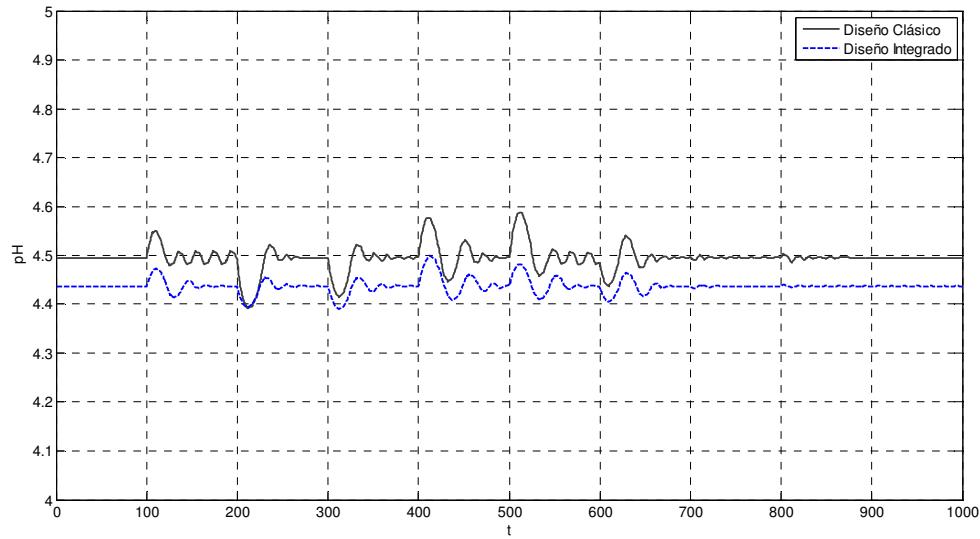


Figura 5.12. Respuesta a lazo cerrado (pH) de las plantas obtenidas

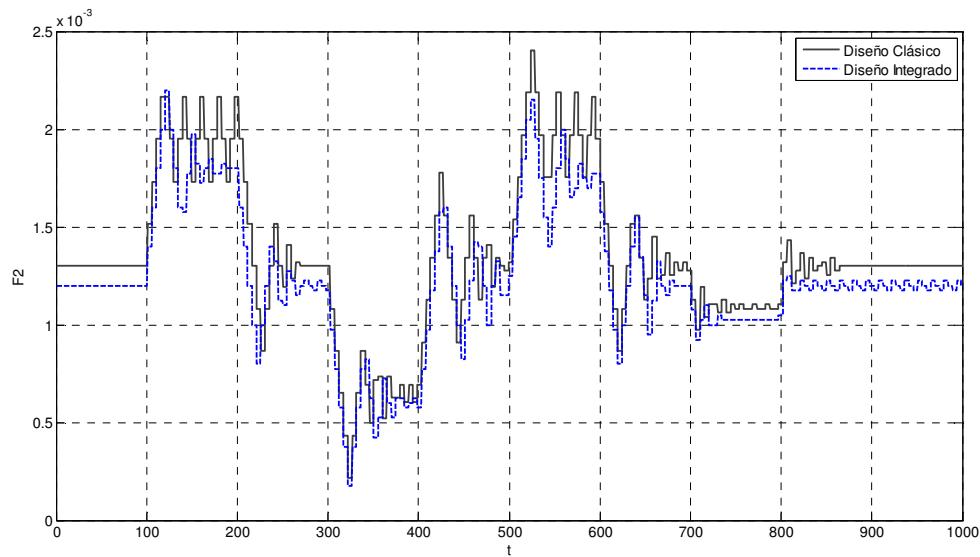


Figura 5.13. Señal de control para el diseño obtenido en ambos escenarios.

6. CONCLUSION

En este capítulo, se hace el diseño integrado de la planta sulfitadora de un proceso de refinación del azúcar de caña. Se presenta una aproximación novedosa al problema considerando tanto la controlabilidad de estado (basada en métricas de controlabilidad práctica para sistemas no lineales) como la controlabilidad de salida. Esta propuesta se traduce en un problema de optimización no lineal donde se imponen restricciones sobre las condiciones de operación de la planta, las medidas de controlabilidad de estado y algunos índices de desempeño a lazo cerrado, mientras se minimizan los costos de inversión, operación y control. El problema de optimización se resuelve usando el algoritmo genético de codificación real presentado en el capítulo IV.

La controlabilidad de estado estudiada se usa en primer término para definir cualitativamente la concepción del lazo de control, y el apareamiento más conveniente variable manipulada-variable controlada, que garantiza la ausencia de grados de libertad insatisfechos y el rango completo de la matriz de controlabilidad no lineal. Además, los requisitos de la controlabilidad de estado también establecen límites o rangos posibles en las variables de diseño, que se usan como restricciones en el cálculo del diseño integrado.

Los resultados ponen en evidencia la gran ventaja de la metodología integrada respecto a la metodología clásica de diseño ya que se obtienen simultáneamente las dimensiones de la planta, el punto de trabajo, y un controlador sintonizado de manera de garantizar los requerimientos de controlabilidad impuestos.

CAPÍTULO VI

DISEÑO INTEGRADO DEL PROCESO DE FANGOS ACTIVADOS

1. INTRODUCCIÓN

La toma de conciencia a nivel mundial respecto a la protección del medio ambiente ha aumentado la demanda de tecnologías eficientes para el tratamiento de aguas residuales. En las plantas de tratamiento de aguas, el proceso de fangos activados (“*Activated Sludge Process*”) es la tecnología más utilizada para remover compuestos orgánicos y nutrientes de las aguas residuales, ya que tiene la capacidad de producir efluentes de alta calidad al menor costo y puede adaptarse a aguas residuales de cualquier tipo.

Existen múltiples razones para considerar el proceso de fangos activados como una aplicación interesante para el estudio de la metodología de diseño integrado. El tratamiento de aguas residuales es un proceso sometido a estrictas regulaciones ambientales ya que la calidad del efluente es primordial y cualquier inconveniente en la operación puede ocasionar importantes pérdidas en la producción y severos daños ambientales. Sin embargo, a pesar de su importancia, el tratamiento de aguas es una industria que no produce retribución económica, lo que la hace poco atractiva. Por esta razón, es fundamental diseñar plantas de tratamiento capaces de satisfacer las exigentes regulaciones ambientales al menor costo posible. Esto se logra con la aplicación de estrategias de control y operación eficientes, resolviendo el compromiso economía - desempeño dinámico mediante el *diseño integrado*.

El proceso de fangos activados constituye un fuerte desafío. Las tareas de control son difíciles debido a la complejidad y no linealidad de los sistemas biológicos presentes y al amplio rango de las constantes de tiempo de los procesos involucrados que pueden ser de minutos, días o semanas. Además, es normal que ocurran grandes perturbaciones en el caudal y en la composición del flujo de entrada a tratar (Pons et al, 1999).

En este capítulo se lleva a cabo el diseño integrado del proceso de fangos activados, aprovechando la posibilidad de evaluar distintas condiciones de operación y estrategias de control en escenarios distintos.

A continuación en la sección 2 se describe el proceso de fangos activados de manera general, se introduce el modelo matemático del proceso y se exponen las características de la operación de la planta a considerar. En las secciones 3 y 4, se presenta el *diseño integrado* del proceso para dos casos de estudio que difieren en el modelo final utilizado, donde además se plantean distintas estrategias de control, y diferentes criterios de *controlabilidad*.

2. EL PROCESO DE FANGOS ACTIVADOS

El objetivo principal de las plantas de tratamiento de agua es reducir la cantidad de materia orgánica y de nutrientes en el agua que se descarga al ambiente. En estas plantas se efectúan diversos tipos de tratamientos, físicos, químicos y/o biológicos para llevar el agua residual a las condiciones deseadas. El tratamiento biológico es el más importante, y se basa en el proceso de degradación de residuos que ocurre espontáneamente en la naturaleza, donde los microorganismos aprovechan las sustancias presentes en el agua para incorporarlas a su metabolismo celular y obtener energía para sus funciones vitales.

Entre los diferentes tipos de tratamientos biológicos desarrollados, el proceso de fangos activados es el más utilizado para aguas residuales urbanas. Existen distintas versiones del proceso. El esquema básico, diseñado para reducir la materia orgánica, consiste en uno o varios tanques de aireación (bio-reactores), seguidos por uno o varios tanques de sedimentación o decantadores secundarios, como se muestra en la figura 6.1.

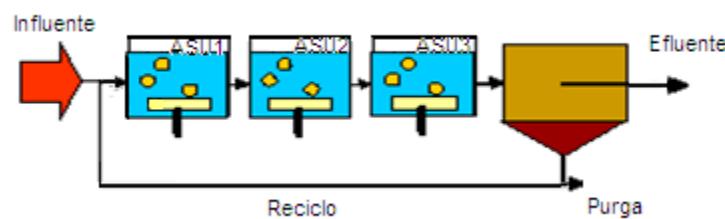


Figura 6.1. Esquema básico del proceso de fangos activados

En el bio-reactor se mezclan los fangos activados con el agua a tratar y tiene lugar el proceso de degradación biológica en condiciones aeróbicas, los microorganismos (o biomasa) satisfacen sus requerimientos de energía y como resultado se produce su crecimiento. Transcurrido el tiempo necesario para que se lleve a cabo la reacción, la mezcla pasa al decantador secundario o clarificador donde se separa el lodo del agua tratada. Parte de la biomasa se recircula hacia los tanques de aireación para mantener la cantidad de microorganismos en los bio-reactores en un alto nivel, mientras un pequeño porcentaje es removido como fango en exceso. Dependiendo de las características de la planta, el agua separada en el decantador continúa con su depuración o es descargada al ambiente. Después de este proceso el agua contiene aproximadamente 10% del material de desecho (Knobelsdorf, 2005; Gutierrez, 2000; Björn, 2003).

Este esquema básico del proceso ha ido modificándose a lo largo del tiempo para adaptarse a los requerimientos cada vez más exigentes en la depuración de las aguas. Es muy común la presencia en aguas residuales, junto a la materia orgánica, de nutrientes como compuestos nitrogenados y fósforo, y en la actualidad se exige la eliminación conjunta de ambos contaminantes (Fernández-Alba et al, 2006). Esto se consigue incluyendo en la planta zonas con aireación y sin aireación, creando una secuencia de zonas aerobias, anóxicas y anaerobias.

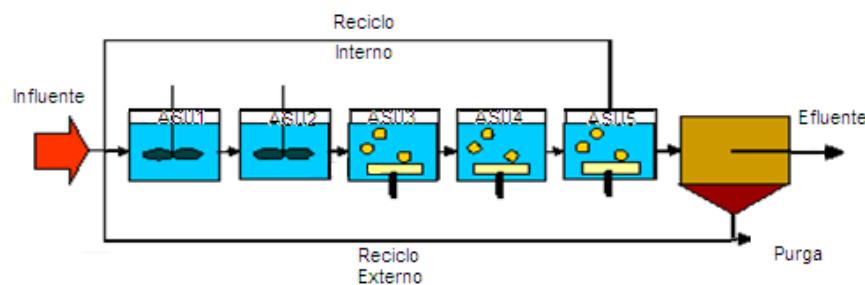


Figura 6.2. Esquema básico del proceso de fangos activados con nitrificación y desnitrificación

La eliminación biológica del nitrógeno, ocurre en dos procesos sucesivos, la nitrificación y la desnitrificación. Se inicia el proceso con una zona anóxica donde la materia orgánica del agua residual actúa como fuente de carbono en el proceso de desnitrificación, y el amoníaco es oxidado a nitritos y nitratos. Sigue una zona aireada donde se lleva a cabo el proceso de

nitrificación y se producen nitratos. (Knobelsdorf, 2005; Fernández-Alba et al, 2006). El agua en tratamiento debe contener suficiente carbono orgánico para llevar a cabo la desnitrificación. Usualmente es necesario recircular parte del efluente del último reactor de nitrificación. El esquema de este proceso se muestra en la figura 6.2.

2.1 MODELO DEL PROCESO DE FANGOS ACTIVADOS

En el marco del programa europeo “*COST ACTION 682 y 684*”, entre los años 1998 y 2004 se desarrolló la herramienta *BSMI* (“*Benchmark Simulation Model 1*”), que consiste en una planta de referencia del proceso de fangos activados para el estudio y comparación de diferentes estrategias de control aplicables a este proceso.

La estructura del BMS1 es relativamente simple, está compuesta por: 5 reactores biológicos, dos de los reactores son anaeróbicos y los otros tres son aeróbicos. Los dos primeros operan de forma anóxica para llevar a cabo el proceso de desnitrificación, mientras que los tres restantes se mantienen aireados y en ellos tiene lugar el proceso de nitrificación. Seguido de los reactores de lodos activados, se encuentra un decantador. El comportamiento de los bioreactores está representado por el Modelo del Proceso de Fangos Activados nro. 1 (*ASMI*) y el decantador o clarificador se modela usando una función de velocidad de sedimentación de Tacaks *et al* (1991). En la figura 6.3 se muestra una representación del proceso correspondiente al BSM1.

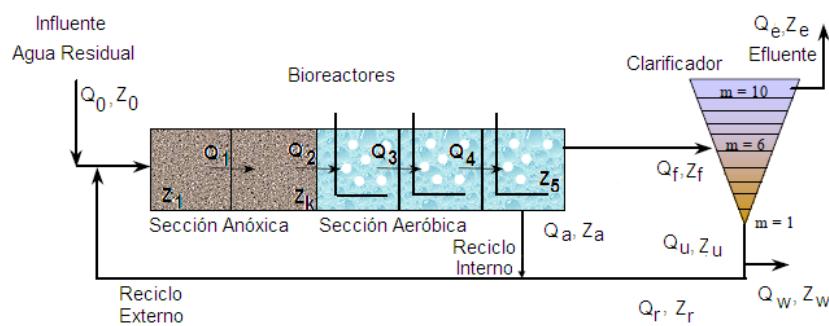


Figura 6.3. Esquema general del *BSMI*

El conjunto de ecuaciones del *BSMI* está dado por las relaciones entre los flujos, los balances de masa en cada reactor, el modelo de los bioreactores (*ASMI*) y el modelo del clarificador, como se explica a continuación:

a) *Balances entre flujos:*

De acuerdo a la figura 6.3, las relaciones entre los flujos en los puntos de mezcla o separación son:

$$Q_l = Q_a + Q_0 + Q_r \quad (6.1)$$

$$Q_f = Q_e + Q_u \quad (6.2)$$

$$Q_u = Q_r + Q_w \quad (6.3)$$

$$Q_k = Q_{k-1} \quad (6.4)$$

donde:

Q_0 : caudal de entrada

Q_a : reciclo interno o reciclo de nitrógeno

Q_r : reciclo externo

Q_e : caudal del efluente

Q_w : caudal de purga

Q_w : caudal a la salida del decantador

b) *Balances de masa:*

La forma general de los balances de masa en cada reactor es:

Para el primer reactor ($k=1$):

$$\frac{dZ_i(1)}{dt} = \frac{1}{V_1} [Q_0 \cdot Z_{ai} + Q_r Z_{ri} + Q_0 \cdot Z_{0i} - Q_l \cdot Z_i(1) + r_i(1) \cdot V_1] \quad (6.5)$$

En el reactor $k=2$ a 5:

$$\frac{dZ_i(k)}{dt} = \frac{1}{V_k} [Q_{k-1} \cdot Z_i(k-1) - Q_k \cdot Z_i(k) + r_i(k) \cdot V_k] \quad (6.6)$$

Para el caso especial del oxígeno ($Z=S_O$):

$$\frac{dS_O(k)}{dt} = \frac{1}{V_k} [Q_{k-1} \cdot S_O(k-1) - Q_k \cdot S_O(k) + r(k) \cdot V_k + K_{La}(k)(S_{O,Sat} - S_O(k))] \quad (6.7)$$

donde:

Z_i : concentración del compuesto i

Z_{ai} : concentración del compuesto i en el reciclo de nitrógeno

Z_{0i} : concentración del compuesto i en el influente

Z_{ri} : concentración del compuesto i en el reciclo externo

k : índice correspondiente al número de reactor

V_k : volumen de cada reactor

S_O : concentración de oxígeno

r_i : tasa de reacción del compuesto i .

El volumen de cada compartimiento en el BSM1 es:

$$V_1=V_2=1000\text{m}^3$$

$$V_3=V_4=V_5=1333\text{m}^3$$

c) Reactores biológicos:

Aunque luego se utilicen modelos simplificados del proceso, en el caso de los bio-reactores conviene presentar primero el modelo completo. En éste se definen los compuestos a considerar como variables de estado, su tasa de reacción (r_i) y los procesos biológicos involucrados, los cuales se introducen en la expresión general para los balances de masa dados por las ecuaciones (6.5 a 6.7).

Se escoge el Modelo del Proceso de Fangos Activados nro. 1 (“Activated Sludge Model No 1-ASM1”) desarrollado por la IAWPCR (“International Association on Water Pollution Research and Control”) para representar el proceso en los reactores biológicos, ya que cuenta con la aceptación general de la comunidad de investigación y la industria de plantas de

tratamiento de aguas residuales. Este modelo es la base del ASM2 que considera además la remoción biológica de fósforo y el ASM3 donde se reconoce la importancia del almacenamiento de polímeros en la conversión heterótrofa.

Tabla 6.1. Variables de estado descritas por el ASM1

Descripción	Variable
Concentración de materia orgánica inerte soluble (gr COD/m ³)	S_I
Concentración de sustrato fácilmente biodegradable (gr COD/m ³)	S_S
Concentración de materia orgánica inerte insoluble (gr COD/m ³)	X_I
Concentración de sustrato de degradación lenta (gr COD/m ³)	X_S
Concentración de biomasa heterótrofa activa (gr COD/m ³)	$X_{B,H}$
Concentración de biomasa autótrofa activa (gr COD/m ³)	$X_{B,A}$
Concentración de productos insolubles que provienen de decaimiento de biomasa (gr COD/m ³)	X_P
Concentración de oxígeno disuelto (gr/m ³)	S_O
Concentración de nitrógeno en forma de nitratos y nitritos (gr N/m ³)	S_{NO}
Concentración de nitrógeno en forma de $\text{NH}_4^+ + \text{NH}_3$ (gr N/m ³)	S_{NH}
Concentración de nitrógeno orgánico soluble biodegradable (gr N/m ³)	S_{ND}
Concentración de nitrógeno orgánico insoluble biodegradable (gr COD/m ³)	X_{ND}
Alcalinidad (gr COD/m ³)	S_{alk}

En el ASM1 se consideran las trece variables de estado que se presentan en la tabla 6.1. y ocho procesos biológicos (ρ_j) que se explicarán más adelante. Los parámetros y constantes cinéticas se presentan en la tabla 6.2.

Las tasas de reacción observadas para cada variable, en función de los procesos biológicos básicos contemplados en el modelo, se describen a continuación.

- Materia orgánica inerte soluble (S_I):

$$r_I(k) = 0 \quad (6.8)$$

- Sustrato fácilmente biodegradable (S_S):

$$r_2(k) = -\frac{1}{Y_H} \rho_1(k) - \frac{1}{Y_H} \rho_2(k) + \rho_7(k) \quad (6.9)$$

- Materia orgánica inerte insoluble (X_I):

$$r_3(k) = 0 \quad (6.10)$$

- Sustrato de degradación lenta (X_S):

$$r_4(k) = (1 - f_p) \cdot \rho_4(k) + (1 - f_p) \cdot \rho_5(k) - \rho_7(k) \quad (6.11)$$

- Biomasa heterótrofa activa ($X_{B,H}$):

$$r_5(k) = \rho_1(k) + \rho_2(k) - \rho_4(k) \quad (6.12)$$

- Biomasa autótrofa activa ($X_{B,A}$):

$$r_6(k) = \rho_3(k) - \rho_5(k) \quad (6.13)$$

- Productos insolubles de la biomasa muerta (X_P):

$$r_7(k) = f_p \cdot \rho_4(k) + f_p \cdot \rho_5(k) \quad (6.14)$$

- Oxígeno (S_O):

$$r_8(k) = -\frac{1 - Y_H}{Y_H} \rho_1(k) - \frac{4.57 - Y_A}{Y_A} \rho_3(k) \quad (6.15)$$

- Nitratos y nitritos (S_{NO}):

$$r_9(k) = -\frac{1 - Y_H}{2.86 \cdot Y_H} \rho_2(k) + \frac{1}{Y_A} \rho_3(k) \quad (6.16)$$

- Amonio (S_{NH}):

$$r_{10} = -i_{XB} \cdot \rho_1(k) - i_{XB} \cdot \rho_2(k) - \left(i_{XB} + \frac{1}{Y_A} \right) \rho_3(k) + \rho_6(k) \quad (6.17)$$

- Nitrógeno orgánico soluble (S_{ND}):

$$r_{11}(k) = -\rho_6(k) + \rho_8(k) \quad (6.18)$$

- Nitrógeno orgánico insoluble (X_{ND}):

$$r_{12}(k) = (i_{XB} - f_p \cdot i_{XP}) \cdot \rho_4(k) + (i_{XB} - f_p \cdot i_{XP}) \cdot \rho_5(k) - \rho_8(k) \quad (6.19)$$

- Alcalinidad ($Salk$):

$$r_{13}(k) = -\frac{i_{XB}}{14} \cdot \rho_1(k) + \left(\frac{1 - Y_H}{14 \cdot 2.86 \cdot Y_H} - \frac{i_{XB}}{14} \right) \cdot \rho_2(k) - \left(\frac{i_{XB}}{14} + \frac{1}{7 \cdot Y_A} \right) \cdot \rho_3(k) + \frac{1}{14} \rho_6(k) \quad (6.20)$$

Tabla 6.2. Valores de las constantes biológicas del *ASMI*.

