



# ESCUELA POLITÉCNICA NACIONAL

## FACULTAD DE CIENCIAS

### NUEVA METODOLOGÍA PARA LA DETECCIÓN DE RESULTADOS INCONSISTENTES EN ESTUDIOS INTERLABORATORIO UTILIZANDO ANÁLISIS DE DATOS FUNCIONALES

TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE  
INGENIERO MATEMÁTICO

#### PROYECTO DE INVESTIGACIÓN

GÉNESIS CHABELI MORENO ARMIJOS

genesis.moreno@epn.edu.ec

CRISTIAN IVÁN SOLÓRZANO MAYA

cristian.solorzano@epn.edu.ec

Director: DR. MIGUEL ALFONSO FLORES SÁNCHEZ

miguel.flores@epn.edu.ec

QUITO, ENERO 2022

## **DECLARACIÓN**

Nosotros, GÉNESIS CHABELI MORENO ARMIJOS & CRISTIAN IVÁN SOLÓRZANO MAYA, declaramos bajo juramento que el trabajo aquí escrito es de nuestra autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y se han consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedemos nuestros derechos de propiedad intelectual, correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su reglamento y por la normatividad institucional vigente.

## **CERTIFICACIÓN**

Certifico que el presente trabajo fue desarrollado por GÉNESIS CHABELI MORENO ARMIJOS & CRISTIAN IVÁN SOLÓRZANO MAYA, bajo mi supervisión.

---

Dr. MIGUEL ALFONSO FLORES SÁNCHEZ

Director del Proyecto

## **AGRADECIMIENTOS**

Agradecemos a nuestros padres Enrique & Nohemí e Isabel & Carlos, por ser los principales promotores de nuestros sueños, por confiar y creer en nuestras expectativas, por los consejos, valores y principios que nos han inculcado y sobre todo por su apoyo incondicional al no permitirnos decaer cuando momentos difíciles se presentaron; valoramos con mucha gratitud la motivación y cariño constante que nos ha impulsado a concluir una nueva etapa importante para nuestras vidas.

Finalmente queremos expresar nuestro más sincero agradecimiento por la confianza y apoyo al PhD. Miguel Flores, principal colaborador y mentor durante todo este proceso quien con su dirección, enseñanza y guía permitió el desarrollo de este trabajo. Además de las nuevas experiencias, consejos y oportunidades que nos brindó para alcanzar nuestros primeros aportes en la investigación científica; sin duda todas las sugerencias brindadas desde su formación y experticia nos ha permitido tener nuevas visiones y enfoques hacia donde encaminar nuestro desarrollo profesional.

## **DEDICATORIA**

*A nuestros familiares y amigos, quienes han mostrado un apoyo incondicional a lo largo de nuestra carrera profesional.*

# Índice general

<b>Resumen</b>	<b>XIV</b>
<b>Abstract</b>	<b>XVI</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Introducción . . . . .	1
1.2. Objetivos . . . . .	3
1.2.1. Objetivo General . . . . .	3
1.2.2. Objetivo Específico . . . . .	3
1.3. Justificación . . . . .	3
<b>2. Enfoque FDA para la detección de atípicos en Estudios Interlaboratorio</b>	<b>7</b>
2.1. Introducción Estudios Interlaboratorio . . . . .	7
2.1.1. Consideraciones generales para realizar un Estudio Interlaboratorio . . . . .	8
2.1.2. Tratamiento de datos atípicos en Estudios Interlaboratorio . . . . .	9
2.2. Técnicas FDA aplicadas en un Estudio Interlaboratorio . . . . .	15
2.2.1. Análisis de Datos Funcionales . . . . .	15
2.2.2. Técnicas de remuestreo bootstrap suavizado . . . . .	18

2.2.3.	Método de detección de atípicos basado en proyecciones aleatorias . . . . .	20
2.2.4.	Pruebas ANOVA Funcionales . . . . .	24
2.3.	Casos de estudio FDA . . . . .	27
2.3.1.	Datos Termogravimétricos TG . . . . .	27
2.3.2.	Confort Térmico de superficies comerciales de empresa textil en Ciudad de Panamá . . . . .	30
2.3.3.	Aplicación de técnicas FDA a Datos Termogravimétricos. . . . .	32
<b>3.</b>	<b>Metodología de detección de atípicos basados en técnicas FDA para los estadísticos funcionales <math>H</math> y <math>K</math> de Mandel, utilizando proyecciones aleatorias</b>	<b>41</b>
3.1.	Precisión, repetibilidad y reproducibilidad . . . . .	42
3.2.	Estadísticos funcionales $H$ y $K$ de Mandel . . . . .	44
3.2.1.	Planteamiento del contraste de hipótesis de repetibilidad y reproducibilidad para el caso funcional . . . . .	45
3.3.	Metodología propuesta para la detección de laboratorios atípicos con base en proyecciones aleatorias (PA) . . . . .	49
3.3.1.	Metodología funcional para la detección de laboratorios atípicos con base en la norma $L_2$ . . . . .	49
3.3.2.	Metodología funcional para la detección de laboratorios atípicos con base en proyecciones aleatorias (PA) . . . . .	50
<b>4.</b>	<b>Validación y aplicación de la nueva metodología</b>	<b>53</b>
4.1.	Validación del nuevo enfoque FDA . . . . .	53
4.1.1.	Descripción del estudio de simulación . . . . .	54
4.1.2.	Estudio de simulación bajo la hipótesis nula de repetibilidad y reproducibilidad . . . . .	55



4.1.3. Estudio de simulación bajo las hipótesis alternativas de repetibilidad y reproducibilidad . . . . .	58
4.2. Aplicación de la nueva metodología a casos reales de estudio . . . . .	66
4.2.1. Estudio interlaboratorio usando análisis termogravimétrico . . . . .	67
4.2.2. Identificación de sensores anómalos en superficies comerciales . . . . .	71
<b>5. Conclusiones y Recomendaciones . . . . .</b>	<b>75</b>
5.1. Conclusiones . . . . .	75
5.2. Recomendaciones y futuras líneas de estudio . . . . .	78
<b>A. Anexo I: Códigos implementados en R . . . . .</b>	<b>79</b>
A.1. Técnicas FDA aplicadas a ILS . . . . .	79
A.1.1. Análisis exploratorio FDA . . . . .	79
A.1.2. Bootstrap Suavizado . . . . .	80
A.1.3. Bandas de confianza Bootstrap . . . . .	81
A.1.4. Detección de valores atípicos FDA basado en el cálculo de la profundidad de datos usando proyecciones aleatorias. . . . .	82
A.1.5. Pruebas ANOVA Funcionales . . . . .	82
A.2. Estudio de simulación . . . . .	83
A.2.1. Generación aleatoria de datos funcionales TG . . . . .	83
A.2.2. Generación de proyecciones aleatorias y profundidades univariantes de Tukey . . . . .	84
A.2.3. Generación de estadísticos funcionales H y K y su distribución Bootstrap . . . . .	85
A.2.4. Escenarios de reproducibilidad (Estadístico H) . . . . .	88
A.2.5. Escenarios de repetibilidad (Estadístico K) . . . . .	91

A.3. Aplicación caso real de estudio curvas TG . . . . .	94
<b>Bibliografía</b>	<b>96</b>
<b>Anexos</b>	<b>99</b>

# Índice de figuras

2.1. Curva TG correspondiente al oxalato de calcio y su proceso de degradación . . . . .	29
2.2. Curva TG correspondiente a cada laboratorio . . . . .	30
2.3. Curvas de temperatura obtenidos por sensores en local comercial (comfort térmico) . . . . .	31
2.4. Media y varianza funcional Global de curvas TG . . . . .	33
2.5. Bandas de confianza bootstrap curvas TG . . . . .	33
2.6. Media y varianza funcional de curvas TG para cada laboratorio . . . .	34
2.7. Resultados ANOVA Funcional un factor aplicado a data TG . . . . .	36
2.8. Resultados estadísticos $h$ y $k$ de Mandel a partir de datos IDT . . . . .	38
2.9. Resultados estadísticos $h$ y $k$ de Mandel a partir de datos IDT (segundo paso de degradación) . . . . .	39
4.1. Medias curvas TG simuladas y estadístico H . . . . .	59
4.2. Resultados obtenidos de los estadísticos $d^H$ para la metodología PA Y $L_2$ . . . . .	60
4.3. Varianzas curvas TG simuladas y estadístico K . . . . .	61
4.4. Resultados obtenidos de los estadísticos $d^K$ para la metodología PA Y $L_2$ . . . . .	62
4.5. Curvas de potencia metodología PA . . . . .	64

4.6. Curvas de potencia metodología $L_2$ . . . . .	65
4.7. Resultados obtenidos para el estadístico $d_{PA}^H$ . . . . .	68
4.8. Estadísticos $d_{PA}^H$ para la muestra de los laboratorios . . . . .	68
4.9. Histogramas obtenidos para el estadístico $d_{L_2}^H$ . . . . .	69
4.10. Resultados obtenidos para el estadístico $d_{PA}^K$ . . . . .	70
4.11. Resultados obtenidos para el estadístico $d_{L_2}^K$ . . . . .	70
4.12. Media y varianza funcional globales de Sensores térmicos . . . . .	72
4.13. Media y varianza funcional de Sensores térmicos . . . . .	73
4.14. Resultados obtenidos para el estadístico $d_{PA}^H$ . . . . .	73
4.15. Resultados obtenidos para el estadístico $d_{PA}^K$ . . . . .	74

# Índice de tablas

2.1. Descripción y etiquetas correspondiente a cada laboratorio. . . . .	29
2.2. Profundidades de curvas TG obtenidas mediante proyecciones aleatorias . . . . .	35
2.3. Resultados método ANOVA funcional Vía proyecciones aleatorias .	37
2.4. Resultados estadísticos $h$ y $k$ de Mandel a partir de datos IDT con normas ASTM . . . . .	38
2.5. Resultados estadísticos $h$ y $k$ de Mandel a partir de característica IDT en el segundo paso de degradación con normas ASTM . . . . .	39
4.1. Potencias estadístico $d^H$ para metodologías $PA$ Y $L_2$ bajo la hipótesis nula de reproducibilidad . . . . .	57
4.2. Potencias estadístico $d^K$ para metodologías $PA$ Y $L_2$ bajo la hipótesis nula de repetibilidad . . . . .	57
4.3. Potencias estadístico $d^H$ para metodologías $PA$ Y $L_2$ bajo la hipótesis alternativa de reproducibilidad . . . . .	63
4.4. Potencias estadístico $d^H$ para metodologías $PA$ Y $L_2$ bajo la hipótesis alternativa de repetibilidad . . . . .	63

# Resumen

El presente trabajo de investigación propone el desarrollo e implementación de una nueva metodología para la detección de laboratorios atípicos en Estudios Interlaboratorio, a partir de técnicas de análisis de datos funcionales y del uso de los estadísticos  $h$  y  $k$  de Mandel, encargados del análisis de la variabilidad de las mediciones de estudio.

El trabajo de investigación se desarrolla en cinco capítulos. El primero de ellos, habla sobre los objetivos y alcances del estudio. En el Capítulo 2 se presenta un enfoque del análisis de datos funcionales para la detección de atípicos en estudios interlaboratorio. En este capítulo se introduce formalmente a los estudios interlaboratorio y se proporciona una descripción general del análisis de datos funcionales, así como definiciones y técnicas que pueden ser aplicadas en un estudio interlaboratorio, tales como técnicas remuestreo Bootstrap suavizado, ANOVA funcional, método de detección de atípicos basado en nociones de profundidad de datos y proyecciones aleatorias. Adicionalmente se presentan los conjuntos de datos reales que serán analizados durante el desarrollo del presente trabajo, y en uno de ellos se aplicará de forma práctica las técnicas funcionales antes mencionadas.

En el Capítulo 3 se presenta a detalle la nueva metodología propuesta para la detección de laboratorios atípicos en Estudios Interlaboratorio en base a técnicas del análisis de datos funcionales, proyecciones aleatorias y la extensión funcional de los estadísticos  $h$  y  $k$  de Mandel. A partir de estos estadísticos se realizará el planteamiento de contraste de las Hipótesis de Repetibilidad y Reproducibilidad para el caso funcional. La nueva metodología tiene como idea clave trasladar la resolución de un problema funcional a un caso más simple de estudio, el caso real.

En el Capítulo 4 se presenta la validación y aplicación de la nueva metodología propuesta, así mismo, se realiza una comparación y evaluación de los principales resultados obtenidos frente a resultados de estudios previos relacionados al uso de

datos funcionales en el marco de los ILS.

La metodología propuesta es corroborada y aplicada en un conjunto de datos reales de carácter funcional, es decir, curvas experimentales de medidas obtenidas por unidad de tiempo. La validación consiste en un estudio de simulación de curvas térmicas, específicamente, curvas termogravimétricas, mientras que la aplicación se realiza en dos casos de estudios reales. El primer caso de estudio corresponde a técnicas de análisis térmico de termogravimetría en el marco de un ILS, y el segundo a un estudio de confort térmico, mediante el análisis de mediciones de temperatura obtenidas por sensores colocados un área comercial de un edificio de la ciudad de Panamá. Este último se realiza con la finalidad de mostrar que la nueva propuesta podría determinar la precisión de cualquier método de medición, y aportar soluciones no solo en los diferentes laboratorios de análisis sino a problemas reales de diferentes industrias.

Finalmente, en el Capítulo 5 se muestran las principales conclusiones y recomendaciones obtenidas durante el desarrollo de la investigación.

# Abstract

This research work proposes the development and implementation of a new methodology for detection of atypical laboratories in Interlaboratory Studies, based on functional data analysis techniques and the use of Mandel's  $h$  and  $k$  statistics, which are in charge of the analysis of variability of study measurements.

This work is developed in five chapters. The first one talks about the objectives and scope of the topic. Chapter 2 presents an approach to the analysis of functional data for the detection of outliers in interlaboratory studies. This chapter formally introduces interlaboratory studies, repeatability and reproducibility, and given the functional nature of the data to be worked with, is provided a general description of functional data analysis, as well as certain definitions and techniques that can be applied to an interlaboratory study, such as smoothed Bootstrap resampling techniques, functional ANOVA, outlier detection method based on notions of data depth and random projections. Additionally, this chapter presents two sets of real data to be analyzed, and in one of them the aforementioned functional techniques will be applied in a practical way.

Chapter 3 presents in detail the new methodology proposed for the detection of atypical laboratories in Interlaboratory Studies based on functional data analysis techniques, random projections and the functional extension of Mandel's  $h$  and  $k$  statistics. From these statistics, the contrasting approach of the repeatability and reproducibility hypotheses for the functional case will be carried out. The new methodology has as a key idea transferring the resolution of a functional problem to a simpler case study, the real case.

Chapter 4 presents the validation and application of the new proposed methodology, additionally, a comparison and evaluation of the main results obtained against the results of previous studies related to the use of functional data in the framework of ILS is made. The proposed methodology is validated and applied to real data sets



of a functional nature, that is, experimental curves of measurements obtained per unit of time. The validation consists of a simulation study of thermal curves, specifically, thermogravimetric curves, while the application is carried out in two real case studies. The first study case corresponds to thermogravimetry thermal analysis techniques within the framework of an ILS, and the second to a study of thermal comfort, through the analysis of temperature measurements obtained by sensors placed in a commercial area of a building in Panama City. The latter is carried out in order to show that the new proposal could determine the precision of any measurement method and provide solutions not only in the different analysis laboratories but also to real problems in different industries.

Finally, Chapter 5 shows the main conclusions and recommendations obtained during the development of the research.

# Capítulo 1

## Introducción

### 1.1. Introducción

Los laboratorios de ensayos analíticos están sujetos a la implementación de sistemas de control de la calidad con el fin de alcanzar la excelencia en los resultados de la información proporcionada en este tipo de estudios. Estos sistemas son evaluados internamente por el mismo laboratorio, y externamente, a través de los denominados Estudios Interlaboratorio (de aquí en adelante, ILS por sus siglas en el idioma inglés *Interlaboratory Studies*), cuyo objetivo principal es la detección de laboratorios atípicos, es decir, aquellos laboratorios que presentan resultados que difieren al resto.

Un ILS entonces es un proceso de control organizado y gestionado por organismos independientes, uno de ellos es la *American Society for Testing and Materials* (ASTM International), que se ocupa de tareas y procedimientos propios de un ILS, como el diseño de un ILS, la identificación de muestras, el análisis estadístico de los resultados aportados por los laboratorios, entre otras funciones. De igual manera, se puede mencionar a la *International Organization for Standardization* (ISO), encargada de la definición de estándares de calidad para los materiales empleados en un proceso experimental.

En las normas establecidas por estas dos entidades, normas ASTM E-691 e ISO 5725, se establece inicialmente una metodología clásica paramétrica de carácter univariante para la detección de laboratorios atípicos en un ILS, esta metodología supone principalmente la normalidad de los datos de estudio, y con ello enfatiza en el aná-

lisis de dos factores importantes. El primero corresponde a la determinación de la precisión de un método de medición mediante el empleo de los denominados estudios de Reproducibilidad y Repetibilidad (R&R), y el segundo al análisis estadístico del estudio de precisión, en donde se utilizan pruebas de consistencia y contrastes de hipótesis basándose en los estadísticos  $h$  y  $k$  de Mandel [1].

Continuando con los enfoques univariantes, en colaboración con investigadores de la universidad de la Coruña y el actual director de tesis, propusimos una nueva alternativa no paramétrica univariante, para la detección de laboratorios atípicos en un ILS, con base en técnicas de remuestreo bootstrap; este estudio fue publicado recientemente en un artículo científico en el año 2021 [2]. El objetivo de la investigación fue hallar las aproximaciones de las distribuciones teóricas de los estadísticos  $h$  y  $k$  de Mandel, mediante remuestreo bootstrap, para no prescindir del supuesto de normalidad de los datos establecido por las normas ASTM E-691 en el enfoque tradicional.

Los estudios y metodologías antes mencionadas son de gran utilidad cuando se aplican a datos o características de análisis cuyos valores son escalares, sin embargo, en la actualidad, existen casos de estudios experimentales en los que se emplean instrumentos de medida cada vez más sofisticados, generando grandes volúmenes de datos rápidos y complejos que, haciendo referencia al campo del Big data, son difíciles de procesarlos con métodos tradicionales y por tanto, un enfoque univariante dentro de un ILS no llega a ser suficiente ni adecuado para analizar este tipo de datos. Adicional a lo descrito, cuando los datos masivos corresponden a medidas obtenidas por unidad de tiempo, conciernen a datos funcionales y se requiere emplear técnicas del *Análisis de datos funcionales* (FDA), con el objetivo de aprovechar toda la información disponible y evitar generar resultados variables y poco confiables que pueden presentarse cuando los datos no son tratados como tales.

En el ámbito FDA, una primera y única propuesta fue planteada por *Flores et al.* en [3], donde se desarrolla una metodología para la detección de laboratorios atípicos con base en la extensión funcional de los estadísticos  $h$  y  $k$  de Mandel y el uso de la norma  $L_2$ . Tomando como referencia [3], con la finalidad de ampliar las metodologías de estudio y contar con herramientas propias para los ILS en el ámbito funcional, el presente trabajo de investigación propone una nueva metodología que toma como punto focal la detección de anomalías y de laboratorios atípicos en un ILS, haciendo uso de técnicas FDA, extensión funcional de los estadísticos univariantes  $h$  y  $k$  de Mandel y el método de proyecciones aleatorias.

## **1.2. Objetivos**

### **1.2.1. Objetivo General**

Desarrollar una nueva metodología a partir del análisis de datos funcionales que permita detectar laboratorios atípicos en Estudios Interlaboratorio.

### **1.2.2. Objetivo Específico**

1. Extender conceptos y técnicas estadísticas univariantes, empleados en los estudios de repetibilidad y reproducibilidad, para la detección de laboratorios atípicos dentro de un ILS, al caso funcional.
2. Programar un algoritmo en el software estadístico R para la detección de atípicos, en el contexto FDA, utilizando técnicas de remuestreo Bootstrap y métodos de análisis de profundidad de datos.
3. Crear un escenario de simulación de un Estudio Interlaboratorio con datos funcionales para validar el estudio de inconsistencias y detección de atípicos con base en la nueva metodología propuesta.
4. Realizar un Estudio Interlaboratorio, aplicando la nueva metodología propuesta, a dos casos reales de estudio correspondientes a técnicas de análisis térmico como la termogravimetría y a un estudio de confort térmico de un área comercial de la ciudad de Panamá.
5. Comparar y evaluar los resultados de la metodología propuesta frente a los resultados de estudios previos relacionados al uso de datos funcionales en el marco de los ILS.

## **1.3. Justificación**

En la actualidad, nos encontramos viviendo la cuarta revolución industrial o también llamada industria 4.0. En el contexto global de esta industria surgen nuevos retos relacionados con la metrología, estos entre otros consisten en la obtención de una mayor cantidad de datos y con ello un aumento en su complejidad de análisis.

El empleo de instrumentos de medida cada vez más sofisticados, capaces de proporcionar un mayor número de medidas por unidad de tiempo, hacen que los datos sean de carácter funcional o multivariado, ante esto es natural pensar que surja el desarrollo de nuevas herramientas, técnicas y metodologías que permitan el análisis adecuado de este tipo de datos.

El desarrollo de metodologías dentro del análisis de datos funcionales (FDA), resulta ser de gran utilidad para una gran variedad de técnicas experimentales en diversas áreas como en la química aplicada, la física y la ingeniería, en donde los datos a menudo son de alta dimensionalidad. Este es el caso de gran parte de los datos experimentales obtenidos mediante diferentes y nuevos instrumentos de medición (sensores, espectros, termogramas, etc.).

En [4] se enuncian cuatro razones prácticas e importantes para considerar el análisis de datos desde una perspectiva funcional, las cuales se resumen a continuación:

- Las observaciones funcionales se presentan en contextos aplicados cada vez con mayor frecuencia a medida que las instalaciones automatizadas de recopilación de datos se vuelven accesibles para más investigadores. Además, los procedimientos de suavizado e interpolación pueden producir representaciones funcionales de conjuntos finitos de observaciones.
- Es más natural pensar en algunos problemas de modelado en términos funcionales, aunque solo se disponga de un número finito de observaciones.
- Los objetivos de un análisis pueden ser de naturaleza funcional, por ejemplo, si se usaran datos finitos para estimar una función completa, sus derivadas o los valores de otros funcionales.
- Tener en cuenta consideraciones como la fluidez de los datos multivariados que surgen de los procesos funcionales puede tener importantes implicaciones para sus análisis.

Estudios previos como [5], [6] y [7] muestran la aplicación directa de técnicas FDA, como método de detección de datos funcionales atípicos en base al cálculo de medidas de profundidad, sin embargo, el uso de técnicas FDA ha sido poco común en el marco de los ILS, por lo tanto, surge la necesidad de desarrollar una nueva metodología, desde el contexto FDA, enfocadas a desarrollar herramientas propias de un ILS, para la detección de laboratorios atípicos.

Si bien otros estudios han contribuido al campo del análisis de datos funcionales, la presente investigación pretende aportar una nueva propuesta funcional para contrastar la hipótesis nula de reproducibilidad de igualdad de medias y la hipótesis nula de repetibilidad de igualdad de varianzas, de la extensión funcional de los estadísticos de contraste  $h$  y  $k$  de Mandel y el método de proyecciones aleatorias.

Un Estudio Interlaboratorio se desarrolla bajo la hipótesis nula de que todas las observaciones de los diferentes laboratorios vienen de la misma población, lo que equivale a decir, en el contexto de un modelo ANOVA, que las medias y las desviaciones estándar de todos los laboratorios son aproximadamente iguales para cada nivel de medición [8]. En este caso, se justifica el planteamiento de las hipótesis de Reproducibilidad y Repetibilidad, con fundamento en la extensión funcional de los estadísticos univariantes  $h$  y  $k$  de Mandel.

Por otro lado, en [9], se muestra una propiedad importante que resulta del uso de proyecciones aleatorias, esto es la preservación de la distribución de los datos. Dicha propiedad hace de las proyecciones aleatorias una herramienta importante dentro de la estadística funcional. La principal ventaja es que permite aplicar técnicas y metodologías univariantes en los datos, una vez que estos han sido proyectados, evitando así la aplicación de técnicas complejas en datos que se encuentran en un espacio de dimensión mayor que uno; además si se cumplen determinadas condiciones de regularidad, es posible inferir conclusiones en el espacio de partida.

Por lo tanto, crear una metodología que utilice los estadísticos funcionales  $H$  y  $K$  y que emplee proyecciones aleatorias para convertir estos estadísticos en escalares, se convierte en un proceso factible para realizar ILS con datos de carácter funcional; además al aplicar técnicas de remuestreo Bootstrap Suavizado a curvas y estableciendo un nivel de significancia  $\alpha$ , se puede hallar la distribución y los valores críticos de los estadísticos unidimensionales, por lo que es posible hacer un contraste de hipótesis para detectar diferencia de grupos o laboratorios. Con esto se justifica que la nueva metodología permitirá la detección de laboratorios atípicos dentro de un ILS.

En la práctica, esta nueva propuesta llega a ser de gran utilidad para aquellas industrias que trabajan con mediciones de carácter funcional, que buscan que sus procesos sean precisos y se encuentren bajo control. Además, realizar un ILS con un enfoque funcional, beneficiará no solo a los laboratorios o industrias que participen de este tipo de estudios, sino que la nueva metodología propuesta podría determinar la

precisión de cualquier método de medición, es decir, pretende aportar posibles soluciones a problemas reales de diferentes industrias y en los diferentes laboratorios de análisis.

Finalmente, considerando que en la actualidad tanto las normas ISO-5725 (1994) como ASTM-E691(2004) no contemplan metodologías para realizar Estudios Interlaboratorio cuando los resultados son curvas o datos funcionales [10], se pretende con la presente investigación, dar un posible aporte en el control estadístico de calidad, el cual podría ser considerado por los estándares de la ASTM.

## Capítulo 2

# Enfoque FDA para la detección de atípicos en Estudios Interlaboratorio

### 2.1. Introducción Estudios Interlaboratorio

Los Estudios Interlaboratorio son estudios colaborativos en los que intervienen un conjunto o grupo de laboratorios, encargados de analizar un material específico bajo las mismas condiciones, considerando un proceso experimental bien definido. El estudio se realiza con el objetivo de evaluar la consistencia de los resultados proporcionados por los conjuntos de laboratorios [11].

En [12], se identifican tres tipos de ILS dependiendo de la finalidad con la que se analiza la información o los datos de estudio, pueden identificarse los siguientes estudios:

- Estudios de rendimiento de métodos o estudios de la precisión de un método de medición.
- Estudios de rendimiento o competencia de laboratorio.
- Estudios de certificación de materiales.

La presente investigación toma como punto focal el análisis de un ILS como un estudio de precisión de un método de medición.



El procedimiento para realizar un ILS como un estudio de precisión de un método de medición, se encuentra detallado en las normas ISO 5725 (*Accuracy - Trueness and precision - of measurement methods and results*) [13], y ASTM E-691 (*Standard practice for conducting an interlaboratory study to determinate the precision of a test method*) [14] y se profundizará en la sección 3.1 de la presente memoria de Tesis, mediante el estudio de la precisión, repetibilidad y reproducibilidad. A continuación, se presentan brevemente algunas consideraciones generales para realizar un ILS.

### **2.1.1. Consideraciones generales para realizar un Estudio Interlaboratorio**

Si bien existen varios enfoques para realizar un ILS, algunos problemas son comunes a todos los tipos de estudio. Según resultados presentados por *Hund et al.* en [15], existen ciertas consideraciones generales para realizar un ILS, las mismas engloban diferentes problemáticas que se pueden presentar. Estas consideraciones pueden resumirse de la siguiente manera:

#### *1. Requisitos preliminares.*

Los participantes que formen parte del estudio interlaboratorio deben anticipar todos los sistemas de aseguramiento de calidad dentro del laboratorio.

#### *2. Definiciones.*

Se debe considerar las definiciones de precisión de repetibilidad y reproducibilidad establecidas en las normas ISO [13]. Según estas directrices, un estudio de precisión tiene como finalidad evaluar la repetibilidad y reproducibilidad de un método de análisis.

#### *3. Condiciones de repetibilidad.*

Las condiciones de repetibilidad se definen como las condiciones en las que los resultados de análisis experimentales son independientes y se obtienen utilizando un mismo método en muestras idénticas analizadas por un mismo laboratorio y por un mismo operador, haciendo uso de un mismo equipo en cortos intervalos de tiempo [10].

#### 4. *Condiciones de reproducibilidad.*

Las condiciones de reproducibilidad se definen como las condiciones en las que los resultados de los análisis experimentales se obtienen con el mismo método aplicado a muestras idénticas, por diferentes laboratorios con diferentes operadores usando instrumentos diferentes [10].

#### 5. *Selección del material*

Los materiales o niveles de medición deben ser seleccionados de tal manera que la muestra sea representativa, es importante considerar su factibilidad del material para ser medido.

#### 6. *Homogeneidad de materiales*

Se debe verificar la homogeneidad de materiales, de lo contrario se podrán presentar errores significativos en los resultados. La precisión de los métodos en mejor en cuanto mayores sean los requisitos de homogeneidad de los materiales.

#### 7. *Replicación de experimentos*

Es importante establecer un número de mediciones o réplicas adecuado para el material de estudio. La norma ASTM E-691 [14] recomienda tomar como mínimo tres réplicas de cada medición, la norma ISO 5725 [13] no hace referencia a un número específico de réplicas por laboratorio, pero indica que al menos cada laboratorio debe entregar dos réplicas.

#### 8. *Detección de laboratorios atípicos.*

Dentro de los principales objetivos de un ILS, la detección de laboratorios atípicos es fundamental, pues resulta importante identificar aquellos laboratorios que proporcionan resultados diferentes, además de detectar y excluir los datos inconsistentes. La finalidad radica en lograr estimar correctamente la precisión y exactitud de los métodos de estudio empleados por los laboratorios [16].

### 2.1.2. **Tratamiento de datos atípicos en Estudios Interlaboratorio**

Previo a la aplicación de técnicas que permiten identificar los laboratorios atípicos, es importante identificar los tipos de atípicos que pueden presentarse en un ILS.

Las normas ASTM E-691 e ISO 5725 indican que en un ILS **puede** presentarse y diferenciarse tres tipos de valores atípicos (tipo I, II y III) [15].

Los valores atípicos de **tipo I** corresponden a las observaciones o datos atípicos presentes en las mediciones y réplicas obtenidas dentro de un mismo laboratorio. Los de **tipo II** hacen referencia a las medias atípicas obtenidas cuando se comparan las medias de los diferentes laboratorios para determinar si existen diferencias en las **mismas, y por último** los valores atípicos de **tipo III** hacen referencia a las desviaciones estándar atípicas obtenidas cuando se identifican diferencias al comparar las medidas de dispersión y desviaciones de los diferentes laboratorios [15]. Es importante recalcar que el objetivo de un ILS es detectar laboratorios atípicos y no únicamente observaciones atípicas, por lo que el punto focal de estudio son los valores atípicos de tipo II y tipo III.

### **Pruebas de consistencia**

Los métodos utilizados para la detección de laboratorios atípicos han consistido principalmente en técnicas estadísticas univariantes relacionadas con la estimación de la variabilidad presente en el proceso experimental o de medición, además de la aplicación de varios test de contrastes de hipótesis univariantes y métodos gráficos.

