



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Dpto. de Estadística e Investigación Operativa  
Aplicadas y Calidad

Modelado y predicción de la evolución de la temperatura  
máxima de vaina en un transitorio LBLOCA

Trabajo Fin de Máster

Máster Universitario en Ingeniería de Análisis de Datos, Mejora de  
Procesos y Toma de Decisiones

AUTOR/A: Cárcel Robredo, Jaime

Tutor/a: Martón Lluch, Isabel

Cotutor/a: Sánchez Galdón, Ana Isabel

CURSO ACADÉMICO: 2023/2024

## **Agradecimientos**

En primer lugar, me gustaría agradecer tanto la ayuda como la disponibilidad a mis tutoras Isabel y Ana. Gracias por aceptar mi propuesta para ser mis tutoras del trabajo y gracias por ayudarme siempre en lo que ha sido necesario.

Por otro lado, me gustaría dar las gracias a mis padres y a mi hermano, los cuales siempre han estado ahí cuando lo he necesitado y siempre me han apoyado tanto en los buenos como en los malos momentos, tanto personales como académicos.

Y, por último, solo me queda dar las gracias a mis compañeros del máster, por haberme brindado un año increíble, ya que sin ellos no hubiese sido lo mismo. Gracias a todos.

# Resumen

Actualmente, los análisis de accidentes de plantas nucleares se realizan mediante la aplicación de la metodología BEPU (“Best Estimate Plus Uncertainty”) basada en el uso de códigos termohidráulicos de simulación BE (“Best Estimate”) junto con el tratamiento de las incertidumbres que se introducen en el proceso de análisis. Estos códigos tienen un elevado coste computacional con tiempos de simulación elevados por lo que, una práctica, ampliamente aceptada es la sustitución del código por un metamodelo (modelo surrogado). El objetivo del presente trabajo final de máster es extender la metodología BEPU al caso de un output funcional, es decir, analizando la evolución de la variable de seguridad considerada a lo largo del transitorio en función de las condiciones bajo las cuales se inicia el accidente objeto de estudio. De esta forma, es posible el desarrollo de modelos dinámicos que representen la evolución de las variables relevantes en función del tiempo, permitiendo una evaluación más precisa y realista. El caso de aplicación se centra en el estudio de un transitorio LBLOCA (“Loss of Coolant Accident”), comúnmente estudiado en el contexto BEPU, en un reactor de agua a presión (PWR) considerando como variable de seguridad la evolución de la máxima temperatura de vaina en el núcleo del reactor en función del tiempo.

# Resum

Actualment, les anàlisis d'accidents de plantes nuclears es realitzen mitjançant l'aplicació de la metodologia BEPU (“Best Estimate Plus Uncertainty”) basada en l'ús de codis termo hidràulics de simulació BE (“Best Estimate”) juntament amb el tractament de les incerteses que s'introdueixen en el procés d'anàlisi. Estos codis tenen un elevat cost computacional amb temps de simulació elevats pel que, una pràctica, àmpliament acceptada és la substitució del codi per un metamodel (model subrogat). L'objectiu del present treball final de màster és estendre la metodologia BEPU al cas d'un output funcional, és a dir, analitzant l'evolució de la variable de seguretat considerada al llarg del transitori en funció de les condicions sota les quals s'inicia l'accident objecte d'estudi. D'esta manera, és possible el desenvolupament de models dinàmics que representen l'evolució de les variables rellevants en funció del temps, permetent una avaluació més precisa i realista. El cas d'aplicació se centra en l'estudi d'un transitori LBLOCA (“Loss of Coolant Accident”), comunament estudiat en el context BEPU, en un reactor d'aigua a pressió (PWR) considerant com a variable de seguretat l'evolució de la màxima temperatura de baina en el nucli del reactor en funció del temps.

# **Abstract**

Currently, nuclear power plant accident analyses are conducted using the Best Estimate Plus Uncertainty (BEPU) methodology, which employs best estimate thermal-hydraulic simulation codes (BE codes) along with the treatment of uncertainties introduced in the analysis process. These codes involve high computational costs with lengthy simulation times, leading to a widely accepted practice of replacing the code with a metamodel (subrogated model). The objective of this master's thesis is to extend the BEPU methodology to the case of a functional output, specifically analyzing the evolution of the safety variable considered throughout the transient as a function of the initial conditions under which the accident of interest is initiated. This approach allows for the development of dynamic models that depict the evolution of relevant variables over time, enabling a more precise and realistic evaluation. The application case focuses on studying a Loss of Coolant Accident (LBLOCA), commonly examined within the BEPU context, in a Pressurized Water Reactor (PWR), with the maximum fuel cladding temperature evolution over time serving as the safety variable of interest.

## Tabla de contenido

|   |    |
|---|----|
| 1. Introducción y justificación   | 1  |
| 2. Contexto. Metodología BEPU   | 4  |
| 2.1 Descripción del transitorio   | 5  |
| 2.2 Criterios de aceptación, FOMs y límites de seguridad  | 6  |
| 2.3 Identificación de la fenomenología física relevante y selección de las variables TH de entrada apropiadas.                  | 6  |
| 2.4 Descripción del modelo TH   | 7  |
| 3. Materiales y métodos   | 10 |
| 3.1. Materiales   | 10 |
| 3.2. Métodos  | 12 |
| 3.2.1 Análisis de datos funcionales   | 12 |
| 3.2.2 Representación en bases de funciones  | 14 |
| 3.2.3 Regresión funcional   | 16 |
| 3.2.4 Selección de variables predictoras  | 17 |
| 3.2.5 Método de corrección de error   | 18 |
| 3.2.6 Intervalos de tolerancia  | 19 |
| 4. Resultados caso de aplicación  | 24 |
| 4.1 Análisis descriptivo  | 25 |
| 4.2 Selección de variables según los tamaños de muestra   | 27 |
| 4.2.1 Resultados de la selección de variables   | 27 |
| 4.3 Modelo de regresión funcional   | 31 |
| 4.4 Construcción de intervalos de tolerancia de Wilks   | 36 |
| 4.5 Comparación de la predicción con los intervalos de tolerancia   | 38 |
| 5. Conclusiones y trabajo futuro  | 40 |
| 6. Referencias  | 41 |
| Anexos  | 44 |
| Anexo 1. Código RStudio para la selección de variables y obtención de coeficientes (paquete “refund”)                           | 44 |
| Anexo 2. Evolución temporal de los coeficientes del modelo de regresión funcional   | 47 |
| Anexo 3. Código de obtención de intervalos de tolerancia para un tamaño de muestra $n=93$ (Método de Wilks 95/95 para un E.O 1) | 50 |
| Anexo 4. Relación del trabajo con los Objetivos de Desarrollo Sostenible de la Agenda 2030                                      | 52 |

## Índice de figuras

|  |    |
|--|----|
| Figura 1. Etapas metodología BEPU (Sánchez Sáez, 2017).  | 4  |
| Figura 2. Vista de SNAP del primario del typical PWR realizado para TRACE (Sánchez Sáez, 2017).                          | 8  |
| Figura 3. Vista de SNAP del secundario del typical PWR realizado para TRACE (Sánchez Sáez, 2017).                        | 9  |
| Figura 4. Evolución temporal de la máxima temperatura de vaina en 450 simulaciones.                                      | 11 |
| Figura 5. Etapas caso de aplicación.   | 24 |
| Figura 6. Función de la media muestral.  | 25 |
| Figura 7. Función de la varianza muestral.   | 26 |
| Figura 8. Box-plot funcional.  | 27 |
| Figura 9. Evolución temporal de los coeficientes para n=450.   | 28 |
| Figura 10. Evolución R2 en el conjunto de test.  | 31 |
| Figura 11. Evolución de los coeficientes para los cuatro primeros inputs.  | 33 |
| Figura 12. Evolución temporal de la PCT frente a su predicción.  | 34 |
| Figura 13. Evolución del percentil 95 y la media de la distribución del error.   | 35 |
| Figura 14. Comparación PCT y predicciones sin y con corrección de error.   | 35 |
| Figura 15. Banda de confianza 95% obtenida a partir del modelo de regresión funcional con corrección de error (n=10000). | 36 |
| Figura 16. Comparación intervalos de tolerancia 95/95 (n=93, 124 y 153).   | 37 |
| Figura 17. Zoom de los intervalos de tolerancia para los 3 E.O.  | 37 |
| Figura 18. Comparación banda de confianza vs Wilks.  | 38 |

## Índice de tablas

|  |    |
|--|----|
| Tabla 1. Variables de entrada seleccionadas.   | 10 |
| Tabla 2. Tamaños de muestra (n) para los diferentes estadísticos intervalo bilateral de orden en el método de Wilks. | 23 |
| Tabla 3. Variables seleccionadas para el modelo de regresión funcional.  | 29 |
| Tabla 4. Comparación de las variables seleccionadas con el modelo de regresión funcional y un estudio previo.        | 30 |
| Tabla 5. Máximo valor del límite superior de tolerancia obtenido con el método de Wilks.                             | 38 |
| Tabla 6. Valor máximo de la evolución de la PCT para los diferentes métodos analizados.                              | 39 |
| Tabla 7. Relación del trabajo con los Objetivos de Desarrollo Sostenible de la Agenda 2030.                          | 52 |

## Lista de acrónimos

ADF – Análisis de Datos Funcionales

ADS – Análisis Determinista de Seguridad

BE – Best Estimate

BEMUSE - Best Estimate Methods plus Uncertainty and Sensitivity Evaluation

BEPU – Best Estimate Plus Uncertainty

EO – Estadístico de Orden

FFT – Transformada Rápida de Fourier

FoM – Figuras de Mérito

IAEA - Agencia Internacional de la Energía Atómica

LBLOCA – Large Break Loss of Coolant Accident

LIT – Límite Inferior de Tolerancia

LPIS – Inyección de seguridad a baja presión

LST – Límite Superior de Tolerancia

mas – Muestreo aleatorio simple

PCT – Temperatura Máxima de la Vaina

PREMIUM - Post-BEMUSE Reflood Model Input Uncertainty Methods

PIRT - Phenomena Identification and Ranking Table

PWR – Reactor de Agua a Presión

STL – Nivel Estándar de Tolerancia

TRACE - TRAC/RELAP Advanced Computational Engine



## 1. Introducción y justificación

Los análisis deterministas de seguridad (ADS) de accidentes de plantas nucleares se han realizado, tradicionalmente, con modelos computacionales e hipótesis de carácter conservador. A finales de los años 80, se empezó a autorizar internacionalmente el uso de modelos e hipótesis realistas con estimación de incertidumbres, denominados BEPU (“Best Estimate Plus Uncertainty”). Desde entonces, las metodologías BEPU de análisis de accidentes han ido aumentando su presencia en el contexto de la seguridad nuclear.

En el año 2009, la IAEA (Agencia Internacional de la Energía Atómica) emitió la guía “Deterministic Safety Analysis for Nuclear Power Plants. Specific Safety Guide SSG-2” (IAEA, 2009). En esta guía se plantean diferentes opciones para realizar los ADS. Estas opciones dependen del nivel de conservadurismo asociado a los supuestos relativos, entre otros, al tipo de cálculo computacional, la disponibilidad de sistemas y las condiciones iniciales y de contorno. Las opciones 1 y 2 son conservadoras y se utilizan desde la aparición de los primeros diseños de centrales nucleares. En la actualidad, se siguen utilizando todavía en la fase de diseño y de licenciamiento de las centrales. Sin embargo, desde hace algunos años, se está tendiendo a utilizar modelos más realistas, por lo que se está optando por la opción 3, la cual se basa en el uso de códigos “Best Estimate” (BE) junto con el tratamiento de incertidumbres que se introducen en el proceso de análisis, dando lugar a la metodología BEPU. Algunos de los factores más importantes que justifican la tendencia de utilizar metodología BEPU (Petruzzi, D'Auria, Gianotti, & Ivanov, 2005) son:

- Cálculo más realista de los márgenes de seguridad de la planta para su licenciamiento.
- Mejoras en las respuestas de emergencia ya que dichas respuestas pueden basarse en supuestos más realistas.
- Identificación de los parámetros que producen un mayor impacto en la incertidumbre de la salida, lo cual sirve de guía para el desarrollo de futuros modelos y/o experimentos.

El presente trabajo final de máster se enmarca en la metodología BEPU, las cuales se fundamentan, en su mayoría, en la modelización probabilista de la incertidumbre y su propagación utilizando diferentes métodos. Uno de los métodos utilizados es el Monte Carlo puro el cual se basa en el muestreo aleatorio simple (m.a.s.) de los inputs (modelizados como variables aleatorias) y la evaluación mediante un código computacional de las denominadas Figuras de Mérito (FoM), las cuales corresponden a determinadas variables de seguridad como, por ejemplo, la temperatura máxima de vaina. Habitualmente, los códigos utilizados tienen un alto coste computacional por lo que una estimación precisa de la función de densidad de la FoM de interés resulta prohibitiva.

Una alternativa al método de Monte Carlo, comúnmente aceptada por los organismos reguladores, en el contexto BEPU, consiste en que en el análisis de incertidumbres del comportamiento de las variables de seguridad sea suficiente con demostrar estadísticamente que se cumplen con los límites establecidos por el organismo regulador con un nivel de confianza del 95% en el 95% de los casos, criterio que es conocido como nivel estándar de tolerancia 95/95 (STL 95/95) y permite la evaluación de las variables de seguridad respecto a los límites establecidos. El método más extendido para la

estimación del STL 95/95 es el método no paramétrico de Wilks (Wilks, 1941) que permite obtener el STL con una determinada cobertura/confianza ( $\gamma/\beta$ ) utilizando un número reducido de simulaciones del código de simulación.

Otra alternativa consiste en la sustitución del código computacional por un metamodelo (también denominado modelo sustitutivo o emulador). El uso de un metamodelo presenta la ventaja de poder realizar un número elevado de simulaciones con un bajo coste computacional permitiendo reemplazar al código en la propagación de incertidumbres. En la literatura, en el contexto de la metodología BEPU, se han realizado estudios utilizando diferentes tipos de metamodelos como, por ejemplo, modelos de mezclas finitas (Delgado Manríquez, Cid Cid, & Leguey Galán, 2013), modelos aditivos generalizados (Sánchez Sáez, 2017), redes neuronales (Sallehuddin & Diab, 2021) o máquinas de vectores de soporte (Najar & Wang, 2022). En (Freixa, y otros, 2024) se presenta una revisión exhaustiva de la aplicación de metodologías BEPU en España y del uso de metamodelos.

Tradicionalmente, estos metamodelos se utilizan para la predicción de una FoM escalar (p.e. la temperatura máxima de vaina). Si el objetivo es la estimación de la evolución temporal de la variable de seguridad en función de un conjunto de inputs inciertos el problema resulta más complejo. En este contexto, una opción es la sustitución del código de simulación por un modelo de regresión con respuesta funcional. El análisis de datos funcionales y, en concreto, los modelos de regresión funcional, es una de las áreas de estadística que ha despertado un gran interés en los últimos años dada su aplicabilidad a campos tan diversos como la medicina, la biología, la economía y la ingeniería (Morris, 2015). (Ramsay & Silverman, 2005) introducen las bases metodológicas del análisis de datos funcionales, así como la extensión de los métodos y técnicas de la estadística clásica a datos funcionales.

En este contexto, el objetivo del presente trabajo final de máster es extender el análisis de incertidumbre, en la metodología BEPU, considerando como la FoM la evolución temporal de la temperatura máxima de vaina en el núcleo en un transitorio LBLOCA (Large Break Loss of Coolant Accident) que consiste en una rotura de doble guillotina en la rama fría en un reactor de agua a presión (PWR). La selección de un LBLOCA está motivada por tratarse de un transitorio de referencia en la literatura siendo el analizado en proyectos internacionales como BEMUSE (Best Estimate Methods plus Uncertainty and Sensitivity Evaluation) (Perez, y otros, 2011) o PREMIUM (Post-BEMUSE Reflood Model Input Uncertainty Methods) (Reventós, y otros, 2016).

En relación con los objetivos específicos, para la consecución del objetivo general, los mismos son:

- Analizar la aplicabilidad de la regresión con respuesta funcional como metamodelo en el análisis de incertidumbre.
- Comparar los resultados obtenidos con la regresión con respuesta funcional y otras técnicas empleadas en el análisis de incertidumbre como, por ejemplo, Wilks.
- Evaluar la capacidad de predicción y generalización del modelo de regresión funcional frente a la variabilidad temporal de la temperatura de la vaina.
- Identificar y seleccionar las variables más significativas que presentan un mayor efecto sobre la temperatura de la vaina durante el transitorio LBLOCA, teniendo en

cuenta los momentos críticos, identificados tras el estudio de la evolución temporal de la PCT.

Mediante el presente TFM se pretende también contribuir al alcance de varios de los ODS (Objetivos de Desarrollo Sostenible) definidos por la ONU para 2030. En concreto, las aportaciones del TFM están directamente relacionadas con ODS nº7: Energía asequible y no contaminante, al mejorar la seguridad y eficiencia en la operación de plantas nucleares; ODS nº9: Industria, innovación e infraestructura, al desarrollar y aplicar metodologías avanzadas en el análisis de seguridad nuclear; y el ODS nº13: Acción por el clima, ayudando a reducir las emisiones de gases de efecto invernadero, dando paso a una energía más sostenible y menos contaminante.

Además, el trabajo ha sido realizado en base a los conocimientos aprendidos en el Máster Universitario de Ingeniería de Análisis de Datos, Mejora de procesos, y Toma de Decisiones, en la Universitat Politècnica de València, concretamente a lo estudiado en las asignaturas de Fiabilidad, Disponibilidad y Mantenibilidad [32051] (proporcionó los fundamentos y metodologías necesarias para evaluar la fiabilidad y disponibilidad de sistemas, aspectos cruciales en el análisis de seguridad de plantas nucleares) y Análisis de Datos Funcionales [35129] (ofreció las herramientas necesarias para aplicar modelos de regresión funcional, permitiendo el análisis detallado de la evolución temporal de variables críticas como la temperatura máxima de la vaina en transitorios nucleares).

El proyecto está estructurado en las siguientes secciones. En la sección 2 se presenta el contexto e incluye una introducción a la metodología BEPU y la descripción del transitorio objeto de análisis.

A continuación, en la sección 3, se detalla la metodología utilizada, incluyendo la implementación del análisis de incertidumbre y la aplicación de la regresión funcional.

Posteriormente, en la sección 4, se presentan los resultados del caso de aplicación, donde se estudiará los resultados de la simulación del transitorio, centrándose en la evolución temporal de la temperatura máxima de la vaina (PCT). Una vez completado el análisis, se procede a realizar una predicción basada en la regresión funcional obtenida, comparando así su eficacia frente a otras técnicas, y evaluando de esta manera la precisión y robustez del modelo.

Por último, en la sección 6 se presentan las conclusiones del estudio.

## 2. Contexto. Metodología BEPU

Como se ha comentado en el apartado anterior, el presente trabajo, se centra en la aplicación de la metodología BEPU la cual se basa en el uso de códigos de simulación BE junto con el tratamiento de las incertidumbres que se introducen en el proceso de análisis. En la Figura 1 se muestra de forma esquemática las diferentes etapas que constituyen la metodología BEPU.