Parámetro	Unidad	Valor
Y_A	gr COD formado/(gr N oxidado)	0.24
Y_H	gr COD formado/(gr COD oxidado)	0.67
f_p	adimensional	0.08
i_{XB}	gr N/(gr COD) en la biomasa	0.08
i_{XP}	gr N/(gr COD) en los productos insolubles	0.06
μ_H	1/d	4.0
μ_A	1/d	0.5
K_s	gr COD/m ³	10.0
$K_{O,H}$	gr (COD)/m ³	0.2
$K_{O,A}$	gr (COD)/m ³	0.4
K_{NO}	gr NO ₃ -N.m	0.5
K_{NH}	gr NH ₃ -N.m	1.0
K_X	gr COD lentamente biodegradable/(grCOD)	0.1
b_H	d ⁻¹	0.3
b_A	d ⁻¹	0.05
η_g	adimensional	0.8
η_h	adimensional	0.8
k_h	gr COD lentamente biodegradable/(grCOD.d)	3.0
k_a	m ³ /(grCOD.d)	0.05

Los ocho procesos biológicos que se llevan a cabo en los reactores de acuerdo al modelo son:

- ρ_1 crecimiento aeróbico de la biomasa heterótrofa

$$\rho_1(k) = \mu_H \cdot \left(\frac{Ss(k)}{K_s + Ss(k)} \right) \cdot \left(\frac{S_o(k)}{K_{O,H} + S_o(k)} \right) X_{B,H}(k) \quad (6.21)$$

- ρ_2 crecimiento anóxico de la biomasa heterótrofa

$$\rho_2(k) = \mu_H \cdot \left(\frac{Ss(k)}{K_s + Ss(k)} \right) \cdot \left(\frac{K_{O,H}}{K_{O,H} + S_o(k)} \right) \left(\frac{S_{NO}(k)}{K_{NO} + S_{NO}(k)} \right) n_g X_{B,H}(k) \quad (6.22)$$

- ρ_3 crecimiento aeróbico de la biomasa autótrofa

$$\rho_3(k) = \mu_A \cdot \left(\frac{S_{NH}(k)}{K_{NH} + S_{NH}(k)} \right) \cdot \left(\frac{S_o(k)}{K_{O,A} + S_o(k)} \right) X_{B,A}(k) \quad (6.23)$$

- ρ_4 muerte de la biomasa heterótrofa

$$\rho_4(2) = b_H \cdot X_{B,H}(k) \quad (6.24)$$

- ρ_5 muerte de la biomasa autótrofa

$$\rho_5(2) = b_A \cdot X_{B,A}(k) \quad (6.25)$$

- ρ_6 amonificación del nitrógeno orgánico soluble

$$\rho_6(k) = k_a \cdot S_{ND}(k) \cdot X_{B,H}(k) \quad (6.26)$$

- ρ_7 hidrólisis de los compuestos orgánicos

$$\rho_7(k) = k_h \cdot \left(\frac{\frac{X_s(k)}{X_{B,H}(k)}}{K_X + \frac{X_s(k)}{X_{B,H}(k)}} \right) \cdot \left[\left(\frac{S_o(k)}{K_{O,H} + S_o(k)} \right) + \eta_h \left(\frac{K_{O,H}}{K_{O,H} + S_o(k)} \right) \left(\frac{S_{NO}(k)}{K_{NO} + S_{NO}(k)} \right) \right] X_{B,H}(k) \quad (6.27)$$

- ρ_8 hidrólisis de los compuesto inorgánicos

$$\rho_8(k) = k_h \cdot \left(\frac{\frac{X_s(k)}{X_{B,H}(k)}}{K_X + \frac{X_s(k)}{X_{B,H}(k)}} \right) \cdot \left[\left(\frac{S_O(k)}{K_{O,H} + S_O(k)} \right) + \eta_h \left(\frac{K_{O,H}}{K_{O,H} + S_O(k)} \right) \left(\frac{S_{NO}(k)}{K_{NO} + S_{NO}(k)} \right) \right] X_{B,H}(k) \left(\frac{X_{ND}(k)}{X_S(k)} \right) \quad (6.28)$$

d) Clarificador

En la plataforma *BSMI* la dinámica del decantador se representa con el modelo de Tacaks *et al.* (1991) suponiendo una unidad no reactiva de diez capas.

En los casos de estudio que se tratan en este capítulo, no se utilizan las ecuaciones del *BSMI* para el clarificador sino que se limita el modelo a una unidad de tres capas, utilizando el modelo presentado por Moreno *et al.* (1992), como se resume a continuación.

En la primera capa del decantador:

$$A \cdot l_d \frac{dX_d}{dt} = Q_e \cdot X_b - Q_e \cdot X_d - A \cdot vs(X_d) \quad (6.29)$$

donde:

l_d altura de la capa superior,

A es el área del decantador,

X_d concentración de biomasa en la capa superior,

Q_e caudal de agua limpia a la salida del decantador,

X_b es la concentración de biomasa en la capa intermedia

$vs(X_d)$ es la velocidad másica de sedimentación del lodo dependiente de la concentración de biomasa en la capa actual.

En la segunda capa:

$$A \cdot l_b \frac{dX_b}{dt} = Q_f \cdot X_f - Q_e \cdot X_b - Q_u \cdot X_b + A \cdot vs(X_d) - A \cdot vs(X_b) \quad (6.30)$$

donde:

l_b es la altura de la capa intermedia,

X_f es la concentración de biomasa que entra al decantador,
 Q_u es el flujo de lodo activo a la salida del decantador.

En tercera capa del decantador:

$$A \cdot l_r \frac{dX_r}{dt} = Q_u \cdot X_b - Q_u \cdot X_r + A \cdot vs(X_b) \quad (6.31)$$

con:

l_r la altura de la última capa,

X_b concentración de biomasa en la capa inferior.

Para cada capa del clarificador, la velocidad de sedimentación está dada por la siguiente expresión, cuyos parámetros (nnr y aar) se determinan por ajuste a datos experimentales:

$$vs(Z_j) = nnr \cdot Z_j \cdot e^{(-aar \cdot Z_j)} \quad (6.32)$$

Finalmente, el decantador también puede representarse, de manera muy simplificada como se hace en algunas aplicaciones, como un punto de mezcla:

$$\frac{dX_{B,rec}}{dt} = \frac{1}{V_{dec}} [Q_f \cdot (X_{B,A} + X_{B,H}) - (Q_a + Q_w) X_{B,rec}] \quad (6.33)$$

donde:

V_{dec} es el volumen del decantador

$X_{B,rec}$ es la biomasa total recirculada

2.2 OPERACIÓN DEL PROCESO DE FANGOS ACTIVADOS

El funcionamiento eficiente del proceso de fangos activados depende de las siguientes restricciones de operación (Fernández-Alba et al, 2006; Mulas, 2006):

- La *concentración de oxígeno disuelto*. Se utilizan aireadores superficiales o difusores para suministrar a los microorganismos la cantidad de oxígeno requerida para oxidar la materia orgánica y mantener la calidad de los fangos. Si se suministra oxígeno en exceso, además del alto consumo de energía, puede afectarse la calidad de los fangos.

En la práctica la concentración de oxígeno en los tanques de aireación debe mantenerse alrededor de 2gr/m³. En la zona anóxica sólo se necesita una cantidad mínima de oxígeno para satisfacer los requerimientos de mezclado.

- Para evitar un excesivo consumo de aire en la zona aeróbica, la *concentración de nitratos* en la zona anóxica debe mantenerse entre 1 y 3 grN/m³ en los sistemas con reciclo de nitrógeno.
- La *edad de los fangos* es un parámetro que mide el tiempo medio que permanecen los fangos en el decantador. Esta magnitud suele tener un valor de 3-15 días en condiciones convencionales de operación.
- La *carga másica* conocida también como relación *A/M*. Se refiere a la relación entre la masa de materia orgánica alimentada al bio-reactor por unidad de tiempo y la cantidad de microorganismos (biomasa) disponibles en el sistema.
- El *tiempo de retención celular* es el tiempo de permanencia del agua a tratar en los reactores para que se efectúen las reacciones, el cual debe estar entre 2 y 5 horas aproximadamente.
- Las *relaciones entre los caudales* del proceso que dependen de la legislación ambiental en particular.

Los límites para estas restricciones se presentan en la tabla 6.3.

Tabla 6.3. Límites sobre las restricciones de operación

Variable	Valor
Oxígeno disuelto sección aeróbica (DO^n)	$1.5\text{gr}/\text{m}^3 \leq DO^n \leq 4\text{gr}/\text{m}^3$
Oxígeno disuelto sección anóxica (DO^d)	$0.05\text{gr}/\text{m}^3 \leq DO^d \leq 0.5\text{gr}/\text{m}^3$
Nitratos y nitritos en la sección anóxica ((S_{NO}^d))	$1\text{ grNO}/\text{m}^3 \leq S_{NO}^d \leq 3\text{ grNO}/\text{m}^3$
Edad de los fangos (SA)	$3\text{d} \leq SA \leq 15\text{d}$
Carga másica (A/M)	$0.05\text{ grCOD/grSS} \leq A/M \leq 1\text{ grCOD/grSS}$
Tiempo de retención celular	$2\text{hr} \leq TRC \leq 5\text{hr}$

La concentración de materia orgánica en el reactor afecta el proceso de desnitrificación. Por esto, en algunas ocasiones, se utiliza una fuente de carbón externa ($S_{dossage}$) para garantizar que la cantidad de carbono que entra al sistema es suficiente para alimentar a las bacterias durante el proceso de desnitrificación.

La *calidad del agua* a ser descargada es el objetivo de la operación y determina la eficiencia del proceso. Se mide en función de los parámetros característicos del efluente, que deben mantenerse dentro de ciertos límites que dependen de la legislación regional. En la tabla 6.4 se presentan algunos valores sugeridos en Alex et al. (2008) y Mulas (2006).

Las variables de estado definidas en el *ASMI* son las variables más importantes que intervienen en el proceso, pero no siempre pueden ser medidas en aplicaciones prácticas. Por esto, los parámetros relacionados con la calidad del efluente y otras características del proceso se expresan en función de ciertas variables compuestas como la *demandा química de oxígeno (COD)*, la *demandা biológica de oxígeno (BOD)*, el *nitrógeno total (Nt)* y la producción total de fangos (*TSS*) (Mulas, 2006):

$$COD = (S_I + S_S + X_I + X_S + X_{B,A} + X_{B,H} + X_P) \text{ grCOD/m}^3 \quad (6.34)$$

$$BOD = 0.25 \cdot ((S_S + X_S) + (1 - 0.08)(X_{B,A} + X_{B,H})) \text{ grCOD/m}^3 \quad (6.35)$$

$$TSS = 0.75 \cdot (X_S + X_I + X_P) + 0.9 \cdot (X_{B,A} + X_{B,H}) \text{ grSS/m}^3 \quad (6.36)$$

$$Nt = S_{NO} + S_{NH} + S_{ND} + X_{ND} + i_{XB} \cdot (X_{B,H} + X_{B,A}) + i_{XP} \cdot (X_P + X_I) \text{ grN/m}^3 \quad (6.37)$$

Tabla 6.4. Límites sobre la calidad del efluente

Variable	Valor
<i>Nitrógeno total (Nt)</i>	<18 grN/m ³
<i>Demandа Química de Oxígeno (COD)</i>	<125 grCOD/m ³
<i>Amonio (S_{NH})</i>	<4 grN/m ³
<i>Nitratos y nitritos (S_{NO})</i>	<10 grN/m ³
<i>Producción total de fangos(TSS)</i>	<30grSS/m ³

Parte del atractivo del proceso de fangos activados y específicamente de la plataforma *BSM1* es la posibilidad de probar diferentes estrategias de control en un sistema caracterizado por un comportamiento dinámico complejo.

Entre los objetivos de control se plantea, satisfacer los requerimientos sobre la calidad del efluente, rechazar las perturbaciones que afectan el correcto funcionamiento del proceso y lograr una operación eficiente al menor costo.

Una de las principales razones por las que se necesitan estrategias eficientes de control en este proceso, es el hecho de que está sometido a fuertes perturbaciones en el caudal y la concentración de materia orgánica y compuestos nitrogenados en el flujo de entrada. Como resultado, es difícil lograr la operación en estado estacionario, por lo que la planta se mantiene en estado transitorio la mayor parte del tiempo.

En el *BSM1* (Alex et al, 2008) se dispone de perfiles de perturbaciones característicos en distintas condiciones como clima seco, clima lluvioso y tormenta. En los casos de estudio se consideran el conjunto de datos utilizado para representar las perturbaciones en la concentración de sustrato y caudal de entrada para clima seco, los cuales se muestran en las figuras 6.4 y 6.5.

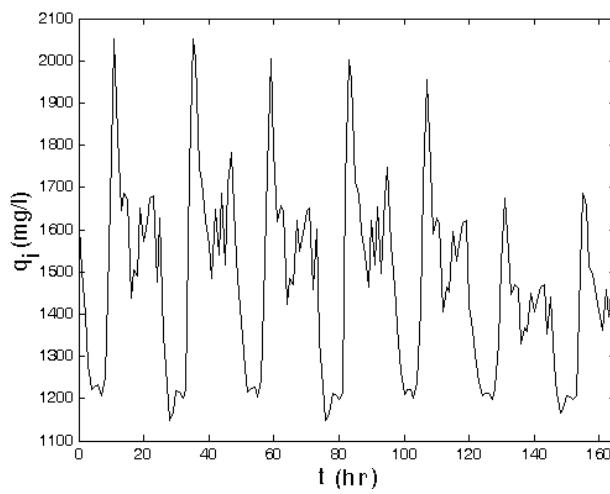


Figura 6.4. Perfil de perturbaciones en el caudal del influente en condiciones de clima seco
(Alex et al, 2008)

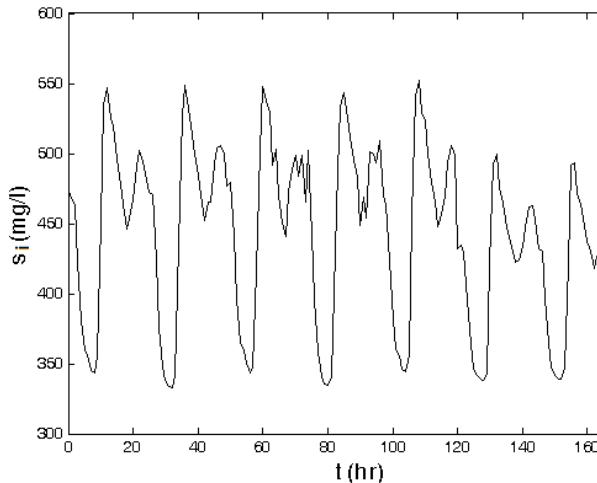


Figura 6.5. Perfil de perturbaciones en la concentración de sustrato en el influente en condiciones de clima seco (Alex et al, 2008)

En la literatura se han propuesto distintas estrategias de control para el proceso de fangos activados. Es importante aclarar que algunas de estas estrategias son idealizaciones, ya que en la práctica no es posible implementarlas debido a la dificultad para medir y/o manipular las variables involucradas.

En Inglidsen (2002) se propone un esquema en cascada para controlar la concentración de amonio (S_{NH}) en el efluente, el lazo externo afecta la concentración de oxígeno disuelto (S_O) en el reactor aeróbico, que se regula a su vez manipulando el K_{La} con el lazo interno. También proponen controlar el nitrógeno total manipulando el caudal de recirculación interno (Q_a).

En el *BSMI*, Alex et al (2008), proponen dos lazos de control, uno para regular la concentración de nitratos S_{NO} manipulando el caudal de recirculación interno (Q_a) y otro para controlar la concentración de oxígeno disuelto (S_O) manipulando el K_{La} . Para minimizar el efecto de las perturbaciones en el caudal del influente se puede introducir un esquema por alimentación adelantada (“*feedforward*”) que permita variar la concentración de oxígeno disuelto (S_O) cuando se presenten perturbaciones en el caudal de entrada.

En Gutiérrez (2000) se controla la concentración de sustrato manipulando el caudal de recirculación externa (recirculación de fangos). Respecto a esta estrategia, en Mulas (2006) se señala que es una práctica común mantener una relación fija entre el reciclo externo y el flujo de entrada a la planta permitiendo sólo un pequeño margen de variación para mantener el

balance de masa total de fangos. Se recomienda que la relación Q_r/Q_0 esté comprendida entre 1.5 y 0.75.

El desempeño dinámico del proceso de fangos activados está directamente relacionado con los costos de operación. En la plataforma Benchmark se presentan algunos índices de desempeño, para un periodo de operación T , que permiten cuantificar el gasto de energía del proceso, como la energía de bombeo y la energía de aireación. Este consumo eléctrico es importante dentro de los costos de operación del proceso y se puede incluir mediante un factor de costo.

La energía de bombeo (PE) depende los flujos de recirculación y de purga en el sistema, los cuales pueden variar debido a las acciones de control:

$$PE = \frac{1}{T} \int_{t_0}^{t_f} (0.004 \cdot Q_a(t) + 0.008 \cdot Q_w(t) + 0.05 \cdot Q_w(t)) dt \quad [kWh/d] \quad (6.38)$$

con los caudales expresados en m^3/d .

La energía de aireación (EA) depende del equipo utilizado, la ubicación de los difusores y el tamaño de las burbujas. En el *BSMI* se presenta la siguiente relación obtenida para un equipo particular.

$$AE = \frac{S_{O,sat}}{T \cdot 1.8 \cdot 1000} \int_{t_0}^{t_f} \sum_{k=1}^{k=5} V_k \cdot K_{Lak}(t) dt \quad [kWh/d] \quad (6.39)$$

con K_{La} expresado en $1/días$.

La calidad del efluente (EQ) es la suma ponderada de las cargas en el efluente de los compuestos con mayor efecto sobre la calidad del agua. Este índice se usa para determinar la penalización económica por las descargas fuera de especificación, y está dado por:

$$EQ = \frac{1}{T \cdot 1000} \int_{t_0}^{t_f(\text{days})} [2 \cdot SS_e + COD_e + 30 \cdot Nt + 10 \cdot S_{NO,e} + 2 \cdot BOD_e] Q_e dt \quad [Kg \text{ polución}/d] \quad (6.40)$$

con las concentraciones expresadas en gr/m^3 y el caudal en m^3/d .

Partiendo de la descripción general del proceso de fangos activados, su modelo detallado y las características de su operación, se propone a continuación el planteamiento del problema de diseño integrado para dos aplicaciones particulares para las cuales se utilizan modelos distintos del proceso:

- **Modelo 1.** Se considera el diseño del proceso enfocado únicamente en la reducción de materia orgánica en el efluente, por lo que se descartan los procesos de nitrificación y desnitrificación y los relacionados con los compuestos insolubles. Se considera la síntesis del proceso y se evalúan esquemas de control PI y esquemas avanzados en distintos escenarios que dependen de los criterios de *controlabilidad* utilizados. Este ejemplo también es utilizado para evaluar el desempeño del algoritmo genético en comparación a otros métodos estocásticos y deterministas de optimización.
- **Modelo 2.** Se considera el *BSMI* simplificado con una estructura fija de un reactor aeróbico, uno anóxico y el decantador. Este último representado por simplicidad como un punto de mezcla, con recirculación de fangos constante. Se lleva a cabo la remoción de materia orgánica (sustrato) y la eliminación de nitrógeno, pero sólo se toman en cuenta los procesos relacionados con los compuestos solubles cuya dinámica es significativa en períodos de 14 días, como se propone en Ingildsen (2002).

