







**Universidad de Valladolid**



**ESCUELA DE INGENIERÍAS  
INDUSTRIALES**

**UNIVERSIDAD DE VALLADOLID**

**ESCUELA DE INGENIERIAS INDUSTRIALES**

**Grado en Organización Industrial**

**DISEÑO DE SENSORES SOFTWARE CON  
TÉCNICAS BASADAS EN INTELIGENCIA  
COMPUTACIONAL Y ESTADÍSTICA**

**Autor:**

**Zapatero Campo, Javier**

**Tutor:**

**De la Fuente Aparicio, María Jesús  
Departamento de Ingeniería de  
Sistemas y Automática**

**Valladolid, Julio 2017.**





# ÍNDICE DE CONTENIDOS

<b>1 INTRODUCCIÓN Y OBJETIVOS DEL TRABAJO</b>	<b>9</b>
1.1 Introducción	10
1.2 Objetivos del trabajo	11
1.3 Organización del trabajo	12
<b>2 CONTROL ESTADÍSTICO DE PROCESOS</b>	<b>15</b>
2.1 Introducción	16
2.2 Variabilidad de un proceso	17
2.2.1 Causas comunes	19
2.2.2 Causas especiales	21
2.3 Proceso de inspección	22
2.3.1 Mediciones de la calidad	22
2.3.2 Muestreo	23
2.3.2.1 Distribuciones de muestreo	24
2.3.2.2 Gráficas de control	25
2.4 Control estadístico de procesos univariante	29
2.5 Detección y diagnóstico de fallos	31
2.5.1 Detección de fallos	31
2.5.2 Diagnóstico de fallos	32
2.5.3 Métodos de FDD existentes	32
2.5.3.1 Métodos basados en señal	33
2.5.3.2 Métodos analíticos	35
2.5.3.3 Métodos basados en conocimiento	36
2.5.4 Requerimientos de diseño	39
2.6 Mínimos cuadrados parciales (PLS)	42
2.6.1 Regresión PLS univariada	42
2.6.1.1 Cálculo de las estadísticas $T^2$ y $Q$	46
2.6.2 Regresión PLS dinámica (DPLS)	49
2.6.3 Regresión PLS no lineal (NPLS)	50
2.6.4 Regresión PLS con extracción de características (PLSEX)	50
2.6.4.1 Método de correlación	51



<b>2.7 Redes neuronales artificiales (RNA)</b>	<b>53</b>
2.7.1 Neuronas	53
2.7.2 Estructura de las RNA	57
2.7.2.1 Una capa de neuronas	58
2.7.2.2 Múltiples capas de neuronas	59
2.7.3 Tipos de redes neuronales	61
2.7.4 Aprendizaje de las RNA	62
2.7.5 Perceptrón	63
2.7.5.1 Perceptrón multicapa	65
2.7.6 Ventajas de las RNA	66

<b>3 DESCRIPCIÓN DE LA PLANTA</b>	<b>69</b>
-----------------------------------	-----------

<b>3.1 Introducción</b>	<b>70</b>
<b>3.2 Azúcar de remolacha</b>	<b>70</b>
3.2.1 Historia del azúcar	71
<b>3.3 Proceso productivo del azúcar de remolacha</b>	<b>75</b>
3.3.1 Recolección y entrega de la remolacha	76
3.3.2 Almacenamiento y transporte	76
3.3.3 Lavado de la remolacha	77
3.3.4 Extracción del jugo	78
3.3.5 Tratamiento de la pulpa	78
3.3.6 Depuración y filtrado del jugo	79
3.3.7 Evaporación del jugo	79
3.3.8 Generación de vapor y electricidad	79
3.3.9 Cristalización y centrifugado	80
3.3.10 Secado del azúcar	80
3.3.11 Producción de melazas	81
3.3.12 Almacenamiento y embalaje	81
3.3.13 Tratamiento de aguas	81
<b>3.4 Proceso de evaporación</b>	<b>81</b>
3.4.1 Descripción general	82
3.4.2 Descripción de los equipos utilizados	83
3.4.2.1 Evaporadores	83
3.4.2.2 Balones de expansión	87



3.4.2.3	Saturador	88
3.4.2.4	Otros equipos	88
3.4.3	Funcionamiento y lazos de control	88

<b>4 ANÁLISIS DE RESULTADOS</b>	<b>91</b>
---------------------------------	-----------

4.1	Introducción	92
4.2	Toma de datos de la planta	93
4.2.1	Variables del proceso	94
4.2.2	Grupos de datos	96
4.2.3	Fallos provocados en los datos	97
4.3	Estimación de la variable y DDF con PLS	98
4.3.1	Mínimos cuadrados parciales	99
4.3.2	Aplicación PLS genérico	102
4.3.2.1	Entrenamiento PLS	102
4.3.2.2	Prueba sin fallo PLS	107
4.3.2.3	Prueba con fallo PLS	112
4.3.3	Aplicación PLS EX	119
4.3.3.1	Entrenamiento PLS EX	119
4.3.3.2	Prueba sin fallo PLS EX	121
4.3.3.3	Prueba con fallo PLS EX	123
4.3.4	Aplicación PLS dinámico	127
4.3.4.1	Entrenamiento PLS dinámico	128
4.3.4.2	Prueba PLS dinámico	129
4.3.5	Aplicación PLS no lineal	133
4.3.5.1	Entrenamiento PLS no lineal	134
4.3.5.2	Prueba PLS no lineal	136
4.3.6	Combinaciones de métodos PLS	138
4.3.6.1	PLS dinámico EX	139
4.3.6.2	PLS dinámico no lineal	141
4.3.6.3	PLS no lineal EX	142
4.3.6.4	PLS dinámico no lineal EX	144
4.4	Estimación de la variable con RNA	147
4.4.1	Red neuronal	147

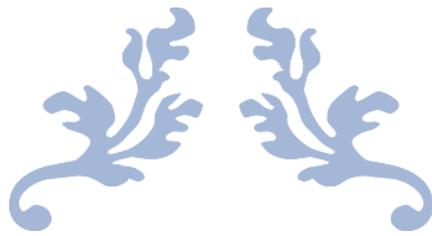


4.4.2	Red neuronal EX	152
4.4.3	Red dinámica	154
4.4.4	Red dinámica EX	159
4.5	Comparativa y conclusiones	162
4.5.1	Comparativa PLS	162
4.5.2	Comparativa RNA	164

<b>5 CONCLUSIONES Y FUTURAS LÍNEAS DE TRABAJO</b>	<b>165</b>
---	------------

5.1	Conclusiones	166
5.2	Futuras líneas de trabajo	169

<b>6 BIBLIOGRAFÍA</b>	<b>170</b>
-----------------------	------------



---

# INTRODUCCIÓN Y OBJETIVOS DEL TRABAJO

---

CAPÍTULO 1





## **1.1 INTRODUCCIÓN**

La seguridad en la producción es uno de los principales objetivos de la industria moderna. Las técnicas de control actuales permiten que los procesos funcionen con un alto grado de autonomía, sin embargo, la aparición de causas especiales en los procesos puede provocar que estos dejen de operar de manera óptima o incluso que la seguridad de los operarios, o del medio ambiente, se pueda ver comprometida. Por estos motivos, el diseño de sistemas de detección y diagnóstico de fallos pueden ser vistos como un paso hacia la fiabilidad y seguridad en la industria.

Otro de los objetivos fundamentales de la industria es la calidad, es decir, producir productos con la calidad adecuada para poder satisfacer las demandas de los clientes. Normalmente, esta medida de la calidad se evalúa midiendo algunas variables a la salida del proceso, sin embargo, el uso de técnicas para la estimación de las variables de calidad en línea mientras el proceso está funcionando, puede ser una herramienta útil para la predicción de la calidad, ya que, si esta predicción se realiza de manera temprana y se corrige, se pueden ahorrar muchas pérdidas en la producción.

Las técnicas estadísticas multivariantes son probablemente unas de las técnicas basadas en datos más utilizadas en la detección y diagnóstico de fallos en procesos. Dentro de estas, el análisis de componentes principales (PCA, por sus siglas en inglés) se ha aplicado en la monitorización de sistemas complejos mostrando resultados muy satisfactorios debido a su efectividad y su simplicidad. Sin embargo, los métodos de detección y diagnóstico basados en PCA pueden presentar problemas a la hora de monitorizar procesos con varios modos de operación. Además, la fase de aislamiento de fallos no se realiza de una forma completa.



PCA utiliza datos históricos del proceso para construir modelos empíricos del mismo que son capaces de extraer las principales tendencias del proceso. Esta extracción se basa en el estudio de las relaciones entre las variables medidas del proceso. De esta forma, cuando ocurre un fallo, se genera un cambio en estas relaciones que puede ser detectado. Otra de las técnicas estadísticas multivariantes ampliamente utilizadas en la industria es la regresión por mínimos cuadrados parciales (PLS, por sus siglas en inglés), que será la técnica utilizada para la realización de este trabajo, es un método de regresión basado en componentes principales. PLS también puede usarse tanto para el diseño de técnicas de detección y diagnóstico de fallos, como para la estimación de variables no medibles, es decir, en el diseño de sensores software.

Por otro lado, tenemos las redes neuronales artificiales (RNA) (también conocidas como sistemas conexionistas), que son modelos computacionales basados en un conjunto de neuronas simples, que trabajan de forma análoga al comportamiento observado en los axones de las neuronas en el sistema nervioso de los seres vivos. Estas redes neuronales artificiales son capaces de aproximar cualquier función no lineal, y en particular la relación entre las variables del proceso y la variable de calidad que se quiere estimar, por tanto, en este trabajo se utilizarán estas redes neuronales para diseñar otro sensor software, y poder comparar los resultados entre ambos métodos: PLS y RNA.

## **1.2 OBJETIVOS DEL TRABAJO**

El objetivo principal de este trabajo consiste en el diseño de sensores software para la estimación de variables de calidad de un proceso y al mismo tiempo para la detección y diagnóstico de fallos en procesos continuos mediante el uso de la regresión por mínimos cuadrados parciales y las redes neuronales artificiales.



La planta en la que se van a probar y utilizar estos métodos, concretamente en su sección de evaporación, es una planta azucarera real.

En primer lugar, se desarrollará un programa basado en el método PLS (*Partial Least Squares*) para la estimación de la variable de calidad y para la detección de fallos que se producen en la sección de evaporación de la planta azucarera, de forma que, en el momento en el que se produzca un error en alguna de las secciones del proceso, el sistema de detección de fallos detectará que se está trabajando en una situación anómala, e indicará al operario que se ha producido un fallo. Cuando esto ocurra, el programa indicará además en que instante se produjo el fallo y cuál es la variable que está provocando el fallo, la cual estará relacionada con un dispositivo específico, lo que facilitará la tarea de reparación al operador encargado del funcionamiento del proceso, y que haga que el proceso vuelva a funcionar bajo control estadístico, es decir, que siga funcionando en la zona de operación que asegura la calidad del producto obtenido. Después se harán modificaciones sobre el método PLS clásico, añadiéndole dinámica, haciéndole no lineal, etc. para ver su comportamiento en ambos aspectos: estimación de la variable de calidad y detección y diagnóstico de fallos. Posteriormente, los resultados obtenidos con PLS para la estimación de la calidad en el proceso, se compararán con los resultados obtenidos tras utilizar redes neuronales artificiales para tal fin, explicando las ventajas y/o puntos débiles de cada uno.

## **1.3 ORGANIZACIÓN DEL TRABAJO**

A continuación, se expondrán brevemente cada uno de los capítulos con los que va a contar este trabajo:

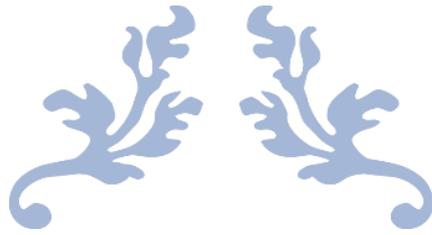


- **Capítulo 1: Introducción y objetivos del trabajo:** En el presente capítulo se introducen las razones que motivan la realización de este trabajo, las herramientas que van a utilizarse para su desarrollo, así como los objetivos fijados a alcanzar en el mismo. También se va a añadir un índice que muestre la organización de la memoria.
- **Capítulo 2: Control estadístico de procesos:** En este capítulo se tratará en profundidad el SPC (*Statistical Process Control*) desde sus orígenes hasta su finalidad, y también se desarrollarán los métodos utilizados para el diseño del sensor software para estimación de calidad, y para detección y diagnóstico de fallos. Estos son, PLS (*Partial Least Squares*) y derivados del mismo, y Redes Neuronales Artificiales.
- **Capítulo 3: Descripción de la planta:** En este capítulo se hará un repaso histórico de la producción de azúcar española, además de estudiar los elementos que componen la sección de evaporación de la planta sobre la cual se ha realizado este trabajo y su función dentro del proceso, lo cual es importante a la hora de determinar el elemento que necesita ser cambiado, reparado o recalibrado al detectar un fallo en el proceso.
- **Capítulo 4: Análisis de resultados:** En este capítulo se procederá al análisis de los resultados arrojados por cada método descrito anteriormente, al ser aplicados a los datos recogidos en la planta. En primer lugar, cada método constará de una fase de entrenamiento de datos, y posteriormente se probarán con nuevos datos y se compararán los resultados obtenidos entre ellos.
- **Capítulo 5: Conclusiones y futuras líneas de trabajo:** En este capítulo se analizarán todos los resultados obtenidos en los apartados anteriores a modo de



reflexión personal sobre el trabajo realizado. Del mismo modo, se expondrán futuras líneas de trabajo y mejora.

- **Capítulo 6: Bibliografía:** Por último, en este capítulo se presenta el listado de la bibliografía utilizada para la realización de este trabajo.



---

# CONTROL ESTADÍSTICO DE PROCESOS

---

CAPÍTULO 2





## **2.1 INTRODUCCIÓN**

En este capítulo hablaremos de los fundamentos de este trabajo, explicando punto por punto cada uno de los conceptos en los que se basa el diseño del sensor software, qué motiva la necesidad de diseñarlo, y qué técnicas se han utilizado para conseguir que funcione con éxito.

El Control Estadístico de Procesos (en inglés *Statistical Process Control, SPC*) se define como la capacidad de determinar el resultado de un proceso mediante técnicas estadísticas, y de esta manera hacerlo predecible a lo largo del tiempo. En otras palabras, SPC es un método para mejorar de forma continua la calidad de un proceso basado en la reducción sistemática de las causas de variabilidad que más influyen en el producto final.

En este contexto del control estadístico de procesos, un sistema se puede definir como un conjunto de componentes interactivos con un propósito particular. Un proceso es el conjunto de procedimientos que transforman las entradas en salidas con un cierto valor añadido. En este contexto, las variaciones pueden definirse como cambios en el valor de una característica medida del proceso. Este valor se considera óptimo cuando está dentro de un rango de tolerancias. La variación en esta característica del proceso, incluso dentro de los rangos de tolerancia, es la responsable de pérdidas económicas debido a la baja calidad de los productos. Sin embargo, es importante enfatizar que la variabilidad es algo inherente a los procesos, y es imposible eliminarla totalmente, pero puede ser reducida. Por tanto, el objetivo fundamental de la calidad es producir productos con la mínima variación posible. Hay por tanto que encontrar las fuentes de variabilidad de los procesos y eliminarlas para obtener productos con alta calidad.



Para lograrlo, utilizaremos herramientas que nos permitan distinguir entre causas comunes y causas especiales de variación. Una vez identificadas, debemos eliminar las causas especiales, ajenas al desarrollo natural del proceso, para ser capaces de anticiparnos al resultado al ser afectado únicamente por las causas comunes (aleatorias) de variación, las cuales conocemos. De esta manera, sabremos si el proceso está dentro de su variabilidad aleatoria, o, por el contrario, está fuera de control produciendo errores asignables a una causa determinada.

La Gestión de la Calidad Total (TQM) es un aspecto fundamental de numerosas empresas, incorporando métodos de mejora continua a sus procesos de producción.

Una vez seamos capaces de medir y comparar nuestros productos, podremos valorar el grado de conformidad alcanzado respecto a las especificaciones del cliente.

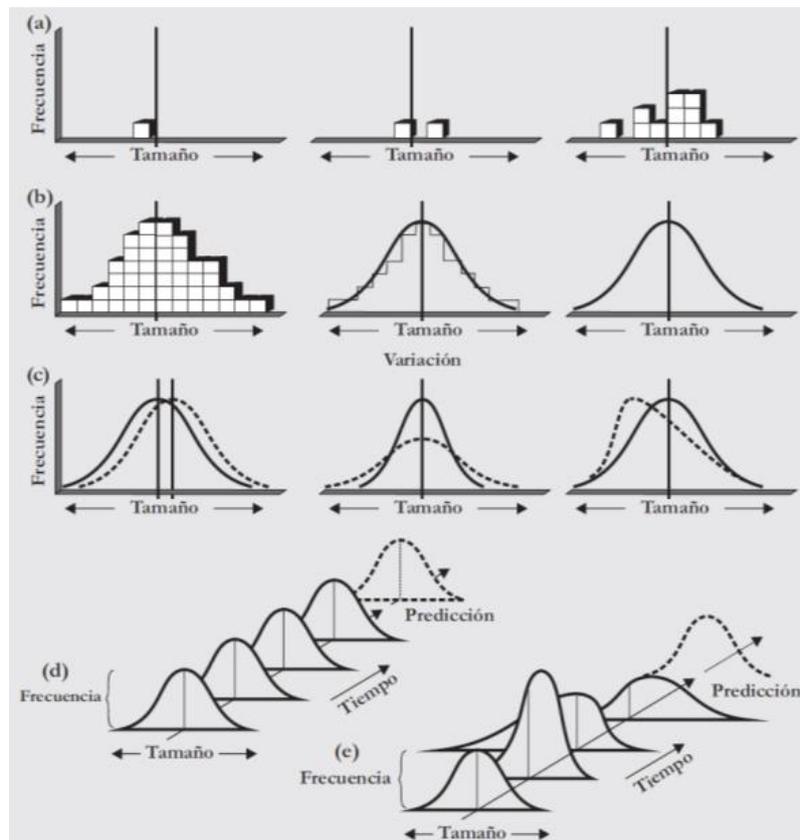
## **2.2 VARIABILIDAD DE UN PROCESO**

Los procesos de producción incluyen numerosas fuentes de variación, aunque estén desarrollándose de forma prevista, por tanto, podemos decir que no pueden existir dos productos o servicios idénticos.

Para determinar la variabilidad de un proceso, primero -ver figura 2.1 (a)- se toma una serie de pequeñas muestras y se las coloca en una escala proporcional (el eje horizontal). Después, en el eje vertical, se indica el número de veces que ocurrieron (frecuencia). Eventualmente, después de un número de muestras, se obtienen las distribuciones mostradas en la figura 2.1 (b). Estas distribuciones, como podemos apreciar en la figura 2.1 (c), difieren en relación a lo que revelaron las muestras. Si las causas



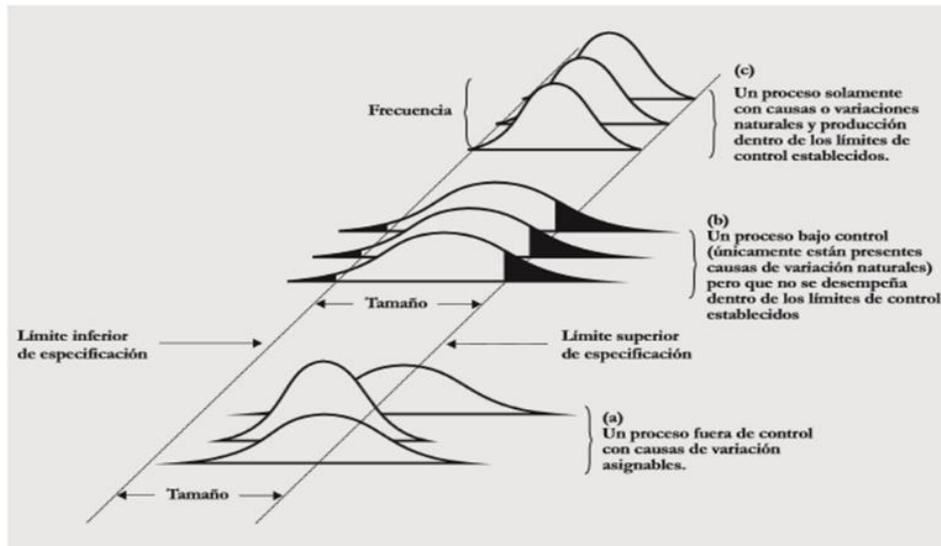
de variación encontradas son meramente naturales, las distribuciones que se obtienen tomarán una forma similar a las de la figura 2.1 (d). De otra forma, si las causas que aparecen pertenecen a variables especiales (causas no esperadas por parte del proceso), se obtendrán distribuciones inesperadas, como las que muestra la figura 2.1 (e).



**Figura 2.1:** Variaciones naturales y especiales. (a) Las muestras varían; (b) pero forman un patrón que, si es estable, es llamado distribución. (c) Las distribuciones pueden diferir en la medida de la tendencia central, variación, forma, o cualquier combinación de estas. (d) Si solo se presentan causas de variación naturales, la salida del proceso forma una distribución que es estable a lo largo del tiempo y es predecible. (e) Si se presentan causas de variación especiales, la salida del proceso no es estable a lo largo del tiempo y no es predecible. (Fuente: Control estadístico de procesos. Carro Paz, Roberto y González Gómez, Daniel)



El objetivo es, por tanto, eliminar las variaciones especiales y mantener el proceso bajo control. La figura 2.2 muestra tres tipos de salidas del proceso; en la figura 2.2 (a) podemos observar un proceso fuera de control (proceso salvaje); la figura 2.2 (b) muestra un proceso bajo control pero que no está dentro de los límites de control establecidos (proceso domado, bajo control pero incapaz); y la figura 2.2 (c) muestra un proceso bajo control y dentro de los límites establecidos (proceso entrenado).



**Figura 2.2:** Control del proceso: Tres tipos de salida del proceso. (Fuente: Control estadístico de procesos. Carro Paz, Roberto y González Gómez, Daniel)

## **2.2.1 CAUSAS COMUNES**

Existen dos categorías básicas de variación en los procesos: causas comunes y causas especiales. Las primeras son fuentes de variación puramente aleatorias, no identificables e imposibles de evitar mientras se utilice el procedimiento actual.

Su variabilidad es inherente al proceso, tienen carácter permanente, y su efecto da lugar a una pauta de variabilidad



estable o predecible, cuantificada por la capacidad del proceso. Si sólo hay causas comunes, el proceso está bajo control estadístico. Su solución pasa por modificar el sistema, lo cual incumbe a la Dirección.

Como hemos mencionado anteriormente, a partir de las muestras recogidas obtendremos una distribución, la cual se caracteriza por su media, su expansión y su forma.

1. La media es la suma de las observaciones dividida entre el número total de observaciones:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Donde:  $x_i$  = observación de una característica de calidad,  $n$  = número total de observaciones y  $\bar{x}$  = media.

2. La varianza es una medida de la dispersión de las observaciones en torno a la media. Dos medidas utilizadas comúnmente en la práctica son el rango y la desviación estándar. El rango es la diferencia entre la observación más grande contenida en una muestra y la más pequeña. La desviación estándar es la raíz cuadrada de una población, basada en una muestra, y se obtiene mediante la siguiente fórmula:

$$\delta = \sqrt{\frac{\sum x_i - \bar{x}}{n - 1}}$$

,o bien:

$$\delta = \sqrt{\frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n - 1}}$$



Donde:  $\delta$  = desviación estándar de una muestra,  $n$  = número total de observaciones de la muestra,  $\bar{x}$  = media y  $x_i$  = observación de una característica de calidad

Si el rango o la desviación estándar tienen valores relativamente pequeños, implica que las observaciones están agrupadas alrededor de la media.

3. Dos formas comunes en las distribuciones de procesos son la simétrica y la asimétrica o sesgada. Una distribución simétrica presenta el mismo número de observaciones ubicadas por encima y por debajo de la media. Sin embargo, una distribución asimétrica presenta un número mayor de observaciones que están ubicadas por encima o por debajo de la media.

Si la variabilidad del proceso se debe únicamente a causas comunes de variación, generalmente estaremos ante una distribución simétrica, donde la mayor parte de las observaciones se localiza alrededor del centro.

### **2.2.2 CAUSAS ESPECIALES**

En esta categoría se engloban las causas de variación no inherentes al proceso, que deben ser identificadas y eliminadas. Son de carácter puntual, y su efecto da lugar a una pauta de variabilidad errática o impredecible.

Si aparecen causas especiales, el proceso estará fuera de control estadístico. Únicamente estará bajo control estadístico cuando la localización, expansión o forma de su distribución no varía a lo largo del tiempo. Una vez un proceso está bajo control estadístico, podremos, mediante técnicas SPC, detectar cuándo



surgen causas especiales, y finalmente, eliminarlas. Generalmente pueden eliminarse mediante actuaciones locales a cargo de los operarios o encargados del proceso.

## **2.3 PROCESO DE INSPECCIÓN**

Un error comúnmente cometido por numerosas empresas es el tratar de separar las unidades defectuosas antes de que lleguen al cliente, ocasionándoles grandes pérdidas. Esto puede evitarse mediante una inspección previa combinada con control estadístico, para vigilar la calidad y ser capaces de detectar y corregir la presencia de anomalías en el proceso.

Para poder aplicar SPC a un proceso, debemos elegir cómo vamos a medir las características de calidad, qué tamaño de muestra se va a utilizar, y en qué fases del proceso será necesario realizar inspecciones.

### **2.3.1 MEDICIONES DE LA CALIDAD**

Para detectar anomalías en un producto, los encargados de realizar la inspección deben contar con la capacidad necesaria para medir los rasgos característicos de la calidad, la cual puede evaluarse de dos formas.

La primera consiste en **medir los atributos** o las características del producto o servicio que sean medibles rápidamente para saber si la calidad es aceptable. Este método permite a los inspectores decidir si un producto o servicio cumple o no con las especificaciones. Los atributos se usan frecuentemente cuando



las especificaciones de calidad son complicadas y la medición mediante variables resulta difícil o costosa.

La ventaja del recuento de atributos es que requiere de un menor esfuerzo y coste que el caso de medición de variables. La desventaja es que, aún cuando la información obtenida a través de los atributos puede ser suficiente para revelar que la calidad ha cambiado, no es de mucha utilidad para conocer la magnitud del cambio.

La segunda forma de evaluar la calidad, como hemos mencionado anteriormente, consiste en la **medición de variables**, es decir, las características del producto o servicio que podemos medir, como pueden ser el peso, el volumen, la longitud o el tiempo.

La ventaja que ofrece este método es que si algún producto o servicio no cumple con sus especificaciones de calidad, el inspector podrá saber cuál es el valor exacto de la diferencia. Como desventaja, podemos hablar de que este método generalmente necesita del uso de equipos especiales, así como una mayor destreza del empleado, los procedimientos son más rigurosos y requieren de un mayor tiempo y esfuerzo.

### 2.3.2 MUESTREO

Este procedimiento, llamado **inspección completa**, consiste en revisar la calidad de todos los productos o servicios en cada una de las etapas del proceso. De esta manera, se asegura que una unidad defectuosa no pasará a la siguiente operación. Se utiliza cuando los costes causados por un defecto al pasar de una estación de trabajo a otra, o al cliente final, son mayores que los costes derivados de la inspección.



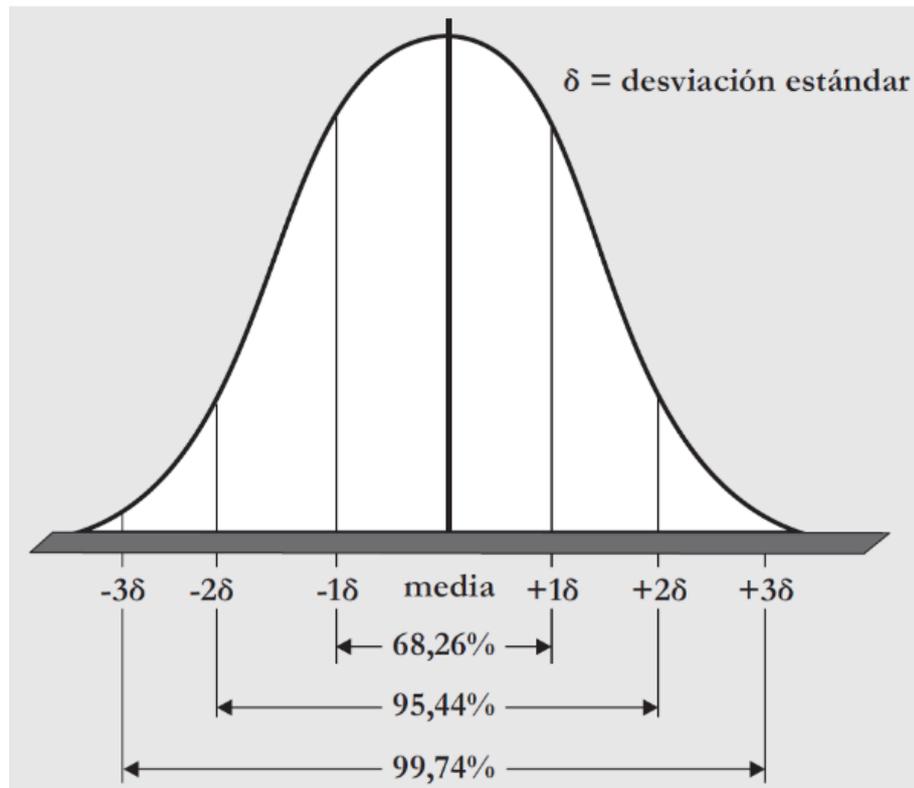
Sin embargo, si este proceso de inspección recae únicamente en las personas encargadas, algunos fallos pueden pasar inadvertidos debidos a un error humano o algún defecto en el aparato de medición. Por tanto, el uso de equipos de inspección automatizados que registren, resuman y muestren los datos, es rentable a largo plazo para la empresa.

En el plan de muestreo se especifican: el **tamaño de la muestra** (cantidad de observaciones seleccionadas al azar), el **intervalo de tiempo** (tiempo que transcurre entre dos muestras consecutivas), y las **reglas de decisión** (determinan cuándo es necesario entrar en acción).

### **2.3.2.1 DISTRIBUCIONES DE MUESTREO**

El objetivo del muestreo es encontrar una variable que sirva para medir cierta característica de calidad de la muestra. Esta medida se usará para evaluar el rendimiento del proceso.

Un número suficiente de muestras recogidas podrá aproximarse a seguir una distribución normal. En la figura 2.3 se presentan los porcentajes de valores que están ubicados dentro de diferentes rangos en una distribución normal. Por ejemplo, se puede observar que el 68,26% de la muestra tendrá valores dentro de una desviación estándar de  $\pm 1$  de la media. A partir de esta distribución, podemos calcular la probabilidad de que el resultado de una muestra quede fuera de unos límites establecidos, lo cual servirá para la construcción y empleo de gráficas de control.



**Figura 2.3:** Distribución normal. (Fuente: Control estadístico de procesos. Carro Paz, Roberto y González Gómez, Daniel)

### 2.3.2.2 GRÁFICAS DE CONTROL

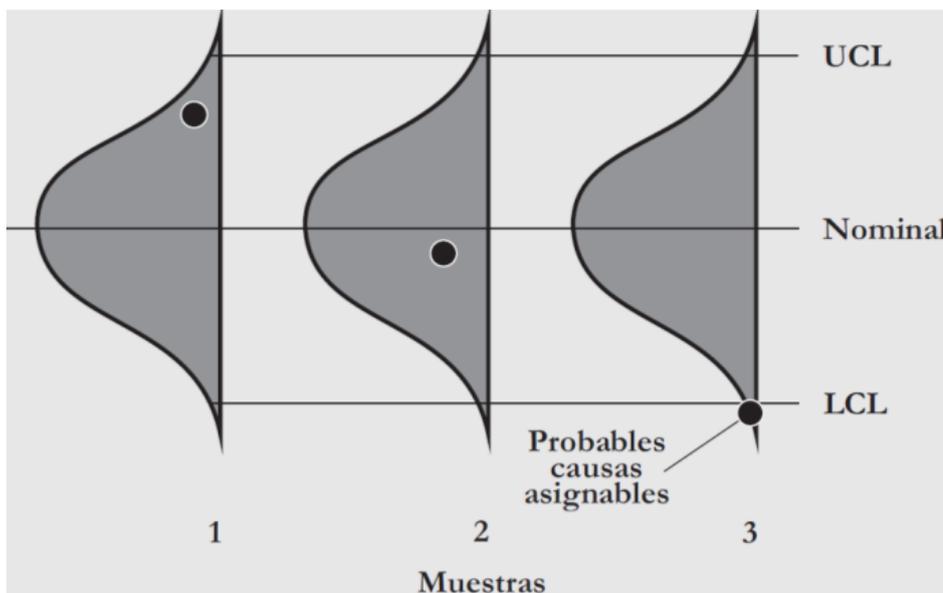
Son diagramas mediante los cuales podemos medir y trazar la gráfica de una determinada característica de calidad tomada de la muestra, por unidad de tiempo. Es la herramienta más usada para monitorizar el estado de un proceso.

La gráfica de control tiene un valor nominal o **línea central**, que es el valor objetivo a alcanzar por medio del proceso, y dos **límites de control** basados en la distribución de muestreo de la medida de calidad, mediante los cuales se define el estándar de funcionamiento a alcanzar.



Así mismo, permiten detectar la presencia de factores asignables (especiales) que desvían la producción del estándar a alcanzar. El valor más grande representa el límite superior de control (en inglés, **UCL**, *upper control limit*) y el valor más pequeño, el límite inferior de control (**LCL**, *lower control limit*).

La figura 2.4 muestra cómo se relacionan los límites de control con la distribución. Una estadística de muestras ubicada entre el UCL y el LCL indica que el proceso muestra causas comunes de variación; en cambio, una estadística ubicada fuera de los límites de control indica que el proceso presenta causas especiales de variación.



**Figura 2.4:** Relación de los límites de control con la distribución de muestreo y observaciones tomadas de tres muestras. (Fuente: Control estadístico de procesos. Carro Paz, Roberto y González)

Las observaciones que se encuentran fuera de los límites de control no siempre denotan una mala calidad, sino que puede deberse a un cambio en el proceso, lo cual hace que la estadística de calidad tomada en esa muestra descienda por



debajo del LCL de la gráfica de control. Para asegurar que se ha producido mejoría debido al cambio introducido, será necesario construir una nueva gráfica de control.

Los encargados de monitorizar un proceso utilizan las gráficas de control de la siguiente forma:

1. Se toma una muestra aleatoria del proceso, se mide la característica de calidad y se calcula una medida variable o de atributos.
2. Si la distribución se va fuera de los límites de control, se busca una causa especial.
3. Se elimina la causa si esta degrada la calidad, o se incorpora si se consigue una mejora. Tras ello, se ha de reconstruir la gráfica de control.
4. Repetir periódicamente el procedimiento.

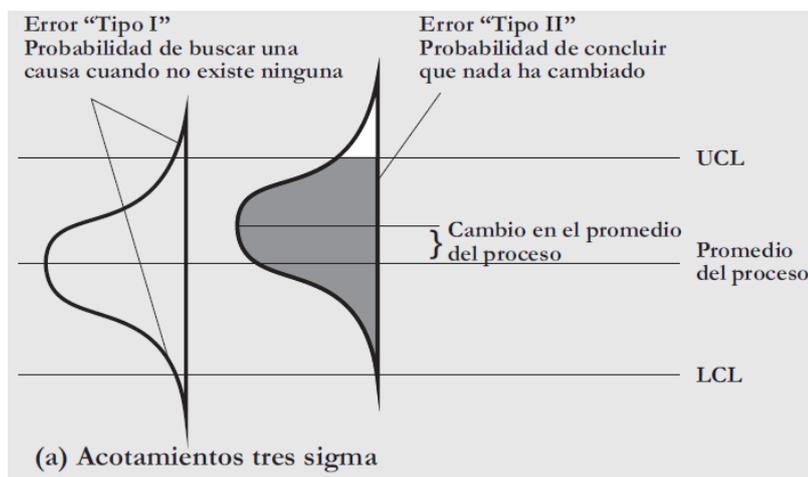
Debido a que están basadas en distribuciones de muestreo, las gráficas de control no son herramientas perfectas, y pueden presentar dos tipos de errores:

- **Error “Tipo I”** (rechazar un lote de buena calidad) se comete cuando se llega a la conclusión de que el proceso está fuera de control basándose en un resultado de muestra fuera de los límites de control, aunque se tratase de un efecto puramente aleatorio.
- **Error “Tipo II”** (aceptar un lote de mala calidad) se produce cuando se cree que un proceso está bajo control



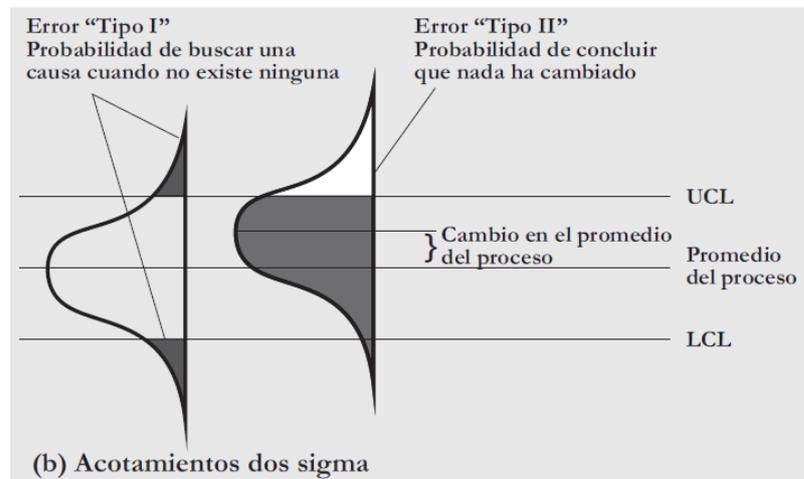
presentando solamente causas comunes de variación, cuando realmente está fuera de control estadístico.

En la figura 2.5 (a) se establecieron los límites de control de la distribución de muestreo para tres desviaciones estándar con respecto a la media, lo que llamamos **acotamiento tres sigma**. Este tipo de acotamiento implica que el coste de buscar las causas especiales es grande en relación al coste de no detectar a tiempo algún cambio en la media del proceso.



**Figura 2.5 (a):** Relación de la expansión de los límites de control con los errores “Tipo I” y “Tipo II”, acotamiento tres sigma. (Fuente: Control estadístico de procesos. Carro Paz, Roberto y González Gómez, Daniel)

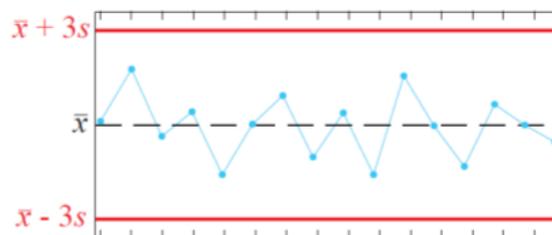
En la figura 2.5 (b), la gráfica de control cuenta únicamente con **acotamiento dos sigma**. Podemos deducir de las figuras que la probabilidad de cometer un error de Tipo I es mucho mayor con acotamiento tres sigma, mientras que, por el contrario, los errores de Tipo II se ven aumentados en acotamiento dos sigma. Con acotamientos como este y más pequeños, será necesario realizar una búsqueda más intensiva para encontrar las causas especiales.



**Figura 2.5 (b):** Relación de la expansión de los límites de control con los errores "Tipo I" y "Tipo II", acotamiento dos sigma.  
 (Fuente: Control estadístico de procesos. Carro Paz, Roberto y González Gómez, Daniel)

## 2.4 CONTROL ESTADÍSTICO DE PROCESOS UNIVARIANTE

Las gráficas de control univariantes permiten monitorizar una única variable sin tener en cuenta al resto de variables del proceso. La figura 2.6 muestra un ejemplo de una gráfica de control de Shewhart para una variable  $x$ , donde su media es " $\bar{x}$ " y su desviación estándar " $s$ ". Estas gráficas no tienen en cuenta la correlación existente entre las variables del proceso.

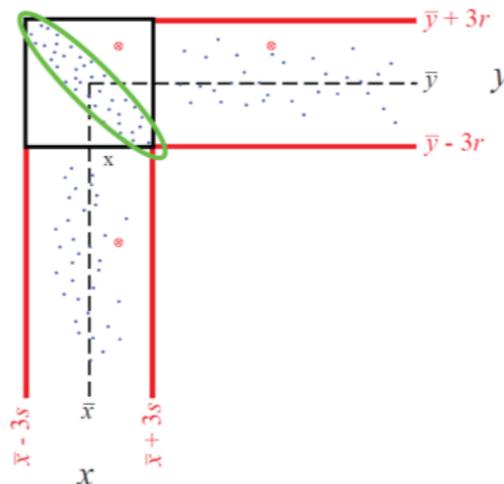


**Figura 2.6:** Gráfico de control Shewhart. (Fuente: Monitoring, fault detection and estimation in processes using multivariate statistical techniques. García Álvarez, Diego)



Cuando aplicamos gráficas de control univariante a sistemas multivariantes, con cientos de variables, los resultados no son los esperados, dado que se comete un error o anomalía en la operación, causando falsas alarmas en cortos periodos de tiempo, o simultáneamente. Esta situación se debe a que las variables del proceso están correlacionadas, y una causa especial puede afectar más que una variable en el mismo tiempo.

Un ejemplo de esto puede verse en la figura 2.7 (Nomikos y McGregor, 1995), donde dos variables (“x” e “y”) presentan un comportamiento normal si están siendo inspeccionadas mediante gráficas de control univariante. Sin embargo, ambas variables están correlacionadas y la muestra expuesta en cruz rompe la correlación. Esta ruptura de la correlación puede ser debida a una causa especial. Por esta razón, la industria moderna requiere técnicas que tengan en cuenta la correlación entre variables, y considerar la planta entera, de manera que sea posible diseñar modelos que puedan predecir con exactitud el comportamiento real de la planta.



**Figura 2.7:** Dos variables correlacionadas. (Fuente: Monitoring, fault detection and estimation in processes using multivariate statistical techniques. García Álvarez, Diego)



El análisis de componentes principales (PCA, *principal component analysis*) y los mínimos cuadrados parciales (PLS, *partial least squares*) son dos técnicas de control estadístico de procesos multivariante que pueden cumplir esos requisitos.

## **2.5 DETECCIÓN Y DIAGNÓSTICO DE FALLOS**

Las técnicas modernas de control han solucionado numerosos problemas en la industria. Sin embargo, la aparición de una causa especial en el proceso puede significar que no está funcionando bajo control. Muchas veces, estas causas especiales se deben a errores en el proceso o en los componentes del mismo. Según Blanke et al. (2003), un fallo se define como *una desviación inesperada de al menos una característica, propiedad o parámetro del sistema de su condición aceptable, común o estándar*. Pueden darse por numerosas causas, como obstrucciones en tubos, desajustes en sensores, disminución de catalizadores, cambios extremos en concentraciones o cambios bruscos en la temperatura ambiente, entre muchos otros. Es por ello necesario dar un paso más allá en los procesos automatizados, desarrollando métodos de detección y diagnóstico de fallos automáticos (FDD).

### **2.5.1 DETECCIÓN DE FALLOS**

La detección de fallos es cada vez un aspecto más importante en los procesos industriales, ayudando a evitar, por una parte, daños físicos y materiales dentro de la propia planta industrial, y por otra, reduciendo pérdidas económicas al producirse un cambio



inesperado en una variable previamente controlada. El proceso no se degrada, con el consiguiente ahorro de materias primas. Se busca, por tanto un sistema de detección que ofrezca de forma eficiente seguridad y fiabilidad al proceso industrial.