Según [10], existen dos casos que deben ser evaluados para determinar la presencia de laboratorios atípicos.

- *Caso 1:* cuando las medidas realizadas por un laboratorio presentan desviaciones significativas con respecto a las obtenidas por otros laboratorios.
- *Caso 2:* cuando se identifican observaciones atípicas dentro de un mismo laboratorio en específico.

A continuación, se describen los estadísticos y las pruebas clásicas recomendadas por la norma ISO 5725 (ISO-5725, 1994) y la norma ASTM E-691 (ASTM-E691, 2004), para evaluar los casos 1 y 2.

### **Estadísticos $h$ y $k$ de Mandel**

Cuando se pretende identificar laboratorios inconsistentes o atípicos (Caso 1), se puede considerar la aplicación de test de contrastes de hipótesis de diferencia de

valores medios (tipo II) y desviaciones estándar (tipo III) en base a los estadísticos  $h$  y  $k$  de Mandel. Estos estadísticos corresponden a técnicas y herramientas gráficas univariantes que permiten evaluar la calidad de los laboratorios comparando la variabilidad presente en los procesos de medición de un laboratorio individual con el resto de los laboratorios [12]. Si bien existen otros métodos gráficos para la detección de laboratorios atípicos como, los gráficos de control, diagramas de box plot [17], gráficos de Youden [18], entre otros, el presente estudio profundizará únicamente en los estadísticos  $h$  y  $k$  de Mandel ya que son los cimientos claves para el desarrollo de la presente investigación.

### Estadístico $h$ de Mandel

Sea  $(x_1, x_2, \dots, x_L)$  una muestra de  $L$  observaciones, donde  $L$  es el número de laboratorios. Los  $x_l$  son modelados como realizaciones de variables aleatorias independientes e idénticamente distribuidas  $X_l$  con  $l = 1, \dots, L$ ; acorde a una distribución normal  $N(\mu, \sigma^2)$ . Para el caso en el que  $L$  laboratorios realizan  $n$  réplicas sobre un material, se define el estadístico  $h$  de Mandel para el laboratorio  $l$ –ésimo mediante la expresión:

$$h_l = \frac{d_l}{S_{\bar{x}}} = \frac{\bar{x}_l - m}{\sqrt{\sum_{l=1}^L \frac{(\bar{x}_l - m)^2}{L-1}}}.$$

Donde:

- $\bar{x}_l = \frac{\sum_{i=1}^n x_i}{n}$  es la media de la magnitud medida a partir de todas las réplicas realizadas sobre el mismo material en cada laboratorio  $l = 1, \dots, L$ .
- $m = \frac{\sum_{l=1}^L \bar{x}_l}{L}$  es el promedio de las medias obtenidas por todos los laboratorios (media global).
- $S_{\bar{x}}$  es la desviación estándar muestral de  $\bar{x}_l$ .

En este caso nos referimos a un enfoque paramétrico ya que asume la normalidad de los datos de estudio; y por ende la distribución y los valores críticos del estadístico pueden ser aproximados.

Según Wilrich [19], la distribución del estadístico  $h$  aproxima a una distribución  $t$  de Student con  $L - 2$  grados de libertad.

$$h \approx t_{L-2}$$



En consecuencia, el valor crítico del estadístico  $h$  de Mandel, se define mediante:

$$h_{l;1-\frac{\alpha}{2}} = \frac{(L-1)t_{L-2;1-\frac{\alpha}{2}}}{\sqrt{L(t_{L-2;1-\frac{\alpha}{2}}^2 + L-2)}}$$

Donde:

- $t_{L-2;1-\frac{\alpha}{2}}$ , corresponde al cuantil  $(1 - \frac{\alpha}{2})$  de la distribución  $t$  de Student con  $v = L - 2$  grados de libertad.

Con lo descrito anteriormente, un laboratorio es detectado como inconsistente mediante el estadístico  $h$  cuando el valor observado del estadístico  $h_l$ , para algún laboratorio  $l$ , es mayor que el valor crítico a un nivel de significación  $\alpha$ , es decir, cuando  $h_l \geq h_{l;1-\frac{\alpha}{2}}$  o  $h_l \leq h_{l;\frac{\alpha}{2}}$ .

### Estadístico $k$ de Mandel

Para el caso en el que se dispongan de  $l = 1, 2, \dots, L$  laboratorios de ensayo que realicen  $n$  réplicas de un determinado experimento sobre un mismo material, se define el estadístico  $k$  de Mandel para el laboratorio  $l$ -ésimo mediante la expresión.

$$k_l = \frac{S_i}{S_r} = \frac{\sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}}}{\sqrt{\frac{\sum_{i=1}^L s_i^2}{L}}}$$

De manera similar en [19], se puede ver a detalle la aproximación de la distribución teórica y valor crítico del estadístico  $k$ , bajo las siguientes consideraciones:

Sean  $(S_1^2, S_2^2, \dots, S_L^2)$  las  $L$  varianzas muestrales, cada una basada en  $n$  valores observados. Bajo la hipótesis de que los valores de las observaciones  $x_{ji}; j = 1, \dots, L; i = 1, \dots, n$  son realizaciones de variables aleatorias  $X_{ji}$  donde cada  $j$  es idénticamente distribuida e independiente acorde a una distribución normal  $N(\mu_i, \sigma^2)$ , la expresión  $S_j^2/\sigma^2; j = 1, \dots, L$  sigue una distribución  $\chi^2/v$  con  $v = n - 1$  grados de libertad.

El estadístico  $k$  aproxima a una distribución  $F$  de Snedecor con  $v_1 = (L - 1)(n - 1)$  y  $v_2 = n - 1$  grados de libertad.

$$k \approx F_{v_1;v_2}$$

El valor crítico del estadístico  $k$  de Mandel, viene dado por:

$$k_{l,n;1-\alpha} = \sqrt{\frac{L}{\left(1 + \frac{L-1}{F_{E_1,n_2;\alpha}}\right)}},$$

donde  $F_{v_1,v_2;\alpha}$  es el  $\alpha$ -cuantil de la distribución  $F$  con  $v_1 = (L-1)(n-1)$  y  $v_2 = n-1$  grados de libertad.

Bajo estas consideraciones, un laboratorio es detectado como inconsistente mediante el estadístico  $k$  cuando el valor observado del estadístico  $k_l$ , para algún laboratorio  $l$ , es mayor que el valor crítico a un nivel de significación  $\alpha$ , es decir, cuando  $k_l \geq k_{l,1-\alpha}$ .

### Test de Cochran y Grubbs

Cuando se pretende identificar las observaciones atípicas obtenidas por un laboratorio (Caso 2, Tipo I), es necesario detectar dichas observaciones para analizar su influencia en todo el conjunto de datos dentro del laboratorio; y por ende concluir si efectivamente el laboratorio es inconsistente o no.

Es importante considerar que si dentro del análisis se encuentran observaciones atípicas para un único material, esto no implica que el laboratorio sea atípico. Sin embargo, si los valores inconsistentes se presentan en diferentes materiales dentro de un mismo laboratorio, entonces esto podría ser una señal de que la varianza intra-laboratorio es demasiado alta o que existe un alto sesgo en el laboratorio y por ende todas las observaciones correspondientes a ese laboratorio deben ser omitidas del análisis [8].

Dentro de las pruebas más utilizadas para la detección de observaciones atípicas, destacan los Test de Cochran mediante contrastes de varianzas y Grubbs para contraste de medias. En un ILS, estos test son utilizados como complemento a los estadísticos  $h$  y  $k$  de Mandel y en conjunto permiten detectar atípicos tipo II y tipo III.

### Test de Grubbs

Si se desea determinar si la observación  $X_{\max} = \max(X_1, \dots, X_n)$  o  $X_{\min} = \min(X_1, \dots, X_n)$  son valores atípicos, se utiliza la prueba de Grubbs [20]. Los estadísticos de esta prueba vienen dados respectivamente por las siguientes expresiones:

$$G = \max_{i=1,\dots,n} H_i = \max_{i=1,\dots,n} \frac{X_i - \bar{X}}{S} = \frac{X_{\max} - \bar{X}}{S}$$

$$G = \max_{i=1,\dots,n} -H_i = \min_{i=1,\dots,n} \frac{X_i - \bar{X}}{S} = \frac{X_{\min} - \bar{X}}{S}$$

La aproximación del valor crítico para esta prueba se detalla en [19], y se define de la siguiente manera:

$$g_{n;1-\alpha} \approx \frac{(n-1)t_{n-2;1-\frac{\alpha}{n}}}{\sqrt{n(n-2+t_{n-2;1-\frac{\alpha}{n}})}},$$

donde  $t_{n-2;1-\frac{\alpha}{n}}$  es el cuantil  $(1 - \frac{\alpha}{n})$  de la distribución  $t$  con  $v = n - 2$  grados de libertad.

Cuando intervienen  $L$  laboratorios y  $n$  réplicas en el estudio, se define el estadístico  $g_{L;1-\alpha}$  considerando las medias de los resultados de cada laboratorio y la media global de los promedios de los laboratorios.

Si un laboratorio es detectado con resultados atípicos, aplicando el estadístico  $h$ , o la prueba de Grubbs a diferentes niveles dentro de un laboratorio, indicaría que el sesgo del laboratorio es alto [8].

### Test de Cochran

La prueba de Cochran [21], es una prueba de valores atípicos de una cola porque examina solamente el valor más alto en una serie de varianzas [8]. Como se recalcó anteriormente, si un laboratorio tiene varios valores atípicos para diferentes materiales, entonces, el laboratorio puede ser considerado como atípico y se excluye del estudio. Para determinar, mediante la prueba de Cochran, si la mayor varianza  $S_{\max}^2 = \max(S_1^2, S_2^2, \dots, S_L^2)$  es un valor atípico, se utiliza el estadístico de prueba:

$$C = \frac{S_{\max}^2}{\sum_{j=1}^L S_j^2} = \frac{1}{L} \max_{i=1,\dots,L} \frac{S_j^2}{\bar{S}^2},$$

para esta prueba, en [19], el valor crítico se aproxima mediante la siguiente expresión:

$$c_{l,n;1-\alpha} \approx \frac{1}{(1 + (L-1)F_{v_1,v_2;\alpha_L})},$$

donde  $F_{v_1,v_2;\frac{\alpha}{L}}$  es el  $\frac{\alpha}{L}$ -cuantil de la distribución  $F$  con  $v_1 = (L-1)(n-1)$  y  $v_2 = n-1$  grados de libertad.

Los estadísticos y las pruebas clásicas univariantes mencionados son recomendados por la norma ISO 5725 y ASTM E-691. Sin embargo, para el ámbito funcional estas normas no han establecido ninguna propuesta metodológica para la detección de laboratorios atípicos en ILS. Entre los pocos resultados; una primera propuesta, dentro del análisis de datos funcionales, fue presentada por Flores en [3, 10], a partir de la extensión funcional de los estadísticos  $h$  y  $k$  de Mandel y el uso de la norma  $L_2$ . Este estudio sirve como referente para realizar un estudio comparativo con la nueva propuesta desarrollada en el presente trabajo de investigación y mediante la comparación de los métodos, identificar los posibles escenarios donde se presentan los mejores resultados en la aplicación de cada propuesta. Las metodologías y los casos de aplicación establecidos en el ámbito FDA se amplían en las secciones 3.3 y 2.3 respectivamente.

## 2.2. Técnicas FDA aplicadas en un Estudio Interlaboratorio

Esta sección tiene como finalidad dar una introducción general al análisis de datos funcionales y definir algunas técnicas y herramientas FDA, con la finalidad de establecer un marco teórico referencial, a partir del cual se desarrolla la nueva metodología propuesta para la detección de laboratorios atípicos en ILS. Una vez definido el marco teórico se presenta a detalle dos casos reales de estudio donde se aplica la nueva propuesta metodológica y en particular en uno de ellos, se realiza un estudio preliminar en base a técnicas y herramientas FDA que permite obtener las primeras nociones del comportamiento del caso de estudio, para ello se utilizan paquetes enfocados en el Análisis de datos funcionales del software estadístico R.

### 2.2.1. Análisis de Datos Funcionales

En esta sección se presenta una breve introducción al análisis de datos funcionales, así como también las principales definiciones sobre las que se desenvuelve la presente investigación, tales como: variable aleatoria funcional, espacio funcional, medidas de tendencia central, medidas de dispersión, y profundidad de curvas.



En cuanto a métodos y técnicas de estadística inferencial, en el ámbito funcional, se dan a conocer aquellas que permiten detectar diferencias entre grupos y curvas atípicas, como el método de detección de curvas atípicas mediante técnicas de remuestreo Bootstrap, análisis de la profundidad de los datos con proyecciones aleatorias y ANOVA funcional. La aplicación y uso de estas técnicas han sido agilizadas gracias a los paquetes disponibles en el software estadístico R tales como fda, fda.usc e ILS.

A continuación, se muestran algunas definiciones, en las cuales se representan los datos funcionales correspondientes a los casos de estudio reales que abarca esta tesis.

## Definiciones

### Espacio Funcional

El espacio de Hilbert separable,  $L^2(T)$ , está definido por las funciones cuadrado-integrables en el intervalo  $T = [a, b] \subset \mathbb{R}$  :

$$L^2(T) = \left\{ \mathcal{X} : T \rightarrow \mathbb{R}; \int_T \mathcal{X}(t)^2 dt < \infty \right\}$$

con un producto interno definido por:

$$\langle \mathcal{X}(t), \mathcal{Y}(t) \rangle = \int_T \mathcal{X}(t) \mathcal{Y}(t) dt$$

### Variable Funcional

Una variable aleatoria  $X$  es una variable funcional si toma valores en  $L^2(T)$ .

### Datos Funcionales

Se define un conjunto de datos funcionales  $S_n = \{\mathcal{X}_i\}_{i=1}^n$  o  $\mathcal{X}_1, \dots, \mathcal{X}_n$ , como la observación de  $n$  variables aleatorias funcionales  $\mathbf{X}_1, \dots, \mathbf{X}_n$  con igual distribución que  $\mathbf{X}$ .

Los datos funcionales se refieren a datos para los cuales la  $i$ -ésima observación es una función real,  $x_i(t)$  con  $i = 1, \dots, n$  ;  $t \in T$ , donde  $T$  es un intervalo real y, por lo

tanto, cada  $x_i$  es un punto en algún espacio funcional, en este caso, espacio de Hilbert. En la práctica, los valores observados son considerados como discretizaciones.

### **Análisis exploratorio con datos funcionales**

En [22] se definen las medidas de dispersión y de tendencia para el análisis exploratorio de datos en el ámbito funcional de la siguiente manera:

#### ***Media Funcional***

Sea  $\{\mathcal{X}_1, \dots, \mathcal{X}_n\}$ , un conjunto de  $n$  datos funcionales definidos en  $t \in T$ ;  $T \subset \mathbb{R}$ . La media funcional muestral se define como sigue:

$$\bar{\mathcal{X}}(t) = \frac{\sum_{i=1}^n \mathcal{X}_i(t)}{n}.$$

#### ***Varianza Funcional***

Bajo las mismas consideraciones; se define la varianza funcional como:

$$s^2(t) = \frac{\sum_{i=1}^n (\mathcal{X}_i(t) - \bar{\mathcal{X}}(t))^2}{(n-1)},$$

y la desviación estándar

$$s(t) = \sqrt{s^2(t)}.$$

### **Medidas de profundidad**

La noción de profundidad fue introducida inicialmente para datos multivariados con el fin de generalizar las ideas de estadísticas de orden en dimensiones mayores a uno. Luego la profundidad en el contexto del análisis funcional se definió como sigue a continuación:

Sea  $X$  una variable aleatoria con función de distribución  $F \in \mathcal{R}^d$ . Una medida de profundidad estadística es una función que asigna a cada punto  $x$  un valor no negativo  $D_F(x) \in [0, 1]$ .

$D_F(x)$  representa el grado de centralidad del dato  $x$  respecto a la distribución  $F$ . Los valores cercanos a uno son considerados centrales o más profundos y los valores de la muestra más pequeños y cercanos a cero corresponden a los datos de menor profundidad. El dato que maximiza  $D_F(x)$ , es la mediana de la muestra.

Entre las medidas de profundidad multivariadas más utilizadas se tienen, por ejemplo, la profundidad de Tukey [23]; la profundidad de Liu [24] y la profundidad de Zuo y Serfling [25] que muestran propiedades claves a satisfacer por una medida de profundidad, definiciones, propiedades y aplicaciones de las profundidades multivariadas. Adicionalmente, Fraiman y Muniz [5] y Cuevas [6, 7] extienden la definición de profundidad, al caso funcional.

En la subsección 2.2.3 se realiza una breve revisión del uso de la profundidad funcional **en base a** proyecciones aleatorias para la detección de curvas atípicas, un enfoque primordial a partir del cual se desarrolla la nueva propuesta metodológica para la detección de laboratorios atípicos en un ILS.

### 2.2.2. Técnicas de remuestreo bootstrap suavizado

El método de estimación bootstrap, basado en remuestreo, permite cuantificar la incertidumbre de alguna estadística de interés cuando la distribución muestral de la misma es desconocida [26]. El método fue introducido por Efron en 1979 [27].

El bootstrap clásico consiste en muestrear a partir de la función de distribución empírica  $F$ , en cambio, cuando la distribución poblacional  $F$  es continua, entonces se puede hacer uso de la función de densidad asociada con alguna estimación de tipo **kernel**. En otras palabras, para el caso funcional se debe considerar una versión suavizada  $\bar{F}$  y para ello se requiere utilizar un estimador de la función de densidad y remuestrear sobre dicho estimador.

El esquema de bootstrap para datos funcionales propuesto por Cuevas en [6] y [26] se presenta de la siguiente manera.

1. Sean  $X_1(t), X_2(t), \dots, X_n(t)$  los datos funcionales muestrales originales y sea la estadística de interés  $T = T(X_1(t), \dots, X_n(t))$ .
2. Se obtienen  $B$  realizaciones a partir de remuestreo, como  $\mathbf{X}^b = (X_1^b(t), X_2^b(t), \dots, X_n^b(t))$  con  $b = 1, \dots, B$ .

3. Generar  $Z_i^b, i = 1, \dots, n$ , independientes de  $X_i^b$ , tales que  $Z_i^b \sim K_d$  y hacer

$$Y_i^b = X_i^b + Z_i^b$$

Esto es, se obtienen muestras suavizadas como  $\mathbf{Y}^b = (Y_1^b, Y_2^b, \dots, Y_n^b), b = 1, \dots, B$ . Nótese que generalmente  $X_i^b$  es una versión discretizada del dato funcional, esto quiere decir que  $X_i^b = (X_i^b(t_1), \dots, X_i^b(t_d))$  y por lo tanto la variable  $Z^b$  es  $d$ -dimensional, donde  $d$  depende del número de puntos de observación, así  $Z_i^b = (Z_i^b(t_1), \dots, Z_i^b(t_d))$ .

4. Finalmente se determina la estadística  $\hat{T}^b = T(Y_1^b(t), Y_2^b(t), \dots, Y_n^b(t))$  correspondiente a la  $b$ -ésima muestra bootstrap.

Este método puede ser fácilmente aplicado con la función *boot.sim.set* del paquete *ILS* [28].

### Bandas de confianza bootstrap

A partir del método bootstrap suavizado se puede construir las bandas de confianza bootstrap que tienen como objetivo evaluar la precisión de un estimador de localización [10].

El procedimiento de bootstrap suavizado para la construcción de bandas de confianza se describe en [6] de la siguiente manera:

1. Se obtienen  $B$  muestras bootstrap a partir de  $\mathcal{S}_n : \mathcal{S}_n^b = \{Y_i^b\}_{i=1}^n$  donde  $Y_i^b(t) = \mathcal{X}_i^b(t) + Z(t)$ , con  $b = 1, 2, \dots, B$  e  $i = 1, 2, \dots, n$  donde  $\mathcal{X}_i^b(t)$  es seleccionado al azar y  $Z(t)$  corresponde a una variable distribuida normalmente con media cero y matriz de covarianza  $\gamma \Sigma_{\mathcal{S}_n}$ . Donde  $\Sigma_{\mathcal{S}_n}$  es la matriz de varianza-covarianza de la muestra y  $\gamma$  es un parámetro de suavización que controla la cantidad de la variabilidad de las nuevas muestras bootstrap (varios autores recomiendan  $\gamma = 0.05$ ).
2. A partir de la muestra original,  $\mathcal{X}_1(t), \dots, \mathcal{X}_n(t)$ , y, de las sucesivas muestras bootstrap  $Y_1^b, \dots, Y_n^b$ , se calcula las estimaciones del estadístico de interés,  $\hat{T}(\mathcal{S}_n)$  y  $\hat{T}(\mathcal{S}_n^b)$ , respectivamente.
3. Luego, se obtienen las distancias  $\{d(\hat{T}(\mathcal{S}_n), \hat{T}(\mathcal{S}_n^b))\}_{b=1}^B$  utilizando la norma  $L^2$ .

4. Finalmente, una banda de confianza bootstrap, correspondiente a un nivel de confianza  $(1 - \alpha)$  es calculada a partir del cuantil  $(1 - \alpha)$  de las distancias  $\{d(\hat{T}(\mathcal{S}_n), \hat{T}(\mathcal{S}_n^b))\}_{b=1}^B$ .

La función *fdata.bootstrap* del paquete *fda.usc* [29], permite estimar la media, mediana, la moda y la varianza marginal funcional y calcular sus respectivas bandas de confianza.

### 2.2.3. Método de detección de atípicos basado en proyecciones aleatorias

En estudios previos, se ha realizado una primera aproximación al contexto FDA, donde se aplican técnicas del análisis exploratorio y procedimientos de detección de valores atípicos en base al cálculo de profundidad de datos [30]. Además, se han planteado los enfoques más comunes para el cálculo de la profundidad funcional, como son la profundidad mediana [5], la profundidad modal [6] y la profundidad de proyecciones aleatorias [7]. La aplicación directa de estas técnicas FDA permite identificar aquellas curvas atípicas, además de contrastar si las medias funcionales o las desviaciones estándar de los laboratorios son iguales.

Este estudio se enfoca en la detección de curvas atípicas en base al cálculo de la profundidad funcional usando proyecciones aleatorias. A continuación, se define el concepto de proyecciones aleatorias y en particular se detalla la metodología propuesta en [30] que permite detectar curvas atípicas.

#### Proyecciones aleatorias

Las proyecciones aleatorias son usadas principalmente como una técnica para reducir la dimensionalidad de una curva. En [26] se define como sigue a continuación:

Sea  $X_1(t), X_2(t), \dots, X_n(t)$  una curva funcional con  $t$  discretizaciones de tiempo en un intervalo  $[a, b]$ . Sea  $w$  una dirección aleatoria independiente de  $X_i$  y sea  $D$  una medida de profundidad univariante. Considerando que cada dato funcional  $X_i$  pertenece al espacio de Hilbert  $L^2[a, b]$ , las proyecciones unidimensionales de cada dato

hacia la dirección  $w$  se obtienen mediante el producto interno,

$$PA_i^w = \langle w, X_i \rangle = \int_a^b w(t) X_i(t) dt ; i = 1, \dots, n.$$

Esta definición lleva a una medida aleatoria de profundidad basada en el rango de las proyecciones sobre una dirección aleatoria. Entonces, se define la medida de profundidad proyección aleatoria PAD para el  $i$ -ésimo dato como:

$$PAD_i = PAD(\mathcal{X}_i) = D(PA_i^w)$$

Una profundidad poblacional univariante dotada por  $D$  de un dato  $X_i$  con respecto a un elemento aleatorio  $w$ , puede definirse como el cuantil correspondiente al valor  $\langle w, X_i \rangle$  en la distribución de  $\langle w, X \rangle$ . La medida de profundidad univariante empleada en la presente metodología corresponde a la profundidad de Tukey.

La profundidad de Tukey de un punto  $x$ , se define en [5] como:

$$TD(x) = \inf_S \{F(S) : x \in S\}$$

donde  $S$  es un semi espacio de Hilbert (generalización de un espacio de Hilbert en el análisis funcional).

Para el caso unidimensional, la profundidad de Tukey considera como profundidad más alta al cuantil del valor de la media del conjunto de datos analizado, por lo tanto, se define de la siguiente manera:

$$TD(x) = \min \{F(x), 1 - F(x^-)\}$$

Donde  $F$  corresponde a la función de distribución empírica.

Teniendo en cuenta la definición anterior, para este caso la profundidad de una proyección aleatoria para cada  $i=1, \dots, n$ ; se define explícitamente como:

$$TD(PA_i^w) = \min \{F(PA_i^w), 1 - F(PA_i^w)\}$$

donde,

$$D(PA_i^w) = F(PA_i^w) = \frac{1}{n} \sum_{k=1}^n \mathbb{1}(PA_k^w \leq PA_i^w).$$

### *Consideraciones generales al usar proyecciones aleatorias*

El uso de proyecciones aleatorias puede influir directamente en la estabilidad del proceso en el que son empleadas; uno de los principales inconvenientes se debe a que resume la información de toda una curva en un valor proyectado unidimensional, lo que implica cierta pérdida de información. Sin embargo, como se mencionó anteriormente, su utilidad radica en que permite aplicar técnicas y metodologías univariantes en los datos, una vez que estos han sido proyectados, evitando así la aplicación de técnicas complejas en datos que se encuentran en un espacio de dimensión mayor que uno.

Otro inconveniente, por ejemplo, se da cuando las proyecciones aleatorias intervienen como parte de un proceso de reducción de dimensionalidad de datos funcionales, para realizar una prueba de significación en el campo real. Al usar proyecciones elegidas al azar estamos aceptando inestabilidad del proceso; es decir, puede que la hipótesis nula sea rechazada inicialmente y en ocasiones posteriores no, además de disminuir la potencia de la prueba. Para solucionar este inconveniente, en [31] y [32] se propone tomar un número de proyecciones aleatorias mayor a 1 y en base a ellas se redefine los criterios de rechazos de las pruebas aplicadas.

En [33] se comprueba que el aumento del número de proyecciones aleatorias permite aumentar la potencia de la prueba; con esto es razonable pensar que el número de direcciones aleatorias debe coincidir con el número de puntos en los que se han medido las trayectorias. Sin embargo, dentro del mismo estudio se realiza un estudio de simulación donde se generalizó para la mayoría de los casos el siguiente resultado con respecto al número de curvas a ser proyectadas.

- Si el número de curvas  $N$  es suficientemente grande entonces el número de proyecciones óptimas puede aproximarse por el límite inferior de  $3 \times \min(N, 50)$ .
- Si el número de curvas  $N$  es pequeño entonces el número de proyecciones óptimas puede aproximarse por el límite superior de  $3 \times \min(N, 50)$ .

Cuando se cuenta con una colección de direcciones aleatorias  $\{W_l\}_{l=1}^K$ , la medida de profundidad para datos funcionales se calcula haciendo uso de todas las proyecciones consideradas en esas direcciones y viene representada por el valor medio de las profundidades de los puntos proyectados, por lo tanto, la profundidad de

proyección aleatoria de una curva  $\mathcal{X}_i$  se define en [7] por:

$$PAD(\mathcal{X}_i) = \frac{1}{K} \sum_{l=1}^K D(PA_i^{wl}).$$

### Metodología para la detección atípicos basado en profundidades

La metodología para la detección de curvas atípicas presentada en esta sección fue propuesta en [30]. En el ámbito de los datos funcionales un dato es considerado como atípico cuando tiene una profundidad baja en relación con el resto de las curvas, lo que implica una distinción en magnitud y forma [10]. La curva con profundidad más alta corresponde a la mediana funcional y es el punto central referencial a partir del cual son obtenidas las profundidades del resto de curvas. Este resultado se aplicará como parte de la metodología para la detección de laboratorios atípicos en un ILS.

El presente estudio tomará únicamente el enfoque en base a proyecciones aleatorias. La metodología de detección de atípicos en base a profundidades considera como parámetro un porcentaje de recorte de curvas profundas y varios tipos de profundidad como la mediana, modal o proyecciones aleatorias. Ésta se describe en [3] y [30] en los siguientes pasos:

1. Se obtiene el cálculo de las profundidades funcionales  $\{D(\mathcal{X}_i)\}_{i=1}^n$  correspondientes al conjunto de datos,  $\mathcal{X}_1(t), \dots, \mathcal{X}_n(t)$ ; en este caso empleando la medida de profundidad con proyecciones aleatorias.
2. A partir de los datos funcionales originales  $\{\mathcal{X}_i\}_{i=1}^n$ , con  $b = 1, \dots, B$ ; se obtienen  $B$  remuestras con el método bootstrap suavizado  $Y^b(t), \dots, Y_n^b(t)$ . El objetivo es estimar el valor de corte  $C$  tal que  $P(D(\mathcal{X}) < C) = \alpha$ ; para ellos es necesario aproximar la distribución de la profundidad de los datos originales, mediante el remuestreo.

Es importante considerar que cada remuestra debe ser generada de manera consistente, por lo que esta metodología establece un porcentaje de recorte de curvas menos profundas que son excluidas de los datos originales previo al remuestreo.

3. Posteriormente se realiza el cálculo del valor  $C^b$  con  $b = 1, \dots, B$  para cada remuestra bootstrap, este valor correspondiente al percentil empírico de la



distribución de las profundidades,  $\{D(Y_i^b)\}_{i=1}^n$ . El valor del percentil  $C^j$  es elegido para que en ausencia de valores atípicos, el porcentaje de las observaciones que son correctas pero a la vez son mal etiquetadas e identificadas como valores atípicos sean aproximadamente iguales a una pequeña proporción  $\alpha$ .

4. Finalmente, se obtiene un valor representativo  $C$ , correspondiente a la mediana de los diferentes  $B$  valores de  $C^b$ . Una curva es considerada como atípica si cumple con el siguiente criterio  $D(\mathcal{X}_{i_j}(t)) \leq C$ , y es eliminada de la muestra original.

La aplicación de esta metodología se ve facilitada gracias a la librería *fda.usc* [29] incluida en el software estadístico R que incluye la función *outliers.depth.trim*, de igual manera, el paquete *ILS* contiene la función *outliers.ils* [28]. Ambas opciones se basan en la metodología ya descrita [30]. Los parámetros que reciben estas funciones son similares y en particular la presente investigación hace uso de la segunda opción mencionada.

#### 2.2.4. Pruebas ANOVA Funcionales

Las pruebas ANOVA funcionales o FANOVA permiten determinar la diferencia que existe entre los resultados obtenidos por un conjunto de grupos, en este caso los datos corresponden a curvas funcionales. En esta sección se habla de dos pruebas FANOVA. La primera es una prueba ANOVA funcional para un factor y la segunda es una prueba ANOVA funcional vía proyecciones aleatorias.

