



UNIVERSIDAD DE JAÉN

**ESCUELA POLITÉCNICA SUPERIOR DE JAÉN
DEPARTAMENTO DE INGENIERÍA
CARTOGRÁFICA, GEODÉSICA Y
FOTOGRAMETRÍA**

TESIS DOCTORAL

**CONTROL DE CALIDAD POSICIONAL EN
CARTOGRAFÍA POR ELEMENTOS LINEALES**

**PRESENTADA POR:
ANTONIO TOMÁS MOZAS CALVACHE**

**DIRIGIDA POR:
DR. D. FRANCISCO JAVIER ARIZA LÓPEZ**

JAÉN, 15 DE OCTUBRE DE 2007

ISBN 978-84-8439-404-4

Nombre y apellidos del autor:

ANTONIO TOMÁS MOZAS CALVACHE

Título de la Tesis Doctoral:

CONTROL DE CALIDAD POSICIONAL EN CARTOGRAFÍA POR ELEMENTOS LINEALES

I.S.B.N.:

978-84-8439-404-4

Fecha de Lectura:

15 DE OCTUBRE DE 2007

Centro y Departamento en que fue realizada la lectura:

ESCUELA POLITÉCNICA SUPERIOR DE JAÉN

Departamento de Ingeniería Cartográfica, Geodésica y Fotogrametría

Composición del Tribunal/Dirección de la Tesis:

Dirección de la Tesis Dr. D. Francisco Javier Ariza López

Presidente/a del Tribunal Dr. D. Manuel Sánchez de la Orden

Vocales Dr. D. José L. Berné Valero

Dr. D. Sebastián Mas Mayoral

Dr. D. Alan D. J. Atkinson Gordo

Secretario/a Dr. D. Tomás Fernández del Castillo

Calificación Obtenida:

SOBRESALIENTE CUM LAUDE



UNIVERSIDAD DE JAÉN

tesis doctoral



Resumen

Tradicionalmente, el control de la calidad posicional de los productos cartográficos se ha venido realizando mediante el uso de técnicas basadas en puntos. Sin embargo, las nuevas tecnologías han permitido proponer nuevas formas de control, como es el uso de elementos lineales. Hasta la fecha, se han presentado algunas propuestas metodológicas sobre este tema que se han limitado al desarrollo de una metodología básica muy concreta y su ensayo con pequeños conjuntos de datos.

En esta tesis doctoral, se propone una metodología completa para la realización de dicho control, que engloba los métodos propuestos, analizándolos y complementándolos. Para ello, se presentan las principales decisiones adoptadas en el control posicional por elementos lineales, utilizando para ello herramientas como el control automatizado de los mismos, la simulación, la caracterización y similitud entre funciones de distribución, etc. aplicándolas, por un lado, a dos bases de datos cartográficas reales, cuyas vías de comunicación han sido controladas con fuentes de mayor exactitud (GPS), y por otro lado, un conjunto de líneas sintéticas con errores simulados para estudiar su comportamiento individualizado.

Así mismo, con los resultados obtenidos se ha estudiado la muestra a emplear obteniendo importantes conclusiones sobre el tamaño, caracterización y densidad de puntos de elementos lineales a observar.

Abstract

Usually, the positional quality control of cartographic products is developed by means of control points. Nowadays, new technologies have raised the possibility of using linear elements for this positional quality control. There have been several previous studies which presents singular methodologies applied to a reduced group of data. These methods have been developed independently without been tested together.

In this PhD, the general methodology to develop this control is presented, using and complementing previous methodologies. We also present the main decisions taken about this control, using tools like an automatic control, simulations, characterizations and the similarity of distribution functions. These decisions have been achieved by means of two real cartographic databases, whose roads have been surveyed using GPS cinematic observations, and with a series of synthetic lines with simulated errors to study their particular behaviour.

From the results' analysis we have studied the sample of lines to use, and obtained interesting conclusions about the size, category and vertexes density of the linear elements to observe.



UNIVERSIDAD DE JAÉN

CONTROL DE CALIDAD POSICIONAL EN CARTOGRAFÍA POR ELEMENTOS LINEALES

ANTONIO TOMÁS MOZAS CALVACHE

tesis doctoral

*Nunca vayas por el camino trazado
porque sólo conduce hacia donde otros han ido ya.*

Alexander Graham Bell

Agradecimientos

En primer lugar deseo mostrar mi agradecimiento al director de la tesis doctoral, el profesor Francisco Javier Ariza, por el apoyo que me ha prestado para su realización y la confianza mostrada al apostar por mi en este proyecto.

También es justo, agradecer a todos aquellos que me han apoyado en su realización, y en especial:

A mis compañeros del Departamento de Ingeniería Cartográfica, Geodésica y Fotogrametría de la Universidad de Jaén en general, y más concretamente, a mis amigos Manuel Antonio Ureña, José Luis García, Juan José Ruíz y José Luis Pérez por su colaboración y apoyo.

A la profesora Virtudes Alba del Departamento de Estadística e Investigación Operativa de la Universidad de Jaén y el profesor Alan D. J. Atkinson del Departamento de Expresión Gráfica de la Universidad de Extremadura por su colaboración.

A mis amigos y familiares que han soportado mis malos momentos.

Por último, quiero agradecer muy especialmente a mis padres y a Paqui por su paciencia, cariño y apoyo.

Índice

Índice

Capítulo I. Introducción	10
I.1. Hipótesis de partida y objetivos	14
I.2. Organización del documento	15
I.3. Proyecto de investigación asociado	15
Capítulo II. Antecedentes	17
II.1. Calidad en cartografía	17
II.2. Exactitud Posicional	20
II.3. Requerimientos de exactitud	22
II.4. Exactitud posicional basada en elementos lineales	23
II.4.1. Banda épsilon, de error y Banda-g	23
II.4.2. Métodos de control de exactitud posicional en elementos lineales	30
II.4.2.1 Método de banda épsilon	30
II.4.2.2 Método del orlado simple	33
II.4.2.3 Método del doble orlado	37
II.4.2.4 Método basado en las distancias de Hausdorff	44
II.4.2.5 Método MPS y MDS de Veregin	47
II.4.2.6 Método de correspondencia de puntos	50
II.4.3. Errores de orlados	52
II.5. Muestreo	55
II.5.1. Muestreo aleatorio simple	56
II.5.2. Muestreo estratificado	57
II.5.3. Muestreo por conglomerados	59
II.6. Variables aleatorias y distribuciones de probabilidad	60

Capítulo III. Metodología y Material.....	66
III.1. Introducción	67
III.2. Procedimiento General.....	68
III.3. Elementos Lineales	72
III.3.1. Caracterización	72
III.3.2. Zona de estudio	73
III.3.3. Diseño de muestras	73
III.3.4. Conclusiones	74
III.4. Generación de Bases de Datos de elementos lineales.....	74
III.4.1. Obtención de las bases de datos a controlar	74
III.4.2. Obtención de las bases de datos de control	75
III.4.3. Depuración, comparación y edición de las bases de datos.....	76
III.4.4. Tramificación	77
III.4.5. Enriquecimiento	77
III.4.6. Identificación de elementos homólogos	78
III.4.7. Eliminación de información redundante.....	79
III.4.8. Selección de elementos	80
III.4.9. Conclusiones	81
III.5. Control automatizado por elementos lineales.....	81
III.5.1. Método de las Distancias de Hausdorff (MDH)	83
III.5.2. Método de la Banda Épsilon o Skidmore (MSK).....	84
III.5.3. Método del Orlado Simple (MOS)	86
III.5.4. Método del Orlado Doble (MOD).....	88
III.5.5. Conclusiones	90
III.6. Caracterización y similitud entre funciones de distribución	91
III.7. Determinación del tamaño de la muestra	94
III.8. Aplicación de los métodos de control a líneas sintéticas.....	95
III.9. Aplicación a bases cartográficas	97
III.9.1. Material utilizado	97
III.9.2. Zona de estudio	99
III.9.3. Elementos de control (GPS).....	94
III.9.4. Conclusiones	98

Capítulo IV. Resultados y Discusión	110
IV.1. Introducción	111
IV.2. La herramienta CPLin	112
IV.3. Resultados obtenidos en la aplicación a líneas sintéticas	114
IV.3.1. Resultados del Método de Distancias de Hausdorff (MDH)	114
IV.3.2. Resultados del Método de la Banda Épsilon o Skidmore (MSK).....	115
IV.3.3. Resultados del Método de Orlado Simple (MOS)	116
IV.3.4. Resultados del Método de Orlado Doble (MOD).....	119
IV.3.4.1 Superficie Común	119
IV.3.4.2 Desplazamiento Promedio	122
IV.3.4.3 Oscilación	124
IV.3.5. Interpretación y análisis	127
IV.4. Aplicación a bases de datos cartográficas. Caracterización de elementos lineales.....	127
IV.4.1. Longitud del tramo.....	131
IV.4.2. Zona de Trabajo.....	132
IV.4.3. Tipo de vía	133
IV.4.4. Orientación.....	134
IV.4.5. Sinuosidad	136
IV.4.6. Interpretación y análisis	138
IV.5. Resultados obtenidos con la aplicación a bases de datos cartográficas	138
IV.5.1. Resultados del Método de las Distancias de Hausdorff (MDH)	138
IV.5.2. Resultados del Método de la Banda Épsilon o Skidmore (MSK).....	141
IV.5.3. Resultados del Método del Orlado Simple (MOS)	144
IV.5.4. Resultados del Método del Orlado Doble (MOD).....	147
IV.5.5. Interpretación y análisis	150
IV.6. Determinación de la muestra	152
IV.6.1. Criterios para la determinación de la muestra	152
IV.6.1.1 Zona de trabajo	154
IV.6.1.2 Tipo de vía	156
IV.6.1.3 Longitud de tramo	157
IV.6.1.4 Orientación	159

IV.6.1.5 Sinuosidad	161
IV.6.1.6 Interpretación y análisis	163
IV.6.2. Estimación del tamaño de la muestra	163
IV.6.2.1 Método MDH.....	164
IV.6.2.2 Método MSK.....	165
IV.6.2.3 Método MOS	166
IV.6.2.4 Método MOD	167
IV.6.2.5 Pruebas con el tamaño de muestra	169
IV.6.2.6 Interpretación y análisis	172
IV.6.3. Determinación de la densidad de vértices de control	172
IV.6.3.1 Método MDH.....	173
IV.6.3.2 Método MSK.....	174
IV.6.3.3 Método MOS	174
IV.6.3.4 Método MOD	176
IV.6.3.5 Interpretación y análisis	178
IV.7. Conclusiones	178
 Capítulo V. Conclusiones	 180
V.1. Objetivos alcanzados	181
V.2. Conclusiones acerca de los resultados obtenidos	182
V.2.1. Metodología	182
V.2.2. Aplicación metodológica	183
V.3. Perspectivas de futuro	184
 Referencias	 186
 Anexos	 194
Anexo I. Test de exactitud posicional basados en puntos	195
AI.1. Introducción	195
AI.2. Test NMAS.....	196
AI.3. Test EMAS	198
AI.4. Test ASPRS.....	202

AI.5. Test NSSDA	204
AI.6. Test de la fórmula de Koppe	205
AI.7. Test del USGS de categorías de exactitud de MDE.....	207
Anexo II. Especificaciones de las BDGs	209
AII.1. Mapa Topográfico de Andalucía	210
AII.2. Base Cartográfica Numérica.....	211
Anexo III. Ajuste a distribuciones de probabilidad	215
AIII.1. Método MOS	216
AIII.2. Método MOD. Superficie Común.....	217
AIII.3. Método MOD. Desplazamiento promedio.....	218
Anexo IV. Tablas de Resultados	219
AIV.1. Resultados del Método MDH. Base de datos MTA10	220
AIV.2. Resultados del Método MDH. Base de datos BCN25	220
AIV.3. Resultados del Método MSK. Base de datos MTA10	221
AIV.4. Resultados del Método MSK. Base de datos BCN25	221
AIV.5. Resultados del Método MOS. Base de datos MTA10	222
AIV.6. Resultados del Método MOS. Base de datos BCN25	224
AIV.7. Resultados del Método MOD. Base de datos MTA10	226
AIV.8. Resultados del Método MOD. Base de datos BCN25	229
Índices	232
Índice de Figuras	233
Índice de Tablas	240
Acrónimos	243

Capítulo I.

Introducción

Capítulo I. Introducción

El desarrollo actual de la cartografía y la creciente demanda de los productos cartográficos digitales por parte de los usuarios son consecuencia lógica de la evolución de las nuevas tecnologías de la información.

El triunfo de estos productos digitales sobre los tradicionales está basado en la mejora de las prestaciones, la universalización de los productos ofertados y la reducción de costes tanto para el productor como para el usuario. Ejemplos de esta espectacular expansión son la aparición de las infraestructuras de datos espaciales, integradores a través de Internet de datos, metadatos, servicios e información geográfica, la cartografía asociada a navegadores G.P.S. cada vez más difundidos entre la población o aplicaciones de Internet como callejeros o visualizadores (Google Earth).

Estas nuevas tecnologías han irrumpido con fuerza en todas las fases del proceso cartográfico abarcando desde la toma de datos, la producción cartográfica, el soporte y la difusión del producto en sí, etc.

Actualmente, se esta produciendo una demanda cartográfica sin precedentes, la cual exige, simultáneamente de los productores un aumento de calidad de la cartografía.

La calidad es definida por la Real Academia Española como la propiedad o conjunto de propiedades inherentes a algo, que permiten juzgar su valor. El concepto de calidad ha variado a lo largo del tiempo pasando de la adecuación de un producto a

las especificaciones establecidas en el principio de la industrialización a la definición más aceptada y completa que actualmente se refiere como a la totalidad de las características de un producto o servicio tal que le confieren su aptitud para satisfacer unas necesidades explícitas e implícitas. De este modo, se define la calidad de un producto o servicio como su aptitud o idoneidad para el uso. La norma ISO 9000:2000 adopta la calidad como el grado en el que un conjunto de características inherentes cumple con los requisitos (necesidades o expectativas establecidas, generalmente implícitas u obligatorias) (Ariza, 2002a).

La calidad en la cartografía no puede ser descrita por un índice unívoco, sino que resulta necesario que cada una de las componentes de los datos geográficos tenga asociada un valor de su bondad o calidad (Shi, 1994). Estas componentes son:

- Posición espacial (x, y, z)
- Atributos ($a_1, a_2, \dots$)
- Tiempo en el que suceden ($t_1, t_2, \dots$)

La tesis doctoral que este documento presenta abordará el aspecto posicional de esta calidad y más concretamente la componente planimétrica (x, y). Las características propias del dato geográfico hacen que se trate de forma independiente las componentes planimétrica y altimétrica. Según Atkinson (2005) esto se debe a:

- Los métodos empleados en la captura difieren sustancialmente.
- Generalmente, la incertidumbre en la captura de la información altimétrica es superior a la planimétrica.
- Tradicionalmente, la altimetría se ha “representado” mediante simbología, debiendo “leerse” en vez de medirse. Por ejemplo, para obtener el valor de la componente altimétrica entre dos curvas de nivel habrá que realizar una interpolación en función de la distancia planimétrica del punto de interés suponiendo que el terreno es constante y homogéneo entre ambas curvas. Por el contrario, la planimetría se puede “medir” al representarse en su totalidad (el nivel de representación vendrá dado por la escala, el tipo de cartografía y la utilización que se haga de la información).

El aspecto posicional de los productos cartográficos es una de sus componentes de calidad más importantes y por ello queda recogido como un elemento de la calidad en la reciente normativa ISO 19113 (ISO, 2002). Esta componente se convierte en un elemento clave para la interoperabilidad de las bases de datos (Church et al., 1998), más aún si pensamos en los actuales avances en las Infraestructuras de Datos Espaciales (IDE) (Mas, 2002).

El control de la calidad posicional de los productos cartográficos se ha venido realizando mediante la estimación de la exactitud posicional de los mismos usando técnicas basadas en puntos. Estas técnicas o métodos de control son procesos de base estadística en los que a través de una muestra reducida de puntos bien identificados en la cartografía y en el terreno se realiza una estimación para todo el producto.

Existen numerosas metodologías para la estimación o control de la exactitud posicional por elementos puntuales. Las más conocidas se describen en el Anexo I.

Estas metodologías puntuales, se basan en disponer de dos conjuntos de puntos homólogos (cartografía-terreno) que se encuentren perfectamente definidos e identificados. Las técnicas de control de la exactitud posicional por elementos puntuales fueron desarrolladas debido a que los métodos de trabajo y levantamiento clásicos se centran en la obtención de puntos individuales y de relevancia sobre el terreno. Sin embargo, la aparición de la tecnología G.P.S. permite la posibilidad de levantar curvas con un muestreo suficientemente denso como para ser consideradas continuas a la resolución de la Base de Datos Geográfica. Debido a esta ventaja, asociada a que los elementos lineales se conforman como el conjunto más numeroso dentro de una base de datos geográfica –según Cuenin (1972) cerca del 80%-, ha hecho reflexionar en los últimos años a la comunidad científica sobre la posibilidad de uso de estos elementos en el control posicional. De esta forma han surgido diversas propuestas que serán analizadas en el capítulo II de este documento.

La alternativa basada en elementos lineales supone el uso de una geometría más compleja que la puntual en la que las líneas no pueden ser reducidas a puntos. Esto es así dado que, si bien puede realizarse un case de elementos homólogos a nivel de líneas, los puntos que en secuencia las conforman no están bien definidos, ni en el terreno, ni sobre las propias líneas. La fuerza del método de evaluación lineal proviene justamente del uso de una geometría lineal basada en muchos puntos de calidades posicionales inferiores a las que podrían obtenerse sobre un conjunto reducido de puntos bien definidos. En este caso, se prefiere cantidad y forma a unos pocos puntos de mayor calidad posicional, esperando con ello mejor representatividad del hecho posicional a evaluar y buscando igualmente agilidad en la captura.

La aplicación práctica de cualquiera de las propuestas de control posicional por elementos lineales requiere de manera indiscutible disponer de herramientas de software específicas y adecuadas para poder llevarlas a cabo. Estas herramientas deben disponer de capacidades de emparejamiento de elementos lineales homólogos, así como para derivar la medida, o medidas, que requiera la metodología a aplicar. En este documento se describirán las bases computacionales para el desarrollo de herramientas de control posicional de la cartografía mediante elementos lineales.

Indiscutiblemente estos planteamientos deben estar sustentados en un estudio con datos geográficos reales. La aplicación de las diferentes metodologías de control posicional por elementos lineales se realiza sobre dos bases de datos geográficas, utilizando para dicho control datos de elementos lineales, en nuestro caso carreteras, obtenidos de una observación G.P.S.. A lo largo del documento también se plantea cómo realizar la observación de estos elementos, los errores cometidos, su proceso de obtención, los criterios para el muestreo de datos, la edición de las observaciones y de las bases de datos a controlar y la materialización de un software para la realización y estudio del control posicional.

I.1. Hipótesis de partida y objetivos

Tradicionalmente se ha realizado el control posicional de la cartografía mediante la comparación de elementos puntuales. En este trabajo se propone realizar este estudio mediante elementos lineales.

La hipótesis inicial de trabajo consiste en la posibilidad de realizar un control posicional de la cartografía más preciso, más barato, más riguroso y comparable; mediante el uso de elementos lineales.

Para ello se emplea:

- Cartografía vectorial disponible en formato digital.
- Datos de elementos lineales obtenidos mediante la observación con G.P.S.
- Mecanismos de simulación y programación.

El método a emplear consiste en:

- Revisión bibliográfica: Con el fin de conocer el estado de la cuestión y de obtener el punto de partida de la investigación y diferenciarla de otras existentes dentro de este ámbito.
- Obtención de datos: Se obtendrán los elementos lineales con fuentes de mayor exactitud a la cartografía a estudiar.
- Diseño: Se crearán algoritmos que determinen un valor de calidad posicional global de la cartografía, empleando algoritmos propios o basados en la estadística.
- Análisis: Se realizará un análisis, estudiando las posibles casuísticas que podamos encontrar, etc.
- Resultados: Obtención de los resultados y conclusiones del proceso proponiendo un método o métodos de obtención de la exactitud posicional en cartografía a partir de elementos lineales, así como sus ventajas e inconvenientes para su posterior empleo.

El objetivo fundamental de esta tesis doctoral es el control de calidad posicional en cartografía por elementos lineales, desarrollando una metodología que pueda complementar a los actuales métodos de control basados en elementos puntuales.

La investigación se lleva a cabo en las siguientes líneas:

- Análisis y desarrollo de las medidas evaluadoras de la exactitud posicional en cartografía por elementos lineales.
- Desarrollo de mecanismos para la obtención de la exactitud posicional en cartografía para aplicarlos a cualquier base geográfica en entorno vectorial a partir de los valores obtenidos con fuentes de mayor exactitud, comprobando los resultados.

- Desarrollo de sistemas que permitan obtener un valor global de exactitud posicional en cartografía.

Para el desarrollo de estas líneas se emplearán algoritmos propios y otros obtenidos en la revisión. Como fuente de datos se utiliza cartografía convencional a diferentes escalas y datos sobre elementos lineales obtenidos de fuentes de mayor exactitud (GPS).

I.2. Organización del documento

El presente documento se encuentra dividido en 5 capítulos.

- Capítulo I: Se corresponde con la presente introducción.
- Capítulo II. Antecedentes: Este apartado permite conocer el estado de la cuestión. Se profundiza en los conceptos e investigaciones previas en las que se apoya el estudio objeto de esta tesis doctoral.
- Capítulo III. Material y metodología: Se describe el material y datos utilizados, la obtención de los mismos, el muestreo realizado y la metodología propuesta para la obtención de la calidad posicional mediante elementos lineales.
- Capítulo IV. Resultados y discusión: Junto a los principales resultados obtenidos se presentan la interpretación y el análisis derivado de los mismos.
- Capítulo V. Conclusiones: Se finaliza el cuerpo principal de este documento con las conclusiones del trabajo.
- Posteriormente se presentan las referencias bibliográficas utilizadas y los anexos que complementan otras partes del documento.

I.3. Proyecto de investigación asociado

La investigación que se desarrolla en esta tesis doctoral se enmarca dentro del proyecto de investigación “Control de calidad posicional de Cartografía por elementos lineales (ConPoCar)” del Ministerio de Educación y Ciencia (BIA2003-02234), del cual es el director de esta tesis doctoral el investigador responsable.

Capítulo II.

Antecedentes

Capítulo II. Antecedentes

En el capítulo anterior se hace una breve introducción sobre el concepto de calidad, que en este apartado es desarrollada con mayor profundidad. Como se ha comentado, nuestro trabajo de investigación se centra en el control posicional por elementos lineales. Esta novedosa metodología se presenta frente a la tradicional basada en elementos puntuales. En este capítulo se describen los principales estándares y métodos para el control de calidad posicional en cartografía, comenzando con los métodos tradicionales basados en puntos (Anexo I) para continuar con los métodos que utilizan elementos lineales.

II.1. Calidad en cartografía

Como se ha comentado anteriormente, el dato geográfico se caracteriza por varias componentes: posicional, de atributos y temporal. Una base de datos geoespacial (en adelante BDG) es un modelo de la realidad, por lo que es interesante que este modelo sea lo más exacto posible, aunque en su proceso de obtención aparezcan fuentes de imprecisión, a lo que se debe añadir los errores incorporados tras el complejo proceso de manipulación e integración de datos de muy diversa procedencia (Ariza, 2002a).

El dato geográfico debe entenderse como un vector en el espacio de las componentes que lo definen, extrapolando las elipses de error o de incertidumbre típicas del aspecto posicional al resto de componentes (Veregin, 1994) (Figura II. 1).

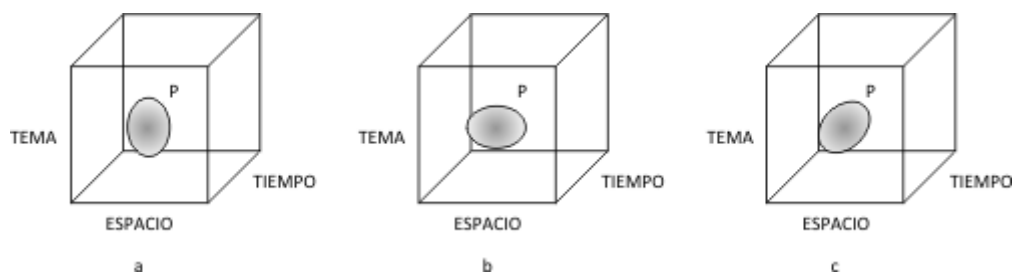


Figura II. 1. Elipsoide de inexactitud alrededor de un punto: a) inexactitud temática dominante, b) inexactitud espacial dominante, c) inexactitud temporal dominante

Según Ariza (2002a), calidad y error son dos conceptos íntimamente ligados, aunque el primero sea algo más relativo y utilitario. Por ejemplo, unos datos de mala calidad para un trabajo de precisión pueden resultar aceptables para otro con menores exigencias. Es aquí donde reaparece el concepto de “idoneidad para el uso” inherente al de calidad. El error es un aspecto complementario al de calidad con la salvedad de que ésta se relaciona con el uso, mientras que el error lo hace con la realidad (diferencia con el valor real).

En general, se puede hablar de errores groseros, sistemáticos y accidentales.

- Error grosero: Equivocación evitable. Su control y detección ha de derivarse del método de trabajo utilizado.
- Error sistemático: Se genera por causas permanentes, por lo que puede eliminarse si se determina su fuente.
- Error accidental: Se genera por un fallo aleatorio, por lo que no presenta una ley fija. Se compensan parcialmente al incrementar el número de observaciones o medidas.

Otro aspecto fundamental a tener en cuenta es la diferencia entre exactitud y precisión. Diversos autores asumen la exactitud como la aproximación de los resultados respecto a los considerados como verdaderos. Esta consideración se realiza ante el desconocimiento del valor verdadero, por lo que se relaciona con el obtenido de una media de observaciones o de una fuente de mayor fiabilidad. La precisión, estadísticamente, es una medida del agrupamiento de los datos obtenidos por repetición de una medida. Por tanto, exactitud y precisión son dos términos totalmente distintos: Una medida puede ser muy precisa debido a que todas las observaciones realizadas tengan un grado de dispersión muy pequeño y sin embargo estar muy alejadas del valor real (exactitud). Desde el punto de vista informático (Burrough, 1986; Morant y Lerma, 1999) la precisión puede entenderse como el número de cifras significativas que soporta un tipo de dato (entero, real corto, real largo, etc.).

Según Veregin (1999), una de las principales características del dato geográfico es la resolución, existiendo tantas resoluciones como dimensiones o componentes:

- Resolución espacial: Tamaño del menor elemento que puede ser distinguido o separado de sus partes constituyentes o del conjunto al

que pertenecen. En un sistema vectorial se refiere a la mínima unidad cartografiada y en un sistema raster lo hace a la dimensión de la celda.

- Resolución temática: Se distingue en función de la naturaleza cualitativa o cuantitativa del dato. En el primero se refiere a la adecuación o precisión de la definición de las clases, mientras que en caso de ser cuantitativa su definición es similar a resolución espacial.
- Resolución temporal: Se corresponde con la duración temporal del muestreo que deberá ser inferior a la del fenómeno a tomar.

Los componentes básicos de la calidad de los datos considerados por varios autores e instituciones (NCDCCS, 1988; USGS, 1994; Morrison, 1995; Ariza, 2002a) son:

- Exactitud posicional: Proximidad entre las coordenadas dadas y la reales.
- Exactitud del atributo: Similar a la anterior pero referida al valor del atributo.
- Exactitud temporal: Adecuación de la fecha de la BDG a la del análisis concreto. Normalmente se considera como la actualidad de los datos.
- Consistencia lógica: Trata de las reglas lógicas que ha de cumplir la estructura de los datos, sus atributos, sus formatos físicos. Por ejemplo, en una estructura topológica, las líneas deben empezar y acabar en nodos, las intersecciones entre líneas deben ser nodos, etc.
- Compleción: Mide lo completa que es la base de datos. La selección de elementos depende de la selección temática, así como de la generalización.
- Genealogía o linaje: Está referido a las fuentes, procesos de captura, métodos de análisis, sistemas de referencia, parámetros de transformación, resolución de los datos, etc.

Se pueden considerar diferentes formas de realizar un seguimiento de la exactitud:

- Estimaciones de experto: Aportan conclusiones orientativas y subjetivas.
- Estudio de la coherencia lógica.
- Comparación con las fuentes de las cuales ha sido obtenido.
- Comparación con fuentes de mayor exactitud: Consiste en el control de los datos con fuentes más exactas o con trabajos de campo confrontando estadísticamente la BDG con la fuente de mayor exactitud.

II.2. Exactitud Posicional

Por exactitud posicional se entiende la diferencia en la localización de las coordenadas geográficas de un cierto objeto sobre la superficie terrestre y su representación en la base de datos (Giordano y Veregin, 1994).

Esta definición se complementa con la aportada por Ariza (2002a) que delimita la exactitud posicional como el grado de conformidad con que se representan las posiciones horizontales y verticales de los objetos de una BDG en relación a la posición verdadera de los mismos sobre el terreno.

Un cierto grado de imprecisión en el dato geográfico es siempre inevitable. La exactitud posicional, como componente de la calidad cartográfica, hace referencia a exactitud planimétrica y altimétrica. Los requerimientos de exactitud de un producto están relacionados íntimamente con su uso (calidad como idoneidad para el uso).

Tradicionalmente, la exactitud posicional de una base de datos geográfica, se ha determinado mediante el análisis de una muestra significativa de puntos. A lo largo del Anexo I se plantean los métodos más importantes.

La utilización de esta metodología basada en puntos, que se ha venido usando por numerosos productores de cartografía hasta el momento, presenta en la actualidad una alternativa con el uso de elementos lineales. Esta novedad surge de la inquietud de la comunidad científica y la investigación realizada en los últimos años sobre cómo se comporta los errores posicionales en estos elementos y por otro lado del desarrollo de nuevas metodologías más precisas en la toma de estos elementos (toma de carreteras con G.P.S.).

La exactitud posicional puede ser caracterizada por dos parámetros como precisión y desplazamiento. El primero se corresponde con la dispersión en los errores, describiendo el grado de conformidad intrínseco de unos datos. El desplazamiento es una discrepancia sistemática media entre las posiciones representadas y las verdaderas.

Según el DOD (1990) se pueden definir los siguientes conceptos referidos a la exactitud posicional de aplicación en cartografía:

- Exactitud absoluta: Estadístico que da la incertidumbre de un punto respecto a un datum. Esta definición implica la necesidad de considerar los efectos de todas las fuentes de error tanto sistemáticas como aleatorias. La exactitud absoluta se establece en términos de dos componentes: horizontal y vertical.
- Exactitud relativa: Estadístico que da la incertidumbre de la posición de dos puntos después de haber eliminado los efectos de los errores comunes a ambos. La exactitud relativa es independiente del datum del producto ya que se define como el error en las componentes del vector entre dos puntos. También se define en componentes horizontal y vertical.
- Exactitud absoluta horizontal: Evaluación estadística de todos los errores sistemáticos y aleatorios encontrados en la determinación de la

posición horizontal de un punto respecto a un datum especificado. Se suele expresar como error circular al 90% de probabilidad.

- Exactitud absoluta vertical: Evaluación estadística de los errores sistemáticos y aleatorios encontrados en la determinación de la altitud de un punto respecto al nivel medio del mar. Se suele expresar como error lineal al 90% de probabilidad.
- Exactitud relativa horizontal: Evaluación estadística de todos los errores aleatorios encontrados en la determinación de la posición horizontal de un dato horizontal respecto a otro. Se expresa como un error circular sobre una distancia con el 90% de probabilidad.
- Exactitud relativa vertical: Evaluación estadística de todos los errores aleatorios encontrados en la determinación de la elevación de un dato horizontal respecto a otro. Se expresa como un error circular sobre una distancia con el 90% de probabilidad.

Según Caspary y Scheuring (1993) la exactitud planimétrica puede derivarse de la raíz del error cuadrático medio (RMSE) que presentan las coordenadas (X_m , Y_m) de un punto respecto a sus coordenadas sobre el terreno (X_t , Y_t). El par de coordenadas debe estar referido al mismo datum y proyección. Las coordenadas del terreno tienen que ser obtenidas por métodos precisos y tener errores inferiores a los que exija la cartografía. La ecuación Ec. II. 1 presenta la raíz del error cuadrático medio para la coordenada "X" de una muestra de "n" datos.

$$RMSE_X = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_{mi} - X_{ti})^2}$$

Ec. II. 1

Análogamente se obtiene el valor para la componente "Y".

El RMSE planimétrico vendrá dado por la componente cuadrática del RMSE de la coordenada "X" y de la coordenada "Y" (Ec. II. 2).

$$RMSE = \sqrt{RMSE_X^2 + RMSE_Y^2}$$

Ec. II. 2

Algunos autores como Caspary y Scheuring (1993) afirman que el uso del RMSE es una simplificación del error posicional. En este contexto, surge la definición del error posicional como la combinación de las probabilidades lineales en dos dimensiones. El error sigue una distribución normal circular, por lo que para un punto en una posición (X_p, Y_p), la probabilidad de cualquier otro punto (x,y) de estar sobre la verdadera posición viene dada por:

$$f(x, y) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{(x - X_p)^2 + (y - Y_p)^2}{2\sigma^2}\right]$$

Ec. II. 3

De esta forma se presentan algunas medidas de errores circulares y sus probabilidades correspondientes (Maling, 1989):

Nombre	Probabilidad	Desviación
Error circular estándar (CSE)	0.3935	1.0 σ
Error circular probable (CPE)	0.5	1.1774
Error circular medio (CMSPE)	0.6321	1.4142 σ
Estándar de exactitud circular del mapa (CMAS)	0.9	2.1460 σ
Error 3.5 sigma	0.9978	3.5 σ

Tabla II. 1. Definiciones de errores circulares. Probabilidades

Estos errores se relacionan con las componentes unidimensionales mediante:

$$\begin{aligned} CSE &= \pm 0.5 (\sigma_x + \sigma_y) \\ CPE &= \pm 0.59 (\sigma_x + \sigma_y) \\ CMAS &= \pm 1.073 (\sigma_x + \sigma_y) \end{aligned}$$

Ec. II. 4

II.3. Requerimientos de exactitud

Los criterios de exactitud se establecen normalmente en función de la escala de representación. Según López-Cuervo (1996), el límite de percepción del ojo humano es de 0.25 milímetros, con un error no superior a 0.2 milímetros. Por tanto el error acumulado que tiene un punto no debe superar este umbral trasladado a la escala del producto. Por tanto, para una escala de módulo "M", el umbral de exactitud posicional es 0.2 x M. Estadísticamente, se debe tomar una tolerancia en función del grado de confianza con el que se desee tener los puntos del producto (Ariza, 2002a). En la Tabla II. 2 se indican las desviaciones tolerables para que el 90% de los puntos se encuentren a una distancia inferior a 0.43 milímetros de su posición correcta en el mapa (Sevilla, 1991):

Escala del mapa	Desviación Típica (metros)
1:10.000	2.5
1:5.000	1.25
1:2000	0.5
1:1000	0.25
1:500	0.125
1:200	0.05
1:100	0.025

Tabla II. 2. Precisiones planimétricas usuales (Sevilla, 1991)

II.4. Exactitud posicional basada en elementos lineales

La presente tesis doctoral se basa fundamentalmente en el estudio de la exactitud posicional por elementos lineales. Tradicionalmente, se ha utilizado para este menester métodos y estándares basados en elementos puntuales. Estas metodologías han sido descritas y analizadas en numerosos trabajos como los de Veregin (1989b), Giordano y Veregin (1994), Ariza (2002a), Atkinson (2005) y Nero (2005) por lo que se incluye un resumen de las mismas en el Anexo I.

En el apartado que nos ocupa, se plantea un recorrido por los principales métodos que utilizan elementos lineales y que son fundamentales en el estudio realizado en esta tesis doctoral.

De forma preliminar, se introducirá el concepto de banda épsilon, banda de error y *Banda-g* para continuar con la presentación de los distintos métodos. Esta descripción se realiza de forma similar a la utilizada para los test de elementos puntuales y está basada en la estructura usada por Giordano y Veregin (1994) y ampliada por Ariza (2002a) y Atkinson y Ariza (2002).

II.4.1. Banda épsilon, de error y Banda-g

El concepto de banda épsilon, planteado por Perkal (1956) ha sido utilizado para describir la inexactitud de la posición de líneas digitalizadas, donde la posición verdadera de una línea se presenta dentro de un cierto desplazamiento (*épsilon*) de la posición medida (Zhang y Goodchild, 2002). Es definida para un segmento cualquiera, por las dos líneas paralelas a la línea más probable y tangentes a los círculos de error de los puntos inicial y final del segmento (Figura II. 2).

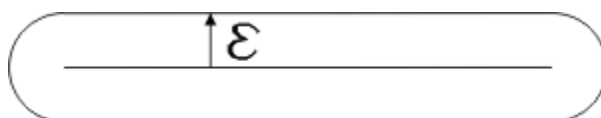


Figura II. 2. Banda Épsilon

La anchura de esta banda de indeterminación será, por tanto, el doble del error (ε) (Figura II. 2). El área contenida dentro de la banda de indeterminación viene dada por:

$$\text{Área} = 2L\sigma_p + \pi\sigma_p^2$$

Ec. II. 5

donde L es la longitud del segmento y σ_p es el RMSE en los vértices.

Según Giordano y Veregin (1994), la banda ε es un tipo de distribución centrado sobre la línea. Se asume que la probabilidad de la localización real de la línea sobre el terreno contenida dentro de la banda es igual a 1. La amplitud de la banda está determinada por el valor de ε , que define el grado de incertidumbre existente en la localización real de la línea. De esta forma, cuanto mayor es el valor de ε , mayor es la incertidumbre.

La banda ε puede ser utilizada para expresar una tolerancia específica en las distintas operaciones realizadas en un SIG, como las operaciones de *overlay*, entendiendo que el valor de ε pueda variar de un objeto a otro (Zhang y Goodchild, 2002).

Caspary y Scheuring (1993) critican el modelo previamente expuesto ya que supone que los errores de los puntos internos a la línea son los mismos que en los extremos y constantes a lo largo del segmento. Según su planteamiento, los errores de los puntos intermedios son inferiores a los de los puntos extremos.

Las coordenadas de un punto intermedio al segmento formado por los puntos A y B serán:

$$\begin{aligned} X_i &= X_A + \frac{X_E - X_A}{l} l_i \\ Y_i &= Y_A + \frac{Y_E - Y_A}{l} l_i \\ \text{para } 0 \leq l_i \leq l \end{aligned}$$

Ec. II. 6

donde l_i es la longitud del tramo del punto A al punto intermedio y l es la longitud total del segmento.

Suponiendo que los errores en los puntos de inicio y fin son independientes e iguales, el error medio cuadrático de los puntos interiores del segmento es:

$$\begin{aligned} \sigma_{xi}^2 &= \left(1 - \frac{2l_i}{l} + \frac{2l_i^2}{l^2}\right) \sigma^2 = \sigma_{yi}^2 \\ \sigma_{Pi} &= \sigma_{xi} \sqrt{2} \end{aligned}$$

Ec. II. 7

La Figura II. 3 muestra la banda de error posicional definido por σ_{pi} . La desviación estándar en la mitad del segmento $l_i = l/2$ es $1/\sqrt{2}$ veces el valor en los extremos del segmento. Los semicírculos centrados en los puntos finales del segmento muestran el error posicional de A y B. Esta descripción bidimensional puede ser también aplicada a la componente Z.

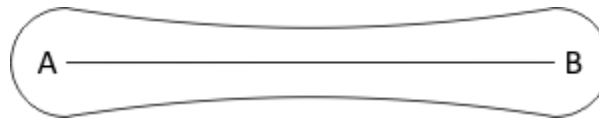


Figura II. 3. Banda de error

Las bandas de error en este caso no son paralelas, el error mínimo es alcanzado a la mitad del tramo. La superficie definida por la banda será:

$$\text{Área} = \pi\sigma^2 + 2L\sigma \int \sqrt{1-2x-2x^2} dx \approx \pi\sigma^2 + 1.62L\sigma$$

Ec. II. 8

Este valor supone aproximadamente el 80% de la superficie de la banda de error de lados paralelos.

Sin embargo, Caspary (1992), presenta una simplificación a la expresión Ec. II. 8 al proponer reemplazar los lados curvos de la banda por tramos rectos (Figura II. 4) manteniendo como punto medio el anteriormente obtenido.



Figura II. 4. Simplificación bandas no paralelas

En este caso la superficie vendrá dada por:

$$\text{Área} = 4 \frac{L}{2} \sigma \frac{1 + \frac{1}{\sqrt{2}}}{2} = \pi\sigma^2 + 1.71L\sigma$$

Ec. II. 9

Al comparar el resultado de esta simplificación, se puede observar una sobreestimación con respecto al resultado real (Ec. II. 9 frente a Ec. II. 8).

En el modelo de banda de error para los puntos finales de un segmento recto plantea la determinación de las posiciones de los mismos a partir de una muestra

aleatoria de posibles situaciones, siguiendo para ello una distribución normal, por lo que asume errores circulares en los extremos (Alesheikh y col., 1999).

Diversos autores (Alesheikh y col., 1999; Zhang y Goodchild, 2002; Ariza, 2002a) describen estudios de errores posicionales utilizando técnicas estadísticas y simulación, en los que se obtiene otra forma de la banda de indeterminación, en la que los semicírculos de los extremos sufren un achatamiento (Figura II. 5).

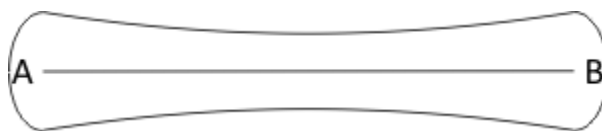


Figura II. 5. Banda obtenida por simulación

De manera análoga al estudio individualizado de segmentos, se estudia lo que ocurre con poligonales (concatenación de segmentos) y polígonos (concatenación y cerramiento de poligonales).

Para conocer el comportamiento estadístico cuando las bandas de error de dos segmentos se superponen, se recurre a la simulación. Los resultados empíricos indican que este comportamiento depende en gran medida del ángulo de confluencia de los segmentos (Figura II. 6).

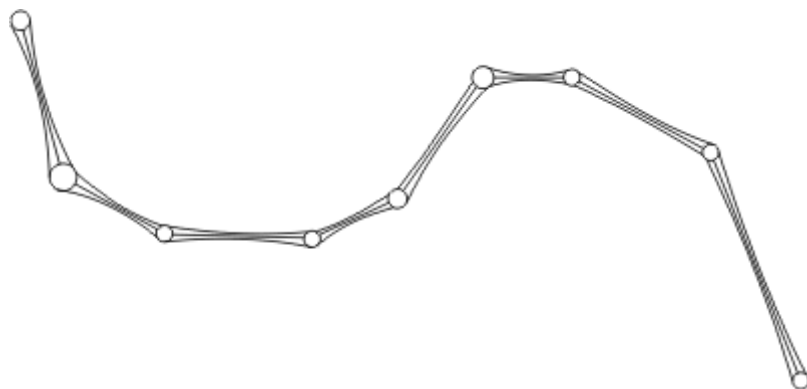


Figura II. 6. Comportamiento de la banda de error

En cualquier caso, para un elemento lineal (L), la banda ϵ de (L) es el interior de la línea que equidista ϵ unidades de (L). En la Figura II. 7 aparecen los tipos de banda ϵ de polígonos (Leung y Yan, 1998).

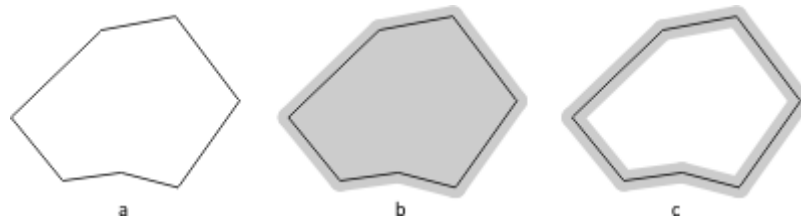


Figura II. 7. Banda épsilon en poligonales y polígonos

donde en la Figura II. 7(a) aparece la línea poligonal, en la Figura II. 7(b) la banda épsilon del polígono y en la Figura II. 7(c) la banda épsilon de la poligonal.

Más recientemente, Shi y Liu (2000) presentan un modelo basado en un proceso estocástico con la asunción (ya vista previamente) de que los puntos extremos del segmento siguen una distribución normal bidimensional. Para un punto del segmento (punto A) la matriz de varianza-covarianza viene dada por la expresión Ec. II. 10:

$$\sum_{AiAi} = \begin{bmatrix} \sigma_{x_i}^2 & \sigma_{x_i y_i} \\ \sigma_{y_i x_i} & \sigma_{y_i}^2 \end{bmatrix}$$

Ec. II. 10

Un segmento se puede definir como el conjunto de puntos infinitos que se expresan como combinación lineal de los puntos extremos del mismo. Shi y Liu (2000) indican esta correlación a través la siguiente matriz:

$$\sum_{AA}(l_1, l_2) = \begin{bmatrix} \sigma_X^2(l_1, l_2) & \sigma_{XY}(l_1, l_2) \\ \sigma_{XY}(l_1, l_2) & \sigma_Y^2(l_1, l_2) \end{bmatrix}$$

Ec. II. 11

La función de probabilidad para un segmento AB se muestra en la Figura II. 8.

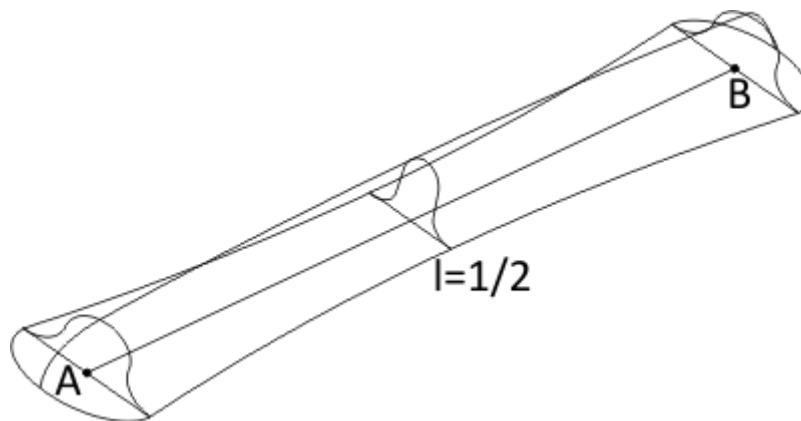


Figura II. 8. Función de densidad de probabilidad del segmento AB

Si se corta la superficie de esta distribución de probabilidad por un plano que es paralelo al plano XY y se proyecta sobre este, se puede obtener un grupo de bandas concéntricas. La banda que es cortada por el plano que pasa a través de las elipses de error en los puntos extremos es definida por Shi y Liu (2000) como la *Banda-g*. La forma y tamaño de esta banda dependerán de la superficie de la función de distribución de densidades (Figura II. 8). En la Figura II. 9 aparece una representación para un segmento de la *Banda-g*.

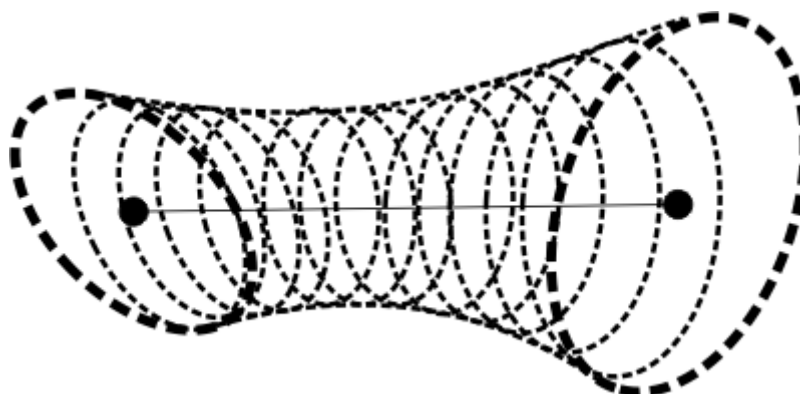


Figura II. 9. Visualización de la *Banda-g* de un segmento

Otros estudios más recientes, como los realizados por Cheung y Shi (2004), apoyan el concepto de región de confianza. Esta región es un área en la cual se registra la posición media de un punto con un grado de confianza (por ejemplo del 95%). Si se construyen estas regiones, teniendo en cuenta las dos componentes de las coordenadas del punto (X, Y), se obtendrá una elipse como representación de esta región teniendo en cuenta que los datos de cada punto procedan de una distribución normal (Figura II. 10). Como consecuencia, se asume que la incertidumbre en el punto sigue una distribución simétrica (simetría en las elipses).



Figura II. 10. Región de confianza en puntos extremos

Tanto una poligonal como un polígono son definidos por su eje o borde respectivamente. Se puede definir, por tanto, un modelo del error posicional a través de un modelo de su eje o borde (Leung y Yan, 1998). Para ello, se tiene en cuenta la banda de incertidumbre de estos ejes o bordes. Según esto, y aplicando el concepto de *Banda-g* de Shi y Liu (2000) se puede obtener, por ejemplo, una representación como la indicada en la Figura II. 11.

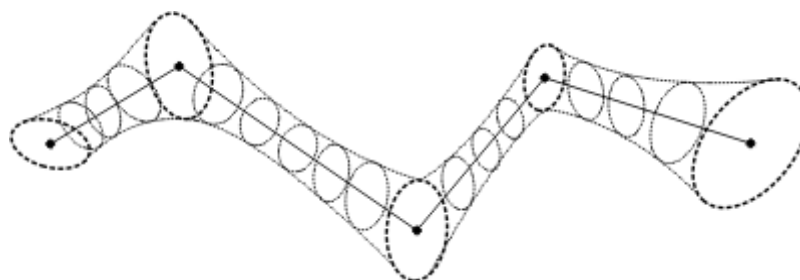


Figura II. 11. *Banda-g* en una poligonal

El modelo de *Banda-g* propuesto por Shi y Liu (2000) para describir la incertidumbre usando regiones de confianza para los puntos que forman la línea y los puntos intermedios de los segmentos puede simplificarse en la práctica, según Cheung y Shi (2004), ya que es imposible construir un modelo con infinitos número de regiones de confianza para toda la línea. De forma alternativa, proponen una muestra de algunos puntos a los que denominan *puntos simbólicos* (Figura II. 11) y que son obtenidos aleatoriamente sobre la línea inicial.

En general, en los últimos años, se han presentado 2 modelos (Alesheikh y col. (1999); Shi y Liu ,2000; Zhang y Goodchild, 2002):

- Modelos de errores de región de confianza: Consisten en una banda alrededor de un segmento lineal medido fuera de la cual la posición real de un segmento es falsa para una probabilidad superior a un nivel predefinido de confianza (modelo determinista).
- Modelos de bandas de error: Consisten en una banda alrededor de la posición real de un segmento. La distribución y los errores de un punto arbitrario en el segmento puede venir dado (modelo probabilista).

El modelo determinista es cuestionado, sin embargo, por Alesheikh y col. (1999) debido a que no se plantea un modelo de la distribución del error dentro de la banda y que la posición verdadera es falsa fuera de la misma.

La Tabla II. 3 resume los conceptos tratados en este apartado.

Modelo	Fuente	Descripción	Tipo distribución	Observaciones
Banda ϵ	Perkal (1956) Chrisman (1982) Blakemore (1984)	Modelo de error posicional para elementos lineales. Un segmento es envuelto en cada lado por un área de anchura constante ϵ . El resultado representa un orlado alrededor de la línea.	Distribución normal circular en los puntos extremos pero sin distribución de errores dentro de la banda.	Modelo ampliamente utilizado para incertidumbre de elementos lineales.
Banda de error	Dutton (1992) Caspary, Scheuring (1992) Shi (1994)	Extensión del Modelo de banda ϵ . Se elimina la banda de anchura constante cambiándose por una banda de error que tiene anchura mínima en el centro del segmento.	Asume distribuciones normales del error en los puntos extremos.	Introduce la noción de compresión de las funciones de densidad en los extremos y en el centro del segmento.
Banda-g	Shi, Liu (2000)	Descripción más genérica del modelo de banda de error, permitiendo la interrelación entre los 2 puntos extremos. La distribución bidimensional de los puntos extremos de la Banda-g viene determinada por las elipses de error de los puntos extremos y su superficie de distribución espacial de densidades.	Los errores de los extremos de los segmentos siguen una distribución normal bidimensional.	Basado en un proceso estocástico. La Banda-g no es simétrica. Las elipses de error poseen ejes y orientaciones variables.

Tabla II. 3. Resumen de los modelos de error posicional para elementos lineales

II.4.2. Métodos de control de exactitud posicional en elementos lineales

II.4.2.1 Método de banda ϵ

Este método está basado en el concepto de banda ϵ previamente analizado, y fue desarrollado por Skidmore y Turner (1992). La metodología es analizada por Giordano y Veregin (1994), así como por Ariza (2002a) y Atkinson y Ariza (2002).

- Aplicaciones corrientes: Hasta el momento el test presenta interés teórico.
- Método de comparación: Con fuentes de mayor exactitud o estimaciones deductivas.
- Componente posicional: Horizontal.
- Clase de elementos: Línea o área.
- Correspondencia con un estándar de exactitud predefinido: No se ha elaborado un estándar hasta la fecha.
- Descripción: La banda ϵ es un tipo de distribución centrado sobre una línea o borde de un polígono. Se asume que la probabilidad de la posición real de la línea sobre el terreno contenida dentro de ϵ sea 1. La anchura de la banda está determinada por el valor de ϵ ,

que define el grado de incertidumbre relacionada con la posición real de la línea. Cuanto más grande es ϵ , mayor es la incertidumbre.

- Procedimiento: La estimación de los valores de ϵ puede realizarse de 2 formas:

ϵ puede ser estimado de manera deductiva utilizando el nivel de error (conocido o presupuesto) introducido en las diversas fases de la codificación del dato.

Estimar ϵ comparando el dato con fuentes de mayor exactitud. En caso de coincidir los puntos extremos de la línea del mapa y de la fuente de mayor exactitud el procedimiento a utilizar es:

1. Selección de una muestra de líneas en la base de datos.
2. Seleccionar las mismas líneas en la otra fuente de mayor exactitud.
3. Unir los puntos finales de la línea tanto en el mapa con en la otra fuente.
4. Medir el alejamiento (la diferencia entre la posición de la línea del mapa y de la fuente de mayor exactitud, AB y AC en la Figura II. 12) a intervalos regulares. El alejamiento debe medirse perpendicular a la línea que conecta los puntos finales de la línea del mapa y la fuente de mayor exactitud (Figura II. 12).
5. Los alejamientos resultantes forman una distribución estadística del error. Se pueden calcular diversos parámetros de la distribución, incluso el error medio, el error medio cuadrático, el error mínimo y el máximo. Si se asume una distribución centrada en torno a la línea, ϵ define una zona que tiene probabilidad igual a 1 de contener la posición real de la línea sobre el terreno. El error máximo puede ser utilizado como error de ϵ .

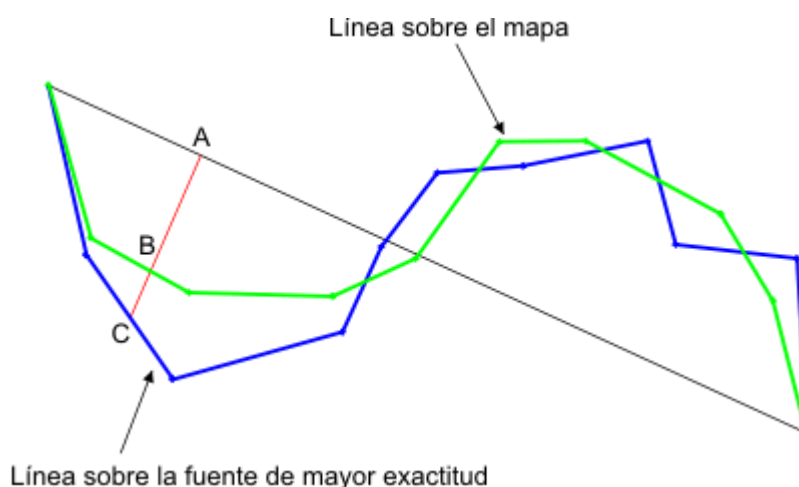


Figura II. 12. Método de banda ϵ . Puntos finales coincidentes

En caso de no coincidir los puntos extremos de la línea del mapa y de la fuente de mayor exactitud el procedimiento a utilizar es:

1. Selección de una muestra de líneas en la base de datos.
2. Seleccionar las mismas líneas en la otra fuente de mayor exactitud.
3. Medir el área total de desplazamiento entre la línea del mapa y la línea de mayor exactitud (Figura II. 13).
4. Para cada línea, dividir el área por la longitud de la línea sobre el mapa para calcular un valor aproximado de epsilon.
5. El valor resultante forma una distribución estadística de los valores de epsilon. Se pueden calcular diversos parámetros de la distribución, incluso la media o el error medio cuadrático de epsilon.

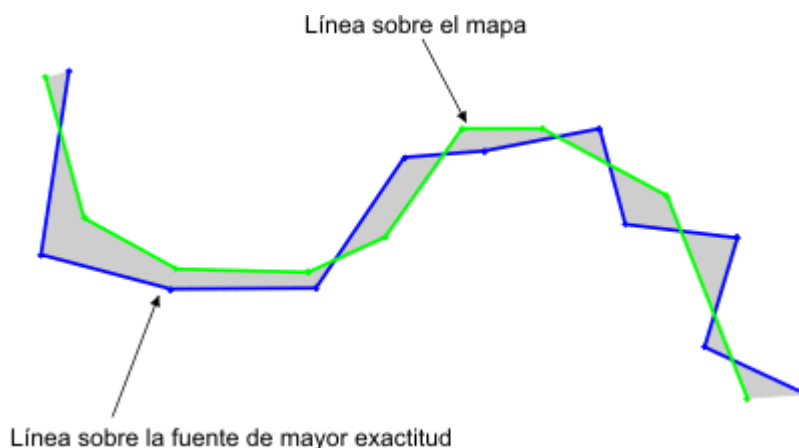


Figura II. 13. Método de banda epsilon. Puntos finales no coincidentes

- **Discusión:** Como indican Giordano y Veregin (1994) es importante subrayar que se pueden encontrar dos versiones opuestas de la banda epsilon. La primera, donde la banda está centrada sobre la línea del mapa y se presume que contiene la posición real de la línea sobre el terreno; y la segunda, que parte de la posición de la línea, sobre la cual se superpone la banda epsilon presumiendo que la línea del mapa está en su interior.

Atkinson y Ariza (2002) subrayan que esta metodología resulta interesante al ser la primera en desarrollarse. Su implementación no supondría gran esfuerzo dentro de un SIG. Sin embargo, no se tienen en cuenta errores a priori del mapa ni de la fuente de mayor exactitud, ni se menciona ningún nivel de confianza con el que obtener los resultados.

Veregin (2000), plantea una metodología basada en la banda ϵ que se estudia en el apartado II.4.2.5 Método MPS y MDS de Veregin.

II.4.2.2 Método del orlado simple

Esta metodología está basada en los trabajos de Goodchild y Hunter (1997). Atkinson y Ariza (2002) analizan este método:

- Aplicaciones corrientes: Posible implementación en aplicaciones SIG.
- Método de comparación: Con fuentes de mayor exactitud.
- Componente posicional: Horizontal.
- Clase de elementos: Línea.
- Correspondencia con un estándar de exactitud predefinido: No se ha elaborado un estándar hasta la fecha.
- Descripción: El método se fundamenta en la generación de orlados o buffers alrededor de la línea de la fuente de mayor exactitud, contabilizando el porcentaje de línea del mapa que se encuentra dentro de este orlado (Figura II. 14).

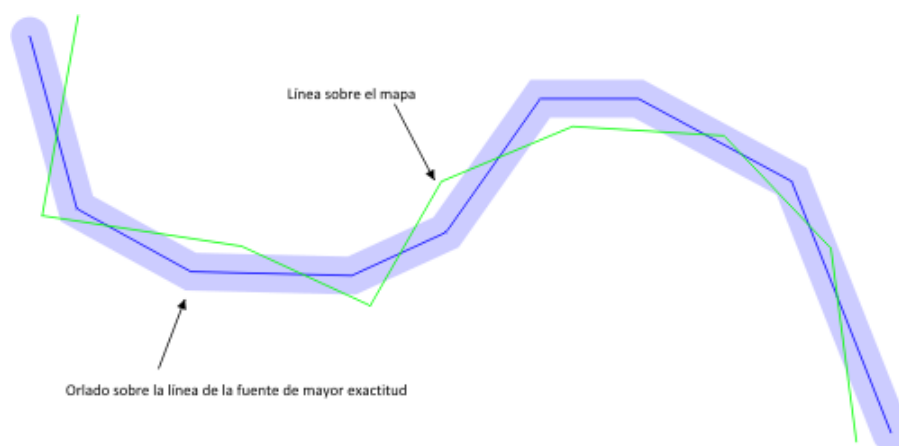


Figura II. 14. Orlado simple

- Procedimiento:
 1. Seleccionar una muestra de líneas en el mapa.
 2. Seleccionar estas mismas líneas en la fuente de mayor exactitud.
 3. Generar un orlado de anchura X metros sobre las líneas de la fuente de mayor exactitud.
 4. Contabilizar el percentil (%) de la longitud de la línea del mapa que se encuentra dentro del orlado.
 5. Incrementar la anchura del orlado y repetir el proceso desde el punto 3.

Como consecuencia se obtiene una distribución de probabilidades de pertenencia de las líneas de mapa al orlado sobre la fuente de mayor exactitud, de forma, que se puede discriminar para un percentil que se establezca como umbral la anchura de indeterminación (curva con puntos negros en Figura II. 15). La curva resultante se aproxima a la curva de una distribución gaussiana (curva con puntos blancos en Figura II. 15) (Goodchild y Hunter, 1997).

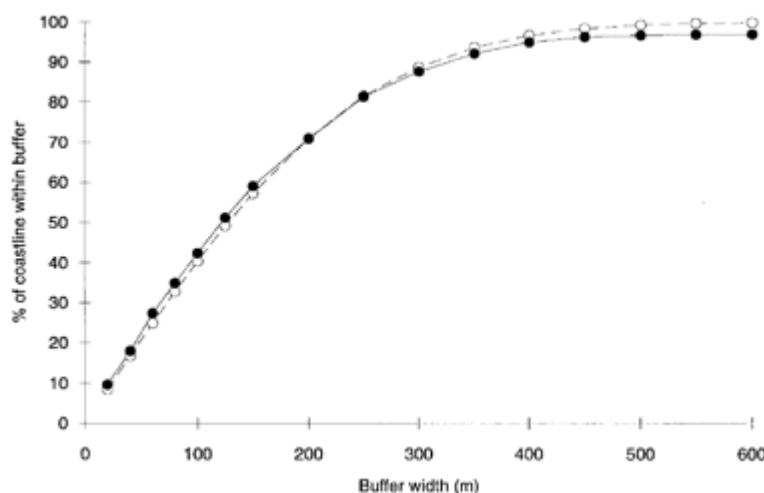


Figura II. 15. Resultados obtenidos por Goodchild y Hunter, 1997

- **Discusión:** Según Atkinson y Ariza (2002), este método resulta bastante robusto y fácil de implementar en un SIG. Sin embargo, no tiene en cuenta a qué lado de la línea de la fuente de mayor exactitud se encuentra la línea del mapa, por lo que no es posible detectar posibles sistematismos. El resultado del tratamiento de errores obtenido es a nivel global en cuanto a la calidad geométrica del mapa.

Una vez descrito el método, presentamos algunos ejemplos o aplicaciones del mismo, comenzando con el ensayo realizado por sus autores.

Goodchild y Hunter (1997) ensayan estos estudios sobre varias líneas de costa de Australia obtenidas del *Digital Chart of the World* (DCW) a escala 1:100.000 comparándolas con las mismas líneas procedentes de un mapa topográfico local a escala 1:25.000. De esta forma, aplican una transformación de coordenadas para referir ambas cartografías a la misma proyección y datum. Los puntos finales de las líneas del mapa de mayor exactitud se enlazaron a los pertenecientes a la cartografía a controlar. Como resumen, se obtuvieron 247 km en el mapa topográfico que se correspondían con 179 km del DCW. La diferencia fundamentalmente proviene de la existencia de estuarios de ríos en el mapa topográfico que no aparecen sobre el DCW.

