

NUEVAS APORTACIONES DEL ANÁLISIS
DE DATOS FUNCIONALES EN EL
CONTROL ESTADÍSTICO DE PROCESOS

Miguel Flores Sánchez

Tese de doutoramento UDC/2018

Departamento de Matemáticas



NUEVAS APORTACIONES DEL ANÁLISIS DE DATOS FUNCIONALES EN EL CONTROL ESTADÍSTICO DE PROCESOS

Miguel Flores Sánchez

Tese de doutoramento UDC/2018

Directores:

Salvador Naya Fernandez

Javier Tarrío Saavedra

Departamento de Matemáticas



UNIVERSIDADE DA CORUÑA



Los abajo firmantes hacen constar que son los directores de la Tesis Doctoral titulada “Nuevas Aportaciones del Análisis de Datos Funcionales en el Control Estadístico de Procesos”, desarrollada por Miguel Flores Sánchez en el ámbito del programa de doctorado de Estadística e Investigación Operativa ofertado por el Departamento de Matemáticas de la Universidade da Coruña, dando su consentimiento para que su autor, cuya firma también se incluye, proceda a su presentación y posterior defensa.

A Coruña, 11 de Diciembre de 2018.

Directores:

Salvador Naya Fernandez

Javier Tarrío Saavedra

Doctorando:

Miguel Flores Sánchez

A la memoria de mi mamá, Nelly Sánchez Correa.

Agradecimientos

En primer lugar desearía expresar mi gratitud a los directores de esta tesis, el profesor Salvador Naya y el profesor Javier Tarrío Saavedra por su apoyo, confianza y dedicación sin la cual este trabajo no habría salido adelante. Ha sido todo un privilegio poder contar con su guía y ayuda, pero sobre todo, con su amistad. Quiero extender el agradecimiento también al profesor Luis Escobar quien me recomendó estudiar en este programa de Doctorado y me presento al profesor Salvador Naya, además al profesor Rubén Fernández Casal por ser un amigo y mentor en el campo de la estadística computacional.

También deseo agradecer al tribunal encargado de juzgar el Seminario de Tesis por su exhaustiva revisión, sin sus acertados comentarios no se podría contar con la versión actual de esta Tesis.

Durante mis estancias en A Coruña hice muchos amigos de varios países, pero también tuve la fortuna de contar con la amistad de mi paisano Sergio Castillo, el cual me brindó su apoyo moral y anímico. A todos mis amigos, gracias por compartir momentos de viajes, fiestas y comidas. Siempre los tengo presentes en mi corazón.

De manera especial quiero agradecer a la familia que tengo en A Coruña, siempre los recuerdo con mucho cariño. Gracias querido Romel Capelo por brindarme tú amistad y permitirme ser un miembro más de tú familia.

A partir de muchas conversaciones amenas con los profesores Javier Tarrío Saavedra y Rubén Fernández Casal sobre temas de investigación relacionados a mi tesis y sus líneas de investigación, se generaron varios trabajos de titulación, los cuales dirigí para obtener grados y postgrados en Ecuador. Quisiera agradecerles, a nombre de mis dirigidos, por compartir sus conocimientos.

En la última etapa del desarrollo de mi tesis conté con el apoyo incondicional

de muchos amigos que me permitió obtener resultados de los estudios de simulación en pocos días, los cuales sin su ayuda hubieran durado algunos meses. Gracias por ejecutar los diferentes escenarios de simulación en sus computadoras personales. Gracias a los amigos de República Dominicana y a los amigos de Ecuador.

Sin el apoyo y cariño de toda mi familia, especialmente de mis hermanas y sobrinos este trabajo no hubiera sido posible. Gracias por estar siempre en los momentos donde más se necesita el amor de una familia.

Finalmente, quiero hacer constar mi agradecimiento a la Escuela Politécnica Nacional por haberme concedido las facilidades para poder culminar con éxito mi tesis doctoral.

Prólogo

Esta memoria propone nuevas metodologías en el control estadístico de la calidad, en especial técnicas adaptadas a datos obtenidos mediante sensores y que se engloban dentro del análisis de datos funcionales.

Todas las metodologías, técnicas y programas propuestos han sido desarrolladas para aportar posibles soluciones a problemas reales en la industria y en los laboratorios de análisis. Los cuales tienen su punto de partida en la necesidad de resolución de problemas de detección de anomalías relacionadas con el consumo eléctrico, confort térmico y calidad del aire en edificios comerciales y públicos.

En particular, las empresas NERXUS, asesoría estadística en eficiencia energética, FRIDAMA, de instalaciones y control de clima, y Σ qus, de plataformas web y gestión de Big Data en edificios, contactaron con el autor y directores de la presente tesis para mostrar sus problemas y proporcionar datos y casos prácticos que ilustran su utilidad. Por otro lado, el germen del desarrollo de las nuevas aproximaciones FDA (Functional Data Analysis) a los estadísticos h y k de Mandel es la participación de los directores Salvador Naya y Javier Tarrío en diversos estudios ILS, que permitió el contacto con miembros de la American Society for Testing and Materials o ASTM International, y que a su vez transmitieron su interés por el desarrollo de nuevas metodologías de detección de laboratorios atípicos desde el contexto FDA, dada la naturaleza funcional de gran parte de los datos experimentales obtenidos actualmente mediante la nueva instrumentación (espectros, termogramas, etc.). Así, uno de los objetivos finales del desarrollo de estas aproximaciones FDA es su posible inclusión dentro de los estándares de la ASTM.

Para aportar soluciones a los problemas planteados, se propone el desarrollo, programación y aplicación de aproximaciones, desde la perspectiva del análisis de

datos funcionales (FDA), de herramientas básicas del control estadístico de la calidad (CEC) como son, por un lado, los gráficos de control, técnica más importante de la etapa Controlar dentro de la metodología Seis Sigma de mejora de procesos y, por otro lado, los estadísticos h y k de Mandel, desarrollados para la detección de laboratorios atípicos.

En el Capítulo 1, se introducen los conceptos principales del CEC, rama de la estadística que es la temática principal de esta memoria de tesis, haciendo mención especial a los estudios R&R, los estudios interlaboratorio y los gráficos de control. También se incluye una pequeña introducción a la meteorología 4.0 y a los retos de la industria debido a los avances en la sensorización y adquisición de datos.

Dada la naturaleza funcional de los datos que definen los problemas a resolver en esta tesis, en el Capítulo 2 se proporciona también una descripción somera del análisis de datos funcionales, ámbito en el que se desarrollan las metodologías propuestas. En el Capítulo 2, se introducen aquellos casos de estudio reales cuya resolución se busca mediante el desarrollo de las metodologías propuestas en esta tesis, además de describir las bases de datos.

En el Capítulo 3, de la presente tesis se ha diseñado un procedimiento para la detección de laboratorios que proporcionan resultados inconsistentes dentro de un estudio interlaboratorio. La principal aportación de dicho procedimiento es que permite la identificación de laboratorios que presentan resultados atípicos con respecto a los demás a partir de las curvas experimentales obtenidas originalmente, ya sean espectros o termogramas, entre otros tipos, previamente a la extracción de características representativas o aplicación de otro tipo de pre-procesado de datos.

En el Capítulo 4, se propone un gráfico de control para datos funcionales basado en la detección de atípicos mediante el cálculo de profundidad de datos funcionales, el remuestreo bootstrap y la adaptación de los gráficos de control no paramétricos de rangos. En particular, se propone una metodología de control que proporciona herramientas para llevar a cabo las llamadas Fases I y II de la construcción de un gráfico de control.

La metodología propuesta es aplicada al caso real de estudio, que es la detección de anomalías en la eficiencia energética que se aplicó al estudio de distintos locales de una conocida empresa textil gallega radicadas en Ciudad de Panamá, siendo

el consumo energético la variable crítica para la calidad (CTQ) del sistema HVAC (Heating, Ventilating and Air Conditioning). En resumen, la metodología propuesta, que combina técnicas FDA y multivariantes, validadas en diversos trabajos, adapta el concepto de gráfico de control para el caso particular de datos funcionales, representando una alternativa novedosa para el control de instalaciones en las que los datos son tomados en continuo (medidas horarias, pentaminutales, minutales, etc.) por sensores.

En el Capítulo 5, se mostrarán las principales conclusiones y las líneas futuras de investigación en el campo de los estudios interlaboratorio y los gráficos de control para datos funcionales.

Una parte importante de esta memoria de tesis es el desarrollo de aplicaciones informáticas que proporcionen a los profesionales de la industria y servicios, comunidad científica y usuarios en general una forma sencilla y eficaz de aplicar las metodologías propuestas en este trabajo. Para ello, todas las rutinas programadas han sido incluidas en dos nuevas librerías: por un lado la librería *Interlaboratory Study* o ILS y, por el otro, la librería *qcr* (*Quality Control Review*).

Para describir su contenido y uso se han agregado a esta memoria dos apéndices, el Apéndice A describe la librería ILS que facilita el desarrollo de estudios interlaboratorio cuando los datos tienen un carácter funcional, permitiendo también la detección de laboratorios atípicos cuando los datos son escalares (incorpora las herramientas clásicas en este tipo de estudios), mientras que en el Apéndice B se describe la funciones que se encuentran desarrolladas en el paquete *qcr*, pone a disposición de los usuarios las herramientas para construir gráficos de control para datos univariante, multivariantes y funcionales.

Objetivos

El objetivo de los trabajos de investigación que componen la presente tesis doctoral es, fundamentalmente, la propuesta de nuevas herramientas estadísticas, además de la modificación de técnicas existentes, para el control de la calidad de procesos en la industria, por un lado, y en los laboratorios, por otro. En particular, se han desarrollado, modificado y aplicado metodologías estadísticas alternativas para la detección de anomalías y datos atípicos específicamente adaptadas para los nuevos tipos de datos que el actual desarrollo de las ciencias de la computación y sensorización proporciona, como son los datos de alta dimensión. Por tanto, el carácter y la finalidad de esta tesis doctoral es fundamentalmente práctica, buscando mejorar el desempeño de técnicas estadísticas en aplicaciones relacionadas con la industria en la era digital o Industria 4.0. Dentro de este mismo objetivo práctico, mediante los contenidos que a continuación se muestran en esta tesis, se pretende hacer fácilmente accesibles a los usuarios todas las herramientas desarrolladas en el presente proyecto. Con tal fin, se ha hecho especial hincapié en programar herramientas computacionales mediante el software estadístico R, de fácil acceso y aplicación para los usuarios de la industria y academia. En particular, se han desarrollado dos librerías en R: la librería quality control review (**qcr**), enfocada en el desarrollo de gráficos de control y análisis de capacidad de procesos, y el paquete interlaboratory study (**ILS**), diseñado para la detección de laboratorios que proporcionan datos no consistentes (atípicos), en el ámbito de los estudios interlaboratorio, también llamados ILS.

Objetivos Especificos

Teniendo en cuenta la motivación principal de este trabajo, se puede definir la siguiente serie de objetivos específicos:

- Propuesta de nuevas metodologías de detección de atípicos consistentes en la aproximación desde el Análisis de Datos Funcionales (FDA) de los estadísticos h y k de Mandel para la detección de laboratorios que proporcionan datos no consistentes dentro de un estudio interlaboratorio. Mediante estas aproximaciones FDA se pretende dar respuesta al problema de detección de laboratorios atípicos cuando los resultados experimentales que proporcionan son datos funcionales (curvas). Actualmente se emplean procedimientos de detección de atípicos univariantes que obligan a resumir la información de la curva en una única variable, como son los estadísticos h y k univariantes. Las aproximaciones FDA propuestas pretenden proporcionar una solución, desde la perspectiva del control estadístico de calidad (CEC), a uno de los nuevos problemas que plantea la Metrología 4.0 (en lo que se refiere a la complejidad de los datos).
- Desarrollo de gráficos de control que permitan controlar la variabilidad de procesos y servicios, y así detectar anomalías, a partir de datos funcionales. De hecho, se busca construir gráficos de control para datos funcionales basados en la detección de atípicos a partir del cálculo de la profundidad de datos funcionales propuesta por [Febrero-Bande *et al.* \(2007\)](#) y [Hyndman y Shang \(2010\)](#), por un lado, y en el gráfico de control de rangos o r propuesto por Regina Liu para datos multivariantes ([Liu, 1995](#)). La estimación de los límites de calibración (Fase I de los gráficos de control) se llevaría a cabo gracias a la aplicación de métodos de detección de atípicos basados en el cálculo de profundidad de datos funcionales y el remuestreo bootstrap, mientras que la Fase II o monitorización del proceso se realizaría mediante la aplicación de los gráficos de control no paramétricos de rangos, extendidos al caso funcional.
- Programar herramientas computacionales que permitan una aplicación sencilla y eficaz de nuevas técnicas estadísticas desarrolladas en los contextos FDA, multivariante y univariante, con el objeto de facilitar su uso tanto en el ámbito de los Estudios Interlaboratorio como en el de control y análisis de la capacidad de procesos en la industria.

Abstract

This thesis report proposes new methodologies in the field of statistical quality control (SQC), specifically those techniques adapted to data obtained by sensors, continuously monitored with respect to time, and which are included in the framework of the branch of statistics called Functional Data Analysis (FDA).

All the methodologies, techniques and programs proposed have been developed to provide possible solutions to real problems in the industry and in the laboratories of experimental analysis. All these problems have their starting point in the need to find solutions for the detection of anomalies related to electrical consumption, thermal comfort and air quality in commercial and public buildings, on the one hand, and, on the other hand, the detection of atypical sensors or laboratories in the framework of interlaboratory studies.

In particular, NERXUS company, statistical consultancy in energy efficiency, FRIDAMA, of facilities related to climate control, and Σ qus, developer of web platforms and management of Big Data in buildings, contacted the author and directors of this PhD thesis to show their problems through practical case studies related to the detection of anomalies in fully sensorized buildings. Given the functional nature of the data, a new methodology of FDA control charts based on the concept of data depth was proposed. On the other hand, the germ of the development of new FDA approaches to Mandel's h and k statistics for the detection of atypical laboratories within an interlaboratory study (ILS) is the participation of the directors Salvador Naya and Javier Tarrío in several ILS studies, which allowed to contact with members of the American Society for Testing and Materials or ASTM International (organization of international standards that develops and publishes voluntary agreements of technical standards for materials, products, systems or services). In

fact, those professionals in charge of interlaboratory studies organized by the ASTM conveyed their interest in the development of new methodologies for detecting atypical laboratories from the FDA context, given the functional nature of a large part of the experimental data currently obtained through the new available instrumentation (spectra, thermograms, etc.). As a result of this contact, it was established that one of the final objectives of this thesis was the development of these FDA approaches for possible inclusion within the ASTM standards.

In order to provide solutions to the problems raised, the author has proposed the development, programming and application of approximations, from the FDA perspective, of basic tools of SQC such as, on the one hand, the control charts, the most important and representative tool of the "Control" stage of Six Sigma methodology of process improvement, and, on the other hand, the Mandel's h and k statistics, developed for the detection of atypical laboratories, taking into account that their graphical output is the most used technique in this field for the detection of laboratories whose data are anomalous.

Specifically, in Chapter 1, the main concepts of the SQC (the branch of statistics in which this thesis is circumscribed), with special mention to reproducibility and repeatability (R & R) studies, interlaboratory studies and control charts, are introduced. It also includes a small introduction to metrology 4.0 and some of the new challenges of industrial statistics, in terms of process control and monitoring, which recent advances in sensorization and data acquisition propose.

Given the functional nature of the data defining the problems to be solved in this thesis, the Chapter 2 provides a brief description of functional data analysis, an area in which the proposed methodologies are developed. Also in the Chapter 2, we introduce those real case studies whose resolution is sought through the development and application of the methodologies proposed in this thesis, in addition to the description of the databases.

In the Chapter 3 of this thesis, a procedure has been designed for the detection of laboratories that provide inconsistent results (with respect to the other laboratories) within an interlaboratory study. The main contribution of this procedure is the identification of laboratories that obtain atypical results, with respect to those obtained by others, from the original experimental curves, whether spectra or thermograms,

among other types, and preventing to apply extraction processes of representative features or the application of another type of data pre-processing.

In the Chapter 4, a control chart is proposed for functional data based on the detection of atypicals by calculating the depth of functional data, the bootstrap resampling and the adaptation of nonparametric range control charts. In particular, a control methodology that provides tools to perform the so-called Phases I and II of the construction of a control chart is proposed. The proposed methodology is applied to the real study case: the detection of anomalies in energy efficiency in stores of a well-known Galician textile company based in Panama City, with energy consumption being the critical to quality variable (CTQ) of the heat, ventilation, and air conditioning system (HVAC). In summary, the proposed methodology, which combines FDA and multivariate techniques (validated in different works) adapts the concept of control chart for the particular case of functional data. It represents a novel alternative for the control of those processes defined by data continuously monitored over time (hourly, pentaminutal, minutal, etc.) through sensors.

In Chapter 5, the main conclusions and future lines of research are shown, both in the field of interlaboratory studies and in the field of control charts for functional data.

It is also important to note that an important part of this thesis is the development of computer applications that provide professionals in the industry and services, scientific community and users in general a simple and effective way to apply the methodologies proposed in this work. For this, all the programmed routines have been included in two new libraries: on the one hand, the Interlaboratory Study library or ILS and, on the other hand, the `qcr` (Quality Control Review) library.

To describe its content and use, two appendices have been added to this report. The Appendix A describes the ILS library that facilitates the development of interlaboratory studies when the data are functional, also providing the detection of atypical laboratories when the data are scalar (ILS incorporates the classic univariate tools in this type of studies). The Appendix B focuses on the description of the functions that are developed in the `qcr` package (that it is still in the development and expansion phase) making available to users those tools to build control charts for univariate, multivariate and functional data, in addition to include techniques for

process capacity analysis.

Resumo

Nesta memoria de tese propóñense novas metodoloxías no eido do control estatístico da calidade (CEC), especificamente, apórtanse novas técnicas adaptadas a datos obtidos mediante o uso de sensores, monitorizados continuamente con respecto do tempo, e que se atinxen dentro da rama da estatística denominada análise de datos funcionais (FDA).

Todas as metodoloxías, técnicas e programas propostos foron deseñados para proporcionar solucións alternativas a problemas reais na industria e nos laboratorios de análise. Todos estes problemas teñen o seu cerne na necesidade de atopar solucións para a detección de anomalías relacionadas co consumo eléctrico, o confort térmico ea calidade do aire nos edificios comerciais e públicos, por un lado, e a detección de sensores ou laboratorios atípicos no marco dos estudos interlaboratorio, por outro.

En particular, as empresas NERXUS, consultoría estatística sobre a eficiencia enerxética, FRIDAMA, adicada ás instalacións de climatización e Σ qus, que desenvolve plataformas web e xestión de Big Data en edificios, contactaron co autor e directores desta memoria de tese para mostrar os seus problemas e proporcionar datos a través de casos de estudo prácticos relacionados coa detección de anomalías en edificacións totalmente sensorizadas. Dada a natureza funcional dos datos, propúxose unha nova metodoloxía de gráficos de control FDA baseada no concepto de profundidade de datos. Por outra banda, o xerme do desenvolvemento dos novos enfoques FDA para os coñecidos estatísticos h e k de Mandel (ata agora só definidos para o caso escalar) para a detección de laboratorios atípicos dentro dos estudos interlaboratorio é a participación dos directores Salvador Naya e Javier Tarrío en varios estudos ILS, o que permitiu o contacto con membros da merican Society for Testing and Materials ou ASTM (organización de estándares internacionais que desenvolve e

publica acordos voluntarios de normas técnicas tanto para materiais, produtos, sistemas ou servizos). De feito, os responsábeis a cargo dos estudos ILS organizados pola ASTM transmitiron o seu interese no desenvolvemento de novas metodoloxías para a detección de laboratorios atípicos no contexto FDA, dada a natureza funcional de gran parte dos datos experimentais obtidos actualmente grazas a nova instrumentación dispoñible (espectros, termogramas, etc.). Como consecuencia inmediata deste contacto, fixouse coma un dos obxectivos finais da presente tese de doutoramento o desenvolvemento destes enfoques FDA para a súa posible inclusión dentro das normas ASTM.

Co galo de proporcionar solucións aos problemas propostos, proponse o desenvolvemento, programación e posta en marcha de aproximacións, desde a perspectiva da análise de datos funcionais (FDA), de ferramentas básicas do control estatístico da calidade como son, por unha banda, os gráficos de control, con certeza a técnica estatística máis importante da etapa “Controlar”, no marco da metodoloxía Seis Sigma e, por outra banda, os estatísticos h e k de Mandel, desenvolvidos para a detección de laboratorios atípicos, sendo a súa representación gráfica a técnicas máis utilizada no ámbito dos ILS para a detección de laboratorios cuxos datos son anómalos.

Especificamente, no Capítulo 1, introdúcense os principais conceptos da CEC, a rama da estatística na que se circunscribe a presente memoria de tese, facendo especial mención a estudos de reproducibilidade e repetibilidade (&R), os estudos ILS e os gráficos de control. Tamén se inclúe unha breve introdución á metroloxía 4.0 e aos novos retos da industria industrial, no que se refire ao control e monitorización de procesos, que os recentes avances na sensorización e adquisición de datos propoñen.

Dada a natureza funcional dos datos que definen os problemas a resolver nesta tese, no Capítulo 2 se proporciona unha breve descrición da análise de datos funcionais, unha área na que se desenvolven as metodoloxías propostas. Tamén no Capítulo 2, se introducen os casos reais de estudo cuxa resolución se busca mediante o desenvolvemento e aplicación das metodoloxías propostas nesta tese, ademais de describir as bases de datos.

No capítulo 3 da presente tese deseñouse un procedemento para a detección de laboratorios que proporcionan resultados inconsistentes (con respecto aos demais laboratorios) dentro dun estudo interlaboratorio. A principal contribución deste pro-

cedemento é a identificación de laboratorios que obteñen resultados atípicos, con respecto ao obtidos polos demais laboratorios, a partir das curvas experimentais orixinais, sexan espectros ou termogramas, entre outros tipos, evitando os procesos de extracción de características representativas ou a aplicación de outro tipo de pre-procesamento de datos.

No capítulo 4 propónse un gráfico de control para datos funcionais para a detección de anomalías baseado no cálculo da profundidade de datos funcionais, a remostraxe bootstrap e adaptación dos gráficos de control non paramétricos de rangos. En particular, propónse unha metodoloxía de control que proporciona ferramentas para levar a cabo as chamadas Fases I e II da construción dun gráfico de control. A metodoloxía proposta aplícase ao estudo de casos reais, como son a detección de anomalías en eficiencia enerxética en varias tendas dunha coñecida empresa téxtil galega baseada na Cidade de Panamá, onde o consumo de enerxía é a variable crítica para a calidade (CTQ) do sistema de climatización (heat, ventilation, and air conditioning, HVAC). En resumo, a metodoloxía proposta, que combina técnicas FDA e multivariantes, validadas en diversos estudos, adapta o concepto de gráfico de control para o caso particular dos datos funcionais, o que supón unha nova alternativa para controlar instalacións onde se toman datos mediante monitorización continua (horarias, pentaminutales, minutais, etc.) a través de sensores.

No capítulo 5 mostráanse as principais conclusións e futuras liñas de investigación, tanto no eido dos ILS coma no dos gráficos de control para datos funcionais.

É importante destacar que unha parte significativa desta tese de doutoramento consiste no desenvolvemento de programas que fornezan aos profesionais da industria e servizos, á comunidade científica e a usuarios en xeral de ferramentas para aplicar as metodoloxías propostas dun xeito sinxelo e eficaz. Para iso, as rutinas programadas foron incluídas en dúas novas librerías do software estatístico R: por un lado a librería Interlaboratory Study ou ILS e, polo outro, a librería qcr (Quality Control Review).

Para describir o seu contido e uso, engadíronse a esta memoria dous apéndices, o Apéndice A, que describe a librería ILS, sendo esta a que facilita o desenvolvemento dos estudos interlaboratorio cando os datos teñen un carácter funcional, permitindo tamén a detección de laboratorios atípicos cando os datos son escalares (incorpora as ferramentas clásicas neses estudos). O Apéndice B describe as funcións imple-

mentadas no paquete `qcr` , que aínda está en a desenvolverse e ampliarse, poñendo a disposición dos usuarios as ferramentas precisas para elaborar gráficos de control datos univariantes, multivariantes e funcionais, ademais de incluír técnicas para a análise da capacidade de procesos.

Resumen

En esta memoria de tesis se proponen nuevas metodologías en el ámbito del control estadístico de la calidad (CEC), específicamente se aportan técnicas adaptadas a datos obtenidos mediante sensores, monitorizados continuamente respecto al tiempo, y que se engloban dentro del análisis de datos funcionales (FDA).

Todas las metodologías, técnicas y programas propuestos han sido desarrollados para aportar posibles soluciones a problemas reales en la industria y en los laboratorios de análisis. Todos estos problemas tienen su punto de partida en la necesidad de encontrar soluciones para la detección de anomalías relacionadas con el consumo eléctrico, confort térmico y calidad del aire en edificios comerciales y públicos, por un lado, y la detección de sensores o laboratorios atípicos en el marco de los estudios interlaboratorio (ILS), por otro.

En particular, las empresas NERXUS, asesoría estadística en eficiencia energética, FRIDAMA, de instalaciones y control de clima, y Σ qus, de plataformas web y gestión de Big Data en edificios, contactaron con el autor y directores de la presente tesis para mostrar sus problemas a través de casos de estudio prácticos relacionados con la detección de anomalías en edificaciones totalmente sensorizadas. Dada la naturaleza funcional de los datos, se propuso una nueva metodología de gráficos de control FDA basada en el concepto de profundidad de datos. Por otro lado, el germen del desarrollo de nuevas aproximaciones FDA a los estadísticos h y k de Mandel para la detección de laboratorios atípicos dentro de un estudio interlaboratorio es la participación de los directores Salvador Naya y Javier Tarrío en diversos estudios ILS, que permitió el contacto con miembros de la American Society for Testing and Materials o ASTM International (organización de normas internacionales que desarrolla y publica acuerdos voluntarios de normas técnicas ya sea para materiales,

productos, sistemas o servicios). De hecho, los responsables a cargo de estudios interlaboratorio organizados por la ASTM nos transmitieron su interés por el desarrollo de nuevas metodologías de detección de laboratorios atípicos desde el contexto FDA, dada la naturaleza funcional de gran parte de los datos experimentales obtenidos actualmente mediante la nueva instrumentación disponible (espectros, termogramas, etc.). Como consecuencia de este contacto, se estableció que uno de los objetivos finales de la presente tesis fuera el desarrollo de estas aproximaciones FDA para su posible inclusión dentro de los estándares de la ASTM.

Para aportar soluciones a los problemas planteados, se propone el desarrollo, programación y aplicación de aproximaciones, desde la perspectiva FDA, de herramientas básicas del control estadístico de la calidad como son, por un lado, los gráficos de control, la técnica más importante y representativa de la etapa “Controlar”, dentro de la metodología Seis Sigma de mejora de procesos, y, por otro lado, los estadísticos h y k de Mandel, desarrollados para la detección de laboratorios atípicos, siendo su representación gráfica la técnica más usada en este ámbito para la detección de laboratorios cuyos datos son anómalos.

Específicamente, en el Capítulo 1, se introducen los conceptos principales del CEC, rama de la estadística en la que se circunscribe la presente memoria de tesis, haciendo mención especial a los estudios de reproducibilidad y repetibilidad (R&R), los estudios interlaboratorio y los gráficos de control. También se incluye una pequeña introducción a la metrología 4.0 y a algunos de los nuevos retos de la estadística industrial, en lo que se refiere a control y monitorización de procesos, que los recientes avances en la sensorización y adquisición de datos proponen.

Dada la naturaleza funcional de los datos que definen los problemas a resolver en esta tesis, en el Capítulo 2 se proporciona una descripción somera del análisis de datos funcionales, ámbito en el que se desarrollan las metodologías propuestas. También en el Capítulo 2, se introducen aquellos casos de estudio reales cuya resolución se busca mediante el desarrollo y aplicación de las metodologías propuestas en esta tesis, además de describir las bases de datos.

En el Capítulo 3 de la presente tesis se ha diseñado un procedimiento para la detección de laboratorios que proporcionan resultados inconsistentes (con respecto a los demás laboratorios) dentro de un estudio interlaboratorio. La principal aportación

de dicho procedimiento es la identificación de laboratorios que obtienen resultados atípicos, con respecto a los obtenidos por los demás, a partir de las curvas experimentales originales, ya sean espectros o termogramas, entre otros tipos, evitando tener que aplicar procesos de extracción de características representativas o la aplicación de otro tipo de pre-procesado de datos.

En el Capítulo 4, se propone un gráfico de control para datos funcionales basado en la detección de estados anómalos mediante el cálculo de la profundidad de datos funcionales, el remuestreo bootstrap y la adaptación de los gráficos de control no paramétricos de rangos. En particular, se propone una metodología de control que proporciona herramientas para llevar a cabo las llamadas Fases I y II de la construcción de un gráfico de control. La metodología propuesta es aplicada al caso real de estudio, que es la detección de anomalías en la eficiencia energética en diversas tiendas de una conocida empresa textil gallega radicadas en Ciudad de Panamá, siendo el consumo energético la variable crítica para la calidad (CTQ) del sistema de climatización (heat, ventilation, and air conditioning, HVAC). En resumen, la metodología propuesta, que combina técnicas FDA y multivariantes, validadas en diversos trabajos, adapta el concepto de gráfico de control para el caso particular de datos funcionales, representando una alternativa novedosa para el control de instalaciones en las que los datos son tomados en continuo (medidas horarias, pentaminutales, minutales, etc.) por sensores.

En el Capítulo 5, se muestran las principales conclusiones y las líneas futuras de investigación, tanto en el ámbito de los estudios interlaboratorio como en el campo de los gráficos de control para datos funcionales.

Es también importante destacar que una parte significativa de esta memoria de tesis consiste en el desarrollo de aplicaciones informáticas que proporcionen a los profesionales de la industria y servicios, comunidad científica y usuarios en general una forma sencilla y eficaz de aplicar las metodologías propuestas en este trabajo. Para ello, todas las rutinas programadas han sido incluidas en dos nuevas librerías: por un lado la librería Interlaboratory Study o ILS y, por el otro, la librería qcr (Quality Control Review).

Para describir su contenido y uso se han añadido a esta memoria dos apéndices. El Apéndice A describe la librería ILS que facilita el desarrollo de estudios inter-

laboratorio cuando los datos tienen un carácter funcional, permitiendo también la detección de laboratorios atípicos cuando los datos son escalares (incorpora las herramientas clásicas en este tipo de estudios). El Apéndice B se centra en la descripción de las funciones que se encuentran desarrolladas en el paquete `qcr`, aún en fase de desarrollo y ampliación, poniendo a disposición de los usuarios herramientas para construir gráficos de control para datos univariantes, multivariantes y funcionales, además de incluir técnicas para el análisis de capacidad de procesos.

Índice general

Agradecimientos	VIII
Resumen	XXVIII
Lista de Figuras	XXXVI
Lista de Tablas	XLI
1. Introducción	1
1.1. Industria 4.0 y Metrología 4.0	3
1.1.1. Industria 4.0	4
1.1.2. Metrología 4.0	6
1.2. Control estadístico de la calidad y metodología Seis Sigma	7
1.2.1. Proceso bajo control	7
1.2.2. Metodología Seis sigma	10
1.3. Estudios interlaboratorio	14
1.3.1. Pruebas de consistencia	16
1.4. Gráficos de control	20
1.5. Retos actuales del Control Estadístico de la Calidad	25
1.5.1. Escenarios de monitorización	25
1.5.2. Monitorizado de bases de datos de alta dimensión	26
2. Técnicas estadísticas FDA y casos de estudio reales	30
2.1. Análisis de datos funcionales	30
2.1.1. Definiciones	31

2.1.2.	Análisis exploratorio	32
2.1.3.	Bandas de confianza Bootstrap	37
2.1.4.	Método de detección de atípicos para datos funcionales	38
2.1.5.	ANOVA Funcional	39
2.2.	Casos de estudio	40
2.2.1.	Estudio interlaboratorio a partir de datos termogravimétricos y calorimétricos	40
2.2.2.	Control de la eficiencia energética y confort térmico de super- ficies comerciales de empresas textiles en Ciudad de Panamá	46
3.	Nuevo enfoque FDA para detección de atípicos en Estudios Inter- laboratorios (ILS)	51
3.1.	Introducción	51
3.1.1.	Consideraciones generales para realizar un estudio interlabo- ratorio	52
3.1.2.	Técnicas FDA aplicadas en un Estudio Interlaboratorio	55
3.2.	Metodología propuesta para la detección de laboratorios atípicos	70
3.3.	Validación del nuevo enfoque FDA	73
3.3.1.	Estudio bajo la hipótesis nula del comportamiento de los es- tadísticos de contraste	74
3.3.2.	Hipótesis de reproducibilidad	76
3.3.3.	Hipótesis de repetibilidad	78
3.3.4.	Efectos del tamaño de la muestra (número de resultados en cada laboratorio)	80
3.3.5.	Incumplimiento de la hipótesis de repetibilidad y reproducibi- lidad simultáneamente	81
3.3.6.	Comparación entre el enfoque FDA y el escalar	82
3.4.	Aplicación a datos reales: curvas termogravimétricas	83
4.	Propuesta de una metodología de gráficos de control basada en datos funcionales	90
4.1.	Introducción	90

4.1.1.	Alternativas del CEC cuando se incumplen las hipótesis básicas de los gráficos de control	93
4.1.2.	Construcción de un gráfico de control: Fase I, Fase II, muestra de calibrado y monitorizado	94
4.1.3.	Importancia de la Fase I en la construcción de un gráfico de control	96
4.2.	Metodología	98
4.2.1.	Procedimiento de construcción de un gráfico de control para la FASE I (estabilización)	99
4.2.2.	Procedimiento de construcción de un gráfico de control para la FASE II (monitorización)	102
4.3.	Estudio de simulación	105
4.3.1.	Medición y comparación del rendimiento del gráfico de control propuesto para la Fase I	108
4.3.2.	Medición y comparación del rendimiento del gráfico de control propuesto para la Fase II	119
4.4.	Aplicación a datos reales	120
5.	Conclusiones y líneas futuras	131
5.1.	Estudios Interlaboratorio	131
5.2.	Gráficos de control	133
5.3.	Líneas futuras	134
A.	Control estadístico de la calidad con el paquete qcr	137
A.1.	Introducción	138
A.2.	Creando un objeto qcs	143
A.3.	Conjuntos de datos en el paquete qcr	145
A.4.	Una alternativa: gráficos de control no paramétrico basados en la profundidad de los datos	146
A.4.1.	Profundidad de datos	147
A.4.2.	Estadísticos derivados de la profundidad de datos	147
A.4.3.	El gráfico r	148

B. El paquete ILS para realizar estudios intelaboratorios	150
B.1. Introducción	151
B.2. Ejemplos de Estudios Interlaboratorio	153
B.2.1. Estudio clínico de medida de glucosa en sangre	153
B.2.2. Caracterización de materiales mediante análisis termogravimétrico	158
B.3. Estudios Interlaboratorio: Enfoque estándar	160
B.3.1. ILS: Estudio de glucosa	162
B.4. Estudios Interlaboratorio: Nuevo enfoque FDA	166
B.4.1. ILS: Estudio Termogravimétrico	166
Bibliografía	171

Índice de figuras

1.1. Evolución de la Industria desde los inicios de la Primera Revolución Industrial a la Cuarta Revolución Industrial.	5
1.2. a) Modificaciones por descentrado. b) Modificaciones por aumento de la dispersión.	9
1.3. Proporción de elementos no conformes según el criterio Seis Sigma cuando el proceso está centrado o tiene una desviación con respecto a los límites de especificación.	11
1.4. Etapas de la metodología Seis Sigma: ciclo DMAIC.	13
1.5. Técnicas de CEC en relación con las etapas del proceso Seis Sigma.	13
1.6. Paquetes en R relacionados con el CEC con sus aplicaciones.	14
2.1. Curva TG correspondiente al oxalato de calcio en rojo. La curva TG proporciona información acerca del proceso global de degradación, en este caso en ambiente oxidante. Adicionalmente se muestra la curva de flujo de calor (en azul) y la relación de la temperatura con el tiempo (verde).	42
2.2. Muestra de oxalato de calcio en un crisol justo antes de ser introducido en el horno del analizador TG.	42
2.3. En este trabajo se han utilizado dos analizadores simultáneos que realizan análisis TG y DSC “simultáneamente” (instrumentos de medida). Panel izquierdo: analizador simultáneo SDT de TA Instruments, con la celda, crisoles y brazos sujetores en horizontal. Panel derecho: analizador simultáneo STA de Rheometrics Inc, con celda del horno y balanza en vertical.	43

2.4.	Curvas TG correspondientes a cada laboratorio. Han sido obtenidas a partir del ensayo de muestras de oxalato de calcio.	44
2.5.	Curvas DSC correspondientes a cada laboratorio. Han sido obtenidas a partir del ensayo de muestras de oxalato de calcio.	45
2.6.	Curvas diarias de consumo energético en sistemas HVAC.	47
3.1.	Se muestran las bandas de confianza para la media global de las curvas TG a un nivel de significación del 95 % y la varianza global funcional. Media y varianza funcional globales se estiman teniendo en cuenta todas las curvas	57
3.2.	En el panel izquierdo se muestran las estimaciones de la varianza y media funcional correspondientes a las curvas TG de oxalato de calcio obtenidas por cada laboratorio. En el panel derecho se observan las medias funcionales correspondientes al tercer proceso de degradación	57
3.3.	Distribución Bootstrap del estadístico del ANOVA Funcional	60
3.4.	Curvas DSC obtenidas por seis laboratorios a partir de muestras de Oxolate de Calcio.	61
3.5.	Curvas DSC, estimación de la media funcional con sus bandas de confianza al 95 %, y estimación de la varianza funcional correspondiente a las muestras oxalato de calcio.	62
3.6.	Estimación de las medias y varianzas funcionales de las curvas DSC obtenidas por cada laboratorio a partir de muestras de oxalato de calcio.	63
3.7.	(a) Distribución Bootstrap del estadístico del ANOVA Funcional . . .	66
3.8.	(a) Gráfico de barras correspondiente a los valores del estadístico h de cada laboratorio obtenido a partir de medidas de IDT. (b) Gráfico de barras que corresponde a los valores del estadístico k de cada de laboratorio obtenido a partir de medidas de IDT.	68
3.9.	(a) Gráfico de barras correspondiente a los valores del estadístico h de cada laboratorio obtenido a partir de la pérdida de masa en el segundo paso. (b) Gráfico de barras que corresponde a los valores del estadístico k de cada laboratorio obtenido a partir de la pérdida de masa en el segundo paso.	69
3.10.	(a) Distribución de p -valores para el estadístico d_H bajo la hipótesis nula, (b) p -valores para el estadístico d_K bajo la hipótesis nula.	76

3.11. (a) Medias teóricas para curvas TG simuladas bajo la hipótesis nula $H_0 : m_0 = 0.5$ (línea continua) y la alternativa $H_1 : m_1 = m_0(1 + \delta_H)$ ($\delta_h = 0.005$, la línea a trazos), (b) el estadístico funcional $H(t)$ para cada uno de los 6 laboratorios simulados bajo la nula (línea continua) y un laboratorio bajo la alternativa (línea a trazos), (c) la distribución estimada por remuestreo bootstrap del estadístico d_H , y (d) el estadístico d_H para cada muestra de laboratorio. 77

3.12. (a) Varianzas teóricas para las curvas TG simuladas bajo la hipótesis nula $H_0 : v_0 = 5$ (línea continua) y la alternativa, $H_1 : v_1 = v_0(1 + \delta_K)$ ($\delta_k = 2$, la línea a trazos), (b) el estadístico funcional $K(t)$ para cada uno de los 6 laboratorios simulados bajo la nula (línea continua) y un laboratorio bajo la alternativa (línea a trazos), (c) la distribución estimada por remuestreo bootstrap del estadístico d_K , y (d) el estadístico d_k para cada muestra de laboratorio. 79

3.13. (a) Curvas de potencia (proporción de rechazo) para el estadístico d_H correspondiente a $n = 10$ y $n = 20$ muestras por laboratorio, (b) curvas de potencia para el estadístico d_K correspondiente a $n = 10$ y $n = 20$ muestras por laboratorio. Las líneas segmentadas corresponden a $n = 20$, mientras que las líneas continuas corresponden a $n = 10$. . . 80

3.14. (a) Potencias de prueba correspondientes al caso univariante (5tw % en peso y 10tw % en peso) en trazos continuos y la aproximación FDA para el estadístico h en trazo discontinuo, con $n = 10$, (b) Potencias de prueba correspondientes al caso univariado (5tw % en peso y 10tw % en peso) en trazos continuos y un enfoque FDA del estadístico h en trazo discontinuo, con $n = 20$ 83

3.15. (a) El histograma del estadístico de prueba d_H y c_H para $\alpha = 0.01$ (primera iteración), (b) Los estadísticos d_H para la muestra de los laboratorios (primera iteración), El histograma del estadístico de prueba d_H y c_H para $\alpha = 0.01$ (cuarta iteración), (b) Los estadísticos d_H para la muestra de los laboratorios (cuarta iteración). 86

3.16. (a) Estadístico funcional $H(t)$ con bandas de confianza del 99 % (primera iteración), (b) Estadístico funcional $H(t)$ con bandas de confianza del 99 % (cuarta iteración). 86

3.17. (a) Estadístico funcional K para cada laboratorio y bandas de confianza para $\alpha = 0.01$ (primera iteración), (b) estadístico de prueba d_K para cada laboratorio y nivel crítico correspondiente a $\alpha = 0.01$ (primera iteración), (c) Estadístico funcional K para cada laboratorio y bandas de confianza $\alpha = 0.01$ (segunda iteración), y (d) estadístico de prueba d_K para cada laboratorio y nivel crítico correspondiente a $\alpha = 0.01$ (segunda iteración).	88
4.1. Proceso de monitorización de una variable que afecta a la calidad de un producto o servicio mediante la aplicación de gráficos de control. .	97
4.2. (a) Medias funcionales y (b) cambios en la media del proceso en forma y magnitud.	106
4.3. Escenarios en los que se considera independencia entre curvas y para cambios en la media funcional con respecto a su magnitud (a) y forma (b). En el caso de dependencia, los paneles (c) y (d) muestran escenarios de simulación en los que se observan cambios de magnitud y forma, respectivamente en la media funcional.	108
4.4. Detección de consumos energéticos en HVAC atípicos mediante la aplicación de boxplot a cada uno de los consumos horarios.	122
4.5. Curvas diarias de consumo energético en instalaciones HVAC para la tienda de Panamá City. Se ha representado las curvas pertenecientes a la muestra de calibrado, a partir de las cuales se estiman los límites de control naturales (Fase I).	123
4.6. Panel izquierdo: Curvas diarias de consumo energético en instalaciones HVAC para la tienda de Panamá City, identificando aquellas curvas detectadas como anomalías, además de medidas de posición funcionales. Panel derecho: Gráfico de control correspondiente a la Fase I, mediante el cual se detectan las curvas anómalas, eliminado así las causas asignables de variación. Se muestran las profundidades de cada una de las curvas de consumo diario y el límite de control natural.	125

4.7. Proceso bajo control (consumo energético en instalaciones HVAC estabilizado). Panel izquierdo: Curvas diarias de consumo energético en instalaciones HVAC para la tienda de Panamá City, indicando la envolvente correspondiente al 95% de curvas más profundas. Panel derecho: gráfico de control para la Fase I basado en la profundidad de datos funcionales.	126
4.8. Muestra de referencia vs. monitorizada para el mes de octubre	126
4.9. Muestra de referencia vs. monitorizada para el mes de noviembre . . .	127
4.10. Gráfico rank: Muestra de referencia vs. monitorización para el mes de noviembre	128
A.1. Tablas de control implementadas en el paquete.	140
A.2. Comparación entre los principales paquetes de R dedicados al control estadístico de calidad y el paquete.	143
B.1. Mediciones de la concentración de glucosa en sangre utilizadas para controlar la diabetes.	155
B.2. Las siguientes medidas de variabilidad se calculan para cada material a partir de las medidas de todos los laboratorios. S es la desviación global de todos los laboratorios, S_r es la desviación de la repetibilidad, S_B la desviación entre las medias de los laboratorios y S_R la desviación, en función de la reproducibilidad.	157
B.3. Curvas de TG obtenidas de oxalato de calcio.	160
B.4. Estadística descriptiva funcional: medias y varianzas funcionales para cada laboratorio, así como la media y la varianza global de laboratorio de las curvas de TG obtenidas a partir de oxalato de calcio.	161
B.5. Estadísticos k por laboratorio y material correspondientes a las medidas de glucosa en sangre disponibles en la base de datos Glucosa. . .	165
B.6. Estadísticos h por laboratorio y material.	167
B.7. Conjunto de datos TG: Estadísticos funcionales $H(x)$ y $K(x)$ para cada laboratorio. Estadísticos de contraste d_H y d_K para cada laboratorio.	168
B.8. Conjunto de datos DSC: Estadísticos funcionales $H(x)$ y $K(x)$ para cada laboratorio. Estadísticos de contraste d_H y d_K para cada laboratorio.	170

Índice de cuadros

1.1. ANOVA para la estimación de σ_r^2 y σ_L^2	16
2.1. Etiqueta y descripción de cada uno de los laboratorios.	46
3.1. Profundidades utilizando el método de proyección aleatoria correspondientes a las diferentes curvas TG obtenidas de muestras de oxalato de calcio. Las curvas TG correspondientes a las profundidades marcadas en negrita se consideran atípicas (cuando la profundidad de una curva TG es menor que el percentil 10 de la distribución de las profundidades, 0.0908)	59
3.2. Profundidades utilizando la medida modal correspondientes a las diferentes curvas TG obtenidas a partir de muestras de oxalato de calcio. Las curvas TG correspondientes a las profundidades marcadas en negrita han sido identificadas como valores atípicos (si la profundidad de una curva TG es menor que el percentil 10 de la distribución de las profundidades, 2.9245)	59
3.3. Profundidades utilizando el método por proyección aleatoria correspondientes a las diferentes curvas DSC obtenidas de muestras de oxalato de calcio. Las curvas DSC correspondientes a las profundidades marcadas en negrita se definen como valores atípicos (si la profundidad de una curva DSC es menor que el percentil 10 de la distribución de las profundidades, 0.0798).	64

3.4. Profundidades utilizando el método modal correspondientes a las diferentes curvas DSC obtenidas de muestras de oxalato de calcio. Las curvas DSC correspondientes a las profundidades marcadas en negrita se definen como valores atípicos (si la profundidad de una curva DSC es menor que el percentil 10 de la distribución de las profundidades, 3.199)	65
3.5. Media y desviación estándar del IDT de los 7 laboratorios estudiados	67
3.6. Resultados correspondientes a los datos IDT (primer paso) al aplicar la norma ASTM 691	68
3.7. Resultados utilizando como característica la pérdida de masa en el segundo proceso de degradación y el ASTM 691	69
3.8. Proporción de rechazo (%) del estadístico d_H y d_K , con $\alpha = 0.01$ y $n = 10, 20, 30$	75
3.9. Porcentaje de rechazo p_v y p_{vm} del estadístico d_K , con $\alpha = 0.01$ y $n = 10$	81
3.10. Proporción de rechazo estadístico d_K (con $\alpha = 0.01$) en comparación con la proporción de rechazo del estadístico k univariante cuando se calcula a partir de las variables escalares tiempo/temperatura a pérdida del 5tw % y 10tw % en peso de la masa de la muestra.	84
4.1. Resultados de la estimación de la tasa de falsas alarmas (Error I) para el caso de independencia entre curvas	110
4.2. Porcentajes de \hat{p}_c y \hat{p}_f para los casos M_1 (expresión 4.5) y M_2 (expresión 4.6) bajo el supuesto de independencia entre curvas.	111
4.3. Resultados para el caso dependencia de la estimación de la tasa de falsas alarmas (error de tipo I) para valores de ρ que van de 0.1 hasta 0.9 en pasos de 0.2.	113
4.4. Resultados de los valores empíricos de \hat{p}_f y \hat{p}_c con $\rho = 0.1$ para el caso dependencia entre curvas.	114
4.5. Resultados de los valores empíricos de \hat{p}_f y \hat{p}_c con $\rho = 0.3$ para el caso dependencia entre curvas.	115
4.6. Resultados de los valores empíricos de \hat{p}_f y \hat{p}_c con $\rho = 0.5$ para el caso dependencia entre curvas.	116

4.7. Resultados de los valores empíricos de \hat{p}_f y \hat{p}_c con $\rho = 0.7$ para el caso dependencia entre curvas.	117
4.8. Resultados de los valores empíricos de \hat{p}_f y \hat{p}_c con $\rho = 0.9$ para el caso dependencia entre curvas.	118
4.9. Potencia del gráfico de control $1 - \beta$ para el caso M_1 (expresión 4.5) en el escenario de independencia entre curvas.	119
4.10. Potencia del gráfico de control $1 - \beta$ para el caso M_1 (4.5) en el escenario de dependencia	120
A.1. PCR de primera a cuarta generación, USL es el límite de especificación superior, LSL es el límite de especificación más bajo, μ es la media real, $\hat{\mu}$ es la media estimada y $\hat{\sigma}$ es la desviación estándar estimada. Tablas de control implementadas en el paquete.	141
A.2. Cuadros de control de Shewhart disponibles en el paquete	144
A.3. Conjuntos de datos incluidos en el paquete	146
B.1. Funciones incorporadas en el paquete ILS para realizar Estudios Interlaboratorio con un enfoque escalar y funcional.	161

Capítulo 1

Introducción

En este Capítulo se presenta una introducción general a los conceptos y metodologías estadísticas e ingenieriles desarrolladas en esta memoria de tesis para la resolución de problemas en el control de la eficiencia energética de edificios y la detección de laboratorios (instrumentos de medida o sensores) que proporcionan resultados atípicos, en el contexto de los estudio interlaboratorio, cuando los datos experimentales que se obtienen son datos funcionales.

Estos dos problemas se pueden enmarcar en el ámbito de la Metrología 4.0 y, por extensión, en la Industria 4.0. En ambos casos de estudio, para su resolución, se requiere generar, implementar y utilizar nuevas metodologías estadísticas dentro de sistemas de control de calidad, ya sea para mantener y mejorar la competitividad en las industrias de bienes y servicios o bien para estimar correctamente una determinada característica de una sustancia, medir la precisión de un procedimiento de medida o el desempeño de un laboratorio, instrumento de medida o sensor. Por consiguiente, en este capítulo se incluye una breve reseña acerca de los conceptos de Industria y Metrología 4.0, además de su relación con los casos de estudio analizados y las soluciones propuestas.

En el contexto global de la Industria 4.0, entendida como toda acción orientada a la digitalización de procesos industriales, surgen nuevos retos relacionados con la metrología, que tiene por objeto la obtención de la medida de una magnitud, con su incertidumbre. Estos, entre otros, consisten en la obtención de una mayor cantidad de datos y un aumento en la complejidad de los mismos, debido en gran parte al

empleo de sensores e instrumentos de medida cada vez más sofisticados y capaces de proporcionar un mayor número de medidas por unidad de tiempo correspondientes a un número mayor de variables críticas para la calidad de los procesos y servicios estudiados. En consecuencia, los nuevos retos que definen a la Metrología 4.0 están relacionados con el Big Data y el tratamiento de datos más complejos, entre los que se encuentran los datos funcionales.

Llegados a este punto, es importante destacar que todos los procedimientos estadísticos propuestos para proporcionar soluciones a dichos problemas pertenecen al ámbito del análisis de datos funcionales (FDA). Esto se debe a que los datos obtenidos en los casos de estudio que componen esta tesis, ya sea por sensores en edificios ya sean las curvas experimentales obtenidas en laboratorios, pueden considerarse de naturaleza funcional.

En definitiva, el desarrollo de nuevas metodologías, en este caso FDA, es de vital importancia para el manejo de los nuevos paradigmas y tipos de datos que la digitalización de la industria de productos y servicios proporciona. Su desarrollo y aplicación es absolutamente necesaria, no sólo para garantizar la calidad de productos, servicios, procedimientos, instrumentos y laboratorios, sino también para incrementar la eficiencia de los procesos de medición, reducir el coste de estudios incompletos y mejorar en el análisis de resultados.

Teniendo en cuenta el ámbito en el que se desarrollan los casos de estudio analizados, el tipo de datos que los definen y las metodologías desarrolladas para darles una solución, previamente a la descripción y evaluación de los procedimientos estadísticos propuestos, es necesario describir los conceptos de Industria y Metrología 4.0, el análisis de datos funcionales, el control estadístico de la calidad, los estudios interlaboratorio como un caso particular de los estudios de reproducibilidad y repetitividad y los gráficos de control como herramienta para controlar la variabilidad de un proceso y detectar anomalías. Todos estos conceptos se desarrollan como secciones del presente capítulo de introducción.

1.1. Industria 4.0 y Metrología 4.0

El concepto Industria 4.0, o cuarta revolución industrial, representa una nueva era para la organización de la producción. Uno de sus objetivos es la puesta en marcha de la llamada “fábrica inteligente” (smart factory), capaz de una mayor adaptabilidad a las necesidades y a los procesos de producción, así como a una asignación más eficiente de los recursos, abriendo así vías para una nueva era industrial. Industria 4.0 representa un nuevo hito en el desarrollo industrial que sin duda marcará, y ya está marcando, importantes cambios sociales en los próximos años. Sus principales características son hacer un uso intensivo de la simulación de procesos, emplear el llamado Internet de las cosas, el Big Data, la Simulación, aparte de otras tecnologías punta con el fin de desarrollar plantas industriales con cadenas de producción mucho mejor comunicadas entre sí y más competitivas. En el caso particular de esta tesis doctoral, el problema del control de la variabilidad y detección de anomalías en la eficiencia energética y confort térmico de superficies comerciales, el concepto Industria 4.0 está en la toma de datos de todas las variables críticas para la calidad de los sistemas de climatización (HVAC, Heat, Ventilation and Air Conditioning) que promuevan el control automático de dichas instalaciones. Se incluye por tanto la tecnología de sensores, la plataforma web para el control y monitorización de variables (gestión de Big Data) y las herramientas de análisis estadístico para visualización, detección de anomalías y decisión (analytics).

Por otro lado, definiendo Metrología como la ciencia que se ocupa de la obtención y expresión del valor de las magnitudes empleando para ello instrumentos, métodos y medios apropiados, con la exactitud requerida en cada caso, la Metrología 4.0 se define como el conjunto de nuevas tendencias de la Metrología para la medición avanzada e inteligente con el fin de satisfacer las necesidades del entorno productivo en aras de conseguir una producción cada vez más eficiente. Así, los sensores inteligentes son un ejemplo de la Metrología 4.0, que permiten realizar medidas on-line en planta, para optimizar la respuesta de los procesos de fabricación en la propia línea de producción, asegurando la calidad del producto, reduciendo su ciclo productivo y, lo que es muy importante, los costes correspondientes.