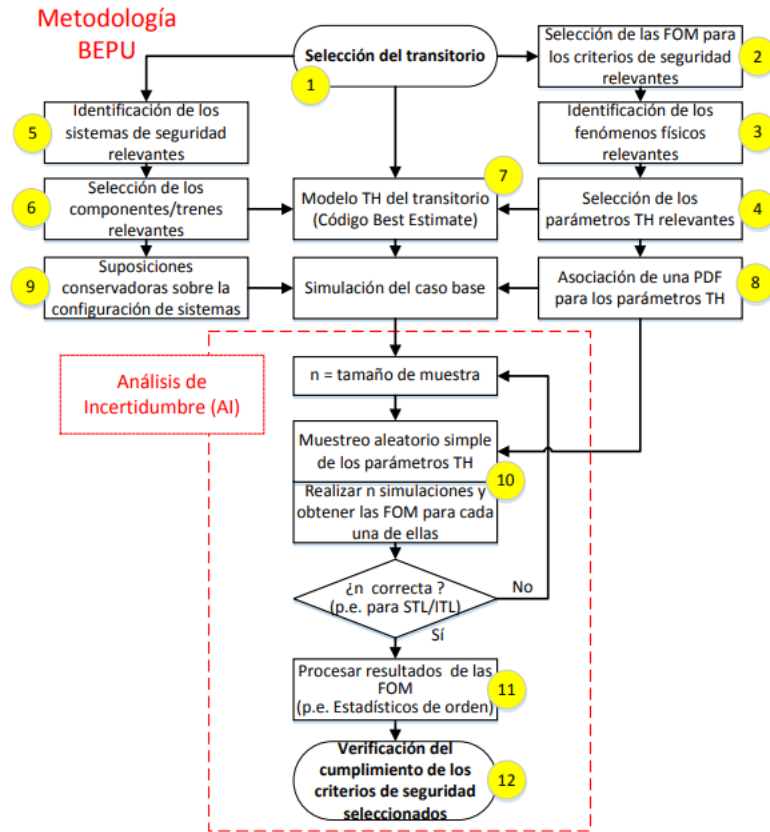


Figura 1. Etapas metodología BEPU (Sánchez Sáez, 2017).

Como se observa en la Figura 1, la metodología se inicia con la selección del transitorio, es decir del accidente, objeto de estudio. Para ello se recopila toda la información que pueda ser de interés: datos experimentales, estudios anteriores sobre el mismo tipo de secuencia accidental o de transitorios similares, etc.

Una vez identificado el transitorio de interés, el siguiente paso consiste en establecer los criterios de seguridad aplicables a dicho transitorio y definir las FOM (variable(s) de seguridad de salida) asociadas. Esto implica identificar las variables de seguridad y los parámetros medibles relacionados con el transitorio.

Después de establecer los criterios de seguridad, se procede a identificar los fenómenos físicos más relevantes que ocurren durante el transitorio seleccionado. Este paso es esencial para determinar los parámetros de entrada cruciales para el análisis de incertidumbre. Al comprender estos fenómenos, se pueden tomar decisiones más informadas sobre cómo modelizar el transitorio y qué aspectos son más críticos para la seguridad.

A partir de los fenómenos físicos identificados, se seleccionan las variables de entrada (predictores) relevantes para el análisis y se considera la incertidumbre asociada a estos parámetros para propagarla hacia las FOM previamente seleccionadas. Esto implica evaluar qué variables son críticas para el comportamiento del sistema y cómo su variabilidad puede afectar a la seguridad.

Una vez identificados los parámetros relevantes, se procede al desarrollo de un modelo termohidráulico del transitorio utilizando un código de simulación. Es importante asignar valores nominales a los parámetros de entrada (caso base) y realizar simulaciones para calibrar y validar el modelo con datos experimentales u otros estudios previos.

Para reflejar la incertidumbre asociada a los parámetros de entrada, se asigna una distribución de probabilidad a cada uno de ellos utilizando estudios previos o la opinión de expertos en el campo. En ausencia de información, es común asignar una distribución uniforme con valores cercanos al valor nominal del parámetro.

Los parámetros de entrada se muestrean aleatoriamente utilizando las distribuciones de probabilidad asignadas. Este proceso se repite tantas veces como simulaciones se vayan a realizar con el código BE. El muestreo aleatorio permite explorar la variabilidad de los parámetros y su impacto en las variables de seguridad seleccionadas.

Una vez completadas las simulaciones, se procede al procesamiento de los resultados mediante la aplicación de técnicas estadísticas para el análisis de incertidumbre y obtener información sobre las FOMs. Dicho análisis puede realizarse mediante la aplicación de diferentes técnicas como el método no paramétrico de Wilks o el uso de metamodelos.

Finalmente, se verifica si las FOMs obtenidas cumplen con los criterios de seguridad establecidos por el organismo regulador.

## 2.1 Descripción del transitorio

El transitorio que se analiza en el estudio es un accidente de pérdida de refrigerante a través de una rotura grande LBLOCA (Large Break Loss of Coolant Accident) en un reactor PWR. La selección de este transitorio es consecuencia de la importancia que el mismo tiene en el análisis de seguridad de las centrales nucleares y en el cual se centran la mayoría de los estudios BEPU realizados hasta el momento. Concretamente, en este trabajo, se ha elegido el mismo transitorio que se estudió en el proyecto internacional BEMUSE (Best Estimate Methods plus Uncertainty and Sensitivity Evaluation) (Perez , y otros, 2011). El transitorio en cuestión consiste en una rotura en doble guillotina en la rama fría de una central PWR de 4 lazos, en concreto la planta Zion, construida por la empresa Westinghouse Electric Corporation, situada en el norte de Illinois (EE. UU.). Una descripción detallada del transitorio analizado puede ser consultado en (Perez , y otros, 2011).

## 2.2 Criterios de aceptación, FOMs y límites de seguridad

Los criterios de seguridad que aplican a un LBLOCA de acuerdo con la USNRC (Acceptance criteria for emergency core cooling systems for light-water nuclear nuclear power reactors, 2007) son:

- Temperatura máxima de vaina (PCT). Esta temperatura se mide en el combustible del reactor y no debe sobrepasar los 1477 K para ninguna barra.
- Oxidación máxima de la vaina. El grosor de la vaina tras la oxidación no debe exceder 0.17 veces el grosor total del revestimiento de la vaina antes de la oxidación.
- Generación máxima de hidrógeno. La cantidad total de hidrógeno generada tras la reacción química de oxidación que se produce entre la vaina y el agua no debe superar el 0.01 del total de hidrógeno que se generaría tras reaccionar todo el metal de las vainas que envuelven las barras de combustible.

Para la realización del análisis de incertidumbre de este trabajo, teniendo en cuenta los criterios de seguridad citados, se escoge como variable de seguridad la temperatura máxima de vaina (PCT). A diferencia de estudios previos, en este trabajo se analiza la evolución temporal de la temperatura máxima de vaina en el núcleo en función del tiempo.

## 2.3 Identificación de la fenomenología física relevante y selección de las variables TH de entrada apropiadas.

Considerando la variable de seguridad de interés en un LBLOCA, resulta necesario identificar los parámetros significativos que puedan influir en la misma. La selección de los parámetros requiere del análisis de estudios previamente realizados por expertos del área. Esta información se encuentra en las tablas PIRT (Phenomena Identification and Ranking Table). Estos análisis se han llevado a cabo en el contexto de LBLOCA en reactores de agua a presión con configuración similar de 4 lazos.

Aunque no existe un consenso respecto a las variables que representan los fenómenos en cada fase considerada, muchos de estos parámetros son compartidos en la mayoría de los PIRT realizados. Los primeros estudios PIRT abordaron la mayoría de las variables que se han aceptado comúnmente, siguiendo el método original CSAU (USNRC, 1989b). En comparación, estudios PIRT más recientes, como los de AREVA (Martin & O'Dell, 2005) y Westinghouse/EPRI (Young & Bajorek, 1998), han introducido condiciones iniciales específicas de la planta y otros parámetros adaptados a cada instalación. Además, se han considerado los resultados de estudios más recientes como BEMUSE (Perez , y otros, 2011) y (NEA., 2015). En base a estos estudios, se han seleccionado un total de 45 inputs (predictores), las cuales han demostrado su relevancia y se encuentran descritos en (Sáez, 2017).

## 2.4 Descripción del modelo TH

El modelo termohidráulico consiste en un reactor nuclear de tipo Westinghouse de 4 lazos, comúnmente denominado "typical PWR". Este modelo se basa en la central nuclear de Zion, con una potencia nominal de 3250 MW térmicos, y se han seguido las directrices y recomendaciones del proyecto BEMUSE (Perez , y otros, 2011).

El modelo termohidráulico empleado ha sido adaptado para simular de manera integral el sistema primario y secundario, incluyendo los sistemas operativos y de seguridad esenciales. Esto implica cuatro inyecciones de seguridad, tanto a alta como a baja presión, independientes para cada uno de los cuatro lazos, así como cuatro acumuladores ubicados en las ramas frías de los lazos del primario, situados a la entrada de la vasija.

El código de simulación utilizado es el TRACE (TRAC/RELAP Advanced Computational Engine) (Belaid, Freixa, & Zerkak, 2010) que es un código diseñado para el análisis de LBLOCA tanto para reactores de tipo presurizado como de ebullición. Se trata de un código determinista que, bajo los mismos parámetros de entrada, proporcionará el mismo valor de las FOM de interés.

En lo que respecta a los sistemas y funciones de seguridad del primario reproducidos en el modelo TRACE, se incluyen:

- Acumuladores: Tres acumuladores independientes, uno en cada rama fría de los lazos intactos, controlados por la presión de su respectivo lazo.
- Inyección de seguridad a baja presión (LPIS): Similar a los acumuladores, se implementa una inyección en rama fría por lazo intacto, controlada por la presión de su respectivo lazo.

En cuanto al sistema secundario:

- Secundario de los generadores de vapor de cada uno de los 4 lazos, con descargas independientes.
- Se controla la inyección de agua de alimentación para mantener el nivel del secundario de los generadores de vapor, así como la presión.
- Se regula la válvula de descarga de vapor a la turbina para alcanzar la temperatura de rama fría en el primario requerida.

Las Figuras 2 y 3 muestran una vista del primario y del secundario, respectivamente, del modelo termohidráulico empleado, utilizando el programa SNAP (Applied Programming Technology, 2012).

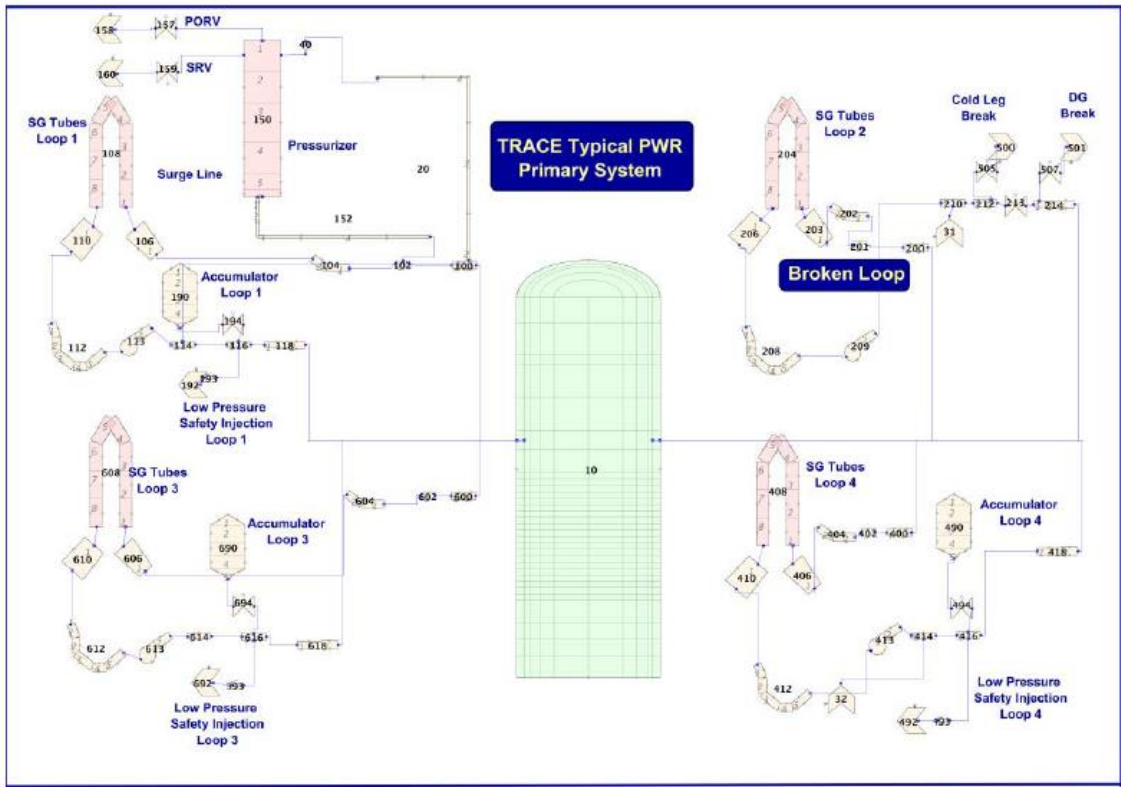


Figura 2. Vista de SNAP del primario del typical PWR realizado para TRACE (Sánchez Sáez, 2017).



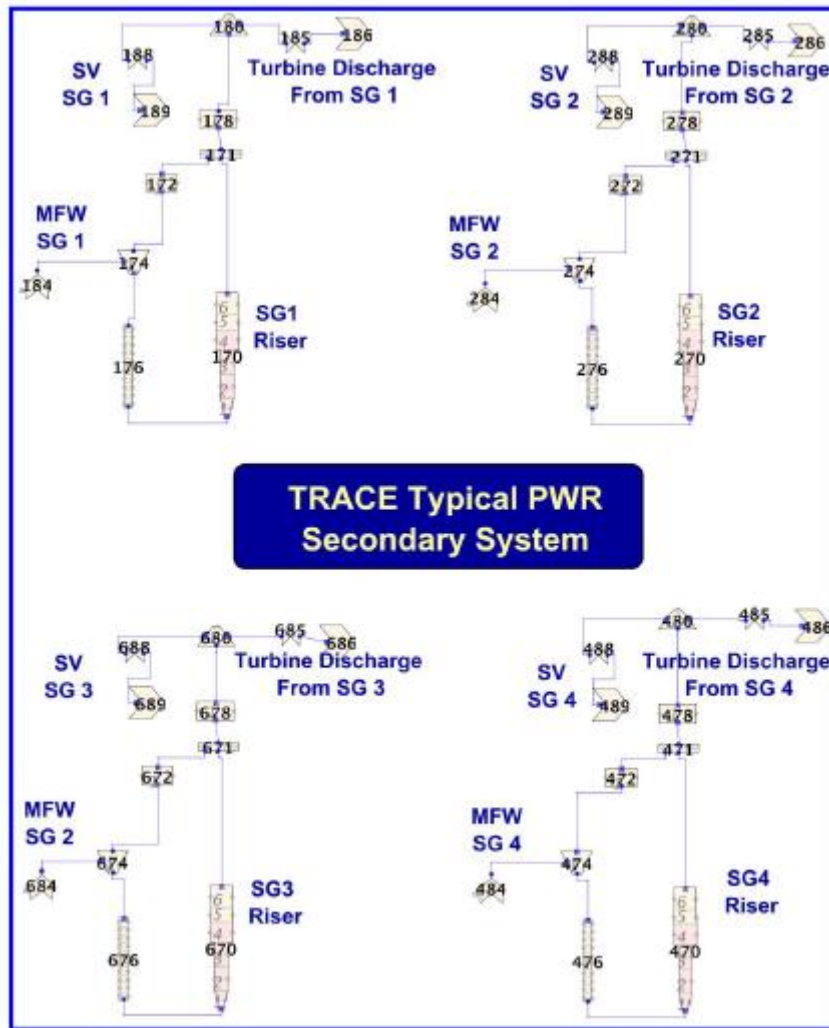


Figura 3. Vista de SNAP del secundario del typical PWR realizado para TRACE (Sánchez S ez, 2017).

### 3. Materiales y métodos

#### 3.1. Materiales

El estudio se ha realizado a partir de un total de 450 simulaciones del transitorio LBLOCA obtenidas mediante el código termohidráulico TRACE (Belaid, Freixa, & Zerkak, 2010). Para la obtención de las simulaciones se han muestreado las variables de entrada mediante muestreo aleatorio simple (m.a.s.) de acuerdo con la distribución de probabilidad de las mismas. El número de variables de entrada consideradas es, como se ha indicado anteriormente, de 45, las cuales han sido seleccionadas a partir de estudios previos.

A continuación, en la Tabla 1, se muestran las 45 variables de entrada consideradas, junto con el modelo de distribución de probabilidad considerado, el valor mínimo y el valor máximo.

Tabla 1. Variables de entrada seleccionadas.

| Parámetros con distribución normal    |   | Variable   | Min  | Max  |
|---------------------------------------|---|------------|------|------|
| <b>INPOW</b>                          | Initial thermal power                                 | <b>X1</b>  | 0.98 | 1.02 |
| <b>PEAKF</b>                          | Peaking factor  | <b>X2</b>  | 0.95 | 1.05 |
| <b>HGSAV</b>                          | Hot gap size average                                  | <b>X3</b>  | 0.8  | 1.2  |
| <b>HGSHR</b>                          | Hot gap size hot rod                                  | <b>X4</b>  | 0.8  | 1.2  |
| <b>RPOWM</b>                          | Residual power multiplier                             | <b>X5</b>  | 0.92 | 1.08 |
| <b>UO2TC</b>                          | UO <sub>2</sub> Thermal Conductivity                  | <b>X6</b>  | 0.9  | 1.1  |
| <b>UO2SH</b>                          | UO <sub>2</sub> Specific Heat                         | <b>X7</b>  | 0.98 | 1.02 |
| <b>ACCPR</b>                          | Accumulator pressure setpoint                         | <b>X8</b>  | -0.2 | 0.2  |
| <b>ACCLT</b>                          | Accumulator liquid temperature                        | <b>X9</b>  | -10  | 10   |
| <b>LPISQ</b>                          | LPI mass flow rate                                    | <b>X10</b> | 0.95 | 1.05 |
| <b>MASFR</b>                          | Initial total primary mass flow rate                  | <b>X11</b> | 0.96 | 1.04 |
| <b>CLTEM</b>                          | Initial cold legs average temperature                 | <b>X12</b> | -2   | 2    |
| Parámetros con distribución lognormal |   | Variable   | Min  | Max  |
| <b>ACCKL</b>                          | Accumulator loss coeff.                               | <b>X13</b> | 0.5  | 2    |
| <b>KSURG</b>                          | Surge line coss coeff.                                | <b>X14</b> | 0.5  | 2    |
| Parámetros con distribución uniforme  |   | Variable   | Min  | Max  |
| <b>KDWNC</b>                          | Downcomer loss coeff.                                 | <b>X15</b> | 0.95 | 1.05 |
| <b>KLCPL</b>                          | Low core plate loss coeff.                            | <b>X16</b> | 0.95 | 1.05 |
| <b>KBRCL</b>                          | Break cold leg loss coeff.                            | <b>X17</b> | 0.95 | 1.05 |
| <b>CHFC1</b>                          | Vessel liquid wall friction multiplier                | <b>X18</b> | 0.95 | 1.05 |
| <b>CHFC2</b>                          | Wall friction multiplier                              | <b>X19</b> | 0.95 | 1.05 |
| <b>GTHCO</b>                          | Gap Conductance coeff.                                | <b>X20</b> | 0.9  | 1.1  |
| <b>CLCON</b>                          | Clad thermal conductivity                             | <b>X21</b> | 0.9  | 1.1  |
| <b>CLDEN</b>                          | Clad density  | <b>X22</b> | 0.95 | 1.05 |
| <b>CONTP</b>                          | Maximum containment pressure                          | <b>X23</b> | 0.85 | 1.15 |
| <b>CONVL</b>                          | Single phase liquid to wall heat transfer coeff.      | <b>X24</b> | 0.9  | 1.1  |
| <b>CONVV</b>                          | Single phase vapor to wall heat transfer coeff.       | <b>X25</b> | 0.9  | 1.1  |
| <b>FILMB</b>                          | Dispersed flow film boiling heat transfer coeff.      | <b>X26</b> | 0.9  | 1.1  |
| <b>TMIFS</b>                          | Film to transition boiling T <sub>min</sub> criterion | <b>X27</b> | 0.9  | 1.1  |
| <b>CRIFH</b>                          | Critical Heat Flux (CHF) multiplier                   | <b>X28</b> | 0.8  | 1.2  |
| <b>FLOSS</b>                          | Form loss coeff.                                      | <b>X29</b> | 0.9  | 1.1  |
| <b>DRAG4</b>                          | Interfacial drag (droplet) coeff.                     | <b>X30</b> | 0.9  | 1.1  |

| Parámetros con distribución uniforme |   | Variable   | Min  | Max |
|--------------------------------------|---|------------|------|-----|
| <b>DRAG5</b>                         | Interfacial drag (dispersed flow film boiling) coeff. | <b>X31</b> | 0.9  | 1.1 |
| <b>DRAG7</b>                         | Interfacial drag (inverted annular flow) coeff.       | <b>X32</b> | 0.9  | 1.1 |
| <b>CCFLC</b>                         | Wallis c for counter current flow limitation          | <b>X33</b> | 0.9  | 1.1 |
| <b>ACCLV</b>                         | Accumulator liquid volume                             | <b>X34</b> | -0.5 | 0.5 |
| <b>LIQA</b>                          | Liquid to wall inverted annular heat transfer coeff.  | <b>X35</b> | 0.9  | 1.1 |
| <b>UHTEM</b>                         | Initial upper head temperature                        | <b>X36</b> | 0    | 10  |
| <b>LINT2</b>                         | Liquid to interface annular-mist heat transfer coeff. | <b>X37</b> | 0.9  | 1.1 |
| <b>LINT3</b>                         | Liquid to interface transition heat transfer coeff.   | <b>X38</b> | 0.9  | 1.1 |
| <b>VINT3</b>                         | Vapor to interface transition heat transfer coeff.    | <b>X39</b> | 0.9  | 1.1 |
| <b>VINT4</b>                         | Vapor to interface stratified heat transfer coeff.    | <b>X40</b> | 0.9  | 1.1 |
| <b>SBHTC</b>                         | Subcooled boiling heat transfer coeff.                | <b>X41</b> | 0.9  | 1.1 |
| <b>DNBHC</b>                         | Departure from nucleate boiling/CHF                   | <b>X42</b> | 0.9  | 1.1 |
| <b>WH001</b>                         | Fuel thermal conductivity before burst coeff.         | <b>X43</b> | 0.9  | 1.1 |
| <b>WH002</b>                         | Cladding metal-water reaction rate coeff.             | <b>X44</b> | 0.9  | 1.1 |
| <b>WH004</b>                         | Burst temperature coeff.                              | <b>X45</b> | 0.9  | 1.1 |

La evolución temporal, durante el transitorio, de la temperatura máxima de vaina para las 450 simulaciones realizadas se muestran en la Figura 4. En dicha figura se observa la similitud en la forma de la evolución de la temperatura, aunque se observa que la máxima temperatura de vaina se alcanza a diferentes valores del tiempo debido a las diferentes condiciones de entrada de cada simulación.

