

## Influencia de la adición de nano y microsíllice en la estabilidad térmica de una resina epoxi. Aplicaciones del ANOVA funcional

Influence of Nano and Micro Silica Addition on Thermal Stability of an Epoxy Resin. Use of Functional ANOVA

JAVIER TARRÍO<sup>a</sup>, SALVADOR NAYA<sup>b</sup>

DEPARTAMENTO DE MATEMÁTICAS, ESCUELA POLITÉCNICA SUPERIOR, UNIVERSIDAD DE A  
CORUÑA, FERROL, ESPAÑA

---

### Resumen

El principal objeto de este trabajo ha sido el empleo de una nueva técnica que aúna el análisis funcional y el diseño de experimentos: ANOVA funcional para un factor tratamiento. Mediante esta herramienta se ha medido la influencia que tiene la adición de una mezcla de nano y micro partículas de humo de sílice, subproducto de la industria del silicio, en la degradación térmica de una resina epoxi. Para tal fin se ha realizado un diseño de experimentos con un factor tratamiento (cantidad de humo de sílice) a tres niveles diferentes. Los datos se han obtenido mediante el empleo del Análisis Termogravimétrico (TG), dando como resultado cinco curvas o trayectorias de degradación por nivel. El ANOVA funcional aprovecha toda la información de cada curva o dato funcional.

El empleo del ANOVA funcional a partir de los datos TG ha dado como resultado que la cantidad de humo de sílice influye significativamente en la estabilidad térmica del compuesto y de la resina epoxi involucrada en cada muestra. Estos hechos pueden ser un indicio de la interacción entre la fase orgánica e inorgánica del material.

**Palabras clave:** análisis de datos funcionales, ANOVA, diseño de experimentos.

### Abstract

The main purpose of this work has been the use of a new technique that combines the functional analysis and design of experiments: a one way functional ANOVA. Using this tool is intended to measure the influence of the addition of a mixture of nano and micro particles of silica fume, by-product

---

<sup>a</sup>Profesor titular. E-mail: jtarrío@udc.es

<sup>b</sup>Profesor titular. E-mail: salva@udc.es

of silicon industry, on the thermal degradation of epoxy resin. To achieve this, it has been performed a design of experiments with a treatment factor (amount of fumed silica) at three different levels. The data were obtained using thermogravimetric analysis (TG), resulting in five curves or degradation pathways by level. The functional ANOVA uses all information of each curve or functional data.

The use of functional ANOVA from TG data has resulted in the amount of silica fume significantly affect the thermal stability of the compound and the epoxy resin involved in each sample. These facts may be indicative of the interaction between the organic and the inorganic phase.

**Key words:** ANOVA, Experimental design, Functional data analysis.

## 1. Introducción

En los últimos años se ha producido un gran avance en técnicas estadísticas asociadas al incremento y mejora de las aplicaciones informáticas. Algunas de estas técnicas han surgido para dar respuesta al problema que supone el trabajar con grandes bases de datos, que en ocasiones corresponden a observaciones de una variable aleatoria tomadas a lo largo de un intervalo continuo. Un ejemplo de estas nuevas técnicas es lo que se conoce como estadística o análisis de datos funcionales (FDA, por sus siglas en inglés).

La estadística de datos funcionales aborda muchos de los problemas de la estadística clásica: descripción de variables, modelización, clasificación, inferencia, etc. Habría que mencionar la publicación del libro de Ramsay & Silverman (1997) como el pionero en la divulgación de este tipo de estudios, si bien muchas de las aplicaciones a problemas concretos en áreas tan dispares como la medicina o la ingeniería se presentan en una segunda obra de los mismos autores (Ramsay & Silverman 2005). Otros dos textos importantes para el acercamiento a este estudio son, por un lado, el libro de de Ferraty & Vieu (2006), obra referencia en el enfoque no paramétrico del FDA, y el más reciente y generalista de Ferraty & Romain (2010).

En este trabajo se propone la aplicación de algunas de estas nuevas técnicas del campo de la estadística de datos funcionales en el estudio de las propiedades de un nuevo material compuesto para la industria: las resinas poliméricas epoxi con refuerzo de nano y micropartículas de sílice. El material en cuestión es un nanocompuesto de matriz polimérica (resina epoxi) y carga inorgánica (partículas de humo de sílice).

Las resinas epoxi son un polímero termoestable ampliamente utilizado. Algunas de las ventajas que exhiben las resinas epoxi frente a otros tipos de resinas poliméricas son, entre otras, su baja contracción de curado, la práctica ausencia de volátiles durante el mismo, resistencia mecánica y durabilidad, adhesión, resistencia química y aislamiento eléctrico. Además, estas resinas son compatibles con un gran número de materiales, por lo que son ampliamente utilizadas como matriz de materiales compuestos. De hecho, aunque su uso está de alguna forma limitado debido a su alta rigidez, provocada por una densa estructura de entrecruzamien-

to, sus propiedades mecánicas pueden mejorarse mediante la adición de partículas inorgánicas (Harsch, Karger-Kocsis & Holst 2007). La forma, volumen, tamaño, características superficiales y dispersión de las partículas dentro de la matriz condicionan altamente las propiedades mecánicas del material compuesto resultante (Lee & Lichtenhan 1999, Mehta, Mirabella, Bafna & Rufener 2004, Hsiue, Chen & Liu 2000, Shao-Yun, Xi-Qiao, Bernd & Yiu-Wing 2008). Particularmente el uso de nano-partículas provoca un mayor aumento del módulo de Young que el uso de micropartículas, incluso con una menor cantidad de las mismas (Zhang, Zhang, Friedrich & Eger 2006). Los nanocompuestos de matriz orgánica y carga inorgánica han mostrado ser capaces de proporcionar aumentos simultáneos en propiedades tales como la estabilidad térmica, el retardo a la llama, la temperatura de transición vítrea y la estabilidad dimensional, además de la disminución de la constante dieléctrica (Zhang et al. 2006, Liu, Wei, Hsu & Ho 1999, Preghenella, Pegoretti & Migliaresi 2005, Yousefi, Lafleur & Gauvin 1997). Debido a la mejora de sus propiedades mecánicas, el uso de materiales compuestos de matriz epoxídica es una de las opciones más utilizadas en la construcción de materiales ligeros para aplicaciones estructurales en la industria aeronáutica (aviones, transbordadores, satélites) y de los vehículos de transporte en general, como por ejemplo en automóviles y embarcaciones de altas prestaciones (Petrie 2006). Particularizando en este estudio, se ha buscado obtener mejoras en las propiedades de una determinada matriz epoxídica, siendo el primer paso para obtener nuevos materiales compuestos con aplicación, tanto en las industrias anteriormente mencionadas como en el sector de la construcción, fabricación de adhesivos o mismo en la industria eléctrica y electrónica, además de cumplir con una labor necesaria como es el reciclaje de un importante desecho industrial como es el humo de sílice.

