



UNIVERSIDADE DA CORUÑA
DEPARTAMENTO DE ENXEÑERÍA
INDUSTRIAL II

NUEVAS APLICACIONES DE LA
ESTIMACIÓN PARAMÉTRICA Y NO
PARAMÉTRICA DE CURVAS AL ANÁLISIS
TÉRMICO

Tesis doctoral

Salvador Naya Fernández

Abril 2003



UNIVERSIDADE DA CORUÑA
DEPARTAMENTO DE ENXEÑERÍA
INDUSTRIAL II

3

NUEVAS APLICACIONES DE LA
ESTIMACIÓN PARAMÉTRICA Y NO
PARAMÉTRICA DE CURVAS AL ANÁLISIS
TÉRMICO

Tesis doctoral

Salvador Naya Fernández

Abril 2003

Índice General

Agradecimientos	i
Prólogo	iii
1 El Análisis Térmico	1
1.1 Introducción	1
1.2 Breve reseña histórica	2
1.3 Técnicas de Análisis Térmico	5
1.3.1 Análisis Térmico Diferencial (DTA)	6
1.3.2 Calorimetría Diferencial de Barrido (DSC)	6
1.3.3 Análisis Termogravimétrico (TGA)	7
1.3.4 Análisis Térmico Simultáneo (STA)	7
1.3.5 Análisis Mecanodinámico (DMA) y Termomecánico-dinámico (DMTA)	8
1.3.6 Técnicas complementarias	9
1.4 Campos de aplicación del Análisis Térmico	10
1.4.1 Cerámicos y vidrios	10
1.4.2 Metales y aleaciones	11
1.4.3 Polímeros	12
1.4.4 Materiales heterogéneos de base polimérica (mezclas, maderas y otros productos naturales)	13
1.4.5 Ejemplo de aplicación de varias técnicas de Análisis Térmico	14
1.5 Comportamientos en Análisis Térmico	15
1.5.1 Evaporación del Anhídrido Hexahidroftálico	15
1.5.2 Descomposición térmica de minerales	16
1.5.3 Comportamiento de una superaleación a altas temperaturas	16
1.5.4 Fusión y cristalización de un polímero cristalino	16
1.5.5 Pirólisis de un polímero termoestable	16
1.5.6 Degradación de un termoestable en aire	17
1.5.7 Determinación de la temperatura de transición vítrea (T _g)	17
1.5.8 Dos transiciones vítreas en una sola muestra	17
1.5.9 Comparación de maderas	17
1.5.10 Comparación de residuos forestales	18
1.6 Gráficas de distintos comportamientos térmicos	18

2	Modelos de regresión	25
2.1	Introducción al suavizado de datos en Análisis Térmico	25
2.2	Introducción a los modelos de regresión	26
2.3	Modelos de diseño aleatorio y de diseño fijo	27
2.4	Modelos de regresión según el tipo de variables	28
2.4.1	Modelos con variables reales	28
2.4.2	Modelos con variables vectoriales y funcionales	28
2.5	Modelos paramétricos	29
2.5.1	Modelos paramétricos lineales	29
2.5.2	Modelos paramétricos no lineales	31
2.6	Modelos no paramétricos	31
2.7	Métodos de suavización tipo núcleo	33
2.7.1	Otros estimadores no paramétricos de la regresión	35
2.8	Estimador polinómico local (RPL)	38
2.8.1	Ventajas del estimador de regresión polinómico local	40
2.8.2	Desarrollo matricial del modelo	41
2.9	Estimación con datos dependientes	42
3	Estimación paramétrica de curvas TGA	45
3.1	Estudio cinético	45
3.2	Modelos cinéticos tipo Arrhenius	46
3.2.1	Cálculo de los parámetros de la ecuación de Arrhenius	47
3.2.2	Cálculo del orden n de la reacción	48
3.2.3	Método de Friedman	49
3.2.4	Otros modelos propuestos para la cinética	50
3.3	Métodos numéricos en análisis cinético	51
3.3.1	Inconvenientes de los métodos actuales	52
3.4	Modelo paramétrico con mezcla de logísticas	53
3.4.1	Estimación de los parámetros del modelo	56
3.5	Aplicaciones del método de ajuste logístico	57
3.5.1	Ajuste de un modelo logístico para el anhídrido hexa- hidroftálico	57
3.5.2	Ajuste para una muestra de madera de ciprés	57
3.5.3	Ajuste para una muestra de madera de eucalipto	60
3.5.4	Ajuste para una muestra de PVC	60
3.6	Interpretación del modelo mezcla de logísticas	61
3.6.1	Ajuste del modelo con 2 componentes	62
3.6.2	Ajuste del modelo con 3 componentes	63
3.7	Ventajas del método propuesto	63
4	Estimación no paramétrica de curvas	65
4.1	Introducción	65
4.1.1	Inconvenientes del modelo logístico	65
4.1.2	Inconvenientes del software actual	66
4.2	Estimación no paramétrica de curvas DSC y TGA	67
4.2.1	Estimador polinómico local para curvas TGA y DSC	68

4.2.2	Elección de la función núcleo y del grado del polinomio	69
4.3	Estimación de la ventana óptima	70
4.3.1	Tipos de error	70
4.3.2	Métodos para la estimación de la ventana óptima	71
4.4	Metódo plug-in en dos etapas bajo dependencia	73
4.4.1	Estimación de la suma de autocovarianzas	73
4.4.2	Selección de las ventanas piloto	76
4.5	Aspectos Computacionales	78
4.6	Elección de ventana global o local	79
4.7	Aplicación a un caso práctico	81
4.8	Estudio de simulación	82
4.8.1	Generación de datos	84
4.8.2	Estimación en los extremos (efecto frontera)	84
4.8.3	Resultados de las simulaciones	86
4.8.4	Algunos resultados gráficos	88
5	Clasificación de materiales	99
5.1	Introducción	99
5.2	Clasificación de datos	99
5.2.1	Reglas de clasificación	101
5.2.2	Regla de máxima verosimilitud	101
5.2.3	Regla Bayes	102
5.3	Modelo no paramétrico de regresión funcional	102
5.3.1	Selección de la ventana de suavizado	103
5.3.2	Elección de la seminorma	104
5.4	Programa para clasificar materiales	105
5.4.1	Cálculo de las distancias entre las curvas	105
5.4.2	Determinación de la ventana de validación cruzada	105
5.4.3	Clasificación de un nuevo material	106
5.5	Aplicación a muestras de maderas	106
5.5.1	Características específicas de las muestras a analizar	106
5.6	Aplicación a muestras de PVC	109
5.6.1	Características específicas de las muestras a analizar	109
5.7	Estudios de simulación	110
5.7.1	Pasos para las simulaciones	111
5.7.2	Muestras simuladas	111
5.7.3	Resultados de las simulaciones	118
6	Conclusiones y líneas futuras	121
6.1	Conclusiones y objetivos alcanzados	121
6.1.1	Método paramétrico de mezcla de funciones logísticas	121
6.1.2	Suavizado automático de curvas de Análisis Térmico	122
6.1.3	Método de clasificación de materiales	123
6.2	Líneas futuras	124
6.2.1	Extensión del modelo de mezcla de logísticas	124
6.2.2	Suavización de espectros tridimensionales	125

6.2.3	Ampliación del método de clasificación de materiales . . .	126
6.2.4	Algunas salidas gráficas	128
Anexos		131
Bibliografía		160

Agradecimientos

Deseo hacer mención, en unas breves líneas, de mi agradecimiento hacia algunas personas e instituciones que con su ayuda me han permitido llevar a buen término esta monografía.

En primer lugar, expreso mi más profundo agradecimiento a los directores de esta tesis, los profesores Ricardo Cao Abad y Ramón Artiaga Díaz, por sus aportaciones a la misma, así como por sus enseñanzas y constantes estímulos. Tan importante como la ayuda científica que de ellos he recibido, ha sido la amistad que durante estos años me han brindado y que sinceramente agradezco.

Reconozco toda la ayuda prestada por mis compañeros de la Universidade da Coruña, tanto los del área de Estadística e Investigación Operativa del Departamento de Matemáticas, como los del área de Ciencia de los Materiales del Departamento de Ingeniería Industrial II.

Quiero agradecer expresamente la colaboración prestada por el profesor Mario Francisco Fernández, quien gentilmente me cedió sus trabajos y programas de ordenador para iniciar esta investigación. Asimismo, deseo agradecer la amable colaboración del profesor Iñaki López de Ullibari, por su ayuda con la utilización del Software S-plus y la del profesor Alejandro Quintela, por haberme iniciado en el campo de la estimación no paramétrica de curvas.

Del mismo modo debo agradecer los consejos y observaciones recibidos del profesor José Luis Mier, que han sido de gran importancia para enriquecer el resultado final de esta memoria.

Es justo reconocer la ayuda brindada por Javier Trillo Enríquez y Carlos Dominguez Salgueiro -antiguos alumnos de la Facultad de Informática de la Universidade da Coruña- en la elaboración de los programas informáticos con los que se han realizado varios ajustes y simulaciones empleados en esta memoria, y a los que tuve el placer de codirigir en sus proyectos de fin de carrera de la titulación de Ingeniería Informática.

Por último, quiero hacer constar que este trabajo se ha beneficiado de los proyectos de investigación PGIDT99PXI16602A de la Xunta de Galicia y del BFM2002-00265 del Ministerio de Ciencia y Tecnología.

Prólogo

El Análisis Térmico de materiales requiere del manejo de una gran cantidad de datos, para lo que hace uso de varias técnicas estadísticas. La actual utilización de potentes ordenadores conectados a sofisticados aparatos de laboratorio permite analizar toda una serie de variables, que hace poco tiempo sería imposible. Sin embargo, este análisis de datos no siempre emplea los mejores métodos estadísticos. Así, por ejemplo, en el estudio de curvas procedentes del estudio térmico, el interés suele recaer en el ajuste de una determinada variable a lo largo del tiempo. En estos casos se requiere formular un modelo matemático adecuado que explique la posible relación funcional existente entre las variables de estudio, y que tenga en cuenta el efecto de la dependencia de los datos.

La primera finalidad de este trabajo es la de construir un modelo que formalice matemáticamente la complejidad existente en la realidad de los datos de estudios de laboratorio, concretamente el comportamiento de variables usadas en el estudio térmico de materiales. Una vez que el modelo sea formulado, y validado adecuadamente, éste podrá utilizarse para estimar y predecir curvas o características de interés. Otro fin, y no menos importante cuestión, es la programación de los modelos presentados mediante algoritmos computacionales, que permitan su incorporación a otros programas de software existentes y que, en la actualidad, carecen de este tipo de herramientas en este contexto.

Para este primer objetivo emplearemos diferentes métodos del *Análisis de Regresión*, que es la parte de la estadística que se encarga de tratar problemas como los comentados anteriormente. Concretamente, en esta memoria abordaremos el problema de la implementación de estimadores paramétricos y no paramétricos de tipo polinómico local bajo condiciones de dependencia. Además, estudiaremos el caso del ajuste, no sólo de la función de regresión, sino también de determinadas funciones asociadas a ella como son la primera y segunda derivadas, que aportan información adicional al estudio térmico de materiales.

Un último objetivo de esta memoria es la clasificación de materiales a partir de curvas procedentes de su análisis térmico. Concretamente, presentaremos un método de tipo no paramétrico funcional para la discriminación entre distintas muestras. El empleo de este método permitirá la comparación de espectros para clasificar adecuadamente una nueva muestra dentro de una gama de materiales conocida. La novedad del modelo de clasificación radica en el uso de toda la información de la muestra, mediante el empleo de modelos de regresión

denominados de dimensión infinita o funcionales, por cuanto usan curvas como variables regresoras.

La memoria consta de seis capítulos. En el primero de ellos presentamos el Análisis Térmico de materiales mediante una breve introducción histórica, donde hacemos hincapié en la importancia del estudio térmico en la aparición de nuevos materiales. Introducimos las distintas técnicas empleadas en este campo y se enfatiza la importancia que tiene el correcto ajuste de las curvas mostrando una serie de ejemplos de experimentos analizados en función de la información que muestran las correspondientes gráficas.

En el segundo capítulo analizamos los diferentes modelos de regresión haciendo un repaso a las distintas técnicas de inferencia funcional. Destacamos la importancia que tiene la adecuada elección del parámetro de suavizado en el caso de la regresión no paramétrica, y desarrollamos con mayor detalle el modelo no paramétrico de tipo polinómico local, que será el elegido para el ajuste de nuestros datos.

El capítulo 3 introduce la cinética de la descomposición térmica, exponiendo los diferentes métodos cinéticos y sus problemas. A continuación planteamos un método alternativo de tipo *paramétrico logístico* que aplicaremos a distintas muestras.

En el capítulo 4 proponemos un método de tipo no paramétrico para el suavizado de curvas, aplicado a la mayoría de los espectros obtenidos en Análisis Térmico. Este nuevo método, con selector del parámetro de suavizado de tipo *plug-in en dos etapas*, permite realizar un ajuste automático de los datos para lo que tiene en cuenta la expresión del parámetro óptimo de suavizado.

En el capítulo 5 exponemos un método para la clasificación de materiales basado en el *análisis no paramétrico discriminante*. Como ejemplo lo aplicaremos a muestras de madera y de PVC que, debido a su naturaleza, presentan una dificultad añadida de clasificación.

En el capítulo 6 recogemos las conclusiones de mayor relevancia a las que hemos llegado con nuestro estudio, proponiendo nuevas líneas de investigación: por una parte aplicaciones de los modelos propuestos, y por otra, ampliación de los modelos de suavización de curvas y de clasificación de materiales mediante el empleo de espectros tridimensionales de infrarrojos. Finalmente, se completa la memoria con un anexo final que contiene algunos de los cálculos empleados en el desarrollo de este trabajo y los esquemas de los programas informáticos desarrollados.

Capítulo 1

El Análisis Térmico

1.1 Introducción

El Análisis Térmico se ocupa del estudio de las propiedades físicas y termodinámicas de los materiales en función de la temperatura o del tiempo. La ICTAC (Confederación Internacional para el Análisis Térmico y la Calorimetría) lo define como: “conjunto de técnicas mediante las cuales, una propiedad física o química de un material es medida en función de la temperatura o del tiempo”.

La Calorimetría, que es una de las técnicas de Análisis Térmico con mayor tradición, se interesa por la medición cuantitativa de un intercambio de calor, es decir, estudia la energía que se transfiere de un cuerpo a otro debida a la diferencia de temperatura.

El Análisis Térmico forma parte de la Ciencia de los Materiales, entendiendo ésta como la parte del saber que trata del estudio de la composición, estructura y procesado de los materiales con objeto de aplicarlos para distintos fines tecnológicos. Así, el COSMAT (Committee on the Survey of Material Science and Engineering of the United States), define la Ciencia de los Materiales como la parte del conocimiento necesaria para el progreso en los materiales, y en particular, para resolver problemas tecnológicos complejos que, a menudo, no coinciden con los objetivos de las disciplinas tradicionales.

La Ciencia de los Materiales se caracteriza por ser interdisciplinar, abarcando no sólo campos tradicionales como la Química o la Física, sino también la mayoría de las ramas de la ingeniería, que están involucradas en su desarrollo.

Desde los primeros calorímetros, hasta los sofisticados aparatos de laboratorio asociados a potentes programas de ordenador de hoy día, se han producido toda una serie de avances científicos en este campo, propiciados, en gran medida, por la revolución informática. Para acercarnos al estudio actual intentaremos, mediante un recorrido histórico, resaltar algunos de los avances que se han ido produciendo en el campo de la Ciencia de los Materiales, vinculándolos a sus propiedades térmicas.

1.2 Breve reseña histórica

Uno de los grandes avances de la humanidad fue el descubrimiento de que se podía alterar la naturaleza de los materiales mediante la aplicación de calor. Los cerámicos fueron los primeros materiales inorgánicos en ser alterados por el ser humano.

La siguiente destreza (tecnología, diríamos ahora) para manejar materiales inorgánicos fue probablemente la selección y molienda de colores minerales para pigmentos, que fueron utilizados con carácter decorativo o ceremonial. En el antiguo Egipto se utilizaban las propiedades térmicas para la fabricación de pigmentos, como el clásico azul egipcio. Este compuesto, cuya fórmula es $\text{CaCu}(\text{Si}_4\text{O}_{10})$, se puede sintetizar fácilmente en el laboratorio mediante una termobalanza calentando una mezcla de cuarzo, óxido cúprico (CuO), calcita (CaCO_3) y un agente fundente (Na_2CO_3 , bórax o PbO). Para su obtención es necesario controlar adecuadamente la temperatura de calentamiento, pues sin la temperatura adecuada la reacción es muy lenta, dando lugar a un producto muy impuro, finamente dividido, sin el color azul intenso de los antiguos grabados.

En China, durante el denominado período Heian (794-1192 de nuestra era), fue famosa su cerámica, asociada a la ceremonia del té. Se trataba de piezas que incorporaban metales para su ornamentación. También, en Japón, a finales del siglo XVI, aparece la cerámica rakú, una porcelana vidriada con plomo, que se cocía en un horno a bajas temperaturas. Estas piezas se fabrican a partir de la caolinita, cuya fórmula molecular es $\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$. Para su elaboración se requiere el conocimiento de sus propiedades térmicas. Hoy en día, la caolinita, sigue usándose para el mismo fin, teniendo gran importancia en la industria.

En la fabricación de morteros para la construcción de las pirámides también se tenía un amplio conocimiento de la influencia del calor en los materiales. Este hecho lo confirman autores como Wallace (1865), quien demostró que los morteros de las grandes pirámides tenían un 81.5% de yeso (CaSO_4) y un 9.5% de carbonato de calcio, lo que confirma el conocimiento de sus propiedades térmicas.

En la construcción de las catedrales medievales, normalmente, se preparaba el mortero mezclando óxido de calcio (CaO), que se obtenía calcinando la caliza, con agua y agregados. En ciertas catedrales, como la de Chartres y Bourges, en Francia, se encontraron morteros de caliza y yeso; en principio se atribuía el yeso al efecto de la contaminación atmosférica por SO_2 , sin embargo, autores como Adams (véase Adams, Kneller y Dollimore (1992)) demostraron que su existencia se debe al proceso de fabricación del mortero.

Al descubrimiento del endurecimiento de las arcillas por el calor siguió el descubrimiento y uso de los metales. Evidentemente, los primeros metales en ser usados fueron aquellos que se encuentran en su estado nativo como el cobre y el oro (cuarto milenio antes de Cristo).

La importancia del descubrimiento y tratamiento térmico de los metales queda patente en la denominación de los períodos de la prehistoria: Edad de Bronce, Edad de Hierro. Uno de los criterios más utilizados para distinguir los distintos metales es, sin lugar a dudas, la temperatura a la cual ocurren sus

transformaciones características. La Edad del Cobre precedió a la Edad del Hierro debido a que el cobre funde a una temperatura (1083 °C) inferior a la del hierro (1537°C).

La humanidad, que se enfrentó primero al problema de la fundición, no se encontró con materiales aislados que funden a temperaturas bien definidas. Por el contrario, tuvo frente a sí a todos los materiales, y sin saberlo fue testigo de hechos como estos: a 100°C, en algunos materiales aparecen capas de óxido; las piritas y el óxido de plata empiezan a descomponerse a 330°C, mientras que el estaño y el plomo puros ya se fundieron; 500°C es la temperatura a la que recristalizan y se reblandecen el cobre y el bronce; al alcanzar 600°C, los barros de alfarería se endurecen, algunos de ellos presentan un acabado vítreo y algunos vidrios ya se pueden moldear a esta temperatura.

Fue así, de manera empírica, como se fueron descubriendo materiales, y la manera de producir cambios en ellos. La importancia de muchos de estos materiales y procesos debió pasar inadvertida hasta que algún hecho repetido de manera casual o premeditadamente les dio la relevancia que ahora tienen. Tal vez, el caso más sobresaliente corresponde al zinc, que fue utilizado largo tiempo y sin saberlo en una aleación (el latón), mucho antes de que se le descubriera como metal. Digno también de atención resulta el estaño, cuyos orígenes son sumamente nebulosos. Sin embargo, su importancia en el desarrollo de la humanidad llega al punto de marcar, en una aleación con el cobre, toda una era: la Edad del Bronce.

El acero tiene su origen en el calentamiento de hierro y carbono. La fabricación del acero fue sencilla; sin embargo, el reconocimiento de que se trataba de una aleación y no un metal puro llegó mucho más tarde. Así, a diferencia del bronce en el que la mezcla de cobre y estaño se hacía de modo consciente, el acero se producía de modo accidental ya que el carbono se absorbía en el hierro sin conocimiento del metalurgista.

La presencia del carbono en el acero y en el hierro colado no fue detectada hasta 1774. Aunque la fecha en que se empezaron a controlar los tratamientos térmicos del acero es todavía incierta, son resaltables aplicaciones como las espadas hechas con acero de Damasco o las tradicionales espadas (*kiris*) construidas en Japón, durante el dominio de los guerreros samurais (período Edo: 1616-1868 de nuestra era); la calidad de estas armas demuestra el alto grado de conocimiento del tratamiento térmico de los metales.

El desarrollo de los polímeros ha estado muy ligado al estudio de sus propiedades térmicas. Antecedentes de los polímeros sintéticos de hoy día son el asta natural y los caparazones de tortuga. Desde el antiguo Egipto se venía usando el asta natural y los caparazones de tortuga para la fabricación de ornamentos y utensilios, siendo en el Medioevo donde el oficio de astero alcanzaría una gran relevancia en toda Europa.

Un polímero natural, antecedente de los plásticos sintéticos, es la *goma laca*, traída a Europa desde la India, por Marco Polo, en torno al año 1290. La goma laca se obtenía a partir de un líquido segregado por un insecto llamado *lac*, en su proceso de reproducción, en los árboles de la especie *rhus succedanea*: mediante la aplicación de calor este polímero natural podía moldearse para la fabricación

de piezas. La industria de la goma laca, que alcanzó su apogeo en la China de la dinastía Ming, se mantuvo hasta que en 1930 aparecen los polímeros sintéticos superando, con creces, sus cualidades.

Otro polímero natural, que tuvo gran repercusión en el pasado, es la *gutapercha*, material que se obtenía a partir de los árboles *Palaquium gutta* originarios de la península de Malasia. En 1843, William Montgomery informó de que en este territorio, se utilizaba gutapercha para la fabricación de mangos de cuchillos. Su informe condujo al nacimiento de la Gutta Percha Company, que se mantuvo activa hasta 1930. La gutapercha es un material que al ser calentado puede estirarse en tiras largas y, dado que era de fácil moldeo, se utilizó para la fabricación de diversas piezas. Su inercia ante el ataque de productos químicos la convirtió en un excelente aislante para cables y conductores eléctricos. Así, la primera conexión por cable telegráfico bajo el agua, que se hizo en el Canal de la Mancha, debe su éxito al aislamiento con gutapercha. Asimismo, la gutapercha fue empleada para proteger el primer cable transatlántico que se tendió en 1866.

En 1862, Alexander Parkes, un hombre de negocios inglés, con gran experiencia en la manipulación de polímeros naturales, como la gutapercha o el caucho, anunció el descubrimiento de un nuevo material al que denominó *parkesina*: obtuvo un material sólido que, al calentarse, se volvía maleable. Aseguraba que se trataba de un material que compartía las propiedades del marfil, la concha de tortuga, el asta natural, el caucho, la gutapercha, etc., y que habría de sustituir a dichos materiales. Fundó una empresa para comercializar dicho material pero sus expectativas no se cumplieron ya que este nuevo plástico se contraía en exceso.

Otro intento por descubrir materiales con mejores condiciones que los naturales lo dio Jhon Hyatt, quien en 1870 patentó el *celuloide*, que se obtenía mediante la mezcla de piroxilina en polvo con goma de alcanfor pulverizada. El celuloide resultó ser un buen producto para el moldeo, fabricándose distintos artículos como peines, espejos, cepillos, etc. Sin embargo, su nombre quedó asociado a la fabricación de películas cinematográficas. Un inconveniente de este material era su inflamabilidad, lo que produjo varios accidentes en salas de proyección.

Un cambio sustancial en el desarrollo de los materiales lo supuso la modificación de materias naturales. Un ejemplo claro es el procesado del caucho. Este compuesto se obtiene del látex natural que se encuentra en la savia de algunos árboles, como el *hevea brasiliensis*. El caucho es muy sensible a la temperatura: en condiciones climatológicas cálidas se ablanda y en ambiente frío se vuelve rígido. En 1939, Charles Goodyear descubrió que al calentar la mezcla de azufre en polvo con caucho, el compuesto resultante era más resistente y duro, menos sensible a la temperatura y más elástico. Apareció así la vulcanización de la goma de caucho, que se convertiría en un importante polímero industrial.

En 1907, el químico Leo Baekeland, que dedicó su trabajo a la búsqueda de un sustituto para la goma laca, descubrió un material plástico que denominó *baquelita*. Este material lo obtuvo a partir del fenol y del formaldehído, productos que provenían de la industria química, lo que marcó una diferencia con

otros plásticos naturales modificados.

Con la baquelita comenzó la nueva era de los plásticos, demostrando que era posible reproducir en laboratorio lo que los insectos lac o los árboles del caucho hacían en la naturaleza. La baquelita dejó abierta la puerta para el desarrollo de los polímeros sintéticos, muchos de ellos adaptados para satisfacer requisitos específicos. De hecho, Leo Baekeland previó el potencial sin límites de los plásticos fenólicos utilizando el símbolo infinito (∞) para representar sus usos. Así, a mediados del pasado siglo fueron apareciendo otros polímeros de gran uso en la industria, como el caso del policloruro de vinilo (1927), el poliestireno (1938), el poliéster (1942), el polietileno (1943), las resinas epoxídicas (1947) o el polipropileno (1957). Las propiedades y aplicaciones de estos materiales están ligadas a su comportamiento térmico.

Con respecto a la invención de los aparatos de laboratorio, cabe destacar, que los primeros estudios usando un calorímetro parten de Lavoisier y Laplace (1783). También son resaltables la aparición del primer calorímetro con bomba inventado por Berthelot (1864) y los calorímetros adiabáticos de Dickinson (1915) y de Washburn (1933). Es en la década de los 50, a partir de la aparición de la bomba rotatoria de Sunner (1949), cuando da comienzo la era moderna en esta ciencia. Como fechas destacables en esta nueva visión del Análisis Térmico, puede mencionarse que el primer equipo DTA comercial aparece en el año 1960 y el DSC en el año 1964. En 1969 se publica el primer número de la revista *Journal of Thermal Analysis* y el año siguiente aparece la revista *Thermochemical Acta*, que hoy en día siguen siendo importantes referentes de publicaciones en este campo.

1.3 Técnicas de Análisis Térmico

Los estudios actuales en el campo del Análisis Térmico se apoyan en el manejo de potentes aparatos de laboratorio. Estos instrumentos permiten obtener series de puntos que relacionan algún tipo de variable dependiente (flujo de calor, pérdida de masa, etc...) frente a la temperatura o el tiempo. Esta gran cantidad de datos requiere un adecuado tratamiento estadístico con el fin de suavizar adecuadamente las distintas gráficas y hacer una interpretación correcta de las mismas, que es uno de los objetivos de este trabajo. La aplicación de estas herramientas estadísticas depende del tipo de análisis requerido en cada experimento. En esta sección se introducirán las principales técnicas empleadas en el estudio térmico de materiales.

Las técnicas de Análisis Térmico suelen clasificarse en función del aparato que se utilice. Para ello se acostumbra a nombrarlas mediante sus siglas: DTA (Differential Thermal Analysis), DSC (Differential Scanning Calorimetry), TGA (Thermal Gravimetric Analysis), STA (Simultaneous Thermal Analysis), DMA (Dynamic Mechanical Analysis) y DMTA (Dynamic Mechanical Thermal Analysis). Es frecuente el uso de estas siglas, no sólo para denominar a la técnica en cuestión, sino también para nombrar el aparato que se utiliza. A continuación describimos las características básicas de cada una de estas técnicas.

1.3.1 Análisis Térmico Diferencial (DTA)

La técnica DTA permite medir la diferencia de temperatura entre una muestra y un material de referencia cuando ambos son sometidos a un programa térmico. Los resultados obtenidos mediante el empleo de esta técnica permiten observar procesos endotérmicos y exotérmicos, así como cambios en la capacidad calorífica de la muestra. Esto ayuda a identificar puntos de interés como la temperatura de transición vítrea (T_g), en aquellos materiales susceptibles de sufrir un proceso de cristalización, como los polímeros. También se pueden identificar procesos endotérmicos como la temperatura de fusión y procesos exotérmicos como la cristalización o el curado.

El inconveniente que presentan las técnicas DTA es que no permiten cuantificar la entalpía del proceso ni la capacidad calorífica.

1.3.2 Calorimetría Diferencial de Barrido (DSC)

La técnica DSC es similar a la DTA, con la importante diferencia de que en este caso, en lugar de medir la diferencia de temperatura entre la muestra y una referencia, se mide la diferencia de flujo de calor que experimentan ambas al ser sometidas al mismo programa de temperatura. La evolución de la muestra puede obtenerse en función de la temperatura o del tiempo, según el programa establecido.

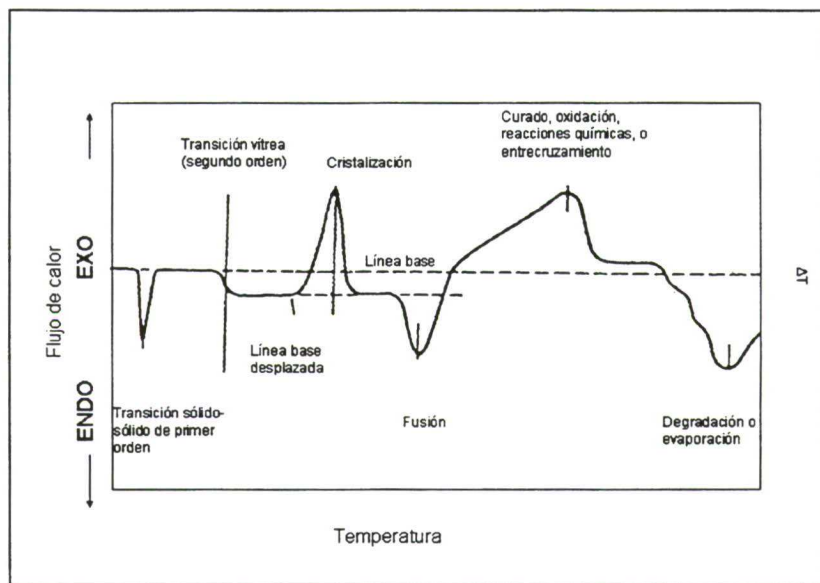


Figura 1.1: Eventos que pueden observarse en una curva DSC.

Con esta técnica se pueden identificar los mismos eventos que con la DTA y,

además, permite medir cuantitativamente tanto las entalpías como las capacidades caloríficas (véase figura 1.1).

1.3.3 Análisis Termogravimétrico (TGA)

La técnica TGA permite medir la masa de la muestra mientras está siendo sometida a un programa térmico (isotérmico o en rampa) en atmósfera controlada. Normalmente, a esta técnica se le denomina termogravimetría, llamándose termogramas a los gráficos resultantes.

Se puede utilizar un *método dinámico* si se estudia el comportamiento de la muestra mientras es sometida a un programa de temperaturas con una determinada velocidad de calentamiento. Si, por el contrario, se mantiene fija la temperatura y se estudia la variación de la masa con el tiempo, se hablará de *método isotermo*. En cualquier caso, siempre es posible reflejar el cambio de masa de la muestra como una serie temporal, ya que, aún en el caso dinámico, se conoce la evolución de la temperatura con el tiempo.

La curva que representa la velocidad de variación de masa, es la derivada de la curva TGA, y se denomina DTG (Derivative Termogravimetry). Es también de gran interés, ya que permite determinar puntos críticos, como son los instantes de velocidad máxima de pérdida de masa. Además, la velocidad de la pérdida de masa del material se utiliza para los estudios cinéticos, es decir, el análisis de la velocidad con que ocurren las transformaciones en función de la temperatura.

Una termobalanza o equipo TGA está compuesta por una microbalanza, un horno y un procesador de temperaturas; también dispone de un circuito de gas de purga (normalmente aire o N_2). Añadido a este aparato suele utilizarse un ordenador con su correspondiente programa de procesado de datos.

Las aplicaciones más típicas de esta técnica incluyen la determinación cuantitativa de componentes que se volatilizan a temperaturas características, tales como la humedad. También es un método idóneo para evaluar la estabilidad térmica de materiales.

1.3.4 Análisis Térmico Simultáneo (STA)

La técnica STA combina las dos principales técnicas de análisis térmico, DSC y TGA, simultáneamente, sobre una misma muestra. Presenta la ventaja de eliminar la variación debida a diferencias de formulación, orientación o ambiente, siendo más rápida y más eficiente en cuanto a costes.

Como puede observarse en la figura 1.2, se representan las curvas DSC y TGA frente al tiempo, obtenidas a partir de una muestra de oxalato de calcio en un aparato STA. Se ha mantenido una velocidad de calentamiento constante y se observa cómo cada uno de los escalones que presenta la gráfica TGA, corresponde a diferentes cambios experimentados por la muestra. En el primer escalón se aprecia la pérdida de agua, en el segundo escalón se pierde CO , en el tercero CO_2 y, al finalizar, se obtiene óxido de calcio.

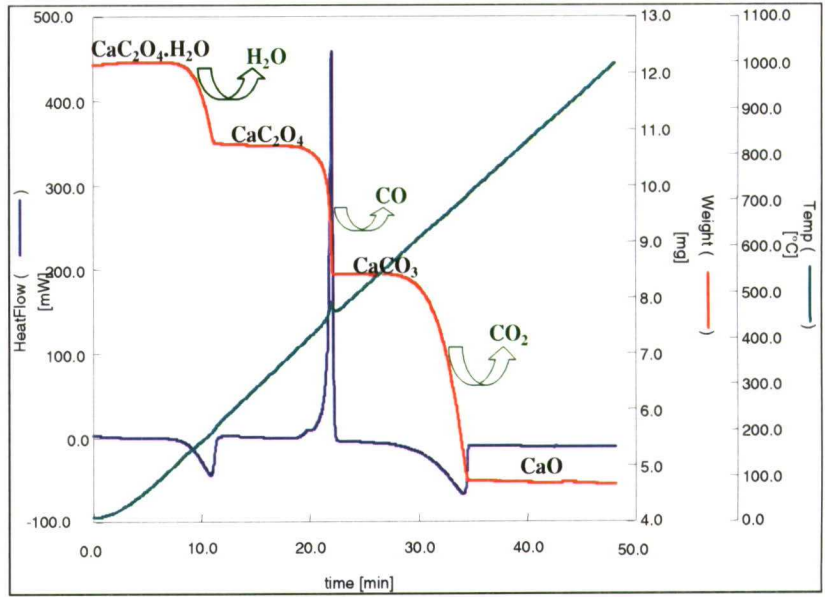


Figura 1.2: Curva TGA (en rojo) y DSC (en azul) del oxalato de calcio.

1.3.5 Análisis Mecanodinámico (DMA) y Termomecánico-dinámico (DMTA)

El estudio dinámico se basa en el hecho de que al aplicar una fuerza sinusoidal, variable en función del tiempo, t , con una determinada frecuencia angular, ϖ , produce una deformación también variable sinusoidalmente.

El DMTA combina la reología de sólidos con el análisis térmico y permite la medida de propiedades mecánicas, mientras se somete la muestra a un programa térmico. Esta técnica ofrece la posibilidad de trabajar a temperatura constante, obteniendo la respuesta en función del tiempo, o de otras variables, como puede ser la amplitud de la deformación o la frecuencia. Además, existe la posibilidad de operar en modo estacionario y, en este caso, los resultados también se pueden representar en función del tiempo.

En lo que se refiere al empleo de esta técnica (DMTA), las mediciones más típicas son la temperatura de transición vítrea y temperaturas características de otras relajaciones secundarias. En estos casos se suele variar la temperatura en forma de rampa.

La tensión (σ) y la deformación (ξ) pueden expresarse mediante:

$$\begin{aligned}\sigma(t) &= \sigma_0 \sin(\varpi t + \delta) \\ \xi(t) &= \xi_0 \sin \varpi t,\end{aligned}\tag{1.1}$$

donde σ_0 y ξ_0 son la amplitud de la tensión y de la deformación máximas,

respectivamente, siendo δ una constante que representa la magnitud de la diferencia de fase entre la tensión aplicada y la deformación producida, dependiente de cada material a estudio.

Las ecuaciones de la expresión (1.1) suelen relacionarse, teniendo en cuenta la teoría lineal de la viscoelasticidad, mediante la expresión:

$$\begin{aligned}\sigma(t) &= E^*(\omega)\xi(t) \\ E^*(\omega) &= E' + iE''.\end{aligned}$$

El valor $E^*(\omega)$ se denomina módulo dinamomecánico complejo, siendo E' el módulo de almacenamiento (razón entre la tensión y la deformación) y E'' el módulo de pérdidas (componente desfasada 90° de la razón antes citada).

La razón entre las dos componentes E'' y E' es la tangente del ángulo de desfase, denominado factor de pérdidas, o $\tan \delta$, donde:

$$\tan \delta = \frac{E''}{E'}.$$

El módulo de almacenamiento es uno de los parámetros más utilizados para estudiar los cambios del material en función de la temperatura o del tiempo.

1.3.6 Técnicas complementarias

En general se considera que las distintas técnicas analíticas que existen son complementarias unas de otras, ya que permiten obtener diferentes propiedades del material sujeto a estudio. Sin embargo, existen tres técnicas que normalmente se combinan con el TGA para identificar los componentes en los que se ha descompuesto la muestra por efecto del tratamiento térmico: cromatografía de gases (GC), espectrometría de infrarrojos por transformada de Fourier (FTIR) y espectrometría de masas (MS). La GC permite separar las componentes de una muestra gaseosa, las cuáles pueden analizarse posteriormente mediante MS o FTIR.

La utilización combinada de estas técnicas con el TGA o el STA permite conocer, simultáneamente, en qué sustancias se descompone la muestra en cada instante del tratamiento térmico y cuáles son los cambios térmicos asociados a cada proceso de transformación.

Usada de forma independiente, la técnica FTIR es el mejor complemento al análisis térmico: permite identificar estructuras moleculares en la muestra, y, a partir de esta identificación, es posible, por ejemplo, determinar cinéticas del curado. Consiste básicamente en hacer un barrido de frecuencias en la zona IR del espectro, de forma que el haz de luz IR atraviese la muestra. Finalmente, a los datos obtenidos se les aplica una transformada de Fourier. Se obtiene un espectro de absorvancias en función de la longitud de onda o de la frecuencia, que es característica de cada compuesto, y que permite identificar grupos funcionales, siendo usada para la clasificación de materiales.

Otra técnica complementaria es la microscopía electrónica de barrido (SEM) que proporciona imágenes de alta resolución espacial, y posibilita la observación

y estudio de superficies muy reducidas de cualquier parte del material. Si además se cuenta con un sistema de análisis de energías dispersas de rayos X (EDX), se pueden identificar los elementos en las áreas de observación. La imagen se genera por la interacción de un haz de electrones que barre un área determinada sobre la superficie de la muestra. El uso más frecuente de esta técnica es la observación de la microestructura de las composiciones de las mezclas y el estudio de las superficies de fractura.

Como prueba complementaria al análisis térmico, a veces suelen hacerse ensayos mecánicos, que tienen por finalidad medir propiedades como la elasticidad o la plasticidad. La elasticidad manifiesta la capacidad de recuperar la forma primitiva al cesar el esfuerzo deformante, mientras la plasticidad manifiesta la capacidad de adquirir deformaciones permanentes. Estas propiedades se analizan mediante ensayos de tracción y flexión, que pueden ser estáticos o estacionarios. La mayoría de estos ensayos mecánicos se encuentran normalizados, y aparecen publicados en normas internacionales como las ISO o las ASTM. Para cada material existe una versión de estas normas, que se relacionan mediante códigos, concretamente para los polímeros se aplican la ISO/DIS 527 o la ASTM D638M en los ensayos de tensión.

1.4 Campos de aplicación del Análisis Térmico

Las aplicaciones del análisis térmico abarcan la mayoría de los materiales. Los materiales para la ingeniería suelen clasificarse en tres grandes grupos: cerámicos, metales y polímeros. Esta clasificación se basa en la diferencia que hay en el enlace entre átomos y grupos atómicos. A esta clasificación suele añadirse un cuarto grupo, los materiales compuestos, que resultan de integrar dos o más materiales distintos en otro con propiedades diferentes a las que tendría cada uno por separado, pero manteniéndose los elementos integrantes en fases separadas.

1.4.1 Cerámicos y vidrios

La mayoría de los cerámicos de uso industrial incluyen compuestos inorgánicos con más de dos elementos, constituidos por, al menos, un elemento metálico y uno no metálico (C, N, O, P o S).

El óxido de aluminio (Al_2O_3) es uno de los materiales cerámicos más importantes en ingeniería, que posee dos grandes ventajas con respecto al aluminio metálico: en primer lugar es químicamente estable en una gran variedad de ambientes severos, en segundo lugar tiene una temperatura de fusión significativamente mayor (2020°C) que el aluminio metálico (660°C); esto hace del Al_2O_3 un refractario bastante común, esto es, un material resistente a altas temperaturas, ampliamente utilizado en la construcción de hornos.

La caolinita, que se usa desde antiguo en la fabricación de piezas cerámicas, se sigue usando hoy en día, tanto en la industria (porcelanas, gres, loza, electrocerámica), como en la construcción (fabricación de aislantes y cementos).

Los cerámicos, al igual que los metales, que trataremos más adelante, tienen la propiedad de ser cristalinos, lo que significa que sus átomos están unidos según una distribución regular y repetida. Mediante técnicas de procesamiento bastante simples, muchos cerámicos pueden fabricarse de forma no cristalina, es decir, sus átomos están dispuestos de forma irregular y aleatoria, denominándose vidrios. La mayoría de los vidrios comunes son silicatos; el vidrio utilizado para ventanas está compuesto por, aproximadamente, el 72% de peso de sílice (SiO_2), siendo el resto óxido de sodio (Na_2O) y óxido de calcio (CaO). La importancia de los vidrios en ingeniería reside en su capacidad para transmitir la luz visible (así como la radiación ultravioleta e infrarroja) y su inercia química.

Los recientes desarrollos en la tecnología de los cerámicos están haciendo aumentar su empleo en aplicaciones estructurales, eliminando su inherente fragilidad y aumentando su resistencia a la fractura.

Finalmente, cabe resaltar un desarrollo reciente en este grupo de materiales: los vitrocerámicos. La base de los vitrocerámicos está en la propiedad que tienen algunas formulaciones de vidrio (como los aluminosilicatos de litio) de poder desvitrificarse totalmente (esto es, experimentan una transformación del estado vítreo al estado cristalino) mediante un tratamiento térmico adecuado, que los hace más resistentes a la fractura por cambios de temperatura.

1.4.2 Metales y aleaciones

Los materiales metálicos se caracterizan por el enlace denominado metálico, en el cual existe un gran número de electrones deslocalizados que no pertenecen a un átomo en particular sino al conjunto de ellos. Son buenos conductores de electricidad y del calor, tienen interesantes propiedades mecánicas como ductilidad o conformabilidad, resistencia mecánica, resistencia al impacto y alta rigidez. Aunque ocasionalmente se utilizan en forma pura, se prefiere, normalmente, el empleo de sus combinaciones, denominadas aleaciones, para mejorar ciertas propiedades deseadas o permitir una combinación de las mismas.

Los metales forman un grupo de materiales de gran uso industrial, con un desarrollo ligado a sus propiedades térmicas. La mayoría de los metales empleados en la industria son aleaciones férricas y, en la práctica, estas aleaciones se dividen en dos grandes categorías dependiendo de la cantidad de carbono: los aceros, que generalmente tienen un contenido en carbono comprendido entre un 0.05% y un 2% de su peso, y las fundiciones, que contienen entre un 2% y un 4.5% en peso de carbono. Además, pueden emplearse otro tipo de elementos distintos al carbono y al hierro, lo que daría lugar a aceros y fundiciones aleadas: una concentración de un 5% en peso de elementos aleantes distintos del carbono marcaría la frontera entre aceros de baja aleación y aceros de alta aleación.

Los materiales metálicos, al tener su microestructura condicionada, no sólo por la temperatura, sino también por las condiciones de elaboración (velocidades de calentamiento y enfriamiento), se prestan a la aplicación de las técnicas de análisis térmico. Así, se pueden determinar temperaturas de fusión, transformaciones alotrópicas, cambios en el comportamiento magnético, variación del calor específico, estabilidad térmica, etc.

En la década de los setenta es cuando comienza a aplicarse, con carácter más o menos general, el estudio térmico a los metales. En este tiempo, ya se habían producido grandes avances sobre el comportamiento de los mismos, fundamentalmente en el caso de las aleaciones de mayor demanda industrial, como los aceros empleados en construcción naval, o de aleaciones ligeras como las de aluminio. Instituciones del prestigio de la ASM (American Society for Metals) han elaborado y publicado atlas de curvas TTT (Temperatura, Tiempo, Transformación) y TTC (Temperatura, Tiempo, Continuo) para los aceros comerciales, aplicando para ello técnicas de metalografía y dilatometría (técnica que mide la variación en las dimensiones de la muestra frente a la temperatura).

Entre las aplicaciones específicas de las técnicas de análisis térmico cabe destacar la aplicación de técnicas DSC a materiales con memoria de forma para la determinación de temperaturas características y cambios de entalpía. Otra gran aplicación del análisis térmico se encuentra en el estudio del acero estructural, utilizado por ejemplo en la construcción naval. Es bien conocido el efecto que tienen las bajas temperaturas en el comportamiento del acero, al menos desde el fatídico suceso del Titanic, que puede estar relacionado con la baja calidad del acero empleado (véase Naya (2001)).

Las aleaciones más empleadas en ingeniería son: las de aluminio, magnesio, titanio, níquel, cinc y las de cobre. Un subgrupo dentro de la aleaciones lo constituyen las superaleaciones que contienen una amplia gama de metales con resistencia extraordinaria a altas temperaturas (incluso por encima de 1000°C). Son destacables las superaleaciones de base cobalto y de base níquel (véase Artiaga et al. (2002)).

1.4.3 Polímeros

Los polímeros han cobrado importancia con posterioridad a los metales y cerámicos. Su nombre proviene de la composición de los vocablos griegos *poly* (muchos) y *meros* (partes). Son sustancias de alto peso molecular, constituidas por repetición múltiple de unidades estructurales iguales obtenidas por reacción de compuestos sencillos, llamados monómeros, que se unen formando cadenas. Los cambios de temperatura provocan que estas cadenas cambien de conformación o incluso se degraden, lo que hace de gran interés el estudio de la influencia del calor en los cambios de estructura y, por tanto, justifica, de forma especial, su estudio térmico.

Según su estructura molecular se clasifican en termoplásticos, termoestables y elastómeros.

Los termoplásticos, como el polietileno $(-(-C_2H_4)-)_n$, están formados por cadenas lineales o ramificadas. Las fuerzas intermoleculares son débiles y se reducen por efecto del calor, volviéndose moldeables: el calentamiento hace que se ablanden.

Los termoestables, como el poliéster, están compuestos por macromoléculas con un grado de reticulación muy elevado. A temperatura ambiente son duros y a diferencia de los termoplásticos, arden con el fuego, pero no se ablandan.

Algunos elastómeros tienen un grado de reticulación bajo. A temperatura ambiente se comportan como el caucho y no son fundibles ni solubles. Otros son del tipo termoplástico y se les suele conocer como termoelastómeros.

Otra posible clasificación se obtiene teniendo en cuenta el mecanismo de polimerización o procedimiento utilizado para su obtención: polimerización por adición (como el poliestireno), por apertura de anillo (como el polieter) o por condensación (como el nylon o el poliester).

También es posible agruparlos según su origen en naturales como el caucho, o polímeros sintéticos como el polipropileno.

Una propiedad importante de algunos polímeros es la posibilidad de pasar del estado vítreo al cristalino (cristalización), convirtiéndose en un material más rígido. Mediante las técnicas de DSC es posible encontrar el punto crítico en el que se produce esta cristalización, denominado temperatura de cristalización (T_c).

Desde que los polímeros han comenzado a ser utilizados masivamente por la industria (como el polietileno), han suscitado gran interés por parte de la comunidad científica. La preocupación por el envejecimiento y reciclaje de estos materiales es bien evidente hoy día. Por efecto del entorno pueden experimentar cambios irreversibles en su estructura, y, por este motivo, el estudio de su degradación es un tema de gran actualidad como puede verse en el trabajo de Pagès et al. (1996).

1.4.4 Materiales heterogéneos de base polimérica (mezclas, maderas y otros productos naturales)

Las tres categorías anteriores de materiales empleados en ingeniería (cerámicos y vidrios, metales y polímeros) engloban una gran cantidad de ejemplos. Sin embargo, existen materiales que no pueden ser catalogados en estas tres clases, y que llamaremos materiales compuestos, y quizá, los mejores ejemplos los constituyen el plástico reforzado con fibra de vidrio o carbono y el hormigón.

El hormigón es un material compuesto de gran uso en la construcción, que se forma por la mezcla del clínker de cemento con agua y yeso, pudiendo agregarse también cuarzo, caliza, dolomita y escorias. El clínker se forma por reacción de óxido de calcio (CaO), sílice (SiO_2), alúmina (Al_2O_3) y óxido de hierro al ser calentado a una temperatura de 1500°C aproximadamente. La aplicación del análisis termogravimétrico a muestras de cemento nos proporciona las diferentes etapas en que este se descompone.

Los plásticos reforzados con fibras (FRP) están formados por una serie de fibras de vidrio o carbono embebidas en una matriz polimérica. Su utilización se ha extendido en los últimos tiempos debido a sus excelentes propiedades. El plástico reforzado con fibra de vidrio reúne lo mejor de sus componentes por separado: la alta resistencia de las fibras se combina con la ductilidad de la matriz polimérica para producir un material resistente que es de gran aplicación en la industria.

Desde la década de los cincuenta, en que se empezaron a explotar comercialmente, las resinas epoxídicas han experimentado un continuo desarrollo, siendo

hoy una de las familias de polímeros de condensación más importantes y constituyen el material de matriz más utilizado en los materiales compuestos. Como características de sus altas prestaciones caben destacar sus propiedades de adhesión, resistencia, baja contracción, protección frente a la corrosión y versatilidad en el procesado.

El estudio de estos materiales se dirige últimamente a mejorar las formulaciones, con el fin de acomodar en lo posible las características de las resinas conocidas a nuevas aplicaciones en las que entran en competencia con otros materiales como las poliamidas. Las tendencias actuales se interesan por la mejora de las propiedades térmicas, resistencia a la humedad, resistencia a la llama, nuevos sistemas de curado y mejora de las propiedades mecánicas.

Las resinas epoxi tienen aplicación en campos tan diversos como los recubrimientos, adhesivos, fabricación de embarcaciones deportivas, matrices de materiales compuestos o encapsulación de componentes electrónicos.

Para la obtención de materiales compuestos, estas resinas se combinan con una enorme variedad de refuerzos (fibras de vidrio, carbono, aramida, etc.), por distintas técnicas: laminado manual, con saco a vacío, bobinado de filamentos, moldeado en prensa, etc.

En la microencapsulación de componentes electrónicos cada vez se exigen mejores propiedades de resistencia a la temperatura, y en este tipo de aplicación las propiedades termomecánicas cobran una especial relevancia. Así, por ejemplo, el comportamiento del adhesivo que une el microprocesador Pentium con el disipador de calor, debe ser adecuado para soportar temperaturas moderadamente altas, durante largos períodos de tiempo, y adaptarse a los cambios dimensionales de las componentes a unir, que se producen debido a la diferencia de temperatura entre el estado de “encendido” y “apagado”.

La madera es un excelente ejemplo de un material natural con propiedades mecánicas útiles debido a su estructura reforzada de fibras. Sería interesante poder identificar el tipo de madera a partir de su estudio mediante análisis térmico.

1.4.5 Ejemplo de aplicación de varias técnicas de Análisis Térmico

Se ha elegido una resina epoxi como ejemplo de material cuya caracterización requiere varias técnicas de análisis térmico. Se trata de un material termoendurecible, por lo que será útil el estudio del proceso de curado y las propiedades del material ya curado. Las distintas fases y técnicas aplicables se describen a continuación:

1. Determinar mediante DSC la entalpía de curado de las distintas formulaciones. Esta característica, aunque en algunos casos puede no ser muy importante, en otros puede ser crítica (adhesivos para explosivos, encapsulación de microchips).