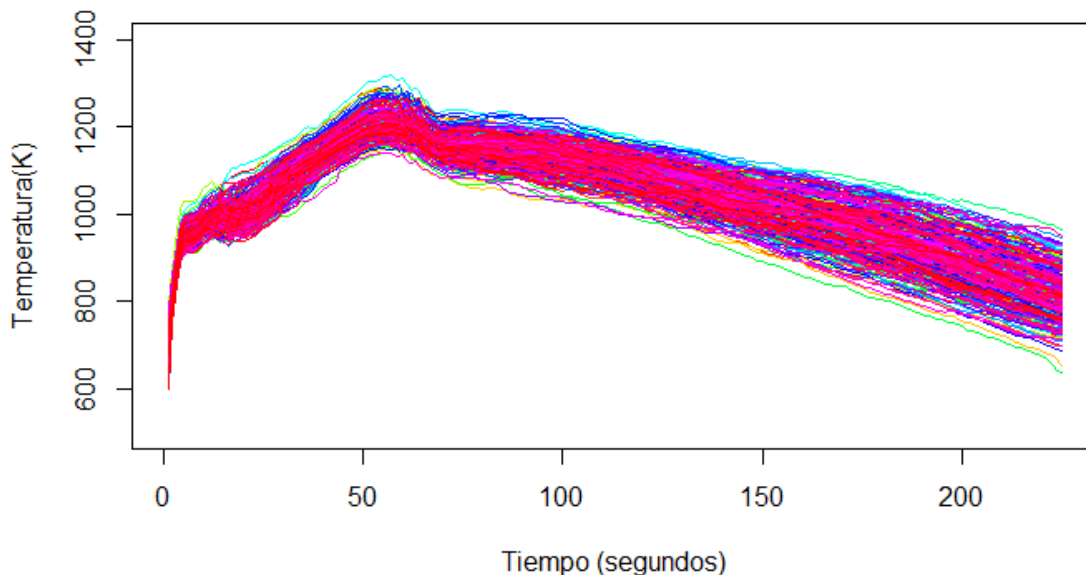


Figura 4. Evolución temporal de la máxima temperatura de vaina en 450 simulaciones.

## 3.2. Métodos

Como se ha descrito en la introducción, el objetivo del trabajo se centra en la propagación de la incertidumbre en el contexto de simulación mediante códigos termohidráulicos considerando como FoM la evolución temporal de la temperatura máxima de vaina en el núcleo en un transitorio.

Los códigos tienen un elevado coste computacional con tiempos de simulación elevados por lo que, una práctica, ampliamente aceptada es la sustitución del código por un metamodelo. Habitualmente, el metamodelo se utiliza para el análisis de incertidumbre de variables de seguridad escalares (temperatura pico de la vaina, máxima oxidación, etc.). El objetivo de los métodos que se utilizan en este trabajo es extender el uso del metamodelo a variables funcionales, es decir, al análisis de incertidumbre considerando la evolución temporal de la variable de seguridad de interés a lo largo del transitorio.

El análisis de datos funcionales (ADF) es un campo que, desde su concepción por (Ramsay & Silverman, 1997), ha experimentado un rápido crecimiento como consecuencia de su amplia aplicabilidad en áreas como la medicina, biología, química e ingeniería, entre otras. El ADF incluye técnicas estadísticas diseñadas para tratar conjuntos de datos creados mediante la observación de funciones o curvas. Si bien los datos subyacentes son curvas, los datos observados suelen ser una representación discretizada de las mismas. Así, cada curva queda representada por un vector de dimensión finita. En este contexto, el ADF desarrolla herramientas que extienden las técnicas clásicas de la estadística a este tipo de datos.

En los apartados siguientes se presentan los fundamentos del ADF y la regresión con datos funcionales, en concreto, la regresión con respuesta funcional.

### 3.2.1 Análisis de datos funcionales

Los pioneros en la conceptualización de datos funcionales fueron Ramsay y Dalzell en 1991 (Ramsay & Dalzell, 1991), y Ramsay y Silverman en 1997 (Ramsay & Silverman, 1997). La noción de datos funcionales surge cuando los valores de una variable pueden ser naturalmente observados como una curva o una función suave.

De acuerdo con la definición formal presentada por Ferraty y Vieu en 2006 (Ferraty & Vieu, 2006), los datos funcionales se pueden caracterizar de la siguiente manera:

**Definición 1:** Se trata de una variable aleatoria  $X$  que toma valores en un espacio funcional  $\varepsilon$  (un espacio normado o seminormado completo).

**Definición 2:** Un conjunto de datos funcionales  $\{X_1, \dots, X_n\}$  representa la observación de  $n$  variables funcionales  $X_1, \dots, X_n$  idénticamente distribuidas.

En sus primeras etapas, los datos funcionales se conceptualizaban como observaciones independientes en un proceso estocástico en un espacio Hilbert  $L^2(T)$ . En este enfoque, las variables se identificaban como curvas en un intervalo  $T = [0, T]$ , y se destacaba su capacidad para suavizarse.

Este planteamiento ha llevado al ADF a un desarrollo significativo, convirtiéndose en una herramienta metodológica valiosa en diversas áreas de investigación, como en ciencias de la salud, neurociencia, economía y en la ingeniería. En el contexto del ADF, se aplican

numerosas metodologías que, al combinarse con análisis clásicos, permiten su aplicación en modelos de regresión, análisis multivariantes, clustering, estadística espacial o estudio de series temporales, entre otros.

Es crucial comprender que, inicialmente, se trabaja con datos discretos que deben convertirse en datos funcionales. Si los datos observados no presentan errores, se realiza una interpolación directa. Por otro lado, si se identifican errores en los datos discretos, se requiere aplicar un proceso de suavizado para corregir cualquier tipo de irregularidad.

Aunque el análisis de datos funcionales, introducido por Ramsay y Silverman en 1997 (Ramsay & Silverman, 1997), es una metodología relativamente reciente, se enfrenta a desafíos similares a otras metodologías. De manera general, este enfoque tiene objetivos diversos que se resumen en los siguientes puntos:

- Representar los datos de manera que permita obtener más detalles e información.
- Analizar los datos para resaltar características de interés.
- Explorar fuentes cruciales de patrones y variaciones entre los datos.
- Explicar la variación en un resultado o variable dependiente utilizando información de entrada o variables independientes.
- Comparar dos o más conjuntos de datos en términos de ciertos tipos de variación, donde estos conjuntos pueden contener diferentes réplicas de las mismas funciones o diferentes funciones para un conjunto común de réplicas.

Cada uno de estos objetivos tiene su propio proceso. Para abordar estos desafíos, se puede adoptar una estrategia exploratoria, confirmatoria o predictiva. La metodología exploratoria se centra en examinar de manera general la información que se puede obtener de los datos. Esta técnica permite descubrir aspectos nuevos e interesantes sobre las observaciones, brindando nuevas posibilidades de estudio. En contraste, la metodología confirmatoria se orienta más hacia la inferencia estadística y está sujeta a reglas estadísticas más estrictas. Estas normas pueden estar relacionadas con la estructura de los datos, el tipo de hipótesis que se pretende evaluar y la formulación misma. Por otro lado, los estudios predictivos son menos comunes y se enfocan en utilizar datos disponibles para afirmar o predecir información no observada.

Además de Ramsay y Silverman, diversos autores han contribuido al desarrollo y expansión de la teoría que rodea el análisis de datos funcionales. Jacques y Preda (2014) (Jacques & Preda, 2014) ofrecen una síntesis metodológica completa en un artículo que aborda esta temática.

Si consideramos una variable funcional aleatoria  $X$  (en este caso, la  $X$  se puede interpretar tanto como variable respuesta, como variable explicativa), la cual es una variable aleatoria con valores en un espacio infinito-dimensional, los datos funcionales que representan un conjunto de observaciones serán  $\{X_1, \dots, X_n\}$  de  $X$ . El modelo subyacente de  $X_i$  generalmente es un conjunto de muestras de variables aleatorias extraídas de la misma distribución que  $X$ .

Un enfoque aceptado para este tipo de datos es considerarlos como un proceso estocástico, donde  $X = \{X_t\}_t \in T$  toma valores en un espacio de Hilbert  $L^2$  de funciones

definidas en algún conjunto  $T$ . Por lo general,  $T$  representa un intervalo de tiempo, longitudes de onda o cualquier otro subconjunto continuo de  $\mathbb{R}$ .

La principal dificultad al tratar con datos funcionales radica en que se supone que las observaciones pertenecen a un espacio infinito-dimensional, mientras que, en la práctica, solo se han muestreado las curvas observadas en un conjunto finito de puntos temporales. De hecho, es común tener solo observaciones discretas  $X_{ij}$  para cada ruta de muestra  $X_i(t)$  en un conjunto finito de nodos  $\{t_{ij} : j = 1, \dots, m_i\}$ . Debido a esto, el primer paso en el ADF suele ser la reconstrucción de la forma funcional de los datos a partir de observaciones discretas. La solución más común para este problema es considerar que las trayectorias pertenecen a un espacio de dimensión finita generado por alguna base de funciones.

### 3.2.2 Representación en bases de funciones

Una base de funciones se compone de funciones conocidas, denotadas como  $\phi_k$ , que son matemáticamente independientes entre sí y poseen la ventaja de describir de manera específica, mediante una suma ponderada o combinación lineal, un número determinado de funciones  $K$  (Erich, Aicke, & Mario, 2014).

A modo de ejemplo, consideremos el sistema de funciones de base más común, la colección de monomios, utilizada para construir series de potencias, como se muestra a continuación:  $1, t, t^2, t^3, t^4$ , y así sucesivamente.

Esta colección puede representarse como un proceso de bases en una función  $x$  para una expansión lineal, como se indica a continuación:

$$x(t) = \sum_{k=1}^K c_k \phi_k(t) \quad [ 1 ]$$

donde  $c$  denota el vector de longitud  $K$  de coeficientes  $c_k$  y  $\phi$  es el vector funcional cuyos elementos funcionales son las bases  $\phi_k$ . También podemos expresar la ecuación anterior de la siguiente forma:

$$x = c' \phi = \phi' c \quad [ 2 ]$$

La expansión de bases nos permite representar la dimensión infinita de funciones en términos de un espacio de vectores finitos  $c$ . La dimensión de esta expansión se denomina  $K$ . Una interpolación real se logra cuando  $K = n$ , donde  $n$  es el número de observaciones. En este contexto, debemos elegir los coeficientes  $c_k$  para obtener  $x(t_j) = y_j$  para cada  $j$ . Así, la suavización de los valores  $y_j$ , en lugar de interpolarlos, está determinada por el número  $K$  de funciones básicas, un parámetro que seleccionamos según las características de los datos disponibles.

El propósito de las bases es capturar las mismas características de las funciones que están siendo estimadas. Cuanto menor sea  $K$  y más adecuadamente representen las bases a las funciones, más grados de libertad tendremos para probar las hipótesis (en caso de trabajar en esta área).

Existen diversas bases disponibles, como polinomios, series de Fourier o B-spline, cada una con sus propias ventajas y desventajas. Algunas pueden no ser apropiadas para ciertas aplicaciones, y aún no existe un marco universal.

Siguiendo la bibliografía, destaca la síntesis realizada por Curtis y Bárcenas Sánchez (2016) en su publicación "Estudio de Perfiles de Olas vía Análisis de Datos Funcionales (Bárcenas Curtis, 2016). Al convertir datos discretos en funciones, se abordan dos etapas:

1. Elegir una base adecuada según la naturaleza de los datos.
2. Determinar los coeficientes que definen la función como una combinación lineal de esas funciones base.

En la primera etapa, se consideran diversos tipos de bases, y se selecciona la más adecuada en función de diferentes criterios. Por ejemplo, se sugiere utilizar una base de Fourier para datos periódicos o una base B-spline para datos no periódicos. El objetivo es lograr una buena estimación con un número relativamente pequeño de funciones base ( $K$ ), y se valora que las derivadas proporcionen información relevante para el análisis.

A continuación, se presentan los principales tipos de bases disponibles y sus características.

Una de las funciones más populares y ampliamente utilizadas en el análisis de datos funcionales es la función de Fourier, donde las bases son senos y cosenos de frecuencia creciente:

$$\phi_1(t) = 1 \quad [ 3 ]$$

$$\phi_2(t) = \sin(\omega t) \quad [ 4 ]$$

$$\phi_3(t) = \cos(\omega t) \quad [ 5 ]$$

$$\phi_4(t) = \sin(2\omega t) \quad [ 6 ]$$

$$\phi_5(t) = \cos(2\omega t) \quad [ 7 ]$$

y así sucesivamente, donde  $\omega = P2\pi$  y  $P$  es el período de oscilación.

Esta base se utiliza y define especialmente cuando los datos disponibles son periódicos o cercanamente periódicos. Si los valores de  $t_j$  son iguales en el espacio  $T$  y el período es igual a la longitud del intervalo  $T$ , entonces la base es ortogonal. En este sentido, se obtiene el producto de la matriz diagonal  $\phi^T \phi$ , que puede calcularse igual que la identidad si dividimos las funciones de base por las constantes adecuadas ( $\sqrt{n}$  para  $j = 0$  y  $\sqrt{n}/2$  para todo  $j$ ).

La Transformada Rápida de Fourier (FFT) permite encontrar todos los coeficientes de manera extremadamente eficiente, especialmente cuando  $n$  es una potencia de 2 y los argumentos están igualmente espaciados. En este caso, se pueden calcular tanto los coeficientes  $c_k$  como todos los  $n$  valores suavizados en  $x(t_j)$  en operaciones de complejidad  $O(n \log n)$ . Esta eficiencia ha convertido a la serie de Fourier en la base

tradicional de elección para series de tiempo prolongado. Sin embargo, técnicas más recientes, como B-splines o wavelets, han demostrado ser igualmente eficaces e incluso superiores en ciertos contextos.

La base B-spline fue propuesta por de Boor en 1977 (de Boor, 1977) y se basa en una serie de bases tipo  $\phi_k(t)$  que deben cumplir:

- Cada una de las  $\phi_k(t)$  es una función spline definida por un orden  $m$  y una sucesión de nodos  $T$ .
- Cualquier combinación lineal que se realice con esta serie de funciones base  $\phi_k(t)$  será un spline.
- Dentro de cualquier función spline definida por su orden  $m$  con nodos  $\tau$ , esta puede expresarse como una combinación lineal de este tipo de funciones.

La base de B-splines se utiliza principalmente en datos no periódicos cuya función es suave. Los coeficientes se pueden calcular de manera rápida con una aproximación muy precisa, y esto puede lograrse incluso con un número pequeño  $K$  de funciones.

### 3.2.3 Regresión funcional

La regresión funcional es una extensión de la regresión clásica en la cual las repuestas y/o los predictores son datos funcionales. Los modelos de regresión funcional se clasifican, habitualmente, en tres tipos: a) respuesta y predictores funcionales, b) respuesta escalar con predictores funcionales y c) respuesta funcional con predictores escalares. En el presente estudio, dadas las características de la respuesta y de los predictores, es de aplicación el último tipo, regresión que también es conocida como ‘Function-on-scalar regression’ o regresión con respuesta funcional.

Dados los predictores escalares  $x_{ij}$  para  $i = 1, \dots, N, j = 1, \dots, J$  y la respuesta funcional  $y_i(t)$ , el modelo de regresión lineal con respuesta funcional viene dado por:

$$y_i(t) = \beta_0(t) + \sum_{j=1}^J x_{ij} \beta_j(t) + \varepsilon_i(t), \quad i = 1, \dots, N \quad [8]$$

donde  $\beta_0(t)$  es la función intercepto,  $\beta_j(t)$  es la función de coeficientes de regresión para el predictor  $j$  y  $\varepsilon_i(\cdot)$  es la función error de media cero y covarianza  $\Sigma(s, t) = \text{cov}(\varepsilon_i(s), \varepsilon_i(t))$ .

Alternativamente, el modelo puede expresarse en forma matricial como:

$$\mathbf{y}(t) = \mathbf{Z}(t)\boldsymbol{\beta}(t) + \boldsymbol{\varepsilon}(t) \quad [9]$$

La función de coeficientes  $\beta_j(t)$  se puede expandir como

$$\beta_j(t) = \sum_{l=1}^{\infty} \beta_{jl} \phi_l(t) \quad [10]$$

para  $j = 1, \dots, J$  siendo  $\{\phi_l(\cdot)\}_{l=1}^{\infty}$  un conjunto de funciones base preespecificadas. Truncando al orden  $m$ , el modelo de regresión de respuesta funcional dado por Ecuación (9) puede ser expresado como:



$$y_i(t) = \sum_{l=1}^m b_{0l} \phi_l(t) + \sum_{j=1}^J \sum_{l=1}^m x_{ij} b_{jl} \phi_l(t) + T_i(t) + \varepsilon_i(t) \quad [11]$$

siendo  $T_i(t)$  el error de truncación. Habitualmente, se utilizan B-splines para aproximar las funciones de coeficientes variables.

Si los datos se recopilan en los instantes de tiempo  $\{t_{ik}\}_{k=1}^{n_i}$ , el problema de estimación de las funciones  $\beta_j(t)$  se reduce a la estimación de  $\mathbf{B}=(b_{jl})_{(j+1) \times m}$ , siendo el objetivo minimizar

$$\frac{1}{n} \sum_{i=1}^n \sum_{k=1}^{n_i} \left[ y_i(t_{ik}) - \sum_{l=1}^m b_{0l} \phi_l(t) - \sum_{j=1}^J \sum_{l=1}^m x_{ij} b_{jl} \phi_l(t_{ik}) \right]^2 \quad [12]$$

con alguna penalización en los coeficientes  $\mathbf{B}$ , con la finalidad de evitar un sobreajuste en la estimación. Este es el enfoque que se adopta, por ejemplo, en (Ramsay & Silverman, 2005) o en (Reiss, Huang, & Mennes, 2010) que presentan un método automático para la estimación de los coeficientes, teniendo en cuenta errores correlacionados y utilizando mínimos cuadrados generalizados.

