



**UNIVERSIDAD SAN FRANCISCO DE QUITO**

**Colegio de Posgrados**

Análisis de Datos Meteorológicos del Valle de los Chillos  
usando Datos Funcionales

**Edgar Ramiro Guerrón Varela**

**Carlos Jiménez Mosquera, PhD., Director de Tesis**

**Tesis de grado presentada como requisito para la obtención  
del título de Magíster en Matemáticas Aplicadas**

**Quito, abril 2015**

**Universidad San Francisco de Quito**

**Colegio de Posgrados**

**HOJA DE APROBACIÓN DE TESIS**

**Análisis de Datos Meteorológicos del Valle de los Chillos usando Datos  
Funcionales**

**Edgar Ramiro Guerrón Varela**

Carlos Jiménez Mosquera, Ph.D.  
Director de Tesis

---

Eduardo Alba, Ph.D.  
Miembro del Comité de Tesis

---

Carlos Jiménez Mosquera, Ph.D.  
Director de la Maestría en Matemática Aplicada y  
Miembro del Comité de Tesis

---

Santiago Gangotena, Ph.D.  
Decano de la Escuela de Ciencias

---

Víctor Viteri Breedy, Ph.D.  
Decano del Colegio de Posgrados

---

Quito, abril 2015

## ©Derechos de Autor

Por medio del presente documento certifico que he leído la Política de Propiedad Intelectual de la Universidad San Francisco de Quito y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo de investigación quedan sujetos a lo dispuesto en la Política.

Así mismo, autorizo a la USFQ para que realice la digitalización y publicación de éste trabajo de investigación en el repositorio virtual, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Firma: \_\_\_\_\_  
Nombre: Edgar Ramiro Guerrón Varela  
C. I.: 1707312698

Quito, abril 2015

## **DEDICATORIA**

A mi esposa Gina, mis hijos Kathy, Ramiro, Ginita y a mi madre Ruth Clemencia,

quienes con su amor, paciencia y entrega, depositaron su confianza en mi, y fueron pilares fundamentales para llegar a cumplir satisfactoriamente éste proyecto de vida.

## **AGRADECIMIENTOS**

A Dios, quien ha sido mi fortaleza en mis momentos de debilidad y mi guía para tomar decisiones correctas.

A mi Madre, quien con su ejemplo has sabido forjar en mi un espíritu de trabajo, honestidad y respeto.

A mi Director de Tesis, Carlos Jiménez Ph.D, quien con paciencia, conocimientos y ayuda ha permitido que llegue a cumplir mi meta.

Al Ing. Marcelo Arce, mi compañero y amigo, quién colaboró de manera desinteresada con toda la información necesaria para el desarrollo de éste proyecto.

## RESUMEN

El análisis de datos funcionales (ADF) es un área de la estadística que tiene como objetivo principal estudiar variables cuyos datos varíen en un espacio continuo, dentro de éstas se encuentran las que dependen del tiempo. Actualmente se puede disponer de una gran cantidad de datos debido a que la mayoría de ellos su recolección se hace de manera automática, por lo tanto ajustar éstos datos a una función o dato funcional es una de las características del análisis funcional, éstos datos una vez convertidos en funciones permiten que se puedan extraer sus derivadas y explicar de mejor manera los fenómenos en estudio. La Estación Meteorológica de la Hacienda El Prado, ubicada en el valle de los Chillos, ha registrado sistemáticamente y de manera analógica y automática una gran variedad de datos meteorológicos, con éstos datos lo que se pretende es modelar el comportamiento de las variables atmosféricas en dicho valle y su área de influencia. En éste análisis se usa software estadístico de libre distribución R, con énfasis en el uso de las librerías `fda` y `fda.usc`. Su desarrollo se basará en el análisis exploratorio de datos, la selección del tipo de base, el número de bases, la construcción de un modelo basado en las series de Fourier y un modelo basado en la construcción de un operador funcional lineal. En éste estudio se ha seleccionado dos variables para estudiar que son las temperaturas diaria mínima máxima. Con esto se pretende dejar una puerta abierta a otros estudios seleccionando otras variables, o escogiendo otra estación meteorológica y poder establecer comparaciones.

## ABSTRACT

The analysis of functional data (ADF) is a statistical area that has as main objective to study variables whose data vary in a continuous space, and within these variables you can find the time-dependent. Currently you can have a lot of data because most of them the gathering is done automatically, therefore to adjust this data to a function or functional data is one of the characteristics of the functional analysis. These data once converted into functions allow its derivatives to be extracted and explain in a better way how the phenomena under study works. The Meteorological Station Hacienda El Prado, located in the valley of Chillos has systematically analog and automatically registered a variety of meteorological data. With these data what is aimed to do is to model the behavior of the atmospheric variables in that Valley and its area of influence. In this statistical analysis free software R is used, with emphasis on the use of the `fda` and `fda.usc` librarys. Its development is based on the exploratory data analysis, selection of the base type, the number of bases, the building of a system based on Fourier series and a model based on the construction of a linear functional operator. In this study two variables have been selected to be studied, they are the daily minimum temperature and daily maximum temperature. The intention of the before mentioned is to give the door open for other studies by selecting other variables, or choosing another weather station and to establish comparisons.

# Contenido

<b>1</b>	<b>INTRODUCCIÓN AL PROBLEMA</b>	<b>15</b>
1.1	Antecedentes . . . . .	17
1.2	El Problema . . . . .	19
1.3	Hipótesis . . . . .	20
1.4	Pregunta de Investigación . . . . .	21
1.5	Contexto y marco teórico . . . . .	21
1.5.1	El propósito del estudio . . . . .	22
1.5.2	El significado del estudio . . . . .	22
1.6	Presunciones de los autores del estudio . . . . .	23
1.7	Supuestos del estudio . . . . .	23
<b>2</b>	<b>ANÁLISIS CON DATOS FUNCIONALES</b>	<b>24</b>
2.1	Introducción . . . . .	24
2.2	Conceptos Básicos: . . . . .	26
2.3	Análisis Exploratorio de los Datos . . . . .	27
2.4	Configuración del Sistema de Bases para la Construcción de Funciones . . . . .	28
2.5	Series de Fourier para Datos Periódicos y Funciones . . . . .	29
2.6	Suavizamiento mediante un Operador Diferencial Lineal . . . . .	31

	10
<b>3 CASO DE ESTUDIO - SERIE DATOS DIARIOS</b>	<b>33</b>
3.1 Exploración de Datos . . . . .	33
3.2 Análisis de Datos Funcionales para registros diarios . . . . .	35
3.2.1 Análisis de Normalidad Temperatura Mínima . . . . .	35
3.2.2 Regresión de la temperatura mínima sobre una base de Fourier . . . . .	36
3.2.3 Modelo de ajuste usando Ldf para la temperatura mínima . . . . .	37
3.3 Análisis Temperatura Máxima . . . . .	39
3.3.1 Análisis de Normalidad Temperatura Máxima . . . . .	39
3.3.2 Regresión de la temperatura máxima sobre una base de Fourier . . . . .	39
3.3.3 Modelo de ajuste usando Ldf para la temperatura máxima . . . . .	40
3.4 Análisis de comparación entre temperaturas . . . . .	41
3.4.1 Componentes de la serie temporal de la diferencia de temperaturas . . . . .	44
3.4.2 Comparación entre las temperaturas usando los modelos generados . . . . .	44
<b>4 CASO DE ESTUDIO - SERIE DATOS PROMEDIOS</b>	<b>46</b>
4.1 Análisis Exploratorio de Datos . . . . .	47
4.2 Ajuste de modelo con bases Fourier . . . . .	47
4.3 Modelo de ajuste usando Ldf para la temperatura promedio mínima . . . . .	50
4.4 Modelo de ajuste usando Ldf para la temperatura promedio mínima . . . . .	51
4.5 Correlación entre las variables . . . . .	52
4.5.1 Correlación entre temperaturas . . . . .	52
4.5.2 Matriz de Correlación entre las variables . . . . .	52
<b>5 LIMITACIONES DEL ESTUDIO</b>	<b>54</b>
<b>6 RECOMENDACIONES PARA ESTUDIOS FUTUROS</b>	<b>55</b>

<b>7</b>	<b>REFERENCIAS BIBLIOGRÁFICAS</b>	<b>57</b>
<b>8</b>	<b>ANEXOS</b>	<b>60</b>
8.1	ANEXO A: DEFINICIÓN DE TÉRMINOS . . . . .	60
8.1.1	Definiciones en Análisis de Datos Funcionales: . . . . .	63
8.1.2	Espacio de Funciones de Cuadrado Integrable: . . . . .	65
8.2	ANEXO B: CÓDIGO R - DATOS FUNCIONALES . . . . .	72

# Lista de Figuras

1.1	Ubicación Geográfica de la Estación Meteorológica Hacienda El Prado . . . . .	17
1.2	Estación Meteorológica Hacienda El Prado . . . . .	18
1.3	Zona de influencia de la Estación Meteorológica Hacienda El Prado . . . . .	19
2.1	Sistemas Bases de Fourier . . . . .	31
3.1	Series de variables meteorológicas diarias (1998 - 2014) . . . . .	34
3.2	QQ - Plot Temperatura Mínima . . . . .	35
3.3	Modelo ajustado a una Base de Fourier de la Temperatura Mínima . . . . .	37
3.4	Modelo de ajuste de la temperatura mínima usando el operador diferencial lineal	38
3.5	QQ - Plot Temperatura Máxima . . . . .	39
3.6	Modelo ajustado a una Base de Fourier de la Temperatura Máxima . . . . .	40
3.7	Modelo de ajuste de la temperatura máxima usando el operador diferencial lineal	41
3.8	Diagramas de caja temperatura máxima y mínima promedio diaria . . . . .	42
3.9	Comparación entre la temperatura máxima y mínima . . . . .	43
3.10	Componentes Serie temporal diferencia de temperaturas . . . . .	44
3.11	Diferencia de temperaturas entre modelos ajustados . . . . .	45
4.1	Series variables meteorológicas promedio diarias . . . . .	46
4.2	Modelo ajustado con series de Fourier - 5 funciones . . . . .	48
4.3	Modelo ajustado con series de Fourier - 10 funciones . . . . .	49

4.4	Modelo ajustado con Ldf - Temperatura Mínima . . . . .	50
4.5	Modelo ajustado con Ldf - Temperatura Mínima . . . . .	50
4.6	Modelo ajustado con Ldf - Temperatura Mínima - Temperatura Máxima . . . . .	51
4.7	Comparación modelos ajustado Ldf - Temperatura máxima . . . . .	51
4.8	Comparación modelos de la temperatura máxima y mínima ajustado Ldf . . . . .	52

# Lista de Tablas

3.1	Resumen descriptivo de la base de datos diarios (1998 - 2014) . . . . .	34
3.2	Resumen descriptivo modelo temperatura mínima usando Lfd . . . . .	38
4.1	Resumen de la información con la serie con los promedios por día . . . . .	47
4.2	Resumen de la regresión con cinco funciones base . . . . .	48
4.3	Resumen de la regresión con diez funciones base . . . . .	49
4.4	Matriz de Correlación de la serie con datos promedios por día . . . . .	52

# Capítulo 1

## INTRODUCCIÓN AL PROBLEMA

Un problema de actualidad, de mucha discusión a nivel nacional e internacional es ¿hacia donde vamos con el cambio climático?, ¿es cierto el cambio?, ¿hay como medir?, ¿qué parámetros meteorológicos están cambiando?, ¿podemos pronosticar alguno de los parámetros?.

Sin duda hay una gran cantidad de interrogantes al respecto, como hay también varias metodologías para estudiar el clima, una de las más empleadas es Series de Tiempo y actualmente una metodología es el Análisis de Datos Funcionales.

Éstos estudios tienen gran importancia en la agricultura, se tiene un panorama más claro o menos incierto acerca de cuándo sembrar, en un proyecto de construcción cuando planear el inicio de la misma, la fundición de una losa, etc, en el área turística, para que fecha realizar un viaje o posponer por causas meteorológicas, vuelos de aviones, y otras actividades.

En Ecuador quien es el encargado del monitoreo climático y su predicción es el Instituto Nacional de Meteorología e Hidrología, quien posee 1156 Estaciones Meteorológicas activas, entre convencionales y automáticas, alrededor de todo el país.<sup>1</sup>, en la cual se desarrollan

---

<sup>1</sup><http://www.serviciometeorologico.gob.ec/red-de-estaciones-meteorologicas/>

modelos físico-matemáticos empleando software como WRF (Weather Research and Forecasting) y MM5 (Fifth-Generation Penn State/NCAR Mesoscale Model), con los cuales se realizan predicciones a tres días en intervalos de tres, seis y veinte y cuatro horas, pronósticos estacionales de clima empleando modelos estadísticos con CPT (Climate Predictability Tool) para pronósticos mensuales y trimestrales y finalmente un software a prueba el CWRF para hacer pronósticos probabilísticos<sup>2</sup>.

La Empresa Municipal de agua potable del Distrito Metropolitano de Quito, monitorea variables metereologicas a través de 32 estaciones, para controlar la dotación de agua potable para el Distrito de la ciudad.

Para el presente estudio se trabajará con la información de la Estación Meteorológica localizada en el Instituto Superior Andino -IASA-<sup>3</sup> de la carrera de Agropecuaria de la Universidad de las Fuerzas Armadas -ESPE-<sup>4</sup>, ubicada en la Hacienda el Prado, Cantón Sangolquí, Provincia de Pichincha, República del Ecuador, con coordenadas geográficas, Longitud: 78° 24' 44" , Latitud: 0° 23' 20" , Altitud: 2748 m.s.n.m., proporcionada por el Ing. Marcelo Arce, responsable de la mencionada estación

---

<sup>2</sup><http://www.serviciometeorologico.gob.ec/modelos-hidrometeorologicos/>

<sup>3</sup><http://iasa1.espe.edu.ec/>

<sup>4</sup><http://www.espe.edu.ec/>

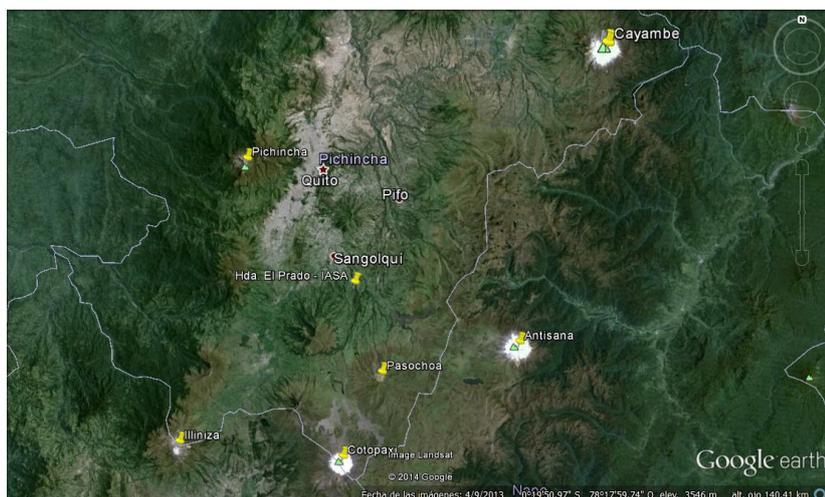


Figura 1.1: Ubicación Geográfica de la Estación Meteorológica Hacienda El Prado

## 1.1 Antecedentes

La Carrera Agropecuaria de la Universidad de las Fuerzas Armadas -ESPE- fue fundada con el nombre de Instituto Agropecuario Superior Andino, conocido como -IASA- el 4 de abril de 1994, y la Estación Meteorológica del IASA empezó su funcionamiento el 1 de enero de 1998, estación de tipo convencional, la cual registra temperatura, heliofanía, humedad relativa, presión atmosférica, velocidad del viento, ésta estación no consta como Estación en el INAMHI debido a que está junto a la Estación digital de EMAPA que se trata de un tipo de estación automática, pero que su funcionamiento empezó el 2003, y que sufre frecuentes daños en el sistema informático debido a que la zona presenta meteroros eléctricos continuos lo cual ocasiona su paralización, y no haya series continuas.



Figura 1.2: Estación Meteorológica Hacienda El Prado

La Estación de la Hacienda el Prado (IASA), desde su inicio está bajo la responsabilidad del Ing. Marcelo Arce Carriel. Por su ubicación geográfica, se destaca su importancia por la influencia en la zona, se encuentra localizada en el Valle de los Chillos, sector San Fernando y que corresponde a la Cuenca Hidrográfica del Río Pita y a la subcuenca del Río Pillacoto o Santa Rita, y su radio abarca aproximadamente unos siete kilómetros, y ésto se debe a que la topografía de la zona es muy irregular. Existen otras estaciones que podría complementar o referenciar el estudio las de Tumbaco y Machachi.

La zona presenta características muy peculiares ya que está influenciada por la zona de convergencia intertropical del Atlántico en el abra de Pifo hacia la Amazonía, hacia el sur por el volcán Cotopaxi, hacia el Norte el declive correspondiente al valle de los Chillos, igualmente las cadenas montañosas cuyos elevaciones cercanas son el Antisana, Rumiñahui. Todas éstas condiciones hace que se genere un microclima donde se producen cambios bruscos en precipitación (granizadas y aguaceros fuertes y continuos), humedad relativa promedio mayor al 50 por ciento y cambios de temperatura en las mínimas y máximas.