Para poder detectar fallos existen varios métodos. Se va a aplicar un sistema de detección de fallos basado en análisis estadístico multivariante llamado PLS (mínimos cuadrados parciales). Este método permitirá la detección de los errores que se produzcan en la planta de forma eficiente.

## **2.5.2 DIAGNÓSTICO DE FALLOS**

La diagnosis de fallos es un sistema complementario a la detección, es decir, un sistema que permanece en *stand-by* mientras la planta funciona sin fallos, pero que se activa cuando alguno se detecta. Su objetivo es diagnosticar el fallo, que significa indicar al operario de la planta encargado del proceso qué elemento provocó el fallo.

Si se diagnostica a qué elemento pertenece el fallo, o lo que es lo mismo, cuál es el elemento que falla, se ahorrará tiempo y con ello recursos económicos.

## **2.5.3 MÉTODOS DE FDD EXISTENTES**

Un esquema de monitorización está basado en la idea de transformar los datos tomados de un proceso en medidas útiles que permitan a los operarios determinar el estado en el que se encuentra el proceso, y si fuera necesario, diagnosticar los fallos



encontrados. La realización de esta tarea se ha abordado fundamentalmente desde tres perspectivas diferentes, que son las siguientes:

- Métodos basados en señal.
- Métodos analíticos.
- Métodos basados en conocimiento.

Sería importante resaltar que estos métodos no plantean soluciones competitivas, sino complementarias. La experiencia muestra que la realidad es demasiado compleja para poder ser diagnosticada desde un único punto de vista.

### **2.5.3.1 MÉTODOS BASADOS EN SEÑAL**

Dentro de esta clase, existen diferentes técnicas que podemos utilizar:

- **Alarmas:** Las variables que medimos las comparamos con valores límite constantes y previamente establecidos. De esta forma nos indicarían el fallo. Generalmente, estos niveles de alarma no podrían adaptarse a la dinámica del sistema con la restricción que ello supone. Esto significa que en algún caso, el sistema puede presentar algunas anomalías sin necesidad de alcanzar los valores límite preestablecidos de las alarmas, aunque estos niveles de alarma podrían ser recalibrados periódicamente, y además, podríamos modificarlos en función de alguna condición de funcionamiento específica.
- **Técnicas estadísticas multivariantes:** Las variables que medimos en un proceso están relacionadas con otras



variables del mismo (correlación espacial) y con otras medidas de la misma variable tomadas en diferentes tiempos (correlación temporal), especialmente cuando los intervalos de muestreo son pequeños. Las técnicas de monitorización basadas en técnicas estadísticas multivariantes se han desarrollado a causa de la necesidad de manejar correlaciones espaciales. Algunas de estas técnicas de monitorización serían las siguientes:

- Mínimos cuadrados parciales (PLS): Se detallará más adelante en este capítulo.
  - Análisis del discriminante de Fisher (FDA): Es una técnica de clasificación.
  - Análisis de componentes principales (PCA): Es una técnica que reduce la dimensionalidad en términos de la varianza de los datos recogidos en una muestra.
- **Redundancia física:** Establecemos múltiples componentes físicos (actuadores, sensores y componentes del sistema). Todos estos componentes están distribuidos espacialmente por toda la planta, y con ello garantizamos una cierta protección en presencia de algún fallo localizado. Gracias a este análisis de consistencia de los datos y a la comparación de las señales de salida de elementos (supuestamente idénticos), nos permiten detectar la presencia de fallos, así como la localización y el aislamiento de los mismos. Tiene una desventaja, y es que no es capaz de detectar los fallos que afectan a todos los elementos de la misma manera. Por otro lado, el coste económico y espacial es elevado, al duplicar todos los elementos.



- **Análisis en frecuencia:** Algunas variables, como pueden ser las vibraciones en una máquina rotativa, pueden ser estudiadas en el dominio de la frecuencia, aportándonos una gran cantidad de información sobre el estado de operación de la máquina. Cada fallo se puede representar también por un patrón de frecuencias y amplitudes característico, lo que nos permite realizar su diagnóstico.

### **2.5.3.2 MÉTODOS ANALÍTICOS**

- **Redundancia analítica:** Es un método para la detección de fallos que emplea un modelo de la planta. Este método se utiliza porque las medidas que produce un sistema se derivan de un mismo estado dinámico, a pesar de que los sensores sean diferentes, y por tanto, están funcionalmente relacionadas a través de su modelo. Estas técnicas de FDD están basadas en modelos que constan de los siguientes pasos:
  - Generación de residuos.
  - Toma de decisión basada en esos residuos.

En el primer paso, los residuos que nominalmente son cero (se desvían del cero si se ha producido un fallo en particular), se generan mediante un algoritmo determinado en el que se procesan las entradas y salidas del sistema. De unos métodos a otros, las técnicas para generar los residuos son diferentes, pero generalmente consiste en calcular la diferencia entre la salida de la planta y la salida calculada por un modelo de la misma. Existen diferentes formas para calcular el modelo del proceso. Entre las más conocidas, tenemos las siguientes:



- Filtros de Kalman.
- Ecuaciones de paridad .
- Observadores de estado .
- Estimación de parámetros.

En el segundo paso, tenemos la toma de decisión, es decir, para determinar si hay o no fallos, examinaremos los residuos. El proceso de toma de decisión puede consistir en hacer un simple test de umbral en los valores instantáneos, en la media móvil de los residuos, o también pueden utilizarse los métodos de la teoría estadística de decisión.

Una de las principales características de la generación de residuos es su robustez (insensibilidad) a los errores que se pueden dar en el modelado, los cuales son inevitables. Debido a las perturbaciones externas, al ruido, a la incertidumbre de los parámetros de la planta y a las dinámicas no modeladas del sistema, se producen falsas alarmas o una incorrecta identificación de los fallos. Debido a los anteriores factores, se ha conducido a la determinación de umbrales apropiados a cada residuo, así como al diseño de generadores de residuos robustos. En el segundo paso también es posible introducir la robustez, lo que significa hacer una evaluación robusta de los residuos.

### **2.5.3.3 MÉTODOS BASADOS EN CONOCIMIENTO**

Existen diferentes técnicas que podemos utilizar:

- **Sistemas expertos:** Esta técnica nos sirve como complemento a las técnicas anteriormente vistas, lo cual



significa que podemos utilizar toda la información proveniente de ellas en la detección de fallos, para poder analizar todos los síntomas y realizar un diagnóstico completo. El conocimiento de los expertos es expresado en forma de reglas, las cuales relacionan los síntomas con las causas que lo han provocado, representando todo esto en el sistema experto, pudiendo ofrecer soluciones o recomendaciones.

El **diccionario de fallos** y el **árbol de decisión** son las dos formas más usuales de diagnóstico que existen. El árbol de decisión consiste en un conjunto estructurado de reglas lógicas del tipo “Si-síntoma-Y-síntoma-ENTONCES-conclusión”. Cada conclusión puede a su vez servir de síntoma en la siguiente regla, y así sucesivamente hasta que se alcance una conclusión final, mientras que el diccionario de fallos contiene una lista de causas y efectos.

Estas técnicas tienen el problema de que se requiere la existencia de uno o más expertos en el sistema estudiado. Este experto debe conocer a la perfección el funcionamiento del sistema nominal, así como las posibles situaciones anómalas que puedan presentarse para cada uno de los posibles fallos o situaciones extrañas que pudieran producirse. Esto provoca una limitación de las posibilidades de desarrollo para sistemas de reciente implantación. En estos sistemas se desconoce su comportamiento anómalo, y las relaciones entre posibles síntomas y sus causas son hipótesis difíciles de validar.

- **Redes neuronales:** En diagnóstico, las redes neuronales admiten varias aplicaciones. La principal consiste en clasificar los diferentes estados del sistema a partir de



las variables medidas. Estos estados los podemos agrupar en normales y defectuosos, permitiendo clasificar estos últimos para los diferentes tipos de fallos. Para realizar el aprendizaje de los distintos modos de operación del sistema, así como los modelos de fallos, se realizan mediante la presentación de muestras de ejemplos. Esto realmente está clasificado dentro de las técnicas basadas en reconocimiento de patrones.

El inconveniente del modelo de redes neuronales es que si no existen datos de fallo o si son muy escasos, no puede aprender este modo de funcionamiento. La limitación que tiene una red neuronal es que su capacidad de generalización podría llevar a clasificar modos de fallo no aprendidos, como modos de funcionamiento normal. No obstante, se puede superar esta limitación si la simulación de los modos de fallo por computador es lo suficientemente fiable.

Otra aplicación que se le puede dar a las redes neuronales en la detección y diagnóstico de fallos, es calcular un modelo nominal de la planta. Puesto que una red neuronal artificial es capaz de encontrar las diferentes relaciones funcionales que representan el proceso, filtrando de esta manera el ruido. La comparación entre la salida del proceso y la salida de la red neuronal, nos permite detectar posibles fallos.

- **Sistemas basados en lógica difusa:** Todas estas técnicas las podemos utilizar tanto para la evaluación de los residuos (introduciendo reglas difusas basadas en el conocimiento del sistema), como en la generación de residuos, ya que de nuevo, un sistema lógico difuso es un aproximador universal de funciones, en particular las



funciones que relacionan las distintas variables del proceso, es decir, calculando un modelo difuso del proceso. La comparación entre la salida del sistema lógico difuso y la salida del proceso permite la detección de posibles fallos.

- **Búsqueda causal:** Está basada en el seguimiento de los fallos hasta su causa. Estas causas son representadas mediante grafos directos dirigidos, donde las variables de estado están representadas por los nodos, y las relaciones que existen entre ellas se representan mediante las ramas, es decir, las restricciones que se tienen que cumplir. Una búsqueda por el grafo nos dará la causa final del fallo.

## **2.5.4 REQUERIMIENTOS DE DISEÑO**

Las características que debe de poseer un método de detección y diagnóstico de fallos son las siguientes:

- *Tiempo bajo de detección ( $t_d$ ):* Es el tiempo que tarda el esquema de monitorización en detectar el fallo desde que se produjo. Normalmente se minimiza para un índice de falsas alarmas fijo.
- *Alto grado de detecciones correctas ( $P_d$ ):* Todos los fallos son detectados.
- *Índice bajo de falsas alarmas ( $P_f$ ):* Cuando una alarma indica un fallo que realmente no ha ocurrido.



- *Buena capacidad de aislamiento:* Es la capacidad para diferenciar los distintos fallos que se pueden dar, y depende de los test estadísticos que se empleen y de la estructura de las ecuaciones del sistema.
- *Sensibilidad:* Es una medida que determina la magnitud del fallo necesaria para poder aislar dicho fallo bajo condiciones preestablecidas. Dependerá de las características del ruido y de las magnitudes de los coeficientes de las ecuaciones del sistema. La sensibilidad está relacionada con el tiempo necesario para detectar y localizar el fallo.
- *La robustez:* Es la característica que trata de aislar fallos cuando existen errores de modelado y/o perturbaciones desconocidas. Por lo tanto, es una valoración de cómo se comporta el método de diagnóstico. En dichos casos, este aspecto es de una gran importancia al ser prácticamente inevitable que se de algún error en el modelado o el desconocimiento de las perturbaciones que afectan al sistema.

A la hora de analizar la eficacia de un esquema de FDI (Detección e Identificación de fallos), podemos recurrir a provocar un determinado fallo, manteniéndolo y observando cómo reacciona el sistema de detección. Si el fallo provocado produce una respuesta nula por parte del módulo de monitorización, entonces podemos decir que el fallo ha sido omitido, lo cual puede ser aceptable si se trata de un fallo sin importancia. En el caso de que el fallo tuviera un gran impacto sobre el correcto funcionamiento del sistema sería completamente inaceptable.

Cuando ha ocurrido un fallo y este ha sido detectado con éxito, es de gran importancia la rapidez con la que se detecta. Por



ejemplo, en las aplicaciones aeroespaciales, si hay un fallo que persiste de forma continuada durante varias fracciones de segundo sin que éste sea corregido, puede tener como consecuencia la destrucción de la misión del sistema en operación, o la del propio sistema. Sin embargo, en otros sistemas a veces es mejor tener una detección más fiable de fallos con poca magnitud, aunque esto conlleve que el tiempo de detección sea mayor. A la hora de combinar estos factores debe de existir un compromiso entre ambas opciones.

Un fallo no siempre se produce repentinamente y pueden persistir durante mucho tiempo. También existen fallos que se desarrollan lentamente o son de pequeña magnitud, a los cuales se les conoce como fallos incipientes. Por ejemplo, medidas con sesgo (error que ocurre de forma sistemática), fallos intermitentes, o, desviaciones de media móvil de un instrumento ("drifts"), en estos casos la rapidez juega un papel secundario respecto a la sensibilidad del esquema de FDI. En otras aplicaciones, tanto la rapidez como la sensibilidad de detección forman un papel esencial, en cuyo caso los esquemas de detección son más complejos, siendo a veces necesaria una combinación de redundancia analítica y de dispositivos físicos para lograr los dos objetivos.

Las falsas alarmas son generalmente el factor indicativo del mal funcionamiento de un esquema FDI. Incluso podemos decir que un nivel bajo de falsas alarmas durante la operación normal de la planta es intolerable ya que esto nos conduce a una falta de confianza en el sistema de detección al que se refiere.



## 2.6 MÍNIMOS CUADRADOS PARCIALES (PLS)

La regresión por mínimos cuadrados parciales (PLS), fue introducida por Herman Wold (1975). La demanda por esta metodología y la evidencia de que trabaja bien, van en aumento y así, la metodología PLS está siendo aplicada en muchas ramas de la ciencia.

El PLS es una técnica de reducción de la dimensionalidad maximizando la covarianza entre la matriz de predicción  $X$  y la matriz predicha  $Y$  para cada componente del espacio reducido, con  $X$ : matriz de variables de predicción, de orden  $n \times p$ , e  $Y$ : matriz de variables dependientes, de orden de  $n \times q$ . Donde la matriz  $X$  está constituida por los datos que se recogen en planta. Las columnas son las diferentes variables que estamos manejando, y las filas son los datos observados a lo largo del tiempo. La matriz  $Y$  es la matriz que contiene la o las variables que queremos calcular o predecir, y que son normalmente una medida de la calidad del proceso.

### 2.6.1 REGRESIÓN PLS UNIVARIADA

Es el caso de aplicación de regresión PLS (de la Fuente, 2009), se realiza cuando tenemos datos de operación normal de la planta, y se opera con la matriz  $Y$  con una sola columna. Puede ser visto como una transformación de las variables de predicción matriz  $X$ , considerando su relación con el vector de respuestas  $Y$  de orden  $n \times 1$ , obteniéndose como resultado una matriz,  $T = (T_1, \dots, T_p)$  de orden  $n \times p$ .

$$X = TP^T + E \quad (1)$$



Esta matriz producto  $T P^T$  se puede descomponer en la suma del producto de los vectores  $t_j$  (columna  $j$  de  $T$ ) y los vectores de carga  $p_j$  (columna  $j$  de  $P$ ).

$$X = \sum_{j=1}^a t_j p_j^T + E \quad (2)$$

De la misma forma,  $Y$  se puede descomponer en una matriz  $U \in R^{n \times a}$  y una matriz de carga  $Q \in R^{p \times a}$  más una matriz de residuos  $F \in R^{n \times p}$ :

$$Y = UQ^T + F = \sum_{j=1}^a u_j q_j^T + F \quad (3)$$

La técnica PLS relaciona el vector estimado  $\hat{u}_j$  de  $Y$  con el vector  $t_j$  de  $X$  de la siguiente forma:

$$\hat{u}_j = b_j t_j \quad (4)$$

con  $b_j$  el coeficiente de regresión. Si se pone esta relación en forma matricial, se obtiene:

$$\hat{U} = TB \quad (5)$$

Con  $B \in R^{a \times a}$  la matriz de regresión diagonal con  $B_{jj} = b_j$ . Sustituyendo esta ecuación en la ecuación (3) se obtiene:

$$Y = TBQ^T + F \quad (6)$$

Siendo  $F$  la matriz de error de predicción. La matriz  $B$  la calculamos de forma que el máximo valor singular de  $F$ ,  $\|F\|_2$  sea mínimo. Los vectores  $t_j$  y  $\hat{u}_j$  se calculan para cada factor PLS, es decir, con  $j=1, 2, \dots, a$  de tal forma que la covarianza entre  $X$  e  $Y$  sea máxima para cada factor. Se puede hacer este mismo procedimiento para calcular la regresión PLS multivariada, es decir, un modelo para cada columna de  $Y$ , es decir, para cada



salida, obteniéndose en dicho caso:  $y_i = T_i B_i q_i^T + f_i$ , donde  $y_i \in R^n$  es la columna  $i$  de  $Y$ ,  $T_i \in R^{n \times a}$ ,  $B_i \in R^{a \times a}$  es la matriz de regresión,  $q_i \in R^a$  es el vector de carga y  $f_i \in R^n$  es el error de predicción.

La entrada de datos corresponde a las matrices  $X$  e  $Y$ , las cuales han sido centradas y normalizadas, por columnas con varianza igual a uno y media igual a cero. El proceso se inicializa dando el valor de  $E_0 = X$ ,  $F_0 = Y$ , y el valor de  $u_1 = Y$ . A continuación, realizaremos varias iteraciones a las variables  $w$ ,  $t$ ,  $u$ ,  $q$ . Hasta que el resultado sea convergente, que se calcula comparando el vector  $t_j$  con su valor en la iteración anterior.

$$w_j = \frac{E_{j-1}^T u_j}{\|E_{j-1}^T u_j\|_2} \quad (7)$$

$$t_j = E_{j-1} w_j \quad (8)$$

$$q_j = \frac{F_{j-1}^T t_j}{\|F_{j-1}^T t_j\|_2} \quad (9)$$

$$u_j = F_{j-1} q_j \quad (10)$$

Las cuatro ecuaciones anteriores convergen cuando  $t_j$  menos el  $t_{j-1}$  es menor a una cierta tolerancia de error.

Una vez obtenidos los vectores  $w$ ,  $t$ ,  $q$ ,  $u$ , pasamos a calcular el valor de la variable  $p_j$ , cuya fórmula sería la siguiente:

$$p_j = \frac{E_{j-1}^T t_j}{t_j^T t_j} \quad (11)$$



El siguiente paso sería sacar los valores definitivos de  $w$ ,  $p$  y  $t$ , escalándolos por la norma de  $p_j$ , de la anterior iteración. Las fórmulas para realizarlo serían las siguientes:

$$t_{j,nuevo} = \frac{p_{j,anterior}}{\|p_{j,anterior}\|_2} \quad (12)$$

$$t_{j,nuevo} = t_{j,anterior} \|p_{j,anterior}\|_2 \quad (13)$$

$$w_{j,nuevo} = w_{j,anterior} \|p_{j,anterior}\|_2 \quad (14)$$

Una vez realizado esto, podremos calcular el vector de regresión  $b_j$ , que sería el siguiente:

$$b_j = \frac{u_j^T t_j}{t_j^T t_j} \quad (15)$$

Finalmente, tenemos las matrices residuales, las cuales nos servirán para continuar realizando de nuevo otra iteración. Las fórmulas serán las siguientes:

$$E_j = E_{j-1} - t_j p_j^T \quad (16)$$

$$F_j = F_{j-1} - b_j t_j q_j^T \quad (17)$$

Una vez realizada esta operación, volvemos a repetir toda la iteración hasta que obtengamos una matriz de las variables  $w$ ,  $t$ ,  $q$  y  $p$ , cuyas columnas dependen del número de vectores con los que deseemos trabajar.



### **2.6.1.1 CÁLCULO DE LAS ESTADÍSTICAS T<sup>2</sup> Y Q**

Una vez realizado el pls univariado, vamos a calcular el error mediante dos métodos para ver cuándo se produce un fallo (Geladi, y Kowalski, 1996). El primer método que vamos a analizar es el método Q.

**Método Q:** Para realizar este método, seleccionamos la matriz X del método PLS univariante y normalizamos la primera fila con una varianza igual a uno y media igual a cero, donde obtendremos como resultado la fila normalizada  $X_{nn}$ . Una vez normalizada, aplicamos respectivamente las siguientes expresiones:

$$T_n = X_{nn}w \quad (18)$$

$$X_{ns} = T_n p^T \quad (19)$$

$$Error = X_{nn} - X_{ns} \quad (20)$$

$$Q = Error \times Error^T \quad (21)$$

Una vez obtenido el error en la primera fila de la matriz X, realizamos el mismo proceso con las siguientes filas del mismo modo, hasta obtener el error de todas las filas, con lo que al final tendremos una matriz de error Q. Se detecta un fallo cuando en un instante determinado, la estadística Q supera un umbral. Este mismo procedimiento se puede realizar también con datos nuevos recogidos de la planta. Una vez tomados los datos, se construye la primera fila de una nueva matriz  $X_{new}$ , calculamos un nuevo Q como se indica anteriormente, y comparando Q con su umbral podemos saber si en el instante actual la planta actúa en condiciones normales de operación o ha ocurrido un fallo.



**Método T<sup>2</sup>:** Para realizar este método, previamente seleccionamos la matriz X del proceso PLS univariado y realizamos el mismo procedimiento de normalización que se ha realizado previamente en el método Q, tomando como antes fila por fila. Después, utilizamos el T<sub>n</sub> calculado previamente en el método Q (ecuación 18).

Una vez realizados todos estos pasos, hacemos tantas iteraciones como vectores PLS hayamos obtenido en el método PLS univariado, en la siguiente fórmula:

$$T_{na} = \frac{T_n}{\sqrt{S_j}} \quad (22)$$

Donde S<sub>j</sub> es la varianza de la columna de la matriz t<sub>j</sub>, previamente calculada en el método PLS univariado. Una vez obtenido T<sub>na</sub>, aplicamos la siguiente fórmula para calcular la estadística T<sup>2</sup>:

$$T^2 = T_{na} \times T_{na}^T \quad (23)$$

Realizaremos sucesivamente este proceso con las siguientes filas de la matriz X, hasta obtener el error de todas las filas de dicha matriz. Con lo que al final, obtendremos una matriz T<sup>2</sup>. Se detecta un fallo cuando en un instante determinado, la estadística T<sup>2</sup> supera un umbral.

En nuestro caso, utilizamos ambos métodos de forma conjunta para la detección de fallos, ya que hay algunos casos en los cuales un método no es suficiente, y de esta forma aseguramos una mejor detección.

Para diagnosticar un fallo detectado, debemos preguntarnos qué variables de la matriz de datos original son las responsables del fallo detectado, es decir, qué variable o variables han contribuido



a que cualquiera o ambos de los estadísticos  $T^2$  y  $Q$  hayan superado sus respectivos umbrales durante un número determinado de instantes consecutivos. Para ello, utilizaremos diagramas de contribución.

En el caso del estadístico  $T^2$ , para una observación con fallo en primer lugar debemos determinar los  $n$  scores  $t_i$  que superen el umbral determinado por el usuario, los responsables de que el proceso esté fuera de control. Tras ello, debemos calcular la contribución de cada variable  $X_j$  a ese score  $t_i$  que está fuera de control, de la siguiente forma:

$$t_i = X_{tfalloj} \times R_{ji} \quad (24)$$

$$T_i^2 = t_i^2 \quad (25)$$

Si;  $T_i^2 > Umbral_{T^2}$  (26)

, entonces;  $cont_{ij} = \frac{t_i}{\delta_i} R_{ji} X_{tfalloj}$  (27)

Y calculamos la contribución total de la variable  $j$ -ésima:

$$CONT_j = \sum_{i=1}^n cont_{ij} \quad (28)$$

Por último, representamos  $CONT_j$ .

Para el estadístico  $Q$ , lo que haremos será, en el instante en el que se produce el fallo, representamos el valor al cuadrado del residuo para cada variable en ese instante de tiempo, obtenido de la siguiente forma (para el caso de  $Q_x$ ):

$$r_i = (x_i - \hat{x}_i)^2 \quad (29)$$

Los valores del residuo para cada variable en el instante del fallo han sido convertidos, en caso de ser negativos, a su valor absoluto para trabajar con ellos.



## 2.6.2 REGRESIÓN PLS DINÁMICA (DPLS)

Basándonos en el método PLS univariado estudiado en los apartados anteriores, se ha desarrollado un algoritmo de detección de fallos que tiene en cuenta la historia reciente de la planta. El método diseñado es similar al PLS univariado, con la diferencia de la matriz de predicción  $X$ .

Normalmente, el método PLS univariado tiene tantas columnas como variables se deseen estudiar. El nuevo método dinámico tendrá el triple de columnas más una. Esto se debe a que además de los valores de las variables en el momento actual, se incluirán también en la matriz de predicción las variables hasta dos instantes pasados, además de un instante pasado de la variable estimada  $Y$ , de forma que la nueva matriz de predicción  $X$  tendrá la siguiente estructura:

$$X = [x(t), y(t - 1), x(t - 1), x(t - 2)] \quad (30)$$

Siendo  $x$  la matriz de predicción que se elige cuando se aplica el método PLS.

Una vez formada la nueva matriz de predicción  $X$ , se continúa de igual forma a como lo hace PLS en el apartado anterior.

La finalidad de este nuevo método de detección es mejorar los resultados que se han de obtener con PLS, ya que es de suponer que al poseer más información de las variables al realizar el método PLS, la predicción que se obtenga de la variable de calidad será mejor, y por tanto los errores que obtendremos como diferencia entre el valor actual de la planta y el valor predicho por el método de regresión DPLS serán más pequeños, por otro lado, también se espera una mayor diferencia de las estadísticas  $T^2$  y  $Q$



respecto al umbral, y por tanto, que aumente la habilidad de detección del método.

### **2.6.3 REGRESIÓN PLS NO LINEAL (NPLS)**

Hasta ahora, en PLS habíamos utilizado una regresión lineal mediante la proyección a un nuevo espacio de las variables de predicción y la variable de salida. En este caso, nos centraremos en dotar de carácter no lineal a una parte de la matriz de predicción  $X$ , con el objetivo de analizar cómo esta modificación influye en los resultados arrojados.

Para hacer esto posible, simplemente se utiliza conocimiento del proceso con el que se está trabajando, en concreto se necesita conocer como varía la variable de calidad que se está prediciendo con respecto a las variables medibles del proceso. Estas relaciones son no lineales, y pueden ser producto entre dos variables, raíces cuadradas, exponenciales, logaritmos, etc. y se utilizan variables nuevas calculadas utilizando estas expresiones para construir la matriz  $X$  de regresión del método PLS.

$$X = [x(t), z(t)] \quad (31)$$

Siendo  $x$  la matriz de predicción que se elige cuando se aplica el método de PLS, y  $z(t) = x_1(t) * x_2(t)$  o  $z(t) = \exp(x_1(t))$ , etc.

### **2.6.4 REGRESIÓN PLS CON EXTRACCIÓN DE CARACTERÍSTICAS (PLSEX)**

La extracción de características es un método de reducción de dimensionalidad, utilizado para transformar un conjunto de datos



con espacio de gran dimensión, en uno con menor dimensión. En el caso de PLS, esta transformación será lineal.

La técnica lineal utilizada, realiza un análisis lineal de los datos y los proyecta a un espacio de dimensión menor, de tal forma que la varianza de los datos en la nueva dimensión se maximiza. En la práctica, se construye la matriz de correlación de datos, calculando en ella los autovectores, que corresponden a los valores propios más grandes. Una vez calculados, pueden utilizarse para reconstruir en gran medida la varianza de los datos originales.

### **2.6.4.1 MÉTODO DE CORRELACIÓN**

El coeficiente de correlación de Pearson se utiliza para medir la relación lineal entre dos variables aleatorias cuantitativas. Este coeficiente es independiente de la escala de medida de las variables, por lo que puede trabajar con variables normalizadas.

En el caso de estudio de dos variables aleatorias X e Y sobre una muestra, el coeficiente de correlación de Pearson se representa como  $\rho_{x,y}$ , y se calcula a partir de la siguiente ecuación:

$$\rho_{x,y} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{E[(X-\mu_x)(Y-\mu_y)]}{\sigma_x \sigma_y} \quad (32)$$

Donde:  $\sigma_{xy}$  es la covarianza de (X, Y),  $\sigma_x$  es la desviación estándar de la variable X,  $\sigma_y$  es la desviación estándar de la variable Y.

Análogamente, se puede calcular este coeficiente sobre un estadístico muestral, representado por  $r_{xy}$ :



$$r_{xy} = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{(n-1) s_x s_y} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}} \quad (33)$$

El valor del índice de correlación  $r$  está comprendido en el intervalo  $[-1, 1]$ . Cuanto más se aproxime a estos límites, mayor será la correlación existente entre las variables, la cual puede ser tanto positiva como negativa.

Una vez calculado este índice de correlación de cada variable de la matriz de predicción  $X$  con la variable de la matriz predicha  $Y$ , seleccionaremos aquellas variables cuyo índice supere un umbral establecido, y utilizaremos el mismo procedimiento que en PLS univariante con la nueva matriz de predicción  $X$  con características ya extraídas.



## **2.7 REDES NEURONALES ARTIFICIALES (RNA)**

Las redes de neuronas artificiales (RNA) son dispositivos programados de tal manera que funcionen como un proceso automático inspirado en la forma en la que trabaja el sistema nervioso de los seres vivos. Se trata de un sistema de interconexión de neuronas en una red que, operando de forma paralela, consiguen resolver problemas relacionados con el reconocimiento de formas o patrones, predicción, codificación, control y optimización, para producir un estímulo de salida.

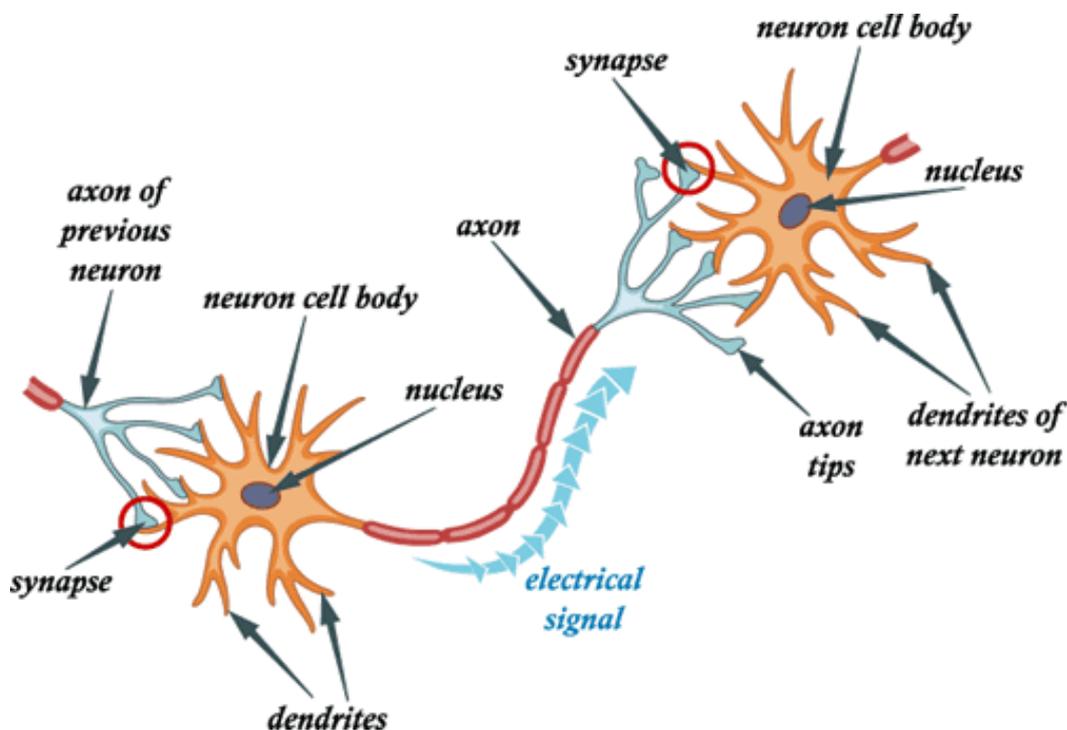
### **2.7.1 NEURONAS**

La neurona es el elemento fundamental para que el sistema nervioso entre en funcionamiento. Generalmente, una neurona consta de un cuerpo celular más o menos esférico del que salen una rama principal, el axón, y varias ramas más cortas llamadas dendritas. Una de las características de las neuronas es su capacidad para comunicarse. Las dendritas y el cuerpo celular reciben señales de entrada; el cuerpo celular combina e integra y emite señales de salida. El axón transmite dichas señales a los terminales axónicos, que distribuyen información a un nuevo conjunto de neuronas.

Las señales que se utilizan son de dos tipos; eléctrica y química. La señal generada por la neurona y transportada a lo largo del axón es un impulso eléctrico, mientras que la señal que se transmite entre los terminales axónicos de una neurona y las dendritas de la otra es de origen químico.



Para construir una similitud directa entre la actividad sináptica y la analogía con las redes neuronales artificiales, se pueden considerar que las señales que llegan a la sinapsis son las entradas a la neurona. Estas son ponderadas (atenuadas o simplificadas) a través de un parámetro denominado peso asociado a la sinapsis correspondiente. Estas señales de entrada pueden excitar a la neurona (sinapsis con peso positivo) o inhibirla (peso negativo). El efecto es la suma de las entradas ponderadas, siendo esta señal del tipo todo/nada, es decir, la neurona se activa o no se activa. En la figura 2.8 podemos observar el esquema de componentes y funcionamiento una neurona real.



**Figura 2.8:** Esquema de componentes de una neurona del sistema nervioso.

Fuente: study.com

En las RNA, el elemento principal de las redes es la neurona artificial, igual que en sistema nervioso. Una neurona artificial (figura 2.9) es un procesador elemental en el que se procesa una



o varias entradas, donde puede verse aumentada su intensidad debido al parámetro característico de conexión denominado peso. Las conexiones determinan la influencia entre neuronas. Finalmente, se produce una respuesta o salida única.

Los elementos básicos de una neurona artificial son:

- **Entradas:** Estas capas reciben la información desde el exterior (entradas  $W_i$  a la neurona). La entrada neta a cada unidad puede escribirse de la siguiente manera:

$$neta_t = \sum_{t=1}^n W_t X_t = \vec{X}\vec{W} \quad (34)$$

- **Pesos:** Generalmente, una neurona recibe muchas y múltiples entradas simultáneas. Cada entrada tiene su propio peso relativo, el cual proporciona la importancia de la entrada dentro de la función de agregación de la neurona. Las entradas con mayor peso, tienen mayor efecto sobre el procesamiento de la neurona al combinarse para producir la respuesta neuronal. Estos pesos pueden adaptarse dentro de la red, y constituyen la medida de la fuerza de una conexión de entrada.
- **Salidas:** Cada elemento de procesamiento tiene permitida una única salida  $Y_t(t)$  que puede estar asociada con otras neuronas. La salida es equivalente al valor resultante de la función de activación.

$$Y_t(t) = F_t(a_t(t)) = a_t(t) \quad (35)$$

En la figura 2.9 está representado el esquema de una neurona artificial:

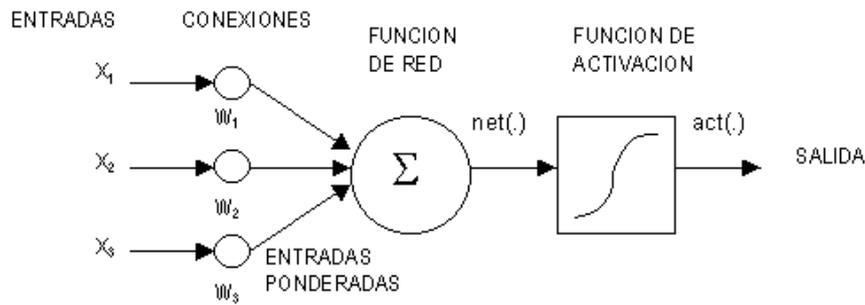


Figura 2.9: Esquema de una red neuronal artificial (RNA).

- **Función de activación:** Entre las neuronas que forman una red neuronal artificial, existe un conjunto de conexiones que sirven de unión. Cada neurona transmite señales a aquellas que estén conectadas a su salida. Cada neurona o unidad  $U_j$  está asociada a una función de activación ( $F_{act}$ ) que transforma la entrada neta que la neurona presenta como resultado de estar conectada con otras que le suministran información, en el valor de salida  $Y_j$ .

$$Y_j(t) = F_{act,j}(neta_j(t)) \quad (36)$$

En la tabla de la figura 2.10 podemos ver los diferentes tipos de función de activación más usuales.



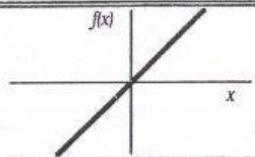
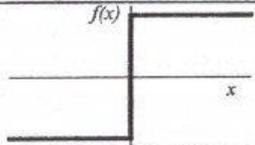
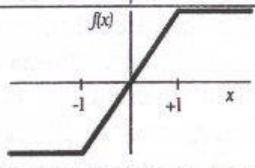
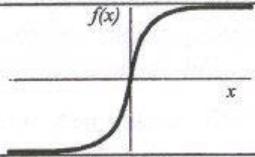
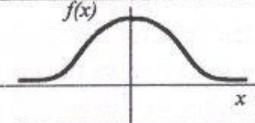
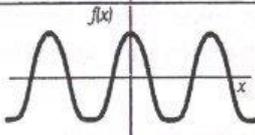
	Función	Rango	Gráfica
<b>Identidad</b>	$y = x$	$[-\infty, +\infty]$	
<b>Escalón</b>	$y = \text{sign}(x)$ $y = H(x)$	$\{-1, +1\}$ $\{0, +1\}$	
<b>Lineal a tramos</b>	$y = \begin{cases} -1, & \text{si } x < -l \\ x, & \text{si } -l \leq x \leq +l \\ +1, & \text{si } x > +l \end{cases}$	$[-1, +1]$	
<b>Sigmoidea</b>	$y = \frac{1}{1+e^{-x}}$ $y = \text{tgh}(x)$	$[0, +1]$ $[-1, +1]$	
<b>Gaussiana</b>	$y = Ae^{-Bx^2}$	$[0, +1]$	
<b>Sinusoidal</b>	$y = A \text{sen}(\omega x + \varphi)$	$[-1, +1]$	

Figura 2.10: Tipos de Función de Activación y su representación. (Fuente: [www.ibiblio.org](http://www.ibiblio.org)).

### 2.7.2 ESTRUCTURA DE LAS RNA

La combinación de dos o más neuronas forman una capa, y una o más capas forman una red. Generalmente, los elementos de una misma capa tienen la misma función de activación. Cada capa puede recibir información de una o más capas, y después de procesarla puede enviar su salida a una o más capas.



### 2.7.2.1 UNA CAPA DE NEURONAS

El caso más general a tener en cuenta para analizar una capa de neuronas es considerar que las neuronas de la capa poseen múltiples entradas. En la figura 2.11 vemos un esquema de dicha arquitectura, donde se representa a una capa de  $S$  neuronas con  $R$  entradas y vemos que cada neurona recibe el mismo vector de entrada  $P$  de  $R$  componentes y una "bias"  $b$ , y que cada conexión del vector de entrada con las neuronas posee un vector de pesos  $W$ . De esta forma, la  $i$ -ésima neurona realiza la operación  $W(i)*P+b(i)$  y produce la salida  $a(i)=F(W(i)*P + b(i))$ . La salida de la capa será un vector de  $S$  componentes.

El elemento  $w(i,j)$  de la matriz  $W$  representa el peso de la conexión entre la neurona  $i$ -ésima y la componente  $j$ -ésima del vector de entrada. La "bias" de la  $i$ -ésima neurona es  $b(i)$ , y la salida de la correspondiente neurona es  $a(i)$ . El número de neuronas de una capa no tiene por qué ser igual que el número de componentes del vector de entrada. Generalmente  $R \neq S$ .

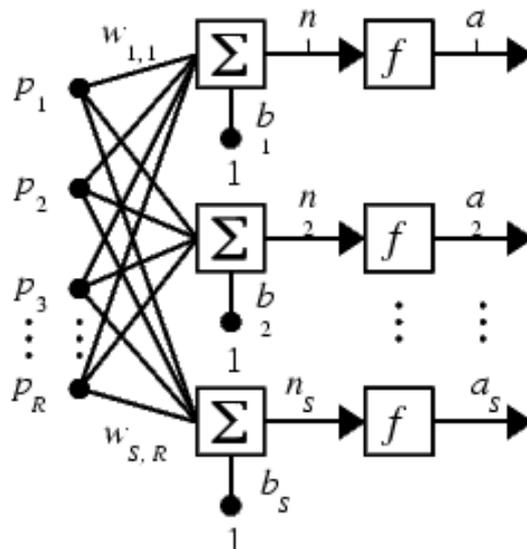


Figura 2.11: Estructura de una red neuronal de una capa.



### 2.7.2.2 MÚLTIPLES CAPAS DE NEURONAS

Una red neuronal está formada por una o más capas de neuronas, y como hemos dicho antes, cada capa puede recibir información de una o más capas y enviarla a una o más capas. Una red típica consta de tres tipos de capas:

**Capa de entrada:** Recibe directamente la información de fuentes externas a la red.

**Capa oculta:** No tiene contacto directo con el exterior, y las neuronas que la componen pueden estar interconectadas de diferentes formas.

**Capa de salida:** Es la que transfiere la información al exterior.

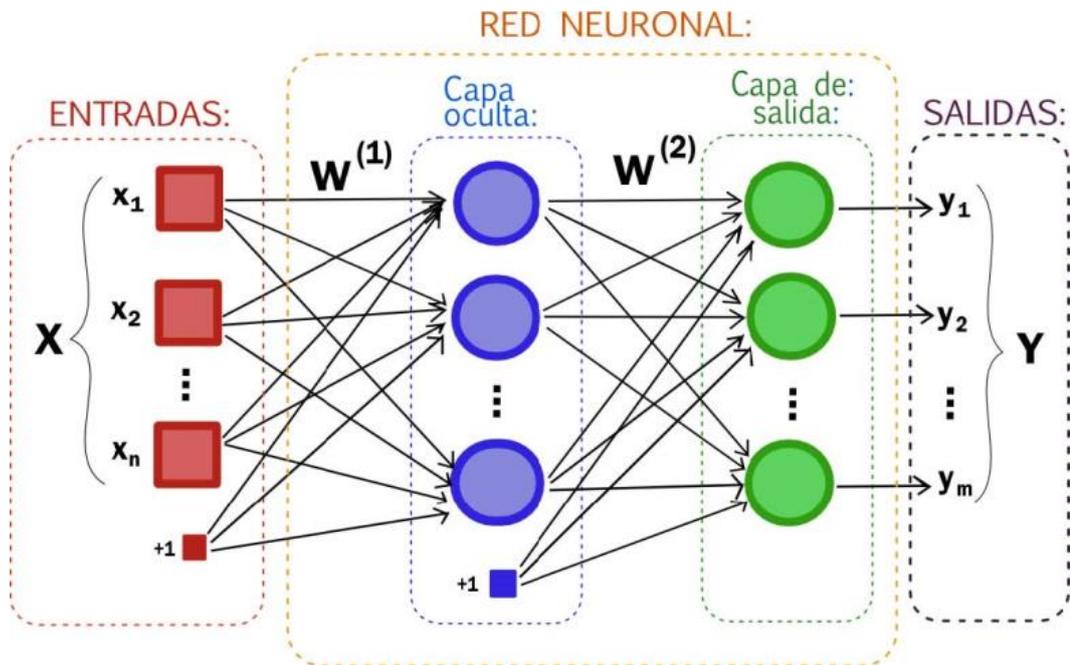


Figura 2.12: Estructura de una red neuronal de tres capas. (Fuente: [ceres.ugr.es](http://ceres.ugr.es))



### 2.7.3 TIPOS DE REDES NEURONALES

La salida de cada elemento de proceso (de cada neurona) puede ser conectada a través de los pesos correspondientes con sus propias entradas, con capas precedentes, con capas sucesivas y con el mundo exterior. Entonces, desde el punto de vista del flujo de la señal de una capa a otra podemos establecer una clasificación y distinguir entre dos tipos de redes:

1. Si las entradas y las señales de las capas intermedias se propagan siempre hacia adelante, la red se denomina **estática o red feedforward**. En este caso toda la información se dirige siempre hacia la salida y no hay caminos de vuelta permitidos.
2. En cambio, si algunos estados intermedios o salidas de la red son devueltos hacia atrás decimos que la red es **dinámica o red recurrente**. Aquí las señales son reutilizadas, es decir, las salidas actuales del sistema dependen de las salidas anteriores, y ésta es una característica muy importante de la que carece el anterior tipo de red.