La formulación particular y los resultados de la aplicación de la metodología de *Diseño Integrado* se presentan a continuación.

3. DISEÑO INTEGRADO DEL PROCESO DE FANGOS ACTIVADOS. MODELO 1.

El interés en este caso es fundamentalmente ilustrar la metodología de *diseño integrado* contemplando la síntesis de la planta y la inclusión de un esquema de control avanzado. Por esta razón, se considera una simplificación del modelo del proceso donde no tienen lugar la nitrificación, desnitrificación ni hidrólisis de la materia orgánica, de modo que la planta consta únicamente de reactores aeróbicos, el clarificador y el reciclo externo.

El modelo del proceso para este caso de estudio es una variante simplificada del *BSMI*, desarrollado específicamente para la planta de tratamiento de aguas de Manresa (España) en Moreno et al (1992) y constituido por las siguientes ecuaciones para los reactores:

$$\frac{dx}{dt} = \mu_{\max} y \frac{sx}{(K_s + s)} - K_d \frac{x^2}{s} - K_c x + \frac{1}{v_l} (x_i \cdot q_i + x_r \cdot q_r - q \cdot x) \quad (6.41)$$

$$\frac{ds}{dt} = -\mu_{\max} \frac{sx}{(K_s + s)} + f_{kd} K_d \frac{x^2}{s} + F_{kd} K_c x + \frac{1}{v_l} (s_i \cdot q_i + s \cdot q_r - q \cdot s) \quad (6.42)$$

$$\frac{dc}{dt} = K_{La} F k_1 (c_s - c) - K_{01} \mu_{\max} \frac{x \cdot s}{(K_s + s)} - \frac{q}{v_l} c \quad (6.43)$$

donde:

x (mg/l): biomasa heterótrofa ($X_{B,H}$)

s (mg/l): sustrato soluble fácilmente biodegradable (S_S)

c (mg/l): es la notación utilizada para el oxígeno disuelto (S_O)

v_l (m^3): es el volumen del bio-reactor

Fk_1 es un factor de aireación (adimensional) proporcional a la velocidad de las turbinas. En este modelo el K_{La} permanece constante y la variación en la transferencia de oxígeno se relaciona con este factor.

En el decantador, las ecuaciones 6.29 a 6.31 se desarrollan para esta planta específica, obteniéndose:

$$A \cdot l_d \frac{dx_d}{dt} = q_{sal} x_b - q_{sal} x_d - A \cdot vs(x_d) \quad (6.44)$$

$$A \cdot l_b \frac{dx_b}{dt} = q x_l - q_{sal} x_b - q_2 x_b + A \cdot vs(x_d) - A \cdot vs(x_b) \quad (6.45)$$

$$A \cdot l_r \frac{dx_r}{dt} = q_2 x_b - q_2 x_r + A \cdot vs(x_b) \quad (6.46)$$

donde:

x_d (mg/l): concentración de biomasa en la primera capa del decantador

x_b (mg/l): concentración de biomasa en la segunda capa del decantador

x_r (mg/l): concentración de biomasa en la tercera capa del decantador y en la recirculación

A (m^2): área transversal del decantador

La notación utilizada para los parámetros restantes y sus valores correspondientes se presentan en la tabla 6.5.

Tabla 6.5. Parámetros biológicos, físicos y operacionales para el proceso de fangos activados seleccionado

Notación	Parametro	Valor
μ_{max}	Velocidad específica de crecimiento máximo	0.1824
y	Fracción de sustrato metabolizado que se convierte en biomasa	0.5948
fkd	Fracción de biomasa muerta que se transforma en sustrato	0.2
K_d	Constante de actividad celular de los microorganismos	$5.5\text{e}^{-5} \text{ h}^{-1}$
K_s	Coeficiente de saturación de sustrato	300
K_c	Constante de mortandad de la biomasa	$1.333\text{e}^{-4} \text{ h}^{-1}$
K_{La}	Coeficiente de masa global para la transferencia de oxígeno	0.7 h^{-1}
K_{O1}	Coeficiente de relación entre el crecimiento celular y la tasa de consumo de oxígeno	0.0001
c_s	Concentración de saturación de oxígeno en los bioreactores	8 mg/l
xi	Concentración de biomasa en el influente	80 mg/l
si	Concentración de sustrato en el influente	366.67 mg/l
qi	Flujo de entrada	$1300 \text{ m}^3/\text{h}$
q	Flujo de entrada al bioreactor	m^3/h
qr	Reciclo	m^3/h
vs	Velocidad de sedimentación de los fangos en el decantador	
nnr	Parámetro empírico para el ajuste de la velocidad de sedimentación	3.1563
aar	Parámetro empírico para el ajuste de la velocidad de sedimentación	-0.000785
ld	Altura de la primera capa del decantador	2m
lb	Altura de la segunda capa del decantador	1m
lr	Altura de la tercera capa del decantador	0.5m

3.1. FORMULACION GENERAL DEL PROBLEMA DE OPTIMIZACION

En el planteamiento del problema de *diseño integrado* se considera la posibilidad de seleccionar la estructura de la planta entre uno o dos bio-reactores seguidos por un decantador para introducir la *síntesis del proceso*. Ambas alternativas se representan en la *superestructura* que se muestra en la figura 6.4.

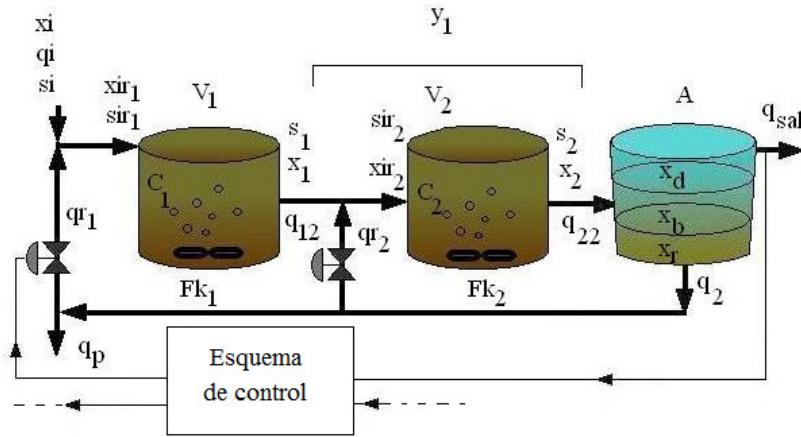


Figura 6.4. Superestructura del proceso de fangos activados para eliminación de materia orgánica

A partir de la superestructura y haciendo uso del modelo, se derivan una serie de ecuaciones diferenciales y algebraicas que, junto con la función objetivo y el esquema de control, se traducen en un problema de *optimización mixto-entero con restricciones no lineales y requerimientos sobre índices de desempeño dinámicos*. El conjunto de variables de decisión está dado por: la estructura del proceso definida por la variable binaria (y_1), las dimensiones de la planta, el punto de trabajo y los parámetros del sistema de control. El objetivo es minimizar una función general, que incluye costos de inversión y de operación, respecto a las variables de decisión.

En el problema se plantean restricciones de operación que surgen de los requerimientos impuestos sobre la carga másica, la edad de los fangos, la capacidad hidráulica y las relaciones entre los flujos. Además, se imponen restricciones sobre el modelo del proceso que garantizan un punto inicial de operación en estado estacionario.

El análisis de *controlabilidad* depende de la estrategia de control seleccionada y de los criterios establecidos para evaluar el desempeño dinámico y su efecto sobre las restricciones del proceso. Las medidas de *controlabilidad* se pueden introducir en el problema como objetivos de optimización o como restricciones.

Se presenta a continuación la formulación general del problema de *diseño integrado* sin especificar la técnica de control a utilizar ni las restricciones de *controlabilidad*, pues éstas se describen de manera individual en cada uno de los escenarios propuestos para abordar el problema de *diseño integrado*.

Función objetivo.

$$f = p_1 \cdot v_1^2 + p_1 \cdot v_2^2 + p_2 \cdot A + p_3 \cdot Fk_1^2 + p_3 \cdot Fk_2^2 + p_4 \cdot q_2^2 \quad (6.47)$$

donde:

v_1, v_2 son los volúmenes de los reactores 1 y 2 respectivamente,

Fk_1 y Fk_2 son los factores de aireación para cada reactor,

A es el área transversal del decantador

q_2 es el flujo total de reciclo,

p_1, p_2, p_3 son pesos que determinan la importancia relativa de cada factor.

Los tres primeros términos están asociados a los costos de construcción, luego aparecen los costos operacionales relacionados con la energía utilizada por las turbinas de aireación, y finalmente los costos de bombeo (purga y reciclo). Esta es una representación general y muy simple de la función objetivo para este caso de estudio específico. No representa costos reales sino un índice proporcional a éstos, por lo que es adimensional.

Condiciones lógicas del problema.

Es necesario imponer condiciones lógicas que garanticen la coherencia matemática del modelo para cada posible configuración de la planta. Dependiendo de la existencia o no del segundo reactor (representada por la variable binaria y_1), se establece lo siguiente:

- Si el segundo reactor no existe $y_1=0$, entonces $v_2=0, x_2=x_1, s_2=s_1, c_2=c_1, Fk_2=0, qr_2=0$.

- Si el segundo reactor existe $y_I=1$, entonces v_2, x_2, s_2, Fk_2 y qr_2 toman valores adecuados dentro del rango correspondiente a cada variable.

Para ello se definen las variables de holgura W_1, W_2, W_3, W_4 , de manera de anular las ecuaciones y restricciones que se desactivan cuando el segundo reactor no existe ($y_I=0$), sin afectar las mismas cuando sí existe ($y_I=1$).

Restricciones sobre los balances de masa.

Se imponen restricciones sobre los balances de masa en los tanques de aireación y el decantador desarrollados para la *superestructura* del proceso presentada en la figura 6.4, con la finalidad de garantizar un punto de trabajo inicial en equilibrio.

$$\left| v_1 \frac{dx_1}{dt} \right| = \left| \mu_{\max} Y \frac{s_1 x_1}{(K_s + s_1)} v_1 - K_d \frac{x_1^2}{s_1} v_1 - K_c x_1 v_1 + (x_i \cdot q_i + x_r \cdot qr_1 - q_{12} \cdot x_1) \right| \leq \varepsilon \quad (6.48)$$

$$\left| v_1 \frac{ds_1}{dt} \right| = \left| -\mu_{\max} \frac{s_1 x_1}{(K_s + s_1)} v_1 + f_{kd} K_d \frac{x_1^2}{s_1} v_1 + f_{kd} K_c x_1 v_1 + (s_i \cdot q_i + s_1 \cdot qr_1 - q_{12} \cdot s_1) \right| \leq \varepsilon \quad (6.49)$$

$$\left| v_1 \frac{dc_1}{dt} \right| = \left| K_{la} F k_1 (c_s - c_1) v_1 - K_{01} \mu_{\max} \frac{s_1 x_1}{(K_s + s_1)} v_1 - q_{12} c_1 \right| \leq \varepsilon \quad (6.50)$$

$$\left| v_2 \frac{dx_2}{dt} \right| = \left| \mu_{\max} Y \frac{s_2 x_2}{(K_s + s_2)} v_2 - K_d \frac{x_2^2}{s_2} v_2 - K_c x_2 v_2 + (x_1 \cdot q_{12} + x_r \cdot qr_2 - q_{22} \cdot x_2) \right| \leq \varepsilon \quad (6.51)$$

$$\left| v_2 \frac{ds_2}{dt} \right| = \left| -\mu_{\max} \frac{s_2 x_2}{(K_s + s_2)} v_2 + f_{kd} K_d \frac{x_2^2}{s_2} v_2 + f_{kd} K_c x_2 v_2 + (s_1 \cdot q_{12} + s_2 \cdot qr_2 - q_{22} \cdot s_2) \right| \leq \varepsilon \quad (6.52)$$

$$\left| v_2 \frac{dc_2}{dt} \right| = \left| K_{la} F k_2 (c_s - c_2) v_2 - K_{01} \mu_{\max} \frac{s_2 x_2}{(K_s + s_2)} v_2 - q_{22} c_2 + W_1 \right| \leq \varepsilon \quad (6.53)$$

$$\left| A L_d \frac{dx_d}{dt} \right| = \left| q_{sal} x_b - q_{sal} x_d - A \cdot nnr \cdot x_d \exp(aar \cdot x_d) \right| \leq \varepsilon \quad (6.54)$$

$$\left| AL_b \frac{dx_b}{dt} \right| = \left| q_{22}x_2 - q_{sal}x_b + A \cdot nnr \cdot x_d \exp(aar \cdot x_d) - A \cdot nnr \cdot x_b \exp(aar \cdot x_b) \right| \leq \varepsilon \quad (6.55)$$

$$\left| AL_r \frac{dx_r}{dt} \right| = \left| q_2x_b - q_2x_r + A \cdot nnr \cdot x_b \exp(aar \cdot x_b) \right| \leq \varepsilon \quad (6.56)$$

Las relaciones entre los caudales vienen dadas por:

$$q_{12} = q_i + qr_1 \quad (6.57)$$

$$q_{22} = q_{12} + qr_2 \quad (6.58)$$

$$q_2 = qr_1 + qr_2 + q_p \quad (6.59)$$

Para el caso de la estructura de la planta con un reactor y un decantador ($y_l=0$), las ecuaciones (6.51) a (6.53) se desactivan de acuerdo a las condiciones lógicas antes descritas. Cuando $y_l=0 \Rightarrow v_2=0$ y $qr_2=0$, entonces se anulan directamente los términos multiplicados por v_2 en todas las ecuaciones, los términos restantes en las ecuaciones (6.51) y (6.52) desaparecen al quedar $q_{22} = q_{12}$ en (6.58), finalmente, para anular el último término de la ecuación (6.54) se requiere de una variable de holgura (W_l):

$$W_l = (1 - y_1)q_{22} \cdot c_2 \quad (6.60)$$

Restricciones de operación.

Se establecen las siguientes restricciones en función de los límites entre los cuales se deben mantener algunos parámetros para garantizar la operación adecuada de la planta.

- Tiempo de residencia en los reactores:

$$2.5 \leq \frac{v_1}{q_{12}} \leq 8 \quad (6.61)$$

$$2 \leq \frac{v_2 + (1 - y_1) \cdot W_l}{q_{22}} \leq 6 \quad (6.62)$$

Esta última restricción se descompone en dos inecuaciones:

$$v_2 + (1 - y_1) \cdot W_2 - 6 \cdot q_{22} \leq 0 \quad (6.62a)$$

$$2 \cdot q_{22} - v_2 + (1 - y_1) \cdot W_2 \leq 0 \quad (6.62b)$$

donde W_2 se utiliza para ajustar la relación a la estructura actual de la planta siendo $W_2 = 6 \cdot q_{22}$ para (6.62a) y $W_2 = -2 \cdot q_{22}$ para (6.62b).

- Carga másica en los tanques de aireación (A/M):

$$0.001 \leq \frac{q_i s_i + qr_1 s_2}{v_1 x_1} \leq ML_{\max} \quad (6.63)$$

$$0.001 \leq \frac{q_{12} s_1 + qr_2 s_2 - (1 - y_1) W_3}{v_2 x_2} \leq ML_{\max} \quad (6.64)$$

donde:

$$W_3 = (Qr_1 + Q_i) \cdot s_1 \quad (6.65)$$

- Edad de los fangos en el decantador (SA):

$$SA_{\min} \leq \frac{v_1 x_1 + v_2 x_2 + AL_r x_r}{q_p x_r 24} \leq 10 \quad (6.66)$$

- Capacidad hidráulica:

$$\frac{q_{22}}{A} \leq 1.5 \quad (6.67)$$

- Relación entre el caudal de purga y el reciclo:

$$0.03 \leq \frac{q_p}{q_2} \leq Rp_{\max} \quad (6.68)$$

- Relación entre el reciclo y el caudal de entrada:

$$Rr_{\min} \leq \frac{q_2}{q_i} \leq 0.9 \quad (6.69)$$

Los límites ML_{max} , SA_{min} , Rp_{max} y Rr_{min} se usan para establecer rangos o condiciones de operación distintas. En este trabajo se proponen dos escenarios basados en distintos valores de estos límites (Tabla 6.5). El primero es el conjunto de valores presentado originalmente en Moreno et al (1992), el segundo es una relajación de los límites de las restricciones, presentada en Revollar et al. (2010).

Habiendo formulado matemáticamente la síntesis y el dimensionamiento del proceso, a continuación se realiza, con distintas estrategias, la incorporación del análisis de controlabilidad y la definición y ajuste en el procedimiento de diseño integrado. Primero se lleva a cabo la *síntesis del proceso* basada sólo en criterios económicos, luego el *diseño integrado* con un controlador PI y finalmente el *diseño integrado* con un esquema avanzado de control (Controlador predictivo-*MPC*). En cada caso se desarrolla, desde la formulación de las restricciones y objetivos de controlabilidad hasta la resolución del problema utilizando como herramienta los algoritmos genéticos.

Tabla 6.6. Límites de las restricciones de operación

Parámetros	Conjunto 1	Conjunto 2
ML_{max}	0.6	0.12
SA_{min}	3	2
Rp_{max}	0.07	0.7
Rr_{min}	0.5	0.05

3.2. CASOS DE ESTUDIO

3.2.1. Síntesis económica del proceso

La *síntesis económica* del proceso permite obtener una planta de referencia cuyas características están determinadas únicamente por criterios económicos. Además, es el escenario más apropiado para evaluar el desempeño de los métodos de optimización, ya que

no intervienen otros elementos, como la linealización del modelo o la simulación de la respuesta dinámica, que puedan afectar la eficiencia de los algoritmos de resolución.

Así, para llevar a cabo la síntesis económica del proceso, se resuelve el problema de optimización descrito por las ecuaciones (6.47) a la (6.69). En la resolución se evalúa el desempeño del algoritmo genético (“*Genetic Algorithm*”- *GA*) propuesto en este trabajo, junto con otro método estocástico de optimización como es el enfriamiento simulado (“*Simulated Annealing*”-*SA*), y un método determinista clásico como el de *ramificación y acotamiento* (“*Branch and Bound*”-*B&B*). Dentro del procedimiento de *B&B* se utiliza el algoritmo de programación secuencial cuadrática (“*Sequential Quadratic Programming*”-*SQP*) para resolver los problemas de optimización no lineal que surgen en los diferentes nodos. En Francisco et al. (2005), Dominguez (2007) y Revollar et al. (2010) se describe con mayor detalle cómo se emplean los algoritmos *B&B* y *SA* en la resolución de este problema de diseño integrado.

Algoritmo genético.

Se usa un algoritmo genético de codificación mixta como el descrito en el Capítulo IV. Se aplican los operadores de selección por ruleta y cruce aritmético, con una probabilidad de cruce de 85%, con la finalidad de combinar únicamente los genes que representan la misma variable. La mutación es *aleatoria* con tasa variable entre 0.1 y 0.02, para incrementar la exploración del espacio de búsqueda en las primeras generaciones.

El cromosoma contiene las variables de decisión del problema:

$$[x_{1n}, x_{2n}, s_{1n}, s_{2n}, c_{1n}, c_{2n}, x_{dn}, x_{bn}, x_{rn}, qr_{1ss}, qr_{2n}, qp_n, Fk_{1n}, Fk_{2n}, v_{1n}, v_{2n}, A_n, y_l]$$

donde el sub-índice n indica que se trata de una variable normalizada entre el valor máximo y mínimo de su rango factible. Para facilitar la aplicación de los operadores de cruce y mutación sobre el cromosoma mixto, las variables binarias ocupan las últimas posiciones (solo y_l en este caso).