En este trabajo, los compuestos de humo de sílice-resina epoxi fueron preparados y caracterizados mediante el empleo de técnicas termogravimétricas (TG), usuales en la evaluación de la estabilidad térmica de un material (Tarrio-Saavedra, López-Beceiro, Naya & Artiaga 2008). Para ello se ha escogido una resina epoxi no convencional, basada en el trimetilolpropano, especialmente indicada para la manufactura de materiales compuestos. Por otro lado, el humo de sílice empleado es un subproducto de la industria del silicio y ferrosilicio. Se forma en lo alto de los hornos de inducción de arco sumergido, siendo su forma de obtención diferente a los procesos convencionales para  $\text{SiO}_2$  sintético (Mohammad & Simon 2006). Por las características especiales del método de producción, el humo de sílice empleado tiene una pureza variable, dependiente de las condiciones de operación en los hornos. En todo caso, la proporción en peso de sílice nunca resulta menor del 95%. También es variable el tamaño de partícula, de hecho el humo de sílice consta de nano y micropartículas, teniendo en cuenta el criterio de Schadler (2003) (diámetro menor de 100 nm implica que es una nanopartícula). Esta especial distribución de tamaños hace esperar una posible combinación de efectos micro y nano.

Parte importante de este estudio es la determinación del grado de interacción del humo de sílice con la resina epoxi, relacionado con la interfase orgánica-inorgánica. Esta se define como aquella región que comienza en el punto de la carga que difiere del resto de la carga y termina en el punto de la matriz en el que

sus propiedades son las mismas que en el resto de la matriz (Schadler 2003). La existencia de esta interfase influye en propiedades tales como la estabilidad térmica o la temperatura de transición vítrea. De hecho, la variación de dichas propiedades puede tomarse como un índice de su existencia (Tarrío-Saavedra et al. 2008).

El resto del artículo se ha estructurado del siguiente modo. En la sección siguiente se presenta el modelo ANOVA funcional, seguidamente, en una sección dedicada a la parte experimental, se analizan los datos, curvas TG obtenidas en el estudio térmico de los nanomateriales, en otra sección se realiza un ajuste de los mismos en una base adecuada, con el fin de suavizarlos. A continuación se aplica el modelo ANOVA funcional a estas curvas termogravimétricas. Finalmente en una última sección se presentan los resultados, las conclusiones y se proponen nuevas líneas de investigación.

## 2. El modelo ANOVA funcional

Los datos funcionales, llamados también datos longitudinales, aparecen asociados a procesos monitorizados continuamente en el tiempo. Es decir, cuando se mide una variable en un conjunto discreto y finito de puntos ordenados; considerando que esa variable sigue una relación funcional continua. En este contexto, una variable aleatoria  $X$  se dice una variable funcional si toma valores en un espacio funcional  $F$  (normado o seminormado completo). Un caso particular se produce cuando la variable funcional  $X$  pertenece a un espacio de Hilbert, como es el caso de las funciones continuas en un intervalo (Ramsay & Silverman 1997). Este es el caso de las curvas de Análisis Termogravimétrico, conocidas como curvas TGA o TG, que serán las utilizadas en este trabajo, y que cumplen la propiedad de pertenecer al espacio de funciones medibles en un intervalo real, es decir,  $X \in L_2([0, T])$  (Naya, Cao & Artiaga 2003).

Como la variable de interés, masa de la muestra, se ha evaluado cada segundo, los datos pueden ser considerados funcionales. Dada la naturaleza funcional de los datos se propone la construcción de un ANOVA funcional (Cuevas, Febrero & Fraiman 2004). Una de las ventajas que presenta frente a un ANOVA clásico consiste en el aprovechamiento de toda la información de la curva en lugar de recurrir a indicadores puntuales sobre la misma.

Siguiendo la nomenclatura de Cuevas et al. (2004), cada dato funcional se puede representar como  $X_{ij}(t)$  con  $t \in [a, b]$ , extraídos de  $k$  procesos  $L_2$ ,  $X_i(t)$ , siendo  $i$  el subíndice que indica el nivel del factor y  $j$  el número de réplica ( $j = 1, 2, \dots, n_i$  e  $i = 1, 2, \dots, k$ ). Como, en este caso, la temperatura es directamente proporcional al tiempo (10 °C/min), puede considerarse a “ $t$ ” como los valores de la propia temperatura.

La media correspondiente a cada nivel o muestra independiente tiene la forma  $E(X_i(t)) = m_i(t)$ , mientras la covarianza entre dos instantes determinados dentro de una curva, en el caso más restrictivo de existencia de heterocedasticidad, se

identifica como  $K_i(s, t)$ :

$$K_i(s, t) = \sum_{j=1}^{n_i} \frac{(X_{ij}(s) - \bar{X}_{i\cdot}(s)) (X_{ij}(t) - \bar{X}_{i\cdot}(t))}{n_i - 1} \quad (1)$$

Por tanto, en este contexto funcional, el objetivo del contraste asociado al ANOVA funcional es probar que:

$$H_0 : m_1 = m_2 = \dots = m_k \quad (2)$$

El estadístico implementado por Cuevas et al. (2004) para contratar esta hipótesis es el siguiente:

$$V_n = \sum_{i < j} n_i \|\bar{X}_{i\cdot} - \bar{X}_{j\cdot}\|^2 \quad (3)$$

La utilización del estadístico (3) evita el requerimiento de la hipótesis de homocedasticidad en un ANOVA usual, siempre que se cumplan las propiedades siguientes:

- a)  $n_i, n \rightarrow \infty$  suponiendo  $p_i > 0$  para  $i = 1, \dots, k$ ,  $p_i = \frac{n_i}{n}$ .
- b) las observaciones  $X_{ij}(t)$  con  $j = 1, \dots, n_i$ , se corresponden con  $k$  muestras independientes de tamaño  $n_i$  procedentes de  $k$  procesos  $L_2$  con media cero y covarianza  $K_i(s, t)$ , se puede afirmar que la distribución de  $V_n$  bajo la hipótesis  $H_0$  coincide con la del estadístico:

$$V = \sum_{i < j} \|C_i(t) - C_{ij} Z_j(t)\|^2 \quad (4)$$

Donde  $C_{ij} = (p_i/p_j)^{1/2}$  y  $Z_1(t), \dots, Z_j(t)$  son procesos gaussianos independientes de media cero y covarianza  $K_i(s, t)$ .

El cálculo de la distribución del estadístico (4) bajo la hipótesis nula puede hacerse mediante la aplicación de un bootstrap paramétrico y del método de Monte Carlo, obteniéndose el valor del cuantil  $V_\alpha$ , que permite resolver el contraste, así si  $V_n > V_\alpha$  se rechazará la hipótesis  $H_0$  siendo el test significativo, es decir, las medias de los distintos niveles del factor serían diferentes.

### 3. Experimentación y obtención de datos

#### 3.1. Materiales

Los materiales empleados en este estudio están formados por una resina epoxi, que actúa como matriz del material compuesto. Este polímero está basado en el diglicidil eter de trimetilolpropano, Triepox GA, manufacturado por la empresa GAIRESA, SA. El agente de curado utilizado es una amina aromática, 1,3-bencenodimetanamina, suministrada por Aldrich. Triepox GA es una resina altamente tixotrópica que posee, además de una baja densidad, la capacidad de curar a temperatura ambiente en ausencia de plastificantes o aditivos.