El esquema de las redes recurrentes es diferente de los vistos anteriormente. En la figura 2.13 vemos un esquema para una red neuronal recurrente.

Las redes recurrentes son algo más potentes debido a que son secuenciales, a diferencia de las redes feedforward que son combinacionales. Muchos problemas han sido resueltos usando las redes tradicionales (por tradicionales entendemos las *feedforward*). Estas redes producen una salida que se basa exclusivamente en la entrada actual. Pero también existen algunos problemas donde es necesario tener en cuenta la información pasada además de la actual, problemas como el



control de sistemas no lineales e identificación, y modelos en tiempo continuo (en contraposición a los modelos basados en el tiempo discreto fácilmente tratados con redes *feedforward*). Esta realimentación de las salidas hacia las entradas permite a este tipo de redes presentar su comportamiento temporal, de forma que sus salidas pueden ser estables y converger o bien oscilar indefinidamente.

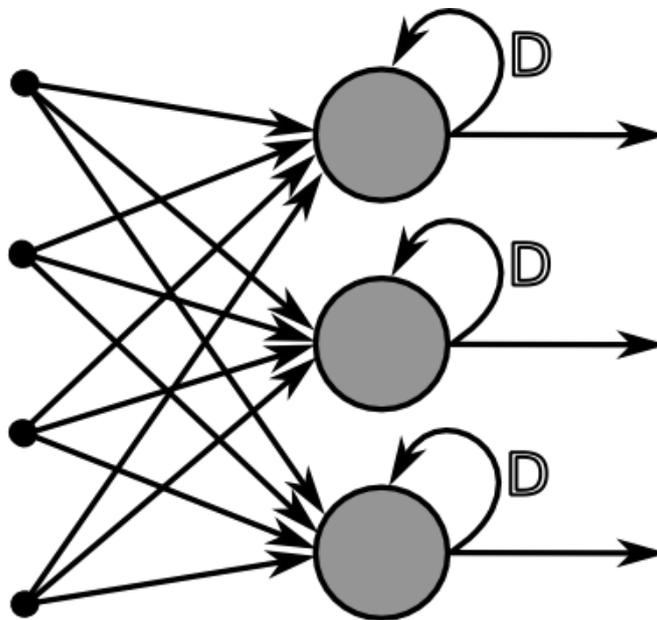


Figura 2.13: Esquema de una red recurrente. (Fuente: [github.com](https://github.com))

Las redes recurrentes generalmente están provistas de un vector con las condiciones iniciales, de forma que este vector se utiliza como primera entrada al sistema (representando el estado inicial) y que una vez que el sistema está en marcha, las salidas que produce son usadas como entradas para los demás instantes. Vemos pues que las redes recurrentes poseen una memoria que permite el aprendizaje de patrones temporales, disponemos de una capa de contexto que proporciona a la red la habilidad de



reconocer secuencias y de distinguirlas en función del contexto en el que se encuentren.

Esta cualidad de poder almacenar información para referencias futuras hace que estas redes sean útiles trabajos como:

- Reconocimiento de secuencias.
- Reproducción de secuencias. Ciclos.
- Asociación temporal (a cada entrada le corresponde una salida u otra dependiendo de las entradas anteriores).

## **2.7.4 APRENDIZAJE DE LAS RNA**

El aprendizaje se define como el mecanismo que hace que los pesos tomen los valores necesarios para desempeñar la tarea deseada. Los mecanismos de aprendizaje y determinación de pesos constituyen uno de los puntos críticos en la investigación con redes neuronales.

- **Aprendizaje supervisado:** Se presenta a la red un conjunto de patrones, junto con la salida deseada, e iterativamente, esta ajusta sus pesos hasta que la salida tiende a ser la deseada, utilizando para ello información detallada del error que se comete en cada paso. De este modo, la red es capaz de estimar relaciones entrada/salida sin necesidad de proponer una cierta forma funcional de partida.
- **Aprendizaje no supervisado o auto-organizado:** Se puede describir como la estimación de la función de densidad de probabilidad  $p(x)$  que describe la distribución de patrones del espacio de entrada, lo cual nos sirve para



reconocer regularidades en el conjunto de entradas, extraer rasgos o agrupar patrones según su similitud.

- **Aprendizaje híbrido:** En este caso ambos tipos de aprendizaje coinciden en una red, generalmente en distintas capas de neuronas.
- **Aprendizaje reforzado:** Está situado entre el aprendizaje supervisado y el auto-organizado. Como en el primero, se utiliza información sobre el error cometido, pero en este caso existe una única señal de error, que representa un índice global del rendimiento de la red (indica lo bien o mal que está actuando). Como en el caso del auto-organizado, no se suministra explícitamente la salida deseada.

## **2.7.5 PERCEPTRÓN**

Fue el primer modelo de red neuronal artificial, desarrollado en 1958 por Rosenblatt. Despertó un gran interés en los años 60 debido a su capacidad para aprender a reconocer patrones sencillos.

Está formado por varias neuronas para recibir las entradas a la red, y una neurona de salida que es capaz de decidir cuándo una entrada a la red pertenece a una de las dos clases que es capaz de reconocer (Figura 2.14). La neurona de salida del Perceptrón realiza la suma ponderada de las entradas, resta el umbral y pasa el resultado a una función de transferencia de tipo escalón. La regla de decisión es responder +1 si el patrón presentado pertenece a la clase A, o -1 si el patrón pertenece a la clase B.

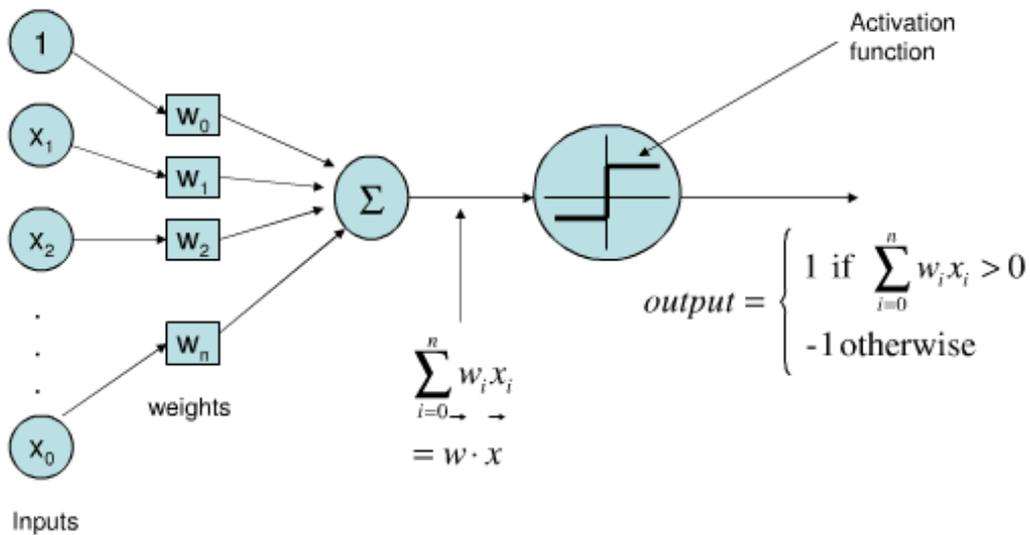


Figura 2.14: El Perceptrón. (Fuente: tex.stackexchange.com)

Al constar solamente de una capa de entrada y otra de salida con una única neurona, tiene una capacidad de representación bastante limitada. Solo es capaz de discriminar patrones muy sencillos y linealmente separables en el plano.

La separabilidad lineal limita a las redes con solo dos capas a la resolución de problemas en los cuales el conjunto de puntos (valores de entrada) sean separables geoméricamente. En el caso de dos entradas, la separación se lleva a cabo mediante una línea recta. Para tres entradas, la separación se realiza mediante un plano en el espacio tridimensional y así sucesivamente hasta el caso de N entradas, en el cual el espacio N-dimensional es dividido en un hiperplano.

El hecho de que el perceptrón solo es capaz de representar funciones linealmente separables se deduce de la expresión:

$$\text{salida} = \vec{W} \cdot \vec{E} \tag{37}$$



Donde  $\vec{W}$  es el vector de los pesos de las conexiones con la neurona de salida y  $\vec{E}$  es el vector de las entradas a la neurona de salida. Este produce un 1 si  $\vec{W} - \vec{E} > 0$ . Esto significa que todo el espacio de entradas se divide en dos a lo largo de un límite definido por  $\vec{W} - \vec{E} = 0$ , es decir, un plano en el espacio de entrada con coeficientes definidos por los pesos.

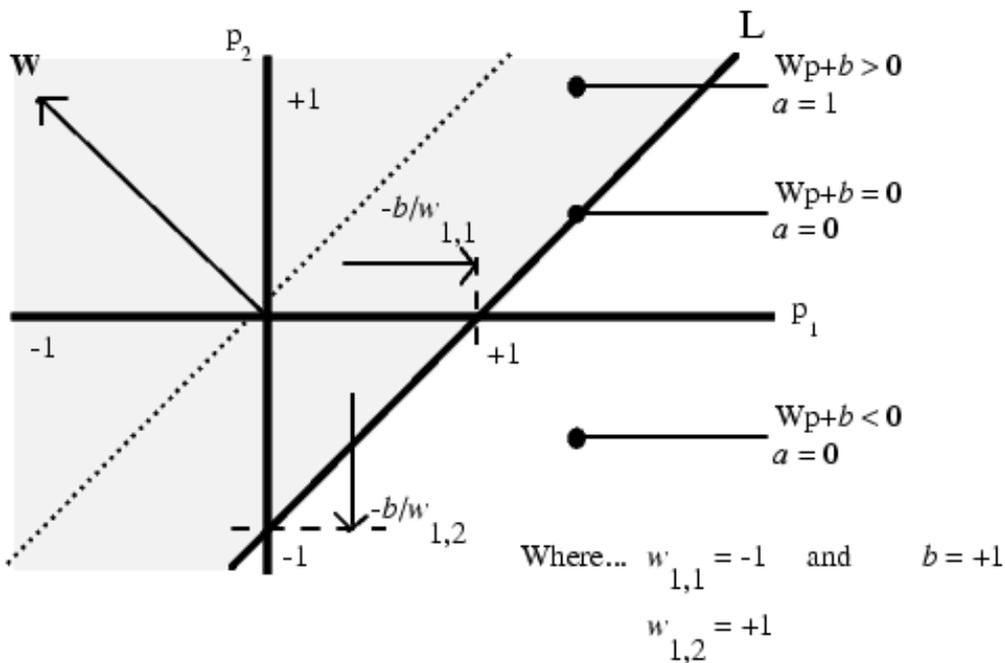


Figura 2.15: Regla de aprendizaje.

El algoritmo de aprendizaje del Perceptrón es de tipo **supervisado**, por lo que necesita que sus resultados sean evaluados y se realicen las modificaciones de los pesos que sean necesarias.

### 2.7.5.1 PERCEPTRÓN MULTICAPA

El Perceptrón multicapa o red multicapa con conexiones hacia adelante es una generalización del Perceptrón simple, que surge



como consecuencia de las limitaciones en la arquitectura del anterior, referidas a su problema de la separabilidad. En 1969, Minsky y Papert (*Minsky and Papert, 1969*) mostraron que la combinación de varios perceptrones simples podía resultar una solución adecuada para tratar ciertos problemas no lineales. Sin embargo, los autores no presentaron una solución al problema de cómo adaptar los pesos de la capa de entrada a la capa oculta, pues la regla de aprendizaje del Perceptrón simple no puede aplicarse en este escenario.

A pesar de ello, la idea de combinar varios perceptrones sirvió de base para estudios posteriores, que dieron lugar al desarrollo de un método de retropropagación de errores medidos en la salida de la red hacia neuronas ocultas. Se ha demostrado que el Perceptrón multicapa es un aproximador universal, en el sentido de que cualquier función continua sobre un compacto de  $\mathbb{R}^n$  puede aproximarse con un Perceptrón multicapa, con al menos una capa oculta de neuronas.

La habilidad del Perceptrón multicapa para aprender a partir de un conjunto de ejemplos, aproximar relaciones no lineales, filtrar ruido en los datos, etc. hace que sea un modelo adecuado para abordar problemas reales. En la actualidad, es una de las arquitecturas más utilizadas en la resolución de problemas.

## **2.7.6 VENTAJAS DE LAS RNA**

Debido a que están basadas en la estructura del sistema nervioso, las redes neuronales artificiales (RNA) tienen numerosas ventajas, tales como:

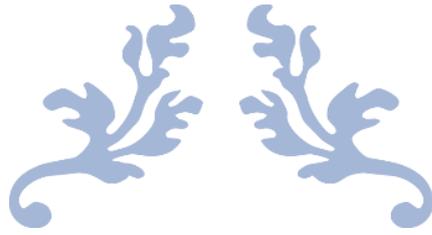
- **Aprendizaje:** Tienen la habilidad de aprender mediante una etapa denominada de aprendizaje, la cual consiste



en proporcionar a la RNA datos como entradas, a la vez que se le indica cuál es la respuesta (salida) esperada.

- **Auto organización:** Una RNA crea su propia representación de la información en su interior, descargando al usuario de esto.
- **Tolerancia a fallos:** Debido a que las RNA almacenan la información de forma redundante, pueden seguir respondiendo aceptablemente aunque se dañen parcialmente.
- **Tiempo real:** La estructura de una RNA es paralela, por lo que si es implementada en computadoras o dispositivos electrónicos, se pueden obtener respuestas en tiempo real.





---

# DESCRIPCIÓN DE LA PLANTA

---

CAPÍTULO 3





## **3.1 INTRODUCCIÓN**

En este apartado se va a describir la planta sobre la cual se ha realizado el estudio. Con ello, estudiaremos en profundidad su funcionamiento y la validez de los métodos que hemos utilizado, dentro de un sistema real.

En primer lugar, se tratará de dar una idea general del proceso de producción del azúcar a partir de la remolacha, y más concretamente, se hablará de la etapa de evaporación que tiene lugar en el proceso. Además, se expondrá una descripción del funcionamiento considerado normal y de las diferentes variables en las que puede ocurrir un fallo, con el objetivo de dar por válidos los métodos de detección y diagnóstico de fallos estudiados.

Los datos obtenidos para llevar a cabo nuestro estudio provienen de una fábrica azucarera real, concretamente de la sección de evaporación. Teniendo la posibilidad de utilizar datos del proceso de una planta real, resulta más adecuado realizar el estudio de los fallos disponiendo de un modelo matemático que nos permita introducir fallos a voluntad, con el objetivo de estudiar la respuesta de los métodos de análisis.

## **3.2 AZÚCAR DE REMOLACHA**

Podemos denominar comúnmente como azúcar a la sacarosa, cuya fórmula química es  $C_{12}H_{22}O_{11}$ , también llamado azúcar común o azúcar de mesa. La sacarosa es un disacárido formado por una molécula de glucosa y una de fructosa, que se obtiene principalmente de la caña de azúcar o de la remolacha. En ámbitos industriales, se usa la palabra azúcar o azúcares para designar los diferentes monosacáridos y disacáridos, que generalmente tienen sabor dulce, aunque, por extensión, se refiere a todos los hidratos de carbono.



El azúcar es una importante fuente de calorías en la dieta alimenticia moderna, pero es frecuentemente asociado a calorías vacías, debido a la completa ausencia de vitaminas y minerales.

Además de azúcar de mesa, en el proceso productivo de la remolacha azucarera, se obtienen subproductos como son la pulpa y la melaza. Importantes en la producción de alcoholes o piensos para el ganado.

### **3.2.1 HISTORIA DEL AZÚCAR**

El azúcar cristalizado era ya conocido en Persia en el siglo IV a.C., y provenía seguramente de la India, donde se extraía de una variedad salvaje de caña.

El cultivo de la remolacha se desarrolla en Francia y España durante el siglo XV, inicialmente, se daba la recolección de esta planta por sus hojas, que probablemente equivalían a las espinacas y acelgas. A partir de entonces la raíz ganó popularidad, especialmente la de la variedad roja conocida como remolacha.

En 1747, el científico alemán Andreas Marggraf demostró que los cristales de sabor dulce obtenidos del jugo de la remolacha eran iguales a los de la caña de azúcar. En 1801, se construyó la primera fábrica de azúcar en Cunern, Baja Silesia (Alemania). La incipiente industria azucarera basada en la remolacha tal vez no hubiera resistido la competencia con la caña de azúcar como materia prima si no hubiera sido por los bloqueos ingleses al continente europeo, lo que obligó a la búsqueda de nuevos recursos.

En 1811, por decreto de Napoleón, 32.000 Hectáreas de remolacha debían ser cultivadas en la Europa bajo su dominio para suplir el déficit de azúcar causado por el bloqueo naval inglés, que impedía el ingreso de azúcar de caña proveniente de



las colonias francesas, contribuyendo de este modo al establecimiento de las fábricas. En pocos años se construyeron más de cuarenta fábricas de azúcar de remolacha, distribuidas desde el norte de Francia, Alemania, Austria, Rusia y Dinamarca.

La introducción de una verdadera industria remolachero-azucarera en España se inicia a finales del siglo XIX, cuando comienza a tener una mayor importancia la remolacha, en detrimento de la caña de azúcar. La fuerte implantación de esta industria en nuestra vecina Europa, así como la pérdida de la economía colonial constituirá el inicio de una industria propia.

Alrededor de la década de los 70 el sector consolida su actividad en España. Se impulsa una ampliación y mejora de las fábricas, lo que permite una mayor productividad y automatización de las mismas. Además de la implantación de modelos de mercado que basan el precio de la remolacha en su riqueza sacárica, consiguiendo un aumento de la productividad agrícola. Se hundieron las zonas del Ebro y Vega Granadina, potenciando la cuenca del Duero, región en la que nos encontramos.

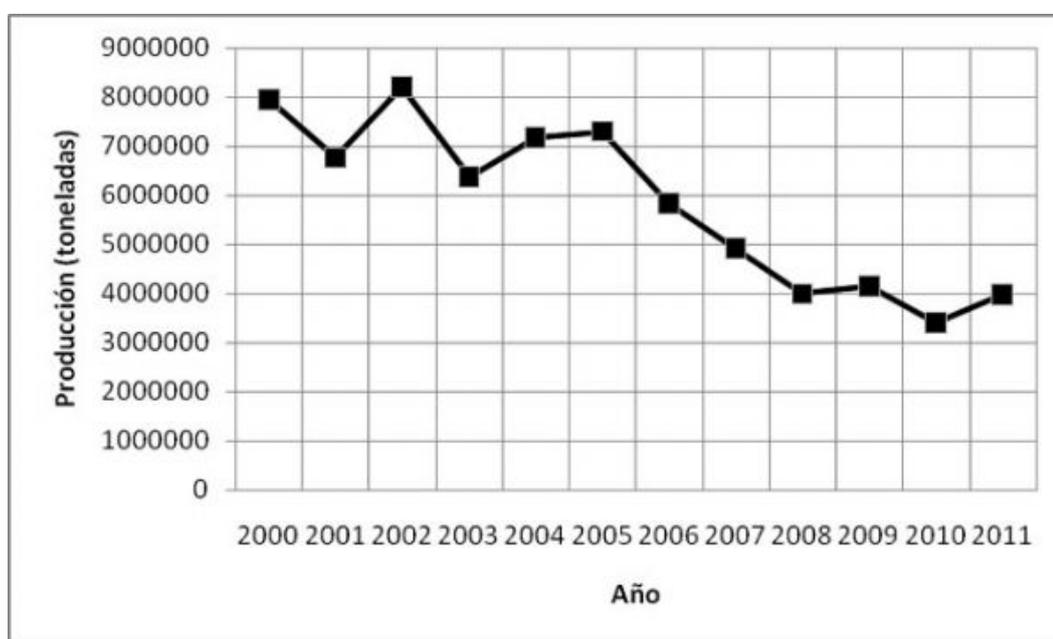
Caracterizada tradicionalmente por la trashumancia y teniendo en cuenta su baja y desordenada competitividad inicial la industria azucarera se ha ido trasladando por casi todo el territorio nacional, y es a principios de siglo cuando mejor se encuentra el sector en nuestro país, capaces de competir y adaptarnos al mercado europeo.

Es entonces cuando se da un mayor auge azucarero en Europa, recién estrenado siglo, se cultiva tres veces más azúcar de remolacha que cinco años atrás y en cifras absolutas de producción, se supera a la caña de azúcar, debido tanto a la modernización del cultivo como a la disminución de la producción de remolacha forrajera. Casi el 90% del azúcar que se consume en Europa durante estos años es de producción interna.

Todo cambia en estos últimos años, en los cuales se produce una tendencia de descenso de la producción. Se trata de uno de los



sectores que más abandono está sufriendo debido a los cambios legislativos producidos por la Europa comunitaria. Muchos agricultores están decidiendo apostar por cultivos que ofrezcan más rentabilidad. Y es que, a la vista de las últimas previsiones de campaña de remolacha azucarera de invierno, concentrada en la zona norte, todo apunta a que se mantendrá el descenso productivo, que se inició en 2006 con la reforma de la Organización Común de Mercado (OCM) del azúcar, y la apuesta de la Unión Europea por el abandono del cultivo a cambio de ayudas compensatorias.



**Figura 3.1:** Evolución temporal de la producción de azúcar en España.

Fuente: [www.infoagro.com](http://www.infoagro.com)

En la figura 3.1 podemos ver representada la evolución temporal de la producción en toneladas de azúcar en plantas españolas, desde 2000 hasta 2011.

Hoy en día solo quedan en pie cinco de las 12 fábricas azucareras que había en todo el territorio nacional. Alrededor de



unos 7000 agricultores han abandonado el sector en los últimos años. Sufriendo por tanto un decremento de la producción de casi medio millón de toneladas desde la campaña 2005-2006.

La temporada 2009-2010 por tanto, informó acerca de una caída de la producción como de la superficie cultivada. Así, según los resultados de la Encuesta sobre Superficies y Rendimientos de Cultivos, la superficie sembrada de remolacha azucarera en España se redujo en 2009 en un siete por ciento con respecto a 2008, hasta situarse en unas 50.000 hectáreas, de las que la mayor parte corresponden a la zona norte del país (unas 38.000), y el resto en Andalucía.

En la actualidad, los cultivadores de la Unión Europea sembrarán el año que viene la última remolacha bajo el sistema de cuotas de producción, que expirará el 30 de septiembre de 2017 tras más de 50 años de existencia. Junto con las cuotas, desaparecerán también los precios mínimos garantizados y la limitación de las importaciones de azúcar. En España, el sector acoge la apertura total del mercado azucarero con “cierto vértigo” e incertidumbre, en un contexto de precios bajos a escala mundial y con mucho stock de azúcar en Europa y en grandes países productores como Brasil, Tailandia o India. En cualquier caso, tanto los agricultores como la industria transformadora consideran que están preparados para afrontar el final del régimen de cuotas.

En el nuevo escenario, la principal debilidad competitiva de los productores españoles es el alto coste que conlleva el regadío, un coste que no soportan otros competidores comunitarios, como Francia. Por su parte, la industria azucarera tiene un problema de dimensión tras la reestructuración derivada de las últimas reformas de la Organización Común de Mercados (OCM). En concreto, España tiene asignadas 498.480 toneladas de azúcar, que se reparten entre cuatro fábricas de AB Azucarera



Iberia (378.480 toneladas) y una fábrica de la cooperativa ACOR (120.000 toneladas). Los datos al respecto, han sido obtenidos en diversas páginas web.

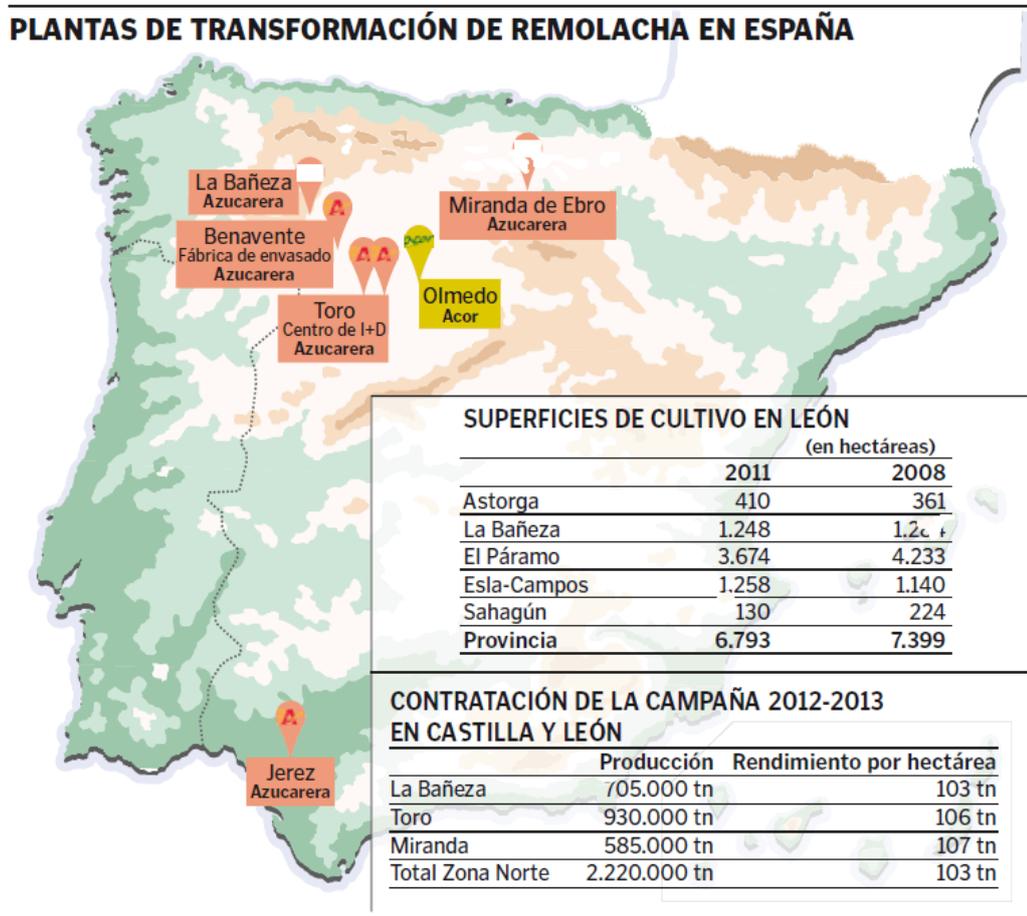


Figura 3.2: Plantas de transformación de remolacha en España.  
Fuente: www.diariodeleon.com

### **3.3 PROCESO PRODUCTIVO DEL AZÚCAR DE REMOLACHA**

En este apartado trataremos de dar una visión general del proceso productivo al que es sometida la remolacha para obtener azúcar, principalmente, junto con otros subproductos. El proceso



puede ser diferente dependiendo de la fábrica, pero en rasgos generales es siempre igual.

### **3.3.1 RECOLECCIÓN Y ENTREGA DE LA REMOLACHA**

La cosecha de la remolacha comienza en otoño o principios de invierno. Se realiza extrayendo la raíz de la planta del suelo, donde se encuentra la mayor riqueza de la planta. Esta es transportada a las fábricas por camiones para que comience su procesado.

La remolacha puede ser descargada en seco o con utilización de agua. La descarga del camión se realiza siempre en seco con camiones basculantes o plataformas basculantes. Una cantidad insignificante de remolacha en España llega a la fábrica con vagones (vagones de fondo plano), se descarga con la ayuda de agua a presión.

### **3.3.2 ALMACENAMIENTO Y TRANSPORTE DE LA REMOLACHA**

Se dispone de un sistema de almacenamiento de remolacha controlado, con el fin de entrelazar la discontinuidad de la descarga de camiones con el proceso continuo de producción de azúcar que se inicia en el lavadero de remolacha. En instalaciones denominadas de “transporte húmedo”, se extrae la remolacha del silo de almacenamiento por transporte hidráulico (arrastre de remolacha en corriente de agua).

Las instalaciones de transporte en seco utilizan cintas transportadoras que llevan la remolacha de forma continua y dosificada del punto de almacenamiento, denominado “punto



fijo”, al lavadero. Las cantidades almacenadas dependen en gran medida de la climatología y la capacidad productiva de la región, en la medida de lo posible se ha de mantener siempre abastecida la fábrica. El almacenamiento siempre se realiza a la intemperie, independientemente del clima de la región donde se encuentre la fábrica.

El transporte en seco posee ciertas ventajas de tipo técnico en relación al transporte húmedo, ya que da un menor consumo de agua y electricidad y además reduce las pérdidas de sacarosa al no mantener el producto en contacto directo con el agua. En contra punto, las instalaciones de transporte húmedo poseen un menor tiempo de amortización. Además, el circuito de transporte entre el silo y el lavadero es capaz de eliminar un alto contenido de piedras y materia vegetal.

### **3.3.3 LAVADO DE LA REMOLACHA**

Es la operación encargada de eliminar la tierra y materia vegetal que todavía pueda contener la materia prima. Se lleva a cabo de múltiples formas, mediante lavaderos fijos con álabes rotativos en su interior o con lavadoras rotativas con álabes fijos (tambores). La idea es hacer circular la remolacha y el agua de lavado en diferente dirección para así separar toda la materia indeseada.

La materia desechada (piedras, tierra, hojas, pequeños rabillos, etc.) es separada del agua de lavado mediante cribas y demás procesos mecánicos. El agua de lavado es reciclada para transporte o llevada a las balsas de decantación. Los desechos vegetales son o bien incorporados de nuevo al proceso, en el caso de que puedan ser aprovechados, o lavados y utilizados como alimento para el ganado.



Es necesario añadir al agua de lavado y al agua de transporte productos que eviten la aparición de espumas que puedan obstaculizar el funcionamiento del proceso de fabricación.

### **3.3.4 EXTRACCIÓN DEL JUGO**

Cuando termina la etapa de lavado, la remolacha es transportada a unos molinos llamados corta-raíces, donde esta es cortada de forma determinada, llamando a estos cortes cosetas, para aumentar así su superficie efectiva para el intercambio.

La materia prima ya cortada se lleva al difusor donde la sacarosa que se extrae pasa al agua que circula en dirección contraria. Esta operación es realizada con la mezcla a una temperatura de unos 73 °C, a la cual se da el paso de la sacarosa a la parte líquida. Dando como resultado una separación de la pulpa, de un primer jugo azucarado no depurado. Existen tres tipos diferentes de difusores, horizontal fijo, horizontal rotativo y difusor vertical.

### **3.3.5 TRATAMIENTO DE LA PULPA**

Esta pulpa con alto contenido en agua es sometida a varios procesos de prensado mecánico, seguidos de otros procesos de secado térmico que permite reducir su humedad en un 88%. De este modo se le confieren a la pulpa mejores propiedades de conservación y manipulación.

La pulpa de la remolacha es un importante producto como materia prima destinada a la alimentación animal. Hay veces que se le añaden a la pulpa diferentes productos para mejorar su valor nutricional.



### **3.3.6 DEPURACIÓN Y FILTRADO DEL JUGO**

El jugo azucarado proveniente del difusor tiene un contenido de materia seca cercano al 15%. Contiene macromoléculas y partículas en suspensión que deben ser eliminadas para hacer posibles las fases posteriores del proceso de fabricación del azúcar. Esto se logra por medio de un proceso de depuración con lechada de cal que consiste en fijar los componentes que no son azúcar en el carbonato cálcico, que convierte el compuesto obtenido en insoluble (espumas de cal). La mezcla puede entonces filtrarse para obtener así un jugo de mayor pureza, dependiendo de las zonas y las campañas.

### **3.3.7 EVAPORACIÓN DEL JUGO**

El jugo azucarado resultante de la depuración tiene un contenido de materia seca que todavía se encuentra en torno al 15%. Por ello, es necesario aumentar este contenido de materia seca para cristalizar el azúcar. El proceso de evaporación permite aumentar la materia seca del 15% a más del 68%. Está basado en un principio de intercambio de calor entre el jugo y el vapor producido en las calderas. Este proceso se explicará más detenidamente, ya que es el que someteremos a estudio en nuestro trabajo.

### **3.3.8 GENERACIÓN DE VAPOR Y ELECTRICIDAD**

La industria azucarera precisa de altas cantidades de energía en forma de vapor y electricidad. La mayor parte de la energía requerida para el proceso de la fabricación de azúcar se obtiene quemando gas o fuel en una caldera que lo convierte, por medio de una planta de cogeneración, en vapor y electricidad.



El vapor producido a alta presión en las calderas es introducido a su vez en una turbina de un generador para abastecer las necesidades eléctricas de la fábrica. Este vapor que proviene del escape de estas turbinas, a una menor presión será el utilizado en el primer efecto de evaporación.

El factor de rendimiento del combustible utilizado en estos sistemas combinados de calor y energía puede llegar al 80%. Esta eficacia no es alcanzada en ningún otro diseño de aprovechamiento de la energía dado en la industria.

### **3.3.9 CRISTALIZACIÓN Y CENTRIFUGADO**

Tras la evaporación, el jugo espeso se cuece para evaporar algo más del agua que contiene y para comenzar la cristalización, que se logra añadiendo azúcar impalpable. Esta operación produce un magma que contiene cristales de azúcar y de jarabe (mieles).

El proceso de cristalización se lleva a cabo en las tachas, recipientes sometidos a vacío y calentados con vapor. El proceso de centrifugado se lleva a cabo en un separador sólido-líquido. Se somete a este proceso al magma procedente de las tachas, separándose el azúcar del jarabe que lo rodea.

De esta forma obtenemos el azúcar de primera, que pasa al proceso de secado y enfriado. Por otro lado, el jarabe resultante vuelve a ser incorporado a la siguiente fase de cristalización, repitiéndose esto hasta tres veces, el azúcar de peor calidad obtenido de la segunda y tercera fase es disuelto junto al jarabe de evaporación. Tras la última fase, queda una solución de azúcar residual, las melazas.

### **3.3.10 SECADO DEL AZÚCAR**

Tras el centrifugado, el azúcar debe secarse y enfriarse para evitar la condensación del agua y para que el azúcar blanco tenga una manipulación satisfactoria. Se utilizan secaderos de lecho



fluido para el secado y enfriado, así como los secaderos de tambor rotativo y secadero de bandejas.

### **3.3.11 PRODUCCIÓN DE MELAZAS**

Las melazas, que constituyen la fracción líquida final de la cristalización, son empleadas en su totalidad, ya sea como añadidos a la pulpa o para producción de alcoholes, levaduras, etc.

### **3.3.12 ALMACENAMIENTO Y EMBALAJE**

Tras la cristalización, el secado y el enfriado, el azúcar está listo para ser almacenado o envasado.

### **3.3.13 TRATAMIENTO DE AGUAS**

El agua se toma directamente del río mediante las correspondientes instalaciones de bombeo y se envía a las balsas de almacenamiento de las fábricas donde, tras tratamientos de adecuación, es utilizada en los distintos procesos de fábrica.

Las fábricas azucareras disponen de circuitos cerrados de agua, aprovechando al máximo los métodos de reciclado y reutilización.

## **3.4 PROCESO DE EVAPORACIÓN**

En esta sección pasaremos a explicar cómo funciona el proceso de evaporación en una fábrica real. Se dará una idea general del funcionamiento y finalidad de dicha sección, seguida de una descripción de los equipos que en ella intervienen. Finalmente se dará una explicación de los procesos de control empleados para mantener su correcto funcionamiento.



### 3.4.1 DESCRIPCIÓN GENERAL

La sección de evaporación es una de las más significativas dentro de una fábrica azucarera. Su importancia radica en su alto consumo de vapor y el consiguiente gasto energético y económico que este acarrea. Compuesta de procesos sencillos, su control y precisión es de vital importancia puesto que su consumo de vapor lleva asociado efectos en otras zonas del proceso, ya que este es reutilizado para así reducir consumos de energía durante el proceso.



**Figura 3.3:** Esquema general de la fase de evaporación. Imagen extraída del Capítulo 6 de la Tesis Doctoral; Entornos distribuidos para simuladores de procesos.

El jugo procedente de la depuración es una disolución azucarada con una gran cantidad de agua y poco contenido en materia seca (llamada comúnmente Brix<sup>1</sup>). El objeto de la evaporación es aumentar el contenido de materia seca evaporando así, gran parte del agua que contiene el jugo. Un esquema general de la sección de evaporación se muestra en la figura 3.3. La

---

<sup>1</sup> Los grados Brix, °Bx, miden el cociente total de sacarosa disuelta en un líquido. Su uso más común se da en porcentaje.



evaporación se realiza por medio de un sistema de “evaporación de efecto múltiple” en varias cajas o evaporadores.

En esta sección usamos vapor producido en calderas industriales a alta presión, utilizado para la generación de energía eléctrica en turbinas y posteriormente, ya a una menor presión, reutilizado como fluido calefactor en esta sección. El sistema de “evaporación de efecto múltiple” tiene por objeto reducir el consumo de vapor reutilizando los vapores que se producen en cada efecto como fuente de calor para el siguiente. Ello conlleva la reducción progresiva de la presión de los sucesivos efectos, operándose en el último por debajo de la presión atmosférica y por ello a temperatura menor de 100° C.

La evacuación de los vapores de este último efecto se realiza provocando vacío en la última caja de evaporación. Esto se realiza mediante condensación barométrica con agua fría que está en circuito cerrado tras enfriarse en las torres de refrigeración. El jugo concentrado que sale de la instalación de evaporación se denomina jarabe. En éste se disuelven azúcares de los bajos productos para la obtención de otro jarabe denominado jarabe estándar, que una vez filtrado está listo para el proceso de cristalización.

## **3.4.2 DESCRIPCIÓN DE LOS EQUIPOS UTILIZADOS EN LA PLANTA**

### **3.4.2.1 EVAPORADORES**

Son los equipos en los que se produce la evaporación del jugo. Su objetivo es concentrar un sólido disuelto en un líquido. Éste consta de una carcasa donde se condensa el vapor y de una cámara de separación donde tiene lugar la ebullición de la



solución a concentrar, la cual, al perder disolvente aumenta su proporción de soluto.

El interés de este sistema de múltiple efecto radica, fundamentalmente, en la mejor utilización del vapor. En estos sistemas se llama primer efecto a los evaporadores que reciben directamente el vapor de calderas, segundo efecto a los que reciben el vapor procedente del primer efecto y así sucesivamente. Así mismo, el vapor que sale de cada efecto se denomina vahos, por lo que el vapor que sale del primer efecto se llamará “vahos de primera”.

En cuanto a la circulación de cada uno de los fluidos, el vapor circula desde el efecto de mayor presión al de menor presión. La solución a concentrar lleva la misma dirección que el vapor circulante (flujo directo). Este sistema implica una mayor sencillez.

La circulación de la solución a concentrar a través de los diversos evaporadores se debe a la diferencia de presión de cada cámara de concentración, para lo cual es necesario aplicar vacío al último efecto. Los flujos caloríficos entre la cámara de condensación (carcasa) y la cámara de concentración se originan debido a la diferencia de temperatura existente entre ambas.

Respecto a la selección del tipo de evaporador, se suele seleccionar aquel que tiene un coeficiente de transferencia de calor más alto en las condiciones de operación deseadas. En nuestro caso, se han descrito dos tipos de evaporadores, Robert y de película descendente o Fallstrom.

En los **evaporadores tipo Robert** (Figura 3.4) el jugo entra por la parte baja de estos, una vez allí comienza a hervir debido a la condensación del vapor introducido. Se compone de un cuerpo cilíndrico vertical con los tubos sujetos entre dos placas tubulares horizontales.

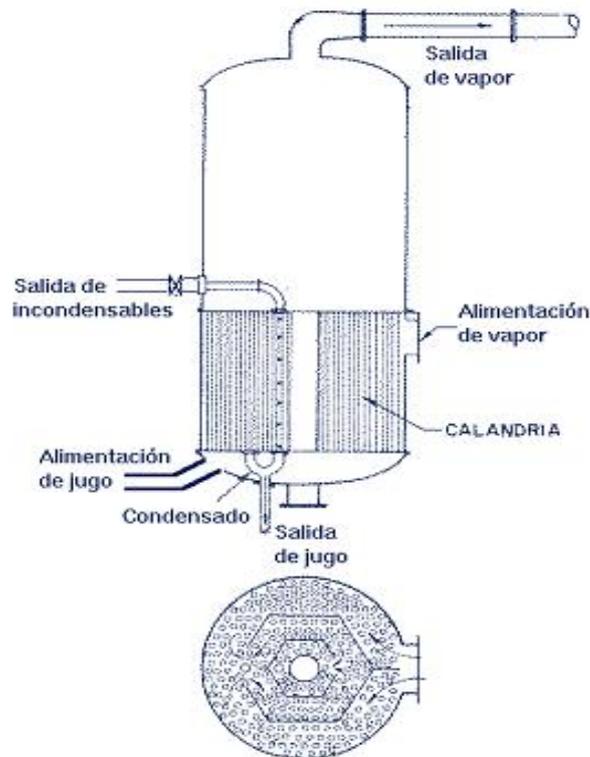
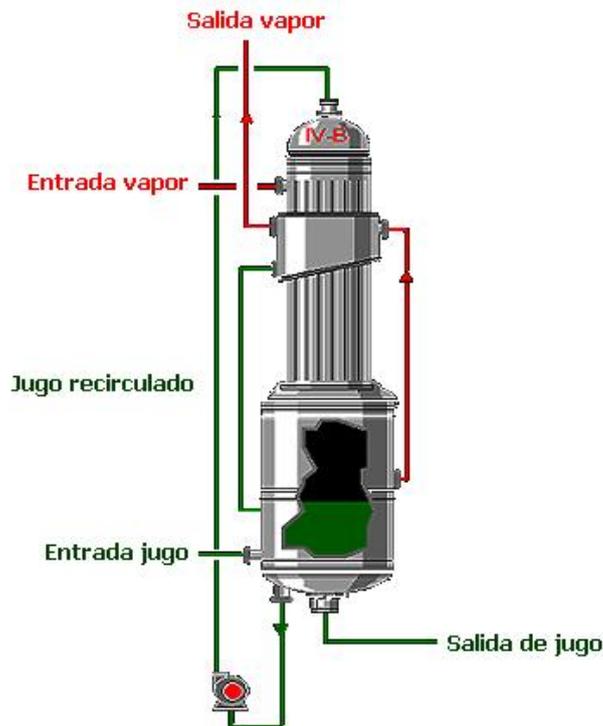


Figura 3.4: Evaporador de tipo Robert.

El jugo termina rebosando por un gran tubo central de bajada que permite obtener una buena circulación de la solución hacia el exterior, lo cual es una gran ventaja, mientras que su mayor desventaja es la cantidad de espacio que ocupan.

En los **evaporadores de película descendente o Fallstrom** (Figura 3.5), el líquido a concentrar se bombea mediante unos pulverizadores, recirculándolo, desde un depósito inferior y entra por la parte superior al igual que los vahos que abandonan el equipo mientras que el líquido concentrado lo hace por el fondo de la misma.

