

Identificación de sustancias explosivas mediante la parametrización/caracterización de curvas planas



Alejandro Gracia Benito
Trabajo fin del grado de Matemáticas
Universidad de Zaragoza

Agradecimientos

Agradecer sinceramente a mis directores Javier Martínez Torres y Jorge Martín Morales la confianza depositada en mí para la elaboración de este trabajo y por el apoyo mostrado durante este periodo. Agradecer también a Luis Rández García su colaboración.

«Solo lo difícil es estimulante.»
José Lezama Lima.

Sacrificio.

Abstract

In Security and Defense, the identification and quantification of live explosives is an unresolved technological challenges. Nowadays, the most useful and efficient method is the smell of dogs, but they have clear limitations. This is why there are efforts in developing innovative, fast and effective platforms multi-sensors to ensure early detection of explosives. These systems are called “nanoSNIFFER” and they are made by a lot of sensors.

The main purpose of this work is to develop a method to encode a family of curves in \mathbb{R}^2 depending on a set of parameters (maximum 10). These curves correspond to different experimental tests with different amounts of explosive obtained through electrical signals of “nanoSNIFFER”. Each of them has between 300 and 320 points. The curve with 25 ppm (parts per million) will be the curve example.

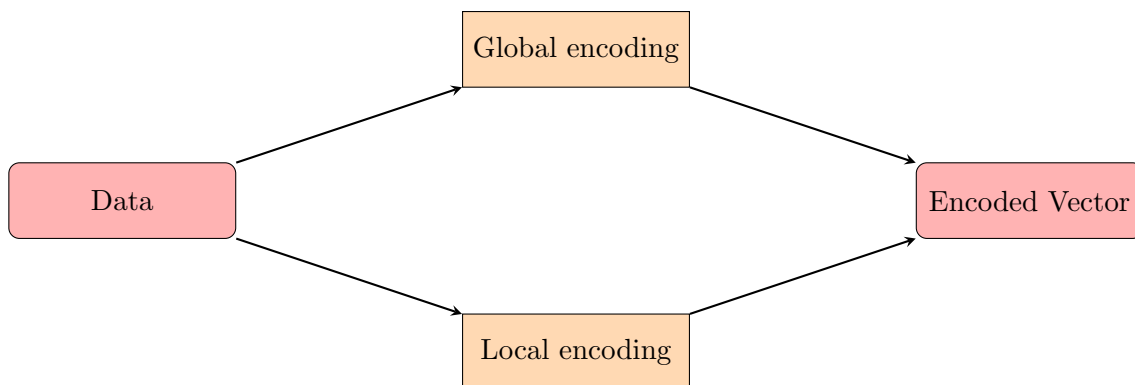


Figure 1: Encoding process.

As shown in Figure 1, for each curve, it shows a comparison between the global encoding and the local one, taking into account the physical meaning of each piece.

Global encoding

For global encoding, the following generic nonlinear functions are used to adjust data.

Parameters	Method	Function
2	exp1	$a_0 \exp(a_1 x)$
4	exp2	$\mathbf{exp1} + a_2 \exp(a_3 x)$
1+3	fou1	$a_0 + a_1 \cos(wx) + b_1 \sin(wx)$
1+5	fou2	$\mathbf{fou1} + a_2 \cos(2wx) + b_2 \sin(2wx)$
1+7	fou3	$\mathbf{fou2} + a_3 \cos(3wx) + b_3 \sin(3wx)$
3	gauss1	$a_0 \exp(-(x - a_1)^2 / b_1^2)$
6	gauss2	$\mathbf{gauss1} + a_2 \exp(-(x - a_3)^2 / b_2^2)$
9	gauss3	$\mathbf{gauss2} + a_4 \exp(-(x - a_5)^2 / b_3^2)$

For all curves, the best method is **fou3** obtaining a value of $R^2 \geq 0.9964$ for the smoothed data and for the original ones of $R^2 \geq 0.9544$. This method uses 8 parameters which is below the set threshold.

Thus, the encoding of the curve example with **fou3** is:

$$(83.08, 0.00269, 0.003321, -0.0001061, 0.001823, 0.0003441, 0.0007918, 0.503)$$

with a goodness of fit $R^2 = 0.9982$.

Local encoding

Each curve can be divided in three pieces taking into account the physical meaning: the first piece corresponds to the heating of the sensor, the second one corresponds to the introduction of the explosive and the actual experiment. The last one correspond to the removal of the explosive and sensor recovery. See Figure 3.5.

If each piece is encoded separately with previous methods, 24 parameters are required for coding with a value of $R^2 \geq 0.99$ (smoothed data). Therefore, this encoding does not improve the global one.

However, the idea of dividing the curves is quite interesting because the central piece is actually the important one, so only this piece will be encode. This one is encoded with a custom function called “*particularizada*” which is:

$$f(x, \theta) = a \exp(-bx) + c \sin(dx), \quad b \geq 0.$$

As noted, this function has half the parameters as **fou3**, i.e. only 4 parameters: (a, b, c, d) . In addition, it provides a goodness of fit of all curves higher than that obtained with other methods of four parameters like **fou1** and **exp2**.

So, the encoding of the curve example is:

$$(83.09, 1.871 \cdot 10^{-5}, 0.000978, 0.8789)$$

and $R^2 = 0.9959$.

Therefore, it has found a function that encodes each curve with only 4 parameters with a value of $R^2 \geq 0.9864$ (smoothed data) which is clearly advantageous to global encoding because it uses half the parameters.

Contents

Abstract	vii
1 Introducción	1
1.1 Motivación	1
1.2 Objetivo	2
1.3 Metodología	2
2 Suavizado y Regresión	5
2.1 Suavizado	5
2.1.1 Filtro <code>sgolay</code> (Savitzky-Golay)	5
2.1.2 Filtros <code>lowess</code> y <code>loess</code> (locally weighted scatter plot smooth)	6
2.2 Regresión	7
2.2.1 Regresión lineal	10
2.2.2 Regresión no lineal	11
3 Codificación de las curvas	15
3.1 Caso de estudio	15
3.2 Suavizado de la curva	16
3.3 Codificación global	16
3.4 Codificación global por trozos	19
3.5 Codificación del trozo central	20
3.5.1 Codificación particularizada	20
3.6 Codificación sin suavizado	22
4 Conclusiones	25
Bibliografía	27
Anexo	29
A.1 Suavizado	29
A.2 Comparación codificación global para todas las curvas	29
A.3 Comparación codificación global, global por trozos y local para cada curva	30
A.3.1 25 ppm	31
A.3.2 30 ppm	31
A.3.3 35 ppm	32
A.3.4 40 ppm	33
A.3.5 45 ppm	33
A.3.6 50 ppm	34
A.3.7 55 ppm	35
A.3.8 60 ppm	35
A.3.9 65 ppm	36

A.3.10	70 ppm	37
A.3.11	75 ppm	37
A.3.12	80 ppm	38
A.3.13	85 ppm	39
A.3.14	90 ppm	39
A.3.15	95 ppm	40
A.3.16	100 ppm	41
A.4	Codificación particularizada de cada una de las curvas	41

Capítulo 1

Introducción

En el ámbito de la Seguridad y Defensa, la identificación y cuantificación de explosivos «*in-situ*» y en tiempo real es uno de los retos tecnológicos pendientes de resolver. Los desafíos en este ámbito están relacionados con la amplia variedad de sustancias explosivas existentes, incluyendo los artefactos explosivos improvisados (IED). Si bien los fabricantes de explosivos incorporan en la formulación de los mismos sustancias relativamente volátiles o «*taggants*» que facilitan su identificación para controlar su tráfico ilegal, los cada vez más frecuentes IED caen fuera de esta categoría. En teoría, cualquier esquema de análisis químico podría ser utilizado, y de hecho, casi todos los métodos instrumentales conocidos ya han sido investigados para la detección de explosivos. Sin embargo, muchos de los análisis clásicos no son aplicables debido al tiempo requerido, y solamente son usados como ensayos de confirmación a posteriori, a su coste o a la formación especializada del usuario final. A día de hoy, podemos concluir que no existe una solución que resuelva los requerimientos de sensibilidad, selectividad, reproducibilidad, fiabilidad, sencillez, rapidez y coste para detección de IED en todas sus variantes y escenarios posibles.

En el ámbito de este proyecto y en relación a la aplicación de técnicas matemáticas para la automatización de procesos para la detección de compuestos químicos, se encuentran dos áreas de investigación bien diferenciadas:

- Estimación de la localización de la fuente de un determinado compuesto.
- Identificación de un determinado compuesto a partir de señales eléctricas de uno o más sensores.

En este trabajo, se da la información necesaria para la codificación de las sustancias para la posterior identificación mediante señales eléctricas de uno o más sensores.

Esta introducción así como el desarrollo del presente trabajo está enmarcado dentro del proyecto *Microsistemas basados en nanoestructuras con propiedades específicas de absorción y plasmón superficial, para detección de explosivos ocultos y agentes de guerra química* (CTQ2013-49068-C2-2-R) cuyo investigador principal es Miguel Ángel Urbitzondo Castro del CUD.

1.1 Motivación

En la actualidad, el método más útil y eficiente para la identificación y cuantificación de explosivos es el olfato de los perros adiestrados, pero tienen limitaciones obvias, de ahí los esfuerzos en el desarrollo de plataformas multisensoras innovadoras, rápidas y eficaces, con un principio de funcionamiento similar al olfato de los perros para garantizar la detección temprana de explosivos. Estos sistemas se denominan «*narices electrónicas*» y están constituidos

por multitud de sensores donde se combinan principios de transducción muy variados con el mismo fin.

1.2 Objetivo

El objetivo principal de este trabajo es obtener un método para codificar una familia de curvas en \mathbb{R}^2 dependiendo de un conjunto de parámetros, de tal modo que en caso de presentarse una nueva curva, el modelo sea capaz de codificarla según las curvas anteriormente dadas.

Para cada una de las sustancias explosivas, se dispone de varias curvas correspondientes a las diferentes pruebas experimentales con distintas cantidades de explosivo obtenidas gracias a los múltiples sensores que componen las llamadas «*narices electrónicas*».

La idea es mostrar para cada una de estas curvas planas, una comparación entre la parametrización global y la parametrización por trozos, teniendo en cuenta el sentido físico de cada uno de ellos.

Para ello, realizaremos lo que se llama un «*bombardeo fit*» con diferentes funciones de ajuste intentando encontrar un equilibrio entre el número de parámetros utilizados y el coeficiente de correlación entre la curva plana dada y la curva de ajuste, es decir, se intentará encontrar un coeficiente de correlación lo más próximo a la unidad pero, minimizando el número de parámetros que definen la curva.

1.3 Metodología

Para la codificación de tales curvas se han utilizado dos vías, regresión mediante funciones no lineales genéricas y regresión mediante una función particularizada para el trozo central de la curva que es el más representativo. Se ha obtenido con cada una de estas vías el coeficiente de correlación entre los datos obtenidos experimentalmente y los datos regresionados, que nos indicará lo buena que es la caracterización, junto con el vector de parámetros que la caracteriza. El proceso completo viene dado en la Figura 1.1.

Como se observa en dicha figura, se realizará un primer estudio sobre el suavizado y el troceado de las curvas para después proceder a la posterior codificación de la curva mediante las dos vías citadas.

Notar que se sitúa en diez el número máximo de parámetros puesto que estos datos servirán para una posterior clasificación de las curvas incrementándose el tiempo de identificación notablemente si se supera tal máximo.

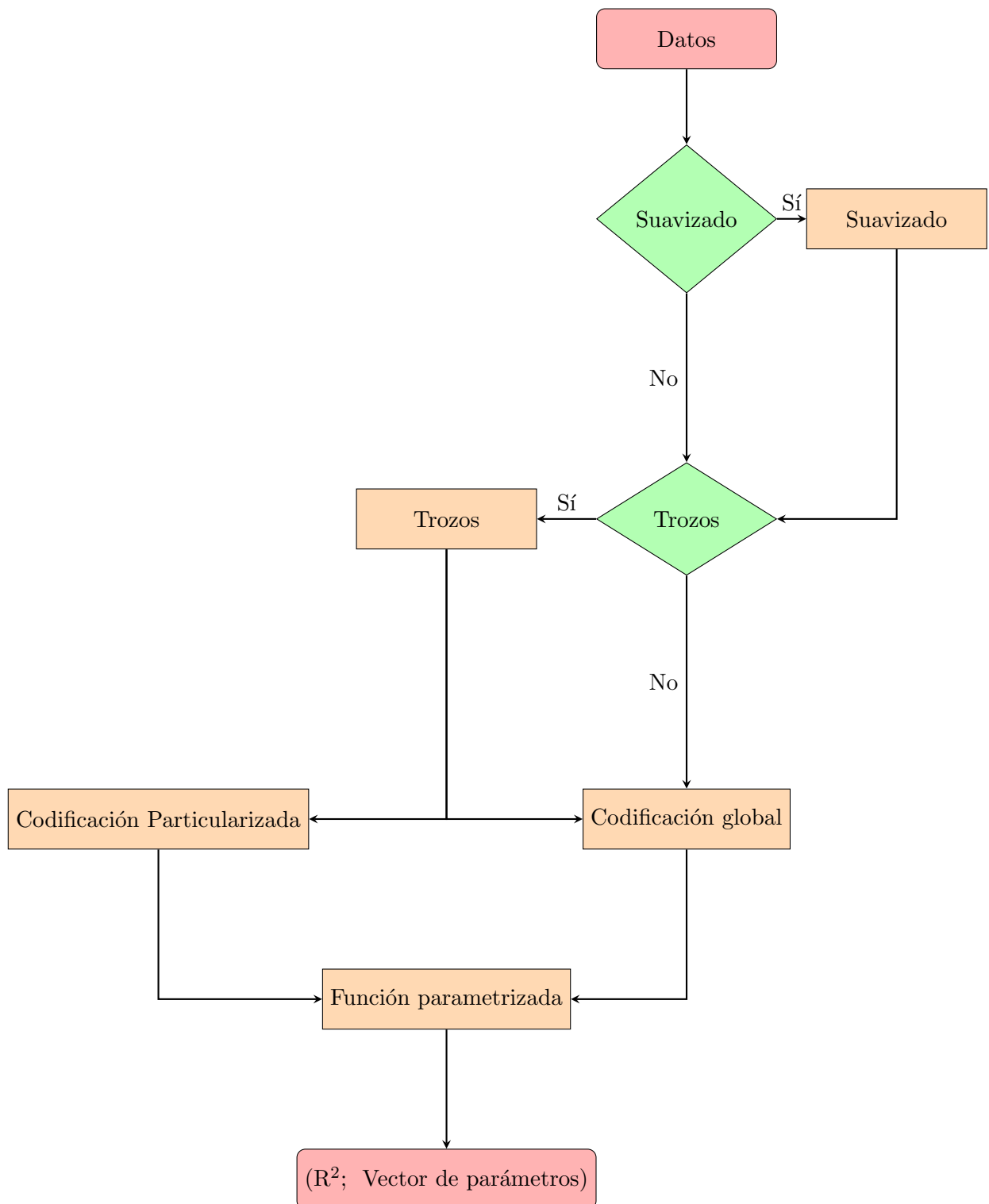


Figura 1.1: Vías utilizadas para la caracterización de la curva.

Capítulo 2

Suavizado y Regresión

En este capítulo se exponen las herramientas matemáticas utilizadas en el proceso de codificado de nuestras curvas. Dichas herramientas son de carácter general y no solo sirven para dicho trabajo. La estructura del capítulo se divide en dos partes: suavizado y regresión. En la parte de regresión el proceso a seguir es el estudio del modelo teórico general seguido por el estudio del caso lineal para finalmente terminar con el caso no lineal.

2.1 Suavizado

En esta sección, vamos a hablar de diferentes filtros digitales utilizados en el proceso de suavizado: `sgolay`, `loess`, `lowess`, `rloess` y `rloess`.

