



Universidad de Valladolid



**ESCUELA DE INGENIERÍAS
INDUSTRIALES**

UNIVERSIDAD DE VALLADOLID

ESCUELA DE INGENIERIAS INDUSTRIALES

Grado en Ingeniería en Tecnologías Industriales

Diseño de sensores software para el control de calidad de un proceso

Autor:

Güemes García, Enrique

Tutor:

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

Valladolid, Junio de 2018



RESUMEN

Actualmente en cualquier proceso de fabricación la calidad de un producto se presenta como una característica indispensable. Para comprobar el cumplimiento de dicha característica todo proceso industrial está expuesto a un seguimiento continuo de sus parámetros. Este trabajo se plantea con el objetivo principal de diseñar un sensor software capaz de estimar diferentes variables en un control de calidad en una planta depuradora de aguas residuales en función de los valores de otras variables. Para ello se aplicarán diferentes métodos basados en regresión lineal multivariable y así poder evaluar cuál ofrece una mejor predicción. Adicionalmente, se introducirá un método de detección de fallos, basado en PCA (Principal Component Analysis) que permitirá detectar si existe un fallo en la planta, el instante y las variables responsables de dicho fallo.

PALABRAS CLAVE

Sensor Software

Calidad

Regresión Lineal Multivariable

Predicción

Detección y Diagnóstico de fallos



Índice de contenidos



Índice

Índice de contenidos	7
1. Introducción y objetivos	11
1.1 Introducción.....	13
1.2 Objetivos	14
1.3 Organización de la memoria.....	14
2. Estadística aplicada a un proceso de calidad	17
2.1 Introducción.....	19
2.2 Sensor software	20
2.2.1 Introducción	20
2.2.2 Tipos de sensores software	21
2.2.3 Desarrollo de un sensor software	21
2.2.3.1 Primera inspección de datos.....	21
2.2.3.2 Selección de datos históricos e identificación de estados	22
2.2.3.3 Preprocesamiento de los datos	22
2.2.3.4 Selección de modelo, entrenamiento y validación	22
2.2.3.5 Mantenimiento del sensor software	23
2.3 Regresión lineal multivariable.....	23
2.3.1 Mínimos Cuadrados (LS).....	25
2.3.2 Método Ridge	27
2.3.3 Método Lasso	27
2.3.4 Método PLS (Partial Least Squares).....	29
2.3.5 PCA (Análisis de Componentes principales).....	30
2.3.6 Método dinámico	32
2.4 Calidad de un proceso.....	33
2.4.1 Evolución histórica.....	33
2.4.2 Dimensiones de la calidad	34
2.4.3 Gestión de la calidad	36
2.4.4 Aplicación de Métodos Estadísticos a la calidad	37
2.4.5 Procesos de inspección	38
2.4.6 Normativa sobre calidad	40
2.5 Variabilidad de un proceso	41
2.5.1 Causas comunes	42
2.5.2 Causas especiales.....	43
2.6 Detección de fallos	44

2.6.1	Introducción	44
2.6.2	Métodos de detección de fallos.....	44
2.6.2.1	Métodos basados en señal	45
2.6.2.2	Métodos analíticos	45
2.6.2.3	Métodos basados en conocimiento.....	46
2.6.3	Requerimientos de diseño	47
2.6.4	Control estadístico de procesos (SPC) para detección de fallos	48
2.6.5	Detección de fallos usando PCA.....	50
3.	Depuradora de aguas residuales	53
4.	Resultados	63
4.1	Introducción.....	65
4.2	Toma de datos de la planta	65
4.3	Preprocesamiento de los datos	66
4.4	Estimación de la variable DQO en el efluente.....	67
4.4.1	Modelo estático.....	67
4.4.1.1	Método de mínimos cuadrados.....	67
4.4.1.2	Método Ridge	70
4.4.1.3	Método Lasso	73
4.4.1.4	Método PLS.....	78
4.4.1.5	Método PCA	80
4.4.1.6	Combinación de todos los métodos	82
4.4.1.7	Comparativa entre métodos aplicados al modelo estático	85
4.4.2	Modelo dinámico.....	87
4.4.2.1	Método de mínimos cuadrados.....	87
4.4.2.2	Método Ridge	89
4.4.2.3	Método Lasso	91
4.4.2.4	Método PLS.....	95
4.4.2.5	Método PCA	98
4.4.2.6	Combinación de todos los métodos	100
4.4.2.7	Comparativa entre métodos aplicados al modelo dinámico	103
4.4.3	Conclusiones	104
4.5	Detección y diagnóstico de fallos.....	106
4.5.1	Introducción	106
4.5.2	Pasos previos	106
4.5.3	Estudio del fallo	107
4.5.3.1	Modelo estático	107
4.5.3.1.1	Predicción por PCA	107



4.5.3.1.2 Detección y diagnóstico con estadístico T^2	108
4.5.3.1.3 Detección y diagnóstico con estadístico Q	110
4.5.3.2 Modelo dinámico	112
4.5.3.2.1 Predicción por PCA	112
4.5.3.2.2 Detección y diagnóstico con estadístico T^2	113
4.5.3.2.3 Detección y diagnóstico con estadístico Q	115
4.5.4 Conclusiones	117
5. Conclusiones y futuras líneas de trabajo	121
5.1 Conclusiones.....	123
5.2 Futuras líneas de trabajo.....	124
6. Bibliografía.....	127



Introducción y objetivos



1. INTRODUCCIÓN Y OBJETIVOS

1.1. Introducción

La calidad de un producto o de un proceso industrial está constituida por el conjunto de propiedades y características del mismo, que le confieren su aptitud para satisfacer las necesidades establecidas o implícitas. También puede definirse la calidad como la idoneidad o utilidad del producto respecto de la aplicación para la que ha sido previsto.

La calidad se puede contemplar desde tres puntos de vista. En primer lugar desde la perspectiva del cliente, que es la calidad deseada por el cliente en función de que un producto cumpla sus necesidades y expectativas, en segundo lugar se encuentra la calidad de proyecto o de diseño, que es la medida en que las características establecidas en el mismo logran que el producto diseñado proporcione esa utilidad, y por último, la calidad de fabricación o de concordancia, que hace referencia a que el producto real fabricado satisfaga las especificaciones del proyecto.

Con el objetivo de lograr la calidad deseada, todo proceso industrial está expuesto a un seguimiento de sus parámetros para poder comprobar que dicho proceso funciona bajo lo establecido o, por lo contrario, poder detectar anomalías o fallos que se produzcan. Sin embargo, no todos los parámetros se pueden medir u obtener fácilmente, ya sea porque a veces no existen medidores comerciales para esas variables o porque son muy caros y difíciles de mantener. Es por esta razón por la que se considera necesario el uso de los sensores software. El uso de esta técnica nos permite evitar pérdidas económicas en la inversión de costosos aparatos de medición y en productos mal fabricados tras no haber podido analizar el fallo rápidamente.

Estos sensores software nos permiten predecir los valores de variables difíciles de medir, basando su estudio en datos predictores (otras variables) que sean relativamente fáciles de obtener y nos permita alcanzar un valor del dato objetivo. Hay dos tipos fundamentalmente de sensores software, los basados en el conocimiento y los basados en datos. Los basados en el conocimiento, o también llamados modelos de caja blanca, son sensores que se basan en las leyes físicas que rigen el comportamiento del proceso. En cambio los sensores basados en datos, o también llamados modelos de caja negra, se basan únicamente en los valores que se disponen a la entrada y a la salida del sistema. Para implementar estos sensores se pueden utilizar diferentes técnicas como es el caso de la regresión lineal múltiple. En este último tipo de sensores se va a basar este trabajo.

La seguridad en la producción es otro 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 fallos 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.

Las técnicas estadísticas multivariantes son probablemente una 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, y será la técnica usada en este trabajo para detectar fallos en la planta.

Un sensor software tiene como principal ventaja que nos permite estudiar un proceso a lo largo de sus diferentes etapas, no solo a la entrada y a la salida. Esto nos permite evaluar cómo evoluciona el proceso y, a través de las técnicas de detección y diagnóstico de fallos poder encontrar fácilmente un fallo y en qué etapa se encuentra, para así poder tomar las medidas pertinentes para solucionarlo y recuperar el estado normal de nuestro proceso.

1.2. Objetivos

Este Trabajo de Fin de Grado tiene como objetivo principal el diseño de un sensor software que sea capaz de estimar las diferentes variables en un control de calidad en una planta depuradora de aguas residuales en función de los valores de otras variables. Para ello se aplicarán diferentes métodos basados en regresión lineal múltiple con la finalidad de observar la respuesta que nos ofrece cada uno y así poder evaluar cuál nos proporciona la mejor solución. Adicionalmente, se introducirá un método de detección de fallos, basado en PCA, que permitirá detectar fallos en el proceso. De esta forma el sensor software además de darnos el valor actual de la variable que se esté calculando, deberá decir si está trabajando en una situación anómala. También deberá proporcionar la información del instante en el que se produjo el fallo y cuál es la variable que lo está provocando. Todo esto nos proporcionará una información de gran valor de cara a reparar el fallo y volver a trabajar con normalidad, asegurando la calidad del sistema.

1.3 Organización de la memoria

Esta memoria está organizada de la siguiente forma:

- **Capítulo 2: Métodos de estudio.** En este capítulo se explica de forma teórica los diferentes métodos que se van a llevar cabo para el diseño de un sensor software, así como el método de detección y diagnóstico de fallos que se va a emplear en este trabajo.



- **Capítulo 3: Depuradora de aguas residuales.** Se describe el funcionamiento de una planta depuradora de aguas residuales así como sus etapas, y el modelo de planta sobre el que se va a realizar el estudio.
- **Capítulo 4: Resultados.** Se exponen los resultados obtenidos en la depuradora de aguas residuales tras aplicar cada uno de los métodos anteriormente expuestos, comparando los resultados hallados inicialmente en la fase de entrenamiento, así como los hallados posteriormente en la fase que denominamos test.
- **Capítulo 5: Conclusiones y futuras líneas de trabajo.** Se presentan en este capítulo las conclusiones finales que se han alcanzado tras la realización del trabajo, así como las diferentes posibilidades con las que continuar la investigación.
- **Bibliografía.** Se muestran las referencias bibliográficas que se han utilizado para hacer este trabajo.



Estadística aplicada a un proceso de calidad



2. ESTADÍSTICA APLICADA A UN PROCESO DE CALIDAD

2.1 Introducción

Actualmente, en cualquier tipo de proceso de fabricación automatizado es cada vez más relevante la aplicación de las características de fiabilidad, disponibilidad, seguridad y protección medioambiental. Esto es debido a un aumento de la demanda en productos con alta calidad. Para poder alcanzar estos objetivos se presenta necesario la elaboración de sistemas modernos de control más complejos y con algoritmos de control más sofisticados.

Los controladores clásicos permiten controlar que los procesos operen satisfactoriamente compensando los efectos de las perturbaciones y los cambios que ocurren a lo largo del proceso. Sin embargo hay controladores que no pueden manejar adecuadamente ciertos cambios, por lo que es necesario la elaboración de sistemas de supervisión más sofisticados que permitan aumentar la eficiencia, mantenibilidad, disponibilidad y fiabilidad de la planta.

Para poder lograr el cumplimiento de los objetivos mencionados relativos a la fiabilidad de los sistemas de control, es necesario dotar a dichos sistemas de un esquema de monitorización con sensores software capaces de controlar la calidad del proceso así como detectar la aparición de diferentes fallos o el inicio de la degradación de un proceso. Esto se presenta necesario para evitar posibles daños o pérdidas en el producto final, con la finalidad de que los componentes dañados o que producen el fallo sean reemplazados pudiendo así reducir las pérdidas posibles lo antes posible. Por lo tanto, la monitorización debe asegurar el éxito de las operaciones planeadas detectando de forma eficaz las anomalías en su comportamiento.

Se considera que el valor de una característica medida es óptimo cuando se encuentra dentro de un rango de tolerancias. Es imposible que una característica tenga siempre el mismo valor considerándose la variabilidad como algo inherente a los procesos. Por lo tanto, eliminar al completo la variabilidad de un proceso resulta imposible por lo que se presenta como objetivo fundamental de la calidad producir productos con la mínima variación posible. Para llevar a cabo esto se debe encontrar las fuentes de variabilidad de los procesos y eliminarlas para obtener productos con alta calidad.

La variabilidad se puede producir por dos causas principalmente, causas comunes y causas especiales. Las causas comunes son aquellas relacionadas directamente con el proceso, mientras que las especiales son aquellas que se consideran ajenas al desarrollo natural del proceso. Es importante saber identificarlas, ya que las causas especiales deben ser eliminadas para ser capaces de anticiparnos al resultado únicamente afectado por causas comunes (aleatorias) de variación. De esta forma, se podrá analizar si el proceso está dentro del rango de su variabilidad aleatoria, o por lo contrario, está fuera de control produciendo errores asignables a una causa determinada.

2.2 Sensor software

2.2.1 Introducción

A la hora de definir qué es un sensor software podemos empezar por analizar cada palabra por separado.

La palabra sensor nos hace referencia a un dispositivo que está capacitado para detectar acciones o estímulos externos transformando magnitudes físicas o químicas en magnitudes eléctricas. En definitiva, podemos decir que un sensor es un artefacto que permite al ser humano obtener información del entorno e interactuar con ella.

La palabra software se refiere al equipamiento lógico e intangible de una computadora digital que comprende el conjunto de los componentes necesarios para hacer posible la realización de tareas específicas; en contraposición a los componentes físicos del sistema, llamados hardware.

Por lo tanto, podemos decir que un sensor software es una herramienta que tras la inclusión de unos datos correctamente tomados es capaz de predecir el comportamiento de variables relacionadas o afectadas por esos datos y poder sacar óptimas conclusiones al realizar experimentos de forma virtual con éstos. La instalación de este sistema nos presenta el problema de la posible falta de fiabilidad de información a la hora de medir ciertos parámetros, pero también nos elimina algunos problemas frecuentes de los habituales sistemas de medida en general. Estos son algunos de estos problemas:

- El tiempo total necesario.
- El gasto producido debido a la realización física del estudio.
- Posibles fallos del hardware (sondas con errores de medición que se vean afectadas por otras variables).
- Problemas de mantenimiento.

Actualmente son utilizados en un amplio rango de actividades, pudiéndonos encontrar sensores de pH, humedad, sensores de biomasa, sensores de dióxido de carbono o sensores de temperatura.

Concretamente en nuestro campo de estudio, una planta de aguas residuales, se presenta necesario la automatización del proceso. Debido a las características biológicas del proceso de tratamiento de lodos activados, es difícil medir distintas variables por medio de sensores en línea, ya sea o porque no existen medidores comerciales para esas variables o tienen un alto coste siendo difíciles de mantener. En estos casos los datos se obtienen en el laboratorio mediante análisis químicos en los que se puede sacar un único dato al día o hasta cada semana. Al tener una frecuencia tan baja, estas variables no pueden ser utilizadas para el control y monitorización de la planta en la línea.

Muchas de las variables necesarias para el control y la monitorización de la planta no se pueden medir en línea, añadiendo además la condición de la poca fiabilidad a la hora de realizar la medida. Por lo tanto, se considera necesario el desarrollo de sensores software que nos permita reemplazar a los sensores hardware o utilizarse en paralelo para proporcionarnos una comparativa de resultados, y así comprobar si los sensores hardware realizaran algún intervalo de tiempo de mediciones erróneas.

2.2.2 Tipos de sensores software

Se pueden distinguir principalmente dos tipos de sensores software, los llamados sensores basados en conocimiento y los sensores basados en datos. Los primeros son también conocidos como modelos de caja blanca porque se conocen todas las leyes físicas y químicas que rigen el comportamiento del proceso. En cambio, los sensores basados en datos son conocidos como los modelos de caja negra porque no se tiene conocimiento sobre el proceso basándose únicamente en las observaciones empíricas del mismo. Entre estos dos tipos hay muchas combinaciones posibles denominándose modelos híbridos, o también modelos de caja gris.

Los modelos basados en conocimiento están desarrollados principalmente con el propósito de la planificación y el desarrollo de las plantas de proceso, basándose en las ecuaciones que describen los principios físicos y químicos. Muchas veces este tipo de sensor se rechaza porque requiere un gran conocimiento experto sobre el proceso, lo que no siempre está disponible.

Los modelos basados en datos han emergido como un tipo de modelo muy atractivo por su gran apoyo a la hora de realizar un diagnóstico, pronóstico y decisión sobre los métodos disponibles para el operador. Estos modelos están basados en medidas reales, las cuales son recogidas y grabadas en el registro histórico de datos de cada planta.

2.2.3 Desarrollo de un sensor software

A continuación en este apartado se explica la metodología que se sigue a la hora de desarrollar un sensor software como se ve en la Figura 2.1 (Kadlec et al., 2009).

2.2.3.1 Primera inspección de datos

El objetivo de este primer paso es obtener una visualización del conjunto de la estructura de los datos e identificar algunos problemas obvios que deberían ser eliminados en esta etapa inicial, como por ejemplo desechar algunas variables que tengan un valor constante. A la hora de realizar el modelo debería ser siempre evaluado y comparado con modelos alternativos en posteriores fases del desarrollo. Se debe prestar especial atención en el modelo la variable objetivo, que debe comprobarse si puede ser modelada.

2.2.3.2 Selección de datos históricos e identificación de estados históricos

En esta etapa, se seleccionan los datos que van a ser utilizados para el entrenamiento y la evaluación del modelo. Las partes estacionarias de los datos deben ser identificadas y seleccionadas. En la mayoría de los casos la modelización del proceso solo será correcta con los estados estacionarios del proceso, centrándose principalmente en los datos más representativos.

2.2.3.3 Preprocesamiento de los datos

Esta etapa es importante a la hora de desarrollar el sensor software, ya que es el paso que requiere de una gran cantidad de trabajo manual y conocimiento de un experto. Esta fase de preprocesamiento está realizada generalmente de una forma iterativa con el objetivo de procesar los datos hasta que el desarrollador del modelo considera que éstos están preparados para el entrenamiento y la evaluación del modelo actual. Un típico ejemplo de preprocesamiento de los datos es la normalización de los datos a media cero y varianza uno, debido a que esta transformación de los datos es necesaria para la aplicación de algunos métodos como el PCA.

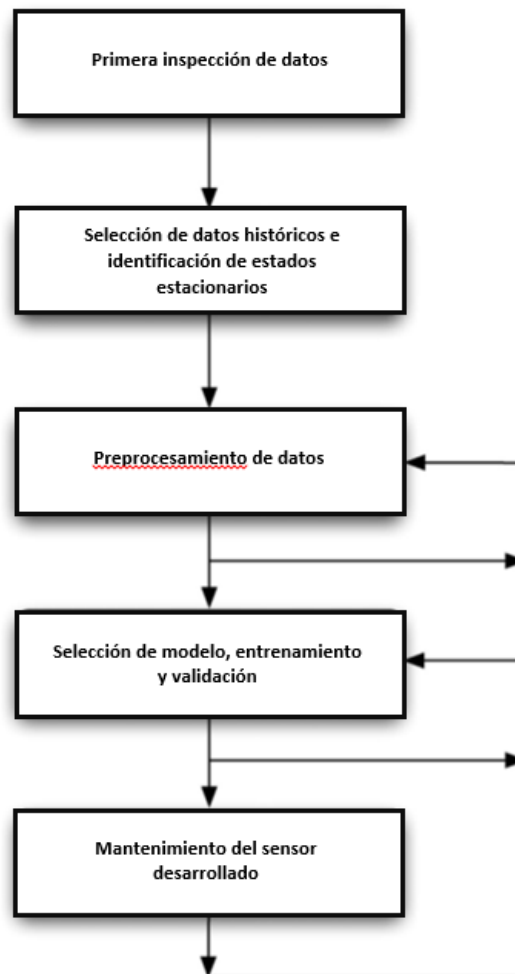


Figura 2.1. Pasos en el desarrollo de un sensor software

2.2.3.4 Selección de modelo, entrenamiento y validación

Para el resultado final del sensor software esta etapa es considerada la más crítica porque el resultado se va a ver muy influenciado por el tipo del modelo elegido, considerándose esta etapa crucial en el desarrollo del sensor software. Esta etapa puede resultar complicada debido a que no hay una teoría unificada para la elección del modelo, ya que muchas veces se elige de forma aleatoria. Muchas veces esta elección se realiza por la experiencia o la preferencia personal del desarrollador, lo que puede convertirse en una desventaja para el sensor final.

A pesar de la ausencia de un enfoque común teórico a la hora de elegir el modelo, hay algunas técnicas que se pueden abordar para llevar a cabo esta importante decisión. Una posibilidad es comenzar con una estructura o tipo de modelo simple,



e ir gradualmente incrementando la complejidad del modelo siempre que pueda ser observada una evolución en el desarrollo.

Una vez que se han aplicado diferentes modelos al proceso, se puede realizar una combinación de los diferentes métodos asignando más valor porcentual a un método o el mismo valor porcentual a cada método. De esta forma se obtiene otro modelo que muchas veces nos proporciona grandes resultados sobre la predicción final.

Después de encontrar la estructura óptima del modelo y haber entrenado, el sensor software debe ser evaluado con datos independientes a los de entrenamiento. Hay varias herramientas posibles para evaluar el modelo, siendo el error mínimo cuadrado (MSE) la evaluación numérica más popular. Consiste en comparar el promedio de la diferencia al cuadrado entre los valores predichos y los correctos. Otra forma de analizar la respuesta obtenida es a través de una representación visual de las predicciones. Una desventaja de este método es que la valoración final es algo subjetiva, ya que no proporciona valores matemáticos, solo interpretaciones visuales.

2.2.3.5 Mantenimiento del sensor software

Después de desarrollar el sensor software, éste debe ser mantenido y sintonizado debido a las variaciones y otros cambios que pueden sufrir los datos. Es necesario debido a que los sensores software se deterioran con el paso del tiempo y deben ser compensados para adaptarse a los cambios que puedan presentar los modelos.

Actualmente la mayoría de los sensores software no presentan ningún mecanismo automatizado para su automantenimiento, ya que incluir esto representa un coste significativo para la aplicación. El mantenimiento se presenta imprescindible a los cambios de datos o de etapas de un proceso. Uno de los problemas principales que se nos presenta es que no hay una medida objetiva para evaluar la calidad del sensor, y comprobar su correcto funcionamiento, ya que depende de una valoración subjetiva de una persona sobre una interpretación visual. Como principal solución para esta fase se recomienda hacer una adaptación continua o periódica de los principales componentes base.

A pesar de los métodos existentes para la adaptación de los sensores software automatizados, el operador del modelo todavía juega un importante rol con su juicio y su conocimiento a la hora de decidir qué métodos son utilizados para la adaptación y realizar un correcto mantenimiento.

2.3 Regresión lineal multivariable

Los tipos de sensores software que se van a aplicar en este trabajo son aquellos sensores basados en datos, y en particular, los métodos de estudio que se van a aplicar están basados en sistemas de regresión lineal multivariable (MLS). Estas técnicas son las más conocidas y aplicadas del análisis multivariable. La regresión

lineal múltiple se aplica tanto a datos de corte transversal, es decir, observaciones referidas a un mismo momento del tiempo, como a datos de series temporales. En el contexto de análisis multivariable los datos que se utilizan suelen ser predominantemente de corte transversal, basándose dicho modelo en diferentes hipótesis básicas y propiedades probabilísticas para poder realizar inferencias.

El modelo de regresión múltiple que se va a presentar se considera que la variable predicha es una función lineal de k regresores y de una perturbación aleatoria, existiendo además un regresor ficticio correspondiente al término ficticio. Designando por Y_t a la predicción, por $X_{1t}, X_{2t}, \dots, X_{kt}$ a los regresores y por u_t a la perturbación aleatoria, el modelo teórico de regresión lineal viene dado, para la observación genérica t -ésima, por la siguiente expresión:

$$Y_t = b_1 \cdot X_{1m} + b_2 \cdot X_{2m} + \dots + b_n \cdot X_{nm} + u_n \quad m = 1, 2, \dots, T$$

Siendo T el tamaño de la muestra y dando valores a m desde $m=1$ hasta $m=T$, se obtiene el siguiente sistema de ecuaciones:

$$Y_1 = b_1 \cdot X_{11} + b_2 \cdot X_{21} + \dots + b_n \cdot X_{n1} + u_1$$

$$Y_2 = b_1 \cdot X_{12} + b_2 \cdot X_{22} + \dots + b_n \cdot X_{n2} + u_2$$

.....

$$Y_T = b_1 \cdot X_{1m} + b_2 \cdot X_{2m} + \dots + b_n \cdot X_{nm} + u_n$$

El sistema de ecuaciones anterior se puede expresar de forma más compacta utilizando notación matricial. Así, vamos a denominar

$$y = \begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix} \quad X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1m} \\ X_{21} & X_{22} & \dots & X_{2m} \\ \dots & \dots & \dots & \dots \\ X_{n1} & X_{n2} & \dots & X_{nm} \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix} \quad u = \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{bmatrix}$$

El modelo de regresión lineal múltiple expresado en notación matricial es el siguiente:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix} = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1m} \\ X_{21} & X_{22} & \dots & X_{2m} \\ \dots & \dots & \dots & \dots \\ X_{n1} & X_{n2} & \dots & X_{nm} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{bmatrix}$$

Si tenemos en cuenta las denominaciones dadas a vectores y matrices, el modelo de regresión lineal múltiple se puede expresar de forma compacta de la siguiente forma:

$$y = Xb + u$$

Donde \mathbf{y} es un vector $n \times 1$, \mathbf{X} es una matriz $n \times m$, \mathbf{b} es un vector $n \times 1$, y \mathbf{u} es un vector $n \times 1$. El correspondiente modelo ajustado será el siguiente:

$$\hat{y} = X\hat{b}$$

El vector de residuos es igual a la diferencia entre valores observados y ajustados, es decir,

$$\hat{u} = y - \hat{y} = y - X\hat{b}$$

El parámetro b_i representa el efecto del aumento de una unidad del regresor x_i sobre la respuesta, y, cuando el resto de los regresores permanecen constantes.

Se pueden presentar tres casos en función del valor de n y de m .

- **$m > n$** : Significa que hay más variables que muestras lo que implica que el sistema tiene infinitas soluciones.
- **$m = n$** : Hay solo una solución, caso poco probable en problemas reales.
- **$m < n$** : El sistema tiene solución, siendo esto lo más habitual.

2.3.1 Mínimos Cuadrados (LS)

El método de Mínimos Cuadrados (Least Squares) es posiblemente el método más sencillo, el que menos complejidades presenta. Se basa en la ecuación expuesta en el apartado anterior $Y = Xb + u$, donde b son los coeficientes desconocidos que se quieren estimar (Uriel y Aldás, 2005).

Para calcular b empezaremos calculando el vector de residuos, que es igual a la diferencia entre valores observados y valores ajustados, es decir,

$$\hat{u} = y - \hat{y} = y - X\hat{b}$$

Denominando S a la suma de los cuadrados de los residuos, se tiene que:

$$S = \hat{u}'\hat{u} = [\hat{u}_1 \quad \hat{u}_2 \quad \dots \quad \hat{u}_m] \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \dots \\ \hat{u}_m \end{bmatrix} = \sum_{m=1}^T \hat{u}_m^2$$

Se obtiene que:

$$S = (y - X\hat{b})'(y - X\hat{b})$$

Derivando S con respecto al vector de coeficientes mínimo-cuadráticos, \hat{b} , e igualando a cero se obtiene el siguiente sistema de ecuaciones:

$$X'X\hat{b} = X'y$$

Al sistema anterior se le denomina genéricamente sistema de ecuaciones normales del hiperplano. Cuando $n = 2$, se obtiene el sistema de ecuaciones

normales de la recta; cuando $n = 3$, se obtiene el sistema de ecuaciones normales del plano; finalmente, cuando $n > 3$, se obtiene específicamente el sistema de ecuaciones normales de hiperplano, el cual no es susceptible de ser representado físicamente. En notación matricial expandida, el sistema de ecuaciones normales es el siguiente:

$$\begin{bmatrix} \sum_{m=1}^T X_{1m}^2 & \sum_{m=1}^T X_{2m} & \dots & \sum_{m=1}^T X_{nm} \\ \sum_{m=1}^T X_{2m} & \sum_{m=1}^T X_{2m}^2 & \dots & \sum_{t=1}^T X_{2m}X_{nm} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{m=1}^T X_{nm} & \sum_{m=1}^T X_{nm}X_{2m} & \dots & \sum_{m=1}^T X_{nm}^2 \end{bmatrix} \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \vdots \\ \hat{b}_n \end{bmatrix} = \begin{bmatrix} \sum_{m=1}^T Y_m \\ \sum_{m=1}^T X_{2m}Y_m \\ \vdots \\ \sum_{m=1}^T X_{nm}Y_m \end{bmatrix}$$

Para poder resolver el sistema respecto a \hat{b} unívocamente, se debe cumplir que el rango de la matriz $X'X$ sea igual a n . Si se cumple esta condición, se pueden premultiplicar ambos miembros como se muestra a continuación:

$$[X'X]^{-1}X'X\hat{b} = [X'X]^{-1}X'y$$

Con lo cual se obtiene la expresión del vector de estimadores mínimo cuadráticos:

$$\hat{b} = [X'X]^{-1}X'y$$

Al aplicar este método se nos pueden presentar los siguientes problemas

- Si $n < m$: La parte de la solución $X^T X$ no tiene inversa, y por lo tanto, el problema no tiene solución.
- Si $n > m$: Se puede calcular la solución generalmente pero puede ocurrir que algunas de las variables que forman X estén correlacionadas apareciendo así un problema de colinealidad, y siendo la matriz $X^T X$ casi singular. En estos casos la solución de mínimos cuadrados es muy poco robusta, llegando a ser muy sensible frente al ruido en los datos.

En muchas ocasiones, la colinealidad en una matriz se puede detectar a través de:

- La matriz de correlación de las variables explicativas.
- La diagonal de la inversa de esta matriz.
- Los autovalores de la matriz $X'X$.

La colinealidad se puede tratar eliminando regresores del modelo que tengan una alta correlación con el resto, disminuyendo de esta forma el número de parámetros que hay que estimar. Esta solución es considerada la más sencilla, pero no se puede utilizar en todos los casos.



2.3.2 Método Ridge

Tras estudiar el método de Mínimos Cuadrados, visto en el apartado anterior, surge como mejora este método de regresión Ridge, (Ryan, 2009) con el objetivo de aliviar la colinealidad que se puede presentar entre las variables predictoras. Para ello este método añade un término de penalización, k , haciendo que la matriz $X^T X$ del método anterior tenga inversa, y por tanto, el problema pueda proporcionarnos una solución más robusta.

Si antes minimizábamos la siguiente función:

$$\min SS(B) = \|Y - BX\|_2^2$$

Ahora se deberá minimizar la siguiente función para obtener la respuesta:

$$\min PSS_2(B) = \|Y - BX\|_2^2 + k\|B\|_2^2$$

La solución a este problema, y la expresión que nos proporciona este estimador será:

$$B = (X'X + kI)^{-1}X'Y$$

Donde Y y las columnas de X son tales que $X'X$ es una matriz de correlaciones, y $X'Y$ es un vector de correlaciones.

Si se asigna un valor de $k = 0$ se eliminaría el término de penalización y el método de regresión utilizado se convertiría en el modelo de mínimos cuadrados. En este método el valor asignado a la variable k se presenta crucial de cara a obtener una predicción de mejor o peor calidad. Varios métodos han sido propuestos para seleccionar la variable k , y todos estos métodos han llegado a la conclusión de que k es una variable aleatoria. Sin embargo, la justificación inicial para utilizar la regresión de Ridge que había sido dada por Hoerl y Kennard en 1970, sus creadores, estaba basada en que la asignación del valor k no era una variable aleatoria. Esto ha sido una de las razones por lo que se ha presentado una situación de controversia respecto a este método.

El método utilizado para hallar el valor de k es utilizar diferentes valores de k hasta que los parámetros estimados tienden a estabilizarse. El problema que se presenta con este método de calcular k es que se está realizando de una forma muy subjetiva, ya que dos personas podrían tener un criterio diferente sobre cuando se estabilizan los parámetros estimados.

Como consideración final cabe destacar que para poder aplicar este método es necesario que la matriz X esté centrada en cero y su varianza sea igual a uno.

2.3.3 Método Lasso

El Método Lasso, cuyo nombre viene de las siglas en inglés Least Absolute Shrinkage and Selection Operator, es una técnica propuesta por el canadiense Robert Tibshirani motivado por el objetivo de encontrar una técnica de regresión lineal que lograra estimaciones y predicciones, y que realizase una selección de variables.

Es una técnica de regresión similar a Ridge que presenta una pequeña diferencia en la penalización que conlleva importantes consecuencias. En concreto, a partir de cierto valor del parámetro de complejidad el estimador de Lasso produce estimaciones nulas para algunos coeficientes y no nulas para otros, con lo cual Lasso realiza una especie de selección de variables en forma continua. Lasso reduce la variabilidad de las estimaciones por la reducción de los coeficientes y al mismo tiempo produce modelos interpretables por la reducción de algunos coeficientes a cero.