Los resultados obtenidos muestran que para un 90% el valor obtenido es de 330 m.

Kagawa, Sekimoto y Shibaski (1999) en su estudio comparativo de varias metodologías, incluyen el método del orlado. El estudio se realiza sobre datos de

carreteras en la ciudad de Sagamihara (Japón). Los datos a controlar proceden de un vuelo aéreo a escala 1:2.500 y los datos de las líneas de mayor exactitud proceden de carreteras digitalizadas a escala 1:500. La longitud total de la carretera es de 706 m y los resultados obtenidos revelan para un 95% de longitud dentro del orlado una anchura del mismo de 1.99 m.

Van Niel y McVicar (2002) realizan un estudio comparativo de evaluación de exactitud posicional con una metodología basada en puntos y otra en líneas. El método basado en puntos utiliza los estándares NMCA¹ (Australia) y NSSDA (ver A1.5 en Anexo I), mientras que la metodología que utiliza elementos lineales sigue el test del orlado simple de Goodchild y Hunter (1997). El estudio se realiza en el sur de Nueva Gales del Sur (Australia) y se utiliza como cartografía a controlar las carreteras de la *Digital Topographic Data Base* (DTDB) a escala 1:50.000, siendo la tolerancia en Australia para estas escalas de 25 m.

Para el control se utilizan los ejes de 466 km de carreteras obtenidos por GPS diferencial (DGPS). Estas líneas son transformadas del sistema de coordenadas AMG84² al propio de la cartografía (AMG66), suponiendo una exactitud de las líneas DGPS de 1.5 m.

Los resultados presentados por Van Niel y McVicar (2002) son:

- NCMA (puntos): Cerca de la mitad de las intersecciones de carreteras testadas superaban la tolerancia de 25 m por lo que no cumplían el estándar australiano.
- NSSDA (puntos): Estima que el 95% de las posiciones tenían un error menor o igual a 105 m (cuatro veces superior a la tolerancia exigida).
- Orlado (líneas): En este caso, se realizan orlados con un salto de 25 m desde 25 a 325 m, indicando que sólo un 78 % de las líneas estarían dentro de los 25 m marcados como tolerancia. Para un 90% de las líneas el error sube a 40 m.

Johnston, Timlin, Szafoni y Casanova (1999) presentan la aplicación de esta metodología a las carreteras y arroyos de la base de datos geográfica del ITAM³ en Fort Hood (Texas- EE.UU.). Como base de datos de control utilizan el USGS DRG (Data Raster Graphics) y varias ortofotografías de la zona. Los resultados obtenidos tras aplicar el método de orlado simple se muestran en la Figura II. 16.

¹ NMCA: National Mapping Council of Australia

² AMG84: Australian Map Grid 1984

³ ITAM: Integrated Training Area Management

Positional Accuracy Testing For Linear Features			
Feature:		Streams	
Source:		USGS 7.5 minute DRG's	
Sampling Method:		Watershed sample	
Buffer Size	Frequency	Sum of Lengths	% of Line Within Buffer
0	378	268161.690766	100.00
2	4	114.615718	0.04
4	1201	47327.676870	17.65
6	1214	64820.976922	24.17
8	1227	80805.187532	30.13
10	1207	97192.324484	36.24
12	1191	112763.993813	42.05
14	1155	126707.492506	47.25
16	1122	140313.475707	52.32
18	1071	153986.080295	57.42
20	1008	166923.304629	62.25
22	978	177492.675907	66.19
24	912	188901.756244	70.44
26	868	198021.496041	73.84
28	834	205620.451366	76.68
30	776	213363.604876	79.57
32	746	219925.996425	82.01
34	693	226568.977396	84.49
36	651	231707.049227	86.41
38	606	236466.753069	88.18
40	579	240139.369723	89.55
60	431	258441.593583	96.38
80	386	262380.255792	97.84
100	380	263056.353341	98.10

Figura II. 16. Resultados método de orlado simple Johnston y col. (1999)

Mozas (2003) y Ariza y Mozas (2005) presentan los resultados de un proyecto piloto en el que se estudian las carreteras de una hoja topográfica 1:25.000 correspondiente al Mapa Topográfico Nacional, en la zona de Jaén (Hoja MTN 947). Como fuente de mayor exactitud se emplea los resultados de un levantamiento cinemático con GPS de las mismas. Una vez procesados los datos obtenidos en la observación y transformados al mismo sistema de coordenadas que la cartografía a controlar, se procesan las líneas con una aplicación informática creada por los autores.

Mozas (2003) presenta una estructura para la planificación y ejecución de esta metodología que abarca desde la toma de datos con la observación GPS hasta la obtención de resultados. Se trata de un documento base para esta tesis doctoral.

El test de orlado simple se aplica a 136 kms de carreteras en su globalidad y a varias selecciones de líneas atendiendo a criterios como el tipo de carretera, el ángulo origen-fin de la línea, la sinuosidad aparente, la zona o la longitud con la finalidad de detectar alguna relación entre los resultados dependiendo de estas selecciones.

Tras aplicar la metodología previamente descrita obtienen los resultados presentados gráficamente en la Figura II. 17 donde en el eje de abscisas se muestra la anchura del orlado y en el eje de ordenadas el porcentaje de inclusión en el mismo.

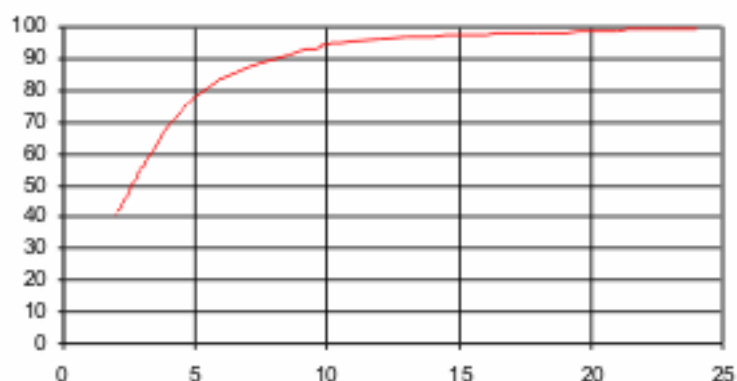


Figura II. 17. Resultados globales test orlado Hoja MTN 947. Mozas (2003)

En la Figura II. 17 se puede observar la distribución de probabilidades y su curva resultante. Para un nivel del 90% se obtiene un valor de orlado de 8 m.

Al realizar el estudio seleccionando subconjuntos de elementos según una serie de criterios, los autores aportan otros resultados, como que las carreteras autonómicas se comportan peor que el resto de elementos. Este resultado tenía su causa en las recientes mejoras en su trazado que sufrieron estas carreteras y que no habían sido incluidas en la cartografía.

Sin embargo, no se encontraron diferencias significativas al dividir el estudio en criterios como la longitud del tramo, la orientación o la sinuosidad.

Desde otro punto de vista, Mas (2006) presenta un trabajo en el que la incertidumbre posicional de los bordes de polígonos son estudiados usando una aproximación de la banda ϵ . Para ello genera un orlado sobre los bordes de polígonos de una cartografía y realiza un proceso de *overlay* sobre los mismos bordes procedentes de una época distinta. Teniendo en cuenta las zonas de incertidumbre generadas por los orlados, elimina las variaciones aparecidas en los bordes dentro de los orlados, con lo que obtiene una importante reducción en cuanto al número de polígonos discrepantes y el área de los mismos.

II.4.2.3 Método del doble orlado

El método del doble orlado es ideado por Tveite y Langaas (1999) que lo denominan B.O.S. (*Buffer Overlay Statistics Method*). Este test puede ser usado para evaluar: la anchura de las bandas ϵ , el desplazamiento medio, el nivel de generalización y la compleción. El método ha sido verificado con éxito sobre distintos tipos de líneas (carreteras, ferrocarriles y líneas de costa) del DCW, frente a cartografía 1/250.000 del *Norwegian Mapping Authority* (NMA) y es analizado por Atkinson y Ariza (2002) resumiendo sus principales características:

- Aplicaciones corrientes: Posible implementación en aplicaciones SIG.
- Método de comparación: Con fuentes de mayor exactitud.

- Componente posicional: Horizontal.
- Clase de elementos: Línea.
- Correspondencia con un estándar de exactitud predefinido: No se ha elaborado un estándar hasta la fecha.
- Descripción: El método se fundamenta en la generación de orlados, aunque es algo más complicado que el método de orlado simple. Esto es debido a que el proceso de orlado (buffer) se realiza sobre varias líneas: en la cartografía a analizar (X) y en la fuente de mayor exactitud (Q). Los orlados generados son denominados XB y QB respectivamente. Según las situaciones que estos orlados plantean al intersectar en el espacio, se pueden distinguir 4 tipos de áreas (Figura II. 18):
 - Tipo 1: Zona exterior que no está incluido en ningún orlado: $\overline{XB} \cap \overline{QB}$
 - Tipo 2: Zona interior al orlado de línea de la fuente de mayor exactitud y exterior a orlado de la línea del mapa: $\overline{XB} \cap QB$
 - Tipo 3: Zona interior al orlado de la línea del mapa y exterior a orlado de la línea de la fuente de mayor exactitud: $XB \cap \overline{QB}$
 - Tipo 4: Zona interior a los dos orlados (intersección): $XB \cap QB$

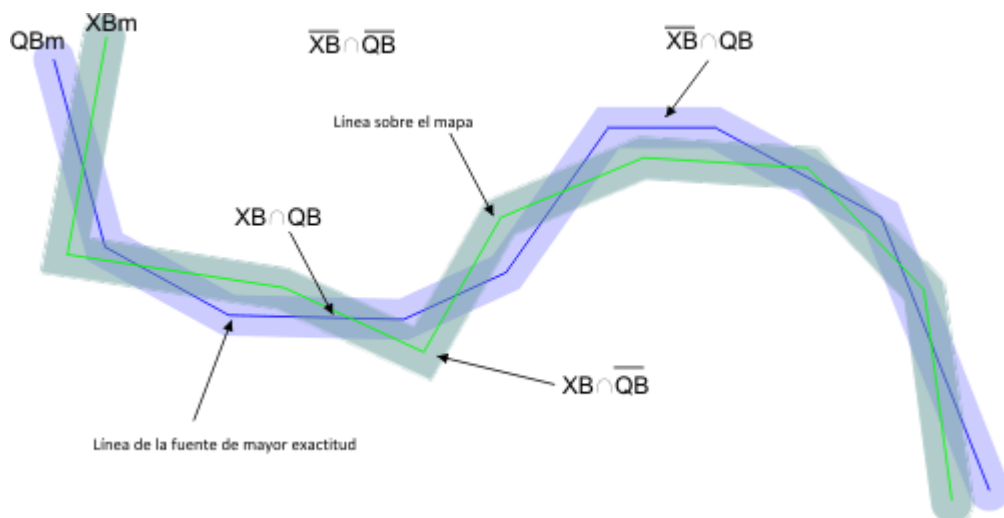


Figura II. 18. Método de doble orlado

Si las líneas son muy similares, el área intersección $XB \cap QB$ será dominante a las otras, pero si las líneas presentan más diferencias, el área de los otros tipos se incrementa en función del tamaño de los desplazamientos. El área $XB \cap QB$ comparada al total de áreas de los orlados XB y QB puede ser usada como una medida de la exactitud posicional de la línea. Si las líneas son muy similares en su forma y posición, el área tipo 4 domina sobre las demás, mientras que si aumentan las áreas tipo 2 y 3 en función del tamaño del orlado será un

síntoma de que existe variación en tamaño y posición. Cuando son similares en forma pero difieren en posición (existe un desplazamiento), se puede obtener una estimación de la exactitud posicional cuando el área de tipo 4 ($XB \cap QB$) se aproxima al 50% de las áreas de XB o QB . En la Figura II. 19 se aprecia los resultados gráficos obtenidos por Tveite y Langaas (1999) para las distintas superficies propuestas.

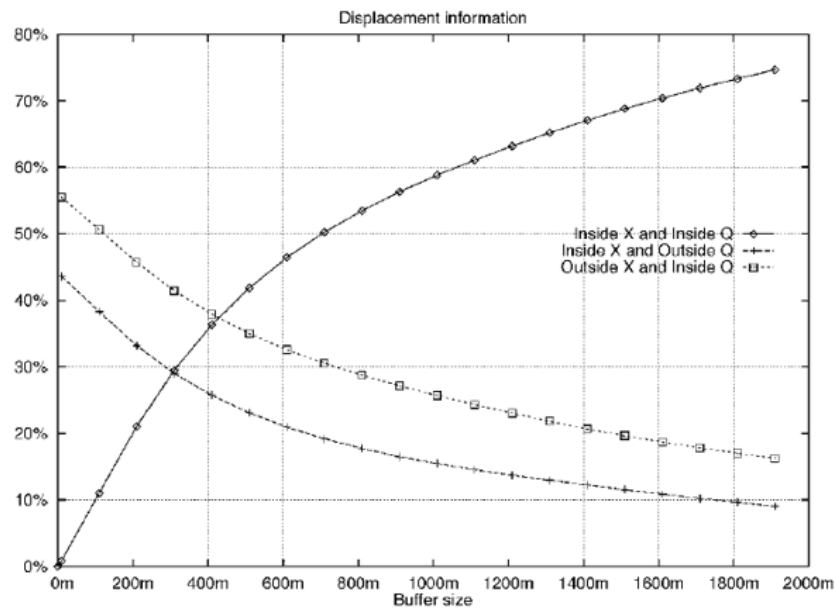


Figura II. 19. Doble orlado: Comportamiento áreas. Tveite y Langaas (1999)

Tveite y Langaas (1999) presentan algunas medidas evaluadoras de la exactitud posicional:

- a) Desplazamiento promedio: El desplazamiento promedio para cada tamaño de orlado (i) se calcula con la expresión:

$$DE_i = \frac{\pi}{2} \frac{2bs_i \text{Área}(\overline{XB_i} \cap \overline{QB_i})}{\text{Área}(XB_i)} = \pi bs_i \frac{\text{Área}(\overline{XB_i} \cap \overline{QB_i})}{\text{Área}(XB_i)}$$

Ec. II. 12

donde bs_i es la anchura del orlado.

Esta medida se usa para determinar la media de los desplazamientos de una línea Q de mayor exactitud sobre otra línea X .

Si se representan gráficamente los resultados del desplazamiento promedio (Figura II. 20) cabe esperar que la curva resultante de aplicar la expresión Ec. II. 12 a varias distancias de orlado aumente regularmente hasta que el tamaño del orlado alcance el desplazamiento medio de las líneas, entonces la curva debería comenzar a estabilizarse. La forma de

la curva por lo tanto, dará una indicación del desplazamiento promedio. También se puede realizar una evaluación de la banda epsilon media de la línea X.

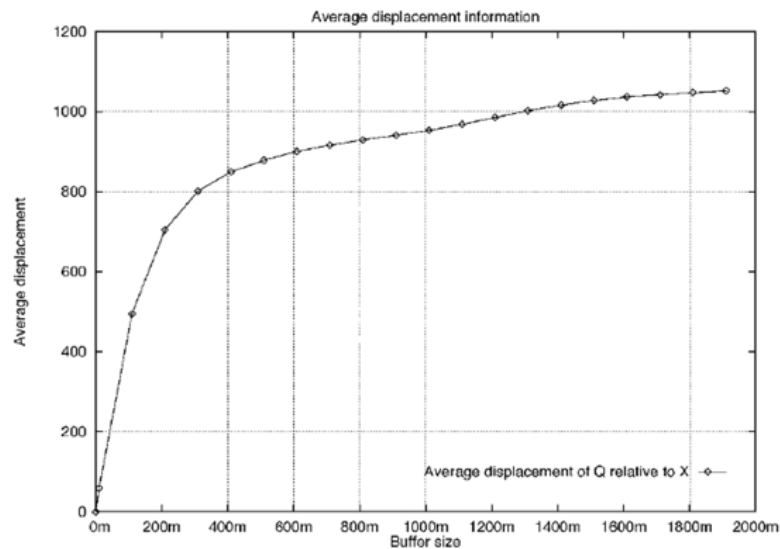


Figura II. 20. Doble orlado: Desplazamiento promedio. Tveite y Langaas (1999)

b) Oscilación: La medida de la oscilación se calcula con la expresión:

$$O_i = \frac{\# \text{Áreas}(\overline{XB_i} \cap \overline{QB_i})}{\text{Long}(X)}$$

Ec. II. 13

Donde $\# \text{Áreas}()$ es el número de áreas y la $\text{Long}(X)$ es la longitud de la línea a controlar (se usa para normalizar el resultado).

Esta medida es la más útil para medir oscilaciones aleatorias. Puede ser usada como una indicación de la existencia de sistematismos (probablemente haya sistematismos si el número de oscilaciones es bajo para lo esperable en dos líneas independientes).

Sin embargo, la mejor medida para encontrar sistematismos es la ecuación modificada:

$$O_i = \frac{\# \text{Áreas}(\overline{XB_i} \cap \overline{QB_i})}{\text{Long}(X)}$$

Ec. II. 14

Gráficamente (Figura II. 21) se obtiene una curva que disminuye regularmente hasta que tiende a estabilizarse al aumentar la anchura de los orlados.

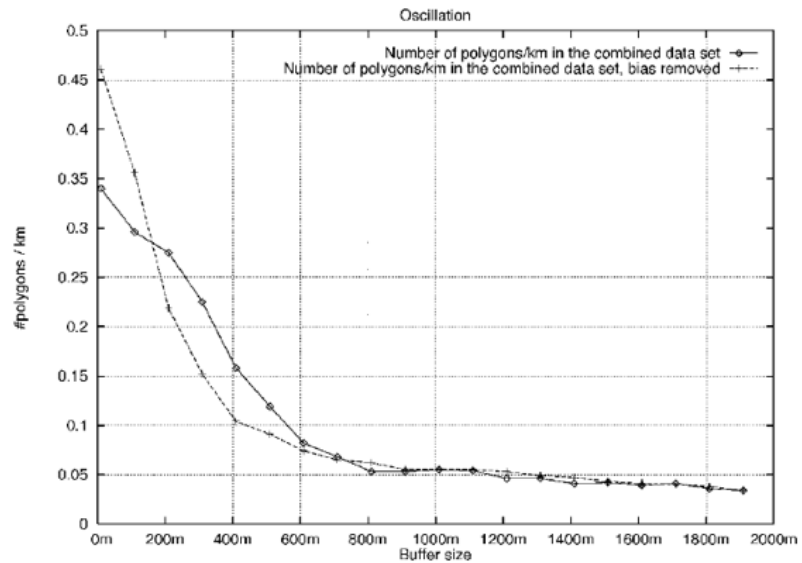


Figura II. 21. Doble orlado: Oscilación. Tveite y Langaas (1999)

- c) **Compleción:** Como se indica previamente en este capítulo, la completión mide lo completa que es la base de datos. La medida de la completión se calcula con la expresión:

$$C = \frac{Long(Q \text{ dentro } XB)}{Long(Q)}$$

Ec. II. 15

donde $Long(Q \text{ dentro } XB)$ es la longitud de la línea de mayor exactitud que está dentro del orlado de la línea del mapa (X) y $Long(Q)$ es la longitud de la línea de mayor exactitud.

Si se observa la expresión Ec. II. 15, esta medida es similar a la obtenida con el método del orlado simple, salvo que en aquel caso se aplicaba el orlado sobre la línea de mayor exactitud.

Gráficamente cabe esperar una curva que aumenta hasta que el tamaño del orlado alcanza el desplazamiento promedio, donde la curva se estabiliza (Figura II. 22).

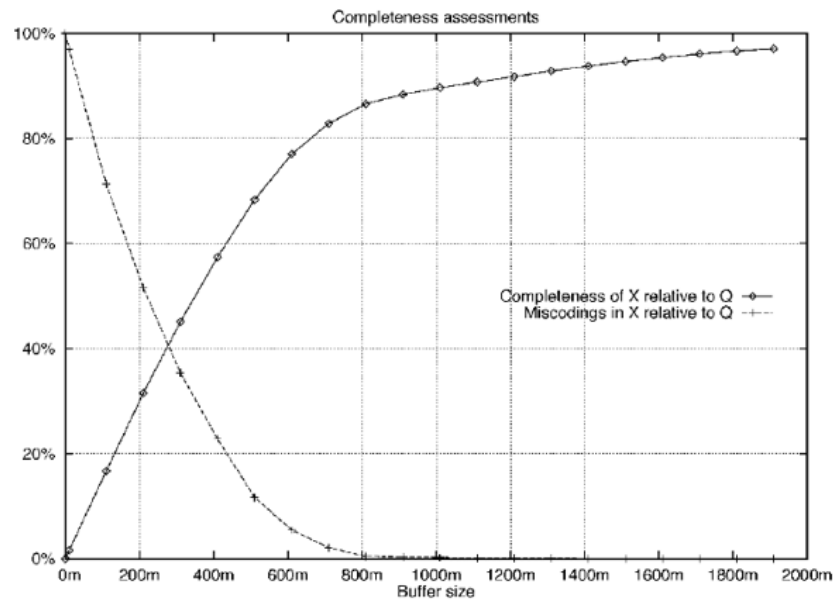


Figura II. 22. Doble orlado: Compleción y omisiones. Tveite y Langaas (1999)

- d) Sin codificar: Indica la falta de información de la línea del mapa (X) comparada con la línea de la fuente de mayor exactitud (Q):

$$SCod = \frac{Long(X \text{ fuera } QB)}{Long(X)}$$

Ec. II. 16

donde $Long(X \text{ fuera } QB)$ es la longitud de la línea del mapa que está fuera del orlado de la línea de la fuente de mayor exactitud y $Long(X)$ es la longitud de la línea del mapa.

Gráficamente cabe esperar una curva que disminuye hasta que el tamaño del orlado alcanza el desplazamiento promedio (Figura II. 22).

- Procedimiento: Para aplicar el método se sigue un proceso iterativo:
 1. Seleccionar una muestra de líneas en el mapa (X).
 2. Seleccionar estas mismas líneas en la fuente de mayor exactitud (Q).
 3. Calcular la longitud de los conjuntos de líneas $Long(X)$ y $Long(Q)$.
 4. Para un orlado de anchura bs_i metros se realiza:
 - Orlado de los elementos X y Q de las bases de datos obteniendo los polígonos XB_i y QB_i de tamaño $2bs_i$.
 - Proceder a la superposición de los polígonos generando el nuevo conjunto de polígonos $XBQB_i$.

- Calcular las áreas de tipo 2, 3, y 4 y los estadísticos de interés (desplazamiento promedio, oscilación compleción y omisión).
- 5. Incrementar el tamaño de orlado ($bs_i + \Delta$) y repetir el proceso desde el punto 4.
- 6. Obtener y analizar los resultados, estudiando los gráficos obtenidos.
- 7. Incrementar la anchura del orlado y repetir el proceso desde el punto 3.
- Discusión: Al igual que el método del orlado simple, se realiza un proceso iterativo aumentando la anchura del orlado, pero en este caso sobre las líneas de las dos fuentes. Los requerimientos del software son más complejos que en el caso del orlado simple, aunque los resultados ofrecen un análisis más exhaustivo sobre la tipología de los errores, su magnitud y la compleción de la base de datos.

Tveite y Langaas (1999) aplican esta metodología a varios grupos de líneas:

- Líneas procedentes del N250 (Mapa nacional a escala 1:250.000) de la *Norwegian Mapping Authority* (NMA) de Noruega.
- Líneas procedentes del *Digital Chart of the World* (DCW) producido por la *Defence Mapping Agency* (DMA) de los Estados Unidos a escala 1:1.000.000.
- Líneas procedentes del *World Vector Shoreline* (WVS), producidos por la *Defence Mapping Agency* (DMA) de los Estados Unidos a escala 1:250.000.

Las líneas del N250 son utilizadas como la fuente de mayor exactitud (Q) en todos los casos. Se realizaron 20 iteraciones comenzando con orlados de 10 m e incrementando con 100 m en cada iteración. El estudio se realiza sobre datos de carreteras, líneas de costa y ferrocarriles, obteniendo unos los resultados acordes a lo previsto en la metodología (Tveite y Langaas, 1999).

Goeman, Martínez-Fonte, Bellens y Gautama (2005a y 2005b) presentan los resultados de aplicar esta metodología a tramos de carreteras obtenidos de forma automática a partir de imágenes Quickbird comparándolas con la digitalización manual de estas vías. La métrica utilizada es el desplazamiento promedio, obteniendo los resultados que se muestran en la Figura II. 23 correspondientes a la curva representada en negro (D1).

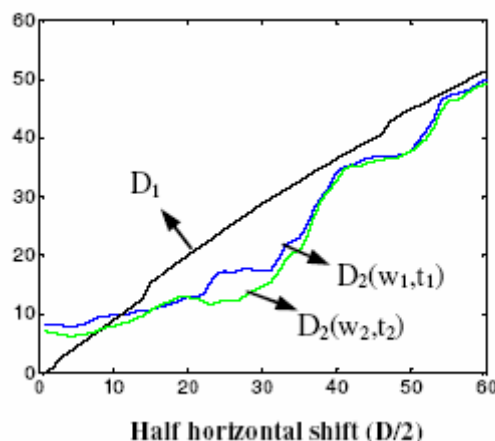


Figura II. 23. Resultados método doble orlado. Goeman y col. (2005)

II.4.2.4 Método basado en las distancias de Hausdorff

Esta metodología se basa en los trabajos de Abbas, Grussenmeyer y Hottier (1995) y permite evaluar la exactitud planimétrica a partir de las distancias de Hausdorff de pares de objetos. Atkinson y Ariza (2002) resumen sus principales características:

- Aplicaciones corrientes: Posible implementación en el control de procesos de generalización y calidad posicional.
- Método de comparación: Con fuentes de mayor exactitud.
- Componente posicional: Horizontal.
- Clase de elementos: Línea.
- Correspondencia con un estándar de exactitud predefinido: No se ha elaborado un estándar hasta la fecha.
- Descripción: El método propuesto utiliza la distancia de Hausdorff como herramienta matemática entre una pareja de líneas. Se basa en el cálculo de las distancias de Hausdorff sobre pares de líneas homólogas, a las cuales se les ha eliminado cualquier detalle discordante, y se han normalizado los valores finales en función del RMSE (Ec. II. 1 y Ec. II. 2) y el factor de generalización. En este caso se tiene muy en cuenta un aspecto no tratado hasta el momento: el nivel de generalización de los elementos lineales.
- Procedimiento:
 1. Seleccionar una muestra de líneas en la base de datos (Q).
 2. Seleccionar las mismas líneas en otra fuente de mayor exactitud (X).
 3. Regularización de las líneas Q y X. Para ello, deberemos eliminar aquellos detalles discordantes producidos por aspectos como

tamaño, forma, complejidad, número de segmentos en los elementos lineales, ..., que perturbarían el resultado final. Su eliminación se logra mediante un proceso de simulación sobre cada pareja de líneas a partir de valores sobre la exactitud del mapa y de la referencia.

4. Cálculo de las distancias de Hausdorff d_H . Ésta se obtiene mediante el cálculo de d_1 como el máximo de las distancias menores del mapa a la referencia ($Q \rightarrow X$) y d_2 como el máximo de las distancias menores de la referencia al mapa ($X \rightarrow Q$) (Figura II. 24). Sobre ellas se define la distancia d_H como el máximo de entre las dos componentes anteriores:

$$d_H = \text{Max}\{d_1, d_2\}$$

Ec. II. 17

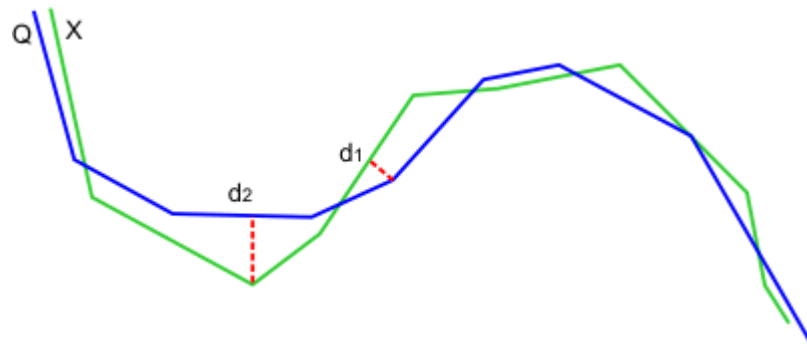


Figura II. 24. Método basado en las distancias de Hausdorff

A partir de los valores de distancia obtenidos, Abbas, Grussenmeyer y Hottier (1995) introducen la medida del *Índice de Generalización* según la siguiente expresión:

$$i = \frac{d_2 - d_1}{d_2 + d_1} \in [-1, 1]$$

Ec. II. 18

5. Normalización de valores. Para normalizar los valores, las distancias iniciales d_1 , d_2 y d_H se dividen por las cantidades d_{10} , d_{20} y d_h obtenidas para cada pareja de objetos tras una simulación basada en la perturbación de los vértices de los elementos lineales con valores de RMSE determinados a priori ($RMSE_{01}$ para el mapa y $RMSE_{02}$ para la referencia), y multiplicados por la discrepancia a priori $\sqrt{RMSE_{10}^2 + RMSE_{20}^2}$. En estas simulaciones se considera que el error cartográfico ($RMSE_1$) es la suma de dos errores: un error puntual $\sqrt{(1-G^2)RMSE_1}$ y un error de generalización $G * RMSE_1$, donde G es un factor de generalización

6. Obtención de estimadores:

- a) Porcentaje de acuerdo: como el porcentaje de líneas del mapa trazadas en acuerdo con la referencia (ratio entre las líneas regularizadas del mapa y las líneas iniciales).
 - b) RMSE planimétrico: en base a la estimación de precisión alcanzada para el conjunto de líneas regularizadas.
- Discusión: Según Atkinson y Ariza (2002) este método se basa en el cálculo de las distancias de Hausdorff en vez de orlados sobre las líneas, por lo que se simplifican los cálculos necesarios. No obstante, los resultados tan solo nos ofrecen información sobre el porcentaje de acuerdo entre ambas líneas (cartografía y fuente de mayor exactitud) y la cuantía del RMSE.

Abbas, Grussenmeyer y Hottier (1995) detectan que la media de las distancias d_1 es generalmente menor que la media de las distancias d_2 , es decir, la distancia del mapa a la realidad (fuente de mayor exactitud) es más pequeña que la distancia de la realidad al mapa.

Por ejemplo, los autores presentan para el caso de una muestra de 266 ejes de carreteras y 101 ejes de vías férreas de la BDCarto (IGNF) sobre una zona rectangular de 6x8 km. Los ejes están formados por tramos con longitudes entre 100 y 1000 m, se determinó un 95.5% de acuerdo y un RMSE de 10.67 m con una precisión de 0.35 m en las primeras, y un 95.5% de acuerdo y un RMSE de 11.64 en las vías férreas.

Kagawa, Sekimoto y Shibaski (1999) en su estudio comparativo de varias metodologías, incluyen el método de basado en las distancias de Hausdorff. El estudio se realiza sobre datos de carreteras en la ciudad de Sagamihara (Japón). Los datos a controlar proceden de un vuelo aéreo a escala 1:2.500 y los datos de las líneas de mayor exactitud proceden de carreteras digitalizadas a escala 1:500. Las distancias obtenidas son $d_1 = 4.3$ m y $d_2 = 5.8$ m.

Skopeliti y Tsoulos (2000a, 2000b) realizan un estudio de la exactitud posicional utilizando esta metodología. Los datos utilizados proceden de un mapa digitalizado a escala 1:100.000 y más concretamente, se trata de líneas de costa de la isla de Ithaki (Grecia). El estudio realizado consiste en segmentar las líneas categorizándolas según tres niveles de sinuosidad: suaves, sinuosas y muy sinuosas. Posteriormente, se digitalizan las mismas líneas utilizando una imagen raster del mapa. Esta digitalización es realizada por 18 operadores. Las líneas digitalizadas son contrastadas con las originales utilizando la metodología de la distancia de Hausdorff.

Los resultados presentan que el grupo de líneas pertenecientes a la categoría muy sinuosa, muestran la mayor desviación horizontal. Las líneas sinuosas presentan valores más moderados de desviación horizontal y las suaves obtienen los números más bajos.

Por otro lado, Vauglin (1999) presenta una alternativa a esta metodología, con la finalidad de representar gráficamente un modelo de incertidumbre posicional. Para ello utiliza las funciones generadas a partir de las distancias mínimas de cada punto de

cada línea con respecto a la otra. Esta metodología es denominada por el autor con el término: *discrepancia geométrica*.

II.4.2.5 Método MPS y MDS de Veregin

Veregin (2000) plantea una metodología para estudiar los efectos de la generalización (simplificación) y más concretamente del algoritmo de *Douglas-Peucker*⁴ en la exactitud posicional de elementos lineales. El autor propone dos variantes, el MPS o estándar de máxima proporción, y el MDS o estándar de máxima distorsión. Atkinson y Ariza (2002) plantean sus principales características:

- Aplicaciones corrientes: Posible implementación en el control de procesos de generalización y calidad posicional.
- Método de comparación: Con fuentes de mayor exactitud.
- Componente posicional: Horizontal.
- Clase de elementos: Línea.
- Correspondencia con un estándar de exactitud predefinido: El MPS pretende asemejarse al NMAS (ver Test NMAS en Anexo I punto AI.2) en su aplicación a los elementos lineales. El MDS se asemeja al test MDE del USGS (ver Test del USGS de categorías de exactitud de MDE en Anexo I punto AI.7). También trabaja con la superficie comprendida entre las líneas, al igual que ocurre con el test de banda epsilon (Métodos de control de exactitud posicional en elementos lineales)
- Método de banda epsilon II.4.2.1)
- Descripción: La metodología propuesta está basada en el cálculo de la *distancia de distorsión uniforme* (DDU). Esta distancia viene definida por la superficie diferencia de entre las líneas por unidad de longitud, de forma que se calculan las áreas entre las dos líneas en relación a la longitud de la línea del mapa:

$$DDU = \frac{\sum S(Q, X)}{Long_X}$$

Ec. II. 19

Siendo $S(Q, X)$ la superficie formada entre la línea del mapa y la de la fuente de mayor exactitud y $Long_X$ la longitud total de la línea del mapa.

⁴ Algoritmo Douglas-Peucker: Algoritmo de simplificación de elementos lineales publicado en 1973 incluido en la mayoría de los software comerciales y empleado por la mayor parte de las instituciones cartográficas.

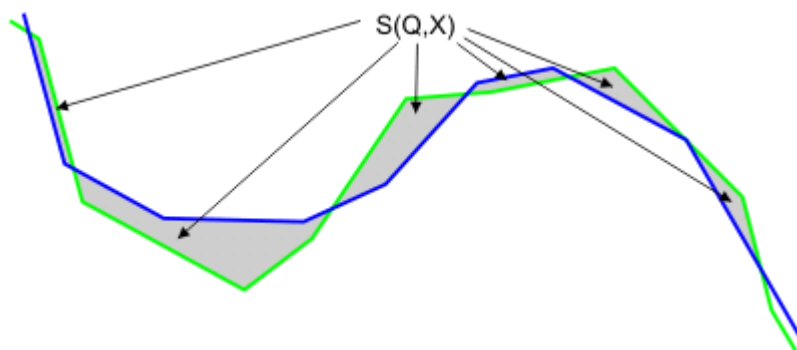


Figura II. 25. Test MPS y MDS de Veregin

Si se aplican varios anchos de banda al algoritmo de *Douglas-Peucker*⁵ sobre una base de datos lineales, cuando se calcula la *distancia de distorsión uniforme* (DDU) se obtienen una distribución de frecuencias acumuladas.

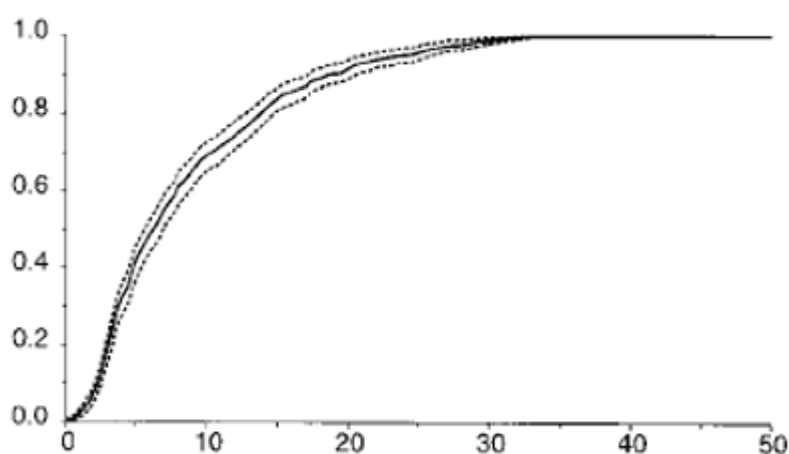


Figura II. 26. Test MPS y MDS: Distribución frecuencias acumuladas. Veregin (2000)

En la Figura II. 26 aparece la distribución de frecuencias acumuladas de las carreteras estudiadas por Veregin (2000) con 50 m de anchura de banda. La línea discontinua representa el límite de confianza a un 95%, en el eje "X" se representa la distancia de distorsión uniforme y en el eje "Y", la proporción acumulada de líneas.

- Procedimiento:
 1. Seleccionar una muestra de líneas en el mapa (X).

⁵ El ancho de banda en el algoritmo Douglas-Peucker es la tolerancia que permite eliminar los puntos de la línea que se encuentre dentro de la misma.

2. Seleccionar estas mismas líneas en la fuente de mayor exactitud (Q).
3. Calcular la distancia DDU para diferentes tolerancias o anchos de banda del algoritmo de *Douglas-Peucker*.
4. Construir los diagramas de frecuencia acumulada en función de los intervalos de confianza para cada ancho de banda.
5. Estimación de los estándares MPS y MDS:

- a) MPS: Estándar de Máxima Proporción: Se limita el porcentaje de elementos lineales que superen una distorsión mayor que un determinado umbral. Las curvas de frecuencia acumulada se obtienen con una muestra significativa del tipo de líneas a controlar.

Según el estándar NMAS, para una escala 1:100.000, la tolerancia marcada es de 50.8 m sobre el terreno. Para la muestra empleada, con una tolerancia de 129 m en el algoritmo de *Douglas-Peucker* (para $1-\alpha = 0.95$) al menos el 90% de las líneas tienen una posición que dista menos de 50.8 m de su posición original.

- b) MDS: Estándar de Máxima Distorsión: Se determina un umbral de distorsión a priori que no debe ser superado. Empleando diferentes tolerancias sobre varias líneas se procede a detectar para cada una la mayor tolerancia que se puede alcanzar para un nivel de calidad deseado.

Por ejemplo, con un umbral 1.25 veces la tolerancia marcada por el NMAS ($1.25 \times 50.8 = 63.5$ m) y un ancho de banda de 129 m se pueden detectar aquellas líneas con distorsiones mayores a 50.8 m. Para cada una de ellas, la anchura de banda resultante del algoritmo *Douglas-Peucker*, obtienen una distorsión máxima de 63.5. Sobre la muestra empleada, esta distorsión se obtiene con 89 m.

- Discusión: Aunque su fundamento es similar al test de banda ϵ , esta metodología añade el concepto de diagrama de frecuencia acumulada. De esta forma, se trabaja dentro de un determinado nivel de confianza, obteniendo, por tanto, el porcentaje de líneas que se encuentran dentro de un determinado umbral de error, así como los máximos niveles de generalización que se pueden aplicar a dicha línea para que su error posicional sea admisible. Esta herramienta permite aplicar esta metodología al control posicional en procesos de generación de elementos lineales. Sin embargo, el test no ofrece información acerca de la procedencia del error (sistemático o accidental).

Veregin (2000) realiza su estudio sobre datos del *United States Geological Survey* (USGS) y más concretamente del *Digital Line Graphs* (DLGs) a escala 1:100.000.

Este producto cubre diferentes temáticas (hidrografía, vías de comunicación, ...) de las que se seleccionan 30 líneas aleatoriamente de carreteras y arroyos. Se aplica sobre esta selección el algoritmo de simplificación de líneas de *Douglas-Peucker* con varios anchos de banda desde 2 m hasta 250 m.

Por ejemplo, considerando una amplitud de banda de 2m para la base de datos de arroyos, aproximadamente el 85 % de líneas tiene una distorsión de 0.05 m o menos, aproximadamente el 95 % tiene una distorsión de 0.1 m o menos, etc.

En los resultados se puede destacar que para cualquier valor de anchura de banda, la distorsión media y máxima son menores que esta anchura.

Veregin (2000) detecta importantes diferencias en los resultados de las líneas procedentes de carreteras y de arroyos, debido fundamentalmente a que las carreteras no presentan meandros, por lo que no están tan afectadas por valores bajos de anchura de banda. Las carreteras tienen 'un margen dinámico' mucho más estrecho en el sentido que valores de amplitud de banda bajos tienen poco efecto, pero los niveles máximos de simplificación son alcanzados usando sólo valores moderados de amplitud de banda.

II.4.2.6 Método de correspondencia de puntos

Kagawa, Sekimoto y Shibaski (1999) presentan un estudio comparativo de varias metodologías. Entre ellas, utilizan el denominado por ellos como método de *correspondencia de puntos* (*Point Correspondence Method*) y que está basado en los trabajos de Ramírez, Bidoshi, Douglas, Phuyal y Szakas (1996).

- Aplicaciones corrientes: Posible implementación en el control de procesos de generalización y calidad posicional.
- Método de comparación: Con fuentes de mayor exactitud.
- Componente posicional: Horizontal.
- Clase de elementos: Línea.
- Correspondencia con un estándar de exactitud predefinido: No se ha elaborado un estándar hasta la fecha.
- Descripción: Se generan puntos sobre la línea del mapa (X) con un intervalo constante y conectan estos puntos con otros generados sobre la línea de mayor exactitud (Q). De esta forma aparecen los *vectores virtuales de error* (Figura II. 27). Los puntos virtuales generados en la línea de la fuente de mayor exactitud son determinados minimizando la suma los cuadrados de las longitudes de estos vectores de error.

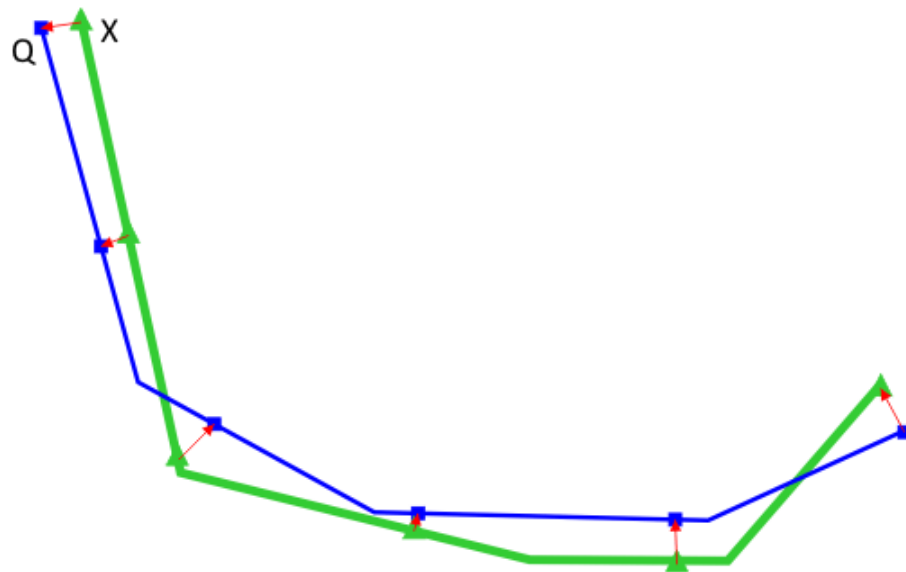


Figura II. 27. Método de correspondencia de puntos

- Procedimiento:
 1. Generar los puntos virtuales a lo largo de la línea a controlar con un intervalo constante.
 2. Desde estos puntos, calcular una aproximación de los vectores de error a la línea de referencia.
 3. Determinar o ajustar los vectores de referencia minimizando la suma de los cuadrados de las longitudes.
 4. Durante este procedimiento, establecer condiciones para ordenar los puntos finales de los vectores de error en ambas líneas.
 5. Generar un gráfico ordenado de magnitud de errores en cada punto estableciendo el sentido positivo del eje "Y" para los vectores que apuntan a la izquierda siguiendo la dirección de la medida. De esta forma se obtiene una idea de la tendencia de los errores.
 6. Calcular estadísticos sobre los valores de los errores obtenidos.
- Discusión: Como los errores son evaluados en términos de mínimos cuadrados y asumiendo que los mismos siguen una distribución normal o gaussiana, el cuadrado de los errores se corresponde con la probabilidad de los mismos. Con este método, se pueden estimar sistematismos fácilmente. Sin embargo, cuando no hay puntos, el algoritmo determinará puntos interpolados. El método es interesante aunque exige un intervalo de puntos constante pequeño, por lo que otro posible problema es que los puntos que se utilizan provienen de la interpolación de puntos medidos (vértices de la línea original) pero no existen realmente.

Kagawa, Sekimoto y Shibaski (1999), en su estudio de carreteras en la ciudad de Sagami-hara (Japón) aplican esta metodología. Para una carretera de 706 m generan 100 puntos a intervalos constantes y sus vectores a la línea de referencia (mayor exactitud) minimizando la suma de los cuadrados de las longitudes de los vectores. Los resultados presentan un sistematismo hacia la izquierda en el sentido de medida y un error medio de 0.68 m.

Schmidley (1996) presenta una metodología relacionada con la anterior, basada en la comparación de 2 segmentos a través de la distancia entre los puntos finales de los tramos. Una línea es dividida en tramos de una distancia fija y el número de divisiones es trasladado a la segunda línea. Los tramos son enlazados ordenadamente calculando las distancias entre los puntos finales. Ramírez y otros (1996) aplican este método a distintas líneas (carreteras, ferrocarriles, límites de parcelas, hidrografía y límites administrativos), utilizando una longitud fija para dividir la primera línea de 60 m.

II.4.3. Errores de orlados

Los orlados o búfers sobre elementos lineales, como se ha visto anteriormente, son una herramienta utilizada para estudiar la exactitud posicional de estos elementos. Un orlado es una banda de anchura constante alrededor de una línea. Las aplicaciones de esta figura en el control de exactitud posicional se han estudiado anteriormente en este capítulo, sin embargo, es interesante contrastar algunos conceptos, como el estudio de sus errores que, generalmente, han sido desarrollados en los últimos años.

Shi, Cheung y Zhu (2003) estudian un modelo de propagación de errores en el control por orlados en el que se tienen en cuenta el error posicional de la línea y los propios del búfer. El error posicional de un búfer es estimado por vía del error de comisión, el error de omisión, y el área discrepante entre las posiciones 'verdaderas' y 'medidas' del búfer. Basado en la suposición de distribución de errores de un elemento espacial y de la anchura del búfer, se pueden expresar los valores esperados del error de comisión, el error de omisión, y el área discrepante como múltiples integrales. Estos modelos analíticos se solucionan por un método numérico que proporciona una aproximación a los valores esperados.

Los errores de comisión y de omisión son adoptados de los estudios de exactitud temática en los procesos de clasificación en teledetección. Sin embargo, cuando la posición de dos búfers medidos no intersecta con la verdadera posición, el error de comisión y de omisión no muestran qué búfer tiene mayor error. En estos casos se define la distancia entre el búfer "verdadero" y los medidos como el *área de discrepancia* para indicar el error posicional del búfer.

Shi y Cheung (2000) definen el error de omisión como la no existencia de una parte del área del búfer esperado y el error de comisión como la identificación de una parte del área como una parte del búfer esperado cuando en realidad no es parte de este búfer (Figura II. 28). Estos conceptos son ampliados posteriormente por Shi, Cheung y Zhu (2003).

Sadahiro (2005) define el error de omisión como al que se corresponde a la situación donde un punto de un búfer está observado incorrectamente fuera del búfer

verdadero, mientras que el error de comisión se produce cuando se sitúa un punto exterior a un búfer en su interior (Figura II. 28).

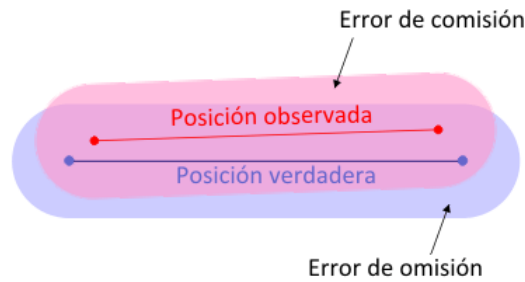


Figura II. 28. Error de comisión y omisión en búfer

En un elemento puntual al igual que en un segmento, el error de comisión se determina restando la intersección del búfer actual y el búfer medido del búfer medido. El error de omisión se obtiene restando la intersección del actual y el búfer medido del búfer actual. Los valores de los errores para las diferentes figuras según Shi, Cheung y Zhu (2003) son:

- Puntos:

$$E_C = \iint f(x_1, y_1, w) \text{Area} (A_{P^*}(1) - A_P(1)) dx_1 dy_1 dw$$

$$E_O = \iint f(x_1, y_1, w) \text{Area} (A_P(1) - A_{P^*}(1)) dx_1 dy_1 dw$$

Ec. II. 20

$A_P(1)$ se refiere a un grupo de puntos dentro de un búfer alrededor de un punto esperado $P(1) = (\mu_{x1}, \mu_{y1})$ y anchura de búfer μ_w .

$A_{P^*}(1)$ se refiere a un grupo de puntos dentro de un búfer alrededor de un punto medido $P^*(1) = (x_1, y_1)$ con anchura de búfer w .

$f(x_1, y_1, w)$ se refiere a la función de densidad de probabilidades de las variables (x_1, y_1, w) .

- Segmentos:

$$E_C = \int \dots \int f(x_1, y_1, x_2, y_2, w) \text{Area} (A_{L^*}(1) - A_L(1)) dx_1 dy_1 dx_2 dy_2 dw$$

$$E_O = \int \dots \int f(x_1, y_1, x_2, y_2, w) \text{Area} (A_L(1) - A_{L^*}(1)) dx_1 dy_1 dx_2 dy_2 dw$$

Ec. II. 21

$A_L(1)$ se refiere a un grupo de puntos dentro de un búfer alrededor de un punto esperado “i” de $L(1)$.

$A_{L^*}(1)$ se refiere a un grupo de puntos dentro de un búfer alrededor de un punto medido “i” de $L^*(1)$.

$f(x_1, y_1, x_2, y_2, w)$ se refiere a la función de densidad de probabilidades de las variables (x_1, y_1, x_2, y_2, w) .

- Polilíneas:

$$E_C = \int \dots \int f(x_1, y_1, \dots, x_{NP}, y_{NP}, w) \text{Area} (A_{Mult-L^*} - A_{Mult-L}) dx_1 dy_1 \dots dx_{NP} dy_{NP} dw$$

$$E_O = \int \dots \int f(x_1, y_1, \dots, x_{NP}, y_{NP}, w) \text{Area} (A_{Mult-L} - A_{Mult-L^*}) dx_1 dy_1 \dots dx_{NP} dy_{NP} dw$$

Ec. II. 22

A_{Mult-L} se refiere a un grupo de puntos dentro de un búfer alrededor de un línea esperada que tiene “NP” nodos.

A_{Mult-L^*} se refiere a un grupo de puntos dentro de un búfer alrededor de un línea medida que tiene “NP” nodos.

$f(x_1, y_1 \dots x_{NP}, y_{NP}, w)$ se refiere a la función de densidad de probabilidades de las variables $(x_1, y_1 \dots x_{NP}, y_{NP}, w)$.

- Polígonos:

$$E_C = \int \dots \int f(x_1, y_1, \dots, x_{NP}, y_{NP}, w) \text{Area} (A_{R^*} - A_R) dx_1 dy_1 \dots dx_{NP} dy_{NP} dw$$

$$E_O = \int \dots \int f(x_1, y_1, \dots, x_{NP}, y_{NP}, w) \text{Area} (A_R - A_{R^*}) dx_1 dy_1 \dots dx_{NP} dy_{NP} dw$$

Ec. II. 23

A_R se refiere a un grupo de puntos dentro de un búfer alrededor de un polígono esperado que tiene “NP” nodos.

A_{R^*} se refiere a un grupo de puntos dentro de un búfer alrededor de un polígono medido que tiene “NP” nodos.

$f(x_1, y_1 \dots x_{NP}, y_{NP}, w)$ se refiere a la función de densidad de probabilidades de las variables $(x_1, y_1 \dots x_{NP}, y_{NP}, w)$.

El área de discrepancia para un punto es la superficie diferencia entre el búfer esperado y el medido. La discrepancia de búfer de un segmento puede ser determinada por la discrepancia de todos los puntos esperados y medidos del segmento (Figura II. 29).

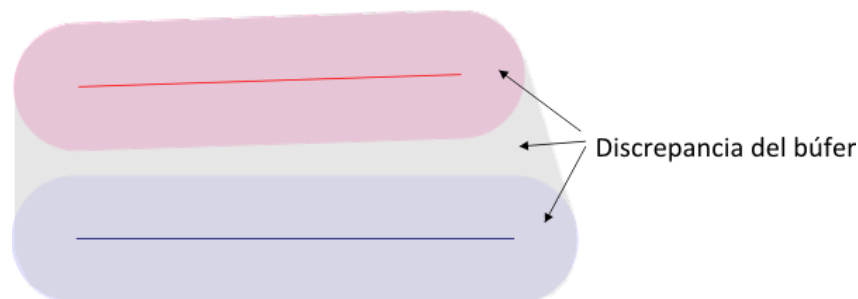


Figura II. 29. Discrepancia del búfer

En este apartado, se han presentado estudios más recientes referentes a los errores existentes y su propagación en los orlados sobre elementos lineales. Hasta el momento, estos planteamientos constituyen una importante base teórica para describir el comportamiento del error en los orlados, no habiendo, aplicación de los mismos sobre datos reales hasta la fecha.

Sin embargo, los resultados de este estudio están muy relacionados con una metodología ya implementada de control de calidad posicional. En concreto, se trata del Método del orlado doble, donde se obtienen valores de errores mediante diferencias entre áreas de orlados.

Los conceptos de error de omisión y comisión pueden igualmente trasladarse al método del orlado doble. De hecho, son resultantes del mismo como la superficie de un orlado que no intersecta con el otro y viceversa.

II.5. Muestreo

Resulta obvio que la forma ideal de realizar un control sobre un proceso es analizar el comportamiento de todos y cada uno de los elementos que intervienen. Sin embargo, este planteamiento resulta inviable debido al gran esfuerzo que supone, por lo que generalmente se recurre al muestreo. La ventaja fundamental que plantea la realización de muestreos es el ahorro económico que se obtiene al reducir los trabajos de comprobación de la calidad a un número más reducido de elementos.

Obteniendo una muestra de un conjunto de elementos “n”, se abarca el control de calidad del total de elementos “N”. Además se introduce unos valores límites de aceptación previamente definidos que aplicados a la muestra permita tomar la decisión de aceptación o rechazo del producto (Ariza, 2002a).

Sin embargo, el muestreo, al trabajar con un planteamiento estadístico, introduce riesgos. Se puede hablar del riesgo del productor y del cliente o usuario:

- Riesgo del productor: Probabilidad de que se le rechace un trabajo bueno.

- Riesgo del cliente o usuario: Probabilidad de aceptar como bueno un trabajo que realmente es malo.

La determinación de los tamaños muestrales se realiza bajo ciertas hipótesis estadísticas sobre la población origen, basándose en ciertas propiedades a determinar y en la precisión y riesgo que se asumen. La componente económica es fundamental teniendo en cuenta tanto el coste de los trabajos de muestreo como el coste de una mala decisión. (Ariza, García y Amor, 2004).

El muestreo puede aplicarse de diferentes formas (discontinuo, porcentual, estadístico,...), seleccionando los elementos de varias maneras (extracción aleatoria, estratificada, por conglomerados,...) pudiendo afectar a distintos elementos (variables, atributos, mixtas,...).

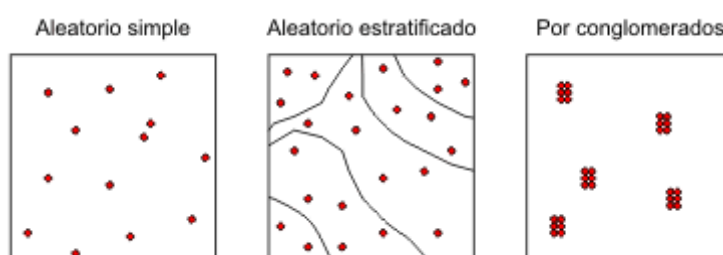


Figura II. 30. Métodos de selección de muestras

La aplicación del muestreo estadístico a los elementos puntuales para el control de calidad posicional ha sido ampliamente utilizada hasta la fecha, ya que consiste en trasladar la metodología existente para la sección de muestras a estos elementos. Sin embargo, para el caso de elementos lineales, no existen recomendaciones sobre el muestreo a realizar para la selección de líneas.

Por otro lado, las metodologías de control posicional por elementos lineales previamente descritas no incluyen recomendaciones acerca de la selección de muestras, limitándose, a la presentación de un método de control sin profundizar en la selección de elementos necesarios para su aplicación.

Por tanto, la elección de una metodología de selección de muestras para elementos lineales es un punto pendiente de estudio.

A lo largo de este punto se describen los principales métodos de selección de muestras con la finalidad de establecer una base de partida para su posible implementación y análisis sobre los elementos lineales.

De forma general, la selección de muestras es ampliamente descrita por numerosos autores: Cochran (1977), Pérez (1999), etc.

II.5.1. Muestreo aleatorio simple

Se trata de un procedimiento de selección con probabilidades iguales consistente en la obtención de la muestra, unidad a unidad, aleatoriamente (Pérez, 1999). Dependiendo de la existencia o no de reposición se pueden considerar:

- Muestreo sin reemplazamiento: Dada una población finita, se obtiene una muestra unidad a unidad y sin reposición. Todas las posibles muestras tienen la misma probabilidad de ser elegidas. En este caso el tamaño de la muestra para estimar la media y el total para muestreos sin reemplazamiento viene dado por la Ec. II. 24 y la Ec. II. 25 (Ariza, García y Amor, 2004).

Total	Media
$n = \frac{n_0}{1 + \frac{n_0}{N}}$ $\text{donde : } n_0 = \frac{N^2 k^2 S^2}{\varepsilon^2}$ Ec. II. 24	$n = \frac{n_0}{1 + \frac{n_0}{N}}$ $\text{donde : } n_0 = \frac{k^2 S^2}{\varepsilon^2}$ Ec. II. 25

Donde: N es el tamaño de la población, S^2 es la varianza, ε es el error máximo admisible, k es una constante en función del nivel de confianza.

- Muestreo con reemplazamiento: En este caso se parte de una población finita, de la que se obtiene una muestra unidad a unidad, pero con reposición de cada unidad a la población después de cada extracción. Para este tipo de muestreo el tamaño muestral vendrá dado por las expresiones Ec. II. 26 y Ec. II. 27.

Total	Media
$n = \frac{N^2 k^2 S^2}{\varepsilon^2}$ Ec. II. 26	$n = \frac{k^2 S^2}{\varepsilon^2}$ Ec. II. 27

Donde: N es el tamaño de la población, S^2 es la varianza, ε es el error máximo admisible y k es una constante en función del nivel de confianza.

II.5.2. Muestreo estratificado

Este método se utiliza cuando se desea que ciertas clases aparezcan en el muestreo, por lo que supone un cierto conocimiento a priori de la población. La población heterogénea es dividida en “ L ” subpoblaciones (estratos) lo más homogéneas posibles. Cada estrato tiene un tamaño poblacional N_h , no existiendo

superposición de los mismos, por lo que la suma de los estratos forma la totalidad de la población. En este tipo de muestreo de cada estrato se extrae una muestra de tamaño n_h .

Cada estrato tiene un peso dado por $W_h = N_h/N$ y una fracción de muestreo dada por $f_h = n_h/N_h$ y sus respectivas medias poblacionales, medias muestrales, varianzas poblacionales y varianzas muestrales.

Los tamaños muestrales se definen por lo general de 4 formas posibles:

- Afijación⁶ uniforme: Todos los n_h son iguales (Ec. II. 28).

$$n_h = \frac{n}{L}$$

Ec. II. 28

- Afijación proporcional: El valor de cada n_h es proporcional al tamaño del estrato, por lo que todas las unidades tienen la misma probabilidad de ser seleccionadas (Ec. II. 29).

$$n_h = N_h * cte$$

Ec. II. 29

- Afijación de mínima varianza: Los n_h se eligen de manera que se minimice la varianza para n fijo (Ec. II. 30).

$$n_h = \frac{N_h S_h}{\sum_{h=1}^L N_h S_h} n$$

Ec. II. 30

Donde S_h es la desviación típica del estrato.

- Afijación óptima: Los n_h se eligen de manera que se minimice la varianza para un coste fijo. Por tanto habrá que definir una función de coste (C).

$$C = C_0 + \sum_{h=1}^L n_h C_h$$

Ec. II. 31

⁶ Afijación: Reparto, asignación, adjudicación adscripción o distribución del tamaño muestral entre los diferentes estratos (Pérez, 1999).

$$n_h = \frac{N_h S_h \frac{1}{\sqrt{C_h}}}{\sum_{h=1}^L N_h S_h \frac{1}{\sqrt{C_h}}}$$

Ec. II. 32

El tamaño de la muestra vendrá dado por la Ec. II. 33 y la Ec. II. 34.

Total	Media
$n = \frac{\sum_{h=1}^L \frac{N_h^2 S_h^2}{w_h}}{\frac{\varepsilon^2}{k^2} \sum_{h=1}^L N_h S_h^2}$	$n = \frac{\sum_{h=1}^L \frac{W_h^2 S_h^2}{w_h}}{\frac{\varepsilon^2}{k^2} \sum_{h=1}^L W_h S_h^2}$
Ec. II. 33	Ec. II. 34

II.5.3. Muestreo por conglomerados

Al realizar un muestreo en una población grande o muy dispersa, los costes del muestreo y la duración del trabajo de campo pueden aumentar extraordinariamente si se utilizan los diseños anteriormente citados.

En el muestreo por conglomerados las unidades de muestreo comprenden dos o más unidades que constituyen el objeto de estudio. El método está basado en la división de la población en conglomerados, de los cuales se selecciona cierto número. La muestra estará compuesta por todos los elementos que componen el conglomerado seleccionado. Por lo tanto, para que la muestra sea representativa, los elementos que componen el conglomerado deben ser muy heterogéneos.

En general, a igual tamaño de muestra, el muestreo por conglomerados es menos eficiente que el muestreo estratificado, e incluso que el aleatorio simple. Sin embargo, a igual coste de muestreo el diseño por conglomerados permite obtener tamaños muestrales mayores que los otros diseños, y por lo tanto aumentar la eficacia real de un estudio por muestreo. La eficacia se mejorará tomando un número elevado de conglomerados y eligiéndolos de forma que sean lo más heterogéneos posible.

En principio, la división de la población en conglomerados es un problema previo y se realiza en función de criterios económicos, por lo que si consideramos conocidos los conglomerados y sus tamaños se obtendrá un número de conglomerados que vendrá dado por la expresión Ec. II. 35.

$$m = \frac{MS_c^2}{S_c^2 + M \frac{\varepsilon^2 \bar{N}^2}{k^2}}$$

Ec. II. 35

Donde: M es el total de conglomerados, S_c^2 es la varianza por conglomerados, ε es el error máximo admisible, k es una constante en función del nivel de confianza y $\bar{N} = N/M$ es la media poblacional de la variable.

Se pueden distinguir varios tipos de muestreo por conglomerados:

Muestreo Unietápico: En el que se seleccionan los conglomerados recogiendo la información del 100% de las unidades elementales contenidas en cada conglomerado.

Muestreo Bietápico: En el que una vez seleccionados los conglomerados, se realiza una segunda elección sobre los elementos pertenecientes a cada uno de ellos.