2. Determinar mediante TGA la estabilidad térmica de la resina utilizando distintas proporciones de diluyente. La determinación se realizará tanto en atmósfera oxidante como inerte, lo que permite seleccionar las formulaciones más

adecuadas, desde el punto de vista de la estabilidad térmica, para aplicaciones de elevada exigencia térmica como la encapsulación de componentes electrónicos.

3. Caracterización mecanodinámica de los materiales obtenidos. El estudio de las propiedades mecánicas a distintas temperaturas es de vital importancia para predecir el comportamiento en servicio de estos materiales.

4. Establecer correlaciones entre las propiedades térmicas y mecánicas con la composición. Estas correlaciones nos permitirán diseñar formulaciones específicas para aplicaciones concretas, teniendo en cuenta el tipo de condiciones mecánicas que deberá soportar el material, el ambiente en el que deberá trabajar y los intervalos de temperaturas y picos térmicos que se pueden presentar en servicio.

En tales circunstancias, un objetivo primordial es la utilización del análisis térmico simultáneo (STA) para realizar un curado dinámico y comprobar la estabilidad térmica de dichos materiales en ambientes oxidantes y no oxidantes. Con tal fin, el análisis de las curvas DSC y TGA, obtenidas mediante la citada técnica, permite observar no sólo cómo afecta la cantidad de diluyente a la estabilidad térmica, sino también clarificar cómo se incorpora el diluyente a la estructura tridimensional entrecruzada. En este aspecto es fundamental el correcto ajuste de las curvas, mediante técnicas de suavización muy precisas, objeto de capítulos posteriores en esta memoria, pues los datos registrados por el calorímetro se hallan sujetos a error experimental.

1.5 Comportamientos en Análisis Térmico

Existe un número reducido de eventos típicos individuales que pueden aparecer combinados de distintas maneras dependiendo del material que se analiza.

Con vistas a considerar la mayor parte de las situaciones que pueden darse en análisis térmico, se hizo una revisión de los experimentos realizados a lo largo de los últimos seis años en el laboratorio de Ciencia de Materiales de la Escuela Politécnica Superior de Ferrol, extrayéndose unos cuantos ejemplos, con la intención de representar distintas casuísticas reales que se han observado y que merecen ser tenidas en cuenta, con el fin de justificar el adecuado ajuste de los datos que abordaremos en los próximos capítulos de esta memoria.

1.5.1 Evaporación del Anhídrido Hexahidroftálico

En la figura 1.3 se observa la aplicación de la técnica STA a una muestra de anhídrido hexahidroftálico. La curva DSC (en azul) indica la existencia de dos procesos endotérmicos, y la correspondiente curva TGA (en rojo) confirma que en el segundo proceso hay una pérdida de masa de la muestra. Además, la forma del segundo pico, mínimo absoluto de la curva DSC, indica que la máxima velocidad se manifiesta al final del proceso. Todo ello permite afirmar que el primer pico corresponde a una fusión, mientras que el segundo es una evaporación.

1.5.2 Descomposición térmica de minerales

El estudio de un mortero románico (procedente de la catedral de Mondoñedo) nos servirá como ejemplo de la descomposición térmica de minerales.

En la figura 1.4 pueden observarse las gráficas DSC (azul), TGA (rojo) y DTGA (verde), correspondiente a una muestra de mortero, en función de la temperatura. Se aprecia cómo se descompone el carbonato cálcico en torno a 700°C. Además, en una primera parte (zonas A y B), se puede observar la pérdida de agua debida a la hidratación del yeso alrededor de 120°C (zona B), y a temperaturas inferiores pérdidas de masa debidas, posiblemente, a la humedad no ligada (zona A). Finalmente, se aprecian otras tres zonas diferenciadas: una parte de estabilidad de la muestra (zona C), una pérdida rápida de peso (zona D) y una última zona de desintegración de la muestra (zona E).

1.5.3 Comportamiento de una superaleación a altas temperaturas

En la figura 1.5 se representan las curvas DSC (azul) y TGA (rojo) obtenidas en el estudio de una superaleación de base níquel. El estudio STA en el intervalo de 50°C a 1400°C permitió evaluar la capacidad calorífica del material mediante la comparación de los datos del flujo de calor de la muestra con los del zafiro, que se utilizó como material estándar de referencia.

Este estudio permitió observar aumentos de peso a temperaturas superiores a 1000°C, que hemos atribuido a un proceso de oxidación por el oxígeno del aire. También se han detectado dos procesos endotérmicos que indican fusiones producidas en algunas fases del experimento.

1.5.4 Fusión y cristalización de un polímero cristalino

En la figura 1.6 se detalla un proceso de fusión del polímero cristalino PET. La fusión y cristalización son fenómenos irreversibles que presentan los polímeros cristalinos. En la correspondiente curva DSC se puede observar que el área del primer pico de cristalización es inferior a la del pico de fusión, lo cual permite deducir que el material se encontraba parcialmente cristalizado al iniciar el experimento. Este dato proporciona información de interés sobre la historia térmica del material. Se han señalado, en la propia figura, puntos de interés que se pueden observar a partir del estudio DSC.

1.5.5 Pirólisis de un polímero termoestable

La pirólisis es la conversión de la muestra en otra sustancia por medio del calor, y, en sentido estricto, debe de realizarse en ausencia de oxígeno. Este proceso puede conducir a moléculas de menor masa, debido a la fisión térmica, o puede desembocar en un aumento del peso molecular por medio de reacciones intermoleculares, dependiendo de las condiciones elegidas (véase Conesa (2000)).

En la figura 1.7 puede observarse la degradación térmica de una resina epoxi en atmósfera inerte. En este caso, al no haber presencia de oxígeno, es normal que predominen los procesos endotérmicos de pirólisis, es decir, las macromoléculas rompen en moléculas más pequeñas, algunas de las cuales se volatilizan a las temperaturas del experimento. Este es un proceso que consume calor.

1.5.6 Degradación de un termoestable en aire

Comparamos en este caso la misma resina epoxi del caso anterior, pero en corriente de aire seco. Puede observarse en la figura 1.8 que buena parte del proceso degradativo (pérdida de masa) va acompañado de un pico exotérmico. Esto se debe a que los “trozos” de molécula que se volatilizan arden a esas temperaturas en presencia de oxígeno.

1.5.7 Determinación de la temperatura de transición vítrea (T_g)

La temperatura de transición vítrea (T_g), entre otras cosas y con ciertas matizaciones, indica un cambio de comportamiento de la muestra de rígido a blando. Los polímeros presentan un comportamiento rígido y frágil a temperaturas inferiores a la T_g, mientras que a temperaturas superiores su comportamiento es más blando y tenaz.

La T_g puede determinarse por varias técnicas, siendo la más extendida la DSC. Sin embargo, el DMA permite localizar más fácilmente esa relajación del material (propriadamente no se considera una transición, a pesar del nombre).

En la figura 1.9 puede observarse como aparece la T_g del PVC por las dos técnicas mencionadas, normalmente la T_g aparece a temperaturas inferiores por DSC que por DMA, debido a que las dos técnicas detectan cambios de material a distinta escala.

1.5.8 Dos transiciones vítreas en una sola muestra

En la figura 1.10 se puede observar una muestra procedente de un trozo de una hoja de transparencia de las usadas para impresoras de chorro de tinta (concretamente la catalogada como 639-31-A4 de la casa Staedtler). En la gráfica de $\tan \delta$ se pueden observar dos picos que se pueden atribuir a dos T_gs. Cada T_g corresponde a un polímero: el PET y un recubrimiento superficial en una de las caras, que permite que se fije la tinta.

1.5.9 Comparación de maderas

Las maderas, como material lignocelulósico, presentan complejos mecanismos de reacción térmica. Normalmente, en materiales de los denominados biomasa, como el caso de la madera, suelen estudiarse por separado sus componentes: hemicelulosa, celulosa y lignina.

La pirólisis de celulosa ha sido ampliamente investigada y, sin embargo, no se conocen completamente la cinética ni el mecanismo de esta pirolisis, debido probablemente a la compleja naturaleza de la reacción. En la figura 1.11 se aprecian las curvas TGA de dos muestras de madera (eucalipto y castaño), que como puede verse presentan una forma muy similar, en cuanto a escalones, aunque a diferentes tiempos. La clasificación de estas y otras maderas y de otro tipo de materiales será abordada, con detalle, en el capítulo 5 de esta memoria.

1.5.10 Comparación de residuos forestales

Las técnicas DSC y TGA permiten evaluar las propiedades térmicas de los residuos forestales, con intención de evaluar parámetros necesarios para la lucha contra los incendios. Además, dado que cada tipo de residuo presenta un comportamiento térmico diferente, es posible identificar las distintas clases a partir de estas técnicas (véase figura 1.12).

Así, en la figura 1.13 puede observarse la comparación de dos muestras procedentes de residuos forestales de bosques canadienses. Se aprecia que, aunque el comportamiento general es muy parecido (en los dos casos predomina la celulosa y la lignina), existen diferencias que permiten distinguir ambas muestras.

1.6 Gráficas de distintos comportamientos térmicos

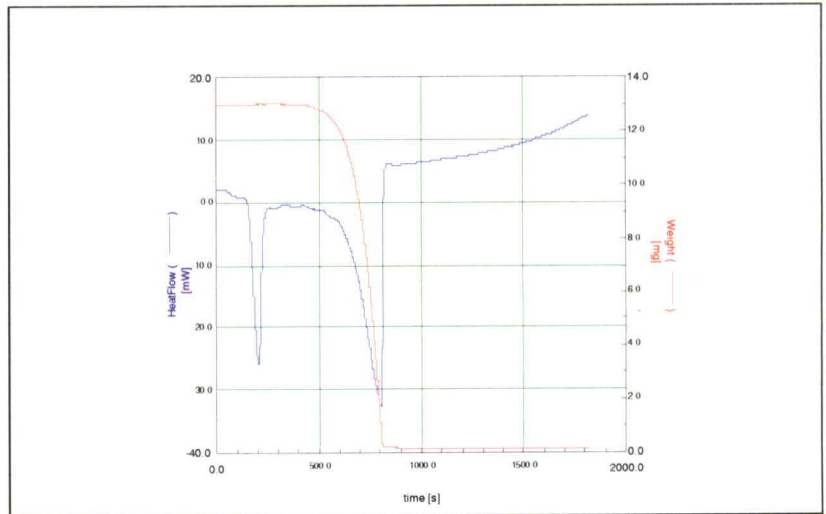


Figura 1.3: Fusión y evaporación del hexahidroftálico.

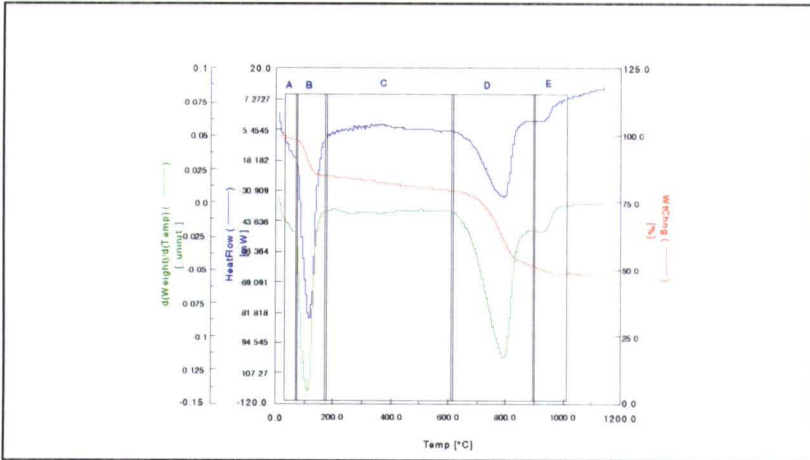


Figura 1.4: Estudio del mortero de la catedral de Mondoñedo.

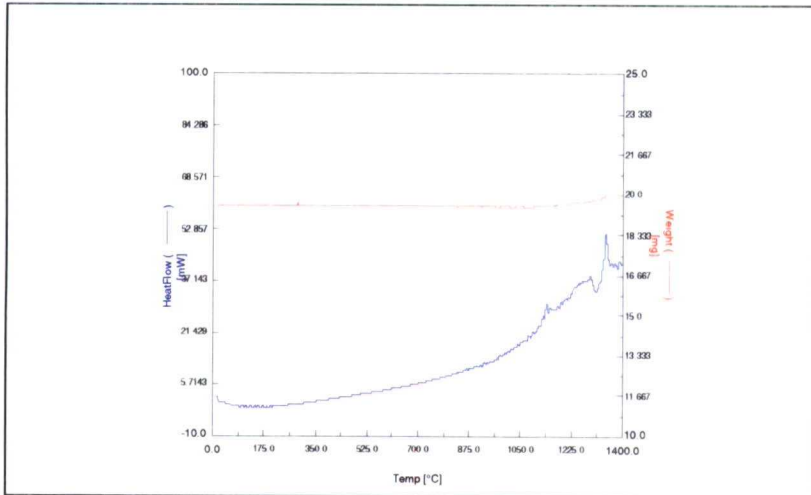


Figura 1.5: Superaleación de base níquel.

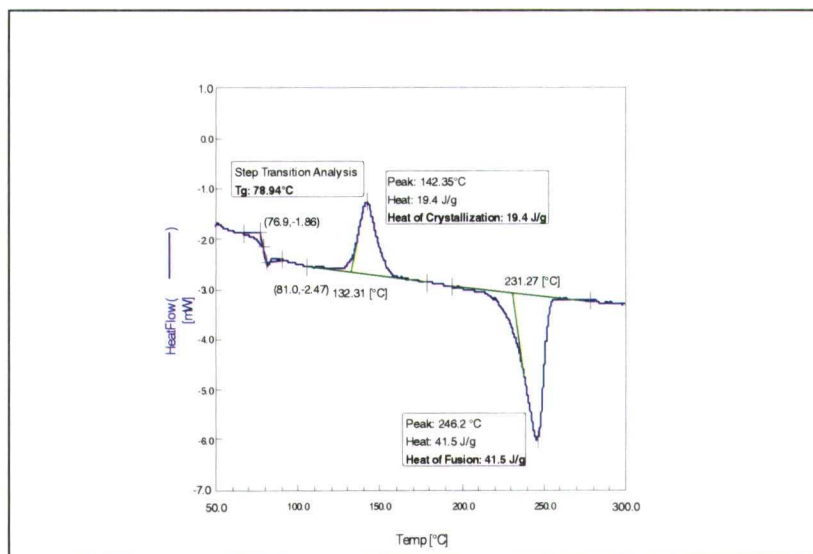


Figura 1.6: Proceso de fusión del PET.

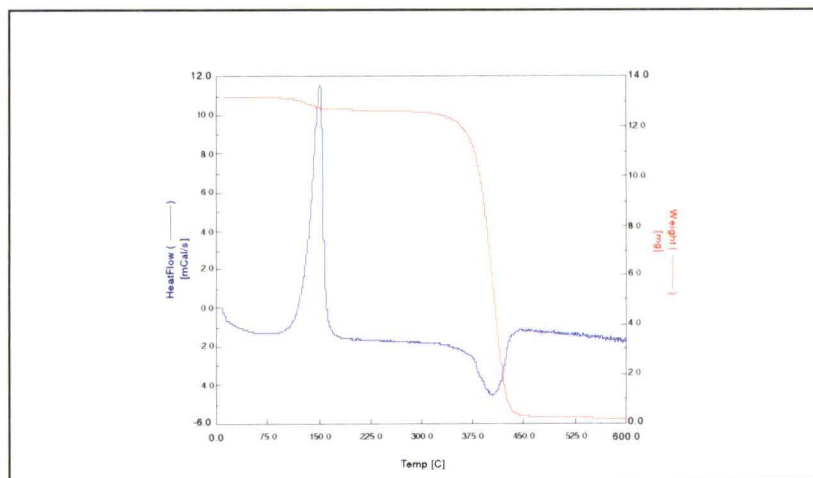


Figura 1.7: Pirólisis de un polímero termoestable.

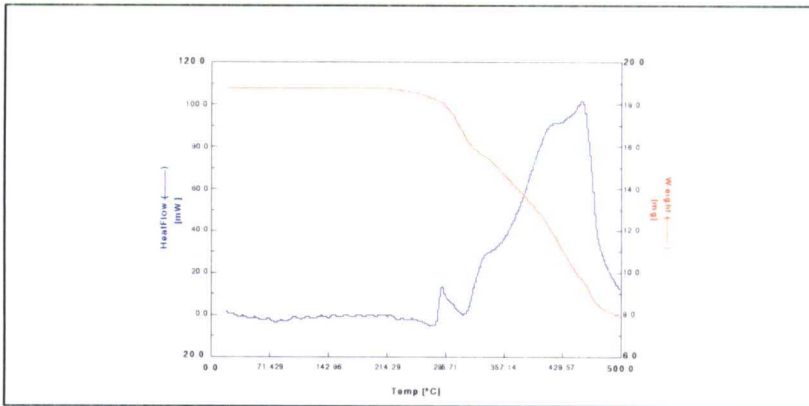


Figura 1.8: Degradación de un termoestable en aire.

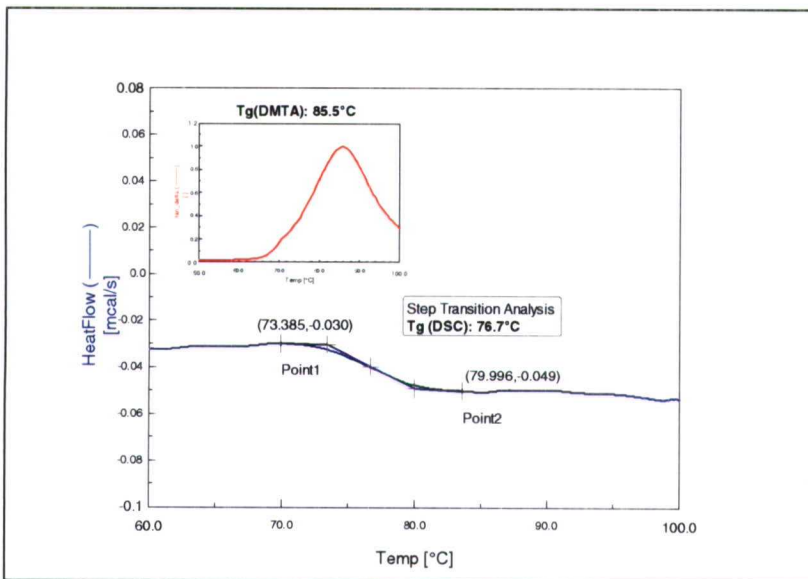


Figura 1.9: Detalle de Tg con DMA y DSC.

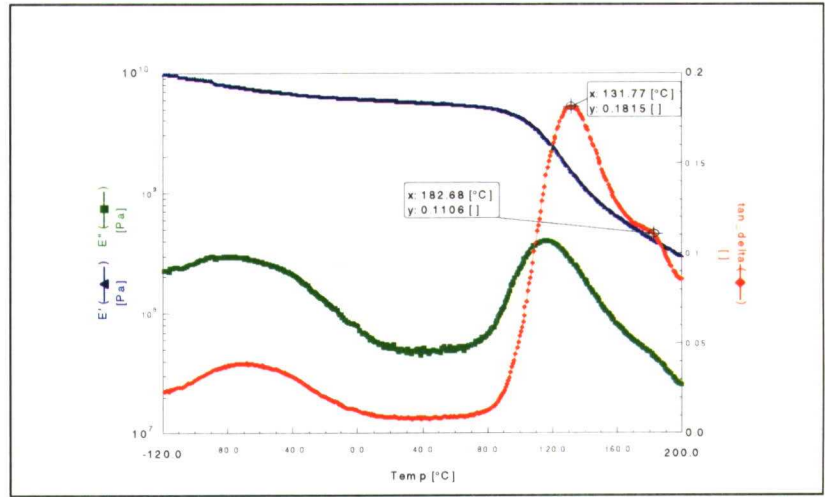


Figura 1.10: Detalle de dos transiciones vítreas.

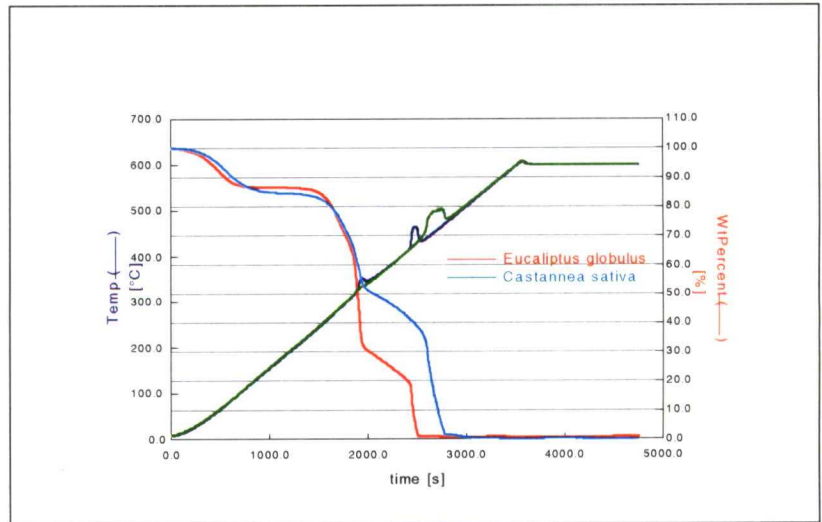


Figura 1.11: Comparación de dos muestras de madera.

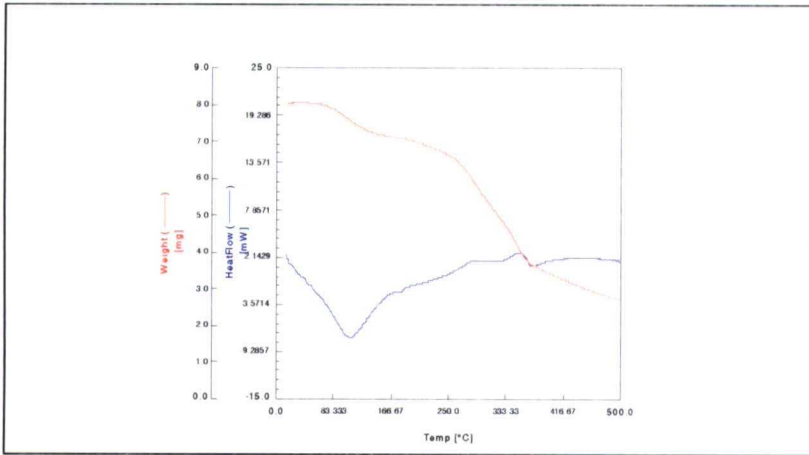


Figura 1.12: Curva TGA (en rojo) y DSC (en azul) correspondientes a una muestra de residuos forestales del Canadá.

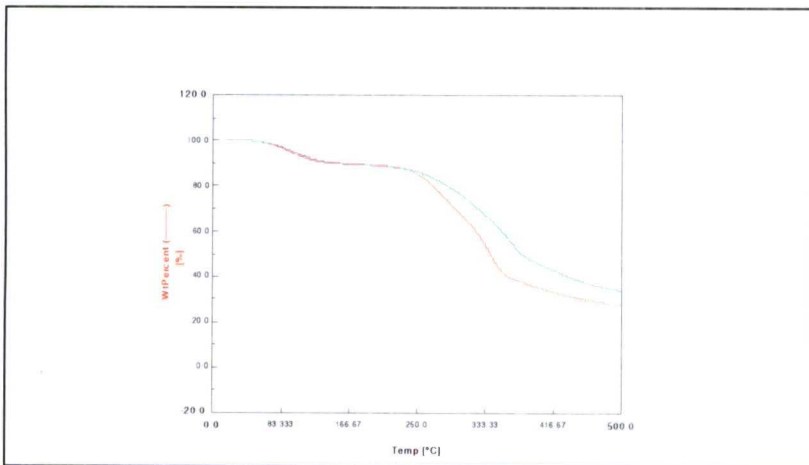


Figura 1.13: Dos curvas TGA correspondientes a dos muestras distintas de residuos forestales del Canadá.

Capítulo 2

Modelos de regresión

2.1 Introducción al suavizado de datos en Análisis Térmico

Como se ha puesto de manifiesto en el capítulo 1 de esta memoria, la información que aportan las distintas gráficas usadas en análisis térmico es crucial y, por tanto, estas señales deben tener la mayor calidad posible. Para este fin, es necesario el empleo correcto de las técnicas estadísticas de cara a encontrar el mejor método de suavización de las curvas DSC y TGA, así como las de sus derivadas.

Sin embargo, la mayoría del software existente para el tratamiento de estos datos no tiene en cuenta muchos de los avances aparecidos dentro del campo de la regresión, como es el caso de considerar la dependencia de los errores o el buscar un suavizado óptimo de las curvas y de sus derivadas.

Como primer ejemplo podemos ver, en la figura 2.1, directamente la salida gráfica proporcionada por uno de los programas habituales en el tratamiento de datos procedentes de análisis térmico. Concretamente, se ha usado el software RSI Orchestrator (Rheometric Scientific (2001)), que constituye uno de los estándares para el estudio de experimentos basados en el STA. Dicho software, al igual que otros de distintas marcas, no ofrece métodos de suavización todo lo precisos que serían deseables, especialmente en lo referente al suavizado de la primera y segunda derivadas de las curvas DSC y TGA.

En la figura 2.1 podemos observar en rojo la curva TGA, obtenida a partir de un experimento realizado con una muestra de oxalato de calcio. En azul, a su vez, la primera derivada de la curva TGA, pudiendo apreciar claramente cómo presenta un clarísimo problema de suavizado lo que limita, en gran medida, su análisis posterior a fin de obtener algún tipo de conclusión.

El software RSI Orchestrator incorpora diferentes métodos de suavización de datos: reducción de datos aplicando medias móviles, técnicas de suavizado FFT (Fast Fourier Transform) y técnicas de suavización no paramétricas de tipo lineal local.

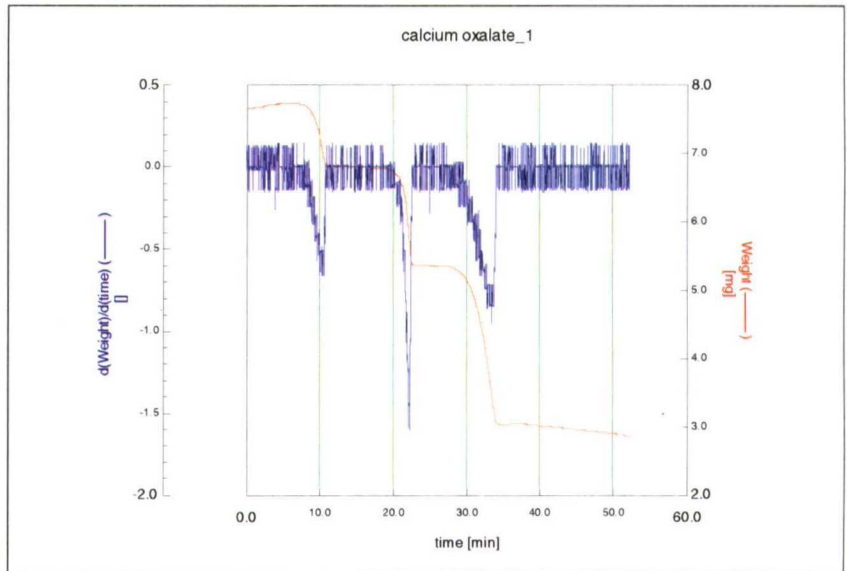


Figura 2.1: Curva TGA (en rojo) y su primera derivada (en azul) para el oxalato de calcio.

Este programa incluye como posibilidad la modificación manual del parámetro de suavizado o parámetro ventana (h) a través del teclado. En nuestra opinión esto puede mejorarse, incluyendo la posibilidad de la estimación del valor óptimo del parámetro ventana, basándose única y exclusivamente, en los propios datos de la función a suavizar.

Una prioridad de esta memoria es encontrar un algoritmo que permita estimar matemáticamente la relación entre el tiempo (o la temperatura), denotada por la variable X , y el peso o la diferencia de calor, denotado por la variable Y . Este problema se engloba en el contexto de un modelo general de regresión mediante la expresión $Y = m(X) + \varepsilon$, donde m es la función de regresión de Y dado X , y ε es una variable aleatoria que representa el error de observación.

Este capítulo tiene como objetivo introducir los diferentes tipos de modelos de regresión. Se abordará con mayor intensidad, en una última sección, el modelo de regresión no paramétrico de tipo polinómico local, que será el elegido para la suavización de las curvas procedentes de Análisis Térmico, entre ellas las curvas DSC y TGA.

2.2 Introducción a los modelos de regresión

El análisis clásico de regresión estudia las asociaciones cuantitativas entre una o varias variables explicativas X (aleatorias o no), y una variable aleatoria respuesta o dependiente Y , de modo que, a partir de un conjunto de observa-

ciones $\{(X_i, Y_i)\}_{i=1}^n$, la relación de regresión puede modelizarse como:

$$Y_i = m(X_i) + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (2.1)$$

donde $m(\cdot)$ representa la función de regresión desconocida y ε el error de observación que se supone de media cero. Uno de los objetivos del análisis de regresión es obtener una aproximación razonable de la función desconocida m , reduciendo el error de las observaciones y centrándose en la respuesta media de Y , condicionada a X , lo que se puede expresar como una esperanza condicionada:

$$m(x) = E(Y|X = x). \quad (2.2)$$

En nuestro caso, al modelizar curvas procedentes de experimentos de análisis térmico, va a ser de interés aproximar ciertas características de la función de regresión como, por ejemplo, sus derivadas.

El proceso de estimar la función desconocida $m(\cdot)$ se puede llevar a cabo de diferentes formas, en función del tipo de datos disponibles y del enfoque elegido, dando nombre a los modelos: diseño fijo o aleatorio, modelos paramétricos y no paramétricos, regresión de variables reales, vectoriales o funcionales.

2.3 Modelos de diseño aleatorio y de diseño fijo

La forma de generar los datos $\{(X_i, Y_i)\}_{i=1}^n$ da lugar a dos tipos de situaciones. En la primera de ellas, las variables regresoras, X_i , son estocásticas y se conoce como modelo de diseño aleatorio. Los datos corresponden a variables aleatorias idénticamente distribuidas que pueden ser independientes o dependientes, según la naturaleza del problema.

En el llamado modelo de diseño fijo la variable X no es estocástica, sino que está controlada por el experimentador y suelen denotarse, con la intención de diferenciarla de las de diseño aleatorio, como x_i .

En el supuesto de que esta variable explicativa, x_i , $i = 1, 2, \dots, n$, esté equiespaciada en el intervalo $[a, b]$ (sin pérdida de generalidad, el $[0, 1]$) y, por tanto, $x_i = \frac{i}{n}$, $i = 1, 2, \dots, n$, se llama modelo de regresión de diseño fijo equiespaciado. Existe también la opción de considerar la posibilidad de diseño fijo no equiespaciado, con una determinada función de densidad del diseño f , lo que significa que:

$$\int_{x_{i-1}}^{x_i} f(t)dt = \frac{1}{n}, \quad i = 1, 2, \dots, n; \quad x_0 = -\infty.$$

Ejemplos de modelos de regresión de diseño fijo son las series de tiempo con tendencia y, como caso particular, las curvas DSC y TGA objeto de nuestro estudio.

Obsérvese que cuando las variables explicativas están controladas por el experimentador, la aleatoriedad de las respuestas proviene, únicamente, de los

errores, $\{\varepsilon_i\}_{i=1}^n$, exigiéndose que sean variables aleatorias de media cero y frecuentemente de varianza constante σ^2 (homocedasticidad). Estas variables pueden ser independientes o bien presentar algún tipo de dependencia, lo que generaría un problema añadido a la estimación.

2.4 Modelos de regresión según el tipo de variables

El modelo de regresión $Y = m(X) + \varepsilon$ depende del espacio funcional al que pertenezca m y de la clasificación de las variables X e Y . La clasificación de los modelos en base al tipo de variables sobre las que se realiza el estudio considera: modelos con variables reales, con variables vectoriales y con variables funcionales.

2.4.1 Modelos con variables reales

Es el caso más simple en que asumimos que la variable explicativa X es una variable real. Las distintas técnicas estadísticas se clasifican, como se comentó anteriormente, en modelos paramétricos y no paramétricos que estudiaremos con más detalle en la siguiente sección. También tendríamos una serie de modelos intermedios como, por ejemplo, los semiparamétricos.

Como aplicación de estos modelos veremos el ajuste a curvas DSC y TGA en los capítulos 3 y 4 de esta memoria en el que se considera que la variable X es real: tiempo o temperatura de experimentos de análisis térmico.

2.4.2 Modelos con variables vectoriales y funciones

Es el supuesto de que la variable X sea un vector, es decir X toma valores en \mathbb{R}^p , m será una función de \mathbb{R}^p en \mathbb{R} .

Si la variable X es de tipo funcional, es decir, toma valores en un espacio de funciones de dimensión infinita, E , el modelo se dirá funcional o de dimensión infinita. Este tipo de modelos funcionales, bastante más recientes, está recibiendo una atención creciente por parte de la comunidad investigadora, no sólo en lo que se refiere a la regresión funcional sino, de forma más general, a todos los problemas que tienen que ver con las variables de tipo funcional.

Existen dos razones que influyen en el interés reciente por las variables de tipo funcional: el desarrollo de nuevos e interesantes instrumentos teóricos y el inmenso potencial para este tipo de modelos gracias al avance informático (tratamiento de imágenes, espectrometría, econometría, etc). Son destacables las aportaciones teóricas de las monografías de Ramsay y Silverman (1997, 2002) y las de Ferraty, Núñez-Antón y Vieu (2001). Como ejemplo de aplicación de estas técnicas destacan las de Besse et al. (1997) y Ferraty y Vieu (2002a, 2002b y 2003).

Como aplicación de regresión funcional se estudiará, en el capítulo 5 de esta memoria, la clasificación de materiales. Se tomará como variable X las distintas

curvas TGA (funciones) procedentes de los experimentos de laboratorio, y como variable Y , indicadores de pertenencia a los diferentes grupos en los que pueden clasificarse los materiales a estudio.

2.5 Modelos paramétricos

Un modelo de regresión paramétrico es aquel en que la curva de regresión teórica m pertenece a una familia indicada por un parámetro finito dimensional (por ejemplo, un polinomio). Una vez realizada esta suposición, las técnicas paramétricas se basan en utilizar la muestra para estimar los parámetros desconocidos (como los coeficientes del polinomio) usando, por ejemplo, el método de mínimos cuadrados.

Los modelos de regresión de tipo paramétrico vienen usándose desde antiguo para encontrar relaciones entre distintas variables estadísticas. Algunos autores, como el historiador Carl Boyer, datan sus principios en el cálculo de órbitas de planetas desarrollado por Gauss en el año 1800, quien, por vez primera utiliza el método de mínimos cuadrados. Sin embargo el nombre de regresión lo utiliza por vez primera Francis Galton en sus trabajos de Biología a finales del siglo XIX.

A partir de los primeros estudios de regresión fueron muchos los estadísticos que aportaron resultados importantes, resaltando Karl Pearson (1925) que introduce el método de ajuste ortogonal por mínimos cuadrados. En 1930 Wiener y Kolmogorov demuestran que el método de los mínimos cuadrados proporciona estimadores óptimos, aunque será el conocido teorema de Gauss-Markov el fundamento teórico principal que justifica la utilización de este método. Finalmente, Ronald Fisher (1935) estudia las transformaciones del coeficiente de correlación, dando métodos para la construcción de intervalos de confianza y contrastes de hipótesis para los parámetros del modelo.

El método de mínimos cuadrados consiste en encontrar los valores de aquellos parámetros que mejor ajustan la nube de puntos de los datos muestrales, minimizando la suma de las distancias al cuadrado de la expresión siguiente:

$$\sum_{i=1}^n (y_i - m(x_i))^2. \quad (2.3)$$

La elección de $m(x)$ estará marcada por el problema particular que se analice. Para cualquier especificación funcional: $m(x) = m(a_0, a_1, a_2, \dots, a_p; x)$, siendo los coeficientes $a_0, a_1, a_2, \dots, a_p$, los parámetros del modelo.

Una posible clasificación de los modelos paramétricos los divide en modelos lineales y no lineales, entendiendo que la linealidad se observará en la expresión de la variable explicativa X .

2.5.1 Modelos paramétricos lineales

Los modelos lineales serán aquellos en que se supone que la media de la distribución de la variable Y , para cada valor x fijo, crece linealmente con x . Esta

hipótesis debe ser contrastada siempre, ya que condiciona toda la construcción del modelo. La utilidad de este modelo radica en que muchas relaciones no lineales se pueden convertir en lineales transformando las variables adecuadamente.

La metodología a seguir para la aplicación de este modelo es la siguiente:

- Estimación de los parámetros y cálculo de las propiedades de los estimadores.
- Cálculo de intervalos de confianza y realización de contrastes sobre los parámetros.
- Diagnóstico del modelo mediante el análisis de las hipótesis del modelo por medio de los residuos (diferencia entre valores pronosticados y observados).

Como caso particular del modelo lineal, está el de suponer un modelo polinómico de grado uno para la relación; en este caso, basta escoger a y b de tal forma que se haga mínima la cantidad siguiente:

$$\sum_{i=1}^n (y_i - a - bx_i)^2. \quad (2.4)$$

Al optimizar la suma de errores cuadráticos en la predicción dada en (2.4), el óptimo se alcanza en

$$\hat{b} = \frac{S_Y}{S_X} r_{XY}, \quad \hat{a} = \bar{y} - \hat{b}\bar{x},$$

que se llaman coeficientes de la regresión lineal, y donde S_X y S_Y son, respectivamente, las desviaciones típicas de las variables X e Y , siendo r_{XY} el coeficiente de correlación lineal entre ambas variables.

Otros ejemplos de modelos de tipo polinómico son el que considera una regresión cuadrática, donde se asumirá que $m(x) = a + bx + cx^2$, y el de predicción cúbica donde $m(x) = a + bx + cx^2 + dx^3$. En general, se puede definir para polinomios de cualquier grado, p , fijo; este caso, de suponer una regresión polinómica de grado p , se puede expresar de forma matricial como $Y = \vec{a}^T X$, donde:

$$\vec{a} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^p \\ 1 & x_2 & x_2^2 & \cdots & x_2^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^p \end{pmatrix}, \quad Y = \begin{pmatrix} y_0 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

La solución al problema de estimación paramétrica de minimización de (2.3), está ampliamente estudiada pudiendo expresarse el vector de parámetros estimados en forma matricial como:

$$\widehat{\vec{a}} = (X^t X)^{-1} X^t Y \quad (2.5)$$

Un estudio detallado de este tipo de modelos puede verse, por ejemplo, en Peña (1987).

2.5.2 Modelos paramétricos no lineales

También existen otros modelos, ya no de tipo lineal, en donde la expresión de la relación de la variable Y no se adapta a la hipótesis de linealidad. En nuestro estudio utilizaremos modelos no lineales para el ajuste de curvas TGA de diversos materiales. Concretamente, se aplicará un modelo logístico dado por la expresión siguiente:

$$m(x) = \frac{e^{P(x)}}{1 + e^{P(x)}} \quad (2.6)$$

Este tipo de modelos, supone que $P(x)$ es un polinomio de grado p , y para la estimación de los parámetros se hace uso del método de mínimos cuadrados no lineales. Los fundamentos teóricos de este método pueden verse en Gay (1984).

La dificultad para muchos de estos modelos no lineales consiste en el cálculo de los parámetros por mínimos cuadrados, debido a que en ocasiones es necesaria la utilización de métodos numéricos que pueden no converger a la solución buscada, siendo necesario partir de unos valores previos para el algoritmo de aproximación.

Uno de los métodos más usados, en el caso no lineal, para el cálculo de los valores de los parámetros es el método de Levenberg-Marquart. Este método trata de generar una secuencia de aproximación al punto mínimo basada en el algoritmo de la "trust region", siendo tratado en textos como el de Gay (1984).

La suposición, a priori, de una cierta forma funcional de la curva desconocida es muy útil si la función realmente se ajusta a esa forma supuesta pero, en muchos casos, esto puede ser muy distinto de la realidad con lo que se estarán cometiendo grandes errores. Así, se podría estar ajustando una recta de regresión a una nube de puntos que no tengan relación lineal, sino parabólica.

2.6 Modelos no paramétricos

Los métodos de estimación no paramétrica de la regresión intentan estimar directamente la forma de la función de ajuste m , sin necesidad de especificar, a priori, su estructura. La filosofía de estos métodos se basa en la idea de dejar que los datos hablen por si mismos.

Estas técnicas estiman la curva punto a punto, con lo que, en principio, presentan la desventaja del gran coste computacional que tienen, si bien esto es, hoy en día, un asunto menor debido al gran avance informático.

Las técnicas de estimación no paramétricas empiezan a usarse a partir de 1940, pero es a partir de 1980 cuando su expansión y uso empieza a hacerse patente. Publicaciones como las de Härdle (1990), Fan y Gijbels (1996) y Wand y Jones, entre otras muchas, no hacen más que confirmar el desarrollo de esta nueva metodología.

Se puede citar un gran número de ventajas de los estimadores no paramétricos, así Härdle (1990) menciona como puntos a destacar los siguientes: proporcionan un método versátil para explorar una relación general entre las variables, predicen observaciones sin referencia a un modelo paramétrico fijo, proporcionan una herramienta para encontrar datos atípicos y constituyen un método flexible de sustitución para valores perdidos o interpolación de valores adyacentes de la variable explicativa.

Los estimadores no paramétricos de la función de regresión suelen tener una estructura lineal en la variable respuesta. La idea de estos estimadores no paramétricos lineales se basa en realizar promedios locales: si la función de regresión, m , se considera *suave*, entonces las observaciones con variable respuesta asociada con los X_i cercanos a un punto x contendrán información acerca del valor de $m(x)$, y su promedio, debidamente ponderado, podrá utilizarse para construir un estimador de $m(x)$:

$$\hat{m}_n(x) = \frac{1}{n} \sum_{i=1}^n W_{ni}(x) Y_i, \quad (2.7)$$

donde $W_{ni}(x)$, $i = 1, 2, \dots, n$, es una sucesión de pesos que depende de las observaciones X_1, X_2, \dots, X_n .

Imponemos que los pesos, $W_{ni}(x)$, $i = 1, 2, \dots, n$, sean positivos y estén normalizados para todo x , es decir,

$$\frac{1}{n} \sum_{i=1}^n W_{ni}(x) = 1,$$

el estimador $\hat{m}_n(x)$ se puede obtener como solución de un problema de mínimos cuadrados ponderado localmente, es decir, se puede escribir $\hat{m}_n(x)$ como una solución del siguiente problema de minimización:

$$\min_{\theta} \frac{1}{n} \sum_{i=1}^n W_{ni}(x) (Y_i - \theta)^2. \quad (2.8)$$

Esta expresión, como promedios locales, es equivalente al procedimiento de calcular un estimador de mínimos cuadrados ponderados localmente.

El estimador $\hat{m}_n(x)$, propuesto en (2.7), es un promedio de las observaciones en un entorno del punto de interés. El modo de promediar las observaciones está controlado por los pesos $W_{ni}(x)$, $i = 1, 2, \dots, n$, los cuales, a su vez, dependen de un parámetro de suavizado que regula el tamaño del entorno de x donde

se considerarán las observaciones X_i utilizadas para estimar $m(x)$. Si dicho parámetro de suavizado es grande, el promedio local se realizará con un número grande de observaciones y la curva obtenida estará *sobresuavizada*, dando lugar a un estimador, $\hat{m}_n(x)$, sesgado. Por el contrario, si el parámetro de suavizado es pequeño, podría ocurrir que el número de observaciones utilizadas para la construcción de \hat{m}_n fuera también muy pequeño, lo cual aumentará la varianza de \hat{m}_n y producirá una curva demasiado *rugosa*.

Una finalidad de esta memoria es realizar una buena selección del parámetro de suavizado, para tratar de encontrar el punto justo entre *sobresuavización* e *infrasuavización*, en el contexto de datos procedentes de análisis térmico.

2.7 Métodos de suavización tipo núcleo

En general, la función de pesos $W_{ni}(x)$ viene caracterizada por una forma y una escala. Dependiendo de cómo se considere esta forma, se obtienen diferentes tipos de estimadores.

Una forma, conceptualmente simple, de generar la sucesión de pesos es utilizar una función de densidad, función núcleo, a la que se suele denotar por K (kernel), con un parámetro de escala, parámetro de suavizado o ventana, que ajusta el tamaño y la forma de las ponderaciones cerca del punto x .

La función núcleo, K , suele ser una función continua, acotada, simétrica y con integral igual a uno. La tabla 2.1 muestra alguna de las funciones núcleo más comúnmente utilizadas:

Núcleo	$K(u)$
Uniforme	$(1/2) I(u \leq 1)$
Cuártico	$(15/16) (1 - u^2)^2 I(u \leq 1)$
Gausiano	$(2\pi)^{-1/2} \exp(-u^2/2)$
Epanechnikov	$(3/4) (1 - u^2) I(u \leq 1)$

Tabla 2.1: Ejemplo de funciones tipo núcleo.

Entre los estimadores más usados, dentro de los de tipo núcleo, están los que se presentan en los siguientes apartados.

Estimador de Nadaraya-Watson (NW)

Este estimador, que fue propuesto por Nadaraya (1964) y Watson (1964), tiene como sucesión de pesos núcleo:

$$W_{NW}(x) = \frac{1}{\hat{f}_h(x)} K_h(X_i - x), \quad (2.9)$$

donde $K_h(u) = \frac{1}{h} K(u/h)$ y $\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(X_i - x)$ es el denominador estimador núcleo de la densidad de Rosenblatt-Parzen (Rosenblatt (1956) y

Parzen (1962)); por tanto, se obtiene como expresión para este estimador, que denotaremos por \hat{m}_{NW} ,

$$\hat{m}_{NW}(x) = \frac{\sum_{i=1}^n K_h(X_i - x) Y_i}{\sum_{i=1}^n K_h(X_i - x)}. \quad (2.10)$$

La forma de los pesos, (2.9), está determinada por la forma de la función núcleo K , y el tamaño de éstos está dado por el parámetro de escala $h = h_n$. Obsérvese que la normalización de los pesos mediante $\hat{f}_h(x)$ hace posible adaptarlos a la intensidad local de las variables X_i , a la vez que garantiza que su suma es uno.

Estimador de Gasser-Müller (GM)

Este estimador propuesto por Gasser y Müller (1979) tiene como pesos los siguientes:

$$W_{GM}(x) = n \int_{s_{i-1}}^{s_i} K_h(u - x) du, \quad (2.11)$$

donde $s_0 = -\infty$, $s_n = \infty$, $x_{i-1} \leq s_i \leq x_i$, $i = 1, 2, \dots, n-1$, se eligen entre los datos, y los pares $\{(x_i, Y_i)\}_{i=1}^n$ se han ordenado previamente de manera creciente con respecto a la variable X .

La expresión de este estimador será, por tanto:

$$\hat{m}_{GM}(x) = \sum_{i=1}^n \left(\int_{s_{i-1}}^{s_i} K_h(u - x) du \right) Y_i. \quad (2.12)$$

Este estimador fue diseñado originalmente para el supuesto de diseño fijo, aunque puede emplearse en el caso de diseño aleatorio sin más que sustituir x_i por $X_{(i)}$, el i -ésimo estadístico ordenado de la variable explicativa, e Y_i por $Y_{(i)}$, la correspondiente variable respuesta.

Estimador de Priestley-Chao (PC)

El estimador de Priestley y Chao (1972) está muy relacionado con los dos anteriores, siendo su sucesión de pesos:

$$W_{PC}(x) = n(x_i - x_{i-1}) K_h(x_i - x), \text{ con } x_0 = 0, \quad (2.13)$$

donde también se suponen ordenados los datos según la variable X en el contexto de diseño fijo en el intervalo $[0, 1]$. Es fácil interpretar (2.13) en términos de (2.9), tomando $\hat{f}_h(x) = (n(x_i - x_{i-1}))^{-1}$ para $x \in (x_{i-1}, x_i]$.

2.7.1 Otros estimadores no paramétricos de la regresión

Si bien los estimadores no paramétricos tipo núcleo, presentados en la sección anterior, han sido los más utilizados y estudiados, otras muchas técnicas no paramétricas de estimación de la función de regresión han sido definidas e investigadas en la literatura especializada. En esta sección se explican someramente tres de ellas: los estimadores de series ortogonales, los métodos spline y los métodos basados en redes neuronales.

Finalmente abordaremos, con mayor detalle en la siguiente sección, el estimador polinómico local, que será el que utilizaremos en nuestro problema de ajuste de curvas calorimétricas.

Estimadores de series ortogonales

Esta técnica parte de suponer que la función de regresión puede representarse como una serie de Fourier:

$$m(x) = \sum_{j=0}^{\infty} \beta_j \varphi_j(x), \quad (2.14)$$

donde $\{\varphi_j\}_{j=0}^{\infty}$ es una base de funciones conocida y $\{\beta_j\}_{j=0}^{\infty}$ son los coeficientes de Fourier desconocidos. Una vez que las funciones base se han fijado, el problema de estimar m se reduce a estimar los coeficientes de Fourier $\{\beta_j\}_{j=0}^{\infty}$. Dada una muestra finita de tamaño n , únicamente un subconjunto finito de tales coeficientes puede ser efectivamente estimado.

Por comodidad en la representación, se supone que la variable explicativa, X , tiene por soporte el intervalo $[-1, 1]$ y que las observaciones $\{Y_i\}_{i=1}^n$ se han recogido en puntos $\{X_i\}_{i=1}^n$, en este intervalo. Si el sistema de funciones $\{\varphi_j\}$ constituye una base ortonormal en $[-1, 1]$, es decir,

$$\int_{-1}^1 \varphi_j(x) \varphi_k(x) dx = \delta_{jk} = \begin{cases} 0 & \text{si } j \neq k \\ 1 & \text{si } j = k \end{cases}$$

entonces, el coeficiente de Fourier β_j puede ser calculado como:

$$\begin{aligned} \beta_j &= \sum_{k=0}^{\infty} \beta_k \delta_{jk} \\ &= \sum_{k=0}^{\infty} \beta_k \int_{-1}^1 \varphi_k(x) \varphi_j(x) dx \\ &= \int_{-1}^1 m(x) \varphi_j(x) dx. \end{aligned} \quad (2.15)$$

Obsérvese que en la integral dada en (2.15) aparecen no sólo las funciones base conocidas, sino también la función de regresión desconocida, $m(x)$. Si esta

función m se estima de un modo razonable entonces automáticamente se tiene un estimador para β_j .

Para estimar m se parte de $\{A_i\}_{i=1}^n$, un conjunto de intervalos disjuntos tales que

$$\bigcup_{i=1}^n A_i = [-1, 1],$$

y

$$X_i \in A_i, \quad i = 1, 2, \dots, n.$$

Con este cambio de valores, la fórmula (2.15) se puede aproximar de la siguiente manera:

$$\begin{aligned} \beta_j &= \sum_{i=1}^n \int_{A_i} m(x) \varphi_j(x) dx \\ &\approx \sum_{i=1}^n m(X_i) \int_{A_i} \varphi_j(x) dx, \end{aligned} \quad (2.16)$$

si para cada i , el intervalo A_i está concentrado en torno a X_i . Sustituyendo en (2.16) $m(X_i)$ por Y_i , se tiene un estimador, $\hat{\beta}_j$, de β_j , dado por:

$$\hat{\beta}_j = \sum_{i=1}^n Y_i \int_{A_i} \varphi_j(x) dx.$$

Puesto que sólo se dispone de un número finito de observaciones, no se pueden estimar a la vez todos los coeficientes de Fourier. Si se consideran $N(n)$ términos en la representación (2.14), la función de regresión se aproxima por:

$$\hat{m}_N(x) = \sum_{j=0}^{N(n)} \hat{\beta}_j \varphi_j(x). \quad (2.17)$$

Este estimador se denomina *estimador de series ortogonales* de m . Se trata también de un suavizador lineal de la variable respuesta, con pesos:

$$W_{Ni}(x) = n \sum_{j=0}^{N(n)} \left(\int_{A_i} \varphi_j(u) du \right) \varphi_j(x),$$

y el parámetro de suavizado en este caso es $N(n)$; el número de coeficientes de Fourier que aparecen en (2.17).

Métodos Spline

Los métodos tipo Spline se basan en seleccionar como estimador de la función de regresión, la función, g , que minimice el funcional:

$$S_\lambda(g) = \sum_{i=1}^n (Y_i - g(X_i))^2 + \lambda \int (g''(x))^2 dx, \quad (2.18)$$

donde λ denota el parámetro de suavización. Este parámetro representa una forma de compensar el error residual, dado por el primer sumando de (2.18), y la variabilidad de la curva g , dado por el segundo sumando de esta igualdad.