La diferencia que presenta Lasso con Ridge es que a la hora de minimizar la función para hallar el estimador b Lasso utiliza la norma L_1 para el término penalización, mientras que Ridge utiliza la norma L_2 .

La norma L_1 de un vector es:

$$\|v\|_1 = \sum_{i=1}^n |v_i|$$

La norma L_2 de un vector es:

$$\|v\|_2 = \sqrt{\sum_{i=1}^n v_i^2}$$

Tras ver las diferencias existentes entre las normas L_1 y L_2 , la función a minimizar para obtener el estimador b será:

$$\min PSS_1(B, k) = \|Y - BX\|_2^2 + k\|B\|_1$$

El método Lasso suele presentar un modelo con buena precisión pero también presenta ciertas limitaciones:

- En el caso $p > n$, Lasso selecciona a lo sumo n variables antes de saturarse, debido a la naturaleza del problema de optimización convexa. Esto parece ser una limitación para un método de selección de variables. Además, Lasso no está bien definido a menos que el límite de la norma L_1 de los coeficientes sea menor que un cierto valor.
- Si hay n grupos de variables entre las cuales las correlaciones por parejas son muy altas, entonces Lasso tiende a seleccionar solo una variable del grupo, sin importarle cuál de ellas selecciona.
- En el caso $n > p$, si existe una alta correlación entre predictores, se ha observado que en general, la predicción a través de regresión Ridge resulta más óptima que la obtenida a través de Lasso.

Algunos estudios comparativos entre los métodos Lasso y Ridge nos permiten llegar a diferentes conclusiones:

- 1) El método Lasso al realizar una selección de variables presenta una gran ventaja frente al método Ridge, ya que produce modelos más simples y

más interpretables que implican un único subconjunto de predictores. Sin embargo no hay un método que sea dominante sobre el otro.

- 2) En general, se puede esperar que Lasso obtenga un mejor resultado en un entorno en el que un número relativamente pequeño de predictores tengan coeficientes sustanciales, y el resto de predictores tengan coeficientes muy pequeños o iguales a cero.

2.3.4 Método PLS (Partial Least Squares)

El método de regresión PLS fue introducido alrededor de 1975 por el estadístico Herman Wold para un tratamiento de cadenas de matrices y aplicaciones en econometría, y posteriormente continuado por su hijo Svante Wold en estudios relacionados con la quimiometría.

El Método de mínimos cuadrados parciales (PLS) es una técnica de reducción de la dimensionalidad que encuentra una regresión lineal mediante la proyección de las variables de predicción y las variables observables en un nuevo espacio de menor dimensión que maximiza la covarianza entre estas variables. Los elementos de X son las observaciones (variables del proceso) y los elementos de Y suelen ser las medidas de la calidad del producto (Puijganer et al., 2006).

El modelo subyacente general de PLS es:

$$X = TP^T + E$$

$$Y = UQ^T + F$$

- X es una matriz $n \times m$ de predictores.
- Y es una matriz $n \times p$ de respuestas.
- T y U son proyecciones, también conocidos como *scores*, de las matrices X e Y respectivamente.
- P y Q son matrices de carga ortogonales, también conocidas como *loadings*, de las matrices de X e Y respectivamente.
- E y F son los residuos, los términos de error.

Para obtener la matriz B hay varios algoritmos que se pueden utilizar para este modelo, pero quizás el más utilizado es el algoritmo NIPALS, que es el que se va a utilizar en este trabajo. A continuación se explican los pasos que sigue el algoritmo NIPALS:

- 1- Como primer paso se deben normalizar las matrices X e Y a media cero y varianza unidad para cada variable.
- 2- Iniciamos realizando la columna u_i .

$$X_i = X$$

$$Y_i = Y$$

$$H_i = 1$$

$$u_i = \text{cualquier columna de } Y$$

- 3- Se itera hasta la convergencia comparando u_i con su valor en la iteración anterior:

$$w_i = \frac{X_i^T u_i}{\|X_i^T u_i\|} \rightarrow t_i = X_i w_i$$

$$q_i = \frac{Y_i^T t_i}{\|Y_i^T t_i\|} \rightarrow u_i = Y_i q_i$$

- 4- Calcular:

$$p_i = \frac{X_i^T u_i}{\|X_i^T u_i\|}$$

$$r_i = H_i w_i$$

$$H_{i+1} = H_i - r_i p_i^T$$

- 5- Se calculan los nuevos residuos:

$$X_{i+1} = X_i - t_i p_i^T$$

$$Y_{i+1} = Y_i - t_i q_i^T$$

- 6- Volver al paso 3, con $i=i+1$, hasta encontrar todos los factores principales, es decir, esto se repite hasta que $i=\min(m,n)$ o hasta que se calculen el número adecuado "a" de factores PLS. Este orden de reducción se calcula usando validación cruzada.

Una vez calculado el algoritmo NIPALS, podemos determinar el valor de B que es el objetivo final de este método ya que $Y = X \cdot B$:

$$B = W(P^T W)^{-1}(T^T T)^{-1}T^T Y$$

Algunos de los factores importantes que motivan la utilización del método PLS son los siguientes:

1. El modelo de regresión PLS es un potente método de regresión lineal, que considera la multicolinealidad en las variables explicativas y acepta un número muy grande de variables.
2. El modelo resultante predice la respuesta a partir de un conjunto de variables linealmente dependientes x_1, x_2, \dots, x_n .
3. Durante el desarrollo del modelo, un número relativamente pequeño de componentes PLS son calculados y utilizados para la regresión.
4. El número de componentes PLS determina la complejidad del modelo y puede ser optimizado para tener un alto rendimiento en la predicción.

2.3.5 PCA (Análisis de Componentes Principales)

Esta técnica fue inicialmente desarrollada por Pearson a finales del siglo XIX y posteriormente fue estudiada por Hotelling en los años 30 del siglo XX. Sin embargo, hasta la aparición de los ordenadores no se empezó a popularizar.

El Análisis de Componentes Principales (PCA) es una técnica estadística de síntesis de la información, o reducción de la dimensión (número de variables) perdiendo la menor cantidad de información posible. Cabe destacar que las variables que proporcionan mayor información son aquellas que poseen una mayor varianza (Puijganer et al., 2006).

El método PCA construye una transformación lineal que escoge un nuevo sistema de coordenadas para el conjunto original de datos en el cual la varianza de mayor tamaño del conjunto de datos es capturada en el primer eje (llamado el Primer Componente Principal), la segunda varianza más grande es el segundo eje, y así sucesivamente. Para construirse esta transformación lineal debe construirse primero la matriz de covarianza o matriz de coeficientes de correlación. Debido a la simetría de esta matriz existe una base completa de vectores propios de la misma. La transformación que lleva de las antiguas coordenadas a las coordenadas de la nueva base es precisamente la transformación lineal necesaria para reducir la dimensionalidad de datos. Además las coordenadas en la nueva base dan la composición en factores subyacentes de los datos iniciales.

Una de las ventajas de PCA a la hora de reducir la dimensionalidad de un grupo de datos es que retiene aquellas características del conjunto de datos que contribuyen más a su varianza, manteniendo un orden de bajo nivel de los componentes principales e ignorando los de alto nivel. El objetivo es que esos componentes de bajo orden a veces contienen el aspecto más importante de esa información.

A continuación se muestra el método para el cálculo de PCA a partir de la matriz X de tamaño $n \times m$, que representa n observaciones de cada una de las m variables del proceso a considerar.

$$X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1m} \\ X_{21} & X_{22} & \dots & X_{2m} \\ \dots & \dots & \dots & \dots \\ X_{n1} & X_{n2} & \dots & X_{nm} \end{bmatrix}$$

Los vectores de carga pueden calcularse mediante la descomposición de valores singulares (SVD):

$$\frac{1}{\sqrt{n-1}} X = U \Sigma V$$

Donde $U \in R^{n \times m}$ y $V \in R^{m \times m}$ son unas matrices unitarias ($V^T V = I$, I la matriz identidad) y $\Sigma \in R^{n \times m}$ es una matriz diagonal que contiene los valores singulares reales no negativos de X en orden decreciente ($\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m \geq 0$). Los vectores de carga son los vectores columna ortonormales de la matriz V , y la varianza del conjunto de los datos proyectados sobre la columna i de V es σ_i^2 .

El hecho de resolver la ecuación anterior es equivalente a calcular los valores propios de la matriz de covarianza de X : $A = X^T X$.

$$A = \frac{1}{n-1} X^T X = V \Lambda V^T$$

Con $\Lambda = \Sigma^T \Sigma \in R^{m \times m}$ una matriz diagonal conteniendo los valores propios reales no negativos en orden decreciente de la matriz A ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$) y cuyo valor es $\lambda_i = \sigma_i^2$. Si se eligen los vectores propios correspondientes a los “a” valores propios más grandes y se almacenan en P con $P \in R^{m \times a}$, la proyección de los datos observados X, en este espacio de dimensión reducida se encuentra en la matriz T.

$$T = XP$$

Definiendo la columna i-ésima de T (conjunto de datos transformado), se cumplen las siguientes propiedades:

- La varianza esta ordenada:

$$\text{var}(t_1) \geq \text{var}(t_2) \geq \dots \geq \text{var}(t_a)$$

- Centrado en la media

$$\text{media}(t_i) = 0; \forall i$$

- Descomposición ortogonal

$$t_i^T t_j = 0; \forall i \neq j$$

- No existe ninguna otra expansión ortogonal de a componentes que capture más variación de los datos.

Finalmente, una vez calculado la matriz T se hallará B mediante la siguiente fórmula:

$$B = (T^T T)^{-1} T^T Y$$

2.3.6 Método Dinámico

Basándonos en los métodos anteriormente vistos, se ha desarrollado un nuevo modelo que tiene en cuenta la historia reciente de la planta. Este nuevo método no introduce ninguna diferencia en los algoritmos que se utilizan para hallar la predicción, sino que la diferencia se encuentra en la matriz de predicción X.

Normalmente la matriz de predicción X tiene tantas columnas como variables se desean estudiar. Sin embargo, el nuevo método dinámico contará con más columnas. 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 de l instantes pasados, además de l instantes pasados de la variable estimada Y. De esta forma la nueva matriz de predicción tendrá la siguiente forma:

$$X = [X(t), X(t - 1), Y(t - 1), X(t - 2), Y(t - 2), \dots, X(t - l), Y(t - l)]$$

Siendo X la matriz de predicción que se elige cuando se aplican los diferentes métodos, continuando el procedimiento de la misma forma que antes.

La finalidad de este nuevo método es introducir la dinámica del sistema, y por tanto, mejorar los resultados que se han de obtener con los métodos estáticos, ya que es de suponer que al poseer más información de las variables, 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 dinámico serán más pequeños.

2.4 Calidad de un proceso

La calidad es definida como el conjunto de características de un producto o servicio que le confieren la aptitud para satisfacer las necesidades del cliente. Es considerada actualmente como un factor imprescindible en los mercados que cada día son más competitivos. Hoy por hoy, la calidad no es opcional si se pretende sobrevivir como compañía en un entorno globalizado, y el no cumplimiento de ciertas especificaciones puede significar el primer paso hacia la salida del mercado.

2.4.1 Evolución histórica

El concepto de calidad ha ido evolucionando con el paso del tiempo, pudiendo dividir su evolución histórica en 5 etapas:

1) Industrialización

El término calidad parte en el siglo XIX, en los años de la Revolución Industrial, cuando el trabajo manual es reemplazado por el trabajo mecánico. En la Primera Guerra Mundial, las cadenas de producción adquieren mayor complejidad y simultáneamente surge el papel del inspector, que era la persona encargada de supervisar la efectividad de las acciones que los operarios realizaban. Es el primer gesto de control de calidad.

2) Control estadístico

La segunda etapa se sitúa entre 1930 y 1950. Las compañías ya no sólo dejan ver su interés por la inspección, sino también por los controles estadísticos. Estos procesos se vieron favorecidos por los avances tecnológicos de la época. Se pasó de la inspección a un control más global.

3) Primeros sistemas

Entre 1950 y 1980, las compañías descubren que el control estadístico no es suficiente. Hace falta desglosar los procesos en etapas y, tras un período de observación, detectar los fallos que se originen en ellas. En estos años surgen los primeros sistemas de calidad y las compañías ya no

dan prioridad a la cantidad de productos obtenidos, sino que se focalizan en la calidad.

4) Estrategias

A partir de los años 80 y hasta mediados de los 90, la calidad se asume como un proceso estratégico. Este es quizá uno de los cambios más significativos que ha tenido el concepto, pues a partir de este momento se introducen los procesos de mejora continua. La calidad, que ahora ya no es impulsada por inspectores sino por la dirección, se contempla como una ventaja competitiva. Además, toma como centro de acción las necesidades del cliente. Los Sistemas de Gestión se consolidan y la implicación del personal aumenta.

5) Calidad total

A partir de los años 90 y hasta la fecha, la distinción entre producto y servicio desaparece. No hay diferencias entre el artículo y las etapas que lo preceden; todo forma parte de un nuevo concepto que entra en escena: la Calidad Total, es decir, el proceso en su conjunto. Adicionalmente, la figura del cliente adquiere mayor protagonismo que en la etapa anterior y su relación con el artículo, que ahora llega incluso a etapas de posventa, se convierte en el principal indicador de calidad. Los sistemas se perfeccionan y se adaptan.

2.4.2 Dimensiones de la calidad

Tras la evolución que ha experimentado el término de calidad la figura del cliente se ha situado como un concepto clave a la hora de evaluar la calidad. Todo producto o servicio admite distintos niveles de valoración, frente a los cuales el cliente tiene tendencia a reaccionar de una manera diferente. A la hora de realizar esta valoración Garvin en 1984 estableció ocho *dimensiones de la calidad* independientes (Sebastian et al., 1998). Estas dimensiones son:

- **Características principales:** Corresponden a aquellas características que establecen las prestaciones funcionales más importantes del producto o servicio. Tiene en cuenta ciertos rasgos medibles como puede ser la velocidad o el consumo de un vehículo. Sin embargo esta dimensión no puede emplearse de manera individual ya que distorsionaría el concepto de calidad.
- **Características secundarias:** Son peculiaridades o prestaciones adicionales del producto o servicio que no contribuyen a satisfacer las necesidades básicas del cliente, pero que complementan el equipamiento o las prestaciones de aquél. Cabe destacar que en ocasiones la frontera entre características principales y secundarias es muy difusa y subjetiva según el cliente.
- **Conformidad con especificaciones:** Se refiere al grado en que las características primarias y secundarias del producto o servicio se ajustan a las especificaciones establecidas en el diseño o en las normas aplicables

vigentes. Esta dimensión resulta consecuencia directa de los recursos, tecnología y procedimientos puestos en juego durante la fabricación del producto, a diferencia de las características primarias y secundarias cuya responsabilidad recae más directamente en las fases de concepción y desarrollo del producto o servicio.

- **Comportamiento en servicio:** Esta dimensión puede evaluarse en términos de fiabilidad o de disponibilidad y en ocasiones llega a ser la dimensión preponderante a la hora de caracterizar la calidad del producto o servicio.
- **Duración:** La duración o caducidad técnica de un producto viene expresada por la vida del mismo; esto es, el periodo de tiempo existente entre su puesta en servicio y el instante en que se alcanza un determinado criterio de inutilidad. Dicho criterio de inutilidad puede corresponder al deterioro catastrófico del producto, a la aparición de una avería que no admite una reparación convencional o a la llegada de una situación en la que la frecuencia de las averías y los costes de reparación alcanzan valores tan importantes que afectan sensiblemente a la rentabilidad en el uso del producto considerado.
- **Servicio y asistencia técnica:** El éxito comercial de gran cantidad de productos se basa en el servicio asociado a su asesoría, venta e instalación. También constituye un elemento importante para configurar la calidad de un bien el grado de efectividad de la asistencia técnica postventa, entendiéndose ésta tanto a nivel de garantía, como de mantenimiento, reparación e incluso de sustitución.
- **Estética:** Corresponde al acabado formal del producto o servicio y suelen ser aquellas características que no son importantes desde un punto de vista funcional pero que resultan esenciales en la comercialización del mismo, sobre todo a la hora de atraer las preferencias de los compradores. En ocasiones la dimensión estética puede invertir la apreciación de la calidad de un bien por parte de consumidores individuales e incluso a nivel de toda la sociedad.
- **Percepción:** La calidad de un producto o servicio no es únicamente una propiedad intrínseca de él, también debe ser percibida por sus usuarios y por los potenciales compradores. De hecho, no resulta nada sencillo determinar cómo es percibida la calidad por los consumidores.

Estudiando las ocho dimensiones indicadas de la calidad, cabe destacar la consideración de lo relativo que puede resultar tipificar la calidad de una entidad, incluso en casos de un producto físico de baja complejidad.

A pesar de la importancia correspondiente al cliente, el concepto de calidad engloba desde que el diseño de un producto es llevado a fábrica hasta que éste es entregado al cliente para su uso, por lo tanto se puede decir que incluye la satisfacción del cliente, pero también se aplica tanto al producto como a la organización.

2.4.3 Gestión de la calidad

Analizando los diferentes enfoques y aproximaciones del concepto de calidad, podemos decir que un sistema de gestión de calidad se compone básicamente de tres partes en el ámbito productivo: Planificación de la calidad, control de la calidad y mejora continua de la calidad.

- **Planificación de la calidad:** consiste en realizar el desarrollo de una estrategia anticipada que asegure que los productos y servicios que se crean tengan la capacidad de satisfacer las necesidades y expectativas de los clientes. Este proceso comprende la identificación, clasificación y ponderación de las características de calidad, así como los objetivos, requisitos y restricciones de la misma. Para obtener las anteriores condiciones se lleva a cabo una estrategia de calidad que comienza con la identificación del posible cliente y determinar sus necesidades y finaliza con la elaboración de un proceso optimizado y estandarizado que se adapte a sus preferencias y necesidades.
- **Control de calidad:** este proceso participa en la caracterización de los nuevos productos en sus diferentes fases de desarrollo y en el establecimiento de las especificaciones de calidad. En esta fase se desarrollan y se coordinan la ejecución de los métodos de ensayo en busca de las características de calidad deseadas tanto en las materias primas y materiales como en los productos intermedios y finales. Para elaborar un control de calidad se siguen los siguientes pasos:
 1. Elegir el aspecto o característica a controlar del proceso.
 2. Desarrollar un objetivo.
 3. Determinar una unidad de medida.
 4. Elaborar un sensor o forma de medir la característica de control.
 5. Medir la característica durante el proceso o al final de éste.
 6. Analizar las diferencias resultantes entre el rendimiento real y el esperado.
 7. Realizar las acciones necesarias para obtener los valores que se habían marcado como objetivo.
- **Mejora continua de la calidad:** hoy en día una empresa no puede conformarse con la calidad obtenida por un producto. Esto significa quedarse estancado en un mercado en el que todas las empresas están en continua mejora, por lo que una empresa siempre deberá estar en constante estudio y mejora de la calidad de su producto o proceso.

El modo de obtener con éxito el control de calidad conlleva la participación activa de todos los trabajadores de una empresa en la mejora del desarrollo, diseño y fabricación del producto. Para alcanzar estos objetivos se implantan programas,

mecanismos, herramientas y diferentes técnicas que persigan y garanticen una mejora en la calidad de sus productos. Los estudios estadísticos se presentan fundamentales a la hora de decidir qué técnicas son las adecuadas para su implantación.

2.4.4 Aplicación de Métodos Estadísticos a la calidad

La utilización de estos Métodos Estadísticos es uno de los principales puntos de apoyo, por lo que se hace necesario un mejor conocimiento de los mismos por parte del personal que trabaja en relación directa con la calidad, como puede ser el de recepción de materiales, el de producción o el que se dedica específicamente al control de calidad.

Está demostrado que efectuar una inspección total de la producción, además de producir un elevadísimo coste, no siempre proporciona la confianza deseada. Factores tales como la monotonía, por la repetitividad del proceso de inspección, o el cansancio de quien la realiza, llevan consigo, en la práctica, a la obtención de hasta un quince por ciento de productos aceptados o rechazados incorrectamente, es decir:

- Productos defectuosos aceptados como buenos.
- Productos no defectuosos rechazados.

La estadística proporciona técnicas rápidas, sencillas y económicas que permiten obtener resultados con una fiabilidad muy superior a la indicada anteriormente. Las áreas de la Estadística más utilizadas en Control de Calidad son:

- La Estadística Descriptiva que trata la recogida y análisis de los datos.
- La Estadística inductiva o inferencial, cuyo objetivo central es el estudio de técnicas que permitan obtener conclusiones sobre la población a estudio a partir de los datos de una muestra.

La conocida rueda de Deming (Figura 2.2), muestra el papel central que juega la Estadística en un programa de mejora de la calidad.

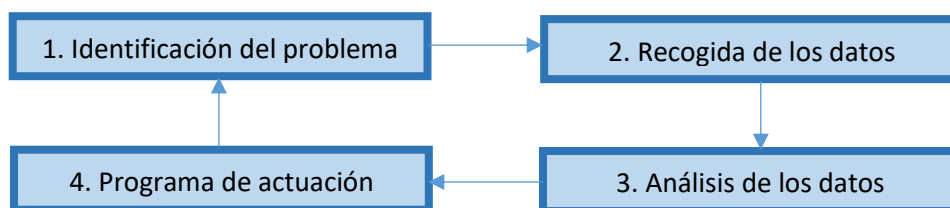


Figura 2.2. Rueda de Deming

Lo más frecuente es que los datos que se analizan y se estudian sean numéricos, pero no siempre es así. Si se estudian características no medibles o cualitativas se citarán como atributos y si son medibles se citarán como variables. En muchos casos una tabla o una representación gráfica de estos datos obtenidos nos informa con un simple golpe de vista de la distribución de los mismos.

A partir de los datos obtenidos se definen los parámetros estadísticos que son medidas que recogen la información muestral y que se clasifican en parámetros de centralización (se sitúan en el centro de los datos) y parámetros de dispersión (miden las desviaciones de los datos respecto de algún parámetro de centralización). Los más utilizados en control estadístico de calidad son la media y la mediana como parámetros de centralización y el rango (o recorrido) y la desviación típica como parámetros de dispersión.

En muchos casos es preciso utilizar más de una variable para describir el proceso, se utilizan entonces las técnicas estadísticas de descripción conjunta de varias variables. Como herramienta de relación gráfica se estudia el diagrama de dispersión, se introducen la covarianza y el coeficiente de correlación lineal y se obtiene la recta de regresión de mínimos cuadrados.

Aunque una buena descripción de los datos es necesaria, el objetivo principal del método estadístico es obtener conclusiones acerca de la población a partir de los datos de la muestra. Una población es el conjunto de individuos que se desea estudiar, una muestra es el subconjunto de elementos de la población que se estudia realmente.

Una de las áreas de la Estadística que resultan fundamentales en calidad es la inferencia estadística. Se trata, en definitiva, de obtener información sobre la población a partir de los datos de la muestra. Se introducen los estimadores puntuales y se describe el método de estimación por intervalos de confianza como base para el entendimiento posterior de los gráficos de control.

2.4.5 Proceso de Inspección

El procedimiento de inspección de un producto, servicio o proceso debe considerarse como una parte integrante del control de producción y de calidad. Este procedimiento consiste en ciertas comprobaciones de la calidad de los productos que van resultando a través de un cierto proceso de fabricación, ya sea en la aceptación de materiales procedentes de proveedores, en la vigilancia de productos intermedios o en la recepción de productos terminados, a fin de que exista un determinado nivel de calidad de los mismos.

Existen diferentes procedimientos de inspección, todos ellos tendentes a garantizar y proteger los intereses tanto del consumidor como del proveedor, mediante el establecimiento de soluciones que comprometan y satisfagan a ambos.

Una inspección puede realizarse tanto a magnitudes de calidad variable como a atributos de la misma. Dado que en el primer caso la metrología y sus diversos patrones ofrecen una solución bastante exacta de las inspecciones, es en el segundo caso donde tiene sentido que se aplique.

Además, la consideración de que una magnitud variable esté o no incluida en un cierto intervalo de tolerancia aceptable, puede considerarse como un atributo, que



puede cumplir una de sus dos posibilidades, aceptable o defectuoso, respectivamente.

La inspección es un proceso operativo en el que básicamente pueden distinguirse tres etapas: planificación, ejecución y control.

La primera etapa consiste en la definición de los puntos en los que tendrá lugar la inspección, estableciendo sus pautas e instrucciones; incluye también la designación del personal que efectuará la misma y los medios necesarios para su ejecución. El responsable de la planificación suele ser el director de calidad.

La planificación de la inspección comienza por el análisis de las especificaciones del producto: materiales, dimensiones, acabado, etc., y con estos datos se puede definir su sistema y forma en que se va a realizar dicha inspección.

La etapa de ejecución consiste en el desarrollo de la planificación en lo referente a la verificación del producto. Comprende la interpretación de la especificación, la medida real del producto y la comparación de ambas.

Comparar el resultado de la medición con las especificaciones correspondientes no siempre es fácil, decidir la aceptación o el rechazo resulta sencillo si el límite de aceptabilidad está bien definido, como ocurre en las características variables, pero si éstas son atributos se pueden crear decisiones subjetivas si los criterios no están claramente establecidos.

La tercera etapa corresponde al control de resultados obtenidos en la inspección. Esta etapa permite no solo la detección de las no conformidades sino que además permite establecer las acciones correctivas necesarias.

En la mayoría de los casos es necesario conservar los resultados de la inspección, para lo que se precisa un archivo donde queden registrados los resultados de la misma y el procedimiento seguido para obtenerlos; con ello se podrá llegar a conocer la evolución de la calidad a lo largo del tiempo.

Para medir la exactitud de las inspecciones se utiliza al inspector de comprobación que reexamina todos los productos para identificar las unidades correctas que se han rechazado, las unidades defectuosas que se han aceptado y el número real de defectos del conjunto. Con estos datos se puede calcular el porcentaje defectuoso identificado correctamente, que es lo que se conoce como exactitud del inspector.

Se pueden definir dos formas de actuar en el proceso de inspección:

- Inspección al 100 por 100: consiste en examinar todos y cada uno de los elementos de un conjunto. Este método es útil para pequeñas cantidades y se realiza sobre dimensiones o características críticas. En estos casos está generalizado el uso de instrumentos de medida automáticos, con el fin de reducir al mínimo el coste de mano de obra.
- Inspección por muestreo: consiste en la inspección de un grupo reducido de elementos pertenecientes al conjunto total. Este método resulta más

rápido, más exacto y más económico que la inspección al 100 por 100 y resulta muy útil en conjuntos de gran tamaño e imprescindible en ensayos destructivos.

No se puede definir un método mejor que otro, ya que cada caso es único. Por lo tanto, al final el aspecto económico es muchas veces el factor determinante a la hora de elegir un método de inspección u otro.

2.4.6 Normativa sobre calidad

Con el fin de establecer un marco común para poder valorar la Gestión de la Calidad se creó la norma ISO 9001, vigente desde los años 80. El ISO 9001 es una normativa de carácter voluntario que establece los requisitos mínimos que las organizaciones deben cumplir para garantizar la calidad de sus servicios y productos. Sin embargo, estos requisitos han ido cambiando y la norma ha ido incorporando mejoras y avances en función de las necesidades y, sobre todo, de la propia evolución del concepto de calidad empresarial.

La Organización Internacional de Estandarización (ISO) surgió ante la necesidad de elaborar normas que tuviesen un alcance mundial, sobre todo tras el aumento del comercio que se registró a partir de los años 50. El primer paquete normativo ISO 9000 se publicó en el año 1987 en el que se incluyeron tres modelos normativos orientados a unificar los criterios para la implementación de un sistema de Gestión de Calidad en las empresas. Dichos modelos fueron ISO 9001, ISO 9002 e ISO 9003. Sin embargo, para una mejor aplicación de la norma, en el año 2000 ISO llevó a cabo una revisión del estándar y decidió unificar el paquete normativo en el estándar ISO 9001, que es desde entonces el principal referente en temas de calidad a nivel internacional.

El estándar ISO 9001 es actualmente la norma más prestigiosa y mejor valorada en el terreno de la Gestión de Calidad. Aunque no es el único estándar vigente en el mundo, cada día son más las organizaciones que se acogen a sus requisitos para optimizar sus procesos internos y mejorar la calidad de sus productos y servicios. No obstante, dicho reconocimiento no se ha forjado de la noche a la mañana. Todo lo contrario, ha sido el resultado de un largo proceso evolutivo en el que ISO ha tenido que atender las distintas necesidades de las organizaciones y los cambios en el contexto empresarial de las últimas décadas. En concreto, los avances tecnológicos y la tecnificación de los procesos han provocado que los estándares sean revisados periódicamente para adaptarse a cada momento

La norma ISO de Calidad proporciona múltiples beneficios para la gestión de la empresa, así como de cara al mercado y los clientes, ya que cada vez son más los clientes que requieren proveedores certificados. Además esta norma se considera una plataforma ideal desde la que avanzar hacia otras certificaciones de sistemas de gestión del medio ambiente, seguridad o responsabilidad social.

2.5 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 2.3 (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 2.3 (b). Estas distribuciones, como podemos apreciar en la figura 2.3 (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 2.3 (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 2.3 (e).

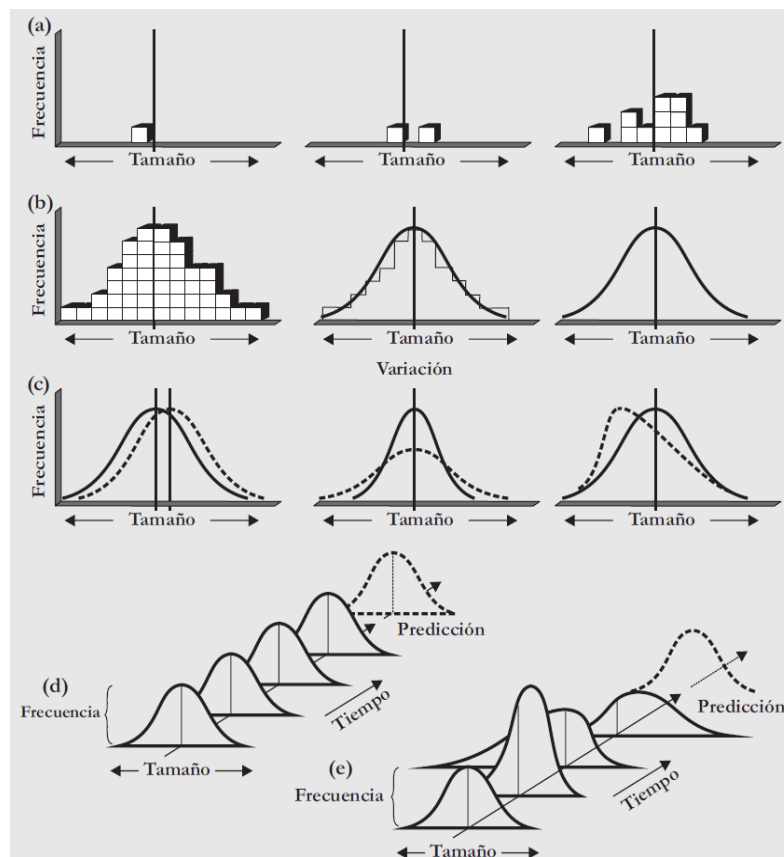


Figura 3. Variaciones naturales y especiales. (a) Las muestras varían; (b) pero forman un patrón que, si es estable, es llamado distribución. (c) Las distribuciones pueden diferir en la medida de la tendencia central, variación, forma, o cualquier combinación de estas. (d) Si solo se presentan causas 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 y González, 2014)

El objetivo es, por tanto, eliminar las variaciones especiales y mantener el proceso bajo control. La figura 2.4 muestra tres tipos de salidas del proceso; en la figura 2.4 (a) podemos observar un proceso fuera de control (proceso salvaje); la figura 2.4 (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 2.4 (c) muestra un proceso bajo control y dentro de los límites establecidos (proceso entrenado).

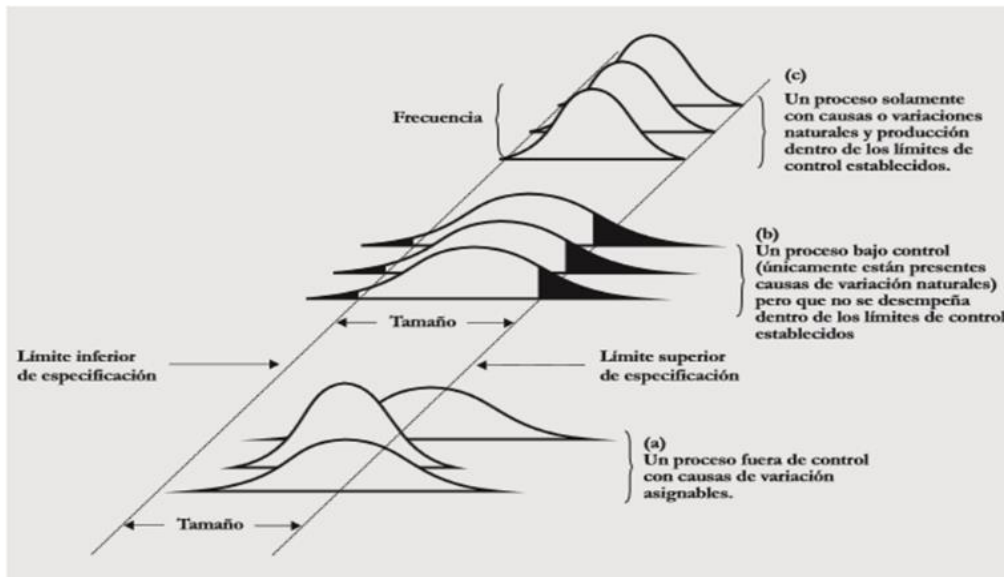


Figura 2.4. Control del proceso: Tres tipos de salida del proceso. (Fuente: Control estadístico de procesos. Carro y González, 2014)

2.5.1 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 δ = 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.