1.1.1. Industria 4.0

El concepto de Industria 4.0 surge motivado por la última gran crisis de la década de los 80 cuando Alemania, preocupada por la falta de competitividad de sus empresas, empieza a mirar con objetividad a sus más directos competidores, el mercado asiático. De esta comparación surge la necesidad de dar un vuelco a la industria para hacerla más competitiva. Esta primera aproximación al problema, que se plantea en la cumbre de Hannover de 2011, será seguida por el resto de países europeos con mayor o menor celeridad. En Estados Unidos se han desarrollado iniciativas propias para el impulso de la Industria 4.0 como el proyecto “Smart Manufacturing Leadership Coalition” (Davis *et al.*, 2012), orientado a las modalidades de la fabricación industrial del futuro, mientras que multinacionales como General Electric trabajan en proyectos propios como el Industrial Internet.

Como puede verse en la Figura 1.1, que muestra los principales cambios metodológicos en función del tiempo, la evolución de la industria se puede describir mediante una sucesión de etapas asociadas a grandes cambios tecnológicos o “revoluciones industriales”. De este modo, la primera revolución o Industria 1.0 se correspondería con la aparición de la máquina de vapor, mientras que la Industria 2.0 está caracterizada por la incorporación de la producción en cadena, hito en la organización y automatización del trabajo introducido por las primeras industrias del automóvil (como es el caso de la Ford).

Siguiendo la secuencia mostrada en la Figura 1.1, el mayor cambio producido en la industria ha sido sin duda la implementación de la electrónica y la informática de forma masiva, lo que podría definirse como Industria 3.0. De ahí a la situación actual, que denominamos Industria 4.0, han pasado muy pocos años, pues estas evoluciones se han dado en intervalos de tiempo cada vez más cortos, como también puede verse en la Figura 1.1.

El objetivo de la Industria 4.0 es la optimización de los procesos de producción, para lo que se incide en el desarrollo y aplicación de los denominados 9 pilares tecnológicos de la Industria 4.0: Big Data y Análisis de Datos, Robots Autónomos, Simulación, Sistemas para la Integración Vertical y Horizontal, Internet de las Cosas Industrial o IIoT, Ciberseguridad, Cloud Computing, Fabricación Aditiva y Realidad Aumentada (Davis *et al.*, 2012). La presente tesis doctoral busca aportar soluciones

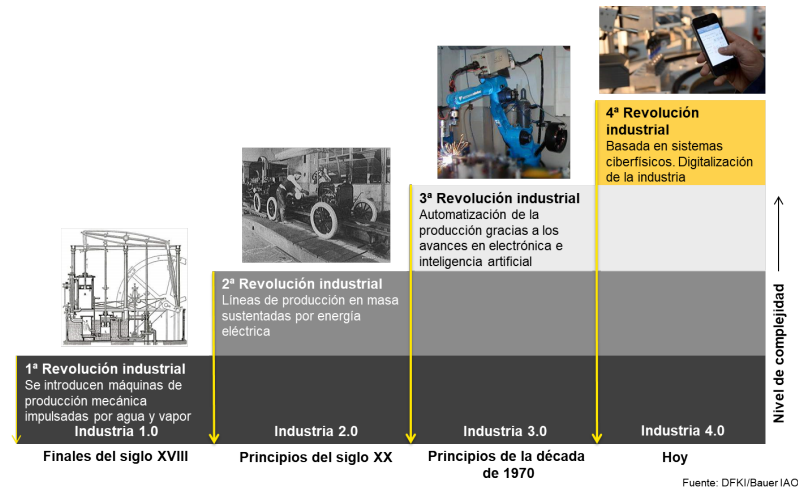


Figura 1.1: Evolución de la Industria desde los inicios de la Primera Revolución Industrial a la Cuarta Revolución Industrial.

a la problemática originada por el uso masivo de datos debida a la digitalización creciente de los procesos productivos, a su vez consecuencia de los avances en instrumentación y sensórica. Por tanto, de entre estas 9 nuevas tecnologías que configuran la industria 4.0, esta tesis está principalmente relacionada con el Big Data y Análisis de Datos, además de (de una manera indirecta) con el IIOT, teniendo en cuenta que la aplicación de esta tecnología aporta una parte importante de los datos reales tratados en este trabajo.

Los requerimientos de la Industria 4.0 de digitalización de los procesos productivos y productos conlleva el traslado de todas las actividades relacionadas con la metrología del laboratorio a la planta productiva. Es aquí donde aparece el concepto de Metrología 4.0, que se define por la adquisición rápida de datos (mediante tecnologías de sensores sin contacto), el procesamiento automático de las medidas y la toma automática de decisiones mediante modelos, permitiendo la comunicación del sistema metrológico con el sistema de control del proceso (Mutilba y Kortaberria, 2016). Para satisfacer estas exigencias, los principales suministradores de equipamiento metrológico comienzan a comercializar soluciones que integran sistemas metrológicos, hardware y software (Imkamp *et al.*, 2014; Mutilba y Kortaberria, 2016). Estos sistemas pretenden adaptar los procedimientos de medición offline a la

fabricación (in-process) con el fin de controlar los procesos en planta, con un mayor conocimiento de las variables críticas para la calidad del proceso. Son sistemas que requieren la recogida de datos, su procesado y la toma de una decisión inmediata teniendo en cuenta la monitorización de las características críticas para la calidad del proceso.

1.1.2. Metrología 4.0

Los principales retos que implica el desarrollo de la Metrología 4.0, entendida como un conjunto de técnicas favorecedoras de la digitalización de los procesos productivos, son los siguientes:

- Conseguir la trazabilidad de las medidas bajo condiciones en planta no tan controladas como en laboratorio (condiciones ambiente, geometría de las máquinas, calibración, entre otras). Esto puede originar un aumento en la incertidumbre de las mediciones (Mutilba y Kortaberria, 2016; Schmitt *et al.*, 2016).
- Medir y aplicar medidas correctivas en tiempo real.
- Gestión de los datos masivos o Big Data resultantes, que implica el desarrollo de software.
- Desarrollar algoritmos de detección de anomalías, toma de decisiones y mantenimiento predictivo.
- Interconectar todos los sistemas de medición de una planta productiva mediante plataformas ad hoc diseñadas para la gestión integral. Mediante estas plataformas se centraliza la información, se gestionan los datos de medición desde un único punto de acceso, controlando el proceso productivo.
- Contar con un perfil profesional que permita trasladar la inteligencia de las personas a los procesos productivos, mediante el desarrollo de algoritmos diseñados para identificar la información útil para la toma de decisiones.

En la actualidad, se está desarrollando el binomio fabricación-medición, es decir, en la medición dentro de la propia máquina herramienta o instalación que presta el servicio. Un ejemplo de este desarrollo es el proyecto europeo Traceable In-process Metrology (T.I.M.) (Mutilba y Kortaberria, 2016; Schmitt y Peterek, 2015).

En la presente tesis doctoral, se presenta un problema directamente relacionado con la Industria y Metrología 4.0. Este es el control de la eficiencia energética y confort térmico en superficies comerciales. En particular, se pretende estudiar las instalaciones HVAC de superficies comerciales de una conocida empresa textil española, situadas en la Ciudad de Panamá. De hecho, se ha desarrollado una plataforma web de gestión de Big Data denominada Σ qus, especialmente diseñada para la gestión de las mencionadas superficies comerciales, dotándolas de aplicaciones tipo Supervisory Control and Data Acquisition (SCADA), definidas como sistemas de monitorización a partir de la información prestada por sensores en planta (consumo energético, temperaturas, calidad del aire, radiación solar, etc.) y actuación o consigna. Esta plataforma actualmente está demandando la incorporación de herramientas estadísticas que permitan la detección de anomalías de forma automática y la identificación de causas asignables a partir de datos continuamente monitorizados.

1.2. Control estadístico de la calidad y metodología Seis Sigma

En esta tesis doctoral se han propuesto soluciones a los problemas planteados (detección de laboratorios atípicos en ILS, detección de anomalías en la eficiencia energética de edificaciones) mediante la aplicación de técnicas pertenecientes al control estadístico de la calidad o CEC. Es frecuente denominar al control estadístico de la calidad, control estadístico de procesos (CEP o SPC del inglés), esta versión proviene de la orientación hacia los procesos industriales, aunque lógicamente la idea de proceso abarca mucho más.

1.2.1. Proceso bajo control

Un concepto muy importante en el control de calidad es el de proceso bajo control. Para explicar su significado es necesario introducir previamente lo que se entiende por causas de variabilidad asignables y no asignables.

Se llamarán causas asignables (especiales o esporádicas) de variabilidad a las que producen efectos predecibles que aumentan la variabilidad. Normalmente son

muy pocas, pero de mucha influencia. Por ejemplo un fallo en una máquina, una mezcla de materias primas en una proporción inadecuada, etc. Por el contrario, se llamarán causas no asignables (comunes o aleatorias) a aquellas que aparecen con efectos combinados, no predecibles de antemano e inherentes a la incertidumbre del proceso productivo. Suele haber muchas causas no asignables, cada una de las cuales tiene un efecto individual muy pequeño. Ejemplos de estas son la heterogeneidad de las materias primas, la precisión de las máquinas, la de los elementos de medida o la destreza de los operarios, entre otras.

Estudiando con detalle un proceso, es posible eliminar todas las causas asignables de variabilidad, para que así, la única variabilidad existente se deba a causas no asignables. En tal caso, diremos que el proceso se encuentra bajo control (o en estado de control). En la práctica, no existen procesos que se encuentren espontáneamente bajo control. A menudo son necesarios muchos esfuerzos para conseguir llevar un proceso a un estado de control. Ese es, precisamente, uno de los objetivos del control de procesos.

La existencia de tolerancias obliga a que las unidades que no cumplan este requisito sean consideradas como defectuosas o, en términos más usuales, no conformes. La producción de este tipo de unidades, o su rectificación, conlleva un coste que podría ser eliminado; pero también la inspección y la búsqueda de las causas de esta situación presenta un coste monetario cada vez mayor. Surge así el reto de intentar reducir la proporción de unidades no conformes hasta un punto en el que el incremento del coste de inspección y control iguale al incremento de ahorro por la disminución del número de unidades rechazadas.

La inspección o medición de algunas características de calidad (como pueden ser la resistencia o tiempo de vida útil) son de carácter destructivo, por lo que no pueden ser llevadas a cabo sobre todas las unidades; incluso aunque las pruebas no fuesen destructivas, la inspección del 100 % de la producción puede ser improcedente económicamente. Aparece entonces la necesidad de garantizar la viabilidad de la producción de forma general mediante la inspección de una parte significativa de ésta y, por tanto, la necesidad de determinar qué proporción de unidades debe ser inspeccionada, cómo seleccionar dichas unidades y cómo garantizar que los resultados obtenidos son extrapolables a la generalidad de la producción.

Las causas asignables dan lugar a dos tipos distintos de modificaciones: modificaciones por descentrado y modificaciones por aumento de variabilidad. Las primeras son las que se producen cuando se presenta una traslación del valor medio de la característica de calidad, mientras que en las segundas lo que ocurre es un aumento de la variabilidad (véase la Figura 1.2):

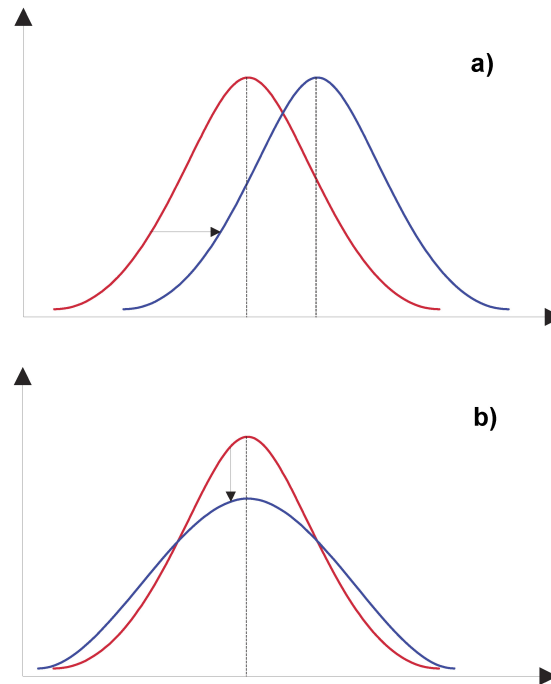


Figura 1.2: a) Modificaciones por descentrado. b) Modificaciones por aumento de la dispersión.

Las modificaciones por descentrado del proceso se observan sobre todo en procesos de fabricación automáticos mientras que en los semiautomáticos son más frecuentes las modificaciones por dispersión. En la práctica, ambos tipos de modificaciones se mezclan y confunden, dando lugar en cualquier caso a un incremento del número de piezas fuera de especificación.

Las técnicas de control de calidad en curso de fabricación tienen por objeto mantener estable la variabilidad propia del proceso y detectar con la mayor eficacia posible la aparición de causas asignables de fluctuación. La manera de controlar un proceso es mediante el control de la variabilidad.

1.2.2. Metodología Seis sigma

Seis Sigma es, sin duda, la herramienta de mayor implantación para la Gestión de la Calidad hoy en día. El término Seis Sigma que hace referencia a la costumbre de representar la desviación típica con esta letra griega sigma (σ), prima la importancia del control de la variabilidad, siendo pues una metodología que tiene una fuerte base estadística. Seis Sigma es una metodología de mejora de procesos que persigue la reducción de la variación, los defectos y otros tipos de errores (Cano *et al.*, 2012).

Seis Sigma nace en los años 80 en la empresa Motorola, de la mano del Dr. Mikel Harry, como un intento para conseguir mejoras sustanciales en la calidad del producto. En sus primeros años Seis Sigma se circunscribía a la mejora de procesos, pero entendida como mejora radical y profunda: una verdadera “ruptura” (según la propia terminología Seis Sigma) en los métodos de producción tradicionales. Entre los objetivos de Seis Sigma está alcanzar unos niveles de calidad en los procesos y productos que no superen los 3.4 defectos por millón.

En la década de 1990, Jack Welch, presidente de General Electric, decidió utilizar Seis Sigma consiguiendo resultados económicos espectaculares. Desde entonces, Seis Sigma se ha convertido en una de las herramientas de mejora más utilizadas, habiendo sido adoptadas, además de Motorola y General Electric, por compañías como Polaroid, Toshiba, Honeywell, City Bank, American Express, Movistar o Iberia, entre muchas otras. Aunque el concepto Seis Sigma fue acuñado por Motorola los pasos de los que consta su aplicación fueron llevados a cabo por primera vez en USA ya en 1985 en la empresa Florida Power and Light cuando decidió solicitar el Premio Deming (Voehl, 2000).

Como se comentó una de las características de esta metodología está en buscar los “cero defectos”. En términos prácticos puede decirse que un proceso de producción tiene un nivel de calidad Seis Sigma cuando a largo plazo presenta poco más de tres DPMO (defectos por millón de oportunidades). Este “nivel sigma” de un proceso mide la distancia entre la media y los límites superior e inferior de la especificación correspondiente (ver Figura 1.3).

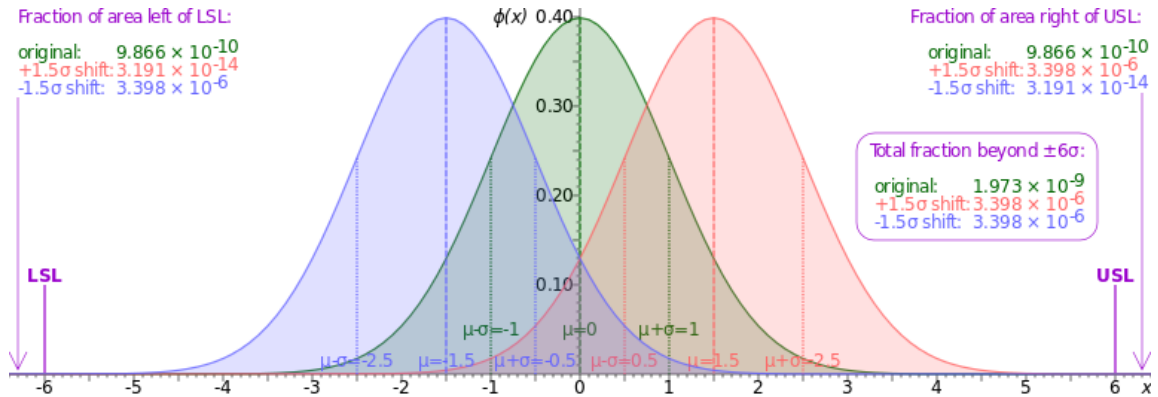


Figura 1.3: Proporción de elementos no conformes según el criterio Seis Sigma cuando el proceso está centrado o tiene una desviación con respecto a los límites de especificación.

Ha sido habitual considerar suficiente que un proceso tuviese una desviación de $\pm 3\sigma$, lo que viene a significar que dicho proceso era capaz de producir sólo 2.7 defectos por cada mil oportunidades (equivalente a 2700 DPMO), lo que equivale a trabajar con un nivel de confianza del 99.73%. Hoy en día este nivel de calidad es inaceptable para muchos procesos. Por ejemplo en los procesos que involucren varias componentes, si se supone que la probabilidad de fallo de cada componente es independiente, en el caso de contarse con 100 de estas componentes, cada una con un nivel de confianza del 99.73%, el nivel de confianza final del sistema sería inferior al 77% ($0.9973^{100} = 0.763$), lo que supondría que un 23% de los productos fuesen defectuosos. Teniendo en cuenta muchos productos, como los automóviles, los barcos, aviones, etc. están compuestos por más de un millón de componentes estos niveles de calidad son claramente inadmisibles.

Más recientemente, y como consecuencia natural de la aplicación de la metodología, el concepto Seis Sigma se ha extendido hasta convertirse en una verdadera filosofía de gestión global de la empresa. Una de sus características es que promueve la utilización de herramientas y métodos estadísticos de manera sistemática y organizada, que permite a las empresas alcanzar considerables ahorros económicos a la vez que mejorar la satisfacción de sus clientes, todo ello en un período de tiempo muy corto.

Los cambios se consiguen traduciendo las necesidades de los clientes al lenguaje de las operaciones y definiendo los procesos y las tareas críticas que hay que realizar de manera excelente. Se fijan objetivos que además de buscar la reducción de defectos relacionan esta mejora en ganancias a corto plazo para la empresa o institución que los aplique.

Para su puesta en práctica, aunque requiere de la implicación de todo el personal de la empresa, suelen formarse equipos técnicos que se gestionan siguiendo cierta similitud con los grados de las artes marciales. Así es frecuente que un equipo Seis Sigma esté liderado por un cinturón negro (black belt) con varios técnicos conocedores del proceso que harían el papel de cinturones verde (green belts). En el artículo de Hoerl (2001) pueden verse las fases y exigencias de formación de estos equipos Seis Sigma. Los elementos clave que soportan la filosofía Seis Sigma son los siguientes:

- (a) Conocimiento de los requerimientos del cliente.
- (b) Dirección basada en datos y hechos.
- (c) Mejora de procesos.
- (d) Implicación de la dirección.

Entre sus características se encuentra el uso de una serie de etapas, el denominado ciclo DMAIC, para conseguir la mejora de los procesos (ver Figura 1.4).

Hoy en día la filosofía Seis Sigma está siendo de gran aplicación en todo tipo de organizaciones y se ha visto ampliado mediante las técnicas de optimización basadas en eliminar lo innecesario, lo que se ha dado en llamar “Lean” (vocablo que podría traducirse como “eliminar grasa”), surge así la metodología Lean Six Sigma, que está siendo de gran aplicación auspiciada por la gran crisis económica mundial.

La metodología Seis Sigma, como ya se ha comentado, proporciona el esquema de aplicación de las diversas técnicas del CEC, de acuerdo con cada etapa del procedimiento de reducción de variabilidad. Por ello, en este capítulo, se describirán las principales técnicas del CEC en el marco de las etapas del ciclo DMAIC en las cuales se suelen aplicar. En la Figura 1.5 se enumeran las técnicas del control estadístico de la calidad susceptibles de aplicarse en cada una de las etapas del procedimiento Seis Sigma. Es importante destacar que una misma técnica del CEC puede aplicarse en más de una etapa del ciclo DMAIC.

Para la aplicación de las técnicas del CEC indicadas en la Figura 1.5, es necesari-

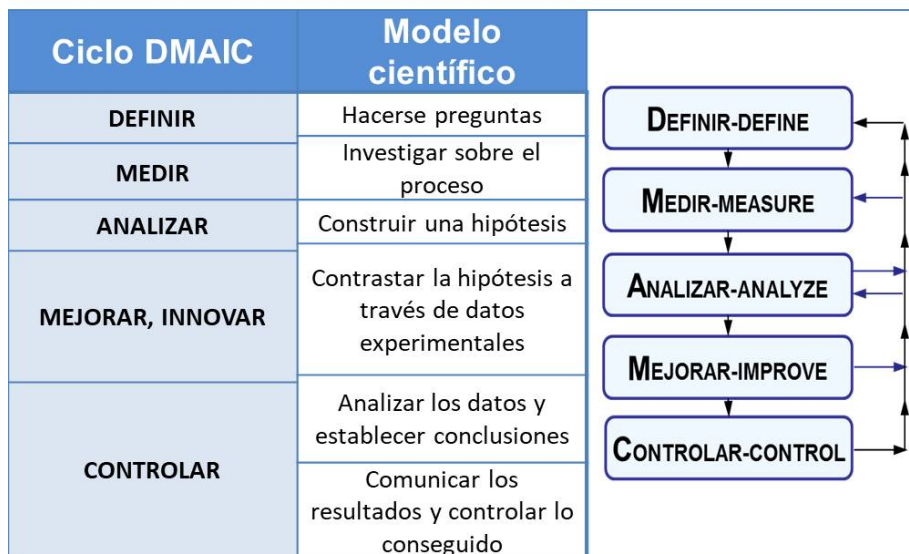


Figura 1.4: Etapas de la metodología Seis Sigma: ciclo DMAIC.

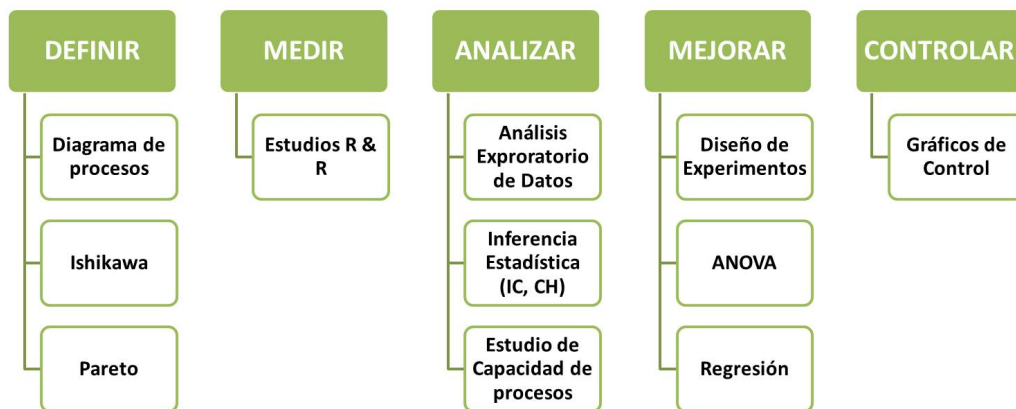


Figura 1.5: Técnicas de CEC en relación con las etapas del proceso Seis Sigma.

rio disponer de herramientas computacionales adecuadas. Entre otros programas, el software estadístico R proporciona una amplia variedad de funciones para la aplicación de técnicas CEC, disponibles en una serie de librerías entre las que destacan el *qcc* (Strucca, 2004), *qualityTools* (Roth, 2012), *SixSigma* (Cano *et al.*, 2012), *ICC* (Wolak *et al.*, 2012) y *qcr* (quality control review) (Flores, 2016), este último desa-

rollado por el autor de la presente tesis doctoral y que se describe en el Apéndice A. La Figura 1.6 muestra las principales utilidades que proporciona cada librería.

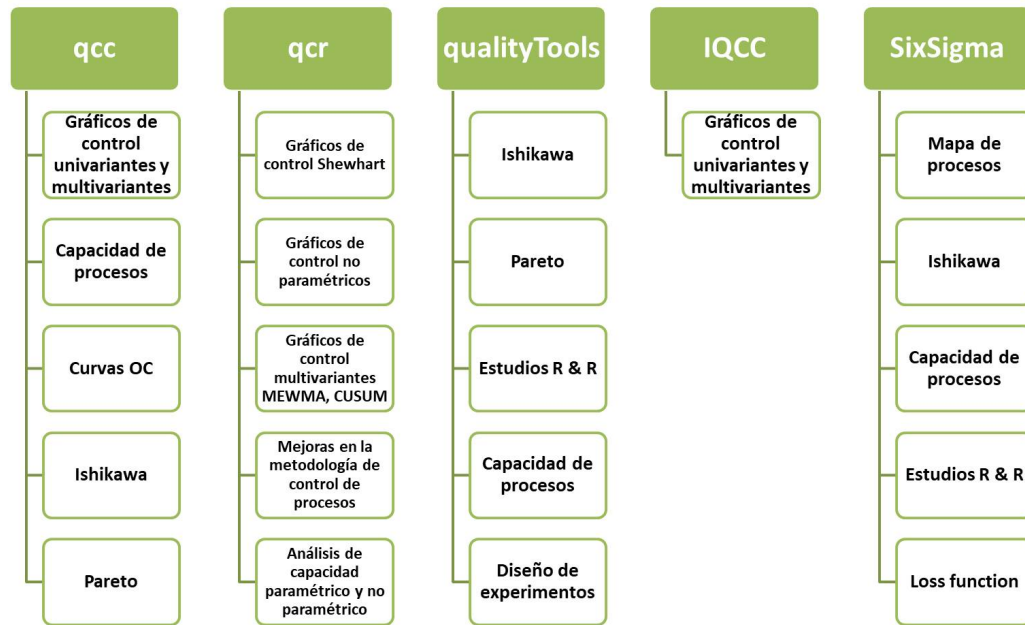


Figura 1.6: Paquetes en R relacionados con el CEC con sus aplicaciones.

1.3. Estudios interlaboratorio

Una parte fundamental de esta tesis doctoral, en concreto el Capítulo 3, está relacionada con la propuesta de nuevas técnicas estadísticas para la detección de laboratorios no consistentes o atípicos mediante el análisis de datos funcionales en el contexto de los estudios interlaboratorios. Los estudios interlaboratorio o ILS se realizan con diferentes objetivos y, en consecuencia, requieren diferentes métodos de evaluación y tratamiento estadístico. Las motivaciones principales para realizar un estudio interlaboratorio, que a su vez definen los diferentes tipos de ILS, son las siguientes:

- ILS colaborativos para evaluar el desempeño de un método de análisis expe-

rimental. Por ejemplo: la evaluación de un nuevo procedimiento experimental para obtener la cantidad de una determinada proteína en tejidos a partir de datos proporcionados por resonancia magnética molecular, o mismo por espectroscopía del infrarrojo cercano.

- Los estudios de desempeño de laboratorio en particular o pruebas de competencia para llevar a cabo un determinado análisis. Por ejemplo, evaluación de un laboratorio para poder llevar a cabo la datación por el procedimiento del carbono 14, C_{14} . Para ello se comparan los resultados obtenidos por el laboratorio que pretende certificarse en este procedimiento con los obtenidos por aquellos laboratorios de referencia ya certificados.
- Evaluación de procedimientos o métodos estándar, es decir, aquellos procedimientos experimentales que serán incluidos en normas de referencia, ya sean ISO (internacionales), DIN (alemanas), ASTM (americanas), etc.
- ILS para la certificación de las propiedades de materiales de referencia. Por ejemplo, para estimar la temperatura de transición vítrea de un nuevo material compuesto que se pretende introducir en la construcción de aeronáutica.

Una de las tareas más importantes dentro de un estudio ILS es la evaluación de la consistencia de los resultados obtenidos por los laboratorios participantes. En otras palabras, detectar aquellos laboratorios atípicos. La detección de estos laboratorios es sumamente importante para poder estimar de forma correcta la precisión de un método de medida, el valor de una propiedad física de un material de referencia o comparar si, teniendo en cuenta la variabilidad debida a la repetitividad y reproducibilidad, los resultados de un laboratorio son comparables a los de referencia. La detección de laboratorios atípicos siempre se ha llevado a cabo mediante contrastes de hipótesis univariantes como son el test de Cochran, el de Grubbs, Youden o, los contrastes a partir de los estadísticos h y k de Mandel (Hund *et al.*, 2000a; Mandel, 1985), dos de los más importantes debido a que proporcionan también útiles herramientas gráficas. El uso de estos estadísticos está normalizado según diferentes normativas de referencia como son la ASTM E691 y la ISO 5725 (ISO-5725, 1994).

Una parte importante de las aportaciones de la presente memoria de tesis se ocupa del desarrollo y aplicación de extensiones funcionales o FDA para los estadísticos escalares h y k de Mandel, descritas en el Capítulo 3.

1.3.1. Pruebas de consistencia

El modelo estadístico básico, propuesto en la norma ISO-5725 (1994), y que permite la estimación de la exactitud y la precisión de un método analítico es:

$$y = \mu + B + \epsilon \quad (1.1)$$

Donde μ es la media global para el material analizado, $B \sim N(0, \sigma_L^2)$ es la componente de sesgo de los laboratorios asumiendo condiciones de repetibilidad, y $\epsilon \sim N(0, \sigma_r^2)$ el error aleatorio que ocurre en cada medida en condiciones de repetibilidad. La varianza de reproducibilidad σ_R^2 viene dada por:

$$\sigma_R^2 = \sigma_L^2 + \sigma_r^2. \quad (1.2)$$

En la Tabla 1.1, se muestra el análisis de la varianza de efectos aleatorios de un solo factor (ANOVA), donde x_{lj} es el resultado j -ésimo del laboratorio l , \bar{x}_l es la media de los n_l resultados del laboratorio l , \bar{x} es el promedio global de los resultados de los p laboratorios,

$$\bar{n} = \frac{1}{p-1} \left(N - \frac{\sum_{l=1}^p n_l^2}{N} \right),$$

y $N = \sum_{l=1}^p n_l$.

La varianza de repetibilidad σ_r^2 se estima mediante S_r^2 , que es la varianza dentro del laboratorio. Por otro lado, la varianza entre laboratorios σ_L^2 , relacionada con el sesgo de laboratorio, se estima usando S_L^2 :

$$S_L^2 = \frac{MS_L - S_r^2}{\bar{n}},$$

cuando $MS_L \geq S_r^2$ ($S_L^2 = 0$ de lo contrario).

Tabla 1.1: ANOVA para la estimación de σ_r^2 y σ_L^2 .

Source	Media cuadrática	Estimación de
Laboratorio	$MS_L = \frac{\sum_{l=1}^p n_l (\bar{x}_l - \bar{x})^2}{p-1}$	$\sigma_r^2 + \bar{n}\sigma_L^2$
Residuo (repetibilidad)	$S_r^2 = \frac{\sum_{l=1}^p \sum_{j=1}^{n_l} (x_{lj} - \bar{x}_l)^2}{N-p}$	σ_r^2

Antes de estimar σ_r^2 y σ_L^2 , primero se deberían excluir del ILS los laboratorios que presenten resultados no consistentes, para esto se deben realizar pruebas de consistencia e identificación de resultados atípicos (ISO-5725, 1994).

Existen dos posibles tipos de escenarios a evaluar para determinar la presencia de laboratorios atípicos o que proporcionan resultados no consistentes. El primero es que los resultados de algún laboratorio se desvíen de los demás en términos de precisión, es decir, cuando las medidas realizadas por un laboratorio se desvían significativamente con respecto a las obtenidas por otros laboratorios. El segundo, tiene que ver con la identificación de resultados atípicos en un laboratorio. A continuación se describen los estadísticos y las pruebas recomendadas por la norma ISO 5725 (ISO-5725, 1994) y la norma ASTM E-691 (ASTM-E691, 2004).

Estadístico h de Mandel y prueba de Grubbs

Sea $\{x_1, x_2, \dots, x_p\}$ una muestra de p observaciones. Los x_l , $l = 1, \dots, p$, se modelan como realizaciones de variables aleatorias X_l , $l = 1, \dots, p$, distribuyéndose de forma idéntica e independiente según la distribución normal $N(\mu, \sigma^2)$. Denotamos:

$$\bar{X} = \frac{\sum_{l=1}^p X_l}{p},$$

como la media de los X_l , y:

$$S^2 = \frac{1}{p-1} \sum_{l=1}^p (X_l - \bar{X})^2,$$

como la varianza muestral de X_l .

El estadístico h de Mandel (Mandel, 1985) se define como:

$$H_l = \frac{X_l - \bar{X}}{S}, \quad l = 1, \dots, p, \quad (1.3)$$

que tiene la misma distribución para todo $l = 1, \dots, L$. El valor crítico es (Wilrich, 2013):

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

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

Para el caso de p laboratorios que obtienen n réplicas para cada uno, el estadístico h se obtiene como:

$$h_l = \frac{\bar{x}_l - \bar{x}}{S_B}, \quad l = 1, \dots, p,$$

donde $S_B = \sqrt{\sum_{l=1}^p \frac{(\bar{x}_l - \bar{x})^2}{p-1}}$, \bar{x}_l es la media de los n resultados del laboratorio l , y \bar{x} es la media global de los resultados de los p laboratorios. Un laboratorio es detectado como inconsistente cuando el valor observado del estadístico h_l es mayor que el valor crítico, es decir, cuando $h_l \geq h_{p;1-\frac{\alpha}{2}}$.

Por otro lado, si se desea determinar si la observación $X_{\text{máx}} = \max\{X_1, \dots, X_p\}$ es una observación atípica, se puede utilizar el test de Grubbs (Grubbs y Beck, 1972). El estadístico correspondiente para esta prueba se define mediante la siguiente expresión:

$$G_{\text{máx}} = \max_{l=1, \dots, p} H_l = \max_{l=1, \dots, p} \frac{X_l - \bar{X}}{S} = \frac{X_{\text{máx}} - \bar{X}}{S}.$$

Si se desea determinar si la observación $X_{\text{mín}} = \min\{X_1, \dots, X_p\}$ es un atípico, el estadístico de prueba es:

$$G_{\text{mín}} = \max_{l=1, \dots, p} \{-H_l\} = \max_{l=1, \dots, p} \frac{\bar{X} - X_l}{S} = \frac{\bar{X} - X_{\text{mín}}}{S}.$$

El valor crítico para esta prueba puede ser aproximado (Wilrich, 2013) por:

$$g_{p;1-\alpha} \approx \frac{(p-1)t_{p-2;1-\frac{\alpha}{p}}}{\sqrt{p\left(p-2+t_{p-2;1-\frac{\alpha}{p}}^2\right)}}.$$

Para el caso especial en el que se obtienen n réplicas por cada uno de los p laboratorios, el estadístico de Grubbs se define análogamente. Las observaciones deben ser reemplazadas por la media de los resultados correspondientes a cada laboratorio, mientras que la media de las observaciones también se reemplaza por la media global obtenida como promedio de las medias de los laboratorios.

Si un laboratorio se identifica como un valor atípico, después de aplicar el estadístico h y el contraste de Grubbs a diferentes niveles dentro de un laboratorio,

esto es una evidencia de la presencia de un alto sesgo en el laboratorio (debido a un error sistemático en la calibración del aparato de medida, o errores en las ecuaciones utilizadas para calcular los resultados, entre otras causas).

Estadístico k de Mandel y prueba de Cochran

Asúmase que los valores observados x_{lj} , $l = 1, \dots, p$, $j = 1, \dots, n$, son realizaciones de variables aleatorias X_{lj} independientes e idénticamente distribuidas según una distribución normal $N(\mu_j, \sigma^2)$ para cada réplica j y laboratorio l , y denótese por S_l^2 la varianza muestral de las réplicas del laboratorio l para un material dado:

$$S_l^2 = \sum_{j=1}^n \frac{(X_{lj} - \bar{X}_l)^2}{n-1}.$$

Entonces $(n-1)S_l^2/\sigma^2$, $l = 1, \dots, p$, sigue una distribución χ_v^2 con $v = n-1$ grados de libertad. El estadístico k de Mandel (Mandel, 1985) se define como:

$$K_l = \frac{S_l}{S_r}, \quad l = 1, \dots, p,$$

donde

$$S_r^2 = \frac{\sum_{l=1}^p S_l^2}{p},$$

con la misma distribución para todo $l = 1, \dots, p$. El valor crítico (Wilrich, 2013) correspondiente es:

$$k_{p,n;1-\alpha} = \sqrt{\frac{p}{1 + \frac{p-1}{F_{v_1, v_2; \alpha}}}},$$

donde $F_{v_1, v_2; \alpha}$ es el cuantil correspondiente a un nivel de significación α de la distribución F con $v_1 = (p-1)(n-1)$ y $v_2 = n-1$ grados de libertad. Un laboratorio es detectado como inconsistente cuando el valor de la observación del estadístico k_l es mayor que el valor crítico, esto es, $k_l \geq k_{p,n;1-\alpha}$.

Por otro lado, para determinar si $S_{max}^2 = \max \{S_1^2, S_2^2, \dots, S_p^2\}$ es un atípico, se utiliza la prueba de Cochran:

$$C = \frac{S_{max}^2}{\sum_{l=1}^p S_l^2} = \frac{1}{p} \max_{l=1, \dots, p} \frac{S_l^2}{S_r^2}.$$

El valor crítico de esta prueba es aproximado (Wilrich, 2013) por:

$$c_{p,n;1-\alpha} \approx \frac{1}{1 + (p-1)F_{v_1, v_2; \frac{\alpha}{p}}}.$$

El test de Cochran solo evalúa el valor más alto en una serie de varianzas. Si se detecta un laboratorio como un valor atípico, usando el estadístico k o con la prueba de Cochran, esto indica que la varianza dentro del laboratorio es alta (esto puede ser debido a la falta de familiaridad con el método analítico, diferencias de apreciación entre operadores, equipo inadecuado, equipo en mal estado o ejecución descuidada), en cuyo caso, el total de los resultados obtenidos por este laboratorio deben rechazarse y retirarse del estudio.

La detección de laboratorios inconsistentes debe repetirse hasta que dejen de detectarse valores atípicos. Sin embargo, las pruebas de consistencia deben usarse con precaución, ya que si este proceso se lleva a cabo en exceso, podría conducir a una identificación falsa de valores atípicos.

Es también importante destacar que en el Capítulo 3 de esta tesis se proponen nuevas extensiones funcionales para los estadísticos h y k . De hecho, una parte importante de las técnicas experimentales en química analítica y física aplicada dan como resultado curvas definidas en el tiempo o la frecuencia.

1.4. Gráficos de control

Los gráficos de control representan la herramienta más importante en el análisis de las variaciones de los procesos de producción o servicios. Teniendo en cuenta que en esta tesis se han adoptado los gráficos de control no paramétricos de rangos al caso particular del control y monitorización de datos funcionales, es necesario introducir previamente las principales características del este tipo de herramientas.

Un gráfico de control es un gráfico de dos dimensiones cuyo eje de las ordenadas (Y) representa un estadístico estimado a partir de las observaciones de la variable (o estadístico en general) que estamos monitorizando, denominada variable CTQ (Critical to Quality). Los valores de la característica a medir se representan secuencialmente con respecto al tiempo. Los valores de la variable pueden ser individuales

o, más comúnmente, los promedios referidos a grupos de tamaño 4, 5, etc., llamados muestras racionales. El eje de las abscisas (X) de la gráfica muestra los números de identificación para el conjunto de ítems (individuales o grupos). Los valores de la variable se representan con puntos unidos mediante líneas rectas, para poder identificar patrones indicadores de cambios significativos en el rendimiento del proceso. Por ejemplo, se tiene los gráficos de tipo Shewhart en honor a su creador, Walter Shewhart, que los empleó por vez primera en sus trabajos en la Bell Telephone. Gráficamente, se caracterizan por:

- Una línea central (CL). Media teórica de las observaciones de la muestra. Los valores monitorizados varían en torno a esta media.
- Límite de control inferior (LCL), por debajo del cual es muy poco probable que se produzcan realizaciones de la variable.
- Límite de control superior (UCL). Los LCL y UCL son simétricos respecto al CL si la distribución de probabilidad de la variable es simétrica (se suele utilizar la normal).

Utilidades de los gráficos de control:

- Evitar que el proceso esté fuera de control, detectando las causas asignables a cada variación de la característica medida y tomando las medidas al respecto.
- Para no hacer ajustes cuando no se necesitan. La mayoría de los procesos de producción dan a los operadores margen de maniobra para hacer ajustes en los equipos que están utilizando. Los gráficos de control pueden indicar cuando los ajustes son necesarios y cuando no lo son.
- Para determinar el rango natural (límites de control) de un proceso y para comparar este rango con sus límites especificados. Si el rango de los límites de control es más amplio que el de los límites especificados, el proceso va a generar productos defectuosos y tendrá que ser ajustado.
- Dar a conocer la capacidad y estabilidad del proceso, entendiendo como capacidad de proceso a su adecuación para ofrecer productos dentro de los límites especificados continuamente en el tiempo.

- Para llevar a cabo el proceso de monitorizado del proceso y así evitar los defectos en el producto final.
- Para facilitar la planificación de la asignación de recursos de producción. Las variaciones de una característica de calidad determinan la cantidad de defectos. Tener información para predecir estas variaciones ayuda a asignar los recursos.

El funcionamiento de un gráfico de control puede relacionarse con las ideas que sustentan los contrastes de hipótesis. Como es conocido una hipótesis estadística es una conjetura sobre una o varias características de interés.

En el caso del control de calidad la hipótesis nula sería suponer que el proceso está bajo control en cada una de las muestra seleccionadas. Rechazar la hipótesis nula implica asumir como correcta una hipótesis complementaria, que en el contexto de los contrastes se denomina hipótesis alternativa y que suele denotarse por H_1 . Un punto que esté fuera de los límites de control puede interpretarse como un incumplimiento de la hipótesis nula y una aceptación de la alternativa, es decir, dar por contrastado que el proceso ha salido de control.

En un contraste de hipótesis, a la decisión de rechazar la hipótesis nula cuando ésta es cierta se la denomina error de tipo I, en este contexto de control es más habitual denominarlo riesgo del vendedor, mientras que al error cometido al no rechazar la hipótesis nula cuando ésta es falsa se la denomina riesgo del comprador (error de tipo II en el caso de los contrastes estadísticos). Las cuatro posibles situaciones son:

		Situación real:	
		H ₀ es cierta	H ₀ es falsa
No se Rechaza H ₀	DECISIÓN CORRECTA	Error tipo II o riesgo del comprador	
Se Rechaza H ₀	Error tipo I o Riesgo del vendedor	DECISIÓN CORRECTA	

La construcción de un gráfico de control es equivalente al trazado de las regiones de aceptación de una secuencia de pruebas de hipótesis a lo largo del tiempo. Por ejemplo, el gráfico \bar{x} es un gráfico de control que se usa para supervisar el proceso con la media μ . Traza las medias muestrales, \bar{X} , de los subgrupos de las observaciones $\{X_1, X_2, \dots\}$ y es equivalente a probar las hipótesis $H_0 : \mu = \mu_0$ versus $H_\alpha : \mu \neq \mu_0$ (para un valor objetivo μ_0) llevado a cabo a lo largo del tiempo, usando \bar{X} como el

estadístico de prueba. Aquí se asume que $\{X_1, X_2, \dots\}$ son las medidas de muestra de una característica de calidad particular siguiendo una distribución F con media μ y desviación estándar σ . Cuando no hay pruebas suficientes para rechazar H_0 , se dice que el proceso está bajo control; de lo contrario, se dice que está fuera de control. La decisión de aceptar o rechazar H_0 se basa en el valor de la media de la muestra \bar{x} observada en cada intervalo de tiempo (Liu y Tang, 1996).

Como ya se ha comentado, el gráfico contiene una línea central que representa el valor promedio de la característica de calidad correspondiente al estado en control y dos líneas horizontales, llamadas límite de control superior (UCL) y límite de control inferior (LCL). La región entre los límites de control es la región de aceptación de H_0 . Como consecuencia, el proceso se considera fuera de control, cuando una media de subgrupos observada cae fuera de los límites, esto sugiere que el proceso puede haber sido afectado por algunas causas especiales. Incluso si todos los puntos se trazan dentro de los límites de control, si se comportan de forma sistemática y no aleatoria, esto podría ser una indicación de que el proceso está fuera de control. Los métodos para buscar secuencias o patrones no aleatorios se pueden aplicar a gráficos de control como una ayuda para detectar condiciones fuera de control (Montgomery, 2009). Con los límites de control elegidos correctamente, una tabla de control puede detectar un cambio de una distribución de buena calidad a una mala.

Sea w un estadístico muestral que mida alguna característica de calidad de interés, y supongamos que la media de w es μ_w y la desviación estándar de w es σ_w . Luego, la línea central, el límite superior de control y el límite inferior de control se convierten en:

$$\text{UCL} = \mu_w + T\sigma_w,$$

$$\text{CL} = \mu_w,$$

$$\text{LCL} = \mu_w - T\sigma_w,$$

donde T es la “distancia” de los límites de control desde la línea central, expresada en unidades de desviación estándar.

Al diseñar un gráfico de control, se debe especificar el tamaño de la muestra y la frecuencia de muestreo, pequeños cambios en el proceso serán más fáciles de detectar con muestras grandes. Tomar muestras grandes con frecuencia sería la situación

deseable, pero no es factible desde un punto de vista económico. El uso del ARL (Average Run Length) es una forma de evaluar la elección del tamaño de muestra y la frecuencia de muestreo. El ARL es el número de puntos que, en promedio, se trazarán en una tabla de control antes de que se indique una condición fuera de control. Si las observaciones en el proceso no están correlacionadas, el ARL se puede calcular para cualquier control Shewhart como

$$ARL = \frac{1}{p},$$

donde p es la probabilidad de que cualquier punto supere los límites de control. El rendimiento de los gráficos de control se puede evaluar utilizando esta ecuación, aunque a veces es conveniente expresar este rendimiento en términos de ATS (Average Time to Signal), que se define como $ATS = ARL \times h$ donde h es el intervalo de tiempo fijo en horas, en el que se toman muestras.

En los contrastes de hipótesis se mide la potencia del test mediante la curva de potencia, sin embargo, en el caso de los gráficos de control también se utiliza la inversa de esta curva de potencia que se denomina curva característica de operación (curva OC de las siglas Operating Characteristics) que representa visualizar en un gráfico el error de tipo II (también llamado riesgo β) en función de distintos valores para la hipótesis alternativa (Montgomery, 2009). Es frecuente que dicha curva se construya representando este error en función de la magnitud del cambio que se pretende estudiar (normalmente expresado en unidades de desviación estándar).

En el diseño de gráficos de control una de las decisiones importantes es la especificación del tamaño muestral y la frecuencia de muestreo. En general, muestras grandes facilitan detectar cambios pequeños pero cuando existe poca variabilidad en el proceso optar por muestras pequeñas puede ser un acierto. Para la selección del tamaño de cada muestra racional o submuestra suele hacerse uso de las curvas OC. Si se representan para distintos tamaños de muestra, estas curvas pueden ser un instrumento eficaz para analizar el poder del gráfico.

Por otra parte la frecuencia de muestreo es otro factor importante a tener en cuenta, lo habitual es considerar cuestiones económicas y del propio proceso para encontrar un equilibrio entre el tamaño de las muestras y la frecuencia.

Se suelen emplear distintos procedimientos para la construcción de los subgru-

pos racionales. El más frecuente se aplica cuando el objetivo del gráfico de control es detectar cambios pequeños en el proceso. En este caso, se buscará minimizar la variabilidad dentro de cada muestra al tiempo que se maximiza entre muestras distintas. Este primer enfoque proporcionará buenas estimaciones de la variabilidad en el caso del control por variables.

1.5. Retos actuales del Control Estadístico de la Calidad

En esta sección se muestran los escenarios o situaciones actuales en las que el análisis estadístico de datos y, en particular, el control estadístico de la calidad, se tienen que desenvolver y aportar soluciones. Estos nuevos escenarios han sido originados por la toma de datos masivos o Big Data, en gran medida generados por el reciente desarrollo de la tecnología IIOT (Stankovic, 2014). Uno de estos escenarios es el correspondiente a procesos definidos por datos funcionales, también denominados perfiles (por ser la relación entre dos variables la que define la calidad del proceso) en el ámbito del control estadístico de calidad. De hecho, en el Capítulo 4 de esta tesis doctoral, se propone una metodología de gráficos de control para las Fases I y II que, basados en los conceptos de profundidad de datos y gráficos de control de rangos, son aplicables para controlar la calidad de procesos definidos por datos funcionales.

1.5.1. Escenarios de monitorización

En la actualidad, ya sea en la industria, administración, finanzas, servicios, I+D+I y sociedad en general, la adquisición de datos cada vez más complejos es una tendencia, por lo que el desarrollo y aplicación de técnicas estadísticas específicas es una necesidad cada vez más acuciante. El Big Data no sólo obedece al tamaño, sino a la complejidad de los datos, a la velocidad a la que se toman, número y tipo de variables diferentes, entre otras características. De forma general, se puede definir el Big Data como conjuntos de datos grandes, diversos, complejos y/o longitudinales generados a partir de una amplia gama de instrumentos de medida y sensores

(Megahed y Jones-Farmer, 2013). Según aumenta el tamaño, velocidad de recogida y variedad en las bases de datos, los requerimientos de nuevas técnicas de análisis aumentan, siendo el control estadístico de procesos una de las ramas de la estadística afectadas y que más puede aportar a este nuevo paradigma (Woodall y Montgomery, 2014).

Los principales desarrollos dentro del campo del control estadístico de la calidad en relación al nuevo tipo de datos, están relacionados con la monitorización espacio-temporal, monitorizado de perfiles, uso de datos autocorrelacionados y el monitoreo de datos de alta dimensión. Se requieren herramientas aplicables a datos tomados a lo largo del tiempo, que sean capaces de detectar rápidamente cambios en el proceso especificado debido a causas asignables. Debido a estas características, muy a menudo las bases de datos, aparte de ser complejos conjuntos multivariantes, no cumplirán las hipótesis básicas para poder aplicar gráficos de control estándar de forma fiable. En resumen, se pretende dar respuesta a los siguientes retos:

1. Proponer alternativas cuando se incumplen las hipótesis estadísticas básicas de los gráficos de control tradicionales (Shewhart, CUSUM, EWMA).
2. Desarrollar técnicas estadísticas que permitan la reducción eficiente de la dimensionalidad y que exploten la estructura de correlación poblacional.

1.5.2. Monitorizado de bases de datos de alta dimensión

En los últimos años, la cantidad de variables monitorizadas y su complejidad han crecido, y seguirá creciendo, exponencialmente (Jensen *et al.*, 2006). A estos nuevos “data streams” se los suele denominar también flujos de datos de alta dimensión (Woodall y Montgomery, 2014). Hasta hace bien poco, el análisis de estos data streams se hacía bajo la hipótesis de la independencia dentro y fuera de este flujo de datos, por lo que se detectaban cambios en un solo flujo o cambios simultáneos en todos ellos, pero últimamente se está trabajando en el monitoreo de múltiples data streams teniendo en cuenta la autocorrelación existente y los efectos estacionales (Epprecht, 2015). Por otro lado, en este tipo de datos, la probabilidad de falsa alarma no se puede estudiar por ARL debido a que, con múltiples flujos, fácilmente se puede detectar un estado fuera de control a cada instante. Una alternativa fiable

en la que se está trabajando es el False Discovery Rate (FDR), mediante el que se controla la proporción de señales que se corresponden con falsas alarmas. De hecho, se han desarrollado ya varias alternativas de gráficos de control que implementan esta técnica (Spiegelhalter *et al.*, 2012) y se esperan más en el futuro. La principal característica es que se convierten los estadísticos del gráfico de control en p -valores (Li *et al.*, 2013). Estos flujos de datos acontecen hoy en día en una gran variedad de campos, desde el financiero a la salud pública, pasando por los datos originados por la Industria 4.0. Dada su alta dimensionalidad, actualmente se propone la aplicación de gráficos de control multivariantes como es el T^2 de Hotelling, técnicas de clasificación supervisada, de regresión multivariante y de series de tiempo multivariadas combinadas a su vez con técnicas de reducción de dimensión como son el análisis de componentes principales, análisis factorial y los mínimos cuadrados parciales (Bisgaard, 2012; Mei, 2010; Woodall y Montgomery, 2014). A modo de conclusión, la monitorización de grandes flujos de datos o bases de datos de alta dimensión requiere todavía una labor de investigación intensiva que permita resolver el problema de detección de anomalías en tiempo real, desde la perspectiva de clasificación.

También dentro del análisis multivariante, se pueden encontrar otras dos líneas de intensa investigación en SPC. Estas son el control espacio-temporal y el control de perfiles. En lo que respecta al control espacio temporal, se añade la novedad en el SPC de tener en cuenta las variaciones espaciales en el proceso. Este análisis es fundamental, por ejemplo, para el control de piezas medidas por escáner de láser en la industria manufacturera, naval, etc. Aunque es un campo de investigación incipiente, se pueden encontrar interesantes aportaciones en Wells *et al.* (2012), en el que se propone una metodología para el control de calidad o monitorización de procesos mediante imágenes, siendo ésta una muy interesante alternativa en un área de vital desarrollo; de hecho, como indican Woodall y Montgomery (2014), todavía no existen gráficos de control para imágenes propiamente dichos. Según indica Megahed y Jones-Farmer (2013), el control por imágenes puede considerarse una extensión del control de perfiles y, por tanto, afrontarse a partir de la perspectiva del análisis de datos funcionales.

Por otro lado, está el cada vez más importante campo del control de perfiles, que puede afrontarse desde la perspectiva del análisis de datos funcionales (Megahed y

Jones-Farmer, 2013; Woodall y Montgomery, 2014). La monitorización de perfiles es un conjunto de técnicas estadísticas aplicadas cuando la calidad de un proceso se mide por la relación funcional entre dos variables (Megahed y Jones-Farmer, 2013; Woodall y Montgomery, 2014). Suelen involucrar el ajuste de modelos de regresión (lineales, no lineales, no paramétricos, semiparamétricos) y la monitorización, mediante gráficos T^2 de Hotelling u otras alternativas multivariantes de los parámetros de ajuste. Aparte de las referencias relacionadas con el control de imágenes, pueden nombrarse algunas aplicaciones interesantes como las incluidas en Noorossana et al (Noorossana *et al.*, 2011) y el desarrollo de gráficos de control no paramétricos para perfiles cuando la relación entre variables no se puede ajustar por modelos de regresión lineal o no lineal (Qiu *et al.*, 2010).

Capítulo 2

Técnicas estadísticas FDA y casos de estudio reales

En este capítulo, se realiza una breve introducción del Análisis de Datos Funcionales (FDA) con la finalidad de contar con un marco teórico que permita tener un punto de referencia para el desarrollo de nuevas metodologías para resolver dos problemas concretos: estudios interlaboratorio y monitorización de variables mediante gráficos de control cuando estas tienen un carácter funcional.

Así también, en la segunda sección de este capítulo, se presentan los casos de estudio, donde se aplican las metodologías desarrolladas (Flores *et al.*, 2018b) y que han sido implementadas en las librerías del software estadístico R ILS (Flores *et al.*, 2018a) y `qcr`, que describen problemas reales relacionados con la Metrología 4.0 y la Industria 4.0.