El humo de sílice ha sido proporcionado por Ferroatlántica I+D, España. Se obtiene como un subproducto en la producción del silicio en hornos de arco eléctrico sumergido. Este proceso implica la reducción de cuarzo de alta pureza, a temperaturas superiores a 1800°C. El humo de sílice se forma cuando el gas SiO, resultante de la reducción de cuarzo, se mezcla con el oxígeno en la parte superior del horno, dando lugar a la producción de partículas esféricas de sílice. Es un polvo fino de color variable, desde casi negro a ligeramente blanquecino, en función de su contenido en carbono. Su tamaño medio de partícula es de 0.15 micras y el 41.9% de las partículas tienen un diámetro inferior a 0.2 micras. El área superficial se sitúa en torno a  $20 \text{ m}^2\text{g}^{-1}$ . En cuanto a la composición química, el humo de sílice consta de SiO<sub>2</sub> amorfo de pureza variable.

### 3.2. Obtención de las muestras

Las muestras han sido preparadas para contenidos del 0, 10, y 20% en peso de humo de sílice. Tanto la resina como el endurecedor se han mezclado en una proporción estequiométrica. Para la obtención de los compuestos correspondientes al 10 y al 20% en humo de sílice, las mezclas de sílice y resina se han agitado durante 15 minutos con el fin de obtener una distribución lo más uniforme posible. A continuación, se ha aplicado un tratamiento de ultrasonidos durante 5 minutos y a temperatura ambiente para dispersar a los aglomerados de sílice. La pasta así obtenida se ha vertido en un molde de silicona con cavidades de dimensiones  $0.8 \times 4 \times 30$  mm. En este emplazamiento, las muestras han sido curadas a temperatura ambiente durante 24 horas, aplicando posteriormente un postcurado a 90°C durante 2 horas.

Los experimentos TGA se han llevado a cabo por medio de una termo-balanza STA 1500, de Rheometric Scientific. Todas las muestras se han sometido a una rampa de calentamiento de 10°C/min, en un rango de temperaturas comprendido entre 20 y 600°C. Todos los experimentos se han realizado bajo atmósfera oxidante, manteniendo un flujo de aire de  $50 \text{ mL}\cdot\text{min}^{-1}$ .

## 4. Diseño de experimentos

Con el objeto de evaluar el efecto de la adición de humo de sílice en la degradación térmica del material resultante, se ha realizado un diseño de experimentos con datos funcionales. Se ha elegido un diseño completamente aleatorizado de una vía para probar las posibles diferencias en la respuesta al variar el nivel del factor. Se ha tomado como factor de estudio o variable explicativa el contenido en sílice de cada muestra, con tres niveles (0%, 10% y 20% en peso de humo de sílice). Se han efectuado entre 5 y 7 experimentos o réplicas por cada nivel con el fin de caracterizar la influencia de la cantidad de humo de sílice en la estabilidad térmica de los materiales compuestos resultantes: 7 correspondientes a la resina epoxi sin carga, otros 7 para un 10% en peso de humo de sílice y 5 para un 20% (ver figura 1). El número de réplicas elegidas se ha marcado principalmente con el fin de llegar a un compromiso aceptable entre la representación adecuada de la variabilidad dentro

de cada nivel y el tiempo de experimentación total requerido. Además, de entre los 19 experimentos efectuados en total, 4 han sido realizados variando ligeramente las condiciones experimentales. Mediante un posterior análisis de profundidad de datos funcionales se pretende identificar las curvas TG obtenidas mediante condiciones experimentales ligeramente diferentes. En el caso de poder identificarlas correctamente -probándose la utilidad del cálculo de profundidad de datos- se procede a eliminarlas para no introducir mayor variabilidad al modelo, con lo que quedarían 5 curvas por nivel (correspondiéndose con un diseño balanceado).

Para poder comparar convenientemente todos los datos, se ha expresado la masa de cada muestra en tanto por ciento de la cantidad inicial. Así, todas las curvas comienzan con un valor del 100 % en el eje de ordenadas.

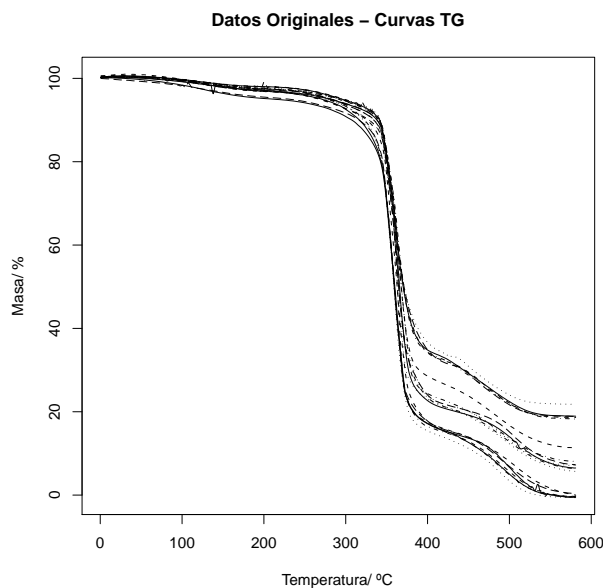


FIGURA 1: Datos experimentales: curvas TG de partida.

Cada uno de los experimentos se corresponde con un dato funcional donde se representa la masa de la muestra en función de la temperatura a la que está sometida. Como ya se ha indicado, cada muestra es calentada a  $10^{\circ}\text{C}/\text{min}$  en un rango de temperaturas que va de 20 a  $600^{\circ}\text{C}$ . Se observa que al final de cada ensayo, a  $600^{\circ}\text{C}$ , se degrada por completo la fase orgánica (resina epoxi), quedando únicamente la masa añadida de humo de sílice, mucho más resistente térmicamente. Es importante señalar que la masa de la muestra se representa en %, es decir, se asigna a la masa inicial el 100 % representándose todos los valores siguientes con respecto a esta. Cada curva consta de un número variable de puntos entorno a 3480, uno por segundo, dependiendo de la temperatura ambiente a la que se encuentre la máquina de ensayo.

La variable respuesta o dependiente es una variable funcional donde cada dato es una curva que representa la masa del material en función de la temperatura a la que está sometido. Para su obtención se ha programado un aumento constante en la temperatura de  $10^{\circ}\text{C}/\text{min}$ . Todas las curvas son decrecientes, pues el material se degrada, es decir, pierde masa, según aumenta la temperatura a la que está sometido. En sí, desde el punto de vista de un estudio de fiabilidad de los materiales, cada curva representa el camino particular de degradación de cada muestra ensayada. La finalidad de esta experimentación es dar respuesta a las siguientes preguntas: ¿son las medias de las curvas TG (caminos de degradación de cada muestra) diferentes para distintos niveles del factor cantidad de sílice? Si la respuesta se responde afirmativamente, las diferencias en cuanto a estabilidad térmica observadas en la figura 4 tendrían una significación estadística.