2.5.2 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.

2.6 Detección de fallos

2.6.1 Introducción

La gran demanda por producir productos de alta calidad ha provocado que características como fiabilidad, disponibilidad, seguridad y protección medioambiental adquieran cada vez más importancia.

Los controladores estándar están diseñados para mantener los procesos operando de forma satisfactoria, pero a veces aparecen causas especiales que se traducen en un incorrecto funcionamiento del proceso. Estas causas especiales son causadas por errores en el proceso o en los componentes del mismo. Para evitar esta situación o prever la posible degradación se presenta necesario dotar a los sistemas de control de un esquema de monitorización capaz de detectar fallos incipientes o el inicio de la degradación de un proceso para evitar el daño o futuras pérdidas.

La monitorización del proceso a la hora de detectar un fallo consiste en los siguientes pasos:

- 1) Detección de la presencia de un fallo a partir de la información extraída de los datos medidos del proceso (módulo de detección).
- 2) Identificación del fallo: identificar las variables más relevantes para diagnosticarlo.
- 3) Diagnóstico del fallo: módulo que someta las características y los fallos producidos en la planta a un procedimiento de clasificación con la finalidad de determinar el tipo de fallo presente, la localización de este, su tamaño y la causa que lo produjo.
- 4) Recuperación del proceso que se trata de eliminar los efectos del fallo.

Cabe destacar la diferencia entre los términos fallo y avería. Fallo es un cambio inesperado del sistema que lleva una degradación del comportamiento global del mismo pero que no tiene por qué representar un defecto de ninguno de sus componentes. En cambio, avería se refiere a un comportamiento catastrófico del sistema lo que conlleva una interrupción permanente de la capacidad para mantener una función requerida por el sistema.

2.6.2 Métodos de detección de fallos

Existen tres métodos diferentes para transformar los datos tomados de un proceso en medidas útiles, y en caso de ser necesario, diagnosticar los fallos encontrados. Estos métodos son: métodos basados en señal, métodos analíticos y métodos basados en conocimiento.

Estos métodos nos proporcionan una información complementaria, ya que resulta demasiado complejo analizar un caso desde un único punto de vista.

2.6.2.1 Métodos basados en señal

Como se muestra a continuación existen diferentes técnicas que pueden ser utilizadas:

- **Alarmas:** Las variables medidas son comparadas con valores límites constantes y previamente establecidos. Existe la posibilidad de que el sistema pueda presentar alguna anomalía sin llegar a alcanzar los valores límite prestablecidos debido a que los niveles de alarma no pueden adaptarse a la dinámica del sistema.
- **Técnicas estadísticas multivariantes:** las variables medidas en un proceso están correlacionadas con otras variables del proceso (correlación espacial) y con otras medidas de la misma variable tomadas en diferentes tiempos (correlación temporal). Debido a la necesidad de manejar correlaciones espaciales se han desarrollado algunas de estas técnicas como pueden ser el método de Mínimos Cuadrados Parciales (PLS) o el Análisis de Componentes Principales (PCA), vistos anteriormente.
- **Redundancia física:** La distribución espacial por la planta de diferentes componentes físicos, como actuadores, sensores y componentes del sistema, garantizan una cierta protección en presencia de algún fallo localizado. Este sistema de comparación de las señales de salida de elementos permite tanto detectar la presencia de fallos como su localización y así aislarlos. La desventaja que presenta es que no es capaz de detectar fallos que afecten a todos los elementos de la misma manera. Además de elevarse el coste económico y el uso del espacio al duplicar todos los elementos.
- **Análisis en frecuencia:** ciertas variables como las vibraciones de una máquina rotativa pueden ser estudiadas en el dominio de la frecuencia, aportándonos así una gran cantidad de información sobre el estado de operación de la misma. Al representar cada fallo por un patrón de frecuencias y amplitudes característico permite realizar su diagnóstico.

2.6.2.2 Métodos analíticos

Estos métodos basados en la redundancia analítica emplean un modelo de la planta. Las medidas que se producen se derivan de un mismo estado dinámico, por lo que están funcionalmente relacionadas a través de su modelo. Estas técnicas basadas en modelos constan de un primer paso que consiste en la generación de los residuos, y un segundo paso en el que se toma una decisión basada en los residuos.

En el primer paso, los residuos que nominalmente son cero (se desvían del cero si se ha producido un fallo particular), se generan mediante un algoritmo determinado en el que se procesan las entradas y salidas del sistema. Generalmente la técnica para generar los residuos consiste en calcular la diferencia entre la salida de la planta y la salida calculada por un modelo de la

misma. Algunas de las formas más conocidas para calcular el modelo son: los filtros de Kaman, las ecuaciones de paridad, los observadores de estado y la estimación de parámetros.

El segundo paso consiste en determinar la existencia o no de fallos. Este proceso puede consistir en un sencillo test de umbral en los valores instantáneos, en la media móvil de los residuos, o pueden utilizarse los métodos de la teoría estadística de decisión.

La principal característica de la generación de residuos es su robustez (insensibilidad) a los errores que se pueden dar en el modelado, los cuales son inevitables. Frente a la posibilidad de producirse falsas alarmas o una incorrecta identificación de los fallos, es aconsejable determinar umbrales apropiados para cada residuo y diseñar generadores de residuos robustos.

2.6.2.3 Métodos basados en conocimiento

Existe la posibilidad de utilizar las diferentes técnicas que se muestran a continuación:

- **Sistemas expertos:** esta técnica sirve de complemento a las técnicas vistas anteriormente pudiendo utilizar así la información para poder analizar todos los síntomas y realizar un diagnóstico completo. Este sistema representa el conocimiento de los expertos en forma de reglas, las cuales relacionan los síntomas con las causas que lo han provocado, y poder ofrecernos soluciones o recomendaciones.
Las dos formas más comunes de diagnosis son el diccionario de fallos y el árbol de decisión. El diccionario de fallos contiene una lista de causas y efectos, mientras que el árbol de decisión consiste en un conjunto estructurado de reglas lógicas del tipo “Si-síntoma-Y-síntoma-ENTONCES-conclusión”. Cada conclusión puede servir como síntoma en la siguiente regla y así sucesivamente hasta alcanzar la conclusión final. Esta técnica presenta el problema de que requiere la existencia de uno o más expertos en el sistema estudiado. Este experto debe conocer tanto el funcionamiento como los posibles fallos o situaciones extrañas que pueda presentar el sistema nominal. Esta situación trae como consecuencia una limitación a la hora de desarrollar 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 diagnosis, las redes neuronales admiten diferentes aplicaciones. La principal consiste en clasificar los diferentes estados del sistema a partir de las variables medidas. Estos estados se pueden 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 que presenta este modelo es que puede no aprender el modo de funcionamiento si no existen datos de fallo o si éstos son muy escasos. La limitación que presenta es que su capacidad de generalización podría llevar a clasificar modos de fallo no aprendidos como modos de funcionamiento normal.

Una red neuronal también 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:** estas técnicas pueden ser utilizadas para la evaluación de los residuos (introduciendo reglas difusas basadas en el conocimiento del sistema) y para la generación de residuos, ya que un sistema lógico difuso es un aproximador universal de funciones, concretamente funciones que relacionan las distintas variables del proceso, es decir, calculando un modelo difuso del proceso. Finalmente la aparición de fallos se detecta comparando la salida del sistema lógico difuso y la salida del proceso.
- **Búsqueda causal:** consiste en el seguimiento de los fallos hasta llegar a sus causas, que son representadas a través de grafos directos dirigidos. En los grafos 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. Para encontrar la causa final del fallo solo habrá que realizar una búsqueda a través del grafo.

2.6.3 Requerimientos de diseño

En todo método de detección y diagnóstico de fallos se deben cumplir las siguientes características:

- **Tiempo bajo de detección (t_d):** es el tiempo que transcurre desde que se produjo el fallo hasta que el esquema de monitorización lo detecta. Normalmente se minimiza para un índice de falsas alarmas fijo.
- **Alto grado de detecciones correctas (P_d):** todos los fallos que ocurren deben ser detectados.
- **Índice bajo de falsas alarmas (P_f):** hace referencia a aquellas alarmas que indican la aparición de un fallo que no ha ocurrido.
- **Buena capacidad de aislamiento:** es la capacidad de distinguir los diferentes fallos. Depende de los test estadísticos empleados 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 magnitudes de los coeficientes de las ecuaciones del sistema y de las características del ruido. La sensibilidad está íntimamente relacionada con el tiempo necesario de detección y localización del fallo.

- **Robustez:** es una valoración del comportamiento del método de diagnóstico. Trata de aislar fallos cuando existen errores de modelado y/o perturbaciones desconocidas. Este aspecto es de gran importancia al ser casi inevitable la aparición del algún error en el modelado o el desconocimiento de las perturbaciones que afectan al sistema.

Una forma de analizar la eficacia de un Sistema de Detección e Identificación de Fallos es provocar un determinado fallo, mantenerlo y observar cómo reacciona el sistema de detección. Si se observa que el fallo produce una respuesta nula por parte del módulo de motorización, se podrá considerar que el fallo ha sido omitido, lo cual será aceptable si se trata de un fallo sin importancia. En el caso contrario, esta situación se consideraría totalmente inaceptable.

Una de las características que se presenta imprescindible es la rapidez con la que se detecta la existencia de un fallo en el sistema. En función del sistema al que se aplique el sistema de detección de fallos se le otorgará más importancia a la fiabilidad en la detección de fallos de pequeña magnitud o al tiempo de detección de dicho fallo.

Los fallos no siempre se producen de repente y persisten durante un tiempo prolongado. Existen también fallos que se desarrollan lentamente o son de pequeña magnitud, los cuales son conocidos como fallos incipientes. En estos casos la sensibilidad del sistema de detección de fallos juega un papel primordial, por delante de la rapidez. Algunos de estos fallos incipientes son medidas con sesgo (errores que ocurren de forma sistemática), fallos intermitentes o desviaciones de media móvil de un instrumento (“drifts”).

2.6.4 Control estadístico de procesos (SPC) para detección de fallos

Las gráficas de control permiten medir y trazar la evolución de una determinada característica de la calidad de un proceso por unidad de tiempo. Suele ser la herramienta más utilizada a la hora de monitorizar el estado de un proceso.

Toda grafica 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. Estos límites permiten detectar la presencia de factores especiales, que son los fallos a detectar en el área de la detección y diagnóstico de fallos, 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).

Aunque existen diferentes criterios a la hora de calcular los valores límite, existe un criterio que es bastante utilizado hallando estos valores con la siguiente ecuación:

$$UCL = media + 3 \cdot desviación\ estándar$$

$$LCL = media - 3 \cdot desviación\ estándar$$

En la figura 2.5 se aprecia cómo se relacionan los límites de control con la distribución. Un valor de la muestra ubicado en la región situada entre los dos límites, superior e inferior, nos indica que el proceso muestra causas comunes de variación, mientras que un valor que se encuentre fuera de los límites de control indicaría que el proceso presenta causas especiales de variación.

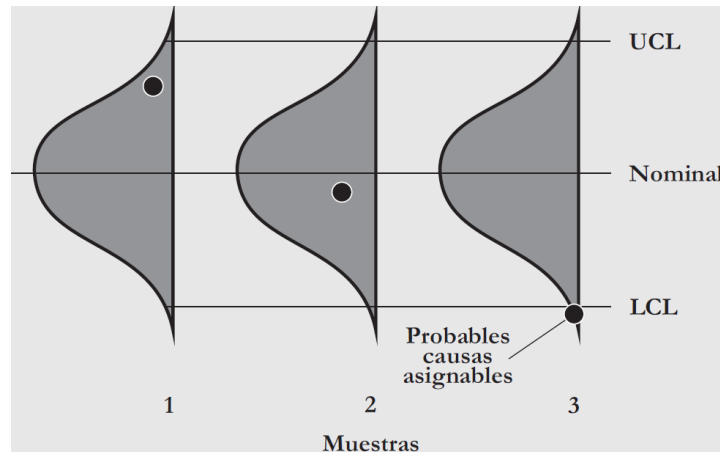


Figura 2.5. Relación de los límites de control con la distribución de muestreo y observaciones de tres muestras tomadas (Fuente: Control estadístico de procesos. Carro y Gonzalez, 2014)

Una observación que se encuentra fuera de los límites de control no significa que la calidad sea baja en todos los casos. Esto puede deberse a un cambio del proceso, lo que provoca que la estadística de calidad tomada en esa muestra se ubique fuera del rango establecido por los límites. En estos casos, se presentará necesario construir una nueva gráfica de control.

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

- 1) Tras tomar una muestra aleatoria del proceso, se mide la característica de calidad y se calcula una medida variable o de atributos.
- 2) En el caso de que la distribución se sitúe fuera de los límites de control, se busca una causa especial.
- 3) Se elimina la causa en el caso de que ésta degrade la calidad, o se incorpora en el caso de que aporte una mejora. Tras ello, se reconstruirá la gráfica de control.
- 4) Se repetirá el proceso periódicamente.

Las gráficas de control al estar basadas en distribuciones de muestreo no son herramientas perfectas y pueden presentarnos dos tipos de errores:

- **Error "Tipo I"**
Se comete este tipo de error cuando se rechaza un lote de buena calidad. Esto puede deberse a que el proceso se encuentre fuera de los límites de control debido a un efecto puramente aleatorio.

- **Error “Tipo II”**

Se comete este error al 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.

2.6.5 Detección de fallos usando PCA

En muchos casos para la detección de fallos usando el método PCA se utiliza la estadística de Hotelling's o T^2 , que se puede calcular directamente a partir de la representación en componentes principales (De la Fuente, 2010), dado un vector de observación x :

$$T^2 = xP\Lambda_a^{-1}P^T x^T$$

Donde P incluye solo los vectores de carga asociados con los “ a ” valores singulares más grandes y Λ_a contiene las primeras “ a ” filas y columnas de Λ . La estadística T^2 nos proporciona una medida de las variaciones del proceso, y en el caso de que esta medida supere un umbral indicará que el sistema se encuentra fuera de control, lo que significa que hay un fallo. El umbral para la estadística T^2 se halla de la siguiente forma:

$$T_a^2 = \frac{(n^2 - 1)a}{n(n - a)} F_\alpha(a, n - a)$$

Donde:

- F_α es la distribución de Fisher-Snedecor, con a y $n-a$ grados de libertad.
- a es el número de componentes principales seleccionado.
- n es el número de observaciones.
- α es el nivel de significancia que especifica el grado de compromiso entre el nivel de falsas alarmas y el nivel de alarmas no detectadas en el sistema. Algunos valores típicos son $\alpha = 0,01$ ó $0,05$.

La estadística T^2 además es útil para detectar operaciones del proceso fuera de sus condiciones normales de operación.

A pesar de esto, monitorizar el sistema solo con los primeros “ a ” componentes de PCA no es suficiente, y por tanto los “ $n-a$ ” valores restantes también deberán ser monitorizados. Estos valores se monitorizarán de una forma más robusta utilizando la estadística Q , que se halla de la siguiente forma:

$$r = (X - \hat{X})$$

$$Q = rr^T$$

Donde r es el vector de residuos calculado como la proyección de la observación x en el espacio PCA. Si un proceso está bajo control, el valor de Q será pequeño.

Para evaluar si está fuera de control, se calcula un límite superior de control para esta estadística basado en los datos históricos de la siguiente forma:

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

Con $\theta_i = \sum_{j=a+1}^n \sigma_j^{2i}$ y $h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}$.

Tras detectar la existencia de un fallo el siguiente paso será determinar la causa del fallo, es decir, qué variable o variables son las que han provocado dicho fallo. Para averiguar esto con el estadístico T^2 se sigue el siguiente procedimiento:

- 1) Comprobar los vectores normalizados $(t_i / \sigma_i)^2$ para la observación x y determinar el número r de vectores principales responsables (con $r < a$) del estado fuera de control. (Por ejemplo, los vectores tales que $(t_i / \sigma_i)^2 > 1/\alpha$ (T_α^2), recordando que t_i es la proyección de la observación x en el vector de carga i -ésimo, y σ_i es el correspondiente valor singular).
- 2) Calcular la contribución de cada variable x_j a los t_i fuera de control:

$$\text{cont}_{i,j} = \frac{t_i}{\sigma_i^2} p_{i,j} (x_j - \mu_j)$$

- 3) Cuando $\text{cont}_{i,j}$ es negativo, se iguala a cero.
- 4) Calcular la contribución total de la variable x_j :

$$\text{CONT}_j = \sum_{i=1}^r \text{cont}_{i,j}$$

- 5) Hacer una gráfica con los valores de CONT_j , para todas las posibles m variables, x_j .

A través de los valores de CONT_j se podrá observar las variables del proceso que realizan una mayor aportación al fallo, y por tanto serán estas variables las culpables del fallo.

Para determinar la variable o variables que han provocado el fallo con el estadístico Q , se calculará el valor de cada residuo al cuadrado en el instante que se ha detectado el fallo de la siguiente forma:

$$\text{residuo} = (X - \hat{X})^2$$

El residuo o residuos del instante en el que se ha producido el fallo que tengan el valor más alto respecto de los demás será la variable o variables responsables del fallo.



Depuradora de aguas residuales

3. DEPURADORA DE AGUAS RESIDUALES

Una estación depuradora de aguas residuales (Figura 3.1) es una planta cuya función principal es recoger las aguas residuales que vierten una población o una industria con el objetivo de reducir la concentración de contaminantes presentes mediante ciertos tratamientos y procesos para devolver al medio natural las aguas en el mejor estado posible. Algunos de los agentes contaminantes más importantes aparecen en la siguiente tabla junto con la razón por los que se busca eliminar:

Contaminantes	Razón de la importancia
Sólidos en suspensión	Estos sólidos en suspensión pueden provocar el desarrollo de depósitos de fango y de condiciones anaerobias en el entorno acuático.
Materia orgánica biodegradable	Se compone principalmente por proteínas, carbohidratos y grasas animales. Su vertido sin el debido tratamiento puede llevar al agotamiento de los recursos naturales de oxígeno.
Patógenos	Pueden transmitir enfermedades contagiosas.
Nutrientes	Entre estos nutrientes destacan el nitrógeno, el fósforo y el carbono, que pueden favorecer el crecimiento de una vida acuática no deseada, pudiendo también provocar la contaminación del agua subterránea si se vierte en cantidades excesivas.
Contaminantes prioritarios	Son compuestos orgánicos o inorgánicos característicos por su toxicidad o mutagenicidad.
Materia orgánica refractaria	Esta materia tiende a resistir los métodos convencionales de tratamiento. Algunos ejemplos típicos son agentes tensoactivos, fenoles y pesticidas agrícolas.
Metales pesados	Muchos de estos metales son añadidos al agua residual en ciertas actividades industriales, y deben ser eliminados si se pretende reutilizar el agua residual.
Sólidos inorgánicos disueltos	Estos sólidos son el calcio, sodio y los sulfatos, que se añaden al agua de suministro para el uso del agua, y es posible que se deban eliminar si se va a reutilizar esta agua residual.

Para llevar a cabo el tratamiento se utilizan diferentes operaciones físicas, químicas y biológicas, realizándose en diferentes etapas. Generalmente estas etapas se dividen de la siguiente forma:

- Pretratamiento
- Etapa primaria
- Etapa secundaria
- Etapa terciaria
- Desinfección del agua tratada
- Tratamiento de fangos

La etapa de pretratamiento consistiría en medir y regular el caudal de agua que llega a la planta así como extraer los sólidos flotantes grandes, arenas, grasas o aceites con la finalidad de evitar problemas en tratamientos posteriores (obstrucciones, depósitos de arena, rotura de equipos...).

La etapa primaria tiene como principal objetivo eliminar los sólidos en suspensión mediante un proceso de sedimentación por gravedad o separación de las partículas en suspensión mediante burbujas de aire. El método más utilizado suele ser el de sedimentación por gravedad donde se separa por efecto de la gravedad las partículas suspendidas cuyo peso específico sea mayor que el del agua para obtener un efluente clarificado y un fango con elevada concentración de sólidos. Este proceso se suele realizar con estanques de sedimentación cuya forma puede ser circular, cuadrada o rectangular.

La etapa secundaria pretende eliminar la materia orgánica mediante un proceso de naturaleza biológica seguido de otro proceso de decantación. El proceso biológico es un tratamiento de descontaminación natural realizado por seres vivos, concretamente microorganismos ya presentes en el agua residual. Estos microorganismos utilizan la materia orgánica como sustrato o fuente



Figura 3.1. Estación depuradora de aguas residuales de Vitoria

de alimento para obtener energía, pudiéndose producir este método de forma aeróbica o anaeróbica. La forma anaeróbica necesita un sistema de aireación para aportar oxígeno a la reacción química. El resultado de este proceso es la formación de una biomasa bacteriana que posteriormente se precipitará en el decantador. El lugar donde se pone en contacto la biomasa con el agua residual para llevar a cabo el proceso se denomina reactor biológico o biorreactor. Los microorganismos se pueden encontrar suspendidos en el agua, fijados a un medio inerte, haciendo pasar el agua por este lecho, o distribuidos en un sistema mixto, que es una combinación de los métodos.



En el decantador secundario hay un flujo lento de agua, de forma que la biomasa, es decir, los flóculos bacterianos producidos en el reactor, sedimentan. Este sedimento producido está formado principalmente por bacterias denominándose fango activo.

La etapa terciaria consiste en otro tratamiento más con la finalidad de aumentar la calidad del agua residual. Generalmente esta etapa es específica de cada planta, ya que en función de la procedencia o del destino del agua se realizará un tratamiento u otro.

La siguiente etapa sería la desinfección del agua tratada, siendo esta etapa realmente importante si estas aguas van a tener un uso recreacional, piscicultura o van a ser una fuente para consumo humano. A pesar de haber realizado todos los tratamientos anteriores, hay microorganismos patógenos que siguen sobreviviendo. Para eliminarlos se pueden realizar diferentes procedimientos, como la cloración, la ozonización, la bromación o la radiación ultravioleta. El proceso más utilizado es la cloración por ser barata y muy efectiva.

Una vez realizadas estas etapas, es necesario tratar los fangos producidos en la etapa primaria o secundaria. Esta etapa es la que denominaríamos como tratamiento de fangos.

Estos lodos se caracterizan por ser un residuo extremadamente líquido (más de un 95% de agua). Su composición no es fija, depende de la carga de contaminación del agua residual inicial y de los tratamientos que se realizan sobre ésta, así como su eficiencia. Los lodos contienen amplia diversidad de materias suspendidas o disueltas, algunas de ellas con cierto valor agrónomo o micronutrientes esenciales para las plantas como pueden ser nitrógeno (N), fósforo (P) y potasio (K), y otras sustancias con potencial contaminante como materiales pesados, patógenos y diferentes contaminantes orgánicos.

El propósito final es reducir el contenido en agua y los patógenos de los lodos producidos. Con el objetivo de asegurar la estabilidad de la materia orgánica se aplican diferentes tratamientos como:

- Digestión aeróbica
- Digestión anaerobia
- Compostaje
- Acondicionamiento químico
- Tratamiento físico

Según datos del Registro Nacional de Lodos, tienen como destino principalmente su utilización como fertilizante agrícola (aproximadamente un 80%). Actualmente se está logrando disminuir en gran medida los lodos destinados al vertedero (cerca de un 8%), mientras que lo destinado a incinerarse está incrementando su valor (en torno a un 4%). Otro destino con menos importancia son los suelos no agrícolas.

La digestión anaerobia consiste en la degradación de la materia orgánica presente en ausencia de oxígeno molecular convirtiéndose principalmente en un 65-70% en metano (CH_4), un 25-30% en dióxido de carbono (CO_2) y en menor cantidad nitrógeno, hidrógeno o ácido sulfúrico. Este proceso se desarrolla en un reactor completamente cerrado donde se introducen los lodos de forma continua o intermitente.

La digestión aeróbica se realiza en un estanque abierto realizándose la digestión de la materia orgánica por bacterias aerobias. El producto final es dióxido de carbono (CO_2) y agua (H_2O). La principal ventaja de este proceso es que no se produce metano ni malos olores, y se consiguen menores concentraciones de DBO (Demanda Biológica de Oxígeno).

El compostaje es el proceso de digestión aeróbica del fango mezclado con madera u otros desechos orgánicos, con el objetivo de obtener un producto utilizable como fertilizante.

El acondicionamiento químico tiene como objetivo aumentar la concentración de los lodos reduciendo su contenido en agua. Esta etapa también es conocida como espesado, habiendo principalmente tres técnicas de llevarla a cabo: espesado por sedimentación o gravedad, espesado por flotación y espesado por centrifugado.

Una vez concluidos estos procesos, los fangos siguen teniendo una alta cantidad de agua. Para ello se han diseñado dos métodos principales: secado por aire y secado mecánico. El secado por aire puede realizarse con lechos de arenas, lagunas de fangos, lechos adoquinados y eras de secado. En cambio, el secado mecánico se lleva a cabo a través de filtros de banda, filtros prensa, filtros de vacío y centrífugas.

Finalmente, en función de su cantidad de materias tóxicas se destina a terrenos agrícolas, o por lo contrario, si es elevado, a vertederos o ser incinerados.

Después de haber visto los procesos que se pueden llevar a cabo en una planta depuradora de aguas residuales cabe destacar que no siempre se realizan las mismas etapas, depende en gran parte del origen de estas aguas, ya que su composición es diferente. Algunos de los posibles orígenes son los siguientes:

- Aguas residuales domésticas o aguas negras: proceden de las aguas sanitarias de los hogares de los seres humanos. Suelen contener gran cantidad de materia orgánica y microorganismos, así como restos de jabones, detergentes, lejía y grasas.
- Aguas blancas: tienen su procedencia de la atmósfera en forma de precipitaciones, o del riego y limpieza de lugares públicos.
- Aguas residuales industriales: proceden de los procesos realizados en fábricas y establecimientos industriales, siendo su composición muy variable dependiendo de cada actividad industrial. Pueden contener aceites, detergentes, ácidos, grasas y productos de origen mineral, químico, vegetal o animal.

- Aguas residuales agrícolas: son todas aquellas aguas procedentes de las labores agrícolas en las zonas rurales. Muchas veces estas aguas no tienen un tratamiento previo.

Por lo tanto, a la hora de diseñar la instalación resulta muy importante conocer el origen de las aguas residuales así como el destino del agua tratada. Conociendo el destino se podrán tener en cuenta los niveles máximos de contaminantes permitidos.

Existen varios modelos de simulación de depuradoras de aguas residuales que varían en función de los diferentes elementos que contengan, como pueden ser digestores, decantadores, sistemas de tratamiento de fangos, y según el entorno en el que se haya programado, como son Matlab, FORTRAN, etc. Concretamente en el estudio que vamos a realizar en este trabajo se va a utilizar un modelo de depuradora conocido como BSM2 (Benchmark Simulation Model 2), que está considerado como uno de los modelos más completos. En la figura 3.2 se muestra un esquema del modelo BSM2:

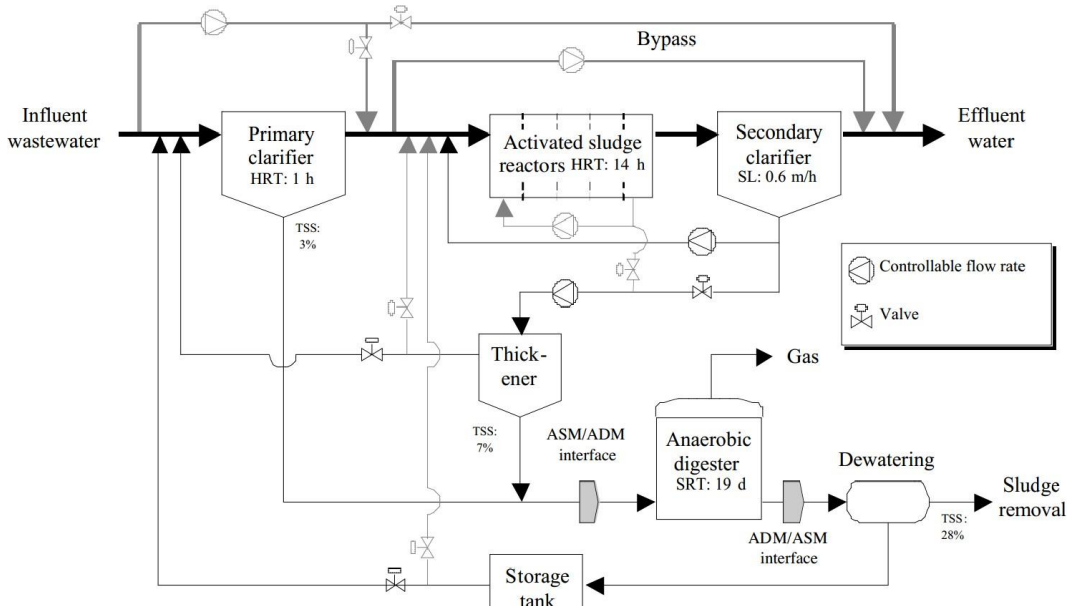


Figura 3.2. Esquema del modelo BSM2 de una depuradora de aguas residuales

Este modelo se divide en varios bloques compuestos por diferentes elementos como se puede ver en la figura 3.2. Este modelo se inicia con la entrada de un efluente de agua residual al sistema pasando en primer lugar por un decantador primario (Primary clarifier), que realiza la primera separación de materia en suspensión. Una vez superado este primer paso, el caudal pasará a un bloque formado por cinco reactores químicos distribuidos en serie (Activated sludge reactors) en los cuales, a través de procesos biológicos descomponen la materia orgánica reduciendo la demanda química de oxígeno del agua residual y amonificando el nitrógeno orgánico. Estos reactores se dividen en dos grupos, uno primero formado por los dos primeros reactores que realizan el proceso biológico a través de bacterias y otros microorganismos de forma anaerobia, y otro segundo grupo formado por los tres últimos reactores que realizan el proceso biológico de forma aerobia. A la salida de estos reactores se sitúa un segundo decantador

(Secondary clarifier), encargado de reducir la cantidad de sólidos en suspensión en el líquido, obteniendo un agua preparada para abandonar la depuradora.

Por otro lado, los sólidos retirados del agua residual se envían a un bloque en el que serán tratados ya que no se pueden extraer del sistema al medio ambiente en tal estado.

Una pequeña parte de los sólidos se envían directamente a la entrada de los reactores en serie explicados anteriormente. El resto de sólidos deben ser tratados en el bloque de tratamiento de fangos en los que se someten a diversos tratamientos con los objetivos de reducir la cantidad de materia orgánica presente generando gas combustible a través del digestor anaeróbico (Anaerobic Digester) y de reducir la cantidad de agua que transportan (Thickener y Dewatering). Finalmente estos fangos tratados son almacenados en un tanque (Storage Tank) para después ser reintroducidos en el sistema, mientras que otra parte de éstos se retira con la finalidad de eliminarlos en vertederos, etc.

Este modelo cuenta con un sistema de control que regula su funcionamiento. Este sistema mide el nivel de oxígeno disuelto en el líquido del reactor 4 del bloque de reactores (Activated sludge reactors). Otro control con el que cuenta el sistema es la posibilidad de variar los caudales Q_T (caudal de retorno del decantador secundario a los reactores), Q_{int} (caudal de retorno del reactor 5 al reactor 1) y Q_w (caudal de entrada al “espesador” de lodos, Thickener). Además cuenta con la existencia de algunas tuberías que permiten desviar el flujo de agua en el caso de que las condiciones lo requieran, para que no atraviese determinados elementos de la red (tuberías “by-pass”).

A lo largo del proceso se van a recoger datos en diferentes puntos de nuestro sistema. Estos datos pueden ser tanto unidimensionales como multidimensionales. En nuestro caso, aunque habrá algunas variables de una dimensión, como puede ser el tiempo, la mayoría de las variables van a ser multidimensionales proporcionándonos diferentes datos sobre el estado del fluido en un determinado punto. El estado del fluido lo vamos a medir con 7 variables en diferentes puntos a lo largo del proceso. Las siete variables consideradas son:

1) DQO

Esta primera medida nos indicará la Demanda Química de Oxígeno de nuestro caudal. La DQO es un parámetro que mide la cantidad de sustancias susceptibles de ser oxidadas por medios químicos que hay disueltas o en suspensión en una muestra líquida. Estas sustancias son materiales solubles inertes, materias inertes y sustratos biodegradables. Se utiliza para medir el grado de contaminación y se expresará en $\frac{g \text{ DQO}}{m^3}$.