##### ANOVA de un factor

El método de Análisis de la Varianza Funcional (FANOVA) propuesto por [33] se define bajo las siguientes consideraciones:

Sea  $\mathcal{X}_{ij}(t), i = 1, \dots, L; j = 1, \dots, n_i$ ;  $L$  grupos de funciones aleatorias independientes definidas en un intervalo compacto  $T = [a, b]$  con  $n_i$  curvas en cada uno de ellos. Denotamos a  $\bar{X}_i$  como la media funcional muestral para el grupo  $i, t \in T$ . El modelo FANOVA de una vía para  $k$  muestras independientes de datos funcionales contrasta

las siguientes hipótesis

$$H_0 : \mu_1(t) = \dots = \mu_L(t)$$
$$H_1 : \exists k, m \quad \mu_k(t) \neq \mu_m(t),$$

donde el estadístico de contraste, que permite medir la variabilidad entre grupos, es el siguiente:

$$V_n = \sum_{k < j} n_i \|\bar{X}_k - \bar{X}_j\|$$



La introducción de este estadístico evita el supuesto de homocedasticidad requerido en el ANOVA usual, siempre y cuando se cumplan lo siguiente:

- Que  $n_i, n \rightarrow \infty$ , suponiendo  $p_i = n_i/n > 0$ , para  $i = 1, \dots, L$ .
- Las observaciones  $X_{ij}(t)$  corresponden a  $L$  muestras independientes de tamaño  $n_i$ , con  $i = 1, \dots, L$ , provenientes de  $L$  procesos con media cero y covarianza  $\text{Cov}_i(s, t)$ .

Para calcular el valor crítico del estadístico  $V_n$ , se utilizan técnicas de remuestreo (bootstrap suavizado), lo que permite, considerar la posible heterocedasticidad entre grupos. La hipótesis nula es rechazada si el valor observado es  $V_n > c_v$ ; donde  $c_v$  es el valor crítico calculado a partir de la distribución bootstrap.

### ANOVA funcional vía proyecciones aleatorias

Esta alternativa considera la aplicación de proyecciones aleatorias y fue propuesto por Alberto Cuesta en 2010 [34]. El procedimiento se resume en dos etapas:

1. Proyectar en direcciones aleatorias los datos funcionales a datos univariados.
2. Resolver un ANOVA simple y apropiado en el caso real.

En resumen, el estudio propuesto por *Cuesta et al.* [34] se sustenta en un resultado importante, el cual establece que si se tienen dos distribuciones diferentes, alguna de ellas determinada por sus momentos, y se elige aleatoriamente una marginal de ellas, entonces las distribuciones de probabilidad son diferentes [26]. Adicionalmente se consideran  $k > 1$  proyecciones aleatorias y se contrasta la hipótesis nula para

cada proyección, ajustando a través de bootstrap los  $p$ -valores obtenidos (prueba no paramétrica).

El caso de estudio a resolver es un ANOVA funcional para dos factores, el mismo que puede ser extendido a varios factores y se enuncia de la siguiente manera:

Sean números  $R, S \in \mathbb{N}$  y supóngase que para todos  $r = 1, \dots, R$  y  $s = 1, \dots, S$ , existen  $X_i^{r,s}, i = 1, \dots, n_{r,s} \in \mathbb{N}$  funciones aleatorias tales que

$$X_i^{r,s} = m(t) + f^r(t) + g^s(t) + h^{r,s}(t) + \gamma(t)Y_i^{r,s} + \epsilon_i^{r,s}(t), \quad t \in [0, 1]$$

donde:

- $m$  es una función no aleatoria y describe la forma completa del proceso.
- $f^r, g^s, h^{r,s}$  son funciones no aleatorias que contienen los efectos del primero y segundo factor y la interacción entre los mismos respectivamente.
- $\gamma$  y  $Y_i^{r,s}$  son funciones no aleatoria y aleatoria respectivamente.
- $\epsilon_i^{r,s}$  corresponden a trayectorias aleatorias independientes y centradas. Además, para cada  $r, s$  fijas,  $\epsilon_i^{r,s}, i = 1, \dots, n_{r,s}$  son idénticamente distribuidas.

Generalmente, las hipótesis a contrastar son las siguientes:

- El primer factor no tiene influencia.  $H_0^A : f^1 = \dots = f^R = 0$ .
- El segundo factor no tiene influencia.  $H_0^B : g^1 = \dots = g^S = 0$ .
- No hay interacción entre los factores.  $H_0^I : h^{1,1} = \dots = h^{R,S} = 0$ .
- La covariable no tiene influencia.  $H_0^C : \gamma = 0$ .

Dado que el procedimiento bootstrap muchas veces requiere de cómputo intensivo, la metodología también permite aplicar como criterio de rechazo el procedimiento de Bonferroni, en la literatura se dice que este procedimiento es conservador, así que también propone una mejora a este método con una corrección de la tasa de descubrimiento falso. Estas dos alternativas se encuentran detalladas en [34] y se resumen como sigue a continuación:

Sea  $U_1, \dots, U_k$  direcciones aleatorias donde se proyectan los datos; sean  $p_1, \dots, p_k$  los  $p$ -valores de las pruebas de proyecciones aleatorias de los datos en las direcciones  $U_1, \dots, U_k$ , respectivamente. El procedimiento de Bonferroni lleva a rechazar la hipótesis  $H_0$  con un nivel  $\alpha$  si:

$$\min_j p_j \leq \frac{\alpha}{k}$$

Por otro lado, si se prueban  $k$  hipótesis diferentes, entonces la corrección con FDR es la proporción de hipótesis rechazadas erróneamente. En este caso en particular las hipótesis probadas son equivalentes y la FDR coincide con el nivel del procedimiento. Así, dados los  $p$ -valores ordenados  $p_{(1)}, \dots, p_{(k)}$ , bajo esta metodología, se rechaza la hipótesis nula para cada nivel si:

$$\alpha \geq \inf \left\{ \frac{k}{i} p_{(i)}, i = 1, \dots, k \right\}$$

La aplicación de esta metodología se ve facilitada por la función *fanova.RPm* de la librería *fda.usc* en el software estadístico R.

## 2.3. Casos de estudio FDA

Esta sección tiene como objetivo dar a conocer los casos prácticos de estudio, analizados en la presente memoria de tesis. Los casos se desarrollan en el ámbito de la Química Analítica y el confort térmico en edificaciones. Los conjuntos de datos se desarrollan en el ámbito FDA y fueron analizados en un estudio previo [3], donde se prueba una primera aproximación FDA de los estadísticos  $h$  y  $k$  de Mandel.

### 2.3.1. Datos Termogravimétricos TG

El primer conjunto de datos trata sobre técnicas analíticas de termo análisis en muestras de oxalato de calcio. La termogravimetría es una técnica de análisis térmico que proporciona información sobre la estabilidad térmica de un material mediante la medición de la pérdida de masa en función del tiempo o la temperatura [10]. Los datos corresponden a mediciones reales, obtenidos por nuestro actual director de trabajo de titulación PhD. Miguel Flores y colaboradores expertos en el área de Química Analítica.

El conjunto de datos fue analizado a profundidad en [35]. Según los autores; el material de estudio elegido corresponde al oxalato de calcio monohidrato con un 99 % de pureza ( $C_aC_2O_4 \cdot H_2O$ ), este material es comúnmente utilizado como referencia la termogravimetría porque se conoce de forma precisa sus características físicas, facilitando así su estudio. Adicionalmente presenta como característica tres pasos o reacciones de degradación, que ocurren en intervalos de temperatura y tiempo muy cortos. Una de las razones de emplear oxalato de calcio se debe al hecho de que las variaciones (o variabilidad esperada) de las curvas TG en función de la muestra se supone prácticamente despreciable con respecto a otras fuentes de variación [10].

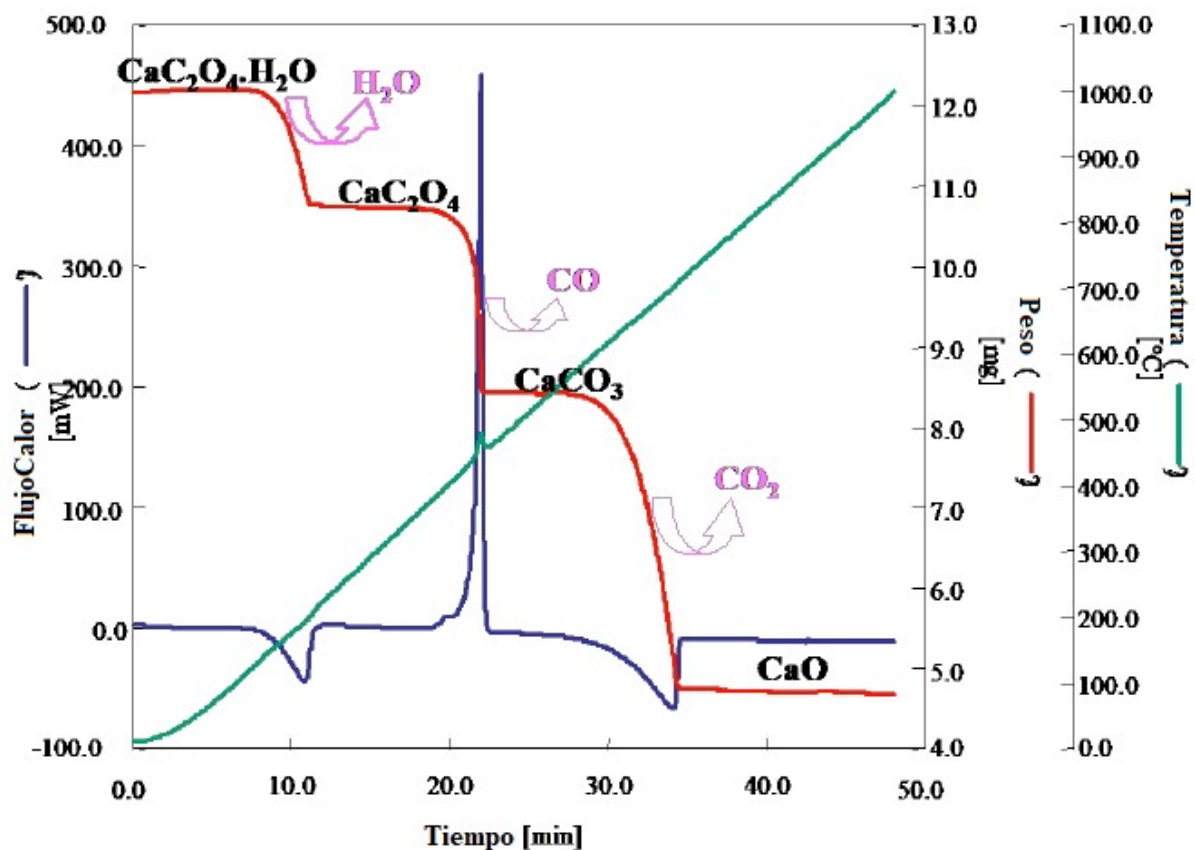
La Figura 2.1 muestra las tres reacciones químicas de las que se compone la degradación termooxidativa del oxalato de calcio con sus reactivos y productos, donde cada reacción de pérdida de masa mostrada en la curva TG se muestra como un escalón.

Los datos termogravimétricos TG fueron obtenidos mediante la emulación de un estudio interlaboratorio en donde intervienen 7 laboratorios que utilizan analizadores simultáneos con diferentes programas de calibración incluyendo una calibración antigua y sesgada (2 grados Celsius desplazados del punto de fusión del zinc) con los que se emulan los laboratorios posiblemente defectuosos. Los instrumentos de análisis simultáneos corresponden a un TA Instruments SDT 2960 y a un analizador simultáneo Rheometric STA 1500.

De cada muestra ensayada, se obtuvo como resultado una curva TG que proporcionan información acerca de los procesos de degradación sufridos por el material, en forma de escalones de pérdida de peso.

El ensayo experimental se compone de 4 laboratorios con resultados homogéneos desde el punto de vista de la repetibilidad y reproducibilidad (R & R) y 2 laboratorios que generan resultados inconsistentes; es decir, sus resultados fueron obtenidos por los instrumentos de medida defectuosos. La Tabla 2.1 muestra en resumen la descripción de cada uno de los laboratorios.

Los datos presentan 105 muestras de oxalato de calcio mediante análisis termogravimétrico (TG), donde se obtienen curvas TG que muestran la pérdida de masa de



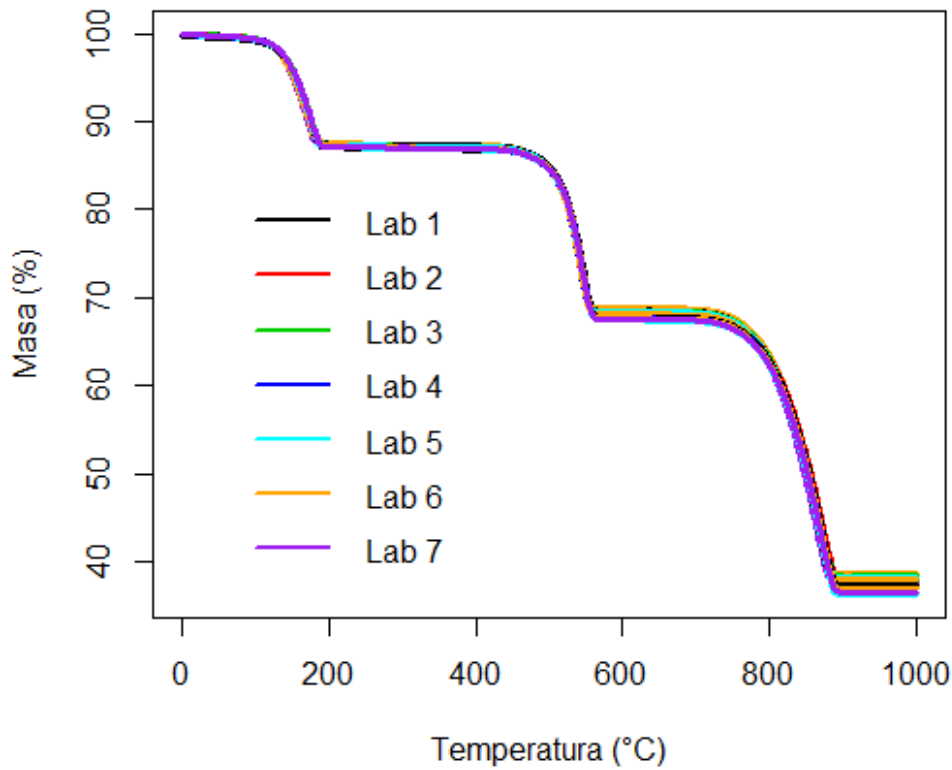
**Figura 2.1:** La curva TG correspondiente al oxalato de calcio y su proceso de degradación se muestra en color rojo. La curva de flujo de calor se muestra en azul y la relación de la temperatura con el tiempo en verde.

Fuente: Tomado de [10].

Laboratorio 1	Analizador STA con una calibración antigua
Laboratorio 2	Núcleo de laboratorios que proporcionan datos consistentes con instrumento SDT.
Laboratorio 3	
Laboratorio 4	
Laboratorio 5	
Laboratorio 6	Analizador SDT con una calibración antigua.
Laboratorio 7	Analizador SDT con una calibración sesgada 2°C.

**Tabla 2.1:** Descripción correspondiente a cada laboratorio.

oxalato dependiendo del tiempo en que las muestras son calentadas a una tasa de temperatura constante, estas condiciones de ensayo se eligieron atendiendo a la experiencia de los analistas y puede encontrarse los detalle en [35]. El conjunto de datos está compuesto por 15 curvas TG de 1000 observaciones o discretizaciones obtenidas por cada uno de los diferentes laboratorios. Las Curvas TG correspondientes a cada laboratorio puede verse en la Figura 2.2.



**Figura 2.2:** Curvas TG obtenidas a partir de oxalato de calcio correspondientes a cada laboratorio.

Este conjunto de datos es de fácil acceso ya que se encuentra adjunto en la librería *ILS* [28] implementada en el software estadístico R y son analizados nuevamente en el capítulo 4 con la aplicación de la nueva metodología propuesta.

### 2.3.2. Confort Térmico de superficies comerciales de empresa textil en Ciudad de Panamá

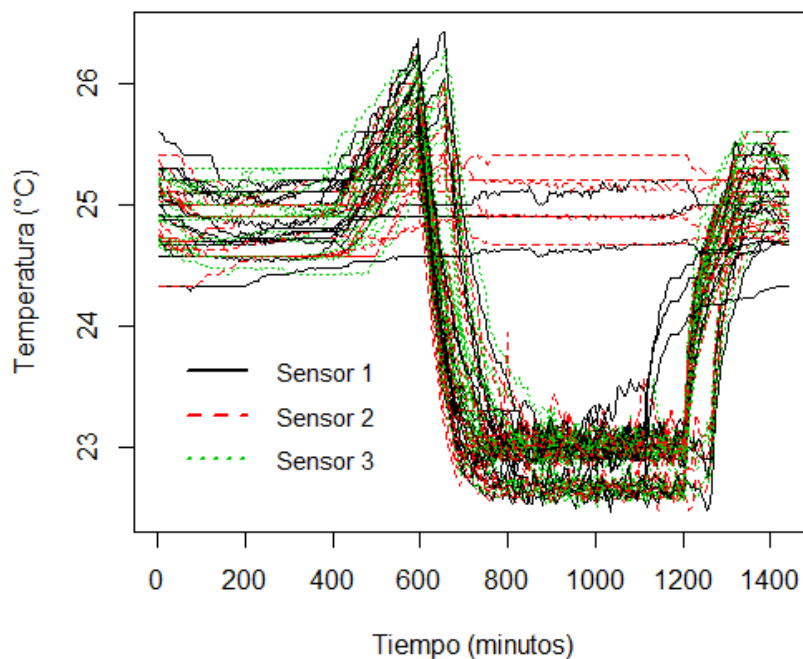
El segundo caso de estudio corresponde al análisis del confort térmico de una tienda ubicada en un centro comercial de la ciudad de Panamá. Para ello se trabaja con un conjunto de datos que presenta mediciones de temperatura obtenidas por tres sensores redundantes colocados en la misma zona comercial. Los datos fueron usados en [3] y estudiados a profundidad en [35]. Las mediciones fueron recopiladas por  $\Sigma qus$ , desarrollador de plataformas web y la empresa de consultoría estadística

Nexus para datos energéticos [36].

Se trata de un caso de estudio, que permite asegurar la correcta medición de la temperatura en la tienda de la empresa textil gallega, mediante la aplicación de un método automático que permita la detección de inconsistencias, es decir, se desea identificar aquellos sensores que podrían presentar anomalías y con ello la implementación de medidas correctivas de ser necesarias.

Las mediciones obtenidas por los 3 sensores corresponden al registro de la temperatura de la tienda (desde agosto 8, 2017 al 6 de octubre de 2017). El objetivo es detectar si los resultados obtenidos por los tres sensores son semejantes. En particular, el sensor 1 es utilizado como referencia para el correcto funcionamiento (por los responsables de mantenimiento), por tanto, se pretende evaluar si los sensores 2 y 3 proporcionan medidas consistentes.

La Figura 2.3 muestra los datos originales medidos en una frecuencia de 5 min durante el horario de atención de la tienda que es entre las 10 y 21 horas, durante ese tiempo la instalación de ventilación y aire acondicionado está en funcionamiento.



**Figura 2.3:** Curvas de temperatura obtenidas por 3 sensores diferentes en frecuencia de 5 min.



### 2.3.3. Aplicación de técnicas FDA a Datos Termogravimétricos.

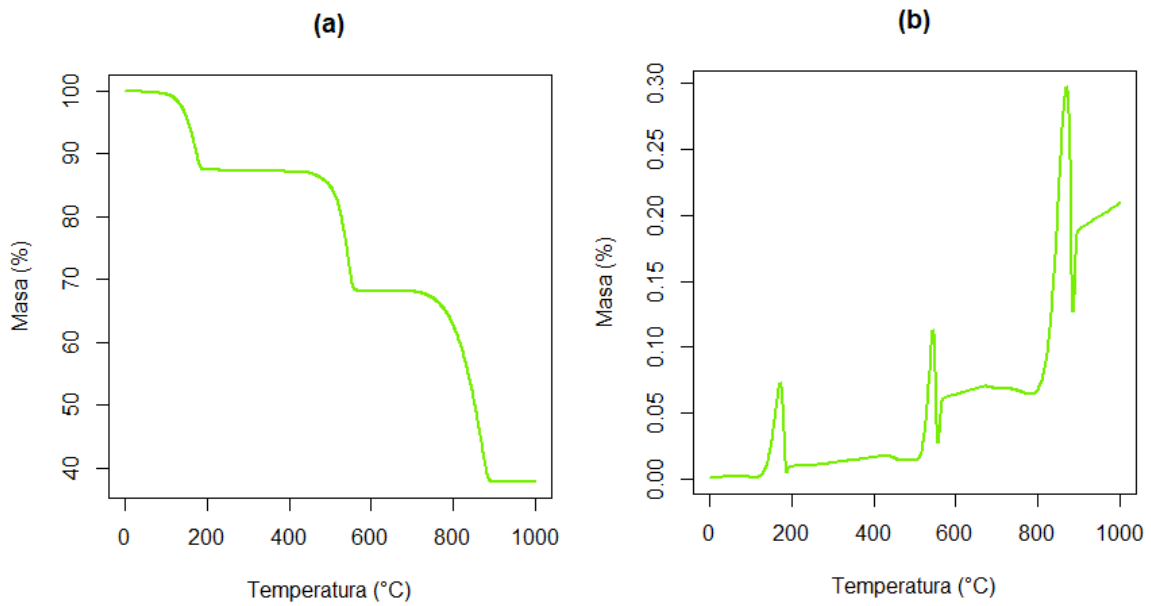
Con el fin de contar con un análisis de datos preliminar en el ámbito del Análisis de datos Funcional que permita utilizar toda la información de la curva para la detección de datos inconsistentes, en esta sección se muestra los resultados obtenidos al aplicar algunas técnicas funcionales detalladas en la Sección 2.2. Para ello se considerarán datos relacionados a la Química Analítica, que corresponden a curvas en datos termogravimétricos. A continuación, se muestra la aplicación de las siguientes técnicas y procedimientos en un ILS.

- Análisis exploratorio FDA.
- Detección de valores atípicos FDA basado en el cálculo de la profundidad de datos usando Proyecciones aleatorias.
- ANOVA Funcional.

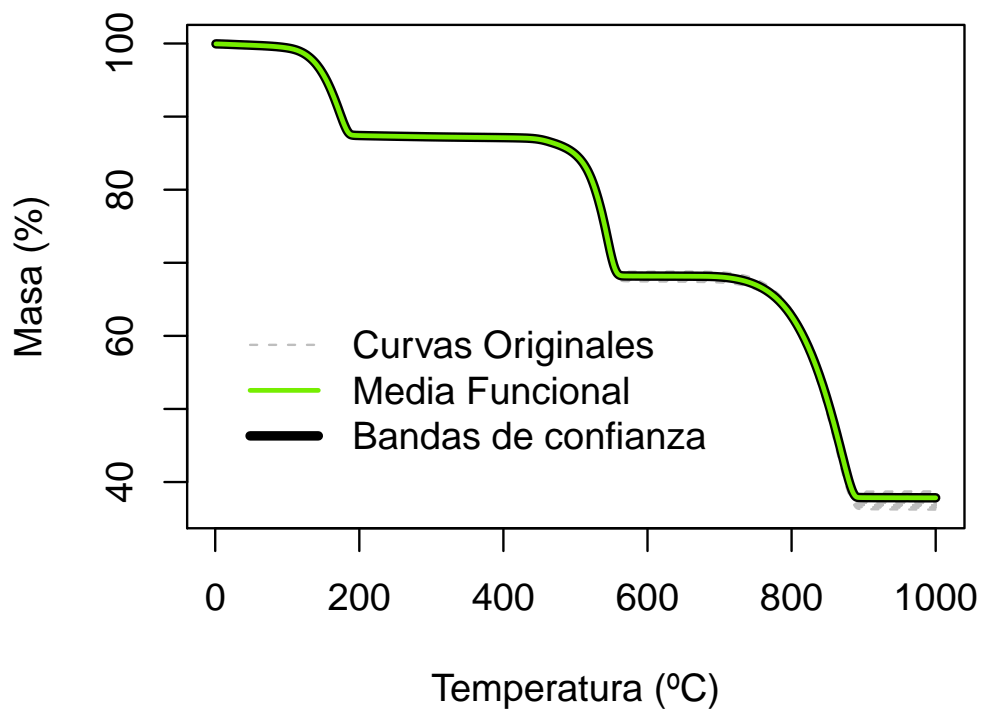
#### Análisis Exploratorio FDA

El análisis exploratorio permite obtener información de la variabilidad y la tendencia de los datos mediante el estudio de la media y la varianza funcionales globales.

La Figura 2.4 muestra la estimación de la media funcional global y la varianza funcional global de las curvas TG. En la media funcional global, se puede notar claramente 3 procesos de degradación bien definidos y separados en distintos intervalos de temperatura; por otro lado la varianza global, indica que el nivel de la varianza es muy alto en el último paso de la degradación. Adicionalmente en la Figura 2.5 se incluyen estimaciones de las bandas de confianza bootstrap para la media global, a un nivel de significación del 95 %; se puede notar que la amplitud de las bandas es muy estrecha, lo que puede indicar una baja variabilidad en la estimación de la media funcional global.

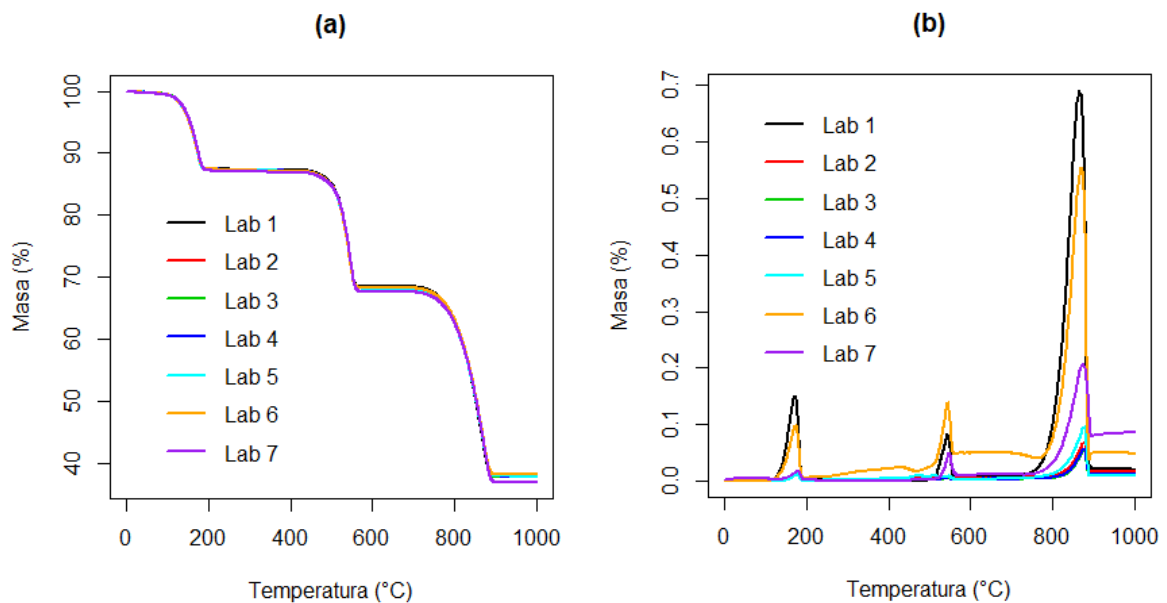


**Figura 2.4:** Estadísticos funcionales globales para curvas TG. (a) Media funcional global y (b) varianza funcional global



**Figura 2.5:** Se muestran las bandas de confianza bootstrap de curvas TG (línea continua negra), media funcional global (línea continua verde) y curvas TG originales (línea cortada gris)

Las medias y varianzas funcionales para cada laboratorio se muestran en la Figura 2.6, donde se evidencia un comportamiento similar en las medias funcionales de los laboratorios 2, 3 y 4, sin embargo para los laboratorios 1, 6 y 7, las diferencias entre las medias funcionales son notorias cuando la temperatura es alta, es decir, en el último proceso de degradación. Con respecto a la estimación de varianza funcional, la Figura 2.6 muestra una alta variabilidad conforme la temperatura aumenta. Las variaciones funcionales más altas se evidencian en los laboratorios 1 y 6, y con menor intensidad en el laboratorio 7, específicamente en el último proceso de degradación, donde las temperaturas son máximas.



**Figura 2.6:** Estadísticos funcionales de curvas TG para cada laboratorio. (a) Medias funcionales y (b) varianzas funcionales.

### Medidas de profundidad en base a proyecciones aleatorias

Continuando con el análisis, es importante también detectar datos atípicos sin considerar el laboratorio del que provengan, esto permite obtener resultados más objetivos sobre los laboratorios que pueden ser inconsistentes. Para ello se aplica el método de detección de curvas menos profundas con el uso de proyecciones aleatorias. Los parámetros con los que se determinaron las profundidades atípicas fueron los siguientes:  $B = 500$  remuestras (obtenidas mediante Bootstrap suavizado) con parámetro de suavizamiento  $\gamma = 0.05$ , nivel de significación  $\alpha = 0.1$  correspondiente al cuantil que permite identificar los resultados atípicos.

La Tabla 2.2 muestra en negrita los valores de profundidad más pequeños obtenidos mediante el método de proyecciones aleatorias. Estos valores corresponden a la profundidad de las curvas TG que fueron detectadas como atípicas. Como se puede ver, los valores de profundidad de las curvas corresponden a los laboratorios 1, 6 y 7. Por lo tanto, podemos suponer que estos laboratorios proporcionan curvas TG diferentes a las obtenidas por los 4 laboratorios restantes.

Estos resultados tienen coherencia de acuerdo con los sistemas de medición que fueron utilizados en cada uno de los laboratorios, de acuerdo a la descripción de los datos, precisamente los laboratorios 1, 6 y 7 fueron obtenidos con sistemas de calibración incorrectos. Esta primera aproximación mediante el análisis exploratorio de datos permite intuir que los laboratorios 1, 6 y 7 podrían ser atípicos y, por tanto, los laboratorios en conjunto pueden ser susceptibles de ser eliminados de un ILS.

Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7
0,1402	0,3055	0,4554	0,2154	0,3850	<b>0,0941</b>	0,1307
<b>0,0619</b>	0,3518	0,2124	0,3531	0,2878	<b>0,0634</b>	0,1916
0,1512	0,2802	0,3343	0,4396	0,2987	0,2406	0,1261
0,1141	0,3364	0,3941	0,3364	0,3053	0,1573	<b>0,0783</b>
<b>0,0916</b>	0,3680	0,3299	0,2924	0,2234	0,2299	0,1648
0,2318	0,2937	0,2890	0,3335	0,3941	0,1640	0,1960
0,1141	0,2124	0,2670	0,3196	0,1731	<b>0,0941</b>	0,1916
0,1783	0,2440	0,3478	0,4326	0,3406	<b>0,0510</b>	0,1156
0,2318	0,2722	0,2648	0,3531	0,2853	0,2124	0,1756
0,2318	0,1990	0,4038	0,4396	0,3518	0,1402	0,0996
<b>0,0895</b>	0,3170	0,4554	0,4010	0,2722	0,1528	<b>0,0636</b>
0,1512	0,4396	0,3053	0,2244	0,2472	0,1364	0,0996
0,2469	0,4162	0,3478	0,3170	0,3170	0,3324	<b>0,0352</b>
0,1575	0,3196	0,3358	0,3535	0,2451	0,1745	0,0975
0,2390	0,2853	0,3899	0,4162	0,4038	0,2335	<b>0,0493</b>

**Tabla 2.2:** Profundidades de curvas TG obtenidas mediante proyecciones aleatorias. Los valores en negrita corresponden a las profundidades más bajas.