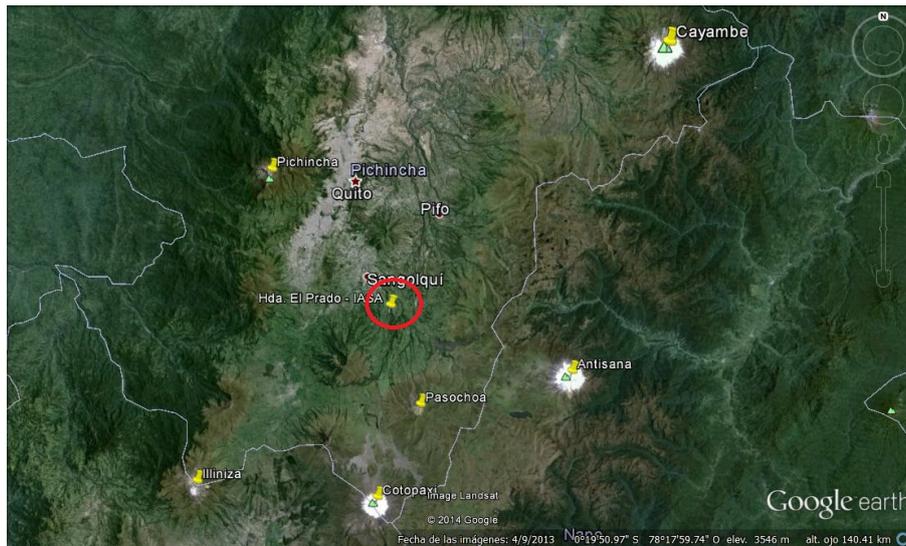


Figura 1.3: Zona de influencia de la Estación Meteorológica Hacienda El Prado

Para éste estudio se cuenta con los registros diarios desde el primero de enero de 1998 hasta el 31 de junio del 2014, también se disponen de los registros promedios mensuales para las mismas fechas y la última serie de datos corresponde a los promedios de cada día del año, y ésto se los hace conociendo que se tiene una serie que se considera periódica anual.

## 1.2 El Problema

Todos los días se escucha o se lee por los diferentes medios de comunicación acerca del cambio climático, que la temperatura del planeta está cambiando, que está subiendo unos cuantos grados, que el hielo en los polos se está derritiendo, el efecto invernadero, etc.. ¿Pero podemos decir en qué medida ha crecido éste incremento? o ¿cuál se su variabilidad?, posiblemente podamos decir que si ha cambiado, pero no sabemos en que medida.

El tratar de comprender el clima y sus diferentes variables es un tema de actualidad, la predicción o pronósticos diarios en la prensa escrita, radial y televisiva, por lo general nos informan de como va a estar el tiempo en corto tiempo, tres horas, seis horas o pueden ser de días e inclusive meses, sin embargo desconocemos cómo ha ido cambiando éstas variables climáticas o meteorológicas, ya que no se encuentran muchos estudios específicos por zonas en el país con respecto a éstos cambios, al igual la información relativa al mismo es insuficiente.

Éste estudio al usará los datos de la Estación metereológica de la Hacienda el Prado, intenta presentar de una manera clara que verdaderamente hay un cambio climático, así como también pretende dejar una metodología para poder replicar en cualquier otra estación metereológica.

### **1.3 Hipótesis**

La hipótesis de éste análisis es que el clima está cambiando en éstos últimos años, hay un aumento progresivo y sostenido de la temperatura, la diferencia de la temperatura entre la temperatura mínima y la máxima es mayor.

Las causad del problema las conocemos de sobra, la emisión de agentes contaminantes que generan el efecto invernadero, la emisión de  $CO_2$  producto de la combustión de los derivados del petróleo en el uso de la transportación, la combustión de gas licuado en las cocinas de uso doméstico, etc.

Las soluciones de igual manera las conocemos, el contribuir en la mejor administración de los recursos que generan contaminación ambiental y de manera específica al aire.

## 1.4 Pregunta de Investigación

La pregunta que pretende validar éste estudio es: ¿Porqué podemos afirmar que hay un cambio climático?

Para lo cual responderemos otras preguntas

¿Cuánto ha variado la temperatura mínima en los últimos catorce años?

¿Cuánto ha variado la temperatura máxima en los últimos catorce años?

¿En cuánto ha variado la temperatura máxima en los últimos catorce años?

## 1.5 Contexto y marco teórico

Para el presente estudio se empleará los registros de variables meteorológicas de la estación meteorológica de la Hacienda El Prado ubicada en la ciudad de Sangolquí.

Usaremos una rama de la Estadística llamada Análisis de Datos Funcionales, que se puede considerar relativamente nueva, que actualmente se encuentra en desarrollo y con una amplia gama de aplicaciones.

En ésta usaremos en el área rama de la estadística en desarrollo, con la que se empelará un análisis descriptivo de la situación de la temperatura, con apoyo de Series temporales, que colaboraría en la verificación de nuestras hipótesis.

En lo referente al Análisis de Datos funcionales, se empleará software estadístico de distribución

libre libre y actualmente el software libre de mayor descargas R, y de manera especial e empleará la librería fda, desarrollada para éste fin por los profesores J. O. Ramsay, Hadley Wickham, Spencer Graves y Giles Hooker.

### **1.5.1 El propósito del estudio**

El propósito del estudio es presentar, describir y analizar las variables meteorológicas que influyen en el tiempo, de manera especial trabajar con las temperaturas mínima y máxima, de forma que las personas quienes puedan disponer de éste documento logren hacer conciencia que éste cambio climático es verdadero, y puedan poner su granito de arena a fin de contribuir a tomar medidas urgentes, y procurar detener éste aumento de la temperatura, a la vez sean portadores de un mensaje logren difundir acerca de éste tema.

### **1.5.2 El significado del estudio**

El estudio deja abierto las posibilidades para que otras personas puedan continuar, ampliar o servir de apoyo con otros proyectos, ya sea analizando otras variables o contrastando con otras estaciones meteorológicas, así como también para otras investigaciones relacionadas con la parte ambiental o agrícola, sin dejar de lado otras relacionadas con el Análisis de Datos Funcionales inclusive Series Temporales.

No se ha logrado encontrar otros estudios de éste tipo en el Ecuador, se puede considerar que es el primer estudio en su género, abriendo las puertas para otros estudios como se mencionó anteriormente.

## **1.6 Presunciones de los autores del estudio**

Hay dos presunciones básicas en el estudio la primera es la confiabilidad de los datos de las variables meteorológicas, los cuales han sido registrados diariamente manualmente o de manera analógica a lo largo de los catorce años y medio, y se ha cuidado que la información no haya sido alterada.

La segunda es la confiabilidad en el uso del software que se está empleando para el diseño de los modelos matemáticos y el cálculo de las medidas estadísticas descriptivas.

## **1.7 Supuestos del estudio**

Al tratarse de variables meteorológicas, el primer supuesto que se considerará es que éstas periódicas, cíclicas o recurrentes.

Suponemos que éstas van a oscilar dentro de un rango razonable y procurar eliminar los datos atípicos.

Suponemos que las variables son independientes

# Capítulo 2

## ANÁLISIS CON DATOS FUNCIONALES

### 2.1 Introducción

El Análisis de Datos Funcionales (ADF o FDA), es una parte de la Estadística relativamente nueva, desarrollada por los profesores Jim O. Ramsey y Bernard W. Silverman en el año de 1997, quienes introducen un nuevo concepto en el análisis de datos multivariado, este concepto consiste en hacer que los datos en vez de ser tratados como vectores, se los convierte en curvas o funciones viniendo de ahí su nombre. Las observaciones por lo general vienen de procesos estocásticos, por lo tanto dependen del tiempo que es una variable continua.

El ADF presenta algunas ventajas sobre el análisis clásico de Series de Tiempo, ya que para éste último es necesario que los intervalos de tiempo para los cuales fueron tomadas las mediciones de las observaciones deben ser equiespaciados, en otras palabras los intervalos de tiempo iguales, además deben ser procesos estacionarios, que para el ADF no es condicional.

Tradicionalmente el análisis de funciones básicamente ha sido estudiado por los matemáticos dentro de Análisis Funcional, en la actualidad el desarrollo tecnológico ordenadores cada vez con mayor rapidez de procesamiento y la mayor capacidad de almacenamiento, ha permitido

que se pueda procesar una mayor cantidad de datos y poder así lograr construir modelos más eficientes y eficaces, y así poder describir, clasificar, construir modelos de regresión y predecir fenómenos que anteriormente eran muy complicados estudiarlos.

El ADF rama relativamente nueva en el área de la Estadística permite estudiar y analizar la información contenida en funciones, curvas de superficie, o cualquier elemento que varíe en un continuo, así podemos usar las herramientas de la propia matemática y poder construir un conjunto de definiciones, teoremas, corolarios y herramientas que permitan estudiar de manera efectiva.

El análisis de datos funcionales tiene cuatro áreas de desarrollo, la descripción de la información, los modelos de regresión, el análisis de componentes principales y los modelos de clasificación, en éstas tres áreas se enfoca la mayoría de la investigación del ADF con aplicación en medicina (cardiología), finanzas, economía, meteorología como es nuestro caso de estudio.

En la práctica hay un número creciente de situaciones que provenientes de diferentes áreas de ciencias aplicadas (Ambiental, Biometría, Meteorología, Medicina, Química, Economía, etc.), en las cuales los datos que son recolectado son curvas. Sin embargo cuando estamos estudiando un fenómeno, podemos darnos cuenta que que existe un gran conjunto de variables, por ejemplo la temperatura del aire de un cierto lugar en la tierra depende de la posición geográfica (longitud, latitud y altitud), influencia de la tierra y los mares, etapa del año, la hora del día, la presión, el viento, la radiación solar, la humedad, entre otras.

Ventajosamente las herramientas tecnológicas, los nuevos sistemas informáticos, permiten el manejo de ésta información y poder establecer mejores modelos, así logramos observar el

desarrollo del Análisis de Datos Funcionales (FDA), actualmente existen algunas redes de investigación que trabajan en ésta área, tenemos por ejemplo el grupo de investigación en la Universidad de Granada, Universidad de Santiago de Compostela, Universidad de Cornell, Universidad de McGill, entre otras.

Con respecto a las referencias bibliográficas básicas tenemos los aportes de los profesores Ramsay y Silverman [11] con su libro *Functional Data Analysis*, posteriormente un libro de los mismos autores *Applied Functional Data Analysis: Methods and Case Studies* [12] en el cual se detallan algunos casos de estudio y el libro *Functional Data Analysis with R and MATLAB* de Ramsay J.O., y Hooker G. y Graves S [13] un texto que se usan dos programas muy empleados en las carreras de ingeniería como son los citados, sin dejar atrás el aporte los profesores Ferraty y Vieu [4] con su libro *Nonparametric Functional Data Analysis: Theory and Practice*, fuente indispensable para iniciar cualquier estudio de Análisis de Datos Funcionales.

## 2.2 Conceptos Básicos:

Definamos inicialmente que es variable funcional, un dato funcional es cualquier observación que varíe sobre un continuo, generalmente son receptados de forma manual como el caso del registro de la temperatura en un cierto lugar o automática en el mismo caso de la temperatura al hacerlo por medio de un equipo que registren los datos de manera digital en un cierto intervalo  $(t_{min}, t_{max})$ . Así, una observación puede expresarse mediante la familia  $X(t)_j, j = 1, \dots, J$ .

En la actualidad el desarrollo en el área de computación hacen posible crear una la rejilla cada vez más pequeña de tal forma podríamos considerar estas muestras como observaciones de la familia continua  $x = x(t); t \in (t_{min}, t_{max})$ . De tal manera que una variable funcional  $x$  se llama la variable funcional (*f.v.*) si toma valores en un espacio de dimensión infinita (o espacio funcional). Una observación  $x$  de  $x$  se llama dato funcional. [4]

Tenga en cuenta que, cuando  $x$  (resp.  $x$ ) denota una curva aleatorio (resp. Su observación), implícitamente hacemos la siguiente identificación  $x = x(t); t \in T$  (resp.  $x = x(t); t \in T$ ). En esta situación, la característica funcional viene directamente de las observaciones. La situación cuando la variable se está asociada una curva con un conjunto unidimensional  $T \subset \mathfrak{R}$ . Aquí, es importante señalar que la noción de la variable funcional cubre un área más grande que el análisis de curvas. En particular, una variable funcional puede ser una superficie aleatoria, como por ejemplo los niveles de gris de una imagen o un vector de curvas (y en estos casos  $T$  es un conjunto bidimensional  $T \subset \mathfrak{R}^2$ ), o cualquier otra más complicada dimensión infinita objeto matemático.

¿Qué son los conjuntos de datos funcionales?

Un conjunto de datos funcional  $x_1, x_2, \dots, x_n$  es la observación de  $n$  variables funcionales  $x_1, x_2, \dots, x_n$  idénticamente distribuidas como  $x_1$ . [4]

Esta definición abarca muchas situaciones, las más popular son las curvas de conjuntos de datos.

Aunque la mayoría de los ADF analizan observaciones que vienen de un proceso estocástico en tiempo continuo, hay aplicaciones que no necesariamente dependen del tiempo pero si en un espacio continuo.

## 2.3 Análisis Exploratorio de los Datos

En el análisis exploratorio de datos funcionales, lo primero que debemos hacer es construir la forma funcional de las curvas muestrales que resultan de las observaciones discretas

recolectadas. Éste proceso consiste en asumir una expansión de cada curva muestral en términos de una base de funciones y aproximar los coeficientes básicos utilizando un suavizado o interpolación.

En ésta sección trabajaremos con los promedios de las variables meteorológicas de cada día del año, es decir el promedio de todos los días primero de enero, tomados desde el 1 de enero de 1998, hasta el 31 junio del 2014, excluyendo todos los días 29 de enero en el caso de los años bisiestos. Por tanto trabajaríamos con un año tipo.

## **2.4 Configuración del Sistema de Bases para la Construcción de Funciones**

Para poder especificar el sistema de bases para las construcción de las funciones, se hará en dos etapas:

En primer lugar, definimos un conjunto de bloques de construcción funcionales  $\phi_k$  llamas funciones de base.

En segundo lugar, creamos un vector, matriz o arreglo de coeficientes para definir la combinación lineal función de las funciones de base.

Las funciones que queremos modelar suelen a caer en dos categorías principales: periódica y no periódica. Para el caso de las funciones periódicas el sistema base que se usa por lo general es de Fourier, y en el caso de las funciones no periódicas la elección habitual es el sistema de base spline o bsplines o las wavelets. Así también podemos encontrar otras bases como: constante, monómica (polinómica), exponencial, poligonal y potencial. Para éste estudio que

haremos énfasis en el Análisis de la temperatura nos centraremos en el uso de la base de Fourier,

Utilizamos un conjunto de bloques de construcción de funciones  $\phi_k, k = 1, \dots, K$  llamada funciones base, las cuales se combinan linealmente. Es decir, una función  $x(t)$  definida de esta manera se expresa en notación matemática como:

$$x(t) = \sum_{k=1}^K c_k \phi_k(t) = \mathbf{c}' \phi_{\mathbf{k}}(\mathbf{t})$$

llamada expansión de la función base. Los parámetros;  $c_1, c_2, \dots, c_K$  son los coeficientes de la expansión. La expresión de la matriz en el último término de la ecuación anterior utiliza  $c$  para representar el vector de  $K$  coeficientes y  $\phi$  denotar un vector de longitud  $K$  que contiene las funciones base.

Por lo general haremos la consideración que tenemos una muestra  $N$  de funciones, por lo tanto la expresión  $x_i(t) = \sum_{k=1}^K c_k \phi_k(t)$ , donde  $i = 1, 2, \dots, N$ , que en notación matricial viene dada por:

$$\mathbf{x}(\mathbf{t}) = \mathbf{C}\phi(\mathbf{t})$$

## 2.5 Series de Fourier para Datos Periódicos y Funciones

Como mencionamos, hay muchas funciones que se repiten cada cierto intervalo de tiempo, éste intervalo se conoce con el nombre de período  $T$ . Una serie de serie de Fourier viene dada por la suma de los siguiente términos:

$$\phi_1(t) = 1$$

$$\phi_2(t) = \text{sen}(\omega t)$$

$$\phi_3(t) = \text{cos}(\omega t)$$

$$\phi_4(t) = \text{sen}(2\omega t)$$

$$\phi_5(t) = \text{cos}(2\omega t)$$

$$\vdots = \vdots$$

donde la constante  $\omega$  conocida como frecuencia está relacionada con el período  $w = 2\pi/T$

$$x(t) = c_0 + c_1 \text{sen}(\omega t) + c_2(\text{cos}\omega t) + c_3 \text{sen}(2\omega t) + c_4(2\text{cos}\omega t) + \dots$$

Como podemos observar las funciones seno y coseno viene dispuestas en parejas, las cuales el argumento de las mismas están creciendo en uno cada vez.

En resumen para poder establecer la serie de Fourier debemos conocer dos cosas, la primera es el número de funciones y la segunda el período.

En la figura 2.1 que se muestra a continuación se tiene una selección de funciones base de Fourier, con diferente número de bases.

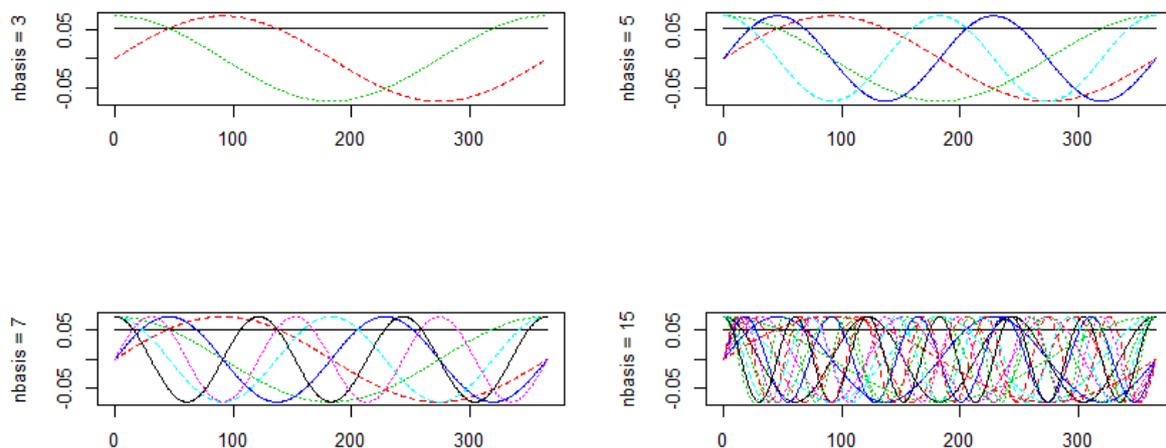


Figura 2.1: Sistemas Bases de Fourier

Sobre éstas base se hace la proyección y así generamos el modelo de regresión con bases de Fourier.

