

INSTITUTO POLITÉCNICO NACIONAL

Escuela Superior de Física y Matemáticas



*Análisis estadístico para
mediciones obtenidas con un
receptor de GPS civil*

T E S I S

que para obtener el grado de

Licenciado en Física y Matemáticas

presenta:

Glenda Patricia Guevara Hernández

Asesor de Tesis Externo:

Dr. Miguel Nakamura Savoy

Asesor de Tesis Interno:

M. en C. Emigdio Salazar Cordero

México, Distrito Federal, Mayo de 2007.

Análisis estadístico para mediciones obtenidas con un receptor de GPS civil

por

Glenda Patricia Guevara Hernández

TESIS

Para obtener el grado de

Licenciado en Física y Matemáticas

Escuela Superior de Física y Matemáticas

Instituto Politécnico Nacional

México, Distrito Federal

Mayo, 2007

©2007, Glenda Patricia Guevara Hernández

Dedicatoria

A mis padres, Ignacia y Manuel, por el amor y confianza que me han brindado en todo momento.

“Golpea tu corazón, que en él está tu genio” – Stuart Mill

Agradecimientos

A mi asesor, Dr. Miguel Nakamura Savoy, por todo su apoyo e infinita paciencia para conmigo y el desarrollo de este trabajo. Al Centro de Investigación en Matemáticas (CIMAT) por los recursos económicos (beca tesis) que me proporcionó durante algunos meses. Al Dr. Jesús Armando Domínguez Molina por compartir conmigo sus conocimientos sobre el SWP. A los topógrafos que me apoyaron fijando tres puntos, los cuales son una pieza primordial en esta tarea.

A mis padres por el apoyo incondicional que me proporcionan en cada momento. A mi novio, Antonio Eduardo Solís Díaz, por el aliento que me da para seguir adelante. A Esteban Bracamontes Orozco por el ánimo que me brinda. A las familias Bück González y Castro Sandoval por su hospitalidad en mi estancia en la ciudad de Guanajato para llevar al cabo el presente trabajo.

A mis profesores de la Escuela Superior de Física y Matemáticas porque de ellos aprendí mucho, en especial al M. en C. Emigdio Salazar Cordero y a el Lic. Eduardo Virueña Silva.

A Dios por permitirme concluir con éxito esta etapa de mi vida.

Análisis estadístico para mediciones obtenidas con un receptor de GPS civil

por

Glenda Patricia Guevara Hernández

Lic. en Física y Matemáticas, Escuela Superior de Física y

Matemáticas

Instituto Politécnico Nacional, 2007

Resumen

El presente trabajo tiene como propósito fundamental investigar de manera empírica lo que significa el concepto llamado *exactitud* (“accuracy”), proporcionado por un receptor de GPS¹ en el momento de realizar una medición de posición, así como saber qué tan “confiables” son las mediciones que puede proporcionar el mismo instrumento para medir las distancias entre dos puntos. Para ello se realizó un experimento en el cual se obtuvo una serie de mediciones en tres puntos fijos de control utilizando un receptor de GPS de uso civil. Se controlaron diversos factores tales como la visibilidad del cielo y el tiempo de encendido, y se registraron varios datos asociados con las mediciones de posición geográfica.

Concibiendo la variabilidad de datos como un fenómeno probabilístico, se utilizaron algunas técnicas de análisis descriptivo de datos tales como gráficas de dispersión, medidas descriptivas, matrices de correlación, y funciones de distribución

¹GPS (por sus siglas en inglés de Global Positioning System)

empírica. Se exploraron modelos probabilísticos para describir la variabilidad observada, y en particular, para descubrir el significado legítimo del parámetro de *exactitud*. Para esto último, se utilizaron pruebas de normalidad y técnicas de estimación de densidad bivariada utilizando el método de kernel.

De los resultados obtenidos se puede concluir que el valor de la *exactitud* proporcionado por el receptor de GPS sí cumple con ser un cierto cuantil de la distribución de errores medidos en línea recta sobre el plano. Este cuantil coincide con una de las interpretaciones documentadas en instructivos de uso acerca del error, aunque se concluye que no hay distribuciones normales involucradas, y por lo tanto, dicho cuantil no tiene la justificación que convencionalmente se ofrece. Por otra parte, también se concluye que el receptor de GPS es un instrumento muy inexacto en cuanto a la medición de distancias entre dos puntos en la escala en que se desarrolló este trabajo. Además, las mediciones de las distancias entre pares de puntos no poseen distribución normal. Así, el GPS no se comporta como lo hacen los instrumentos de medición típicos, en el sentido de que su error esté conformado por una acumulación de un número grande de pequeñas perturbaciones.

Contenido

Índice de Figuras	ix
1 Introducción	1
1.1 Definición del GPS	2
1.2 Segmentos del GPS	5
1.3 Funcionamiento del GPS	5
1.4 Objetivos	9
2 Desarrollo Experimental	12
2.1 Definición de puntos	12
2.2 Factores que conllevan los objetivos planteados	13
2.3 Toma física de las mediciones	15
2.4 Base de Datos	18
3 Análisis Estadístico	20
3.1 Análisis exploratorio inicial para ciclos	20

Contenido

3.2	Análisis estadístico para 143 ciclos	25
3.2.1	Mediciones de posición	25
3.2.2	Mediciones de distancias entre puntos	45
3.3	Conclusiones	54
Apéndices		58
A Estimadores kernel de densidad		59
A.0.1	Definición del estimador kernel multivariado	59
A.0.2	Estimador kernel univariado ($d = 1$)	65
B Base de Datos		69
Bibliografía		73

Índice de Figuras

1.1	<i>Constelación de satélites del Sistema de Posicionamiento Global (GPS).</i>	6
2.1	<i>Ubicación de los tres puntos experimentales en terreno del Centro de Investigación en Matemáticas (CIMAT).</i>	14
2.2	<i>eTrex Summit. Receptor de GPS que se utiliza para el desarrollo del experimento.</i>	16
3.1	<i>Representación gráfica de ciclos, donde cada triángulo representa un ciclo. Las posiciones medidas durante un mismo ciclo forman los vértices de cada triángulo.</i>	21
3.2	<i>Las 147 mediciones de tres puntos experimentales, obtenidas durante el periodo septiembre - diciembre de 2003. Las coordenadas de cada punto son: (longitud, latitud) en UTM.</i>	22
3.3	<i>Histogramas que describen las mediciones de las distancias entre dos puntos, donde Distancia $P1P2$, es la distancia entre los punto $P1$ y $P2$, Distancia $P2P3$ es la distancia entre los puntos $P2$ y $P3$ y Distancia $P3P1$ es la distancia entre los puntos $P3$ y $P1$, en un mismo ciclo.</i>	24

Índice de Figuras

3.4	<i>Representación gráfica de 143 ciclos, donde cada triángulo representa un ciclo. Las posiciones medidas durante un mismo ciclo forman los vértices de cada triángulo.</i>	26
3.5	<i>Representación de las nubes de mediciones en los tres puntos.</i>	27
3.6	<i>Estadísticas descriptivas de las marginales en los tres puntos.</i>	28
3.7	<i>Correlación que existe entre marginales de manera gráfica y numérica.</i>	29
3.8	<i>Gráficas Cuantil-Cuantil e histogramas de las marginales en el Punto 1.</i>	30
3.9	<i>Gráficas Cuantil-Cuantil e histogramas de las marginales en el Punto 2.</i>	31
3.10	<i>Gráficas Cuantil-Cuantil e histogramas de las marginales en el Punto 3.</i>	32
3.11	<i>Pruebas de normalidad de las marginales en cada punto.</i>	33
3.12	<i>Contornos de las funciones de densidad bivariada estimadas por el método de kernel, donde $h=5.5$ (ancho de banda).</i>	34
3.13	<i>Funciones de distribución empírica de los 143 ciclos.</i>	35
3.14	<i>Estadística descriptivas de las distancias que resultan de la diferencia entre las mediciones de cada punto y su centroide en los 143 ciclos.</i>	35
3.15	<i>Estadísticas descriptivas de radios agrupados de acuerdo a la exactitud que tienen.</i>	36
3.16	<i>Funciones de distribución empírica para los radios en cada punto que tienen exactitud igual a 4 metros.</i>	37

Índice de Figuras

3.17	<i>Funciones de distribución empírica para los radios en cada punto que tienen exactitud igual a 5 metros.</i>	38
3.18	<i>Funciones de distribución empírica para los radios en cada punto que tienen exactitud igual a 6 metros.</i>	39
3.19	<i>Funciones de distribución empírica para los radios en cada punto que tienen exactitud igual a 7–9 metros.</i>	40
3.20	<i>Pruebas de normalidad de las distancias que resultan de la diferencia entre las mediciones de cada punto y su centroide.</i>	41
3.21	<i>Estadísticas descriptivas de radios agrupados según el orden en que se registraron durante el recorrido de cada ciclo.</i>	42
3.22	<i>Funciones de distribución empírica de los radios del Punto 1, agrupados según el orden en que se registraron las mediciones de este punto durante el recorrido de cada ciclo.</i>	43
3.23	<i>Funciones de distribución empírica de los radios del Punto 2, agrupados según el orden en que se registraron las mediciones de este punto durante el recorrido de cada ciclo.</i>	44
3.24	<i>Funciones de distribución empírica de los radios del Punto 3, agrupados según el orden en que se registraron las mediciones de este punto durante el recorrido de cada ciclo.</i>	45
3.25	<i>Estadísticas descriptivas de las distancias entre pares de puntos, medidas ciclo por ciclo.</i>	46
3.26	<i>Diferencia entre el valor real y el promedio experimental para cada distancia entre dos puntos en términos del error estándar.</i>	47

Índice de Figuras

3.27	<i>Valores y gráficas de la correlación e histogramas de cada una de las distancias entre dos puntos.</i>	48
3.28	<i>Gráficas Cuantil-Cuantil de las mediciones de las distancias entre pares de puntos.</i>	49
3.29	<i>Pruebas de normalidad de las distancias entre pares de puntos. . . .</i>	49
3.30	<i>Funciones de densidad estimadas por el método de kernel, para las mediciones de las distancias entre pares de puntos.</i>	50
3.31	<i>Funciones de distribución empírica para las mediciones de distancias entre pares de puntos.</i>	51
3.32	<i>Funciones de distribución empírica de los errores de la Distancia P1P2 agrupados según el punto inicial de cada ciclo.</i>	52
3.33	<i>Funciones de distribución empírica de los errores de la Distancia P2P3 agrupados según el punto inicial de cada ciclo.</i>	53
3.34	<i>Funciones de distribución empírica de los errores de la Distancia P3P1 agrupados según el punto inicial de cada ciclo.</i>	54
3.35	<i>Estadísticas descriptivas básicas de los errores en las distancias entre pares de puntos.</i>	55
B.1	<i>Muestra de la Base de Datos empleada en este trabajo. La muestra contiene 30 ciclos.</i>	70
B.2	<i>Continuación.</i>	71
B.3	<i>Continuación.</i>	72

Capítulo 1

Introducción

Actualmente el Sistema de Posicionamiento Global (GPS), una tecnología de vanguardia basada en señales de radio vía satélite para identificar un lugar sobre la Tierra, tiene muchas aplicaciones. No obstante la alta tecnología utilizada, las mediciones de una posición están sujetas a ciertos errores producidos por diversos factores. En efecto, en la práctica se observa que dos mediciones obtenidas con un receptor de GPS sobre un mismo punto son diferentes entre sí. Este hecho sugiere que esta situación bien pudiera describirse con modelos estadísticos y probabilísticos, y que la variabilidad observada no es más que una manifestación de un fenómeno aleatorio.

En literatura especializada sobre GPS—aun en algunos instructivos de operación de receptores civiles—hay la admisión implícita de que en cualquier medición existen errores aleatorios, ya que se habla de errores cuadráticos medios y otros conceptos estadísticos o probabilísticos. Sin embargo, no es claro el significado exacto de la medida que se ofrece a manera de cuantificar la *exactitud*. Se habla informalmente de “5 metros de *exactitud*”. Si bien dicha aseveración entrecomillada tiene un origen probabilístico, no se hace explícito si se trata de una distancia típica, un intervalo

Capítulo 1. **Introducción**

probabilístico de confianza, o alguna otra cantidad similar.

En esta tesis se abordará una colección de datos de mediciones de GPS con un enfoque estadístico, con el objetivo primordial de verificar el significado de la frase “5 metros de *exactitud*”. El trabajo involucra un experimento diseñado de manera explícita para obtener mediciones bajo condiciones controladas, para luego analizar los datos con diversas técnicas estadísticas.

En este capítulo se precisará qué es el Sistema de Posicionamiento Global. Así mismo, se da un panorama general del tipo de información que puede proporcionar un receptor de GPS. En las segunda y tercera secciones se habla de los segmentos que conforman al GPS y del funcionamiento del mismo, respectivamente. Finalmente, se exponen con más detalle los objetivos a estudiar en este trabajo y se resumen brevemente las conclusiones que fueron obtenidas.

1.1 **Definición del GPS**

El Sistema de Posicionamiento Global, permite conocer la posición de un punto fijo o móvil situado en cualquier parte de la superficie terrestre, mediante una constelación de satélites. French (1996) realiza la descripción del GPS como sigue:

“El **Sistema de Posicionamiento Global** es una base espacial de navegación y un sistema de posicionamiento que fue diseñado por el ejército de Estados Unidos para permitir a un soldado o a un grupo de soldados determinar de manera autónoma su posición dentro de 10 ó 20 metros de la posición real. El concepto de autonomía fue importante en el sentido de que era necesario diseñar un sistema que le permitiera al soldado ser capaz de determinar dónde estaba sin tener que utilizar cualquier otro radio para las transmisiones. En otras palabras, con un

Capítulo 1. **Introducción**

solo receptor de dirección o sentido único cuyo uso no pudiera ser descubierto por potencias hostiles”.

Como ya se mencionó, el GPS en sus inicios fue de uso exclusivo para la milicia de Estados Unidos, sin embargo, para el año de 1984 se expandió el uso del GPS a los civiles por orden del Presidente Ronald Reagan. Esta orden como consecuencia de que un caza soviético destruyó a un avión en pleno vuelo de una aerolínea coreana, que pasó de manera accidental a espacio aéreo prohibido. Lo anterior permitió a la aviación y a otros medios de transportación una mejor precisión en sus sistemas de navegación (French, 1996).

En la actualidad el GPS ha evolucionado de tal manera que cuenta con una cobertura mundial, disponible las 24 horas del día. En general un receptor de GPS proporciona la siguiente información bajo cualquier condición climática:

- Posición. La posición y la *exactitud* del punto donde se encuentra el receptor, y cuando se está en movimiento, la dirección de desplazamiento, la velocidad, la distancia recorrida, la altitud y la hora.
- Mapa. Una representación virtual en el cual se pueden localizar waypoints (puntos de referencia), trazado de la ruta que se recorre y la posición.
- Navegación. Muestra el destino, la distancia hasta el mismo, el rumbo teórico, el rumbo real, la velocidad y el tiempo estimado de arribo.

Es importante hacer notar la diferencia entre *exactitud* y precisión, dos conceptos que es común confundir, inclusive por fabricantes y usuarios de instrumentos de medición. Como ejemplo de esta confusión, si el receptor utilizado a lo largo del presente trabajo (eTrex Summit) se configura en idioma inglés, se utiliza la palabra “accuracy” en su pantalla. Pero si se elige el idioma español, se usa la palabra

Capítulo 1. **Introducción**

“precisión”. La precisión en el GPS es el grado o el tamaño de la unidad en que una posición puede medirse y la exactitud son los grados que hay de diferencia entre la posición medida y la posición real. La *exactitud* del GPS usualmente se describe a través de alguna medida estadística de error del sistema (French, 1996). Por ejemplo, en el manual del usuario del receptor GPS que se empleó en este trabajo, con relación a precisión y *exactitud*, éste dice literalmente (Manual del receptor eTrex Summit (ver Figura 2.2):

“Notificación: El GPS es operado por el gobierno de Estados Unidos, el cual es responsable por su exactitud y mantenimiento. El sistema está sujeto a cambios que podrían afectar la *exactitud* y desempeño de todo el equipo del GPS. Aunque el eTrex Summit tiene una precisión electrónica del NAVAID (por sus siglas en inglés de NAVigation AID) puede ser mal usado o mal interpretado y, por consiguiente, volverse inseguro”.

Por otra parte, menciona que una de las funciones del Etrex Summit es proporcionar la *exactitud* de la posición de un punto dado con la especificación de “15 metros RMS (por sus siglas en inglés de Root Mean Square)”, por lo que está sujeta a la degradación de la exactitud a 100 metros 2DRMS¹ bajo el programa de *Disponibilidad Selectiva* que controla el Departamento de Defensa de los Estados Unidos. Uno de los objetivos del presente trabajo es verificar la interpretación de este concepto, incluyendo la noción de DRMS para cuantificarlo.

¹DRMS (por sus siglas en inglés de Distance Root Mean Square) es una medida estadística de la *exactitud* del GPS. El valor RMS de las distancias desde la ubicación verdadera de la posición fija en una colección de medidas. 2DRMS es el radio de un círculo que contiene por lo menos el 95% de las posiciones fijas en una colección dada en cualquier lugar y es el término más común usado del error del GPS (French, 1996).

1.2 Segmentos del GPS

El Sistema de Posicionamiento Global está constituido por tres segmentos: el segmento de **espacio**, el segmento de **control** y el segmento del **usuario**. Los segmentos de espacio y de control son operados por el ejército y administrados por la fuerza aérea de los Estados Unidos. Básicamente, el segmento de control mantiene la integridad de los satélites y los datos que transmiten. El segmento de espacio está compuesto por una constelación de 24 satélites en órbitas aproximadamente a 20,200 km de la Tierra. Por otro lado, el segmento del usuario está compuesto por todas aquellas personas que han adquirido un receptor de GPS (un receptor de GPS permite recibir en la Tierra constantemente información de posición y tiempo en cualquier lugar del planeta, de día o de noche y bajo cualquier condición meteorológica).

Los satélites transmiten señales de radio que permiten la determinación precisa de la posición en cualquier parte del mundo. Cada satélite transmite su señal en dos diferentes frecuencias de radio: 1.57542 GHz y 1.2276 GHz. A su vez, los satélites están distribuidos de manera equidistante en seis órbitas (ver Figura 1.1), desfasadas entre sí en 60° con 55° de inclinación hacia arriba del Ecuador. El periodo orbital de cada satélite es de 12 horas aproximadamente (11 horas y 58 minutos). Esto significa que cada satélite recorre dos veces su órbita en 24 horas.

1.3 Funcionamiento del GPS

El funcionamiento del sistema GPS se basa en el principio de la **trilateración**. Este principio determina la posición del receptor de GPS midiendo las distancias que hay entre éste y tres o más satélites. Por tanto, para calcular la posición de un punto es necesario que el receptor de GPS determine con exactitud la distancia que lo separa de los satélites.

Capítulo 1. *Introducción*

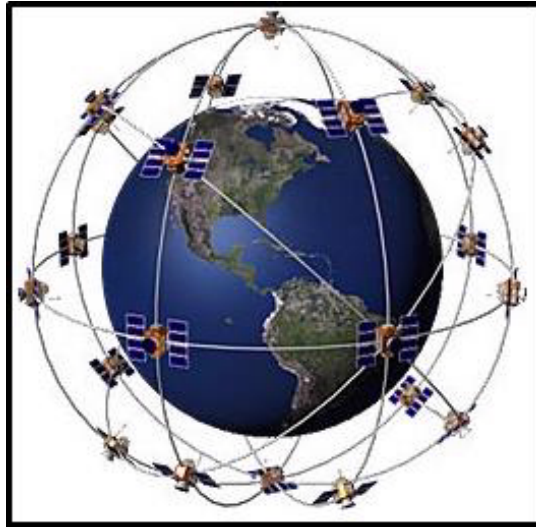


Figura 1.1: *Constelación de satélites del Sistema de Posicionamiento Global (GPS).*

Desde el momento en que el receptor de GPS detecta una señal de radio transmitida por un satélite desde su órbita, se genera una esfera virtual o imaginaria que envuelve al satélite. El propio satélite actúa como centro de la esfera cuya superficie se extenderá hasta el punto o lugar donde se encuentre situada la antena del receptor; por tanto, el radio de la esfera será igual a la distancia que separa al satélite del receptor. A partir de ese instante el receptor GPS medirá las distancias que lo separan de tres satélites o más. Para ello tendrá que calcular el tiempo que demora cada señal en viajar desde los satélites hasta el punto donde éste se encuentra situado. Considerando que las señales de radio viajan a la velocidad de la luz, es decir a 300 mil kilómetros por segundo, las distancias se pueden calcular fácilmente tomando en cuenta que $D = V \times T$, donde D es la distancia, V es la velocidad de la luz y T es el tiempo que se tarda en viajar una señal del satélite al receptor.

Como ya se dijo en el párrafo anterior, cuando el receptor detecta el primer satélite se genera una esfera virtual o imaginaria, cuyo centro es el propio satélite. El radio de la esfera es igual a la distancia que separa al satélite del receptor. Este último

Capítulo 1. **Introducción**

supone entonces que se encuentra situado en un punto cualquiera de la superficie de la esfera, que aún no puede precisar. Al calcular la distancia hasta un segundo satélite, se genera otra esfera virtual. La esfera anteriormente creada se interseca con esta otra y se crea un círculo imaginario. En ese instante ya el receptor reconoce que sólo se puede encontrar situado en cualquiera de los puntos que conforman al círculo. Luego, cuando el receptor calcula la distancia a un tercer satélite se genera una tercera esfera virtual. Esta esfera corta el círculo anterior en dos puntos; uno de los puntos queda ubicado en el espacio y el otro en la superficie de la Tierra. El receptor toma como posición correcta el punto situado en la superficie de la Tierra. Una vez que el receptor ejecuta lo anterior ya puede mostrar en su pantalla los valores correspondientes a las coordenadas de su posición, es decir, la latitud y la longitud. Teniendo información de un cuarto satélite, la cuarta esfera coincidirá con las tres anteriores en un único punto, y es en este momento cuando el receptor puede determinar una posición tridimensional, 3D (latitud, longitud y altitud).

En la práctica se recomienda utilizar el receptor de GPS en un lugar con cielo abierto para tener una mejor recepción de las señales transmitidas por los satélites. Durante el desarrollo del experimento se observó que es posible captar la señal de 4 a 8 satélites.