De este modo tenemos un evaporador en el que el material a tratar fluye hacia abajo, en forma de película descendente, por las paredes interiores de un haz de tubos calentados con vapor de agua.



**Figura 3.5:** *Evaporador de película descendente. Imagen extraída del Capítulo 6 de la Tesis Doctoral; Entornos distribuidos para simuladores de procesos.*

La principal ventaja de estos evaporadores es que proporcionan una película más delgada y rápida por lo que su coeficiente de transmisión de calor es bastante alto. Además, permiten operar con pequeñas diferencias de temperatura, funcionando mejor con líquidos viscosos. Por todo esto este tipo de evaporadores se suele utilizar en los últimos efectos de la evaporación.

A la hora de modelar estos dispositivos, se han tomado unas suposiciones generales que hacen este más sencillo desde el punto de vista matemático y además eliminan ciertas suposiciones que podrían llevarnos a error.

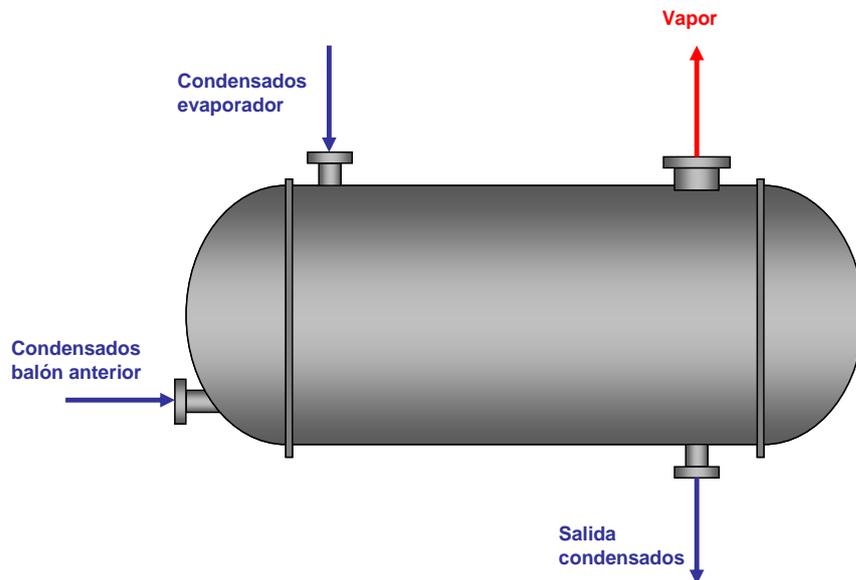
- En el interior de todas las cámaras existe mezcla perfecta.



- El vapor que condensa en las cámaras está saturado, es decir, no se considera la posibilidad de que el vapor llegue sobrecalentado a las cámaras de condensación cediendo calor sensible además de calor latente.
- Se supone que no existen pérdidas de energía.
- Se considera que los vapores se comportan como gases ideales.
- Se modelan la fase líquida y la fase vapor por separado.

### **3.4.2.2 BALONES DE EXPANSIÓN**

En ellos se produce la expansión de los condensados de la evaporación, lo que permite el aprovechamiento de los vahos producidos.



**Figura 3.6:** Balones de expansión. Imagen extraída del Capítulo 6 de la Tesis Doctoral; Entornos distribuidos para simuladores de procesos.



### **3.4.2.3 SATURADOR**

En este equipo se produce la saturación del vapor sobrecalentado mediante el aporte de agua.

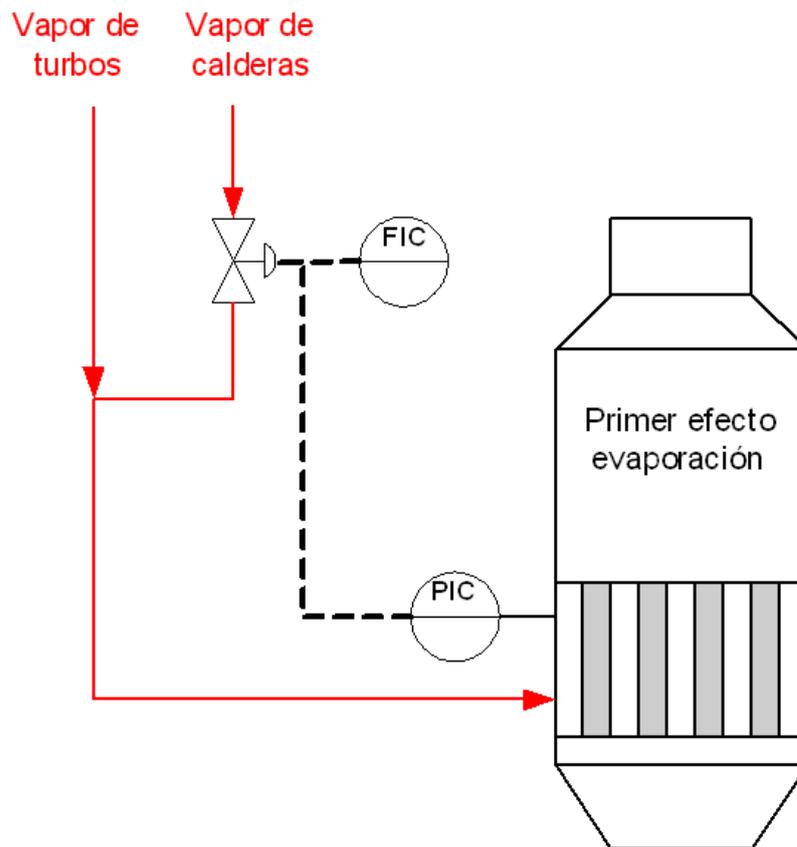
### **3.4.2.4 OTROS EQUIPOS**

Existen además otra serie de equipos auxiliares utilizados en esta sección, mezcladores, acumuladores de vapor, intercambiadores de calor, etc.

## **3.4.3 DETALLES DE FUNCIONAMIENTO Y LAZOS DE CONTROL**

La sección de evaporación posee diferentes lazos de control, a los cuales se les encarga la tarea de obtener a la salida de esta etapa, una solución con el mayor contenido posible de materia seca. Lo que se traduce como principal objetivo de control, mantener la obtención de un jarabe con un brix elevado.

Podemos mantener un brix elevado en el tiempo, a través de la cantidad de agua que se esté evaporando del jugo procedente de la etapa de depuración. La cantidad de agua evaporada depende directamente del caudal y la presión del vapor introducido en el primer efecto y el vacío existente en el último efecto.



**Figura 3.7:** Control del vapor de entrada en el primer efecto. Imagen extraída del Capítulo 6 de la Tesis Doctoral; Entornos distribuidos para simuladores de procesos.

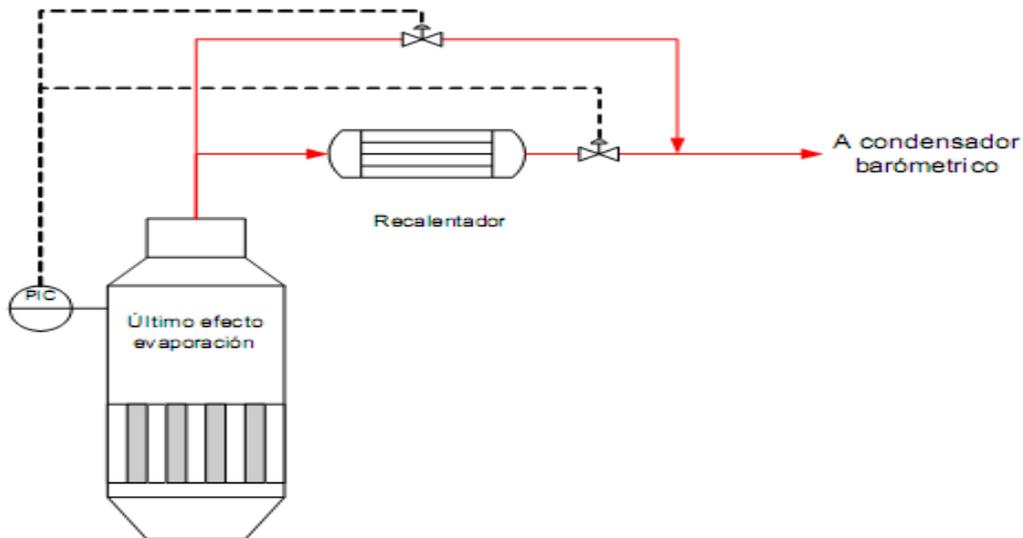
**El vapor de entrada del primer efecto** está compuesto de una parte a alta presión procedente directamente de las calderas y otra parte a menor presión, que, aunque también procede de las calderas, ha pasado previamente por las turbinas.

La utilización de vapor directo de las calderas, a alta presión resulta mucho menos económico, de forma que se controla este caudal mediante una válvula, que se mantiene lo más cerrada posible y así utilizar de forma mayoritaria vapor a una menor presión (Figura 3.7).

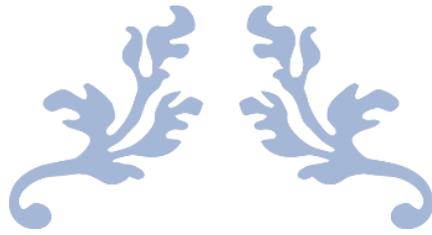


También se controlan la cantidad de aguas condensadas que pasan de los balones de expansión cero al uno, equivalente a controlar el caudal de condensados del primer efecto. Se mide el caudal que está circulando entre los dos balones y se actúa sobre la válvula situada en la línea de vapor de calderas.

**Controlando el vacío en el último efecto**, nos aseguramos de que el jugo circula correctamente, ya que este fluye debido a la diferencia de presión entre los efectos. Este regulador debe ser capaz de asegurar que el vapor que llega al condensador barométrico sea mínimo.



**Figura 3.8:** Control del vapor de entrada en el último efecto. Imagen extraída del Capítulo 6 de la Tesis Doctoral; Entornos distribuidos para simuladores de procesos.



---

# ANÁLISIS DE RESULTADOS

---

CAPÍTULO 4





## **4.1 INTRODUCCIÓN**

En este capítulo se presentará la aplicación de las técnicas utilizadas para el diseño del sensor software, utilizado para estimar el valor de la materia seca (%MS) como variable de calidad del proceso, así como para la detección y diagnóstico de fallos en el proceso de producción de materia seca en la industria azucarera. Los sensores de materia seca comerciales son, en general, caros e imprecisos, por lo que resulta interesante estudiar y desarrollar sensores software para estimar esta medida. En concreto, en este trabajo, se ha estimado el contenido en materia seca del jarabe de salida de la sección de evaporación. Para ello, contamos principalmente de dos métodos. Uno utiliza una estimación basada en una regresión de mínimos cuadrados parciales (PLS), y diferentes métodos derivados del mismo. Los relacionaremos y compararemos con el segundo método, en el que se utilizan redes neuronales artificiales.

El objetivo es diseñar un sensor software para estudiar la calidad del proceso, la cual, para ser posible medirla, será relacionada con la concentración de sacarosa (*Brix*) en el último evaporador.

Posteriormente se analizarán e interpretarán los resultados obtenidos tras aplicar a los datos recogidos en la planta el sensor software para estimar la variable de calidad, así como se implementará un sistema de detección y diagnóstico de fallos que nos indicará si el proceso está bajo control estadístico y por tanto si las variables usadas para estimar la variable de calidad tienen errores o están en condiciones normales de operación. Se tratará en profundidad cómo se ha realizado la implementación en Matlab de los métodos propuestos, así como las diferencias entre los resultados arrojados por cada uno, comparándolos posteriormente según su robustez. Para ello, se introducirán algunos ejemplos concretos de anomalías en la planta para



determinar si el sensor es capaz por un lado de dar una estima correcta de la variable de calidad y por otro de realizar una correcta detección del fallo, y dar un primer diagnóstico acerca de las causas que lo han provocado, es decir, de las variables que se han ido fuera de su funcionamiento correcto.

## **4.2 TOMA DE DATOS DE LA PLANTA**

Una vez se ha explicado las características y el funcionamiento de la planta en el Capítulo 3, es momento de exponer la forma de obtener los datos de la misma. Estos datos han sido recogidos directamente del proceso en funcionamiento normal de la planta.

La muestra recogida debe tener un volumen de datos lo más grande posible, ya que en un volumen limitado podrían no darse diferentes situaciones que el sensor pueda aprender, pudiendo arrojar falsas alarmas al analizar un nuevo conjunto de datos. Por ello, se ha recogido una muestra de 34560 datos de la planta en funcionamiento normal, es decir, una situación exenta de cualquier fallo o funcionamiento anómalo. Esto no supondrá un problema para el sensor a la hora de tratar los datos, debido a que tanto con los métodos de regresión por mínimos cuadrados parciales (PLS) como con las Redes Neuronales Artificiales (RNA) se consigue reducir notablemente la dimensionalidad de los datos.

Es imprescindible que la recolección de datos sea espaciada en el tiempo y registrada a intervalos regulares, de forma que se refleje la mayor normalidad en los mismos. También tenemos que tener en cuenta que los datos de arranque de la planta pueden arrojar resultados que indiquen un comportamiento anómalo del proceso, aunque realmente no sea así.



## 4.2.1 VARIABLES DEL PROCESO

En total contamos con 52 variables de la planta, 51 de ellas pueden verse en la tabla 4.1, la 52 es la variable de salida que queremos predecir, el contenido de materia seca en el sexto evaporador (CS6):

Nomenclatura		Variables	Unidades
1	CS1	Contenido de materia seca (evaporador 1)	g/100 g
2	P1	Presión de vapor (evaporador 1)	bar
3	P2	Presión de vapor (evaporador 2)	bar
4	P3	Presión de vapor (evaporador 3)	bar
5	P4	Presión de vapor (evaporador 4)	bar abs
6	P5	Presión de vapor (evaporador 5)	bar abs
7	P6	Presión de vapor (evaporador 6)	mbar abs
8	T1	Temperatura del jugo (entrada al evaporador 1)	°C
9	T2	Temperatura del jugo (entrada al intercambiador de calor R10)	°C
10	T3	Temperatura del jugo (entrada al intercambiador de calor R13)	°C
11	T4	Temperatura del jugo (entrada al intercambiador de calor R14)	°C
12	T5	Temperatura del jugo (entrada al intercambiador de calor R3)	°C
13	T6	Temperatura del jugo (entrada al intercambiador de calor R3B)	°C
14	T7	Temperatura del jugo (entrada al intercambiador de calor R8)	°C
15	T8	Temperatura del jugo (entrada al intercambiador de calor R9)	°C
16	T9	Temperatura del jugo (salida del tanque de jugo)	°C
17	T10	Temperatura del jugo (salida del intercambiador de calor R10)	°C
18	T11	Temperatura del jugo (salida del intercambiador de calor R11)	°C
19	T12	Temperatura del jugo (salida del intercambiador de calor 12)	°C
20	T13	Temperatura del jugo (salida del intercambiador de calor R3)	°C
21	T14	Temperatura del jugo (salida del intercambiador de calor R3A)	°C
22	T15	Temperatura del jugo (salida del intercambiador de calor R3B)	°C
23	T16	Temperatura del jugo (salida del intercambiador de calor R4)	°C



24	T17	Temperatura del jugo (salida del intercambiador de calor R5)	°C
25	T18	Temperatura del jugo (salida del intercambiador de calor R6)	°C
26	T19	Temperatura del jugo (salida del intercambiador de calor R7)	°C
27	T20	Temperatura del jugo (salida del intercambiador de calor R8)	°C
28	T21	Temperatura del jugo (salida del intercambiador de calor R9)	°C
29	T22	Temperatura de vapor (evaporador 1)	°C
30	T23	Temperatura de vapor (evaporador 2)	°C
31	T24	Temperatura de vapor (evaporador 3)	°C
32	T25	Temperatura de vapor (evaporador 4)	°C
33	T26	Temperatura de vapor (evaporador 5)	°C
34	T27	Temperatura de vapor (evaporador 6)	°C
35	T28	Temperatura de vapor (caldera)	°C
36	W1	Flujo másico del jugo (al intercambiador de calor R10)	T/h
37	W2	Flujo másico del jugo (al intercambiador de calor R3)	T/h
38	W3	Flujo másico del jugo (al intercambiador de calor R3B)	t/h
39	W4	Flujo másico del jugo (que entra al intercambiador de calor R4)	t/h
40	W5	Flujo másico del jugo (que entra al intercambiador de calor R8)	t/h
41	W6	Flujo másico del jugo (que entra al intercambiador de calor R9)	t/h
42	W7	Flujo másico del jugo (dentro del tanque de jugo)	t/h
43	W8	Flujo másico del jugo (que sale del evaporador 6)	t/h
44	W9	Flujo de vapor (que sale de las calderas)	t/h
45	T29	Temperatura del jugo (salida del evaporador 1)	°C
46	T30	Temperatura del jugo (salida del evaporador 2)	°C
47	T31	Temperatura del jugo (salida del evaporador 3)	°C
48	T32	Temperatura del jugo (salida del evaporador 4)	°C
49	T33	Temperatura del jugo (salida del evaporador 5)	°C
50	T34	Temperatura del jugo (salida del evaporador 6)	°C
51	P7	Presión de vapor (desde las calderas)	bar

**Tabla 4.1:** Resumen de las variables tomadas de la planta. La variable de salida será el contenido de materia seca a la salida del evaporador 6

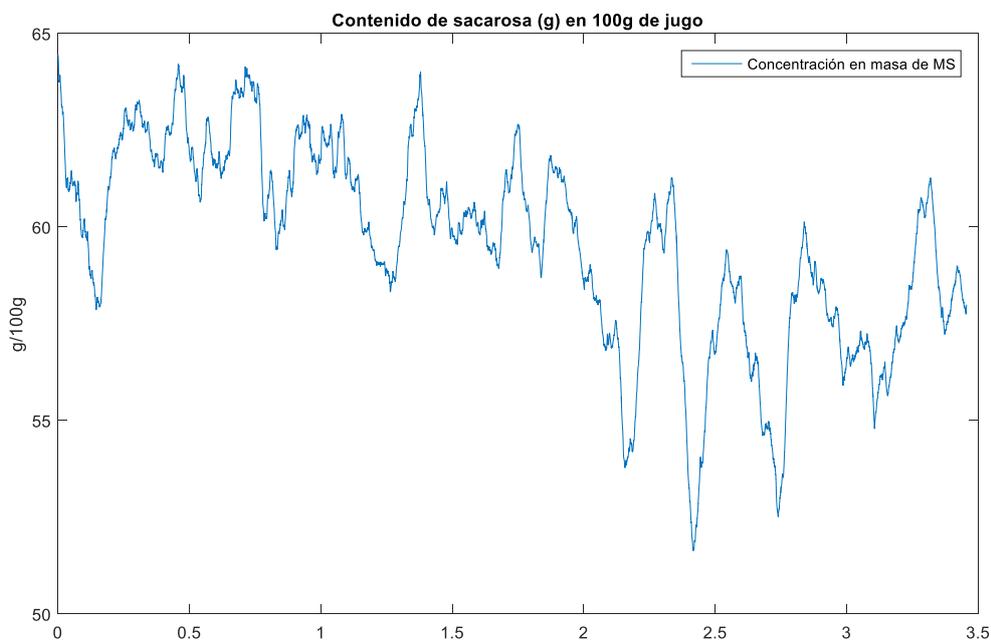
En nuestro caso, la variable que relacionaremos con la calidad del proceso, para poder medir esta última, será CS6, es decir, el contenido en gramos de materia seca existente en el jugo en el



evaporador 6. Es decir, su concentración en masa. Esta variable será almacenada en una matriz que llamaremos “MatrizY”, la matriz predicha (o dependiente).

La matriz de predicción (o matriz independiente), estará compuesta por el resto de variables, que en nuestro caso corresponden a temperaturas, presiones y flujos másicos en las diferentes fases del proceso. La llamaremos “MatrizX”.

En la figura 4.1 se muestran las observaciones recogidas de la variable CS6:



**Figura 4.1:** Concentración en masa de sacarosa en el jugo contenido en el evaporador 6.

## 4.2.2 GRUPOS DE DATOS

Para el desarrollo de los métodos utilizados para el control estadístico del proceso, el rango total de observaciones se ha dividido en dos grupos.



El primero, denominado “Datos de entrenamiento”, será utilizado para que el sistema “aprenda” el comportamiento normal del proceso. El segundo, denominado “Datos de prueba”, serán los datos sobre los cuales probaremos la robustez de cada método, y gracias a ello podremos compararlos entre ellos para determinar las ventajas e inconvenientes de cada uno sobre los demás.

Durante el proceso de entrenamiento con sus datos, es posible que el método se acostumbre a un patrón constante y aprenda “de memoria”, pudiendo verse incapaz de reaccionar ante un cambio inesperado en una prueba. Para evitar esto, se han elegido rangos específicos de datos para cada grupo. Hay que tener en cuenta que esta decisión puede provocar en el futuro falsas alarmas debido a un cambio brusco en medidas aparentemente consecutivas.

En la figura 4.2 se pueden apreciar los rangos de datos seleccionados tanto para entrenar, como para probar cada uno de los métodos que utilizaremos en este capítulo.

### **4.2.3 FALLOS PROVOCADOS EN LOS DATOS**

Para poder comprobar la robustez de los diferentes métodos, se alterará en algún instante de tiempo, alguna de las variables del proceso, de forma que provoque una anomalía controlada.

De esta manera se puede observar si el método ha detectado el fallo, en qué instante de tiempo lo ha detectado, y a su vez, que indique cuál es la variable que está provocando el fallo. De esta manera facilitaremos el trabajo al operario a cargo del proceso, indicando dónde está el error para que pueda solventarlo con rapidez para volver a un funcionamiento normal del sistema.

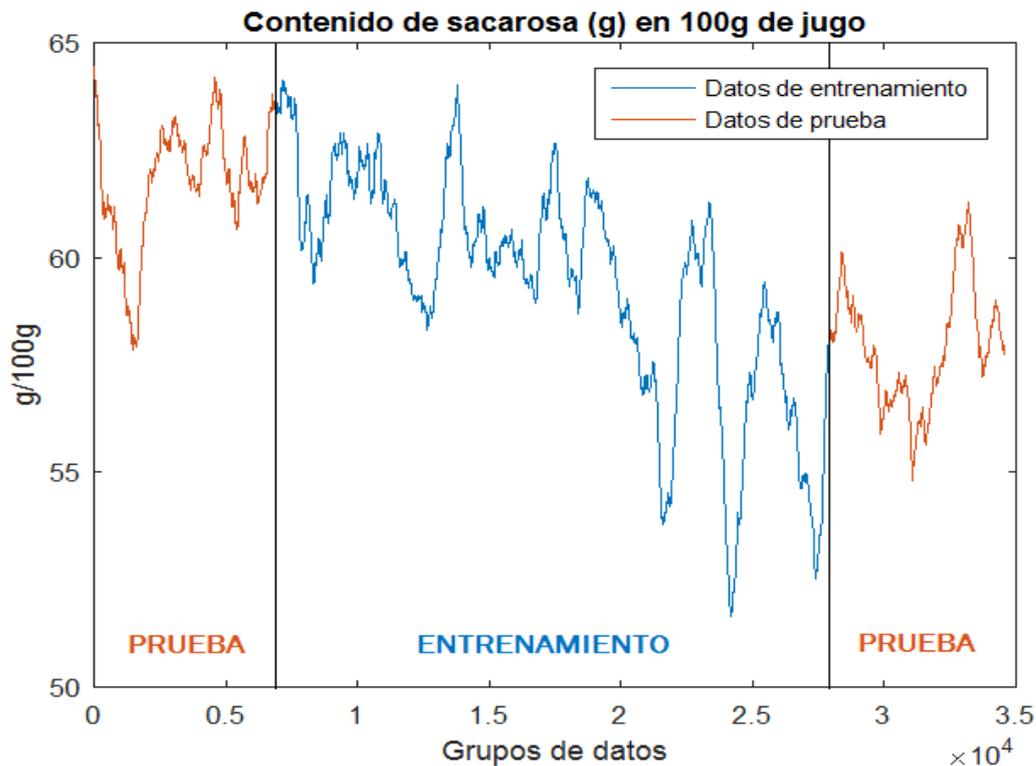


Figura 4.2: Rango de datos de entrenamiento y prueba.

### 4.3 ESTIMACION DE LA VARIABLE Y DETECCIÓN Y DIAGNÓSTICO DE FALLOS CON PLS

En este apartado se explicará cómo se ha adaptado cada uno de los métodos basados en regresión por mínimos cuadrados parciales para su aplicación en el entorno requerido. Además, se reflejarán los resultados obtenidos al ejecutarlos con los datos obtenidos de la planta, y posteriormente se realizará una comparación para determinar sus ventajas e inconvenientes, y facilitar la elección del método más óptimo para aplicarse.



Los métodos derivados de PLS utilizados en este trabajo fueron los siguientes:

- Mínimos Cuadrados Parciales (PLS)
- Mínimos Cuadrados Parciales dinámico (DPLS)
- Mínimos Cuadrados Parciales no lineal (NPLS)
- Mínimos Cuadrados Parciales con extracción de características (PLSEX)
- Otras combinaciones de los anteriores

### **4.3.1 MÍNIMOS CUADRADOS PARCIALES**

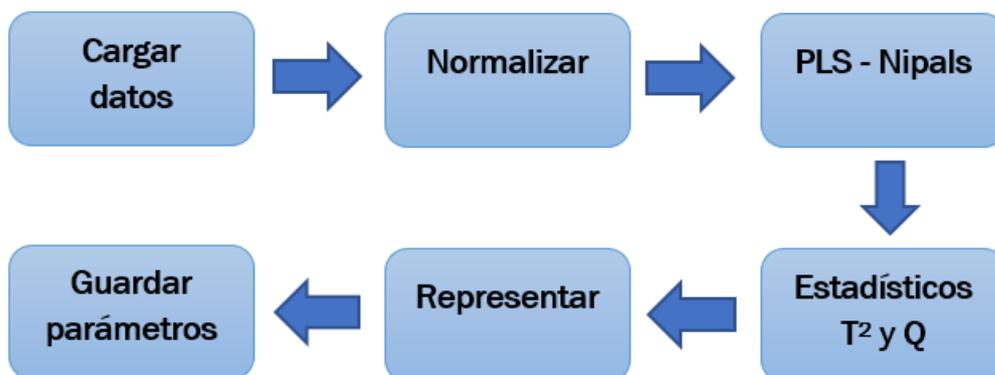
La regresión por mínimos cuadrados parciales es el método de detección y diagnóstico de fallos basado en PLS más genérico de los que se han programado. Los siguientes métodos se han basado en la programación del primero.

En los diagramas de flujo representados en las figuras 4.3 (a) y 4.3 (b) se muestra la estructura que se ha seguido a la hora de implementar PLS mediante Matlab. Los demás métodos seguirán la misma estructura, variando únicamente el conjunto de datos que componen la matriz de predicción. Los pasos que se han seguido en el **proceso de entrenamiento** son los siguientes:

1. Se carga el fichero con todos los datos recogidos de la planta, y a partir de ellos se crea una primera matriz, que será la de **entrenamiento**, con los datos seleccionados para tal fin.
2. Para poder trabajar con los datos, se **normalizan** con media cero y varianza unitaria, de esta forma se consigue que todas las variables tengan la misma importancia en el método independientemente de su valor.



3. Con ayuda del criterio de información de Akaike (**AIC**), se elegirá el número óptimo de variables con el que se va a trabajar.
4. Introduciremos los datos de entrenamiento y el número de variables seleccionadas gracias al AIC en una modificación del **algoritmo de NIPALS** para calcular los parámetros que utiliza el método PLS.
5. Con esos parámetros, el programa nos pedirá que introduzcamos el nivel de significancia deseado ( $\alpha$ ), y procederá a calcular tanto los **residuos** como los umbrales para las estadísticas **T<sup>2</sup>** y **Q**.
6. Los resultados obtenidos se representarán mediante gráficas para facilitar su análisis.
7. Se guardan los parámetros y umbrales que son necesarios para la aplicación sobre los datos de prueba. Estos son; las configuraciones de normalización de las matrices de entrenamiento de X e Y, los parámetros calculados por el algoritmo de NIPALS ( $\beta$ , P, R y  $\delta$ ), y los umbrales de T<sup>2</sup> y Q.



**Figura 4.3 (a):** Diagrama de flujo del proceso de entrenamiento de datos.



Tras haber entrenado el método, llega el turno de probarlo con un conjunto nuevo de datos. De esta forma, podrá comprobarse si el método tan solo es capaz de estimar el valor de la variable de calidad y de detectar fallos en datos similares a los que ha aprendido, o si se desenvuelve con eficacia en situaciones diferentes a las que ha entrenado. Los pasos seguidos en el **proceso de prueba** de datos son los siguientes:

1. De la misma forma que en el entrenamiento, se carga el fichero con los datos iniciales, pero esta vez cargamos los parámetros guardados durante el entrenamiento. Se crea la **matriz de prueba**, con los datos seleccionados para ello.
2. Se **normalizan** los datos de prueba con la misma configuración utilizada en la transformación de los datos de entrenamiento.
3. Calcularemos tanto la nueva estimación de la variable de calidad, como los nuevos **residuos** y las nuevas estadísticas  $T^2$  y  $Q$ .
4. **Detección de fallos:** si alguno de los métodos, tanto  $T^2$  como  $Q$ , superan su umbral se detecta una situación anómala y si lo hace durante un periodo determinado de observaciones consecutivas, detectará e indicará un fallo, y el instante de tiempo en el que ha sido detectado.
5. **Diagnóstico de fallos:** Si el sistema ha detectado algún fallo, mostrará qué variables son las que aportan una mayor contribución a dicho fallo.



6. Los resultados obtenidos se representarán mediante gráficas para facilitar su posterior análisis.

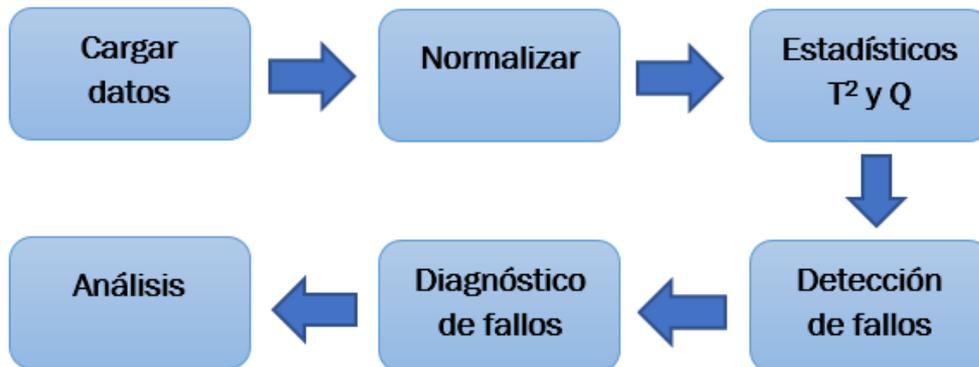


Figura 4.3 (b): Diagrama de flujo del proceso de prueba de datos.

## 4.3.2 APLICACIÓN PLS GENÉRICO

En este apartado comenzaremos el análisis del método PLS genérico. Primero, utilizaremos los datos de entrenamiento para que el método “aprenda”, es decir, se familiarice con los datos analizados.

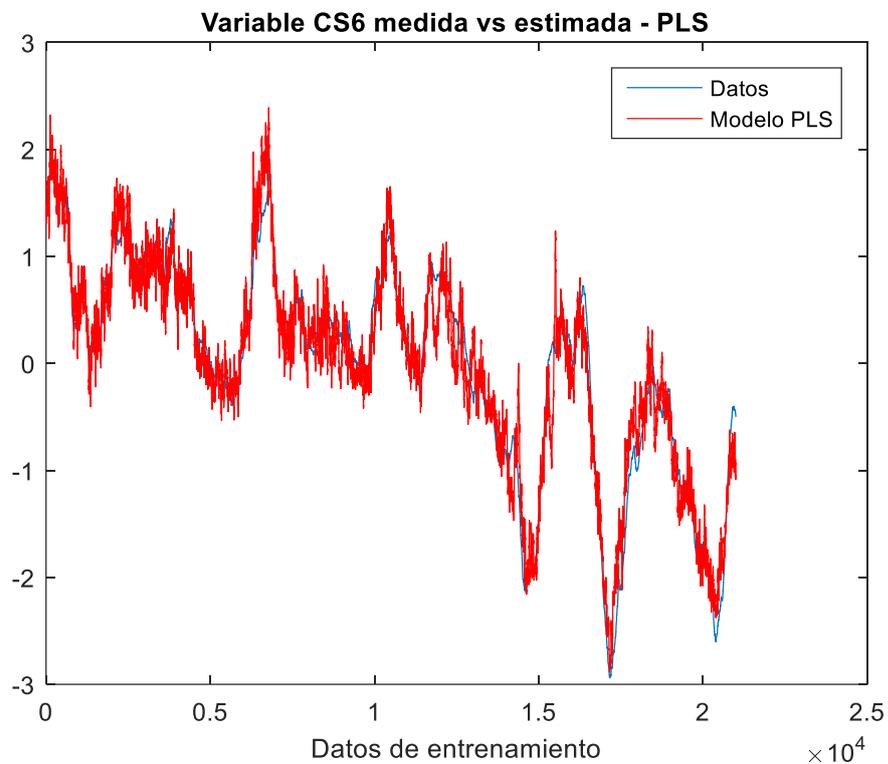
### 4.3.2.1 ENTRENAMIENTO CON PLS

Ejecutaremos el programa PLS con los datos de entrenamiento. Al hacerlo, este pedirá al usuario por pantalla que introduzca el nivel de significación deseado para el análisis. En este caso, he introducido un  $\alpha=0,01$ . He considerado 0,01 como un valor suficientemente pequeño para que el método sea confiable, y correr menor riesgo de error en el proceso.



En este caso, el programa va a trabajar con 21000 datos, y 50 variables en la matriz X. Como hemos explicado anteriormente, la regresión por Mínimos Cuadrados Parciales es un método que reduce la dimensionalidad de los datos. Para determinar el número óptimo de variables con las que vamos a trabajar, he utilizado el criterio de información Akaike. El número elegido ha sido 12, así pues, el algoritmo de NIPLAS modificado va a trabajar con la matriz de entrenamiento X, que será nuestra matriz de predicción ( $21000 \times 50$ ), la matriz de entrenamiento Y, que será la matriz predicha ( $21000 \times 1$ ) y con 12 variables en total.

En la figura 4.4 podemos ver y comparar la variable CS6 (la variable predicha, la concentración en masa final de sacarosa). En azul, representada según los datos recogidos, y en rojo, cómo ha aprendido el modelo PLS, es decir, la variable Y frente a la variable  $\hat{Y}$ .

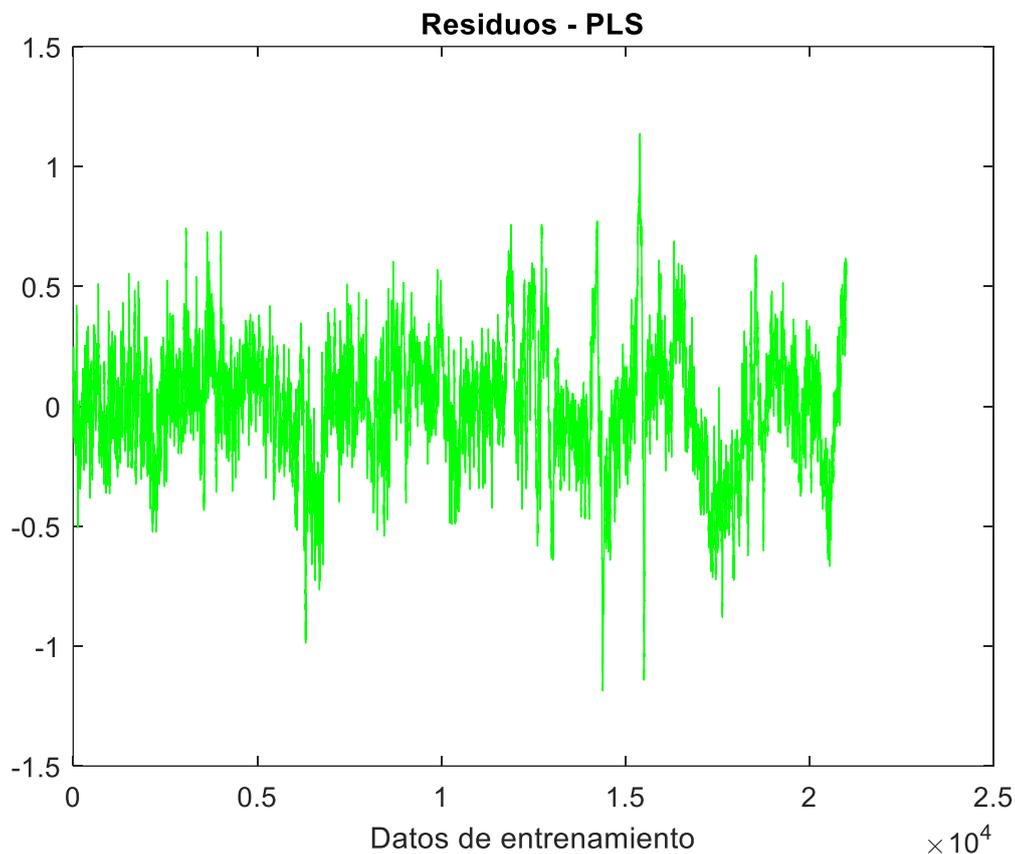


**Figura 4.4:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo PLS (entrenamiento).



Podemos ver que la aproximación es bastante buena. El modelo presenta un poco de ruido en forma de un número moderado de picos, pero en general podemos concluir que ha entrenado bastante bien, teniendo en cuenta que ha reducido el número de variables con las que se trabaja de 50 a 12.

Podemos reforzar esta conclusión fijándonos en los residuos, representados en la figura 4.5. Vemos que se mueven en torno a errores de  $\pm 0.5$ , llegando en alguna observación a  $\pm 1$ , aunque no parece preocupante.

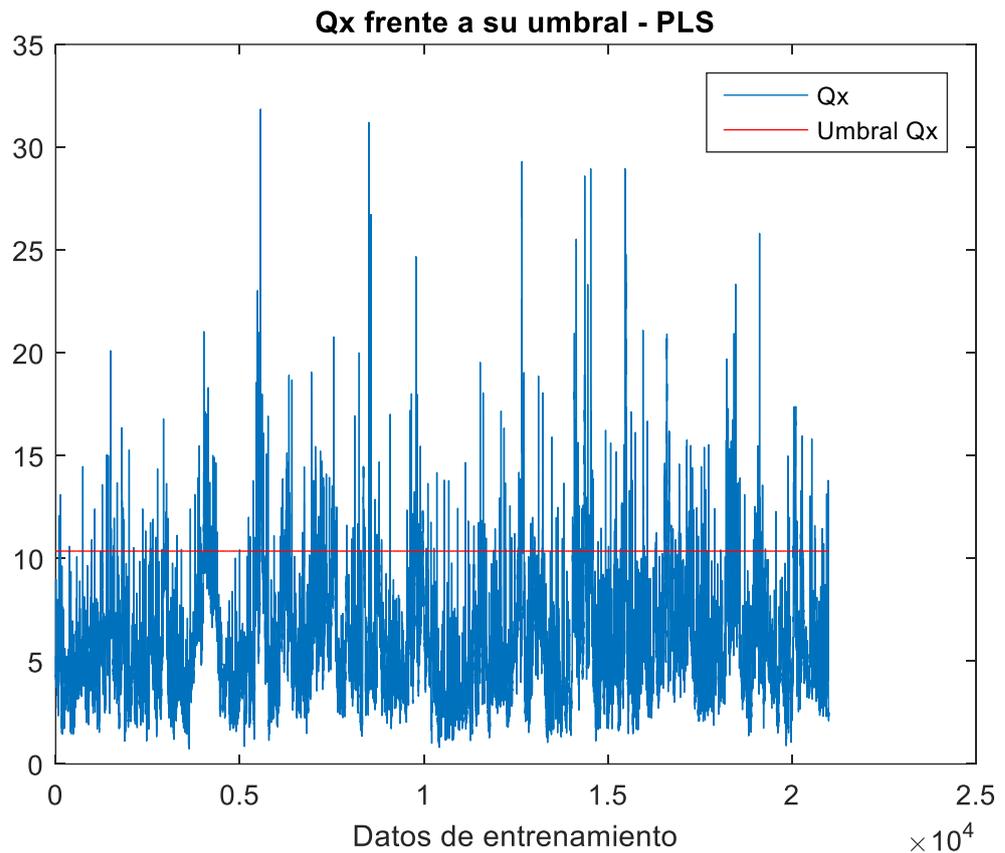


**Figura 4.5:** Residuos obtenidos al aplicar el método PLS (entrenamiento).



El siguiente paso es el cálculo de las estadísticas  $Q$  y  $T^2$ . De ellos, sacaremos los umbrales que nos permitirán detectar y diagnosticar fallos, en caso de haberlos, a la hora de probar el método con los nuevos datos.

Primero calcularemos la estadística  $Q$ , en la cual tendremos la relacionada a la matriz de predicción,  $Q_x$ , y la relacionada a la matriz predicha,  $Q_y$ . Esta estadística se usa para monitorizar el espacio de los residuos. En la figura 4.6 se representa el estadístico  $Q_x$  para cada una de las observaciones. También representaremos el umbral, previamente calculado, para poder observar mejor si tenemos muchos valores que lo superen.



**Figura 4.6:** Estadístico  $Q_x$  frente a su umbral  $UmbralQ_x$  (entrenamiento).



Como podemos observar, la mayor parte de los datos se encuentra bajo el umbral. Aun así, debido a la existencia de perturbaciones que pudieran dar lugar a falsas alarmas, he decidido aumentar ligeramente este umbral de  $Q_x$  para la prueba y los siguientes métodos. Aunque el representado en la figura 4.6 es el umbral teórico, este puede verse ajustado si fuera necesario hasta un punto en el que el método no arroje falsas alarmas, pero que permita detectar un fallo real, de haberlo.

Ahora pasamos al cálculo del estadístico  $Q_y$ , el cual está representado en la figura 4.7. En este caso, podemos ver cómo ninguno de los valores del estadístico supera el umbral. Esto es lógico, ya que la variable  $Y$  es la variable que pretendemos calcular con este método, por lo que sería contradictorio que esta presentase algún error. El umbral de este estadístico no lo modificaremos, será el que utilicemos con los datos de prueba.

Por último, nos queda calcular el estadístico  $T^2$ , o estadística de Hotelling, que se utiliza en el espacio de una dimensión “ $a$ ” determinada, para detectar comportamientos anómalos del sistema cuando traspasan un umbral (Figura 4.8). Como en el caso del estadístico  $Q_x$ , se ha decidido ajustar el umbral teórico para evitar falsas alarmas a la hora de realizar la prueba del método con los datos correspondientes.

En el instante en que se detecta un fallo en la planta, se calcula la contribución de cada variable a que las estadísticas  $Q$  y  $T^2$  hayan superado su umbral, es decir, la contribución de cada variable del proceso al fallo, la mayor contribución nos dirá que variable o variables son las responsables del fallo, es decir, diagnosticará el fallo.