#### 4.1. Suavización de los datos mediante bases splines

La representación de un dato funcional en una base ortonormal proporciona ventajas desde el punto de vista teórico y práctico sirviendo de puente entre la inevitable discretización del dato funcional para su tratamiento y su verdadera forma funcional. En este caso se representa cada dato funcional discretizado según una base finita, de modo que se obtiene una forma explícita para la función (Ramsay & Silverman 1997, Ferraty & Vieu 2006, Cuevas, Febrero & Fraiman 2006):

$$X(t) \approx \sum_{k=1}^K c_k \phi_k(t) \quad (5)$$

Donde la base  $\{\phi_k\}$ , con  $k \in \mathbb{N}$ , es un conjunto de funciones conocidas e independientes tales que cualquier función puede ser aproximada, mediante la combinación lineal de  $K$  de ellas (elementos de la base) y siendo  $c_k$  constantes.

Eligiendo una base apropiada se obtiene una forma más suave del dato funcional. Además, presenta la ventaja de poder ser derivable. De entre todos los tipos de base disponibles, se ha elegido ensayar una bases de b-splines y b-splines penalizados. Esto es debido a las características de los datos de partida. Ambas bases proporcionan la estructura flexible requerida. Las bases de Fourier, sin embargo, no serían tan apropiadas, ya que los datos no presentan una aparente trayectoria periódica. Además, dada la ya de por sí suavidad de los datos de partida, no parece tampoco adecuado ensayar bases wavelets. El ajuste según una base b-splines tiene la expresión:

$$S(t) \approx \sum_{k=1}^{m+L-1} c_k B_k(t, \tau) \quad (6)$$

Donde  $m$  corresponde al orden del polinomio,  $\tau$  son los valores de los  $L$  puntos de corte de los subintervalos (nodos) y  $k$  es el número de elementos de la base. Mientras que la suma de residuos del ajuste según una base b-spline penalizada responden a la expresión:



$$PRSS = \|Y - X\beta\|^2 + \lambda \int (f''(x))^2 dx \quad (7)$$

Donde  $Y$  son los datos a ajustar,  $X\beta$  es el ajuste según la base b-spline,  $f''(x)$  es la derivada segunda de dicho ajuste y  $\lambda$  es el parámetro de suavización (que penaliza la derivada segunda, es decir, restringe la varianza interna que pueda tener el ajuste). Más información al respecto se puede encontrar en los trabajos de Ferraty & Vieu (2006), además de los ya mencionados de Ramsay & Silverman(1997, 2005).

Para elegir el número óptimo de elementos de la base se ha implementado el criterio de Validación Cruzada Generalizada (GCV, por sus siglas en inglés). Se ha ensayado un número de elementos de la base que oscila entre 4 y 480. Se ha elegido el número de elementos que se corresponde con el GCV mínimo en cada dato funcional. Acto seguido se ha elegido el número de elementos de la base mínimos de entre todos los mínimos 19 GCV (uno por dato funcional). De esta forma se disminuye el riesgo de sobresuavizado. El resultado de la minimización de la expresión anterior puede verse en la figura 2 y en la tabla 1.

TABLA 1: Número de elementos de las bases óptimos según el criterio GCV.

Profundidad	GCV óptimo	N° de elementos	GCV (80)
b-splines	$2.0e - 07$	375	$3.1e - 03$
b-splines penal ( $\lambda = 0.5$ )	$3.1e - 04$	182	$6.4e - 03$

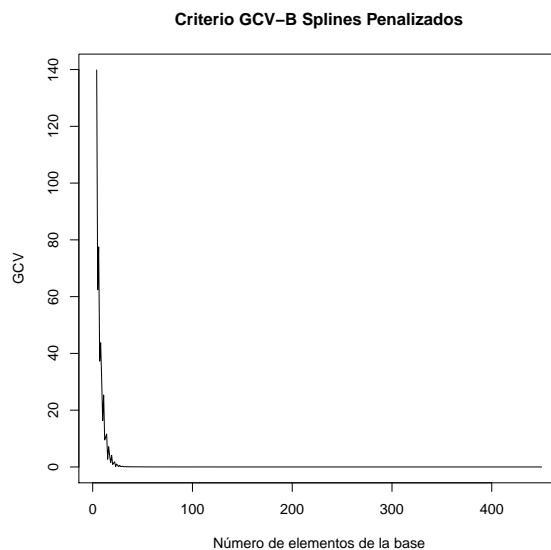


FIGURA 2: Para un dato funcional determinado, GCV frente al número de bases para una base b-spline penalizada.

En cuanto a los b-splines, el GCV óptimo resulta demasiado pequeño y el número de bases demasiado grande, con lo cual se corre el riesgo de interpolar los datos. En cambio se obtiene un GCV aceptable para una base de 80 elementos. Se consigue un ajuste más suave sin apartarse de la trayectoria de los datos originales.

Empleando b-splines penalizados se obtiene un valor para GCV mayor, con un número de elementos en la base también menor. Pero se observa que para 80 elementos se obtiene un GCV aceptable. La prueba está en la figura 2: para un número de elementos de la base de 80 se consigue seguir fielmente la trayectoria de los datos, suavizando además las discontinuidades experimentales. Asimismo, en la figura 2 se observa que el GCV cae abruptamente hasta estancarse para un valor determinado del número de elementos de la base. Este hecho apoya la decisión de optar por una base más pequeña, correspondiente a un número de elementos de la base más cercano al comienzo de ese estancamiento.

Si se disminuyera todavía más el número de elementos de la base, se incurriría en un error inaceptable. Los ajustes se alejan de los datos allí donde es crítico que los reproduzcan fielmente: en los cambios de pendiente de los escalones.

Se ha elegido una base b-spline penalizada de orden cuatro, con 80 elementos, como base de ajuste. El ajuste es muy bueno y, muy ligeramente, más suave que en una base de b-splines. Además, el número reducido de datos no origina un coste computacional muy grande.

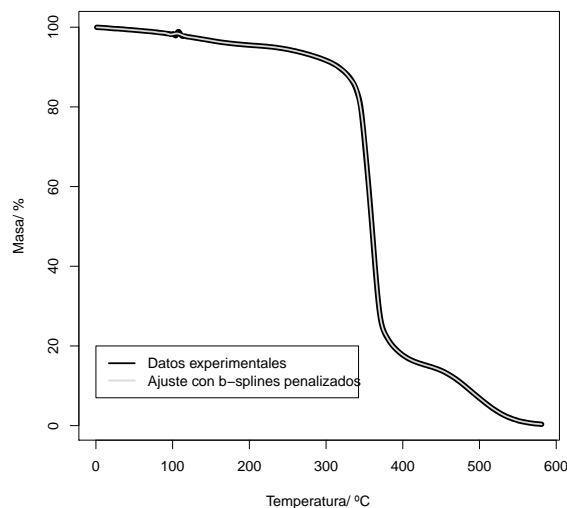


FIGURA 3: Curva TGA de resina epoxi y ajuste con una base b-spline penalizada con 80 elementos.