Existen restricciones en el uso del GPS y por tal motivo se ofrecen dos niveles de navegación y posicionamiento: El *Servicio de Posicionamiento Preciso* (PPS sus siglas en inglés), y el *Servicio de Posicionamiento Estándar* (SPS sus siglas en inglés). El PPS ofrece una *exactitud* muy alta en lo que respecta al posicionamiento, velocidad y tiempo. Está diseñado principalmente para el uso militar y algunos usuarios autorizados, aunque bajo ciertas condiciones puede ser empleado por aquellas personas que tengan el equipo especializado. El SPS ofrece una línea base de *exactitud* que es mucho menor a la proporcionada por el PPS, sin embargo, tiene la ventaja de estar disponible para los usuarios civiles en todo el mundo.

Capítulo 1. **Introducción**

En general, la falta de *exactitud* en las mediciones de un receptor de GPS (en particular, de un GPS civil) puede atribuirse a distintas fuentes de error, por ejemplo: errores humanos, errores técnicos, o por desajustes. Algunas de esas fuentes de error son provocadas, como el ruido en el receptor (dicho término se refiere a la degradación de las señales que recibe un receptor de GPS), un claro ejemplo de las consecuencias debidas a lo que se le denomina como *Disponibilidad Selectiva*.

La *Disponibilidad Selectiva* es una de las fuentes más importantes de error en el GPS, produciendo alrededor de 70 metros de diferencia en la posición con respecto a la posición original. Dicha falta de *exactitud* en el GPS es intencional, ya que es degradada por el Departamento de Defensa estadounidense. El propósito intencionado de hacer esto es limitar la *exactitud* de los GPS civiles a un 95% de probabilidad en 100 metros o menos. Por tanto, el objetivo principal de este trabajo es investigar el significado exacto de la frase anterior, la cual estaría relacionada con la medida de *exactitud* que un receptor de GPS típicamente proporciona. Las condiciones de error en un receptor de GPS se pueden presentar de diversas formas.

La unidad básica de error en un receptor es DRMS ². El valor DRMS representa la distancia dentro de la cual se encontraría el 68% de un grupo de posiciones. Esto, quizás, es mejor pensarlo como un 68% de probabilidad de que cualquier posición arbitraria resulte estar dentro de los límites definidos por el valor DRMS. Por consiguiente se tiene que hay un 32% de oportunidad de que una posición resulte más lejana que el valor RMS. Por otro lado, 2DRMS, significa que hay un 95% de probabilidad que una medición arbitraria resulte estar dentro de los límites establecidos por la distancia error RMS. Así, se alega que el DRMS es simplemente la desviación estándar de una distribución normal.

Existen otras formas en las que se puede presentar el valor del error a las cuales el usuario debe estar atento. El CEP (por sus siglas en inglés de Circular Error

²DRMS (por sus siglas en inglés de Distance Root Mean Square)

Capítulo 1. **Introducción**

Probable), representa el 50% de probabilidad de que una posición arbitraria resulte estar dentro de la cantidad numérica dada en 2D en el plano horizontal. De nuevo, por consiguiente se obtiene que hay un 50% de oportunidad de que una medición salga de los límites de la cantidad numérica dada. El SEP (por sus siglas en inglés de Spherical Error Probable) es similar excepto que hay un 50% de probabilidad de que esté a una distancia esféricamente dada, o en 3D.

El término ampliamente aceptado es 2DRMS, o el 95% de probabilidad. Los folletos de venta frecuentemente citan los valores de CEP o SEP para presentar sus productos en la mejor manera; aun cuando no es deshonesto, pone en dificultad al usuario para entender lo que los números realmente significan (French, 1996).

1.4 **Objetivos**

Cuando se tiene un receptor de GPS en la mano puede observarse que al adquirir varias mediciones en un mismo punto, éstas difieren unas de otras. Por lo anterior, se pensó en hacer un análisis estadístico para investigar y cuantificar las nociones de variabilidad y *exactitud* en dicho aparato.

Los objetivos son:

- Investigar empíricamente el concepto de *exactitud* en un receptor de GPS.
- Observar los posibles efectos que se pudieran presentar en las mediciones (de posición de puntos y distancias entre dos puntos) adquiridas por el receptor de GPS debido al movimiento que sufre dicho aparato al trasladarlo físicamente de un punto a otro.
- Estudiar los efectos causados por cambios en el campo de visión de el receptor de GPS.

Capítulo 1. **Introducción**

- Describir probabilísticamente las mediciones de posición obtenidas con un receptor de GPS.
- Analizar el comportamiento probabilístico de mediciones de la distancia entre dos puntos realizadas con un receptor de GPS.

Para abordar estos objetivos, en el presente trabajo se diseñó un experimento físico de toma de mediciones con un receptor de GPS civil, el cual se describe en el siguiente capítulo, y se analizaron los resultados con diversas técnicas estadísticas descriptivas y probabilísticas (Capítulo 3).

Las conclusiones principales que se obtuvieron del análisis estadístico y probabilístico son las siguientes:

- Se observó que el valor llamado *exactitud* proporcionado por el receptor de GPS sí se relaciona con la varianza en las mediciones obtenidas, aunque no de la manera que uno podría esperar si la cuantificación numérica de *exactitud* fuera en efecto el DRMS.
- Las distancias que resultan de la diferencia entre las mediciones de cada punto y su centroide no se distribuyen normalmente. Sin embargo, se cumple el hecho de que dentro de los límites establecidos por dos veces el valor de la *exactitud* que proporciona el receptor de GPS caen aproximadamente el 95% de las mediciones.
- Las mediciones de los puntos (longitud, latitud) obtenidas con el receptor no se distribuyen normalmente pues las marginales de cada punto tienen colas más pesadas que ésta y son ligeramente asimétricas.
- Respecto a las mediciones de las distancias entre pares de puntos se tiene que tampoco se distribuyen normalmente y que las funciones de densidad que

Capítulo 1. **Introducción**

describen son ligeramente asimétricas. Asimismo se observó que los valores de las mediciones de las distancias entre pares de puntos, en promedio, no son próximos a los valores reales de las distancias entre pares de puntos, sugiriendo entonces que el proceso de medición es sesgado.

Capítulo 2

Desarrollo Experimental

En este capítulo se reseña el desarrollo del experimento que se llevó al cabo para poder dar respuesta a los objetivos planteados en el capítulo anterior. Así, con los datos obtenidos en dicho experimento se construyó una tabla. Con esta base datos se realizará (en el siguiente capítulo) el análisis estadístico y probabilístico.

2.1 Definición de puntos

En el Centro de Investigación en Matemáticas (CIMAT) se fijaron 3 puntos (marcados con ayuda de topógrafos profesionales), considerando las características físicas para que, en cada punto donde se exponga el receptor de GPS, haya distinta visión de las señales de radio transmitidas por los satélites y que además al trasladar el receptor de un punto a otro se cause pérdida momentánea de las señales.

Uno de los puntos, denominado el Punto 1 (P1), se ubicó en la azotea de uno de los edificios del CIMAT (edificio de posgrados). La ubicación de este punto permite al receptor de GPS tener una amplia visión de las señales transmitidas por los satélites. Otro de los puntos, denominado el Punto 2 (P2), se ubicó en las canchas de

Capítulo 2. *Desarrollo Experimental*

basquetbol (en un extremo de la misma). Por la ubicación de este punto, el receptor de GPS tiene una visibilidad media de las señales de radio debido a la interferencia de señales por parte del edificio de la Facultad de Matemáticas (FAMAT) que se encuentra a un costado de la cancha. Por último, un tercer punto, denominado Punto 3 (P3), se ubicó a un lado de la caseta de vigilancia sur del CIMAT. En este punto la visibilidad que tiene el receptor de GPS de las señales de radio es mínima pues, prácticamente, la caseta obstruye la mitad del hemisferio de satélites.

Por otra parte, cabe mencionar que los topógrafos también midieron y proporcionaron la distancia real que existe entre cada par de puntos obteniendo los siguientes resultados (ver la Figura 2.1):

- La distancia que hay del Punto 1 al Punto 2 es de 35.431 metros.
- La distancia que hay del Punto 2 al Punto 3 es de 70.947 metros.
- La distancia que hay del Punto 3 al Punto 1 es de 62.848 metros.

2.2 Factores que conllevan los objetivos planteados

Por la forma en que se realizó el experimento, directa o indirectamente se controlaron factores que pudieran tener efecto sobre la calidad de mediciones. Estos factores son:

- Fecha. Se anotó la fecha en que se registró cada una de las mediciones obtenidas con el receptor de GPS para poder identificar aquellos posibles eventos que se presentarán y pudiesen provocar alteraciones en los datos proporcionados por el mismo.

Capítulo 2. *Desarrollo Experimental*

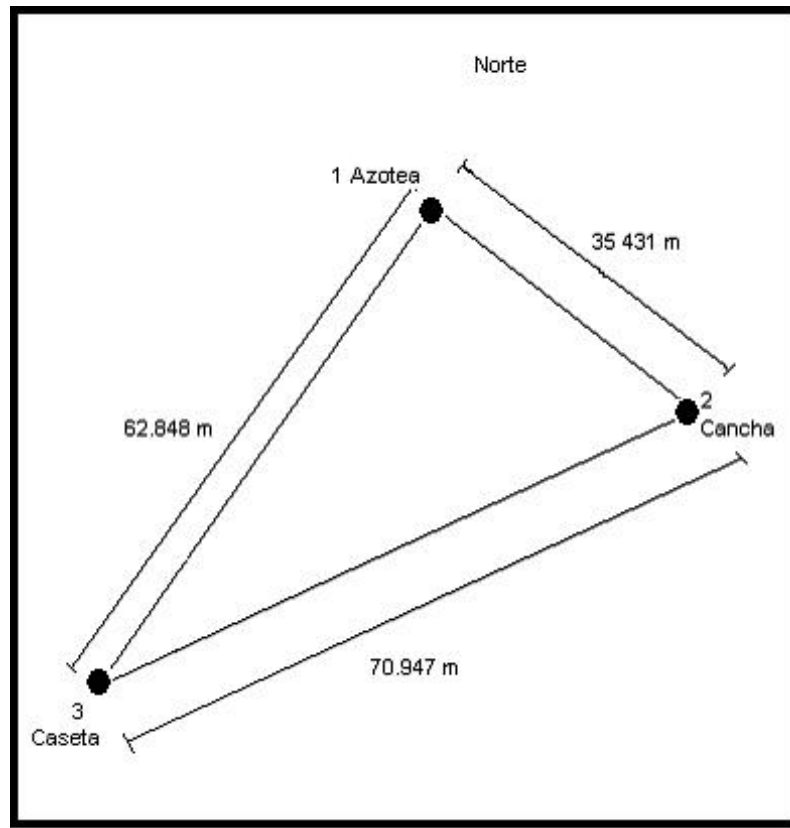


Figura 2.1: *Ubicación de los tres puntos experimentales en terreno del Centro de Investigación en Matemáticas (CIMAT).*

- Hora. Los satélites cambian de posición debido a que recorren su órbita en aproximadamente 12 horas, por lo que se consideró sustancial registrar la hora en que se toma una medición, pues el receptor de GPS recibe la señal de distintos satélites en diferentes tiempos. En un mismo día se registraron mediciones a distintas horas.
- Visibilidad. Es importante considerar la visibilidad que tiene el receptor de GPS de las señales transmitidas por los satélites, pues el valor de la *exactitud* que proporciona el receptor depende del número de satélites de los cuales esté captando su señal y de la intensidad de las mismas. En cada punto donde

Capítulo 2. *Desarrollo Experimental*

se expuso al receptor se registró el dato llamado *exactitud* que proporciona el mismo.

- Tiempo de encendido. Se refiere al tiempo transcurrido entre la primera sintonización de satélites y el tiempo que transcurre antes de tomar la medición en un punto arbitrario. Al estimar este factor se desea analizar si hay cambios en la información que proporciona el receptor de GPS por el hecho de permanecer más tiempo encendido.
- Orden de medición. Se registró el punto en donde se comenzaron a obtener las mediciones. Se consideró importante este factor pues en los tres puntos el receptor de GPS tiene distinta visibilidad de los satélites.
- Movimiento. El receptor de GPS se transportó (encendido), de un punto a otro, a velocidad constante. Al considerar este factor se desea examinar si hay diferencia en la información que proporcionó el receptor de GPS al estar en reposo y la información que proporcionó inmediatamente después de estar en movimiento.
- Pérdida de señal. Se trasladó al receptor de GPS de un punto a otro y cada vez que se realizaba este movimiento el receptor perdía la señal de los satélites por algunos segundos. Se consideró importante observar si existe alguna modificación en la información que proporcionó el receptor de GPS al obtener la medición de un segundo punto.

2.3 Toma física de las mediciones

El receptor de GPS que se utilizó para adquirir las mediciones de un punto arbitrario es un navegador personal Garmin eTrex Summit (ver Figura 2.2), el cual se puede portar fácilmente en la mano. El eTrex tiene una función completa de GPS, compás

Capítulo 2. Desarrollo Experimental

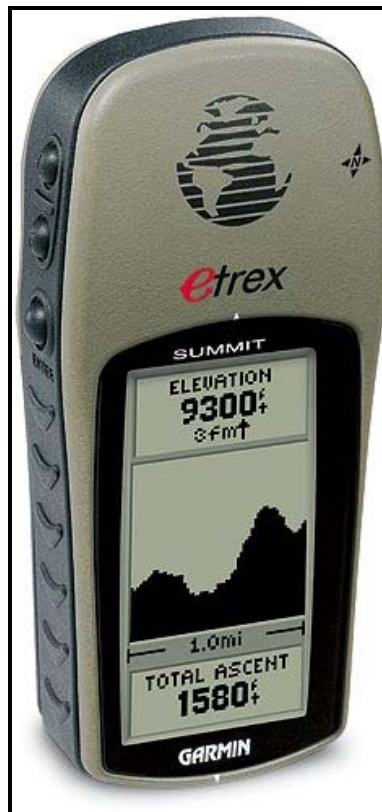


Figura 2.2: *eTrex Summit*. Receptor de GPS que se utiliza para el desarrollo del experimento.

electrónico y altímetro–barómetro. Cuando el Etrex Summit está en movimiento proporciona más datos como: la dirección de desplazamiento, la velocidad o rapidez, la distancia restante para alcanzar el destino, *etc.* Además, el Etrex Summit también puede proporcionar datos como lo son: el lugar donde se está, el lugar donde se estuvo y el lugar donde se irá. Asimismo se puede saber cómo regresar al lugar de dónde se partió.

Para obtener físicamente las mediciones de los puntos que se especifican en la Sección 2.1 se define el concepto de un *ciclo*. Antes de tomar las mediciones de un primer punto se enciende al receptor de GPS y no se apaga hasta que se hayan

Capítulo 2. *Desarrollo Experimental*

tomado las mediciones propias del último punto. El punto donde comienza el ciclo se elige al azar (con un dado) para evitar cualquier tendencia a iniciar el ciclo en algún punto en particular, después se sigue un recorrido en sentido de las manecillas del reloj recorriendo los otros dos puntos. El ciclo esta completo cuando se han tomado las mediciones de los tres puntos que se fijaron. Cabe mencionar que el receptor de GPS se posiciona físicamente sobre cada uno de los puntos que fueron marcados con ayuda de los topógrafos.

Antes de anotar las coordenadas del primer punto, el receptor de GPS permanece encendido cinco minutos aproximadamente. Enseguida se registra fecha, hora y el número del punto donde se comienza, así como la longitud, la latitud y la *exactitud* del mismo. Después, se translada al receptor al siguiente punto, una vez estando ahí se deja expuesto y fijo por espacio de un minuto y se registra la longitud, la latitud, la *exactitud* y la distancia entre este punto y el anterior. Finalmente, en el último punto se realiza el mismo procedimiento que en el anterior registrando la longitud, la latitud, la *exactitud* y las distancias entre este punto y los otros dos. Así, un ciclo queda conformado por la fecha, la hora, el punto de comienzo, la longitud, la latitud y la *exactitud* de cada punto, así como las distancias que se miden entre los mismos.

En total se registraron las mediciones de 147 ciclos, en el periodo comprendido del 30 de septiembre al 12 de diciembre de 2003, registrando cada día los ciclos entre las 11:00 y las 19:00 horas. Se tomaron en intervalos de una hora aproximadamente tratando de que coincidieran en la misma hora todos días para tener así mayor control sobre los factores que se consideraron al inicio del experimento.

Como ya se mencionó en el Capítulo 1, el receptor de GPS proporciona la longitud y latitud en grados y minutos, por lo cual, para poder graficar puntos sobre un plano y representar apropiadamente las distancias entre pares de puntos, se transformaron los datos de posición a metros bajo una proyección llamada “*Proyección Transversal de Mercator Universal*” (UTM por sus siglas en inglés) cuya característica principal

Capítulo 2. *Desarrollo Experimental*

es ser *conforme*, es decir, que conserva ángulos (localmente), además de que conserva la forma y la escala es constante a lo largo del meridiano central. Por otra parte, físicamente los puntos están próximos entre sí (la máxima distancia que hay entre ellos es menor a 100 metros), por lo que bajo la transformación, visualmente no cambian de posición (Kennedy and Kopp, 2000)¹. La hora también fue transformada para fines de registro, a una escala decimal, debido a que el receptor de GPS la proporciona en sistema sexagesimal.

2.4 Base de Datos

La tabla de datos (Figura B.1, Figura B.2 y Figura B.3) que se presenta en el Apéndice B muestra una porción de los datos recabados, donde cada renglón es un ciclo y las columnas están conformadas por la siguiente información:

- Número del ciclo. ----- (Ciclo)
- Número del punto donde comienza el ciclo. ----- (Sorteo)
- Fecha del día en que se registraron los ciclos. ----- (Fecha)
- Hora en que comenzó el ciclo (hora en escala decimal). -----(Hora)
- *Exactitud* del Punto 1 en metros. -----(Exac 1)
- Longitud del Punto 1 en grados decimales. -----(Lon 1)
- Latitud del Punto 1 en grados decimales. -----(Lat 1)

¹Existen calculadoras en línea para realizar la conversión de la longitud y la latitud, que son proporcionadas por el receptor de GPS en grados y minutos, a metros bajo la proyección UTM. Un ejemplo, en la página: http://www.posc.org/Epicentre.2_2/DataModel/ExamplesofUsage/cs34utm.html.

Capítulo 2. *Desarrollo Experimental*

- *Exactitud* del Punto 2 en metros. - - - - - (Exac 2)
- Longitud del Punto 2 en grados decimales. - - - - - (Lon 2)
- Latitud del Punto 2 en grados decimales. - - - - - (Lat 2)
- *Exactitud* del Punto 3 en metros. - - - - - (Exac 3)
- Longitud del Punto 3 en grados decimales. - - - - - (Lon 3)
- Latitud del Punto 3 en grados decimales. - - - - - (Lat 3)
- Distancia entre el Punto 1 y Punto 2 en metros. - - - - - (Distancia P1P2)
- Distancia entre el Punto 2 y Punto 3 en metros. - - - - - (Distancia P2P3)
- Distancia entre el Punto 1 y Punto 3 en metros. - - - - - (Distancia P3P1)
- Longitud del Punto 1 en metros bajo la proyección UTM. - - (Lonutm 1)
- Latitud del Punto 1 en metros bajo la proyección UTM. - - (Latutm 1)
- Longitud del Punto 2 en metros bajo la proyección UTM. - - (Lonutm 2)
- Latitud del Punto 2 en metros bajo la proyección UTM. - - (Latutm 2)
- Longitud del Punto 3 en metros bajo la proyección UTM. - - (Lonutm 3)
- Latitud del Punto 3 en metros bajo la proyección UTM. - - (Latutm 3)

Capítulo 3

Análisis Estadístico

En este capítulo se presenta el análisis estadístico y probabilístico que se realiza a los datos experimentales obtenidos con el receptor de GPS civil. En la primera sección, se lleva al cabo un análisis exploratorio de la base de datos. La unidad experimental es el ciclo descrito en el Capítulo 2 (Sección 2.3), y para cada ciclo se analizarán resultados acerca de la posición y de las distancias entre pares de puntos. En la segunda sección se desarrolla un análisis estadístico más concreto para dar respuesta a los objetivos planteados en el Capítulo 1, Sección 1.4. Por último, en la tercera sección se enlistan las conclusiones relevantes de lo observado.