En el algoritmo se permite hacer cruces entre cromosomas con valores diferentes de y_l , aunque representen estructuras con distinto número de bio-reactores. El resultado de la

aplicación del operador de cruce aritmético sobre la variable binaria se redondea a los valores binarios cero o uno. Luego, para que no se produzcan incoherencias numéricas en las soluciones candidatas producto del cruce, las restricciones lógicas se introducen directamente en el cromosoma. Como los valores mínimos de v_2 y qr_2 son cero, cuando $y_l=0 \Rightarrow v_2=0$ y $qr_2=0$, se reconstruye el cromosoma para que quede de la forma:

$$[x_{ln}, x_{In}, s_{ln}, c_{ln}, c_{In}, x_{dn}, x_{bn}, x_{rn}, qr_l, 0, qp_n, Fk_{ln}, 0, v_{ln}, 0, A_n, 0]$$

Para tratar las restricciones de operación se penaliza la función objetivo con un término proporcional a la desviación de las mismas multiplicado por cierto factor de penalización R (ec. 4.24).

Se genera una población inicial de 200 individuos, las variables continuas se introducen aleatoriamente, y las binarias se fijan de manera de tener igual cantidad de individuos para cada estructura del proceso. El número máximo de iteraciones es de 1500.

Resultados.

Se realizan diez experimentos con cada método, *B&B*, *GA* y *SA*, variando el punto inicial o población inicial, según corresponda. Se estudia la sensibilidad de los métodos estocásticos a la variación de sus parámetros ajustables, como el factor de penalización en la función de adaptación para el caso de los algoritmos genéticos y el parámetro de control c para el enfriamiento simulado.

Resolviendo el problema para el conjunto de restricciones más exigentes (conjunto 1 en la tabla 6.6) con el método determinista (*B&B*), se obtuvo una solución óptima con una función objetivo de 0.17 en todos los experimentos. Este valor se toma como referencia, y respecto a éste, los métodos estocásticos alcanzan soluciones con una desviación promedio de 5% en caso del *GA* y 7% con el *SA*.

Tabla 6.7. Parámetros de la planta obtenida como resultado de la *síntesis económica* y dimensionamiento del proceso de fangos activados

Parámetros	Conjunto de restricciones 1	Conjunto de restricciones 2
Costo	0.17	0.04
Numero de bio-reactores	1	2
$s_1 \text{ (mg/l)}$	87.75	196.67
$s_2 \text{ (mg/l)}$	-	92.43
$V_1 \text{ (m}^3\text{)}$	5278.8	3709.6
$V_2 \text{ (m}^3\text{)}$	-	3102.5
$A \text{ (m}^2\text{)}$	1787.9	1456.4
$qr_1 \text{ (m}^3/\text{h)}$	630.5	167.0
$qr_2 \text{ (m}^3/\text{h)}$	-	11.6
Fk_1	0.076	0.005
Fk_2	-	0.22

El resultado óptimo, correspondiente a la solución por el métodos de ramificación y acotamiento (*B&B*) se presenta en la tabla 6.7 donde se muestran los parámetros característicos de las plantas obtenidas al realizar la síntesis económica. Se puede observar que la variación en los límites de las restricciones afecta significativamente la estructura y dimensiones de la planta. Esto muestra que se obtienen alternativas estructurales distintas con grandes diferencias en los costos de construcción y operación, al imponer diferentes límites.

3.2.2. Diseño integrado con un controlador PI

Formulación.

La formulación del *diseño integrado* del proceso con un controlador PI se basa en la *superestructura* presentada en la figura 6.7. El diseño debe definir la estructura de la planta, las dimensiones y punto de operación óptimos económico, y que además se cumplan ciertas condiciones de desempeño dinámico, por lo que se integra el diseño de un controlador PI en el problema.

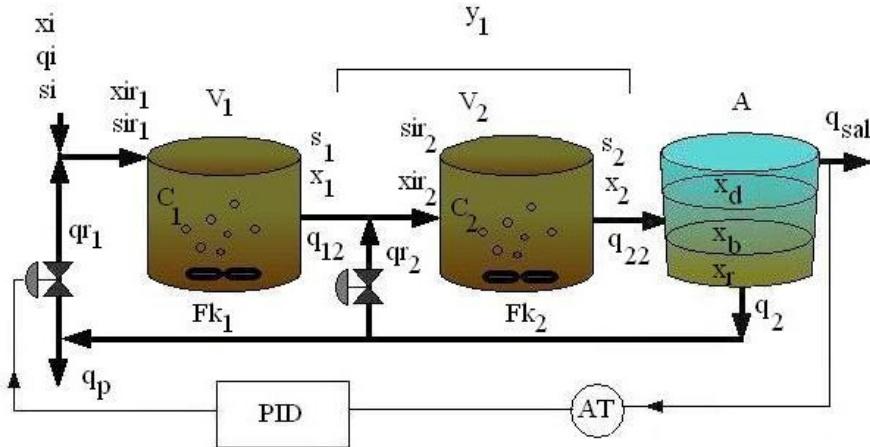


Figura 6.7. Superestructura del proceso de fangos activados para eliminación de materia orgánica con un PI

El objetivo de control es mantener la concentración de sustrato a la salida (s_1 o s_2) dentro de los límites permitidos a pesar de las fuertes perturbaciones a las que está sometido el proceso. Se trata de las perturbaciones características del *BSMI* sobre el caudal q_i y concentración de sustrato s_i descritas en las figuras 6.4 y 6.5.

Se considera la estrategia de control propuesta en Gutiérrez (2000) para esta planta, donde se usa el caudal de recirculación de fangos (qr_1) como variable manipulada. Entonces, la ley de control del PI para este caso es:

$$qr_1 = qr_{1ss} + K_p \cdot (s_{ref} - s_{(1+y_1)}) + \frac{1}{\tau_{i_to}} \int_{t_0}^t (s_{ref} - s_{(1+y_1)}) d\tau \quad (6.70)$$

donde:

s_{ref} es el valor deseado para la concentración de sustrato a la salida, que corresponde al punto de operación óptimo,

qr_{1ss} es el caudal de recirculación en estado estacionario, correspondiente al punto de equilibrio $s_{(1+y_1)} = s_{ref}$.

Para tener una medida del efecto de la estrategia de control sobre la calidad de la operación se introduce la integral del cuadrado del error (ISE) como índice de desempeño dinámico:

$$ISE = \int_{t=0}^{T_{\max}} \left(s_{ref} - s_{(t+y_1)} \right)^2 \cdot dt \quad (6.71)$$

donde:

T_{\max} es el tiempo de simulación

Al minimizar el ISE se minimiza la desviación de la concentración de sustrato a la salida respecto al punto de operación deseado en presencia de perturbaciones, por lo que se impone una cota superior al ISE como restricción de *controlabilidad*:

$$ISE \leq ISE_{\max} \quad (6.72)$$

La formulación del problema de diseño integrado para este caso particular se presenta de forma esquemática en la figura 6.8.

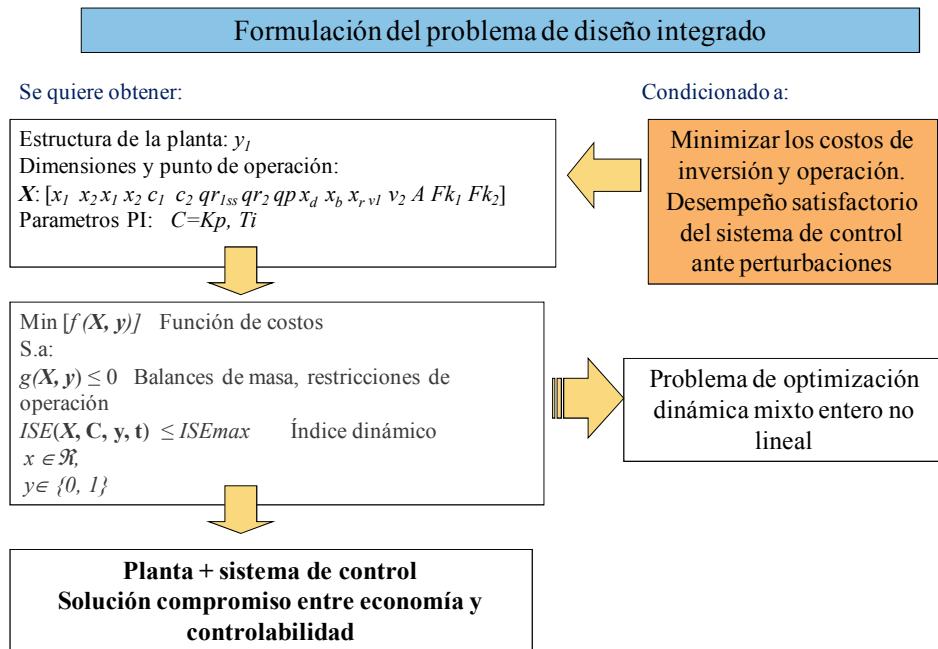


Figura 6.8. Formulación del problema de *diseño integrado* del proceso de fangos activados con un PI

Algoritmo genético.

Se implanta una versión del algoritmo muy similar a la desarrollada para el caso económico tratado anteriormente. Se usan los mismos operadores, técnica de reconstrucción y parámetros. El cromosoma varía porque se introducen los coeficientes del controlador en el genoma:

$$[x_{1n}, x_{2n}, s_{1n}, s_{2n}, c_{1n}, c_{2n}, x_{dn}, x_{bn}, x_{rn}, qr_{1ss}, qr_{2n}, qp_n, Fk_{1n}, Fk_{2n}, v_{1n}, v_{2n}, A_n, Kp_n, Ti_n, y_l]$$

Se aplican además algunas técnicas heurísticas para mejorar el rendimiento del algoritmo genético. Una de ellas garantiza que las soluciones obtenidas sean estables, aunque este requerimiento no está realmente contemplado en la formulación del problema. Consiste en impedir la simulación de las soluciones candidatas que presenten autovalores positivos de la matriz de estados, para evitar problemas numéricos. Cuando esto ocurre se asigna arbitrariamente un valor grande al *ISE* (cien veces mayor que el ISE_{max}) para penalizar fuertemente estos cromosomas.

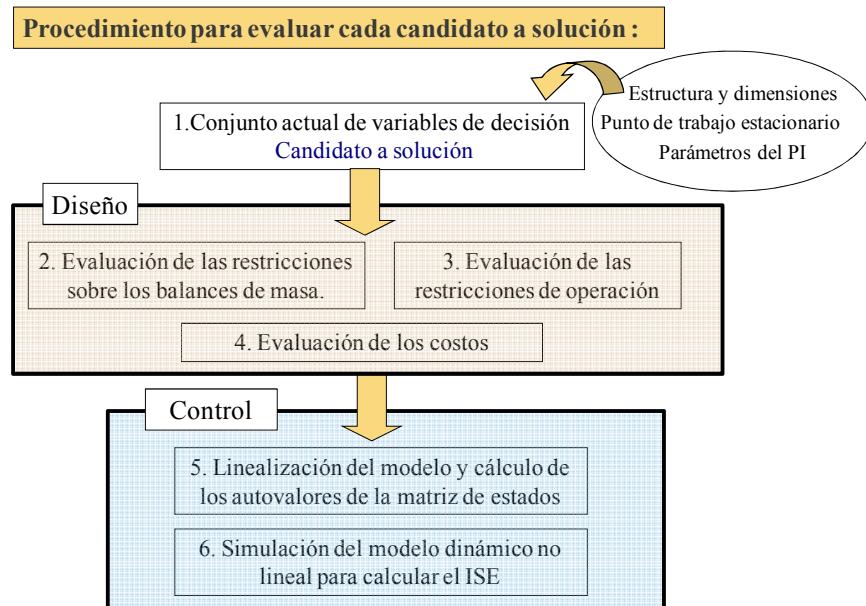


Figura 6.9. Evaluación de las soluciones candidatas durante la ejecución del algoritmo genético para el diseño integrado con un PI

Por otra parte, ya que al principio es difícil encontrar soluciones candidatas que satisfagan la condición de *controlabilidad*, para evitar la convergencia prematura a soluciones con costo óptimo pero desempeño dinámico deficiente, se agregan, solo durante las primeras generaciones, los diez cromosomas con mejores valores del *ISE* al grupo de los seleccionados para la siguiente generación.

La forma de evaluar cada solución candidata en este caso de estudio, es un poco más compleja, porque es necesaria la linealización del modelo para el cálculo de los autovalores de lazo cerrado y también la simulación del modelo dinámico no lineal para calcular el índice de desempeño (*ISE*). En la figura 6.9 se describe el procedimiento para evaluar cada cromosoma de la población.

Resultados.

En la tabla 6.10 se reportan los parámetros característicos de la mejor solución encontrada después de realizar diez pruebas con el algoritmo genético para cada conjunto de restricciones.

Igual que en el caso de optimización económica, en el escenario de restricciones de operación relajadas (*conjunto de restricciones 2*) se obtiene naturalmente una planta menos costosa. Sin embargo, las condiciones más exigentes dadas por el *conjunto de restricciones 1* permiten lograr una mayor remoción de materia orgánica (s_2) a costa de mayor volumen de reacción y flujo de recirculación y una mayor penalización sobre el desempeño dinámico porque el *ISE* es mayor. En cuanto a la estructura óptima, en ambos casos se obtiene un sistema con dos bio-reactores.

Si se compara la solución obtenida aquí, con la económicamente óptima reportada en la tabla 6.7, se encuentra un incremento significativo en los costos, ya que es necesario mayor volumen para atenuar el efecto de las perturbaciones. En consecuencia se obtiene, en ambos escenarios, una solución que es el compromiso entre los menores costos y un desempeño dinámico aceptable frente a perturbaciones, cuantificado por el *ISE* razonable. La figura 6.11 muestra la respuesta dinámica frente a las perturbaciones, de la planta óptima económicamente

a lazo abierto, y la figura 6.12 muestra la respuesta de lazo cerrado obtenida al aplicar el *diseño integrado*, en donde se observa la evidente mejora alcanzada en este último caso.

Tabla 6.10. Parámetros de la planta obtenida como resultado del *diseño integrado* del proceso de fangos activados con un controlador PI

Parámetros	Conjunto de restricciones 1	Conjunto de restricciones 2
<i>Costo</i>	0.25	0.17
$s_1 \text{ (mg/l)}$	69.0	104.5
$s_2 \text{ (mg/l)}$	27.0	32.7
$V_1 \text{ (m}^3\text{)}$	6531.6	4526.6
$V_2 \text{ (m}^3\text{)}$	3048.5	3403.4
$A \text{ (m}^2\text{)}$	2517.1	2208.8
$qr_1 \text{ (m}^3/\text{h)}$	643.5	265.3
$qr_2 \text{ (m}^3/\text{h)}$	0.46	1.52
Fk_1	0.18	0.27
Fk_2	0.13	0.23
K_p	-2.6	-39.3
T_i	137.8	138.8
<i>ISE</i>	1875.9	1046.7

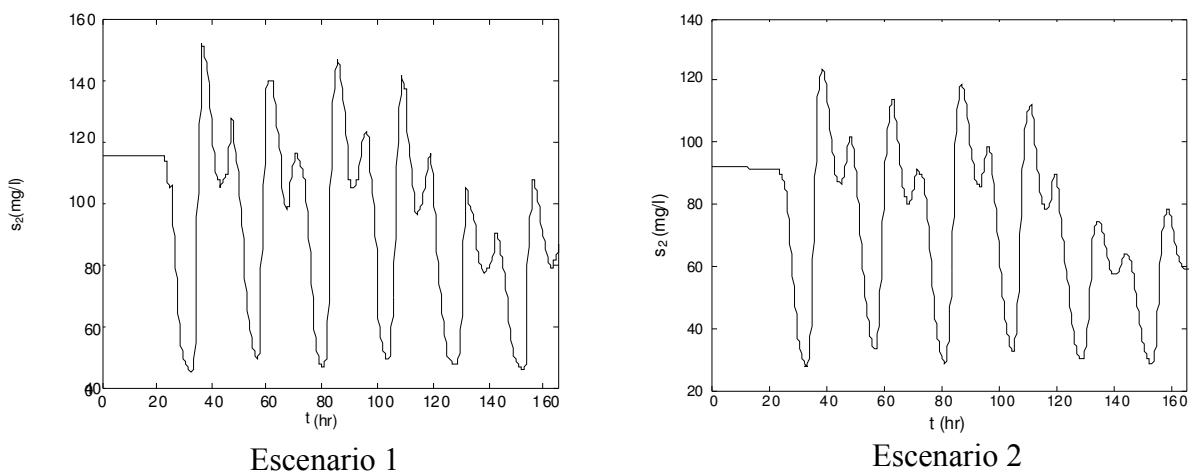


Figura 6.11. Respuesta dinámica de la planta resultante de la síntesis económica.

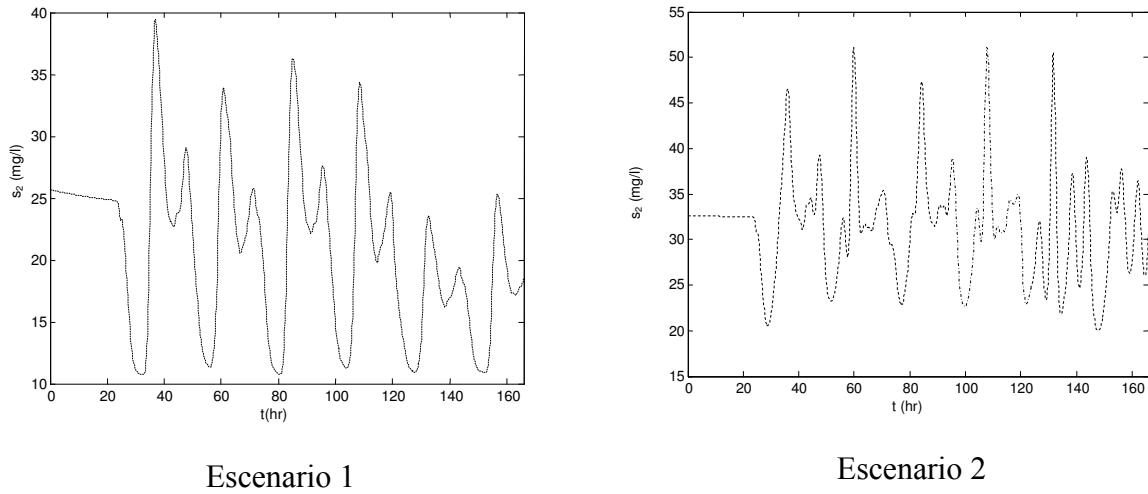


Figura 6.12. Respuesta dinámica de la planta resultante del diseño integrado con control PI

3.2.3. Diseño integrado con un esquema de control predictivo (MPC)

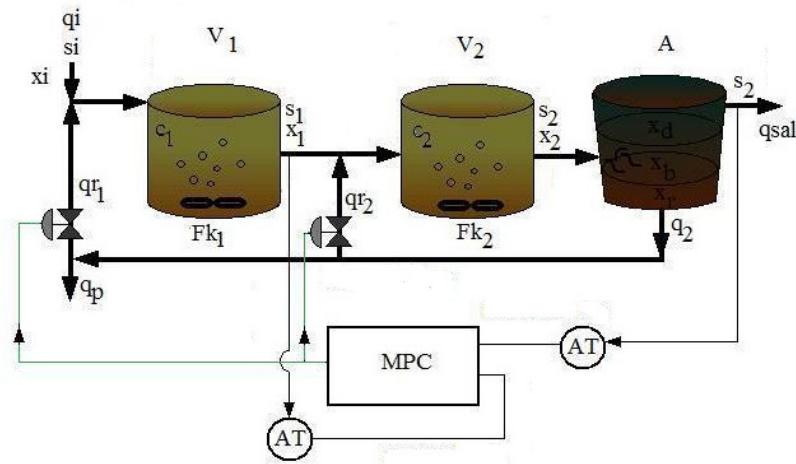


Figura 6.13. Superestructura del proceso de fangos activados para eliminación de materia orgánica con un controlador predictivo multivariante MPC

Formulación.

El planteamiento del diseño integrado con el controlador predictivo multivariable basado en modelos (MPC) y el estudio de la controlabilidad con las normas de sensibilidad que se presentan ahora, se toman de las propuestas de Francisco y Vega (2006 y 2009). Detalles han sido publicados en Vega et al. (2007), Francisco et al. (2009) y Revollar et al. (2010).