2.1.1 Filtro `sgolay` (Savitzky-Golay)

El filtro de Savitzky-Golay (codificado como `sgolay`) fue descrito por primera vez en 1964 por Abraham Savitzky y Marcel J. E. Golay: *Smoothing and differentiation of data by simplified least squares procedures* [13]. Este método se basa en el cálculo de una regresión polinómica local (de grado k), con al menos $k + 1$ puntos equiespaciados, para determinar el nuevo valor de cada punto (ver Figura 2.1). El resultado será una función similar a los datos de entrada, pero suavizada. Existen métodos para calcular desde la primera a la quinta derivada. En la Figura 2.2 se muestra el suavizado para una función con dicho filtro mediante un polinomio de grado 4.

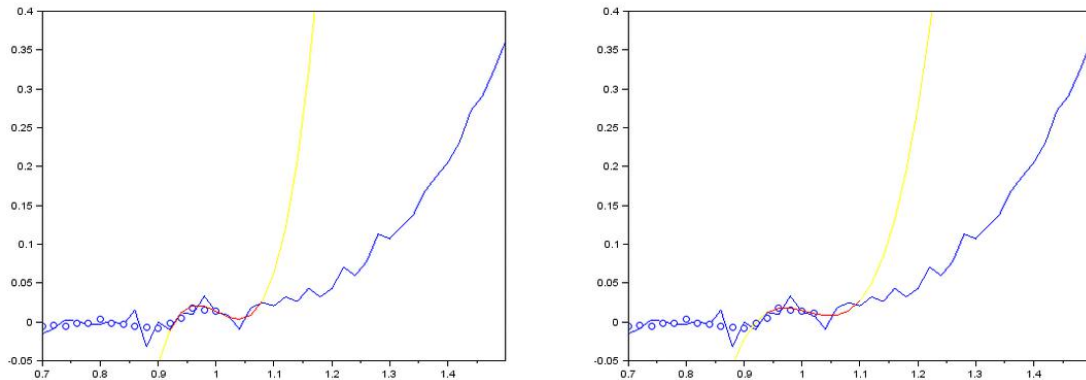


Figura 2.1: Ejemplo regresión local por un polinomio de grado 3.

La principal ventaja de esta aproximación es que tiende a preservar características de la distribución inicial tales como los extremos relativos o el ancho de los picos, que normalmente desaparecen con otras técnicas como la media desplazada¹.

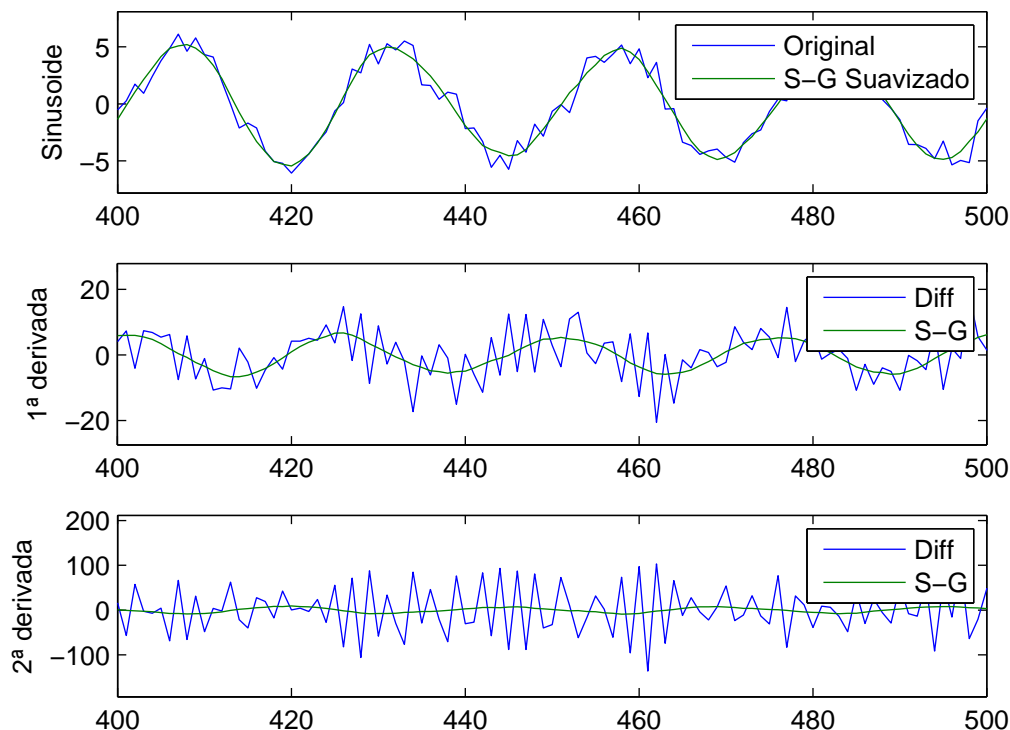


Figura 2.2: Suavizado de una función, de su primera y segunda derivada con el filtro `sgolay`.

2.1.2 Filtros `lowess` y `loess` (locally weighted scatter plot smooth)

El proceso de suavizado se considera local porque, al igual que el método de media desplazada, cada valor suavizado se determina por los datos vecinos. El proceso se pondera por una función de regresión que se define para dichos puntos. Además de esta función de ponderación, se puede utilizar una función de peso robusto, lo que hace que el proceso sea resistente a los «*outliers*». Finalmente, los métodos se diferencian por el modelo utilizado en la regresión: `lowess` utiliza un polinomio lineal, mientras que `loess` utiliza un polinomio cuadrático.

Métodos robustos: `rlowess` y `rloess`

Si los datos contienen «*outliers*», los valores suavizados pueden distorsionarse, y no reflejar el comportamiento de la mayor parte de los puntos vecinos. Para evitar este problema, se pueden suavizar los datos utilizando un procedimiento robusto que no está influenciado por una pequeña fracción de los «*outliers*».

En la Figura 2.3 se puede observar que los valores suavizados no quedan distorsionados por los «*outliers*» al utilizar el método robusto, reflejando mejor el comportamiento de todo el conjunto de puntos.

¹Para saber más sobre el funcionamiento del proceso de la media desplazada visitar:

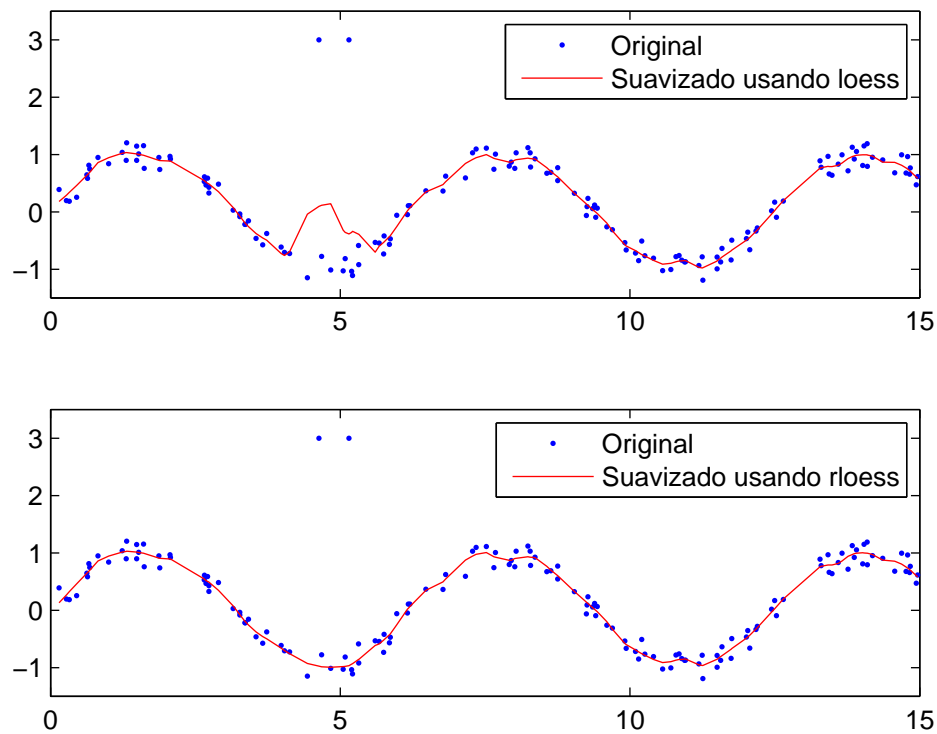


Figura 2.3: Comparación entre el método loess y rloess.

Se puede encontrar más información sobre estos filtros en numerosos libros. A continuación se citan algunos de ellos: *An Evaluation of Smoothing Filters for Gas Sensor Signal Cleaning* [2], *Data Filtering Technique for Neural Networks Forecasting* [15] y *Filters for RSSI-based measurements in a Device-free Passive Localisation Scenario* [4].

En la siguiente sección se va a proceder al análisis de la regresión tanto lineal como no lineal. Aunque existe numerosa literatura clásica como por ejemplo *Introduction to Mathematical Statistics* [7], dicho análisis se basará en el libro *Introducción a la Estadística* [10].

2.2 Regresión

Antes de comenzar con el estudio del modelo teórico general se dará una breve reseña histórica sobre el término. Para más información consultar *Hereditary Genius, An Inquiry Into Its Laws and Consequences* [5].

El término «regresión» fue acuñado por Francis Galton en el siglo XIX para describir un fenómeno biológico. Galton observó que las alturas de los descendientes de ancestros altos tienden a regresar hacia un promedio normal de altura (un fenómeno conocido como regresión a la media). Para Galton, la regresión solo tenía este significado biológico, pero su trabajo se extendió más tarde por Udney Yule y Karl Pearson a un contexto estadístico más general.

<http://es.mathworks.com/help/curvefit/smoothing-data.htm#lbs5ojev>

Se denomina regresión al proceso general de ajustar una función a unos datos. El criterio de minimización empleado para tal ajuste es fundamental a la hora de determinar la curva de regresión, siendo habitual utilizar el método de mínimos cuadrados. Según que la función a ajustar sea lineal o no lineal en los parámetros hablaremos de regresión lineal o regresión no lineal, respectivamente. En ambos casos, el objetivo es el mismo: encontrar las mejores estimaciones de los parámetros y cuantificar la precisión de los mismos.

Matemáticamente, se puede describir un modelo de regresión entre las variables X e Y como:

$$Y = f(X, \theta) + \varepsilon(\theta)$$

donde θ es un vector de parámetros del modelo de regresión que se estima con los valores de los elementos de la muestra y $\varepsilon(\theta)$ es un término de error cometido al explicar la variable Y con la curva dada por la función $f(X, \theta)$.

El correspondiente muestral de este modelo teórico de regresión está dado por una curva de regresión de Y sobre X definida como:

$$y_i = f(x_i, \theta^*) + e_i(\theta^*)$$

donde θ^* define la curva que minimiza la distancia entre los puntos (x_i, y_i) y la curva. Los elementos $e_i(\theta^*)$ se denominan residuos y están definidos por:

$$e_i = y_i - \hat{y}_i, \quad i = 1, \dots, n$$

donde y_i es la ordenada e $\hat{y}_i = f(x_i, \theta^*)$ denota la observación estimada correspondiente al modelo de regresión establecido.

Como se ha expuesto al principio de la sección, el criterio de minimización empleado para ajustar la curva a los datos es fundamental para determinar la curva de regresión. Nos centraremos en el estudio del criterio de mínimos cuadrados.

Mínimos cuadrados

Al igual que antes, se comenzará con unos datos históricos. Para saber más sobre Gauss y su aportación a la Estadística consultar *Gauss y la Estadística* [8].

El método de mínimos cuadrados tiene una larga historia que se remonta a los principios del siglo XIX. En junio de 1801, Zach, un astrónomo que Gauss había conocido dos años antes, publicaba las posiciones orbitales del cuerpo celeste Ceres, un nuevo pequeño planeta descubierto por el astrónomo italiano G. Piazzi en ese mismo año. Desafortunadamente, Piazzi solo había podido observar 9 grados de su órbita antes de que este cuerpo desapareciese tras el sol. Zach publicó varias predicciones de su posición incluyendo una de Gauss que difería notablemente de las demás. Cuando Ceres fue redescubierto por Zach en diciembre de 1801 estaba casi exactamente donde Gauss había predicho. Aunque todavía no había revelado su método, Gauss había obtenido el método de mínimos cuadrados. El francés Legendre desarrolló el mismo método de forma independiente en 1805.

Como su nombre indica, el objetivo del método es obtener los valores de las componentes del vector de parámetros θ^* que minimicen la suma de los cuadrados de los residuos:

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Sustituyendo el valor de \hat{y}_i en la expresión anterior se obtiene:

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - f(x_i, \theta_0))^2$$

Dado que el objetivo es minimizar esta función con respecto a las componentes del vector θ^* , el siguiente paso es calcular la primera derivada con respecto a cada una de las componentes.

Supongamos el vector $\theta^* = (\theta_1^*, \dots, \theta_k^*) \in \mathbb{R}^k$. Así, se tiene la siguiente expresión:

$$F_j = \sum_{i=1}^n \frac{\partial}{\partial \theta_j^*} (y_i - f(x_i, \theta^*))^2, \quad j = 1, \dots, k.$$

Para obtener los valores de θ_j^* hay que resolver el sistema, posiblemente no lineal:

$$F_j = 0, \quad j = 1, \dots, k \quad (2.1)$$

que recibe el nombre de ecuaciones normales.

Bondad de ajuste

Una vez obtenida la curva de regresión queda por determinar la calidad del ajuste. Para realizar esto, se define el coeficiente de determinación R^2 . Este término es función del cociente de la variabilidad de los residuos y la variabilidad total de los datos de la muestra y_1, \dots, y_n .

Definición 2.2.1. Se define el coeficiente de determinación R^2 como:

$$R^2 = 1 - \frac{s_R^2}{s_Y^2}$$

que representa la proporción de la variación de Y explicada por el modelo de regresión, donde:

$$s_R^2 = \frac{1}{n} \sum_{i=1}^n e_i^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Proposición 2.2.2. El coeficiente de determinación satisface las siguientes propiedades:

- $0 \leq R^2 \leq 1$.
- Si $R^2 = 1$, el ajuste de la curva a los datos es perfecto.
- Si $R^2 = 0$, el ajuste es nulo, es decir, la variación de X no explica, en absoluto, la variación de Y .
- $R^2 = r_{XY}^2$, donde r_{XY}^2 es el coeficiente de correlación muestral.

Finalmente, también se puede definir el coeficiente de determinación ajustado.

Definición 2.2.3. Se define el coeficiente de determinación ajustado, \tilde{R}^2 , como:

$$\tilde{R}^2 = 1 - \frac{s_R^2}{s_Y^2},$$

donde:

$$\tilde{s}_R^2 = \frac{1}{n-k} \sum_{i=1}^n e_i^2 = \frac{1}{n-k} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Esta medida es de utilidad en modelos de regresión donde el número de variables explicatorias es mayor que uno. En el análisis de regresión simple no tiene mayor relevancia y es reemplazado por R^2 .

En lo que sigue, se detallará cómo resolver el sistema (2.1) para el caso lineal y el no lineal.

2.2.1 Regresión lineal

La regresión lineal es aquella cuya curva de regresión $y = f(X, \theta) + \varepsilon(\theta)$ viene dada por una función $f(X, \theta)$ lineal en los parámetros y cuyo objetivo es estudiar la relación causal lineal entre dos variables.

Así, el modelo de regresión lineal simple viene dado por:

$$Y = \beta_0 + \beta_1 X + \varepsilon(\beta_0, \beta_1)$$

En tal caso, se llega a:

$$y_i = b_0 + b_1 x_i + e_i, \quad i = 1, \dots, n$$

donde b_0 y b_1 definen la recta que minimiza la distancia entre los puntos de la muestra y la recta.