2) O_2

Nos indicará la cantidad de oxígeno disuelto que hay en el agua residual. Se expresará en gramos de oxígeno disuelto metro cúbico ($\frac{g \text{ } O_2}{m^3}$).

3) Salk

Esta variable nos indicará el valor de la alcalinidad del agua, que nos expresa la capacidad del agua para neutralizar ácidos. Esta propiedad se

debe principalmente a la presencia de ciertas sales de ácidos débiles, aunque también puede contribuir la presencia de bases débiles y fuertes. No debe confundirse esta medida con el pH.

4) N

Hace referencia a la cantidad de nitrógeno que contiene el agua residual. Este nitrógeno aparece en diferentes formas como pueden ser nitritos, nitratos, amonios o nitrógeno orgánico soluble asociado con Sólidos Suspendidos. Esta variable se expresará en $\frac{g N}{m^3}$.

5) SS

Esta variable hace referencia a la cantidad de Sólidos Suspendidos presentes en un agua residual y que pueden ser separados por diferentes procedimientos. Se expresará como la cantidad en gramos de sólidos suspendidos que habrá por metro cúbico de agua ($\frac{g SS}{m^3}$).

6) Caudal

El caudal se corresponde con la cantidad de agua que pasa por un lugar en una cierta cantidad de tiempo, es decir, corresponde a un volumen de agua. Se expresará como $\frac{m^3}{d}$.

7) Temperatura

Nos indicará en grados centígrados la temperatura del agua residual.



Resultados

4. RESULTADOS

4.1 Introducción

En este capítulo se presentará la aplicación de las diferentes técnicas utilizadas para el diseño del sensor software que se ha marcado como objetivo. El sensor software será utilizado para hallar la variable DQO (Demanda Química de Oxígeno) del proceso de una depuradora de aguas residuales, como variable de calidad del proceso. El sensor software tendrá también la misión de detectar y realizar el diagnóstico de posibles fallos que se presenten a lo largo del proceso.

El objetivo es diseñar un sensor software para estudiar la calidad del proceso. Para ello utilizaremos diferentes técnicas y se compararán los resultados obtenidos por cada una de ellas.

Después se analizarán los resultados obtenidos observando cuál nos ofrece una mejor estimación para nuestra variable de calidad. Además 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.

Finalmente también se incluirá en este capítulo la metodología seguida a la hora de realizar la simulación en Matlab, que es el programa informático que nos permitirá introducir nuestro estudio. Matlab además nos permite ver las diferencias presentes entre cada método comparándolos 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 encuentran fuera de su funcionamiento correcto.

4.2 Toma de datos de la planta

Tras haber enumerado en el capítulo anterior las variables que se van a medir a lo largo del proceso, cabe destacar que los datos que se van a utilizar en el estudio han sido recogidos directamente del proceso en un funcionamiento normal de la planta depuradora de aguas residuales.

La muestra de datos que se debe tomar se presenta muy importante de cara al estudio a realizar. La importancia reside en que la muestra debe tener un gran tamaño para así poder recoger diferentes situaciones del proceso. Una muestra de reducido tamaño podría provocar que el sensor no aprendiese diferentes situaciones lo que conllevaría a la posibilidad de que saltasen falsas alarmas al analizar un nuevo conjunto de datos.

La situación en la que se recogerán los datos para nuestro estudio será una situación exenta de cualquier fallo o funcionamiento anómalo. Además también se considera recomendable no recoger datos de una situación de arranque de la

planta porque en muchas ocasiones los datos no están estabilizados. Se considera imprescindible que la recolección de datos esté correctamente espaciada y registrada a intervalos regulares, de forma que se refleje la mayor normalidad de los mismos.

La muestra de datos que se va a recoger en la planta depuradora de aguas residuales contendrá datos de 58465 instantes diferentes. En cada instante se realizarán medidas de siete variables en veinte puntos diferentes del proceso de la planta, por lo tanto, por cada instante se recogerán 140 datos.

4.3 Preprocesamiento de los datos

Para programar los diferentes métodos en Matlab se ha seguido una misma estructura a la hora de realizar el tratamiento de datos y la valoración de resultados. Esta estructura no incluye las ecuaciones propias de cada método, lo cual sería la parte independiente de cada programa.

El primer paso del programa es construir la matriz de datos con la que se aplicará posteriormente el método matemático. Esto se realiza con el objetivo de eliminar variables que apenas carezcan de importancia o que hasta incluso nos puedan falsear los datos. Otro de los objetivos de esta operación es reducir el tamaño de la muestra de datos con el fin de reducir también el tiempo de ejecución del programa y no someterle a una gran carga de datos.

Para disminuir el tamaño de la muestra de datos primero vamos a calcular la varianza de cada variable medida en todos los instantes en los que se ha recogido datos. Una vez hallada la varianza se eliminarán de la matriz de datos aquellas variables que tengan una varianza menor de 0,0001. Las variables que se eliminan son aquellas que permanecen constantes o apenas cambian a lo largo del proceso productivo ya que éstas no nos aportan información de cara a la predicción que se desea realizar.

El siguiente paso para reducir el tamaño de la muestra de datos es hallar el coeficiente de correlaciones de la muestra con la variable que queremos predecir. Una vez obtenido el coeficiente de correlación de cada variable se establecerá un valor umbral, el cual marcará el límite para introducir la variable o no en nuestra matriz de datos final. Este valor umbral se suele elegir en función de los valores obtenidos y del número de variables que se quieran introducir en la matriz final. Se puede decir por lo tanto que este valor es elegido en función del usuario que realice la estimación.

A continuación se ha decidido normalizar a media cero y varianza uno la muestra de datos. Esta decisión se ha tomado debido a que algunos métodos necesitan esta condición para asegurar un correcto funcionamiento.

Antes de aplicar el método estadístico, el rango de observaciones en la matriz se decide dividir en dos grupos. Un primer grupo, denominado "Datos de entrenamiento", será utilizado para que el sistema "aprenda" el comportamiento



en condiciones normales del proceso. El segundo grupo de datos, denominado “Datos de prueba”, serán los datos sobre los cuales se probarán la robustez de cada método, y así poder comparar entre ellos para observar y determinar las ventajas y desventajas de cada uno sobre los demás.

Tras este paso se procederá ejecutar el método estadístico, que es diferente en cada caso. Con este método se halla el vector de valores b que nos permite obtener los valores de predicción multiplicando el vector b por la matriz de datos recogidos de la planta.

Después de hallar el vector b , se multiplicarán los dos grupos de observaciones por dicho vector con el objetivo de comparar la predicción calculada con los valores que disponemos de nuestra muestra inicial. De esta forma se podrá ver si se obtiene una buena respuesta para los datos que llamamos de “entrenamiento” y probar el vector de valores b con unos datos nuevos, los “datos de prueba”. Para ver esta comparación se calcularán las diferencias entre los valores reales y los valores obtenidos por el método de predicción, además de acompañarlos con una gráfica para observarlo de forma gráfica.

4.4 Estimación de la variable DQO en el efluente

4.4.1 Modelo estático

En este apartado se probarán los diferentes modelos y se verán los resultados obtenidos tanto para los datos de entrenamiento como para los datos de test. Finalmente a la hora de reducir el tamaño de la muestra de datos se ha decidido incluir solo aquellas variables que tuviesen un coeficiente de correlación mayor de 0,3 con la variable DQO que queremos hallar. De esta forma se obtiene una matriz de la dimensión 58465×14 , indicando que solo tendremos en cuenta 14 variables en la matriz con la que vamos a trabajar.

4.4.1.1 Método de Mínimos Cuadrados

Para hallar el vector solución b con el método de Mínimos Cuadrados se aplicará la función *regress*, que incluye el programa Matlab, para relacionar la muestra de datos de entrenamiento con los valores del vector Y , también de entrenamiento, que hemos tomado.

En las figuras 4.1a y 4.1b se puede ver la variable DQO calculada según los datos de entrenamiento. Se puede observar que los datos predichos se ajustan bastante bien a los datos reales.

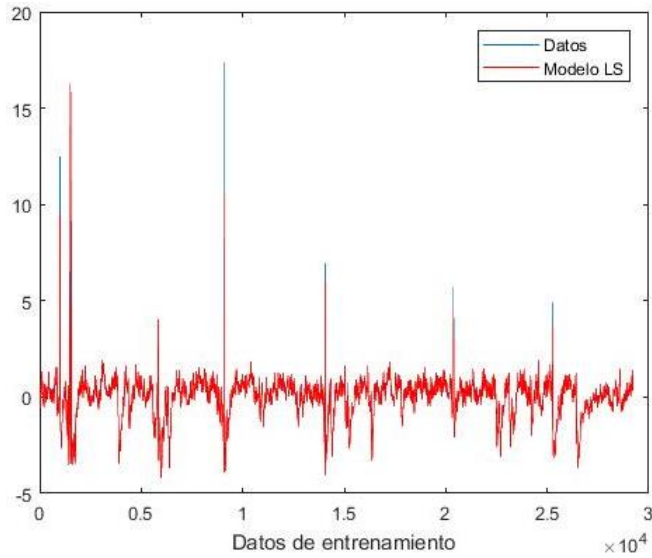


Figura 4.1a. Variable DQO de entrenamiento obtenido por mínimos cuadrados.

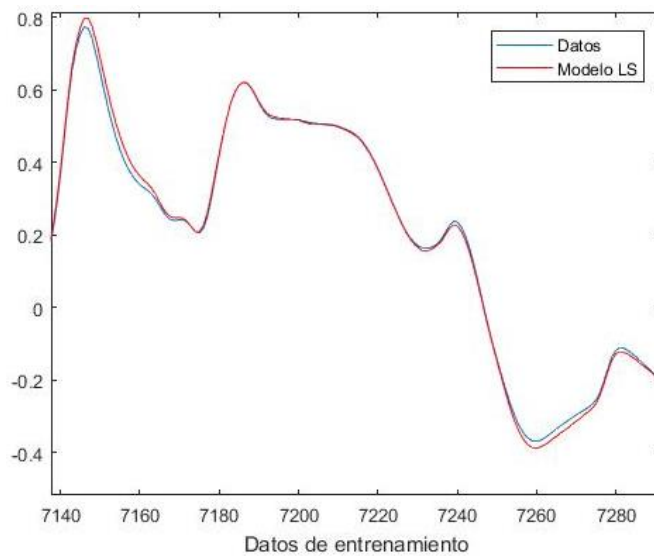


Figura 4.1b. Detalle de la variable DQO de entrenamiento obtenido por mínimos cuadrados.

A continuación en las figuras 4.2a y 4.2b se puede ver la variable DQO calculada con los nuevos datos que llamamos de validación. Como se puede observar la solución obtenida también es buena, aunque a simple vista parece que se ajusta menos que los resultados obtenidos con los datos de entrenamiento.

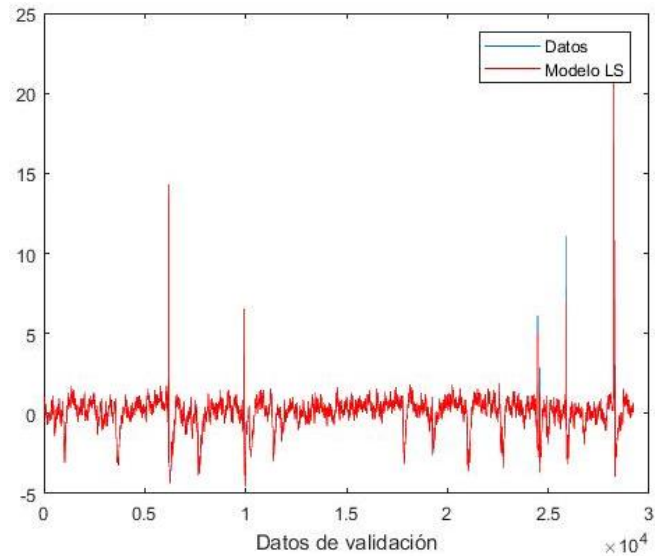


Figura 4.2a. Variable DQO de validación obtenido por mínimos cuadrados.

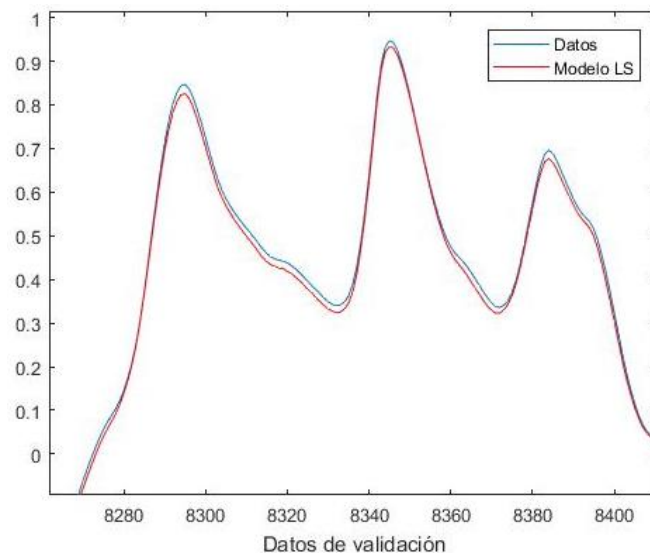


Figura 4.2b. Detalle de la variable DQO de validación obtenido por mínimos cuadrados

En las dos predicciones se puede observar la existencia de algunos picos que representan ruido, pero al tratarse de un número reducido de picos no debe ser algo preocupante.

Cabe destacar que al ejecutar esta aproximación el programa Matlab nos da un aviso en forma de *Warning*, que nos señala que la muestra de datos tiene un rango deficiente de cara a la precisión de la máquina. Esto nos hace referencia a que la colinealidad es alta en la muestra, y aunque en este caso nos proporcione una solución, existe la posibilidad de que con otra muestra de datos el método no sea capaz de calcular una estimación. Por lo tanto sería aconsejable buscar otro método que nos evite la aparición de este posible problema.

Finalmente, queda valorar de una forma objetiva la calidad de la predicción calculada y ver cómo se comporta frente a la introducción de nuevos datos. Para

ello, se calculan el valor medio de los residuos tanto de los datos de entrenamiento como de los datos de validación. Además en la figura 4.3 se representa una gráfica con el valor de los residuos en cada instante de los datos de validación calculados.

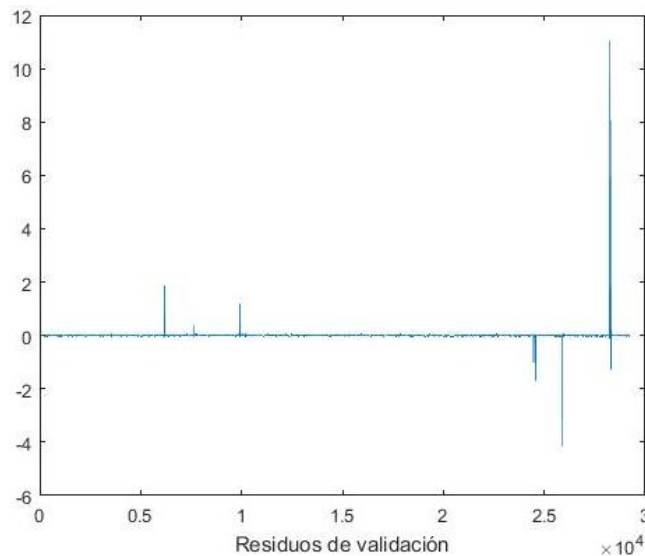


Figura 4.3. Residuos existentes entre la variable predicha por mínimos cuadrados y la variable real

Valor medio de residuo entrenamiento = 0,1516

Valor medio de residuo validación = 0,2674

Tras observar la figura en la que se representan los residuos se puede observar que existe un número moderado de picos, lo cual no debería provocar ninguna preocupación ya que la mayoría de residuos se encuentran en valores muy próximo a cero. Tras valorar los valores medios de los residuos se corrobora la observación que se había apreciado gráficamente destacando que el método no “aprende” tan bien el modelo como los valores obtenidos en la fase de entrenamiento, es decir, el método no trabaja tan bien con datos de una muestra nueva.

4.4.1.2 Método Ridge

Como se vio en el capítulo dos, el método Ridge es una modificación del método de mínimos cuadrados que surgió con la finalidad de eliminar el posible problema de colinealidad que se puede dar. Por lo tanto con la aplicación de este método se espera eliminar este problema que precisamente se ha presentado antes, aunque esto no quiere decir que los valores de la estimación vayan a ser de mejor calidad.

En este método se plantea el problema de elegir un valor de k que proporcione una predicción de gran calidad, para ello se propone probar diferentes valores de k hasta que se estabilicen los valores predichos. Tras realizar diferentes pruebas se alcanzó la decisión de asignar el valor de 0,001 a k .

Para aplicar este método se utilizará la función *ridge* que nos proporciona el programa Matlab.

En las figuras 4.4a y 4.4b se puede ver la variable DQO según los datos de entrenamiento. Visualmente parece que se ha obtenido una buena aproximación de los datos predichos en comparación con los datos reales.

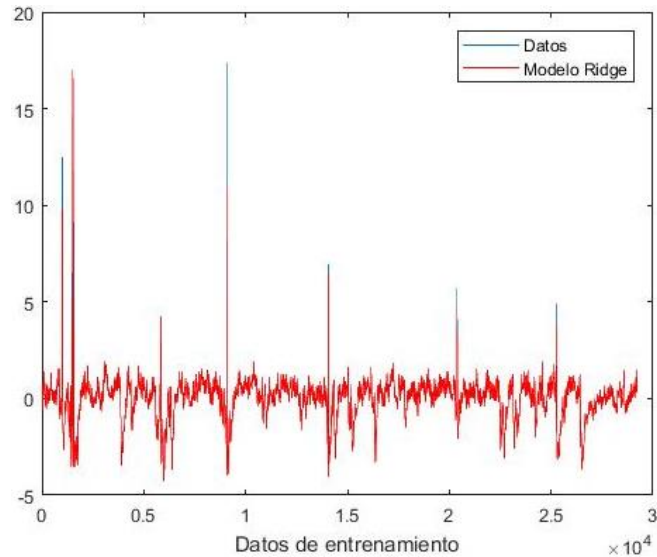


Figura 4.4a. Variable DQO de entrenamiento obtenido por Ridge.

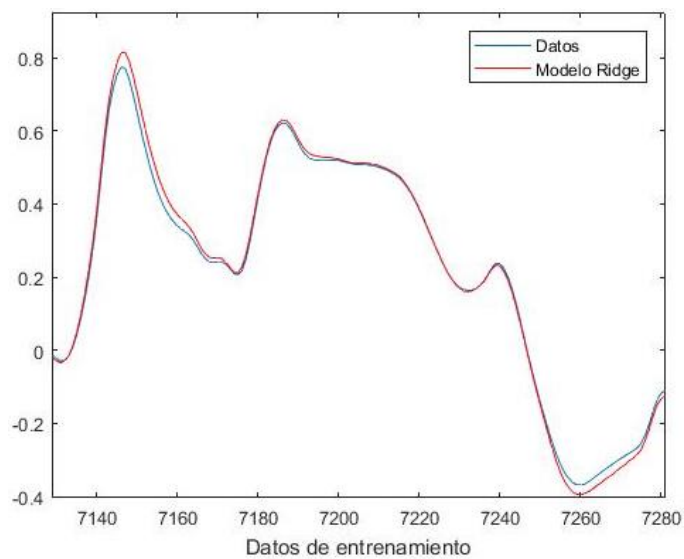


Figura 4.4b. Detalle de la variable DQO de entrenamiento obtenido por Ridge.

A continuación en las figuras 4.5a y 4.5b se puede ver la variable DQO calculada con los nuevos datos que llamamos de validación. A simple vista, la solución obtenida también parece buena.

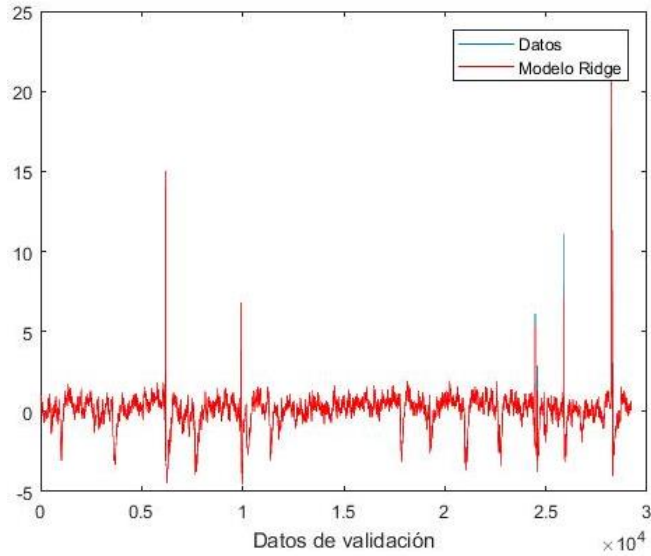


Figura 4.5a. Variable DQO de validación obtenido por Ridge.

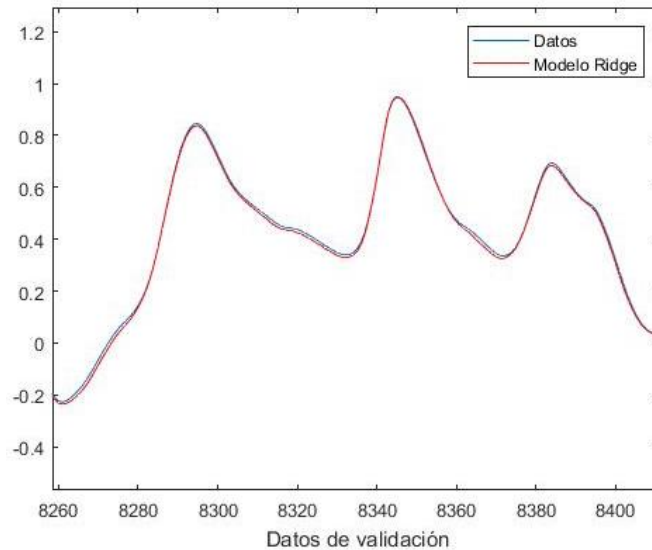


Figura 4.5b. Detalle de la variable DQO de validación obtenido por Ridge.

Después de aplicar el método Ridge, se puede observar que los resultados obtenidos han sido buenos, pero sobre todo lo importante es que se ha conseguido eliminar el problema relacionado con la colinealidad que aparecía aplicando el método anterior. En último lugar queda analizar los residuos obtenidos, lo cual se apreciará a través de la figura 4.6 y los residuos medios obtenidos.

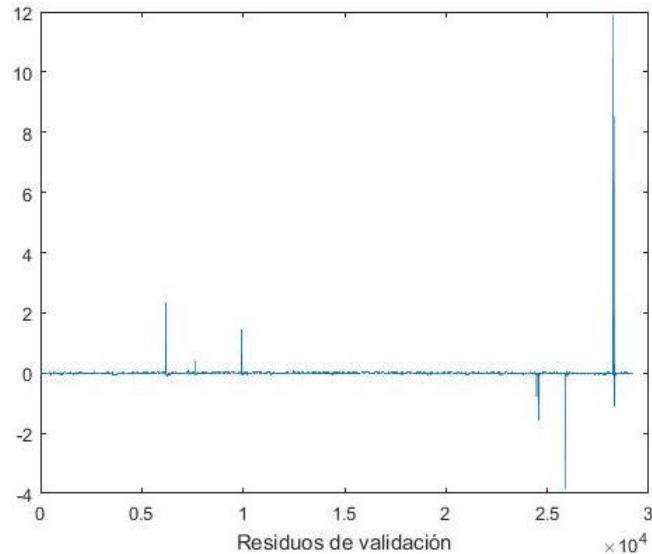


Figura 4.6. Residuos existentes entre la variable predicha por Ridge y la variable real.

Valor medio de residuo entrenamiento = 0,1534

Valor medio de residuo validación = 0,2897

Observando la gráfica se puede destacar que la diferencia entre el valor predicho y el valor real en todos los casos es prácticamente cero, lo cual indica una gran calidad de los resultados obtenidos. En lo que respecta a los valores medios calculados, éstos son bajos pero sobre todo son muy parecidos a los datos obtenidos con el anterior método de mínimos cuadrados. Como conclusión final destacamos que este método ofrece unos resultados muy similares al método de mínimos cuadrados. Sin embargo, a la hora de escoger entre uno u otro método, el elegido deberá ser Ridge ya que elimina el problema de la colinealidad y garantiza al cien por ciento la existencia de una solución.

4.4.1.3 Método Lasso

El siguiente método utilizado en este trabajo es el método Lasso, que es parecido al método Ridge. El objetivo de aplicar este método es comprobar si con la utilización de menos variables se puede obtener resultados mejores o similares a la hora de realizar estimaciones.

La curiosidad de este método es que mientras en otros métodos se obtiene un vector b , en éste se obtiene una matriz en la que cada columna reduce sus valores respecto a los de la columna anterior hasta que todas las variables de una columna son cero, siendo ésta la última columna de la matriz.

Para aplicar este método en Matlab, se utiliza la función `lasso` que tiene implementada el programa. La matriz b obtenida tras aplicar la función contiene 14 filas con 100 columnas, de las cuales haremos diferentes pruebas con ella mostrando a continuación los resultados obtenidos con 4 columnas de dicha matriz.

La primera columna que se utilizará será la columna 10, la cual solo utilizará 11 variables de las 14, otorgando el valor de 0 a las otras tres variables. En las figuras 4.7a y 4.7b se puede apreciar la variable DQO obtenida con los datos de entrenamiento, mientras que en las figuras 4.8a y 4.8b está calculada en función de los datos de validación.

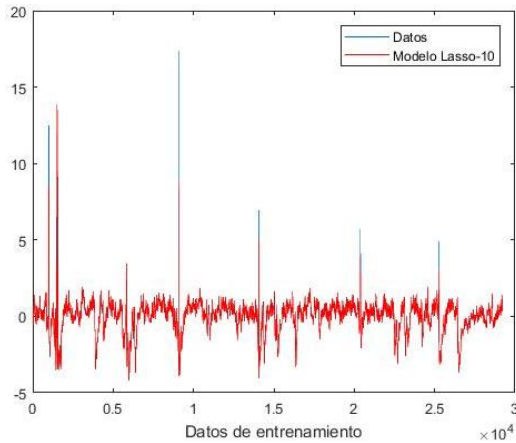


Figura 4.7a. Variable DQO de entrenamiento obtenido por Lasso-10.

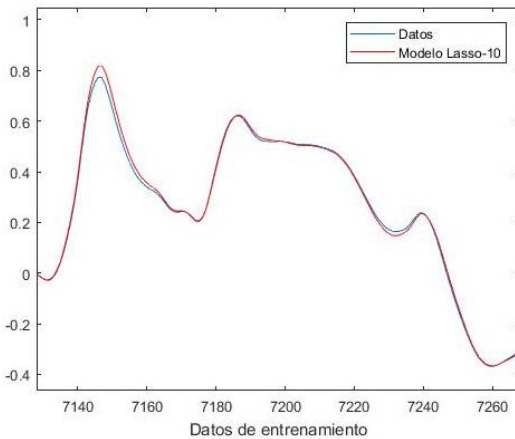


Figura 4.7b. Detalle de la variable DQO de entrenamiento obtenido por Lasso-10.

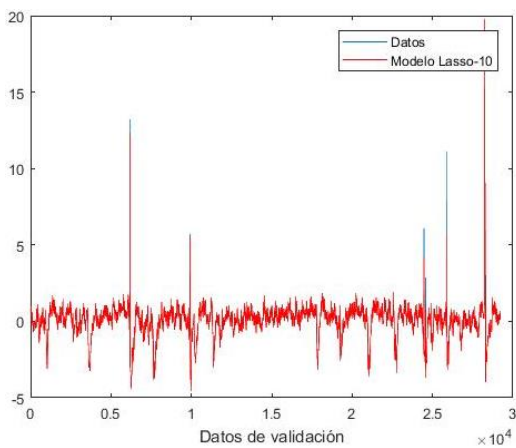


Figura 4.8a. Variable DQO de validación obtenido por Lasso-10.

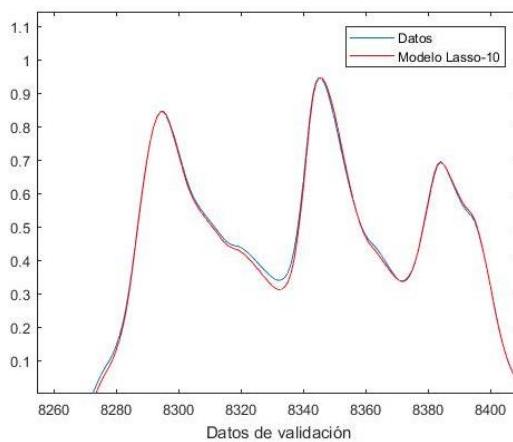


Figura 4.8b. Detalle de la variable DQO de validación obtenido por Lasso-10.

La siguiente columna que se va a utilizar de la matriz de valores b será la columna 25, la cual solo utilizará 10 variables y además habrá reducido el valor de las variables que se van a utilizar. En las figuras 4.9a y 4.9b se puede apreciar la variable DQO obtenida con los datos de entrenamiento, mientras que en las figuras 4.10a y 4.10b está calculada en función de los datos de validación.

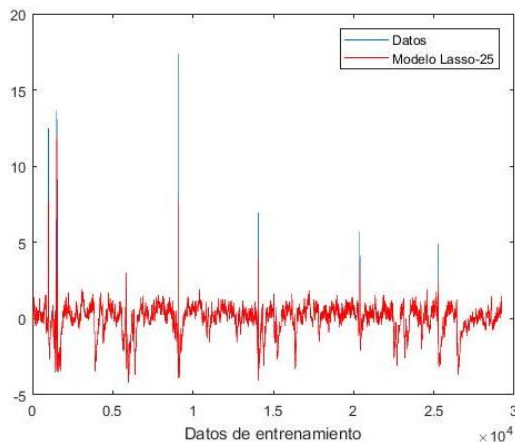


Figura 4.9a. Variable DQO de entrenamiento obtenida por Lasso-25.

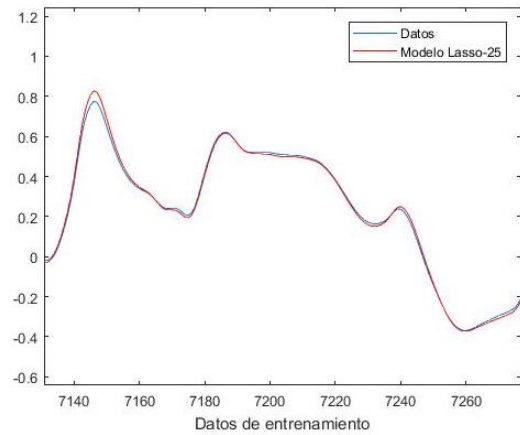


Figura 4.9b. Detalle de la variable DQO de entrenamiento obtenida por Lasso-25.

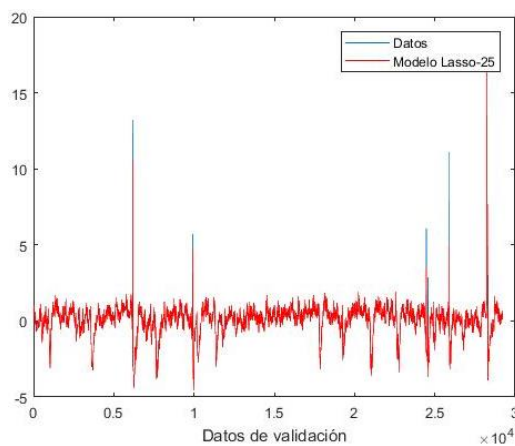


Figura 4.10a. Variable DQO de validación obtenida por Lasso-25.

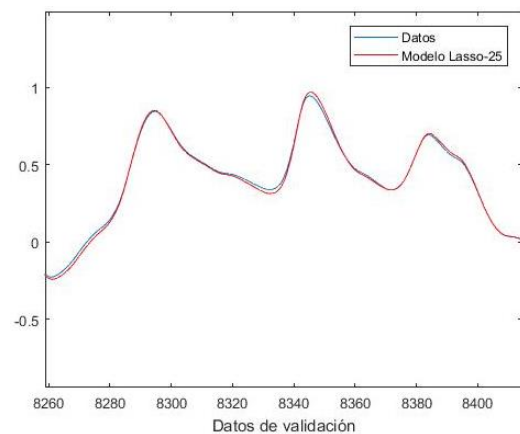


Figura 4.10b. Detalle de la variable DQO de validación obtenida por Lasso-25.

La tercera columna que se va a utilizar de la matriz de valores b será la columna 50, la cual ya presenta una importante reducción de variables que utiliza. En este caso solo se utilizarán 6 variables para calcular la predicción. Las figuras 4.11a y 4.11b nos presenta la variable DQO utilizando los datos de entrenamiento, mientras que en las figuras 4.12a y 4.12b está calculada la variable DQO en función de los datos de validación.