## 2.6 Suavizamiento mediante un Operador Diferencial Lineal

Un operador diferencial lineal identificado con su siglas Ldf, es una metodología de conseguir un modelo de datos funcionales que tengan la característica de poder calcula las derivas de dicha función. La suavización de éstas funciones se soporta usando la clase Lfd, que simplemente expresa un operador diferencial lineal, en el caso particular usaremos es el operador aceleración armónica que utilizaremos con funciones de base de Fourier para suavizar los datos periódicos.

La notación  $Lx$  se define como una aplicación de un operador diferencial lineal  $L$  a una función

de  $x$ .

Por ejemplo tenemos la aceleración,  $Lx = D^2x$ , o algo más sofisticado como la aceleración armónica  $L = \omega^2 D + D^3$ .

Un forma general se puede escribir de la forma:

$$Lx(t) = \beta_0(t)x(t) + \beta_1(t)Dx(t) + \dots + \beta_{m-1}(t)D^{m-1}x(t) + D^m x(t)$$

donde los conocidos coeficientes función operador diferencial lineal  $\beta_j(t), j = 0, \dots, m - 1$  son constantes o son funciones.

# Capítulo 3

## CASO DE ESTUDIO - SERIE DATOS DIARIOS

El primer caso que se trabajará es con el conjunto de datos de registros diarios desde enero de 1998 hasta junio de 2014, que corresponden a quince años seis meses. Inicialmente se hará una descripción de los datos de manera gráfica de las series temporales respectivas, sus medidas descriptivas y posteriormente usaremos el análisis de datos funcionales.

### 3.1 Exploración de Datos

Antes de empezar a realizar cualquier análisis, debemos primeramente tener una visión general de las variables que se desea estudiar, por lo tanto se presenta las series temporales de las variables meteorológicas en cuestión.

Al observar éstas series en la figura 3.1, podemos apreciar que la temperatura mínima y máxima tienen al parecer un comportamiento estacionario, la precipitación tiene un comportamiento de ondas wavelets, las demás variables es mucho más irregular , considerando además que hay una cantidad considerable de datos perdidos, visualmente podemos distinguir que en el caso de la humedad relativa hay una disminución de la variabilidad de la misma.

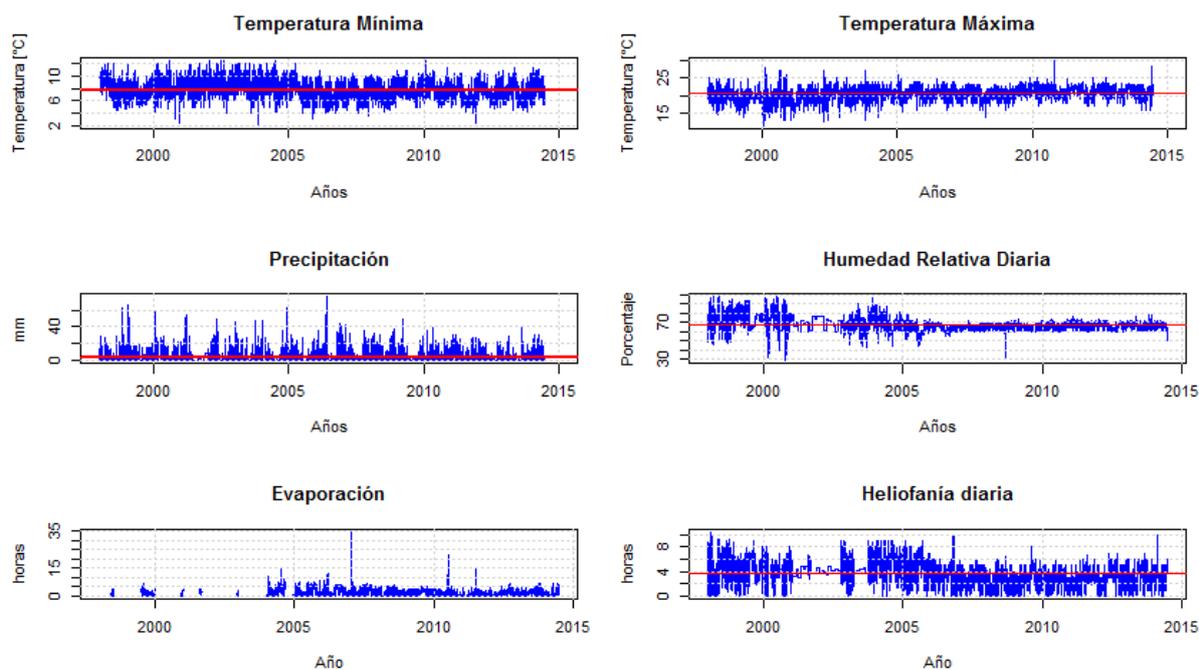


Figura 3.1: Series de variables meteorológicas diarias (1998 - 2014)

Dentro del análisis exploratorio inicial, la descripción gráfica de las series de tiempo es fundamental, ya que se puede a simple vista tener una idea de como varía la información proporcionada, sin embargo no se debe dejar de lado sus respectivas medidas descriptivas, las cuales se deja a continuación.

	Temp. Mínima	Temp. Máxima	Precip.	Evap.	Helio.	Hum. Relativa
Mínimo :	2.00	11.5	0.00	0.00	0.00	28.3
1st Qu. :	6.70	19.5	0.00	0.00	2.50	64.0
Mediana :	7.70	20.9	0.30	1.20	4.00	67.0
Media :	7.76	20.5	3.75	1.73	3.66	68.0
3rd Qu. :	9.00	22.0	5.21	3.00	4.60	71.0
Máximo :	12.50	29.8	76.00	34.55	10.50	98.3
Varianza :	2.27	3.13	41.94		3.27	48.46

Tabla 3.1: Resumen descriptivo de la base de datos diarios (1998 - 2014)

En el resumen estadístico de la tabla 3.1, podemos observar que en cada una de las variables meteorológicas hay una variabilidad bastante alta, hay varios datos atípicos, a pesar que la base

de datos fue examinada previamente y haber eliminado los datos que eran muy atípicos o que eran no lógicos, ésto se fue corrigiendo en el proceso de digitalización de los datos, al detectar éstas incongruencias.

## 3.2 Análisis de Datos Funcionales para registros diarios

Como se mencionó previamente las variables seleccionadas para éste estudio son la temperatura mínima y la temperatura máxima y como éstas se encuentran correlacionadas. Por tanto empezaremos estudiando la temperatura mínima diaria.

### 3.2.1 Análisis de Normalidad Temperatura Mínima

Un primer paso que hay que hacer en éste tipo de análisis es ver si los datos tienen una tendencia o comportamiento normal, para lo cual se ha decidido emplear el test gráfico QQ-plot, que compara los cuantiles reales con los cuantíles teóricos, y al observar que los puntos de comparación siguen una recta, podemos decir que los datos tienen una tendencia normal.

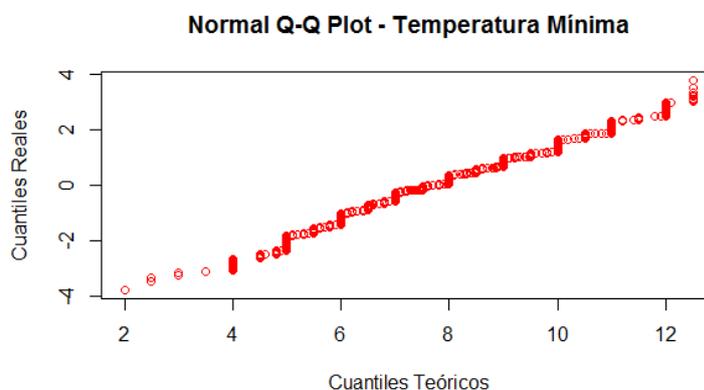


Figura 3.2: Test de Normalidad QQ - Plot Temperatura Mínima

Y como intuitivamente se esperaba y al ver la figura 3.2, podemos indicar que los datos de la temperatura mínima tienen una tendencia normal.

### **3.2.2 Regresión de la temperatura mínima sobre una base de Fourier**

Para realizar éste tipo de regresión, primeramente se debe seleccionar el tipo de bases que vamos a emplear para convertir los datos discretos a funciones o mejor datos funcionales. Al tratarse de una serie periódica se elige bases de Fourier, en el caso de las series no periódicas se recomienda usar splines y wavelets como puede ser el caso de la precipitación.

El número de bases o número de funciones base se recomienda que sea impar, como se trató en el capítulo 2, se hicieron varios intentos para encontrar el mejor modelo de regresión, logrando encontrar el óptimo con un número de bases igual a 169, para todo el rango desde el 1 de enero de 1998 al 31 de diciembre del 2013, en un total de 5110 días-

Luego de seleccionar la base se procedió a realizar un primer paso de aplicar un filtro a los datos, procurando eliminar los datos atípicos, para lo cual se eligió hacer éste suavizado con siete días hacia atrás y siete días adelante.

Una vez configurado el sistema base y los datos a los cuales se les iba a aplicar ésta base, se procedió hacer la regresión respectiva, y su resultado se puede mostrar en la figura 3.3 que se muestra a continuación.

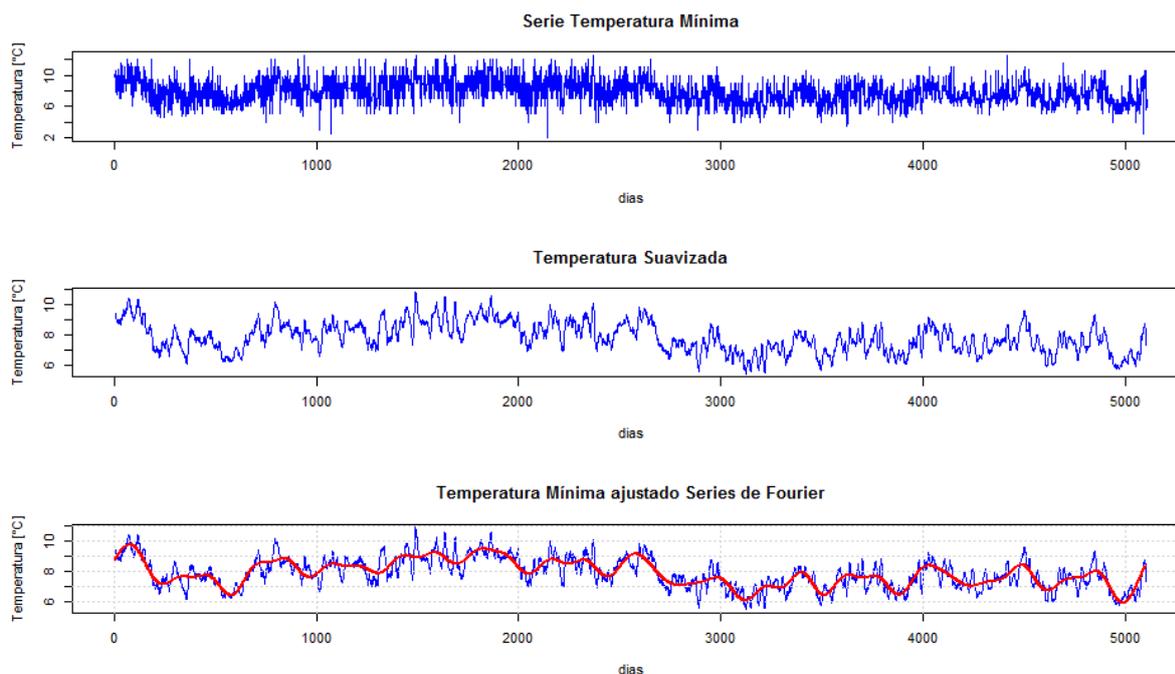


Figura 3.3: Modelo ajustado a una Base de Fourier de la Temperatura Mínima

Para éste modelo de regresión se obtuvo un valor del coeficiente de determinación de  $R^2 = 0,71$  con 5040 grados de libertad, un coeficiente  $F - statistic = 2.2e + 02$ , y un valor  $p - value < 2 \cdot 10^{-16}$ .

### 3.2.3 Modelo de ajuste usando Ldf para la temperatura mínima

Trabajaremos de igual manera con la misma base de datos, usaremos el operador diferencial lineal, el cual a parte de poder establecer un modelo de regresión, lo que podemos obtener es una función de datos funcionales, la cual aprovecha una de las características fundamentales del Análisis de Datos Funcional como es el de poder obtener las derivadas de estas funciones.

El la figura 3.4 se muestra el modelo calculado, su primera derivada y la segunda derivada de la temperatura mínima diaria.

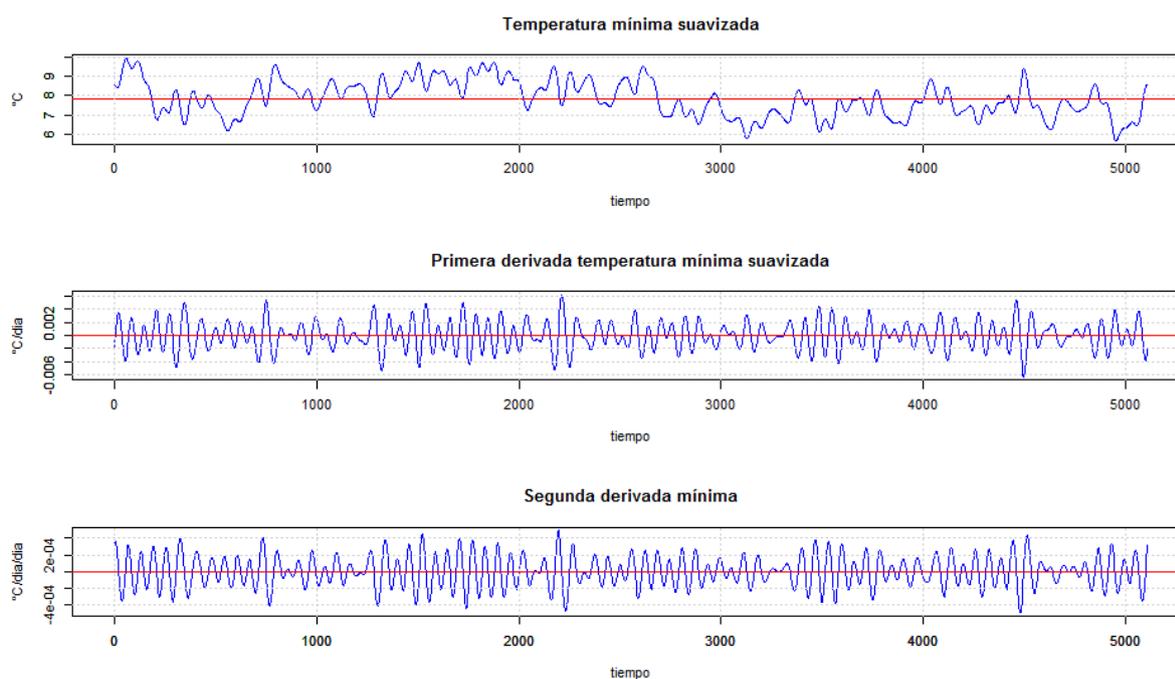


Figura 3.4: Modelo de ajuste de la temperatura mínima usando el operador diferencial lineal

De éste modelo entro otras características podemos obtener sus medidas descriptivas, las cuales con Las medidas descriptivas que se generan de éste modelo son:

Temperatura Mínima	
Mínimo :	2.00
Media :	7.83
Máximo :	12.50
Varianza :	2.32
Varianza :	1.53

Tabla 3.2: Resumen descriptivo modelo temperatura mínima usando Lfd

Ésta característica especial de los datos funcionales de los cuales puedo obtener las derivadas, ayuda para poder entender de mejor manera los fenómenos a estudiar, así tenemos que la primera derivada puede en física entenderse con la rapidez, en economía podría representar un costo marginal en el análisis marginal, o en cualquier otra rama de estudio una razón de cambio, igual la segunda derivada tiene varias aplicaciones.

De esa manera los datos simples y discretos se pueden convertir en datos funcionales, que además al tenerlos como datos funcionales hay un gran ahorro de memoria y de tiempo de procesamiento.

### 3.3 Análisis Temperatura Máxima

El análisis que se realizará para la temperatura máxima es similar al de la temperatura mínima, primeramente analizaremos la normalidad de los datos con el test QQ-plot.

#### 3.3.1 Análisis de Normalidad Temperatura Máxima

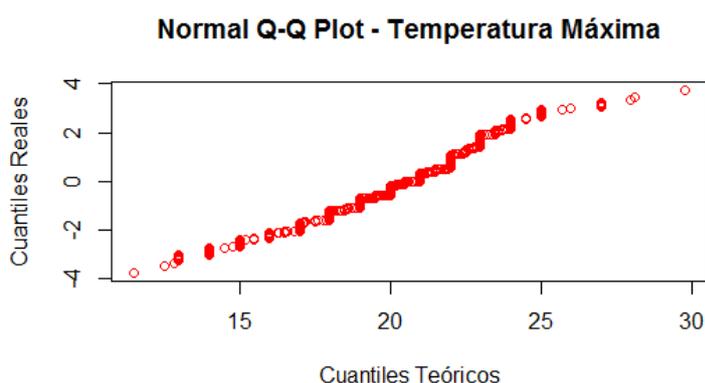


Figura 3.5: Test de Normalidad QQ - Plot Temperatura Máxima

Podemos seguir trabajando ya que se puede considerar a los datos de la temperatura máxima con tendencia normal.