En el caso de la regresión lineal la solución del problema de mínimos cuadrados es única.

Partiendo de la ecuación:

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - (b_0 + b_1 x_i))^2$$

y resolviendo las ecuaciones normales se llega a la solución exacta:

$$b_0 = \bar{y} - b_1 \bar{x}$$

$$b_1 = \frac{\left(\frac{1}{n} \sum_{i=1}^n x_i y_i \right) - \bar{y} \bar{x}}{\left(\frac{1}{n} \sum_{i=1}^n x_i^2 \right) - \bar{x}^2} = \frac{s_{XY}}{s_X^2}$$

donde s_{XY} y s_X^2 son la covarianza y varianza muestral de X , respectivamente.

Numéricamente es mejor resolverlo mediante la factorización QR ya que este es un algoritmo estable (*Algorithms for the QR-Decomposition* [6]). A continuación ilustraremos con un ejemplo sencillo el porqué de tal afirmación:

Ejemplo 2.2.4. Considerar la matriz:

$$A = \begin{pmatrix} 1 & 1 \\ \delta & 0 \\ 0 & \delta \end{pmatrix}$$

Si $0 < \delta \ll 1$, las dos columnas de A son casi idénticas pero linealmente independientes. Las ecuaciones normales hacen que la situación se empeore, ya que la matriz de coeficientes es:

$$A^T A = \begin{pmatrix} 1 + \delta^2 & 1 \\ 1 & 1 + \delta^2 \end{pmatrix}$$

y si $|\delta| < 10^{-8}$, la matriz anterior es singular en aritmética de doble precisión.

Una vez encontrados los valores de los parámetros que minimizan la distancia entre los puntos y la recta de regresión para el caso lineal, se abordará el problema no lineal.

2.2.2 Regresión no lineal

La regresión no lineal es aquella cuya curva $y = f(X, \theta) + \varepsilon(\theta)$ de regresión viene dada por una función $f(X, \theta)$ no lineal respecto a θ . A diferencia de la regresión lineal, el problema de encontrar los parámetros óptimos ya no tiene por qué tener una solución única ni de tipo explícito puesto que el sistema (2.1) es no lineal. En su lugar, hay que utilizar métodos iterativos, que tratan de buscar con diferentes estrategias el mínimo de la suma de los cuadrados de los residuos.

Método de Newton

Considerar el sistema de n -ecuaciones no lineales con n -incógnitas en notación vectorial:

$$F(x) = (F_1(x), F_2(x), \dots, F_n(x)) = 0 \quad (2.2)$$

La solución de este sistema de ecuaciones implica encontrar $\alpha \in \mathbb{R}^n$ tal que $F(\alpha) = 0$. Sea x_r una aproximación a la misma. Linealizando en x_r y suponiendo que $F(x_{r+1}) = 0$:

$$0 \simeq F(x_r) + J(x_r)(x_r - x_{r+1}) \quad (2.3)$$

define el vector x_{r+1} , donde $J(x_r)$ es la matriz Jacobina de F evaluada en x_r .

Resolviendo el sistema lineal (2.3) se obtiene:

$$x_{r+1} = x_r - J^{-1}(x_r)F(x_r)$$

que es el método de Newton para n -variables. La formula anterior proporciona una relación de recurrencia a partir de un vector de partida x_0 , que convergerá a la raíz si x_0 se encuentra en la región de convergencia de la misma y, si además, la matriz $J(x_r)$ es regular $\forall r$.

En general, se procede actualizando el valor de x_r por ξ_r ,

$$x_{r+1} = x_r + \xi_r \quad (2.4)$$

siendo ξ_r solución del sistema lineal:

$$J(x_r)\xi_r = -F(x_r) \quad (2.5)$$

Observación 2.2.5. *Las propiedades más notables del método de Newton son:*

- *Cuando converge, suele ser muy rápido (convergencia cuadrática).*
- *Es adecuado para problemas con matriz jacobiana «sparse».*

Sin embargo, el método de Newton no está libre de inconvenientes. Un punto clave es la elección del valor inicial x_0 de manera que esté «suficientemente cerca» de la solución buscada. En la práctica, dicha semilla inicial no está disponible por lo que se buscan maneras de alargar el dominio de atracción del método de Newton. Además, en cada iteración del método se requiere una evaluación de la matriz jacobiana, $J(x_r)$, y resolver el sistema de ecuaciones lineales (2.3).

Se realizará a continuación un estudio sobre el primero de los inconvenientes.

Aumentando la convergencia local del método de Newton

Dada una semilla inicial x_0 , se obtiene una sucesión de aproximaciones $\{x_j\}_{j \in \mathbb{N}}$ a la raíz α mediante las ecuaciones (2.4) - (2.5).

El sistema lineal (2.5) define una dirección $\xi \in \mathbb{R}^n$ llamada dirección de Newton y, entonces en (2.4) se da un paso en esta dirección para actualizar el punto x_r al punto x_{r+1} .

Si x_0 no está suficientemente cerca de la raíz α , pueden surgir dos dificultades: bien que el paso en la dirección de Newton sea demasiado grande (ver método de Newton amortiguado), o bien, que la dirección de Newton no sea buena. Se procederá a profundizar más sobre esta segunda circunstancia.

Que la dirección de Newton no sea buena conduce a que el punto x_j no esté suficientemente cerca de α . Esto puede ser debido a que la matriz jacobiana sea numéricamente singular (ver Definición (2.2.6)). En muchos problemas, se puede encontrar una matriz jacobiana numéricamente singular, debido a que x_0 está lejos de α . Para intentar solucionar este obstáculo, habrá que alterar la dirección de Newton.

Se sigue dando unas breves ideas del método de Newton amortiguado necesarias para el desarrollo de la alteración de la dirección de Newton. Para mayor información consultar *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations* [1].

Método de Newton amortiguado

La idea de este método es controlar el tamaño del paso tomado en la dirección de Newton ξ . Por tanto, con ξ dado como solución de (2.5), se modifica (2.4) como sigue:

$$x_{r+1} = x_r + \lambda_r \xi_r, \quad 0 < \lambda_r \leq 1 \quad (2.6)$$

Es claro que tanto λ_r como ξ_r dependen de la iteración en la que se encuentran. Así, se sigue de (2.3) que:

$$\xi_r = -J^{-1}(x_r)F(x_r) \quad (2.7)$$

Ahora, el problema está en cómo escoger el factor de amortiguación λ_r . Para este propósito, se introduce la función objetivo $g : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por:

$$g(x) = \frac{1}{2} \|F(x)\|_2^2 = \frac{1}{2} \sum_{j=1}^n F_j(x)^2 \quad (2.8)$$

que satisface las condiciones:

- $g(x) \geq 0$ y $g(\alpha) = 0 \Leftrightarrow F(\alpha) = 0$.
Luego α minimiza $g(x)$ si $F(\alpha) = 0$, lo que implica que el valor de λ es mejor cuanto menor sea $g(x + \lambda \xi)$.
- La dirección de Newton es de descenso respecto de $g(x)$, es decir:

$$\xi^T \nabla g < 0$$

donde ξ es la dirección de Newton y ∇g es el gradiente de la función objetivo.

Veamos que nuestra función objetivo g definida en (2.8) satisface estas dos condiciones.

Dado que $\nabla g = J^T F$, la función g satisface:

$$\nabla g^T \xi = -2g < 0 \quad (2.9)$$

es estrictamente menor a menos que $x = \alpha$. El significado de tomar una dirección de descenso es que, para $\lambda > 0$ suficientemente pequeño, la serie de Taylor da:

$$g(x_r + \lambda \xi) = g(x_r) + \lambda \xi^T \nabla g(x_r) + O(\lambda^2 |\xi|^2) < g(x_r) \quad (2.10)$$

Luego en la siguiente iteración, x_{r+1} dado por (2.6), debe satisfacer:

$$g(x_{r+1}) < g(x_r) \quad (2.11)$$

es decir, se requiere que g sea monótona decreciente hacia la solución α .

Así, el problema del método de Newton amortiguado es decidir qué punto de la semirrecta $x_r + \lambda\xi$ debe ser escogido para la siguiente iteración.

Una vez vistas estas breves ideas del método de Newton amortiguado, se procede al análisis de la alteración de la dirección de Newton.

Alteración de la dirección de Newton

Definición 2.2.6. Se dice que una matriz A es numéricamente singular si su condicionamiento verifica $K(A)\varepsilon \approx 1$ siendo ε la unidad de redondeo.

Para cuantificar de un modo realista la singularidad numérica, es conveniente usar la función objetivo. Primero se escribe la serie de Taylor para $F(x + \lambda\xi)$ con ξ como en (2.7) como sigue:

$$F(x + \lambda\xi) = (1 - \lambda)F(x) + \int_0^\lambda [J(x + t\xi) - J(x)] \xi dt \quad (2.12)$$

En el caso en el cual $J(x_r)$ es numéricamente singular, la dirección de Newton no es útil, y hay que buscar maneras para alterarla.

El enfoque más sencillo es considerar la minimización de la función objetivo (2.8) directamente, es decir, incrustar el problema de la ecuación no lineal en un problema no lineal de mínimos cuadrados. El método de máxima pendiente sugiere entonces una alternativa para la dirección de Newton generalizada, elegir ξ_r como la dirección de máxima pendiente:

$$\xi_r = -J_r^T F_r \quad (2.13)$$

en $x = x_r$, y entonces usar una búsqueda por semirrectas para determinar la siguiente iteración. Este método también aprovecha la posible estructura «*sparse*» de J . Desafortunadamente, se sabe que la convergencia es más lenta. La experiencia dice que donde la dirección de Newton falla, el método de máxima pendiente no es de mucha ayuda.

Una estrategia más flexible es tomar una combinación de la dirección de Newton y de la de máxima pendiente. La dirección de Newton se expresa primero en términos de un paso de Gauss-Newton, así llamada para el problema de mínimos cuadrados minimizando (2.8). Esto es, si J tiene rango máximo entonces las ecuaciones normales para el problema de mínimos cuadrados:

$$\min |J(x_r)\xi + F(x_r)|^2 \quad (2.14)$$

dan:

$$\xi = -(J^T J)^{-1} J^T F. \quad (2.15)$$

Entonces, si el rango de J no es máximo se toma una combinación de (2.15) y de (2.13):

$$\xi = -(J^T J + \mu I)^{-1} J^T F \quad (2.16)$$

en $x = x_r$, con μ un parámetro de control positivo. Este se conoce como el método de Marquardt. Notar que una matriz semidefinida positiva $J^T J$ se convierte en definida positiva añadiendo μI , $\mu > 0$.

La gran cuestión con respecto al método de Marquardt es como controlar el parámetro μ . Un enfoque eficaz es usar un modelo «*trust-region*», donde la solución de (2.14) se elige sujeta al requisito de que su norma no sea demasiado grande. Por tanto se requiere:

$$|\xi| \leq \delta \quad (2.17)$$

donde δ es un parámetro de control que representa la región en la cual el modelo cuadrático, cuyo mínimo está dado por la solución de (2.14), modela adecuadamente la función objetivo. Este enfoque también proporciona una manera natural de decidir cuando se altera la dirección de Newton, es decir, cuando la dirección de Newton no satisface (2.17).

Capítulo 3

Codificación de las curvas

En este capítulo se muestra una codificación de las curvas planas correspondientes a la sustancia nitrotolueno haciendo uso del método de Marquardt descrito en el capítulo anterior para la resolución de ecuaciones no lineales.

3.1 Caso de estudio

Para la elaboración del presente trabajo se dispone de un total de 16 curvas correspondientes a la sustancia nitrotolueno con diferentes partes por millón (ppm de aquí en adelante), empezando con 25 ppm y terminando con 100 ppm con una diferencia de 5 ppm entre cada curva. El tamaño de cada una de las muestras está comprendido entre 300 y 320 puntos.

En todo el capítulo se tomará como curva ejemplo la correspondiente a 25 ppm mostrada en la Figura 3.1. La codificación del resto de las curvas se puede ver en el Anexo.

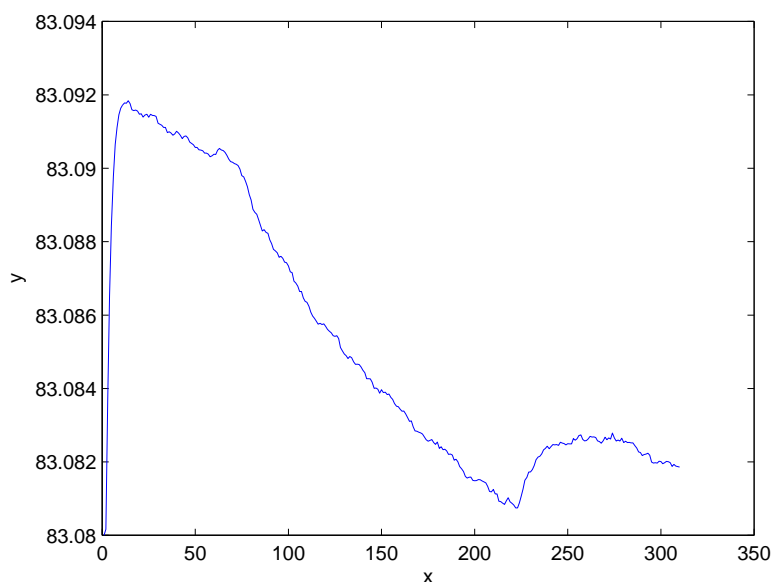


Figura 3.1: Curva correspondiente a 25ppm Nitrotolueno.

3.2 Suavizado de la curva

Se han estudiado diferentes filtros para la realización del suavizado de la curva, codificados como `sgolay`, `rlowess` y `rloess`. Para estudiar cuál es el mejor para nuestras curvas se se tomará aquel que mayor valor de R^2 proporcione. En la siguiente tabla se muestran los datos obtenidos con cada uno de los filtros anteriormente citados para la curva ejemplo:

Filtro	R^2
<code>sgolay</code>	0.9752
<code>rloess</code>	0.9446
<code>rloess</code>	0.9518

A continuación, se muestra un gráfico con el valor del R^2 obtenido al suavizar todas las curvas con los diferentes filtros. Los datos numéricos se encuentran en el Anexo.

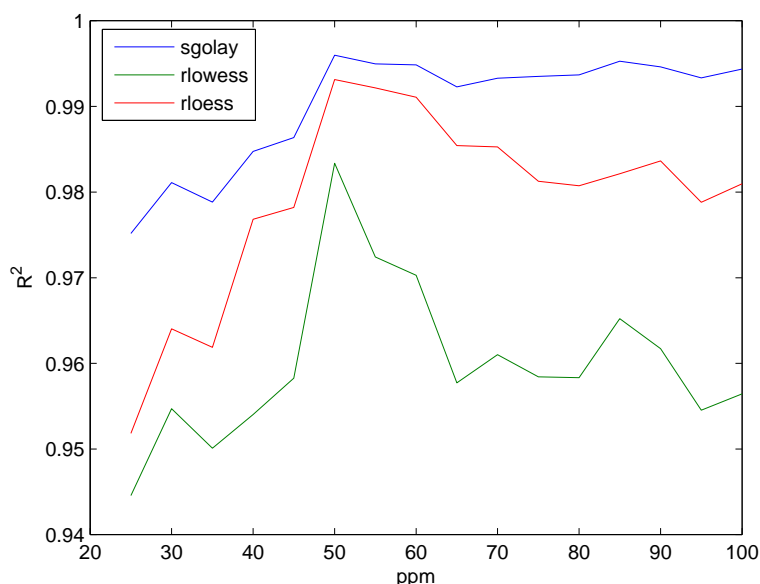


Figura 3.2: Suavizado con diferentes filtros.