Muestreo Polietápico: En el que una vez seleccionados los conglomerados, se realizan varias sub-selecciones de conglomerados y así sucesivamente hasta alcanzar la última etapa donde se seleccionan los elementos pertenecientes a cada uno de ellos.

II.6. Variables aleatorias y distribuciones de probabilidad

Los métodos de control de calidad posicional por elementos lineales basados en orlados presentados previamente (Método del orlado simple y Método del doble orlado) generan como resultado funciones de distribución relativas a la anchura de orlado que se utilice. El análisis de estos resultados debe incluir, por tanto, el estudio de su posible ajuste a una distribución de probabilidad estandarizada.

El análisis de la similaridad de dos funciones de distribución se utiliza para establecer comparaciones entre las mismas. Se realiza a través del Test de Kolmogorov-Smirnov, analizado más ampliamente en el apartado de metodología (Capítulo III.6). Con esta técnica, es posible identificar si dos funciones de distribución son similares.

Como base para el ajuste planteado, en este punto, se describe el concepto de variable aleatoria y cuales son las distribuciones de probabilidad estandarizadas más habitualmente utilizadas.

Una variable aleatoria se define como el resultado numérico de un experimento aleatorio. Matemáticamente, X es una función medible $X : \Omega \rightarrow \Re$ que da un valor numérico, del conjunto de los reales, a cada suceso (espacio Ω) del espacio muestral del experimento.

Se dice que una variable aleatoria X es discreta si el número de valores que puede tomar es contable y si estos pueden arreglarse en una secuencia que corresponde con los enteros positivos, mientras que una variable aleatoria continua

es, aquella que, toma valores en uno o más intervalos de la recta de los reales (Canavos, 2003).

La variable aleatoria discreta X representa los resultados de un espacio muestral de forma que es posible desarrollar una función matemática que asigne una probabilidad a cada realización de la variable aleatoria. Esta función se denomina función de probabilidad de la variable aleatoria X . Por otro lado, la distribución de probabilidad se refiere a la colección de valores de la variable aleatoria y a la distribución de probabilidades de éstos, teniendo en cuenta la existencia de la función de distribución acumulativa de X . La variable aleatoria discreta X está caracterizada por la función de probabilidad puntual que determina la probabilidad puntual de que $X = x$, y por la función acumulativa $F(x)$, que representa la suma de las probabilidades puntuales (Canavos, 2003).

En el caso discreto, se asignan probabilidades positivas a los valores puntuales de la variable aleatoria, siendo la suma de todas uno. En el caso continuo, este caso no es posible, siendo la probabilidad de que una variable aleatoria continua X tome un valor x igual a cero. Es lógico, por tanto, trabajar con intervalos, más que con puntos.

La distribución de probabilidad de una variable aleatoria continua X está caracterizada por la función de densidad de probabilidad $f(x)$. La función $f(x)$ presenta una gráfica con una curva límite que se obtiene para un número muy grande de observaciones y para una amplitud de intervalo muy pequeña. La escala vertical es elegida de forma que el área total bajo la curva sea igual a uno.

Basadas en este concepto de distribución continua de probabilidad, se presentan a continuación algunas distribuciones utilizadas habitualmente en disciplinas como la ingeniería, las ciencias o la economía (Canavos, 2003):

- Distribución normal o gaussiana: Es la más importante y de mayor uso ya que las distribuciones de muchas muestras tienden hacia este tipo a medida que se aumenta el tamaño de la muestra. Gráficamente, se presenta como una curva simétrica con forma de campana, que se extiende sin límite en las direcciones positiva y negativa. Una variable aleatoria se encuentra normalmente distribuida si su función de densidad de probabilidad está dada por la Ec. II. 36.

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

Ec. II. 36

donde $-\infty < x < \infty$; $-\infty < \mu < \infty$; $\sigma > 0$.

Los parámetros μ y σ determinan la función de densidad de probabilidad.

La Figura II. 31 muestra un ejemplo de la función de densidad de probabilidad normal.

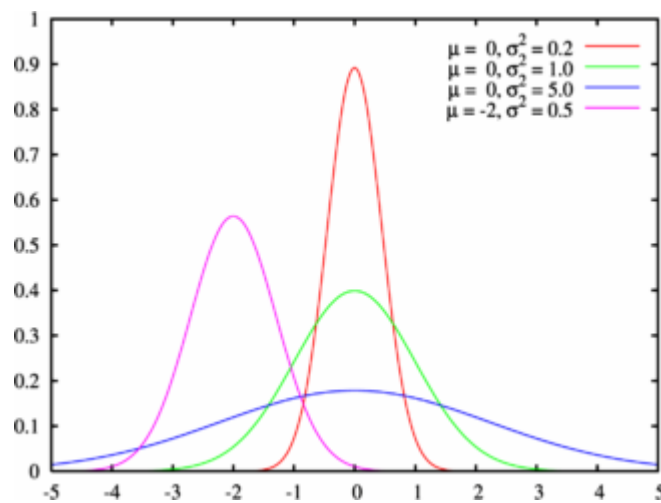


Figura II. 31. Función de densidad de probabilidad. Distribución normal

- Distribución uniforme: Supone que una variable aleatoria toma valores de un intervalo finito, de manera que éstos se encuentran distribuidos igualmente sobre el intervalo, por lo que la probabilidad de que ésta tome un valor en cada subintervalo de igual longitud es la misma. Una variable aleatoria se encuentra distribuida uniformemente sobre el intervalo (a, b) si su función de densidad de probabilidad está dada por la Ec. II. 37.

$$f(x; a, b) = \begin{cases} \frac{1}{(b-a)} & a \leq x \leq b \\ 0 & \text{para otro valor} \end{cases}$$

Ec. II. 37

La Figura II. 32 muestra un ejemplo de la función de densidad de probabilidad uniforme.

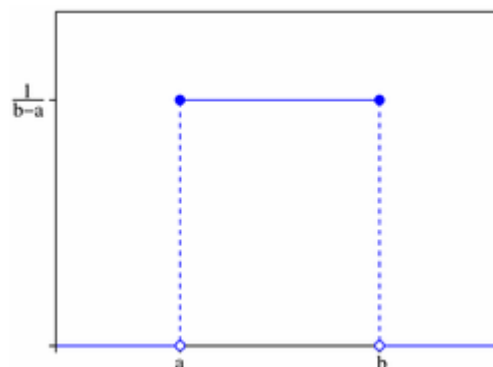


Figura II. 32. Función de densidad de probabilidad. Distribución uniforme

- Distribución beta: Una variable aleatoria presenta una distribución beta si su función de densidad de probabilidad está dada por la Ec. II. 38.

$$f(x; \alpha, \beta) = \begin{cases} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} & 0 < x < 1; \alpha, \beta > 0 \\ 0 & \text{para otro valor} \end{cases}$$

Ec. II. 38

En función de los parámetros α y β se obtendrán diferentes perfiles (Figura II. 33).

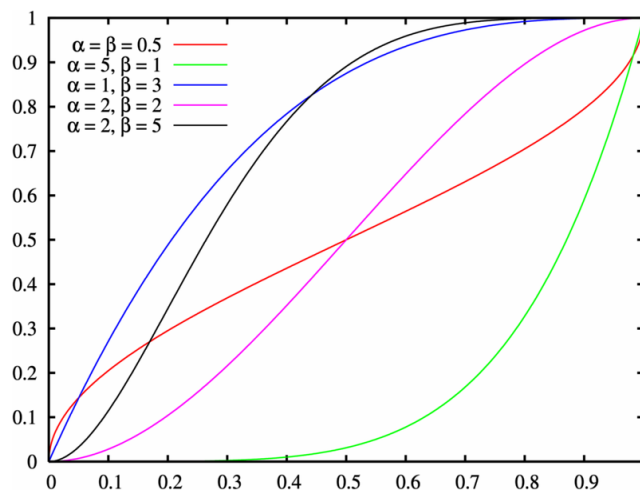


Figura II. 33. Función de densidad de probabilidad. Distribución beta

- Distribución gamma: La función de densidad de una distribución gamma viene dada por la Ec. II. 39.

$$f(x; \alpha, \theta) = \begin{cases} \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} \exp(-x/\theta) & x > 0; \alpha, \theta > 0 \\ 0 & \text{para otro valor} \end{cases}$$

Ec. II. 39

En función del parámetro α y θ se obtendrán diferentes perfiles (Figura II. 34).

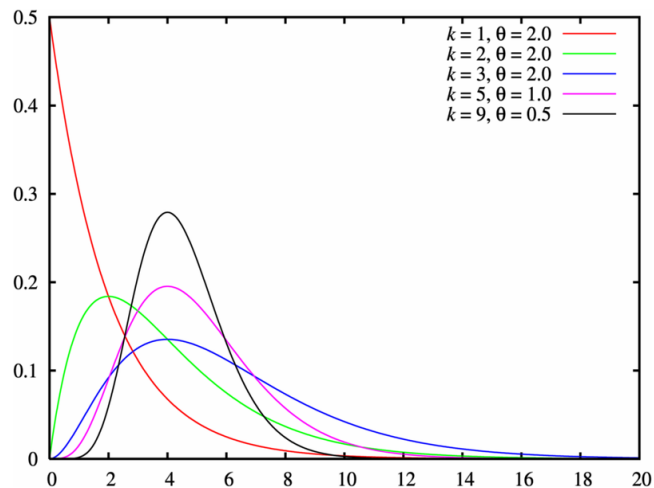


Figura II. 34. Función de densidad de probabilidad. Distribución gamma

- Distribución de Weibull: Una variable aleatoria presenta una distribución Weibull si su función de densidad de probabilidad está dada por la Ec. II. 40.

$$f(x; \alpha, \theta) = \begin{cases} \frac{\alpha}{\theta^\alpha} x^{\alpha-1} \exp\left[-(x/\theta)^\alpha\right] & x > 0; \alpha, \theta > 0 \\ 0 & \text{para otro valor} \end{cases}$$

Ec. II. 40

En función de los parámetros forma (α) y escala (θ) se obtendrán diferentes perfiles (Figura II. 35).

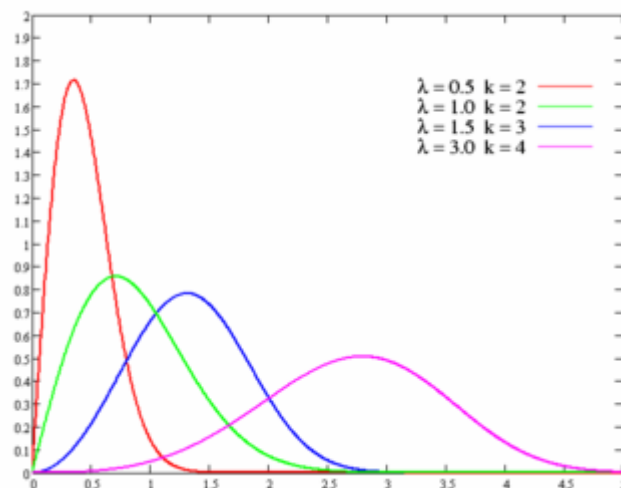


Figura II. 35. Función de densidad de probabilidad. Distribución Weibull

- Distribución exponencial: La función de densidad de una distribución exponencial viene dada por la Ec. II. 41.

$$f(x; \lambda) = \begin{cases} \lambda \exp(-\lambda x) & x \geq 0 \\ 0 & \text{para otro valor} \end{cases}$$

Ec. II. 41

La Figura II. 36 muestra un gráfico de la función de densidad de probabilidad exponencial según el parámetro λ .

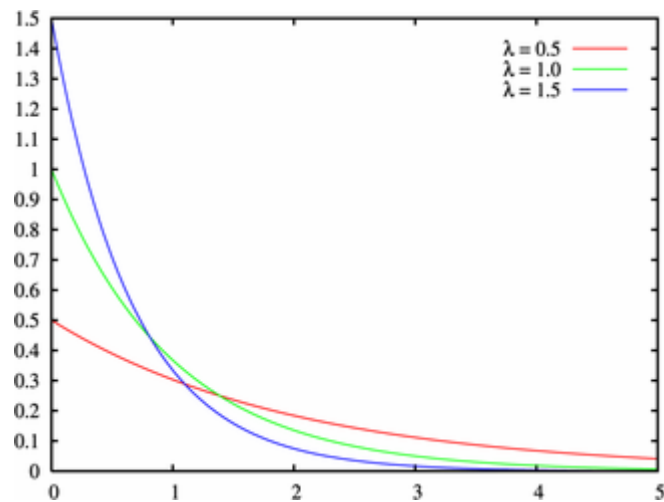


Figura II. 36. Función de densidad de probabilidad. Distribución exponencial

Capítulo III.

Metodología y Material

Capítulo III. Metodología y Material

III.1. Introducción

A lo largo del presente capítulo se aborda con profundidad los aspectos metodológicos de la tesis doctoral y se describen los materiales utilizados con los que se ha ensayado la metodología propuesta.

Con anterioridad, se han presentado algunos de los conceptos básicos y el estado actual de la cuestión en lo relativo al control de calidad posicional en cartografía. Como ya se apuntó en la introducción del presente documento, tradicionalmente, se ha realizado el control de calidad en cartografía a través de elementos puntuales presentes en la misma. En la actualidad, está apareciendo en la comunidad científica una fuerte inquietud por el estudio y la aplicación de nuevas metodologías de control a través de elementos lineales. Dentro de esta línea de investigación es donde se desarrolla la presente tesis doctoral.

En el capítulo anterior se han comentado y analizado las primeras propuestas metodológicas aportadas por los investigadores que trabajan en el control de calidad cartográfica con elementos lineales. Estas aportaciones se han limitado al desarrollo de una metodología básica muy concreta y su ensayo con pequeños conjuntos de datos. En el capítulo que nos ocupa, se propone una metodología completa para la realización de dicho control, que engloba los métodos presentados previamente, complementándolos y dejando la posibilidad de incorporar nuevos. Para ello, se describen las principales decisiones adoptadas en el control posicional por elementos lineales, utilizando para ello herramientas como el control automatizado de los

mismos, la simulación, la caracterización y similitud entre funciones de distribución, etc.

La metodología de control de calidad posicional propuesta en esta tesis doctoral se centra en la componente horizontal de las bases de datos geográficas, no entrando en el estudio de la componente vertical del dato geográfico.

El comportamiento de la componente vertical es tratado tradicionalmente de forma independiente y podrá ser objeto de futuras investigaciones a partir de la base desarrollada en esta tesis doctoral. De hecho, los métodos de control por elementos lineales propuestos por la comunidad científica (descritos en el capítulo anterior) estudian exclusivamente esta componente horizontal.

Todos estos aspectos metodológicos obtienen su refuerzo científico con la aplicación práctica de los mismos, para lo cual se han utilizado por un lado dos bases de datos cartográficas reales, cuyas vías de comunicación han sido controladas con fuentes de mayor exactitud (GPS), y por otro lado un conjunto de líneas sintéticas con errores simulados para estudiar su comportamiento individualizado. Estos conjuntos de datos utilizados para la aplicación metodológica son presentados también en el presente capítulo.

III.2. Procedimiento General

En la Figura III. 1 se presenta un esquema general del procedimiento seguido.

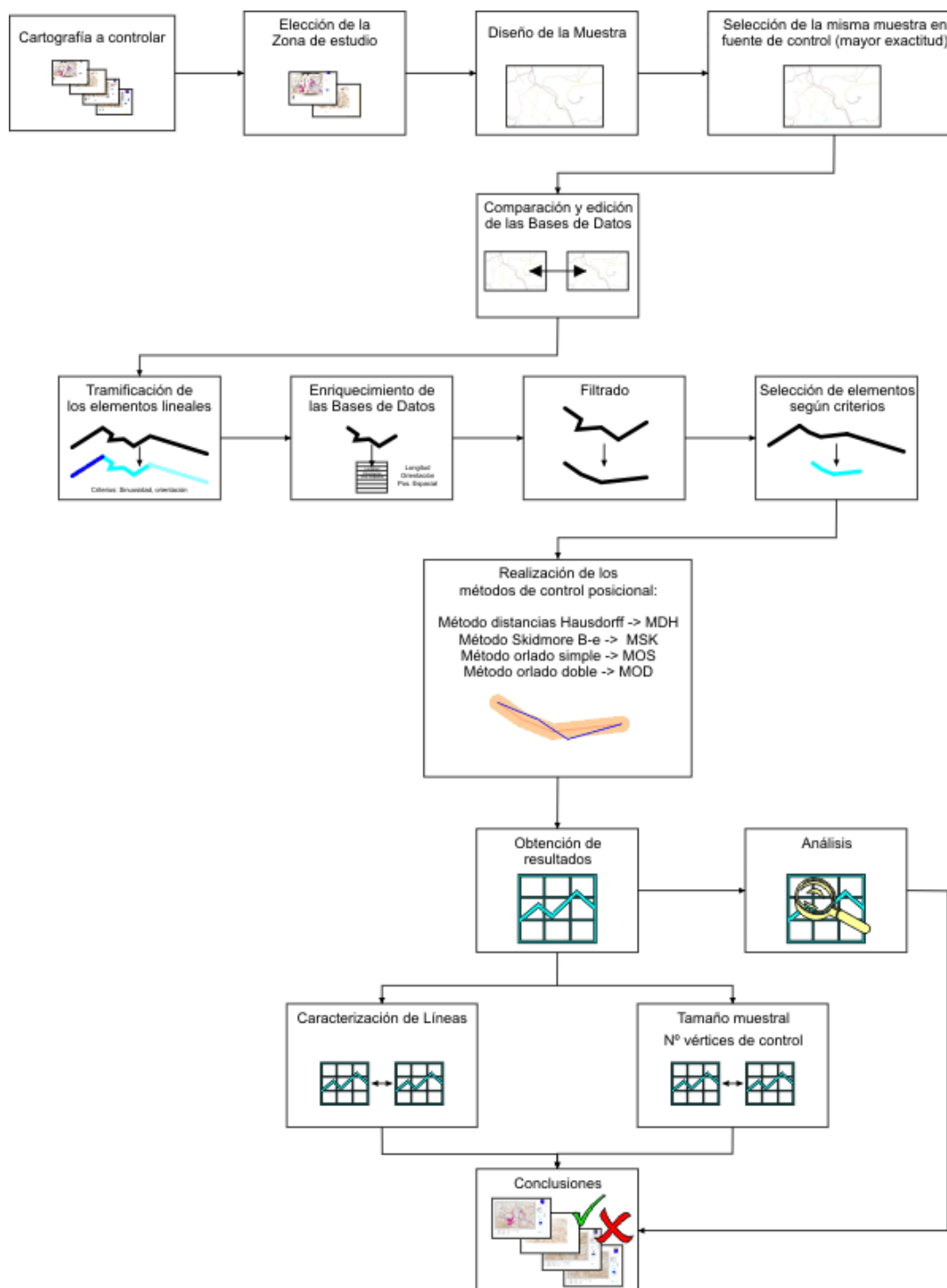


Figura III. 1. Procedimiento General

Como se observa en la Figura III. 1, la metodología propuesta abarca desde la planificación del proyecto de control, la obtención y edición de las bases de datos a utilizar, la realización del control automatizado a través de elementos lineales y el análisis de los resultados. El proceso mostrado se ha desarrollado a través de una herramienta informática CPLin (Control Posicional de Líneas) que permite implementarlo de forma semiautomática, por lo que algunos de los procesos indicados son consecuencia de la optimización de dicha herramienta. CPLin se encuentra ampliamente descrito por Mozas, Ureña y Ariza (2007a, 2007b).

Aunque en la Figura III. 1 el proceso sigue un recorrido secuencial, se ha descrito esta ordenación al resultar la más lógica para su exposición, lo que no impide que algunas fases presentadas en el procedimiento expuesto puedan realizarse de forma previa o simultánea. Por ejemplo, en la fase de análisis, el estudio de resultados parciales puede llevar a la realimentación del proceso.

El control de las bases cartográficas propuesto se realiza sobre vías de comunicación, y más concretamente sobre carreteras. La elección de este tipo de líneas está basada en las características de las mismas: son elementos que se distribuyen de forma más o menos homogénea sobre cualquier territorio, poseen mayor exactitud temporal que otro tipo de líneas como por ejemplo los caminos, y resulta relativamente sencilla su toma en campo (con observaciones cinemáticas GPS).

La metodología propuesta se basa en el control de una cartografía digital existente por elementos lineales, que son a su vez capturados con una observación cinemática GPS de la que se obtienen estos mismos elementos con mayor exactitud.

El proceso comienza con la elección de la cartografía a controlar. De esta fase se pueden obtener datos importantes como la escala de la misma, el número y distribución de los elementos lineales a utilizar, etc.

La elección de la zona de estudio debe tener presente el número y tipo de elementos a controlar. En la mayoría de las ocasiones, el propio proyecto determinará cual es la zona a estudiar, aunque siempre se debe tener en cuenta las características de la zona con objeto de:

- Facilitar la posible captura y postprocesado de las líneas de control con GPS.
- Garantizar la actualidad de la cartografía.
- Tener una adecuada red viaria a controlar (tipo de vías, etc).

Con el diseño de la muestra se pretende seleccionar una serie de elementos cuyo comportamiento pueda ser extrapolado al total de forma que no sea necesario realizar este control al conjunto.

Lógicamente, los elementos seleccionados en la muestra deben existir en la fuente de mayor exactitud con la que se realiza el control, o en caso de materializarlo mediante una observación en campo (GPS), la calidad de dicha medida debe quedar fuera de cualquier duda. Por este motivo, se justifica una fase previa de planificación de la observación en la que la elección de la zona de trabajo pueda cambiar en función de las variables propias de la observación GPS. La metodología de toma de elementos

lineales para el control de calidad cartográfico mediante una observación GPS es ampliamente descrita por Mozas (2003), Ariza y Mozas (2005), Ruiz y col. (2006a, 2006b y 2007).

Con independencia de la forma de adquisición de la muestra seleccionada en una fuente de mayor exactitud, el resultado es una serie de líneas que tienen su correspondencia con las líneas de la muestra cartográfica previamente obtenida. Ambos conjuntos de líneas, al ser tratados digitalmente, se estructuran en dos bases de datos.

La base de este proceso de control consiste fundamentalmente en la comparación de dos bases de datos de elementos lineales. Estas bases de datos suelen estar referidas a periodos de tiempo diferentes, por lo que pueden existir tramos variantes debido a esta diferencia temporal. En la fase de comparación de elementos (Figura III. 1), se editan ambas bases de datos para que las líneas que se corresponden en una y otra no posean este tipo de errores temporales. De igual forma, se dota a las líneas de una cierta coherencia en los inicios y fin de las mismas. Como resultado de esta comparación son obtenidas identificaciones unívocas del elemento lineal en cada base de datos. A partir de este momento, a los elementos relacionados de una y otra base de datos los llamaremos elementos “homólogos”.

De forma complementaria a la fase de comparación y edición se propone la posibilidad de tramificar los elementos, en la que se busca obtener conjuntos con características similares. Con este fin, se utilizan argumentos como pueden ser la sinuosidad, la orientación o la longitud de la línea para realizar esta tramificación. Las bases de datos deben seguir un proceso de enriquecimiento con la finalidad de categorizarlas según estos argumentos. En general, se tiene en cuenta:

- Zona dentro de la serie cartográfica completa.
- Tipo de vía.
- Longitud.
- Orientación de los puntos inicio-fin.
- Sinuosidad.

En la fase de filtrado se pretende eliminar la información redundante de las bases de datos. Se trata, por tanto, de una fase encaminada a reducir el tiempo invertido en ulteriores procedimientos como puede ser el proceso de control posicional. El filtrado a realizar esta basado en el establecimiento de un umbral mínimo de distancia entre dos puntos consecutivos de la misma línea.

La selección de elementos según criterios permite elaborar conjuntos homogéneos de elementos de forma que puedan ser caracterizados según los argumentos comentados previamente.

El paso con mayor importancia dentro del procedimiento descrito consiste en la aplicación de las metodologías de control posicional por elementos lineales a las líneas seleccionadas de las bases de datos a controlar.

Las metodologías de control por líneas propuestas son:

- Método de las Distancias de Hausdorff (MDH).
- Método de Skidmore o Banda Épsilon (MSK).
- Método del Orlado Simple (MOS).
- Método del Orlado Doble (MOD).

Como consecuencia de la aplicación automatizada de estos métodos, se obtienen una serie de resultados que han de ser analizados en su conjunto, ya que cada metodología ofrece una información distinta y complementaria del comportamiento posicional de la cartografía estudiada. Con la obtención de los resultados de distintas selecciones homogéneas de líneas según criterios se puede realizar una caracterización de las líneas a través de las funciones de distribución obtenidas.

Estos resultados también pueden utilizarse para inferir un tamaño de muestra poblacional para cada metodología y para estimar la cadencia mínima de toma de puntos en bases de datos de control obtenidas con observaciones GPS.

El proceso finaliza con la obtención de una serie de conclusiones a partir del análisis realizado previamente.

III.3. Elementos Lineales

Indiscutiblemente el uso de elementos lineales implica la existencia de una serie de características innatas a estos que condiciona la forma en que se realiza el control posicional. Estas características singulares influirán en todos los aspectos, desde la selección de la zona y el muestreo, la forma de trabajar con las bases de datos de líneas y el tipo de control que se realiza sobre estos elementos.

III.3.1. Caracterización

La estimación de la exactitud posicional utilizando elementos puntuales puede entenderse como una medida de la similitud posicional entre el conjunto de puntos de la cartografía y el conjunto de control. Esta similitud se reduce a un valor que informa del nivel de discrepancia posicional entre ambas, sin considerar la existencia de formas o relaciones. Este nivel de discrepancia suele determinarse a través de una distancia entre puntos.

En el caso de elementos lineales, se trabaja con una geometría más compleja que la de los elementos puntuales, ya que además de la posición, se debe considerar la topología y forma de estos elementos. El uso de la distancia no es inmediato ya que no es lo mismo la distancia de un elemento a otro que del segundo al primero. La distancia también varía según se considere sólo los vértices de la línea o cualquier posición intermedia.

La similitud indica la existencia de semejanza o analogía entre los elementos, apareciendo una repetición exacta, o aproximada, de un patrón entre los mismos. Otro concepto a tener en cuenta es la signatura, que se define como una marca o nota puesta en una cosa para distinguirla de otras (Real Academia Española, 1999). De esta forma, la medida de la similitud entre dos elementos se puede realizar comparando

sus firmas mediante el uso de alguna función de distancia adecuada, de forma que cuanto mayor sea la distancia, menos similares puedan considerarse los elementos comparados.

Para el estudio de la similitud se ha introducido dentro de la comparación de elementos lineales, por un lado, la extensión directa del concepto de distancia, y por otro, la comparación de firmas. Este concepto se basa en comparar las funciones de distribución que los resultados del control generan como firma o caracterización.

III.3.2. Zona de estudio

La zona de estudio presenta una gran importancia, ya que incurrir en una mala elección puede condicionar en gran medida la viabilidad del estudio y poner en entredicho los resultados obtenidos. En general, habrá que tener en cuenta las siguientes variables:

- Cartografía:
 - Actualidad.
 - Formato digital.
 - Disponibilidad en la zona de trabajo.
- Captura GPS:
 - Distancia a estaciones de referencia.
 - Topografía circundante.
 - Vegetación circundante. Oclusiones.
 - Hora de toma. Situación de la constelación.
- Red Viaria:
 - Configuración geométrica de la red viaria (radial, diagonal, paralela, etc.).
 - Tipología de las vías de comunicación (nacionales, autonómicas, locales, etc.).
 - Tráfico.
 - Obras.

III.3.3. Diseño de muestras

El diseño de muestras constituye una fase crítica. El trabajo con muestreo permite inferir el comportamiento de toda la población a partir de las muestras seleccionadas, por lo que éstas deben ser representativas de la población.

Lo primero que tendrá que tenerse en cuenta es el establecimiento de la población de interés. Ésta, no consiste en definir los elementos que la forman, sino en tener un listado completo de los mismos, en lo que se denomina el marco de muestreo (Ruiz, Mozas y Ariza, 2006).

Una vez conocida la población, se debe estimar su tamaño. En una serie cartográfica, se pueden adoptar planteamientos como considerar la población compuesta por el conjunto de hojas que forman la serie, por celdas territoriales de tamaño determinados o por las entidades que aparecen en la población.

Otros aspectos importantes son:

- Característica a estimar: Hace referencia al parámetro de la población que se quiere determinar. Puede referirse a un valor medio, proporción o suma de valores, etc.
- Nivel de confianza: Probabilidad del error que corresponde al rechazo de elementos que realmente cumplen la condición impuesta.
- Precisión requerida en los trabajos: Error máximo admisible en el proceso.
- Esquema de muestreo: Cómo se orienta la captura de la muestra. A los elementos lineales se le pueden aplicar todas las alternativas comunes (ver Capítulo II):
 - Muestreo aleatorio simple.
 - Muestreo sistemático.
 - Muestreo estratificado.
 - Muestreo por conglomerados.

III.3.4. Conclusiones

Se ha presentado a lo largo del punto III.3 una pequeña caracterización de los elementos lineales con los que se trabaja. Es importante tener en cuenta las características propias de estos elementos y cuestiones como el estudio de la similitud entre ellos para poder establecer comparaciones entre los mismos.

A continuación, se han presentado cuales son los principales interrogantes conducentes a la elección de la zona de estudio, dependiendo de una serie de variables y de la definición de la muestra a seleccionar teniendo en cuenta el tamaño poblacional, el nivel de confianza, etc.

III.4. Generación de Bases de Datos de elementos lineales

El estudio de elementos lineales está basado en su tratamiento digital, por lo que el conjunto de estos elementos se encuentra almacenado y estructurado en bases de datos. Las mismas están formadas por la secuencia de puntos que forma cada una de las líneas.

III.4.1. Obtención de las bases de datos a controlar

La base de datos de la cartografía a controlar debe ser seleccionada teniendo en cuenta los aspectos previamente comentados. Es importante contar con una base

de datos con elementos lineales de vías de comunicación que presente gran coherencia lógica⁷ lo que ayudará en gran medida en la fase posterior de depuración.

Otro aspecto importante es conocer el sistema de coordenadas en el que se encuentra la base de datos cartográfica ya que en la siguiente fase, se propone realizar una transformación de los datos de control al sistema de referencia de la cartografía.

III.4.2. Obtención de las bases de datos de control

Los elementos de la base de datos de control deben proceder de otra base de datos cartográfica más exacta o, como se propone en este caso, utilizar líneas procedentes de observaciones cinemáticas GPS.

Aunque la adquisición de estos elementos de control viene ampliamente descrita en Ruiz, Mozas y Ariza (2006), se incluyen a continuación algunos aspectos importantes para tener una idea más aproximada del comportamiento de esta base de datos. En la Figura III. 2 se muestra un esquema general del proceso de adquisición y generación de la base de datos de control GPS.

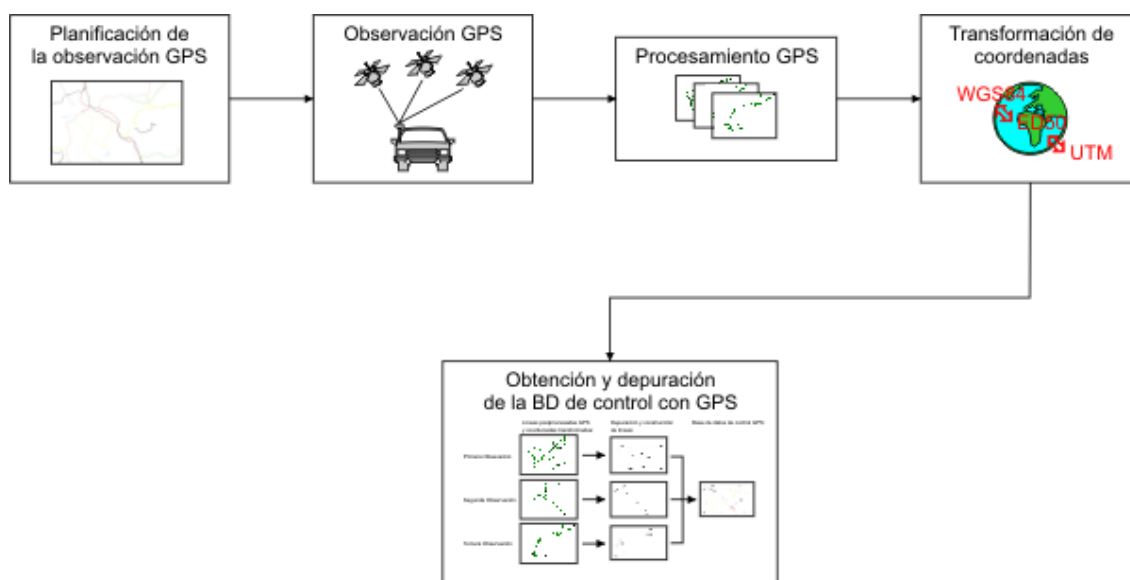


Figura III. 2. Esquema del proceso de obtención de la BD de control

La adquisición de la base de datos de control GPS se realiza a partir de una observación cinemática del arcén en cada sentido de la vía. La toma de puntos se realiza con un intervalo de tiempo que dependerá de la velocidad del vehículo y las necesidades del proyecto (escala de la cartografía). Al final del presente capítulo se propone una metodología para la determinación de la cadencia mínima de vértices dentro de los elementos de control. Para su procesamiento se emplea código diferencial a partir de estaciones de referencia cercanas a la zona de observación. El proceso

⁷ Coherencia lógica: Trata de las reglas lógicas que ha de cumplir la estructura de los datos, sus atributos, sus formatos físicos.

comienza con la planificación de la observación, en la que se determina las carreteras a tomar, se estudia la configuración de la constelación GPS, la velocidad del vehículo, etc. (Ruiz, Mozas y Ariza, 2006). A continuación, se realiza la observación GPS y se procesan los datos obtenidos. Con la secuencia de puntos obtenida se realizan una serie de transformaciones de coordenadas hasta llegar a tener el mismo marco de referencia que la cartografía a controlar.

Por último, se generan las líneas a partir de la secuencia de puntos ordenados teniendo en cuenta para identificar dos puntos contiguos el utilizar aquellos cuya hora de toma durante la observación muestre una diferencia igual al intervalo de toma de puntos durante dicha observación.

Con esto, se obtiene un conjunto de elementos formados por 2 líneas aproximadamente paralelas que definen cada tramo de carretera ya que cada vía es observada en los dos sentidos de la marcha. Para determinar el eje de la vía se realiza una interpolación a partir de las líneas observadas (Figura III. 3). Este procedimiento es descrito por Ruiz, Mozas y Ariza (2006).

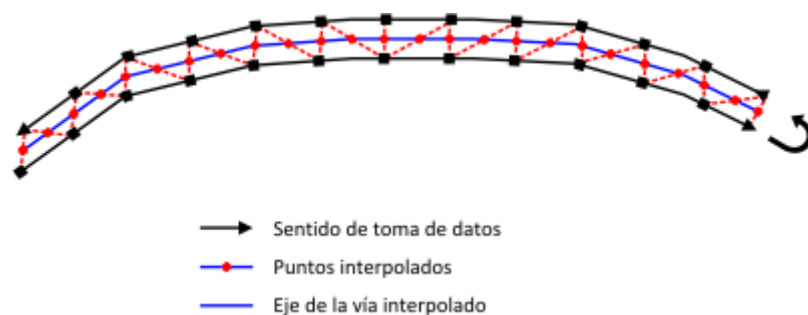


Figura III. 3. Determinación de ejes de líneas de control

La base de datos de control estará formada por los ejes previamente determinados.

III.4.3. Depuración, comparación y edición de las bases de datos

Hasta el momento, se cuenta con dos bases de datos independientes que representan la muestra de carreteras realizada en fases anteriores. Evidentemente, estas dos bases de datos presentarán numerosas discrepancias tanto en el número de elementos, en la longitud de los tramos, en la división de los mismos, etc.

En la fase que nos ocupa, se realiza una depuración de estas diferencias, realizando una comparación de las dos bases de datos. Las variaciones encontradas deben ser editadas y modificadas.

Como ejemplo, se puede indicar que se eliminarán tramos que aparecen sólo en una de las bases de datos (Figura III. 4.a), las líneas que forman los enlaces viarios o túneles (Figura III. 4.b), los tramos en los que se detecta visualmente algún error propio de la observación GPS (multicamino, pérdidas de señal, etc.) (Figura III. 4.c) o la aparición de tramos reformados en la observación GPS pero que en la cartografía presentan un trazado antiguo (Figura III. 4.d). Estos casos suelen aparecer en zonas

curvas de las vías y se identifican con claridad al aparecer suavizado el trazado de la línea.

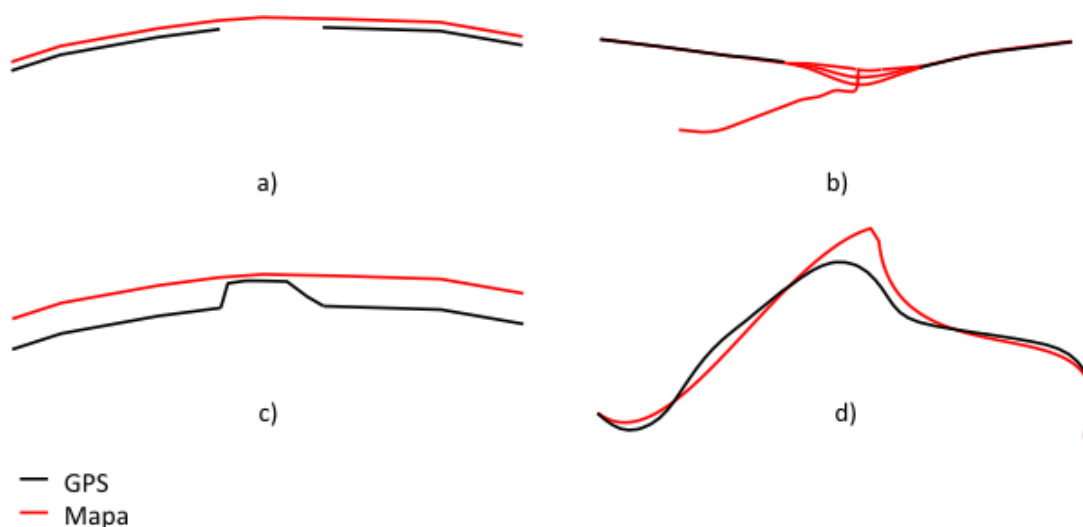


Figura III. 4. Eliminación de tramos

Por otro lado, los tramos en una y otra base de datos deben ser similares por lo que el proceso de edición incluye la reconstrucción de la coherencia lógica de los elementos.

III.4.4. Tramificación

Las bases de datos generadas se encuentran tramificadas como consecuencia del proceso de edición realizado previamente. Sin embargo, puede resultar interesante para ciertos estudios proceder a la división de las líneas en tramos siguiendo una serie de criterios como la orientación, la longitud o la sinuosidad de la línea. Con esta tramificación se pretende obtener elementos homogéneos dentro de las bases de datos para categorizar los resultados obtenidos.

III.4.5. Enriquecimiento

El enriquecimiento de las bases de datos consiste en agregar atributos a cada línea que describan los aspectos más característicos de la misma. Esta técnica es utilizada y descrita por numerosos autores (Plazanet (1996), Ariza y García (2004)).

El disponer de estas bases de datos enriquecidas, permite generar extracciones aleatorias y simular procesos de control por elementos lineales que cumplan ciertas características. Este proceso se realiza de forma automática y favorece la toma de decisiones, reduciendo los procesos de computación.

En esta fase del proceso se incluyen datos como el tipo de vía, la longitud de la línea, la orientación desde el punto inicial al final de la línea o la sinuosidad de la misma, etc.

III.4.6. Identificación de elementos homólogos

Las bases de datos de elementos de la cartografía y de control se encuentran separadas. La identificación de elementos homólogos permite relacionar un elemento de una base de datos con un elemento de la otra. Esta identificación será correcta si estos elementos relacionados se corresponden con el mismo tramo de la realidad.

La identificación de elementos homólogos consiste, por tanto, en encontrar un elemento lineal en las dos bases de datos (cartografía y control) y reordenar los puntos de una de ellas para que su secuencia tenga el mismo orden en las dos bases de datos. Esta reordenación tiene una finalidad computacional, ya que al estar las líneas en el mismo orden resulta más sencillo recorrerlas en procesos automatizados más complejos.

La estrategia de identificación de elementos homólogos se basa en comparar la distancia entre los puntos inicial y final de cada elemento lineal de la base de datos de control con todos los puntos de un elemento de la base de datos cartográfica. Para un elemento lineal de la base de datos de control se obtienen los puntos de la base de datos cartográfica que se encuentran más próximos al punto inicial y al punto final de control y que pertenecen a un mismo elemento de la base cartográfica.

Si la distancia entre los puntos homólogos es inferior a un umbral preestablecido, la identificación es positiva y las dos líneas son consideradas como homólogas. En el caso de que alguna distancia sea superior al umbral, la línea se desecha.

La ocurrencia más sencilla que puede presentarse es cuando los puntos más próximos se corresponden con los puntos inicial y final del elemento de la base de datos cartográfica (Figura III. 5).

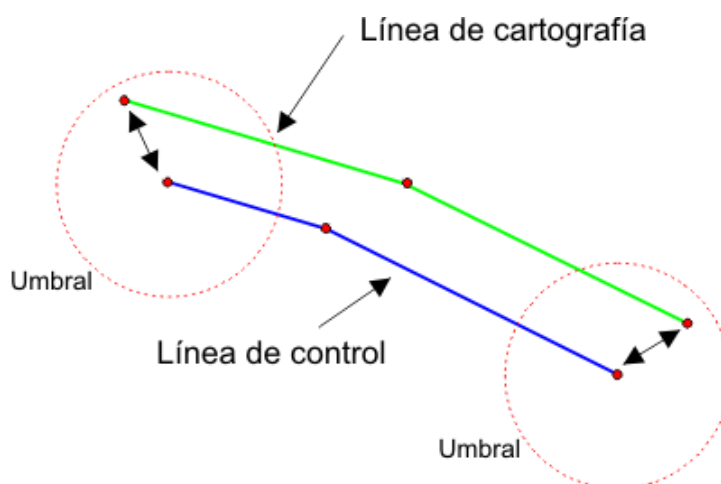


Figura III. 5. Identificación de líneas homólogas. Caso 1

Lo más probable, si no se ha realizado un proceso exhaustivo de revisión y edición de líneas anteriormente, es que el punto más próximo de la base de datos cartográfica sea un punto no extremo de la misma línea. En este caso, se selecciona

este punto como punto extremo (Figura III. 6), eliminando los puntos anteriores a este (si se trata del punto inicial) o los posteriores (si se trata del punto final).

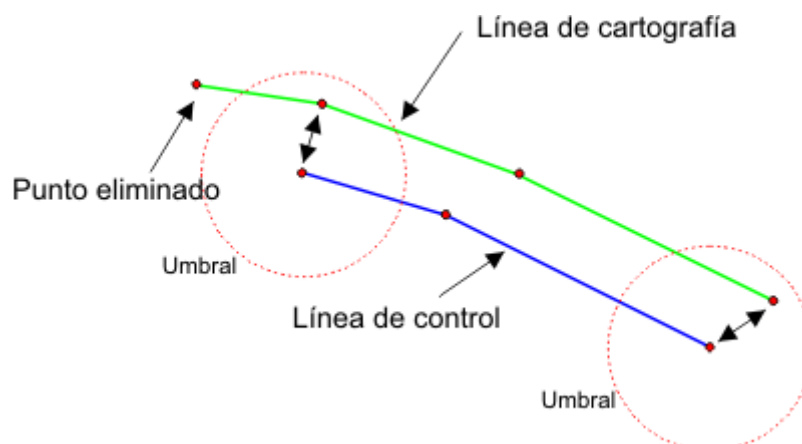


Figura III. 6. Identificación de líneas homólogas. Caso 2

Cuando se han relacionado todos los elementos lineales de la base de datos de control con la cartográfica, se debe reordenar una de las bases de datos. Lógicamente, los elementos de control, van a ser cuantitativamente más complejos en su estructura (mayor número de puntos), por lo que se recomienda realizar este reordenado sobre la base de datos cartográfica. Este proceso, se realiza cuando haya elementos lineales en la cartografía cuyos puntos no estén dispuestos en el mismo sentido que en la base de datos de control. En estos casos, el primer punto de la línea de la base de datos de control se corresponde con el último de la base cartográfica y viceversa. Para realizar el reordenamiento, los puntos de ese elemento lineal de la base cartográfica invierten su orden.

III.4.7. Eliminación de información redundante

Después de obtener las bases de datos de control y a controlar es recomendable aplicar dos procesos. El primero de ellos consiste en la comprobación de las entidades suministradas, eliminando las que no vayan a ser empleadas durante el proceso de control posicional. El segundo radica en asegurar la consistencia de los datos y reducir el coste computacional del resto de procesos eliminando aquellos vértices que no aporten información sobre la curva. Esta información redundante aparece, por ejemplo, en bases de datos de elementos procedentes de observaciones GPS, donde, durante la fase de observación, pueden tomarse puntos poco espaciados. Esta ocurrencia se puede dar, por ejemplo, al iniciar y detener la marcha el vehículo con el que se recorre los elementos lineales durante la observación GPS.

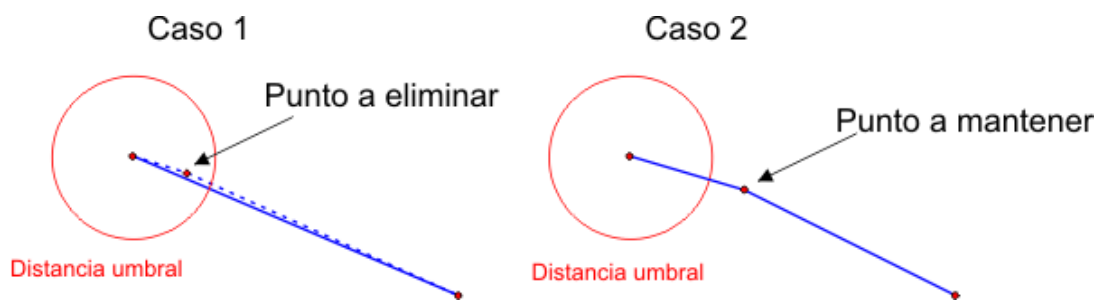


Figura III. 7. Filtrado por distancia

Existen muchos métodos de filtrado posibles (uno de cada, por distancia, por distancia y angularidad, etc.), de hecho, la mayor parte de los métodos de generalización de elementos lineales, siempre que dentro de su proceso se elimine alguno de los vértices tratados, son del tipo de filtrado.

De entre todos ellos, y debido al método general de captura empleado, en el proceso desarrollado se propone la realización de un filtrado por distancia. Este tipo consiste en obtener la distancia entre un punto y el anterior de una línea y comprobar si este valor es superior o inferior a un umbral que se establece. En caso de ser superior este punto se mantiene y se repite el proceso con el siguiente punto de la línea. En caso contrario, se elimina ese punto y se continúa con el siguiente punto (Figura III. 7).

Se realiza un filtrado independiente de las dos bases de datos, de forma, que se puede establecer una distancia umbral diferente para cada una de ellas. La elección de esta distancia dependerá de la escala de la cartografía, teniendo en cuenta el error que se puede cometer, sin que éste tenga repercusión en la calidad de la misma y del error en la toma de datos de los elementos de control.

III.4.8. Selección de elementos

La selección de elementos lineales que interesen para realizar el control de calidad se presenta con anterioridad al mismo. Para la selección se utilizan diversos criterios que clasifiquen los elementos de la base de datos:

- Tipo: Autovías, nacionales, autonómicas y locales.
- Longitud del elemento.
- Orientación general del elemento.
- Selección por zona.
- Selección aleatoria.

La importancia de realizar este proceso de selección radica en poder obtener resultados del control posicional sobre conjuntos de elementos con cualidades homogéneas, buscando una caracterización de los mismos.

III.4.9. Conclusiones

Se ha presentado a lo largo del punto III.4 la metodología de generación de bases de datos de elementos lineales. En general, se tratan de procesos sencillos que se realizan de forma semiautomática, y que permiten aplicar posteriormente los métodos de control posicional de una forma automática a cualquier conjunto de datos. Se puede considerar, por tanto, una fase preparatoria de gran importancia para las fases sucesivas.

La obtención de cada base de datos es un aspecto importante en cualquier proyecto automatizado de control posicional, ya que fases como la depuración y edición de los elementos afecta en buena medida a los métodos de control posicional.

Aspectos como la identificación automática de elementos lineales homólogos también dependen fundamentalmente de lo satisfactorio que resulte el proceso de edición y comparación de las bases de datos.

Por otro lado, se ha propuesto dentro de esta parte del procedimiento un filtrado de elementos conducente a eliminar la información redundante existente en las bases de datos.

Por ultimo, se han presentado herramientas que permitan la selección de elementos dependiendo de características propias de cada uno que se incorporaron en las bases de datos dentro del proceso de enriquecimiento de las mismas.

III.5. Control automatizado por elementos lineales

Las metodologías de control posicional por elementos lineales han ido evolucionando a lo largo de los últimos años. Desde el punto de vista posicional la banda de indeterminación o banda épsilon (Blakemore, 1983) es una aproximación que permite un planteamiento complementario a los tradicionales estudios basados en puntos. Existen diversas aplicaciones de control por elementos lineales bastante elementales y como fue explicado en el capítulo anterior, hay propuestas de métodos de control basados en los mismos (Giordano y Veregin, 1994; Ariza, 2002; Atkinson y Ariza 2002) o en métricas basadas por ejemplo en la distancia de Hausdorff (Abbas, Grussenmeyer y Hottier, 1995).

Como se indica en el capítulo anterior, Skidmore y Turner (1992) presentan un método basado en la idea de banda épsilon como forma para determinar la incertidumbre posicional. La estimación se realiza comparando las líneas homólogas y cuantificando la distancia media entre las mismas. Goodchild y Hunter (1997) proponen un método basado en la realización de un orlado de cierta anchura alrededor de los elementos de control y determinando el porcentaje de longitud incluida dentro de esa orla. Tveite y Langaas (1999) estudian una metodología más compleja basada en la realización de orlados en las dos líneas homólogas (la línea a controlar y la línea de control), trabajando con todas las posibles áreas que los cruces entre estos orlados generan.

Sin embargo, estas investigaciones se limitan por lo general a mostrar ideas básicas que se utilizan aplicándose a pequeños ejemplos, usando cartografía a

pequeña escala, tanto para controlar, como para el control o utilizando una única metodología de control.

Trabajo	Método	Kms de líneas	Cartografía	Control
Goodchild y Hunter (1997)	MOS	247	1:100.000	1:25.000
Kagawa, Sekimoto y Shibaki (1999)	MOS, MDH	0.706	1:2.500	1:500
Van Niel y McVicar (2002)	MOS	466	1:50.000	GPS
Mozas (2003)	MOS	136	1:25.000	GPS
Tveite y Langaas (1999)	MOD	-	1:1.000.000	1:250.000
Skopeliti y Tsoulos (2000a, 2000b)	MDH	-	1:100.000	-
Goeman, Martínez-Fonte, Bellens y Gautama (2005a y 2005b)	MOD	-	Quickbird	Quickbird
Johnston, Timlin, Szafoni y Casanova (1999)	MOS	270	-	-

Tabla III. 1. Comparativa estudios con elementos lineales

En la Tabla III. 1 se muestra un resumen de algunas aplicaciones realizadas con elementos lineales que apoyan las ideas previamente comentadas. Está justificado por tanto, el desarrollo de un sistema que realice un estudio conjunto y completo de todos los métodos, se aplique a cartografía a media escala y con un gran número de datos. Es necesario también, realizar un contraste riguroso de los resultados de estos métodos.

Los métodos de control propuestos son:

- Método de las Distancias de Hausdorff (MDH).
- Método de la Banda Épsilon o Skidmore (MSK).
- Método del Orlado Simple (MOS).
- Método del Orlado Doble (MOD).

Y se corresponden con los investigados por Abbas, Grussenmeyer y Hottier (1995), Skidmore y Turner (1992), Goodchild y Hunter (1997) y Tveite y Langaas (1999) respectivamente.

Estos métodos, están resumidos en el capítulo II, por lo que en el punto que nos ocupa se procede a ampliar su descripción presentando, además, la metodología de automatización de los mismos a través de la herramienta informática CPLin.

III.5.1. Método de las Distancias de Hausdorff (MDH)

El método MDH está basado en el cálculo de las distancias entre pares de líneas. Más concretamente, se calcula el máximo de las distancias menores de la línea de la cartografía a la línea de control y viceversa. De estos valores, se selecciona el máximo. El cálculo de la distancia en ambos sentidos (mapa a control y viceversa) es debido a la asimetría de este tipo de medida. Esta métrica se puede considerar como una extensión directa de la euclídea al llevarla a elementos lineales. Se constituye en una opción con una interpretación física inmediata ya que supone un desplazamiento máximo de carácter puntual.

La automatización e implementación de esta metodología resulta bastante sencilla (Figura III. 8). Basta con recorrer la línea de la cartografía (X) calculando para cada vértice la mínima distancia a la línea de control (Q). A continuación se realiza el proceso contrario, recorriendo la línea de control y calculando las distancias mínimas de sus vértices a la línea a controlar. Con las distancias mínimas obtenidas se propone obtener el valor máximo (máximo entre $d1$ y $d2$), lo que supone la aplicación de forma estricta del concepto de distancia de Hausdorff.

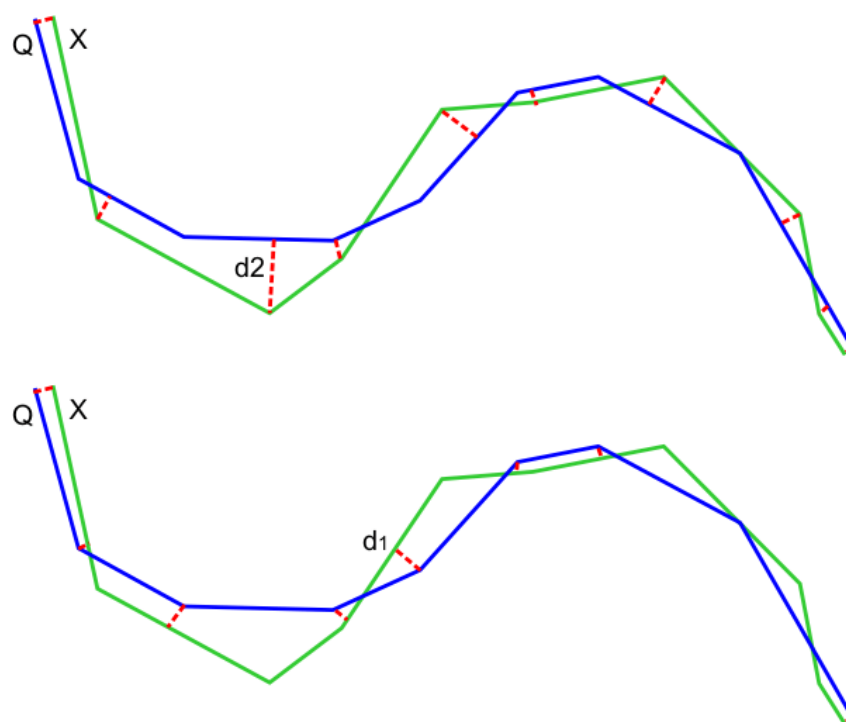


Figura III. 8. Automatización Método MDH

Además, se propone obtener la media de las distancias menores de la cartografía a la línea de control y viceversa, ya que supone otro valor, indicativo del comportamiento de las líneas, que complementa al valor máximo que marca la distancia de Hausdorff. El conjunto de valores máximo y medio de distancia para cada elemento aporta más información que el uso exclusivo del valor de Hausdorff.

Los resultados obtenidos aplicando esta metodología a todas las líneas se expresan como sigue:

- Se promedian los valores obtenidos D_{h_j} , ponderando la longitud de cada una de estas líneas L_j con lo que se obtiene un valor promedio $\overline{D_h}$ y una medida de dispersión (desviación s) del conjunto de líneas.

$$\overline{D_h} = \frac{\sum_{j=1}^n D_{h_j} L_j}{L_T}$$

Ec. III. 1

$$s = \sqrt{\frac{\sum_{j=1}^n (D_{h_j} - \overline{D_h})^2 L_j}{L_T}}$$

Ec. III. 2

siendo L_T la longitud total de líneas.

- De forma análoga se promedian los valores para las medias de las distancias menores.
- Se plantea una representación gráfica con los valores de la funciones de distribución de los resultados obtenidos.

III.5.2. Método de la Banda Épsilon o Skidmore (MSK)

El método MSK considera los elementos lineales como tales. En la versión del método con los puntos no coincidentes (ver Capítulo II), que tiene la ocurrencia más lógica en casos reales, la medida se deriva del cálculo de la superficie encerrada por las líneas de la cartografía y de control homólogas. La interpretación física resulta bastante sencilla, ya que supone un desplazamiento promedio de las líneas.

La automatización e implementación de este método es algo más complejo que el método MDH (Figura III. 9), ya que implica la construcción de los polígonos intermedios generados por el posible cruce de las líneas homólogas. En caso de no existir cruces bastará con unir los puntos de inicio y de fin de las líneas homólogas entre sí para obtener el polígono intermedio.

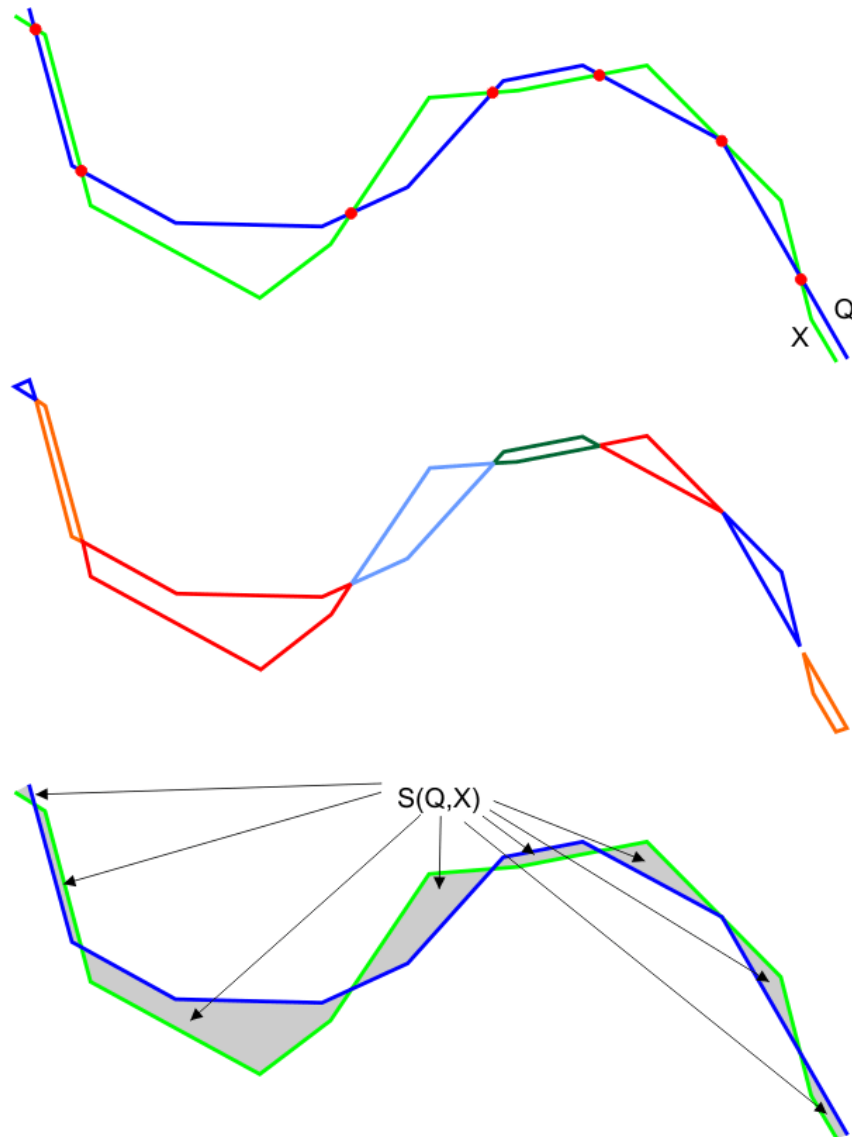


Figura III. 9. Automatización Método MSK

Una vez generado el/los polígono/s intermedio/s se procede al cálculo de su/s superficie/s S_i y se divide entre la longitud de la línea L obteniendo un valor único que indica el desplazamiento promedio obtenido (Ec. III. 3).

$$Dp = \frac{\sum_{i=1}^{i=n} S_i}{L}$$

Ec. III. 3

Para el conjunto de líneas se plantean los siguientes resultados:

- Se promedian los valores obtenidos de forma análoga al caso anterior (Ec. III. 1 y Ec. III. 2), ponderando la longitud de cada una de estas líneas

con lo que se obtiene un valor promedio y una medida de dispersión (desviación) del conjunto de líneas.

- Se plantea una curva con los valores de la funciones de distribución de los resultados obtenidos.
- Se estudia el número de cruces producidos entre las dos líneas homólogas.

III.5.3. Método del Orlado Simple (MOS)

El método MOS se basa en cálculos sobre elementos lineales. En este método se realiza un orlado sobre el elemento de control y se calcula el porcentaje de longitud del elemento homólogo de la cartografía que se encuentra en el interior del orlado para distintas anchuras del mismo. La idea del orlado puede expresarse como la extensión del concepto de incertidumbre al caso lineal.

Para un par de elementos homólogos se obtienen dos parámetros: La anchura del orlado y la longitud o el porcentaje de la longitud que se encuentra dentro del mismo.

La implementación y automatización de este método es mucho más compleja que las anteriores (Figura III. 10). Supone, para cada anchura de orlado, realizar a partir de la línea de control una línea a la mitad de la anchura de orlado a cada lado de la línea original (generación del orlado descrita por Mozas, Ureña y Ariza (2007b)). Una vez construido el orlado, se realiza una intersección con la línea homóloga de la cartografía contabilizando la longitud de la misma L_B que se encuentra dentro del orlado Goodchild y Hunter (1997).

$$OS_j = \frac{L_B}{L_j} 100$$

Ec. III. 4

Siendo L_j la longitud total de la línea.

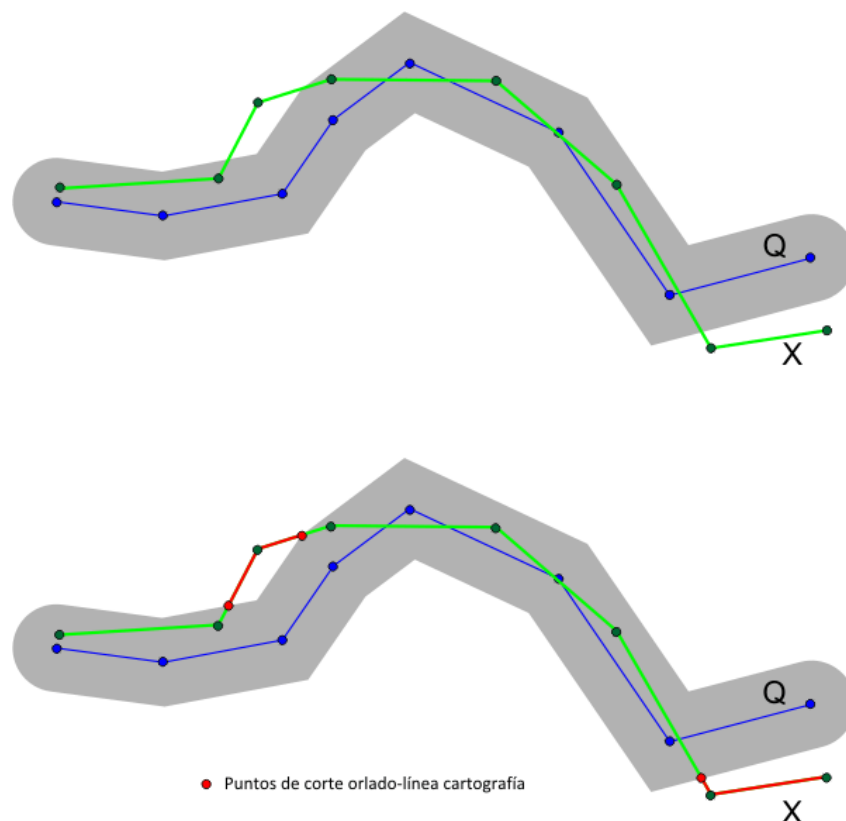


Figura III. 10. Automatización Método MOS

El resultado (Ec. III. 4) es un valor de porcentaje de longitud para cada anchura de orlado por lo que para cada elemento se puede generar una función de distribución de estos valores.