### Pruebas ANOVA Funcionales

Finalmente, se aplica pruebas ANOVA Funcionales para determinar si existen diferencias entre las medias de los resultados correspondientes a los 7 laboratorios  $\mu_l(t)$ ,

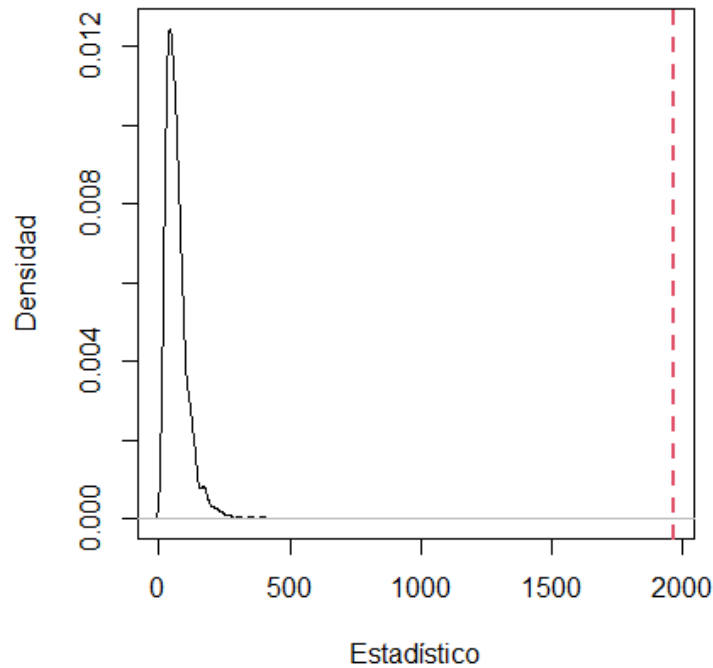
es decir las hipótesis pueden enunciarse de la siguiente manera:

$$H_0 : \mu_1(t) = \cdots = \mu_L(t)$$

$$H_1 : \exists k, m \quad \mu_k(t) \neq \mu_m(t)$$

Se consideran entonces los Método FANOVA un factor y ANOVA funcional vía Proyecciones aleatorias. Para estos casos, la variable porcentaje de pérdida de masa en una curva TG es la variable de respuesta, mientras que el laboratorio al que pertenecen los resultados se considera como una variable cualitativa o factor. Se utilizan las funciones `anova.onefactor` y `fanova.RPm` del paquete `fda.usc` para aplicar el contraste de igualdad de medias funcionales para cada caso respectivamente.

Los resultados obtenidos con el método ANOVA de un factor se presentan en la Figura 2.7 donde se muestra la distribución bootstrap del estadístico de contraste y la ubicación de su valor en la muestra (1965.399), identificado por una línea segmentada. El cálculo del p-valor se obtuvo a partir de 1000 remuestras bootstrap, y en este caso fue menor que 0.01, con lo cual se rechaza la hipótesis nula.



**Figura 2.7:** Distribución bootstrap del estadístico del método ANOVA funcional un factor

Por otro lado, los resultados obtenidos con el método ANOVA funcional vía proyecciones aleatorias se presentan en la Tabla 2.3, donde se muestran los resultados de los p-valores obtenidos bajo los criterios de Bon Ferroni (BF) y Tasa de descubrimiento Falso (FDR). A partir de 1000 remuestras bootstrap y de diferentes cantidades de proyecciones aleatorias  $RP = \{5, 30\}$ , los p-valores obtenidos son menores que el nivel de significación establecido  $\alpha = 0.01$ , por ende se rechaza la hipótesis nula de que no existen diferencias entre las medias de los laboratorios.

RP	p-valor Método BF	p-valor Método FDR
5	0,00	0,00
30	0,00	0,00

**Tabla 2.3:** p-valores obtenidos mediante el método ANOVA funcional Vía proyecciones aleatorias

Los resultados que arrojan las pruebas ANOVA funcionales son los esperados e indican que existen diferencias entre las medias obtenidas por cada laboratorio, lo que implica que existen laboratorios que presentan resultados significativamente diferentes a los de los demás. Según el análisis descriptivo y la prueba de detección de atípicos en base a profundidades, estos corresponden a los laboratorios 1, 6, y 7 que fueron considerados como inconsistentes. Por lo tanto, bajo este análisis, los tres laboratorios mencionados podrían ser considerados como atípicos y por ende ser eliminados del estudio.



### Comparación con Normas ASTM E-691

Es importante recalcar que si los datos son de carácter funcional, estos deben ser tratados como tales, por lo que se debe evitar utilizar enfoques escalares para el análisis de curvas. En este caso, utilizar un enfoque escalar puede llevar a la pérdida de información relevante para el estudio, generando errores de precisión y, de hecho, dependiendo de la característica univariante extraída de la curva, el resultado de la prueba puede variar.

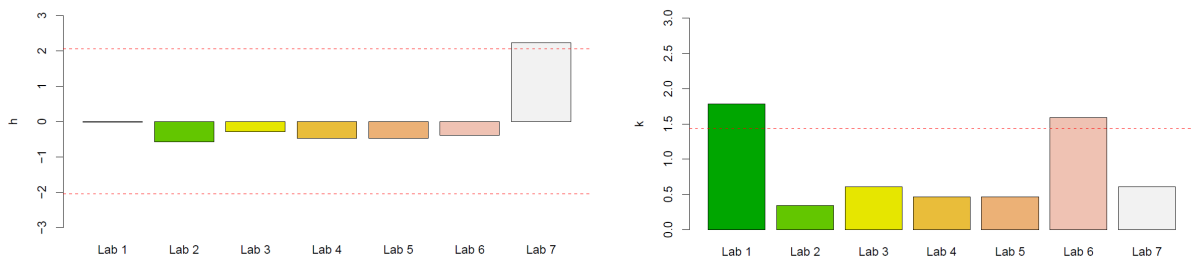
Considerando el mismo conjunto de datos termogravimétricos, en [10] se realizó una aplicación de la norma ASTM E691 y se tomó como dato univariante aquella característica que resume toda la información de la curva TG. La característica utilizada es la temperatura inicial de descomposición (IDT); definida como la temperatura a la que se pierde el 5 % de la masa total (primer paso de degradación) y está relacionada con la estabilidad térmica obtenida a partir de pruebas experimentales.

En dicho estudio, haciendo uso de los datos (IDT), se calculan los estadísticos  $h$  y  $k$  de Mandel univariantes y sus valores críticos  $h_c$  y  $k_c$  mediante el uso de la función *mandel.fqcs* de la librería *ILS*, en el software estadístico R [28] (Ver 2.4). Considerando los criterios para determinar si un laboratorio es detectado como atípico (también detallados en la sección 2.1), según las normas ASTM E-691, la Figura 2.8 muestra gráficamente como los laboratorios 1, 6 y 7 fueron detectados como atípicos según el contraste de los estadísticos estadístico  $h$  y  $k$ .

Número de laboratorios del ILS	7
Valor crítico estadístico $h$ : $c_h$	2.05, -2.05
Valor crítico estadístico $k$ : $c_k$	1.44
Número de laboratorios atípicos identificados	3
Laboratorios atípicos	Lab 1, Lab 6 y Lab 7

**Tabla 2.4:** Resultados obtenidos al aplicar la Norma ASTM E-691 en los datos correspondientes a la característica IDT en el primer paso de degradación.

Fuente: Adaptado de [10].



**Figura 2.8:** Gráfico correspondiente a los estadístico  $h$  y  $k$  obtenidos para cada laboratorio (panel izquierdo y derecho respectivamente) y sus valores críticos (líneas punteadas en color rojo)

Fuente: Tomado de [10].

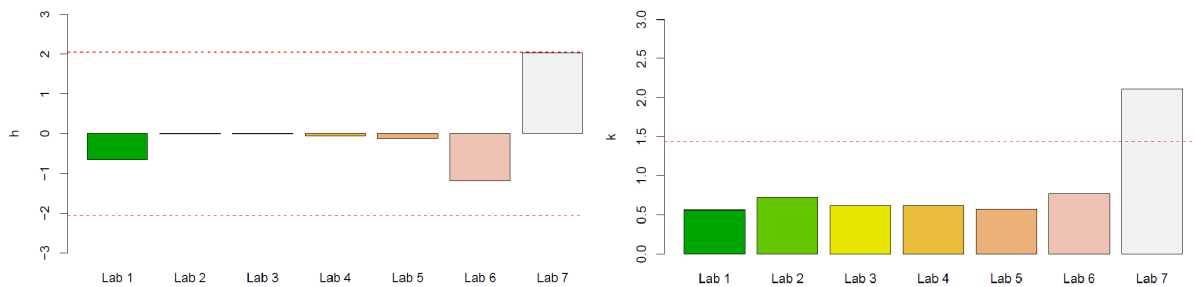
Adicionalmente para demostrar como los resultados pueden variar dependiendo de la característica univariante elegida, en el mismo estudio, se analizó una segunda característica correspondiente a la pérdida de masa en el segundo proceso de degradación en curvas TG, en este caso únicamente el laboratorio 7 fue detectado como atípico (Ver 2.5 y Figura 2.9). El proceso es iterativo, es decir, para cerciorarse que no hay más laboratorios atípicos se realizó una segunda iteración excluyendo el laboratorio 7 y finalmente la metodología no detecta ningún laboratorio adicional [3].

En este sentido, los laboratorios detectados como atípicos según las dos características univariantes analizadas difieren, además los resultados obtenidos en el apartado anterior con la aplicación de técnicas FDA indicaron que los laboratorios 1 6 y 7 son

	Iteración 1	Iteración 2
Número de laboratorios del ILS	7	6
$h_c$	2.05, -2.05	1.92, -1.92
$k_c$	1.44	1.43
Número de laboratorios atípicos identificados	1	0
Laboratorios atípicos	Lab 7	—

**Tabla 2.5:** Resultados obtenidos al aplicar la Norma ASTM E-691 en los datos correspondientes a la característica IDT en el segundo paso de degradación.

Fuente: Adaptado de [10].



**Figura 2.9:** Gráfico correspondiente a los estadístico  $h$  y  $k$  obtenidos para cada laboratorio (panel izquierdo y derecho respectivamente) y sus valores críticos (líneas punteadas en color rojo)

Fuente: Tomado de [10].

los posibles candidatos a ser atípicos; por lo que los resultados también difieren con respecto al enfoque escalar.

A manera de conclusión podemos decir que este estudio se citó con la finalidad de realizar una comparación entre los resultados, del análisis de curvas TG, proporcionados por la metodología clásica paramétrica establecida por la norma ASTM E691 [10] y los resultados obtenidos mediante la aplicación de técnicas FDA, detalladas en la Sección 2.3 del presente estudio. Tomando en cuenta ambos resultados se deja en evidencia como las técnicas funcionales utilizan toda la información presente en las curvas TG, para identificar laboratorios que proporcionan resultados inconsistentes, mientras que el estudio ILS univariante, identifica los laboratorios atípicos a partir de una pequeña parte de la información contenida en las curvas. Se puede ver que en los resultados anteriores (Tablas 2.4, 2.5), el número de valores atípicos identificados en un ILS tradicional en base a las normas ASTM E-691, depende de la característica extraída de las curvas **TG es este** caso IDT y pérdida de masa en el segundo proceso de degradación. Por lo tanto, se puede concluir que no todas las características resumen correctamente la información de toda una curva, lo que podría generar resultados poco fiables.



Como se mencionó al inicio de la presente sección, el uso de las técnicas FDA permite identificar aquellos laboratorios, posiblemente atípicos, que presentan resultados inconsistentes. Sin embargo, para contar con herramientas propias que provengan de procedimientos utilizados en un ILS, en el Capítulo 3 se desarrolla una nueva metodología para la detección de laboratorios atípicos en un ILS a partir de las extensiones funcionales de los estadísticos  $h$  y  $k$  de Mandel y las profundidades de curvas con base en proyecciones aleatorias.

## Capítulo 3

# Metodología de detección de atípicos basados en técnicas FDA para los estadísticos funcionales $H$ y $K$ de Mandel, utilizando proyecciones aleatorias

La nueva metodología FDA para la detección de laboratorios atípicos en ILS se apoya en la aplicación de técnicas FDA con base en proyecciones aleatorias como la reducción de dimensiones, la determinación de la profundidad de datos y el uso de las extensiones funcionales de los estadísticos  $h$  y  $k$  de Mandel que se los define formalmente en la sección 3.2.

En este capítulo se presenta la nueva propuesta metodológica sustentada en proyecciones aleatorias, denominada como metodología PA de aquí en adelante; además se aborda la metodología presentada por Flores [10], donde se ha realizado una primera aproximación en el contexto FDA para la de detección de laboratorios atípicos con base en el cálculo de la distancia de curvas haciendo uso de la norma  $L_2$ , a la cual desde este momento nos referimos como metodología  $L_2$ . Se precisa el alcance de ambas metodologías, con el objetivo de comparar y evaluar el desempeño de la nueva propuesta desarrollada ante previos enfoques funcionales.

Se busca entonces introducir un marco referencial para metodologías que permitan

determinar laboratorios atípicos. En la sección 3.1 se detallan los conceptos de precisión de un método de medición realizado en función de los estudios de repetibilidad y reproducibilidad, los mismos que se fundamentan en el uso de los estadísticos  $h$  y  $k$  de Mandel. Posteriormente en la sección 3.2 se define formalmente la extensión funcional de los estadísticos  $h$  y  $k$  de Mandel y se plantean las hipótesis de reproducibilidad y repetibilidad basándose en los estadísticos de contraste funcionales. Finalmente, en la sección 3.3 se detalla tanto la metodología  $L_2$ , analizada en estudios previos [3], así como la metodología PA propuesta en el presente trabajo de investigación.

### 3.1. Precisión, repetibilidad y reproducibilidad

Bajo los lineamientos de las normas ASTM E-691 e ISO 5725 numerosas organizaciones realizan estudios ILS para verificar el cumplimiento de ciertos requisitos de calidad de un material específico, por lo cual, es importante contar con procedimientos que permitan determinar si los resultados obtenidos no se ven afectados por los laboratorios que las realizan. Considerando que los resultados deben ser exactos y precisos es necesario contrastar los supuestos de reproducibilidad y repetibilidad.

Al evaluar los resultados de ensayos o pruebas de medición obtenidas por un grupo de laboratorios, en circunstancias idénticas, con técnicas similares e incluso materiales iguales, por lo general, los resultados arrojados no son equivalentes entre ellos y esto se atribuye a errores aleatorios que son inherentes en cada procedimiento. Sin embargo, existen factores que no pueden estar completamente en control y que contribuyen a la variabilidad de los resultados como los sistemas de medición, la calibración del equipo e incluso condiciones externas como temperatura, humedad, etc.

Con la finalidad de mitigar las variaciones que pueden llegar a producir estos factores y de admitir solamente los errores inherentes al proceso experimental, resulta indispensable estandarizar procedimientos para determinar la exactitud de las metodologías (diferencias entre resultados obtenidos durante la medición y valores esperados).

La precisión de un sistema de medición o de un procedimiento experimental en [37] se define como la variabilidad estadística existente en las mediciones realizadas de una misma magnitud en una misma muestra. Esta variabilidad está relacionada con

los conceptos de reproducibilidad y repetibilidad (R & R). En efecto, al llevar a cabo un ILS, se realizan estudios R & R de las medidas obtenidas para evaluar si la variabilidad depende del laboratorio que las realiza (alta variabilidad de reproducibilidad) o si la variación en las mediciones realizadas por cada laboratorio resulta significativa (alta variabilidad de repetibilidad) [2].

Tomando como referencia la norma ISO 5725-2, los estudios acerca de la precisión definida en términos de varianza o desviación típica de las mediciones tomadas de una **determinada** magnitud dentro de un **determinado método** de análisis son parte fundamental de los ILS [37], puesto que los estudios de repetibilidad y reproducibilidad resultan ser dos condiciones que garantizan la precisión de los resultados.

Según la ASTM international [14] la repetibilidad se define como la precisión bajo condiciones donde las medidas independientes son obtenidas con el mismo método, materiales, laboratorio, operador y usando instrumentos similares de medida en cortos intervalos de tiempo entre mediciones. Por otra parte, se define a la reproducibilidad como la precisión bajo condiciones donde las mediciones se obtienen con el mismo método aplicado a materiales iguales, en distintos laboratorios, con diferentes operadores e instrumentos de medida [38].

Dentro del marco de los ILS, al hablar de estudios de repetibilidad nos referimos a los estudios intralaboratorios aplicados a un procedimiento experimental bien definido que busca estimar errores en las mediciones de un laboratorio apoyándose en una muestra de reiteradas mediciones. Al tratarse del mismo laboratorio, se garantiza que se mantienen constantes los factores que pueden influir en la variabilidad de resultados.

Por otro lado, en el caso de estudios de reproducibilidad en un ILS, usando el mismo procedimiento experimental se busca estimar el error en las mediciones de un laboratorio **basados** en un conjunto de mediciones tomadas por diferentes laboratorios. En este caso, los factores (instrumentos de medición, condiciones ambientales, etc.) cambian de un laboratorio a otro y por lo general contribuyen de manera apreciable a la variabilidad de los resultados de las pruebas.

Finalmente, los estudios (R & R) permiten realizar el planteamiento de las hipótesis de repetibilidad y reproducibilidad en sección 3.2.2 para evaluar la consistencia de los datos y la precisión de los resultados obtenidos al implementar un estudio ILS. Las hipótesis de repetibilidad y reproducibilidad hacen uso de los estadísticos de contraste  $h$  y  $k$  de Mandel en el caso univariante, pero al plantear una nueva meto-

dología FDA se requiere presentar la extensión funcional de los estadísticos  $h$  y  $k$  de Mandel en la siguiente sección.

### 3.2. Estadísticos funcionales $H$ y $K$ de Mandel

En el control de la calidad de sistemas de medición son recientes las propuestas de metodologías que permiten manejar y analizar los datos cuando son curvas. En este caso es importante contar con herramientas propias de un ILS que engloben toda la información de la curva y que sean extensiones naturales de procedimientos comúnmente utilizados para la detección de laboratorios atípicos. Como se observó en el capítulo anterior, los resultados obtenidos al realizar un ILS con un enfoque escalar, mediante las normas ASTM, no representaron correctamente la información de las curvas; por lo cual se requiere introducir las aproximaciones funcionales de los estadísticos de Mandel con el propósito de que se considere toda la información de la curva y lograr que los resultados sean más precisos y menos variables.

En el Capítulo 2 se expuso que los estadísticos  $h$  y  $k$  de Mandel tienen como objetivo detectar laboratorios que generan resultados atípicos respecto a otros laboratorios que participan del estudio ILS basándose en el análisis de la variabilidad presente en proceso de mediciones univariantes. En el ámbito FDA, la extensión funcional de estos estadísticos  $h$  y  $k$  cumplen con el mismo propósito, es decir detección de laboratorios inconsistentes. A continuación, se definen las extensiones funcionales de los estadísticos  $h$  y  $k$  de Mandel.

#### Estadístico funcional $H$ de Mandel

En el caso univariante, el estadístico  $h$  de Mandel permite verificar si existen diferencias entre las medias de las observaciones obtenidas por cada laboratorio. La extensión funcional para el estadístico  $h$  de Mandel cumple con la misma función y se define en [10] de la siguiente manera:

Sea  $\{X_1^l(t), \dots, X_n^l(t)\}$  un conjunto de observaciones obtenidas por  $L$  laboratorios  $l = 1, \dots, L$  donde  $t$  son discretizaciones de tiempo en un intervalo  $[a, b]$ .

La extensión funcional del estadístico  $h$  de Mandel, denotado por  $H_l(t)$ , está dado

por:

$$H_l(t) = \frac{\bar{X}_l(t) - \bar{X}(t)}{\sqrt{\sum_{l=1}^L \frac{(\bar{x}_l(t) - m(t))^2}{L-1}}}; \quad l = 1, \dots, L$$



donde  $\bar{X}_l(t)$  es la media funcional de cada laboratorio calculada punto a punto y  $m(t)$  es la curva de medias global (media de medias).

Es decir, que  $m(t) = \sum_{l=1}^L \frac{\mu_l(t)}{L}$  siendo  $\mu_l(t)$  la media poblacional funcional calculada para cada laboratorio  $l$ .

### Estadístico funcional K de Mandel

El estadístico  $K$  de Mandel, es utilizado para contrastar la variabilidad interna de un laboratorio individual con el resto de los laboratorios. Bajo las mismas consideraciones, de un conjunto de  $L$  laboratorios que obtienen  $n$  curvas o mediciones en un intervalo de tiempo, en [10] se definió la extensión funcional del estadístico  $k$  de Mandel denotado por  $K_l(t)$  de la siguiente manera:

$$K_l(t) = \frac{S_l(t)}{\sqrt{\bar{S}^2(t)}}; l = 1, \dots, L$$

donde,  $\bar{S}^2(t) = \frac{1}{L} \sum_{l=1}^L S_l^2(t)$  y  $S_l(t)$  es la varianza funcional de cada laboratorio calculada punto a punto.

### 3.2.1. Planteamiento del contraste de hipótesis de repetibilidad y reproducibilidad para el caso funcional

Una vez definidos los estadísticos funcionales  $H$  y  $K$ , se realiza el planteamiento de los contrastes de hipótesis de repetibilidad y reproducibilidad haciendo uso de los estadísticos de contraste funcionales.

Iniciamos entonces planteando el supuesto de que todos los laboratorios participantes del ILS generan observaciones consistentes tanto en media como varianza por lo que se asevera que las mediciones provienen de la misma distribución. Para contrastar la hipótesis planteada, dentro del marco de los estudios  $R\&R$ , es necesario establecer un punto de partida que fundamente la detección de inconsistencias. Para ello se establece la hipótesis nula de reproducibilidad de igualdad de medias y

la hipótesis nula de repetibilidad de igualdad de varianzas, de los estadísticos de contraste funcionales.

Notar que la hipótesis nula de reproducibilidad está asociada al estadístico funcional  $H$ , mientras que la hipótesis de repetibilidad al estadístico  $K$ . Ahora, para el contraste de hipótesis se requiere aproximar la distribución empírica de los estadísticos de prueba que corresponden a la proyección unidimensional de los estadísticos funcionales  $H$  y  $K$ . Estos estadísticos de contraste son obtenidos mediante técnicas de proyecciones aleatorias y medidas de profundidad (detalladas en el Capítulo 2), herramientas base para el desarrollo de la nueva metodología propuesta en el presente estudio.

### **Hipótesis nula de reproducibilidad**

La hipótesis nula de reproducibilidad establece que no existe diferencia entre las medias poblacionales funcionales  $\mu_l(t)$  calculadas para cada laboratorio  $l$ .

$$H_0 : \mu_1(t) = \mu_2(t) = \dots = \mu_L(t)$$

Para contrastar la hipótesis de reproducibilidad, con base en la metodología PA, se requiere de un estadístico de prueba. Este estadístico corresponde a la proyección unidimensional del estadístico  $H$  y es denotado por  $d_{PA}^H$ .

El estadístico  $d_{PA}^H$ , se define a partir de la proyección unidimensional del estadístico funcional  $H$ ; los conceptos relacionados a proyecciones aleatorias fueron detallados en la sección 2.2.3. El estadístico  $d_{PA}^H$  se define de la siguiente manera:

Sea el estadístico funcional  $H_l(t)$  de Mandel, evaluado para cada uno de los  $l$  laboratorios presentes en el estudio ILS, y sea  $w(t)$  una dirección aleatoria independiente de  $H_l(t)$ . Entonces una técnica para la reducción de dimensionalidad para la curva  $H_l(t)$  denotada por  $PA_l^w$  se define como el producto interno entre el estadístico funcional  $H_l$  sobre la dirección aleatoria  $w$ :

$$PA_l^w = \langle w, H_l \rangle = \int_a^b w(t) H_l(t) dt ; l = 1, \dots, L$$

donde  $L$  es el número de laboratorios presentes en el ILS.

Posteriormente se requiere obtener la medida de profundidad univariante de cada

proyección unidimensional. En este caso se considera la profundidad de Tukey  $TD$ :

$$TD(PA_l^{w_i}) = \min \{F(PA_l^{w_i}), 1 - F(PA_l^{w_i})\}$$

donde la dirección  $w_i \in \{W_i\}_{i=1}^N$  y  $F$  corresponde a la función de distribución empírica, se define:

$$F(PA_l^{w_i}) = \frac{1}{L} \sum_{k=1}^L \mathbb{1}(PA_k^{w_i} \leq PA_l^{w_i}).$$



En este caso se trabaja con una colección de direcciones aleatorias  $\{W_i\}_{i=1}^N$  como sugieren estudios previos [31], [32] y [33], con el fin de evitar problemas de estabilidad del proceso y pérdida de información al reducir la dimensionalidad; por lo tanto, la medida de profundidad se calcula haciendo uso de todas las proyecciones consideradas. En consecuencia, el valor medio de las profundidades de los puntos proyectados determina una profundidad para el estadístico  $H$ . Finalmente el estadístico de prueba  $(d_{PA})_l^H$  corresponde a la medida de profundidad basada en proyecciones aleatorias para el estadístico funcional  $H$  y el  $l$ -ésimo laboratorio. El estadístico  $(d_{PA})_l^H$  se define como sigue:

$$(d_{PA})_l^H = PAD(H_l) = \frac{1}{N} \sum_{i=1}^N F(PA_l^{w_i}),$$

de forma que los valores más pequeños de  $(d_{PA})_l^H$  corresponden a laboratorios no consistentes.

### Hipótesis nula de repetibilidad

La hipótesis nula de repetibilidad establece que no hay diferencia entre las desviaciones de las medidas tomadas por cada laboratorio, es decir:

$$H_0 = \sigma_1^2(t) = \sigma_2^2(t) = \dots = \sigma_L^2(t)$$

Donde  $l = 1, 2, 3, \dots, L$  es el número de laboratorios, y  $\sigma_l^2(t)$  es la varianza funcional poblacional correspondiente a cada laboratorio  $l$ .

Ahora, para contrastar la hipótesis nula de repetibilidad, se define el estadístico de



prueba  $d_{PA}^K$  con las mismas consideraciones anteriores como:

$$(d_{PA})_I^K = PAD(K_I) = \frac{1}{N} \sum_{i=1}^N D(PA_I^{w_i}),$$

de forma que los valores más pequeños de  $(d_{PA})_I^K$  corresponden a laboratorios no consistentes.

En este caso los estadísticos de prueba  $d_{PA}^H$  y  $d_{PA}^K$  fueron obtenidos bajo los conceptos de profundidad de datos basados en proyecciones aleatorias. Con el fin de reducir la variabilidad, la profundidad de cada dato resulta ser el promedio de las profundidades obtenidas para un gran número de direcciones aleatorias. Los estadísticos de prueba representan entonces una profundidad, por lo que valores de profundidad muy pequeños pueden implicar que el laboratorio sea considerado como atípico si sobrepasan el valor crítico obtenido a partir de un nivel de significación  $\alpha$ . Es preciso enfatizar que las curvas menos profundas tendrán una profundidad cercana a 0, mientras que los datos más profundos tendrán una profundidad cercana a 0.5 ya que el valor más profundo corresponde al valor medio de las curvas.

Por otro lado, la metodología  $L_2$  propuesta por Flores *et al.* en [10]; también considera el contraste de las hipótesis de repetibilidad y reproducibilidad. La diferencia con la metodología  $PA$  radica principalmente en la manera en cómo se definen los estadísticos de contraste. Así en [3], se definen a los estadísticos de contraste con base en la norma  $L_2$ , de la siguiente manera:

$$(d_{L_2})_I^H = \|H_I(t)\| = \left( \int_a^b H_I(t)^2 dt \right)^{\frac{1}{2}},$$

$$(d_{L_2})_I^K = \|K_I(t)\| = \left( \int_a^b K_I(t)^2 dt \right)^{\frac{1}{2}}$$

El estadístico  $(d_{L_2})_I^H$  permite contrastar la hipótesis de reproducibilidad y el estadístico de prueba  $(d_{L_2})_I^K$  permite contrastar la hipótesis de repetibilidad.

En este caso la norma  $L_2$  representa la distancia de las curvas de los estadísticos  $H$  y  $K$  con respecto al origen. Por lo que valores de distancias muy altos pueden implicar que el laboratorio sea considerado como atípico si sobrepasan el valor crítico obtenido a partir de un nivel de significación  $\alpha$ .

### 3.3. Metodología propuesta para la detección de laboratorios atípicos con base en proyecciones aleatorias (PA)

En secciones previas, se ha presentado la extensión funcional de los estadísticos  $H$  y  $K$  de Mandel además de plantear las hipótesis nulas de reproducibilidad y repetibilidad, y definir los estadísticos de prueba para cada hipótesis.

Como ya se mencionó, existe una propuesta metodológica funcional para la detección de laboratorios atípicos en ILS basándose en la norma  $L_2$  [10]. Esta metodología sirve como referente para el desarrollo de la nueva propuesta funcional basada en proyecciones aleatorias, Metodología PA. En este sentido se formula un algoritmo en función métodos de Bootstrap suavizado y técnicas de reducción de dimensionalidad que, en conjunto al uso de los estadísticos  $H$  y  $K$  funcionales, permita contrastar las hipótesis de repetibilidad y reproducibilidad.


A continuación, se detallan las Metodologías  $L_2$  y la nueva propuesta metodológica PA.

#### 3.3.1. Metodología funcional para la detección de laboratorios atípicos con base en la norma $L_2$

A manera de resumen, la metodología con base en la norma  $L_2$ , propuesta por Flores *et al.*, aplica un algoritmo basado en técnicas de remuestreo bootstrap, para estimar la distribución de los estadísticos de contraste  $(d_{L_2})_I^H$  y  $(d_{L_2})_I^K$  y con ello sus valores críticos a un nivel de significancia  $\alpha$ ; los mismos que en conjunto permitirán contrastar las hipótesis de repetibilidad y reproducibilidad.

El algoritmo para la metodología  $L_2$  consta de los siguientes pasos:

1. Agrupar todas las mediciones en un solo conjunto, y eliminar las curvas atípicas para generar un conjunto de datos consistente.
2. Se obtienen  $B$  réplicas bootstrap de tamaño  $n \times L$  del conjunto de datos. Las  $B$  muestras bootstrap se generan a partir de la hipótesis nula y posteriormente se asignan de manera aleatoria a cada laboratorio.