Puede observarse en la figura anterior que para todas las curvas consideradas, el mejor filtro es el codificado de `sgolay`. En lo que sigue, cuando se hable de datos suavizados se entenderá que han sido suavizados con este filtro.

3.3 Codificación global

Primero, se abordará la codificación global de nuestras curvas mediante la primera de las dos vías, la regresión mediante funciones no lineales genéricas. Para esta codificación global, se realizará un «*bombardeo fit*» con las funciones mostradas en la siguiente tabla:

Parámetros	Método	$f(X, \theta)$
2	exp1	$a_0 \exp(a_1 x)$
4	exp2	exp1 + $a_2 \exp(a_3 x)$
1+3	fou1	$a_0 + a_1 \cos(wx) + b_1 \sin(wx)$
1+5	fou2	fou1 + $a_2 \cos(2wx) + b_2 \sin(2wx)$
1+7	fou3	fou2 + $a_3 \cos(3wx) + b_3 \sin(3wx)$
3	gauss1	$a_0 \exp(-(x - a_1)^2/b_1^2)$
6	gauss2	gauss1 + $a_2 \exp(-(x - a_3)^2/b_2^2)$
9	gauss3	gauss2 + $a_4 \exp(-(x - a_5)^2/b_3^2)$

Estas funciones han sido escogidas puesto que son estándar en el software utilizado (Matlab).

Se muestra la calidad del ajuste con cada una de las funciones anteriores para la curva ejemplo, mediante el coeficiente de correlación entre los datos suavizados y los datos procedentes de la regresión, R^2 . No hay que olvidar que el objetivo del presente trabajo es encontrar un equilibrio entre dicho coeficiente de correlación y la dimensión del vector de parámetros θ .

Método	R^2	\tilde{R}^2
exp1	0.8046	0.8040
exp2	0.8994	0.8984
fou1	0.9503	0.9499
fou2	0.9941	0.9940
fou3	0.9982	0.9981
gauss1	0.7838	0.7824
gauss2	0.9920	0.9918
gauss3	0.9895	0.9892

Se observa que el mejor ajuste se corresponde con el método de función fou3, es decir, con:

$$f(X, \theta) = a_0 + a_1 \cos(wx) + b_1 \sin(wx) + a_2 \cos(2wx) + b_2 \sin(2wx) + a_3 \cos(3wx) + b_3 \sin(3wx)$$

donde $\theta = (a_0, a_1, \dots, w)$. Dicho método utiliza un total de 8 parámetros que está por debajo del umbral fijado inicialmente en un máximo de 10 parámetros. Gráficamente, el ajuste es el mostrado en la Figura 3.3.

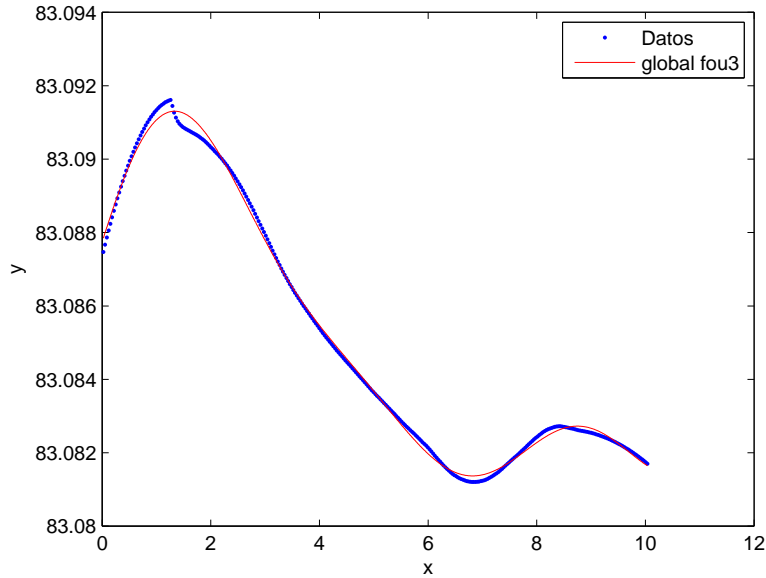
Como se ha expuesto en el capítulo anterior, no se puede dar un valor exacto para los parámetros que componen el vector θ , por lo que a continuación se muestran los valores de los parámetros junto con su intervalo de confianza al 95 %.

Así, con la semilla inicial $\theta_0 = (83.08, 0, 0, 0, 0, 0, 0, 0.3136)$, se obtienen los siguientes resultados para los parámetros $(a_0, a_1, b_1, a_2, b_2, a_3, b_3, w)$:

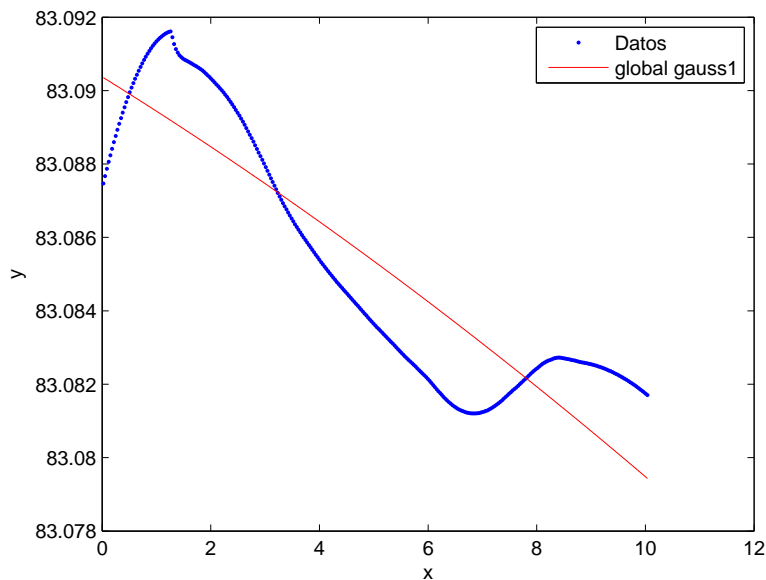
$$\begin{aligned} a_0 &= 83.08 \\ a_1 &= 0.00269 \in [0.002647, 0.002734] \\ b_1 &= 0.003321 \in [0.003287, 0.003355] \\ a_2 &= -0.0001061 \in [-0.0001556, -5.654 \cdot 10^{-5}] \\ b_2 &= 0.001823 \in [0.001701, 0.001944] \\ a_3 &= 0.0003441 \in [0.0002716, 0.0004165] \\ b_3 &= 0.0007918 \in [0.0007568, 0.0008268] \\ w &= 0.503 \in [0.495, 0.5109] \end{aligned}$$

Por tanto, la codificación para esta curva será la siguiente:

$$(R^2, \theta) = (0.9982; (83.08, 0.00269, 0.003321, -0.0001061, 0.001823, 0.0003441, 0.0007918, 0.503))$$

Figura 3.3: Ajuste con `fou3`.

Se refleja el ajuste de la curva ejemplo con el método codificado como `gauss1`. Como este método es el que menor valor de R^2 proporciona, la relación entre los datos suavizados y la curva de ajuste no parece adecuada (ver Figura 3.4).

Figura 3.4: Ajuste con `gauss1`.

Los valores de los parámetros se obtienen mediante métodos iterativos. En estos métodos, la semilla inicial juega un papel muy importante a la hora de hallar la solución buscada. Si se cambia la semilla inicial y se parte de $\theta_0 = \bar{0}$, los valores de los parámetros varían de

manera significativa:

$$\begin{aligned}
 a_0 &= 1.224 \cdot 10^{11} \in [-5.576 \cdot 10^{12}, 5.821 \cdot 10^{12}] \\
 a_1 &= 0.00 \in [-8.077 \cdot 10^{11}, 8.077 \cdot 10^{11}] \\
 b_1 &= 0.00 \in [-8.62 \cdot 10^{18}, 8.62 \cdot 10^{18}] \\
 a_2 &= 0.00 \in [-4.331 \cdot 10^{12}, 4.331 \cdot 10^{12}] \\
 b_2 &= 0.00 \in [-6.831 \cdot 10^{18}, 6.831 \cdot 10^{18}] \\
 a_3 &= -1.224 \cdot 10^{11} \in [-4.462 \cdot 10^{12}, 4.217 \cdot 10^{12}] \\
 b_3 &= -5.178 \cdot 10^4 \in [-1.823 \cdot 10^{18}, 1.823 \cdot 10^{18}] \\
 w &= 1,49 \cdot 10^{-08} \in [-2.761 \cdot 10^{-7}, 3.059 \cdot 10^{-7}]
 \end{aligned}$$

Además, el valor de R^2 que se obtiene partiendo del vector nulo como semilla inicial es 0.8738. Luego, tanto el valor de R^2 como los intervalos de confianza de cada uno de los parámetros indican que este ajuste no es lo suficientemente bueno.

3.4 Codificación global por trozos

Se puede dividir cada una de las curvas en tres trozos como se muestra en la Figura 3.5. Estos tres trozos se rigen por el sentido físico del problema:

- El primer trozo corresponde con el encendido y calentamiento de la «nariz electrónica».
- El segundo trozo corresponde a la introducción de la sustancia y al comienzo real del experimento.
- El tercer y último trozo corresponde a la retirada de la sustancia y la recuperación de la «nariz electrónica».

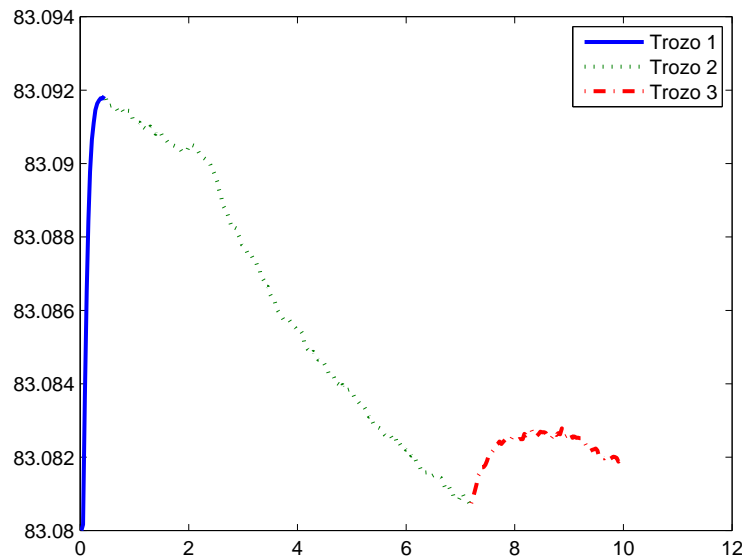


Figura 3.5: Trozos en la gráfica correspondiente a 25ppm.

Por lo que se puede pensar en una codificación global por trozos ajustando cada uno de ellos por separado.

Así, en la siguiente tabla se muestra el valor de R^2 para cada uno de los trozos de la curva ejemplo utilizando los mismos métodos que para la codificación global.

Método	Trozo 1	Trozo 2	Trozo 3
exp1	0.7662	0.9836	0.0200
exp2	0.9704	0.9873	0.8465
fou1	0.9736	0.9952	0.8499
fou2	0.9863	0.9977	0.9630
fou3	0.9988	0.9991	0.9945
gauss1	0.7738	0.9823	0.8561
gauss2	0.9737	0.9960	0.9150
gauss3	0.9737	0.9960	0.9434

Lo deseable sería que para esta codificación global por trozos se usase el mínimo número de parámetros posibles en cada uno de los trozos. Pero como se observa en la tabla anterior, para obtener un ajuste con una calidad superior a 0.99 en cada uno de ellos se necesita un total de $8 + 8 + 8 = 24$ parámetros. Por lo que esta codificación global por trozos no mejora la codificación global obtenida anteriormente puesto que bastaba con 8 parámetros para obtener una calidad de ajuste de 0.9982.

3.5 Codificación del trozo central

Gracias al sentido físico del problema descrito con anterioridad, todos los esfuerzos se van a basar en codificar el trozo central puesto que se corresponde con la introducción de la sustancia.

De la misma manera que en la codificación global, se mostrará en una tabla la calidad del ajuste de este tramo central con cada una de las funciones no lineales que hemos ido describiendo a lo largo de nuestro trabajo. Así, la tabla sería la siguiente:

Método	R^2	\tilde{R}^2
exp1	0.9836	0.9836
exp2	0.9874	0.9872
fou1	0.9952	0.9951
fou2	0.9977	0.9976
fou3	0.9991	0.9991
gauss1	0.9823	0.9821
gauss2	0.9972	0.9971
gauss3	0.9979	0.9979

Como se puede observar, todos los ajustes obtienen un valor de $R^2 > 0.98$ lo cual es muy alto, pero el mejor método de ajuste sigue correspondiendo con **fou3** (ver Figura 3.6). Se muestra también el gráfico correspondiente al ajuste con **gauss1** al igual que en la codificación global.

En este caso donde solo codificamos el segundo trozo de la curva, con este método de **fou3** se necesitan un total de 8 parámetros, lo cual no proporciona ninguna mejora con respecto a la codificación global, puesto que utiliza el mismo número de parámetros.

3.5.1 Codificación particularizada

Ahora, se podría pensar que cualquiera de los métodos anteriores nos sirve para obtener una buena codificación de esta curva, puesto que como se ha indicado, todos los ajustes

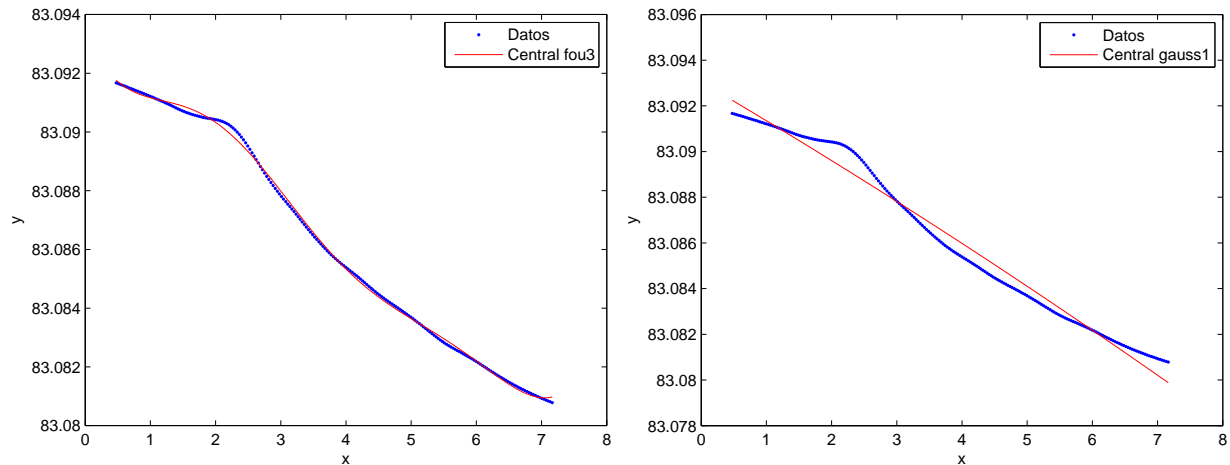


Figura 3.6: Ajuste trozo central fou3 y gauss1.

proporcionan un valor de $R^2 > 0.98$. A pesar de esto, viendo la forma de la gráfica en el intervalo central, se intenta obtener una función $f(x, \theta)$ distinta a las anteriores con un número menor de parámetros que consiga una calidad de ajuste mejor a la que se muestra en la tabla anterior mediante los métodos clásicos.

Así, con estas premisas, se busca una función $f(x, \theta)$ con cuatro parámetros de forma que mejora en muchos de los casos la calidad de ajuste obtenida con los métodos ya expuestos. Tal función es la siguiente:

$$f(x, \theta) = a \exp(-bx) + c \sin(dx), \quad b \geq 0$$

Como se ve, es una combinación de una función trigonométrica y una exponencial con cuatro parámetros, es decir, el vector θ es (a, b, c, d) . Dicha función se denota función particularizada (por comodidad a la hora de referenciarla).