De forma simultánea al cálculo del porcentaje de longitud se propone otra medida como es el porcentaje de vértices de la línea de la cartografía incluidos dentro del orlado.

Para el conjunto de líneas de un proyecto se plantean los siguientes resultados:

- Se promedian los valores obtenidos, ponderando el porcentaje de longitud OS_j de cada una de estas líneas j dentro de cada orlado de anchura D con lo que se obtiene una función de distribución promedio para el conjunto de los datos.

$$OS_D = \frac{\sum_{j=1}^n OS_j L_j}{L_T}$$

Ec. III. 5

- A partir de esta función se genera una curva gráfica representativa de su comportamiento.
- Se plantea otra curva similar con la medida del porcentaje de vértices incluidos en los orlados.

III.5.4. Método del Orlado Doble (MOD)

El método MOD es el más complejo de los aplicados ya que supone la realización de orlados sobre ambas líneas y contabilizar las superficies obtenidas según los cruces producidos. En este caso, las incertidumbres se aplican tanto a los elementos de control como a los cartográficos.

El proceso automatizado (Figura III. 11) consiste en generar el orlado para una anchura concreta sobre el elemento de control de la misma forma que se realiza en el método MOS. Con posterioridad, se realiza el orlado con la misma anchura sobre el elemento a controlar. Cada uno de los orlados generados puede ser considerado un polígono, por lo que se puede calcular su superficie. Se calculan las posibles intersecciones de los orlados obteniendo los polígonos comunes a ambos. A estos polígonos comunes se le calcula la superficie, con lo que tiene una medida sobre la superficie común para esa distancia de orlado. Bastará con realizar el cálculo de la diferencia entre la superficie del orlado del elemento de control con la superficie común para obtener otra medida propuesta en la metodología original y análogamente con la superficie del orlado del elemento a controlar.

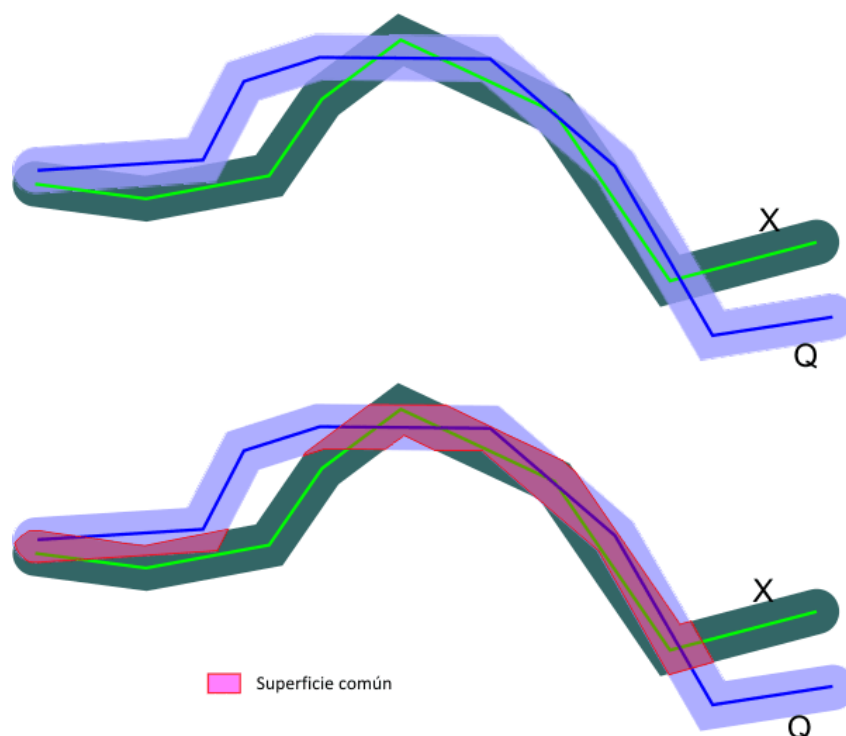


Figura III. 11. Automatización Método MOD

En definitiva se cuenta con las siguientes medidas:

- Superficie total común de los orlados para cada anchura.
- Superficie total incluida en los orlados de control y no perteneciente a las zonas comunes de los mismos para cada anchura.

- Superficie total incluida en los orlados de la cartografía y no perteneciente a las zonas comunes de los mismos para cada anchura.

Además, el proceso ofrece como resultado un valor cuantitativo del número de áreas interiores vacías generadas por ambos orlados.

Las medidas propuestas permiten calcular los valores descritos en el método original como son el desplazamiento promedio o la oscilación (ver Capítulo II).

Para el conjunto de líneas se plantean los siguientes resultados:

- Superficie Común:
 - Se calcula el porcentaje de superficie común SC_j para cada elemento y para cada anchura de orlado.

$$SC_j = \frac{2 \sum_{i=1}^n SC_i}{SO_{j \text{ Control}} + SO_{j \text{ Cartografía}}}$$

Ec. III. 6

- Se promedian los valores obtenidos de superficie común, ponderando el resultado con la longitud L_j de cada una de estas líneas dentro de cada orlado D con lo que se obtiene una función de distribución promedio para el conjunto de los datos.

$$SC_D = \frac{\sum_{j=1}^n SC_j L_j}{L_T}$$

Ec. III. 7

- A partir de esta función, se genera una curva gráfica representativa de su comportamiento.
- Desplazamiento promedio:
 - Se calcula el desplazamiento promedio (Capítulo II) para cada elemento y para cada anchura de orlado D .

$$DP_j = \pi D \frac{SO_{j \text{ Control}} - SC_j}{SO_{j \text{ Cartografía}}}$$

Ec. III. 8

- Se promedian los valores obtenidos de desplazamiento promedio, ponderando el resultado con la longitud de cada una de estas líneas dentro de cada orlado con lo que se obtiene una función de distribución promedio para el conjunto de los datos (análogamente a Ec. III. 7).
- Se obtiene una representación gráfica de la función de distribución obtenida.
- Oscilación:
 - Se calcula el valor de la oscilación (Capítulo II) para cada elemento y para cada anchura de orlado.

$$O_j = \frac{N_{Areas}}{L_j}$$

Ec. III. 9

Donde N_{Areas} es el número de áreas interiores vacías generadas por ambos orlados.

- Se promedian los valores obtenidos de oscilación de cada una de estas líneas dentro de cada orlado con lo que se obtiene una función de distribución promedio para el conjunto de los datos (análogamente a Ec. III. 7).
- A partir de esta función, se genera una curva gráfica representativa de su comportamiento.

III.5.5. Conclusiones

Los métodos previamente comentados no se han aplicados de forma exhaustiva y con abundantes datos reales para la evaluación de la exactitud posicional. Se tratan de propuestas aplicadas de forma experimental (Tabla III. 1).

La propuesta metodológica de esta tesis doctoral incluye la utilización de estas cuatro metodologías para el control posicional. La implementación de las mismas de forma automatizada se realiza a través de una herramienta informática (CPLin), de la que se derivan las medidas propuestas por los autores de las distintas técnicas además de otros valores propuestos en la presente investigación.

El uso que se realiza de estas cuatro metodologías debe ser entendido conjuntamente a la hora del análisis de los resultados, aunque en la fase de cálculo cada una se realice de forma independiente. Con esto, se pretende no establecer un valor único de la exactitud posicional, sino realizar un enfoque más amplio que abarque las medidas resultantes de estas metodologías de forma que, cada valor aporte información sobre el comportamiento de los elementos lineales utilizados.

Como novedad a las metodologías originales se le han incorporado el cálculo de los siguientes valores:

- Método MDH:
 - Cálculo de distancias medias en el método MDH.
 - Cálculo de promedio global del proyecto y medida de la dispersión.
- Método MSK:
 - Cálculo de promedio global del proyecto y medida de la dispersión.
 - Estudio de los cruces entre las líneas homólogas.
- Método MOS:
 - Cálculo y estudio de la función de distribución global del proyecto.
 - Cálculo del porcentaje de vértices incluido en el orlado.
- Método MOD:
 - Cálculo y estudio de la función de distribución global del proyecto para la superficie común, el desplazamiento promedio y la oscilación.

III.6. Caracterización y similitud entre funciones de distribución

Como se ha explicado en el apartado anterior, las técnicas MDH y MSK generan como resultado valores escalares de media y desviación con los que las medidas de exactitud derivadas de los elementos lineales utilizando estos métodos quedan perfectamente caracterizadas. De esta forma, con sólo dos parámetros, asumiendo la normalidad, se pueden establecer relaciones con diversos niveles de confianza.

El trabajo con los resultados procedentes de las técnicas MOS y MOD es más complejo, ya que sobre ellos no se pueden establecer distribuciones parametrizadas ni parametrizables. Por ello, se propone el uso de las funciones de distribución empíricas como signature de forma que permitan realizar medidas de similitud entre ellas.

Para estos resultados, la determinación de si distintas circunstancias como la longitud, zona, orientación, sinuosidad, etc., presentan un mismo comportamiento, ha de realizarse sobre las funciones de distribución obtenidas.

Para realizar la comparación entre dos funciones de distribución para obtener una cuantificación de su similitud se propone utilizar el test de Kolmogorov-Smirnov (Gibbons y Chakraborti, 1992).

Este test permite determinar, si dos funciones de distribución de dos conjuntos de datos son o no similares con cierto nivel de significación.

En la Figura III. 12 (Press y Teukolsky, 1992), se muestra un ejemplo del criterio base del Test de Kolmogorov-Smirnov para el caso de la distancia entre una función de distribución teórica (curva suave) y una función de distribución observada (curva en escalera).

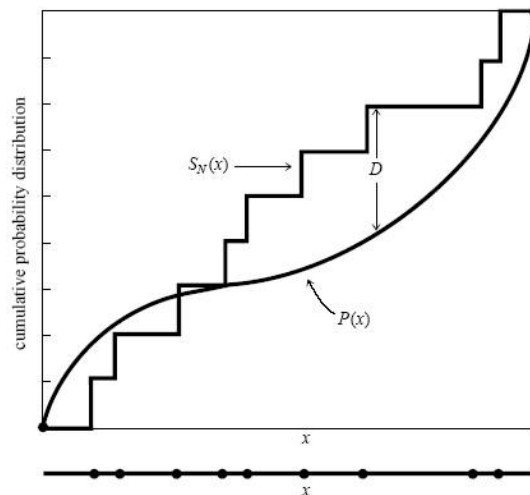


Figura III. 12. Ejemplo del criterio base del TKS. Press y Teukolsky (1992)

Según Gibbons y Chakraborti (1992) el test de Kolmogorov-Smirnov se basa en comparar las funciones de distribución empíricas asociadas a sendas muestras aleatorias de las respectivas variables. Para realizar este contraste se comprueba la siguiente hipótesis nula:

$$H_0 : F_X(x) = F_Y(x)$$

$$H_1 : F_X(x) \neq F_Y(x)$$

Ec. III. 10

La hipótesis nula supone los dos conjuntos de datos provenientes de la misma distribución y la hipótesis alternativa lo contrario.

Según la Figura III. 12 con el Test de Kolmogorov-Smirnov se pretende determinar la distancia máxima $D_{KS \text{ Observada}}$ existente entre las dos funciones de distribución que se analizan. La similitud entre dos curvas que representan sendas funciones de distribución queda definida por el valor que alcanza una distancia establecida como tolerancia ($D_{KS \text{ Tolerancia}}$) que se determina, entre otros, en función de: el nivel de significación que se adopta y el número de casos que definen las funciones de distribución. Bajo las condiciones anteriores, si $D_{KS \text{ Observada}} < D_{KS \text{ Tolerancia}}$, se acepta la hipótesis nula y se puede afirmar que las dos funciones de distribución son semejantes.

Otra manera de trabajar, que es la adoptada en nuestro caso, consiste en obtener los valores de probabilidad $p\text{-valor}$ que le corresponden a una $D_{KS \text{ Observada}}$. En este caso, cuanto menor sea $D_{KS \text{ Observada}}$ tanto mayor será el $p\text{-valor}$, lo cual puede ser interpretado como mayor confianza en la similitud de las funciones de distribución analizadas. Este $p\text{-valor}$ puede determinarse analíticamente utilizando diversas propuestas como las presentadas en Press y Teukolsky (1992) y Gibbons y Chakraborti (1992). La herramienta informática CPLin calcula los valores de $D_{KS \text{ Observada}}$ y $p\text{-valor}$ para un estadístico D .

Gibbons y Chakraborti (1992) realizan el siguiente planteamiento:

Dadas las muestras aleatorias $\{X_i\}_{1 \leq i \leq N_1}$, $\{Y_j\}_{1 \leq j \leq N_2}$, de tamaños N_1 y N_2 y considerando $S_{N_1}(x)$ y $S_{N_2}(x)$ las funciones de distribución empíricas asociadas a dichas muestras, el estadístico de contraste será:

$$D = \max_x |S_{N_1}(x) - S_{N_2}(x)|$$

Ec. III. 11

Donde x recorre todos los valores de la muestra de tamaño $N_1 + N_2$ que se tendría si la hipótesis nula es cierta.

La distribución nula asintótica del estadístico D es:

$$\lim_{N_1, N_2 \rightarrow \infty} P \left[\sqrt{\frac{N_1 N_2}{N_1 + N_2}} D \geq d \right] = Q_{KS}(d) = 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2 d^2}$$

Ec. III. 12

Para un valor observado D_{obs} , el p-valor asociado se obtiene como

$$p = P[D \geq D_{obs}] = Q_{KS} \left(\sqrt{\frac{N_1 N_2}{N_1 + N_2}} D_{obs} \right)$$

Ec. III. 13

Sin embargo, para implementar este test en una aplicación informática habrá que realizar una serie de aproximaciones.

Un valor aproximado del p-valor según Press y Teukolsky (1992) vendrá dado por:

$$p = P[D \geq D_{obs}] \approx Q_{KS} \left(\left[\sqrt{N_e} + 0.12 + 0.11 / \sqrt{N_e} \right] D_{obs} \right)$$

Ec. III. 14

Donde N_e se calcula de la forma:

$$N_e = \frac{N_1 N_2}{N_1 + N_2}$$

Ec. III. 15

Para calcular $Q_{KS}(d)$ se realiza una aproximación con 100 iteraciones de la Ec. III. 12.

$$Q_{KS}(d) \approx 2 \sum_{i=1}^{100} (-1)^{i-1} e^{-2i^2 d^2}$$

Ec. III. 16

En cuanto al estadístico Z se corresponde con el valor de distancia tipificada pudiendo calcularse de esta forma:

$$Z_{KS} = \left(\sqrt{\frac{N_1 + N_2}{N_1 N_2}} D_{obs} \right)$$

Ec. III. 17

En definitiva, para comprobar si dos funciones de distribución son similares se aplica el Test de Kolmogorov-Smirnov. Esta prueba permite caracterizar los resultados procedentes de la metodología de control posicional, ya que, como se ha expuesto anteriormente, estos resultados son obtenidos en forma de funciones de distribución no parametrizables.

III.7. Determinación del tamaño de la muestra

La determinación del tamaño de la muestra puede ser obtenida a través de la simulación.

La simulación es una herramienta de gran valor para el ingeniero e investigador. Sus campos de aplicación son múltiples y cada vez más importantes y frecuentes en el desarrollo de aplicaciones cartográficas de control debido a las ventajas que presenta para el estudio deductivo de ciertas formulaciones que de otra manera serían inabordables (Ariza, 2002).

La simulación está relacionada con la idea de sistema, mostrando mayor interés cuando la operación con el sistema real no es posible. En estos casos, se procede a realizar un modelo de tipo lógico-matemático de forma que, se establecen unas leyes físicas o estadísticas del comportamiento del sistema. El estudio, análisis, validación y verificación del modelo se puede realizar con métodos analíticos, numéricos o mediante simulación, constituyendo ésta una herramienta importante y más sencilla para casos en los que el sistema presente mayor complejidad.

Según Ríos, D. y Ríos, S. (1997), se puede decir que la simulación es simplemente la construcción de un programa de ordenador que describiendo el comportamiento de un sistema mediante un modelo permite obtener conclusiones de valor para apoyar la toma de decisiones. Se trata de una metodología de análisis basada en un soporte informático y en la teoría de sistemas (Ariza, 2002).

Con la simulación, en este caso, se puede probar el comportamiento del sistema antes de comprometer recursos para su implementación. Al realizar el proceso por ordenador, el análisis permite jugar con las distintas variables para estudiar cómo interactúan.

La simulación realizada permite la determinación del tamaño de la muestra de elementos lineales. Para ello se utiliza la técnica bootstrapping.

El bootstrapping se basa en que muchas veces puede ser mejor extraer conclusiones sobre las características de la población estrictamente a partir de la muestra que se maneja, que haciendo asunciones quizás poco realistas sobre la población. El bootstrapping implica remuestreo de los datos obtenidos en una muestra, con reemplazamiento, muchas veces para generar una estimación empírica de la distribución muestral completa. El bootstrapping puede considerarse como un tipo especial de simulación denominada simulación basada en los datos.

La técnica consiste en extraer aleatoriamente muestras y analizar las estimaciones derivadas de cada una de ellas frente al valor poblacional correspondiente y conocido. El proceso de simulación para la determinación del tamaño de la muestra se basa en la construcción de muestras de la población con paquetes de una longitud determinada, de las que se obtiene con cada método de control posicional una función de distribución. Esta función se compara con la de la población total mediante las técnicas vistas en el punto III.6. Este proceso se repite un número de veces (normalmente se realizan 1000 iteraciones) para cada paquete de longitud. A continuación se incrementa el tamaño del paquete de longitud, de forma que se llega a un valor de longitud para el cual los resultados de la prueba de Kolmogorov-Smirnov verifican la similitud entre las funciones de distribución comparadas. Esta longitud puede suponerse como la muestra mínima para este método de control.

Junto con el estudio del tamaño muestral también se incorpora el cálculo de la densidad de puntos de toma para realizar el control a partir de líneas obtenidas con observaciones GPS de las vías de comunicación. Este intervalo mínimo se calcula eliminando puntos o lo que es lo mismo, manteniendo “uno de tantos” puntos de las bases de datos de control y comparando los resultados con los obtenidos de la base de datos completa mediante el Test de Kolmogorov-Smirnov (III.6). De esta forma, se verifica la similitud entre las funciones de distribución comparadas. La técnica de mantenimiento de uno de tantos puntos se repite hasta obtener un valor de “tantos” que marque el intervalo mínimo de observación. La utilización de esta técnica de eliminación de puntos (manteniendo un punto de tantos) se fundamenta en simular las condiciones de una variación en la captura de puntos en campo, condicionadas por un aumento de velocidad del vehículo que redunde en la mejora de rendimientos o un aumento de la seguridad vial durante la toma, etc.

III.8. Aplicación de los métodos de control a líneas sintéticas

Los métodos de control posicional por elementos lineales utilizados en la presente tesis doctoral se aplican a una serie de líneas en las que se simulan errores que pueden darse en la realidad. Estos errores simulados, responden a desplazamientos sistemáticos de una línea con respecto a otra, variaciones producidas por un ruido aleatorio o errores provocados por la inexactitud posicional de los puntos iniciales y finales de la línea.

Con esta técnica, se pretende conocer cómo responde cada método a estos posibles errores habituales en cualquier base de datos geográfica, con la idea de poder reconocer qué métodos son más apropiados para realizar el control y lo que es más

importante, si es posible interpretar la existencia de un error conocido estudiando los resultados de los cuatro métodos globalmente.

Las líneas sintéticas se presentan en la Figura III. 13.

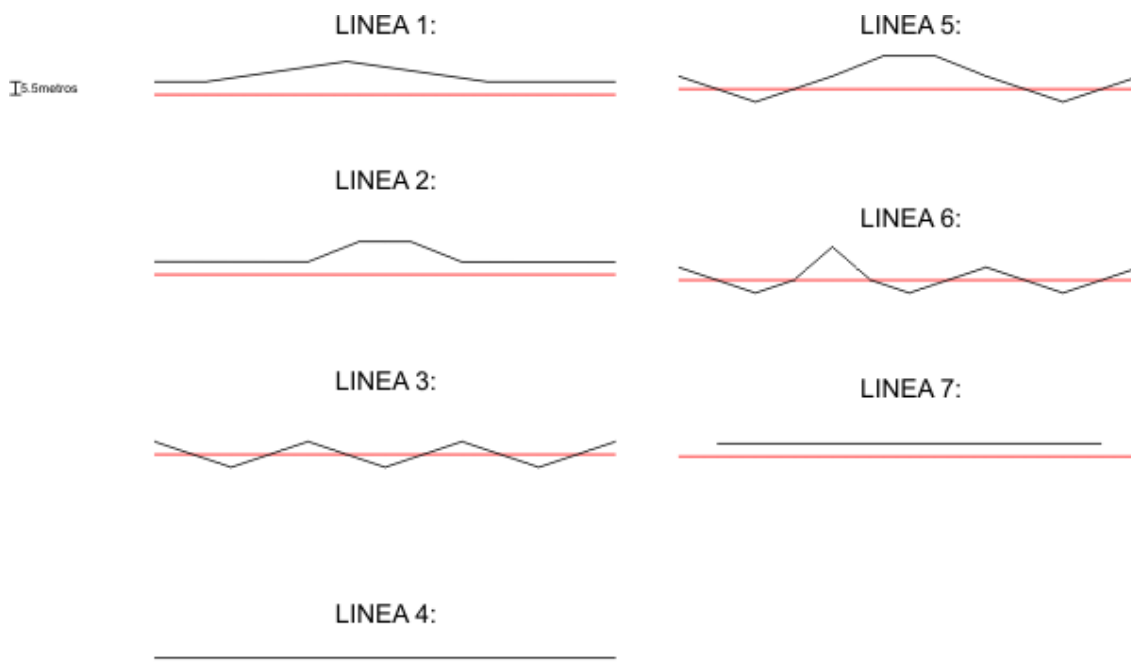


Figura III. 13. Líneas sintéticas

Como se puede observar en la Figura III. 13, en las líneas de color negro se ha simulado un error. En todas ellas se incluye un desplazamiento sistemático de 5.5 m. Además, a cada una se ha introducido otro error individualizadamente. Estos errores simulados se corresponden con:

- Línea 1: Presenta un error puntual que no se mantiene. Gráficamente se muestra con un pico.
- Línea 2: Presenta el mismo error puntual que caso anterior pero se mantiene a lo largo de una cierta longitud. Gráficamente se muestra con un pico que se mantiene una cierta longitud.
- Línea 3: Presenta un error tipo zigzag simulando el producido por errores aleatorios. Gráficamente se muestra con una línea zigzagueando sobre la otra.
- Línea 4: Presenta únicamente el desplazamiento sistemático propuesto en todos los casos.
- Línea 5: Se corresponde con la suma del caso 2 y 3. Se incluye un error puntual que se mantiene cierto tramo y un error zigzagueante. Gráficamente se muestra con una línea zigzagueando sobre la otra y un tramo con un error puntual que se mantiene.

- Línea 6: Se corresponde con la suma del caso 1 y 3. Se incluye un error puntual que no se mantiene y un error zigzagueante. Gráficamente se muestra con una línea zigzagueando sobre la otra y un tramo con un error puntual que no se mantiene.
- Línea 7: Presenta un error en la edición de los puntos iniciales y finales de la línea.

Con estos casos prácticos, se estudia como se comporta cada método de control al aplicarse sobre los errores simulados. Este tipo de estudio supone una novedosa herramienta para conocer dicho comportamiento.

III.9. Aplicación a bases cartográficas

A partir de la metodología propuesta, se realiza una aplicación práctica de la misma con casos reales, para ensayar el funcionamiento de la misma. Resulta evidente, que el ensayo con datos reales genera unos resultados que dependen de los datos empleados, aunque en el caso de nos ocupa la gran extensión de las carreteras utilizadas para el control posicional hace que la aplicación, a priori, resulte bastante completa.

Los datos utilizados para aplicación práctica proceden del proyecto de investigación “ConPoCar” del Ministerio de Educación y Ciencia (BIA2003-02234) mencionado anteriormente y dentro del cual se desarrolla esta tesis doctoral.

III.9.1. Material utilizado

Los datos empleados son, por un lado, datos digitales de bases cartográficas oficiales en España y por otro, datos obtenidos en campo con observaciones GPS, cuyas características se estudian en el punto III.4.2.

Las bases cartográficas utilizadas son:

- MTA: Es el Mapa Topográfico de Andalucía a escala 1:10.000 en su versión vectorial y estructurada como Base de Datos Geográfica (BDG). Se trata del principal producto del Instituto de Cartografía de Andalucía (ICA). Sus especificaciones están establecidas parcialmente en Gutiérrez y Sampayo (1981), y para el modelo vectorial en ICA (2002). Puede encontrarse una descripción general en Corral y col. (2000, 2001). En el Anexo II se incluye una ficha que describe en mayor medida este producto. Se utilizan las capas correspondientes a las vías de comunicación.
- BCN: Se trata de la Base Cartográfica Numérica a escala 1:25.000 producida por el Instituto Geográfico Nacional (IGN) y que cubre, prácticamente, todo el territorio nacional. Esta serie cartográfica procede de los flujos de producción del MTN25 y cuyas especificaciones están establecidas parcialmente en CSG (1985, 1992). En Mapping (2007) puede encontrarse una descripción general de esta BDG. En el Anexo II se incluye una ficha que describe en mayor medida este

producto. Igualmente, se ha trabajado con las capas correspondientes a las vías de comunicación.

Los dos conjuntos de datos son bases de datos vectoriales, estructuradas en capas, por lo que la extracción de las vías de comunicación es inmediata. Para el estudio que nos ocupa, se utilizan las vías asfaltadas.

Por tanto, no se realiza proceso degradador de la calidad posicional de los elementos a utilizar. Lo interesante de utilizar estos dos productos recae en que se tratan de las series nacionales y autonómicas oficiales en Andalucía. La serie 1:10.000 del MTA pertenece a la clasificación de escalas grandes mientras que la 1:25.000 se considera entre las medias.

En lo que respecta a los elementos lineales presentes interesa conocer la densidad de puntos utilizada en su definición geométrica. La Tabla III. 2 y la Tabla III. 3 presentan unas estadísticas básicas al respecto, las cuáles indican que existe una clara diferencia entre las dos bases de datos en cuanto a la distancia media, con unos 8 m de diferencia entre ambas. Sin embargo, para los valores medios, la diferencia determinada no viene a indicar una relación de escalas 10:25 como la que se deriva de la aplicación de la Ley de Töpfer y Pillewizer (1966). Los valores mínimos, máximos y desviaciones son siempre inferiores para el MTA. Los valores que se presentan por zonas no vienen a indicar la existencia de un patrón común en los dos conjuntos de datos.

Zona	Casos[n]	Media[m]	Mínimo[m]	Máximo[m]	Desviación[m]
Almería	197365	46.61	2.66	136.20	31.057
Cádiz	137084	26.29	4.30	213.38	25.456
Córdoba-Jaén	352101	53.44	6.60	141.16	24.935
Huelva	199834	53.45	9.47	202.25	27.624
Málaga	171217	21.18	2.72	70.83	13.120
Sevilla	152936	31.97	7.16	146.09	25.959
Total	1210537	41.98	2.66	213.38	28.497

Tabla III. 2. Resumen estadístico de distancia media entre vértices de los elementos lineales en el MTA

Zona	Casos[n]	Media[m]	Mínimo[m]	Máximo[m]	Desviación[m]
Almería	239347	45.80	4.43	234.05	32.268
Cádiz	148068	69.89	10.88	256.42	29.969
Córdoba-Jaén	356605	55.13	3.40	236.85	33.005
Huelva	194754	43.01	6.22	160.01	26.999

Málaga	173315	37.27	6.44	147.28	22.774
Sevilla	155448	49.98	9.73	191.58	23.882
Total	1267537	50.16	3.40	256.42	30.782

Tabla III. 3. Resumen estadístico de distancia media entre vértices de los elementos lineales en la BCN

Resulta interesante conocer el comportamiento respecto a la exactitud posicional planimétrica de estas dos bases cartográficas. Trabajos como el de Sevilla (1991) y normas como ASPRS (1989) establecen precisiones estándar generales para diversas escalas cartográficas. No hay constancia de la existencia de ningún estudio riguroso, oficial y público, que informe de las exactitudes de ninguna de las dos por lo que se va a asumir que cumplen con dos criterios clásicos de producción cartográfica, es decir, la regla del límite de percepción visual (normalmente establecido en 0.2mm) y la compilación a escala doble de la final.

En el caso del MTA las hipótesis anteriores hay que matizarlas desde el conocimiento de su linaje. Este producto proviene de un mapa analógico a escala 1:10.000, escaneado y luego vectorizado, por lo que, en el mejor de los casos sus propiedades métricas podrían ser las propias de la publicación, es decir, las correspondientes a una 1:10.000. La BCN es un producto digital desde la restitución original, por lo que se puede suponer que la calidad posicional de la misma se mantiene en todo el proceso. En este caso, se puede considerar una escala equivalente 1:12.500. Dada la similitud final entre ambas bases cartográficas se considera una única escala y equivalente a 1:10.000. Según esto, y considerando la producción sobre papel a la que los dos productos iban orientados en un principio, se puede establecer en 4m (0.4×10000) el valor del error máximo admisible (López-Cuervo, 1980), entendiendo por máximo una probabilidad lineal del 90%, lo anterior llevaría a considerar desviaciones estándar en el producto del orden de 2.41m, lo que está en el orden de las recogidas en referencias como Sevilla (1991).

III.9.2. Zona de estudio

La zona de trabajo utilizada es la que se presenta en la Figura III. 14. Como se puede observar, se conforma como un conjunto de seis bloques a partir de la distribución del territorio en hojas cartográficas de la serie MTN25. Los bloques son: Almería (AL), Cádiz (CA), Córdoba-Jaén (CoJa), Huelva (HU), Málaga (MA), Sevilla (SE).

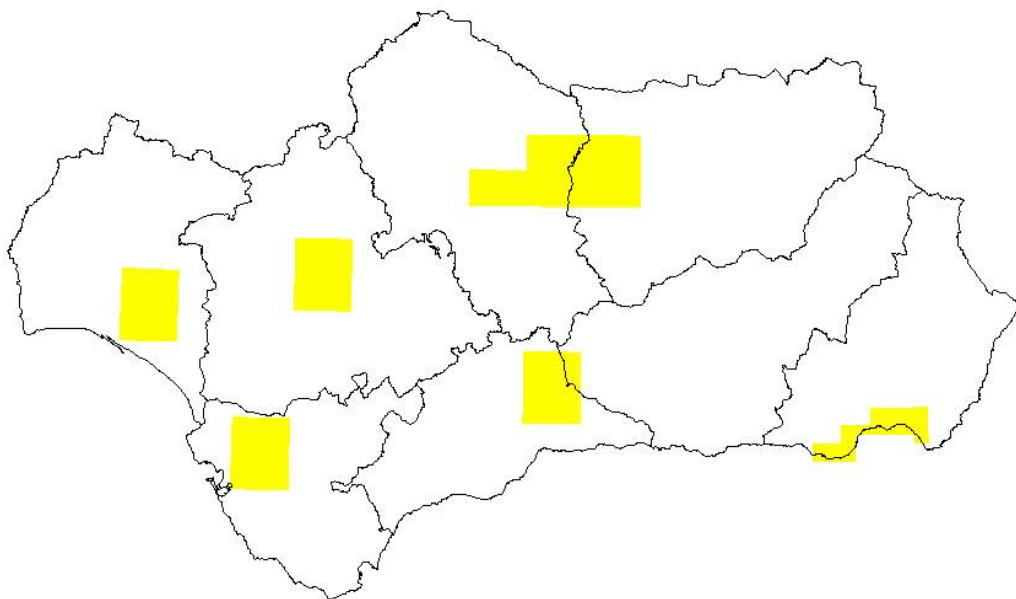


Figura III. 14. Zona de estudio

Las hojas utilizadas se presentan en la Tabla III. 4.

Zona de Almería			
EL EJIDO (1058)		ALMERIA / SUR (1045)	
1058-II	Roquetas de Mar	1059-I	Almería Sur
1058-III	San Agustín	1059-II	Cabo de Gata
1058-IV	Las Marinas	1045-III	Almería
-	-	1045-IV	El Alquíán
Zona de Cádiz			
JEREZ DE LA FRONTERA (1048)		PATERNA DE RIVERA (1062)	
1048-I	Nueva Jarilla	1062-I	El Portal
1048-II	Gibalbín	1062-II	La Barca de la F.
1048-III	Jerez de la Frontera	1062-III	Puerto Real
1048-IV	Jédula	1062-IV	Paterna de Rivera
Zona de Córdoba-Jaén			
MONTORO (903)		ANDUJAR (904)	

903-I	Ermita de S. Judas	904-I	La Goleta
903-II	Casas de la Onza	904-II	Peñallana
903-III	Montoso	904-III	Andujar
903-IV	San Julián	904-IV	Villanueva de la Reina
CORDOBA (923)		BUJALANCE (924)	
923-I	Cerro Muriano	924-I	El Carpio
923-II	Villafranca de C.	924-II	Villa del Río
923-III	Córdoba	924-III	Bujalance
923-IV	Los Cansinos	924-IV	Cañete de las Torres
ARJONA (925)			
925-I	Arjona	925-III	Porcuna
925-II	La Higuera	925-IV	Villardompardo
Zona de Huelva			
LA PALMA DEL CONDADO (982)		MOGUER (1000)	
982-I	Beas	1000-I	Moguer
982-II	Embalse Curumbel	1000-II	Rociana del Cond.
982-III	Trigueros	1000-III	Casa de Peñuelas
982-IV	La Palma	1000-IV	Los Cabezudos
Zona de Málaga			
ARCHIDONA (1024)		COLMENAR (1039)	
1024-I	Archidona	1039-I	V. del Rosario
1024-II	Estación de Salinas	1039-II	Periana
1024-III	Huertas del Río	1039-III	Casabermeja
1024-IV	V. del Trabuco	1039-IV	Colmenar
Zona de Sevilla			
LORA DEL RIO (963)		ALCALA DE GUADAIRA (985)	

	Almería	Cádiz	Córdoba Jaén	Huelva	Málaga	Sevilla
Organismo	IGN	IGN	IGN UJA	IGN	IGN	IGN
Distancia Media [m]	5500	23000	33500 56000	32000	31000	117000

Tabla III. 5. Resumen de características descriptivas de la zona de estudio

En cuanto a la cota media, cuatro zonas están por debajo de los 100 m y se corresponden con zonas costeras y llanuras fluviales bajas y llanas. La zona de Almería es una zona compleja que incluye llanuras costeras como el Campo de Dalías y la desembocadura del Río Andarax y el pie de monte de la Sierra de Alhamilla. La zona de Cádiz se corresponde con un área que incluye marismas, y campiñas bajas alomadas. La extensa área de Córdoba-Jaén se presenta como una franja alargada que incluye el valle del Guadalquivir y las estribaciones de Sierra Morena, al norte, y terrenos de campiña al sur. La zona de Huelva está atravesada por el Río Tinto y presenta un relieve muy suave. La zona norte de Málaga es un área de campiña alta y ondulada que evoluciona hacia una serranía; la zona sur se corresponde ya con un área abrupta de sierras de media altura. Finalmente, la zona de Sevilla incluye el valle del Guadalquivir a la altura de Alcolea del Río y un poco de las estribaciones de Sierra Morena al norte, y las campiñas de Carmona hacia el sur.

Las curvas hipsométricas correspondientes al territorio de Andalucía (línea continua) y al conjunto de las zonas de estudio (línea a trazos) son las que se representan en la Figura III. 15. Como se aprecia las zonas de trabajo presentan una clara predominancia de altitudes bajas. Ello es debido a zonas costeras como Almería pero principalmente a las amplias zonas del bajo valle del Guadalquivir (Huelva y Sevilla).

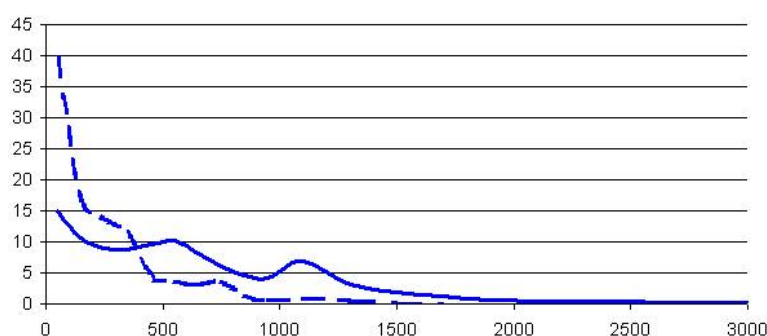


Figura III. 15. Curvas hipsométricas de la zona de estudio

En relación al litoral, se pueden considerar una zona plenamente costera (Almería) y tres zonas con influencia diversa y parcial (Cádiz, Huelva y Málaga). Las zonas de Córdoba-Jaén y Sevilla se consideran interiores.

La red viaria se caracteriza tanto por la estructura como la densidad de vías de comunicación. En relación a la estructura las configuraciones que se presentan son muy diversas y, en algún caso, hasta difícil de caracterizar. La estructura de la red en Almería es muy horizontal condicionada por su paralelismo a la costa. En Cádiz se trata de una estructura dispersa, algo radial descentrada y laxa. La zona de Córdoba y Jaén presenta un comportamiento complejo, con algunas subestructuras radiales y un claro eje horizontal. En Huelva se aprecia claramente una distribución radial y un perímetro circular, mientras que para el caso de Málaga se trata de una red sin una forma particular. Finalmente, en Sevilla se aprecia una estructura multiradial.

Presenta gran interés también, la relación de cada una de estas zonas con las estaciones de referencia GPS oficiales utilizadas en el procesamiento de las observaciones. Se incluye en la Tabla III. 5 para cada zona el nombre de la localidad en la se sitúa la estación de referencia operativa más cercana con capacidad de suministrar registros con frecuencia de un segundo. Todas las estaciones, salvo la de Jaén, son titularidad del IGN. La distancia media que aparece se refiere al total del área de cada una de las subzonas de trabajo.

III.9.3. Elementos de control (GPS)

Una vez descritas las series cartográficas a controlar y la zona de estudio donde se enmarca la aplicación metodológica de la tesis doctoral, se procede a presentar los elementos lineales utilizados para el control de dicha cartografía.

El control se realiza a partir del levantamiento del eje de las vías de comunicación representadas en las hojas cartográficas utilizadas. La adquisición de estos elementos de control viene ampliamente descrita en Ruiz, Mozas y Ariza (2006) y resumida en el punto III.4.2 por lo que a continuación se expondrá una descripción de los mismos. Las Figura III. 16, Figura III. 17, Figura III. 18, Figura III. 19, Figura III. 20 y Figura III. 21 presentan los ejes obtenidos con la observación GPS una vez realizada la fase de edición de la base de datos (ver punto III.4.2). Los datos procedentes de la observación GPS tienen una exactitud posicional de 1.5 m. El cálculo de errores de esta metodología se encuentra descrito en Ruiz, Mozas y Ariza (2006).

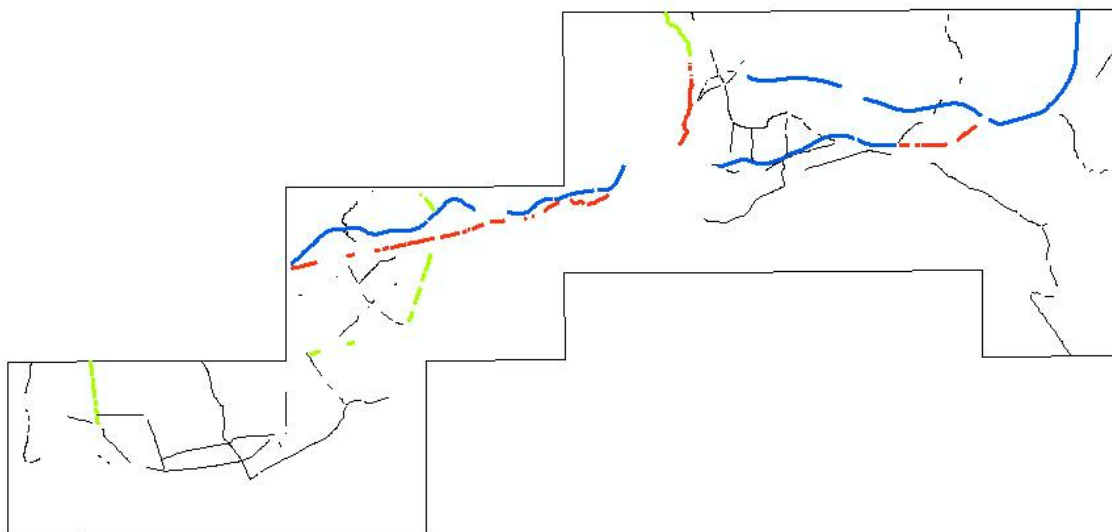


Figura III. 16. Líneas de control GPS. Almería

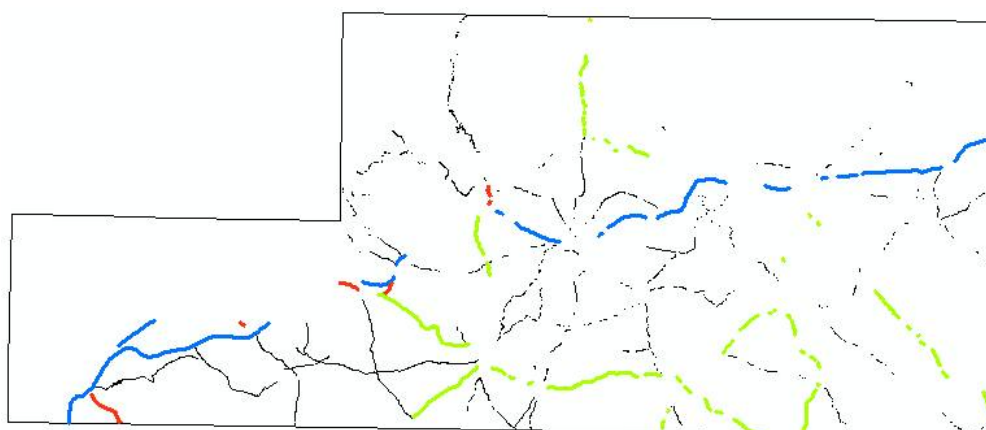


Figura III. 17. Líneas de control GPS. Córdoba-Jaén

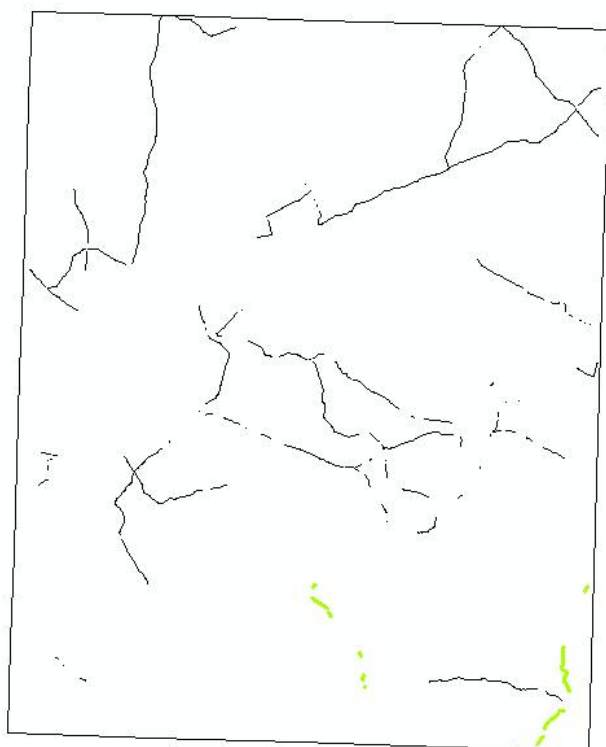


Figura III. 18. Líneas de control GPS. Cádiz

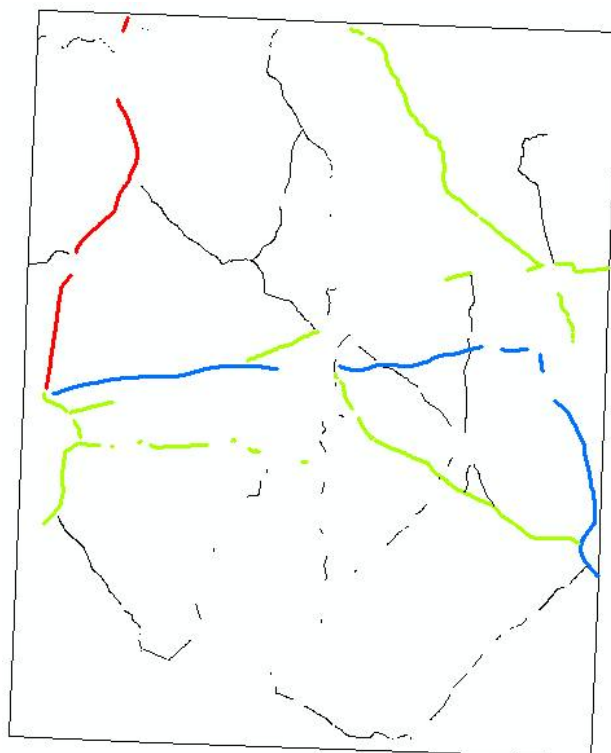
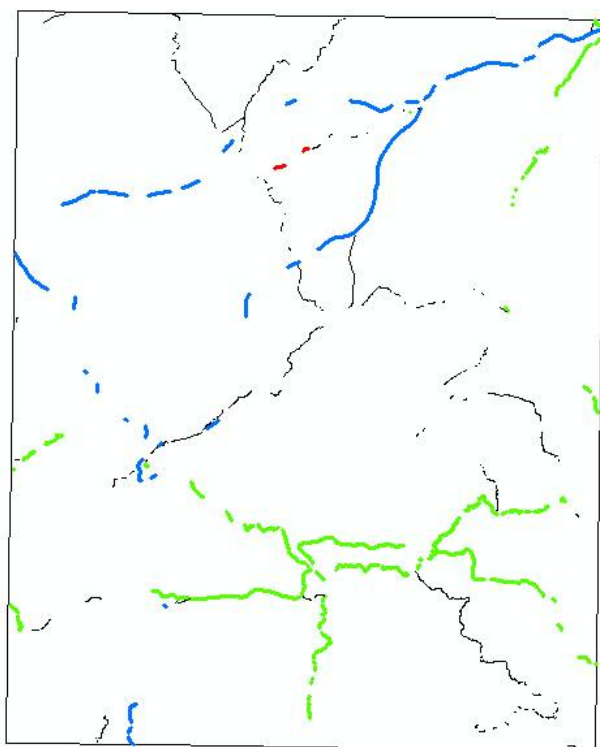
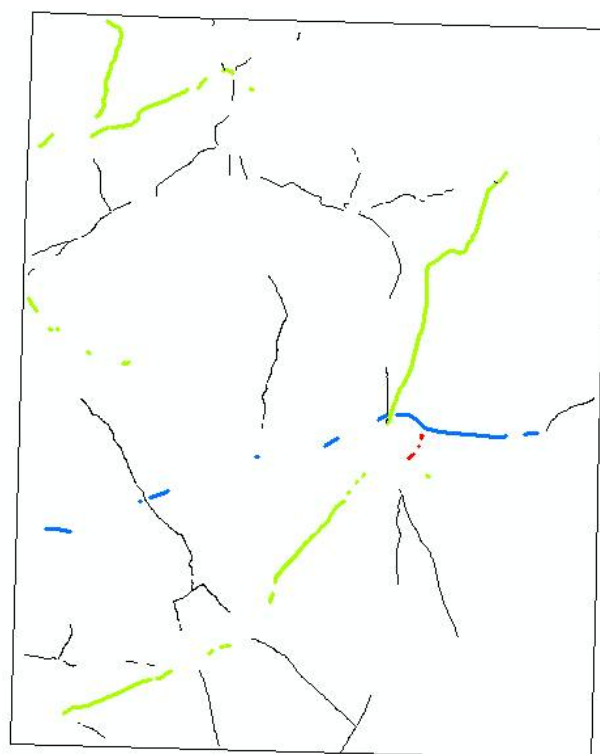
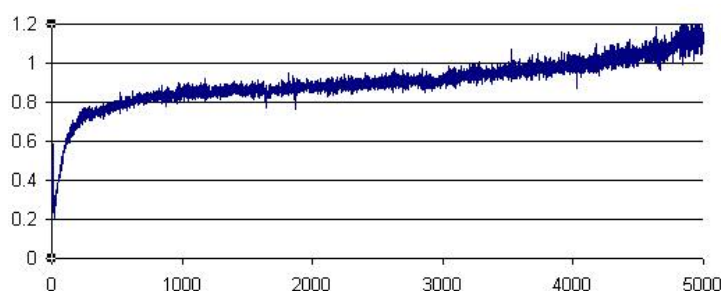


Figura III. 19. Líneas de control GPS. Huelva

**Figura III. 20. Líneas de control GPS. Málaga****Figura III. 21. Líneas de control GPS. Sevilla**

Las líneas GPS son procesadas en código diferencial. En total son algo más de 450000 puntos capturados, incluyendo sólo aquellos puntos con calidades mejores a 1.5 m.

El comportamiento espacial de los valores de calidad posicional queda representado por el variograma experimental que muestra la Figura III. 22. Este variograma se ha creado para parejas de puntos sobre la misma cadena-gps y con distancias hasta de 5000 m. En ella se observa claramente el efecto pepita y la evolución sostenida hasta alcanzar una meseta con tendencia creciente relativamente estable entorno a los 0.8 m en un alcance de unos 500 m.

**Figura III. 22. Variograma para las diferencias de calidad posicional planimétrica de los datos GPS**

En la Figura III. 23 se presenta la evolución del coeficiente de correlación entre valores de calidad posicional GPS para puntos que se encuentran a X saltos o tomas. Se recuerda que estas tomas se efectúan con una frecuencia de 1 Hz, es decir, cada segundo, y que por ello, a la velocidad de cruce del vehículo de captura cada salto supone, en término medio unos diez m. La Figura III. 23 muestra una autocorrelación alta (superior al 90%) para saltos cercanos, el 80% se alcanza a unos 10 saltos ($\approx 100\text{m}$) y para 100 saltos ($\approx 1000\text{m}$) los valores ya son cercanos al 50%.

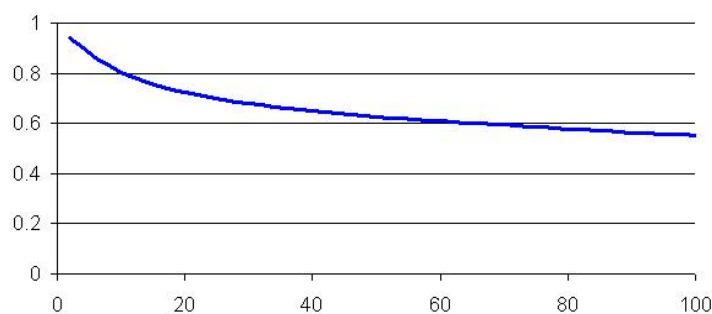


Figura III. 23. Autocorrelaciones de los valores de calidad GPS

La distancia entre puntos de los elementos lineales observados con GPS depende en gran medida de la velocidad de cruce en la captura. El valor medio del espaciado entre puntos es de 9.88 m (desviación 3.196 m) lo que equivale a una velocidad media de cruce materializada durante los trabajos de captura de 35.56 km/h. La distribución de estos valores en la población es la que se presenta en la Figura III. 24 y la Figura III. 25, tanto para las distancias como para las velocidades, que dependen directamente de las anteriores dado que la frecuencia temporal de muestreo es un factor fijo.

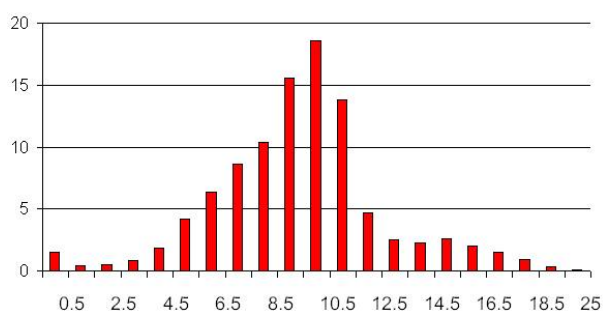


Figura III. 24. Histograma relativo al espaciado entre puntos

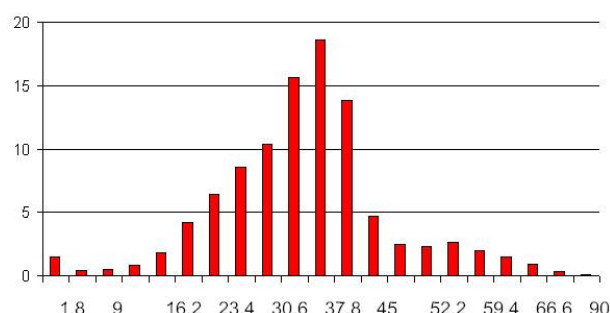


Figura III. 25. Histograma relativo a la velocidad entre puntos

III.9.4. Conclusiones

En este apartado se ha descrito el material y la zona de estudio aplicada en esta tesis doctoral. Las bases de datos cartográficas a controlar se corresponden con las series mayores completas a nivel nacional y autonómico. El control se realiza con los

ejes interpolados de las líneas observadas con GPS (correspondientes a los arcenes de ambos sentidos de la marcha) sobre las vías de comunicación de las zonas seleccionadas. Los datos GPS utilizados para realizar el control han sido descritos, exponiendo sus principales características.

Capítulo IV.

Resultados y Discusión

Capítulo IV. Resultados y Discusión

IV.1. Introducción

El capítulo IV se ocupa de los resultados obtenidos con la aplicación de la metodología de control de exactitud posicional con elementos lineales. Una vez presentado el método y el material de trabajo en el capítulo anterior, se desarrolla la aplicación sobre datos reales.

Este ensayo supone un resultado en sí mismo de esta investigación ya que como se expuso anteriormente, las aplicaciones realizadas en este campo por la comunidad científica hasta el momento se circunscriben a pequeñas bases de datos a las que se aplica un único método concreto de control.

Por otro lado, se incorporan datos procedentes de observaciones GPS cinemáticas, cuestión que marca considerablemente la base de datos de control utilizada incorporando novedosos planteamientos a estudiar.

En el capítulo que nos ocupa, se exponen también los resultados obtenidos de aplicar la metodología propuesta a líneas sintéticas, así como los provenientes de los estudios sobre el tamaño de la muestra o la cadencia mínima de toma de la observación GPS, etc.

La metodología propuesta, como se mencionado con anterioridad, se aplica de forma semiautomática mediante la herramienta informática CPLin (Control Posicional

por Líneas). Los diferentes módulos presentes en este software reciben los resultados postprocesados procedentes de la observación GPS, realizando las ediciones y depuraciones necesarias a las bases de datos de control y a controlar, y los procesos de enriquecimiento, identificación de elementos homólogos, tramificación, filtrado y selección de elementos. Evidentemente esta herramienta aplica de forma automática los métodos de control propuestos (MDH, MSK, MOS, MOD) y permite realizar los análisis planteados de las funciones de distribución resultantes.

La forma de plasmar los resultados obtenidos en los siguientes puntos es: Primero se muestran para cada método las tablas, gráficos, explicaciones, etc. necesarios para su presentación, dejando un apartado final para su interpretación y análisis.

IV.2. La herramienta CPLin

La mayor complejidad de las metodologías de control por elementos lineales frente a las metodologías basadas en puntos hace necesario el disponer una serie de herramientas adecuadas a este propósito.

Como primer resultado de esta tesis doctoral se presenta la herramienta informática que ha permitido, de forma semiautomática, la implementación de la metodología propuesta y la obtención de resultados analíticos procedentes de la aplicación de la misma sobre las bases cartográficas descritas en el capítulo anterior.

La aparición de esta herramienta supone una gran novedad en el campo del control posicional de bases de datos cartográficas a partir de elementos lineales, ya que, entre otras características hay que destacar que CPLin está exclusivamente planificado y desarrollado con esta finalidad.

Por otro lado, CPLin permite la aplicación de gran parte de las metodologías de control por elementos lineales aparecidas en los últimos tiempos, permitiendo el trabajo con abundantes líneas procedentes de bases de datos cartográficas y de control, optimizado, en este último caso, para la utilización de datos procedentes de observaciones cinemáticas GPS.

La herramienta CPLin es descrita ampliamente en Mozas, Ureña y Ariza (2007b), sin embargo, a continuación se presentan las principales características y funcionalidades de esta aplicación:

- Módulo de generación de base de datos de control a partir de GPS:
 - Lectura de puntos procedentes de una observación cinemática con GPS una vez procesados.
 - Transformación del marco de referencia de los datos:
 - Cálculo de altura ortométrica.
 - Transformación del sistema WGS84 al sistema ED50.
 - Transformación de coordenadas geográficas a la proyección UTM.

- Construcción de las entidades lineales a partir de la sucesión horaria de toma de cada punto.
- Módulo de generación de la base de datos cartográfica.
- Enriquecimiento de las bases de datos.
- Filtrado de elementos y eliminación de vértices de las bases de datos.
- Identificación de elementos homólogos en ambas bases de datos.
- Selección de elementos.
- Realización de test posicionales por elementos lineales:
 - Método de las Distancias de Hausdorff (MDH).
 - Método de Skidmore o Banda Épsilon (MSK).
 - Método del Orlado Simple (MOS).
 - Método del Orlado Doble (MOD).
- Obtención y exportación de resultados
- Comparación de funciones de distribución obtenidas mediante el test de Kolmogorov-Smirnov.

Como se puede observar, el esquema del programa CPLin se desarrolla de forma similar a la metodología presentada en el capítulo anterior.

En la Figura IV. 1 se presenta una captura, a modo de ejemplo, de una ventana de CPLin.

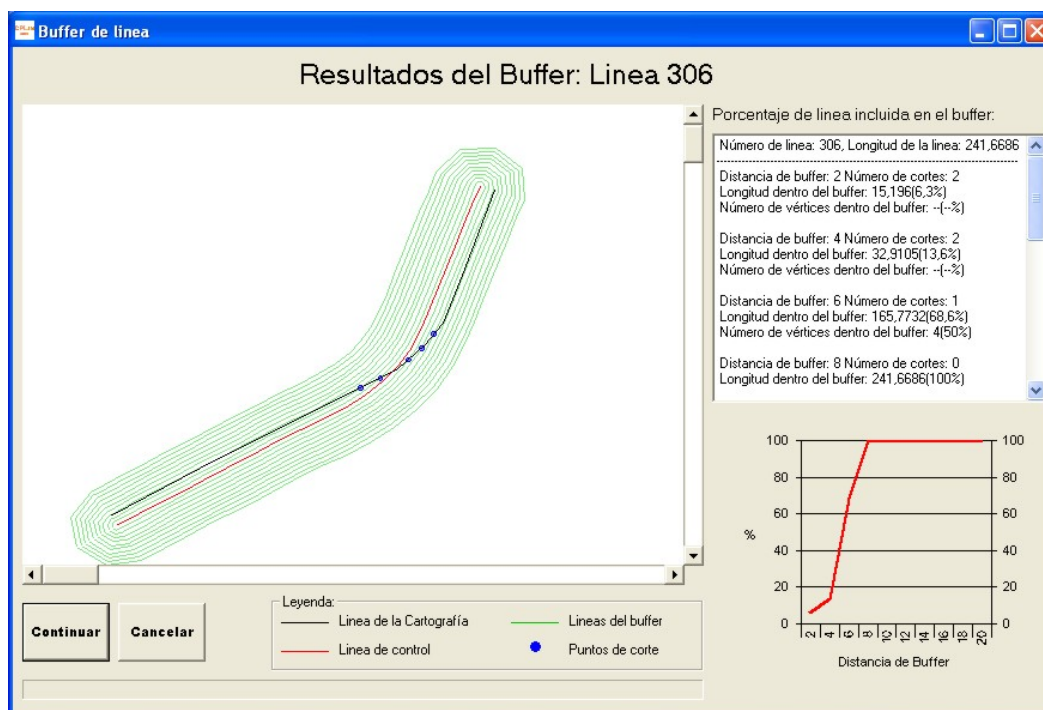


Figura IV. 1. Ejemplo de aplicación del método MOS con CPLin

La programación del software CPLin ha sido realizada con Microsoft Visual Basic 6.0. La elección de esta herramienta de desarrollo de aplicaciones se debe a su capacidad para generar soluciones informáticas de forma sencilla, aprovechando los recursos que ofrece al programador entre los que se puede destacar la facilidad de conexión a bases de datos Microsoft, en las que se ha asentado el sistema de gestión de datos utilizado por CPLin. Otra característica importante a destacar es la funcionalidad e interoperatividad de este lenguaje en plataformas Windows, ampliamente distribuidas en el mercado actual.

Aunque la herramienta CPLin tiene un gran potencial y se encuentra plenamente en funcionamiento, desde su nacimiento, esta aplicación fue concebida como un proyecto abierto a las novedades e innovaciones existentes en cada momento, por lo ha sido diseñada para que pueda adaptarse e incorporar las nuevas técnicas y procedimientos evaluadores que la comunidad científica pudiera promover dentro del campo del control posicional por elementos lineales.

IV.3. Resultados obtenidos en la aplicación a líneas sintéticas

El resultado de aplicar la metodología propuesta a las líneas sintéticas ya presentadas en el capítulo anterior se muestra a continuación para cada método de control.

IV.3.1. Resultados del Método de Distancias de Hausdorff (MDH)

La Tabla IV. 1 muestra los resultados de aplicar el método MDH a las líneas sintéticas planteadas.

Línea	Imagen	Longitud [m]	Máx. Pt. Cartog. -> Control [m]	Máx. Pt. Control -> Cartog. [m]	Media Pt. Cartog. -> Control [m]	Media Pt. Control -> Cartog. [m]	Distancia Hausdorff [m]
1		198	14.204	14.060	8.088	8.030	14.204
2		198	14.204	14.204	7.382	7.217	14.204
3		198	5.5	5.218	2.825	2.680	5.5
4		198	5.5	5.5	5.5	5.5	5.5
5		198	14.204	14.204	5.600	5.327	14.204
6		198	14.204	10.765	3.532	3.130	14.204
7		165	17.393	5.5	6.633	5.5	17.393

Tabla IV. 1. Resultados Método MDH líneas sintéticas

Los resultados revelan la existencia de valores más altos de distancia de Hausdorff en las líneas 7, 1, 2, 5 y 6, mientras que las líneas 3 y 4 presentan valores menores. Estos resultados son lógicos debido a la existencia en estas líneas de errores puntuales que tienen gran relevancia en esta medida. Los valores de distancia

obtenidos desde los vértices de la línea de la cartografía a la línea de control son en todos los casos iguales o superiores a las distancias entre los vértices de control y la línea de la cartografía.

En cuanto a los valores medios de distancia obtenidos, las diferencias en todos los casos son inferiores al anterior resultado, debido a que los efectos de los errores puntuales se compensan a lo largo de la línea. El mejor comportamiento se da en la línea 3 y los máximos en las líneas 1 y 2.

La línea 4 (con un desplazamiento sistemático) tiene valores iguales de distancia máxima y media. Sin embargo, para el resto de líneas los valores obtenidos de las distancias máximas y medias varían. De esta forma, se observa en los valores máximos obtenidos una gran sensibilidad a cualquier desplazamiento puntual, ya sea o no mantenido una cierta longitud. También estos valores máximos son muy sensibles a la falta de precisión en la determinación de los inicios y finales de las líneas (línea 7). Sin embargo, los valores medios de distancias tienen menor sensibilidad a estos errores puntuales, ya que este error se compensa a lo largo de la línea. Como es lógico, los valores medios son menores cuando existe un desplazamiento aleatorio simulado, mostrándose superiores cuando se simula un desplazamiento sistemático.