Las librerías r “refund” (Crainiceanu, y otros, 2024) y “fda” (Ramsay, Wickham, Graves, & Hooker, 2024) tienen implementadas diferentes funciones en el contexto de regresión con respuesta funcional.

### 3.2.4 Selección de variables predictoras

La propagación de la incertidumbre a través de un metamodelo que sustituya el código computacional requiere, habitualmente, el ajuste de modelos con un elevado número de variables predictoras y tamaño de muestra reducido. Por ello, resulta de interés el uso de técnicas que permitan realizar una selección de las variables predictoras más importantes. Entre las diferentes técnicas de selección de variables se ha seleccionado la regularización.

El modelo de selección 'grMCP' (Minimax Concave Penalty) (Zhang, 2010) es una técnica de regularización utilizada en la regresión funcional para la selección de variables. Esta técnica destaca por su capacidad para manejar variables predictoras dispersas y variables que cambian con el tiempo. A diferencia de otros modelos de regularización como el Lasso, que tiende a sobreajustar los coeficientes de las variables seleccionadas, 'grMCP' mantiene un sesgo mínimo mientras selecciona eficazmente las variables relevantes, lo cual hace que se trate de un método particularmente útil en situaciones donde se desea una selección de variables precisa y robusta sin introducir un sesgo excesivo en la estimación.

El estimador MCP se obtiene a través de la siguiente ecuación (Choon, 2012):

$$\hat{\beta}_j^{MCP} = \arg \min_p [ \|Y - X\beta\|^2 + \sum_{j=1}^p P_{\lambda,\gamma}^{MCP} ] \quad [ 13 ]$$

donde  $\sum_{j=1}^p P_{\lambda,\gamma}^{MCP}$  es la función de penalización MCP. Dicha función toma la siguiente forma dependiendo del valor de  $\beta$ :

$$p_{\lambda}(|\beta|) = \begin{cases} \lambda \left( |\beta| - \frac{|\beta|^2}{2\lambda\gamma} \right), & |\beta| < \lambda\gamma \\ \frac{\lambda^2\gamma}{2}, & |\beta| \geq \lambda\gamma \end{cases} \quad [ 14 ]$$

Donde  $\gamma > 1$  (Huang, Breheny, Lee, Ma, & Zhang, 2016)

### 3.2.5 Método de corrección de error

Un problema en el uso de metamodelos, en el contexto BEPU, es que predicen la variable dependiente en la media de la distribución, por lo que las predicciones que se obtienen no son conservadoras. En la literatura se han propuesto diferentes estrategias para la obtención de predicciones conservadoras. En este trabajo se utiliza la distribución del error de predicción, diferencia entre el valor real y el valor predicho, para garantizar el grado de conservadurismo requerido.

En el caso de la regresión funcional, en cada instante de tiempo, el error de predicción se distribuye de cierta manera. Analizar esta distribución ayuda a entender la magnitud y la variabilidad del error. Dos estadísticos útiles son:

- **Media del error:** Indica si hay un sesgo sistemático en las predicciones.
- **Percentil 95 del error:** Representa el valor por debajo del cual se encuentran el 95% de los errores, proporcionando información sobre los errores extremos.

El objetivo final es ajustar las predicciones para que sean conservadoras, es decir, para evitar subestimar el valor real de la variable de seguridad analizada. Esto es particularmente importante en contextos donde la subestimación puede tener consecuencias graves (por ejemplo, la temperatura en un reactor nuclear).

El **percentil 95 de la distribución del error** se utiliza para hacer una corrección de la predicción con el objetivo de obtener una predicción conservadora.

De forma que, el proceso de corrección sería el siguiente:

1. **Cálculo de los errores de predicción:** Para cada instante de tiempo, se calcula la diferencia entre los valores observados,  $y_K(t)$ , y los valores predichos,  $y'_K(t)$ :

$$e_K(t) = y_K(t) - y'_K(t) \quad [ 15 ]$$

siendo  $K$  el número de simulaciones del transitorio consideradas.

2. **Obtención de la distribución del error:** A partir de estos errores, se calcula la media y el percentil 95 para cada instante de tiempo. La media del error permite analizar el sesgo, mientras que el percentil 95 permite corregir el error en la predicción.

$$\text{Media del error } (t) = \frac{\sum_{k=1}^K e_k(t)}{K} \quad [16]$$

$$\text{Percentil 95} = \frac{(95 \cdot K)}{100} \quad [17]$$

3. **Ajuste de las predicciones:** Las predicciones obtenidas con el modelo de regresión se ajustan sumando el percentil 95 de la distribución del error,  $e_{95}$ , en cada instante de tiempo:

$$y''(t) = y'(t) + e_{95}(t) \quad [18]$$

### 3.2.6 Intervalos de tolerancia

En el marco de la metodología BEPU se encuentra extendido el uso del método no paramétrico de Wilks para la estimación de intervalos o límites de tolerancia en el análisis de incertidumbres.

Un intervalo de tolerancia (Krishnamoorthy & Mathew, 2009), se define como un rango, derivado de una muestra aleatoria, que se espera que contenga una proporción específica de la población muestreada con una determinada confianza. Específicamente en el contexto de análisis de seguridad mediante códigos termohidráulicos, se utiliza el intervalo unilateral de tolerancia 95/95.

Considerando una variable aleatoria continua  $X$  con función de distribución acumulada (CDF)  $F(x) = P(X \leq x)$ , para un  $\gamma (0 < \gamma < 1)$  especificado, la inversa de la CDF se define como:

$$F^{-1}(\gamma) = \mathbf{inf}\{x: F(x) \geq \gamma\} \quad [19]$$

Siendo  $F^{-1}(\gamma)$  el valor de  $x$  para el cual  $F(x) = P(X \leq x) = \gamma$ .

Tomando una muestra aleatoria  $X = (X_1, X_2, \dots, X_n)$  de la población, para definir un intervalo de tolerancia se requiere especificar un nivel de cobertura  $\gamma$  y un nivel de confianza  $\beta$ . En aplicaciones prácticas,  $\gamma$  y  $\beta$  toman valores en el conjunto  $\{0.90, 0.95, 0.99\}$ . El intervalo se construirá utilizando una muestra aleatoria  $X$  y debe contener al menos un  $\gamma\%$  de la población o más con un nivel de confianza  $\beta$ . Formalmente, un intervalo de tolerancia unilateral  $(\gamma, \beta)$  de la forma  $(-\infty, U(X))$  satisface la condición:

$$P\{P(X \leq U(X) | X) \geq \gamma\} = \beta \quad [20]$$

Esto implica que  $U(X)$  se determina de manera que al menos un  $\gamma\%$  de la población sea menor o igual a  $U(X)$  con un nivel de confianza  $\beta$ . El intervalo  $(-\infty, U(X))$  se denomina intervalo de tolerancia unilateral y  $U(X)$  es el límite superior de tolerancia unilateral.

La ecuación (14) puede expresarse como:

$$P\{q_\gamma < U(X)\} = \beta \quad [21]$$

Siendo  $q_\gamma$  el percentil  $\gamma$ .

El método no paramétrico más extendido para la estimación de intervalos de tolerancia en análisis BEPU es, como se ha indicado anteriormente, el método de Wilks caracterizado por la utilización de muestreo aleatorio simple y el uso de estadísticos de orden como límites de tolerancia.

Dada una variable aleatoria escalar y continua,  $X$ , con función de distribución (CDF) continua, y una muestra aleatoria simple de tamaño  $n$  de valores de  $X$ . La llamada *muestra ordenada* se obtiene colocando los elementos muestrales de menor a mayor

$$X_{1:n} < X_{2:n} < \dots < X_{r:n} < \dots < X_{n:n} \quad [22]$$

de manera que el estadístico de orden  $r$  es el  $r$ -ésimo valor más alto de la muestra. Por ejemplo, el estadístico de orden  $n$  es el máximo muestral, y el estadístico de orden 1 es el mínimo muestral.

El resultado básico obtenido por Wilks es que la cobertura del intervalo comprendido entre los estadísticos de orden  $r$  y  $s$ , con  $r < s$ ,

$$\mu_X(X_{r:n}, X_{s:n}) = PR_X\{X_{r:n} < X < X_{s:n}\} \equiv F_X(X_{s:n}) - F_X(X_{r:n}) \quad [23]$$

es, a su vez, un estadístico que sigue la distribución beta con parámetros  $s - r$  y  $n - s + r + 1$ , sea cual sea la distribución de  $X$ . En la ecuación anterior se considera que  $X$  es independiente de los dos estadísticos de orden; es decir, representa una “nueva extracción” al azar de  $X$ , independiente de la muestra aleatoria simple de la que se toman los estadísticos de orden.

Un corolario del resultado de Wilks es que el estadístico

$$W_{r:n} \equiv F_X(X_{r:n}) \quad [24]$$

Para  $r=1, \dots, n$  sigue la distribución beta  $(r, n-r+1)$ .

El resultado obtenido por Wilks permite utilizar los estadísticos de orden como límites de tolerancia, para intervalos de tolerancia uni y bilaterales.

La condición para que  $(X_{r:n}, X_{s:n})$  sea un intervalo bilateral de tolerancia de nivel  $(\beta, \gamma)$  para  $X$  es

$$PR_S\{PR_X\{X_{r:n} < X < X_{s:n}\} \geq \beta\} \geq \gamma \quad [25]$$

Teniendo en cuenta el resultado de Wilks, se convierte en

$$PR_S\{B_{s-r, n-s+r+1} \geq \beta\} \geq \gamma \quad [26]$$

Donde  $B_{l,m}$  representa la variable beta con parámetros l y m. La ecuación anterior es la llamada fórmula de Wilks. Es una desigualdad que relaciona los órdenes r y s, el tamaño de muestra n y el nivel de tolerancia  $(\beta, \gamma)$ .

La fórmula de Wilks se puede expresar en términos de los cuantiles de la distribución beta

$$B_{s-r, n-s+r+1}^{1-\gamma} \geq \beta \quad [27]$$

Fijado el nivel de tolerancia y los órdenes r y s, estos últimos posiblemente como funciones de n, existe un mínimo valor entero de n que satisface, al que llamaremos “tamaño mínimo de muestra”:

$$N_m(\beta, \gamma, s, r) \equiv \text{Min}\{n \in Z^+ / B_{s-r, n-s+r+1}^{1-\gamma} \geq \beta\} \quad [28]$$

Esta ecuación se puede minimizar frente a la elección de los órdenes r y s. Ese mínimo de mínimos (MM), dependiente sólo del nivel de tolerancia, se obtiene cuando  $r=1, s=n$ ; es decir, cuando se utilizan los extremos muestrales para construir el intervalo. Entonces,

$$PR_S\{B_{s-r, n-s+r+1} \geq \beta\} \geq \gamma \quad [29]$$

se reduce a

$$PR_S\{B_{n-1,2} \geq \beta\} \geq \gamma \quad [30]$$

y definimos:

$$N_{mm}(\beta, \gamma) \equiv \text{Min}\{n \in Z^+ / B_{n-1,2}^{1-\gamma} \geq \beta\} \quad [31]$$

La fórmula de Wilks puede particularizarse al caso de intervalos unilaterales de tolerancia. La condición para que el estadístico de orden s sea un límite superior de tolerancia  $(\beta, \gamma)$  de X es:

$$PR_S\{PR_X\{X \leq X_{s:n}\} \geq \beta\} \geq \gamma \quad [32]$$

que se puede escribir

$$PR_S\{F_X(X_{s:n}) \geq \beta\} \geq \gamma \quad [33]$$

es decir,

$$PR_S\{F_X(B_{s,n-s+1}) \geq \beta\} \geq \gamma \quad [34]$$

Dicha fórmula de Wilks es la correspondiente al caso de un límite superior de tolerancia. Coincide con  $PR_S\{B_{s-r, n-s+r+1} \geq \beta\} \geq \gamma$  cuando  $r=0$ . Fijado el nivel de tolerancia y el orden s (quizá como función de n), se puede buscar el valor mínimo de n que cumple

dicha ecuación; recíprocamente, fijado  $n$ , puede buscarse el mínimo valor de  $s$  que satisface la expresión. En términos de los cuantiles de la distribución beta, la fórmula

$$PR_S\{PR_X\{X \leq X_{s:n}\} \geq \beta\} \geq \gamma \quad [35]$$

se escribe

$$B_{s, n-s+1}^{1-\gamma} \geq \beta \quad [36]$$

Y el tamaño de muestra es

$$N_m(\beta, \gamma, s, 0) \equiv \text{Min}\{n \in Z^+ / B_{s, n-s+1}^{1-\gamma} \geq \beta\} \quad [37]$$

El tamaño de dicha ecuación se puede minimizar posteriormente frente a  $s$ . El mínimo de mínimos se obtiene para  $s = n$ ; es decir, cuando se utiliza el máximo muestral como LST. En este caso, ésta asume una forma muy simple:

$$1 - \beta^n \geq \gamma \quad [38]$$

y de aquí se deduce el tamaño mínimo de muestra

$$N_{mm} = \left\lceil \frac{\log(1-\gamma)}{\log \beta} \right\rceil \quad [39]$$

El paréntesis parcial designa la función techo (redondeo hacia arriba).

Es fácil comprobar que el tamaño mínimo de muestra para intervalo bilateral, obtenido de la ecuación

$$N_m(\beta, \gamma, s, r) \equiv \text{Min}\{n \in Z^+ / B_{s-r, n-s+r+1}^{1-\gamma} \geq \beta\} \quad [40]$$

, es mayor que el de intervalo unilateral (para un mismo nivel de tolerancia) derivado de

$$N_m(\beta, \gamma, s, 0) \equiv \text{Min}\{n \in Z^+ / B_{s, n-s+1}^{1-\gamma} \geq \beta\} \quad [41]$$

Cualquier enfoque utilizado para establecer límites superiores de tolerancia, como el método de Wilks y otros, puede igualmente aplicarse para determinar límites inferiores. Es importante tener en consideración que un límite inferior de tolerancia (LIT) para una variable  $V$  se obtiene simplemente invirtiendo el signo de un límite superior de tolerancia (LST) correspondiente, manteniendo el mismo nivel de tolerancia, pero aplicando la operación  $(-V)$ .

El método de Wilks ha generado interés en la incorporación de estadísticos de orden en la construcción de límites de tolerancia no paramétricos. La estimación de estos límites de tolerancia mediante el método de Wilks 95/95 se realizan para los estadísticos de primer, segundo, tercer y cuarto orden, cuyos tamaños de muestra para intervalos de tolerancia bilaterales se muestran en la Tabla 2.

*Tabla 2. Tamaños de muestra (n) para los diferentes estadísticos intervalo bilateral de orden en el método de Wilks.*

| <b>Estadístico de orden</b> | <b>n</b> |
|-----------------------------|----------|
| <b>1</b>                    | 93       |
| <b>2</b>                    | 124      |
| <b>3</b>                    | 153      |
| <b>4</b>                    | 181      |

#### 4. Resultados caso de aplicación

En este apartado, se presentan los resultados de los diferentes métodos utilizados para la modelización de la evolución temporal de la PCT durante un accidente en un reactor nuclear de agua a presión (PWR). El accidente considerado, como se ha comentado anteriormente, es un LBLOCA.

El análisis temporal permitirá el estudio del comportamiento de la PCT en función de la activación de los diferentes sistemas de seguridad, así como la caracterización de, por ejemplo, el tiempo en que se alcanza la máxima temperatura de vaina. Asimismo, facilitará el estudio de cómo las condiciones iniciales de planta, es decir, las condiciones en el instante en que ocurre el accidente influyen en el valor de la PCT en los diferentes instantes de tiempo considerados.

En la Figura 5 se presentan las diferentes etapas del caso de aplicación. Tal como se observa en dicha figura, en primer lugar, se realiza un análisis descriptivo de las 450 simulaciones de la evolución temporal de la PCT. Posteriormente, con el objetivo de reducir el número de inputs, se ha aplicado un método de selección de variables en el contexto de regresión con respuesta funcional. Utilizando los predictores seleccionados se han obtenido modelos de regresión funcional penalizada flexible para diferentes tamaños de muestras. A partir del modelo obtenido y utilizando las predicciones en el conjunto de test se obtiene la distribución del error en cada instante de tiempo. El modelo de regresión funcional y la distribución de error se utilizan para la estimación de una banda de confianza a partir de una muestra de 10000 predictores seleccionados aleatoriamente. Finalmente, se compara la banda de confianza, obtenida con el modelo de regresión funcional, con los límites de los intervalos de tolerancia estimados mediante el método de Wilks.

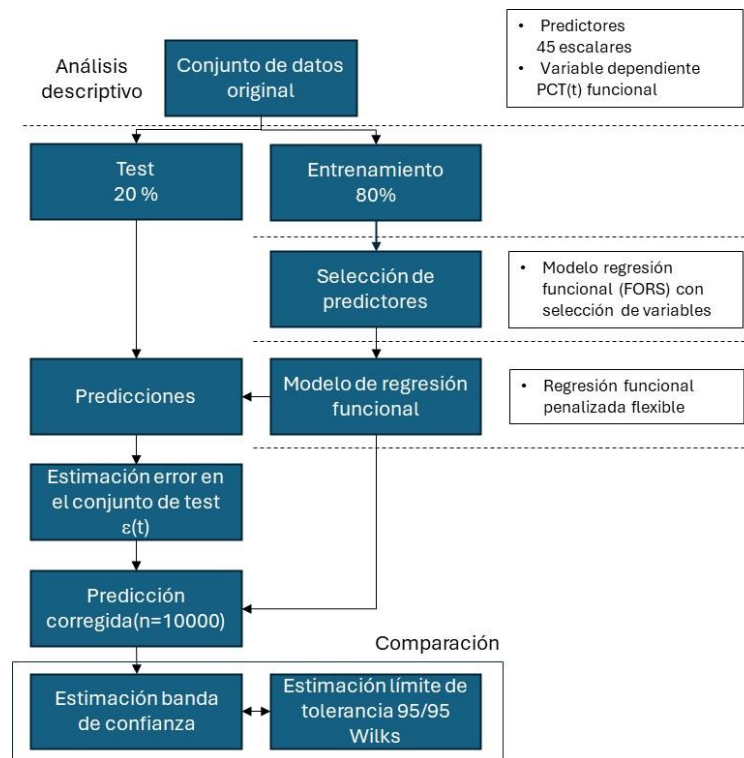


Figura 5. Etapas caso de aplicación.



A continuación, se muestran los resultados obtenidos en cada una de las etapas consideradas.

#### 4.1 Análisis descriptivo

En este apartado se realiza un análisis descriptivo de las curvas muestrales para analizar su comportamiento frente al tiempo.

En primer lugar, se ha calculado la media de todas las curvas para observar la tendencia central. Para ello, se emplea la función *mean.fda()* de la librería *fda*. En la Figura 6 se presenta su evolución. Tal y como se puede observar, durante los primeros segundos del transitorio se produce un incremento de la temperatura de la vaina, aumentando en los primeros 5 segundos alrededor de unos 300 K. Además, la temperatura media de la vaina alcanza su valor máximo a los 57 segundos del inicio del transitorio, un punto a partir del cual, se produce un descenso constante hasta el final de la simulación.