La superestructura utilizada para la planta con el controlador MPC se muestra en la figura 6.13. El cálculo de la ley de control se formula como un problema de optimización por minimización del funcional de error V , sujeto a restricciones sobre las entradas, la predicción de las salidas y las variaciones en las variables manipuladas, como sigue:

$$V(k) = \sum_{i=H_w}^{H_p} (\hat{y}(k+i|k) - r(k+i|k))^2 \cdot W_y(i) + \sum_{i=0}^{H_c-1} (\Delta\hat{u}(k+i|k))^2 \cdot W_u(i) \quad (6.73)$$

con:

$$\begin{aligned} u_{\min} < \hat{u}(k+i|k) &< u_{\max} \quad i = 0, \dots, H_c \\ y_{\min} < \hat{y}(k+i|k) &< y_{\max} \quad i = H_w, \dots, H_p \\ \Delta u_{\min} < \Delta\hat{u}(k+i|k) &< \Delta u_{\max} \quad i = 0, \dots, H_c \end{aligned}$$

donde:

\hat{y} : salida estimada,

r : trayectoria de referencia,

Δu : incrementos de la señal de control,

H_p : horizonte de predicción,

H_w : instante a partir del cual se penaliza el error de control. $H_w \geq l$,

H_c : horizonte de control, a partir del cual los Δu son nulos, es decir:

$$\hat{u}(k+i|k) = \hat{u}(k+H_c-1|k) \quad \forall i \geq H_c$$

$W_y(i)$: matriz diagonal que contiene los pesos aplicados a los errores de control a lo largo del horizonte de predicción,

$W_u(i)$: matriz diagonal que contiene los pesos aplicados a los esfuerzos de control.

Las matrices W_y y W_u dependen del instante de muestreo i , posibilitando la penalización variante del error y del esfuerzo de control.

Esta formulación matemática resulta en un problema de optimización cuadrático (QP), que se resuelve en cada instante de muestreo k y permite determinar la secuencia de acciones de control $\Delta\hat{u}(k+i|k)$ óptima que se debe aplicar sobre el sistema en el instante de tiempo

$k+1$. El modelo de predicción para el MPC es el modelo discreto linealizado en espacio de estado del proceso, con las salidas, entradas manipuladas y perturbaciones dadas por:

$$y = [s_1 \ s_2]^T,$$

$$u = [qr_1 \ qr_2]^T \text{ y}$$

$$d = [si \ qi]^T.$$

Resolviendo el problema anterior sin restricciones correspondiente a la optimización del controlador MPC definido hasta aquí, se obtiene una ley de control lineal, y puede ser expresada en el dominio de Laplace de acuerdo al diagrama de bloques del sistema nominal que se muestra en la figura 6.14, como:

$$u(s) = K_1(r(s) - y(s)) + K_2 d(s) \quad (6.74)$$

donde:

K_1 y K_2 son las funciones de transferencia entre la señal de control y las entradas $r(s)$ y $d(s)$, respectivamente, y dependen de los parámetros de sintonía del MPC.

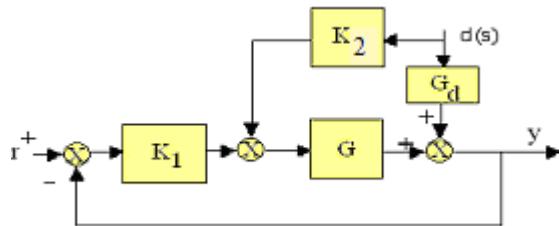


Figura 6.14. Sistema nominal en lazo cerrado

Nótese que la estructura propuesta del MPC es un sistema que combina retroalimentación y alimentación adelantada.

Teniendo en cuenta la ley de control (6.74) y la función de transferencia del sistema en lazo abierto, la respuesta del sistema en lazo cerrado es:

$$y(s) = \frac{G_0 K_1}{1 + G_0 K_1} r(s) + \frac{1}{1 + G_0 K_1} \tilde{d}(s) \quad (6.75)$$

donde:

$\tilde{d}(s)$ son las perturbaciones filtradas $\tilde{d}(s) = (G_0 K_2 + G d_0) d(s)$,

G_0 y $G d_0$ son las funciones de transferencia nominales.

La ecuación (6.75) puede ser expresada sustituyendo las funciones de sensibilidad y sensibilidad complementarias (definidas en el capítulo II):

$$y(s) = T_0(s) r(s) + S_0(s) R_{d0}(s) d(s) \quad (6.76)$$

donde:

$$S_0 = \frac{1}{1 + G_0 K_1} \quad (6.77)$$

$$T_0 = \frac{G_0 K_1}{1 + G_0 K_1} \quad (6.78)$$

$$R_{d0}(s) = G d_0(s) - K_2 G_0(s) \quad (6.79)$$

Para plantear el problema de sintonía automática con el sistema nominal, se considera la función de sensibilidad $S_0(s)R_{d0}(s)$ entre las perturbaciones medibles y la salida. También se considera la función de sensibilidad al control $M_0(s)$ (entre las perturbaciones medibles y las señales de control cuando la referencia es cero) dada por:

$$M_0(s) = \frac{u(s)}{d(s)} = \frac{K_2 - K_1 G d_0}{1 + G_0 K_1} \quad (6.80)$$

Por simplicidad, solo se presenta el caso una entrada-una salida, pero cuando existen varias entradas o salidas, se pueden calcular las matrices de sensibilidad del sistema en lazo cerrado y el trabajo se puede generalizar.

Se trata de realizar el ajuste óptimo de los parámetros (sintonía) del MPC como un problema de sensibilidad mixta que tiene en cuenta rechazo a perturbaciones y esfuerzos de control en la misma función. La optimización se plantea como:

$$\min_c \|N_0\|_\infty = \max_\omega |N_0(j\omega)| \quad (6.81)$$

donde:

c contiene los parámetros de sintonía del MPC: H_p , H_c y W_u .

La forma de capturar las distintas condiciones que deben imponerse sobre la función S en el dominio de la frecuencia, es la definición de un peso frecuencial con forma adecuada (W_p) que pondere la función S . Este peso, se elige de manera que su módulo inverso sea menor al espectro inverso de perturbaciones. Entonces:

$$N_0 = \begin{pmatrix} W_p \cdot S_0 \cdot R_{d0} \\ W_{esf} \cdot s \cdot M_0 \end{pmatrix}, \text{ y} \quad (6.82)$$

donde:

$W_p(s)$ y $W_{esf}(s)$ son pesos adecuados para conseguir las especificaciones de desempeño en lazo cerrado en base al espectro de perturbaciones, y para reducir los esfuerzos de control, respectivamente.

Las restricciones impuestas en el marco del problema, considerando perturbaciones normalizadas son:

$$\|N_0\|_\infty < 1 \quad (6.83)$$

$$\|W_p \cdot S_0\|_\infty < 1 \quad (6.84)$$

para penalizar adecuadamente los esfuerzos de control y garantizar un buen rechazo a perturbaciones, y

$$\|M\|_1 < u_{\max} \quad (6.85)$$

para restringir el máximo valor de la variable manipulada respecto a una cota u_{\max} .

Nótese que en la función objetivo (6.82) se incluyen los esfuerzos de control en vez de las magnitudes del control al considerar la derivada de la función de transferencia M_0 . La dependencia en s de las funciones de transferencia N_0 , S_0 y M_0 se omite por brevedad.

El análisis de *controlabilidad* realizado basado en el problema de sensibilidad mixta permite establecer restricciones en el problema de *diseño integrado*, donde las funciones de transferencia dependen de las salidas s_1 , s_2 y las variables manipuladas qr_1 , qr_2 , considerando las variaciones en s_i presentadas en la figura 6.5 como perturbación.

En la ley de control del MPC, la penalización es relativa al error, pues se considera que los pesos asociados al error de control $W\mathbf{y}$ son iguales a uno mientras que los pesos asociados al los esfuerzos de control $W\mathbf{u}$ (cambios en las variables manipuladas) son las variables de decisión ajustables.

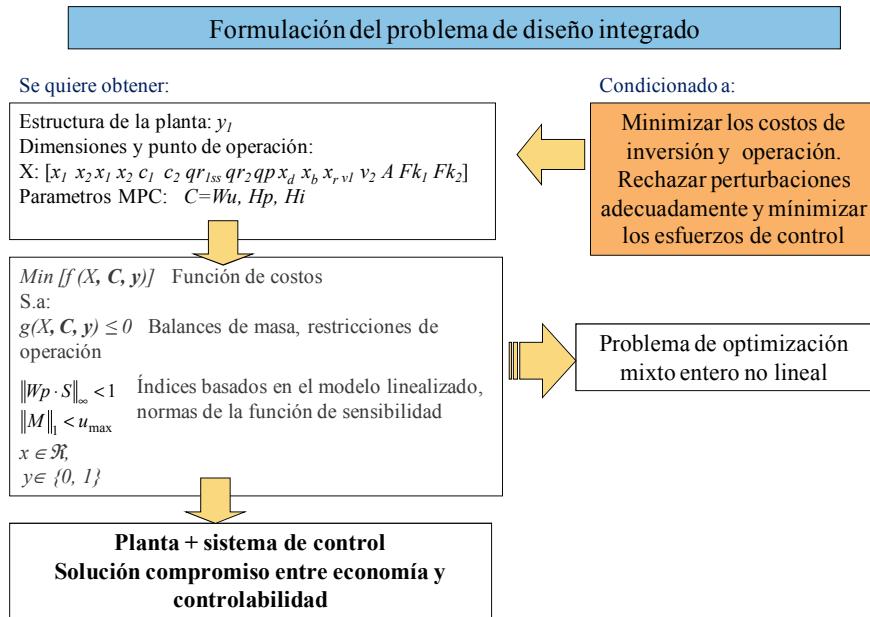


Figura 6.15. Formulación del problema de *diseño integrado* del proceso de fangos activados con un MPC

En la superestructura se observa que la selección de la estructura de control depende directamente de la configuración de la planta para $y_1=0 \Rightarrow v_2=0$ y $qr_2=0$ y el sistema de control es de una entrada una salida, se controla s_1 manipulando qr_1 y se tiene solo un peso asociado al

error (Wu). Para $y_I=1 \Rightarrow v_2 \neq 0$ y $qr_2 \neq 0$, y se tiene un sistema de control multivariable con los pareamientos s_1-qr_1 , s_2-qr_2 y \mathbf{Wu} es una matriz diagonal.

Esta formulación se describe de forma esquemática en la figura 6.15.

Algoritmo genético.

El cromosoma utilizado, que incluye los parámetros del controlador MPC, es:

$$[x_{1n}, x_{2n}, s_{1n}, s_{2n}, c_{1n}, c_{2n}, x_{dn}, x_{bn}, x_{rn}, qr_{1ss}, qr_{2n}, qp_n, Fk_{1n}, Fk_{2n}, v_{1n}, v_{2n}, A_n, Wqr_{1n}, Wqr_{2n}, Hp, Hc, y_I]$$

donde:

Wqr_{1n} y Wqr_{2n} denotan los pesos asociados a los esfuerzos de control. Conforman una matriz diagonal cuando $y_I=1$ y es un escalar cuando $y_I=0 \Rightarrow Wqr_{2n}=0$.

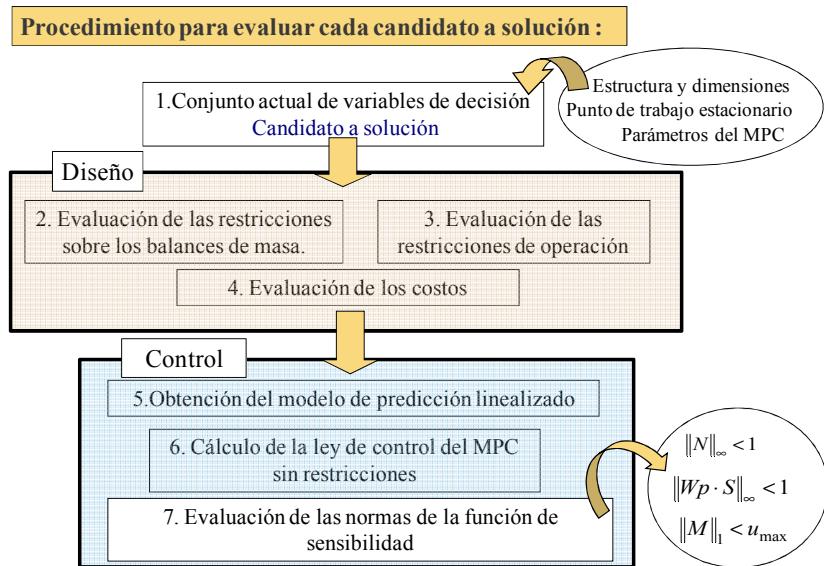


Figura 6.16. Evaluación de las soluciones candidatas durante la ejecución del algoritmo genético para el diseño integrado con un MPC

El algoritmo genético presenta las mismas características y operadores que en el caso anterior, solo fueron necesarias algunas variaciones en los parámetros ajustables. En realidad,

esta aplicación es más sencilla desde el punto de vista del gasto computacional porque no es necesario simular la respuesta del modelo no lineal. Además, al suponer un MPC sin restricciones, el cálculo de la ley de control se hace de forma bastante rápida con la herramienta MPC de MATLAB.

El problema se resuelve con una población de 100 individuos y un número de iteraciones máximo de 1000.

En la figura 6.16 se muestra el procedimiento para evaluar cada cromosoma durante la ejecución del algoritmo genético.

Resultados.

Se plantean dos casos de estudio, utilizando las mismas restricciones de operación del conjunto 2 de la tabla 6.5, pero considerando diferentes demandas sobre el desempeño dinámico en función del rechazo a perturbaciones y diferente valor máximo que puede alcanzar la variable manipulada en presencia de perturbaciones. Los resultados se presentan en la tabla 6.11. En el caso 1 se considera $\|M\|_1 < 1000$ y un peso $Wp_{caso1}(s) = \frac{8s + 19.2}{s + 0.0001}$. En el caso 2 se tiene la restricción $\|M\|_1 < 450$ y un peso $Wp_{caso2}(s) = \frac{4.4s + 10.56}{s + 0.0001}$.

Se observa cómo la variación en los requerimientos sobre las normas de *controlabilidad* produce nuevamente un cambio de estructura y condiciones de operación de la planta. En el caso 2 el resultado de la optimización es una planta con un mejor desempeño ante perturbaciones, ya que el peso Wp_1 seleccionado es más restrictivo para la función de sensibilidad S_0 , mientras que la cota más relajada sobre $\|M\|_1$ permite tener un rango más amplio de acción sobre la variable manipulada para rechazar perturbaciones. Los valores de los pesos $Wesf$ para las funciones de sensibilidad de qr_1 y qr_2 se establecen en:

$$Wesf_{qr1}(s) = \frac{0.0117s + 0.14}{s + 0.0004} \quad Wesf_{qr2}(s) = \frac{0.0183s + 0.22}{s + 0.0004}$$

Tabla 6.11. Parámetros de la planta obtenida como resultado del *diseño integrado* del proceso de fangos activados con un controlador MPC

	Caso 1	Caso 2
V_1	5409.2	3442.7
V_2	0	2819.2
A	1253.1	1147.0
$s_1 \text{ o } s_2$	107.03	93.86
Wu	[0.009 0.274]	[0.347 0.052]
Hp	7	7
Hc	2	4
$\ N\ _\infty$	0.74	0.96
$\ M\ _1$	725.7	281.3
$\ Wp \cdot S\ _\infty$	0.63	0.93
$rango(s)$	17.7	36.5
$max(qr)$	563.9	254.5

En la figura 6.17 se muestra la función de sensibilidad para ambos casos de estudio junto con los pesos Wp (que representan la magnitud de las perturbaciones), y el inverso del espectro de perturbaciones en el influente. Aquí, se observa mayor separación entre el peso Wp_{caso2} y la función de sensibilidad, que en el caso 1, evidenciando peor desempeño ante perturbaciones en el caso 2. Esto se puede observar también en la respuesta dinámica con el modelo lineal presentada en la figura 6.18.

Si se comparan estos resultados con los obtenidos para el sistema con controlador PI, es notable la ventaja del esquema MPC al obtenerse plantas más reducidas, aunque se debe tener presente que se usan índices de distintos tipos para el diseño. Sin embargo, el uso de las funciones de sensibilidad para imponer restricciones y evaluar la controlabilidad hace práctica la resolución del problema con el MPC que sería de otro modo muy engorrosa.

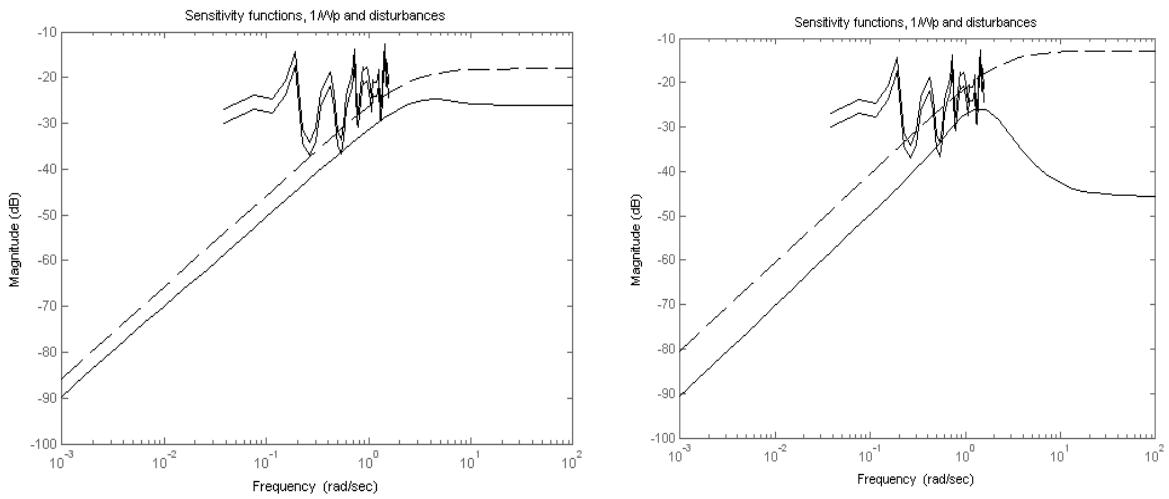


Figura 6.17. Magnitud de la función de sensibilidad (S_0) para los casos 1 y 2 respectivamente, junto con los pesos W_p (línea punteada) y el inverso del espectro de perturbaciones en el influente.

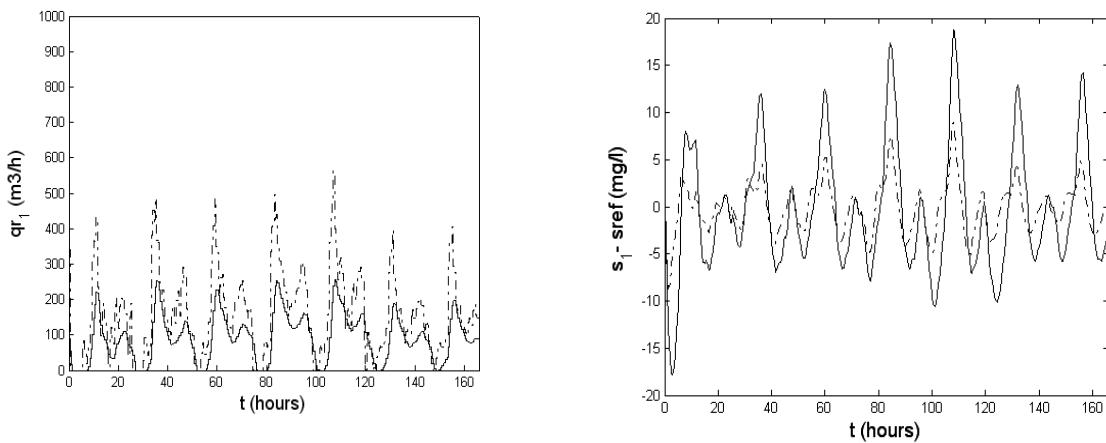


Figura 6.18. Comparación de la respuesta dinámica de las plantas resultantes de la *síntesis económica* y del *diseño integrado* en el escenario 2

4. DISEÑO INTEGRADO DEL PROCESO DE FANGOS ACTIVADOS PARA LA ELIMINACION DE MATERIA ORGANICA Y NUTRIENTES. MODELO 2.

La aplicación de la metodología de diseño integrado al proceso de fangos activados se hace más interesante si se consideran los procesos de nitrificación y desnitrificación. Para esto, se utiliza una variante simplificada del BSM1 encontrada en Ingildsen (2002) y en Halvarson (2003), donde se toman en cuenta únicamente las variables significativas en una escala media de tiempo. Los procesos más lentos, generalmente los relacionados con compuestos insolubles como la hidrólisis y la amonificación del nitrógeno orgánico no se consideran y se supone que la biomasa heterótrofa y autótrofa no varía, por lo que el número de estados en cada reactor se reduce de 13 a 4.