El problema de minimizar $S_\lambda(\cdot)$ sobre la clase de todas las funciones diferenciables en el intervalo $[a, b] = [X_{(1)}, X_{(n)}]$ tiene una única solución, $\hat{m}_\lambda(x)$, que se denomina *spline cúbico* (véase Eubank (1988)). Este estimador tiene las siguientes propiedades:

1. $\hat{m}_\lambda(x)$ es un polinomio cúbico entre dos valores sucesivos de la variable X .
2. En cada observación, X_i , la curva $\hat{m}_\lambda(x)$ y sus primeras dos derivadas son continuas, pero no así la tercera derivada.
3. En los puntos extremos $X_{(1)}$ y $X_{(n)}$ la segunda derivada de $\hat{m}_\lambda(x)$ es cero.

Estas propiedades se siguen de la elección de la *función de penalización*, $\int (g''(x))^2 dx$, en términos de la segunda derivada. Si esta penalización se midiese por derivadas de mayor orden, las propiedades cambiarían.

La dificultad conceptual de las técnicas de suavización tipo spline está en el hecho de que \hat{m}_λ está definida implícitamente como la solución de un problema de minimización funcional, por lo que puede resultar complicada la obtención de propiedades estadísticas del estimador.

Métodos basados en redes neuronales

Las redes neuronales surgen a mediados del pasado siglo sobre la idea de imitar al cerebro humano, con la intención de simular su inteligencia y así poder tomar decisiones tras un período de aprendizaje.

Una red neuronal está constituida por un conjunto de neuronas, también llamadas *nodos*, distribuidas en *capas*. La capa de entrada es aquella en que los datos se presentan en la red; las variables que conforman estos datos reciben el nombre de *inputs*; la capa de salida será aquella que devuelve la respuesta y recibe el nombre de *output*. Las capas intermedias se denominan *capas ocultas*. La información que proporcionan las diferentes observaciones se almacena en las conexiones que comunican los nodos, llamadas pesos. Es frecuente el uso de redes neuronales que consideran que los nodos de una capa están conectados sólo con los de la capa inmediatamente superior, considerando la primera capa la de entrada, y como última, la de salida.

El funcionamiento de la red es el siguiente: se dispone de una serie de datos (situaciones del pasado) y asociados a ellos la respuesta deseable de la red (*training set*). La red, observando estos datos aprende de ellos (*entrenamiento o aprendizaje*), para finalmente evaluarse en nuevos datos (*validation set*).

Las redes neuronales pueden usarse para estimar la regresión (véase, por ejemplo, Broomhead y Lowe (1988)). Su fundamento teórico se apoya en el teorema de Kolmogoroff (1957), que afirma que una función de k -variables, m , puede expresarse como suma de funciones unidimensionales, g_m , evaluadas en combinaciones lineales de funciones, φ_j , dependientes de las coordenadas del vector. Una expresión de este resultado es la siguiente:

$$m(x_1, x_2, \dots, x_k) = \sum_{j=1}^{2k+1} g_m \left(\sum_{i=1}^k \lambda_i \varphi_j(x_i) \right).$$

En algunos casos, las redes neuronales presentan ventajas frente a los estimadores tipo núcleo tradicionales al no estar sometidas a la rigidez de una función núcleo fija para toda la muestra. Sin embargo, una desventaja es la dificultad de implementación computacional para la estimación de la regresión y sus derivadas.

2.8 Estimador polinómico local (RPL)

La idea del estimador RPL se basa en realizar ajustes locales ponderados por pesos tipo núcleo, utilizando mínimos cuadrados de funciones polinómicas. Esta idea se aplicaba desde hacía tiempo en el contexto de series de tiempo, (Macauley (1931)), y fue posteriormente introducida en el contexto de la regresión no paramétrica por Stone (1977) y Cleveland (1979), este último con la técnica LOWESS, que combina las ideas de los estimadores polinómico locales y la de estimadores robustos.

En los últimos años los estimadores de tipo polinómico local han ganado una gran popularidad entre la comunidad estadística, debido a las buenas propiedades tanto teóricas como aplicadas que han sido mostradas en numerosos artículos. Es a partir de los trabajos de Fan y Gijbels, en la década de los 90, cuando este tipo de estimadores han vuelto a llamar la atención, ya que se ha demostrado que presentan numerosas ventajas sobre los estimadores núcleo más clásicos. En este sentido cabe destacar: Fan (1992, 1993), Fan y Gijbels (1992, 1995), Fan et al. (1997) y Rupert y Wand (1994).

La idea básica del estimador polinómico local es la siguiente: si se supone que existe la derivada de orden $p + 1$ de la función de regresión en un punto $x \in \mathbb{R}$, utilizando un desarrollo Taylor de la función de regresión, se puede aproximar $m(t)$ localmente por un polinomio de grado p ,

$$m(t) \approx \underbrace{m(x)}_{\beta_0(x)} + \underbrace{m'(x)}_{\beta_1(x)}(x - x_0) + \cdots + \underbrace{\frac{m^{(p)}(x)}{p!}}_{\beta_p(x)}(x - x_0)^p. \quad (2.19)$$

Por tanto, se puede estimar el vector $\vec{\beta}(x) = (\beta_0(x), \beta_1(x), \dots, \beta_p(x))^t$, donde $\beta_j(x) = m^{(j)}(x)/(j!)$ con $j = 0, 1, \dots, p$, minimizando la función

$$\Psi(\vec{\beta}(x)) = \sum_{t=1}^n \left(Y_t - \sum_{j=0}^p \beta_j(x)(X_t - x)^j \right)^2 \omega_{n,t}, \quad (2.20)$$

donde $\omega_{n,t} = n^{-1}K_h(x_t - x)$ son los pesos. Si $\hat{\beta}_j(x)$, $j = 0, 1, \dots, p$, denota la solución del problema anterior de mínimos cuadrados locales ponderados, está claro que a partir de la aproximación (2.19), el estimador de $m(x)$ es $\hat{\beta}_0(x)$.

Si el polinomio a ajustar es una constante, es decir $p = 0$, el estimador de la función de regresión obtenido coincide con el de Nadaraya-Watson (2.10). Otro caso interesante es aquel en el que el polinomio a ajustar localmente es una recta, es decir $p = 1$; en este caso el estimador recibe el nombre de estimador local lineal (RLL) y ha sido ampliamente estudiado en los trabajos de Fan y Gijbels antes mencionados. Su expresión analítica es:

$$\hat{m}_{RLL}(x) = \frac{1}{n} \sum_{i=1}^n W_{ni,RLL}(x) Y_i, \quad (2.21)$$

donde

$$W_{ni,RLL}(x) = K_h(X_i - x) \left\{ \frac{s_{2,n}(x) - (X_i - x)s_{1,n}(x)}{s_{0,n}(x)s_{2,n}(x) - s_{1,n}^2(x)} \right\},$$

siendo

$$s_{j,n}(x) = \frac{1}{n} \sum_{t=1}^n (X_t - x)^j K_h(X_t - x), \quad j = 0, 1, 2.$$

Para estimar la función de regresión, el orden del polinomio a ajustar suele tomarse 1, o, en ocasiones, 3. En general, el ajuste de orden impar mejora el ajuste de orden par. Así, el estimador polinómico local lineal funciona mejor que el ajuste constante local, y el de orden 3 proporciona mejores resultados que el de orden 2, en el sentido de que al pasar de un grado par de polinomio al siguiente impar provoca una reducción en el sesgo, sin pérdida en términos de varianza.

Como grado p para el polinomio, Fan y Gijbels (1996) recomiendan que sea impar y exceda en, al menos, una unidad al orden de la derivada a estimar. Así,

si se quiere estimar la función de regresión, bastaría con elegir un polinomio de grado 1, siendo necesario uno de grado 3 para estimar la segunda derivada. En general, si se quiere estimar la derivada de grado j de la función de regresión se tomará el grado p del polinomio de forma que $p - j$ sea 1 ó 3. En general, aumentar el grado del polinomio, para un mismo parámetro de suavizado, provoca una reducción en el sesgo y un aumento en la varianza así como una mayor carga computacional.

Con respecto a la selección de la función núcleo de asignación de pesos (función núcleo K), se puede establecer teóricamente en el contexto de la regresión local la recomendación de utilizar una función de asignación de pesos no negativos, simétrica y con integral la unidad.

Gasser, Müller y Mammitzsch (1985) recomiendan tomar como función núcleo la de Epanechnikov, demostrando que minimiza el error cuadrático medio (MSE) asintótico.

De todos modos, estos aspectos, grado del polinomio y función núcleo, adquieren menor relevancia si se consiguen controlar el sesgo y la varianza mediante una correcta elección del parámetro ventana. En nuestro caso, dedicaremos el capítulo 4 de esta memoria a la selección de la ventana para el ajuste de curvas TGA y DSC.

2.8.1 Ventajas del estimador de regresión polinómico local

Las ventajas de los estimadores polinómicos locales con respecto a otras familias de estimadores están en sus excelentes propiedades teóricas y prácticas. Otros estimadores basados en funciones núcleo, entre los cuales podemos incluir el estimador de Nadayara-Watson o el estimador de Gasser-Müller, presentan inconvenientes, como pueden ser la aparición de un sesgo indeseable y el incremento de la varianza al trabajar con experimentos de diseño aleatorio.

Podemos resumir sus ventajas en los siguientes apartados:

- Este estimador es óptimo en el sentido de minimización asintótica del error. Stone (1980, 1982) estudió las tasas de convergencia en la regresión no paramétrica y demostró que la regresión local es óptima en estos términos. Por otra parte, Fan (1993) estableció resultados sobre la eficiencia del estimador local lineal en sentido minimax.
- Desde un punto de vista meramente computacional, el método sigue presentado unas características verdaderamente atractivas, dada su simplicidad de cálculo en relación a otros estimadores. Siempre es deseable el incremento de la velocidad de los cálculos, lo cual cobra aún mayor interés si éstos son realmente intensos (cálculo de la ventana de estimación o parámetro h). A este respecto, Fan y Marron (1994) establecieron que el método de ajuste polinómico local es, cuando menos, tan rápido computacionalmente como cualquiera de los otros métodos conocidos.
- Ausencia de problemas en los límites del soporte de la función a ajustar, problema conocido como *efecto frontera*, que surge con otro tipo de esti-

madores y consiste en un aumento del sesgo en los puntos extremos del soporte de la función a estimar. El problema se presenta en aquellos casos en los cuales nos encontramos con que, para la estimación de $m^{(j)}(x_0)$, constituyendo x_0 un punto cercano a los límites del intervalo de definición de la función, el vecino local $x_0 \pm h$ (siendo h el parámetro ventana), caiga fuera del ámbito de definición de la función a ajustar. Es por ésto que muchos supuestos válidos para puntos interiores no se pueden aplicar de forma tan inmediata a puntos limítrofes, derivando todo ello en un incremento del sesgo en dichos puntos. Se ha trabajado en el desarrollo de dos aproximaciones que solucionen este problema: métodos basados en funciones núcleo específicas y métodos de reflexión. Pero ninguno de estos métodos alcanza la simplicidad y eficiencia de la corrección automática del efecto frontera proporcionada por el ajuste polinómico local. El método polinómico local evita tener que usar funciones núcleo específicamente diseñadas para evitar esta situación, comportándose el sesgo de igual modo en puntos interiores del soporte que en puntos de la frontera.

- Por último, destacar que la técnica de estimación polinómico local no sólo permite obtener estimadores de la función de regresión, sino también de sus derivadas. Así, teniendo en cuenta (2.19), un estimador para $m^{(j)}(x)$ sería $(j!) \hat{\beta}_j(x)$. Esto contrasta con las versiones más complicadas de estimadores de las derivadas de la función de regresión basadas en los estimadores de NW o GM.

Este estimador también presenta pequeños inconvenientes, como los apuntados en Seifert y Gasser (1996), donde se comenta que la varianza condicional de este estimador se va a infinito cuando en el entorno de puntos en que se realiza el promedio ponderado no hay al menos $p + 3$ puntos; incluso con ese número de datos no se puede asegurar que dicha varianza esté acotada. Es decir, en diseños con pocos datos, o datos agrupados en ciertas zonas del diseño, la varianza del estimador se hace muy grande. Estos autores resuelven este problema presentando dos modificaciones de los polinomios locales: un incremento local de la ventana en las regiones con pocos datos del diseño, y el uso de estimadores de tipo contraído (ridge), combinados con las técnicas polinómico locales.

2.8.2 Desarrollo matricial del modelo

A continuación incluimos una versión matricial del modelo, aproximación que utilizaremos para la implementación computacional.

Denotaremos por:

$$X = \begin{pmatrix} 1 & (X_1 - x_0) & \dots & (X_1 - x_0)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (X_n - x_0) & \dots & (X_n - x_0)^p \end{pmatrix},$$

$$Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \quad \hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_p \end{pmatrix},$$

y finalmente llamemos W a la matriz diagonal de dimensiones $n \times n$ que proporciona los pesos asignados a cada punto:

$$W = \text{diag}\{K_h(X_i - x_0)\}.$$

A partir de dicha notación matricial, el problema de mínimos cuadrados ponderados puede ser expresado del siguiente modo:

$$\min_{\beta} (Y - X\beta)^t W (Y - X\beta), \quad \text{con } \beta = (\beta_0, \dots, \beta_p)^t.$$

El vector solución puede obtenerse a través del siguiente cálculo matricial:

$$\hat{\beta} = (X^t W X)^{-1} X^t W Y \quad (2.22)$$

2.9 Estimación con datos dependientes

Cuando los datos presentan dependencia, como es el caso de curvas procedentes de análisis térmico, se pueden adaptar los resultados obtenidos en independencia, estimando previamente el grado de autocorrelación de las observaciones.

El problema que aparece en estos casos es que si realmente los datos son correlados y esta correlación no se tiene en cuenta (es decir, se tratan los datos como si fueran independientes), las estimaciones no paramétricas de la función de regresión pueden ser bastante pobres; en estos casos las ventanas obtenidas pueden diferir mucho de las mejores posibles, puesto que la correlación existente entre las observaciones puede ser interpretada como parte de la curva de regresión. Por ejemplo, si la correlación fuera positiva y no se tuviera en cuenta este hecho, las ventanas obtenidas seguramente serán bastante más pequeñas de lo que deberían de ser, con lo que la estimación no paramétrica de la regresión será una curva muy *variable*.

Se puede considerar que el modelo de regresión es del tipo:

$$Y_{i,n} = m(X_{i,n}) + \varepsilon_{i,n}, \quad i = 1, 2, \dots, n,$$

donde los errores, $\{\varepsilon_{in}\}$, están correlados y los puntos del diseño, $\{X_{i,n}\}$, pueden ser fijos o aleatorios.

En Hart y Vieu (1990) se estudia un método de selección del parámetro de suavizado para el estimador NW, bajo la suposición de dependencia de tipo α -mixing. En Quintela (1992) se realiza un estudio somero sobre el cálculo del parámetro de suavización bajo distintas condiciones de dependencia.

En relación al estimador polinómico local, Härdle (1986) y Härdle y Tsybakov (1997), estiman la media condicional y la varianza condicional en series de tiempo. Masry y Fan (1997) obtienen propiedades asintóticas de sesgo, varianza y normalidad del estimador polinómico local suponiendo que la muestra $\{(X_i, Y_i)\}$ es de tipo mixing. Bajo estas mismas hipótesis, en Masry (1996a) se generalizan los resultados al caso multivariante, y en Masry (1996b) se obtienen propiedades de consistencia fuerte.

Roussas (1989, 1990) y Roussas et al. (1992) estudian propiedades de consistencia y normalidad asintótica de suavizadores no paramétricos, en el caso de diseño fijo y errores de tipo mixing.

Vilar-Fernández y Vilar-Fernández (1998, 2000) obtienen propiedades asintóticas para una versión recursiva del estimador polinómico local, y estudian propiedades de consistencia para este estimador bajo dependencia.

Finalmente, mencionar la memoria correspondiente a la tesis doctoral del profesor Francisco-Fernández (2001), como un referente sobre la estimación polinómica local en diseño fijo con dependencia.

Capítulo 3

Estimación paramétrica de curvas TGA

3.1 Estudio cinético

El estudio cinético de los materiales por termogravimetría dinámica se basa en llevar a cabo reacciones de descomposición térmica a temperatura variable, con el fin de analizar la velocidad a la que el material pierde peso. La pérdida de peso de la muestra puede considerarse como la suma de diferentes procesos de descomposición isotérmicos, como queda de manifiesto en textos clásicos como el de Turi (1997).

Uno de los objetivos del estudio cinético es conocer si existen uno o varios procesos y cuáles son las constantes de reacción características de éstos. La utilidad industrial de estos estudios es el aprovechamiento de los productos químicos que se puedan generar, el aprovechamiento energético, o simplemente, la eliminación de determinados materiales.

Al aumentar la velocidad de calefacción existe un desplazamiento de las curvas TGA que ha sido ampliamente estudiado en la literatura (véase, por ejemplo, Turi (1997)). Hay diversos argumentos para explicar estos desplazamientos. Algunos autores consideran que este desplazamiento puede ser explicado por las expresiones matemáticas de las leyes de la cinética. Obviamente, un modelo matemático que represente un conjunto de experimentos debe ser capaz de explicar tales desplazamientos a varias velocidades de calefacción (Conesa (2000)). Sin embargo, es posible ajustar las mismas curvas TGA a diferentes expresiones, dependiendo del modelo cinético elegido, lo cual representa un serio problema.

En principio, los métodos cinéticos suelen clasificarse en endotermos e isotermos, con una o con varias velocidades de calentamiento. También cabe la clasificación en métodos derivativos o integrales, dependiendo del procedimiento que se emplee para el cálculo de los parámetros.

La utilización de la derivada de la curva TGA (DTG), en lugar de la propia

curva, tiene la ventaja de que un pequeño cambio en la TGA es aumentado en su derivada. Otras ventajas de la termogravimetría dinámica, con respecto al estudio en procesos isotermos, según puede verse en trabajos como el de Carrasco y Costa (1989), son las siguientes:

- Tiempos de manipulación inferiores permitiendo, con un solo experimento, determinar los parámetros cinéticos.
- Sólo se utiliza una muestra, con lo cual se evitan los problemas inherentes al cambio de muestra necesarios en los procesos isotérmicos.
- La cinética global aparente puede determinarse en una gran gama de temperaturas de forma continua.
- En los procesos isotermos se requiere un precalentamiento previo, antes de alcanzarse la temperatura de reacción. Durante este período transitorio pueden obtenerse conversiones no nulas, de tal forma que se falsearía el análisis cinético.
- Es posible determinar con precisión la temperatura umbral de descomposición.

Uno de los problemas de los métodos empleados para el estudio cinético es la dificultad de ajustar los datos al modelo usado, normalmente de tipo Arrhenius.

3.2 Modelos cinéticos tipo Arrhenius

La forma clásica de estudiar la descomposición térmica de la muestra se basa en suponer que la velocidad de pérdida de peso de las reacciones de descomposición térmica dependen de la masa y de la temperatura, cuya expresión general sería:

$$m' = dm/dt = -f(m)k(T)q(m, T). \quad (3.1)$$

En la expresión (3.1) la función $f(m)$ se toma como m_t^n , lo que comúnmente se interpreta diciendo que la pérdida de peso de la muestra obedece a una *cinética de orden n*, donde m_t es la masa de la muestra en el tiempo t .

Esta suposición del orden de reacción es aplicable a reacciones homogéneas y a algunas reacciones heterogéneas que se llevan a cabo en fase condensada. Sin embargo, algunos polímeros que se degradan siguiendo un mecanismo de escisión aleatoria de enlaces, no poseen un orden de reacción constante, de tal forma que sería erróneo el uso de esta expresión (3.1), dado que los parámetros estarían sobreestimados (véase Conesa (2000)).

En todo caso, la hipótesis fundamental de la relación (3.1) se basa en suponer que la pérdida de masa sigue la Ley de Arrhenius representada por la ecuación siguiente:

$$k(T) = A \exp\left(-\frac{E_a}{RT}\right). \quad (3.2)$$

La ecuación de Arrhenius (3.2) expresa la variación de la velocidad de reacción (k) con la temperatura (T), y parte de la suposición de que la velocidad de descomposición del proceso aumenta exponencialmente con la temperatura. La constante A , llamada constante preexponencial, es independiente de la temperatura; E_a es la denominada energía de activación y R es la constante de los gases y proviene del papel que juega en la ley de los gases perfectos ($pV = nRT$). El nombre se debe a que, en principio, esta ecuación fue propuesta por el químico sueco Svante August Arrhenius para la velocidad de reacción química en los gases.

Finalmente, el término $q(m, T)$ de la expresión (3.1) indica la interacción entre la masa y la temperatura, y, aunque en muchos procesos se considera igual a uno, en ocasiones no se puede hacer esta suposición y su estimación puede resultar complicada. Estas interacciones entre masa y temperatura pueden estar originadas por los siguientes fenómenos (Carrasco y Costa (1989)):

- Modificación de las propiedades físicas de la muestra durante la descomposición térmica.
- Existencia de reacciones competitivas debidas a que la energía de activación y el factor preexponencial puedan variar con la conversión.

Existe algún otro método alternativo a la ley de Arrhenius, como el que se basa en suponer que la energía de activación E_a no es constante, sino que puede representarse por una distribución de probabilidad normal, como el presentado por Carrasco y Costa (1989).

3.2.1 Cálculo de los parámetros de la ecuación de Arrhenius

Con el objeto de exponer los problemas que presenta el cálculo de los parámetros cinéticos, haremos a continuación una revisión de los modelos más utilizados para este estudio.

Para la determinación de los distintos parámetros existentes en la ecuación de Arrhenius (3.2), es necesario llevar a cabo una serie de transformaciones que se basan en expresar dicha ecuación en forma logarítmica:

$$\ln(k(T)) = \ln A - \frac{E_a}{R} \frac{1}{T}. \quad (3.3)$$

Esta ecuación (3.3) permite, mediante una representación de $\ln(k(T))$ frente al inverso de la temperatura absoluta ($\frac{1}{T}$), calcular los parámetros deseados. En la práctica, basta con calcular la recta de regresión de $\ln(k(T))$ frente al

inverso de la temperatura absoluta ($\frac{1}{T}$) para, al menos, dos experimentos (hay que determinar dos constantes de la recta). La pendiente de dicha recta es la estimación del cociente $-\frac{E_a}{R}$, mientras que la ordenada en el origen es una estimación de $\ln A$.

Habitualmente se usa este método para la determinación de la energía de activación a partir de la pendiente de la recta de regresión. Este hecho tiene el fundamento teórico en que la energía de activación, que tiene unidades de energía por mol, es posible expresarla mediante la propia ecuación de Arrhenius, sin más que dividir E_a y R entre el número de Avogrado (N_{AV}), obteniéndose:

$$k(T) = C \exp(-q/kT), \quad (3.4)$$

donde $q = E_a/N_{AV}$ es la energía de activación por unidad de escala atómica y $k = R/N_{AV}$ es la constante de Boltzmann ($13.8 \times 10^{24} \text{ J/K}$).

Además, la expresión (3.4) permite realizar una comparación con el extremo de alta energía de la distribución de Maxwell-Boltzmann de energías moleculares de los gases, dada por:

$$p \propto \exp\left(-\frac{\Delta E}{kT}\right). \quad (3.5)$$

La ecuación de Maxwell-Boltzmann (3.5) expresa la probabilidad p de encontrar una molécula con una energía ΔE mayor que la energía media característica correspondiente a una temperatura T . Este hecho es el que permite interpretar la energía de activación E_a como la barrera energética que debe de ser superada mediante la activación térmica. Aunque la ecuación (3.5) se desarrolló inicialmente para gases, también es aplicable a sólidos.

3.2.2 Cálculo del orden n de la reacción

En la expresión (3.1), una de las suposiciones más simples para la función f es suponer que $f(m) = m_t^n$, cinética de orden n . Para el cálculo del orden de reacción, n , se supone que la masa en el instante t , m_t , está relacionada con la conversión mediante la ecuación:

$$m_t = m_0(1 - \alpha). \quad (3.6)$$

Existen diferentes formas lineales para la expresión de la función f , destacando los métodos de Freeman y Carol, Newkirk, Horowitz y Metzger, Coats y Redfern, Vachuska y Voboril y Friedman (Turi (2000)). Conesa (2000) enumera los siguientes inconvenientes de estos métodos:

- Se manipulan muchos datos obtenidos en el equipo, por lo que el reiterado abuso de logaritmos de logaritmos puede esconder los verdaderos datos de la señal TGA.

- Estos métodos son válidos para descomposiciones que tengan lugar en un proceso simple. Si se observa más de un proceso, es necesario realizar una deconvolución "visual" antes de aplicar el método.
- En lugar de la relación experimental $T = f(t)$, se utiliza la velocidad de calefacción programada $T = T_0 + \beta t$, donde β es la velocidad real de calefacción: este hecho produce situaciones alejadas de la realidad, por cuanto que la temperatura experimental (leída por el termopar) puede desviarse puntualmente de la programada.
- La mayoría de estos métodos (todos los citados anteriormente excepto el de Friedman) utilizan sólo los datos obtenidos a una velocidad de calefacción.

3.2.3 Método de Friedman

A partir de experimentos realizados a varias velocidades de calefacción, el método de Friedman (1964) se basa en la representación del valor $\log \frac{d\alpha}{dt}$ frente a $1/T$ para diferentes valores de α ; se buscará el mejor ajuste lineal, siendo la pendiente de la recta una estimación de E_a/R .

Cuando el análisis termogravimétrico se lleva a cabo en condiciones de programación lineal con la temperatura, esta última y el tiempo de reacción están vinculados por la velocidad de calentamiento:

$$\beta = \frac{dT}{dt}. \quad (3.7)$$

Puesto que la masa en el instante t , m_t , está relacionada unívocamente con la conversión, se sigue que:

$$f(m) = h(\alpha)m_0^n. \quad (3.8)$$

Teniendo en cuenta las ecuaciones (3.1, 3.2, 3.6 y 3.7), y previa separación de variables, se obtiene la expresión siguiente:

$$\int_{\alpha=0}^{\alpha} \frac{d\alpha}{h(\alpha)} = \frac{Am_0^{n-1}}{\beta} \int_{T(\alpha=0)}^T \exp\left(\frac{-E_a}{RT}\right) dT. \quad (3.9)$$

En la ecuación (3.9) la integral no posee primitiva analítica. Utilizando el método numérico de aproximación de Coats-Redfern, basado en la expresión truncada del desarrollo en serie de Taylor, se obtiene como resultado:

$$\int_{T(\alpha=0)}^T \exp\left(\frac{-E_a}{RT}\right) dT = \frac{RT^2}{E_a} \left(1 - \frac{2RT}{E_a}\right) \exp\left(\frac{-E_a}{RT}\right). \quad (3.10)$$

Además, la integral que depende de la conversión en la ecuación (3.9), tiene una primitiva que puede expresarse como:

$$\int_{\alpha=0}^{\alpha} \frac{d\alpha}{h(\alpha)} = g(\alpha). \quad (3.11)$$

Introduciendo estos resultados (3.10 y 3.11) en la expresión (3.9), y linealizando la ecuación resultante, se llega a la expresión final siguiente:

$$\log \left[\frac{g(\alpha)}{T^2(1 - 2RT/E_a)} \right] = \log \left[\frac{A^*R}{\beta E_a} \right] - \frac{E_a}{R} \frac{1}{T},$$

donde $A^* = Am_0^{n-1}$.

Finalmente, mediante regresión lineal, como se expuso anteriormente, puede calcularse la energía de activación y el factor preexponencial a partir de la pendiente y de la ordenada en el origen de dicha recta, respectivamente. Cabe resaltar que para conocer la ordenada se necesita conocer la energía de activación, lo que precisa de un método iterativo.

Además, la linealización de la ecuación cinética engendra errores propios de la estimación de la recta de regresión. Una forma de atenuar dichos errores es usar un método de regresión lineal ponderada, en el cual se aplica un factor o peso multiplicativo a los puntos experimentales.

Autores como Várhegyi (1987) afirman que los parámetros cinéticos son muy sensibles a los datos del principio y del final de la curva TGA, mientras que su sensibilidad es mínima en las proximidades de la zona de máxima velocidad de reacción lo que justificaría utilizar pesos en el ajuste lineal ponderado del tipo de:

$$w = \left(\frac{\partial y}{\partial \alpha} \right)_T^{-2} = \left(\frac{g(\alpha)}{g'(\alpha)} \right)^2. \quad (3.12)$$

En cualquier caso, siempre puede ocurrir que dentro del material puedan tener lugar varios mecanismos simultáneamente, y cada mecanismo podría tener una energía de activación E_a característica, por lo que resulta difícil encontrar los parámetros para este tipo de modelos basados en la ley de Arrhenius.

Los diferentes métodos para el cálculo de dichos parámetros necesitan de aproximaciones numéricas, para integrales o derivadas, lo cual permite clasificar estos métodos en integrativos y derivativos.

3.2.4 Otros modelos propuestos para la cinética

Existen otros modelos para la cinética que partiendo de la ecuación de Arrhenius llegan a distintas expresiones. A continuación expondremos, someramente, algunos de los más resaltables.

La mayoría de estas modificaciones de la ecuación general (3.1) pueden obtenerse a partir del modelo propuesto por Sesták y Berggren (1971), cuya forma general sigue la ecuación:

$$\frac{d\alpha}{dt} = k(\alpha)^m(1-\alpha)^n[-\ln(1-\alpha)]^p, \quad (3.13)$$

donde n , m y p son constantes. Obsérvese que para $m = p = 0$ se reduce a la expresión vista anteriormente para el cálculo del orden de reacción.

Braun y Burnham (1987) introducen un nuevo parámetro q , en la expresión (3.13) obteniendo una nueva variante, dada por:

$$\frac{d\alpha}{dt} = k((1-q)(1-\alpha))^m(1-\alpha)^n \quad (3.14)$$

Para el cálculo de los parámetros de las ecuaciones de los distintos modelos es necesario estimar valores mediante una regresión lineal. Así, hay autores que estiman estos números mediante el ajuste de $\log(\alpha)$ frente a $1/T$, como el caso de Ozawa (1965), y en otros casos se opta por la estimación de estos parámetros mediante el método Kissinger, que se basa en obtener el ajuste mediante regresión lineal de $\log(\beta/T_{\max}^2)$ frente a $1/T_{\max}$.

Otros modelos, de tipo derivativo, como los basados en la ecuación de Freeman y Carroll (1958), asumen la expresión $f(\alpha) = (1-\alpha)^n$, y buscan relacionar $1-\alpha$ con $1/T$ mediante la ecuación:

$$\Delta \ln \left(\frac{d\alpha}{dT} \right) = n \Delta \ln(1-\alpha) - \left(\frac{E_a}{RT} \right) \Delta \left(\frac{1}{T} \right). \quad (3.15)$$

3.3 Métodos numéricos en análisis cinético

A partir de 1985 se extiende el empleo de los denominados métodos numéricos de análisis cinético, auspiciados sobre todo por la aparición de programas informáticos acoplados a calorímetros. Estos métodos emplean algoritmos numéricos como los de Euler o el de Runge-Kutta (Turi (2000)).

La idea de estas aproximaciones numéricas parte de una expresión para $f(m)$, y se integran las ecuaciones para unos determinados valores de los parámetros cinéticos; se compara la curva calculada con la experimental mediante una función objetivo, y se calculan aquellos valores que optimizan dicha función objetivo mediante el empleo de métodos como el *simplex flexible*.

El criterio más utilizado es el método de mínimos cuadrados, siendo sus expresiones más habituales:

$$\min \sum_i \sum_j (m_{\text{cal}(i,j)} - m_{\text{exp}(i,j)})^2, \quad (3.16)$$

$$\min \sum_i f_i \sum_j \left(\frac{dm_{\text{cal}(i,j)}}{dt} - \frac{m_{\text{exp}(i,j)}}{dt} \right)^2. \quad (3.17)$$

Al primer método (3.16) se le denomina integral y al segundo método (3.17) diferencial. En ambas expresiones $m_{cal(i,j)}$ y $m_{exp(i,j)}$ representan los valores calculados con los diferentes modelos cinéticos y los valores experimentales de la masa de la muestra. El factor f_i es un parámetro de escala, incluido para asegurar que todos los experimentos tienen el mismo peso en la función objetivo.

Para obtener el mínimo de las expresiones anteriores (3.16) y (3.17) suele emplearse el método del simplex flexible. Se trata de una técnica de optimización para funciones no lineales desarrollada por Nedler y Mead en 1965, que no necesita utilizar las derivadas para el cálculo del mínimo, sino que evalúa la función en cada uno de los $n + 1$ vértices de un hiperpoliedro, y va eliminando aquellos vértices que dan valores más altos, la iteración de este proceso permite obtener el mínimo buscado.

Existen también propuestas que no se ajustan al modelo inicial propuesto en (3.1). En este sentido cabe mencionar los *métodos cinéticos de tipo no paramétrico* (NPK) propuestos por Nomen y Sampere. Este método se basa en suponer que la reacción se puede expresar mediante el producto de dos funciones independientes, $f(\alpha)$ y $f'(T)$. La función $f(\alpha)$ representa la cinética y $f'(T)$ la dependencia de la temperatura. El modelo se expresaría mediante una matriz A , de rango $n \times m$ con $A_{i,j} = f(\alpha_i)f'(T_j)$, según la expresión $A = ab^t$, siendo los vectores $a = \{f(\alpha_1), f(\alpha_2), \dots, f(\alpha_n)\}^t$ y $b = \{f'(T_1), f'(T_2), \dots, f'(T_m)\}^t$. Este método tiene dos inconvenientes: por un lado necesita un número muy elevado de cálculos para la estimación de los parámetros mediante algoritmos de aproximación y, por otra parte, tiene que tomar varios experimentos a distintas temperaturas T_1, T_2, \dots, T_n para la determinación de un modelo.

Finalmente, destacar el modelo propuesto por Carrasco y Costa (1989) para el caso del estudio cinético del poliestireno. Si bien realizan un estudio estadístico detallado del ajuste de la ecuación mediante contrastes de hipótesis, modelo ANOVA, parten de la misma idea basada en la ecuación de Arrhenius, lo que hace un modelo difícil de calcular, con una metodología laboriosa.

3.3.1 Inconvenientes de los métodos actuales

Los distintos métodos mencionados anteriormente tienen como gran inconveniente proporcionar resultados diferentes para los parámetros de un mismo modelo. Este hecho fue manifestado por varios científicos, como el caso de Vyazovkin (1996), como el mayor inconveniente en la aplicación de métodos tipo Arrhenius.

Este problema se ha puesto de manifiesto en diversos artículos, pero fue sin duda en el congreso internacional de la ICTAC (Confederación Internacional de Análisis Térmico y Calorimetría) celebrado en Filadelfia, USA, en agosto de 1996, donde se ha dado a conocer a la comunidad científica. En este foro, se puso en evidencia la diferencia en el ajuste de modelos cinéticos a una misma muestra dependiendo del método usado. Se propuso la aplicación a diferentes experimentos con carbonato de calcio (CC) a participantes invitados, y los resultados obtenidos se recogen en una serie de artículos de Brown et al. (2000) en la revista *Thermochemica Acta*. En ellos se constatan las grandes diferen-

cias entre los valores obtenidos, tanto para la energía de activación (E_a) como para la constante preexponencial A . Cabe resaltar que en estos trabajos participaron también los científicos Nomen y Sampere, quienes aplicaron a los datos su método no paramétrico, resultando estimaciones con grandes diferencias con respecto a métodos más clásicos tipo Freedman, Ozawa, Kissinger y otros. Otros resultados similares pueden verse en Arnold et al. (1982) y en Doyle (1961).

Las diferencias entre los parámetros encontrados por cada autor para un mismo experimento corroboran la poca eficacia de los modelos existentes para este tipo de ajustes, lo que nos llevó a la búsqueda de un modelo alternativo que será expuesto en este capítulo.

3.4 Modelo paramétrico con mezcla de logísticas

En esta sección proponemos descomponer la curva TGA en diferentes funciones logísticas (modelo con mezcla de logísticas), basándonos en la idea de que cada una representa la cinética de descomposición de los distintos materiales simples de los que está compuesta la muestra inicial. Proponemos la siguiente expresión para modelizar las pérdidas de peso del material $Y(t)$, con el tiempo t :

$$Y(t) = \sum_{i=1}^k w_i f(a_i + b_i t), \quad (3.18)$$

$$f(t) = \frac{\exp t}{1 + \exp t},$$

donde w_i son parámetros que representan pérdidas de peso en cada proceso; a_i son parámetros de localización y b_i representa la velocidad de pérdida de peso, para los k escalones que presente la función $Y(t)$.

La idea inicial parte de buscar una función matemática que se ajuste lo mejor posible a los datos reales obtenidos para las curvas termogravimétricas. Las funciones candidatas a estimar estas curvas ($t, (Y(t))$), deben de verificar que para valores grandes de t ($t \rightarrow \infty$) las respuestas ($Y(t)$) deben tender a 0, lo que obliga a que los parámetros b_i tienen que ser negativos, mientras que al inicio del proceso ($t = 0$) la función debe tender al peso de la muestra en cada momento; por tanto, los valores w_i representan, aproximadamente, las pérdidas de masa en el i -ésimo proceso. Los distintos procesos estarían marcados por los "escalones" de cada experimento.

Para una primera aproximación al método propuesto veamos un caso de ajuste a un material, como el anhídrido hexahidroftálico, que sólo presenta un escalón, lo que simplifica la expresión (3.18) (ver figura 3.1). Supondremos, sin pérdida de generalidad, que el peso de la muestra es la unidad ($w = 1$), y que $a = 12$ y $b = -4$. Las expresiones de este modelo y de sus derivadas son las siguientes:

$$\begin{aligned}
 Y(t) &= \frac{e^{12-4t}}{1 + e^{12-4t}} & (3.19) \\
 Y'(t) &= -4 \frac{e^{12-4t}}{(1 + e^{12-4t})^2} \\
 Y''(t) &= -16e^{12-4t} \frac{-1 + e^{12-4t}}{(1 + e^{12-4t})^3}
 \end{aligned}$$

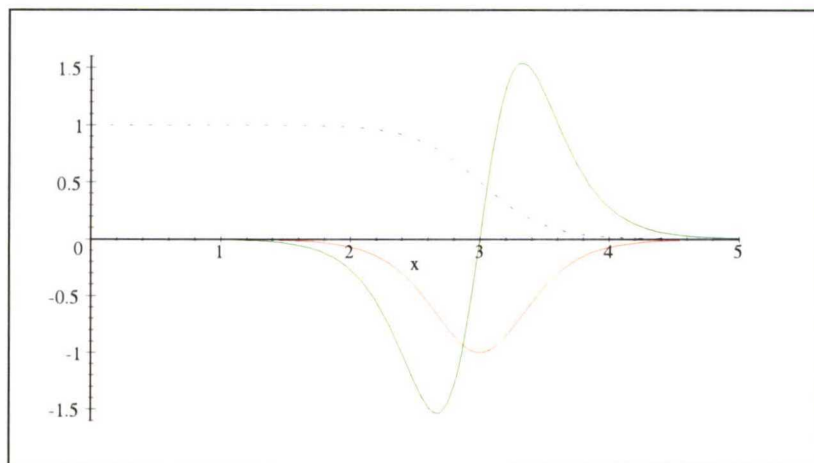


Figura 3.1: Modelo logístico para un sólo escalón con su primera derivada (rojo) y su segunda derivada (verde).

En el supuesto, más general, de materiales con varios escalones será necesario un ajuste a cada trozo de curva según la expresión (3.18). Partiendo de una muestra de PVC, que presenta 4 escalones (por tanto $k = 4$), una primera modelización, observando la nube de puntos, nos lleva a un modelo del tipo:

$$Y(t) = 5 \frac{e^{12-4t}}{1 + e^{12-4t}} + 4 \frac{e^{14-2t}}{1 + e^{14-2t}} + 7 \frac{e^{43-4t}}{1 + e^{43-5t}} + \frac{e^{16-t}}{1 + e^{16-t}} \quad (3.20)$$

En la figura 3.2 se puede observar el modelo de cuatro componentes (3.20), donde se representan las distintas funciones logísticas $Y_i(t)_{i=1}^4$ con distintos colores (granate, rojo, verde y azul) correspondientes a cada una de las 4 ecuaciones logísticas.

Una vez estimados los parámetros de la curva TGA, resulta inmediato el cálculo de las distintas derivadas y, por tanto, de sus gráficas.

En las figuras 3.3 y 3.4 podemos observar la estimación de la curva TGA y de su primera y segunda derivada.

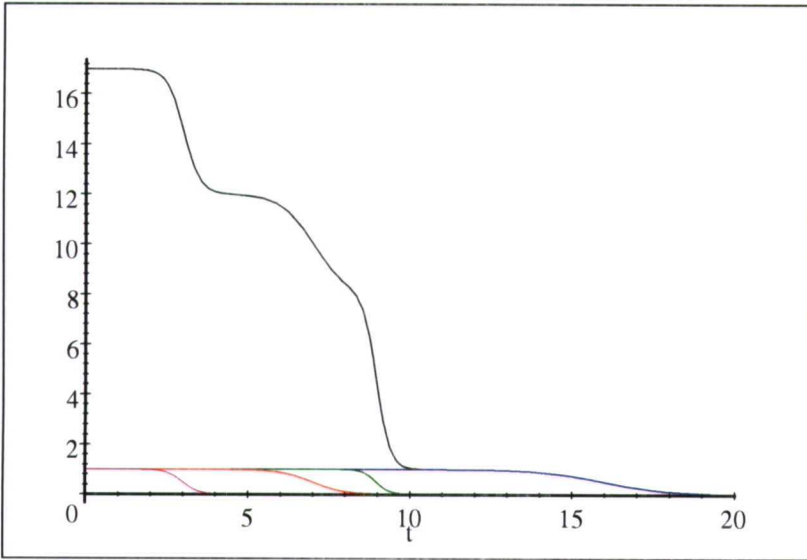


Figura 3.2: Modelo logístico para el PVC y sus cuatro componentes.

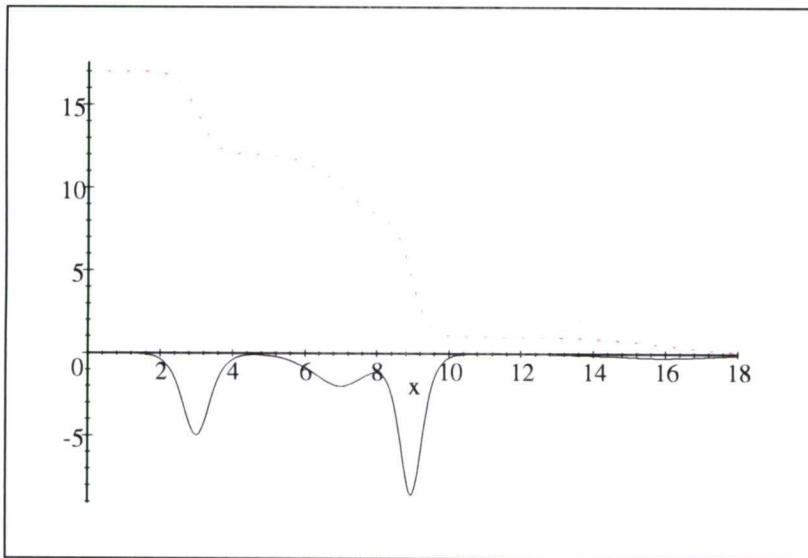


Figura 3.3: Estimación de la TGA (en rojo) y de su primera derivada (en negro).

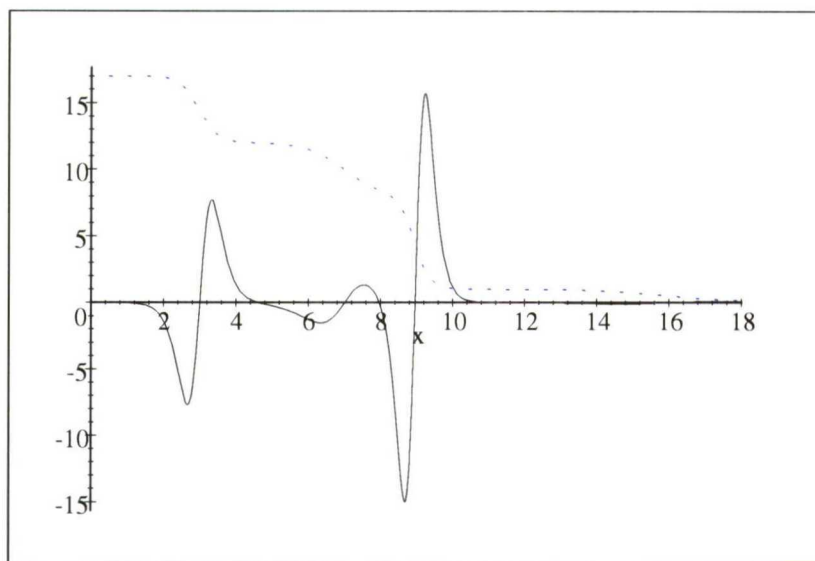


Figura 3.4: Estimación de la curva TGA (en azul) y de su segunda derivada (en negro).

3.4.1 Estimación de los parámetros del modelo

Para el ajuste de los datos a una curva logística es necesario calcular los valores de los parámetros que intervienen en la ecuación, y para ello hacemos uso de algún programa estadístico, en nuestro caso el S-plus y, en particular, usaremos el módulo de regresión no lineal.

El algoritmo de regresión no lineal para estimar los parámetros aplica el método de mínimos cuadrados no lineales. Los fundamentos teóricos de este procedimiento, pueden verse en Gay (1984), y se basan en el método de Levenberg-Marquart para generar la secuencia de aproximación al punto mínimo basada en el algoritmo de la “trust region”, siendo discutido en los libros de Chambers y Hastie (1992) y en Dennis et al. (1981).

Uno de los problemas a la hora del ajuste es partir de puntos iniciales para los distintos parámetros a estimar. Para este fin puede observarse la propia curva TGA, e intentar estimar el punto en que presenta cambios de curvatura (escalones). Con el fin de encontrar unos parámetros iniciales para el algoritmo de aproximación haremos uso de la función logit ($\text{logit}(u) = \log \frac{u}{1-u}$). La idea es ajustar una recta a los puntos $\text{logit}(Y_i(t)/w_i)$, para cada tramo de valores en que la función tiene escalones. La ordenada en el origen de este ajuste lineal es el valor para iniciar el algoritmo para estimar el parámetro buscado (a_i) y la pendiente de dicho ajuste lineal será el valor para iniciar el algoritmo para estimar el parámetro b_i . El motivo de este ajuste lineal se explica mediante la expresión (3.21):

$$\begin{aligned}
 Y_i(t) &= w_i f(a_i + b_i t) = w_i \frac{\exp(a_i + b_i t)}{1 + \exp(a_i + b_i t)}, & (3.21) \\
 \frac{\frac{Y_i(t)}{w_i}}{1 - \frac{Y_i(t)}{w_i}} &= \frac{Y_i(t)}{w_i - Y_i(t)} = \exp(a_i + b_i t), \\
 \text{logit} \left(\frac{Y_i(t)}{w_i} \right) &= \log \left(\frac{\frac{Y_i(t)}{w_i}}{1 - \frac{Y_i(t)}{w_i}} \right) = a_i + b_i t.
 \end{aligned}$$

3.5 Aplicaciones del método de ajuste logístico

Vamos a tratar diferentes tipos de materiales de cara a validar el método propuesto en la sección anterior a distintos casos. Primeramente ajustaremos el modelo propuesto a una muestra de anhídrido hexahidroftálico que presenta un solo escalón. A continuación, se usará para el ajuste a muestras de madera (ciprés y eucalipto), y finalmente ajustaremos los datos reales de un experimento con PVC, curva que hemos usado para introducir el modelo (3.2) y que presenta 4 escalones.

3.5.1 Ajuste de un modelo logístico para el anhídrido hexahidroftálico

En el caso de tratarse de una muestra de hexahidroftálico, puesto que este tipo de material presenta una curva TGA de un solo escalón, se necesitará una sola función logística para modelizar los datos del experimento. Es decir, será suficiente con una ecuación del tipo (3.19).

Una vez representada la curva TGA ajustamos la ecuación (3.18) a los datos de este experimento obteniendo el ajuste lineal, por mínimos cuadrados, que nos proporcionan los parámetros a_1 y b_1 .

En el caso del hexahidroftálico resulta la siguiente expresión

$$Y(t) = 12.93 \frac{\exp(17.48 - 0.024t)}{1 + \exp(17.48 - 0.024t)}. \quad (3.22)$$

La representación del modelo (3.22) puede verse en la gráfica 3.5, en la que se aprecian los datos reales de la curva TGA (en rojo) y el ajuste logístico (en azul).

3.5.2 Ajuste para una muestra de madera de ciprés

Ajustaremos un modelo logístico con dos componentes para una curva TGA correspondiente a una muestra de madera de ciprés.

En primer lugar observando la curva de la figura 3.7 descubrimos los distintos escalones que marcarán los subintervalos para el ajuste lineal de la función logit.

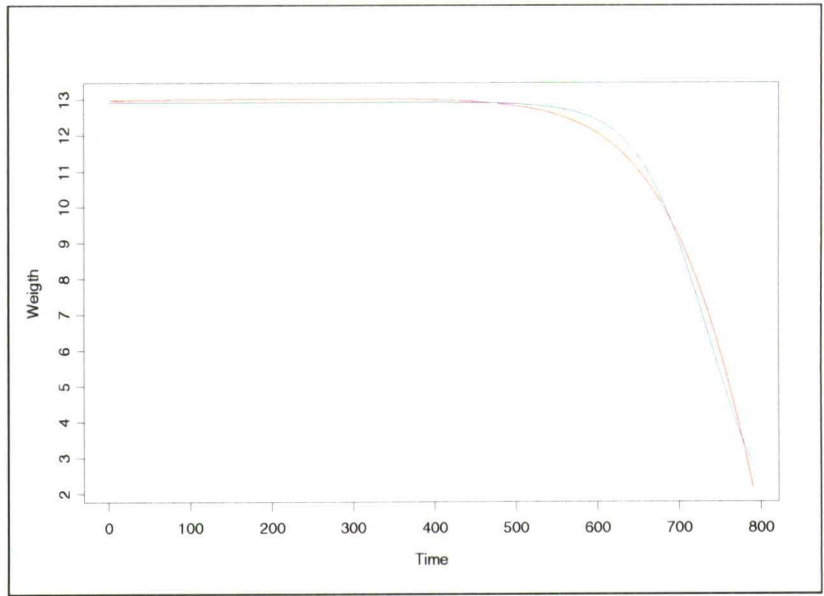


Figura 3.5: Ajuste logístico (azul) de los datos de hexahidroftálico (rojo).

En este caso, el primer escalón abarca los tiempos de 0 a 700, y, para hacer el ajuste lineal, tendremos en cuenta que en valores próximos al 0 y a 700 pueden resultar valores con el logaritmo de 0 por lo que eliminamos 10 puntos en cada lado del intervalo. Ajustamos linealmente una recta entre 10 y 690 resultando: $w_1 = 8.5$, $a_1 = 5.12$ y $b_1 = 0.012$ (ver figura 3.6). Estos serán los valores que se utilizarán para iniciar el modelo.

Se procede de igual forma para el siguiente intervalo [700:1640], en donde presenta el otro escalón, resultando como valores del ajuste lineal de la función logit $a_2 = 9.175879$, $b_2 = -0.004551135$. Los valores w_1 y w_2 corresponden a las diferencias entre el máximo valor y el mínimo valor de $Y(t)$ en cada subintervalo.

Finalmente, con estos valores iniciales se ajusta el modelo (6.1) resultando:

	$i = 1$	$i = 2$
w_i	10.5352	89.905
a_i	3.791	12.636
b_i	-0.00765	-0.0057

Tabla 3.1: Parámetros del modelo mezcla de logísticas para el ciprés.

La gráfica correspondiente al ajuste logístico dado en la tabla 3.1, que se obtuvo a partir de los datos del ciprés, se puede observar en la figura 3.7.

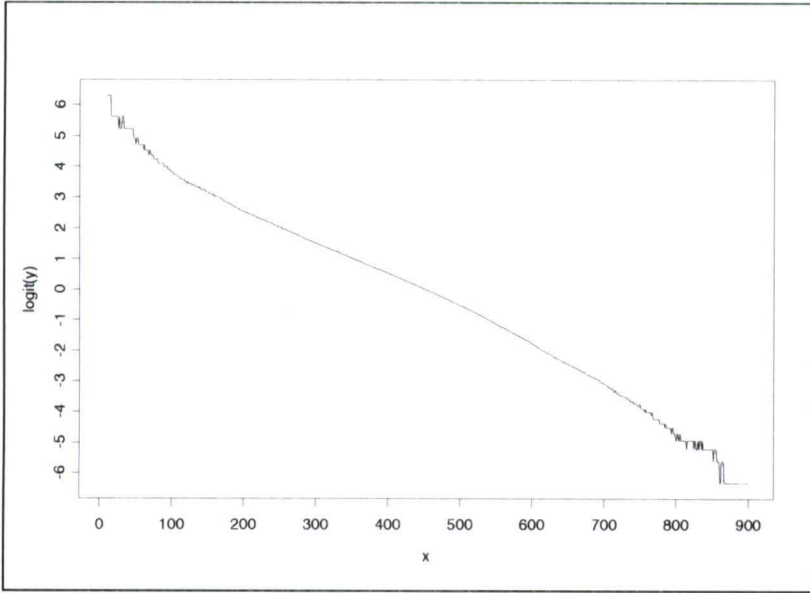


Figura 3.6: Ajuste lineal en el intervalo $[0,700]$.