2.1. Análisis de datos funcionales

En esta sección, en primer lugar, se dan algunas definiciones importantes en el contexto FDA, tales como las correspondientes a variable funcional (Ferraty y Vieu, 2006) y, dentro del análisis exploratorio funcional, las medidas funcionales de tendencia central y variabilidad (Ramsay y Silverman, 2005). El concepto de profundidad de datos también es definido en términos de cómo un dato está centrado con respecto a un conjunto de puntos que pertenecen a una determinada población.

En el contexto del FDA, la curva más profunda es identificada como la más cercana al centro, y es utilizada como una estimación de la mediana funcional (Fraiman y Muniz, 2001). Tres de los más comunes enfoques para calcular la profundidad funcional son la profundidad de Fraiman y Muniz o profundidad mediana (Fraiman y Muniz, 2001), la profundidad modal (Cuevas *et al.*, 2006), y la profundidad de proyecciones aleatorias (Cuevas *et al.*, 2007).

Por otro lado, también se explican técnicas de estadística inferencial para contrastar diferencias entre grupos y detección de datos atípicos. El uso del concepto de profundidad de datos es utilizado en Febrero-Bande *et al.* (2007) para proponer un método de detección de atípicos. Otro método que se explica es el Análisis de la Varianza Funcional (FANOVA), que tiene como objetivo verificar si alguna variable continua (por ejemplo, temperatura, pérdida de masa, módulo, etc.) es significativamente diferente dependiendo del valor de uno o más factores, variables definidas en algunos niveles específicos (por ejemplo, cantidad de sílice en tres niveles diferentes, dos operadores diferentes, cuatro velocidades de calentamiento diferentes, etc.).

La propuesta de un ANOVA con respuesta funcional y una variable explicativa cualitativa (grupos) fue presentado en Cuevas *et al.* (2004) y para el caso de más de un factor, utilizando un procedimiento de proyección aleatoria, se propuso un Anova multifactor en Cuesta-Albertos y Febrero-Bande (2010). Algunas aplicaciones y otros enfoques se puede encontrar en Brumback y Rice (1998), Ramsay y Silverman (2002, 2005), Ramsay y Hooker (2009), Fan y Lin (1998), Shen y Faraway (2004).

El uso de estas técnicas estadísticas FDA se ha facilitado por el desarrollo de diversos paquetes implementados en el software estadístico R tales como `fda` y `fda.usc`. Este hecho ha facilitado su uso y la generalización de estas técnicas.

2.1.1. Definiciones

Ferraty y Vieu (2006) definen una variable aleatoria funcional como una variable aleatoria que toma valores en un espacio de funciones, es decir, un espacio infinito dimensional (espacio funcional). A continuación, se muestran, entre otras, la definición del espacio funcional, en el cual se representan los datos funcionales correspondientes a los casos de estudio reales mostrados en esta tesis doctoral.

Definición 1: Sea $L^2(T)$, el espacio de Hilbert separable definido por las funcio-

nes de cuadrado integrable en el intervalo $T = [a, b] \subset \mathbb{R}$:

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

con un producto interno definido por:

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

Definición 2: Una variable aleatoria \mathbf{X} es una variable funcional si toma valores en $L^2(T)$.

Definición 3: Un conjunto de datos funcionales $S_n = \{\mathcal{X}_i\}_{i=1}^n$ (también denotado como $\mathcal{X}_1, \dots, \mathcal{X}_n$) es la observación de n variables funcionales $\mathbf{X}_1, \dots, \mathbf{X}_n$ con igual distribución que \mathbf{X} .

Un dato funcional es la observación de la variable aleatoria a lo largo de un intervalo (comúnmente de tiempo, también de frecuencia) continuo fijo. En la práctica, se supone que estos valores observados son discretizaciones.

2.1.2. Análisis exploratorio

En Ramsay y Silverman (2005) se consideran las siguientes definiciones de las funciones muestrales descriptivas de la media y varianza calculadas a partir de un conjunto de datos funcionales $S_n = \{\mathcal{X}_1, \dots, \mathcal{X}_n\}$, definidos en $t \in T \subset \mathbb{R}$. La media funcional muestral se define como:

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

Al igual que se pueden definir medidas de centralidad para datos funcionales punto a punto, para la estimación de la variabilidad se cuenta con la varianza y la desviación estándar:

- Varianza: $s^2(t) = (n-1)^{-1} \sum_{i=1}^n (\mathcal{X}_i(t) - \bar{\mathcal{X}}(t))^2$
- Desviación estándar: $s(t) = \sqrt{s^2(t)}$

En Febrero-Bande *et al.* (2007), se menciona que la media funcional muestral es un estimador del centro de la distribución funcional (media funcional poblacional,

que es la trayectoria que minimiza la suma de las distancias a todos los puntos). Además de la definición anterior de la media muestral punto a punto, $\bar{\mathcal{X}}$, existe otra definición en la que se utiliza una base $\{\phi_k\}_{k \in \mathbb{N}}$ y una función de distancia d en L^2 . De manera general, los estimadores funcionales media, mediana y varianza se define, respectivamente, como:

- Media: $\operatorname{argmax}_{a \in \mathcal{S}_n} \sum_{i=1}^n d(\mathcal{X}_i, a)^2$
- Mediana: $\operatorname{argmax}_{a \in \mathcal{S}_n} \sum_{i=1}^n d(\mathcal{X}_i, a)$
- Varianza: $\frac{1}{n} \sum_{i=1}^n d(\mathcal{X}_i, a)^2$

En el caso que se necesite describir otras características de los datos como son la forma, suavidad y rugosidad, que es parte del análisis exploratorio, es necesario utilizar una medida de proximidad. Como se menciona en Ferraty y Vieu (2006), en algunos casos, cuando nos encontramos en un espacio de dimensión infinita, la mejor elección para una medida de proximidad es una semi-métrica.

Según comentan estos autores, considerar la distancia entre derivadas es una de las alternativas para medir la proximidad entre curvas. Siendo $\{\mathcal{X}_i(t)\}_{i=1}^n$ realizaciones i.i.d. de la variable aleatoria funcional $X(t)$, y \mathcal{X}_i y $\mathcal{X}_{i'}$ dos curvas observadas, esta familia de semi-métricas se puede parametrizar como se muestra en la siguiente expresión:

$$d_q^{deriv}(\mathcal{X}_i, \mathcal{X}_{i'}) = \sqrt{\int \left(\mathcal{X}_i^{(q)}(t) - \mathcal{X}_{i'}^{(q)}(t) \right)^2 dt} \quad (2.1)$$

donde, $\mathcal{X}_i^{(q)}$ representa la q -ésima derivada de \mathcal{X}

A continuación, se definen las tres funciones de profundidad funcionales más usadas (Modal, Fraiman y Muniz y Proyección Aleatoria) y su uso en la definición de medidas robustas de centralidad y variabilidad.

Medidas de profundidad

Las medidas de profundidad se introdujeron originalmente en el análisis de datos multivariados para medir la centralidad de un punto $X \in \mathbb{R}^d$, con respecto a una nube de datos generada a partir de una distribución de probabilidad F en \mathbb{R}^d .

Las profundidades proporcionan una manera de ordenar puntos en el espacio Euclídeo desde el centro hacia el exterior, de tal manera que los puntos más cercanos al centro tienen una mayor profundidad que los puntos localizados en los extremos que tienen una menor profundidad. Entre las aplicaciones directas de tener este orden están la definición de medidas de ubicación y la detección de valores atípicos.

Entre algunos ejemplos bien conocidos de medidas de profundidad se tienen la de Tukey (Tukey, 1975) y la profundidad simplicial (Liu, 1990). Véase Zuo y Serfling (2000) para un extenso análisis de las definiciones, propiedades y aplicaciones de las profundidades multivariadas. En Fraiman y Muniz (2001) y en Cuevas *et al.* (2006, 2007) se ha extendido la definición de profundidad, al caso funcional.

Una de las ideas básicas del cálculo de la profundidad de datos funcionales, es medir la centralidad de una curva dada \mathcal{X}_i dentro de las curvas $\mathcal{X}_1, \dots, \mathcal{X}_n$, generadas a partir del proceso estocástico S que toma valores en $L^2([a, b])$, el espacio de funciones continuas cuadrado integrables definidas en el intervalo $[a, b] \subset \mathbb{R}$. En esta subsección se realiza una breve revisión de las profundidades funcionales que se utilizan en esta monografía.

La profundidad de Fraiman y Muniz (Fraiman y Muniz, 2001), es una de las primeras medidas de profundidad introducida para datos funcionales.

Definición 4: Sea $\mathcal{S}_n = \{\mathcal{X}_i(t)\}_{i=1}^n$ i.i.d., realizaciones de la variable aleatoria funcional $\mathcal{X}(t)$ con dominio $T = [a, b]$ y sea D una medida de profundidad en \mathbb{R} . Para cada $t_0 \in T$, se considera $z_i(t_0) = D(\mathcal{X}_i(t_0))$ la profundidad univariante del dato i en t_0 con respecto a $\{\mathcal{X}_i(t_0)\}_{i=1}^n$. Entonces se define la profundidad de FM para el i -ésimo dato como:

$$FM_i = FMD(\mathcal{X}_i) = \int_a^b z_i(t) dt$$

En el presente trabajo se utilizara la siguiente medida de profundidad univariante (Fraiman y Muniz, 2001):

$$FM1 : z_i^{FM1}(t) = 1 - \left| \frac{1}{2} - F_{n,t}(\mathcal{X}_i(t)) \right| \quad (2.2)$$

con la correspondiente función de distribución empírica:

$$F_{n,t}(\mathcal{X}_i(t)) = \frac{1}{n} \sum_{k=1}^n I(\mathcal{X}_k(t) \leq \mathcal{X}_i(t)),$$

donde $I(\cdot)$ es la función indicadora. Entonces, la profundidad funcional de Fraiman y Muniz de una curva \mathcal{X} con respecto al conjunto $\mathcal{X}_1, \dots, \mathcal{X}_n$ está dada por:

$$FMD(\mathcal{X}_i) = \int_a^b z_i^{FM1}(t) dt$$

Definición 5: Sea $\mathcal{S}_n = \{\mathcal{X}_i\}_{i=1}^n$ i.id. realizaciones de la variable aleatoria X y $K : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ una función kernel asimétrica y h , un parámetro ventana. Entonces, la profundidad modal funcional MD para el i -ésimo dato, se define como (Cuevas *et al.*, 2006):

$$MD_i = MD(\mathcal{X}_i) = \sum_{j=1}^n K\left(\frac{\|\mathcal{X}_i - \mathcal{X}_j\|}{h}\right)$$

La profundidad modal es una medida de cuántos datos hay en un vecindario de la curva y se asemeja al estimador densidad del kernel no paramétrico pero con la principal diferencia de que el parámetro ventana no necesariamente tiende a cero a medida que n aumenta. El objetivo de esta medida de profundidad es proporcionar rangos que no cambien para anchos de banda lo suficientemente grandes.

Desde el punto de vista práctico, hay que elegir una norma funcional, una función núcleo y un h . En Cuevas *et al.* (2006), se recomienda utilizar la norma L^2 .

Además, se recomienda un núcleo asimétrico de Gauss:

$$K(t) = \frac{2}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right), t > 0$$

El parámetro h recomendado es el décimo quinto percentil de la distribución empírica de $\|\mathcal{X}_i - \mathcal{X}_j\|$, con $i, j = 1, \dots, n$. Dado que el interés está en los valores alrededor del centro de la distribución, que no son muy sensibles a la elección del h , se recomienda únicamente que el h no sea muy pequeño. Véase Cuevas *et al.* (2006) para obtener más información sobre la elección del h en estas situaciones.

De igual forma, se cuenta con la extensión de la medida de profundidad modal funcional para el caso p -variado funcional MD^p .

Finalmente, en (Cuevas *et al.*, 2007) consideran una medida de la profundidad de los datos funcionales bajo proyecciones. La idea básica es proyectar cada curva funcional a lo largo de una dirección aleatoria.

Definición 6: Sea $\mathcal{S}_n = \{\mathcal{X}_i(t)\}_{i=1}^n$ i.i.d., realizaciones de la variable aleatoria funcional $\mathcal{X}(t)$ con dominio $T = [a, b]$, D una medida de profundidad univariante y sea $h \in \mathcal{H}$ una realización aleatoria que pertenece a un proceso independiente \mathcal{H} , y $P_i^h = \langle h, \mathcal{X}_i \rangle \in \mathbb{R}$, es la proyección de \mathcal{X}_i a lo largo de la dirección h , es decir:

$$P_i^h = \langle h, \mathcal{X}_i \rangle = \int_a^b h(t) \mathcal{X}_i(t) dt$$

Entonces, se define la medida de profundidad proyección aleatoria RPD para el i -ésimo dato como:

$$RPD_i = RPD(\mathcal{X}_i) = D(P_i^h)$$

Cuando se cuenta con una colección de proyecciones aleatorias $\{h_l\}_{l=1}^M$, la medida de profundidad se calcula usando todas las proyecciones (debe ser homogénea para diferentes proyecciones). Con esto, si usamos un número suficiente de proyecciones aleatorias, el valor medio de las profundidades de los puntos proyectados define una profundidad para los datos funcionales. Véase Cuevas *et al.* (2007) para una exposición detallada de la profundidad de proyección aleatoria. Las direcciones aleatorias pueden ser generadas por un proceso gaussiano definido en el intervalo $C([a, b])$. Entonces, la profundidad de proyección aleatoria de una curva \mathcal{X}_i se define por:

- $RPD(\mathcal{X}_i) = \frac{1}{M} \sum_{l=1}^M D(P_i^{h_l})$, según Cuevas *et al.* (2007).
- $RTD(\mathcal{X}_i) = \min_l D(P_i^{h_l})$, como se muestra en Cuesta-Albertos y Nieto-Reyes (2008).

A partir de las medidas de profundidad funcional se deriva la definición de medidas robustas de localización. En Fraiman y Muniz (2001) se cuenta con una propuesta de la media recortada y la mediana funcional como un caso particular de esta.

Definición 7: Sea $\mathcal{X}_{(1)}, \dots, \mathcal{X}_{(n)}$, tales que $\mathcal{X}_{(1)}$ es la curva más profunda y $\mathcal{X}_{(n)}$ es la curva menos profunda. Se define una medida robusta de la media funcional como la media de la mayoría de las curvas más profundas $n - [\alpha n]$, como sigue:

$$FTM_\alpha(\mathcal{X}_1, \dots, \mathcal{X}_n) = \frac{1}{n - [\alpha n]} \sum_{i=1}^{n - [\alpha n]} \mathcal{X}_{(i)}$$

donde α es tal que $0 \leq \alpha \leq (n-1)/n$ y $[\cdot]$ denota la parte entera. La media recortada funcional (*FTM*) tomando $\alpha = 0$ es la media funcional.

Si solo una curva alcanza el valor más profundo, este representaría el parámetro mediana funcional y se denota por:

$$FMED(\mathcal{X}_1, \dots, \mathcal{X}_n) = \mathcal{X}_{(1)}$$

Caso contrario, la mediana funcional se define como la media de las curvas que maximiza la profundidad.

De igual forma se tiene una medida robusta para la varianza marginal utilizando el concepto de profundidad. Se define a continuación:

$$\hat{\sigma}_{TSD,\alpha}(t) = \left(\frac{1}{n - [\alpha n]} \sum_{i=1}^{n-[\alpha n]} (\mathcal{X}_{(i)}(t) - FTM_{\alpha}(t))^2 \right)^{\frac{1}{2}} \quad (2.3)$$

donde, $\mathcal{X}_1(t), \dots, \mathcal{X}_n(t)$ son los datos funcionales observados en el tiempo t , y α es tal que $0 \leq \alpha \leq (n-1)/n$ y $[\cdot]$ denota la parte entera.

2.1.3. Bandas de confianza Bootstrap

A partir del cálculo de bandas de confianza se puede evaluar la precisión de un estimador de localización, tal como la media funcional. Para su construcción se utiliza el procedimiento bootstrap suavizado. Este procedimiento tiene en cuenta la estructura de covarianza de los datos (Cuevas *et al.*, 2006).

La función *fdata.bootstrap*, incluida en el paquete *fda.usc*, permite estimar la media, mediana, la moda y la varianza marginal funcional y calcular sus respectivas bandas de confianza. A continuación se detalla el procedimiento de bootstrap suavizado para la construcción de bandas de confianza:

1. Se obtienen b muestras bootstrap a partir de \mathcal{S}_n : $\mathcal{S}_n^{*j} = \{\mathcal{X}_i^*\}_{i=1}^n$ donde $\mathcal{X}_i^{*(j)}(t) = \mathcal{X}_i^*(t) + Z(t)$, con $j = 1, 2, \dots, b$ e $i = 1, 2, \dots, n$ donde $\mathcal{X}_i^*(t)$ es seleccionado al azar y $Z(t)$ es una variable distribuida normalmente con media igual a 0 y matriz de covarianza $\gamma \Sigma_{\mathcal{S}_n}$, donde $\Sigma_{\mathcal{S}_n}$ es la matriz de varianza-covarianza de la muestra y γ es un parámetro de suavización que controla la cantidad de la variabilidad de las nuevas muestras bootstrap (se recomienda $\gamma = 0.05$).

2. A partir de la muestra original, $\mathcal{X}_1(t), \dots, \mathcal{X}_n(t)$, y, a partir de las sucesivas muestras bootstrap $\mathcal{X}_1^{*(j)}, \dots, \mathcal{X}_n^{*(j)}$, se calcula las estimaciones requeridas, $\hat{\theta}(\mathcal{S}_n)$ y $\hat{\theta}(\mathcal{S}_n^{*j})$, respectivamente.
3. A continuación, se obtienen las distancias $\left\{d(\hat{\theta}(\mathcal{S}_n), \hat{\theta}(\mathcal{S}_n^{*j}))\right\}_{j=1}^b$ utilizando la norma L^2 .
4. Finalmente, una banda de confianza bootstrap, correspondiente a un nivel de confianza $(1-\alpha)$ es calculado a partir del cuantil $(1-\alpha)$ de $\left\{d(\hat{\theta}(\mathcal{S}_n), \hat{\theta}(\mathcal{S}_n^{*j}))\right\}_{j=1}^b$.

2.1.4. Método de detección de atípicos para datos funcionales

El procedimiento propuesto en Febrero-Bande *et al.* (2007) para la detección de atípicos funcionales mediante la profundidad de datos se describe en esta sección con la finalidad de usar este contraste en la detección de atípicos en un estudio ILS. Un valor atípico se define por un valor bajo de profundidad. Un dato tiene una profundidad baja o es menor que la de los otros puntos, si difiere en magnitud o forma del resto de curvas. Por otro lado, la mediana funcional de un conjunto de curvas tiene el mayor valor de profundidad. Por esta razón, el análisis de la profundidad de datos se convierte en un método para la detección de valores atípicos.

Por otro lado, en Hyndman y Shang (2010), proponen dos métodos gráficos para la detección de datos atípicos: bagplot y boxplot. Ambos métodos se basan en el uso de las componentes principales robustas. Otra forma de detectar observaciones atípicas lo propone en Hyndman y Ullah (2007), el cual se basa en el análisis de las componentes principales robustas y la integración del error cuadrático desde un modelo lineal.

La librería *fda.usc* proporciona las funciones *outliers.depth.pond* y *outliers.depth.trim*, que permiten detectar valores atípicos mediante la profundidad de curvas. A continuación se describe el algoritmo implementado en estas funciones para detectar la presencia de los valores atípicos:

1. Se calculan las profundidades funcionales correspondientes al conjunto de datos, $\mathcal{X}_1(t), \dots, \mathcal{X}_n(t)$. Por ejemplo, empleando las funciones de profundidad

Fraiman Muniz, modal o proyecciones aleatorias, obteniendo así $\{D(\mathcal{X}_i)\}_{i=1}^n$.

2. A continuación, se obtienen b remuestras con el método bootstrap suavizado $\mathcal{X}_1^{*(j)}(t), \dots, \mathcal{X}_n^{*(j)}(t)$ a partir de las curvas originales $\{\mathcal{X}_i\}_{i=1}^n$, con $j = 1, \dots, b$. Con la finalidad de estimar el corte C a partir de la distribución de la profundidad de los datos originales, tal que $P(D(\mathcal{X}) < C) = \alpha$.

Para que esta estimación no se vea afectada por la presencia de valores atípicos cada remuestra es obtenida ya sea con una probabilidad proporcional a su profundidad, o a partir del (1-trim) % de las curvas más profundas (trim es el porcentaje de curvas menos profundas que se descartan de la muestra original).

3. Se obtiene un valor C^j con $j = 1, \dots, b$ para cada remuestra bootstrap correspondiente al percentil empírico de la distribución de las profundidades, $\{D(\mathcal{X}_i^{*j})\}_{i=1}^n$. La selección del percentil C^j es tal que, en ausencia de valores atípicos, el porcentaje de las observaciones correctas mal etiquetadas como valores atípicos son aproximadamente iguales a una pequeña proporción, por ejemplo $\alpha = 1\%$.
4. Se obtiene C como la mediana de los diferentes b valores de C^j . Si $\mathcal{X}_{i_1}(t), \dots, \mathcal{X}_{i_k}(t)$ son k curvas tal que $D(\mathcal{X}_{i_j}(t)) \leq C$, entonces son consideradas curvas atípicas y normalmente se eliminan de la muestra.

2.1.5. ANOVA Funcional

En la presente sección describiremos el método de Análisis de la Varianza Funcional (FANOVA) propuesto en Cuevas *et al.* (2004). Consideramos L grupos de funciones aleatorias independientes $\mathcal{X}_{ij}(t), i = 1, \dots, L; j = 1, \dots, n_i$ definidas en un intervalo compacto $T = [a, b]$. Denotamos a \bar{X}_i como la media funcional muestral para el grupo $i, t \in T$. Analizamos las siguientes hipótesis para el modelo FANOVA de una vía para k muestras independientes de datos funcionales.

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

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

Para contrastar la hipótesis nula de igualdad de funciones medias de datos funcionales, basada en la versión asintótica de la prueba FANOVA, se utiliza el siguiente estadístico que mide la variabilidad entre grupos:

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

Para calcular el valor crítico del estadístico V_n , se utilizan técnicas de remuestreo, dado que no hay una distribución de referencia. Esto permite, considerar la posible heterocedasticidad entre grupos. De este modo, se rechaza la hipótesis nula si el valor observado de V_n de la muestra es mayor que la aproximación bootstrap del valor crítico calculado a partir de la distribución bootstrap.

2.2. Casos de estudio

En esta sección se describen todos los casos de estudio que han originado el desarrollo de las metodologías estadísticas introducidas en la presente memoria de tesis. Es importante indicar que el objetivo principal de esta tesis ha sido el dar solución a diferentes casos prácticos planteados en el ámbito de la Química Analítica y de la eficiencia energética en edificaciones.

2.2.1. Estudio interlaboratorio a partir de datos termogravimétricos y calorimétricos

Las nuevas aproximaciones FDA a los estadísticos univariantes h y k de Mandel han sido desarrolladas debido a la necesidad de desarrollo de metodologías FDA para llevar a cabo estudios interlaboratorio o ILS, mediante los cuales, entre otras utilidades, se estima la precisión de un método de medida en laboratorio. Este tipo de metodología es aplicable, no sólo a laboratorios de medida, sino a todo tipo de sensores, en el amplio marco de los estudios R & R. Para evaluar el desempeño de estas nuevas aportaciones en el ámbito FDA, se ha emulado un estudio interlaboratorio a partir de varios instrumentos de medida, variando también las condiciones de calibración. Es importante destacar que los datos a partir de los cuales se han evaluado las metodologías propuestas son datos reales y obtenidos por el propio autor

y directores de tesis. A partir de pequeñas variaciones en el instrumento de medida y su calibración, se han emulado los distintos laboratorios, en este caso 7.

Siendo relativamente sencillo acceder a estudios interlaboratorio en el que la variable medida es univariante, no sucede lo mismo cuando los datos son funcionales. La razón es que se suele extraer una o varias características de interés de las curvas a partir de las cuales se hace el análisis estadístico; como hasta la fecha no existían métodos estadísticos para tratar las curvas completas en el ámbito de los ILS, no se incluyen los datos funcionales en las bases de datos de referencia. Por tanto, con el fin de tener una base de datos real con la que poder evaluar el procedimiento de detección de laboratorios atípicos que aquí se propone, se ha emulado la forma de trabajar de 7 laboratorios variando ligeramente la calibración de los instrumentos, los operadores y el analizador.

Dada la experiencia de los directores de la presente tesis en el ámbito del análisis térmico, añadido al hecho que el desarrollo de técnicas FDA para los ILS arranca de la propuesta de miembros de la ASTM relacionados con este ámbito de la Química Analítica, se ha optado por evaluar las técnicas propuestas a partir de un estudio ILS con datos provenientes del análisis térmico; en particular, los correspondientes al análisis termogravimétrico (TG) y calorimétrico (DSC) de muestras de oxalato de calcio. Se ha elegido el oxalato como material debido a que es comúnmente utilizado como material de referencia en termogravimetría y calorimetría, conociendo de forma precisa sus características físicas. La termogravimetría es una técnica de análisis térmico que proporciona información sobre la estabilidad térmica del material mediante la medición de la pérdida de masa en función del tiempo o la temperatura.

En lo que respecta al material estudiado, se ha elegido el oxalato de calcio monohidrato, con un 99% de pureza, proporcionado por la empresa Panreac, pudiéndose formular como $C_aC_2O_4 \cdot H_2O$. Presenta la característica de que su degradación conste de tres pasos o reacciones de degradación, tendiendo lugar en intervalos de temperatura y tiempo muy estrechos. 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. La Figura 2.1 muestra las tres reacciones químicas de las que se compone la degradación termooxidativa del oxalato de calcio, incluyendo reactivos

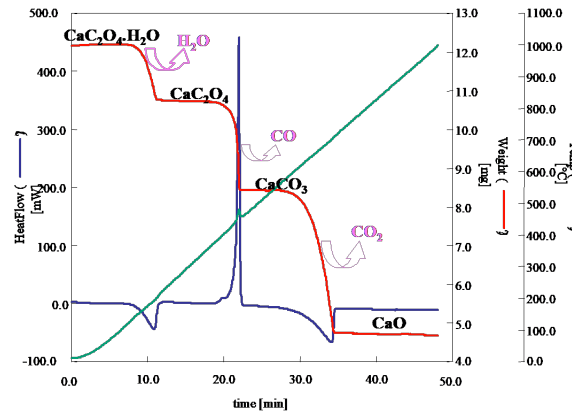


Figura 2.1: Curva TG correspondiente al oxalato de calcio en rojo. La curva TG proporciona información acerca del proceso global de degradación, en este caso en ambiente oxidante. Adicionalmente se muestra la curva de flujo de calor (en azul) y la relación de la temperatura con el tiempo (verde).

y productos y, lo que es más importante, relacionando cada reacción con la pérdida de masa mostrada en la curva TG correspondiente (cada reacción aparece como un escalón en la curva TG).



Figura 2.2: Muestra de oxalato de calcio en un crisol justo antes de ser introducido en el horno del analizador TG.

En la Figura 2.2 se muestra el detalle de un crisol en el que se ha introducido una cantidad de oxalato de calcio, previamente a su introducción en el horno del

analizador simultáneo TG-DSC.

Para simular un ILS común, se emularon 7 laboratorios diferentes combinando dos instrumentos de análisis TG con diferentes calibraciones. De este modo, cada laboratorio emulado ensayó 15 muestras de oxalato cálcico mediante análisis termogravimétrico, por lo tanto, se usaron 105 muestras en total.

Los instrumentos de análisis utilizados fueron dos analizadores simultáneos DSC-TG, en particular un TA Instruments SDT 2960 y, alternativamente, un analizador simultáneo Rheometric STA 1500 (ver Figura 2.3). De cada muestra ensayada, se obtuvo como resultado una curva TG y otra DSC, que proporcionan información acerca de los procesos de degradación sufridos por el material, en forma de escalones de pérdida de peso, en el caso de las curvas TG, y en forma de picos (el área del pico es la cantidad de energía intercambiada en cada etapa de degradación) en el caso de las curvas DSC.

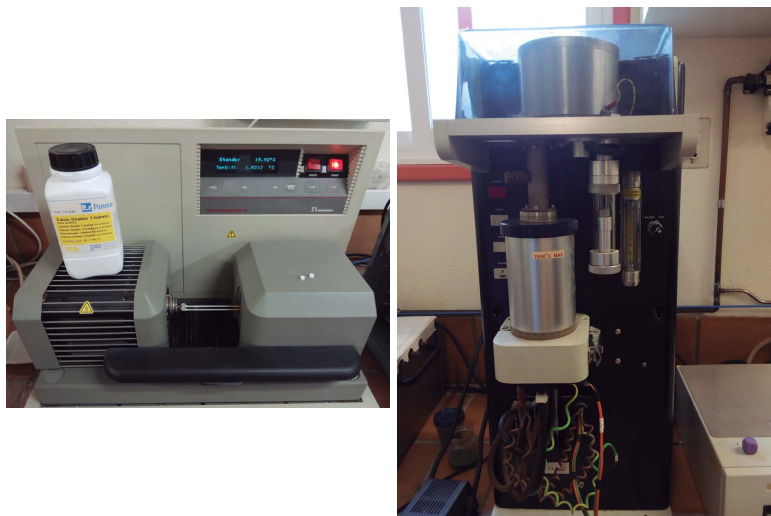


Figura 2.3: En este trabajo se han utilizado dos analizadores simultáneos que realizan análisis TG y DSC “simultáneamente” (instrumentos de medida). Panel izquierdo: analizador simultáneo SDT de TA Instruments, con la celda, crisoles y brazos sujetores en horizontal. Panel derecho: analizador simultáneo STA de Rheometrics Inc, con celda del horno y balanza en vertical.

Las condiciones de ensayo se eligieron atendiendo a la experiencia de los analistas

de laboratorio. De este modo, cada muestra fue calentada a una velocidad constante de $20\text{ }^{\circ}\text{C min}^{-1}$, en un rango comprendido entre 20 y $900\text{ }^{\circ}\text{C}$, dentro de una atmósfera oxidante de 50 ml min^{-1} de aire. La Figura 2.4 muestra las curvas TG correspondientes a los 7 laboratorios estudiados, mientras que la Figura 2.5 muestra las curvas DSC.

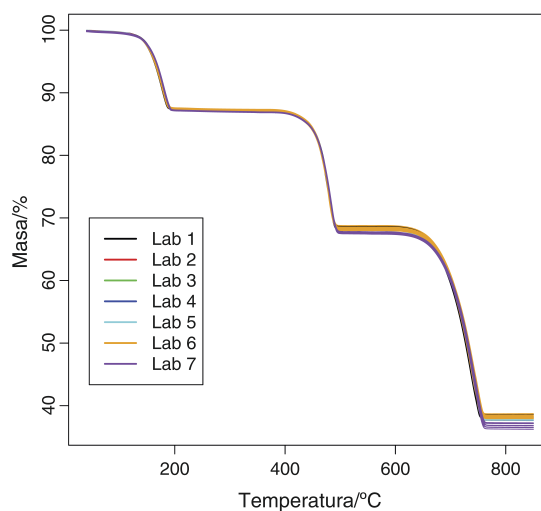


Figura 2.4: Curvas TG correspondientes a cada laboratorio. Han sido obtenidas a partir del ensayo de muestras de oxalato de calcio.

Siendo relativamente sencillo acceder a estudios interlaboratorio en el que la variable medida es univariante, no sucede lo mismo cuando los datos son funcionales. La razón es que se suele extraer una o varias características de interés de las curvas a partir de las cuales se hace el análisis estadístico; como hasta la fecha no existen métodos estadísticos para tratar las curvas completas en el ámbito de los ILS, no se incluyen los datos funcionales en las bases de datos de referencia.

Por tanto, con el fin de tener una base de datos real con la que poder evaluar el procedimiento de detección de laboratorios atípicos que aquí se propone, se ha emulado la forma de trabajar de 7 laboratorios variando ligeramente la calibración de los instrumentos, los operadores y el analizador.

Este diseño de experimentos consta de un grupo de 4 laboratorios cuyos resultados son homogéneos desde el punto de vista de la reproducibilidad y repetibilidad. Esto

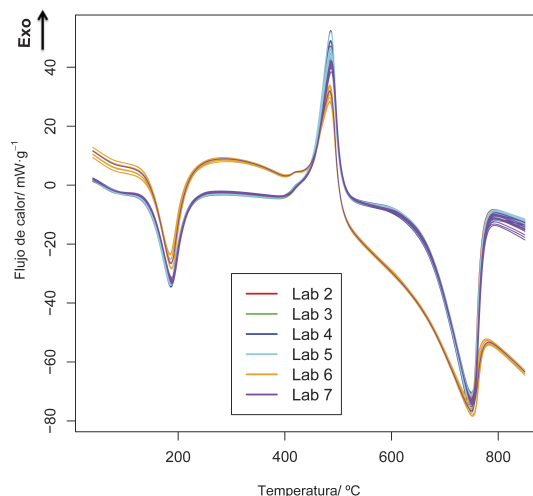


Figura 2.5: Curvas DSC correspondientes a cada laboratorio. Han sido obtenidas a partir del ensayo de muestras de oxalato de calcio.

se ha conseguido utilizando el mismo instrumento SDT, con la misma calibración, y con 4 operarios diferentes que ensayaron las muestras en diferentes momentos. Además, se han obtenidos resultados asignados a un laboratorio para los que se ha utilizado una calibración de temperatura 2°C sesgada con respecto a la verdadera temperatura fusión de zinc (patrón comúnmente usado para tal fin). También se simuló el desempeño de otro laboratorio a partir de los datos obtenidos utilizando un instrumento de análisis con una calibración antigua (de dos años). Por último, se ha diseñado un séptimo laboratorio cuyos resultados presentan sesgo con respecto a los demás, a la par que se caracterizan por una mayor variabilidad, debido a que para su obtención se ha utilizado un instrumento de análisis diferente, en este caso un Rheometric STA 1500, además de utilizarlo con un calibración antigua. La Tabla 2.1 asigna una etiqueta a cada laboratorio, además de describir cuales son sus características principales. Según se muestra, la nueva metodología de detección de laboratorios inconsistentes debería detectar como atípicos los laboratorios 1, 6 y 7.

Estos datos serán analizados en el Capítulo 3, en el cual se presentarán las nuevas aproximaciones FDA de los estadístico h y k de Mandel, que posteriormente se

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

Tabla 2.1: Etiqueta y descripción de cada uno de los laboratorios.

aplicarán a este conjunto de datos.

2.2.2. Control de la eficiencia energética y confort térmico de superficies comerciales de empresas textiles en Ciudad de Panamá

En esta sección se presenta un caso de estudio real de control y monitorización de la eficiencia energética y confort térmico en superficies comerciales de una empresa textil gallega recientemente radicada en la Ciudad de Panamá. Esta empresa ha contratado una plataforma web inteligente, en este caso Σ qus (Barbeito *et al.*, 2017), para la gestión de todos los datos que caracterizan el funcionamiento de las instalaciones HVAC y de iluminación en cada una de las tiendas inauguradas en los tres centros comerciales. Una vez obtenidos los datos de consumo energético, temperaturas, nivel de CO₂ y humedades relativas, entre otras variables CTQ, la necesidad de la empresa es disponer de un método automático para la detección de anomalías, detección de causas asignables de las mismas e implementación de medidas correctivas. En particular, la empresa, a través de la compañía desarrolladora de plataformas web Σ qus, ha planteado el problema de detección de anomalías en el consumo energético de climatización. Para ello ha proporcionado los datos correspondientes con medidas horarias comprendidas entre el 1 de agosto y el 27 de noviembre de 2017. En la Figura 2.6, se representan las curvas de consumo HVAC desde las 0:00 horas hasta las 23:00. La forma de las curvas de consumo diario se debe a las siguientes características de funcionamiento de las instalaciones: todas las instalaciones de ofimática, iluminación y HVAC de la tienda de estudio están apa-

gadas hasta las 9:00 o 10:00 de la mañana. En el arranque, se produce un pico de consumo debido a las características de las máquinas HVAC. Acto seguido, a partir de las 12:00 en adelante, el consumo permanece relativamente constante hasta las 20:00, 21:00 o 22:00 de la tarde, en la que la tienda cierra sus puertas y apaga todas las máquinas de las diferentes instalaciones. Desde que comienza el apagado hasta que se deja de consumir energía, pasan alrededor de 1 o 2 horas, cayendo el consumo a una tasa de cambio constante. Como se puede observar, los datos resultantes pueden considerarse datos funcionales.

Es importante destacar que la hora de apertura y cierre depende en gran medida del tipo de día, es decir, si es un día de diario, un sábado (se cierra una hora más tarde) o un domingo (se abre una hora más tarde con respecto al horario de lunes a viernes). Además, dependiendo del mes, la forma de las curvas de consumo y su dispersión puede variar; de hecho, se observa que en noviembre el pico de consumo correspondiente al arranque, a las 11:00, prácticamente es imperceptible. Esto se debe a procesos de regulación llevados a cabo en las instalaciones de HVAC en dicho mes.

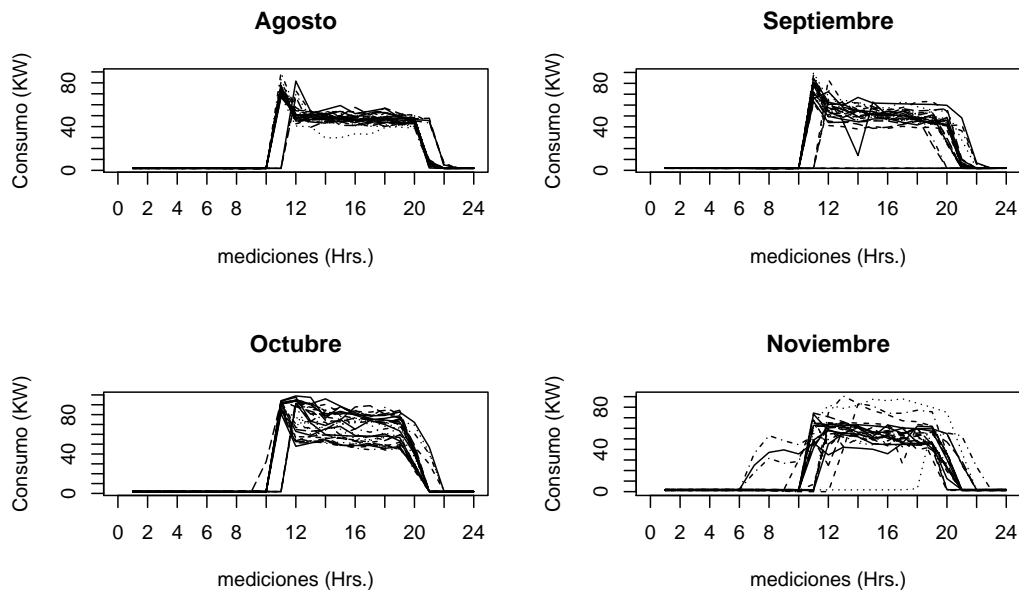


Figura 2.6: Curvas diarias de consumo energético en sistemas HVAC.

De un estudio preliminar de los datos, se puede observar que en el mes de agosto no se registran anomalías, sin embargo tanto en este mes como en los siguientes meses, se puede observar en la Figura 2.6 que existen curvas de consumo que difieren en su forma, esto es debido a que normalmente la tienda tiene un horario fijo de lunes a viernes, y los sábados cierran una hora más tarde y los domingos abren una hora más tarde.

En el rango de tiempo al que pertenecen los datos, los ingenieros de mantenimiento y los trabajadores de FRIDAMA, Σqus y Nerxus, han detectado una serie de anomalías en las instalaciones HVAC y, no sólo eso, sino que han identificado cuáles son las causas asignables. Por lo tanto, se parte de un caso de estudio con datos controlados a partir de los cuáles se puede evaluar la metodología de gráficos de control a partir de datos funcionales que se presenta en este capítulo. En particular, se han detectado las siguiente anomalías con sus causas asignables:

- El día 11 de septiembre hubo una bajada en el consumo del aire acondicionado hacia el medio día.
- Los días 21, 22 y 30 de septiembre estuvo cerrado el centro comercial, por lo que no hubo consumo y la temperaturas fueron altas (Panamá tiene clima tropical y los sistemas HVAC son en exclusiva de aire acondicionado) y el consumo de energía muy bajo.
- El día 27 de septiembre se hicieron pruebas de mantenimiento en las instalaciones de la tienda, actividades que quedaron patentes en la forma de las curvas de consumo energético y temperatura.
- El día 29 de septiembre se apagaron las instalaciones de la tienda una hora antes de lo habitual.
- A partir del 19 de septiembre el aire acondicionado se apaga media hora antes, es decir, se produce un cambio de regulación en el sistema HVAC, hecho que se ve plasmado en las curvas de consumo energético y temperatura.
- En octubre, a mediados de este mes, se registra una fuga en el circuito del aire acondicionado. A partir de ese momento el consumo energético empezó a subir.
- El 1 de noviembre se hizo una reparación provisional, que provocó que el consumo bajara y también cambiara un poco su comportamiento con respecto a

antes de producirse la avería. Por ejemplo, ya no se registra un pico de arranque a las 11:00, como se apreciaba en los meses anteriores. Finalmente, entre el 17 y 20 de noviembre el consumo volvió a aumentar nuevamente.

Debido a la diferencia de horario de apertura de la tienda entre los días de lunes a viernes y los días sábado y domingos, se consideran las curvas de consumo sólo de lunes a viernes para aplicar el procedimiento de gráficos de control. De igual modo podría aplicarse la metodología propuesta para la detección de anomalías los sábados, por un lado, y los domingos.

Capítulo 3

Nuevo enfoque FDA para detección de atípicos en Estudios Interlaboratorios (ILS)

3.1. Introducción

El uso de técnicas FDA es poco común en las tareas de control estadístico de la calidad. Específicamente, en el control de la calidad de sistemas de medición no se tienen propuestas para tratar los datos cuando son curvas. Por ejemplo, para la realización de estudios interlaboratorio en el campo de la Química Analítica, donde se generan los datos descritos en el Capítulo 3, en la sección 3.2.1, se realizan estudios univariantes en los que se requiere la extracción previa de una característica representativa de las curvas. Utilizar un enfoque escalar puede llevar a perder información importante de un proceso, de hecho, dependiendo de la característica extraída, el resultado de la prueba podría ser diferente.

Con la finalidad de contar con una primera aproximación al contexto FDA que permita utilizar toda la información de la curva y tener un análisis previo a la aplicación de la metodología propuesta en esta memoria, se han aplicado algunas técnicas funcionales. En la siguiente subsección se detalla el uso de las siguientes técnicas y procedimientos en un Estudio Interlaboratorio donde se cuenta con curvas DSC y TG:

- Análisis exploratorio FDA.
- Método de detección de valores atípicos FDA basado en el cálculo de la profundidad de datos (Modal, FM y Proyecciones aleatorias).
- ANOVA Funcional.

El análisis exploratorio nos permite tener una estimación de la tendencia y variabilidad funcional para cada laboratorio y tener una idea de qué laboratorios podrían ser inconsistentes. Por otro lado, el ANOVA permite contrastar si las medias funcionales de los laboratorios son iguales. En el caso, para la detección de resultados atípicos individuales para cada laboratorio se utiliza el cálculo de la profundidad funcional y test de detección de resultados atípicos.

La aplicación de estas técnicas FDA permite identificar qué laboratorios presentan resultados inconsistentes. Sin embargo, para contar con herramientas propias para el ILS que sean extensiones naturales de procedimientos utilizados en la detección de laboratorios inconsistentes en un ILS, se desarrolla una nueva metodología para realizar un estudio interlaboratorio a partir de las extensiones de los estadísticos h y k , que se expone en la sección 3.2.

Esta nueva metodología es validada a través de un estudio de simulación (sección 3.3) y su aplicación está compuesta de un conjunto de datos experimentales (sección 3.4). En lo que respecta al estudio de simulación, está compuesto de una amplia gama de escenarios en los que se tienen de simuladas curvas térmicas. Específicamente, curvas termogravimétricas (TG), aprovechando que han sido objeto de estudio, desde la perspectiva FDA, en diversos trabajos relacionados con la Química Analítica.

En la siguiente sección se exponen las consideraciones generales para realizar un estudio interlaboratorio y, posteriormente, se realiza un estudio ILS utilizando técnicas FDA.

3.1.1. Consideraciones generales para realizar un estudio interlaboratorio

Aunque los enfoques para los estudios interlaboratorio son bastante diferentes, algunos problemas son comunes a todos los tipos de estudios.

1. Requisitos preliminares. Todos los participantes en estudios interlaboratorio deberían establecer de antemano todos los sistemas posibles de aseguramiento de la calidad dentro del laboratorio.
2. Definiciones. Las definiciones de repetibilidad, reproducibilidad y diferentes tipos de precisión intermedia están dadas por la correspondiente norma ISO.
3. Condiciones de repetibilidad. Estas son las condiciones en las que se obtienen resultados experimentales independientes, utilizando el mismo método, en muestras idénticas, en el mismo laboratorio y por el mismo operador que usa el mismo equipo en cortos intervalos de tiempo.
4. 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.
5. Selección de material: valor verdadero y valor asignado del contenido. Los materiales para una prueba interlaboratorios deben seleccionarse de tal manera que las muestras sean representativas del tipo de material que generalmente se analiza con respecto a la matriz y el rango de concentración. Si bien es posible proporcionar tales muestras para materiales manufacturados, las muestras naturales a menudo no están disponibles en el rango de concentración requerido. Para algunas sustancias, como la sangre, se encuentran disponibles los denominados materiales de control de calidad.
6. Homogeneidad de materiales. La homogeneidad de los materiales de prueba es un requisito fundamental para todos los estudios interlaboratorio.
7. Replicación de experimentos. La AOAC (Association of Official Agricultural Chemists) sugiere realizar una replicación de dos niveles, lo que significa que las dos muestras muestran una pequeña diferencia en la composición y la concentración. La norma ISO también recomienda una distribución de muestra codificada ciega para evitar la manipulación de la estimación de repetibilidad.

8. Detección de atípicos. Tanto en ensayos de colaboración como en pruebas de sesgo, los procedimientos de detección de valores atípicos juegan un papel fundamental (Dvorkin, 1994; Kelly, 1990; Uhlig y Lischer, 1998), que es el detectar los laboratorios que proporcionan resultados significativamente diferentes de los demás y descartar los datos inconsistentes que ellos proveen. Así se podrá estimar de una forma más fiable la precisión de un nuevo método de medida o el valor de una determinada propiedad de un material, por poner dos ejemplos.

Los ILS establecen varias tipologías de atípicos. Estos se pueden clasificar según el nivel de significación del test de detección o según la causa que los origina. Según el nivel de significación, los valores identificados como valores atípicos en un nivel de significación entre 1 y 5 % se consideran “rezagados”, y solo aquellos valores que son significativos a un nivel de significación del 1 % se consideran como valores atípicos y, por lo tanto, se eliminan. En lo que respecta a la causa que los originan, se establecen tres tipos. El tipo I se corresponden con observaciones atípicas entre las diferentes réplicas que hacen todos los laboratorios, el tipo II hace referencia a la diferencia entre medias correspondientes a los diferentes laboratorios. mientras que el tipo III de atípicos tiene en cuenta sus diferencias en dispersión, desviación estándar (Hund *et al.*, 2000b). En los estudios ILS, son más importantes por útiles los métodos que detectan atípicos tipo II y III, es decir, que detectan laboratorios o sensores atípicos, no observaciones atípicas.

Se han aplicado muy diversos test para la detección de atípicos en estudios ILS, siendo todos ellos, hasta la fecha, de naturaleza univariante. Las normas internacionales generalmente proponen implementar test de atípicos basados en la varianza (pruebas unilaterales) previamente a la aplicación de los test que detectan diferencias entre las medias de cada laboratorio (Hund *et al.*, 2000b). La prueba de Cochran es, con mucho, el contraste basado en la varianza más utilizado en estudios interlaboratorio (Cochran, 1941; Hund *et al.*, 2000b). Aparte de este, también suele emplearse el test F para comparar las varianzas dentro de cada laboratorio con respecto a la varianza de repetibilidad (Hund *et al.*, 2000b). Con respecto a los contrastes de detección de laboratorios atípicos basados en la diferencia de valores medios, el más utilizado es el test de Grubbs

(para valores atípicos sencillos o dobles) (Grubbs, 1950; Grubbs y Beck, 1972) y el contraste de Graf y Henning (Wilrich, 2013). Además, se han propuesto algunas alternativas robustas como alternativa a los contrastes clásicos. Estas están basadas en el cálculo de la mediana de las desviaciones absolutas con respecto a la mediana mediana (MEDA) (Davies, 1988), en la estimación de la media robusta y el cálculo de desviación estándar (Committee, 1989), y en la biweight function de Tukey (Davies, 1988).

De entre los test existentes para la detección de atípicos, han cobrado especial relevancia en los estudios interlaboratorio aquellos que proporcionan herramientas gráficas que permiten la identificación de laboratorios inconsistentes de una forma simple y totalmente intuitiva. Por otro lado, también se propone el uso de diagramas como los box-plot (Miller y Miller, 1988), los gráficos de Youden (Youden, 1959) y los gráficos de control, entre otras alternativas. Entre los diferentes métodos gráficos existentes, destacan los diagramas de barras relacionados con los contrastes para la detección de atípicos h y k de Mandel (Minkkinen, 1995). El estadístico h tiene en cuenta la variabilidad intralaboratorio, es decir, las diferencias entre los valores medios de variable CTQ obtenidos por cada laboratorio con respecto a la media global, mientras que el estadístico k estima la variabilidad intralaboratorio comparando las varianzas de repetibilidad correspondientes a cada laboratorio. Por lo tanto, estos estadísticos se emplean para detectar valores atípicos a partir de la diferencia de medias (tipo II) y a partir de las diferencias entre desviaciones estándar (tipo III) (Hund *et al.*, 2000b). Actualmente, su uso ha sido generalizado por los protocolos que rigen el desarrollo de ILS colaborativos y de sesgo, entre otros (ASTM-E691, 2004), siempre combinado con la aplicación de otros contrastes como las pruebas de Cochran, Grubbs y F.

3.1.2. Técnicas FDA aplicadas en un Estudio Interlaboratorio

En el presente estudio, se ha utilizado el software estadístico R (R Core Team, 2018), considerado lengua franca de la investigación estadística. El software estadísti-

co R es la opción más adecuada para llevar a cabo estudios de FDA ya que cuenta con paquetes tales como el `fda.usc` (Febrero-Bande y Oviedo de la Fuente, 2012) que permite aplicar las técnicas FDA explicadas en el Capítulo 2.

Además del estudio FDA, también se calculan los estadísticos h y k univariantes para mostrar que los resultados de un estudio ILS pueden ser diferentes dependiendo de la característica (aquella que resume toda la información de la curva) elegida.

Estudio descriptivo y FANOVA de las curvas TG

El análisis exploratorio consiste en la estimación de la media y la varianza funcional global y para cada laboratorio. Esto permite obtener información acerca de la tendencia y variabilidad de los resultados (ver Figuras 3.1 y 3.2). Además, para la media global de las curvas TG se calculan las bandas de confianza a un nivel del 95 %.

Como se puede observar en la Figura 3.1, la estimación de la media funcional global de las curvas TG para oxalato de calcio evidencia tres procesos de degradación en condiciones oxidantes bien definidos y separados. Además, en esta figura se presenta la estimación de las bandas de confianza cuya amplitud es muy estrecha. El hecho de que la amplitud de las bandas de confianza sea pequeña, es un indicativo de una baja variabilidad en la estimación de la media funcional global.

Con respecto a la estimación de la media funcional para cada laboratorio, en la Figura 3.2 se puede apreciar que la diferencia entre las medias funcionales de los laboratorios del 2 al 4 es mínima, tanto así que se solapan. Por otro lado, se observan diferencias entre las medias de los laboratorios 1, 6 y 7 a temperaturas altas.

Además, se puede observar que la estimación de la varianza funcional de las curvas TG para cada laboratorio va en aumento cuando la temperatura también aumenta, y que en cada proceso de degradación existen picos de variabilidad. Igualmente, el nivel de la varianza es muy alto en el último paso de la degradación. Es decir, la variabilidad del proceso medido es mayor a temperaturas más altas.

Las variaciones funcionales anormalmente altas se observan para los laboratorios 1 y 6. Estos picos de varianza se producen a las temperaturas donde se alcanza la tasa máxima de degradación (para cada proceso de degradación). Por otro lado, se observan también pequeños picos en los laboratorios restantes. Sin embargo, un au-

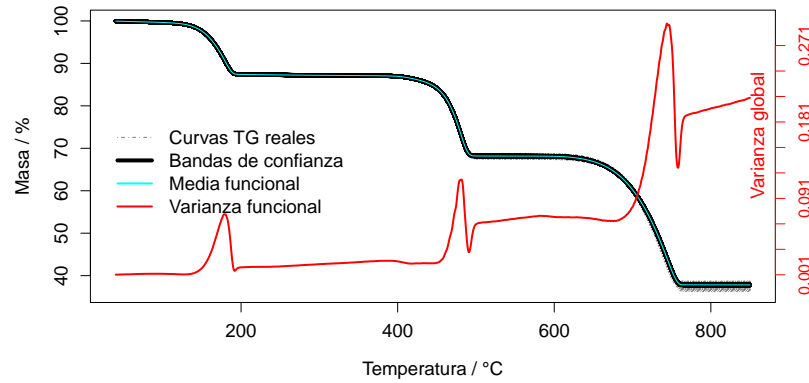


Figura 3.1: Se muestran las bandas de confianza para la media global de las curvas TG a un nivel de significación del 95 % y la varianza global funcional. Media y varianza funcional globales se estiman teniendo en cuenta todas las curvas

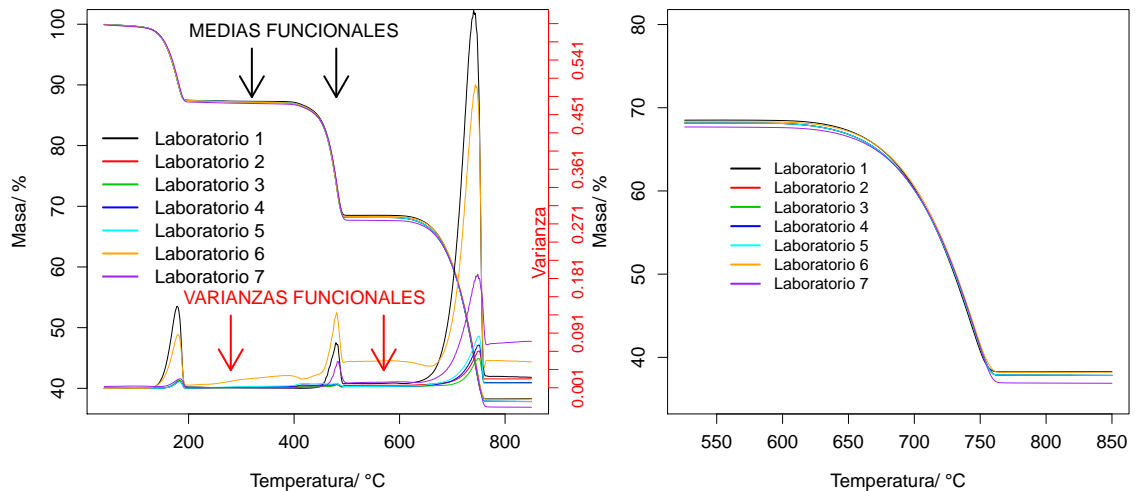


Figura 3.2: En el panel izquierdo se muestran las estimaciones de la varianza y media funcional correspondientes a las curvas TG de oxalato de calcio obtenidas por cada laboratorio. En el panel derecho se observan las medias funcionales correspondientes al tercer proceso de degradación

mento en sus varianzas funcionales se observa especialmente en la última etapa de la degradación. Se puede concluir que los resultados obtenidos por los laboratorios 1, 6 y 7 podrían ser valores atípicos y, por tanto, si se confirma, susceptibles de ser eliminados de las estimaciones ILS finales. Es importante destacar que está es una primera aproximación a la resolución del problema utilizando el análisis exploratorio. Más adelante se describirán los contrastes propuestos para la identificación de laboratorios atípicos (en lugar de observaciones atípicas como se muestra en esta sección)

Todos estos resultados son razonables, ya que los laboratorios identificados como atípicos se han generado a partir de calibraciones incorrectas (laboratorios 1, 6 y 7) y una máquina de ensayo diferente (laboratorio 1). Por otra parte, parece que el laboratorio caracterizado por un sesgo (diferencia entre la media de cada laboratorio y la media global) más grande en las curvas TG es aquel al que le corresponde una calibración de la temperatura incorrecta (ver Figuras 3.1 y 3.2), mientras que la variabilidad se ve afectada principalmente por el uso calibraciones antiguas y por la utilización de una máquina de ensayo diferente, en este caso un dispositivo STA.