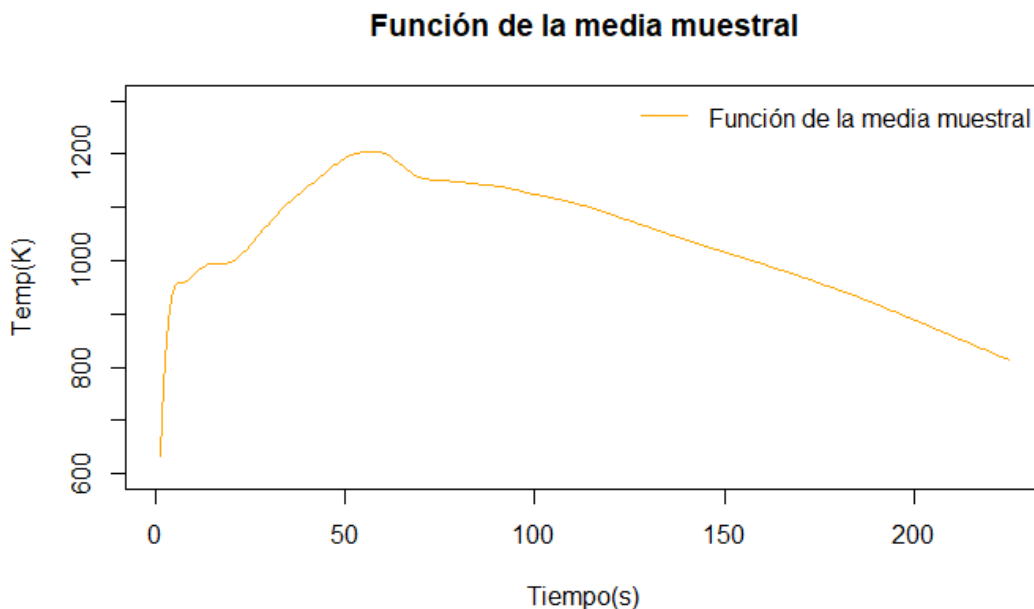
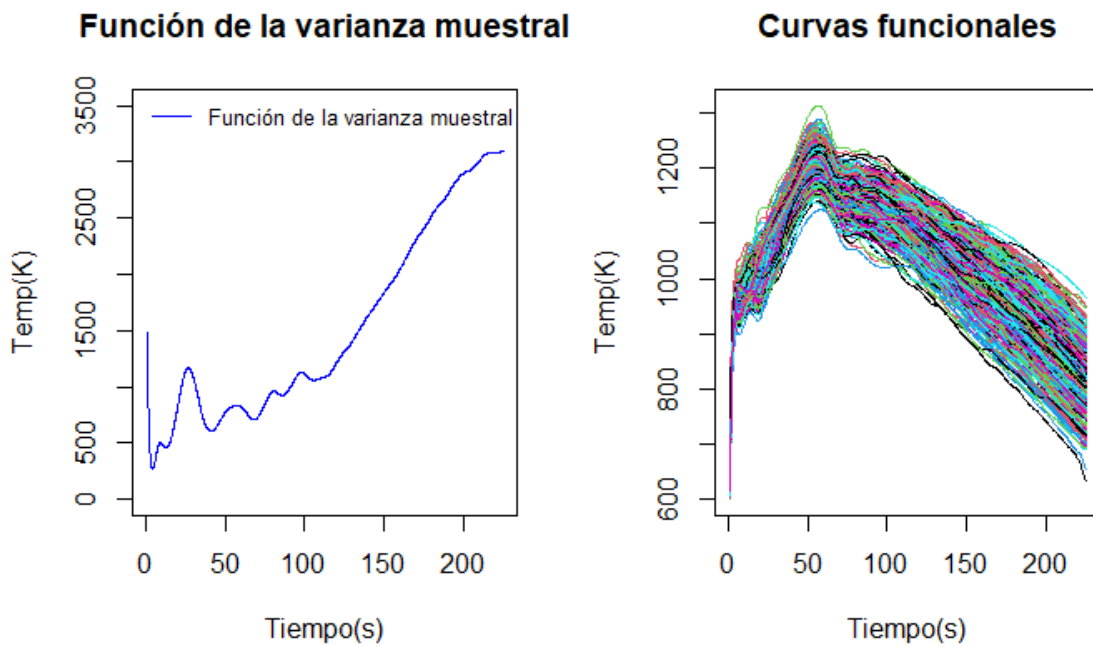


Figura 6. Función de la media muestral.

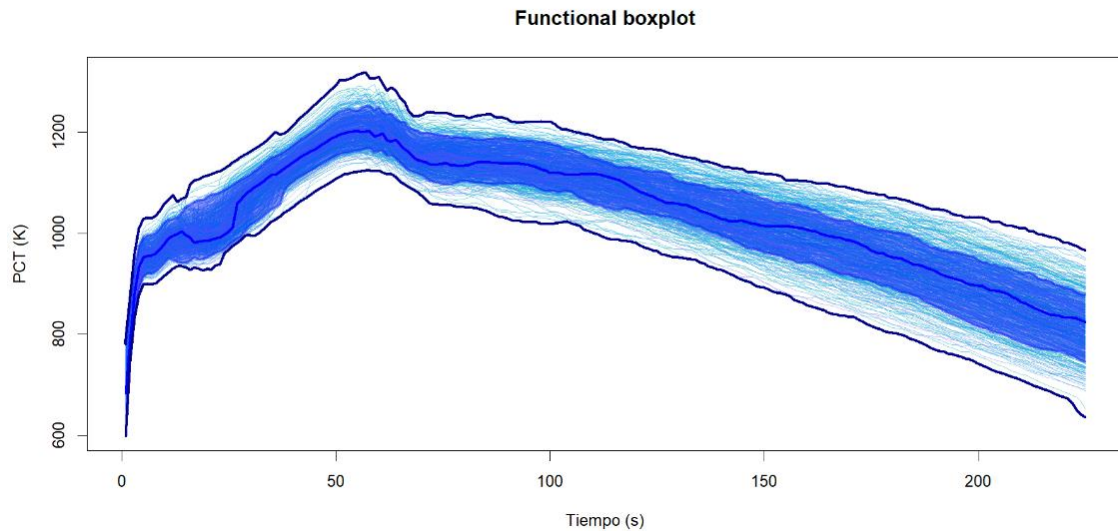
Por otro lado, en la Figura 7 se muestra la evolución temporal de la varianza de la muestra, la cual se ha calculado mediante la función *std.fda()* de la librería *fda*.



*Figura 7. Función de la varianza muestral.*

Como se observa en la figura, al inicio del transitorio la varianza comienza siendo considerablemente baja, pues al inicio del transitorio, las simulaciones parten de temperaturas similares. A continuación, tras 5 segundos de subida drástica de la temperatura, la varianza se observa que disminuye, alcanzando en este caso el valor más bajo de todo el transitorio. A partir de este momento, y tal como se observa en el gráfico de la derecha, correspondiente a las curvas funcionales de las 450 simulaciones, se observa como aumenta progresivamente la variabilidad, alcanzando el valor más alto de la varianza al final del transitorio.

A continuación, en la Figura 8 se muestra el box-plot funcional correspondiente a las 450 simulaciones del transitorio que permite, entre otros, la detección de curvas atípicas. El box-plot funcional es una herramienta exploratoria informativa. Las estadísticas descriptivas del diagrama de box-plot funcional son: la envolvente de la región central del 50% que es análoga al rango intercuartílico y proporciona información relativa a la dispersión del 50% central de las curvas, la curva mediana y la envolvente máxima del conjunto de datos (excepto los valores atípicos). Además, permite la detección de valores atípicos mediante la regla de 1.5 veces la región central del 50%, análoga a la regla de los diagramas box plot clásicos. Tal y como se observa en la figura, no se observan curvas atípicas en el conjunto de simulaciones.



*Figura 8. Box-plot funcional.*

#### 4.2 Selección de variables según los tamaños de muestra

En este apartado se lleva a cabo la selección de las variables de entrada utilizando la base de datos de la que se dispone, con un tamaño de muestra de 450 simulaciones mediante el método 'grMCP' de la librería 'refund' (Zhang, 2010).

##### 4.2.1 Resultados de la selección de variables

A continuación, en la Figura 9 se puede observar la evolución temporal de los coeficientes (sin tener en cuenta el intercepto) obtenida tras utilizar el método de penalización 'grMCP' para el tamaño de muestra considerado. En el Anexo 1 se puede consultar el código de RStudio empleado.

### Curvas de Coeficientes a lo largo del tiempo (n=450)

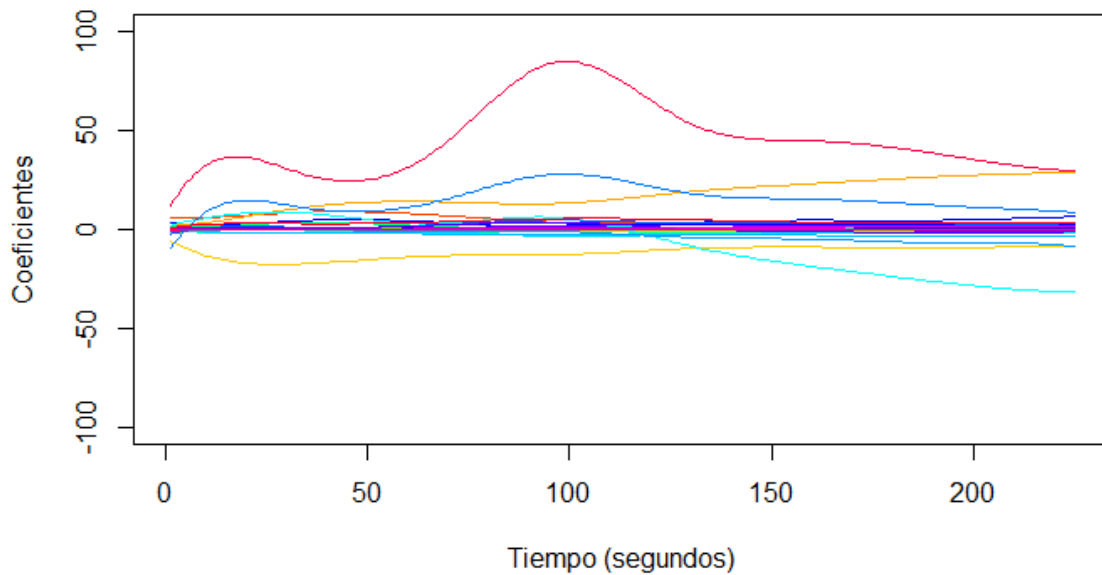


Figura 9. Evolución temporal de los coeficientes para n=450.

A la vista de la evolución temporal, se puede observar claramente como una de las variables presenta un coeficiente positivo relativamente superior al resto, por lo que su efecto positivo será el más relevante. Se trata de la variable WH002 (coeficiente de velocidad de reacción metal-agua del revestimiento). Por otro lado, siguiendo con los coeficientes positivos, se aprecian también tanto la curva de color azul como la curva de color naranja, correspondientes a las variables RPOWM (multiplicador de potencia residual) y TMIFS (criterio de temperatura mínima para la transición de ebullición de película a ebullición de transición), respectivamente. Esta última es la etapa entre la ebullición de película (donde la superficie está cubierta por una película de vapor) y la ebullición nucleada (donde se forman burbujas en la superficie)).

Y, por el contrario, se observa una variable que presenta un efecto negativo sobre la PCT, y no es otra que la curva de color amarillo, correspondiente a la variable UO2TC (conductividad térmica del UO2).

Sin embargo, en vista del modelo base y de la representación gráfica de las 450 curvas, la selección de las variables se realiza en base al momento crítico del transitorio el cual corresponde al punto donde se alcanza la temperatura máxima de la vaina (alrededor de los 50-60 segundos de transitorio). Este punto crítico se considera fundamental ya que representa el escenario más determinante para la seguridad del reactor.

Se seleccionan un total de 27 variables significativas, las cuales se muestran en la Tabla 3 en orden descendente según su impacto sobre la variable respuesta, independientemente de su efecto positivo o negativo. Para ello, se seleccionó el punto crítico mencionado, y se calcularon los valores de los coeficientes correspondientes a cada variable, en valor absoluto, observando así el grado de impacto de cada una de ellas sobre la variable respuesta, independientemente de si su efecto era positivo o negativo, lo cual se verá posteriormente.

Tabla 3. Variables seleccionadas para el modelo de regresión funcional.

| Variables seleccionadas |              |   |
|-------------------------|--------------|---|
| X44                     | <b>WH002</b> | Cladding metal-water reaction rate coeff.             |
| X6                      | <b>UO2TC</b> | UO2 Thermal Conductivity                              |
| X5                      | <b>RPOWM</b> | Residual power multiplier                             |
| X27                     | <b>TMIFS</b> | Film to transition boiling Tmin criterion             |
| X2                      | <b>PEAKF</b> | Peaking factor  |
| X1                      | <b>INPOW</b> | Initial thermal power                                 |
| X23                     | <b>CONTP</b> | Maximum containment pressure                          |
| X45                     | <b>WH004</b> | Burst temperature coeff.                              |
| X19                     | <b>CHFC2</b> | Wall friction multiplier                              |
| X25                     | <b>CONVV</b> | Single phase vapor to wall heat transfer coeff.       |
| X11                     | <b>MASFR</b> | Initial total primary mass flow rate                  |
| X24                     | <b>CONVL</b> | Single phase liquid to wall heat transfer coeff.      |
| X21                     | <b>CLCON</b> | Clad thermal conductivity                             |
| X33                     | <b>CCFLC</b> | Wallis c for counter current flow limitation          |
| X28                     | <b>CRIHF</b> | Critical Heat Flux (CHF) multiplier                   |
| X26                     | <b>FILMB</b> | Dispersed flow film boiling heat transfer coeff.      |
| X4                      | <b>HGSHR</b> | Hot gap size hot rod                                  |
| X13                     | <b>ACCKL</b> | Accumulator loss coeff.                               |
| X14                     | <b>KSURG</b> | Surge line coss coeff.                                |
| X22                     | <b>CLDEN</b> | Clad density  |
| X12                     | <b>CLTEM</b> | Initial cold legs average temperature                 |
| X7                      | <b>UO2SH</b> | UO <sub>2</sub> Specific Heat                         |
| X31                     | <b>DRAG5</b> | Interfacial drag (dispersed flow film boiling) coeff. |
| X30                     | <b>DRAG4</b> | Interfacial drag (droplet) coeff.                     |
| X37                     | <b>LINT2</b> | Liquid to interface annular-mist heat transfer coeff. |
| X36                     | <b>UHTEM</b> | Initial upper head temperature                        |
| X9                      | <b>ACCLT</b> | Accumulator liquid temperature                        |

Desde el punto de vista de la seguridad el punto crítico en la evolución del transitorio es el máximo valor que se alcanza en la PCT. Por este motivo, resulta de interés comparar las variables seleccionadas con el modelo de regresión funcional con las seleccionadas en un estudio previo cuyo objetivo era la selección y clasificación de las variables predictoras considerando únicamente la máxima PCT. En la Tabla 4 se muestra una comparación de

las variables seleccionadas en ambos estudios. Como se observa en dicha tabla de las 27 variables seleccionadas con el modelo de regresión 20 coinciden con las variables seleccionadas en el estudio previo. Existe un conjunto de variables que fueron seleccionadas en el estudio previo y que, sin embargo, no son seleccionadas en el modelo de regresión funcional. No obstante, estas variables no se encuentran entre las que fueron identificadas como más importantes en dicho estudio.

Tabla 4. Comparación de las variables seleccionadas con el modelo de regresión funcional y un estudio previo.

| Parámetros   |   | Regresión funcional | Estudio Previo |
|--------------|---|---------------------|----------------|
| <b>INPOW</b> | Initial thermal power                                 | •                   | •              |
| <b>PEAKF</b> | Peaking factor  | •                   | •              |
| <b>HGSAV</b> | Hot gap size average                                  |                     |                |
| <b>HGSHR</b> | Hot gap size hot rod                                  | •                   | •              |
| <b>RPOWM</b> | Residual power multiplier                             | •                   | •              |
| <b>UO2TC</b> | UO <sub>2</sub> Thermal Conductivity                  | •                   | •              |
| <b>UO2SH</b> | UO <sub>2</sub> Specific Heat                         | •                   | •              |
| <b>ACCPR</b> | Accumulator pressure setpoint                         | •                   | •              |
| <b>ACCLT</b> | Accumulator liquid temperature                        | •                   | •              |
| <b>LPISQ</b> | LPI mass flow rate                                    | •                   |                |
| <b>MASFR</b> | Initial total primary mass flow rate                  | •                   |                |
| <b>CLTEM</b> | Initial cold legs average temperature                 | •                   |                |
| <b>ACCKL</b> | Accumulator loss coeff.                               |                     | •              |
| <b>KSURG</b> | Surge line loss coeff.                                | •                   | •              |
| <b>KDWNC</b> | Downcomer loss coeff.                                 | •                   | •              |
| <b>KLCPL</b> | Low core plate loss coeff.                            |                     |                |
| <b>KBRCL</b> | Break cold leg loss coeff.                            | •                   |                |
| <b>CHFC1</b> | Vessel liquid wall friction multiplier                |                     |                |
| <b>CHFC2</b> | Wall friction multiplier                              | •                   | •              |
| <b>GTHCO</b> | Gap Conductance coeff.                                |                     |                |
| <b>CLCON</b> | Clad thermal conductivity                             |                     |                |
| <b>CLDEN</b> | Clad density  | •                   | •              |
| <b>CONTP</b> | Maximum containment pressure                          | •                   |                |
| <b>CONVL</b> | Single phase liquid to wall heat transfer coeff.      |                     |                |
| <b>CONVV</b> | Single phase vapor to wall heat transfer coeff.       | •                   | •              |
| <b>FILMB</b> | Dispersed flow film boiling heat transfer coeff.      | •                   | •              |
| <b>TMIFS</b> | Film to transition boiling T <sub>min</sub> criterion | •                   | •              |
| <b>CRIHF</b> | Critical Heat Flux (CHF) multiplier                   | •                   | •              |
| <b>FLOSS</b> | Form loss coeff.                                      |                     |                |
| <b>DRAG4</b> | Interfacial drag (droplet) coeff.                     | •                   | •              |
| <b>DRAG5</b> | Interfacial drag (dispersed flow film boiling) coeff. |                     |                |
| <b>DRAG7</b> | Interfacial drag (inverted annular flow) coeff.       |                     | •              |
| <b>CCFLC</b> | Wallis c for counter current flow limitation          |                     | •              |
| <b>ACCLV</b> | Accumulator liquid volume                             |                     | •              |
| <b>LIQIA</b> | Liquid to wall inverted annular heat transfer coeff.  |                     | •              |
| <b>UHTEM</b> | Initial upper head temperature                        | •                   |                |
| <b>LINT2</b> | Liquid to interface annular-mist heat transfer coeff. |                     |                |

| Parámetros   |   | Regresión funcional | Estudio Previo |
|--------------|---|---------------------|----------------|
| <b>LINT3</b> | Liquid to interface transition heat transfer coeff. | •                   | •              |
| <b>VINT3</b> | Vapor to interface transition heat transfer coeff.  | •                   | •              |
| <b>VINT4</b> | Vapor to interface stratified heat transfer coeff.  |                     |                |
| <b>SBHTC</b> | Subcooled boiling heat transfer coeff.              |                     | •              |
| <b>DNBHC</b> | Departure from nucleate boiling/CHF                 |                     | •              |
| <b>WH001</b> | Fuel thermal conductivity before burst coeff.       |                     |                |
| <b>WH002</b> | Cladding metal-water reaction rate coeff.           | •                   | •              |
| <b>WH004</b> | Burst temperature coeff.                            | •                   |                |

### 4.3 Modelo de regresión funcional

Una vez realizada la selección de variables se entrena un modelo de regresión funcional utilizando únicamente las variables predictoras seleccionadas. Utilizando la librería ‘refund’ se han obtenido modelos de regresión con respuesta funcional y predictores escalares que posteriormente permitirán realizar predicciones sobre la evolución temporal de la temperatura de la vaina en función de los 27 predictores seleccionados en la etapa anterior.

Las 450 simulaciones del código termohidráulico se han dividido en un 80% para el entrenamiento del modelo y un 20% para su validación. El modelo obtenido explica un 97.8% de la variabilidad presente en los datos en el conjunto de entrenamiento, reduciéndose a un 80 % en el conjunto de test.

En la Figura 10 se presenta la evolución temporal del  $R^2$  en el conjunto de test. Como se observa en la figura la capacidad predictiva del modelo se incrementa tras los primeros segundos del transitorio siendo superior al 80% excepto en el periodo correspondiente a aproximadamente 60-65 segundos donde la capacidad predictiva es ligeramente inferior con un valor de  $R^2$  en torno al 75%.