En la siguiente tabla se muestran los resultados obtenidos comparándolos con los métodos fou1 y exp2 (todos ellos con cuatro parámetros).

Curva	Particularizada	fou1	exp2
25 ppm	0.9959	0.9952	0.9874
30 ppm	0.9970	0.9953	0.9878
35 ppm	0.9952	0.9929	0.9841
40 ppm	0.9916	0.9862	0.9757
45 ppm	0.9928	0.9872	0.9763
50 ppm	0.9907	0.9851	0.9737
55 ppm	0.9911	0.9810	0.9733
60 ppm	0.9891	0.9807	0.9684
65 ppm	0.9893	0.9792	0.9653
70 ppm	0.9882	0.9758	0.9654
75 ppm	0.9885	0.9760	0.9638
80 ppm	0.9872	0.9736	0.9638
85 ppm	0.9879	0.9755	0.9616
90 ppm	0.9868	0.9729	0.9607
95 ppm	0.9866	0.9714	0.9630
100 ppm	0.9864	0.9714	0.9677

Para dichos métodos, la función particularizada obtiene una mejora en la calidad de ajuste. Además, se observa que el error relativo es casi lineal en ppm/10 (ver Figura 3.7).

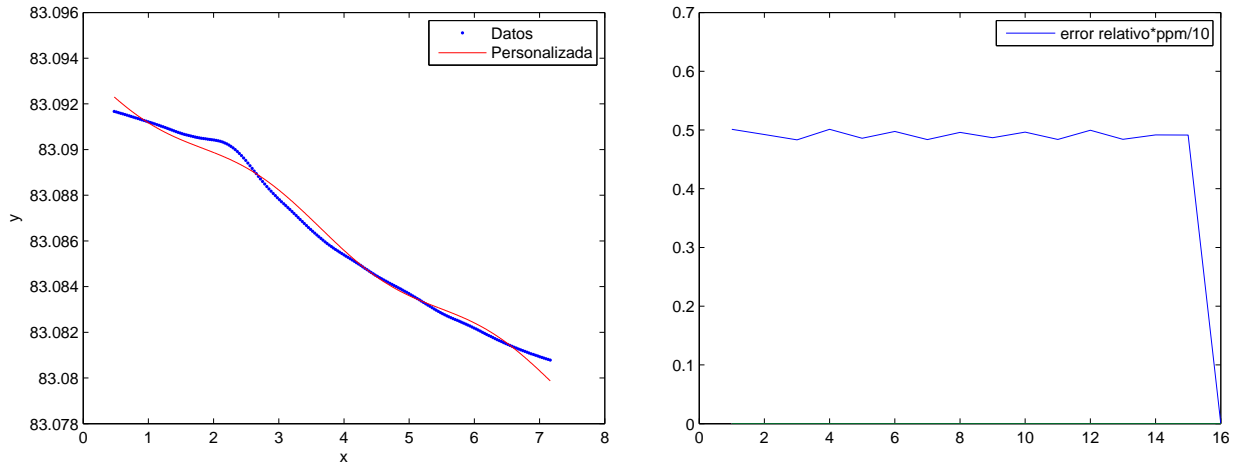


Figura 3.7: Ajuste tramo central particularizada y error relativo*ppm/10.

Volviendo a nuestra curva ejemplo se obtienen los siguientes valores para los parámetros (a, b, c, d) partiendo de la semilla inicial $\theta_0 = (0.5465, 0.0902, 0.3516, 0.9157)$, indicando los intervalos de confianza al 95 %:

$$\begin{aligned} a &= 83.09 \\ b &= 1.871 \cdot 10^{-5} \in [1.824 \cdot 10^{-5}, 1.918 \cdot 10^{-5}] \\ c &= 0.000978 \in [0.0008861, 0.00107] \\ d &= 0.8789 \in [0.8647, 0.8931] \end{aligned}$$

Así, la codificación para la curva ejemplo es:

$$(0.9959; (83.09, 1.871 \cdot 10^{-5}, 0.000978, 0.8789))$$

3.6 Codificación sin suavizado

Hasta ahora, todo el proceso de codificación de las curvas se ha realizado con los datos suavizados. En esta sección se muestran los datos obtenidos al codificar las curvas con los datos originales.

Para la parametrización global de la curva ejemplo se obtienen los siguientes resultados:

Método	Global
exp1	0.7552
exp2	0.8459
fou1	0.8939
fou2	0.9423
fou3	<u>0.9544</u>
gauss1	0.7357
gauss2	0.9388
gauss3	0.9393

Se observa que el mayor valor de R^2 lo sigue proporcionando el método fou3 pero este valor es mucho menor al obtenido en la codificación global tras el suavizado (0.9982). En general, para esta curva ejemplo hay una diferencia entre los valores de R^2 de 0.05 aproximadamente. Sin embargo, a lo largo de la familia la diferencia va disminuyendo como se puede observar en la Figura 3.8.

Para la codificación del trozo central se han obtenido los siguientes datos:

Método	Central
exp1	0.9829
exp2	0.9866
fou1	0.9944
fou2	0.9969
fou3	0.9983
gauss1	0.9815
gauss2	0.9952
gauss3	0.9968

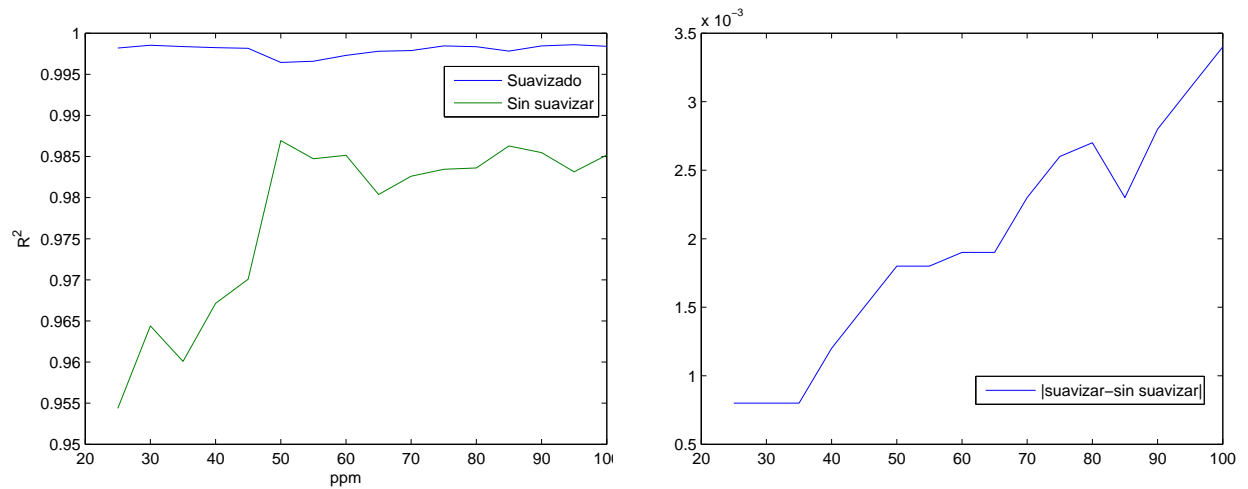


Figura 3.8: Comparación de R^2 entre ajuste global con suavizado y sin suavizado. Diferencia en valor absoluto entre los valores de R^2 para el trozo central.

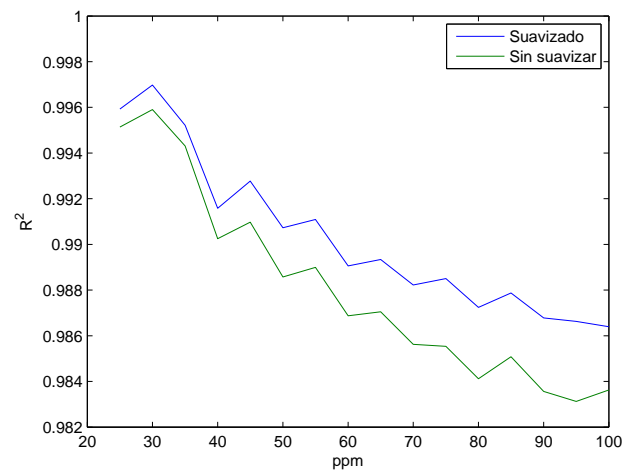


Figura 3.9: R^2 trozo central ajuste particularizado.

En la Figura 3.8 se muestra la diferencia en valor absoluto entre el valor de R^2 que se obtiene al suavizar y el que se obtiene al no suavizar los datos de la familia completa con

fou3. Se observa que dicha diferencia aumenta conforme la curva tiene mayor cantidad de ppm.

Por último, en la Figura 3.9 se muestra la diferencia entre los valores de R^2 para el trozo central pero con el ajuste mediante la función particularizada. Al igual que para el ajuste del trozo central con **fou3**, a lo largo que crecen las ppm la diferencia entre el valor de R^2 al suavizar y no suavizar se acentúa.

Capítulo 4

Conclusiones

En vista a los resultados del capítulo anterior en el cual se ha realizado una codificación tanto global como por trozos de una serie de curvas en \mathbb{R}^2 correspondientes a la sustancia nitrotolueno se obtienen las siguientes conclusiones.

La codificación global de las curvas se obtiene con un total de 8 parámetros que son los correspondientes a la función de ajuste del método `fou3`. Con esta codificación se obtiene una calidad de ajuste para la curva ejemplo con los datos suavizados mediante el método `sgolay` de $R^2 = 0.9982$ (ver Figura 3.3). Si se realiza el ajuste con los datos originales, es decir, sin suavizar el método `fou3` sigue siendo el que mayor calidad de ajuste proporciona aunque siendo algo menor (para la curva ejemplo 0.9544). Además, en la Figura 3.8 se observa que a medida que las ppm aumentan, la diferencia en la calidad de ajuste entre los datos suavizados y los datos originales disminuye aunque siendo mucho más regular la correspondiente a los datos suavizados.

Esta codificación inicial se realiza sin tener en cuenta el sentido físico de la curva. Si se tiene en cuenta, se pueden diferenciar tres trozos en cada una ellas (ver Figura 3.5), por lo que se procede a realizar una codificación de cada uno de los trozos para posteriormente unirlos y obtener así una codificación global. Para obtener una codificación global de esta manera cuya calidad de ajuste esté por encima de 0.99 se necesitan 8 parámetros en cada uno de los trozos. Con lo cual, esta codificación global por trozos no es lo suficientemente buena puesto que utiliza el triple de parámetros que la codificación global. Sin embargo esta idea de trocear la curva gracias al sentido físico ha sido de gran ayuda para la codificación de las curvas puesto que gracias a ello solo se codifica el trozo central.

Inicialmente, para dicha codificación del trozo central se utilizan los mismos métodos que para la codificación global. Con esta estrategia la codificación de las curvas mantiene el mismo número de parámetros que en la codificación global puesto que el método que mejor ajustaba a las curvas sigue siendo `fou3`. Citar también que la diferencia en la calidad de ajuste entre los datos suavizados y originales crece a medida que aumentan las ppm (ver Figura 3.8). Esto no supone una mejoría respecto a la codificación global puesto que aunque el ajuste de la curva es más correcto al codificar solo el trozo central se necesitan el mismo número de parámetros. Por tanto, se busca una función de ajuste $f(x, \theta)$ con un número menor de parámetros de manera que no se pierda calidad en el ajuste. Esa función es:

$$f(x, \theta) = a \exp(-bx) + c \sin(dx), \quad b \geq 0$$

Con esta función, llamada función particularizada, se obtiene una codificación de las curvas con un total de 4 parámetros con una calidad en el ajuste por encima de 0.9864 superando a los métodos utilizados en el presente trabajo con el mismo número de parámetros (`fou1`, `exp2`). Se observa la diferencia entre los valores de R^2 obtenidos al ajustar los datos suavizados y sin suavizar en la Figura 3.9.

Esta codificación particularizada del trozo central tiene una clara ventaja frente a la codificación global: se obtiene una codificación con una calidad muy semejante con la mitad de los parámetros. Por tanto, se puede concluir que gracias a la codificación particularizada se obtiene una codificación de las curvas con un número total de 4 parámetros para cada una de las curvas.

En una visión más general del proyecto, puesto que este presente trabajo solo se centra en la codificación de las curvas, se exponen las posibles líneas futuras para la continuación del mismo:

- Elaboración de un estudio de tales características con nuevos datos de la misma sustancia obtenidos con sensores más eficientes.
- Codificación de más sustancias además de la estudiada en este trabajo.
- Búsqueda de mejores funciones particularizadas para tal codificación descrita en el punto anterior.
- Clasificación de las curvas para el estudio mediante redes neuronales.

Bibliografía

- [1] U.M. Ascher, T.M.M. Mattheij, R.D. Russell; *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, Classics in Applied Mathematics, SIAM, 1995.
- [2] E. Bassey, J. Whalley, P. Sallis; *An Evaluation of Smoothing Filters for Gas Sensor Signal Cleaning*, IARIA, 2014.
- [3] R.L. Burden, J.D. Faires; *Análisis numérico*, International Thomson Editores, 2002.
- [4] G. Deak, K. Curran, J. Condell; *Filters for RSSI-based measurements in a Device-free Passive Localisation Scenario*, 2010.
- [5] F. Galton; *Hereditary Genius, An Inquiry Into Its Laws and Consequences*, Cosimo Classics, 2005.
- [6] W. Gander; *Algorithms for the QR-Decomposition*, 2003.
- [7] R.V. Hogg, J.W. McKean, A. Craig; *Introduction to Mathematical Statistics*, Pearson Education Limited, 7ª Edición, 2013.
- [8] P. Ibarrola; *Gauss y la Estadística*.
- [9] D. Kincaid, W. Cheney; *Análisis Numérico: Las Matemáticas del Cálculo Científico*, Addison-Wesley Iberoamericana, 1994.
- [10] J. Martinez, J. Olmo, M. Rodríguez, A. Fiasconaro; *Introducción a la Estadística*, Centro Universitario de la Defensa, 2012.
- [11] M. Penkova; *Métodos iterativos eficientes para la resolución de sistemas no lineales*, Universidad Politécnica de Valencia, 2011.
- [12] S. Plaza, J.M. Gutiérrez; *Dinámica del método de Newton*, Universidad de La Rioja, Servicio de Publicaciones, 2013.
- [13] A. Savitzky, M.J.E. Golay; *Smoothing and differentiation of data by simplified least squares procedures*, Analytical Chemistry 36, 1964.
- [14] R.W. Schafer; *What is a Savitzky-Golay Filter?*, 2011.
- [15] W. Wettayaprasit, N. Laosen, S. Chevakidagarn; *Data Filtering Technique for Neural Networks Forecasting*, 2007.

Anexo

En este anexo, se presentan los resultados obtenidos para las 16 curvas de la sustancia nitrotolueno gracias al potencial de la metodología conseguida en el presente trabajo. De la misma manera que en el Capítulo 3, se muestra una codificación global, global por trozos y local del trozo central de cada curva indicando el R^2 entre los datos obtenidos experimentalmente y los datos regresionados.

A.1 Suavizado

A continuación se presenta el valor de R^2 correspondiente al suavizado global de todas las curvas con los filtros digitales codificados de `sgolay`, `rloess` y `rloess`. Se hace uso de los métodos robustos para hacer frente a los posibles «*outliers*».