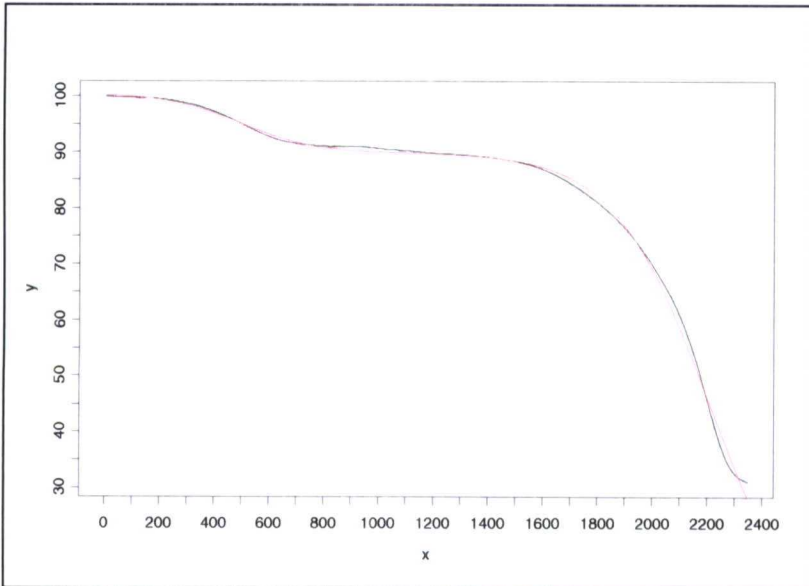


Figura 3.7: Ajuste logístico (en rojo) para los datos de ciprés (en negro), donde x es el tiempo e y el peso.

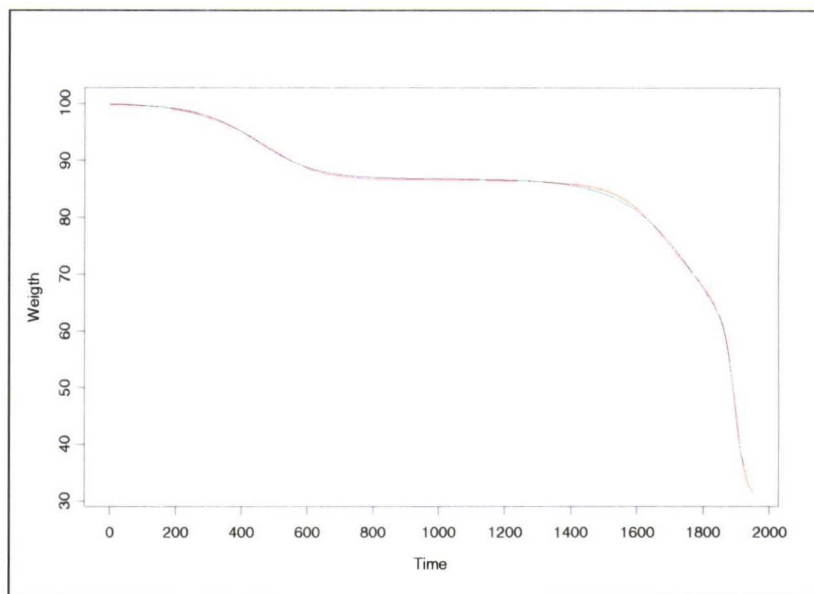


Figura 3.8: Ajuste logístico (en rojo) para una muestra de eucalipto (en azul).

3.5.3 Ajuste para una muestra de madera de eucalipto

Ajustaremos ahora el modelo logístico con tres componentes para una curva TGA correspondiente a una muestra de madera de eucalipto. Al igual que en el caso anterior del ciprés, observando la figura 3.8 descubrimos los distintos escalones que marcarán los tres subintervalos para el ajuste lineal de la función logit. Finalmente, resultan los parámetros siguientes:

	$i = 1$	$i = 2$	$i = 3$
w_i	13.047	41.094	22.534
a_i	5.067	5.458	162.17
b_i	-0.0113	-0.0085	-0.0856

Tabla 3.2: Parámetros del ajuste con 3 logísticas para el eucalipto.

La gráfica correspondiente al ajuste logístico dado en la tabla 3.2 para los datos del eucalipto en el intervalo $[0, 1850]$, se puede observar en (3.8).

3.5.4 Ajuste para una muestra de PVC

Finalmente ajustaremos el modelo logístico con cuatro componentes para una curva TGA correspondiente a una muestra de PVC. En este caso ajustaremos los datos de la muestra inicial que hemos usado para la introducción del modelo (3.20). Después de efectuar las cuatro regresiones lineales con la función logit, se obtienen los valores para el inicio del algoritmo de mínimos cuadrados

no lineales. A partir de estos valores iniciales, mediante las correspondientes iteraciones se obtienen los parámetros del modelo que se observan en la tabla 3.3.

	$i = 1$	$i = 2$	$i = 3$	$i = 4$
w_i	5.274	6.856	2.283	2.063
a_i	14.481	36.721	0.635	22.05
b_i	-0.011	-0.012	-0.00047	-0.013

Tabla 3.3: Parámetros del ajuste con cuatro logísticas para el PVC.

La gráfica correspondiente a este modelo de cuatro escalones de PVC es la figura 3.9. En ella se puede apreciar el buen ajuste de los datos (en rojo) al modelo (en azul).

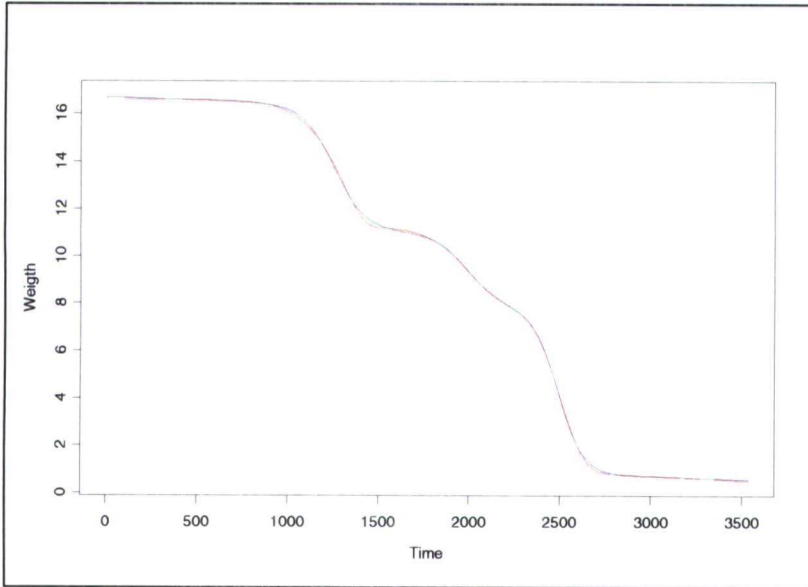


Figura 3.9: Ajuste logístico (azul) de los datos de PVC (rojo).

3.6 Interpretación del modelo mezcla de logísticas

De cara a interpretar las distintas componentes logísticas en las que se puede descomponer el modelo (3.18) analizaremos una muestra de residuos forestales. Los residuos foratales, al igual que la madera, son materiales pertenecientes al grupo de biomasas. La dificultad de interpretar los procesos que experimentan estos materiales al ser calentados mediante los modelos cinéticos clásicos nos ha llevado a la aplicación de nuestro método.

Muestra de residuo forestal del Canadá.

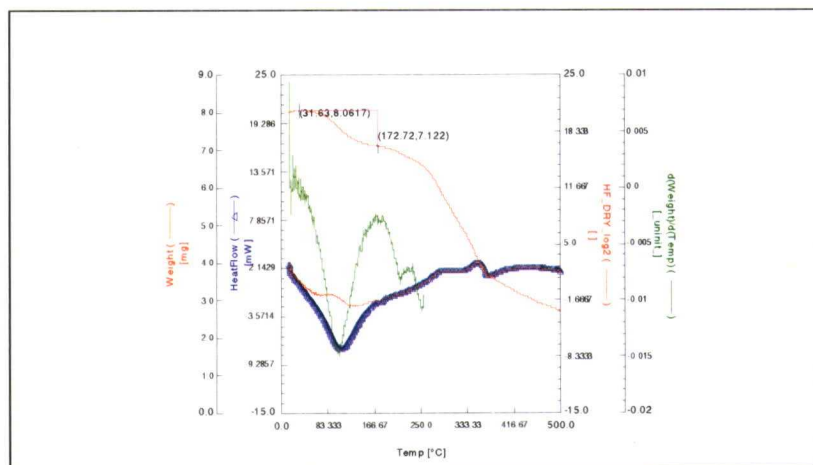


Figura 3.10: Gráficas correspondientes a residuo forestal del Canadá, TGA (en rojo), DSC (en azul) y DTGA (en verde).

Estamos interesados en el estudio del material hasta los 170°C (véase figura 3.10), en el que ya se ha producido un primer proceso, asociado normalmente a la volatilización del agua existente en la muestra. Sin embargo, pensamos que aparte del agua se pierden otras sustancias, pues el estudio de la primera derivada de la curva TGA en este primer tramo de temperaturas, denota dos o tres puntos críticos.

Con el fin de discernir entre los dos casos posibles: que se pierda agua y otra sustancia o que se pierda agua y otras dos sustancias, aplicaremos el ajuste del modelo (3.18) y analizaremos los resultados obtenidos.

3.6.1 Ajuste del modelo con 2 componentes

Suponiendo que en el primer tramo se pierde agua y otra sustancia, sus pesos vendrían dados por las constantes w_1 y w_2 , mientras que los valores iniciales de (a_1, b_1) y (a_2, b_2) corresponderían al modelo lineal que ajusta la función logit (3.21). Los parámetros finales para este primer caso son:

	$i = 1$	$i = 2$
w_i	7.18428	0.87873
a_i	12.1785	8.17382
b_i	-0.00678	-0.01081

Tabla 3.4: Parámetros para el modelo de dos logísticas.

3.6.2 Ajuste del modelo con 3 componentes

Si suponemos ahora que en el primer tramo se pierde agua y otras dos sustancias, los parámetros para este segundo caso se muestran en la tabla 3.5.

	$i = 1$	$i = 2$	$i = 3$
w_i	-0.18012	7.61532	0.79192
a_i	0.94879	3.47303	8.13485
b_i	-0.00587	-0.00073	-0.01075

Tabla 3.5: Parámetros para el modelo de tres logísticas.

Obsérvese cómo al imponer tres componentes, la primera componente resulta con constante $w_1 = -0.18012$, que indicaría una ganancia de peso de la muestra al principio del experimento. Este hecho puede apreciarse visualizando los propios datos e indicaría un pequeño error de la termobalanza que sería detectado por nuestro modelo al intentar el ajuste con tres componentes. Sin embargo, de cara a la interpretación, este hecho no debe de ser tenido en cuenta, pues el error residual de los dos modelos es prácticamente el mismo: 0.00023 y 0.00016, respectivamente, para el primer y el segundo modelo.

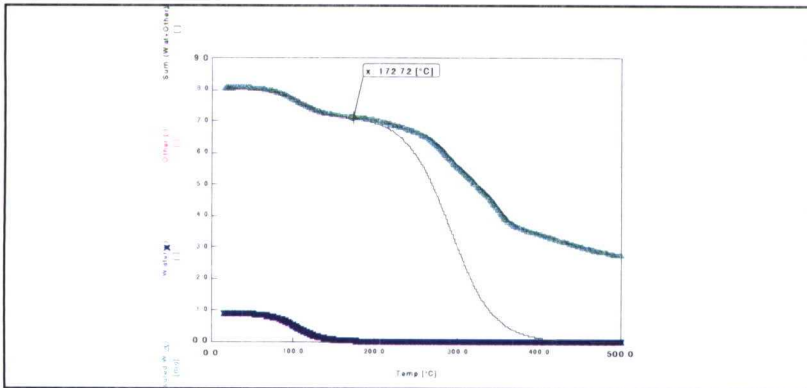


Figura 3.11: Parte ajustada de la curva TGA (hasta 172.72 °C).

Finalmente, se elegirá el modelo con dos logísticas cuyos parámetros figuran en la tabla 3.4. La interpretación de estos resultados nos permite concluir que puesto que la masa de la muestra que se pierde en este tramo analizado es de 8.0409 mg, deducimos que $w_1 = 7.18428$ mg corresponden a la pérdida de agua, mientras que $w_2 = 0.87873$ corresponden a otras sustancias. En la figura 3.11 se puede apreciar la incorporación del ajuste logístico a la parte correspondiente (hasta 172.72 °C) de la curva TGA .

3.7 Ventajas del método propuesto

Podríamos enumerar las siguientes ventajas del modelo propuesto:

1. Nuestro método cubre toda la gama de conversiones. El modelo (3.18) puede ser usado para el estudio de cualquier tipo de material, tanto si está regido por una sola ley cinética como si lo está por varias.

2. Permite representar tanto la velocidad de pérdida de masa como el ajuste de esta con el tiempo mediante una ecuación única para cada proceso.

3. No se basa en una ley preestablecida tipo Arrhenius, lo que permite compararlo con otros modelos al uso.

4. Proporciona medios estadísticos que permitirán mediante contrastes de significación valorar la bondad del ajuste del modelo propuesto.

6. Permite la aplicación de modelos cinéticos clásicos del tipo Arrhenius a cada una de las funciones de degradación individuales para su comparación con materiales ya estudiados.

7. Se reproduce perfectamente el carácter asintótico al final de cada proceso degradativo.

Resumiendo, el método que proponemos de ajuste de las curvas TGA, basado en la composición de tantas funciones logísticas como procesos distintos sufra la muestra, supone un modelo alternativo a los existentes para la explicación de la cinética. Este método nos parece más simple, en cuanto a la fácil interpretación, que los anteriores basados en la ley de Arrhenius.

En cuanto a su validación, una primera versión del modelo, concretamente la aplicación a procesos simples, lo hemos presentado en el V Congreso Gallego de Estadística e Investigación Operativa (Naya, Cao y Artiaga 2001) y en al 5th Mediterranean Conference on Calorimetry and Thermal Analysis (Artiaga, Cao y Naya 2001). La aplicación al hexahidroftálico fue enviada a la revista *Thermochimica Acta* siendo recomendada su publicación (Naya, Cao y Artiaga (2003a)). El modelo con mezcla de logísticas, con aplicaciones a distintos casos, se ha enviado a la revista *Journal of Chemometrics* (Naya, Cao, Artiaga, Barbadillo y López de Ullibarri (2003)).

Capítulo 4

Estimación no paramétrica de curvas

4.1 Introducción

El modelo paramétrico propuesto en el capítulo anterior para curvas TGA es aplicable a un gran número de experimentos. Sin embargo, en aquellos materiales que presenten caídas bruscas de peso, el modelo paramétrico logístico no se ajusta bien. Por este motivo, en este capítulo, propondremos un ajuste de tipo no paramétrico para estas curvas. Además, será aplicable al resto de señales, como el caso de las DSC, DMTA e, incluso, para espectros procedentes de estudios mediante técnicas de infrarrojos.

Los primeros métodos usados para el suavizado de curvas DSC o TGA se basaban en suavizado mediante medias móviles, como el propuesto por Savitzky y Golay (1964). Hoy en día el software disponible para Análisis Térmico permite el suavizado mediante métodos no paramétricos. Sin embargo, en lo relativo a la estimación de las derivadas, suelen utilizarse métodos de derivación numérica lo que provoca mucho ruido en la señal, y, por consiguiente, un pésimo ajuste.

4.1.1 Inconvenientes del modelo logístico

Aunque en la mayoría de los casos estudiados se puede encontrar un buen ajuste del modelo logístico (figura 3.9) a los datos reales, existen muestras en donde los escalones presentan curvaturas “poco suaves” dificultando el buen ajuste. Como ejemplos de estos casos en que el ajuste no funciona del todo bien, veremos los datos del eucalipto y del hexahidroftálico usados en el capítulo anterior con la muestra completa.

En el caso de la curva TGA correspondiente a una muestra de madera de eucalipto se observa que presenta un caída brusca de peso, alrededor de 2400, lo que hace que el ajuste en esta parte de la curva no sea del todo bueno, como puede apreciarse en la figura 4.1.

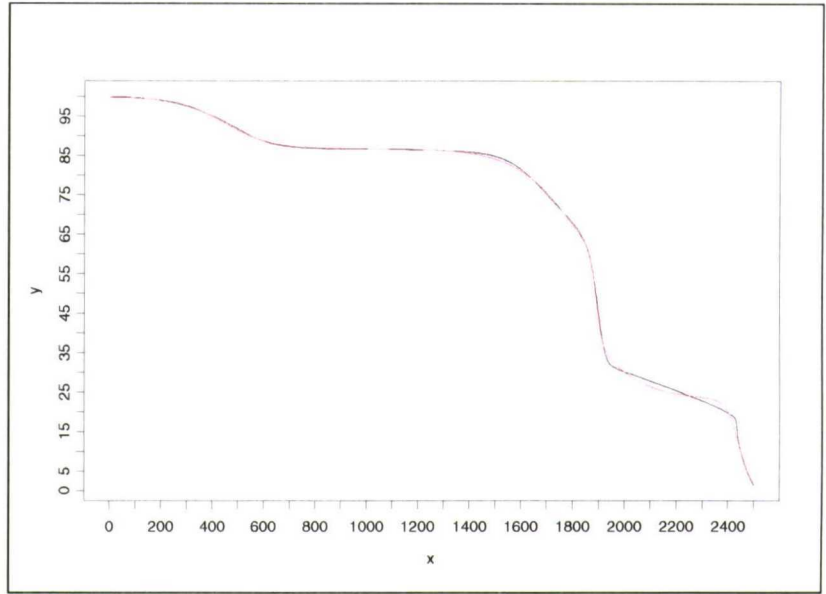


Figura 4.1: Ajuste paramétrico logístico (en rojo) de toda la muestra de eucalipto (en negro).

En el caso en que la curva TGA sufra una caída brusca de peso, el ajuste logístico no recoge bien esta caída, este inconveniente se manifiesta con mayor claridad en la estimación de la primera derivada. Como ejemplo podemos observar la figura 4.2, en la que se destaca la primera derivada obtenida por el método logístico y la estimada mediante un programa convencional como es el Orchestrator[®] de Rheometric Scientific Incorporation[®].

La muestra analizada, que corresponde a anhídrido hexahidroftálico, tiene como ecuación de ajuste del modelo logístico ($\hat{m}_l(t)$) y como primera derivada ($\hat{m}'_l(t)$).

$$\hat{m}_l(t) = 13.3168 \frac{\exp(16.7382 - 0.0134742t)}{1 + \exp(16.7382 - 0.0134742t)},$$

$$\hat{m}'_l(t) = 0.17943 \frac{\exp(16.7382 - 0.0134742t)}{(1 + \exp(16.7382 - 0.0134742t))^2}.$$

4.1.2 Inconvenientes del software actual

El software actual para el suavizado de datos procedentes de análisis térmico permite realizar estimaciones no paramétricas, tanto de la regresión como de sus derivadas. Sin embargo, estas estimaciones, en algunos casos presentan serias deficiencias, que son más destacables en el caso de la estimación de las derivadas.

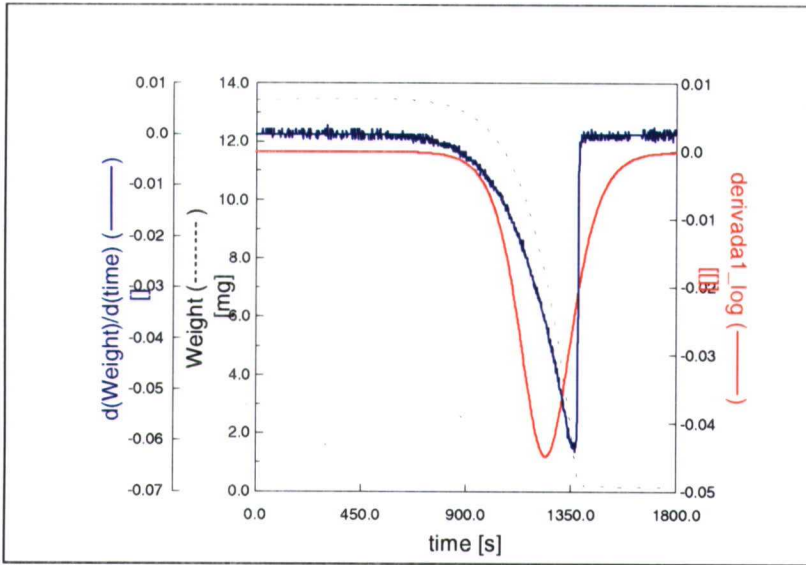


Figura 4.2: Curva TGA (en negro), primera derivada estimada por método logístico (en rojo) y primera derivada estimada por Orchestrator (en azul).

En la figura 4.3 podemos ver la curva TGA y su primera derivada, para una muestra de oxalato de calcio, usando el programa Orchestrator[®] de Rheometric Scientific Incorporation[®]. Este programa permite un suavizado a mano de la curva, lo que suele conllevar un sobresuavizado de la señal. Por otra parte, en el caso de utilizar los métodos no paramétricos aportados por dicho programa, estos no tienen en cuenta la dependencia de los datos.

En este capítulo nos centraremos en el cálculo de la ventana óptima mediante un método plug-in en dos etapas y finalmente abordaremos los problemas para la adaptación de dicho método a un programa de ordenador. Esta implementación se ha realizado en lenguaje C++, con la ayuda del informático Javier Trillo, quien lo presentó como Proyecto de Fin de Carrera (véase Trillo (2001)).

Por otra parte, estamos a la espera de integrar este módulo en el propio programa Orchestrator[®], ya que la empresa nos ha manifestado su interés por disponer de este método de ajuste automático de datos para incorporarlo a la nueva versión del programa.

4.2 Estimación no paramétrica de curvas DSC y TGA

Como se indicó en el capítulo 2, los métodos de regresión no paramétrica para estimar la función m no necesitan especificar *a priori* ningún modelo paramétrico. La idea es asumir que m es una función suave que puede aproxi-

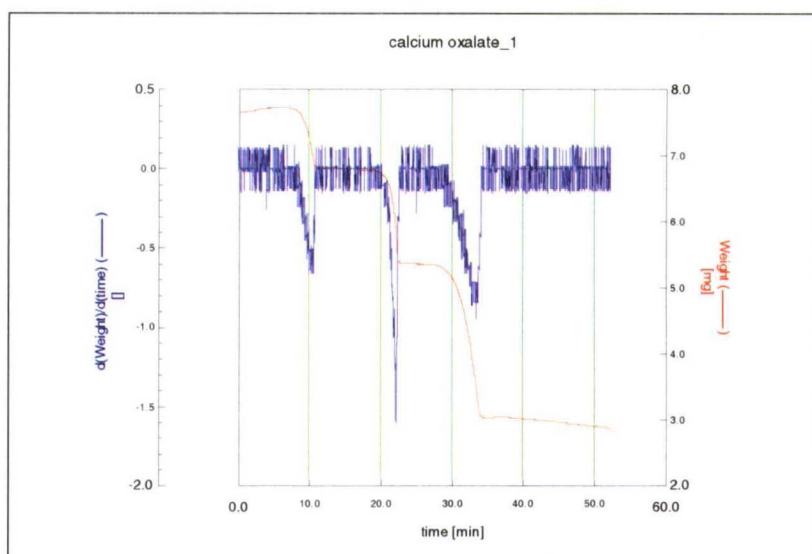


Figura 4.3: Curva TGA (en rojo) y su primera derivada (en azul) para el oxalato de calcio.

marse, en cada x , por $\hat{m}(x)$, haciendo medias ponderadas de las Y -observaciones más próximas al punto x , lo que expresamos como:

$$\hat{m}(x) = \sum_{i=1}^n W_{ni}(x) Y_i, \text{ con } \sum_{i=1}^n W_{ni}(x) = 1, \quad (4.1)$$

donde $W_{ni}(x)$ es el peso de la i -ésima observación correspondiente al punto x . Estos pesos dependen del parámetro de suavizado h y de la función núcleo K .

En nuestro caso, de los distintos métodos no paramétricos expuestos en el capítulo 2, elegiremos el método de estimación polinómico local por sus buenas propiedades para la estimación de las derivadas y su fácil implementación. Además usaremos la versión de este estimador para datos dependientes, ya que, analizados diferentes experimentos, comprobamos la dependencia existente entre los datos. Concretamente, hemos verificado que los errores siguen un modelo autorregresivo de orden 1.

4.2.1 Estimador polinómico local para curvas TGA y DSC

El estimador polinómico local introducido por Stone (1977) y Cleveland (1979) ha sido ampliamente estudiado por otros autores como el caso de Ruppert y Wand (1994) y Fan y Gijbels (1995). Este estimador, como se puso de manifiesto en el capítulo 2, estima la función de regresión ($j = 0$) y sus derivadas ($j = 1, 2, \dots, p$), localmente en cada punto x , mediante la expresión:

$$\widehat{m}^{(j)}(x) = j! \cdot \widehat{\beta}_j, \quad j = 0, 1, 2, \dots, p, \tag{4.2}$$

donde

$$\widehat{\beta} = (\widehat{\beta}_0, \dots, \widehat{\beta}_p) = \arg \min_{\beta} (Y - X\beta)^t W (Y - X\beta),$$

$$X = \begin{pmatrix} 1 & (X_1 - x_0) & \dots & (X_1 - x_0)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (X_n - x_0) & \dots & (X_n - x_0)^p \end{pmatrix}, \quad Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix},$$

y $W = \text{diag}\{K_h(x_i - x)\}$ es la matriz $n \times n$ de pesos. La expresión para el vector estimado $\widehat{\beta}$ es:

$$\widehat{\beta} = (X^t W X)^{-1} X^t W Y. \tag{4.3}$$

4.2.2 Elección de la función núcleo y del grado del polinomio

Para el cálculo del estimador polinómico local (4.2) es necesario elegir una función núcleo o kernel, K , el grado del polinomio, p , y la ventana, h . La elección de K y p es un problema secundario con respecto al parámetro de suavizado h . Fan y Gijbels (1996) recomiendan el uso del núcleo de Epanechnikov, puesto que minimiza el error cuadrático medio asintótico para la ventana óptima. También sugieren que se elija p en función del orden de la derivada a estimar, j , de forma que $p - j$ sea impar. Por tanto, tomaremos $p = 1$ para estimar la función de regresión, y $p = 3$ para la estimación de la segunda derivada de la función de regresión. En general, basándonos en el hecho de que el sesgo decrece y la varianza crece con p , una recomendación práctica es tomar $p - j = 1$ ó 3 . El problema de la elección de la ventana lo trataremos con mayor detalle en un apartado posterior.

Un inconveniente del estimador polinómico local es que la varianza condicional tienda a infinito, cuando la ventana usada para la estimación (con un núcleo de soporte compacto) no contenga más de $p + 3$ puntos. Este problema, apuntado por Seifert y Gasser (1996), puede resolverse aumentando el parámetro de suavizado en estos valores, tema que trataremos al final del capítulo.

En el supuesto de tener datos dependientes, como es el caso que nos ocupa, el estimador polinómico local clásico puede también usarse, teniendo en cuenta que el error cuadrático medio asintótico depende de la suma de las covarianzas de los errores. Una alternativa para el caso de dependencia ha sido propuesta por Francisco-Fernández y Vilar-Fernández (2001), mediante una adaptación de las ideas presentadas en Fan y Gijbels (1996) para el caso de datos dependientes.

4.3 Estimación de la ventana óptima

Siempre que se realiza una estimación no paramétrica, como se puso de manifiesto en sección anterior y en el capítulo 2 de esta memoria, la elección del parámetro de suavizado o parámetro ventana es crucial para el correcto ajuste de los datos. La selección de este parámetro decidirá las propiedades que tendrá el estimador usado.

Una elección de la ventana demasiado pequeña dará lugar a estimadores con mucha varianza, mientras que si el valor es demasiado grande proporcionará estimadores con mucho sesgo. Por tanto, en la selección correcta de dicho parámetro se tendrá que llegar a una situación intermedia que permita minimizar el error cuadrático medio (suma de la varianza y del cuadrado del sesgo). En nuestro caso usaremos el método plug-in.

4.3.1 Tipos de error

Puesto que los diferentes métodos para elegir el grado de suavizado se basan en minimizar algún tipo de error cometido en la estimación, es importante revisar los diferentes tipos de error. En todo caso, nos centraremos en esta sección, en aquellos errores, para la estimación de la función de regresión, útiles en la selección de la ventana.

En primer lugar tenemos el error cuadrático medio (*Mean Squared Error*) para la estimación en un punto (error local) definido en cada punto x como:

$$MSE_x(\hat{m}_h) = E(\hat{m}_h(x) - m(x))^2. \quad (4.4)$$

Para el caso de la estimación de la derivada v -ésima se tiene:

$$MSE_x(\hat{m}_h^{(v)}) = E(\hat{m}_h^{(v)}(x) - m(x)^{(v)})^2. \quad (4.5)$$

Descomponiendo el cuadrado y aplicando las propiedades de la media y varianza se obtiene la expresión del *MSE*, como función del sesgo y de la varianza del estimador del modo siguiente:

$$MSE_x(\hat{m}_h) = (E(\hat{m}_h(x)) - m(x))^2 + Var(\hat{m}_h(x)).$$

El error cuadrático integrado *ISE* (*Integrated Squared Error*) se define como:

$$ISE(\hat{m}_h) = \int (\hat{m}_h(x) - m(x))^2 w(x) dx.$$

Una vez definida la medida del error local para cada punto x , es fácil extender ésta al caso global, sin más que considerar la integral de las diferencias al cuadrado en cada punto, ponderada por alguna función de pesos no negativa w , y luego tomar la media de estas integrales obteniendo el error cuadrático medio integrado *MISE* (*Mean Integrated Squared Error*), que viene a ser el valor esperado del *ISE*, y que tendrá la fórmula siguiente:

$$MISE(\hat{m}_h) = E \int (\hat{m}_h^{(v)}(x) - m(x)^{(v)})^2 w(x) dx, \quad (4.6)$$

donde w es una función de ponderación que permite otorgar distinto peso a cada punto en el que se está realizando la estimación. Este error de medida fue introducido por Rosenblatt (1956).

El error cuadrático promedio, ASE (*Average Square Error*) se define como:

$$ASE(\hat{m}_h) = \frac{1}{n} \sum_{i=1}^n (\hat{m}_h(X_i) - m(X_i))^2 w(X_i).$$

También es posible definir otros tipos de error considerando en vez de criterios cuadráticos, otras normas como por ejemplo la norma L^1 , que daría lugar al valor absoluto integrado IAE y su correspondiente valor esperado $MIAE$.

La importancia del $MISE$ está en que permite un estudio asintótico mediante la descomposición del error en suma de la varianza y del sesgo, lo que justifica la influencia de la ventana h en el suavizado de los datos, en el sentido de que ventanas pequeñas producen infrasuavización (poco sesgo y mucha varianza) y ventanas grandes producen sobresuavización (poca varianza y mucho sesgo).

4.3.2 Métodos para la estimación de la ventana óptima

Para la selección de la ventana se puede recurrir a dos vías. Una consiste en estimar de manera consistente alguna medida de error y seleccionar como parámetro aquel que minimice dicha estimación; esta idea es usada, por ejemplo, por los métodos de validación cruzada. La otra vía se basa en encontrar una expresión asintótica de una medida de error, para después obtener como parámetro aquel que minimice los términos de mayor orden de dicha expresión. En este caso, la expresión de esta medida va a depender de cantidades desconocidas lo que obligará a una primera estimación de estas. Los métodos basados en esta idea son los denominados métodos plug-in, debido a que es necesario una estimación previa y luego introducir esta estimación (“enchufarla”) para el cálculo final de la ventana óptima.

Los métodos de validación cruzada (*cross-validation*), propuestos por Rude-mo (1982) y Bowman (1984) para la estimación de la densidad, se basan en seleccionar el valor de h que minimice una estimación del $ISE(\hat{m}_h)$. Aunque en un principio estos métodos fueron usados para el caso de independencia de los errores se pueden utilizar mediante una ligera modificación, propuesta por Hart y Vieu (1990), para el caso de dependencia.

Los métodos de tipo plug-in se basan en elegir aquel parámetro ventana que minimice, en cada punto x , el error cuadrático medio, $MSE_x(h)$, tanto asintótico como teórico, entre la estimación no paramétrica de la v -ésima derivada de la función de regresión y su valor en el punto x :

$$MSE_x(h) = E \left(\hat{m}_h^{(v)}(x) - m(x)^{(v)} \right)^2. \quad (4.7)$$

Así, utilizando el $MSE_x(h)$, se obtendría la ventana óptima local, y para la obtención de una ventana global para toda la muestra, consideraremos el $MISE$, que tiene en cuenta las distancias entre el estimador y la verdadera curva del modo siguiente:

$$MISE(h) = E \int \left(\widehat{m}_h^{(v)}(x) - m(x)^{(v)} \right)^2 w(x) dx, \quad (4.8)$$

para alguna función de pesos w . La expresión (4.8) puede fácilmente descomponerse en la suma de la integral de la varianza y en la integral del cuadrado del sesgo.

En el supuesto de diseño fijo, independencia en la estructura de los errores y tomando $v + p$ impar, Fan y Gijbels (1996) expresan el sesgo y la varianza para el estimador polinómico local de la forma:

$$Sesgo(m^{(v)}(x)) = h_n^{p+1-v} \frac{m^{(p+1)}(x)}{(p+1)!} v! B_v (1 + o(1)), \quad (4.9)$$

$$Var(m^{(v)}(x)) = \frac{1}{nh_n^{2v+1}} \frac{c(\varepsilon)}{f(x)} (v!)^2 V_v (1 + o(1)), \quad (4.10)$$

donde f es la densidad del diseño de los datos y los valores B_v y V_v dependen de la función núcleo usada para $v = 0, 1, \dots, p$ (véase Ruppert, Sheather y Wand (1995) para más detalles). Usando las expresiones (4.9) y (4.10) el parámetro de suavizado que minimiza la expresión asintótica del error dada en (4.8) puede fácilmente calcularse mediante la expresión:

$$h_{AMISE} = C_{\nu,p}(K) \left(\frac{\sigma^2}{n \int (m^{(p+1)}(x))^2 w(x) f(x) dx} \right)^{\frac{1}{2p+3}},$$

donde σ^2 es la varianza del error (supuesto homocedástico) y $C_{\nu,p}$ son constantes que pueden obtenerse mediante:

$$C_{\nu,p}(K) = \left[\frac{(p+1)!^2 (2\nu+1) \int K_\nu^{*2}(t) dt}{2(p+1-\nu) \left\{ \int t^{p+1} K_\nu^*(t) dt \right\}^2} \right]^{\frac{1}{2p+3}},$$

con $K_\nu^*(t) = \left(\sum_{\ell=0}^p S^{\nu\ell} t^\ell \right) K(t)$, $S^{\nu\ell}$ son los elementos de la matriz S^{-1} y $S = (\mu_{j+\ell})_{j,\ell=0}^p$, con $\mu_j = \int u^j K(u) du$.

En el caso de dependencia de los errores se obtienen fórmulas similares basándose en expresiones paralelas del sesgo (4.9) y de la varianza (4.10). Así, para el error cuadrático medio asintótico (véase Francisco-Fernández y Vilar-Fernández (2001) para más detalles) se obtiene la expresión siguiente del $AMSE(h)$:

$$AMSE(h) = \left(h_n^{p+1-v} \frac{m^{(p+1)}(x)}{(p+1)!} v! B_v \right)^2 + \frac{1}{nh_n^{2v+1}} \frac{c(\varepsilon)}{f(x)} (v!)^2 V_v, \quad (4.11)$$

donde $c(\varepsilon) = \sum_{k=-\infty}^{+\infty} c(k)$ y $c(k)$ es la autocovarianza de orden k de los errores ε_i .

Por tanto, la ventana óptima asintótica (en el sentido del *AMSE*) para estimar la ν -ésima derivada de la función de regresión es:

$$h_{opt,L} = C_{\nu,p}(K) \left(\frac{c(\varepsilon)}{n(m^{(p+1)}(x))^2 f(x) dx} \right)^{\frac{1}{2p+3}}. \quad (4.12)$$

Una expresión similar para (4.11) puede obtenerse para el *AMISE* (h), y la ventana óptima (en el sentido del *AMISE*) será:

$$h_{opt,G} = C_{\nu,p}(K) \left(\frac{c(\varepsilon)}{n \int (m^{(p+1)}(x))^2 w(x) f(x) dx} \right)^{\frac{1}{2p+3}} \quad (4.13)$$

Las fórmulas anteriores son válidas si $\nu + p$ es impar. En el supuesto de que $\nu + p$ sea par, las expresiones para las ventanas óptimas son:

$$h_{opt,L} = C_{\nu,p}(K) \left(\frac{c(\varepsilon)}{n(m^{(p+2)}(x))^2 f(x) dx} \right)^{\frac{1}{2p+5}}, \quad (4.14)$$

$$h_{opt,G} = C_{\nu,p}(K) \left(\frac{c(\varepsilon)}{n \int (m^{(p+2)}(x))^2 w(x) f(x) dx} \right)^{\frac{1}{2p+5}}. \quad (4.15)$$

4.4 Método plug-in en dos etapas bajo dependencia

Para la estimación de la ventana con el método polinómico local bajo dependencia, las ventanas asintóticamente óptimas local y global están dadas en las expresiones (4.12) y (4.13) ó (4.14) y (4.15). Para la estimación de estas expresiones uno de los problemas consiste en estimar $c(\varepsilon)$, la suma de autocovarianzas, y la $(p + 1)$ -ésima derivada de la función de regresión. Estos dos aspectos serán tratados a continuación.

4.4.1 Estimación de la suma de autocovarianzas

Aunque $c(\varepsilon)$ puede estimarse a través de la densidad espectral de los errores ε_i , nuestro primer método será algo más simple. Asumiremos, en base a distintos modelos de muestras estudiados, que los errores ε_i siguen un proceso autorregresivo de orden 1 (*AR*(1)) con coeficiente de autocorrelación ρ .

Método 1

En un primer método utilizaremos que, en un $AR(1)$, $c(\varepsilon)$ puede expresarse, en términos de la varianza del error, σ^2 , y del coeficiente de autocorrelación, como:

$$c(\varepsilon) = \sum_{k=-\infty}^{+\infty} c(k) = \sum_{k=-\infty}^{+\infty} \sigma^2 \rho^k = \sigma^2 \left(1 + 2 \sum_{k=1}^{+\infty} \rho^k \right) = \sigma^2 \frac{1 + \rho}{1 - \rho}. \quad (4.16)$$

Podemos calcular los residuos $\hat{\varepsilon}_i = Y_i - \hat{m}_h(x_i)$, usando un parámetro de suavizado preliminar h , y después hallar la varianza estimada, que vendrá dada por:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (\hat{\varepsilon}_i - \bar{\varepsilon})^2,$$

donde $\bar{\varepsilon} = \frac{1}{n} \sum_{i=1}^n \hat{\varepsilon}_i$. El estimador del primer coeficiente de autocorrelación que usaremos es:

$$\hat{\rho} = \frac{\sum_{i=1}^{n-1} (\hat{\varepsilon}_i - \bar{\varepsilon})(\hat{\varepsilon}_{i+1} - \bar{\varepsilon})}{\sum_{i=1}^n (\hat{\varepsilon}_i - \bar{\varepsilon})^2}.$$

Finalmente, estos valores se usarán para el cálculo del estimador de $c(\varepsilon)$:

$$\hat{c}(\varepsilon) = \hat{\sigma}^2 \frac{1 + \hat{\rho}}{1 - \hat{\rho}}. \quad (4.17)$$

Método 2

El segundo método es también muy simple. Se basa en el hecho de que los datos que aparecen en la mayoría de los experimentos termogravimétricos presentan valores contiguos muy próximos de la variable explicativa, en los que la función m no fluctúa demasiado.

Puesto que los ε_i no son observables, calculamos las diferencias de orden uno, $Y_i - Y_{i-1}$. Debido a la cercanía entre los datos x_i y x_{i-1} , podemos asumir que $m(x_i) - m(x_{i-1}) \simeq 0$, y por tanto la serie de las diferencias $Y_i - Y_{i-1}$ es aproximadamente igual a la serie $e_i = \varepsilon_i - \varepsilon_{i-1}$, obteniendo $Y_i - Y_{i-1} \simeq e_i$, donde $e_i = \varepsilon_i - \varepsilon_{i-1}$.

Asumiendo también que los ε_i siguen un proceso $AR(1)$, mediante la expresión: $\varepsilon_i = \rho \varepsilon_{i-1} + a_i$ (siendo a_i ruido blanco), escribiendo la varianza de e_i , σ_e^2 , y el primer coeficiente de autocorrelación de e_i , ρ_e , mediante propiedades de los modelos $AR(1)$ se sigue que:

$$\sigma_e^2 = \rho^2 \sigma_\varepsilon^2 + \sigma_a^2, \text{ lo que implica que } \sigma_\varepsilon^2 = \frac{\sigma_a^2}{1 - \rho^2},$$

$$\sigma_e^2 = \text{Var}(\varepsilon_i - \varepsilon_{i-1}) = \text{Var}((\rho - 1)\varepsilon_{i-1} + a_i) = (\rho - 1)^2 \sigma_\varepsilon^2 + \sigma_a^2.$$

De las expresiones anteriores se obtiene:

$$\sigma_e^2 = 2(1 - \rho) \sigma_\varepsilon^2,$$

o, equivalentemente:

$$\sigma_\varepsilon^2 = \frac{\sigma_e^2}{2(1 - \rho)}. \tag{4.18}$$

El primer coeficiente de autocorrelación puede obtenerse teniendo en cuenta que:

$$\begin{aligned} \gamma_e &= Cov(e_i, e_{i+1}) = E((\varepsilon_i - \varepsilon_{i-1})(\varepsilon_{i+1} - \varepsilon_i)) \\ &= E[((\rho - 1)\varepsilon_{i-1} + a_i)((\rho - 1)\varepsilon_i + a_{i+1})] = \\ &= (\rho - 1)^2 E(\varepsilon_i \cdot \varepsilon_{i-1}) + (\rho - 1) E(\varepsilon_i \cdot a_i) = (\rho - 1)^2 \gamma_\varepsilon + (\rho - 1) \sigma_a^2 = \\ &= (\rho - 1)^2 \rho \sigma_\varepsilon^2 + (\rho - 1) \sigma_\varepsilon^2 (1 - \rho^2) = \sigma_\varepsilon^2 (\rho - 1) (1 - \rho), \end{aligned}$$

así que:

$$\begin{aligned} \rho_e &= \frac{\gamma_e}{\sigma_e^2} = \frac{\sigma_\varepsilon^2 (\rho - 1) (1 - \rho)}{2(1 - \rho) \sigma_\varepsilon^2} = \frac{\rho - 1}{2}, \\ \rho &= 1 + 2\rho_e. \end{aligned} \tag{4.19}$$

En la práctica, aunque asumimos que $m(x_i) - m(x_{i-1}) \simeq 0$, para alguna ventana piloto g , podemos calcular un estimador de la regresión polinómico local con la intención de estimar las diferencias:

$$\tilde{e}_i = Y_i - Y_{i-1} - (\hat{m}_g(x_i) - \hat{m}_g(x_{i-1}))$$

Entonces dados los estimadores

$$\begin{aligned} \hat{\sigma}_e^2 &= \frac{1}{n-1} \sum_{i=2}^n \tilde{e}_i^2 - (\bar{e})^2, \quad \bar{e} = \frac{1}{n-1} \sum_{i=2}^n \tilde{e}_i \\ \hat{\rho}_e &= \frac{\frac{1}{n-2} \sum_{i=2}^{n-1} \tilde{e}_i \tilde{e}_{i+1}}{\hat{\sigma}_e^2}, \end{aligned}$$

y considerando las expresiones (4.18) y (4.19) se obtiene:

$$\hat{\rho}_\varepsilon = \hat{\rho} = 1 + 2\hat{\rho}_e, \tag{4.20}$$

$$\hat{\sigma}_\varepsilon^2 = \frac{\hat{\sigma}_e^2}{2(1 - \hat{\rho}_\varepsilon)}. \tag{4.21}$$

Finalmente, sustituyendo (4.20) y (4.21) en (4.16) se obtiene el estimador para la suma de autocovarianzas:

$$\widehat{c(\varepsilon)} = \sum_{k=-\infty}^{+\infty} \widehat{c(k)} = \hat{\sigma}_\varepsilon^2 \frac{1 + \hat{\rho}}{1 - \hat{\rho}} \tag{4.22}$$

4.4.2 Selección de las ventanas piloto

El método plug-in requiere de la estimación de cantidades desconocidas en las fórmulas (4.12) y (4.13), mediante parámetros de suavizado que denominaremos $h_{PI,L}$ y $h_{PI,G}$ (ventana piloto local y ventana piloto global). Para un valor fijo, ν , el método necesita estimar $m^{(\nu)}$ usando un polinomio de grado p .

La estimación de $c(\varepsilon)$, según se ha puesto de manifiesto en la sección previa, requiere de la elección de una ventana preliminar h_1 , para el cálculo de los residuos. Su elección se tratará en el párrafo siguiente. Además, la ventana plug-in también necesita estimar la $(p+1)$ -ésima derivada de la función de regresión, m . La estimación polinómica local necesita, para estimar la ν_1 -ésima derivada ($\nu_1 = p+1$), el empleo de un polinomio de grado $p_1 = p+2$. Esto requiere la elección de una ventana preliminar piloto, $h_2^{(1)}$. Para su determinación se observa que, en la fórmula para las ventanas (local o global) asintóticamente óptimas para esta nueva situación, aparecen de nuevo dos cantidades desconocidas: $c(\varepsilon)$, no problemática, como veremos, y la derivada (p_1+1) -ésima de m (es decir $m^{(p+3)}$). La idea del método plug-in en dos etapas consiste en observar que la fórmula de la ventana asintóticamente óptima para este nuevo problema (ventana pre-piloto), $h_2^{(0)}$, proporciona un valor:

$$h_2^{(0)} = C_2 \delta n^{-\frac{1}{2p_1+3}} = C_2 \delta n^{-\frac{1}{2p+7}}, \quad (4.23)$$

donde δ es un factor de escala y C_2 es una constante que se obtiene de forma heurística a partir de los datos.

Supongamos un determinado experimento termogravimétrico con datos equiespaciados, o casi equiespaciados. Elegiremos $\delta = (x_n - x_1)/(n-1)$. Para la elección de h_1 (ventana piloto para los residuos) usaremos un estimador lineal ($p_1 = 1$) y estableceremos su valor inicial, utilizando directamente la expresión de su valor asintóticamente óptimo:

$$h_1 = C_1 \delta n^{-\frac{1}{2p_1+3}} = C_1 \delta n^{-\frac{1}{5}}, \quad (4.24)$$

para alguna constante C_1 que obtendremos de forma heurística.

Para la obtención de las constantes C_1 y C_2 utilizaremos una parte de una curva DSC correspondiente al experimento con oxalato de calcio. Tomando $n = 950$ y usando como ventana inicial $h_1 = 6$, calculamos los residuos para la estimación de la suma de autocovarianzas, seleccionando distintos valores de la ventana prepiloto $h_2^{(0)}$ para la obtención de las ventanas mediante el procedimiento plug-in en dos etapas.

Los resultados para este ejemplo pueden verse en las tablas 4.1 y 4.2. Estas muestran que la influencia de la ventana prepiloto $h_2^{(0)}$ en la elección de la ventana piloto, $h_2^{(1)}$, es muy pequeña. Cuando estimamos la ventana piloto, partiendo de una ventana prepiloto 10 veces mayor (pasamos de 10 a 100) el valor de $h_2^{(1)}$ queda multiplicado por 4, teniendo un aumento de 1.5 veces en la ventana plug-in h_{PI} . Para la estimación de la primera y segunda derivadas, la ventana plug-in es todavía más estable con respecto a la selección de la ventana

prepiloto. Hemos comprobado que resultados similares se obtienen con otras muestras de material y con otros valores iniciales para h_1 , fijando el valor de $h_2^{(0)}$.

	m	m	m'	m'
$h_2^{(0)}$	$h_2^{(1)}$	$h_{PI,G}$	$h_2^{(1)}$	$h_{PI,G}$
	$\nu = 2, p = 3$	$\nu = 0, p = 1$	$\nu = 3, p = 4$	$\nu = 1, p = 2$
10	6.860839	1.01116	4.56602	2.6356
20	9.592215	1.07833	7.61298	3.2715
30	11.928195	1.13428	9.67675	3.5414
40	14.038808	1.18431	11.5646	3.7621
50	16.379445	1.23940	12.8962	3.9113
100	26.760712	1.47904	22.5909	4.9377

Tabla 4.1. Ventanas piloto en dos etapas para el método plug-in para la estimación de $m^{(\nu)}$, $\nu = 0, 1$, con $h_1 = 6$.

	m''	m''
$h_2^{(0)}$	$h_2^{(1)}$	$h_{PI,G}$
	$\nu = 4, p = 5$	$\nu = 2, p = 3$
10	4.3589	2.7163
20	9.0604	4.6653
30	11.9025	5.3950
40	13.7104	5.7591
50	16.0595	6.2005
100	28.3775	8.3594

Tabla 4.2. Ventanas piloto en dos etapas para el método plug-in para la estimación de m'' , con $h_1 = 6$.

Observando las figuras 4.4, 4.5 y 4.6 se puede apreciar cómo se estabiliza el valor de la ventana en la segunda etapa en función del aumento de las ventanas de la primera etapa. Del estudio de estos resultados experimentales obtenemos que un buen valor inicial es tomar $h_1 = 6$. En el mismo orden, valores de $h_2^{(0)} \simeq 30$, para m , $h_2^{(0)} \simeq 28$, para m' y $h_2^{(0)} \simeq 30$, para m'' , resultan adecuados para la ventana prepiloto $h_2^{(0)}$.

La tabla 4.3 muestra las fórmulas heurísticas a utilizar para el cálculo de las ventanas piloto h_1 y prepiloto $h_2^{(0)}$ en función del tamaño muestral (n) y de la distancia media entre observaciones contiguas ($\delta = (x_n - x_1)/(n - 1)$).

ν	0	1	2
C_1	24	24	24
C_2	64	52	51
h_1	$C_1 \delta n^{-\frac{1}{5}}$	$C_1 \delta n^{-\frac{1}{5}}$	$C_1 \delta n^{-\frac{1}{5}}$
$h_2^{(0)}$	$C_2 \delta n^{-\frac{1}{5}}$	$C_2 \delta n^{-\frac{1}{11}}$	$C_2 \delta n^{-\frac{1}{13}}$

Tabla 4.3. Valores utilizados para C_1 , C_2 , h_1 y $h_2^{(0)}$ para estimar $m^{(\nu)}$.

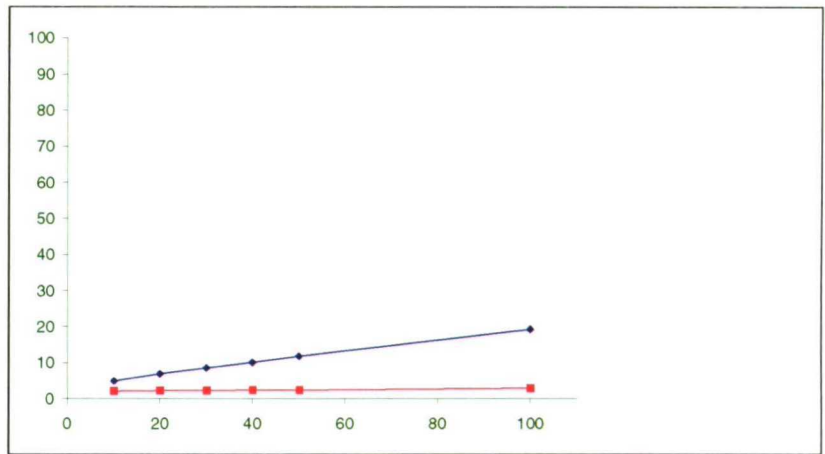


Figura 4.4: Relación entre las ventanas en dos etapas para el caso de la regresión con $p = 3$. Ventana primera etapa (en azul) y ventana segunda etapa (en rosa).