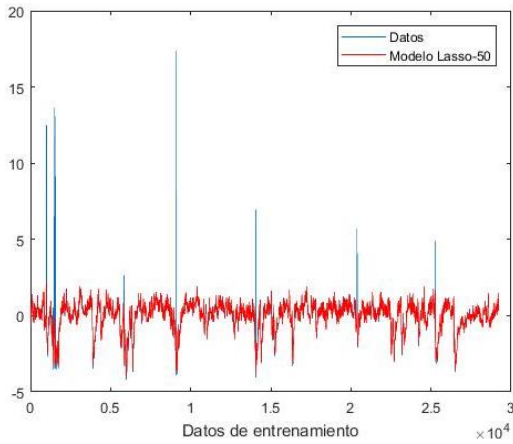


Figura 4.11a. Variable DQO de entrenamiento obtenido por Lasso-50.

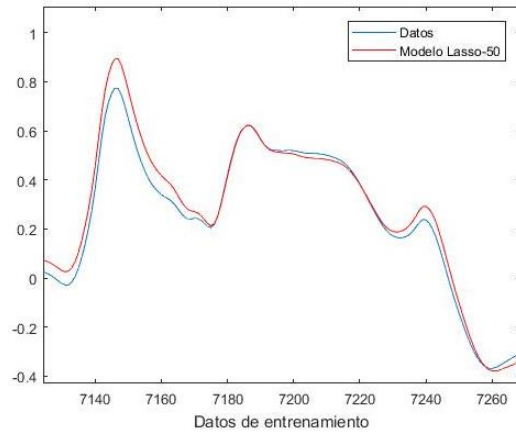


Figura 4.11b. Detalle de la variable DQO de entrenamiento obtenido por Lasso-50.

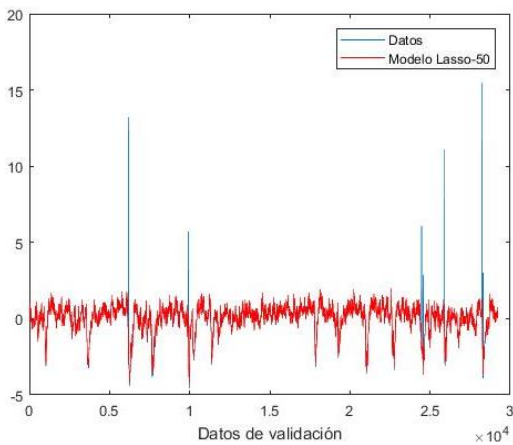


Figura 4.12a. Variable DQO de validación obtenido por Lasso-50.

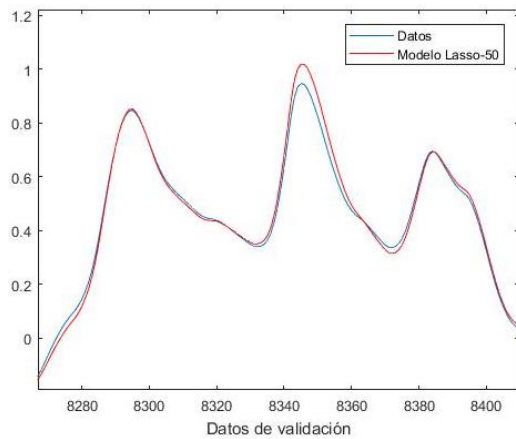


Figura 4.12b. Detalle de la variable DQO de validación obtenido por Lasso-50.

Finalmente la última columna que vamos a utilizar de la matriz de valores b será la columna 70, que únicamente tiene una variable con valor distinto de cero. En las figuras 4.13a y 4.13b la variable DQO está calculada con los datos de entrenamiento, y en las figuras 4.14a y 4.14b la variable predicha está calculada con los datos de validación.

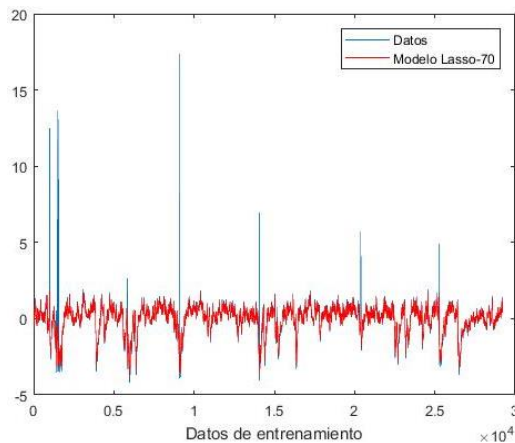


Figura 4.13a. Variable DQO de entrenamiento obtenida por Lasso-70.

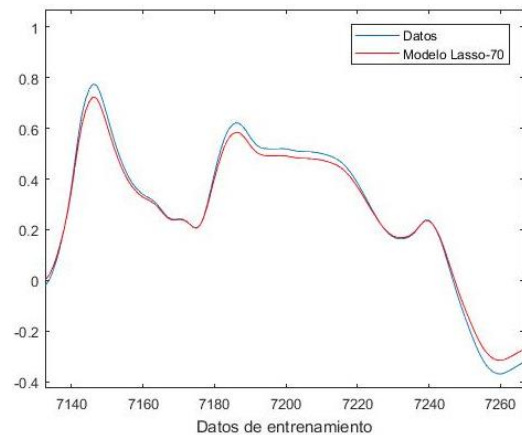


Figura 4.13b. Detalle de la variable DQO de entrenamiento obtenida por Lasso-70.

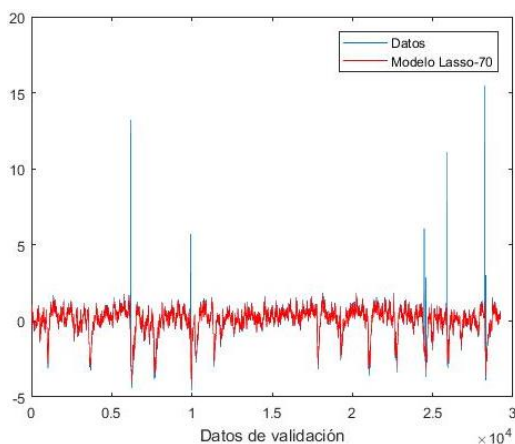


Figura 4.14a. Variable DQO de validación obtenida por Lasso-70.

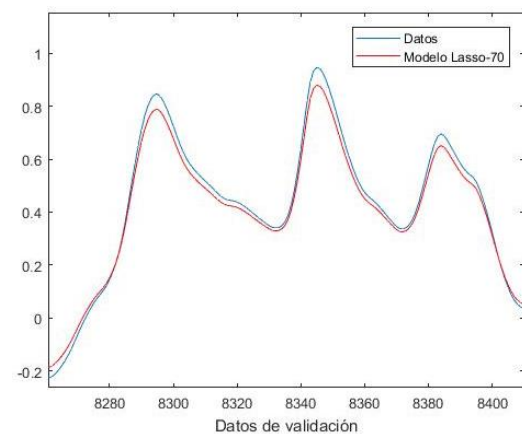


Figura 4.14b. Detalle de la variable DQO de validación obtenida por Lasso-70.

Una vez obtenidas las variables predichas en los diferentes casos que se ha decidido estudiar con Lasso se pueden sacar diferentes conclusiones. Las predicciones realizadas por Lasso-10 y Lasso 25 parecen de gran calidad, incluso se podría destacar que posiblemente sea mejor la aproximación de Lasso-25 en la que se utilizan menos variables. Por otro lado las estimaciones realizadas con Lasso-50 y Lasso-70 son peores predicciones, pero entre ellas dos no existe gran

diferencia y eso que en el último caso solo utiliza una variable. Tras esto último se puede concluir que la única variable utilizada con la columna 70 de la matriz de Lasso tiene un gran peso en el resultado final.

A pesar de estas conclusiones conviene analizar los residuos como se ha realizado con el resto de métodos para obtener una conclusión más objetiva. Los residuos se presentan a continuación:

- Valor medio de residuo entrenamiento Lasso-10 = 0,1653
- Valor medio de residuo validación Lasso-10= 0,2009

- Valor medio de residuo entrenamiento Lasso-25 = 0,1893
- Valor medio de residuo validación Lasso-25= 0,1652

- Valor medio de residuo entrenamiento Lasso-50 = 0,4803
- Valor medio de residuo validación Lasso-50= 0,4445

- Valor medio de residuo entrenamiento Lasso-70 = 0,4908
- Valor medio de residuo validación Lasso-70= 0,4524

Con la obtención de los valores medios de los residuos se pueden corroborar algunas de las conclusiones obtenidas tras la representación gráfica de los valores, y sacar también otras nuevas conclusiones. Mientras que los resultados de Lasso-10 y Lasso-25 resultan parecidos, se puede destacar una diferencia importante, y es que Lasso-25 saca unas aproximaciones mejores con datos nuevos. De hecho el resultado obtenido es muy bueno, tratándose del mejor resultado hasta el momento. Respecto a Lasso-50 y Lasso-70 aunque presentan peores resultados también cabe destacar el hecho de que presenta una mejor estimación con datos nuevos. Por lo tanto, la conclusión principal del método Lasso es que se suele obtener mejores conclusiones con datos nuevos, es decir, que “aprende” muy bien el modelo. También resulta interesante que muchas veces no por utilizar más datos se obtiene una mejor predicción.

4.4.1.4 Método PLS

El método de Mínimos Cuadrados Parciales o PLS (Partial Least Squares) consiste principalmente en una reducción del sistema para lo cual existen diferentes criterios, como el test del porcentaje de la varianza, test scree o la validación cruzada. En este caso se utilizará el test del porcentaje de la varianza, a través del cual se selecciona el número a de variables a utilizar de forma que explique un porcentaje de la varianza total.

La matriz s de tamaño 14×14 a la cual se aplicará el test de porcentaje de la varianza es la siguiente:

Observando las dos representaciones gráficas, tanto de la predicción obtenida con los datos de entrenamiento como con los datos de validación, se puede concluir que las predicciones logradas no son de gran calidad. Existe una gran diferencia entre la predicción y los valores recogidos de datos. Los valores medios de los residuos obtenidos son los siguientes:

Valor medio de residuo entrenamiento = 0,4704

Valor medio de residuo validación = 0,4234

Estos valores corroboran que la predicción obtenida no es de gran calidad. Lo único favorable de este método es que trabaja mejor con datos nuevos que con los datos conocidos.

4.4.1.5 Método PCA

El método PCA (Principal Component Analysis) tiene como objetivo reducir las variables a utilizar en un nuevo sistema de coordenadas. Para determinar el número de variables a utilizar se tomará el valor a obtenido en el método PLS anterior, cuyo valor es igual a cinco.

En las figuras 4.17a y 4.17b está representada la variable DQO tomada como dato junto con la predicha con los datos de entrenamiento, y en las figuras 4.18a y 4.18b se observa la variable DQO calculada en función de los datos de validación.

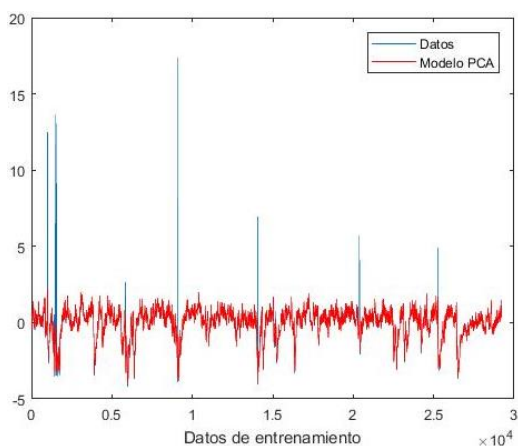


Figura 4.17a. Variable DQO de entrenamiento obtenido por PCA.

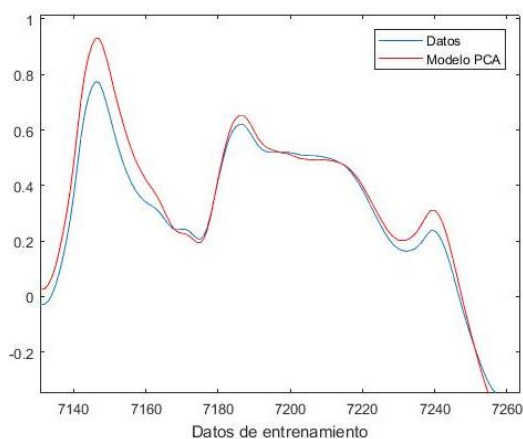


Figura 4.17b. Detalle de la variable DQO de entrenamiento obtenido por PCA.

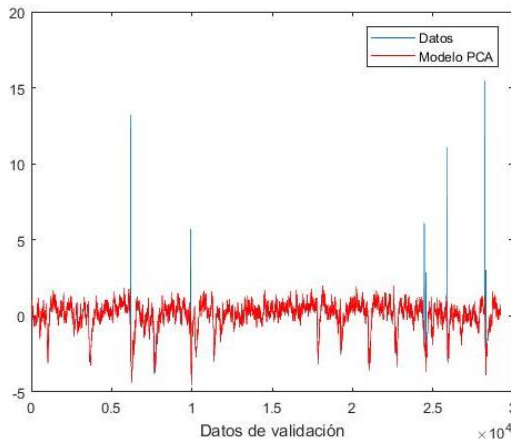


Figura 4.18a. Variable DQO de validación obtenido por PCA.

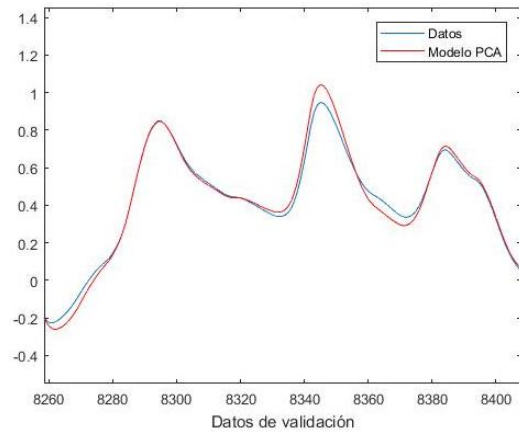


Figura 4.18b. Detalle de la variable DQO de validación obtenido por PCA.

Observando los resultados obtenidos en las figuras se aprecia una buena aproximación pero de peor calidad respecto a métodos anteriores utilizados. Para confirmar esta conclusión se procede a calcular y representar gráficamente los residuos medios del método aplicado. La figura 4.19 muestra la diferencia del valor predicho con el dato recogido de la planta de cada instante tomado.

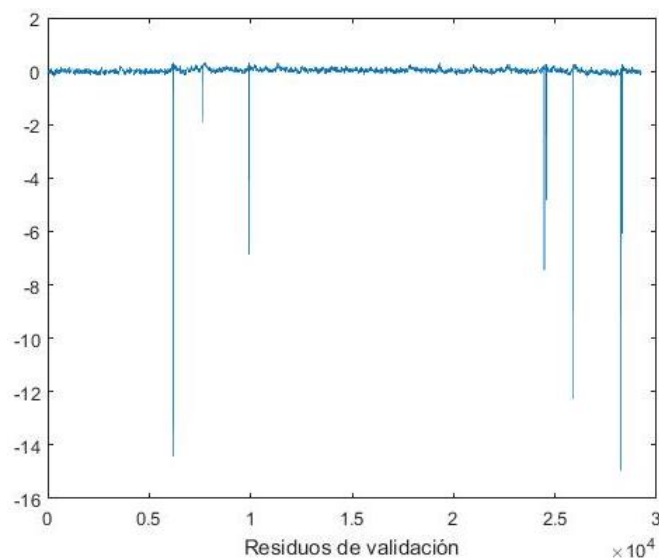


Figura 4.19. Residuos existentes entre la variable predicha por PCA y la variable real.

Valor medio de residuo entrenamiento = 0,4818

Valor medio de residuo validación = 0,4469

Si el valor del residuo fuese igual a cero en todos los datos recogidos el trazo de la figura 4.19 sería una línea recta de trazo fino. Sin embargo en este caso se puede observar como hay varios picos pequeños y algunos más grandes. Esto último verifica que la predicción varía bastante con los valores recogidos de datos, aspecto que también se aprecia con los valores medios de los residuos obtenidos.

4.4.1.6 Combinación de todos los métodos

Tras realizar estos diferentes métodos surge la idea de diseñar un modelo que tenga en cuenta todos los métodos calculados anteriormente. Este nuevo método consistirá en asociar un porcentaje a cada respuesta predicha obtenida en los diferentes métodos anteriores. El objetivo de este nuevo método es probar si trabaja mejor tras combinar los diferentes métodos, y observar cómo varía la respuesta en función del peso que se le otorgue a cada método.

En primer lugar se ha decidido realizar aplicar un peso igual a cada método en la respuesta final, es decir un 20% a cada método de forma que la respuesta se obtendrá a partir de la siguiente ecuación:

$$\hat{Y} = 0,2 \cdot \hat{Y}_{\text{mínimos cuadrados}} + 0,2 \cdot \hat{Y}_{\text{Ridge}} + 0,2 \cdot \hat{Y}_{\text{Lasso-25}} + 0,2 \cdot \hat{Y}_{\text{PLS}} + 0,2 \cdot \hat{Y}_{\text{PCA}}$$

Respecto al método Lasso ya que se tenía más de una respuesta posible se ha optado por escoger la respuesta obtenida con la columna 25 de la matriz de valores b ya que era la opción con la que se obtenía una mejor predicción en los datos de validación.

En las figuras 4.20a y 4.20b se encuentra representada la variable DQO predicha en función de los datos de entrenamiento junto con la variable recogida de datos reales. En las figuras 4.21a y 4.21b se puede visualizar la misma variable pero calculada con los datos de validación

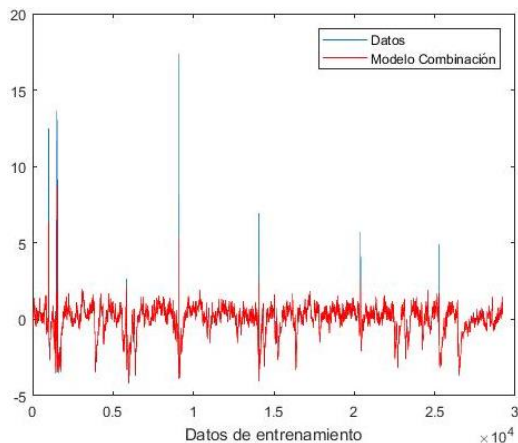


Figura 4.20a. Variable DQO de entrenamiento obtenido por combinación equitativa de métodos.

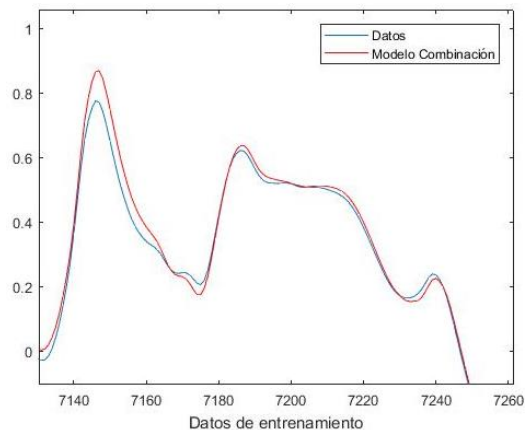


Figura 4.20b. Detalle de la variable DQO de entrenamiento obtenido por combinación equitativa de métodos.

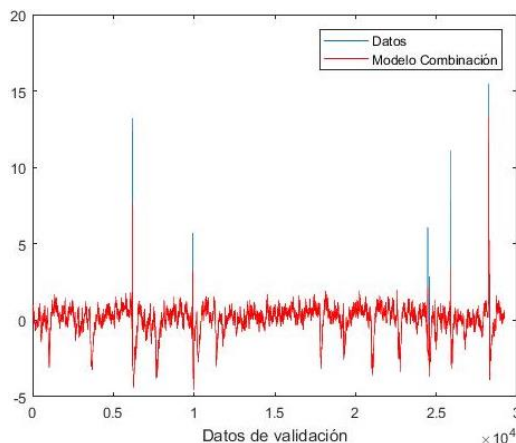


Figura 4.21a. Variable DQO de validación obtenido por combinación equitativa de métodos.

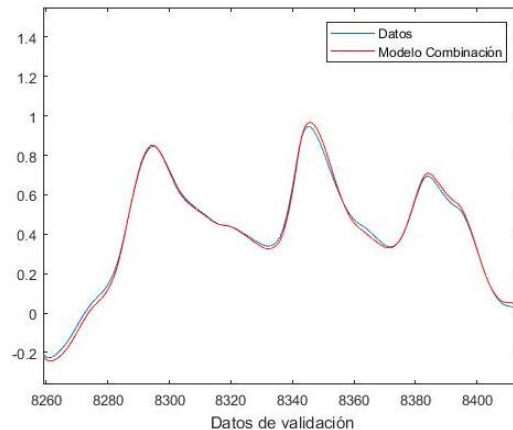


Figura 4.21b. Detalle de la variable DQO de validación obtenido por combinación equitativa de

Tras utilizar este método se observa una predicción de buena calidad con los datos de validación, mientras que con los datos de entrenamiento parece que se ajusta peor el sistema. Se puede decir que el sistema “aprende” bastante bien el sistema, esto también puede ser debido a que tres de los métodos incluidos en la combinación trabajaban mejor con datos nuevos que con los datos de entrenamiento.

El valor medio de los residuos obtenido de las predicciones realizadas anteriormente son los siguientes valores:

Valor medio de residuo entrenamiento = 0,2496

Valor medio de residuo validación = 0,1613

Los valores medios de los residuos obtenidos corroboran las impresiones a las que se había llegado anteriormente. Lo más destacable es que la predicción realizada con los datos de validación trabaja mejor que con los datos de entrenamiento. De hecho el valor medio del residuo de datos de validación es el valor más bajo que se ha obtenido, es decir, se trata de la mejor predicción realizada hasta el momento.

Visto los buenos resultados logrados con la combinación de métodos otorgando el mismo peso a cada método ahora se realizará una combinación de los métodos otorgando diferente peso a cada método. En este reparto de porcentaje se decidirá otorgar un mayor peso a aquellos métodos que trabajan mejor con datos de validación y un menor peso a aquellos métodos con los que se ha obtenido una peor predicción. En la siguiente tabla se puede observar el valor medio de residuo obtenido y el peso que se le ha decidido otorgar a cada método:

Método	Valor medio de residuo	Porcentaje designado
Mínimos cuadrados	0,2674	20%
Ridge	0,2897	20%
Lasso-25	0,1652	40%
PLS	0,4234	10%
PCA	0,4469	10%

Tras asignar el porcentaje para cada método, la ecuación a utilizar para obtener la predicción será:

$$\hat{Y} = 0,2 \cdot \hat{Y}_{\text{mínimos cuadrados}} + 0,2 \cdot \hat{Y}_{\text{Ridge}} + 0,4 \cdot \hat{Y}_{\text{Lasso-25}} + 0,1 \cdot \hat{Y}_{\text{PLS}} + 0,1 \cdot \hat{Y}_{\text{PCA}}$$

Una vez planteada la ecuación con la que se calculará la nueva predicción, a continuación se simulará el nuevo método. Por una parte en las figuras 4.22a y 4.22b estará representada la variable DQO con los datos de entrenamiento junto con la variable recogida de los datos reales, en las figuras 4.23a y 4.23b estará representada la misma variable calculada con los datos de validación.

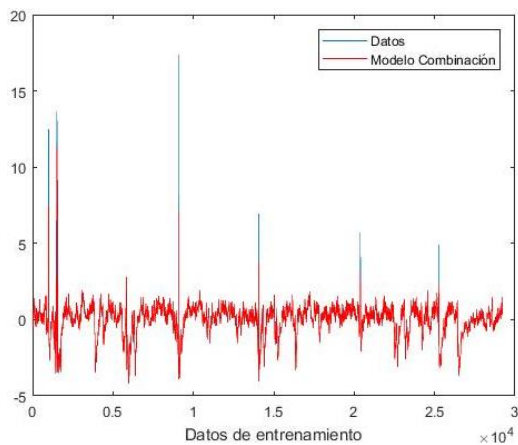


Figura 4.22a. Variable DQO de entrenamiento obtenido por combinación de métodos.

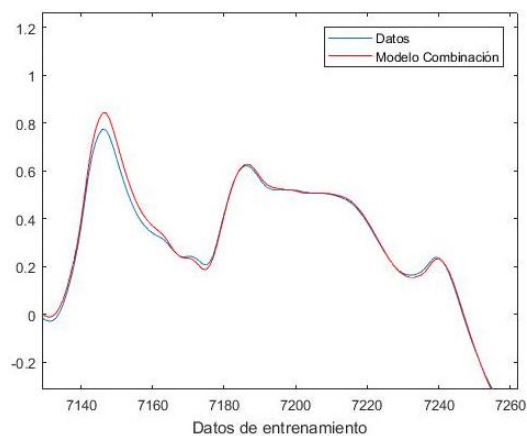


Figura 4.22b. Detalle de la variable DQO de entrenamiento obtenido por combinación de métodos.

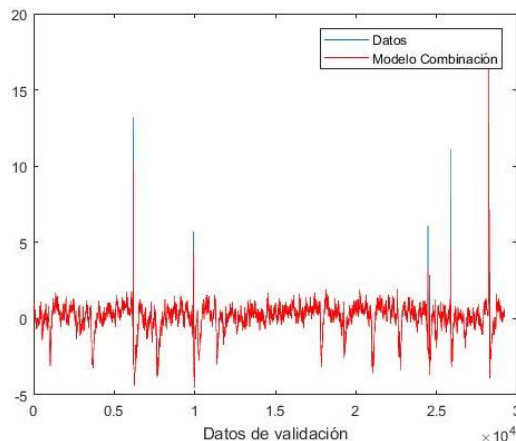


Figura 4.23a. Variable DQO de validación obtenida por combinación de métodos.

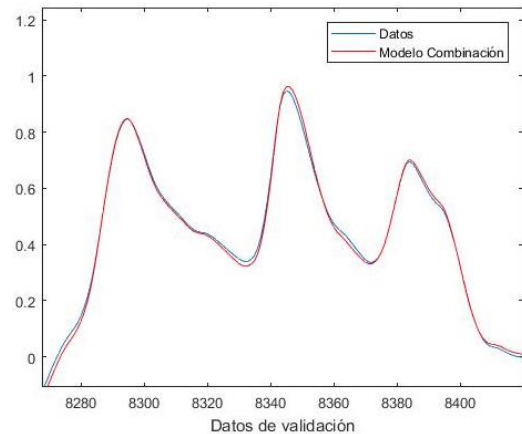


Figura 4.23b. Detalle de la variable DQO de validación obtenida por combinación de métodos.

A primera vista parece que los resultados obtenidos tras utilizar esta combinación diferente de los métodos son similares a los logrados con la combinación equitativa de cada método. Aun así parece que esta predicción se ajusta un poco más a los datos reales, así que para poder sacar una conclusión final se hallan los valores medios de los residuos.

Valor medio de residuo entrenamiento = 0,2004

Valor medio de residuo validación = 0,1563

Efectivamente esta combinación que otorga un peso diferente a cada método proporciona una predicción de mejor calidad que la combinación equitativa realizada antes. Esto se podría decir que era previsible, ya que se otorga más peso al resultado final del método que trabaja mejor y menos peso a los métodos que trabajan peor. A pesar de esto tampoco existe una gran diferencia entre los resultados de una combinación y otra.

Otro hecho que cabe destacar es que con las dos combinaciones de métodos propuestas se obtiene una mejor predicción que utilizando únicamente el método que nos ofrece una mejor respuesta hasta el momento. Lo lógico era pensar que al incluir otros métodos que trabajaban peor empeorase la mejor predicción calculada hasta el momento. Esto seguramente suceda debido a que no todos los métodos utilizan el mismo número de variables a la hora de hallar el resultado final.

4.4.1.7 Comparativa entre métodos aplicados al modelo estático

Después de aplicar los diferentes métodos propuestos sobre el modelo estático, se considera adecuado realizar una comparación entre ellos de forma que se pueda apreciar las ventajas y desventajas de cada método.

Como el estudio aplicado se basa en el valor final de la predicción, se ha decidido hacer una tabla con los diferentes valores medios de los residuos obtenidos, tanto

de los datos de entrenamiento como de los datos de validación. La tabla se presenta a continuación en la que a través de un vistazo se observan todos los residuos.

Método aplicado	Valor medio del residuo en entrenamiento	Valor medio del residuo en validación
Mínimos cuadrados	0.1516	0.2674
Ridge	0.1534	0.2897
Lasso-10	0.1653	0.2009
Lasso-25	0.1893	0.1652
Lasso-50	0.4803	0.4445
Lasso-70	0.4908	0.4524
PLS	0.4704	0.4234
PCA	0.4818	0.4469
Combinación equitativa de métodos	0.2496	0.1613
Combinación porcentual de métodos	0.2004	0.1563

El valor medio del residuo obtenido con los datos de validación será el dato que cobrará importancia a la hora de valorar si un método es mejor que otro, debido a que siempre se trabajará con datos nuevos a la hora de hacer una predicción. Por lo tanto el método que ofrece una mejor predicción es el que se ha propuesto como combinación de los diferentes métodos utilizados. De los dos métodos calculados por combinación el que funciona mejor es al que se le ha asignado un porcentaje diferente a cada método en función de la calidad del método.

Cabe destacar también de las predicciones obtenidas por los métodos de combinación que se obtienen mejores predicciones que el resto, cuando de primera se podría pensar que al combinar métodos que funcionan bien con otros que funcionan peor, la predicción final tendría una calidad intermedia entre las buenas y malas predicciones.

Los datos predichos por estos dos métodos que usan la combinación tampoco difieren mucho del obtenido por el denominado Lasso-25, que utiliza la columna 25 de la matriz b de Lasso. Lo que resulta más llamativo de este método es que trabaja mejor con menos variables que Lasso-10, que a priori se supone que contiene más información. Lasso-10 trabaja con los datos de entrenamiento mejor

que con Lasso-25, pero éste último “aprende” mejor el sistema y ofreciendo una buena predicción con datos nuevos.

Observando los valores medios de los residuos de todos los métodos también llama la atención que la mayoría de los métodos trabaja mejor con datos nuevos a excepción de los métodos de Mínimos cuadrados, Ridge y Lasso-10. Esto es favorable, debido a que se supone que los métodos siempre trabajarán en la planta depuradora de aguas residuales con datos nuevos.

Finalmente, decir que aunque los métodos de Mínimos cuadrados y Ridge trabajen prácticamente igual, ofreciendo una predicción parecida, siempre será mejor utilizar el método Ridge. Al utilizar el método de Mínimos cuadrados existe siempre la posibilidad que aparezca un problema debido a la colinealidad de la muestra de datos que consigue eliminar el método Ridge. De esta forma si se utiliza siempre Ridge se garantiza tener una predicción.

4.4.2 Modelo dinámico

Todos los métodos utilizados hasta este momento han sido métodos estáticos. En este apartado se realizará la predicción de la variable deseada, DQO, teniendo en cuenta instantes anteriores del proceso denominándose método dinámico. La principal diferencia consiste en que el tamaño de la matriz de datos aumentará, y con ello la información que se posee.

El objetivo de este método es que al contar con más información del pasado proporcione una mejor predicción. Además, como los procesos industriales son dinámicos, introducir la dinámica del sistema en los métodos de predicción mejorara el resultado. Para ello a la nueva matriz de datos se aplicarán los mismos métodos que al sistema estático pudiendo así observar las diferencias que aparecen entre los dos.

Para la nueva matriz de datos se realizará también la reducción de tamaño inicial antes de aplicar los diferentes métodos, como se ha realizado con los modelos estáticos. Tras disminuir el tamaño de la matriz, éste será de 58465×57 , es decir, se trabajará con 57 variables en este caso. Al igual que con el modelo estático los datos se dividirán en dos grupos, datos de entrenamiento y datos de validación.

4.4.2.1 Método de mínimos cuadrados

El primer método que se aplicará al nuevo conjunto de datos sobre el que se va a trabajar va ser el método de mínimos cuadrados. En las figuras 4.24a y 4.24b está representada la variable DQO predicha con los datos de entrenamiento, y en las figuras 4.25a y 4.25b la variable predicha con los datos de validación.

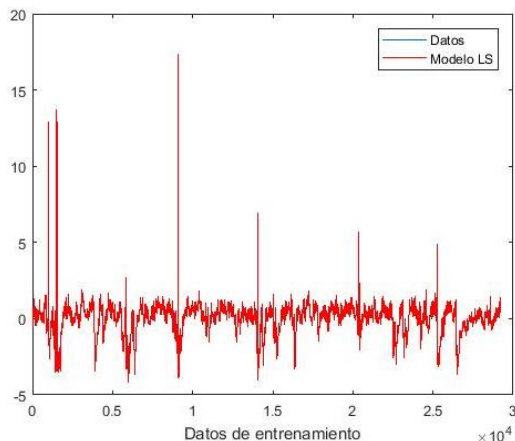


Figura 4.24a. Variable DQO de entrenamiento obtenido por mínimos cuadrados dinámico.