Curva	<code>sgolay</code>	<code>rloess</code>	<code>rloess</code>
25 ppm	0.9752	0.9446	0.9518
30 ppm	0.9811	0.9547	0.9640
35 ppm	0.9788	0.9501	0.9619
40 ppm	0.9847	0.9540	0.9768
45 ppm	0.9864	0.9583	0.9782
50 ppm	0.9960	0.9834	0.9931
55 ppm	0.9950	0.9724	0.9922
60 ppm	0.9949	0.9703	0.9911
65 ppm	0.9923	0.9577	0.9854
70 ppm	0.9933	0.9610	0.9853
75 ppm	0.9935	0.9584	0.9813
80 ppm	0.9937	0.9583	0.9807
85 ppm	0.9953	0.9652	0.9821
90 ppm	0.9946	0.9617	0.9836
95 ppm	0.9933	0.9545	0.9788
100 ppm	0.9944	0.9564	0.9810

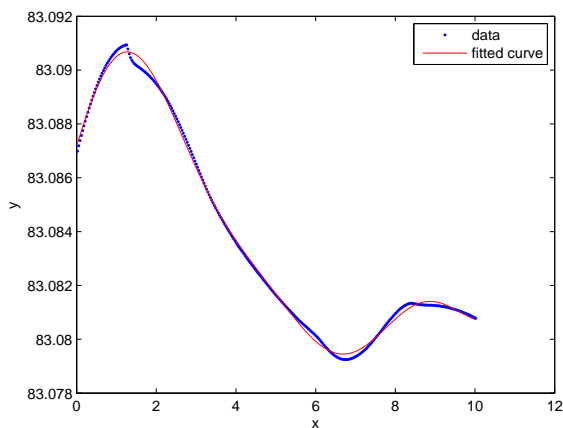
Para todas ellas, el mejor filtro, entendido como el que mayor valor de R^2 proporciona es `sgolay` (ver Figura 3.2).

A.2 Comparación codificación global para todas las curvas

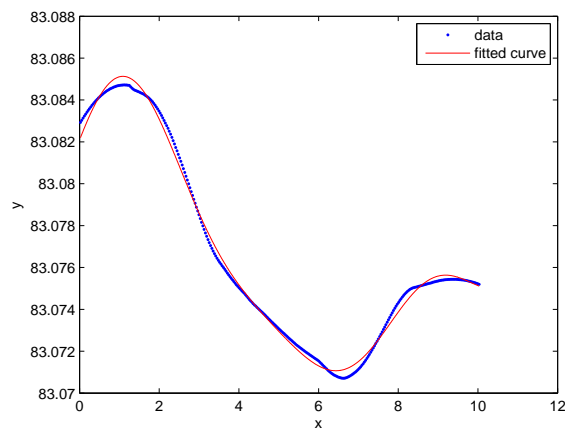
En la siguiente tabla se muestra el valor de R^2 obtenido con todos los métodos descritos en el Capítulo 3 para la familia de curvas correspondientes a la sustancia nitrotolueno. Notar que estos valores se han obtenido con los datos suavizados mediante el filtro codificado de `sgolay` (observar apartado anterior).

Curva	exp1	exp2	fou1	fou2	fou3	gauss1	gauss2	gauss3
25 ppm	0.8046	0.8994	0.9503	0.9941	0.9982	0.7838	0.9920	0.9895
30 ppm	0.7573	0.9044	0.9540	0.9948	0.9985	0.7566	0.9926	0.9912
35 ppm	0.7059	0.8859	0.9465	0.9932	0.9984	0.7052	0.9903	0.9886
40 ppm	0.6271	0.8752	0.9389	0.9913	0.9982	0.6257	0.9876	0.9891
45 ppm	0.5992	0.8821	0.9447	0.9916	0.9982	0.5983	0.9885	0.9886
50 ppm	0.5746	0.9059	0.9539	0.9902	0.9964	0.5738	0.9894	0.9895
55 ppm	0.5437	0.9065	0.9546	0.9893	0.9966	0.5431	0.9870	0.9875
60 ppm	0.4857	0.8843	0.9478	0.9882	0.9973	0.4771	0.9859	0.9866
65 ppm	0.4952	0.8802	0.9439	0.9878	0.9978	0.4942	0.9857	0.9895
70 ppm	0.4815	0.8863	0.9439	0.9859	0.9979	0.4810	0.9836	0.9895
75 ppm	0.4734	0.8852	0.9440	0.9848	0.9985	0.4728	0.9821	0.9892
80 ppm	0.4595	0.8879	0.9468	0.9846	0.9984	0.4590	0.9823	0.9888
85 ppm	0.4656	0.8937	0.9487	0.9855	0.9978	0.4508	0.9831	0.9890
90 ppm	0.4603	0.8932	0.9487	0.9835	0.9984	0.4448	0.9813	0.9895
95 ppm	0.4667	0.8861	0.9437	0.9836	0.9986	0.4664	0.9814	0.9912
100 ppm	0.4524	0.8884	0.9465	0.9834	0.9984	0.4367	0.9812	0.9939

Para todas ellas, el mejor método de ajuste global es el correspondiente con **fou3** obteniéndose valores extremos para R^2 en las curvas correspondientes a 30 y 50 ppm. Estos valores son: 0.9985 (max) y 0.9964 (min), respectivamente. El resultado gráfico de dichos ajustes correspondientes a los casos extremos es el mostrado en la Figura A.1. Ambos ajustes, incluyendo el correspondiente al de valor mínimo de R^2 , son muy buenos, puesto que el valor mínimo, que es 0.9964, es un valor muy cercano a 1 encontrando así el equilibrio buscado entre el número de parámetros y el valor de R^2 .



(a) Ajuste correspondiente a 30ppm.



(b) Ajuste correspondiente a 50ppm.

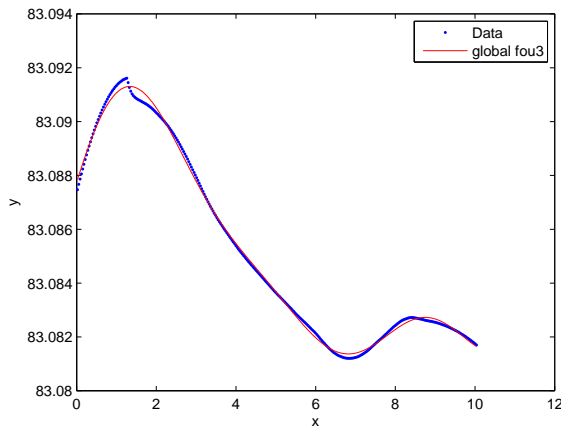
Figura A.1: Ajuste de los casos extremos de R^2 .

A.3 Comparación codificación global, global por trozos y local para cada curva

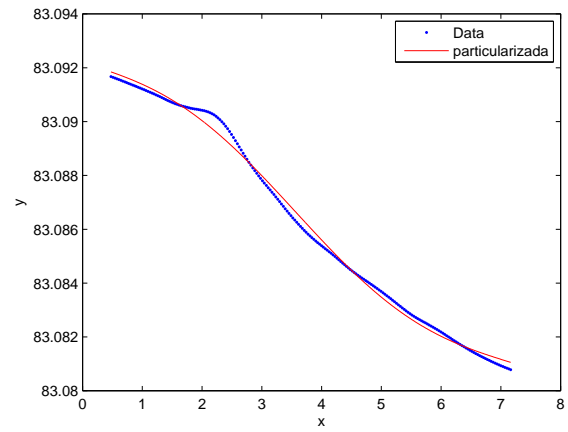
Ahora, se presenta una comparación entre el ajuste global, global por trozos y local para cada una de las curvas indicando el vector de caracterización para cada una de ellas en la sección final.

A.3.1 25 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.8046	0.7662	0.9836	0.0200	0.9836
exp2	0.8994	0.9704	0.9873	0.8465	0.9873
fou1	0.9503	0.9736	0.9952	0.8499	0.9952
fou2	0.9941	0.9863	0.9977	0.9630	0.9977
fou3	0.9982	0.9988	0.9991	0.9945	0.9991
gauss1	0.7838	0.7738	0.9823	0.8561	0.9823
gauss2	0.9920	0.9737	0.9960	0.9150	0.9960
gauss3	0.9895	0.9737	0.9960	0.9434	0.9960
Particularizado	-	-	-	-	0.9959



(a) Ajuste global.



(b) Ajuste particularizado.

Figura A.2: Comparación ajuste global y particularizado 25 ppm.

A.3.2 30 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.7573	0.7917	0.9796	0.2813	0.9796
exp2	0.9044	0.9719	0.9878	0.6112	0.9878
fou1	0.9540	0.9746	0.9953	0.8048	0.9953
fou2	0.9948	0.9864	0.9984	0.9796	0.9984
fou3	0.9985	0.9986	0.9993	0.9947	0.9993
gauss1	0.7566	0.7980	0.9786	0.8230	0.9786
gauss2	0.9926	0.9746	0.9982	0.9436	0.9982
gauss3	0.9912	0.9746	0.9985	0.9862	0.9985
Particularizado	-	-	-	-	0.9970

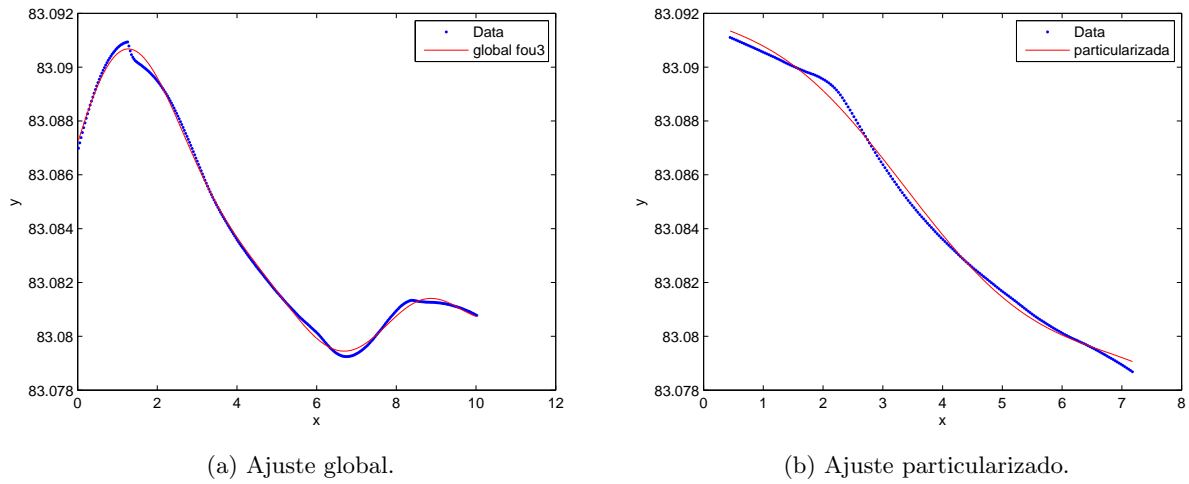


Figura A.3: Comparación ajuste global y particularizado 30 ppm.

A.3.3 35 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.7059	0.7716	0.9731	0.5379	0.9731
exp2	0.8859	0.9682	0.9842	0.8809	0.9842
fou1	0.9465	0.9713	0.9929	0.8977	0.9929
fou2	0.9932	0.9845	0.9976	0.9962	0.9976
fou3	0.9984	0.9981	0.9988	0.9989	0.9988
gauss1	0.7052	0.7792	0.9717	0.8986	0.9717
gauss2	0.9903	0.9714	0.9970	0.9840	0.9970
gauss3	0.9886	0.9714	0.9984	0.9936	0.9984
Particularizado	-	-	-	-	0.9952

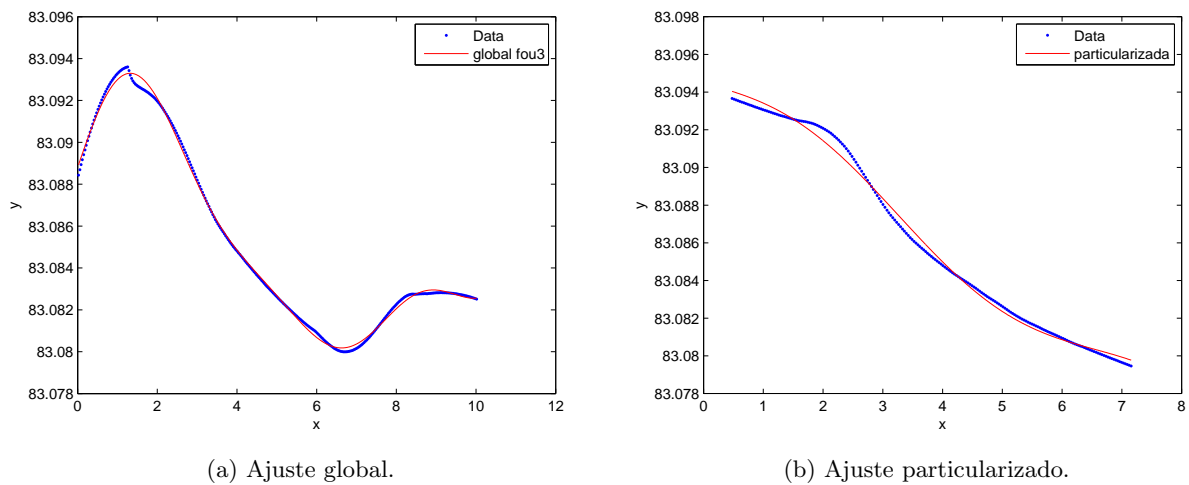
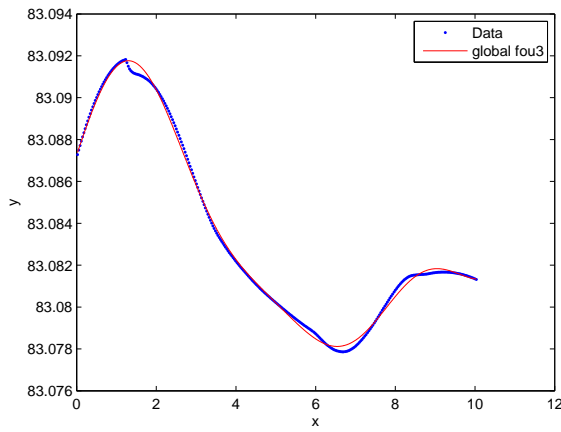


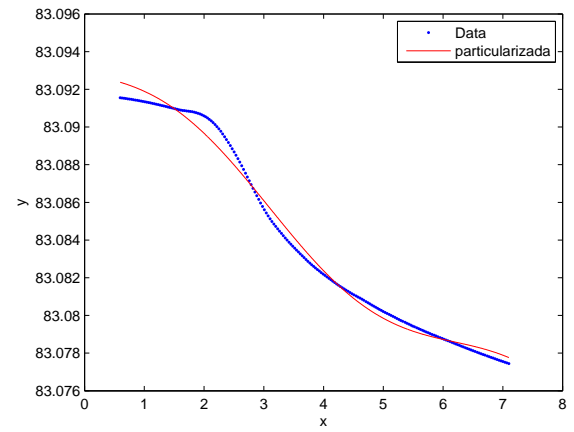
Figura A.4: Comparación ajuste global y particularizado 35 ppm.

A.3.4 40 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.6271	0.6516	0.9546	0.5979	0.9546
exp2	0.8752	0.9441	0.9758	0.8898	0.9758
fou1	0.9389	0.9489	0.9862	0.9211	0.9862
fou2	0.9913	0.9912	0.9971	0.9977	0.9971
fou3	0.9982	0.9989	0.9983	0.9993	0.9983
gauss1	0.6257	0.9493	0.9535	0.9211	0.9535
gauss2	0.9876	0.9493	0.9947	0.9931	0.9947
gauss3	0.9891	0.9493	0.9979	0.9978	0.9979
Particularizado	-	-	-	-	0.9916



(a) Ajuste global.