En esta simplificación del BSM1, el sistema completo se aproxima a un reactor anóxico y uno aeróbico, con un volumen equivalente al volumen total del los compartimientos anóxicos y aeróbicos del Benchmark. Se supone el proceso de desnitrificación sólo tiene lugar en el tanque anóxico y el proceso de nitrificación ocurre únicamente en el tanque aireado. El decantador se aproxima a un punto de mezcla y el reflujo externo se supone constante.

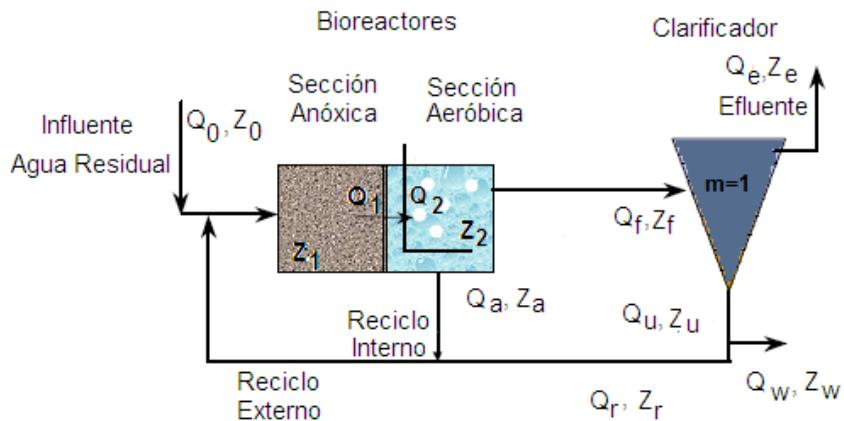


Figura 6.19. Esquema general del *BSM1* simplificado

Retomando las ecuaciones 6.5 a 6.28, que representan los balances de masa, tasas de reacción y procesos biológicos involucrados en el BSM1, se obtiene el conjunto de ecuaciones

diferenciales para el modelo simplificado, donde el índice 1 está referido al tanque anóxico y el índice 2 está referido al tanque aeróbico.

En el reactor anóxico:

$$\frac{dS_{NH}(1)}{dt} = \frac{1}{V_1} \left[Q_0 \cdot S_{NH} \cdot in + Q_a S_{NH}(2) - (Q_0 + Q_a) \cdot S_{NH}(1) - i_{xb} \cdot \rho_1(1) - i_{xb} \cdot \rho_2(1) - \left(i_{xb} + \frac{1}{Y_A} \right) \cdot \rho_3(1) \right] \quad (6.86)$$

$$\frac{dS_{NO}(1)}{dt} = \frac{1}{V_1} \left[Q_a S_{NO}(2) - (Q_0 + Q_a) S_{NO}(1) - \frac{1 - Y_H}{2.86 Y_H} \rho_2(1) + \frac{1}{Y_A} \cdot \rho_3(1) \right] \quad (6.87)$$

$$\frac{dS_s(1)}{dt} = \frac{1}{V_1} \left[Q_0 \cdot S_{s,in} + Q_a S_s(2) - (Q_0 + Q_a) S_s(1) - \frac{1}{Y_H} \rho_1(1) - \frac{1}{Y_H} \rho_2(1) \right] \quad (6.88)$$

$$\frac{dS_o(1)}{dt} = \frac{1}{V_1} \left[Q_a \cdot S_o(2) - (Q_0 + Q_a) S_o(1) - \left[\frac{1 - Y_H}{Y_H} \rho_1(1) + \left(\frac{4.57}{Y_A} + 1 \right) \rho_3(1) \right] \right] \quad (6.89)$$

En el reactor aeróbico:

$$\frac{dS_{NH}(2)}{dt} = \frac{1}{V_2} \left[(Q_0 + Q_a) S_{NH}(1) - (Q_0 + Q_a) S_{NH}(2) - i_{xb} \cdot \rho_1(2) - \left(i_{xb} + \frac{1}{Y_A} \right) \cdot \rho_3(2) \right] \quad (6.90)$$

$$\frac{dS_{NO}(2)}{dt} = \frac{1}{V_2} \left[(Q_0 + Q_a) S_{NO}(1) - (Q_0 + Q_a) S_{NO}(2) - \frac{1 - Y_H}{2.86 Y_H} \rho_2(2) + \frac{1}{Y_A} \cdot \rho_3(2) \right] \quad (6.91)$$

$$\frac{dS_s(2)}{dt} = \frac{1}{V_2} \left[(Q_0 + Q_a) S_s(1) - (Q_0 + Q_a) S_s(2) - \frac{1}{Y_H} \rho_1(2) - \frac{1}{Y_H} \rho_2(2) \right] \quad (6.92)$$

$$\frac{dS_o(2)}{dt} = \frac{1}{V_2} \left[(Q_0 + Q_a) \cdot S_o(1) - (Q_0 + Q_a) \cdot S_o(2) - \left[\frac{1 - Y_H}{Y_H} \rho_1(2) + \frac{4.57 - Y_A}{Y_A} \rho_3(2) \right] \right] \\ + K_L a_1 (S_{O,sat} - S_o(2)) \quad (6.93)$$

En el primer reactor el único proceso biológico importante es el crecimiento anóxico de la biomasa heterótrofa, relacionado con la desnitrificación:

$$\rho_2(1) = \mu_H \cdot \left(\frac{Ss(1)}{K_s + Ss(1)} \right) \cdot \left(\frac{K_{O,H}}{K_{O,H} + So(1)} \right) \left(\frac{S_{NO}(1)}{K_{NO} + S_s(1)} \right) n_g X_{B,H} \quad (6.94)$$

En el segundo reactor, donde se tiene una mayor concentración de oxígeno, se consideran los crecimientos aeróbicos de la biomasa heterótrofa y autótrofa, relacionados con la nitrificación:

$$\rho_1(2) = \mu_H \cdot \left(\frac{S_S(2)}{K_s + S_S(2)} \right) \cdot \left(\frac{S_O(2)}{K_{O,H} + S_O(2)} \right) X_{B,H} \quad (6.95)$$

$$\rho_3(2) = \mu_A \cdot \left(\frac{S_{NH}(2)}{K_{NH} + S_{NH}(2)} \right) \cdot \left(\frac{S_O(2)}{K_{O,A} + S_O(2)} \right) X_{B,A} \quad (6.96)$$

4.1. FORMULACION GENERAL DEL PROBLEMA DE OPTIMIZACION

En este caso de estudio se tiene una estructura fija de la planta y se conocen las dimensiones de los reactores. En este modelo simplificado la dimensión de los bio-reactores es equivalente a lo propuesto en el Benchmark, entonces: $V_1=1000\text{m}^3$ y $V_2=1333\text{m}^3$.

El objetivo de control es regular la concentración de nitratos en el primer reactor alrededor de un valor óptimo mediante la manipulación de la recirculación de nitrógeno, y mantener la concentración de oxígeno disuelto en el segundo reactor.

Se formula el diseño integrado del proceso para hallar el punto de operación y los parámetros de un controlador MPC, que minimicen los costos de operación relacionados con la energía de bombeo y la calidad del efluente, y que además permitan lograr un desempeño dinámico óptimo el cual es cuantificado con el ISE.

Las variables de decisión son: el caudal de recirculación de nitrógeno Q_a , la dosificación de carbono $S_{dossage}$, la concentración de oxígeno disuelto en el reactor aeróbico $S_o(2)$ y los parámetros de sintonía del controlador, es decir, los pesos Wy y Wu y los horizontes de predicción y de control, H_p y H_c , estos últimos variables enteras.

Objetivos de optimización.

Los costos de inversión (EUR) sólo dependen de los costos de construcción de la estación de bombeo:

$$Cost_{invest} = 2334 \cdot Q_a^{0.637} \quad (6.97)$$

Los costos de operación (EUR/d) se relacionan con los índices de desempeño que proveen una medida de la calidad de la operación, como la calidad del efluente (EQ), que es un indicador de la cantidad de polución producida por unidad de tiempo y la energía de bombeo (PE -kWh/d). En este modelo el decantador se representa como un punto de mezcla y las concentraciones que van al efluente se calculan como una fracción respecto a la salida del decantador.

$$EQ = \frac{50}{T \cdot 1000} \int_0^{tf(\text{days})} \left[\begin{aligned} & (X_{B,A} + X_{B,H} + S_S(2)) + \\ & 0.25 \cdot (S_S(2) + (1 - 0.08) \cdot (X_{B,A} + X_{B,H})) \\ & + 0.75 \cdot (X_{B,A} + X_{B,H}) + S_{NH}(2) + i_{xb} (X_{B,A} + X_{B,H}) \end{aligned} \right] Q_e dt \quad (6.98)$$

$$PE = \frac{1}{T} \int_0^{tf} Q_r dt \quad (6.99)$$

Los factores de costo asociados al EQ y a la energía de bombeo son 50EUR/Kg y 25 EUR/kW (Vanrolleghem and Gillot, 2002).

Entonces, los costos asociados a la operación de la planta están dados por:

$$Cost = w_1 \cdot Cost_{invest} + w_2 \cdot Cost_{EQ} + w_3 \cdot Cost_{PE} \quad (6.100)$$

donde w_1, w_2, w_3 son pesos ajustados de acuerdo a los requerimientos de la optimización.

Una medida similar a la integral del error (ISE) se usa para medir el desempeño dinámico de la planta con el controlador. En este caso:

$$ISE = \int_0^{tf} (S_{NOss}(1) - S_{NO}(1)(t))^2 \cdot dt \quad (6.101)$$

Restricciones del proceso.

Se establecen los siguientes límites sobre la concentración de los compuestos nitrogenados:

$$S_{NH}(2) \leq 4mg/l \quad (6.102)$$

$$S_{Ntot}(2) \leq 24mg/l \quad (6.103)$$

donde:

$S_{Ntot}(2)$: concentración total de compuestos de nitrogenados a la salida del reactor.

La edad de los fangos en este caso se calcula como:

$$SA = \frac{1}{\mu_A \frac{S_{NH}}{S_{NH} + K_{NH}} \frac{S_O}{S_O + K_{OA}} \frac{V_2}{V_1 + V_2} - b_A} \quad (6.104)$$

Se propone como restricción:

$$4 \text{ days} \leq SA \leq 10.6 \text{ days}, \quad (6.105)$$

y para la concentración de oxígeno disuelto en el reactor aeróbico:

$$1 \leq S_O(2) \leq 3 \quad (6.106)$$

Estrategia y restricciones de control.

El objetivo de control es mantener la concentración de nitrato en el reactor anóxico $S_{NO}(1)$ alrededor del punto de operación óptimo manipulando el reciclo de nitrógeno (Q_a). Otra variable a controlar es el oxígeno disuelto en el reactor aeróbico $S_o(2)$, lo cual se logra manipulando la transferencia de aire al reactor dada por K_{La} .

El proceso está sometido a las perturbaciones en el flujo de entrada Q_0 y la concentración de material orgánica a la entrada S_{si} , caracterizadas por Alex et al. (2008) para clima seco, y descritas en las figuras 6.4 y 6.5.

Se imponen las siguientes restricciones sobre los parámetros del controlador MPC:

$$4 \leq Hp \leq 12 \quad (6.107)$$

$$1 \leq Hc \leq 3 \quad (6.108)$$

$$0 \leq Wu \leq 0.8; \quad Wy = 1 \quad (6.109)$$

Se pueden establecer diferentes condiciones para resolver el problema aquí planteado. A continuación se presenta un ejemplo.

4.2. CASO DE ESTUDIO

Por simplicidad, se propone una estrategia de control de una entrada una salida, para regular $S_{NO}(1)$ alrededor del punto de operación óptimo manipulando el reciclo de nitrógeno (Q_a). En este caso, se supone que la concentración de oxígeno disuelto en el reactor aeróbico $S_o(2)$, se mantiene fija en el valor de referencia dada por la optimización mediante un lazo de control no especificado, por lo tanto, $\frac{dS_o(2)}{dt} = 0$, en la ecuación 6.93.

La ley de control del MPC esta dada por el problema definido por la ec. 6.73. Se utiliza el modelo linealizado en espacio de estados como modelo de predicción. Para facilitar el cálculo, se supone que no hay restricciones en 6.73 y se utiliza la herramienta de diseño de MPC de Matlab®.

Algoritmo genético.

El cromosoma propuesto para este problema contiene variables continuas y enteras:

$$[S_{sdossage}, Q_a, S_o(2), Wu, Hp, Hc]$$

Se utiliza un algoritmo genético de codificación real en su variante multiobjetivo presentada en la sección 5 del capítulo V. En efecto se consideran dos objetivos de optimización, los costos de inversión y de operación dados en la ec. (6.100) (f_1), y el índice de controlabilidad ISE dado en la ec. (6.101) (f_2). Igual que anteriormente, las violaciones de las restricciones se tratan con una función de penalización. En el proceso de selección de los padres, la población se ordena en función de cada objetivo por separado por lo que se tiene una lista ordenada para cada objetivo. Se descartan los peores individuos de cada lista (25% del total), luego, se identifican los cromosomas que representan las mejores soluciones

compromiso mediante una función de adaptación basada en la posición de los individuos en cada lista ordenada.

Los operadores utilizados son: cruce aritmético y mutación constante con una tasa de 0.1. Se usa selección por torneo para cada objetivo por separado. La población es de 20 individuos y se requieren 40 iteraciones como máximo y 10 ensayos para cada caso.

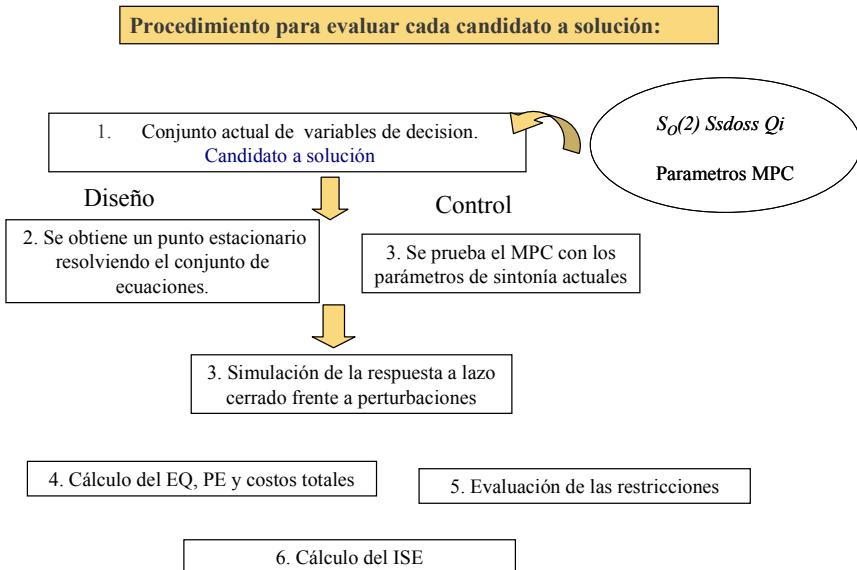


Figura 6.20. Evaluación de las soluciones candidatas durante la ejecución del algoritmo genético para el diseño integrado con un MPC

Resultados.

Se proponen dos casos de estudio, el caso de estudio 1 consiste en la minimización de los costos de inversión y operación sujetos a una restricción sobre el desempeño dinámico ($ISE \leq 30000$). En el caso de estudio 2, tanto los costos como el desempeño dinámico son objetivos de la optimización.

En la tabla 6.12 se muestran los parámetros e índices de desempeño de las plantas obtenidas.

Tabla 6.12. Parámetros de la planta obtenida como resultado del *diseño integrado* del proceso de fangos activados con un controlador MPC monovariable.

	Caso 1	Caso 2
<i>Costos de inversión (Eur)</i>	57859	103680
<i>ISE</i>	26466	5824
<i>PE (kWh/d)</i>	15.56	38.6
<i>EQ (Kg/d)</i>	25433	25450
<i>SA (d)</i>	4	4
Variables de Decisión		
<i>Qi</i>	155.27	385.92
<i>S_{sdossage} mg/l</i>	51.47	26.6
<i>So(2) mg/l</i>	1.19	2.33
<i>Wu</i>	0.75	0.23
<i>Wy</i>	1	1
<i>Hp</i>	7	12
<i>Hc</i>	2	1
Salida en estado estacionario		
<i>S_{NO(1)} mg/l</i>	2.28	7.10
<i>S_{NH(2)} mg/l</i>	0.96	0.81
<i>S_{s(2)} mg/l</i>	0.25	0.12

La manera de plantear los objetivos de la optimización influye significativamente sobre los diseños obtenidos. La optimización de ambos objetivos lleva a plantas de mayor tamaño pero con mejor desempeño dinámico. El principal problema al imponer cotas sobre los índices de desempeño dinámico es que se necesita conocer al priori el comportamiento de la planta para establecer un límite adecuado.

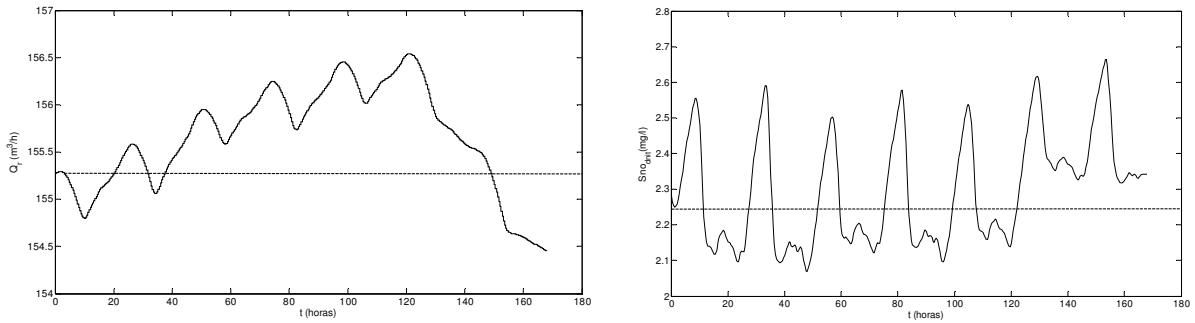


Figura 6.21. Señal de control y respuesta de la planta obtenida en el Caso 1.

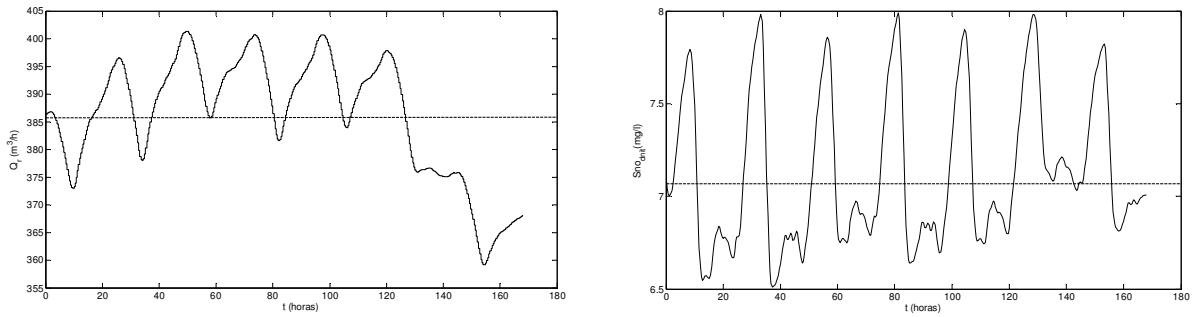


Figura 6.22. Señal de control y respuesta de la planta obtenida en el Caso 2.

5. CONCLUSIÓN

El diseño integrado del proceso de fangos activados se lleva a cabo usando un modelo matemático simple con la eliminación de materia orgánica únicamente, y usando un modelo más complejo donde se incluye la eliminación de materia orgánica y nutrientes. Los casos de estudio planteados se seleccionan para evaluar distintas aplicaciones de la metodología donde se contempla la síntesis solamente económica del proceso, la síntesis y diseño con controladores PI y la síntesis con controladores avanzados MPC, considerando la evaluación de índices de controlabilidad basados en el modelo dinámico no lineal e índices basados en la evaluación de normas de sensibilidad robusta. También se evalúa el diseño de una estructura fija con un controlador MPC considerando una formulación multiobjetivo, donde se optimizan por separado índices dinámicos de desempeño de la planta e índices de control.