Ya de por sí, los datos presentan una forma muy suave, sin grandes discontinuidades. Se observa que los dos ajustes de la figura 3 reproducen fielmente la

trayectoria de los datos reales, obviando las pocas discontinuidades debidas al error experimental.

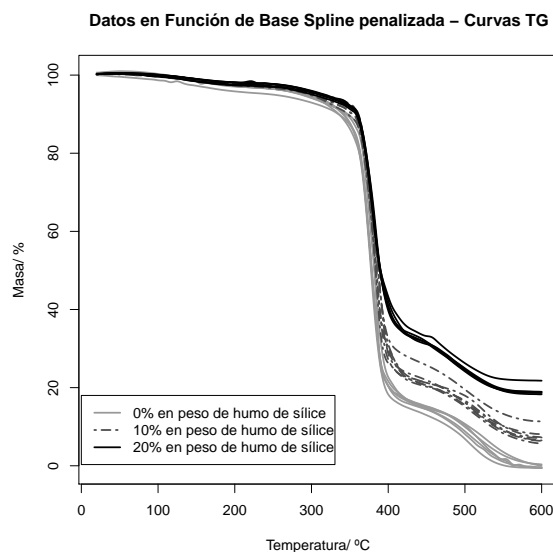


FIGURA 4: Dato experimental (curva TGA de resina epoxi) y ajuste con una base b-spline penalizada con 80 elementos.

## 4.2. Estudio de la profundidad de los datos

El concepto de profundidad, aplicado al análisis de datos funcionales, indica qué tan central es una curva o dato funcional particular en relación con un conjunto de curvas pertenecientes a una población. Siguiendo este criterio, pueden ordenarse los diferentes datos funcionales pertenecientes a una muestra de una población dada: se identifican como más profundos aquellas curvas más cercanas al centro (dato más profundo definido como mediana), mientras que se identifica como datos atípicos aquellas curvas que se generan por un proceso estocástico diferente al resto (Cuevas et al. 2006). En el presente estudio se ha calculado la profundidad asociada a cada dato funcional por tres métodos referenciados cuando se tratan datos de dimensión infinita (FDA Cuevas, Febrero & Fraiman (2007)). El resultado obtenido nos da una idea completa de la profundidad de cada curva TG con respecto al total.

El primer método utilizado es la profundidad de Fraiman & Muniz (2001), que está basado en la idea de profundidad funcional integrada:

$$I_i = \int_0^1 D_i(t) dt \quad (8)$$

donde  $D_i(t) = 1 - |\frac{1}{2} - F_{n,t}(x_i(t))|$  es la profundidad univariante correspondiente a la curva  $x_i(t)$  con  $t \in [0, 1]$  y  $F_{n,t}$  la distribución empírica de la muestra

$x_1(t), \dots, x_n(t)$ . Los datos funcionales  $x_i(t)$  se ordenan según el valor  $I_i$  correspondiente.

El método de la profundidad moda es una alternativa al uso de la profundidad de Fraiman y Muniz en el contexto del FDA. Según este procedimiento, se define la población de profundidades moda relativas a un dato  $z$  mediante la expresión:

$$f_h(z) = E(K_h(\|z - X\|)) \quad (9)$$

para el caso de dimensión infinita y

$$\hat{f}_h(z) = \frac{1}{n} \sum_{i=1}^n K_h(\|z - X_i\|) \quad (10)$$

para el caso de una muestra compuesta de  $n$  curvas donde  $X$  es el elemento aleatorio que describe la población,  $\|\cdot\|$  es, por lo general, la norma  $L_2$  y  $K_h$  un kernel reescalado que presenta la forma  $K_h = \frac{1}{h} K(\frac{\cdot}{h})$ , siendo  $K(t)$  una función kernel (el kernel gaussiano es el más utilizado) y  $h$  un parámetro de ajuste. La profundidad asignada a cada dato funcional  $z$  coincide con la moda de  $X$ , que es el producto de la maximización, en  $z$ , de la función  $f_h(z)$ .

Por último, a continuación se muestra el método de proyecciones aleatorias que consiste en, a partir de una muestra  $X_1, X_2, \dots, X_n$ , tomar una dirección aleatoria  $a$  independiente de  $X_i$  y proyectar los datos a lo largo de esa dirección. De este modo, la profundidad de cada dato funcional  $X_i$  es definida como la profundidad univariante correspondiente a su proyección unidimensional. En el caso particular de datos funcionales, suponiendo que  $X_i$  pertenece a un espacio de Hilbert, las proyecciones de cada dato se obtienen mediante el producto interior  $\langle a, S \rangle = \int_0^1 a(t)X(t) dt$ . Para más información, consultar Cuevas et al. (2007).

Cada nivel del factor aparece, en la figura 4, relativamente bien diferenciado de los restantes. Las diferencias se observan sobre todo a altas temperaturas. La resina se va degradando paulatinamente de forma que al final sólo va quedando el humo de sílice añadido. La resina sin humo de sílice se degrada completamente (en tono gris claro y trazo continuo) mientras que todavía queda un entorno a un 10 % y un 20 % de la masa para los grupos que tienen un 10 y un 20 % en carga de humo de sílice.

A simple vista se pueden observar tres escalones en las curvas. El primero, no siempre perceptible, se corresponde a la pérdida de sustancias volátiles y humedad (se localiza a temperaturas alrededor de 100 °C).

Es singularmente importante el segundo escalón de las gráficas, ya que indica en realidad el primer proceso de degradación. Es aquí donde la resina epoxi empieza a perder masa, y lo hace de forma abrupta. Se observa que cuando se pasa de 0 % a 10 % en humo de sílice, la masa remanente al comienzo de este escalón es ligeramente mayor a una temperatura dada (por ejemplo, en torno a 320 °C). Esto quiere decir que se ha conseguido aumentar aparentemente la estabilidad térmica del material compuesto resultante. En cambio, cuando se pasa de 10 % a 20 % de humo de sílice, este aumento es mucho más ligero.

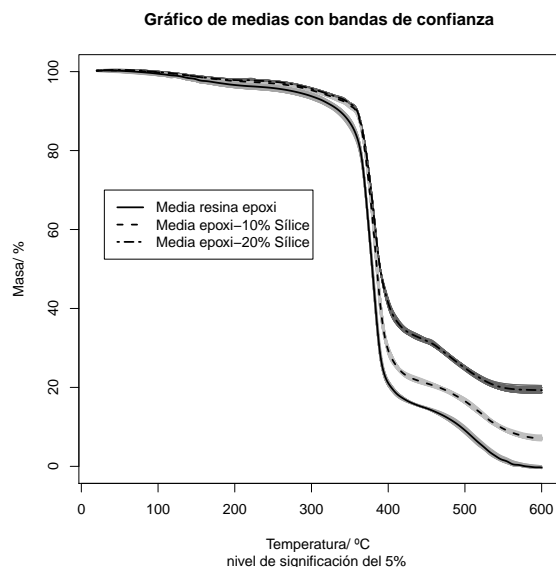


FIGURA 5: Datos funcionales suavizados con una base b-spline penalizada de 80 elementos.