#### 3.3.2 Regresión de la temperatura máxima sobre una base de Fourier

El modelo de regresión para la temperatura máxima será prácticamente una réplica del procedimiento realizado para la construcción del modelo de regresión para la temperatura

mínima, el la cual se eligió el mismo número de bases, debido a que posteriormente se deben comparar las dos temperaturas y ver si en realidad ha existido la variación de la temperatura que se planteó en la hipótesis.

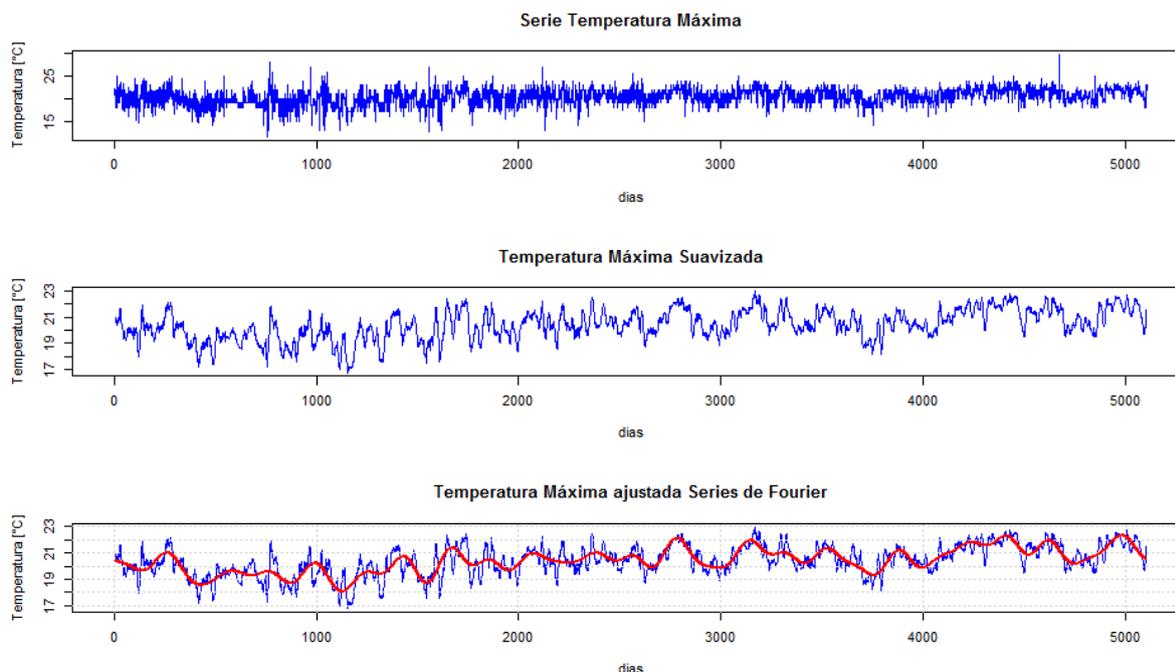


Figura 3.6: Modelo ajustado a una Base de Fourier de la Temperatura Máxima

El modelo generado tiene un coeficiente de determinación  $R^2 = 0,65$ , ligeramente menor al de la temperatura mínima, pero cabe aclarar que son 5040 grados de libertad, que hace que esto sea más confiable, un coeficiente de Fisher  $F - statistic = 1.7e+02$  y un  $p - value :< 2e - 16$ , que hace que se pueda decir que el modelo está muy bien ajustado.

### 3.3.3 Modelo de ajuste usando Ldf para la temperatura máxima

Repetiendo el mismo caso, realizamos el procedimiento como para la temperatura mínima, se puede ver el código en R en el anexo B, para ver a detalle todo el proceso para poder conseguir

transformar los datos a datos funcionales usando el operador diferencial lineal "Lfd".

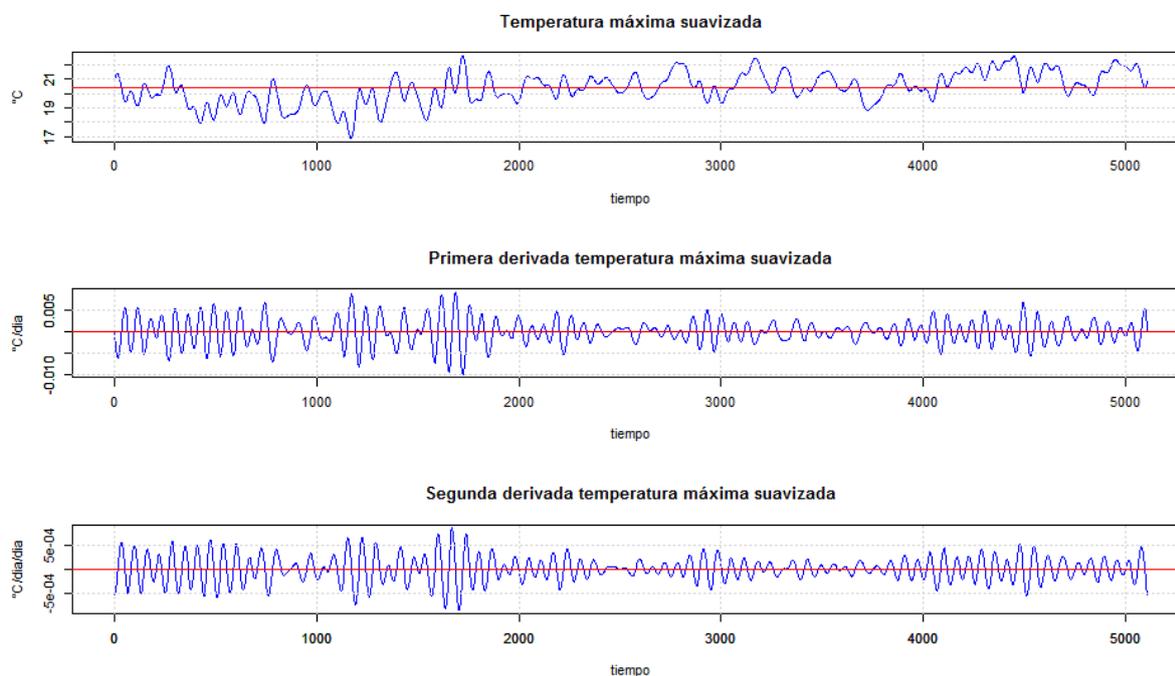


Figura 3.7: Modelo de ajuste de la temperatura máxima usando el operador diferencial lineal

Estos modelos son mucho más interesantes y aplicables, ya que se pueden generar otros modelos más sofisticados como los modelos con personalización, que no son motivo del presente estudio, pero que se deja al lector con la inquietud y pueda investigar mas a fondo.

### 3.4 Análisis de comparación entre temperaturas

Se ha visto la utilidad práctica del Análisis de Datos Funcionales para la generación de modelos de datos funcionales, pero el tema más importante de éste estudio es el poder verificar si la temperatura está cambiando, para lo cual se ha obtenido ayuda de las series de tiempo para concluir con ésta aseveración.

Primeramente presentemos algunos un par de gráficos estadísticos para poder visualizar la variabilidad de las variables meteorológicas en estudio.

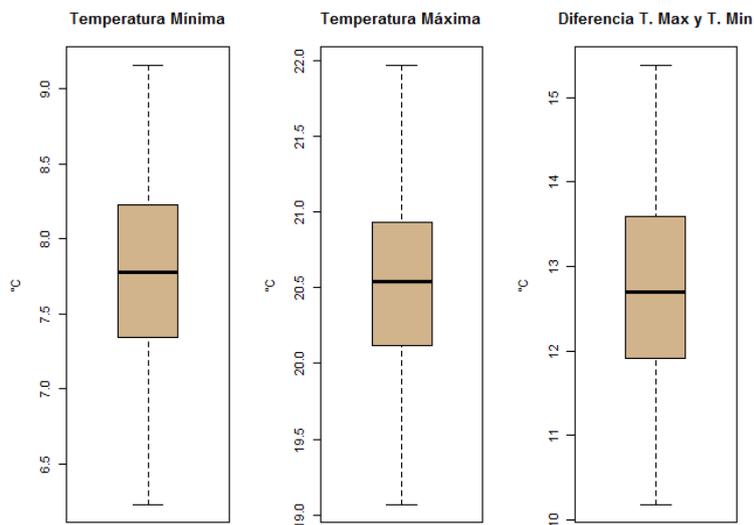


Figura 3.8: Diagramas de caja temperatura máxima y mínima promedio diaria

Al observar en los diagramas de cajas podemos apreciar la variabilidad de las temperaturas mínima, máxima y la diferencia, y se puede decir que el rango es amplio, sin embargo ésta no es concluyente ya que no me proporciona información en que instantes se producen las mayores diferencias.

Ahora emplearemos figura 3.9 que contiene cuatro gráficas, las superiores que corresponden a la componente de la tendencia de las series de tiempo de las temperaturas máxima y mínima, las mismas que resultan de la descomposición en sus respectivas componentes de las mismas.

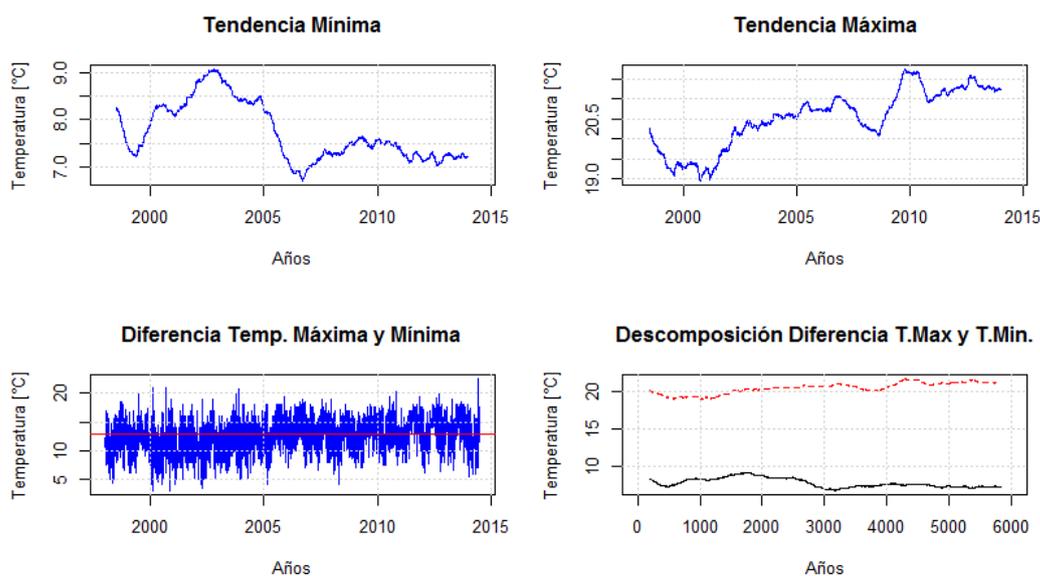


Figura 3.9: Comparación entre la temperatura máxima y mínima

En éstas gráficas podemos observar que el aumento de la temperatura máxima diaria es sostenido en el tiempo, mientras que la temperatura mínima en cambio tiene una tendencia a la baja.

En la gráfica inferior izquierda de la misma figura 3.9, se refiere a la serie de tiempo de la diferencia de temperatura entre la máxima y la mínima diaria, observando que ésta diferencia va aumentando paulatinamente.

Y por último se tiene la gráfica donde se encuentran las tendencias de la temperatura máxima y mínima observándose que la brecha entre las dos va aumentando, que dicha diferencia ha aumentado hasta en dos grados centígrados desde que se obtuvieron los primeros registros.

### 3.4.1 Componentes de la serie temporal de la diferencia de temperaturas

Nuevamente en éste análisis nos apoyamos en el análisis de series de tiempo para poder descomponer la serie generada de la diferencia entre las temperaturas máxima y mínima, en la que tenemos cada una de sus componentes, como la estacionariedad, la tendencia y el residuo.

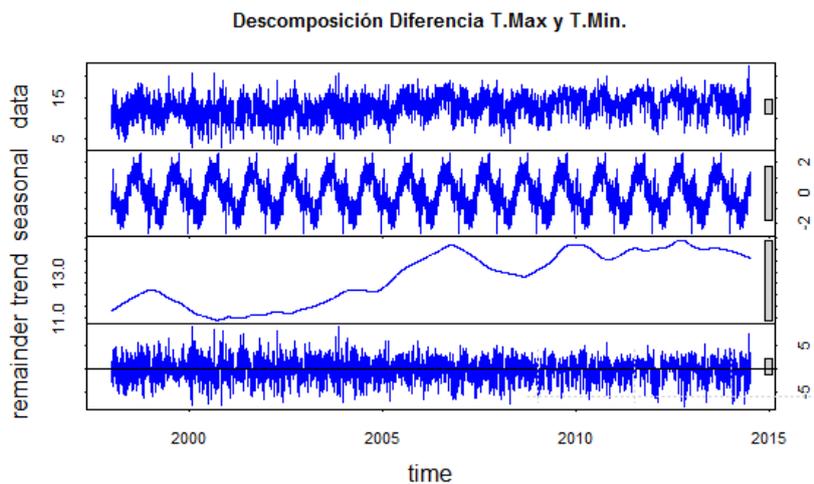


Figura 3.10: Componentes Serie temporal diferencia de temperaturas

En éste cálculo se puede apreciar mucho más claro que la tendencia a ir aumentando ésta diferencia es mayor, se puede decir que ahora los días en la Hacienda El Prado son más calurosos y las noches son más frías, un fenómeno que no está circunscrito a esa zona de estudio. Por lo cual es recomendable que se haga un estudio comparativo con otras estaciones meteorológicas.

### 3.4.2 Comparación entre las temperaturas usando los modelos generados

Ahora se calculará las diferencias entre los modelos sobre las bases de Fourier obtenidas previamente, posteriormente éste modelo se lo graficará sobre las diferencias de las series temporales.

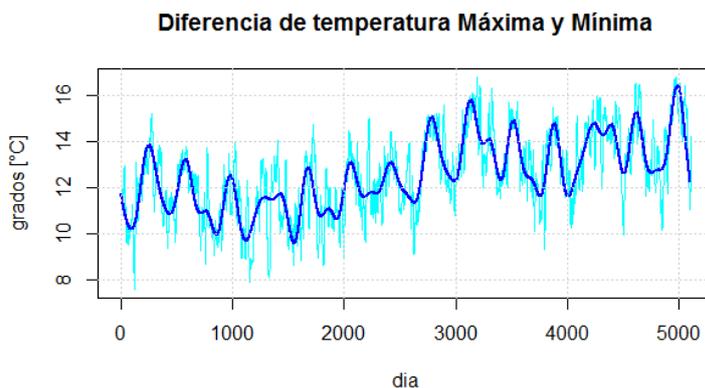


Figura 3.11: Diferencia de temperaturas entre modelos ajustados

Nuevamente encontramos con lo que hemos visualizado perviamente, la diferencia entre las temperaturas máximas y mínimas diarias está aumentando, su variabilidad es mayor.

Las consecuencias de éste aumento de temperatura son gravísimas, se altera completamente el ecosistema, las plantas y animales empiezan a emigrar a lugares más altos, debido a que buscan su hábitat que han vivido por millones de años, el problema es cuando dichas plantas no encuentran lugares más altos, éstas tienden a desaparecer, devoradas por el cambio climático o por las plagas que no existían por condiciones climáticas y ahora les afecta, sin considerar los demás efectos, como los deshielos de los nevados, y los glaciares, etc.

# Capítulo 4

## CASO DE ESTUDIO - SERIE DATOS PROMEDIOS

En éste segundo caso emplearemos los datos promedios diarios, se formaría lo que se llamaría un año típico o promedio, estos se obtuvieron sacando el promedio de cada fecha del año de los años que se disponen, así la primera observación corresponde al promedio de las variables meteorológicas de todos los primeros de enero de los años en estudio.