4.5 Aspectos Computacionales

Uno de los problemas cuando se intenta implementar el cálculo del estimador polinómico local es que la matriz X^tWX sea singular. Esto puede ocurrir cuando el núcleo es de soporte compacto, el diseño es equiespaciado y la ventana es pequeña. Consideremos un punto concreto, x_0 , donde la estimación se ha llevado a cabo con una ventana h . Asumiremos que el soporte de la función núcleo es el intervalo $[-1, 1]$, entonces sólo usaremos los valores x_i 's pertenecientes al intervalo $[x_0 - h, x_0 + h]$ para la obtención de la estimación en el punto x_0 .

En estas situaciones, nos encontramos con que el número de puntos utilizados para la estimación de x_0 puede ser demasiado pequeño, y por tanto, resultar imposible el cálculo del vector de soluciones del problema de mínimos cuadrados locales:

$$\hat{\beta} = (X^tWX)^{-1} X^tWY.$$

Seifert y Gasser (1996) han estudiado el caso en que $\det(X^tWX) = 0$ (caso en que no podríamos obtener $(X^tWX)^{-1}$), dando las siguientes condiciones para el estimador polinómico local:

1. Para la estimación en el punto x_0 se requiere, al menos, $p + 1$ puntos en el intervalo $[x_0 - h, x_0 + h]$. Esta condición es más o menos restrictiva al incrementar el valor del grado del polinomio p .
2. Para garantizar que la varianza del estimador sea finita, es necesario tomar al menos $p + 3$ datos en el intervalo $[x_0 - h, x_0 + h]$.

En los métodos de regresión polinómica local, cuando trabajamos con funciones núcleo de soporte compacto (como son el núcleo de Epanechnikov y el

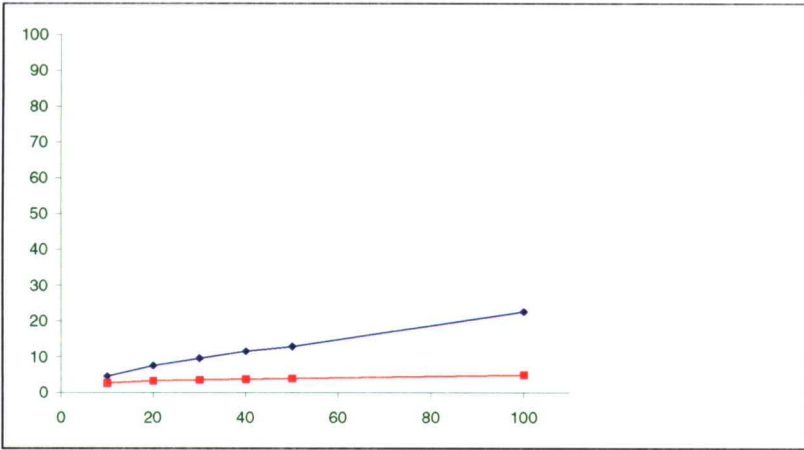


Figura 4.5: Relación entre las ventanas en dos etapas para el caso de la derivada primera con $p = 4$. Ventana primera etapa (en azul) y ventana segunda etapa (en rosa).

cuártico), nos podemos encontrar en la situación de que la varianza sea infinita. No obstante, el empleo de este tipo de funciones sigue resultando atractivo, ya que facilitan la labor de interpretación del modelo, además de permitir el desarrollo de algoritmos más rápidos en comparación con los algoritmos convencionales (en un orden de 50-100 veces en muchos de nuestros ejemplos con datos reales).

Con el objetivo de respetar las afirmaciones de Seifert y Gasser, hemos incluido en la rutina de suavizado, comprobaciones del tamaño mínimo requerido para el parámetro de suavizado h , a fin de que permita considerar los $p + 3$ puntos exigidos en el intervalo $x_0 \pm h$, utilizado para la estimación del propio punto x_0 .

4.6 Elección de ventana global o local

A lo largo de este capítulo, hemos comentado la posibilidad de usar dos tipos de ventanas plug-in: local y global. La ventaja de usar ventanas de tipo global es que requiere menos tiempo computacional para la estimación. En algunos casos, la ventana global puede presentar, como se ha comentado en la sección anterior, problemas de tipo frontera. En estos puntos frontera se puede optar por el empleo de ventanas de tipo local.

En principio, la elección de una ventana de tipo local parece mucho más correcta para la estimación de la regresión y sus derivadas, al considerar un suavizado distinto en cada punto a estimar. Sin embargo, el algoritmo computacional para su cálculo requiere más tiempo de ejecución, debido a que tiene que calcular una ventana nueva para cada valor considerado.

Es evidente que, usando un mismo parámetro de suavizado (ventana glo-

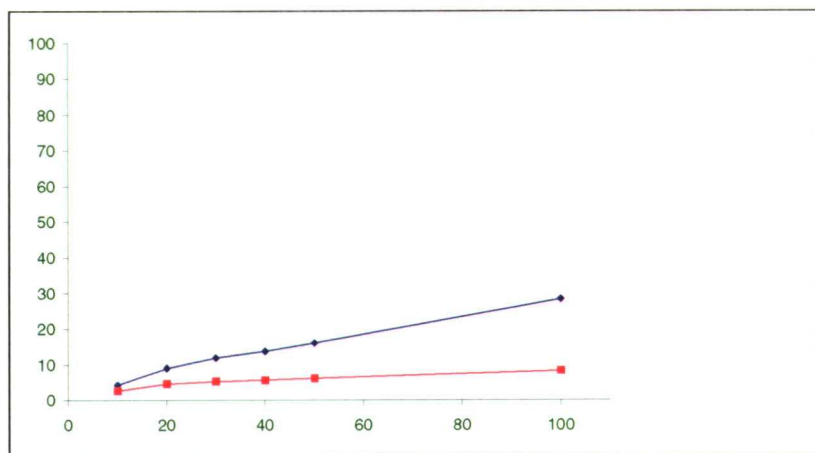


Figura 4.6: Relación entre las ventanas en dos etapas para el caso de la derivada segunda con $p = 5$. Ventana primera etapa (en azul) y ventana segunda etapa (en rosa).

bal) en lugar de uno diferente para cada punto se ahorra tiempo. Sin embargo, es necesario considerar algunos aspectos con respecto a la programación del algoritmo de la ventana global en aras de hacerlo más eficiente para el caso de datos equiespaciados, en el que nos centraremos en este apartado. En este caso, la matriz $(X^t W X)^{-1}$ no cambia cuando el estimador es evaluado en cada punto interior del diseño, x , tal que $x_1 < x - h$ y $x + h < x_n$. La razón es que las distancias entre x y los x_i 's que caen dentro del intervalo $[x - h, x + h]$ no cambian cuando $x \in (x_1 + h, x_n - h)$ y x es uno de los x_j de la muestra. Esto lleva a que la matriz $(X^t W X)_i^{-1}$ usada para calcular el estimador polinómico local para x_i no cambia para $i = \ell, \ell + 1, \dots, u - 1, u$ donde $\ell = \lceil h/\delta \rceil + 2$ y $u = n - \lceil h/\delta \rceil - 1$. Para cada i fuera de este rango, será necesario calcular explícitamente la matriz $(X^t W X)_i^{-1}$.

Con el objeto de agilizar los cálculos computacionales para los estimadores en cada punto $x = x_p$, obsérvese que cada elemento (i, j) -ésimo de la matriz $(X^t W X)$ se puede expresar como:

$$(X^t W X)_{ij} = \sum_{r=1}^n (x_r - x_p)^i W_r^{(p)} (x_r - x_p)^j = \sum_{r=1}^n (x_r - x_p)^{i+j} W_r^{(p)}, \quad (4.25)$$

donde $W_r^{(p)} = K_h(x_r - x_p)$ es el r -ésimo elemento de la diagonal de W . Usaremos una función núcleo con soporte en $[-1, 1]$, es decir, que verifique $K(u) = 0$, $\forall u \notin [-1, 1]$ y obtenemos que

$$\left| \frac{x_r - x_p}{h} \right| > 1 \implies W_r^{(p)} = 0,$$

o, equivalentemente:

$$W_r^{(p)} \neq 0 \implies x_r - x_p \in [-h, h] \iff (r - p)\delta \in [-h, h]$$

$$\iff r \in \left[p - \left\lceil \frac{h}{\delta} \right\rceil, p + \left\lceil \frac{h}{\delta} \right\rceil \right].$$

Definiendo los índices $s_1 = p - \lceil \frac{h}{\delta} \rceil$ y $s_2 = p + \lceil \frac{h}{\delta} \rceil$ podemos obtener una expresión para el cálculo rápido del (i, j) -ésimo elemento de la matriz $(X^t W X)$:

$$\sum_{r=s_1}^{s_2} (x_r - x_p)^{i+j} W_r^{(p)} = \frac{\delta^{i+j}}{h} \sum_{k=-\lceil \frac{h}{\delta} \rceil}^{\lceil \frac{h}{\delta} \rceil} k^{i+j} K\left(\frac{k\delta}{h}\right).$$

De forma similar, el término:

$$(X^t W Y)_i = \sum_{r=1}^n (x_r - x_p)^i W_r^{(p)} Y_r,$$

puede escribirse como:

$$\sum_{r=s_1}^{s_2} (x_r - x_p)^i W_r^{(p)} Y_r = \frac{\delta^i}{h} \sum_{k=-\lceil \frac{h}{\delta} \rceil}^{\lceil \frac{h}{\delta} \rceil} k^i K\left(\frac{k\delta}{h}\right) Y_{p+k}.$$

Esta implementación reduce el número de cálculos necesarios para obtener el estimador del orden de $O(n)$ hasta $O\left(\frac{h}{\delta}\right)$. Esta reducción es especialmente importante para ventanas pequeñas. En la práctica, en la mayoría de experimentos con datos termogravimétricos, el tiempo de reducción computacional ha sido del orden de 10 a 20 veces.

4.7 Aplicación a un caso práctico

Nuestro objetivo primordial era obtener unos resultados de suavización que superen los proporcionados por el software existente, que, como comentamos al principio de esta memoria, presenta importantes problemas de suavización.

A fin de constatar las mejoras que presenta nuestra rutina de suavizado, realizaremos una comparación de los resultados que obtenemos una vez aplicado nuestro algoritmo de suavización, en contraste con los obtenidos previamente. Para ello utilizaremos la muestra de oxalato de calcio.

En la figura 4.7, podemos observar, como ya se ha comentado, los graves problemas de suavizado en la estimación de la primera derivada, que en el caso de la segunda derivada todavía se incrementan más, no siendo posible apreciar ninguna tendencia. El software disponible (como el caso del Orchestrator), permite modificar manualmente el parámetro de suavizado y obtener una curva, en apariencia, mucho más precisa.

A simple vista parece que el Orchestrator nos ha permitido obtener una curva en base a la cual, podamos esta vez sí, obtener algún tipo de conclusión con un cierto rigor. Realmente, ésto no es así, observando con detalle esta curva (figura 4.8) se pueden apreciar graves deficiencias.

El problema que se presenta, deriva del hecho de haber realizado una sobresuavización de la curva (parámetro de suavizado demasiado grande). La gravedad de este hecho, queda resaltado de forma especialmente relevante en las zonas destacadas en la figura 4.9, ya que podemos observar como la profundidad de los picos se ha visto claramente reducida en comparación con la primera infrasuavización (figura 4.8). Ésto cobra una gran importancia debido a que el área comprendida en dichas zonas, resaltadas en la figura 4.9, aporta una información en cuanto a las pérdidas o liberaciones de energía durante el desarrollo del experimento.

En las figuras 4.10 y 4.11 presentamos nuestros resultados de suavización, mediante el empleo de ventana global o local. La primera de ellas permite visualizar la suavización obtenida mediante la variante de ventana global. Podemos constatar cómo el problema de sobresuavización de los picos ha sido resuelto y afirmar, a su vez, que la curva no presenta las graves deficiencias de suavización obtenidas mediante el Orchestrator en primera instancia.

En el supuesto del empleo de ventana local (figura 4.11), podemos ver que de nuevo los picos no presentan problemas de sobresuavización, y se obtiene una aparente mejora en la suavización de las zonas más planas de la curva con respecto a la variante de ventana global.

Finalmente, este hecho lo hemos contrastado con los picos de la curva DSC correspondiente, comprobando que en nuestro caso sí son del mismo orden de amplitud. Además, como comprobación adicional, hemos comparado la estimación proporcionada con nuestro método con el ajuste paramétrico logístico visto en el capítulo 3 de esta memoria, para muestras susceptibles de ser ajustadas correctamente por este modelo paramétrico, observando la aparente igualdad de las gráficas resultantes.

4.8 Estudio de simulación

En esta sección realizaremos un estudio de simulación con el objeto de estudiar cómo se comporta el estimador no paramétrico propuesto. Analizaremos diferentes grados de dependencia y estudiaremos las ventajas e inconvenientes del empleo de la ventana local o global en el suavizado.

Puesto que queremos simular diferentes grados de dependencia entre los errores, y estos sólo dependen del aparato que se esté usando, no del tipo de curva que se analice, centraremos nuestro estudio en la simulación de curvas TGA, para las cuales hemos visto en el capítulo 2 cómo obtener un modelo paramétrico adecuado, el de suma de funciones logísticas.

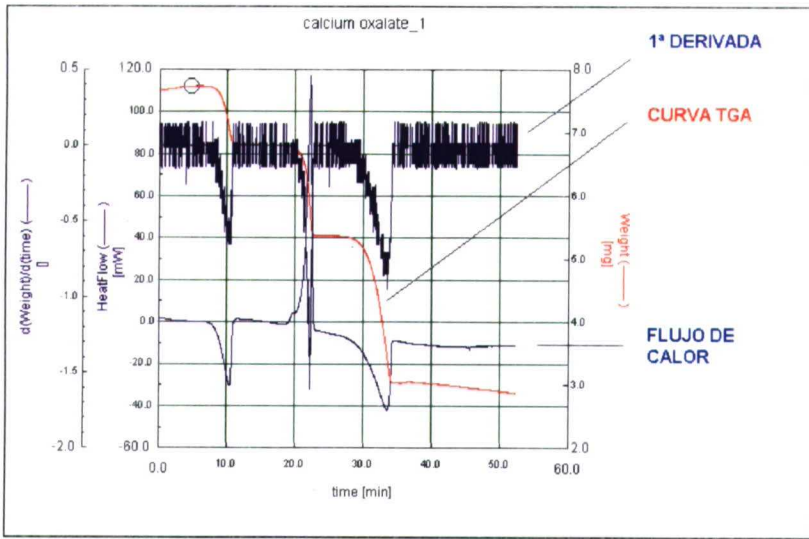


Figura 4.7: Gráficas del experimento con oxalato de calcio.

Con el fin de generar un modelo para la dependencia de los errores se aplicará un método similar al empleado en otros trabajos, concretamente, aplicaremos el también usado por Cao, Quintela y Vilar-Fernández (1993), entre otros.

Entre los distintos materiales que podríamos modelizar nos centraremos en aquellos que tienen una curva con varios escalones, como es el caso de algunos polímeros. Concretamente simularemos diferentes muestras para una función que refleja la forma que tiene la curva de PVC. En cualquier caso la función usada para la modelización paramétrica de una curva TGA de un material simple se modelizará como:

$$f(x) = \frac{e^x}{1 + e^x}$$

Partiendo de esta función f podemos obtener la función m mediante una composición de funciones f con diferentes pesos, mediante la expresión:

$$m(x) = \sum_{i=1}^k w_i f(a_i + b_i x) \tag{4.26}$$

donde: k es el número de componentes logísticas empleadas, w_i es el peso de la i -ésima componente, a_i es el parámetro de localización y b_i es el parámetro que indica la velocidad de la caída del peso en cada proceso.

En la figura 4.12, se representa la gráfica obtenida para tratar de aproximarse a una nube de puntos del termograma de PVC, resulta de la combinación de cuatro funciones logísticas:

$$m(x) = 5f(12 - 4x) + 4f(14 - 2x) + 7f(43 - 4.8x) + 1f(16 - x) \tag{4.27}$$

A fin de poder simular la primera y segunda derivada de la función a suavizar obtenemos a su vez la primera y segunda derivada de la función dada por (4.12), obteniendo las expresiones siguientes para la primera y segunda derivadas de m :

$$m'(x) = \sum_{i=1}^k w_i b_i f'(a_i + b_i x) \quad (4.28)$$

$$m''(x) = \sum_{i=1}^k w_i b_i^2 f''(a_i + b_i x) \quad (4.29)$$

A partir de las expresiones (4.28) y (4.29) se obtienen las gráficas de la primera y segunda derivada de la función m , respectivamente (ver figuras 4.13 y 4.14).

4.8.1 Generación de datos

El modelo empleado para generar los experimentos de simulación sigue la pauta $y_i = m(x_i) + \varepsilon_i$ siendo $m(x_i)$ la función suma de logísticas que hemos creado mediante la expresión (4.27), y donde ε_i son los errores del modelo.

Por tanto, a fin de obtener una simulación de la curva con ruido, hemos de seguir los siguientes pasos:

1. Siendo n el tamaño de la muestra, generar los valores x_1, x_2, \dots, x_n , a partir del valor de δ , que representa la distancia entre dos puntos consecutivos. De este modo podemos establecer:

$$x_i = \delta \cdot i, \quad i = 1, \dots, n$$

2. Una vez obtenidos dichos valores x_1, x_2, \dots, x_n , y a partir de la especificación de k componentes logísticas, podremos simular la señal $m(x_i)$ para $i = 1, 2, \dots, n$, según (4.27).
3. Se obtiene los valores de la señal con ruido; $Y_i = m(x_i) + \varepsilon_i$, donde ε_i son simulados imitando el error producido por el aparato STA. Estudios previos han permitido ajustar dicho error a un modelo $AR(1)$: $\varepsilon_i = \rho \varepsilon_{i-1} + a_i$, con $\rho \in (-1, 1)$ y los a_i son valores independientes simulados de una $N(0, \sigma_a^2)$, con $\sigma_a^2 = \sigma^2 (1 - \rho^2)$.

La simulación de cada componente a_i se ha llevado a cabo haciendo uso del método del Teorema Central del Límite (véase Cao (2002)).

4.8.2 Estimación en los extremos (efecto frontera)

Como comentamos en la sección 4.8, nos hemos encontrado con un problema a la hora de estimar los valores de la señal en los puntos extremos del intervalo, problema que ha sido sometido a un profundo estudio por parte de Seifter y

Gasser (1996), y que encuentra su fundamento teórico en la escasez de puntos a partir de los cuales obtener la estimación en un punto dado, x_0 , cuando, debido al valor del parámetro de suavizado h , $x_0 \pm h$ no contiene suficientes puntos, lo que se denomina efecto frontera.

Seifter y Gasser afirman que en dichos casos, con el fin de garantizar una varianza finita en la estimación, se ha de exigir la existencia de al menos $p + 3$ puntos en el intervalo de suavizado. Sin embargo, aunque imponiendo esta restricción garantizamos una varianza finita, ésta puede ser elevada, viéndose gravemente afectado el valor del *AMSE* (Average Mean Squared Error) que obtenemos mediante:

$$\widehat{AMSE} = \frac{1}{n} \sum_{i=1}^n \overline{MSE}(m(x_i)),$$

donde:

$$\overline{MSE}(m(x_i)) = \frac{1}{T} \sum_{j=1}^T (\hat{m}_j(x_i) - m(x_i))^2,$$

siendo n el número de instantes de muestreo de la señal, T el número de simulaciones realizadas y \hat{m}_j el estimador de la función de regresión construido con la j -ésima muestra simulada.

Teniendo en cuenta que el *MSE* se puede obtener a su vez a partir de:

$$MSE(m(x_i)) = [\text{Sesgo}(\hat{m}(x_i))]^2 + \text{Var}(\hat{m}(x_i)),$$

es obvio que la varianza introducida por los puntos extremos influye de forma clara en el valor obtenido del *AMSE*.

Con el objetivo de que los valores estimados en los puntos extremos de las señales (véase figura 4.18), no nos impidan realizar un estudio riguroso de los resultados obtenidos mediante las simulaciones, hemos incluido la posibilidad de no considerar dichos puntos en el cálculo del *AMSE* que denotaremos como *AMSE*^{sin}. Como consecuencia, en las simulaciones hemos permitido la posibilidad de fijar el número de puntos extremos a descartar (d) y expresar, por tanto, el *AMSE*^{sin} de la forma:

$$AMSE^{\text{sin}} = \frac{1}{(n - 2d)} \sum_{i=1+d}^{n-d} MSE(m(x_i))$$

Realizamos una ampliación del área indicada para poder observar con mayor detalle el problema reseñado. La figura 4.19 muestra los primeros 50 puntos, lo que nos permite apreciar cómo aparecen en el extremo inicial unos valores de la estimación claramente atípicos. De igual modo, podemos observar que el mismo tipo de problema se produce en los puntos extremos por la derecha de la señal (figura 4.20).

4.8.3 Resultados de las simulaciones

Hemos simulado un total de $T = 500$ muestras, de tamaño $n = 800$ con el fin de estudiar la influencia de los dos selectores de ventana propuestos en el capítulo 4. La varianza del error la hemos tomado como $\sigma^2 = 0.05$ y los diferentes valores del coeficiente de correlación se han variado desde -0.9 hasta 0.9 . Además se ha considerado que los datos están generados mediante un diseño fijo equiespaciado, con $\delta = 0.025$. El criterio usado para medir el ajuste ha sido el del promedio del error cuadrático medio:

$$AMSE = \frac{1}{n} \sum_{i=1}^n E \left((\hat{m}(x_i) - m(x_i))^2 \right) = E \left(\frac{1}{n} \sum_{i=1}^n (\hat{m}(x_i) - m(x_i))^2 \right),$$

que se ha estimado mediante la expresión:

$$\widehat{AMSE} = \frac{1}{n} \sum_{i=1}^n \frac{1}{T} \sum_{j=1}^T (\hat{m}_j(x_i) - m(x_i))^2,$$

donde $\hat{m}_j(x_i)$ es el estimador de la regresión en el i -ésimo punto del diseño usando la j -ésima muestra simulada. Para medir la influencia de la eliminación de puntos en torno a la frontera, hemos considerado una modificación mediante la versión \widehat{AMSE}^{\sin} :

$$\widehat{AMSE}^{\sin} = \frac{1}{(n-2d)} \sum_{i=1+d}^{n-d} MSE(m(x_i)),$$

donde d es el número de puntos que eliminamos en cada extremo. En la práctica tomaremos $d = 10$.

Emplearemos, además, la siguiente notación: $\widehat{AMSE}_{L,1}$ y $\widehat{AMSE}_{L,2}$ representan el error para el estimador polinómico local obtenido con el método plug-in en dos etapas y con ventana local incorporando la estimación de la suma de autocovarianzas mediante los métodos 1 y 2 (véanse (4.17) y (4.22)), respectivamente, introducidas en la sección 4.4. En el caso de usar ventana global emplearemos la notación $\widehat{AMSE}_{G,1}$ y $\widehat{AMSE}_{G,2}$. Similar notación usaremos para el caso de eliminar puntos $\widehat{AMSE}_{L,1}^{\sin}$, $\widehat{AMSE}_{L,2}^{\sin}$, $\widehat{AMSE}_{G,1}^{\sin}$ y $\widehat{AMSE}_{G,2}^{\sin}$.

En las tablas 4.4 a 4.9 se puede apreciar la tendencia decreciente del $AMSE$ y de la versión sin los extremos al disminuir el valor del coeficiente de correlación.

Las diferencias en tiempo de cálculo entre las ventanas global y local, son apreciables ya que la ventana global necesita menos tiempo de CPU (del orden de 6 veces menos) que la ventana local.

ρ	$\widehat{AMSE}_{L,1}$	$\widehat{AMSE}_{L,2}$
0.9	0.009300	0.009334
0.6	0.001036	0.001048
0.3	0.000381	0.000381
0	0.000203	0.000205
-0.3	0.000134	0.000133
-0.6	0.000097	0.000105
-0.9	0.000075	0.000078

Tabla 4.4. $AMSE$ para la ventana plug-in tipo local con los dos métodos.

ρ	$\widehat{AMSE}_{G,1}$	$\widehat{AMSE}_{G,2}$
0.9	0.010242	0.010230
0.6	0.001397	0.001396
0.3	0.000567	0.000564
0	0.000319	0.000317
-0.3	0.000210	0.000209
-0.6	0.000151	0.000151
-0.9	0.000122	0.000122

Tabla 4.5. $AMSE$ para las ventanas plug-in globales con los dos métodos sin puntos extremos.

ρ	$\widehat{AMSE}_{L,1}^{\sin}$	$\widehat{AMSE}_{L,2}^{\sin}$
0.9	0.009376	0.009357
0.6	0.001015	0.001028
0.3	0.000369	0.000368
0	0.000195	0.000197
-0.3	0.000128	0.000127
-0.6	0.000092	0.000101
-0.9	0.000071	0.000070

Tabla 4.6. $AMSE$ para las ventanas plug-in locales con los dos métodos sin puntos extremos.

ρ	$\widehat{AMSE}_{G,1}^{\sin}$	$\widehat{AMSE}_{G,2}^{\sin}$
0.9	0.010080	0.010072
0.6	0.001392	0.001393
0.3	0.000562	0.000565
0	0.00316	0.000317
-0.3	0.000208	0.000209
-0.6	0.000151	0.000151
-0.9	0.000119	0.000119

Tabla 4.7. $AMSE$ para las ventanas plug-in globales con los dos métodos sin puntos extremos para la primera derivada.

ρ	$\widehat{AMSE}'_{L,1}$	$\widehat{AMSE}'_{L,2}$
0.9	0.185815	0.354937
0.6	0.038992	0.166503
0.3	0.014042	0.092983
0	0.007527	0.059680
-0.3	0.004930	0.043556
-0.6	0.003606	0.036175
-0.9	0.002838	0.058530

Tabla 4.8: Resultados para ventanas locales al estimar la primera derivada.

ρ	$\widehat{AMSE}''_{L,1}$	$\widehat{AMSE}''_{L,2}$
0.9	7.921071	8.485327
0.6	1.903320	2.016023
0.3	0.704224	0.952018
0	0.392530	0.536148
-0.3	0.268913	0.385504
-0.6	0.205573	0.310128
-0.9	0.169449	0.285341

Tabla 4.9: Resultados para ventanas locales al estimar la segunda derivada.

Observando los valores de las tablas anteriores (tablas de 4.4 a 4.9) se puede apreciar la disminución del $AMSE$ al pasar de dependencia grande entre los datos ($\rho = 0.9$) a independencia ($\rho = 0$). Para el caso de dependencias negativas se observa cómo disminuye al aumentar el valor absoluto de la correlación.

Como consecuencia del estudio de simulación hemos tomado las siguientes decisiones:

- Empleo del método 1 para la estimación de la suma de autocovarianzas en detrimento del método 2.
- Posibilidad de trabajar con ventana local o global, en función de los resultados que se quieran obtener, es decir, basándose en un criterio de precisión frente a tiempo de ejecución.

4.8.4 Algunos resultados gráficos

Mostramos a continuación una serie de gráficas, a modo de ejemplo, donde podemos observar las funciones que obtenemos mediante simulación, partiendo del modelo (4.26) y variando los valores de los parámetros ρ y σ^2 , es decir, variando el factor de dependencia o correlación de orden uno y la varianza del error.

En la gráfica de la figura 4.15 hemos utilizado un valor de $\rho = 0.25$ con $\sigma^2 = 0.007$. Como vemos en la vista general de la función, el error introducido es inapreciable, por lo que realizaremos una ampliación de la curva en el punto indicado.

A fin de observar la relevancia de la dependencia en el modelo de simulación, analizamos a continuación los efectos producidos por la modificación de dicho parámetro. Inicialmente hemos tomado un valor de $\rho = 0.9$, que indica una dependencia positiva muy elevada, lo que contribuye a que el error en un punto mantenga la tendencia del anterior en este caso la curva con error se mantiene por debajo de la original a lo largo de todo el intervalo de definición. En el caso contrario, es decir tomando un valor del coeficiente de correlación negativo, $\rho = -0.9$, que indica una dependencia negativa elevada, observemos cómo la curva con error oscila en torno a la original a lo largo de todo el intervalo de definición (figura 4.17). Por otra parte la influencia de la varianza en el modelo también queda patente en la zona ampliada de las figuras 4.16 y 4.17, observamos como es lógico, un incremento considerable del ruido al incrementar la varianza.

Finalmente, presentamos otras salidas gráficas en las que se reflejan parte de los resultados de las simulaciones comentados y ya expuestos en las tablas 4.4 a 4.9.

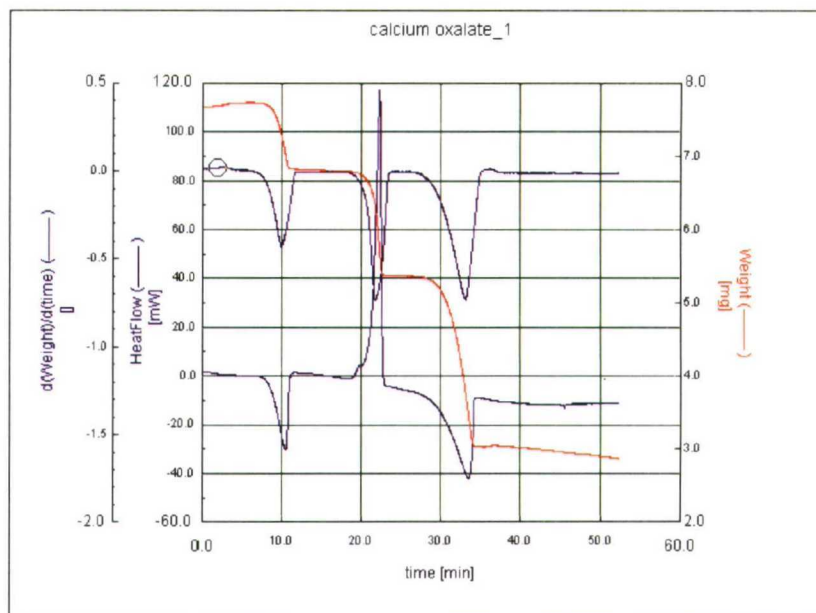


Figura 4.8: Suavizado manual de los datos con el software existente hasta el momento.

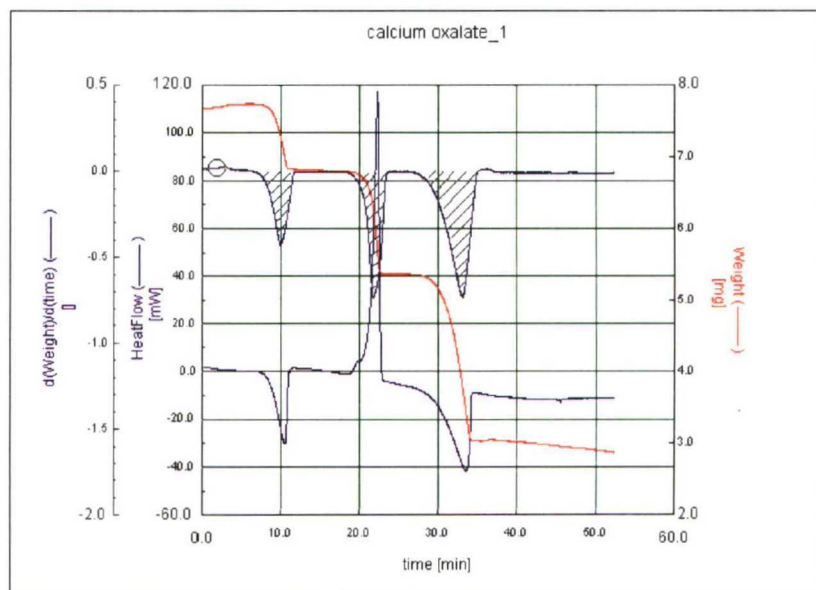


Figura 4.9: Zonas de interés en la derivada de la curva TGA.

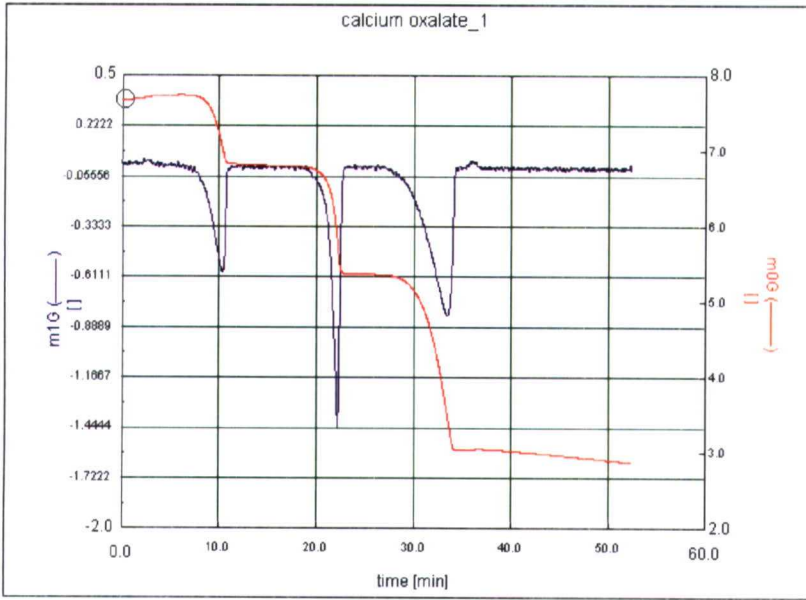


Figura 4.10: Suavizado de la primera derivada de la curva TGA con ventana global.

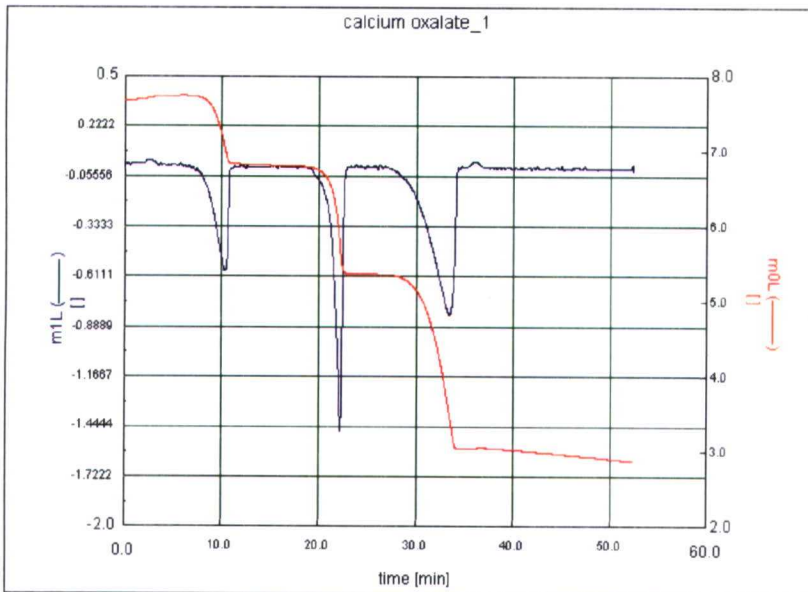


Figura 4.11: Estimación de la primera derivada con ventana local.

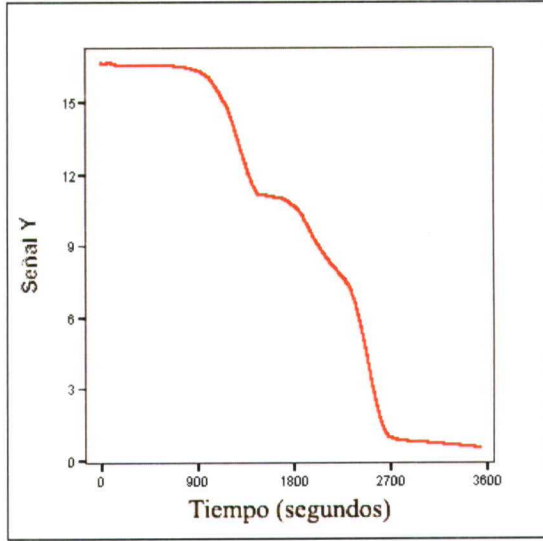


Figura 4.12: Función simulada tipo PVC.

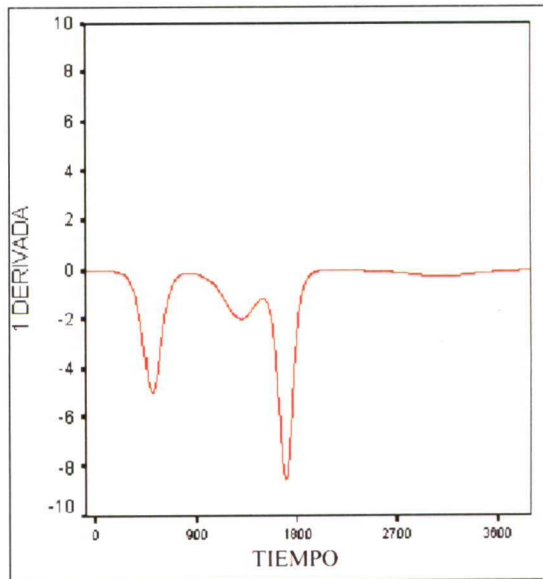


Figura 4.13: Primera derivada de la función simulada.

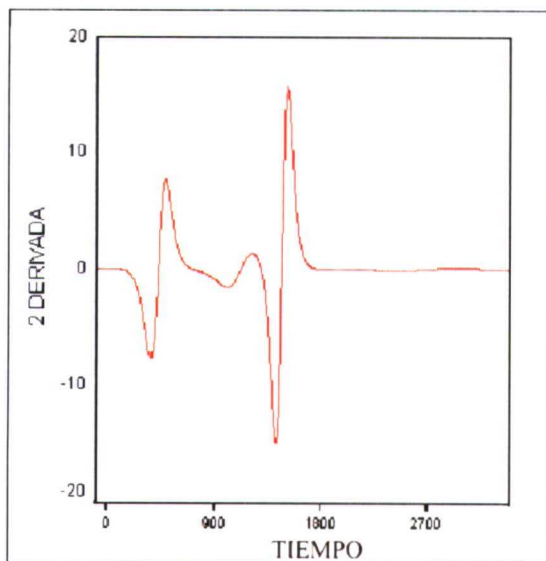


Figura 4.14: Segunda derivada de la función simulada.

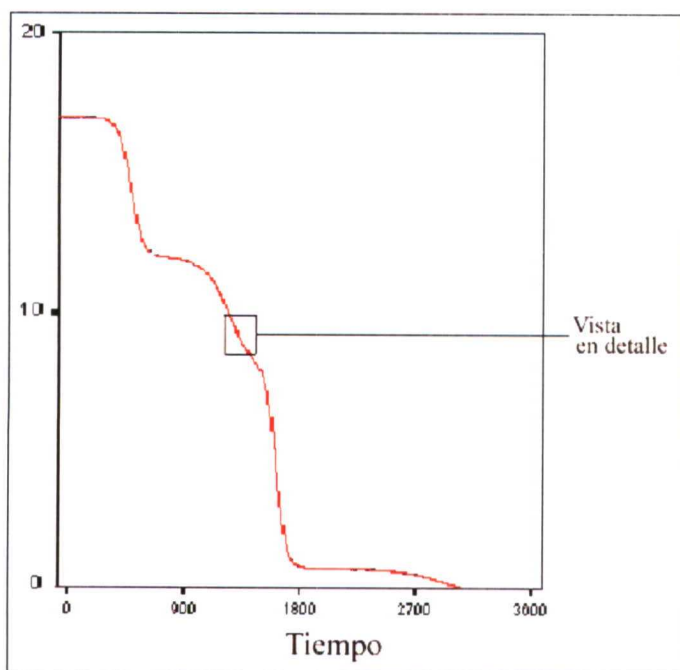


Figura 4.15: Datos simulados con ruido.

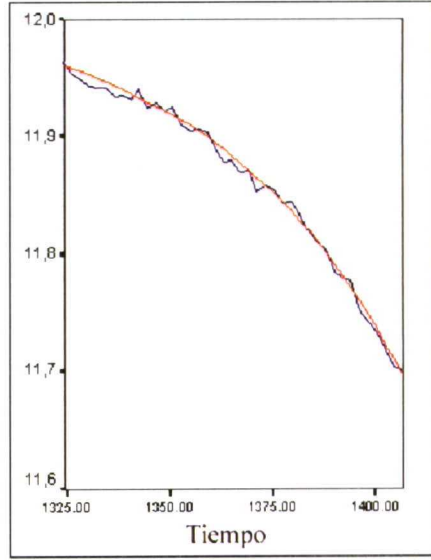


Figura 4.16: Zona ampliada de la curva con ruido ($\rho = 0.25$ y $\sigma^2 = 0.007$).

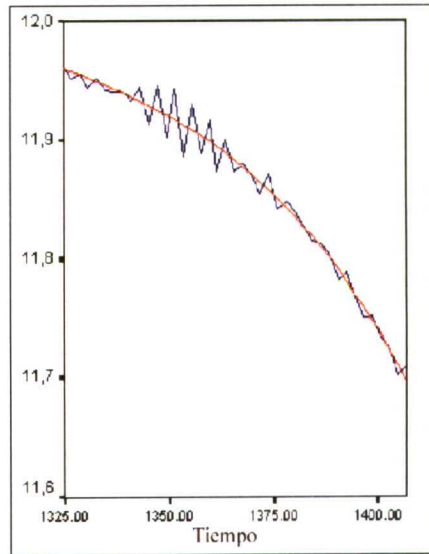


Figura 4.17: Zona ampliada con dependencia negativa ($\rho = -0.9$ y $\sigma^2 = 0.05$).

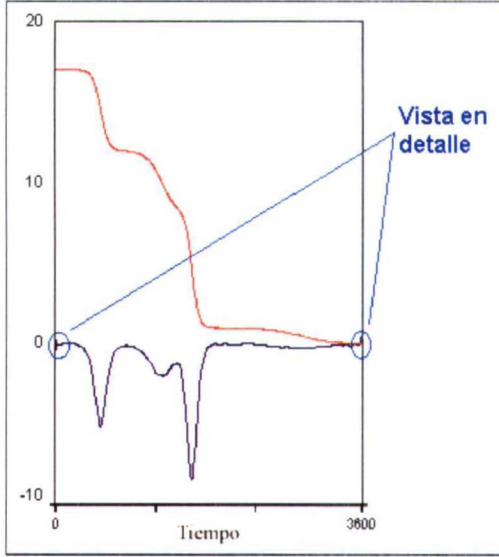


Figura 4.18: Función simulada y su primera derivada.

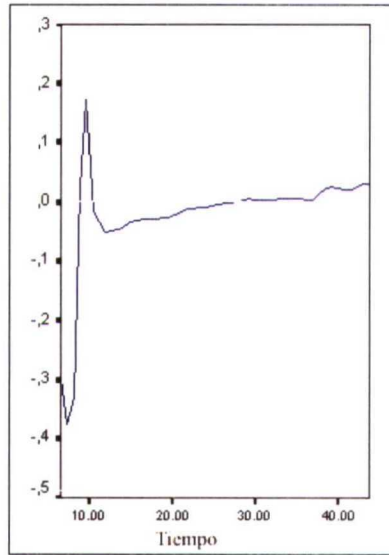


Figura 4.19: Detalle de la zona del inicio (efecto frontera).

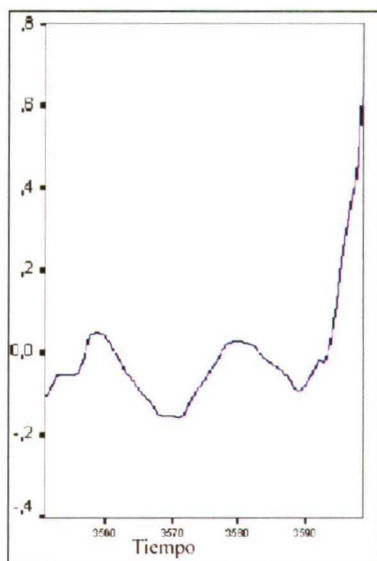


Figura 4.20: Detalle del final de la curva (efecto frontera).

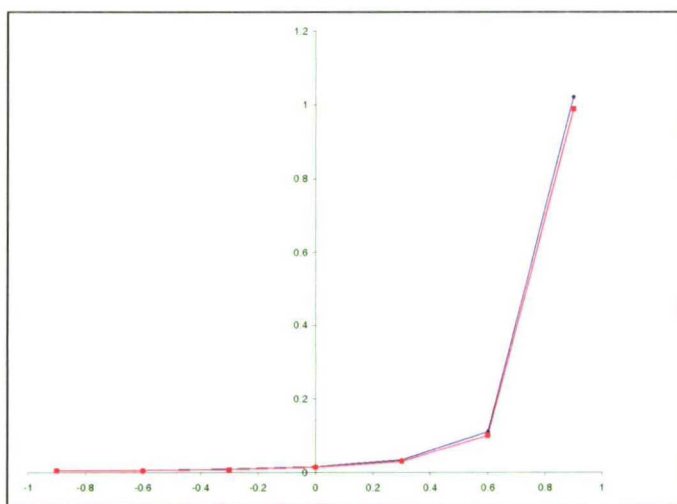


Figura 4.21: Comparación de los dos métodos con ventanas locales para la estimación de la regresión. En rosa *AMSE* con el método 2 y en azul *AMSE* con el método 1.

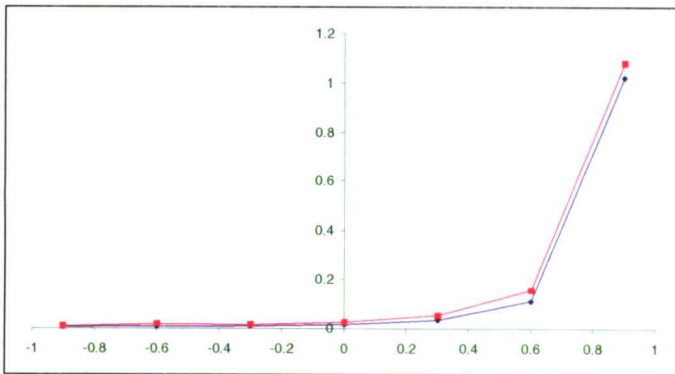


Figura 4.22: Comparación del $AMSE$ para ventana local frente a global para la estimación de la regresión. En rosa $AMSE$ con ventana global y en azul $AMSE$ con ventana local.

Capítulo 5

Clasificación de materiales

5.1 Introducción

Uno de los problemas a los que se enfrentan los investigadores es el de clasificación de los materiales y, para este fin, suelen usar curvas procedentes de espectrometría. Los inconvenientes que tiene el manejo de estas curvas es, por un lado, el elevado coste de estos equipos, y por otro, la dificultad de clasificar ciertos materiales debido al requerimiento de que las muestras a estudio deben ser lo más homogéneas posibles.

Puesto que las curvas TGA permiten medir la masa de la muestra mientras está siendo sometida a un programa térmico, describen el comportamiento del material al ser calentado, siendo únicas para cada clase de material. Esta propiedad nos llevó a la idea de usar estas señales para la clasificación de los materiales.

En este capítulo haremos una introducción a distintas técnicas de clasificación con el objeto de presentar un método que nos permita la clasificación de materiales. En el contexto de clasificación, necesitamos tener una información de muestras de las que se conozca su procedencia para, a partir de ellas, poder clasificar otras nuevas. En nuestro caso, trabajaremos con distintas muestras de madera y de PVC, si bien, podría extrapolarse el estudio a cualquier material susceptible de ser tratado mediante análisis térmico. Presentaremos un método de clasificación basado en la regresión no paramétrica funcional, utilizando como datos las curvas TGA. Finalmente, con el fin de estudiar el comportamiento del método de discriminación propuesto, se realizará un estudio de simulación.

5.2 Clasificación de datos

El problema general de clasificación estudia cómo asignar una observación a una clase de un conjunto finito de ellas. Este problema recibe diferentes nombres según la perspectiva de estudio: en informática se conoce como teoría

de aprendizaje, en ingeniería se suele denominar reconocimiento de patrones y en estadística se engloba dentro del problema de clasificación.

En el campo de la informática el problema de clasificación se engloba dentro de la teoría del aprendizaje, dividiéndolo en dos fases: *entrenamiento* y *aprendizaje*. En la fase de entrenamiento, se diseña un extractor de características para representar los patrones de entrada y se entrena al clasificador con un conjunto de datos, llamados de entrenamiento, de forma que el número de patrones mal identificados se minimice. En el modo de reconocimiento, el clasificador ya entrenado toma como entrada un nuevo elemento desconocido y lo asigna a una de las clases. En este campo suele emplearse técnicas denominadas de neurocomputación cuya base está en la teoría de las redes neuronales artificiales (véase Werbos (1991)).

En el contexto de la ingeniería un patrón se define (Watanabe (1985)) como una entidad a la que se le puede dar un nombre y que está representada por un conjunto de propiedades medibles; como ejemplo de patrón podría considerarse las curvas TGA. Un sistema de reconocimiento de patrones tiene dos objetivos: identificar el patrón como miembro de una clase predefinida (*clasificación supervisada*) o asignar el patrón a una clase todavía no definida (*clasificación no supervisada o clustering*). En cualquier caso, el diseño de un sistema de reconocimiento de patrones requiere tres fases: adquisición y procesamiento de datos, extracción de características y toma de decisiones.

En el contexto de la estadística, las técnicas de clasificación se estudian dentro del análisis discriminante. El análisis discriminante se utiliza para clasificar individuos en grupos o poblaciones a partir de los valores de un conjunto de variables sobre los individuos a los que se pretende clasificar (Uriel (1991)). La pertenencia a uno u otro grupo se introduce en el análisis mediante una variable categórica, que toma tantos valores como grupos existentes, jugando también el papel de variable dependiente. Al igual que ocurría en el campo de la regresión, el análisis discriminante puede hacerse en un contexto paramétrico o no paramétrico.

En el análisis discriminante paramétrico existen varios métodos, destacando, entre ellos, el análisis discriminante lineal de Fisher, el análisis factorial discriminante (FDA), el análisis discriminante con penalización (PDA) propuesto por Hastie et al. (1995) o su versión generalizada propuesta por Marx y Eilers (1999).

Dentro de los modelos de tipo no paramétrico, que no necesitan de la estimación previa de ningún parámetro, destacaremos el método discriminante tipo núcleo, que está dando excelentes resultados en su aplicación a la clasificación de curvas, como constatan Ferraty y Vieu (2003). Estos autores aplican varias técnicas a 160 espectros, obtenidos del análisis de muestras de carne, con el fin de clasificarlos por su contenido en grasa. En su trabajo comparan métodos conocidos, como el FDA o el PDA, con el método no paramétrico discriminante (NPCD), resultando éste mucho mejor, en cuanto a presentar menor probabilidad de clasificación incorrecta.

En nuestro caso, puesto que hemos visto que la estimación paramétrica de las curvas TGA puede no ser adecuada para algunos materiales, y teniendo en

cuenta que disponemos de un procedimiento automático para la estimación no paramétrica de dichas curvas (expuesto en el capítulo 4), hemos optado por un modelo no paramétrico de clasificación.

En cualquier caso, el modelo matemático se puede expresar en términos de una variable aleatoria funcional, X , que toma valores en el espacio de posibles curvas TGA, y una variable aleatoria Y que toma valores en un conjunto finito de clases $\{0, 1, \dots, G\}$.

5.2.1 Reglas de clasificación

Una regla de clasificación es una función ξ que asigna a cada observación una clase. Cuando la observación consta de d componentes se trata de:

$$\xi : R^d \rightarrow \{0, 1, \dots, G\}.$$

La calidad de una determinada regla de clasificación se medirá mediante la probabilidad de cometer un error en la clasificación:

$$L(\xi) = P(\xi(X) \neq Y).$$

Se tratará de minimizar $L(\xi)$. En el caso en que tengamos sólo dos posibles clases, $G = 1$, y definiendo $\eta(x) = P(Y = 1/X = x)$ la *regla de clasificación de Bayes*, ξ^* , definida como:

$$\xi^*(x) = \begin{cases} 1 & \text{si } \eta(x) > \frac{1}{2} \\ 0 & \text{si } \eta(x) \leq \frac{1}{2} \end{cases}$$

resulta ser óptima, en el sentido de que para cualquier otra posible regla ξ se verifica que $L(\xi) \geq L(\xi^*)$. Al valor $L(\xi^*)$ suele denominarse *error Bayes*.

Obsérvese que en el caso de dos clases, $G = 1$, se tiene que:

$$\eta(x) = E(Y/X = x) = P(Y = 1/X = x),$$

es una función de regresión. Si pudiésemos estimar esta función de regresión tendríamos resuelto el problema de la clasificación.