3.1 Análisis exploratorio inicial para ciclos

Los datos que se presentan en la Figura 3.1 son los que conforman a los ciclos. Cada ciclo esta compuesto por las mediciones obtenidas por un receptor de GPS (en particular, por un eTrex Summit) en tres puntos distintos y las mediciones de las distancias entre pares de puntos. El orden en que se obtienen las mediciones es en el sentido de las manecillas de un reloj. En total se obtuvieron mediciones para 147

Capítulo 3. Análisis Estadístico

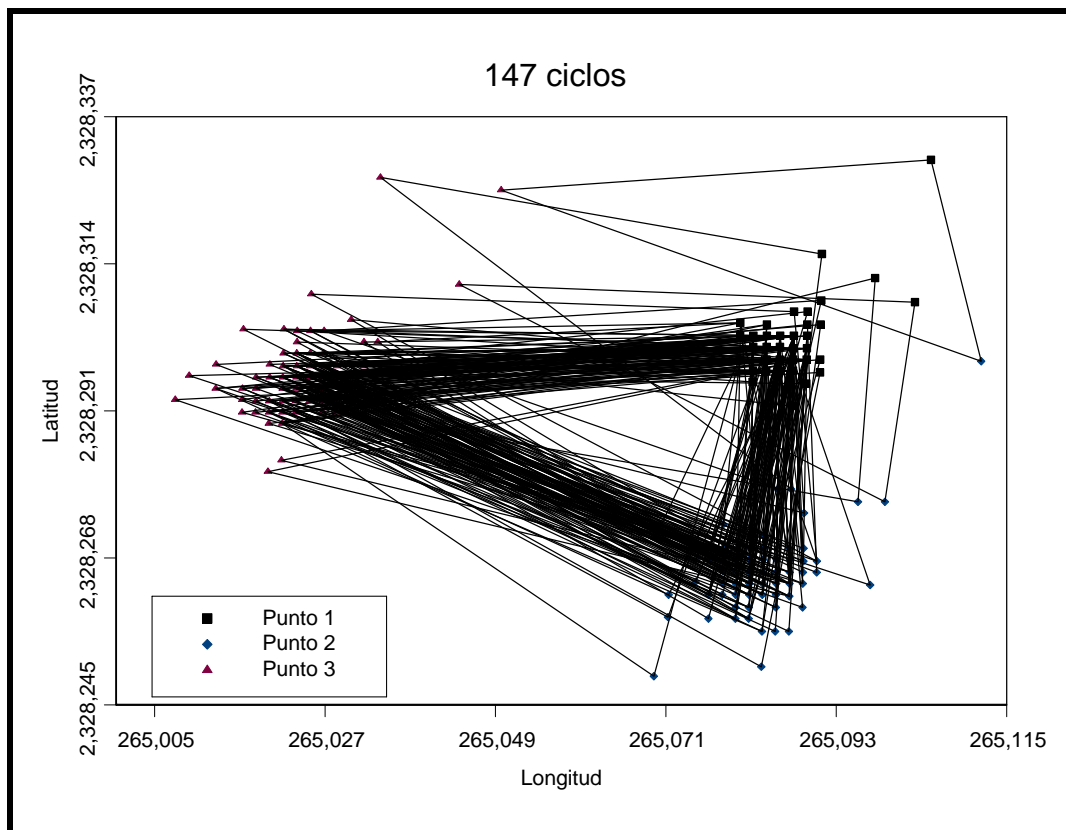


Figura 3.1: Representación gráfica de ciclos, donde cada triángulo representa un ciclo. Las posiciones medidas durante un mismo ciclo forman los vértices de cada triángulo.

ciclos, y el punto donde inicia cada ciclo fue elegido al azar. Las coordenadas de las mediciones de los puntos que conforman a cada ciclo en todas las gráficas que se estudien de aquí en adelante estarán bajo la proyección UTM, donde la unidad de medición está dada en metros. Por convención internacional, esta proyección hace referencia a ciertos meridianos de referencia y al Ecuador, y es por esta razón que los órdenes de magnitud de longitud y latitud se encuentran en los miles de metros.

La Figura 3.1 pone en evidencia que sí hay cierta correlación inducida por el momento en que se mide cada ciclo, pues se observa una especie de translación en los triángulos graficados. En la figura se puede apreciar con mayor claridad

Capítulo 3. *Análisis Estadístico*

la trayectoria que se recorrió en cada ciclo. En su mayoría los ciclos están muy próximos unos de otros, y los que pudieran percibirse como atípicos tienen una marcada tendencia, puesto que se encuentran desplazados primordialmente hacia la parte superior derecha.

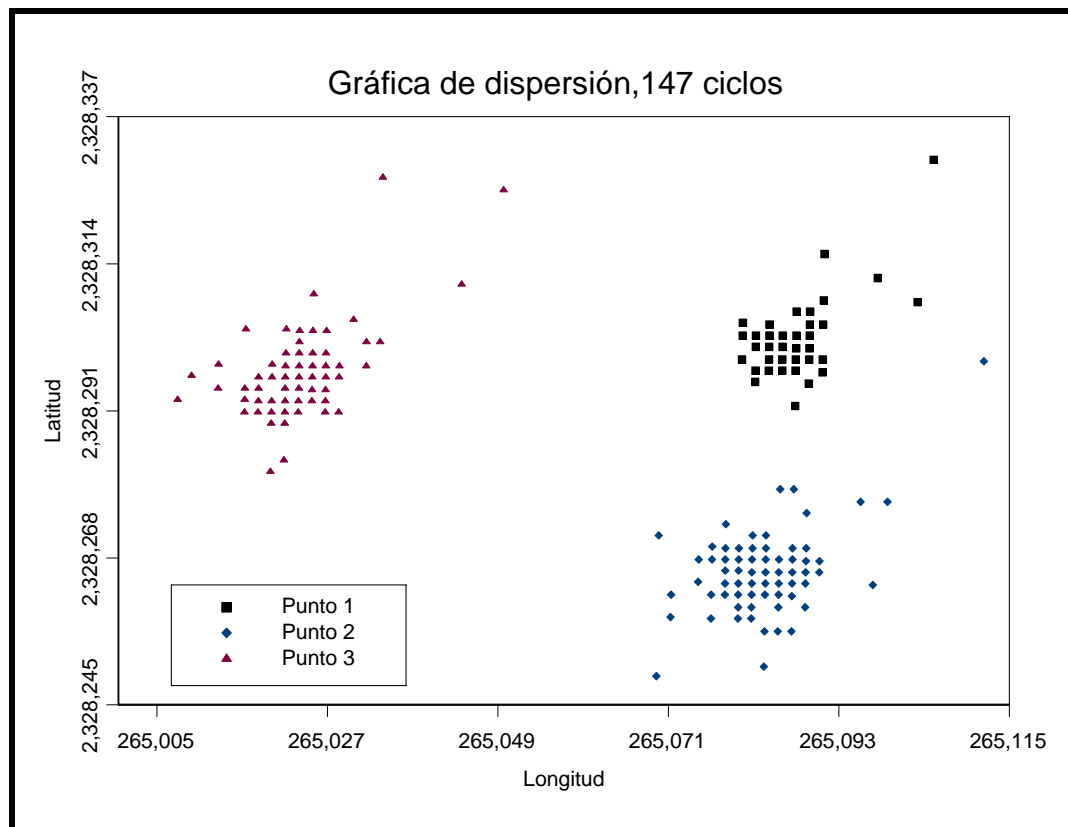


Figura 3.2: Las 147 mediciones de tres puntos experimentales, obtenidas durante el periodo septiembre - diciembre de 2003. Las coordenadas de cada punto son: (longitud, latitud) en UTM.

En la gráfica de dispersión que se muestra en la Figura 3.2 se pueden observar las nubes que forman las mediciones individuales en cada punto. En esta figura nuevamente se percibe, como se mencionó, que algunas de las mediciones en los tres puntos se cargan de manera atípica hacia la esquina noreste de la gráfica. Por otra parte, también se observa que las mediciones en Punto 2 y Punto 3 tienen mayor

Capítulo 3. *Análisis Estadístico*

variabilidad en comparación con las del Punto 1. Esto es de esperarse, puesto que la visibilidad que tiene el receptor de GPS para los satélites en el Punto 1 es mayor que la de los otros dos.

De estos resultados gráficos se observa que las nubes de mediciones en los tres puntos describen distribuciones bivariadas que tienen marginales asimétricas. Los tres pares de distribuciones marginales son asimétricas por la derecha, haciendo notar que el par de distribuciones marginales más asimétrico es el que describe la nube de mediciones del Punto 1, y que a su vez es el que tiene menor varianza en comparación con los otros dos puntos.

En la Figura 3.3 se muestran los histogramas que describen las mediciones de las distancias entre pares de puntos de los ciclos. En esta figura se observa que los histogramas descritos por las mediciones de las Distancia P1P2 y Distancia P3P1 son ligeramente asimétricas hacia la derecha, mientras que el descrito por las mediciones Distancia P2P3 es asimétrico por la izquierda. El histograma de las mediciones de la Distancia P3P1 tiene mayor concentración de mediciones alrededor de la media que los histogramas que describen las mediciones de las otras dos distancias entre pares de puntos. Es interesante señalar que no obstante que en la medición de posiciones de los tres puntos se detectan valores aberrantes o atípicos, no necesariamente se producen valores atípicos en las mediciones de distancias.

Como consecuencia de los primeros análisis gráficos que se han realizado, se aprecia que hubo algún o algunos factores externos que provocaron cierta alteración en algunas de las mediciones. Estudiando un poco sobre los fenómenos naturales que se pudieron presentar durante el periodo en que fueron tomadas las mediciones (30 septiembre–12 diciembre de 2003), cabe señalar que existió un fenómeno de tormenta solar durante los últimos días del mes de octubre de 2003. El periódico *El Universal* del jueves 30 de octubre de 2003 reporta lo siguiente:

Capítulo 3. *Análisis Estadístico*

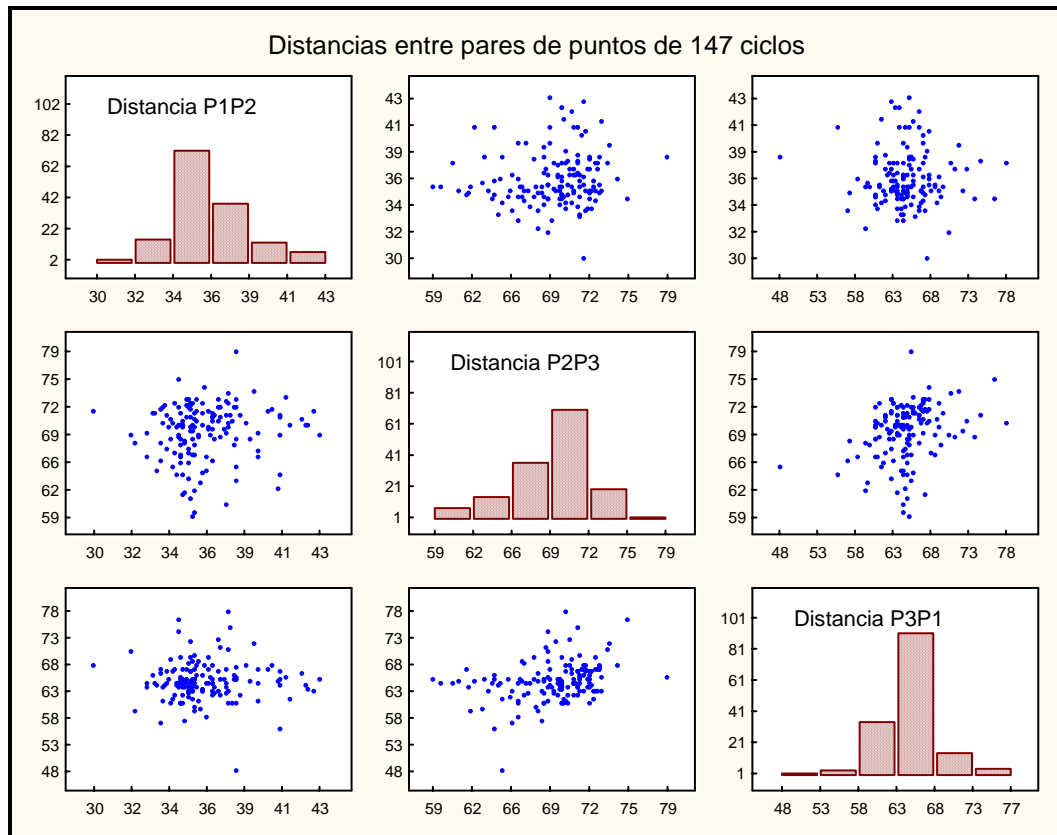


Figura 3.3: *Histogramas que describen las mediciones de las distancias entre dos puntos, donde Distancia P1P2, es la distancia entre los punto P1 y P2, Distancia P2P3 es la distancia entre los puntos P2 y P3 y Distancia P3P1 es la distancia entre los puntos P3 y P1, en un mismo ciclo.*

“Una violenta eyección de plasma solar azotó ayer la Tierra, pero sólo afectó algunas comunicaciones aéreas y obligó a los operadores de redes eléctricas a reducir las transmisiones de energía por precaución.

El centro, que actúa como la agencia oficial de Estados Unidos observadora del clima en el espacio que alerta a las compañías de energía, aerolíneas y de redes de comunicaciones de potenciales amenazas del espacio, emitió su primera alerta hace una semana”.

Capítulo 3. *Análisis Estadístico*

Por otra parte el periódico El Mundo (de España) del mismo día dice:

“El Sol ha desencadenado la segunda tempestad geomagnética en una semana, una de las mayores tormentas solares de las últimas décadas, que puede afectar a las comunicaciones por satélite. Llegó a la Tierra alrededor de las 18:00 de ayer miércoles (18:00 horas en España, 11:00 horas en México).

La tormenta puede causar trastornos en las comunicaciones por vía satélite, la navegación aérea y las redes de distribución de energía eléctrica durante un espacio de tiempo de entre 18 y 24 horas”.

Se considera que este fenómeno alteró la información proporcionada por el receptor de GPS en esos días, por lo que se decidió prescindir de 4 ciclos registrados durante los días 29 y 30 de octubre del 2003, los cuales tienen, en efecto, una explicación como atípicos con base en un fenómeno físico explicado.

3.2 Análisis estadístico para 143 ciclos

A partir de esta sección se trabajará con 143 ciclos, que son los que resultan después de eliminar a los que se consideraron como atípicos. Con éstos se realizará un análisis estadístico más exhaustivo, incluyendo estadística descriptiva y algunas consideraciones de densidades de probabilidad.

3.2.1 Mediciones de posición

La gráfica que se presenta en la Figura 3.4 muestra la trayectoria que sigue cada uno de los ciclos con los que, como ya se mencionó, se trabajará de aquí en adelante. Es

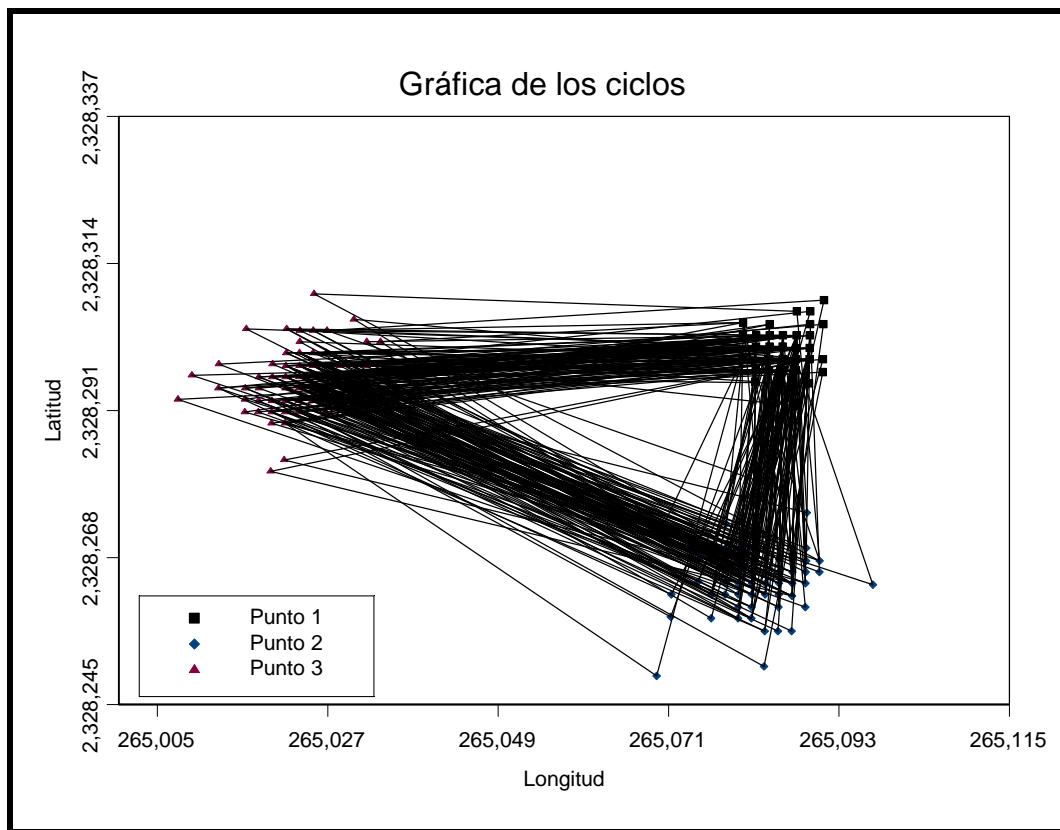


Figura 3.4: *Representación gráfica de 143 ciclos, donde cada triángulo representa un ciclo. Las posiciones medidas durante un mismo ciclo forman los vértices de cada triángulo.*

reiterativa de la Figura 3.1, pero se incluye aquí por completez. En esta gráfica se observa con mayor claridad que las mediciones del Punto 1 están más concentradas alrededor de su media que las mediciones en los otros dos puntos.

En la Figura 3.5 se representan las mediciones de posición en una gráfica de dispersión. Nuevamente, al igual que en la gráfica anterior, se aprecia que las mediciones del Punto 1 están más concentradas entre sí y que las mediciones más dispersas son las del Punto 2, pues tiene algunas mediciones más alejadas de la media. Con la eliminación de las mediciones atípicas, ahora es más claro que la distribución bivariada

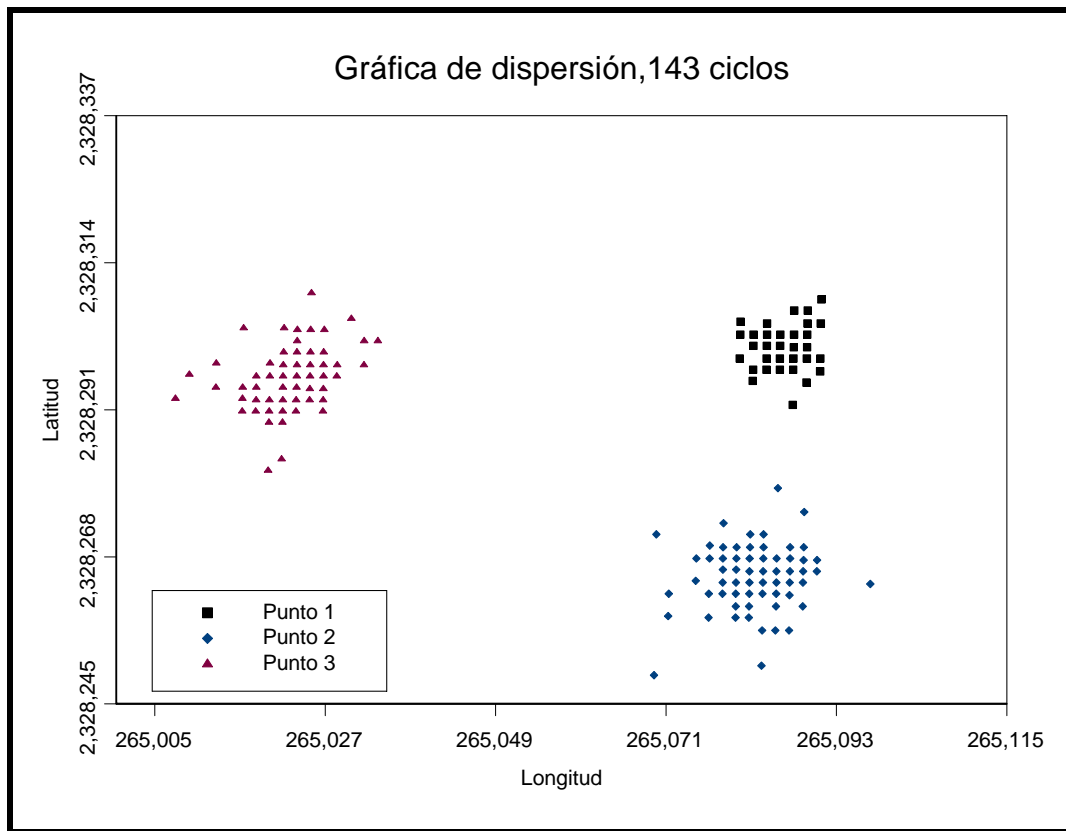


Figura 3.5: *Representación de las nubes de mediciones en los tres puntos.*

de las mediciones es simétrica en algún sentido, al no existir mediciones flagrantes hacia una dirección preferencial.

En la tabla que se exhibe en la Figura 3.6 se presentan algunas estadísticas descriptivas (media, varianza, desviación estándar, asimetría y kurtosis) para las marginales de cada punto. De los resultados se observa que las funciones de distribución bivariada que describen los puntos tienen marginales que son ligeramente asimétricas hacia la izquierda y son más puntiagudas en comparación a una que tiene distribución normal. La ausencia de valores aberrantes cambia el sentido de la asimetría. Por otra parte, la función de distribución bivariada que forman las mediciones del

Capítulo 3. *Análisis Estadístico*

		Media	Varianza	Desv. Est.	Asimetría	Kurtosis
Punto 1	Longitud	265086.2612	5.4825	2.3415	-0.5582	0.1897
	Latitud	2328300.4565	5.5596	2.3579	0.0652	1.2652
Punto 2	Longitud	265082.4248	18.3388	4.2824	-0.1794	1.1209
	Latitud	2328264.9223	16.7479	4.0924	-0.4521	2.0658
Punto 3	Longitud	265022.8164	13.9311	3.7324	-0.8204	3.2082
	Latitud	2328296.1150	15.6429	3.9551	-0.1992	1.7099

Figura 3.6: *Estadísticas descriptivas de las marginales en los tres puntos.*

Punto 1 tiene mayor concentración de mediciones alrededor de su media que las que se forman con las mediciones de los otros dos puntos. De nuevo, esto era de esperarse porque la visibilidad que tiene el receptor de GPS de los satélites en cada punto es distinta, siendo ésta mayor en el Punto 1.

De lo dicho en el párrafo anterior se puede pensar que se tienen ya ciertos indicios de que las marginales en los tres puntos no se distribuyen normalmente. Sin embargo, para abordar de manera más explícita este punto, más adelante se expondrán algunas pruebas de normalidad.

La Figura 3.7 muestra los valores de la correlación que hay entre las marginales de los tres puntos, y de ella se observa que la correlación que existe entre las marginales de cada punto en Punto 1 y Punto 2 es mínima. Sin embargo, la correlación que existe entre las marginales del Punto 3 sí es significativa, lo cual sugiere que las mediciones del receptor de GPS sí tienen un efecto de dependencia. Esto es concordante con la noción, en la Figura 3.4, en la que se observa un efecto de translación completa de los triángulos, hacia arriba o hacia la derecha.