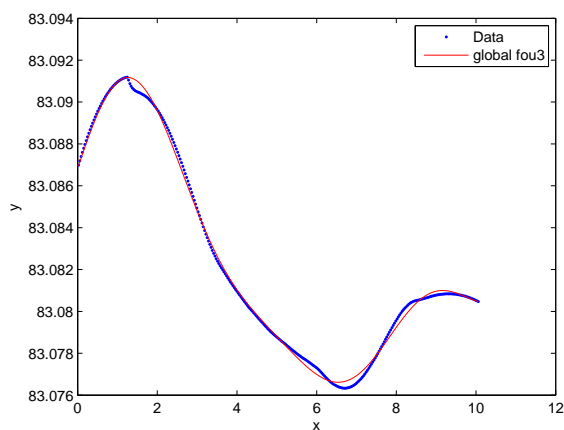


(b) Ajuste particularizado.

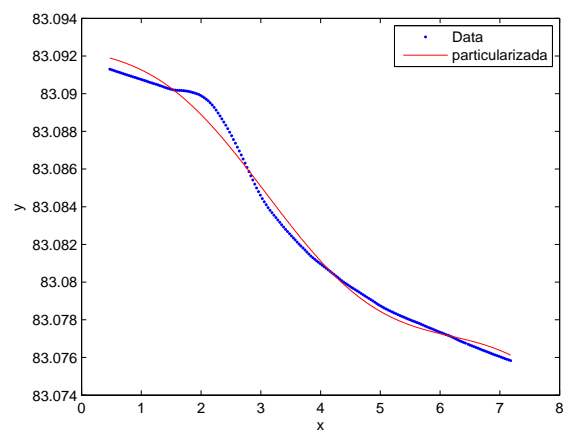
Figura A.5: Comparación ajuste global y particularizado 40 ppm.

A.3.5 45 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.5992	0.7737	0.9554	0.6567	0.9554
exp2	0.8821	0.9688	0.9763	0.9241	0.9763
fou1	0.9447	0.9730	0.9872	0.9480	0.9872
fou2	0.9916	0.9855	0.9964	0.9941	0.9964
fou3	0.9982	0.9973	0.9980	0.9993	0.9980
gauss1	0.5983	0.7812	0.9544	0.9482	0.9544
gauss2	0.9885	0.9731	0.9949	0.9895	0.9949
gauss3	0.9886	0.9731	0.9926	0.9976	0.9926
Particularizado	-	-	-	-	0.9928



(a) Ajuste global.

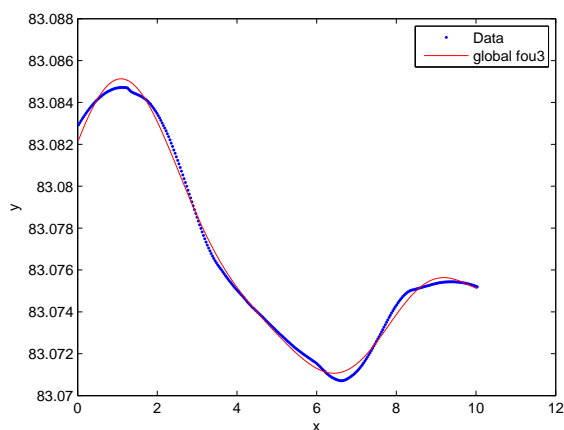


(b) Ajuste particularizado.

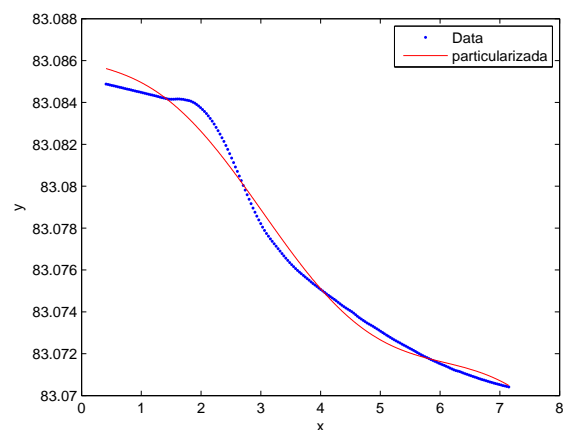
Figura A.6: Comparación ajuste global y particularizado 45 ppm.

A.3.6 50 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.5746	0.8090	0.9520	0.6791	0.9520
exp2	0.9059	0.9715	0.9737	0.9249	0.9737
fou1	0.9539	0.9736	0.9851	0.9412	0.9851
fou2	0.9902	0.9852	0.9958	0.9965	0.9958
fou3	0.9964	0.9992	0.9973	0.9993	0.9973
gauss1	0.5738	0.8180	0.9513	0.9413	0.9513
gauss2	0.9894	0.9490	0.9932	0.9909	0.9932
gauss3	0.9895	0.9725	0.9966	0.9968	0.9966
Particularizado	-	-	-	-	0.9907



(a) Ajuste global.

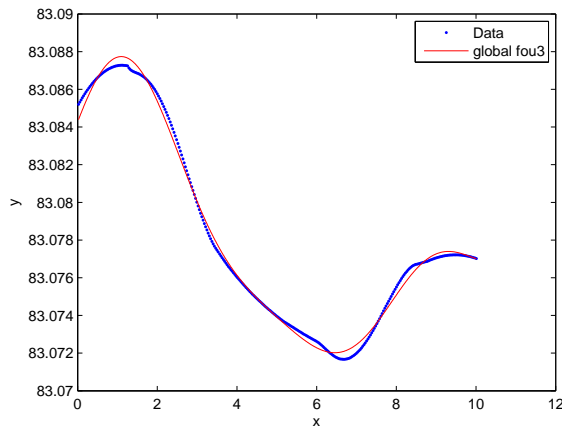


(b) Ajuste particularizado.

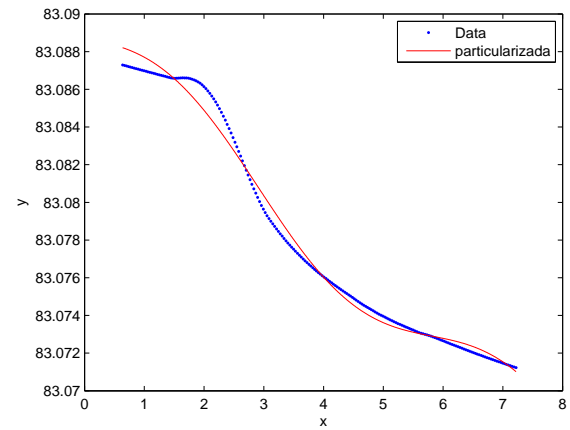
Figura A.7: Comparación ajuste global y particularizado 50 ppm.

A.3.7 55 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.5437	0.5339	0.9339	0.7273	0.9339
exp2	0.9065	0.8508	0.9733	0.9340	0.9733
fou1	0.9546	0.8683	0.9810	0.9449	0.9810
fou2	0.9893	0.9995	0.9967	0.9933	0.9967
fou3	0.9966	0.9998	0.9971	0.9994	0.9971
gauss1	0.5431	0.5763	0.9331	0.9450	0.9331
gauss2	0.9870	0.8691	0.9850	0.9874	0.9850
gauss3	0.9875	0.8691	0.9970	0.9990	0.9970
Particularizado	-	-	-	-	0.9911



(a) Ajuste global.



(b) Ajuste particularizado.

Figura A.8: Comparación ajuste global y particularizado 55 ppm.

A.3.8 60 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.4857	0.7359	0.9373	0.7640	0.9373
exp2	0.8843	0.9091	0.9684	0.9624	0.9684
fou1	0.9478	0.9103	0.9807	0.9733	0.9807
fou2	0.9882	0.9508	0.9963	0.9993	0.9963
fou3	0.9973	0.9899	0.9972	0.9995	0.9972
gauss1	0.4771	0.7564	0.9362	0.9734	0.9362
gauss2	0.9859	0.9101	0.9930	0.9974	0.9930
gauss3	0.9866	0.9102	0.9971	0.9974	0.9971
Particularizado	-	-	-	-	0.9891

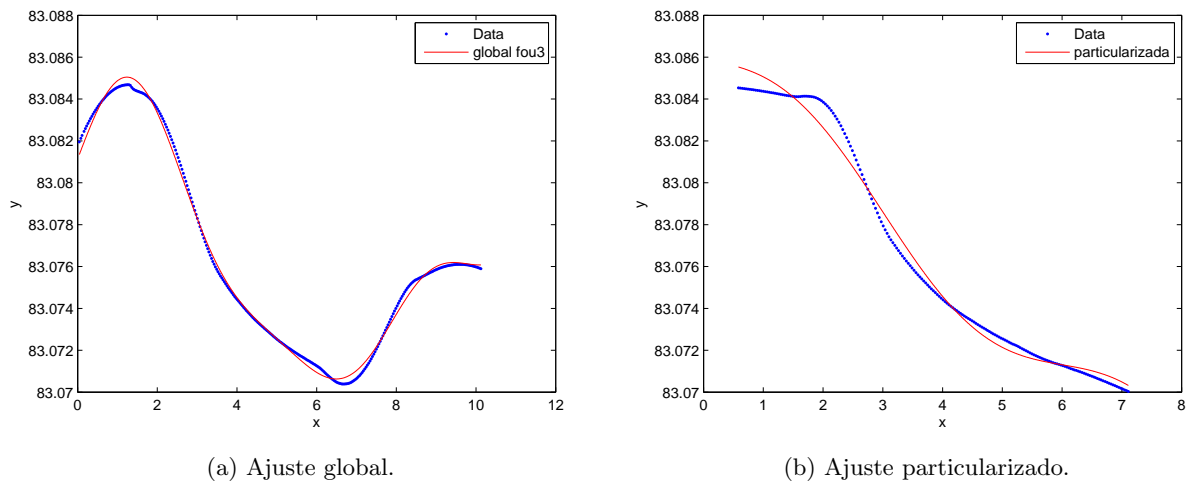


Figura A.9: Comparación ajuste global y particularizado 60 ppm.

A.3.9 65 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.4952	0.8090	0.9301	0.7807	0.9301
exp2	0.8802	0.9725	0.9653	0.9647	0.9653
fou1	0.9439	0.9761	0.9792	0.9737	0.9792
fou2	0.9878	0.9868	0.9952	0.9990	0.9952
fou3	0.9978	0.9982	0.9968	0.9997	0.9968
gauss1	0.4942	0.8135	0.9292	0.9739	0.9292
gauss2	0.9857	0.9761	0.9917	0.9969	0.9917
gauss3	0.9895	0.9761	0.9949	0.9983	0.9949
Particularizado	-	-	-	-	0.9893

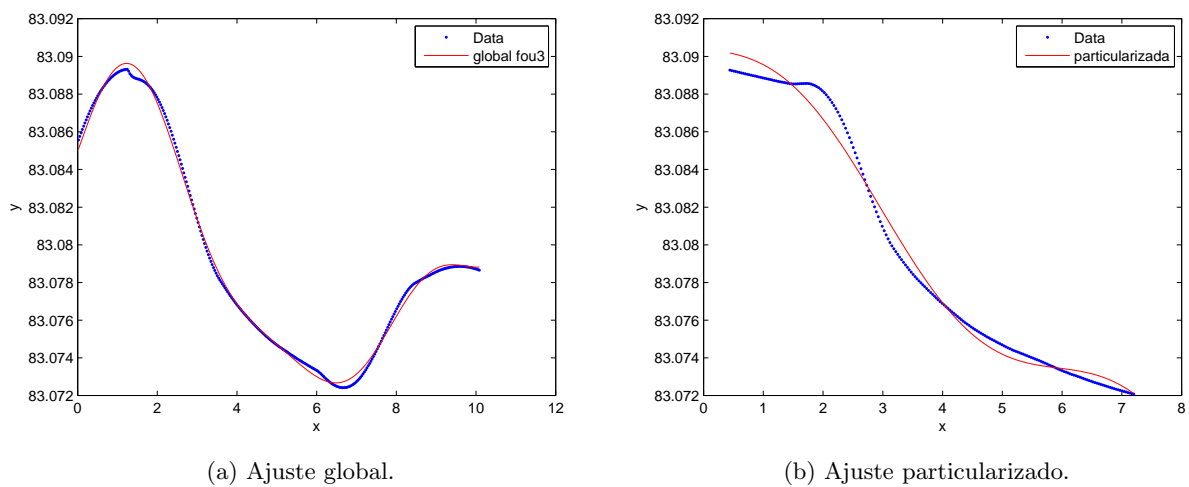
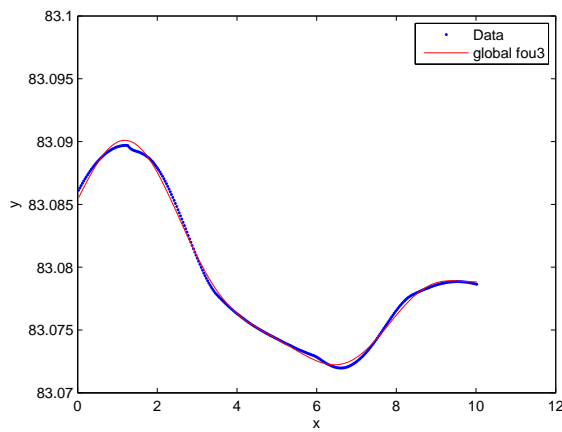


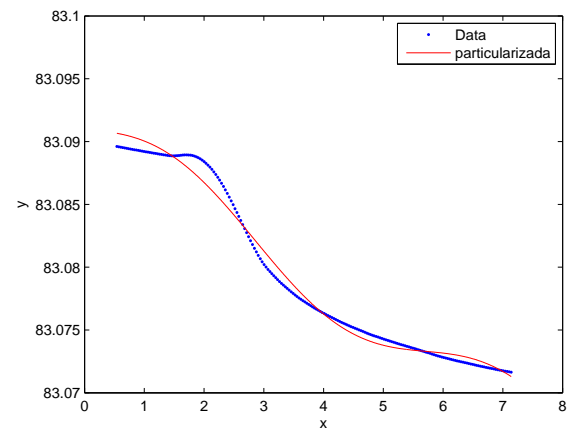
Figura A.10: Comparación ajuste global y particularizado 65 ppm.

A.3.10 70 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.4815	0.6870	0.9217	0.8035	0.9217
exp2	0.8863	0.9466	0.9654	0.9691	0.9654
fou1	0.9439	0.9559	0.9758	0.9820	0.9758
fou2	0.9859	0.9933	0.9952	0.9987	0.9952
fou3	0.9979	0.9985	0.9964	0.9996	0.9964
gauss1	0.4810	0.7031	0.9208	0.9822	0.9208
gauss2	0.9836	0.9563	0.9902	0.9974	0.9902
gauss3	0.9895	0.9563	0.9802	0.9985	0.9802
Particularizado	-	-	-	-	0.9882



(a) Ajuste global.

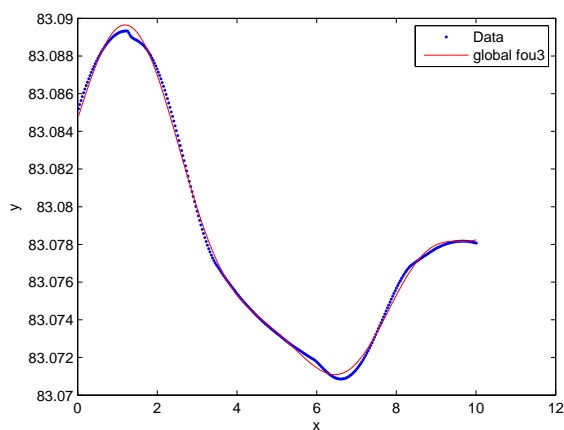


(b) Ajuste particularizado.

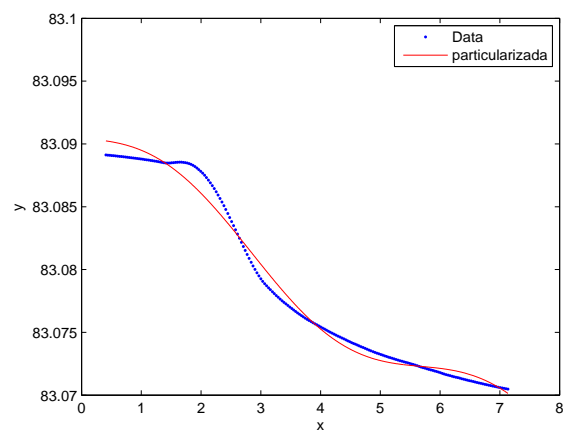
Figura A.11: Comparación ajuste global y particularizado 70 ppm.