3. Para cada muestra bootstrap se calculan los estadísticos funcionales bootstrap notados por  $H_l^*(t)$  y  $K_l^*(t)$  y sus **corrientes** estadísticos de prueba  $(d_{L_2}^*)_l^H$  y  $(d_{L_2}^*)_l^K$ . 
4. Se aproximan los valores críticos  $c_{L_2}^H$  y  $c_{L_2}^K$  de los estadísticos de contraste  $(d_{L_2}^*)_l^H$  y  $(d_{L_2}^*)_l^K$  por el percentil empírico  $100(1 - \alpha)/L$  de la distribución de las  $L \times B$  réplicas bootstrap correspondientes. En los ILS, los valores críticos para los contrastes de detección de atípicos se calculan usualmente para un nivel de significación de  $\alpha = 0.01$ .
5. Posteriormente se calculan los estadísticos funcionales  $H_l(t)$  y  $K_l(t)$  y sus correspondientes estadísticos de prueba  $(d_{L_2})_l^H$  y  $(d_{L_2})_l^K$ , para cada laboratorio  $l = 1 \dots L$  de la muestra original.
6. **Criterio de rechazo:** Se contrastan las hipótesis nulas de reproducibilidad y repetibilidad siendo rechazadas si  $(d_{L_2})_l^H = \|H_l(t)\| > c_{L_2}^H$ , y  $(d_{L_2})_l^K = \|K_l(t)\| > c_{L_2}^K$ , respectivamente. El procedimiento se repite de forma iterativa hasta que no se detectan laboratorios atípicos.

### 3.3.2. Metodología funcional para la detección de laboratorios atípicos con base en proyecciones aleatorias (PA)

La presente investigación pretende aportar una nueva propuesta funcional para contrastar la hipótesis nula de reproducibilidad de igualdad de medias y la hipótesis nula de repetibilidad de igualdad de varianzas. Por lo tanto, **basados** en las ideas planteadas en la metodología  $L_2$ , se desarrolla una metodología que utilice los estadísticos de contraste  $d_{PA}^H$  y  $d_{PA}^K$  y que emplee proyecciones aleatorias para convertir estos estadísticos en escalares y realizar un contraste univariante. Al igual que la metodología  $L_2$ , la metodología  $PA$  aplica un algoritmo basado en técnicas de remuestreo bootstrap, para estimar la distribución de los estadísticos de contraste  $(d_{PA})_l^H$  y  $(d_{PA})_l^K$  y con ello sus valores críticos notados por  $c_{PA}^H$  y  $c_{PA}^K$  a un nivel **se** significancia  $\alpha$ .

El algoritmo para la metodología  $PA$  consta de los siguientes pasos:

1. Agrupar las mediciones en un conjunto de datos y eliminar las curvas atípicas mediante **en base al** método de recorte de las profundidades más bajas haciendo uso de proyecciones aleatorias (detallado en el capítulo anterior sección

2.2.3). Se establece un porcentaje de recorte del 0.01 para aquellas curvas con menor profundidad. Bajo la hipótesis nula de que todos los laboratorios tienen resultados consistentes, se agrupan todas las curvas que no fueron eliminadas en un solo conjunto de datos.

2. Se utiliza el bootstrap suavizado bajo las especificaciones del Capítulo 2 para obtener muestras bootstrap de tamaño  $n \times L$  a partir del conjunto de datos de donde fueron descartados previamente las curvas atípicas. Las réplicas bootstrap se generan a partir de la hipótesis nula (todos los laboratorios tienen resultados consistentes) y posteriormente se asignan aleatoriamente a los laboratorios.
3. Para cada muestra bootstrap, se calculan los estadísticos funcionales  $H_l^*(t)$  y  $K_l^*(t)$  y delimitando un número de direcciones aleatorias, se genera la colección de proyecciones aleatorias para el cálculo de los estadísticos de prueba,  $(d_{PA}^*)_l^H$  y  $(d_{PA}^*)_l^K$ , para cada laboratorio  $l = 1, \dots, L$ .
4. Se establece un nivel de significación de  $\alpha = 0.01$ , valor usual en los ILS, y se procede a calcular el percentil empírico  $100(\alpha)/L$  de la distribución bootstrap, aproximando así los valores críticos  $c_{PA}^H$  y  $c_{PA}^K$  de los estadísticos de contraste  $(d_{PA}^*)_l^H$  y  $(d_{PA}^*)_l^K$ .
5. Posteriormente se calculan los estadísticos funcionales  $H_l(t)$  y  $K_l(t)$  y sus correspondientes estadísticos de prueba  $(d_{PA})_l^H$  y  $(d_{PA})_l^K$ , para cada laboratorio  $l = 1 \dots L$  de la muestra original.
6. **Criterio de rechazo:** Se contrastan las hipótesis nulas de reproducibilidad y repetibilidad siendo rechazadas si  $(d_{PA})_l^H < c_{PA}^H$ , y  $(d_{PA})_l^K < c_{PA}^K$ , respectivamente. El procedimiento se repite de forma iterativa hasta que no se detectan laboratorios atípicos.

Por último, una vez calculados los estadísticos  $(d_{PA})_l^H$  y  $(d_{PA})_l^K$  basados en la profundidad de proyecciones aleatorias y los valores críticos  $c_{PA}^H$  y  $c_{PA}^K$ , se contrastan las hipótesis nulas de reproducibilidad y repetibilidad, siendo rechazadas si  $(d_{PA})_l^H = \frac{1}{N} \sum_{i=1}^N D(PA_l^{w_i}) < c_{PA}^H$ , y  $(d_{PA})_l^K < c_{PA}^K$ , respectivamente. El procedimiento se repite de forma iterativa hasta que no se detectan laboratorios atípicos.

Los resultados de las metodologías para detección de atípicos basados en las hipótesis de contraste, usualmente pueden ser presentados de manera gráfica, intuitiva

y de fácil entendimiento. Lo cual, hacen de los estadísticos de Mandel una de las herramientas más utilizadas para evaluar la calidad de los sistemas de medición de los laboratorios en un ILS. La aplicación de las metodologías en los casos de estudio real se presenta en el siguiente capítulo.

## Capítulo 4

# Validación y aplicación de la nueva metodología

En el presente capítulo se realiza la validación de la metodología propuesta en un estudio de simulación de curvas termogravimétricas, considerando para ello diferentes escenarios simulados para los estadísticos  $H$  y  $K$  funcionales. El desarrollo de este capítulo incluye una comparación entre las metodologías en base a los estadísticos  $H$  y  $K$  para detección de laboratorios atípicos en ILS, considerando tanto la norma  $L_2$  (sección 3.3.1) [3] y la actual propuesta en base a proyecciones aleatorias (sección 3.3.2).

En la actualidad estos son los únicos enfoques presentados en el ámbito funcional para detección de atípicos en ILS, la comparación tiene como finalidad identificar los casos o escenarios en los que cada metodología presenta los mejores resultados.

Posterior a la validación se realiza la aplicación metodológica en los conjuntos de datos reales detallados en el capítulo 2, sección 2.3.

### 4.1. Validación del nuevo enfoque FDA

La validación mediante un estudio de simulación tiene como finalidad evaluar el desempeño de la nueva metodología FDA ante la metodología propuesta por Flores en [3]. Se consideran entonces dos escenarios de simulación que consisten en la generación de datos consistentes e inconsistentes respectivamente para evaluar la

potencia de la prueba de los estadísticos de contraste para las hipótesis de repetibilidad y reproducibilidad. Cada escenario está compuesto por  $L$  laboratorios y cada laboratorio genera  $n$  muestras.

El primero de ellos pretende estudiar el comportamiento de las metodologías cuando los datos son generados bajo la hipótesis nula, es decir, cuando los datos son consistentes, esto implica calcular la proporción de rechazo para verificar, si estableciendo un nivel de significación, este rechaza más veces que el nivel fijado.

Por otro lado, el segundo escenario pretende analizar el comportamiento de las metodologías ante el incumplimiento de las hipótesis de repetibilidad y reproducibilidad. Para ello, los datos simulados de curvas TG son generados con inconsistencias a partir de ligeras desviaciones su media y varianza en relación a sus valores teóricos bajo la hipótesis nula. En este caso nuevamente se calcula la proporción de rechazo y con ello se analiza la capacidad de las metodologías para identificar estas inconsistencias.

#### 4.1.1. Descripción del estudio de simulación

El estudio de simulación consiste en la generación de curvas TG. El comportamiento de este tipo de curvas ha sido abordado a profundidad en estudios previos como [39, 40], donde se indica que dicho comportamiento puede ser simulado considerando el modelo logístico generalizado, con esto, las curvas TG obtenidas por cada laboratorio, son simuladas mediante el siguiente proceso gaussiano:

$$Y(t) = \mu(t) + \sigma(t)\epsilon(t); \quad t \in [0, 1],$$

donde:

- $t \in [0, 1]$  es el intervalo simulado de tiempo/temperatura medido en unidades arbitrarias (u.a.)

La media corresponde al proceso logístico generalizado,

$$\mu(t) = \frac{c}{(1 + \exp(b(t - m)))^{\frac{1}{\tau}}}$$

La varianza determinista es:

$$\sigma(t)^2 = c_0 \left( 5 + v \left( 1 - \left( \frac{t}{0.5} - 1 \right)^2 \right)^3 \right).$$

El estudio de simulación descrito también fue abordado en [3] y con la finalidad de generar curvas similares al caso de estudio real, se realiza una réplica considerando parámetros específicos, que son determinados considerando el significado del modelo logístico generalizado. La representación de cada parámetro se detalla a continuación.

- $\epsilon$ : proceso estacionario de segundo orden definido con media 0 y cuyo elemento de la matriz de covarianza es  $\sigma_{ij} = \exp(-|t_i - t_j| / 0.3)$ .
- $m$ : representa el tiempo o la temperatura correspondiente a la máxima pendiente de la curva o punto de inflexión
- $c$ : valor inicial de la masa de la muestra.
- $b$ : representa la velocidad de cambio (en este caso velocidad de degradación térmica).
- $\tau$ : representa el grado de asimetría de las curvas.

En particular, para establecer la parte fija de la media teórica de la hipótesis nula; se considera un proceso de degradación simétrico ( $\tau = 1$ ) definido por una masa inicial de ( $c = 1$   $c_0 = 10^{-6}$ ) y la velocidad de degradación  $b = 10$ . Finalmente, la media para la hipótesis nula es:

$$\mu_0(t) = \frac{1}{1 + \exp(10(t - m))}.$$

#### 4.1.2. Estudio de simulación bajo la hipótesis nula de repetibilidad y reproducibilidad

En los escenarios de simulación, las hipótesis nulas de reproducibilidad y repetibilidad se formulan en términos de  $m_0$  y  $v_0$  respectivamente.

Para este caso en particular, los datos serán generados consistentemente para cada uno de los laboratorios, cuando la hipótesis nula de repetibilidad sea  $H_0 : v_0 = 5$  ;



es decir, cuando:

$$\sigma(t)^2 = c_0 \left( 5 + 5 \left( 1 - \left( \frac{t}{0,5} - 1 \right)^2 \right)^3 \right)$$

Por otro lado, considerando la hipótesis nula de reproducibilidad los resultados serán consistentes si  $H_0 : m_0 = 0.5$ ; es decir, cuando:

$$\mu(t) = \frac{1}{1 + \exp(10(t - m_0))}.$$

Bajo las consideraciones establecidas para las hipótesis nulas de repetibilidad y reproducibilidad, se simulan curvas TG correspondientes a  $L = 7$  laboratorios consistentes cada uno con  $n = 10, 20, 30$  réplicas, con el fin de determinar el porcentaje de rechazo bajo las hipótesis nulas, obtenidos por los dos métodos. Para ello se realizaron un total de 1000 simulaciones Monte Carlo. Cada simulación se compuso de 500 muestras bootstrap para estimar la distribución de los estadísticos  $d^H$  y  $d^K$  correspondientes a las dos metodologías, luego se obtiene el valor crítico de la prueba,  $c^H$  y  $c^K$ , a un nivel de significación  $\alpha = 0.01$ , que es el valor recomendado para un ILS.

La hipótesis nula de reproducibilidad debe ser probada mediante el contraste de los estadísticos  $d^H$  a partir de la estimación del estadístico funcional  $H(t)$ . De forma similar la hipótesis nula de repetibilidad debe ser probada considerando los estadísticos  $d^K$  y  $K(t)$ . Adicionalmente, teniendo en cuenta que dentro del algoritmo de cada metodología, previo al remuestreo, es necesario realizar la eliminación de datos atípicos mediante el método de medidas de profundidad. En ambos casos se utilizó el mismo tipo de profundidad.

La Tabla 4.1 muestra el porcentaje de rechazo bajo  $H_0$ , obtenido por cada uno de los dos métodos  $PA$  y  $L_2$ ; cuando se generan 7 laboratorios consistentes con diferentes tamaños de muestras ( $n = \{10, 20, 30\}$ ). A un nivel de significación  $\alpha$  del 1 % utilizando el estadístico  $d_{PA}^H$  y  $d_{L_2}^H$  para contrastar la hipótesis de reproducibilidad. Además, para la metodología  $PA$  se consideran varias proyecciones aleatorias ( $RP = \{30, 50\}$ ).

Se observa que la metodología  $PA$  no rechaza más veces que el nivel fijado  $\alpha$  y conforme aumenta el tamaño de la muestra, la potencia de la prueba converge al nivel  $\alpha$  establecido. Además, cuando se considera un numero de direcciones aleatorias más alto ( $RP = 50$ ) los valores de potencia incrementan sin sobrepasar el nivel fijado  $\alpha$ .

Por otro lado, la metodología  $L_2$  presenta valores de potencia ligeramente más altos e incluso aumentan conforme el tamaño de la muestra aumenta, sin lograr estabilizarse en el nivel  $\alpha = 0.01$ , la **potencia** de la prueba llega a superar ligeramente el nivel de significación  $\alpha$  establecido en todos los casos.

$L$	$n$	$m_0$	Metodología PA		Metodología $L_2$
			$pot_{PA} (RP=30)$	$pot_{PA} (RP=50)$	$pot_{L_2}$
7	10	0.5	0.008429	0.008429	0.010143
7	20	0.5	0.008714	0.011000	0.012714
7	30	0.5	0.009429	0.009714	0.012857

**Tabla 4.1:** Valores de potencias obtenidos bajo la hipótesis nula de reproducibilidad de los estadísticos  $d_{PA}^H$  y  $d_{L_2}^H$ .

Ahora, la Tabla 4.2 muestra el porcentaje de rechazo bajo  $H_0$  para  $\alpha = 0.01$ , obtenido por cada uno de los dos métodos  $PA$  y  $L_2$ ; utilizando el estadístico  $d_{PA}^K$  y  $d_{L_2}^K$  para contrastar la hipótesis de repetibilidad. Se puede ver que los valores de **potencia** para los dos métodos superan ligeramente, en casi todos los casos, al nivel de significación establecido. Sin embargo para el caso  $PA$ , cuando se aumenta el número de direcciones aleatorias ( $RP = 50$ ), las **potencias** disminuyen más rápidamente convergiendo hacia el nivel  $\alpha$ ; con esto, la metodología  $PA$  no llega a superar el nivel  $\alpha$ . No obstante, la metodología  $L_2$  supera ligeramente el nivel de significación establecido en todos los casos y a pesar de que la **potencia** disminuye conforme aumenta el tamaño de muestra cuando  $n = 30$  aún supera el nivel  $\alpha = 0.01$ .

$L$	$n$	$v_0$	Metodología PA		Metodología $L_2$
			$pot_{PA} (RP=30)$	$pot_{PA} (RP=50)$	$pot_{L_2}$
7	10	5	0.010429	0.013143	0.015714
7	20	5	0.012571	0.012714	0.012714
7	30	5	0.009429	0.009000	0.013571

**Tabla 4.2:** Valores de potencias obtenidos bajo la hipótesis nula de repetibilidad de los estadísticos  $d_{PA}^K$  y  $d_{L_2}^K$ .

Con los resultados descritos para el estudio de simulación bajo la hipótesis nula (Tabla 4.1 y Tabla 4.2); se puede concluir que ambas metodologías estiman cerca-namente el nivel de significación  $\alpha$ . Siendo ligeramente más precisa la metodología  $PA$ , ya que los valores de **potencia** para los estadísticos de contraste  $d^H$  y  $d^K$  logran estabilizarse sin superar el nivel  $\alpha = 0.01$ . Por otro lado, los valores de **potencia** de los estadísticos  $d^H$  y  $d^K$  para la metodología  $L_2$  superaron el nivel establecido en todos los **casos, en otras palabras;** la metodología  $L_2$  rechazó ligeramente más de lo

que debía en relación a la metodología *PA* y por ende la metodología *PA* permite estimar de forma más precisa al nivel de significación  $\alpha$ .

### 4.1.3. Estudio de simulación bajo las hipótesis alternativas de repetibilidad y reproducibilidad

El estudio de simulación permite estudiar el comportamiento de la potencia de la prueba bajo el incumplimiento de las hipótesis de repetibilidad y reproducibilidad.

#### Hipótesis alternativa de reproducibilidad

En este escenario de simulación, la hipótesis nula de reproducibilidad se formula en términos de  $m_0, \delta_h$ .

Los datos serán generados bajo la hipótesis alternativa de reproducibilidad cuando

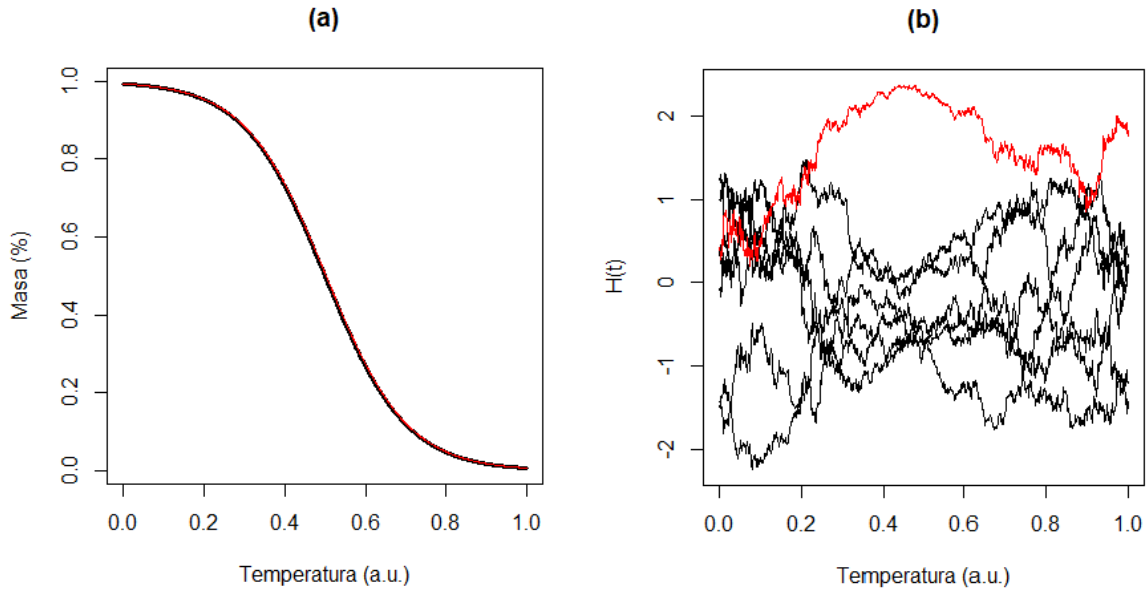
$$H_1 : m_1 = m_0 (1 + \delta_H)$$

con  $-0.005 \leq \delta_H \leq 0.005$ .

Las diferencias establecidas entre la media teórica y la media alternativa son ligeras. La Figura 4.1.a muestra la gráfica para la media teórica cuando  $m_0 = 0.5$  y  $m_1 = m_0(1 + 0.005)$  con  $\delta_H = 0.005$ .

En este sentido, se simulan curvas TG correspondientes a  $L = 7$  laboratorios, cada uno con  $n = \{10, 20\}$  réplicas. Del total de laboratorios, 6 son simulados bajo  $H_0$ ; es decir, son consistentes, mientras que los resultados del séptimo laboratorio, no lo son ya que se obtienen bajo  $H_1$ .

Las variaciones inducidas en la hipótesis alternativa permiten evaluar la capacidad de la nueva metodología para detectar correctamente los laboratorios inconsistentes mediante el estudio de la proporción de rechazo o la potencia de la prueba. Para ello nuevamente, la hipótesis nula de reproducibilidad debe ser probada mediante el contraste del estadístico  $d^H$  para la metodología *PA* y la metodología  $L_2$  a partir de la estimación del estadístico funcional  $H(t)$ . El estadístico funcional  $H(t)$  para cada laboratorio simulado se puede ver en la Figura 4.1.b.



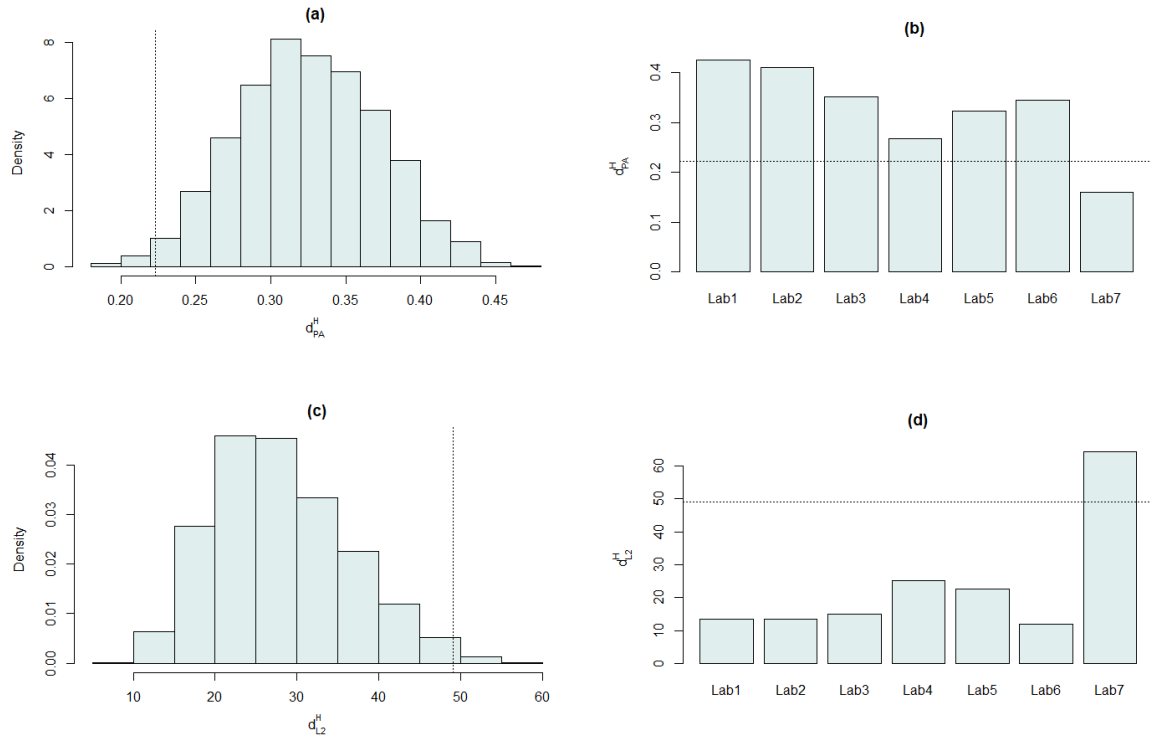
**Figura 4.1:** (a) Medias teóricas para curvas TG simuladas bajo la hipótesis nula  $H_0 : m_0 = 0.5$  (línea negra) y la alternativa  $H_1 : m_1 = m_0(1 + \delta_H)$  ( $\delta_H = 0.005$  línea roja). (b) Estadístico funcional  $H(t)$  para cada uno de los 6 laboratorios simulados bajo la nula (línea negra) y un laboratorio bajo la alternativa (línea roja).

Se realizó entonces un total de 1000 simulaciones, cada una compuesta de 500 muestras bootstrap para estimar la distribución de los estadísticos  $d^H$  de ambas metodologías (Figura 4.2.a, Figura 4.2.c), luego se obtiene el valor crítico de la prueba,  $c^H$ , representado con una línea vertical punteada y correspondiente a  $\alpha = 0.01$ , que es el valor recomendado para un ILS.

Los resultados para el estudio de simulación bajo la hipótesis de reproducibilidad para el método vía proyecciones aleatorias  $PA$  y norma  $L_2$  se detallan en la Figura 4.2.

Los estadísticos  $h$  y  $k$  de Mandel pueden ser representados gráficamente y permiten identificar de forma intuitiva los laboratorios inconsistentes. La Figura 4.1.b, muestra las realizaciones de la muestra de los estadísticos  $H_l(t)$  con  $l = 1, 2, \dots, 7$ . El estadístico funcional  $H$  proporciona información sobre el intervalo tiempo/temperatura donde los resultados del laboratorio 7 son diferentes de los otros.

Los correspondientes  $d_H^l$  para cada laboratorio, obtenidos para la metodología  $PA$  y  $L_2$ , son mostrados en el gráfico de barras de la Figura 4.2.b y la Figura 4.2.d; en líneas punteadas horizontales se muestra sus valores críticos  $c^H$ . En particular, para una de las 1000 simulaciones realizadas, ambas metodologías identifican al séptimo



**Figura 4.2:** (a) Distribución estimada por remuestreo bootstrap del estadístico  $d_{PA}^H$ . (b) Estadístico  $d_{PA}^H$  para cada muestra de laboratorio. (c) Distribución estimada por remuestreo bootstrap del estadístico  $d_{L_2}^H$ . (d) Estadístico  $d_{L_2}^H$  para cada muestra de laboratorio.

laboratorio (cuyos datos han sido simulados bajo  $H_1$ ) como inconsistente, bajo los siguientes criterios:

- La metodología  $PA$  identifica  $(d_{PA}^H)_7 < c_{PA}^H$ ; representada en la Figura 4.2.b como la línea horizontal punteada que determina el borde por debajo del cual un laboratorio es considerado.
- La metodología  $PA$  identifica  $(d_{L_2}^H)_7 > c_{L_2}^H$ ; representada en la Figura 4.2.d como la línea horizontal punteada que determina el borde por encima del cual un laboratorio es considerado atípico.

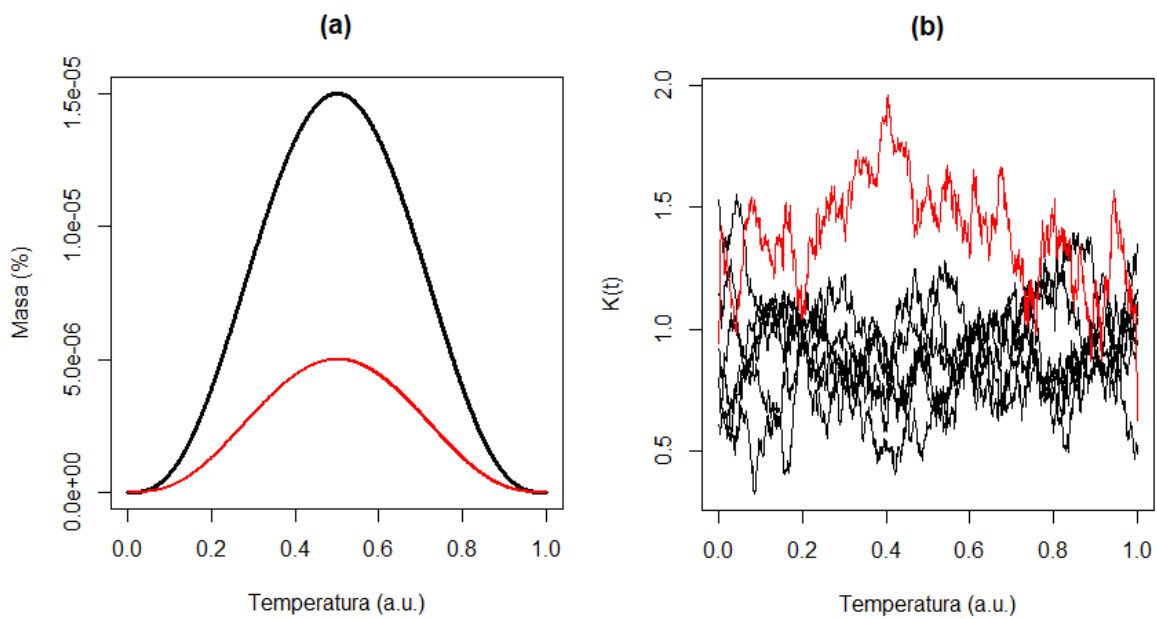
### Hipótesis alternativa de repetibilidad

La hipótesis alternativa de repetibilidad se formulan en términos de  $v_0$ ,  $\delta_k$  y por ende se define por  $H_1 : v_1 = v_0 (1 + \delta_K)$ , donde  $0 \leq \delta_K \leq 2$ .

Las variaciones incluidas en la hipótesis alternativa han sido creadas bajo el criterio de estudios previos [39]. Las diferencias entre la varianza teórica  $v_0 = 5$  y la varianza

alternativa  $v_1 = v_0(1 + 2)$  con  $\delta_K = 2$ , son ligeras y pueden verse en la Figura 4.3.

En este caso, el estudio de simulación se realiza considerando las especificaciones establecidas en la simulación del incumplimiento de la hipótesis nula de reproducibilidad. De igual manera, con el fin determinar la capacidad para determinar inconsistencias, se prueba la hipótesis nula de repetibilidad utilizando los estadísticos de contraste  $d^K$  correspondientes a las metodologías  $PA$  y  $L_2$ , a partir de la estimación del estadístico funcional  $K(t)$ . El estadístico funcional  $K(t)$  para cada laboratorio simulado se puede ver en la Figura 4.3.b.

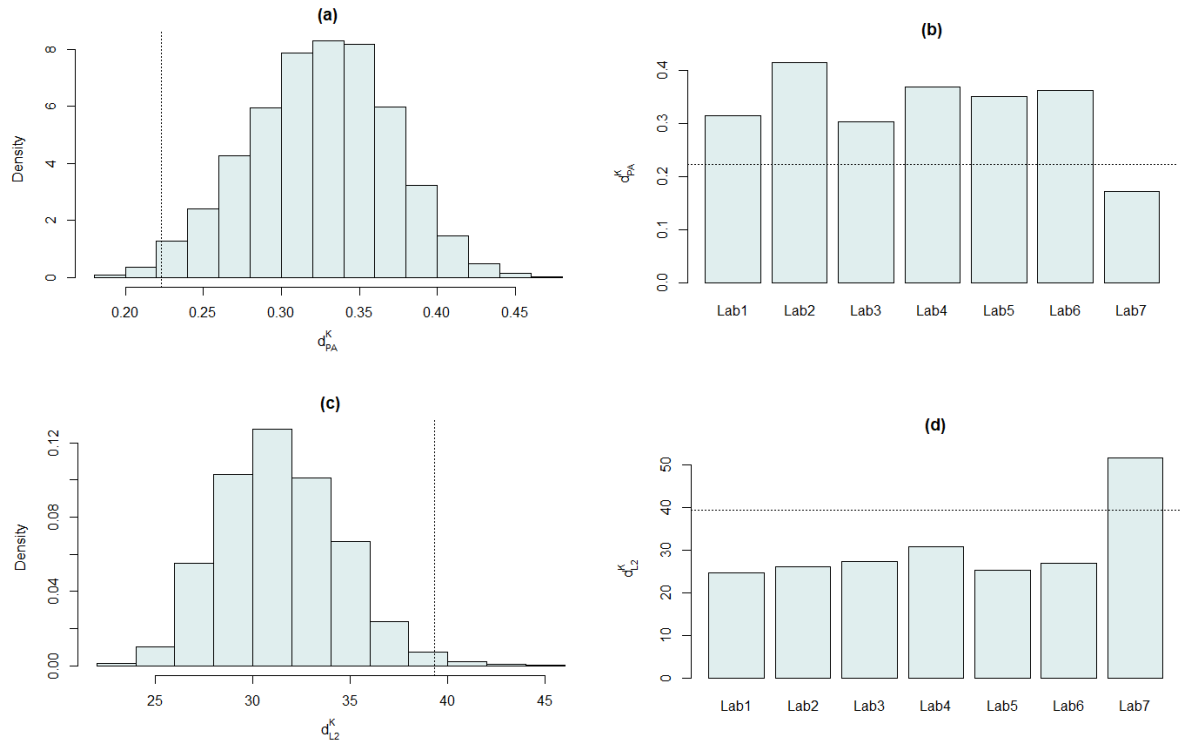


**Figura 4.3:** (a) Varianzas teóricas para curvas TG simuladas bajo la hipótesis nula  $H_0 : v_0 = 5$  (línea negra) y la alternativa,  $H_1 : v_1 = v_0(1 + \delta_K)$  ( $\delta_K = 2$ , línea roja). (b) Estadístico funcional  $K(t)$  para cada uno de los 6 laboratorios simulados bajo la nula (línea negra) y un laboratorio bajo la alternativa (línea roja).