El siguiente paso de este análisis es realizar un estudio de detección de resultados atípicos sin tener cuenta el laboratorio. Esto se hace para obtener información más objetiva sobre qué laboratorio podría proporcionar resultados diferentes. Dependiendo del tipo de profundidad utilizada, el número de curvas TG identificadas como valores atípicos difiere. Por ejemplo, en las tablas 3.1 y 3.2 se muestran las profundidades correspondientes a las diferentes curvas TG, utilizando los métodos de proyecciones aleatorias y modal, respectivamente.

Para determinar las profundidades atípicas, se utilizaron $B = 1000$ remuestras, y un valor del parámetro $\gamma = 0.05$, con un nivel de significación $\alpha = 0.1$ (percentil 10 %) para estimar el cuantil que permita identificar los resultados atípicos.

El mayor número de valores atípicos se obtiene a partir del cálculo de la profundidad modal, mientras que si se miden las profundidades usando el método de proyecciones aleatorias, se obtiene el menor número de identificaciones de las curvas TG atípicas entre los tres métodos utilizados. Por otro lado, la profundidad proyección aleatoria, identifica también algunas de las curvas TG correspondientes a los laboratorios 1, 6 y 7 como datos atípicos.

Tabla 3.1: Profundidades utilizando el método de proyección aleatoria correspondientes a las diferentes curvas TG obtenidas de muestras de oxalato de calcio. Las curvas TG correspondientes a las profundidades marcadas en negrita se consideran atípicas (cuando la profundidad de una curva TG es menor que el percentil 10 de la distribución de las profundidades, 0.0908)

Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7
0.1351	0.2170	0.2540	0.1586	0.2445	0.1212	0.1232
0.0795	0.2282	0.1785	0.2247	0.2084	0.0914	0.1376
0.1361	0.2010	0.2240	0.2401	0.1816	0.1750	0.1135
0.1251	0.2159	0.2330	0.2159	0.2190	0.0995	0.1010
0.1022	0.2361	0.2207	0.2073	0.1770	0.1953	0.1269
0.1188	0.1988	0.2113	0.2204	0.2330	0.1252	0.1334
0.1251	0.1785	0.1900	0.2240	0.1543	0.1212	0.1376
0.1437	0.1769	0.2251	0.2396	0.2220	0.0698	0.1144
0.1188	0.2127	0.1899	0.2247	0.1945	0.1134	0.1169
0.1188	0.1711	0.2464	0.2401	0.2282	0.1351	0.1121
0.1033	0.2309	0.2540	0.2328	0.2127	0.0879	0.0827
0.1361	0.2401	0.2190	0.1597	0.1982	0.1498	0.1121
0.1331	0.2404	0.2251	0.2309	0.2309	0.1631	0.0584
0.0644	0.2240	0.2129	0.2363	0.1721	0.1789	0.1076
0.0996	0.1945	0.2257	0.2404	0.2464	0.1578	0.0793

Tabla 3.2: Profundidades utilizando la medida modal correspondientes a las diferentes curvas TG obtenidas a partir de muestras de oxalato de calcio. Las curvas TG correspondientes a las profundidades marcadas en negrita han sido identificadas como valores atípicos (si la profundidad de una curva TG es menor que el percentil 10 de la distribución de las profundidades, 2.9245)

Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7
2.6027	11.9933	16.1837	6.6202	15.6488	1.4947	1.6548
1.1937	13.5966	8.4040	14.9200	13.4324	1.2635	2.1337
2.3395	13.1208	13.2219	15.3169	8.7713	4.4644	2.3310
1.6619	10.7575	14.4891	10.7575	13.4539	0.7713	2.0267
1.1812	14.6270	11.5517	12.6996	6.9916	5.9693	1.9573
1.7366	10.0207	11.9340	11.8748	14.4891	0.8459	2.0380
1.6619	8.4040	10.9403	14.0598	6.5118	1.4947	2.1337
2.4603	8.3704	14.1373	16.0011	13.1552	0.8514	2.1775
1.7366	11.9572	10.1641	14.9200	9.9451	0.7906	1.3564
1.7366	8.8463	14.7410	15.3169	13.5966	2.6027	2.4958
1.5871	13.1214	16.1837	15.3309	11.9572	0.7170	1.2343
2.3395	15.3169	13.4539	6.7424	12.3862	2.0442	2.4958
1.6871	14.5902	14.1373	13.1214	13.1214	3.8299	0.5465
0.4296	14.0598	12.4551	14.7803	7.7295	3.6674	2.3760
1.4124	9.9451	12.5093	14.5902	14.7410	4.1367	1.0056

Por lo tanto, podemos suponer que estos laboratorios proporcionan curvas TG diferentes a las obtenidas por los 4 laboratorios restantes. Al utilizar el procedimiento de detección de valores atípicos mediante los diferentes métodos de cálculo de profundidades, todas las curvas TG correspondientes a los laboratorios 1 y 7 han sido identificadas como valores atípicos, además casi todas las curvas TG obtenidas por el laboratorio 6. Estas son evidencias objetivas de que los laboratorios 1, 6 y 7 proporcionan diferentes resultados.

Finalmente, se aplica un FANOVA para determinar si existen diferencias entre las medias de los resultados correspondientes a los 7 laboratorios. La variable porcentaje de pérdida de masa en una curva TG es la variable de respuesta, mientras que el laboratorio al que pertenece los resultados se considera como una variable cualitativa o factor. Se utiliza la función `anova.onefactor` del paquete `fda.usc` para aplicar el contraste de igualdad de medias funcionales. En la Figura 3.3 se presenta la distribución bootstrap del estadístico de contraste y la ubicación de su valor en la muestra (146.521), representado por una línea segmentada. A partir de 1000 remuestras se calcula el p-valor, en este caso menor que 0.001, con lo cual se rechaza la hipótesis nula. Por tanto, se obtienen el resultado que se esperaba incluso después de realizar el análisis descriptivo y la prueba de detección de atípicos, es decir, la detección de los laboratorios 1, 6, y 7 como inconsistentes.

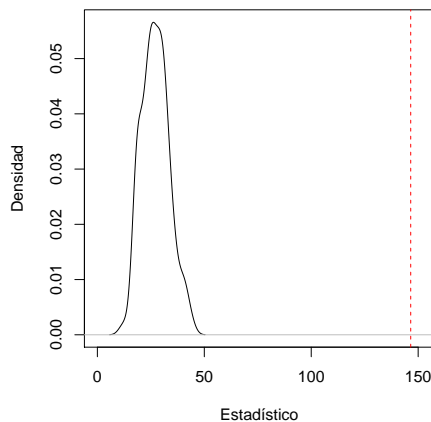


Figura 3.3: Distribución Bootstrap del estadístico del ANOVA Funcional

Estudio descriptivo y FANOVA de las curvas DSC

En este caso, el ILS se realiza a partir de los resultados de seis laboratorios, los laboratorios etiquetados del dos al siete. La Figura 3.4 muestra las curvas DSC obtenidas para cada laboratorio. Como se puede observar, una vez más tres reacciones aparecen bien definidas y separadas, descritas por tres picos del flujo de calor. El primer y tercer pico se corresponde con reacciones químicas endotérmicas, mientras que el segundo se asocia a una reacción de degradación exotérmica. Las curvas DSC se superponen unas sobre otras, excepto las curvas DSC proporcionadas por el laboratorio 6 y algunas correspondientes al laboratorio 7.

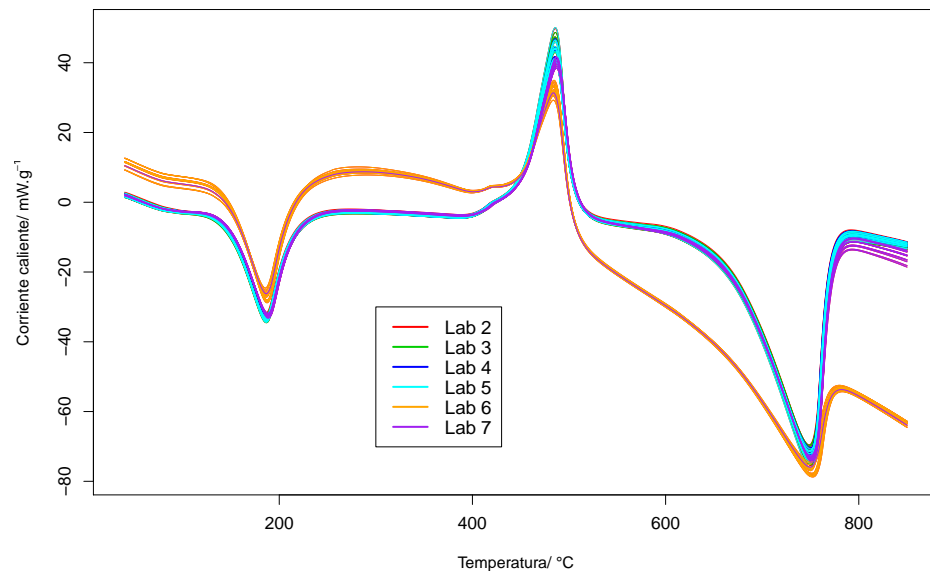


Figura 3.4: Curvas DSC obtenidas por seis laboratorios a partir de muestras de Oxolate de Calcio.

Como se puede observar en la Figura 3.5, la media global funcional de las curvas DSC obtenidas a partir del oxalato de calcio presenta unas relativamente ajustadas bandas de confianza, aunque más anchas que las calculadas para la media funcional de las curvas TG. Aparentemente, las curvas DSC obtenidas por el analizador simultáneo SDT presenta más variabilidad que las curvas TG. La estimación de la

media funcional de las curvas DSC es menos precisa que la obtenida a partir de las curvas TG. Por otro lado, en la estimación de la varianza de las curvas DSC, se observa un aumento de la variabilidad a altas temperaturas.

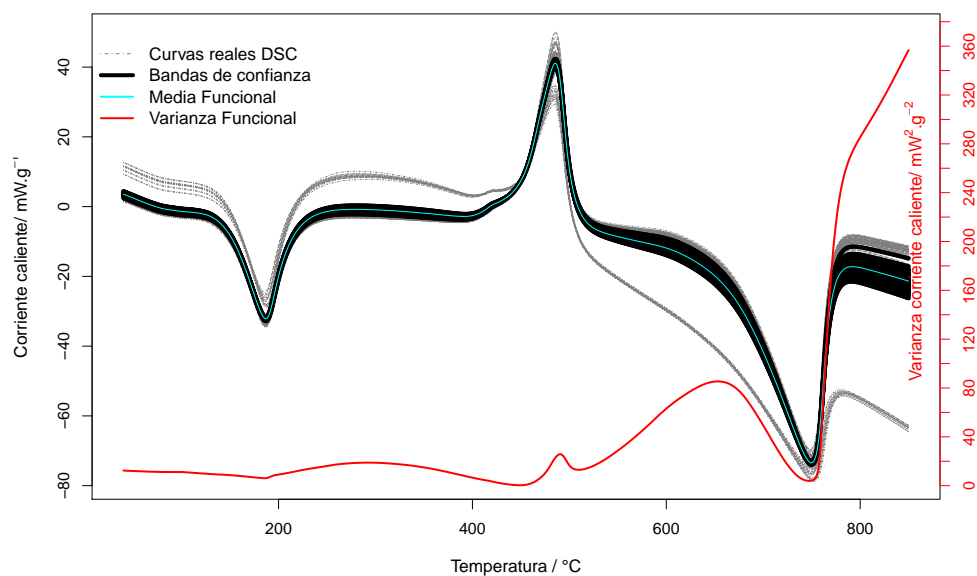


Figura 3.5: Curvas DSC, estimación de la media funcional con sus bandas de confianza al 95 %, y estimación de la varianza funcional correspondiente a las muestras oxalato de calcio.

Las medias y varianzas funcionales de las curvas DSC para cada laboratorio también se calculan para medir la tendencia y la variabilidad de los resultados obtenidos por cada laboratorio usando el mismo procedimiento experimental (Figura 3.6). Como se puede observar, las medias funcionales de las curvas DSC para cada laboratorio son muy similares, excepto la correspondiente al laboratorio 6, que utiliza una vieja calibración (obsoleta).

Las varianzas funcionales para cada laboratorio presentan tres picos positivos correspondientes a los picos endotérmico y exotérmico (indicadores de reacciones de degradación) de las curvas DSC. La variabilidad es mayor a temperaturas donde la velocidad de las reacciones oxidativas es también mayor. Se observan menos diferencias entre las varianzas funcionales de las curvas DSC obtenidas por cada laboratorio

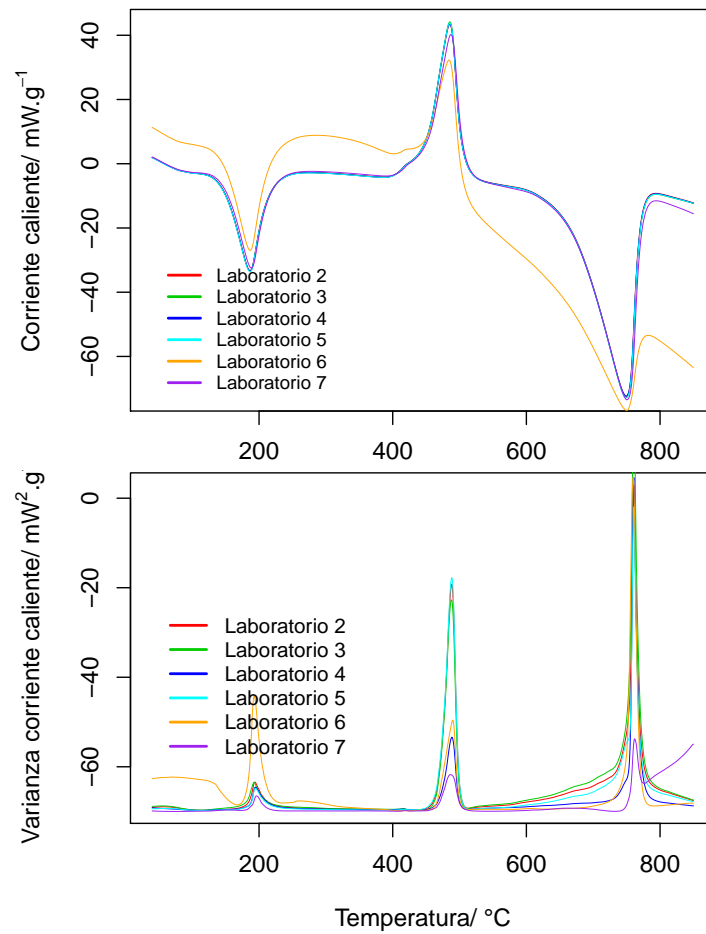


Figura 3.6: Estimación de las medias y varianzas funcionales de las curvas DSC obtenidas por cada laboratorio a partir de muestras de oxalato de calcio.

si estas éstas se comparan con las observadas en el caso de las curvas TG. Sin embargo, parece que las curvas DSC obtenidas por el laboratorio 6 tienden a presentar mayor varianza. Por tanto, los resultados obtenidos por el laboratorio 6 son candidatos a ser valores atípicos. Si esto se confirma, el laboratorio 6 sería eliminado de las estimaciones finales en un ILS.

El segundo paso de este análisis funcional es realizar un estudio de detección de valores atípicos. Las Tablas 3.3 y 3.4 presentan las profundidades correspondientes a las diferentes curvas DSC calculadas utilizando los métodos de profundidad modal y proyección aleatoria, respectivamente.

Tabla 3.3: Profundidades utilizando el método por proyección aleatoria correspondientes a las diferentes curvas DSC obtenidas de muestras de oxalato de calcio. Las curvas DSC correspondientes a las profundidades marcadas en negrita se definen como valores atípicos (si la profundidad de una curva DSC es menor que el percentil 10 de la distribución de las profundidades, 0.0798).

Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7
0.1270	0.1775	0.2131	0.1400	0.0819	0.1976
0.1984	0.2291	0.1484	0.1963	0.0831	0.2030
0.2143	0.2319	0.2016	0.1746	0.0744	0.2016
0.2030	0.1433	0.2110	0.1824	0.1015	0.1666
0.1824	0.2202	0.2372	0.2110	0.0797	0.2041
0.1558	0.2016	0.2143	0.1334	0.0798	0.2096
0.2218	0.1830	0.2202	0.2291	0.0897	0.2016
0.2202	0.2112	0.1307	0.2372	0.1015	0.1831
0.2372	0.1973	0.1437	0.1882	0.0876	0.1686
0.2131	0.1848	0.1746	0.1973	0.0960	0.2041
0.1852	0.2131	0.1778	0.2233	0.0703	0.2032
0.2131	0.2246	0.1766	0.1963	0.0897	0.2016
0.2319	0.1717	0.2319	0.2157	0.1015	0.1746
0.1814	0.2065	0.1568	0.2325	0.1015	0.1844
0.1848	0.1433	0.2291	0.2016	0.0960	0.1779

Cuando se calcula la profundidad por proyección aleatoria, sólo tres curvas DSC correspondientes al laboratorio 6 son detectadas como valores atípicos. Por lo tanto, este laboratorio podría proporcionar curvas DSC diferentes de las obtenidas por los restantes laboratorios. Se obtienen mayores evidencias cuando se calcula la profundidad modal. De hecho, todas las curvas DSC correspondientes al laboratorio 6 se

Tabla 3.4: Profundidades utilizando el método modal correspondientes a las diferentes curvas DSC obtenidas de muestras de oxalato de calcio. Las curvas DSC correspondientes a las profundidades marcadas en negrita se definen como valores atípicos (si la profundidad de una curva DSC es menor que el percentil 10 de la distribución de las profundidades, 3.199)

Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7
6.1005	8.3484	10.6621	7.9054	1.6322	4.8224
10.8817	12.3794	4.9814	11.7619	2.5315	4.9910
12.4427	12.3152	12.1633	9.5054	1.4780	4.3486
11.0360	7.7439	12.3148	9.4926	2.8170	1.4980
9.4926	12.1509	13.8290	12.3148	1.4432	4.7700
9.1415	12.1633	12.4427	2.5768	2.1526	5.4513
12.9456	11.3906	12.1509	12.3794	1.8701	4.3486
12.1509	12.4332	1.8293	13.8290	2.8170	3.1774
13.8290	9.9901	4.5184	9.6577	1.8610	1.8310
10.6621	7.5364	9.5054	9.9901	2.7553	4.7700
10.7022	10.6621	9.7809	13.0506	1.5444	4.9724
10.6621	12.8747	8.0092	10.6083	1.8701	4.3486
12.3152	9.1126	12.3152	12.3090	2.8170	3.7312
8.8647	12.2934	9.1773	14.1676	2.8170	3.9567
9.9893	7.7439	12.3794	12.1633	2.7553	4.1255

identifican como atípicas; además, se detectaron tres curvas atípicas en el laboratorio 7.

Para el caso de las curvas DSC, también se aplica un FANOVA para determinar si existen diferencias entre las medias de los resultados de los 6 laboratorios. La variable porcentaje de pérdida de masa en una curva TG es la variable respuesta, mientras que el laboratorio al que pertenecen los resultados es la variable factor. Se utiliza la función `anova.onefactor` del paquete `fda.usc` para aplicar el contraste de igualdad de medias funcionales. En la Figura 3.7 se presenta la distribución bootstrap del estadístico de contraste, mientras que el valor muestral (2935.824) se representa por una línea segmentada. A partir de 1000 remuestras se calcula el p-valor, que resulta ser menor que 0.001, con lo cual se rechaza la hipótesis nula.

Los resultados del análisis exploratorio, incluyendo los métodos de detección de atípicos ya habría adelantado esta conclusión: la detección de los laboratorios 6 y 7 como proveedores de datos inconsistentes con respecto a los demás.

Es importante destacar que este estudio descriptivo de las curvas TG y DSC

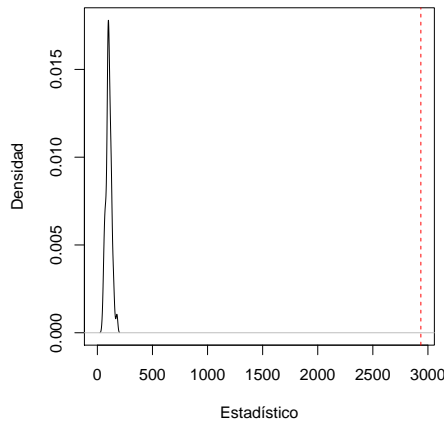


Figura 3.7: (a) Distribución Bootstrap del estadístico del ANOVA Funcional

representa una primera aproximación a la identificación de atípicos, validada empíricamente en el caso de estudio particular que aquí se presenta. Por tanto, el proceso es susceptible de aplicar muchas y variadas mejoras, como por ejemplo, el cálculo de las profundidades de los datos de cada laboratorio con respecto a las curvas de los demás laboratorios. En todo caso, en este mismo capítulo se desarrollan y presentan las nuevas aproximaciones de los estadísticos h y k de Mandell para datos funcionales, mediante los cuales se pretende resolver de forma fiable este problema.

Comparación con la norma ASTM E691

Se ha realizado una comparación entre los resultados obtenidos en el análisis exploratorio funcional y los proporcionados por la metodología tradicional establecida mediante la norma ASTM E691. Para esto se considera una característica específica que está relacionada con la estabilidad térmica, extraída previamente de curvas termogravimétricas (TG), obtenidas a su vez a partir de pruebas experimentales. La característica utilizada es la temperatura inicial de descomposición (IDT), que se define como la temperatura a la que se pierde el 5% de la masa total (primer paso). También se obtuvo otra característica importante de las curvas TG, la pérdida de masa en el segundo proceso de degradación.

Tabla 3.5: Media y desviación estándar del IDT de los 7 laboratorios estudiados

	Media	Desv. Est.	Réplicas
Lab 1	164.4	1.099	15
Lab 2	164.1	0.209	15
Lab 3	164.3	0.371	15
Lab 4	164.2	0.285	15
Lab 5	164.2	0.285	15
Lab 6	164.2	0.979	15
Lab 7	165.7	0.371	15

Es importante destacar que la metodología funcional proporciona estimaciones de la variabilidad y la detección de laboratorios cuyos resultados son inconsistentes, utilizando para ello toda la información presente en las curvas TG; mientras que el estudio ILS univariante proporciona las estimaciones de la media, la varianza y los laboratorios atípicos a partir de una pequeña parte de la información contenida en los datos TG. Esta es la principal diferencia entre los dos enfoques.

En la Tabla 3.5 se presentan la media y la desviación estándar del IDT correspondiente a cada laboratorio, así como el número de réplicas obtenidas por cada laboratorio. Con el fin de identificar los laboratorios que proporcionan resultados atípicos, la norma ASTM E691 propone utilizar los estadísticos h y k de Mandel. Como ya se ha mencionado anteriormente, estos estadísticos se calculan para medir la consistencia de los resultados obtenidos por los laboratorios.

En la Tabla 3.6 se presentan los resultados obtenidos al aplicar la norma ASTM 691 a la característica IDT extraída de las curvas TG. Se puede observar que, al comparar los correspondientes valores de los estadísticos h y k para cada laboratorio con los límites h_c y k_c , los laboratorios 1, 6 y 7 se detectaron correctamente como laboratorios atípicos, es decir, no proporcionan resultados consistentes con respecto a los obtenidos por los laboratorios restantes.

En la Figura 3.8, el laboratorio 7 (correspondiente al uso de una máquina de ensayo SDT con una calibración obsoleta) se identifica como inconsistente utilizando el estadístico h (panel a), mientras que utilizando el estadístico k (panel b), el laboratorio 1 (que utiliza una máquina STA con una calibración errónea) y el laboratorio 6 (SDT con calibración obsoleta) se identifican como atípicos.

Tabla 3.6: Resultados correspondientes a los datos IDT (primer paso) al aplicar la norma ASTM 691

Número de laboratorios (l)	7
Media global \bar{x}	164.4
Desviación estándar Repetibilidad (S_r)	0.6151
Desviación estándar Reproducibilidad (S_R)	0.9651
h_c	2.05, -2.05
k_c	1.44
Número de laboratorios atípicos definidos detectados	3
Laboratorios detectados como atípicos	Lab 1, Lab 6 y Lab 7

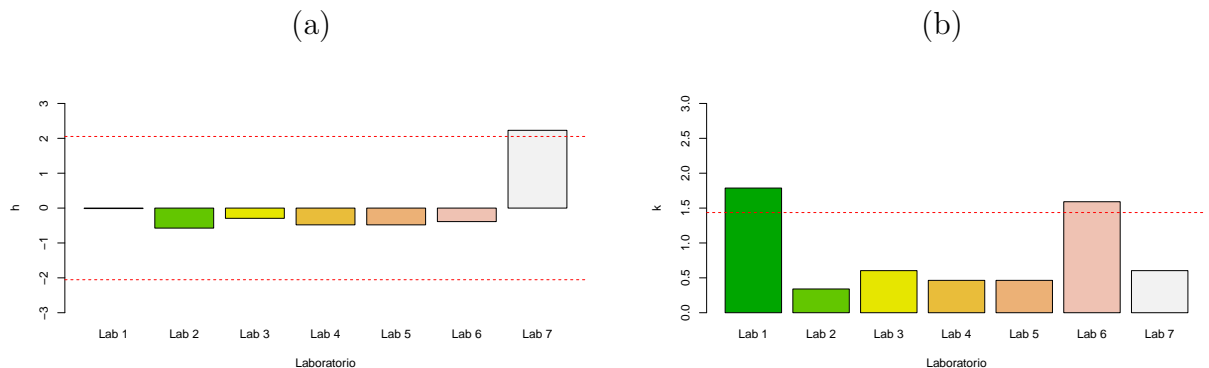


Figura 3.8: (a) Gráfico de barras correspondiente a los valores del estadístico h de cada laboratorio obtenido a partir de medidas de IDT. (b) Gráfico de barras que corresponde a los valores del estadístico k de cada de laboratorio obtenido a partir de medidas de IDT.

Por el contrario, cuando usamos la característica o variable “pérdida de masa en el segundo escalón de degradación”, sólo se logra identificar el laboratorio 7 a través del estadístico k (ver Figura 3.9). Por tanto, los laboratorios mal calibrados 1 y 6 no se detectan. Además, en la Tabla 3.7 se presentan los cálculos realizados para cada iteración del procedimiento indicado en la norma ASTM 691.

Tabla 3.7: Resultados utilizando como característica la pérdida de masa en el segundo proceso de degradación y el ASTM 691

	Iteración 1	Iteración 2
Número de laboratorios (l)	7	6
Media global \bar{x}	30.36	30.28
Desviación estándar Repetibilidad (S_r)	0.0939	0.0613
Desviación estándar Reproducibilidad (S_R)	0.4791	0.3327
h_c	2.05, -2.05	1.92, -1.92
k_c	1.44	1.43
Número de laboratorios detectados	1	0
Laboratorios definidos como atípicos	Lab 7	-

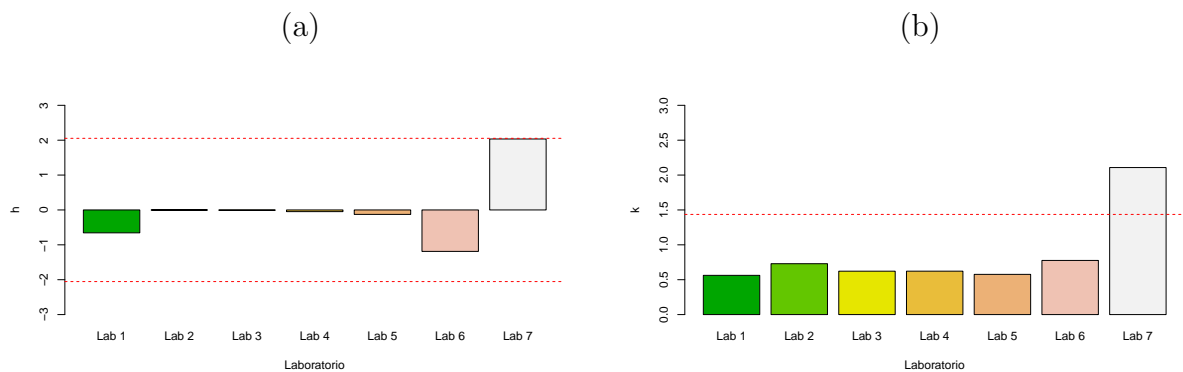


Figura 3.9: (a) Gráfico de barras correspondiente a los valores del estadístico h de cada laboratorio obtenido a partir de la pérdida de masa en el segundo paso. (b) Gráfico de barras que corresponde a los valores del estadístico k de cada laboratorio obtenido a partir de la pérdida de masa en el segundo paso.

A partir de estos resultados podemos concluir que el número de valores atípicos identificados mediante un estudio estándar interlaboratorio depende de la característica extraída de las curvas TG (IDT, la pérdida de masa en el segundo proceso

de degradación, etc.). Por tanto, no todas las características resumen correctamente la información de toda una curva. En este trabajo se proponen técnicas FDA que utilizan la información de toda la curva para identificar los laboratorios no consistentes, en lugar de extraer las características de interés en cada curva experimental. Por otra parte, con la metodología funcional empleada, no es necesario asumir experimentalmente la hipótesis de que los datos estén normalmente distribuidos.

Este procedimiento se puede aplicar en una amplia gama de ILS en los que los resultados pueden ser gestionados como datos funcionales, como es el caso de curvas obtenidas por técnicas de análisis térmico: TG, DSC, análisis mecánico dinámico (DMA), análisis termo mecánico (TMA), y el análisis dieléctrico (DEA).

3.2. Metodología propuesta para la detección de laboratorios atípicos

En esta sección, se presenta la metodología desarrollada para extender la estimación de los estadísticos h y k al caso funcional, $H(t)$ y $K(t)$, y contrastar las hipótesis de repetibilidad y reproducibilidad a través de los estadísticos d_H y d_K . La metodología propuesta se fundamenta en un algoritmo bootstrap implementado para estimar los valores críticos correspondientes a los contrastes de reproducibilidad y repetibilidad c_h y c_k , respectivamente. Los estadísticos de prueba d_H y d_K , se obtienen a partir de los estadísticos $H(t)$ y $K(t)$, permitiendo detectar laboratorios que no proporcionan datos consistentes en un ILS.

Las hipótesis de repetibilidad y reproducibilidad a contrastar se definen a partir del siguiente escenario:

Supongamos que se tomaron un conjunto de observaciones obtenidas por L laboratorios $l = 1, \dots, L$. Se seleccionan n curvas $\{X_1^l(t), \dots, X_n^l(t)\}$ en cada laboratorio l , $l = 1, \dots, L$ (también es válido en el caso de que no todos los laboratorios tomen el mismo número de observaciones). Se calculan los estadísticos funcionales $H_l(t)$ y $K_l(t)$ para cada laboratorio, siendo el punto de partida para contrastar las hipótesis nulas en un estudio R & R.

En primer lugar, la hipótesis nula de reproducibilidad establece que las medias de los resultados de los L laboratorios son iguales, denotándose como:

$$H_0 : \mu_1(t) = \mu_2(t) = \cdots = \mu_L(t), \quad (3.1)$$

donde $\mu_l(t)$, con $l = 1 \dots L$, son las medias poblacionales funcionales para cada laboratorio l . Para contrastar la reproducibilidad de los resultados de los laboratorios, se precisa calcular el estadístico $H(t)$. Este se define como:

$$H_l(t) = \frac{\bar{X}_l(t) - \bar{X}(t)}{S_l(t)}; l = 1, \dots, L,$$

donde $\bar{X}_l(t)$ y $S_l(t)$ son la media y la varianza funcionales calculadas punto a punto y correspondientes al laboratorio l .

Por otro lado, mediante la hipótesis de repetibilidad se afirma que la varianza marginal de los resultados en cada laboratorio l , como se indica a través de la expresión:

$$H_0 : \sigma_1^2(t) = \sigma_2^2(t) = \cdots = \sigma_L^2(t), \quad (3.2)$$

donde $\sigma_l^2(t)$, $l = 1 \dots L$ son las varianzas funcionales teóricas correspondientes a cada laboratorio l . La prueba de repetibilidad está basada en el estadístico $K(t)$, definido como

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

siendo, $\bar{S}^2(t) = \frac{1}{L} \sum_{l=1}^L S_l^2(t)$.

Por otro lado, para contrastar la hipótesis de reproducibilidad, se define el estadístico de prueba d_H como

$$d_H^l = \|H_l(t)\| = \left(\int_a^b H_l(t)^2 dt \right)^{\frac{1}{2}},$$

de forma que los valores más grandes de d_H corresponden a laboratorios no consistentes. De igual forma, para contrastar la hipótesis de repetibilidad, también se define el estadístico $d_K^l = \|K(t)\|$, correspondiéndose los valores grandes de d_K con laboratorios no consistentes.

A continuación se detalla la metodología propuesta para probar si d_H^l y d_K^l son

significativamente grandes, es decir, para contrastar las hipótesis nulas de reproducibilidad y repetibilidad.

El procedimiento bootstrap propuesto pretende reproducir la distribución de estos estadísticos bajo las correspondientes hipótesis nula, definidas por las expresiones 3.1 y 3.2, respectivamente. Suponiendo un nivel de significación α (típicamente $\alpha = 0.01$), el algoritmo consta de los siguientes pasos:

1. Eliminar observaciones atípicas (en este caso curvas), agrupando todas las curvas en un sólo conjunto (hipótesis nula). Para esto se aplica el procedimiento basado en recortar un porcentaje de las curvas con menor profundidad.
2. Se utiliza el bootstrap suavizado con las especificaciones explicadas en el Capítulo 2 para obtener muestras bootstrap de tamaño $n \times L$ a partir del conjunto de datos, una vez que los atípicos hayan sido previamente descartados. 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 los correspondiente estadísticos de prueba, d_H^{l*} y d_K^{l*} , para cada laboratorio $l = 1, \dots, L$.
4. Se aproximan los valores críticos c_H y c_K de los estadísticos de prueba (d_H^{l*} y d_K^{l*}) 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. Adicionalmente, como herramienta gráfica de diagnosis se puede calcular la envolvente de las muestras bootstrap definidas por una norma inferior al valor crítico para el conjunto formado por los estadísticos $H(t)$ y $K(t)$ correspondientes a cada laboratorio. Esta representación gráfica no influye en el cálculo del test.

Es importante señalar que, para cada laboratorio, una vez calculados d_H^{l*} y d_K^{l*} , además de sus respectivos valores críticos, c_H y c_K , se contrastan las hipótesis nulas

de reproducibilidad 3.1 y repetibilidad 3.2, siendo rechazadas si $d_H^l = \|H(t)\| > c_H$, y $d_K^l > c_K$, respectivamente. El procedimiento se repite de forma iterativa hasta que no se detectan laboratorios atípicos.

3.3. Validación del nuevo enfoque FDA

Se consideran dos escenarios de simulación con el fin de evaluar el desempeño de las nuevas aproximaciones FDA para los estadísticos h y k de Mandel. Se obtienen al generar datos cuya media y varianzas presentan desviaciones con respecto a la media y varianza bajo la hipótesis nula (ausencia de laboratorios inconsistentes). En la práctica se sugiere primero realizar el test de repetibilidad (varianzas) y posteriormente el de reproducibilidad (medias).

En consecuencia, el primer escenario consiste en variar la media de un proceso Gaussiano de un laboratorio con respecto a los datos simulados bajo la hipótesis nula, mientras que el segundo se define fijando varios resultados diferentes de la varianza para un laboratorio. Los escenarios mencionados permiten evaluar la potencia de los estadísticos d_H y d_K .

Además, en la sección 3.3.5 se estudia la potencia de la prueba en el caso donde se considera un escenario en el que los resultados de los laboratorios no cumplen la hipótesis de repetibilidad y reproducibilidad simultáneamente.

A partir de los valores simulados se desarrolla un nuevo gráfico de control funcional para los estadísticos $H(t)$ y $K(t)$. El uso de estos gráficos de control proporciona información sobre el intervalo de tiempo/temperatura en el que las curvas se convierten en valores atípicos y, por tanto, sobre el proceso químico o físico y las sustancias implicadas (dependiendo de los datos experimentales analizados).

Cada escenario está compuesto por L laboratorios (cada uno con n muestras). Los resultados (las curvas TG) de cada laboratorio, se simulan a partir de un proceso gaussiano $Y(t) = \mu(t) + \sigma(t)\epsilon(t)$, donde $t \in [0, 1]$ es el intervalo simulado de tiempo/temperatura medido en unidades arbitrarias (u.a.).

La tendencia o media del proceso viene dado por:

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

que corresponde al modelo logístico generalizado, mientras que la varianza determinista es:

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

donde $c_0 = 10^{-6}$ con la finalidad de generar curvas similares al caso de estudio real. Adicionalmente, ϵ es un proceso estacionario de segundo orden definido con media 0 y cuyo elemento de la matriz de covarianza es $\sigma_{ij} = \exp(-|t_i - t_j|/0.3)$.

Para este estudio se considera el modelo logístico generalizado debido a que modela el comportamiento de los datos experimentales (curvas TG) utilizados en este trabajo.

Los escenarios se definen teniendo en cuenta el significado de los parámetros del modelo logístico generalizado. De hecho, el parámetro m representa el tiempo o la temperatura correspondiente a la máxima pendiente de la curva o punto de inflexión, c es el valor inicial de la masa de la muestra (en este caso se fija a 1, en u.a), b se relaciona con la velocidad de cambio (en el caso de las curvas TG, velocidad de degradación térmica), y τ representa el grado de asimetría de las curvas. Para este estudio se considera un proceso de degradación simétrico ($\tau = 1$) definido por una masa inicial de 1 u.a. ($c = 1$), y la velocidad de degradación caracterizada por $b = 10$. Es decir, la media para la hipótesis nula es :

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

3.3.1. Estudio bajo la hipótesis nula del comportamiento de los estadísticos de contraste

La hipótesis nula de repetibilidad se formula en términos de la variabilidad como $H_0 : v_0 = 5$, es decir, los resultados de los laboratorios son consistentes cuando $\sigma(t)^2 = c_0(5 + 5(1 - (\frac{t}{0.5} - 1)^2)^3)$. Por otro lado, la hipótesis nula de reproducibilidad se formula en términos de m . $H_0 : m_0 = 0.5$, es decir, los resultados de los laboratorios son consistentes cuando $\mu(t) = \frac{1}{1 + \exp(10(t - m_0))}$.

En este sentido, se diseña un estudio de simulación con la finalidad de calcular el porcentaje de rechazo bajo la hipótesis nula de repetibilidad y reproducibilidad.

Se realizaron una serie de 1000 simulaciones, cada una compuesta de 500 muestras bootstrap para estimar la distribución de d_H y d_K . Cuando se tienen 7 laboratorios bajo la H_0 , y considerando diferentes valores de $n = 10, 20, 30$ réplicas. En la Tabla 3.8 se muestra el porcentaje de rechazo bajo H_0 para $\alpha = 0.01$ y distintos tamaños muestrales, empleando distintas medidas de profundidad para eliminar atípicos (primer paso del algoritmo descrito al final de la sección 3.2).

Hipótesis	d_H			d_K		
	$n = 10$	$n = 20$	$n = 30$	$n = 10$	$n = 20$	$n = 30$
Moda	1.7	1.3	1.2	3.6	3.8	2.9
FM	1.8	1.5	1.1	3.5	3.8	2.9
RP	1.7	1.6	1.1	3.4	2.8	2.6

Tabla 3.8: Proporción de rechazo (%) del estadístico d_H y d_K , con $\alpha = 0.01$ y $n = 10, 20, 30$.

A un nivel de significación del 1 % utilizando el estadístico d_K para la hipótesis de repetibilidad se rechaza más veces que el nivel fijado, sin embargo cuando el tamaño de la muestra para cada laboratorio aumenta el porcentaje de rechazo disminuye. Por otro lado, para la prueba de reproducibilidad el porcentaje de rechazo para un tamaño de 30 y para las distintas profundidades empleadas (utilizada para la eliminación de atípicos) es similar al nivel de significación de la prueba.

Es importante destacar que el estadístico p -valor es una variable aleatoria distribuida uniformemente en el intervalo $[0, 1]$, $U(0, 1)$, bajo la hipótesis nula (Lehmann, 2005). Por lo tanto, esta hipótesis debe comprobarse con el fin de verificar si el porcentaje de rechazos es la adecuada bajo H_0 para las aproximaciones propuestas por el enfoque FDA para los estadísticos h y k de Mandel.

Por consiguiente, a partir de las simulaciones de la muestras compuestas por diez resultados para 7 laboratorios bajo H_0 se han obtenido los valores observados para d_H and d_K y sus correspondientes p -valores. A partir de estos resultados, mediante la aplicación de la prueba de Kolmogorov-Smirnov a la muestra de los p -valores y fijando un nivel de significación de $\alpha = 0.01$, se obtiene que la hipótesis nula de que la variable aleatoria p -valor están uniformemente distribuidos en $[0, 1]$ no puede ser rechazada. Los resultados de la Figura 3.10 apoyan esta afirmación en la que el

acumulado de la proporción de rechazo de cada prueba se representa en función de diferentes valores de α .

Se puede observar (Figura 3.10.a) que para el estadístico d_H la proporción de rechazo es similar a α . En el caso de la Figura 3.10.b, se puede observar la tendencia es casi igual a la bisectriz, sin embargo, en el caso de valores pequeños de α , se obtienen de rechazo ligeramente mayores a lo esperado.

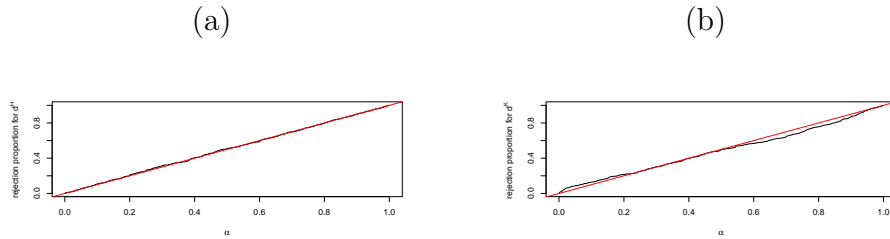


Figura 3.10: (a) Distribución de p -valores para el estadístico d_H bajo la hipótesis nula, (b) p -valores para el estadístico d_K bajo la hipótesis nula.

3.3.2. Hipótesis de reproducibilidad

En el primer escenario de simulación, la hipótesis nula se formula en términos de m . Se define por $H_0 : m_0 = 0.5$, es decir, los resultados de los laboratorios son consistentes cuando $\mu(t) = \frac{1}{1+\exp(10(t-m_0))}$. Por otro lado, la hipótesis alternativa se define por $H_1 : m_1 = m_0(1 + \delta_H)$, donde $-0.005 \leq \delta_H \leq 0.005$. La media teórica para $m_0 = 0.5$ y $m_1 = m_0(1+0.005)$ se presentan en la Figura 3.11.a. Como puede observarse, existen muy ligeras diferencias entre los dos escenarios.

En consecuencia y teniendo en cuenta estudios previos con datos reales, se simulan curvas TG correspondientes a $L = 7$ laboratorios, cada uno con $n = 10$ réplicas. Donde, los datos correspondientes a un grupo de 6 laboratorios son simulados bajo H_0 , mientras que los resultados del séptimo laboratorio se obtienen bajo H_1 .

Considerando que las diferencias entre laboratorios son inducidas, la hipótesis nula de reproducibilidad debe ser probada utilizando el estadístico de prueba d_H a partir de la estimación del estadístico funcional $H(t)$. En este sentido, se realizaron una serie de 1000 simulaciones, cada una compuesta de 500 muestras bootstrap para

estimar la distribución de d_H (Figura 3.11.c). Una vez que se estima la distribución del estadístico d_H mediante el procedimiento bootstrap, se obtiene el valor crítico de la prueba, c_H , correspondiente a $\alpha = 0.01$. Se muestra en la Figura 3.11.c (línea de puntos).

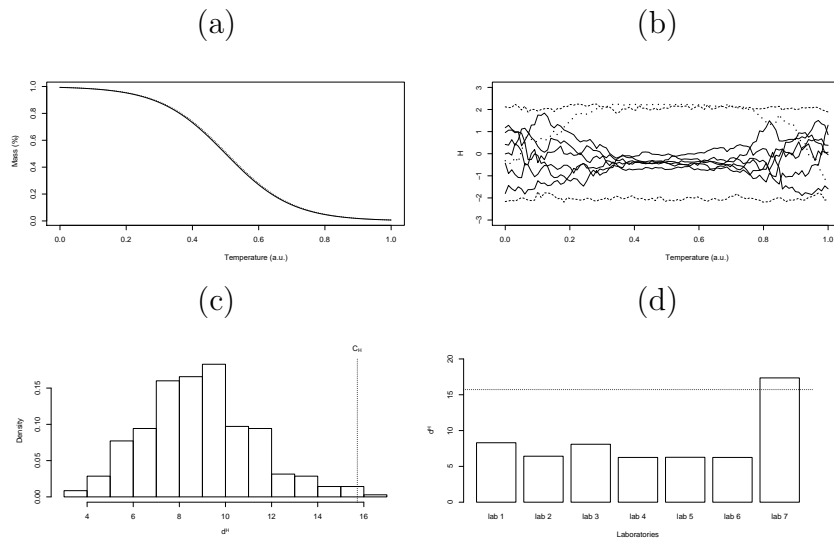


Figura 3.11: (a) Medias teóricas para curvas TG simuladas bajo la hipótesis nula $H_0 : m_0 = 0.5$ (línea continua) y la alternativa $H_1 : m_1 = m_0(1 + \delta_H)$ ($\delta_h = 0.005$, la línea a trazos), (b) el estadístico funcional $H(t)$ para cada uno de los 6 laboratorios simulados bajo la nula (línea continua) y un laboratorio bajo la alternativa (línea a trazos), (c) la distribución estimada por remuestreo bootstrap del estadístico d_H , y (d) el estadístico d_H para cada muestra de laboratorio.

Una de las ventajas de los h y k de Mandel son sus representaciones gráficas, que permiten identificar de forma intuitiva los laboratorios inconsistentes. Por lo tanto, la actual propuesta FDA pretende reproducir e incluso completar los resultados de los gráficos correspondientes al enfoque univariante. En la Figura 3.11.b, se muestran las realizaciones de la muestra de los estadísticos $H_l(t)$ con $l = 1, 2, \dots, 7$.

Los correspondientes d_H^l son mostrados en el gráfico de barras de la Figura 3.11.d, siguiendo el mismo estilo que el gráfico univariado. En esta memoria se propone una nueva forma gráfica de realizar la prueba de hipótesis, donde c_H es la línea horizontal punteada que determina el borde por encima del cual un laboratorio podría considerarse atípico.

Por ejemplo, en una de las 1000 simulaciones, el séptimo laboratorio (cuyos datos han sido simulados bajo H_1) se identifica con éxito como un atípico, $d_H^7 > c_H$ (Figura 3.11.d). La Figura 3.11.b, permite identificar los intervalos de tiempo/temperatura en los que hay inconsistencia. Se puede observar que la sección central de la curva $H_7(t)$, mostrada como una línea segmentada, está fuera de la región definida por los cuantiles funcionales (mostrados en líneas segmentadas) obtenidos a partir de c_H .

Es importante señalar que el enfoque funcional, en la Figura 3.11.b, también proporciona información sobre el intervalo tiempo/temperatura donde los resultados del laboratorio 7 son diferentes de los otros. En este caso, este intervalo corresponde a la zona del escalón TG, donde la pendiente de la curva varía (relacionado con un proceso de degradación simulado). Esta información podría ser útil no sólo para las curvas de análisis térmico sino para analizar resultados obtenidos por otras técnicas analíticas.

3.3.3. Hipótesis de repetibilidad

El segundo escenario de simulación es definido para probar la nueva metodología de detección de laboratorios atípicos teniendo en cuenta la variabilidad intra-laboratorio. Así, la hipótesis nula se formula en términos de la variabilidad, modificando el parámetro v definido anteriormente. Se considera $H_0 : v_0 = 5$, es decir, los resultados de los laboratorios son consistentes cuando $\sigma(t)^2 = c_0(5+5(1-(\frac{t}{\sigma^5}-1)^2)^3)$.

Por otra parte, la hipótesis alternativa se define por $H_1 : v_1 = v_0(1 + \delta_K)$, donde $0 \leq \delta_K \leq 2$. La varianza teórica para $v_0 = 5$ y $v_1 = v_0(1 + 2)$ se presentan en la Figura 3.12.a. Las variaciones han sido elegidas teniendo en cuenta estudios anteriores (Tarrío-Saavedra *et al.*, 2014). Como en el caso del primer escenario, los datos correspondientes a los 6 primeros laboratorios se simulan bajo $H_0 : v_0 = 5$, mientras que las curvas sintéticas del séptimo laboratorio se obtienen asumiendo la hipótesis alternativa, $H_1 : v_1 = v_0(1 + 2)$.

Teniendo en cuenta que se introducen diferencias intra-laboratorios, la hipótesis nula de repetibilidad se prueba utilizando los estadísticos propuestos d_K y $K(t)$. Se realizaron 1000 simulaciones, cada una compuesta de 500 muestras bootstrap para estimar la distribución del estadístico d_K (Figura 3.12.c).

El valor crítico de la prueba, c_K , también se estima para $\alpha = 0.01$ (Figura 3.11.c,

definido por una línea punteada). Además, las realizaciones de la muestra (con $l = 1, 2, \dots, 7$) del estadístico funcional $K_l(t)$ se muestran en la Figura 3.12.b, mientras que los correspondientes d_K^l se representan gráficamente en la Figura 3.12.d. Se proporciona una herramienta gráfica para realizar la hipótesis, donde c_K es una línea horizontal punteada por encima de la cual un laboratorio podría ser considerado como anómalo, como es el caso del séptimo laboratorio, $d_H^7 > c_H$ (Figura 3.12.d).

Como en el escenario anterior, se tiene un ejemplo de las 1000 simulaciones, donde se da el mismo resultado. Se puede observar en la Figura 3.12.b, donde la sección central de la curva $K_7(t)$, mostrada como una línea punteada, está fuera de la región definida por los cuantiles funcionales (mostrados en líneas discontinuas) correspondiente a c_K . Este intervalo corresponde a la región que rodea al punto de inflexión de la curva TG, donde la varianza es mayor.

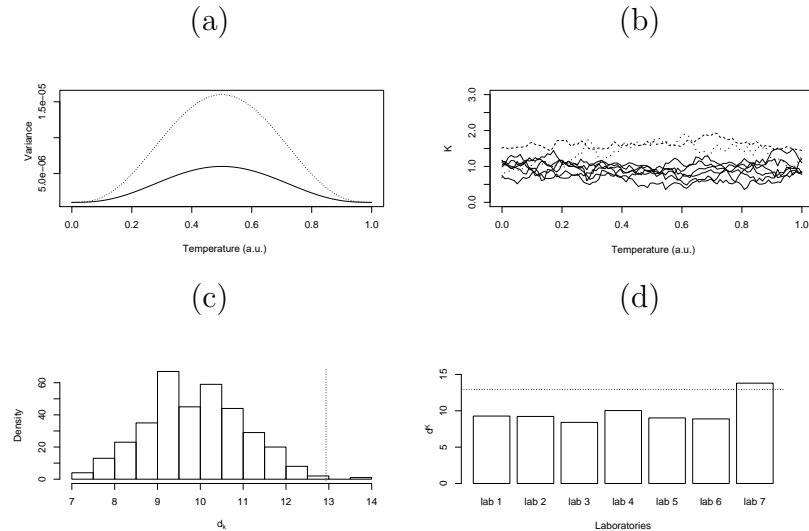


Figura 3.12: (a) Varianzas teóricas para las curvas TG simuladas bajo la hipótesis nula $H_0 : v_0 = 5$ (línea continua) y la alternativa, $H_1 : v_1 = v_0(1 + \delta_K)$ ($\delta_k = 2$, la línea a trazos), (b) el estadístico funcional $K(t)$ para cada uno de los 6 laboratorios simulados bajo la nula (línea continua) y un laboratorio bajo la alternativa (línea a trazos), (c) la distribución estimada por remuestreo bootstrap del estadístico d_K , y (d) el estadístico d_k para cada muestra de laboratorio.

3.3.4. Efectos del tamaño de la muestra (número de resultados en cada laboratorio)

Para cada escenario simulado, se estudia el efecto del tamaño de la muestra para cada laboratorio ($n = 10, 20$) en el desempeño de las nuevas pruebas, utilizando un nivel de significación de $\alpha = 0.01$.

El objetivo es evaluar la consistencia de las dos aproximaciones FDA de los contrastes h y k de Mandel. La Figura 3.13 muestra la proporción de rechazos bajo la hipótesis nula para ambas estadísticas, d_H en el panel (a) y d_K en el panel (b).

Las líneas de la Figura 3.13.a y 3.13.b muestran la potencia de d_H y d_K , respectivamente, correspondientes a dos tamaños de muestra: líneas segmentadas para $n = 20$ y líneas continuas para $n = 10$.

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. Esto está de acuerdo con las recomendaciones de ASTM para el caso univariado. En cualquier caso, incluso utilizando muestras pequeñas ($n = 10$), las aproximaciones FDA propuestas muestran un buen desempeño.

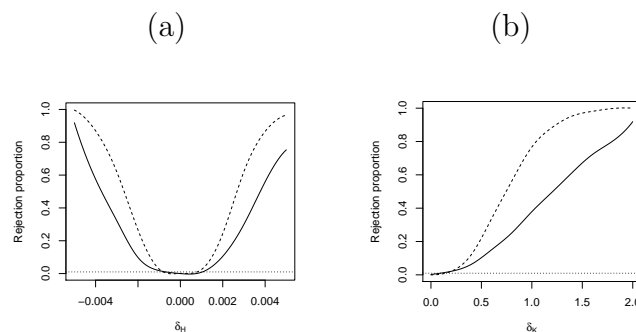


Figura 3.13: (a) Curvas de potencia (proporción de rechazo) para el estadístico d_H correspondiente a $n = 10$ y $n = 20$ muestras por laboratorio, (b) curvas de potencia para el estadístico d_K correspondiente a $n = 10$ y $n = 20$ muestras por laboratorio. Las líneas segmentadas corresponden a $n = 20$, mientras que las líneas continuas corresponden a $n = 10$.

