



Universidad Central de Venezuela  
Facultad de Ciencias  
Escuela en Matemática

## **Estudio de algunas variables meteorológicas en el Lago de Valencia usando series temporales.**

**Autor: Julibeth Tinoco Mantilla**

**Tutor: Dr. José Benito Hernández Chaudary**

Trabajo Especial de Grado presentada ante la  
ilustre Universidad de Central de Venezuela  
como requisito para obtener el título de  
Licenciada en Matemáticas.

Caracas, Abril 2015.

Nosotros, los abajo firmantes, designados por la Universidad Central de Venezuela como integrantes del Jurado Examinador del Trabajo Especial de Grado titulado "**Estudio de algunas variables meteorológicas en el Lago de Valencia usando series temporales.**", presentado por la **Br. Julibeth Tinoco Mantilla**, titular de la Cédula de Identidad **20.677.587**, certificamos que este trabajo cumple con los requisitos exigidos por nuestra Magna Casa de Estudios para optar al título de **Licenciado en Matemática**.

---

**José Benito Hernández**  
**Tutor**

---

**Mairene Colina**  
**Jurado**

---

**Ricardo Rios**  
**Jurado**

*A Dios por darme la constancia y fortaleza para culminarlo. A mi madre por siempre estar en los momentos que más la necesite y enseñarme las herramientas necesarias para el éxito. A mi abuela Julia por ser mi total inspiración y a Gustavo Rojas por ser un gran apoyo en el transcurso de la investigación.*

---

A Dios, a mi familia y a Gustavo Rojas.

---

# Índice general

<b>Contenido</b>	<b>I</b>
<b>Introducción</b>	<b>1</b>
<b>1. Meteorología</b>	<b>4</b>
1.1. Conceptos básicos de meteorología. . . . .	4
1.1.1. La Tierra. . . . .	4
1.2. El clima . . . . .	9
1.2.1. Elementos del clima . . . . .	9
1.2.2. Factores del clima . . . . .	15
1.2.3. Sistema climático. . . . .	19
1.3. Patrones de circulación . . . . .	20
1.3.1. Corrientes Oceánicas . . . . .	20
1.3.2. Patrones de viento . . . . .	22
<b>2. Series temporales.</b>	<b>23</b>
2.0.3. Modelos estadísticos para series de tiempo. . . . .	23
2.0.4. Medidas de dependencia: autocorrelación y correlación cruzada. . . . .	26
2.1. Series de tiempo estacionarias. . . . .	28
2.1.1. Estimación de correlación. . . . .	32
2.2. Regresión de series de tiempo. . . . .	34
2.2.1. Regresión clásica . . . . .	34

---

2.3. El Modelo Aditivo de Componentes de Series de Tiempo. . . . .	35
2.4. Análisis exploratorio de datos. . . . .	37
2.4.1. Eliminación de la tendencia en ausencia de estacionalidad . . . . .	38
2.4.2. Eliminación de la tendencia y la estacionalidad . . . . .	43
<b>3. Análisis de los datos</b>	<b>49</b>
3.1. Temperatura. . . . .	50
3.1.1. Estación Sucre. . . . .	50
3.1.2. Estación Tacarigua. . . . .	56
3.2. Radiación. . . . .	63
3.2.1. Estación Sucre. . . . .	63
3.2.2. Estación Tacarigua. . . . .	72
3.3. Precipitaciones. . . . .	81
3.3.1. Estación Sucre. . . . .	81
3.3.2. Estación Tacarigua. . . . .	84
3.4. Viento. . . . .	87
3.4.1. Estación Sucre. . . . .	87
3.4.2. Estación Tacarigua. . . . .	97
<b>4. Análisis de correlaciones cruzadas</b>	<b>103</b>
4.1. Conclusiones. . . . .	114
<b>Bibliografía</b>	<b>116</b>

---

El lago de Valencia o lago de Tacarigua es el segundo lago en importancia de Venezuela después del lago de Maracaibo y es el cuerpo de agua dulce sin desagüe al mar más grande de Venezuela. Se encuentra emplazado en una fosa tectónica conocida como Graben de Valencia situada entre la cordillera de la Costa y la serranía del Interior. La cuenca es de tipo endorreica <sup>1</sup> y cubre 3,150km<sup>2</sup> lo que equivale al 0,3 % del país.

Este lago es el fenómeno hidrográfico más importante del centro del país, tiene una longitud máxima de 30km, una anchura máxima que alcanza los 20km, una profundidad máxima de 39 metros y una media de 21 metros.

Geológicamente, el lago es bastante joven, aproximadamente de una edad pre-Pleistocena o Pleistoceno medio y su origen se debe a un hundimiento de bloques de la corteza terrestre. A lo largo de su historia el nivel de sus aguas ha fluctuado significativamente. La desaparición del lago de Valencia, al igual que la de todos los lagos del mundo, es un hecho geológico, pero el proceso debió ser casi imperceptible, lo grave fue que el propio hombre lo aceleró durante el siglo XIX y gran parte del siglo XX. La disminución de sus aguas en aquellos años se debió a múltiples causas entre ellas:

- La evaporación excesiva por las altas temperaturas locales.
- La deforestación.
- El fenómeno de Piratería Fluvial, captura del río Tuy (que antes drenaba hacia el lago de Valencia) por el cauce actual de dicho río que drena ahora hacia el Caribe, lo que dio origen al desecamiento progresivo de sus aguas y a la formación de extensas planicies de sedimentos lacustres que convierten a sus riberas en importantes centros agrícolas. Ya Alexander von Humboldt mencionaba la creencia de que el lago podía evaporarse.

Para la década de 1970 el lago resistió prolongados períodos de sequías. Esto unido a la utilización de los afluentes en labores agrícolas y la creciente actividad industrial en la zona ocasionaron un descenso progresivo en el nivel de las aguas del lago, que para el año 1980 su cota llegó a tocar los 401m.s.n.m.

Es de hacer notar que el lago de Valencia antes de sufrir todos estos cambios causados tanto por la naturaleza como por el hombre, vertía el exceso de sus aguas de la época de lluvias a la cuenca del río Orinoco a través de un río emisario que hoy en día es el río Paito, afluente del río Pao que a su vez, es un afluente del Orinoco. El río Paito podría considerarse, en realidad, como la parte superior o cabeceras del río Pao.

---

<sup>1</sup>"En geografía, una cuenca endorreica es un área en la que el agua no tiene salida fluvial hacia el mar. El término tiene raíces griegas, *endo*, "interior" y *rhein*, "fluir". Cualquier lluvia o precipitación que caiga en una cuenca endorreica permanece allí, abandonando el sistema únicamente por infiltración o evaporación, lo cual contribuye a la concentración de sales."

---

---

Sin embargo, en la actualidad se está agravando el problema del aumento del nivel del lago, debido al transvase de las aguas de la cuenca del Orinoco (embalses de Pao Cachinche y Pao La Balsa, en el río Pao) hacia la cuenca endorreica del lago de Valencia, sin haber hecho los canales de drenaje para las aguas sobrantes con el fin de sacarlas de la cuenca antes de que lleguen a un nivel inferior al de salida de la cuenca por gravedad, por el antiguo curso del río Cabriales. Como consecuencia de ello, los niveles de contaminación de sus aguas también han venido aumentando debido a la importante industrialización de las ciudades de Maracay y Valencia, cuyas aguas servidas son vertidas directamente al lago, en gran parte, sin un tratamiento previo, lo que hace poco atractivo el desarrollo de actividades turísticas o recreacionales. Para paliar la situación hace pocos años el río Cabriales que fue desviado hacia el lago en 1979, fue desviado nuevamente, con esto se espera que desciendan el nivel de las aguas. Actualmente están en ejecución los proyectos que permitirán el saneamiento de la cuenca del lago mediante una serie de obras como la adecuación de las plantas de tratamiento de Los Guayos y otras.

En este orden de idea, el tener conocimiento de las condiciones climatológicas en la cuenca del Lago de Valencia resulta muy útil para poder establecer estrategias de saneamiento del mismo, un proyecto que está en funcionamiento actualmente es el proyecto "Alternativas de Saneamiento para el Lago de Valencia a través de un Modelo Matemático Tridimensional". Un proyecto conjunto del CEMVICC - FACYT, (Centro Multidisciplinario de Visualización y Computo Científico de la Facultad de Ciencia y Tecnología, Universidad de Carabobo), la Fundación Universidad de Carabobo, el Gobierno de Carabobo y la Dirección de Hidrografía y Navegación. A finales de la década de los setenta, el Ministerio del Ambiente y de los Recursos Naturales Renovables (MARNR) contrato a la empresa GEOCIN C.A. para realizar un estudio hidrogeológico de la cuenca del Lago de Valencia. Esta empresa ofreció los siguientes servicios: compiló, analizó, interpretó y reportó los datos sobre climatología, hidrología superficial, hidrogeología y flujo en medios porosos y similares en un reporte titulado: Estudio Hidrogeológico Preliminar Cuenca del Lago de Valencia (MARNR-GEOCIN, 1983). Hizo cuatro modelos matemáticos de sistemas de agua subterránea en Maracay, Valencia, San Joaquín y Río Guigue, y creó documentación básica utilizando programas computacionales. El proyecto, que finalizó en 1983, fue el paso inicial en una secuencia de programas ambientales bajo el marco del programa financiado por el Banco Mundial "Programa de Saneamiento Ambiental Integral de la Cuenca del Lago de Valencia". Los problemas que se identificaron durante el estudio fueron los siguientes:

- Vaciado del agua subterránea almacenada debido al sobrebombeo, especialmente cerca de la ciudad de Maracay.
  - Intrusión de aguas negras provenientes del lago dentro del acuífero a lo largo del perímetro del lago y deterioración de las aguas subterráneas utilizadas para la industria de producción de comida, textiles y
-

papel.

- Posible hundimiento de la tierra en áreas con recursos acuíferos sobre-explotados.

Hasta los momentos, la mayoría de estudios realizados en el Lago de Valencia son de tipo físico-químico y biológico, no toman en cuenta la dinámica del lago, como corrientes y transporte de propiedades físicas (temperatura, salinidad, etc.) y su efecto en la biota<sup>2</sup>.

Es posible, estudiar la dinámica del lago con el objetivo de conocer el conjunto de factores que influyen directamente sobre él y consecutivamente ser partícipe del aporte de información científica al ente responsable de la toma de decisiones. En particular, estudiaremos las correlaciones cruzadas entre viento (velocidad), la temperatura, radiación y la precipitación.

Las series de tiempo en la rama de la meteorología, principalmente se usan para mostrar el comportamiento de los elementos del clima durante intervalos de tiempo; Lo cual, se usa a su vez para observar las semejanzas entre los datos de la serie y semejanzas con otras series temporales. Además, se usan en el análisis para la predicción y pronóstico de los elementos climáticos.

Las series temporales son de gran importancia en la meteorología por que se puede tener conocimiento de las condiciones atmosféricas y meteorológicas que están sucediendo o lo que se puede esperar para las próximas horas, días, meses ó años.

El trabajo estará estructurado del siguiente modo: El primer capítulo contendrá la teoría y conceptos básico de meteorología, relacionados con los vientos, las precipitaciones la radiación y las temperaturas. En el capítulo dos se presentará la teoría de series de tiempo, y los conceptos aplicados al estudio meteorológico. En el tercer capítulo realizaremos el análisis de los datos. El cuarto capítulo contendrá el análisis de las correlaciones cruzadas entre la temperatura, el viento, la radiación y la precipitación.

---

<sup>2</sup>"Conjunto de la fauna y la flora de una región".

---

# Meteorología

## 1.1. Conceptos básicos de meteorología.

La meteorología es la ciencia interdisciplinaria, fundamentalmente una rama de la Física de la atmósfera, que estudia el estado del tiempo, el medio atmosférico, los fenómenos allí producidos y las leyes que lo rigen.

A continuación, se mostraran varios conceptos referentes a la meteorología para comprender mejor el comportamiento de los factores climáticos que influyen sobre el lago.

### 1.1.1. La Tierra.

**La Tierra** es un planeta del sistema solar que gira alrededor del Sol en la tercera órbita más interna; al mismo tiempo gira sobre su propio eje cada día.

La Tierra está constituida por :

- **La litosfera:** es la capa sólida superficial de la tierra, caracterizada por su rigidez. Está formada por la corteza terrestre y la zona más externa del manto terrestre, y flota sobre la astenósfera, una capa blanda que forma parte del manto superior. La litosfera suele tener un espesor aproximado de 50 a 300  $km^2$  siendo su límite externo la superficie terrestre.

La litosfera está fragmentada en una serie de placas tectónicas o litosféricas, en cuyos bordes se concentran los fenómenos geológicos endógenos, como el magmatismo (incluido el vulcanismo), la sismicidad o la orogénesis.

- **La hidrosfera:** es el sistema material constituido por el agua que se encuentra bajo y sobre la superficie de la tierra.

La hidrosfera incluye los océanos, mares, ríos, lagos, agua subterránea, el hielo y la nieve. La Tierra es el único planeta del Sistema Solar en el que está presente de manera continuada el agua líquida, que cubre aproximadamente dos terceras partes de la superficie terrestre, con una profundidad promedio de 3,5 km, lo que representa el 97 % del total de agua del planeta. El agua dulce representa 3 % del total y de esta cantidad aproximadamente 98 % está congelada, de allí que solo se tenga acceso al 0,06 % de toda el agua del planeta. El agua migra de unos depósitos a otros por procesos de cambio de estado y de transporte que en conjunto configuran el ciclo hidrológico o ciclo del agua.

La presencia del agua en la superficie terrestre es el resultado de la desgasificación del manto, que está compuesto por rocas que contienen en disolución sólida cierta cantidad de sustancias volátiles, de las que el agua es la más importante. El agua del manto se escapa a través de procesos volcánicos e hidrotermales. El manto recupera gracias a la subducción<sup>3</sup> una parte del agua que pierde a través del vulcanismo.

- **La biosfera:** es el sistema formado por el conjunto de los seres vivos del planeta Tierra y sus relaciones.
- **La atmósfera:** es la capa gaseosa que rodea al planeta tierra, se divide teóricamente en varias capas concéntricas sucesivas. Estas son, desde la superficie hacia el espacio exterior: troposfera, tropopausa, estratosfera, estratopausa, mesosfera y termosfera.

- **La troposfera:** o baja atmósfera es la que está en íntimo contacto con la superficie terrestre y se extiende hasta los 11 km s.n.m. en promedio. Tiene un grosor que varía desde los 8 km. en los polos hasta 16 km. en el ecuador, principalmente debido a la diferencia de presupuesto energético en esos lugares. Abarca el 75 % de la masa de gases totales que componen la atmósfera, el 99 % de la masa de la atmósfera se encuentra bajo los 30 km s.n.m. Consta en particular, del 99 % de dos gases, el Nitrógeno (N<sub>2</sub>, 78 %) y Oxígeno (O<sub>2</sub>, 21 %). El 1 % que resta consta principalmente de Argón (Ar, 1 %) y Dióxido de Carbono (CO<sub>2</sub>, 0,035 %).

El aire de la troposfera incluye vapor de agua en cantidades variables de acuerdo a condiciones locales, por ejemplo, desde 0,01 % en los polos hasta 5 % en los trópicos. La temperatura disminuye con la altura, en promedio, 6,5°C por kilómetro. La mayoría de los fenómenos que involucran el clima ocurren en esta capa de la atmósfera (Kaufmann, 1968), en parte sustentado

---

<sup>3</sup>subducción: deslizamiento del borde de una placa de la corteza terrestre por debajo del borde de otra.

---

por procesos convectivos que son establecidos por calentamiento de gases superficiales, que se expanden y ascienden a niveles más altos de la troposfera donde nuevamente se enfrían. Esta capa incluye además los fenómenos biológicos.

- **La tropopausa:** es la zona de transición entre la troposfera y la estratosfera. Marca el límite superior de la troposfera, sobre la cual la temperatura se mantiene constante antes de comenzar nuevamente a aumentar sobre los 20 km s.n.m. Esta situación térmica evita la convección del aire y confina de esta manera el clima a la troposfera.
- **La estratosfera:** es una de las capas más importantes de la atmósfera, esta se sitúa entre la troposfera y la mesosfera, y se extiende en una capa que va desde los 10 hasta los 50 km de altura aproximadamente. La temperatura aumenta progresivamente desde los  $-55^{\circ}\text{C}$  de la tropopausa hasta alcanzar los  $0^{\circ}\text{C}$  de la estratopausa, aunque según algunos autores puede alcanzar incluso los  $17^{\circ}\text{C}$  o más. Es decir, en esta capa la temperatura aumenta con la altitud, al contrario de lo que ocurre en las capas superior e inferior. Esto es debido principalmente a la absorción de las moléculas de ozono que absorben radiación electromagnética en la región del ultravioleta.

En la parte baja de la estratosfera la temperatura es relativamente estable, y en toda la capa hay muy poca humedad.

La estratosfera es una región en donde se producen diferentes procesos radiactivos, dinámicos y químicos.

- **La estratopausa:** es la capa de transición que está situada entre la mesosfera y estratosfera. La mayor parte del ozono de la atmósfera se sitúa en torno a 22 kilómetros por encima de la superficie de la tierra, en la región próxima a la estratopausa, en la parte superior de la estratosfera. La estratosfera tiene como límite superior la estratopausa, donde está el punto de inflexión de la temperatura, su temperatura se mantiene en torno a  $0^{\circ}\text{C}$ . Los movimientos de aire en esta región son casi en su totalidad horizontales, siguiendo a los vientos de la estratosfera.
  - **La mesosfera:** es la parte de la atmósfera situada por encima de la estratosfera y por debajo de la termosfera. En la mesosfera la temperatura va disminuyendo a medida que se aumenta la altura, hasta llegar a unos  $-80^{\circ}\text{C}$  a los 80.5km aproximadamente. Se extiende desde la estratopausa (zona de contacto entre la estratosfera y la mesosfera) hasta una altura de unos 80 km donde la temperatura vuelve a descender hasta unos  $-70^{\circ}\text{C}$  u  $-80^{\circ}\text{C}$ .
-

La mesosfera es la tercera capa de la atmósfera de la tierra. La temperatura disminuye a medida que se sube, como sucede en la troposfera. Puede llegar a ser hasta de  $-90^{\circ}$  C. Es la zona más fría de la atmósfera.

- **La mesopausa:** es la región de la atmósfera que determina el límite entre una atmósfera con masa molecular constante de otra donde predomina la difusión molecular.

Se sitúa a aproximadamente 90 km de altitud, es la región donde existe la temperatura más baja en la atmósfera, cerca de  $-80^{\circ}$ C.

- **La termosfera:** su extensión comienza aproximadamente entre 80 y 120 kilómetros de la Tierra prolongándose hasta entre 500 y 1000 kilómetros de la superficie terrestre. Dentro de esta capa, la radiación ultravioleta, pero sobre todo los rayos gamma y rayos X provenientes del Sol, provocan la ionización de átomos de sodio y moléculas. En dicho proceso los gases que la componen elevan su temperatura varios cientos de grados, de ahí su nombre.
-

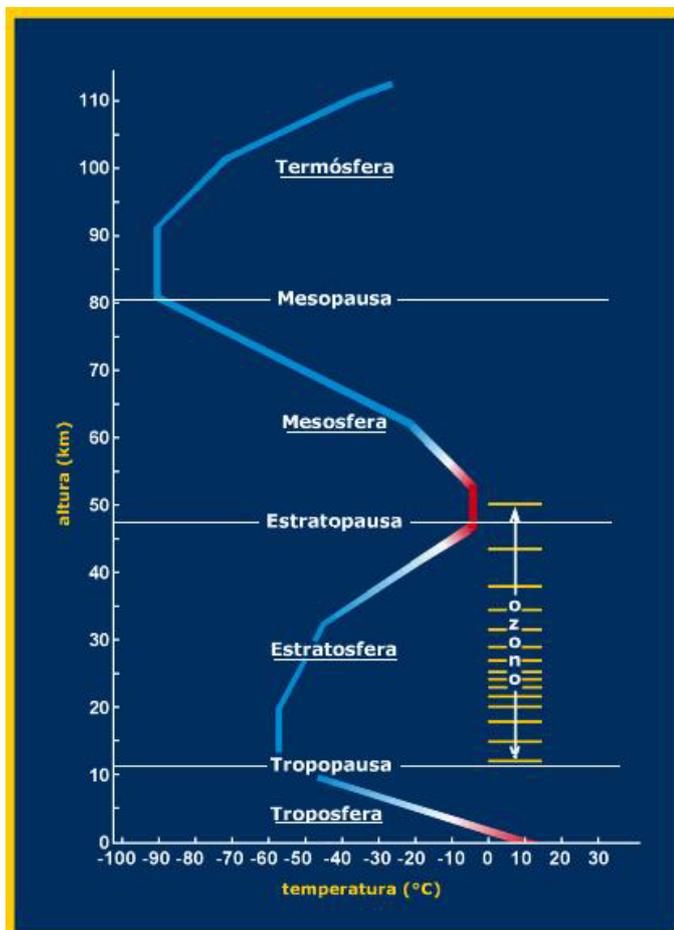


Figura 1.1: Representación gráfica de las capas de la atmósfera terrestre.

Las capas de la tierra se relacionan entre sí produciendo modificaciones profundas en sus características. La ciencia que estudia estas características, las propiedades y los movimientos de las tres capas fundamentales de la tierra, es la Geofísica. En ese sentido, la meteorología es una rama de la geofísica que tiene por objeto el estudio detallado de la envoltura gaseosa de la tierra y sus fenómenos.

Se debe distinguir entre las condiciones actuales y su evolución llamado tiempo atmosférico, y las condiciones medias durante un largo periodo que se conoce como clima del lugar o región. En este sentido, la meteorología es una ciencia auxiliar de la climatología ya que los datos atmosféricos obtenidos en múltiples estaciones meteorológicas durante largo tiempo se usan para definir el clima, predecir el tiempo, comprender la interacción de la atmósfera con otros subsistemas, etc. El conocimiento de las variaciones meteorológicas y el

impacto de las mismas sobre el clima ha sido siempre de suma importancia para el desarrollo de la agricultura, la navegación, las operaciones militares y la vida en general.

## 1.2. El clima

El clima abarca los valores estadísticos sobre los elementos del tiempo atmosférico en una región durante períodos que se consideran suficientemente representativos, de 30 años o más. Estos períodos conviene que sean más largos en las zonas subtropicales y templadas que en la zona intertropical, especialmente, en la faja ecuatorial, donde el clima es más estable y menos variable en lo que respecta a los parámetros meteorológicos. Los parámetros meteorológicos más importantes que integran el concepto de clima son temperatura, humedad, presión, vientos y precipitaciones. Estos valores se obtienen con la recopilación de forma sistemática y homogénea de la información meteorológica, durante períodos que se consideran suficientemente representativos.

### 1.2.1. Elementos del clima

- **Temperatura:** es el grado de calor específico del aire en un lugar y momento determinados. La temperatura de la atmósfera es función de la mayor o menor insolación o radiación solar.

La temperatura se mide con termómetros, los cuales pueden ser calibrados de acuerdo a una multitud de escalas que dan lugar a unidades de medición de la temperatura. En el Sistema Internacional de Unidades, la unidad de temperatura es el kelvin (K), y la escala correspondiente es la escala Kelvin o escala absoluta, que asocia el valor "cero kelvin" ( $0^\circ \text{K}$ ) al "cero absoluto", y se gradúa con un tamaño de grado igual al del grado Celsius. Sin embargo, fuera del ámbito científico el uso de otras escalas de temperatura es común. La escala más extendida es la escala Celsius, llamada "centígrada"; y, en mucha menor medida, y prácticamente sólo en los Estados Unidos, la escala Fahrenheit. También se usa a veces la escala Rankine ( $^\circ\text{R}$ ) que establece su punto de referencia en el mismo punto de la escala Kelvin, el cero absoluto, pero con un tamaño de grado igual al de la Fahrenheit, y es usada únicamente en Estados Unidos, y sólo en algunos campos de la ingeniería.

A continuación, se muestran tres termómetros con diferentes unidades de medición (Celsius, Fahrenheit, kelvin):

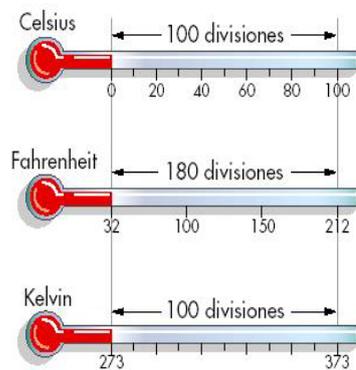


Figura 1.2: Termómetro.

- **Humedad:** es la cantidad de vapor de agua presente en el aire. El vapor procede de la evaporación de los mares y océanos, de los ríos, los lagos, las plantas y otros seres vivos. La cantidad de vapor de agua que puede absorber el aire depende de su temperatura. El aire caliente admite más vapor de agua que el aire frío.

El grado o cantidad de humedad de aire se mide con el higrómetro cuya unidad de medida se señala en porcentaje (%); cuando el higrómetro marca el 100 % se dice que el aire está saturado, es decir, contiene el máximo de humedad y es incapaz de admitir más vapor de agua. Un higrómetro es de la forma;

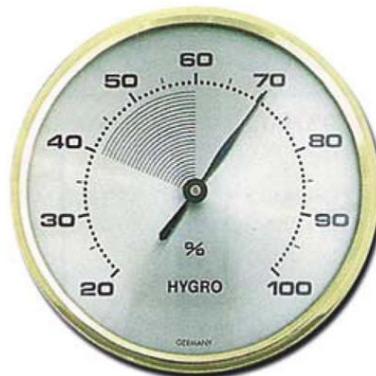


Figura 1.3: Higrómetro.

- **Presión atmosférica:** es la presión que ejerce el peso de las masas de aire en todas direcciones, además, varía con la altitud y con la temperatura.

La unidad de medida de la presión atmosférica que suelen marcar los barómetros se llama hectopascal, de abreviación hPa. Esta unidad significa: hecto: cien; pascales: unidad de medida de presión.

Su instrumento de medición es el barómetro.



Figura 1.4: Barómetro.

- **Viento:** es el movimiento de masas de aire de acuerdo con las diferencias de presión atmosférica. Además, constituye un elemento fundamental en el ciclo hidrológico que, a su vez, resulta imprescindible para sustentar la vida en la Tierra.

Es importante destacar, la medición de la velocidad y dirección del viento que se efectúa con instrumentos registradores llamados anemómetros, que disponen de dos sensores: uno para medir la velocidad y otro para medir la dirección del viento. Las mediciones se registran en anemógrafos.

La unidad de medición de la velocidad del viento es kilómetro por hora (Km/h) y la unidad de medición de la dirección del viento es metro por segundos (m/s).

La siguiente imagen muestra en la parte superior el anemómetro y en la parte inferior el anemógrafo;

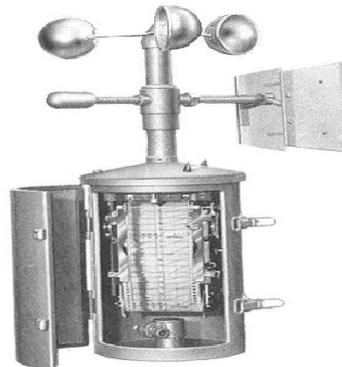


Figura 1.5: Anemógrafo.

- **Precipitación:** es cualquier forma de hidrometeoro que cae de la atmósfera y llega a la superficie terrestre. por medio de las precipitaciones.

La precipitación pluvial se mide en milímetros (mm), que equivale al espesor de la lámina de agua que se formaría, a causa de la precipitación, sobre una superficie plana e impermeable.

Los instrumentos más frecuentemente utilizados para la medición de la lluvia y el granizo son los pluviómetros y pluviógrafos, estos últimos se utilizan para determinar las precipitaciones pluviales de corta duración y alta intensidad. Estos instrumentos deben ser instalados en locales apropiados donde no se produzcan interferencias de edificaciones, árboles, o elementos orográficos como rocas elevadas. Cabe destacar que, en los pluviógrafos se registran las mediciones efectuadas por los pluviómetros.



Figura 1.6: Pluviógrafos.

- **Evaporación:** es un proceso físico que consiste en el paso lento y gradual de un estado líquido hacia un estado gaseoso.

La evaporación es medida en mililitros (ml) y su instrumento de medición se denomina evaporímetro.

El evaporímetro es un tubo de vidrio cerrado por un extremo y abierto por el otro, se llena de agua pura o de lluvia y se tapa el extremo abierto con un disco de papel secante (ostia de papel) sujeto por una arandela y se instala dentro del abrigo meteorológico colgado con la boca hacia abajo, de él se produce la evaporación; y es creciente de arriba hacia abajo; como puede apreciarse en la figura 1.7.



Figura 1.7: Evaporímetro.

- **Nubosidad:** es la cantidad de nubes en la atmósfera y se expresa en octas, u octavos de la bóveda celeste. Ésta es dividida en 8 partes por el operador, quien evalúa el número de esas partes que están cubiertas por las nubes. De este modo se puede estimar el rango de visibilidad del observador. El principal problema asociado con este método consiste en no poder hacer mediciones bajo condiciones de visibilidad muy baja (por ejemplo, en caso de niebla) o la dificultad para estimar la correcta cobertura fraccional del cielo por nubes cercanas al horizonte visual. Gracias a los satélites meteorológicos es posible calcular la nubosidad con mucha más precisión, aunque aun así algunas nubes muy delgadas pueden escapar de la detección satelital.
- **Radiación solar:** es la energía emitida por el sol, que se propaga en todas las direcciones a través del espacio mediante ondas electromagnéticas. Esta energía es el motor que determina la dinámica de los procesos atmosféricos y el clima. La energía procedente del sol es radiación electromagnética proporcionada por las reacciones del hidrógeno en el núcleo del sol por fusión nuclear y emitida por la superficie solar. El sol emite energía en forma de radiación de onda corta. Después de pasar por la atmósfera, donde sufre un proceso de debilitamiento por la difusión, reflexión en las nubes y de absorción por las moléculas de gases (como el ozono y el vapor de agua) y por partículas en suspensión, la radiación solar alcanza la superficie terrestre oceánica y continental que la refleja o la absorbe. La cantidad de radiación absorbida por la superficie es devuelta en dirección al espacio exterior en forma de radiación de onda larga, con lo cual se transmite calor a la atmósfera.

La radiación solar se mide normalmente con un instrumento denominado piranómetro.



Figura 1.8: Piranómetro.

Los tipos de radiación existentes son:

- **Radiación ultravioleta:** es la radiación de menor longitud de onda (360 nanómetro), la cual lleva

mucha energía e interfiere con los enlaces moleculares. Estas ondas son absorbidas por la parte alta de la atmósfera, especialmente por la capa de ozono.

Es importante protegerse de este tipo de radiación, ya que actúa sobre el ADN y esta asociada con el cáncer de piel. Sólo las nubes tipo cúmulos de gran desarrollo vertical atenúan éstas radiaciones prácticamente a cero.

- **Radiación visible:** es la radiación con una longitud de onda que varía entre los 380nm y los 780 nm, por la energía que lleva, es una radiación que nuestros ojos pueden procesar y hacen que sea visible. Dicha radiación no es nociva aunque en cantidades extremas puede provocar deslumbramiento y fotofobia.
- **Radiación infrarroja:** es la radiación que corresponde a longitudes de onda más largas (760nm), y lleva poca energía asociada. Su efecto aumenta la agitación de las moléculas, provocando el aumento de la temperatura. El CO<sub>2</sub>, el vapor de agua y las pequeñas gotas de agua que forman las nubes absorben con mucha intensidad las radiaciones infrarrojas.

### 1.2.2. Factores del clima

Hay una serie de factores que pueden influir sobre estos elementos climáticos; estos son:

- **Latitud:** determina la inclinación con la que caen los rayos del Sol y la diferencia de la duración del día y la noche. Cuanto más directamente incide la radiación solar, más calor aporta a la tierra. Las variaciones de la insolación que recibe la superficie terrestre se deben a los movimientos de rotación (variaciones diarias) y de traslación (variaciones estacionales).
  - **Altitud:** la altura del relieve modifica sustancialmente el clima, en especial en la zona intertropical, donde se convierte en el factor modificador del clima de mayor importancia. Este hecho ha determinado un criterio para la conceptualización de los pisos térmicos, que son fajas climáticas delimitadas por curvas de nivel que generan también curvas de temperatura (isotermas) que se han establecido tomando en cuenta tipos de vegetación, temperaturas y orientación del relieve. Se considera la existencia de cuatro pisos térmicos en la zona intertropical:
    1. **Macrotérmico** (menos de 1 km de altura), con una temperatura que varía entre los 27°C al nivel del mar y los 20°C.
    2. **Mesotérmico** (1 a 3 km): presenta una temperatura entre los 10 y 20° C, su clima es templado de montaña.
-

3. **Microtérnico** (3 a 4,7 km): su temperatura varía entre los 0 y 10° C. Presenta un tipo de clima de páramo o frío.
  4. **Gélido** (más de 4,7 km): su temperatura es menor de 0°C y le corresponde un clima de nieves perpetuas.
- **Orientación del relieve:** la más importante disposición de las cordilleras con respecto a la incidencia de los rayos solares determina dos tipos de vertientes o laderas montañosas: de solana y de umbría.