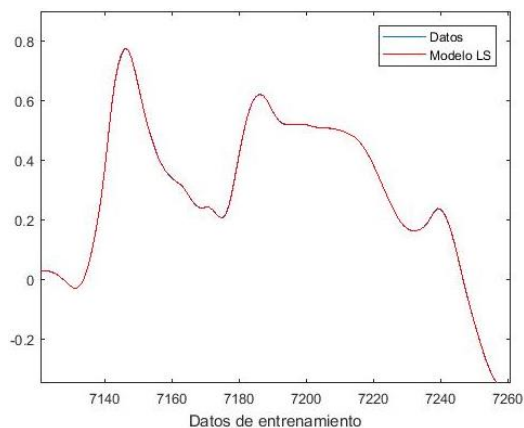


Figura 4.24b. Detalle de la variable DQO de entrenamiento obtenido por mínimos cuadrados dinámico.

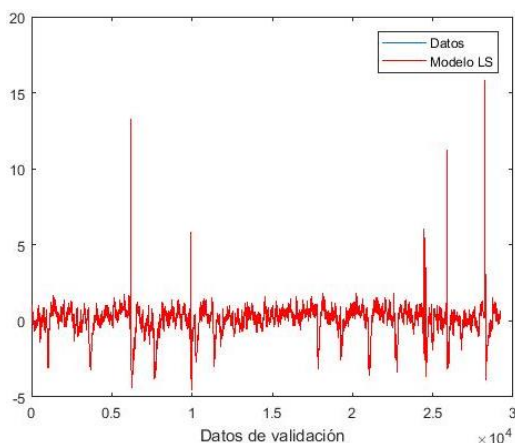


Figura 4.25a. Variable DQO de validación obtenido por mínimos cuadrados dinámico.

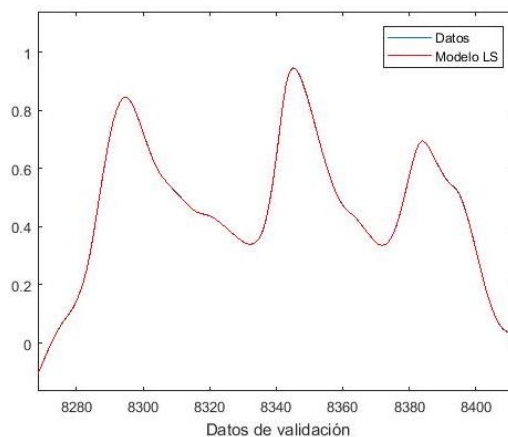


Figura 4.25b. Detalle de la variable DQO de validación obtenido por mínimos cuadrados

Las figuras muestran una aproximación lograda de gran calidad en la que apenas se aprecia diferencia entre la predicción y la variable real. La figura 4.26 muestra el residuo de cada instante tomado entre la predicción y sus datos reales, además los valores medios de los residuos son los siguientes:

Valor medio de residuo entrenamiento = 0,0233

Valor medio de residuo validación = 0,0241

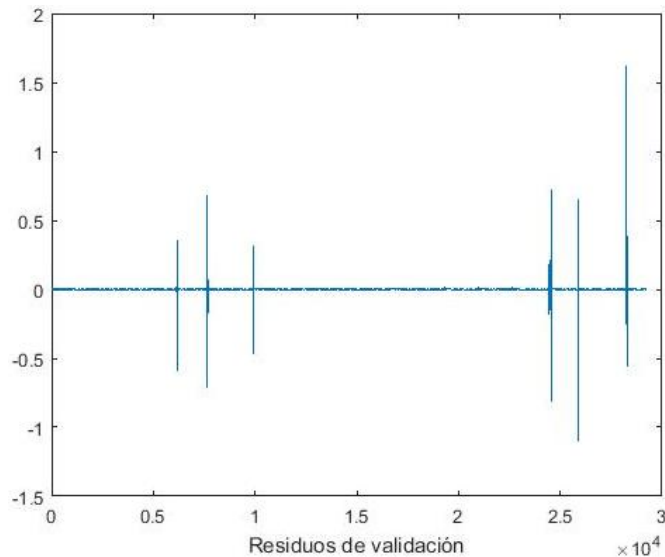


Figura 4.26. Residuos existentes entre la variable predicha por mínimos cuadrados y la variable real.

Efectivamente, tras visualizar los valores medios de los residuos se corrobora que la predicción se aproxima a los valores reales de los datos recogidos. Además en la figura 4.26 se observa muy bien como el valor del residuo es prácticamente una línea recta situada sobre cero a excepción de seis picos, representando a la perfección la igualdad entre la predicción y los valores reales.

También se puede decir que apenas hay diferencias entre el resultado obtenido con los datos de entrenamiento y con los datos de validación.

Como punto negativo que aparece al aplicar este método con el modelo dinámico es que aparece también un Warning relacionado con la colinealidad de los datos de la muestra como aparecía con el modelo estático. Esto indica que aunque en este caso se consigue obtener una predicción, es posible que con otros datos de un error el sistema. Con la finalidad de evitar este posible problema se propondrá el método Ridge a continuación que debería eliminar dicha situación.

4.4.2.2 Método Ridge

A continuación se aplicará el método Ridge con el objetivo de eliminar el problema de colinealidad que ha aparecido al utilizar el método de mínimos cuadrados. En las figuras 4.27a y 4.27b se observa la variable DQO calculada con los datos de entrenamiento, mientras que en las figuras 4.28a y 4.28b aparecerá la misma variable calculada con los datos de validación.

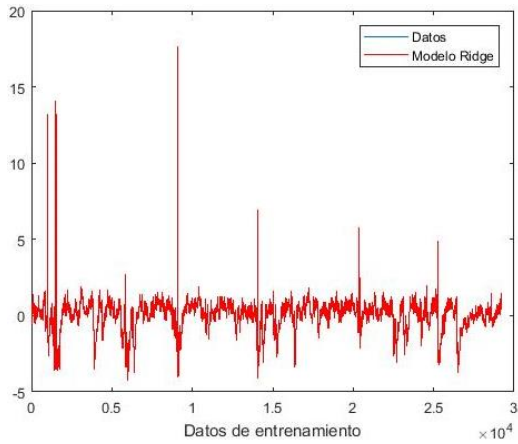


Figura 4.27a. Variable DQO de entrenamiento obtenido por Ridge dinámico.

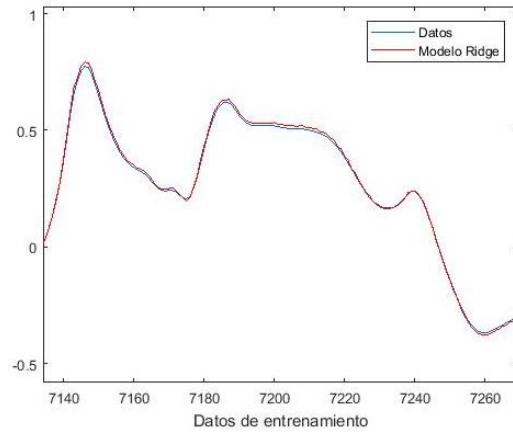


Figura 4.27b. Detalle de la variable DQO de entrenamiento obtenido por Ridge dinámico.

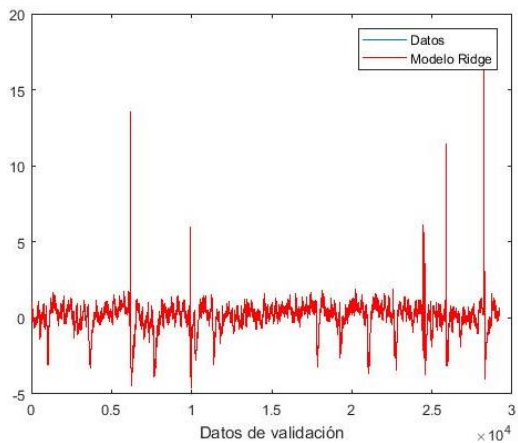


Figura 4.28a. Variable DQO de validación obtenido por Ridge dinámico.

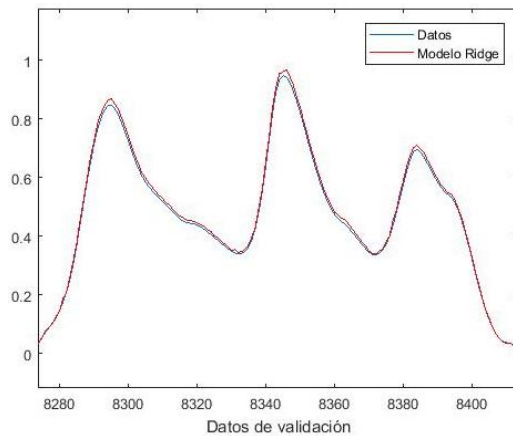


Figura 4.28b. Detalle de la variable DQO de validación obtenido por Ridge dinámico.

En este caso tras aplicar Ridge se obtiene también una predicción de gran calidad que se aproxima prácticamente a la perfección con la variable recogida de los datos. Los residuos de cada instante están representados en la figura 4.29 así como el valor medio de los residuos de entrenamiento y validación a continuación:

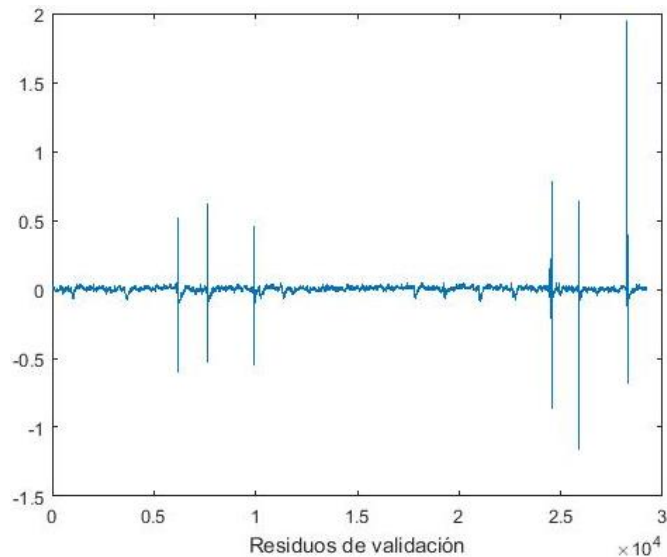


Figura 4.29. Residuos existentes entre la variable predicha por Ridge y la variable real.

Valor medio de residuo entrenamiento = 0,0323

Valor medio de residuo validación = 0,0362

Esta predicción efectivamente es de una gran calidad ajustándose mucho a los datos reales, aunque un poco peor que el método de mínimos cuadrados. En cambio, este método transmite la seguridad de ofrecer una respuesta en todos los casos debido a que elimina el posible problema que podía aparecer en el método de mínimos cuadrados.

4.4.2.3 Método Lasso

Al aplicar la función Lasso a la muestra de datos del modelo dinámico, se obtiene una matriz b , compuesta por 57 filas y 100 columnas. Los valores de cada columna son menores que los valores de la columna anterior, es decir, los valores de las columnas van disminuyendo hasta que en la última columna, la 100, todos los valores son iguales a cero.

La primera predicción que se va a calcular se realizará con la columna 10 de la matriz b . Esta columna a la hora de calcular la predicción solo utiliza 22 variables de las 57 que dispone la muestra de datos. Para referirse a este método que utiliza la columna 10 se denominará Lasso-10. En la figura 4.30a y 4.30b está representada la variable DQO predicha con los datos de entrenamiento, y en la figura 4.31a y 4.31b está representada con los datos de validación.

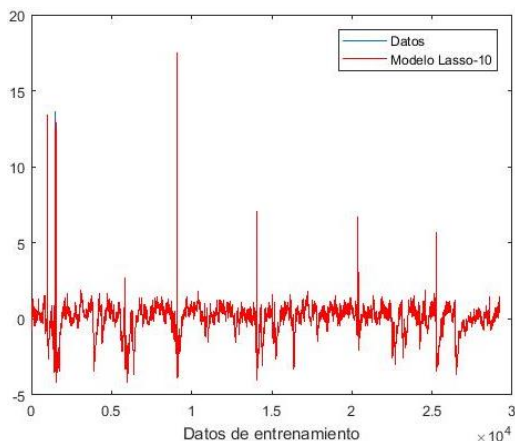


Figura 4.30a. Variable DQO de entrenamiento obtenido por Lasso-10 dinámico.

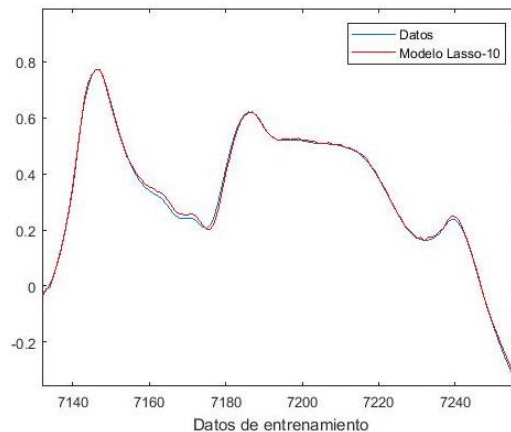


Figura 4.30b. Detalle de la variable DQO de entrenamiento obtenido por Lasso-10 dinámico.

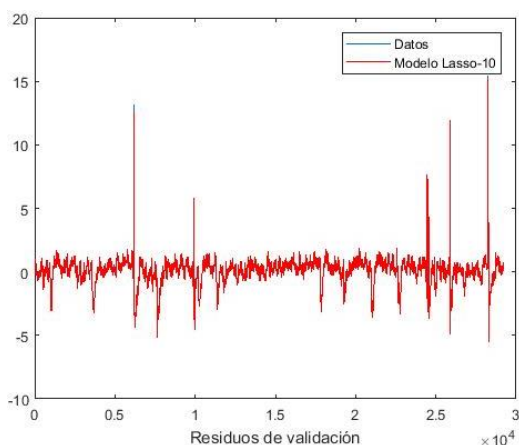


Figura 4.31a. Variable DQO de validación obtenido por Lasso-10 dinámico.

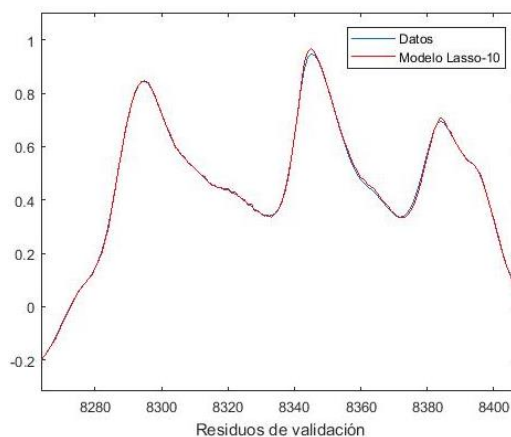


Figura 4.31b. Detalle de la variable DQO de validación obtenido por Lasso-10 dinámico.

La columna que se va a utilizar para realizar la siguiente predicción será la columna número 50. Se pasa de utilizar 22 variables en la predicción anterior a utilizar en este caso únicamente 4, lo que significa una gran disminución en el número de variables usadas. La predicción calculada con los datos de entrenamiento está representada en las figuras 4.32a y 4.32b junto con los datos recogidos de la planta, y en las figuras 4.33a y 4.33b la predicción está calculada con los datos de validación.

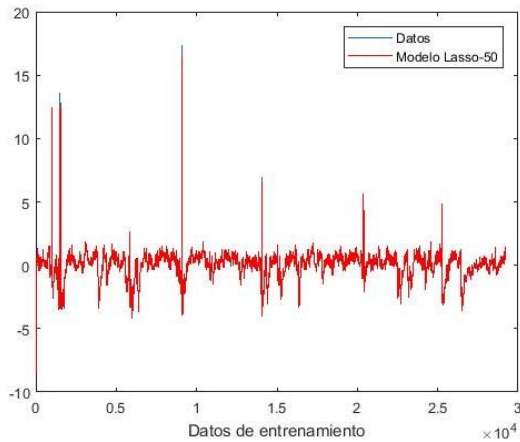


Figura 4.32a. Variable DQO de entrenamiento obtenido por Lasso-50 dinámico.

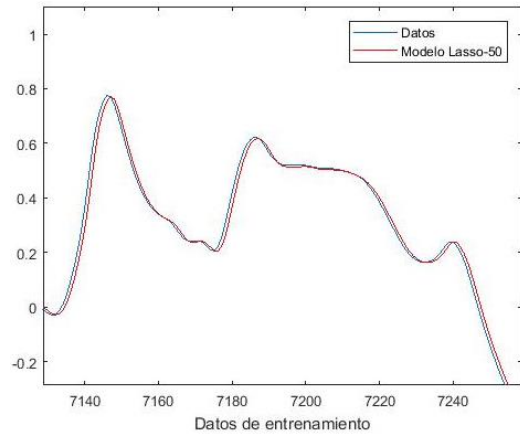


Figura 4.32b. Detalle de la variable DQO de entrenamiento obtenido por Lasso-50 dinámico.

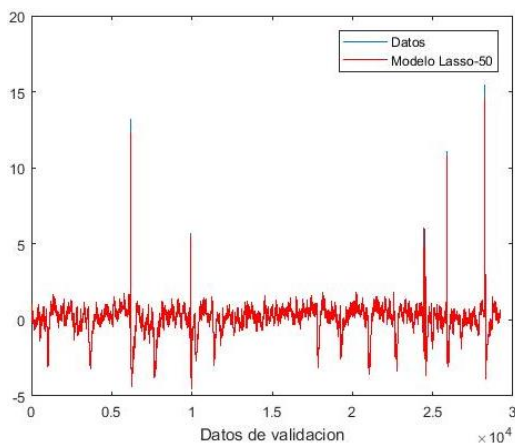


Figura 4.33a. Variable DQO de validación obtenido por Lasso-50 dinámico.

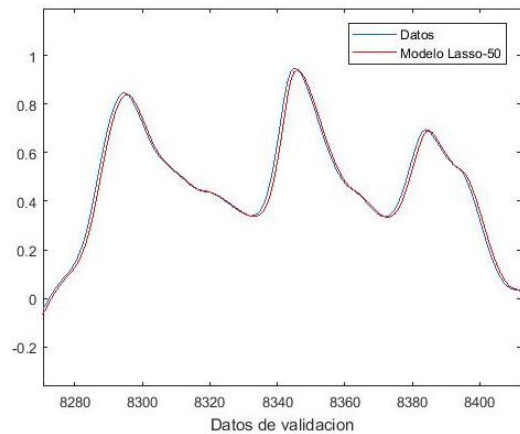


Figura 4.33b. Detalle de la variable DQO de validación obtenido por Lasso-50 dinámico.

Finalmente, la última columna que se utilizará será la columna 74, que utiliza una única variable para calcular la predicción. A continuación se encuentran las figuras 4.34a y 4.34b en las que está representada la predicción calculada con los datos de entrenamiento, y en las figuras 4.35a y 4.35b la variable está calculada con los datos de validación.

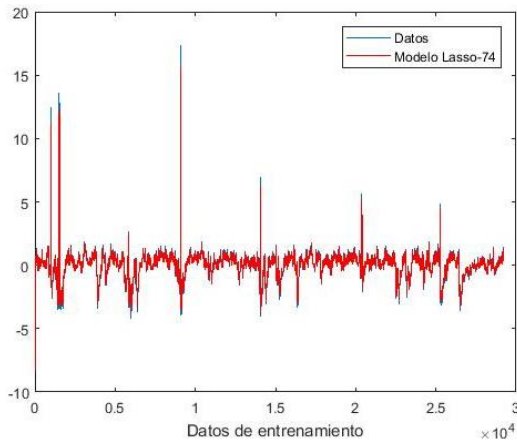


Figura 4.34a. Variable DQO de entrenamiento obtenido por Lasso-74 dinámico.

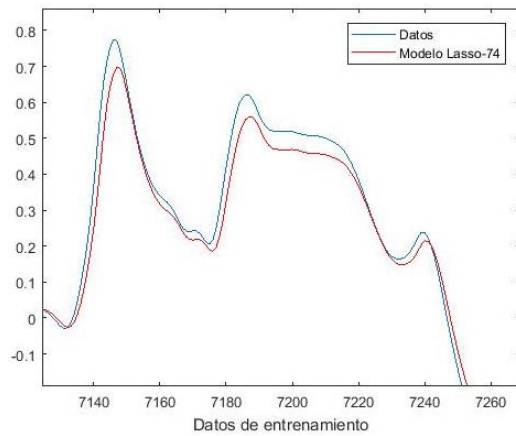


Figura 4.34b. Detalle de la variable DQO de entrenamiento obtenido por Lasso-74 dinámico.

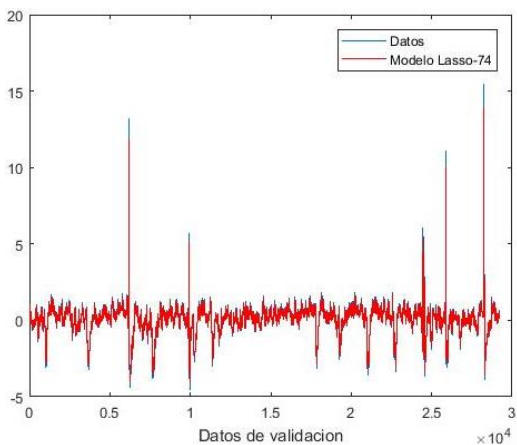


Figura 4.35a. Variable DQO de validación obtenido por Lasso-74 dinámico.

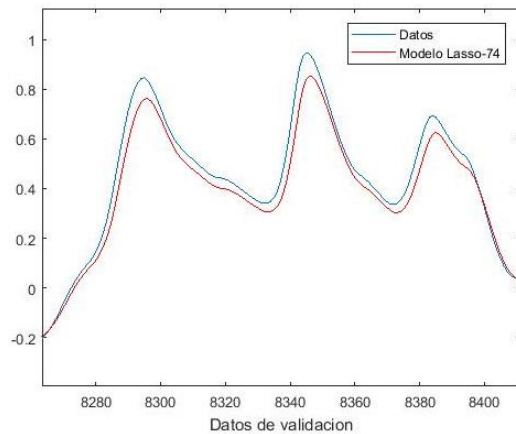


Figura 4.35b. Detalle de la variable DQO de validación obtenido por Lasso-74 dinámico.

Vistas las tres simulaciones se puede observar que a medida que se trabaja con menos variables se obtiene peores predicciones. Para apreciar numéricamente esta observación, se presentan los valores medios de los residuos obtenidos en los tres casos de Lasso que se han simulado:

- Valor medio de residuo entrenamiento Lasso-10 = 0,0707
- Valor medio de residuo validación Lasso-10 = 0,0685
- Valor medio de residuo entrenamiento Lasso-50 = 0,0867
- Valor medio de residuo validación Lasso-50 = 0,0754
- Valor medio de residuo entrenamiento Lasso-74 = 0,1668
- Valor medio de residuo validación Lasso-74 = 0,1447

Las predicciones obtenidas con Lasso son de mejor calidad cuando más variables se utilizan según estos tres casos vistos. La mejor cualidad de este método es que “aprende” bien el modelo, obteniendo mejores resultados con los datos de validación que con los datos de entrenamiento. Finalmente se puede observar como las predicciones obtenidas definen bien el recorrido de los datos recogidos de la planta pero obteniendo valores más pequeños, siempre por debajo de los datos reales en las gráficas de las figuras.

4.4.2.4 Método PLS

Como se ha visto anteriormente el método PLS realiza una reducción de variables a la hora de predecir la variable deseada. Para elegir el número de variables que se van a utilizar en este caso, se seguirá utilizando el test del porcentaje de la varianza como en el modelo estático. Habrá que tener cuenta que el número de variables de la muestra del modelo dinámico no es el mismo que el utilizado antes en el modelo estático. Por lo tanto, si se elige el mismo porcentaje de antes no se trabajará con el mismo número de variables.

Se realizarán tres pruebas de este modelo dinámico con el método PLS, uno en el que se aplicará el mismo porcentaje que en el modelo estático, otro en el que se aplicara el porcentaje adecuado para trabajar con las mismas variables que se han usado en el modelo estático, y otro en el que se aplicará un porcentaje más bajo con el objetivo de trabajar con menos variables.

El primer caso a estudiar se trabajará con el 98,5% como porcentaje. Aplicando este porcentaje se usarán 6 variables, mientras que aplicando este porcentaje en el estático se utilizaba una variable menos, 5 en total. En las figuras 4.36a y 4.36b está representada la variable DQO predicha con los datos de entrenamiento, y en las figuras 4.37a y 4.37b está predicha con los datos de validación.

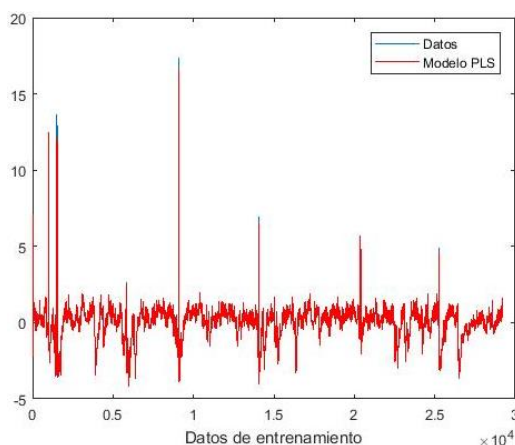


Figura 4.36a. Variable DQO de entrenamiento obtenido por PLS (98,5%) dinámico.

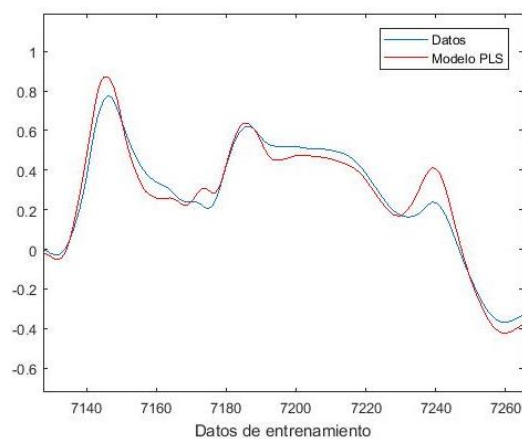


Figura 4.36b. Detalle de la variable DQO de entrenamiento obtenido por PLS (98,5%) dinámico.

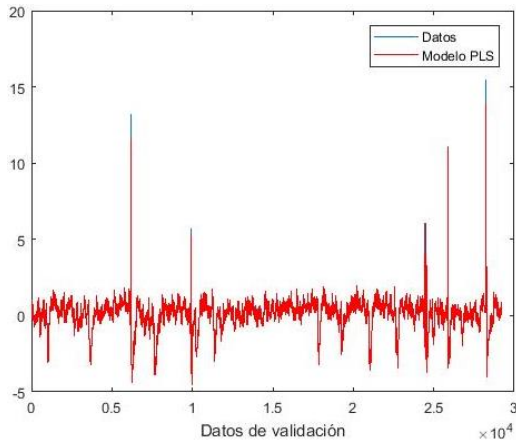


Figura 4.37a. Variable DQO de validación obtenido por PLS (98,5%) dinámico.

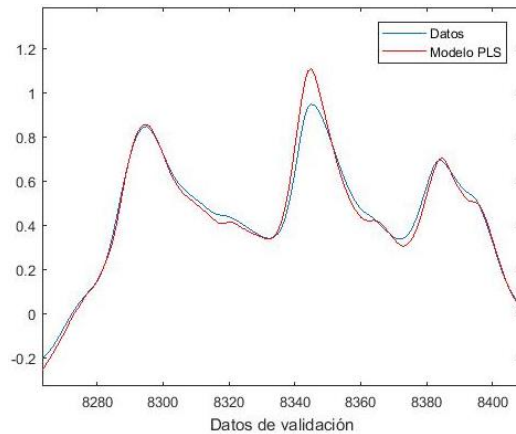


Figura 4.37b. Detalle de la variable DQO de validación obtenido por PLS (98,5%) dinámico.

De esta primera predicción calculada se puede concluir que el método aplicado “aprende” bastante bien el modelo proporcionando un mejor resultado con los datos de validación que con los de entrenamiento. Aun así, la predicción parece de peor calidad que las calculadas con otros métodos del modelo dinámico.

Ahora se realizará la segunda predicción con el método PLS para el que se aplicará un porcentaje de un 97,4 %. Con este porcentaje la predicción trabaja con 5 variables, como se ha llevado a cabo en el modelo estático. En las figuras 4.38a y 4.38b está representada la variable DQO calculada con los datos de entrenamiento, mientras que en las figuras 4.39a y 4.39b está la variable predicha con los datos de validación.

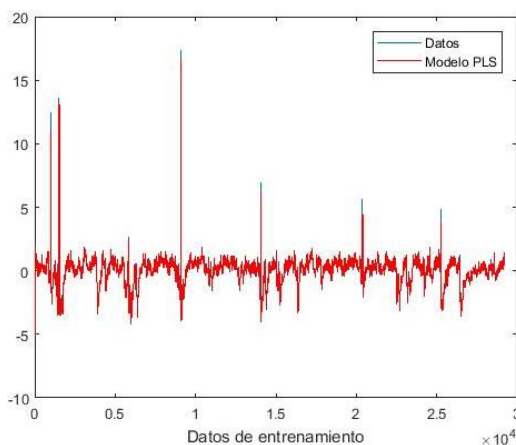


Figura 4.38a. Variable DQO de entrenamiento obtenido por PLS (97,4%) dinámico.

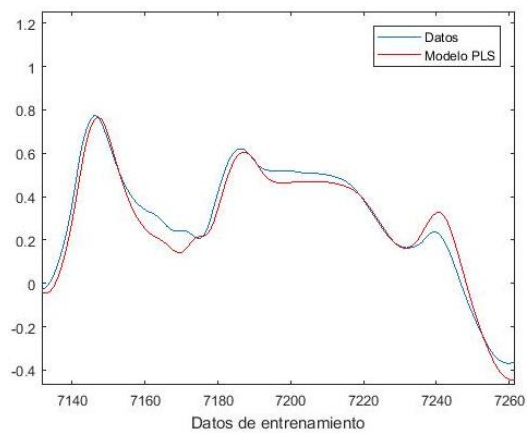


Figura 4.38b. Detalle de la variable DQO de entrenamiento obtenido por PLS (97,4%) dinámico.

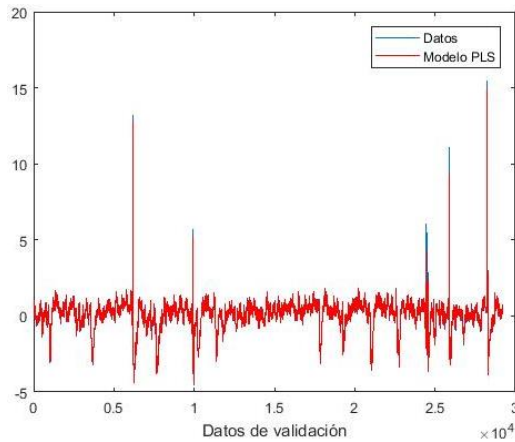


Figura 4.39a. Variable DQO de validación obtenido por PLS (97,4%) dinámico.

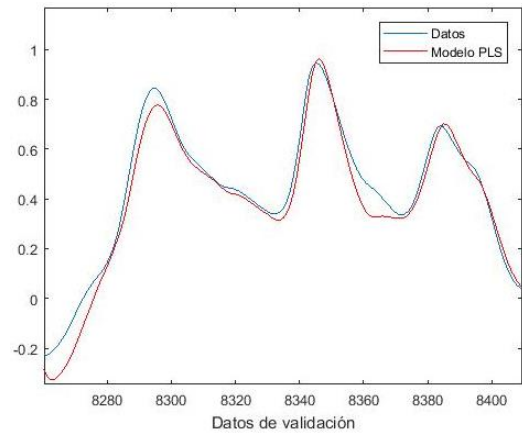


Figura 4.39b. Detalle de la variable DQO de validación obtenido por PLS (97,4%) dinámico.

Se puede observar que al utilizar cinco variables la predicción es peor que en el caso anterior, donde se utilizaban seis variables. Finalmente, queda estudiar la última situación donde se trabajará con el 91%, es decir, se utilizarán únicamente tres variables. Las predicciones obtenidas están representadas en las figuras 4.40 y 4.41.

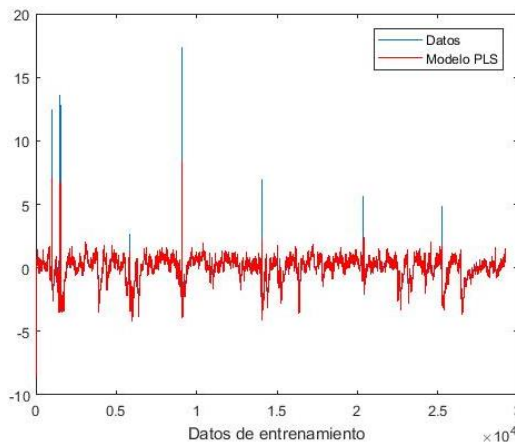


Figura 4.40a. Variable DQO de entrenamiento obtenido por PLS (91%) dinámico.