3.3.5. Incumplimiento de la hipótesis de repetibilidad y reproducibilidad simultáneamente

En las anteriores secciones se han diseñado escenarios de simulación que permiten evaluar el desempeño de la metodología cuya finalidad es contrastar si existen diferencias en los resultados entre laboratorios (reproducibilidad) y dentro de ellos (repetibilidad) de forma separada. En esta sección se plantean escenarios en donde se considera un cambio en la media $H_1 : m_1 = 0.502$, y en la varianza $H_1 : v_1 = v_0(1 + \delta_K)$, con $\delta_K = 0.2, 0.4, 1, 2, \dots, 10$ de forma simultánea. Para calcular el porcentaje de rechazo bajo el incumplimiento de la hipótesis de reproducibilidad, p_m , y repetibilidad, p_v , y las dos simultáneamente p_{vm} , cuando se considera un $\alpha = 0.01$ y $n = 10$, se ha generado una curva bajo cada uno de estos escenarios.

En la Tabla 3.9, se presentan los resultados de contrastar la hipótesis de repetibilidad ($K(t)$), porcentaje de rechazo, tanto para el caso donde se tiene un escenario donde se incumple solo la hipótesis de repetibilidad o ambas. En estos dos casos al aumentar δ_K , el porcentaje de rechazo p_{vm} y p_v aumenta.

H_a		K(t)	
m_1	v_1	p_v	p_{vm}
0.502	6	2	1.6
0.502	7	5.8	4
0.502	10	6.6	5.8
0.502	15	13.6	12.4
0.502	20	32.4	31.6
0.502	25	47.6	46.2
0.502	30	69.8	69.8
0.502	35	73.8	72.8
0.502	40	76.2	77.6
0.502	45	80.6	83.2
0.502	50	89.6	89.6
0.502	55	91.6	92.2

Tabla 3.9: Porcentaje de rechazo p_v y p_{vm} del estadístico d_K , con $\alpha = 0.01$ y $n = 10$.

Para estos mismos escenarios, se pudo observar que el estadístico $H(t)$, cuando se realizan cambios menores en la varianza tiene en promedio $p_m = 82.0$ y $p_{vm} = 82.6$, sin embargo cuando estos cambios van en aumento tanto p_m y p_{vm} no pasan del 11%

de rechazo.

3.3.6. Comparación entre el enfoque FDA y el escalar

En la actualidad los estudios ILS se realizan utilizando enfoques escalares, es decir, se mide y analiza una variable escalar de interés. Las pruebas univariadas para detectar los valores atípicos son una de las herramientas más empleadas en los ILS para detectar aquellos laboratorios que proporcionan resultados inconsistentes.

En este sentido, hay varias pruebas escalares muy populares entre las cuales destacan los estadísticos de Mandel. Sin embargo, hasta el momento no existen extensiones FDA para realizar detección de resultados atípicos de laboratorios cuando los datos experimentales son funcionales, exceptuando por la propuesta desarrollada en esta memoria.

Así, en esta sección se compara la nueva metodología con respecto al enfoque escalar con el objetivo de comprobar si el uso de las extensiones propuestas del enfoque FDA proporcionan ventajas con respecto al caso escalar cuando los datos son curvas (funcionales). Por lo tanto, para poder calcular los h y k de Mandel univariados, se extrae una característica representativa de la curva. En el caso de las curvas TG se estudia habitualmente la temperatura/tiempos hasta la pérdida del 5wt % en peso de la masa inicial (denominado temperatura/tiempos de descomposición inicial, IDT). Para este estudio se consideran las temperaturas/tiempos en las que se alcanza una pérdida del 5wt % y 10wt %.

La consistencia de las pruebas, tanto para el enfoque univariado y funcional, se compara utilizando dos tamaños de muestra, $n = 10$ y $n = 20$. Como puede observarse en la Figura 3.14, la potencia de la prueba univariada depende de la característica que se ha extraído. De hecho, para ambos tamaños de muestra, la potencia de la prueba univariate h es menor pero se puede observar una tendencia que esta va creciendo en función de la característica extraída.

En la Figura 3.14.a, la potencia más alta corresponde a la extensión FDA d_H . Por lo tanto, esto apoya la aplicación del enfoque FDA, teniendo en cuenta el desempeño del caso univariante depende en gran medida de la característica elegida. Además, la figura 3.14.b muestra que la potencia de prueba del estadístico univariante h para el 5wt % en peso y el 10wt %, con $n = 20$, son inferiores a la potencia del enfoque

FDA.

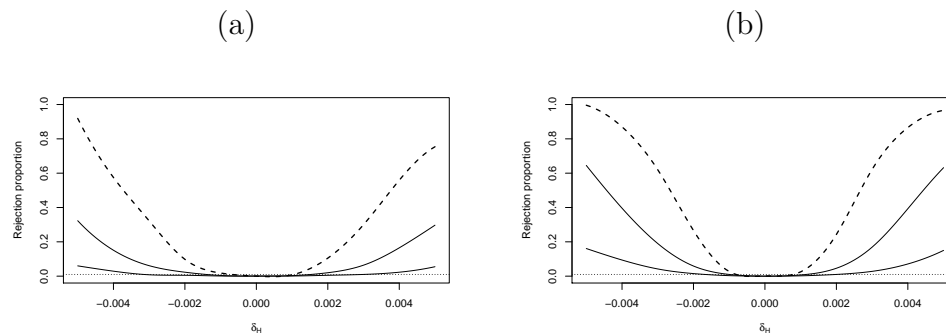


Figura 3.14: (a) Potencias de prueba correspondientes al caso univariante (5tw % en peso y 10tw % en peso) en trazos continuos y la aproximación FDA para el estadístico h en trazo discontinuo, con $n = 10$, (b) Potencias de prueba correspondientes al caso univariado (5tw % en peso y 10tw % en peso) en trazos continuos y un enfoque FDA del estadístico h en trazo discontinuo, con $n = 20$.

Cuando se estudia el estadístico k de Mandel, se obtiene de igual forma que el enfoque FDA proporciona el mejor rendimiento en términos de potencia para todos los tamaños de muestra. De hecho, la Tabla 3.10 muestra la proporción de rechazo para cada caso a partir de la hipótesis nula, δ_k , correspondiente al estadístico k univariante aplicado al tiempo/temperatura correspondiente a la pérdida del 5tw % y 10tw % en peso y a la extensión FDA.

Como puede observarse, la potencia del enfoque FDA es bastante más alta que la correspondiente al estadístico k univariante, independientemente del tamaño de la muestra y la característica extraída (en el caso univariante). Resumiendo, el uso de la extensión FDA se recomienda cuando la variabilidad intra-laboratorio es analizada.

3.4. Aplicación a datos reales: curvas termogravimétricas

Además de los estudios de simulación, la aplicación a datos reales obtenidos mediante técnicas experimentales es necesaria para conocer el desempeño y la utilidad de la metodología propuesta. En este sentido, el procedimiento funcional descrito

n	n = 10			n = 20		
	5wt %	10wt %	FDA	5wt %	10wt %	FDA
0.2	0.001	0.001	0.021	0.000	0.000	0.021
0.4	0.002	0.001	0.063	0.000	0.000	0.121
0.6	0.003	0.002	0.150	0.003	0.001	0.336
0.8	0.003	0.002	0.251	0.003	0.002	0.561
1	0.004	0.005	0.379	0.004	0.006	0.767
1.2	0.004	0.005	0.496	0.005	0.007	0.884
1.4	0.004	0.006	0.614	0.008	0.007	0.953
1.6	0.004	0.009	0.722	0.009	0.011	0.981
1.8	0.006	0.009	0.802	0.009	0.012	0.998
2	0.008	0.012	0.919	0.009	0.014	1

Tabla 3.10: Proporción de rechazo estadístico d_K (con $\alpha = 0.01$) en comparación con la proporción de rechazo del estadístico k univariante cuando se calcula a partir de las variables escalares tiempo/temperatura a pérdida del 5wt % y 10wt % en peso de la masa de la muestra.

anteriormente se aplica a datos experimentales generados a partir de un análisis termogravimétrico del oxalato de calcio, presentado en la sección 2.2.

En este ejemplo, hay 4 laboratorios que proporcionan resultados similares, mientras que los laboratorios 1, 6 y 7 proporcionan diferentes resultados debido a causas diferentes. Estos datos han sido profundamente estudiados en Naya *et al.* (2014). Como se muestra a continuación, los laboratorios que proporcionan resultados inconsistentes son identificados con éxito mediante la aplicación de las extensiones FDA h y k de Mandel. Este enfoque tiene como objetivo utilizar la información de toda la curva, evitando el paso de extracción de características.

El primer paso consiste en estimar los estadísticos funcionales $H(t)$ (3.16.a) y $K(t)$ (3.17.a), con sus respectivos estadísticos de prueba d_H y d_k y, posteriormente, compararlos con respecto a los valores críticos c_H y c_K , definidos como los cuantiles correspondientes a $\alpha = 0.01$ (como es habitual en los estudios ILS), y obtenidos por el procedimiento bootstrap desarrollado en la sección anterior. En el caso de que algún laboratorio l cumpla $d_l^H > c_H$ y/o $d_l^K > c_K$, se identifica como un valor atípico. Además, es importante señalar que la identificación de valores atípicos en ILS es un proceso iterativo, es decir, el procedimiento se aplica consecutivamente hasta que

ningún laboratorio sea identificado como inconsistente.

Cabe recalcar que antes de realizar las réplicas bootstrap se deben eliminar las curvas atípicas, es decir las curvas cuyas profundidades se encuentren entre el 1% de las curvas con profundidad más baja. Así, el 1% de las curvas TG (obtenidas por todos los laboratorios) con la profundidad modal funcional más baja son descartadas. En el presente caso de estudio, se han eliminado curvas TG correspondientes al laboratorio 7. Este es un resultado esperado ya que el laboratorio 7 es uno de los verdaderos valores atípicos que pretendemos detectar.

La Figura 3.15.a muestra el d_H obtenido por el procedimiento bootstrap (utilizando $B = 500$ remuestras), además del valor crítico de c_H correspondiente a $\alpha = 0.01$. La Figura 3.15.b proporciona el gráfico del estadístico de prueba d_H correspondiente a la primera aplicación del enfoque funcional (primera iteración), donde el laboratorio 7 se ha detectado correctamente como anómalo. En la segunda y tercera iteración de la metodología FDA, los laboratorios 1 (STA instrumento con calibración antigua) y 6 (instrumento SDT con calibración antigua) también se detectan con éxito como laboratorios atípicos.

La Figura 3.15.c muestra la distribución del estadístico de prueba d_H en la cuarta iteración, es decir, en la cuarta aplicación del procedimiento FDA. El correspondiente estadístico de prueba d_H se muestra en la Figura 3.15.d, en la cual ningún laboratorio se detecta como atípico. Por lo que se interrumpe el proceso iterativo. Todos los laboratorios atípicos, desde el punto de vista de la reproducibilidad, se detectan con éxito utilizando el enfoque FDA propuesto.

La hipótesis de reproducibilidad también se puede estudiar usando directamente el estadístico funcional H , como se muestra en la Figura 3.16. Al aplicar un proceso iterativo con $\alpha = 0.01$, el laboratorio 7 es detectado como un valor atípico (Figura 3.16.a). El estadístico H proporciona información adicional sobre el intervalo de temperatura en el que el laboratorio 7 se convierte en atípico: la región correspondiente al primer proceso de degradación del oxalato de calcio está fuera de las bandas de confianza del 99%. Los laboratorios 1 y 6 son detectados como valores atípicos en la segunda y tercera iteración de la metodología. La Figura 3.16.b muestra los estadísticos H correspondientes a cada laboratorio en la cuarta iteración. Estos laboratorios corresponden al grupo de laboratorios consistentes, como se esperaba.

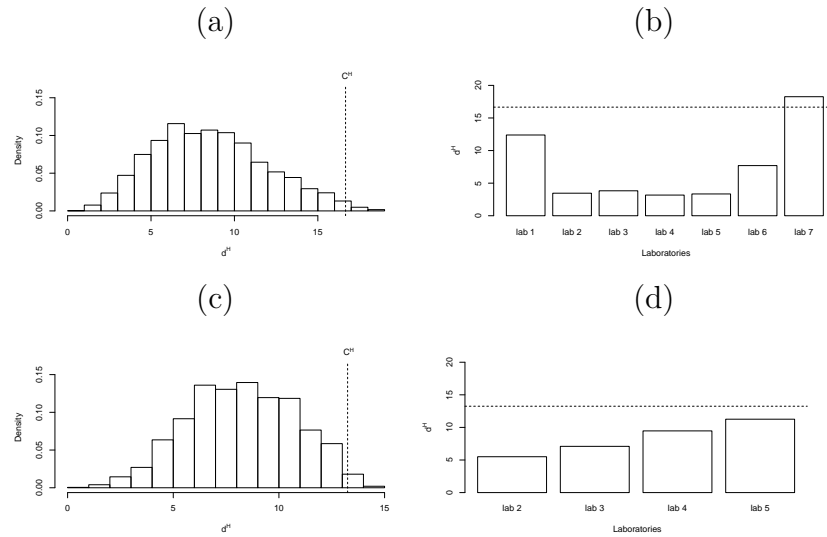


Figura 3.15: (a) El histograma del estadístico de prueba d_H y c_H para $\alpha = 0.01$ (primera iteración), (b) Los estadísticos d_H para la muestra de los laboratorios (primera iteración), El histograma del estadístico de prueba d_H y c_H para $\alpha = 0.01$ (cuarta iteración), (b) Los estadísticos d_H para la muestra de los laboratorios (cuarta iteración).

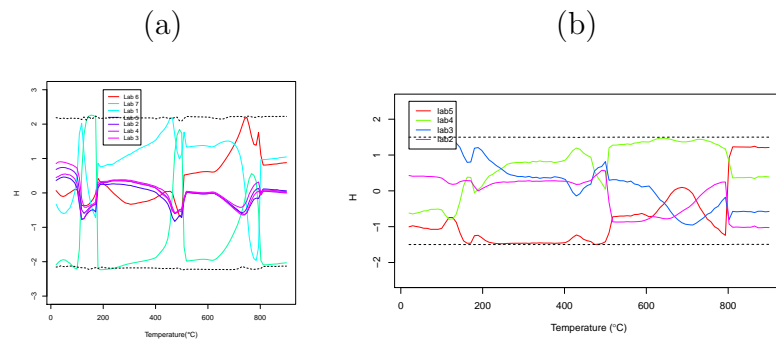


Figura 3.16: (a) Estadístico funcional $H(t)$ con bandas de confianza del 99 % (primera iteración), (b) Estadístico funcional $H(t)$ con bandas de confianza del 99 % (cuarta iteración).

El mismo procedimiento se aplica para contrastar la hipótesis de repetibilidad, pero utilizando el estadístico funcional K y el estadístico de prueba d_K . El laboratorio 6 se detecta como un laboratorio atípico a un nivel de significación de $\alpha = 0.01$. Este resultado se puede observar en la Figura 3.17.a y en la Figura 3.17.b.

Las regiones correspondientes a la primera, segunda y tercera etapas de degradación del oxalato de calcio están fuera de las bandas de confianza de 99%. El proceso iterativo se detiene en esta iteración porque no se detectan laboratorios atípicos en la segunda aplicación de la metodología (ver Figura 3.17.c) y Figura 3.17.d).

Aunque los laboratorios 1 y 7 son candidatos para ser atípicos (Figura 3.17.c), no hay suficientes evidencias con $\alpha = 0.01$ (Figura 3.17). Por lo tanto, a un nivel de confianza de 99% (comúnmente utilizado en ILS) sólo el laboratorio 6, donde una calibración demasiado antigua se había utilizado, se detecta como anómalo, teniendo en cuenta la variabilidad intra-laboratorio (repetibilidad).

En conclusión, los laboratorios 1, 6 y 7 han sido debidamente identificados como atípicos, es decir proporcionan resultados inconsistentes en comparación con los cuatro laboratorios que no presentan anomalías. El estadístico funcional H y el estadístico de prueba d_H han detectado los laboratorios 1, 6 y 7 teniendo en cuenta la hipótesis de reproducibilidad, mientras que los estadísticos K y d_K detectan al laboratorio 6 en el contexto de la hipótesis de repetibilidad. Por lo tanto, las extensiones de los estadísticos h y k al caso FDA propuestos proporcionan una nueva forma útil de realizar la detección de valores atípicos en un ILS cuando se trata de datos funcionales, sin el paso adicional de extracción de características representativas.

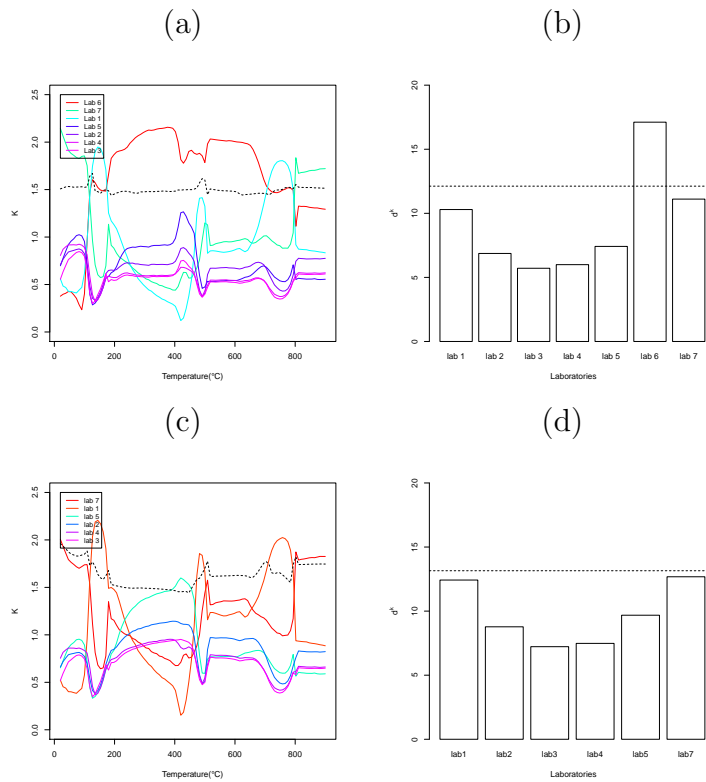


Figura 3.17: (a) Estadístico funcional K para cada laboratorio y bandas de confianza para $\alpha = 0.01$ (primera iteración), (b) estadístico de prueba d_K para cada laboratorio y nivel crítico correspondiente a $\alpha = 0.01$ (primera iteración), (c) Estadístico funcional K para cada laboratorio y bandas de confianza $\alpha = 0.01$ (segunda iteración), y (d) estadístico de prueba d_K para cada laboratorio y nivel crítico correspondiente a $\alpha = 0.01$ (segunda iteración).

Capítulo 4

Propuesta de una metodología de gráficos de control basada en datos funcionales

4.1. Introducción

La detección de anomalías en la industria, y el control de la calidad de productos y servicios en general, ha venido realizándose mediante la aplicación de gráficos de control univariantes y multivariantes. Sin embargo, el problema de la monitorización continua de datos, resultado de los avances en sensorización, obliga a la utilización de herramientas cada vez más sofisticadas que tengan en cuenta la presencia de autocorrelación en las variables críticas para la calidad de un proceso.

En los últimos años, se han aportado soluciones muy diversas como la modificación y aplicación de gráficos EWMA (Chao-Wen y Reynolds Jr, 1999), gráficos de control basados en el ajuste de modelos ARIMA (Alwan y Roberts, 1988), uso de gráficos de control para perfiles entendidos como el control de los parámetros que definen la relación de varias variables críticas para la calidad (Chipman *et al.*, 2010; Noorossana *et al.*, 2011; Qiu *et al.*, 2010; Shiau *et al.*, 2009), y la aplicación de técnicas de aprendizaje máquina (redes neuronales y máquinas de vector soporte) o series de tiempo para la detección de anomalías que se escapan de la idea de los gráficos de control. El avance de este tipo de técnicas se debe en gran parte al hecho de que

los gráficos de control usuales muchas veces no son aplicables al nuevo paradigma de datos obtenidos en la industria.

Muchos de estos nuevos datos, en general curvas, pueden estudiarse como datos funcionales. Este es el caso de los consumos energéticos, temperaturas, humedades relativas, cantidad de CO₂, entre otras variables, medidas en todo tipo de edificios. Estos nuevos datos, que actualmente caracterizan los procesos industriales, requieren el desarrollo de soluciones innovadoras por parte de los investigadores en control estadístico de la calidad a partir de la aplicación de técnicas de análisis de datos funcionales (FDA). Hasta la fecha, se han publicado pocos trabajos (en relación con la importancia del tema y lo frecuente de la monitorización de este tipo de datos) relacionados con el control de la calidad de procesos cuando los datos a monitorizar son variables CTQ funcionales. Entre las investigaciones más destacadas en este contexto, se encuentra el trabajo de Colosimo y Pacella (2010), en el que se realiza un estudio de comparación del desempeño de diferentes gráficos de control para monitorizar datos funcionales, que en la literatura del control estadístico de la calidad se identifican como perfiles. Según Megahed y Jones-Farmer (2013) y Woodall y Montgomery (2014), la monitorización de perfiles engloba el conjunto de técnicas estadísticas (por lo común multivariantes) aplicadas para el control de procesos cuando estos están definidos por la relación funcional entre dos variables. Por otro lado, en Sheu *et al.* (2013) también se propone un gráfico de control para la monitorización de datos funcionales basado en remuestreo bootstrap, mientras que García (2011) y Rodrigo (2013) son dos completas monografías en las que se adaptan los conceptos de gráficos de control para perfiles al contexto de los datos funcionales.

En este capítulo, se pretende dar una solución alternativa a la detección de anomalías relacionadas con la eficiencia energética en edificios, en particular en superficies comerciales como lo son las diversas tiendas de empresas textiles. Una de las principales motivaciones de este capítulo es, por tanto, la resolución de un problema real de este tipo de empresas, en particular de las compañías instaladoras de iluminación, ofimática y ventilación y aire acondicionado (HVAC), como es el caso de FRIDAMA, SL. De hecho, el grupo formado por FRIDAMA (instalaciones), Σqus (plataformas web de gestión de datos masivos) y Nerxus (consultoría estadística para datos en el sector de la energía), plantearon el problema a resolver, definido a su

vez por datos monitorizados continuamente en el tiempo que pueden tratarse como perfiles o datos funcionales. A partir de esta necesidad, en este capítulo se proponen metodologías para construir gráficos de control que permitan controlar los procesos antes mencionados. El control de un proceso mediante gráficos de control consiste de dos fases. La Fase I de estabilización o calibración del proceso y la Fase II de monitorización del proceso.

Para la Fase I, se propone un gráfico de control a partir de la medida de profundidad de un dato funcional y la idea de detección de atípicos, y en la Fase II se propone un gráfico de control no paramétrico de rangos, basado también en el cálculo de la profundidad de datos funcionales, para monitorizar el proceso de interés.

En la siguiente subsección se introducen de manera muy general algunas alternativas de gráficos de control utilizados cuando se incumplen las hipótesis de independencia y normalidad de la variable CTQ, incluyendo entre ellos los gráficos de rangos. Además, se explica en que consisten las fases de control de un proceso y la importancia de la fase de calibrado (Fase I) en la construcción de un gráfico de control. Por otro lado, se exponen los procedimientos para construir gráficos de control correspondientes a la Fase I y a la Fase II, además de mostrar los resultados de su desempeño mediante un estudio de simulación que comprende muy diversos escenarios. La metodología propuesta ha sido desarrollada y programada en **R**, a través de diversas funciones que serán integradas en la librería “Quality Control Review”, **qcr**, implementada por el autor de la presente tesis doctoral.

La metodología propuesta se aplicó mediante una base de datos correspondiente a una instalación HVAC real, cuyas anomalías y sus causas asignables están identificadas por los encargados de su mantenimiento. Como resultado de la aplicación de la herramienta de gráficos de control para datos funcionales propuesta, se detectaron correctamente todos los estados anómalos existentes.

Además, en este trabajo se propone una nueva forma de visualización de los gráficos de control para datos funcionales, incluyendo los datos originales y las curvas correspondientes a los límites de control estimados. Esta representación permite a los usuarios y encargados de mantenimiento relacionar cada anomalía con una causa asignable de forma intuitiva.

Los resultados del estudio de simulación y la aplicación a los datos reales muestran

que esta metodología de gráficos de control es una herramienta útil para la detección de anomalías cuando el proceso está definido por datos funcionales, en particular las curvas diarias de consumo energético en superficies comerciales.

4.1.1. Alternativas del CEC cuando se incumplen las hipótesis básicas de los gráficos de control

En los últimos años se han llevado a cabo diversos análisis de la robustez de los modelos ante alteraciones en la hipótesis de normalidad, explorando el uso de contrastes distintos a los habituales cuando no se cumpla esta hipótesis. En esta línea, se han venido planteando nuevos métodos no paramétricos de control estadístico de la calidad, entre los que se encuentran nuevos gráficos de control. Son de especial interés los trabajos de Regina Liu, que ha desarrollado los gráficos de control r , Q , y S basados en la profundidad de datos y los rank o rangos (Liu, 1990, 1995; Liu y Tang, 1996). A este respecto, es importante destacar que los trabajos de Liu han marcado tendencia, haciendo que gran parte de las alternativas no paramétricas desarrolladas estén basadas en el concepto de rangos. Por otro lado, se ha de subrayar el hecho de que una de las líneas de investigación más importantes del CEC, los gráficos de control de perfiles, están basados, en muchos casos, en la aplicación de modelos de regresión no paramétricos o semiparamétricos (Qiu *et al.*, 2010; Shiau *et al.*, 2009). Es de reseñar también el empleo de técnicas de remuestreo para el cálculo de límites de control natural de diferentes tipos de gráficos de control (Chatterjee y Qiu, 2009; Jones y Woodall, 1998). El trabajo de Chakraborti (Chakraborti, 2011) constituye una completa monografía acerca de las tendencias actuales para la construcción de gráficos de control.

A pesar de sus ventajas en la reducción de los supuestos relacionados con la distribución de probabilidad, los gráficos de control no paramétricos no están ganando terreno en el ámbito industrial y empresarial. Según se apunta en Woodall y Montgomery (2014), esto podría deberse a diversos factores entre los que se encuentran la falta de software específico, tanto comercial como libre, la falta de formación general en estadística no paramétrica que genera inseguridad y desconfianza en los usuarios y la falta de textos de referencia contrastados para la aplicación de métodos no paramétricos en CEC. No obstante, la actividad investigadora en este campo no deja

de crecer, valga como ejemplo el trabajo de Qiu y Li (2011).

Otra de las hipótesis de partida de los gráficos de control es el de independencia de la observaciones. Los datos monitorizados continuamente en el tiempo por diferentes sensores suelen presentar un nivel variable de autocorrelación (mayor cuanto más cercanas estén las observaciones en el tiempo). La aplicación de técnicas estándar cuando se viola la hipótesis de independencia suele traer consigo la detección de un número inaceptable de falsas alarmas (Barbeito *et al.*, 2017). Por ello, el desarrollo y análisis de técnicas de eliminación de la autocorrelación muestral está plenamente justificada. Dentro de estas técnicas, cabe destacar la posible aplicación de modelos de series temporales para eliminar la correlación existente entre las observaciones. Dentro de estas técnicas, la más extendida es la aplicación de modelos de series temporales (ARMA, ARIMA, etc.) para eliminar la correlación existente entre las observaciones y el posterior monitorizado de los residuos mediante gráficos de control (Barbeito *et al.*, 2017; Ledolter y Bisgaard, 2011; Montgomery, 2009).

Por otro lado, Box *et al.* (2011) y Del Castillo (2006) proponen la combinación de gráficos de control con algoritmos de ajuste. Finalmente, las monografías de Knoth y Schmid (2004), Prajapati y Singh (2012) muestran y describen las líneas de investigación más relevantes en el control y monitorizado de datos autocorrelados.

Es importante destacar que este tipo de datos están relacionados con los gráficos de control para datos funcionales o FDA, que serán introducidos más adelante. De hecho, el uso de técnicas FDA permite considerar la autocorrelación de los datos y, mediante técnicas de remuestreo, prescindir de supuestos paramétricos sobre la tendencia y la dependencia.

4.1.2. Construcción de un gráfico de control: Fase I, Fase II, muestra de calibrado y monitorizado

En la Fase I, se estabiliza el proceso, es decir, se deja el proceso bajo control. Esto implica que no hay otras causas asignables presentes en el proceso salvo aquellas debidas a la propia aleatoriedad. Esto es equivalente a afirmar que el proceso permanece estable, permaneciendo invariables los parámetros de la distribución de probabilidad de la característica CTQ. Esto permite estimar los límites de control natural para la variable que describe la calidad del proceso. Para la construcción

del gráfico de control en esta fase los límites de control se estiman utilizando una muestra preliminar. Es por tanto en esta fase, en la que se identifican las principales causas asignables de variación y se toman medidas para corregirlas, hasta estimar los límites de control natural de la variable CTQ correspondiente al proceso bajo control. Los límites de control son completamente diferentes de los límites de especificación (aquellos aceptados por el cliente o fijados por los responsables de calidad, que representan la consigna o target).

Para el caso univariante, los límites de control se calculan como un intervalo de confianza. Se suelen tomar aquellos que distan de la media en tres desviaciones típicas ($\mu \pm 3\sigma$). Para este fin se suelen utilizar datos históricos, es decir, la denominada muestra de calibrado, siendo el contraste de hipótesis (con H_0 : sistema bajo control) sobre una misma población, F . Se recomienda que los límites de control naturales se estimen a partir de una muestra de calibrado constituida por 20 o más observaciones (Montgomery, 2009). El proceso de estimación de los límites de control naturales es como sigue:

1. A partir de la muestra de calibrado, se estiman los límites de control. En los gráficos tipo Shewhart, se construyen a una distancia 3σ de la media o, lo que es lo mismo, a un nivel de confianza del 99.73 % (bajo normalidad).
2. Si existe alguna observación fuera de los límites de control o formando parte de una racha (patrones poco probables en ausencia de causas asignables), se eliminará de los cálculos siempre y cuando se halle la causa asignable que la ha originado (desgaste de una máquina, mala calibración, cambio de turno de operarios, cambio de operario, mal funcionamiento del aparato de medida, etc.).
3. Se vuelven a repetir los pasos 1 y 2 hasta que no haya observaciones fuera de los límites ni rachas. En estas condiciones, se dice que el proceso está bajo control.

Entre las reglas más comunes para la detección de estados fuera de control, están las propuestas por la Western Electric (Electric, 1956). Estas son:

1. Un punto fuera de los límites de control 3σ .

2. Dos puntos de tres consecutivos más allá de los límites de aviso 2σ .
3. Cuatro de cinco puntos consecutivos a una distancia de σ o más de la línea central.
4. Ocho puntos consecutivos a un lado de la línea central.

No se suelen utilizar más de 3 reglas a la vez para evitar incurrir en falsos rechazos de la H_0 (proceso bajo control). Las librerías de R, `qcr` y `qcc` (Flores, 2016; Strucca, 2004) emplean la regla 1 y otra regla que detecta como estado fuera de control la situación en la que más de 6 medidas seguidas están a un lado de la línea central del gráfico, CL. En este punto, es importante mencionar que algunos autores opinan que tener en cuenta las observaciones que forman parte de rachas pueden llevar a la detección de un número inaceptable de falsas alarmas (Woodall, 2000).

Con respecto la propuesta desarrollada en este capítulo, en el caso de monitorización de datos funcionales, se considera solo la regla consistente en que, si un dato funcional cuya profundidad está más allá del límite inferior del gráfico de control para la Fase I, el proceso está fuera de control. Este límite es calculado mediante un procedimiento de remuestreo bootstrap suavizado.

Por otro lado, en la Fase II, se asume que el proceso está bajo control y en cada nueva muestra (muestra de monitorizado) se obtiene el estadístico rango y se representa en el gráfico con el límite de control inferior estimado. El estadístico rango se obtiene a partir de la muestra de calibrado. De la misma forma que en la Fase I, sólo se considera que el proceso tiene anomalías si el valor del estadístico es menor que el límite inferior del gráfico de control. A modo ilustrativo, en la Figura 4.1 muestra el proceso estándar de monitorización de una variable CTQ (Stapenhurst, 2013). El mismo proceso sería válido para el caso de datos funcionales.

4.1.3. Importancia de la Fase I en la construcción de un gráfico de control

La mayor parte de la literatura hace referencia a la Fase II de la estimación de los gráficos de control en detrimento de la Fase I. Por otro lado, actualmente, en el ámbito del Big Data, no se establece una diferenciación entre la Fase I y la Fase II.

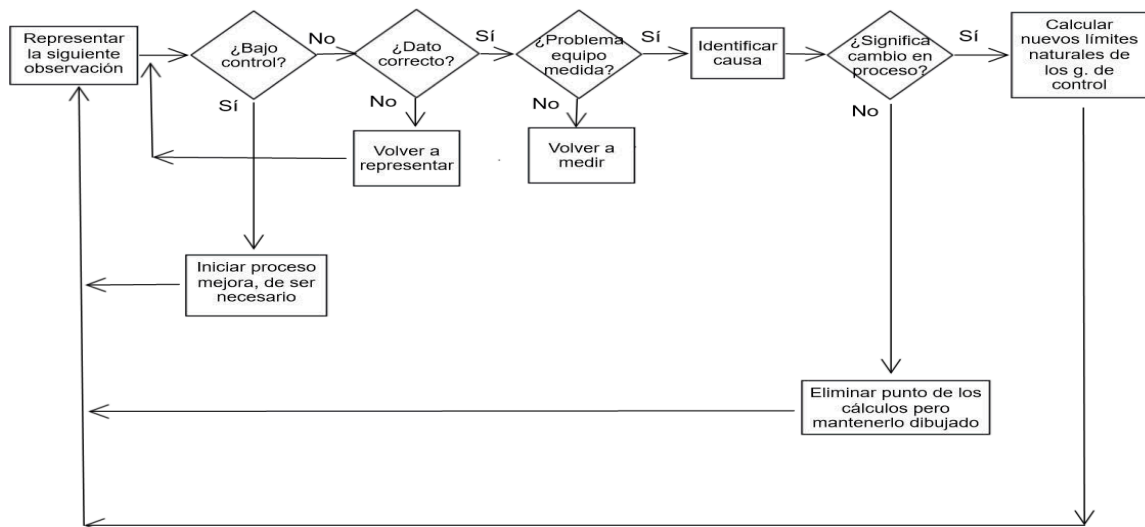


Figura 4.1: Proceso de monitorización de una variable que afecta a la calidad de un producto o servicio mediante la aplicación de gráficos de control.

Y no sólo eso, sino que incluso algunos autores definen la necesidad de llevar a cabo una nueva fase denominada Fase 0, relacionada con el preprocesamiento (limpieza, selección de características, evaluación de la calidad de los datos, estudios de la precisión del sistema de medida, etc.) de los datos (Megahed y Jones-Farmer, 2013). En este trabajo se proporcionan alternativas funcionales para gráficos de control tanto para la Fase I como para la Fase II, clarificando la función de cada uno en el marco actual de datos complejos (en este caso funcionales).

En la Fase I, se emplean técnicas como son los métodos gráficos (multi-vari chart, por ejemplo), metodologías para la detección de outliers y procedimientos para la estimación de los parámetros del gráfico de control, además del diseño de submuestras y, como paso previo, la selección de las características. En definitiva, en la Fase I se estiman los límites de control natural de las variables CTQ, una vez eliminadas las causas asignables (y aplicando las medidas correctivas en el proceso productivo) que se corresponden con las fuentes más importantes de variación. Dentro del nuevo paradigma de datos, según varios autores, se echan en falta un mayor desarrollo de metodologías para la Fase I Megahed y Jones-Farmer (2013). El trabajo de Jones-Farmer *et al.* (2014) proporciona una idea acerca de las recientes aportaciones al

desarrollo de la Fase I de los gráficos de control.

En este capítulo, se adapta el procedimiento de detección de atípicos presentado en [Febrero-Bande *et al.* \(2007\)](#) para construir un gráfico de control para la Fase I, con la finalidad de contar con una muestra de calibrado que provenga de un proceso bajo control y que permita monitorizar el proceso a través de nuevas mediciones mediante un gráfico de control de rangos.

4.2. Metodología

En esta sección, se desarrollan metodologías para el control de calidad mediante gráficos de control de una variable aleatoria funcional, \mathbf{X} , que toma valores en un espacio funcional $\mathcal{E} = L^2(T)$, con $T \subset \mathbb{R}$.

A partir de observaciones de la variable funcional \mathbf{X} , se cuenta con una muestra de calibración y otra de monitorización, que son conjuntos de datos funcionales de tamaño n y m respectivamente, que permiten construir gráficos de control para la Fase I (en el caso de la muestra de calibrado) y Fase II (a partir de la muestra de monitorizado).

Para el diseño de los gráficos de control se considera que si se detecta que el proceso no es estable o no está bajo control, es por causas de variación asignables que corresponden a eventos inusuales y evitables que interrumpen el proceso, es decir, cuando provocan un cambio en los parámetros del modelo subyacente del perfil o dato funcional, las cuales pueden ser eliminadas de los datos conociendo y actuando sobre la causa para así evitar dichas variaciones en el futuro ([Woodall *et al.*, 2004](#)).

Para la propuesta de un método para construir gráficos de control de la calidad, se considera que la probabilidad de que el proceso no sea estable (nivel de significación) es una medida de su rendimiento. Esta es la probabilidad, siempre que se esté dentro de la H_0 , de obtener al menos una medida (valor observado del estadístico) fuera de los límites de control ([Colosimo y Pacella, 2010](#)). En la Fase I, se propone un método y se estima el nivel de significación mientras que en la Fase II, se asume que el proceso está bajo control, es decir, se fija el nivel de significación.

4.2.1. Procedimiento de construcción de un gráfico de control para la FASE I (estabilización)

Como se ha mencionado anteriormente, en la Fase I, se analizan los datos históricos correspondientes a una muestra de calibrado de tamaño n del proceso con la finalidad de evaluar la estabilidad del proceso cuya calidad se evalúa a lo largo del tiempo y estimar los parámetros del gráfico de control (Colosimo y Pacella, 2010).

En la Fase I, al utilizar un gráfico de control se está contrastando la hipótesis de que no existe un cambio de distribución de las observaciones de la variable ordenadas en el tiempo $\{\mathcal{X}_1(t), \mathcal{X}_2(t), \dots, \mathcal{X}_n(t)\}$.

Estos cambios pueden ser puntuales (freaks o bunches) o pueden estar relacionados con un cambio en proceso que se evalúa (observable a través de patrones de cambio repentino o paulatino en la media del proceso). Con respecto a los cambios aislados, es la ocurrencia de al menos una observación de la variable observada que se desvía de la distribución de las otras observaciones (Lanhede, 2015). La hipótesis contrastada en la Fase I es:

$$\begin{aligned} H_0 : \mathcal{X}_i(t) &\stackrel{d}{=} \mathcal{X}_j(t), \forall i, j \in \{1, \dots, n\} \\ H_a : \mathcal{X}_i(t) &\neq \mathcal{X}_j(t), \text{ para algún } i, j \in \{1, \dots, n\} \end{aligned} \quad (4.1)$$

La fase de estabilización de un proceso consiste en aplicar un método iterativo que permita detectar y eliminar aquellas observaciones (en este contexto curvas), que tengan una desviación con respecto a la forma o magnitud de la mayoría de las curvas observadas. Es decir, una curva es un valor atípico si ha sido generada por un proceso estocástico diferente o existe un cambio en tendencia o variabilidad del proceso estocástico con respecto al correspondiente al resto de los datos. La cantidad de valores atípicos se asume como desconocida aunque pequeña.

En este trabajo se proponen gráficos de control para la Fase I a partir de la adaptación del procedimiento propuesto en Febrero-Bande *et al.* (2007) para detección de datos atípicos a través de medidas de profundidad para datos funcionales.

El método de detección de valores atípicos para datos funcionales (Febrero-Bande *et al.*, 2007) considera una curva atípica si su profundidad es menor al cuantil calculado de la distribución de las profundidades estimadas por bootstrap. Es decir, una

curva atípica tendrá una profundidad significativamente baja.

Este procedimiento se puede utilizar con diversos tipos de profundidades funcionales. En la librería `fda.usc` (Febrero-Bande y Oviedo de la Fuente, 2012) se ofrecen las siguientes alternativas: FM, profundidad modal y RP. También, mediante el uso del paquete `fda.usc` se puede realizar el contraste de detección de atípicos propuesto por Febrero-Bande *et al.* (2007), a partir de la estimación del cuantil de la distribución de las profundidades de un conjunto de curvas $\mathcal{X}_1, \dots, \mathcal{X}_n$. En este trabajo, se ha adaptado el procedimiento de detección de atípicos de la librería `fda.usc` con la finalidad de estimar un cuantil de la distribución de las profundidades que juega el papel del límite de control inferior (LCI) para un gráfico de control en la Fase I.

El gráfico de control propuesto para la Fase I se construye y representa a partir de una medida de profundidad (FM, RP ó Modal) y sólo se considera el límite de control inferior (LCI) para detectar si el proceso está fuera de control (la profundidad de una curva es menor al LCI). Además de esta representación, se cuenta con un gráfica que muestra las curvas originales para proporcionar una idea intuitiva acerca del porqué de la identificación de una anomalía (por su forma o magnitud), y así ayudar a identificar causas asignables (por ejemplo, en el caso del consumo de clima en edificios: parada del climatizador, fallo en el contador o sensor, cambio en la regulación de la máquina, condiciones climatológicas adversas, etc.).

En la Fase I, se considera la variable aleatoria funcional \mathcal{X} , de la cual se selecciona una muestra aleatoria $\{\mathcal{X}_1(t), \mathcal{X}_2(t), \dots, \mathcal{X}_n(t)\}$. A partir de esta, con el fin de construir el gráfico de control para la Fase I, se siguen los siguientes pasos:

1. Se calcula la profundidad correspondiente a cada observación respecto al conjunto de datos, $D(\mathcal{X}_i)_{i=1}^n$ y se realiza un gráfico en función de la profundidad y la temporalidad.
2. Se elige el parámetro LCI en función del nivel de significación del gráfico de control, es decir, tal que el porcentaje de falsas alarmas (observaciones bajo control pero erróneamente detectadas como fuera de control) sea pequeño (por ejemplo, $\alpha = 1\%$). Se tiene los siguientes procedimientos para estimar el LCI.

- Procedimiento Bootstrap basado en recorte:

- Reordenar las curvas de acuerdo a sus profundidades de forma decreciente. $\mathcal{X}_{(1)}, \dots, \mathcal{X}_{(N)}$.
 - Se supone que a lo sumo un α % de la muestra se pueden considerar datos atípicos.
 - Se obtienen B muestras mediante un procedimiento bootstrap suavizado a partir del conjunto de datos resultante de descartar el α % de las curvas menos profundas. Sean $\mathcal{X}_i^{*b}, i = 1, \dots, N, b = 1, \dots, B$ estas muestras bootstrap. Para obtener cada réplica bootstrap:
 - Se realiza un muestreo uniforme, i^* de $1, \dots, [N(1 - \alpha)]$.
 - Se genera Z_{i^*} como un proceso gaussiano con media cero y matriz de varianzas y covarianzas $\delta \Sigma_{\mathcal{X}}$ con $\delta \in [0, 1]$. Donde $\Sigma_{\mathcal{X}}$ es la matriz de varianzas y covarianzas de las observaciones $\mathcal{X}_{(1)}, \dots, \mathcal{X}_{([N(1-\alpha)])}$.
 - Finalmente, se obtiene $\mathcal{X}_i^{*b} = \mathcal{X}_{(i^*)} + Z_{i^*}$
 - Para cada $b = 1, \dots, B$, se obtiene C^b , que se identifica con el cuantil empírico del 1 % correspondiente a la distribución de las profundidades, $D(\mathcal{X}_i^{*b})$. El valor final $C = LCI$ es la mediana de los valores $C^b, b = 1, \dots, B$.
- Procedimiento Bootstrap basado en ponderación:
- Se obtienen las profundidades de las curvas $\mathcal{X}_1, \dots, \mathcal{X}_N$.
 - Se obtienen B muestras mediante un bootstrap suavizado, a partir del conjunto de datos original ponderado por su profundidades. Sea $\mathcal{X}_i^{*b}, i = 1, \dots, N, b = 1, \dots, B$ estas muestras bootstrap. Estas réplicas se obtendrían:
 - Se realiza un muestreo ponderado, con i^* de $1, \dots, N$, y con probabilidad proporcional a $D(\mathcal{X}_1), \dots, D(\mathcal{X}_N)$.
 - Se genera Z_{i^*} como un proceso gaussiano con media cero y matriz de varianzas y covarianzas $\delta \Sigma_{\mathcal{X}}$ con $\delta \in [0, 1]$. Donde $\Sigma_{\mathcal{X}}$ es la matriz de varianzas y covarianzas de las observaciones $\mathcal{X}_1, \dots, \mathcal{X}_N$.
 - Finalmente, se obtiene $\mathcal{X}_i^{*b} = \mathcal{X}_{i^*} + Z_{i^*}$
 - Para cada $b = 1, \dots, B$, se obtiene C^b , que es el cuantil empírico correspondiente al 1 % de la distribución de las profundidades, $D(\mathcal{X}_i^{*b})$. El valor final $C = LCI$ es el cuantil β de los valores C^b , con

$$b = 1, \dots, B.$$

3. Si hay alguna curva tal que $D(\mathcal{X}_i) \leq LCI$ para un LCI dado, se considera atípica y el proceso estaría fuera de control.
4. Adicionalmente, se realiza un gráfico que incluye a las curvas originales y la envolvente funcional obtenida a partir del 99% de las réplicas bootstrap más profundas.

Además, una vez que se detectan las curvas atípicas, seguidamente se eliminan y se repite este procedimiento hasta que el proceso esté estable (bajo control), es decir, con total ausencia de datos atípicos.

4.2.2. Procedimiento de construcción de un gráfico de control para la FASE II (monitorización)

La Fase II se ocupa de la monitorización del proceso, con el fin de detectar rápidamente cambios a partir de la muestra de calibrado estabilizada en la Fase I. Para el caso escalar y multivariante, se monitoriza el proceso teniendo como referencia los límites de control estimados en la Fase I (Woodall *et al.*, 2004). En esta fase, se emplea la longitud media de racha (ARL) para evaluar el desempeño de los gráficos de control.

En este contexto, las pruebas de hipótesis que se aplica contrasta si existen desviaciones de la distribución entre los datos obtenidos en la Fase II, también llamados muestra de monitorizado, $\{\mathcal{X}_{n+1}(t), \mathcal{X}_{n+2}(t), \dots, \mathcal{X}_m(t)\}$ y los datos de referencia $\{\mathcal{X}_1(t), \mathcal{X}_2(t), \dots, \mathcal{X}_n(t)\}$ o muestra de calibrado.

En la Fase II, en el caso univariante se asume una distribución F estimada a partir del proceso mediante una muestra de calibrado o datos de referencia. Se asume que F es la distribución de la variable CTQ del proceso cuando éste está bajo control (Fase I). Esta distribución se usa para establecer límites de control que se usarán para monitorizar el proceso en la Fase II. Los límites abarcan un intervalo que debe cubrir nuevas observaciones del proceso con alta probabilidad, dentro del supuesto de que el proceso esté bajo control. En la Fase II se monitoriza una muestra de la

distribución G . Por tanto, en esta etapa los métodos de construcción de gráficos de control se basan en contrastar la hipótesis:

$$\begin{aligned} H_0 : F &= G \\ H_1 : F &\neq G \end{aligned} \tag{4.2}$$

En el contexto FDA no se cuenta con una función de densidad para una variable aleatoria funcional \mathcal{X} para realizar los diferentes contrastes expuestos en la Fase I y Fase II, se utiliza la distribución de la medida de profundidad correspondiente a las curvas que pertenecen a una muestra de datos funcionales.

En este capítulo para la Fase II, se propone emplear los gráficos de control de rangos (Liu, 1995), pero en un contexto FDA. De hecho, se propone el cálculo de profundidades definidas para datos funcionales, a partir de las cuales se calculan los rangos, que son la base a partir de la cual se construyen los gráficos de control r , comúnmente llamados gráficos de rangos.

La adaptación de los gráficos de control de rangos presentado por primera vez en Liu (1995), consiste básicamente en utilizar medidas de profundidad funcional univariante. Se Calcula $\{r_{G_m}(\mathcal{X}_{n+1}), r_{G_m}(\mathcal{X}_{n+2}), \dots, r_{G_m}(\mathcal{X}_m)\}$. El gráfico r consiste en trazar el estadístico de rango en función del tiempo. Cuyo límite de control central $CL = 0.5$ y sirve como un punto de referencia para ver si un patrón o tendencia está surgiendo o se está desarrollando en una secuencia de muestras y el límite inferior es $LCL = \alpha$, donde α es la tasa de falsa alarma.

Los gráficos de control de rangos con medidas de profundidad funcional permiten, además de monitorizar conjuntos de curvas que estén relacionados con la variable de control, monitorizar igualmente las derivadas del conjunto de datos. La información que proporcionan las derivadas de las curvas originales puede ser relevante para detectar anomalías relacionadas con la tasa de cambio de las variables CTQ funcionales.

Para la aplicación práctica de esta propuesta, se utiliza el paquete `qcr` que, a su vez usa funciones del paquete `fda.usc` para el cálculo de profundidades funcionales, además de tener implementados los gráficos no paramétricos propuestos en Liu (1995).

Como se ha mencionado anteriormente, en la Fase II se utilizan las curvas co-

respondientes a la muestra de calibrado de la Fase I, $\{\mathcal{X}_1(t), \mathcal{X}_2(t), \dots, \mathcal{X}_n(t)\}$, con el propósito de detectar cambios, desviaciones en general con respecto al comportamiento del proceso descrito en la Fase I. Para esto, se seleccionan curvas del proceso, preferentemente en tiempo real $\{\mathcal{X}_{n+1}(t), \mathcal{X}_{n+2}(t), \dots, \mathcal{X}_m(t)\}$, y así contrastar la hipótesis de cambio distribución en el proceso.

El procedimiento para realizar gráficos de control para la Fase II sigue el mismo esquema presentado en Liu (1995). Es decir, se utilizan los resultados de Liu, en los que el estadístico rango sigue una distribución asintótica uniforme, este resultado es aplicable al caso funcional debido a la forma en que se construye el estadístico rango de una observación (porcentaje de curvas menos profundas a la observada). Esto permite tener una ventaja computacional en la monitorización de procesos continuos ya que no se ha de estimar el LCI, sino que se fija a un nivel de significancia α .

A continuación se detalla la metodología para el caso univariante funcional, el cual puede ser fácilmente generalizado al caso multivariante funcional.

1. A partir de la muestra de referencia $\{\mathcal{X}_1(t), \mathcal{X}_2(t), \dots, \mathcal{X}_n(t)\}$, se obtienen las profundidades del conjunto de datos, $D(\mathcal{X}_i)_{i=1}^n$, y se obtienen las profundidades de las curvas que componen la muestra de monitorización, $D(\mathcal{X}_j)_{j=n+1}^m$.
2. Se calculan los estadísticos rank, $r_G(\mathcal{X}_{n+1}), \dots, r_G(\mathcal{X}_m)$, de la muestras de monitorización, considerando como muestra de referencia $\{\mathcal{X}_1(t), \mathcal{X}_2(t), \dots, \mathcal{X}_n(t)\}$.

$$r_G(\mathcal{X}) = \frac{\#\{\mathcal{X}_i | D(\mathcal{X}_i) \leq D(\mathcal{X}), i = 1, \dots, n\}}{n}$$

3. Se representan gráficamente los valores del estadístico rango y los límites inferior $LCI = \alpha$ y central $LC = 0.5$ (el valor esperado del estadístico rango).
4. Se procede a monitorizar el proceso. Sí se observa que al menos la profundidad de una curva es tal que $D(\mathcal{X}_j) \leq LCI$, entonces el proceso esta fuera de control.
5. Se realiza un gráfico de control funcional en el que se incluyen las curvas originales, tanto de la muestra de referencia como de monitorización, además de la envolvente funcional obtenida a partir del 99% de las curvas más profundas de la muestra de calibrado o referencia.

4.3. Estudio de simulación

El rendimiento o desempeño de un gráfico de control generalmente se mide mediante su potencia $(1 - \beta)$, definida como la probabilidad de detectar una señal de fuera de control cuando el estado del proceso subyacente efectivamente no está bajo control. Cuando el proceso está bajo control, la probabilidad de que se de una señal de fuera de control se denomina error de tipo I (α) o tasa de falsa alarma.

En general, cuando el proceso esté bajo control, la probabilidad de detectar una observación fuera de control debe ser lo suficientemente pequeña como para evitar un número no asumible de falsas alarmas. Sin embargo, cuando el proceso está realmente fuera de control, la potencia debe ser lo suficientemente alta como para detectar rápidamente el cambio de estado del proceso.

Otro índice común para medir el rendimiento de un gráfico de control es el ARL, que se define como el número promedio de observaciones trazadas antes de que se de una señal de fuera de control. El ARL es igual a $\frac{1}{p}$ (donde p es la probabilidad de tener una señal de fuera de control).

Para evaluar el desempeño de un gráfico de control en la Fase II, a menudo se utilizan el ARL_0 y el ARL_1 , que son el promedio del número de observaciones hasta detectar la primera fuera de control, en los casos en los que el proceso está realmente bajo control ($ARL_0 = \frac{1}{\alpha}$) o realmente no lo está ($ARL_1 = \frac{1}{1-\beta}$). El ARL_1 debe ser lo más bajo posible para aumentar la probabilidad de identificar rápidamente eventos $(1 - \beta)$, potencia de una prueba) que lleven al proceso estar fuera de control.

Como no se conoce la distribución F para calcular la potencia de los gráficos de control se diseña una simulación Monte Carlo. Los diferentes escenarios simulados permiten estimar y comparar la potencia de los gráficos de control para diferentes medidas de profundidad, y para el caso de datos funcionales independientes y dependientes. En esta sección, para evaluar el rendimiento de los gráficos de control propuestos para la Fase I y Fase II, se parte del esquema de simulación diseñado en Febrero-Bande *et al.* (2007).