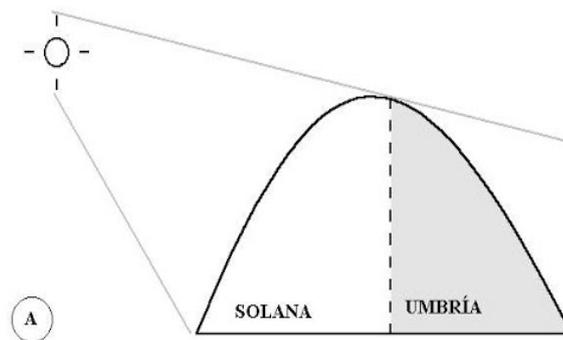


Figura 1.9: Vertientes solana y umbría.

Al norte del Trópico de Cáncer, las vertientes de solana son las que se encuentran orientadas hacia el sur, mientras que al sur del Trópico de Capricornio las vertientes de solana son, obviamente, las que están orientadas hacia el norte. En la zona intertropical, las consecuencias de la orientación del relieve con respecto a la incidencia de los rayos solares no resultan tan marcadas, ya que una parte del año el sol se encuentra incidiendo de norte a sur y el resto del año en sentido inverso.

La orientación del relieve con respecto a la incidencia de los vientos dominantes (los vientos planetarios) también determina la existencia de dos tipos de vertientes: de barlovento y de sotavento.

Barlovento es la ladera de un relieve o región, orientada hacia la dirección del viento. Habitualmente la ladera de barlovento es más húmeda, ya que el aire se ve impulsado a ascender, al hacerlo se enfría y se produce la precipitación.

Por otra parte, sotavento es la ladera o lado de un relieve, protegido del viento dominante, generalmente más seca que la ladera de barlovento, pues las masas de aire ya han descargado la humedad en la ladera de barlovento.

Por ende, llueve mucho más en las vertientes de barlovento porque el relieve da origen a las lluvias orográficas, al forzar el ascenso de las masas de aire húmedo.



Figura 1.10: Vertientes barlovento y sotavento.

- **Masas de agua:** son las que implican una amplia porción de agua singularizada por su temperatura, salinidad y densidad. Las masas de agua se configuran como resultado de los intercambios mar-aire y por las mezclas de aguas de distinta procedencia. La configuración de cada masa no es permanente, sino que se modifica progresivamente.

Las características que tiene el agua y, en particular, la temperatura y salinidad son adquiridas mientras se encuentra a la superficie bajo la influencia de los procesos que afectan esas propiedades (radiación, evaporación, precipitación). Estas características se modifican al mezclarse con otras masas de agua, dejan la superficie para hundirse y difundirse, según su densidad.

Se puede afirmar que las distintas masas de agua se forman de dos maneras:

- Por procesos físicos meteorológicos en la superficie (enfriamiento, calentamiento por radiación, evaporación o precipitación, fusión de hielo, etc.).
- Por mezcla subsuperficial de las masas de agua que se transportan por efecto de las corrientes.

Hay varios tipos de masas de agua, estas pueden ser dulces o saladas y estar en movimiento o estancadas. A menudo el tamaño de esas masas de agua se diferencian entre unas y otras, como pueden hacerlo sus límites. En algunos casos, se diferencian por la vegetación que crece dentro de ellas.

1. Hay tres tipos de **masas grandes de agua salada**. Los océanos son las más grandes, cubriendo las dos terceras partes del planeta. Son lo suficientemente grandes como para rodear los continentes.

Un mar es una gran zona de agua que en muchos casos se conecta con un océano, como el mar Caribe, que se extiende al sudeste de los Estados Unidos. En algunos casos, un mar de agua salada tiene tierra a sus lados, como el mar Caspio. La tierra puede rodear parcialmente un golfo, el cual contiene parte de un océano o un mar, el Golfo de México, en el sudeste de los Estados Unidos, es el principal ejemplo.

2. Las **masas de agua salada más pequeñas** representan bahías que son similares a los golfos, la tierra las rodea parcialmente, pero son más pequeñas. Las ensenadas, donde la tierra abarca al agua en forma de una herradura, contienen menos agua que las bahías. En agua salada, un canal es el tipo de masa de agua que une dos masas, como el Canal de la Mancha. Las lagunas son normalmente bastante poco profundas comparadas con el resto de un área de agua salada, y son cercanas a la costa. Algunas veces, las lagunas tienen un banco de arena o un trecho bajo de tierra entre ellas y el océano principal, por lo que en ellas no se producen olas. Los angostos hilos de agua que unen aguas más extensas son los estrechos.
  3. Las **aguas dulces en movimiento** están representadas por los arroyos los cuales son las masas de agua dulce más pequeñas. Con frecuencia fluyen en corrientes más grandes, que a su vez desembocan en los ríos. En la mayoría de los casos, los ríos desembocan en un océano, mar o lago. Las aguas más pequeñas que van hacia cualquier nivel de otras más grandes son los afluentes.
  4. Los **lagos** por lo general está compuesto por agua dulce y rodeado por tierra en todos sus lados.
  5. Los **humedales** están representados por los pantanos, turberas y marismas que también son masas de agua que se diferencian mayormente por las plantas que crecen en ellos. Un pantano contiene agua en sus límites y en él se desarrollan plantas leñosas como arbustos y árboles. Los pantanos tienen profundidades que van desde unos 5 centímetros hasta unos 30 centímetros, y algunas veces en la estación seca tienen poca cantidad o nada de agua. Las turberas se presentan en climas más fríos y contienen capas de turba y musgo, lo que las hace esponjosas al tacto. Las marismas tienen plantas de tallos más fuertes, como las totoras, y pueden ser de agua dulce o salada.
- **Distancia al mar:** es de los factores fundamentales que definen el clima ya que la lejanía de las grandes masas de agua dificulta que llegue aire húmedo hasta estas regiones. En estas regiones se observa un aumento de la amplitud térmica (diferencia entre la temperatura máxima y la mínima que se registra diariamente) y descenso de las precipitaciones debido a la lejanía de las masas de agua que suministran la humedad necesaria para las lluvias, además del hecho de que las zonas del interior de los continentes son zonas de alta presión o anticiclones donde el aire más pesado tiende a bajar, especialmente durante
-

la noche, manteniendo el aire subsidente <sup>1</sup>, bastante frío aunque muy seco, como puede verse en las regiones desérticas del interior de los continentes, que tienen una enorme amplitud térmica entre el día y la noche: en un mismo día se puede pasar desde el punto de congelación hasta los 40° C o más.

### 1.2.3. Sistema climático.

El sistema climático se considera formado por cinco elementos o cinco subsistemas. La atmósfera (la capa gaseosa que envuelve la tierra), la hidrosfera (el agua dulce y salada en estado líquido de océanos, lagos, ríos y agua debajo de la superficie), la criósfera (el agua en estado sólido), la litosfera (el suelo y sus capas) y la biosfera (el conjunto de seres vivos que habitan la tierra). El clima es consecuencia del equilibrio que se produce en la interacción entre esos cinco componentes.

Debido a que las actividades humanas son de extrema importancia para el estudio del cambio climático, están separadas de la biomasa y son estudiadas de forma individual. Así mismo, si se considera la actividad solar y las actividades humanas, se habla de un gran sistema llamado Sistema Global.

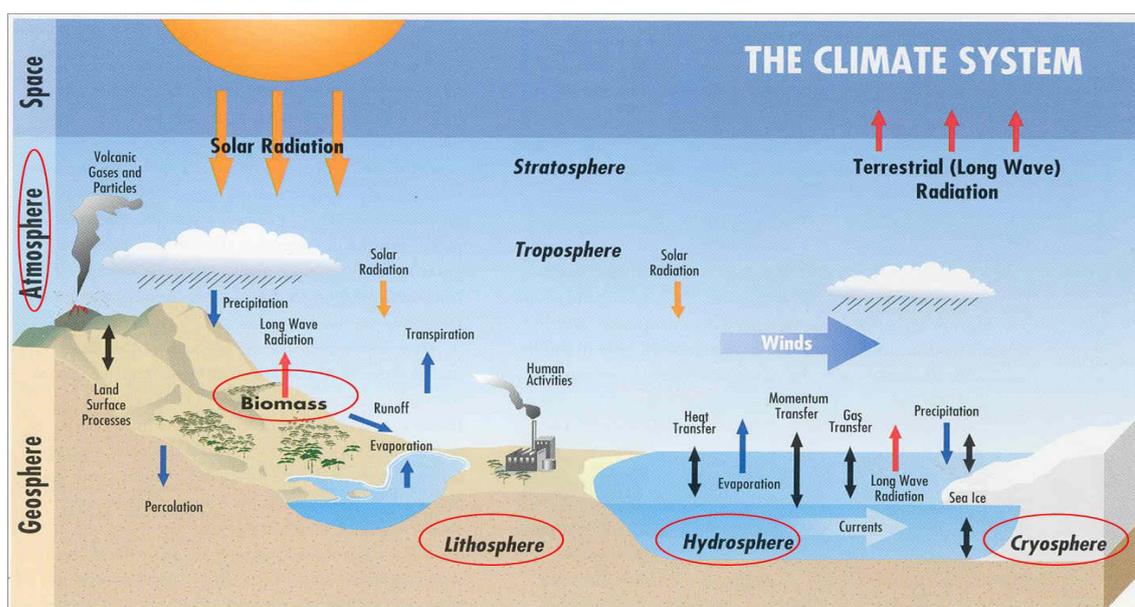


Figura 1.11: Sistema climático.

<sup>1</sup>“Subsidencia se define como el cambio de temperatura en una capa de aire. Es decir, en una capa de aire la parte superior posee una temperatura mas alta que en la base.”

## 1.3. Patrones de circulación

### 1.3.1. Corrientes Oceánicas

El agua salada tiene unas propiedades únicas que la distinguen de otros fluidos. Dentro de las propiedades físicas más importantes están su alto calor específico, su leve conducción de calor y la gran capacidad de disolución. En gran medida estas propiedades dependen de la temperatura, salinidad y presión.

La temperatura promedio del océano es de aproximadamente  $17.5^{\circ}\text{C}$ . La temperatura máxima es de  $36^{\circ}\text{C}$  en el Mar Rojo y la mínima es de  $(-2)^{\circ}\text{C}$  en el Mar de Weddell en la Antártida. La distribución de temperatura de las aguas depende de la radiación solar y de la mezcla de las masas de agua en el océano.

Las aguas cálidas superficiales transmiten el calor a las aguas próximas debajo de ellas formando una zona de productividad, aproximadamente de 200-400 m. A los 1000-1800 m la temperatura disminuye gradualmente y bajo los 1800 m el agua se mantiene fría.

La salinidad de la superficie del agua depende mayormente de la evaporación y la precipitación. En zonas tropicales donde la evaporación es mayor que la precipitación encontramos agua de mayor salinidad a  $> 35\%$ . En las regiones costeras, el agua dulce desemboca cerca de las bocas de los ríos y la salinidad generalmente no excede de  $15 - 20\%$ . En las zonas de los polos, el proceso de congelamiento y derretimiento de los hielos ejerce mayor influencia sobre la salinidad de las aguas superficiales. En el verano del Ártico, encontramos las salinidades más bajas aproximadamente  $29\%$ .

La salinidad promedio del océano es de  $35\%$  pero ésta puede variar dependiendo de la estación, la latitud y la profundidad. En conjunto, la temperatura y la salinidad afecta la densidad del agua. A su vez, la densidad afecta muchos otros parámetros como los procesos de mezcla de las diferentes masas de agua y la transmisión de sonido. Aguas estratificadas evitan la mezcla del agua superficial con el agua de la profundidad, mientras que aguas poco estratificadas favorecen la mezcla.

El calor se mueve por conducción, convección y radiación. La radiación y la conducción son efectivos en la transmisión vertical del calor desde la superficie de la tierra, pero son ineficientes en el sentido horizontal. El agua, al igual que el aire, es un fluido que puede transmitir calor de un lugar a otro. Los meteorólogos tienen diferentes términos para los movimientos horizontales y verticales del fluido, el movimiento en dirección vertical se llama convección hacia arriba y subsistencia hacia abajo, al movimiento en la dirección horizontal se le llama advección. La convección contribuye, junto con la radiación y la conducción, al movimiento del calor en dirección vertical, pero la advección es casi el único proceso que contribuye al transporte horizontal del calor sobre la superficie de la tierra.

---

El agua es cerca de 1000 veces más densa que el aire. Debido al hecho de que la cantidad de energía térmica transportada por un fluido en movimiento es proporcional a su densidad, un volumen de agua transporta cerca de mil veces más calor que el mismo volumen de aire. La tasa de transporte de calor se llama flujo de calor, y es medido en Joules de energía por unidad de área y tiempo, o sea que la tasa a la que éste calor es transportado es proporcional a la rapidez del movimiento (la velocidad del aire en la atmósfera o la velocidad de la corriente en el océano). Debido a que la velocidad del viento es típicamente del orden de 10 m/s y las corrientes de deriva son del orden de los centímetros por segundo, la velocidad del aire es miles de veces más grande que la velocidad de la corriente. Por lo tanto, el aire se mueve miles de veces más rápido que el agua, pero transporta solamente 1/1000 del calor por unidad de volumen, lo que sugiere que el agua es tan importante como el aire en el movimiento del calor alrededor del planeta.

Hay dos tipos de corrientes en el océano: las corrientes superficiales, que constituyen el 10% del agua del océano y se encuentran desde los 400 m hacia arriba y las corrientes de agua profunda o la circulación termohalina que afectan el otro 90% del océano.

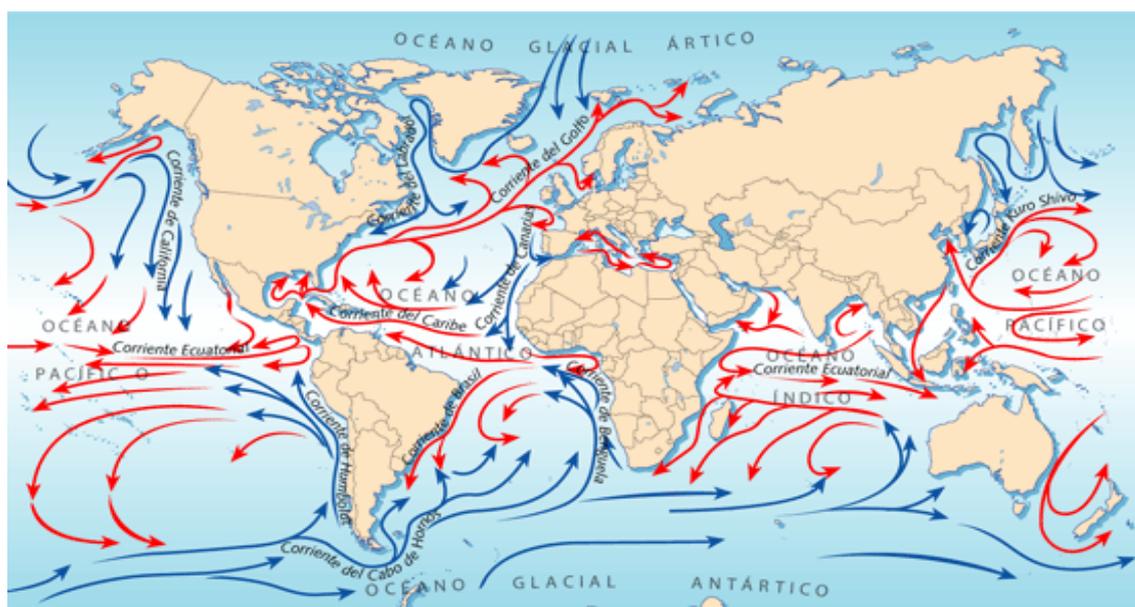


Figura 1.12: Corrientes oceánicas.

### 1.3.2. Patrones de viento

Debido a la rotación de la tierra, todo lo que se mueve en su superficie no sigue una línea recta. Los vientos son los responsables de producir las olas y las corrientes en el océano; a su vez, es el calentamiento solar lo que impulsa los vientos. La mayor energía solar se recibe en el Ecuador, por eso el aire es más caliente en el Ecuador y más frío en los polos. El aire caliente, por ser menos denso, se eleva en el Ecuador, por lo que se forma una baja presión; según, el aire caliente se aleja del Ecuador hacia el norte o hacia el sur, se enfría y se torna más denso y baja, esto ocasiona un gradiente de presión y otra masa de aire tiene que reemplazarlo, ocasionando el viento; entonces, se forma una celda de circulación o de convección.

Cuando el aire caliente del Ecuador asciende se forman las calmas ecuatoriales y al ser reemplazado por aire de latitudes más altas, se forman los vientos alisios; estos, soplan del noreste y sureste desde las altas presiones subtropicales hacia las bajas presiones tropicales del Ecuador. Los otros vientos también son impulsados por la energía solar, pero tienden a ser más variables que los alisios.

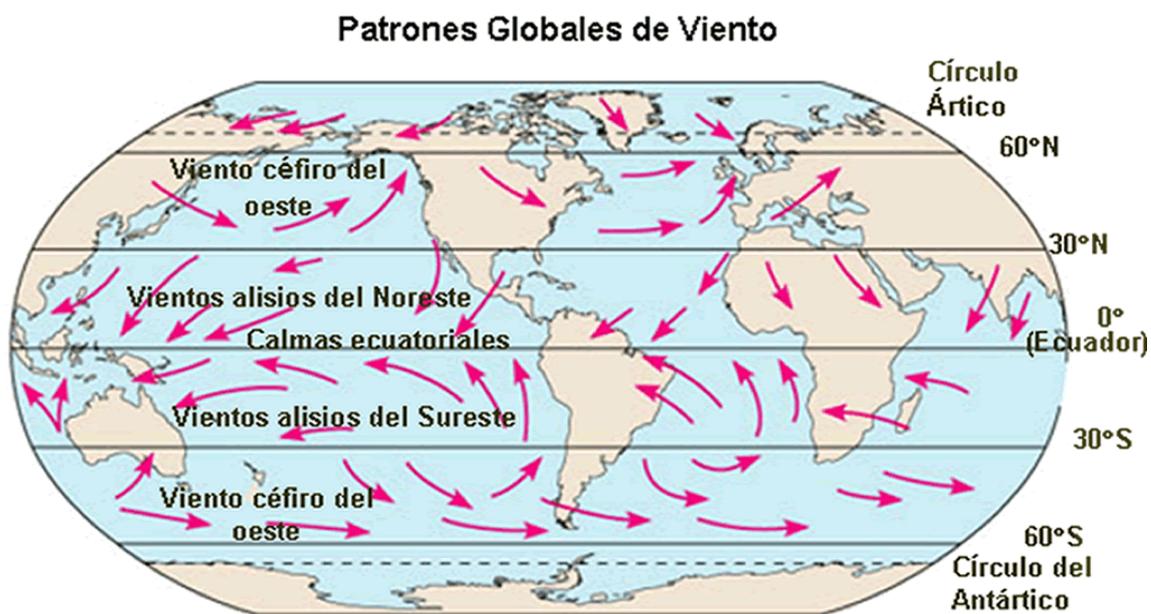


Figura 1.13: Patrones de viento.

## Series temporales.

**Definición 2.0.1** *Una serie de tiempo es un conjunto de observaciones  $x_t$ , cada una registrada a un tiempo específico  $t$ .*

**Definición 2.0.2** *Un modelo de series de tiempo para los datos observados  $x_t$  es una especificación de una distribución conjunta (o posiblemente solo de medias y covarianzas) de una sucesión de variables aleatorias  $x_t$  de las cuales  $x_t$  es una realización.*

### 2.0.3. Modelos estadísticos para series de tiempo.

El objetivo primario en el análisis de Series de Tiempo es desarrollar modelos matemáticos que provean una descripción apropiada para los datos muestrales. Por ende, lo primero que hacemos es utilizar las dos últimas definiciones, para tener un soporte estadístico.

**Definición 2.0.3** *Un proceso estocástico es una familia de variables aleatorias indexadas  $x(w, t)$  ó  $x_t(w)$  donde  $t$  pertenece a un conjunto de índices  $T$  y  $w$  pertenece a un espacio muestral  $\Omega$ . Si  $t = t$  fijo,  $x(w, t)$  es una variable aleatoria. Si  $w = w$  fijo,  $x(w, t)$  es una función de  $t$ , y se llama una realización del proceso. Una serie de tiempo es la realización de un proceso estocástico.*

**Definición 2.0.4** *El ruido blanco o sonido blanco es una señal aleatoria que se caracteriza por el hecho de que sus valores de señal en dos tiempos diferentes no guardan correlación estadística. Como consecuencia de ello, su densidad espectral de potencia es una constante, es decir, su gráfica es plana. Esto significa que la señal contiene todas las frecuencias y todas ellas muestran la misma potencia. Igual fenómeno ocurre con la luz blanca, de allí la denominación.*

Una manera sencilla de generar series de tiempo puede ser considerando una sucesión de variables aleatorias no correlacionadas,  $w_t$  con media 0 y varianza  $\sigma_t^2$ . Las series de tiempo generadas de esta manera son usadas como modelos para ruido en aplicaciones de ingeniería, donde ellas son llamadas ruido blanco, denotaremos este proceso como  $w_t \sim wn(0, \sigma_t^2)$ . La designación blanco se origina de la analogía con luz blanca e indica que todos los posibles periodos de oscilación están presente con igual intensidad.

También se requerirá que el ruido sea una colección de variables aleatorias iid con media 0 y varianza  $\sigma_t^2$ . Distinguiremos este caso diciendo que es ruido blanco independiente o escribiendo  $w_t \sim idd(0, \sigma_t^2)$ . Un muy usado ruido blanco es el ruido blanco gaussiano, donde las  $w_t$  son variables aleatorias normales con media 0 y varianza  $\sigma_t^2$  e identificadas como  $w_t \sim iddN(0, \sigma_t^2)$ .

**Definición 2.0.5** El **promedio móvil o media móvil** es un método utilizado para analizar un conjunto de datos en modo de puntos para crear series de promedios. Así las medias móviles son una lista de números en la cual cada uno es el promedio de un subconjunto de los datos originales.

Una serie de medias móviles puede ser calculada para cualquier serie temporal. Se usa para demanda estable, sin tendencia ni estacionalidad; suaviza las fluctuaciones de plazos cortos, resaltando así las tendencias o ciclos de plazos largos.

Podemos reemplazar las series de ruido blanco  $w_t$  por un promedio móvil que suavice las series. Por **ejemplo**, consideremos la serie mencionada en la definición anterior y reemplacémosla por un promedio móvil como el promedio del valor actual y los vecinos inmediatos anterior y posterior. Esto es:

$$v_t = \frac{1}{3}(w_{t-1} + w_t + w_{t+1}) \quad (2.1)$$

la cual nos da una serie suavizada como se muestra en la siguiente figura. En la parte superior se observa un ruido blanco gaussiano de 500 variables aleatorias generadas con  $\sigma_w^2 = 1$  y en la parte inferior el promedio móvil de tres puntos para la misma serie.

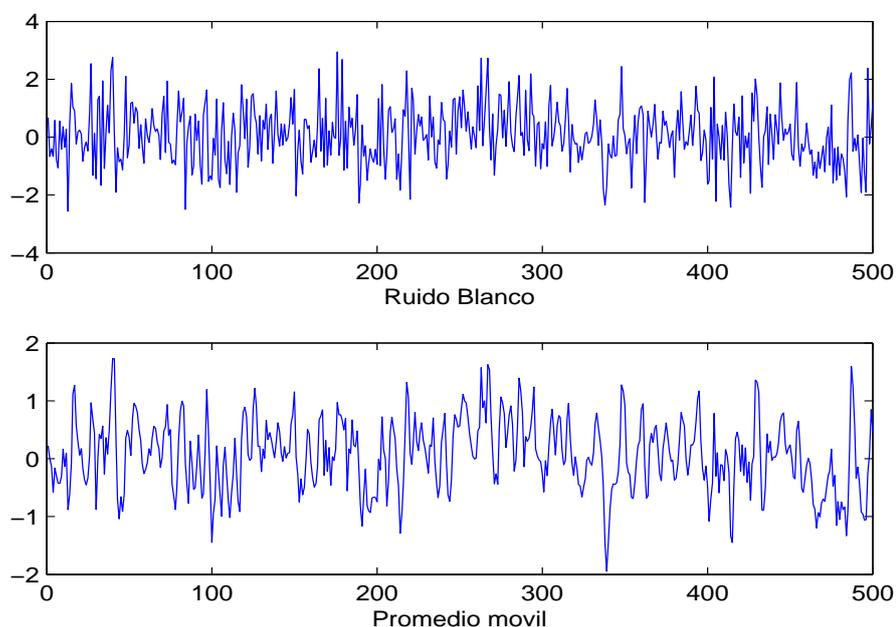


Tabla 2.1: Ruido blanco gaussiano (parte superior) y promedio móvil de 3 puntos (parte inferior).

**Definición 2.0.6** Un *camino aleatorio* es una formalización matemática de la trayectoria que resulta de hacer sucesivos pasos aleatorios. se modela mediante la siguiente expresión:

$$x(t + \tau) = x(t) + \phi(\tau) \quad (2.2)$$

Donde  $x(t)$  define una trayectoria que empieza en la posición  $x(0) = x_0$ ,  $\phi$  es la variable aleatoria que describe la ley de probabilidad para tomar el siguiente paso y  $\tau$  es el intervalo de tiempo entre pasos subsecuentes. A medida que la longitud y dirección de un paso dado depende solo de la posición  $x(t)$  y no de alguna posición previa, se dice que el camino aleatorio posee la Propiedad de Márkov. Comúnmente la distribución del paso será independiente de la posición o del tiempo transcurrido, una propiedad llamada homogeneidad. De cualquier modo, la formulación es extremadamente general. Los paseos aleatorios pueden ocurrir en cualquier número de dimensiones, ser parciales o imparciales, discretos o continuos en el tiempo y/o espacio, y pueden violar la homogeneidad en algún número de formas.

### 2.0.4. Medidas de dependencia: autocorrelación y correlación cruzada.

Una descripción completa de una serie de tiempo, observada como una colección de  $n$  variables aleatorias en puntos de tiempo enteros arbitrarios  $t_1, t_2, \dots, t_n$ , para cada entero positivo  $n$ , es proporcionada por la función de distribución conjunta, evaluada como la probabilidad de que los valores de la serie sean conjuntamente menor que  $n$  constantes  $c_1, c_2, \dots, c_n$ , esto es:

$$F(c_1, c_2, \dots, c_n) = P(x_{t_1} \leq c_1, x_{t_2} \leq c_2, \dots, x_{t_n} \leq c_n). \quad (2.3)$$

Desafortunadamente, la función de distribución multidimensional usualmente no se puede escribir fácilmente a menos que las variables aleatorias tengan distribución normal conjunta, en cuyo caso, la ecuación anterior llega a ser la distribución normal multivariada usual.

Un caso particular en el cual la función de distribución multidimensional es fácil de escribir, será el caso de variables aleatorias normal estándar independientes e idénticamente distribuidas, para lo cual la función de distribución se puede expresar como el producto de las distribuciones marginales, es decir,

$$F(c_1, c_2, \dots, c_n) = \prod_{t_1}^n \Phi(c_t) \quad (2.4)$$

donde,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^x \exp\left(\frac{-z^2}{2}\right) dz \quad (2.5)$$

es la función de distribución normal estándar acumulada.

Aunque la función de distribución multidimensional describa los datos completamente, esto es un instrumento poco manejable para mostrar y analizar datos de series de tiempo. La función de distribución debe ser evaluada como una función de  $n$  argumentos, entonces cualquier graficación de las correspondientes funciones de densidad multivariante es prácticamente imposible. La función de distribución unidimensional

$$F_t(x) = P(x_t \leq x) \quad (2.6)$$

o la correspondiente función de densidad unidimensional

$$f_t = \frac{\partial F_t(x)}{\partial x} \quad (2.7)$$

cuando existen, a menudo son más útiles para determinar si una coordenada en particular de la serie de tiempo tiene una función de densidad conocida, como la distribución normal (gaussiana), por ejemplo.

**Definición 2.0.7** La *función de media* es definida como:

$$\mu_{x_t} = E(x_t) = \int_{-\infty}^{\infty} x f_t dx \quad (2.8)$$

en caso de que exista, donde  $E$  denota el operador usual de esperanza. Cuando no haya confusión sobre a que serie de tiempo nos referimos, escribiremos  $\mu_{x_t}$  como  $\mu_t$ .

Lo importante de comprender sobre  $\mu_t$  consiste en que es una media teórica para la serie de tiempo en un punto particular, donde la media se asume o calcula sobre todos los posibles eventos que podrían haber producido  $x_t$ .

**Definición 2.0.8** La *función de autocovarianza* es definida como producto del segundo momento

$$\gamma_x(s, t) = E[(x_s - \mu_s)(x_t - \mu_t)] \quad (2.9)$$

para todo  $t$  y  $s$ . Cuando no haya confusión en la existencia sobre a que serie nos referimos, escribiremos

$$\gamma_x(s, t) = \gamma(s, t).$$

Note que  $\gamma_x(s, t) = \gamma_x(t, s)$  para todo los puntos  $s$  y  $t$ . La función de autocovarianza mide la dependencia lineal entre dos puntos de la misma serie en diferentes tiempos. La autocovarianza es el promedio de los productos cruzados relacionado con la densidad conjunta  $F(x_s, x_t)$ . Es claro que, para  $s = t$ , la autocovarianza se reduce a la varianza (en el caso finito), dado que

$$\gamma_x(t, t) = E[(x_t - \mu_t)^2]. \quad (2.10)$$

**Definición 2.0.9** La *autocorrelación* de una serie temporal discreta de un proceso  $x_t$  no es más que simplemente la correlación de dicho proceso con una versión desplazada en el tiempo de la propia serie temporal.

Esto es:

$$R(k) = \frac{E[(x_i - \mu)(x_{i-k} - \mu)]}{\sigma^2} \quad (2.11)$$

Donde  $E$  es el valor esperado y  $k$  el desplazamiento temporal considerado (normalmente denominado desfase o paso). Esta función varía dentro del rango  $[-1, 1]$ , donde 1 indica una correlación perfecta (la señal se superpone perfectamente tras un desplazamiento temporal de  $k$ ) y  $-1$  indica una correlación inversa perfecta.

**Definición 2.0.10** La *función de autocorrelación (ACF)* se define como

$$\rho(s, t) = \frac{\gamma(s, t)}{\sqrt{\gamma(s, s)\gamma(t, t)}}. \quad (2.12)$$

La ACF mide la predictibilidad lineal de una serie de tiempo en tiempo  $t$ , digamos  $x_t$  usando solo el valor  $x_s$ . Es fácil de demostrar que  $-1 \leq \rho(s, t) \leq 1$  usando la desigualdad de Cauchy-Schwarz.

Si podemos predecir  $x_t$  exactamente de  $x_s$  a través de la relación lineal  $x_t = \beta_0 + \beta_1 x_s$  entonces la correlación será 1 cuando  $\beta_1 > 0$  y  $-1$  cuando  $\beta_1 < 0$ .

**Definición 2.0.11** La *función de covarianza cruzada* entre dos series  $x_t$  y  $y_t$  se define como

$$\gamma_{xy}(s, t) = E[(x_s - \mu_{x_s})(y_t - \mu_{y_t})]. \quad (2.13)$$

**Definición 2.0.12** La *función de correlación cruzada (CCF)* es definida como

$$\rho_{xy}(s, t) = \frac{\gamma_{xy}(s, t)}{\sqrt{\gamma_x(s, s)\gamma_y(t, t)}}. \quad (2.14)$$

**Definición 2.0.13** La *correlación cruzada o covarianza cruzada* es una medida de la similitud entre dos señales, frecuentemente usada para encontrar características relevantes en una señal desconocida por medio de la comparación con otra que sí se conoce. Es función del tiempo relativo entre las señales, a veces también se la llama producto escalar desplazado, y tiene aplicaciones en el reconocimiento de patrones y en criptoanálisis.