También se realizaron gráficas Cuantil-Cuantil e histogramas para las marginales de cada punto, las cuales se muestran en la Figuras 3.8, 3.9 y 3.10. En las gráficas Cuantil-Cuantil se observa que las mediciones en cada marginal de los tres puntos

Capítulo 3. *Análisis Estadístico*

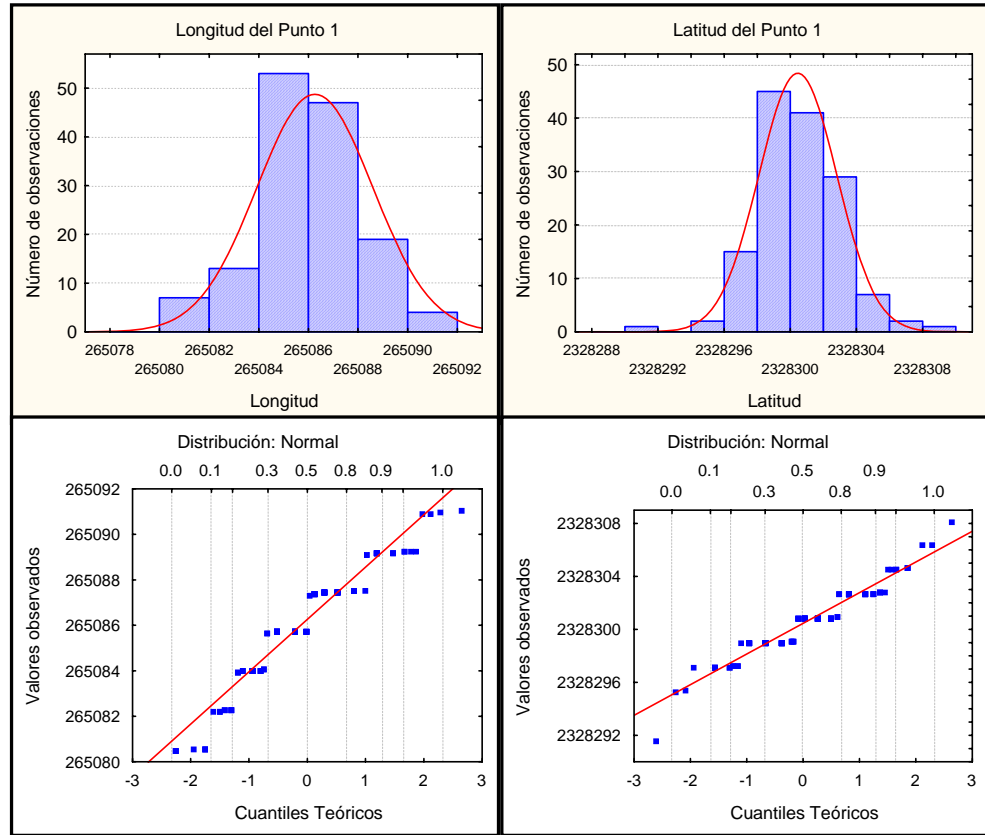


Figura 3.8: *Gráficas Cuantil-Cuantil e histogramas de las marginales en el Punto 1.*

distribuye normalmente. Sin embargo, se puede realizar estimación de densidades bivariadas por el método del kernel (Silverman, 1986, págs. 76–94) para lograr una descripción de ésta. El método del kernel es la suma de “protuberancias” descritas por los datos experimentales, donde la forma de las protuberancias es determinada por la función kernel K y su anchura se establece por el ancho de banda, h . En el caso multivariado en dimensión d , el estimador del kernel basado en n observaciones $\vec{X}_1, \dots, \vec{X}_n$ se define de la manera siguiente:

$$\hat{f}(\vec{x}) = \frac{1}{nh^d} \sum_{i=1}^n K\left\{\frac{1}{h}(\vec{x} - \vec{X}_i)\right\}.$$

Capítulo 3. *Análisis Estadístico*

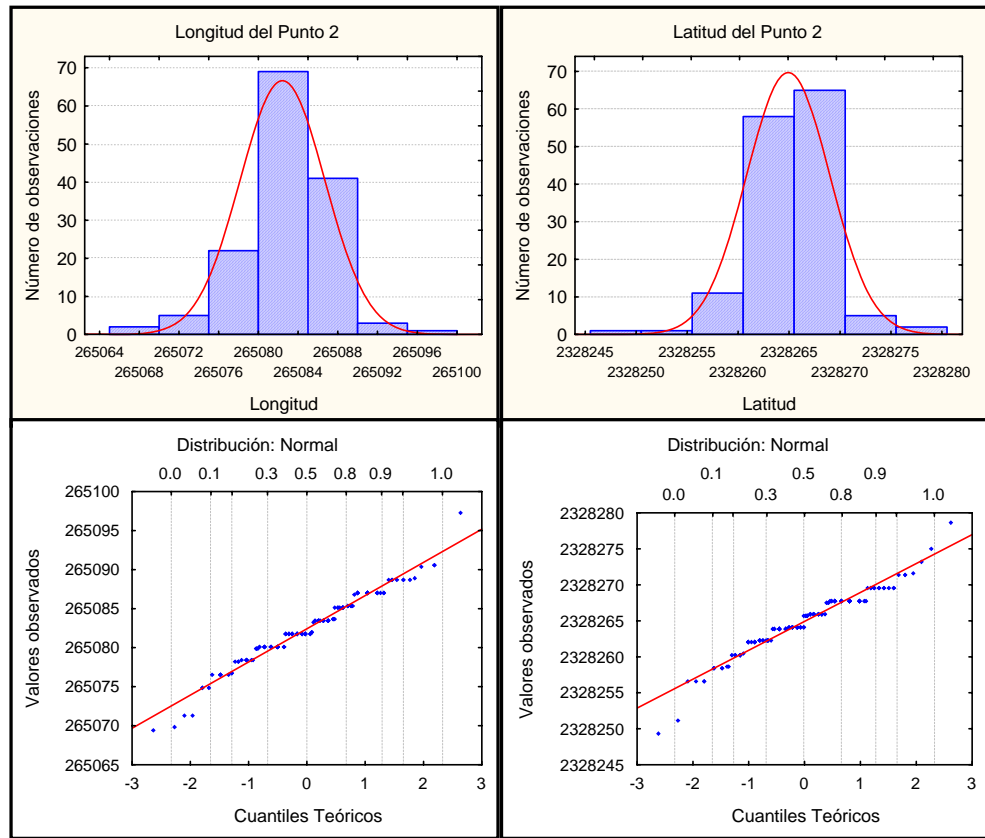


Figura 3.9: Gráficas Cuantil-Cuantil e histogramas de las marginales en el Punto 2.

La función kernel d -dimensional satisface que

$$\int_{R^d} K(\vec{x}) d(\vec{x}) = 1.$$

Por otra parte, frecuentemente la función kernel K es una función de densidad de probabilidad unimodal y radialmente simétrica. Una posible función kernel que cumple con esto es la función de densidad normal multivariada estándar dada por

$$K(\vec{x}) = (2\pi)^{-\frac{d}{2}} \exp\left(-\frac{1}{2}\vec{x}^T\vec{x}\right).$$

Este kernel se utilizó para realizar la estimación de las funciones bivariadas que

Capítulo 3. Análisis Estadístico

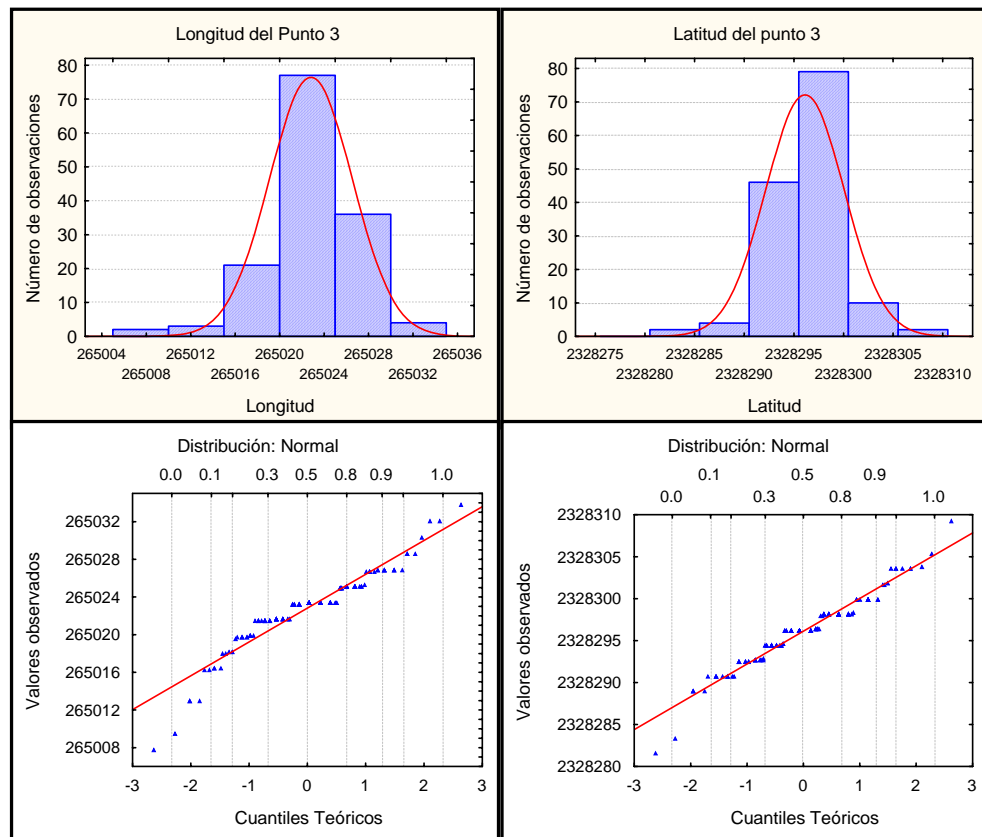


Figura 3.10: Gráficas Cuantil-Cuantil e histogramas de las marginales en el Punto 3.

describen las mediciones de posición. En la Figura 3.12 se representan los contornos de la función de densidad bivariada estimada punto por punto, con un ancho de banda de 5.5. Los contornos en cada nube se observan casi simétricos y con una forma acampanada. Por la concentración de curvas de nivel, puede verse gráficamente el hecho de menor variabilidad sobre el Punto 1.

Significado de *Exactitud*

Si bien los tres puntos experimentales fueron localizados por topógrafos, ellos no proporcionaron las coordenadas. En este sentido, las posiciones reales de los tres

		Pruebas de Normalidad		
		K-S	Lilliefors	Shapiro-Wilk
Punto 1	Longitud	d=0.17552, p<0.01	p<0.01	W=0.92901, p=0.00000
	Latitud	d=0.15547, p<0.01	p<0.01	W=0.93895, p=0.00001
Punto 2	Longitud	d=0.10053, p<0.15	p<0.01	W=0.97083, p=0.00378
	Latitud	d=0.11747, p<0.05	p<0.01	W=0.95348, p=0.00010
Punto 3	Longitud	d=0.16989, p<0.01	p<0.01	W=0.91445, p=0.00000
	Latitud	d=0.14126, p<0.01	p<0.01	W=0.95972, p=0.00034

Figura 3.11: *Pruebas de normalidad de las marginales en cada punto.*

puntos son desconocidos. Por este motivo, para fines de estudiar los conceptos de exactitud, se considerarán como coordenadas reales de los puntos, las coordenadas medias que se obtuvieron con las 143 observaciones. Estas coordenadas promedio se encuentran en la tabla que se presenta en la Figura 3.6, y para fines de estudiar la exactitud, consideraremos que estas coordenadas corresponden a la realidad. Se abordan a continuación estudios para cuantificar la distancia en línea recta que hay entre una medición arbitraria y el valor real. Esta distancia será denominada *radio*, y constituye la noción principal de error que se tratará de relacionar con la medida de exactitud proporcionada por el receptor GPS.

En la Figura 3.13 están representadas las funciones de distribución empírica de radios para cada uno de los tres puntos. De esta figura, se observa que la distribución de errores (radios) no es la misma en cada uno de los tres puntos. Más aún, el Punto 1 tiene la menor media de las tres. Pero esto tiene una explicación sencilla: en cada punto, la exactitud típica vigente a la hora de hacer una medición, es diferente, en virtud de distintos grados de visibilidad. Sin embargo, en cada punto existen mediciones hechas a diversos grados de exactitud, debido a que la exactitud en un momento dado también depende de cuántos, cuáles, y en qué posiciones se encuentran los satélites en el cielo. Por este motivo, se procederá a agrupar radios conforme al

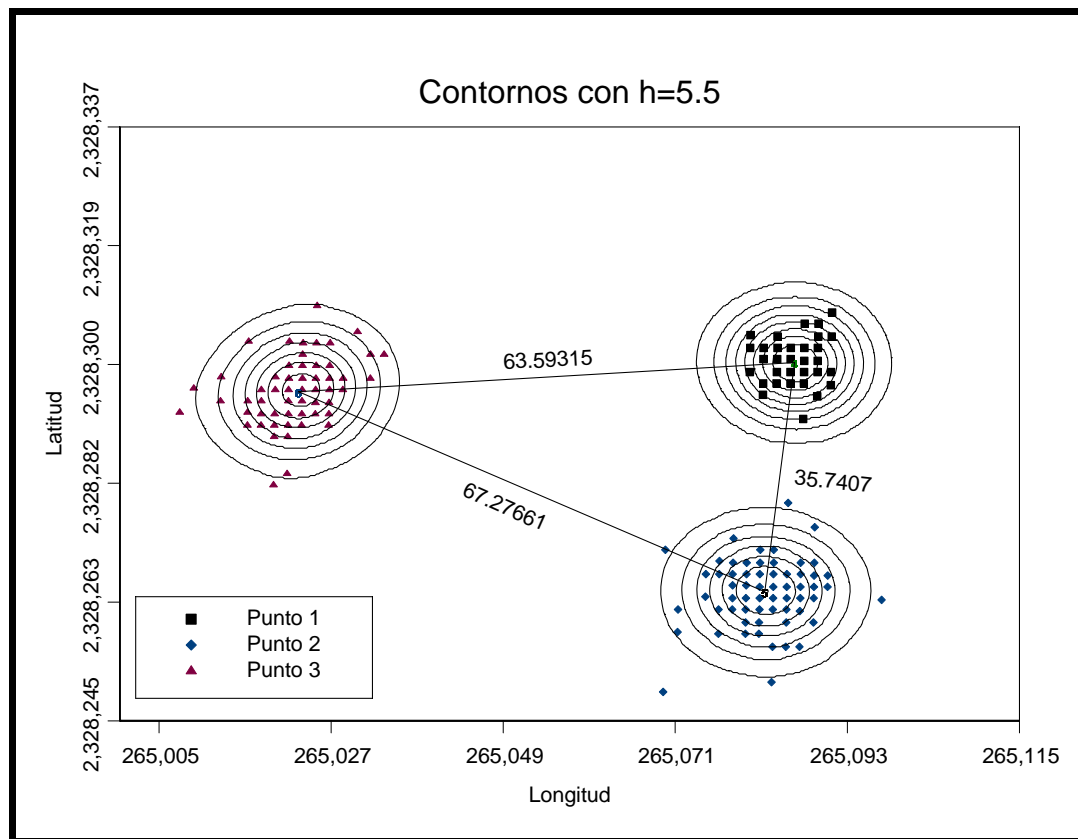


Figura 3.12: *Contornos de las funciones de densidad bivariada estimadas por el método de kernel, donde $h=5.5$ (ancho de banda).*

valor registrado de exactitud en el momento de tomar la medición correspondiente.

En la Figura 3.14 se muestran la media y la desviación estándar de los radios en los tres puntos. En la Figura 3.15 se muestran algunas estadísticas descriptivas de los radios en los tres puntos de acuerdo a su *exactitud*.

En efecto, uno de los objetivos de este trabajo es estudiar el significado del valor de la *exactitud* dado por el receptor de GPS. Así, se presentarán las funciones de distribución empírica de los radios para cada uno de los puntos, en grupos de acuerdo al valor de *exactitud* que se registró en las mediciones. Las Figuras 3.16, 3.17, 3.18 y

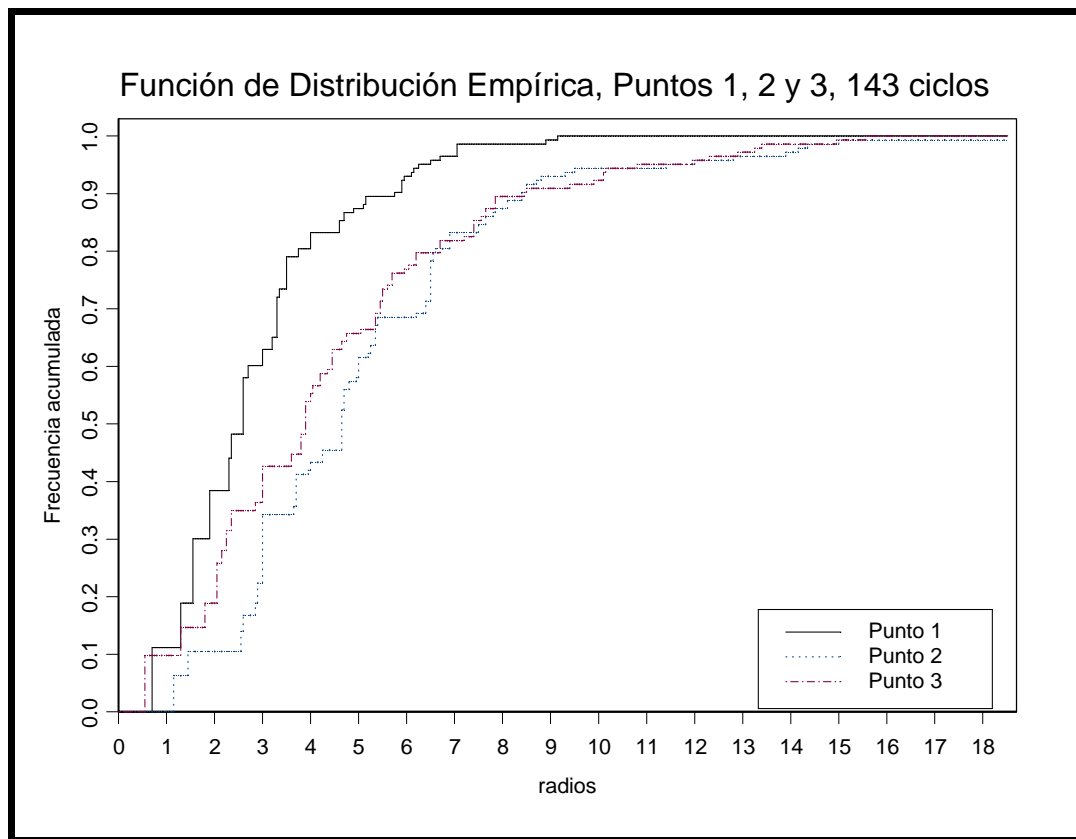


Figura 3.13: *Funciones de distribución empírica de los 143 ciclos.*

3.19 representan las funciones de distribución empírica de los radios en cada punto que tuvieron *exactitud* de 4, 5, 6 y de 7–9 metros, respectivamente.

Radios del	Media	Desv. Est.
Punto 1	2.82069	1.74064
Punto 2	5.03889	3.08503
Punto 3	4.39728	3.17832

Figura 3.14: *Estadística descriptivas de las distancias que resultan de la diferencia entre las mediciones de cada punto y su centroide en los 143 ciclos.*

Capítulo 3. *Análisis Estadístico*

	Radios del	Datos	Media	Desv. Est.
Exactitud 4	Punto 1	47	2.92196	1.83420
	Punto 2	10	6.31210	2.11734
	Punto 3	2	5.19860	0.00000
Exactitud 5	Punto 1	62	2.70351	1.82335
	Punto 2	52	4.80457	2.60398
	Punto 3	20	3.57156	2.94434
Exactitud 6	Punto 1	27	2.76400	1.34055
	Punto 2	46	4.10706	2.26551
	Punto 3	53	4.04871	3.15448
Exactitud 7-9	Punto 1	7	3.22185	1.90425
	Punto 2	33	5.62568	3.81857
	Punto 3	63	4.82674	3.19865

Figura 3.15: *Estadísticas descriptivas de radios agrupados de acuerdo a la exactitud que tienen.*

En estas figuras, se grafican líneas verticales sobre los valores *exactitud* y dos veces *exactitud*, para fines de investigar si hay concordancia entre posibles interpretaciones del término de error y las nociones de DRMS y 2DRMS. En la nota del Capítulo 1, Sección 1.1 se dice que 2DRMS es el radio de un círculo que contiene por lo menos el 95% de mediciones en una colección dada y DRMS es el radio de una circunferencia que contiene el 68% de las mediciones en una colección dada. Si alguno de estos conceptos son legítimos, entonces se esperaría encontrar aproximadamente probabilidad acumulada 0.68 en DRMS, y probabilidad acumulada 0.95 en 2DRMS.