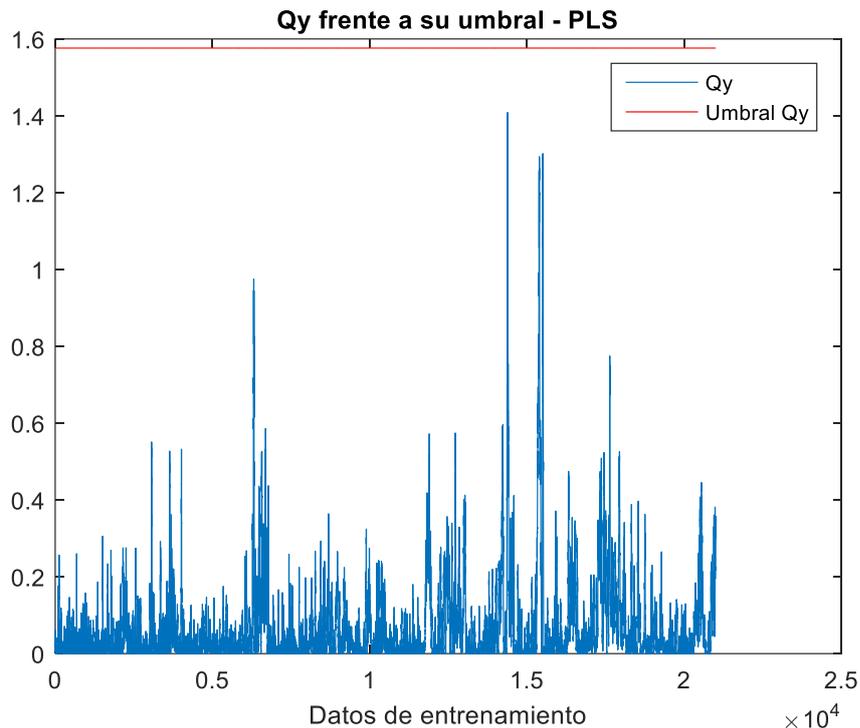


Figura 4.7: Estadístico  $Q_y$  frente a su umbral  $UmbralQ_y$  (entrenamiento).

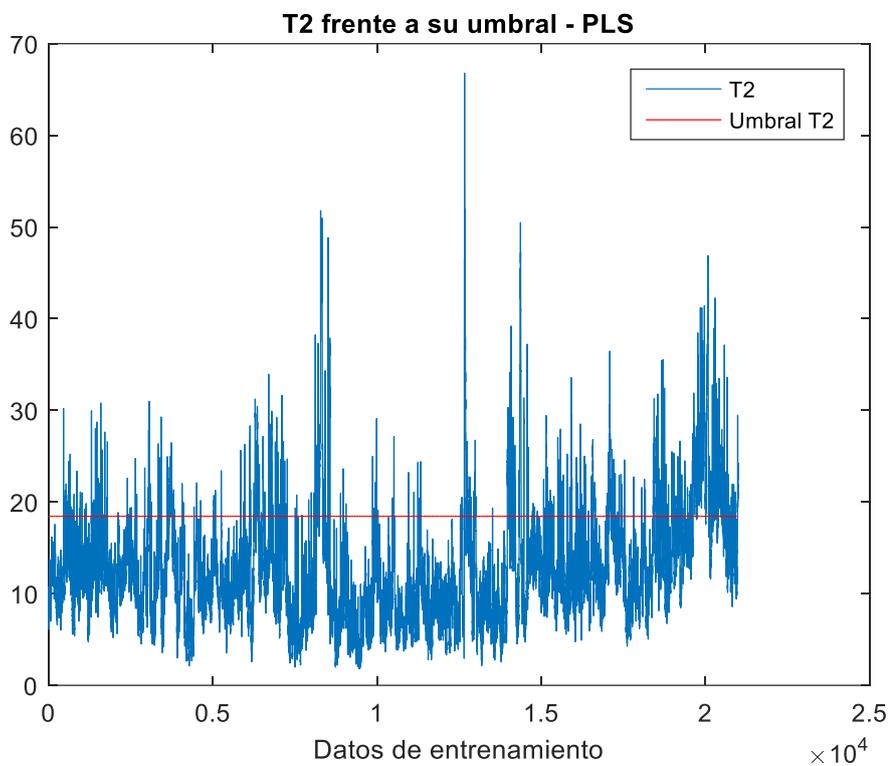
### 4.3.2.2 PRUEBA SIN FALLO CON PLS

Aunque los resultados obtenidos al entrenar los datos parecen convincentes, no podremos comprobar la robustez del método sin probarlo con unos nuevos datos, utilizando la misma configuración que hemos guardado tras el entrenamiento.

De nuevo, representamos la variable medida (figura 4.9) en contra de la variable estimada por el modelo. El salto brusco en torno a la observación número 7000 aproximadamente, es debido a que, como hemos comentado en el apartado 4.2.2, el conjunto de datos utilizados para la prueba consta de dos rangos de datos no consecutivos.



Esto podría suponer una situación complicada para el método, en el que vemos presencia de picos, de ruido, y, aunque la variable modelada no sigue exactamente el patrón de los datos (siempre tendremos un error, por pequeño que sea), la aproximación es bastante buena. Podemos comprobarlo también en la gráfica de residuos, en la figura 4.10, donde podemos ver, los residuos no han variado en exceso, quizá se muevan más en el rango de  $\pm 1$ , pero no es preocupante.

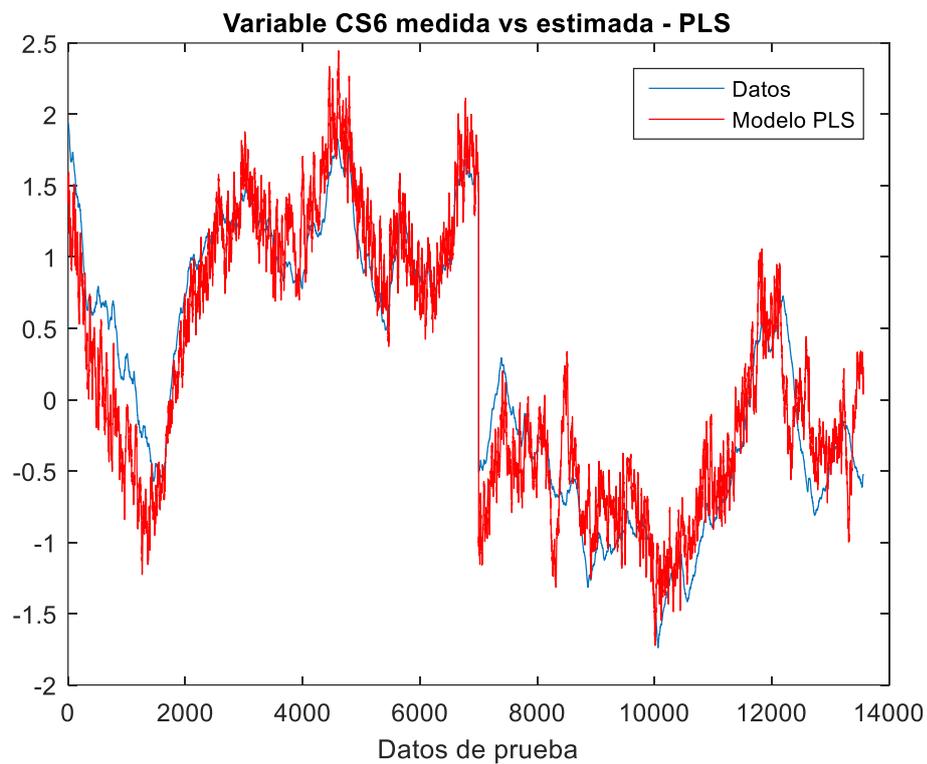


**Figura 4.8:** Estadístico  $T^2$  frente a su umbral  $UmbralT^2$  (entrenamiento).

Pasamos ahora al cálculo de los nuevos estadísticos  $Q$ . En la figura 4.11 se representa el nuevo estadístico  $Q_x$  para los datos de prueba. El umbral guardado en el entrenamiento ha sido ajustado ligeramente. Como puede observarse, en torno a la



observación número 10000, tenemos un punto que se va muy por encima del umbral, cuando realmente no existe ninguna anomalía en el proceso. Con el nuevo umbral, el método evita detectar fallos en este punto y otros como este, manteniendo su eficacia cuando se trate de un error real, como veremos en el siguiente apartado de prueba con error.



**Figura 4.9:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo PLS (prueba).

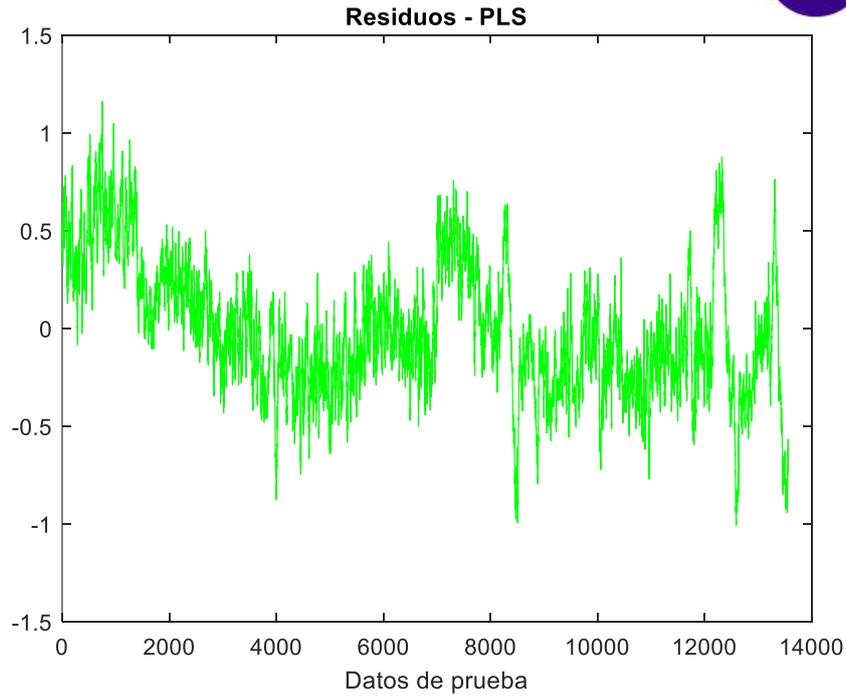


Figura 4.10: Residuos obtenidos al aplicar el método PLS (prueba).

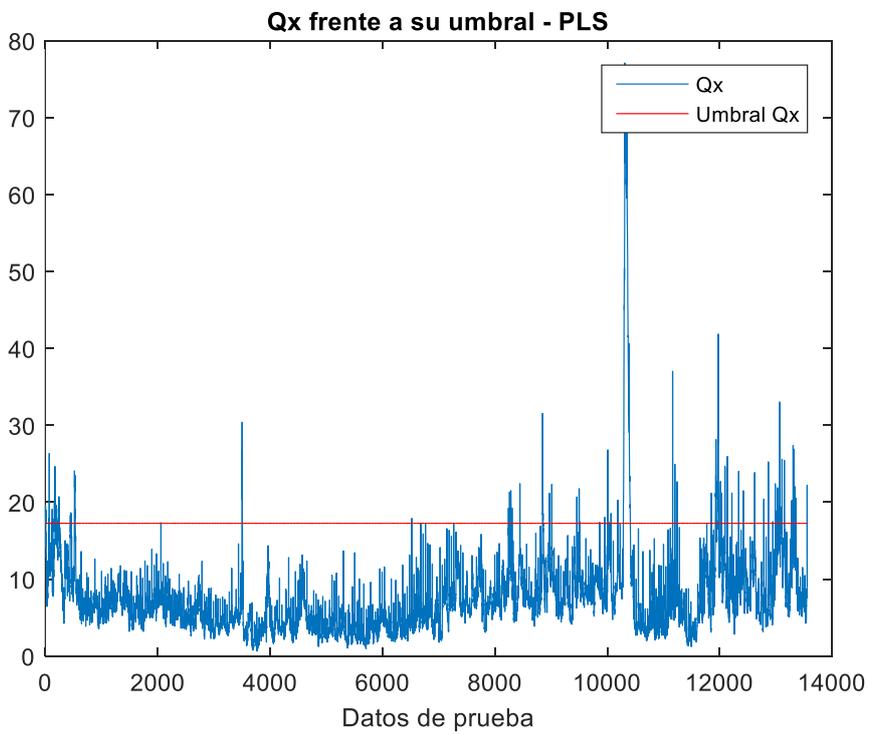
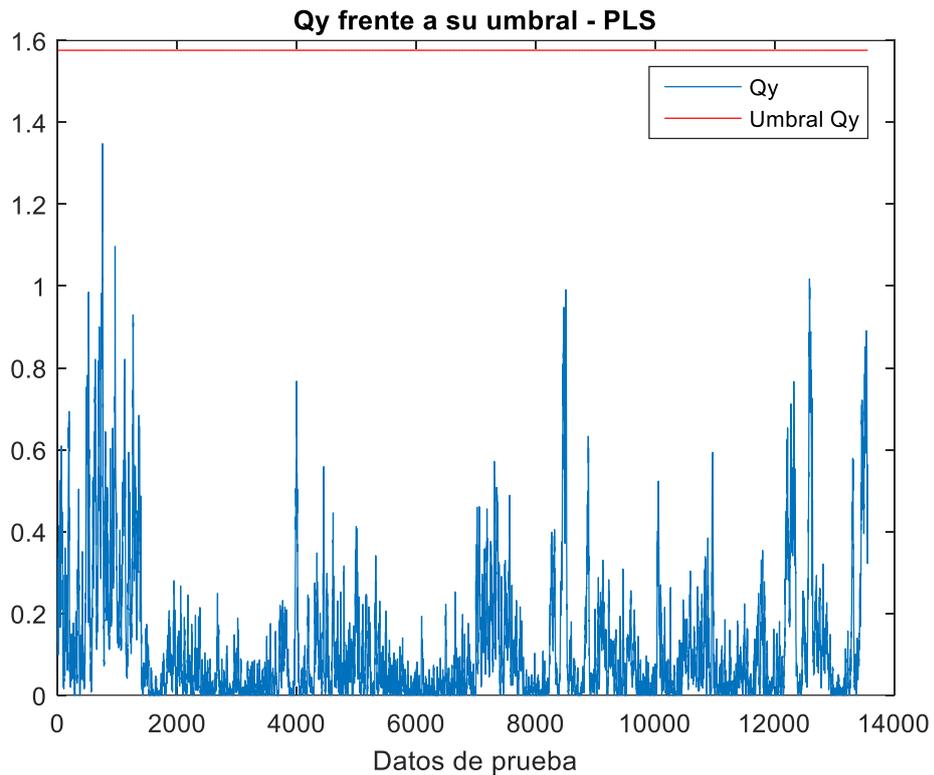


Figura 4.11: Estadístico  $Q_x$  frente a su umbral  $UmbralQ_x$  (prueba).

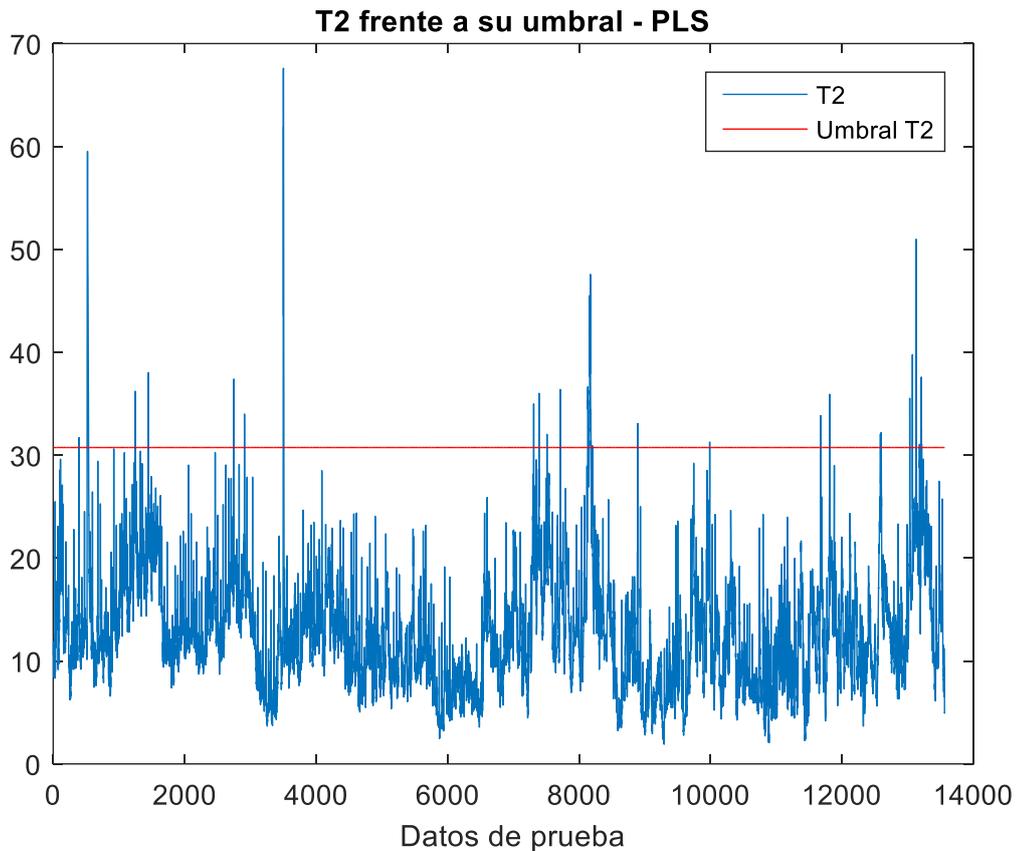


En la figura 4.12 está representado el estadístico  $Q_y$ , en este caso con el mismo umbral calculado al entrenar el método. Como hemos comentado, este estadístico no presentará ningún error en situaciones normales, ya que la variable  $Y$  es la variable de salida, que obtenemos a partir de las variables de la matriz  $X$  de predicción.



**Figura 4.12** Estadístico  $Q_y$  frente a su umbral  $UmbralQ_y$  (prueba).

En el estadístico  $T^2$ , se ha seguido el mismo procedimiento que en el del estadístico  $Q_x$ , ajustando ligeramente su umbral para evitar falsas alarmas en el sistema. Puede apreciarse en la figura 4.13 cómo la amplia mayoría de los valores está dentro del límite especificado, y, como era de esperar, el sistema de detección de fallos no ha sido activado.



**Figura 4.13:** Estadístico  $T^2$  frente a su umbral  $UmbralT^2$  (prueba).

### **4.3.2.3 PRUEBA CON FALLO CON PLS**

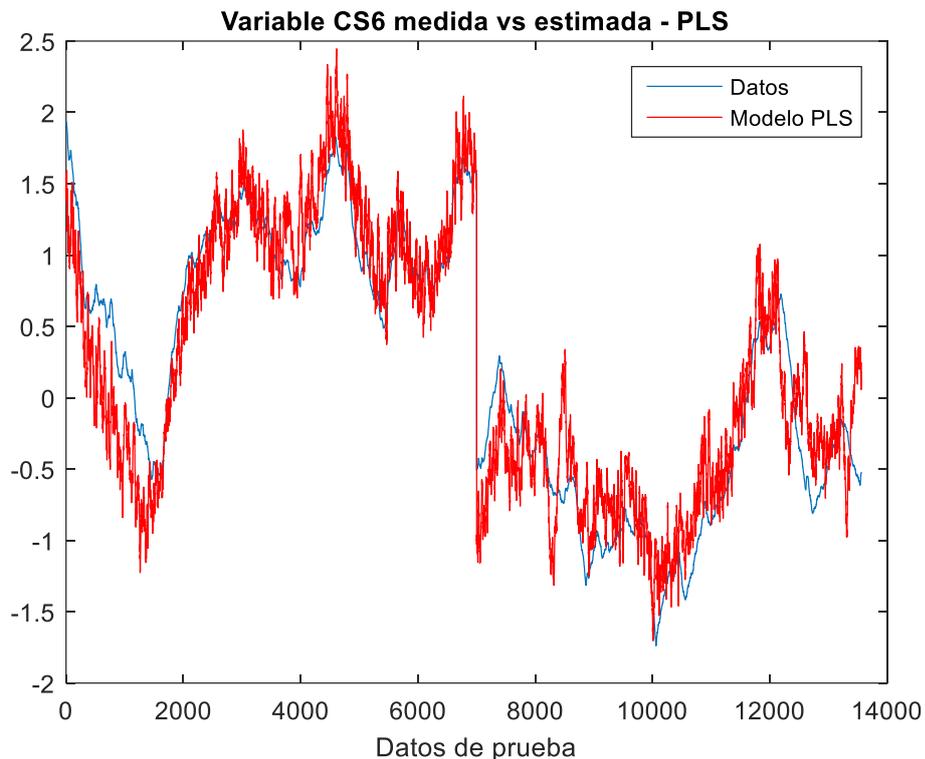
Pasaremos ahora a comprobar la robustez del método, es decir, si es capaz de detectar un fallo provocado en alguna de sus variables. También podremos comprobar si lo detecta en el instante esperado, y si el sistema de diagnóstico de fallos aporta información correcta acerca de cuál es la variable que está provocando el error. De esta manera, se facilitará el trabajo al operario a cargo del proceso encargado de solucionar el error manualmente para que el sistema vuelva a un funcionamiento normal.



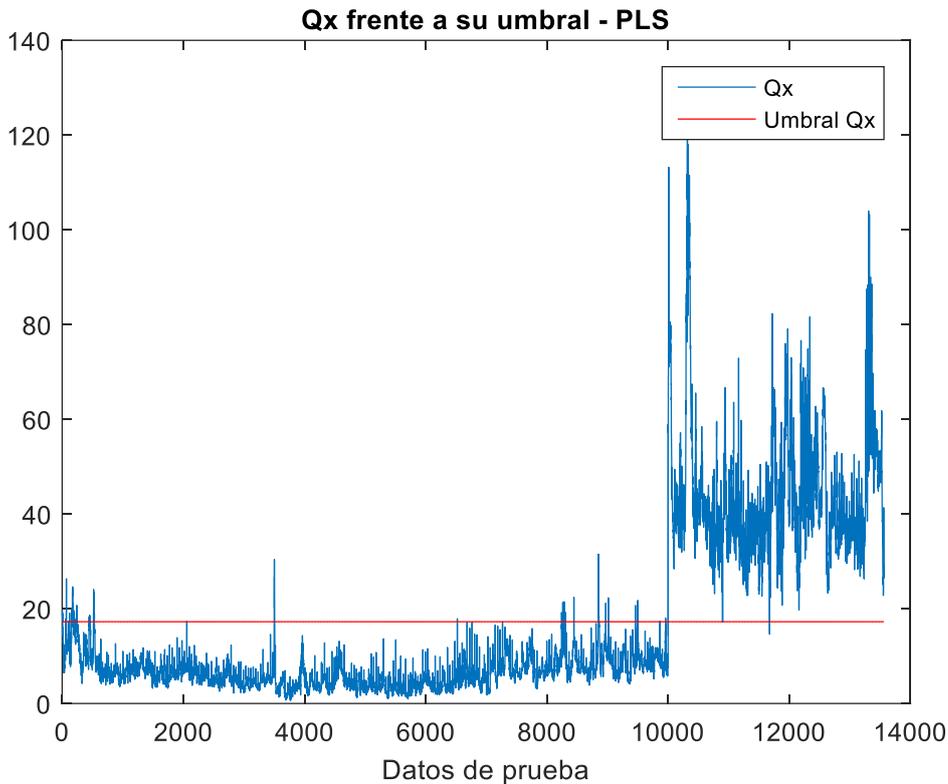
Tras introducir errores en diferentes variables elegidas aleatoriamente, y comprobar que el método las detecta con eficacia, vamos a ver representado uno que sirva como ejemplo.

En este caso, se ha introducido un error en el evaporador 6, en concreto en la variable P6 (ver tabla 4.1), que mide la presión de vapor en ese instante del proceso. El error se ha introducido en la observación número 10000, del total de 13560 que componen la matriz de datos de prueba.

En la figura 4.14 se representa la comparación de la variable de calidad CS6 tomada de la planta frente a la misma variable estimada por PLS, ahora con fallo. Vemos que la prueba apenas sufre cambios al introducir el error, por lo que puede deducirse que se adapta de forma correcta a nuevas situaciones.



**Figura 4.14:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo PLS (prueba con fallo).



**Figura 4.15:** Estadístico  $Q_x$  frente a su umbral  $UmbralQ_x$  (prueba con error).

Podemos apreciar en la figura 4.15 el cambio brusco que sufre el estadístico  $Q_x$  en el instante en el que el error es introducido. Claramente, los valores superan el umbral, y el sistema de detección de fallos se activa, indicando el fallo en el instante 10149. Al activarse el sistema de detección de fallos de  $Q_x$ , a su vez, el sistema de diagnóstico de fallos recibe esta información, y busca cuál es la variable que está provocando el fallo.

En la figura 4.16 puede observarse los resultados arrojados por este sistema de diagnóstico de fallos, donde podemos ver claramente que la variable número 6 predomina sobre las demás, es decir, el sistema indica correctamente dónde está el problema.

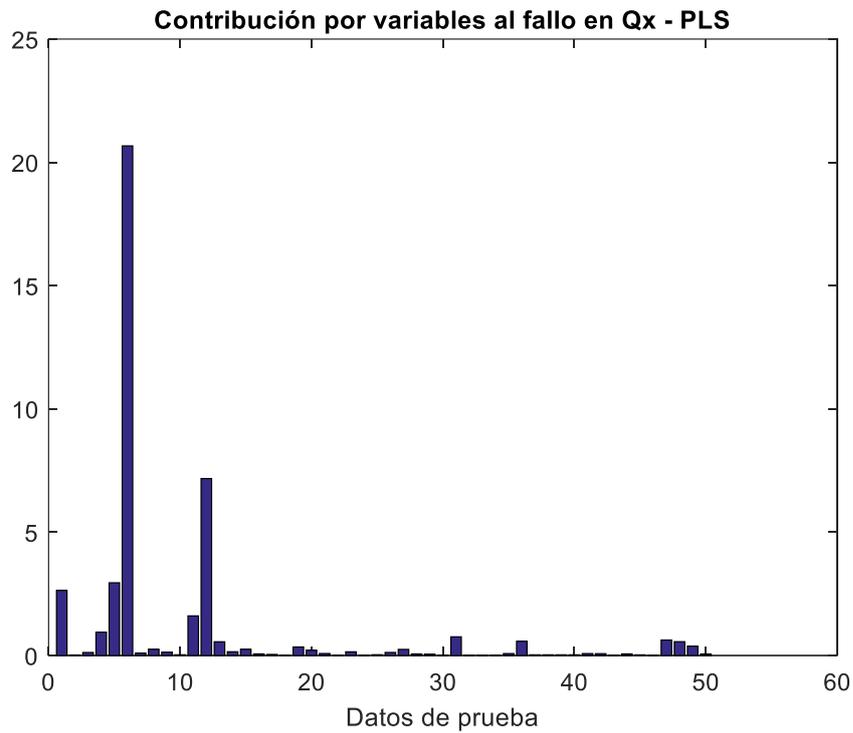


Figura 4.16: Diagrama de contribución al fallo por variables según el sistema de diagnóstico de  $Q_x$  (prueba con error).

Como era de esperar, el estadístico  $Q_y$  no arroja ningún fallo, por lo que hemos comentado anteriormente, no tendría sentido. Esto puede verse representado en la figura 4.17. De aquí en adelante nos centraremos en los estadísticos  $Q_x$  y  $T^2$  a la hora de analizar el procedimiento de detección y diagnóstico de fallos.

Pasaremos ahora a comprobar si el método ha detectado el fallo gracias al estadístico  $T^2$ . En la figura 4.18 podemos observar que esta vez no parece tan claro el instante en el que un amplio número de observaciones supera el umbral, tal como veíamos claramente con el estadístico  $Q_x$ .

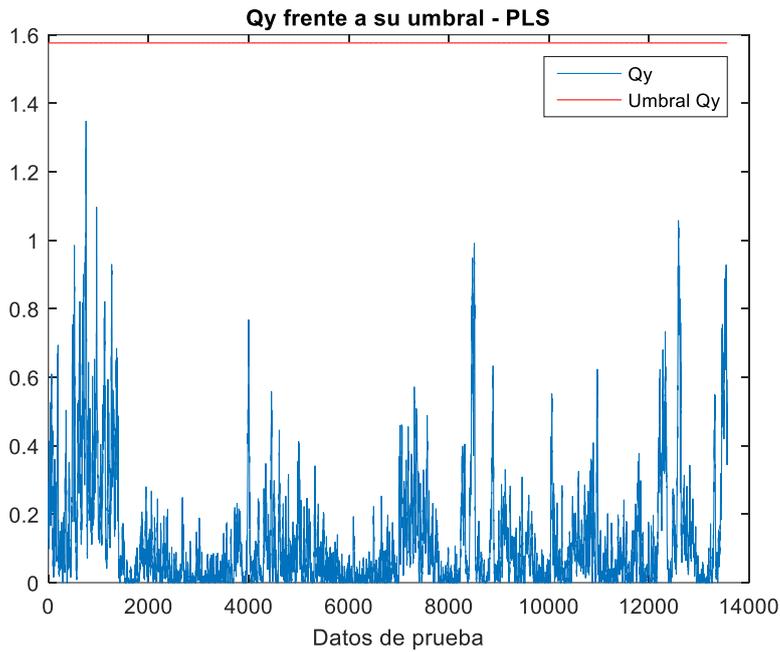


Figura 4.17 Estadístico  $Q_y$  frente a su umbral  $UmbralQ_y$  (prueba con error).

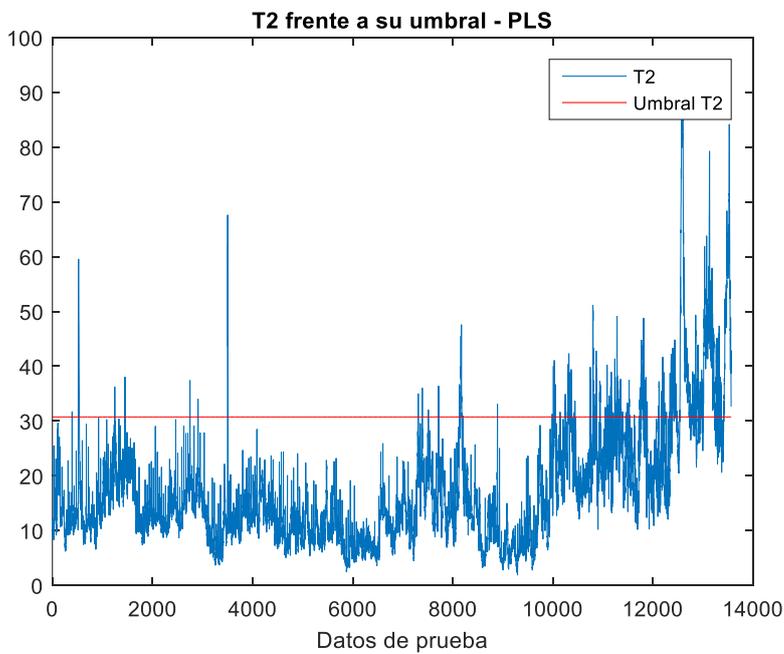
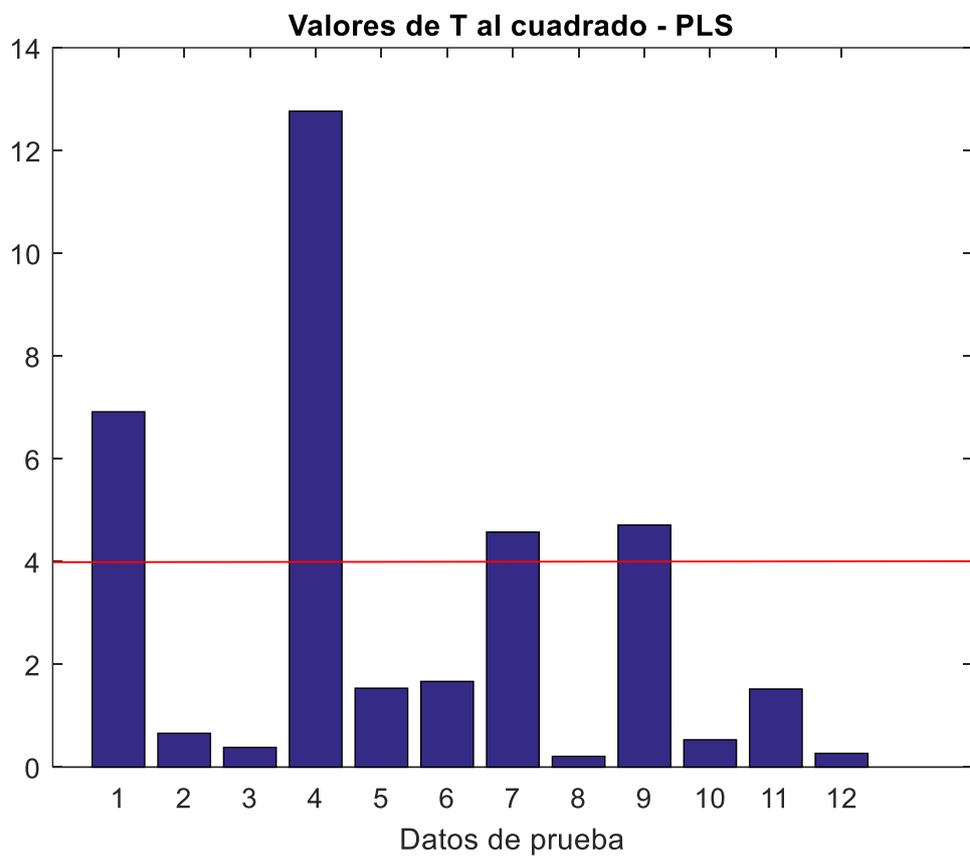


Figura 4.18: Estadístico  $T^2$  frente a su umbral  $UmbralT^2$  (prueba con error).



A priori puede parecer que el método va a tener más problemas para detectar el fallo en este caso. Sin embargo, como podremos apreciar en la figura 4.19, vemos que son varios valores de  $T$  (componentes principales) los que superan el umbral que se ha decidido conveniente. Esto indica que se ha detectado un error provocado por una anomalía en una variable que predomina sobre las demás. En este caso, Matlab nos indica que el error ha sido detectado en el instante 12684, notablemente más tarde de lo esperado, y claramente ha tardado más en detectarlo que el estadístico  $Q_x$  en este caso.

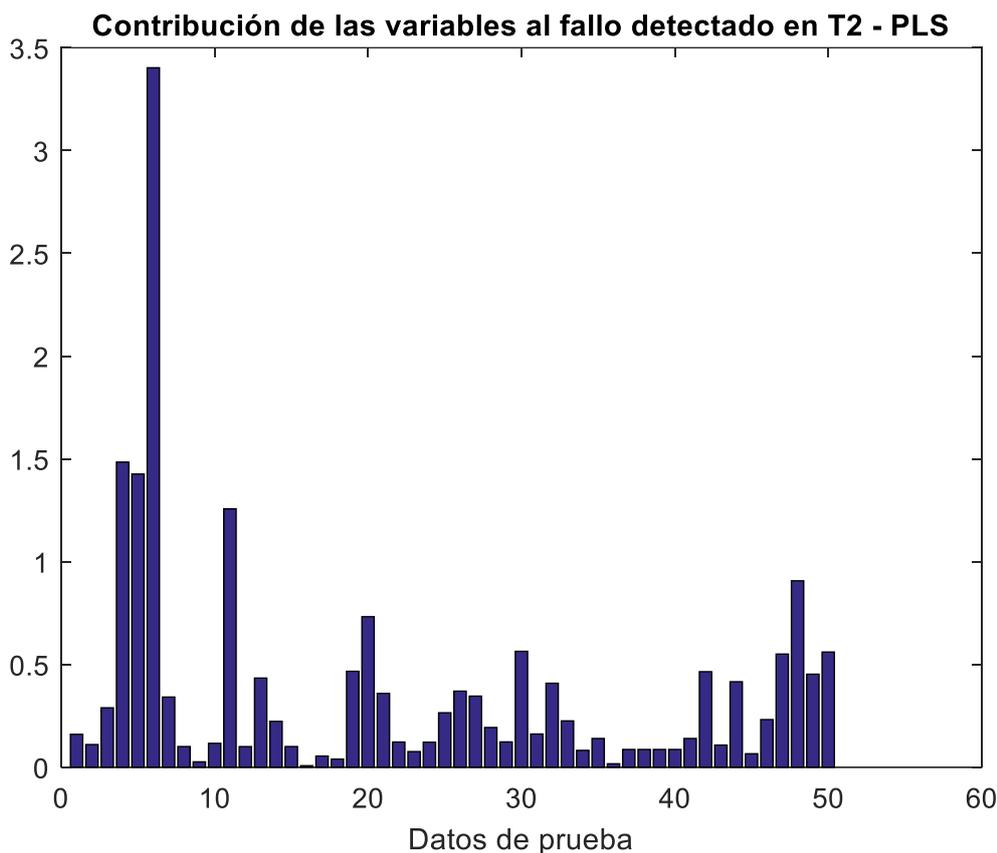


**Figura 4.19:** *Valores de  $T$  que superan el umbral que da lugar a un fallo en el sistema (prueba con error).*



Finalmente, gracias al diagrama de contribución de la figura 4.20, podremos ver si, en este caso, el estadístico  $T^2$  es capaz de determinar cuál es la variable que está provocando el fallo. Y con los resultados, vemos que la variable 6, la variable en la que hemos introducido el error, es la que predomina sobre las demás con claridad, por lo tanto, el diagnóstico ha sido correcto.

Queda entonces concluido el apartado del método genérico PLS. Los resultados obtenidos serán comparados más adelante con los que obtengamos con los métodos derivados de este, que desarrollaremos en los siguientes apartados, así como con las Redes Neuronales Artificiales.



**Figura 4.20:** *Diagrama de contribución al fallo por variables según el sistema de diagnóstico de  $T^2$  (prueba con error).*



### 4.3.3 APLICACIÓN DE PLS CON EXTRACCIÓN DE CARACTERÍSTICAS

En este apartado estudiaremos la modificación del método PLS genérico basado en la reducción de dimensionalidad mediante extracción de características. En un principio, esta extracción de características se ha realizado a través de la función *plsregress* de Matlab, la cual utiliza el algoritmo SIMPLS (válido solamente cuando tenemos una variable dependiente, como es nuestro caso). El procedimiento se basa en centrar X e Y restando sus respectivas medias por columnas, obteniendo así las variables centradas  $X_0$  e  $Y_0$  y calculando un PLS con todos los factores, es decir, sin reducir la dimensionalidad. Esto nos da una matriz P de carga del sistema, donde en cada columna está la relación de cada variable con los componentes PLS calculados. Se calcula el tanto por ciento de influencia de cada variable sobre el modelo PLS, y se eligen sólo las variables con mayor influencia (se extraen características) para calcular el PLS clásico.

#### 4.3.3.1 ENTRENAMIENTO PLS CON EXTRACCIÓN DE CARACTERÍSTICAS

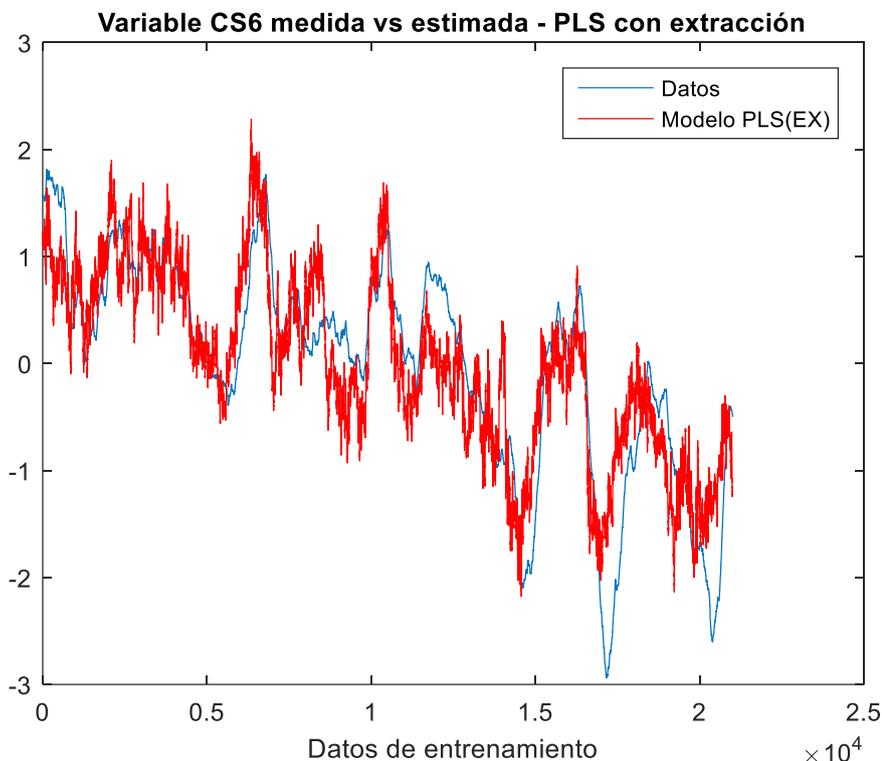
Primero realizaremos el proceso de entrenamiento con los datos correspondientes para ver cómo se adapta el método a trabajar con un número más reducido de variables.

En la figura 4.21 podemos comparar nuestra variable de calidad tomada de los datos, con la variable modelada por PLS(EX). Vemos que, en este caso, el modelo no es demasiado bueno. Hay demasiados picos y el modelo no sigue correctamente el patrón seguido por los datos reales.



Al ver que el modelo no es suficientemente bueno, se ha decidido modificar el procedimiento de extracción de características. Vamos a dejar la función *plsregress* a un lado, y a probar a extraer las características consideradas más significativas para el modelado de nuestra variable de calidad, basándonos en la correlación existente entre cada variable de la matriz X con la variable de salida.

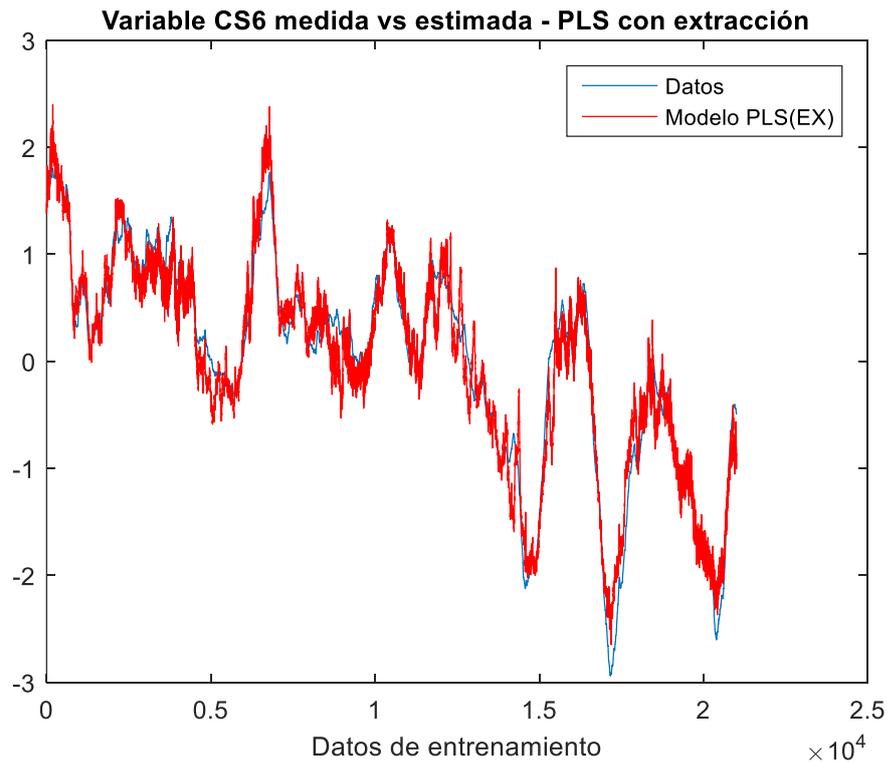
Para ello, calcularemos el coeficiente de correlación entre ambas matrices para cada una de las variables. Tras obtenerla, seleccionaremos las variables que superen un umbral determinado, que serán las de mayor coeficiente de correlación, y aplicaremos el método NIPALS para calcular PLS únicamente a las variables extraídas, es decir, aquellas variables más correlacionadas con la salida.