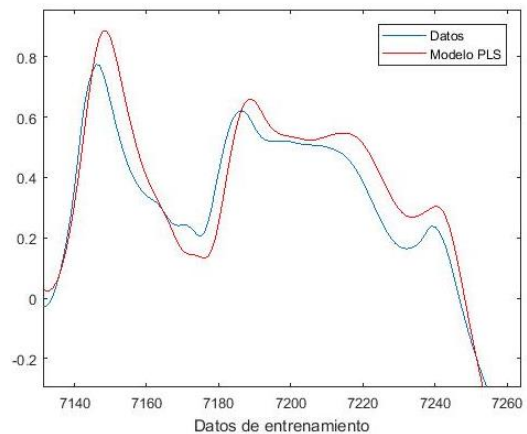


Figura 4.40b. Detalle de la variable DQO de entrenamiento obtenido por PLS (91%) dinámico.

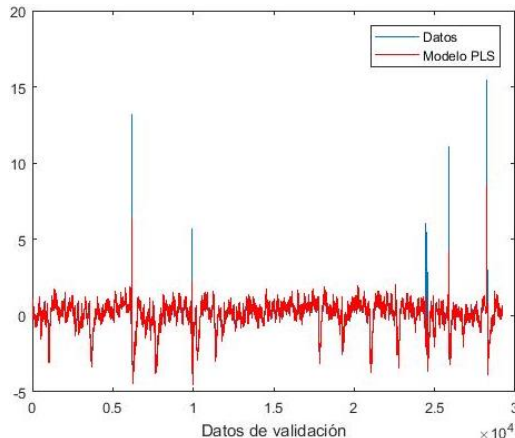


Figura 4.41a. Variable DQO de validación obtenido por PLS (91%) dinámico.

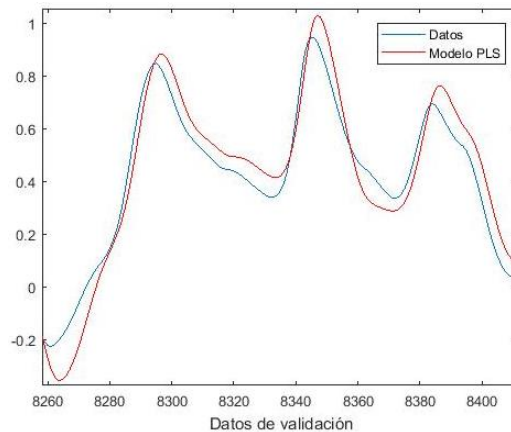


Figura 4.41b. Detalle de la variable DQO de validación obtenido por PLS (91%) dinámico.

Una vez realizadas las tres predicciones con el método PLS sobre el modelo dinámico se llega a la conclusión que cuantas menos variables son utilizadas, peor es la calidad de la predicción. También otro aspecto a destacar es que este método trabaja mejor con los datos de validación que con los datos de entrenamiento. Para corroborar esta conclusión se muestran a continuación el valor medio de los residuos que se ha hallado para cada predicción:

- Valor medio de residuo entrenamiento PLS (98,5%) = 0,1533
- Valor medio de residuo validación PLS (98,5%) = 0,1290

- Valor medio de residuo entrenamiento PLS (97,4%) = 0,2121
- Valor medio de residuo validación PLS (97,4%) = 0,1778

- Valor medio de residuo entrenamiento PLS (91%) = 0,3229
- Valor medio de residuo validación PLS (91%) = 0,2719

La obtención de estos datos permite concluir que la mejor predicción se corresponde con el mayor porcentaje utilizado, que implica usar más variables.

4.4.2.5 Método PCA

Este método consiste principalmente en reducir el tamaño de la muestra de datos. Para elegir el número de variables que se van a utilizar se elegirá el número de variables que resulta de aplicar el 98,5% al test del porcentaje de la varianza usado en el método PLS. Aplicando este porcentaje se obtiene el número de seis variables con las que trabajar.

En las figuras 4.42a y 4.42b se ha realizado la predicción de DQO con los datos de entrenamiento, y las figuras 4.43a y 4.43b se ha calculado en función de los datos de validación.

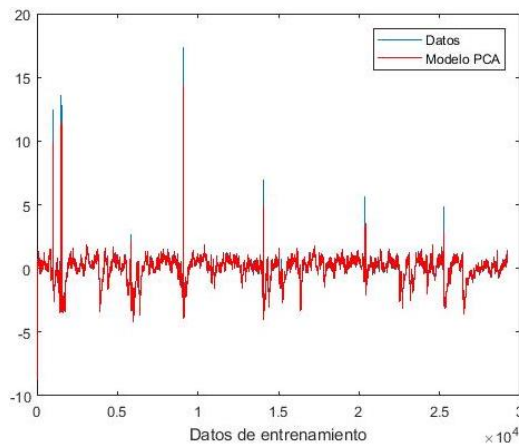


Figura 4.42a. Variable DQO de entrenamiento obtenido por PCA dinámica.

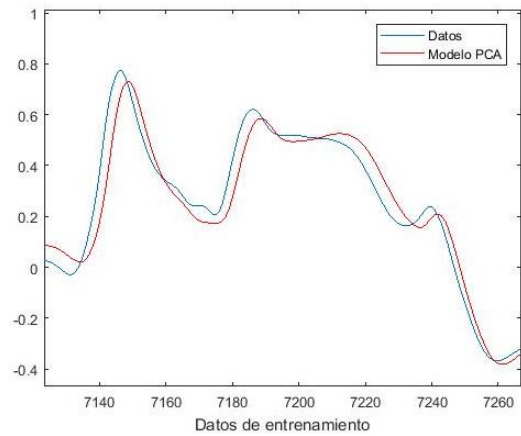


Figura 4.42b. Detalle de la variable DQO de entrenamiento obtenido por PCA dinámica.

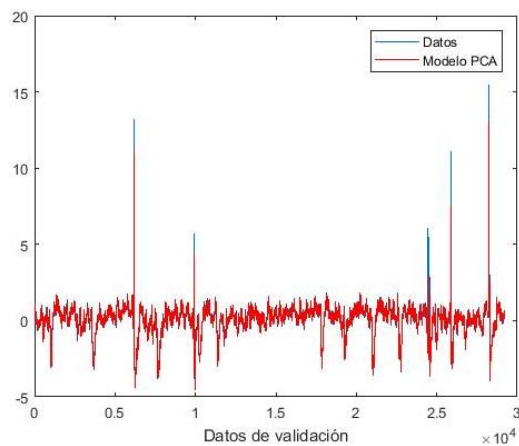


Figura 4.43a. Variable DQO de validación obtenido por PCA dinámico.

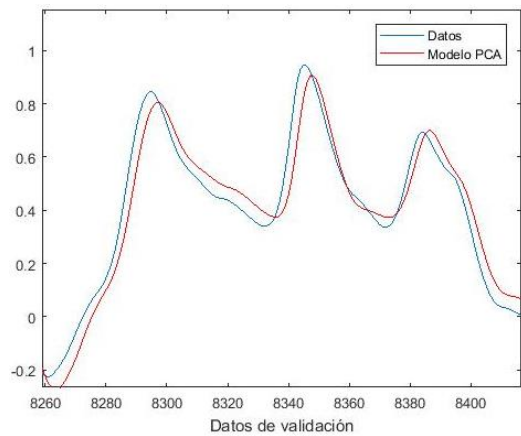


Figura 4.43b. Detalle de la variable DQO de validación obtenido por PCA dinámica.

Una vez representada la variable DQO predicha en función de los datos de entrenamiento y de validación lo siguiente es sacar los valores medios de los residuos para cada caso.

Valor medio de residuo entrenamiento = 0,2815

Valor medio de residuo validación = 0,2285

La predicción obtenida por el método PCA sobre el modelo dinámico no ofrece una gran respuesta no aproximando tanto la solución como otros métodos aplicados. Utilizando más variables seguramente se conseguiría una mejor respuesta pero el objetivo de este método es reducir la dimensión de la muestra de datos por lo que se ha visto conveniente usar solo 6 variables.

4.4.2.6 Combinación de todos los métodos

Para este método que surge de la combinación de los diferentes métodos calculados anteriormente, se van a realizar dos predicciones diferentes aplicando diferentes porcentajes a cada método. Para los casos de los métodos Lasso y PLS en los que se ha realizado más de una predicción se escogerá aquella que proporciona un resultado mejor. Para Lasso será el calculado con la columna 10, y para PLS el calculado con el porcentaje de 98,5%, que utiliza 6 variables.

En un primer lugar, se hará una predicción en la que se aplicará el mismo porcentaje a cada método, y en segundo lugar, se simulará otra situación en la que se aplicará un porcentaje de acuerdo con la calidad de la predicción obtenida anteriormente. En las figuras 4.44a y 4.44b se ha representado la variable DQO predicha con los datos de entrenamiento, mientras que en las figuras 4.45a y 4.45b está representada la predicción calculada con los datos de validación.

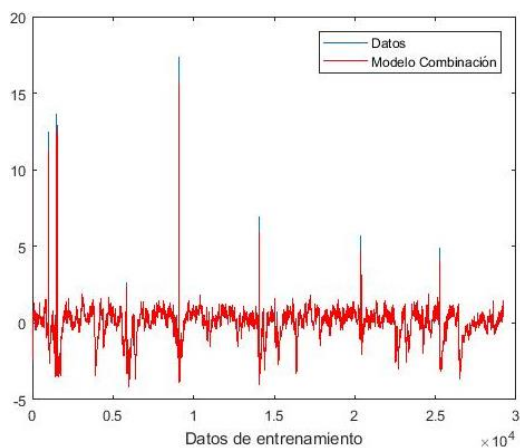


Figura 4.44a. Variable DQO de entrenamiento obtenido por combinación equitativa dinámica.

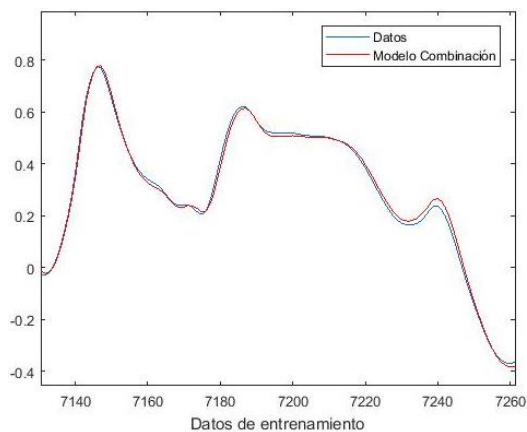


Figura 4.44b. Detalle de la variable DQO de entrenamiento obtenido por combinación equitativa dinámica.

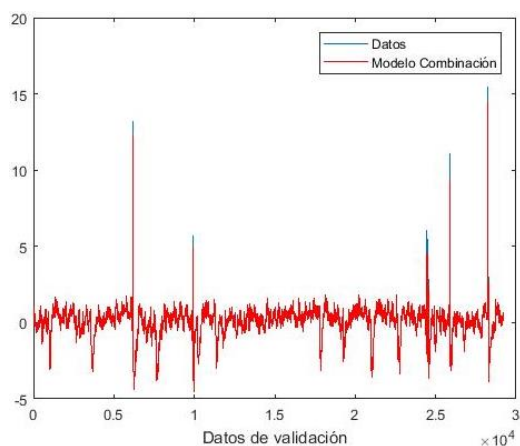


Figura 4.45a. Variable DQO de validación obtenido por combinación equitativa dinámico.

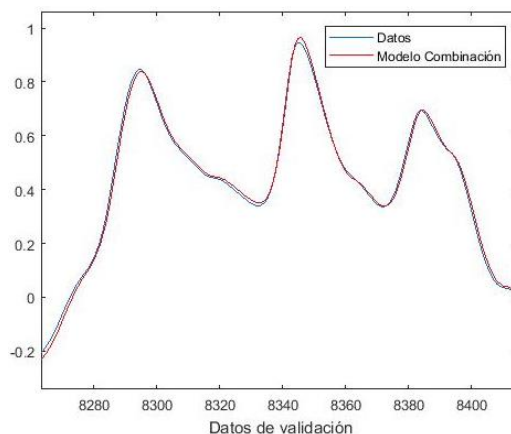


Figura 4.45b. Detalle de la variable DQO de validación obtenido por combinación equitativa

Observando las predicciones obtenidas se aprecia que la predicción es de gran calidad, resultando mejor la realizada con los datos de validación. Esto último indica que el método “aprende” bien el modelo, y es capaz de generalizar cuando se le presentan datos nuevos. A continuación se muestra los valores medios de los residuos de las dos predicciones anteriores:

Valor medio de residuo entrenamiento = 0,0855

Valor medio de residuo validación = 0,0677

Efectivamente estos valores certifican la gran calidad de la predicción obtenida. Si utilizando una designación equitativa del peso de cada método se ha obtenido un resultado tan bueno, asociando un peso mayor a los métodos que trabajan mejor se supone que se obtendrá un resultado mejor. Para decidir el porcentaje a aplicar a cada método se hará en función de los valores medios de residuos de cada uno, como se puede ver en la siguiente tabla.

Método	Valor medio de residuo	Porcentaje designado
Mínimos cuadrados	0,0241	30%
Ridge	0,0362	30%
Lasso-10	0,0685	20%
PLS (98,5%)	0,1290	15%
PCA	0,2285	5%

En las figuras 4.46a y 4.46b se ha representado la variable DQO calculada en función de los datos de entrenamiento aplicando el método combinado en función de los porcentajes anteriores, mientras que en las figuras 4.47a y 4.47b se ha representado la misma variable predicha con el mismo método con los datos de validación.

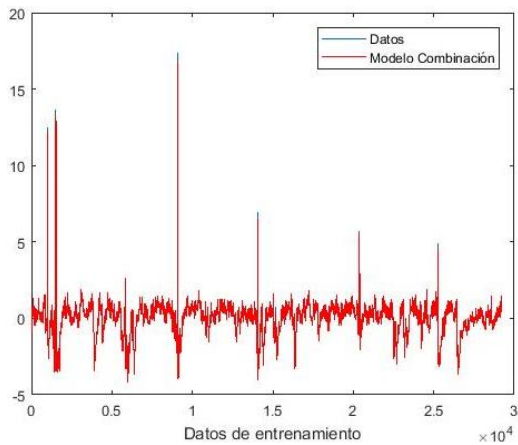


Figura 4.46a. Variable DQO de entrenamiento obtenido por combinación dinámica.

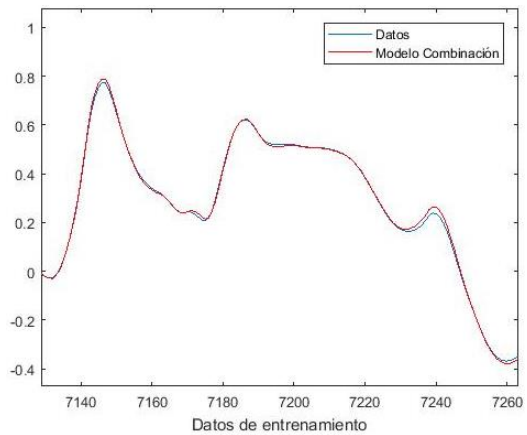


Figura 4.46b. Detalle de la variable DQO de entrenamiento obtenido por combinación dinámica.

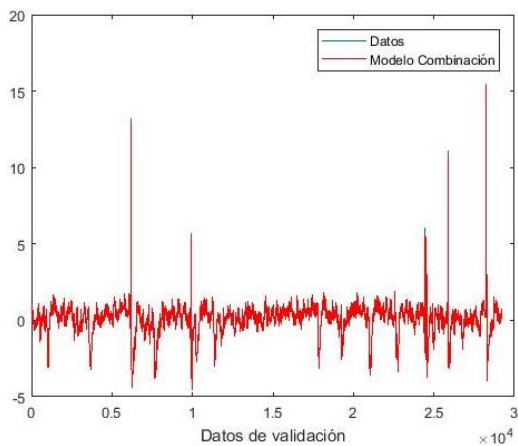


Figura 4.47a. Variable DQO de validación obtenido por combinación dinámica.

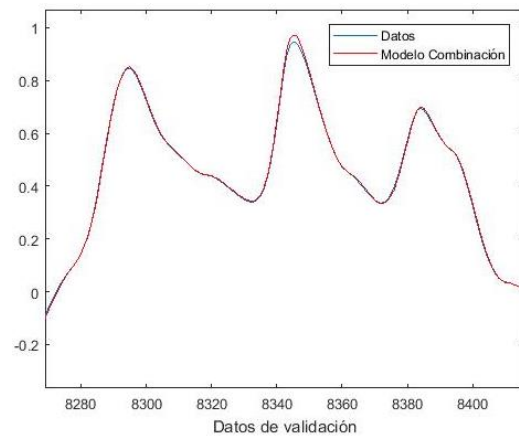


Figura 4.47b. Detalle de la variable DQO de validación obtenido por combinación dinámica.

Las predicciones obtenidas se aproximan bastante bien a los datos recogidos de la planta, por lo que se puede decir que se ha obtenido una gran calidad en la predicción. Los valores medios de los residuos obtenidos son los siguientes:

Valor medio de residuo entrenamiento = 0,0466

Valor medio de residuo validación = 0,0359

Efectivamente la predicción es de gran calidad, destacando que el método trabaja mejor con datos nuevos. Seguramente esto se debe a que la mayoría de los métodos utilizados obtienen mejores predicciones con datos nuevos. Cabe destacar también que la predicción obtenida no supera los resultados obtenidos con el mejor método utilizado para el modelo dinámico, el de mínimos cuadrados, pero si mejora el resto.

4.4.2.7 Comparativa entre métodos aplicados al modelo dinámico

Como se ha realizado con el modelo estático, una vez aplicado los diferentes métodos propuestos al modelo dinámico se considera oportuno realizar una comparativa entre ellos. Con el fin de visualizar las posibles diferencias de una forma rápida se presenta la siguiente tabla que dispone de los valores medios de los residuos obtenidos aplicando los métodos sobre los datos de entrenamiento y los datos de validación.

Método aplicado	Valor medio del residuo en entrenamiento	Valor medio del residuo en validación
Mínimos cuadrados	0.0233	0.0241
Ridge	0.0323	0.0362
Lasso-10	0.0707	0.0685
Lasso-50	0.0867	0.0754
Lasso-74	0.1668	0.1447
PLS (98,5%)	0.1533	0.1290
PLS (97,4%)	0.2121	0.1778
PLS (91%)	0.3229	0.2719
PCA	0.2815	0.2285
Combinación equitativa de métodos	0.0855	0.0677
Combinación porcentual de métodos	0.0466	0.0359

De todos los métodos aplicados sobre el modelo dinámico, el método de Mínimos cuadrados es el que ofrece una mejor predicción. Pero hay que tener en cuenta que al utilizar este método saltaba un Warning, un problema con la posible colinealidad de la muestra de datos. Por lo tanto, no sería recomendable usar este método porque aunque en unos casos ofrezca una predicción de gran calidad, en otros casos es posible que no pudiese ofrecer una solución.

Tras descartar el método de Mínimos cuadrados, hay dos métodos que ofrecen también una buena predicción. Estos métodos son Ridge con un valor medio del residuo de 0.0362 y la combinación porcentual de métodos con un 0.359 como valor medio. Finalmente de estos dos métodos sería más recomendable utilizar Ridge porque el método por combinación también incluye el método de Mínimos

cuadrados pudiendo aparecer el mismo problema de colinealidad anteriormente mencionado.

Tanto los métodos estudiados de Lasso como los de PLS tienen en común que al utilizar menos variables presentan una predicción de peor calidad. Se puede concluir que al poseer más información el método obtiene una predicción más ajustada a los datos reales.

En último lugar destacar que a excepción de los métodos de Mínimos cuadrados y Ridge, el resto de métodos aplicados ofrecen mejores predicciones con datos nuevos como se puede ver en los valores medios de los residuos que se han obtenido.

4.4.3 Conclusiones

Una vez que se han realizado las predicciones de la muestra de datos aplicando diferentes métodos sobre el modelo estático y el dinámico se pueden obtener diferentes conclusiones. Este apartado se encargará de realizar una comparativa entre todo lo obtenido en los apartados anteriormente.

En primer lugar, se realizará una comparación entre los métodos iguales que se han aplicado tanto al modelo estático como al modelo dinámico. Para ello se presenta la siguiente donde aparecen los diferentes valores medios de los residuos obtenidos aplicando cada modelo sobre los datos de validación.

Método aplicado	Modelo estático	Modelo dinámico
Mínimos cuadrados	0.2674	0.0241
Ridge	0.2897	0.0362
Lasso-10	0.2009	0.0685
Lasso-50	0.4445	0.0754
PLS (98,5%)	0.4234	0.1290
PCA	0.4469	0.2285
Combinación equitativa de métodos	0.1613	0.0677
Combinación porcentual de métodos	0.1563	0.0359

Como se puede observar en la tabla todas las predicciones que se obtienen al aplicar el mismo método sobre el modelo dinámico presentan mejores valores medios de los residuos que sobre el modelo estático. Esto era algo que se suponía

que debería pasar debido ya que el modelo dinámico contiene más información, concretamente información de instantes anteriores.

Además cabe destacar la mejora que proporciona el modelo dinámico, mientras que ningún valor medio del residuo del modelo estático baja de 0.1, la mayoría de los valores del modelo dinámico se encuentran por debajo de 0.1. Así que siempre que sea posible será mejor utilizar un modelo dinámico.

También es interesante hacer una comparativa entre los resultados obtenidos por el método PLS en los dos modelos, como se puede apreciar en la siguiente tabla:

Método aplicado	Valor medio del residuo en entrenamiento	Valor medio del residuo en validación
PLS (98,5%) estático	0,4704	0,4234
PLS (98,5%) dinámico	0.1533	0.1290
PLS (97,4%) dinámico	0.2121	0.1778
PLS (91%) dinámico	0.3229	0.2719

Observando los valores obtenidos aplicando el modelo dinámico indican que las predicciones se ajustan más a los datos recogidos de planta que utilizando el modelo estático. Analizando las predicciones obtenidas aplicando el mismo porcentaje, el 98,5%, proporcionan resultados muy diferentes. Esto es debido a que el porcentaje no se realiza sobre el mismo número de variables, y éstas no tienen la misma varianza. Aplicando este mismo porcentaje, el modelo dinámico trabaja con seis variables mientras que el dinámico trabaja únicamente con cinco.

Entonces se planteó aplicar un porcentaje menor al modelo dinámico con el objetivo de trabajar con cinco variables y evaluar si se obtenía el mismo resultado o parecido. Esto se consiguió aplicando el porcentaje de 97,4%, y como se puede ver en la tabla la predicción obtenida seguía mejorando con creces al modelo estático. Aunque se trabaje con el mismo número de variables, esto no quiere decir que sean las mismas variables produciéndose estas diferencias por este motivo.

Después de esto se decidió aplicar un porcentaje de la varianza menor buscando trabajar con menos variables aún. Por lo que se aplicó un 91%, lo que se traducía en trabajar con 3 variables, e incluso así se ha obtenido una predicción mejor que aplicando un mayor porcentaje en el modelo estático.

Finalmente, como conclusión final destacar que el modelo dinámico ofrece siempre una mejor predicción que el modelo estático debido a que posee más información de instantes. Y también se podría destacar que muchas veces importa más la calidad de la información de la que se dispone que la cantidad como se ha podido ver en los métodos de reducción del tamaño de la matriz de datos.

4.5 Detección y diagnóstico de fallos

4.5.1 Introducción

El objetivo principal de este apartado será diseñar un sistema que permita determinar la existencia de un fallo, así como el instante en el que se ha producido. Además como complemento también se determinará las variables que han provocado dicho fallo.

Para realizar diferentes pruebas sobre nuestro sistema se trabajará con una toma de datos de la depuradora de aguas residuales que contendrá un fallo. De este fallo se conoce el instante en el que se produce y las variables que lo han provocado.

Se realizará la detección y diagnóstico de fallos con dos estadísticos que permitirán realizar esta labor con la finalidad de obtener las mismas conclusiones tras aplicarlos y verificar su correcto funcionamiento. Estos dos estadísticos son conocidos como el T^2 y Q.

4.5.2 Pasos previos

Para simular el modelo con fallo en Matlab y realizar la detección y diagnóstico del fallo se realizan unos pasos previos comunes que se explican en este apartado.

La detección y diagnóstico de fallos realizados con los estadísticos T^2 y Q se apoya en el método PCA, por lo que se necesita construir una matriz X que contenga los datos de la planta sin fallo, que funcione con normalidad.

En primer lugar se reducirá el tamaño de dicha matriz de datos. En este caso solo se eliminarán las variables cuya varianza es prácticamente igual a cero, es decir, aquellas variables que se habían eliminado anteriormente con una varianza menor a 0,0001 también serán eliminadas ahora. Tampoco es recomendable eliminar muchas variables, debido a que existiría la posibilidad de eliminar la variable que provoca el fallo, y por lo tanto, no se podría diagnosticar la causa de éste.

Una vez reducido el tamaño de la toma de datos, finalmente se va a trabajar con una matriz de datos que contiene 122 variables por las 140 de las que se disponía en un momento inicial.

El siguiente paso, será normalizar la nueva matriz de datos reducida a media cero y varianza uno. Después se aplicará el método PCA comentado en el tema 2, almacenando para testear el método, el modelo PCA calculado, que consiste en la matriz P y la matriz A que contiene en su diagonal los a valores propios retenidos al realizar el PCA.

Finalmente en este paso previo será necesario calcular los umbrales de las estadísticas T^2 y Q, como se definieron en el capítulo 2 de esta memoria, con los datos normales de operación.

El valor T^2 se hallará de un método simple como se ha visto en el apartado 2.6.4 utilizando la siguiente fórmula:

$$T^2 \text{ umbral} = \text{media} + 3 \cdot \text{desviación estándar}$$

$$T^2 \text{ umbral} = 38,2$$

Para hallar el valor umbral de Q se utilizará la misma fórmula que con T^2 , obteniendo el siguiente valor:

$$Q \text{ umbral} = 21,52$$

Una vez realizados todos estos datos fuera de línea, se está en disposición de recoger nuevos datos de la planta y probar si el método PCA es capaz de detectar fallos en la planta.

4.5.3 Estudio del fallo

A continuación se va a realizar una simulación de una toma de datos que contienen un fallo. En primer lugar se simulará sobre un modelo estático en el que se realizará la predicción de la variable DQO del efluente con alguno de los métodos vistos anteriormente, después se llevará a cabo la detección y diagnóstico del fallo mediante los estadísticos T^2 y Q. Una vez ejecutado el modelo estático, se repetirán estas operaciones sobre un modelo dinámico.

4.5.3.1 Modelo estático

4.5.3.1.1 Predicción por PCA

En primer lugar de cara a estudiar los nuevos datos que contienen un fallo, se decide calcular la predicción de la variable DQO del efluente con el modelo PCA. Esta predicción se realiza con el objetivo de observar si aparece algún comportamiento anómalo en la predicción o por lo contrario, se obtiene una predicción defectuosa.

A continuación en la figura 4.48 se ha representado la predicción obtenida por el método PCA sobre el modelo estático con el predictor b calculado anteriormente. Además también están representados en dicha figura los datos recogidos de planta.

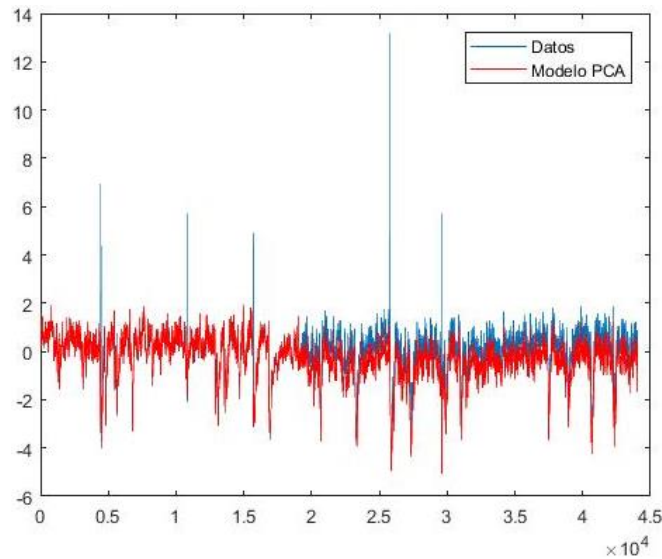


Figura 4.48. Variable DQO predicha por PCA con modelo estático de datos con Fallo

Observando la predicción lograda tras aplicar el método PCA se observa que no se ajusta bien la predicción a los datos recogidos de planta. De hecho, visualmente se aprecia como la predicción se ajusta bastante bien hasta un instante situado entre el 18500-19500 aproximadamente en el que empieza a diferir más la predicción con los datos recogidos de planta.

Se puede suponer que seguramente el fallo se encuentre en ese intervalo de instantes, pero de esta forma no se puede saber con exactitud el momento en el que ha surgido el fallo y las variables que lo han provocado. Por esta razón se presenta necesario la utilización de los estadísticos T^2 y Q anteriormente mencionados.

4.5.3.1.2 Detección y diagnóstico con el estadístico T^2

Para estudiar a fondo el fallo del nuevo conjunto de datos recogidos de la planta depuradora de aguas residuales se utilizara el estadístico T^2 .

Se calcula el valor T^2 para cada instante de la nueva muestra de datos que presumiblemente contiene un fallo. A la hora de decidir si existe o no un fallo se determina que éste existirá si T^2 supera el valor umbral en 20 instantes consecutivos. Esto se decide realizar así porque es posible que algún valor exceda el umbral en algún caso puntual sin que llegue a existir un fallo debido a un error a la hora de tomar la medida u otro motivo.

Tras evaluar la nueva muestra de datos según las condiciones que se han establecido se obtiene que el instante en el que se produce el fallo es el 19206. Este valor concuerda con lo que se podía haber apreciado anteriormente en la figura 4.48, donde se suponía que el fallo se encontraba entre los instantes 18500 y 19500. En la figura 4.49 está representado el valor T^2 calculado de cada instante junto con el valor T^2 umbral.

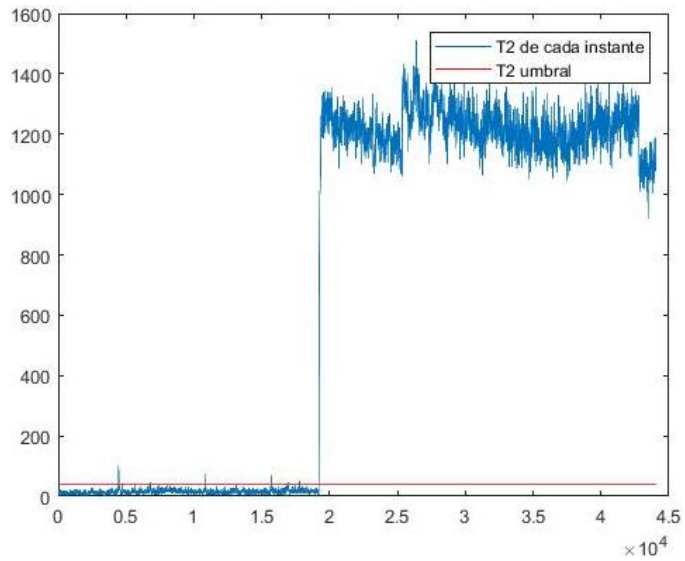


Figura 4.49. Representación de valor T^2 en cada instante.

Una vez determinado el instante en el que se produce el fallo que perturba el funcionamiento de nuestra planta depuradora de aguas residuales, se realizara el estudio de la contribución que aporta cada variable para que ocurra el fallo. La variable o variables que realicen una aportación por encima del resto serán la culpable o culpables de dicho fallo.

Aplicando el método de la contribución sobre el instante 19206, se obtienen una serie de valores que se ha decidido representar sobre un diagrama de barras en la figura 4.50.