El tercer escalón se corresponde con la desaparición del residuo carbonoso resultante de la reacción anterior. Se observa que aquí las diferencias son muy evidentes, debido principalmente a que el humo de sílice no se degrada a esas temperaturas: queda patente, al comparar las muestras, las diferencias que existen de contenido en sílice.

Para mayor sencillez de cálculo del estadístico de contraste, se pretende realizar un diseño balanceado. Por consiguiente, habría que desestimar dos datos funcionales correspondientes al nivel del 0% y otros dos del 10%. De hecho, existen algunas diferencias de experimentación en algunos datos, con lo cual sería interesante detectarlos como atípicos o estudiar simplemente su profundidad.

## 5. Resultados

### 5.1. Cálculo de medidas de profundidad

Se han calculado las profundidades de Fraiman & Muniz (2001), llamada profundidad mediana, la profundidad moda (el punto más profundo sería el equivalente a la moda de los datos) (Cuevas et al. 2006) y la profundidad de proyecciones aleatorias (Cuevas et al. 2007) para los grupos de 0% y de 10% por separado. Es en estos grupos donde interesa reducir el número de datos. Los resultados aparecen en las tablas 2 y 3. En estas tablas se presentan la profundidad mediana (F-M), la profundidad moda y la de proyecciones aleatorias (P-A) para las 14 muestras analizadas (M1, M2, ..., M14).

Todos los criterios coinciden en indicar los datos menos profundos tanto para el 0 % como para el grupo de 10 % de humo de sílice: para el 0 % es la muestra 7 la menos profunda por unanimidad, mientras que para el 10 % es la muestra 12 con diferencia (obsérvense los números en negrita dentro de las tablas).

TABLA 2: Profundidades para las 7 muestras de resina epoxi sin humo de sílice según los 3 criterios.

Prof.	M1	M2	M3	M4	M5	M6	M7
F-M	0.766	0.757	0.746	0.766	<b>0.718</b>	0.799	<b>0.661</b>
Moda	0.645	<b>0.345</b>	0.480	1.000	0.725	0.910	<b>0.000</b>
P-A	0.258	0.234	<b>0.200</b>	0.287	0.247	0.287	<b>0.202</b>

TABLA 3: Profundidades para las 7 muestras de 10 % en peso de humo de sílice según los 3 criterios.

Prof.	M8	M9	M10	M11	M12	M13	M14
F-M	<b>0.751</b>	0.721	0.814	0.755	<b>0.577</b>	0.764	0.833
Moda	0.705	0.751	0.745	0.942	<b>0.000</b>	<b>0.565</b>	1.000
P-A	0.232	0.271	0.260	0.280	<b>0.190</b>	<b>0.198</b>	0.284

Es necesario todavía eliminar otro punto dentro de estos dos grupos pero la unanimidad en los tres métodos ya no existe. Así, en el caso de 0 % se elige como posible atípico una solución de compromiso: la muestra 3, que coincide con el segundo dato menos profundo para las proyecciones aleatorias y con el tercero menos profundo según FM y la profundidad moda. Para el caso de 10 %, se ha elegido el segundo dato menos profundo según la profundidad de proyecciones aleatorias y según la profundidad moda, la muestra 13.

Es de destacar que en el caso de 0 % en humo de sílice, se ha identificado como menos profundo aquel cuyas condiciones experimentales eran más especiales: se ensayó muy poca masa de material con respecto a las demás muestras y, además, partida en dos trozos. La identificación de la muestra 3 puede deberse a que, en este caso, no se efectuó una calibración adicional en peso.

En cuanto a la de 10 %, se ha señalado como dato menos profundo aquel cuya masa en humo de sílice se aparta más de 10 % especificado. La dispersión de la carga sobre la matriz puede ser no del todo homogénea debido a la formación de aglomerados de partículas. Por lo tanto, dependiendo del trozo de muestra elegido puede haber más o menos cantidad de sílice. En definitiva, los puntos menos profundos han resultado ser los correspondientes a las condiciones más extremas de ensayo.

Una vez estudiada la profundidad de los datos, cabría la posibilidad que aquellos menos profundos pudieran clasificarse como atípicos. Para localizarlos se utiliza el siguiente procedimiento (Febrero, Galeano & González-Manteiga 2008):

- Se obtienen las medidas de profundidad por uno de los tres métodos.

- Si las profundidades  $D(x_{ik}) \leq C$  con  $k = 1, \dots, K$ , siendo  $C$  una profundidad prefijada, los datos funcionales  $x_{ik}$  se toman como datos atípicos.
- Se ha de elegir  $C$  de forma tal que el número de observaciones no atípicas clasificadas como atípicas sea aproximadamente igual al 1 %.

Para calcular  $C$  se realizan  $B$  remuestras bootstrap (suavizado) a partir de la muestra inicial (una vez descartadas las  $\alpha$  menos profundas):  $y_{ib}$ , donde  $i = 1, \dots, n$  y  $b = 1, \dots, B$ . Acto seguido se obtiene el cuantil  $C_b$  correspondiente al 1 % de la distribución de las profundidades  $D(y_{ib})$ . El valor  $C$  es la mediana de los valores de  $C_b$ .

Se ha realizado, por tanto, un bootstrap suavizado basado en el recorte (Cuevas et al. 2006) con  $B = 200$  y  $\alpha = 0.05$ . Se ha empleado la profundidad basada en Fraiman y Muniz dando como resultado la no existencia de atípicos. Hay que indicar que en este resultado pesa el hecho de disponer sólo de 5 datos funcionales por nivel.

Finalmente, se ha optado por prescindir de aquellos datos elegidos como menos profundos, así para el 0 % en humo de sílice prescindiremos de las muestra 3 y 7; mientras que para el 10 % en humo de sílice, se eliminarán las muestras 12 y 13. Con los 15 datos funcionales restantes, 5 por nivel, se calculan la media, mediana y varianza de los datos funcionales (ver figura 4, figura 5 y figura 6).

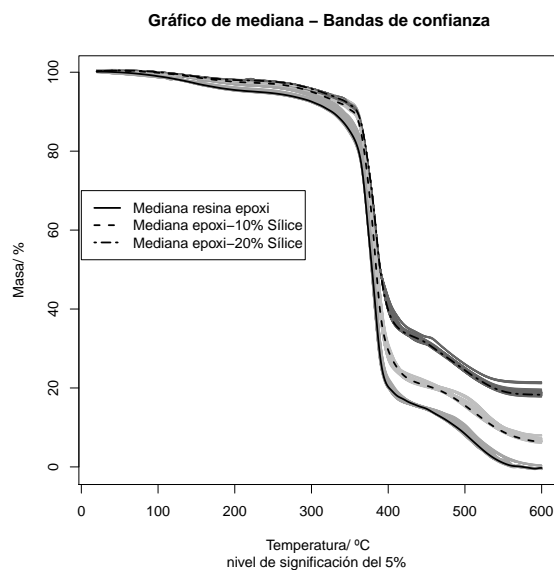


FIGURA 6: Medias para los distintos grupos con bandas de confianza bootstrap.