## 2.1. Series de tiempo estacionarias.

**Definición 2.1.1** Una *serie de tiempo estrictamente estacionaria* es una serie para la cual el comportamiento probabilístico de cada sucesión de valores

$$x_{t_1}, x_{t_2}, \dots, x_{t_k} \quad (2.15)$$

es idéntico a la serie trasladada en el tiempo:

$$x_{t_1+h}, x_{t_2+h}, \dots, x_{t_k+h}. \quad (2.16)$$

Esto es,

$$P(x_{t_1} < c_1, x_{t_2} < c_2, \dots, x_{t_k} < c_k) = P(x_{t_1+h} < c_1, x_{t_2+h} < c_2, \dots, x_{t_k+h} < c_k) \quad (2.17)$$

para todo  $k = 1, 2, \dots$ , todo puntos de tiempos  $t_1, t_2, \dots, t_k$  y números  $c_1, c_2, \dots, c_k$  y todo salto  $h = 0, \pm 1, \pm 2, \dots$

**Definición 2.1.2** Una serie de tiempo débilmente estacionaria  $x_t$ , es un proceso de varianza finita tal que:

1. La función de media  $\mu_t$ , definida anteriormente es constante y no depende del tiempo  $t$ , y
2. La función de covarianza,  $\gamma(s, t)$ , definida anteriormente depende solo de las diferencias de  $s$  y  $t$ ,  $|s - t|$ .

Por consiguiente, usaremos el término **estacionaridad** para referirnos a estacionaridad débil; si un proceso es estacionario en el sentido estricto usaremos el término estrictamente estacionario.

Ya que la función de media  $E(x_t) = \mu_t$  de una serie de tiempo estacionaria es independiente del tiempo  $t$ , escribimos

$$\mu_t = \mu. \quad (2.18)$$

Debido a que la función de covarianza de una serie de tiempo estacionaria,  $\gamma(s, t)$  en tiempos  $s$  y  $t$  depende sólo de la diferencia  $|s - t|$ , podemos simplificar la notación. Sea  $s = t + h$ , donde  $h$  representa el tiempo de traslación o salto, entonces

$$\begin{aligned} \gamma(s, t) &= E[(x_{t+h} - \mu)(x_t - \mu)] \\ &= E[(x_h - \mu)(x_0 - \mu)] \\ &= \gamma(h, 0) \end{aligned} \quad (2.19)$$

no depende del argumento de tiempo  $t$ ; asumiendo que  $\text{var}(x_t) = \gamma(0, 0) < 1$ . De ahora en adelante, por conveniencia, prescindiremos del segundo argumento de  $\gamma(h, 0)$ , es decir, la función de covarianza se denotará  $\gamma(h)$ .

**Definición 2.1.3** La función de autocovarianza de una serie de tiempo estacionaria se escribirá como:

$$\gamma(h) = E[(x_{t+h} - \mu)(x_t - \mu)]. \quad (2.20)$$

**Definición 2.1.4** La función de autocorrelación (ACF) de una serie de tiempo estacionaria será escrita como,

$$\begin{aligned} \rho(h) &= \frac{\gamma(t+h, t)}{\sqrt{\gamma(t+h, t+h)\gamma(t, t)}} \\ &= \frac{\gamma(h)}{\gamma(0)}. \end{aligned} \quad (2.21)$$

La desigualdad de Cauchy-Schwarz muestra nuevamente que  $-1 \leq \rho(h) \leq 1$  para todo  $h$ .

▪ **Propiedades**

1. Para el valor en  $h = 0$ , la función de autocovarianza

$$\gamma(0) = E[(x_t - \mu)^2] \quad (2.22)$$

es la varianza de la serie de tiempo; note que la desigualdad de Cauchy-Schwarz implica que  $|\gamma(h)| \leq \gamma(0)$ .

2. La autocovarianza de una serie estacionaria es simétrica respecto al origen, esto es

$$\gamma(h) = \gamma(-h) \quad (2.23)$$

para todo  $h$ . Esta propiedad se debe a que trasladar la serie por  $h$  significa que

$$\begin{aligned} \gamma(h) &= \gamma(t+h-t) \\ &= E[(x_{t+h} - \mu)(x_t - \mu)] \\ &= E[(x_t - \mu)(x_{t+h} - \mu)] \\ &= \gamma(t - (t+h)) \\ &= \gamma(-h) \end{aligned} \quad (2.24)$$

lo cual muestra como usar la notación para demostrar el resultado.

**Definición 2.1.5** *Dos series de tiempo  $x_t$  y  $x_s$  se dice que son conjuntamente estacionarias si cada una de ellas es estacionaria y la función de correlación cruzada*

$$\gamma_{(xy)}(h) = E[(x_{t+h} - \mu_x)(y_t - \mu_y)]. \quad (2.25)$$

*es una función sólo del salto  $h$ .*

**Definición 2.1.6** *La función de correlación cruzada (CCF) de dos series conjuntamente estacionarias  $x_t$  y  $y_t$  se define como:*

$$\rho_{xy}(h) = \frac{\gamma_{xy}(h)}{\sqrt{\gamma_x(0)\gamma_y(0)}}. \quad (2.26)$$

De nuevo, tenemos el resultado  $-1 \leq \rho_{xy}(h) \leq 1$  lo cual nos permite comparar los valores extremos  $-1$  y  $1$  cuando vemos la relación entre  $x_{t+h}$  y  $y_t$ . La función de correlación cruzada satisface

$$\rho_{xy}(h) = \rho_{xy}(-h) \quad (2.27)$$

cuya demostración es:

$$\begin{aligned} \rho_{xy}(h) &= \rho(t+h-t) \\ &= E[(x_{t+h} - \mu)(x_t - \mu)] \\ &= E[(x_t - \mu)(x_{t+h} - \mu)] \\ &= \rho(t - (t+h)) \\ &= \rho_{xy}(-h) \end{aligned} \quad (2.28)$$

La CCF mide no sólo la fuerza de la relación, sino también su dirección. Esta última propiedad es útil para identificar variables causales. Por esta razón, en la CCF es importante examinar tanto los valores positivos de  $h$  como valores negativos. Para valores negativos de  $h$ , la CCF describe influencia lineal de valores pasados de  $y_t$  sobre  $x_t$ . Para valores positivos de  $h$ , la CCF indica la influencia lineal de los valores pasados de  $x_t$  sobre  $y_t$ . La gráfica de la CCF contra  $h$ , llamado correlograma cruzado es útil para visualizar estas relaciones. La CCF refleja los diversos componentes de frecuencia que tienen en común entre las dos series  $x_t$  y  $y_t$ . Además, también contiene información vital acerca de las fases relativas de los componentes de frecuencia compartidos.

**Definición 2.1.7** Un **proceso lineal**  $x_t$  se define como una combinación lineal de variables aleatorias de ruido blanco  $w_t$ , y está dado por

$$x_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j \omega_{t-j} \quad (2.29)$$

donde los coeficientes satisfacen

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty. \quad (2.30)$$

Para un proceso lineal, podemos demostrar que la función de autocovarianza está dada por

$$\gamma(h) = \sigma_w^2 \sum_{j=-\infty}^{\infty} \psi_{j+h} \psi_j \quad (2.31)$$

para todo  $h \geq 0$ ; recuerde que  $\gamma(-h) = \gamma(h)$ . Finalmente como mencionamos anteriormente, un caso importante en el cual una serie débilmente estacionaria es también estrictamente estacionaria es la serie normal o gaussiana.

**Definición 2.1.8** Un proceso  $x_t$ , se dice que es un **proceso gaussiano** si el  $k$ -ésimo vector dimensional  $\hat{x} = (x_{t_1}, x_{t_2}, \dots, x_{t_k})'$ , para cada conjunto de puntos  $t_1, t_2, \dots, t_k$  y cada entero positivo  $k$  tiene distribución normal multivariada.

Definiendo  $k \times 1$  vector de medias  $\hat{\mu} = (\mu_{t_1}, \mu_{t_2}, \dots, \mu_{t_k})'$  y la  $k \times k$  matriz de covarianza positiva como  $\Gamma = \gamma(t_i, t_j); i, j = 1, \dots, k$ , la función de densidad normal multivariada se puede escribir como:

$$f(\hat{x}) = (2\pi)^{-\frac{k}{2}} |\Gamma|^{-\frac{1}{2}} \exp -\frac{1}{2}(\hat{x} - \hat{\mu})' \Gamma^{-1} (\hat{x} - \hat{\mu}) \quad (2.32)$$

donde  $|\cdot|$  denota el determinante. Esta distribución forma la base para resolver problemas que envuelven inferencia estadística para series de tiempo. Si una serie de tiempo gaussiana  $x_t$  es débilmente estacionaria, entonces  $\mu_t = \mu$  y  $\gamma(t_i, t_j) = \gamma(|t_i - t_j|)$ , de modo que el vector  $\hat{\mu}$  y la matriz  $\Gamma$  son independientes del tiempo. Este hecho implica que todas las distribuciones finitas, de la serie  $x_t$  dependen sólo del salto de tiempo y no del tiempo actual, y por consiguiente la serie debe ser estrictamente estacionaria.

### 2.1.1. Estimación de correlación.

Aunque las funciones teóricas de autocorrelación y correlación cruzada son muy útiles para describir las propiedades de ciertos modelos hipotéticos, la mayoría de los análisis se realizan usando datos muestrales. Esta limitación significa que los puntos muestrales  $x_1, x_2, \dots, x_n$ , sólo nos permite estimar las funciones de media, autocovarianza y correlación.

Desde el punto de vista de la estadística clásica esto plantea un problema porque no tenemos copias iid de  $x_t$  que sean o estén disponibles para estimar las funciones de covarianza y correlación.

**Definición 2.1.9** Sea  $x_1, x_2, \dots, x_n$  una muestra de una serie de tiempo. La **media muestral** de  $x_1, x_2, \dots, x_n$  es

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t. \quad (2.33)$$

La **función de autocovarianza muestral** se define como

$$\hat{\gamma}(h) = n^{-1} \sum_{t=1}^n -h(x_{t+h} - \bar{x})(x_t - \bar{x}). \quad (2.34)$$

con  $\hat{\gamma}(-h) = \hat{\gamma}(h)$  para  $h = 0, 1, \dots, n - 1$ .

La **función de autocorrelación muestral** se define como

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} \quad (2.35)$$

La no-negatividad de la función de autocovarianza  $\gamma(h)$  es una propiedad importante porque nos asegura que la varianza de combinaciones lineales de valores de la serie de tiempo nunca serán negativa y la estimación en la función de autocovarianza muestral conserva las propiedades.

▪ **Propiedad 1. Distribución de la ACF para muestras grandes**

Bajos ciertas condiciones generales, si  $x_t$  es un ruido blanco, entonces para  $n$  grande, la ACF muestral  $\hat{\rho}_x(h)$  para  $h = 1, 2, \dots, H$ , donde  $H$  es un valor fijo arbitrario, es aproximadamente normal distribuida con media cero y desviación estándar dada por

$$\sigma_{\hat{\rho}(h)} = \frac{1}{\sqrt{n}}. \quad (2.36)$$

Basándonos en el resultado anterior, podemos obtener un método basto de evaluación de si los picos en  $\hat{\rho}(h)$  son significativos, por medio de determinar si los picos observados están fuera del intervalo  $\frac{\pm 2}{\sqrt{n}}$  (o mas o menos dos veces el error estándar); para un sucesión de ruido blanco aproximadamente el 95 % de la ACF muestral debe estar entre estos límites.

**Definición 2.1.10** El estimador para la función de covarianza cruzada  $\gamma_{xy}(h)$  es la **función de covarianza cruzada muestral** definida como

$$\hat{\gamma}_{xy}(h) = n^{-1} \sum_{t=1}^n -h(x_{t+h} - \bar{x})(y_t - \bar{y}) \quad (2.37)$$

donde  $\hat{\gamma}_{xy}(-h) = \hat{\gamma}_{yx}(h)$  determina la función para saltos negativos, y la **función de correlación cruzada muestral** es

$$\hat{\rho}_{xy}(h) = \frac{\hat{\gamma}_{xy}(h)}{\sqrt{\hat{\gamma}_x(0)\hat{\gamma}_y(0)}}. \quad (2.38)$$

▪ **Propiedad 2. Distribución de la correlación cruzada para muestras grandes con independencia**

La distribución de  $\hat{\rho}_{xy}(h)$  para muestras grandes es normal con media cero y

$$\sigma_{\hat{\rho}_{xy}} = \frac{1}{\sqrt{n}} \quad (2.39)$$

si al menos uno de los procesos es un ruido blanco independiente.

## 2.2. Regresión de series de tiempo.

Los modelos de regresión son importantes para modelos en el dominio de tiempo y de frecuencia que discutiremos posteriormente. La idea principal depende de poder expresar una serie respuesta  $x_t$  como una combinación lineal de entradas  $z_{t_1}, z_{t_2}, \dots, z_{t_q}$ . La estimación de los coeficientes  $\beta_1, \beta_2, \dots, \beta_q$  de la combinación por mínimos cuadrados proporciona un método para modelar  $x_t$  en términos de las entradas.

### 2.2.1. Regresión clásica

Supongamos que tenemos  $x_t$ , para  $t = 1, 2, \dots, n$  influenciada por una colección de series independientes  $z_{t_1}, z_{t_2}, \dots, z_{t_q}$ , donde consideraremos primero que las entradas son fijas y conocidas. Podemos expresar esta relación como

$$x_t = \beta_1 z_{t_1} + \beta_2 z_{t_2} + \dots + \beta_q z_{t_q} + w_t \quad (2.40)$$

donde  $\beta_1, \beta_2, \dots, \beta_q$  son los coeficientes de regresión fijos y desconocidos,  $w_t$  es un error aleatorio o un proceso de ruido consistente de variables normales *iid* con media cero y varianza  $\sigma_w^2$ .

**Definición 2.2.1** El Criterio de información de Akaike (AIC) está dado por

$$AIC = -2 \ln L_k + 2k \quad (2.41)$$

donde  $L_k$  es la verosimilitud maximizada y  $k$  es el número de parámetros del modelo.

Sin embargo, para el problema de regresión normal, AIC se reduce a

$$AIC = \ln \hat{\sigma}_k^2 + \frac{n + 2k}{n} \quad (2.42)$$

donde, el estimador de máxima verosimilitud para la varianza se expresa como:

$$\hat{\sigma}_k^2 = \frac{RSS_k}{n} \quad (2.43)$$

donde  $RSS_k$  denota la suma residual de cuadrados bajo el modelo con  $k$  coeficientes de regresión.

**Definición 2.2.2** El Criterio de información de Akaike (AIC) con sesgo corregido (AICc) está dado por:

$$AICc = \ln \hat{\sigma}_k^2 + \frac{n + k}{n - k - 2} \quad (2.44)$$

donde  $\hat{\sigma}_k^2$  está expresado como se muestra en la ecuación (2.43),  $k$  es el número de parámetros en el modelo y  $n$  es el tamaño de la muestra.

**Definición 2.2.3** El *Criterio de información de Schwarz (SIC)* está definido como:

$$SIC = \ln \hat{\sigma}_k^2 + \frac{k \ln n}{n} \quad (2.45)$$

usando la misma notación que en la definición 2.2.2. SIC también es llamado *Criterio Bayesiano de Información (BIC)* el cual es un criterio para la selección de modelos entre un conjunto finito de modelos. Se basa, en parte, de la función de probabilidad y que está estrechamente relacionado con el *Criterio de Información de Akaike (AIC)*.

### 2.3. El Modelo Aditivo de Componentes de Series de Tiempo.

Dada una serie  $X_t, t = 1, \dots, n$ , el Modelo Aditivo de Componentes consiste en asumir que  $X_t$  se puede descomponer en tres componentes:

$$X_t = T_t + E_t + \epsilon_t \quad (2.46)$$

donde  $T_t$  es la tendencia,  $E_t$  es la componente estacional y  $\epsilon_t$  es la componente aleatoria o de errores. Las componentes  $T_t$  y  $E_t$  son funciones de  $t$  determinísticas. Su evolución es perfectamente predecible.

La componente  $T_t$  en algunos casos también puede ser una componente estacional, pero de baja frecuencia, o, equivalentemente, una componente con período muy grande. Por ejemplo, en una serie diaria,  $E_t$  puede tener período 30 días, y  $T_t$  período 360 días.

El Modelo Multiplicativo consiste en asumir que  $X_t$  se puede descomponer en tres componentes:

$$X_t = T_t E_t \exp(\epsilon_t) \quad (2.47)$$

En la Figura 2.1 se muestra la idea de la descomposición. Al superponer las series en los gráficos (a), (b) y (c) se obtiene la serie en el gráfico (d).

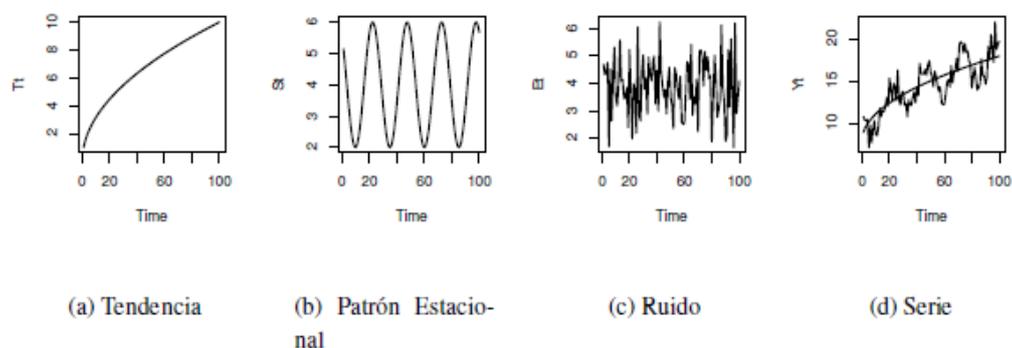


Figura 2.1: Descomposición de la serie  $X_t$

El análisis consiste en modelar y estimar  $T_t$  y  $E_t$  y luego extraerlas de  $X_t$  para obtener  $\hat{\epsilon}_t = X_t - \hat{T}_t - \hat{E}_t$ . La serie  $\hat{\epsilon}_t$  se modela y estima para finalmente reconstruir  $X_t$ ,  $\hat{X}_t = \hat{T}_t + \hat{E}_t + \hat{\epsilon}_t$ , y poder realizar el pronóstico  $\hat{X}_{t+h} = \hat{T}_{t+h} + \hat{E}_{t+h} + \hat{\epsilon}_{t+h}$ , utilizando la información disponible  $X_t, \dots, X_n$  con  $h = 1, 2, \dots, m$ . Sin embargo, puede suceder que la serie  $\hat{\epsilon}_t$  sea incorrelacionada, es decir,  $Corr(\hat{\epsilon}_t, \hat{\epsilon}_{t+s}) = 0$ , para  $s \neq 0$ . En este caso  $\hat{\epsilon}_{t+h} = 0$  para todo  $h > 0$ .

**Definición 2.3.1 (Tendencia).** Se define como una función  $T_t$  de  $t$  que describe la evolución lenta y a largo plazo del nivel medio de la serie. La función  $T_t$  depende de parámetros, que deben estimarse.

A continuación presentamos una lista de posibles modelos para la tendencia  $T_t$ :

1. Lineal

$$T_t = \beta_0 + \beta_1 t. \tag{2.48}$$

2. Cuadrático

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2. \tag{2.49}$$

3. Cúbico

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3. \tag{2.50}$$

4. Exponencial

$$T_t = \exp(\beta_0 + \beta_1 t). \tag{2.51}$$

5. Logístico

$$T_t = \frac{\beta_2}{1 + \beta_1 \exp(-\beta_0 t)}. \quad (2.52)$$

En la tendencia cuadrática podemos observar:

1. Si  $\beta_1, \beta_2 > 0$ ,  $T_t$  es monótona creciente.
2. Si  $\beta_1, \beta_2 < 0$ ,  $T_t$  es monótona decreciente.
3. Si  $\beta_1 > 0$  y  $\beta_2 < 0$ ,  $T_t$  es cóncava.
4. Si  $\beta_1 < 0$  y  $\beta_2 > 0$ ,  $T_t$  es convexa.

**Definición 2.3.2** El modelo *Logarítmico Lineal* o *Log-Lineal* se define como:

$$\ln X_t = \beta_0 + \beta_1 t + \epsilon_t. \quad (2.53)$$

Corresponde a un modelo con tendencia lineal para el logaritmo de  $X_t$ . En (2.53) al tomar exponencial se tiene  $X_t = \exp(\beta_0 + \beta_1 t + \epsilon_t)$ , que es similar al modelo con tendencia exponencial (2.51). Sin embargo, son modelos diferentes y se estiman por métodos diferentes.

## 2.4. Análisis exploratorio de datos.

En el análisis exploratorio de datos, quizás la forma más fácil de trabajar con series no-estacionarias es el modelo de tendencia estacionaria donde el proceso tiene comportamiento estacionario alrededor de una tendencia. Podemos escribir este tipo de modelos como

$$X_t = T_t + Y_t \quad (2.54)$$

donde  $X_t$  son las observaciones,  $T_t$  denota la tendencia y  $Y_t$  es un proceso estacionario.

Por lo general, una tendencia fuerte  $T_t$  puede oscurecer el comportamiento del proceso estacionario  $Y_t$ , como veremos en ejemplos posteriores. De aquí, será una ventaja el que podamos remover la tendencia como un primer paso para un análisis exploratorio de los datos. Los pasos envuelven obtener un estimador razonable del componente de tendencia, llamémoslo  $\hat{T}_t$  y entonces trabajar con el residual

$$\hat{Y}_t = X_t - \hat{T}_t. \quad (2.55)$$

El primer paso en el análisis de cualquier tipo de serie es un gráfico de los datos.

- Si existe alguna aparente discontinuidad en la serie, tal como un cambio súbito en el nivel de la serie, esto puede darnos una idea para el análisis de la serie, un primer paso sería dividir la serie en segmentos homogéneos.
- Si existen observaciones o datos "outliers", estos deben ser estudiados con cuidado para verificar si existe alguna justificación para descartar estas observaciones, como por ejemplo si una observación ha sido registrada de algún otro proceso por error.
- La inspección del gráfico también podría sugerir la representación de los datos como una realización de un proceso, como el modelo clásico de descomposición dado por (2.46).

Si la componente estacional y la componente aleatoria o ruido parecen incrementarse con el nivel del proceso entonces una transformación preliminar de los datos es a menudo usada para hacer que los datos transformados sean compatibles con el modelo (2.46). En esta sección discutiremos algunas técnicas para identificar y eliminar las componentes en (2.46).

Nuestro objetivo es estimar y extraer las componentes determinísticas  $T_t$  y  $E_t$  con la esperanza de que el residual o la componente aleatoria  $\epsilon_t$  llegue a ser un proceso estacionario. Entonces podremos usar la teoría de tales procesos para hallar un modelo probabilístico satisfactorio para el proceso  $\epsilon_t$ , analizar sus propiedades y usarlo en conjunto con  $T_t$  y  $E_t$  para hacer pronósticos y control de  $X_t$ .

Los dos enfoques para la eliminación de las componentes de tendencia y estacional son:

1. Estimación de  $T_t$  y  $E_t$  en el modelo (2.46),
2. Diferencia de los datos  $X_t$ .

Ilustraremos ambos enfoques con varios ejemplos.

### 2.4.1. Eliminación de la tendencia en ausencia de estacionalidad

En ausencia de la componente estacional  $E_t$  el modelo (2.46) llega ser

$$X_t = T_t + \epsilon_t, \quad t = 1, \dots, n \quad (2.56)$$

donde, sin pérdida de generalidad, podemos asumir que  $E(\epsilon_t) = 0$ .

**Método T1: Estimación de  $T_t$  por mínimos cuadrados.** En este procedimiento intentamos ajustar una familia paramétrica de funciones como vimos en la sección 2.3, a los datos eligiendo los parámetros que minimicen  $\sum_t (X_t - T_t)^2$ .

---

Ajustando una función de la forma (2.49) para la población de los datos en el gráfico 2.2 para  $1790 \leq t \leq 1980$  nos da los parámetros estimados

$$\hat{a}_0 = 2,0978; \quad \hat{a}_1 = 2,3349 \times 10^{-3}; \quad \hat{a}_2 = 6,4984 \times 10^{-7}.$$

En la figura 2.2 se puede observar la curva ajustada y los datos originales. Los valores estimados del proceso de ruido  $\epsilon_t$ ,  $1790 \leq t \leq 1980$ , son los residuales obtenidos por sustracción de  $\hat{T}_t = \hat{a}_0 + \hat{a}_1 t + \hat{a}_2 t^2$  de la serie  $X_t$ . La componente de tendencia  $T_t$  nos proporciona un predictor natural de los valores futuros de  $X_t$ . Por ejemplo si deseamos estimar  $\epsilon_{1990}$  por su valor medio, obtenemos

$$T_{1990} = 2,7588 \times 10^8$$

para la población de EE.UU. en 1990. Sin embargo si los residuales  $\hat{\epsilon}_t$  están altamente correlacionados podemos ser capaces de usar esos valores para dar una mejor estimación de  $\epsilon_{1990}$  y por consiguiente de  $X_{1990}$ .

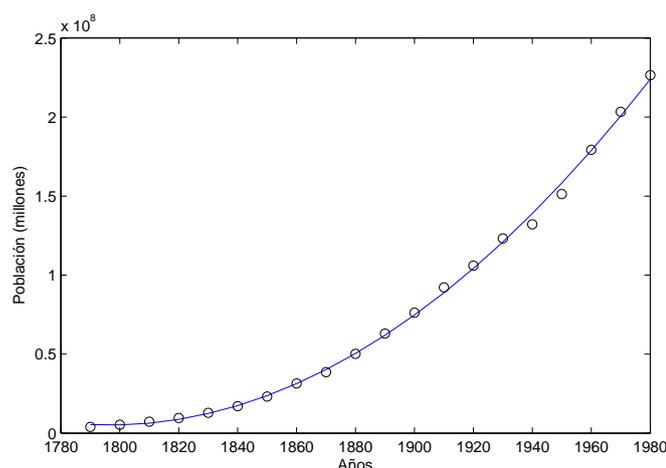


Figura 2.2: Población de los EE.UU., 1790-1980, junto con la parábola ajusta por mínimos cuadrados

**Método T2: Suavizado por medio de un promedio móvil.** Sea  $q$  un entero no negativo y consideremos un promedio móvil de la forma

$$W_t = \frac{1}{2q + 1} \sum_{j=-q}^q X_{t+j} \tag{2.57}$$

de un proceso  $\{X_t\}$  definido por (2.56). Entonces para  $q + 1 \leq t \leq n - q$ ,

$$\begin{aligned} W_t &= \frac{1}{2q+1} \sum_{j=-q}^q T_{t+j} + \frac{1}{2q+1} \sum_{j=-q}^q \epsilon_{t+j} \\ &\simeq T_t \end{aligned} \quad (2.58)$$

suponiendo que  $T_t$  es aproximadamente lineal sobre el intervalo  $[t - q, t + q]$  y que el promedio del término de error sobre este intervalo es cercano a cero.

El promedio móvil entonces nos provee con el estimador

$$\hat{T}_t = \frac{1}{2q+1} \sum_{j=-q}^q X_{t+j}, \quad q+1 \leq t \leq n-q. \quad (2.59)$$

Dado que  $X_t$  es no observado para  $t \leq 0$  o  $t \geq n$  no podemos usar (2.59) para  $t \leq q$  o  $t > n - q$ . Una forma de resolver este problema es haciendo  $X_t = X_1$  para  $t < 1$  y  $X_t = X_n$  para  $t > n$ . A continuación presentamos un ejemplo

**Ejemplo 2.4.1** El gráfico 2.3 muestra las huelgas ocurridas en EE.UU., de 1951 a 1980, según la Oficina de Estadísticas Laborales del Departamento de Trabajo de los EE.UU.

A estos datos le aplicamos un promedio móvil de 5 puntos, la Figura 2.4 muestra la serie suavizada y el término de error estimado  $\hat{\epsilon}_t = X_t - \hat{T}_t$  se muestra en la Figura 2.5. Como era de esperarse ellos no presentan una tendencia clara.



Figura 2.3: Huelgas en EE.UU., años 1951-1980. (Fuente: Oficina de Estadísticas Laborales, Departamento de Trabajo, EE.UU)

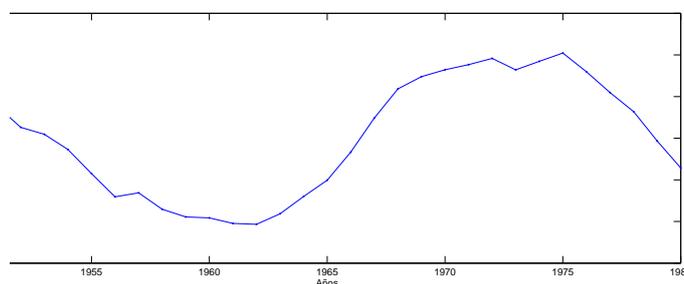


Figura 2.4: Promedio móvil de 5 puntos  $\hat{T}_t$ , de los datos del gráfico 2.3

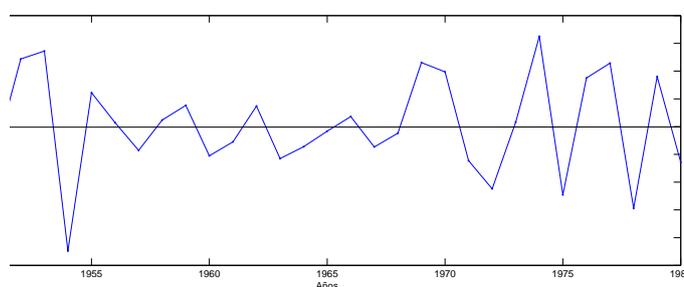


Figura 2.5: Residuales,  $\epsilon_t = X_t - \hat{T}_t$ , después de sustraer el promedio móvil de 5 puntos a los datos de huelgas

Para cada valor fijo  $a \in [0, 1]$ , el promedio móvil de un lado  $\hat{T}_t, t = 1, \dots, n$ , definido por la recursión

$$\hat{T}_t = aX_t + (1 - a)\hat{T}_t, \quad t = 2, \dots, n \tag{2.60}$$

y

$$\hat{T}_1 = X_1,$$

se puede calcular usando la función `filter` de Matlab o usando la opción `sides=1` en la función `filter` de R. Es usual pensar como aplicación de la ecuación (2.60) como un suavizado exponencial, dado que se sigue de la recursión que para  $t \leq 2$ ,  $\hat{T}_t = \sum_{j=0}^{t-2} a(1-a)^j X_{t-j} + (1-a)^{t-1} X_1$ , es un promedio móvil con peso de  $X_t, X_{t-1}, \dots$ , con pesos decreciendo exponencialmente (excepto para el último término).

Es útil pensar en  $\{\hat{T}_t\}$  en (`filter`) como un proceso obtenido de  $\{X_t\}$  por aplicación de un operador lineal o filtro lineal  $\hat{T}_t = \sum_{j=-\infty}^{\infty} a_j X_{t+j}$  con pesos  $a_j = (2q+1)^{-1}, -q \leq j \leq q$ , y  $a_j = 0, |j| > q$ . Este filtro

particular es un filtro de "paso-bajo" ya que toma los datos  $\{X_t\}$  y remueve la componente de rápida fluctuación (o de alta frecuencia)  $\{\hat{\epsilon}_t\}$ , para dejar el término de la tendencia estimada de lenta variación  $\{\hat{T}_t\}$ .

**Método T3: Diferenciación para generar datos estacionarios.** En lugar de intentar remover el ruido por suavizado como en el Método T2, ahora intentaremos eliminar la tendencia por diferenciación.

Definamos primero el operador diferencia  $\nabla$  por

$$\nabla x_t = x_t - x_{t-1} = (1 - B)x_t, \quad (2.61)$$

donde  $B$  es el operador de desplazamiento hacia atrás (*backward shift operator* en inglés),

$$Bx_t = x_{t-1}. \quad (2.62)$$

Las potencias de los operadores  $B$  y  $\nabla$  se definen de manera obvia, esto es,  $B^j(x_t) = x_{t-j}$  y

$$\nabla^j(x_t) = \nabla(\nabla^{j-1}(x_t)), j \geq 1 \text{ con}$$

$\nabla^0(x_t) = x_t$ . Los polinomios en  $B$  y  $\nabla$  se manipulan de la misma manera que las funciones polinómicas de variables reales.

Por ejemplo

$$\begin{aligned} \nabla^2 x_t &= \nabla(\nabla x_t) = (1 - B)(1 - B)x_t = (1 - 2B + B^2)x_t \\ &= x_t - 2x_{t-1} + x_{t-2}. \end{aligned}$$

Si el operador  $\nabla$  se aplica a una función con tendencia lineal  $T_t = at + b$ , entonces obtenemos la función constante  $\nabla T_t = a$ . De la misma manera cada tendencia polinomial de grado  $k$  se puede reducir a una constante por aplicación del operador  $\nabla^k$ .