En la Figura 3.16 se observan las funciones de distribución empírica para los radios de los puntos cuya *exactitud* es de 4 metros. En esta gráfica se observa que para un círculo cuyo radio es igual a 4 metros no se cumple el hecho de que haya el 68% de las mediciones que tienen *exactitud* de 4 metros, pues la función de distribución empírica que describen los radios del Punto 2 el porcentaje es mucho menor; para la función de distribución empírica de los radios en el Punto 1 se tiene que el círculo contiene más del del 80% de las mediciones. La función de distribución empírica de

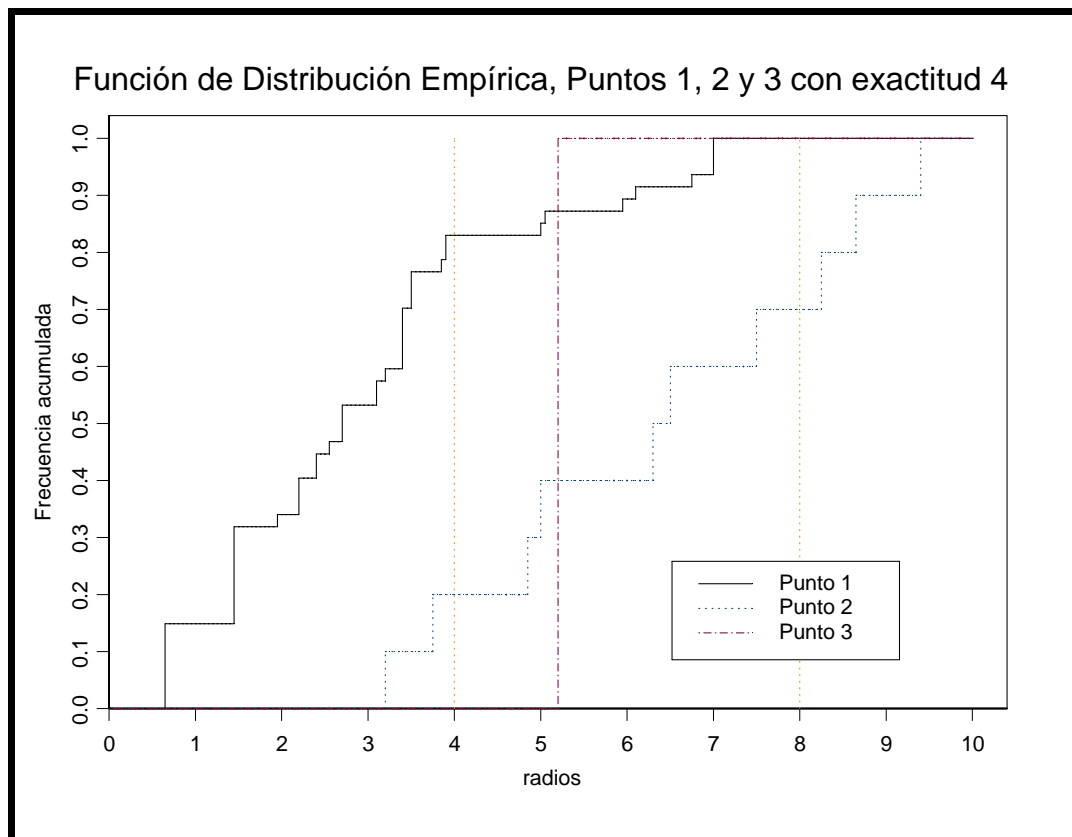


Figura 3.16: *Funciones de distribución empírica para los radios en cada punto que tienen exactitud igual a 4 metros.*

los radios del Punto 3 no se considera en este caso pues sólo resultan dos mediciones. Ahora, si se construye un círculo cuyo radio sea igual a 8 metros (dos veces el valor de la *exactitud*) se observa que para la función de distribución empírica de los radios del Punto 2 no se cumple que el círculo contenga el 95% de las mediciones, pero para la función de distribución empírica de los radios del Punto 1 sí se cumple.

En las gráficas que se exponen en las Figuras 3.17, 3.18 y 3.19 se observa que si se traza el radio con valor 5, 6 y 7 metros, respectivamente, en ninguna de las funciones de distribución empírica de los radios de los puntos se cumple que haya el 68% de las mediciones, pero sí se cumple que si se traza un círculo con radio igual

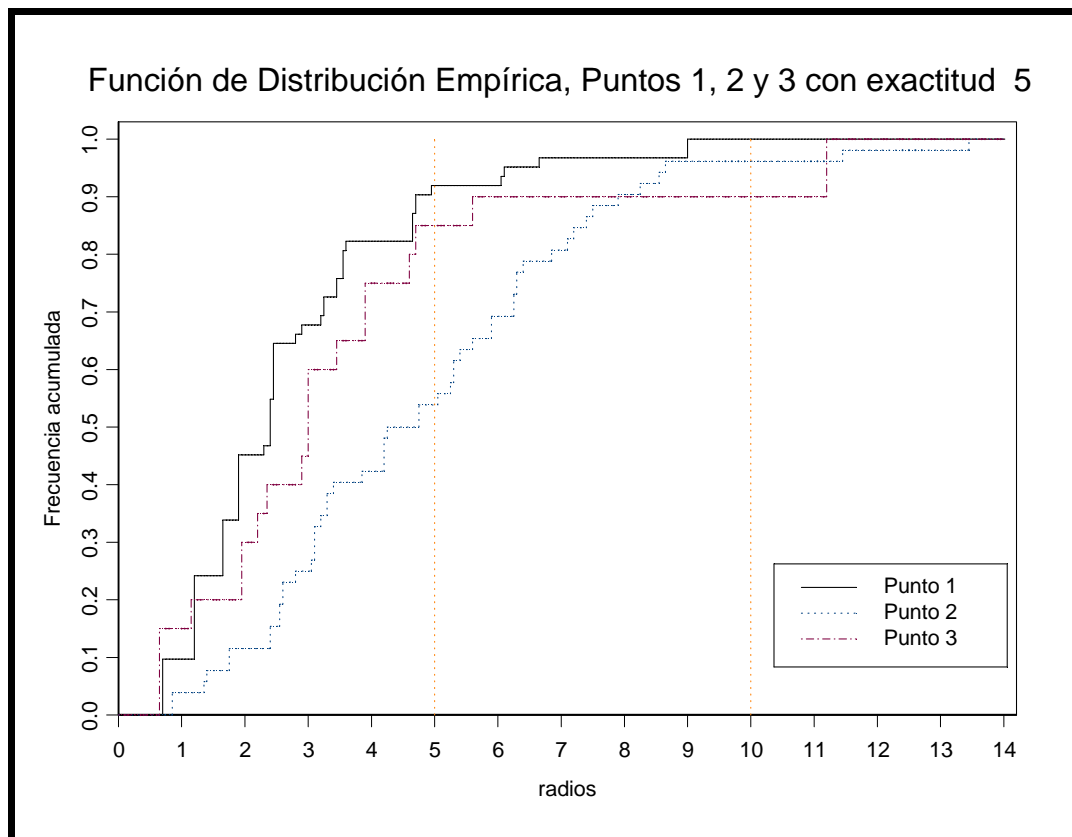


Figura 3.17: *Funciones de distribución empírica para los radios en cada punto que tienen exactitud igual a 5 metros.*

a dos veces la *exactitud* que presentan los radios de los puntos en cada gráfica hay aproximadamente el 95% de las mismas.

Lo anterior sugiere que el valor *exactitud* proporcionado por el GPS en el momento de tomar una medición, no es congruente con ser DRMS. Si tal fuera el caso, se debería haber observado una distribución empírica de aproximadamente 0.68 en ese valor. Sin embargo, el doble de la *exactitud* proporcionada por el GPS sí corresponde aproximadamente con un valor de distribución empírica de 0.95 (salvo el caso Punto 2 en el caso de exactitud 4, aunque debe tomarse en cuenta que en dicho caso sólo resultaron 10 observaciones). Una conclusión general es entonces que el valor

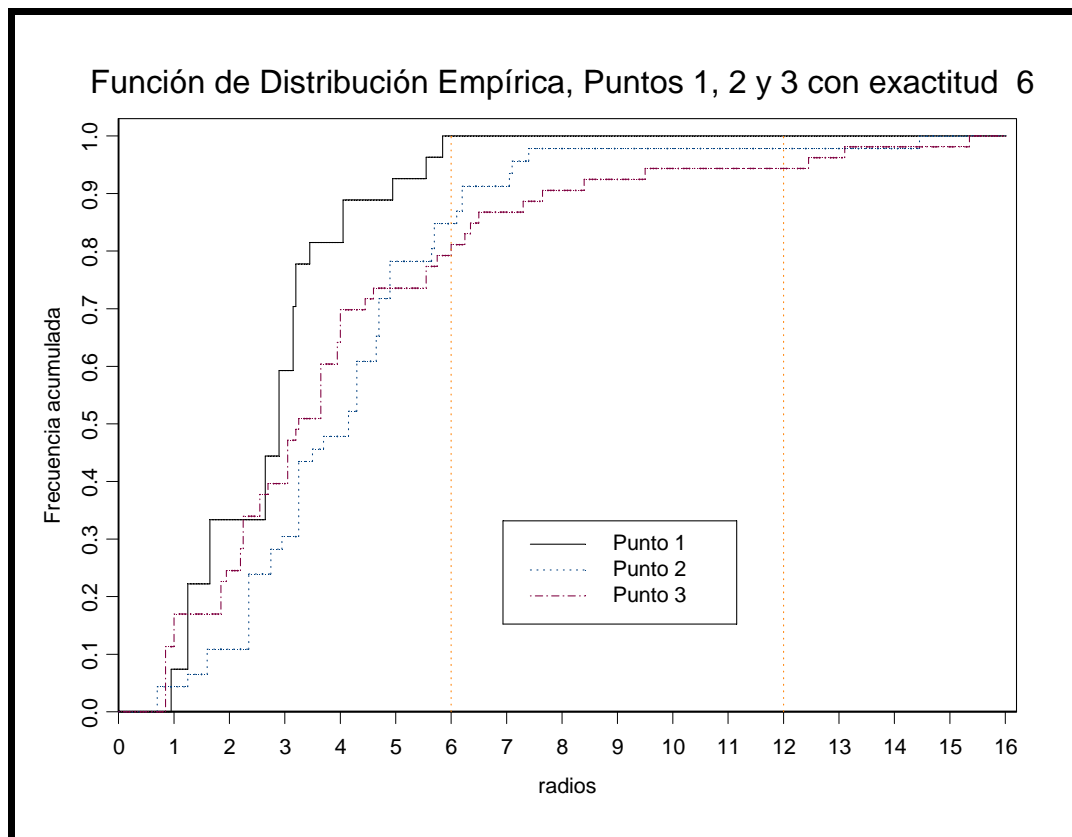


Figura 3.18: *Funciones de distribución empírica para los radios en cada punto que tienen exactitud igual a 6 metros.*

numérico de *exactitud* sí se relaciona con la distribución probabilística de los radios. Pero su interpretación no es DRMS. Esta evidencia sugiere que el valor *exactitud* corresponde a la mitad del cuantil 0.95 de la distribución de errores.

Los valores 0.68 y 0.95 aludidos en las descripciones de DRMS y 2DRMS son fuertemente insinuates de una distribución normal. El hecho de que $2 \times \textit{exactitud}$ sea congruente con el valor 0.95 pero que *exactitud* no sea congruente con 0.68, ya comienza a sugerir que la distribución de radios no es normal, no obstante las citas en la literatura (ver Capítulo 1, Sección 1.3)). Se realizaron algunas pruebas formales de normalidad a las distancias de las mediciones de cada punto a su centroide (ver

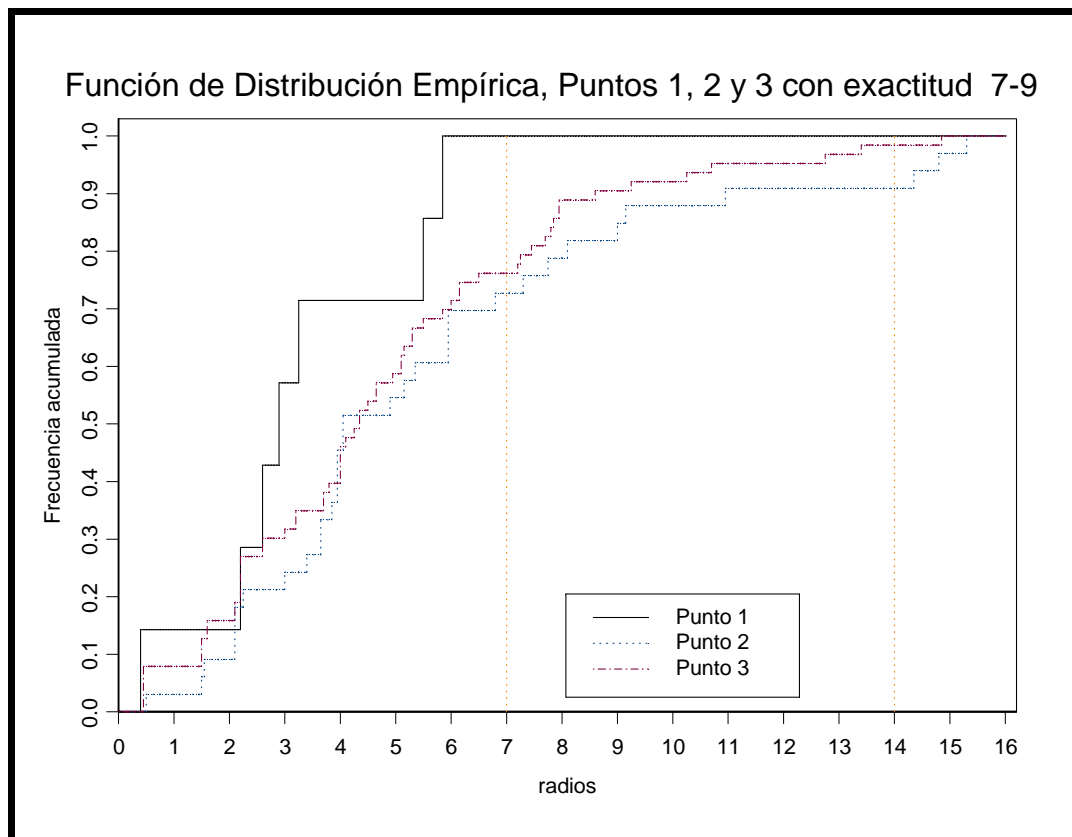


Figura 3.19: *Funciones de distribución empírica para los radios en cada punto que tienen exactitud igual a 7-9 metros.*

Figura 3.20), estas pruebas son las de Kolmogorov-Smirnov, de Lilliefors y las de Shapiro-Wilk, de las cuales se aprecia que las distancias de las mediciones de cada punto a su centroide no se distribuyen normalmente en lo general. Es por esto que no se cumple que en un círculo con radio igual al valor de la *exactitud* registrada haya el 68% de las distancias de las mediciones de cada punto a su centroide y al mismo tiempo se construya otro círculo con radio igual a dos veces la *exactitud* registrada y contenga el 95% de las distancias de las mediciones de cada punto a su centroide. Sin embargo, sólo se cumple que haya aproximadamente el 95% de distancias de las mediciones de cada punto a su centroide en un círculo cuyo radio es igual a dos veces

	Radios del	Pruebas de Normalidad		
		K-S	Lilliefors	Shapiro-Wilk
Exactitud 4	Punto 1	d=0.14580, p>0.20	p<0.05	W=0.90465, p=0.00102
	Punto 2	d=0.13838, p> 0.20	p>0.20	W=0.95665, p=0.74712
	Punto 3	No datos suficientes.		
Exactitud 5	Punto 1	d=0.20177, p<0.05	p<0.01	W=0.84420, p=0.00000
	Punto 2	d=0.11395, p> 0.20	p<0.10	W=0.94344, p=0.01548
	Punto 3	d=0.21105, p>0 .20	p<0.05	W=0.78440, p=0.00051
Exactitud 6	Punto 1	d=0.15526, p> 0.20	p<0.10	W=0.92344, p=0.04782
	Punto 2	d=0.15351, p>0 .20	p<0.01	W=0.84628, p=0.00002
	Punto 3	d=0.20678, p<0.05	p<0.01	W=0.82924, p=0.00000
Exactitud 7-9	Punto 1	d=0.21280, p> 0.20	p> 0.20	W=0.93405, p=0.58587
	Punto 2	d=.17624, p> 0.20	p<0.05	W=0.87905, p=0.00159
	Punto 3	d=0.11367, p> 0.20	p<0.05	W=0.92857, p=0.00128

Figura 3.20: *Pruebas de normalidad de las distancias que resultan de la diferencia entre las mediciones de cada punto y su centroide.*

la *exactitud* dada.

Efectos del movimiento

Otro de los objetivos de este trabajo es observar si el movimiento que experimentó el receptor de GPS al momento de llevarlo de un punto a otro, provocó alteraciones en las mediciones (tanto de posición como de distancias entre dos puntos) que se registraron después del punto de inicio en cada ciclo. Cabe mencionar que al momento de trasladar físicamente de un lugar a otro al receptor también sufría de pérdida de las señales de los satélites. Para investigar este aspecto, se incluye la tabla de la Figura 3.21 que presenta algunas estadísticas descriptivas para los radios, agrupados según el orden de toma que tuvo cada uno de ellos en los ciclos. También se construyeron las gráficas de las Figuras 3.22, 3.23 y 3.24.

En la Figura 3.22 están graficadas las funciones de distribución empírica de las distancias de las mediciones del Punto 1 a su centroide (los radios), agrupados según

Capítulo 3. *Análisis Estadístico*

	Radios en	Datos	Media	Desv. Est.
Punto 1	1er. Lugar	51	3.05118	1.74793
	2do. Lugar	47	2.52395	1.62093
	3er. Lugar	45	2.85787	1.84079
Punto 2	1er. Lugar	45	4.27978	2.08130
	2do. Lugar	51	5.87103	4.27044
	3er. Lugar	47	4.49111	1.83944
Punto 3	1er. Lugar	47	3.22165	2.92256
	2do. Lugar	45	5.61867	3.19544
	3er. Lugar	51	4.27660	2.94313

Figura 3.21: *Estadísticas descriptivas de radios agrupados según el orden en que se registraron durante el recorrido de cada ciclo.*

el orden que le correspondió para ser medido en cada ciclo. En este punto se esperaría que la variabilidad de los radios cuando se registraron en segundo y tercer lugar del recorrido en cada ciclo fuese mayor que cuando se registró en primer lugar debido al movimiento y a la pérdida de señal. Sin embargo, en la gráfica se observa que la variabilidad de los radios fue similar en las tres ordenaciones posibles que tuvo este punto. Esto sugiere que debido a que este punto es donde el receptor tiene mayor visibilidad de los satélites, que el movimiento no produce un efecto de degradación en el receptor debido a una pérdida momentánea de señal.

Este comportamiento contrasta con el que se observa ahora para el Punto 2. En este caso se esperaría que los radios que se registraron en segundo y tercer lugar tuvieran variabilidad semejante, pues son registrados después de que el receptor experimentara movimiento y pérdida de señal. Sin embargo, en la Figura 3.23 se observa que la variabilidad de los radios en el Punto 2 es mayor cuando se registraron en segundo lugar que cuando se registraron en primer y tercer lugar. No hay que despreciar el hecho de que para las mediciones en el Punto 2 cuando se registraron en segundo y tercer lugar, el punto anterior por el que se pasó en el recorrido de cada ciclo es el Punto 1 donde el receptor tiene mayor visibilidad. Aquí, sí hay entonces

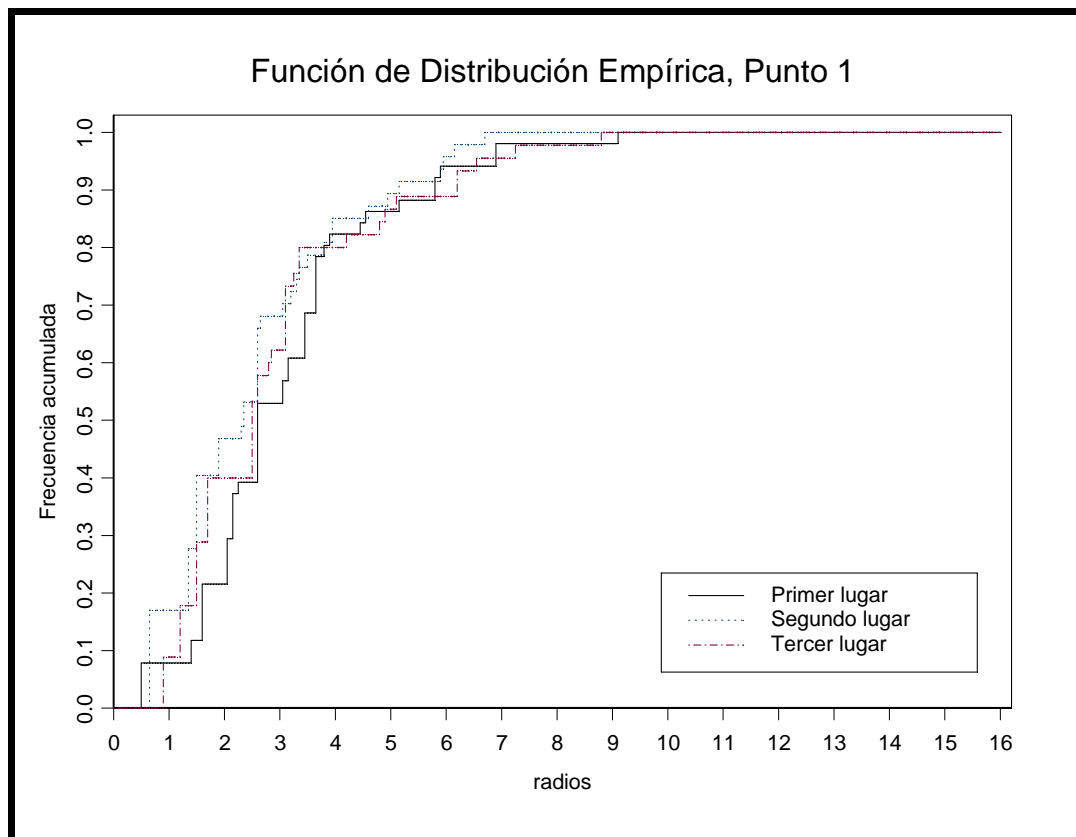


Figura 3.22: *Funciones de distribución empírica de los radios del Punto 1, agrupados según el orden en que se registraron las mediciones de este punto durante el recorrido de cada ciclo.*

una evidencia de que el movimiento está afectando a los valores de las mediciones, pero no es claro cuál es el juego entre visibilidad, pérdida momentánea de señal, y tiempo de encendido.