Mediante bootstrap simple (en el caso de la media) y bootstrap suavizado de parámetro  $h = 0.07$  (en los demás casos), se construyen bandas de confianza para cada uno de los estadísticos. Se ha empleado bootstrap suavizado en aquellos casos

donde el bootstrap simple no ofrece una idea clara de cómo puede ser el intervalo de confianza, siendo necesaria la adición de una componente aleatoria adicional.

Al realizar un bootstrap suavizado se está perturbando la muestra original sumándole un proceso gaussiano multiplicado por un factor de suavizado  $h$  (Cuevas et al. 2006):  $x_i^* = x_j(t) + hz(t)$  donde  $z(t)$  es un proceso gaussiano de media cero y matriz de covarianzas  $\Sigma$ . Así se combate en cierta forma la existencia de tan pocos datos para elaborar las bandas.

Se calculan la media y las medianas (Fraiman & Muniz 2001), tanto en la media como en la mediana se observa que las curvas correspondientes al 20 % en humo de sílice se sitúan por encima de las curvas correspondientes 10 %, que a su vez se sitúan por encima de las que no tienen humo de sílice. Esto es un indicador del aumento de la estabilidad térmica. Se observa que es mayor la diferencia entre los estadísticos de 0 % y de 10 % que entre 10 y 20 %, sobre todo en el escalón principal.

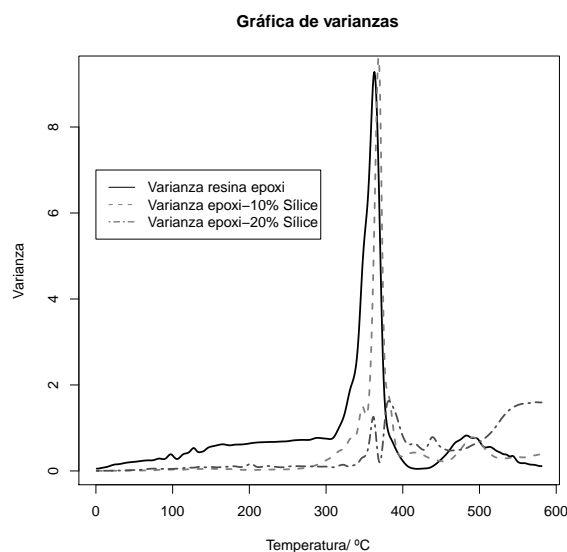


FIGURA 7: Varianza para los distintos grupos.

Se observa una mayor variabilidad general en los datos correspondientes a la resina epoxi sola. La variabilidad disminuye ligeramente para 10 % y, finalmente, es mucho menor para 20 % (esto se puede observar en media y mediana, además de en la varianza de la figura 6). Dicha tendencia puede deberse a la heterogeneidad de las propias muestras o al efecto aprendizaje del operador. De hecho, las primeras muestras que se realizaron corresponden a 0 % y las últimas a 20 % (aquellas que tienen menor variabilidad). Otra posible causa a tener en cuenta es el ensayo de muestras con diferentes contenidos de humedad, para el caso de 0 %.

La gráfica de las varianzas (figura 6) es en particular interesante porque resalta que allí donde difieren las sucesivas muestras es, principalmente, en los dos esca-



nes principales asociados a los dos primordiales procesos de degradación. Además, se observa que, contrariamente a lo que se había enunciado antes, la varianza para una temperatura cercana a los 600°C aumenta en la secuencia: resina epoxi→10% humo de sílice→20% humo de sílice. Esto se debe a que se forman aglomerados de partículas con lo que la distribución no es del todo uniforme. Así, para un 10% teórico, unas muestras tendrán un 8% en sílice, otras un 9% y otras un 11% (figura 5). La varianza es un indicador de la heterogeneidad de la muestra.

## 5.2. Aplicación del ANOVA funcional

Para construir el contraste, se han seguido los pasos marcados por Cuevas et al. (2004), teniendo en cuenta que la hipótesis nula a contrastar de igualdad de medias entre las curvas (7), donde  $m_i$  es la media de los datos funcionales dentro de cada uno de los tres niveles estudiados del factor “cantidad de humo de sílice”.

En primer lugar se construye el estadístico  $V_n$ , donde los tres niveles se corresponden con la adición del 0%, del 10% y del 20% en humo de sílice mientras que la respuesta está representada por la media de los datos funcionales para cada nivel. Con el fin de aproximar por Monte Carlo la distribución del estadístico  $V_n$ , se procede del modo siguiente:

- a) Se calcula la matriz de varianzas covarianzas de los datos muestrales. Si se supone que la estructura de covarianzas es diferente según el nivel del factor, nos encontramos en el caso heterocedástico; más restrictivo que el homocedástico. Suponiendo heterocedasticidad, se calcula una matriz de covarianzas diferente por nivel, siendo su estimador el siguiente:

$$\widehat{K}_i(s, t) = \frac{\sum_{j=1}^{n_i} (X_{ij}(s) - \overline{X}_i(s)) (X_{ij}(t) - \overline{X}_i(t))}{n_i - 1}$$

Siendo  $s$  y  $t$  dos instantes dados dentro de cada dato funcional.

- b) Bajo la hipótesis nula, el estadístico  $V$  sigue una distribución dependiente de  $k$  procesos gaussianos (Cuevas et al. 2004) difícil de conseguir. Pero dicha distribución se puede aproximar por Monte Carlo. Para ello se lleva a cabo el siguiente procedimiento bootstrap paramétrico:
  - b.1) partiendo de la muestra original, se calculan las matrices  $\widehat{K}_i(s_p, t_q)_{1 \leq p, q \leq n}$  y se generan  $B = 2000$  remuestras bootstrap por nivel, generadas según una distribución normal de media cero y matriz de covarianzas  $\widehat{K}_i(s_p, t_q)_{1 \leq p, q \leq n}$ .
  - b.2) se consiguen 2000 valores  $Z_{il}^* = (Z_{il}^*(t_1), \dots, Z_{il}^*(t_m))$  por nivel, con  $l = 1, \dots, 2000$  e  $i = 1, 2, 3$ . Estos valores aproximan las trayectorias continuas de  $Z_i(t)$  por versiones escalonadas evaluadas en una rejilla  $a \leq t_1, \dots, t_m \leq b$ .
  - b.3) se construyen 2000 réplicas de la forma  $\widehat{V}_l = \sum_{i < j} \left\| Z_{il}^*(t) - C_{ij} Z_{jl}^*(t_1) \right\|^2$ , que aproximan la distribución de  $V$  bajo  $H_0$ .

Mediante el proceso anterior se puede calcular  $V_\alpha$  tal que  $P(V > V_\alpha) = \alpha$  bajo la hipótesis nula. Si  $V_n > V_\alpha$ , el contraste resulta significativo.

El resultado de la implementación de este proceso al caso de las curvas TG sin reescalar ha proporcionado los siguientes valores  $V_n = 570315.1$  y  $V_\alpha = V_{0.95} = 1.402.292$ .