Una interpretación de estos resultados muestra que este método puede dar información de la existencia de errores puntuales o falta de precisión en la determinación de los inicios y finales de líneas. Para esto, sólo hay que detectar valores máximos (distancias de Hausdorff) muy por encima de los medios (distancias medias). Sin embargo, esta información sólo permite asegurar la presencia de estos errores, sin que dictamine el número de ellos y la magnitud de los mismos.

IV.3.2. Resultados del Método de la Banda Épsilon o Skidmore (MSK)

Los resultados presentados en la Tabla IV. 2 muestran los valores obtenidos del ensayo del método MSK a las líneas sintéticas.

Línea	Imagen	Longitud [m]	Número de cruces	Superficie [m ²]	Sup/Lon [m]
1		198	0	1615.610	8.160
2		198	0	1472.152	7.435
3		198	6	181.500	0.917
4		198	0	1089.000	5.500
5		198	4	1018.502	5.144
6		198	6	597.371	3.017
7		165	0	998.250	6.050

Tabla IV. 2. Resultados Método MSK líneas sintéticas

Los resultados muestran valores más bajos de desplazamiento (Superficie/Longitud) en las líneas 3 y 6 y mayores para las líneas 1 y 2. El resto de líneas muestran un comportamiento cercano a los 5-6 m de desplazamiento.

Los resultados del método MSK son interpretables como un valor medio del desplazamiento. Los errores aleatorios son compensados, mientras que el resto se comporta de manera parecida a los valores medios del método MDH, por lo que aquellas líneas que simulan desplazamientos sistemáticos (línea 1 y línea 2) obtienen valores superiores a las que simulan desplazamientos aleatorios (línea 3). Por otro lado, la falta de precisión en los puntos inicial y final de la línea (caso 7), provoca un valor superior de desplazamiento.

IV.3.3. Resultados del Método de Orlado Simple (MOS)

La Tabla IV. 3 presenta los resultados de aplicar el método MOS a las líneas sintéticas planteadas para anchuras de orlado que abarcan desde 1 a 20 m incrementadas cada metro.

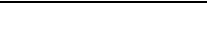
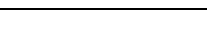
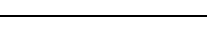
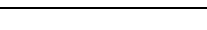
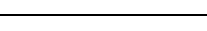
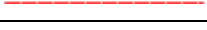
Línea	Imagen	Porcentaje (%) de longitud para cada anchura de orlado									
		1 m	2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m
1		0.0	0.0	0.0	0.0	0.0	42.17	49.22	56.27	63.32	70.37
2		0.0	0.0	0.0	0.0	0.0	66.92	69.61	72.31	75.01	77.71
3		18.18	36.36	54.55	72.73	90.91	100.0	100.0	100.0	100.0	100.0
4		0.0	0.0	0.0	0.0	0.0	100.0	100.0	100.0	100.0	100.0
5		10.12	20.23	30.35	40.47	50.59	57.38	60.86	64.33	67.81	71.28
6		15.95	31.9	47.85	63.8	79.75	88.43	89.84	91.25	92.66	94.07
7		0.0	0.0	0.0	0.0	0.0	84.55	86.99	89.11	90.12	91.13
Línea	Imagen	Porcentaje (%) de longitud para cada anchura de orlado									
		11 m	12 m	13 m	14 m	15 m	16 m	17 m	18 m	19 m	20 m
1		77.41	84.46	91.51	98.56	100.0	100.0	100.0	100.0	100.0	100.0
2		80.41	83.11	85.8	88.5	100.0	100.0	100.0	100.0	100.0	100.0
3		100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
4		100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
5		74.75	78.21	81.7	85.19	100.0	100.0	100.0	100.0	100.0	100.0
6		95.48	96.89	98.3	99.71	100.0	100.0	100.0	100.0	100.0	100.0
7		92.14	93.15	94.16	95.17	96.18	97.19	98.2	99.21	100.0	100.0

Tabla IV. 3. Resultados Método MOS líneas sintéticas

Los resultados obtenidos pueden ser más fácilmente interpretables si se representan gráficamente las funciones de distribución obtenidas (Figura IV. 2, Figura IV. 3, Figura IV. 4, Figura IV. 5, Figura IV. 6, Figura IV. 7 y Figura IV. 8).

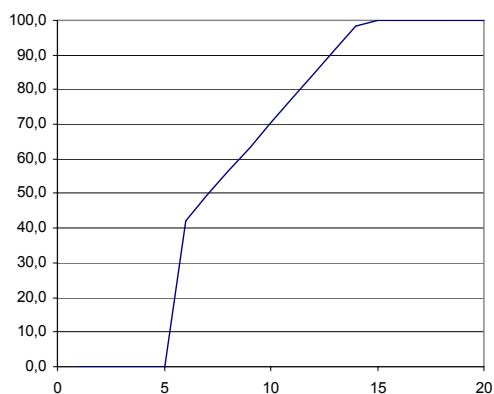


Figura IV. 2. Método MOS. Línea sintética nº1

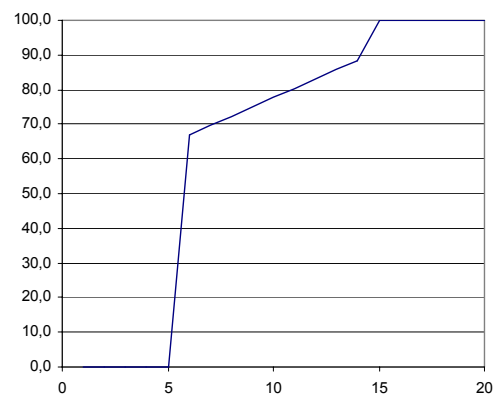


Figura IV. 3. Método MOS. Línea sintética nº2

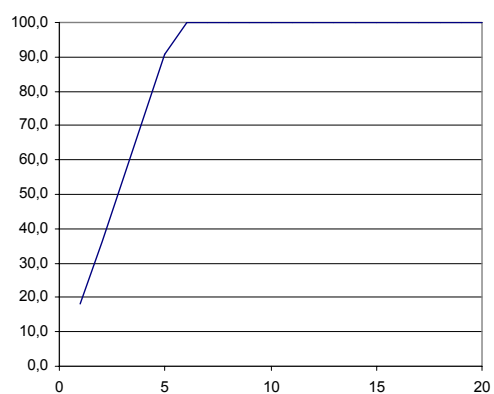


Figura IV. 4. Método MOS. Línea sintética nº3

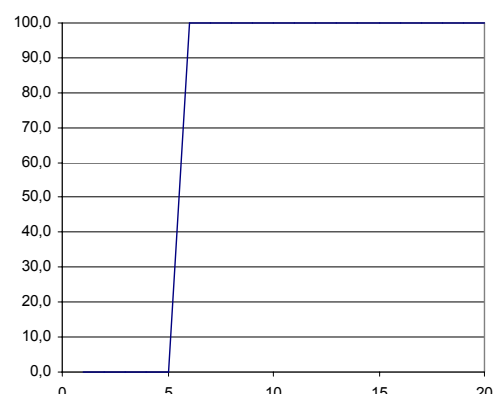


Figura IV. 5. Método MOS. Línea sintética nº4

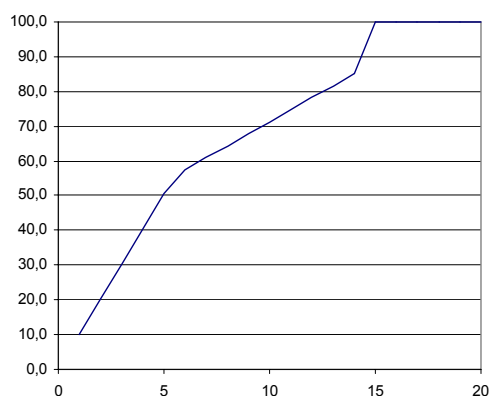


Figura IV. 6. Método MOS. Línea sintética nº5

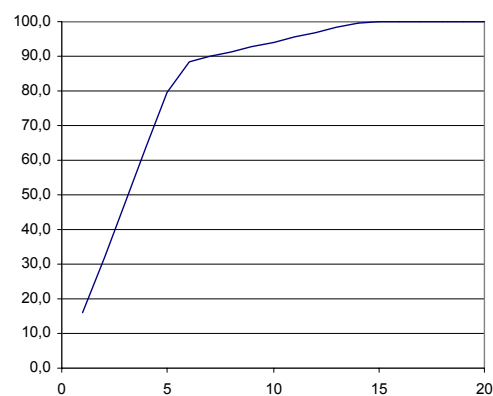


Figura IV. 7. Método MOS. Línea sintética nº6

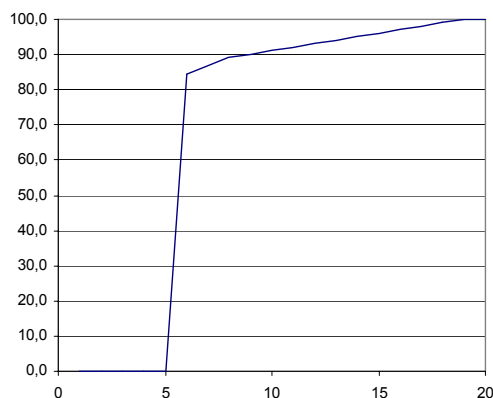


Figura IV. 8. Método MOS. Línea sintética nº7

En general, las líneas alcanzan el valor máximo del 100% de inclusión antes de utilizar la anchura de orlado de 15 m. Sólo para la línea 7 es necesario llegar hasta la anchura de 19 m para conseguir la plena inclusión de la línea en el orlado.

Por otro lado, la línea 4 presenta valores de inclusión global a partir de 6 m y nula inclusión con 5 m o menos de anchura, por lo que en el intervalo de 5 a 6 m se produce el salto de 0 a 100 %. El resto de líneas ofrecen tendencias más suaves con incrementos más fuertes en las anchuras de orlado más bajas y tendencias menos acusadas en valores más altos de anchura de orlado hasta alcanzar el porcentaje máximo.

Las líneas 1, 2, 4 y 7 presentan valores nulos hasta alcanzar la anchura de orlado de 5 m, mientras que el resto parten con porcentajes de entre un 10% y un 20%.

El método MOS aplicado a estas líneas sintéticas presenta resultados más difíciles de interpretar. Si se observan las curvas de las funciones de distribución obtenidas se puede inferir un desplazamiento sistemático cuando los valores de la

función son muy bajos. Por ejemplo, en la línea 1 y la línea 2 la curva arranca a partir del valor del desplazamiento inicial (5.5 m), tendiendo posteriormente a buscar un valor en el que el error simulado sea superado. En la línea 4, que sólo simula un desplazamiento sistemático, una vez alcanzado el valor de dicho desplazamiento, la curva llega al 100%. Por otro lado, las líneas con errores aleatorios simulados, al tener cruces con la línea de control, nunca parten de un porcentaje nulo y presentan curvas menos tendidas en las que se alcanza el 100% de una forma menos brusca. Por último, la línea 7 con errores en los puntos iniciales y finales tiene comportamientos similares a las líneas que simulan desplazamientos.

Los resultados obtenidos con este método aportan información acerca de aquellas líneas que tienen errores sistemáticos presentando curvas con saltos bruscos (línea 4) y aquellas que mayoritariamente poseen errores aleatorios donde las curvas tienen comportamientos más suaves (línea 2 y 5).

IV.3.4. Resultados del Método de Orlado Doble (MOD)

Como ya se describe con anterioridad en el apartado de la propuesta metodológica, la aplicación del método MOD genera tres valores resultantes para cada anchura de orlado analizada: Superficie común, desplazamiento promedio y oscilación.

IV.3.4.1 Superficie Común

La Tabla IV. 4 muestra los resultados correspondientes a la superficie común para las anchuras de orlado comprendidas entre 1 y 20 m incrementadas cada metro.

Línea	Imagen	Porcentaje (%) de superficie común para cada anchura de orlado									
		1 m	2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m
1		0.00	0.00	3.37	15.07	25.00	33.97	42.35	49.59	55.23	59.74
2		0.00	0.00	5.53	21.81	32.75	40.98	47.64	54.07	59.18	63.27
3		18.52	36.77	54.18	65.11	71.72	76.14	79.31	81.69	83.54	85.03
4		0.00	0.00	8.20	30.89	44.55	53.67	60.20	65.10	68.91	71.97
5		12.35	24.53	36.26	44.81	51.09	56.20	60.62	65.35	69.13	72.15
6		16.58	32.91	48.56	58.98	65.87	70.96	74.85	77.78	80.04	81.85
7		0.00	0.00	7.48	28.34	41.02	49.36	55.46	60.07	63.67	66.55
Línea	Imagen	Porcentaje (%) de superficie común para cada anchura de orlado									
		11 m	12 m	13 m	14 m	15 m	16 m	17 m	18 m	19 m	20 m
1		63.43	66.50	69.10	71.33	73.26	74.95	76.44	77.76	78.94	80.01
2		66.61	69.40	71.76	73.78	75.53	77.06	78.42	79.62	80.69	81.66
3		86.24	87.25	88.11	88.85	89.48	90.04	90.53	90.97	91.37	91.72

4		74.46	76.54	78.31	79.82	81.13	82.27	83.29	84.19	84.99	85.72
5		74.63	76.69	78.44	79.94	81.24	82.37	83.38	84.27	85.07	85.78
6		83.33	84.56	85.60	86.48	87.25	87.92	88.51	89.03	89.49	89.91
7		68.92	70.89	72.57	74.02	75.27	76.38	77.35	78.23	79.01	79.69

Tabla IV. 4. Resultados Método MOD (superficie común) líneas sintéticas

Estos resultados tienen su reflejo en las representaciones gráficas que se muestran en la Figura IV. 9, Figura IV. 10, Figura IV. 11, Figura IV. 12, Figura IV. 13, Figura IV. 14 y Figura IV. 15.

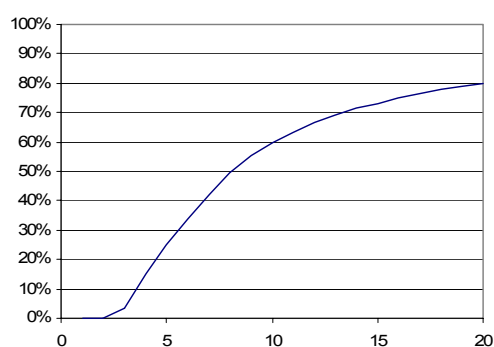


Figura IV. 9. Método MOD(SC). Línea sintética n°1

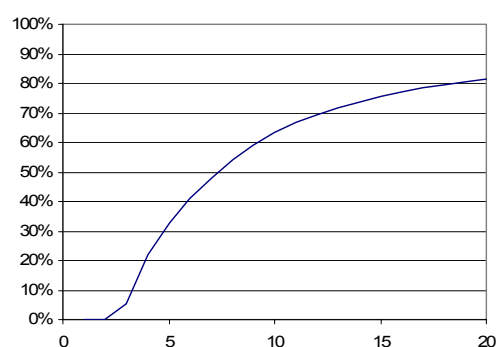


Figura IV. 10. Método MOD(SC). Línea sintética n°2

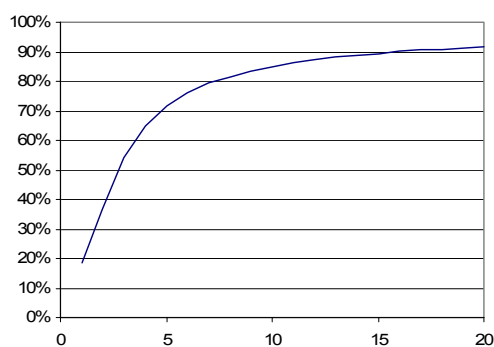


Figura IV. 11. Método MOD(SC). Línea sintética n°3

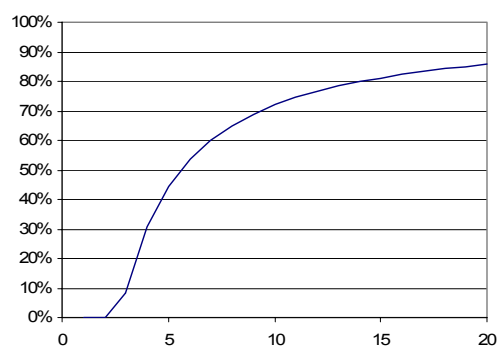


Figura IV. 12. Método MOD(SC). Línea sintética n°4

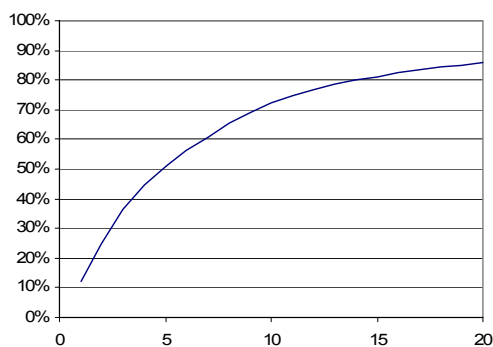


Figura IV. 13. Método MOD(SC). Línea sintética nº5

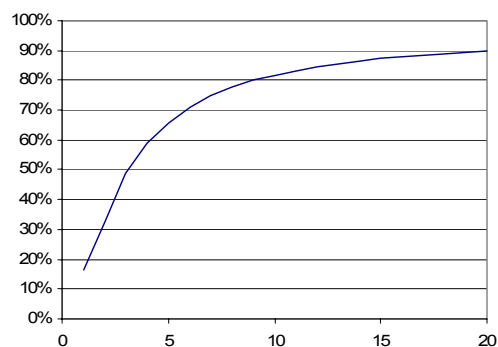


Figura IV. 14. Método MOD(SC). Línea sintética nº6

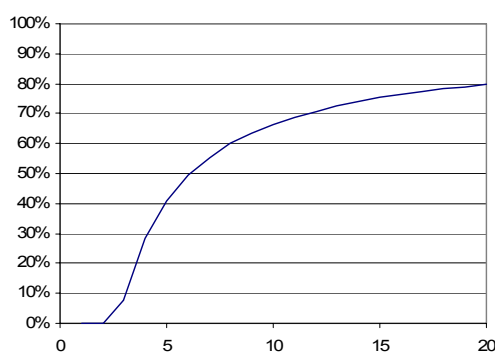


Figura IV. 15. Método MOD(SC). Línea sintética nº7

Los resultados reflejan curvas ascendentes a medida que se aumenta la anchura de orlado tendiendo a estabilizarse para los valores más altos de ésta. Las líneas 1, 2, 4 y 7 parten con valores nulos para las anchuras de 1 y 2 m, teniendo el resto unos valores mínimos en el intervalo de 10 a 20 m. Por otro lado, para una anchura de 20 m se alcanzan valores entre un 80% y un 90% de superficie común.

La línea 7 presenta una evolución parecida al resto, con una curva ascendente hasta alcanzar un valor estable en torno al 80% de superficie común.

Las gráficas correspondientes a la superficie común tienen aspectos similares. Sólo puede significarse diferencias en el valor de partida, ya que las líneas que tienen solo errores sistemáticos parten de valores nulos (línea 1, línea 2 y línea 4), mientras que el resto parte de un valor superior. Por otro lado, la línea 7 con errores en la determinación de los puntos iniciales y finales también presenta este arranque con valores nulos de superficie común, aunque en este caso, la curva sólo alcanza un porcentaje del 80%.

Las similitudes con las curvas obtenidas a partir del método MOS son evidentes, aunque con la diferencia de que la medida de la superficie común se comporta como una línea con tendencia a alcanzar un valor máximo que no podrá ser superado,

mientras que en el método MOS se puede alcanzar el 100% de inclusión en el orlado aumentando la anchura del mismo.

IV.3.4.2 Desplazamiento Promedio

El desplazamiento promedio obtenido tras aplicar el método MOD a las líneas sintéticas se presenta en la Tabla IV. 5 para las anchuras de orlado de 1 hasta 20 m incrementadas cada metro.

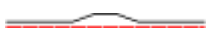
Línea	Imagen	Desplazamiento promedio para cada anchura de orlado [m]									
		1 m	2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m
1		3.122	6.244	9.051	10.603	11.698	12.352	12.575	12.557	12.539	12.521
2		3.09	6.181	8.755	9.647	10.355	10.889	11.252	11.257	11.233	11.21
3		2.415	3.714	3.974	3.961	3.943	3.923	3.902	3.881	3.86	3.841
4		3.142	6.283	8.652	8.685	8.71	8.733	8.753	8.771	8.789	8.807
5		2.607	4.469	5.628	6.461	7.117	7.605	7.929	7.908	7.858	7.81
6		2.365	3.752	4.223	4.383	4.45	4.429	4.356	4.284	4.215	4.148
7		3.764	5.252	7.242	7.261	7.262	7.284	7.281	7.276	7.269	7.266
Línea	Imagen	Desplazamiento promedio para cada anchura de orlado [m]									
		11 m	12 m	13 m	14 m	15 m	16 m	17 m	18 m	19 m	20 m
1		12.503	12.486	12.469	12.452	12.435	12.419	12.402	12.386	12.371	12.355
2		11.186	11.163	11.141	11.118	11.096	11.074	11.053	11.032	11.011	10.99
3		3.822	3.803	3.786	3.769	3.752	3.736	3.721	3.707	3.693	3.679
4		8.825	8.842	8.86	8.877	8.893	8.91	8.926	8.942	8.958	8.973
5		7.762	7.714	7.668	7.622	7.577	7.533	7.489	7.446	7.403	7.362
6		4.085	4.024	3.966	3.91	3.857	3.806	3.758	3.712	3.669	3.628
7		7.263	7.262	7.261	7.262	7.264	7.267	7.271	7.276	7.284	7.303

Tabla IV. 5. Resultados Método MOD (desplazamiento promedio) líneas sintéticas

La representación gráfica de los resultados se presenta en las Figura IV. 16, Figura IV. 17, Figura IV. 18, Figura IV. 19, Figura IV. 20, Figura IV. 21, Figura IV. 22.

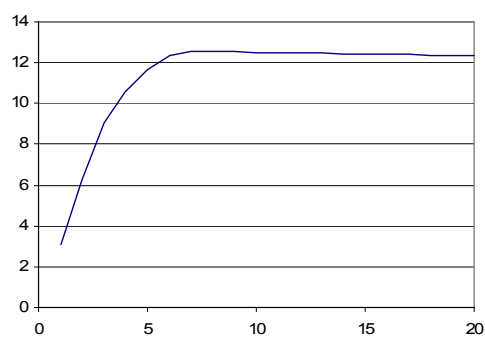


Figura IV. 16. Método MOD(DP). Línea sintética nº1

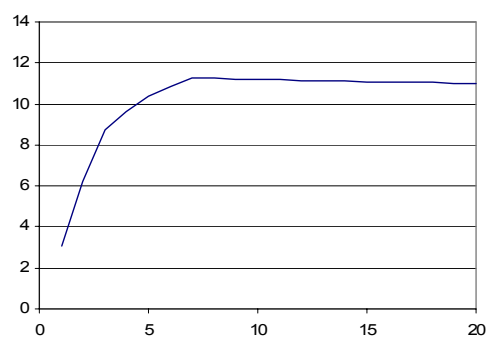


Figura IV. 17. Método MOD(DP). Línea sintética nº2

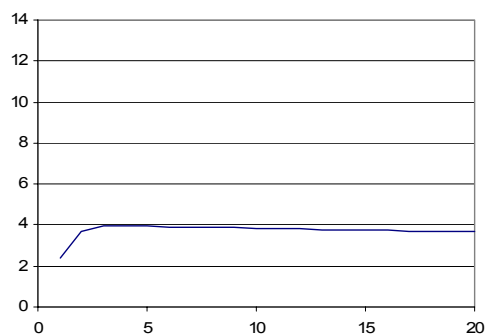


Figura IV. 18. Método MOD(DP). Línea sintética nº3

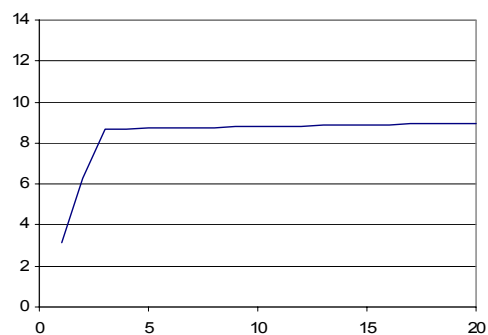


Figura IV. 19. Método MOD(DP). Línea sintética nº4

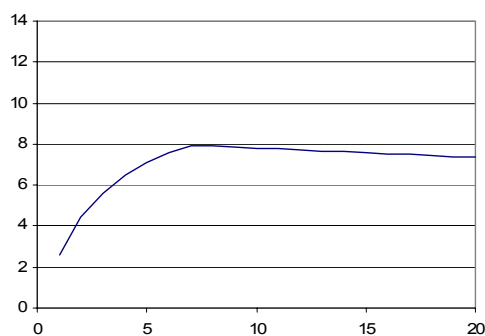


Figura IV. 20. Método MOD(DP). Línea sintética nº5

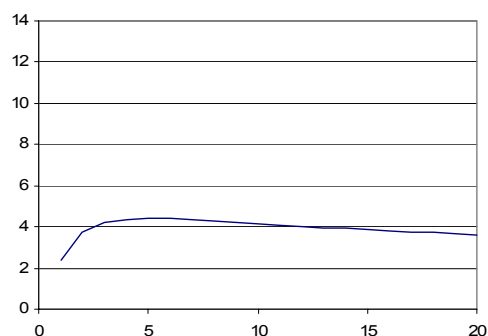


Figura IV. 21. Método MOD(DP). Línea sintética nº6

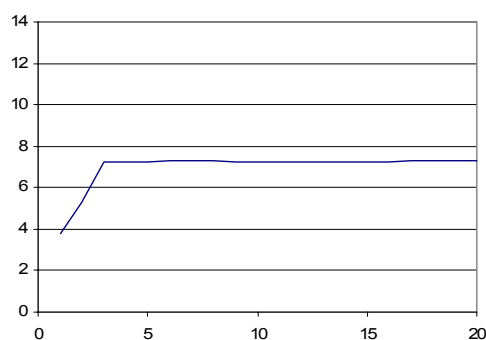


Figura IV. 22. Método MOD(DP). Línea sintética nº7

Los resultados muestran curvas ascendentes a medida que se aumenta la distancia de orlado hasta alcanzar un valor de desplazamiento estabilizado con tendencia a descender en los casos donde la superficie del orlado a controlar se reduce a consecuencia de los quiebros de la propia línea con respecto a la superficie de un orlado producido por una línea sin quiebros de la misma longitud (líneas 4 y 7).

El desplazamiento promedio, como su propia denominación indica, es interpretado como la diferencia promediada entre las dos líneas. Los resultados muestran en todos los casos una curva que alcanza un valor para descender suavemente a continuación. El valor máximo obtenido de este desplazamiento puede ser considerado como el desplazamiento promedio de la línea estudiada y presenta mayores valores en las líneas con desplazamiento sistemático simulado (línea 1, línea 2 y línea 4). El suave descenso de la curva debe interpretarse como el producido por las variaciones ocurridas por el cálculo de las superficies de los extremos de los orlados al tratarse de semicircunferencias. En cuanto a la línea 7, con errores simulados en los puntos inicial y final presenta una curva con un comportamiento ascendente hasta alcanzar el valor de desplazamiento.

IV.3.4.3 Oscilación

Los resultados correspondientes a la oscilación se presentan en la Tabla IV. 6.

Línea	Imagen	Oscilación para cada anchura de orlado [núm. áreas/m]									
		1 m	2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m
1		0.01	0.01	0.015	0.015	0.015	0.015	0.015	0.01	0.01	0.01
2		0.01	0.01	0.015	0.015	0.015	0.015	0.015	0.01	0.01	0.01
3		0.035	0.035	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
4		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
5		0.025	0.025	0.015	0.015	0.015	0.015	0.015	0.01	0.01	0.01
6		0.035	0.035	0.015	0.015	0.015	0.015	0.01	0.01	0.01	0.01

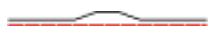
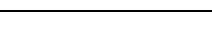
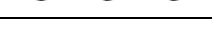
7		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Línea	Imagen	Oscilación para cada anchura de orlado [núm. áreas/m]									
		11 m	12 m	13 m	14 m	15 m	16 m	17 m	18 m	19 m	20 m
1		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
2		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
3		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
4		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
5		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
6		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
7		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Tabla IV. 6. Resultados Método MOD (oscilación) líneas sintéticas

Gráficamente, estos valores muestran las curvas de la Figura IV. 23, Figura IV. 24, Figura IV. 25, Figura IV. 26, Figura IV. 27, Figura IV. 28 y Figura IV. 29.

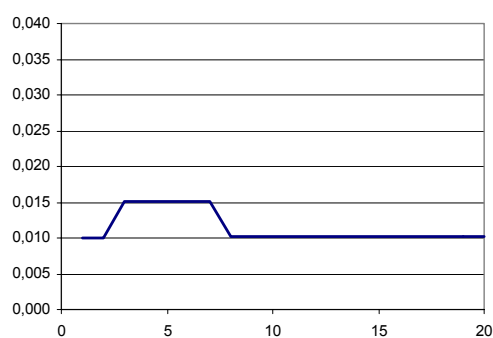


Figura IV. 23. Método MOD(OS). Línea sintética nº1

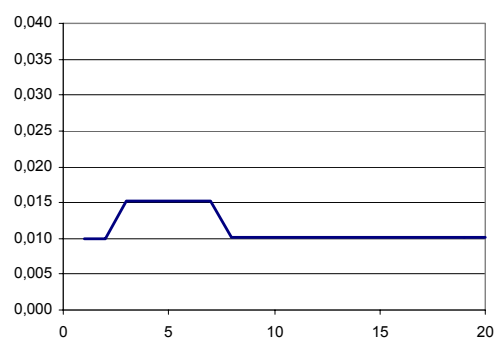


Figura IV. 24. Método MOD(OS). Línea sintética nº2

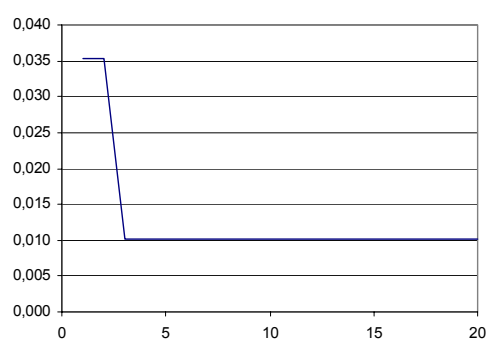


Figura IV. 25. Método MOD(OS). Línea sintética nº3

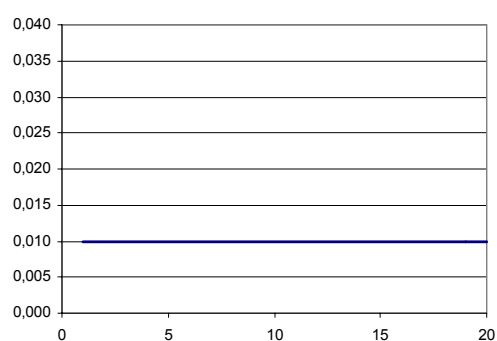


Figura IV. 26. Método MOD(OS). Línea sintética nº4

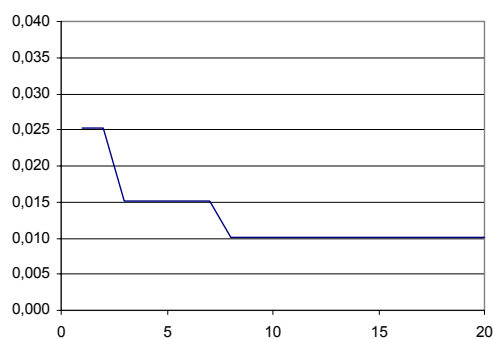


Figura IV. 27. Método MOD(OS). Línea sintética nº5

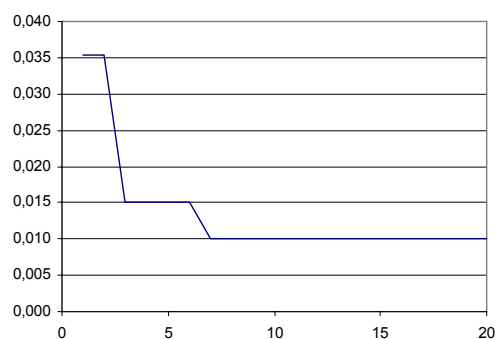


Figura IV. 28. Método MOD(OS). Línea sintética nº6

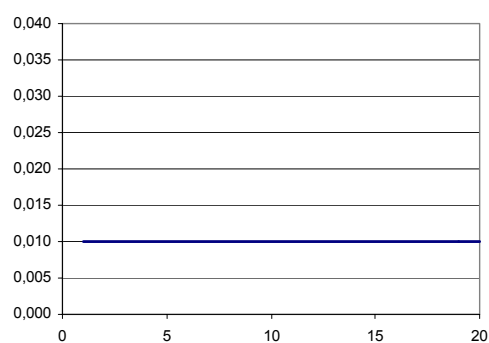


Figura IV. 29. Método MOD(OS). Línea sintética nº7

Los resultados presentan curvas con valores ascendentes hasta que intersectan los orlados para descender después y estabilizarse en 0.01 áreas/m cuando se llega a cierta anchura de orlado en las líneas 1, y 2. En las líneas 4 y 7 el valor se mantiene estable en el mínimo 0.01 ya que no existe oscilación, y en las líneas 3, 5 y 6 se presentan comportamientos descendentes hasta estabilizarse en ese mismo valor mínimo.

La oscilación puede interpretarse como una medida del sesgo del desplazamiento entre las líneas. Presenta unas gráficas con curvas escalonadas con tendencia descendente en las líneas sin desplazamiento sistemático y ascendente a valores pequeños de orlado en las líneas con dicho desplazamiento, tendiendo en todos los casos a un valor estable cuando se aumenta la anchura de orlado.

IV.3.5. Interpretación y análisis

Los resultados obtenidos tras la implementación de la metodología propuesta de control posicional por elementos lineales a las líneas sintéticas muestran un comportamiento diferente según el método utilizado. La aplicación de cada una de las metodologías aporta una importante información de puede ser complementaria sobre el comportamiento de las líneas.

En conclusión, el estudio de estas líneas sintéticas permite adquirir una idea de cómo se comporta cada método con una línea a la que se le ha incorporado un error conocido. En casos reales, las líneas, a priori, van a poseer numerosos desplazamientos, cada uno debido a un error en concreto.

Resulta difícil, identificar el tipo de error existente en casos reales, aunque el estudio presentado en este punto es de utilidad para conocer cómo se comporta cada método y adquirir una visión global de cada uno. Los resultados aportados de los distintos métodos de control de elementos lineales deben estudiarse de forma combinada.

IV.4. Aplicación a bases de datos cartográficas. Caracterización de elementos lineales

En el punto anterior se ha presentado la aplicación de la metodología propuesta de control sobre elementos lineales a líneas sintéticas. Es evidente que un estudio de estas características exige una aplicación a bases de datos reales para ensayar la metodología y comprobar su comportamiento sobre líneas cartográficas.

Como ya fue mencionado en el apartado descriptivo del material en el capítulo anterior, la aplicación no se ha limitado a una sola base de datos y a un escaso espacio geográfico, ya que se ha procurado abarcar territorios de diversas características tanto orográficas, como de altitud media y tipología de vías de comunicación. Se trata, por tanto, de una importante prueba práctica, tanto por el número de elementos implicados, como por el territorio abarcado y las bases cartográficas controladas (Mapa Topográfico de Andalucía a 1:10.000 (MTA) y Base Cartográfica Numérica a 1:25.000 (BCN). Por otro lado, cabe recordar que el control propuesto se realiza a partir de observaciones cinemáticas con GPS.

Con las bases de datos cartográficas y de control presentadas en el apartado de material, se procede a realizar la fase de edición, comparación y depuración de las mismas. A lo largo de este proceso se realiza la edición de los elementos, el enriquecimiento de las bases de datos con información sobre las líneas, la identificación de los elementos homólogos dentro de las bases cartográficas y la de control, el filtrado de las líneas y la selección de las mismas. Los resultados estadísticos tras el proceso de edición y depuración se presentan en la Tabla IV. 7 y la Tabla IV. 8.

Tipo	Red viaria de carreteras de Andalucía ⁸		Red viaria de la zona de trabajo		Líneas tras proceso de edición MTA10		Líneas tras proceso de edición BCN25	
	Longitud [m]	%	Longitud [m]	%	Longitud [m]	%	Longitud [m]	%
Autovías	2294000	9.34	543987	16.58	188175	15.50	203359	16.03
Nacionales	1851000	7.54	154871	4.72	45395	3.74	46332	3.65
Autonómicas	9829000	40.04	640766	19.53	238126	19.62	237186	18.70
Locales	10574000	43.07	1940948	59.16	742100	61.14	781595	61.62
Total	24548000		3280571		1213796		1268473	

Tabla IV. 7. Resultados del proceso de edición y depuración por tipo

Los resultados presentados muestran la evolución producida tras el proceso de edición y depuración. En la Tabla IV. 7, se compara la longitud por tipo de vía de los elementos de la red viaria de Andalucía, de la zona de trabajo utilizada y de las bases de datos depuradas. Las líneas predominantes tras este proceso son las locales, mientras que las nacionales tienen menor presencia.

	Almería [m]	Cádiz [m]	Co-Ja [m]	Huelva [m]	Málaga [m]	Sevilla [m]	Total [m]
Red viaria de la zona de trabajo	393466	508837	1107145	413775	480340	377008	3280571
Líneas tras proceso de edición MTA10	197841	137073	352864	201090	170948	153978	1213796
	50.28%	26.94%	31.87%	48.60%	35.59%	40.84%	37.00%
Líneas tras proceso de edición BCN25	239522	148209	356842	194740	173433	155725	1268473
	60.87%	29.13%	32.23%	47.06%	36.11%	41.31%	38.67%

Tabla IV. 8. Resultados del proceso de edición por zona

La Tabla IV. 8 muestra la evolución tras este proceso por zonas. Se observa un mayor mantenimiento de elementos en Almería, mientras que en Cádiz el número de

⁸ Fuente: Dirección General de Carreteras. Ministerio de Fomento. 2004.

eliminaciones es superior, debido en gran medida, a las dificultades encontradas para la observación GPS debido a la existencia de obras.

En definitiva, se produce una reducción de la base de datos tras el proceso de edición y depuración en torno a un 37% para el MTA10 y un 38% en la BCN25. Las líneas obtenidas se muestran en la Figura IV. 30 para el MTA y en la Figura IV. 31 para la BCN.

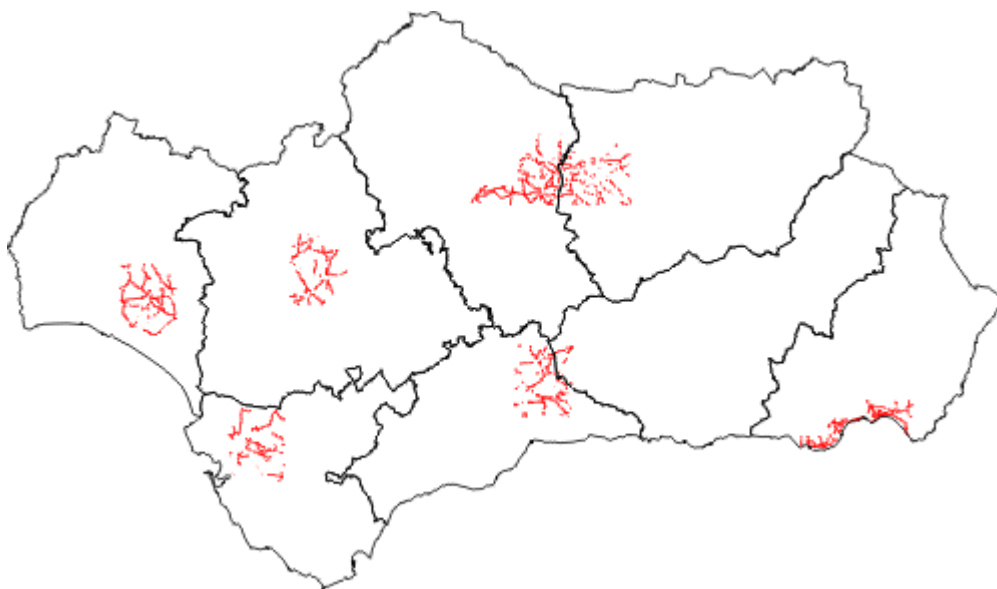


Figura IV. 30. Líneas obtenidas tras el proceso de edición en el MTA.

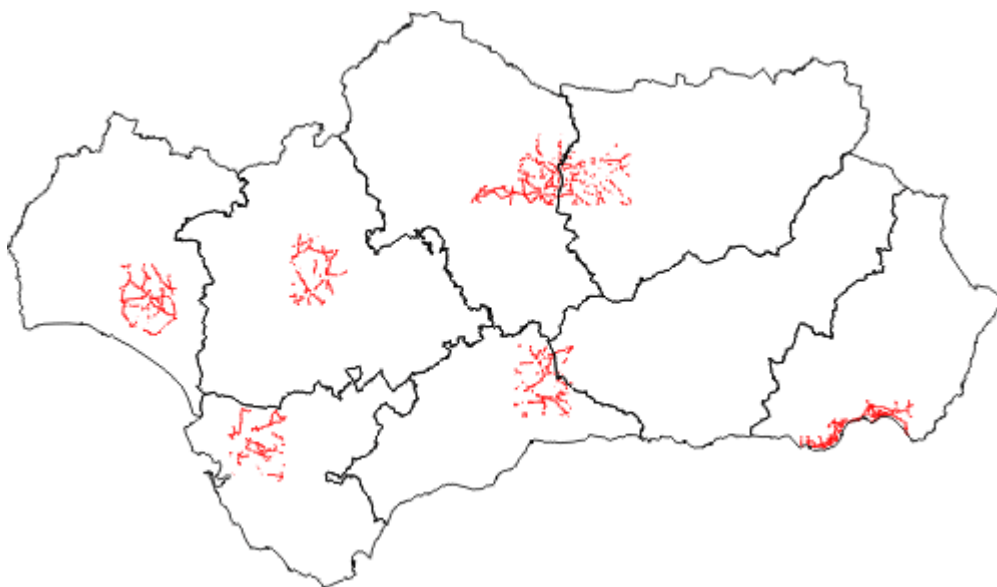


Figura IV. 31. Líneas obtenidas tras el proceso de edición en la BCN

Los rendimientos en la adquisición de elementos lineales con observaciones GPS son analizados por Ruiz y col. (2006a, 2006b y 2007).

Una vez realizada la fase previa de edición de las bases de datos y antes de proceder a presentar los resultados obtenidos con el control automático se procede a estudiar la población en función de diversos criterios como son la longitud del tramo, la zona de trabajo, el tipo de vía, etc.

También presenta interés la caracterización de las dos bases de datos cartográficas en lo relativo a su comportamiento posicional respecto a las líneas GPS. Considerando la evolución de la distancia euclídea de cada punto de una línea, perteneciente a una de las dos bases a controlar, frente a las líneas GPS tomadas como referencia se han obtenido los variogramas relativos a la diferencia de valores entre parejas de puntos de una misma línea, tal como se representan en la Figura IV.32 y en la Figura IV.33 para distancias de hasta 2000m. En el caso del MTA se observa una evolución casi lineal en el origen hasta alcanzar una zona de tendencia o deriva suave y creciente, que se puede considerar como una meseta. Para la BCN la evolución es más compleja y, además, presenta mayor variabilidad. Lo más característico de ésta es la existencia de un pico que rompe la tendencia creciente y genera lo que se conoce como un efecto pozo. Se aprecian otros picos de menor importancia dentro de la gráfica correspondiente a la BCN y todo ello hace pensar en cierto comportamiento periódico.

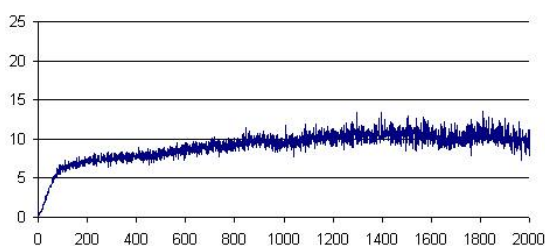


Figura IV.32. Variogramas correspondientes a las diferencias de distancias entre MTA y GPS

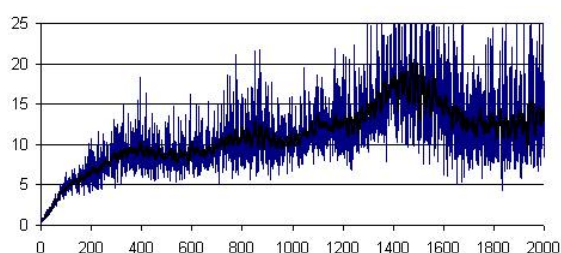


Figura IV.33. Variogramas correspondientes a las diferencias de distancias entre BCN y GPS

Por otro lado, si se analizan los valores de cada punto con los vecinos a una distancia determinada se pueden obtener los correlogramas representados en la Figura IV.34 y la Figura IV.35. Estos correlogramas indican la existencia de un nivel de correlación alto hasta 2 ó 3 saltos a lo largo de la línea, para distancias de 10 saltos los coeficientes descienden a valores de 0.4. Como se observa en las gráficas, las dos bases de datos a controlar presentan un comportamiento bastante similar si bien la BCN presenta mayores correlaciones lo que equivale a mayor suavidad.

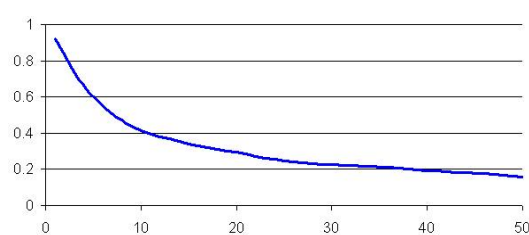


Figura IV.34. Correlogramas correspondientes a los valores de distancias entre puntos del MTA y GPS

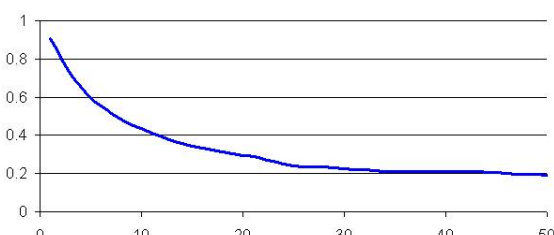


Figura IV.35. Correlogramas correspondientes a los valores de distancias entre puntos de la BCN y GPS

IV.4.1. Longitud del tramo

La longitud del tramo es una característica que viene condicionada por la propia estructura de la red viaria y por el proceso GPS de captura, ya que en cada pérdida de señal del GPS, genera la rotura de los elementos. El número de elementos está en el orden de 1250, lo que supone un tamaño medio de 965m y 995m para el MTA y la BCN, respectivamente.

La Tabla IV. 9 presenta los datos estadísticos básicos para ambas bases de datos.

Cartografía	Casos [n]	Media [m]	Mínimo [m]	Máximo [m]	Suma [km]	Desviación [m]
MTA	1254	965	21	20180	1210	1671.483
BCN	1274	995	18	22700	1267	1805.077

Tabla IV. 9. Distribución relativa de elementos por longitud

La Figura IV. 36 y la Figura IV. 37 presenta los histogramas relativos al número de elementos según su tamaño tanto en el MTA como en la BCN respectivamente. Como puede observarse, se trata de una distribución en donde las mayores frecuencias de ocurrencia se dan en los valores más bajos. El pico más alto (19%) se corresponde con las longitudes en el intervalo [100; 200]m. En total, los tramos con tamaños inferiores a 500 m suman un 58% de los casos debido al troceamiento ya comentado.

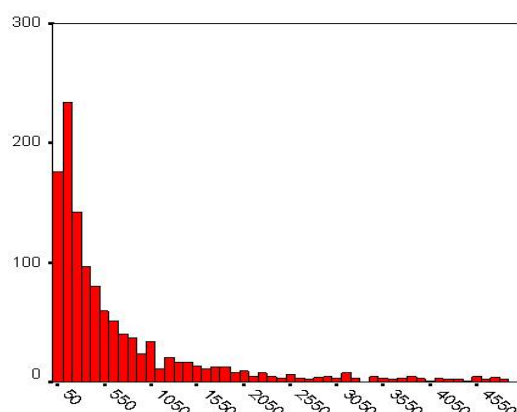


Figura IV. 36. Histogramas relativos a la distribución de casos por longitud. MTA

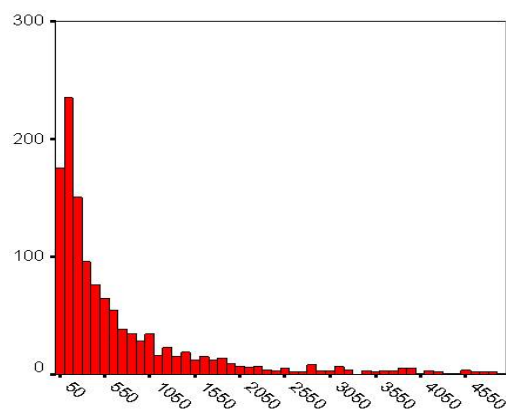


Figura IV. 37. Histogramas relativos a la distribución de casos por longitud. BCN

IV.4.2. Zona de Trabajo

La distribución de las muestras por zonas se presenta en la Figura IV. 38 y la Figura IV. 39. Se observa que Cádiz presenta el mínimo, lo cual se debe tanto a la densidad de su red vial como al hecho de las numerosas obras que impidieron tomar algunos tramos importantes (≈ 50 kms de obras). La zona de Córdoba y Jaén destaca claramente, pero aquí debe tenerse en cuenta que esta zona es bastante más extensa que el resto de los bloques (en el orden de 2.5 a 3.5 veces).

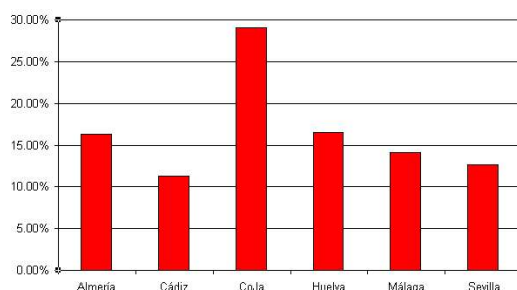


Figura IV. 38. Histogramas relativos a la distribución de longitudes por zona. MTA

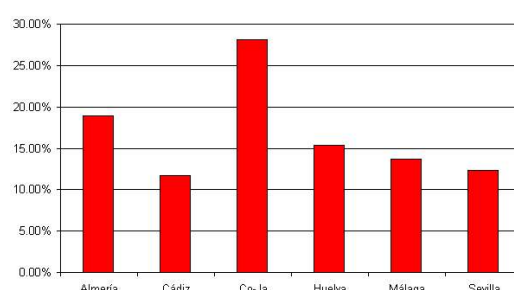


Figura IV. 39. Histogramas relativos a la distribución de longitudes por zona. BCN

La Tabla IV. 10 y la Tabla IV. 11 presentan esa misma distribución en forma numérica. Junto a los valores, en porcentaje, respecto al total de casos y kilómetros, se incluye el número de casos y de kilómetros por zona de trabajo.

Zona	Casos [n]	Media [m]	Suma [km]	% Suma	% Casos
Almería	168	1175	197	16.3	13.4
Cádiz	113	1212	137	11.3	9.0
Co-Ja	505	697	352	29.1	40.3
Huelva	119	1679	200	16.5	9.5
Málaga	254	673	171	14.1	20.3
Sevilla	95	1606	153	12.6	7.6
Total	1254	965	1210	100.0	100.0

Tabla IV. 10. Distribución de los elementos por zonas. MTA

Zona	Casos [n]	Media [m]	Suma [km]	% Suma	% Casos
Almería	175	1367	239	18.9	13.7
Cádiz	117	1262	147	11.7	9.2
Co-Ja	505	706	357	28.1	39.6
Huelva	118	1650	195	15.4	9.3
Málaga	262	662	173	13.7	20.6
Sevilla	97	1604	156	12.3	7.6
Total	1274	995	1267	100.0	100.0

Tabla IV. 11. Distribución de los elementos por zonas. BCN

IV.4.3. Tipo de vía

El tipo de vía se refiere a la categoría administrativa asignada a la vía. La Figura IV. 40 y la Figura IV. 41 muestran que existe una gran abundancia en la categoría local ($\approx 70\%$), y que la categoría de carreteras nacionales queda muy reducida. Los datos se ajustan, sin embargo, al reparto existente en la zona de trabajo (Tabla IV. 7), aunque presenta una reducción de las vías nacionales y un aumento de las locales con respecto a la red viaria global andaluza. En la Tabla IV. 12 y en la Tabla IV. 13 se presentan en forma numérica los valores más importantes.

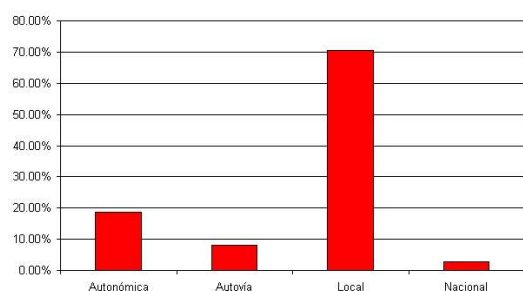


Figura IV. 40. Histogramas relativos a la distribución de longitudes por tipo de vía. MTA

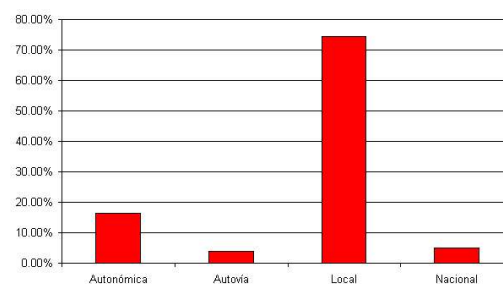


Figura IV. 41. Histogramas relativos a la distribución de longitudes por tipo de vía. BCN

Tipología	Casos [n]	Media [m]	Suma [km]	% Suma	% Casos
Autonómica	249	915	228	18.8	19.9
Autovía	82	1183	97	8.0	6.5
Local	882	966	852	70.4	70.3
Nacional	41	809	33	2.7	3.3
Total	1254	965	1210	100.0	100.0

Tabla IV. 12. Distribución de los elementos por tipo de vía. MTA

Tipología	Casos [n]	Media [m]	Suma [km]	% Suma	% Casos
Autonómica	255	822	207	16.5	20.0
Autovía	63	808	51	4.0	4.9
Local	915	1032	944	74.5	71.8
Nacional	41	1525	63	4.9	3.2
Total	1274	995	1267	100.0	100.0

Tabla IV. 13. Distribución de los elementos por tipo de vía. BCN

IV.4.4. Orientación

La orientación se refiere al ángulo con respecto al norte de la alineación de los puntos inicial y final del elemento lineal. El comportamiento respecto a las orientaciones es casi homogéneo, como se puede apreciar en los histogramas de la Figura IV. 42 y la Figura IV. 43. Esto indica una buena distribución respecto a este factor lo que implica la existencia de aleatoriedad en la muestra. La Tabla IV. 14 presentan los valores numéricos de la distribución según este aspecto.

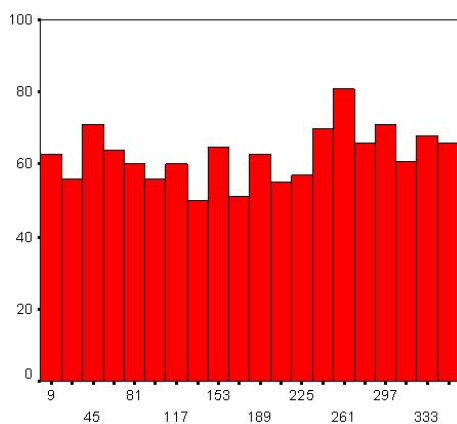


Figura IV. 42. Histograma relativo a la distribución según la orientación. MTA

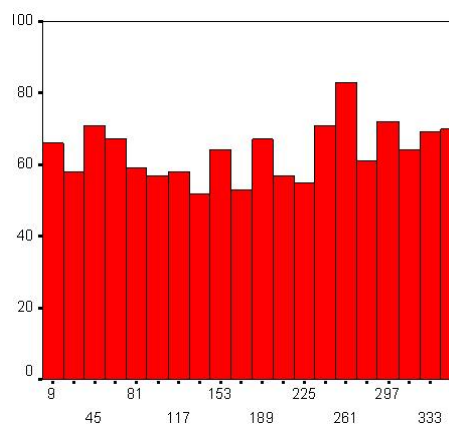


Figura IV. 43. Histograma relativo a la distribución según la orientación. BCN

Orientación	Casos [n]	Media [m]	Suma [km]	% Suma	% Casos
337°-22°	164	841	138	11.4	13.1
22°-67°	155	935	145	12.0	12.4
67°-112°	148	1269	188	15.5	11.8
112°-157°	148	790	117	9.7	11.8
157°-202°	141	1000	141	11.7	11.2
202°-247°	146	829	121	10.0	11.6
247°-292°	184	960	177	14.6	14.7
292°-337°	168	1092	183	15.2	13.4
Total	1254	965	1210	100.0	100.0

Tabla IV. 14. Distribución de los elementos por orientación. MTA

Orientación	Casos [n]	Media [m]	Suma [km]	% Suma	% Casos
337°-22°	168	853	143	11.3	13.2
22°-67°	160	1095	175	13.8	12.6
67°-112°	151	1153	174	13.7	11.9
112°-157°	148	866	128	10.1	11.6
157°-202°	144	1021	147	11.6	11.3

202°-247°	150	917	138	10.9	11.8
247°-292°	179	996	178	14.1	14.1
292°-337°	174	1055	184	14.5	13.7
Total	1274	995	1267	100.0	100.0

Tabla IV. 15. Distribución de los elementos por orientación. BCN

La Figura IV. 44 y la Figura IV. 45 representan la distribución de la orientación por zonas. El valor porcentaje teórico para cada caso corresponde a un 12.5%. En el MTA se puede apreciar zonas con orientaciones mejor distribuidas como Cádiz, Córdoba-Jaén y Huelva con máximos en torno al 15%, y otras donde destaca una orientación sobre el resto, como es el caso de Málaga y Sevilla en el intervalo 202°-247°, y Almería en el intervalo 247°-292° que llegan a tener valores en torno al 20%. Para el caso de la BCN, el comportamiento es semejante, aunque en el caso de Sevilla también aparece destacado el intervalo 67°-112°, con una distribución cercana al 20%.

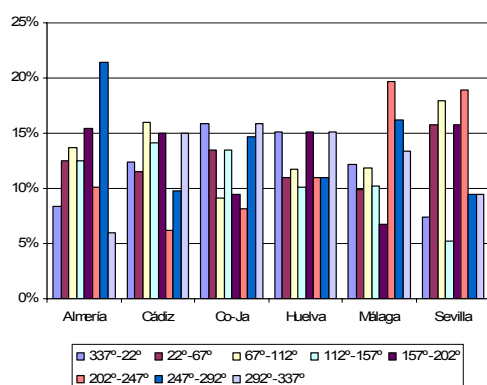


Figura IV. 44. Histograma relativo a la distribución según la orientación por zonas. MTA

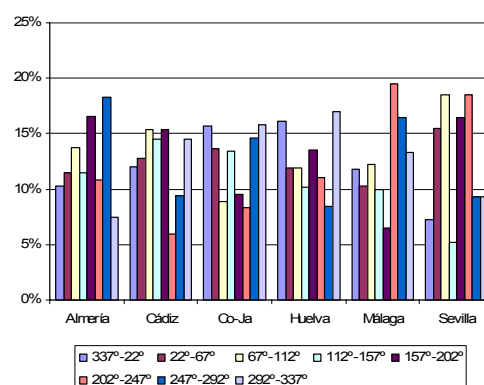


Figura IV. 45. Histograma relativo a la distribución según la orientación por zonas. BCN

IV.4.5. Sinuosidad

Este criterio presenta una vinculación directa con la topografía y la categoría de la vía. Cuanto menor categoría tenga una vía y tanto más accidentado sea un terreno, mayor sinuosidad cabrá esperar en la misma. Por su parte, las vías principales aunque discurren por terrenos accidentados serán menos sinuosas que las vías de menor rango.

Para clasificar los elementos según este criterio se utiliza la metodología propuesta por García (2006) considerando las siguientes categorías: muy suave, suave, sinuosa. La clasificación automática de las vías de comunicación en estas categorías se ha realizado merced a la red neuronal artificial desarrollada en el trabajo de García (2006).

La Figura IV. 46 y la Figura IV. 47 presentan la distribución de casos y kilómetros de longitud entre las distintas categorías de sinuosidad consideradas para las dos bases de datos. De igual forma se presentan los valores numéricos en la Tabla IV. 16 y en la Tabla IV. 17.

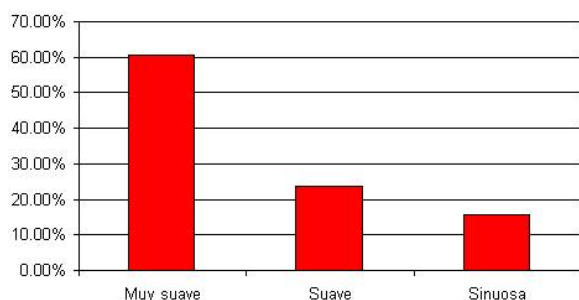


Figura IV. 46. Histograma relativo a la distribución de longitudes según sinuosidad. MTA

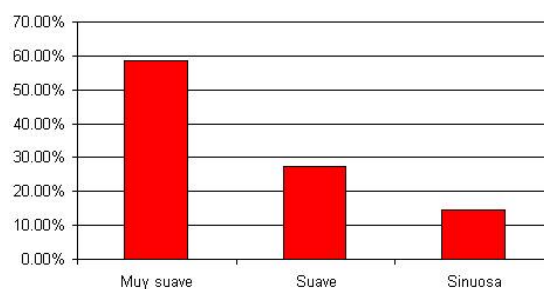


Figura IV. 47. Histograma relativo a la distribución de longitudes según sinuosidad. BCN

Categoría	Casos [n]	Media [m]	Suma [km]	% Suma	% Casos
Muy suave	673	1091	734.6	60.7	53.7
Suave	289	989	286.0	23.6	23.0
Sinuosa	292	648	189.3	15.6	23.3
Total	1254	965	1210	100.0	100.0

Tabla IV. 16. Distribución de los elementos por sinuosidad. MTA

Categoría	Casos [n]	Media [m]	Suma [km]	% Suma	% Casos
Muy suave	683	1083	739.5	58.4	53.6
Suave	292	1183	345.4	27.3	22.9
Sinuosa	299	609	182.2	14.4	23.5
Total	1274	995	1267	100.0	100.0

Tabla IV. 17. Distribución de los elementos por sinuosidad. BCN

Los datos aportados indican un dominio de la categoría muy suave dentro de los elementos de las bases de datos. La justificación a este hecho, evidentemente, procede de las zonas de trabajo utilizadas, ya que, como se indica en el apartado de presentación del material, predominan las zonas llanas donde la suavidad de la red viaria es característica.

IV.4.6. Interpretación y análisis

La caracterización de los elementos lineales utilizados realizada a lo largo de este apartado permite conocer el conjunto de datos que se utiliza en la aplicación práctica de la metodología propuesta para el control posicional. Este conjunto de datos procede de la fase inicial de edición y depuración de las bases cartográficas y de control, por lo que se ve afectado por las posibles eliminaciones o modificaciones realizadas durante esta fase. También se ha cuantificado, en este apartado, la reducción producida en las bases de datos tras esta fase de edición.

A lo largo del punto que nos ocupa, se han presentado porcentajes, longitudes y otras medidas acerca de las distribuciones de elementos lineales según una serie de criterios. La consecuencia fundamental de la toma de datos con GPS para la confección de la base de datos y más concretamente la pérdida de señal durante la observación, es el troceamiento de la base de datos que provoca que la mayor parte de los elementos utilizados se encuentren en el intervalo de longitudes inferior a 1 km. El estudio de las orientaciones revela una gran homogeneidad de los datos. También se ha presentado la distribución por zonas y según el tipo de las vías de comunicación utilizadas. Otro aspecto a tener en cuenta es la sinuosidad, en la que se ha mostrado una clasificación mayoritaria de elementos muy suaves.