5.2.2 Regla de máxima verosimilitud

Supóngase conocida la distribución del vector aleatorio de interés condicionada a cada uno de los grupos: $\vec{X}|_{Y=i}$, $i = 0, 1, \dots, G$, y denótese por $L_i(\vec{x})$ el valor de la función de verosimilitud (función de masa de probabilidad, en el caso discreto, o de densidad, en el caso continuo) del grupo i , es decir $L_i(\vec{x}) = P(\vec{X} = \vec{x}|_{Y=i})$, o bien $L_i(\vec{x}) = f(\vec{x}|_{Y=i})$. La regla discriminante de

máxima verosimilitud consiste en asignar a cada posible futura observación, \vec{x} , el índice j para el cual

$$L_j(\vec{x}) = \max_{i \in \{0, 1, \dots, G\}} L_i(\vec{x}),$$

decidiendo arbitrariamente en caso de empate en el máximo.

5.2.3 Regla Bayes

La regla Bayes se basa en calcular las probabilidades de pertenencia, a posteriori, a cada clase, dada una nueva observación:

$$P(Y = i/X = x) = \frac{f(X/Y = i) p_i}{\sum_{j=0}^G f(X/Y = i) p_j}. \quad (5.1)$$

donde $f(X/Y = i)$ es la función de densidad condicionada de la variable X dado el valor $Y = i$ y p_j es la probabilidad de pertenencia a la clase j de una observación elegida al azar de la población. En el caso discreto las funciones de densidad condicionadas $f(X/Y = i)$ serán sustituidas por las probabilidades condicionadas $P(X = x/Y = i)$.

La clasificación de cada individuo se puede realizar mediante la comparación de las probabilidades “a posteriori” dadas en (5.1). Así, se asignará un individuo al grupo para el cual sea mayor su probabilidad a posteriori.

La regla discriminante de máxima verosimilitud y la regla Bayes pueden extenderse fácilmente a un contexto no paramétrico, sustituyendo los estimadores paramétricos de la función de masa de probabilidad o de densidad del vector aleatorio por estimadores no paramétricos.

Como se ha comentado anteriormente, nuestro interés está en obtener un método que permita la clasificación de materiales a partir de sus termogramas, más concretamente, partiremos de un conjunto de materiales para los que conocemos a qué grupo pertenecen e intentaremos, dado un nuevo material, mediante su curva TGA, asignarlo a una de las clases. Puesto que trabajaremos con curvas como variables clasificadoras y pretendemos usar un método discriminante no paramétrico, haremos una pequeña introducción, en la siguiente sección, a la regresión no paramétrica funcional.

5.3 Modelo no paramétrico de regresión funcional

Dada una variable respuesta aleatoria real Y y otra variable explicativa X que toma valores en un espacio vectorial, en el que se ha definido una seminorma $\|\cdot\|$ deseamos estimar el operador funcional r definido por $Y = r(X) + \varepsilon$, donde ε es una variable real de media cero e independiente de X y r es suave.

Una vez medida la distancia entre curvas mediante una seminorma, cuya elección es un problema de índole no menor, se enfocará la clasificación como una consecuencia del problema de regresión de tipo funcional, en el que para estimar r se hará uso del estimador tipo núcleo $\hat{r}_h(x)$ siguiente:

$$\hat{r}_h(x) = \frac{\sum_{i=1}^n Y_i K\left(\frac{\|x-X_i\|}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{\|x-X_i\|}{h_n}\right)} \quad (5.2)$$

En la expresión anterior $\hat{r}_n(x)$ es una versión del estimador de Nadaraya-Watson para regresión funcional y $h_n(x)$ representa el parámetro ventana que puede depender de cada valor x , en el contexto de estimación local, y que también se puede expresar en su versión global. Este estimador ha sido usado por autores como Ferraty y Vieu (2002a) para el estudio de espectros donde han utilizado el núcleo de Epanechnikov $K(x) = \frac{3}{4}(1-x^2)I_{(|x|\leq 1)}$.

En la formulación del problema, en el contexto de la estimación de la regresión, se considera que la variable Y es de tipo escalar y la variable X es de tipo funcional con una seminorma asociada $\|\cdot\|$. Un caso particular de este modelo es aquel en que X pertenece a un espacio de Hilbert (véase Ramsay y Silverman (1997), como es el de las funciones continuas en un intervalo, caso también de las curvas TGA usadas en este trabajo. ($X \in L^2([0, T])$).

La idea de la estimación consiste en asumir condiciones de regularidad para r y utilizar un suavizado tipo núcleo. Para medir la proximidad entre curvas utilizaremos una seminorma $\|\cdot\|$. Al igual que en la estimación no paramétrica real, en este contexto funcional, es también crucial la adecuada selección de la ventana.

5.3.1 Selección de la ventana de suavizado

Entre los métodos existentes para estimar la ventana óptima está el de validación cruzada de mínimos cuadrados, propuesto por Rudemo (1982) y Bowman (1984), este método consiste en seleccionar el valor de h que minimice una estimación del error cuadrático integrado ($ISE(\hat{r}_h)$).

$$ISE(\hat{r}_h) = \int (\hat{r}_h(x) - r(x))^2 dx.$$

El parámetro ventana será aquel h que minimice la función de validación cruzada general $CV(h)$

$$CV(h) = n^{-1} \sum_{j=1}^n \left(Y_j - \hat{r}_h^{-j}(X_j) \right)^2, \quad (5.3)$$

siendo \hat{r}_h^{-j} el estimador calculado con la muestra sin el dato (X_j, Y_j) . Härdle y Marron (1985) demuestran la optimalidad asintótica de la ventana elegida por este criterio en el caso de una variable explicativa real.

En nuestro caso, puesto que lo que nos interesa es clasificar adecuadamente los materiales, utilizaremos una modificación del método de validación cruzada basado en minimizar la probabilidad de clasificación incorrecta que denotaremos por $\widetilde{CV}(h)$. Esta probabilidad dependerá de la regla de decisión empleada d . Usaremos la regla que maximice la probabilidad estimada de pertenencia a posteriori ($\widehat{r}_h^{(j)}$), lo que equivale al empleo de la regla Bayes no paramétrica.

Estimaremos $\widehat{r}_h^{(0)}, \widehat{r}_h^{(1)}, \dots, \widehat{r}_h^{(G)}$ para una rejilla de valores de h , para finalmente calcular el máximo de estas estimaciones en el punto en cuestión:

$$d_h(x) = \arg \max_{0 \leq j \leq G} \left\{ \widehat{r}_h^{(j)}(x) \right\}.$$

Finalmente, se definirá la función de validación cruzada modificada como

$$\widetilde{CV}(h) = n^{-1} \sum_{j=1}^n 1_{\{Y_j \neq d_h^{-j}(X_j)\}}, \quad (5.4)$$

donde d_h^{-j} denota la regla de clasificación construída sin la j -ésima observación.

La ventana de validación cruzada, que denotaremos por h_{CV} , será aquella que minimice la expresión anterior. Obsérvese que esto trata de estimar aquel parámetro que minimiza la probabilidad de clasificar mal una futura observación.

5.3.2 Elección de la seminorma

Otro problema distinto es encontrar cuál es la “distancia” que debemos usar de cara a que curvas del mismo material no presenten distancias elevadas y sí las tengan pares de curvas procedentes de materiales distintos. En el estimador usado, $\widehat{r}_h(x)$, la distancia entre las funciones, curvas TGA en nuestro ejemplo, se mide a través de una seminorma $\|\cdot\|$.

En nuestro caso la elección de la seminorma dependerá de los datos a usar. En un principio usaremos la norma L_1 , pero en caso de observar problemas de cambios de escala intentaremos ver qué mejoras se obtienen con el uso de las “distancias” entre la derivadas (hasta un orden k) de las curvas TGA, en vez de usar la propia función de regresión, para ello se podrían usar seminormas que dependan de la derivada v -ésima, que podemos agrupar dentro de la familia N_k siguiente:

$$N_k = \{ \|\cdot\|_v ; v = 0, 1, 2, \dots, k \} \text{ con } \|f\|_v = \left(\int \left(f^{(v)}(t) \right)^2 dt \right)^{1/2}.$$

Como caso particular está el caso $v = 0$, denominada norma euclídea o norma L_2 :

$$\|f\|_0 = \left(\int f(t)^2 dt \right)^{1/2}.$$

5.4 Programa para clasificar materiales

Con el objeto de implementar nuestro método de clasificación de curvas es necesario realizar una serie de cálculos que a continuación se detallan.

5.4.1 Cálculo de las distancias entre las curvas

En primer lugar tenemos que hallar las distancias entre todas las curvas a clasificar. Para este fin necesitamos realizar una serie de pasos. En un principio, las diferentes curvas TGA no están evaluadas sobre el mismo conjunto de puntos, por tanto, un primer paso consiste en situar las gráficas en una misma rejilla (o partición). Para el eje de abscisas, que correspondería con la variable tiempo, elegimos $m = 3000$ puntos; para el de ordenadas situamos todas las muestras en función del tanto por ciento de peso, reescalando el peso de la muestra considerada al 100%. Finalmente, después de estudiar alguna situación real, hemos considerado que es suficiente considerar las curvas hasta que pierden el 60% de su masa, por lo que se ha eliminado del estudio la parte final de cada espectro.

A continuación hemos procedido al cálculo de las distancias entre todas las curvas a estudio, para lo que usamos la norma L_1 :

$$\|f - g\| = \int |(f(t) - g(t))| dt \simeq \sum_{i=2}^m |f(t_i) - g(t_i)| (t_i - t_{i-1}),$$

siendo $t_1 < t_2 < \dots < t_m$ los puntos en cuestión elegidos, de forma única, para el conjunto de todas las curvas.

Finalmente, con el objeto de calcular las diferencias $|f(t_i) - g(t_i)|$, se han interpolado los valores de aquellas curvas TGA de la muestra que no estaban evaluadas en cada punto de la partición $t_1 < t_2 < \dots < t_m$.

5.4.2 Determinación de la ventana de validación cruzada

Una vez halladas las distancias para cada par de curvas, esto es, $\|X_i - X_j\|$, para cada $i \neq j$, se procede al cálculo de los valores $\hat{r}_h^{(-i)(k)}$ con el fin de calcular la ventana por el método de validación cruzada:

$$\hat{r}_h^{(-i)(k)} = \frac{\sum_{j=1, j \neq i}^n K \left(\frac{\|X_i - X_j\|}{h} \right) 1_{\{Y_j=k\}}}{\sum_{j=1, j \neq i}^n K \left(\frac{\|X_i - X_j\|}{h} \right)}, \quad (5.5)$$

para cada $k \in \{0, 1, \dots, G\}$, haciendo que h recorra una rejilla de valores. Concretamente hemos elegido diferentes valores para h , entre el mínimo de todas las distancias $\|X_i - X_j\|$ para $i \neq j$, y el triple del máximo de estas distancias, tomados a intervalos de 1.5 veces el anterior (es decir, una partición con paso multiplicativo),

$$h \in \left[\min_{i \neq j} \|X_i - X_j\|, 3 \max \|X_i - X_j\| \right].$$

Finalmente, se elegirá aquel h que minimice la función $\widetilde{CV}(h)$ definida en (5.4), que resulta ser $n^{-1} \sum_{j=1}^n 1_{\{Y_j \neq d_h^{-j}(X_j)\}}$ donde:

$$d_h^{-j}(X_j) = \arg \max_{0 \leq k \leq G} \left\{ \widehat{r}_h^{(-j)(k)} \right\},$$

eligiendo la media entre el mínimo y el máximo valor de la ventana que presenten un valor mínimo de $\widetilde{CV}(h)$ en caso de haber más de uno.

Conviene resaltar que las distancias entre las curvas TGA, $\|X_i - X_j\|$ para $i \neq j$, se calculan una sola vez, y se almacenan en una matriz al objeto de evitar cálculos repetitivos de distancias entre un mismo par de curvas (obsérvese que finalmente tendremos un método iterativo que va “probando” con distintos valores de h). La función núcleo utilizada ha sido la de Epanechnikov.

5.4.3 Clasificación de un nuevo material

Dada una nueva muestra que nos interese clasificar, calcularemos las distancias entre su curva TGA, previamente reescalada, con las demás curvas de nuestra base de datos. Así, si llamamos x a esta nueva curva a clasificar, calcularemos $\|x - X_j\|$, para $j = 1, 2, \dots, n$, a continuación calculamos $\widehat{r}_h^{(k)}(x)$ para cada clase de valores $k \in \{0, 1, \dots, G\}$, eligiendo aquella clase k que maximice $\widehat{r}_h^{(k)}(x)$.

Este método de clasificación se ha implementado en lenguaje C, formando parte del Proyecto de Fin de Carrera de Ingeniería Informática de Carlos Dominguez Salgueiro (Dominguez (2003)).

5.5 Aplicación a muestras de maderas

5.5.1 Características específicas de las muestras a analizar

Las muestras de madera presentan una gran dificultad de clasificación, debido a la gran variabilidad existente dentro de un mismo grupo: humedad diferente, edad, tiempo desde el corte, etc.

Las maderas forman parte del grupo de materiales lignocelulósicos. Sus componentes principales son: celulosa, hemicelulosa y lignina. Los estudios térmicos de cada componente por separado vienen realizándose desde la década de los 60, destacando los trabajos de Chatterjee y Conrad (1970), que sientan las bases de estudios posteriores. Estos autores analizan la pirólisis de la celulosa, que a temperaturas por debajo de 280°C se ve favorecida por la formación de residuos sólidos y gases, mientras que a temperaturas superiores se favorece la formación de alquitranes.

Una cuestión tratada ampliamente es si es posible representar la pirólisis de materiales lignocelulósicos como la suma simple de sus componentes, o si estos componentes interaccionan físicamente, provocando que el material lignocelulósico se comporte de forma única durante su degradación térmica.

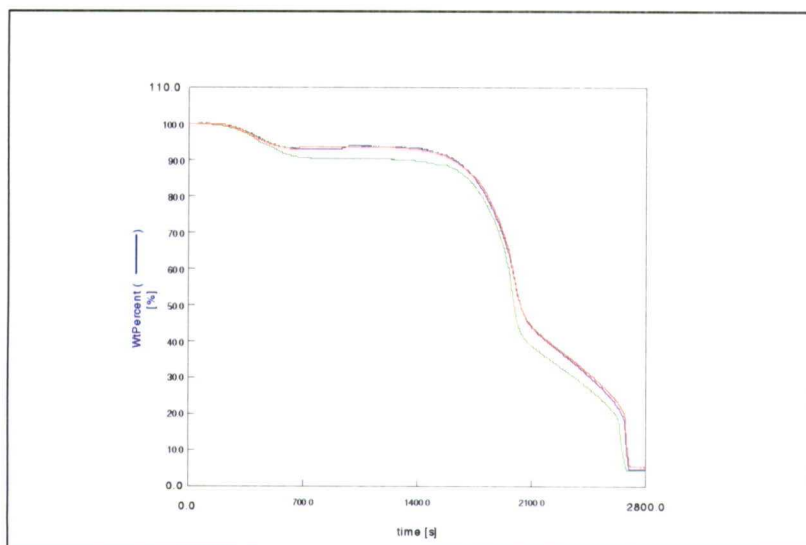


Figura 5.1: Comparación de tres muestras de la misma madera (pino), una con distinta humedad (verde).

Los estudios de la cinética de celulosa, lignina y hemicelulosa por separado han revelado que las interacciones entre las fracciones son importantes, y el comportamiento de la madera en su pirólisis no se puede considerar una adición de sus componentes. Además el proceso de separación de la lignina, celulosa y hemicelulosa puede alterar la estructura del material de partida. En particular Chatterjee y Conrad (1970) afirman que es imposible aislar la lignina de la madera sin cambiar su estructura, incluso utilizando el mismo método, es difícil obtener muestras idénticas. Sin embargo, en el análisis en la termobalanza de materiales lignocelulósicos generalmente aparecen dos o tres picos (o procesos), que pueden identificarse con hemicelulosa, celulosa y lignina, indicando que, aunque existen interacciones entre las fracciones, se mantiene su identidad (véase figura 5.1). Estos comentarios ponen de relieve la dificultad de establecer un buen método para la clasificación de tipos de madera.

Aplicación a un caso práctico

Partiendo de un total de 19 muestras de 7 clases de madera, hemos verificado nuestro programa de clasificación de materiales en base a los espectros de sus curvas TGA. Tenemos distintas muestras de cada tipo de madera que se mencionan en la tabla 5.1.

Grupo	Clase	Muestras
1	castaño	4
2	ciprés	3
3	higuera	4
4	manzano	2
5	olivo	2
6	pino	2
7	roble	2

Tabla 5.1: Muestras de madera estudiadas.

Se han utilizado tres tipos de distancias entre las distintas combinaciones de pares de las 19 muestras: la distancia entre las curvas TGA, la distancia entre las derivadas primeras de estas curvas y la distancia entre las derivadas de segundo orden. Hemos comprobado que no existe gran diferencia en el empleo de estas distancias.

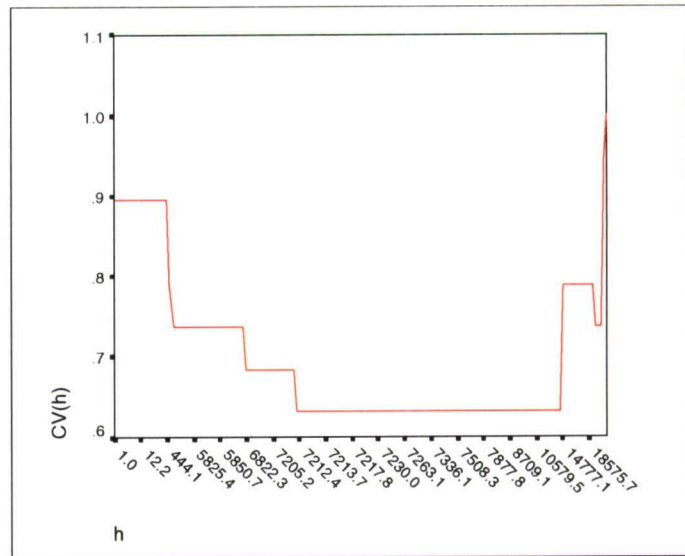


Figura 5.2: Función de validación cruzada para las maderas.

En la figura 5.2 se puede observar la función de validación cruzada para este ejemplo, en la que para la ventana óptima se obtiene un valor 0.64 de $\widetilde{CV}(h)$. En un principio, nos ha parecido que estos resultados no eran lo bueno que cabe esperar, sin embargo, otros métodos de clasificación de materiales existentes, como la utilización de espectrometría, no mejoran estos porcentajes. Uno de los posibles motivos que dificulta aumentar el porcentaje de clasificación correcta es que se disponía de un número muy reducido de experimentos, por lo que hemos verificado nuestro método mediante un estudio de simulación que se detallará en una sección posterior.

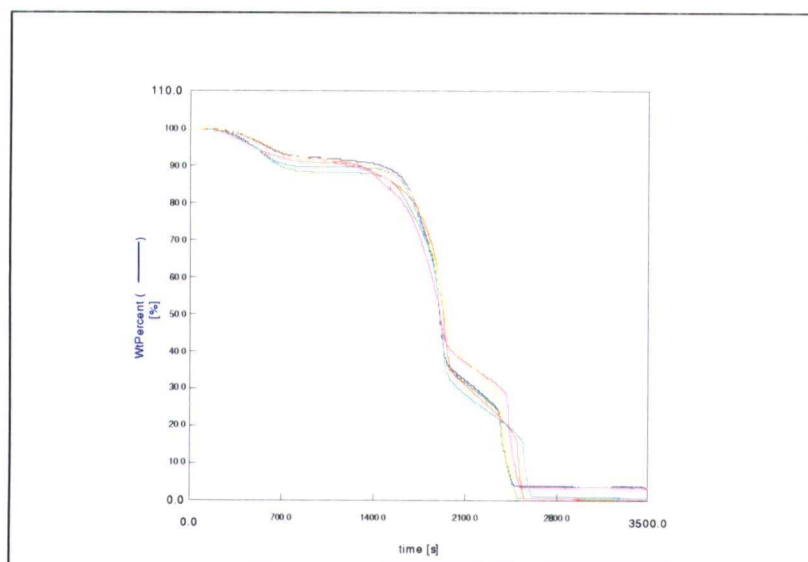


Figura 5.3: Comparación de 6 muestras de madera de dos clases distintas, una formada por las curvas en rosa y amarillo y otra por las restantes.

5.6 Aplicación a muestras de PVC

5.6.1 Características específicas de las muestras a analizar

Como una segunda aplicación del método propuesto de clasificación hemos elegido 16 muestras de PVC, diferenciadas entre rígidos y flexibles, elegidos a partes iguales y que serán las dos clases a considerar. Se tomaron muestras con un peso aproximado de 35 mg, y fueron sometidas a una rampa de calentamiento en aire desde 25°C hasta 600°C a una velocidad constante de 10°C por minuto. Posteriormente se sometieron a una etapa isoterma a 600°C para asegurar su combustión.

En la figura 5.4 se aprecia la diferencia entre una muestra de PVC rígido frente a otro flexible. Se puede observar las dos etapas fundamentales de la degradación del PVC. La primera, proceso primario, se debe a la inestabilidad de determinados átomos de cloro que conduce a la eliminación de la molécula de HCl y la formación de un doble enlace. En la segunda etapa es cuando el polímero residual arde.

Las muestras analizadas permanecieron estables por debajo de los 180°C, siendo, aproximadamente, a esta temperatura cuando se inicia la degradación. Como diferencias entre las muestras de PVC rígido y flexible que se pueden apreciar, mediante el estudio termogravimétrico, cabe resaltar:

- En las curvas TGA de las muestras rígidas se aprecian, con mayor claridad, dos escalones en la segunda etapa de degradación.

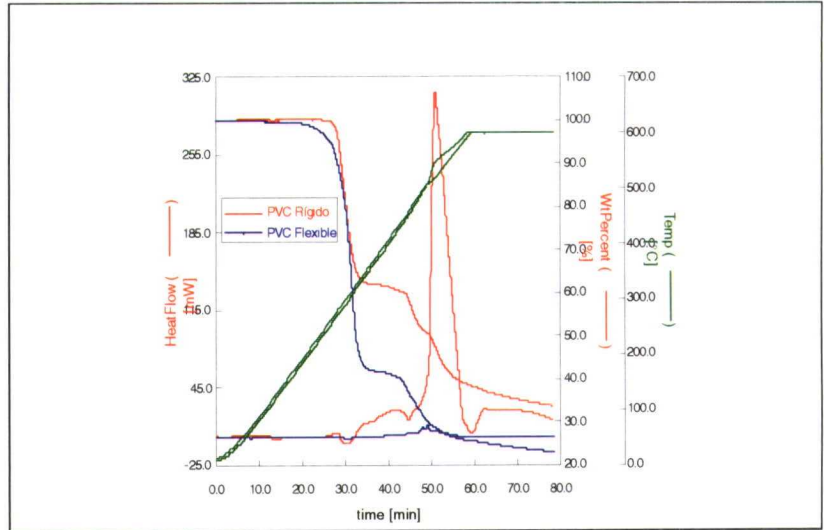


Figura 5.4: Curvas TGA y DSC para dos clases de PVC (en rojo rígido y en azul flexible).

- El calor generado en la combustión está entre 8 y 20 mW en las muestras flexibles y entre 225 y 330 mW en las rígidas.
- La pérdida de peso en la primera etapa varía entre el 40% y el 50% del PVC rígido, mientras en el flexible la variación está entre el 50% y el 60% (tiene mayor contenido de plastificante).

Teniendo en cuenta estas diferencias pensamos que es posible la correcta clasificación analizando el proceso hasta la pérdida del 60% del peso de las muestras (véase figura 5.6).

Resultados obtenidos

En este caso mostramos la función de validación cruzada para las 16 muestras analizadas (véase figura 5.5). Observando que la estimación de la probabilidad de clasificación correcta de una nueva observación, utilizando el método de clasificación no paramétrico con ventana h_{CV} sería del 99.4%.

5.7 Estudios de simulación

Con objeto de verificar el método presentado en la sección anterior hemos llevado a cabo un estudio de simulación. Considerando tres tipos de materiales de difícil clasificación (maderas de ciprés, eucalipto y roble) hemos simulado muestras de curvas TGA, mediante la modelización como mezcla de logísticas,

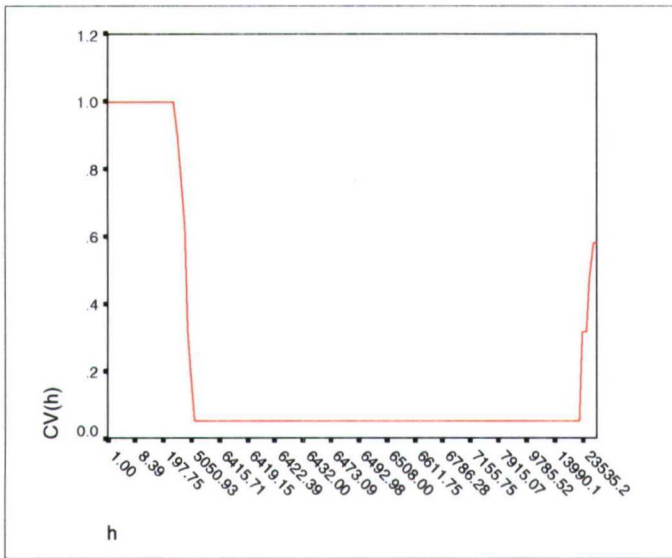


Figura 5.5: Función de validación cruzada para el PVC.

para lo cual fue necesario analizar muestras reales de estas maderas con el fin de estimar sus parámetros (véanse las tablas de 5.1 a 5.9).

5.7.1 Pasos para las simulaciones

Usaremos el modelo de mezcla de logísticas, presentado en el capítulo 3, siendo los pasos para las simulaciones:

1. Para cada tamaño, n , generar datos equiespaciados x_1, x_2, \dots, x_n , usando algún valor $\delta > 0$, que representa el tiempo entre instantes de muestreo del experimento termogravimétrico: $x_i = \delta \cdot i \quad i = 1, \dots, n$.
2. En cada punto x_1, x_2, \dots, x_n , calculamos la señal $m(x_i)$, para $i = 1, \dots, n$, usando el modelo de mezcla de logísticas con los parámetros obtenidos en las tablas 5.10 a 5.12.
3. Consideraremos dos situaciones distintas con respecto a la posible dependencia entre los parámetros (aleatorios) del modelo de mezcla de logísticas simulado, utilizando distintas varianzas y covarianzas entre parámetros.

5.7.2 Muestras simuladas

Simularemos distintas maderas para lo que ajustamos 3 clases (ciprés, eucalipto y roble) mediante un modelo de regresión paramétrico con mezclas de logísticas. Como tomamos muestras con el peso en porcentaje todos los pesos

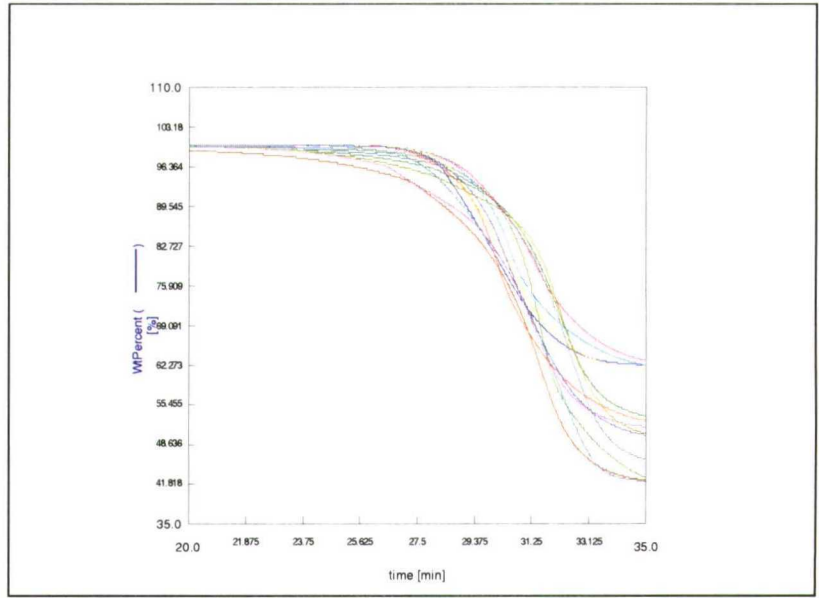


Figura 5.6: Curvas TGA correspondientes a 18 tipos de PVC.

sumarán 100. Supondremos que tenemos $G + 1$ grupos $i = 0, 1, \dots, G$, con $G = 2$ en este caso. Consideraremos como señal para el grupo r -ésimo las funciones $\varphi^{(r)}(x)$ siguientes:

$$\varphi^{(r)}(x) = \sum_{j=1}^{k_{(r)}} w_j^{(r)} f(a_j^{(r)} + b_j^{(r)} x),$$

donde los parámetros del modelo,

$$\left((w_1^{(r)}, a_1^{(r)}, b_1^{(r)}), (w_2^{(r)}, a_2^{(r)}, b_2^{(r)}), \dots, (w_k^{(r)}, a_k^{(r)}, b_k^{(r)}) \right),$$

se suponen variables aleatorias con distribución diferente para cada tipo de madera simulado.

Para el grupo r -ésimo se simulará:

$$(x_1, y_1, z_1, \dots, x_k, y_k, z_k) \stackrel{d}{=} N_{3k} \left(\mu_{(r)}, \Sigma_{(r)} \right),$$

con la condición de que si $\exists j/z_j < 0$ entonces se volverá a simular de nuevo. Luego se define:

$$a_j^{(r)} = x_j, b_j^{(r)} = y_j, w_j^{(r)} = \frac{z_j}{\sum_{l=1}^k z_l} 100, \quad j = 1, 2, \dots, k$$

Se han considerado dos situaciones al respecto de la posible dependencia entre los parámetros del modelo de mezcla de logísticas:

1.- Suponer independencia entre los parámetros del modelo.

$$\Sigma_{(r)} = \begin{pmatrix} \Sigma_{(r)(1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \Sigma_{(r)(2)} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \Sigma_{(r)(k)} \end{pmatrix},$$

con

$$\Sigma_{(r)(j)} = \begin{pmatrix} \sigma_{1,j}^2 & 0 & 0 \\ 0 & \sigma_{2,j}^2 & 0 \\ 0 & 0 & \sigma_{3,j}^2 \end{pmatrix}.$$

2.- Suponer dependencia entre los parámetros del modelo con coeficiente de correlación ρ_j .

$$\Sigma_{(r)} = \begin{pmatrix} \Sigma_{(r)(1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \Sigma_{(r)(2)} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \Sigma_{(r)(k)} \end{pmatrix},$$

con

$$\Sigma_{(r)(j)} = \begin{pmatrix} \sigma_{1,j}^2 & \rho_j \sigma_{1,j} \sigma_{2,j} & \rho_j \sigma_{1,j} \sigma_{3,j} \\ \rho_j \sigma_{1,j} \sigma_{2,j} & \sigma_{2,j}^2 & \rho_j \sigma_{2,j} \sigma_{3,j} \\ \rho_j \sigma_{1,j} \sigma_{3,j} & \rho_j \sigma_{2,j} \sigma_{3,j} & \sigma_{3,j}^2 \end{pmatrix}.$$

A continuación presentamos las tablas con los valores obtenidos del ajuste mediante el modelo de mezcla de logísticas para 3 muestras de madera de cada una de las tres clases consideradas, con objeto de estimar los parámetros de las distribuciones normales multivariantes $N_{3k}(\mu_{(r)}, \Sigma_{(r)})$.

i	1	2	3
w_i	8.5	32.75	47.60
a_i	5.12	105.01	13.45
b_i	-0.012	-0.056	-0.008

Tabla 5.2. Valores de los parámetros de la mezcla de logísticas de la curva TGA para ciprés-1.

i	1	2	3
w_i	10.5	34.65	48.40
a_i	4.7	98.02	15.85
b_i	-0.02	-0.048	-0.015

Tabla 5.3. Valores de los parámetros de la mezcla de logísticas de la curva TGA para ciprés-2.

i	1	2	3
w_i	13.5	37.35	39.57
a_i	5.5	101.5	23.42
b_i	-0.012	-0.058	-0.005

Tabla 5.4. Valores de los parámetros de la mezcla de logísticas de la curva TGA para ciprés-3.

i	1	2	3	4
w_i	13.27	18.81	32.75	33.40
a_i	5.31	14.02	105.01	14.47
b_i	-0.012	-0.008	-0.056	-0.006

Tabla 5.5. Valores de los parámetros de la mezcla de logísticas de la curva TGA para eucalipto-1.

i	1	2	3	4
w_i	12.2	17.5	30.5	34.50
a_i	4.35	13.05	101.5	16.5
b_i	-0.016	-0.007	-0.045	-0.01

Tabla 5.6. Valores de los parámetros de la mezcla de logísticas de la curva TGA para eucalipto-2.

i	1	2	3	4
w_i	14.5	16.5	34.5	35.60
a_i	7.33	13.05	102.5	13.5
b_i	-0.02	-0.01	-0.06	-0.008

Tabla 5.7. Valores de los parámetros de la mezcla de logísticas de la curva TGA para eucalipto-3.

i	1	2	3
w_i	5.12	55.7	38.75
a_i	7.35	78.5	10.01
b_i	-0.02	-0.07	-0.001

Tabla 5.8. Valores de los parámetros de la mezcla de logísticas de la curva TGA para roble-1.

i	1	2	3
w_i	6.6	57.5	33.6
a_i	5.5	87.5	11.5
b_i	-0.021	-0.07	-0.005

Tabla 5.9. Valores de los parámetros de la mezcla de logísticas de la curva TGA para roble-2.

i	1	2	3
w_i	5.8	56.5	30.5
a_i	9.5	67.05	12.5
b_i	-0.008	-0.06	-0.007

Tabla 5.10. Valores de los parámetros de la de mezcla de logísticas de la curva TGA para roble-3.

A partir de estos valores se pueden calcular las medias y varianzas para cada tipo de madera, así como los valores estimados para el coeficiente de correlación empleados para el segundo método.

	<i>Media</i>	<i>Desviación típica</i>
$w_1^{(1)}$	11.9067	2.7464
$a_1^{(1)}$	5.1067	0.4002
$b_1^{(1)}$	-0.0147	0.0046
$w_2^{(1)}$	38.402	2.5176
$a_2^{(1)}$	101.51	3.4954
$b_2^{(1)}$	-0.0540	0.0053
$w_3^{(1)}$	49.6907	5.2152
$a_3^{(1)}$	17.573	5.2036
$b_3^{(1)}$	-0.0093	0.0051

Tabla 5.11. Valores de las medias y desviaciones típicas para las simulaciones de muestras de ciprés.

	<i>Media</i>	<i>Desviación típica</i>
$w_1^{(2)}$	13.578	0.7318
$a_1^{(2)}$	5.6633	1.5211
$b_1^{(2)}$	-0.0160	0.0041
$w_2^{(2)}$	17.9827	1.4784
$a_2^{(2)}$	13.3733	0.5601
$b_2^{(2)}$	-0.0083	0.0015
$w_3^{(2)}$	33.2240	0.9643
$a_3^{(2)}$	103.0033	1.8083
$b_3^{(2)}$	-0.0537	0.0078
$w_4^{(2)}$	35.2153	1.2145
$a_4^{(2)}$	14.8233	1.5309
$b_4^{(2)}$	-0.0080	0.0020

Tabla 5.11. Valores de las medias y desviaciones típicas para las simulaciones de muestras de eucalipto.

	<i>Media</i>	<i>Desviación típica</i>
$w_1^{(3)}$	6.0497	0.8251
$a_1^{(3)}$	7.4507	2.0019
$b_1^{(3)}$	-0.0163	0.0072
$w_2^{(3)}$	58.5593	2.4852
$a_2^{(3)}$	77.6833	10.2494
$b_2^{(3)}$	-0.0667	0.0057
$w_3^{(3)}$	35.3913	3.1471
$a_3^{(3)}$	11.3367	1.2536
$b_3^{(3)}$	-0.0043	0.0031

Tabla 5.12. Valores de las medias y desviaciones típicas para las simulaciones de muestras de roble.

	$w_1^{(1)}$	$a_1^{(1)}$	$b_1^{(1)}$
$w_1^{(1)}$	1	0.653	0.215
$a_1^{(1)}$	0.653	1	0.880
$b_1^{(1)}$	0.215	0.880	1

Tabla 5.13. Matriz de correlaciones para las primeras componentes de ciprés.

	$w_2^{(1)}$	$a_2^{(1)}$	$b_2^{(1)}$
$w_2^{(1)}$	1	-0.038	-0.627
$a_2^{(1)}$	-0.038	1	-0.754
$b_2^{(1)}$	-0.627	-0.754	1

Tabla 5.14. Matriz de correlaciones para las segundas componentes de ciprés.

	$w_3^{(1)}$	$a_3^{(1)}$	$b_3^{(1)}$
$w_3^{(1)}$	1	-0.998	-0.600
$a_3^{(1)}$	-0.998	1	-0.554
$b_3^{(1)}$	-0.600	-0.554	1

Tabla 5.15. Matriz de correlaciones para las terceras componentes de ciprés.

	$w_1^{(2)}$	$a_1^{(2)}$	$b_1^{(2)}$
$w_1^{(2)}$	1	0.993	-0.569
$a_1^{(2)}$	0.993	1	-0.664
$b_1^{(2)}$	-0.569	-0.664	1

Tabla 5.16. Matriz de correlaciones para las primeras componentes de eucalipto.

	$w_2^{(2)}$	$a_2^{(2)}$	$b_2^{(2)}$
$w_2^{(2)}$	1	0.683	0.846
$a_2^{(2)}$	0.683	1	0.189
$b_2^{(2)}$	0.846	0.189	1

Tabla 5.17. Matriz de correlaciones para las segundas componentes de eucalipto.

	$w_3^{(2)}$	$a_3^{(2)}$	$b_3^{(2)}$
$w_3^{(2)}$	1	0.375	-0.987
$a_3^{(2)}$	0.375	1	-0.517
$b_3^{(2)}$	-0.987	-0.517	1

Tabla 5.18. Matriz de correlaciones para las terceras componentes de eucalipto.

	$w_4^{(2)}$	$a_4^{(2)}$	$b_4^{(2)}$
$w_4^{(2)}$	1	0.664	-1
$a_4^{(2)}$	0.664	1	-0.663
$b_4^{(2)}$	-1	-0.663	1

Tabla 5.19. Matriz de correlaciones para las cuartas componentes de eucalipto.

	$w_1^{(3)}$	$a_1^{(3)}$	$b_1^{(3)}$
$w_1^{(3)}$	1	-0.265	0.143
$a_1^{(3)}$	-0.265	1	0.917
$b_1^{(3)}$	0.143	0.917	1

Tabla 5.20. Matriz de correlaciones para las primeras componentes de roble.

	$w_2^{(3)}$	$a_2^{(3)}$	$b_2^{(3)}$
$w_2^{(3)}$	1	-0.470	0.810
$a_2^{(3)}$	-0.470	1	-0.898
$b_2^{(3)}$	0.810	-0.898	1

Tabla 5.21. Matriz de correlaciones para las segundas componentes de roble.

	$w_3^{(3)}$	$a_3^{(3)}$	$b_3^{(3)}$
$w_3^{(3)}$	1	-0.986	0.996
$a_3^{(3)}$	-0.986	1	-0.997
$b_3^{(3)}$	0.996	-0.997	1

Tabla 5.22. Matriz de correlaciones para las terceras componentes de roble.

A partir de los datos de las tablas 5.13 a 5.15 obtuvimos los valores para la estimación del coeficiente de correlación utilizando la mediana de las correlaciones, así para la muestra de ciprés, con correlaciones 0.653, 0.215 y 0.880, utilizamos $\rho = 0.653$.

5.7.3 Resultados de las simulaciones

Hemos simulado una muestra de 90 curvas TGA con igual probabilidad 1/3 de cada tipo de madera, encontrándose la ventana h_{CV} y el mínimo de la función de cross-validation (que se incluye en las dos últimas columnas de la tabla 5.23). Luego se han simulado 1000 nuevas curvas, que con las mismas probabilidades especificadas anteriormente serán de cada tipo, comprobándose a qué grupo las clasifica, cada una de ellas, la regla Bayes estimada no paramétricamente con ventana h_{CV} y confrontándolo con el grupo del que se ha simulado dicha curva. En base a todo ello se ha estimado la probabilidad de clasificación correcta, mediante el porcentaje de esas 1000 curvas que se han clasificado correctamente, tanto bajo dependencia ($PCCDep$) como bajo independencia ($PCCIndep$), que se presenta en la segunda y tercera columna de la tabla 5.23.

Varianza	$PCCIndep.$	$PCCDep.$	$\widetilde{CV}(h_{CV})Indep.$	$\widetilde{CV}(h_{CV})Dep.$
4σ	64.8	72.2	0.38	0.33
2σ	73.6	75.9	0.26	0.22
σ	77.8	82	0.15	0.14
$\sigma/2$	75.2	87	0.07	0.07
$\sigma/4$	89.5	88.5	0.05	0.03
$\sigma/8$	92.4	95.4	0.03	0.01
$\sigma/10$	92.4	97.8	0.01	0.01
$\sigma/20$	94.3	99.8	0.01	0
$\sigma/40$	98.9	100	0	0
$\sigma/50$	99.8	100	0	0

Tabla 5.23: Porcentajes de clasificación correcta (PCC) en función del método utilizado, con dependencia ($PCCDep$) o independencia ($PCCIndep$).

Correlación	PCC	$\overline{CV}(h_{CV})$
1.4ρ	78.6	0.12
ρ	80.7	0.15
$\rho/2$	79.9	0.16
$\rho/4$	78.2	0.16

Tabla 5.24: Influencia del coeficiente de correlación ρ en el método con dependencia.

Como era de prever, el porcentaje de clasificación correcta aumenta según disminuyen las varianzas del modelo, llegando a cotas del 92%–95% con factores de 1/8 de las varianzas originales. Para aumentos de hasta 4 veces la varianza las probabilidades de clasificación correcta bajan hasta valores de un 62%–72%. En general, los resultados son presencia de correlación son algo mejores que bajo independencia, aunque si la correlación es elevada el efecto puede ser contrario (ver tabla 5.24).

Hemos variado las proporciones de pertenencia a cada grupo observando que disminuye el porcentaje de clasificación correcta al tomar las muestras con proporciones más desiguales, como por ejemplo (1/4, 1/4, 1/2), lo que nos indica, que se obtiene una mejor clasificación cuando los grupos de materiales tienen la misma probabilidad de salir elegidos.

Capítulo 6

Conclusiones y líneas futuras

6.1 Conclusiones y objetivos alcanzados

6.1.1 Método paramétrico de mezcla de funciones logísticas

Un primer objetivo que hemos conseguido, y que se detalla en el capítulo 3 de esta memoria, es encontrar un método para el ajuste paramétrico de las curvas TGA. El método propuesto se basa en el ajuste de la curva mediante el empleo de diferentes funciones logísticas (*modelo con mezcla de logísticas*), basándonos en la idea de que cada una representa la cinética de descomposición de los distintos materiales simples de los que está compuesta la muestra inicial. Hemos propuesto la expresión (6.1) para modelizar la evolución de la masa de la muestra $Y(t)$, con el tiempo t :

$$Y(t) = \sum_{i=1}^k w_i f(a_i + b_i t), \quad (6.1)$$
$$f(t) = \frac{\exp t}{1 + \exp t},$$

donde w_i son parámetros que representan las cantidades globales de pérdidas de peso en cada proceso, a_i son parámetros de localización y b_i representa la velocidad de pérdida de peso, para los k escalones que presente la función $Y(t)$.

Las ventajas del método propuesto frente a otros existentes son:

1. Nuestro método cubre toda la gama de conversiones. El modelo (6.1) puede ser usado para el estudio de cualquier tipo de material, tanto si está regido por una sola ley cinética como por varias.

2. Permite representar, tanto la velocidad de pérdida de masa como el ajuste de ésta con el tiempo, mediante una ecuación única para cada proceso.

3. Posibilita el empleo de medios estadísticos que permiten, mediante contrastes de significación, valorar la bondad del ajuste del modelo propuesto.

4. Permite, para su comparación con materiales ya estudiados, la aplicación de modelos cinéticos clásicos del tipo Arrhenius a cada una de las funciones de degradación individuales.

5. Reproduce perfectamente el comportamiento asintótico al final de cada proceso degradativo.

6. Puesto que se obtiene una función paramétrica de ajuste, se pueden calcular las demás derivadas de esta función, sin que presenten el ruido típico de la derivación numérica usada en otros métodos.

Por tanto, este primer método que proponemos de ajuste de las curvas TGA, supone un modelo alternativo a los existentes para la explicación de la cinética.

Finalmente, comentar que esta metodología ha sido presentada en distintos foros científicos como el Congreso Gallego de Estadística e Investigación Operativa o el Congreso TERMEC. Además, una primera versión del método, concretamente el caso de aplicación a procesos de descomposición simples (con un solo escalón) ha sido admitido para su publicación por la revista *Thermochimica Acta* (véase Naya, Cao y Artiaga (2003a)) resaltando el interés como aplicación al Análisis Térmico por uno de los referees de esta revista. La versión completa del modelo de mezcla de logísticas ha sido enviada a la revista *Journal of Chemometrics* (véase Naya, Cao, Artiaga, Barbadillo y López de Ullibarri (2003)) y será presentada en el ICCE10 (véase Artiaga et al. (2003)).

6.1.2 Suavizado automático de curvas de Análisis Térmico

Con los métodos existentes actualmente en el Software comercial, el obtener un buen suavizado se convierte en una cuestión de arte que depende, en gran medida, de la pericia del usuario. Al final no existe un método matemático que indique el grado de bondad del suavizado siendo la “estética” de la curva la que nos puede indicar si hay un infrasuavizado. El problema de sobresuavización es más difícil de identificar visualmente. El método que hemos propuesto resuelve este problema al seleccionar, mediante criterios estadísticos, el grado de suavizado óptimo en cada momento.

Otro de los objetivos alcanzados en esta tesis fue el de diseñar un método para el suavizado de curvas, que permite, mediante el empleo de una ventana plug-in en dos etapas, la elección automática del grado de suavizado, implementando el correspondiente programa informático para su aplicación a datos reales. Este método lo hemos verificado con distintos experimentos y concluimos que mejora sustancialmente otros empleados actualmente en distintos programas de software.

El objetivo era obtener unos resultados de suavización que superasen a los proporcionados por el software existente, que, como comentamos al principio de esta memoria, presenta problemas importantes de suavización.

A fin de constatar las mejoras, que presenta nuestra rutina de suavizado, realizamos, en el capítulo 4 de esta memoria, una comparativa de los resultados que obtenemos una vez aplicado nuestro algoritmo de suavización, en contraste

con los obtenidos previamente. Para ello utilizamos la muestra de oxalato de calcio en la que pudimos observar los graves problemas de suavizado en la estimación de la primera derivada, que en el caso de la segunda derivada todavía se incrementaron más, no siendo posible apreciar ninguna tendencia.

El problema que presentan los programas de suavizado existentes, deriva del hecho de realizar una sobresuavización de la curva (parámetro de suavizado demasiado grande). La gravedad de este hecho, quedaba resaltado en las zonas marcadas en la figura 4.9, ya que podemos observar como la profundidad de los picos se ha visto claramente reducida en comparación con la suavización que proponemos obtenida mediante la adecuada elección de la ventana de suavizado.

En base a todo ello, presentamos nuestros resultados de suavización, incluyendo la utilización de ventana local o ventana global.

Podemos constatar como el problema de sobresuavización de los picos de la figura 4.9, ha sido resuelto y afirmar a su vez que la curva no presenta las graves deficiencias de suavización obtenidas por los programas existentes. Además hemos aplicado nuestro método a otras muestras observando su buen ajuste.

Finalmente, este hecho lo hemos contrastado comparando los picos de la curva DTG con los picos de la curva DSC correspondiente, comprobando que en nuestro caso sí tienen la misma relación de amplitudes. Además, como comprobación adicional, hemos comparado la estimación proporcionada por nuestro método con el ajuste paramétrico logístico visto en el capítulo 3 de esta memoria, para muestras susceptibles de ser ajustadas correctamente por este modelo paramétrico, observando la igualdad de las gráficas resultantes.

Este método para el suavizado automático de curvas, para el caso de la estimación de la primera derivada, lo hemos presentado en el Congreso Nacional de Ingeniería Mecánica (Naya et al. 2002). Un estudio más completo, con los resultados de las simulaciones realizadas, lo hemos enviado a la revista *Journal of Statistical Computation and Simulation* (véase Artiaga, Cao, Naya y Trillo (2003)).

6.1.3 Método de clasificación de materiales

En el capítulo 5 de esta memoria hemos presentado un nuevo método de clasificación de materiales mediante el empleo de regresión funcional.

El método propuesto de clasificación está basado en minimizar la probabilidad de clasificación incorrecta utilizando la regla Bayes.

Hemos aplicado esta técnica a la discriminación de maderas y de otros polímeros, debido a que estos materiales presentan una dificultad, constatada por distintos autores, de clasificación.

Con el fin de verificar el método de reconocimiento de materiales hemos realizado simulaciones imitando el comportamiento de distintos experimentos concluyendo que el porcentaje de clasificación correcta depende de la variabilidad de las muestras de cada tipo de material.

Al igual que el caso del suavizado de curvas, hemos implementado este método de clasificación en un programa en lenguaje C++ (Dominguez (2003)). Además, una aplicación al reconocimiento de maderas ha sido presentado en la

IX Conferencia Española de Biometría (véase Naya, Cao, Artiaga, Dominguez y Losada (2003)) y una aplicación al reconocimiento de PVCs la presentaremos en el Congreso Internacional de Ingeniería de Materiales Compuestos (*ICCE10*) que se celebrará en Nueva Orleans en el próximo mes de julio (Naya, Cao y Artiaga (2003b)).

6.2 Líneas futuras

Todo trabajo de investigación abre el camino a nuevos problemas, así, hemos encontrado una serie de posibles extensiones de algunos de los resultados expuestos en esta memoria, y que podrían constituir futuras líneas de investigación que se enumeran a continuación:

- Extensión del modelo de mezcla de logísticas.
- Suavización de espectros tridimensionales.
- Ampliación del método de clasificación de materiales.

6.2.1 Extensión del modelo de mezcla de logísticas

Como primera línea de futura investigación proponemos extender el modelo de mezcla de logísticas, presentado en el capítulo 3 de esta memoria, mediante la posibilidad de contemplar composiciones de la función logística con polinomios de grado mayor a 1. Suponiendo que la curva TGA tome valores entre un mínimo (A) y un máximo (B) podemos de reescalar la expresión del modelo de regresión logística sin más que considerar la nueva ecuación siguiente:

$$y = \frac{B + A \exp(P_k(x))}{1 + \exp(P_k(x))}, \quad (6.2)$$

donde $P_k(x)$ es un polinomio de grado k . Los parámetros del modelo (6.2) pueden estimarse mediante un ajuste paramétrico del modo siguiente:

$$P_k(x) = \log \left(\frac{B - y}{y - A} \right).$$

El polinomio buscado P_k será el que ajuste la nube de puntos en los que la variable independiente sigue siendo la X , en nuestro caso el tiempo de la reacción o la temperatura, y donde la variable respuesta es ahora el logaritmo del cociente $\frac{B-y}{y-A}$.

Finalmente, proponemos elaborar un contraste, de bondad de ajuste, de tipo no paramétrico para este tipo de modelos basándonos en los propuestos por Vilar-Fernández y González-Manteiga (1995) y Vilar-Fernández y González-Manteiga (2000).

6.2.2 Suavización de espectros tridimensionales

Una segunda línea de futura investigación que proponemos es el estudio del suavizado, mediante métodos no paramétricos, de espectros FTIR tridimensionales obtenidos a partir de los gases de salida del TGA. Estos espectros son funciones de \mathbb{R}^2 en \mathbb{R} en las que se representa la absorvancia respecto a la longitud de onda y al tiempo. Para conseguir un suavizado correcto es necesario hacer uso de técnicas de regresión bidimensional.

Estos modelos de regresión parten de muestras del tipo $\{(\mathbf{X}_i, Y_i)\}_{i=1}^n$ que corresponden con puntos de \mathbb{R}^{d+1} , donde Y_i son las observaciones de la variable respuesta de tipo escalar y \mathbf{X}_i son \mathbb{R}^d -variables independientes que tienen en común una densidad f con soporte en un cierto espacio $\Omega \subseteq \mathbb{R}^d$. La función de regresión, a estimar de modo no paramétrico, se puede expresar en este caso mediante:

$$m(x) = E(Y | \mathbf{X} = \mathbf{x})$$

para un vector dado $\mathbf{x} \in \Omega$ sin la imposición de que m se ajuste a una familia paramétrica de funciones. Asumiremos el modelo:

$$Y_i = m(\mathbf{X}_i) + \varepsilon_i, \quad i = 1, 2, \dots, n,$$

donde

$$E(\varepsilon_i | \mathbf{X}_i) = 0,$$

$$Var(\varepsilon_i | \mathbf{X}_i) = \sigma^2,$$

siendo posible la consideración de dependencia entre los errores.