En Febrero-Bande *et al.* (2007), se generan realizaciones de un proceso estocástico gaussiano de la forma:

$$\mathcal{X}(t) = \mu(t) + \sigma(t) \cdot \epsilon(t) \tag{4.3}$$

donde $\sigma^2(t) = 0.5$ y

$$\mu(t) = \mathbf{E}(\mathcal{X}(t)) = 30t(1-t)^{3/2} \quad (4.4)$$

además $\epsilon(t)$ es un proceso gaussiano $\epsilon(t) \sim GP(\mathbf{0}, \Sigma)$ con media $\mathbf{0}$ y matriz de varianzas y covarianzas igual a

$$\mathbf{E}[\epsilon(t_i) \times \epsilon(t_j)] = e^{-\frac{|t_i - t_j|}{0.3}}$$

adicionalmente, en Febrero-Bande *et al.* (2007) se considera un modelo alternativo para generar las curvas atípicas con $\mu(t) = 30t^{3/2}(1-t)$. En la Figura 4.2, en el panel (a), se presentan las dos medias funcionales, la curva de color negro es la media del proceso sin curvas atípicas mientras que la de color rojo es la media del proceso que genera las curvas atípicas.

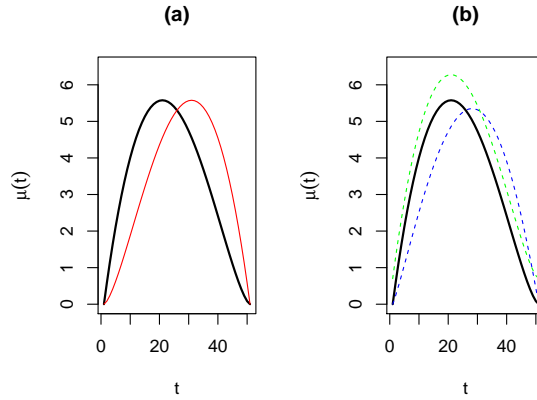


Figura 4.2: (a) Medias funcionales y (b) cambios en la media del proceso en forma y magnitud.

Los gráficos de control propuestos en esta memoria han sido diseñados para monitorizar la media funcional con la finalidad de detectar dos eventos, cambio en la media del proceso en magnitud y cambio en forma, que producen que el proceso no este bajo control. Para el diseño de los gráficos de control para la Fase II, se asume que el proceso esta bajo control, es decir no hay presencia de atípicos. Para generar los escenarios de simulación para cada uno de estos eventos, se han considerado las siguientes medias funcionales.

- Media del modelo con cambio en la magnitud:

$$\mu(t) = 30t(1 - t)^{3/2} + \delta \quad (4.5)$$

donde δ es el cambio que va de 0.4 hasta 2 en pasos de 0.4.

- Media del modelo con cambio en la forma:

$$\mu(t) = (1 - \eta) \cdot 30t(1 - t)^{3/2} + \eta \cdot 30t^{3/2}(1 - t) \quad (4.6)$$

donde η es el cambio que va de 0.2 hasta 1 en pasos de 0.2.

En el panel (b) de la Figura 4.2, la curva de color verde representa la media funcional de un proceso cuando hay cambio en magnitud ($\delta = 0.7$), mientras que la curva de color azul representa la media de un proceso cuando hay cambio en la forma ($\eta = 0.3$).

En Febrero-Bande *et al.* (2007), se consideran los datos funcionales $\mathcal{X}_1, \dots, \mathcal{X}_n$, obtenidos como realizaciones de un proceso estocástico $X(\cdot)$, asumiendo trayectorias continuas en el período de observación $[a, b] = [0, 1]$ e independencia entre las curvas. Por otro lado, también se han considerado escenarios en los que las curvas simuladas están definidas por un grado variable de dependencia. Esto se debe a que gran parte de las aplicaciones prácticas de este tipo de gráficos están relacionadas con datos monitorizados continuamente con respecto al tiempo, formando series temporales funcionales, como es el caso de las curvas de consumo energético diario en superficies comerciales. De este modo, se generan curvas dependientes a partir del modelo $\tilde{Y}_i(t) = \mu(t) + \sigma(t) \cdot \tilde{\epsilon}(t)$, con $\tilde{\epsilon}(t) = \rho \cdot \tilde{\epsilon}_{i-1}(t) + (1 - \rho) \cdot \epsilon_i(t)$. Donde ρ es la medida de correlación entre curvas, y $\sigma(t) = 0.5$, y tanto $\epsilon(t)$ y $\tilde{\epsilon}(t)$ son procesos gaussianos.

Para poder comparar los resultados de las simulaciones entre el caso independiente y dependiente se reescala la varianza de ϵ (queremos que la varianza del error $\tilde{\epsilon}$ sea uno). Es decir, considerando $\sigma_\epsilon^2 = \frac{(1-\rho^2)}{(1-\rho)^2} = \frac{(1+\rho)}{(1-\rho)}$, se tiene $\sigma_{\tilde{\epsilon}}^2 = 1$.

En la Figura 4.3 se presentan los diferentes escenarios considerando los cambios en la media funcional del proceso en forma y magnitud, en los casos de independencia y dependencia entre curvas. Las curvas de color gris representan las realizaciones del proceso cuando éste está bajo control (cuya media es la expresión 4.4). Mientras que la curva de color rojo en cada gráfico representa los escenarios en los que se considera

la presencia de eventos que desestabilizan el proceso, es decir, que el proceso no está bajo control. En la Figura 4.3, los paneles (a) y (b) representan el caso de independencia entre curvas y presencia de eventos definidos por cambios de la media funcional en magnitud y en forma, respectivamente. Por otro lado, en los paneles (c) y (d) se muestran dos casos con dependencia entre curvas, incluyendo también cambios en la magnitud de la media (panel (c)) y en su forma (panel (d)).

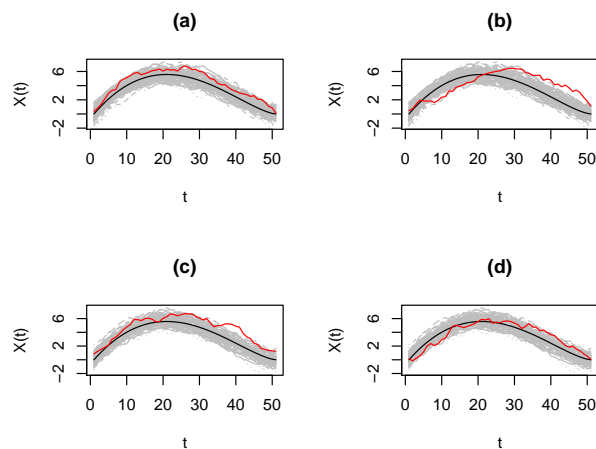


Figura 4.3: Escenarios en los que se considera independencia entre curvas y para cambios en la media funcional con respecto a su magnitud (a) y forma (b). En el caso de dependencia, los paneles (c) y (d) muestran escenarios de simulación en los que se observan cambios de magnitud y forma, respectivamente en la media funcional.

En la siguiente sección se realiza un estudio de simulación con la finalidad de determinar en que condiciones se puede asegurar que el procedimiento bootstrap suavizado funciona cuando se tiene independencia y dependencia entre curvas.

4.3.1. Medición y comparación del rendimiento del gráfico de control propuesto para la Fase I

Se compara el desempeño del gráfico de control a partir de la generación de muestras de calibrado de tamaño 50 y 100 (curvas). Para cada muestra se calculan las diferentes medidas de profundidad funcional descritas en el Capítulo 2 y se aplican los procedimientos robustos (ponderado y recortado) para la estimación del error

de tipo I cuando el proceso está bajo control y la potencia de la prueba cuando está fuera de control. Para la estimación del error tipo I, se replican 1000 veces cada escenario ($n = 50, 100$, asumiendo independencia y dependencia) y para el caso de la estimación de la potencia de la prueba, en cada escenario (asumiendo independencia y dependencia) se genera una curva dentro de la hipótesis alternativa, este procedimiento se repite 1000.

Siguiendo el esquema descrito en Febrero-Bande *et al.* (2007), se consideran curvas observadas en puntos equidistantes, el número de puntos que definen cada curva es 51 en el intervalo $[0, 1]$. A partir de 1000 remuestras ($B = 1000$) y con un procedimiento de recorte del 2.5% (eliminación de las curvas menos profundas), se aplica un procedimiento bootstrap suavizado definido por un factor de suavización $\gamma = 0.05$ con el fin de estimar el cuantil $C = 0.01$ que representa el Límite de Control Inferior (LCI).

Primero se realiza un estudio de simulación para estimar y comparar el error de tipo I (se fija $\alpha = 0.01$) del gráfico de control propuesto considerando los escenarios de independencia y dependencia entre curvas. Posteriormente se realiza un estudio similar para estimar y evaluar la potencia del gráfico de control para detectar señales de fuera de control en diferentes situaciones (independencia, dependencia, tamaño de muestra, cambio en forma o magnitud).

En la Tabla 4.1, se muestran los resultados de la estimación de la tasa de falsas alarmas (error de tipo I) en el escenario de independencia. El porcentaje medio de falsas señales de fuera de control (error de tipo I) detectado por el procedimiento arriba mostrado son muy próximas al 1% nominal para los dos tamaños de muestra considerados. Además se puede observar que cuando n aumenta los porcentajes de error de tipo I están más cerca del nivel nominal. En general, para una muestra de tamaño $n = 100$, los resultados de aplicar el método ponderado se aproximan más a α , sobre todo al utilizar la medida de profundidad Modal. Los resultados obtenidos en las simulaciones son similares a los expuestos en Febrero-Bande *et al.* (2007).

En cualquier proceso, el error de tipo I causa que los costos de producción aumenten, por tal motivo en la gestión de la calidad es imprescindible no sobrestimar esta tasa de error.

En la Tabla 4.2 se presentan los resultados de la simulación con el fin de evaluar

n	Método	FM	RP	Modal
50	Ponderado	1.94	1.89	1.49
	Recortado	1.34	1.95	1.36
100	Ponderado	1.55	1.75	1.25
	Recortado	1.67	2.33	1.76

Tabla 4.1: Resultados de la estimación de la tasa de falsas alarmas (Error I) para el caso de independencia entre curvas

la capacidad del gráfico de control para detectar un cambio en forma ó magnitud de la media funcional del proceso a través de la estimación de su potencia $(1 - \beta)$. El porcentaje de señales fuera de control (valores atípicos) correctamente detectados cuando se contamina la población definida por la expresión 4.4 con los modelos M_1 (expresión 4.5) y M_2 (expresión 4.6), se denota por p_c , mientras que el porcentaje de falsas alarmas (falsos estados fuera de control), p_f . Estos parámetros han sido estimados para cada uno de los escenarios de independencia utilizando la media de los valores empíricos correspondientes, \hat{p}_c y \hat{p}_f .

A partir de los resultados en la Tabla 4.2, para el modelo M_1 donde se consideran cambios en la magnitud el gráfico de control propuesto, se puede concluir que se tiene en general un mejor rendimiento que para el modelo M_2 . Es decir, la tasa de detección verdaderas observaciones fuera de control $(1 - \beta)$ es más alta. Cuando se desea detectar cambios en la forma de la media del proceso, M_2 , la profundidad Modal proporciona los más altos porcentajes de señales de fuera de control detectadas correctamente. Por otro lado, para el caso del modelo M_1 , la profundidad RP presenta porcentajes de detección de verdaderos estados fuera de control más bajos que al usar las profundidades FM y Modal. Con respecto al método robusto, el desempeño es similar en todos los escenarios con excepción del uso de la profundidad RP, para el que se tiene un menor rendimiento del gráfico de control al detectar observaciones que realmente denotan estados fuera de control.

En general, el porcentaje de detección de falsas señales de fuera de control para el escenario de independencia es cercano al 1%. Sin embargo al usar el método recortado, se sobrestima la tasa de detección de falsas señales de fuera de control, pero este porcentaje disminuye cuando aumenta el tamaño de la muestra.

Escenario	Muestra	Profundidad	Método	δ		0.4		0.8		1.2		1.6		2	
				\hat{p}_f	\hat{p}_c	\hat{p}_f	\hat{p}_c	\hat{p}_f	\hat{p}_c	\hat{p}_f	\hat{p}_c	\hat{p}_f	\hat{p}_c	\hat{p}_f	\hat{p}_c
M_1	50	FM	Ponderado	1.80	8.30	1.63	27.80	1.47	56.50	1.48	85.30	1.60	95.70		
			Recortado	1.25	7.70	1.19	26.80	1.19	56.00	1.22	83.00	1.20	95.50		
		RP	Ponderado	1.88	6.80	1.68	20.10	1.49	45.10	1.42	71.90	1.42	87.70		
			Recortado	1.88	7.00	1.89	23.20	2.08	49.30	2.14	75.70	2.20	91.90		
		Modal	Ponderado	1.41	5.70	1.27	19.90	1.09	45.10	1.08	76.70	1.17	94.30		
			Recortado	1.33	6.80	1.32	22.20	1.33	48.60	1.38	77.50	1.39	95.20		
	100	FM	Ponderado	1.51	5.70	1.40	23.30	1.34	53.50	1.34	82.00	1.38	96.20		
			Recortado	1.63	7.40	1.51	25.90	1.32	56.10	1.13	83.60	1.00	96.70		
		RP	Ponderado	1.68	5.00	1.57	18.60	1.50	45.40	1.45	71.40	1.52	90.10		
			Recortado	2.28	7.40	2.22	23.70	2.23	54.20	2.13	78.70	2.08	93.50		
		Modal	Ponderado	1.22	4.20	1.15	17.90	1.08	46.80	1.06	78.90	1.12	94.70		
			Recortado	1.70	5.80	1.62	22.90	1.46	54.20	1.28	83.50	1.13	95.90		
M_2	50	FM	Ponderado	1.91	1.50	1.85	3.40	1.70	11.00	1.46	26.30	1.25	55.00		
			Recortado	1.28	1.90	1.23	4.10	1.19	11.60	1.13	26.70	1.18	50.60		
		RP	Ponderado	1.94	2.40	1.90	3.30	1.75	8.00	1.64	16.90	1.42	32.20		
			Recortado	1.89	2.30	1.83	4.60	1.73	10.10	1.84	22.50	1.92	41.40		
		Modal	Ponderado	1.43	2.80	1.35	10.70	1.19	31.00	1.06	65.30	1.12	91.70		
			Recortado	1.33	4.00	1.31	12.30	1.31	33.70	1.42	68.00	1.48	92.10		
	100	FM	Ponderado	1.53	2.00	1.51	2.70	1.45	12.20	1.35	27.90	1.25	53.00		
			Recortado	1.66	2.50	1.63	4.30	1.57	14.40	1.51	33.40	1.43	58.00		
		RP	Ponderado	1.69	2.70	1.66	4.10	1.63	8.20	1.55	17.30	1.48	34.70		
			Recortado	2.25	2.90	2.19	5.60	2.15	13.60	2.15	25.40	2.21	45.50		
		Modal	Ponderado	1.22	3.30	1.18	11.90	1.12	34.10	1.07	66.80	1.09	89.60		
			Recortado	1.72	4.60	1.65	15.70	1.57	41.70	1.41	72.70	1.22	92.90		

Tabla 4.2: Porcentajes de \hat{p}_c y \hat{p}_f para los casos M_1 (expresión 4.5) y M_2 (expresión 4.6) bajo el supuesto de independencia entre curvas.

Los resultados de la tasa de falsa alarma (error de tipo I) para los escenarios de dependencia entre curvas se muestran en la Tabla 4.3. Es importante destacar que, para los diferentes valores de ρ , se tienen resultados similares al caso de independencia, es decir, los porcentajes medios de falsas señales de fuera de control son cercanos al 1 % nominal en los dos tamaños de muestra considerados. Además, de igual manera, se puede observar que cuando n aumenta los porcentajes de error de tipo I están más cerca del nivel nominal. Sin embargo se observan algunas diferencias en el caso en el que se utiliza la medida de profundidad RP para construir el gráfico de control; en este caso, se tiene una sobrestimación del porcentaje de falsas señales de fuera de control. Adicionalmente, cuando la dependencia es alta ($\rho = 0.9$) la profundidad modal subestima el porcentaje de falsas alarmas.

En las Tablas 4.4, 4.5, 4.6, 4.7 y 4.8 se presentan los resultados de la estimación empírica de p_c y p_f para diferentes valores de ρ que van desde 0.1 hasta 0.7. La potencia (\hat{p}_c) del gráfico de control propuesto para el modelo M_1 (expresión 4.5) tiene, en general, un mejor desempeño cuando se aplica el método ponderado y se aumenta el tamaño de la muestra. Se observa también que, por lo general, el desempeño no difiere del tipo de medida de profundidad que se utilice. Eso sí, el desempeño de los gráficos de control para detectar cambios reales en el proceso, relacionados con diferencias de forma en la media, es mejor si se utiliza la profundidad Modal.

Con respecto a la tasa de falsas señales de fuera de control \hat{p}_f , para el caso del modelo M_1 , cuando se utiliza el método de recorte se obtiene una menor tasa, sin embargo, presenta la desventaja de una menor potencia de detección de correctas señales de fuera de control, como se mencionó anteriormente. En el caso del modelo M_2 , se tienen resultados similares a los escenarios con independencia entre curvas, es decir, que el \hat{p}_f es menor cuando se usa el método de Recortado.

En Raña (2016), se propusieron nuevos métodos para la detección de atípicos para el caso en el que existe una dependencia entre curvas. De los estudios de simulación realizados en este estudio, a distintos grados de dependencia, se evidenciaba que el método de detección de atípicos propuesto en Febrero-Bande *et al.* (2007) era relativamente robusto frente a la presencia de dependencia entre curvas. El estudio de simulación realizado en este capítulo sustenta los resultados obtenidos en el trabajo de Raña (2016) y, como conclusión, justifica el uso de este método dentro de la nueva

ρ	n	Método	FM	RP	Modal
0.1	50	Ponderado	1.92	1.87	1.44
		Recortado	1.34	1.95	1.37
	100	Ponderado	1.56	1.79	1.25
		Recortado	1.68	2.32	1.75
0.3	50	Ponderado	1.85	1.88	1.41
		Recortado	1.31	1.90	1.33
	100	Ponderado	1.54	1.71	1.20
		Recortado	1.65	2.23	1.71
0.5	50	Ponderado	1.80	1.81	1.28
		Recortado	1.26	1.82	1.29
	100	Ponderado	1.46	1.68	1.15
		Recortado	1.61	2.21	1.65
0.7	50	Ponderado	1.66	1.73	0.93
		Recortado	1.25	1.74	1.06
	100	Ponderado	1.42	1.61	0.99
		Recortado	1.56	2.15	1.57
0.9	50	Ponderado	1.39	1.40	0.31
		Recortado	1.00	1.45	0.61
	100	Ponderado	1.14	1.37	0.36
		Recortado	1.28	1.88	0.98

Tabla 4.3: Resultados para el caso dependencia de la estimación de la tasa de falsas alarmas (error de tipo I) para valores de ρ que van de 0.1 hasta 0.9 en pasos de 0.2.

Escenario	Muestra	Profundidad	Método	δ											
				0.4		0.8		1.2		1.6		2			
				$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$		
M1	50	FM	Ponderado	1.64	7.60	1.30	26.60	1.02	59.50	0.99	83.90	1.13	97.20		
			Recortado	1.11	8.00	0.80	23.10	0.47	42.60	0.17	54.20	0.04	53.40		
		RP	Ponderado	1.85	6.30	1.53	20.50	1.10	44.90	0.80	67.90	0.70	85.80		
			Recortado	1.81	8.30	1.55	23.10	1.13	41.80	0.62	56.90	0.33	65.00		
		Modal	Ponderado	1.31	5.30	1.04	20.00	0.75	47.40	0.62	78.80	0.83	95.20		
			Recortado	1.26	6.20	0.94	18.80	0.65	42.20	0.32	62.80	0.10	80.90		
	100	FM	Ponderado	1.46	5.50	1.29	22.40	1.10	57.10	1.07	86.00	1.19	97.50		
			Recortado	1.58	6.80	1.32	24.00	1.07	57.40	0.94	85.50	0.90	97.10		
		RP	Ponderado	1.79	4.80	1.60	20.80	1.36	47.10	1.19	75.00	1.08	92.80		
			Recortado	2.25	6.70	2.06	22.50	1.90	53.60	1.82	81.60	1.87	95.60		
		Modal	Ponderado	1.19	4.00	1.04	18.50	0.88	48.90	0.87	81.80	1.00	96.60		
			Recortado	1.69	5.30	1.46	22.30	1.20	52.70	1.09	83.90	1.07	96.90		
M2	50	FM	Ponderado	η		0.20		0.40		0.60		0.80		1.00	
			Recortado	1.83	0.95	1.69	1.90	1.43	6.25	1.06	12.35	0.77	21.50		
		RP	Ponderado	1.23	1.05	1.15	2.05	1.05	6.10	0.83	12.35	0.61	19.25		
			Recortado	1.92	0.95	1.81	2.30	1.63	4.90	1.33	8.70	1.12	14.20		
		Modal	Ponderado	1.87	1.40	1.77	3.15	1.71	6.05	1.58	11.25	1.33	16.40		
			Recortado	1.37	2.80	1.16	13.20	0.90	34.30	0.69	67.00	0.73	90.70		
		100	FM	Ponderado	1.27	3.40	1.15	13.70	0.90	33.90	0.49	56.60	0.16	73.20	
				Recortado	1.56	0.85	1.46	5.50	1.51	1.30	1.17	13.50	0.98	24.95	
			RP	Ponderado	1.66	1.05	1.58	6.80	1.61	2.00	1.31	16.45	1.15	27.90	
				Recortado	1.84	0.60	1.79	4.80	1.80	1.70	1.52	9.00	1.38	16.60	
			Modal	Ponderado	2.35	1.00	2.25	6.70	2.27	2.50	2.02	12.70	1.96	23.35	
				Recortado	1.21	3.10	1.19	4.00	1.12	9.80	0.88	69.20	0.94	91.80	
				Recortado	1.71	3.90	1.69	5.30	1.56	14.50	1.20	74.10	1.14	93.30	

Tabla 4.4: Resultados de los valores empíricos de $\hat{\rho}_f$ y $\hat{\rho}_c$ con $\rho = 0.1$ para el caso dependencia entre curvas.

Escenario	Muestra	Profundidad	Método	δ		0.4		0.8		1.2		1.6		2	
				$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$		
M1	50	FM	Ponderado	1.58	8.90	1.22	33.60	0.98	69.80	1.07	91.70	1.19	99.30		
			Recorte	1.12	9.20	0.75	26.60	0.37	47.40	0.09	54.30	0.02	49.20		
		RP	Ponderado	1.75	7.10	1.38	26.30	0.97	53.80	0.74	77.60	0.67	90.70		
			Recorte	1.72	7.60	1.42	25.90	0.89	45.90	0.49	60.90	0.20	68.70		
		Modal	Ponderado	1.21	5.80	0.90	24.90	0.67	59.80	0.70	88.30	0.86	98.50		
			Recorte	1.21	6.40	0.89	24.00	0.50	49.40	0.19	73.50	0.04	89.80		
	100	FM	Ponderado	1.47	7.30	1.23	29.60	1.07	67.80	1.13	92.30	1.24	99.20		
			Recorte	1.53	7.60	1.29	31.20	1.04	68.70	0.96	91.30	0.92	99.10		
		RP	Ponderado	1.69	5.50	14.67	24.40	1.21	58.00	1.06	83.90	1.02	95.90		
			Recorte	2.19	8.80	1.99	27.40	1.88	64.80	1.92	89.30	1.85	97.30		
		Modal	Ponderado	1.17	4.60	0.98	23.00	0.86	62.40	0.89	91.00	1.01	98.50		
			Recorte	1.62	6.90	1.38	27.90	1.15	67.10	1.07	91.70	1.06	98.90		
M2	50	FM	Ponderado	1.75	1.30	1.61	1.95	1.31	7.70	0.94	16.85	0.64	26.05		
			Recorte	1.25	1.40	1.15	2.35	0.97	7.70	0.73	15.95	0.51	20.50		
		RP	Ponderado	1.86	1.20	1.75	3.00	1.43	5.50	1.24	10.85	0.96	16.55		
			Recorte	1.79	1.65	1.68	3.40	1.58	7.15	1.46	13.10	1.16	19.40		
		Modal	Ponderado	1.31	3.70	1.08	16.00	0.78	44.90	0.62	79.30	0.81	96.90		
			Recorte	1.22	4.70	1.11	17.50	0.74	42.20	0.29	65.00	0.07	82.50		
	100	FM	Ponderado	1.53	0.90	1.46	2.05	1.30	7.45	1.07	19.50	0.92	32.35		
			Recorte	1.65	1.20	1.57	2.70	1.44	9.40	1.22	22.75	1.12	35.35		
		RP	Ponderado	1.76	0.60	1.73	2.20	1.57	5.10	1.41	12.10	1.27	21.80		
			Recorte	2.23	1.10	2.18	3.30	2.09	8.25	1.91	17.45	1.97	29.35		
		Modal	Ponderado	1.17	3.55	1.07	14.70	0.91	47.20	0.86	83.10	0.98	97.50		
			Recorte	1.64	4.70	1.48	20.50	1.25	52.60	1.13	85.10	1.12	98.20		

Tabla 4.5: Resultados de los valores empíricos de $\hat{\rho}_f$ y $\hat{\rho}_c$ con $\rho = 0.3$ para el caso dependencia entre curvas.

Escenario	Muestra	Profundidad	Método	δ		0.4		0.8		1.2		1.6		2	
				$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$		
M_1	50	FM	Ponderado	1.46	11.50	1.04	43.90	0.84	80.80	1.06	97.40	1.14	99.90		
			Recortado	1.02	10.80	0.63	33.40	0.23	50.70	0.28	53.80	0.00	45.60		
		RP	Ponderado	1.68	9.10	1.21	31.10	0.80	63.00	0.65	85.60	0.60	94.60		
			Recortado	1.73	9.60	1.26	30.70	0.66	52.60	0.28	66.40	0.11	68.80		
		Modal	Ponderado	1.10	7.50	0.76	34.40	0.55	73.00	0.69	95.20	0.82	99.90		
			Recortado	1.05	8.10	0.72	31.40	0.32	62.60	0.06	86.50	0.01	96.40		
	100	FM	Ponderado	1.20	10.00	1.08	39.60	1.01	79.40	1.10	97.30	1.18	99.80		
			Recortado	1.30	10.00	1.18	41.60	0.96	80.20	0.89	97.00	0.84	99.80		
		RP	Ponderado	1.40	10.00	1.39	32.30	1.09	69.20	0.97	90.60	0.97	97.90		
			Recortado	1.80	10.00	1.93	38.20	1.84	77.10	1.80	94.70	1.73	98.80		
		Modal	Ponderado	0.90	10.00	0.86	32.10	0.79	75.80	0.91	96.00	0.98	99.90		
			Recortado	1.30	10.00	1.30	38.30	1.10	79.40	1.06	97.20	1.02	99.90		
M_2	50	FM	Ponderado	1.63	1.40	1.43	3.55	1.02	10.80	0.68	21.60	0.50	33.10		
			Recortado	1.16	1.50	1.02	3.60	0.80	11.20	0.54	18.00	0.32	21.70		
		RP	Ponderado	1.89	1.75	1.65	3.30	1.32	7.30	0.98	14.30	0.75	22.55		
			Recortado	1.79	2.30	1.62	3.80	1.54	9.05	1.26	16.40	0.95	22.90		
		Modal	Ponderado	1.13	5.40	0.86	22.60	0.56	60.50	0.59	91.50	0.85	99.80		
			Recortado	1.10	6.30	0.83	23.40	0.46	54.00	0.11	77.20	0.02	93.80		
	100	FM	Ponderado	0.01	0.01	0.01	0.03	1.13	11.90	0.90	28.25	0.77	38.90		
			Recortado	0.02	0.01	0.01	0.04	1.28	15.25	1.11	30.90	1.04	39.90		
		RP	Ponderado	0.02	0.01	0.02	0.03	1.42	7.65	1.26	16.65	1.08	27.55		
			Recortado	0.03	0.02	0.02	0.05	2.02	12.60	1.91	23.25	1.96	35.00		
		Modal	Ponderado	0.01	0.05	0.01	0.23	0.82	63.00	0.87	92.50	0.99	99.80		
			Recortado	0.02	0.07	0.01	0.28	1.21	68.40	1.12	93.40	1.12	99.60		

Tabla 4.6: Resultados de los valores empíricos de $\hat{\rho}_f$ y $\hat{\rho}_c$ con $\rho = 0.5$ para el caso dependencia entre curvas.

Escenario	Muestra	Profundidad	Método	δ		0.4		0.8		1.2		1.6		2	
				$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$		
M_1	50	FM	Ponderado	1.32	16.00	0.87	57.70	0.85	91.40	1.01	99.30	1.01	100.00		
			Recortado	0.87	13.40	0.42	40.30	0.07	53.60	0.00	47.50	0.00	35.80		
		RP	Ponderado	1.41	13.00	0.97	43.40	0.61	76.70	0.56	92.00	0.53	98.00		
			Recortado	1.51	14.00	0.97	40.70	0.39	61.40	0.14	68.20	0.04	71.30		
		Modal	Ponderado	0.71	14.00	0.44	52.20	0.41	88.90	0.51	99.30	0.53	100.00		
			Recortado	0.82	14.30	0.43	47.40	0.09	80.20	0.01	96.00	0.00	99.90		
	100	FM	Ponderado	1.29	12.00	0.99	54.90	1.05	90.40	1.14	99.50	1.13	100.00		
			Recortado	1.36	13.80	1.08	56.00	0.89	90.20	0.84	99.20	0.76	100.00		
		RP	Ponderado	1.50	9.90	1.20	48.30	0.97	80.80	0.93	96.40	0.96	99.50		
			Recortado	2.00	13.90	1.80	53.10	1.80	87.30	1.72	98.30	1.60	99.90		
		Modal	Ponderado	0.90	9.80	0.68	51.00	0.69	90.00	0.81	99.40	0.83	99.90		
			Recortado	1.43	13.40	1.14	58.10	1.01	92.40	0.97	99.60	0.92	99.90		
M_2	50	FM	Ponderado	1.52	1.80	1.21	6.05	0.76	17.90	0.49	30.70	0.41	39.95		
			Recorte	1.07	2.10	0.85	6.15	0.56	15.30	0.31	22.30	0.18	24.05		
		RP	Ponderado	1.62	2.10	1.36	5.25	0.98	11.40	0.68	19.95	0.56	28.30		
			Recorte	1.59	2.40	1.41	5.80	1.19	13.85	0.92	21.05	0.70	28.15		
		Modal	Ponderado	0.75	8.60	0.47	39.90	0.37	82.80	0.50	99.10	0.61	100.00		
			Recorte	0.87	10.30	0.51	38.30	0.14	73.10	0.01	93.80	0.00	99.30		
	100	FM	Ponderado	2.02	0.00	1.24	6.35	0.99	21.20	0.81	36.45	0.75	45.10		
			Recortado	3.03	0.00	1.36	7.80	1.16	24.30	1.05	37.95	1.03	44.55		
		RP	Ponderado	3.69	0.00	1.50	5.25	1.25	13.00	1.08	24.50	0.99	34.85		
			Recortado	6.28	0.00	1.98	7.20	1.84	18.50	1.89	31.80	1.89	41.60		
		Modal	Ponderado	1.63	0.00	0.72	37.90	0.69	82.70	0.80	98.60	0.90	100.00		
			Recortado	3.48	0.00	1.24	46.70	1.10	87.10	1.05	99.00	1.04	100.00		

Tabla 4.7: Resultados de los valores empíricos de $\hat{\rho}_f$ y $\hat{\rho}_c$ con $\rho = 0.7$ para el caso dependencia entre curvas.

Escenario	Muestra	Profundidad	Método	δ		0.4		0.8		1.2		1.6		2	
				$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$	$\hat{\rho}_f$	$\hat{\rho}_c$		
M_1	50	FM	Ponderado	0.86	33.60	0.64	84.40	0.76	98.70	0.65	99.80	0.52	100.00		
			Recortado	0.45	24.60	0.09	52.30	0.00	42.10	0.00	25.90	0.00	15.40		
		RP	Ponderado	0.88	29.20	0.56	74.40	0.48	93.80	0.49	98.80	0.45	99.80		
			Recortado	0.93	29.70	0.35	60.00	0.07	66.90	0.01	67.30	0.00	65.00		
		Modal	Ponderado	0.13	45.10	0.14	90.10	0.17	99.40	0.15	99.90	0.14	99.90		
			Recortado	0.20	45.80	0.02	85.30	0.00	94.30	0.00	90.50	0.00	86.30		
	100	FM	Ponderado	0.91	28.50	0.77	82.50	0.87	99.20	0.82	100.00	0.73	100.00		
			Recortado	1.04	31.60	0.77	83.60	0.66	98.90	0.56	100.00	0.46	100.00		
		RP	Ponderado	1.09	27.30	0.77	75.60	0.72	95.40	0.78	99.40	0.77	100.00		
			Recortado	1.64	32.50	1.57	82.30	1.51	97.90	1.36	99.90	1.26	100.00		
		Modal	Ponderado	0.25	33.90	0.24	86.80	0.30	99.30	0.28	100.00	0.25	100.00		
			Recortado	0.79	43.10	0.70	91.90	0.63	99.50	0.57	100.00	0.49	100.00		
M_2	50	FM	Ponderado	1.06	6.20	0.63	20.15	0.39	34.45	0.41	42.35	0.52	46.10		
			Recortado	0.67	5.95	0.39	16.30	0.19	22.35	0.09	21.05	0.06	20.05		
		RP	Ponderado	1.12	6.35	0.74	15.15	0.57	26.25	0.56	35.80	0.58	41.30		
			Recortado	1.24	6.95	0.95	16.75	0.59	24.95	0.36	30.25	0.24	33.35		
		Modal	Ponderado	0.15	39.50	0.13	89.30	0.17	99.50	0.19	99.90	0.19	99.90		
			Recortado	0.21	39.90	0.04	83.80	0.00	94.90	0.00	94.70	0.00	90.90		
	100	FM	Ponderado	1.04	4.45	0.77	19.40	0.62	38.20	0.60	46.15	0.66	48.65		
			Recortado	1.17	5.75	0.97	23.20	0.88	40.00	0.88	46.10	0.92	48.30		
		RP	Ponderado	1.23	3.80	1.03	14.65	0.77	26.90	0.70	38.05	0.60	46.80		
			Recortado	1.69	6.05	1.65	19.25	1.65	33.90	1.75	43.40	1.82	46.65		
		Modal	Ponderado	0.24	29.80	0.23	81.60	0.21	99.30	0.20	100.00	0.18	100.00		
			Recortado	0.81	38.40	0.77	87.90	0.73	99.80	0.72	100.00	0.69	100.00		

Tabla 4.8: Resultados de los valores empíricos de $\hat{\rho}_f$ y $\hat{\rho}_c$ con $\rho = 0.9$ para el caso dependencia entre curvas.

propuesta de gráficos de control para la Fase I, incluso en escenarios de dependencia entre curvas.

4.3.2. Medición y comparación del rendimiento del gráfico de control propuesto para la Fase II

Para la Fase II, la etapa de monitorización, se ha propuesto el uso del gráfico de control de rangos que permite monitorizar simultáneamente cambios en la media y la variabilidad de un proceso. Para el caso funcional con el objeto de calcular el estadístico rango, se utilizan las profundidades funcionales FM, RP y Modal.

Para evaluar el desempeño del gráfico de control, se considera un $ARL_0 = \frac{1}{\alpha=0.025}$ (la muestra de monitorización se considera que esta bajo control), y mediante un estudio de Monte Carlo se estima y se compara la potencia del gráfico de control cuando el proceso no está bajo control a partir de la generación de una muestra de calibrado de tamaño $n = 50$.

Siguiendo el esquema de simulación de la Fase I, se consideran curvas observadas en puntos equidistantes, siendo el número de puntos 51 en el intervalo $[0, 1]$. A partir de 1000 remuestras ($B = 1000$) y con un procedimiento de Recortado del 2.5% (eliminación de las curvas menos profundas), se aplica un bootstrap suavizado con un factor de suavización $\gamma = 0.05$ con el fin de estimar y comparar la potencia del gráfico de control para detectar señales de fuera de control cuando se asume un nivel de significación de $\alpha = 0.025$. Además, de igual forma que en la Fase I, considera la simulación de escenarios de independencia y dependencia entre curvas.

δ	FM	RP	Modal
0.5	14.5	14.8	14.8
1	47.6	47.2	48.0
1.5	83.3	80.8	83.0
2	97.8	97.5	97.7

Tabla 4.9: Potencia del gráfico de control $1 - \beta$ para el caso M_1 (expresión 4.5) en el escenario de independencia entre curvas.

En la Tabla 4.9 se presentan los resultados para el caso de independencia, mientras que en la Tabla 4.10 el de dependencia los resultados de la simulación para evaluar la capacidad del gráfico de control para detectar un cambio de la magnitud en la

media funcional del proceso a través de la estimación de su potencia $(1 - \beta)$.

A partir de los resultados de la Tabla 4.9, se tiene que, en general, para detectar un desplazamiento de la media del proceso se puede utilizar cualquier medida de profundidad, ya que obtiene el mismo desempeño en términos de potencia.

ρ	δ	FM	RP	Modal
0.1	0.5	16.70	20.07	17.31
	1	53.87	56.00	53.78
	1.5	86.86	86.68	87.40
	2	98.61	98.07	98.88
0.3	0.5	20.28	23.73	21.84
	1	62.95	64.75	64.06
	1.5	92.98	92.33	93.64
	2	99.56	99.16	99.75
0.5	0.5	25.59	29.41	28.41
	1	73.84	75.12	76.24
	1.5	97.15	96.56	97.94
	2	99.91	99.77	99.98
0.7	0.5	35.47	40.23	43.55
	1	86.14	86.87	90.38
	1.5	99.48	99.11	99.78
	2	100.00	99.99	100.00

Tabla 4.10: Potencia del gráfico de control $1 - \beta$ para el caso M_1 (4.5) en el escenario de dependencia

Los resultados de detectar un desplazamiento de la media del proceso se presentan en la Tabla 4.10. En general, se tiene un desempeño similar al utilizar cualquier medida de profundidad para los diferentes valores de ρ . Aparentemente, el gráfico de control para la Fase II es robusto frente a la existencia de dependencia entre las curvas o datos funcionales.

4.4. Aplicación a datos reales

En esta sección se muestra la utilidad y el desempeño de la nueva metodología de gráficos con control para datos funcionales, evaluados en el caso de estudio de la detección de anomalías correspondientes a la eficiencia energética de una instalación HVAC. En particular, aquella correspondiente a una superficie comercial de una conocida marca gallega de modal, situada en Ciudad de Panamá. Como ya se ha

indicado, este caso de estudio es en sí un estudio controlado en el que se han detectado previamente cuáles son las anomalías que se deberían identificar y sus causas asignables. Para más información acerca de los datos, consúltese la sección 2.2.2.

A continuación se muestra la necesidad del desarrollo y aplicación de metodologías FDA para gráficos de control, teniendo en cuenta la observación de los datos del presente caso de estudio, en particular los correspondientes al mes de agosto. Como se ha mencionado anteriormente, en el mes de agosto no se registraron eventos que desestabilicen el proceso por causas asignables, sin embargo, al utilizar una metodología como la usual para datos escalares (ignorando la autocorrelación entre las variables) se puede detectar un número inaceptable de falsas alarmas. En el caso escalar, una herramienta muy habitual para detectar datos anómalos o atípicos son los boxplot.

En la Figura 4.4, se muestra un enfoque escalar tradicional en la detección de atípicos mediante el uso de boxplot. En el panel izquierdo se presentan los boxplot para cada variable de consumo energético en HVAC por hora, mientras que en el panel derecho se muestran las curvas de consumo energético diario en HVAC, resaltando aquellas que han sido detectadas como atípicas mediante el procedimiento descriptivo basado en la aplicación de gráficos de caja a cada consumo horario. En el procedimiento habitual, se consideran curvas atípicas aquellas curvas que en al menos un boxplot presenten atípicos, pero tiene el inconveniente de que aumenta la probabilidad de error tipo I. Este procedimiento sobre estima el número de atípicos, considera 12 curvas diarias de consumo como datos anómalos. En la siguiente sección se presenta la metodología desarrollada con base a la adaptación de procedimientos FDA para la detección de datos atípicos.

Con base a la información descrita en esa sección se procede primero aplicar el gráfico de control de profundidades para la Fase I y posteriormente se utilizan los gráfico de rango para la monitorización del proceso. La aplicación de estas dos técnicas estadísticas, unida a la herramienta gráfica aportada, constituye el nuevo procedimiento propuesto de gráficos de control cuando los datos a estudiar son funcionales. En general se procede del siguiente modo:

- Se consideran las curvas de consumo energético en HVAC de los correspondientes meses de agosto y septiembre como muestra de calibrado. Se ha tomado un

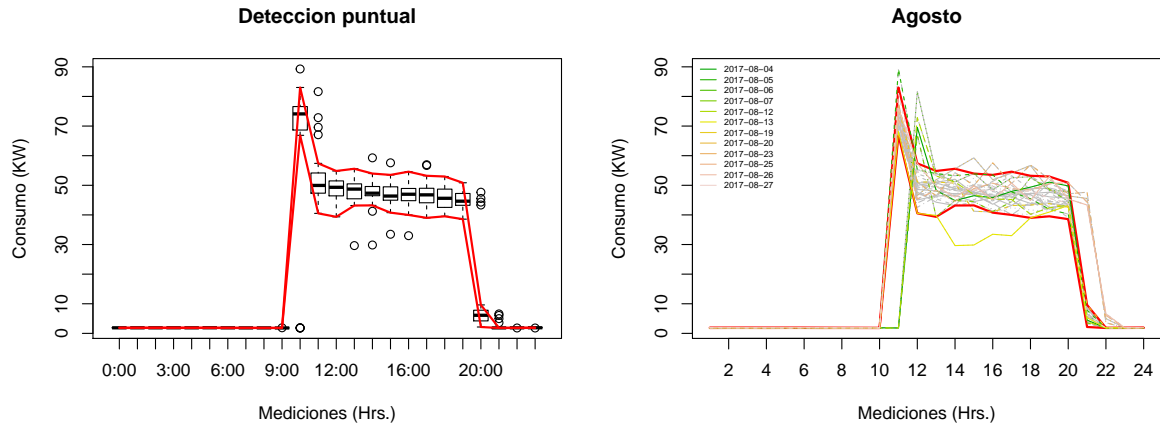


Figura 4.4: Detección de consumos energéticos en HVAC atípicos mediante la aplicación de boxplot a cada uno de los consumos horarios.

rango temporal amplio con el objeto de estimar convenientemente los límites de control naturales del consumo energético en HVAC.

- Para la Fase I, se estabiliza la muestra de referencia mediante el gráfico de control para las profundidades de las curvas. Es decir, se eliminan de los cálculos para estimar los límites de control aquellas curvas detectadas como atípicas mediante métodos de detección de datos anómalos FDA.
- Se monitorizan los meses de octubre, noviembre y diciembre para comprobar que no exista una desviación en el proceso. Si existiese un cambio, es decir la presencia de días de consumo atípicos durante esos meses, entonces se eliminan las fuentes de causas asignables.

En la Figura 4.5, se presentan las curvas de color negro para los días de agosto (23 curvas) y las grises para los días de septiembre (21 curvas). Sólo se consideran días de lunes a viernes. En septiembre se detectaron anomalías (ver Figura 4.5, curvas marcadas en rojo). Entre las anomalías se tienen:

- El día 11 hubo una bajada en el consumo del aire hacia el medio día.
- Los días 21, 22 el establecimiento estuvo cerrado, por lo que no hubo consumo y la temperaturas fueron altas.

- El día 27 de septiembre se realizaron labores de mantenimiento en las instalaciones de la tienda, que se tradujo en un comportamiento anómalo de las curvas de consumo energético.
- El día 29 de septiembre se apagó la instalación de aire acondicionado una hora antes de lo habitual de forma no programada.

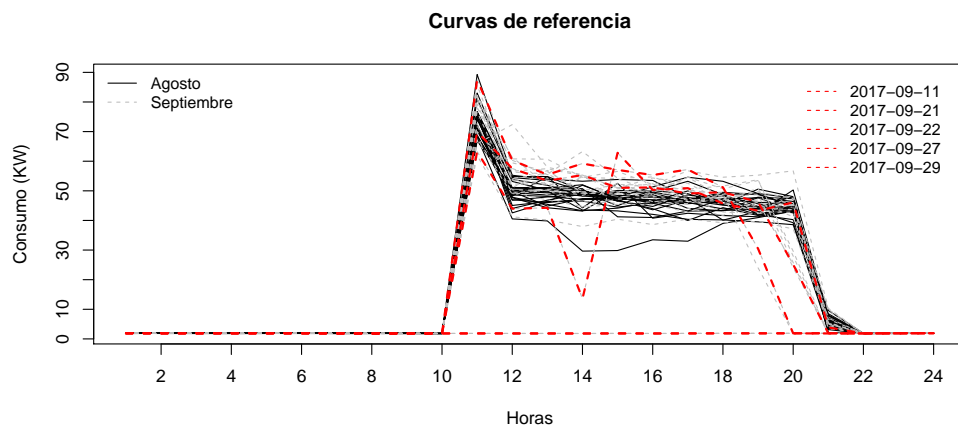


Figura 4.5: Curvas diarias de consumo energético en instalaciones HVAC para la tienda de Panamá City. Se ha representado las curvas pertenecientes a la muestra de calibrado, a partir de las cuales se estiman los límites de control naturales (Fase I).

A partir de esta muestra de referencia y con base a los resultados de las simulaciones presentadas para el gráfico para las profundidades funcionales, se utiliza una medida de profundidad funcional modal y el método para la detección de curvas atípicas ponderado. Además se fija un nivel de significación $\alpha = 0.025$ para la estimación del LCI a partir de $B = 500$ muestras bootstrap y un coeficiente de suavizado $\gamma = 0.8$ (este parámetro es fijado a partir de que la detección de atípicos no dependa de su valor, si se consideran valores menores al valor fijado se detectan un número menor de curvas y si este parámetro toma valores mayores al fijado el número de detección de atípicos no varía) y porcentaje de Recortado $trim = 0.05$ que permite obtener una envolvente del 95% de las curvas más profundas. La ventaja de este procedimiento de detección de curvas atípicas para llevar a cabo la Fase I de la propuesta de gráfico de control para datos funcionales es su flexibilidad para adaptarse

a una amplia variedad de problemas reales, mediante la regulación de sus parámetros (γ y *trim*).

En el panel izquierdo de la Figura 4.6 se tiene, además de las curvas originales en gris, la estimación de la mediana (curva azul) y la media recortada (curva roja) funcionales, además de la envolvente estimada a partir del 95 % de las curvas más profundas, representadas a su vez en color rojo. Las curvas detectadas como atípicas son aquellas curvas representadas en color gris con punteado en negro. Una vez que se identifican las mencionadas curvas de consumo energético atípicas, se procede a eliminarlas de todo cálculo relacionado con la estimación de los límites de control naturales, dado que se han identificado sus causas asignables (causas aparte de la propia aleatoriedad de los datos). Acto seguido, se vuelve a repetir el proceso según un esquema iterativo hasta que no se detecten curvas atípicas asociadas a causas asignables, es decir, hasta que el proceso que se mide (en este caso la eficiencia energética de las instalaciones HVAC) esté bajo control.

En el panel derecho de la Figura 4.6, se muestra el gráfico de control rangos adaptado para ser aplicado a datos funcionales, donde cada punto representa una curva diaria de consumo energético en HVAC, mediante el cálculo de su profundidad. Los primeros 23 puntos representan las curvas de consumo del mes de agosto y los siguientes 21 puntos se corresponden con las curvas de consumo del mes de septiembre.

En esta primera iteración se detectan las curvas identificadas previamente como anómalas (cuya estructura se corresponde a una causa de variación asignable), que son las correspondientes a los días 11, 21, 22 y 29 de septiembre. Sin embargo en esta iteración no se detecta la curva anómala del 27 de septiembre. También, se detectan las curvas correspondientes al 23 de agosto, 7, 21, 20 y 28 de septiembre. Además, se puede observar que a partir de la curva 36 (19 de septiembre) las medidas de profundidad son muy pequeñas (de hecho están fuera del límite de control, detectándose como anomalías). Esto es debido a que a partir de ese día el aire acondicionado se apaga media hora antes.

En la siguiente iteración se detectan como anómalas las siguientes curvas de consumo diarias: 15, 19, 25, 27 del mes de septiembre y el 04 de agosto. En esta iteración se detecta la curva de consumo del día 27 de septiembre, además de las

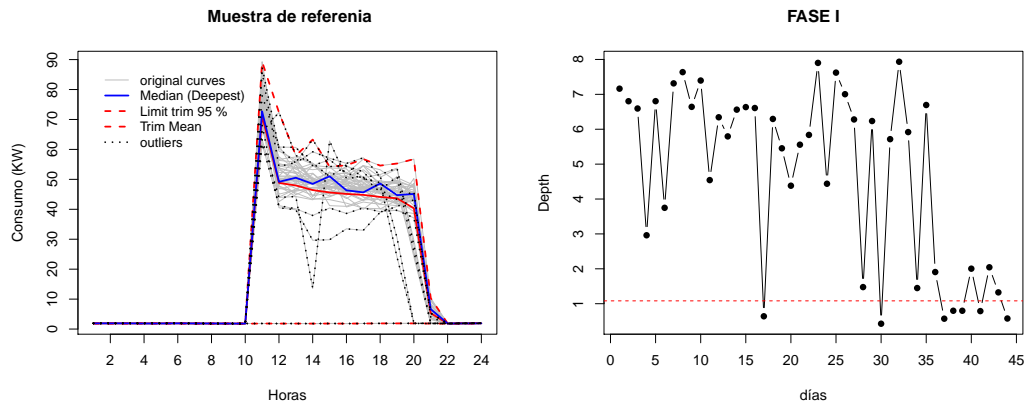


Figura 4.6: Panel izquierdo: Curvas diarias de consumo energético en instalaciones HVAC para la tienda de Panamá City, identificando aquellas curvas detectadas como anomalías, además de medidas de posición funcionales. Panel derecho: Gráfico de control correspondiente a la Fase I, mediante el cual se detectan las curvas anómalas, eliminando así las causas asignables de variación. Se muestran las profundidades de cada una de las curvas de consumo diario y el límite de control natural.

curvas correspondientes a los últimos días del mes de septiembre, para los que se cambió la hora de apagado del sistema HVAC. Finalmente, después de la segunda iteración, no se detectan curvas de consumo como atípicas (Ver Figura 4.7).

El proceso se inició con 44 curvas, de las cuales fueron detectadas como anómalas una cantidad de 9 y 5, correspondientes a la primera y segunda iteración del proceso iterativo anteriormente descrito. De estas 14 curvas de consumo energético, dos corresponden al mes de agosto y 12 al mes de septiembre. Como resultado, la muestra de calibrado resultante está compuesta por días de agosto y septiembre, hasta el 18 de este último mes incluido.

Una vez llevada a cabo la Fase I, se procede a realizar las actividades correspondientes a la Fase II de la construcción de gráficos de control para datos funcionales, también llamada fase de monitorizado. Por tanto, la muestra a monitorizar se compone por los días correspondientes al mes de octubre, en el que las instalaciones de HVAC de la tienda de ropa volvieron a funcionar en horario habitual. Téngase en cuenta que, como ya se ha comentado, a mediados de este mes se registra una fuga en el circuito del aire acondicionado. De hecho, a partir de ese momento el consumo

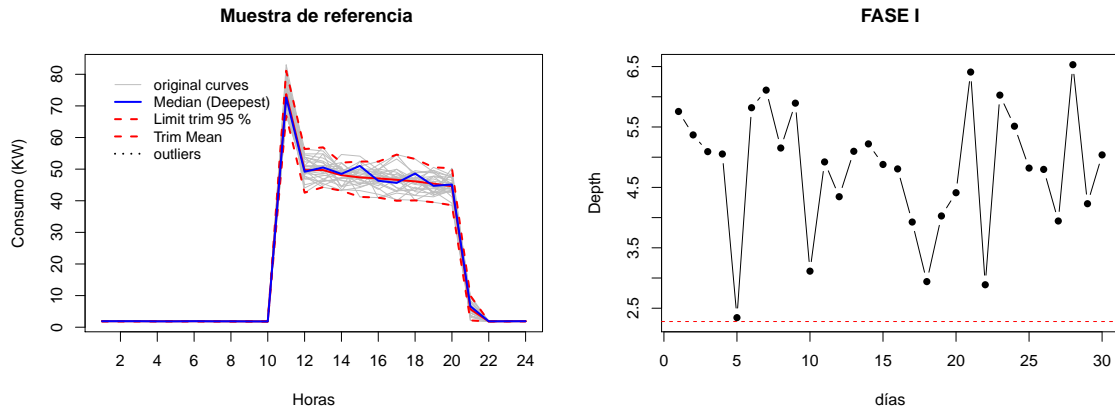


Figura 4.7: Proceso bajo control (consumo energético en instalaciones HVAC estabilizado). Panel izquierdo: Curvas diarias de consumo energético en instalaciones HVAC para la tienda de Panamá City, indicando la envolvente correspondiente al 95% de curvas más profundas. Panel derecho: gráfico de control para la Fase I basado en la profundidad de datos funcionales.

empezó a subir. Este comportamiento se puede observar en la Figura 4.8, en la que, además de la muestra de monitorizado, se representa las curvas de la muestra de calibrado y su envolvente estimada en la Fase I.

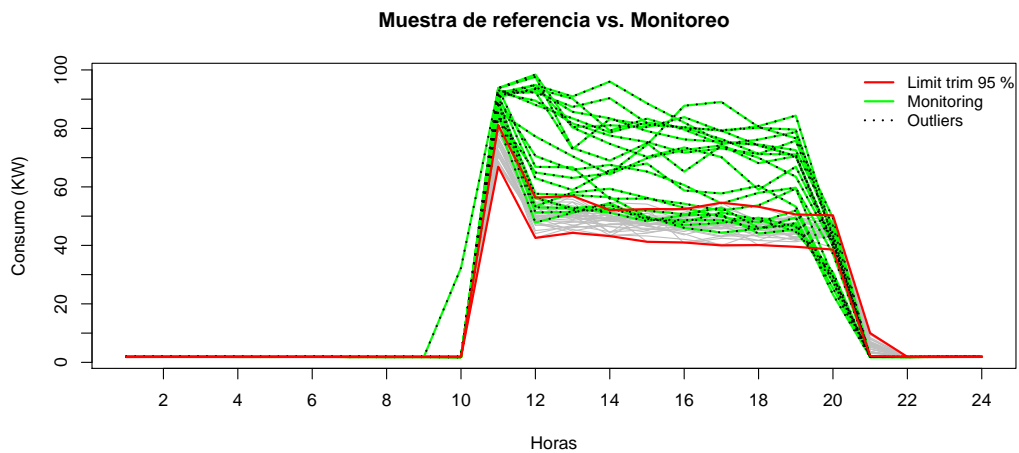


Figura 4.8: Muestra de referencia vs. monitorizada para el mes de octubre

Como se puede apreciar en la Figura 4.8, se observa que el consumo en HVAC, variable crítica de la calidad de la eficiencia energética de las instalaciones HVAC, está fuera de control, siendo la causa asignable la fuga en el sistema de aire acondicionado. Con el objeto de subsanar el funcionamiento anómalo de las instalaciones HVAC, el 1 de noviembre se llevó a cabo una reparación provisional. Estas acciones produjeron una disminución del consumo, además de la atenuación del pico de consumo correspondiente al arranque (obsérvense las curvas antes de las 11:00). Como consecuencia, el consumo de las instalaciones en noviembre es ligeramente diferente al de agosto y septiembre, meses correspondientes a la muestra de calibrado. Sin embargo, este consumo comienza a subir nuevamente entre los días 17 y 20.