En la Figura 3.24 se observa que la variabilidad de los radios en el Punto 3 es mayor que en los otros dos puntos. Esto es de esperarse porque la visibilidad que tiene el receptor de GPS en este punto es ciertamente menor que en los otros dos. Sin embargo, no se observa qué efectos pudo haber tenido el movimiento sobre la variabilidad que presentan los radios de este punto de acuerdo a la posición que tuvo

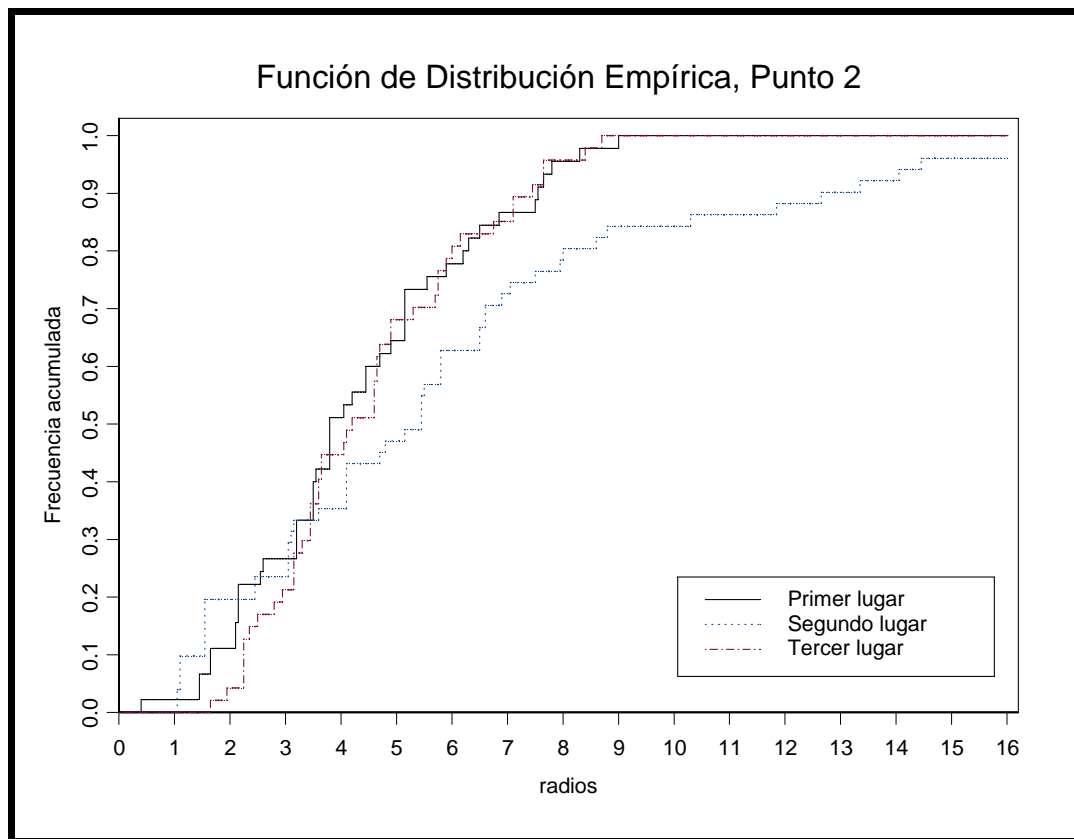


Figura 3.23: *Funciones de distribución empírica de los radios del Punto 2, agrupados según el orden en que se registraron las mediciones de este punto durante el recorrido de cada ciclo.*

éste en cada ciclo. En efecto, las mediciones que se tomaron en tercer lugar tienen menor variabilidad que las que se tomaron en segundo lugar aun cuando se hubiese esperado lo contrario.

La conclusión general respecto a movimiento es que sí parece haber una interacción entre pérdida momentánea de señal, la calidad de la señal justo en el momento de perderla, y la visibilidad presente en el momento de recuperarla. Por otra parte, se verifica que la calidad de la señal que tiene el receptor también tiene que ver con el tiempo de encendido y la visibilidad.

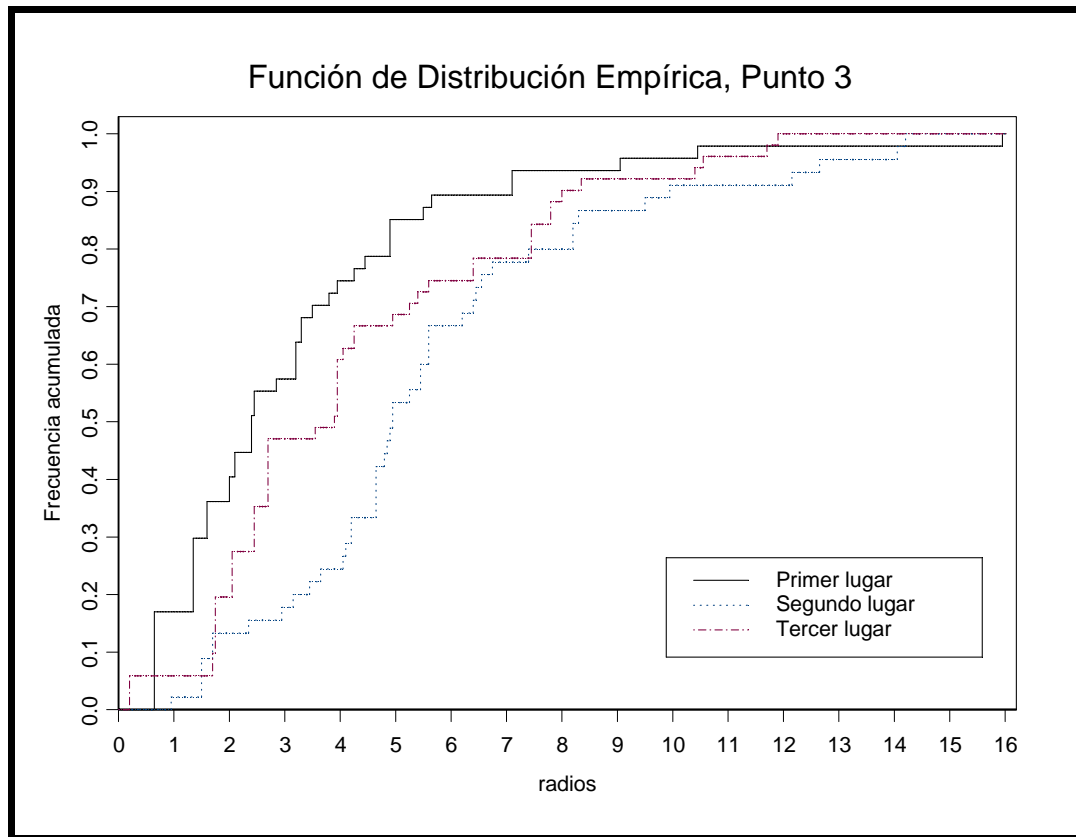


Figura 3.24: *Funciones de distribución empírica de los radios del Punto 3, agrupados según el orden en que se registraron las mediciones de este punto durante el recorrido de cada ciclo.*

3.2.2 Mediciones de distancias entre puntos

En la Figura 3.25 se presentan algunas estadísticas descriptivas (media, varianza, desviación estándar, asimetría y kurtosis) de las distancias entre pares de puntos. En esta tabla se observa que la función de densidad que describen las mediciones en la Distancia P1P2 es ligeramente asimétrica por la derecha y las funciones de densidad que se forman con las mediciones de las Distancias P2P3 y P3P1 son asimétricas por la izquierda. Las mediciones de la Distancia P1P2 son las menos variables comparándolas con las mediciones de las otras dos distancias entre pares de puntos, lo

Capítulo 3. *Análisis Estadístico*

Distancia	Media	Varianza	Desv. Est.	Asimetría	Kurtosis
P1P2	36.30699	4.84347	2.200789	0.766136	1.052986
P2P3	69.14406	10.95023	3.309113	-0.726008	0.738151
P3P1	64.46294	13.18615	3.631274	-0.062928	4.374738

Figura 3.25: *Estadísticas descriptivas de las distancias entre pares de puntos, medidas ciclo por ciclo.*

cual tiene sentido físico porque estos dos puntos son los que por diseño experimental conjuntamente cuentan con la mayor visibilidad. Por otra parte, la función de densidad que describe la Distancia P2P3 es la menos puntiaguda de las tres y la más puntiaguda es la que describen las mediciones de la Distancia P3P1. Sin embargo, conforme a la interpretación que tiene el coeficiente de kurtosis, las funciones de densidad de las mediciones en las tres distancias entre pares de puntos son más puntiagudas que una con distribución normal. Ya a partir de estas estadísticas descriptivas comienza a vislumbrarse que si bien las distribuciones de distancias son casi simétricas, éstas distan de ser distribuciones normales. Comparando las medias de las mediciones de las distancias entre pares de puntos con los valores reales obtenidos por topógrafos, se tiene que la diferencia de la media de las mediciones de Distancia P1P2 y el valor de la distancia real es del orden de un metro (0.876 metros) y la diferencia entre los promedios de las otras dos distancias entre pares de puntos y los valores reales de las mismas es del orden de dos metros. Debido a que el tamaño de muestra es 143, la desviación estándar de la media es el valor anotado en la Tabla 3.25 para “Desv. Est.” dividido entre $\sqrt{143}$. En la Figura 3.26, se indican los errores estándar de la media, y la magnitud del error con respecto al valor verdadero, en unidades de errores estándar. Esta última tabla muestra que como instrumento de medición de la distancia entre dos puntos, el receptor de GPS es un instrumento inexacto, al menos a la escala utilizada en una actividad tal como la construcción.

La Figura 3.27 muestra histogramas que describen las distancias medidas por

Capítulo 3. Análisis Estadístico

Distancia	Valor real de topógrafos	Valor medio experimental	Error Estándar	# Errores Estándar
P1P2	35.431	36.30699	0.184039	4.7
P2P3	70.947	69.14406	0.276722	6.5
P3P1	62.848	64.46294	0.303662	5.3

Figura 3.26: Diferencia entre el valor real y el promedio experimental para cada distancia entre dos puntos en términos del error estándar.

el GPS para cada par de puntos, diagramas de dispersión correspondientes, y coeficientes de correlación estimados. Cada triada de distancias fue obtenida durante la ejecución de un mismo ciclo, como se describió en la Sección 2.3. En la figura se observa que las distancias en general poseen una distribución simétrica (aunque para P1P2 y para P2P3 existen ligeros indicios de asimetría hacia la derecha y la izquierda, respectivamente). Los diagramas de dispersión en esta misma figura sugieren que entre P1P2 y P2P3, o entre P1P2 y P3P1 hay poca correlación, mientras que entre P2P3 y P3P1 hay un indicio de correlación positiva.

Las gráficas Cuantil-Cuantil que se presentan en la Figura 3.28 corresponden a las mediciones de las distancias entre pares de puntos. De estas gráficas se observa que las curvas en general tienen colas más pesadas que una normal.

La tabla que se presenta en la Figura 3.29 muestra algunas pruebas formales de normalidad, las cuales son las de Kolmogorov-Smirnov, de Lilliefors y las de Shapiro-Wilk. Se concluye que, en efecto, se rechaza la hipótesis de normalidad en las tres funciones de densidad que describen las distancias entre pares de puntos. Se realizó una estimación no-paramétrica para cada una de las funciones de densidad de las distancias entre pares de puntos utilizando el método de kernel normal univariado con un ancho de banda de 1.245, las cuales se muestran en la Figura 3.30. En ella se observa de una manera alternativa todo lo descrito en los párrafos anteriores. En términos de funciones de distribución estimadas por vía de distribuciones empíricas,

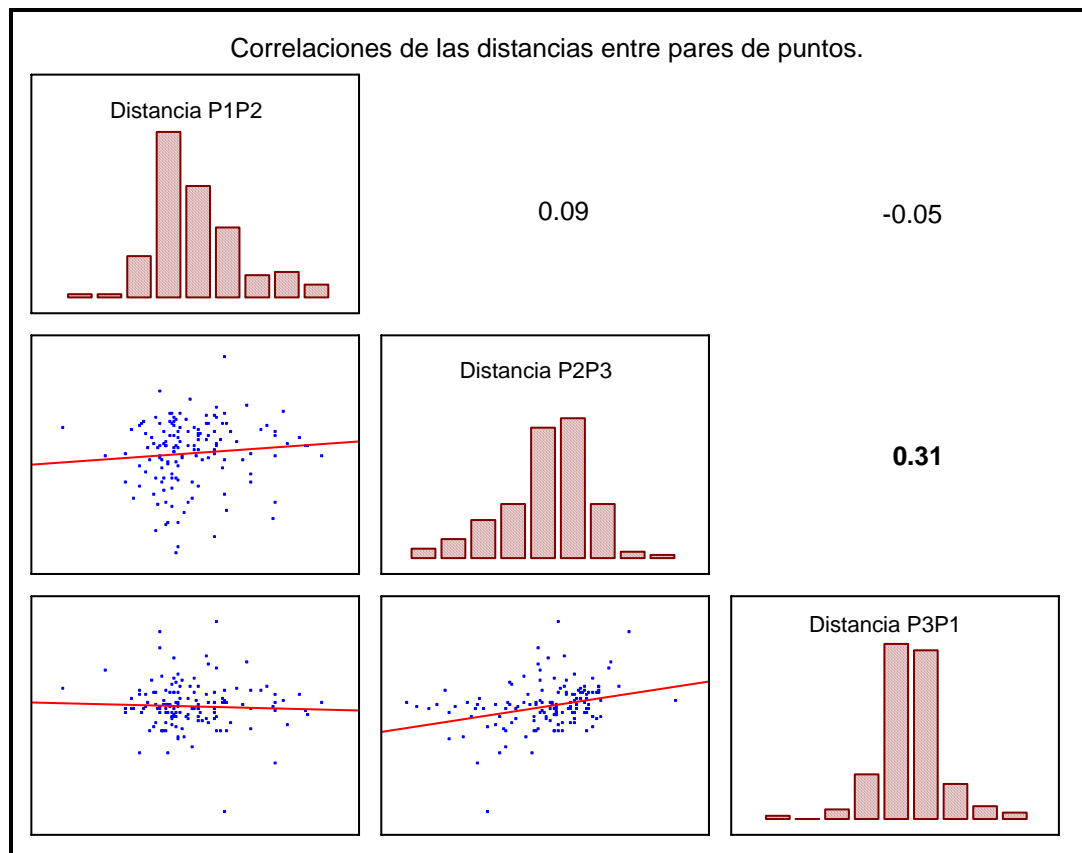


Figura 3.27: *Valores y gráficas de la correlación e histogramas de cada una de las distancias entre dos puntos.*

las diferencias anteriores se pueden observar en la Figura 3.31.

Efectos del movimiento

Como ya se había mencionado, uno de los objetivos de este trabajo es explorar qué posibles efectos tiene el movimiento que sufre el receptor de GPS en el momento de llevarlo de un lugar a otro considerando que por segundos el receptor no recibe la señal de los satélites durante el traslado físico. Para esto se consideraron las diferencias entre mediciones de distancias entre pares de puntos y el valor real de cada distancia.

Capítulo 3. *Análisis Estadístico*

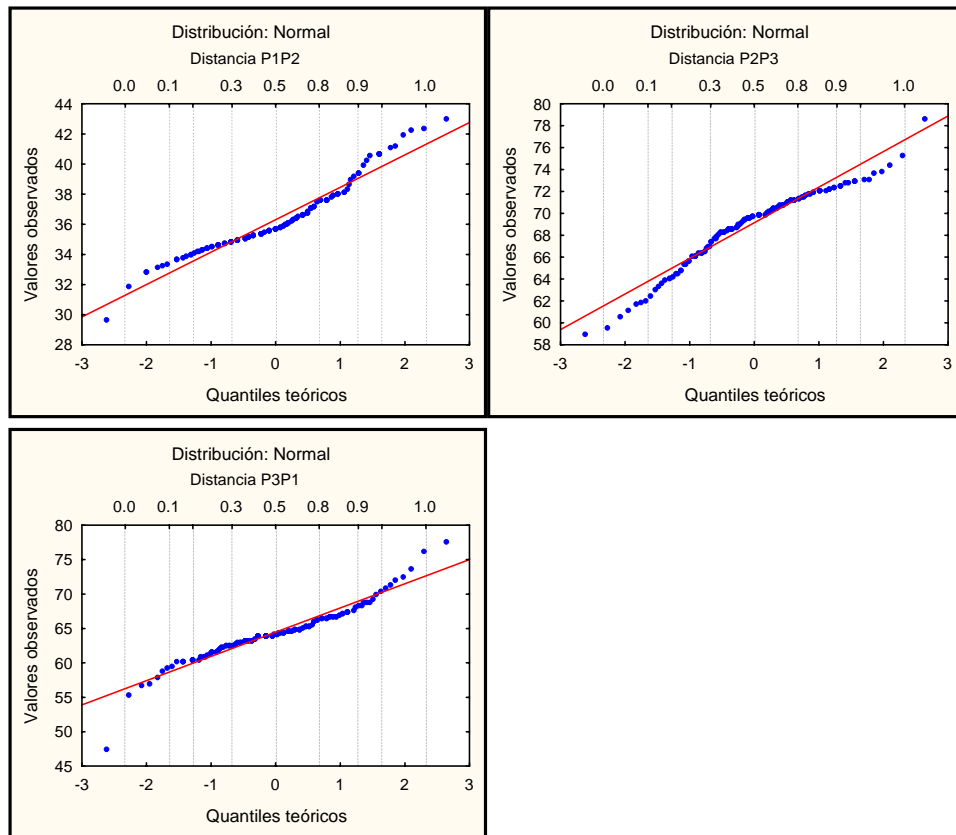


Figura 3.28: *Gráficas Cuantil-Cuantil de las mediciones de las distancias entre pares de puntos.*

Estas cantidades se denominarán *errores*. Se graficaron las funciones de distribución empírica para los errores, agrupando según el orden en que se eligió (al azar) el punto

Distancia	Pruebas de Normalidad		
	K-S	Lilliefors	Shapiro-Wilk
P1P2	d=0.12956, p<0.05	p<0.01	W=0.94122, p=0.00001
P2P3	d=0.11626, p<0.05	p<0.01	W=0.95188, p=0.00007
P3P1	d=0.09855, p<0.15	p<0.01	W=0.93346, p=0.00000

Figura 3.29: *Pruebas de normalidad de las distancias entre pares de puntos.*

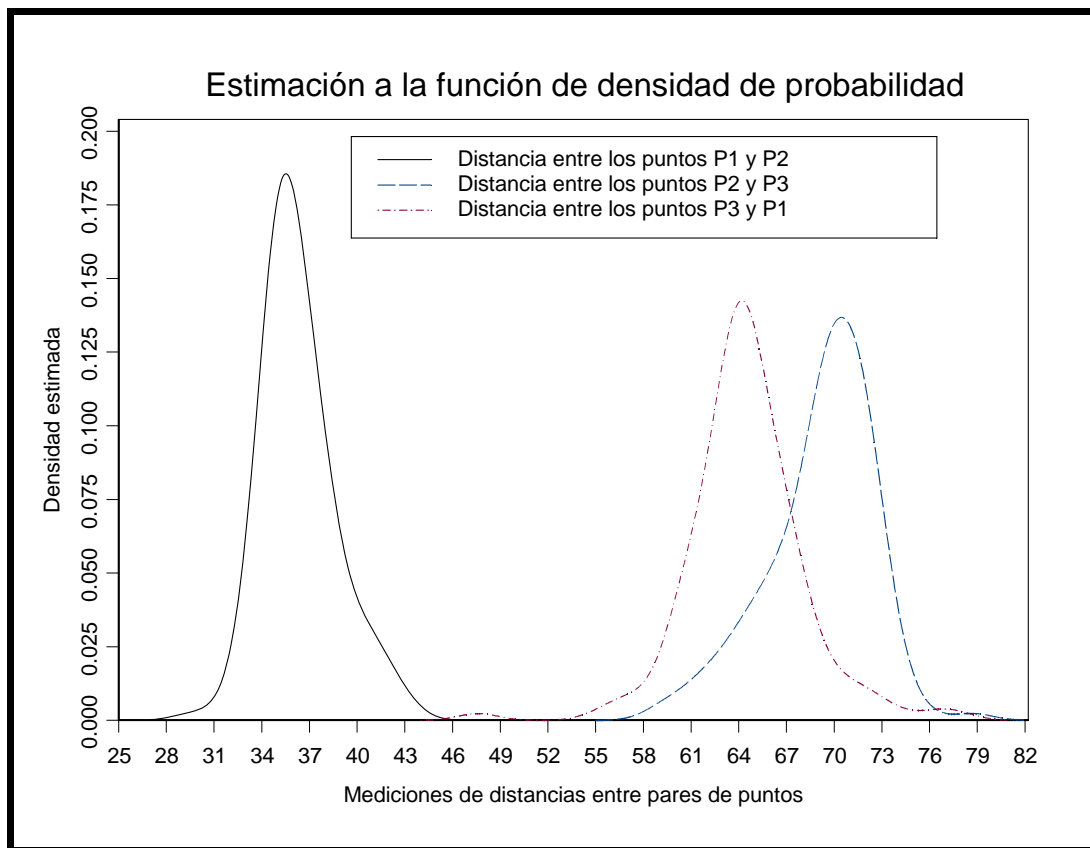


Figura 3.30: *Funciones de densidad estimadas por el método de kernel, para las mediciones de las distancias entre pares de puntos.*

de arranque para cada ciclo. Estos distintos arranques representan diferentes formas en que varió la visibilidad durante el traslado, causado por obstáculos que evitaban una vista despejada del cielo. Estas se muestran en las Figuras 3.32, 3.33 y 3.34.

La Figura 3.32 muestra las funciones de distribución empírica para errores en Distancia P1P2. En ellas se observa que, en general, la variabilidad de errores en las tres curvas es similar, es decir, las variaciones en visibilidad a lo largo del recorrido en el ciclo no tuvieron efectos notables sobre la calidad de las mediciones. Esto podría ser esperado, debido a que estos dos puntos son precisamente los que mejor visibilidad tienen.