Iniciando entonces con el modelo  $X_t = T_t + \epsilon_t$ , donde  $T_t = \sum_{j=0}^k a_j t^j$  y  $\epsilon_t$  es estacionario con media cero, obtenemos

$$\nabla^k X_t = k! a_k + \nabla^k \epsilon_t,$$

un proceso estacionario con media  $k! a_k$ . Esta consideración sugiere la posibilidad, dada una sucesión  $\{X_t\}$  de datos, de aplicar el operador  $\nabla$  repetidamente hasta conseguir una sucesión  $\{\nabla^k X_t\}$  la cual puede ser apropiadamente modelada como una realización de un proceso estacionario. Se encuentra a menudo en la práctica que el orden  $k$  de diferenciación es bastante pequeño, frecuentemente uno o dos.<sup>1</sup>

---

<sup>1</sup>Esto depende del hecho de que muchas funciones pueden ser aproximadas bastante bien, en un intervalo de longitud finita, por un polinomio de grado razonablemente bajo.

Aplicando esta técnica al ejemplo de los 20 datos de población de los EE.UU, hallamos que dos operaciones de diferenciación son suficientes para producir una serie sin aparente tendencia. Los datos diferenciados se muestran en la Figura 2.6. Note que la magnitud de las fluctuaciones en  $\nabla^2 X_n$  se incrementa con el valor de  $n$ . Este efecto se puede suprimir tomando primero logaritmo natural,  $y_n = \ln X_n$  y entonces aplicando el operador  $\nabla^2$  a la serie  $\{y_n\}$ .

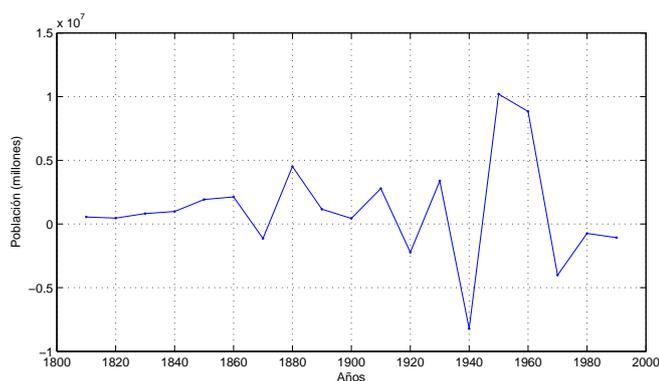


Figura 2.6: Serie con dos diferenciaciones obtenida de los datos de población de la Figura 2.2

## 2.4.2. Eliminación de la tendencia y la estacionalidad

Los métodos descritos para remover la tendencia pueden ser adaptados de manera natural para eliminar tanto la tendencia como la estacionalidad en el modelo general

$$X_t = T_t + E_t + \epsilon_t \quad (2.63)$$

donde  $\epsilon_t = 0$ ,  $E_{t+d} = E_t$  y  $\sum_{j=1}^d E_t = 0$ . Ilustraremos estos métodos con referencia al siguiente ejemplo de accidentes. En la Tabla 2.2 se muestran los datos, y en la Figura 2.7 podemos observar que en los mismos se ve claramente una componente estacional con periodo  $d = 12$ .

	1973	1974	1975	1976	1977	1978
Ene	9007	7750	8162	7717	7792	7836
Feb	8106	6981	7306	7461	6957	6892
Mar	8928	8038	8124	7776	7726	7791
Abr	9137	8422	7870	7925	8106	8129
May	10017	8714	9387	8634	8890	9115
Jun	10826	9512	9556	8945	9299	9434
Jul	11317	10120	10093	10078	10625	10484
Ago	10744	9823	9620	9179	9302	9827
Sep	9713	8743	8285	8037	8314	9110
Oct	9938	9129	8433	8488	8850	9070
Nov	9161	8710	8160	7874	8265	8633
Dic	8927	8680	8034	8647	8796	9240

Tabla 2.2: Accidentes mortales mensuales en EE.UU., años 1973-1978.

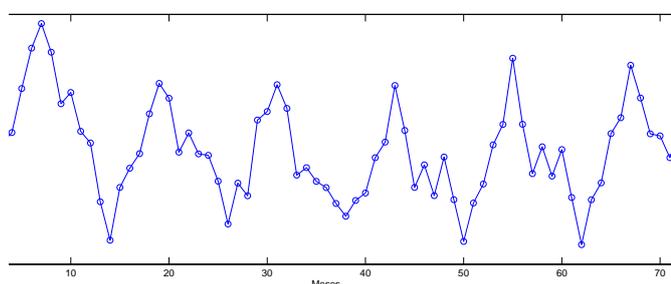


Figura 2.7: Número de accidentes mortales mensuales en EE.UU., 1973-1978

Será conveniente para el primer método indexar los datos por año y mes. Entonces  $X_{j,k}$ ,  $j = 1, \dots, 6$ ,  $k = 1, \dots, 12$  denotará el número de muertes accidentales reportados para el  $k$ -ésimo mes del  $j$ -ésimo año, (1972+j). En otras palabras, definimos

$$X_{j,k} = X_{k+12(j-1)}, j = 1, \dots, 6, k = 1, \dots, 12.$$

**Método E1: Método de la tendencia pequeña.** Si la tendencia es pequeña (como en los datos de accidentes) no es irrazonable suponer que el término de la tendencia es constante, digamos  $T_j$  para el año  $j$ . Dado

que  $\sum_{k=1}^{12} E_k = 0$ , nos lleva al estimador insesgado natural

$$\hat{T}_j = \frac{1}{12} \sum_{k=1}^{12} X_{j,k}, \quad (2.64)$$

mientras que para  $E_k, k = 1, \dots, 12$  tenemos el estimador

$$\hat{E}_t = \frac{1}{6} \sum_{j=1}^6 (X_{j,k} - \hat{T}_j), \quad (2.65)$$

el cual automáticamente satisface el requisito de que  $\sum_{k=1}^{12} \hat{E}_k = 0$ . El término de error estimado para el mes  $k$  del año  $j$  es por supuesto

$$\hat{\epsilon}_{j,k} = X_{j,k} - \hat{T}_j - \hat{E}_k, \quad j = 1, \dots, 6, k = 1, \dots, 12. \quad (2.66)$$

La generalización de (2.64) a (2.66) para datos con estacionalidad con un periodo distinto de 12 es bastante claro.

Las Figuras 2.8, 2.9 y 2.10 muestran respectivamente las observaciones con la tendencia removida  $X_{j,k} - \hat{T}_j$ , la componente estacional estimada  $\hat{E}_t$  y las observaciones con la tendencia y la estacionalidad removida  $\hat{\epsilon}_{j,k} = X_{j,k} - \hat{T}_j - \hat{E}_k$ . En la última no se observa una aparente tendencia o estacionalidad.

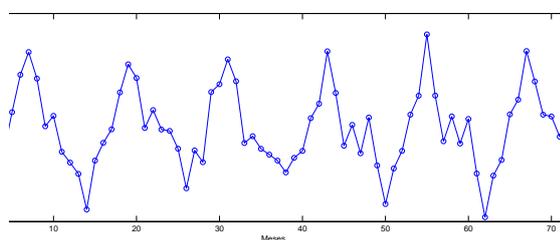


Figura 2.8: Accidentes mortales mensuales luego de sustraer la tendencia estimada por el Método S1.

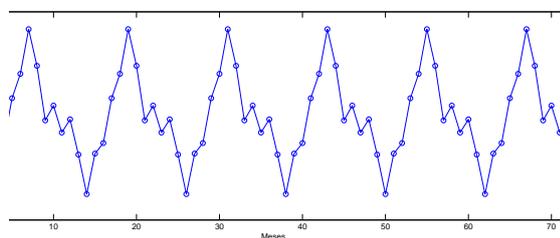


Figura 2.9: La componente estacional de los accidentes mortales mensuales estimada por el Método S1.

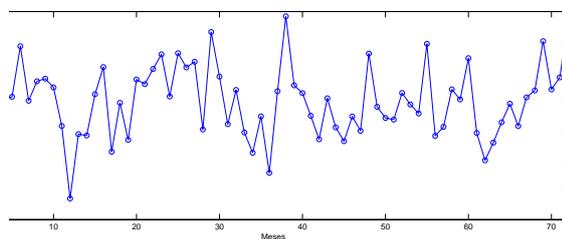


Figura 2.10: Accidentes mortales mensuales luego de sustraer la tendencia y la estacionalidad estimadas por el Método S1.

**Método E2: Estimación por promedio móvil**

La siguiente técnica es preferible al Método S1 ya que no se basa en la suposición de que  $T_t$  es casi constante sobre cada ciclo estacional.

Suponga que tenemos las observaciones  $\{x_1, \dots, x_n\}$ . Se estima primero la tendencia aplicando un filtro de promedio móvil especialmente elegido para eliminar la componente estacional y para amortiguar el ruido. Si el periodo  $d$  es par, digamos  $d = 2q$ , entonces usamos

$$\hat{T}_t = (0,5x_{t-q} + x_{t-q+1} + \dots + x_{t+q-1} + 0,5x_{t+q})/d, q < t \leq n - q. \tag{2.67}$$

Si el periodo es impar, digamos  $d = 2q + 1$ , entonces usamos el promedio móvil simple (2.59). La Figura 2.11 muestra la tendencia estimada  $\hat{T}_t$  para los datos de accidentes mortales obtenido de (2.67). También muestra la tendencia constante a trozos obtenida por el Método S1.

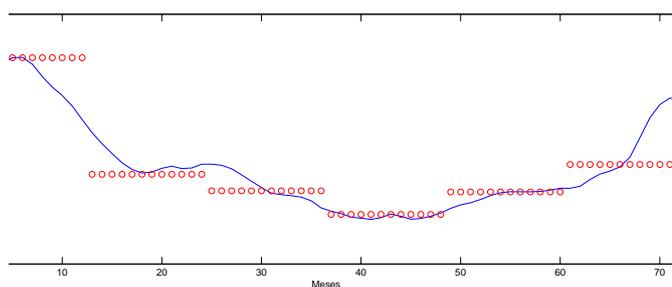


Figura 2.11: Comparación del promedio móvil y la tendencia estimada a trozos de los accidentes mortales mensuales.

El segundo paso, es estimar la componente estacional. Para cada  $k = 1, \dots, d$ , calculamos el promedio  $w_k$

de las desviaciones  $\{(X_{k+jd} - \hat{T}_{k+jd}) : q < k + jd \leq n - q\}$ . Dado que este promedio de desviaciones no necesariamente suma cero, estimamos la componente estacional  $E_k$  como:

$$\hat{E}_k = w_k - \frac{1}{d} \sum_{i=1}^d w_i, \quad i = 1, \dots, d, \quad (2.68)$$

y  $\hat{E}_k = \hat{E}_{k-d}, k > d$ .

Los datos sin la componente estacional se definen entonces como la serie original con la componente estacional removida, es decir,

$$d_t = X_t - \hat{E}_t, \quad t = 1, \dots, n. \quad (2.69)$$

Finalmente, reestimamos la tendencia de  $\{d_t\}$  aplicando un filtro de promedio móvil como se describió para los datos no estacionales o fijando un polinomio a la serie  $\{d_t\}$ . El término del ruido estimado llega a ser entonces

$$\hat{\epsilon}_t = X_t - \hat{E}_t - \hat{T}_t, \quad t = 1, \dots, n.$$

Los resultados de aplicar los Métodos S1 y S2 a los datos de accidentes mortales son casi iguales, dado que en este caso la constante a trozos y el promedio móvil de  $T_t$  están razonablemente cercanos.

Una comparación de los valores estimados de  $E_k, k = 1, \dots, 12$ , obtenido con ambos métodos se muestra en la Tabla 2.3

k	1	2	3	4	5	6	7	8	9	10	11	12
$\hat{E}_t$ (S1)	-7434	-1504	-724	-523	338	808	1665	961	-87	197	-321	-67
$\hat{E}_t$ (S2)	-804	-1522	-737	-526	343	746	1680	987	-109	258	-259	-57

Tabla 2.3: Componentes estacional estimadas para los datos de accidentes mortales.

**Método E3: Diferenciación a paso d.** La técnica de diferenciación la cual aplicamos antes a datos no estacionales se pueden adaptar para lidiar con el caso estacional de periodo  $d$  introduciendo el operador de diferencia de paso  $d$   $\nabla_d$  definido por:

$$\nabla_d X_t = X_t - X_{t-d} = (1 - B^d)X_t. \quad (2.70)$$

Este operador no debe confundirse con el operador  $\nabla^d = (1 - B)^d$  definido por (2.61).

Aplicando el operador  $\nabla_d$  al modelo

$$X_t = T_t + E_t + \epsilon_t,$$

donde  $\{E_t\}$  tiene periodo  $d$ , obtenemos

$$\nabla_d X_t = T_t - T_{t-d} + \epsilon_t - \epsilon_{t-d},$$

lo cual nos da una descomposición de la diferencia  $\nabla_d X_t$  en una componente de tendencia ( $T_t - T_{t-d}$ ) y un término de ruido ( $\epsilon_t - \epsilon_{t-d}$ ). La tendencia ( $T_t - T_{t-d}$ ) se puede eliminar usando los métodos ya descritos, por ejemplo, aplicando alguna potencia del operador  $\nabla$ . La Figura 2.12 muestra el resultado de aplicar el operador  $\nabla_{12}$  a los datos de accidentes mortales. La componente estacional evidente en la Figura 2.7 está ausente en la Figura de  $\nabla_{12} X_t$ ,  $13 \leq t \leq 72$ . Sin embargo todavía parece haber una tendencia decreciente. Si ahora aplicamos el operador  $\nabla$  a  $\nabla_{12} X_t$  y graficamos las diferencias  $\nabla \nabla_{12} X_t$ ,  $t = 14, \dots, 72$  obtenemos el gráfico mostrado en la Figura 2.13, los cuales no tienen una aparente tendencia o componente estacional.

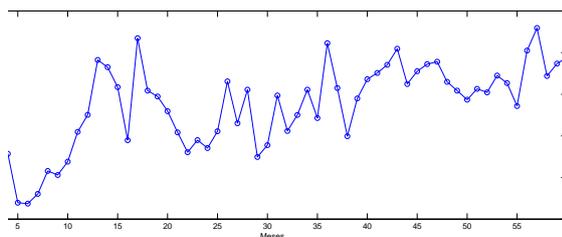


Figura 2.12: La serie diferenciada  $\nabla_{12} X_t$  derivada de la serie de accidentes mortales.

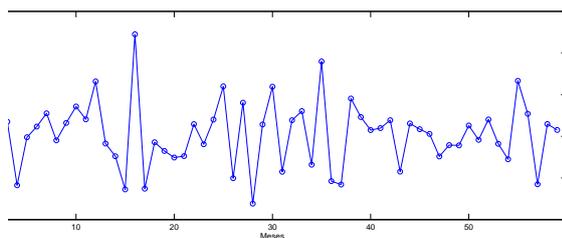


Figura 2.13: La serie diferenciada  $\nabla \nabla_{12} X_t$  derivada de la serie de accidentes mortales.

## Análisis de los datos

En este capítulo vamos a realizar un análisis descriptivo de los datos meteorológicos del Lago de Valencia, los mismos provienen de dos estaciones meteorológicas, que llamaremos "Estación Sucre" y "Estación Tacarigua".

La estación meteorológica Sucre se encuentra ubicada en la ribera NE del lago en Punta Pan de Azúcar, cuyas coordenadas aproximadas son:  $10^{\circ}25'N$ ,  $67^{\circ}71'O$  y la estación meteorológica Tacarigua se encuentra ubicada en la Isla de Tacarigua, cuyas coordenadas aproximadas son:  $10^{\circ}16'N$ ,  $67^{\circ}73'O$ . Ambas estaciones meteorológicas recolectaron datos correspondientes a las temperaturas, la intensidad de la radiación sobre la corteza terrestre, las precipitaciones y las velocidades del viento sobre el lago de Valencia.

En ambas estaciones meteorológicas, los registros fueron tomados cada 5 minutos,teniéndose 288 registros por día; el periodo de estudio es de Noviembre del año 2007 hasta octubre de año 2008. Sin embargo, durante la recolección de los datos hubo perdida de los mismos de modo que no existen registros de todos los días y además existen días con menos de 288 registros.

Para cada mes se dividieron los registros recolectados en dos subconjuntos : día (6am-6pm) y noche (6pm-6am); teniendo así, cada subconjunto 144 registros y una duración de doce horas.

A continuación, el estudio de las temperaturas, la intensidad de la radiación sobre la corteza terrestre, las precipitaciones y las velocidades del viento sobre el lago de Valencia.

### 3.1. Temperatura.

En esta sección se presentará el estudio de la temperaturas en las estaciones metereológicas Sucre y Tacarigua respectivamente.

#### 3.1.1. Estación Sucre.

A continuación, la serie temporal del periodo 2007-2008 :

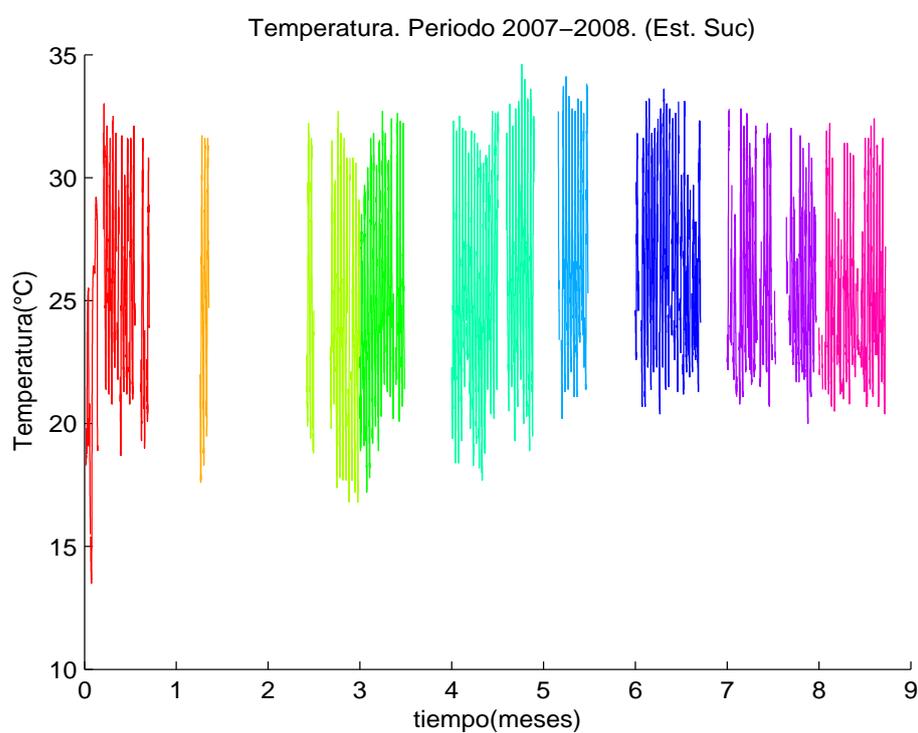


Figura 3.1: Serie temporal mensual de la temperatura. (Est. Sucre)

Los parámetros estadísticos mensuales correspondientes al periodo 2007-2008 , están representados en las siguientes tablas:

Parámetros Estadísticos	Año 2007	
	Noviembre	Diciembre
Min	13,50	17,60
Q1	21,90	20,90
Mediana	25,20	25,20
Media	24,92	24,84
Q3	28,40	28,20
Max	33,00	31,70
D.Est	4,08	4,02
Varianza	16,69	16,20

Tabla 3.1: Temperatura. Estación de Sucre (Año 2007).

Parámetros Estadísticos	Año 2008						
	Enero	Febrero	Marzo	Abril	Junio	Sep	Oct
Min	16,80	17,20	17,70	20,20	20,40	20,00	20,40
Q1	20,80	22,00	22,30	23,70	23,20	22,90	22,40
Mediana	23,60	24,20	25,10	25,90	25,70	24,90	23,80
Media	24,01	24,76	25,41	26,43	26,10	25,48	24,72
Q3	27,30	27,60	28,60	29,20	28,70	27,70	26,60
Max	32,70	32,70	34,60	34,10	33,60	32,80	32,40
D.Est	3,97	3,49	3,84	3,41	3,26	3,07	2,93
Varianza	15,83	12,23	14,76	11,63	10,68	9,43	8,59

Tabla 3.2: Temperatura. Estación de Sucre (Año 2008).

En los meses correspondientes al período 2007-2008, los registros mensuales no se encuentran muy dispersos respecto a la media respectiva de cada mes.

Para corroborar los valores de las tablas anteriores, se muestra el boxplot del periodo 2007-2008:

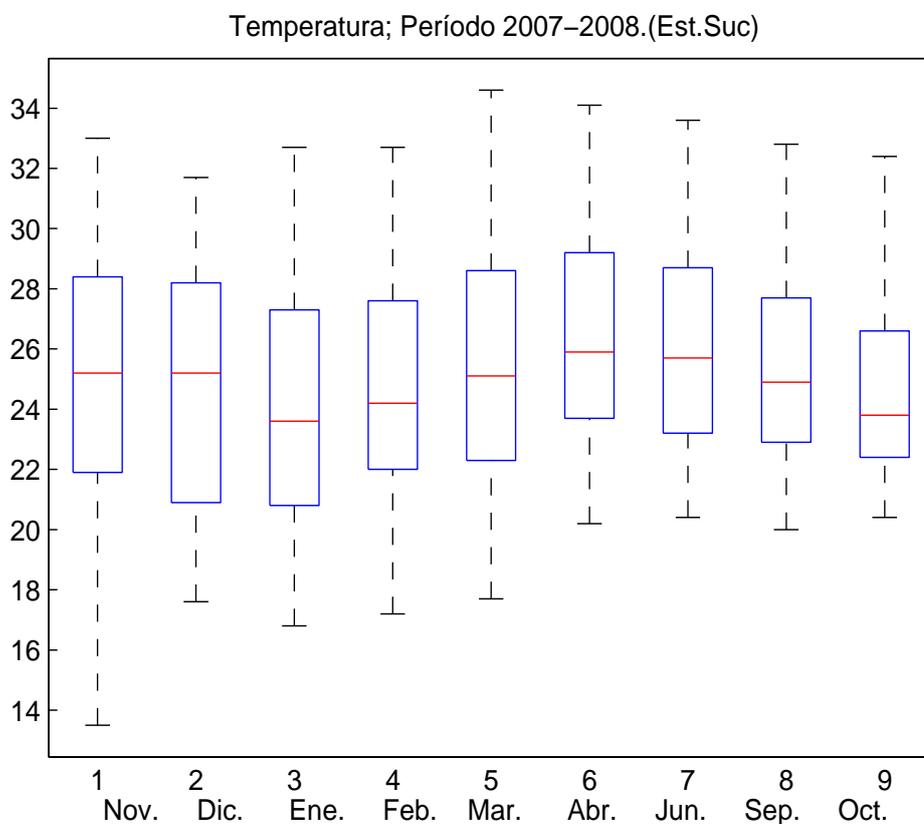


Figura 3.2: Boxplot mensual de la temperatura. (Est. Sucre)

Claramente, se puede observar una distribución uniforme de los datos durante marzo, junio, septiembre y octubre. A diferencia del resto de los meses, en los que la distribución de los datos es irregular.

Tomemos como ejemplo el mes febrero del año 2008: en febrero se registrarón los datos de los primeros 16 días del mes, los cuales están representados en la siguiente serie temporal:

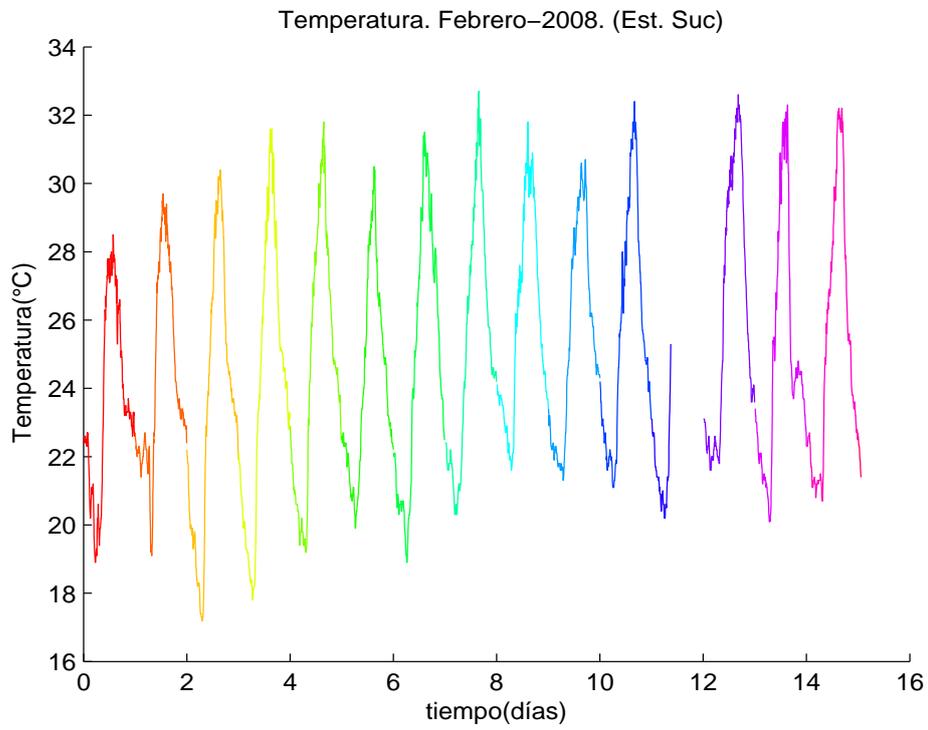


Figura 3.3: Serie temporal de la temperatura correspondiente al mes de febrero del año 2008. (Est. Sucre)

La siguiente tabla muestra los parámetros estadísticos diarios diurnos del mes:

Días	Parámetros Estadísticos							
	Min	Q1	Mediana	Media	Q3	Max	D.Est	Varianza
1	19,10	22,85	26,30	25,01	27,30	28,50	2,94	8,68
2	19,10	22,80	27,30	25,96	28,60	29,70	3,23	10,43
3	17,20	22,77	26,50	25,25	29,10	30,40	4,34	18,88
4	17,80	23,15	27,30	25,87	29,22	31,60	4,37	19,17
5	19,20	24,12	27,80	26,47	29,60	31,80	3,89	15,18
6	19,90	22,95	26,40	25,54	28,22	30,50	3,21	10,32
7	18,90	23,70	27,30	26,42	29,70	31,50	3,93	15,45
8	21,00	24,10	27,80	27,07	29,90	32,70	3,53	12,52
9	21,60	25,80	28,30	27,47	30,10	31,80	3,08	9,49
10	21,30	24,47	28,20	27,01	29,60	30,70	3,08	9,51
11	21,10	25,27	28,20	27,48	30,72	32,40	3,44	11,84
12	20,20	20,60	21,40	21,75	22,82	25,30	1,45	2,10
13	21,80	24,80	29,80	28,38	31,30	32,60	3,61	13,06
14	20,10	24,20	26,80	26,87	30,60	32,30	3,81	14,52
15	20,70	25,05	27,90	27,54	30,95	32,20	3,64	13,26

Tabla 3.3: Temperatura (Febrero/día:6am-6pm)(Est.Suc)

En la mayoría de los días, los registros no se encuentran dispersos respecto a su media; sin embargo, en los días 3 y 4 de febrero los valores de la desviación estándar y la varianza están influenciados indicando así temperaturas uniformemente crecientes o decrecientes durante el día.

La proxima tabla muestra los parámetros estadísticos diarios nocturnos:

Días	Parámetros Estadísticos							
	Min	Q1	Mediana	Media	Q3	Max	D.Est	Varianza
1	18,90	21,10	22,60	22,21	23,30	24,40	1,44	2,07
2	21,40	22,10	22,70	22,81	23,40	25,40	0,89	0,80
3	17,80	19,70	22,10	21,76	23,70	25,90	2,36	5,60
4	18,30	19,90	22,45	22,28	24,60	27,40	2,65	7,03
5	19,40	20,60	22,60	22,76	24,70	27,40	2,32	5,39
6	20,30	21,80	22,50	22,85	24,10	26,30	1,57	2,48
7	19,10	20,60	22,30	23,17	25,80	28,70	2,89	8,36
8	20,30	21,60	23,30	23,54	25,40	28,80	2,40	5,78
9	21,90	23,40	23,80	24,40	25,30	29,10	1,71	2,95
10	21,60	22,20	23,85	23,99	25,80	29,10	1,99	3,99
11	21,10	22,20	23,55	23,83	24,90	28,30	1,84	3,41
12	20,20	20,90	21,40	21,45	21,95	23,10	0,80	0,65
13	21,60	22,30	23,35	24,16	25,40	30,80	2,43	5,92
14	20,80	21,80	23,00	22,95	24,30	24,80	1,26	1,61
15	20,80	21,30	22,60	23,10	24,00	27,90	1,87	3,53

Tabla 3.4: Temperatura (Febrero/noche:6pm-6am)(Est.Suc)

En la totalidad de los días, los registros no se encuentran dispersos respecto a su media; indicando temperaturas regulares en el transcurso de la noche.

Consideremos como ejemplo, el día 2 de febrero del año 2008 mediante la siguiente serie de tiempo:

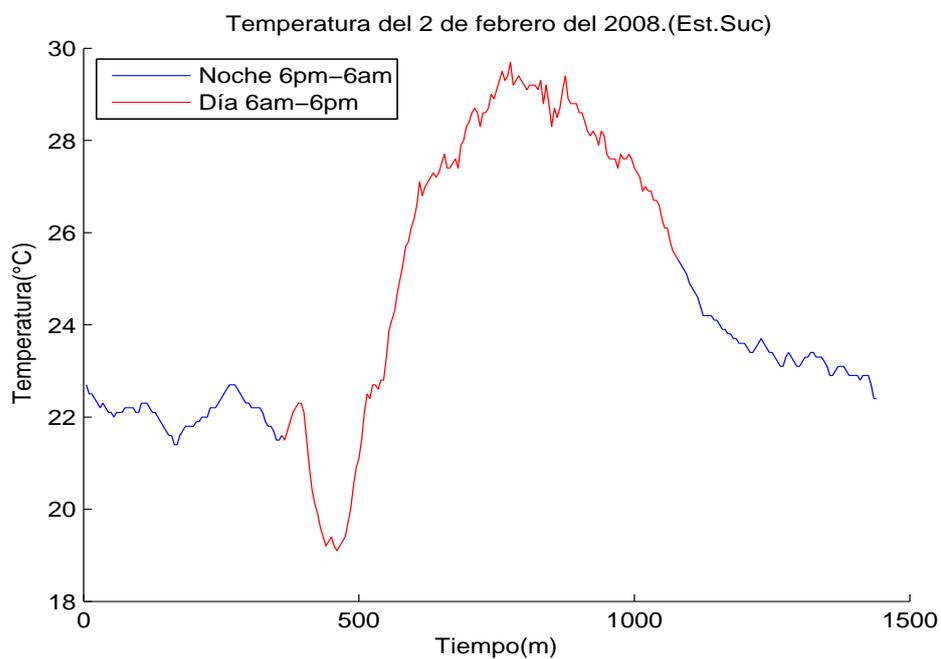


Figura 3.4: Serie temporal de la temperatura correspondiente al 2 de febrero del año 2008. (Est. Sucre)

En esta serie temporal, durante el periodo diurno se destaca un aumento exponencial de las temperaturas luego de un descenso de la misma al inicio de la mañana.