Los resultados para el estudio de simulación bajo la hipótesis alternativa de repetibilidad se muestran en la Figura 4.4:

Las realizaciones de la muestra con  $l = 1, 2, \dots, 7$  laboratorios del estadístico funcional  $K_l(t)$  se muestran en la Figura 4.3.b, mientras la distribución bootstrap de los estadísticos de prueba  $d_K^l$ , para la metodología  $PA$  y la metodología  $L_2$ , se representan gráficamente en la Figura 4.4.a y la Figura 4.4.c, incluyendo sus correspondientes valores críticos  $c^K$ , representados con una línea vertical punteada.

Como en el escenario anterior, se toma en particular una de las 1000 simulaciones, donde se da el mismo resultado para ambas metodologías, es decir, el séptimo labo-



**Figura 4.4:** (a) Distribución estimada por remuestreo bootstrap del estadístico  $d_{PA}^K$ . (b) Estadístico  $d_{PA}^K$  para cada muestra de laboratorio. (c) Distribución estimada por remuestreo bootstrap del estadístico  $d_{L_2}^K$ . (d) Estadístico  $d_{L_2}^K$  para cada muestra de laboratorio.

ratorio es considerado como atípico bajo los siguientes criterios:

- La metodología  $PA$  identifica  $(d_{PA}^K)_7 < c_{PA}^K$ ; representada en la Figura 4.4.b como la línea horizontal punteada que determina el borde por debajo del cual un laboratorio podría considerarse atípico.
- La metodología  $PA$  identifica  $(d_{L_2}^K)_7 > c_{L_2}^K$ ; representada en la Figura 4.4.d como la línea horizontal punteada que determina el borde por encima del cual un laboratorio podría considerarse atípico.

### Efectos de tamaño de muestra

En este apartado se pretende estudiar el efecto del tamaño de la muestra en el desempeño del contraste de hipótesis realizada en las metodologías  $PA$  y  $L_2$ , a partir de los estadísticos  $d^H$  y  $d^K$  utilizando un nivel de significación de  $\alpha = 0.01$ .

En cada escenario simulado, se utilizó un tamaño de muestra  $n = \{10, 20\}$  para cada laboratorio y en particular para la metodología  $PA$  se utilizó un conjunto de direcciones aleatorias, notadas por  $RP$ , con diferente tamaño ( $RP = \{30, 50\}$ ). Con ello, el objetivo es evaluar la consistencia y el desempeño de los estadísticos  $H$  y  $K$  funcionales a través del estudio de la potencia de la prueba. Los valores de potencia obtenidos para los estadísticos  $d^H$  y  $d^K$  se muestran en la Tabla 4.3 y en la Tabla 4.4 respectivamente tanto para la metodología  $PA$  como para la metodología  $L_2$ .

$L$	$n$	$m_a$	Metodología $PA$		Metodología $L_2$
			$pot_{PA} (RP=30)$	$pot_{PA} (RP=50)$	$pot_{L_2}$
7	10	0.4975	0.968	1.000	1.000
7	10	0.4980	0.893	0.868	0.984
7	10	0.4990	0.494	0.438	0.552
7	10	0.50	0.011	0.013	0.014
7	10	0.5010	0.412	0.481	0.563
7	10	0.5020	0.863	0.892	0.978
7	10	0.5025	0.941	0.988	0.998
7	20	0.4975	0.990	1.000	1.000
7	20	0.4980	0.961	0.986	1.000
7	20	0.4990	0.686	0.744	0.864
7	20	0.50	0.007	0.010	0.013
7	20	0.5010	0.653	0.755	0.871
7	20	0.5020	0.975	0.987	1.000
7	20	0.5025	0.991	1.000	1.000

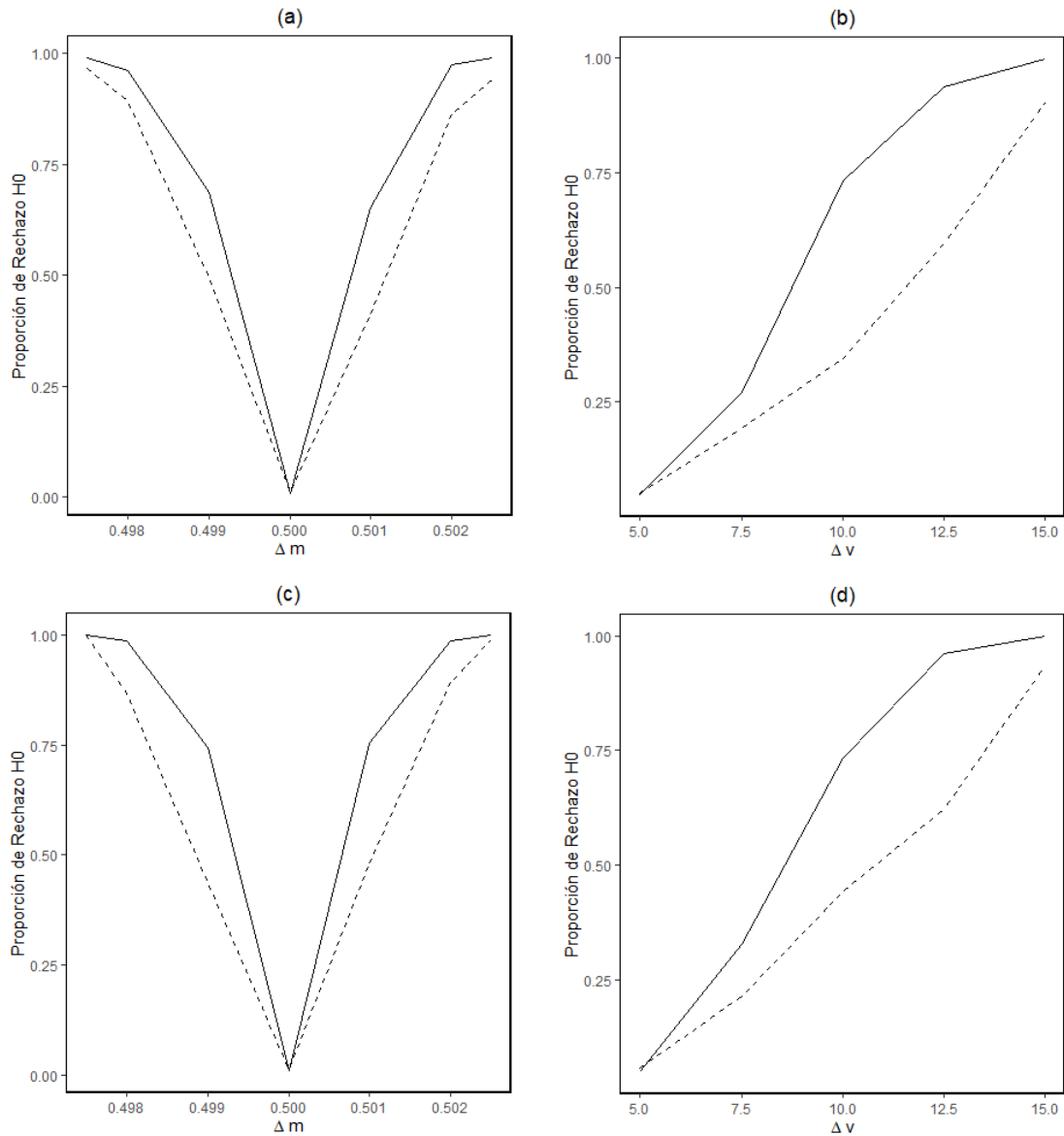
**Tabla 4.3:** Valores de potencias obtenidos bajo la hipótesis alternativa de reproducibilidad de los estadísticos  $d_{PA}^H$  y  $d_{L_2}^H$ .

$L$	$n$	$v_a$	Metodología $PA$		Metodología $L_2$
			$pot_{PA} (RP=30)$	$pot_{PA} (RP=50)$	$pot_{L_2}$
7	10	5.0	0.052	0.057	0.062
7	10	7.5	0.194	0.214	0.281
7	10	10.5	0.344	0.441	0.449
7	10	12.5	0.596	0.621	0.687
7	10	15.0	0.901	0.933	0.936
7	20	5.0	0.048	0.048	0.054
7	20	7.5	0.272	0.327	0.338
7	20	10.5	0.731	0.733	0.756
7	20	12.5	0.937	0.961	0.967
7	20	15.0	0.998	1.000	1.000

**Tabla 4.4:** Valores de potencias obtenidos bajo la hipótesis alternativa de repetibilidad de los estadísticos  $d_{PA}^K$  y  $d_{L_2}^K$ .

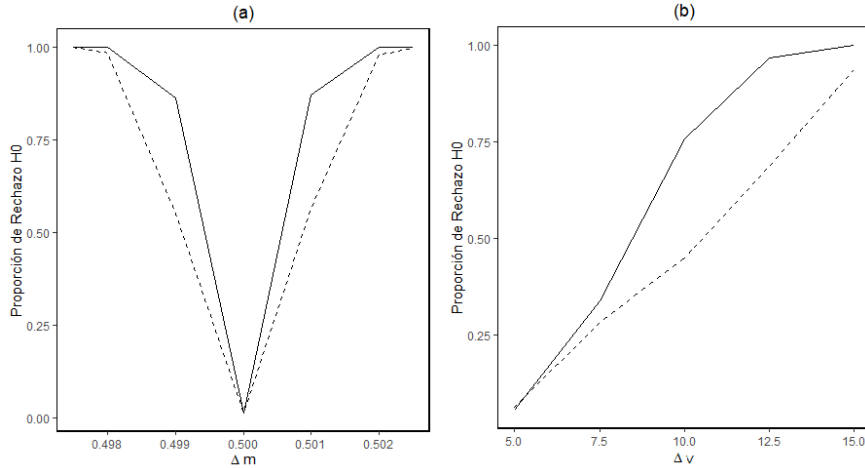


Con respecto a la metodología  $PA$ , la Figura 4.5 muestra la proporción de rechazos bajo la hipótesis nula para las estadísticas,  $d_{PA}^H$  y  $d_{PA}^K$  en los paneles (a) y (b) considerando  $RP = 30$  y en los paneles (c) y (d) considerando  $RP = 50$  direcciones aleatorias. El tamaño de muestra varía entre  $n = 10$  (línea punteada) y  $n = 20$  (línea continua). Los resultados indican que los valores de potencia aumentan, para ambos estadísticos, conforme aumenta el tamaño de la muestra y el número de direcciones aleatorias.



**Figura 4.5:** (a) Curvas de potencia del estadístico  $d_{PA}^H$  con  $RP = 30$  direcciones aleatorias. (b) Curvas de potencia del estadístico  $d_{PA}^K$  con  $RP = 30$  direcciones aleatorias. (c) Curvas de potencia del estadístico  $d_{PA}^H$  con  $RP = 50$  direcciones aleatorias. (d) Curvas de potencia del estadístico  $d_{PA}^K$  con  $RP = 50$  direcciones aleatorias. Los tamaños de muestra  $n = 10$  se representa en línea punteada y  $n = 20$  línea continua

Con respecto a la metodología  $L_2$ , la Figura 4.6 muestra la proporción de rechazos bajo la hipótesis nula para ambas estadísticas,  $d_{L_2}^H$  en el panel (a) y  $d_{L_2}^K$  en el panel (b). Los tamaños de muestra varían en  $n = 10$  líneas punteadas y para  $n = 20$  líneas continuas. Como puede observarse en ambos paneles, un aumento de la muestra implica mayores proporciones de rechazo fuera de la hipótesis nula, es decir, una mayor potencia de prueba.



**Figura 4.6:** (a) Curvas de potencia del estadístico  $d_{L_2}^H$  con  $RP = 30$  direcciones aleatorias. (b) Curvas de potencia del estadístico  $d_{L_2}^K$  con  $RP = 30$  direcciones aleatorias. (c) Curvas de potencia del estadístico  $d_{L_2}^H$  con  $RP = 50$  direcciones aleatorias. (d) Curvas de potencia del estadístico  $d_{L_2}^K$  con  $RP = 50$  direcciones aleatorias. Los tamaños de muestra  $n = 10$  se representa en línea punteada y  $n = 20$  línea continua

Adicionalmente, comparando los valores de potencia obtenidos por cada metodología se puede notar que la metodología  $L_2$  es más sensible a pequeños cambios. Su potencia incrementa más rápidamente, en relación con la propuesta funcional  $PA$ , conforme la media y varianza alternativa se aleja de la media y varianza teórica. Además, los valores de potencia son adecuados y altos incluso en tamaños de muestra pequeños. Por otro lado, la metodología  $PA$  llega a obtener valores de potencia altos y semejantes a los obtenidos por la metodología  $L_2$  cuando se trabaja con un mayor tamaño de muestra y una mayor cantidad de direcciones aleatorias. Por lo tanto, podemos concluir que el desempeño de la metodología  $PA$  es más preciso en casos de estudio donde se trabaje con tamaños de muestra grande y haciendo uso de una mayor cantidad de direcciones aleatorias. Además, los valores de potencia más altos se obtienen para el estadístico  $d_{PA}^H$  por lo que se puede intuir que la metodología  $PA$  es más sensible a detectar cambios mediante el estadístico  $H$ .

## 4.2. Aplicación de la nueva metodología a casos reales de estudio

Una vez realizado el estudio de simulación, se procede con la aplicación usando datos reales, aspecto necesario para determinar el desempeño y la utilidad de la metodología propuesta.

El procedimiento funcional descrito se aplica a dos casos de estudio real. El primero corresponde a los datos experimentales generados a partir de un análisis termogravimétrico del oxalato de calcio y el segundo corresponde a los datos de confort térmico de una tienda textil. Los conjuntos de datos fueron presentados en detalle en la sección 2.3. Como se mencionó anteriormente, ambos casos fueron estudiados previamente en [10]. El primero de ellos con el fin de evaluar el desempeño de la metodología funcional en base a la norma  $L_2$ , y el segundo para aplicar técnicas correspondientes a gráficos de control estadístico en datos funcionales.

Como punto inicial, se presenta, los resultados obtenidos con la aplicación del nuevo método  $PA$  y se citan los resultados correspondientes al estudio de Flores [28] mediante la norma  $L_2$  para el primer caso de estudio con datos termogravimétricos. Para realizar correctamente el estudio comparativo, se consideran los mismos parámetros en ambas metodologías. Los enfoques funcionales mencionados tienen como objetivo común utilizar la información de toda la curva en los datos, evitando previamente el paso de extracción de características univariantes y proporcionar herramientas FDA, propias de los ILS, mediante la extensión funcional de los estadísticos  $H$  y  $K$  de Mandel.

Es importante recalcar que la identificación de valores atípicos dentro de un ILS es un proceso iterativo por lo que debe ser aplicado consecutivamente hasta que en una determinada iteración no se detecten más laboratorios atípicos. De forma general podemos decir que los resultados obtenidos en esta sección coinciden para ambas metodologías; incluso ambos procedimientos utilizan un mismo número de iteraciones en la aplicación del proceso. Además, los laboratorios identificados como anómalos por los estadísticos de prueba  $d^H$  y  $d^K$  también coinciden; estos resultados pueden evidenciarse a detalle en las siguientes secciones.

Por otro lado, se aplica únicamente la metodología  $PA$  al segundo caso de estudio, con el fin de mostrar que su aplicación puede ser útil y ampliarse a cualquier pro-

blema de estudio real, de carácter funcional y no necesariamente a los relacionados con ILS.

#### 4.2.1. Estudio interlaboratorio usando análisis termogravimétrico

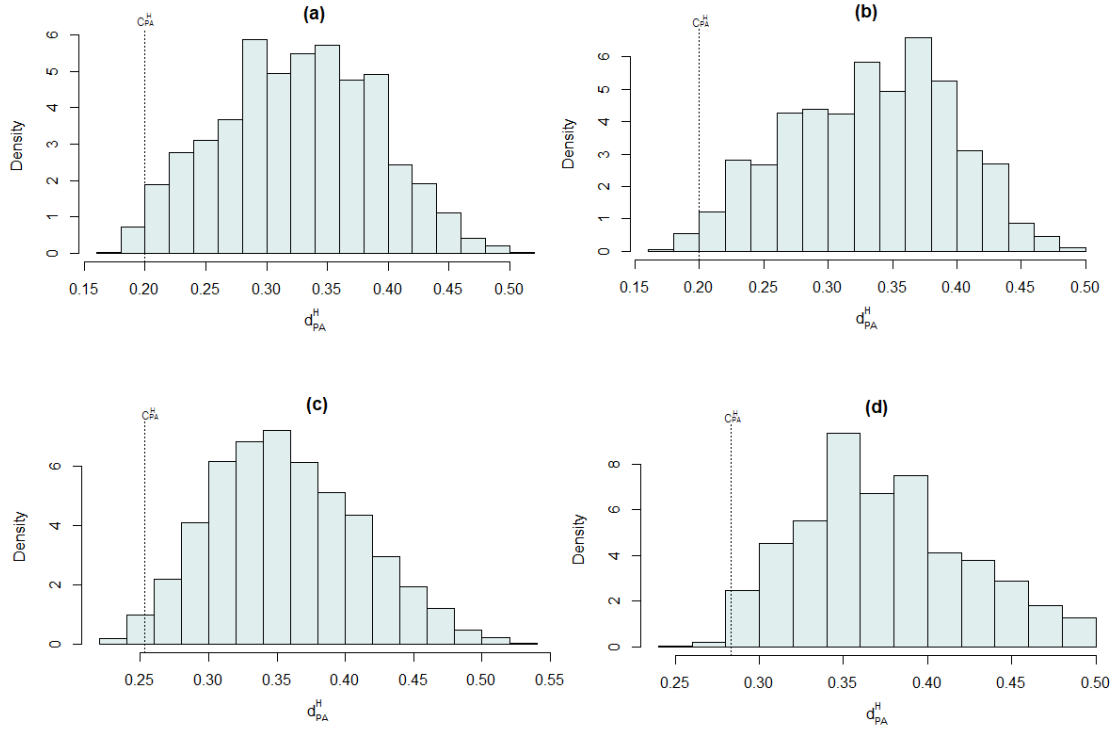
El caso de estudio correspondiente a curvas termogravimétricas fue especificado en el capítulo 2. En resumen, los instrumentos de medida utilizados por cada laboratorio fueron diferentes; el ejemplo indica que hay 4 laboratorios que proporcionan resultados similares y consistentes, mientras que los laboratorios 1, 6 y 7 proporcionan resultados diferentes debido a las condiciones de los sistemas de calibración. El conjunto de datos fue analizado a profundidad en [35].

La aplicación metodológica toma como punto de partida la estimación de los estadísticos funcionales  $H(t)$  y  $K(t)$  con sus estadísticos de prueba  $d_{PA}^H$  y  $d_{PA}^K$ , luego, a partir del procedimiento bootstrap suavizado (utilizando  $B=500$  remuestras) se obtiene los respectivos valores críticos  $c_{PA}^H$  y  $c_{PA}^K$ , definidos como los cuantiles correspondientes al nivel de significación  $\alpha = 0.01$  (por fines comparativos se consideraron los mismos parámetros que los aplicados en la metodología  $L_2$ ). Finalmente, bajo el siguiente criterio, si algún laboratorio  $l$  cumple  $(d_{PA})_l^H < c_{PA}^H$  y/o  $(d_{PA})_l^K < c_{PA}^K$ , entonces el laboratorio se identifica como un valor atípico, en otras palabras, un laboratorio es identificado como atípico cuando su valor de profundidad no alcanza el mínimo valor crítico de profundidad, obtenido a un nivel  $\alpha = 0.01$ .

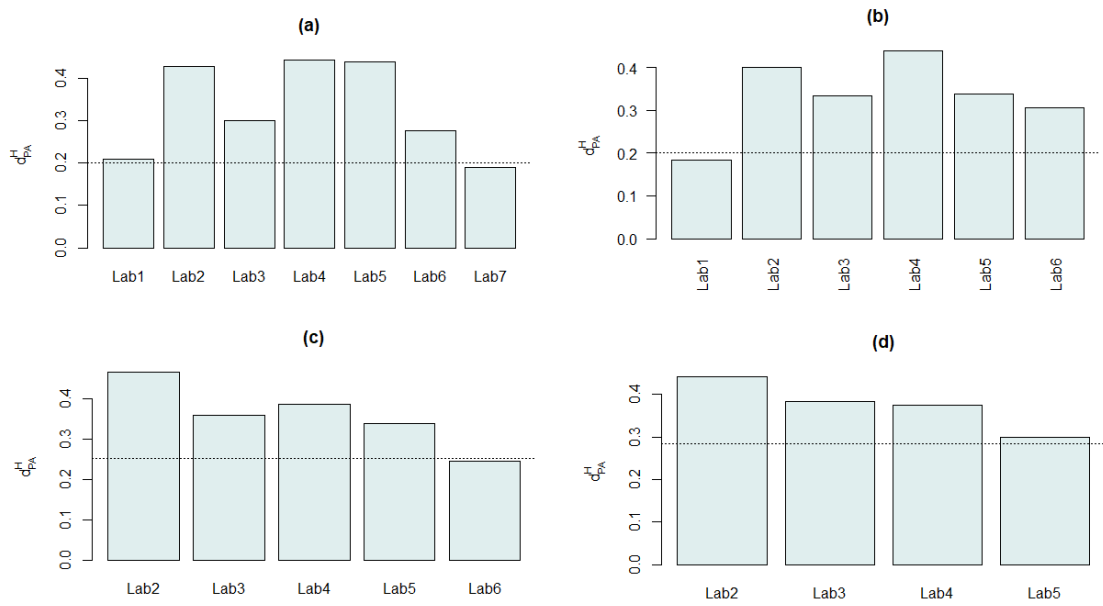
Se realiza entonces, el estudio de reproducibilidad con el contraste de la hipótesis de reproducibilidad en base al estadístico  $H(t)$  y el estadístico de prueba  $d_{PA}^H$ . En cuanto a los resultados, la metodología  $PA$  identifica como atípicos a los laboratorios 1, 6, 7; en un total de 4 iteraciones; de acuerdo con el estudio previo realizado basándose en técnicas FDA de la sección 2.3, los laboratorios que fueron detectados como atípicos eran los esperados.

Con respecto al orden de detección, la Figura 4.8 muestra los laboratorios que no alcanzaron el nivel de profundidad mínima establecido por el cuantil  $c_{PA}^H$ , dentro de cada iteración; así en la primera iteración se identifica como anómalo al laboratorio 7 (instrumento con calibración sesgada), luego en la segunda y tercera iteración se identifican los laboratorios 1 y 6 (instrumento con calibración antigua) respectivamente. Finalmente, en la cuarta iteración no se detecta ningún laboratorio atípico adicional. La Figura 4.7 proporciona el histograma del estadístico  $d_{PA}^H$  en cada ite-

ración y su respectivo valor crítico  $c_{PA}^H$  a un nivel  $\alpha = 0.01$ .

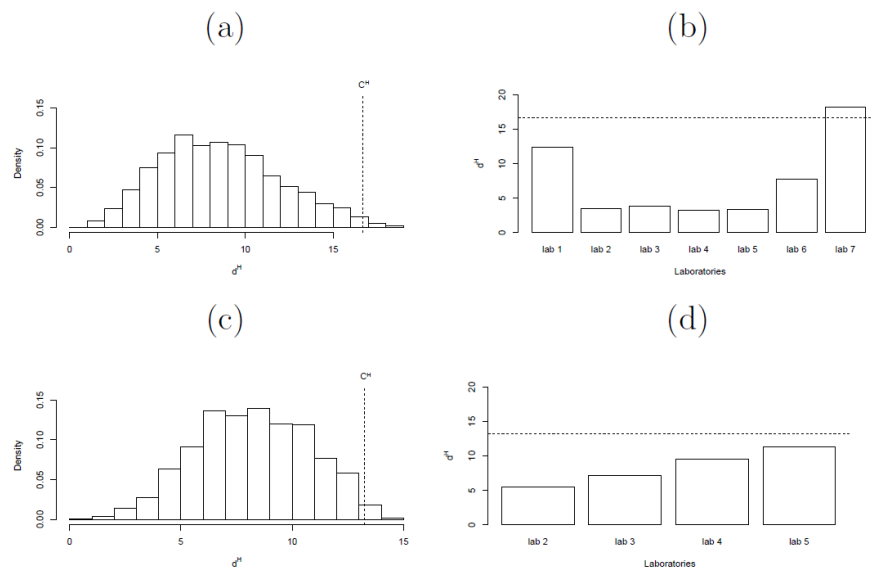


**Figura 4.7:** Histograma del estadístico de prueba  $d_{PA}^H$  y sus valores críticos  $c_{PA}^H$  a un nivel  $\alpha = 0.01$  en cada iteración: (a) primera iteración; (b) segunda iteración; (c) tercera iteración y (d) cuarta iteración.



**Figura 4.8:** Estadísticos  $d_{PA}^H$  para la muestra de los laboratorios en cada iteración: (a) primera iteración; (b) segunda iteración; (c) tercera iteración y (d) cuarta iteración. Los valores críticos se representan con líneas punteadas

Por otro lado, la aplicación del método  $L_2$  en [3] conserva la misma idea, pero teniendo en cuenta que al referirse a una norma, a diferencia del método RP, un laboratorio es detectado como atípico cuando  $(d_{L_2})_l^H > c_{L_2}^H$  y/o  $(d_{L_2})_l^K > c_{L_2}^K$ , es decir, un laboratorio es identificado como atípico cuando su valor de distancia supera el máximo valor crítico a un nivel  $\alpha = 0.01$ . Los resultados obtenidos con el método  $L_2$  fueron similares. Los laboratorios 1, 6 y 7 fueron detectados como atípicos en un total de 4 iteraciones (Ver Figura 4.9). En este caso, al considerar que el estadístico de prueba representa una distancia  $L_2$ , los laboratorios son identificados como atípicos cuando sobrepasan el cuantil  $c_{L_2}^H$ .



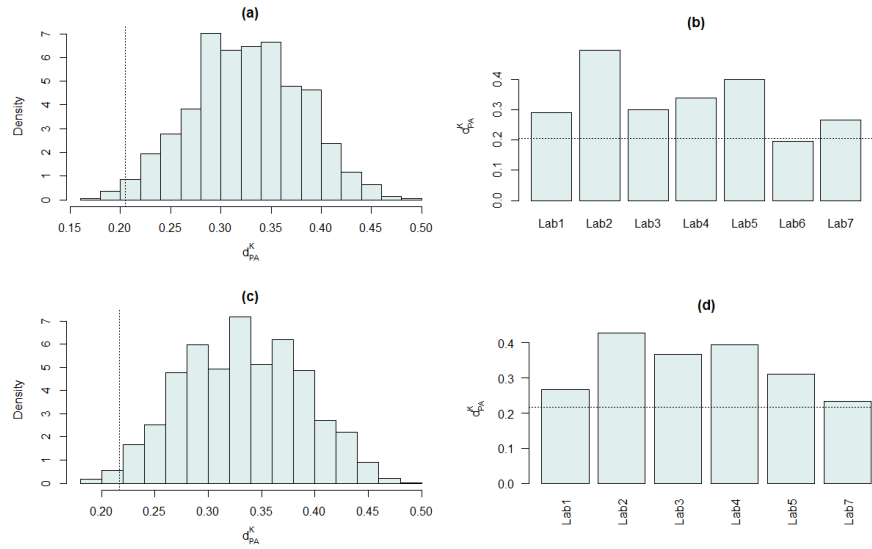
**Figura 4.9:** (a) Histograma del estadístico de prueba  $d_{L_2}^H$  y  $c_{L_2}^H$  para  $\alpha = 0.01$  (primera iteración), (b) Estadísticos  $d_{L_2}^H$  para la muestra de los laboratorios (primera iteración), (c) Histograma del estadístico de prueba  $d_{L_2}^H$  y  $c_{L_2}^H$  para  $\alpha = 0.01$  (cuarta iteración), (d) Estadísticos  $d_{L_2}^H$  para la muestra de los laboratorios (cuarta iteración). Los valores críticos se representan con líneas punteadas.

Fuente: Tomado de [10].

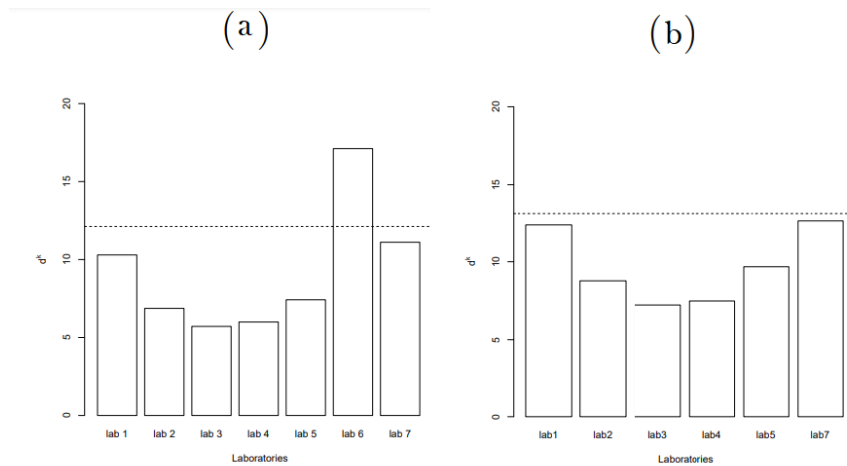
Finalmente, en este ejemplo en particular, podemos notar que para ambas metodologías; desde el punto de vista de la reproducibilidad y desde el estudio del estadístico funcional  $H(t)$ , los laboratorios atípicos se detectaron satisfactoriamente.

Para el estudio de la repetibilidad, se lleva a cabo **mismo** procedimiento aplicando el contraste de hipótesis de repetibilidad en base al  $K(t)$  y el estadístico de prueba  $d_{PA}^K$  al nivel de significancia  $\alpha = 0.01$ . La Figura 4.10 muestra cómo se detecta únicamente el laboratorio 6 en un total de 2 iteraciones, en donde finaliza el proceso iterativo porque no se identifica ningún laboratorio adicional como atípico. La de-

tección de este único laboratorio se debe a que, a pesar de que los laboratorios 1 y 7 son candidatos a ser atípicos, no hay suficientes evidencias a un nivel  $\alpha = 0,01$  de que sean considerados como atípicos en términos de variabilidad intralaboratorio. Los resultados son similares para el estudio de repetibilidad mediante el método  $L_2$  y pueden observarse en la Figura 4.11.