A.3.11 75 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.4734	0.8382	0.9250	0.8230	0.9250
exp2	0.8852	0.9739	0.9638	0.9663	0.9638
fou1	0.9440	0.9765	0.9760	0.9793	0.9760
fou2	0.9848	0.9866	0.9950	0.9992	0.9950
fou3	0.9985	0.9990	0.9964	0.9998	0.9964
gauss1	0.4728	0.8420	0.9240	0.9796	0.9240
gauss2	0.9821	0.8629	0.9912	0.9974	0.9912
gauss3	0.9892	0.9765	0.9829	0.9993	0.9829
Particularizado	-	-	-	-	0.9885



(a) Ajuste global.

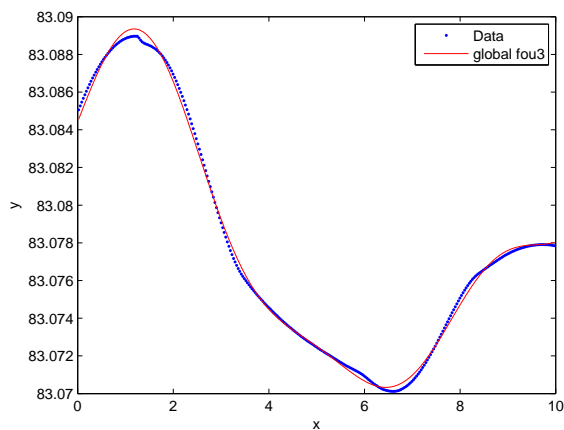


(b) Ajuste particularizado.

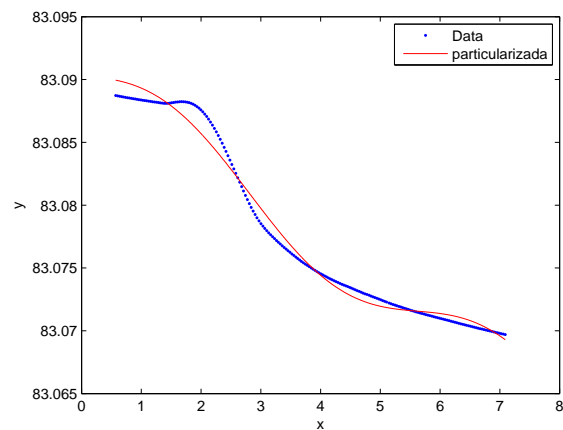
Figura A.12: Comparación ajuste global y particularizado 75 ppm.

A.3.12 80 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.4595	0.6768	0.9166	0.8569	0.9166
exp2	0.8879	0.9474	0.9638	0.9822	0.9638
fou1	0.9468	0.9570	0.9736	0.9869	0.9736
fou2	0.9846	0.9886	0.9951	0.9993	0.9951
fou3	0.9984	0.9985	0.9960	0.9998	0.9960
gauss1	0.4590	0.7015	0.9155	0.9869	0.9155
gauss2	0.9823	0.9573	0.9899	0.9985	0.9899
gauss3	0.9888	0.9573	0.9729	0.9992	0.9729
Particularizado	-	-	-	-	0.9872



(a) Ajuste global.

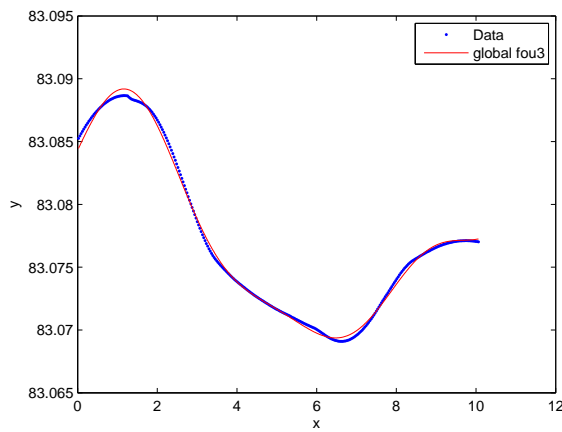


(b) Ajuste particularizado.

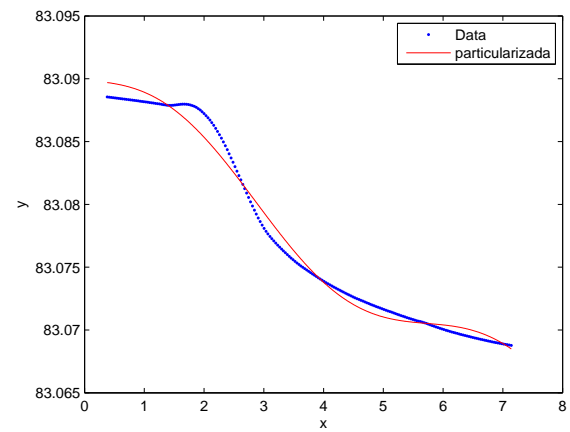
Figura A.13: Comparación ajuste global y particularizado 80 ppm.

A.3.13 85 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.4656	0.8465	0.9234	0.8481	0.9234
exp2	0.8937	0.9908	0.9616	0.9802	0.9616
fou1	0.9487	0.9924	0.9755	0.9878	0.9755
fou2	0.9855	0.9966	0.9945	0.9991	0.9945
fou3	0.9978	0.9995	0.9964	0.9998	0.9964
gauss1	0.4508	0.8503	0.9228	0.9880	0.9228
gauss2	0.9831	0.8707	0.9902	0.9973	0.9902
gauss3	0.9890	0.9924	0.9926	0.9977	0.9926
Particularizado	-	-	-	-	0.9879



(a) Ajuste global.



(b) Ajuste particularizado.

Figura A.14: Comparación ajuste global y particularizado 85 ppm.

A.3.14 90 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.4603	0.8144	0.9197	0.8705	0.9197
exp2	0.8932	0.9739	0.9607	0.9844	0.9607
fou1	0.9487	0.9752	0.9729	0.9891	0.9729
fou2	0.9835	0.9879	0.9941	0.9992	0.9941
fou3	0.9984	0.9991	0.9955	0.9995	0.9955
gauss1	0.4448	0.8189	0.9191	0.9891	0.9191
gauss2	0.9813	0.8440	0.9892	0.9986	0.9892
gauss3	0.9895	0.9752	0.9839	0.9991	0.9839
Particularizado	-	-	-	-	0.9868

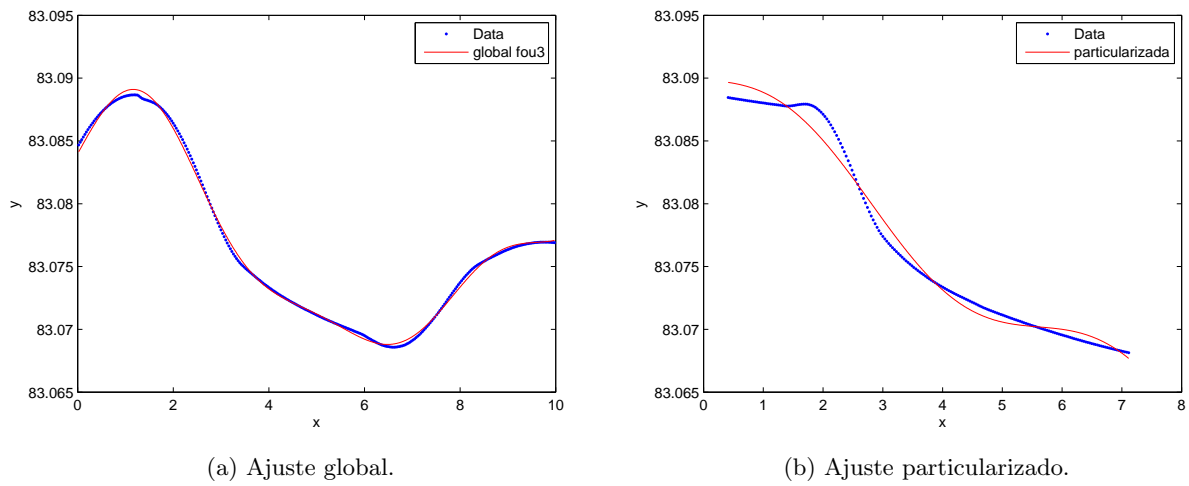


Figura A.15: Comparación ajuste global y particularizado 90 ppm.

A.3.15 95 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.4667	0.6039	0.9071	0.8698	0.9071
exp2	0.8861	0.9157	0.9630	0.9900	0.9630
fou1	0.9437	0.9276	0.9714	0.9911	0.9714
fou2	0.9836	0.9894	0.9950	0.9986	0.9950
fou3	0.9986	0.9981	0.9957	0.9998	0.9957
gauss1	0.4664	0.9283	0.9066	0.9911	0.9066
gauss2	0.9814	0.9283	0.9886	0.9977	0.9886
gauss3	0.9912	0.9283	0.9954	0.9992	0.9954
Particularizado	-	-	-	-	0.9866

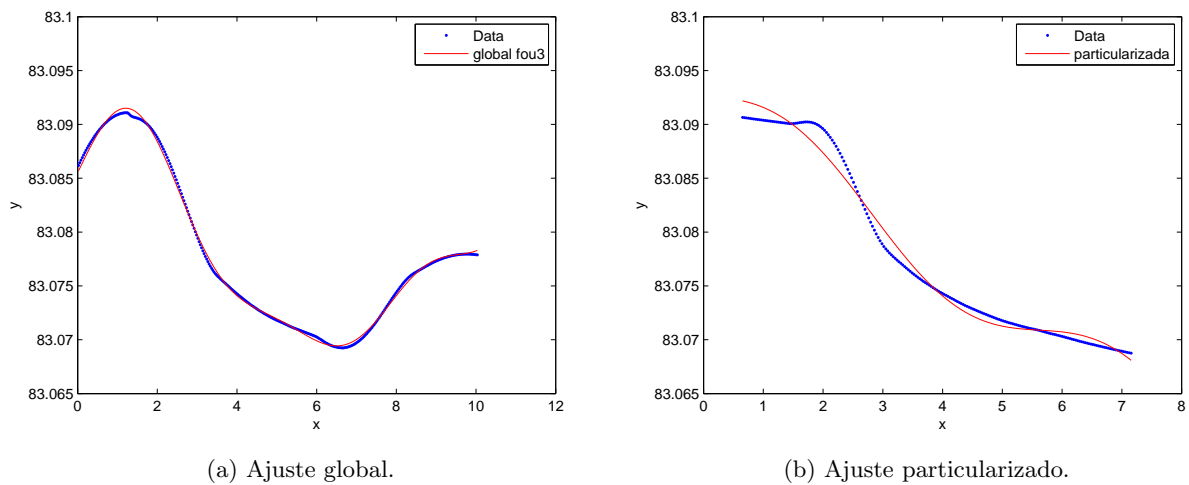
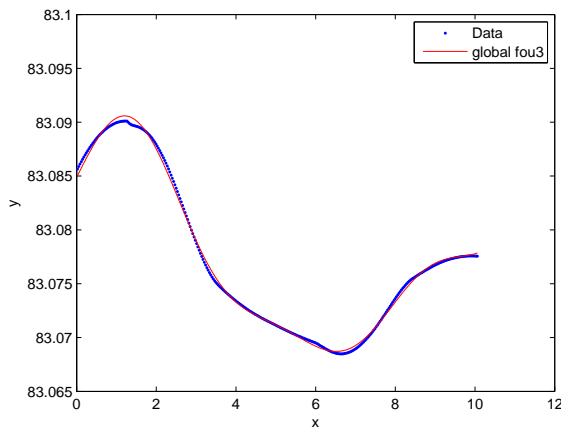


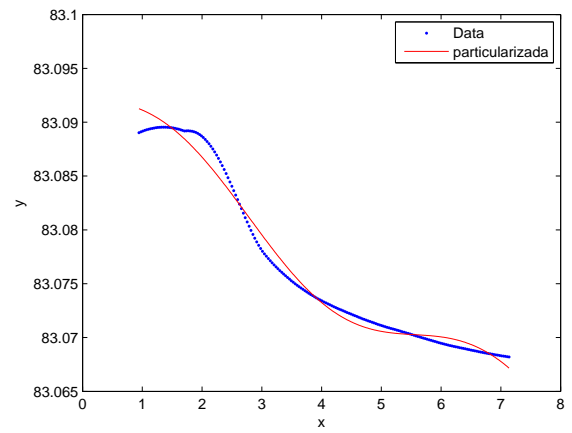
Figura A.16: Comparación ajuste global y particularizado 95 ppm.

A.3.16 100 ppm

Método	Global	Trozo 1	Trozo 2	Trozo 3	Central
exp1	0.4524	0.4453	0.8942	0.8783	0.8942
exp2	0.8884	0.8069	0.9677	0.9902	0.9677
fou1	0.9465	0.8133	0.9714	0.9904	0.9714
fou2	0.9834	0.9932	0.9967	0.9992	0.9967
fou3	0.9984	0.9967	0.9980	0.9999	0.9980
gauss1	0.4367	0.8155	0.8782	0.9905	0.8782
gauss2	0.9812	0.8155	0.9870	0.9983	0.9870
gauss3	0.9939	0.8155	0.9755	0.9991	0.9755
Particularizado	-	-	-	-	0.9864



(a) Ajuste global.



(b) Ajuste particularizado.

Figura A.17: Comparación ajuste global y particularizado 100 ppm.

A.4 Codificación particularizada de cada una de las curvas

El trozo central es el realmente importante en el experimento (ver apartado 3.5 y Figura 3.5). Luego, podemos codificar cada curva caracterizando el trozo central. Para ello se utiliza la función particularizada. Así, se puede formar la siguiente correspondencia:

25 \leftrightarrow (83.09, $1.871 \cdot 10^{-5}$, 0.000978, 0.8789)	65 \leftrightarrow (83.09, $3.435 \cdot 10^{-5}$, 0.002309, 1.059)
30 \leftrightarrow (83.09, $2.224 \cdot 10^{-5}$, 0.001054, 0.9416)	70 \leftrightarrow (83.09, $3.728 \cdot 10^{-5}$, 0.002422, 1.1)
35 \leftrightarrow (83.09, $2.599 \cdot 10^{-5}$, 0.001334, 0.9594)	75 \leftrightarrow (83.09, $3.848 \cdot 10^{-5}$, 0.002496, 1.105)
40 \leftrightarrow (83.09, $2.773 \cdot 10^{-5}$, 0.001585, 1.015)	80 \leftrightarrow (83.09, $4.01 \cdot 10^{-5}$, 0.002616, 1.117)
45 \leftrightarrow (83.09, $2.948 \cdot 10^{-5}$, 0.00169, 1.016)	85 \leftrightarrow (83.09, $4.063 \cdot 10^{-5}$, 0.002719, 1.095)
50 \leftrightarrow (83.09, $2.842 \cdot 10^{-5}$, 0.001614, 1.034)	90 \leftrightarrow (83.09, $4.218 \cdot 10^{-5}$, 0.002772, 1.119)
55 \leftrightarrow (83.09, $3.273 \cdot 10^{-5}$, 0.001942, 1.081)	95 \leftrightarrow (83.09, $4.647 \cdot 10^{-5}$, 0.003159, 1.117)
60 \leftrightarrow (83.09, $2.923 \cdot 10^{-5}$, 0.001846, 1.046)	100 \leftrightarrow (83.09, $4.753 \cdot 10^{-5}$, 0.003214, 1.131)