### 3.1.2. Estación Tacarigua.

A continuación, la serie temporal del periodo 2007-2008 :

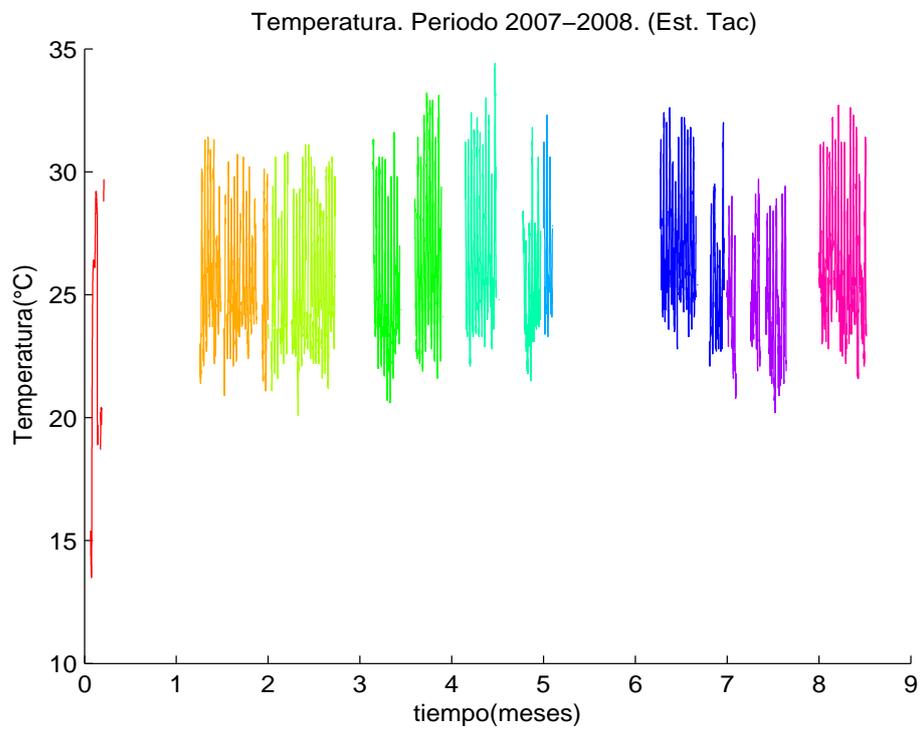


Figura 3.5: Serie temporal mensual de la temperatura. (Est. Tacarigua)

Los parámetros estadísticos mensuales correspondientes al periodo 2007-2008 , están representados en las siguientes tablas:

Parámetros Estadísticos	Año 2007	
	Noviembre	Diciembre
Min	13,50	20,90
Q1	19,70	23,80
Mediana	26,10	24,90
Media	23,29	25,40
Q3	27,07	26,90
Max	29,70	31,40
D.Est	5,05	2,19
Varianza	25,51	4,81

Tabla 3.5: Temperatura (Año 2007)(Est.Tac)

Parámetros Estadísticos	Año 2008						
	Enero	Marzo	Abril	Mayo	Junio	Julio	Agosto
Min	20,10	20,60	21,50	23,30	22,10	20,20	21,60
Q1	23,10	23,37	24,10	24,80	24,80	22,90	24,60
Mediana	24,40	24,80	25,30	25,70	26,00	24,30	25,50
Media	25,12	25,56	25,96	26,76	26,61	24,53	26,15
Q3	27,10	27,60	27,40	28,80	28,30	26,00	27,70
Max	31,10	33,20	34,40	32,30	32,60	29,70	32,70
D.Est	2,46	2,87	2,56	2,38	2,48	2,14	2,42
Varianza	6,07	8,27	6,59	5,70	6,16	4,60	5,87

Tabla 3.6: Temperatura (Año 2008)(Est.Tac).

De acuerdo a las tablas anteriores, los registros mensuales no se encuentran dispersos respecto a su media. Más aún, los valores de la varianza no se ven influenciados indicando la carencia de registros atípicos en la distribución de los datos.

Seguidamente, el boxplot del periodo 2007-2008:



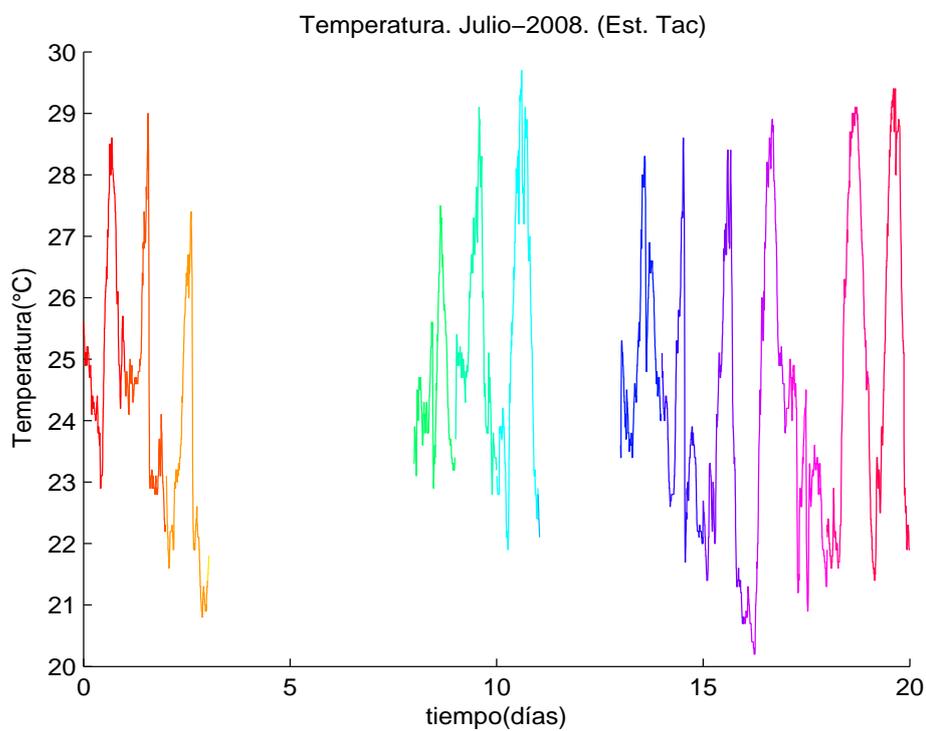


Figura 3.7: Serie temporal de la temperatura correspondiente al mes de julio del año 2008. (Est. Tacarigua)

En el periodo diurno se consideraron los días desde el 1 hasta el 3, desde el 9 hasta el 11 y desde el 14 hasta el 20 de julio; ya que, en los días restantes hubo perdida total de registros.

Los parámetros estadísticos diarios diurnos, se muestran en la siguiente tabla:

Días	Parámetros Estadísticos							
	Min	Q1	Mediana	Media	Q3	Max	D.Est	Varianza
1	22,90	23,97	25,30	25,62	27,70	28,60	1,91	3,68
2	22,80	23,10	24,80	25,07	26,90	29,00	1,84	3,39
3	21,90	23,10	24,20	24,47	26,22	27,40	1,77	3,13
9	22,90	24,00	24,80	25,03	26,02	27,50	1,22	1,48
10	23,90	25,07	26,60	26,41	27,42	29,10	1,35	1,84
11	21,90	25,60	27,70	26,90	28,60	29,70	2,11	4,46
14	23,40	24,60	25,60	25,76	26,80	28,30	1,41	1,99
15	21,70	22,90	23,80	24,48	25,70	28,60	1,82	3,33
16	22,00	24,00	25,10	25,40	26,90	28,40	1,78	3,19
17	22,20	24,27	26,90	26,00	28,12	28,90	2,59	6,71
18	20,90	22,60	23,10	22,85	23,30	24,50	0,87	0,76
19	21,60	24,55	27,60	26,50	28,80	29,10	2,60	6,79
20	22,50	24,77	28,00	26,92	28,80	29,40	2,31	5,34

Tabla 3.7: Temperatura (Julio/día:6am-6pm)(Est.Tac)

En el periodo nocturno, se consideraron desde el 5 hasta el 16 y desde el 25 hasta el 30 de abril; ya que, en los días restantes hubo perdida total de registros.

Los parámetros estadísticos diarios nocturnos, se muestran en la siguiente tabla:

Días	Parámetros Estadísticos							
	Min	Q1	Mediana	Media	Q3	Max	D.Est	Varianza
1	24,10	24,80	25,05	25,22	25,50	27,70	0,83	0,69
2	22,20	23,10	24,20	23,84	24,60	25,00	0,85	0,72
3	20,80	21,20	21,90	21,86	22,20	23,20	0,68	0,46
4	21,40	21,50	21,60	21,59	21,70	21,80	0,14	0,02
9	23,10	23,60	23,90	24,05	24,50	25,80	0,66	0,44
10	22,80	23,70	24,55	24,30	24,90	25,40	0,69	0,48
11	22,10	22,80	23,75	23,99	24,30	28,40	1,46	2,14
12	22,10	22,30	22,40	22,45	22,67	22,80	0,22	0,05
14	23,40	24,10	24,55	24,65	25,10	26,60	0,83	0,69
15	22,00	22,30	23,00	23,18	24,00	25,10	0,87	0,75
16	20,70	21,30	21,70	21,92	22,60	23,40	0,81	0,66
17	20,20	20,70	22,55	22,90	24,90	27,20	2,33	5,43
18	21,30	22,20	23,65	23,43	24,50	25,00	1,19	1,42
19	21,60	22,00	23,70	24,19	26,40	28,70	2,36	5,57
20	21,40	22,00	22,65	23,32	23,80	28,80	1,88	3,55

Tabla 3.8: Temperatura (Julio/noche:6pm-6am)(Est.Tac)

Durante el periodo diurno y nocturno, los registros no se encuentran dispersos respecto a los valores de la media respectivamente.

Como ejemplo, la serie temporal correspondiente al 18 de julio:

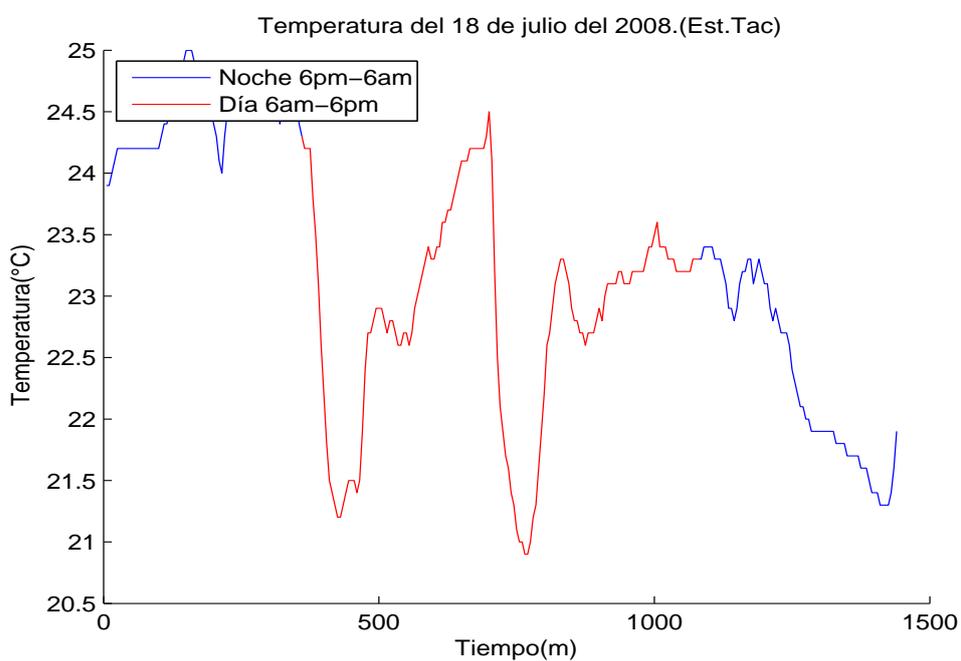


Figura 3.8: Serie temporal de la temperatura correspondiente al 18 de julio del año 2008. (Est. Tacarigua)

A pesar de las destacadas variaciones observadas en el periodo diurno, en las tablas antes mostradas se puede observar que los valores de la desviación estándar y la varianza respectivamente no están influenciados; esto ocurre por que las variaciones no representan valores muy alejados de la media.

## 3.2. Radiación.

En esta sección se presentará el estudio de la radiación en las estaciones metereológicas Sucre y Tacarigua respectivamente.

### 3.2.1. Estación Sucre.

A continuación, la serie temporal del periodo 2007-2008 :

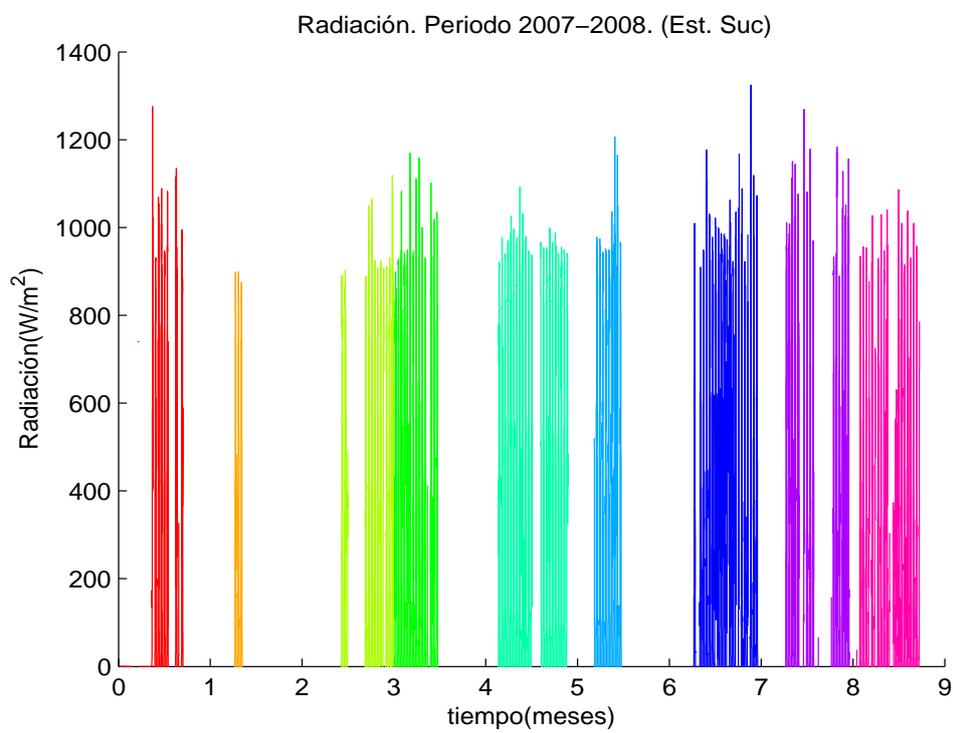


Figura 3.9: Serie temporal mensual de la radiación. (Est. Sucre)

Los parámetros estadísticos mensuales correspondientes al periodo 2007-2008 , están representados en las siguientes tablas:

Parámetros Estadísticos	Año 2007	
	Noviembre	Diciembre
Min	0,00	0,00
Q1	0,00	170,25
Mediana	0,00	530,00
Media	248,94	485,07
Q3	531,50	800,75
Max	1275,00	899,00
D.Est	339,99	324,80
Varianza	1,15e+005	1,05e+005

Tabla 3.9: Radiación. Estación de Sucre (Año 2007).

Parámetros Estadísticos	Año 2008						
	Enero	Febrero	Marzo	Abril	Junio	Sep	Oct
Min	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Q1	168,00	126,00	238,00	176,50	147,00	123,50	76,75
Mediana	495,00	455,00	567,50	509,50	367,00	388,00	310,00
Media	476,25	460,20	527,79	503,00	435,36	450,56	395,49
Q3	794,25	781,00	844,00	826,00	741,75	783,00	719,50
Max	1118,00	1169,00	1092,00	1206,00	1324,00	1269,00	1086,00
D.Est	326,54	335,22	326,13	331,35	317,97	342,90	332,82
Varianza	1,06e+005	1,12e+005	1,06e+005	1,09e+005	1,01e+005	1,17e+005	1,10e+005

Tabla 3.10: Radiación. Estación de Sucre (Año 2008).

En los meses correspondientes al período 2007-2008, los registros se encuentran dispersos respecto a su media; sin embargo no puede asegurarse la existencia de casos atípicos.

Para corroborar los valores de las tablas anteriores, se muestra el boxplot del periodo 2007-2008:

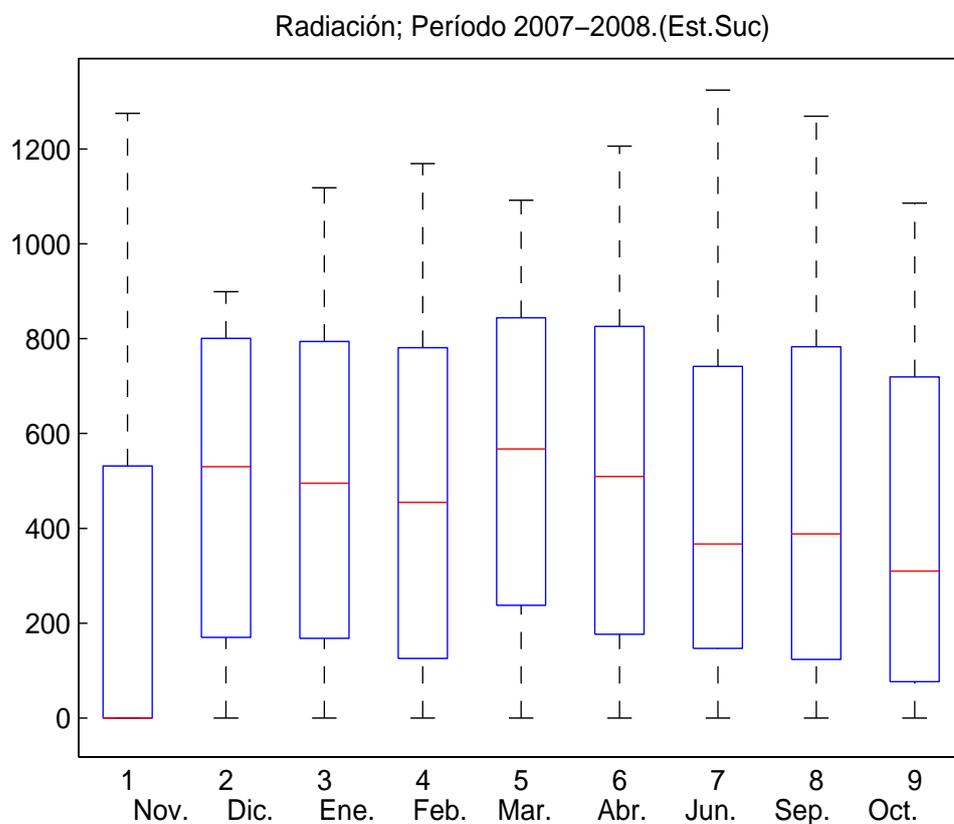


Figura 3.10: Boxplot mensual de la radiación. (Est. Sucre)

Claramente, pueden apreciarse los casos atípicos y la dispersión de los datos respecto a la media de la distribución de la data.

Tomemos como ejemplo, los meses enero y marzo del año 2008: en enero, se registrarán los datos de los primeros 31 días del mes, los cuales están representados en la siguiente serie temporal:

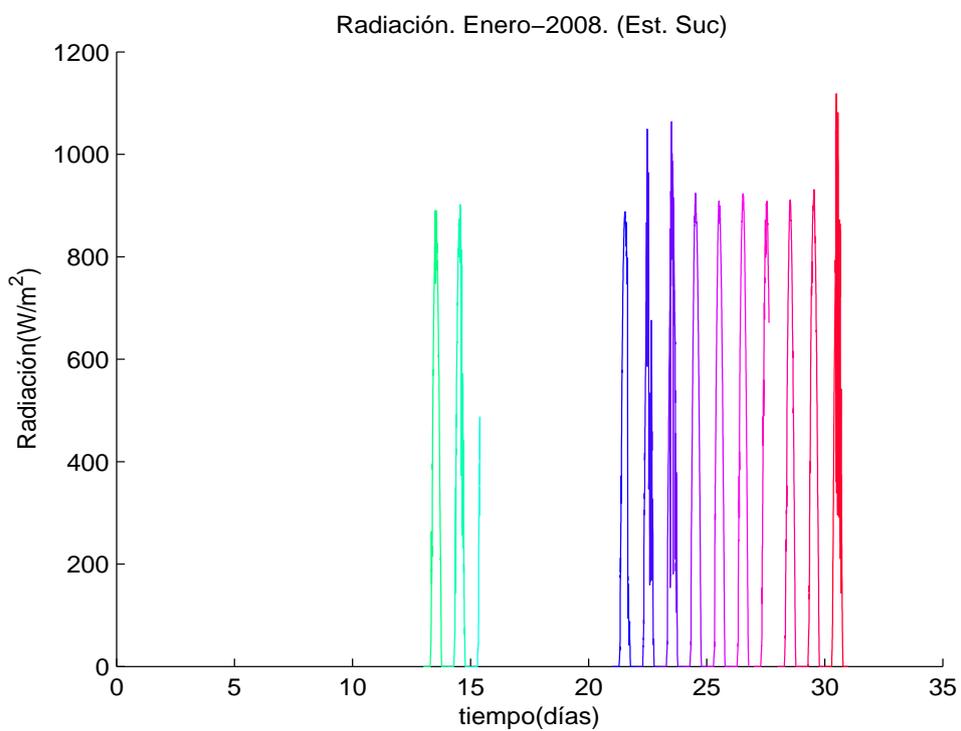


Figura 3.11: Serie temporal de la radiación correspondiente al mes de enero del año 2008. (Est. Sucre)

Se tomaron en cuenta los días desde el 14 hasta el 15, desde el 22 hasta el 27 y desde el 29 hasta el 31 de enero; ya que, en algunos días restantes hubo perdida total de registros y en otros hubo registros nulos.

La siguiente tabla muestra los parámetros estadísticos diarios diurnos del mes:

Días	Parámetros Estadísticos							
	Min	Q1	Mediana	Media	Q3	Max	D.Est	Varianza
14	0,00	210,50	558,00	1,71	776,50	86,00	9,93	98,60
15	0,00	168,25	442,00	1,28	789,50	61,00	7,04	49,58
22	0,00	71,50	505,00	1,02	811,00	41,00	5,62	31,62
23	0,00	167,50	350,00	1,56	628,00	54,00	7,71	59,55
24	0,00	180,25	516,00	1,84	812,25	92,00	10,33	106,85
25	0,00	243,50	578,00	1,89	828,25	80,00	9,72	94,47
26	0,00	253,25	584,00	2,52	816,50	100,00	12,68	160,87
27	0,00	226,75	602,00	3,32	833,50	80,00	12,75	162,66
29	0,00	203,00	551,00	2,15	801,50	84,00	10,57	111,87
30	0,00	207,75	605,00	2,23	827,00	91,00	11,38	129,69
31	0,00	203,75	372,00	2,05	633,75	77,00	10,20	104,14

Tabla 3.11: Radiación (Enero/noche:6pm-6am)(Est.Suc)

En todos los días los valores de la desviación estándar y la varianza se ven muy influenciados respecto a los valores de la media, indicando la existencia de datos atípicos en los registros de datos.

Consideremos como ejemplo, el día 31 de enero mediante la siguiente serie de tiempo:

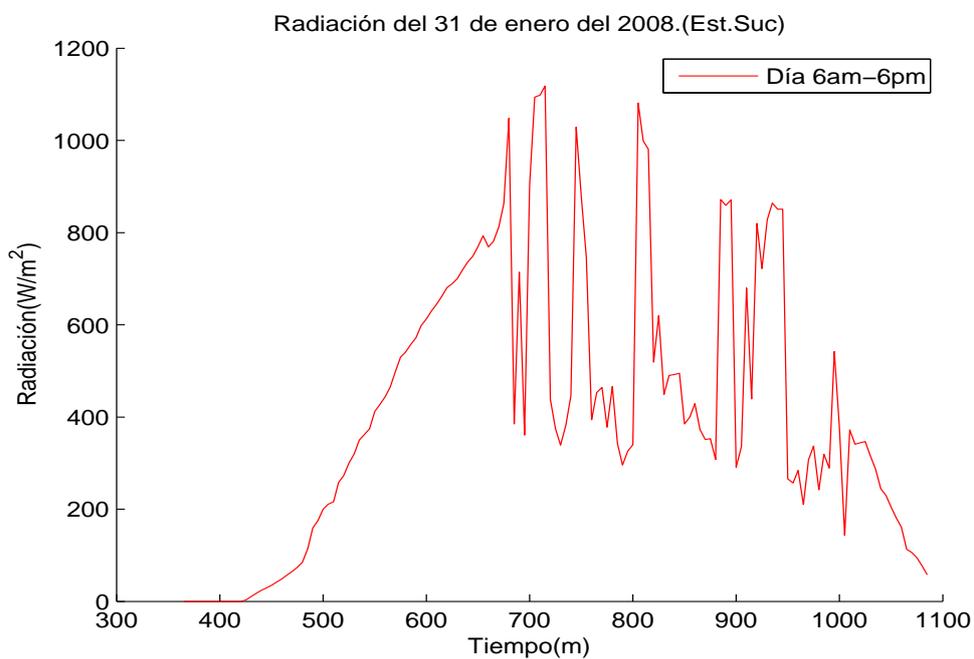


Tabla 3.12: Serie temporal de la radiación correspondiente al 31 de enero del año 2008. (Est. Sucre)

Claramente, se puede observar un comportamiento irregular en la intensidad de la radiación ; Más aún, se pueden apreciar varios picos que indican altas y bajas de la intensidad de la radiación sobre la superficie de terrestre.

En marzo, se registrarón los valores de los primeros 28 días del mes, los cuales están representados en la siguiente serie temporal:

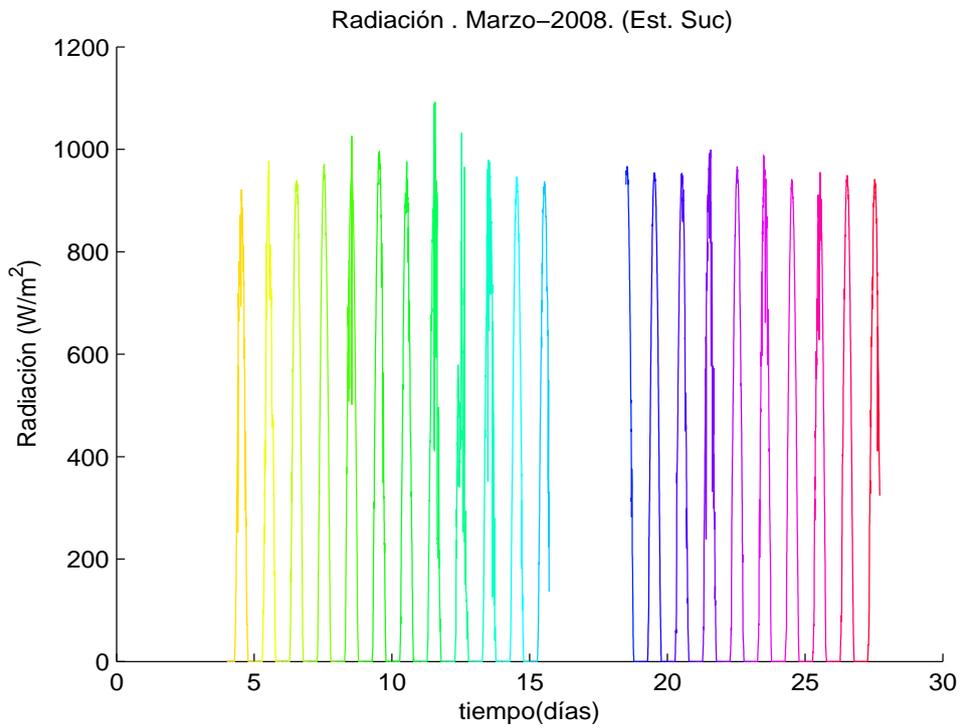


Figura 3.12: Serie temporal de la radiación correspondiente al mes de marzo del año 2008. (Est. Sucre)

Se tomaron en cuenta los días desde el 5 hasta el 15 y desde el 19 hasta el 27 de marzo ; ya que, en algunos días restantes hubo registros nulos y en otros hubo perdida total de registros.

La siguiente tabla muestra los parámetros estadísticos diarios diurnos del mes:

Días	Parámetros Estadísticos							
	Min	Q1	Mediana	Media	Q3	Max	D.Est	Varianza
5	0,00	149,25	497,00	1,73	777,75	55,00	7,83	61,39
6	0,00	182,00	473,00	2,82	764,25	102,00	13,24	175,51
7	0,00	256,25	601,00	3,23	856,50	120,00	15,17	230,29
8	0,00	273,00	618,00	3,34	864,00	117,00	14,99	224,75
9	0,00	271,00	611,00	3,10	864,75	106,00	14,01	196,29
10	0,00	268,50	627,00	3,88	896,00	116,00	7,24	297,22
11	0,00	243,25	565,00	3,06	857,25	111,00	13,82	191,16
12	0,00	201,50	434,00	2,84	849,25	83,00	12,54	157,32
13	0,00	119,75	341,00	2,56	455,00	61,00	10,80	116,75
14	0,00	156,50	590,00	2,04	849,00	67,00	9,63	92,74
15	0,00	246,75	600,00	2,95	834,50	110,00	13,88	192,92
19	0,00	450,75	771,00	5,84	939,50	106,00	19,07	363,82
20	0,00	267,00	627,00	2,90	871,25	91,00	13,30	177,10
21	0,00	236,75	558,00	2,37	860,25	66,00	10,26	105,42
22	0,00	258,75	552,00	2,76	821,00	98,00	12,77	163,28
23	0,00	284,50	621,00	1,91	865,75	48,00	7,97	63,67
24	0,00	267,50	604,00	3,32	853,25	132,00	16,18	262,09
25	0,00	243,75	603,00	2,91	847,75	97,00	13,38	179,26
26	0,00	252,25	576,00	1,03	825,00	32,00	4,53	20,52
27	0,00	249,50	608,00	2,39	855,75	80,00	11,00	121,01

Tabla 3.13: Radiación (Marzo/noche:6pm-6am)(Est.Suc)

En todos los días los valores de la desviación estándar y la varianza se ven muy influenciados respecto a los valores de la media, indicando la existencia de datos atípicos en la data.

Consideremos como ejemplo, el día 22 de marzo mediante las siguiente serie de tiempo:

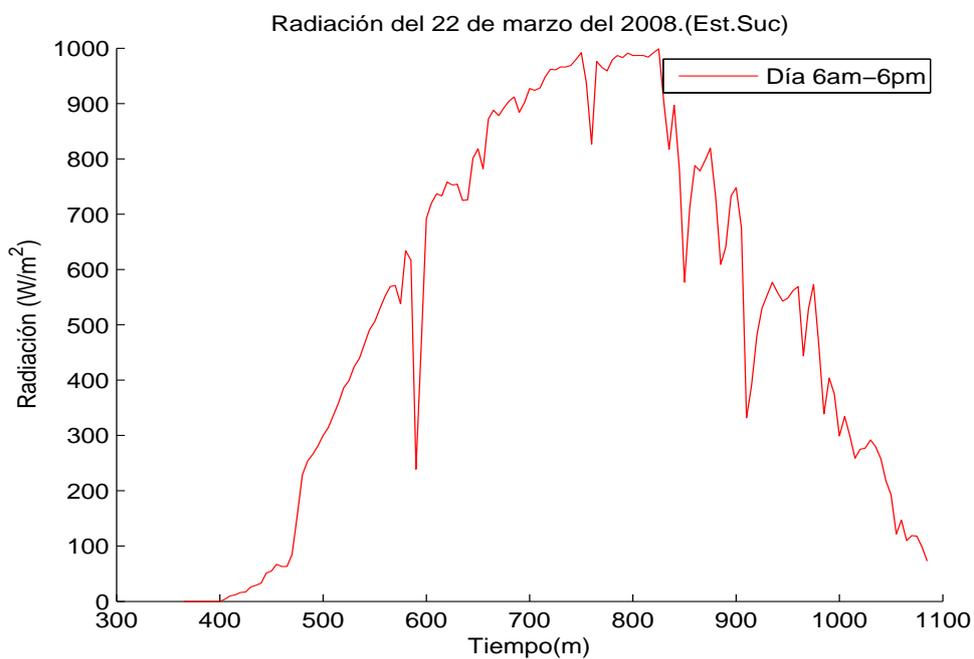


Tabla 3.14: Serie temporal de la radiación correspondiente al 22 de marzo del año 2008. (Est. Sucre)

En esta serie temporal, se puede apreciar intensidades bajas de la radiación repentinamente durante pequeños intervalos de tiempo.