El estimador lineal local de la regresión m será el $\hat{\alpha}$, que es solución para α del siguiente problema de minimización:

$$\min \sum_{i=1}^n \{Y_i - \alpha - \beta^t(\mathbf{X}_i - \mathbf{x})\}^2 K_{\mathbf{H}}(\mathbf{X}_i - \mathbf{x}), \quad (6.3)$$

donde \mathbf{H} es una matriz $d \times d$ simétrica positiva dependiente de n ; K es el núcleo en dimensión d con $\int K(\mathbf{u})d\mathbf{u} = 1$; y $K_{\mathbf{H}}(\mathbf{u}) = |\mathbf{H}|^{-1} K(\mathbf{H}^{-1}\mathbf{u})$. Igual que en el contexto univariante tendremos que calcular el grado de suavización que en este caso multivariante se trata de una matriz \mathbf{H} , y que jugará un papel similar de control del número de valores escogidos en cada entorno del punto a estimar, usados en la estimación de $m(\mathbf{x})$. Al igual que en el caso univariante, la selección de esta matriz va a ser crucial para la correcta estimación de la función de regresión. Asumiendo que $\mathbf{X}_x^t \mathbf{W}_x \mathbf{X}_x$ es no singular, la solución del problema (6.3) es:

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = (\mathbf{X}_x^t \mathbf{W}_x \mathbf{X}_x)^{-1} \mathbf{X}_x^t \mathbf{W}_x \mathbf{Y} \quad \text{con } \mathbf{X}_x = \begin{pmatrix} 1 & (\mathbf{X}_1 - \mathbf{x})^t \\ \vdots & \vdots \\ 1 & (\mathbf{X}_n - \mathbf{x})^t \end{pmatrix},$$

y $\mathbf{W}_x = \text{diag}\{K_{\mathbf{H}}(\mathbf{X}_1 - \mathbf{x}), \dots, K_{\mathbf{H}}(\mathbf{X}_n - \mathbf{x})\}$. El estimador lineal local para $m(\mathbf{x})$ es:

$$\hat{m}(\mathbf{x}; \mathbf{H}) = e_1^t (\mathbf{X}_x^t \mathbf{W}_x \mathbf{X}_x)^{-1} \mathbf{X}_x^t \mathbf{W}_x \mathbf{Y}, \quad (6.4)$$

donde e_1 es el $(d+1) \times 1$ vector que toma el valor 1 en su primera coordenada y todas las demás el valor 0.

En este caso multivariante es también posible encontrar métodos para el cálculo óptimo de la ventana mediante criterios similares a los vistos en el capítulo 3 de esta memoria para el caso unidimensional.

Como un primer caso de estudio hemos ajustado una parte del espectro correspondiente a una muestra de madera de castaño, utilizando el método de validación cruzada generalizado para la obtención del selector de la ventana (véase Liu (2001)). Obteniendo como resultando para la ventana la matriz siguiente:

$$\mathbf{H}_{CV}^t = \begin{pmatrix} 0.035645 & 0.001791 \\ 0.001791 & 0.964132 \end{pmatrix},$$

y como gráfico suavizado del espectro el de la figura 6.1, con su correspondiente gráfico de curvas de nivel dado en la figura 6.2.

6.2.3 Ampliación del método de clasificación de materiales

Una última propuesta de trabajo consiste en la extensión del método de clasificación a otro tipo de curvas, como el caso de las curvas DSC o los espectros obtenidos por técnicas FTIR, puesto que únicamente se necesita disponer de diferentes muestras ya clasificadas. Incluso se podría extender al caso de espectros tridimensionales mediante la adecuada elección de la seminorma (que, obviamente, mediría volúmenes entre las distintas superficies).

Otro aspecto, que nos parece interesante, es encontrar un método que permita la clasificación de maderas independientemente de su grado de humedad. Hemos comprobado que maderas del mismo tipo pero con distinto grado de humedad presentan curvas TGA diferentes, que difieren en un polinomio de grado 2, por lo que un estudio de sus derivadas puede ser de gran utilidad para su clasificación (véase la figura 6.3).

Estas derivadas de las curvas están desplazadas horizontalmente, lo que sugiere que utilizando un cambio de escala se podrían conseguir distancias

pequeñas entre materiales del mismo tipo. Como una línea de futura investigación sugerimos experimentar con distintos casos de maderas, con diferente grado de humedad, mediante la “distancia” siguiente:

$$\|f - g\|_R = \inf_{a,b} \int \left| f''(x) - g''\left(\frac{x-a}{b}\right) \right| dx.$$

Podemos apreciar como las gráficas TGA de la figura 6.3 son distintas entre sí, aunque ofrecen una forma similar. Las dos pertenecen al mismo tipo de madera, pero con una humedad distinta. En las figuras 6.4 y 6.5 podemos observar las primeras y segundas derivadas, respectivamente, de las curvas TGA.

6.2.4 Salidas gráficas

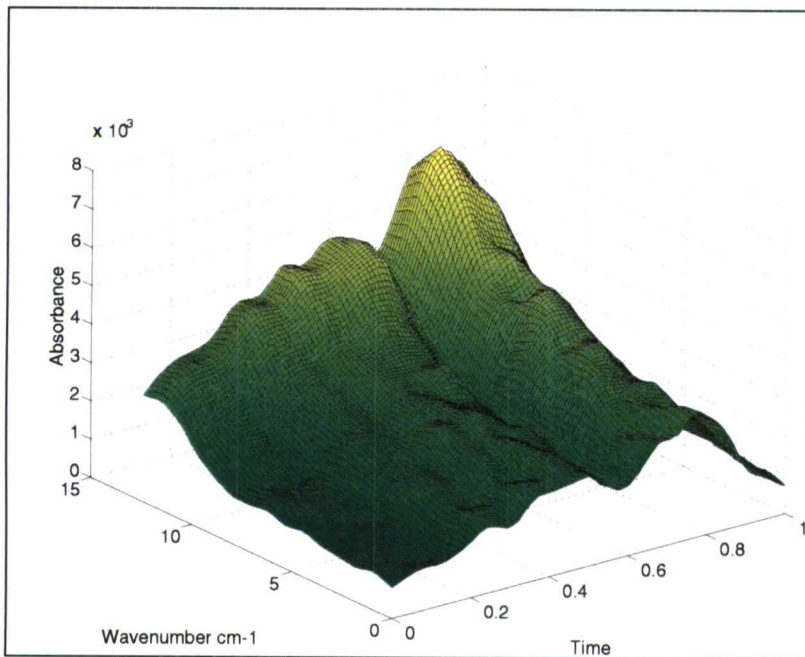


Figura 6.1: Espectro de una muestra de madera de castaño, en donde los valores de los ejes se han reescalado.

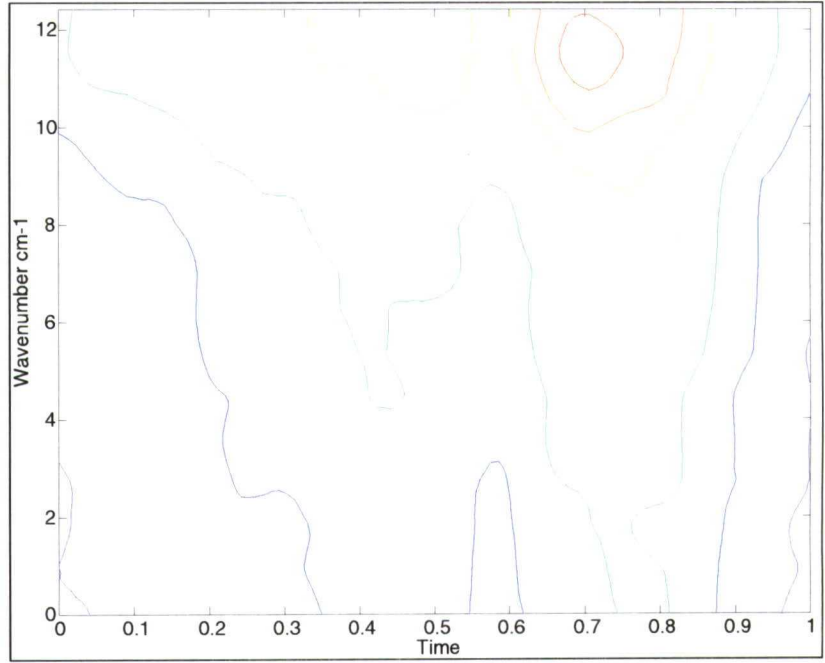


Figura 6.2: Curvas de nivel correspondientes al suavizado del espectro de FTIR.

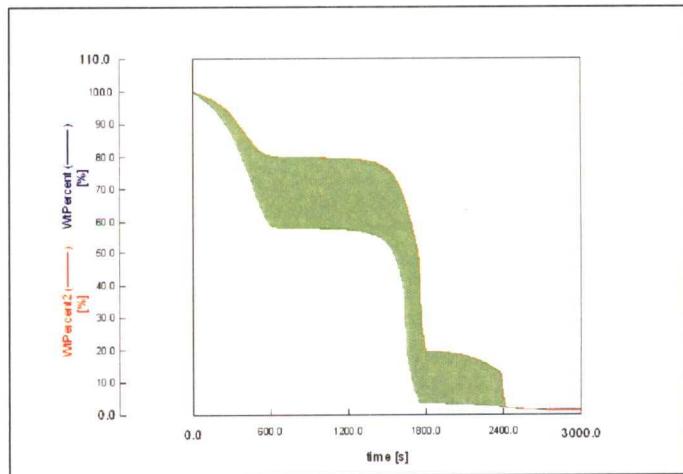


Figura 6.3: Comparación de dos muestras de madera de avellano con distinta humedad (se resalta el área entre sus curvas TGA).

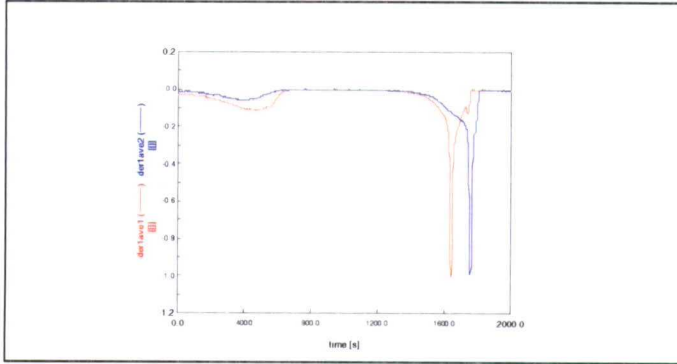


Figura 6.4: Comparación de las primeras derivadas para el avellano.

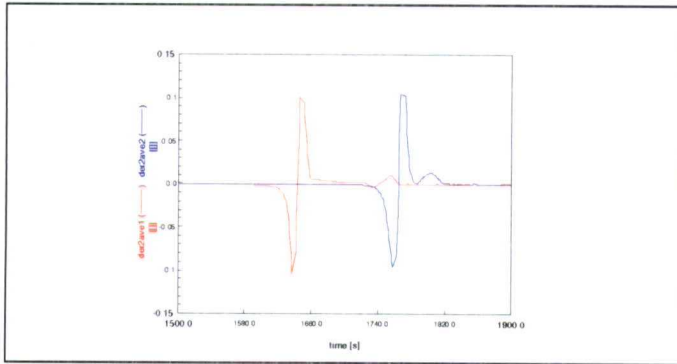


Figura 6.5: Comparación de las segundas derivadas para el avellano.

Anexos

Cálculo de las constantes $C_{\nu,p}$

A continuación, en este primer anexo, incluimos los cálculos realizados para la obtención de las constantes $C_{\nu,p}$ necesarias para estimar el valor óptimo del parámetro de suavizado, a partir de las expresiones dadas en el capítulo 4.

En lo que sigue ν será el orden de la derivada que desamos estimar, y p el orden del polinomio empleado en la estimación.

Si $\nu + p$ es impar, utilizaremos la expresión:

$$C_{\nu,p}(K) = \left[\frac{(p+1)!^2 (2\nu+1) \int K_{\nu}^{*2}(t) dt}{2(p+1-\nu) \left\{ \int t^{p+1} K_{\nu}^*(t) dt \right\}^2} \right]^{\frac{1}{2p+3}}.$$

Si $\nu + p$ es par, dado que $\int t^{p+1} K_{\nu}^*(t) dt$ es igual a cero, emplearemos la fórmula:

$$C_{\nu,p}(K) = \left[\frac{(p+2)!^2 (2\nu+1) \int K_{\nu}^{*2}(t) dt}{2(p+2-\nu) \left\{ \int t^{p+2} K_{\nu}^*(t) dt \right\}^2} \right]^{\frac{1}{2p+5}}.$$

Para el cálculo de estos valores, una vez hallados los valores $\mu_j = \int u^j K(u) du$, son necesarias las siguientes matrices y funciones auxilires:

$$S = (\mu_{j+\ell})_{j=0,\dots,p}^{\ell=0,\dots,p}, \quad S^{-1} = (S^{\nu\ell})_{\nu=0,\dots,p}^{\ell=0,\dots,p}, \quad K_{\nu}^*(t) = \left(\sum_{\ell=0}^p S^{\nu\ell} t^{\ell} \right) K(t),$$

donde K es la función núcleo. En este apartado hemos utilizado el núcleo de Epanechnikov, cuya expresión es $K(t) = \frac{3}{4} (1-t^2) I_{(|t|\leq 1)}$, también hemos repetido los cálculos para el núcleo cuártico, recogiénose al final de esta sección.

- Cálculo de las constantes $C_{\nu,p}(K)$ para $p = 6$

$$\begin{aligned} \mu_0 &= \int_{-1}^1 u^0 K(u) du = 1 \\ \mu_1 &= \int_{-1}^1 u^1 K(u) du = 0 \\ \mu_2 &= \int_{-1}^1 u^2 K(u) du = \frac{1}{5} \\ \mu_3 &= \int_{-1}^1 u^3 K(u) du = 0 \\ \mu_4 &= \int_{-1}^1 u^4 K(u) du = \frac{3}{35} \\ \mu_5 &= \int_{-1}^1 u^5 K(u) du = 0 \end{aligned}$$

$$\begin{aligned}\mu_6 &= \int_{-1}^1 u^6 K(u) du = \frac{1}{21} \\ \mu_7 &= \int_{-1}^1 u^7 K(u) du = 0 \\ \mu_8 &= \int_{-1}^1 u^8 K(u) du = \frac{1}{33} \\ \mu_9 &= \int_{-1}^1 u^9 K(u) du = 0 \\ \mu_{10} &= \int_{-1}^1 u^{10} K(u) du = \frac{3}{143} \\ \mu_{11} &= \int_{-1}^1 u^{11} K(u) du = 0 \\ \mu_{12} &= \int_{-1}^1 u^{12} K(u) du = \frac{1}{65}\end{aligned}$$

A partir de estos valores obtenemos la matriz S :

$$S = \begin{pmatrix} 1 & 0 & \frac{1}{5} & 0 & \frac{3}{35} & 0 & \frac{1}{21} \\ 0 & \frac{1}{5} & 0 & \frac{3}{35} & 0 & \frac{1}{21} & 0 \\ \frac{1}{5} & 0 & \frac{3}{35} & 0 & \frac{1}{21} & 0 & \frac{1}{33} \\ 0 & \frac{3}{35} & 0 & \frac{1}{21} & 0 & \frac{1}{33} & 0 \\ \frac{3}{35} & 0 & \frac{1}{21} & 0 & \frac{1}{33} & 0 & \frac{3}{143} \\ 0 & \frac{1}{21} & 0 & \frac{1}{33} & 0 & \frac{3}{143} & 0 \\ \frac{1}{21} & 0 & \frac{1}{33} & 0 & \frac{3}{143} & 0 & \frac{1}{65} \end{pmatrix}$$

Obtenemos la matriz $(S^{v,l})$ invirtiendo la matriz S :

$$S^{-1} = \begin{pmatrix} \frac{3675}{1024} & 0 & -\frac{40425}{1024} & 0 & \frac{105105}{1024} & 0 & -\frac{75075}{1024} \\ 0 & \frac{3675}{64} & 0 & -\frac{8085}{32} & 0 & \frac{15015}{64} & 0 \\ -\frac{40425}{1024} & 0 & \frac{832755}{1024} & 0 & -\frac{2597595}{1024} & 0 & \frac{2027025}{1024} \\ 0 & -\frac{8085}{32} & 0 & \frac{21945}{16} & 0 & -\frac{45045}{32} & 0 \\ \frac{105105}{1024} & 0 & -\frac{2597595}{1024} & 0 & \frac{8963955}{1024} & 0 & -\frac{7432425}{1024} \\ 0 & \frac{15015}{64} & 0 & -\frac{45045}{32} & 0 & \frac{99099}{64} & 0 \\ -\frac{75075}{1024} & 0 & \frac{2027025}{1024} & 0 & -\frac{7432425}{1024} & 0 & \frac{6441435}{1024} \end{pmatrix}$$

Partiendo de los términos $S^{v,l}$ y de la expresión:

$$K_v^*(t) = \left(\sum_{\ell=0}^p S^{v\ell} t^\ell \right) K(t)$$

obtenemos las funciones $K_v^*(t)$ para $v = 0, 1, \dots, 6$:

$$K_0^*(t) = \left(\frac{3675}{1024} - \frac{40425}{1024}t^2 + \frac{105105}{1024}t^4 - \frac{75075}{1024}t^6 \right) K(t)$$

$$K_1^*(t) = \left(\frac{3675}{64}t - \frac{8085}{32}t^3 + \frac{15015}{64}t^5 \right) K(t)$$

$$K_2^*(t) = \left(-\frac{40425}{1024} + \frac{832755}{1024}t^2 - \frac{2597595}{1024}t^4 + \frac{2027025}{1024}t^6 \right) K(t)$$

$$K_3^*(t) = \left(-\frac{8085}{32}t + \frac{21945}{16}t^3 - \frac{45045}{32}t^5 \right) K(t)$$

$$K_4^*(t) = \left(\frac{105105}{1024} - \frac{2597595}{1024}t^2 + \frac{8963955}{1024}t^4 - \frac{7432425}{1024}t^6 \right) K(t)$$

$$K_5^*(t) = \left(\frac{15015}{64}t - \frac{45045}{32}t^3 + \frac{99099}{64}t^5 \right) K(t)$$

$$K_6^*(t) = \left(-\frac{75075}{1024} + \frac{2027025}{1024}t^2 - \frac{7432425}{1024}t^4 + \frac{6441435}{1024}t^6 \right) K(t)$$

Una vez obtenidas las funciones $K_v^*(t)$ para $v = 0, 1, \dots, 6$ podemos calcular los valores de $\int K_v^{*2}(t) dt$, de $\left(\int t^{p+1} K_v^*(t) dt \right)^2$ y de $\left(\int t^{p+2} K_v^*(t) dt \right)^2$ siendo, en este caso, $p = 6$ y $v = 0, 1, \dots, 6$:

$$\int_{-1}^1 K_0^*(t)^2 dt = \int_{-1}^1 \left(\frac{3675}{1024} - \frac{40425}{1024}t^2 + \frac{105105}{1024}t^4 - \frac{75075}{1024}t^6 \right)^2 K(t)^2 dt = \frac{11025}{4352}$$

$$\begin{aligned} \int_{-1}^1 K_1^*(t)^2 dt &= \int_{-1}^1 \left(\frac{3675}{64}t - \frac{8085}{32}t^3 + \frac{15015}{64}t^5 \right)^2 K(t)^2 dt \\ &= \frac{2205}{64} \end{aligned}$$

$$\begin{aligned} \int_{-1}^1 K_2^*(t)^2 dt &= \int_{-1}^1 \left(-\frac{40425}{1024} + \frac{832755}{1024}t^2 - \frac{2597595}{1024}t^4 + \frac{2027025}{1024}t^6 \right)^2 K(t)^2 dt \\ &= \frac{1964655}{4352} \end{aligned}$$

$$\begin{aligned} \int_{-1}^1 K_3^*(t)^2 dt &= \int_{-1}^1 \left(-\frac{8085}{32}t + \frac{21945}{16}t^3 - \frac{45045}{32}t^5 \right)^2 K(t)^2 dt \\ &= \frac{10395}{16} \end{aligned}$$

$$\begin{aligned} \int_{-1}^1 K_4^*(t)^2 dt &= \int_{-1}^1 \left(\frac{105105}{1024} - \frac{2597595}{1024}t^2 + \frac{8963955}{1024}t^4 - \frac{7432425}{1024}t^6 \right)^2 K(t)^2 dt \\ &= \frac{17342325}{4352} \end{aligned}$$

$$\begin{aligned} \int_{-1}^1 K_5^*(t)^2 dt &= \int_{-1}^1 \left(\frac{15015}{64}t - \frac{45045}{32}t^3 + \frac{99099}{64}t^5 \right)^2 K(t)^2 dt \\ &= \frac{189189}{320} \end{aligned}$$

$$\begin{aligned} \int_{-1}^1 K_6^*(t)^2 dt &= \int_{-1}^1 \left(-\frac{75075}{1024} + \frac{2027025}{1024}t^2 - \frac{7432425}{1024}t^4 \right. \\ &\quad \left. + \frac{6441435}{1024}t^6 \right)^2 K(t)^2 dt = \frac{10405395}{4352} \end{aligned}$$

$$\begin{aligned} \left\{ \int_{-1}^1 t^{p+1} K_0^*(t) dt \right\}^2 &= \left(\int_{-1}^1 t^7 \left(\frac{3675}{1024} - \frac{40425}{1024}t^2 + \frac{105105}{1024}t^4 \right. \right. \\ &\quad \left. \left. - \frac{75075}{1024}t^6 \right) K(t) dt \right)^2 = 0 \end{aligned}$$

$$\begin{aligned} \left\{ \int_{-1}^1 t^{p+1} K_1^*(t) dt \right\}^2 &= \left(\int_{-1}^1 t^7 \left(\frac{3675}{64}t - \frac{8085}{32}t^3 \right. \right. \\ &\quad \left. \left. + \frac{15015}{64}t^5 \right) K(t) dt \right)^2 = \frac{49}{20449} \end{aligned}$$

$$\begin{aligned} \left\{ \int_{-1}^1 t^{p+1} K_2^*(t) dt \right\}^2 &= \left(\int_{-1}^1 t^7 \left(-\frac{40425}{1024} + \frac{832755}{1024}t^2 - \frac{2597595}{1024}t^4 \right. \right. \\ &\quad \left. \left. + \frac{2027025}{1024}t^6 \right) K(t) dt \right)^2 = 0 \end{aligned}$$

$$\begin{aligned} \left\{ \int_{-1}^1 t^{p+1} K_3^*(t) dt \right\}^2 &= \left(\int_{-1}^1 t^7 \left(-\frac{8085}{32}t + \frac{21945}{16}t^3 \right. \right. \\ &\quad \left. \left. - \frac{45045}{32}t^5 \right) K(t) dt \right)^2 = \frac{49}{169} \end{aligned}$$

$$\begin{aligned} \left\{ \int_{-1}^1 t^{p+1} K_4^*(t) dt \right\}^2 &= \left(\int_{-1}^1 t^7 \left(\frac{105105}{1024} - \frac{2597595}{1024}t^2 + \frac{8963955}{1024}t^4 \right. \right. \\ &\quad \left. \left. - \frac{7432425}{1024}t^6 \right) K(t) dt \right)^2 = 0 \end{aligned}$$

$$\begin{aligned} \left\{ \int_{-1}^1 t^{p+1} K_5^*(t) dt \right\}^2 &= \left(\int_{-1}^1 t^7 \left(\frac{15015}{64}t - \frac{45045}{32}t^3 \right. \right. \\ &\quad \left. \left. + \frac{99099}{64}t^5 \right) K(t) dt \right)^2 = \frac{49}{25} \end{aligned}$$

$$\left\{ \int t^{p+1} K_6^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^7 \left(-\frac{75\,075}{1024} + \frac{2027\,025}{1024} t^2 - \frac{7432\,425}{1024} t^4 + \frac{6441\,435}{1024} t^6 \right) K(t) dt \right)^2 = 0$$

$$\left\{ \int t^{p+2} K_0^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^8 \left(\frac{3675}{1024} - \frac{40\,425}{1024} t^2 + \frac{105\,105}{1024} t^4 - \frac{75\,075}{1024} t^6 \right) K(t) dt \right)^2 = \frac{49}{5909\,761}$$

$$\left\{ \int t^{p+2} K_1^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^8 \left(\frac{3675}{64} t - \frac{8085}{32} t^3 + \frac{15\,015}{64} t^5 \right) K(t) dt \right)^2 = 0$$

$$\left\{ \int t^{p+2} K_2^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^8 \left(-\frac{40\,425}{1024} + \frac{832\,755}{1024} t^2 - \frac{2597\,595}{1024} t^4 + \frac{2027\,025}{1024} t^6 \right) K(t) dt \right)^2 = \frac{784}{48\,841}$$

$$\left\{ \int t^{p+2} K_3^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^8 \left(-\frac{8085}{32} t + \frac{21\,945}{16} t^3 - \frac{45\,045}{32} t^5 \right) K(t) dt \right)^2 = 0$$

$$\left\{ \int t^{p+2} K_4^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^8 \left(\frac{105\,105}{1024} - \frac{2597\,595}{1024} t^2 + \frac{8963\,955}{1024} t^4 - \frac{7432\,425}{1024} t^6 \right) K(t) dt \right)^2 = \frac{196}{289}$$

$$\left\{ \int t^{p+2} K_5^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^8 \left(\frac{15\,015}{64} t - \frac{45\,045}{32} t^3 + \frac{99\,099}{64} t^5 \right) K(t) dt \right)^2 = 0$$

$$\left\{ \int t^{p+2} K_6^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^8 \left(-\frac{75\,075}{1024} + \frac{2027\,025}{1024} t^2 - \frac{7432\,425}{1024} t^4 + \frac{6441\,435}{1024} t^6 \right) K(t) dt \right)^2 = \frac{784}{289}$$

Finalmente, sustituiremos los valores obtenidos anteriormente en la expresión:

$$C_{v,p}(K) = \left[\frac{(p+1)!^2 (2v+1) \int K_v^{*2}(t) dt}{2(p+1-v) \left\{ \int t^{p+1} K_v^*(t) dt \right\}^2} \right]^{\frac{1}{2p+3}},$$

o bien en la expresión:

$$C_{v,p}(K) = \left[\frac{(p+2)!^2 (2v+1) \int K_v^{*2}(t) dt}{2(p+2-v) \left\{ \int t^{p+2} K_v^*(t) dt \right\}^2} \right]^{\frac{1}{2p+5}},$$

en función de si $v+p$ es impar o par, respectivamente.

$$C_{0,6}(K) = \left[\frac{8!^2 (0+1) \int_{-1}^1 K_0^{*2}(t) dt}{2(8-0) \left\{ \int_{-1}^1 t^8 K_0^*(t) dt \right\}^2} \right]^{\frac{1}{17}} = \left[\frac{8!^2 (4+1) \frac{1964655}{4352}}{2(8-2) \frac{784}{48841}} \right]^{\frac{1}{17}} = 6.2179$$

$$C_{1,6}(K) = \left[\frac{7!^2 (2+1) \int_{-1}^1 K_1^{*2}(t) dt}{2(7-1) \left\{ \int_{-1}^1 t^7 K_1^*(t) dt \right\}^2} \right]^{\frac{1}{15}} = \left[\frac{7!^2 (2+1) \frac{2205}{64}}{2(7-1) \frac{49}{20449}} \right]^{\frac{1}{15}} = 5.379$$

$$C_{2,6}(K) = \left[\frac{8!^2 (4+1) \int_{-1}^1 K_2^{*2}(t) dt}{2(8-2) \left\{ \int_{-1}^1 t^8 K_2^*(t) dt \right\}^2} \right]^{\frac{1}{17}} = \left[\frac{8!^2 (4+1) \frac{1964655}{4352}}{2(8-2) \frac{784}{48841}} \right]^{\frac{1}{17}} = 6.0419$$

$$C_{3,6}(K) = \left[\frac{7!^2 (6+1) \int_{-1}^1 K_3^{*2}(t) dt}{2(7-3) \left\{ \int_{-1}^1 t^7 K_3^*(t) dt \right\}^2} \right]^{\frac{1}{15}} = \left[\frac{7!^2 (6+1) \frac{10395}{16}}{2(7-3) \frac{49}{169}} \right]^{\frac{1}{15}} = 5.166$$

$$C_{4,6}(K) = \left[\frac{8!^2 (8+1) \int_{-1}^1 K_4^{*2}(t) dt}{2(8-4) \left\{ \int_{-1}^1 t^8 K_4^*(t) dt \right\}^2} \right]^{\frac{1}{17}} = \left[\frac{8!^2 (8+1) \frac{17342325}{4352}}{2(8-4) \frac{196}{289}} \right]^{\frac{1}{17}} = 5.8418$$

$$C_{5,6}(K) = \left[\frac{7!^2 (10+1) \int_{-1}^1 K_5^{*2}(t) dt}{2(7-5) \left\{ \int_{-1}^1 t^7 K_5^*(t) dt \right\}^2} \right]^{\frac{1}{15}} = \left[\frac{7!^2 (10+1) \frac{189 \cdot 189}{320}}{2(7-5) \frac{49}{25}} \right]^{\frac{1}{15}} = 4.878$$

$$C_{6,6}(K) = \left[\frac{8!^2 (12+1) \int_{-1}^1 K_6^{*2}(t) dt}{2(8-6) \left\{ \int_{-1}^1 t^8 K_6^*(t) dt \right\}^2} \right]^{\frac{1}{17}} = 5.5614$$

- Cálculo de las constantes $C_{v,p}(K)$ para $p = 5$

Procedemos de igual modo que para el caso $p = 6$. Calculados los elementos $(\mu_{j+l})_{j,l=0..p}$ se obtiene la matriz S para $p = 5$:

$$S = \begin{pmatrix} 1 & 0 & \frac{1}{5} & 0 & \frac{3}{35} & 0 \\ 0 & \frac{1}{5} & 0 & \frac{3}{35} & 0 & \frac{1}{21} \\ \frac{1}{5} & 0 & \frac{3}{35} & 0 & \frac{1}{21} & 0 \\ 0 & \frac{3}{35} & 0 & \frac{1}{21} & 0 & \frac{1}{33} \\ \frac{3}{35} & 0 & \frac{1}{21} & 0 & \frac{1}{33} & 0 \\ 0 & \frac{1}{21} & 0 & \frac{1}{33} & 0 & \frac{3}{143} \end{pmatrix}$$

$$S^{-1} = \begin{pmatrix} \frac{175}{64} & 0 & -\frac{525}{32} & 0 & \frac{1155}{64} & 0 \\ 0 & \frac{3675}{64} & 0 & -\frac{8085}{32} & 0 & \frac{15015}{64} \\ -\frac{525}{32} & 0 & \frac{3045}{16} & 0 & -\frac{8085}{32} & 0 \\ 0 & -\frac{8085}{32} & 0 & \frac{21945}{16} & 0 & -\frac{45045}{32} \\ \frac{1155}{64} & 0 & -\frac{8085}{32} & 0 & \frac{24255}{64} & 0 \\ 0 & \frac{15015}{64} & 0 & -\frac{45045}{32} & 0 & \frac{99099}{64} \end{pmatrix}$$

$$K_0^*(t) = \left(\frac{175}{64} - \frac{525}{32}t^2 + \frac{1155}{64}t^4 \right) K(t)$$

$$K_1^*(t) = \left(\frac{3675}{64}t - \frac{8085}{32}t^3 + \frac{15015}{64}t^5 \right) K(t)$$

$$K_2^*(t) = \left(-\frac{525}{32} + \frac{3045}{16}t^2 - \frac{8085}{32}t^4 \right) K(t)$$

$$K_3^*(t) = \left(-\frac{8085}{32}t + \frac{21945}{16}t^3 - \frac{45045}{32}t^5 \right) K(t)$$

$$K_4^*(t) = \left(\frac{1155}{64} - \frac{8085}{32}t^2 + \frac{24\,255}{64}t^4 \right) K(t)$$

$$K_5^*(t) = \left(\frac{15\,015}{64}t - \frac{45\,045}{32}t^3 + \frac{99\,099}{64}t^5 \right) K(t)$$

$$\int_{-1}^1 K_0^*(t)^2 dt = \int_{-1}^1 \left(\frac{175}{64} - \frac{525}{32}t^2 + \frac{1155}{64}t^4 \right)^2 K(t)^2 dt = \frac{1575}{832}$$

$$\int_{-1}^1 K_1^*(t)^2 dt = \int_{-1}^1 \left(\frac{3675}{64}t - \frac{8085}{32}t^3 + \frac{15\,015}{64}t^5 \right)^2 K(t)^2 dt = \frac{2205}{64}$$

$$\int_{-1}^1 K_2^*(t)^2 dt = \int_{-1}^1 \left(-\frac{525}{32} + \frac{3045}{16}t^2 - \frac{8085}{32}t^4 \right)^2 K(t)^2 dt = \frac{19\,845}{208}$$

$$\int_{-1}^1 K_3^*(t)^2 dt = \int_{-1}^1 \left(-\frac{8085}{32}t + \frac{21\,945}{16}t^3 - \frac{45\,045}{32}t^5 \right)^2 K(t)^2 dt = \frac{10\,395}{16}$$

$$\int_{-1}^1 K_4^*(t)^2 dt = \int_{-1}^1 \left(\frac{1155}{64} - \frac{8085}{32}t^2 + \frac{24\,255}{64}t^4 \right)^2 K(t)^2 dt = \frac{121\,275}{832}$$

$$\int_{-1}^1 K_5^*(t)^2 dt = \int_{-1}^1 \left(\frac{15\,015}{64}t - \frac{45\,045}{32}t^3 + \frac{99\,099}{64}t^5 \right)^2 K(t)^2 dt = \frac{189\,189}{320}$$

$$\left\{ \int t^{p+1} K_0^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^6 \left(\frac{175}{64} - \frac{525}{32}t^2 + \frac{1155}{64}t^4 \right) K(t) dt \right)^2 = \frac{25}{184\,041}$$

$$\left\{ \int t^{p+1} K_1^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^6 \left(\frac{3675}{64}t - \frac{8085}{32}t^3 + \frac{15\,015}{64}t^5 \right) K(t) dt \right)^2 = 0$$

$$\left\{ \int t^{p+1} K_2^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^6 \left(-\frac{525}{32} + \frac{3045}{16}t^2 - \frac{8085}{32}t^4 \right) K(t) dt \right)^2 = \frac{2025}{20\,449}$$

$$\left\{ \int t^{p+1} K_3^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^6 \left(-\frac{8085}{32}t + \frac{21\,945}{16}t^3 - \frac{45\,045}{32}t^5 \right) K(t) dt \right)^2 = 0$$

$$\left\{ \int t^{p+1} K_4^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^6 \left(\frac{1155}{64} - \frac{8085}{32}t^2 + \frac{24\,255}{64}t^4 \right) K(t) dt \right)^2 = \frac{225}{169}$$

$$\left\{ \int t^{p+1} K_5^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^6 \left(\frac{15\,015}{64}t - \frac{45\,045}{32}t^3 + \frac{99\,099}{64}t^5 \right) K(t) dt \right)^2 = 0$$

$$\left\{ \int t^{p+2} K_0^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^7 \left(\frac{175}{64} - \frac{525}{32}t^2 + \frac{1155}{64}t^4 \right) K(t) dt \right)^2 = 0$$

$$\left\{ \int t^{p+2} K_1^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^7 \left(\frac{3675}{64}t - \frac{8085}{32}t^3 + \frac{15\,015}{64}t^5 \right) K(t) dt \right)^2 = \frac{49}{20\,449}$$

$$\left\{ \int t^{p+2} K_2^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^7 \left(-\frac{525}{32} + \frac{3045}{16}t^2 - \frac{8085}{32}t^4 \right) K(t) dt \right)^2 = 0$$

$$\left\{ \int t^{p+2} K_3^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^7 \left(-\frac{8085}{32}t + \frac{21\,945}{16}t^3 - \frac{45\,045}{32}t^5 \right) K(t) dt \right)^2 = \frac{49}{169}$$

$$\left\{ \int t^{p+2} K_4^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^7 \left(\frac{1155}{64} - \frac{8085}{32}t^2 + \frac{24\,255}{64}t^4 \right) K(t) dt \right)^2 = 0$$

$$\left\{ \int t^{p+2} K_5^*(t) dt \right\}^2 = \left(\int_{-1}^1 t^7 \left(\frac{15\,015}{64}t - \frac{45\,045}{32}t^3 + \frac{99\,099}{64}t^5 \right) K(t) dt \right)^2 = \frac{49}{25}$$

De forma similar calculamos los valores para $C_{v,5}(K)$

$$C_{0,5}(K) = \left[\frac{6!^2 (0+1) \int_{-1}^1 K_0^{*2}(t) dt}{2(6-0) \left\{ \int_{-1}^1 t^6 K_0^*(t) dt \right\}^2} \right]^{\frac{1}{13}} = \left[\frac{6!^2 (0+1) \frac{1575}{832}}{2(6-0) \frac{25}{184\,041}} \right]^{\frac{1}{13}} = 4.7354$$

$$C_{1,5}(K) = \left[\frac{7!^2 (2+1) \int_{-1}^1 K_1^{*2}(t) dt}{2(7-1) \left\{ \int_{-1}^1 t^7 K_1^*(t) dt \right\}^2} \right]^{\frac{1}{15}} = 5.379$$

$$C_{2,5}(K) = \left[\frac{6!^2 (4+1) \int_{-1}^1 K_2^{*2}(t) dt}{2(6-2) \left\{ \int_{-1}^1 t^6 K_2^*(t) dt \right\}^2} \right]^{\frac{1}{15}} = 4.5021$$

$$C_{3,5}(K) = \left[\frac{7!^2 (6+1) \int_{-1}^1 K_3^{*2}(t) dt}{2(7-3) \left\{ \int_{-1}^1 t^7 K_3^*(t) dt \right\}^2} \right]^{\frac{1}{15}} = 5.166$$

$$C_{4,5}(K) = \left[\frac{6!^2 (8+1) \int_{-1}^1 K_4^{*2}(t) dt}{2(6-4) \left\{ \int_{-1}^1 t^6 K_4^*(t) dt \right\}^2} \right]^{\frac{1}{15}} = 4.2029$$

$$C_{5,5}(K) = \left[\frac{7!^2 (10+1) \int_{-1}^1 K_5^{*2}(t) dt}{2(7-5) \left\{ \int_{-1}^1 t^7 K_5^*(t) dt \right\}^2} \right]^{\frac{1}{15}} = 4.878$$

- Cálculo de las constantes $C_{v,p}(K)$ para $p = 4$

Procediendo de forma similar a los caso $p = 6, 5$ obtenemos los siguientes resultados:

$$C_{0,4}(K) = \left[\frac{6!^2 (0+1) \int_{-1}^1 K_0^{*2}(t) dt}{2(6-0) \left\{ \int_{-1}^1 t^6 K_0^*(t) dt \right\}^2} \right]^{\frac{1}{13}} = \left[\frac{6!^2 (0+1) \frac{1575}{832}}{2(6-0) \frac{25}{184041}} \right]^{\frac{1}{13}} = 4.7354$$

$$C_{1,4}(K) = \left[\frac{5!^2 (2+1) \int_{-1}^1 K_1^{*2}(t) dt}{2(5-1) \left\{ \int_{-1}^1 t^5 K_1^*(t) dt \right\}^2} \right]^{\frac{1}{11}} = 3.8565$$

$$C_{2,4}(K) = \left[\frac{6!^2 (4+1) \int_{-1}^1 K_2^{*2}(t) dt}{2(6-2) \left\{ \int_{-1}^1 t^6 K_2^*(t) dt \right\}^2} \right]^{\frac{1}{13}} = 4.5021$$

$$C_{3,4}(K) = \left[\frac{5!^2 (6+1) \int_{-1}^1 K_3^{*2}(t) dt}{2(5-3) \left\{ \int_{-1}^1 t^5 K_3^*(t) dt \right\}^2} \right]^{\frac{1}{11}} = 2.9446$$

$$C_{4,4}(K) = \left[\frac{6!^2 (8+1) \int_{-1}^1 K_4^{*2}(t) dt}{2(6-4) \left\{ \int_{-1}^1 t^6 K_4^*(t) dt \right\}^2} \right]^{\frac{1}{13}} = 4.202$$

- Cálculo de las constantes $C_{v,p}(K)$ para $p = 3$

$$C_{0,3}(K) = \left[\frac{4!^2 (0+1) \int_{-1}^1 K_0^{*2}(t) dt}{2(4-0) \left\{ \int_{-1}^1 t^4 K_0^*(t) dt \right\}^2} \right]^{\frac{1}{9}} = \left[\frac{4!^2 (0+1) \frac{5}{4}}{2(4-0) \frac{1}{441}} \right]^{\frac{1}{9}} = 3.2431$$

$$C_{1,3}(K) = \left[\frac{5!^2 (2+1) \int_{-1}^1 K_1^{*2}(t) dt}{2(5-1) \left\{ \int_{-1}^1 t^5 K_1^*(t) dt \right\}^2} \right]^{\frac{1}{11}} = 3.8565$$

$$C_{2,3}(K) = \left[\frac{4!^2 (4+1) \int_{-1}^1 K_2^{*2}(t) dt}{2(4-2) \left\{ \int_{-1}^1 t^4 K_2^*(t) dt \right\}^2} \right]^{\frac{1}{9}} = 2.8925$$

$$C_{3,3}(K) = \left[\frac{(6)!^2 (6+1) \int K_v^{*2}(t) dt}{2(2) \left\{ \int t^5 K_v^*(t) dt \right\}^2} \right]^{\frac{1}{11}} = 3.539$$

- Cálculo de las constantes $C_{v,p}(K)$ para $p = 2$

$$C_{0,2}(K) = \left[\frac{4!^2 (0+1) \int_{-1}^1 K_0^{*2}(t) dt}{2(4-0) \left\{ \int_{-1}^1 t^4 K_0^*(t) dt \right\}^2} \right]^{\frac{1}{9}} = \left[\frac{4!^2 (0+1) \frac{5}{4}}{2(4-0) \frac{1}{441}} \right]^{\frac{1}{9}} = 3.2431$$

$$C_{1,2}(K) = \left[\frac{3!^2 (2+1) \int_{-1}^1 K_1^{*2}(t) dt}{2(3-1) \left\{ \int_{-1}^1 t^3 K_1^*(t) dt \right\}^2} \right]^{\frac{1}{7}} = 2.2746$$

$$C_{2,2}(K) = \left[\frac{4!^2 (4+1) \int_{-1}^1 K_2^{*2}(t) dt}{2(4-2) \left\{ \int_{-1}^1 t^4 K_2^*(t) dt \right\}^2} \right]^{\frac{1}{9}} = 2.8925$$

- Cálculo de las constantes $C_{v,p}(K)$ para $p = 1$

En este último caso obtenemos:

$$C_{0,1}(K) = \left[\frac{2!^2 (0+1) \int_{-1}^1 K_0^{*2}(t) dt}{2(2-0) \left\{ \int_{-1}^1 t^2 K_0^*(t) dt \right\}^2} \right]^{\frac{1}{5}} = \left[\frac{2!^2 (0+1) \frac{3}{5}}{2(2-0) \frac{1}{25}} \right]^{\frac{1}{5}} = 1.7188$$

$$C_{1,1}(K) = \left[\frac{3!^2 (2+1) \int_{-1}^1 K_1^{*2}(t) dt}{2(3-1) \left\{ \int_{-1}^1 t^2 K_1^*(t) dt \right\}^2} \right]^{\frac{1}{7}} = \left[\frac{3!^2 (2+1) \frac{15}{7}}{2(3-1) \frac{9}{49}} \right]^{\frac{1}{7}} = 2.2746$$

Resumen de resultados obtenidos

A continuación, resumimos, en la tabla 7.1, los valores de las constantes $C_{v,p}$ obtenidos anteriormente para la función núcleo de Epanechnikov. Presentamos, también, los valores calculados con la función núcleo cuártico, cuya expresión es $K(t) = \frac{15}{16} ((1-t^2))^2 I_{(|t|\leq 1)}$.

v	p	$C_{v,p}$ Epanechnikov	$C_{v,p}$ Cuártico
0	1	1.719	2.036
1	1	2.275	2.586
0	2	3.243	3.633
1	2	2.275	2.586
2	2	2.893	3.208
0	3	3.243	3.633
1	3	3.857	4.234
2	3	2.893	3.208
3	3	3.539	3.860
0	4	4.735	5.158
1	4	3.857	4.234
2	4	4.502	4.874
3	4	2.945	3.860
4	4	4.203	4.530
0	5	4.735	5.158
1	5	5.379	5.789
2	5	4.502	4.874
3	5	5.166	5.536
4	5	4.203	4.530
5	5	4.878	5.210
0	6	6.218	6.659
1	6	5.379	5.789
2	6	6.042	6.444
3	6	5.166	5.536
4	6	5.842	6.211
5	6	4.878	5.210
6	6	5.561	5.898

Tabla 7.1: Constantes para el cálculo de la ventana óptima.

Algoritmos computacionales

Para la implementación de los programas de ordenador, hemos utilizado el lenguaje de programación C++ (compilador DJGPP). El principal motivo de esta elección, radica en el hecho de que la casa Rheometric Scientific ha empleado este lenguaje en el desarrollo del software que acompaña a los aparatos STA, sugiriéndonos el empleo de este lenguaje con el objeto de integrar, en un futuro, nuestra rutinas de suavizado en la nueva versión software en desarrollo.

Mostraremos algunos de los diagramas de flujo que nos ayudan a comprender la lógica de las rutinas implementadas y determinar la secuencia de acciones a realizar para generar los algoritmos de suavizado. Nos hemos centrado en aquellos procedimientos o funciones específicas más relevantes de cada uno de los métodos empleados, intentando que los diagramas de flujo reflejen la lógica de los programas generados. Indicar que p representa el grado del polinomio utilizado para la estimación y nu indica el orden de la derivada que queremos obtener.

Programas desarrollados

Hemos creado varios programas cuyas tareas principales y diagramas de flujo se enumeran, resumidamente, a continuación:

- Estimación automática de la curva y sus derivadas con ventana local (figuras 6.6 a 6.10).
- Estimación automática de la curva y sus derivadas con ventana global (figuras 6.11 y 6.16)..
- Clasificación de materiales mediante discriminación no paramétrica (figuras 6.17, 6.18 y 6.19).
- Simulación de curvas TGA de mediante mezcla de funciones logísticas (figura 6.20).

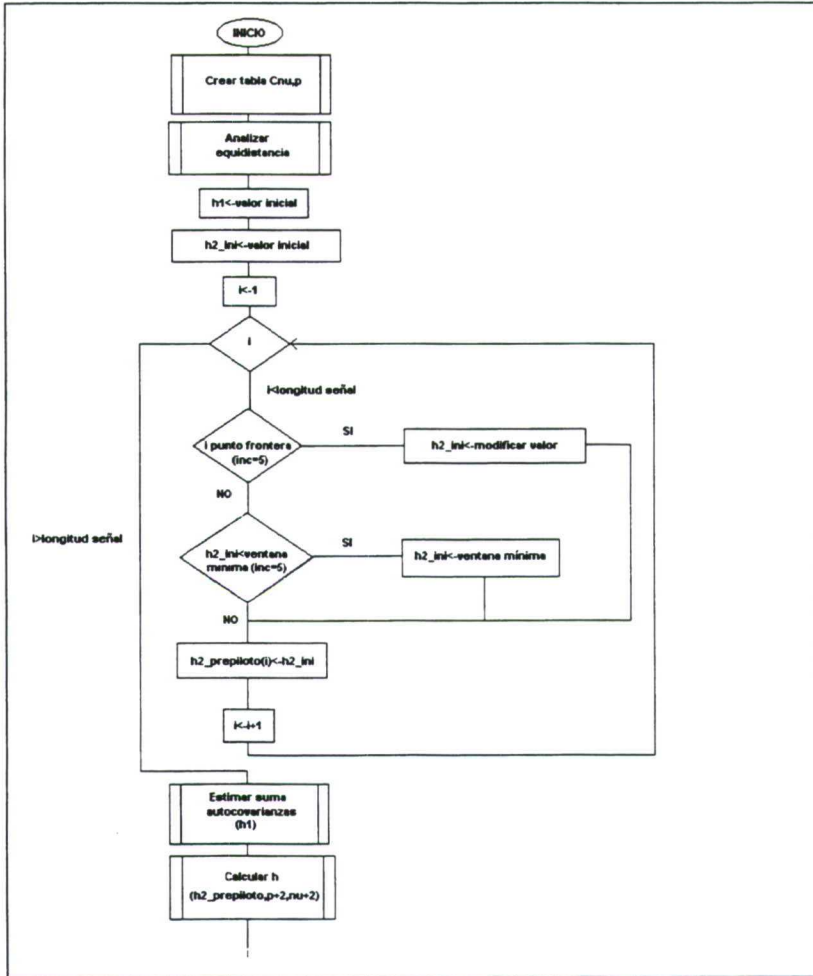


Figura 6.6: Diagrama de flujo para la suavización con ventana local (Parte I).

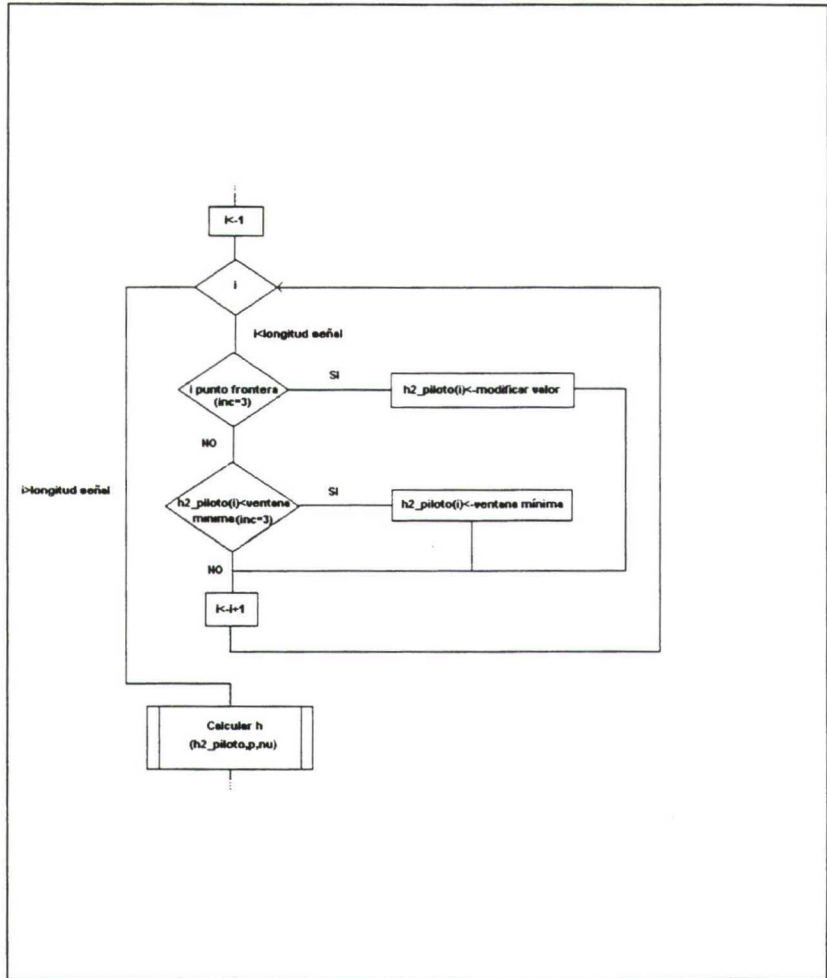


Figura 6.7: Diagrama de flujo para la suavización con ventana local (Parte II).