En todos los casos se formula un problema de optimización no lineal mixta entera que es resuelto de forma satisfactoria utilizando los algoritmos genéticos, que permiten evaluar con facilidad índices dinámicos de desempeño (ISE, PE, EQ). Tal como se puede encontrar en los resultados finales, la estructura y dimensiones de la planta obtenida dependen de las restricciones de controlabilidad impuestas y de la estrategia de control utilizada. En esto se evidencia la versatilidad de la metodología de diseño integrado que puede ser aplicada para lograr diseños que satisfacen el compromiso diseño-control en distintos escenarios con distintos niveles de exigencias y alcances del problema.

El algoritmo genético demostró ser una excelente herramienta en la resolución del problema de optimización en todos los casos de estudio, por lo cual se recomienda su uso como alternativa a los métodos deterministas tradicionales para la resolución de estos complejos problemas.

Una de las mayores ventajas de los algoritmos genéticos en diseño integrado es lo sencillo y compacto de su aplicación, ya que son capaces de seleccionar automáticamente la configuración de la planta y al mismo tiempo los parámetros del proceso y diseño del controlador que mejor satisface las condiciones impuestas, haciendo innecesaria la descomposición del problema numérico.

CAPITULO VII

CONCLUSIONES GENERALES

Este trabajo está dedicado al *diseño integrado* es decir a la síntesis y diseño simultáneos del proceso y el control en plantas químicas, y a la aplicación de algoritmos genéticos para la resolución de los problemas de optimización que surgen de este planteamiento.

Después de una revisión intensiva de la literatura relacionada al diseño integrado y una clasificación de las distintas líneas de investigación, se identificaron elementos mejorables o áreas de mayor interés, y se constató que los algoritmos genéticos por su parte abren la posibilidad de introducir mejoras en muchas de estas áreas, porque en general las mayores limitaciones del diseño integrado radican en la complejidad de los problemas numéricos que se presentan. Los aspectos más resaltantes tratados en este trabajo son: la formulación completa del diseño integrado que incluye la síntesis del proceso, su diseño, y la selección y ajuste de las estrategias de control; el uso de índices basados en el modelo no lineal para la evaluación del desempeño dinámico ante situaciones realistas de perturbaciones; y la aplicación de esquemas de control especiales en lugar de los controladores retroalimentados típicos.

La revisión de las técnicas conocidas para evaluar la controlabilidad, permitió distinguir del análisis de controlabilidad entrada-salida el de controlabilidad de estado, tanto en base a modelos lineales como no lineales. Entre todas, las más apropiadas para cuantificar el desempeño dinámico de la planta que se diseña son las basadas en índices numéricos calculados por integración del modelo dinámico no lineal del proceso, pero tienen el inconveniente del gasto computacional que representa su evaluación. En este trabajo se estudiaron métodos alternativos para evaluar la capacidad de los diseños de ser controlados, en particular, el análisis de controlabilidad de estado no lineal, y también se presentaron aplicaciones de índices basados en normas de control robusto.

Se desarrollaron diferentes aplicaciones de diseño integrado con la intención de demostrar el potencial de los algoritmos genéticos como métodos alternativos para la resolución de estos problemas. En todos los casos, los resultados evidenciaron la ventaja de aplicar esta metodología que permite obtener las dimensiones óptimas de la planta, junto con el punto de operación y un controlador sintonizado para garantizar los requerimientos de controlabilidad impuestos.

El primer caso de estudio, el diseño integrado de un sistema de reactores químicos tipo tanque agitado continuo, se planteó para ilustrar el efecto de la interacción entre la síntesis, el diseño y el control del proceso, y para introducir las características especiales del algoritmo genético propuesto para los problemas de optimización dinámica mixta entera no lineal asociados al diseño.

Se llevó también a cabo el diseño simultáneo de proceso y control de una planta sulfitadora perteneciente a la etapa de clarificación en un ingenio de azúcar de caña, que se tradujo en un problema de optimización no lineal donde se imponen restricciones sobre las condiciones de operación de la planta y las medidas de controlabilidad de estado, mientras se minimizan los costos de inversión, operación y de control, cuantificados mediante la evaluación del modelo dinámico no lineal a lazo cerrado. El problema de optimización se resolvió eficientemente usando el algoritmo genético de codificación real propuesto. En este caso se introdujo el análisis de controlabilidad de estado, en primer término para definir cualitativamente la concepción del lazo de control, y el pareamiento más conveniente variable manipulada-variable controlada, que garantiza la ausencia de grados de libertad insatisfechos y el rango completo de la matriz de controlabilidad no lineal. Además, los requisitos de la controlabilidad de estado también establecieron límites o rangos posibles en las variables de diseño, que se usaron como restricciones en el cálculo del diseño simultáneo de proceso y control.

Esta aplicación, se completó luego con una formulación multi-objetivo del diseño del proceso usando un esquema de control MPC no lineal basado en un algoritmo de enumeración restringida. Se extiendió así la aplicación del algoritmo genético al problema multi-objetivo

evidenciando su flexibilidad para introducir mejoras en el marco del diseño integrado a pesar de la complejidad de los problemas involucrados.

Se abordó también el diseño integrado del proceso de fangos activados en plantas de tratamiento de aguas, considerando dos variantes del caso. En la primera se realizó la síntesis y diseño del proceso usando un modelo bastante simplificado. En la segunda se propuso el diseño con una estructura fija del proceso, pero usando un modelo más realista e índices de desempeño dinámico directamente relacionados con características operacionales del proceso. Se desarrolló y validó una variante del algoritmo genético para la optimización mixta-entera con restricciones que resulta de la formulación del problema de diseño estructural y paramétrico. Este algoritmo tiene como base el algoritmo genético simple, pero contiene modificaciones en su codificación y en algunos operadores de manera que puede manejar simultáneamente variables enteras y variables reales, además de incorporar técnicas heurísticas para mejorar su desempeño en la resolución de este tipo de problemas.

En conclusión, el Diseño Integrado de Procesos, es una metodología en la cual los procedimientos de diseño del proceso y del sistema de control se integran en el marco de un problema de optimización. Las ventajas que ofrece esta metodología y la complejidad de los problemas de optimización que surgen de su formulación, son razones que justifican la búsqueda de métodos de optimización alternativos como los algoritmos genéticos para ser capaces de resolver eficientemente estos problemas.

La aplicación del algoritmo genético propuesto, supone siempre una etapa de afinación de los ajustes de los parámetros, estrategias para tratar restricciones y adaptación de los operadores. Por otra parte, el algoritmo evolutivo por su naturaleza, no exige condiciones de derivabilidad u otras en las funciones a optimizar, ni muestra dependencia de los valores iniciales, por lo que es una excelente herramienta para resolver los problemas del tipo MINLP y MIDO, que surgen del diseño.

La plataforma para la resolución del problema de diseño integrado utilizando algoritmos genéticos es general, puede ser aplicada a problemas complejos con un alto número de variables tanto enteras como reales y puede considerar cualquier esquema de control más complejo de forma directa. Así mismo, es interesante hacer un estudio del efecto de la

incertidumbre en los parámetros del modelo y estudiar distintos escenarios de perturbaciones. En este trabajo no se incluye un análisis formal del efecto de incertidumbre pero en algunos casos se consideró un perfil de perturbaciones complejo en el diseño del proceso de fangos activados.

Se plantea en el futuro, aplicar esta plataforma para la resolución de problemas de mayor escala, específicamente enfocados en lograr la operación óptima de una planta de tratamiento de aguas completa, por lo que un primer paso es extender la aplicación desarrollada utilizando el modelo BSM1 completo. Además, dada la versatilidad de los algoritmos genéticos y el buen desempeño exhibido en los casos de estudio propuestos en este trabajo, se espera utilizarlos como herramienta en otra área de creciente interés en el control de procesos, como es el control total de planta.

REFERENCIAS

- Alex, J. L. (April). Benchmark Simulation Model no. 1 (BSM1). 2008
- H. Alvarez, R. Lamanna, S. Revollar y P.Vega. Metodología para la Obtención de Modelos Semifísicos de Base Fenomenológica Aplicada a una Sulfitadora de Jugo de Caña de Azúcar. Revista Iberoamericana de Automática e Informática Industrial (RIAI). Vol. 6, Nro. 3, pp.10 – 20. 2009
- Bansal V., J.D. Perkins, y E.N. Pistikopoulos. Flexibility analysis y design of dynamic processes with stochastic parameters. Comput. Chem. Eng., 22, Suppl:S817-S820, 1998.
- Bansal, V.; Perkins, J. D.; Pistikopoulos, E. N.; Ross, R.; van Schijndel, J. M. G. Simultaneous design and control optimization under uncertainty. Comput. Chem. Eng., 24, 261–266. 2000
- Barton, P.I.. The Modeling and Simulation of Combined Discrete/Continuous Processes, PhD Thesis, Univ. of London. 1992
- Bahri, P. A.; Bandoni, J. A.; Romagnoli, J. A. Integrated flexibility and controllability analysis in design of chemical processes. *AIChE J.*, 43, 997–1015. 1997
- Chawankul, N.; Ricardez Sandoval, L. A.; Budman, H.; Douglas, P. L. Integration of design and control: A robust control approach using MPC. *Can. J. Chem. Eng.*, 85, 433–446. 2007
- Copp, J.B.,, The COST Simulation Benchmark: Description and Simulator Manual. Office for Official Publications of the European Community. 2002
- Costa, L. y P. Oliveira *Evolutionary algorithms approach to the solution of mixed integer-non-linear programming problems*. Comp. Chem. Eng. **25**, 257. 2001
- Cuthrell J.E. y L.T. Biegler. On the optimization of differential-algebraic process systems. *AIChE J.*, 33(8) :1257-1270, 1987.

De Prada, C., Vega, P. and Alonso, L. Modelling and simulation of a sulfitation tower for adaptive control. Proceedings of the 11th IASTED Conference on Applied Modelling and Simulation. 1984

Deb, K , Anand, A., y Joshi, D A Computationally Efficient Evolutionary Algorithm for Real-Parameter Optimization. KanGAL Report No. 2002003. 2002

Deb, K. Genetic Algorithms for Optimization. KanGAL Report Number 2001002. 2001

DeJong, K. An analysis of the behavior of a class of genetic adaptive systems. Diss. Abstr. Int. B, 36, 5140. 1975

Dimitriadis V.D. y E.N. Pistikopoulos. Flexibility analysis of dynamic systems. Ind. Eng. Chem. Res., 34:4451-4462, 1995.

Dimitriadis V.D., N. Shah, y C.C. Pantelidis. Model-based safety verification of discrete/continuous chemical processes. AIChE Annual Meeting in. San Francisco, 1994.

Dimitriadis, V.D. y E.N. Pistikopoulos. Flexibility Analysis of Dynamic Systems. Ind. Eng. Chem. Res. 34(12), 1451-4462. 1995

Dominguez, S. Revollar, M. Francisco, R. Lamanna, P. Vega. Simultaneous Synthesis and Integrated Design of Chemical Processes Using IMC PID Tuning Methods. International Conference on Chemical and Process Engineering (IcheaP, 2009)

Douglas, J.M. Conceptual Design of chemical Processes. McGraw-Hill. 1988

Elliott, L; Ingham, D.; Kyne, A.; Mera, N.; Porkashanian, M. and Whittaker, S., Reaction mechanism reduction and optimization for modeling aviation fuel oxidation using standard and hybrid genetic algorithms. Computers and Chemical Engineering, 30, 889–900.. 2006

Exler, O., Antelo, L., Egea, J., Alonso, A. Banga, J.. A Tabu search-based algorithm for mixed-integer-nonlinear problems and its applications to integrated process and control system design. Computers and Chemical Engineering, 32, pp 1877-1891. 2008

Fisher, W.R., Doherty, M.F., y Douglas, J.M.. The Interface Between Design and Control
1. Process Controllability. *Ind. Eng. Chem. Res.*, **27**, 597-605. 1988

Fjeld, M., Asbjørnsen, O.A.y Astrom, K. Reaction invariants and their importance in the analysis of eigenvectors, state observability and controllability of the continuous stirred tank reactor. *Chem Eng. Science*. Vol. 29. 1974

Floudas, C. . Nonlinear and Mixed-Integer Optimization. Oxford University Press, New York. 1995

Fogel, D. Editor. Proceedings of the IEEE International Conference on Evolutionary Computation, IEEE Press, Piscataway. NY. 1996

Francisco, M., Revollar S., Lamanna R. y P. Vega. A Comparative Study of Deterministic and Stochastic Optimization Methods for Integrated Design of Processes. Proceedings 16th IFAC World Congress. Praga, 2005

Francisco, M. y Vega, P. Diseño Integrado de procesos de depuración de aguas utilizando Control Predictivo Basado en Modelos. *Rev. Iberoamericana de Automática e Informática Industrial*, Vol. 3, nº 4. pp 88-98. 2006

Francisco M., y Vega, P. Sintonía Automática De Controladores Predictivos Con Estabilidad Garantizada. Jornada de Automatica. Valladolid 2009.

Francisco, S. Revollar, P. Vega, R. Lamanna. Simultaneous Synthesis, Design and Control of Processes Using Model Predictive Control. International Symposium on Advanced Control of Chemical Processes (Adchem, 2009)

Gani,R.; O'Connell,J.P y Lewin, D.R. Relationship between thermodynamic properties and process control. ECCE1, Florence, Italy, May 4-7 pp. 1929-1932. 1997

Gen, M. y R. Chen. Genetic algorithms and engineering optimisation. J. W. and Sons. 2000

Geoffrion, A. Generalized bender decomposition. *Journal of Optimization Theory and Applications*, 10(4), 237–260.1972

Govatsmark, M., y Skogestad, S. Optimal number of stages in distillation with respect to controllability. In J. Grievink & J. Schijndel (Eds.), *ESCAPE-12* (Vol. 12, pp. 499–504). Elsevier Science. 2002

Goldberg, D.F Genetic Algorithm in Search, Optimization and Machine Learning. Addison-Wesley, Reading, MA. 1989

Grossmann, I.E. y D.A.Straub. Recent Developments in the Evaluation and Optimization of Flexible Chemical Processes. Computer-Oriented Process Engineering. Edited by L. Puigjaner and A. Espuna. Amsterdam. 1991

Gutiérrez, G. y P. Vega, , Process synthesis applied to activated sludge processes: A framework with MINLP optimisation models. 15th IFAC Triennial World Congress. Barcelona. 2002a

Gutiérrez, G. y P. Vega, , Integrated design of chemical processes and their control system including closed loop properties for disturbances rejection. 15th IFAC Triennial World Congress. Barcelona. 2002b

Hedrick, J.K. y Girard, A. Control of nonlinear dynamic systems: Theory and applications. Berkeley Press. 2005

Hermann, R. y Krener, A.J. . Nonlinear Controllability and Observability. *IEEE Trans. Aut. Control*, 5. 1977.

Holland, J. . Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor, MI 1975

Jorgesen, S., Gani, R., Andersen, T. R. Towards integration of controllability into plant design. Proceedings of the 7th Mediterranean Conference on Control and Automation (MED99). Haifa, Israel. 1999

Halvarsson, B., , Applications of Coupling Analysis on Bioreactor Models. MSc thesis. Uppsala Universitet. Sweden. 2003

Holt B.R. y M. Morari. Design of resilient processing plants -VI. The effect of right half plane zeros on dynamic resilience. Chem. Eng. Sci., 40, 1985.

Holt B.R.y M. Morari. Design of resilient processing plants -V. The effect of dead time on dynamic resilience. Chem. Eng. Sci., 40, 1985.

Hopkins,L.; Lant,P. y Newell,B..Output structural controllability: A tool for integrated process design and control. J. Proc.Cont. 8 pp. 57-88. 1998

Ierapetritou, M.G., J. Acevedo y E.N. Pistikopoulos. An Optimization Approach for Process Engineering Problems Under Uncertainty. Comput. Chem. Engug. 20, 703-710. 1996

Ingildsen, P. Realising Full-Scale Control in Wastewater Treatment Systems Using In Situ Nutrient Sensors. PhD thesis. Lund University, Sweden. 2002,

Knobelsdorf Miranda, J. Eliminación biológica de nutrientes en un ARU de baja carga orgánica mediante el proceso VIP . Escola Tècnica Superior D'enginyers De Camins Canals I Ports De Barcelona. 2005

Kookos, I. y J. Perkins, An algorithm for simultaneous process design and control, Ind. Eng. Chem. Res. 40, 4079. 2001

Kookos, I. K.; Perkins, J. D. An algorithm for simultaneous process design and control. Ind. Eng. Chem. Res., 40, 4079–4088. 2001

Koza, J. Genetic Programming. MIT Press, Cambridge, MA, 1992.

Kuo, B., Automatic control systems. Fifth edition. Prentice-Hall International Editions. 1987

Lamanna, R., Vega, P., Revollar, S., Alvarez, H. Diseño Simultáneo de Proceso y Control de una Torre Sulfitadora de Jugo de Caña de Azúcar. Revista Iberoamericana de Automática e Informática Industrial (RIAI). Vol. 6, Nro. 3, pp. 32 – 43. 2009

Loeblein, C., y Perkins, J. D. Structural design for on-line process optimization: I. Dynamic economics of mpc. *AIChE J.*, 45(5), 1018–1029. 1999

Luyben,M.L. y Floudas,C.A.. Analyzing the Interaction of design and Control- 1. A Multiobjective Framework and Application to Binary Distillation Synthesis. *Comp. Chem. Eng.* 18 pp. 933-969. 1994

Maciejowsky, J. M. Predictive Control with Constraints. Prentice Hall, 2002.

Marlin T.E. y A.N. Hrymak. Real-time operations optimization of continuous processes. In J.C. Kantor, C.E. Garcia, y B. Carnahan, editors, Proc. of Fifth International Conference on Chemical Process Control, volume 93 of AIChE Symposium Series No. 316, pages 156-164. AIChE and CACHE, 1997.

Meeuse, F., y Grievink, J.. Optimum controllability of distributed systems based on non-equilibrium thermodynamics. In J. Grievink & J. V. Schijndel (Eds.), *ESCAPE* (Vol. 12, pp. 259–264). Elsevier. 2002

Mohideen, M., Perkins, J. D., & Pistikopoulos, E. N. Optimal design of dynamic systems under uncertainty. *AIChE J.*, 42(8), 2251–2272. 1996

Mohideen, M., Perkins, J. D., & Pistikopoulos, E. N. Optimal synthesis and design of dynamic systems under uncertainty. *Computers and Chemical Engineering*, 20(Suppl), S895–S900. 1996

Mohideen, M., Perkins, J. D., & Pistikopoulos, E. N. Integrated framework for design and control. In *Proceedings of the UKACC International Conference on CONTROL '96* (Vol. 427, pp. 918–923), UKACC. 1996

Mohideen, M.J., J.D. Perkins y E.N. Pistikopoulos. Optimal Design of Dynamic Systems Under Uncertainty. *AIChE J.*, 42(8), 2251-2272. 1996

Morari, M: Design of Resilient Processing Plants-III. A General Framework for the Assessment of Dynamic Resilience. *Chem. Engng. Sci.* 38 pp. 1881-1891. 1983

Narraway, L. y Perkins, J. D. Selection of process control structure based on linear dynamic economics. *Industrial & Engineering Chemistry Research*, 32(11), 2681–2692. 1993

Narraway and Perkins, Selection of process control structure based en economics. *Comp. Chem.. Eng.*, 18, S511-S515.1994

Pantelides, C. C.. Modeling, simulation and optimisation of hybrid processes. *Proceedings Workshop Analysis and Design of Event-Driven Operations in Process Systems*, Imperial College, London, 10.-11. 1995

Perkins J.D. y S.P.K. Walsh. Optimization as a tool for design/control integration. *Comput. Chem. Eng.*, 20(4):315-323, 1996.

Petersen K., Jorgensen S.B., y Skogestad S. A Review on Process Controllability, CAPE.Net Status Paper, 1998.