**Figura 4.21:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo PLS(EX) (entrenamiento).



En la figura 4.22 podemos ver la mejoría que ha experimentado el modelo al extraer las variables de esta nueva forma, por tanto, de aquí en adelante utilizaremos este método cuando queramos extraer características con otros métodos.



**Figura 4.22:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo PLS(EX) corregido (entrenamiento).

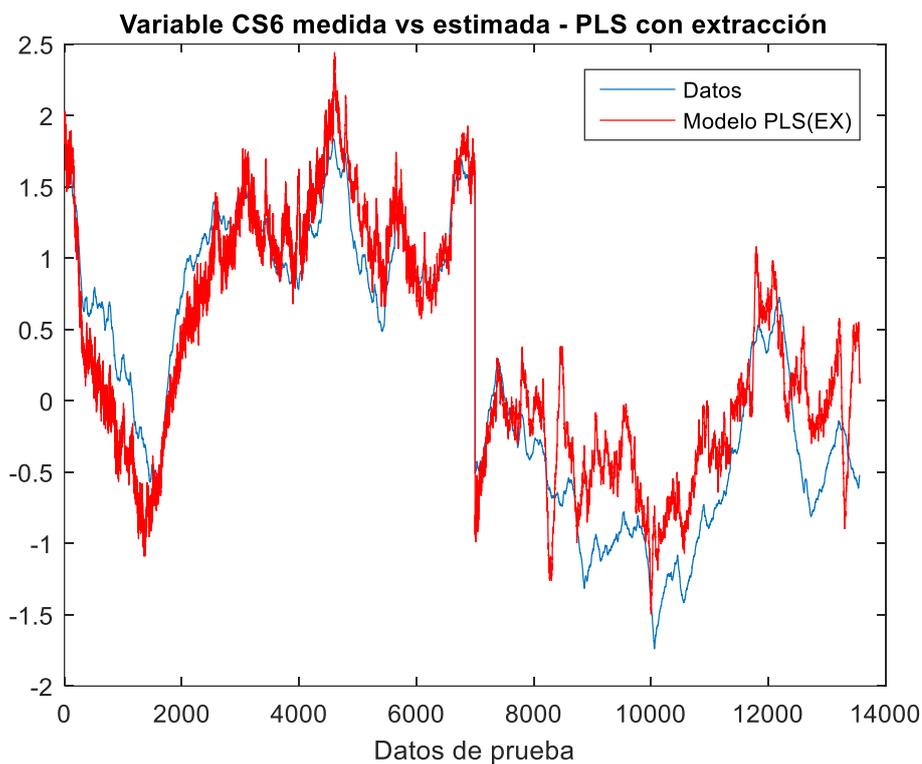
### **4.3.3.2 PRUEBA SIN FALLO CON PLS CON EXTRACCIÓN DE CARACTERÍSTICAS**

En la figura 4.23 podemos ver los resultados de la estimación de nuestra variable de calidad utilizando los datos de prueba con las características extraídas por correlación.



En ella, vemos que, aunque el resultado parece no ser muy convincente, es mucho mejor que el resultado obtenido al extraer mediante *plsregress*. Hay que tener en cuenta que la extracción de características se ha quedado con 16 variables consideradas relevantes, de las 50 que componen la matriz de predicción.

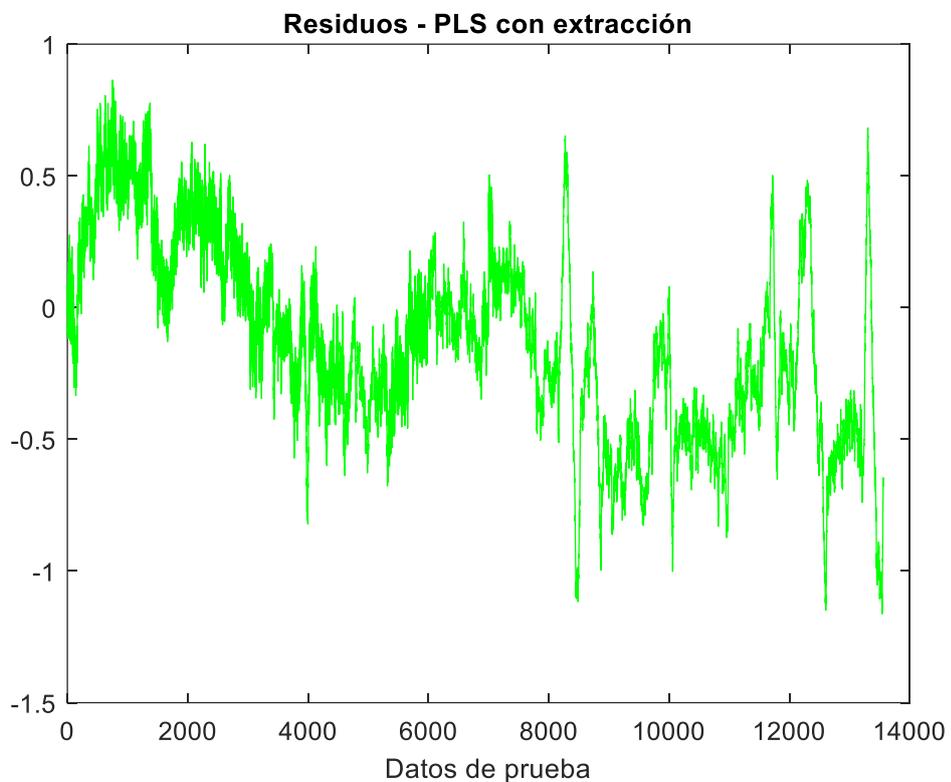
De esas 16 variables, el algoritmo NIPALS ha trabajado únicamente con 14, con lo cual podríamos predecir que el modelo iba a ser peor que el que utiliza un mayor número de variables, ya que cuenta con más información a la hora de realizar el modelado. Sin embargo, este método ofrece otras ventajas, como hemos explicado en el capítulo 2.



**Figura 4.23:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo PLS(EX) corregido (prueba).



En la figura 4.24 se representa la gráfica de residuos para comprobar que no existe ningún problema alarmante. En torno a la observación 8000, los residuos empiezan a oscilar en valores mayores. Esto es debido a cambios bruscos obtenidos en la salida de la planta real, donde vemos valores más bajos de la salida, lo que ocasiona una situación inesperada para el método, que intenta resolver con lo que ha aprendido durante el entrenamiento.



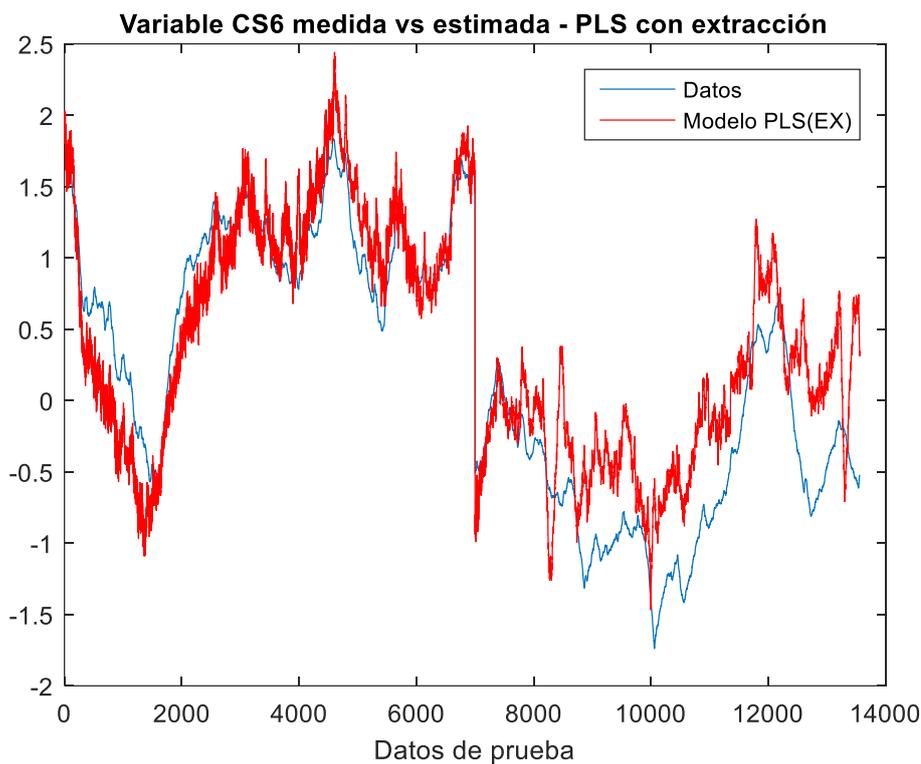
*Figura 4.24: Residuos obtenidos al aplicar PLS(EX) (prueba).*

### **4.3.3.3 PRUEBA CON FALLO CON PLS CON EXTRACCIÓN DE CARACTERÍSTICAS**

Llega la hora de comprobar si el método es capaz de detectar un fallo provocado. En este caso, debemos provocar el fallo en una

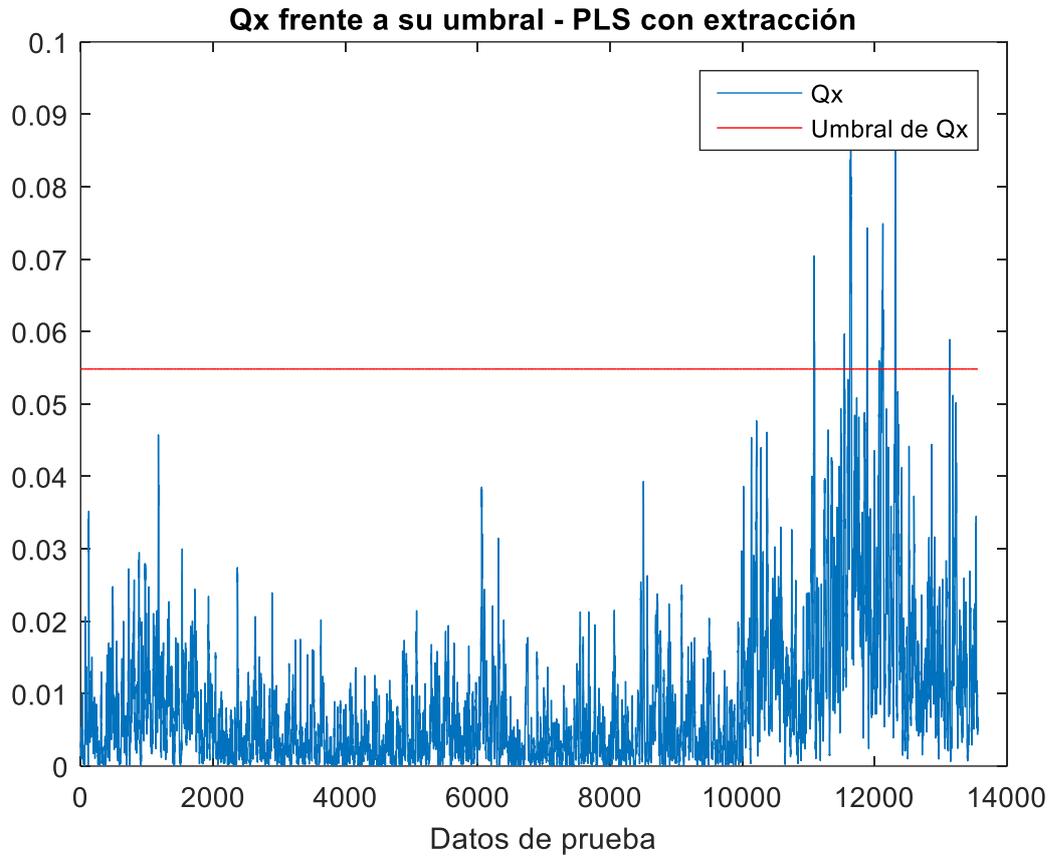


de las variables extraídas, ya que no trabaja con el resto, con lo que no conseguiríamos detección en este caso. Hemos elegido la variable T27 (ver figura 4.1) que corresponde a la temperatura de vapor en el evaporador 6. En la figura 4.25 podemos observar que la salida de la variable de calidad modelada por PLS(EX) no difiere mucho respecto a la prueba sin fallo.



**Figura 4.25:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo PLS(EX) corregido (prueba con error).

Esto es debido a que estamos introduciendo errores muy pequeños, con el fin de comprobar si el método es realmente sólido. En este caso, un incremento de 0.05 en valores del orden de  $10^2$  como son las temperaturas en las que trabaja el proceso parece un incremento insignificante. El error ha sido introducido, de nuevo, en el instante 10000. Comprobaremos ahora si el método es capaz de detectarlo. En este caso, como podemos apreciar en la figura 4.26, el estadístico  $Q_x$  no detecta el error.



**Figura 4.26:** Estadístico  $Q_x$  frente a su umbral  $UmbralQ_x$  en el método PLS(EX) corregido (prueba con error).

Sin embargo, en la siguiente figura (4.27) vemos que el estadístico  $T^2$  lo ha detectado a la perfección y con rapidez, en este ejemplo, en el instante 10149. Por último, comprobaremos que el sistema de diagnóstico de fallos es capaz de señalar con precisión la variable que está provocando el fallo. Como vemos en la figura 4.28, el diagnóstico es perfecto, ya que  $T^2$  nos indica que la variable 10 tiene un peso rotundamente superior en ese fallo respecto al resto de variables extraídas.

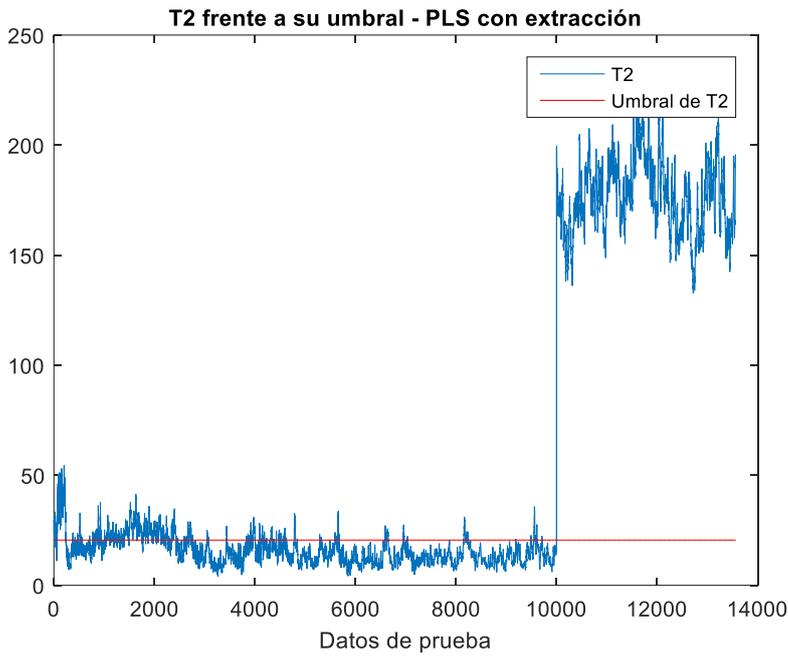


Figura 4.27: Estadístico  $T^2$  frente a su umbral  $UmbralT^2$  en el método PLS(EX) corregido (prueba con error).

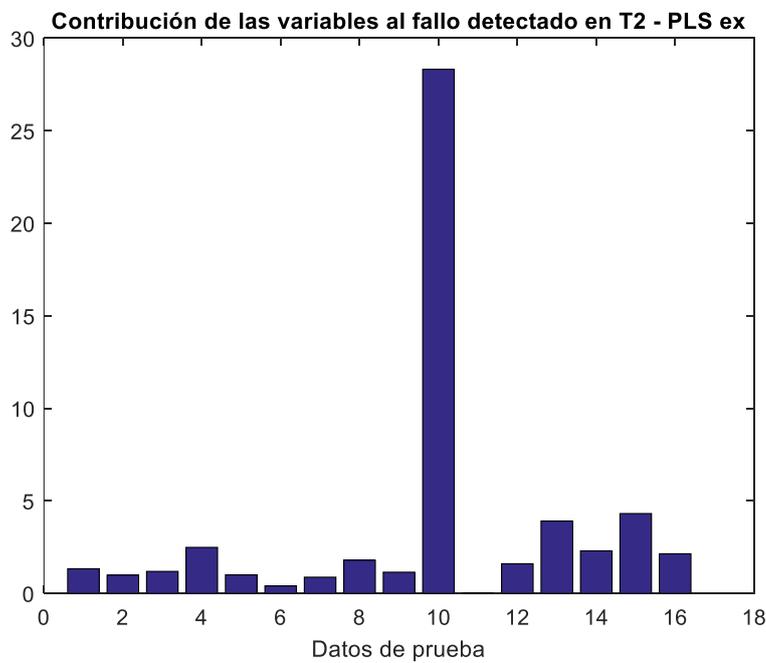


Figura 4.28: Diagrama de contribución al fallo por variables según el sistema de diagnóstico de  $T^2$  en PLS(EX) (prueba con error).



El estadístico  $Q_x$  detectó mejor el error introducido en el PLS genérico, y sin embargo en este caso de PLS(EX) no lo ha detectado, mientras que  $T^2$  lo ha hecho eficazmente. Esto es debido al tipo de dato en el cual provoquemos el fallo. El error introducido en PLS fue en el valor de una presión, la cual está relacionada con valores exponenciales, con lo cual es más fácil de detectar para el estadístico  $Q_x$ .

Por otro lado, el fallo provocado en PLS(EX) fue en el valor de una de las temperaturas del proceso, las cuales están correlacionadas de forma más lineal, con lo que la detección es favorable para el estadístico  $T^2$ . De esta manera, los métodos se complementan para detectar cualquier tipo de error, creando un sistema de detección y diagnóstico de fallos robusto.

### 4.3.4 APLICACIÓN DE PLS DINÁMICO

En este apartado estudiaremos la aplicación del método PLS dinámico en los datos recogidos de la planta. La diferencia respecto al PLS genérico será que este método tiene en cuenta la historia reciente de la planta. En este caso, se ha añadido a la matriz  $X$  de datos, dos retardos de la propia matriz, y un retardo de la matriz  $Y$ . Es decir, la nueva matriz de predicción será una matriz de la forma:

$$X_{\text{dinámica}} = [X(t) \quad Y(t-1) \quad X(t-1) \quad X(t-2)]$$

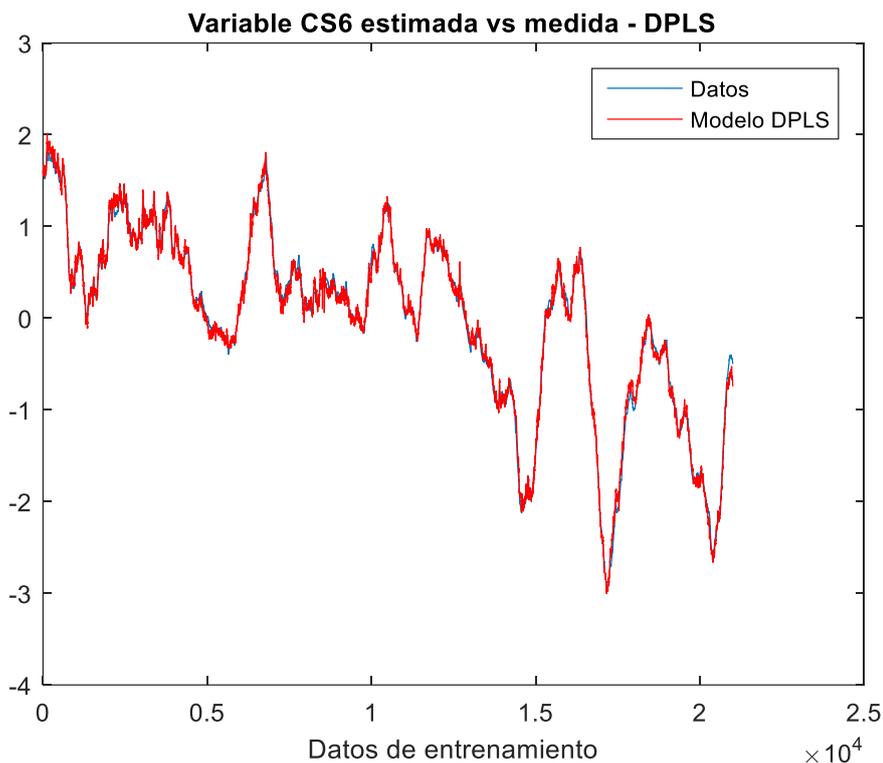
El conjunto de datos no ha sido modificado, simplemente ahora nuestra matriz de predicción contará con un total de 151 variables, es decir, 50 por cada  $X$  introducida en la matriz dinámica, y 1 por la  $Y$  retardada.



### 4.3.4.1 ENTRENAMIENTO CON PLS DINÁMICO

En la figura 4.29 podemos comparar la estimación de la variable de calidad CS6 por el modelo DPLS frente a la variable medida. Vemos que el modelado es muy bueno en este caso, con muy pocos picos que indiquen ruido o perturbaciones, y siguiendo de manera casi exacta el patrón seguido por la variable medida.

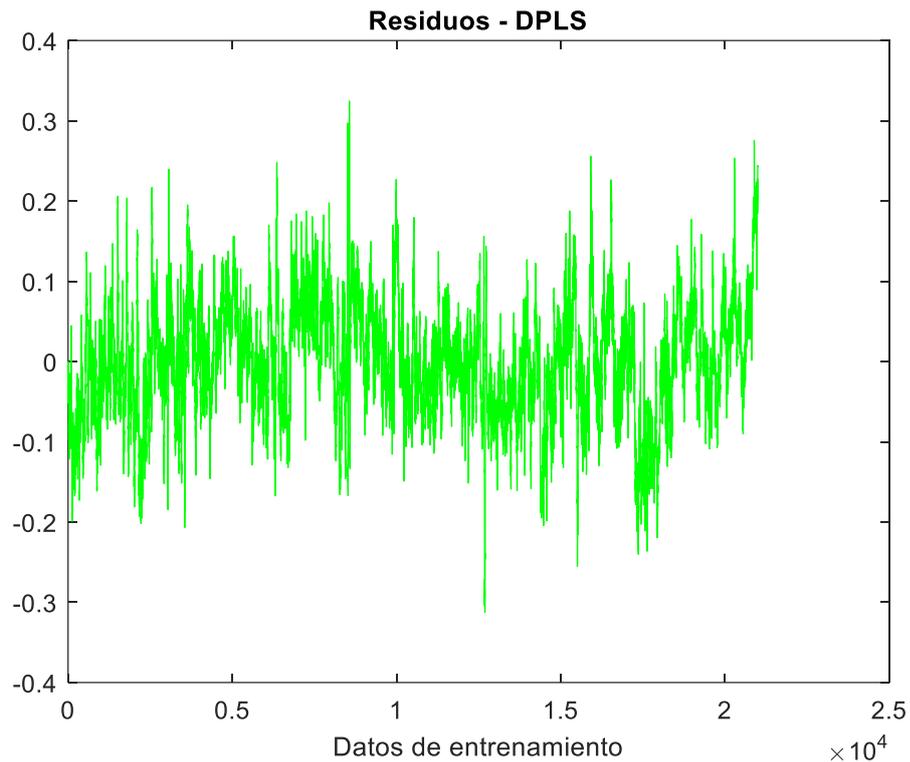
En realidad, es normal que el modelo sea mejor que PLS genérico, ya que el método cuenta con más información para entrenar. Sin embargo, debemos esperar a realizar la prueba para confirmar que el método funciona de forma correcta.



**Figura 4.29:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo DPLS (entrenamiento).



Como podemos ver en la figura 4.30, los residuos han disminuido considerablemente, estando concentrados en su mayoría en un rango de  $\pm 0.2$ , con lo que podemos afirmar que es sustancialmente menor que los residuos arrojados por PLS genérico.



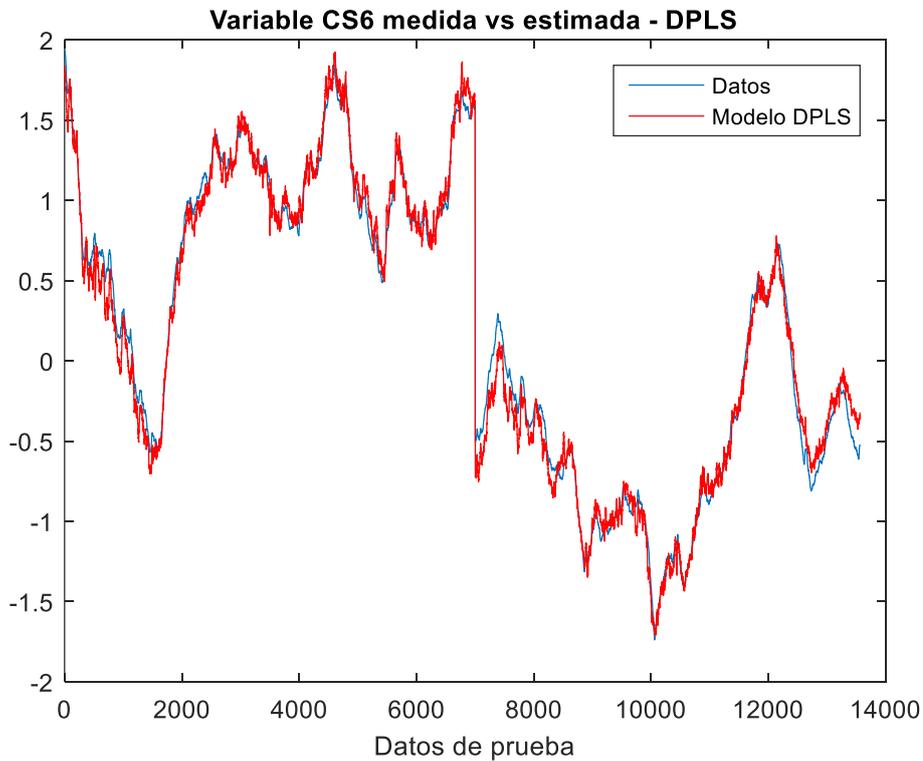
**Figura 4.30:** Residuos obtenidos al aplicar DPLS (entrenamiento).

#### **4.3.4.2 PRUEBA CON PLS DINÁMICO**

Pasaremos ahora a comprobar si realmente el método es suficientemente bueno como para mejorar al PLS genérico, probándolo en los datos correspondientes. En la figura 4.31 se ha representado la estimación de nuestra variable de calidad CS6 frente a su valor medido, y vemos que la simulación ha sido muy buena, desviándose ligeramente en el rango de datos que vienen



tras el salto entre los dos rangos diferentes que componen la matriz de datos de prueba. Sin embargo, se recupera correctamente en los siguientes instantes.



**Figura 4.31:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo DPLS (prueba).

En este caso, se ha decidido introducir un fallo en la variable T4, que corresponde a la temperatura del jugo a la entrada del intercambiador de calor R14. En la figura 4.32 está representado el estadístico  $Q_x$  frente a su umbral, donde podemos ver cómo claramente a partir del instante 10000 aproximadamente (que es el instante en el que el error ha sido introducido), el estadístico toma valores que superan el umbral con claridad. Ha detectado el fallo con claridad, concretamente en el instante 10147.

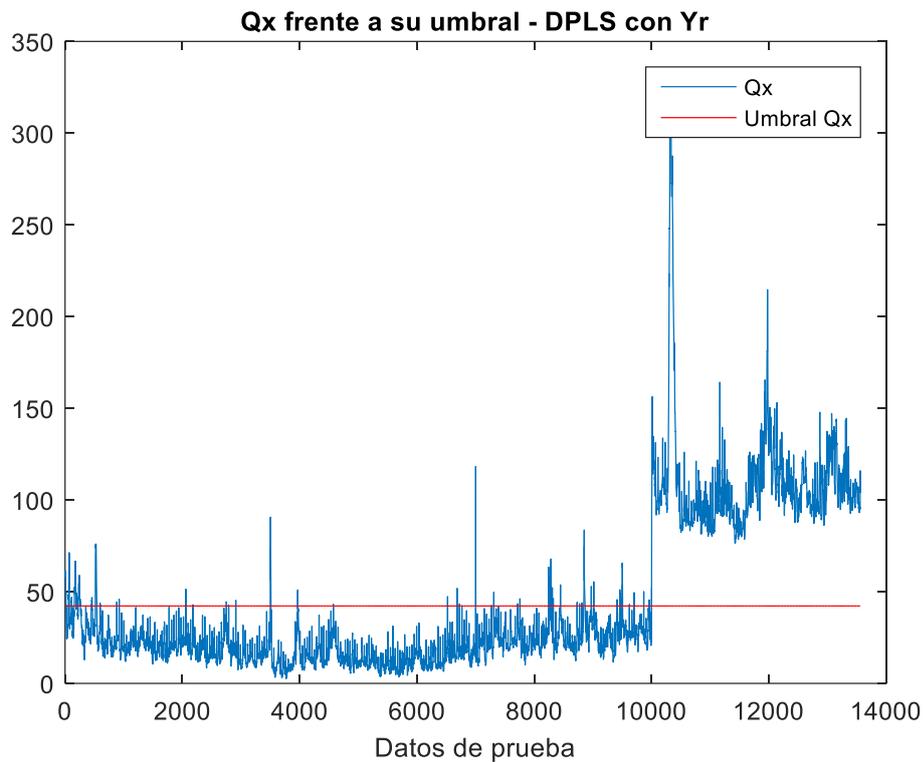


Figura 4.32: Estadístico  $Q_x$  frente a su umbral  $UmbralQ_x$  en el método DPLS (prueba con error).

También tenemos que comprobar si el sistema de diagnóstico del fallo es correcto. En la siguiente figura, la 4.33, vemos cómo, efectivamente el diagrama de contribución muestra la variable 10 como predominante sobre las demás. En nuestro caso, también muestra como predominantes la variable 61, así como la 111, las cuales corresponden a los retardos  $t-1$  y  $t-2$  de esa variable, por lo que es lógico que también nos las esté indicando. Es la misma variable en instantes diferentes, pero el fallo está en todas ellas. Podemos, por tanto, dar el método como eficaz.



Contribución de las variables al fallo detectado en  $Q_x$  - PLS dinámico con  $Y_r$

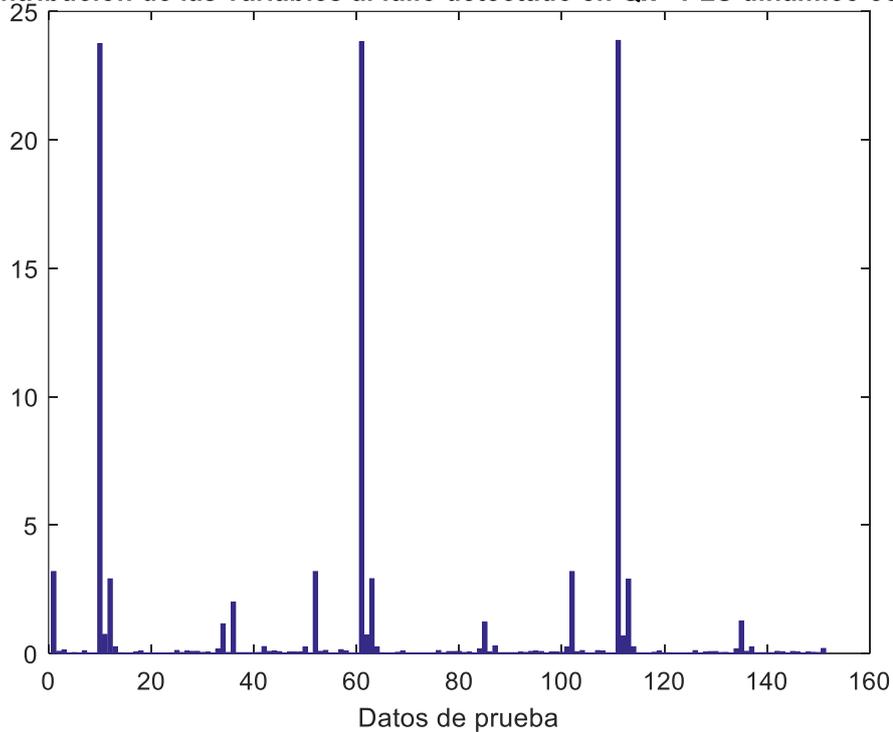


Figura 4.33: Diagrama de contribución al fallo por variables según el sistema de diagnóstico de  $Q_x$  (prueba con error).

Sin embargo, para ese error introducido, el estadístico  $T^2$  no detecta el fallo, como podemos apreciar en la figura 4.34. Como explicamos antes, dependiendo del tipo de dato en el cual provoquemos el error, o el sistema encuentre una anomalía, puede detectarlo un solo método o los dos. Esto no quiere decir que el método sea menos fiable, ya que la detección por ambos estadísticos es complementaria.

Nos quedamos con los resultados obtenidos, dando por eficaz el método, y más adelante procederemos a compararlo con el resto de métodos estudiados.

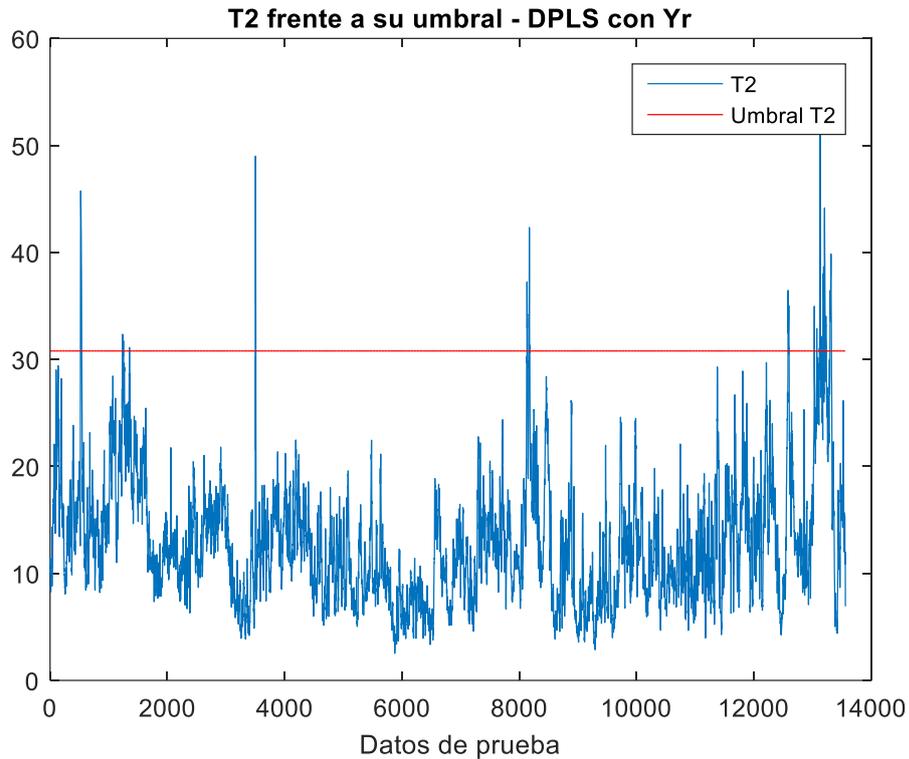


Figura 4.34: Estadístico  $T^2$  frente a su umbral  $UmbralT^2$  en el método DPLS (prueba con error).

### 4.3.5 APLICACIÓN DE PLS NO LINEAL

Hasta ahora, la variable de calidad del proceso era calculada a través de un vector estimado a partir de una regresión lineal. Para transformar esta regresión en no lineal, añadiremos a la matriz de predicción  $X$  componentes no lineales, con el objetivo de ver la evolución del método genérico al ser modificado de esta forma.

Las variables no lineales introducidas han sido modificaciones exponenciales, logarítmicas, o de la forma de raíces cuadradas de las variables presión y temperatura de la matriz de predicción  $X$ . También hemos introducido modificaciones de la variable de calidad como entrada, aunque para ello, hemos de retardarla al



menos un instante, ya que no podemos utilizar para el cálculo la variable que vamos a calcular. Estas relaciones matemáticas utilizadas en las variables se obtienen del conocimiento que se tiene del proceso, y como se puede calcular matemáticamente la relación entre la salida y las entradas.

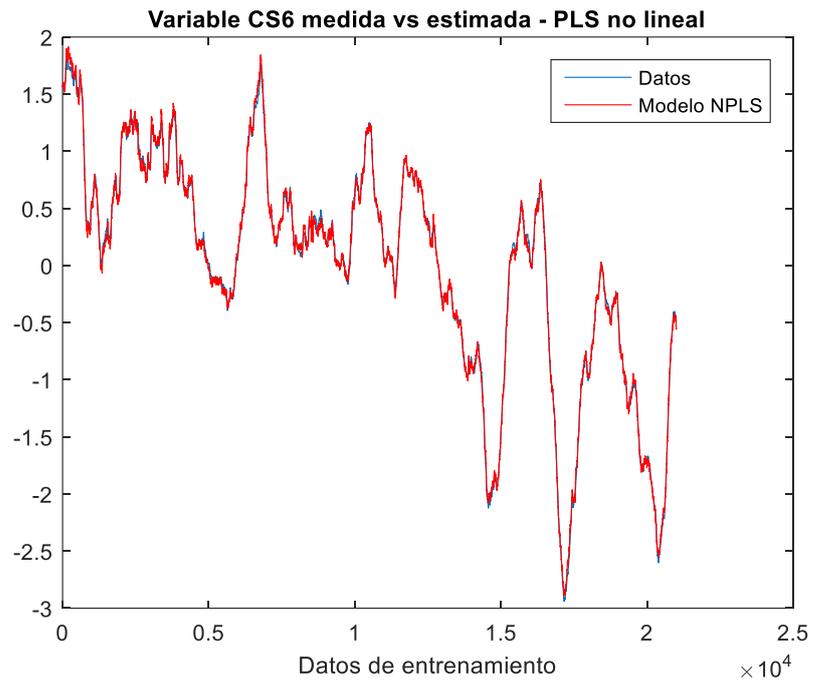
En total, la nueva matriz de predicción no lineal consta de 129 variables, conformadas por la matriz X original más las nuevas variables no lineales.

#### **4.3.5.1 ENTRENAMIENTO CON PLS NO LINEAL**

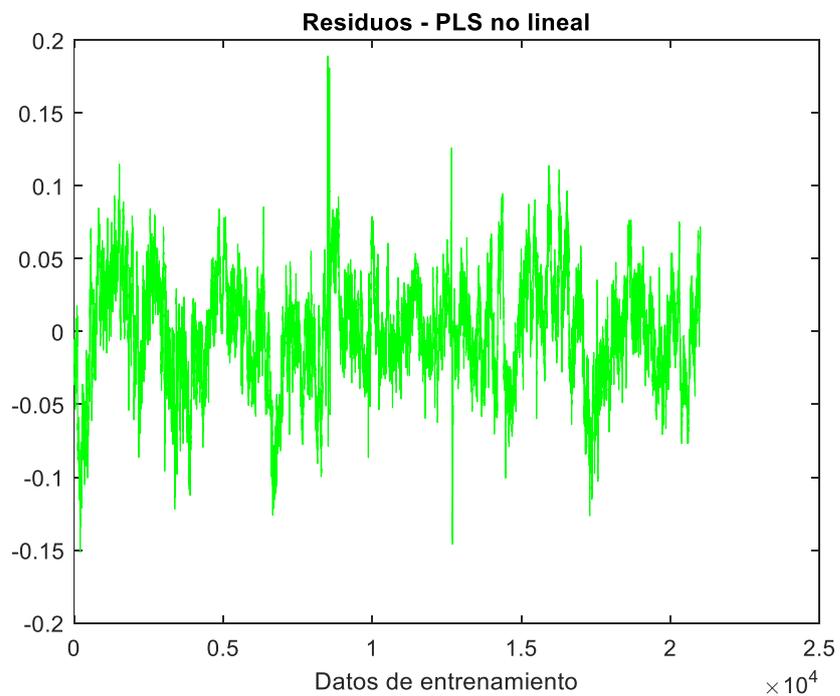
Vamos a comenzar, como en los demás métodos, realizando el proceso de entrenamiento de datos.

En la figura 4.35 podemos observar la comparación entre la variable CS6 real, medida de la planta, y la estimada por NPLS. Vemos cómo el modelado es prácticamente perfecto en este caso, obteniendo resultados casi iguales que los recogidos de la planta.

Esto podemos reafirmarlo fijándonos en los residuos arrojados (figura 4.36) por el método, los cuales esta vez se ven reducidos a un rango de tan solo  $\pm 0.5$ , prácticamente despreciables. A la vista de estos resultados, podemos pensar que este método resultará el más robusto de todos, pero antes, debemos asegurarnos de que no ha aprendido los datos de entrenamiento de memoria, lo que podría dar lugar a un mal resultado con los datos de prueba, es decir no es capaz de generalizar a datos no vistos durante el entrenamiento.



**Figura 4.35:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo NPLS (entrenamiento).



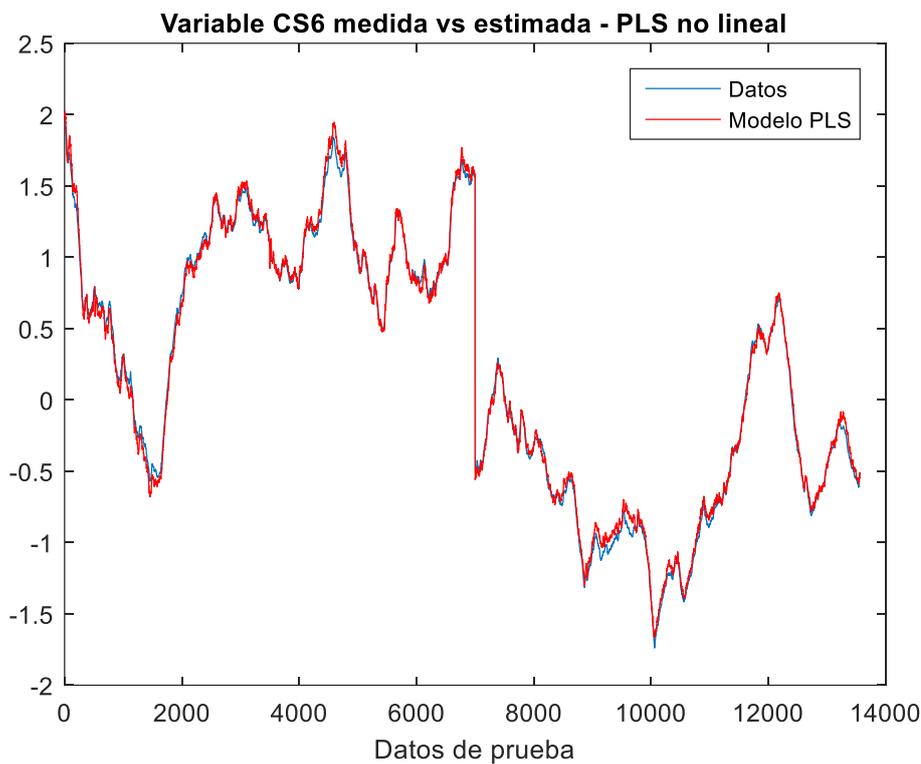
**Figura 4.36:** Residuos obtenidos al aplicar DPLS (entrenamiento).



### 4.3.5.2 PRUEBA CON PLS NO LINEAL

Pasaremos ahora a realizar la prueba del método para comprobar que nuestra hipótesis de situarlo como método más robusto está fundamentada.