Estos cambios se puede apreciar en la Figura 4.9, donde se observa que el consumo energético en HVAC de la muestra de monitorizado está caracterizada por una mayor variabilidad que la muestra de calibrado o referencia. Tanto para las muestras de monitorización para los meses de octubre y noviembre el procedimiento del gráfico de control rank considera que el proceso está fuera de control, en el primer mes por la presencia de la avería y en el segundo mes por el cambio en el proceso, reducción del consumo.

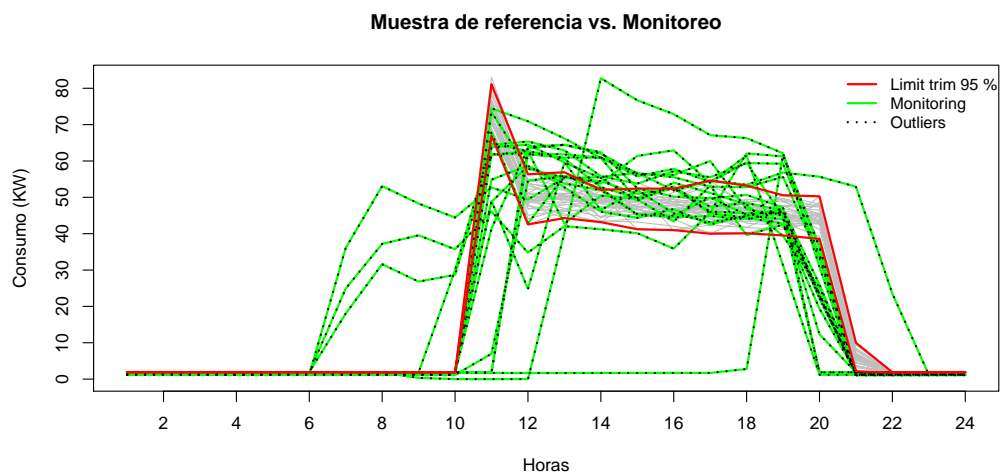


Figura 4.9: Muestra de referencia vs. monitoriada para el mes de noviembre

Debido a que el proceso ha cambiado, se debe considerar una nueva muestra de

referencia e iniciar el proceso de estabilización y de monitorización para el siguiente año ya que diciembre debe presentar altos consumos.

Es decir, considerando las curvas diarias de consumo y sus derivadas. En la Figura 4.10 se presenta el gráfico de rango para las muestras de referencia o calibrado y monitorización del mes de noviembre. En la práctica, realizar la monitorización mediante los gráficos propuestos es muy eficiente, ya que no se necesita estimar el LCI mediante procedimientos de remuestreo. Además, el gráfico de control de rangos se puede adaptar para la monitorización de datos funcionales multivariantes, es decir, cuando se tienen varios tipos de curvas que a la vez definen la calidad de un determinado sistema o proceso. Por ejemplo un caso directo de aplicación es considerar, además de las curvas originales también su derivada.

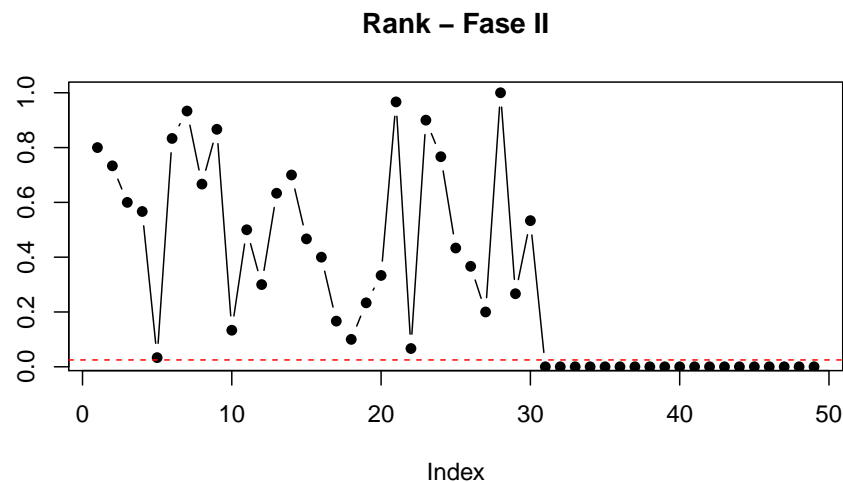


Figura 4.10: Gráfico rank: Muestra de referencia vs. monitorización para el mes de noviembre

La aplicación de los gráficos de rangos, permiten además realizar un proceso de monitorización considerando no sólo las curvas de consumo energético, sino también las curvas de temperatura diaria, humedad relativa diaria, concentración de CO₂ diaria, etc., que caracterizan completamente no sólo la eficiencia energética de las instalaciones del edificio o superficie comercial, sino también su confort térmico y la calidad del aire ambiente. Es decir, realizar una monitorización multivariante. Esta

tarea es una línea de investigación futura que arranca de las investigaciones realizadas en esta tesis doctoral que deberá considerar procedimientos donde se considere que las escalas de unas variables no enmascaren a las otras.

Como se ha podido observar en el estudio de simulación y en la resolución del caso de estudio con datos reales, la nueva propuesta de gráficos para datos funcionales, incluyendo las fases de desarrollo I y II, es de utilidad para detectar anomalías en una amplia variedad de escenarios. En el caso de su aplicación a datos reales, el conjunto de técnicas propuestas está actualmente en proceso de estudio para su implementación en la plataforma web Σ qus y su utilización por la empresa Nerxus, en problemas de detección de alarmas en las instalaciones correspondientes a superficies comerciales. La metodología de gráficos de control que aquí se propone, podría utilizarse para tareas de control, monitorización, detección de anomalías y mejora continua en gran variedad de procesos industriales, monitorización de variables ambientales, industria química y, en general, todo proceso que conlleve la monitorización continua de datos a través del tiempo y cuya naturaleza sea funcional. Esta tarea se llevará a cabo en investigaciones y aplicaciones futuras.

Capítulo 5

Conclusiones y líneas futuras

5.1. Estudios Interlaboratorio

En esta tesis se proponen nuevas metodologías para el control de calidad usando datos funcionales. Estas nuevas propuestas vienen a cubrir la monitorización de datos obtenidos por sensores donde la componente funcional es una característica común, tanto en el contexto de datos de laboratorio como en los datos de sensores. Para el caso de tratar los datos de laboratorio se ha propuesto una extensión para la detección de atípicos y también una nueva propuesta para ILS.

Se ha propuesto extensiones funcionales para los estadísticos h y k de Mandel con el objeto de tratar las curvas (datos funcionales) obtenidas por técnicas experimentales en laboratorio. Estas extensiones han sido desarrolladas para identificar los laboratorios que proporcionan resultados no consistentes o atípicos en un estudio ILS, utilizando la información proporcionada por toda la curva. En consecuencia, se proporciona un enfoque alternativo para realizar la detección de valores atípicos en los laboratorios involucrados en un estudio ILS cuando los datos experimentales son funcionales, frente a los enfoques escalares actuales. El uso de esta nueva metodología propuesta evita el paso de extracción de características a partir de curvas experimentales.

Por otra parte, otro de los objetivos del nuevo procedimiento presentado para la detección de valores atípicos es la identificación directa de los laboratorios inconsistentes. Esto representa una mejora sustancial frente a otras alternativas FDA

previas en las que solo se proporcionaban los datos atípicos independientemente del laboratorio. De este modo, la metodología propuesta ofrece una mejora que permite trabajar con datos funcionales en la misma manera que los enfoques escalares h y k , que son las alternativas más populares en los estudios ILS para la detección de atípicos.

El enfoque FDA propuesto consiste en el cálculo de los estadísticos funcionales H y K (obtenidos puntualmente en el intervalo donde se definen las curvas), de los cuales se obtienen los estadísticos de prueba d_H y d_K a partir de la distancia L_2 . El estadístico d_H se define para probar la hipótesis de reproducibilidad, evaluando la variabilidad entre laboratorios, mientras que la hipótesis de repetibilidad (variabilidad intra-laboratorio) es probada mediante el uso del estadístico d_K .

Las distribuciones de probabilidad de los estadísticos de prueba d_H y d_K se estiman mediante un procedimiento bootstrap y, por lo tanto, los valores críticos c_H y c_K se pueden obtener como los cuantiles correspondientes a un nivel de significación determinado ($\alpha = 0.01$). Obsérvese que, antes de la aplicación del remuestreo bootstrap, el cálculo de la profundidad de los datos funcionales es necesario para eliminar las curvas atípicas. El estudio de simulación realizado con curvas TG ha proporcionado información sobre la validez y la potencia de las nuevas extensiones FDA de los estadísticos h y k , además de los efectos del tamaño de la muestra.

Adicionalmente, se obtiene que los p -valores bajo la hipótesis nula de reproducibilidad o repetibilidad se distribuyen uniformemente, apoyando la validez de la propuesta de los estadísticos de prueba d_H y d_K . Además, todos los laboratorios simulados como atípicos han sido identificados con éxito mediante la aplicación de la propuesta, en la que los estadísticos H y d_H han sido utilizados para detectar cambios entre laboratorios, mientras que, por otro lado, K y d_K se usan para identificar las diferencias intra-laboratorio.

Con respecto a la potencia del test, cuando el número de réplicas por laboratorio no es muy alto ($n = 10$), la potencia del test d_H es mayor que la del estadístico h escalar de Mandel cuando se extraen características como el tiempo/temperatura necesario para perder el 5 % de la masa. Por otro lado, como se esperaba, al aumentar el número de repeticiones de laboratorio, aumenta significativamente la potencia de los estadísticos de prueba. Así también, para los casos observados, se puede verificar

que tanto la potencia de d_K como de d_H siempre son mayores que las correspondientes a los estadísticos escalares k y h , cualquiera que sea la característica extraída.

La aplicación del enfoque FDA a los datos termogravimétricos es capaz de detectar todos los laboratorios atípicos verdaderos. Para esta tarea, el uso de los estadísticos de prueba d_H y d_K o los estadísticos funcionales H y K han resultado alternativas mejores que las metodologías escalares clásicas. Además, el uso de los estadísticos funcionales H y K permite identificar el intervalo de tiempo/temperatura en el que los laboratorios presentan comportamientos atípicos.

5.2. Gráficos de control

Se ha propuesto una nueva alternativa de gráficos de control cuando la variable crítica para la calidad del proceso es funcional. La propuesta incluye alternativas para desarrollar la Fase I, para estabilizar el proceso, y la Fase II, de monitorización, de los gráficos de control. Para el desarrollo de la Fase I, se propone el uso de métodos de detección de atípicos para datos funcionales, de hecho el gráfico de control propuesto esta basado en el método de [Febrero-Bande *et al.* \(2007\)](#), el cual se basa en un método de remuestreo. Por otro lado, para implementar la Fase II, se propone el uso de gráficos tipo rank construidos a partir del concepto de profundidad de datos funcionales. Este gráfico se calcula directamente el LCI mediante el supuesto de que el estadístico rank sigue una distribución uniforme asintóticamente. La combinación de las dos fases, unida al desarrollo de una nueva herramienta gráfica, dan lugar a la metodología propuesta, que ha sido aplicada con éxito en casos de estudios reales. Además se ha realizado un estudio de simulación que permite comparar el desempeño (medido como el porcentaje de rechazo cuando no se cumple la hipótesis nula) de los gráficos dependiendo de la profundidad de datos funcional utilizada.

En lo que respecta al estudio de simulación, se ha comparado el uso de diferentes tipos de profundidades funcionales para el desarrollar la Fase II del gráfico de control propuesto. En caso de datos funcionales univariantes (un solo tipo de curvas), para los tres escenarios, se tiene un mejor desempeño con la medida de profundidad modal combinada con el método de detección de atípicos ponderado.

5.3. Líneas futuras

Las futuras líneas de investigación en el campo de gráficos de perfiles están relacionadas con la necesidad de desarrollar gráficos de control de perfiles con submuestras que contengan más de un perfil, es decir, la alternativa a los gráficos de control Shewhart de la media, teniendo en cuenta la autocorrelación entre perfiles y su estructura de varianzas covarianzas, además de la necesidad de más trabajos para poder determinar el error de estimación en la Fase II de los gráficos de control (Chipman *et al.*, 2010). A este respecto es importante mencionar que, si bien se puede considerar un problema abordable mediante el uso de técnicas FDA, todavía no se han implementado gráficos de control para perfiles que tengan en cuenta la detección de atípicos mediante técnicas FDA propiamente dichas, para la Fase I, y la utilización del concepto de profundidad de datos funcionales para la Fase I y II.

En este trabajo se han aportado nuevas metodologías y herramientas informáticas que pretenden solucionar problemas actuales en la industria y laboratorios de pruebas relacionados con el control de la calidad. Básicamente, se han realizado contribuciones en dos líneas de investigación: gráficos de control e estudios interlaboratorio. A pesar de haber cumplido con los objetivos planteados en esta memoria se han podido identificar varias oportunidades para seguir generando nuevas adaptaciones al caso funcional de herramientas y metodologías escalares y multivariantes útiles en el control de la calidad.

Se está trabajando en la adaptación del gráfico rango para el caso de datos funcionales multivariantes es directa, con esto además de considerar las curvas originales, también se podría considerar su derivada si no se cuenta con otro tipo de curvas. Aparte de esto, se puede realizar un proceso de monitorización considerando diferentes variables funcionales que definen un sistema. Por ejemplo, las curvas de consumo energético, temperatura, humedad relativa, etc. que definen la eficiencia energética y el confort térmico de superficies comerciales.

Además, el desarrollo de índices de capacidad que no se han abordado en esta memoria es otra línea de investigación en la que se tiene interés en realizar investigaciones con la finalidad de proponer metodológicas y herramientas informáticas que permitan su aplicación en un contexto funcional. En este sentido, se abordará su estudio desde la relación con los gráficos de control de la calidad y el uso de medidas

de profundidad funcional.

Apéndice A

Control estadístico de la calidad con el paquete qcr

En este apéndice, se presenta y describe el paquete en R para control estadístico de calidad (SQC). El paquete incluye actualmente un conjunto completo de herramientas SQC univariadas y multivariadas que completa e incrementa las técnicas de SQC disponibles en R. Actualmente, se está trabajando en la implementación de las nuevas herramientas desarrolladas en la memoria de tesis para realizar gráficos de control para datos funcionales.

El objetivo principal del paquete `qcr`, es combinar procedimientos SQC flexibles, tradicionales y nuevos para tratar problemas reales de control de calidad en la industria y la consultoría. Además de integrar diferentes paquetes en R dedicados a SQC (`qcc`, `MSQC`), proporciona nuevas herramientas no paramétricas muy útiles cuando no se cumple la suposición gaussiana. Este paquete proporciona el conjunto más completo de funciones en R para calcular gráficos de control de atributos y variables, desde un punto de vista paramétrico y no paramétrico, de una manera univariante o multivariante. Para ser aplicado en problemas reales de la industria, permite estimar los límites de control naturales y monitorear las variables críticas de una manera práctica y más automática.

El paquete propuesto actualmente calcula gráficos de control univariados estándar de medidas individuales: \bar{x} , S , R , p , np , c , u , EWMA y CUSUM. Además, incluye funciones para realizar gráficos de control multivariante como Hotelling T^2 , MEWMA

y MCUSUM.

Además, nuevas alternativas no paramétricas basadas en la profundidad de datos (Liu, 1995) se implementan en este paquete: gráficos de control r , Q y S . En esta línea, ahora hay disponibles alternativas robustas de índices de capacidad en R a través de la librería `robust`. La librería también permite estimar el conjunto más completo de índices de capacidad univariante y multivariante de primera a cuarta generación, y realizar los gráficos de análisis de capacidad correspondientes.

A.1. Introducción

A lo largo de las últimas décadas, ha aumentado el interés por medir, mejorar y controlar la calidad de los productos, servicios y procedimientos. Esto está conectado a la fuerte relación entre calidad, productividad, prestigio, confianza e imagen de marca. Los procedimientos de implementación del control de calidad estadístico (SQC) están actualmente relacionados con el aumento de la competitividad de las empresas.

El concepto de control de calidad se ha extendido desde las primeras definiciones basadas en la idea de ajustar la producción a un modelo estándar para satisfacer los requisitos del cliente e incluir a todos los participantes. Hoy en día no solo se aplica a productos manufacturados sino a todos los procesos industriales y de servicio.

Con el desarrollo de Six Sigma por Motorola en 1997, el uso de diferentes técnicas de SQC ha sido estandarizado. Six Sigma es una metodología o incluso una filosofía centrada en la reducción de la variabilidad que promueve el uso de métodos y herramientas estadísticas para mejorar los procesos en la industria y los servicios. El objetivo de Six Sigma es alcanzar hasta 3,4 defectos por millón de eventos u oportunidades (DPMO) que cumplan con los requisitos del cliente. La aplicación Six Sigma está compuesta por cinco etapas: Definir, Medir, Analizar, Mejorar y Controlar (DMAIC). La metodología Six Sigma en sí misma indica qué técnica aplicar en cada etapa del proceso de mejora. Las dos herramientas estadísticas exclusivas de SQC son los gráficos de control y el análisis de capacidad de proceso. Por lo tanto, el paquete propuesto se ha desarrollado para tratar estos campos principales en SQC. La etapa de control se caracteriza por el uso de herramientas basadas en la detec-

ción y corrección de errores, los más representativos son los gráficos de control. Los gráficos de control de calidad son la herramienta principal utilizada en el Control estadístico del proceso (SPC) entre todas las metodologías utilizadas en este campo para la recopilación y el análisis de datos para evaluar el rendimiento del proceso y su estado en cualquier momento. Un gráfico de control es un gráfico bidimensional cuyo eje representa la variable o atributo que se está monitoreando. Fueron presentados por Shewhart en 1920 en los laboratorios Bell. La estimación de los límites de control natural de las variables críticas a la calidad (CTQ) se debe a un proceso compuesto por dos fases: en la Fase I, los límites de control se estiman utilizando una muestra preliminar donde las causas de variación son solo al azar. En la fase II, las muestras se trazan en un gráfico de control con los límites naturales obtenidos en el paso anterior. Por lo tanto, cualquier causa de variación asignable en la muestra de monitoreo podría detectarse usando las tablas de control.

Los gráficos de control más utilizados se basan en los supuestos de normalidad e independencia de las variables estudiadas CTQ. Estos gráficos controlan la posición y la dispersión de atributos y variables de CTQ. La figura A.1 muestra algunos de los tipos más importantes de gráficos de control. Estos pueden clasificarse según el tipo de característica que se está controlando (atributo o variable), la dimensión (univariante o multivariante) y suponiendo o no una distribución paramétrica de la variable (paramétrica o no paramétrica). Por un lado, el paquete proporciona gráficos para la media (\bar{x}), desviación estándar (S), rango (R), medidas individuales (I), rangos móviles (MR), proporción de unidades no conformes (p), número de unidades no conformes (np), número de disconformidades por unidad (c), número medio de no conformidades por unidad de control (u), promedio móvil ponderado exponencialmente (EWMA) y gráfico de control de suma acumulada (CUSUM) (los dos últimos se dedicaron a detectar desviaciones de menos de dos desviaciones estándar en muestras racionales o mediciones individuales). Por otro lado, se implementan nuevos gráficos de control basados en el concepto de profundidad de datos y desarrollados por (Liu, 1995) en . Esos son los gráficos de control r, Q y S, las alternativas no paramétricas para mediciones individuales, el gráfico de control medio y el gráfico de control CUSUM, respectivamente. Cuando más de una variable define la calidad del proceso, se aplican gráficos de control multivariante. Si se cumple la suposición gaus-

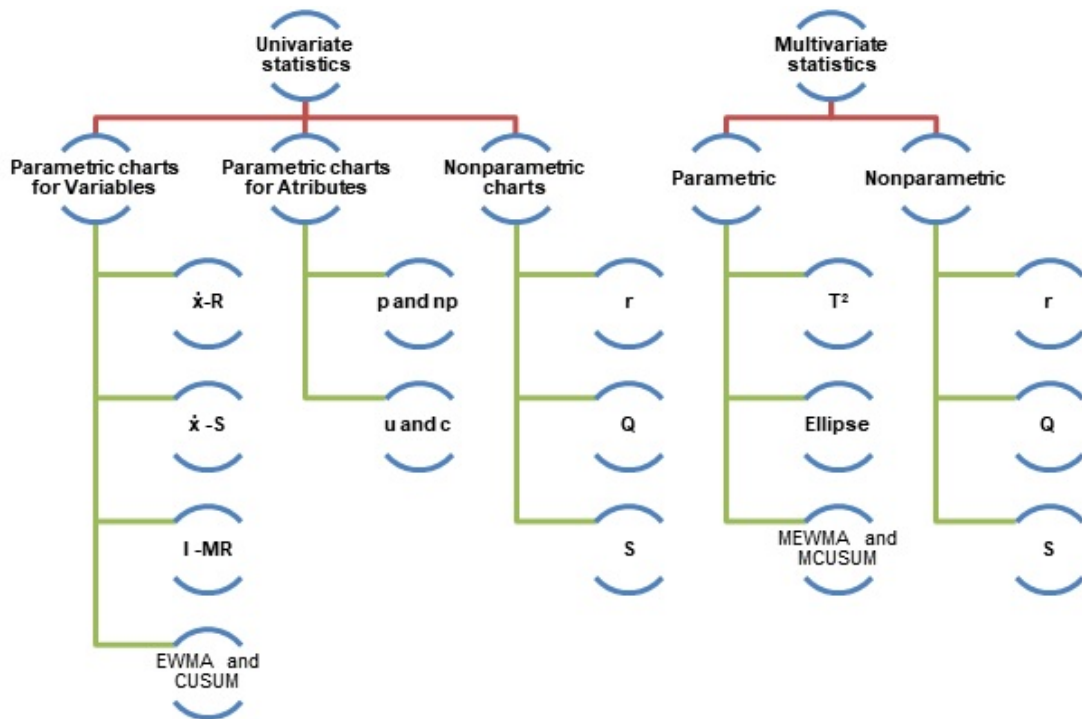


Figura A.1: Tablas de control implementadas en el paquete.

siana, se puede aplicar el gráfico de control de T^2 de Hotelling . Si queremos detectar pequeñas desviaciones, se puede implementar EWMA multivariante (MEWMA) y CUSUM multivariante (MCUSUM). Cuando se asume una distribución paramétrica, se pueden usar gráficos r, Q y S.

Otro componente interesante de SQC es el Análisis de Capacidad de un Proceso (PCA), un dominio muy útil en la industria. Calcula qué tan bien un proceso cumple con las tolerancias definidas por la compañía, los clientes, las regulaciones, etc. mediante la comparación de la tolerancia de especificación con respecto al rango natural de variación de las características CTQ. La capacidad del proceso se mide mediante índices de capacidad, por lo tanto, la Capacidad de Procesamiento (PCR) es una calificación numérica que ayuda a los fabricantes a saber si la salida de un proceso cumple con las especificaciones de ingeniería. Los grandes valores de PCR muestran que el proceso industrial o de servicio es capaz de cumplir los requisitos del cliente. Existen muchas PCR diferentes desarrolladas en las últimas cuatro décadas que requieren la suposición gaussiana para la variable CTQ medida (Boyles, 1991).

Pero muchos procesos en la industria y las aplicaciones reales no cumplen con esta suposición, por lo que podríamos estimar erróneamente la capacidad mediante PCR, por lo tanto, muchos autores han estudiado diferentes alternativas no paramétricas al PCR (Polansky, 2007) tradicional.

Tabla A.1: PCR de primera a cuarta generación, USL es el límite de especificación superior, LSL es el límite de especificación más bajo, μ es la media real, $\hat{\mu}$ es la media estimada y $\hat{\sigma}$ es la desviación estándar estimada. Tablas de control implementadas en el paquete.

Capacidad potencial	$\hat{C}_p = \frac{USL-LSL}{6\hat{\sigma}}$
Capacidad real respecto a los límites de especificación	$\hat{C}_{p,lower} = \frac{\hat{\mu}-LSL}{3\hat{\sigma}}$
	$\hat{C}_{p,upper} = \frac{USL-\hat{\mu}}{3\hat{\sigma}}$
	$\hat{C}_{pk} = \min \left[\frac{USL-\hat{\mu}}{3\hat{\sigma}}, \frac{\hat{\mu}-LSL}{3\hat{\sigma}} \right]$
Desplazamiento de la media respecto al objetivo	$\hat{C}_{pm} = \frac{\hat{C}_p}{\sqrt{1+\left(\frac{\hat{\mu}-T}{\hat{\sigma}}\right)^2}}$
C_{pk} corrección para detectar desviación respecto al destino	$\hat{C}_{pkm} = \frac{\hat{C}_{pk}}{\sqrt{1+\left(\frac{\hat{\mu}-T}{\hat{\sigma}}\right)^2}}$

El motor de desarrollo de las aplicaciones informáticas implementadas en este artículo es el software R. Como es sabido, R es un lenguaje de programación y un entorno para el análisis estadístico. Es un proyecto de software libre cuyos primeros pasos se debieron a Ross Ihaka y Robert Gentleman del Departamento de Estadística de la Universidad de Auckland. Sin embargo, las versiones sucesivas son controladas y desarrolladas por el R Development Core Team, que incluye muchos socios de todo el mundo. R se distribuye libremente bajo los términos de la Licencia Pública General (GNU) y se compila y ejecuta en diferentes plataformas: Unix, Windows y MacOS. Hoy en día hay algunos paquetes R que actualmente proporcionan herramientas de control de calidad a los profesionales. Los usos de cada uno se pueden consultar en Figure A.2.

El paquete qcc (Scrucca, 2004) fue desarrollado por el Profesor Luca Scrucca del Departamento de Economía, Finanzas y Estadística de la Universidad de Perugia. Permite realizar gráficos de control de calidad Shewhart para variables y atributos, así como los diagramas CUSUM y EWMA para variables dependientes continuas.

El análisis multivariante se realiza aplicando el gráfico de control T^2 de Hotelling. Además, tiene funciones implementadas para evaluar y calcular las curvas de operación características OC y los índices de análisis de capacidad de proceso. Los diagramas de Pareto e Ishikawa también se implementan. De lo contrario, el paquete IQCC (Recchia *et al.*, 2010) es mantenido por el profesor Emanuel P. Barbosa del Instituto de Matemáticas de la Universidad Estatal de Campinas. Tiene un menor número de gráficos de control implementados, pero incorpora gráficos multivariantes. El paquete qualityTool (Roth, 2012) fue desarrollado para ayudar al aprendizaje en ciencias de la calidad. La figura A.2 muestra algunas de sus utilidades, e.d. análisis de capacidad. Además la librería SixSigma (Emilio *et al.*, 2017) proporciona funciones alternativas a qualityTools y la posibilidad de implementar mapas de procesos.

Además, hay otros paquetes enfocados en gráficos de control. Los paquetes MCU-SUM (Gráfico de control de suma acumulativa multivariante), MEWMA (Gráfico de control de media móvil móvil ponderado exponencialmente multivariante) y MSQC (Control estadístico de calidad multivariante) han sido desarrollados por Edgar Santos Fernández de la Compañía de Telecomunicaciones de Cuba. El paquete MSQC (Santos-Fernandez, 2016) es un conjunto de herramientas para el control de procesos multivariantes. Contiene las principales alternativas para gráficos de control multivariante como Hotelling (T^2), Chi squared, MEWMA, MCUSUM y gráficos de control de Variancia generalizada. También incluye algunas herramientas para evaluar el supuesto normal multivariado. Cuando no se asume la normalidad de los datos, se pueden usar gráficos no paramétricos. El paquete mnsqc (Control de proceso estadístico multivariante no paramétrico) (Bezener y Qiu, 2011) desarrollado por Martin Bezener y Qiu de la Universidad de Peihua de Minnesota proporciona un procedimiento CUSUM alternativo.

En este apéndice, se presenta la nueva librería, que implementa la mayoría de las herramientas estadísticas para el control de calidad. El objetivo es proporcionar a la comunidad científica y a los profesionales de control de calidad una aplicación informática que permita el manejo simple y eficiente de herramientas estadísticas para el control de calidad, que son herramientas básicas en la metodología Six Sigma: gráficos de control para variables y atributos y cálculo de capacidad de los índices. Es importante enfatizar que el paquete también incluye nuevas aplicaciones tales como

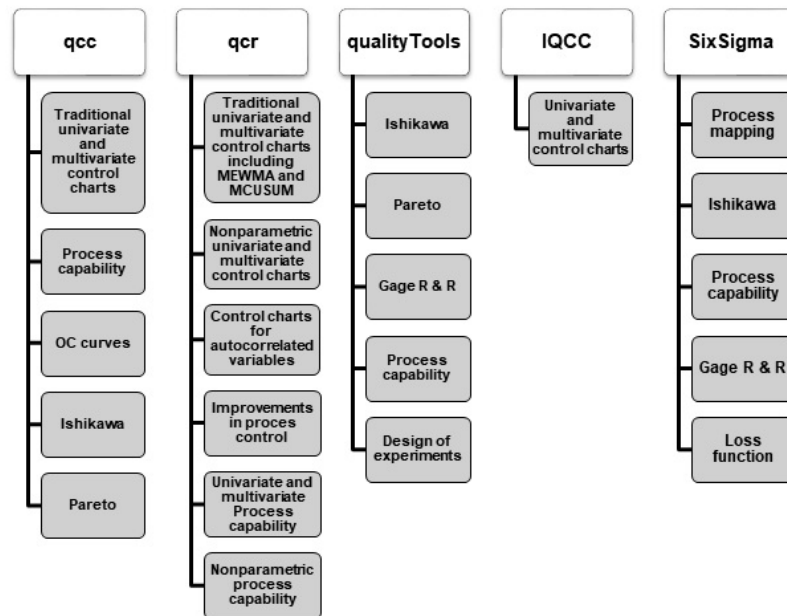


Figura A.2: Comparación entre los principales paquetes de R dedicados al control estadístico de calidad y el paquete.

enfoques no paramétricos de gráficos de control e índices de capacidad, hasta ahora no disponibles en el software R.

A.2. Creando un objeto qcs

Un objeto `qcs` es el punto de partida para la construcción de gráficos de control que en sus formas más simples requiere un vector que contenga datos y una cadena de caracteres que especifique las estadísticas del grupo para calcular. Permite generar gráficos tipo Shewhart y obtener resultados numéricos de interés para un control de calidad de proceso (que involucra datos continuos, de atributos o de recuento). También proporciona (consulte tabla A.2) funcionalidad básica para análisis de control de calidad univariado y multivariante (`mqcs`) además de gráficos de control no paramétricos (`npqcs`) en los que se debe especificar una función de profundidad, como así como índices de capacidad de proceso perimétrico y no paramétrico.

Gráficos tipo Shewhart		
tipo	Graficas de control para variables	
"xbar"	Gráficos \bar{X}	Las medias muestrales de una variable de un proceso continuo se trazan para controlar el promedio del proceso
R"	Gráficos R	Rango muestral de una variable de un proceso continuo se trazan para controlar la variabilidad del proceso.
"S"	Gráficos S	Desviación estándar muestral de una variable continua se trazan para controlar la variabilidad del proceso.
"one"	\bar{X} chart	Valores de muestra uno a la vez de una variable de proceso continua para controlar el nivel medio del proceso
tipo	Cuadros de control para atributos	
"p"	Gráficos p	Se traza una proporción de unidades no conformes, el número de elementos defectuosos sigue una distribución binomial.
"np"	Gráficos np	El gráfico se construye basándose en el promedio del proceso.
"c"	Gráfico c	Se traza una proporción de unidades no conformes, el número de defectos en una población grande sigue una distribución de Poisson.
"u"	Gráfico u	Se traza la media de las unidades no conformes, este gráfico no requiere un número constante de unidades.
"g"	Gráfico g	Cuenta el número de eventos entre errores que ocurren raramente o incidentes no conformes.
tipo	Cuadros de control para atributos y variables	
"cusum"	Gráfico CUSUM	Se trazan sumas acumulativas para observaciones individuales o para los promedios de subgrupos racionales para monitorear la media del proceso.
"ewma"	Gráfico EWMA	Se trazan promedios actuales y de observaciones, donde se da más peso a las observaciones más recientes

Tabla A.2: Cuadros de control de Shewhart disponibles en el paquete

Se devuelve un objeto de clase `qcs`. Las estadísticas de resumen se pueden recuperar usando la función `summary`.

A.3. Conjuntos de datos en el paquete qcr

El paquete contiene nuevas bases de datos (ver tabla A.3), así como conjuntos de datos implementados en otros paquetes dedicados al control estadístico de calidad, tales como:

- **arquería1**: Consiste en una etapa en la que el arquero dispara 72 flechas. La información se da en coordenadas x e y . Se implementa en el paquete MSQC ((Santos-Fernandez, 2016)).
- **circuit**: Número de no conformidades observadas en 26 muestras sucesivas de 100 placas de circuito impreso. Se implementa en el paquete qcc ((Scrucca, 2004)).
- **dowel1**: Diámetro y longitud de un pasador de espiga. Se implementa en el paquete MSQC ((Santos-Fernandez, 2016)).
- **orangejuice**: El concentrado de jugo de naranja congelado se empaca en latas de cartón de 6 onzas. Estas latas se forman en una máquina haciéndolas girar desde cartulina y uniendo un panel inferior de metal. Una lata es entonces inspeccionado para determinar si, cuando se llena, el líquido podría filtrarse ya sea en la costura lateral o alrededor de la articulación inferior. Si esto ocurre, una lata se considera no conforme. Los datos fueron recolectados como 30 muestras de 50 latas cada una a intervalos de media hora durante un período de tres turnos en el que la máquina estaba en operación continua. Se implementa en el paquete qcc ((Scrucca, 2004)).
- **pcmanufact**: Un fabricante de computadora personal cuenta el número de no conformidades por unidad en la línea de ensamblaje final. Recopila datos en 20 muestras de 5 computadoras cada una. Se implementa en el paquete qcc ((Scrucca, 2004)).
- **pistonrings**: Los anillos de pistón para un motor de automóvil se producen mediante un proceso de forjado. El diámetro interior de los anillos fabricados

por el proceso se mide en 25 muestras, cada una de ellas de tamaño 5, extraídas de un proceso que se considera “en control”. Se implementa en el paquete qcc ((Scrucca, 2004)).

Name	Description
counters	Una compañía de agua de Coruña quiere controlar el rendimiento de los mostradores instalados en toda la ciudad. Teniendo en cuenta que hay dos marcas o proveedores de contadores.
employment	Una empresa hotelera hispano-argentina quiere controlar el nivel de empleo en sus establecimientos. Para esto, hará un control continuo que mide la cantidad de ocupantes en términos de porcentaje.
oxidation	Contiene información sobre el nivel de pureza de cada lote de variables Picual. Luego tenemos el tipo de aceite midiendo la temperatura de inicio de oxidación. Tenemos 50 submuestras de aceite con su temperatura para oxidar.
plates	Una empresa química quiere medir el grado de homogeneidad del producto, se tomaron 50 muestras, se realizaron 5 mediciones por placa correspondientes a diferentes áreas de dureza Vickers de piedra artificial.
precision	Un astillero de fabricación de embarcaciones recreativas, destinado a optimizar y controlar las propiedades mecánicas de los modelos de yates de casco. Esto ha llevado a cabo un estudio en el que se utilizó el módulo de elasticidad de la resistencia a la tracción de la resina epoxídica (polímero), después de aplicar diferentes presiones de curado medidas: 0,1 y 10 MPa. Se toman 60 submuestras compuestas por tres medidas tomadas el mismo día.

Tabla A.3: Conjuntos de datos incluidos en el paquete

A.4. Una alternativa: gráficos de control no paramétrico basados en la profundidad de los datos

Los gráficos de control presentadas en esta sección fueron propuestas por Regina Liu ((Liu, 1995)). La idea principal detrás de sus gráficos de control es reducir cada

medida multivariada a un índice univariado, es decir, su clasificación relativa centro-exterior inducida por una profundidad de datos. Estos gráficos permiten la detección simultánea del cambio de ubicación (desplazamiento de la media) y el aumento de escala (cambio en la variabilidad) en un proceso.

Liu propuso y justificó tres tipos de gráficos de control, los gráficos r , Q y S que se pueden ver como generalizaciones multivariadas basadas en profundidad de datos de X , \bar{x} univariadas y gráficos CUSUM respectivamente.

A.4.1. Profundidad de datos

En el análisis multivariado, el término profundidad se refiere al grado de centralidad de un punto relativo a una nube de datos o una distribución de probabilidad. Por lo tanto, es posible definir un orden en el espacio euclidiano multidimensional a través de la profundidad de un punto.

Según (Dyckerhoff, 2004) y expuesta por (Casco *et al.*, 2011), la función de profundidad se puede definir como:

Si una distribución P se establece en R^d , una función de profundidad es una función acotada $D_p : R^d \rightarrow R$ que asigna cada punto de R^d su grado de centralidad con respecto a P . Existen varias funciones de profundidad con las que se pueden realizar gráficos de control.

- Profundidad Simplicial ((Liu, 1990)).
- Profundidad de Mahalanobis (Basada en la distancia de Mahalanobis (Mahalanobis, 1936)).
- Profundidad de Halfspace o Tukey ((Tukey, 1975)).
- Profundidad de verosimilitud ((Fraiman *et al.*, 1997)).
- Profundidad de proyección aleatoria ((Zuo y Serfling, 2000)).

A.4.2. Estadísticos derivados de la profundidad de datos

Sea G una distribución k -dimensional, y sean Y_1, \dots, Y_m m observaciones aleatorias de G . La muestra Y_1, \dots, Y_m es generalmente referida como muestra de referencia en el contexto de control de calidad, y consideradas como las medidas de productos

desarrollados en un proceso en control. Sean $X_1, X_2 \dots$ nuevas observaciones de un proceso de manufactura. Supongamos que los X_i siguen una distribución F si la calidad del producto se ha deteriorado o si el proceso está fuera de control, de otra manera siguen una distribución G .

Sea $D_G(\cdot)$ una noción de profundidad y supongamos que G y F son dos distribuciones absolutamente continuas. Si ordenamos los $D_{G_m}(Y_i)$ ascendentemente y usamos $Y_{[j]}$ para denotar el punto muestral asociado al j -ésimo valor de profundidad más pequeño, entonces $Y_{[1]}, \dots, Y_{[m]}$ son los estadísticos de orden de los Y_i siendo $Y_{[m]}$ el punto más central. Cuanto más pequeño sea el orden (o el rango) de un punto, más alejado estará ese punto con respecto al centro de la distribución $G(\cdot)$.

(Liu, 1995) define el estadístico rango como:

$$r_G(y) = P\{D_G(Y) \leq D_G(y) | Y \sim G\}$$

donde $Y \sim G$ indica que Y sigue una distribución G . Y para el caso donde G es desconocido y solo la muestra $\{Y_1, \dots, Y_m\}$ es conocida como

$$r_{G_m}(y) = \frac{\#\{D_{G_m}(Y_j) \leq D_{G_m}(y), j = 1, \dots, m\}}{m}$$

A.4.3. El gráfico r

Calcule $\{r_G(X_1), r_G(X_2), \dots, r_G(X_n)\}$ or $\{r_{G_m}(X_1), r_{G_m}(X_2), \dots, r_{G_m}(X_n)\}$ si solo Y_1, \dots, Y_m están disponibles pero no G . El gráfico r consiste en trazar la estadística de rango en función del tiempo. Cuyo límite de control central CL = 0.5 y el límite inferior es LCL = α , donde α es la tasa de falsa alarma. El proceso se declara fuera de control si $r_G(\cdot)$ cae bajo LCL. Un pequeño valor de la estadística de rango $r_{G_m}(X)$ significa que solo una proporción muy pequeña de Y_i es más remota que X . Por lo tanto, suponiendo que $X \sim F$, entonces un valor pequeño $r_{G_m}(X)$ sugiere una posible desviación de G a F . Esto puede deberse a un desplazamiento en la ubicación y/o un aumento en la escala de la variable de calidad. Aunque el gráfico r no tiene la línea UCL para hacer que CL sea la línea central de la región bajo control, la línea CL sirve como un punto de referencia para ver si un patrón o tendencia está surgiendo o se está desarrollando en una secuencia de muestras. Si

$r_{Gm}(X)$ es mayor que 0.5, hay evidencia de una disminución en la escala con quizás un cambio de ubicación insignificante. Esto se ve como una mejora en la calidad dada una ganancia en la precisión, y por lo tanto el proceso no se vería como fuera de control.

Apéndice B

El paquete ILS para realizar estudios intelaboratorios

En este trabajo se presenta una librería de R para realizar Estudios Interlaboratorio (ILS). Este paquete permite detectar aquellos laboratorios que proporcionan resultados no consistentes al compararlos con otros. Además permite trabajar simultáneamente con diversos materiales de prueba, desde las perspectivas del análisis de datos univariante y funcional (FDA).

Las técnicas implementadas en el paquete ILS permiten realizar Estudios Interlaboratorios con el enfoque univariante, basado en las normas ASTM E691 e ISO 5725, que consiste en estimar los estadísticos h y k de Mandel para identificar aquellos laboratorios que proporcionan resultados significativamente diferentes, así también, permite probar la presencia de valores atípicos mediante las pruebas de Cochran y Grubbs. Además, proporciona técnicas de Análisis de Varianza (ANOVA), incluyendo las pruebas F y de Tukey para probar las diferencias entre las medias de la variable de prueba correspondientes a diferentes laboratorios y a cada material ensayado.

Una de las novedades de esta librería es la de incorporar herramientas para realizar un ILS a partir de datos funcionales. Por tanto, este paquete permite tener en cuenta la naturaleza funcional de los datos obtenidos por técnicas experimentales correspondientes a la química analítica, física aplicada e ingeniería (espectros, termogramas, etc.).

El paquete ILS, permite estimar los estadísticos funcionales, $H(t)$ y $K(t)$, así

como los estadísticos de prueba d_H y d_K utilizados para contrastar las hipótesis de repetibilidad y reproducibilidad a partir de los valores críticos c_h y c_k estimados mediante el algoritmo bootstrap descrito en (Flores *et al.*, 2017).

B.1. Introducción

Un Estudio Interlaboratorio, se puede definir como procedimientos de control para evaluar el desempeño mediante pruebas colaborativas (ASTM-E691, 2004; ISO-5725, 1994). En un Estudio Interlaboratorio se elige un número adecuado de laboratorios para participar en el experimento, con el objetivo de que éstos analicen las muestras que se les proporcionan y así obtener resultados.

Los laboratorios participantes reciben muestras (homogeneizadas o que los laboratorios deberán homogeneizar) para su análisis, para posteriormente estudiar los resultados conjuntos e individuales de dichos análisis y valorar la consistencia de los resultados obtenidos. Algunos de los factores más comunes causantes de variabilidad son el equipamiento de los laboratorios, los operadores, los materiales, la temperatura ó la humedad, entre otros.

Actualmente se utilizan diversas técnicas estadísticas univariantes para estudiar la consistencia de los resultados de las pruebas procedentes de varios laboratorios. La norma ASTM E691 (normas para estudios de precisión de un método de medición) recomienda aplicar sólo una técnica gráfica a partir de los estadísticos k y h de Mandel (ASTM-E691, 2004), mientras que la norma ISO 5725 (práctica estándar para conducir un Estudio Interlaboratorio para determinar la precisión de un método analítico) recomienda además de la técnica gráfica utilizar las pruebas de Cochran y Grubbs (ISO-5725, 1994).

Adicionalmente, mediante un análisis de varianza (ANOVA), se puede aplicar una prueba F donde el único factor es el laboratorio, para contrastar la hipótesis de igual de medias entre los resultados de los laboratorios. Esto permite estimar la varianza de la repetibilidad y reproducibilidad. Por otro lado, la prueba de Tukey es utilizada para probar las diferencias entre las medias de la variable de prueba correspondientes a diferentes laboratorios y a cada material ensayado.

Para realizar las pruebas de consistencia sobre las hipótesis de repetibilidad y

reproducibilidad, así como la detección de resultados atípicos se debe comparar los valores de los estadísticos con sus correspondientes valores críticos. Si estos son mayores se detecta inconsistencia en los resultados de los laboratorios. La norma ISO 5725-2 proporciona algunos valores críticos dependiendo del número de laboratorios L , del número de mediciones n y nivel de significación α .

En la actualidad tanto las normas ISO 5725 como ASTM E691 no contemplan metodologías para realizar ILS cuando los resultados tienen un carácter funcional, es decir las respuestas de las pruebas son curvas (datos funcionales). El Análisis de Datos Funcionales (FDA, por sus siglas en inglés) es una rama relativamente nueva de la estadística que tiene como unidad de análisis curvas, superficies, volúmenes definidos en un continuo tal como el tiempo o dominio de frecuencia.

El propósito del paquete ILS es facilitar el uso de nuevas herramientas en el contexto FDA. A partir de las propuestas en (Naya *et al.*, 2014) y (Flores *et al.*, 2017) de nuevas técnicas en el campo del Análisis de Datos Funcionales para realizar Estudios Interlaboratorios, se implementan en el paquete ILS funciones para aplicar el método gráfico y analítico pero con un enfoque FDA para los estadísticos h y k de Mandel. Así, como estadísticos de contraste que facilitan la realización de estudios de Repetibilidad y Reproducibilidad cuando los datos tienen un carácter funcional. Adicionalmente, el paquete ILS permite aplicar los métodos sugeridos por las normas para el caso escalar. El paquete ILS está disponible en la Red de Archivos Comprehensive R en <http://CRAN.R-project.org/package=ILS>.

En R, Para el caso escalar, se cuenta con el paquete metRology que permite realizar estimación de incertidumbre de la medición, así como estudios interlaboratorio (Ellison, 2017). Mientras que el paquete multcomp permite realizar análisis de la varianza (ANOVA) mediante los contrastes de la F y Tuckey (Hothorn *et al.*, 2017). En el caso funcional, R cuenta con el paquete fda.usc, en este paquete se pueden aplicar técnicas de detección de datos funcionales, así como aplicar un ANOVA para datos funcionales (Febrero-Bande *et al.*, 2007). El paquete ILS a través del multcomp y fda.usc utiliza los métodos mencionados.

El resto del apéndice se organiza de la siguiente manera. En la sección B.2, se describen dos ejemplos de estudios interlaboratorio en los que se obtuvieron cuatro conjuntos de datos experimentales y posteriormente se presenta el uso del paquete

para resumir dos de estos conjuntos. En la sección B.3, se ilustra la funcionalidad del paquete a través de un procedimiento estándar ILS, utilizando el conjunto de datos Glucosa y en la sección B.4 se emplea los conjunto de datos TG y DSC en el caso funcional.

B.2. Ejemplos de Estudios Interlaboratorio

Un Estudio Interlaboratorio permite evaluar métodos analíticos realizados por los laboratorios, ya sea para la verificación de laboratorios, de un procedimiento experimental concreto o una norma, entre otras aplicaciones. Por ejemplo, el paquete ILS, contiene el conjunto de datos Glucose que corresponden a resultados de un análisis clínico utilizado en la norma ASTM E 691 (ASTM-E691, 2004) para explicar la aplicación de pruebas de consistencia. Así también, a partir de un estudio de las propiedades del material oxalato de calcio, se obtienen tres conjuntos de datos (IDT, TG, DSC) que se incorporan en el paquete. Estos últimos conjuntos de datos, han sido descritos ampliamente en el artículo de (Naya *et al.*, 2014).

B.2.1. Estudio clínico de medida de glucosa en sangre

El conjunto de datos Glucose corresponde a la prueba de glucosa en suero (mediciones de la concentración de glucosa en la sangre utilizada para controlar la diabetes). En el estudio participan ocho laboratorios donde se realizaron cinco pruebas diferentes de muestras de sangre etiquetadas con diferentes referencias, que van desde un contenido de azúcar bajo a uno muy alto. Se obtuvieron tres réplicas para cada muestra.

Cada uno de estos laboratorios mide 5 niveles de concentración diferentes de un determinado material (niveles A, B, C, D y E) y en cada uno de esos niveles se toman tres medidas (3 réplicas). Cada laboratorio proporciona entonces un total de 15 mediciones (3 para cada nivel) teniendo por tanto 120 mediciones.

Para poder acceder a este conjunto de datos, primero hay que cargar el paquete ILS y posteriormente el objeto data frame Glucose, mediante las siguientes instrucciones.

```
R> library(ILS)
```

```
R> data("Glucose", package = "ILS")
```

El primer paso para realizar un análisis con el paquete ILS, es utilizar la función `lab.qcdata` (quality control data) que recibe como argumento un data frame. Donde, la primera columna debe contener la variable respuesta, posteriormente debe contener, el índice de repetición por cada laboratorio, el índice del material con el que se realizó la prueba y finalmente en la cuarta columna, el índice del laboratorio donde se ejecutó el procedimiento.

Se construye el objeto `qcdata` de la clase `lab.qcdata`, posteriormente se resume el conjunto de datos mediante las funciones `summary` y `summary`. En la Figura B.1 se representan los resultados de todos los laboratorios.

```
R> qcdata <- lab.qcdata(Glucose)
```

```
R> summary(qcdata)
```

	x	replicate	material	laboratory
Min.	: 39.02	1:40	A:24	Lab1 :15
1st Qu.:	78.45	2:40	B:24	Lab2 :15
Median	:135.03	3:40	C:24	Lab3 :15
Mean	:149.09		D:24	Lab4 :15
3rd Qu.:	196.66		E:24	Lab5 :15
Max.	:309.40			Lab6 :15

```
R> plot(qcdata, ylab = "Laboratory", xlab = "Glucose concentration in blood")
```

En la Figura B.1, se puede notar que el nivel de azúcar en la sangre aumenta del material A al D y existe más variabilidad entre los resultados por cada laboratorio a partir del material C.

Para realizar el cálculo de los estadísticos para los métodos gráficos y analíticos para el caso escalar se debe utilizar primero la función `lab.qcs` (quality control statistics). Esta función retorna la estimación de las medidas estadísticas (media,

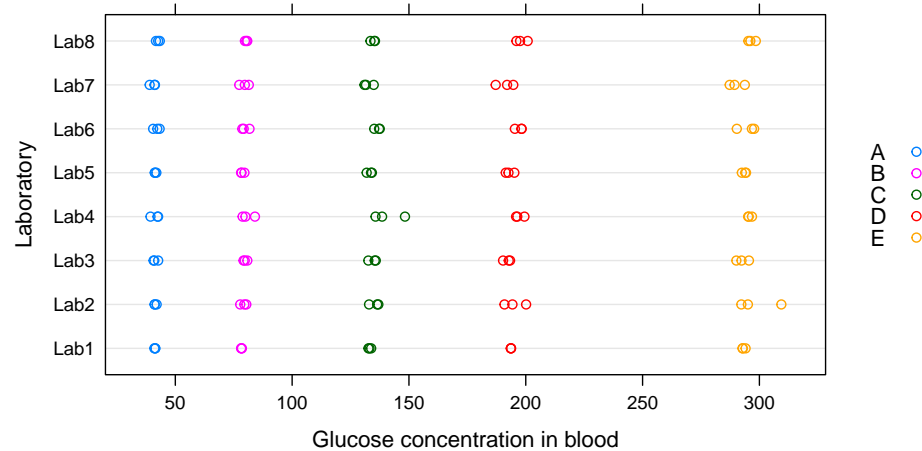


Figura B.1: Mediciones de la concentración de glucosa en sangre utilizadas para controlar la diabetes.

varianza, etc.) necesarias para la estimación de los estadísticos h y k , así como las medidas necesarias para realizar las pruebas de Cochran y Grubbs.

A partir del objeto `qcdata` se utiliza la función `lab.qcs` para crear el objeto `qcstat` que estima simultáneamente tanto la media y desviación global (considerando los resultados de todos los laboratorios), así como la media y desviación para cada laboratorio por tipo de material. Además de la desviación de repetibilidad S_r , S_B la desviación entre las medias de los laboratorios y la desviación de la reproducibilidad S_R para cada material. Para mayor información de la definición y forma de cálculo se puede consultar en (ASTM-E691, 2004; ISO-5725, 1994).

```
R> qcstat <- lab.qcs(qcdata)
R> summary(qcstat)
```

Number of laboratories: 8

Number of materials: 5

Number of replicate: 3

Summary for Laboratory (means):

	Lab1	Lab2	Lab3	Lab4	Lab5	Lab6
A	41.28333	41.44000	41.45000	41.45667	41.46333	42.02000
B	78.31667	79.23333	79.90333	80.96333	78.69000	79.89333
C	133.19667	135.40667	134.59000	140.83000	133.26667	136.61667
D	193.65000	195.10667	192.09000	197.21333	193.05000	197.24333
E	293.25333	298.91667	292.67000	295.82000	293.56333	294.95667

	Lab7	Lab8
A	40.45667	42.57667
B	79.51667	80.34667
C	132.49333	134.71000
D	191.26000	198.12333
E	290.13667	296.62000

Summary for Laboratory (Deviations):

	Lab1	Lab2	Lab3	Lab4	Lab5	Lab6	Lab7
A	0.2230097	0.4850773	1.0608016	1.8117763	0.3666515	1.408119	1.247811
B	0.1582193	1.3268509	0.8303212	2.7660863	0.7754354	1.636592	2.059935
C	0.5909597	2.1679791	1.7287857	6.6200227	1.1987215	1.287025	2.124296
D	0.0600000	4.6824068	1.5932043	1.9365519	1.8826311	1.649616	3.817709
E	0.7266590	9.1869055	2.7101107	0.8835723	0.9543759	4.034282	3.304184

	Lab8
A	0.8224557
B	0.5064912
C	1.0343597
D	2.4637844
E	1.6479078

Summary for Material:

mean	S	S_r	S_B	S_R
------	---	-----	-----	-----

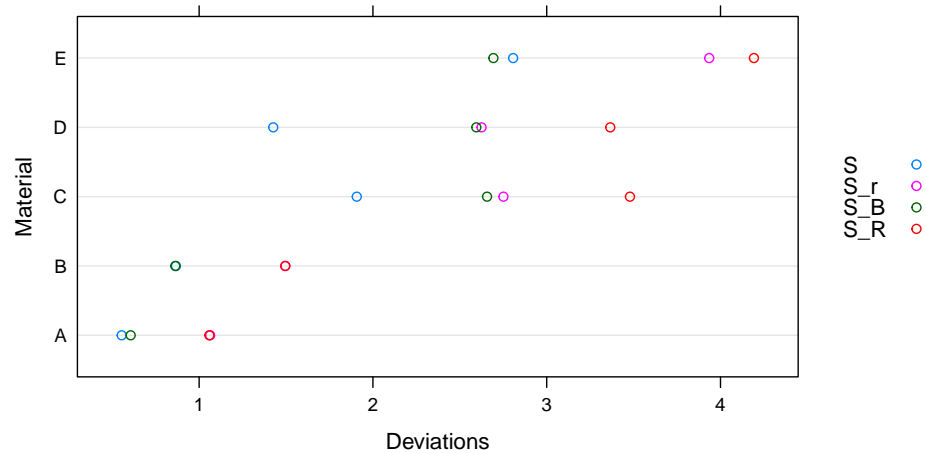


Figura B.2: Las siguientes medidas de variabilidad se calculan para cada material a partir de las medidas de todos los laboratorios. S es la desviación global de todos los laboratorios, S_r es la desviación de la repetibilidad, S_B la desviación entre los medios de los laboratorios y S_R la desviación, en función de la reproducibilidad.

```
A 41.51833 0.5543251 1.063224 0.6061274 1.058783
B 79.60792 0.8664835 1.496071 0.8627346 1.495481
C 135.13875 1.9071053 2.750879 2.6566872 3.478919
D 194.71708 1.4262962 2.625065 2.5950046 3.365713
E 294.49208 2.8067799 3.934974 2.6931364 4.192334
```

```
R> plot(qcstat, xlab = "Deviations", ylab = "Material")
```

En el gráfico de la Figura B.2, se presentan los valores de S, S_r, S_R y S_B para cada material. Se observa mayor presencia de variabilidad a partir del material C. Los materiales C y D tienen mayor variabilidad entre los resultados de los laboratorios (S_R) y dentro de ellos (S_r).