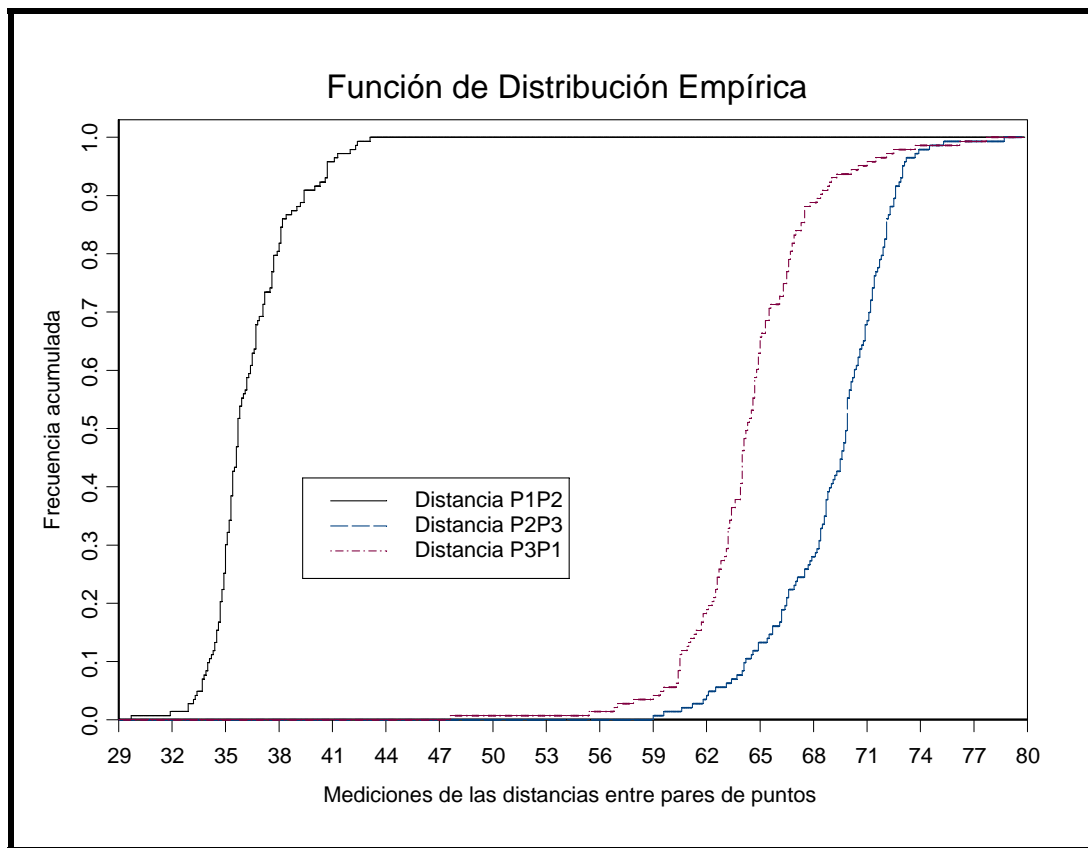


Figura 3.31: *Funciones de distribución empírica para las mediciones de distancias entre pares de puntos.*

En contraste, en la Figura 3.33 se presentan las funciones de distribución empírica para errores en Distancia P2P3. Aquí sí se observa una diferencia en las distribuciones. Una de las tres distribuciones es notoriamente diferente en media y en varianza. Esta distribución corresponde al caso en que la distancia P2P3 fue medida estando el GPS localizado en el punto P3, después de haber sido desplazado a través de los puntos P1 y P2 (en ese orden). Esto sugiere que la pérdida incidental de visibilidad sí tiene algún efecto sobre la calidad de las mediciones. De otra manera, estas tres distribuciones también hubieran resultado ser iguales.

La misma situación se aprecia en la Figura 3.34, que presenta las funciones de

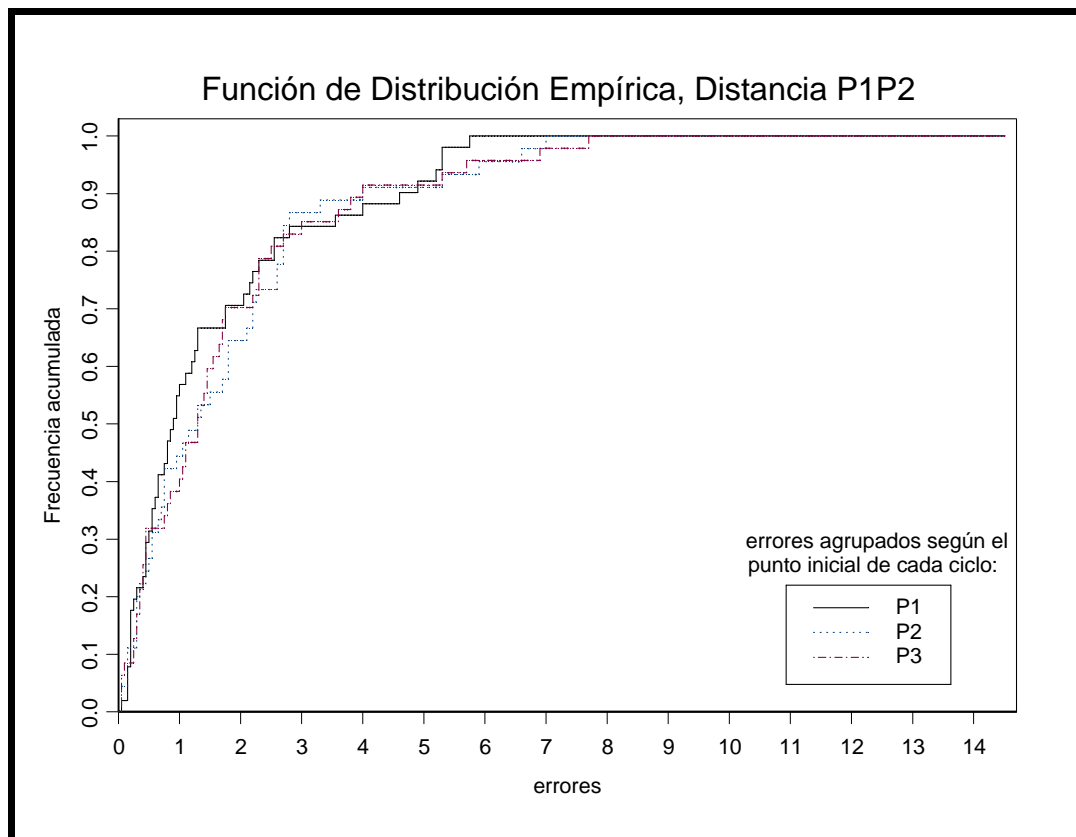


Figura 3.32: *Funciones de distribución empírica de los errores de la Distancia P1P2 agrupados según el punto inicial de cada ciclo.*

distribución empírica de errores en Distancia P3P1. De nuevo, hay una diferencia notable entre distribuciones, y debido a que la única diferencia entre estos tres grupos es el perfil en el que fue variando la visibilidad durante el traslado, se concluye que dicha variación es la que produce las diferencias.

Para complementar las gráficas anteriores, en la Figura 3.35 se presentan algunas estadísticas descriptivas básicas (media y la desviación estándar) de los errores en las distancias entre pares de puntos.

Nuevamente, como en el caso de las mediciones de posición, la conclusión general

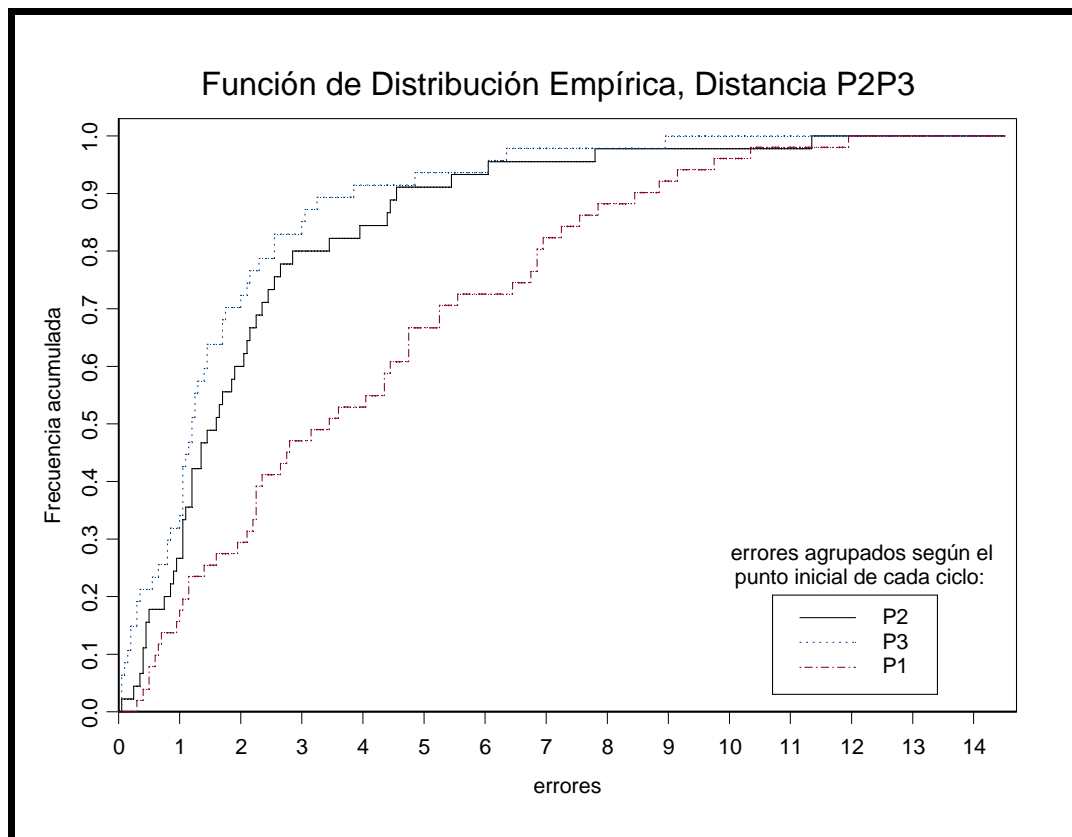


Figura 3.33: *Funciones de distribución empírica de los errores de la Distancia P2P3 agrupados según el punto inicial de cada ciclo.*

respecto al movimiento es que al parecer que sí hay una interacción entre pérdida momentánea de señal, la calidad de la señal justo en el momento de perderla, la dirección en la que se mide una distancia, la visibilidad presente en el momento de recuperarla y el tiempo de encendido que tiene el receptor. Por otra parte, como instrumento de medición de distancias, fue claro que el GPS es un instrumento muy inexacto para aplicaciones que requieran de una escala del orden de magnitud que surgió en el experimento (100m aproximadamente), tales como la industria de la construcción.

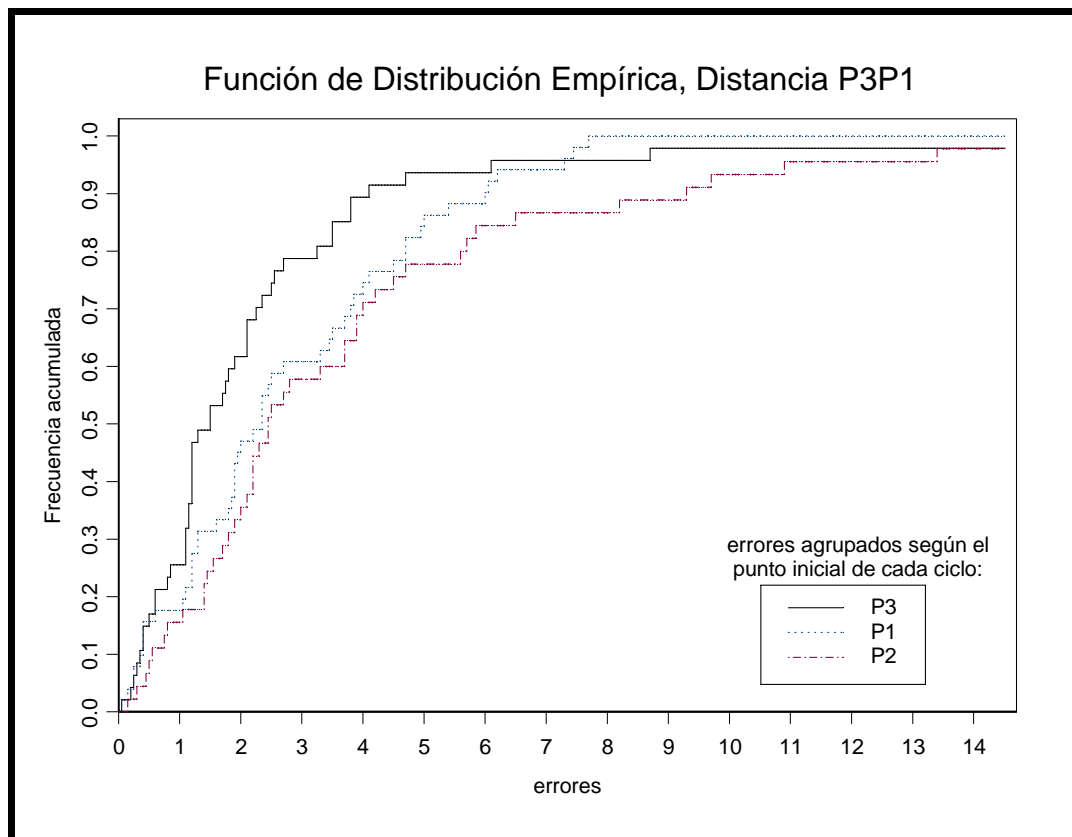


Figura 3.34: *Funciones de distribución empírica de los errores de la Distancia P3P1 agrupados según el punto inicial de cada ciclo.*

3.3 Conclusiones

El primer objetivo planteado fue estudiar de manera empírica el dato llamado *exactitud* que proporciona un receptor de GPS civil. El manual del receptor de GPS (eTrex Summit), que se utilizó para esta tarea, menciona que el error que despliega al proporcionar la exactitud de la posición de un punto dado es 2DRMS asociado a una probabilidad 95%, mientras que la literatura asociada menciona que el error que se presenta en relación a la *exactitud* que proporciona el receptor es un error llamado DRMS, asociado a una probabilidad de 68%. Esto indica una confusión sobre la

Capítulo 3. *Análisis Estadístico*

Distancia	Errores agrupados según el pto. Inicial de cada ciclo	Datos	Media	Desv. Est.
P1P2	P1	51	1.52684	1.61238
	P2	45	1.73158	1.72638
	P3	47	1.71189	1.77436
P2P3	P2	45	2.20220	2.17230
	P3	47	1.71326	1.79951
	P1	51	4.06814	3.04286
P3P1	P3	47	2.16170	2.50672
	P1	51	2.76439	2.06585
	P2	45	3.69547	3.43459

Figura 3.35: *Estadísticas descriptivas básicas de los errores en las distancias entre pares de puntos.*

interpretación correcta que deberá tener el valor desplegado por el equipo.

Se observó—tal y como sería natural anticiparlo—que existe relación entre la visibilidad que tiene el receptor de los satélites en cada punto y la variabilidad que se presentó en cada una de las nubes de mediciones. En efecto, la nube de mediciones que se formó en el Punto 1 tiene menor variabilidad, dado que aquí el receptor tiene una visibilidad total de los satélites. En orden incremental, se encuentra la variabilidad en las mediciones del Punto 2 y del Punto 3, en correspondencia directa con la visibilidad—media y baja respectivamente—de los satélites que tiene el receptor en estos puntos.

Para estudiar más de cerca esta variabilidad, en este trabajo se realizaron análisis con relación a la distribución que describen los *radios* (distancias de una medición a un centroide) tomando en cuenta la *exactitud* que se registró en cada punto. Una primera observación es que dichos radios no se distribuyen normalmente, por lo que se concluye que no son automáticamente consistentes entre sí los requerimientos probabilísticos de 68% para DRMS y 95% para 2DRMS. Si bien sí se cumple que al 95% de probabilidad le corresponde aproximadamente dos veces el valor de la

Capítulo 3. *Análisis Estadístico*

exactitud que reporta el receptor, la *exactitud* por sí misma no coincide con abarcar el 68%. El significado exacto que se ha descubierto es el siguiente: *exactitud* desplegada en el GPS es la mitad del cuantil 95% de la distribución de radios. El valor de la *exactitud* que proporciona el receptor de GPS es una verdad a medias. Sí se podría asociar su valor a una noción de exactitud a través de radios y distancias en línea recta. La exactitud del GPS sí cuantifica el concepto correcto en el sentido de que entre más grande sea, más inexactitud se tiene. Pero, la interpretación adecuada del número no es clara, y de hecho, se encuentran contradicciones en los manuales de uso.

Por otra parte, el movimiento que experimentó el receptor de GPS al momento de trasladarlo de un punto a otro, considerando que durante su traslado físico habían construcciones que causaron que por segundos no recibiera señales de los satélites, sí afectó la distribución de valores de las mediciones tanto de posición como de distancias entre pares de puntos. Se detecta que hay algún tipo de interacción entre la pérdida de señal, la calidad de la señal que tiene el receptor justo antes de perderla, la visibilidad presente al momento de recuperarla, el tiempo de encendido que tiene el receptor y, para el caso de las mediciones en las distancias entre pares de puntos, la dirección en que se realizó la medición. Sin embargo, no es posible estudiar con mayor detalle la naturaleza de estas interacciones, puesto que esto requeriría de un diseño experimental mucho más fino, con mayor control sobre estos factores.

Con base en el análisis estadístico que se realizó para las mediciones de posición se concluye que su distribución no es normal bivariada. En lo general las funciones de densidad que describen los tres pares de marginales tienen colas más pesadas que una normal y son ligeramente asimétricas. Además se observó que sí hay cierta dependencia entre las mediciones de posición obtenidas con el receptor de GPS.

Por otra parte, con relación al empleo de un receptor de GPS como un instrumento para medir distancias entre dos puntos, se concluye que existen algunos incon-

Capítulo 3. *Análisis Estadístico*

venientes, sobre todo a la escala del experimento. De la comparación que se realizó entre los valores reales y los promedios de las mediciones de las distancias entre pares de puntos se observó que el receptor de GPS es un instrumento inexacto, puesto que las diferencias entre la media de las mediciones y la distancia verdadera no es despreciable. En unidades de errores estándar, la Distancia P1P2 tiene 4.7, la Distancia P2P3 tiene 6.5 y la Distancia P3P1 tiene 5.3. Esto significa que el receptor de GPS como un instrumento de medición de distancias tiene un sesgo (bias) muy notable.

Respecto a las mediciones de las distancias entre pares de puntos se obtiene que su distribución no es normal, sino que son descritas por funciones de densidad que difieren de una normal en las colas, aunque en general sí son simétricas. También se observó que puede haber una correlación significativa entre dos mediciones consecutivas de dos segmentos diferentes, lo cual es una característica indeseable en un instrumento de medición. Sin embargo, cabe mencionar que a escalas mayores, por ejemplo para usos de navegación (100km), la medición de distancias con un receptor de GPS tenga una exactitud enteramente útil para esas aplicaciones. Esto es, que el sesgo (bias) pudiera ser prácticamente despreciable.

Apéndices

A Estimadores kernel de densidad

B Base de Datos

Apéndice A

Estimadores kernel de densidad

Se tiene una muestra independiente e idénticamente distribuida (iid) de un vector aleatorio $\vec{X} = \{\vec{X}_1, \dots, \vec{X}_n\}$, con función de densidad $f(\vec{x})$ desconocida. En general la densidad puede ser multivariada, es decir, que \vec{X}_i son vectores aleatorios de dimensión d , aunque es muy común el caso $d = 1$. Se puede estimar a $f(\vec{x})$, para lo cual existen varios estimadores no paramétricos. En este apéndice se presenta un resumen de propiedades teóricas elementales que tiene una clase de estimadores de densidad llamada estimadores kernel¹². Con la notación $\hat{f}(\vec{x})$ denotaremos la función de densidad estimada en el punto \vec{x} .

A.0.1 Definición del estimador kernel multivariado

Dada una muestra de n observaciones vectoriales $\vec{X}_1, \dots, \vec{X}_n$, donde $\vec{X}_i \in \mathbb{R}^d$, definimos al estimador de densidad kernel multivariado con función kernel $K : \mathbb{R}^d \rightarrow \mathbb{R}$

¹La idea de la estimación de densidades kernel fue propuesta por M. Rosenblatt en 1956.

²Para mayores detalles consultar Silverman (1986) y Scott (1992).

Apéndice A. Estimadores kernel de densidad

de la forma

$$\hat{f}(\vec{x}) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{\vec{x} - \vec{X}_i}{h}\right),$$

donde h es ancho de banda o parámetro de suavización. La función kernel $k(\vec{x})$ es generalmente una función de densidad multivariada que satisface

$$\int_{R^d} K(\vec{x})d(\vec{x}) = 1.$$

El estimador kernel puede interpretarse como la suma de protuberancias³ centradas en las observaciones. La función K determina la forma de las protuberancias mientras que el parámetro de suavización h determina su anchura.

Frecuentemente, pero no siempre, K es una función de densidad de probabilidad unimodal radialmente simétrica. Por otra parte, se sigue que $\hat{f}(x)$ es una función de densidad de probabilidad por definición. Además $\hat{f}(x)$ hereda todas las propiedades de continuidad y diferenciabilidad del kernel K , de manera que si, por ejemplo, K es la función de densidad normal, entonces $\hat{f}(x)$ será una curva lisa que tiene derivadas para cualquier orden.

El estimador $\hat{f}(x)$ depende de los datos, del kernel y del ancho de banda; esta dependencia generalmente no se expresa explícitamente. Para cada \vec{x} , $f(\vec{x})$ se puede pensar como un vector aleatorio, debido a que depende de las observaciones $\vec{X}_1, \dots, \vec{X}_n$. Cualquier uso de la probabilidad, esperanza y varianza que involucran $\hat{f}(x)$ es con respecto a su distribución muestral como una estadística basada en estas observaciones aleatorias.

Las funciones kernel más utilizadas en dimensión d son:

- Función de densidad normal multivariada estándar, dada por

$$K(\vec{x}) = (2\pi)^{-\frac{d}{2}} \exp\left(-\frac{1}{2} \vec{x}^T \vec{x}\right).$$

³del inglés *bump*

Apéndice A. Estimadores kernel de densidad

- Función kernel Epanechnikov multivariado, dado por

$$K_e(\vec{x}) = \begin{cases} \frac{1}{2}c_d^{-1}(d+2)(1 - \vec{x}^T\vec{x}) & \text{si } \vec{x}^T\vec{x} < 1 \\ 0 & \text{en otro caso,} \end{cases}$$

donde c_d es el volumen de la esfera unidad de dimensión d : por ejemplo, $c_1 = 2$, $c_2 = \pi$, $c_3 = \frac{4\pi}{3}$, etc.

- Otras funciones kernel útiles para el caso $d = 2$ son:

$$K_2(\vec{x}) = \begin{cases} 3\pi^{-1}(1 - \vec{x}^T\vec{x})^2 & \text{si } \vec{x}^T\vec{x} < 1, \\ 0 & \text{en otro caso} \end{cases}$$

$$K_3(\vec{x}) = \begin{cases} 4\pi^{-1}(1 - \vec{x}^T\vec{x})^3 & \text{si } \vec{x}^T\vec{x} < 1, \\ 0 & \text{en otro caso} \end{cases}$$

El uso de un solo parámetro suavizador h en la definición de kernel multivariado implica que la versión del kernel situado en cada punto del dato multivariado tiene la misma escala en todas las direcciones. En ciertas circunstancias, podría ser más apropiado usar un vector de parámetros suavizadores. Por ejemplo, en el caso en que se tiene que una de las coordenadas direccionales de los puntos es mucho mayor que las otras.