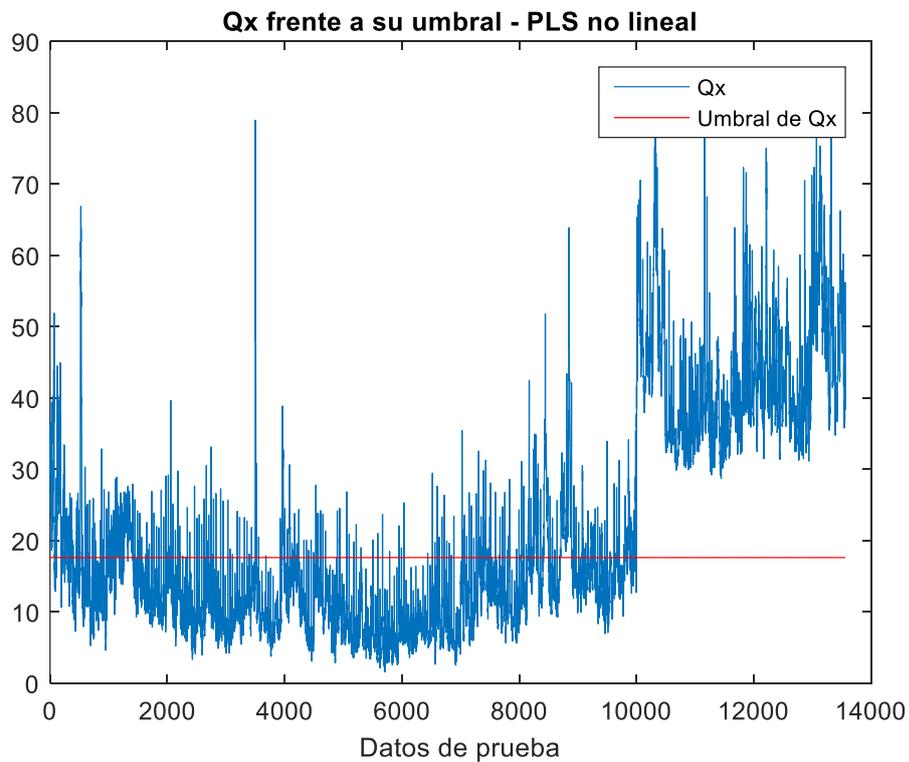
En la siguiente figura, la 4.37 podemos ver la precisión con la que ha sido modelada también la variable de calidad con los datos de prueba, en los que además se ha añadido un fallo en una de las variables de entrada al proceso, prácticamente sin mostrar ningún patrón diferente a la variable medida.



**Figura 4.37:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo NPLS (prueba).



En este caso, el error fue introducido en la variable T4 (Temperatura del jugo a la entrada del intercambiador de calor R14), y de nuevo, el estadístico  $Q_x$  lo detecta sin problemas. El tiempo de detección fue el instante 10149. Podemos ver en la gráfica 4.38 cómo a partir del instante de error inducido, la observación 10000, el valor de  $Q_x$  se dispara muy por encima del umbral.

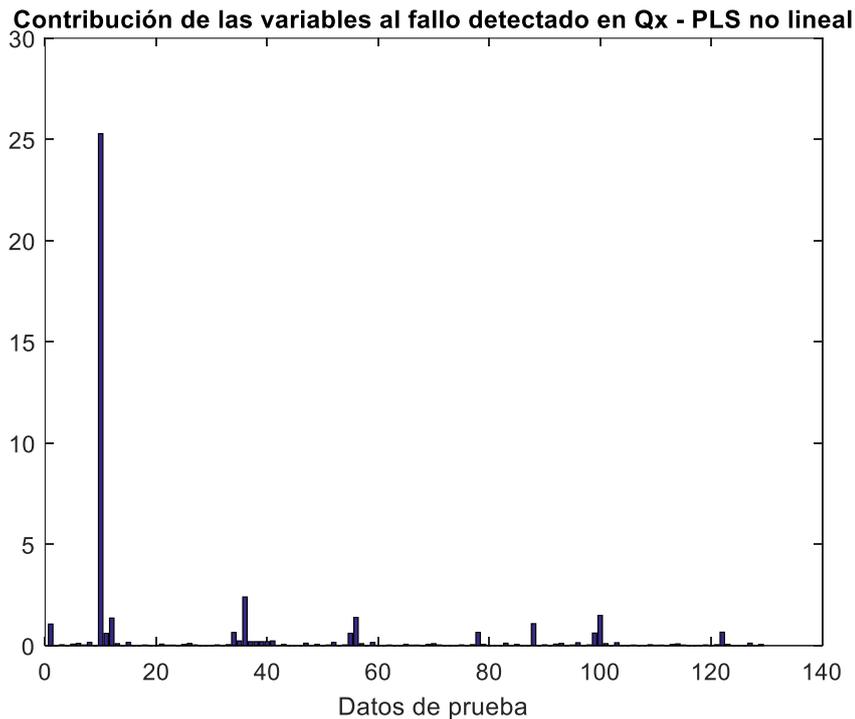


**Figura 4.38:** Estadístico  $Q_x$  frente a su umbral  $UmbralQ_x$  en el método NPLS (prueba con error).

Por último, debemos comprobar si el sistema de diagnóstico de fallos del estadístico  $Q_x$  es capaz de informar correctamente acerca de cuál es la variable más relevante que está provocando el error en el sistema. Esto lo vemos representado en el diagrama de contribución de la figura 4.39, donde podemos ver cómo claramente la variable número 10 predomina sobre las demás, por tanto, el diagnóstico es correcto.



De nuevo, en este caso el estadístico  $T^2$  no fue capaz de detectar este fallo. Como hemos comentado previamente, esto depende del tipo de dato responsable del fallo, en este caso una temperatura, por lo que era de esperar que el estadístico más propenso a detectarlo fuera  $Q_x$ .



*Figura 4.39: Diagrama de contribución al fallo por variables según el sistema de diagnóstico de  $Q_x$  (prueba con error).*

### **4.3.6 APLICACIÓN DE COMBINACIONES DE MÉTODOS PLS**

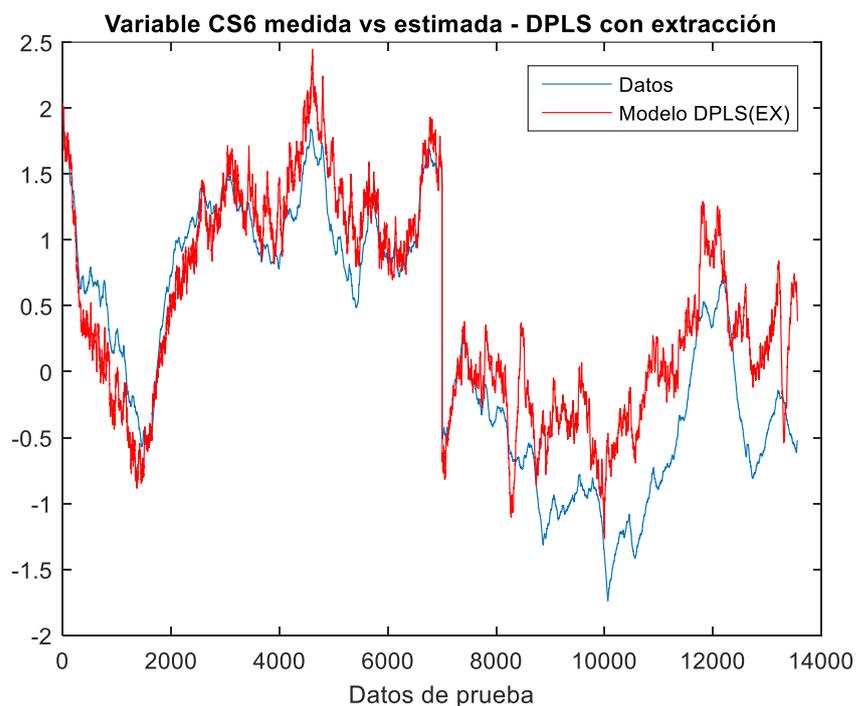
Con este apartado cerraremos el grupo de capítulos referentes al método por regresión PLS y sus derivados (PLSEX, DPLS, NPLS). En este caso trataremos de analizar si una o varias combinaciones de los métodos anteriormente descritos, puede



dar lugar a mejoras respecto al resto, referentes a la detección y diagnóstico de fallos al introducir errores manualmente.

### **4.3.6.1 PLS DINÁMICO CON EXTRACCIÓN DE CARACTERÍSTICAS**

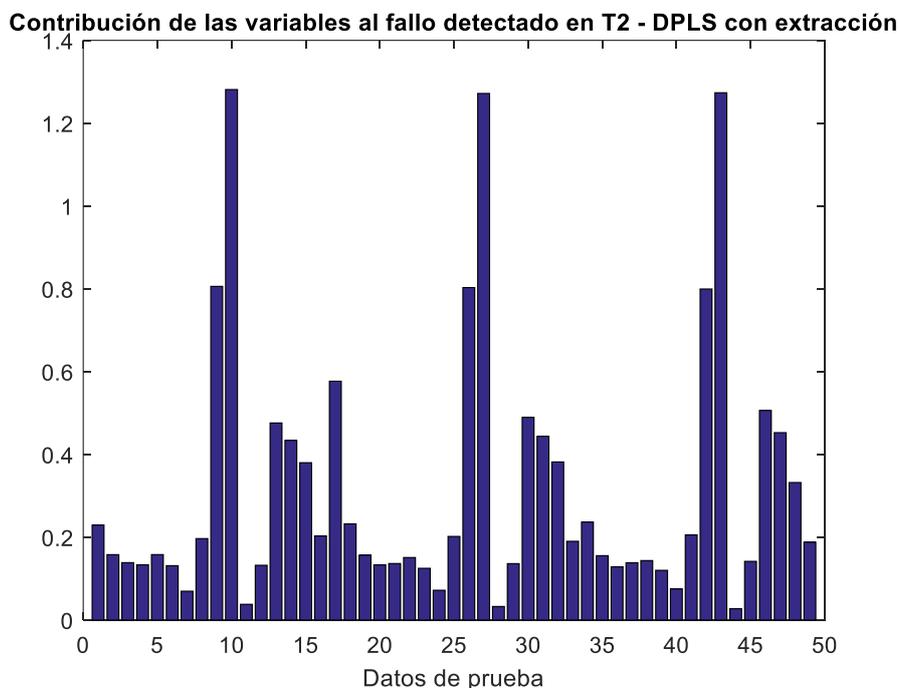
Comenzaremos utilizando nuestro método de extracción de características mediante correlación entre variables, a la matriz X dinámica utilizada en el apartado 4.3.4. En primer lugar, crearemos la matriz dinámica de X mediante dos retardos de X y un retardo de la variable estimada Y. Una vez creada, procederemos a extraer características, y entrenaremos con ese método.



**Figura 4.40:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo DPLS(EX) (prueba con error).



Una vez hayamos entrenado, procederemos a probar el método con los datos correspondientes, provocando en este caso un fallo en la variable T4 (Temperatura del jugo a la entrada al intercambiador de calor R14). En la figura 4.40 podemos comparar nuestra variable de calidad CS6 tomada de los datos, frente a la modelada por el método DPLS(EX). A la vista de los resultados, podemos decir que la aproximación no ha sido demasiado buena, puesto que el fallo se ha introducido en el instante 10000, y sin embargo, vemos que el modelo comienza a desviarse desde el inicio.



**Figura 4.41:** Diagrama de contribución al fallo por variables según el sistema de diagnóstico de  $T^2$  en DPLS(EX) (prueba con error).

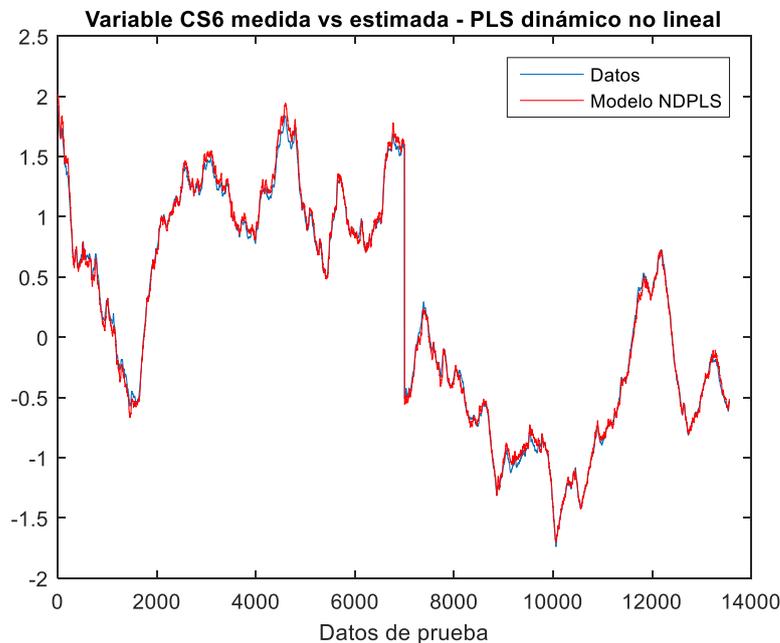
A la hora de detectar el fallo, no hemos observado ningún problema, ya que el estadístico  $T^2$  lo detecta con precisión, en el instante en el que lo hemos introducido. En la figura 4.41 podemos ver el diagrama de contribución generado por el sistema de diagnóstico de fallos. En él, podemos ver claramente



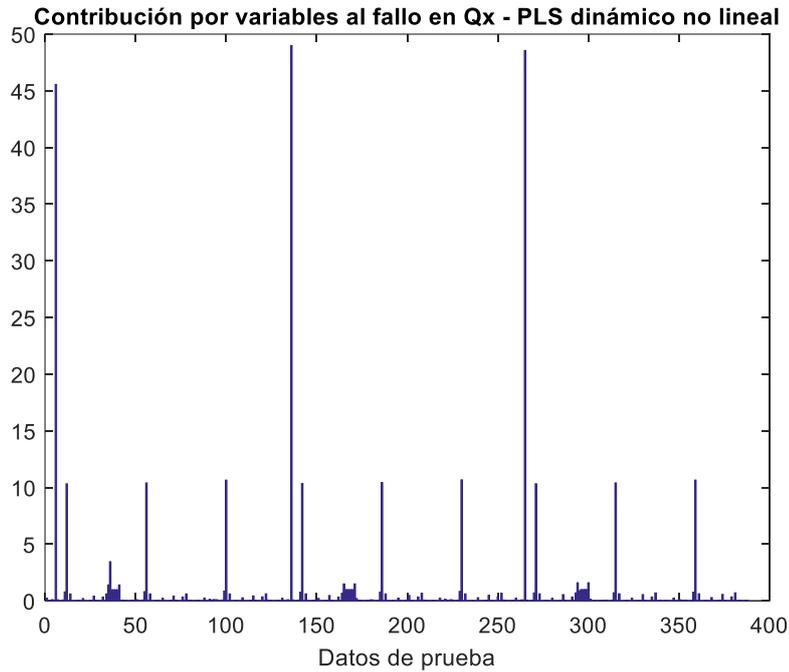
que tres variables predominan sobre las demás. Esto es debido a que la variable T4 en la que hemos introducido el fallo, está retardada dos veces en la matriz dinámica, y el sistema por lo tanto detecta la variable en la que se ha producido el fallo, y sus respectivos retardos.

### **4.3.6.2 PLS DINÁMICO NO LINEAL**

En este caso, a la matriz dinámica X creada con los retardos mencionados anteriormente, la dotaremos de carácter no lineal gracias a las relaciones descritas en este método, creando una matriz que utilice datos pasados, y a su vez datos no lineales como entradas. En la figura 4.42 podemos ver cómo al dotar de carácter no lineal al sistema, la variable de salida es modelada con gran precisión incluso en el proceso de prueba con error.



**Figura 4.42:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo NDPLS (prueba con error).



**Figura 4.43:** Diagrama de contribución al fallo por variables según el sistema de diagnóstico de  $Q_x$  en NDPLS (prueba con error).

Como sucedió anteriormente, en la figura 4.43 podemos ver que el diagrama de contribución obtenido gracias al sistema de diagnóstico de fallos nos muestra varias variables como predominantes. Igual que antes, esto es debido a que las variables se repiten, simplemente son retardadas un instante, por lo que, si una variable tiene un peso importante a la hora de provocar un fallo, también lo tendrán sus análogas retardadas.

### **4.3.6.3 PLS NO LINEAL CON EXTRACCIÓN DE CARACTERÍSTICAS**

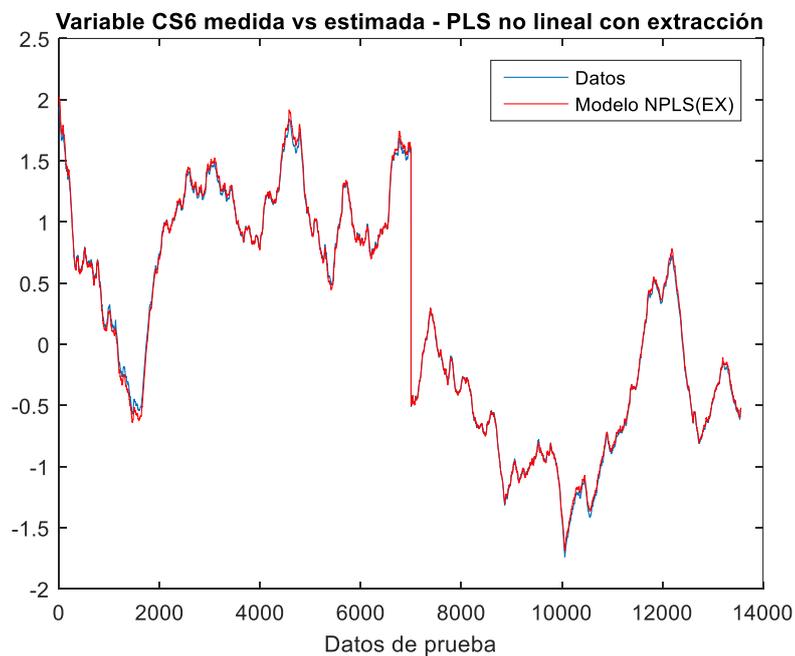
En este caso pasaremos a testear el método de extracción de características mediante correlación entre variables, a la matriz de predicción no lineal de  $X$ . El procedimiento es similar al del PLS dinámico con extracción de características.



En primer lugar, creamos la nueva matriz X con las relaciones no lineales descritas anteriormente, y sobre esa nueva matriz, aplicamos extracción de características para quedarnos con las variables más relevantes y trabajar solamente con ellas.

En la figura 4.44 podemos ver que la estimación de la variable de calidad CS6 es prácticamente perfecta gracias al carácter no lineal de la matriz X de predicción. Esto también nos indica que la extracción de características ha elegido como variable relevante a alguna de las no lineales introducidas, o eso podemos deducir al lograr una estimación tan buena, incluso habiendo introducido ya el fallo.

En la figura 4.45 podemos corroborar esta hipótesis al ver los residuos, que son prácticamente despreciables.



**Figura 4.44:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo NPLS(EX) (prueba con error).

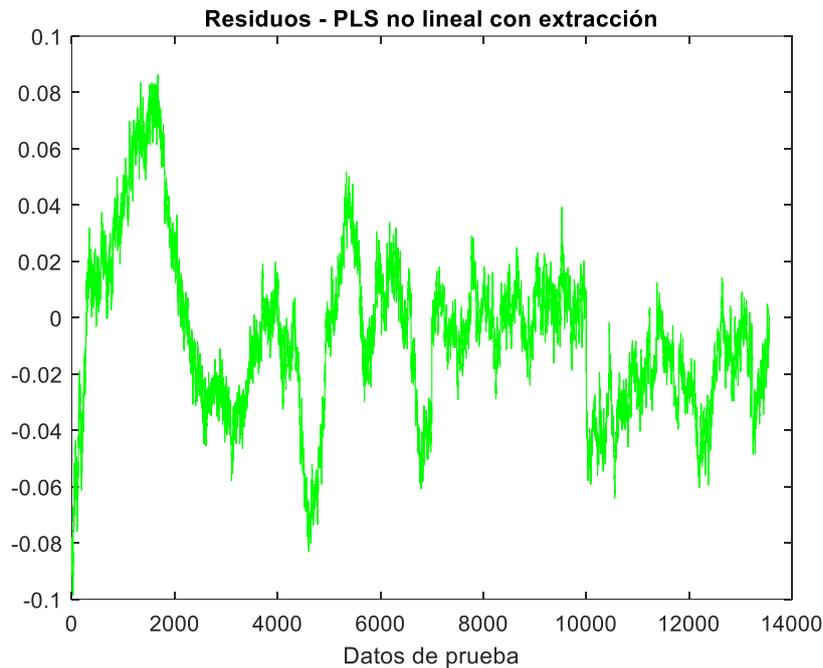


Figura 4.45: Residuos obtenidos al aplicar NPLS(EX) (prueba con error).

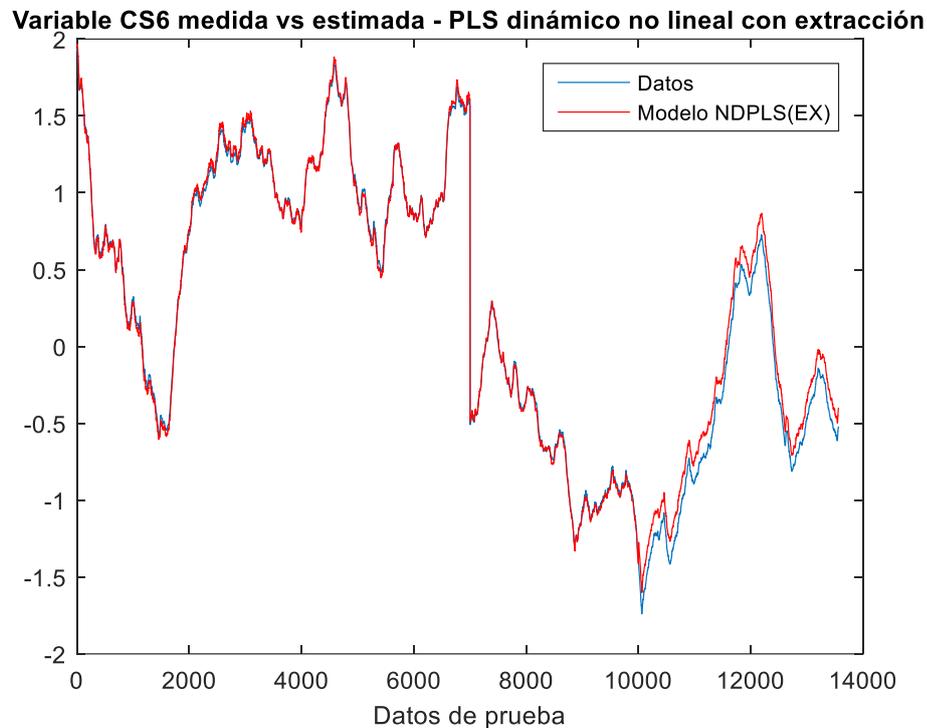
#### 4.3.6.4 PLS DINÁMICO NO LINEAL CON EXTRACCIÓN DE CARACTERÍSTICAS

Para finalizar con el apartado de PLS, probaremos cada uno de los métodos en uno solo. El procedimiento en este caso ha sido similar a los utilizados en PLS dinámico y PLS no lineal, ambos con extracción de características. La diferencia es que primero crearemos nuestra matriz dinámica  $X$  mediante los retardos de  $X$  y el retardo de  $Y$ . Posteriormente, dotaremos a la matriz de carácter no lineal, y una vez hecho esto, procederemos a extraer las variables más significativas mediante el método de correlación ya descrito anteriormente.

El resultado de la estimación de nuestra variable de calidad CS6 puede verse en la figura 4.46, donde vemos que la estimación es casi perfecta hasta el instante en el que el fallo es introducido, la

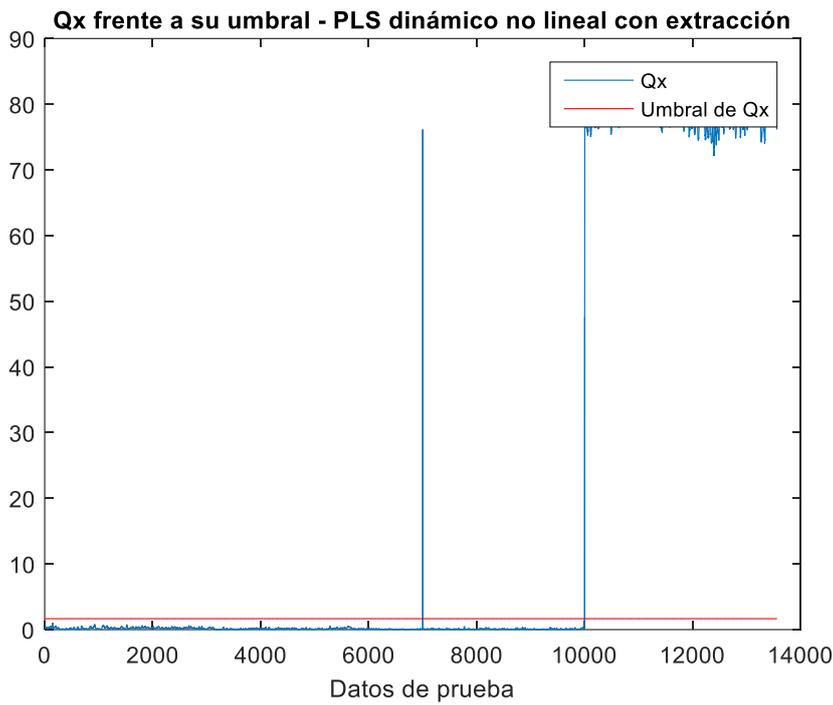


observación 10000, donde la estimación comienza a desviarse ligeramente, aunque no parece preocupante.

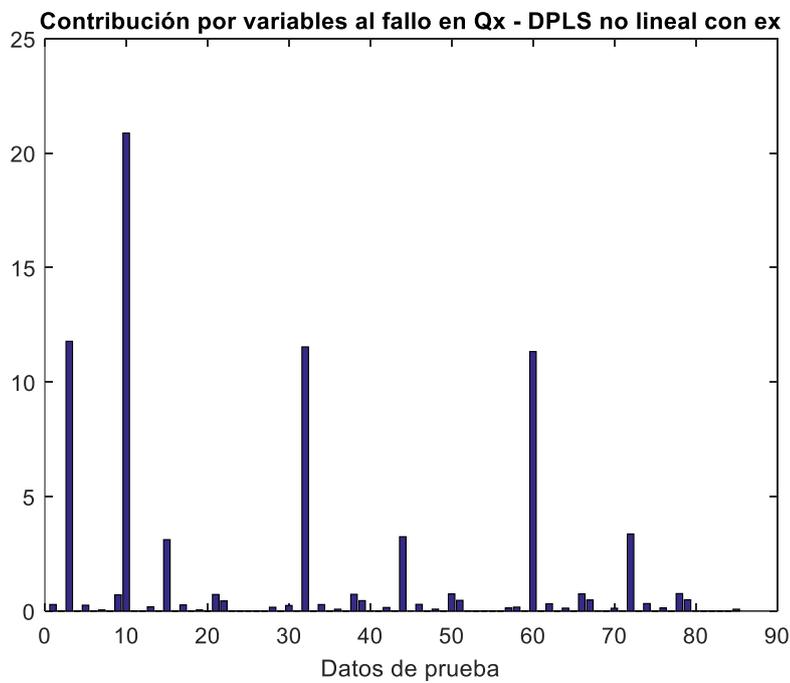


**Figura 4.44:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo NPLS(EX) (prueba con error).

En la figura 4.47 podemos ver cómo el estadístico  $Q_x$  detecta con claridad el fallo en el instante en el que lo hemos introducido, y a su vez, el sistema de diagnóstico de fallos indica que la variable 10 es la responsable de la anomalía. Como nos ha sucedido anteriormente, también nos encontramos otras variables que predominan sobre las demás, debido a que son la misma variable en un instante y dos instantes atrás.



**Figura 4.45:** Estadístico  $Q_x$  frente a su umbral  $UmbralQ_x$  en el método NDPLS(EX) (prueba con error).



**Figura 4.46:** Diagrama de contribución al fallo por variables según el sistema de diagnóstico de  $Q_x$  en NDPLS(EX) (prueba con error).



## 4.4 APLICACIÓN DE ESTIMACIÓN DE LA VARIABLE CON RNA

En este apartado estudiaremos la estimación de la variable de calidad (concentración en masa de sacarosa a la salida del evaporador 6) mediante redes neuronales artificiales. Una vez obtenidos los resultados, procederemos a compararlos con los obtenidos mediante los métodos basados en PLS con el objetivo de estudiar si aportan alguna ventaja sobre los resultados arrojados por estos.

Se ha decidido utilizar para el estudio dos tipos de redes principales, la que llamaremos Red Neuronal (red con propagación hacia adelante) y la Red Dinámica (red retropropagada hacia adelante). Sobre ellas, trataremos de utilizar el método de extracción de características para comprobar cómo varían los resultados al trabajar con menos variables.

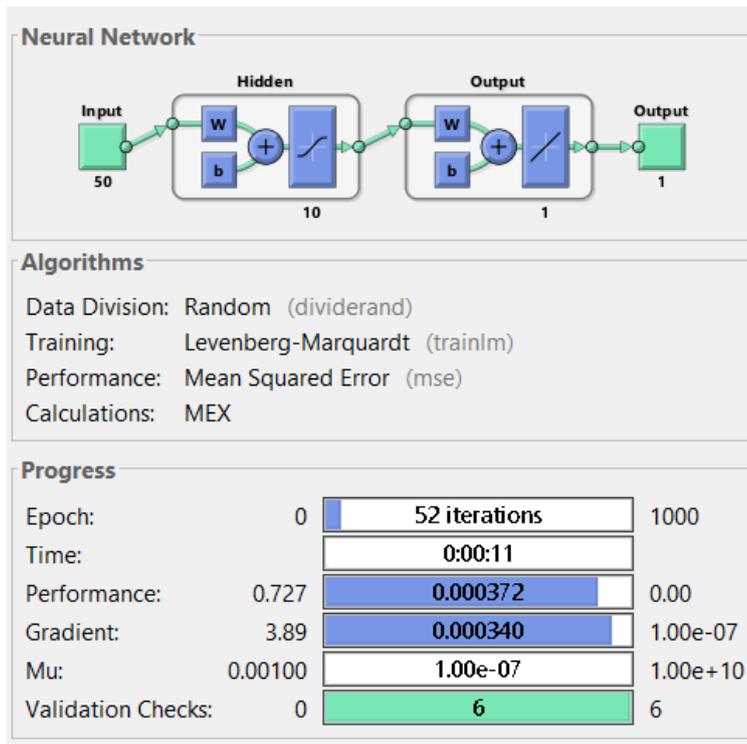
### 4.4.1 RED NEURONAL

En primer lugar, estudiaremos la estimación de la variable CS6 por medio de una Red Neuronal con propagación hacia adelante (*feedforward net*), en concreto una red tipo Perceptrón multicapa, en el cual las salidas de las neuronas van a las siguientes capas, pero no a las anteriores, de modo que no hay bucles de retroalimentación.

El número de neuronas en la capa oculta elegido ha sido cinco. El procedimiento seguido es similar al seguido en los métodos PLS, que es el siguiente: Primero, crearemos la red para entrenarla con los datos correspondientes, y posteriormente, utilizando los datos de prueba trataremos de estimar la variable con los nuevos



datos, y después analizar si la estimación ha sido correcta. En la figura 4.47 podemos ver la estructura de la red *feedforward* y el número de iteraciones que ha seguido la red hasta llegar al modelo estimado, en este caso 52 iteraciones.



**Figura 4.47:** Esquema de la red neuronal y número de iteraciones hasta la estimación (entrenamiento).

Como vemos en la figura 4.48, la iteración elegida como óptima ha sido la 46, arrojando un error cuadrático medio muy bajo, en torno a  $10^{-4}$ .

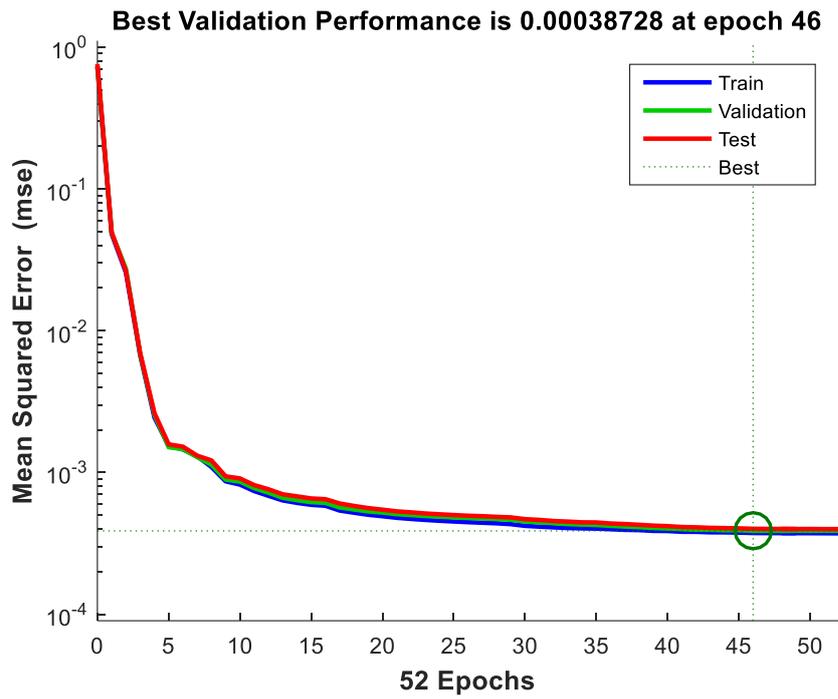


Figura 4.48: Error cuadrático medio por iteración en la red neuronal (entrenamiento).

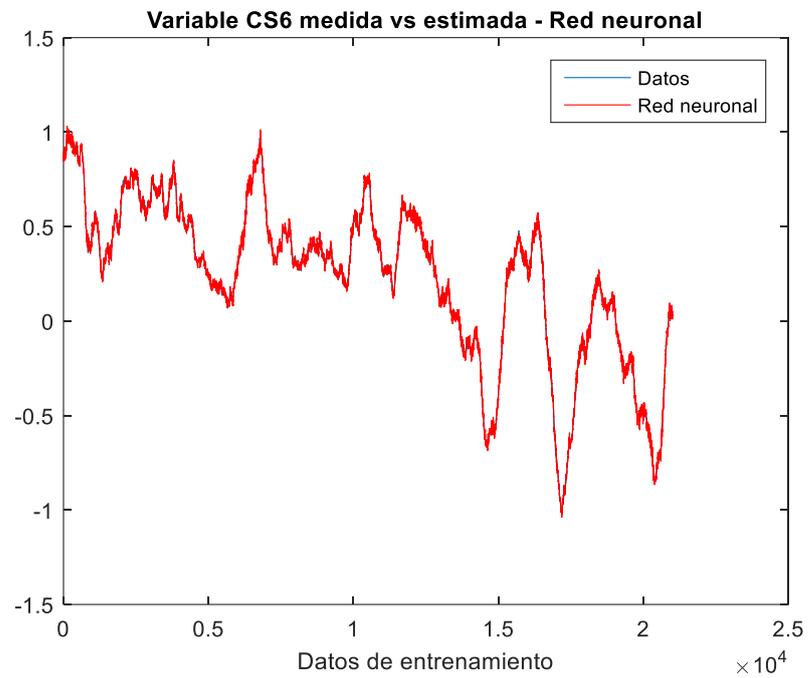
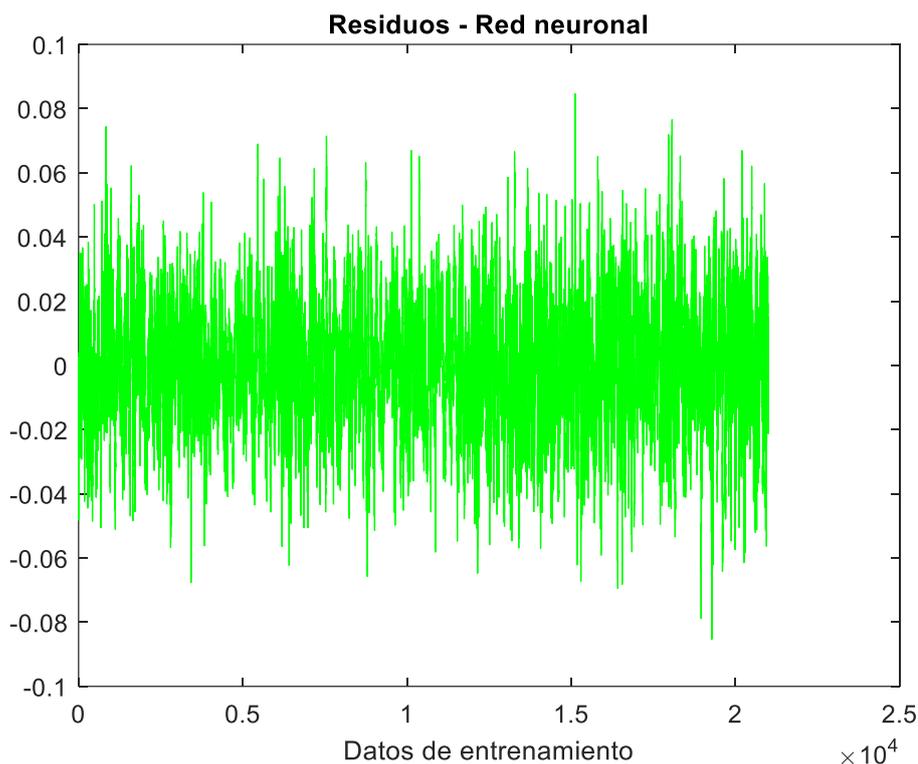


Figura 4.49: Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo Red Neuronal (entrenamiento).



Como podemos observar en la figura 4.49, la estimación de la variable de salida mediante la red neuronal es prácticamente perfecta, casi exacta. El patrón del modelo estimado sigue a la perfección al patrón de los datos extraídos directamente de la planta. Un caso similar sucedía en la aplicación de PLS no lineal, y en este caso, era esperado, ya que las redes neuronales forman modelos no lineales.

En la figura 4.50 vemos en el diagrama de residuos que estos son prácticamente despreciables, y se mantienen constantes a lo largo de todo el entrenamiento. Tanto es así, que hemos tenido que ampliar la gráfica obtenida para poder apreciar oscilaciones, pues inicialmente veíamos una línea recta en el valor 0.

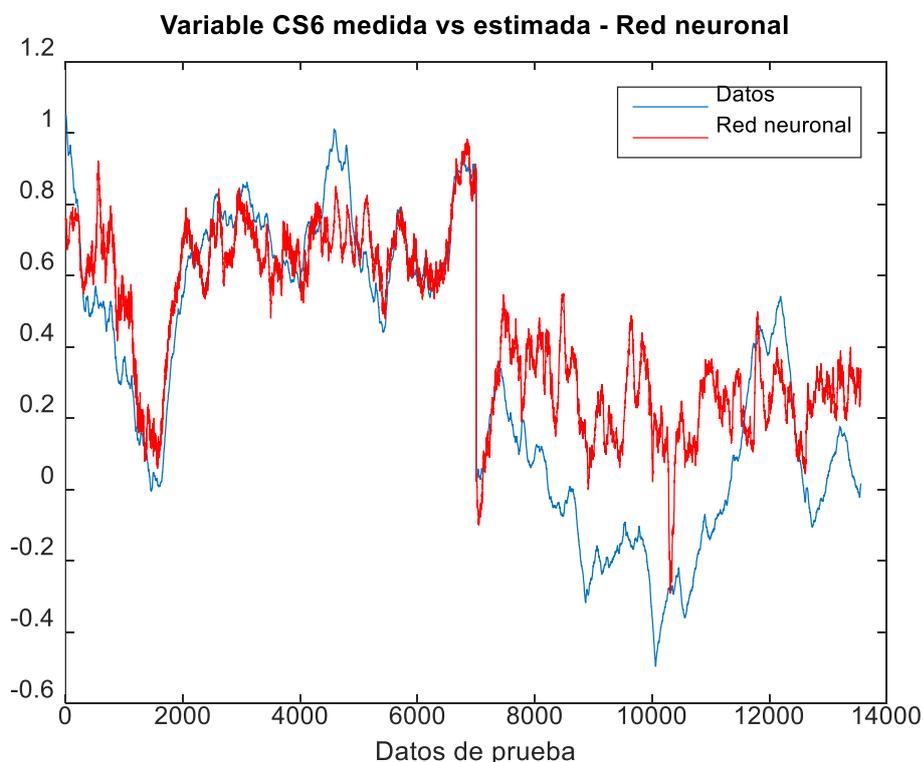


**Figura 4.50:** Detalle del plot de residuos obtenidos al aplicar la Red Neuronal (entrenamiento).



Llega el momento de realizar la prueba del método con los datos correspondientes, para analizar cómo se desenvuelve la red en situaciones anómalas para ella.

En este caso, como podemos apreciar en la figura 4.51, la prueba no ha sido para nada satisfactoria. Vemos que la variable modelada no comienza demasiado mal, pero a partir del instante 7000 aproximadamente, en el cambio de rango de datos, se desestabiliza totalmente, dando lugar a un modelo bastante malo. Esto es debido a que la red cuenta con demasiada información, ya que la matriz X de predicción para el entrenamiento cuenta con 21000 observaciones, lo que da lugar a que la red se aprenda de memoria el patrón seguido en el proceso de entrenamiento, y aunque el modelo de entrenamiento sea capaz de reproducirlo a la perfección, no es capaz de hacerlo al recibir nuevos datos.



**Figura 4.51:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo Red Neuronal (prueba).

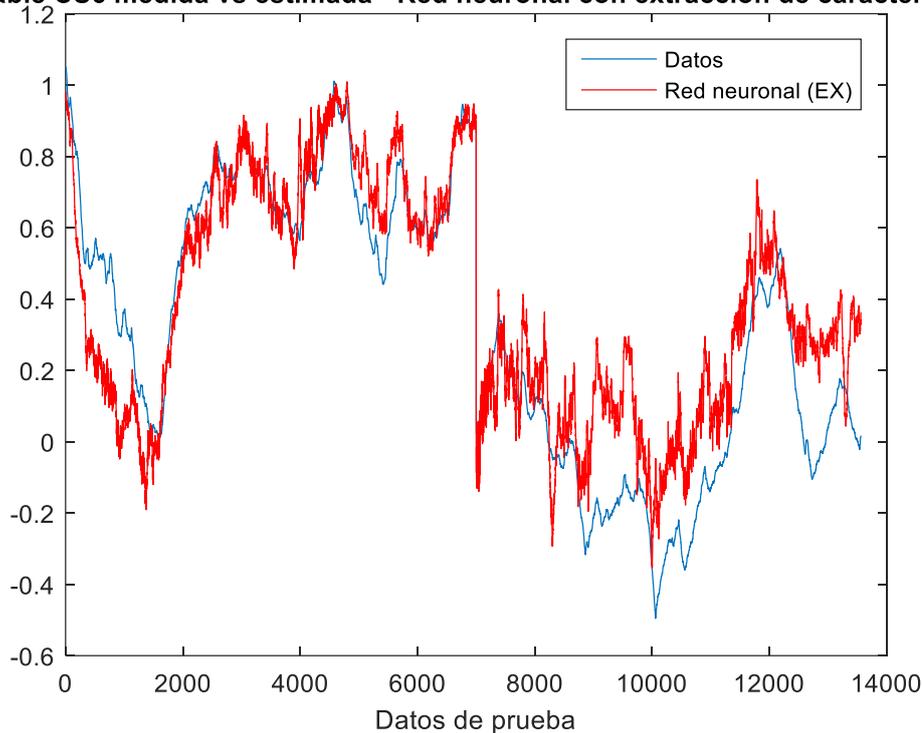


## 4.4.2 RED NEURONAL CON EXTRACCIÓN DE CARACTERÍSTICAS

Vamos a aplicar a la red neuronal descrita en el apartado anterior, la extracción de las variables más relevantes siguiendo el método ya descrito de correlación entre variables. Una vez creada la nueva matriz con las variables extraídas, crearemos la red y entrenaremos con los nuevos datos. El número de neuronas en la capa oculta elegido, como en la anterior, ha sido de cinco.