### 3.2.2. Estación Tacarigua.

A continuación, la serie temporal del periodo 2007-2008 :

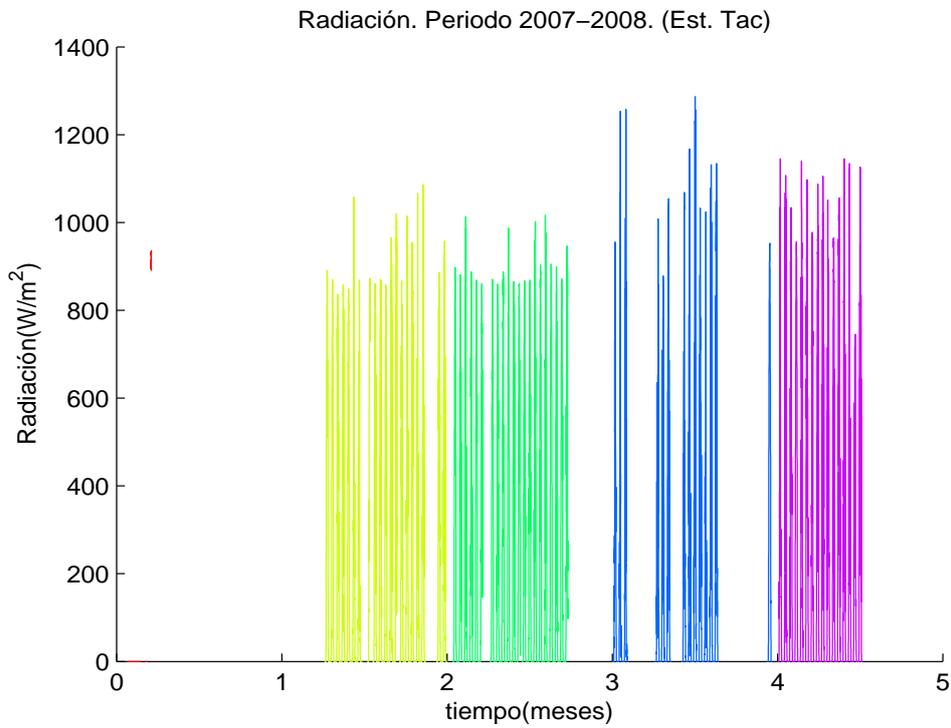


Figura 3.13: Serie temporal mensual de la radiación. (Est. Tacarigua)

Los parámetros estadísticos mensuales correspondientes al periodo 2007-2008, están representados en las siguientes tablas:

Parámetros Estadísticos	Año 2007	
	Noviembre	Diciembre
Min	0,00	0,00
Q1	0,00	69,00
Mediana	0,00	411,00
Media	19,16	414,65
Q3	0,00	722,00
Max	936,00	1085,00
D.Est	131,56	315,91
Varianza	1,73e+004	9,98e+004

Tabla 3.15: Radiación (Año 2007) (Est.Tac)

Parámetros Estadísticos	Año 2008		
	Enero	Julio	Agosto
Min	0,00	0,00	0,00
Q1	92,00	92,00	146,00
Mediana	479,50	278,50	388,00
Media	448,29	363,52	452,39
Q3	765,50	590,00	767,00
Max	1016,00	1286,00	1145,00
D.Est	321,11	308,90	332,13
Varianza	1,03e+005	9,54e+004	1,10e+005

Tabla 3.16: Radiación (Año 2008)(Est.Tac).

De acuerdo a las tablas anteriores, los registros mensuales correspondientes a la radiación sobre la superficie terrestre están muy dispersos respecto a su media.

Seguidamente, el boxplot del periodo 2007-2008:

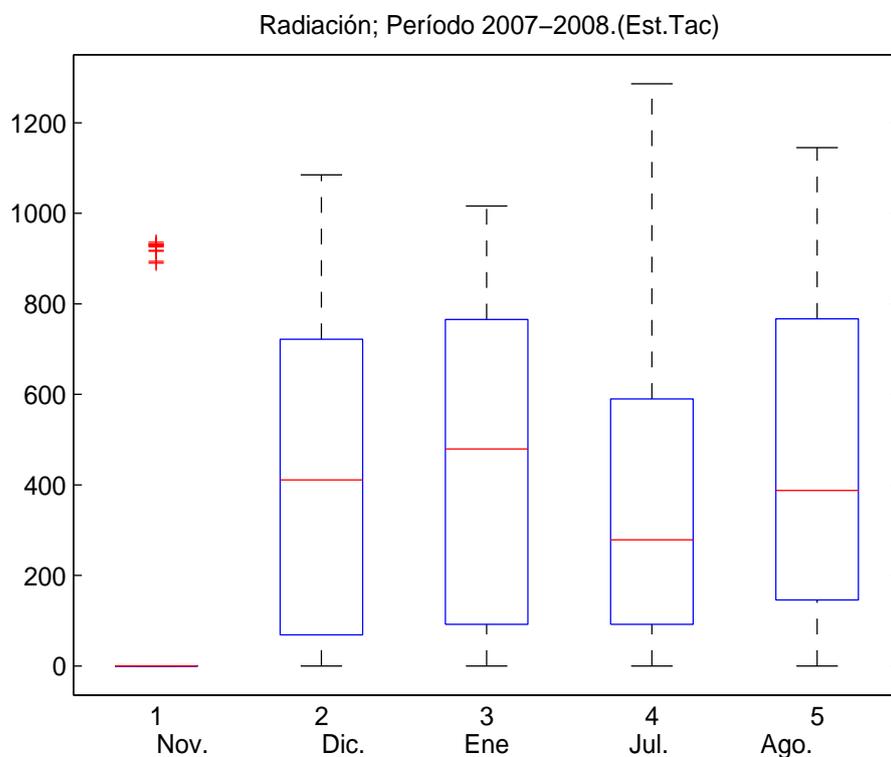


Figura 3.14: Boxplot mensual de la radiación. (Est. Tacarigua)

En la distribución de los registros mensuales, se destacan los meses julio y agosto al poseer la mayor cantidad de registros atípicos.

Tomemos como ejemplo, los meses enero y agosto del año 2008: en enero, se registrarán los datos de los primeros 23 días , los cuales estan representados en la siguiente serie temporal:

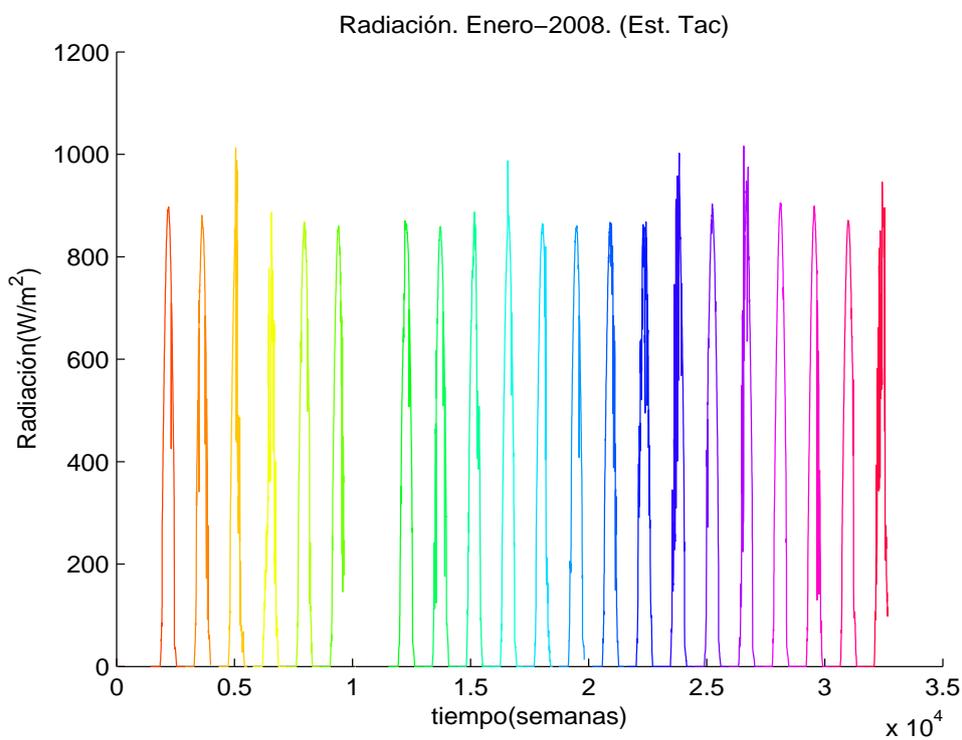


Figura 3.15: Serie temporal de la radiación correspondiente al mes de enero del año 2008. (Est. Tacarigua)

Por perdida total de registros, se considerarán los días desde el 2 hasta el 7 y desde el 9 hasta el 23 de enero. Los parámetros estadísticos diarios diurnos, se muestran en la siguiente tabla:

Días	Parámetros Estadísticos							
	Min	Q1	Mediana	Media	Q3	Max	D.Est	Varianza
2	0,00	76,00	561,00	488,00	808,25	897,00	332,76	1,10e+05
3	0,00	80,00	502,00	455,46	774,25	881,00	324,77	1,05e+05
4	0,00	57,50	323,00	350,40	595,25	1013,00	397,42	9,45e+04
5	0,00	75,00	246,00	300,42	440,25	887,00	252,96	6,39e+04
6	0,00	97,75	536,00	465,31	780,75	868,00	318,66	1,01e+05
7	0,00	218,00	588,00	494,05	779,00	860,00	308,13	9,49e+04
9	0,00	70,50	528,00	463,28	751,075	870,00	316,52	1,00e+05
10	0,00	91,50	509,00	455,44	772,50	859,00	318,00	1,01e+05
11	0,00	109,75	469,00	445,22	738,00	887,00	301,25	9,07e+04
12	0,00	95,75	534,00	476,55	783,25	987,00	322,96	1,04e+05
13	0,00	66,00	418,00	425,93	802,50	865,00	335,69	1,12e+05
14	0,00	140,00	524,00	459,42	769,00	860,00	317,01	1,00e+05
15	0,00	112,50	499,00	458,48	769,25	867,00	311,40	9,69e+04
16	0,00	129,25	472,00	449,12	742,00	868,00	298,64	8,91e+04
17	0,00	73,75	404,00	432,06	746,50	1002,01	337,76	1,14e+05
18	0,00	55,00	565,00	482,04	799,50	903,00	333,76	1,11e+05
19	0,00	44,50	493,00	452,66	818,50	1016,00	353,98	1,25e+05
20	0,00	105,25	575,00	500,58	812,00	905,00	331,21	1,09e+05
21	0,00	110,50	528,00	473,48	795,50	899,00	326,78	1,06e+05
22	0,00	101,50	565,00	472,68	781,25	871,00	324,22	1,05e+05
23	0,00	127,50	381,00	414,34	682,75	946,00	307,19	9,43e+04

Tabla 3.17: Radiación (Enero/día:6am-6pm)(Est.Tac)

Durante el periodo diurno, los registros se encuentran poco dispersos; sin embargo, los valores de la varianza estan muy influenciados respecto a la media.

Como ejemplo, la serie temporal correspondiente al día 12 de enero del año 2008:

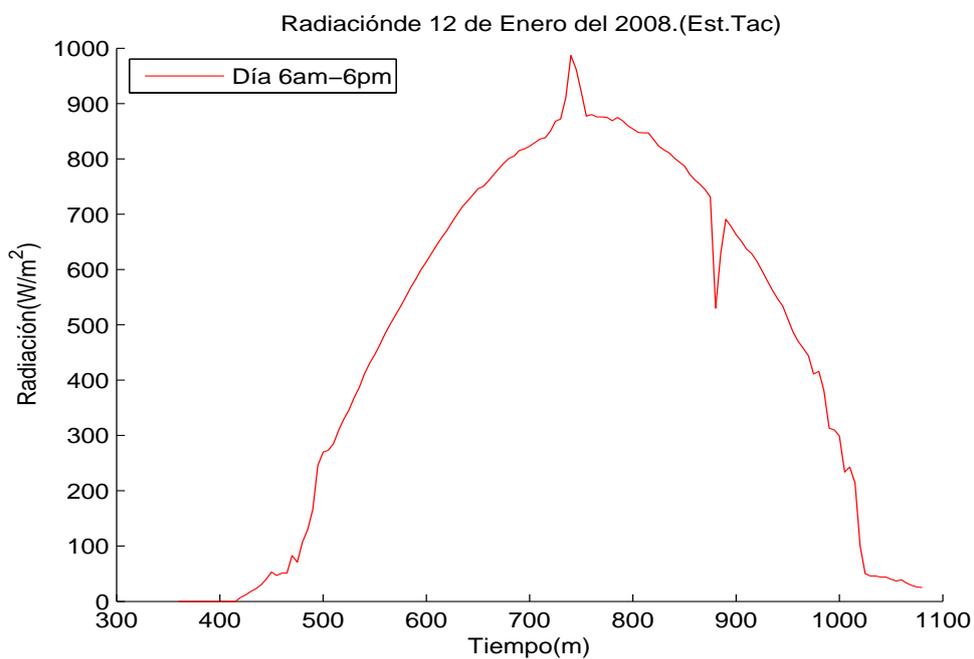


Figura 3.16: Serie temporal de la radiación correspondiente al 12 de enero del año 2008. (Est. Tacarigua)

Durante 12 de enero, a mediados del día se presentó una alta intensidad de la radiación. Esto podría deberse a la presencia o ausencia de nubosidad sobre el Lago de Valencia durante el período en cuestión.

En agosto, se registraron los datos de los primeros 16 días del mes, los cuales están representados en la siguiente serie temporal:

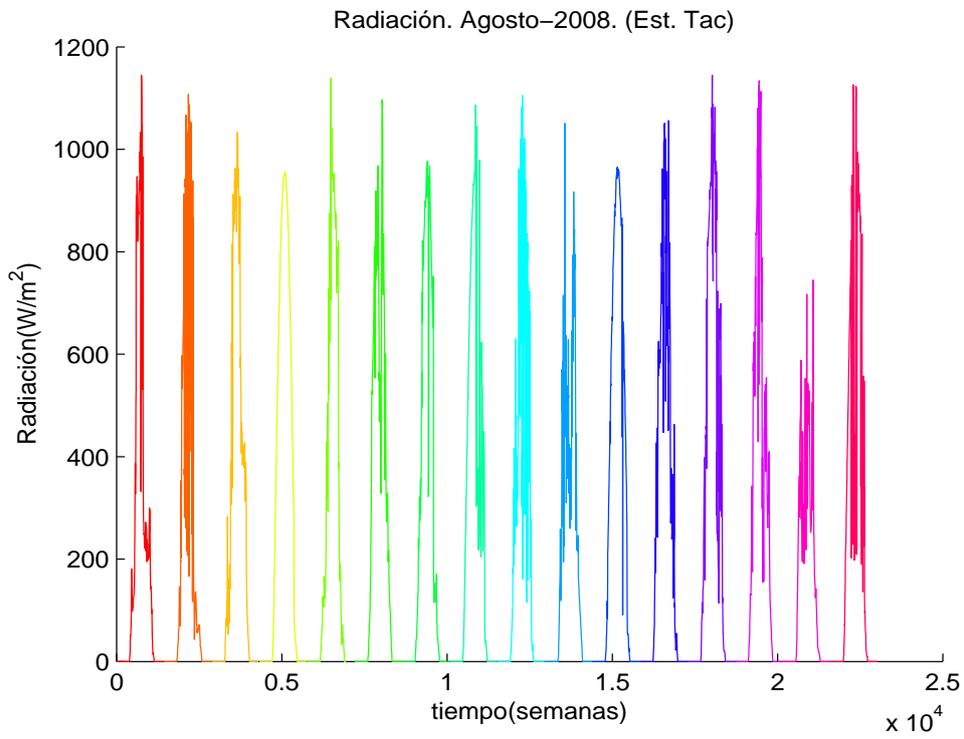


Figura 3.17: Serie temporal de la radiación correspondiente al mes de agosto del año 2008. (Est. Tacarigua)

Los parámetros estadísticos diarios diurnos ,se muestran en la siguiente tabla:

Días	Parámetros Estadísticos							
	Min	Q1	Mediana	Media	Q3	Max	D.Est	Varianza
1	0,00	137,50	238,00	379,68	687,75	1145,00	339,88	1,15e+05
2	0,00	69,00	202,00	366,16	632,75	1107,00	362,05	1,31e+05
3	0,00	130,00	403,00	466,04	842,00	1034,00	337,05	1,13e+05
4	0,00	314,75	643,00	571,35	8,77	956,00	323,44	1,04e+05
5	0,00	112,50	371,00	448,23	826,75	1139,00	367,76	1,35e+05
6	0,00	200,25	518,00	478,37	744,00	1097,00	315,20	9,93e+04
7	0,00	121,50	382,00	479,25	846,75	977,00	360,68	1,30e+05
8	0,00	228,50	397,00	464,72	735,75	1087,00	308,33	9,50e+04
9	0,00	150,00	384,00	442,30	724,00	1105,00	335,90	1,12e+04
10	0,00	158,00	345,00	366,28	539,00	1051,00	244,79	5,99e+04
11	0,00	292,00	623,00	557,51	876,75	965,00	331,99	1,10e+05
12	0,00	226,25	508,00	484,33	700,00	1056,00	315,03	9,92e+04
13	0,00	170,00	539,00	537,77	903,00	1145,00	365,34	1,33e+05
14	0,00	172,50	404,00	448,77	647,50	1134,00	320,28	1,02e+05
15	0,00	136,50	271,00	272,11	390,25	745,00	170,94	2,92e+04
16	0,00	208,75	403,00	475,42	812,25	1126,00	339,47	1,15e+05

Tabla 3.18: Radiación (Agosto/día:6am-6pm)(Est.Tac)

Durante el periodo diurno, los registros se encuentran poco dispersos; Sin embargo, los valores de la varianza estan muy influenciados respecto a la media.

Como ejemplo, la serie temporal correspondiente al 7 de agosto del año 2008:

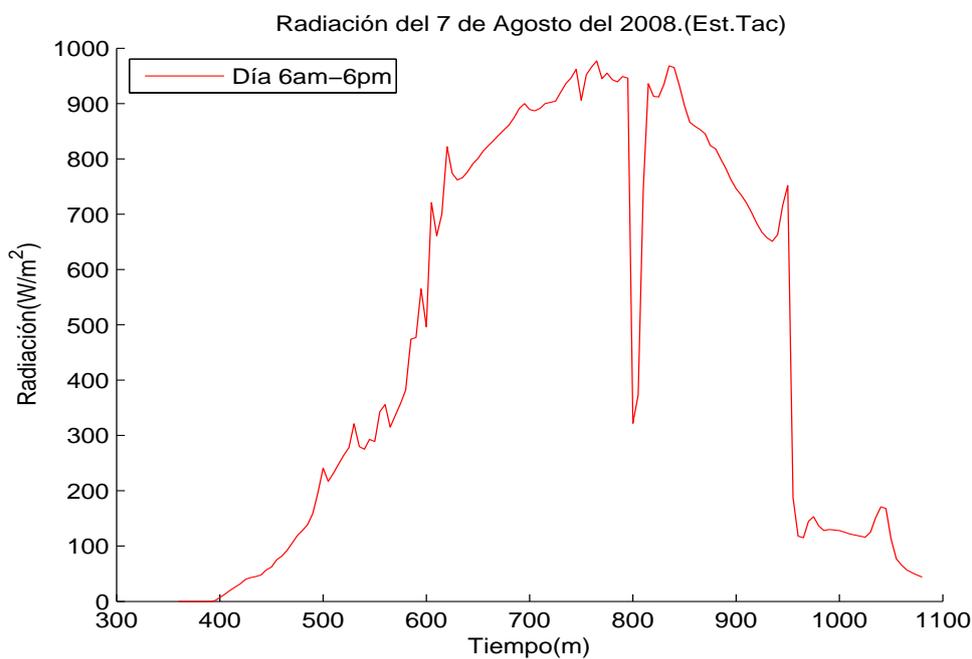


Figura 3.18: Serie temporal de la radiación correspondiente al 7 de agosto del año 2008. (Est. Tacarigua)

Durante el 7 de agosto del año 2008, puede observarse un caso atípico a mediados del medio día; indicando así altas y bajas intensidades de la radiación durante la tarde.

### 3.3. Precipitaciones.

En esta sección se presentará el estudio de las precipitaciones en las estaciones meteorológicas Sucre y Tacarigua respectivamente.

#### 3.3.1. Estación Sucre.

A continuación, la serie temporal del periodo 2007-2008 :

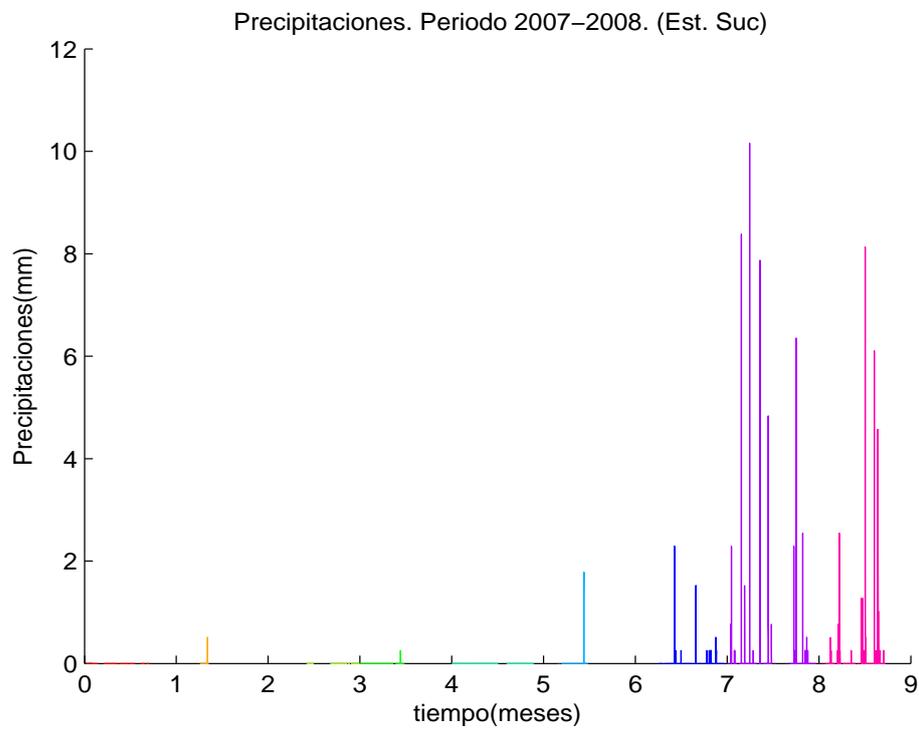


Figura 3.19: Serie temporal mensual de la precipitación. (Est. Suc)

Los parámetros estadísticos mensuales correspondientes al periodo 2007-2008 , están representados en las siguientes tablas:

Parámetros Estadísticos	Año 2007	
	Noviembre	Diciembre
Min	0,00	0,00
Q1	0,00	0,00
Mediana	0,00	0,00
Media	0,00	5,90e-004
Q3	0,00	0,00
Max	0,00	0,51
D.Est	0,00	0,01
Varianza	0,00	3,01e-004

Tabla 3.19: Precipitaciones. Estación de Sucre(Año 2007).

Parámetros Estadísticos	Año 2008						
	Enero	Febrero	Marzo	Abril	Junio	Sep	Oct
Min	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Q1	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Mediana	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Media	0,00	6,01e-005	0,00	0,00	0,00	0,02	0,02
Q3	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Max	0,00	0,25	0,00	1,78	2,29	10,16	8,13
D.Est	0,00	0,00	0,00	0,04	0,07	0,35	0,25
Varianza	0,00	1,50e-005	0,00	0,00	0,00	0,12	0,06

Tabla 3.20: Precipitaciones. Estación de Sucre(Año 2008).

En los meses correspondientes al período 2007-2008, los registros fueron en su mayoría nulos debido a un época de sequía. Para corroborar los valores de las tablas anteriores, se muestra el boxplot del periodo 2007-2008:

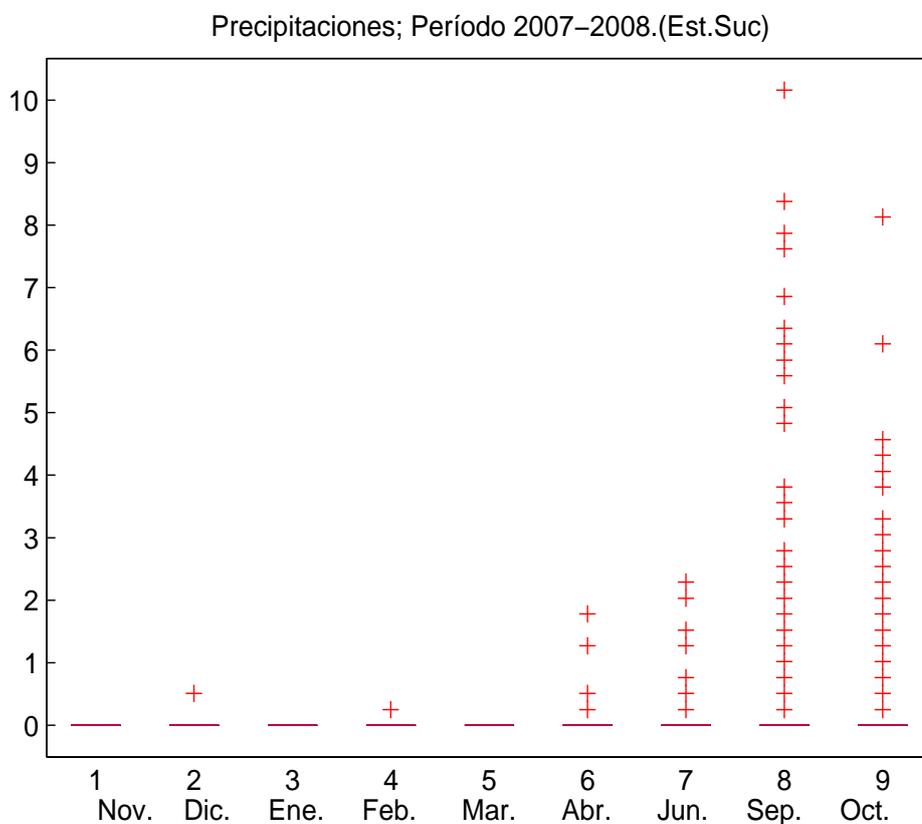


Figura 3.20: Boxplot mensual de la precipitación. (Est. Suc)

En la mayoría de los meses, se puede observar mucha pérdida de registros; sin embargo, en septiembre y octubre la muestra de datos es un poco mayor.

### 3.3.2. Estación Tacarigua.

A continuación, la serie temporal del periodo 2007-2008 :

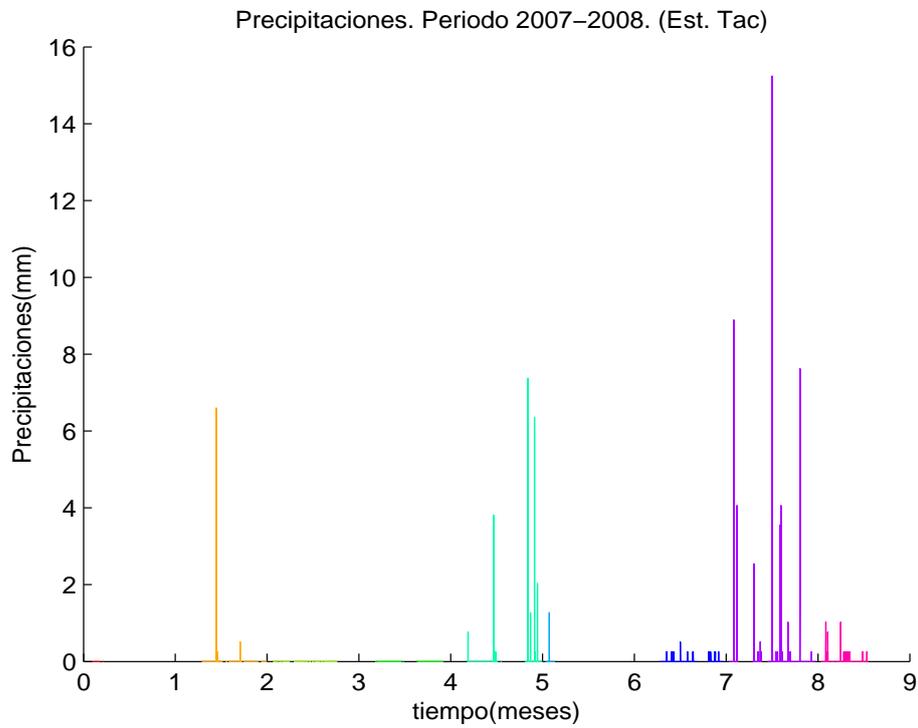


Figura 3.21: Serie temporal mensual de la precipitación. (Est. Tac)

Los parámetros estadísticos mensuales correspondientes al periodo 2007-2008 , están representados en las siguientes tablas:

Parámetros Estadísticos	Año 2007	
	Noviembre	Diciembre
Min	0,00	0,00
Q1	0,00	0,00
Mediana	0,00	0,00
Media	0,00	0,00
Q3	0,00	0,00
Max	0,00	6,60
D.Est	0,00	0,13
Varianza	0,00	0,01

Tabla 3.21: Precipitaciones (Año 2007)(Est.Tac)

Parámetros Estadísticos	Año 2008						
	Enero	Marzo	Abril	Mayo	Junio	Julio	Agosto
Min	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Q1	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Mediana	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Media	0,00	0,00	0,01	0,00	0,00	0,02	0,00
Q3	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Max	0,00	0,00	7,37	1,27	0,51	15,24	1,02
D.Est	0,00	0,00	0,18	0,05	0,01	0,38	0,03
Varianza	0,00	0,00	0,03	0,00	0,00	0,14	0,00

Tabla 3.22: Precipitaciones (Año 2008)(Est.Tac).

Los parámetros estadísticos respectivos a la medición de las precipitaciones durante el periodo 2007-2008 indican que fue un periodo de sequia; Por lo tanto, el estudio mensual de este periodo no será necesario.

Seguidamente, para corroborar los datos de las tablas anteriores se muestra el boxplot de los meses del periodo 2007-2008:

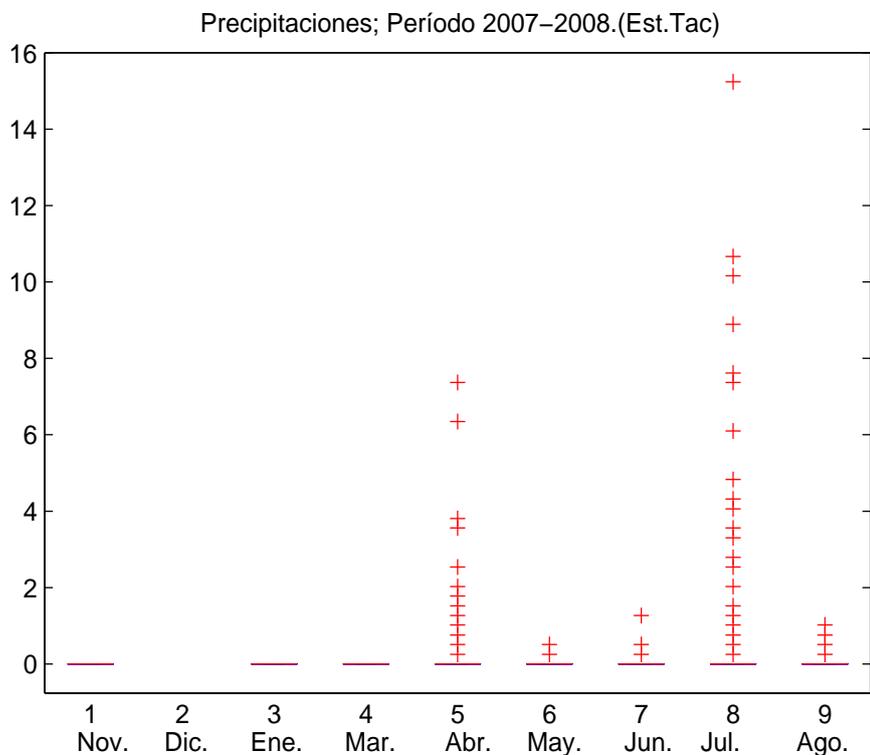


Figura 3.22: Boxplot mensual de la precipitación. (Est. Tac)

En la mayoría de los meses, se puede observar mucha pérdida de registros; debido a esto, no se profundizó más en esta parte.

### 3.4. Viento.

En esta sección se presentará el estudio de las velocidades del viento en las estaciones metereológicas Sucre y Tacarigua respectivamente.

#### 3.4.1. Estación Sucre.

A continuación, la serie temporal del periodo 2007-2008 :

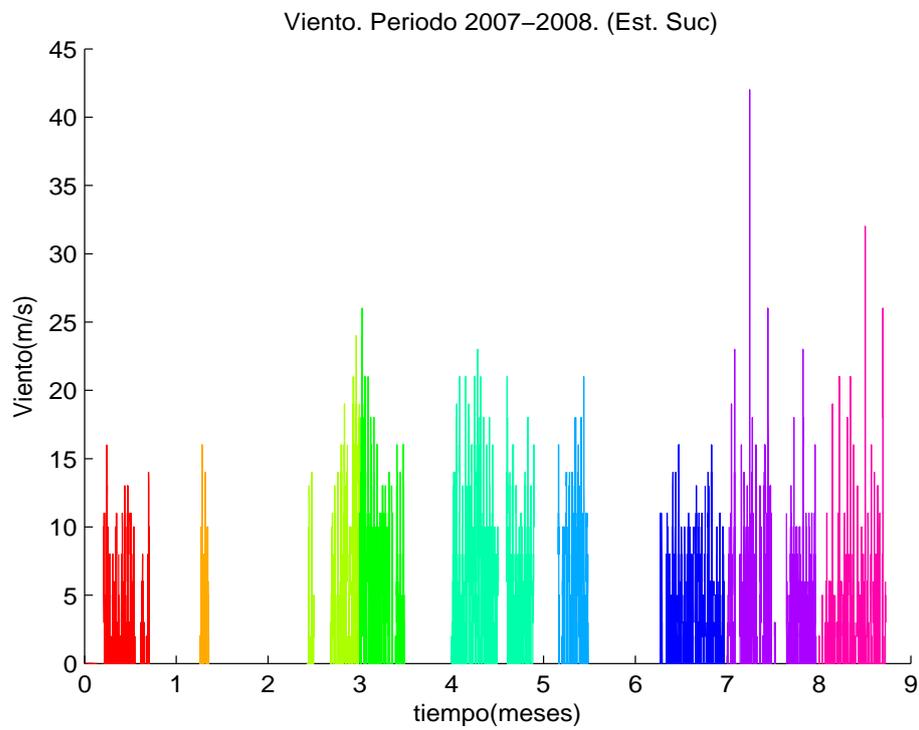


Figura 3.23: Serie temporal mensual del viento. (Est. Suc)

Los parámetros estadísticos mensuales correspondientes al periodo 2007-2008, se muestran en las siguientes tablas:

Parámetros Estadísticos	Año 2007	
	Noviembre	Diciembre
Min	0,00	0,00
Q1	0,00	0,00
Mediana	0,00	2,00
Media	1,89	3,31
Q3	3,00	6,00
Max	16,00	16,00
D.Est	3,11	3,57
Varianza	9,69	12,79

Tabla 3.23: Vientos. Estación de Sucre (Año 2007).