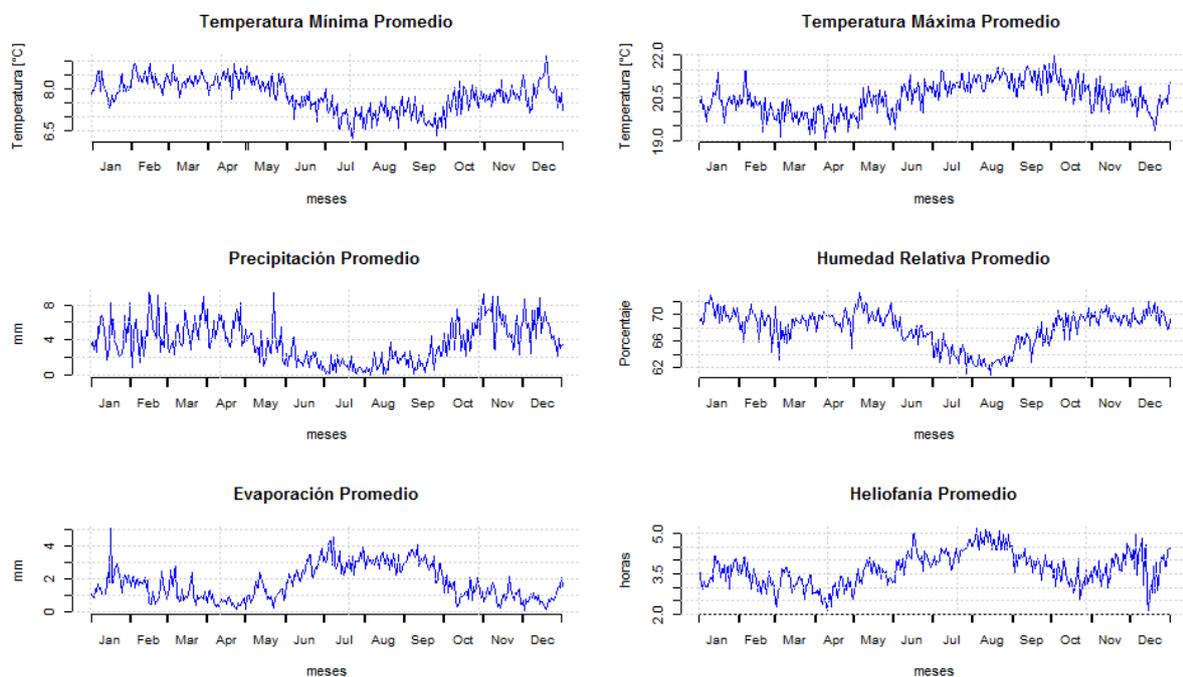


Figura 4.1: Series variables meteorológicas promedio diarias

## 4.1 Análisis Exploratorio de Datos

El análisis de igual manera que en el capítulo 3, empezaremos con una análisis exploratorio de las variables meteorológicas. Primeramente se presentará el gráfico de las series temporales generadas

Los profesores J. O. Ramsay, Spencer Graves y Giles Hooker usan un ejemplo similar a éste en el libro *Functional Data Analysis with R and MATLAB*, páginas 52 y 53.

El objetivo de crear un año típico o promedio es poder analizar si existe una correlación entre cada una de las variables estudiadas.

	Temp. Mínima	Temp. Máxima	Precip.	Evap.	Helio.	Hum. Relativa
Mínimo :	6.23	19.1	0.00	0.06	2.11	60.9
1st Qu.:	7.34	20.1	1.84	0.84	3.21	66.4
Median :	7.78	20.5	3.38	1.53	3.69	68.9
Media :	7.76	20.5	3.70	1.74	3.69	68.0
3rd Qu.:	8.23	20.9	5.32	2.55	4.14	70.0
Máximo :	9.16	22.0	9.43	5.02	5.19	73.4

Tabla 4.1: Resumen de la información con la serie con los promedios por día

## 4.2 Ajuste de modelo con bases Fourier

En ésta sección se han generado los modelos de regresión de todas las variables meteorológicas con bases de Fourier, con cinco funciones base, obteniéndose los siguientes resultados:

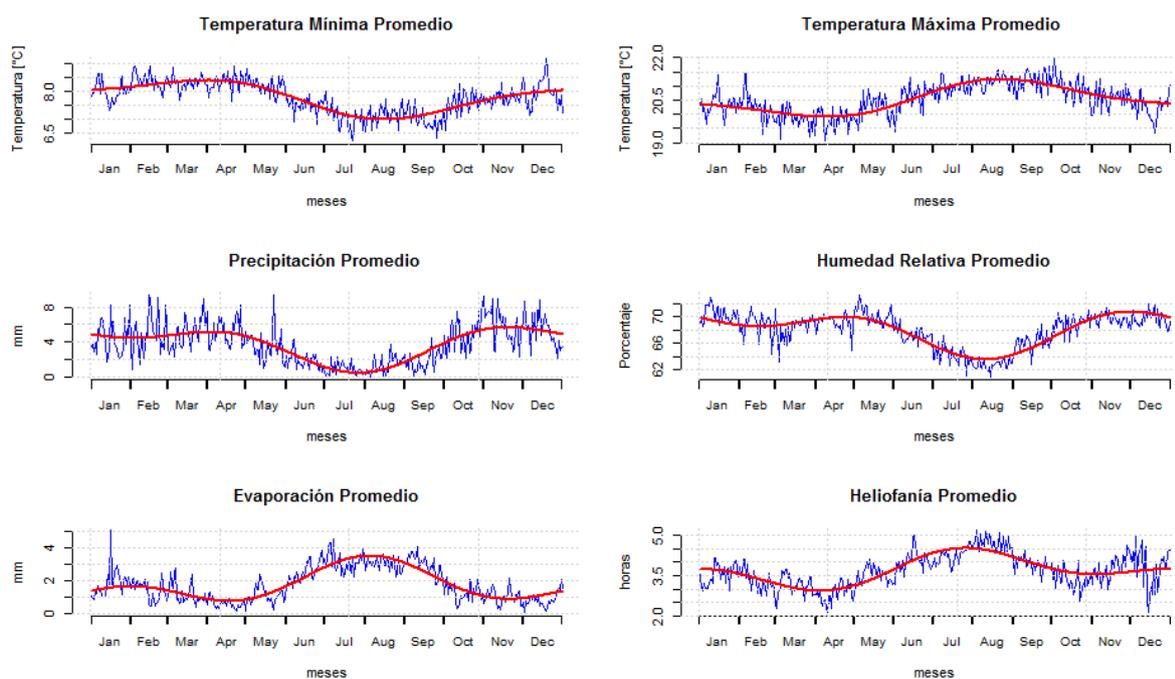


Figura 4.2: Modelo ajustado con series de Fourier - 5 funciones

	$R^2$ .	F-statistic.	p-value.
Temp. Mínima :	0.66	1.7e+02	2e-16
Temp. Máxima :	0.59	1.3e+02	2e-16
Precipitación :	0.55	1.1e+02	2e-16
Humedad Relativa :	0.72	2.3e+02	2e-16
Evaporación :	0.70	2.1e+02	2e-16
Heliofanía :	0.54	1.1e+02	2e-16

Tabla 4.2: Resumen de la regresión con cinco funciones base

Ahora repetiremos el cálculo con diez funciones base y sus gráficas de respuesta son::

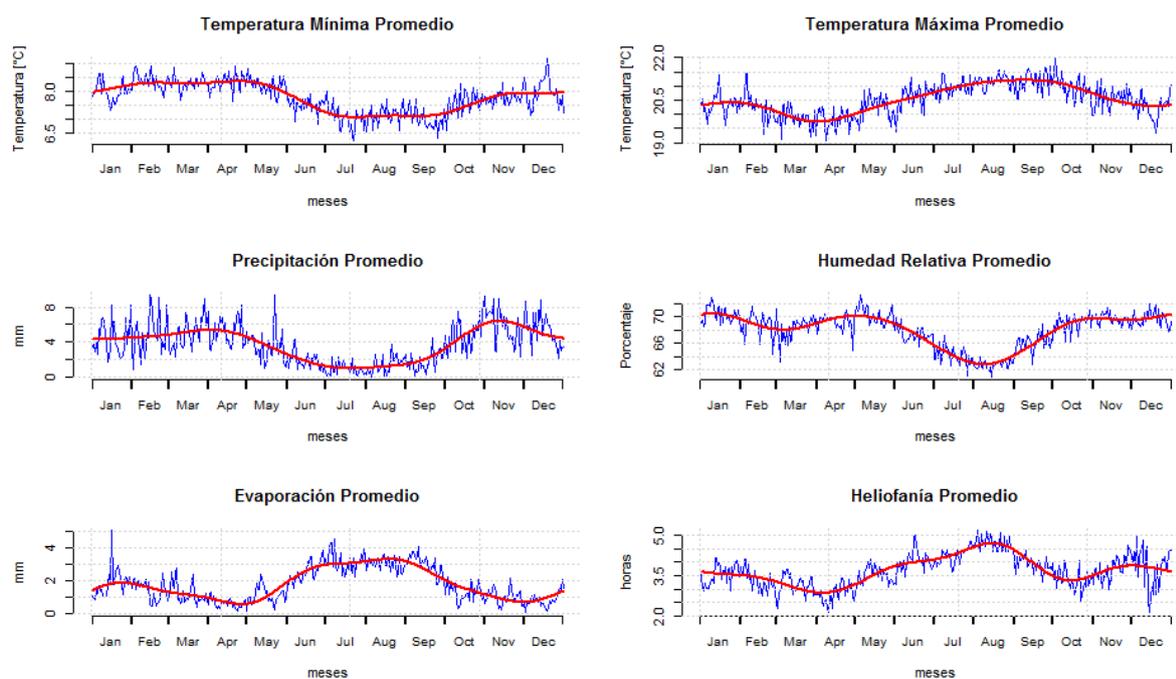


Figura 4.3: Modelo ajustado con series de Fourier - 10 funciones

	$R^2$ .	F-statistic.	p-value.
Temp. Mínima :	0.68	82	2e-16
Temp. Máxima :	0.62	65	2e-16
Precipitación :	0.58	54	2e-16
Humedad Relativa :	0.78	1.4e+02	2e-16
Evaporación :	0.73	1.1e+02	2e-16
Heliofanía :	0.61	61	2e-16

Tabla 4.3: Resumen de la regresión con diez funciones base

Podemos darnos cuenta que los dos modelos están muy bien validados, en el primer caso tendríamos una mejor calificación de acuerdo al test de Fisher, pero en el segundo modelo estamos mejor con el coeficiente de determinación. Con respecto al p-valor todos los valores son iguales.

Para decidir cuál es el mejor modelo si los dos reflejan la misma respuesta, diría que sería mejor manejar con el menor número de funciones.

### 4.3 Modelo de ajuste usando Ldf para la temperatura promedio mínima

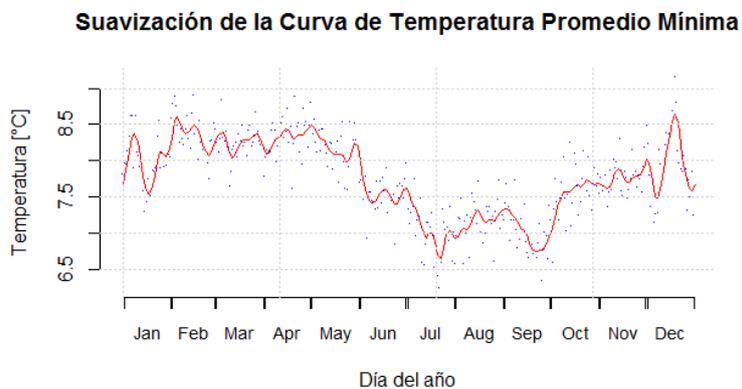


Figura 4.4: Modelo ajustado con Ldf - Temperatura Mínima

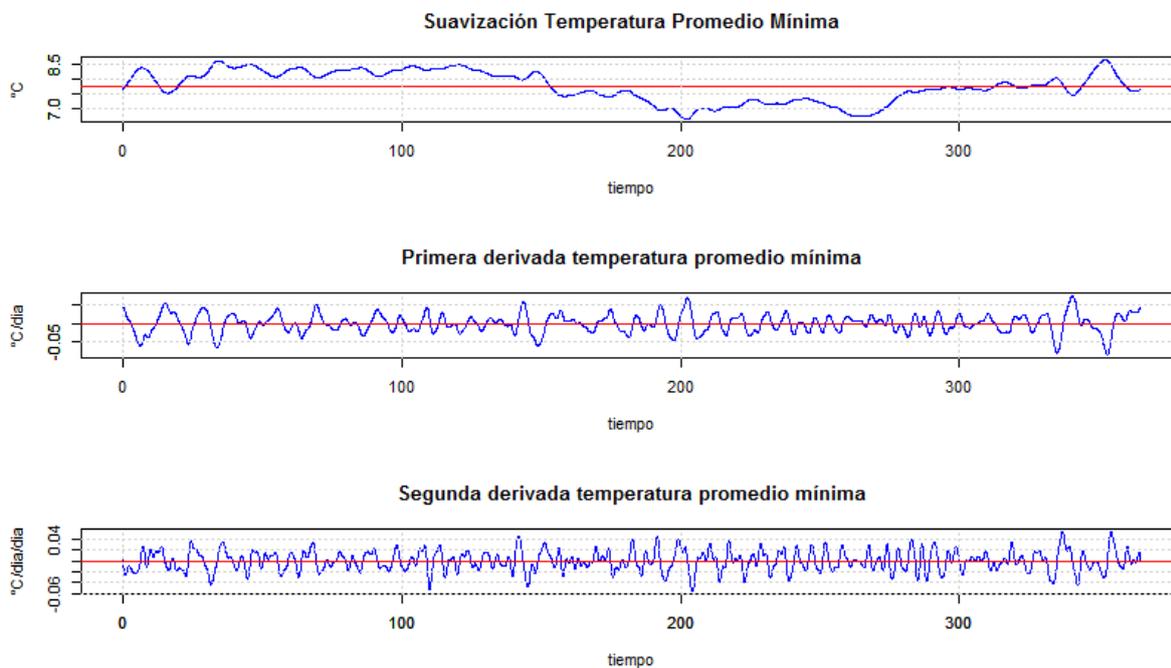


Figura 4.5: Modelo ajustado con Ldf - Temperatura Mínima

## 4.4 Modelo de ajuste usando Ldf para la temperatura promedio mínima

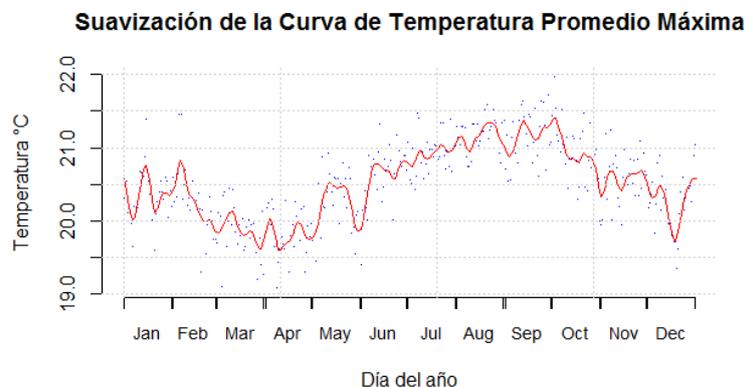


Figura 4.6: Modelo ajustado con Ldf - Temperatura Mínima - Temperatura Máxima

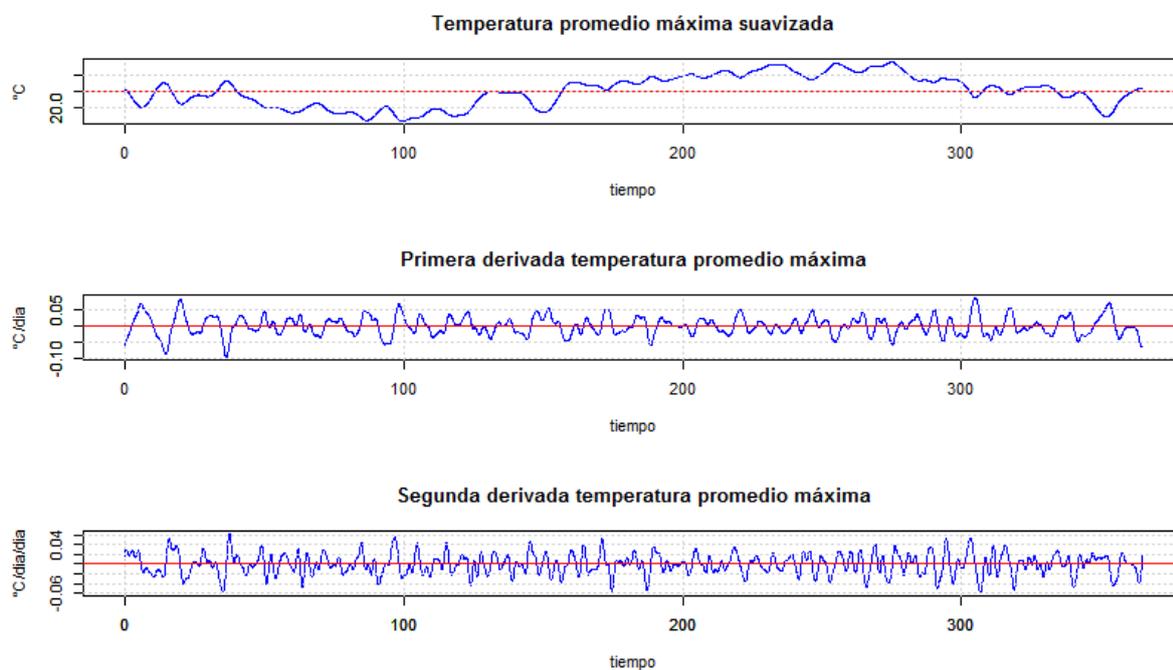


Figura 4.7: Comparación modelos ajustado Ldf - Temperatura máxima

## 4.5 Correlación entre las variables

### 4.5.1 Correlación entre temperaturas

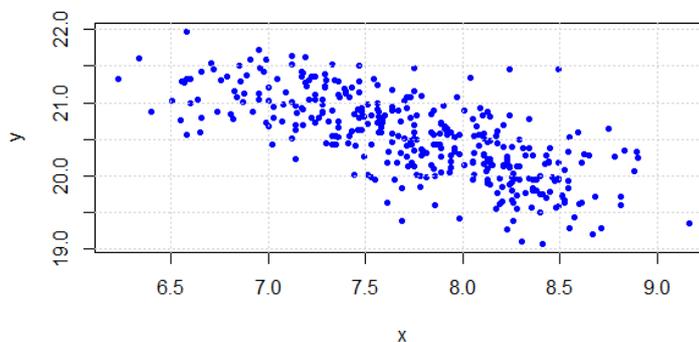


Figura 4.8: Comparación modelos de la temperatura máxima y mínima ajustado Ldf]

Podemos ver que la temperatura mínima y máxima tiene una relación inversa.