Dado que el entrenamiento en las redes neuronales es casi perfecto, trataremos de solventar el problema que surge al probar con nuevos datos mediante variaciones en los datos de entrada, con el objetivo de obtener un método eficiente.

Variable CS6 medida vs estimada - Red neuronal con extracción de características

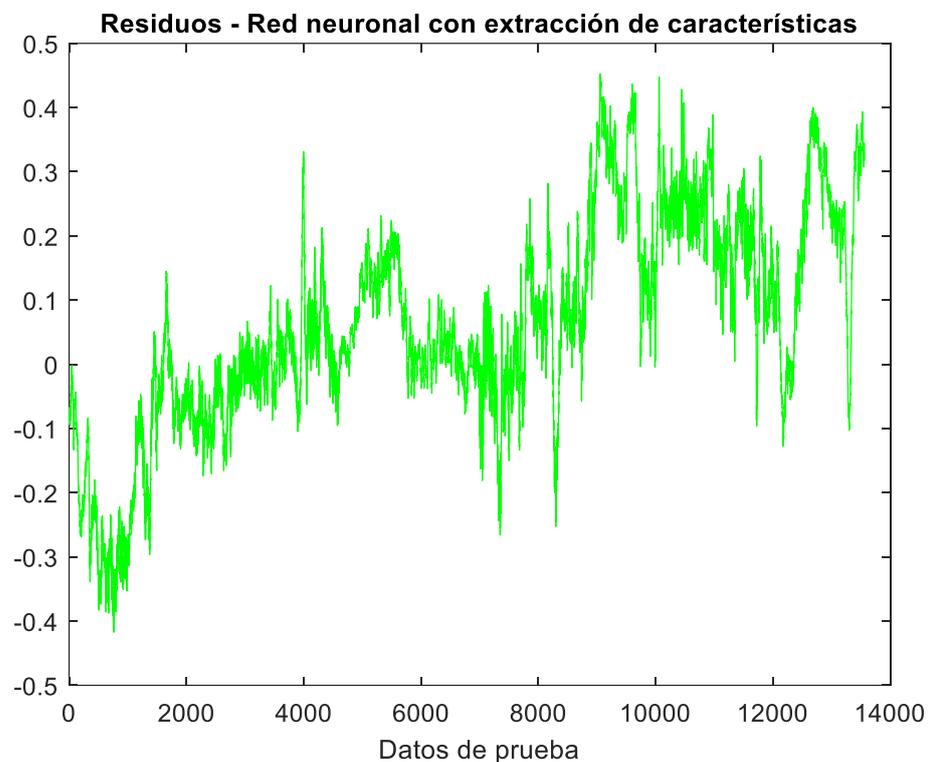


**Figura 4.52:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo Red Neuronal con extracción de características (prueba).



En la figura 4.52 podemos observar la estimación de la variable de calidad CS6 mediante el modelo de red neuronal con extracción de características frente a la variable medida de los datos reales de la planta. Se observa mejoría en el modelado de la variable de salida, aunque dista mucho de un resultado de modelo deseado, ya que aún no es preciso.

Esto lo podemos corroborar fijándonos en la gráfica de los residuos arrojados (figura 4.53) donde podemos ver que el error sigue siendo demasiado alto en comparación al error cometido al realizar el entrenamiento, el cual es prácticamente nulo.



**Figura 4.53:** Residuos obtenidos al aplicar la Red Neuronal con extracción de características (prueba).



### 4.4.3 RED DINÁMICA

En este apartado estudiaremos si aplicando retroalimentación a la red neuronal descrita con anterioridad, somos capaces de mejorar los resultados obtenidos con la misma. Esta red, que llamaremos red dinámica, utiliza información de la salida obtenida en el pasado, como una entrada más para calcular una nueva iteración de la nueva salida.

Para ello, crearemos una nueva matriz de predicción  $X$  siguiendo el mismo procedimiento que en PLS dinámico, es decir, añadiendo a la matriz  $X$  dos retardos de las variables de  $X$ , y un retardo de la variable de salida  $Y$ . Con esa nueva matriz, crearemos la red dinámica con cinco neuronas en la capa ocultas nuevamente, y entrenaremos la red con los nuevos datos.

El esquema de la red dinámica (retroalimentada) es el siguiente:

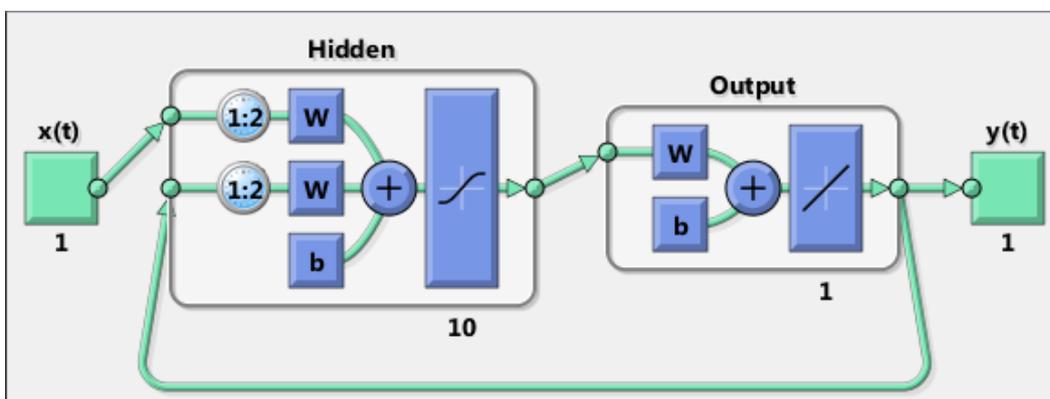


Figura 4.54: Esquema de la red dinámica.

En este caso, al contar con un número de datos de entrada mucho mayor, la red tarda notablemente más tiempo en alcanzar el modelo deseado, en nuestro caso, 125 iteraciones, como muestra la figura 4.55.

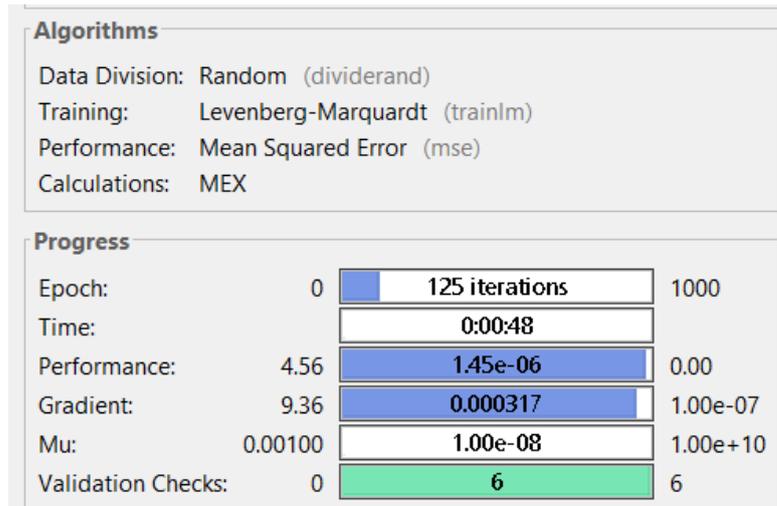


Figura 4.55: Número de iteraciones en la red dinámica hasta obtener el modelo.

Como podemos apreciar en la figura 4.56, el error medio cuadrático se ha reducido notablemente en el proceso de entrenamiento de la red dinámica, llegando al orden de  $10^{-6}$ , un error prácticamente inapreciable.

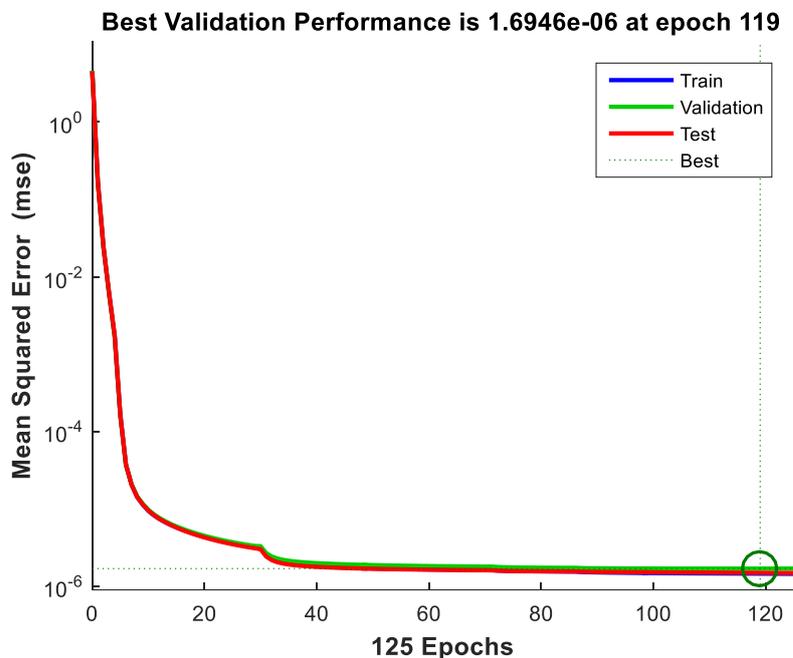
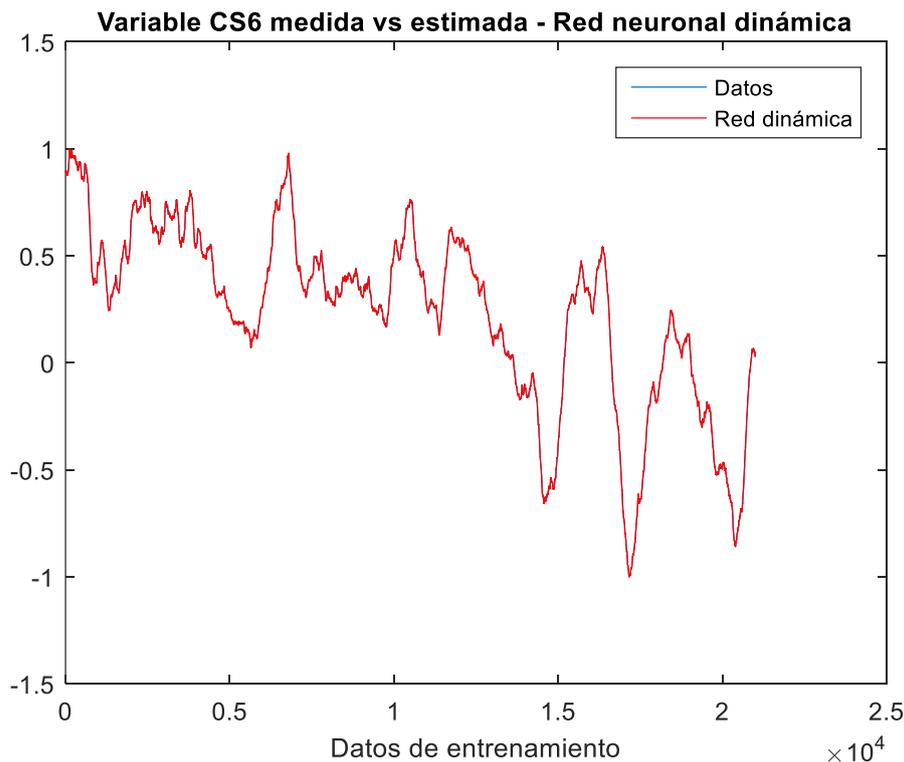


Figura 4.56: Error cuadrático medio por iteración en la red dinámica (entrenamiento).

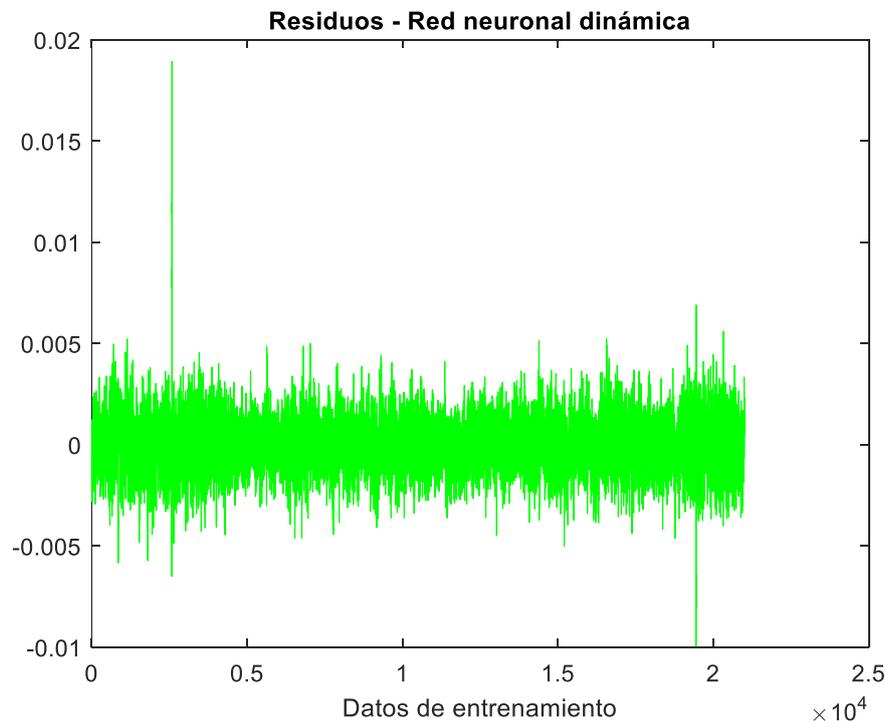


A la vista de esos resultados, podemos esperar, de nuevo, un modelo de entrenamiento prácticamente perfecta, como se muestra a continuación, donde la diferencia entre la estimación de la variable de calidad CS6 mediante la red dinámica y la propia variable tomada de los datos reales de la planta es inapreciable.



**Figura 4.57:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo Red Dinámica (entrenamiento).

Y, como era de esperar, los residuos arrojados por la red dinámica, es decir, la diferencia entre los valores reales y los valores modelados de la salida, son prácticamente nulos. De nuevo, hemos tenido que tomar una sección aumentada del plot para poder reconocer algún tipo de movimiento en torno al valor cero.



**Figura 4.58:** Detalle del plot de *residuos* obtenidos al aplicar la Red Dinámica (entrenamiento).

Sin embargo, como hemos comprobado con anterioridad en la Red Neuronal, estos resultados no significan nada si el método no es capaz de repetir un modelo tan bueno con nuevos datos. Por tanto, procederemos a probar la nueva red con los datos de prueba.

En las figuras 4.59 y 4.60 (residuos) puede observarse cómo esta vez el modelo creado por la red es muy bueno, unos resultados similares a los obtenidos en el entrenamiento, lo cual puede indicarnos que el hecho de utilizar como datos de entrada información obtenida en la salida en instantes anteriores, es decir, la retroalimentación, da lugar a una gran mejora en los resultados obtenidos.

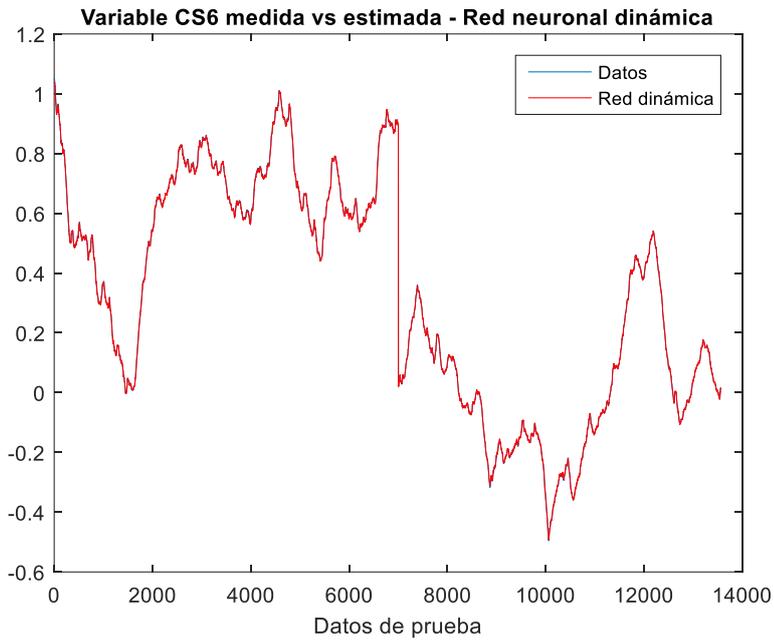


Figura 4.59: Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo Red Dinámica (prueba).

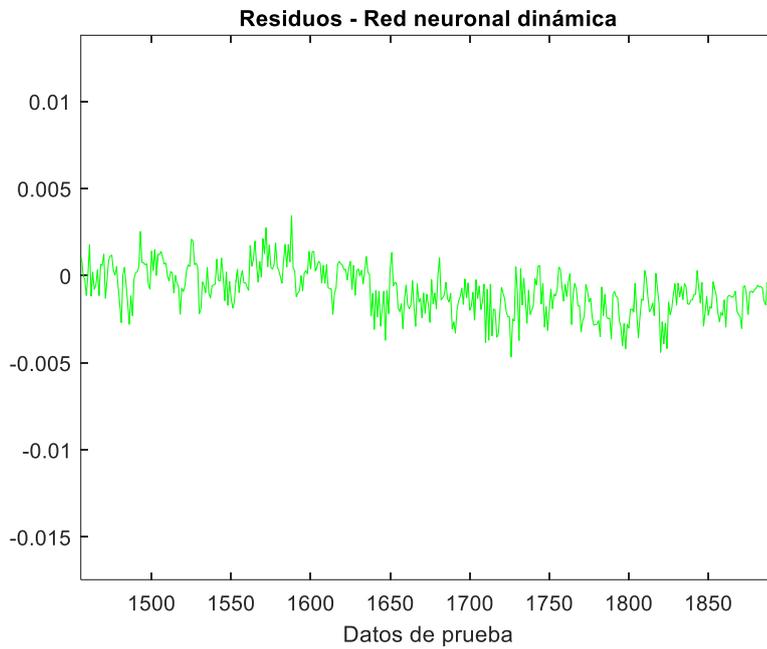


Figura 4.60: Detalle del plot de residuos obtenidos al aplicar la Red Dinámica (entrenamiento).



## 4.4.4 RED DINÁMICA CON EXTRACCIÓN DE CARACTERÍSTICAS

Tras obtener unos resultados satisfactorios utilizando la red dinámica, vamos a probar a extraer las variables más significativas utilizadas en la matriz de predicción, para crear un nuevo modelo con menor cantidad de información para ver si es capaz de modelar la variable de salida de forma apropiada.

Para ello, seguimos el procedimiento utilizado en apartados anteriores, que es el siguiente: Una vez obtenida la matriz dinámica con sus respectivas variables retardadas, utilizaremos extracción de características mediante el método de correlación para quedarnos con las variables de la matriz X más correladas con la salida Y. Tras ello, crearemos una red con cinco neuronas en la capa oculta, y procederemos a entrenarla. El resultado del entrenamiento fue, como en las redes anteriores, satisfactorio, y en tan solo 48 iteraciones (Figura 4.61).

El modelo no es tan bueno como en la red dinámica (Figura 4.62), pero hay que tener en cuenta que la cantidad de información con la que trabaja ahora es mucho menor, por lo que podemos afirmar que el modelo es bastante bueno.

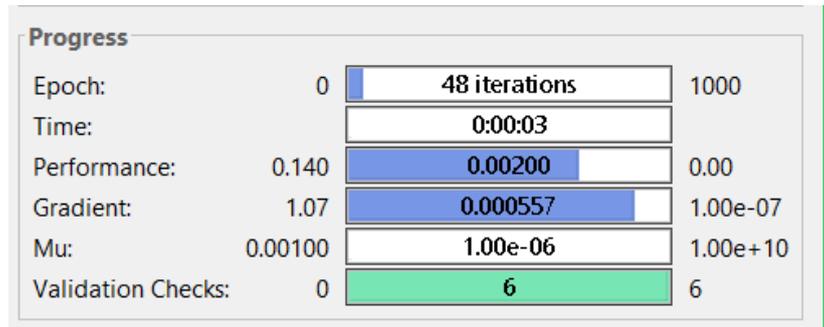
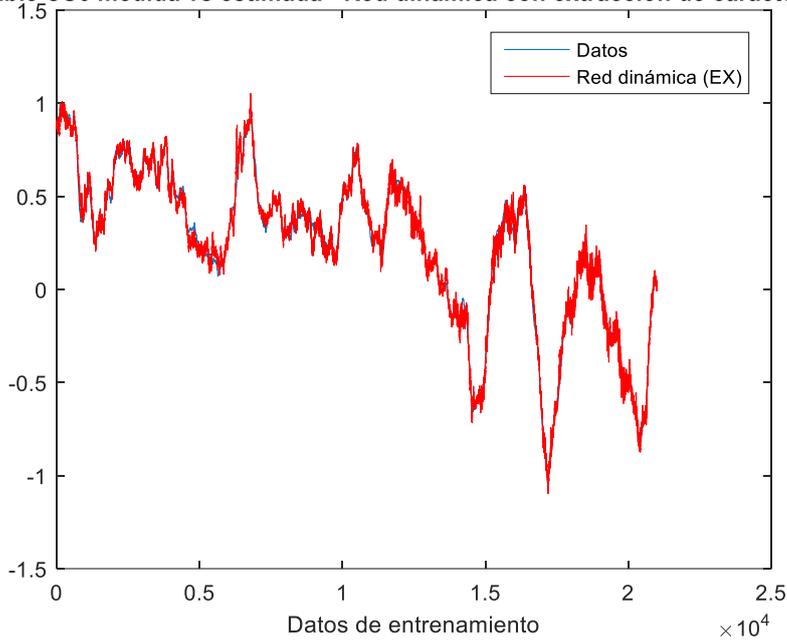


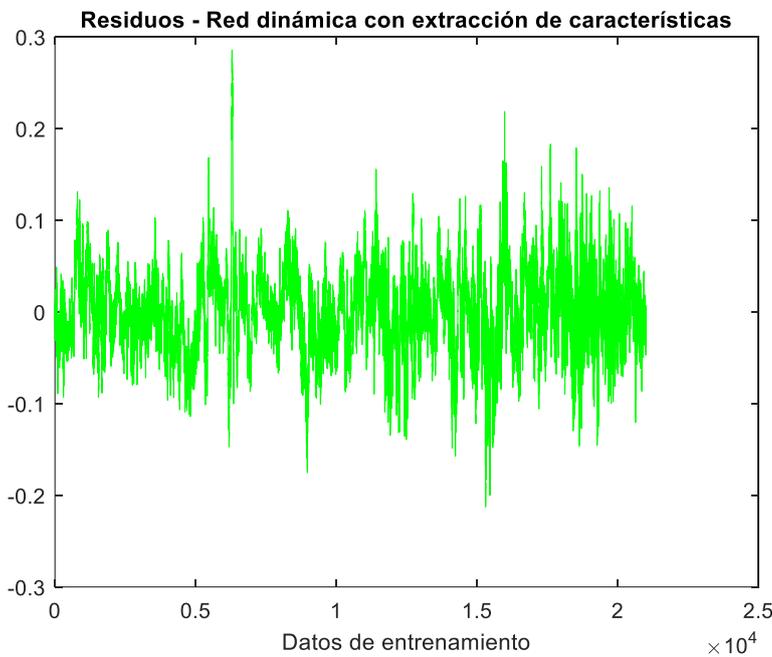
Figura 4.61: Número de iteraciones en la red dinámica con extracción de características hasta obtener el modelo.



**Variable CS6 medida vs estimada - Red dinámica con extracción de características**



**Figura 4.62:** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo Red Dinámica (EX) (prueba).



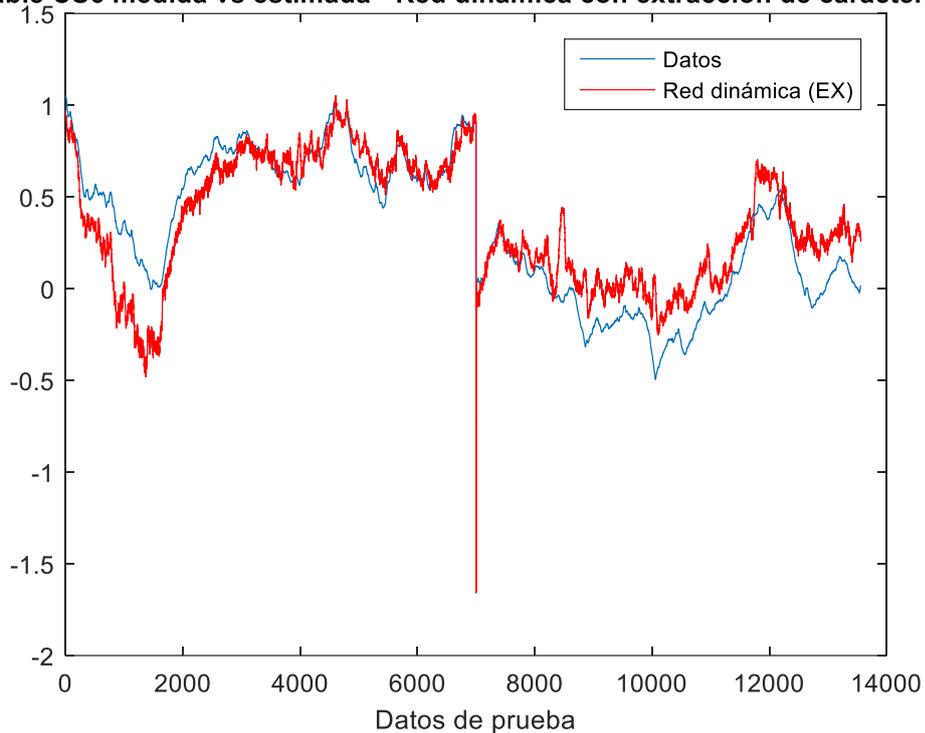
**Figura 4.63:** Residuos obtenidos al aplicar la Red Dinámica con extracción de características (entrenamiento).



Por último, probaremos el modelo en los nuevos datos para comprobar si se observa mejoría respecto a los modelos creados por las redes descritas con anterioridad. En la figura 4.64 vemos cómo efectivamente, el modelo es mucho mejor en este caso que en la Red Neuronal con extracción de características, aunque no podemos considerarlo como bueno, ya que contiene bastantes perturbaciones y no sigue con suficiente exactitud el patrón marcado por los datos reales.

Hay que tener en cuenta que el número de variables con las que trabaja es bastante reducido en comparación a la Red Dinámica, por lo que el modelo era de esperar que no fuese tan bueno, al contar con menor cantidad de información.

**Variable CS6 medida vs estimada - Red dinámica con extracción de características**



**Figura 4.64** Representación de la variable de salida CS6 medida, y de su estimación mediante el modelo Red Dinámica (EX) (prueba).



## 4.5 COMPARATIVA Y CONCLUSIONES

Una vez desarrollados y aplicados todos los métodos, llega el momento de hacer una comparación entre unos y otros para determinar cuál es el más conveniente para el diseño del sensor software.

### 4.5.1 COMPARATIVA PLS

En este apartado realizaremos la comparación entre los métodos PLS. Para poder compararlos, nos fijaremos en el análisis de los residuos de cada método, para obtener el error de estimación de cada uno de ellos. Los errores han sido calculados utilizando la raíz del sumatorio de los residuos al cuadrado ( $rssq$ ), como indica la siguiente fórmula:

$$rssq = \varepsilon = \sqrt{\sum_{i=1}^n \text{residuos}^2}$$

Método	$\varepsilon_{entr}$	$\varepsilon_{test}$	$\varepsilon_{fallo}$	$t_{det}$
PLS	37.0789	39.2111	39.5119	10149
PLS(EX)	35.8663	45.5132	51.9654	10149
DPLS	10.6928	11.6234	12.2118	10149
NPLS	5.3968	5.6793	5.9481	10149
NDPLS	5.1223	5.4007	5.4383	10149
DPLS(EX)	36.0729	44.4688	55.8295	10149
NPLS(EX)	3.6745	4.1026	4.6371	10149
NDPLS(EX)	3.0488	3.6738	8.4035	10149

Tabla 4.1: Error de estimación de cada método PLS (entrenamiento, prueba y prueba con fallo).



En la tabla 4.1 podemos ver los resultados obtenidos. A la vista de los resultados, el método que menor error genera, es decir, el método que mejor realiza el modelado de la variable de calidad que queremos predecir, es, a priori, el PLS dinámico no lineal con extracción de características.

Este método es el que cuenta con mayor cantidad de información, puesto que, aparte de la matriz de predicción  $X$  del método PLS, también cuenta con información pasada de la salida gracias a los retardos, cuenta con carácter no lineal gracias a las ecuaciones introducidas en su matriz de predicción, y además, la extracción de características elige a las variables más significativas y elimina las variables que no aportan información relevante, por lo que trabaja con una matriz de predicción óptima.

El tiempo de detección en todos los métodos ha sido el instante 10149. La explicación a esto es que siempre hemos introducido el error en el instante 10000, y para detectar el error, tanto el estadístico  $Q_x$  como el  $T^2$  deben detectar la anomalía durante 150 instantes consecutivos, y tras ello determinan un fallo como relevante y salta el detector en ese instante, cuando, en realidad, empezaron a contar en el instante 10000, que es exactamente el momento en el que el fallo fue provocado. A la vista de esto, podemos afirmar que el sistema de detección de fallos funciona en cada método con precisión.

En general, los métodos con carácter no lineal arrojan mucho menor error que los demás, tanto en entrenamiento como en test y test con fallo. El método dinámico también parece bastante bueno, sin embargo, al aplicar extracción de características, pierde notablemente su eficacia, y no es capaz de modelar la variable  $Y$  con esa reducción de información provocada por la extracción.



## 4.5.2 COMPARATIVA RNA

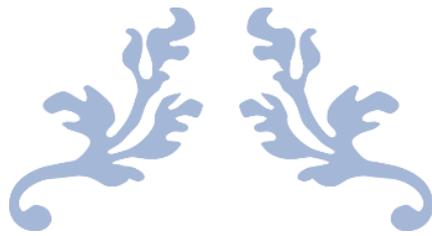
Ahora realizaremos la comparativa de las redes neuronales artificiales utilizadas. De la misma forma que el PLS, nos fijaremos en el análisis de los residuos de cada método, para obtener el error de estimación de cada uno de ellos. Los errores han sido calculados utilizando la raíz del sumatorio de los residuos al cuadrado ( $rssq$ ). En la tabla 4.2 podemos ver los resultados obtenidos.

Tipo de red	$\varepsilon_{entr}$	$\varepsilon_{test}$
Red Neuronal	4.7837	18.4240
Red Neuronal (EX)	8.4475	18.2329
Red Dinámica	0.1751	1.2677
Red Dinámica (EX)	6.4588	22.9229

Tabla 4.2: Error de estimación de cada red neuronal (entrenamiento y prueba).

Como podemos comprobar, ninguna de las redes neuronales arroja errores significativos en el proceso de entrenamiento. Sin embargo, a la hora de probar con nuevos datos, no son capaces de modelar la variable estimada  $Y$  correctamente, y el error obtenido es demasiado alto.

La excepción es la Red Dinámica, donde, tanto el error cometido en el entrenamiento como más adelante en el proceso de prueba, son muy bajos, los más bajos de todos los métodos que hemos utilizado, por lo que podemos elegir esta red como la única con la que a priori queremos trabajar en el diseño de un sensor software.



---

# CONCLUSIONES Y FUTURAS LÍNEAS DE TRABAJO

---

CAPÍTULO 5





## **5.1 CONCLUSIONES**

El objetivo principal de este trabajo ha sido diseñar un sensor software que fuera capaz de ayudar a lograr los objetivos de calidad establecidos en la fase de producción de sacarosa de una industria azucarera. Para ello, se han utilizado tanto técnicas de análisis estadístico (en nuestro caso, tanto un método basado en regresión por mínimos cuadrados parciales, como un método no lineal basado en redes neuronales artificiales) como herramientas de modelado de la variable de calidad que queremos medir.

En el caso de los métodos basados en PLS, se ha implementado un PLS genérico, a partir del cual se han realizado modificaciones para desarrollar otros métodos, como son: PLS dinámico, cuya matriz de datos tiene en cuenta la historia reciente de la planta; PLS no lineal, el cual introduce las relaciones no lineales de variación de la salida respecto de las entradas; y PLS con extracción de características, que consiste en trabajar como matriz de datos solamente con las variables más significativas, es decir, las que guardan mayor relación con la salida. A su vez, estos métodos han sido combinados entre ellos para estudiar si se puede obtener una mejoría en los resultados a partir de una combinación de dos o más métodos.

Se han observado diferencias entre las variables modeladas por cada uno de los métodos, siendo los que arrojaron un menor error los métodos que contaban con carácter no lineal en la matriz de datos utilizada para la predicción de la salida (variable de calidad). También hemos obtenido buenos resultados mediante la aplicación de los métodos dinámicos, ya que contaban con mayor cantidad de información de las variables, lo que ha dado lugar a una mejor predicción.

En un principio, el método teórico utilizado de extracción de características por medio de los scores obtenidos en la regresión PLS, no resultó satisfactorio. El modelo obtenido tras extraer



características mediante este método distaba mucho de los resultados esperados. Al observar esto, se decidió realizar un análisis de correlación entre las variables extraídas y la variable de salida, y se observó que, como era de esperar, las variables que habíamos extraído mediante este método no eran las que guardaban mayor correlación con la variable de salida, dando lugar a un modelo ineficiente. Por ello, se decidió modificar el método de extracción de características, utilizando esta vez únicamente las variables que guardasen mayor correlación con la variable de salida. De esta forma, solventamos el error, y los resultados mejoraron notablemente.

De todas formas, en la mayoría de los métodos, el hecho de extraer las variables más significativas ha conllevado a perder ligeramente la eficiencia del modelo, debido probablemente a la falta de información suficiente por parte de los datos con extracción de características. El único método en el cual se ha obtenido una mejora es, a su vez, el método que menor error ha arrojado en sus resultados. Este es el método PLS no lineal, dinámico, con extracción de características.

La conclusión es que este método es el más completo de todos, ya que su matriz de datos posee carácter no lineal, tiene en cuenta la historia reciente de la salida a la hora de utilizar nuevas entradas, y, además, dado que esta matriz tendría el problema de tener unas dimensiones demasiado grandes para trabajar con ellas, se han extraído las variables más significativas para trabajar únicamente con ellas, y se han mejorado los resultados obtenidos al trabajar con toda la matriz. Por lo tanto, sería el método a elegir para trabajar en un futuro.

Además de la estimación de la variable de calidad a través de los métodos descritos, se ha implementado un sistema de detección y diagnóstico de fallos, con el fin de detectar situaciones anómalas en las que pueda trabajar la planta, y, una vez detectado un posible fallo, que el propio programa sea capaz de informar acerca de la variable que está provocándolo, con el fin



de facilitar la tarea de reparación o recambio del elemento que está provocando el fallo.

Para comprobar la eficacia de este sistema de detección y diagnóstico, se han provocado fallos en diferentes variables del sistema con el objetivo de determinar si el programa era capaz de detectarlos y, en caso de que así fuera, diagnosticarlos correctamente. Cada uno de los métodos, en este caso, detectó y diagnosticó con precisión y rapidez cada fallo introducido, a pesar de ser estos muy leves, tanto que en ocasiones no se llegó a apreciar un cambio drástico en la variable de salida tras provocar el fallo. Es de suponer, por tanto, que un fallo real ocasionado en la planta, que sería de una magnitud mucho mayor, sea detectado sin problemas por el sistema.

En el caso de las redes neuronales artificiales, al ser estas de carácter no lineal, por lo comentado anteriormente, los modelos al entrenar datos salían prácticamente idénticos a los datos reales obtenidos de la planta. Sin embargo, en este caso, el método constaba de tanta cantidad de información que no era capaz de estimar la variable al aplicarlo con nuevos datos. Se probó a variar el número de neuronas en la capa oculta de la red, pero no se obtuvieron mejores resultados. La única red capaz de estimar la variable con nuevos datos de forma correcta fue la red dinámica, contrariamente a lo esperado, ya que el problema de la red neuronal había sido la excesiva cantidad de información proporcionada al método, y la red dinámica posee una cantidad mayor de información. La única explicación lógica posible es que el hecho de tener en cuenta la historia reciente de la salida para crear las nuevas entradas, corrija este error, sin importar que aumente el número de datos de entrada, ya que la información contenida en estos nuevos datos resultó útil para la estimación de la variable de salida.



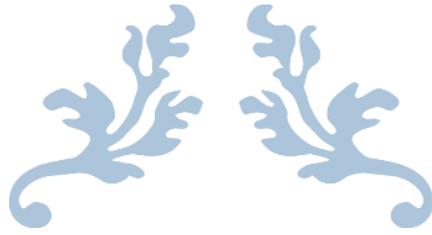
## 5.2 FUTURAS LÍNEAS DE TRABAJO

Los métodos desarrollados para la realización de este trabajo han sido aplicados, en nuestro caso, en la sección de evaporación de una planta azucarera real. Sin embargo, este software no está limitado a la industria azucarera, sino que puede ser implementado en cualquier tipo de industria cuyo objetivo sea el control de la calidad de un producto. Todo lo necesario para la aplicación de este software es una variable de calidad (variable de salida) que se quiera medir y controlar, y las variables incluidas a lo largo del proceso de producción (variables de entrada).

Como posibles futuras mejoras a este software, pueden desarrollarse métodos basados en el análisis de correlación canónica (CVA, *canonical variate analysis*), que es otro método de análisis multivariante de datos. Su principal diferencia con los métodos utilizados es que este método puede ser utilizado para predecir las variables de estado del sistema a partir de múltiples variables independientes.

En el caso de las redes neuronales artificiales, el problema radicaba en que la cantidad de información utilizada para entrenar la red era demasiada para lo que acostumbra a utilizar una red neuronal, por lo que al aplicarla a nuevos datos no era capaz de estimar la variable de manera correcta. Tampoco se logró una gran mejoría en la estimación de la variable al extraer características mediante el método de correlación. Una posible solución a este problema puede ser cambiar de nuevo el método de extracción de características, o utilizar el método de correlación que hemos utilizado anteriormente, pero modificándolo para hacerlo más estricto, es decir, que solo las variables con un coeficiente de correlación especialmente alto sean las variables utilizadas para trabajar en la matriz de datos.





---

# BIBLIOGRAFÍA

---

CAPÍTULO 6



**BIBLIOGRAFÍA**

- [1] Alves, R. (2005). Entornos distribuidos para simuladores de procesos. PhD thesis, Universidad de Valladolid
- [2] Blanke et al 2003 => Blanke, M., Kinnaert, M., Lunze, J., and Staroswiecki, M. (2003).  
Diagnosis and Fault Tolerant Control. Springer.
- [3] Carro Paz, Roberto y González Gómez, Daniel, “Control estadístico de procesos”. Administración de las Operaciones. Facultad de Ciencias Económicas y Sociales. Universidad Nacional de Mar del Plata.
- [4] Cuevas González, S. y Mediavilla Paniagua, H. Capítulo 2, “Redes Neuronales Artificiales”. Proyecto de Fin de Carrera. Departamento de Ingeniería de Sistemas y Automática. Universidad de Valladolid.
- [5] De la Fuente Aparicio, M<sup>a</sup>. J. (2009). Detección y diagnóstico de fallos.
- [6] García Álvarez, Diego (2013), “Monitoring, fault detection and estimation in processes using multivariate statistical techniques”. Tesis Doctoral. Departamento de Ingeniería de Sistemas y Automática. Universidad de Valladolid.
- [7] Geladi P. and Kowalki B. R. (1986). Partial Least-Squares Regression: A Tutorial.
- [8] Isasi Viñuela, Pedro y Galván León, Inés M. (2004), “Redes de Neuronas Artificiales, un enfoque práctico”. Ed. Pearson Prentice Hall.
- [9] Minsky and Papert, 1960 => Minsky M., and Papert. S, 1972 (2nd edition with corrections, first edition 1969) *Perceptrons: An Introduction to Computational Geometry*, The MIT Press, Cambridge MA, ISBN [0-262-63022-2](https://doi.org/10.1017/CBO9780511524804).



- [10] Muñoz González, Samuel (2010), “Diagnosis automática de fallos en una planta real mediante PLS”. Proyecto de Fin de Carrera. Departamento de Ingeniería de Sistemas y Automática. Universidad de Valladolid.
- [11] Nomikos and MacGregor, 1995 => Nomikos, P. and MacGregor, J. (1995). Multivariable SPC Charts for Monitoring Batch Processes. *Technometrics*, 37(1):41–59.
- [12] Puigjaner, L., P. Ollero, C. de Prada y L. Jiménez. (2006), Detección y diagnóstico de fallos, en “Estrategias de modelado, simulación y optimización de procesos químicos”. Ed. Síntesis.
- [13] S. Oakland, John and F. Followell, Roy (1990), “Statistical Process Control, a practical guide”. Ed. Butterworth-Heinemann Ltd.

### **BIBLIOGRAFÍA EN FORMATO DIGITAL**

- [14] Del Valle, Diego. Director Agrícola de Azucarera Ebro. “El sector remolachero/azucarero en España”. Documento informático en PDF.  
[http://upa.es/anuario\\_2007/pag\\_170-175\\_valle.pdf](http://upa.es/anuario_2007/pag_170-175_valle.pdf)
- [15] García-Álvarez, Diego; Merino, Alejandro; Martí, Rubén; de la Fuente, M<sup>a</sup>. Jesús (1998), “Soft sensor design for dry substance content estimation in the sugari industry”. Sugar Technology. Bartens
- [16] Weisstein, Eric W. “Correlation Coefficient”. Wolfram MathWorld. Wolfram Research.
- [17] “GUÍA DE MEJORES TÉCNICAS DISPONIBLES EN ESPAÑA DEL SECTOR AZUCARERO”. Ministerio de Medio Ambiente, Ministerio de Agricultura Pesca y Alimentación, documento informático en PDF.



- [18] “Resumen del procesado de la remolacha en una factoría azucarera”. Texto web:  
<http://acor.es/html/EL%20AZUCAR/EXTRACCION.htm#RECEPCIÓN>
- [19] “Acor y Azucarera, luchan ahora sin cuotas”. Artículo web:  
[http://www.qcom.es/v\\_portal/informacion/informacionver.asp?cod=27823&te=2&idage=30572](http://www.qcom.es/v_portal/informacion/informacionver.asp?cod=27823&te=2&idage=30572)

**IMÁGENES OBTENIDAS EN DIVERSAS PÁGINAS WEB**

- [20] [study.com](http://study.com)
- [21] [www.ibiblio.org](http://www.ibiblio.org)
- [22] [ceres.ugr.es](http://ceres.ugr.es)
- [23] [github.com](http://github.com)
- [24] [tex.stackexchange.com](http://tex.stackexchange.com)
- [25] [www.infoagro.com](http://www.infoagro.com)
- [26] [www.diariodeleon.com](http://www.diariodeleon.com)