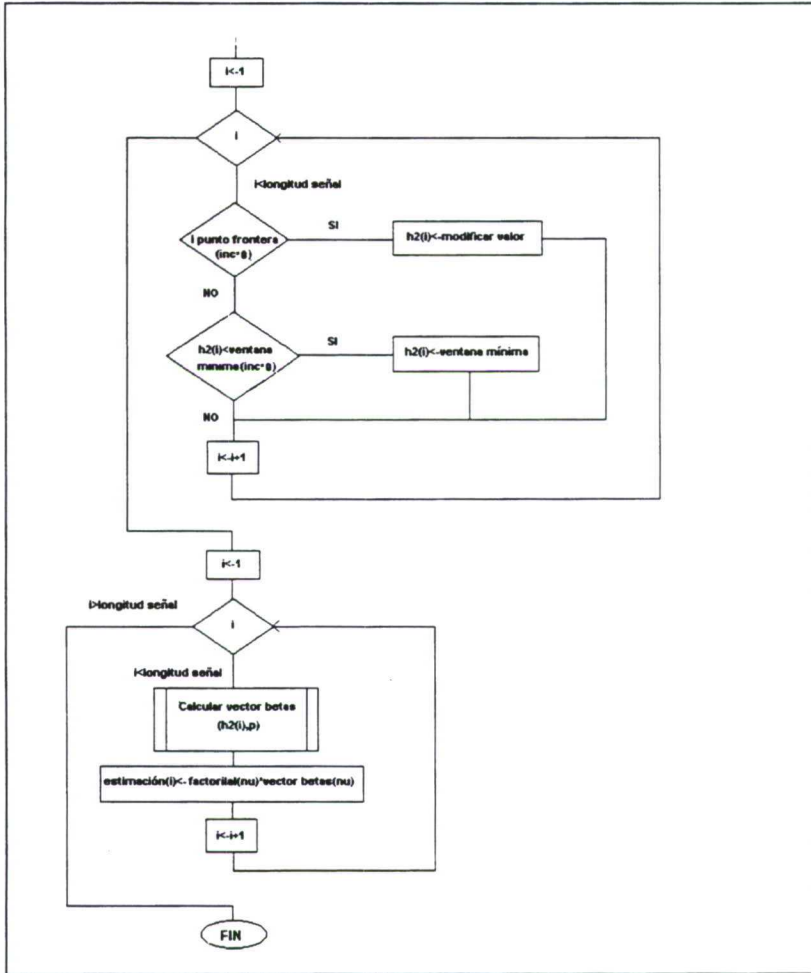


Figura 6.8: Diagrama de flujo para la suavización con ventana local (Parte III).

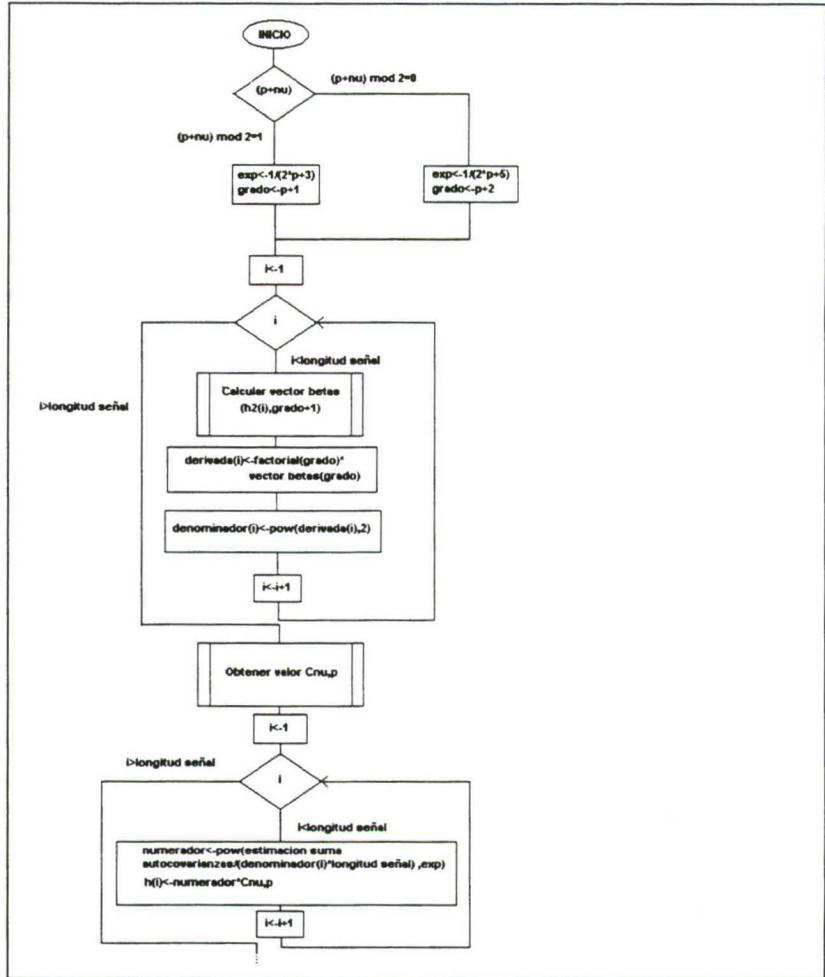


Figura 6.9: Diagrama de flujo para el cálculo de la ventana local (Parte I).

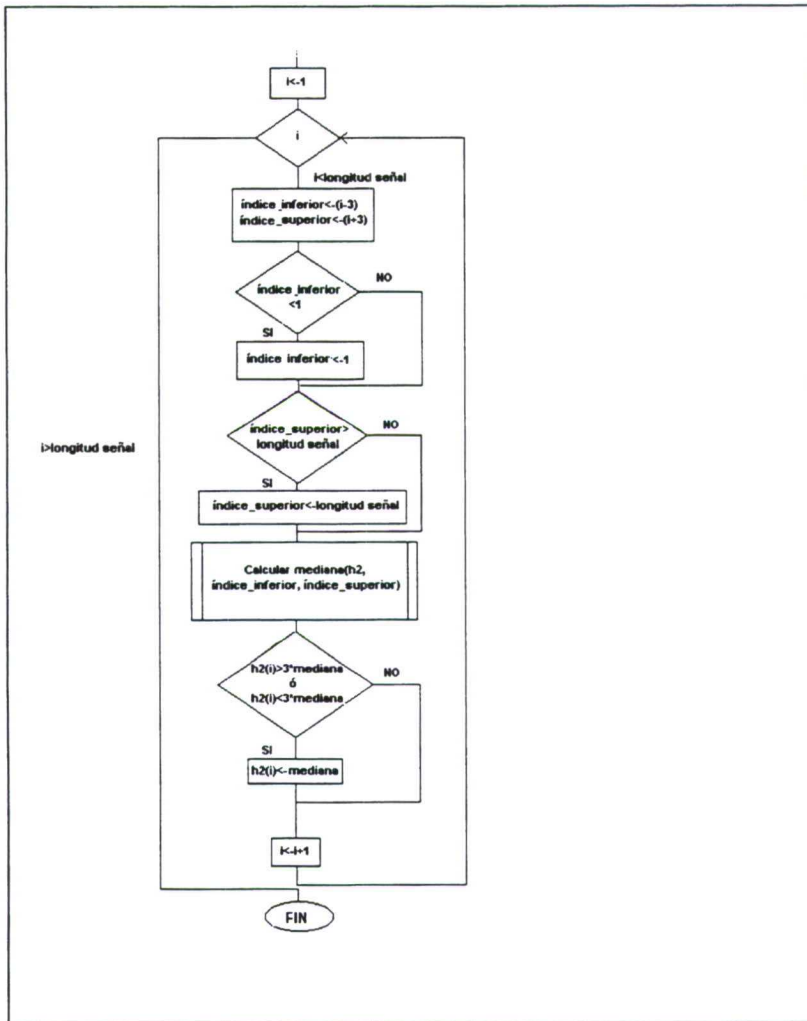


Figura 6.10: Diagrama de flujo para el cálculo de la ventana local (Parte II).

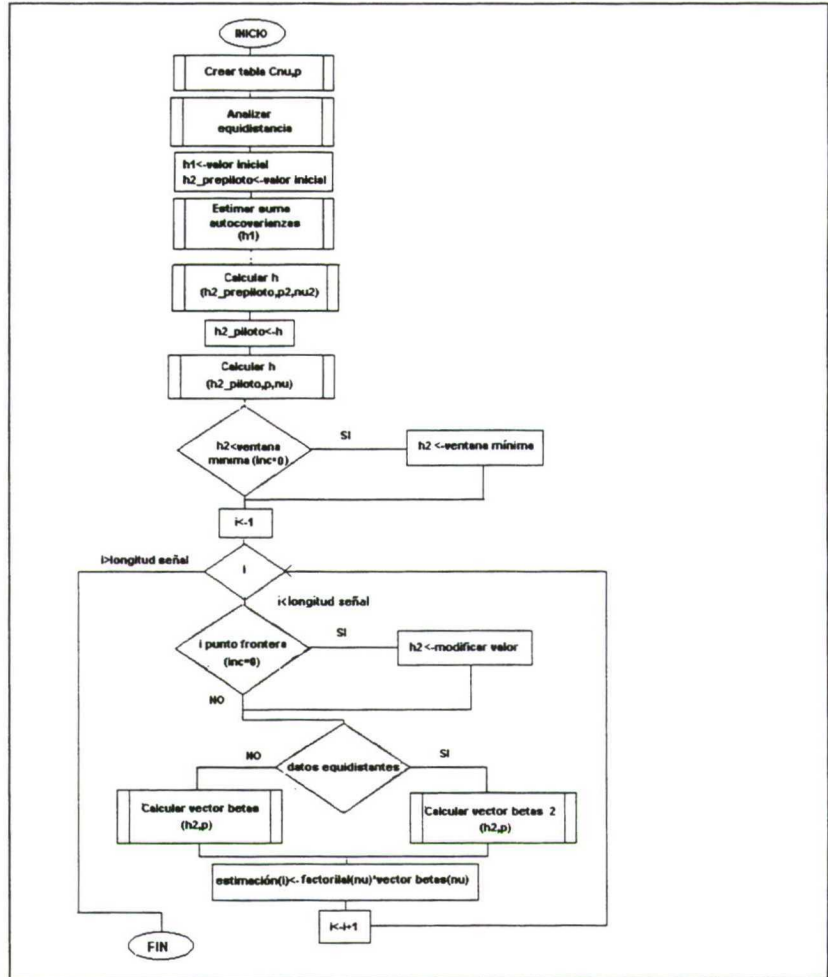


Figura 6.11: Diagrama de flujo para la suavización con ventana global.

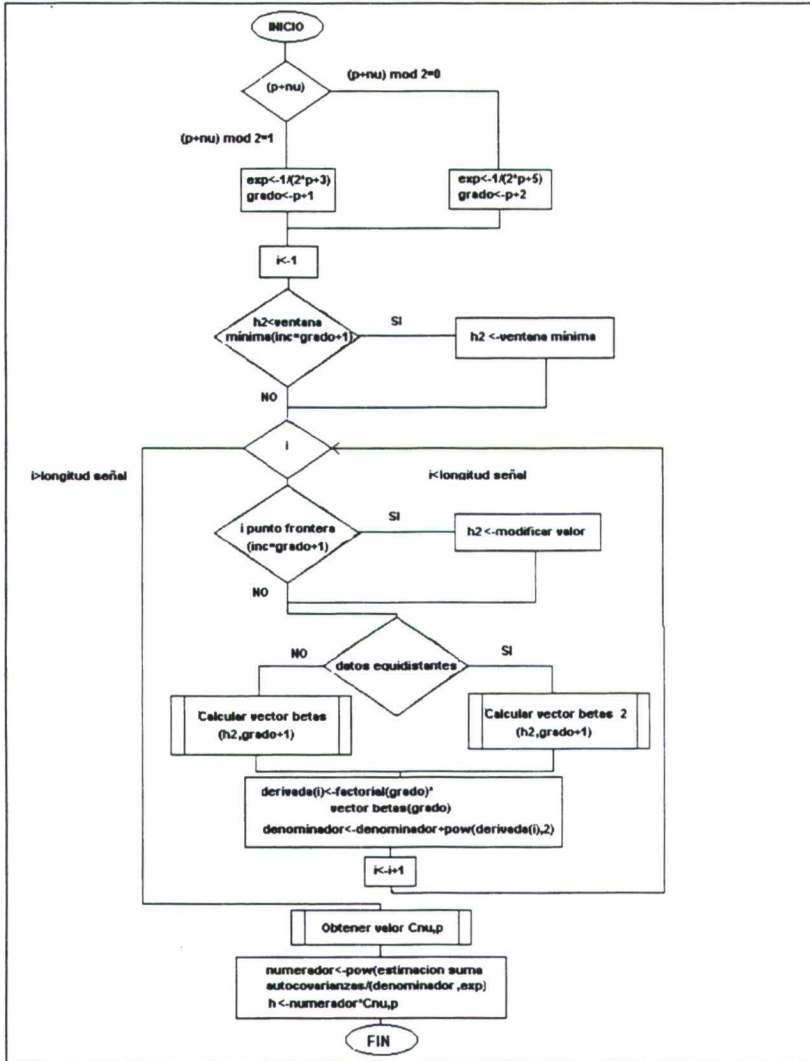


Figura 6.12: Diagrama de flujo para el cálculo de la ventana global.

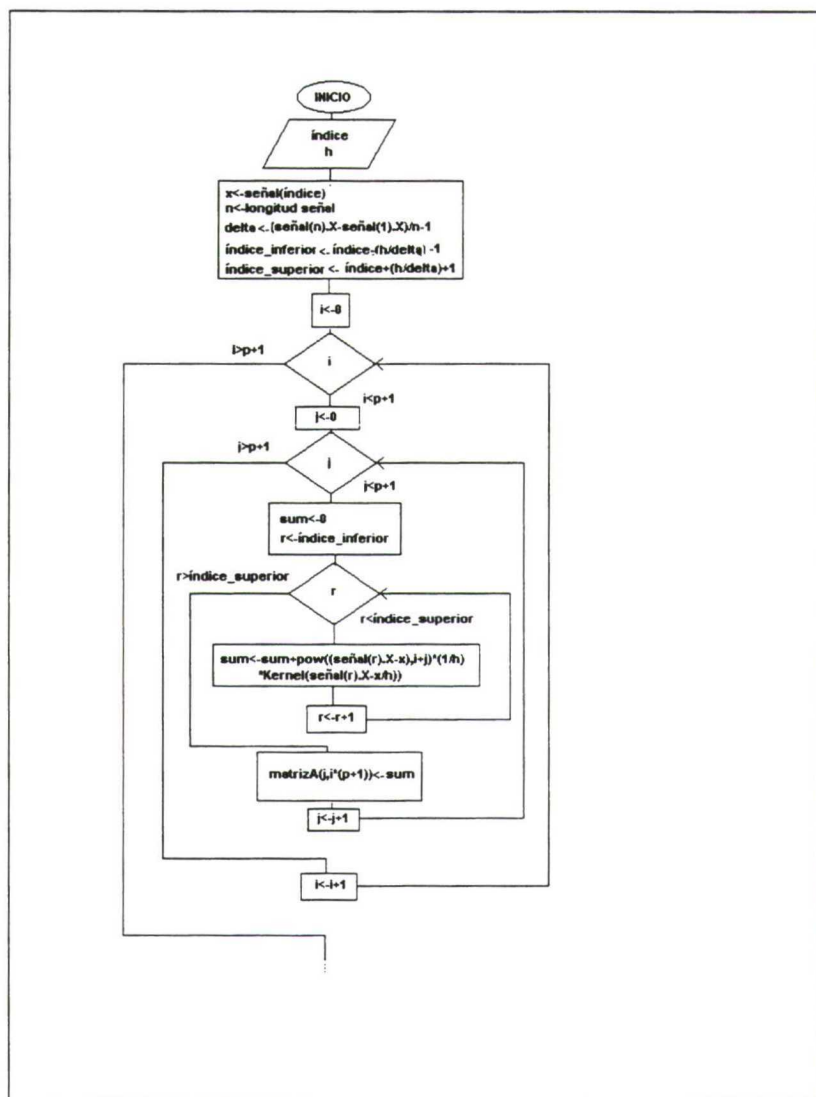


Figura 6.13: Diagrama de flujo para el cálculo del vector de estimaciones (Parte I).

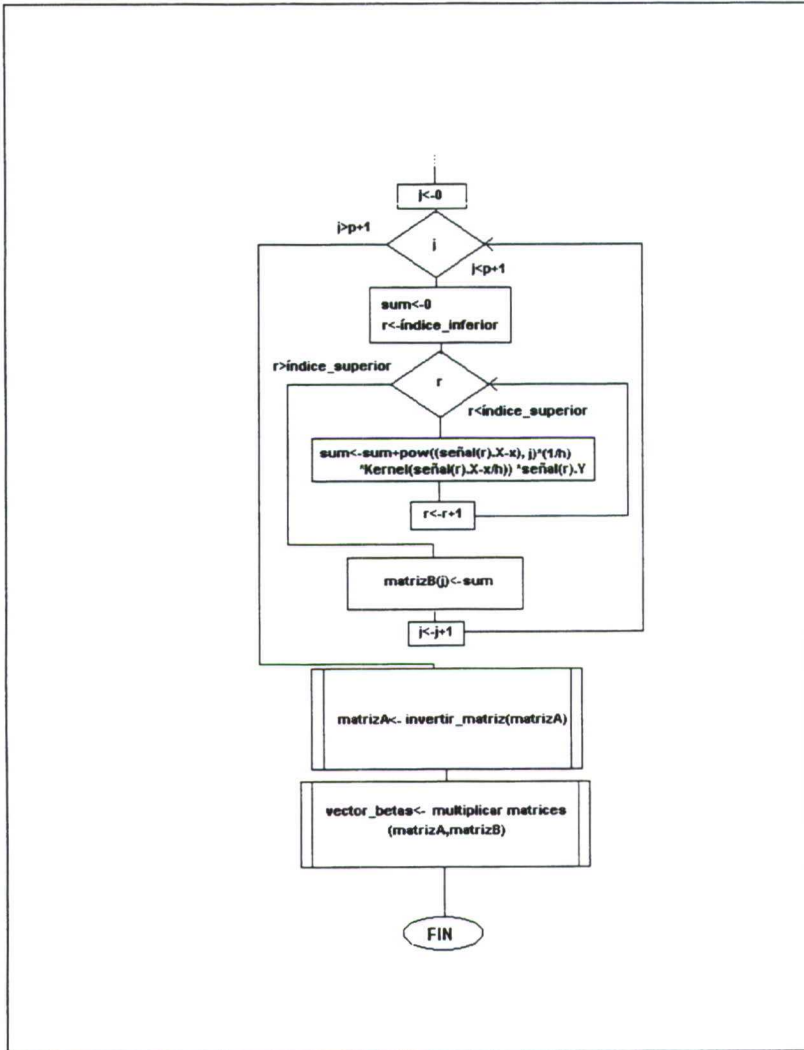


Figura 6.14: Diagrama de flujo para el cálculo del vector de estimaciones (Parte II).

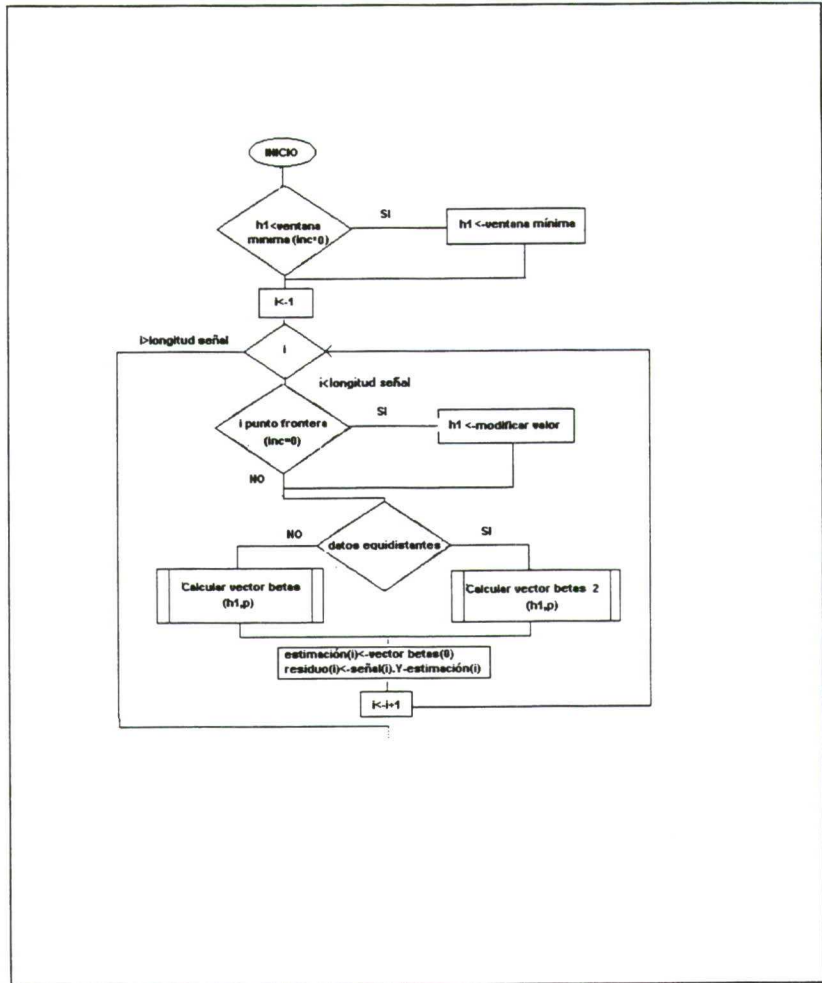


Figura 6.15: Diagrama de flujo para el cálculo del estimador de la suma de autocovarianzas (Parte I).

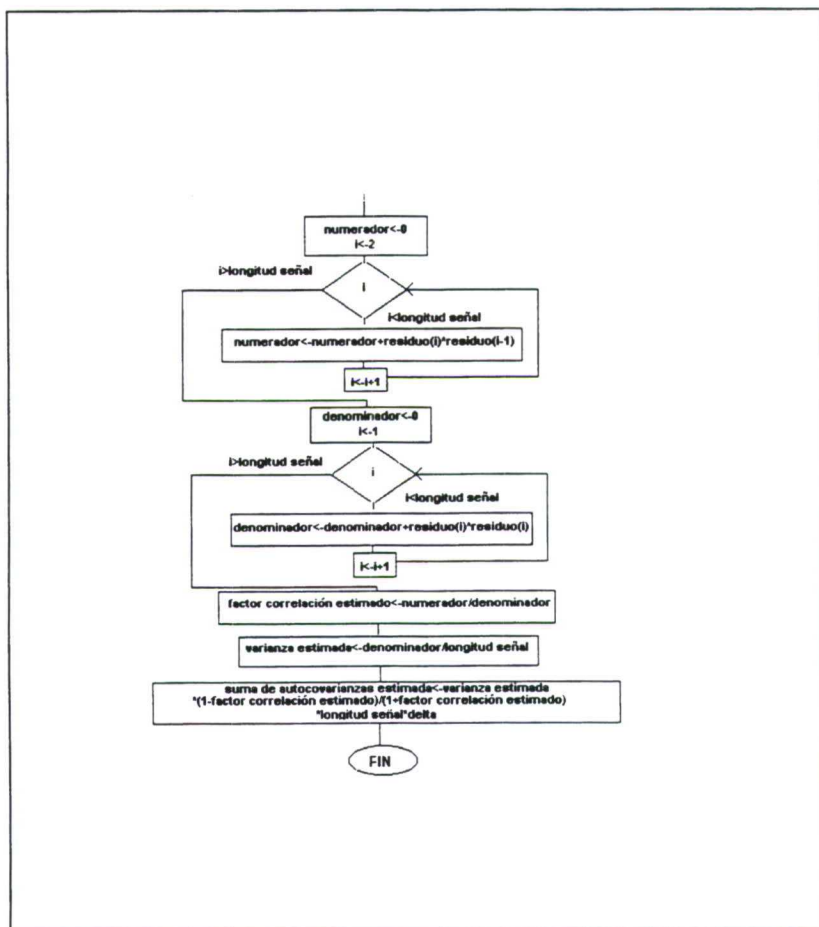


Figura 6.16: Diagrama de flujo para el cálculo del estimador de la suma de autocovarianzas (Parte II).

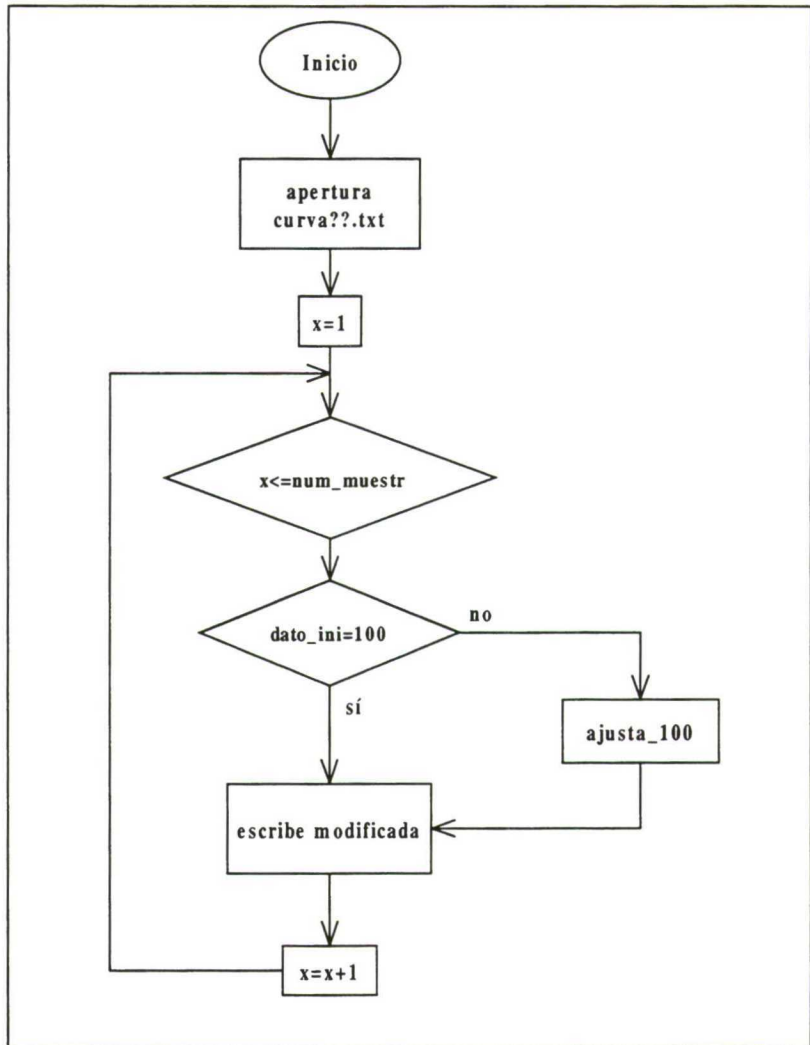


Figura 6.17: Diagrama de flujo para la lectura de datos y manipulaciones iniciales en la clasificación de materiales.

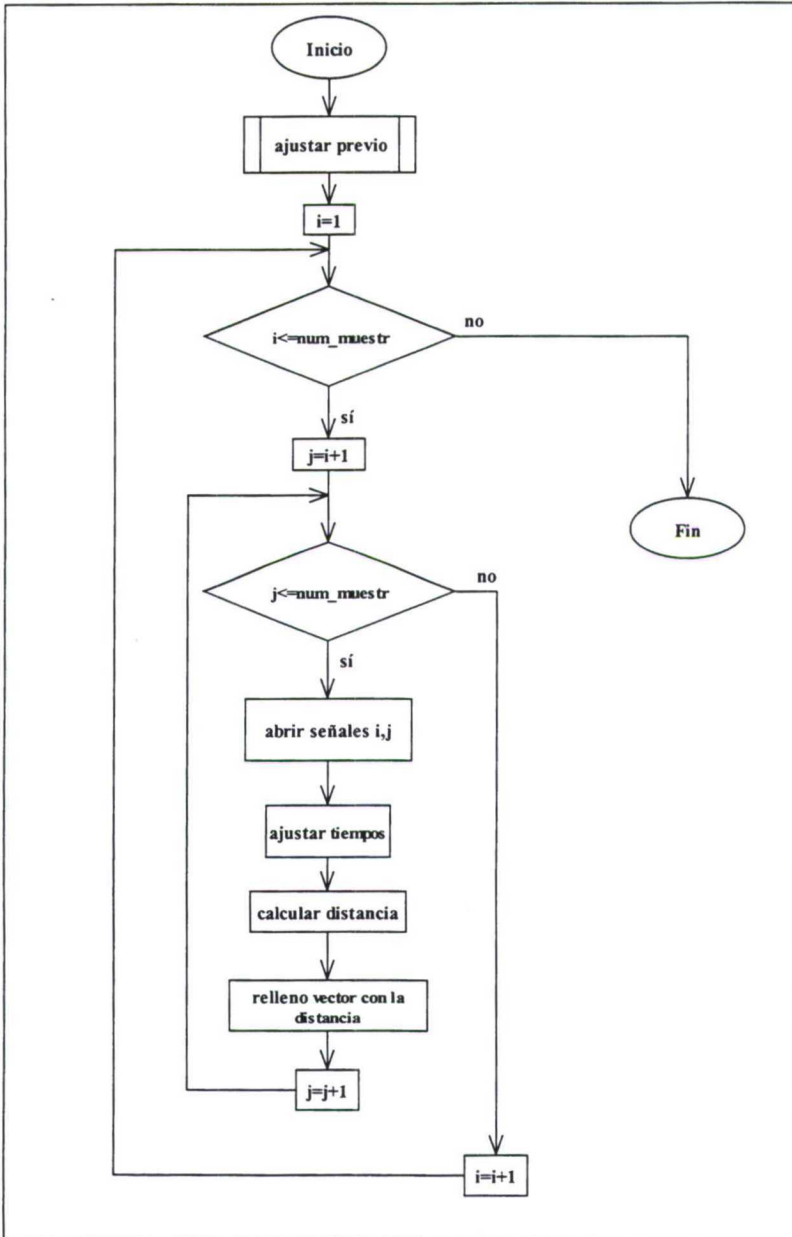


Figura 6.18: Diagrama de flujo para el cálculo de la distancia entre curvas en la clasificación de materiales.

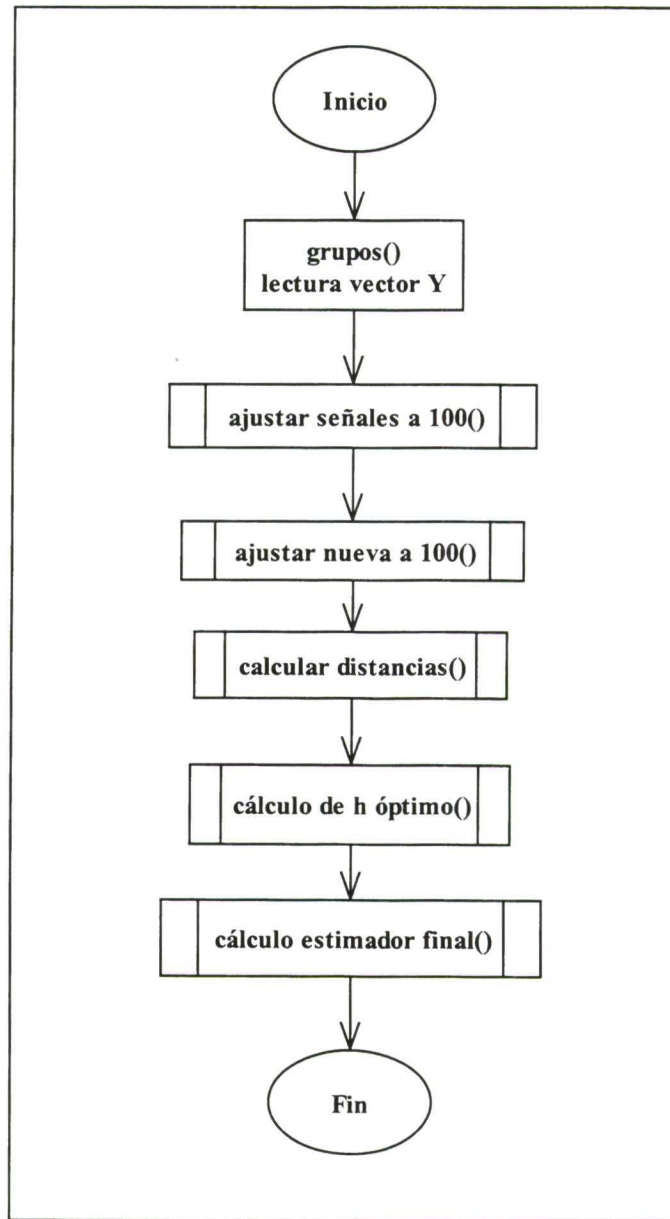


Figura 6.19: Diagrama de flujo general para clasificación de materiales.

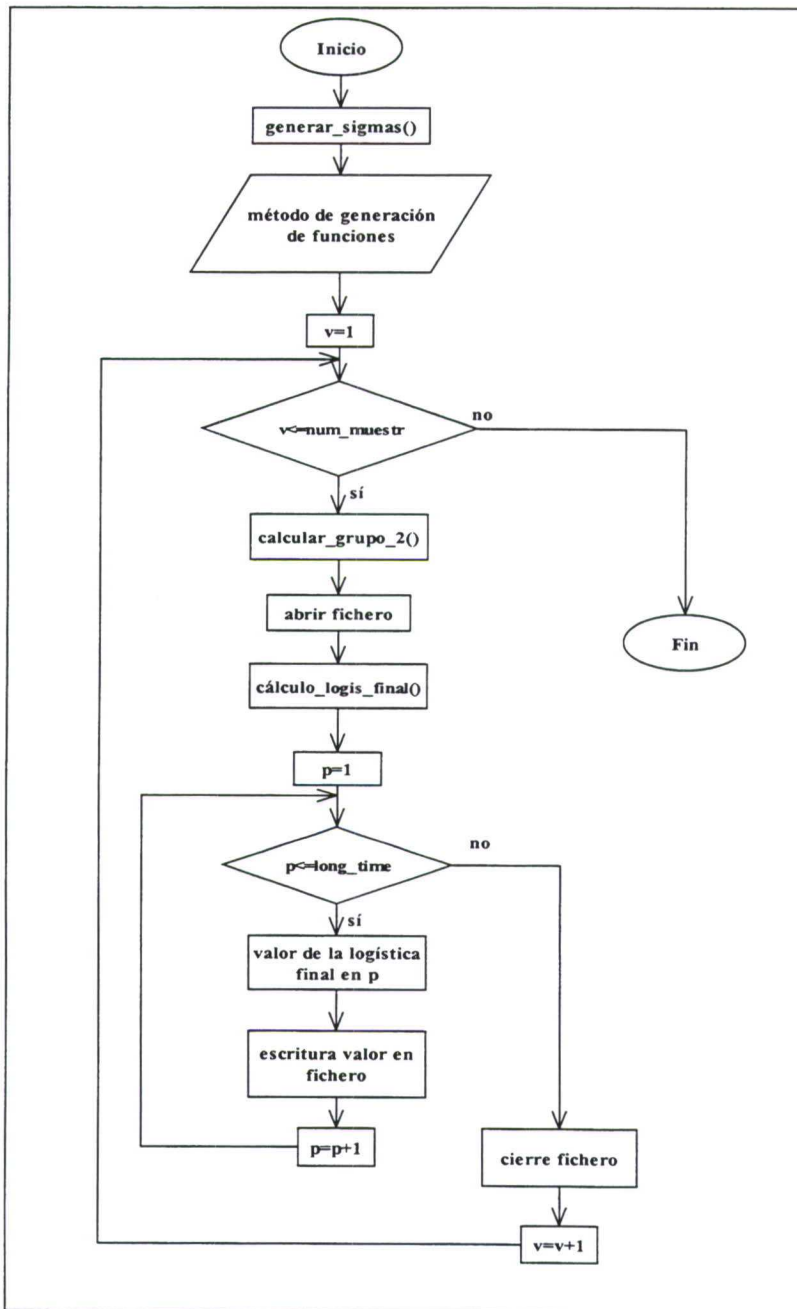


Figura 6.20: Diagrama de flujo para la simulación de curvas TGA.

Bibliografía

- [1] Adams, J., Kneller, W. y Dollimore, D. (1992). Thermal analysis (TA) of lime and gypsum-base medieval mortars. *Thermochimica Acta*, **211**, 93-106.
- [2] Arnold, M., Veress, G. E., Paulik, J. y Paulik, F. (1982) A critical reflection upon the application of the Arrhenius model to non-isothermal thermogravimetric curves. *Thermochimica Acta*, **52**, 67-81.
- [3] Artiaga, R., Cao, R., Naya, S. y Trillo, J. (2003). Nonparametric two-stage plug-in adaptive smoothing for thermal analysis data. *Enviado a la revista Journal of Statistical Computation and Simulation*.
- [4] Artiaga, R., Cao, R. y Naya, S. (2001). Local Polynomial estimation of TGA derivatives using logistic regression for pilot bandwidth selection. *Actas de la 5th Mediterranean Conference on Calorimetry and Thermal Analysis*, 86.
- [5] Artiaga, R., García, A., Varela, A., Mier, J. L., Naya, S. y Graña, M (2002). DMTA study of a nickel-titanium wire. *Journal of Thermal Analysis and Calorimetry*, **70**, 199-207.
- [6] Artiaga, R., Naya, S., Cao, R. García, A. y Mier, J.L. (2003). Application of generalized logistic estimation to epoxies thermal degradation kinetics. *Enviado y aceptado en el ICCE10*.
- [7] Braun, R. L. y Burnham, A. K. (1987). Analysis of chemical reaction kinetics using a distribution of activation energies and simpler models. *Energy & Fuels* **1**, 153-161.
- [8] Barlett, M.S. (1946). On the theoretical specification and sampling properties of autocorrelated time series. *Suppl. Journal of the Royal Statistical Society*, **8**, 27-41.
- [9] Besse, P., Cardot, H y Ferraty, F. (1997). Simultaneous non-parametric regressions of unbalanced longitudinal data. *Computational Statistical and Data Analysis*, **24**, 225-270.

- [10] Bowman, A. W. (1984). An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*, **71**, 353-360.
- [11] Brockwell, P. y Davis, R.A. (1991). *Time Series. Theory and Methods*. Springer, New York.
- [12] Broomhead, D.S. y Lowe, D. (1988). *Multivariable functional interpolation and adaptive networks*. Complex System. New York.
- [13] Brown M.E., Maciejewski M., Vyazovkin S., Nomen R., Sempere J., Burnham A., Opfermann J., Strey R., Anderson H.L., Kemmler A., Keuleers R., Janssens J., Desseyn H.O., Chao L., Tong B., Roduit B., Malek J. and Mitsuhashi T. (2000). Computational aspects of kinetic analysis. *Thermochimica Acta*, **355**, 125-143.
- [14] Cao, R. (2002). *Introducción a la Simulación y a la Teoría de Colas*. Netbiblo, A Coruña.
- [15] Cao, R., Quintela, A., Vilar-Fernández, J.M. (1993). Bandwidth selection in nonparametric density estimation under dependence: a simulation study. *Computational Statistics*, **8**, 313-332.
- [16] Carrasco F. y Costa J. (1989). Modelo Cinético de la descomposición térmica del poliestireno. *Revista de Ingeniería Química*. **3**, 121-129.
- [17] Chambers, J. M. y Hastie, T. J. (1992). *Statistical Models in Statistics. Nonlinear Models*. Wadsworth & Brooks. New York.
- [18] Chatterjee, P. K. y Conrad C. (1970). *Thermal degradation of wood components*. U. S. D. A. Forest Service Research Paper. Madison.
- [19] Cleveland, W.S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, **74**, 829-836.
- [20] Cleveland, W.S. y Devlin, S.J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, **83**, 597-610.
- [21] Conesa J. A. (2000). *Curso Básico de Análisis Térmico*. Editorial Club Universitario. Madrid.
- [22] Conesa J. A. y Marcilla A. (1996). Kinetic study of the thermogravimetric behavior of different rubbers. *Journal of Analytical and Applied Pyrolysis*, **37**, 95-110.
- [23] Dennis, J. E., Gay, D. M. y Welsch, R. E. (1981). An Adaptive Nonlinear Least-Squares Algorithm. *ACM Transactions on Mathematical Software*, **7**, 348-368.

- [24] Dominguez, C. (2003). *Regresión no paramétrica funcional aplicada al reconocimiento de patrones de espectros termogravimétricos*. Proyecto de fin de carrera de Ingeniería Informática. A Coruña.
- [25] Doyle C. D. (1961) Kinetic analysis of thermogravimetric data. *Journal of Applied Polymer Science*, **15**, 285-292.
- [26] Eubank, R.L. (1988). *Spline Smoothing and Nonparametric Regression*. Marcel Dekker. New York.
- [27] Fan, J. (1992). Design-adaptive nonparametric regression. *Journal of the American Statistical Association*, **87**, 998-1004.
- [28] Fan, J. (1993). Local linear regression smoothers and their minimax efficiency. *The Annals of Statistics*, **21**, 196-216.
- [29] Fan, J. y Gijbels, I. (1992). Variable bandwidth and local linear regression smoothers. *The Annals of Statistics*, **20**, 2008-2036.
- [30] Fan, J. y Gijbels, I. (1995). Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. *Journal of the Royal Statistical Society, Series B*, **57**, 2, 371-394.
- [31] Fan, J. y Gijbels, I. (1996). *Local polynomial modelling and its applications*. Chapman and Hall. London.
- [32] Fan, J. y Marron J. S. (1994). Fast implementations of nonparametric curve estimators. *Journal of Computation and Graphical Statistics*, **3**, 35-56.
- [33] Fan, J., Gijbels, I., Hu, T. C. y Huang, L. S. (1996). A study of variable bandwidth selection for local polynomial regression. *Statistica Sinica*, **6**, 113-127.
- [34] Fan, J., Gasser, T., Gijbels, I., Brockmann, M. y Engel, J. (1997). Local polynomial regression: optimal kernels and asymptotic minimax efficiency. *Annals of the Institute of Statistical Mathematics*, **49**, 1, 79-99.
- [35] Ferraty, F. Nuñez-Antón, V. y Vieu P. (2001). *Regression no parametrica: desde la dimension uno hasta la dimension infinita*. Servicio Editorial de la Universidad del Pais Vasco. Bilbao.
- [36] Ferraty, F. y Vieu P. (2001). *Statistical Functionnelle: Modèles de Régression pour Variables Aléatoires Uni, Multi et Infiniment Dimensionnées*. Technical Report, Laboratoire de Statistique et Probabilités. Toulouse.
- [37] Ferraty, F. y Vieu P. (2002a). The functional nonparametric model and application to spectrometric data. *Computational Statistical*, **17**, 4, 545-564.

- [38] Ferraty, F. y Vieu P. (2002b). Nonparametric models for functional data, with application in regression, time series prediction and curve discrimination. *Pendiente de publicación en Computational Statistics and Data Analysis*.
- [39] Ferraty, F. y Vieu P. (2003). Curves Discrimination a Nonparametric Functional Approach. *Pendiente de publicación en Computational Statistics and Data Analysis*.
- [40] Flynn, J. H. y Dickens, B. (1976). Steady-State Parameter-Jump Methods and Relaxation Methods in Thermogravimetry. *Thermochimica Acta*, **15**, 1-16.
- [41] Francisco-Fernández, M. (2001). *La regresión polinómica local en diseño fijo con observaciones dependientes*. Tesis Doctoral. Universidad de Santiago de Compostela.
- [42] Francisco-Fernández, M. y Vilar-Fernández, J.M. (2001). Local polynomial regression estimation with correlated errors. *Communications in Statistics: Theory and Methods*, **30**, 7, 1271-1293.
- [43] Freeman B. y Carroll B. (1958). The application of thermoanalytical techniques to reaction kinetics. The thermogravimetric evaluation of the kinetic of the decomposition of calcium oxalate monhydrate. *Journal of Physical Chemistry*, **62**, 394-397.
- [44] Friedman H. L. (1964). Kinetics of thermal degradation of char forming plastics from thermogravimetry. Application to a phenolic plastic. *Journal of Polymer Science. Part C*, **6**, 183-195.
- [45] Gasser, T. y Müller, H.G. (1979). Kernel estimation of regression functions. In *Smoothing Techniques for Curve Estimation*. Lecture Notes in Mathematics, **757**, 23-68. Springer-Verlag, New York.
- [46] Gasser, T., Müller, H.G. y Mammitzsch, V. (1985). Kernels for nonparametric curve estimation. *Journal of the Royal Statistical Society, Series B*, **47**, 238-252.
- [47] Gay, D. M. (1984). *A trust region approach to linearly constrained optimization in Numerical Analysis*. Springer, Berlin, 171-189.
- [48] Härdle, W. (1986). Approximations to the mean integrated squared error with applications to optimal bandwidth selection for nonparametric regression function estimators. *Journal of Multivariate Analysis*, **18**, 150-168.
- [49] Härdle, W. (1990). *Applied nonparametric regression* Cambridge University Press. London.

- [50] Härdle, W. y Marron, J.S. (1985). Optimal bandwidth selection in non-parametric regression function estimation. *The Annals of Statistics*, **13**, 1465–1481.
- [51] Härdle, W. y Tsybakov, A. (1997). Local polynomial estimators of the volatility function in nonparametric autoregression. *Journal of Econometrics*, **81**, 223–242.
- [52] Hart, J. y Vieu, P. (1990). Data-driven bandwidth choice for density estimation based on dependent data. *The Annals of Statistics*, **18**, 873–890.
- [53] Hastie, T., Buja, A. y Tibshirani, R. (1995). Penalized discriminant analysis. *The Annals of Statistics*, **23**, 73–102.
- [54] Liu, X. (2001). *Kernel smoothing for spatially correlated data*. Tesis doctoral. University of Iowa.
- [55] Macauley, R.R. (1931). *The Smoothing of Time Series*. National Bureau of Economic Research, New York.
- [56] Marx, B. y Eilers, P. (1999). Generalized linear regression on sampled signals and curves: a P-spline approach. *Technometrics*, **41**, 1–13.
- [57] Masry, E. (1996a). Multivariate regression estimation: local polynomial fitting for time series. *Stochastic Processes and their Applications*, **65**, 81–101.
- [58] Masry, E. (1996b). Multivariate local polynomial regression for time series: uniform strong consistency and rates. *Journal of Time Series Analysis*, **17**, 571–599.
- [59] Masry, E. y Fan, J. (1997). Local polynomial estimation of regression function for mixing processes. *Scandinavian Journal of Statistics*, **24**, 165–179.
- [60] Nadaraya, E.A. (1964). Remarks on nonparametric estimates for density functions and regression curves. *Theory of Probability and Applications*, **15**, 134–137.
- [61] Naya, S. (2001). Regresión y Titanic. *Actas V Congreso Galego de Estatística e Investigación Operativa*, 137–140.
- [62] Naya, S., Cao, R., Artiaga, R., Barbadillo, F. y López de Ullibarri, I. (2003). Logistic mixture model vs Arrhenius for kinetic study of degradation of materials by dynamic thermogravimetric analysis. *Enviado a revista Journal of Chemometrics*.
- [63] Naya, S., Cao, R., Artiaga, R., Dominguez, C. y Losada, R. (2003). Regresión no paramétrica funcional aplicada a la clasificación de maderas mediante espectros termogravimétricos. *Actas IX Conferencia española de Biometría*.

- [64] Naya, S., Cao, R., Artiaga, R., López de Ullibarri, I. y García, A. (2002). Suavizado automático de curvas procedentes de Análisis Térmico con errores autocorrelacionados. *Actas del XV Congreso nacional de Ingeniería Mecánica*.
- [65] Naya, S., Cao, R. y Artiaga, R. (2001). Aplicación del modelo logístico generalizado para el ajuste por regresión local a curvas TGA. *Actas V Congreso Galego de Estatística e Investigación Operativa*, 141-144.
- [66] Naya, S., Cao, R. y Artiaga, R. (2003a). Local polynomial estimation of TGA derivatives using logistic regression for pilot bandwidth selection. *Aceptado para su publicación en Thermochimica Acta*.
- [67] Naya, S., Cao, R. y Artiaga, R. (2003b). Functional non-parametric model for materials discrimination by thermal analysis. *Enviado y aceptado en el ICCE10*.
- [68] Ozawa T. (1965). A New Method of Analyzing Thermogravimetric Data. *Bulletin of the Chemical Society of Japan*, **38**, 1881-1886.
- [69] Pagès, P, Carrasco, F., Saurina J. y Colom X. (1996). FTIR and DSC study of HDPE structural changes and mechanical properties variation when exposed to weathering aging during Canadian winter. *Journal of Applied Polymer Science*. **60**, 153-159.
- [70] Parzen, E. (1962). On estimation of a probability density and mode. *Annals of Mathematical Statistics*, **35**, 1065-1076.
- [71] Peña D. (1987). *Estadística modelos y métodos. Modelos lineales y series temporales*. Alianza editorial. Madrid, 187-497.
- [72] Popescu, C. (1984). Variation of the maximum rate of conversión and temperatura with heating rate in non-isothermal kinetics. Part II. *Thermochimica Acta*, **82**, 387-389.
- [73] Priestley, M.B. y Chao, M.T. (1972). Nonparametric function fitting. *Journal of the Royal Statistical Society, Series B*, **34**, 384-392.
- [74] Quintela, A. (1992). *Cálculo del parámetro de suavización en la estimación no paramétrica de curvas con datos dependientes*. Tesis Doctoral. Universidad de Santiago de Compostela.
- [75] Ramsay, J. y Silverman, B. (1997). *Functional Data Analysis*. Springer-Verlag. New York.
- [76] Ramsay, J. y Silverman, B. (2002). *Applied Functional Data Analysis. Methods and Case Studies*. Springer-Verlag. New York.
- [77] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, **27**, 1346-1370.

- [78] Roussas, G.G. (1989). Consistent regression with fixed design points under dependence conditions. *Statistics & Probability Letters*, **8**, 41–50.
- [79] Roussas, G.G. (1990). Nonparametric regression estimation under mixing conditions. *Stochastic Processes and their Applications*, **36**, 107–116.
- [80] Roussas, G., Tran, L. y Ionannides, D. (1992). Fixed design regression for time series: asymptotic normality. *Journal of Multivariate Analysis*, **40**, 262–291.
- [81] Rudemo, M. (1982). Empirical choice of histograms and kernel density estimators. *Scandinavian Journal of Statistics*, **9**, 65–78.
- [82] Ruppert, D. y Wand, P. (1994). Multivariate locally weighted least squares regression. *The Annals of Statistics*, **22**, 1346–1370.
- [83] Ruppert, D., Sheather, S.J. y Wand, M.P. (1995). An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Association*, **90**, 1257–1270.
- [84] Savitzky, A. y Golay, M. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, **36**, 1627–1639.
- [85] Seifert, B. y Gasser, T. (1996). Finite sample variance of local polynomials: analysis and solutions. *Journal of the American Statistical Association*, **91**, 267–275.
- [86] Sesták J. y Berggren G. (1971). Study of the kinetics of the mechanism of solid-state reactions at increasing temperatures. *Thermochimica Acta*, **3**, 1–12.
- [87] Stone, C.J. (1977). Consistent nonparametric regression. *The Annals of Statistics*, **5**, 595–620.
- [88] Stone, C.J. (1980). Optimal rates of convergence for nonparametric estimators. *The Annals of Statistics*, **8**, 1348–1360.
- [89] Stone, C.J. (1982). Optimal global rates of convergence for nonparametric regression. *The Annals of Statistics*, **10**, 1040–1053.
- [90] Trillo, J. (2002). *Implementación de algoritmos de suavización con datos dependientes: aplicaciones termogravimétricas*. Proyecto de fin de carrera de Ingeniería Informática. A Coruña.
- [91] Turi, A. (1997). *Thermal characterization of polymeric materials*. Academic Press. New York.
- [92] Uriel, E. (1995). *Análisis de datos. Series temporales y análisis multivariante*. Editorial AC. Madrid.

- [93] Vilar Fernández, J.M. y González-Manteiga, W. (1995). Goodness-of-fit test for a general linear model under serial error correlation. *Qüestió*, **3**, 337-368.
- [94] Vilar Fernández, J.M. y González-Manteiga, W. (2000). Resampling for checking linear regression models via non-parametric regression estimation. *Computational Statistical and Data Analysis*, **2**, 211-231.
- [95] Vilar-Fernández, J.A. y Vilar-Fernández, J.M. (1998). Recursive estimation of regression functions by local polynomial fitting. *Annals of the Institute of Statistical Mathematics*, **50**, 4, 729-754.
- [96] Vilar-Fernández, J.M. y Vilar-Fernández, J.A. (2000). Recursive local polynomial regression under dependence conditions. *Test.*, **9**, 1, 209-233.
- [97] Vyazovkin S. (1996). A unified approach to kinetic processing of non-isothermal data. *International Journal of Chemical Kinetics*, **28**, 95-101.
- [98] Wand, M.P. y Jones, M. C. (1995). *Kernel Smoothing*. Chapman & Hall. New York.
- [99] Watson, G.S. (1964). Smooth regression analysis. *Sankhya, A*, **26**, 359-372.
- [100] Watanabe, S. (1985). *Pattern Recognition. Human and Mechanical*. Wiley. New York.
- [101] Werbos, P. J. (1991). *Links between ANN and statistical pattern recognition*. Elsevier. Cambridge.