Lógicamente, en el estudio de la caracterización de estos elementos no se han encontrado grandes diferencias entre las dos bases de datos cartográficas, ya que ambas están compuestas de conjuntos de elementos similares.

IV.5. Resultados obtenidos con la aplicación a bases de datos cartográficas

A lo largo de este apartado se presentan los resultados obtenidos al aplicar la metodología de control propuesto a las bases de datos cartográficas oficiales previamente descritas. Como ha sido comentado con anterioridad, la metodología está basada en la aplicación automática de los métodos MDH, MSK, MOS y MOD.

Los resultados se plantean a continuación de forma individualizada para estos métodos, aunque el análisis de los mismos debe realizarse de una forma global.

Estos resultados se presentan de forma resumida en el Anexo IV y en su totalidad en el cd-rom que acompaña al presente documento.

IV.5.1. Resultados del Método de las Distancias de Hausdorff (MDH)

Los resultados del método MDH se resumen en la Tabla IV. 18 y la Tabla IV. 19. En estas tablas se muestran por un lado los valores máximos, mínimos y medios de la distancia de Hausdorff para cada zona y el global así como la media de las distancias medias por elemento.

En el caso del MTA el valor medio determinado para la distancia de Hausdorff es de 10.82 m, comportándose mejor en la zona de Almería (8.79 m) y peor en la de Sevilla (13.22 m). Si se tiene en cuenta la media de las distancias de cada elemento el valor medio es de 4.44 m, siendo el mejor comportamiento el de Córdoba-Jaén (3.79 m) y el peor Sevilla (5.55 m).

Los valores máximos se muestran en el intervalo de 16.75 m a 41.72 m.

Zona	Longitud [m]	Media Hausdorff [m]	Mínimo Hausdorff [m]	Máximo Hausdorff [m]	Desviación Hausdorff [m]	Media distancias [m]	Desviación Dist. Medias [m]
Almería	196929	8.79	0.53	20.68	4.081	3.92	2.344
Cádiz	330712	9.38	0.81	16.75	3.261	3.96	1.633
Co-Ja	136978	9.54	0.40	31.44	4.794	3.79	1.952
Huelva	193703	13.19	1.57	41.72	8.156	4.46	2.116
Málaga	171016	12.06	1.24	25.79	5.309	5.69	3.016
Sevilla	147857	13.22	1.31	25.59	4.466	5.55	2.59
Total	1177196	10.82	0.40	41.72	5.594	4.44	2.403

Tabla IV. 18. Resultados MDH para el MTA

En el caso de la BCN, se muestra para la distancia de Hausdorff un valor medio global de 11.42 m, obteniendo los valores extremos en Huelva (8.17 m) y Sevilla (18.12 m). En cuanto al valor promedio de las distancias medias de cada elemento lineal, se obtiene un valor de 3.36 m en el total, siendo Huelva el valor más bajo (2.54 m) y Málaga el más alto (4.36 m).

Los valores máximos de la distancia de Hausdorff oscilan entre 34.29 m y 80.57 m.

Zona	Longitud [m]	Media Hausdorff [m]	Mínimo Hausdorff [m]	Máximo Hausdorff [m]	Desviación Hausdorff [m]	Media distancias [m]	Desviación Dist. Medias [m]
Almería	237507	9.75	0.76	34.71	7.485	2.89	2.372
Cádiz	349341	11.93	0.55	74.81	16.116	3.18	2.53
Co-Ja	132623	11.64	0.18	51.14	9.224	3.52	2.354
Huelva	184720	8.17	0.80	46.86	5.612	2.54	1.744
Málaga	173315	10.31	1.06	34.29	8.447	4.36	4.265
Sevilla	155448	18.12	0.61	80.57	22.852	3.75	2.756
Total	1232953	11.42	0.18	80.57	12.305	3.36	2.663

Tabla IV. 19. Resultados MDH para la BCN

El comportamiento de las bases de datos controladas es distinto. El MTA posee valores medios mejores que la BCN al igual que los valores máximos. Sin embargo, los valores de las distancias medias por elemento dan valores mucho mejores para la BCN que para el MTA.

Si se observan las funciones de distribución para los máximos de las distancias de Hausdorff (Figura IV. 48 y Figura IV. 49) y para las distancias medias por elemento (Figura IV. 50 y Figura IV. 51), se observa que las segundas tienen un comportamiento mejor que las primeras. Por otro lado, llama la atención el comportamiento de las curvas de máximos entre MTA y BCN, donde la BCN presenta una curva mejor inicialmente. Sin embargo, para niveles de confianza elevados la BCN se comporta peor.

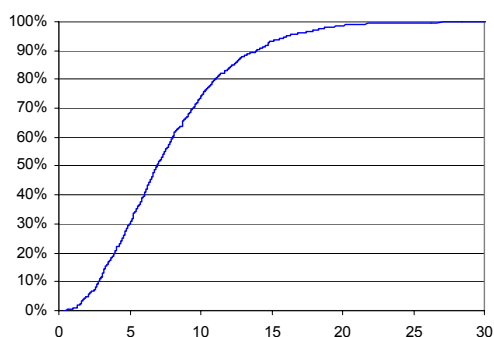


Figura IV. 48. Función de distribución MDH para distancias de Hausdorff. MTA

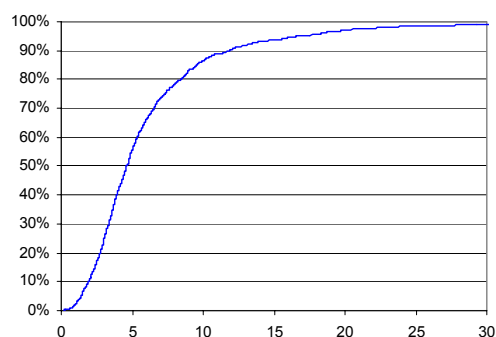


Figura IV. 49. Función de distribución MDH para distancias de Hausdorff. BCN

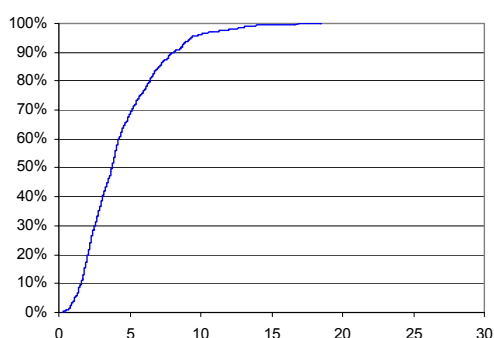


Figura IV. 50. Función de distribución MDH para distancias medias de los elementos. MTA

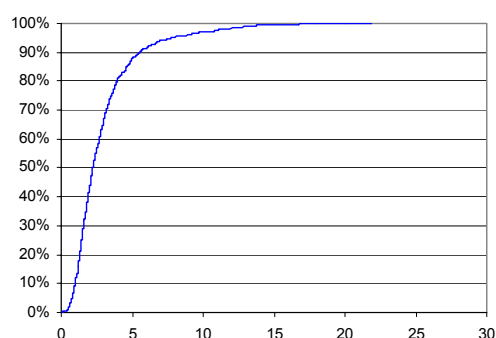


Figura IV. 51. Función de distribución MDH para distancias medias de los elementos. BCN

En la Tabla IV. 20 se muestran los resultados para distintos niveles de confianza, mostrando para un 95% de confianza un valor de 19.99 m para el MTA y de 27.93 m para la BCN en el caso de las distancias de Hausdorff y de 9.32 m para el MTA y de 7.79 m para la BCN para el caso de las medias por elementos.

Tipo	Base de datos	50% [m]	90% [m]	95% [m]	99% [m]
Distancias de Hausdorff	MTA	10.08	17.71	19.99	37.41
	BCN	7.64	21.75	27.93	80.57
Distancias medias por elemento	MTA	3.72	8.1	9.32	13.39
	BCN	2.22	5.57	7.79	13.47

Tabla IV. 20. Resultados para distintos niveles de confianza MDH

El análisis realizado evidencia la sensibilidad del método MDH frente a la presencia de valores atípicos, los cuales, no tienen porqué deberse necesariamente a problemas posicionales. Así por ejemplo, las grandes discrepancias pueden deberse a falta de actualización de la cartografía, errores propios de la fase de edición de la bases de datos, etc.

Los resultados del estudio de la media de las distancias de cada elemento lineal ponen de manifiesto que los valores máximos pueden ser considerados para estimar errores en elementos de forma individualizada.

IV.5.2. Resultados del Método de la Banda Épsilon o Skidmore (MSK)

Los resultados del método MSK se presentan en la Tabla IV. 21 y la Tabla IV. 22. Para el global de la zona de estudio se obtiene un valor de 4.22 m de media para el MTA y 3.11 para la BCN. Las zonas con mejor comportamiento son Almería (3.65 m) para el MTA y Huelva (2.47 m) para la BCN, mientras que las peores son Málaga (5.51 m y 4.15 m) en ambos casos.

Zona	Longitud [m]	Media [m]	Mínimo [m]	Máximo [m]	Desviación [m]
Almería	197365	3.65	0.27	18.30	2.255
Cádiz	137084	3.74	0.29	13.55	1.66
Co-Ja	352101	3.66	0.11	18.95	1.962
Huelva	199834	4.23	0.83	14.85	2.136
Málaga	171217	5.51	0.01	13.89	3.010
Sevilla	152936	5.21	0.11	18.45	2.501
Total	1210537	4.22	0.01	18.95	2.370

Tabla IV. 21. Resultados MSK para el MTA

Zona	Longitud [m]	Media [m]	Mínimo [m]	Máximo [m]	Desviación [m]
Almería	239311	2.52	0.12	19.70	2.277
Cádiz	148068	3.19	0.32	13.13	2.370
Co-Ja	356605	3.24	0.10	14.98	2.076
Huelva	194754	2.47	0.40	22.38	1.657
Málaga	173315	4.15	0.32	14.72	3.837
Sevilla	155448	3.31	0.22	12.46	2.252
Total	1267501	3.11	0.10	22.38	2.488

Tabla IV. 22. Resultados MSK para el BCN

Numéricamente, se observa un mejor comportamiento de la BCN con respecto al MTA. Las funciones de distribución presentadas en la Figura IV. 52 y en la Figura IV. 53 muestran un comportamiento similar. En este caso, la curva que representa el resultado de la BCN se eleva antes y más enérgicamente.

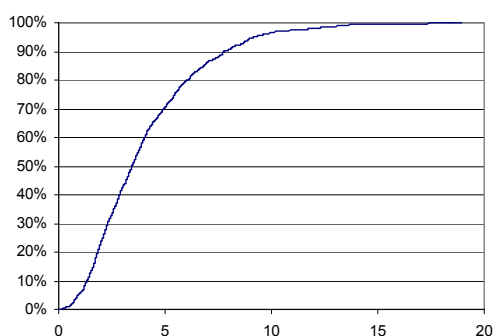


Figura IV. 52. Función de distribución MSK. MTA

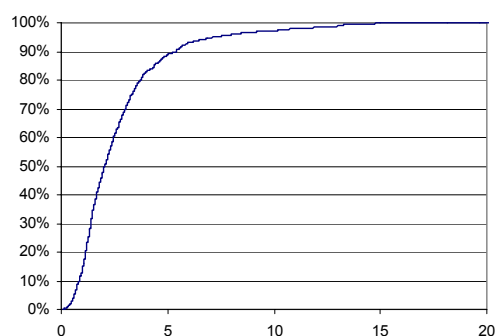


Figura IV. 53. Función de distribución MSK. BCN

En la Tabla IV. 23 se muestran los resultados para distintos niveles de confianza, en los que se confirma lo expuesto anteriormente, presentando para un 95% de confianza un valor de 8.63 m para el MTA y de 8.05 m para la BCN.

Nivel de confianza	50% [m]	90% [m]	95% [m]	99% [m]
MTA	3.72	7.28	8.63	13.62
BCN	2.28	5.78	8.05	13.29

Tabla IV. 23. Resultados para distintos niveles de confianza MSK

Por otro lado, se plantea el estudio del número de cruces que se producen entre las líneas homólogas para estudiar posible existencia de sesgos en los desplazamientos producidos. La Tabla IV. 24 y la Tabla IV. 25 presentan los resultados del número de cruces producidos al aplicar esta metodología a las bases de datos estudiadas.

Zona	Longitud [m]	Número total de cruces	Número de cruces por kilómetro
Almería	197364.552	427	2.16
Cádiz	137084.349	288	2.10
Co-Ja	352100.708	1096	3.11
Huelva	199834.204	480	2.40
Málaga	171217.133	620	3.62
Sevilla	152935.635	260	1.70
Total	1210536.581	3171	2.62

Tabla IV. 24. Resultados del número de cruces. MTA

Zona	Longitud [m]	Número total de cruces	Número de cruces por kilómetro
Almería	239310.571	1177	4.92
Cádiz	148068.121	472	3.19
Co-Ja	356604.888	1473	4.13
Huelva	194754.191	925	4.75
Málaga	173315.215	1224	7.06
Sevilla	155447.534	540	3.47
Total	1267500.519	5811	4.58

Tabla IV. 25. Resultados del número de cruces. BCN

Los resultados del número de cruces ofrecen valores más bajos en el MTA con un valor de 2.62 cruces por kilómetro que en la BCN con 4.58 km, lo que indica que en la BCN existe mayor aleatoriedad.

En definitiva, el método MSK muestra valores medios de los desplazamientos, comportándose de forma menos sensible ante pequeñas variaciones de los datos. En este caso, la base cartográfica BCN se comporta mejor que el MTA. Los resultados obtenidos son similares a los obtenidos mediante la métrica del método MDH con los valores medios de las distancias por elementos. Sin embargo, parece más apropiado utilizar el método MSK, ya que el valor de desplazamiento medio obtenido es producto de toda la longitud de la línea y no de los vértices de la misma.

IV.5.3. Resultados del Método del Orlado Simple (MOS)

Los resultados del método MOS se muestran en forma de función de distribución, donde para una semianchura de orlado se establece el porcentaje de longitud (PL) incluido de la línea cartográfica. Las semianchuras utilizadas en la aplicación de este método son de 1 a 20 m, utilizando un salto de 0.1 m entre 1 y 3 m, de 0.2 m entre 3 y 5 m, de 0.5 m entre 5 y 10 m y de 1 m entre 10 y 20 m. La Tabla IV. 26 y la Tabla IV. 27 muestran los resultados para ciertos niveles de confianza de dichas funciones de distribución.

Zona	50% [m]	90% [m]	95% [m]	99% [m]
Almería	2.9	8	10	16
Cádiz	3.2	8	9	11
Co-Ja	2.9	8.5	10	13
Huelva	3.4	9.5	11	17
Málaga	4.6	12	14	19
Sevilla	4.6	11	14	18
Total	3.4	9	11	16

Tabla IV. 26. Resultados MOS(PL) para el MTA

Zona	50% [m]	90% [m]	95% [m]	99% [m]
Almería	1.4	7	9.5	16
Cádiz	2.2	7	10	17
Co-Ja	2.2	7.5	11	17
Huelva	1.9	5	6.5	13
Málaga	2.4	12	15	20
Sevilla	2.2	7.5	10	19
Total	2	7.5	10	18

Tabla IV. 27. Resultados MOS(PL) para la BCN

La BCN presenta un mejor comportamiento para valores más pequeños del orlado, sin embargo, el MTA se comporta mejor para valores muy altos (99%) del nivel de confianza. De cualquier forma, para un nivel del 95% las dos bases de datos se comportan de forma similar alcanzando un rango de 10 a 11 m.

Para el orlado de 20 m se alcanza un nivel de confianza del 99.73% en el MTA y del 99.37% en la BCN.

De forma análoga, se calcula el porcentaje de inclusión de vértices (PV) del elemento cartográfico sobre los distintos orlados de los elementos de control. La Tabla IV. 28 y la Tabla IV. 29 muestran los resultados para ciertos niveles de confianza de dichas funciones de distribución.

Zona	50% [m]	90% [m]	95% [m]	99% [m]
Almería	3	8.5	11	17
Cádiz	3.4	8	9.5	12
Co-Ja	3	8.5	10	14
Huelva	3.4	9.5	11	17
Málaga	4.8	12	14	19
Sevilla	4.6	12	15	19
Total	3.6	9.5	12	17

Tabla IV. 28. Resultados MOS(PV) para el MTA

Zona	50% [m]	90% [m]	95% [m]	99% [m]
Almería	1.6	8	10	17
Cádiz	2.4	7.5	11	19
Co-Ja	2.2	7.5	11	18
Huelva	1.9	5.5	7	12
Málaga	2.4	11	16	20
Sevilla	2.3	8	12	20
Total	2.1	8	8	20

Tabla IV. 29. Resultados MOS(PV) para la BCN

En la Figura IV. 54 y la Figura IV. 55 se muestran las gráficas con las funciones de distribución del porcentaje de línea dentro de orlado para las dos bases de datos.

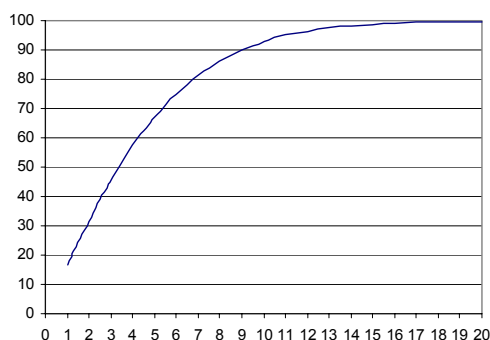


Figura IV. 54. Función de distribución del porcentaje de línea dentro del orlado. MTA

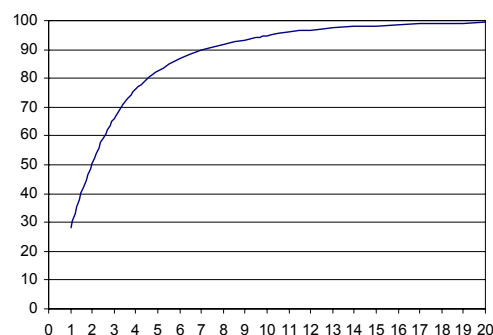


Figura IV. 55. Función de distribución del porcentaje de línea dentro del orlado. BCN

En la Figura IV. 56 y la Figura IV. 57 se muestran las gráficas con las funciones de distribución del porcentaje de vértices dentro de orlado para las dos bases de datos.

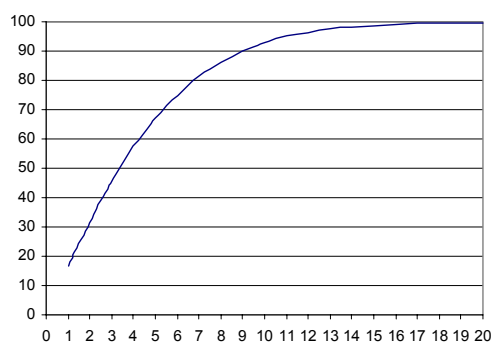


Figura IV. 56. Función de distribución del porcentaje vértices dentro del orlado. MTA

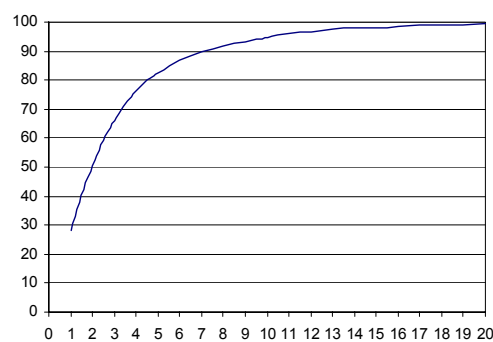


Figura IV. 57. Función de distribución del porcentaje vértices dentro del orlado. BCN

Si se comparan las gráficas anteriores se observa que los resultados obtenidos aplicando el método por porcentaje de longitud del elemento o bien por porcentaje de vértices son similares. Para ello, se utiliza el Test de Kolmogorov-Smirnov, cuyos resultados se muestran en la Tabla IV. 30. Tanto para el MTA como para la BCN se obtienen distancias de Kolmogorov muy pequeñas y un nivel de confianza (p-valor) muy alto para la hipótesis de similitud. Parece lógico por tanto, utilizar el valor original del método basado en contabilizar el porcentaje de longitud dentro del orlado de control.

Porcentaje de Longitud – Porcentaje de vértices	Distancia frecuencial K-S	Z K-S	P-valor
MTA 10	0.019	0.099	1
BCN 25	0.039	0.198	0.999

Tabla IV. 30. Resultados Kolmogorov-Smirnov de resultados de longitud y número de vértices

Las curvas obtenidas han sido chequeadas con el software estadístico SPSS para estudiar el ajuste de las distribuciones obtenidas a algunas distribuciones continuas de probabilidad con el fin de establecer qué comportamiento presentan las funciones obtenidas. Teniendo en cuenta la función de densidad de las curvas resultantes (Figura IV. 58), se ha realizado una selección de las distribuciones más adecuadas para realizar el contraste. Las distribuciones contrastadas son la exponencial y la gamma. Los resultados aportados por el software estadístico en forma de proporciones acumuladas de cada curva respecto a las proporciones acumuladas de las distribuciones de contraste certifican que las funciones de distribución obtenidas con el método MOS no se ajustan a ninguna de ellas (ver gráficos en Anexo III).

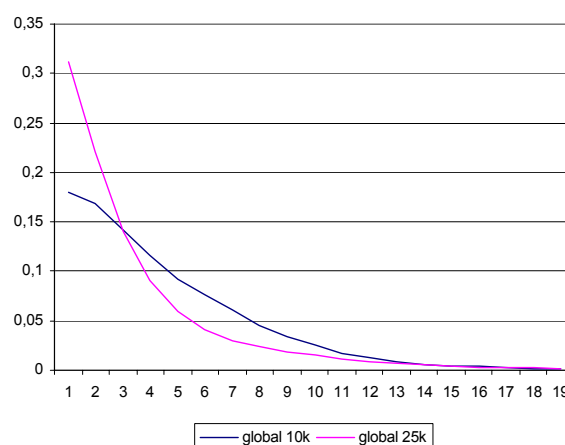


Figura IV. 58. Funciones de densidad de probabilidad de los resultados del método MOS

En definitiva, el método MOS muestra una serie de resultados en forma de funciones de distribución. El comportamiento de dichas funciones es superior en la BCN que en el MTA, aunque éste se comporta mejor para niveles de confianza altos.

IV.5.4. Resultados del Método del Orlado Doble (MOD)

Este método permite extraer de los resultados tres medidas: Superficie común, desplazamiento promedio y oscilación. En la Tabla IV. 31 y la Tabla IV. 32 se presentan valores numéricos resultantes de la aplicación de esta metodología, mostrando las medidas de superficie común y desplazamiento para orlados de 10 y 20 m de semianchura.

Las semianchuras de orlado aplicadas son de 1 m a 20 m con un salto de 1 m.

Zona	Superficie Común		Desplazamiento Promedio [m]	
	10m	20m	10m	20m
Almería	83.27%	91.27%	5.24	5.46
Cádiz	81.43%	90.55%	5.90	5.99
Co-Ja	83.81%	92.07%	4.90	4.97
Huelva	80.53%	90.06%	6.11	6.23
Málaga	74.77%	86.92%	7.96	8.31
Sevilla	74.92%	86.31%	8.28	8.85
Total	80.51%	89.98%	6.18	6.34

Tabla IV. 31. Resultados MOD Superficie común y desplazamiento promedio. MTA

Zona	Superficie Común		Desplazamiento Promedio [m]	
	10m	20m	10m	20m
Almería	94.71%	103.16%	4.72	8.41
Cádiz	87.39%	93.95%	3.72	3.87
Co-Ja	86.14%	91.99%	4.72	5.04
Huelva	91.18%	94.47%	3.0	3.49
Málaga	81.77%	90.45%	5.91	6.12
Sevilla	85.40%	92.64%	4.56	4.58
Total	88.01%	94.60%	4.51	5.38

Tabla IV. 32. Resultados MOD Superficie común y desplazamiento promedio. BCN

Las tres medidas propuestas se expresan en forma de funciones de distribución, que son presentadas a continuación de forma gráfica.

La función de distribución para la medida de la superficie común se muestra en la Figura IV. 59 y la Figura IV. 60. Como se observa la curva del MTA presenta un comportamiento más estable, aunque la curva que describe la función de la BCN obtiene valor superiores a menores semianchos de orlado.

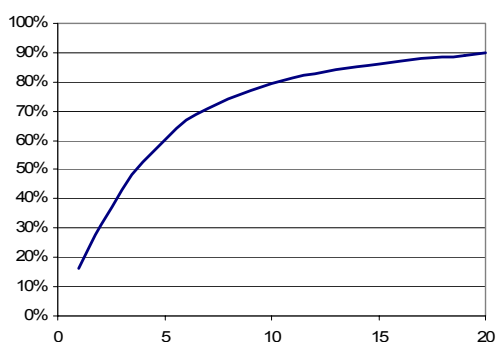


Figura IV. 59. Función de distribución de la superficie común.
MTA

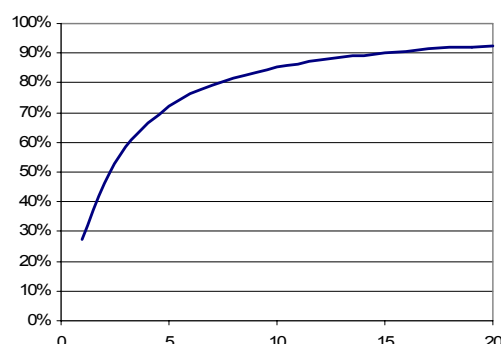


Figura IV. 60. Función de distribución de la superficie común.
BCN

La función de distribución para la medida del desplazamiento promedio se muestra en la Figura IV. 61 y la Figura IV. 62.

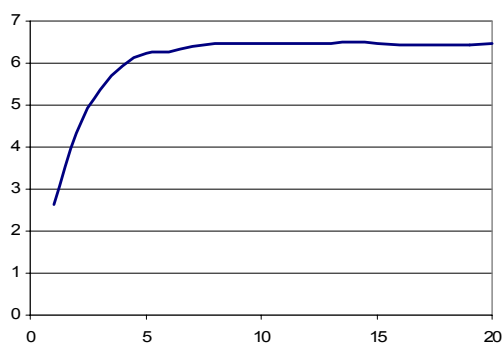


Figura IV. 61. Función de distribución del desplazamiento
promedio. MTA

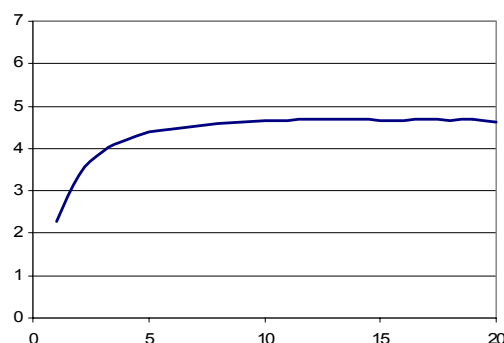


Figura IV. 62. Función de distribución del desplazamiento
promedio. BCN

En cuanto a la medida de la oscilación se presenta la función de distribución de la misma por unidad de longitud en la Figura IV. 63 y en la Figura IV. 64 y el número de oscilaciones por kilómetro en la Figura IV. 65 y la Figura IV. 66 .

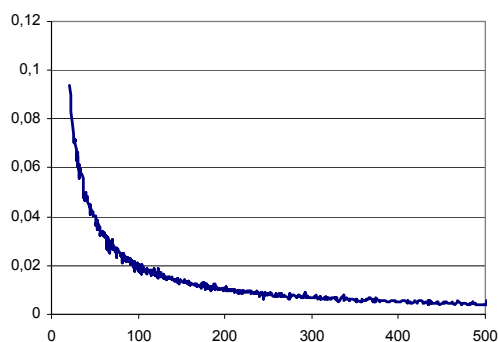


Figura IV. 63. Función de distribución la oscilación por longitud del tramo. MTA

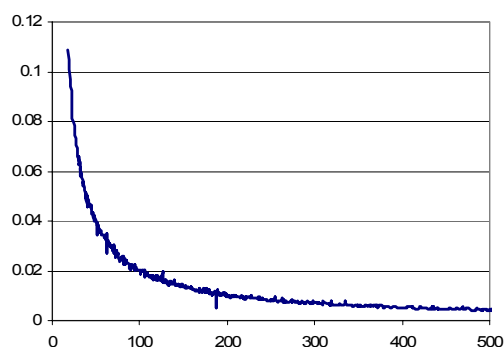


Figura IV. 64. Función de distribución la oscilación por longitud del tramo. BCN

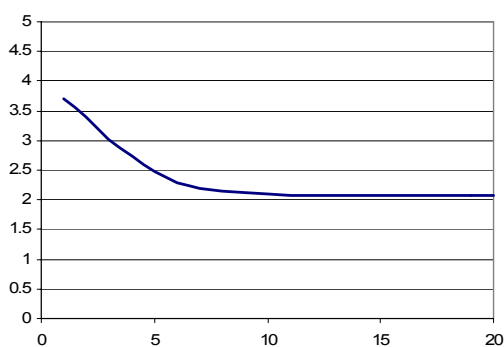


Figura IV. 65. Función de distribución la oscilación por kilómetro. MTA

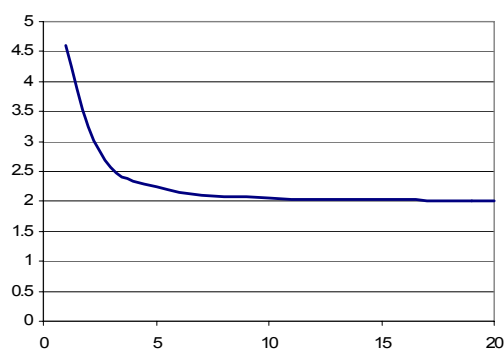


Figura IV. 66. Función de distribución la oscilación por kilómetro. BCN

El método MOD presenta unos resultados con difícil interpretación, aunque de los que se pueden extraer importantes conclusiones. Por un lado, estudiando el desplazamiento promedio y la superficie común se observa que las curvas del MTA tienen comportamientos más suaves o menos bruscos que las curvas generadas por la BCN. Sin embargo, a nivel numérico, la superficie común es menor en el MTA y el desplazamiento promedio es superior en el MTA con respecto a la BCN. Este comportamiento, que en principio parece una contradicción, nos ofrece una importante pista para descubrir cómo se comportan las bases de datos cartográficas utilizadas: la BCN es una base de datos que tiene mayor exactitud posicional que el MTA, aunque esta última contiene menos valores atípicos. Los valores atípicos hacen que las curvas de la BCN sean más irregulares, aunque presenten valores de calidad posicional superiores a los del MTA.

En cuanto al valor de la oscilación, presenta en ambos casos valores similares con curvas descendentes a medida que aumentamos la longitud del tramo, por lo que se puede inferir la no existencia de sesgos entre las bases de datos cartográficas y las correspondientes líneas de control. Las gráficas del número de oscilaciones por kilómetro confirman esta tesis, tendiendo las curvas a un valor cercano a 2, lo que

supone las oscilaciones mínimas presentes en los extremos de las líneas. En este caso el MTA parte de valores más bajos de oscilación, lo que indica una mayor aleatoriedad en la BCN.

Como en el caso del método MOS, las curvas obtenidas (superficie común y desplazamiento promedio) han sido chequeadas con el software estadístico SPSS para estudiar el ajuste a algunas distribuciones continuas de probabilidad. Teniendo en cuenta la función de densidad de las curvas resultantes (Figura IV. 67 y Figura IV. 68), se ha realizado una selección de las distribuciones más adecuadas para realizar el contraste. Las distribuciones contrastadas son la exponencial y la gamma. Los resultados aportados por el software estadístico en forma de proporciones acumuladas de cada curva respecto a las proporciones acumuladas de las distribuciones de contraste certifican que las funciones de distribución obtenidas con el método MOD no se ajustan a ninguna de ellas (ver gráficos en Anexo III).

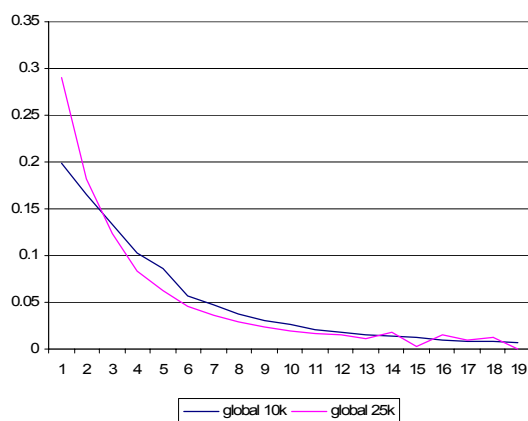


Figura IV. 67. Funciones de densidad de probabilidad de los resultados del método MOD(SC)

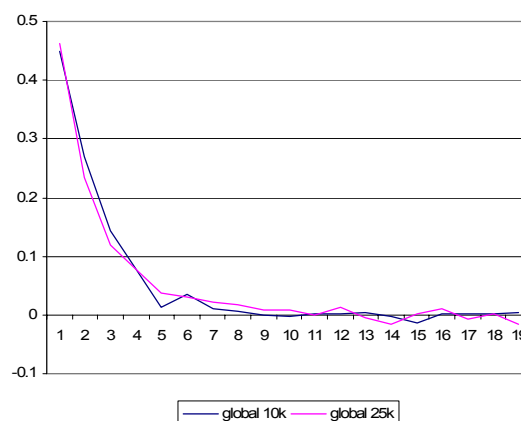


Figura IV. 68. Funciones de densidad de probabilidad de los resultados del método MOD(DP)

IV.5.5. Interpretación y análisis

Cada método utilizado ha aportado resultados con valores numéricos diferentes, lo cual es lógico dado que las métricas y los métodos son distintos. Los cuatro métodos se comportan, por tanto, como alternativas distintas, ya que funcionan y detectan situaciones diferentes. La idea de extender a los elementos lineales los principios del control puntual no es realizable de manera directa debido a la complejidad de estos elementos. Por tanto, es necesario definir más aspectos que en el caso puntual. Al considerar elementos más complejos, como es el caso de las líneas, existen más posibilidades de que una parte contenga errores o falsedades, por lo que es necesaria la existencia de métodos robustos que detecten estos problemas.

El método MDH es una opción que considera los elementos lineales como conjuntos de puntos y su interpretación es la de un valor máximo. Es muy sensible a la presencia de valores atípicos extremos que se producen cuando se dan situaciones de grandes discrepancias geométricas entre los elementos lineales a controlar y de

control. Para este valor máximo, basta con un único punto con gran discrepancia para generar valores altos de la distancia de Hausdorff. Apoyando este valor se ha probado el uso de valores de distancias medias para cada elemento obteniendo en estos casos valores de discrepancias medias. Por tanto, el uso combinado de estas dos métricas permite obtener una estimación de la existencia de valores atípicos.

El método MSK permite obtener una distancia a partir de una superficie diferencia. Constituye una extensión del concepto de distancia entre puntos al aplicarse a una geometría más compleja. El resultado puede considerarse como una distancia media entre las dos líneas. Con este método se consiguen amortiguar los valores atípicos que puedan aparecer entre las líneas, por lo que se muestra complementario a los resultados procedentes del método de las distancias de Hausdorff si no se consideran la variante de obtener las distancias medias de los elementos. En el método MSK, si se producen situaciones de casos puntualmente muy desplazados, que provoquen poca superficie al estar poco extendidos longitudinalmente, éstos tendrán escasa repercusión en el valor medio del elemento, y aún menor sobre el global de elementos. El estudio del número de cruces realizado permite interpretar qué base de datos cuenta con mayor oscilación y por tanto, tiene mayor aleatoriedad, mientras que los valores bajos muestran más desplazamientos sistemáticos.

El método MOS supone una concepción de los resultados desde un punto de vista probabilístico. El concepto de orlado proviene de la extensión a nivel puntual de elipse de error a través, dentro de los elementos lineales, de una banda de incertidumbre o banda ϵ . El resultado debe interpretarse como la probabilidad de que un elemento lineal se encuentre dentro de la banda de una anchura determinada. La probabilidad equivale a la longitud de ese elemento dentro de la banda, expresado en porcentajes. Por tanto, su comportamiento es similar al método MSK en el sentido de amortiguar la presencia de valores atípicos. Para una línea concreta, los valores extremos quedarán amortiguados siempre que la longitud de la línea que los abarca sea poco significativa con respecto a la longitud total de la línea. Los valores obtenidos son superiores a los del método MSK pero inferiores a los de las distancias de Hausdorff, siempre que el valor de confianza no sea muy alto, ya que si se toma un valor de confianza del 100%, el valor obtenido será más sensible a valores extremos.

El método MOD es el más complejo de todos los vistos hasta el momento, tanto en su materialización, como en la interpretación de los resultados. Conceptualmente, supone un paso más sobre el método MOS ya que se incluye la banda de incertidumbre a la línea de control y a controlar. De este concepto se deriva una distancia como cociente de superficies. El valor del desplazamiento promedio tiende a estabilizarse cuando el tamaño de los orlados alcanza el valor de desplazamiento promedio de las líneas. El porcentaje de superficie común no afecta realmente al valor de la estimación, por lo que no existe un sentido de probabilidad como el considerado en el método anterior. Los valores del desplazamiento promedio son similares a los obtenidos con el método MSK si se consideran valores de confianza en torno al 90%. El método MOD puede considerarse parecido al MSK. En cuanto a la medida de la oscilación, ha de considerarse la no existencia de sesgos entre las bases de datos, ya

que las curvas generadas presentan tendencias decrecientes a medida que se aumenta la longitud de los tramos. El estudio de la oscilación es similar al realizado con el número de cruces en el método del MSK.

En definitiva, los cuatro métodos expuestos poseen aspectos funcionales que los hacen distintos, siendo cada uno adecuado para ciertos propósitos concretos. En la Tabla IV. 33 se presenta un resumen de los resultados obtenidos para todos los métodos. La comparación entre los mismos ha de realizarse a partir de varios valores de confianza y sin comparar directamente las cantidades debido a la forma en que se obtienen los distintos resultados.

	MDH				MSK				MOS				MOD			
	50%	90%	95%	99%	50%	90%	95%	99%	50%	90%	95%	99%	50%	90%	95%	99%
MTA	10.0	17.7	19.9	37.4	3.7	7.3	8.6	13.6	3.4	9	11	17		6.3		
BCN	7.6	21.7	27.9	80.6	2.3	5.8	8.1	13.3	2	7.5	10	18			5.4	

Tabla IV. 33. Resumen de resultados numéricos para cada método de control

En cuanto a las bases de datos controladas, teniendo en cuenta los resultados de todos los métodos, se detecta que la BCN tiene un mejor comportamiento posicional, sin embargo, presenta mayores problemas con casos extremos no presentes en el MTA. Además, la BCN cuenta con desplazamientos más aleatorios que el MTA, en el que han de considerarse más sistemáticas estas diferencias.

La presencia de estos valores extremos puede deberse a falta de actualidad de los datos, problemas en la edición de los elementos, etc.

IV.6. Determinación de la muestra

IV.6.1. Criterios para la determinación de la muestra

Un punto importante en la metodología propuesta incluye la necesidad de analizar la posibilidad de que algunos criterios puedan afectar los muestreos de forma que estos pudieran considerarse para el diseño de los mismos. Concretando algo más, se pretende analizar si aspectos como el tipo, la longitud, la orientación o la sinuosidad deben ser tenidos en cuenta a la hora de plantear un muestreo de elementos lineales. Con este apartado, se pretende analizar si realmente las categorías consideradas se conforman como conjuntos con comportamientos diferentes. Para esto, es importante conocer la caracterización de los elementos de las bases de datos BCN, MTA y GPS planteada en el punto IV.4.

Un factor se puede considerar determinante si la existencia de categorías en el mismo conlleva comportamientos significativamente distintos en estas y la población total, lo que obligaría a incluirlo como criterio de estratificación en el diseño de los muestreos. Los criterios a considerar son:

- Tipo de vía de comunicación.

- Tamaño del tramo.
- Orientación del tramo.
- Sinuosidad.
- Zona de trabajo.

La información utilizada para estudiar esta similitud parte de la realización de orlados sobre los elementos GPS levantados y contabilizar la longitud de los elementos de las bases cartográficas que se encuentran incluidos dentro. En concreto, se trata de analizar la similaridad de las funciones de distribución obtenidas con el método MOS de control posicional, ya que la signatura de cada elemento se deriva de la función de orlado respecto a la línea GPS. El planteamiento práctico consiste en utilizar los criterios que etiquetan los elementos lineales para clasificar sus funciones de distribución obteniendo para cada conjunto o criterio una función de distribución que la caracteriza. Posteriormente se realiza el test de Kolmogorov-Smirnov sobre todas las funciones de distribución pertenecientes a las distintas categorías y el total de la población determinando si existen estas signaturas diferenciadas en cada categoría.

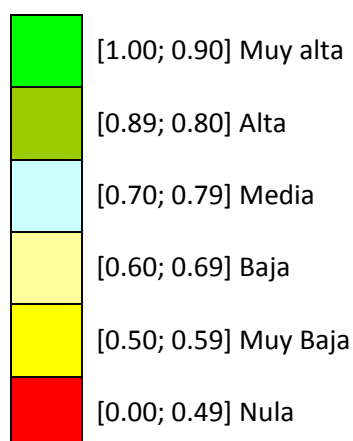
Los resultados obtenidos con esta técnica se presentan a continuación, teniendo en cuenta que se incluyen:

- Distancia frecuencial (D_{obs}): Se corresponde con la diferencia máxima en frecuencias acumuladas entre las dos funciones de distribución. Cuanto menor valor se obtenga existirá mayor similitud.
- Z de Kolmogorov-Smirnov: Procede de la tipificación de la distancia frecuencial bajo el supuesto común de normalidad que se aplica en este tipo de test estadístico.
- P-Valor: Es el nivel de confianza correspondiente a los valores anteriores. Los valores mayores y cercanos a la unidad se corresponden con los menores valores de D_{obs} .

Las tablas de resultados incluyen estos valores de la siguiente forma:

Distancia frecuencial
Z Kolmogorov Smirnov
P- valor

Presentando en función del nivel de confianza los siguientes sombreados.



IV.6.1.1 Zona de trabajo

Las funciones de distribución obtenidas por zona de trabajo se muestran en la Figura IV. 69 y la Figura IV. 70.

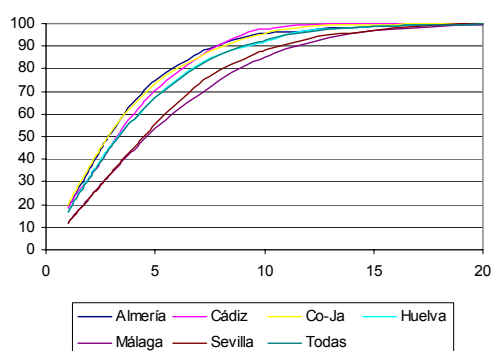


Figura IV. 69. Funciones de distribución MOS por zonas. MTA

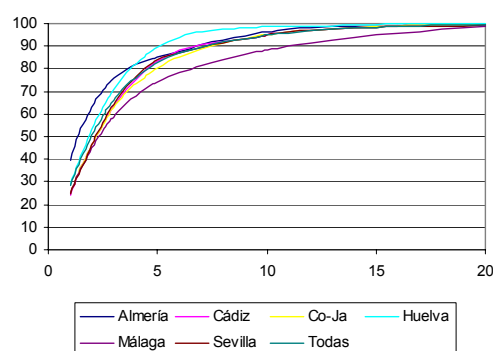


Figura IV. 70. Funciones de distribución MOS por zonas. BCN

Los resultados para el MTA (Tabla IV. 34) presentan un comportamiento diferenciado de las zonas de Málaga y Sevilla ya que las otras presentan mayor grado de similitud, aunque en algunos casos lleguen a ser bastante bajos ($\approx 55\%$). Con respecto al total de la población las zonas de Málaga y Sevilla vuelven a no satisfacer la condición de similaridad.

	Almería	Cádiz	Córdoba	Huelva	Málaga	Sevilla	Todas
Almería		0.157 0.792 0.557	0.078 0.396 0.998	0.078 0.396 0.998	0.235 1.188 0.199	0.216 1.089 0.186	0.098 0.495 0.967
Cádiz			0.137 0.693 0.723	0.157 0.792 0.557	0.176 0.891 0.405	0.176 0.891 0.405	0.157 0.792 0.557
Córdoba				0.078 0.396 0.998	0.235 1.188 0.199	0.235 1.188 0.199	0.098 0.495 0.967
Huelva					0.176 0.891 0.405	0.176 0.891 0.405	0.039 0.198 1
Málaga						0.059 0.297 1	0.176 0.891 0.405
Sevilla							0.176 0.891 0.405

Tabla IV. 34. Resultados similitud por zonas para el MTA

En el caso de la BCN, se detecta una zona con un comportamiento muy distinto que es Almería, y por otro lado, Huelva presenta valores bastante irregulares. Sin embargo, no se presenta un patrón tan acusado como en el caso del MTA.

	Almería	Cádiz	Córdoba	Huelva	Málaga	Sevilla	Todas
Almería		0.196 0.99 0.281	0.216 1.089 0.186	0.137 0.693 0.723	0.255 1.287 0.073	0.196 0.99 0.281	0.176 0.891 0.405
Cádiz			0.059 0.297 1	0.157 0.792 0.557	0.157 0.792 0.557	0.078 0.396 0.998	0.059 0.297 1
Córdoba				0.157 0.792 0.557	0.137 0.693 0.723	0.078 0.396 0.998	0.059 0.297 1
Huelva					0.255 1.287 0.073	0.157 0.792 0.557	0.157 0.792 0.557
Málaga						0.137 0.693 0.723	0.137 0.693 0.723
Sevilla							0.059 0.297 1

Tabla IV. 35. Resultados similitud por zonas para la BCN

IV.6.1.2 Tipo de vía

La Figura IV. 71 y la Figura IV. 72 presentan las funciones de distribución por tipo de vía para el MTA y la BCN. Las categorías son: Autonómicas, autovías, locales y nacionales.

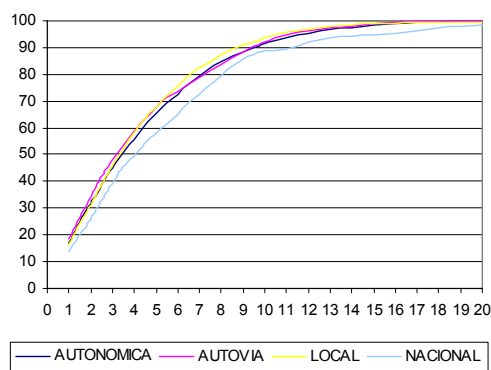


Figura IV. 71. Funciones de distribución MOS por tipo. MTA

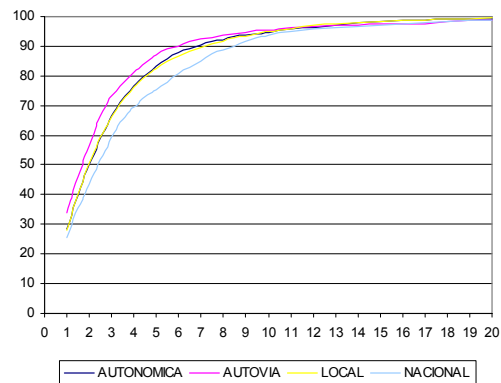


Figura IV. 72. Funciones de distribución MOS por tipo. BCN

Los resultados del Test de Kolmogorov-Smirnov para las funciones de distribución correspondientes a los tipos de vías se presentan en la Tabla IV. 36 y la Tabla IV. 37. Los resultados muestran una confianza alta o muy alta en casi todas las comparaciones, siendo las más bajas en torno al 70%. Los peores comportamientos se corresponden con las vías nacionales, donde el número de elementos es menor según se comprobó en el apartado de caracterización de elementos lineales. El comportamiento del MTA en cuanto a las vías de comunicación puede asumirse para todas las categorías como semejante al de la población global y semejantes entre sí.

	Autonómica	Autovía	Local	Nacional	Todas
Autonómica		0.059	0.039	0.118	0.039
		0.297	0.198	0.594	0.198
		1.000	1.000	0.872	1.000
Autovía			0.078	0.137	0.078
			0.396	0.693	0.396
			0.998	0.723	0.998
Local				0.137	0.020
				0.693	0.099
				0.723	1.000
Nacional					0.118
					0.594
					0.872

Tabla IV. 36. Resultados similitud por tipo de vía para el MTA

En la BCN se produce un caso parecido al MTA en general, sin embargo aparece una confianza muy baja al comparar las vías nacionales con las autovías. Este

comportamiento puede ser debido al escaso número de elementos pertenecientes a ambas categorías.

	Autonómica	Autovía	Local	Nacional	Todas
Autonómica		0.098 0.495 0.967	0.039 0.198 1	0.118 0.594 0.872	0.039 0.198 1
Autovía			0.118 0.594 0.872	0.176 0.891 0.405	0.118 0.594 0.872
Local				0.118 0.594 0.872	0.020 0.099 1
Nacional					0.118 0.594 0.872

Tabla IV. 37. Resultados similitud por tipo de vías para la BCN

IV.6.1.3 Longitud de tramo

Para la longitud de tramo las categorías establecidas son: 0-500 m, 500-1000 m, 1000-1500 m, 1500-2000 m, 2000-3000 m, 3000-5000 m y más de 5000 m. Según los resultados obtenidos se muestran las funciones de distribución representadas en la Figura IV. 73 y en la Figura IV. 74:

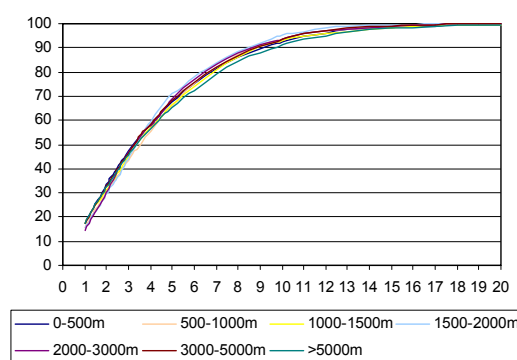


Figura IV. 73. Funciones de distribución MOS por longitud. MTA

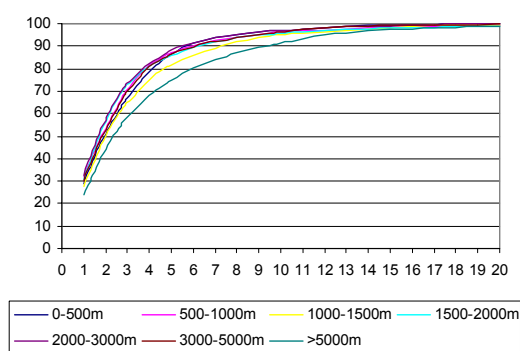


Figura IV. 74. Funciones de distribución MOS por longitud. BCN

Los resultados obtenidos se presentan en la Tabla IV. 38 y en la Tabla IV. 39. Los valores de similitud son altos, contando el MTA con un comportamiento muy superior al obtenido en la BCN. Los resultados del MTA muestran que todas las categorías se comportan de forma similar entre sí y con la población total.

	0-500m	500-1000m	1000-1500m	1500-2000m	2000-3000m	3000-5000m	>5000m	Todas
0-500m		0.059 0.297 1	0.059 0.297 1	0.098 0.495 0.967	0.059 0.297 1	0.059 0.297 1	0.039 0.198 1	0.039 0.198 1
500-1000m			0.059 0.297 1	0.098 0.495 0.967	0.078 0.396 0.998	0.078 0.396 0.998	0.039 0.198 1	0.059 0.297 1
1000-1500m				0.078 0.396 0.998	0.059 0.297 1	0.059 0.297 1	0.039 0.198 1	0.039 0.198 1
1500-2000m					0.059 0.297 1	0.059 0.297 1	0.098 0.495 0.967	0.078 0.396 0.998
2000-3000m						0.039 0.198 1	0.078 0.396 0.998	0.059 0.297 1
3000-5000m							0.078 0.396 0.998	0.039 0.198 1
>5000m								0.039 0.198 1

Tabla IV. 38. Resultados similitud por longitud del tramo para el MTA

Los resultados de la BCN muestran valores altos para la gran mayoría de categorías, aunque presenta algunos problemas con los tramos superiores a 5000 m, en los que aparecen valores más bajos. Sin embargo, el comportamiento respecto al total de líneas ofrece valores altos.

	0-500m	500-1000m	1000-1500m	1500-2000m	2000-3000m	3000-5000m	>5000m	Todas
0-500m		0.098 0.495 0.967	0.098 0.495 0.967	0.137 0.693 0.723	0.098 0.495 0.967	0.059 0.297 1	0.157 0.792 0.557	0.098 0.495 0.967
500-1000m			0.098 0.495 0.967	0.059 0.297 1	0.078 0.396 0.998	0.098 0.495 0.967	0.176 0.891 0.405	0.098 0.495 0.967
1000-1500m				0.118 0.594 0.872	0.137 0.693 0.723	0.098 0.495 0.967	0.118 0.594 0.872	0.039 0.198 1
1500-2000m					0.098 0.495 0.967	0.118 0.594 0.872	0.196 0.99 0.281	0.098 0.495 0.967
2000-3000m						0.059 0.297 1	0.196 0.99 0.281	0.118 0.594 0.872
3000-5000m							0.157 0.792 0.557	0.098 0.495 0.967
>5000m								0.118 0.594 0.872

Tabla IV. 39. Resultados similitud por longitud del tramo para la BCN

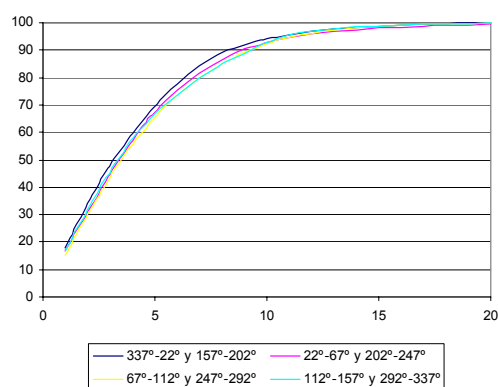
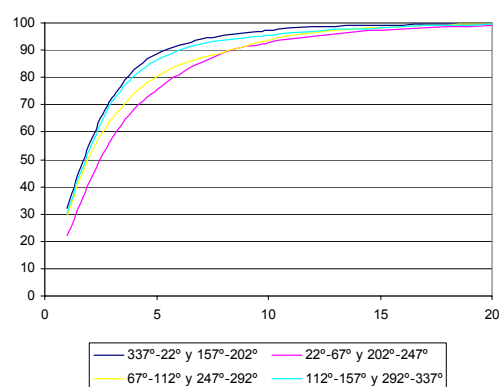
La consideración de longitud de tramo está íntimamente relacionada, además de con la original de la línea en la base de datos, con la tramificación provocada por la observación con GPS de los elementos de control, ya que durante el proceso de captura se producen pérdidas de señal que provocan esta tramificación.

IV.6.1.4 Orientación

La orientación se refiere al ángulo de la línea base (unión del punto inicial y final) del elemento. Las categorías utilizadas son:

- 337° - 22° y 157° - 202°
- 22° - 67° y 202° - 247°
- 67° - 112° y 247° - 292°
- 112° - 157° y 292° - 337°

Las gráficas resultantes de las funciones de distribución para la orientación se presentan en la Figura IV. 75 y la Figura IV. 76.

Figura IV. 75. Funciones de distribución MOS por orientación.
MTAFigura IV. 76. Funciones de distribución MOS por orientación.
BCN

La Tabla IV. 40 y la Tabla IV. 41 presentan los resultados del Test de Kolmogorov-Smirnov para las funciones de distribución procedentes del estudio por orientación.

	337° - 22°	22° - 67°	67° - 112°	112° - 157°	157° - 202°	202° - 247°	247° - 292°	292° - 337°	Todas
337°-22°		0.059 0.297 1	0.118 0.594 0.872	0.098 0.495 0.967	0.098 0.495 0.967	0.157 0.792 0.557	0.078 0.396 0.998	0.078 0.396 0.998	0.078 0.396 0.998
22°-67°			0.137 0.693 0.723	0.059 0.297 1	0.078 0.396 0.998	0.157 0.792 0.557	0.078 0.396 0.998	0.098 0.495 0.967	0.078 0.396 0.998
67°-112°				0.118 0.594 0.872	0.118 0.594 0.998	0.078 0.396 0.998	0.078 0.396 0.998	0.078 0.396 0.998	0.078 0.396 0.998
112°-57°					0.059 0.297 1	0.157 0.792 0.557	0.098 0.495 0.967	0.098 0.495 0.967	0.059 0.297 1
157°-202°						0.157 0.792 0.557	0.059 0.297 1	0.078 0.396 0.998	0.059 0.297 1
202°-247°							0.118 0.594 0.872	0.118 0.594 0.872	0.118 0.594 0.872
247°-292°								0.039 0.198 1	0.078 0.396 0.998
292°-337°									0.059 0.297 1

Tabla IV. 40. Resultados similitud por orientación para el MTA

Para el caso del MTA, los valores altos son mayoritarios por lo que se puede suponer un comportamiento similar y homogéneo, tanto entre las diversas categorías como con el total. La excepción la presenta la clase 202°-247° donde existe un comportamiento peor que en el resto. Este caso atípico está relacionado con la distribución por zonas, ya que la clase 202°-247° se encuentra muy influenciada por líneas de la zona de Málaga (Figura IV. 44), que como se mostró anteriormente presentaba valores desmarcados del resto.

Para el caso de la BCN el comportamiento es muy heterogéneo para el cruce entre categorías, apareciendo valores muy bajos en gran parte de los datos. Sin embargo, respecto al total, los valores resultan más aceptables. Estos casos están relacionados también con la distribución por zonas (Figura IV. 45).

	337° - 22°	22° - 67°	67° - 112°	112° - 157°	157° - 202°	202° - 247°	247° - 292°	292° - 337°	Todas
337°-22°		0.196 0.99 0.281	0.176 0.891 0.405	0.098 0.495 0.967	0.137 0.693 0.723	0.216 1.089 0.186	0.137 0.693 0.723	0.137 0.693 0.723	0.137 0.693 0.723
22°-67°			0.078 0.396 0.998	0.196 0.99 0.281	0.176 0.891 0.405	0.059 0.297 1	0.137 0.693 0.723	0.157 0.792 0.557	0.118 0.594 0.872
67°-112°				0.176 0.891 0.405	0.157 0.792 0.557	0.137 0.693 0.723	0.098 0.495 0.967	0.118 0.594 0.872	0.078 0.396 0.998
112°-157°					0.118 0.594 0.872	0.216 1.089 0.186	0.098 0.495 0.967	0.098 0.495 0.967	0.118 0.594 0.872
157°-202°						0.196 0.99 0.281	0.098 0.495 0.967	0.118 0.594 0.872	0.118 0.594 0.872
202°-247°							0.176 0.891 0.405	0.157 0.792 0.557	0.137 0.693 0.723
247°-292°								0.059 0.297 1	0.059 0.297 1
292°-337°									0.059 0.297 1

Tabla IV. 41. Resultados similitud por orientación para la BCN

IV.6.1.5 Sinuosidad

La sinuosidad se clasifica en tres clases: muy suave, suaves y sinuosa. La Figura IV. 77 y la Figura IV. 78 presentan las curvas representativas de las funciones de distribución obtenidas.

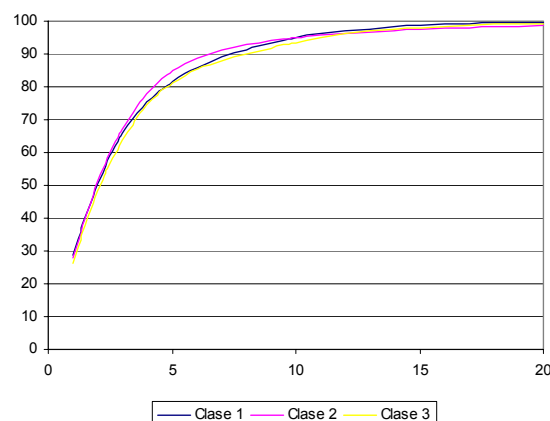
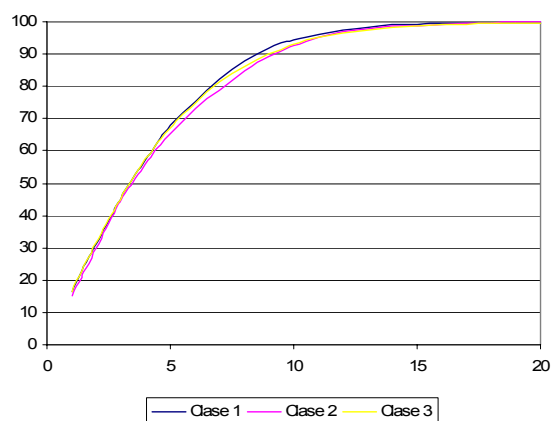


Figura IV. 77. Funciones de distribución MOS por sinuosidad. MTA Figura IV. 78. Funciones de distribución MOS por sinuosidad. BCN

Los resultados del Test de Kolmogorov-Smirnov se muestran en la Tabla IV. 42 y la Tabla IV. 43. Como puede apreciarse los valores obtenidos son bastante altos, por lo que las funciones de distribución obtenidas pueden considerarse similares.

	Clase 1	Clase 2	Clase 3	Todas
Clase 1		0.058	0.039	0.019
		0.297	0.198	0.099
		0.999	0.999	1
Clase 2			0.039	0.039
			0.198	0.192
			0.999	0.999
Clase 3				0.039
				0.192
				0.999

Tabla IV. 42. Resultados similitud por sinuosidad para el MTA

	Clase 1	Clase 2	Clase 3	Todas
Clase 1		0.098	0.059	0.059
		0.495	0.297	0.297
		0.966	1	1
Clase 2			0.078	0.078
			0.396	0.396
			0.999	0.999
Clase 3				0.02
				0.099
				1

Tabla IV. 43. Resultados similitud por sinuosidad para la BCN

IV.6.1.6 Interpretación y análisis

A través de la categorización de los elementos lineales utilizados según una serie de criterios (zona de trabajo, tipo de vía, longitud de tramo, orientación y sinuosidad) con capacidad para afectar a los resultados de los muestreos, se han establecido ciertas clases para cada uno de ellos caracterizado por una función de distribución. Con estas funciones de distribución, se ha realizado un análisis de similitud a través de la prueba de Kolmogorov-Smirnov para las diversas categorías entre sí y con el total de la población.

Los resultados presentan comportamientos distintos para las dos bases cartográficas, siendo siempre más variantes para la BCN, algo coherente con los resultados obtenidos previamente.

La tipología de las vías de comunicación y las longitudes de los tramos parecen no tener relevancia para el MTA y valores algo más dispares para la BCN. Con la orientación, a excepción de una categoría, el resto puede considerarse similares entre sí para el MTA, siendo este criterio más variable para la BCN. La sinuosidad presenta categorías similares en todos los casos.

Considerando los valores obtenidos para el tipo de vía, la longitud del tramo, la orientación y sinuosidad parece no haberse generado para cada una de las clases funciones de distribución distintas del total poblacional, por lo que, no es necesario tenerlas en cuenta a la hora de desarrollar los muestreos.

El único criterio que parece tener cierta influencia es la zona de trabajo y sólo para este criterio se desmarcan las zonas de Málaga y Sevilla. La relevancia de las zonas de trabajo viene a significar la necesidad de aplicar correctamente el criterio de generar muestras con una distribución adecuada.

IV.6.2. Estimación del tamaño de la muestra

La determinación del tamaño de la muestra para el control posicional por elementos lineales se realiza siguiendo la metodología propuesta en el capítulo anterior. Dicha metodología se apoya en un proceso de simulación para obtener una estimación empírica de los tamaños de muestra adecuados.

Si se tienen en cuenta los resultados aportados por las pruebas de similitud entre distintas categorías o conjuntos de líneas agrupadas según una serie de criterios, ninguno de los factores analizados posee repercusión real, por lo que la muestra puede considerarse libre de plantearse en función de ellos. Por lo tanto, se puede establecer un muestreo aleatorio simple como opción a priori.

Los resultados obtenidos con cada uno de los métodos de control son distintos, ya que los métodos MDH y MSK generan unos valores escalares promedio pudiendo seguir una función de distribución que en principio es desconocida, mientras que los métodos MOS y MOD proporcionan funciones de distribución, pero también pueden aportar unos valores escalares medios para un cierto nivel de confianza.