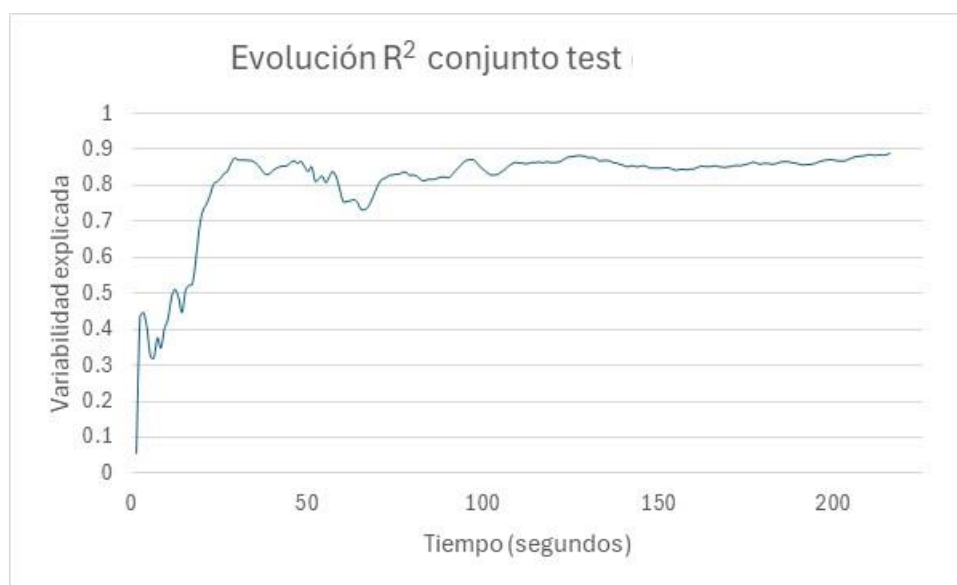
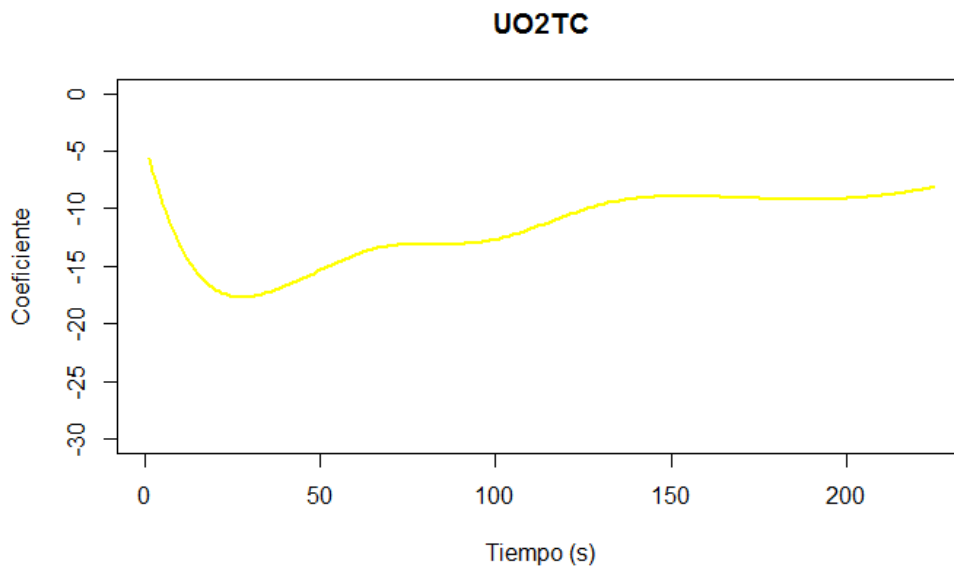
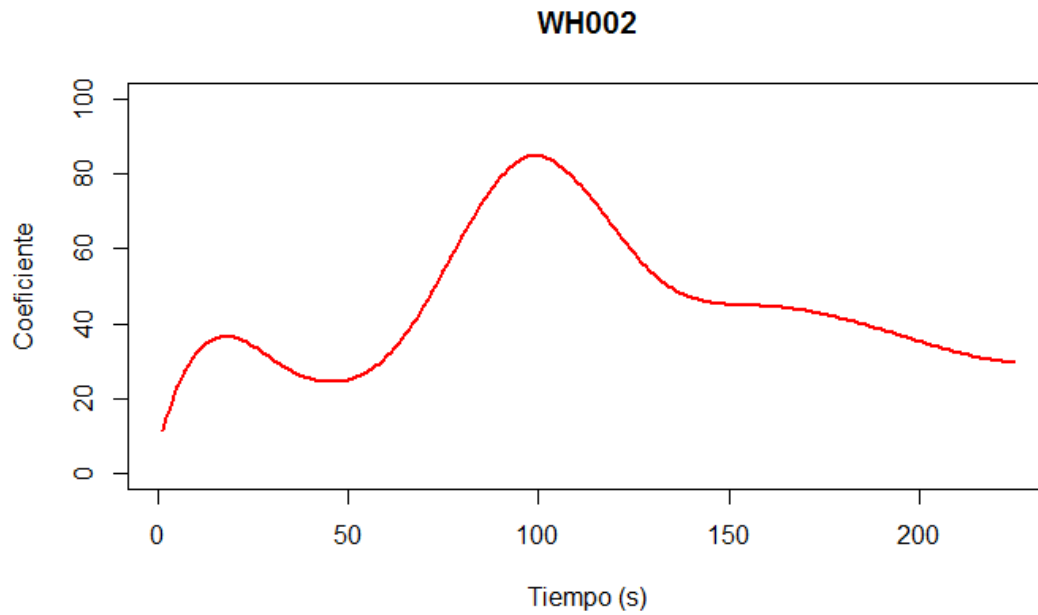
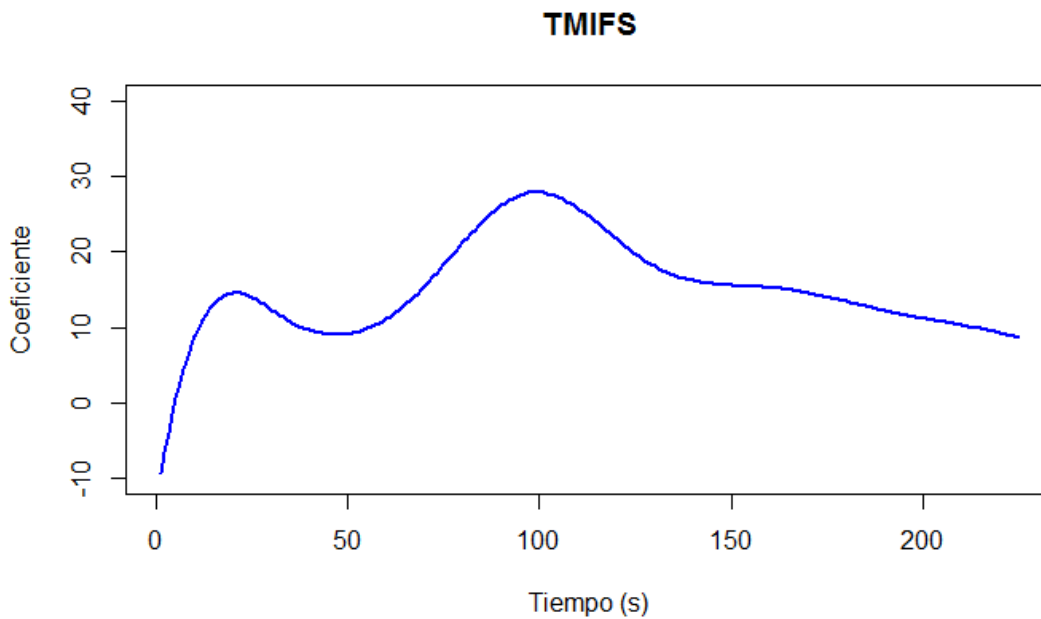
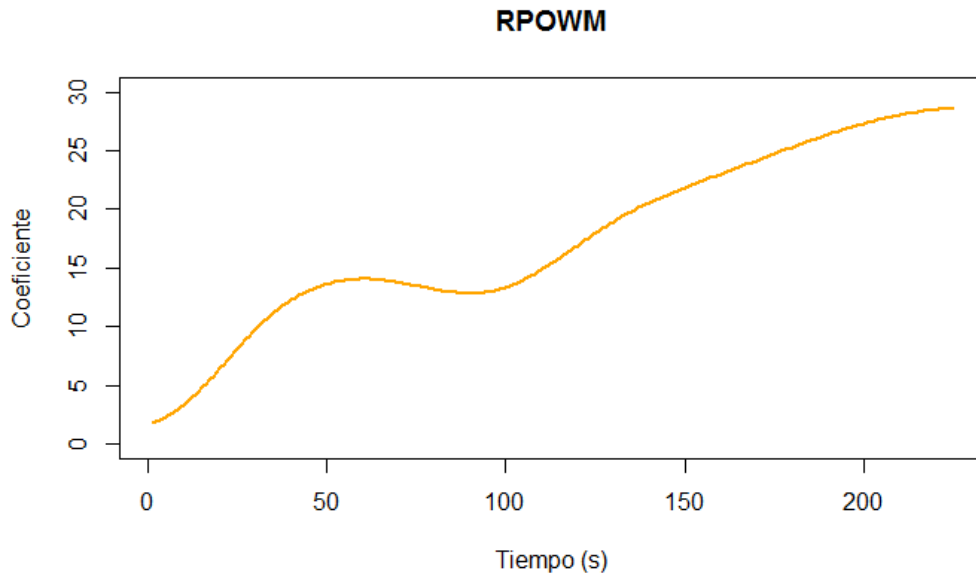


Figura 10. Evolución  $R^2$  en el conjunto de test.

Un aspecto interesante es el análisis de la evolución de los coeficientes del modelo de predicción en función del tiempo. En la Figura 11 se muestra la función de los coeficientes para los cuatros predictores de mayor importancia. En el Anexo 2 se puede consultar la evolución temporal del resto de coeficientes correspondientes a las demás variables significativas.







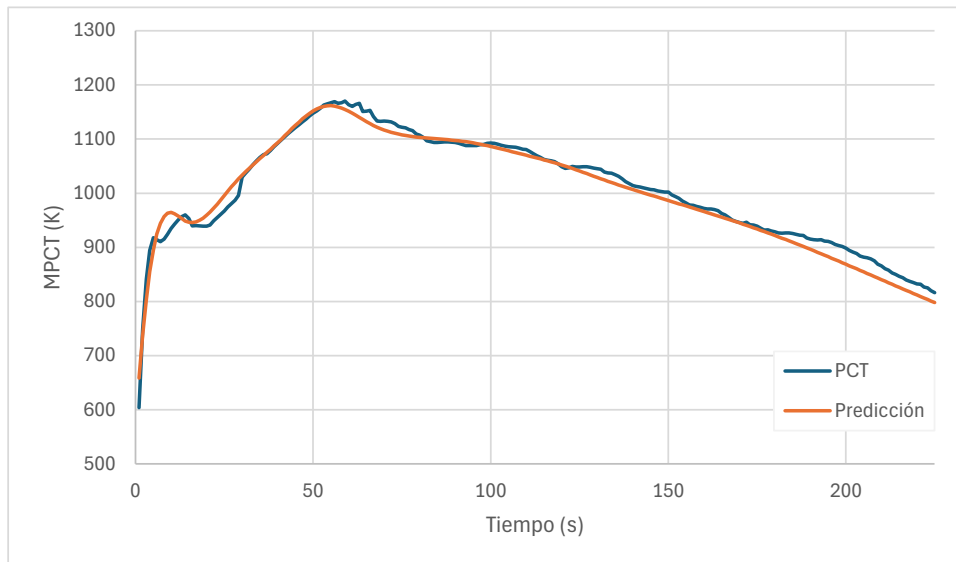
*Figura 11. Evolución de los coeficientes para los cuatro primeros inputs.*

Tal y como se observa en la Figura 11, tanto los coeficientes correspondientes a WH002 como a RPOWM se mantienen positivos durante todo el transitorio, indicando una correlación positiva entre la PCT(t) y dichos coeficientes.

Por otro lado, en el caso de TMIFS, se observan valores negativos y como a partir de los 10 segundos aproximadamente pasan a ser positivos el resto del transitorio, alcanzando el valor más alto en torno a los 100 segundos.

Sin embargo, la función del coeficiente asociado al predictor UO2TC toma valores negativos en todos los instantes temporales.

La Figura 12 muestra, para una determinada simulación, la evolución temporal de la temperatura máxima de vaina en el núcleo en función del tiempo y la predicción obtenida con el modelo de regresión funcional. En dicha figura se observa como en algunos intervalos se obtienen valores de la predicción de temperatura inferiores a los observados.



*Figura 12. Evolución temporal de la PCT frente a su predicción.*

Con el objetivo de obtener predicciones conservadoras se ha aplicado el método presentado en la sección 3.2.5. Para ello, a partir de las predicciones en el conjunto de test se ha obtenido, el percentil 95 y la media de la distribución del error para cada instante de tiempo cuya evolución temporal se muestra en la Figura 13. El percentil 95 del error se utiliza para la corrección de las estimaciones con el objetivo de obtener un modelo que sea conservador. Como se observa en la figura, transcurridos los primeros segundos del transitorio, la media del error oscila en torno a cero excepto en la parte final del transitorio en el cual se subestima el valor de la PCT.

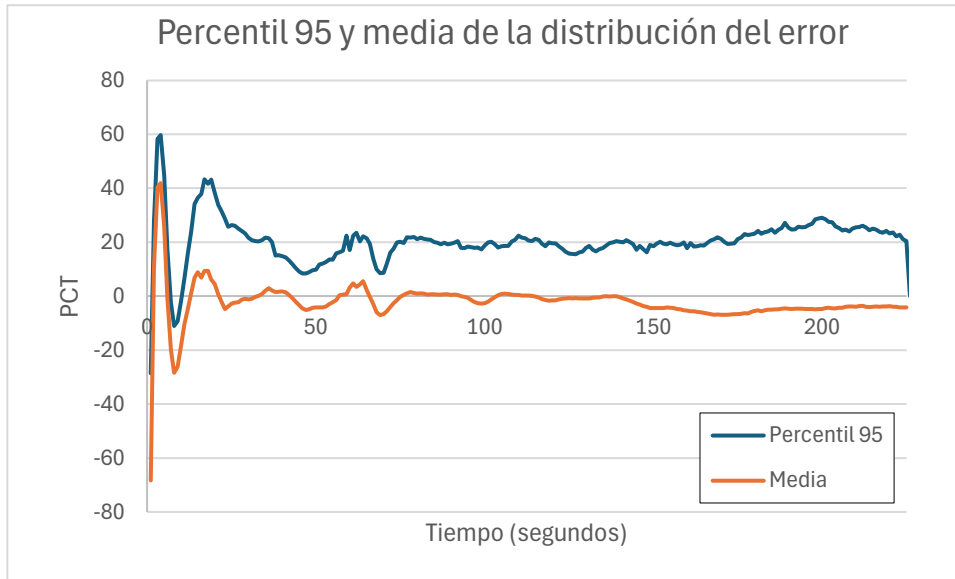


Figura 13. Evolución del percentil 95 y la media de la distribución del error.

En la Figura 14 se presenta, para una simulación del transitorio, la evolución de la PCT, la predicción obtenida con el modelo de regresión funcional y la predicción después de la corrección de error utilizando el percentil 95 pudiendo observar como las estimaciones con la corrección de error son conservadoras lo cual es especialmente relevante en el punto donde se alcanza la máxima PCT. Este comportamiento se observa en el conjunto de predicciones.

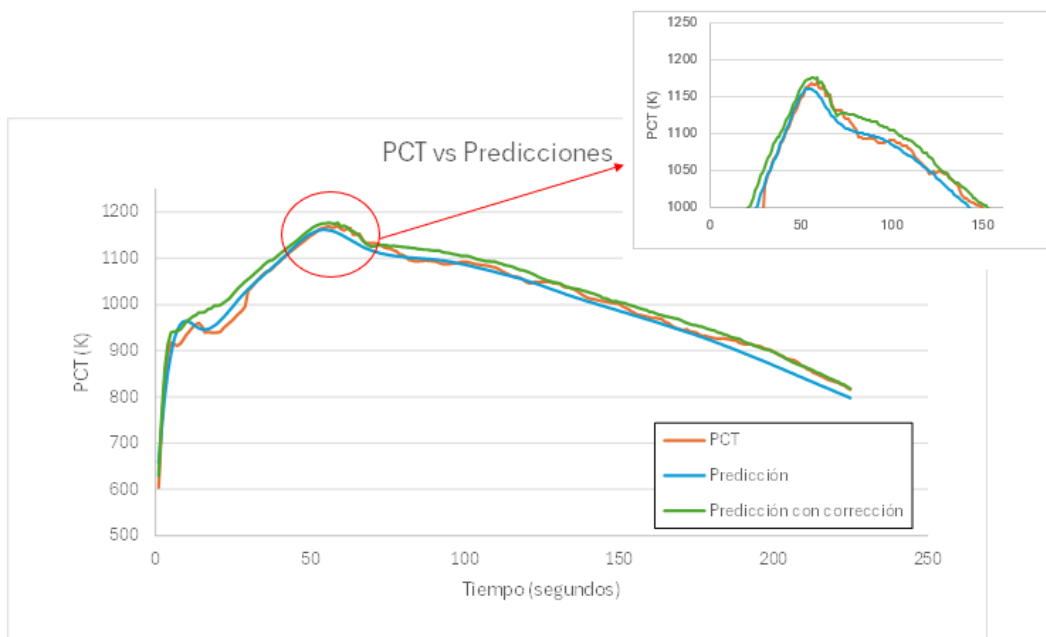


Figura 14. Comparación PCT y predicciones sin y con corrección de error.

A partir del modelo de regresión funcional obtenido se realiza el análisis de incertidumbre obteniendo 10000 simulaciones de la evolución temporal de la PCT muestreando aleatoriamente los 27 inputs seleccionados. A partir de las 10000 simulaciones se obtienen los percentiles 2.5 y 97.5 los cuales, posteriormente, se compararán con la mediana del intervalo de tolerancia bilateral 95/95 obtenido mediante el método de Wilks para diferentes tamaños de muestra 93, 124 y 153 correspondientes a los estadísticos de orden 1, 2 y 3. En la Figura 15 se muestra la evolución temporal de los percentiles 2.5 y 97.5. El valor máximo de la PCT que se alcanza a lo largo del transitorio es de 1270 K.

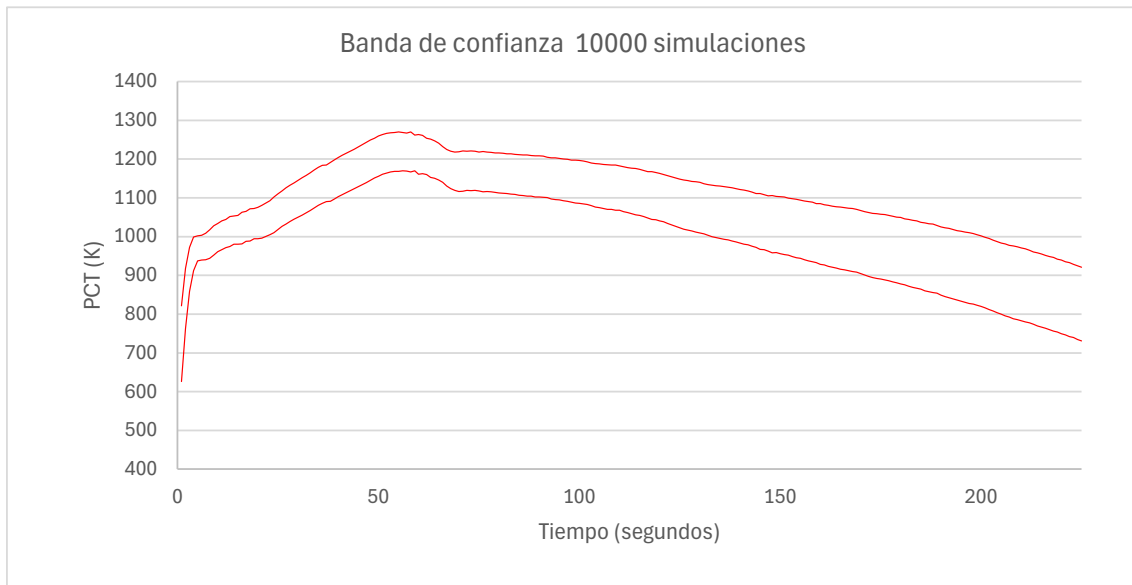


Figura 15. Banda de confianza 95% obtenida a partir del modelo de regresión funcional con corrección de error ( $n=10000$ ).