**Figura 4.10:** Histograma del estadístico de prueba  $d_{PA}^K$  y su valor crítico  $c_{PA}^K$  para  $\alpha = 0.01$  ((a) primera iteración y (b) segunda iteración). Estadísticos  $d_{PA}^K$  para la muestra de los laboratorios ((c) primera iteración, (d) segunda iteración). Los valores críticos  $c_{PA}^K$  se representan con líneas punteadas.



**Figura 4.11:** (a) Estadísticos  $d_{L2}^K$  para la muestra de los laboratorios (primera iteración), (b) Estadísticos  $d_{L2}^K$  para la muestra de los laboratorios (cuarta iteración). Los valores críticos  $c_{L2}^K$  se representan con líneas punteadas.

Fuente: Adaptado de [10].

En conclusión, los laboratorios 1, 6 y 7 han sido identificados como atípicos. El análisis mediante técnicas FDA (Sección 2.3), son coherentes con los resultados obtenidos; es decir estos laboratorios proporcionan resultados diferentes con relación a los demás. En el estudio de reproducibilidad, el estadístico  $H(t)$  y el estadístico de prueba  $d^H$ , en ambos métodos, han detectado los laboratorios 1, 6 y 7, en un mismo número de iteraciones. Por otro lado, el estudio de repetibilidad se realiza mediante el estadístico  $K(t)$  y  $d^K$  donde se detecta al laboratorio 6. En cuanto a las metodologías, ambas funcionan correctamente y para este conjunto de datos en particular proporcionan los mismos resultados en cuando a los laboratorios detectados por cada estadístico y el número de iteraciones realizadas en cada caso.

En este sentido, podemos decir que las extensiones de los estadísticos  $h$  y  $k$  al caso FDA y el uso de proyecciones aleatorias proporcionan una nueva herramienta útil de realizar la detección de laboratorios en un ILS haciendo uso de datos funcionales sin la necesidad de extraer características univariantes específicas.

#### 4.2.2. Identificación de sensores anómalos en superficies comerciales

En esta sección se considera un caso de estudio adicional para la aplicación de la metodología *AP*.

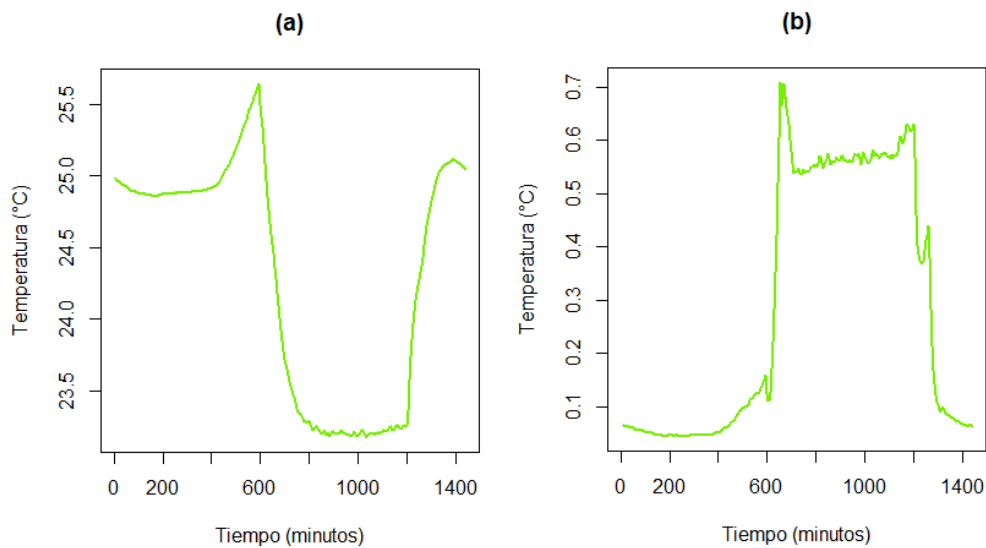
En este ejemplo se muestran los resultados obtenidos de la aplicación del método *PA* al conjunto de datos relacionado al confort térmico de superficies comerciales descrito en el capítulo 2, como se mencionó anteriormente, se realiza con la finalidad de evidenciar que la nueva propuesta funcional, puede ser aplicado para solucionar diversos problemas reales dentro del ámbito funcional y no solo aquellos que se desenvuelven dentro de un ILS.

Para recapitular, en este caso existe un registro de la temperatura de una habitación en una tienda de un edificio comercial de la ciudad de Panamá. Los registros son obtenidos a partir de las mediciones de tres sensores diferentes que se encuentran ubicados dentro de una misma habitación. La Figura 2.4 muestra los datos originales; curvas de temperatura medidas en una frecuencia de 5 min. El horario de la tienda está comprendido entre 10 y 21, cuando la instalación de ventilación y aire acondicionado está en funcionamiento [10].

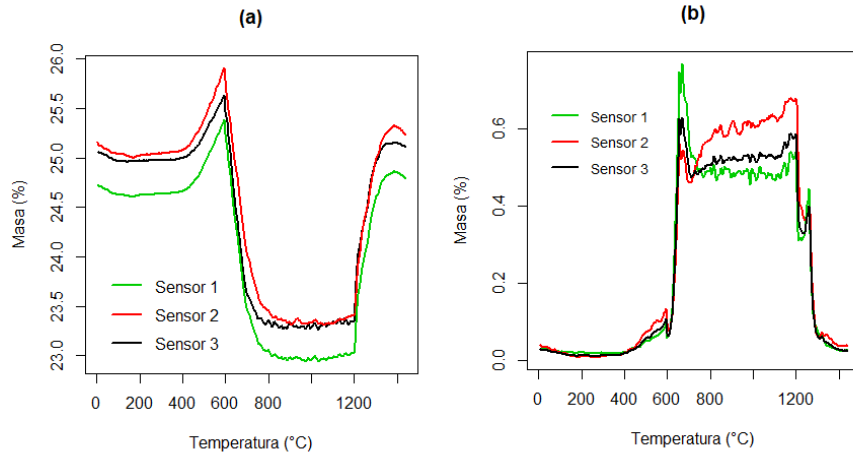


En este ejemplo se pretende detectar si los resultados obtenidos por los tres sensores son similares o existe alguno de ellos que difiera de los demás. En particular, los responsables de mantenimiento indicaron que el sensor 1 se encuentra en correcto funcionamiento por lo que sirve de referencia para analizar el comportamiento de los dos sensores restantes. En ese sentido, el objetivo es evaluar si los sensores 2 y 3 proporcionan mediciones consistentes y semejantes a las obtenidas por el sensor 1. Cabe recalcar que el estudio de este conjunto de datos realizado en [10], concluye que los sensores 2 y 3 deben ser cambiados y no están funcionando correctamente, por lo que se espera que la metodología *AP* proporcione los mismos resultados.

Así, para dar inicio al estudio y tener una noción previa del comportamiento de las curvas, para cada uno de los sensores, se realiza un breve análisis descriptivo FDA de las curvas de temperatura mediante el paquete *ILS*. Los resultados se muestran en la Figura 4.12 y Figura 4.13, en donde se observa que el comportamiento de las curvas de los sensores 2 y 3 son muy similares sin embargo difieren con las estimaciones presentadas por el sensor 1.

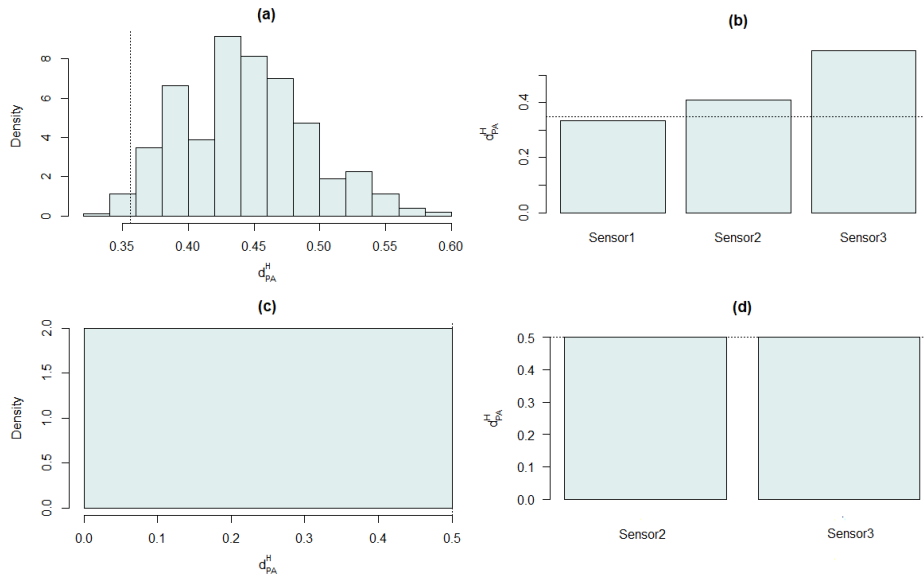


**Figura 4.12:** (a) Media funcional global para sensores térmicos. (b) Varianza funcional global para sensores



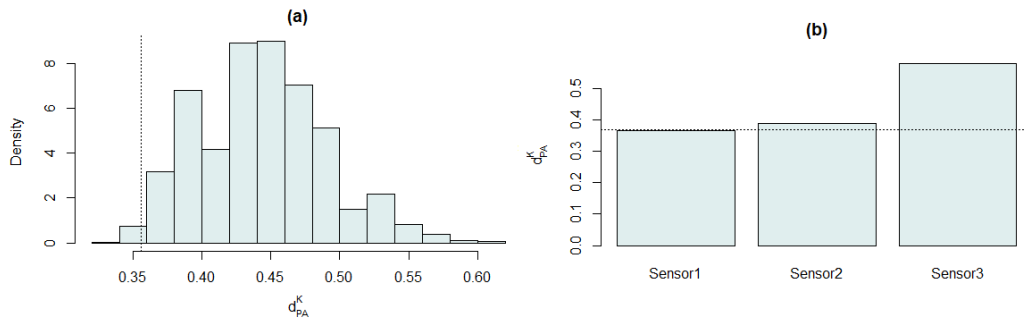
**Figura 4.13:** (a) Media funcional para sensores térmicos. (b) Varianza funcional para sensores.

En la Figura 4.14, se muestra gráficamente las estadísticas escalares de prueba  $d_{PA}^H$  y su correspondiente funcional  $H(t)$  que permiten medir la variabilidad entre laboratorios. Los resultados muestran que el sensor 1 se identifica como un valor atípico si se compara con los sensores 2 y 3. Al tener certeza que este sensor se encuentra en buen funcionamiento, la interpretación da lugar a que tanto el sensor 2 y el sensor 3 proporcionan resultados significativamente diferentes con relación a los referenciales proporcionados por el sensor 1.



**Figura 4.14:** Histograma del estadístico de prueba  $d_{PA}^H$  y su valor crítico  $c_{PA}^H$  para  $\alpha = 0.01$  ((a) primera iteración y (b) segunda iteración). Estadísticos  $d_{PA}^H$  para la muestra de los laboratorios ((c) primera iteración, (d) segunda iteración). Los valores críticos  $c_{PA}^H$  se representan con líneas punteadas.

Por otro lado, para estudiar la variabilidad interna de un mismo sensor en relación con la del resto de sensores se emplea el estadístico  $K(t)$  y su correspondiente  $d_{PA}^K$ , en este caso el sensor 1 estuvo cerca de ser detectado a diferencia de los sensores 2 y 3. Esto demuestra que existen diferencias significativas entre el sensor 1 y los restantes. Los resultados se muestran en la Figura 4.15



**Figura 4.15:** (a) Histograma del estadístico de prueba  $d_{PA}^K$  y su valor crítico  $c_{PA}^K$  para  $\alpha = 0.01$ . (b) Estadísticos  $d_{PA}^K$  para la muestra de los valores críticos  $c_{PA}^K$  se representan con líneas punteadas.

Así, se concluye que los resultados respaldan las declaraciones emitidas por los gerentes de la tienda; es decir los sensores 2 y 3 no tienen un correcto funcionamiento en relación con el sensor 1 que se encuentra en correcto funcionamiento y por lo tanto debió ser analizado y tomarse medidas correctivas de ser necesario.

# Capítulo 5

## Conclusiones y Recomendaciones

### 5.1. Conclusiones

En el presente trabajo de investigación se ha propuesto una nueva metodología para la detección de laboratorios atípicos en estudios interlaboratorio (*ILS*) en el ámbito del análisis de datos funcionales (*FDA*). El propósito es continuar con la línea de estudios realizada previamente por nuestro tutor de trabajo de titulación y proporcionar herramientas útiles para casos específicos en los que la nueva propuesta pueda presentar mejores resultados.

Esta propuesta también abarca el análisis de datos donde la componente funcional prevalece, tanto en el contexto de datos de laboratorios como en los datos de sensores ya que su aplicación puede extenderse a cubrir cualquier problemática real cuando los datos son de carácter funcional, es decir, medidas tomadas a lo largo de un intervalo de tiempo.

El punto focal en el desarrollo del presente trabajo de titulación es el estudio de las extensiones funcionales de los estadísticos univariantes  $h$  y  $k$  de Mandel, porque permiten tratar con mayor precisión los datos funcionales o curvas obtenidas en los ensayos experimentales de laboratorios o por instrumentos que proporcionan grandes cantidades de medidas a lo largo del tiempo.

Las extensiones funcionales de los estadísticos  $H$  y  $K$  de Mandel han sido desarrolladas para identificar los laboratorios que proporcionan resultados inconsistentes o atípicos en un estudio *ILS*. Su importancia radica en que permiten utilizar to-

da la información de la curva, evitando así la extracción previa de características univariantes, factor importante, dado que en la actualidad la mayoría de estudios realizados en el ámbito de los ILS corresponden a enfoques univariantes que no funcionan adecuadamente en este tipo de datos. Los resultados obtenidos por metodologías univariantes ante técnicas FDA mostraron que un laboratorio puede ser identificado como atípico dependiendo de las características extraídas y analizadas, generando así resultados variables además de poco fiables y precisos.

En consecuencia, se propone un enfoque funcional a partir de los estadísticos  $H$  y  $K$  y el uso de proyecciones aleatorias como medida de profundidad de datos. Adicionalmente se realizó un estudio comparativo con el enfoque funcional propuesto previamente en base a la norma  $L_2$  que estudia las distancias de las curvas más lejanas con respecto al origen haciendo uso de los estadísticos funcionales  $H$  y  $K$ .

Considerando los resultados obtenidos, con respecto a la nueva propuesta metodológica  $PA$ , como principales conclusiones podemos enunciar las siguientes:

- Los estadísticos de prueba se obtienen a partir de proyecciones aleatorias. El estadístico  $d_{PA}^H$  permite probar la hipótesis de reproducibilidad, evaluando la variabilidad interlaboratorio, mientras que el estadístico  $d_{PA}^K$  evalúa la variabilidad intralaboratorio. Las distribuciones de probabilidad para estos estadísticos se estiman mediante técnicas de remuestreo bootstrap suavizado y, por lo tanto, los valores críticos  $c_{PA}^H$  y  $c_{PA}^K$  se pueden obtener como los cuantiles correspondientes  $\alpha = 0.01$ .
- Para generar muestras consistentes durante el remuestreo es importante descartar aquellas curvas atípicas; para ello se requiere del cálculo de la profundidad de datos.
- El estudio de simulación realizado mediante curvas TG ha aportado información sobre la validez y potencia de las nuevas extensiones funcionales de los estadísticos  $h$  y  $k$ , además de los efectos que tiene el tamaño de la muestra y de la cantidad de direcciones aleatorias utilizadas durante la proyección de las curvas. El estudio de simulación bajo las hipótesis nulas de reproducibilidad o repetibilidad respaldan que la metodología  $PA$  no rechaza más del nivel de significación establecido. Además, todos los laboratorios atípicos simulados, bajo la hipótesis alternativa, han sido identificados con éxito mediante la aplicación de los estadísticos funcionales  $H$  y  $K$  y sus respectivos estadísticos de contraste  $d_{PA}^H$  y  $d_{PA}^K$ .

- Con respecto a la potencia de prueba, como se esperaba, el aumentar el número de réplicas de laboratorio y la cantidad de proyecciones aleatorias también aumenta la potencia de las pruebas de extensión funcional.
- En cuanto al desempeño de la metodología  $PA$  con respecto a la metodología  $L_2$  se puede concluir que la propuesta  $L_2$  es más sensible a pequeños cambios; pues su potencia incrementa rápidamente conforme la media y varianza alternativa se alejan de la media y varianza teórica. Además, los valores de potencia son adecuados y altos incluso cuando el tamaño de muestra es pequeño. Sin embargo, la metodología  $PA$  llega a obtener valores de potencia altos y semejantes a los obtenidos por la metodología  $L_2$  cuando se trabaja con un mayor tamaño de muestra y una mayor cantidad de direcciones aleatorias. Por lo tanto, el desempeño de la metodología  $PA$  es más preciso en casos en los que se trabaje con tamaños de muestra grande y haciendo uso de una mayor cantidad de direcciones aleatorias. Además, los valores de potencia más altos se obtienen para el estadístico  $d_{PA}^H$  por lo que se puede intuir que la metodología  $PA$  es más sensible a detectar cambios mediante el estadístico  $H$ .
- Finalmente, la aplicación del enfoque FDA a los datos reales termogravimétricos ha demostrado que la metodología  $PA$  propuesta es capaz de detectar todos los laboratorios atípicos reales y se corrobora al coincidir con los resultados obtenidos previamente con la aplicación de la metodología  $L_2$ . Adicionalmente los estadísticos  $H$  y  $K$  también muestran el intervalo de tiempo/temperatura donde los laboratorios se vuelven atípicos, proporcionando información relevante sobre los procesos físicos o químicos; lo que no se podría evidenciar si se analizara únicamente una característica específica. De forma similar sucedió con el segundo caso de estudio correspondiente a la identificación exitosa de sensores de temperatura inconsistentes, por lo que este aporte extiende su aplicación a diversas problemáticas como por ejemplo el confort térmico y la eficiencia energética.

Por último, es importante recalcar que uno de los principales objetivos que se pretendió alcanzar con esta metodología es identificar de forma directa los laboratorios o sistemas de medida atípicos. Esto representa una mejora significativa frente a técnicas FDA que pueden ser aplicadas en los ILS ya que las mismas permiten únicamente detectar observaciones o curvas atípicas para posteriormente analizar su influencia en todo el conjunto de datos dentro del laboratorio o sistema de medida y por ende concluir si efectivamente corresponden a grupos inconsistentes. La

actual propuesta funcional al proporcionar resultados directos en la identificación de laboratorios y sistemas de medida atípicos permite ahorrar tiempo y recursos en investigaciones relacionadas a los ILS y a problemáticas que se desenvuelvan dentro del ámbito funcional.

## 5.2. Recomendaciones y futuras líneas de estudio

El aporte de este trabajo consistió en una nueva propuesta metodológica que pretende solucionar problemas actuales en la industria y laboratorios relacionados con el control de la calidad. Con esto se ha contribuido en la línea de investigación: detección de laboratorios atípicos en estudios interlaboratorio.

En esta memoria se han podido identificar varias oportunidades para seguir generando nuevas adaptaciones al caso funcional de herramientas informáticas y metodologías *FDA* útiles en el control de la calidad. Así se puede realizar la implementación en el paquete estadístico *ILS* del software de código abierto R (creado por el actual tutor de tesis), de las metodologías desarrolladas durante el estudio, con la finalidad de ampliar los recursos que este proporciona.

Como se había mencionado; previamente se realizó un estudio donde se propone una metodología no para métrica, para la detección de laboratorios atípicos, mediante técnicas de remuestreo bootstrap y los estadísticos univariantes  $h$  y  $k$  de Mandel. En este caso se puede realizar un estudio comparativo entre el enfoque no paramétrico univariante y la actual propuesta funcional, esto permitirá medir el desempeño de la propuesta no paramétrica ante la Metodología *PA*, ya que dependiendo de la característica extraída, el poder del enfoque funcional puede ser mayor o no que el enfoque escalar. Estos resultados pueden ser analizados en conjuntos con los obtenidos con base en el enfoque clásico paramétrico y las normas ASTM detallado en [10].

Por otro lado, es importante considerar que los estudios de simulación abarcados en esta memoria requirieron de un alto costo computacional y con ello una alta demanda en su tiempo de ejecución, por lo que se recomienda contar con accesos a servidores que dispongan de una adecuada capacidad de procesamiento y que permitan ejecutar varias instrucciones de una manera óptima y en una menor cantidad tiempo. Es importante conocer el manejo y funcionamiento para tener un mejor desempeño en el desarrollo de las investigaciones.

# Anexos A

## Anexo I: Códigos implementados en R

### A.1. Técnicas FDA aplicadas a ILS

En esta sección se muestra los códigos utilizados para la aplicación de técnicas FDA, descritas en el capítulo 2.

#### A.1.1. Análisis exploratorio FDA

##### Media y Varianza funcional global de curvas TG

```
1 library(ILS)
2 par(mfrow = c(1, 2))
3 xlab <- "Temperature ( C )"
4 ylab <- "Masa (%)"
5 main<- "Media Global"
6 main1<- "Varianza Global"
7 mandel_tg <- mandel.fqcs(fqcdata,nb = 1)
8
9 #Media Funcional global
10 plot(mandel_tg[["fqcs"]][["mean"]],main=main, xlab=xlab, ylab=ylab,col
    = "blue")
11
12 #Varianza funcional global
13 plot(mandel_tg[["fqcs"]][["S2"]],main=main1, xlab=xlab, ylab=ylab,col
    = "blue")
```

##### Media y Varianza Funcional de curvas TG para cada laboratorio



```

1 #Librerias utilizadas: ILS
2 library(ILS)
3
4 #Conjuntos de datos correspondientes a curvas TG
5 data(TG)
6
7 #Media Funcional obtenida para cada laboratorio
8 fqcddata <- ils.fqcddata(TG, p = 7)
9 xlab <- "Temperatura ( C )"
10 ylab <- "Masa (%)"
11 main <- "Curvas TG obtenidas a partir de oxalato de Calcio"
12 plot(x = fqcddata, main = main, xlab=xlab, ylab=ylab, col = 1:7, legend =
    TRUE)
13
14 #Varianza funcional obtenida para cada laboratorio
15 xlab <- "Temperatura ( C )"
16 ylab <- "Varianza"
17 main <- "Varianzas Funcionales curvas TG"
18 mandel_tg <- mandel.fqcs(fqcddata, nb = 1)
19 plot(mandel_tg[["fqcs"]][["S2.i"]], main = main, xlab=xlab, ylab=ylab,
    col = 1:7)

```

## A.1.2. Bootstrap Suavizado

```

1
2 # Paso 1 SACAR AT PICOS ,
3 #BASADOS EN LA PROFUNDIDAD DE PROYECCIONES ALEATORIAS
4
5 out <- outliers.ils(fdata(datos), fdep = depth.RP, trim=trim)$
6 outliers #% CAMBIAR PARA RESULTADOS PAG 75
7 out <- as.numeric(out)
8 if( length(out) != 0){
9     xx <- datos[-out,]
10 }else {
11     xx <- datos
12 }
13 for(ib in 1:B){
14     fdataobj <- fdata(xx)
15     data <- fdataobj[["data"]]
16     varX <- var(data) * smo
17     bmuestra <- fdataobj[sample(1:nrow(fdataobj), size = n*(L+m),
18         replace = TRUE), ]

```

```

18     bmuestra[["data"]] <- bmuestra[["data"]] + mvrnorm(n = n*(L+m),
19     rep(0, ncol(fdataobj)), varX)
20     #data<-paralelizado[[ib]]
21     data<-cbind(bmuestra[["data"]], grupos)

```

### A.1.3. Bandas de confianza Bootstrap

```

1
2 #Librerias utilizadas: ILS
3 library(ILS)
4
5 #Conjuntos de datos correspondientes a curvas TG
6 data(TG)
7 fqcddata <- ils.fqcddata(TG, p = 7 )
8 xlab <- "Temperature ( C )"
9 ylab <- "Masa (%)"
10
11 #Calculamos la media y varianza funcionales punto a punto
12 mandel_tg <- mandel.fqcs(fqcddata,nb = 200)
13
14 #Graficamos la media funcional
15 plot(x = mandel_tg[["fqcs"]][["mean.i"]], main = "(a)", xlab=xlab,
16      ylab=ylab,
17      col = c("#000000", "#ff0000", "#00cd00", "#0000ff", "#00ffff", "#
18      ffa500", "#a020f0"), legend = F, lwd = 2)
19 legend(x = 50 ,y = 85 ,legend = unique(fqcddata$index.laboratory), lty
20      = 1, lwd = 2,
21      col = c("#000000", "#ff0000", "#00cd00", "#0000ff", "#00ffff", "#
22      ffa500", "#a020f0"), cex = 1, box.col = 0)
23
24 #Graficamos la varianza funcional
25 plot(x = mandel_tg[["fqcs"]][["S2.i"]], main = "(b)", xlab=xlab, ylab=
26      ylab,
27      col = c("#000000", "#ff0000", "#00cd00", "#0000ff", "#00ffff", "#
28      ffa500", "#a020f0"), legend = F, lwd = 2)
29 legend(x = 50 ,y = 0.7 ,legend = unique(fqcddata$index.laboratory), lty
30      = 1, lwd = 2,
31      col = c("#000000", "#ff0000", "#00cd00", "#0000ff", "#00ffff", "#
32      ffa500", "#a020f0"), cex = 1, box.col = 0)

```

### A.1.4. Detección de valores atípicos FDA basado en el cálculo de la profundidad de datos usando proyecciones aleatorias.

```
1 #Librerías utilizadas: ILS
2 library(ILS)
3
4 #Profundidades obtenidas con "fdep= depth.RP" para proyecciones
   aleatorias.Se puede usar depth.FM, depth.Mode,etc.
5
6 out.RP<-outliers.ils (fdata(TG),fdep = depth.RP,trim = 0.1)
7 paste("El n mero de outliers detectados con profundidad proyecciones
   aleatorias es: ", out.RP$n.out, sep = "")
8
9 # Tabla de profundidades obtenidas para cada curva en cada laboratorio
10 data.opc<- cbind(out.RP$depth,rep((1:7),c(rep(15,7))))
11 as.data.frame(data.opc) %>% rename(V1=colnames(data.opc)[1],V2 =
   colnames(data.opc)[2]) %>% group_by(V2) %>% mutate(ind = row_
   number()) %>%spread(key=V2, value = V1) %>% dplyr::select(-c(ind))
12
13 # Valor critico de las profundidades a un nivel de significancia del
   10%
14 paste("Las curvas at picas seran aquellas cuyas profundidades son
   menores a:",quantile(data.opc[,1], 0.1))
```

### A.1.5. Pruebas ANOVA Funcionales

#### ANOVA Funcional de un Factor

```
1 #Librerías utilizadas: fda.usc
2 library(fda.usc)
3
4 laboratorios<-as.factor(IDT[,4])
5
6 #ANOVA un factor
7 Anova_un_factor <- fanova.onefactor(object = fdata(TG),group=
   laboratorios, nboot=1000, plot = T)
8 Anova_un_factor$pvalue
9 Anova_un_factor$stat
10 Anova_un_factor$wm
```

#### ANOVA Funcional vía proyecciones aleatorias

```
1
```

```

2 library(fda.usc)
3 library(ILS)
4
5 A<-fdata(TG,1:1000)
6 data=as.data.frame(TG)
7 group=as.factor(IDT[,4])
8 n=nrow(data)
9 RP=c(15,30) #n mero de proyecciones utilizadas
10 model_group=data.frame(group)
11 resultado=fanova.RPm(A,~group,model_group,RP=c(5,30))
12 summary(resultado)

```

## A.2. Estudio de simulación

En esta sección se muestra los códigos implementados para crear funciones y escenarios necesarios para realizar el estudio de simulación, descrito en el capítulo 3. Los paquetes complementarios usados fueron: parallel, ILS, fda.usc, lubridate, nortest, tidyverse.