La idea básica consiste en la extracción aleatoria de muestras y el análisis de las estimaciones derivadas de cada una de ellas frente a la población total (técnica bootstrapp). Dada la extensión de la zona de trabajo en las dos bases de datos, se

puede considerar un conjunto totalmente representativo de la población, por lo que los resultados globales pueden considerarse como elementos comparativos.

Las comparativas se realizan siguiendo el método anteriormente aplicado del Test de Kolmogorov-Smirnov para contrastar la similaridad de las funciones de distribución obtenidas por simulación y la global.

Los pasos seguidos en el proceso de simulación son:

- 1) Se consideran tamaños de muestra T_i , en kilómetros de longitud, de 10 hasta 200 km, con un paso de 10 km.
- 2) Se extraen, mediante un proceso aleatorio, tramos de la población hasta que la suma de la longitud de los mismos alcance el valor de T_i . Si el primero o último tramo completo incluido excede se toma sólo en la parte necesaria y se continúa con el paso 3º.
- 3) Se construye la función de distribución correspondiente a la muestra $M(T_i)$ y se compara con la función de distribución de la población mediante el Test de Kolmogorov-Smirnov obteniendo los valores Z_{KS} y p-valor.
- 4) Se repiten los pasos 2º y 3º 1000 veces.
- 5) Se promedian los valores Z_{KS} y p-valor.
- 6) Se incrementa T_i y se vuelve al paso 1º hasta que T_i llegue a 200 km.

Para los cuatro métodos de control presentados, los resultados se obtienen en función de su forma:

- Valores escalares: Se utiliza una representación que recoge en el eje de abscisas el tamaño de la muestra considerado en kilómetros y en el eje de ordenadas el valor de la estimación, en metros, para cada tamaño de la muestra. La curva central se corresponde con el valor medio estimado, mientras que las curvas extremas (línea discontinua) representan los intervalos correspondientes al considerar la desviación estándar de la estimación por encima y por debajo.
- Funciones de distribución: La representación recoge para el eje de abscisas el tamaño de la muestra (km) y en el eje de ordenadas un valor de probabilidad. En este caso, se presentan de forma independiente las estimaciones del parámetro Z_{KS} y del p-valor.

IV.6.2.1 Método MDH

Los resultados para la simulación propuesta se presentan en la Figura IV. 79 y la Figura IV. 80.

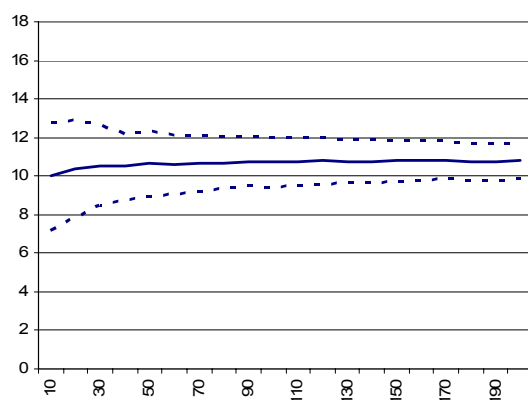


Figura IV. 79. Evolución de la estimación según MDH con el tamaño de la muestra. MTA

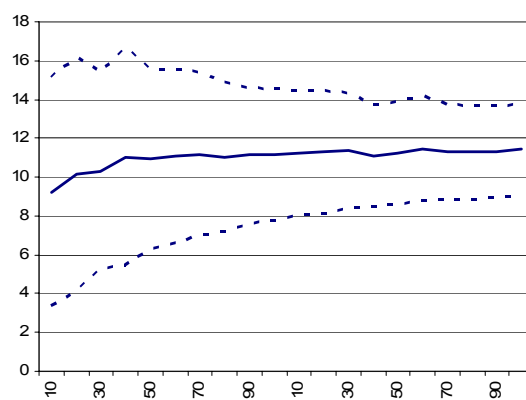


Figura IV. 80. Evolución de la estimación según MDH con el tamaño de la muestra. BCN

Las curvas promedio (línea continua) coinciden con los valores resultantes obtenidos para el total de la población (10.82 m y 11.42 m. apartado IV.5.1) incluso en los menores tamaños de muestra. La línea promediada presenta en el caso del MTA mayor suavidad que en el caso de la BCN, así como las líneas de las desviaciones, aunque estas curvas tienden a juntarse con el aumento del tamaño de muestra.

IV.6.2.2 Método MSK

Análogamente al método mostrado anteriormente, se presentan los resultados de aplicar la simulación propuesta a los valores obtenidos con el método MSK en la Figura IV. 81 y la Figura IV. 82.

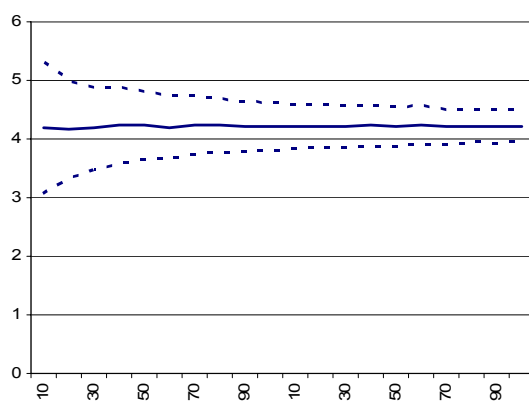


Figura IV. 81. Evolución de la estimación según MSK con el tamaño de la muestra. MTA

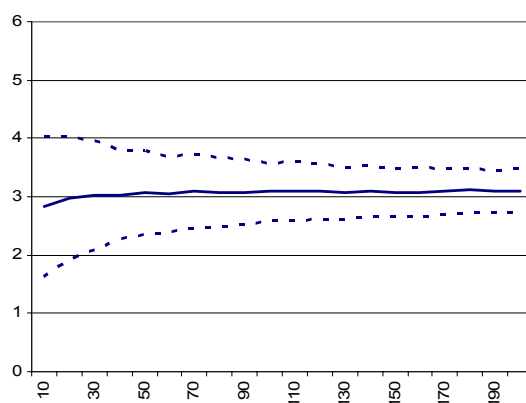


Figura IV. 82. Evolución de la estimación según MSK con el tamaño de la muestra. BCN

Las curvas promedio resultantes se mueven en torno a los valores obtenidos para el total de la población (4.22 m y 3.11 m. apartado IV.5.2). Los valores medios estimados presentan una gran estabilidad en los dos casos, mientras que las

dispersiones muestran una clara tendencia a la disminución al aumentar el tamaño de muestra.

IV.6.2.3 Método MOS

Con el método MOS se procede de forma similar a los casos anteriores presentando los resultados obtenidos de la simulación propuesta a partir de los valores escalares obtenidos. Además, se aportan los resultados procedentes de la aplicación de la simulación a las funciones de distribución, ya que estas son el producto básico de esta metodología.

Los resultados procedentes de la aplicación de la simulación a los valores escalares son obtenidos a partir de la estimación al 95% de confianza y se presentan en la Figura IV. 83 y la Figura IV. 84.

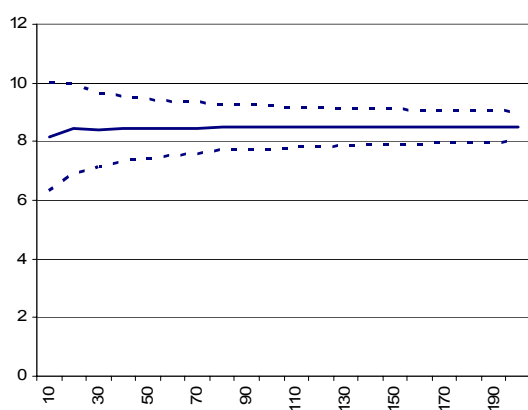


Figura IV. 83. Evolución de la estimación según MOS al 95% con el tamaño de la muestra. MTA

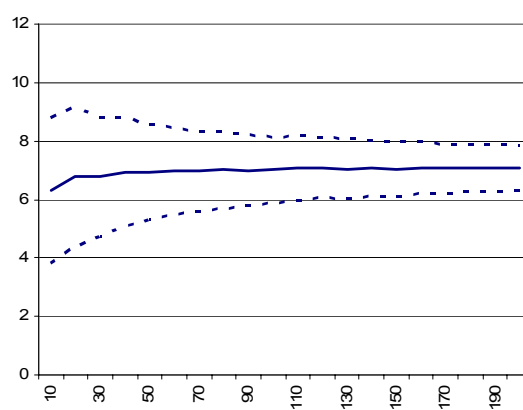


Figura IV. 84. Evolución de la estimación según MOS al 95% con el tamaño de la muestra. BCN

El comportamiento general coincide con los resultados obtenidos en los métodos anteriores, con la curva promedio estable y las desviaciones decrecientes con el tamaño de la muestra, presentando mayor discrepancia los relativos a la BCN.

En lo relativo a la aplicación de la simulación a las funciones de distribución obtenidas con el método MOS, los resultados se presentan tanto para el parámetro Z_{KS} como para el p-valor en la Figura IV. 85, la Figura IV. 86, la Figura IV. 87 y la Figura IV. 88.

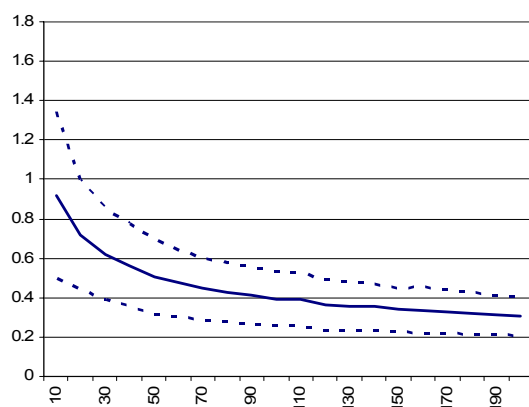


Figura IV. 85. Evolución de la estimación según MOS a la Z_{KS} con el tamaño de la muestra. MTA

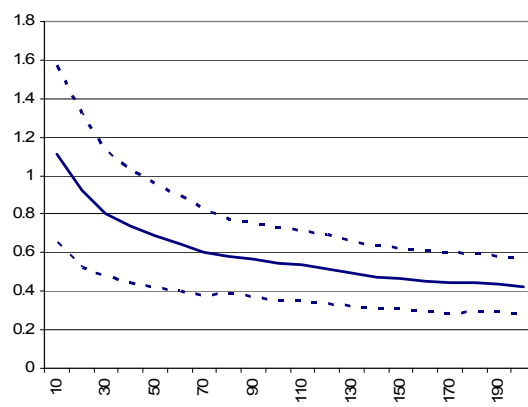


Figura IV. 86. Evolución de la estimación según MOS a la Z_{KS} con el tamaño de la muestra. BCN

Las curvas relativas al estadístico Z_{KS} muestran una evolución decreciente, es decir, un decremento de la distancia entre las funciones de distribución a medida que se aumenta el tamaño de la muestra.

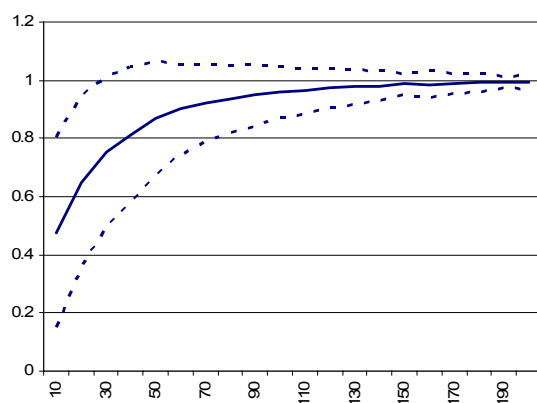


Figura IV. 87. Evolución de la estimación según MOS al p-valor con el tamaño de la muestra. MTA

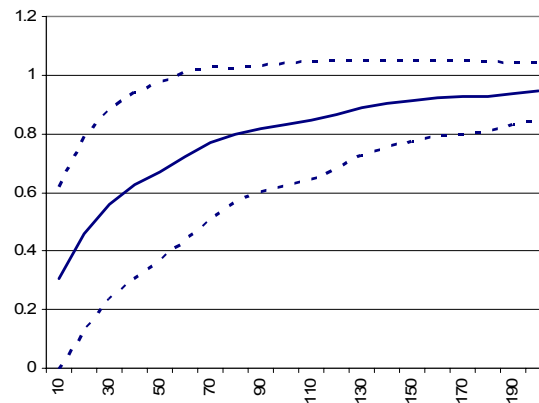


Figura IV. 88. Evolución de la estimación según MOS al p-valor con el tamaño de la muestra. BCN

Las gráficas relativas al p-valor representan la evolución de la confianza en la similitud en las funciones de distribución por lo que las curvas son crecientes conforme se aumenta el tamaño de la muestra tendiendo al 100%. Las curvas permiten deducir que para valores de Z_{KS} inferiores a 0.4 los valores del p-valor se acercan al 100% y que la BCN alcanza los valores más altos de p-valor a tamaños de muestra más altos.

IV.6.2.4 Método MOD

Con el método MOD se aplican las mismas herramientas que en el caso anterior. Por un lado se realiza una estimación del desplazamiento promedio para un

ancho de orlado de 20 m ($\approx 95\%$ de confianza), y por otro se realiza la simulación con las funciones de distribución obtenidas a partir de la medida de la superficie común.

La Figura IV. 89 y la Figura IV. 90 presentan los resultados para la primera métrica indicada.

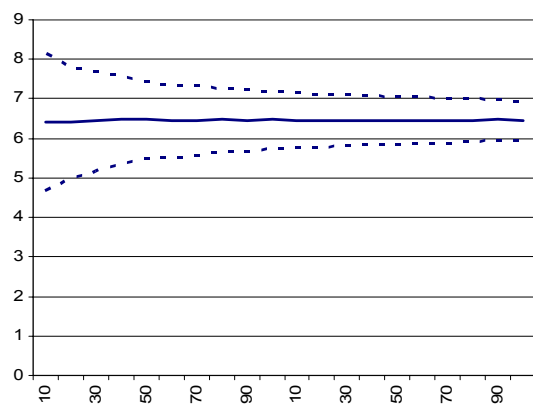


Figura IV. 89. Evolución de la estimación según MOD a 20m con el tamaño de la muestra. MTA

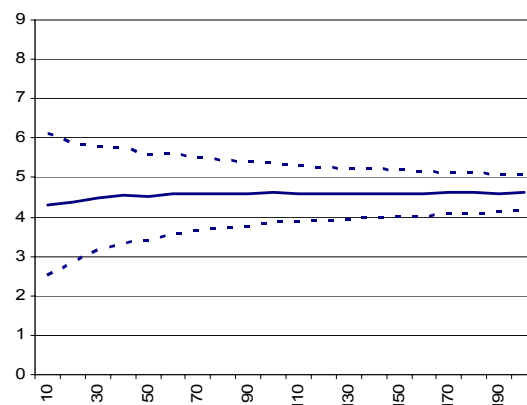


Figura IV. 90. Evolución de la estimación según MOS a 20m con el tamaño de la muestra. BCN

El desarrollo de estas curvas coincide, en líneas generales con los casos anteriores, aunque en este caso existe para pequeños tamaños de muestra mayor desviación en el caso de la BCN.

La aplicación de la simulación a las funciones de distribución presenta los resultados tanto para el parámetro Z_{KS} como para el p-valor que se muestran en la Figura IV. 91, la Figura IV. 92, la Figura IV. 93 y la Figura IV. 94.

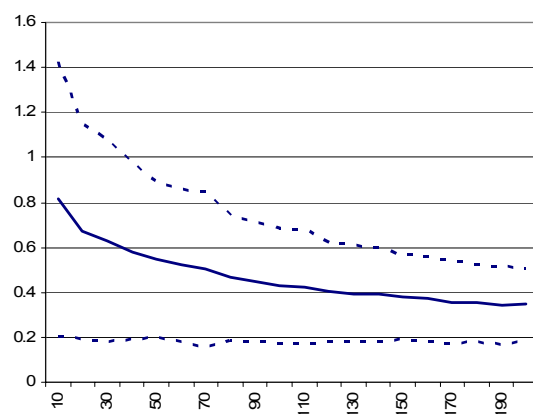


Figura IV. 91. Evolución de la estimación según MOD a la Z_{KS} con el tamaño de la muestra. MTA

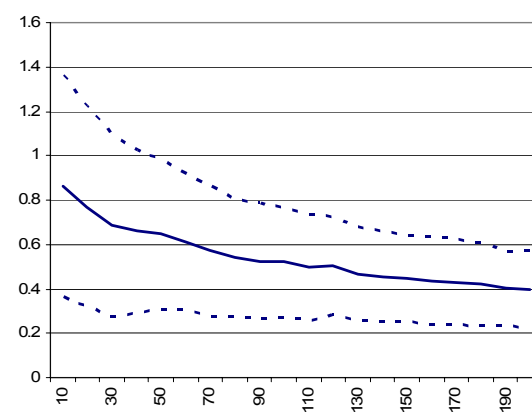


Figura IV. 92. Evolución de la estimación según MOD a la Z_{KS} con el tamaño de la muestra. BCN

Los resultados del estadístico Z_{KS} para el MTA y la BCN están en consonancia con los valores obtenidos para el método MOS. El comportamiento para la BCN es algo superior con respecto al MTA, en el que la curva presenta un comportamiento más estable, aunque en ambos casos desciende con el aumento del tamaño de la muestra.

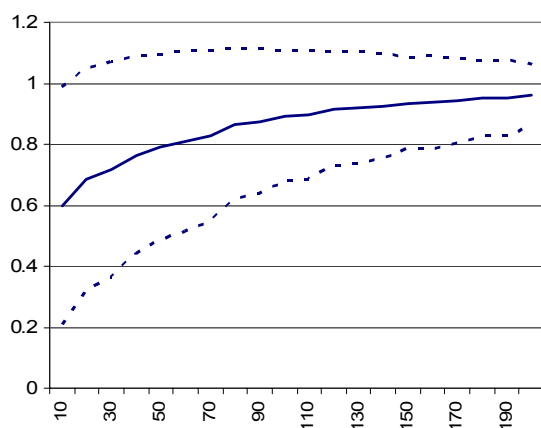


Figura IV. 93. Evolución de la estimación según MOD al p-valor con el tamaño de la muestra. MTA

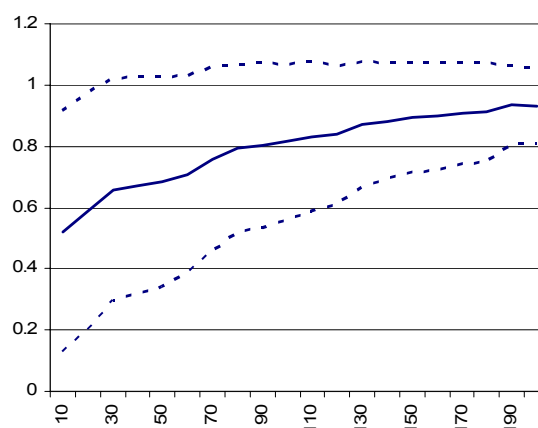


Figura IV. 94. Evolución de la estimación según MOD al p-valor con el tamaño de la muestra. BCN

Para el caso del p-valor ocurre algo similar al estadístico Z_{KS} . El MTA tiene un comportamiento más estable que la BCN. En los dos casos la curva tiende a un p-valor unidad con el aumento del tamaño de muestra.

IV.6.2.5 Pruebas con el tamaño de muestra

Del estudio de los resultados previamente descritos se puede establecer un tamaño de muestra mínimo aproximado de 150-200 kms.

Con éste valor de tamaño de muestra obtenido se realiza una prueba con dos muestras de 200 kms aproximadamente, siguiendo itinerarios posibles de observación en campo. Los resultados de estas pruebas se comparan con los obtenidos para el global de la población para comprobar si ese valor mínimo de tamaño de muestra obtenido es válido.

Las trayectorias planteadas para esta prueba se componen de las líneas descritas en la Figura IV. 95, la Figura IV. 96 y la Tabla IV. 44.



Figura IV. 95. Muestra Almería-Málaga-Sevilla



Figura IV. 96. Muestra Córdoba_Jaén-Cádiz-Huelva

Muestra	Zonas	Número de líneas MTA	Número de líneas BCN	Longitud MTA [m]	Longitud BCN [m]
1	Almería, Málaga, Sevilla	121	136	180822	209676
2	Co-Ja, Cádiz, Huelva	127	125	204565	214794

Tabla IV. 44. Muestras para comprobación del tamaño de la muestra

Los resultados obtenidos con estas muestras para los distintos métodos se presentan en la Tabla IV. 45 y la Tabla IV. 46.

Muestra	MTA			BCN		
	Distancias Hausdorff [m]	Media de distancias [m]	MSK [m]	Distancias Hausdorff [m]	Media de distancias [m]	MSK [m]
Global	10.82	4.44	4.22	11.42	3.36	3.11
1	11.18	4.76	4.53	13.94	3.87	3.44
2	11.15	4.07	3.76	12.43	3.28	3.38

Tabla IV. 45. Resultados muestras para comprobación del tamaño de muestras: MDH Y MSK

En general, la comparación de los valores globales con las medias de las 2 muestras tomadas da valores muy similares en todas las medidas propuestas, siendo la máxima diferencia observable en la medida de las distancias de Hausdorff para la BCN, donde aparece una diferencia de 2.52 m del global respecto a la muestra número 1. Sin embargo, esta diferencia está dentro de los márgenes de dispersión presentados anteriormente (Figura IV. 80).

Para los métodos MOS y MOD se establece un estudio de la similaridad de las funciones de distribución obtenidas mediante la prueba de Kolmogorov-Smirnov. Los resultados se muestran en la Tabla IV. 46, donde se utilizan los criterios de representación presentados en el punto IV.6.1.

	MTA MOS	MTA MOD(SC)	BCN MOS	BCN MOD(SC)
Muestra 1	0.059	0.1	0.078	0.15
	0.297	0.316	0.396	0.474
	1	1	0.996	0.965
Muestra 2	0.078	0.15	0.078	0.1
	0.396	0.474	0.396	0.316
	0.996	0.965	0.996	1

Tabla IV. 46. Resultados muestras para comprobación del tamaño de muestras: MOS Y MOD

La prueba de Kolmogorov-Smirnov presenta valores de similaridad entre las curvas generadas por los métodos MOS y MOD para la superficie común en ambas muestras.

Los resultados de esta prueba permiten confirmar que el tamaño de muestra elegido es apropiado para todos los métodos y medidas.

IV.6.2.6 Interpretación y análisis

El método de simulación empleado se ha mostrado eficaz para resolver de forma automática el problema de mostrar la evolución de los parámetros de interés para distintos tamaños de muestras. Además, el motor de simulaciones permitiría la introducción de criterios de selección para crear muestras estratificadas y su estimación correspondiente. La elección del esquema de muestreo del aleatorio simple resulta adecuada por coherencia con los resultados de los muestreos frente a las estimaciones globales.

Los resultados presentan unos comportamientos lógicos en sus valores medios (tendenciales) y de dispersión. Gráficas como las utilizadas en el análisis pueden considerarse una herramienta adecuada para derivar el tamaño de muestra en función de la precisión que se desee dar al muestreo.

Establecer una recomendación de tamaño de muestra es complejo debido a que cada producto y método (MDH, MSK...) tienen un comportamiento distinto; aunque presenten evoluciones parecidas. Esta determinación debe hacerse en función de la precisión que quiera darse al muestreo.

No obstante, para bases cartográficas que no presenten problemas de actualización ni de cambios de geometría severos respecto a los elementos de control, se puede trabajar con tamaños de muestras en torno a 150-200 kms, obteniendo ya precisiones aceptables. Este tamaño se refiere a las bases de datos definitivas, una vez depuradas y editadas, por lo que se deben tener en cuenta los rendimientos de todo el proceso para establecer el tamaño real de trabajo a ejecutar.

Respecto a las dos bases cartográficas utilizadas se confirma la existencia de un comportamiento distinto. El comportamiento del MTA es mucho mejor que el de la BCN en cuanto a los valores de dispersión, aunque las curvas promedio evidencian que la BCN cuente con valores inferiores.

Con el tamaño de muestra planteado se ha realizado una prueba con dos conjuntos de líneas que siguen un posible itinerario de observación en campo con longitudes aproximadas a los 200 kms. Los resultados comparados con los valores globales para los distintos métodos confirman la adopción de este valor aproximado de 200 kms como tamaño de muestra apropiado.

En definitiva, todo lo anterior sugiere un estudio y recomendación particular para cada caso, donde incluso convendría utilizar un muestreo piloto y una simulación similar a la aquí realizada para orientar los trabajos definitivos.

IV.6.3. Determinación de la densidad de vértices de control

La metodología propuesta de evaluación de la similaridad entre funciones de distribución puede ser aplicable al cálculo de la densidad de vértices en los elementos de control. Los datos procedentes de observaciones GPS pueden contener información redundante debido a que por circunstancias de la toma parece recomendable observar a velocidades bajas y con un intervalo de observación mínimo generalmente de un segundo. Sin embargo, es razonable estudiar la posibilidad de un aumento de la velocidad del vehículo realizando el cálculo de la cadencia mínima de toma de puntos

con la finalidad de mejorar los rendimientos de la toma y lo que es más importante, la seguridad vial del vehículo durante la toma de datos.

Es interesante, por tanto, conocer para cada base cartográfica utilizada, cuál hubiera sido la cadencia de toma suficiente para que las funciones de distribución fueran similares.

De esta forma, se plantea la eliminación de cada elemento lineal de control de un número incremental de vértices para estudiar cuál es el límite de dicha simplificación. La técnica de eliminación simulada, consiste en mantener 1 de cada 2 vértices, 1 de cada 5, 1 de 10, 1 de 15 y 1 de 20 puntos y realizar el control posicional con todos los métodos a estas “nuevas” bases de datos. Esta eliminación sistemática de puntos se realiza de esta forma con la finalidad de simular el efecto producido por un aumento de la velocidad del vehículo. Para mantener la coherencia lógica de las líneas siempre se mantienen los puntos inicial y final de cada elemento.

IV.6.3.1 Método MDH

En este caso, la comparación se realiza a través de la media de los valores máximos de las distancias de Hausdorff, ya que no es posible plantear un análisis de similitud a través de funciones de distribución.

En la Figura IV. 97 y en la Figura IV. 98 se presentan los resultados obtenidos, donde en el eje de abscisas se presenta el resultados de cada caso (1 de cada 2, 1 de cada 5, etc.) y en el eje de ordenadas el valor en metros obtenido, representando la línea horizontal de color rojo el valor sin eliminación.

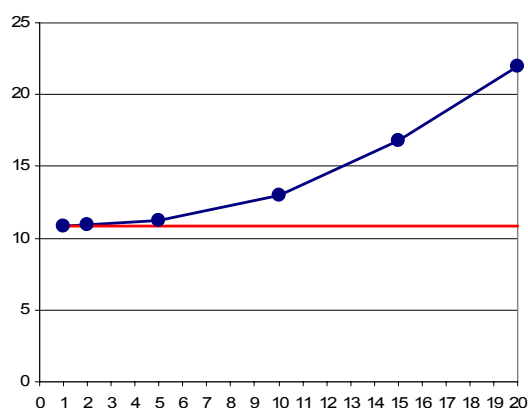


Figura IV. 97. Evolución valores medios MDH con eliminación de vértices. MTA

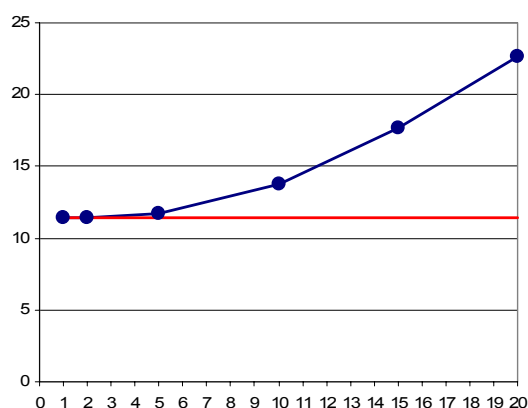


Figura IV. 98. Evolución valores medios MDH con eliminación de vértices. BCN

Los resultados presentan valores medios similares aproximadamente hasta el intervalo de 5 a 10 donde las curvas empiezan a separarse. En el caso del MTA la curva tiende a separarse con anterioridad al caso de la BCN. Por tanto, los resultados para este método se pueden considerar similares hasta un límite de mantenimiento de

puntos de la base de datos de control de 1 de cada 5 (20% de los puntos) a 1 de cada 10 (10% de los puntos).

IV.6.3.2 Método MSK

Para el método MSK se examina también los valores medios obtenidos de los resultados de este método con las bases cartográficas y la base de control con vértices eliminados.

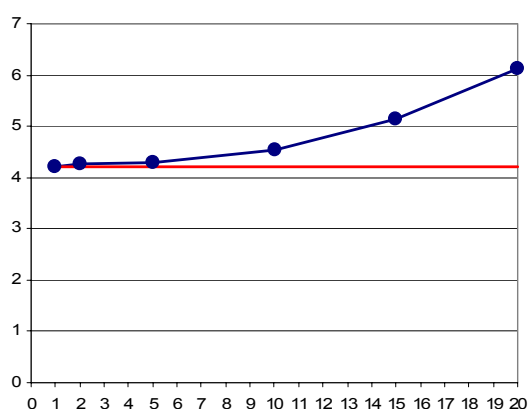


Figura IV. 99. Evolución valores medios MSK con eliminación de vértices. MTA

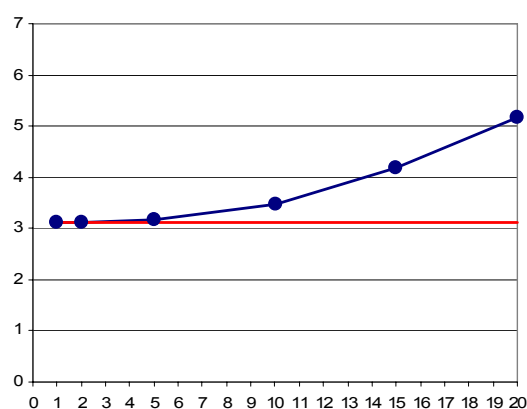


Figura IV. 100. Evolución valores medios MSK con eliminación de vértices. BCN

Los resultados (Figura IV. 99 y Figura IV. 100) presentan valores medios similares hasta el valor 5 aproximadamente. A partir de este punto las curvas empiezan a separarse. Al igual que en el método MDH, la curva del MTA tiende a separarse con anterioridad a la BCN. A partir de los gráficos mostrados, se puede considerar similitud en los resultados para este método con un límite mínimo de mantenimiento de puntos de la base de datos de control de 1 de cada 5 (20% de los puntos).

IV.6.3.3 Método MOS

Los resultados obtenidos con el método MOS permiten estudiar la similitud de sus funciones de distribución (Figura IV. 101 y Figura IV. 102) a través de la técnica expuesta previamente del Test de Kolmogorov-Smirnov. Por tanto, se plantea el estudio de resultados de varias bases de control en la que para cada una se han eliminados vértices incrementalmente.

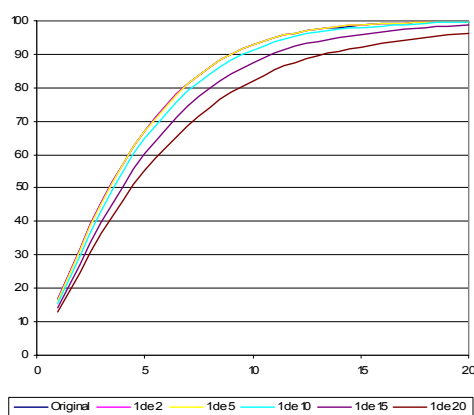


Figura IV. 101. Funciones de distribución MOS con eliminación vértices de control. MTA

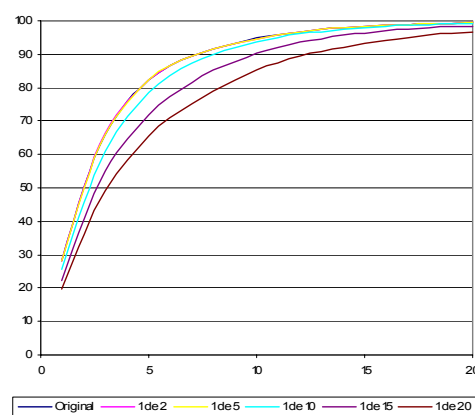


Figura IV. 102. Funciones de distribución MOS con eliminación vértices de control. BCN

Los resultados de la prueba de Kolmogorov-Smirnov se muestran en la Tabla IV. 47 y en la Tabla IV. 48. Los sombreados se establecen en función del nivel de confianza siguiendo los criterios presentados en el punto IV.6.1.

	Original	1 de cada 2	1 de cada 5	1 de cada 10	1 de cada 15	1 de cada 20
Original		0.05 0.158 1	0.05 0.158 1	0.1 0.316 1	0.25 0.791 0.497	0.45 1.423 0.002
1 de cada 2			0.05 0.158 1	0.15 0.474 0.965	0.3 0.949 0.275	0.45 1.423 0.002
1 de cada 5				0.1 0.316 1	0.25 0.791 0.497	0.45 1.423 0.002
1 de cada 10					0.2 0.632 0.771	0.4 1.265 0.05
1 de cada 15						0.25 0.791 0.497

Tabla IV. 47. Resultados similitud con MOS en eliminación de vértices de control para el MTA

	Original	1 de cada 2	1 de cada 5	1 de cada 10	1 de cada 15	1 de cada 20
Original		0.05 0.158 1	0.05 0.158 1	0.05 0.158 1	0.25 0.791 0.497	0.45 1.423 0.002
1 de cada 2			0.05 0.158 1	0.05 0.158 1	0.25 0.791 0.497	0.45 1.423 0.002
1 de cada 5				0.05 0.158 1	0.25 0.791 0.497	0.45 1.423 0.002
1 de cada 10					0.2 0.632 0.771	0.4 1.265 0.05
1 de cada 15						0.25 0.791 0.497

Tabla IV. 48. Resultados similitud con MOS en eliminación de vértices de control para la BCN

Tanto para el MTA como para la BCN se encuentra similitud entre los resultados obtenidos por la base de elementos de control en la que se mantiene 1 de cada 2, de cada 5 y de cada 10 vértices con respecto a la base de datos original.

IV.6.3.4 Método MOD

En cuanto al método MOD los resultados se analizan sobre la superficie común de los orlados. Las funciones de distribución obtenidas para cada base de datos modificada se presentan en la Figura IV. 103 y en la Figura IV. 104.

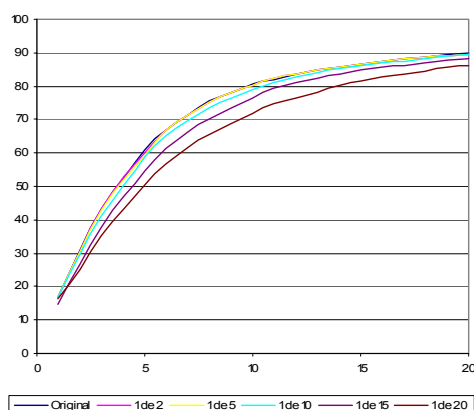


Figura IV. 103. Funciones de distribución MOD(SC) con eliminación vértices de control. MTA

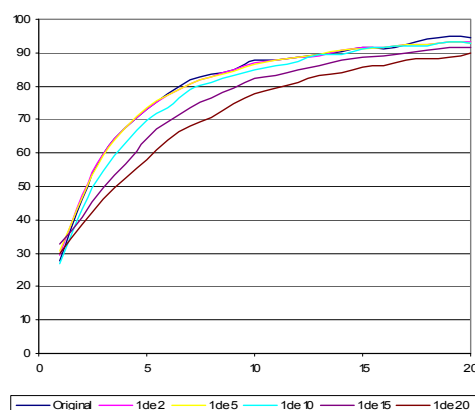


Figura IV. 104. Funciones de distribución MOD(SC) con eliminación vértices de control. BCN

Los resultados de la prueba de Kolmogorov-Smirnov se presentan en la Tabla IV. 49 y en la Tabla IV. 50.

	Original	1 de cada 2	1 de cada 5	1 de cada 10	1 de cada 15	1 de cada 20
Original		0.05	0.05	0.05	0.15	0.3
		0.159	0.159	0.159	0.474	0.949
		1	1	1	0.965	0.275
1 de cada 2			0.05	0.05	0.15	0.3
			0.159	0.159	0.474	0.949
			1	1	0.965	0.275
1 de cada 5				0.05	0.15	0.3
				0.159	0.474	0.949
				1	0.965	0.275
1 de cada 10					0.15	0.3
					0.474	0.949
					0.965	0.275
1 de cada 15						0.2
						0.632
						0.771

Tabla IV. 49. Resultados similitud con MOD en eliminación de vértices de control para el MTA

	Original	1 de cada 2	1 de cada 5	1 de cada 10	1 de cada 15	1 de cada 20
Original		0.15	0.15	0.15	0.25	0.4
		0.474	0.474	0.474	0.790	1.265
		0.965	0.965	0.965	0.497	0.059
1 de cada 2			0.1	0.1	0.2	0.35
			0.316	0.316	0.632	1.107
			1	1	0.771	0.135
1 de cada 5				0.1	0.25	0.35
				0.316	0.790	1.107
				1	0.497	0.135
1 de cada 10					0.25	0.35
					0.790	1.107
					0.497	0.135
1 de cada 15						0.2
						0.632
						0.771

Tabla IV. 50. Resultados similitud con MOD en eliminación de vértices de control para la BCN

Tanto para el MTA como para la BCN pueden considerarse similaridad entre las funciones de distribución de las bases de datos en las que se mantiene 1 de cada 2, de cada 5 y de cada 10 vértices con respecto a la base de datos original, aunque en el caso del MTA también se podría abarcar hasta 1 de cada 15.

IV.6.3.5 Interpretación y análisis

A lo largo de este punto se ha presentado para cada método de control posicional un estudio que permite estimar el límite de eliminación de vértices en la base de datos de control para que los resultados obtenidos sean similares a los originales.

Para los métodos MDH y MSK se ha realizado un análisis de similaridad en función de la representación gráfica obtenida. En estas representaciones se observa que las curvas empiezan a despegarse a partir de mantener entre 1 de cada 5 vértices y 1 de cada 10.

En cambio, con los métodos MOS y MOD se ha aplicado la prueba de similaridad de Kolmogorov-Smirnov para determinar si las funciones de distribución eran similares a la obtenida con la base de datos original. En estos casos, se ha podido avanzar algo más obteniéndose, tanto para la BCN, como para el MTA el límite mínimo de mantenimiento de 1 de cada 10 vértices.

Establecer una recomendación de cadencia de número de vértices es complejo debido a que cada producto cartográfico y método (MDH, MSK...) tienen un comportamiento distinto; aunque presenten evoluciones parecidas. También habrá que tener en cuenta la procedencia de la base de control. En principio, si se obtienen de una observación cinemática con GPS, lo lógico es aprovechar todo el potencial de toma que esta instrumentación permita, para realizar un filtrado posterior. En caso necesario, por tener que observar a más velocidad (caso de autopistas o autovías), se puede aumentar con garantía según los resultados obtenidos previamente o realizar una prueba piloto semejante al trabajo expuesto en este punto para determinar la velocidad máxima a la que podría ir el vehículo sin que se desvirtuaran los resultados.

IV.7. Conclusiones

En el capítulo IV se han presentado los resultados de la presente tesis doctoral. Como resultado fundamental hay que destacar la aplicación de la metodología general propuesta de control de calidad posicional con elementos lineales. Pero los resultados aportados no se limitan a este control. Además, se ha mostrado cómo los métodos se comportan con líneas simples en las que se fuerzan ciertos errores básicos o se ha aplicado una metodología basada en la simulación para determinar el tamaño de la muestra, así como la determinación del número de vértices de control mínimo necesarios para definir las líneas de control.

Gran parte de estos resultados se han basado en la utilización de dos bases de datos de elementos lineales oficiales en el territorio andaluz, y en una base de datos de control obtenida a partir de una observación cinemática GPS sobre vías de comunicación.

En cuanto a la edición y depuración de las bases de datos se ha mostrado como un elemento importante dentro del proceso. Los datos utilizados han sido caracterizados, para estudiar sus propiedades mostrando su gran extensión espacial, así como otras vicisitudes tanto orográficas como de altitud media y tipología de vías

de comunicación. La caracterización realizada permite vincular las medidas implementadas con fines concretos.

En relación a la metodología de control posicional, se han obtenido resultados de la aplicación de los cuatro métodos propuestos, aplicando todas las métricas propuestas tanto las originales como las ampliadas. Los métodos aplicados son el Método de las Distancias de Hausdorff, el Método de Skidmore, el Método del Orlado Simple y el Método del Orlado Doble.

La aplicación conjunta de estos métodos a un mismo conjunto de datos ha permitido el análisis conjunto de los resultados sobre una misma casuística base. Los métodos aplicados son opciones funcionalmente distintas, no intercambiables, pero si pueden ser tratados de forma conjunta de manera que se complementen.

Los resultados comparados de todos los métodos arrojan superior exactitud posicional en la base de datos BCN sobre el MTA, pero, sin embargo, el MTA se comporta mejor en cuanto a los errores extremos que la BCN.

Los métodos de control posicional por elementos lineales han sido ensayados con líneas simples a las que se les había simulado un error de tipo aleatorio o sistemático, obteniendo importantes resultados del comportamiento de cada error con cada método.

Los resultados obtenidos han sido obtenidos a través de una herramienta informática desarrollada en el marco de esta tesis doctoral. Sin esta aplicación no hubiera sido posible realizar la aplicación de la metodología propuesta, máxime si se tiene en cuenta el gran volumen de datos utilizado. La mayor parte de las fases de esta metodología se han realizado de forma automática a través de la misma.

Mediante métodos de simulación y bootstrapp se ha derivado la evolución de las estimaciones, y sus precisiones, para tamaños de muestras entre 0 y 200 kms para todos los métodos analizados, obteniéndose un valor aproximado de 150 a 200 kms como tamaño de muestra mínimo. Este valor ha sido ensayado mediante la selección de varios itinerarios posibles para la toma en campo obteniendo, en todos los casos, valores aceptables que apoyan este tamaño de muestra. En cualquier caso, esta longitud puede ser considerada como un punto de partida a la hora de establecer el tamaño de la muestra, justificándose según cada caso, la realización de una prueba piloto similar a la aquí realizada.

También se ha analizado la posible influencia de diversos factores en la determinación de la muestra, como la longitud del tramo, la orientación, la tipología de la vía o la sinuosidad mediante la comparación de las funciones de distribución del error demostrando que ninguna de estas categorías presenta comportamientos sustancialmente diferentes entre unas y otras o del total poblacional.

Los resultados obtenidos demuestran que se dispone de un método operativo de trabajo que supone una alternativa viable, bien definida y competitiva.

Capítulo V.

Conclusiones

Capítulo V. Conclusiones

En este capítulo se presentan las principales conclusiones de la investigación realizada en la tesis doctoral. Para comenzar, se analiza el grado de cumplimiento de los objetivos planteados sobre la hipótesis de partida, siguiendo con la exposición de las conclusiones derivadas de los resultados obtenidos, tanto a nivel metodológico, como los procedentes de su aplicación, para finalizar con el planteamiento de las posibles perspectivas de futuro que se abren dentro de esta línea de investigación como consecuencia de esta tesis doctoral.

V.1. Objetivos alcanzados

Los objetivos planteados en la hipótesis de partida han obtenido un alto grado de cumplimiento, alcanzando, con éxito, el objetivo propuesto consistente en la realización de una metodología de control de calidad posicional en cartografía por elementos lineales.

Más concretamente, se puede indicar que se han alcanzado los siguientes objetivos:

- Análisis y desarrollo de medidas evaluadoras de la exactitud posicional en cartografía por elementos lineales, reforzando los indicadores existentes con nuevas métricas.
- Desarrollo de una herramienta para la obtención de la exactitud posicional en cartografía para su aplicación a datos vectoriales a partir de elementos lineales de control obtenidos con fuentes de mayor exactitud.
- Comprobación de los resultados obtenidos, para lo que se ha aplicado el método a varias bases de datos cartográficas.
- Obtención y comparación de valores globales de exactitud posicional en cartografía a partir de los resultados obtenidos con cada método de control.

V.2. Conclusiones acerca de los resultados obtenidos

En este punto, se exponen las conclusiones que se pueden extraer de los resultados obtenidos, ampliamente descritos en los capítulos anteriores, dividiendo su descripción en aspectos metodológicos y su aplicación.

V.2.1. Metodología

El planteamiento metodológico realizado en la presente tesis doctoral constituye un resultado a la vez que una importante innovación, ya que la metodología adoptada permite la realización de un control posicional mediante el uso de gran parte de las medidas evaluadoras por elementos lineales existentes actualmente, integrándolas y ampliándolas en un mismo sistema de control.

Los métodos actuales de control, que vieron la luz de forma individual, en esta tesis doctoral han sido estudiados conjuntamente, comparándolos y analizando los resultados obtenidos de forma conjunta, ya que, como quedó patente en esta investigación, la información aportada por cada uno resulta complementaria.

Dicho control se plantea mediante una aplicación informática desarrollada para este propósito que compara las bases de datos a controlar con elementos homólogos obtenidos a partir de observaciones GPS de vías de comunicación (carreteras) de forma automática.

El uso de elementos lineales para el control posicional en cartografía se ha mostrado con éxito en esta investigación, presentando gran potencial debido por una parte, a que la fuente de datos formada por estos elementos es mayoritaria en las bases cartográficas, y por otra, a la posibilidad novedosa de implementar el control a partir de elementos obtenidos con fuentes de mayor precisión mediante observaciones cinemáticas GPS de forma rápida y sencilla. Los rendimientos obtenidos mediante esta técnica constituyen una mejora respecto a los aportados por otras técnicas más convencionales. El uso de este tipo de elementos queda, a su vez, plenamente justificado por la gran distribución espacial de los mismos y la facilidad de observación de estas técnicas GPS.

El procedimiento de control posicional propuesto abarca desde la obtención de las bases de datos de control y a controlar hasta la implementación de los métodos de control a utilizar y la obtención de resultados, lo que también le constituye como característica el ser un sistema integrado de tareas.

Los métodos utilizados para el control posicional utilizados son: El Método de las Distancias de Hausdorff (MDH), el Método de Skidmore o Banda Épsilon (MSK), el Método del Orlado Simple (MOS) y el Método del Orlado Doble (MOD).

Los resultados generados con estos métodos vienen expresados en forma de valores escalares (MDH y MSK) y funciones de distribución (MOS y MOD). La comparación entre los primeros es inmediata, sin embargo, las funciones de distribución han de ser comparadas mediante el análisis de su similitud. Con este fin, se utiliza la prueba de Kolmogorov-Smirnov, que indica el grado de similitud entre dos funciones de distribución.

La prueba de Kolmogorov-Smirnov permite realizar estudios sobre la muestra de elementos lineales a utilizar a partir del análisis de la similitud entre distribuciones y el uso de herramientas de simulación. Este procedimiento de determinación de la muestra no ha sido realizado con este tipo de elementos hasta la fecha, por lo que constituye una importante innovación de esta tesis doctoral y un punto de partida para posteriores estudios que utilicen elementos lineales.

V.2.2. Aplicación metodológica

Como se indica con anterioridad, la aplicación de la metodología se ha realizado principalmente a través de una herramienta informática desarrollada con esta finalidad (CPLin). Esta aplicación supone una importante innovación dentro del control posicional cartográfico a través de elementos lineales, ya que hasta el momento, no existía ninguna herramienta diseñada y optimizada exclusivamente para este control. CPLin realiza todas las tareas establecidas en el procedimiento de control propuesto de forma prácticamente automática siendo transparente al investigador en todos los procesos, lo que permite la comprobación de los resultados obtenidos en cada fase.

Las bases de datos utilizadas para el ensayo de la metodología han constituido un gran banco de pruebas, tanto por su tamaño (superando los 1200 kms de carreteras), como por la importancia de las mismas, ya que se han utilizando las 2 bases de datos a nivel nacional y regional más representativas (Base Cartográfica Numérica 1:25.000 y Mapa Topográfico de Andalucía 1:10.000).

Los métodos de control presentados individualmente por los investigadores hasta la fecha fueron probados originalmente sobre pequeñas bases de datos, por lo que el ensayo realizado en esta tesis doctoral supone un gran refuerzo de su aplicabilidad, obteniendo resultados positivos en todos los casos. El uso de esta gran cantidad de elementos también ha permitido caracterizarlos según una serie de criterios y estudiar como se comportaba cada caso.

Los resultados de la aplicación a dos bases de datos reales suponen en sí mismos una comprobación de la viabilidad de la metodología propuesta, ya que el ensayo de los distintos métodos ha permitido establecer una comparación de los mismos, propiciando su concepción como resultados complementarios. Los métodos

propuestos son opciones funcionalmente distintas pero pueden ser tratados de forma conjunta de forma que se complementen.

Del estudio sobre las dos bases de datos reales se extrae que la BCN posee una mayor exactitud posicional que el MTA, aunque éste tiene un mejor comportamiento en cuanto a los errores extremos que la BCN. En general, los resultados aportados por los distintos métodos muestran valores de exactitud posicional esperados (Tabla IV. 33) para cada base de datos en función de sus características.

A partir de los resultados obtenidos con el control posicional, se han realizado una serie de simulaciones para la determinación de la muestra:

- Se ha analizado la posible influencia de diversos factores en la determinación de muestra (tipo de vía, zona, longitud, orientación o sinuosidad) no encontrando comportamientos sustancialmente diferentes que aconsejen tener en cuenta alguno de estos factores.
- Se ha encontrado y comprobado un tamaño de muestra mínimo recomendable en estas bases de datos en torno a 150 kms mediante procesos de simulación.
- Se ha cuantificado la densidad de vértices de control necesaria para cada base de datos estudiada con el fin de establecer el tamaño de la base de datos de control y derivar de este valor el intervalo de toma de puntos o la velocidad del vehículo durante la observación, lo que implica una reducción del tamaño de esta base de datos y una mejora en la seguridad vial de la observación GPS.

El estudio de la similaridad de las funciones de distribución a través de la prueba de Kolmogorov-Smirnov ha demostrado su eficacia como herramienta en estos estudios de determinación de la muestra.

El estudio de líneas sintéticas con errores simulados ha permitido conocer cómo se comportan los resultados de cada método de control para cada uno de estos desplazamientos simulados, de forma que, los resultados obtenidos con líneas más complejas (casos reales), pueden ser interpretados para cada método de control a partir de la experiencia aportada en el estudio de estas líneas sintéticas.

V.3. Perspectivas de futuro

El trabajo desarrollado en la presente tesis doctoral puede y debe tener continuidad en el futuro, ya que existe actualmente una importante inquietud en la comunidad científica en este tema. Por otro lado, es importante dar continuidad a las aportaciones que esta tesis plantea a esta línea de investigación.

Para ello, se proponen las siguientes líneas de trabajo:

- Implementación de nuevas medidas evaluadoras de control por elementos lineales e incorporación de dichas medidas al software de control desarrollado.

- Mejora y optimización de la herramienta CPLin para su distribución a la comunidad científica u otros usuarios, de forma que ésta pueda ser mejorada a partir de las aportaciones que puedan realizarse.
- Aplicación del control a nuevas bases de datos geográficas de diferentes zonas y escalas para comparar los resultados de las mismas con los obtenidos en la presente tesis doctoral.
- Implementación de la metodología a otra tipología de elementos lineales, como ferrocarriles, caminos o hidrografía, de forma que puedan caracterizarse y compararse con los resultados procedentes de las carreteras.
- Análisis de la componente temporal del dato geográfico a partir de los resultados obtenidos mediante el control posicional por elementos lineales.
- Desarrollo y posible adaptación de las medidas evaluadoras a la componente vertical de las bases de datos geográficas. Estudio de cómo pueden interactuar las exactitudes de las componentes horizontal y vertical por elementos lineales.

Referencias

- [1] **ABBAS, I.; GRUSSENMEYER, P.; HOTTIER, P. (1995).** *Contrôle de la planimétrie d'une base de données vectorielles : une nouvelle méthode basée sur la distance de Hausdorff : la méthode du contrôle linéaire.* Bul. S.F.P.T., Nº 137, pp. 6-11.
- [2] **ALESHEIKH, A.; BLAIS, J.; CHAPMAN, M.; KARIMI, H. (1999).** *Rigorous Geospatial Data Uncertainty Models for GISs.*, en *Spatial Accuracy Assessment: Land Information Uncertainty in Natural Resources*. Ed. Ann Arbor Press. Chelsea – Michigan (EE.UU.), capítulo 24, pp. 195-202.
- [3] **ARIZA, F. J. (2002a).** *Calidad en la Producción Cartográfica*, ed Ra-Ma. Madrid.
- [4] **ARIZA, F. J. (2002b).** *Modelización, Simulación y Toma de Decisiones con Sistemas de Información Geográfica.* Curso de Doctorado, E.P.S. Universidad de Jaén, Jaén.
- [5] **ARIZA, F. J.; GARCÍA, J. L. (2004).** *Segmentación de elementos lineales mediante filtrado de Douglas-Peucker aplicado a las áreas efectivas de Visvalingam-Whyhatt.* Revista Mapping. Octubre 2004.
- [6] **ARIZA, F. J.; MOZAS, A. T. (2005).** *Methodology for the development of a sampling recommendation when using linear elements for positional quality control.* Proceedings of XXII International Cartography Conference. La Coruña. España. 2005.
- [7] **ARIZA, F. J.; GARCÍA, J. L.; AMOR, R. (2004).** *Casos prácticos de calidad en la producción cartográfica (1ª edición. Diciembre 2004).* Publicaciones de la Universidad de Jaén. Jaén.
- [8] **ASCI, (1983).** *Map Uses, scales and accuracies for engineering and associated proposes.* American Society of Civil Engineers, Committee on Cartographic Surveying, Surveying and Mapping Division, New York.
- [9] **ASPRS (1989).** *Accuracy standard for large scale maps.* PE&RS. vol. 56, nº 7.
- [10] **ATKINSON, A. D. J.; ARIZA, F. J. (2002).** *Nuevo enfoque para el análisis de la calidad posicional en cartografía mediante estudios basados en la geometría lineal.* XIV Congreso Internacional de Ingeniería Gráfica. Santander, España, 5-7 de junio.
- [11] **ATKINSON, A. D. J. (2005).** *Control de calidad posicional en cartografía: Análisis de los principales estándares y propuesta de mejora.* Tesis doctoral. Universidad de Jaén.
- [12] **BLAKEMORE, M. (1983).** *Generalisation and Error in Spatial Data Bases.* Cartographica, vol. 21, nº 2/3.
- [13] **CANAVOS, G. (2003).** *Probabilidad y estadística: Aplicaciones y Métodos.* Ed. McGraw-Hill. Madrid.

- [14] **CASPARY, W.; SCHEURING, R. (1993).** *Positional accuracy in spatial databases.* Computer, Environment and Urban Systems, vol. 17, pp. 103-110.
- [15] **CHURCH, R.; CURTIN, K.; FOHL, P.; FUNK, C.; GOODCHILD, M.; KYRIAKIDIS, P.; NORONHA, V. (1998).** *Positional distortion in geographic data sets as a barrier to interoperability.* Technical Papers ACSM. American Congress on Surveying and Mapping. Bethesda.
- [16] **CHEUNG, C. K.; SHI, W. (2004).** *Estimation of the positional uncertainty in line simplification in GIS.* The Cartographic Journal. The British Cartographic Society. Vol. 41, nº 1, pp. 37-45. Junio 2004.
- [17] **CLARKE, K. (1996).** *Words From the Chair.* Cartography Specialty Group Newsletters. Association American Geographers. Vol. 16, nº 2, 1-2.
- [18] **COCHRAN, W.G. (1977).** *Sampling Techniques.* John Willey & Sons. Nueva York.
- [19] **CORRAL, V.; GARCÍA, A. (2000).** *El Mapa Topográfico de Andalucía 1:10.000.* Mapping, Nº 59, pp: 14-15.
- [20] **CORRAL, V.; MARTÍNEZ, M.; SÁCHEZ F. J.; TORRECILLAS, C. (2001).** *El Mapa Digital de Andalucía 1:10.000 del Instituto de Cartografía de Andalucía.* Mapping. Nº 67, pp: 6-16.
- [21] **CSG (1985).** *Normas cartográficas para la ejecución del Mapa Topográfico Nacional escalas E25k y E50k.* Ministerio de Obras Públicas y Transportes, Instituto Geográfico Nacional. Consejo Superior Geográfico. Madrid.
- [22] **CSG (1992).** *Normas y Condiciones Técnicas para los Levantamientos Cartográficos a Grandes Escalas.* Ministerio de Obras Públicas y Transportes, Instituto Geográfico Nacional. Consejo Superior Geográfico. Madrid.
- [23] **CUENIN, R. (1972).** *Cartographie Generale. Tome I. Notions Générales et principes d'élaborations.* Ed. Eyrolles. Paris.
- [24] **DOD (1990).** *Military Standard. Mapping, Charting & Geodesy Accuracy. MIL-STD 600001.* Defense Mapping Agency. Department Of Defense.
- [25] **FGDC (1998).** *Geospatial Positioning Accuracy Standards, National Standard for Spatial Data Accuracy (FGDC-STD-007.3)*
- [26] **GARCIA, J. L. (2006)** *Automatización de los procesos de segmentación y clasificación de vías de comunicación en generalización cartográfica.* Tesis doctoral. Universidad de Jaén.
- [27] **GIBBONS, J. D.; CHAKRABORTI, S. (1992).** *Nonparametric Statistical Inference.* 3ª edición. Ed. Marcel Dekker, New York.

- [28] **GIORDANO, A.; VEREGIN, H. (1994).** *Il controllo di qualità nei sistema informativi territoriali*. Ed. El Cardo, Venecia, Italia.
- [29] **GOEMAN, W. ; MARTINEZ-FONTE, L. ; BELLENS, R. ; GAUTAMA, S. (2005a).** *Automated verification of road network data by VHR satellite images using road statistics*. ISPRS Hannover Workshop 2005: High Resolution Earth Imaging for Geospatial Information Proceedings.
- [30] **GOEMAN, W. ; MARTINEZ-FONTE, L. ; BELLENS, R. ; GAUTAMA, S. (2005b).** *Using image statistics for automated quality assessment of urban geospatial data*. ISPRS 3rd International Symposium Remote Sensing and Data Fusion Over Urban Areas. Tempe 14-16 March 2005.
- [31] **GONZÁLEZ-MATESANZ, F. J.; DALDA-MOURÓN, A. (2003).** *Modelos de transformación entre ED50 y ETRS89*. Ed. Instituto Geográfico Nacional. Madrid.
- [32] **GONZÁLEZ-MATESANZ, F. J.; QUIROS, J.; CANO, M.; SANCHEZ SOBRINO, J.; DALDA-MOURÓN, A. (2004).** *El cambio de datum ED50 – ETRS89. Métodos y resultados*. VIII Congreso Nacional de Topografía y Cartografía TOPCART 2004. Madrid.
- [33] **GOODCHILD, M. F.; GOPAL, S. (1994).** *The Accuracy of Spatial Databases*. Ed. Taylor & Francis, London.
- [34] **GOODCHILD, M. F.; HUNTER, G. J. (1997).** *A simple positional accuracy for linear features*. Int. Journal Geographical Information Science, Vol. 11, nº 3, pp. 299-306.
- [35] **GUSTAFSON, G.; LOON, J. (1982).** *Contour Accuracy and National Map Accuracy Standards*. Surveing and Mapping. Vol. 42, nº 4.
- [36] **GUTIÉRREZ, J. L.; SAMPAYO, A. (1981).** Serie Manuales, nº 2. Centro de Estudios de Ordenación del Territorio y Medio Ambiente. Madrid.
- [37] **JOHNSTON, D. M.; TIMLIN, D. M.; SZAFONI, D. L.; CASANOVA, J. J. (1999).** *Quality Assurance/ Quality Control Procedures for ITAM GIS Databases*. Geographic Modeling Systems Lab. University of Illinois. <http://www.gis.uiuc.edu>.
- [38] **KAGAWA, Y.; SEKIMOTO, Y.; SHIBASKI, R. (1999).** *Comparative study of positional accuracy evaluation of line data*. 20th Asian Conference on Remote Sensing (ACRS). Hong Kong, China. Poster session 4.
- [39] **KUMAR, N. (2000).** *Democratization and Automation in Cartography*. The Cartographic Journal, 37(1), 65-77.
- [40] **LEUNG, Y.; YAN, J. (1998).** *A locational error model for spatial features*. Int. Journal Geographical Information Science, Vol. 12, nº 6, pp. 607-620.

- [41] **LÓPEZ-CUERVO, S. (1980).** *Fotogrametría*. Ed. Egraf. Madrid.
- [42] **LÓPEZ-CUERVO, S. (1996).** *Topografía (2ª Edición)*. Ed. Mundi Prensa. Madrid.
- [43] **MALING, D. (1989).** *Measurements from Maps*. Pergamon Press. Oxford.
- [44] **MAPPING (2007).** *Número especial: Las Normas ISO en la Cartografía*. Mapping, nº 117.
- [45] **MAS, J. F. (2006).** *Reducing positional error in spatio-temporal analyses*. Proceedings of Accuracy 2006: 7th Internacional Symposium on Spatial Accuracy Assessment in Natural Resources and Environment Sciences. Instituto Geográfico Português. Lisboa (Portugal), pp. 281-285.
- [46] **MAS, S. (2002).** *Propuesta para el establecimiento de la Infraestructura de Datos Espaciales de España (IDEE)*. Mapping nº 77, abril 2002.
- [47] **MERCHANT, D. (1987).** *A Spatial Accuracy Specification for Large Scale Topographic Maps*. PE&RS. vol. 53.
- [48] **MORRISON, J. (1995).** *Spatial data quality. Elements of spatial data quality*. Edited by Stephen C. Guptill and Joel L. Morrison. International Cartographic Association. Pergamon Press, Oxford. pp. 1-12.
- [49] **MOZAS, A. T. (2003).** *Memoria del periodo de investigación tutelado*. Universidad de Jaén. Departamento de Ingeniería Cartográfica, Geodésica y Fotogrametría.
- [50] **MOZAS, A. T.; UREÑA, M. A.; ARIZA, F. J. (2007a).** *CPLin: Una herramienta para el control posicional de la cartografía mediante elementos lineales*. Revista Mapping nº 116, pp. 81-87. Febrero 2007.
- [51] **MOZAS, A. T.; UREÑA, M. A.; ARIZA, F. J. (2007b).** *Bases computacionales para el desarrollo de herramientas de control posicional de la cartografía mediante elementos lineales*. Revista Mapping nº 116, pp. 22-29. Febrero 2007.
- [52] **NCDCDS (1989).** *Spatial data transfer Standard (SDTS)*. Nacional Comitee for Digital Cartographic Data Standars.
- [53] **NERO, M.A. (2005).** *Propuestas para el control de calidad de bases cartográficas con énfasis en la componente posicional*. Tesis Doctoral, Escuela Politécnica de la USP, Sao Paulo.
- [54] **PÉREZ, C. (1999).** *Técnicas de Muestreo Estadístico*. Ed. Ra-Ma. Madrid.
- [55] **PERKAL, J. (1956).** *On epsilon length*. Bulletin de l'Academie Polonaise des Sciences, 4. pp. 399-403.
- [56] **PLAZANET, C. (1996).** *Enrichissement des bases de données géographiques: analyse de la géométrie des objets lineaires pour la généralisation*

cartographique (application aux routes). Ph. D. Dissertation. Director: Jacques Désarménien. France: Université de Marne la Vallée.