Rechemberg, I.. Evolutionstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution, Frommann-Holzboog, Stuttgart, Germany. 1973

Revollar S., Lamanna R. y Vega P. Aplicación de los Algoritmos Genéticos en la Solución del Diseño Integrado de Sistemas de Reactores Químicos. *Proceedings IBERAMIA-2002*. F. Garijo, J.C. Riquelme, M. Toro (editors), pp. 239-248. Sevilla, Spain. 2002

Revollar S., Lamanna R. y Vega P. Síntesis y diseño integrado de un sistema de reactores químicos con criterios de lazo abierto. IV Congreso de Automatización y Control, CAC'2003. Mérida, Venezuela. 2003

Revollar, S, Vega, P. y R. Lamanna, Algorithmic synthesis and integrated design of chemical reactor systems using genetic algorithms. *Proceedings of World Automation Conference Vol 17*, 493. Sevilla. 2004

Revollar, S., Lamanna, P. Vega, M. Libretti, L. Lozada, J. Náñez. Aplicación de un Algoritmo Evolutivo en el Diseño Integrado de una Planta de Fangos Activados. V Congreso de Automatización y Control, CAC05. 2005

Revollar S., Lamanna R. Vega P. Algorithmic Synthesis for Activated Sludge Processes Using Genetic Algorithms. Proceedings 7th World Congress of Chemical Engineering. Glasgow, Julio 2005.

Revollar, S., Lamanna y P. Vega. Genetic Algorithms for Simultaneous Design and Control of Continuous Stirred Tank Reactor System. CPC7. Lake Louise. Canada, 2006.

Revollar, S., D. Dominguez, Z. Ramírez, H. Alvarez, R. Lamanna, P. Vega. Diseño Integrado de la Planta de Sulfitacion en un Ingenio Azucarero. XIII Congreso Latinoamericano de Control Automático/VI Congreso Venezolano de Automatización y Control. Mérida-Venezuela Noviembre de 2008.

Revollar, S., M. Francisco, P. Vega, R. Lamanna. Genetic Algorithms for the Synthesis and Integrated Design of Processes Using Advanced Control Strategies. International Symposium on Distributed Computing and Artificial Intelligence 2008 (DCAI'08).

Revollar, S., M. Francisco, P. Vega, R. Lamanna. Stochastic Optimization for the Simultaneous Synthesis and Control System Design of an Activated Sludge Process. Latin American Applied Research. 40: 137-146. 2010.

Revollar, S., R. Lamanna, P. Vega, M. Francisco. Simultaneous Synthesis, Design and Control of Processes Using Model Predictive Control. Multi-objective Genetic Algorithms for the Integrated Design of Chemical Processes using Advanced Control Techniques. 20th European Symposium on Computer Aided Process Engineering (ESCAPE20). Julio 2010.

Ricardez Sandoval, L. A.; Budman, H. M.; Douglas, P. L. Simultaneous design and control of processes under uncertainty: A robust modelling approach. *J. Process Control*, 18, 735–752. 2008

Ricardez-Sandoval, L. A.; Budman, H. M.; Douglas, P. L. Integration of design and control for chemical processes: A review of the literature and some recent results. *Annu. Rev. Control*, 33, 158–171. 2009

Ricardez-Sandoval, L. A.; Budman, H. M.; Douglas, P. L. Simultaneous design and control of chemical processes with application to the Tennessee Eastman process. *J. Process Control*, 19, 1377–1391. 2009

Sakizlis, V., Perkins, J. D., y Pistikopoulos, E. N. Parametric controllers in simultaneous process and control design optimization. *Industrial & Engineering Chemistry Research* 42(20), 4545–4563. 2003

Sakizlis, V.; Perkins, J. D.; Pistikopoulos, E. N. Recent advances in optimization-based simultaneous process and control design. *Comput. Chem. Eng.*, 28, 2069–2086. 2004

Samuelson, P.. Control of Nitrogen Removal in Activated Sludge Processes. Uppsala Dissertations from the Faculty of Science and Technology (PhD). Uppsala University. Sweden. 2005

Samuelsson, P., Halvarsson, B, Carlsson, B., Cost-Efficient Operation of a Denitrifying Activated Sludge Process – An initial Study. Technical reports from the Department of Information Technology. Upsala Universitet. 2005

Schweiger, C. A.; Floudas, C. A. Interaction of design and control: optimization with dynamic models. In *Optimal Control*; Hager, W. H., Padarlos, P. M., Eds.; Kluwer Academic Publishers: The Netherlands,; pp 388-435. 1998

Seferlis, P., y Grievink, J. Process design and control screening based on economic and static controllability criteria. *Computers and Chemical Engineering*, 25, 177–188. 2001

Skogestad S. y M. Morari. Design of resilient processing plants -IX. The effect of model uncertainty on dynamic resilience. *Chem. Eng. Sci.*, 42, 1987.

Skogestad,S. y Larsson T.. A review of Plantwide Control. Internal report EU CAPENET project (TWG 5). 1998

Summanwar, V., Jayaraman, V., Kulkarni, H., Kusumakar, H., Gupta, K. And J. Rajesh.. Solution of constrained optimization problems by multiobjective genetic algorithm. *Computers and Chemical Engineering*, 26, 1481-1492. 2002

Swaney R.E. y I.E. Grossmann. An index for operational flexibility in chemical process design. Part II. Computational algorithms. AIChE J., 31, 1985.

Swaney R.E.y I.E. Grossmann. An index for operational flexibility in chemical process design. Part I. Formulation and theory. AIChE J., 31, 1985.

Swartz, C. L. E. The use of controller parametrization in the integration of design and control. In *The Integration of Process Design and Control*; Seferlis, P., Georgiadis, M. C., Eds.; Elsevier: Amsterdam. pp 239-263. 2004.

Tlacuahuac-Flores, A y Biegler, L., "Integrated Control and Process Design During Optimal Polymer Grade Transitions Operations". *Computers and Chemical Engineering*, 31, 11, pp 2823-2837. 2008

Tsai, M. y Ch. Chang, Water usage and treatment network design using genetic algorithms Ind. Eng. Chem. Res. 40, 4874. 2001

Vega, P. Formulaciones adaptivas para reguladores PID. Tesis Doctoral. Universidad de Valladolid. 1987

Vega, P., Francisco, M., and E. Sanz, Norm based approach for automatic tuning of Model Predictive Controllers. Proceedings of ECCE-6, Copenhagen. 2007

Vu T.T.L. , P.A. Bahri, y J.A. Romagnoli. Operability considerations in chemical processes: A switchability analysis. Comput. Chem. Eng., 21, Suppl:S143-S148, 1997.

Walas, S. Chemical Process Equipment. Selection and Design. Butterworth-Heinemann, USA. 1990

Walsh, S., & Perkins, J. D. Application of integrated process and control system design to waste water neutralisation. Computers and Chemical Engineering, 18(Suppl), S183–S187. 1994

Walsh, S., & Perkins, J. D. Operability and control in processsynthesis and design. In Anderson J. L. (Ed.), Advances in chemical engineering—Process synthesis (pp. 301–402). Academic Press. 1996

Weitz O. y D.R. Lewin. Dynamic controllability and resiliency diagnosis using steady state process Howsheet data. Comp'l't. Chem. Eng., 20(4):325-355, 1996.

White V., J.D. Perkins, y D.M. Espie. Switchability analysis. Comput. Chem. Eng., 20(4):469-474, 1996.

Wright, Alden H. Genetic algorithms for real parameter optimization. Foundations of Genetic Algorithms 1991

Wolff, E., Estudies on control of integrated plants. PhD Thesis University of Throndheim. The Norwegian Institute of Technology. 1994

APENDICE I

ANÁLISIS DE CONTROLABILIDAD EN DISEÑOS PRELIMINARES

En una etapa preliminar a la síntesis y diseño integrado de la planta depuradora se realizó una evaluación de la controlabilidad y del comportamiento dinámico de diseños ya establecidos. Esto se llevó a cabo con la intención de definir dos aspectos muy importantes para el proceso de síntesis y diseño integrado.

- La técnica de diseño de controladores a usar
- Rangos de los índices de controlabilidad para emplearse posteriormente.

Se escogieron cinco diseños con características distintas para tener un conocimiento más amplio del comportamiento de la planta en distintas condiciones, entre ellas se encuentran 3 plantas con 2 reactores y un decantador y 2 plantas con un reactor y un decantador. Las plantas empleadas en esta sección son producto de la optimización por medio de algoritmos genéticos desarrollada por Revollar et al (2004).

Para el análisis de controlabilidad se siguió el siguiente procedimiento:

- Obtener el modelo linealizado del proceso.
- Cálculo de la matriz de las ganancias relativas (RGA).
- Cálculo del número de la condición de la planta (γ) y condición ante la perturbación (γ_p).
- Cálculo de los controladores SS y SIMC.
- Simulación y evaluación del desempeño dinámico del sistema a lazo cerrado, mediante los siguientes índices:
 - Pico de resonancia de la función de sensibilidad $S(j\omega)$ (M_s)
 - Integral del error cuadrático y absoluto (ISE e IAE)
 - Energía de bombeo (PE), utilizada como índice para medir el efecto sobre la señal de control.

A continuación se mostrará, el análisis detallado de controlabilidad para una planta de las plantas seleccionadas de manera de mostrar el procedimiento empleado en cada una de sus etapas. Seguidamente se presentan los resultados compilados para el resto de los casos de estudio.

Caso 1: Planta con 2 reactores

Tabla 1. Valores de las variables de la planta AG1

X1	X2	S1	S2	C1	C2	Xd	Xb	Xr
1677	1663	98	30	4.95	4.37	140	947	6273

Qr1	Qr2	Qp	Fk1	Fk2	V1	V2	A	y
351	380	16.6	0.48	0.52	5199	3457	2784	1

AM =	X1	S1	C1	X2	S2	C2	Xd	Xb	Xr
	-0.2929	0.3607	0	0	0	0	0	0	0.0675
	-0.0445	-0.9001	0	0	0.0675	0	0	0	0
	0	-0.0001	-0.6536	0	0	0	0	0	0
	0.4776	0	0	-0.5839	0.666	0	0	0	0.1099
	0	0.4776	0	-0.0153	-1.347	0	0	0	0
	0	0	0	0	-0.0001	-0.9515	0	0	0
	0	0	0	0	0	0	-1.4888	0.2305	0
	0	0	0	0.7295	0	0	2.5165	-1.0843	0
	0	0	0	0	0	0	0	1.2465	-0.5371

BM =	Qr1	Qr2	Qp	Fk1	Fk2
	0.884	0	0	0	0
	-0.0131	0	0	0	0
	-0.001	0	0	2.1328	0
	0.004	1.3335	0	0	0
	0.0197	0	0	0	0
	-0.0013	-0.0013	0	0	2.5351
	0	0	-0.1485	0	0
	0.25	0.25	0	0	0
	-3.8118	-3.8118	-3.8118	0	0

$$B_p = \begin{array}{c|cc} & Si & Qi \\ \hline 0 & -0.3072 \\ 0.25 & 0.0517 \\ 0 & -0.0010 \\ 0 & 0.0040 \\ 0 & 0.0197 \\ 0 & -0.0013 \\ 0 & 0.1485 \\ 0 & 0.2500 \\ 0 & 0 \end{array}$$

$$C = \begin{array}{c|ccccccc|c} X_1 & S_1 & C_1 & X_2 & S_2 & C_2 & X_d & X_b & X_r \\ \hline 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{array}$$

D=0

e) *Cálculo de la RGA*

Para el cálculo de la Matriz de las Ganancias Relativas (RGA) es preciso calcular antes la matriz de transferencia del sistema. (Ecuación 2.23)

$$G = \begin{array}{c|ccccc|c} Qr1 & Qr2 & Qp & Fk1 & Fk2 & & \\ \hline -0.09 & 0.0645 & 0.6819 & 0 & 0 & & S1 \\ -0.0015 & 0 & 0 & 3.2633 & 0 & & C1 \\ 0.0622 & 0.1625 & -21.4669 & 0 & 0 & & X2 \\ -0.018 & 0.021 & 0.4862 & 0 & 0 & & S2 \\ -0.0013 & -0.0013 & 0 & 0 & 2.6643 & & C2 \end{array}$$

Resultando de este modo la siguiente RGA :

$$RGA = \begin{array}{c|ccccc|c} Qr1 & Qr2 & Qp & Fk1 & Fk2 & & \\ \hline 2.1854 & -1.0532 & -0.1322 & 0 & 0 & & S1 \\ 0 & 0 & 0 & 1 & 0 & & C1 \\ 0.0485 & 0.2342 & 0.7173 & 0 & 0 & & X2 \\ -1.2339 & 1.819 & 0.4149 & 0 & 0 & & S2 \\ 0 & 0 & 0 & 0 & 1 & & C2 \end{array}$$

Tal como se puede observar en la matriz obtenida se cumple que la sumatoria de los elementos de cada fila es igual a 1. Respecto a la interacción entre los posibles pares de variable controlada-manipulada se puede identificar que existe un claro acoplamiento entre las

concentraciones de Oxígeno y sus respectivos factores de aireación, sin recibir influencia de las variables restantes. Esto se concluye a través del coeficiente que relaciona cada uno de los pares.

Esta independencia entre los entre los pares de variables no se encuentra para las demás involucradas, se puede ver como cada una de ellas posee una influencia importante sobre las demás, impidiendo de esta forma que se puedan establecer lazos de control independientes entre si. Por este motivo, y debido a la importancia que posee la concentración de sustrato de salida (en este caso S2), se establecerá esta variable con el caudal de reflujo 1 (Qr1) como su pareja asociada.

Por lo tanto, de aquí en adelante sólo se considerarán estas variables para los estudios de caracterización de la planta y por consiguiente el estudio de las perturbaciones también estará referido a dichas variables.

De este modo se obtiene la siguiente matriz propia, incluyendo además las columnas derivadas del efecto de las perturbaciones:

$$G_0 = \begin{array}{ccccc} & \text{Qr1} & \text{Fk1} & \text{Fk2} & \text{Si} \\ \hline & -0.018 & 0 & 0 & 0.035 \\ & -0.0015 & 3.2633 & 0 & 0 \\ & -0.0013 & 0 & 2.6643 & 0 \end{array} \quad \begin{array}{c} \text{Qi} \\ \hline 0.0584 \\ -0.0015 \\ -0.0013 \end{array} \quad \begin{array}{c} \text{S2} \\ \text{C1} \\ \text{C2} \end{array}$$

De esta matriz se extrae la matriz propia reducida de la planta:

$$G_1 = \begin{array}{ccc} & \text{Qr1} & \text{Fk1} \\ \hline & -0.018 & 0 \\ & -0.0015 & 3.2633 \\ & -0.0013 & 0 \end{array} \quad \begin{array}{c} \text{Fk2} \\ \hline 0 \\ 0 \\ 2.6643 \end{array} \quad \begin{array}{c} \text{S2} \\ \text{C1} \\ \text{C2} \end{array}$$

Y los vectores correspondientes a cada perturbación:

$$GdS = \left| \begin{array}{c|c} \text{Si} & \\ \hline 0.035 & \text{S2} \\ 0 & \text{C1} \\ 0 & \text{C2} \end{array} \right| \quad GdQ = \left| \begin{array}{c|c} \text{Qi} & \\ \hline 0.0584 & \text{S2} \\ -0.0015 & \text{C1} \\ -0.0013 & \text{C2} \end{array} \right|$$

Con la matriz propia reducida podemos recalcular una nueva RGA ahora con el juego nuevo de variables:

	Qr1	Fk1	Fk2	
RGA2 =	1	0	0	S2
	0	1	0	C1
	0	0	1	C2

Tal como lo refleja la matriz RGA2, existe un acoplamiento perfecto entre las variables seleccionadas, por lo tanto las mismas se emplearán para el cálculo posterior de las condiciones. Las variables S2 - Qr1, se emplearán como la pareja variable controlada-variable manipulada en los cálculos del controlador SIMC que se efectuarán a más adelante. En el caso del controlador SS se miden todas las variables de salida, pero igualmente dicha pareja es la preponderante.

Una vez establecidas las matrices con las cuales se trabajará a lo largo de esta investigación, se procede al cálculo de la Condición de la planta y la Condición de la planta ante la perturbación.

Antes de calcular las diversas condiciones, es importante tomar en cuenta que las mismas son dependientes del escalamiento (De Prada et. al, 2006) y para ello Skogestad (2005) plantea la siguiente metodología a través de una matrices de escalamiento.

En primer lugar se establecen dichas matrices:

Para las salidas se plantea una matriz basada en el error o desviación máxima permitida para cada variable.

Dy = De =	50	0	0	S2
	0	1.5	0	C1
	0	0	1	C2

Para las entradas se plantea una matriz formada por los valores máximos permitidos de cada variable.

$$Du = \begin{array}{ccc|c} & 3500 & 0 & 0 \\ & 0 & 1 & 0 \\ & 0 & 0 & 1 \end{array} \quad \begin{array}{l} Qr1 \\ Fk1 \\ Fk2 \end{array}$$

Finalmente para las perturbaciones se emplea el valor máximo probable que pudiese llegar a alcanzar dicha perturbación.

$$DdS = 550 \quad DdQ = 2100$$

Establecidas dichas matrices de escalamiento, se procede a calcular las matrices de del proceso ya escaladas:

$$\begin{aligned} G1s &= De^{-1} \cdot G1 \cdot Du \\ GdSs &= De^{-1} \cdot GdS \cdot DdS \\ GdQs &= De^{-1} \cdot GdQ \cdot DdQ \end{aligned} \quad (1)$$

Resultando la matriz de la planta

$$G1s = \begin{array}{ccc|c} & Qr1 & Fk1 & Fk2 \\ & -1.2613 & 0 & 0 \\ & -3.4068 & 2.1755 & 0 \\ & -4.6536 & 0 & 2.6643 \end{array}$$

Y las perturbaciones:

$$GdSs = \left| \begin{array}{c} Si \\ 0.3853 \\ -0.0108 \\ -0.0044 \end{array} \right| \quad GdQs = \left| \begin{array}{c} Qi \\ 2.4534 \\ -2.0455 \\ -2.8039 \end{array} \right|$$

Una vez obtenidas dichas matrices, se procede al cálculo de las condiciones respectivas:

f) Condición de la planta

Los valores singulares de $G1s$ son

$$\sigma = \left| \begin{array}{c} 6.3967 \\ 2.3519 \\ 0.4860 \end{array} \right|$$

De los cuales se extrae la condición de la planta a través de la ecuación 2.19

$$\gamma = 13.1631$$

g) Condición ante la perturbación

Se calcula la condición ante cada una de las perturbaciones tal como lo establece la ecuación 2.22 y se elige a la mayor de ellas

$$\gamma_p = \begin{vmatrix} 12.9963 \\ 9.4557 \end{vmatrix}$$

$$\gamma_p = 12.9963$$

Como se puede apreciar, la condición de la planta resultó un valor bajo, lo cual indica la bondad de dicha planta, sin embargo a partir de la condición de la perturbación se puede señalar que las mismas tendrán un efecto bastante notable en el funcionamiento de la planta dado que los mismos son elevados y en el caso de la mayor (correspondiente a la entrada de sustrato, Si) es casi igual al valor de la condición de la planta lo cual indica que para la planta será muy difícil rechazar dichas perturbaciones a lazo abierto.

h) Parámetros de lazo cerrado

Una vez concluidos los cálculos de los parámetros de lazo abierto, se proceden a calcular los diversos indicadores de lazo cerrado.

Para dichos indicadores, es necesario contar con un controlador. El primero corresponde al controlador calculado mediante la metodología SIMC y el segundo en la retroalimentación de estados (SS).

Una vez calculados los diversos controladores, se visualiza la respuesta simulada a través del paquete SFA y se calculan los índices respectivos (IAE, ISE, PE), para los casos con controladores PID (SIMC y el propuesto junto con la planta) se calcula adicionalmente el parámetro M_s^1 .

Respecto a la respuesta simulada, y a los parámetros extraídos de la misma, caben mencionar algunos aspectos importantes para su mejor compresión:

- En primer lugar, se deja alcanzar el estado estacionario con las entradas constantes.
- A partir de $t = 500$ h se introducen simultáneamente tanto el caudal como el sustrato inicial perturbado.
- De esta zona es de donde se extrae el comportamiento durante 1 ciclo completo de perturbaciones (168 h) y se calculan los respectivos parámetros.

La Tabla 2 muestra un resumen de los diferentes índices para cada caso y seguidamente en

Tabla 2. Índices de controlabilidad para la planta AG1

Parámetros de Lazo Abierto			
γ			13.1631
γ_p			12.9963
Parámetros de Lazo Cerrado			
	AG	SIMC	SS
Ms	1.1806	1.4296	
IAE	1778.1	1451	1422.5
ISE	28503	21126	17580
PE	680.5	1154.6	899.7
			1.9664
			-6.9059
			-0.0277
			0.1209
K	-9.77	-9.6176	8.8342
			-4.7192
			-0.1864
			-0.3956
			0.5170
Ti (o Ki)	54.4	1.5668	-1.0841
Td	0	0.3888	0