Parámetros Estadísticos	Año 2008						
	Enero	Febrero	Marzo	Abril	Junio	Sep	Oct
Min	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Q1	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Mediana	2,00	3,00	3,00	2,00	2,00	2,00	0,00
Media	4,34	4,42	4,64	3,70	2,51	2,66	1,90
Q3	8,00	8,00	8,00	8,00	5,00	5,00	3,00
Max	24,00	26,00	23,00	21,00	16,00	42,00	32,00
D.Est	4,97	4,91	4,82	4,50	3,10	3,47	3,04
Varianza	24,71	24,19	23,26	20,32	9,65	12,06	9,26

Tabla 3.24: Vientos. Estación de Sucre (Año 2008).

En los meses correspondientes al período 2007-2008, se presentaron vientos con velocidades irregulares; los cuales se estudiarán con más detalle más adelante.

Para corroborar los valores de las tablas anteriores, se muestra el boxplot del período 2007-2008:

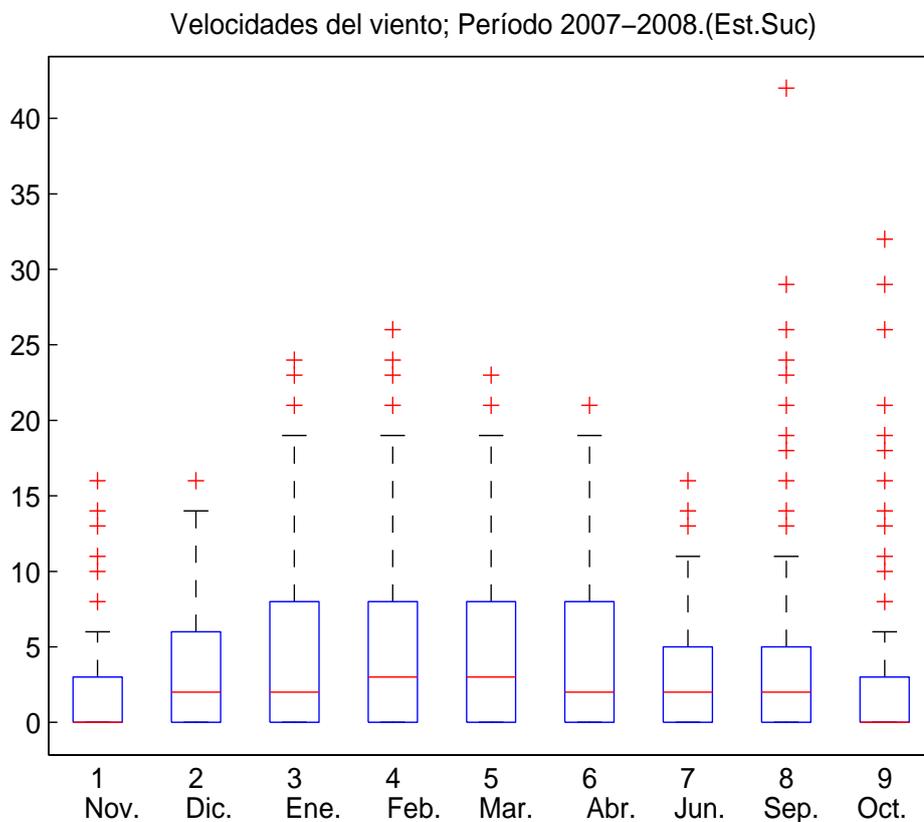


Figura 3.24: Boxplot mensual del viento. (Est. Suc)

Se puede observar la existencia de varios datos atípicos , lo que explica los valores de la varianza respecto a los valores de la media en las tablas anteriores.

Tomemos como ejemplo, los meses febrero y marzo del año 2008: en febrero, se registrarón los valores de los primeros 16 días del mes, los cuales están representados en la siguiente serie temporal:

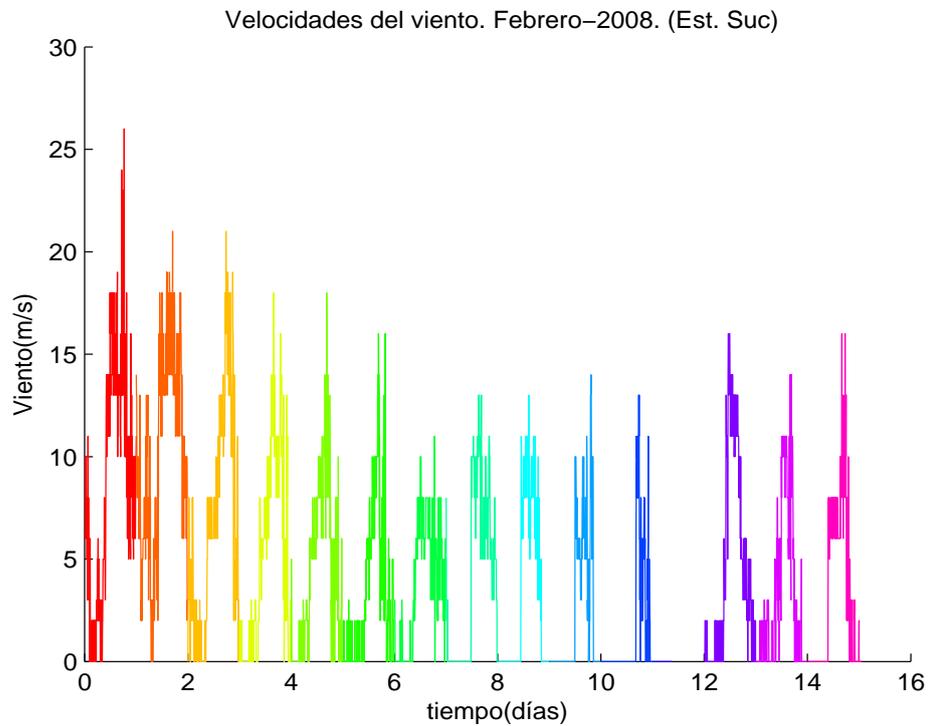


Figura 3.25: Serie temporal del viento correspondiente al mes de febrero del año 2008. (Est. Suc)

Se tomaron en cuenta los días desde el 1 hasta el 11 y desde 13 hasta el 15 de febrero; ya que , en los días restantes hubo perdida total de registros y otros días hubo registros nulos.

La siguiente tabla muestra los parámetros estadísticos diarios diurnos del mes:

Días	Parámetros Estadísticos							
	Min	Q1	Mediana	Media	Q3	Max	D.Est	Varianza
1	0,00	5,00	13,00	10,84	14,00	24,00	5,71	32,65
2	0,00	6,00	14,00	11,35	16,00	21,00	5,53	30,61
3	0,00	5,00	8,00	7,28	11,00	21,00	4,84	23,48
4	0,00	3,00	6,00	6,60	10,00	18,00	4,09	16,74
5	0,00	5,00	6,00	6,89	10,00	18,00	3,58	12,82
6	0,00	1,50	5,00	4,86	8,00	16,00	3,99	15,95
7	0,00	2,00	6,00	4,75	6,00	10,00	2,96	8,77
8	0,00	0,00	6,00	4,76	10,00	13,00	4,68	21,90
9	0,00	0,00	6,00	5,29	10,00	13,00	4,51	20,34
10	0,00	0,00	2,00	3,15	6,00	11,00	3,41	11,68
11	0,00	0,00	0,00	1,36	0,00	13,00	3,36	11,33
13	0,00	2,00	10,00	7,72	13,00	16,00	5,30	28,14
14	0,00	2,00	5,00	5,48	10,00	14,00	4,20	17,69
15	0,00	0,00	6,00	5,60	8,00	16,00	4,05	16,47

Tabla 3.25: Velocidades del viento (Febrero/día:6am-6pm)(Est.Suc)

En la mayoría de los días, la varianza está influenciada respecto a su media; indicando, la existencia de valores atípicos en los registros de datos. La próxima tabla muestra los parámetros estadísticos diarios nocturnos.

Días	Parámetros Estadísticos							
	Min	Q1	Mediana	Media	Q3	Max	D.Est	Varianza
1	0,00	2,00	6,00	6,93	10,00	26,00	5,24	27,52
2	2,00	6,00	10,00	9,52	13,00	18,00	3,91	15,35
3	0,00	2,00	3,00	6,29	11,00	19,00	5,87	34,56
4	0,00	0,00	1,00	3,46	6,00	16,00	4,73	22,45
5	0,00	0,00	0,00	1,82	3,00	11,00	2,64	6,98
6	0,00	0,00	0,00	1,96	3,00	16,00	3,16	10,04
7	0,00	0,00	2,00	2,32	5,00	11,00	2,63	6,92
8	0,00	0,00	2,00	3,02	5,00	11,00	3,31	11,01
9	0,00	0,00	0,00	1,01	0,00	10,00	2,26	5,11
10	0,00	0,00	0,00	1,30	0,00	14,00	3,24	10,54
11	0,00	0,00	0,00	1,45	3,00	11,00	2,41	5,84
13	0,00	0,00	0,00	1,57	3,00	6,00	1,87	3,52
14	0,00	0,00	0,00	0,63	2,00	5,00	1,08	1,18
15	0,00	0,00	0,00	0,85	0,00	10,00	1,86	3,46

Tabla 3.26: Velocidades del viento (Febrero/noche:6pm-6am)(Est.Suc)

En la mayoría de los días, los valores de la desviación estándar y la varianza están alterados respecto a su media; indicando la existencia de casos atípicos durante la noche.

En marzo, se registrarán los valores de los primeros 28 días, los cuales están representados en la siguiente serie temporal:

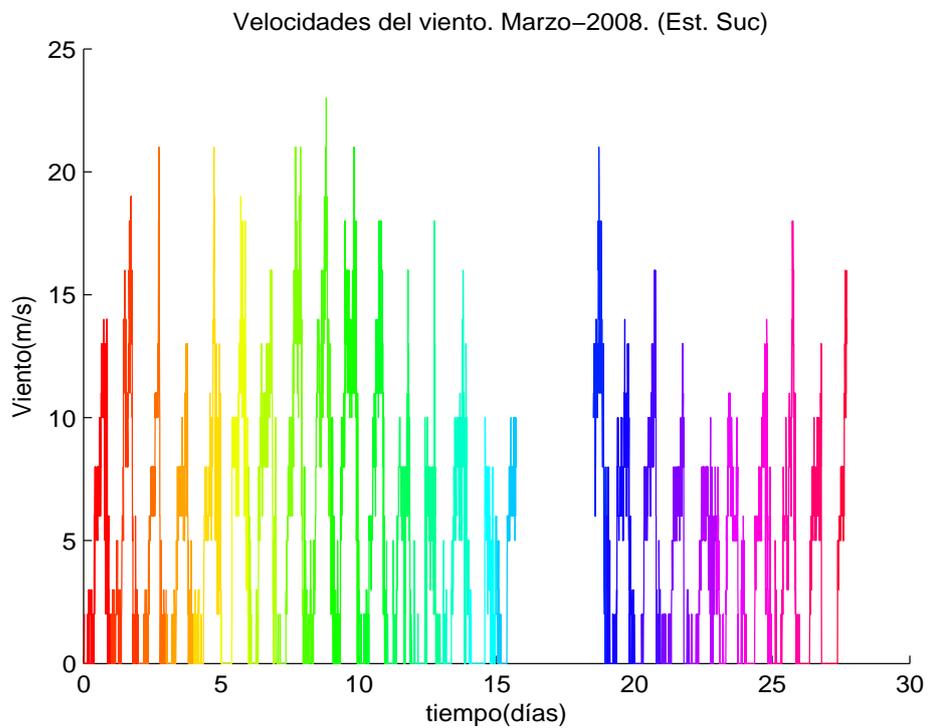


Figura 3.26: Serie temporal del viento correspondiente al mes de marzo del año 2008. (Est. Sucre)

Se tomaron en cuenta los días desde el 1 hasta el 16 y desde 19 hasta el 28 de marzo; ya que, en los días restantes hubo pérdida total de registros. La siguiente tabla muestra los parámetros estadísticos diarios diurnos del mes:

Días	Parámetros Estadísticos							
	Min	Q1	Mediana	Media	Q3	Max	D.Est	Varianza
1	0,00	2,00	6,00	6,28	10,00	14,00	4,00	16,05
2	0,00	3,00	10,00	8,80	13,00	19,00	5,59	31,28
3	0,00	4,50	6,00	6,95	11,00	21,00	4,57	20,88
4	0,00	3,00	6,00	6,15	8,00	13,00	3,28	10,76
5	0,00	5,00	6,00	6,40	8,00	21,00	4,66	21,79
6	0,00	0,00	8,00	7,42	10,00	19,00	4,97	24,71
7	0,00	6,00	10,00	7,83	11,00	13,00	3,83	14,69
8	0,00	3,00	8,00	7,70	11,00	21,00	5,38	28,97
9	0,00	3,00	10,00	8,45	13,00	19,00	5,25	27,56
10	0,00	8,00	11,00	9,71	13,00	18,00	4,69	22,06
11	0,00	5,00	6,00	7,42	11,00	18,00	4,48	20,13
12	0,00	2,75	6,00	5,24	8,00	10,00	2,92	8,53
13	0,00	0,00	3,00	3,82	6,00	18,00	3,52	12,40
14	0,00	3,00	6,00	5,73	8,00	13,00	3,68	13,55
15	0,00	0,00	0,00	2,46	6,00	10,00	3,38	11,45
16	0,00	0,00	6,00	5,25	8,00	10,00	3,47	12,08
19	6,00	10,00	11,00	12,30	14,00	21,00	2,94	8,67
20	0,00	5,00	6,00	6,21	8,00	14,00	3,42	11,72
21	0,00	6,00	8,00	7,12	10,00	16,00	3,82	14,66
22	0,00	4,50	6,00	5,27	8,00	13,00	2,91	8,52
23	0,00	3,00	5,00	4,70	6,00	8,00	2,55	6,52
24	0,00	5,00	6,00	5,88	8,00	11,00	3,07	9,43
25	0,00	2,75	5,00	5,17	8,00	13,00	3,58	12,82
26	0,00	3,00	6,00	5,78	8,00	18,00	4,07	16,61
27	0,00	2,00	6,00	4,86	8,00	10,00	3,15	9,98
28	0,00	0,00	6,00	5,66	8,00	16,00	4,50	20,26

Tabla 3.27: Velocidades del viento (Marzo/día:6am-6pm)(Est.Suc)

En la mayoría de los días, la varianza está influenciada respecto a su media; indicando, la existencia de valores atípicos en los registros de los datos. Sin embargo, en los días 12, 19, 22 y 23 de marzo los valores de

la desviación estándar y la varianza no están influenciados indicando vientos con velocidades regulares. La proxima tabla muestra los parámetros estadísticos diarios nocturnos.

Días	Parámetros Estadísticos							
	Min	Q1	Mediana	Media	Q3	Max	D.Est	Varianza
1	0,00	0,00	0,00	2,93	5,00	14,00	4,34	18,89
2	0,00	0,00	2,00	2,19	3,00	16,00	2,93	8,64
3	0,00	0,00	0,00	1,30	2,00	16,00	2,71	7,37
4	0,00	0,00	0,00	1,47	2,00	13,00	3,19	10,18
5	0,00	0,00	2,00	4,04	8,00	18,00	4,73	22,44
6	0,00	0,00	0,00	4,39	10,00	18,00	5,55	30,84
7	0,00	0,00	2,00	4,36	8,00	16,00	5,19	27,00
8	0,00	0,00	0,00	5,69	13,00	21,00	6,97	48,65
9	0,00	0,00	5,00	7,14	14,00	23,00	6,87	47,25
10	0,00	2,00	5,50	6,78	13,00	21,00	6,10	37,28
11	0,00	0,00	2,00	4,35	11,00	18,00	5,78	33,41
12	0,00	0,00	0,00	2,21	2,00	16,00	3,95	15,60
13	0,00	0,00	0,00	0,73	0,00	13,00	2,32	5,38
14	0,00	0,00	2,00	3,66	5,00	16,00	4,57	20,95
15	0,00	0,00	0,00	1,28	2,00	8,00	1,85	3,43
16	0,00	0,00	0,00	0,72	2,00	6,00	1,28	1,64
19	0,00	5,00	10,00	9,01	13,00	18,00	5,26	27,70
20	0,00	0,00	0,00	2,00	3,00	13,00	3,20	10,24
21	0,00	0,00	0,00	1,33	0,00	16,00	3,27	10,72
22	0,00	0,00	0,00	1,26	2,00	13,00	2,92	8,56
23	0,00	0,00	0,00	1,93	3,00	10,00	2,42	5,86
24	0,00	0,00	0,00	1,65	3,00	8,00	2,01	4,07
25	0,00	0,00	0,00	2,23	3,00	14,00	3,63	13,23
26	0,00	0,00	0,00	2,55	2,00	18,00	4,82	23,26
27	0,00	0,00	0,00	0,84	0,00	13,00	2,74	7,50
28	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Tabla 3.28: Velocidades del viento (Marzo/noche:6pm-6am)(Est.Suc)

En la mayoría de los días, los valores de la desviación estándar y la varianza están alterados respecto a su media; indicando la existencia de casos atípicos durante la noche.

### 3.4.2. Estación Tacarigua.

A continuación, la serie temporal del periodo 2007-2008 :

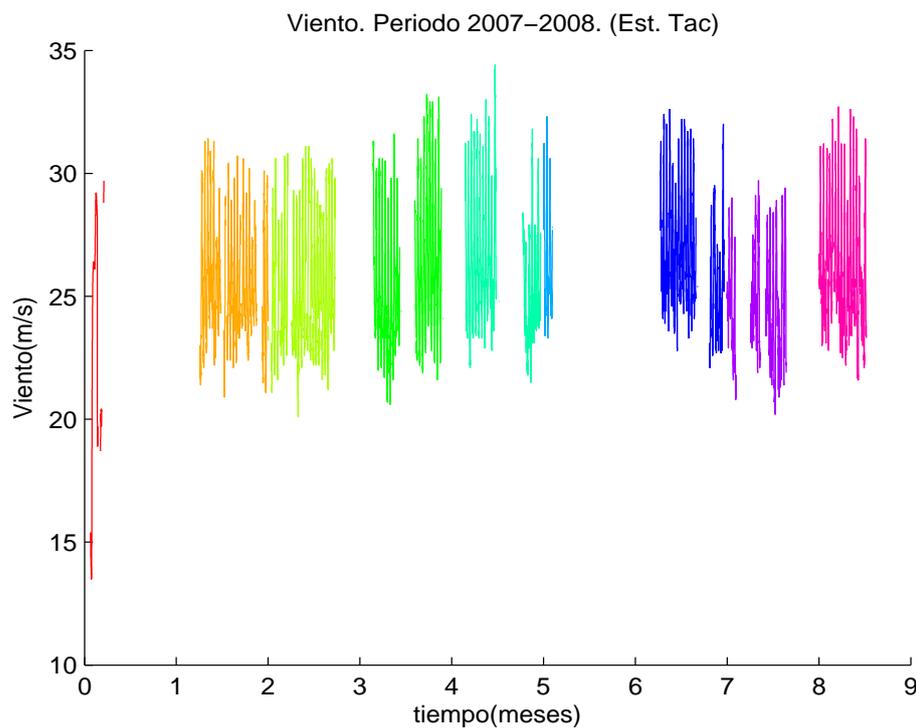


Figura 3.27: Serie temporal mensual del viento. (Est. Tacarigua)

Los parámetros estadísticos mensuales correspondientes al periodo 2007-2008, están representados en las siguientes tablas:

Parámetros Estadísticos	Año 2007	
	Noviembre	Diciembre
Min	0,00	0,00
Q1	0,00	5,00
Mediana	0,00	8,00
Media	0,05	8,80
Q3	0,00	13,00
Max	6,00	39,00
D.Est	0,50	5,58
Varianza	0,25	31,13

Tabla 3.29: Velocidades del viento (Año 2007) (Est.Tac)

Parámetros Estadísticos	Año 2008						
	Enero	Marzo	Abril	Mayo	Junio	Julio	Agosto
Min	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Q1	6,00	1,30	0,40	0,00	0,40	0,00	0,00
Mediana	10,00	2,20	1,80	0,90	2,20	2,00	3,00
Media	10,39	2,59	2,00	1,46	2,23	3,07	4,63
Q3	14,00	3,60	3,10	2,70	3,60	4,50	8,00
Max	37,00	8,90	10,30	6,30	8,90	37,00	40,00
D.Est	6,02	1,59	1,61	1,55	1,84	4,21	5,09
Varianza	36,31	2,55	2,59	2,40	3,40	17,78	25,95

Tabla 3.30: Velocidades del viento (Año 2008) (Est.Tac).

En la tabla 27, el mes de noviembre del año 2007 tiene una media muy baja indicando vientos muy lentos y pérdida de información. A diferencia del mes de diciembre del año 2007 y del mes de enero del año 2008, con medias más destacadas y varianzas elevadas indicando vientos menos lentos y algunos casos atípicos.

Los meses restantes, representan vientos lentos; sin embargo, los meses de julio y agosto del año 2008 poseen varianzas influenciadas respecto a su media respectivamente indicando así casos atípicos.

Seguidamente, para corroborar los valores de las tablas anteriores ; el boxplot del periodo 2007-2008:

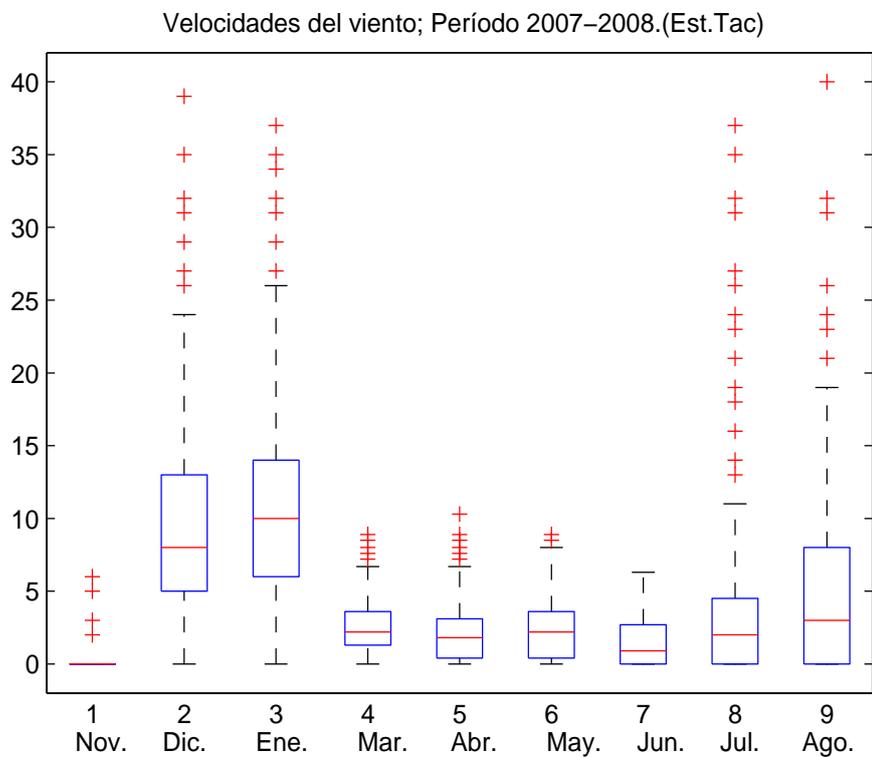


Figura 3.28: Boxplot mensual del viento. (Est. Tacarigua)

Tomemos como ejemplo, el mes de diciembre del año 2008: en diciembre, se registrarón los datos de los primeros 31 días del mes, los cuales estan representados en la siguiente serie temporal:

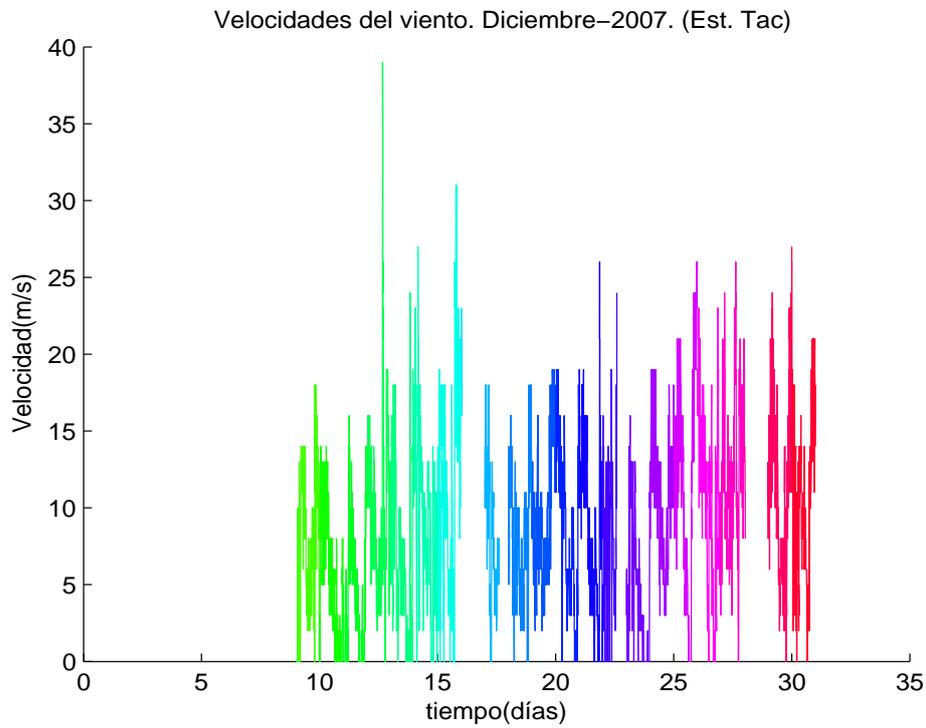


Figura 3.29: Serie temporal del viento correspondiente al mes de diciembre del año 2008. (Est. Tacarigua)

Debido a perdida de registros, se consideraron los días desde el 10 hasta el 16, desde el 19 hasta el 22, desde el 24 hasta el 28 y desde 30 hasta el 30 de diciembre.

La siguiente tabla, muestra los parámetros estadísticos diarios diurnos del mes:

Días	Parámetros Estadísticos							
	Min	Q1	Mediana	Media	Q3	Max	D.Est	Varianza
10	2,00	5,00	6,00	7,38	10,00	14,00	3,22	10,37
11	0,00	3,00	5,00	5,99	8,00	14,00	3,34	11,21
12	0,00	2,75	6,00	5,64	8,00	16,00	3,80	14,45
13	2,00	5,00	8,00	8,83	10,00	39,00	6,06	36,83
14	0,00	3,00	5,00	4,56	6,00	10,00	2,61	6,81
15	0,00	6,00	8,00	8,30	11,00	18,00	3,90	15,21
16	0,00	3,00	6,00	8,10	11,00	26,00	5,93	35,23
19	2,00	5,00	6,00	6,56	8,00	13,00	2,56	6,58
20	2,00	5,75	8,00	7,24	10,00	16,00	2,48	6,19
21	0,00	5,00	6,00	7,20	10,00	16,00	3,55	12,62
22	0,00	5,00	6,00	7,34	10,00	16,00	3,95	15,67
24	0,00	2,00	5,00	4,69	6,00	13,00	3,48	12,14
25	3,00	6,00	8,00	7,95	10,25	14,00	2,88	8,29
26	0,00	2,00	6,00	7,60	13,00	21,00	5,91	35,04
27	0,00	3,00	10,00	7,99	13,00	18,00	5,13	26,35
28	0,00	9,50	11,00	11,53	13,00	26,00	5,07	25,73
30	0,00	5,00	8,00	8,45	13,00	19,00	4,36	19,04
31	0,00	5,00	8,00	7,85	10,00	16,00	3,65	13,33

Tabla 3.31: Velocidades del viento(Diciembre/día:6am-6pm)(Est.Tac)

En la mayoría de los días del mes, los registros no se encuentran tan dispersos respecto a su media. Sin embargo, durante el 13, 16, 26, 27 y 28 de diciembre, los valores de la varianza están influenciados indicando así registros atípicos en la distribución de los datos.

La próxima tabla muestra los parámetros estadísticos diarios nocturnos del mes:

Días	Parámetros Estadísticos							
	Min	Q1	Mediana	Media	Q3	Max	D.Est	Varianza
10	0,00	2,00	8,00	7,44	13,00	18,00	5,61	31,51
11	0,00	0,00	5,00	4,94	8,00	14,00	4,56	20,82
12	0,00	0,00	2,00	3,27	5,00	14,00	4,32	18,71
13	0,00	8,00	11,00	10,64	13,00	19,00	4,63	21,52
14	0,00	3,00	10,00	8,65	13,00	24,00	5,67	32,21
15	0,00	3,00	10,00	9,84	14,00	27,00	6,89	47,60
16	0,00	10,00	14,00	13,78	18,00	31,00	6,82	46,64
19	0,00	6,00	10,00	9,41	13,00	18,00	4,89	23,92
20	2,00	8,00	11,00	11,76	16,00	19,00	4,46	19,90
21	0,00	3,00	11,00	9,56	14,00	19,00	6,20	38,46
22	0,00	5,00	11,00	10,19	14,00	26,00	5,69	32,37
24	0,00	0,00	0,00	3,69	8,00	16,00	4,89	23,92
25	6,00	11,00	13,00	12,88	14,00	19,00	3,05	9,34
26	6,00	13,00	15,00	15,83	19,00	26,00	4,77	22,81
27	0,00	8,00	13,00	12,17	16,00	23,00	5,25	27,66
28	2,00	10,00	14,00	13,73	18,00	24,00	4,43	19,65
30	0,00	11,00	14,00	14,11	18,00	26,00	5,30	28,17
31	0,00	10,00	13,00	12,55	16,00	27,00	5,41	29,32

Tabla 3.32: Velocidades del viento(Diciembre/noche:6pm-6am)(Est.Tac)

En la mayoría de los días del mes, los valores de la varianza están influenciados indicando así registros atípicos en la distribución de los datos a excepción del 25 de diciembre, el cual posee registros poco dispersos respecto a su media.

## Análisis de correlaciones cruzadas

En este capítulo se realizará un análisis de correlaciones entre las variables climáticas ya estudiadas en el capítulo 3, se hará el análisis entre las siguientes series:

- Viento-temperatura.
- Radiación-viento.
- Viento-precipitación.
- Temperatura-radiación.
- Temperatura-precipitación.
- Radiación-precipitación.

Este estudio se hará para los datos de ambas estaciones meteorológicas, con paso máximo  $h=48$ . Cada paso corresponde a 5 minutos de tiempo.

Para el estudio de las correlaciones cruzadas de la radiación con el resto de las variables, solo se consideró el lapso de tiempo: desde las 6:00am hasta las 6:00pm; ya que, la radiación solo puede medirse durante el día mediante las ondas emitidas por el sol.