### A.2.1. Generación aleatoria de datos funcionales TG

```

1
2 #FUNCION DE GENERACION ALEATORIA DE DATOS FUNCIONALES TG
3 library(fda.usc)
4
5 TG.Simu<-function(n_curvas,mu,v,tt,c=1,b=10, tao = 1,c0=10^-6,scale=1,
6   theta= 0.3){
7   tt1<-c(seq(0,1,len=tt))
8   vmu<-fdata(c/((1+exp(b*(tt1-mu)))^(1/tao)), argvals = tt1)
9   vsigma<-c0*(c0+(v*(1-((tt1/0.5)-1)^2)^3))
10  Y<-matrix(nrow = 0,ncol = tt)
11  vepsilon<-rproc2fdata(n_curvas,mu = rep(0,tt),t=tt1,sigma="
12    vexponential",par.list=list("scale"=scale,"theta" = theta))
13  for (i in 1:n_curvas)
14  { o<-sqrt(vsigma) * vepsilon[["data"]][i,]
15    X<-c(vmu[["data"]])+o
16    Y<-rbind(Y,X)}
17  return(Y)
18 }

```

## A.2.2. Generación de proyecciones aleatorias y profundidades univariantes de Tukey

```
1 #RPM:FUNCION PARA GENERAR PROYECCIONES ALEATORIAS Y PROFUNDIDADES DE
  TUKEY.
2
3 library(ILS)
4 library(fda.usc)
5 library(tidyverse)
6
7 RPM=function(object,data.fac=as.data.frame(as.factor(rep(seq(1:7),1)))
  ,RP=min(30,ncol(object)),      zproj=NULL,par.zproj=list(norm=TRUE)
  , dfunc = "TD1",par.dfunc = list(),      w=rep(1,ncol(
  object)), xeps = 1e-15){
8
9   if (class(object)[1] %in% c("matrix","data.frame")) dataM=as.matrix(
  object)
10  else if (is.fdata(object)) dataM=object[["data"]]#Si es funcional
11  lfac=unlist(lapply(data.fac,is.factor))
12  min.data.fac<-min(table(data.fac[,lfac]))
13  if (min.data.fac==0) warning("Contingency table of factor levels (
  data.fac argument) contains 0 counts values")
14
15  nrow=nrow(dataM)
16  ncol=ncol(dataM)
17  nprRP=max(RP)
18
19  if ((class(object) %in% c("matrix","data.frame")) & is.null(zproj))
  zproj=projMV
20  if ((is.fdata(object)) & is.null(zproj)) zproj=rproc2fdata
21  if (is.fdata(zproj) & is.fdata(object)) {
22    if (nrow(zproj)>=nprRP) {
23      z=zproj[1:nprRP]
24    } else {
25      stop(paste("Not enough functions in zproj",length(zproj)," to
  compute ",nprRP," projections"))
26    }
27  } else if (is.matrix(zproj) & (class(object)[1] %in% c("matrix","
  data.frame"))){
28    if (nrow(zproj)>=nprRP){
29      z=zproj[1:nprRP,] } else {
30      stop(paste("Not enough rows in zproj",nrow(zproj)," to compute
  ",nprRP," projections"))
31    }
  }
```

```

32 } else if (is.function(zproj)){
33   # EN EL CASO QUE EL OBJETO ES FDATA Y ZPROJ UNA FUNCION, SE GENERA
    AQUI LAS DIRECCIONES z
34   if (is.fdata(object)) {z=do.call(zproj,modifyList(list(n=nprRP,t=
    object$argvals, sigma=1, norm=TRUE),par.zproj))}
35   else if (class(object)[1] %in% c("matrix","data.frame")) {z=do.
    call(zproj,modifyList(list(n=nprRP,m=ncol(object)),par.zproj))}
36 } else {stop("Parameter zproj is neither an fdata object or a
    function")}]
37
38 #Implementaci n Profundidad de Tukey
39
40 n <- nrow(object)
41 dep = matrix(0, ncol = nprRP, nrow = n)
42 proyeccion <-matrix(0,ncol = nprRP, nrow = n)
43 if (dfunc %in% c("TD1")) {
44   Fn <- list()
45   for (j in 1:nprRP) {
46     valor = inprod.fdata(object, z[j])
47
48     x<-valor
49     n1 <- nrow(x)
50     m1 <- ncol(x)
51     d1 <- matrix(NA, nrow = n1, ncol = m1)
52     for (i in 1:m1) {
53       Fn[[j]] = ecdf(valor)
54       d1[,i] = pmin(Fn[[j]](x), (1 - Fn[[j]](x - xeps)))
55     }
56     proyeccion[,j]<-valor
57     dep[, j]<- apply(d1, 1, mean, na.rm = TRUE)[1:n1]
58   }
59   dep = apply(dep, 1, mean)
60   proyeccion = apply(proyeccion,1, mean)
61 } else { stop("No onedimensional valid function")}]
62
63 return(list(direcciones=z, dep =dep, proyeccion=proyeccion))
64 }

```

### A.2.3. Generaci3n de estadísticos funcionales H y K y su distribuci3n Bootstrap

```

1 #FUNCION QUE CREA LOS ESTADISTICOS H Y K FUNCIONALES, SU DISTRIBUCION

```

## BOOTSTRAP y RETORNA SUS PROYECCIONES ALEATORIAS Y PROFUNDIDADES UNIVARIANTES

```

2
3 MANDEL_FDA_BOOT<-function(datos = TG, L=6, m=1, B=2, n=15,tt=1000,
  zproj=NULL,grupos=c(rep((1:(L+m)),c(rep(n,(L+m))))), RP=30, smo
  =0.05){
4
5 ##### Paso 1  DETECCION ATIPICOS #####
6   out <- outliers.ils(fdata(datos), fdep = depth.RP, trim=0.01)$
   outliers
7   out <- as.numeric(out)
8   if( length(out) != 0){
9     xx <- datos[-out,]
10  }else {
11    xx <- datos
12  }
13
14 ##### Paso 2  REMUESTRAS CON BOOTSTRAP SUAVIZADO #####
15   Hi.boot.data<-matrix(nrow = 0,ncol = tt+1)
16   Ki.boot.data<-matrix(nrow = 0,ncol = tt+1)
17   dhi.boot.data<-matrix(nrow = 0,ncol = 1)
18   dki.boot.data<-matrix(nrow = 0,ncol = 1)
19   rphi.boot.data<- matrix(nrow = 0,ncol = 1)
20   rpki.boot.data<- matrix(nrow = 0,ncol = 1)
21
22   for(ib in 1:B){
23     fdataobj<-fdata(xx)
24     data <- fdataobj[["data"]]
25     varX <- var(data) * smo
26     bmuestra <- fdataobj[sample(1:nrow(fdataobj), size = n*(L+m),
  replace = TRUE), ]
27     bmuestra[["data"]] <- bmuestra[["data"]] + mvrnorm(n = n*(L+m),
  rep(0, ncol(fdataobj)), varX)
28     data<-cbind(bmuestra[["data"]], grupos)
29
30     #Se obtiene el estadistico H Y k para cada remuestra
31     media<-matrix(nrow = (L+m),ncol = tt+1)
32     desviacion<-matrix(nrow = (L+m),ncol = tt+1)
33     desviacion_s <-matrix(nrow=1,ncol=tt+1)
34
35     for (i in 1:(L+m)) {
36       media[i,]<-apply(data[data[,tt+1] == i,],2,mean)
37       df<-(data[data[,tt+1] == i,]-(matrix(1,nrow =n, ncol = 1) %*%
  matrix(media[i,],nrow =1, ncol = tt+1)))^2

```

```

38     desviacion[i,]<-sqrt(apply(df,2,sum)/n)
39 }
40
41 global_mean<-(apply(media,2,mean))
42 global_sd <-(apply((desviacion)^2,2,sum))/(L+m)
43 dfs<-(media-(matrix(1,nrow =(L+m), ncol = 1) %*%      matrix(global_
mean,nrow =1, ncol = tt+1)))^2
44 desviacion_s<-sqrt(apply(dfs,2,sum)/(L+m-1))
45
46 Hi.boot<-matrix(nrow = (L+m),ncol = tt+1)
47 Ki.boot<-matrix(nrow = (L+m),ncol = tt+1)
48 dhi.boot<- matrix(nrow = (L+m),ncol =1)
49 dki.boot<- matrix(nrow = (L+m),ncol =1)
50 rphi.boot<- matrix(nrow = (L+m),ncol =1)
51 rpki.boot<- matrix(nrow = (L+m),ncol =1)
52
53 #Se forma los estadísticos H y K
54 for (i in 1:(L+m)) {
55     Hi.boot[i,]<-(media[i,]-global_mean)/desviacion_s
56     Ki.boot[i,]<-(desviacion[i,])/sqrt(global_sd)
57 }
58 obj_Hi.boot<-fdata(Hi.boot[,1:tt])
59 proj.dhi.boot<-RPM(object=obj_Hi.boot,
60 data.fac=as.data.frame(as.factor(rep(seq(1:(L+m)),1))),
zproj=zproj, par.zproj=list(norm=TRUE), RP=RP)
61
62 dhi.boot<-proj.dhi.boot$dep      rphi.boot<-proj.dhi.boot$proyeccion
63
64 obj_Ki.boot<-fdata(Ki.boot[,1:tt])
65 proj.dki.boot<-RPM(object=obj_Ki.boot, data.fac=as.data.frame(as.
factor(rep(seq(1:(L+m)),1))), zproj=zproj, par.zproj=
list(norm=TRUE), RP=RP)
66 dki.boot<-proj.dki.boot$dep      rpki.boot<-proj.dki.boot$proyeccion
67
68 Hi.boot.data<-rbind(Hi.boot.data,Hi.boot)
69 dhi.boot.data<-rbind(dhi.boot.data,t(t(dhi.boot)))
70 rphi.boot.data<-rbind(rphi.boot.data,t(t(rphi.boot)))
71
72 Ki.boot.data<-rbind(Ki.boot.data,Ki.boot)
73 dki.boot.data<-rbind(dki.boot.data,t(t(dki.boot)))
74 rpki.boot.data<-rbind(rpki.boot.data,t(t(rpki.boot)))
75 }
76 return(list(Hi.boot = Hi.boot.data , Ki.boot = Ki.boot.data, sample
= data, dhi.boot=dhi.boot.data,      dki.boot=dki.boot.data,rphi.

```



```

boot=rphi.boot.data,rpki.boot=rpki.boot.data,levels = seq(1:7)))
77 }

```

## A.2.4. Escenarios de reproducibilidad (Estadístico H)

La siguiente función fue implementada para crear los escenarios de reproducibilidad bajo la hipótesis nula y alternativa.

```

1 #FUNCION QUE SIMULA ESCENARIOS DE REPRODUCIBILIDAD
2 #SE REQUIERE PARALELIZAR LA FUNCION, POR LO QUE SE CREA PREVIAMENTE LA
  FUNCION PARALELIZAR
3 .paralelizar <- function(x,grupos = c( rep((1:(L+m)),c(rep(n,(L+m))))))
  ,L=L, m=m, tt=tt, n=n, B=B ){
4   data<-cbind(x, grupos) # ADJUNTAMOS INDICES
5   media<-matrix(nrow = (L+m),ncol = (tt+1))
6   desviacion<-matrix(nrow = (L+m),ncol = (tt+1))
7   desviacion_s <-matrix(nrow=1,ncol=(tt+1))
8   Hi.sample.tt1<-matrix(nrow=(L+m),ncol=(tt+1))
9
10  for (i in 1:(L+m)) {
11    media[i,]<-apply(data[data[, (tt+1)] == i,],2,mean)
12    df<-(data[data[,tt+1] == i,]-(matrix(1,nrow =n, ncol = 1) %*%
13      matrix(media[i,],nrow =1, ncol = tt+1)))^2
14    desviacion[i,]<-sqrt(apply(df,2,sum)/n)
15  }
16
17  global_mean<-(apply(media,2,mean)) #media de medias,una curva
18  global_sd <-(apply((desviacion)^2,2,sum))/(L+m)
19  dfs<-(media-(matrix(1,nrow =(L+m), ncol = 1) %*%      matrix(global
20    _mean,nrow =1, ncol = (tt+1))))^2
21  desviacion_s<-sqrt(apply(dfs,2,sum)/(L+m-1))
22
23  for (i in 1:(L+m)) {
24    Hi.sample.tt1[i,]<-(media[i,]-global_mean)/desviacion_s
25  }
26
27  Hi.sample<-Hi.sample.tt1[,1:tt]
28  Obj_Hi.sample<-fdata(Hi.sample)
29  dhi.sample<-RPM(object=Obj_Hi.sample,RP=30,data.fac=as.data.frame(as
30    .factor(rep(seq(1:L+m),1))),          zproj=NULL,par.zproj=
31    list(norm=TRUE),                      w=rep(1,ncol(object)))
32  z<-dhi.sample$direcciones

```

```

29 dhi.sample$proyeccion<-cbind(dhi.sample$proyeccion,seq(1:(L+m)))
30 dhi.sample$dep<-cbind(dhi.sample$dep,seq(1:(L+m)))
31
32 Mandel<-MANDEL_FDA_BOOT(datos = x, L , m, B , n ,tt,zproj=z, RP=30,
    grupos=grupos )
33 fqcdata <- ils.fqcdata(x, p = L+m)
34 mandel.x <- mandel.fqcs(fqcdata,nb = B)
35 dhi.sample.L2 <- mandel.x[["norm.h"]]
36 dhi.boot.L2 <- mandel.x[["distboot.h"]]
37
38 return(list(dhi.sample.proyeccion = dhi.sample$proyeccion,dhi.sample
    .dep = dhi.sample$dep, dhi.boot = Mandel[["dhi.boot"]], rphi.boot
    =Mandel[["rphi.boot"]],dhi.boot.L2 = dhi.boot.L2, dhi.sample.L2 =
    dhi.sample.L2))
39 }
40
41 firstdistboot.H <- function(samples = NULL, mc = 1,n = 30, L= 6, m =
    1,mean = 0.5, mean_a = 0.045, sd =5, B = 500, alpha = 0.01,semilla
    = NULL, tt = 1000, grupos=c( rep((1:(L+m))),c(rep(n
    ,(L+m)))))){
42
43 if (!is.null(semilla)) set.seed(semilla)
44 Hi.sample.data<-matrix(nrow = 0,ncol = tt)
45 dhi.sample.data <-matrix(nrow = 0,ncol =2)
46 rphi.sample.data <-matrix(nrow = 0,ncol =2)
47 Hi.boot.data<-matrix(nrow = 0,ncol = tt)
48 dhi.boot.data<-matrix(nrow = 0,ncol =1)
49 rphi.boot.data<-matrix(nrow = 0,ncol =1)
50 dhi.sample.data.L2 <-matrix(nrow = 0,ncol = (L+m))
51 dhi.boot.data.L2 <-matrix(nrow = 0,ncol = 1)
52
53 if (is.null(samples)){
54     samples <- vector(mode = "list",length = mc)
55 }
56
57 print(paste( " Generar Muestra MC ", now(),sep=" "))
58
59 for(j in 1:mc){
60     if (is.null(samples[[j]])){
61         xl<-TG.Simu(n_curvas = (L*n),mu=mean,v=sd,tt=tt)
62         xm<-TG.Simu(n_curvas = (m*n),mu=mean_a,v=sd,tt=tt)
63         x <- rbind(xl,xm)
64         samples[[j]] <- x
65     }

```

```

66     else{
67         x <- samples
68     }
69 }
70
71 print(paste( "inicia el paralel ", now(), sep=" "))
72 paralelizado <- parLapply(cl, samples, grupos = grupos, L=L, m=m, tt=
    tt, n=n, B=B , paralelizar)
73 print(paste( "termina el paralel ", now(), sep=" "))
74
75 for(j in 1:mc){
76     dhi.sample.data<-rbind(dhi.sample.data, paralelizado[[j]][["dhi.
        sample.dep"]])
77     rphi.sample.data<-rbind(rphi.sample.data, paralelizado[[j]][["dhi.
        sample.proyeccion"]])
78     dhi.boot.data<-rbind(dhi.boot.data, paralelizado[[j]][["dhi.boot"
        ]])
79     rphi.boot.data<-rbind(rphi.boot.data, paralelizado[[j]][["rphi.boot
        "]])
80     #comparaci n con norma L2
81     dhi.sample.data.L2<-rbind(dhi.sample.data.L2, paralelizado[[j]][["
        dhi.sample.L2"]])
82     dhi.boot.data.L2<-rbind(dhi.boot.data.L2, as.matrix(paralelizado[[j
        ]][["dhi.boot.L2"]]))
83 }
84
85 dhi.sample.data <- as.data.frame(dhi.sample.data) %>%
86     rename(V1 = colnames(dhi.sample.data)[1] , V2 = colnames(dhi.
        sample.data)[2]) %>%
87     group_by(V2) %>% mutate(ind = row_number()) %>% spread(key=V2,
        value = V1) %>% dplyr::select(-c(ind))
88
89 rphi.sample.data <- as.data.frame(rphi.sample.data) %>%
90     rename(V1 = colnames(rphi.sample.data)[1] , V2 = colnames(rphi.
        sample.data)[2]) %>%
91     group_by(V2) %>% mutate(ind = row_number()) %>% spread(key=V2,
        value = V1) %>% dplyr::select(-c(ind))
92
93 return(list(dhi.boot=dhi.boot.data, dhi.sample=dhi.sample.data, rphi.
        sample=rphi.sample.data, rphi.boot=rphi.boot.data, dhi.boot.L2=dhi.
        boot.data.L2, dhi.sample.L2 = dhi.sample.data.L2))
94 }

```

### A.2.5. Escenarios de repetibilidad (Estadístico K)

La siguiente función fue implementada para crear los escenarios de repetibilidad bajo la hipótesis nula y alternativa.

```
1 # FUNCION QUE SIMULA ESCENARIOS DE REPRODUCIBILIDAD
2 library(parallel)
3 library(ILS)
4 library(fda.usc)
5 library(lubridate)
6 library(nortest)
7 library(tidyverse)
8
9 #Previamente se realiza la paralelización de la función.
10 paralelizar <- function(x, grupos = c( rep((1:(L+m)), c(rep(n, (L+m))))),
11                          L=L, m=m, tt=tt, n=n, B=B ){
12   data<-cbind(x, grupos) # ADJUNTAMOS INDICES
13   media<-matrix(nrow = (L+m), ncol = (tt+1))
14   desviacion<-matrix(nrow = (L+m), ncol = (tt+1))
15   desviacion_s <-matrix(nrow=1, ncol=(tt+1))
16   Ki.sample.tt1<-matrix(nrow=(L+m), ncol=(tt+1))
17
18   for (i in 1:(L+m)) {
19     media[i,]<-apply(data[data[, (tt+1)] == i,], 2, mean)
20     df<-(data[data[, tt+1] == i,]-(matrix(1, nrow = n, ncol = 1) %*%
21     matrix(media[i,], nrow = 1, ncol = tt+1)))^2
22     desviacion[i,]<-sqrt(apply(df, 2, sum)/n)
23   }
24
25   global_mean<-(apply(media, 2, mean)) #media de medias, una curva
26   global_sd <-(apply((desviacion)^2, 2, sum))/(L+m) # S_r
27   dfs<-(media-(matrix(1, nrow =(L+m), ncol = 1) %*%      matrix(
28     global_mean, nrow = 1, ncol = (tt+1))))^2
29   desviacion_s<-sqrt(apply(dfs, 2, sum)/(L+m))
30
31   for (i in 1:(L+m)) {
32     Ki.sample.tt1[i,]<-(desviacion[i,])/sqrt(global_sd)
33   }
34
35   Ki.sample<-Ki.sample.tt1[, 1:tt]
36   Obj_Ki.sample<-fdata(Ki.sample)
37   dki.sample<-RPM(object=Obj_Ki.sample, RP=30, data.fac=as.data.frame(as
38     .factor(rep(seq(1:(L+m)), 1))),      zproj=NULL, par.zproj=list(norm
39     =TRUE),                                w=rep(1, ncol(object)))
```

```

35
36 z<-dki.sample$direcciones
37 dki.sample$proyeccion<-cbind(dki.sample$proyeccion,seq(1:(L+m)))
38 dki.sample$dep<-cbind(dki.sample$dep,seq(1:(L+m)))
39
40 Mandel<-MANDEL_FDA_BOOT(datos = x, L, m, B, n, tt,zproj=z, RP=30,
    grupos=grupos )
41
42 #Comparaci n con norma L2
43 fqcddata <- ils.fqcddata(x, p = L+m)
44 mandel.x <- mandel.fqcs(fqcddata,nb = B)
45 dki.sample.L2 <- mandel.x[["norm.k"]]
46 dki.boot.L2 <- mandel.x[["distboot.k"]]
47
48 return(list(dki.sample.proyeccion =dki.sample$proyeccion,dki.sample.
    dep = dki.sample$dep,dki.boot = Mandel[["dki.boot"]], rpki.boot=
    Mandel[["rpki.boot"]],dki.boot.L2 = dki.boot.L2, dki.sample.L2 =
    dki.sample.L2))
49 }
50
51 firstdistboot.K <- function(samples = NULL, mc = 1,n = 30, L= 6, m =
    1,mean = 0.5, mean_a = 0.045, sd =5, B = 500, alpha = 0.01,semilla
    = NULL, tt = 1000, grupos=c( rep((1:(L+m)),c(rep(
    n,(L+m)))))){
52
53 if (!is.null(semilla)) set.seed(semilla)
54
55 Ki.sample.data<-matrix(nrow = 0,ncol = tt)
56 dki.sample.data <-matrix(nrow = 0,ncol =2)
57 rpki.sample.data <-matrix(nrow = 0,ncol =2)
58 Ki.boot.data<-matrix(nrow = 0,ncol = tt)
59 dki.boot.data<-matrix(nrow = 0,ncol =1)
60 rpki.boot.data<-matrix(nrow = 0,ncol =1)
61 dki.sample.data.L2 <-matrix(nrow = 0,ncol =L+m)
62 dki.boot.data.L2 <-matrix(nrow = 0,ncol = 1)
63
64 if (is.null(samples)){
65     samples <- vector(mode = "list",length = mc)
66 }
67
68 print(paste( " Generar Muestra MC ", now(),sep=" "))
69
70 for(j in 1:mc){
71     if (is.null(samples[[j]])){

```

```

72     xl<-TG.Simu(n_curvas = (L*n),mu=mean,v=sd,tt=tt)
73     xm<-TG.Simu(n_curvas = (m*n),mu=mean,v=sd_a,tt=tt)
74     x <- rbind(xl,xm)
75     samples[[j]] <- x
76   }
77   else{
78     x <- samples
79   }
80 }
81
82 print(paste( "inicia el paralel ", now(),sep=" "))
83 paralelizado <- parLapply(cl, samples,grupos = grupos, L=L, m=m, tt=
84   tt, n=n, B=B , paralelizar)
85
86 print(paste( "termina el paralel ", now(),sep=" "))
87
88 for(j in 1:mc){
89
90   dki.sample.data<-rbind(dki.sample.data,paralelizado[[j]][["dki.
91     sample.dep"]])
92   rpki.sample.data<-rbind(rpki.sample.data,paralelizado[[j]][["dki.
93     sample.proyeccion"]])
94   dki.boot.data<-rbind(dki.boot.data,paralelizado[[j]][["dki.boot"
95     ]])
96   rpki.boot.data<-rbind(rpki.boot.data,paralelizado[[j]][["rpki.boot
97     "]])
98   dki.sample.data.L2<-rbind(dki.sample.data.L2,paralelizado[[j]][["
99     dki.sample.L2"]])
100   dki.boot.data.L2<-rbind(dki.boot.data.L2,as.matrix(paralelizado[[j
101     ]][["dki.boot.L2"]]))
102 }
103
104 dki.sample.data <- as.data.frame(dki.sample.data) %>%
105   rename(V1 = colnames(dki.sample.data)[1] , V2 = colnames(dki.
106     sample.data)[2]) %>%
107   group_by(V2) %>% mutate(ind = row_number()) %>% spread(key=V2,
108     value = V1) %>% dplyr::select(-c(ind))
109
110 rpki.sample.data <- as.data.frame(rpki.sample.data) %>%
111   rename(V1 = colnames(rpki.sample.data)[1] , V2 = colnames(rpki.
112     sample.data)[2]) %>%
113   group_by(V2) %>% mutate(ind = row_number()) %>% spread(key=V2,
114     value = V1) %>% dplyr::select(-c(ind))

```

```

104  return(list(dki.boot=dki.boot.data,dki.sample=dki.sample.data, rpki.
      sample=rpki.sample.data,rpki.boot=rpki.boot.data,          dki.
      boot.L2=dki.boot.data.L2,dki.sample.L2 = dki.sample.data.L2))

```

### A.3. Aplicación caso real de estudio curvas TG

En esta sección se muestra el código utilizado para obtener los resultados de la aplicación de la nueva metodología al caso real de estudio en curvas TG, descrito en el Capítulo 4. Dado que el proceso es iterativo se mostrará el código utilizado para la primera iteración.

```

1
2  library(parallel)
3  library(ILS)
4  library(fda.usc)
5  library(lubridate)
6  library(nortest)
7  library(tidyverse)
8
9  #Parametros
10 mc <- 1
11 L <- 7
12 m <- 0
13 nl <- length(L)
14 N <- c(15)
15 nn <- length(N)
16 mu_a <- mu_0 <- 0.5
17 nma <- length(mu_a)
18 alpha <- c(0.01)
19 na <- length(alpha)
20 B = 1000
21 sd<-5
22
23 # Paralelizacion
24 numCores <- detectCores()
25 cl <- makeCluster(numCores)
26 clusterExport(cl,c("MANDEL_FDA_BOOT","outliers.ils","fdata","TG","
      depth.RP","now","fdata.bootstrap","RPm","rename","group_by", "
      mutate", "spread", "select","apply"),envir =.GlobalEnv)
27 clusterEvalQ(cl,c(library(ILS),library(tidyverse), library(fda.usc),
      library(lubridate), library(nortest)))

```

```

28
29 #Resultados para el estadístico H
30 resultados_H <- firstdistboot.H(samples = TG, mc = mc,n = N[1], L= L,
    m = m, mean = mean, mean_a = mean_a, sd =sd, B = B, alpha = alpha
    [1], semilla = NULL, tt = 1000)
31 resultados_H_TG<-resultados_H
32
33 #Resultados para el estadístico K
34 resultados_K <- firstdistboot.K(samples = TG, mc = mc,n = N[1], L= L,
    m = m, mean = mean, mean_a = mean_a, sd =sd, B = B, alpha = alpha
    [1], semilla = NULL, tt = 1000)
35 resultados_K_TG<-resultados_K
36
37 stopCluster(cl)

```



# Bibliografía

- [1] John Mandel. A new analysis of interlaboratory test results. In *ASQC Quality Congress Transaction—Baltimore*, pages 360–366, 1985.
- [2] Miguel Flores, Génesis Moreno, Cristian Solórzano, Salvador Naya, and Javier Tarrío-Saavedra. Robust bootstrapped mandel’s h and k statistics for outlier detection in interlaboratory studies. *Chemometrics and Intelligent Laboratory Systems*, page 104429, 2021.
- [3] Miguel Flores, Javier Tarrío-Saavedra, Ruben Fernandez-Casal, and Salvador Naya. Functional extensions of mandel’s h and k statistics for outlier detection in interlaboratory studies. *Chemometrics and Intelligent Laboratory Systems*, 176:134–148, 2018.
- [4] James O Ramsay and CJ Dalzell. Some tools for functional data analysis. *Journal of the Royal Statistical Society: Series B (Methodological)*, 53(3):539–561, 1991.
- [5] Ricardo Fraiman and Graciela Muniz. Trimmed means for functional data. *Test*, 10(2):419–440, 2001.
- [6] Antonio Cuevas, Manuel Febrero, and Ricardo Fraiman. On the use of the bootstrap for estimating functions with functional data. *Computational statistics & data analysis*, 51(2):1063–1074, 2006.
- [7] Antonio Cuevas, Manuel Febrero, and Ricardo Fraiman. Robust estimation and classification for functional data via projection-based depth notions. *Computational Statistics*, 22(3):481–496, 2007.
- [8] Bossano Roberto. *Aplicación de una nueva metodología basada en el método bootstrap para la detección de resultados atípicos en estudios interlaboratorio: un caso práctico en laboratorios clínicos*. 2019.

- [9] Juan Antonio Cuesta-Albertos, Ricardo Fraiman, and Thomas Ransford. A sharp form of the cramer-wold theorem. *Journal of Theoretical Probability*, 20(2):201–209, 2007.
- [10] Miguel Flores Sánchez. Nuevas aportaciones del análisis de datos funcionales en el control estadístico de procesos. 2018.
- [11] EA Maier and Ph Quevauviller. *Interlaboratory studies and certified reference materials for environmental analysis: the BCR approach*. Elsevier, 1999.
- [12] Yvan Vander Heyden and Johanna Smeyers-Verbeke. Set-up and evaluation of interlaboratory studies. *Journal of Chromatography A*, 1158(1-2):158–167, 2007.
- [13] Organització Internacional per a la Normalització. *Accuracy (trueness and Precision) of Measurement Methods and Results*. International Organization for Standardization, 1994.
- [14] Gerald B Heyes. Astm e691–87 standard practice for conducting an interlaboratory study to determine the precision of a test method, 1993.
- [15] Edelgard Hund, D Luc Massart, and Johanna Smeyers-Verbeke. Interlaboratory studies in analytical chemistry. *Analytica Chimica Acta*, 423(2):145–165, 2000.
- [16] Patrick C Kelly. Outlier detection in collaborative studies. *Journal of the Association of Official Analytical Chemists*, 73(1):58–64, 1990.
- [17] Jane C Miller and James N Miller. Basic statistical methods for analytical chemistry. part i. statistics of repeated measurements. a review. *Analyst*, 113(9):1351–1356, 1988.
- [18] WJ Youden. Industrial quality control. Vol XV, (11), 1959.
- [19] Peter-T Wilrich. Critical values of mandel’sh and k, the grubbs and the cochrane test statistic. *ASTA Advances in Statistical Analysis*, 97(1):1–10, 2013.
- [20] Frank E Grubbs. Sample criteria for testing outlying observations. *The Annals of Mathematical Statistics*, pages 27–58, 1950.
- [21] William G Cochran. The distribution of the largest of a set of estimated variances as a fraction of their total. *Annals of Eugenics*, 11(1):47–52, 1941.

- [22] James O Ramsay and Bernard W Silverman. *Applied functional data analysis: methods and case studies*, volume 77. Springer, 2002.
- [23] John W Tukey. Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians, Vancouver, 1975*, volume 2, pages 523–531, 1975.
- [24] Regina Y Liu. On a notion of data depth based on random simplices. *The Annals of Statistics*, pages 405–414, 1990.
- [25] Yijun Zuo and Robert Serfling. General notions of statistical depth function. *Annals of statistics*, pages 461–482, 2000.
- [26] Bárcenas Roberto. *Estudio de perfiles de olas vía Análisis de Datos Funcionales*. 2014.
- [27] Bradley Efron. Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics*, pages 569–593. Springer, 1992.
- [28] Miguel Flores, Rubén Fernández-Casal, Salvador Naya, Javier Tarrío-Saavedra, and Roberto Bossano. IIs: An r package for statistical analysis in interlaboratory studies. *Chemometrics and Intelligent Laboratory Systems*, 181:11–20, 2018.
- [29] Manuel Febrero Bande, Manuel Oviedo de la Fuente, Pedro Galeano, Alicia Nieto, Eduardo Garcia-Portugues, and Maintainer Manuel Oviedo de la Fuente. Package ‘fda. usc’. *CRAN Repository*, 2020.
- [30] Manuel Febrero-Bande and Manuel Oviedo de la Fuente. Statistical computing in functional data analysis: The r package fda. usc. *Journal of statistical Software*, 51(1):1–28, 2012.
- [31] Juan Antonio Cuesta-Albertos, Ricardo Fraiman, Antonio Galves, Jesús Garcia, and Marcela Svarc. Identifying rhythmic classes of languages using their sonority: a kolmogorov-smirnov approach. *Journal of Applied Statistics*, 34(6):749–761, 2007.
- [32] Juan Antonio Cuesta-Albertos, Eustasio del Barrio, Ricardo Fraiman, and Carlos Matrán. The random projection method in goodness of fit for functional data. *Computational Statistics & Data Analysis*, 51(10):4814–4831, 2007.
- [33] Antonio Cuevas, Manuel Febrero, and Ricardo Fraiman. An anova test for functional data. *Computational statistics & data analysis*, 47(1):111–122, 2004.

- [34] JA Cuesta-Albertos and Manuel Febrero-Bande. A simple multiway anova for functional data. *Test*, 19(3):537–557, 2010.
- [35] Salvador Naya, Antonio Meneses, Javier Tarrío-Saavedra, Ramón Artiaga, Jorge López-Beceiro, and Carlos Gracia-Fernández. New method for estimating shift factors in time–temperature superposition models. *Journal of thermal analysis and calorimetry*, 113(2):453–460, 2013.
- [36] Inés Barbeito, Sonia Zaragoza, Javier Tarrío-Saavedra, and Salvador Naya. Assessing thermal comfort and energy efficiency in buildings by statistical quality control for autocorrelated data. *Applied energy*, 190:1–17, 2017.
- [37] William Navidi. *Estadística para ingenieros científicos*. Number 519.5 N325. McGraw Hill Interamericana, 2006.
- [38] Larry Barrentine. *Concepts for RR Studies*. ASQ Quality Press, 2003.
- [39] Javier Tarrío-Saavedra, Jorge López-Beceiro, Salvador Naya, Mario Francisco-Fernández, and Ramón Artiaga. Simulation study for generalized logistic function in thermal data modeling. *Journal of Thermal Analysis and Calorimetry*, 118(2):1253–1268, 2014.
- [40] F Barbadillo, A Fuentes, S Naya, R Cao, J Mier, and R Artiaga. Evaluating the logistic mixture model on real and simulated tg curves. *Journal of thermal analysis and calorimetry*, 87(1):223–227, 2007.
- [41] R Kenett and G Shmueli. Clarifying the terminology that describes scientific reproducibility. *Nature Methods*, 12:699, 2015.
- [42] R Cao. Aplicaciones y nuevos resultados del método bootstrap en la estimación no paramétrica de curvas. 1990.