#### 4.4 Construcción de intervalos de tolerancia de Wilks

El método de Wilks es ampliamente utilizado en el sector nuclear por lo que, como se ha mencionado anteriormente, se compararán los resultados obtenidos con el modelo de regresión funcional y el método de Wilks para diferentes estadísticos de orden. En esta sección se presentan las etapas y los resultados de dicho método.

Partiendo de la base de datos inicial, se extrae de forma aleatoria un subconjunto de  $n$  muestras a partir de las 450 simulaciones de las que se dispone, con los datos de entrada de los valores de la temperatura alcanzada para cada instante del transitorio. Este muestreo aleatorio se realiza 1000 veces para cada tamaño de muestra  $n$  con el objetivo de obtener las estimaciones de los límites de tolerancia bilaterales 95/95 para el método de Wilks.

Las estimaciones de los límites de tolerancia bilaterales 95/95 se han realizado para los estadísticos de primer, segundo y tercer orden, cuyos tamaños de muestra correspondientes son  $n = \{93, 124, 153\}$  y con el objetivo de comparar dichos intervalos se ha tomado como referencia la mediana de los 1000 intervalos obtenidos mediante Bootstrap para los diferentes tamaños de muestra considerados. En la Figura 16 se

muestra la evolución de las medianas de los intervalos. (Consultar en el Anexo 3 el código utilizado para el cálculo del límite de tolerancia bilateral 95/95 para un estadístico de primer orden).

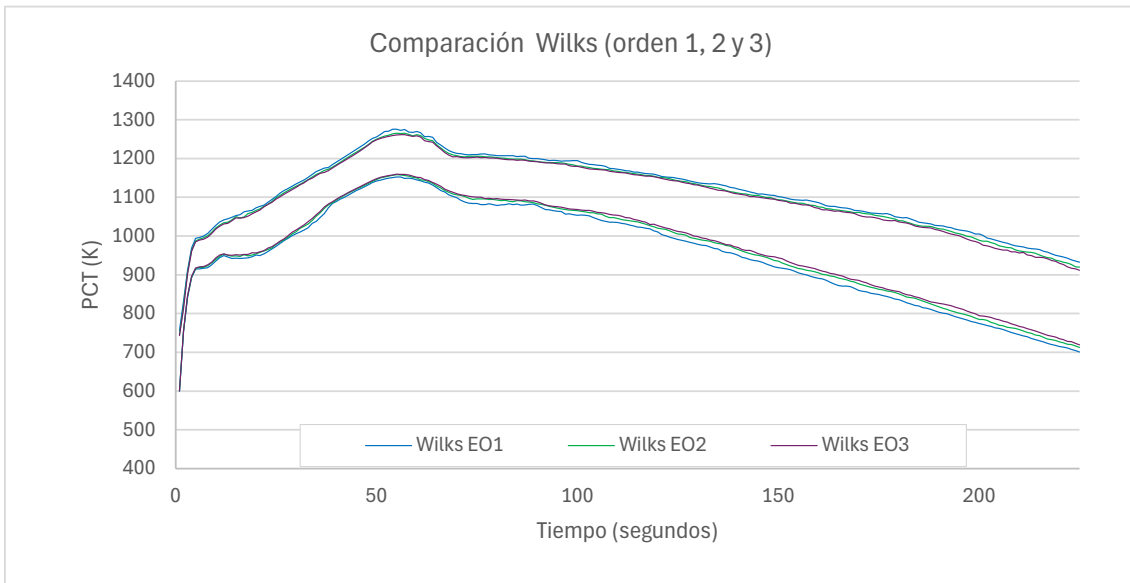


Figura 16. Comparación intervalos de tolerancia 95/95 ( $n=93, 124$  y  $153$ ).

Aparentemente, los resultados obtenidos para los diferentes estadísticos de orden son muy similares. Haciendo ‘zoom’ a los intervalos de tolerancia superiores en la zona crítica del transitorio se pueden apreciar las diferencias entre los mismos y el mayor conservadurismo del método de Wilks utilizando el estadístico de orden 1, tal y como se muestra en la Figura 17.

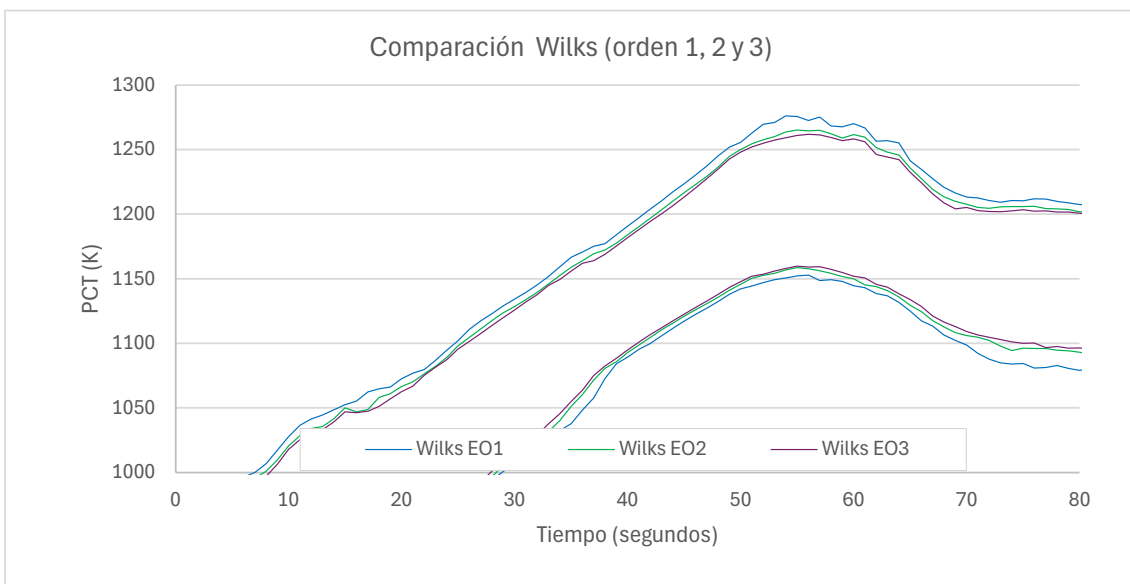


Figura 17. Zoom de los intervalos de tolerancia para los 3 E.O.

Las máximas temperaturas de los intervalos de tolerancia superiores para cada uno de los estadísticos de orden se muestran en la Tabla 5. En dicha tabla se observa como se reduce el valor máximo de la PCT al aumentar el orden en el método de Wilks. Así, en el caso del método de Wilks con estadístico de orden 1 ( $n=93$ ) la mediana toma un valor igual a

1275.97 K mientras que en el caso de utilizar el estadístico de orden 3 (n=153) la mediana toma un valor igual a 1261.92 K.

Tabla 5. Máximo valor del límite superior de tolerancia obtenido con el método de Wilks.

|                   | Wilks orden 1 | Wilks orden 2 | Wilks orden 3 |
|-------------------|---------------|---------------|---------------|
| <b>Max PCT(K)</b> | 1275.97K      | 1265.139      | 1261.92       |

#### 4.5 Comparación de la predicción con los intervalos de tolerancia

En este apartado se compara la banda de confianza obtenida con las 10000 predicciones del modelo de regresión funcional en el apartado 4.3 y la estimación de los intervalos de tolerancia mediante el método de Wilks obtenidos en el apartado 4.4 para diferentes estadísticos de orden. En la Figura 18 se representan los resultados obtenidos.

Como se observa en la figura, respecto al límite superior, excepto en la zona donde se alcanza el máximo valor de la PCT, los valores obtenidos con el modelo de regresión funcional presentan temperaturas superiores a las obtenidas con el método de Wilks. En la zona correspondiente a la máxima PCT los valores de temperatura se encuentran entre los obtenidos con el método de Wilks para estadístico de orden 1 y 2. Si se observa el límite inferior el modelo de regresión funcional es más conservador obteniéndose valores más elevados de la PCT que con el método de Wilks.

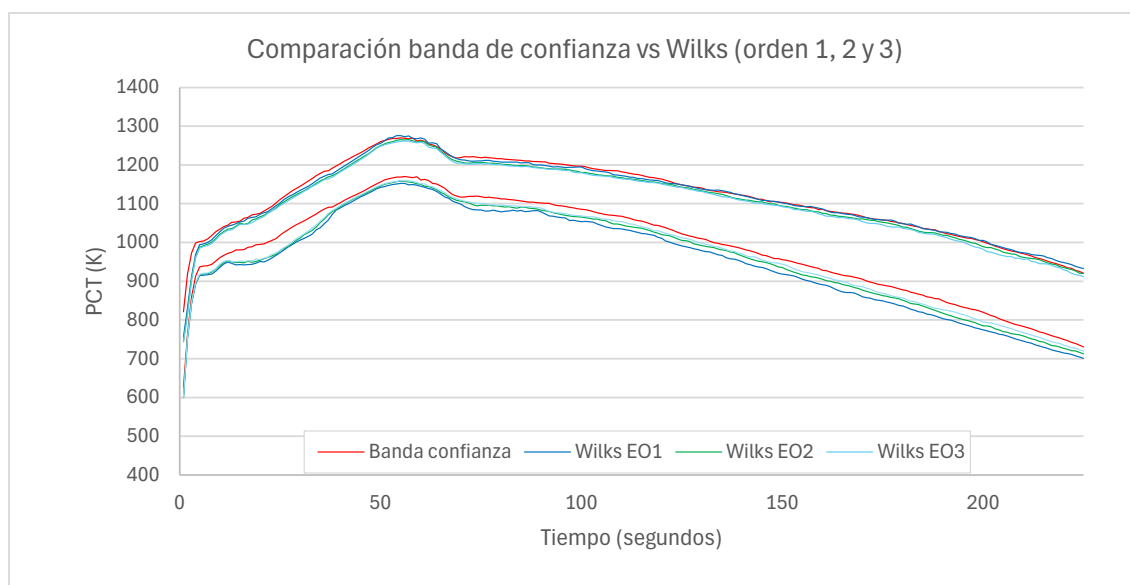


Figura 18. Comparación banda de confianza vs Wilks.

En la Tabla 6 se muestran los valores correspondientes al máximo de la evolución de PCT para los diferentes métodos analizados.

Tabla 6. Valor máximo de la evolución de la PCT para los diferentes métodos analizados.

|                   | <b>Regresión funcional</b> | <b>Wilks orden 1</b> | <b>Wilks orden 2</b> | <b>Wilks orden 3</b> |
|-------------------|----------------------------|----------------------|----------------------|----------------------|
| <b>Max PCT(K)</b> | 1270.115                   | 1275.973             | 1265.139             | 1261.912             |

Como se observa en la tabla y como se ha comentado anteriormente, el máximo valor obtenido con el modelo de regresión funcional es de 1270 K, valor que se encuentra entre los correspondientes al método de Wilks con estadístico de orden 1 y 2. A pesar de que la temperatura máxima de vaina alcanzada por el método de Wilks de Orden 1 es ligeramente superior a la obtenida en la predicción realizada por el modelo de regresión funcional, no se tiene en cuenta, ya que durante la mayor parte del tiempo del transitorio la predicción presenta valores superiores al resto de métodos, lo cual reafirma su nivel de conservadurismo. Además, en ningún caso se alcanzan los 1477K que no deben sobrepasarse, de acuerdo con los criterios de seguridad que aplican a un LBLOCA.

## 5. Conclusiones y trabajo futuro

El presente trabajo final de máster se ha enfocado en el estudio de la aplicabilidad de la regresión funcional en el análisis de incertidumbre de códigos termohidráulicos. A diferencia de la mayoría de los estudios realizados hasta el momento, que consideran la figura de mérito (FOM) como un escalar, en este trabajo se considera una FOM dependiente del tiempo, en concreto, la evolución temporal de la temperatura máxima de vaina en el núcleo. La metodología se ha aplicado a un escenario de rotura grande en una de las ramas frías de una central de tipo PWR de 4 lazos y los resultados obtenidos se han comparado con el método de Wilks, habitualmente utilizado en el análisis de incertidumbre de simulaciones termohidráulicas.

En relación con los resultados, el algoritmo de selección de variables ha demostrado ser una buena técnica para reducir el número de parámetros con los que se realiza el estudio de incertidumbres e identificar los parámetros de entrada más relevantes.

Se ha podido comprobar como la implementación de la regresión funcional como metamodelo, para la predicción de la evolución temporal de la temperatura máxima de vaina, es un método que proporciona estimaciones precisas. Además, permite predecir la FOM con mucho menor coste computacional que el código termohidráulico TRACE. No obstante, con el objetivo de obtener predicciones conservadoras, lo que resulta de interés en el contexto del análisis de incertidumbre de códigos termohidráulicos, se ha realizado una corrección de las predicciones obtenidas con el metamodelo a partir de la distribución del error. La comparación de la banda de incertidumbre, obtenida con el metamodelo ajustado con el error, y el intervalo de tolerancia bilateral de Wilks ha puesto de manifiesto que las estimaciones obtenidas son conservadoras. Así, se puede concluir que la sustitución del código mediante un metamodelo de regresión funcional ajustado con el error puede ser una buena alternativa para el análisis de incertidumbre.

Como continuación de este trabajo de máster y como en cualquier otro proyecto, existen diversas líneas de investigación que quedan abiertas y en las que es posible continuar trabajando. Como trabajo futuro se plantean diferentes líneas que pueden ampliar el presente estudio. En primer lugar, realizar un estudio de simulación del efecto del tamaño de muestra utilizado en el entrenamiento del modelo de regresión funcional. Este estudio permitiría evaluar la influencia del tamaño de muestra en la precisión y robustez de las predicciones del modelo. En segundo lugar, estudiar la aplicabilidad de otros métodos como la regresión cuantil o modelos aditivos generalizados para datos funcionales. Por último, ampliar el estudio a otros accidentes con pérdida de refrigerante como un escenario de rotura pequeña (SBLOCA) o de mediana rotura (MBLOCA).



## 6. Referencias

- Applied Programming Technology, I. (2012). *Symbolic Nuclear Analysis Package (SNAP). User's Manual. Version 2.2.1.* Applied Programming Technology, Inc., Bloomsburg.
- Bárcenas Curtis, R. (2016). *Estudio de Perfiles de Olas vía Análisis de Datos Funcionales.*
- Belaid, S., Freixa, J., & Zerkak, O. (2010). *Analysis of the Test OECD-PKL2 G7. 1 with the Thermal-Hydraulic System Code TRACE.* Washington DC: US Nuclear Regulatory Commission. NUREG/IA-0432.
- Choon, C. (2012). *Minimax concave bridge penalty function for variable selection.*
- Crainiceanu, C., Reiss, P., Goldsmith, J., Huang, L., Huo, L., Scheipl, F., . . . Xiao, L. (2024). *Refund: regression with functional data, version 0.1-35.*
- de Boor, C. (1977). *Package for Calculating with B-Splines. SIAM Journal on Numerical Analysis.*
- Delgado Manríquez, C., Cid Cid, A., & Leguey Galán, S. (2013). *Introducción al muestreo en poblaciones finitas.*
- Erich, N., Aicke, H., & Mario, U. (2014). *On Weak Tractability of the Clenshaw-Curtis Smolyak Algorithm.*
- Ferraty, F., & Vieu, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice (Springer Series in Statistics).* Springer-Verlag, Berlin.
- Freixa, J., Barrachina, T., Berna, C., Bocanegra, R., Carlos, S., Castro, E., & Villanueva, J. (2024). *Spanish contribution to the development and application of best estimate plus uncertainty methodologies: Past, present and future.* Nuclear Engineering and Design, 417, 112837.
- Huang, J., Breheny, P., Lee, S., Ma, S., & Zhang, C.-H. (2016). *The Mnet method for variable selection.*
- IAEA. (2009). *Specific Safety Guide No.SSG-2. Deterministic Safety Analysis for Nuclear Power Plants.* Vienna: International Atomic Energy Agency.
- Jacques, J., & Preda, C. (2014). *Model-based clustering for multivariate functional data.*
- Krishnamoorthy, K., & Mathew, T. (2009). *Statistical tolerance regions.*
- Martin, R., & O'Dell, L. (2005). *AREVA's realistic large break LOCA analysis methodology.*
- Morris, J. (2015). *Functional regression* . Annual Review of Statistics and Its Application, 2, 321-359.
- Najar, M., & Wang, H. (2022). *Comparative Machine Learning Study for Estimating Peak Cladding Temperature in AP1000 Under LOFW.* American Society of Mechanical Engineers: International Conference on Nuclear Engineering (Vol. 86397, p. V005T05A022.

- NEA. (2015). *Post-BEMUSE Reflood Model Input Uncertainty Methods (PREMIUM) Benchmark Phase II: Identification of Influential Parameters*. Nuclear Safety NEA/CSNI/R(2014)14.
- Perez , M., Reventos, F., Batet, L., Guba, A., Tóth, I., & Pascal, B. (2011). Uncertainty and sensitivity analysis of a LBLOCA in a PWR nuclear power plant: Results of the phase V of the BEMUSE programme.
- Petruzzi, A., D'Auria, F., Gianotti, W., & Ivanov, K. (2005). *Methodology of Internal Assessment of Uncertainty and Extension to Neutron Kinetics/Thermal-Hydraulics Coupled Codes*.
- Pourgol-Mohamad, M., Modarres, M., & Mosleh, A. (2009b). *Integrated Methodology for Thermal–Hydraulics Uncertainty Analysis with Application*. Nuclear Technology, 165, 333-359.
- Ramsay, J., & Dalzell, C. (1991). *Some Tools for Functional Data Analysis*. Journal of the Royal Statistical Society: Series B (Methodological).
- Ramsay, J., & Silverman, B. (1997). *Functional Data Analysis*. Springer, New York.
- Ramsay, J., & Silverman, B. (2005). *Fitting differential equations to functional data: Principal differential analysis*. Springer.
- Ramsay, J., Wickham, H., Graves, S., & Hooker, G. (2024). *fda: Functional Data Analysis*. R package version 6.1.8.
- Reiss, P., Huang, L., & Mennes, M. (2010). *Fast function-on-scalar regression with penalized basis expansions*. The international journal of biostatistics, 6(1).
- Reventós, F., A de Alfonso, Zhang, J., Kovtonyuk, A., Mendizabal, R., & Dethioux, A. (2016). *PREMIUM, a benchmark on the quantification of the uncertainty of the physical models in the system thermal-hydraulic codes: methodologies and data review*.
- Sallehuddin, W., & Diab, A. (2021). *Using machine learning to predict the fuel peak cladding temperature for a large break loss of coolant accident*. Frontiers in Energy Research, 9, 755638.
- Sánchez Sáez, F. (2017). *Aportaciones al análisis determinista de seguridad de las centrales nucleares mediante metodología Best Estimate*. Valencia: Universitat Politècnica de València.
- USNRC. (1989a). *RG 1.157 BEST-ESTIMATE CALCULATIONS OF EMERGENCY CORE COOLING SYSTEM PERFORMANCE*. Washington DC: United States Nuclear Regulatory Commission.
- USNRC. (1989b). *Quantifying Reactor Safety Margins. Application of Code Scaling, Applicability, and Uncertainty Evaluation Methodology to a Large-Break, Loss-of-Coolant Accident*. Washington DC: United States Nuclear Regulatory Commission.

- USNRC. (2007). *Acceptance criteria for emergency core cooling systems for light-water nuclear power reactors*. Washington DC: United States Nuclear Regulatory Commission.
- Wilks, S. (1941). *Determination of sample sizes for setting tolerance limits*. The Annals of Mathematical Statistics.
- Young, M., & Bajorek, S. (1998). *Application of code scaling applicability and uncertainty methodology to the large break loss of coolant*. Nuclear Engineering and Design, 186, 39-52.