B.2.2. Caracterización de materiales mediante análisis termogravimétrico

En un estudio realizado en (Naya *et al.*, 2014), se analizaron 105 muestras de oxalato de calcio mediante técnicas termogravimétricas (TG), obteniéndose 105 curvas TG que muestran la pérdida de masa de oxalato en función de la temperatura cuando se calientan las muestras de oxalato a $20^{\circ}\text{C}/\text{min}$. Además se analizaron 90 muestras de oxalato de calcio mediante técnica térmica calorimétrica diferencial de barrido (DSC), obteniéndose 90 curvas DSC que determinan a partir de un dispositivo SDT la diferencia de energía entre una referencia y la muestra de oxalato. Esto permite observar, el intercambio de energía entre muestra y referencia en función de la temperatura. La temperatura varía linealmente con el tiempo con pendiente de $20^{\circ}\text{C}/\text{min}$. A partir de los resultados se generaron dos conjuntos de datos TG obtenidos de 7 diferentes laboratorios y DSC de 6 diferentes laboratorios. En cada laboratorio se analizaron 15 curvas evaluadas en 1000 observaciones.

Adicionalmente, a partir de las curvas TG se obtuvo un tercer conjunto de datos IDT (temperatura de descomposición inicial), es un parámetro definido por la temperatura a la que se produce una pérdida de material del 5% de su peso cuando se calienta usando una velocidad constante, que se compone de los valores de IDT del oxalato de calcio obtenido de 7 diferentes laboratorios que analizan 15 muestras de oxalato cada uno. Este último conjunto de datos es un ejemplo de datos escalares que se obtiene considerando sólo una característica de la curva TG. Al no considerar toda la información de la curva se puede llegar a conclusiones equivocadas durante un ILS.

Los laboratorios 1,6, y 7 presenten resultados no consistentes. Para el laboratorio 1 se utilizó un analizador térmico simultáneo (STA) con un programa de calibración desfasado. En el laboratorio 6 se utiliza un analizador SDT simultáneo con una calibración antigua y finalmente en el laboratorio 7 se utiliza un analizador simultáneo SDT con un sesgo en temperatura con respecto a los valores reales (2°C desplazada con respeto al punto de fusión del zinc).

Para la estimación de estadísticos funcionales que son utilizados para aplicar los métodos gráficos y analíticos se procede de igual forma que en el caso escalar. El paquete ILS consta de las clase `ils.fqcdata` (functional quality control data) que

construye un objeto que tiene la estructura de un data.frame, cada fila representa una el resultado de una prueba. El tamaño del data frame es $n \times l$, donde n es el número de réplicas en cada laboratorio y l el número de laboratorios que participan en el estudio. Se implementaron funciones específicas para realizar gráficos y resúmenes de este tipo de objetos.

Posteriormente, se utiliza la función `ils.fqcs` (functional quality control statistical) que calculan las medidas estadísticas media, varianza, etc. funcionales necesarias para la estimación de los estadísticos $H(t)$ y $K(t)$ y los estadísticos de contraste d_H y d_K .

Para construir un objeto de la clase `ils.fqcdata`, primeramente se debe definir la rejilla donde se han obtenido las observaciones, que en este caso son 1000 puntos que representan temperaturas que van de 40°C a 850°C. En la Figura B.3, se presenta las curvas TG. A partir del objeto `fqcdata` se contruye el objeto `fqcstat`.

```
R> data(TG, package = "ILS")
R> delta <- seq(from = 40 ,to = 850 ,length.out = 1000 )
R> fqcdata <- ils.fqcdata(TG, p = 7, argvals = delta)
R> main <- "TG curves obtained from calcium oxalate"
R> xlab <- "Temperature (C)"
R> ylab <- "Mass (%)"
R> plot(x = fqcdata, main = main, xlab = xlab , ylab = ylab,
+       legend = TRUE,x.co = 20, y.co = 90)
```

```
R> fqcstat <- ils.fqcs(fqcdata)
R> summary(fqcstat)
```

```
Number of laboratories: 7
Number of replicates: 15
```

```
R> xlab <- "Temperature (C)"
R> ylab <- "Mass (%)"
```

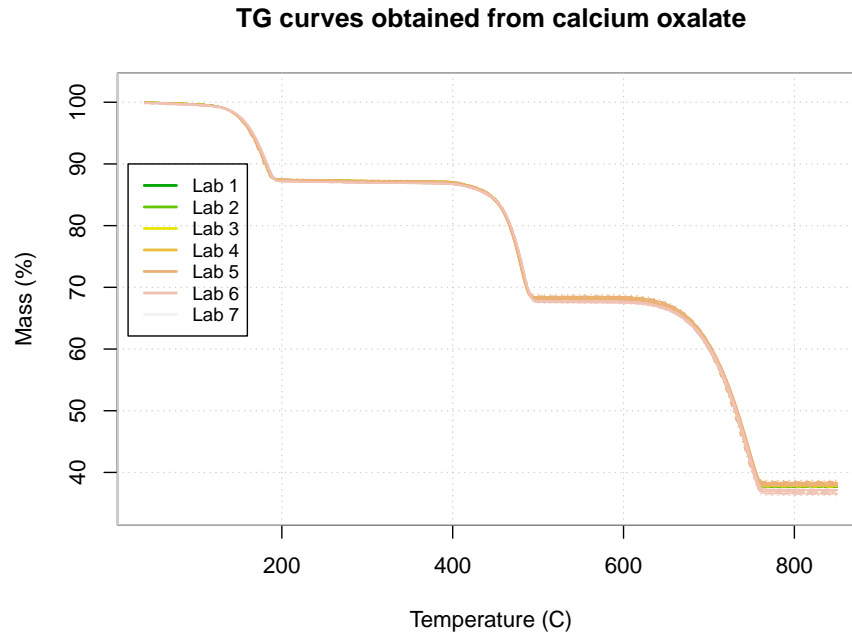


Figura B.3: Curvas de TG obtenidas de oxalato de calcio.

```
R> plot(fqcstat, xlab = xlab, ylab = ylab)
```

La función `plot` permite realizar un panel en el que están cuatro gráficos, en la primera fila se tiene las medias y varianzas funcionales por cada laboratorio, mientras que en la segunda fila se tienen la media y varianza funcional global. En la Figura B.4, se presentan las diferentes medias y varianzas funcionales para cada laboratorio así como la media y la varianza global para el conjunto de datos TG.

B.3. Estudios Interlaboratorio: Enfoque estándar

El paquete ILS consta de dos grupos de funciones (Tabla B.1) para aplicar los enfoques escalar y funcional de detección de laboratorios atípicos en un ILS. Existen procedimientos gráficos y analíticos (contrastos) para este fin.

Como ya se comentó, entre las metodologías utilizadas para evaluar la consistencia de los resultados de los laboratorios, se tienen los estudios R & R, que permiten

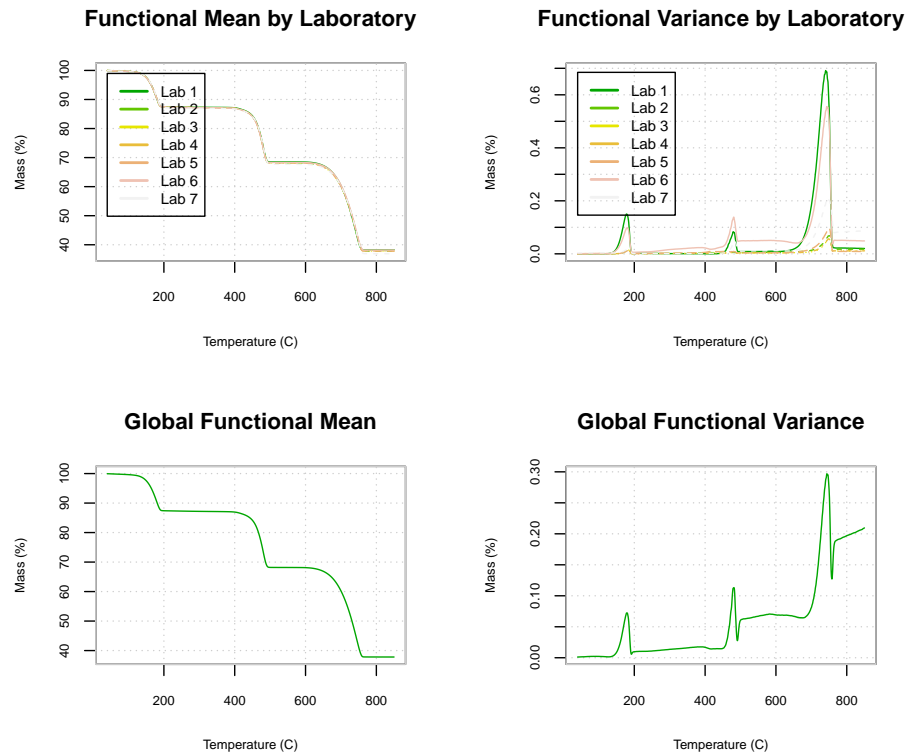


Figura B.4: Estadística descriptiva funcional: medias y varianzas funcionales para cada laboratorio, así como la media y la varianza global de laboratorio de las curvas de TG obtenidas a partir de oxalato de calcio.

Enfoque	Técnica	Función	Descripción
Escalar	Gráfica	h.qcs, k.qcs	Estadísticos h y k de Mandel
	Test	test.cochran,	Pruebas de Cochran
		test.grubbs	Pruebas de Grubbs
		lab.aov	ANOVA
Funcional	Gráfica	h.fqcs, k.fqcs	Estadísticos funcionales $H(x)$ y $K(x)$ de Mandel
	Test	mandel.fqcs	Test funcional de Mandel

Tabla B.1: Funciones incorporadas en el paquete ILS para realizar Estudios Interlaboratorio con un enfoque escalar y funcional.

cuantificar la variabilidad entre laboratorios (reproducibilidad) y la variabilidad entre resultados (reproducibilidad). La repetibilidad es la variabilidad entre los resultados

de las pruebas independientes obtenidas por cada laboratorio individual, aquí se evalúa la variación que se produce por el sistema de medida, ya que las pruebas las obtiene un único operario en cada laboratorio y sin que transcurra demasiado tiempo entre varias mediciones. Por otra parte, la reproducibilidad se refiere a la variabilidad entre los resultados de las pruebas individuales obtenidas en los diferentes laboratorios. Con esto, se puede determinar el sesgo que se produce debido a los operarios o trabajadores del laboratorio en cuestión.

La regla decisión para detectar un laboratorio como inconsistente es comparar el valor del estadístico h o k si es mayor que el valor crítico calculado con un nivel de significación generalmente de 0.5 %, que es el recomendado por la norma ASTM E-691.

Por otro lado, el paquete ILS permite realizar la prueba de Cochran para examinar la consistencia intralaboratorio y el prueba de Grubbs para examinar la consistencia interlaboratorio. La prueba de Grubbs también se puede usar como prueba de consistencia de los resultados medidos en un laboratorio usando materiales idénticos. Estas pruebas son recomendados por la norma ISO 5725-2.

Empleando el paquete ILS se puede realizar un ANOVA y pruebas de comparación de medias. Sin embargo, antes de aplicar esta técnica, primero se deberían excluir del ILS los laboratorios que presenten resultados no consistentes, para esto se deben realizar pruebas de consistencia e identificación de resultados atípicos.

Existen dos posibles tipos de escenarios a evaluar para determinar la presencia de valores atípicos, el primero es que los resultados de algún laboratorio se desvíe de los demás en precisión, lo que significa que ofrecen un trabajo menor o mayor que los demás o la media de los resultados. El segundo, tiene que ver con la identificación de resultados atípicos en un laboratorio es decir en un nivel dado. A continuación se describen los estadísticos y las pruebas recomendadas por la norma ISO 5725 y ASTM E-691.

B.3.1. ILS: Estudio de glucosa

En esta sección, se va a utilizar los objetos `qcdata` y `qcstat` de las clases `lab.qcdata` y `lab.qcs` creados en la subsección B.2.1 a partir del conjunto de datos `Glucose`.

En primera instancia, se realizará un análisis de la variabilidad dentro de cada

laboratorio. Para este fin, se utiliza el estadístico k (k.qcs) y la prueba de Cochran (cochran.test) que permite identificar si hay algún laboratorio con resultados no consistentes. Posteriormente, se utiliza el estadístico h (h.qcs) y la prueba de Grubbs (grubbs.test) para efectuar un análisis para evaluar la variabilidad entre laboratorios.

Las siguientes sentencias, permiten crear un objeto k.qcs y realizar el gráfico correspondiente de los estadísticos k para cada laboratorio y material (ver Figura B.5).

```
R> k <- k.qcs(qcdata, alpha = 0.005)
R> plot(k)
R> summary(k)
```

```
Number of laboratories: 8
Number of materials: 5
Number of replicate: 3
Critical value: 2.06084
Beyond limits of control:
```

	A	B	C	D	E
Lab1	TRUE	TRUE	TRUE	TRUE	TRUE
Lab2	TRUE	TRUE	TRUE	TRUE	FALSE
Lab3	TRUE	TRUE	TRUE	TRUE	TRUE
Lab4	TRUE	TRUE	FALSE	TRUE	TRUE
Lab5	TRUE	TRUE	TRUE	TRUE	TRUE
Lab6	TRUE	TRUE	TRUE	TRUE	TRUE
Lab7	TRUE	TRUE	TRUE	TRUE	TRUE
Lab8	TRUE	TRUE	TRUE	TRUE	TRUE

```
R> cochran.test(qcdata)
```

```
Test Cochran
```

```
Critical value: 0.5157
```

Alpha test: 0.00625

	Smax	Material	C	p.value
1	Lab4	A	0.20033869	1
2	Lab4	B	0.15447962	1
3	Lab4	C	0.10935197	1
4	Lab2	D	0.08493741	1
5	Lab2	E	0.07416440	1

En la Figura B.5, la línea discontinua es el valor crítico obtenido a un nivel de significación de 0.005. A partir de esto, los materiales 5 del laboratorios 2 y el material 3 del laboratorio 4 presentan resultados inconsistentes dado que los correspondientes valores de los estadísticos k son mayores que el valor crítico calculado a partir de $L = 8, n = 15$ y $\alpha = 0.005$.

La función `k.qcs` calcula los siguientes objetos:

- k : El estadístico k para cada laboratorio y por material.
- `k.critical`: El valor crítico para el α definido en la función `k.qcs`.
- `violations`: Matriz de dimensión $L \times R$, número de laboratorios por número de materiales.

En la matriz `violations` se almacenan valores lógicos resultado de las comparaciones entre los valores críticos y los valores del estadístico k . Si esta comparación es FALSE, el laboratorio es consistente, es decir el valor crítico es menor que el valor del estadístico. Para este ejemplo el valor crítico es 2.06.

La función `cocharn.test` permite realizar la prueba de Cochran, ningún laboratorio con la varianza máxima para cada material es considerado inconsistente. Ya que el valor crítico es 0.52 y los p valores en cada material no supera el 5% de significancia.

Se procede a utilizar las funciones `h.qcs` y `plot(h)` para estimar y graficar respectivamente los estadísticos h para cada laboratorio y material. Posteriormente, se aplica la prueba de Grubbs. El valor crítico es 2.15, a partir de este resultado se observa en la Figura 1.3 que los laboratorios 4,7 y 8 presentan resultados no consistentes a un nivel de significancia $\alpha = 0.005$. Así también, al utilizar la prueba de

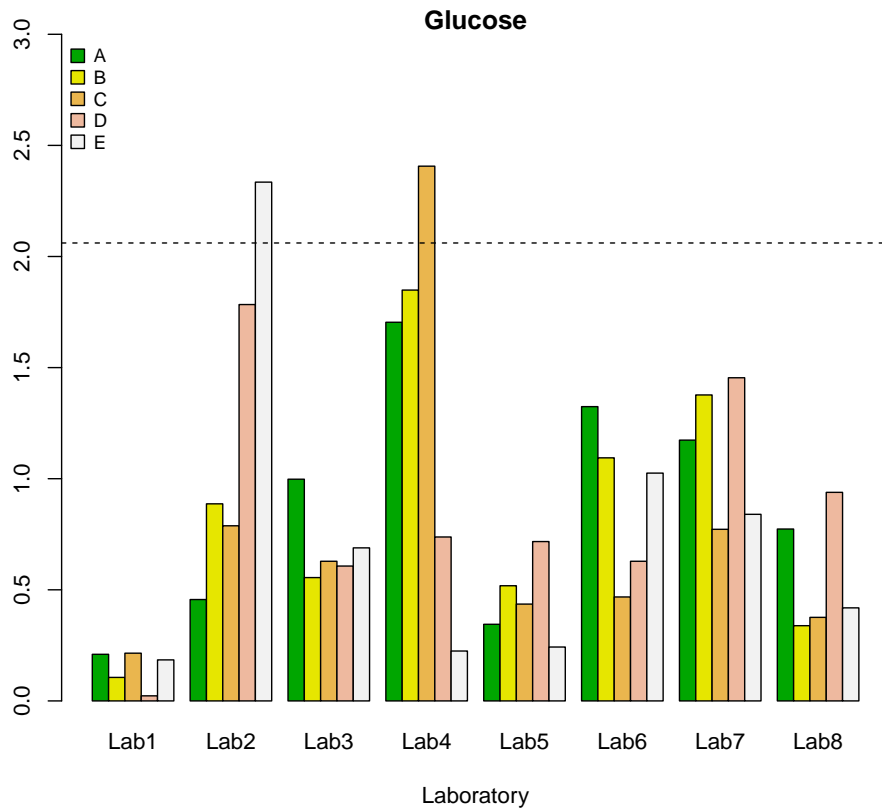


Figura B.5: Estadísticos k por laboratorio y material correspondientes a las medidas de glucosa en sangre disponibles en la base de datos Glucosa.

Grubbs se detectan laboratorios con resultados muy extremos, es decir muy grandes y muy pequeños.

```
R> h <- h.qcs(qcdata, alpha = 0.005)
R> plot(h)
R> summary(h)
```

```
Number of laboratories: 8
Number of materials: 5
Number of replicate: 3
Critical value: 2.152492
```

Beyond limits of control:

	A	B	C	D	E
Lab1	TRUE	TRUE	TRUE	TRUE	TRUE
Lab2	TRUE	TRUE	TRUE	TRUE	TRUE
Lab3	TRUE	TRUE	TRUE	TRUE	TRUE
Lab4	TRUE	TRUE	FALSE	TRUE	TRUE
Lab5	TRUE	TRUE	TRUE	TRUE	TRUE
Lab6	TRUE	TRUE	TRUE	TRUE	TRUE
Lab7	TRUE	TRUE	TRUE	FALSE	TRUE
Lab8	TRUE	TRUE	TRUE	FALSE	TRUE

```
R> grubbs.test(qcdata)
```

Test Grubbs

Critical value: 2.032

Alpha test: 0.00625

	Material	Gmax	G.max	p.value.max	Gmin	G.min	p.value.min
1	A	Lab8	1.909	9.986e-02	Lab7	1.915	9.690e-02
2	B	Lab4	1.564	3.708e-01	Lab1	1.490	4.568e-01
3	C	Lab4	2.984	2.220e-16	Lab7	1.387	5.937e-01
4	D	Lab8	2.388	4.186e-04	Lab7	2.424	8.644e-05
5	E	Lab2	1.576	3.577e-01	Lab7	1.552	3.846e-01

B.4. Estudios Interlaboratorio: Nuevo enfoque FDA

B.4.1. ILS: Estudio Termogravimétrico

Se aplicán las técnicas desarrolladas en (Flores *et al.*, 2017) e implementadas en el paquete ILS para comprobar si se detectan laboratorios inconsistente ya sea por resultados atípicos en la variabilidad intralaboratorios o interlaboratorio. Como se

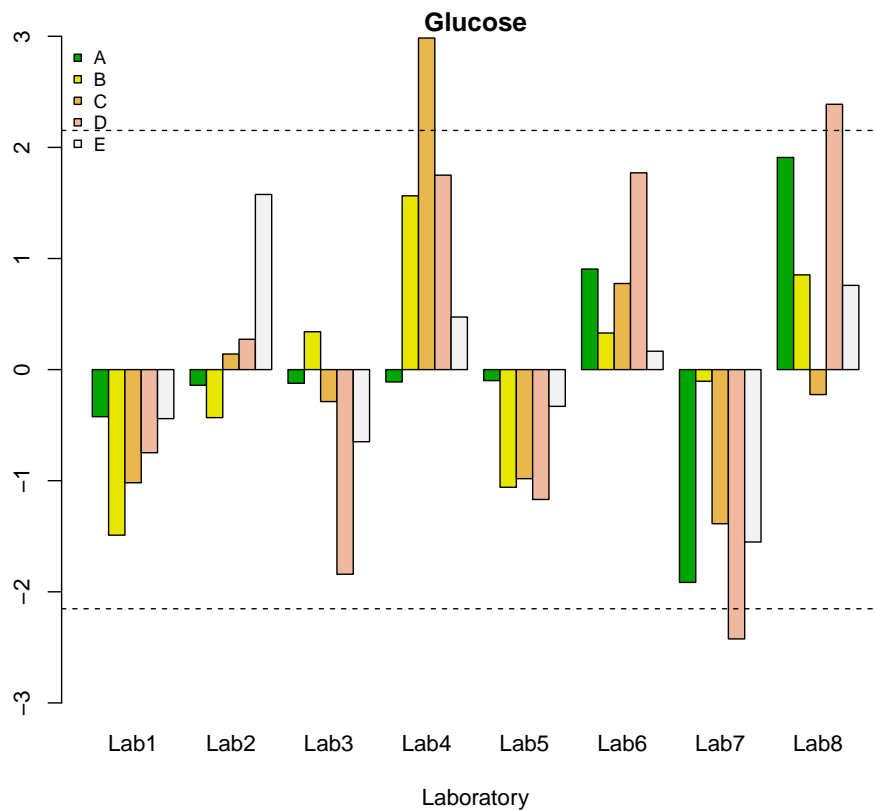


Figura B.6: Estadísticos h por laboratorio y material.

había mencionado los laboratorios, 1, 5 y 6 difieren sus resultados de los demás laboratorios. Se utilizan los conjuntos de datos descritos en la subsección B.2.2, el conjunto de datos TG, contiene resultados de pruebas termogravimétricas de 7 laboratorios, mientras que el conjunto de datos DSC, contiene resultados de 6 laboratorios (se excluye el laboratorio 1).

El primer paso consiste en estimar los estadísticos funcionales $H(t)$ y $K(t)$, mediante la función `mandel.fqcs`, posteriormente se realizan sus respectivas gráficas en el espacio funcional definido. En la Figura B.7 se presentan tanto los estadísticos $K(t)$ y $H(t)$ para cada laboratorio, así como los estadísticos de contraste d_K y d_H . El límite de control entre líneas cortas se construye a un nivel de significación $\alpha = 0.01$, corresponden a los valores críticos c_K y c_H . El siguiente código hace referencia al uso del paquete ILS para tratar el conjunto de datos TG.

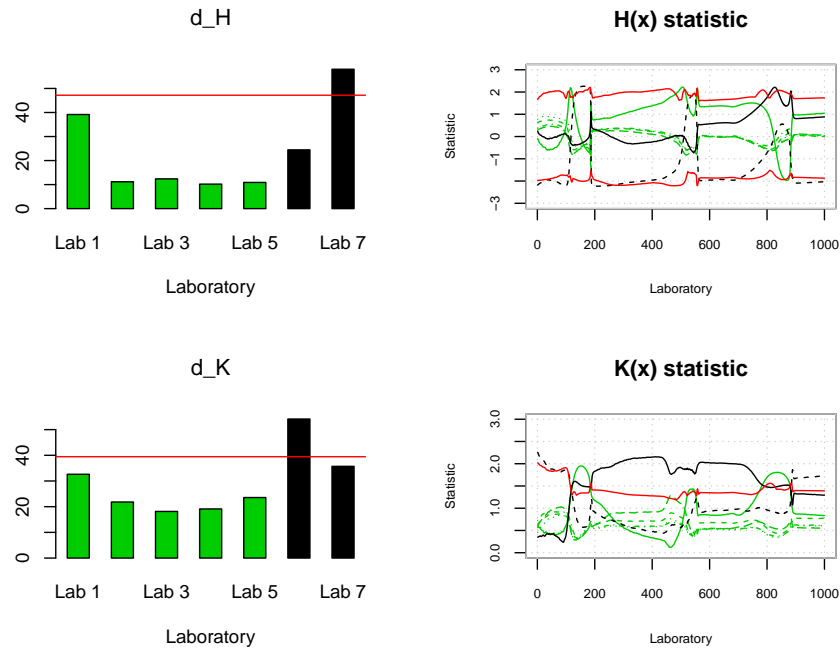


Figura B.7: Conjunto de datos TG: Estadísticos funcionales $H(x)$ y $K(x)$ para cada laboratorio. Estadísticos de contraste d_H y d_K para cada laboratorio.

```
R> mandel.tg <- mandel.fqcs(fqcdata,nb = 10)
R> plot(mandel.tg,legend = T,col=c(rep(3,5),1,1))
```

La hipótesis de reproducibilidad se prueba usando directamente el estadístico de contraste d_H y la de repetibilidad el estadístico de contraste d_K . Mientras que los estadísticos funcionales $K(x)$ y $H(x)$ permiten tener una interpretación gráfica en el dominio de los resultados (curvas). En la Figura B.7 se muestran los estadísticos $K(x)$ y $H(x)$ para cada laboratorio con $\alpha = 0.01$. Los laboratorios 1, 6 y 7 son detectados como laboratorios atípicos en la primera iteración de la metodología. La región correspondiente a la primera, segunda y tercera etapas de degradación del oxalato de calcio está fuera de las bandas de confianza de 99%.

Adicionalmente, en la Figura B.7 los estadísticos de contraste d_H y d_K son graficados y comparados con respecto a sus correspondientes valores críticos c_H y c_K (límites de control), definidos como los cuantiles correspondientes a $\alpha = 0.01$. Para el

estadístico d_H se concluye que el laboratorio 7 no cumple la hipótesis de reproducibilidad. En (Flores *et al.*, 2017) se comprueba mediante un proceso iterativo detectar a los laboratorios 1,6, y 7 como inconsistentes. Mientras que para el caso de la hipótesis de repetibilidad con el estadístico de contraste d_K se detecta el laboratorio 6.

Finalmente, se usa del paquete ILS para presentar los resultados para el conjunto de datos DSC. En la Figura B.8 no se rechaza la hipótesis de repetibilidad pero si la de reproducibilidad para el laboratorio 6.

```
R> data(DSC, package = "ILS")
R> fqcddata.dsc <- ils.fqcddata(DSC, p = 6,
+ index.laboratory = paste("Lab",2:7), argvals = delta)
R> mandel.dsc <- mandel.fqcs(fqcddata.dsc,nb = 10)
R> plot(mandel.dsc,legend = F,col=c(rep(3,4),1,3))
```

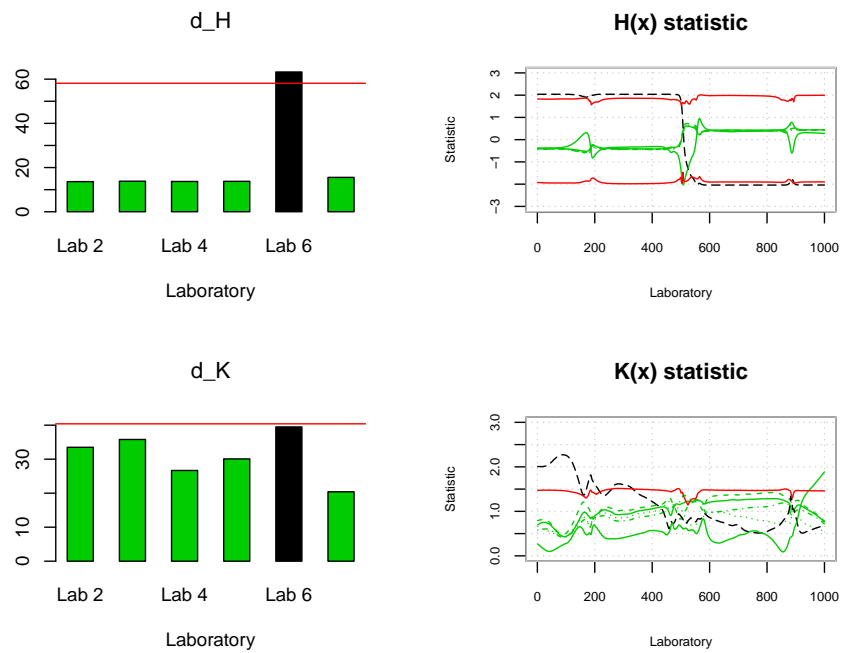


Figura B.8: Conjunto de datos DSC: Estadísticos funcionales $H(x)$ y $K(x)$ para cada laboratorio. Estadísticos de contraste d_H y d_K para cada laboratorio.

Bibliografía

- Alt, F. y Smith, N. (1988). 17 Multivariate Process Control. *Handbook of Statistics*, 7:333–351.
- Alwan, L. C. y Roberts, H. V. (1988). Time-series modeling for statistical process control. *Journal of Business & Economic Statistics*, 6(1):87–95.
- Artiaga, R., López-Beceiro, J., Tarrío-Saavedra, J., Gracia-Fernández, C., Naya, S., y Mier, J. (2011). Estimating the reversing and non-reversing heat flow from standard DSC curves in the glass transition region. *Journal of Chemometrics*, 25(6):287–294.
- ASTM-E691 (2004). *Practice for conducting and interlaboratory study to determine the precision of a test method*. West Conshohocken, USA.
- Barbeito, I. H., Zaragoza, S., Tarrío-Saavedra, J., y Naya, S. (2017). Assessing thermal comfort and energy efficiency in buildings by statistical quality control for autocorrelated data. *Applied Energy*, 190:1–17.
- Bezener, M. y Qiu, P. (2011). Package `mnspec`. *Transactions*, 40:664–677.
- Bisgaard, S. (2012). The future of quality technology: From a manufacturing to a knowledge economy & from defects to innovations. *Quality Engineering*, 24(1):30–36.
- Bissell, A. (1990). How reliable is your capability index?. *Applied Statistics*, pp. 331–340.
- Box, G. E., Luceño, A., y Paniagua-Quinones, M. d. C. (2011). *Statistical control by monitoring and adjustment.*, volumen 700. John Wiley & Sons.

- Boyles, R. (1991). The Taguchi capability index. *Journal of Quality Technology*, 23:17–26.
- Brumback, B. y Rice, J. (1998). Smoothing spline models for the analysis of nested and crossed samples of curves. *Journal of American Statistical Association*, 93:961–94.
- Cano, E., Martínez-Moguerza, J., y Redchuk, A. (2012). *Six Sigma with R: Statistical engineering for process improvement*. Springer Science & Business Media.
- Cascos, I., López, A., y Romo, J. (2011). Data depth in multivariate statistics. *Boletín de Estadística e Investigación Operativa*, 27(3):151–174.
- Chakraborti, S. (2011). Nonparametric (Distribution-Free) Quality Control Charts. *Encyclopedia of Statistical Sciences*.
- Chao-Wen, L. y Reynolds Jr, M. R. (1999). EWMA control charts for monitoring the mean of autocorrelated processes. *Journal of Quality Technology*, 31(2):166.
- Chatterjee, S. y Qiu, P. (2009). Distribution-free cumulative sum control charts using bootstrap-based control limits. *The Annals of Applied Statistics*, pp. 349–369.
- Chipman, H., MacKay, R., y Steiner, S. (2010). Comment on Nonparametric profile monitoring by mixed effects modeling. *Technometrics*, 52(3):280–283.
- Chou, Y.-M., Owen, D., y Borrego, S. (1990). Lower confidence limits on process capability indices. *Journal of Quality Technology*, 22(3):223–229.
- Cochran, W. (1941). The distribution of the largest of a set of estimated variances as a fraction of their total. *Annals of Eugenics*, 11:47–52.
- Colosimo, B. M. y Pacella, M. (2010). A comparison study of control charts for statistical monitoring of functional data. *International Journal of Production Research*, 48(6):1575–1601.
- Committee, A. M. (1989). Robust Statistics - How Not to Reject Outliers. Part 1. Basic concepts. *Analyst*, 114(12):1693–1697.

- Crosier, R. (1988). Multivariate generalizations of cumulative sum quality-control schemes. *Technometrics*, 30(3):291–303.
- Cuesta-Albertos, J. y Febrero-Bande, M. (2010). A simple multiway ANOVA for functional data. *Test*, 19:537–57.
- Cuesta-Albertos, J. y Nieto-Reyes, A. (2008). The random Tukey depth. *Computational Statistics and Data Analysis*, 52:4979–4988.
- Cuevas, A., Febrero-Bande, M., y Fraiman, R. (2004). An ANOVA test for functional data. *Computational Statistics and Data Analysis*, 47:111–22.
- Cuevas, A., Febrero-Bande, M., y Fraiman, R. (2006). On the use of the bootstrap for estimating functions with functional data. *Computational Statistics and Data Analysis*, 51:1063–74.
- Cuevas, A., Febrero-Bande, M., y Fraiman, R. (2007). Robust estimation and classification for functional data via projection based depth notions. *Computational Statistics*, 22:481–96.
- Daly, P. W. (2007). *Natural Sciences Citations and References*.
- Davies, P. (1988). Statistical evaluation of interlaboratory tests. *Fresenius'Zeitschrift für analytische Chemie*, 331(5):513–519.
- Davis, J., Edgar, T., Porter, J., Bernaden, J., y Sarli, M. (2012). Smart manufacturing, manufacturing intelligence and demand-dynamic performance. *Computers & Chemical Engineering*, 47:145–156.
- Del Castillo, E. (2006). Statistical process adjustment: a brief retrospective, current status, and some opportunities for further work. *Statistica Neerlandica*, 60(3):309–326.
- Deleryd, M. y Vännman, K. (1999). Process capability plots – a quality improvement tool. *Quality and Reliability Engineering International*, 15(3):213–227.
- Dvorkin, V. (1994). Data processing in the interlaboratory test by analysis of covariance. *Chemometrics and Intelligent Laboratory Systems*, 22(1):127–146.

- Dyckerhoff, R. (2004). Data depths satisfying the projection property. *Allgemeines Statistisches Archiv*, 88(2):163–190.
- Electric, W. (1956). Statistical quality control handbook. *Easton: The Mack Printing Company*.
- Ellison, S. (2017). *metRology: Support for Metrological Applications*. R package version 0.9-26-2.
- Emilio, L., Martínez-Moguerza, J., Prieto, M., y Redchuk, A. (2017). Six Sigma Tools for Quality Control and Improvement. *R package*.
- Epprecht, E. K. (2015). *Statistical Control of Multiple-Stream Processes: A Literature Review*, pp. 49–64. Springer International Publishing, Cham.
- Fan, J. y Lin, K. (1998). Functional ANOVA models for proportional hazards regression. *Journal of American Statistical Association*, 93:1007–21.
- Febrero-Bande, M., Galeano, P., y González-Manteiga, W. (2007). Outlier detection in functional data by depth measures, with application to identify abnormal NOx levels. *Environmetrics*, 19(4):331–345.
- Febrero-Bande, M. y Oviedo de la Fuente, M. (2012). Statistical computing in functional data analysis: the R package *fda.usc*. *Journal of Statistical Software*, 51:1–28.
- Ferraty, F. y Vieu, P. (2006). *Nonparametric functional data analysis*. Berlin: Springer-Verlag.
- Flores, M. (2016). *qcr: Quality Control Review*. R package version 1.0.
- Flores, M., Fernández-Casal, R., Naya, S., Tarrío-Saavedra, J., y Bossano, R. (2018a). ILS: An R package for statistical analysis in Interlaboratory Studies. *Chemometrics and Intelligent Laboratory Systems*, 181:11 – 20.
- Flores, M., Naya, S., Tarrío-Saavedra, J., y Fernández-Casal, R. (2017). *Functional Statistics and Related Fields.*, capítulo Functional data analysis approach of Mandel’s h and k statistics in Interlaboratory Studies., pp. 123–130. Springer.

- Flores, M., Tarrío-Saavedra, J., Fernández-Casal, R., y Naya, S. (2018b). Functional extensions of Mandel's h and k statistics for outlier detection in interlaboratory studies. *Chemometrics and Intelligent Laboratory Systems*, 176:134 – 148.
- Fraiman, R., Liu, R., y Meloche, J. (1997). Multivariate density estimation by probing depth. *Lecture Notes-Monograph Series*, pp. 415–430.
- Fraiman, R. y Muniz, G. (2001). Trimmed means for functional data. *Test*, 10:419–440.
- Francisco-Fernández, M., Tarrío-Saavedra, J., Mallik, A., y Naya, S. (2012). A comprehensive classification of wood from thermogravimetric curves. *Chemometrics and Intelligent Laboratory Systems*, 118:159–172.
- Francisco-Fernández, M., Tarrío-Saavedra, J., Naya, S., López-Beceiro, J., y Artiaga, R. (2015). Classification of wood using differential thermogravimetric analysis. *Journal of Thermal Analysis and Calorimetry*, 120(1):541–551.
- Francisco-Fernández, M., Tarrío-Saavedra, J., Naya, S., López-Beceiro, J., y Artiaga, R. (2017). Statistical classification of early and late wood through the growth rings using thermogravimetric analysis. *Journal of Thermal Analysis and Calorimetry*, 127(1):499–506.
- García, D. R. (2011). Cartas de control para datos funcionales. *CIMAT*.
- Gracia-Fernández, C., Davies, P., Gómez-Barreiro, S., López, B. J., Tarrío-Saavedra, J., y Artiaga, R. (2010a). A vitrification and curing study by simultaneous TMDSC-photocalorimetry. *Journal of Thermal Analysis and Calorimetry*, 102(3):1057–1062.
- Gracia-Fernández, C., Gómez-Barreiro, S., López-Beceiro, J., Saavedra, J. T., Naya, S., y Artiaga, R. (2010b). Comparative study of the dynamic glass transition temperature by DMA and TMDSC. *Polymer Testing*, 29(8):1002–1006.
- Gracia-Fernández, C., Tarrío-Saavedra, J., López-Beceiro, J., Gómez-Barreiro, S., Naya, S., y Artiaga, R. (2011). Temperature modulation in PDSC for monito-

- ring the curing under pressure. *Journal of Thermal Analysis and Calorimetry*, 106(1):101–107.
- Grubbs, F. (1950). Sample criteria for testing outlying observations. *Annals of Mathematical Statistics*, 21(1):27–58. cited By 385.
- Grubbs, F. y Beck, G. (1972). Extension of sample sizes and percentage points for significance tests of outlying observations. *Technometrics*, 14(4):847–854.
- Hawkins, D. (1991). Multivariate quality control based on regression-adjusted variables. *Technometrics*, 33(1):61–75.
- Healy, J. (1987). A note on multivariate CUSUM procedures. *Technometrics*, 29(4):409–412.
- Hothorn, T., Bretz, F., Westfall, P., Heiberger, R., Schuetzenmeister, A., y Scheibe, S. (2017). *multcomp: Simultaneous Inference in General Parametric Models*. R package version 1.4-7.
- Hubele, N. F. y Vannman, K. (2004). The Effect of Pooled and Un-pooled Variance Estimators on Cpm When Using Subsamples. *Journal of quality technology*, 36(2):207.
- Hund, E Massart, D. y Smeyers-Verbeke, J. (2014). Interlaboratory studies in analytical chemistry. *Analytica Chimica Acta*, 423:145–165.
- Hund, E., Massart, D., y Smeyers-Verbeke, J. (2000a). Interlaboratory studies in analytical chemistry. *Analytica Chimica Acta*, 423:145–65.
- Hund, E., Massart, D., y Smeyers-Verbeke, J. (2000b). Interlaboratory studies in analytical chemistry. *Analytica Chimica Acta*, 423(2):145–165.
- Hyndman, R. J. y Shang, H. L. (2010). Rainbow plots, bagplots, and boxplots for functional data. *Journal of Computational and Graphical Statistics*, 19(1):29–45.
- Hyndman, R. J. y Ullah, M. S. (2007). Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics and Data Analysis*, 51(10):4942 – 4956.

- Imkamp, D., Gabbia, A., y Berthold, J. (2014). *Challenges and Trends in Manufacturing Metrology-VDI/VDE Roadmap*. Universitätsbibliothek Ilmenau.
- ISO-5725 (1994). *International Standard ISO 5725-4-1994: Accuracy (Trueness and precision) of measurement methods and results — Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method*. Geneva, Suiza.
- Janeiro-Arocas, J., Tarrío-Saavedra, J., López-Beceiro, J., Naya, S., López-Canosa, A., Heredia-García, N., y Artiaga, R. (2016). Creep analysis of silicone for Podiatry applications. *Journal of the Mechanical Behavior of Biomedical Materials*, 63:456–469.
- Jensen, W. A., Jones-Farmer, L. A., Champ, C. W., y Woodall, W. H. (2006). Effects of parameter estimation on control chart properties: a literature review. *Journal of Quality Technology*, 38(4):349.
- Jones, L. y Woodall, W. (1998). The performance of bootstrap control charts. *Journal of Quality Technology*, 30(4):362.
- Jones-Farmer, L., Woodall, W., Steiner, S., y Champ, C. (2014). An overview of phase I analysis for process improvement and monitoring. *Journal of Quality Technology*, 46(3):265.
- Kelly, P. (1990). Outlier detection in collaborative studies. *Journal of the Association of Official Analytical Chemists*, 73(1):58–64.
- Knoth, S. y Schmid, W. (2004). Control charts for time series: a review. *Frontiers in Statistical Quality Control*, 7:210–236.
- Lanhede, D. (2015). *Statistical Process Control: Evaluation and implementation of Methods for Statistical Process Control at GE Healthcare*. Tesis de máster, Umeå University, Sweden.
- Ledolter, J. y Bisgaard, S. (2011). Challenges in constructing time series models from process data. *Quality and Reliability Engineering International*, 27(2):165–178.

- Lehmann, EL y Romano, J. P. (2005). *Prueba de hipótesis estadísticas*. Springer, Nueva York, tercera edición.
- Li, Z., Qiu, P., Chatterjee, S., y Wang, Z. (2013). Using p values to design statistical process control charts. *Statistical Papers*, 54(2):523–539.
- Liu, R. (1990). On a notion of data depth based on random simplices. *The Annals of Statistics*, 18(1):405–414.
- Liu, R. Y. (1995). Control Charts for Multivariate Processes. *Journal of the American Statistical Association*, 90(432):1380–1387.
- Liu, R. Y. y Tang, J. (1996). Control charts for dependent and independent measurements based on bootstrap methods. *Journal of the American Statistical Association*, 91(436):1694–1700.
- López-Beceiro, J., Artiaga, R., Gracia, C., Tarrío-Saavedra, J., Naya, S., y Mier, J. (2010). Comparison of olive, corn, soybean and sunflower oils by PDSC. *Journal of Thermal Analysis and Calorimetry*, 104(1):169–175.
- López-Beceiro, J., Pascual-Cosp, J., Artiaga, R., Tarrío-Saavedra, J., y Naya, S. (2011). Thermal characterization of ammonium alum. *Journal of Thermal Analysis and Calorimetry*, 104(1):127–130.
- Lowry, C., Woodall, W., y Champ, Cand Rigdon, S. (1992). A multivariate exponentially weighted moving average control chart. *Technometrics*, 34(1):46–53.
- Lysiak-Pastuszak, E. (2004). Interlaboratory analytical performance studies: a way to estimate measurement uncertainty. *Oceanologia*, 46(3):427–438.
- Mahalanobis, P. (1936). On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*, pp. 49–55.
- Mallik, A., Tarrío-Saavedra, J., Francisco-Fernández, M., y Naya, S. (2011). Classification of wood micrographs by image segmentation. *Chemometrics and Intelligent Laboratory Systems*, 107(2):351–362.

- Mandel, J. (1985). A new analysis of interlaboratory test results. *ASQC Quality Congress Transaction-Baltimore*, pp. 60–366.
- Megahed, F. y Jones-Farmer, L. (2013). A Statistical Process Monitoring Perspective on “Big Data”, *Frontiers in Statistical Quality Control*.
- Mei, Y. (2010). Efficient scalable schemes for monitoring a large number of data streams. *Biometrika*, 97(2):419–433.
- Menczel, J. y Prime, R. (2014). *Thermal analysis of polymers: Fundamentals and applications*. John Wiley & Sons.
- Miller, J. y Miller, J. (1988). *Statistics for analytical chemistry*. John Wiley and Sons, New York, NY.
- Minkkinen, P. (1995). Estimation of variance components from the results of interlaboratory comparisons. *Chemometrics and Intelligent Laboratory Systems*, 29(2):263–270.
- Montgomery, D. (2009). *Introduction to Statistical Quality Control*. John Wiley & Sons (New York).
- Mutilba, U. y Kortaberria, G. (2016). El rol de la Metrología en la Industria 4.0. *e-medida*, 11:1–3.
- Naya, S., Meneses, A., Tarrío-Saavedra, J., Artiaga, R., López-Beceiro, J., y Gracia-Fernández, C. (2013). New method for estimating shift factors in time–temperature superposition models. *Journal of Thermal Analysis and Calorimetry*, 113(2):453–460.
- Naya, S., Tarrío-Saavedra, J., López-Beceiro, J., Francisco-Fernández, M., Flores, M., y Artiaga, R. (2014). Statistical functional approach for interlaboratory studies with thermal data. *Journal of Thermal Analysis and Calorimetry*, 118(2):1229–1243.
- Newton, I. y Campbell, N. (1997). A Re-formulation of Gravity with Respect to Really Cool Models. *Journal of Funny Physics*, 35:39–78.

- Noorossana, R., Saghaei, A., y Amiri, A. (2011). *Statistical analysis of profile monitoring.*, volumen 865. John Wiley & Sons.
- Pearn, W. y Chen, K. (1997). A practical implementation of the process capability index Cpk. *Quality Engineering*, 9(4):721–737.
- Pignatiello, J. y Runger, G. (1990). Comparisons of multivariate CUSUM charts. *Journal of Quality Technology*, 22(3):173–186.
- Polansky, A. (2007). Process Capability Indices, Nonparametric. *Encyclopedia of Statistics in Quality and Reliability*.
- Prajapati, D. y Singh, S. (2012). Control charts for monitoring the autocorrelated process parameters: a literature review. *International Journal of Productivity and Quality Management*, 10(2):207–249.
- Qiu, P. y Li, Z. (2011). On nonparametric statistical process control of univariate processes. *Technometrics*, 53(4):390–405.
- Qiu, P., Zou, C., y Wang, Z. (2010). Nonparametric profile monitoring by mixed effects modeling. *Technometrics*, 52(3):265–277.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Raña, P. (2016). *Pointwise forecast, confidence and prediction intervals in electricity demand and price*. Tesis doctoral, Universidade da Coruña, da Coruña.
- Ramsay, J. y Hooker, G. (2009). *Functional data analysis with R and Matlab*. New York: Springer.
- Ramsay, J. y Silverman, B. (2002). *Applied functional data analysis*. New York: Springer-Verlag.
- Ramsay, J. y Silverman, B. (2005). *Functional data analysis*. New York: Springer-Verlag.
- Recchia, D., Barbosa, E., y Gonçalves, E. d. J. (2010). IQCC: Improved Quality Control Charts. *R package version 0.5*.

- Ríos-Fachal, M., Gracia-Fernández, C., López-Beceiro, J., Gómez-Barreiro, S., Tarrío-Saavedra, J., Ponton, A., y Artiaga, R. (2013). Effect of nanotubes on the thermal stability of polystyrene. *Journal of Thermal Analysis and Calorimetry*, 113(2):481–487.
- Ríos-Fachal, M., Tarrío-Saavedra, J., López-Beceiro, J., Naya, S., y Artiaga, R. (2014). Optimizing fitting parameters in thermogravimetry. *Journal of Thermal Analysis and Calorimetry*, 116(3):1141–1151.
- Rodrigo, O. P. (2013). Monitoreo de datos funcionales. *CIMAT*.
- Roth, T. (2012). qualityTools: Statistics in Quality Science. *R package*.
- Roth, T. (2016). *qualityTools: Statistics in Quality Science*. R package version 1.55 <http://www.r-qualitytools.org>.
- Runger, G. y Prabhu, S. (1996). A Markov chain model for the multivariate exponentially weighted moving averages control chart. *Journal of the American Statistical Association*, 91(436):1701–1706.
- Rüssmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., y Harnisch, M. (2015). Industry 4.0: The future of productivity and growth in manufacturing industries. *Boston Consulting Group*, 9.
- Santos-Fernandez, E. (2016). Package MSQC. *R package*.
- Schmitt, R. y Peterek, M. (2015). Traceable measurements on machine tools - Thermal influences on machine tool structure and measurement Uncertainty. *Procedia CIRP*, 33:576–580.
- Schmitt, R., Peterek, M., Morse, E., Knapp, W., Galetto, M., Härtig, F., Goch, G., Hughes, B., Forbes, A., y Estler, W. (2016). Advances in Large-Scale Metrology—Review and future trends. *CIRP Annals-Manufacturing Technology*, 65(2):643–665.
- Scrucca, L. (2004). qcc: an R package for quality control charting and Statistical Process Control. *dim (pistonrings)*, 1(200):3.

- Sebio-Puñal, T., Naya, S., López-Beceiro, J., Tarrío-Saavedra, J., y Artiaga, R. (2012). Thermogravimetric analysis of wood, holocellulose, and lignin from five wood species. *Journal of Thermal Analysis and Calorimetry*, 109(3):1163–1167.
- Shen, Q. y Faraway, J. (2004). An F test for linear models with functional responses. *Statistica Sinica*, 14:1239–57.
- Sheu, S.-H., Ouyoung, C.-W., y Hsu, T.-S. (2013). Phase II statistical process control for functional data. *Journal of Statistical Computation and Simulation*, 83(11):2144–2159.
- Shiau, J.-J., Huang, H.-L., Lin, S.-H., y Tsai, M.-Y. (2009). Monitoring nonlinear profiles with random effects by nonparametric regression. *Communications in Statistics—Theory and Methods*, 38(10):1664–1679.
- Spiegelhalter, D., Sherlaw-Johnson, C., Bardsley, M., Blunt, I., Wood, C., y Grigg, O. (2012). Statistical methods for healthcare regulation: rating, screening and surveillance. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175(1):1–47.
- Stankovic, J. A. (2014). Research directions for the internet of things. *IEEE Internet of Things Journal*, 1(1):3–9.
- Stapenhurst, T. (2013). *Mastering Statistical Process Control*. Routledge.
- Strucca, L. (2004). qcc: an R package for Quality Control charting and Statistical Process Control. *R News*, 4/1:11–17.
- Tarrío-Saavedra, J., Francisco-Fernández, M., Naya, S., López-Beceiro, J., Gracia-Fernández, C., y Artiaga, R. (2013). Wood identification using pressure DSC data. *Journal of Chemometrics*, 27(12):475–487.
- Tarrío-Saavedra, J., López-Beceiro, J., Álvarez, A., Naya, S., Quintana-Pita, S., García-Pardo, S., y García-Sabán, F. (2015). Lifetime estimation applying a kinetic model based on the generalized logistic function to biopolymers. *Journal of Thermal Analysis and Calorimetry*, 122(3):1203–1212.

- Tarrío-Saavedra, J., López-Beceiro, J., Naya, S., y Artiaga, R. (2008). Effect of silica content on thermal stability of fumed silica/epoxy composites. *Polymer Degradation and Stability*, 93(12):2133–2137.
- Tarrío-Saavedra, J., López-Beceiro, J., Naya, S., Francisco-Fernández, M., y Artiaga, R. (2014). Simulation study for generalized logistic function in thermal data modeling. *Journal of Thermal Analysis and Calorimetry*, 118(2):1253–1268.
- Tarrío-Saavedra, J., López-Beceiro, J., Naya, S., Gracia, C., y Artiaga, R. (2010a). Controversial effects of fumed silica on the curing and thermomechanical properties of epoxy composites. *Express Polymer Letters*, 4(6):382–395.
- Tarrío-Saavedra, J., Naya, S., Francisco-Fernández, M., López-Beceiro, J., y Artiaga, R. (2010b). Functional nonparametric classification of wood species from thermal data. *Journal of Thermal Analysis and Calorimetry*, 104(1):87–100.
- Tong, L.-I. y Chen, J.-P. (1998). Lower confidence limits of process capability indices for non-normal process distributions. *International Journal of Quality & Reliability Management*, 15(8/9):907–919.
- Tukey, J. (1975). Mathematics and the picturing of data. En *Proceedings of the international congress of mathematicians*, volumen 2, pp. 523–531.
- Uhlig, S. y Lischer, P. (1998). Statistically-based performance characteristics in laboratory performance studies. *Analyst*, 123(2):167–172.
- Vännman, K. (1995). A unified approach to capability indices. *Statistica Sinica*, pp. 805–820.
- Vännman, K. (2001). A graphical method to control process capability. En *Frontiers in Statistical Quality Control 6*, pp. 290–311. Springer - Verlag Berlin Heidelberg GmbH.
- Wells, L., Megahed, F., Camelio, J., y Woodall, W. H. (2012). A framework for variation visualization and understanding in complex manufacturing systems. *Journal of Intelligent Manufacturing*, 23(5):2025–2036.

- Wilrich, P. (2013). Critical values of Mandel's h and k, the Grubbs and the Cochran test statistic. *AStA Advances in Statistica Analysis*, 97:1–10.
- Wolak, M., Fairbairn, D., y Paulsen, Y. (2012). Guidelines for Estimating Repeatability. *Methods in Ecology and Evolution* 3(1):129-137.
- Woodall, W. (2000). Controversies and contradictions in Statistical Process Control (with Discussions). *Journal of Quality Technology*, 32:341–350.
- Woodall, W., J Spitzner, D., C Montgomery, D., y Gupta, S. (2004). Using Control Charts to Monitor Process and Product Quality Profiles. *Journal of Quality Technology*, 36.
- Woodall, W. y Montgomery, D. (2014). Some current directions in the theory and application of statistical process monitoring. *Journal of Quality Technology*, 46(1):78.
- Woodall, W. y Ncube, M. (1985). Multivariate CUSUM quality-control procedures. *Technometrics*, 27(3):285–292.
- Youden, W. (1959). *Industrial Quality Control*.
- Zuo, Y. y Serfling, R. (2000). General notions of statistical depth function. *Annals of Statistics*, pp. 461–482.