Una vez calculadas las correlaciones cruzadas diarias de cada combinación se determinaron los  $h$  óptimos para el mes, en función de las máximas correlaciones observadas para cada día en que hubo correlación; luego, se promediaron dichos valores de los  $h$  óptimos ( $\mu_{h_o}$ ); la tabla 4.1 muestra los  $\mu_{h_o}$  correspondientes a la estación Sucre:

Correlaciones	$\mu_{h_o}$								
	Nov	Dic	Ene	Feb	Mar	Abr	Jun	Sep	Oct
Vie-Tem	9	6	6	3	7	5	5	10	9
Rad-Vie	9	20	21	23	22	23	14	13	18
Vie-Pre	NaN	20	NaN	19	NaN	4	25	20	13
Tem-Rad	29	48	22	15	29	27	14	21	33
Tem-Pre	NaN	1	NaN	37	NaN	25	17	26	26
Rad-Pre	NaN	48	NaN	48	NaN	48	34	6	25

Tabla 4.1: Medias de los  $h$  óptimos correspondientes a cada correlación cruzada y a cada mes del periodo 2007-2008. (Est.Suc)

En la tabla 4.1, los valores *NaN* indican que para ese mes la correlación cruzada en las variables no fue posible calcularla, debido a la falta de datos para el mes de dicha combinación. Además, es importante recalcar que para algunas correlaciones cruzadas en las variables, no se puede determinar un patrón de comportamiento de los valores  $h_o$  en algunos meses específicos; ya que, dichos meses presentan una escasa cantidad de días para los cuales se calculó la correlación cruzada. Dichos valores de  $\mu_{h_o}$  corresponden a:

- La correlación cruzada en las variables viento y temperatura, para el mes de diciembre del año 2007.
- La correlación cruzada en las variables viento y precipitación, para los meses de diciembre del año 2007 y febrero y abril del año 2008.
- La correlación cruzada en las variables temperatura y radiación, para el mes de diciembre del año 2007.
- La correlación cruzada en las variables temperatura y precipitación, para los meses de diciembre del año 2007 y febrero y abril del año 2008.
- La correlación cruzada en las variables radiación y precipitación, para los meses de diciembre del año 2007 y febrero, abril, junio y septiembre del año 2008.

En el comportamiento de los valores de  $h_o$  correspondientes a la correlación cruzada en las variables viento y temperatura para todos los meses del periodo 2007-2008 (excepto diciembre); en su mayoría, se observó que las máximas correlaciones de cada día se alcanzaron en un paso  $h_o = 0$ . Para corroborar el análisis hecho, el boxplot 4.1 muestra la distribución de los pasos  $h$ :

Distribución de los pasos "h" de la correlación cruzada entre viento y temperatura.

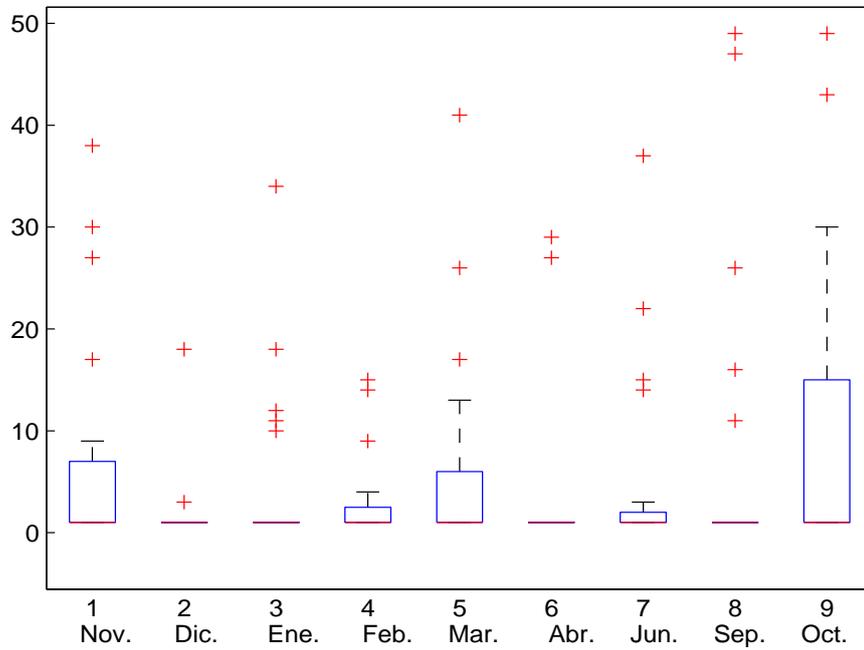


Figura 4.1: Boxplot de distribución de los pasos  $h$  correspondientes a la correlación cruzada en las variables viento y la temperatura. (Est.Suc)

Tomemos como ejemplo los gráficos de dispersión correspondientes al 10 de abril del año 2008:

Gráfico de dispersión entre viento y temperatura con paso promedio  $h=5$  y recta de regresión

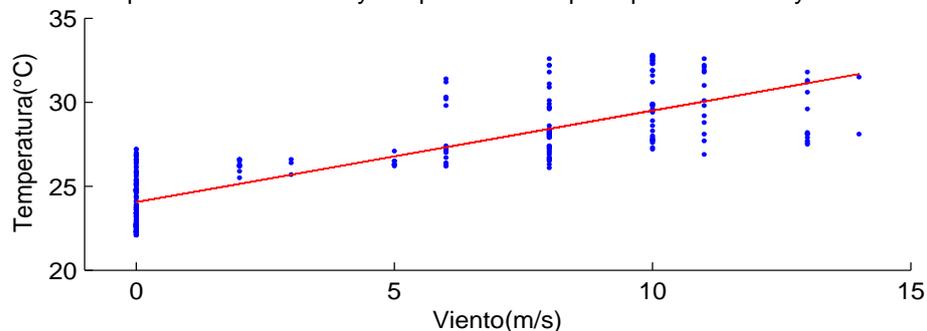


Gráfico de dispersión entre viento y temperatura con paso óptimo  $h=0$  y recta de regresión.

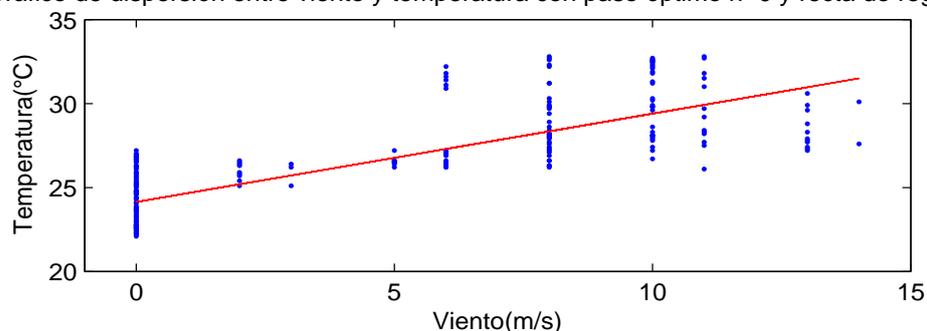


Figura 4.2: Gráficos de dispersión y rectas de regresión con paso promedio  $h = 5$  y paso óptimo  $h = 0$  de la correlación cruzada entre el viento y la temperatura (10/04/2008) (Est.Suc)

En ambos gráficos de dispersión existe una relación positiva entre ambas variables: conforme aumenta la velocidad del viento, también aumenta la temperatura. También, puede observarse que no existe un instante en que el viento sea muy veloz y la temperatura sea muy baja y tampoco un instante en que la temperatura sea muy alta y el viento sea muy lento.

Las pendientes de las rectas de regresión de ambos gráficos de dispersión son positivas, indicando que, existe una correlación positiva o directa en las variables. Sin embargo, puede observarse que la recta de regresión del primer gráfico de dispersión está más inclinada que la recta de regresión del segundo gráfico de dispersión; indicando, que existe una mejor correlación para el paso promedio del mes  $\mu_{h_0} = 5$  que para el paso en que se alcanzó la máxima correlación ese día  $h_0 = 0$ .

En cuanto al patrón de comportamiento de los valores de los  $h_0$  correspondientes a la correlación cruzada en las variables viento y radiación para los meses de noviembre del año 2007 y junio y octubre del año 2008; mayormente, se observó que las correlaciones máximas diarias se alcanzaron en un paso  $\mu_{h_0} = 10$ . Respecto a

los meses de enero y febrero del año 2008 las máximas correlaciones diarias se alcanzaron en un paso  $\mu_{h_0} = 21$  y  $\mu_{h_0} = 23$ , respectivamente. Por último, para los meses de diciembre del año 2007 y marzo, abril y septiembre del año 2008 no se pudo determinar un patrón de comportamiento de los valores  $h_0$ , pues los valores de  $h_0$  en que se alcanzaron las máximas correlaciones diarias fueron muy diferentes entre sí. Para corroborar el análisis hecho, el boxplot 4.3 muestra la distribución de los pasos  $h$ :

Distribución de los pasos "h" de la correlación cruzada entre viento y radiación.

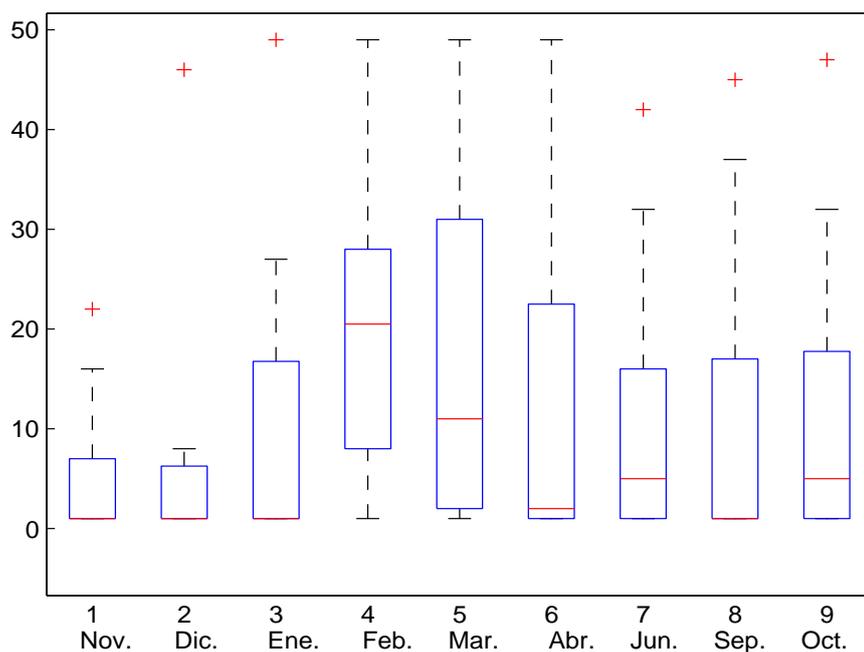


Figura 4.3: Boxplot de distribución de los pasos  $h$  correspondientes a la correlación cruzada en las variables viento y la radiación. (Est.Suc)

Tomemos como ejemplo los gráficos de dispersión correspondientes al 22 de junio del año 2008:

Gráfico de dispersión entre viento y radiación con paso promedio  $h=14$  y recta de regresión.

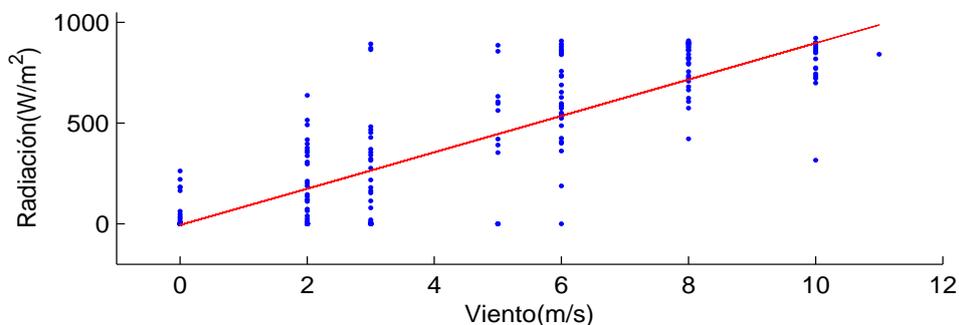


Gráfico de dispersión entre viento y radiación con paso óptimo  $h=4$  y recta de regresión.

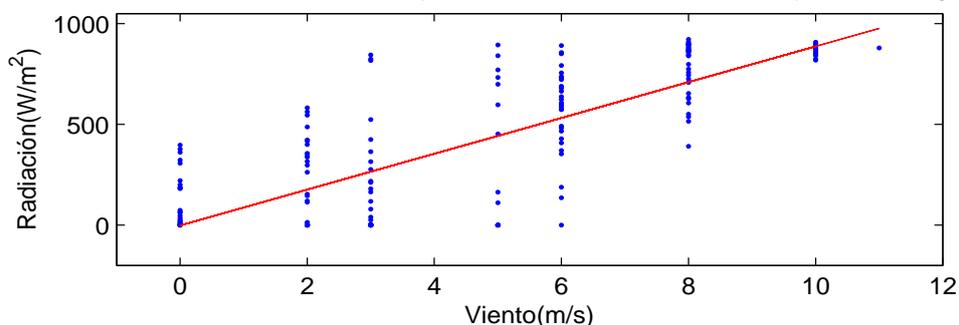


Figura 4.4: Gráficos de dispersión y rectas de regresión con paso promedio  $h = 14$  y paso óptimo  $h = 4$  correspondiente a la correlación cruzada en las variables viento y la radiación (22/06/2008) (Est.Suc)

En ambos gráficos de dispersión existe una relación positiva entre ambas variables: conforme aumenta la velocidad del viento, también aumenta la radiación. También, puede observarse que no existe un instante en que el viento sea muy veloz y la radiación sea muy baja. Sin embargo, puede observarse, dos casos atípicos en que el viento posee una velocidad regular y la radiación es muy baja, esto podría ser por nubosidad presente sobre el lago.

Las pendientes de las rectas de regresión de ambos gráficos de dispersión son positivas, indicando que, existe una correlación positiva o directa en las variables.

En el primer gráfico de dispersión, para el paso promedio del mes  $\mu_{h_0} = 14$  se puede observar que los datos están menos dispersos que en el segundo gráfico de dispersión para el paso en que se alcanzó la máxima correlación ese día  $h_0 = 0$ ; es decir, existe una mejor correlación en las variables para el paso promedio del mes  $\mu_{h_0} = 14$ .

En el comportamiento de los  $h_o$  correspondiente a la correlación cruzada en las variables viento y precipitación para los meses de octubre y septiembre del año 2008; en su mayoría, se observó que las máximas correlaciones diarias se alcanzaron en un paso  $\mu_{h_o} = 3$  y  $\mu_{h_o} = 2$ , respectivamente.

En el comportamiento de los  $h_o$  correspondiente a la correlación cruzada en las variables temperatura y radiación para los meses de noviembre y diciembre del año 2007 y marzo, abril y octubre del año 2008; en su mayoría, se observó que las máximas correlaciones diarias se alcanzaron en un paso  $h_o = 48$ , respectivamente. Por otra parte, para los meses de enero, febrero, junio y septiembre del año 2008 en el comportamiento de los  $h_o$  se observó en su mayoría que las máximas correlaciones diarias se alcanzaron en un paso  $h_o = 0$ , respectivamente.

En cuanto a la correlación cruzada en las variables temperatura y precipitación para los meses de junio, septiembre y octubre del año 2008 no se pudo determinar un patrón de comportamiento de los valores de  $h_o$ , pues los valores de  $h_o$  en que se alcanzaron las máximas correlaciones diarias fueron muy diferentes entre sí.

Por último, en el patrón de comportamiento de los valores de  $h_o$  correspondientes la correlación cruzada en las variables radiación y precipitación para el mes de octubre del año 2008; en su mayoría, se observó que las máximas correlaciones diarias se alcanzaron en un paso  $\mu_{h_o} = 46$ .

A continuación, la tabla 4.2 muestra los  $\mu_{h_o}$  correspondientes a la estación Tacarigua:

Correlaciones	$\mu_h$								
	Nov	Dic	Ene	Mar	Abr	May	Jun	Jul	Ago
Vie-Tem	NaN	19	12	12	11	24	16	27	31
Rad-Vie	NaN	30	26	NaN	NaN	NaN	NaN	28	22
Vie-Pre	NaN	15	NaN	NaN	10	21	33	27	24
Tem-Rad	9	44	48	NaN	NaN	NaN	NaN	23	41
Tem-Pre	NaN	23	NaN	NaN	17	29	21	14	34
Rad-Pre	NaN	NaN	NaN	NaN	NaN	NaN	NaN	31	33

Tabla 4.2: Medias de los  $h$  óptimos correspondientes a cada correlación cruzada y a cada mes del periodo 2007-2008. (Est. Tac)

En la tabla 4.2, los valores *NaN* indican que para ese mes la correlación cruzada en las variables consideradas no fue posible calcularla, debido a la falta de datos para el mes de dicha combinación. Además, es importante recalcar que al igual que ocurrió con los datos de la estación Sucre, para algunas correlaciones cruzadas en las variables consideradas, no se puede determinar un patrón de comportamiento de los valores

$h_o$  en algunos meses específicos; ya que, dichos meses presentan una escasa cantidad de días para los cuales se calculó la correlación cruzada. Dichos valores de  $\mu_{h_o}$  corresponden a:

- La correlación cruzada en las variables viento y temperatura, para el mes de mayo del año 2008.
- La correlación cruzada en las variables viento y precipitación, para los meses de diciembre del año 2007 y mayo del año 2008.
- La correlación cruzada en las variables temperatura y radiación, para el mes de noviembre del año 2007.
- La correlación cruzada en las variables temperatura y precipitación, para los meses de diciembre del año 2007 y mayo del año 2008.

En el patrón de comportamiento de los valores de  $h_o$  correspondientes a la correlación cruzada en las variables viento y temperatura para los meses de enero, marzo, abril, junio, julio y agosto del año 2008; en su mayoría, se observó que las máximas correlaciones diarias se alcanzaron en un paso  $\mu_{h_o} = 2$ ,  $\mu_{h_o} = 3$ ,  $\mu_{h_o} = 1$ ,  $\mu_{h_o} = 2$ ,  $\mu_{h_o} = 0$ ,  $\mu_{h_o} = 46$  y  $\mu_{h_o} = 47$ , respectivamente. Por otra parte, respecto al mes de diciembre del 2007 no se pudo determinar un patrón comportamiento de los valores de  $h_o$ , pues los valores de  $h_o$  en que se alcanzaron las máximas correlaciones diarias fueron muy diferentes entre sí. Para corroborar el análisis hecho, el boxplot 4.5 muestra la distribución de los pasos  $h$ :

---

Distribución de los pasos "h" de la correlación cruzada entre viento y temperatura.

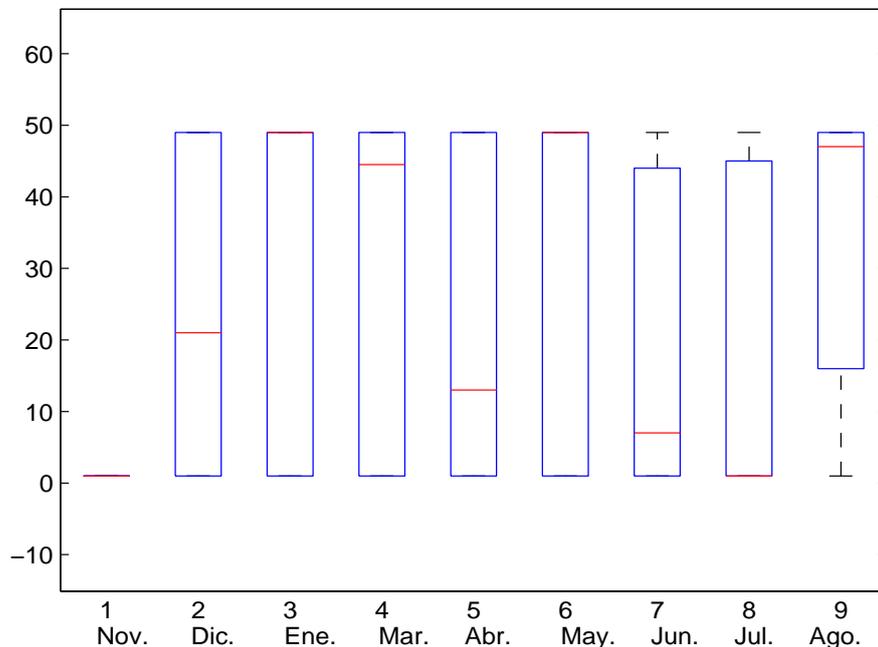


Figura 4.5: Boxplot de distribución de los pasos  $h$  correspondientes a la correlación cruzada en las variables viento y la temperatura. (Est.Tac)

En cuanto a la correlación cruzada en las variables viento y radiación para los meses de diciembre del año 2007 y enero, julio y agosto del año 2008 no se pudo determinar un patrón comportamiento de los valores de  $h_0$ , pues los valores de  $h_0$  en que se alcanzaron las máximas correlaciones diarias fueron muy diferentes entre sí. Para corroborar el análisis hecho, el boxplot 4.6 muestra la distribución de los pasos  $h$ :

Distribución de los pasos "h" de la correlación cruzada entre viento y radiación.

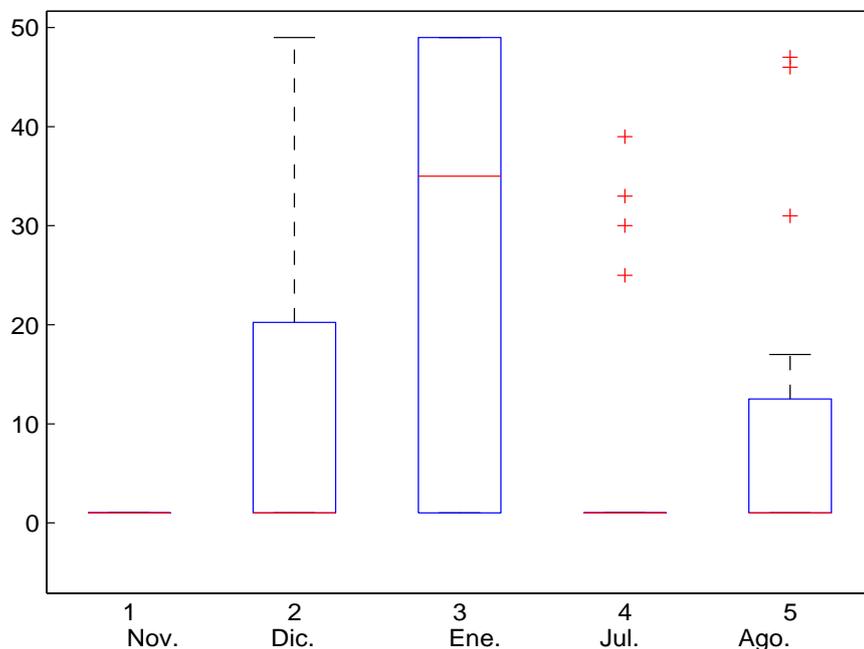


Figura 4.6: Boxplot de distribución de los pasos  $h$  correspondientes a la correlación cruzada entre el viento y la radiación. (Est.Tac)

En el comportamiento de los valores de  $h_0$  correspondientes a la correlación cruzada en las variables viento y precipitación para los meses de abril, junio y agosto del año 2008; en su mayoría, se observó que las máximas correlaciones diarias se alcanzaron en un paso  $\mu_{h_0} = 4$ ,  $\mu_{h_0} = 35$  y  $\mu_{h_0} = 34$ , respectivamente. Por otra parte, respecto al mes de julio del año 2008 no se pudo determinar un patrón comportamiento de los valores de  $h_0$ , pues los valores de  $h_0$  en que se alcanzaron las máximas correlaciones diarias fueron muy diferentes entre sí.

En el patrón de comportamiento de los valores de  $h_0$  correspondientes a la correlación cruzada en las variables temperatura y radiación para los meses de diciembre del año 2007 y enero y agosto del año 2008; en su mayoría, se observó que las máximas correlaciones diarias se alcanzaron en un paso  $\mu_{h_0} = 47$ ,  $\mu_{h_0} = 48$  y  $\mu_{h_0} = 47$ , respectivamente. Por otra parte, respecto al mes de julio del año 2008 no se pudo determinar un patrón comportamiento de los valores de  $h_0$ , pues los valores de  $h_0$  en que se alcanzaron las máximas correlaciones diarias fueron muy diferentes entre sí.

Tomemos como ejemplo el gráfico de dispersión correspondiente al 11 de enero del año 2008:

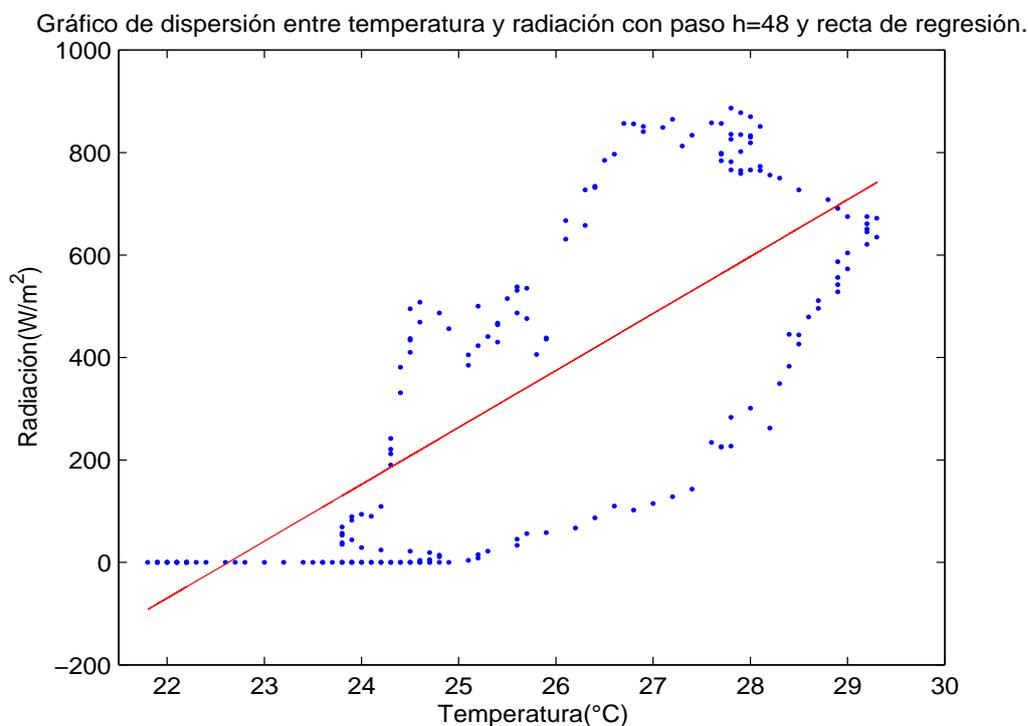


Figura 4.7: Gráfico de dispersión y recta de regresión con paso  $h = 48$  correspondiente a la correlación cruzada en las variables temperatura y la radiación (11/01/2008) (Est.Tac)

En los gráficos de dispersión existe una relación positiva en las variables: conforme aumenta la temperatura, también aumenta la radiación. También, puede observarse que no existe un instante en que la temperatura sea muy alta y la radiación sea muy baja y tampoco un instante en que la radiación sea muy alta y el temperatura sea baja.

La pendiente de la recta de regresión es positiva, es decir, existe una correlación positiva o directa en las variables. Por otra parte, el paso promedio del mes es igual al paso en que se alcanzó la máxima correlación ese día  $\mu_{h_0} = h_0$ .

En el patrón de comportamiento de los valores de  $h_0$  correspondientes a la correlación cruzada en las variables temperatura y precipitación para el mes de agosto del año 2008; en su mayoría, se observó que las máximas correlaciones diarias se alcanzaron en un paso  $\mu_{h_0} = 40$ . Por otra parte, respecto a los meses de abril, junio y julio del año 2008 no se pudo determinar un patrón comportamiento de los valores de  $h_0$ , pues los

valores de  $h_o$  en que se alcanzaron las máximas correlaciones diarias fueron muy diferentes entre sí.

Por último, en el patrón de comportamiento de los valores de  $h_o$  correspondientes la correlación cruzada en las variables radiación y precipitación para el mes de julio del año 2008; en su mayoría, se observó que las máximas correlaciones diarias se alcanzaron en un paso  $\mu_{h_o} = 43$ . En cuanto al mes de agosto del año 2008, no se pudo determinar un patrón comportamiento de los valores de  $h_o$ , pues los valores de  $h_o$  en que se alcanzaron las máximas correlaciones diarias fueron muy diferentes entre sí.

## 4.1. Conclusiones.

Del análisis descriptivo podemos concluir lo siguiente:

- El comportamiento de la temperatura durante el periodo 2007-2008 fue bastante regular; siendo noviembre y diciembre los meses menos cálidos, y enero, febrero, marzo, abril, junio, septiembre y octubre los meses más cálidos.
- El comportamiento de la radiación fue muy irregular durante el periodo estudiado; ya que, se manifestaron muchos casos atípicos. Los cuales, pueden ser causados por la nubosidad presente sobre el lago de Valencia.
- Respecto a la variable precipitación, no fue posible su estudio; ya que, hubo escasos registros de datos durante todo el periodo 2007-2008. Causados por errores en el aparato de medición y pocas precipitaciones sobre el lago de Valencia.
- Mediante los resultados obtenidos en la estación Sucre; se concluye que, el comportamiento de las velocidades del viento fue regular. Las máximas velocidades del viento se alcanzaron en los meses de enero, febrero, marzo y abril.
- Por otra parte, mediante los resultados obtenidos en la estación Tacarigua; se concluye que, el comportamiento de las velocidades del viento fue irregular durante los meses considerados para esa estación; ya que, las máximas y mínimas velocidades del viento fueron muy variadas entre sí.

Del análisis de correlaciones cruzadas entre las diferentes variables estudiadas podemos concluir lo siguiente:

- Conociendo el comportamiento de la temperatura; se concluye que, las velocidades del viento tendrán un comportamiento igual de regular a una diferencia futura aproximada de 45 minutos.
-

- Respecto a la estación Sucre se concluye que, el comportamiento de la radiación es directamente proporcional al comportamiento de las velocidades del viento a una diferencia no mayor de dos horas. Por otra parte, respecto a la estación Tacarigua no se puede concluir nada, ya que no se pudo determinar un patrón de comportamiento para el paso  $h$  correspondiente a las velocidades del viento.
  - Respecto al comportamiento de la temperatura no se puede concluir nada, ya que la radiación presentó un comportamiento irregular durante el periodo 2007-2008.
  - En cuanto al comportamiento de los elementos climáticos en función de la precipitación no se puede concluir nada por la falta de información; hecho que podría ser causado por problemas con el instrumento de medición.
  - Es de mucha utilidad poder llegar a estimar el comportamiento de un elemento climático ya conociendo el comportamiento de otro, para poder llevar a cabo los procesos de saneamiento sobre el lago de Valencia; ya que, conociendo el comportamiento de los elementos climáticos se podría estudiar cómo influyen en la problemática y así considerar diferentes alternativas como soluciones posibles.
-

# Bibliografía

- [1] PETHER J. BROCKWELL y RICHARD A. DAVIS, *Time Series: Theory and Methods*, segunda edición, Springer, Estados Unidos de América: New York, 1991.
- [2] FUNDACIÓN ESPAÑOLA PARA LA CIENCIA Y LA TECNOLOGÍA; *Meteorología y Climatología*, FECYT (Fundación Española para la Ciencia y la Tecnología), Europa, España, 2004.
- [3] JAUME EDITOR ARNAU I GRAS, *Disen os de series temporales : te ´cnicas de ana ´lisis.*, Universitat de Barcelona, Europa: Barcelona, 2001.
- [4] SOLANAS, A., SALAFRANCA, L., FAUQUET, J. y NÚÑEZ, M. I., *Estadística descriptiva en Ciencias del Comportamiento*, Thompson, Europa: Madrid, 2005.
- [5] BOX, G.E.P., JENKINS, G.M. y REINSEL, G.C., *Time Series-Forecasting and control*, tercera edición, Prentice Hall, 1994.
- [6] REINSEL, G.C., *Elements of Multivariate Time Series Analysis*, segunda edición, Springer, Estados Unidos de América: New York, 1997.
- [7] SHUMWAY, R.H. y STOFFER, D.S., *Time Series Analysis and its Applications. With R. Examples*, segunda edición, Springer, Estados Unidos de América: New York, 2006.
- [8] EDILBERTO GUEVARA PÉREZ, JORGE E. GUEVARA BELLO y ESTHER GARCÍA, *Régimen climático y patrón espacial de las lluvias en la cuenca del Lago de Valencia, Venezuela*, Revista de Climatología, Venezuela: Edo. Carabobo, 2008.

- 
- [9] INSTITUTO NACIONAL DE METEOROLOGÍA E HIDROLOGÍA, *Boletín Climatológico de la República Bolivariana de Venezuela. Febrero 2003.*, Ministerio del Poder Popular para el ambiente, Venezuela: D.F., 2003.
- [10] BOLAND, J., *Time-series analysis of climatic variables.*, Fuel and Energy Abstracts., 1996.
- [11] GLUHOVSKY, A y AGEE, E., *On the analysis of atmospheric and climatic time series. Journal of Applied Meteorology and Climatology.*, 6(7), 1125-1129., 2007.
-