### 4.5.2 Matriz de Correlación entre las variables

Con ésta serie de datos generada, se puede empelar para sacar algunas conclusiones.

	T.Minima	T. Maxima	Precip.	Evap.	Heliofania	H.Relativa
T. Minima	1.00	-0.72	0.52	-0.66	-0.60	0.63
T .Maxima	-0.72	1.00	-0.46	0.58	0.61	-0.55
Precip.	0.52	-0.46	1.00	-0.72	-0.51	0.62
Evap.	-0.66	0.58	-0.72	1.00	0.62	-0.70
Heliofania	-0.60	0.61	-0.51	0.62	1.00	-0.60
Hum. Relativa	0.63	-0.55	0.62	-0.70	-0.60	1.00

Tabla 4.4: Matriz de Correlación de la serie con datos promedios por día

Podemos observar que la correlación entre la temperatura máxima y mínima es relativamente alta, pero a la vez inversa.

Otras conclusiones que la humedad relativa y la Evaporación son inversa también y es muy lógico, ya que si la humedad relativa es alta, es muy poco probable que se sude, ya que estaría saturado el aire.

# Capítulo 5

## LIMITACIONES DEL ESTUDIO

Todos los proyectos de investigación traen consigo limitaciones en el estudio del mismo, en el presente trabajo las limitantes más sobresalientes fueron:

La información ha sido una de las limitantes más fuertes, a pesar de la colaboración del Ing. Marcelo Arce, responsable de la Estación Meteorológica, la mayor parte de la información se encontraba en papel y no en forma digital, para lo cual se procedió a digitar los registros existentes de cada una de las variables meteorológicas.

La segunda limitante que se considera es que se carece de información completa de datos de algunas variables, razón por la que no se puede considerar esas variables en el estudio de los modelos.

El Análisis de Datos es una rama de la Estadística relativamente nueva que se encuentra en desarrollo, la información es limitada, hay muy poca información detallada acerca de los procesos que se efectúan, y de la literatura revisada se centran en el uso del software.

Con en todo proyecto el factor tiempo es una limitante bastante fuerte.

# Capítulo 6

## RECOMENDACIONES PARA ESTUDIOS FUTUROS

Se recomienda continuar con el estudio de las otras variables en especial la variable precipitación usando el método de las wavelets, ya que es un modelo recomendado para variables no periódicas.

Otra recomendación muy importante es replicar éste análisis con otras estaciones cercanas a la zona en estudio, como la de Machachi, Tumbaco y Quito, y poder tratar de establecer una comparación e influencia entre las mismas.

A pesar que existe software empleado por el INAMHI, para poder pronosticar a corto, mediano y largo plazo, no hay estudios detallados como el que se ha desarrollado en éste estudio, en especial para un área de impacto.

Sería muy conveniente se desarrolle un curso de Análisis de Datos Funcionales en el país, ya que no hay muchos docentes que conozcan del tema u otros profesionales.

Hacer hincapié sobre el cambio climático, en estos quince últimos años la temperatura se ha incrementado en dos grados centígrados.

Completar este estudio utilizando penalización con rugosidad.

Otra componente que se podría trabajar es el análisis de componente principales.

# **Capítulo 7**

## **REFERENCIAS BIBLIOGRÁFICAS**

# Bibliografía

- [1] Ayllón T. (2003). Elementos de Meteorología y Climatología. Trillas.
- [2] Gonzáles L.B. (2010). Nociones de Análisis Funcional. Departamento de Análisis Matemático - Universidad de Sevilla.
- [3] Cowpertwait P.S. y Metcalfe A.V. (2009). Introductory Time Series with R. Springer-Verlag.
- [4] Ferraty, F. and Vieu, P. (2006). Nonparametric Functional Data Analysis: Theory and Practice. Springer-Verlag.
- [5] Machiwal D. y Jha M.K. (2012). Hidrologic Time Series Analysis: Theory and Practice. Springer-Verlag
- [6] Mudelsee M. (2010). Climate Time Series Analysis: Classical Statistical and Bootstrap Methods. Springer-Verlag.
- [7] Payá R. (2010). Apuntes de Análisis Funcional. Departamento de Análisis Matemático - Universidad de Granada.
- [8] Posso M. (2011).Proyectos, Tesis y Marco Lógico: Planes e Informes de Investigación. IEPI
- [9] Pruscha H. (2013). Statistical Analysis of Climate Series: Analyzing, Plotting, Modeling and Predicting with R. Springer-Verlag.

- [10] Rincón L. (2014). *Introducción a la Probabilidad*. Ciudad Universitaria - UNAM.
- [11] Ramsay J.O., y Silverman B.W. (1997, 2005). *Functional Data Analysis (First and Second Edition)*. Springer-Verlag.
- [12] Ramsay J.O., y Silverman B.W. (2002). *Applied Functional Data Analysis: Methods and Case Studies*. Springer-Verlag.
- [13] Ramsay J.O., y Hooker G. y Graves S. (2009). *Functional Data Analysis with R and MATLAB*. Springer-Verlag.
- [14] Shumway R.H. y Stoffer D.S. (2010). *Time Series Analysis and its Applications: With R examples. (Third Edition)*. Springer-Verlag.
- [15] Valderrama M.J., Aguilera A.M. y Ocaña F.A. (2000). *Predicción Dinámica Mediante Análisis de Datos Funcionales*. Ed. La Muralla.

# Capítulo 8

## ANEXOS

### 8.1 ANEXO A: DEFINICIÓN DE TÉRMINOS

#### **Meteorología:**

La palabra meteorología proviene de las raíces griegas meteoros :elevado en el aire, y logos; estudio. La meteorología estudia las causas , estructura, naturaleza, evolución y relación de los fenómenos atmosféricos, así como las condiciones dominantes en un periodo corto que es llamado tiempo meteorológico.

#### **Climatología:**

La climatología es una ciencia que requiere de la información de cada una de las variables meteorológicas de grandes periodos para obtener valores medios, normales, frecuencia, etc., que permitan detectar y aprecia la interrelación de los fenómenos atmosféricos con la geografía del lugar y de esta manera, determinar el clima.

La meteorología y climatología son disciplinas que comparten el mismo objeto de estudio, la atmósfera; sin embargo, difieren en cuanto al objeto formal; esto es, su orientación y método son distintos

**Atmósfera:**

La atmósfera es la envoltura gaseosa que rodea a la tierra el nombre proviene de las raíces griegas atmos = vapor de agua y sphaira = esfera. La atmósfera es una mezcla gaseosa no una combinación , ya que sus componentes no reaccionan químicamente entre si.

**Radiación solar:**

Medida de la radiación solar

**Heliofanía:**

Es la medida en horas de la radiación solar o también llamado brillo solar, para registrar las horas de insolación en un lugar determinado se usan los heliógrafos ; el mas usado es el heliógrafo de Campbell Stokes.

**Temperatura del aire:**

Es el grado sensible de calor y se debe principalmente a la radiación calorífica de onda larga que emite la superficie del planeta. Comúnmente se le denominará Temperatura, siendo registradas por lo general la mínima y la máxima del día.

**Presión atmosférica:**

Una manera muy sencilla de definir la presión atmosférica es el peso de una columna de aire que se extiende desde un plano horizontal donde se mide, hasta el límite exterior de la atmósfera. Ésta se mide con los barómetros.

**Humedad absoluta del aire:**

Es la cantidad de gramos de vapor existente en un metro cubico de aire. Por lo tanto la humedad absoluta constituye la densidad del vapor de agua existente en el aire.

**Humedad específica del aire:**

Es la cantidad de gramos de vapor de agua contenidos en un kilo gramo de aire húmedo; es decir, en una mezcla de aire seco y vapor de agua.

**Humedad relativa del aire:**

Es la relación, expresada en porcentaje, entre la cantidad de vapor de agua realmente existente en la atmósfera y la que existiría si el aire estuviera saturado a la misma temperatura. Esta relación se obtiene a partir de la tensión de vapor actual y la tensión máxima de saturación. La humedad relativa del aire se calcula por medio de psicrómetro, el cual consta de dos termómetros; uno llamado del bulbo seco o ambiente y otro de bulbo húmedo.

**Evaporación:**

La evaporación es el proceso de cambio de estado de cualquier sustancia del estado líquido al estado gaseoso, al agua le favorece la rapidez de evaporación varios las elevadas temperaturas, el aire muy seco y los vientos fuertes.

**Precipitación:**

Es la caída de agua de las nubes, en estado líquido o sólido para que se produzca es necesario que las pequeñas gotas de las nubes se unan para formar gotas más grandes que venzan la resistencia del aire y puedan caer. Para medir la cantidad de precipitación se usan los pluviómetros

**Una variable aleatoria funcional:**

Es una variable aleatoria  $X$  definida en un espacio de probabilidad  $(\Omega, A, P)$  es funcional si toma valores en un espacio funcional, es decir en un espacio normado o seminormado completo.

**Espacio de Hilbert:**

Es un espacio vectorial dotado de una estructura métrica, el cual es un espacio normado, con norma completa asociada a su producto escalar, determinada por la igualdad  $\|x\|^2 = (x|x)$ .

**Espacio Normado:**

Es un espacio vectorial  $X$  dotado de una norma  $\|\cdot\|$

**Espacio de probabilidad:**

Es una terna  $(\Omega, F, P)$  en donde  $\Omega$  es un conjunto arbitrario,  $F$  es una  $\sigma$ -álgebra de subconjuntos  $\Omega$  y  $P$  es una medida de probabilidad definida sobre  $F$ .

**8.1.1 Definiciones en Análisis de Datos Funcionales:****Producto Escalar:**

El producto escalar euclidiano de elementos  $x$  e  $y$  del espacio vectorial  $E$ , denotado como  $\langle x, y \rangle$ , se define como:

$$\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathfrak{R}$$

$$(x, y) \rightarrow \langle x, y \rangle$$

El producto escalar se puede interpretar como medida de cantidad de asociación entre dos elementos del espacio vectorial en el que trabajamos, también como la medida de la proyección de un vector sobre el otro.

**Norma:**

La norma de un elemento  $x$  del espacio, se define como:  $\|x\| = \sqrt{\langle x, x \rangle}$

**Ortogonalidad entre dos elementos:**

Dos elementos  $\langle x, y \rangle$  del espacio son ortogonales si:  $\langle x, y \rangle = 0$

**Ángulo entre dos elementos:**

El ángulo  $\theta$  entre los elementos  $x$  e  $y$  lo definimos como:  $\theta = \arccos \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}$

**Distancia entre dos elementos:**

La distancia entre los elementos  $x$  e  $y$  se define como:  $d_{x,y} = \|x - y\| = \sqrt{\langle x - y, x - y \rangle}$

**Media:**

La media se define como la cantidad obtenida que resulta al operar producto escalar:

$$\bar{x} = \frac{1}{N} \langle x, \mathbf{1} \rangle, \text{ donde: } \mathbf{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \text{ y } N = \langle \mathbf{1}, \mathbf{1} \rangle = \|\mathbf{1}\|^2.$$

**Varianza:**

La varianza se obtiene al operar:

$$S_x^2 = \frac{1}{N} \cdot \langle x - \bar{x} \cdot \mathbf{1}, x - \bar{x} \cdot \mathbf{1} \rangle = \frac{1}{N} \|x - \bar{x} \cdot \mathbf{1}\|^2,$$

donde:  $\bar{x} \cdot \mathbf{1} = (\bar{x}, \bar{x}, \dots, \bar{x})^t$

**Covarianza:**

La covarianza entre los elementos  $x$  e  $y$  se obtiene al operar:

$$S_{x,y} = \frac{1}{N} \cdot \langle x - \bar{x} \cdot \mathbf{1}, y - \bar{y} \cdot \mathbf{1} \rangle$$

### Correlación:

La correlación entre  $x$  e  $y$  se define como la cantidad que resulta del producto escalar:

$$r_{x,y} = \frac{S_{x,y}}{S_x S_y} = \frac{\langle x - \bar{x} \cdot \mathbf{1}, y - \bar{y} \cdot \mathbf{1} \rangle}{\|x - \bar{x} \cdot \mathbf{1}\| \cdot \|y - \bar{y} \cdot \mathbf{1}\|} = \cos(x - \bar{x} \cdot \mathbf{1}, y - \bar{y} \cdot \mathbf{1})$$

La correlación entre  $x$  e  $y$  es la proporción de la magnitud explicada por la relación de direcciones entre  $x$  e  $y$  con respecto al total de la magnitud de estos dos elementos ( $S_x \cdot S_y$ )

## 8.1.2 Espacio de Funciones de Cuadrado Integrable:

Sea el espacio de funciones de cuadrado integrable:

$$L^2 = \{f : \mathfrak{R} \rightarrow \mathfrak{R} : \int_{\mathfrak{R}} f^2(t) dt < \infty\}$$

El espacio de funciones de cuadrado integrable  $L^2$  es un espacio de Hilbert.

### Producto escalar euclidiano en $L^2$ :

El producto escalar euclidiano en  $L^2$  se define como:

$$\langle x, y \rangle = x'y = \int x(t)y(t)dt$$

### Producto Escalar General en el Espacio $L^2$ :

El producto escalar general en el espacio  $L^2$  se define como:

$$\langle x, y \rangle_W = x'y = \int w(t)x(t)y(t)dt \quad \text{donde } w(t) \in L^2$$

El producto escalar general pondera a unos intervalos de  $t$  con mayor peso y disminuir a otros por medio de una función  $w(t)$ .

### Norma en $L^2$ :

La  $L^2$ -norma de  $x(t)$  se define como:  $\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{\|x\|^2} = \sqrt{\int x(t)^2 dt}$  Al calcular la norma nos proporciona la magnitud del elemento  $x(t)$  dentro del espacio que se opera.

### Ángulo entre Funciones:

El ángulo  $\theta$  entre las funciones  $x(t)$  e  $y(t)$  se define como:

$$\theta = \arccos \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} = \frac{\int x(t)y(t)dt}{\sqrt{\int x(t)^2 dt \cdot \int y(t)^2 dt}}$$

### Ortogonalidad entre Funciones:

Dos funciones  $x(t)$  e  $y(t)$  son ortogonales si  $\langle x, y \rangle = \int x(t) \cdot y(t) dt = 0$

### Media de la Función:

La media de la función  $x(t)$  o llamada también valor medio de la función  $x(t)$ , se define como:

$$\bar{x} = \frac{1}{T} \langle x, \mathbf{1} \rangle = \frac{1}{\int_0^T \mathbf{1}(t)^2 dt} \cdot \int_0^T x(t) \cdot \mathbf{1}(t) dt$$

Donde la función  $\mathbf{1}(t) = 1 \forall t \in \mathfrak{R}$

La media de  $x(t)$  ( $\bar{x}$ ) va a representar la tendencia central y el nivel medio de todos los valores de  $x(t)$ .

### **Función Valor Medio:**

La función valor medio de  $x(t)$ , que denotaremos por  $x \cdot \mathbf{1}(t)$ , es aquella cuyo valor  $\forall t \in [0, T]$  es el de la media de  $x(t) = \bar{x}$ .

Esta definición va a tener sentido a la hora de expresar la varianza y la covarianza de una función.

Hay que observar que  $(\bar{x} \in \mathfrak{R})$  es diferente al concepto de función valor medio.

### **Propiedad**

Sea  $x(t)$  una función continua definida en el intervalo  $[a, b]$  y sea  $\bar{x}$  su valor medio. Definimos ahora la función valor medio  $f_{\bar{x}}(t) = \bar{x} \forall t \in [a, b]$ . Entonces se cumple que:

$$\int_a^b x(t) dt = \int_a^b f_{\bar{x}}(t) dt$$

Es decir, el área que encierra la función  $x(t)$  es la misma que el área que encierra la función de valor constante igual al valor medio de la función en el intervalo  $[a, b]$ .

### **Varianza de una función:**

La varianza de  $x(t)$ , que denotamos con  $S_{x(t)}^2$  se define como:

$$S_{x(t)}^2 = \frac{1}{T} \langle x - \bar{x} \cdot \mathbf{1}, x - \bar{x} \cdot \mathbf{1} \rangle = \frac{1}{\int_0^T dt} \cdot \int_0^T (x(t) - \bar{x} \cdot \mathbf{1}(t))^2 dt$$

Esta cantidad representa la variación cuadrática media de todos los valores de la función respecto a su valor medio.

**Covarianza de una función:**

La covarianza entre dos funciones  $x(t)$  e  $y(t)$ , que denotamos con  $S_{x(t),y(t)}$  se define como:

$$S_{x(t),y(t)} = \frac{1}{T} \langle x - \bar{x} \cdot \mathbf{1}, y - \bar{y} \cdot \mathbf{1} \rangle = \frac{1}{\int_0^T dt} \cdot \int_0^T (x(t) - \bar{x} \cdot \mathbf{1}(t)) \cdot (y(t) - \bar{y} \cdot \mathbf{1}(t)) dt$$

**Correlación de una función:**

La correlación entre dos funciones  $x(t)$  e  $y(t)$ , se denota con  $r_{x(t),y(t)}$ , y se define como:

$$r_{x(t),y(t)} = \frac{S_{x(t),y(t)}}{S_{x(t)}S_{y(t)}} = \frac{\frac{1}{T} \langle x - \bar{x} \cdot \mathbf{1}, y - \bar{y} \cdot \mathbf{1} \rangle}{\frac{1}{T} \langle x - \bar{x} \cdot \mathbf{1}, x - \bar{x} \cdot \mathbf{1} \rangle \frac{1}{T} \langle y - \bar{y} \cdot \mathbf{1}, y - \bar{y} \cdot \mathbf{1} \rangle}$$

$$r_{x(t),y(t)} = \frac{\frac{1}{\int_0^T dt} \cdot \int_0^T (x(t) - \bar{x} \cdot \mathbf{1}(t)) \cdot (y(t) - \bar{y} \cdot \mathbf{1}(t)) dt}{\frac{1}{\int_0^T 1^2(t) dt} \cdot \int_0^T (x(t) - \bar{x} \cdot \mathbf{1}(t))^2 dt \cdot \frac{1}{\int_0^T 1^2(t) dt} \cdot \int_0^T (y(t) - \bar{y} \cdot \mathbf{1}(t))^2 dt}$$

**Función media muestral:**

Sea  $x_1(t), x_2(t), \dots, x_N(t)$  una muestra de funciones de una función aleatoria  $x(t)$ , definida en  $[0, T]$ .