Fukunaga (1972) intuitivamente sugiere que primero se pre-escriban los datos transformandolos linealmente para obtener una matriz unidad de covarianzas, luego usar un kernel radialmente simétrico y finalmente transformarlos de nuevo a su versión anterior. Esto es equivalente a usar un estimador de densidad de la forma:

$$\hat{f}(\vec{x}) = \frac{(\det S)^{-\frac{1}{2}}}{nh^d} \sum_{i=1}^n k \left\{ h^{-2} (\vec{x} - \vec{X}_i)^T S^{-1} (\vec{x} - \vec{X}_i) \right\},$$

donde la función k está dada por

$$k(\vec{x}^T\vec{x}) = K(\vec{x}),$$

Apéndice A. Estimadores kernel de densidad

y S es la matriz de covarianzas muestral de los datos.

Varias medidas se han propuesto para cuantificar la discrepancia del estimador de densidad $\hat{f}(\vec{x})$ de la verdadera densidad $f(\vec{x})$. Cuando se considera la estimación de un sólo punto \vec{x} , una medida natural es el MSE⁴, definido por:

$$\text{MSE}_{\vec{x}}(\hat{f}) = \text{E} \left\{ \hat{f}(\vec{x}) - f(\vec{x}) \right\}^2.$$

Por las propiedades elementales estándar de la media y la varianza,

$$\text{MSE}_{\vec{x}}(\hat{f}) = \left\{ \text{E}\hat{f}(\vec{x}) - f(\vec{x}) \right\}^2 + \text{var}\hat{f}(\vec{x}),$$

la suma del cuadrado del sesgo y la varianza de \vec{x} . Veremos que, como en muchas ramas de la estadística hay una cierta relación entre el sesgo y la varianza. El sesgo puede reducirse a expensas de que la varianza aumente y viceversa, ajustando h .

La primera medida (Rosenblatt,1956), ampliamente usada para evaluar la exactitud *global* de $\hat{f}(\vec{x})$ como un estimador de $f(\vec{x})$ es el MISE⁵, definido por:

$$\text{MISE}(\hat{f}) = \text{E} \int_{\mathbb{R}^d} \left\{ \hat{f}(\vec{x}) - f(\vec{x}) \right\}^2 d(\vec{x}).$$

Es importante notar que, ya que el integrando es no negativo, el orden de integración y la esperanza en la expresión anterior puede invertirse para dar la forma alternativa

$$\begin{aligned} \text{MISE}(\hat{f}) &= \int_{\mathbb{R}^d} \text{E} \left\{ \hat{f}(\vec{x}) - f(\vec{x}) \right\}^2 d(\vec{x}) \\ &= \int_{\mathbb{R}^d} \text{MSE}_{\vec{x}}(\hat{f}) d(\vec{x}) \\ &= \int_{\mathbb{R}^d} \left\{ \text{E}\hat{f}(\vec{x}) - f(\vec{x}) \right\}^2 d(\vec{x}) + \int_{\mathbb{R}^d} \text{var}\hat{f}(\vec{x}) d(\vec{x}), \end{aligned}$$

que da el MISE como la suma del sesgo cuadrado integrado y la varianza integrada.

⁴por *mean square error*

⁵por *mean integrated square error*

Propiedades de aproximación para sesgo y varianza

Suponga que el kernel K es una función de densidad de probabilidad radialmente simétrica y que la densidad desconocida $f(\vec{x})$ es acotada y tiene segundas derivadas continuas. Defina las constantes α y β como

$$\alpha = \int t_1^2 K(\vec{t}) d(\vec{t}),$$

y

$$\beta = \int K(\vec{t})^2 d(\vec{t}).$$

Usando la forma multidimensional del teorema de Taylor, se obtienen las aproximaciones

$$\text{sesgo}_h(\vec{x}) \approx \frac{1}{2} h^2 \alpha \nabla^2 f(\vec{x}),$$

y

$$\text{var} \hat{f}(\vec{x}) \approx n^{-1} h^{-d} \beta f(\vec{x}).$$

Combinando estos resultados se obtiene una aproximación para el MISE dada por

$$\frac{1}{4} h^4 \alpha^2 \int \{\nabla^2 f(\vec{x})\}^2 d(\vec{x}) + n^{-1} h^{-d} \beta.$$

Propiedades asintóticas

Los estimadores de densidad considerados se construyen de las primeras n observaciones en una sucesión independiente e idénticamente distribuida de variables aleatorias $\vec{X}_1, \dots, \vec{X}_n$ con densidad $f(\vec{x})$. Con esto se está diciendo implícitamente que el ancho de banda h depende de alguna manera del tamaño de la muestra.

Consistencia La consistencia de la estimación de $f(\vec{x})$ a un punto \vec{x} fue estudiada por Parzen (1962). Suponga que el kernel K es una función de Borel acotada, que satisface

$$\int |K(\vec{t})| d(\vec{t}) < \infty, \quad \int K(\vec{t}) d(\vec{t}) = 1, \quad \text{y} \quad |K(\vec{t})| \longrightarrow 0 \quad \text{cuando} \quad |\vec{t}| \longrightarrow \infty.$$

Estas condiciones las cumplen cualquiera de las funciones kernel que son usuales.

Suponga que, el ancho de banda h_n satisface

$$h_n \longrightarrow 0 \quad \text{y} \quad nh_n \longrightarrow \infty \quad \text{cuando} \quad n \longrightarrow \infty.$$

Bajo estas condiciones se muestra que, por ser $f(\vec{x})$ continua en \vec{x} ,

$$\hat{f}(\vec{x}) \longrightarrow f(\vec{x}) \quad \text{en probabilidad cuando} \quad n \longrightarrow \infty.$$

Las condiciones anteriores para h_n son típicas para cuando se requiere consistencia. Ello implica que, el ancho de banda se hace más pequeño cuando el tamaño de la muestra aumenta, y debe converger a cero, pero a un ritmo más lento que n^{-1} .

Consistencia Uniforme La consistencia uniforme, que es convergencia en probabilidad de $\sup |\hat{f}(\vec{x}) - f(\vec{x})|$ a cero, ha sido considerada por varios autores, por ejemplo Parzen (1962), y Silverman (1978b). Suponga que el kernel $K(\vec{x})$ es acotado, tiene variación acotada y satisface

$$\int |K(\vec{t})| d(\vec{t}) < \infty \quad \text{y} \quad \int K(\vec{t}) d(\vec{t}) = 1,$$

y que el conjunto de discontinuidades de K tiene medida cero de Lebesgue. Suponga que

$$f(\vec{x}) \quad \text{es uniformemente continua en} \quad (-\infty, +\infty)$$

y el ancho de banda h_n satisface

$$h_n \longrightarrow 0 \quad \text{y} \quad nh_n(\log n)^{-1} \longrightarrow \infty \quad \text{cuando} \quad n \longrightarrow \infty.$$

Apéndice A. Estimadores kernel de densidad

Bertrand-Retali (1978) muestran, por un ingenioso pero complicado argumento, que se da el caso en que, con probabilidad 1,

$$\sup_{\vec{x}} \left| \hat{f}(\vec{x}) - f(\vec{x}) \right| \longrightarrow 0 \text{ cuando } n \longrightarrow \infty.$$

Además, las condiciones $f(\vec{x})$ es uniformemente continua en $(-\infty, +\infty)$ y $h_n \longrightarrow 0$ y $nh_n (\log n)^{-1} \longrightarrow \infty$ cuando $n \longrightarrow \infty$, son tanto necesarias como suficientes para la consistencia uniforme.

Normalidad asintótica Suponga que $f(\vec{x})$ tiene segunda derivada. La normalidad asintótica del estimador kernel

$$\left(\hat{f}_n(\vec{x}) - f(\vec{x}) \right) \xrightarrow{d} N \left(\frac{c^2}{2} \nabla^2 f(\vec{x}) \alpha, \frac{1}{nh^d} f(\vec{x}) \beta \right) \text{ cuando } n \longrightarrow \infty.$$

Suponiendo que el término del sesgo es despreciable con respecto a la desviación estándar, el intervalo de confianza para $f(\vec{x})$ es de la forma

$$\left[\hat{f}_n(\vec{x}) - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{f}_n(\vec{x})\beta}{nh^d}}, \hat{f}_n(\vec{x}) + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{f}_n(\vec{x})\beta}{nh^d}} \right].$$

El valor $z_{1-\frac{\alpha}{2}}$ es el cuantil $(1 - \frac{\alpha}{2})$ de la distribución normal estándar.

A.0.2 Estimador kernel univariado ($d = 1$)

Ahora, a manera de ilustración, se presenta el desarrollo de como se obtienen las aproximaciones del sesgo y la varianza así como también una tabla que muestra las funciones kernel K más utilizadas, todo esto en el caso en que $d = 1$, ya que el uso del estimador kernel univariado es muy común.

Propiedades de aproximación en sesgo y varianza

Suponga que el kernel K es una función simétrica que satisface,

$$\int K(t)dt = 1, \quad \int tK(t)dt = 0, \quad \text{y} \quad \int t^2K(t)dt = k_2 \neq 0,$$

y que la función de densidad $f(x)$ desconocida tiene derivadas continuas en todos los órdenes requeridos.

El sesgo en la estimación de $f(x)$ no depende directamente del tamaño de la muestra, pero sí depende del ancho de banda h . Por supuesto, si h se elige como función de n , entonces el sesgo dependerá indirectamente de n , lo cual se escribe como

$$\begin{aligned} \text{sesgo}_h(x) &= E\hat{f}(x) - f(x) \\ &= \int h^{-1}K \{(x - y)h^{-1}\} f(y)dy - f(x). \end{aligned}$$

El sesgo también depende del kernel K , pero esta dependencia no se expresa explícitamente.

Sea $y = x - th$. Usando el supuesto de que K integra 1, tenemos

$$\begin{aligned} \text{sesgo}_h(x) &= \int K(t)f(x - th)dt - f(x) \\ &= \int K(t) \{f(x - th) - f(x)\} dt. \end{aligned}$$

Dada la expansión de la serie de Taylor

$$f(x - th) = f(x) - ht f'(x) + \frac{1}{2}h^2 t^2 f''(x) + \dots$$

tal que, por la suposición hecha para K ,

$$\begin{aligned} \text{sesgo}_h(x) &= -hf'(x) \int tK(t)dt + \frac{1}{2}h^2 f''(x) \int t^2K(t)dt + \dots \\ &= \frac{1}{2}h^2 f''(x)k_2 + \text{términos de mayor orden en } h. \end{aligned}$$

Apéndice A. Estimadores kernel de densidad

Integrando el sesgo cuadrado, requerido para MISE, se tiene

$$\int \text{sesgo}_h(x)^2 dx \approx \frac{1}{4} h^4 k_2 \int f''(x)^2 dx.$$

Luego, la varianza se expresa como

$$\begin{aligned} \text{var} \hat{f}(x) &= n^{-1} \int h^{-2} K \{(x-y)h^{-1}\}^2 f(y) dy - n^{-1} \{f(x) + \text{sesgo}_h(x)\}^2 \\ &\approx n^{-1} h^{-1} \int f(x-th) K(t)^2 dt - n^{-1} \{f(x) + O(h^2)\}^2, \end{aligned}$$

usando que $y = x - ht$ en la integral, y la aproximación para el sesgo. Suponga que h es pequeña y n es grande, y expanda a $f(x - ht)$ como una serie de Taylor para obtener

$$\begin{aligned} \text{var} \hat{f}(x) &= n^{-1} h^{-1} \int \{f(x) - ht f'(x) + \dots\} K(t)^2 dt + O(n^{-1}) \\ &= n^{-1} h^{-1} f(x) \int K(t)^2 dt + O(n^{-1}) \\ &\approx n^{-1} h^{-1} f(x) \int K(t)^2 dt, \end{aligned}$$

donde $f(x)$ es una función de densidad de probabilidad. Integrando sobre x se da la aproximación

$$\int \text{var} \hat{f}(x) dx \approx n^{-1} h^{-1} f(x) \int K(t)^2 dt.$$

Las funciones de kernel univariado más utilizadas se presentan en la siguiente tabla.

Apéndice A. Estimadores kernel de densidad

Kernel	$K(t)$
Epanechnikov	$\frac{3}{4} \left(1 - \frac{1}{5}t^2\right) (\sqrt{5})^{-1}$ para $ t < \sqrt{5}$, 0 en otro caso
Biweight	$\frac{15}{16} (1 - t^2)^2$ para $ t < 1$, 0 en otro caso
Triangular	$1 - t $ para $ t < 1$, 0 en otro caso
Gausiana	$(2\pi)^{-\frac{1}{2}} \exp\left(-\frac{t^2}{2}\right)$
Rectangular	$\frac{1}{2}$ para $ t < 1$, 0 en otro caso
Triweight	$\frac{35}{32} (1 - t^2)^3$ para $ t < 1$, 0 en otro caso
Coseno	$\frac{\pi}{4} \cos \frac{\pi}{2}t$ para $ t < 1$, 0 en otro caso.

Apéndice B

Base de Datos

Para fines ilustrativos, en la tabla siguiente se presenta una muestra de la base de datos que se empleó en este trabajo. Se trata de los primeros 30 ciclos obtenidos del experimento.¹

¹La base de datos completa esta incluida en el CD que esta adjunto al final de este escrito.

Apéndice B. Base de Datos

Ciclo	Sorteo	Fecha	Hora	Exac 1	Lon1	Lat 1	Exac 2	Lon 2
1	1	30/09/03	16.2125	6	-101.2606	21.0418	13	-101.2607
2	1	30/09/03	17.2367	4	-101.2605	21.0418	5	-101.2605
3	2	30/09/03	18.3019	5	-101.2605	21.0418	6	-101.2605
4	3	30/09/03	19.0533	5	-101.2605	21.0418	6	-101.2605
5	2	02/10/03	17.0675	4	-101.2605	21.0418	5	-101.2605
6	1	03/10/03	12.4056	6	-101.2605	21.0418	6	-101.2606
7	3	03/10/03	13.1208	7	-101.2605	21.0418	6	-101.2605
8	1	03/10/03	14.3283	4	-101.2605	21.0418	5	-101.2605
9	2	03/10/03	15.3164	5	-101.2605	21.0418	5	-101.2605
10	1	03/10/03	16.1378	5	-101.2605	21.0419	7	-101.2605
11	3	03/10/03	17.1772	4	-101.2605	21.0419	4	-101.2606
12	2	03/10/03	18.1625	5	-101.2605	21.0418	4	-101.2605
13	3	03/10/03	19.1350	6	-101.2605	21.0418	7	-101.2606
14	3	04/10/03	14.5486	4	-101.2605	21.0418	5	-101.2605
15	2	04/10/03	15.6264	6	-101.2605	21.0418	9	-101.2605
16	1	04/10/03	16.3525	4	-101.2605	21.0418	6	-101.2605
17	3	04/10/03	17.2417	4	-101.2605	21.0418	6	-101.2605
18	1	04/10/03	18.0208	5	-101.2605	21.0418	6	-101.2605
19	1	07/10/03	13.0553	6	-101.2605	21.0418	5	-101.2605
20	3	07/10/03	14.9400	5	-101.2605	21.0418	4	-101.2606
21	1	07/10/03	15.6058	9	-101.2605	21.0418	8	-101.2606
22	2	08/10/03	13.3872	4	-101.2605	21.0418	5	-101.2605
23	3	08/10/03	14.2600	4	-101.2605	21.0418	5	-101.2605
24	1	08/10/03	15.6381	6	-101.2605	21.0418	5	-101.2605
25	1	08/10/03	16.5353	5	-101.2605	21.0418	6	-101.2605
26	3	08/10/03	17.6861	4	-101.2605	21.0418	5	-101.2605
27	2	10/10/03	12.7708	5	-101.2605	21.0418	7	-101.2605
28	2	10/10/03	13.5008	5	-101.2605	21.0418	5	-101.2605
29	1	10/10/03	14.2653	5	-101.2605	21.0418	5	-101.2605
30	3	10/10/03	15.7661	5	-101.2605	21.0418	5	-101.2606

Figura B.1: Muestra de la Base de Datos empleada en este trabajo. La muestra contiene 30 ciclos.

Apéndice B. Base de Datos

Lat 2	Exac 3	Lon 3	Lat 3	Distancia P1P2	Distancia P2P3	Distancia P3P1
21.0413	9	-101.2611	21.0417	40.6	62.5	64.4
21.0415	6	-101.2611	21.0418	33.7	67.8	66.1
21.0415	5	-101.2611	21.0417	34.9	66.5	64.5
21.0415	12	-101.2611	21.0418	33.8	72.0	63.2
21.0415	6	-101.2612	21.0418	37.2	68.5	71.0
21.0415	8	-101.2611	21.0418	34.9	65.4	64.7
21.0415	6	-101.2611	21.0418	35.1	67.9	64.6
21.0415	6	-101.2611	21.0418	36.7	71.4	64.7
21.0415	6	-101.2611	21.0418	38.1	78.7	65.1
21.0416	8	-101.2611	21.0418	35.0	64.1	64.7
21.0415	6	-101.2611	21.0418	43.1	68.8	64.9
21.0415	9	-101.2611	21.0418	35.3	72.1	63.3
21.0415	11	-101.2612	21.0417	35.4	72.6	68.9
21.0415	6	-101.2611	21.0418	36.5	69.5	64.1
21.0415	7	-101.2611	21.0418	37.5	69.5	64.8
21.0415	5	-101.2611	21.0418	34.2	68.3	60.5
21.0415	5	-101.2611	21.0418	35.2	72.2	64.7
21.0415	5	-101.2611	21.0418	35.6	72.5	61.0
21.0415	6	-101.2611	21.0418	34.5	61.2	64.6
21.0415	8	-101.2611	21.0418	35.0	69.9	63.9
21.0415	7	-101.2611	21.0418	35.3	63.7	62.6
21.0415	9	-101.2611	21.0418	38.2	72.1	60.4
21.0415	5	-101.2611	21.0418	35.0	69.9	63.4
21.0416	6	-101.2611	21.0417	35.6	68.2	62.8
21.0415	5	-101.2611	21.0418	34.8	67.5	69.0
21.0415	4	-101.2611	21.0418	35.4	70.8	64.0
21.0415	6	-101.2612	21.0417	34.7	68.6	73.7
21.0415	8	-101.2612	21.0417	35.9	68.7	68.4
21.0415	5	-101.2611	21.0418	35.6	69.9	63.2
21.0415	9	-101.2611	21.0418	39.4	66.1	60.6

Figura B.2: *Continuación.*

Apéndice B. Base de Datos

Lonutm 1	Latutm 1	Lonutm 2	Latutm 2	Lonutm 3	Latutm 3
265080.53407	2328299.11776	265069.43165	2328249.43562	265021.46555	2328288.87961
265089.19754	2328298.99497	265083.52864	2328265.84899	265023.30290	2328296.23720
265087.49100	2328300.86507	265081.79594	2328265.87355	265024.98326	2328292.52156
265085.75831	2328300.88962	265086.96787	2328263.95434	265023.35523	2328299.92827
265085.75831	2328300.88962	265081.76979	2328264.02801	265016.47678	2328303.71760
265087.43869	2328297.17399	265080.03709	2328264.05257	265023.32906	2328298.08273
265087.46484	2328299.01953	265081.84826	2328269.56462	265023.30290	2328296.23720
265087.51716	2328302.71060	265088.75288	2328267.62085	265023.32906	2328298.08273
265089.27600	2328304.53158	265090.45942	2328265.75076	265023.30290	2328296.23720
265091.06101	2328308.19810	265085.44443	2328278.74319	265026.87294	2328303.57022
265089.30216	2328306.37712	265080.08940	2328267.74364	265023.32906	2328298.08273
265085.73215	2328299.04409	265088.70057	2328263.92978	265023.30290	2328296.23720
265085.73215	2328299.04409	265080.03709	2328264.05257	265016.34596	2328294.48992
265085.75831	2328300.88962	265083.47632	2328262.15792	265021.57021	2328296.26176
265085.78446	2328302.73516	265083.50248	2328264.00345	265019.83751	2328296.28632
265085.73215	2328299.04409	265085.23518	2328263.97889	265025.03559	2328296.21263
265087.46484	2328299.01953	265086.94172	2328262.10880	265023.30290	2328296.23720
265085.73215	2328299.04409	265086.96787	2328263.95434	265021.57021	2328296.26176
265089.19754	2328298.99497	265081.87441	2328271.41016	265025.06175	2328298.05817
265085.75831	2328300.88962	265078.35671	2328267.76820	265021.59637	2328298.10730
265085.75831	2328300.88962	265076.54554	2328262.25615	265023.32906	2328298.08273
265087.49100	2328300.86507	265086.94172	2328262.10880	265026.79445	2328298.03361
265085.75831	2328300.88962	265081.76979	2328264.02801	265023.30290	2328296.23720
265087.51716	2328302.71060	265088.85750	2328275.00300	265025.00943	2328294.36710
265089.22369	2328300.84051	265083.58095	2328269.54006	265021.59637	2328298.10730
265085.75831	2328300.88962	265085.28749	2328267.66997	265023.32906	2328298.08273
265087.51716	2328302.71060	265083.55479	2328267.69452	265016.29363	2328290.79884
265084.07793	2328304.60526	265081.82210	2328267.71908	265016.31980	2328292.64438
265084.02562	2328300.91418	265083.55479	2328267.69452	265021.59637	2328298.10730
265087.49100	2328300.86507	265074.83900	2328264.12624	265026.76828	2328296.18807

Figura B.3: *Continuación.*

Bibliografía

- [1] **French, G. T.** (1996). *Understanding the GPS: An Introduction to the Global Positioning System What It Is and How It Works*, First Edition, GeoResearch.
- [2] **GARMIN** (2001). *eTrex Summit Personal Navigator (Owner's Manual and Reference Guide)*.
- [3] **Kennedy, M. and Kopp, S** (1994). *Understanding Map Projections*, GIS by ESRI.
- [4] **Silverman, B. W.** (1986). *Density Estimation for Statistics and Data Analysis, First Edition*, London: Chapman and Hall.
- [5] **Scott, D.W.** (1992). *Multivariate Density Estimation: Theory, Practice and Visualization*, New York: John Wiley & Sons.
- [6] **El Universal:** http://www2.eluniversal.com.mx/pls/impreso/web_histo_internacional.despliega?var=29495&var_sub_actual=a&var_fecha=30-OCT-03
- [7] **El Mundo:** <http://www.elmundo.es/elmundo/2003/10/29/ciencia/1067463129.html>