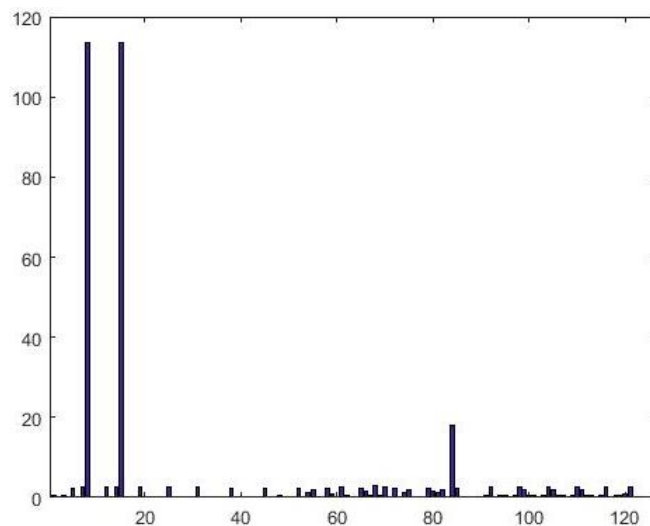


Figura 4.50. Contribución de cada variable al Fallo

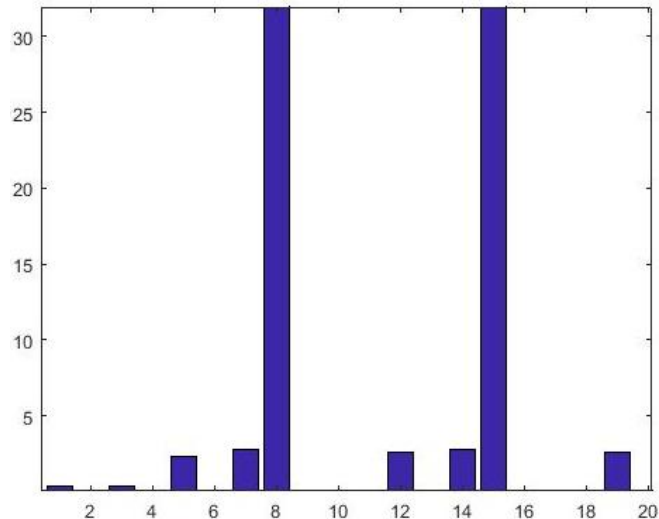


Figura 4.51. Zoom de la figura 4.50

Finalmente tras aplicar el método de la contribución de cada variable y su posterior representación en un diagrama de barras se observa como las variables 8 y 15 tienen un valor T^2 muy superior al resto de las variables. Por lo tanto estas variables son las culpables de que se haya producido dicho fallo. Concretamente estas variables son las medidas siguientes:

Variable 8 = Alcalinidad de la salida superior del decantador primario

Variable 15 = Alcalinidad de la salida inferior del decantador primario

4.5.3.1.3 Detección y diagnóstico con el estadístico Q

Después de analizar el fallo existente con el estadístico T^2 ahora es el turno de evaluarlo con el estadístico Q para comprobar que se obtienen las mismas conclusiones.

Se deben calcular los valores Q de todos los instantes de la muestra de datos que contiene el fallo. Se establecerá además el mismo criterio que con el estadístico T^2 para calcular el instante en el que se ha producido el error, es decir, deberá haber 20 instantes consecutivos con un valor Q superior al Q umbral.

Aplicando las condiciones establecidas se determina que el fallo se ha producido en el instante 19205, dato que concuerda con lo obtenido anteriormente. En la figura 4.52 se ha representado el valor Q calculado de cada instante comparado con el valor Q umbral, que se encuentra sobre la línea horizontal de cero debido a los altos valores de Q tras el fallo.

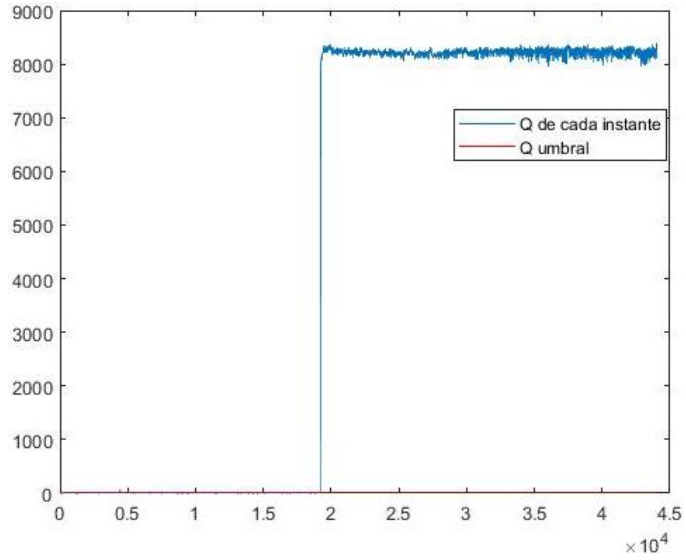


Figura 4.52. Representación de valor Q en cada instante.

El siguiente paso tras determinar el instante en el que se ha producido el fallo será localizar las variables que lo han provocado. Para ello se calculará el cuadrado del residuo de cada variable en el instante del fallo y se representará gráficamente en la figura 4.53 mediante un diagrama de barras. La variable o variables que tengan un residuo demasiado grande en comparación con el resto serán aquellas que han provocado el fallo.

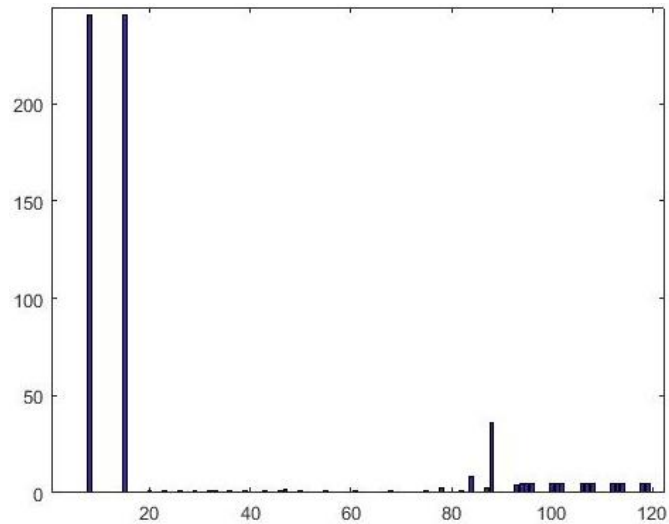


Figura 4.53. Residuos de cada variable

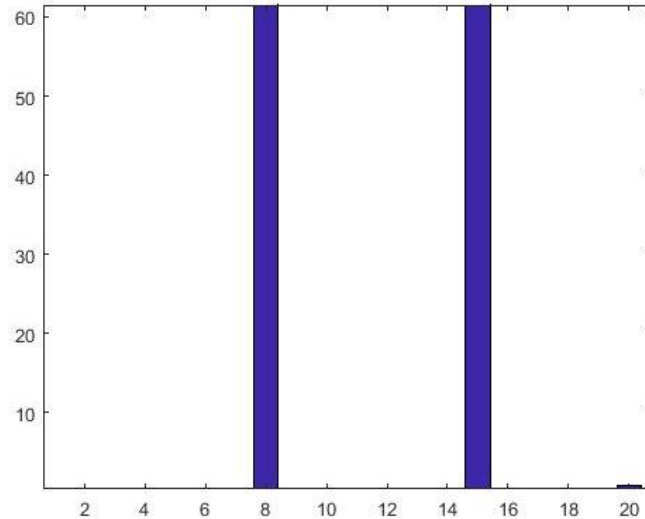


Figura 4.54. Zoom de la figura 4.53

Finalmente tras observar la representación gráfica de las figuras 4.53 y 4.54 se determina que las variables culpables del fallo que se está estudiando son las variables número 8 y 15 que corresponden con las variables de alcalinidad de las salidas superior e inferior del decantador primario.

4.5.3.2 Modelo dinámico

4.5.3.2.1 Predicción por PCA

Como se ha realizado con el modelo estático antes de analizar el posible fallo que se encuentra en la nueva toma de datos se decide calcular la predicción de la variable DQO del efluente con el método PCA. Esta predicción se realiza con la finalidad de ver cómo afecta el fallo a la hora de realizar la predicción.

En la figura 4.55 se ha representado la variable DQO predicha por el método PCA y la variable DQO recogida de los datos.

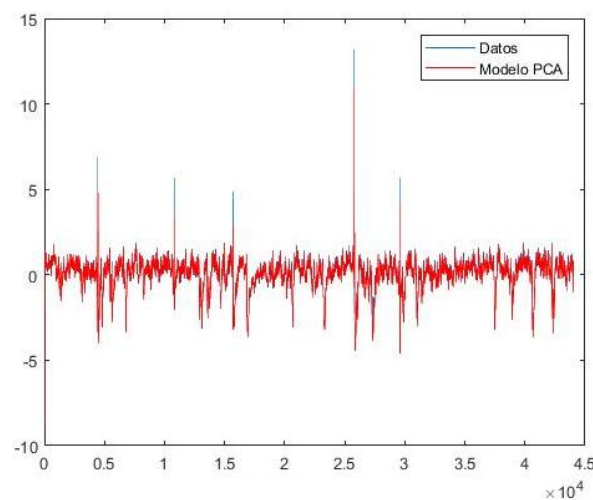


Figura 4.55. Variable DQO predicha por PCA con modelo dinámico de datos con Fallo

Como se puede ver la predicción es buena, y no presenta ninguna anomalía especial a lo largo de la predicción. Para asegurar esto, en la figura 4.56 se ha decidido hacer zoom sobre el instante donde ocurre el fallo para comprobar que a partir de dicho momento la predicción se sigue comportando igual.

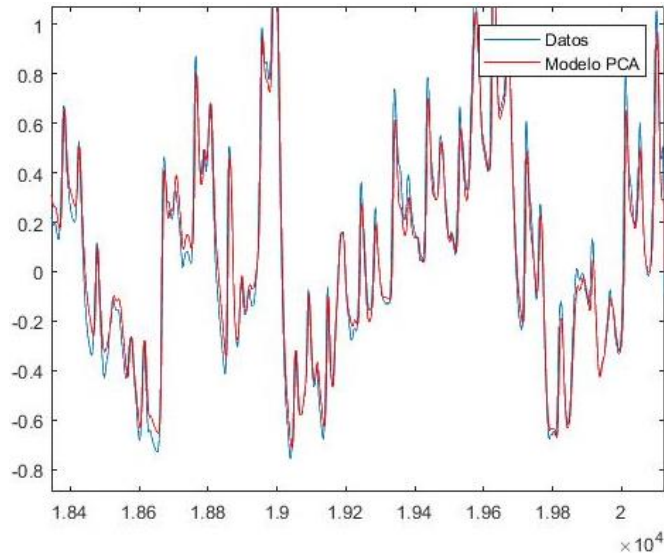


Figura 4.56. Zoom de la figura 4.55

Tras esta última representación gráfica se observa que la predicción se ajusta igual antes y después del instante en el que se produce el fallo. Esto sucede debido a que al realizar la reducción de variables con el método PCA, este método no utilice las variables que provocan el fallo. De hecho el valor medio del residuo obtenido de esta predicción es 0,1737. Debido a lo bien que se ajusta la predicción con el modelo dinámico se presenta necesario aplicar los estadísticos T^2 y Q para poder detectar la existencia del fallo.

4.5.3.2.2 Detección y diagnóstico con el estadístico T^2

Debido a que al realizar la predicción por el método PCA no se ha obtenido ninguna conclusión relacionada con el fallo que contiene la muestra de datos, se aplicará en primer lugar el estadístico T^2 .

Como primer paso se marcará el valor del umbral T^2 , el cual será el mismo valor que utilizado con el modelo estático. Este valor de T^2 es igual a 38,2.

Tras aplicar el método se ha obtenido que el fallo se produce en el instante 19206. Además se ha representado el valor T^2 obtenido en cada instante de la nueva toma de datos junto con el valor umbral T^2 en la figura 4.57. En esta figura se aprecia también el instante en el que se produce el fallo.

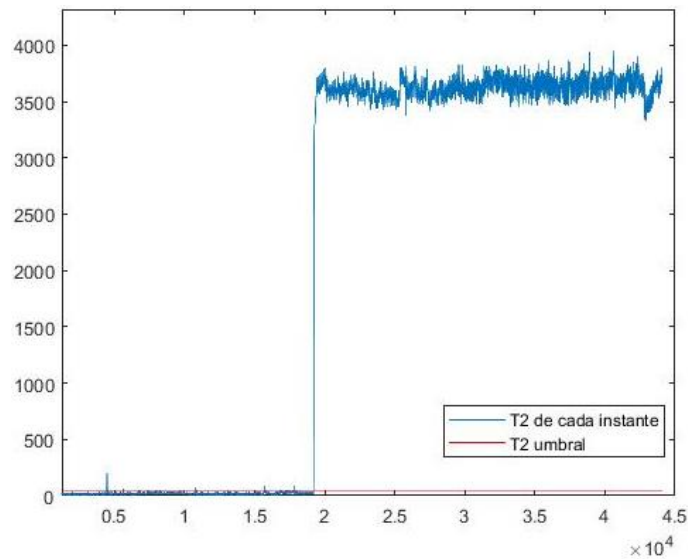


Figura 4.57. Representación de valor T^2 en cada instante.

Existe un pequeño conjunto de valores en torno al valor 4000 que superan también el valor marcado como umbral pero no representan un fallo. En cambio, todos los valores tras el instante 19206 en el que transcurre el fallo superan el valor del umbral.

Una vez que se conoce el instante en el que transcurre el fallo, el último paso que queda es identificar que variable o variables son las culpables de la existencia de dicho fallo. Para ello se aplicará un método de estudio que nos proporciona lo que aporta cada variable al fallo. En la figura 4.58 se ha representado mediante un diagrama de barras la contribución que realiza cada variable del método dinámico.

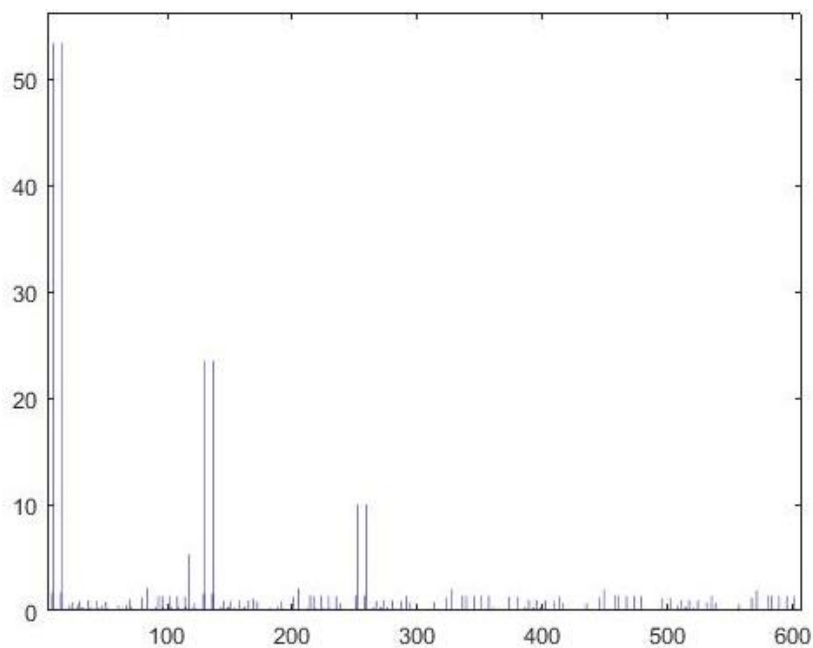


Figura 4.58. Contribución de cada variable al Fallo

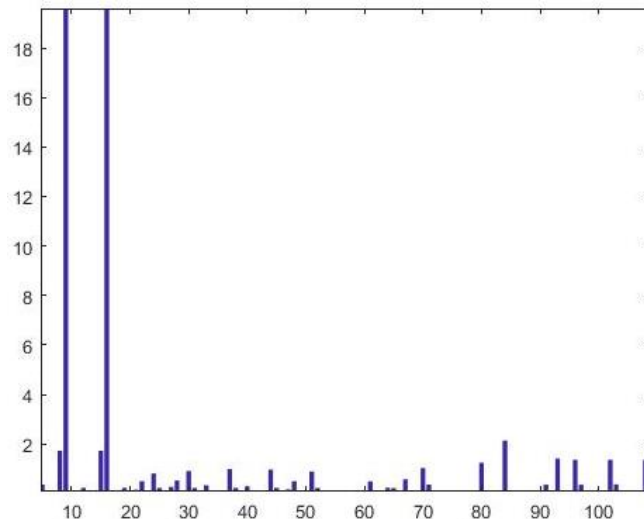


Figura 4.59. Zoom de la figura 4.58

Finalmente tras realizar este paso se pueden determinar las variables culpables del fallo. Debido al gran número de variables con las que trabaja el modelo dinámico, el grosor de barras del diagrama de la figura 4.58 es muy fino. Aun así se puede ver como hay dos variables que contribuyen al fallo en una proporción notable comparado con el resto de variables. Estas dos variables son las número 8 y 15, que se corresponden con las variables de alcalinidad de las salidas del decantador primario superior e inferior, respectivamente

4.5.3.2.3 Detección y diagnóstico con el estadístico Q

Tras estudiar el fallo con el estadístico T^2 , ahora se realizará lo mismo con el estadístico Q con la finalidad de comprobar los resultados obtenidos anteriormente.

El umbral que se usará será el mismo valor Q umbral utilizado con el modelo estático, es decir, Q umbral será igual a 21,52. Una vez definido este valor, se decide ejecutar el programa planteado en Matlab obteniendo que el fallo se produce en el instante 19205. Se ha representado en la figura 4.60 el valor Q de cada instante donde se puede apreciar como a partir del 19205 hay un importante cambio en el valor de este estadístico.

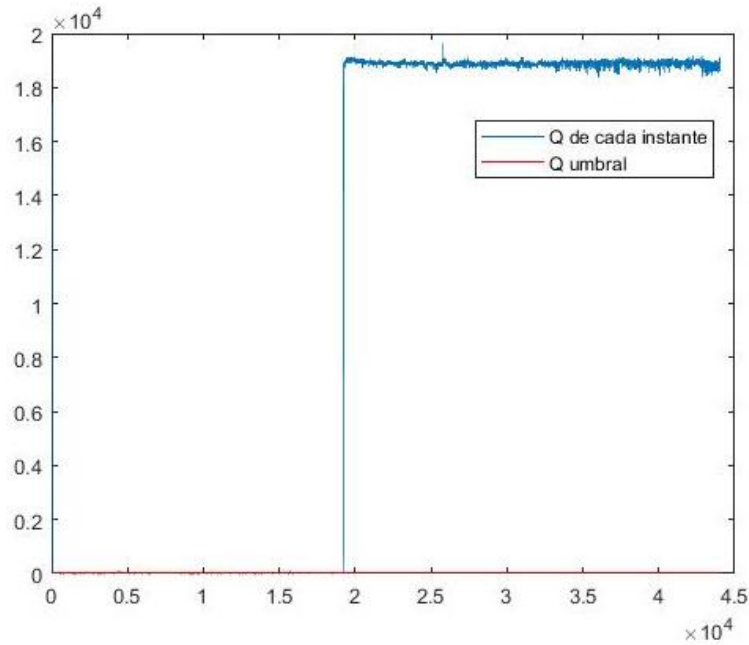


Figura 4.60. Representación de valor Q en cada instante.

El valor Q umbral se encuentra sobre la línea de cero, debido a que su valor es demasiado pequeño en comparación con los valores Q obtenidos tras el fallo. Incluso así se aprecia bien la diferencia que experimenta el valor Q de cada instante tras el instante 19205 en el que se produce el fallo.

Finalmente, solo queda identificar las variables culpables de dicho fallo. Para ello hallaremos el residuo de cada variable y se representará en la figura 4.61 mediante un diagrama de barras.

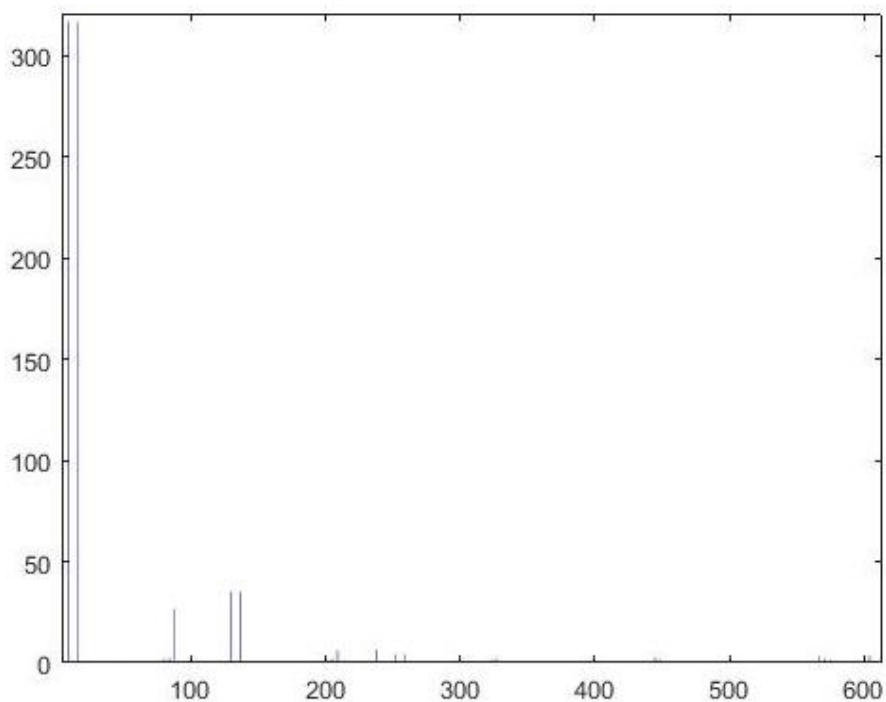


Figura 4.61. Residuos de cada variable

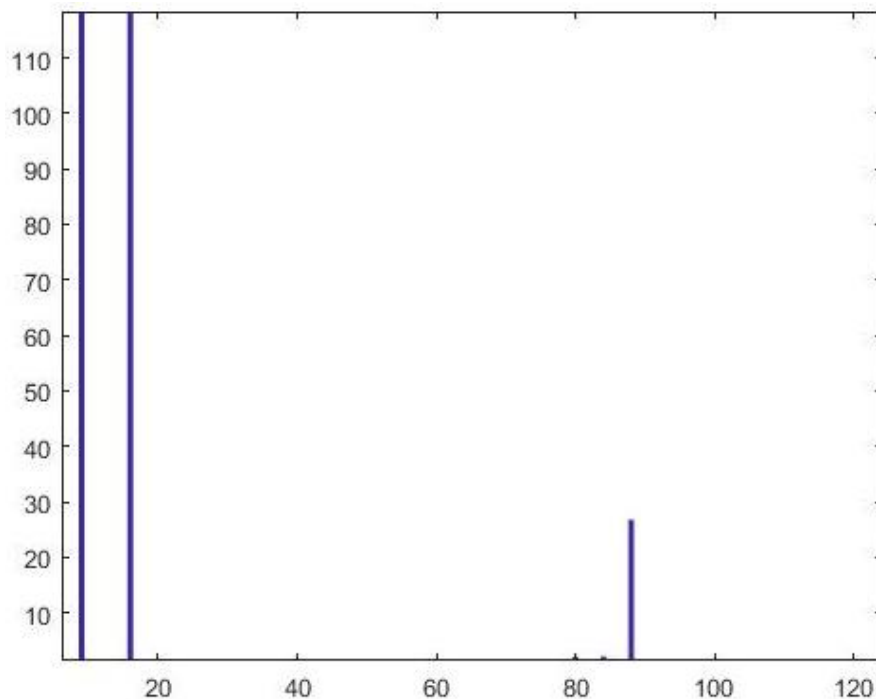


Figura 4.62. Zoom de la figura 4.61

Al igual que había pasado con el estadístico T^2 debido al gran número de variables con las que trabaja el modelo dinámico apenas el diagrama de barras sale de trazo muy fino. Aun así se percibe que hay dos variables muy por encima del resto, y como se puede apreciar en la figura 4.62 en la que se ha realizado el zoom se corresponden con las variables 8 y 15. La variable 8 se corresponde precisamente con la variable alcalinidad de la salida superior del decantador primario, y la variable con la variable alcalinidad de la salida inferior del mismo decantador.

4.5.4 Conclusiones

Tras realizar la detección y diagnóstico del fallo que se encontraba en la muestra de datos con dos estadísticos diferentes, se ha llegado a una serie de conclusiones.

En primer lugar, antes de detectar y diagnosticar el fallo se ha calculado la predicción de los datos nuevos con los predictores b hallados de la muestra de datos que no contenían fallo. Se ha calculado la predicción tanto con el modelo estático como con el modelo dinámico con el objetivo de ver si el fallo afectaba al resultado final de la predicción obtenida.

La predicción obtenida del modelo estático con PCA refleja un cambio en la aproximación de ésta a partir del instante en el que transcurre el fallo, mientras que la predicción obtenida con PCA con el modelo dinámico no refleja ningún comportamiento anómalo. Por lo tanto, se puede concluir que existe la posibilidad de que los métodos que implican una reducción de variables no utilicen las variables que han producido el fallo y la predicción no se vea afectada.

El siguiente aspecto a analizar es la detección del instante en el que se ha producido el fallo. Por una parte, el estadístico T^2 ha detectado el instante 19206 como origen del fallo tanto en el modelo estático como el dinámico. Por otra parte, el estadístico Q ha detectado el instante 19205 como origen del fallo en los dos modelos.

Tras la detección del instante en el que se ha producido el fallo se puede determinar que independientemente del modelo, dinámico o estático, se obtiene el mismo instante con cada estadístico. El instante obtenido con un estadístico no es igual al obtenido con el otro, pero al tratarse de instantes consecutivos se puede determinar que el resultado proporcionado por los dos estadísticos es de buena calidad.

El último aspecto a analizar es la identificación de las variables culpables del fallo. Tanto con un estadístico como con el otro se han obtenido las mismas variables, concretamente las variables 8 y 15 que se corresponden con las variables referentes a la alcalinidad de la salida superior y la salida inferior del decantador primario.

Adicionalmente se ha realizado la detección y diagnóstico de dos fallos más, un primer fallo relacionado con una reducción en el oxígeno del reactor 4, y el otro fallo que varía los caudales de la planta depuradora simulando fugas en las tuberías. La detección de dichos fallos ha sido satisfactoria en ambos casos, obteniendo los mismos resultados con el estadístico T^2 y con Q. A modo resumen se presenta la siguiente tabla donde se indica el instante en el que se ha producido cada fallo y las variables responsables de ello.

Fallo	Instante inicial del fallo	VARIABLES responsables
Reducción de oxígeno en el reactor 4	19210	Variable 44 = O_2 del reactor 4
Variaciones Caudal	19204	- Variable 85 = DQO del espesador de lodos - Variable 92 = DQO del digestor anaeróbico





Conclusiones y futuras líneas de trabajo



5. CONCLUSIONES Y FUTURAS LÍNEAS DE TRABAJO

5.1 Conclusiones

Este trabajo se había marcado como objetivo principal diseñar un sensor software que fuera capaz de obtener una variable en un proceso de calidad en una planta depuradora de aguas residuales. Para ello, se han utilizado diferentes técnicas pertenecientes al análisis estadístico de regresión multivariable sobre dos tipos de modelos. Los modelos utilizados han sido el modelo estático y el modelo dinámico, que incluye datos de instantes pasados de la planta depuradora.

Entre las técnicas de regresión lineal multivariable utilizadas se incluyen métodos como pueden ser Mínimos cuadrados, Ridge, Lasso, PLS o PCA. Además se ha propuesto una combinación de dichos métodos con el objetivo de estudiar otra posibilidad y obtener una más que posible mejora.

Tras obtener las diferentes predicciones se observó como todos los métodos no trabajan igual presentando diferencias entre ellos. Quizá la diferencia más notable es la hallada entre los métodos aplicados sobre el modelo estático y el modelo dinámico. Esta diferencia consiste en que el mismo método aplicado sobre un modelo dinámico siempre proporciona una predicción de mayor calidad que sobre un modelo estático. Por lo tanto, se presenta muy importante a la hora de realizar una predicción tener información de instantes pasados ya que mejora notablemente la calidad de ésta.

Otra conclusión a la que se ha llegado es que no en todos los métodos cuantas más variables sean utilizadas en la predicción, mejor será la calidad de ésta. Ya que ha habido métodos en los que utilizando menos variables se han obtenido mejores predicciones que en otros que usaban más. Por lo que se puede concluir que en muchos casos importa más la calidad de las variables que se utilizan que la cantidad.

También a destacar que a excepción de los métodos de Mínimos cuadrados y Ridge el resto de métodos realizaba mejores predicciones con datos nuevos. Esto se considera importante debido a que en la planta depuradora de aguas residuales las predicciones siempre se realizarán con datos nuevos.

El otro objetivo de este trabajo era que el sensor fuese capaz de realizar la detección y el diagnóstico de un fallo existente en una muestra de datos. Para ello se introdujo una muestra de datos que incluía un fallo conocido. Con el fin de detectar el instante en el que se había originado el fallo y las variables que lo habían provocado se han utilizado dos estadísticos, T^2 y Q , que han permitido llevar a cabo esta labor.

Al utilizar estos dos estadísticos se han obtenido las mismas conclusiones, que coincidían con el error que contenía la nueva muestra de datos. Es decir, que los dos estadísticos proporcionan una detección y diagnóstico fiable de fallos, pudiéndose utilizar cualquiera de las dos técnicas.

5.2 Futuras líneas de trabajo

El análisis estadístico multivariable tiene una gran progresión de futuro, ya que estos métodos son capaces de manejar una gran cantidad de datos. Esto es algo muy importante en la industria del futuro y ya en la presente también, donde cada vez se recogen una mayor cantidad de datos, pudiendo esto llegar a ser un problema.

El siguiente paso sería aplicar estos métodos en otras plantas, como pueden ser una planta azucarera con el objetivo de ver cómo reaccionan éstos ante un nuevo escenario. Además también resultaría interesante aplicar estos métodos a una planta real, ya que toda la prueba que se ha llevado a cabo ha sido de forma teórica y no siempre en la práctica todo se cumple de la misma forma que en la teoría.

En el caso de que los resultados obtenidos en la planta real fuesen satisfactorios, el sistema debería ser implementado en todos los puntos posibles del proceso.

Otra futura línea de trabajo sería aplicar diferentes técnicas estadísticas a las utilizadas, con el objetivo de comparar resultados e implantar la técnica que nos proporcione la mejor opción. Todas las técnicas utilizadas en este trabajo han sido técnicas lineales y, por lo tanto, se podría utilizar técnicas no lineales, como son las redes neuronales.





Bibliografía



6. BIBLIOGRAFÍA

En esta sección de la memoria se muestra la bibliografía consultada. Se ordena por orden alfabético del primer apellido del primer autor:

Libros y documentos:

1. Carro Paz, Roberto; González Gómez, Daniel. *Administración de las Operaciones – Diseño y selección de procesos*. Universidad Nacional de Mar del Plata
2. Departamento de Ingeniería Química (2014). *Apuntes de la asignatura: Tecnología Ambiental y de Procesos, Tecnologías de tratamiento de aguas residuales*. Ingeniería en Tecnología Industriales.
3. De la Fuente, María Jesús (2010). *Apuntes de la asignatura: Métodos de detección de fallos para el control y supervisión de procesos*. Máster en Investigación en Ingeniería de Procesos y Sistemas Industriales.
4. Katsuhiko, Ogata (1998). *Ingeniería de control moderna*. Ed. Pearson, University of Minnesota
5. Montgomery, Douglas (2004). *Control Estadístico de la calidad*. Limusa Wiley, México D.F.
6. Puigjaner, Luis; Ollero, Pedro; De Prada, César; Jiménez, Laureano (2006). *Estrategias de modelado, simulación y optimización de procesos químicos*. Editorial Síntesis, Madrid
7. Ryan, Thomas P. (2009). *Modern Regression Methods*. John Wiley & Sons, New Jersey
8. Sebastián Pérez, Miguel Ángel; Bargeño Fariñas, Vicente; Novo Sanjurjo, Vicente (1998). *Gestión y control de calidad*. Universidad Nacional de Educación a Distancia, Madrid
9. Uriel, Ezequiel; Aldás, Joaquín (2005). *Análisis Multivariante Aplicado*. Paraninfo, Madrid
10. Vilar Barrio, José Francisco (2005). *Control Estadístico de los procesos*. Fundación Confemetal Editorial, Madrid.

Artículos y publicaciones científicas:

11. Kadlec, Petr; Gabrys, Bogdan; Strandt, Sibylle (2009). *Data-Driven Soft Sensors in the process industry*. Computers and Chemical Engineering, Elsevier, Alemania.
12. Sánchez, A.; De la Fuente, María Jesús; Sainz-Palmero, Gregorio (2015) *Fault detection in Wastewater Treatment Plants using Distributed PCA Methods*. IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA), Luxemburgo.

Páginas web: comprobadas a 11/07/2018

13. Lodos de depuración de aguas residuales. Gobierno de España. Ministerio de Agricultura, Pesca y Alimentación- Ministerio para la Transición Ecológica.
Disponible en:
<http://www.mapama.gob.es/es/calidad-y-evaluacion-ambiental/temas/prevencion-y-gestion-residuos/flujos/lodos-depuradora/>
14. Norma ISO 9001: Sistema de Gestión de Calidad. Historia y evolución del concepto de Gestión de Calidad. ISOTools.
Disponible en:
<https://www.isotools.org/2016/01/30/historia-y-evolucion-del-concepto-de-gestion-de-calidad/>
15. Sistema de la Gestión de Calidad ISO 9001. AENOR
Disponible en:
http://www.aenor.es/aenor/certificacion/calidad/calidad_9001.asp#.Wt5owvkgWM8
16. Task Group on Benchmarking of Control Strategies for WWTPs. IWA (International Water Association)
Disponible en:
<http://apps.ensic.inpl-nancy.fr/benchmarkWWTP/>