La función media muestral de  $x(t)$  se define como:  $\bar{x}(t) = \frac{1}{N} \sum_{i=1}^N x_i(t)$

Si fijamos el valor de  $t$  en un punto concreto y evaluamos ahí todas las funciones entonces obtenemos una muestra de tamaño  $N$  de la cual podemos extraer su media. La función media muestral de  $x(t)$  nos va a dar la media de esos valores en ese valor de  $t$  de forma explícita en una sola función.

**Función varianza muestral:**

Sea  $x_1(t), x_2(t), \dots, x_N(t)$  una muestra de funciones de una función aleatoria  $x(t)$ , definida en  $[0, T]$ .

La función varianza muestral de  $x(t)$  se define como:

$$Var_{x(t)}(t) = \frac{1}{N-1} \sum_{i=1}^N (x_i(t) - \bar{x}(t))^2$$

**Función desviación estándar muestral:**

La función desviación estándar muestral de  $x(t)$  se define como:

$$Stdev_{x(t)}(t) = \sqrt{Var_{x(t)}(t)}$$

De la misma manera que la función media muestral de  $x(t)$  nos indica la tendencia central de las funciones en un  $t$  dado, las funciones varianza muestral y desviación estándar muestral de  $x(t)$  nos cuantifican el valor medio al cuadrado y el valor medio respectivamente de las desviaciones respecto a la media en  $t$ .

**Función covarianza muestral:**

Sea  $x_1(t), x_2(t), \dots, x_N(t)$  una muestra de funciones de una función aleatoria  $x(t)$ , definida en  $[0, T]$ .

La función covarianza muestral de  $x(t)$  entre  $t_1$  y  $t_2$  será:

$$Cov_{x(t)}(t_1, t_2) = \frac{1}{N-1} \sum_{i=1}^N (x_i(t_1) - \bar{x}(t_1)) \cdot (x_i(t_2) - \bar{x}(t_2))$$

**Función correlación muestral:**

Sea  $x_1(t), x_2(t), \dots, x_N(t)$  una muestra de funciones de una función aleatoria  $x(t)$ , definida en  $[0, T]$ .

La función correlación muestral de  $x(t)$  entre  $t_1$  y  $t_2$  será:

$$Corr_{x(t)}(t_1, t_2) = \frac{Cov_{x(t)}(t_1, t_2)}{\sqrt{Var_{x(t)}(t_1) \cdot Var_{x(t)}(t_2)}}$$

Estas dos medidas nos van a indicar la magnitud de la relación entre el comportamiento de la función  $x(t)$  en el valor  $t_1$ , y el comportamiento de  $x(t)$  en el valor  $t_2$ .

**Función covarianza cruzada:**

Sean dos funciones aleatorias  $x(t)$  e  $y(t)$  y sean:

- $x_1(t), x_2(t), \dots, x_N(t)$  muestra de  $x(t)$
- $y_1(t), y_2(t), \dots, y_N(t)$  muestra de  $y(t)$

entonces la función de covarianza cruzada entre  $x(t)$  e  $y(t)$  en  $t_1, t_2$  se define como:

$$Cov_{x(t), y(t)}(t_1, t_2) = \frac{1}{N-1} \sum_{i=1}^N [x_i(t_1) - \bar{x}(t_1)] \cdot [(y_i(t_2) - \bar{y}(t_2))]$$

**Correlación Cruzada:**

Sean dos funciones aleatorias  $x(t)$  e  $y(t)$  y sean:

- $x_1(t), x_2(t), \dots, x_N(t)$  muestra de  $x(t)$

- $y_1(t), y_2(t), \dots, y_N(t)$  muestra de  $y(t)$

La función de correlación cruzada de  $x(t)$  e  $y(t)$  en  $t_1, t_2$  se define como:

$$Corr_{x(t), y(t)}(t_1, t_2) = \frac{Cov_{x(t), y(t)}(t_1, t_2)}{\sqrt{Var_{x(t)}(t_1) \cdot Var_{y(t)}(t_2)}}$$

## 8.2 ANEXO B: CÓDIGO R - DATOS FUNCIONALES

```
#####

##                                                                 ##
## Análisis Climatológico Hacienda El Prado - IASA 1 ##
##                                                                 ##
##           Empleado Datos Funcionales           ##
##                                                                 ##
##           Ramiro Guerrón Varela           ##
##                                                                 ##
##           Carlos Jiménez Mosquera           ##
##                                                                 ##

#####

#####

## Cargar las librería requeridas ##

#####

library(tseries)

library(astsa)

library(TSA)

library(it-smr)

library(fda)

library(fda.usc)

library(stats)

library(fpca)
```

```
library(knitr)

library(fBasics)

library(foreign)

library(xtable)

library(plot3D)

#####

## Seleccionar el directorio de trabajo ##

#####

setwd("C:/Users/TESIS/Desarrollo Tesis/TESIS IASA")

#####

## Lectura de Datos ##

#####

rm(list = ls())

DatosIasa=read.csv(

file="Registros_Meterologicos_diarios_IASA_1998_2014.csv",

sep=";",dec=",",header=TRUE)

# (DatosIasa)

head(DatosIasa)

attach(DatosIasa)
```

```
nombres=c(names(DatosIasa))
```

```
nombres
```

```
D.T.Min.S=ts(Temp.Minima,start = c(1998, 1),
             freq = 365)
```

```
D.T.Max.S=ts(Temp.Maxima,start = c(1998, 1),
             freq = 365)
```

```
D.Precip.S=ts(Precipitacion,start = c(1998, 1),
              freq = 365)
```

```
D.Hum.Rel.S=ts(Humedad.Relativa,start = c(1998, 1),
               freq = 365)
```

```
D.Evap.S=ts(Evaporacion,start = c(1998, 1),
            freq = 365)
```

```
D.Helio.S=ts(Heliofania,start = c(1998, 1),
             freq = 365)
```

```
#####
```

```
## Ploteo de las series de tiempo ##
```

```
#####
```

```
par(mfrow=c(3,2))
```

```
plot.ts(D.T.Min.S,main="Temperatura Mnima",
        xlab="Aos",ylab="Temperatura [C]",
        col="blue",xlim=c(1998,2015))
```

```
grid();abline(h=mean(D.T.Min.S),col="red",lwd=2)
```

```
plot.ts(D.T.Max.S,main="Temperatura Máxima",
        xlab="Años",ylab="Temperatura [C]",
        col="blue",xlim=c(1998,2015))

grid()

abline(h=mean(D.T.Max.S),col="red")

plot.ts(D.Precip.S,main="Precipitación",
        xlab="Años", ylab="mm",col="blue",
        xlim=c(1998,2015));grid()

abline(h=mean(D.Precip.S),col="red",lwd="2")

Lista=which(D.Precip.S==0)

length(Lista)

Precip2=ts(D.Precip.S,start = c(1998, 1),
          freq = 365)

Precip2=replace(Precip2,Lista,0.1)

Humedad.R.S=ts(Humedad.Relativa,
              start = c(1998, 1),freq = 365)

plot.ts(Humedad.R.S,xlab="Años",ylab="Porcentaje",
        main="Humedad Relativa",col="blue");grid()

abline(h=mean(Humedad.R.S),col="red")

plot.ts(D.Evap.S,xlab="Año",ylab="horas",
        main="Evaporación",col="blue")
```

```

grid();abline(h=mean(D.Evap.S),col="red")

plot.ts(D.Helio.S,xlab="Año",ylab="horas",
        main="Heliofanía",col="blue")

grid();abline(h=mean(D.Helio.S),col="red")

#####

## Diferencias Máximo - Mínima Diario ##

#####

par(mfrow = c(2, 2))

Descom.Min=decompose(D.T.Min.S)

plot(Descom.Min$trend,main="Tendencia Mínima",
     xlab="Años", ylab="Temperatura [C]",
     col="blue");grid()

Descom.Max=decompose(D.T.Max.S)

plot(Descom.Max$trend,main="Tendencia Máxima",
     xlab="Años",ylab="Temperatura [C]",
     col="blue");grid()

Dif.Max.Min=ts(D.T.Max.S-D.T.Min.S,
              start=c(1998,1), frequency=365)

plot.ts(Dif.Max.Min,xlab="Años",
        ylab="Temperatura [C]",

```

```

    main="Diferencia Temp. Máxima y Mínima",
    col="blue");grid()

abline(h=mean(Dif.Max.Min), col="red")

mat3=cbind(Descom.Min$trend,Descom.Max$trend)

## mat3

ti=c("Descomposición Diferencia T.Max y T.Min.")

matplot(mat3,type="l",xlab="Años",
        ylab="Temperatura [C]",
        main=ti)

grid()

plot(stl(Dif.Max.Min,s.window="periodic"),
     col=c("blue"),
     main="Descomposición Diferencia de Temperaturas ")

grid()

#####

## Definición de un sistema de funciones base ##

#####

rango=c(0,365*14)

# rango

dia=1:5110

## Seleccionamos un día cada cinco

```

```
dia5=seq(0,5110,5)

#dia5

length(dia5)

#####

## Creamos las funciones base ##

#####

## usaremos series de Fourier

## ?create.fourier.basis

par(mfrow=c(2,2))

fbasis=create.fourier.basis(rango,3,365)
plot(fbasis,ylab="nbasis = 3")

fbasis=create.fourier.basis(rango,5,365)
plot(fbasis,ylab="nbasis = 5")

fbasis=create.fourier.basis(rango,7,365)
plot(fbasis,ylab="nbasis = 7")

fbasis=create.fourier.basis(rango,15,365)
plot(fbasis,ylab="nbasis = 15")
```

```
fbasis=create.fourier.basis(rango,169,5110)

# plot(fbasis,ylab="nbasis = 365")

## ?eval.basis

fb.valores=eval.basis(dia5,fbasis)

# fb.valores

dim(fb.valores)

par(mfrow=c(2,2))

plot(dia5,fb.valores[,2],type="p",pch=20,col="red")
plot(dia5,fb.valores[,3],type="p",pch=20,col="red")
plot(dia5,fb.valores[,7],type="p",pch=20,col="red")
plot(dia5,fb.valores[,15],type="p",pch=20,col="red")

fbasis=create.fourier.basis(rango,169,5110)

# plot(fbasis,ylab="nbasis = 365")

## ?eval.basis

fb.valores=eval.basis(dia5,fbasis)

length(fbasis)

Xmat=eval.basis(dia,fbasis)
```

```
dim(Xmat)

# Seleccionamos el número de funciones de la serie
# de Fourier

Xmat2=Xmat[,1:56]

## Generación del modelo y su gráfica suavizada
## y ajustada

par(mfrow=c(3,2))

## Análisis de la Temperatura Promedio Diaria

## Primero chequeamos las distribución para ver su
## normalidad

par(mfrow=c(1,1))

qqnorm(D.T.Min.S,
main="Normal Q-Q Plot - Temperatura Mínima",
xlab="Cuantiles Reales",ylab="Cuantiles Teóricos",
datax=TRUE,col="red")

## Realizamos una regresión sobre la serie de Fourier
```

```
par(mfrow=c(3,1))

D.T.Min.S.2<-D.T.Min.S[1:5110]

plot(D.T.Min.S.2,type="l",
main="Serie Temperatura Mnima",
      ylab="Temperatura [C]", xlab="dias",col="blue")

D.T.Min.S.3<-filter(D.T.Min.S.2, filter = rep(1/15, 15))

plot(D.T.Min.S.3,type="l",
main="Temperatura Suavizada",
      ylab="Temperatura [C]", xlab="dias",col="blue")

length(D.T.Min.S.3)

modelo.temp.min=lm(D.T.Min.S.3~Xmat2)

length(modelo.temp.min$fitted)

#plot(modelo.temp.min$fitted)

length(dia[3:5098])

plot(dia,D.T.Min.S.3, cex=0.5,col="blue",
pch=16,type="l",
      main="Temperatura Mnima ajustado Series de Fourier",
      xlab="dias", ylab="Temperatura [C]"); grid()

lines(dia[3:5098],modelo.temp.min$fitted,col=2,lwd=2)
```

```

print(summary(modelo.temp.min),2)

## xtable(print(summary(modelo.temp.min),2))

#####

## Temperatura Máxima ##

#####

par(mfrow=c(1,1))

qqnorm(D.T.Max.S,

main="Normal Q-Q Plot - Temperatura Máxima",

xlab="Cuantiles Reales",

ylab="Cuantiles Teóricos",

datax=TRUE,col="red")

par(mfrow=c(3,1))

D.T.Max.S.2<-D.T.Max.S[1:5110]

plot(D.T.Max.S.2,type="l",

main="Serie Temperatura Máxima",

ylab="Temperatura [C]", xlab="dias",col="blue")

D.T.Max.S.3<-filter(D.T.Max.S.2, filter = rep(1/15, 15))

plot(D.T.Max.S.3,

main="Temperatura Máxima Suavizada",

ylab="Temperatura [C]", xlab="dias",col="blue")

```

```

length(D.T.Max.S.2)

modelo.temp.Max=lm(D.T.Max.S.3~Xmat2)

length(modelo.temp.Max$fitted)

#plot(modelo.temp.Max$fitted,
main="Temperatura Máxima ajustada Series de Fourier",
#   ylab="Temperatura [C]", xlab="dias",col="blue")

length(dia[3:5098])

plot(dia,D.T.Max.S.3, cex=0.5,col="blue",
pch=16,type="l",
     xlab="dias",ylab="Temperatura [C]",
     main="Temperatura Máxima ajustada Series de Fourier")

grid()

lines(dia[3:5098],modelo.temp.Max$fitted,col=2,lwd=2)

print(summary(modelo.temp.Max),2)

## xtable(print(summary(modelo.temp.Max),2))

par(mfrow=c(1,1))

plot(dia[3:5098],modelo.temp.Max$fitted, col="blue",
     main="Comparación entre temperaturas máxima y mínima",
     ylab="Temperatura [C]",ylim=c(0,22),lwd=1 )

```

```

lines(dia[3:5098],modelo.temp.min$fitted,col="blue")

grid()

plot(dia,c(D.T.Max.S.3-D.T.Min.S.3),
main="Diferencia de temperatura Máxima y Mínima",
type="l",col="cyan",ylab="grados [C]")

lines(dia[3:5098],
c(modelo.temp.Max$fitted-modelo.temp.min$fitted),
col="blue",lwd=2 )

grid()

#####

## Modelo para penalización de rugosidad ##

#####

## Objeto Operador diferencial lineal (Lfd)

## ?Lfd

## Dos maneras comunes de Generalizar Objetos

## Lfd, la primera

## con int2Lfd

## ?int2Lfd

## int2Lfd penaliza algunas derivadas

```

```

curv.Lfd=int2Lfd(2)

## curv.Lfd

## La segunda forma con vec2Lfd

## ?vec2Lfd

## Configurar un operador de aceleración armónica

Lcoef<-c(0, (2*pi/365)^2, 0)

Lcoef

# plot(Lcoef)

harmLfd=vec2Lfd(Lcoef, rangeval=c(0, 365))

## harmLfd

## strharmLfd)

# plot(harmLfd)

hmat <- vec2Lfd(matrix(Lcoef, 1), c(0, 365))

all.equal(harmLfd, hmat)

#####

## Definir un Objeto Parámetro Funcional fdPar ##

#####