## Anexos

Anexo 1. Código RStudio para la selección de variables y obtención de coeficientes (paquete “refund”)

```
## Refund package for Scalar-on-function regression
`` {r}
install.packages("refund")
library(refund)
install.packages("readxl")
install.packages("openxlsx")
library(readxl)
library(openxlsx)
...
`` {r}
Y436 <- read_excel("D:\\Jaime\\Y436_funcional.xlsx")
Y436<-as.matrix(Y436)
X436 <- read_excel("D:\\Jaime\\X436_funcional.xlsx")

#Se estandarizan las covariatas
X436 = scale(X436)
X436 = as.data.frame(X436)

# Formo una nueva matriz de datos juntando ambas matrices
datos436=X436
datos436$Y436=Y436
...

`` {r}
# Aplico la función fosr.vs() del paquete 'refund'
fit.fosr.vs_436=fosr.vs(Y436~., data = datos436, method="grMCP")
grMCP_436 = fit.fosr.vs_436$coefficients
...

```

```

```{r}
grMCP_436 = as.data.frame(grMCP_436)
# Escribir la matriz en un archivo Excel
write.xlsx(grMCP_436,"coef436.xlsx")
...
```{r}
coefs = coef(fit.fosr.vs_436)
tiempo <- 1:225
as.numeric(tiempo)
rango_coefs <- c(-100, 2000)
# Se establece el área de trazado
plot(tiempo, coefs[1,], type = "l", xlab = "Tiempo (segundos)", ylab = "Coeficientes", col
= "blue", main = "Curvas de Coeficientes a lo largo del Tiempo", ylim = rango_coefs)
# Se agregan líneas para las curvas restantes
for (i in 2:46) {lines(tiempo, coefs[i,], col = rainbow(46)[i])}
legend("topright", legend = rownames(coefs), col = 1:nrow(coefs), lty = 1, cex=0.2)
#Plot sin el intercepto
rango_coefs = c(-100,100)
plot(tiempo, coefs[2,], type = "l", xlab = "Tiempo (segundos)", ylab = "Coeficientes", col
= "blue", main = "Curvas de Coeficientes a lo largo del tiempo (n=450)", ylim =
rango_coefs)
# Se agregan líneas para las curvas restantes
for (i in 3:46) {lines(tiempo, coefs[i,], col = rainbow(46)[i])}
...
```{r}
grMCP_436 = grMCP_436[-1,]
# Extraer los coeficientes en el segundo 55
coef_second_55 = grMCP_436[, 55]
# Calcular el valor absoluto de los coeficientes
abs_coef_second_55 = abs(coef_second_55

```

**# Crear un data frame con los nombres de las variables y sus coeficientes**

```
coef_data = data.frame(Variable = rownames(grMCP_436), Coefficient =  
coef_second_55, AbsoluteCoefficient = abs_coef_second_55)
```

**# Ordenar el data frame por el valor absoluto de los coeficientes en orden descendente**

```
coef_data_sorted = coef_data[order(-coef_data$AbsoluteCoefficient), ]
```

**# Mostrar el resultado**

```
print(coef_data_sorted)
```

```
```\n
```

```
```\n{r}
```

```
coef_data_sorted = as.data.frame(coef_data_sorted)
```

```
write.xlsx(coef_data_sorted, "VS450.xlsx")```\n
```

**# Extraer los coeficientes**

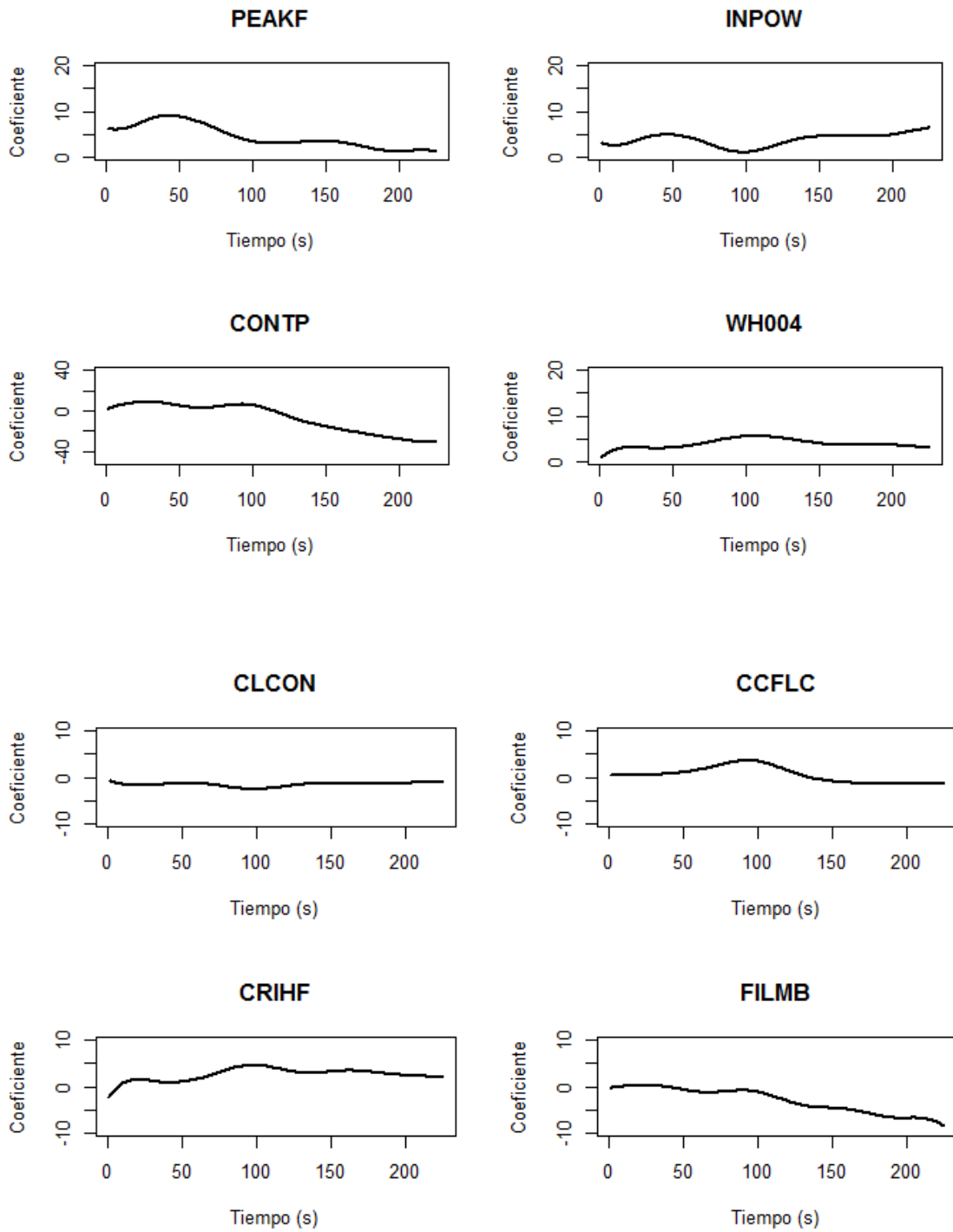
```
significant_coefs <- coefs[28, ]
```

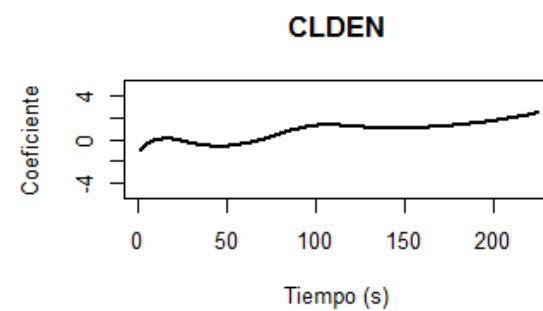
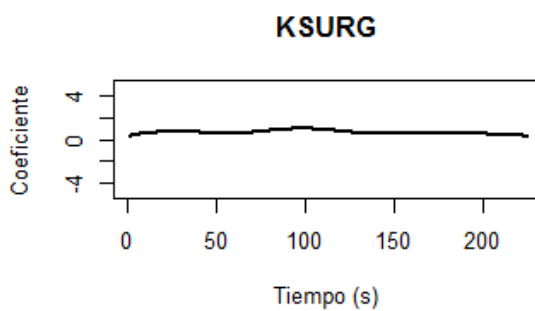
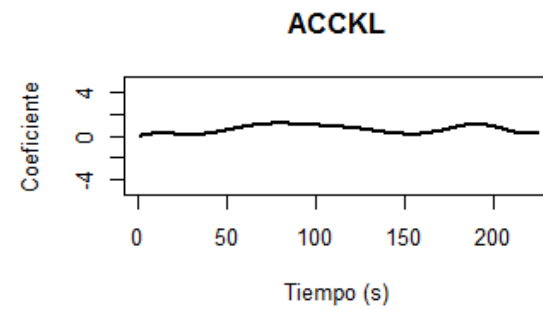
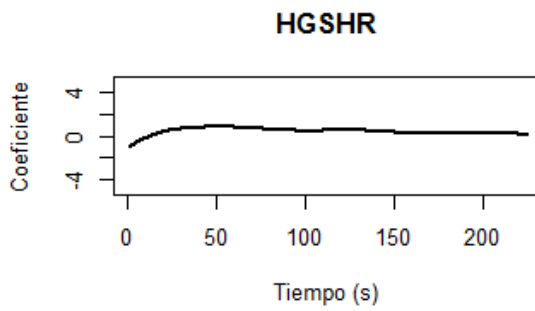
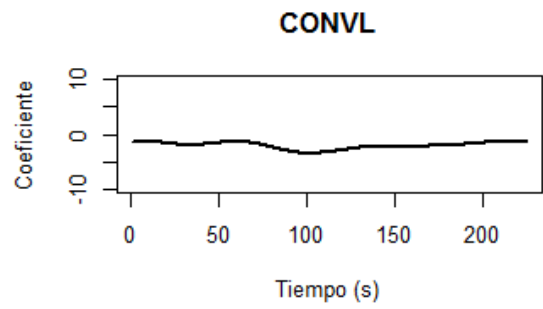
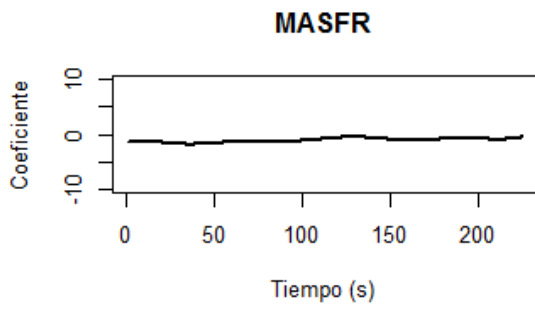
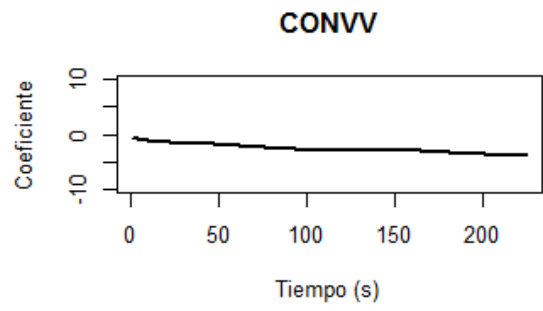
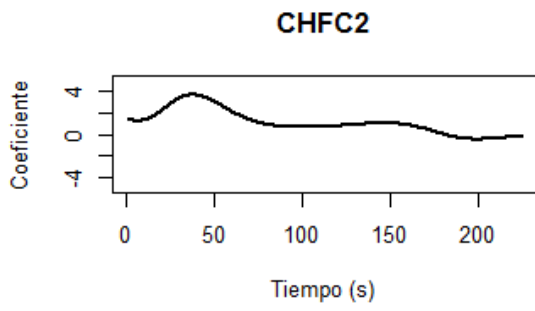
**# Crear el gráfico**

```
plot(tiempo, significant_coefs, type = "l", xlab = "Tiempo (s)", ylab = "Coeficiente",  
col = "blue", main = "TMIFS", ylim = c(-10,40), lwd=2)
```

```
```\n
```

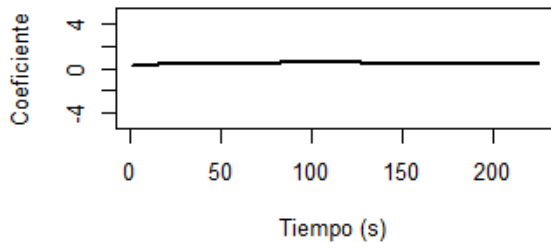
Anexo 2. Evolución temporal de los coeficientes del modelo de regresión funcional



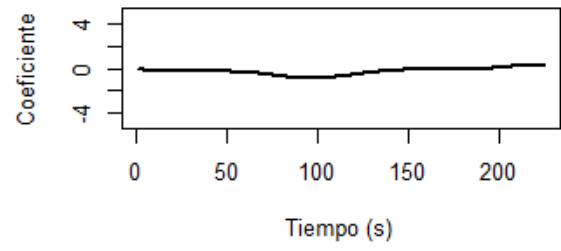




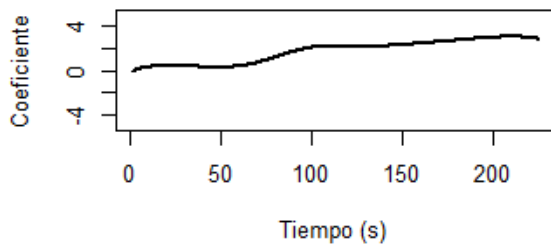
**CLTEM**



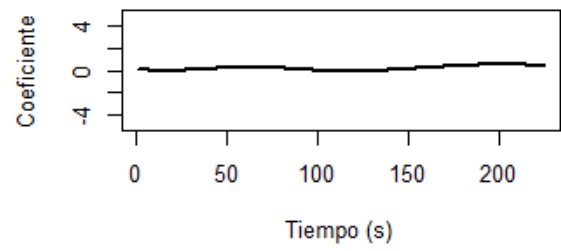
**UO2SH**



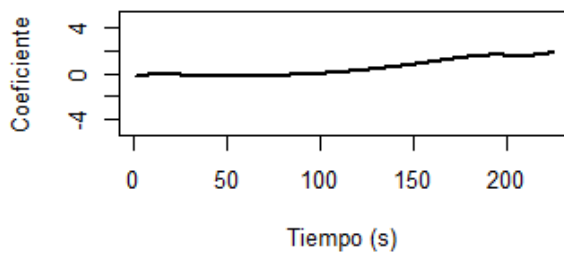
**DRAG5**



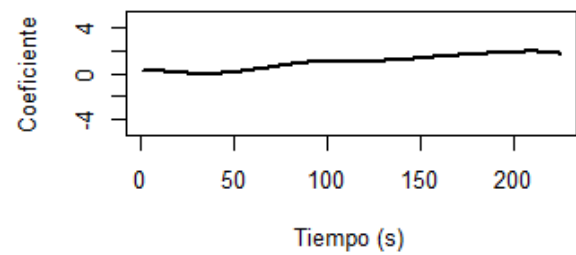
**DRAG4**



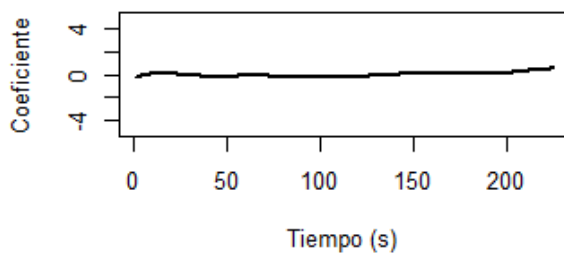
**LINT2**



**UHTEM**



**ACCLT**



Anexo 3. Código de obtención de intervalos de tolerancia para un tamaño de muestra n=93 (Método de Wilks 95/95 para un E.O 1)

```
`` {r}
data = read.csv("D:\\Jaime\\Resultados\\Y_450.csv", sep=";")
``
`` {r}
# Inicializar matrices para almacenar los intervalos de tolerancia
int_tol_inferiores_1 <- matrix(NA, nrow = 450, ncol = 225)
int_tol_superiores_1 <- matrix(NA, nrow = 450, ncol = 225)
# Generar 1000 muestras aleatorias y calcular los intervalos de tolerancia para cada segundo
for (j in 1:1000) {
  # Muestreo aleatorio simple para seleccionar una muestra de tamaño n=93
  muestra_aleatoria <- sample(1:1000, size = 93)
  # Seleccionar la muestra de la matriz de datos
  data_muestra <- data[muestra_aleatoria, ]
  # Calcular intervalos de tolerancia para cada segundo de esta muestra
  for (i in 1:225) {
    # Paso 1: Calcular los estadísticos de orden para el segundo específico
    estadisticos_ordenados <- sort(data_muestra[, i])
    primer_minimo <- estadisticos_ordenados[1]
    primer_maximo <- estadisticos_ordenados[length(estadisticos_ordenados)]
    # Paso 2: Calcular los percentiles 2.5 y 97.5
    percentil_2.5_Wilks <- primer_minimo
    percentil_97.5_Wilks <- primer_maximo
    # Paso 3: Asignar los límites de tolerancia para este segundo específico
    int_tol_inferiores_1[j, i] <- percentil_2.5_Wilks
    int_tol_superiores_1[j, i] <- percentil_97.5_Wilks
  }
}
}
```

```

```{r}
# Reordenar los valores de la matriz de intervalos inferiores de menor a mayor
int_tol_inferiores_1 <- apply(int_tol_inferiores_1, 2, sort)
# Reordenar los valores de la matriz de intervalos superiores de menor a mayor
int_tol_superiores_1 <- apply(int_tol_superiores_1, 2, sort)
# Calcular la mediana de cada columna de la matriz de intervalos inferiores
mediana_inferiores_1 <- apply(int_tol_inferiores_1, 2, median)
# Calcular la mediana de cada columna de la matriz de intervalos superiores
mediana_superiores_1 <- apply(int_tol_superiores_1, 2, median)
```
```{r}
# Se grafica el primer límite inferior
plot(1:225, mediana_inferiores_1, type = "l", col = "red", xlab = "Tiempo (segundos)",
ylab = "Temperatura", ylim = range(int_tol_inferiores_1, int_tol_superiores_1))
# Se agrega la línea para el primer límite superior
lines (1:225, mediana_superiores_1, type = "l", col = "blue")
# Se agregan las leyendas
legend("topright", legend = c("Límite inferior", "Límite superior"), col = c("red", "blue"),
lty = 1)
```

```

Anexo 4. Relación del trabajo con los Objetivos de Desarrollo Sostenible de la Agenda 2030

Tabla 7. Relación del trabajo con los Objetivos de Desarrollo Sostenible de la Agenda 2030.

| Objetivos de Desarrollo Sostenibles             | Alto | Medio | Bajo | No procede |
|---|------|-------|------|------------|
| ODS 1. Fin de la pobreza                        |      |       |      | •          |
| ODS 2. Hambre cero                              |      |       |      | •          |
| ODS 3. Salud y bienestar                        |      |       |      | •          |
| ODS 4. Educación de calidad                     |      |       |      | •          |
| ODS 5. Igualdad de género                       |      |       |      | •          |
| ODS 6. Agua limpia y saneamiento                |      |       |      | •          |
| ODS 7. Energía asequible y no contaminante      | •    |       |      |            |
| ODS 8. Trabajo decente y crecimiento económico  |      | •     |      |            |
| ODS 9. Industria, innovación e infraestructuras | •    |       |      |            |
| ODS 10. Reducción de las desigualdades          |      |       |      | •          |
| ODS 11. Ciudades y comunidades sostenibles      |      | •     |      |            |
| ODS 12. Producción y consumo responsables       |      |       |      | •          |
| ODS 13. Acción por el clima                     | •    |       |      |            |
| ODS 14. Vida submarina                          |      |       |      | •          |
| ODS 15. Vida de ecosistemas terrestres          |      |       |      | •          |
| ODS 16. Paz, justicia e instituciones sólidas   |      |       |      | •          |
| ODS 17. Alianzas para lograr objetivos          |      |       |      | •          |

\*Explicados en el apartado 1 del trabajo