Por lo tanto, como  $V_n \gg V_\alpha$  (siendo el  $P$ -valor =  $5e - 6 < 0.05$ ), el test resulta significativo para un nivel de significación del 0.05, al menos una de las medias funcionales por factor es diferente a las demás. Desde el punto de vista ingenil podría afirmarse que la adición de humo de sílice provoca cambios en las medias funcionales de las curvas TG. Se podría concluir que aumenta significativamente la estabilidad térmica del material compuesto al aumentar la cantidad de sílice.

## 6. Conclusiones y líneas futuras

Como una primera conclusión de este estudio, podría decirse que la aplicación de técnicas de estadística funcional resulta determinante para el estudio de este tipo de datos. Además, el concepto de profundidad resulta muy adecuado para identificar ligeros cambios en las condiciones experimentales, posibilitando eliminar muestras anómalas.

Por otra parte el contraste ANOVA funcional empleado ha resultado significativo, lo cual avala el resultado del análisis realizado mediante el estudio de la profundidad de los datos. Es decir, el aumento en la estabilidad térmica del material compuesto, observado en el análisis descriptivo o, lo que en este caso es lo mismo, las diferencias provocadas en la respuesta funcional al aumentar la cantidad de humo de sílice, tienen significación estadística.

La variabilidad observada en el último tramo de las curvas, creciente conforme se aumenta la cantidad de humo de sílice, ha resultado ser un buen indicador del grado de dispersión de esta carga inorgánica dentro de la matriz polimérica.

Como línea de futura investigación sería interesante realizar un análisis similar con los datos de las curvas TGA realizando una sustracción matemática de la cantidad de humo de sílice añadida y el posterior reescalado con objeto de caracterizar el camino de degradación de la resina epoxi pura, dentro de cada muestra de material compuesto.

También resultaría de gran interés la realización de un estudio similar con las derivadas de estas curvas TGA, las denominadas DTG, ya que permitirá analizar si la velocidad de degradación térmica es diferente en función del humo de sílice añadido. Además, podrían plantearse estudios de interacción entre los factores estudiados.

## Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Educación y Ciencia MTM2008-00166. Los autores quieren expresar su especial agradecimiento a Manuel Febrero Bande por sus interesantes aportaciones para la aplicación de los programas en R utilizados (fda.usc), así como a Aldana González Montoro, por sus valiosos comentarios. Además, también agradecen a los árbitros que revisaron el trabajo sus interesantes sugerencias.

[Recibido: abril de 2010 — Aceptado: febrero de 2011]

## Referencias

- Cuevas, A., Febrero, M. & Fraiman, R. (2004), 'An anova test for functional data', *Computational Statistics and Data Analysis* **47**, 111–122.
- Cuevas, A., Febrero, M. & Fraiman, R. (2006), 'On the use of the bootstrap for estimating functions with functional data', *Computational Statistics and Data Analysis* **51**, 1063–1074.
- Cuevas, A., Febrero, M. & Fraiman, R. (2007), 'Robust estimation and classification for functional data via projection-based depth notions', *Computational Statistics* **22**(3), 481–496.
- Febrero, M., Galeano, P. & González-Manteiga, W. (2008), 'Outlier detection in functional data by depth measures, with application to identify abnormal NOx levels', *Environmetrics* **19**, 331–345.
- Ferraty, F. & Romain, Y. (2010), *The Oxford Handbook of Functional Data Analysis*, Oxford University Press, Oxford, England.
- Ferraty, F. & Vieu, P. (2006), *Nonparametric Functional Data Analysis: Theory and Practice*, Series in Statistics, Springer, Berlin, Alemania.
- Fraiman, R. & Muniz, G. (2001), 'Trimmed means for functional data', *Test* **10**(2), 419–440.
- Harsch, M., Karger-Kocsis, J. & Holst, M. (2007), 'Influence of fillers and additives on the cure kinetics of an epoxy/anhydride resin', *European Polymer Journal* **43**, 1168–1178.
- Hsiue, G. H., Chen, J. K. & Liu, Y. L. (2000), 'Synthesis and characterization of nanocomposite of polyimide-silica hybrid film nanoaqueous sol-gel process', *Journal of Applied Polymer Science* **76**, 1609–1618.
- Lee, A. & Lichtenhan, J. D. (1999), 'Thermal and viscoelastic property of epoxy-clay and hybrid inorganic-organic epoxy nanocomposites', *Journal of Applied Polymer Science* **73**, 1993–2001.

- Liu, Y. L., Wei, W. L., Hsu, K. Y. & Ho, W. H. (1999), 'Thermal stability of epoxy-silica hybrid materials by thermogravimetric analysis', *Thermochimica Acta* **412**, 139–147.
- Mehta, S., Mirabella, F. M., Bafna, A. & Rufener, K. (2004), 'Thermoplastic olefin/clay nano-composites: morphology and mechanical properties', *Journal of Applied Polymer Science* **92**, 928–936.
- Mohammad, A. & Simon, G. P. (2006), Rubber-clay nanocomposites, in M. Yiu-Wing & Y. Zhong-Zhen, eds, 'Polymer Nanocomposites', Woodhead Publishing Limited.
- Naya, S., Cao, R. & Artiaga, R. (2003), 'Local polynomial estimation of TGA derivatives using logistic regression for pilot bandwidth selection', *Thermochimica Acta* **6**, 319–322.
- Petrie, E. M. (2006), *Epoxy Adhesive Formulations*, McGraw-Hill.
- Preghenella, M., Pegoretti, A. & Migliaresi, C. (2005), 'Thermo-mechanical characterization of fumed silica-epoxy nanocomposites', *Polymer* **46**, 12065–12072.
- Ramsay, J. O. & Silverman, B. W. (1997), *Functional Data Analysis*, Springer.
- Ramsay, J. O. & Silverman, B. W. (2005), *Applied Functional Data Analysis*, Springer.
- Schadler, L. S. (2003), Polymer-based and polymer-filled nanocomposites, in P. M. Ajayan, L. S. Schadler & P. V. Braun, eds, 'Nanocomposite Science and Technology', Wiley-VCH, Weinheim, pp. 77–135.
- Shao-Yun, F., Xi-Qiao, F., Bernd, L. & Yiu-Wing, M. (2008), 'Effects of particle size, particle/matrix interface adhesion and particle loading on mechanical properties of particulate-polymer composites', *Composites: Part B* **39**, 933–961.
- Tarrio-Saavedra, J., López-Beceiro, J., Naya, S. & Artiaga, R. (2008), 'Effect of silica content on thermal stability of fumed silica/epoxy composites', *Polymer Degradation and Stability* **93**, 2133–2137.
- Yousefi, A., Lafleur, P. G. & Gauvin, R. (1997), 'Kinetic studies of thermoset cure reactions: a review', *Polymer Composites* **18**, 157–168.
- Zhang, H., Zhang, Z., Friedrich, K. & Eger, C. (2006), 'Property improvements of in situ epoxy nanocomposites with reduced interparticle distance at high nanosilica content', *Acta Materials* **54**, 1833–1842.