```

```
## ?fdPar

lambda=1e1

curv.fdPar=fdPar(fbasis, curv.Lfd, lambda)

# curv.fdPar

## strcurv.fdPar)

## ?smooth.basis

# dia

par(mfrow=c(1,1))

length(dia)

length(D.T.Min.S[1:5110])

tempSmooth1.Min= smooth.basis(dia, c(D.T.Min.S.2),

curv.fdPar)

names(tempSmooth1.Min)

plot(tempSmooth1.Min$fd,

      main="Suavización de la Curva de Temperatura Mínima",

      xlab="Día del año", ylab="Temperatura C",

      col="blue", bg="TAN")

grid()

# tempSmooth1$argvals

length(D.T.Min.S.2)
```

```
par(mfrow=c(1,1))

plotfit.fd(D.T.Min.S.2,dia,tempSmooth1.Min$fd,

col="cyan")

lines(tempSmooth1.Min$fd,col=4)

# Examinar los parámetros del ajuste de la curva

names(tempSmooth1.Min)

## summary(tempSmooth1.Min)

tempSmooth1.Min$df

#tempSmooth1.Min$fd

tempSmooth1.Min$gcv

tempSmooth1.Min$beta

tempSmooth1.Min$SSE

tempSmooth1.Min$y2Map

# tempSmooth1.Min$argvals

# tempSmooth1.Min$y

mean(tempSmooth1.Min$y)

max(tempSmooth1.Min$y)

min(tempSmooth1.Min$y)

var(tempSmooth1.Min$y)

sd(tempSmooth1.Min$y)

## summary(tempSmooth1.Min$y)
```

```
par(mfrow=c(3,1))

temp.suav.fd<-tempSmooth1.Min$fd

plot(temp.suav.fd,
      main="Temperatura mínima suavizada",
      ylab="C",xlab="tiempo",col="blue")
axisIntervals(4); axis(2);grid()
abline(h=mean(tempSmooth1.Min$y),col="red");grid()

d.temp.suav.fd<-deriv.fd(temp.suav.fd)
plot(deriv.fd(d.temp.suav.fd),
      main="Primera derivada temperatura mínima suavizada",
      xlab="tiempo",col="blue", ylab="C/dia")
axisIntervals(4); axis(2);grid()
abline(h=0,col="red")

d2.temp.suav.fd<-deriv.fd(d.temp.suav.fd)
plot(deriv.fd(d2.temp.suav.fd),
      main="Segunda derivada mínima",
      xlab="tiempo",col="blue", ylab="C/dia/dia")
axisIntervals(4); axis(1);grid()
abline(h=0,col="red")

## Suavización Temperatura Máxima
```

```
par(mfrow=c(1,1))

tempSmooth2.Max = smooth.basis(dia,c(D.T.Max.S.2),
curv.fdPar)

names(tempSmooth2.Max)

plot(tempSmooth2.Max$fd,

      main="Temperatura máxima suavizada",
      xlab="Día del año",ylab="Temperatura C",
      col="blue")

grid()

par(mfrow=c(3,1))

temp.suav.fd<-tempSmooth2.Max$fd

plot(temp.suav.fd,

      main="Temperatura máxima suavizada",
      ylab="C",xlab="tiempo",col="blue")

axisIntervals(4); axis(2);grid()

abline(h=mean(tempSmooth2.Max$y),col="red");grid()

d.temp.suav.fd<-deriv.fd(temp.suav.fd)

plot(deriv.fd(d.temp.suav.fd),

      main="Primera derivada temperatura máxima suavizada",
      xlab="tiempo",col="blue", ylab="C/dia")

axisIntervals(4); axis(2);grid()

abline(h=0,col="red")
```

```

d2.temp.suav.fd<-deriv.fd(d.temp.suav.fd)

plot(deriv.fd(d2.temp.suav.fd),

     main="Segunda derivada temperatura máxima suavizada",

     xlab="tiempo",col="blue",ylab="C/dia/dia")

axisIntervals(4); axis(1);grid()

abline(h=0,col="red")

mean(tempSmooth2.Max$y)

is.fd(tempSmooth2.Max$fd)

max(tempSmooth2.Max$y)

min(tempSmooth2.Max$y)

var(tempSmooth2.Max$y)

sd(tempSmooth2.Max$y)

summary(tempSmooth2.Max$y)

#####

## Un año típico ##

#####

## En esta sección trabajaremos con los datos de

## los promedios día calendario

rm(list = ls())

DatosIasaProm=read.csv(

```

```
file="DatosMeterologicosDiariosMediosIASA
1998-2014-04.csv", sep=";", dec=", ",
header=TRUE)

# View(DatosIasaProm)

attach(DatosIasaProm)

head(DatosIasaProm)

D.Prom.T.Min.S=ts(DatosIasaProm$Temp.Minima,
start=c(0,1), freq=1)

D.Prom.T.Max.S=ts(DatosIasaProm$Temp.Maxima)

D.Prom.Precip.S=ts(DatosIasaProm$Precipitacion)

D.Prom.Humedad.Relativa.S=
ts(DatosIasaProm$Humedad.Relativa)

D.Prom.Evap.S=ts(DatosIasaProm$Evaporacion)

D.Prom.Helio.S=ts(DatosIasaProm$Heliofania)

par(mfrow=c(3,2))

plot.ts(D.Prom.T.Min.S, axes=FALSE, cex=0.5,
col="blue",
      pch=16, type="l", xlab="meses",
      ylab="Temperatura [C]",
      main="Temperatura Mnima Promedio")

grid(); axisIntervals(1); axis(2); grid()

plot.ts(D.Prom.T.Max.S, axes=FALSE,
```

```
cex=0.5,col="blue",  
    pch=16,type="l",xlab="meses",  
    ylab="Temperatura [C]",  
    main="Temperatura Máxima Promedio ")  
grid(); axisIntervals(1); axis(2);grid()  
  
plot.ts(D.Prom.Precip.S, axes=FALSE,cex=0.5,  
col="blue",pch=16,  
    type="l", xlab="meses",ylab="mm",  
    main="Precipitación Promedio ")  
axisIntervals(1); axis(2);grid()  
  
plot.ts(D.Prom.Humedad.Relativa.S,  
axes=FALSE,cex=0.5,  
    col="blue",pch=16, type="l",xlab="meses",  
    ylab="Porcentaje",  
    main="Humedad Relativa Promedio")  
axisIntervals(1); axis(2);grid()  
  
plot.ts(D.Prom.Evap.S, axes=FALSE,  
cex=0.5,col="blue",  
    pch=16,type="l", xlab="meses",ylab="mm",  
    main="Evaporación Promedio")  
axisIntervals(1); axis(2);grid()  
  
plot.ts(D.Prom.Helio.S, axes=FALSE,
```

```

cex=0.5,col="blue",

    pch=16,type="l", xlab="meses",ylab="horas",

    main="Heliofanía Promedio")
axisIntervals(1); axis(2);grid()

#####

## Definición de un sistema de funciones base ##

#####

## Todas las bases requieren un rango, en este
## caso sería 365 días del año

rango=c(0,365)

dia=1:365

## Seleccionamos un día cada cinco

dia5=seq(0,365,5)

dia5

length(dia5)

## Creamos las funciones base, usaremos series
## de Fourier

## con 365 funciones base como máximo.

par(mfrow=c(2,2))

fbasis=create.fourier.basis(rango,5,365)

```

```
fb.valores=eval.basis(dia5,fbasis)

dim(fb.valores)

fbasis=create.fourier.basis(rango,365)

length(fbasis)

Xmat=eval.basis(dia,fbasis)

dim(Xmat)

# Seleccionamos el número de funciones
# de la serie de Fourier

Xmat2=Xmat[,1:5]

## Generación del modelo y su gráfica
## suavizada y ajustada

par(mfrow=c(3,2))

## Generación del modelo y su gráfica
## suavizada y ajustada

## Primero chequeemos las distribución
## para ver su normalidad

par(mfrow=c(3,2))
```

```

## Realizamos una regresión sobre la
## serie de Fourier

modelo.temp.min=lm(D.Prom.T.Min.S~Xmat2)

plot(dia,D.Prom.T.Min.S, axes=FALSE,
cex=0.5,col="blue",
      pch=16,type="l", xlab="meses",
      ylab="Temperatura [C]",
      main="Temperatura Mínima Promedio")

grid(); axisIntervals(1); axis(2);grid()

lines(dia,modelo.temp.min$fitted,col=2,lwd=2)

print(summary(modelo.temp.min),2)

## xtable(print(summary(modelo.temp.min),2))

modelo.temp.max=lm(D.Prom.T.Max.S~Xmat2)

plot(dia,D.Prom.T.Max.S, axes=FALSE,
cex=0.5,col="blue",
      pch=16,type="l",xlab="meses",
      ylab="Temperatura [C]",
      main="Temperatura Máxima Promedio ")

grid(); axisIntervals(1); axis(2);grid()

lines(dia,modelo.temp.max$fitted,col=2,lwd=2)

summary(modelo.temp.max)

print(summary(modelo.temp.max),2)

```

```

modelo.Precip=lm(D.Prom.Precip.S~Xmat2)

plot(dia,D.Prom.Precip.S, axes=FALSE,
cex=0.5,col="blue",pch=16,
      type="l", xlab="meses",ylab="mm",
      main="Precipitación Promedio ")

axisIntervals(1); axis(2);grid()

lines(dia,modelo.Precip$fitted,col=2,lwd=2)

summary(modelo.Precip)

## xtable(print(summary(modelo.Precip),2))

modelo.Humedad.Relativa=lm(
D.Prom.Humedad.Relativa.S~Xmat2)

plot(dia,D.Prom.Humedad.Relativa.S,
axes=FALSE,cex=0.5,
      col="blue",pch=16, type="l",xlab="meses",
      ylab="Porcentaje",
      main="Humedad Relativa Promedio")

axisIntervals(1); axis(2);grid()

lines(dia,modelo.Humedad.Relativa$fitted,
col=2,lwd=2)

summary(modelo.Humedad.Relativa)

## xtable(print(summary
## (modelo.Humedad.Relativa),2))

modelo.Evap=lm(D.Prom.Evap.S~Xmat2)

plot(dia,D.Prom.Evap.S, axes=FALSE,

```

```

cex=0.5,col="blue",

    pch=16,type="l", xlab="meses",ylab="mm",

    main="Evaporación Promedio")
axisIntervals(1); axis(2);grid()
lines(dia,modelo.Evap$fitted,col=2,lwd=2)
summary(modelo.Evap)

## xtable(print(summary(modelo.Evap),2))

modelo.Helio=lm(D.Prom.Helio.S~Xmat2)
plot(dia,D.Prom.Helio.S, axes=FALSE,

cex=0.5,col="blue",

    pch=16,type="l", xlab="meses",ylab="horas",

    main="Heliofanía Promedio")
axisIntervals(1); axis(2);grid()
lines(dia,modelo.Helio$fitted,col=2,lwd=2)
summary(modelo.Helio)

## xtable(print(summary(modelo.Helio),2))

print(summary(modelo.temp.min),digits=2)
print(summary(modelo.temp.max),digits=2)
print(summary(modelo.Precip),digits=2)
print(summary(modelo.Humedad.Relativa),

digits=2)

print(summary(modelo.Evap),digits=2)
print(summary(modelo.Helio),digits=2)

```

```
## Tratamos de mejorar el modelo aumentando
## a 10 términos la serie de Fourier

Xmat3=Xmat[,1:10]

## Xmat3

par(mfrow=c(3,2))

modelo1.temp.min=lm(D.Prom.T.Min.S~Xmat3)
plot(dia,D.Prom.T.Min.S, axes=FALSE,
cex=0.5,col="blue",
      pch=16,type="l", xlab="meses",
      ylab="Temperatura [C]",
      main="Temperatura Mínima Promedio")
grid(); axisIntervals(1); axis(2)

lines(dia,modelo1.temp.min$fitted,col=2,lwd=2)

summary(modelo1.temp.min)

modelo1.temp.max=lm(D.Prom.T.Max.S~Xmat3)
plot(dia,D.Prom.T.Max.S, axes=FALSE,
cex=0.5,col="blue",
      pch=16,type="l", xlab="meses",
      ylab="Temperatura [C]",
      main="Temperatura Máxima Promedio")
grid(); axisIntervals(1); axis(2)
```

```
lines(dia,modelo1.temp.max$fitted,col=2,lwd=2)
```

```
modelo1.Precip=lm(D.Prom.Precip.S~Xmat3)
```

```
plot(dia,D.Prom.Precip.S, axes=FALSE,
```

```
cex=0.5,col="blue",
```

```
    pch=16,type="l", xlab="meses",ylab="mm",
```

```
    main="Precipitación Promedio")
```

```
axisIntervals(1); axis(2);grid()
```

```
lines(dia,modelo1.Precip$fitted,col=2,lwd=2)
```

```
modelo1.Humedad.Relativa=lm(
```

```
D.Prom.Humedad.Relativa.S~Xmat3)
```

```
plot(dia,D.Prom.Humedad.Relativa.S,
```

```
axes=FALSE,cex=0.5,
```

```
    col="blue",pch=16, type="l",
```

```
    xlab="meses",ylab="Porcentaje",
```

```
    main="Humedad Relativa Promedio")
```

```
axisIntervals(1); axis(2);grid()
```

```
lines(dia,modelo1.Humedad.Relativa$fitted,
```

```
col=2,lwd=2)
```

```
modelo1.Evap=lm(D.Prom.Evap.S~Xmat3)
```

```
plot(dia,D.Prom.Evap.S, axes=FALSE,
```

```
cex=0.5,col="blue",
```

```
    pch=16,type="l", xlab="meses",ylab="mm",
```

```
    main="Evaporación Promedio")
```

```

axisIntervals(1); axis(2);grid()

lines(dia,modelo1.Evap$fitted,col=2,lwd=2)

modelo1.Helio=lm(D.Prom.Helio.S~Xmat3)
plot(dia,D.Prom.Helio.S, axes=FALSE,
cex=0.5,col="blue",
      pch=16,type="l", xlab="meses",
      ylab="horas",
      main="Heliofanía Promedio")
axisIntervals(1); axis(2);grid()
lines(dia,modelo1.Helio$fitted,col=2,lwd=2)

print(summary(modelo1.temp.min),digits=2)
print(summary(modelo1.temp.max),digits=2)
print(summary(modelo1.Precip),digits=2)
print(summary(modelo1.Humedad.Relativa),
digits=2)
print(summary(modelo1.Evap),digits=2)
print(summary(modelo1.Helio),digits=2)

#####
## Construcción de un modelo para penalización de rugosidad ##
#####

curv.Lfd=int2Lfd(2)

## curv.Lfd

```

```

Lcoef<-c(0, (2*pi/365)^2, 0)

Lcoef

harmLfd=vec2Lfd(Lcoef, rangeval=c(0, 365))

lambda=1e1

curv.fdPar=fdPar(fbasis, curv.Lfd, lambda)

par(mfrow=c(1, 1))

plot(D.Prom.T.Min.S, type="p", pch=".",
main="Suavización de la Curva de Temperatura
Promedio Mínima",
xlab="Día del año", ylab="Temperatura [C]",
col="blue", axes=FALSE)

tempSmooth1 = smooth.basis(d
ia, c(D.Prom.T.Min.S), curv.fdPar)
names(tempSmooth1)

lines(tempSmooth1$fd,
main="Suavización de la Curva de Temperatura
Promedio Mínima",
xlab="Día del año", ylab="Temperatura C",
col="red", bg="TAN")

axisIntervals(1); axis(2); grid()

```

```
length(D.Prom.T.Min.S)

par(mfrow=c(1,1))

plotfit.fd(D.Prom.T.Min.S,dia,tempSmooth1$fd,
col="cyan")

lines(tempSmooth1$y,col=4)

## Examinar los parámetros del ajuste de la curva

names(tempSmooth1)

## summary(tempSmooth1)

tempSmooth1$df
# tempSmooth1$fd
tempSmooth1$gcv
tempSmooth1$beta
tempSmooth1$SSE
tempSmooth1$y2Map
# tempSmooth1$argvals
# tempSmooth1$y

mean(tempSmooth1$y)
max(tempSmooth1$y)
min(tempSmooth1$y)
```

```
var(tempSmooth1$y)

sd(tempSmooth1$y)

summary(tempSmooth1$y)

par(mfrow=c(3,1))

temp.suav.fd<-tempSmooth1$fd

plot(temp.suav.fd

,main="Suavización Temperatura Promedio Mínima",

      ylab="C",xlab="tiempo",col="blue")

axisIntervals(4); axis(2);grid()

abline(h=mean(tempSmooth1$y),col="red");grid()

d.temp.suav.fd<-deriv.fd(temp.suav.fd)

plot(deriv.fd(d.temp.suav.fd),

main="Primera derivada temperatura

promedio mínima",

xlab="tiempo",col="blue", ylab="C/dia")

axisIntervals(4); axis(2);grid()

abline(h=0,col="red")

d2.temp.suav.fd<-deriv.fd(d.temp.suav.fd)

plot(deriv.fd(d2.temp.suav.fd),

main="Segunda derivada temperatura

promedio mínima",

xlab="tiempo",col="blue",ylab="C/dia/dia")
```

```
axisIntervals(4); axis(1);grid()

abline(h=0,col="red")

## Suavización Temperatura Máxima

par(mfrow=c(1,1))

plot(D.Prom.T.Max.S,type="p",pch=".",
main="Suavización de la Curva de Temperatura
Promedio Máxima",
      xlab="Día del año",ylab="Temperatura C",
      col="blue",axes=FALSE)

tempSmooth2 = smooth.basis(
dia,c(D.Prom.T.Max.S),curv.fdPar)
names(tempSmooth2)
lines(tempSmooth2$fd,col=2)
axisIntervals(1); axis(2);grid()

par(mfrow=c(3,1))

temp.suav.fd<-tempSmooth2$fd

plot(
temp.suav.fd,
main="Temperatura promedio máxima suavizada",
      ylab="C",xlab="tiempo",col="blue")
```

```
axisIntervals(4); axis(2);grid()

abline(h=mean(tempSmooth2$y), col="red");grid()

d.temp.suav.fd<-deriv.fd(temp.suav.fd)

plot(deriv.fd(d.temp.suav.fd),

main="Primera derivada temperatura promedio máxima",

xlab="tiempo", col="blue", ylab="C/dia")

axisIntervals(4); axis(2);grid()

abline(h=0, col="red")

d2.temp.suav.fd<-deriv.fd(d.temp.suav.fd)

plot(deriv.fd(d2.temp.suav.fd),

main="Segunda derivada temperatura promedio máxima",

xlab="tiempo", col="blue", ylab="C/dia/dia")

axisIntervals(4); axis(1);grid()

abline(h=0, col="red")

mean(tempSmooth2$y)

is.fd(tempSmooth2$fd)

max(tempSmooth2$y)

min(tempSmooth2$y)

var(tempSmooth2$y)

sd(tempSmooth2$y)

summary(tempSmooth2$y)
```

```
## Visualización de datos en tres dimensiones

par(mfrow=c(1,1))

library(plot3D)

scatter2D(tempSmooth1$y,tempSmooth2$y,
type="p",pch=20,
      col="blue");grid()

scatter3D(1:365,theta=45,phi=30,
tempSmooth1$y,
      tempSmooth2$y,type="p",pch=20,
      zlim=range(tempSmooth2$y))

scatter3D(1:365,theta=45,phi=30,dia,
tempSmooth1$y,type="p",
      pch=20, zlim=range(tempSmooth1$y),
      bty = "f",
      ticktype = "detailed",xlim = c(1, 365),
      ylim = c(1,365))
```