- [57] **PRESS, W. H.; TEUKOLSKY, S. A. (1992).** *Numerical Recipes in C: the art of scientific computing*. Ed. Cambridge University Press.
- [58] **RAMÍREZ, J. R.; BIDOSHI, P. D. K.; DOUGLAS, T.; PHUYAL, B.; SZAKAS, J. (1996).** *A Quality Assessment of Spatial Data Acquired for the Rickenbacker Air National Guard Base Conversion Task*. The Ohio State University Center for Mapping.
- [59] **RÍOS, D.; RÍOS, S. (1997).** *Simulación, Métodos y aplicaciones*. Ed. Ra-Ma. Madrid.
- [60] **RUIZ, J. J.; MOZAS, A. T.; ARIZA, F. J. (2006).** *Metodología de adquisición GPS de elementos lineales para su uso en el control de la calidad posicional en cartografía*. XVIII Congreso Internacional de Ingeniería Gráfica. Barcelona, España, 1-2 de junio.
- [61] **RUIZ, J. J.; MOZAS, A. T.; ARIZA, F. J.; ATKINSON, A. DJ. (2006).** *Sample design criteria and GPS acquisition of linear elements for use in positional quality control in cartography*. 5th. European Congress on Regional Geoscientific Cartography and Information Systems. Barcelona. España. Junio 2006.
- [62] **RUIZ, J. J.; MOZAS, A. T.; ARIZA, F. J. (2007).** *Muestreo, observación GPS y rendimientos obtenidos en la adquisición de vías de comunicación para el control de calidad posicional de la cartografía mediante elementos lineales*. Revista Topografía y Cartografía, vol. 141. pp. 48-57.
- [63] **SADAHIRO, Y. (2005).** *Buffer operation on spatial data with limited accuracy*. Transactions in GIS. Vol. 9(3), pp. 323-344.
- [64] **SCHMIDLEY, R.W. (1996).** *Framework for the control of quality in automated mapping*. Dissertation, Ohio State University.
- [65] **SEVILLA, M. J. (1991).** *Criterios de precisión cartográfica*. Revista Catastro, año III, nº 8.
- [66] **SHI, W.; LIU, W. (2000).** *A stochastic process-based model for the positional error of line segments in GIS*. International Journal Geographical Information Science. Vol. 14, nº 1, pp. 51-66.
- [67] **SHI, W.; CHEUNG, C. K. (2000).** *Measure Error Propagation for Buffer Spatial Analysis in GIS using Gaussian Quadrature Numerical Method*. Accuracy 2000: Proceedings of the 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences. Ed. Delft University Press. Amsterdam (Holanda), pp. 569-576.

- [68] **SHI, W.; CHEUNG, C. K.; ZHU, C. (2003).** *Modelling error propagation in vector-based buffer analysis*. International Journal Geographical Information Science. Vol. 17, nº 3, pp. 251-271.
- [69] **SKIDMORE, A. K.; TURNER, B. J. (1992).** *Map Accuracy Assessment Using Line Intersect Sampling*. Photogrammetric Engineering and Remote Sensing, 58 (10): 1453-1457.
- [70] **SKOPELITI, A.; TSOULOS, L. (2000a).** *Estimating positional accuracy of linear features*. Athens 2000 Workshop 4-7 "Spatial Information Management - Experiences and Visions for the 21st Century". International Federation of Surveyors. Session 5 III, nº 8. Octubre 2000.
- [71] **TÖPFER, F.; PILLEWIZER, W. (1966).** *The Principles of Selection*. The Cartographic Journal. Vol. 3, nº 1.
- [72] **TSOULOS, L.; SKOPELITI, A. (2000b).** *Developing a Model for Quality Assessment of Linear Features*. Accuracy 2000: Proceedings of the 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences. Ed. Delft University Press. Amsterdam (Holanda), pp. 655-658.
- [73] **TVEITE, H.; LANGAAS, S. (1995).** *Proceedings from the 5th Scandinavian Research Conference on GIS*. 12-14 June 1995. Trondheim (Norway). J. T. Bjørke ed., pp. 145-154.
- [74] **TVEITE, H.; LANGAAS, S. (1999).** *An accuracy assessment method for geographical line data sets based on buffering*. Int. Journal Geographical Information Science. Vol. 13, nº 1, pp. 27-47.
- [75] **USBB (1947).** *United States National Map Accuracy Standards*. U.S. Bureau of the Budget.
- [76] **USGS (1994).** *Standards for Digital Line Graphs*. National Mapping Program. Technical Instructions. USGS.
- [77] **VAN NIEL, T. G.; McVICAR, T. R. (2002).** *Experimental evaluation of positional accuracy estimates from a linear network using point- and line-based testing methods*. International Journal Geographical Information Science. Vol. 16, nº 5, pp. 455-473.
- [78] **VAUGLIN, F. (1999).** *Use of Variograms to Represent Spatial Uncertainty of Geographic Linear Features*, en *Spatial Accuracy Assessment: Land Information Uncertainty in Natural Resources*. Ed. Ann Arbor Press. Chelsea – Michigan (EE.UU.), capítulo 19, pp. 157-163.
- [79] **VEREGIN H. (1989a).** *A Taxonomy of Errors in Spatial Data Bases*. Technical Paper 89-12, NCGIA, Santa Barbara. 115 p.

- [80] **VEREGIN, H. (1989b).** *Accuracy of Spatial Databases: Annotated Bibliography.* NCGIA. Santa Bárbara.
- [81] **VEREGIN H. (1999).** *Data quality measurement and assessment.* NCGIA Core Curriculum. NCGIA, Santa Barbara, California.
- [82] **VEREGIN H. (2000).** *Quantifying positional error induced by line simplification.* Int. Journal Geographical Information Science, Vol. 4, Nº 2, pp. 113-130.
- [83] **WIKIPEDIA.ORG (2001-07).** Wikimedia Foundation. <http://es.wikipedia.org>.
- [84] **ZHANG, J.; GOODCHILD, M. (2002).** *Uncertainty in Geographical Information.* Ed. Taylor & Francis. London, New York.

Anexos

Anexo I

Test de exactitud posicional basado en puntos

Al.1. Introducción

En este anexo se presentan algunos de los test de control posicional basados en puntos más utilizados por los productores de cartografía a nivel mundial.

Para esquematizar cada uno de los test que se describen se utiliza la estructura desarrollada por Giordano y Veregin (1994):

- Aplicaciones corrientes: Entidades cartográficas que utilizan el test corrientemente.
- Método de comparación: Indica la metodología utilizada para valorar la exactitud posicional (con respecto a otra fuente).
- Componente posicional: Especifica si el test se aplica a la componente vertical u horizontal.
- Clases de elementos: Elementos puntuales, lineales o superficiales.
- Correspondencia con un estándar de exactitud predefinida: Indica si el test se ha construido de modo que se pueda evaluar la exactitud

posicional de un grupo de elementos con un límite aceptable definido por el usuario.

- Descripción: Presentación del procedimiento.
- Procedimiento: Indica paso por paso como resolver el test.
- Discusión: Se evalúan las ventajas e inconvenientes de este procedimiento.

AI.2. Test NMAS

El test NMAS (*Nacional Map Accuracy Standard*) (USBB, 1947) ha sido ampliamente utilizado por muchas instituciones y organismos cartográficos del mundo. La principal característica a destacar es su simplicidad.

Según la estructura propuesta anteriormente, las principales características de este test se resumen a continuación (Giordano y Veregin, 1994):

- Aplicaciones corrientes: El NMAS se ha utilizado en la cartografía del *United States Geological Survey* (USGS) desde 1947.
- Método de comparación: Con fuentes de mayor exactitud.
- Componente posicional: Horizontal y vertical. La componente horizontal no se divide en "X" e "Y". Se asume cierto grado de interdependencia entre la componente horizontal y vertical.
- Clase de elementos: Puntos.
- Correspondencia con un estándar de exactitud predefinido: Genera una indicación del tipo adecuado / no adecuado.
- Descripción: El test establece si el mapa analizado está dentro del límite de error predefinido para la componente horizontal y vertical del error posicional. El error se determina confrontando las coordenadas de una muestra de puntos del mapa con las coordenadas de los mismos sobre la fuente de mayor exactitud.
- Procedimiento:
 1. Selección de, al menos, 20 puntos de fácil determinación sobre el mapa. Un punto es de fácil determinación si su localización es sencilla sobre el mapa y sobre el terreno.
 2. Determinar las coordenadas del punto sobre el mapa.
 3. Determinar las coordenadas del punto sobre la fuente de mayor exactitud.
 4. Determinar si se alcanza el estándar predefinido de exactitud horizontal. Por ejemplo, para el USGS, el estándar especifica que un máximo del 10 % de los puntos seleccionados puede superar un error horizontal máximo de 0.08 centímetros para escalas superiores a 1:20.000 o de 0.05 centímetros para escala

menores a 1:20.000. El error es definido como la diferencia entre la posición del punto del mapa y la de la fuente de mayor exactitud. El error máximo está referido a unidades sobre el mapa, por lo que, bastará con multiplicar por el módulo de la escala para obtener su cuantificación sobre el terreno.

$$e_{x_i} = X_{t_i} - x_{m_i}$$

$$e_{y_i} = Y_{t_i} - y_{m_i}$$

Ec. Al. 1

5. Determinar si se alcanza el estándar predefinido de exactitud vertical. El estándar especifica que un 10% de los puntos seleccionados pueden haber cometido un error vertical superior de la mitad de la equidistancia de las curvas de nivel. El error en la componente vertical puede ser corregido variando la posición horizontal de los puntos en una magnitud máxima al error horizontal aceptable.

$$e_{z_i} = Z_{t_i} - z_{m_i}$$

Ec. Al. 2

- **Discusión:** La ventaja del test NMAS es su simplicidad: el resultado viene expresado en términos de adecuado o no adecuado, por lo que es fácilmente interpretable por el usuario.

Las desventajas de su utilización son, sin embargo, numerosas. La principal es la omisión de una medida de la inexactitud presente en el mapa ya que existe un solo indicador de adecuación o no de la misma. La diferencia del nivel de exactitud puede ser notable y no es determinable. Como consecuencia, resulta difícil estimar cuanto puede costar obtener una cartografía de mayor exactitud. Otra limitación del test consiste en el método de selección de puntos utilizado: introduciendo el requisito de fácil determinación, se introduce en las medidas un cierto grado de dependencia del mismo del observador externo. Otro aspecto discutible es la posibilidad de corregir el error vertical de un punto actuando sobre su componente horizontal (con la limitación de no superar el error horizontal permisible), lo que provoca que el estándar sea más fácilmente alcanzable en superficies de fuerte pendiente. En definitiva, el test no ofrece una medida de la inexactitud del punto individualmente.

Atkinson (2005) destaca las siguientes consideraciones:

- Es muy permisivo con cartografía a escalas mayores a 1:20.000. Independientemente del número de puntos de control, los errores han de ser muy elevados y numerosos para que sean detectados, por lo que el riesgo del productor es prácticamente nulo, mientras que el del usuario es muy alto.

- Es muy restrictivo con escalas inferiores a 1:20.000 (en las que influye mucho más aspectos como la generalización o la simbología). Así, para aceptar una base de datos geoespacial a dicha escala, deberá ser un 15% más exacta que lo estrictamente recomendable para su escala, empleándose al menos 120 puntos de control.
- Solo se tiene en consideración el porcentaje de valores (máximo del 10%) que se encuentran por encima de una determinada tolerancia en función de su escala de representación.
- Si la cartografía no cumple el estándar, es difícil estimar cuanto costaría producir una cartografía que sí cumpliera con él y viceversa.

En definitiva, Atkinson (2005) no recomienda su utilización.

AI.3. Test EMAS

El test EMAS (*Engineering Map Accuracy Standard*) fue desarrollado por la *American Society of Civil Engineers*, el *American Congress on Surveying and Mapping* y la *American Society of Photogrammetry*. Las características más representativas de este estándar son (Giordano y Veregin, 1994):

- Aplicaciones corrientes: El EMAS especifica la exactitud de los mapas topográficos a gran escala.
- Método de comparación: Con fuentes de mayor exactitud.
- Componente posicional: Horizontal y vertical. Las componentes "X" e "Y" se evalúan por separado. El error vertical y el horizontal son analizados individualmente.
- Clase de elementos: Puntos.
- Correspondencia con un estándar de exactitud predefinido: Basa su hipótesis a partir del límite de error aceptable para los errores aleatorios y sistemáticos.
- Descripción: El test es una expresión estadística de la exactitud calculada sobre la componente de error de las coordenadas "X", "Y" y "Z" para un muestra de al menos 20 puntos. La localización de los puntos sobre la carta se compara con la localización de los mismos sobre la fuente de mayor exactitud.

Se efectúa un test sobre la media, bajo el supuesto de varianza poblacional desconocida, para determinar si el error sistemático se encuentra entre los límites aceptables, y un test sobre la varianza para determinar si el error aleatorio es aceptable (Ariza, García y Amor, 2004).

- Procedimiento:
 1. Selección de una muestra de "n" puntos, con " $n \geq 20$ ".

2. Se calcula el error de cada punto en cada componente:

$$e_{x_i} = \bar{x}_i - x_i$$

$$e_{y_i} = \bar{y}_i - y_i$$

$$e_{z_i} = \bar{z}_i - z_i$$

Ec. Al. 3

donde $\bar{x}_i, \bar{y}_i, \bar{z}_i$ se corresponden con las coordenadas del punto sobre el terreno (o la fuente de mayor exactitud) y x_i, y_i, z_i son las coordenadas del punto "i" sobre el mapa.

3. Se calcula el error medio en cada componente:

$$\bar{e}_x = \frac{1}{n} \sum_{i=1}^n e_{x_i}$$

$$\bar{e}_y = \frac{1}{n} \sum_{i=1}^n e_{y_i}$$

$$\bar{e}_z = \frac{1}{n} \sum_{i=1}^n e_{z_i}$$

Ec. Al. 4

4. Calcular la desviación típica muestral en cada componente:

$$s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (e_{x_i} - \bar{e}_x)^2}$$

$$s_y = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (e_{y_i} - \bar{e}_y)^2}$$

$$s_z = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (e_{z_i} - \bar{e}_z)^2}$$

Ec. Al. 5

5. Efectuar el test de cumplimiento del estándar para cada componente con el fin de estudiar si el error sistemático se encuentra entre los límites aceptables. Para determinar el error sistemático se utiliza el estadístico:

$$t_x = \frac{\bar{e}_x \sqrt{n}}{s_x}$$

$$t_y = \frac{\bar{e}_y \sqrt{n}}{s_y}$$

$$t_z = \frac{\bar{e}_z \sqrt{n}}{s_z}$$

Ec. Al. 6

si se verifica que:

$$\begin{aligned} |t_x| &\leq t_{n-1, \alpha/2} \\ |t_y| &\leq t_{n-1, \alpha/2} \\ |t_z| &\leq t_{n-1, \alpha/2} \end{aligned}$$

Ec. Al. 7

el mapa cumple el test. El valor $t_{n-1, \alpha/2}$ es una medida teórica de la distribución *t de Student*, con “n-1” grados de libertad.

6. Efectuar el test de χ^2 para determinar el grado de variabilidad del error aleatorio:

$$\begin{aligned} \chi_x^2 &= \frac{s_x^2(n-1)}{\sigma_x^2} \\ \chi_y^2 &= \frac{s_y^2(n-1)}{\sigma_y^2} \\ \chi_z^2 &= \frac{s_z^2(n-1)}{\sigma_z^2} \end{aligned}$$

Ec. Al. 8

donde $\sigma_x, \sigma_y, \sigma_z$ son las varianzas máximas aceptables para el error en cada componente.

Si se verifica que:

$$\begin{aligned} |\chi_x^2| &< \chi_{n-1, \alpha}^2 \\ |\chi_y^2| &< \chi_{n-1, \alpha}^2 \\ |\chi_z^2| &< \chi_{n-1, \alpha}^2 \end{aligned}$$

Ec. Al. 9

el mapa cumple el test.

7. El mapa cumplirá el test EMAS si todas las componentes estudiadas han superado los dos test correspondientes.

- **Discusión:** El test EMAS presenta dos importantes diferencias respecto al test NMAS. En el test EMAS expresa la diferencia en términos de distancia sobre el terreno, mientras que en el NMAS la unidad de medida es la distancia sobre el mapa. La segunda es, que el límite del error aceptable puede ser definido por el usuario: hecho que es de gran importancia, ya que el estándar de exactitud puede ser definido en base a las exigencias del usuario.

La desviación típica muestral s_x, s_y, s_z es similar a la raíz del error cuadrático medio (Ec. II. 1). Ambos valores son iguales cuando el error medio ($\bar{e}_x, \bar{e}_y, \bar{e}_z$) es igual a 0.

Atkinson (2005) indica que se trata de unos de los estándares que realiza un análisis más exhaustivo y detallado. Aún así, plantea que originalmente, el nivel de significación en el contraste para la detección de sistematismos (*t-Student*) era del 10% para cada componente. De forma similar, el nivel de significación para cada componente en el estudio de la varianza (*Chi-cuadrado*) era del 5%. Todo ello hacía que el estándar fuera muy restrictivo.

AI.4. Test ASPRS

El test ASPRS fue desarrollado por la *American Society of Photogrammetry and Remote Sensing* (Merchant, 1987; ASPRS, 1989). Las características más representativas indicadas en Ariza (2002a), son

- Aplicaciones corrientes: El ASPRS especifica la exactitud de los mapas topográficos a gran escala.
- Método de comparación: Con fuentes de mayor exactitud.
- Componente posicional: Horizontal y vertical. Las componentes “X” e “Y” se evalúan individualmente. El error vertical y el horizontal son analizados por separado.
- Clase de elementos: Puntos.
- Correspondencia con un estándar de exactitud predefinido: Para los mapas de mejor calidad (clase I), se establecen límites a la raíz del error cuadrático medio RMSE para las componentes horizontales en función de la escala (Tabla AI. 1.). En la componente vertical se considera que el RMSE es de 1/3 del intervalo entre curvas de nivel, aunque hay que considerar que para los puntos acotados el RMSE será de 1/6 de dicha equidistancia. Las limitaciones para los mapas de clase II y clase III se obtienen multiplicando los valores dados para la clase I por dos y por tres.

Escala típica del producto	RMSE límite (metros)
1:50	0.0125
1:100	0.025
1:200	0.05
1:500	0.125
1:1000	0.25
1:2000	0.5
1:4000	1
1:5000	1.25

1:10000	2.5
1:20000	5

Tabla AI. 1. Requerimientos de exactitud planimétrica. Mapa Clase I.

- Descripción: Se realiza el cálculo de la raíz del error cuadrático medio RMSE para comparar con los valores límite (Tabla AI. 1.). Para ello, se utiliza un muestreo de, al menos, 20 puntos de fácil determinación. La localización de los puntos se compara en las dos fuentes.
- Procedimiento:
 1. Selección de una muestra de “n” puntos, con “n≥20”.
 2. Detección y corrección de errores groseros, inconsistencias, etc.
 3. Cálculo del RMSE para cada componente:

$$RMSE_X = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_{mi} - X_{ti})^2}$$

$$RMSE_Y = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_{mi} - Y_{ti})^2}$$

$$RMSE_Z = \sqrt{\frac{1}{n} \sum_{i=1}^n (Z_{mi} - Z_{ti})^2}$$

Ec. AI. 10

donde X_{ti}, Y_{ti}, Z_{ti} son las coordenadas del punto “i” sobre el terreno, X_{mi}, Y_{mi}, Z_{mi} las coordenadas sobre el mapa, y “n” el tamaño de la muestra.

4. Comprobación del cumplimiento del estándar utilizando los límites establecidos (Tabla AI. 1.).
 5. Se incluye una leyenda para informar de que el producto ha superado el estándar.
- Discusión: El test ASPRS, en relación al test EMAS, supone una simplificación en el cálculo y la eliminación de la hipótesis sobre el comportamiento del error, ya que no siempre se comporta con distribuciones de tipo normal.

Atkinson (2005) plantea que el test ASPRS es un estándar pionero en cuanto a que establece las bases de diferentes tipos o categorías de cartografía en lo que se refiere a su calidad. Sin embargo, recomienda trabajar, al menos, con 150 puntos de control y una cartografía un 10% más exacta que la exigida para la *Clase I*.

AI.5. Test NSSDA

El estándar NSSDA (*Nacional Standard for Spatial Data Accuracy*) está propuesto por el *Federal Geographic Data Committee* (FGDC, 1998) y se puede emplear en datos analógicos y/o digitales. Esta norma sustituye al test NMAS y al EMAS, siendo de obligado cumplimiento para las agencias federales de los Estados Unidos que realicen labores de producción cartográfica (Ariza, García y Amor, 2004). Las características más representativas indicadas en Ariza (2002a), son:

- Aplicaciones corrientes: Las agencias federales de los Estados Unidos que realicen labores de producción de datos cartográficos analógicos y/o digitales han de cumplir con los estándares del FGDC de acuerdo con el estándar NSSDA.
- Método de comparación: Con fuentes de mayor exactitud.
- Componente posicional: Analiza tanto la componente horizontal como la vertical. Las componentes “X” e “Y” se evalúan conjuntamente.
- Clase de elementos: Punto.
- Correspondencia con un estándar de exactitud predefinido: Se basa en el RMSE de la muestra. El propio usuario es quien decide los umbrales de aceptación.
- Descripción: Con la obtención del valor de RMSE para las coordenadas planimétricas (X, Y) por un lado y altimétrica (Z) por otro, calcula el error real de la muestra analizada en función de un determinado nivel de confianza impuesto por el usuario (generalmente el 95%). El NSSDA muestra un índice de calidad de la cartografía en unidades reales sobre el terreno.
- Procedimiento:
 1. Selección de una muestra de “n” puntos, con “n≥20”, perfectamente definidos sobre la cartografía y la fuente de mayor exactitud.
 2. Detección y corrección de errores groseros, inconsistencias, etc.
 3. Cálculo del $RMSE_X$, $RMSE_Y$ para las componentes X e Y:

$$RMSE_X = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_{mi} - X_{ti})^2}$$

$$RMSE_Y = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_{mi} - Y_{ti})^2}$$

Ec. AI. 11

4. Cálculo del $RMSE_r$ para la componente posicional (XY):

$$RMSE_r = \sqrt{RMSE_X^2 + RMSE_Y^2}$$

Ec. AI. 12

5. Cálculo del coeficiente de exactitud posicional a un 95% de confianza según 2 opciones:

a) Caso $RMSE_X = RMSE_Y$:

$$Exactitud_r = 1.7308 * RMSE_r$$

Ec. Al. 13

b) Caso $RMSE_X \neq RMSE_Y$:

$$Exactitud_r = 1.22385 * (RMSE_X + RMSE_Y)$$

Ec. Al. 14

6. Cálculo del $RMSE_Z$ para la componente Z:

$$RMSE_Z = \sqrt{\frac{1}{n} \sum_{i=1}^n (Z_{mi} - z_{ti})^2}$$

Ec. Al. 15

7. Cálculo del coeficiente de exactitud posicional a un 95% de confianza para la componente Z:

$$Exactitud_z = 1.96 * RMSE_z$$

Ec. Al. 16

- Discusión: El test NSSDA es un test que permite obtener un índice de calidad pero no indica si el producto es aceptado o rechazado. El resultado del test tendrá que aparecer como metadato en el producto, por lo que ha de ser el usuario el que establezca los límites de aceptación en función de sus necesidades. Según Ariza (2002a) este test tiene unas propiedades estadísticas intermedias entre el test NMAS y el test EMAS. Sin embargo, no se realiza un estudio de presencia de sistematismos, por lo que se centra en el estudio de la precisión de los datos. Realiza el control de las componentes horizontales X e Y conjuntamente.

Atkinson (2005) resalta que se trata de unos de los estándares más actuales y sitúa al usuario como responsable final en la utilización de la base de datos geoespacial. Como resultado del estudio de análisis de este test, concluye que si se emplea el número mínimo de puntos de control recomendado (20 puntos), el nivel de confianza en la obtención del valor de la exactitud horizontal es del 89% y no del 95% como indica el estándar. Para alcanzar este valor, el tamaño muestral ha de llegar a 95 o 100 puntos. Atkinson (2005) recomienda no descuidar el estudio preventivo de sistematismos mediante un contraste de tipo *t-Student*.

Al.6. Test de la fórmula de Koppe

Las características más representativas son (Giordano y Veregin, 1994):

- Aplicaciones corrientes: Se propone este test como alternativo al NMAS.

- Método de comparación: Con fuentes de mayor exactitud.
- Componente posicional: Vertical.
- Clase de elementos: Punto.
- Correspondencia con un estándar de exactitud predefinido: No está previsto.
- Descripción: El test utiliza coeficientes derivados empíricamente para calcular los efectos de la pendiente del terreno sobre el error medio vertical. En el test NMAS, el error de tipo vertical puede reducirse actuando sobre la componente horizontal del error posicional. De esta forma, un cierto grado de error horizontal en un punto corresponde a un mayor grado de error vertical cuando la pendiente es muy fuerte. Esto implica que el estándar de exactitud vertical del NMAS está menos vinculado al crecimiento de la pendiente.
- Procedimiento:
 1. Selección de una muestra puntos.
 2. Determinar la altitud (Z) del punto sobre el mapa.
 3. Determinar la altitud (Z) de los mismos puntos sobre la fuente de mayor exactitud.
 4. Medir el error vertical de cada punto.
 5. Medir el valor de la pendiente para cada punto.
 6. Efectuar el análisis de regresión usando la tangente del valor de la pendiente como variable independiente y el error vertical como variable dependiente.
 7. El error vertical medio se calcula utilizando:

$$\bar{e}_v = \pm(A + B \tan \alpha)$$

Ec. Al. 17

siendo “A” la variable independiente de la regresión, “ α ” la pendiente del terreno y “B” el coeficiente de regresión.

8. El error horizontal medio puede ser definido transformando el error vertical medio:

$$\bar{e}_h = \pm(B + A \cot \alpha)$$

Ec. Al. 18

- Discusión: El test de la fórmula de Koppe permite obtener una relación que une los comportamientos vertical y horizontal pudiendo ser aplicado para comparar la exactitud de curvas de nivel entre dos mapas, estimar el error que cabe esperar al interpolar o establecer una tolerancia de control para controlar la calidad de una cartografía (Gustafson y Loon, 1982).

AI.7. Test del USGS de categorías de exactitud de MDE

El test USGS (*United States Geological Survey*) de categorías de exactitud de Modelos Digitales de Elevaciones (MDE) presenta las siguientes características (Giordano y Veregin, 1994):

- Aplicaciones corrientes: Utilizado por el USGS para definir la exactitud de sus modelos digitales del terreno.
- Método de comparación: Con fuentes de mayor exactitud.
- Componente posicional: Vertical.
- Clase de elementos: Puntos.
- Correspondencia con un estándar de exactitud predefinido: Se basa en el empleo de umbrales para el error máximo y de la raíz del error cuadrático medio (RMSE).
- Descripción: El test se basa en la cuantificación de los niveles de exactitud, que posteriormente se utilizan para determinar la categoría de exactitud del MDE.
- Procedimiento:
 1. Selección de una muestra de al menos 20 puntos.
 2. Cálculo del $RMSE_Z$ para la componente Z:

$$RMSE_Z = \sqrt{\frac{1}{n} \sum_{i=1}^n e_{zi}^2}$$

Ec. AI. 19

3. Definición de las categorías de exactitud para el MDE:

Nivel MDE	Valor máximo RMSE	Límite error vertical
Nivel 1	15 m	50 m
Nivel 2	7 m	El doble del intervalo de las curvas de nivel de la fuente de mayor exactitud.
Nivel 3	7 m	El intervalo de las curvas de nivel de la fuente de mayor exactitud.

Tabla AI. 2. Categorías de Exactitud para MDE. USGS.

- Discusión: Al contrario que ocurre en el test EMAS, las categorías de exactitud de este test están definidas en términos de nivel total de error (mediciones del valor del RMSE) como relativas a los puntos aislados. En

el test EMAS sólo se cuantifica a nivel general. Por tanto, el test USGS de categorías de exactitud de MDE es más preciso en cuanto que en él los errores groseros, aparte de su influencia en el RMSE, contribuyen a determinar la categoría de exactitud de la cartografía. La importancia de esto radica sobre todo en cuando es notorio que las metodologías utilizadas para la recogida y la creación de la base de datos introducen notables errores (Ariza, 2002a).

Anexo II

Especificaciones de los BDGs

All.1. Mapa Topográfico de Andalucía

- Definición espacial: Son 2745 hojas que cubren todo el territorio de la Comunidad Autónoma de Andalucía. Cada hoja ocupa un dieciseisavo de una hoja de MTN50, lo que resulta de dividir ésta en 4x4 filas y columnas. La superficie viene a ser, de forma aproximada, unas 3400 hectáreas, unos 7.7 por 4.6 kms de Este a Oeste y de Norte a Sur.
- Definición nominal: Las hojas de la serie toman su numeración del MTN50, a la que le siguen dos números que hacen referencia a la columna y la fila en la que se inserta la hoja del MTN10 dentro del MTN50, por ejemplo, la hoja 947-4-1 del MTA10, se refiere a la hoja 947 del MTN50, correspondiente a Jaén, y dentro de ésta a la columna cuarta, primera fila.
- Definición métrica y Cartografía matemática: Elipsoide de Hayford. Datum de Postdam. Origen de coordenadas geográficas en el Ecuador y Meridiano de Greenwich. Origen de altitudes, nivel medio del mar en Alicante. Proyección UTM sobre los husos 29 y 30.
- Definición temática y gráfica:
 - Geografía física:
 - Relieve: curvas de nivel maestras (50m) y sencillas (10m), puntos acotados al metro, vértices geodésicos de 1º, 2º y 3º, vértices topográficos y puntos de apoyo.
 - Hidrografía: línea del nivel del mar, corrientes de agua naturales (ríos, arroyos, cauces,...), canales, acequias, estanques, manantiales, pozos.
 - Usos del suelo: límites de parcelas aparentes de secano, regadío, huerta, viñedo, frutales, olivar, eucaliptal, erial, prados y dehesas, monte bajo, monte alto, repoblación forestal, masas de árboles y cortafuegos.
 - Geografía humana:
 - Núcleos de población y edificaciones, muros, paredes o tapias, edificios, alambradas y cercas metálica.
 - Líneas eléctricas de alta tensión, transformadores, torres metálicas y postes.
 - Vías de comunicación, Autopistas, autovías, carreteras, pistas, caminos carreteros, vías pecuarias, cañadas, sendas, cortafuegos, ferrocarriles.
 - Divisiones administrativas: Límite de provincia, límite de término municipal.
- Linaje (proceso de producción): Las fuentes del MTA10 son los vuelos fotogramétricos en Blanco y Negro ejecutados a E25k y E20k, tanto en la fase de realización inicial como en las actualizaciones periódicas. No se recogen elementos con representación menor al milímetro sobre el mapa.
- Índices de calidad: Altimetría expresada mediante curvas de nivel de 10 m de equidistancia y curvas maestras cada 50 m. Puntos acotados redondeados al metro. La planimetría tienen una exactitud media de 3 m. La versión digital se ha obtenido por escaneo de los originales en color con una resolución de 300 dpi. Estos ficheros han sido georeferenciados tomando 8 puntos de control.

- **Formato de distribución:** En formato A1 (60x84 cm) en soporte indeformable o como copia en papel. Los originales son en color pero las copias son todas en blanco y negro. También se dispone digitalmente en formato ráster en 8 CD-ROMs, uno por cada provincia de la comunidad andaluza, con tantos ficheros independientes como hojas. Los ficheros se corresponde con el escaneo de las hojas del MTA10 analógico, a las que se han eliminado los elementos marginales, permitiendo de esta forma el case de las diversas hojas para formar un continuo. El formato de los ficheros es el TIFF, que se acompaña con los datos para su correcta georreferenciación. Los ficheros están comprimidos. La única información existente sobre cada hoja digital es la que se refiere a la fecha de la ejecución de la hoja original. Se incorpora un visualizador denominado Mulhacén con algunas capacidades elementales.
- **Usos:** El objetivo del MTA10 era conseguir una cartografía básica, homogénea y a gran escala del territorio andaluz, dado que, en ese momento, sólo se disponía de cartografías de media escala de los organismos productores nacionales. Se ha utilizado para la obtención de cartografía temática como mapas de carreteras, mapas de parques naturales, etc., así como para la planificación territorial y urbana, análisis de infraestructuras, evaluación de recursos naturales y protección del medio ambiente, en la Comunidad Autónoma de Andalucía.
- **Disponibilidad:** Toda Andalucía.

AII.2. Base Cartográfica Numérica

- **Definición espacial:** Cubre todo el territorio español, incluyendo el territorio peninsular, Islas Baleares, Islas Canarias, Ceuta y Melilla. Está distribuida según la división en 4133 hojas del Mapa Topográfico Nacional E25k (MTN25).
- **Definición nominal:** Sigue la distribución de las Hojas del MTN50. Cada hoja del MTN50 se divide en cuatro cuartos que se corresponden con las hojas editadas del MTN25. La referencia de los cuartos es I, II, III y IV correspondiendo I a la primera fila primera columna y IV segunda fila, segunda columna. Cada cuarto lleva el nombre de la población de mayor número de habitantes (Nomenclátor INE). En caso de no existir población, se le dará el nombre del accidente geográfico más destacado. En cualquier caso, el nombre de uno de los cuartos del MTN25 debe coincidir con el de la hoja MTN50 correspondiente. Por ejemplo, la Hoja 984 del MTN50 se corresponde con Sevilla, en el MTN25 la denominación de los cuartos es: Hoja 984-I Olivares, Hoja 984-II La Algaba, Hoja 984-III Camas y Hoja 984-IV Sevilla.

Dentro de cada hoja se tienen dos ficheros, uno de planimetría y otro de altimetría, cuyos nombres comienzan por P o por Z respectivamente. Por ejemplo, P0123C2.dgn, Z0123C2.dgn.

- **Definición métrica y cartografía matemática:** El Sistema de Referencia es ED50, el elipsoide el de Hayford y el Datum de Potsdam. La proyección cartográfica utilizada es la UTM, las coordenadas de cada hoja se encuentran en el huso correspondiente.

- Definición temática y gráfica: Contiene toda la información que se representa en el MTN25, pero estructurada y preparada para su gestión en un SIG.
 - Geografía física:
 - Relieve: Curvas de nivel maestras (50m) y sencillas (10m), intercaladas y de depresión, puntos acotados al metro, vértices geodésicos de 1^{er} orden y de la ROI, desmontes y terraplenes, vertederos y hoyas.
 - Hidrografía: Línea del nivel del mar y batimétricas, corrientes de agua permanentes y no permanentes, cauces secos y aluviones, canales y acequias (>3m, 1-3m, < 1m), conducciones subterráneas. También se incluyen piscinas, abrevaderos, depósitos (elevados, a nivel de suelo, subterráneos) pozos, fuentes, manantiales, depuradoras y sifones.
 - Usos del suelo: Bosques (frondosas (-F), coníferas (-C), mixtos (M)), Matorral (Mt), regadíos (arrozales (A), caña azúcar (C-A), platanar (B), regadío permanente (Re), regadío con frutales (Re-P)), cultivos arbóreos (olivar (P-O), cítricos (P-N), frutales (P-V), almendros (P-A)), Viñas (viña (V), viña-olivar (V-O), viña con frutales (V-P)), cultivos mixtos (viña y terreno claro, olivar y terreno claro, viña/olivar y terreno claro (V-X; O-X); terrenos incultivables (marismas (Mm), dunas y playas (D), roquedos (R), terrenos volcánicos (L)), Terrenos claros (X) y terrenos claros con árboles (X-A); parques y jardines (J), Prados y Pastizales, Árboles aislados, Árboles en fila.
 - Geografía humana:
 - Núcleos de población y edificaciones aisladas, corrales y ruinas, castillos, molinos de viento y agua, camping, y restos arqueológicos, muros, muros de contención y alambradas. Edificios religiosos (cristianos, musulmán y otros), cementerios y cruces aisladas, torres de observación, estaciones de observación, aduanas y faros.
 - Líneas eléctricas de alta tensión, centrales eléctricas hidráulicas, térmicas, eólicas solares, gas, nucleares.
 - Vías de comunicación, Autopistas, autovías, carreteras nacionales, carreteras autonómicas de 1^{er}, 2^o y 3^o orden y otras, carreteras en construcción, pistas y viales de urbanización, caminos y sendas, vías pecuarias y estaciones de servicio. Ferrocarriles de ancho normal, doble y sencillo, de vía estrecha doble y sencilla, electrificados, en construcción, abandonados, teleféricos, telesillas y telesquí, estaciones y túneles y pasos a nivel.
 - Minas, Canteras, y minas a cielo abierto.
 - Divisiones administrativas: Límite de provincia, límite de término municipal.
 - Los temas o capas de información son: tema 01 División Administrativa; tema 02 Relieve; tema 03 Hidrografía y costas;

tema 05 Edificaciones y construcciones; tema 06 Vías de comunicación; tema 07 Conducciones y transmisiones; tema 10 Vértices geodésicos.

- Método (linaje): Grupo de datos 1: Planimetría de BCN25. Incluye Hidrografía, Cultivos y Vegetación, Edificaciones y Construcciones, Vías de Comunicación, Conducciones y Transmisiones. Fuente: MTN25 del IGN.
 - Procesos: Restitución fotogramétrica asistida por ordenador. Formación por trabajo de campo y trabajo interactivo en entorno MicroStation con comandos de usuario y programas Fortran de desarrollo propio; edición para trazado en entorno MicroStation con comandos de usuario y programas Fortran de desarrollo propio; selección de geometría que representa entes del mundo real, desechando símbolos puntuales, lineales, superficiales y todo aquello cuyo fin exclusivo es el trazado del mapa (cuadrícula, leyenda, etc.). El tratamiento geométrico consiste en la eliminación de puntos superfluos y repetidos; eliminación de bucles y vueltas atrás; eliminación de líneas repetidas, incluidas y solapadas; resolución de intersecciones entre elementos; resolución de anclajes; detección y resolución, si procede, de extremos libres; unificación de elementos del mismo código y nombre dentro de tolerancia; tratamiento de elementos perimetrales; creación de los contornos relativos a cascos de población, urbanizaciones, edificaciones dispersas, barrios, campus universitarios, etc.; case geométrico entre hojas, y cases entre hojas limítrofes.
 - Grupo de datos 2: División Administrativa de BCN25. Incluye Líneas Límite Municipales, Provinciales y Autonómicas.
 - Fuente: Base de Datos de Líneas Límite (BDLL).
 - Procesos: Incorporación y armonización con BCN25.
 - Grupo de datos 3: Curvas de nivel de BCN25.
 - Fuente: Curvas de nivel del MDT25.
 - Procesos: Incorporación y armonización con BCN25.
 - Grupo de datos 4: Vértices Geodésicos de BCN25.
 - Fuente: Base de Datos de Vértices Geodésicos.
 - Procesos: Incorporación y armonización con BCN25.
- Índices de calidad: Exactitud geométrica: Comprendida entre 2.5 m y 7.5 m, dependiendo del código. Estimación realizada tomando en cuenta los errores teóricos cometidos en la fase de restitución fotogramétrica y los desplazamientos a que se someten algunos objetos por excesiva proximidad con otros objetos. El orden de mayor a menor exactitud en los datos es: vértices geodésicos (tema 10), límites administrativos (tema 1), hidrografía (tema 3), comunicaciones (tema 6) y poblaciones (tema 5). Exactitud semántica: Estimada, tasa de errores menor del 4%. Compleción: Estimada, tasa de errores menor del 4%. Estimadas a partir del tanto por ciento de errores detectados en la BCN200.
- Usos: Las principales aplicaciones son la gestión de redes de infraestructuras (agua, electricidad, gas, teléfono); sistemas de navegación para automóvil;

sector editorial (atlas, mapas); planificación (administración central y autonómica).

- Disponibilidad: Toda España.

Anexo III

Ajustes a distribuciones de probabilidad

AIII.1. Método MOS

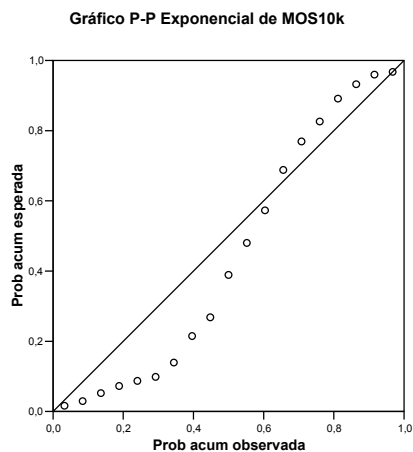


Figura AIII. 1. Ajuste a distribución exponencial - MOS. MTA

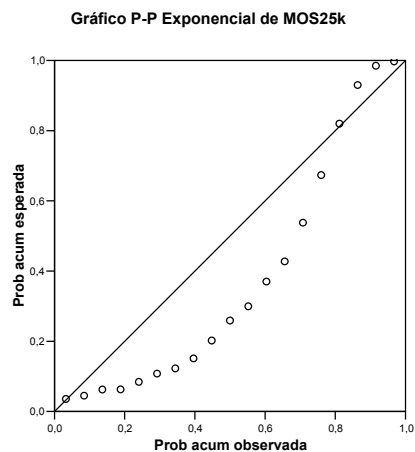


Figura AIII. 2. Ajuste a distribución exponencial - MOS. BCN

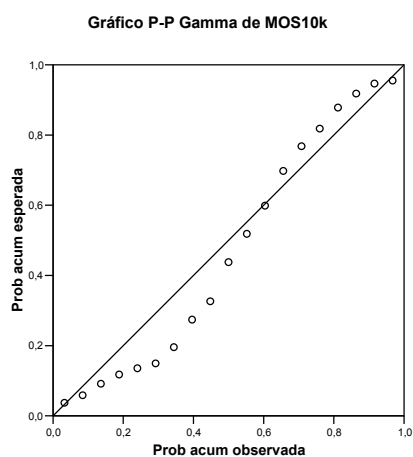


Figura AIII. 3. Ajuste a distribución Gamma - MOS. MTA

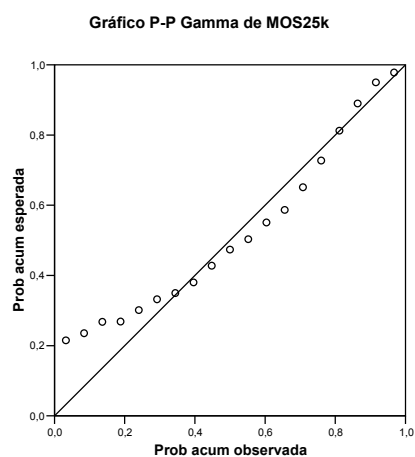


Figura AIII. 4. Ajuste a distribución Gamma - MOS. BCN

AIII.2. Método MOD. Superficie Común

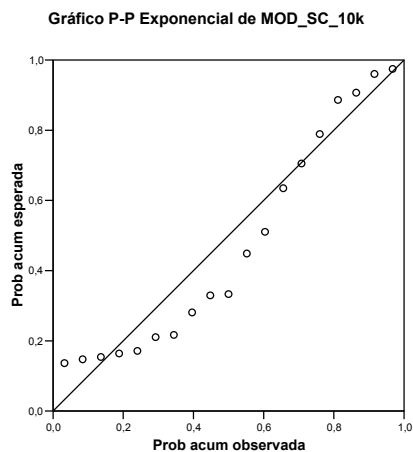


Figura AIII. 5. Ajuste a distribución exponencial - superficie común. MTA

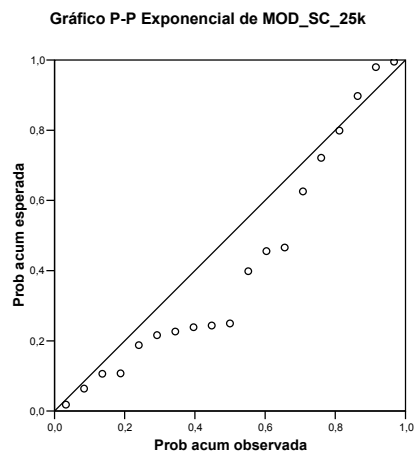


Figura AIII. 6. Ajuste a distribución exponencial - superficie común. BCN

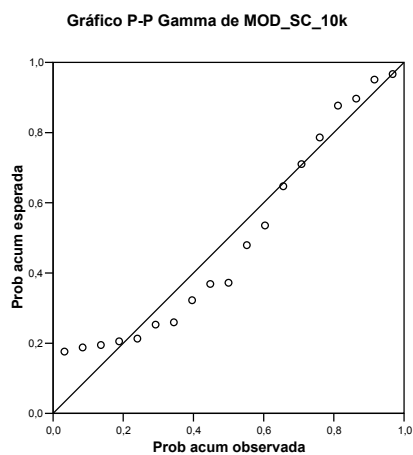


Figura AIII. 7. Ajuste a distribución Gamma - superficie común. MTA

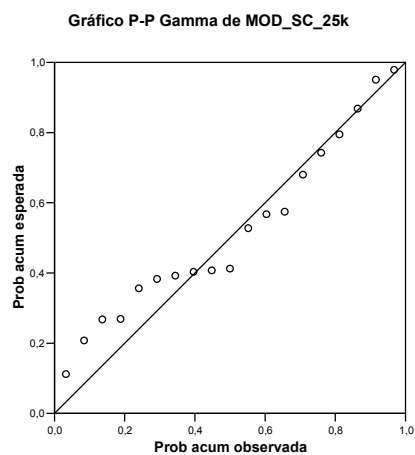


Figura AIII. 8. Ajuste a distribución Gamma - superficie común. BCN

AIII.3. Método MOD. Desplazamiento promedio

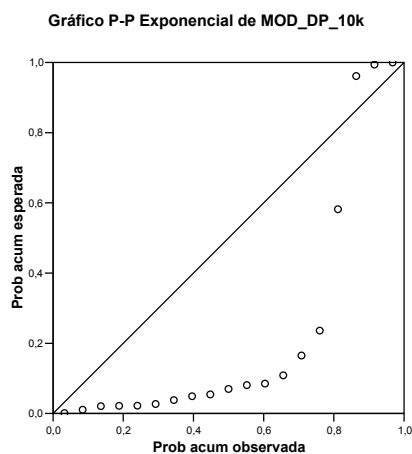


Figura AIII. 9. Ajuste a distribución exponencial - desplazamiento promedio. MTA

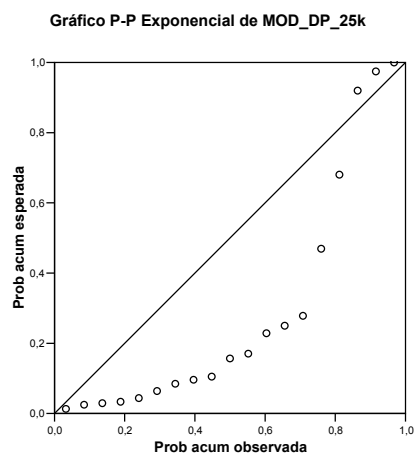


Figura AIII. 10. Ajuste a distribución exponencial - desplazamiento promedio. BCN

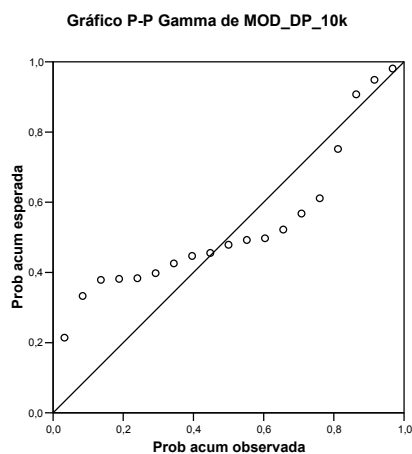


Figura AIII. 11. Ajuste a distribución Gamma - desplazamiento promedio. MTA

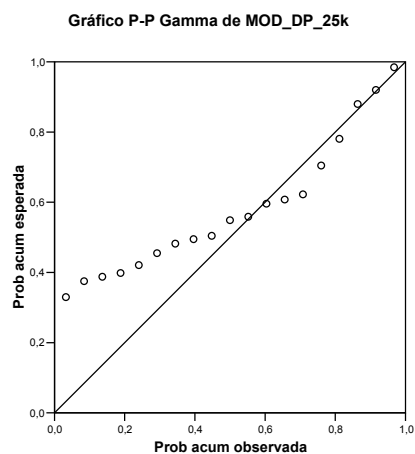


Figura AIII. 12. Ajuste a distribución Gamma - desplazamiento promedio. BCN

Anexo IV

Tablas de Resultados

AIV.1. Resultados del Método MDH. Base de datos MTA10

Zona	Longitud [m]	Distancia de Hausdorff		Valores medios	
		Media [m]	Desviación [m]	Media [m]	Desviación [m]
Almería	196929.215	8.786	4.081	3.915	2.344
Cádiz	136978.493	9.383	3.264	3.956	1.633
Co-Ja	330711.542	9.537	4.794	3.785	1.952
Huelva	193703.266	13.195	8.156	4.461	2.116
Málaga	171015.589	12.060	5.309	5.689	3.016
Sevilla	147857.472	13.221	4.466	5.552	2.590
Total	1177195.58	10.825	5.595	4.437	2.403

Tabla AIV. 1. Resultados del Método MDH. Base de datos MTA10

AIV.2. Resultados del Método MDH. Base de datos BCN25

Zona	Longitud [m]	Distancia de Hausdorff		Valores medios	
		Media [m]	Desviación [m]	Media [m]	Desviación [m]
Almería	237506.881	9.753	7.485	2.892	2.372
Cádiz	132623.133	11.932	16.116	3.178	2.530
Co-Ja	349340.549	11.640	9.224	3.525	2.354
Huelva	184719.876	8.171	5.613	11.274	1.744
Málaga	173315.215	10.306	8.448	4.365	4.265
Sevilla	155447.534	18.121	22.852	3.747	2.756
Total	1232953.19	11.418	12.305	3.364	2.663

Tabla AIV. 2. Resultados del Método MDH. Base de datos BCN25

AIV.3. Resultados del Método MSK. Base de datos MTA10

Zona	Longitud [m]	Media Superficie/Longitud [m]
Almería	197364.552	3.656
Cádiz	137084.349	3.747
Co-Ja	352100.708	3.661
Huelva	199834.204	4.232
Málaga	171217.133	5.511
Sevilla	152935.635	5.215
Total	1210536.58	4.222

Tabla AIV. 3. Resultados del Método MSK. Base de datos MTA**AIV.4. Resultados del Método MSK. Base de datos BCN25**

Zona	Longitud [m]	Media Superficie/Longitud [m]
Almería	239310.571	2.522
Cádiz	148068.121	3.192
Co-Ja	356604.888	3.244
Huelva	194754.191	2.466
Málaga	173315.215	4.146
Sevilla	155447.534	3.312
Total	3946392.39	3.114

Tabla AIV. 4. Resultados del Método MSK. Base de datos BCN25

AIV.5. Resultados del Método MOS. Base de datos MTA10

Orlado [m]	Porcentaje de línea dentro del orlado por zonas						
	Almería	Cádiz	Co-Ja	Huelva	Málaga	Sevilla	Todas
1	18.24	18.38	19.52	16.51	11.76	11.93	16.63
1.1	20.15	19.83	21.36	18.09	12.87	13.05	18.20
1.2	21.90	21.32	23.18	19.77	14.01	14.03	19.75
1.3	23.61	22.88	25.00	21.45	15.15	15.11	21.31
1.4	25.30	24.44	26.84	23.10	16.57	16.23	22.91
1.5	26.96	25.82	28.60	24.67	17.48	17.32	24.37
1.6	28.74	27.05	30.30	26.33	18.55	18.39	25.86
1.7	30.41	28.21	31.98	27.91	19.70	19.48	27.31
1.8	32.08	29.39	33.61	29.37	20.88	20.55	28.74
1.9	33.97	30.65	35.35	30.70	22.05	21.55	30.21
2	35.52	31.92	37.02	32.05	23.12	22.70	31.61
2.1	37.28	33.17	38.53	33.55	24.21	23.87	33.02
2.2	39.09	34.49	40.13	35.00	25.33	25.07	34.48
2.3	41.00	35.90	41.70	36.45	26.49	26.19	35.96
2.4	42.73	37.20	43.35	38.06	27.63	27.29	37.43
2.5	44.28	38.77	44.94	39.72	28.73	28.41	38.90
2.6	45.95	40.17	46.40	41.21	29.79	29.53	40.29
2.7	47.51	41.59	47.83	42.49	30.80	30.71	41.62
2.8	49.03	43.23	49.32	43.82	31.79	31.85	43.00
2.9	50.39	44.97	50.76	45.13	32.84	32.97	44.34
3	51.73	46.54	52.17	46.41	33.92	34.05	45.65
3.2	54.50	49.48	54.94	48.79	35.96	36.19	48.19
3.4	57.21	52.11	57.45	51.05	37.92	38.55	50.61
3.6	59.87	54.97	59.81	53.30	39.86	40.66	52.97
3.8	62.69	57.66	61.94	55.36	41.94	42.80	55.25

Orlado [m]	Porcentaje de línea dentro del orlado por zonas						
	Almería	Cádiz	Co-Ja	Huelva	Málaga	Sevilla	Todas
4	65.26	59.94	64.06	57.36	43.92	44.93	57.43
4.2	67.34	62.19	66.00	59.30	45.88	47.03	59.45
4.4	69.43	64.49	67.85	61.39	47.91	49.27	61.50
4.6	71.48	66.50	69.67	63.51	50.00	51.36	63.50
4.8	73.31	68.25	71.47	65.55	51.84	53.34	65.37
5	74.79	70.08	73.18	67.46	53.58	55.39	67.13
5.5	78.29	74.85	76.93	71.56	57.70	60.67	71.26
6	81.28	78.62	80.08	75.25	61.68	64.79	74.79
6.5	84.20	82.29	82.88	78.79	65.29	69.34	78.16
7	86.86	85.47	85.47	81.90	68.59	73.37	81.20
7.5	89.09	88.52	87.60	84.67	71.89	76.84	83.89
8	90.90	91.02	89.65	86.73	75.45	79.93	86.29
8.5	92.55	93.04	91.53	88.31	78.48	82.23	88.32
9	93.86	95.06	92.93	89.81	80.98	84.17	90.02
9.5	94.85	96.65	94.33	90.91	83.20	86.35	91.54
10	95.52	97.74	95.67	92.17	85.31	88.05	92.88
11	96.29	99.07	97.56	95.17	88.80	90.74	95.03
12	97.16	99.61	98.49	96.65	91.40	92.98	96.40
13	97.88	99.82	99.11	97.48	94.01	94.78	97.46
14	98.43	99.92	99.37	98.12	95.69	95.88	98.11
15	98.77	99.96	99.52	98.53	96.81	96.87	98.57
16	99.08	99.97	99.63	98.96	97.54	97.97	98.96
17	99.33	100.00	99.70	99.22	98.44	98.72	99.29
18	99.51	100.00	99.78	99.41	98.90	99.40	99.53
19	99.64	100.00	99.81	99.57	99.14	99.75	99.66
20	99.70	100.00	99.83	99.66	99.32	99.87	99.73

Tabla AIV. 5. Resultados del Método MOS. Base de datos MTA10

AIV.6. Resultados del Método MOS. Base de datos BCN25

Orlado [m]	Porcentaje de línea dentro del orlado por zonas						
	Almería	Cádiz	Co-Ja	Huelva	Málaga	Sevilla	Todas
1	39.56	24.23	25.07	28.93	25.03	24.95	28.28
1.1	42.78	26.54	27.39	31.50	27.37	27.06	30.79
1.2	45.67	28.91	29.62	33.87	29.63	29.28	33.18
1.3	48.47	31.48	31.89	36.27	31.87	31.27	35.57
1.4	51.00	33.85	34.12	38.68	33.99	33.16	37.84
1.5	53.46	36.18	36.32	41.18	35.96	35.30	40.11
1.6	55.59	38.41	38.48	43.60	37.88	37.47	42.28
1.7	57.59	40.54	40.51	46.27	39.78	39.77	44.43
1.8	59.43	42.53	42.54	48.75	41.58	42.13	46.50
1.9	61.39	44.52	44.51	51.09	43.31	44.29	48.52
2	63.26	46.40	46.40	53.32	44.98	46.30	50.44
2.1	65.01	48.30	48.24	55.36	46.60	48.31	52.29
2.2	66.62	50.26	50.01	57.30	48.17	50.37	54.09
2.3	68.05	52.07	51.73	59.29	49.80	52.37	55.83
2.4	69.37	53.89	53.47	61.16	51.24	54.58	57.54
2.5	70.65	55.49	55.23	62.96	52.65	56.35	59.14
2.6	71.82	57.14	56.89	64.61	54.14	58.08	60.70
2.7	72.91	58.74	58.50	66.22	55.44	59.70	62.16
2.8	73.88	60.26	60.07	67.91	56.67	61.44	63.61
2.9	74.84	61.86	61.49	69.43	57.76	62.90	64.94
3	75.64	63.35	62.79	70.87	58.87	64.29	66.17
3.2	77.18	66.14	65.19	73.54	60.98	67.44	68.55
3.4	78.69	68.53	67.47	75.95	62.93	69.96	70.70
3.6	79.89	70.79	69.63	78.20	64.71	72.06	72.64
3.8	80.97	72.97	71.54	80.27	66.42	74.24	74.46

Orlado [m]	Porcentaje de línea dentro del orlado por zonas						
	Almería	Cádiz	Co-Ja	Huelva	Málaga	Sevilla	Todas
4	81.98	75.19	73.18	82.06	68.02	76.37	76.13
4.2	82.79	77.19	74.76	83.78	69.45	78.43	77.67
4.4	83.54	79.14	76.20	85.34	70.84	80.25	79.10
4.6	84.24	80.77	77.58	86.82	72.06	81.51	80.36
4.8	84.79	82.11	78.81	88.18	73.29	82.59	81.47
5	85.27	83.46	80.07	89.51	74.28	83.65	82.54
5.5	86.54	86.02	83.13	92.07	76.60	85.68	84.90
6	87.95	88.12	85.38	93.97	78.38	86.90	86.73
6.5	89.19	89.72	87.30	95.65	79.90	88.15	88.31
7	90.56	90.78	88.85	96.56	81.29	89.48	89.62
7.5	91.72	91.49	90.36	97.00	82.67	90.54	90.74
8	92.84	92.13	91.47	97.47	84.05	91.47	91.71
8.5	93.88	92.83	92.55	97.83	85.46	92.68	92.69
9	94.65	93.46	93.50	98.11	86.47	93.36	93.44
9.5	95.50	94.08	94.18	98.36	87.40	94.00	94.11
10	96.26	94.77	94.84	98.53	88.37	94.78	94.77
11	97.36	96.04	96.12	98.84	89.90	96.00	95.90
12	98.31	97.15	96.97	98.95	91.10	96.94	96.74
13	98.69	97.83	97.55	99.06	92.48	97.71	97.36
14	98.84	98.36	98.03	99.27	93.91	98.01	97.85
15	98.93	98.68	98.53	99.57	95.03	98.29	98.27
16	99.06	98.87	98.90	99.71	95.97	98.51	98.60
17	99.17	99.12	99.17	99.74	96.59	98.72	98.84
18	99.28	99.31	99.40	99.76	97.40	98.90	99.09
19	99.36	99.38	99.59	99.78	97.99	99.02	99.26
20	99.43	99.46	99.69	99.79	98.53	99.08	99.40

Tabla AIV. 6. Resultados del Método MOS. Base de datos BCN25

AIV.7. Resultados del Método MOD. Base de datos MTA10

Orlado [m]	Superficie común por zonas [%]						
	Almería	Cádiz	Co-Ja	Huelva	Málaga	Sevilla	Todas
1	18.04	17.52	19.21	16.44	11.67	11.60	18.04
2	34.78	31.84	35.56	31.31	22.73	22.77	34.78
3	47.97	44.66	48.05	43.68	33.22	33.61	47.97
4	58.28	54.94	57.32	54.51	42.16	43.44	58.28
5	66.33	62.88	64.48	61.52	49.94	51.56	66.33
6	71.35	68.87	71.36	67.85	56.53	58.43	71.35
7	75.45	73.27	74.30	72.08	61.96	63.58	75.45
8	78.39	76.59	77.43	75.47	66.32	67.72	78.39
9	80.73	79.18	79.90	78.10	69.89	71.14	80.73
10	82.78	81.25	81.90	80.24	72.93	73.98	82.78
11	84.33	82.95	83.67	81.93	75.49	76.32	84.33
12	85.69	84.36	85.01	83.41	77.49	78.27	85.69
13	86.78	85.56	86.15	84.66	79.22	79.93	86.78
14	87.72	86.58	87.14	85.74	80.53	81.41	87.72
15	88.53	87.47	87.99	86.67	82.01	82.64	88.53
16	89.24	88.25	89.12	87.49	83.13	83.71	89.24
17	89.87	88.93	89.75	88.21	84.07	84.66	89.87
18	90.42	89.54	90.31	88.86	84.94	85.50	90.42
19	90.92	90.09	90.82	89.43	85.72	86.25	90.92
20	91.38	90.58	91.28	89.96	86.43	86.82	91.38

Tabla AIV. 7.Resultados del Método MOD(Superficie Común). Base de datos MTA10

Orlado [m]	Desplazamiento promedio por zonas [m]						
	Almería	Cádiz	Co-Ja	Huelva	Málaga	Sevilla	Todas
1	2.57	2.60	2.54	2.62	2.78	2.79	2.57
2	4.09	4.29	4.05	4.31	4.87	4.87	4.09
3	4.90	5.23	4.91	5.30	6.31	6.28	4.90
4	5.24	5.68	5.36	5.71	7.29	7.14	5.24
5	5.28	5.85	5.58	6.04	7.89	7.64	5.28
6	5.39	5.89	5.39	6.05	8.22	7.87	5.39
7	5.39	5.90	5.65	6.13	8.40	8.05	5.39
8	5.42	5.91	5.67	6.16	8.50	8.16	5.42
9	5.44	5.91	5.68	6.18	8.55	8.21	5.44
10	5.40	5.92	5.68	6.20	8.55	8.23	5.40
11	5.40	5.92	5.64	6.23	8.52	8.24	5.40
12	5.38	5.93	5.64	6.24	8.53	8.25	5.38
13	5.38	5.93	5.65	6.25	8.54	8.26	5.38
14	5.39	5.94	5.65	6.26	8.62	8.25	5.39
15	5.39	5.94	5.65	6.27	8.54	8.26	5.39
16	5.39	5.95	5.46	6.28	8.54	8.27	5.39
17	5.39	5.95	5.46	6.28	8.57	8.28	5.39
18	5.40	5.95	5.46	6.29	8.58	8.29	5.40
19	5.40	5.96	5.47	6.29	8.59	8.30	5.40
20	5.39	5.96	5.47	6.29	8.60	8.38	5.39

Tabla AIV. 8. Resultados del Método MOD(Desplazamiento Promedio). Base de datos MTA10

Orlado [m]	Oscilación por zonas [oscilaciones/km]						
	Almería	Cádiz	Co-Ja	Huelva	Málaga	Sevilla	Todas
1	3.17	2.84	4.62	2.86	5.17	2.48	3.70
2	2.73	2.86	4.11	2.41	5.13	2.41	3.39
3	2.30	2.46	3.73	2.02	4.75	2.19	3.02
4	2.06	2.20	3.39	1.71	4.48	1.95	2.73
5	1.87	1.86	3.22	1.53	3.89	1.73	2.47
6	1.83	1.71	3.02	1.40	3.52	1.54	2.29
7	1.78	1.66	2.93	1.29	3.26	1.45	2.19
8	1.74	1.66	2.89	1.26	3.11	1.37	2.13
9	1.75	1.65	2.88	1.24	3.02	1.34	2.11
10	1.71	1.65	2.88	1.21	3.01	1.28	2.09
11	1.70	1.65	2.87	1.21	2.98	1.26	2.08
12	1.70	1.65	2.87	1.20	2.98	1.26	2.08
13	1.70	1.65	2.87	1.20	2.97	1.26	2.08
14	1.70	1.65	2.87	1.20	2.97	1.26	2.08
15	1.70	1.65	2.87	1.20	2.97	1.26	2.08
16	1.70	1.65	2.87	1.20	2.97	1.26	2.08
17	1.70	1.65	2.87	1.20	2.97	1.26	2.08
18	1.70	1.65	2.87	1.20	2.97	1.26	2.08
19	1.70	1.65	2.87	1.20	2.97	1.26	2.07
20	1.70	1.65	2.87	1.20	2.97	1.26	2.08

Tabla AIV. 9. Resultados del Método MOD(Oscilación). Base de datos MTA10

AIV.8. Resultados del Método MOD. Base de datos BCN25

Orlado [m]	Superficie común por zonas [%]						
	Almería	Cádiz	Co-Ja	Huelva	Málaga	Sevilla	Todas
1	36.87	24.03	24.42	28.12	24.15	24.06	36.87
2	56.58	43.42	43.02	48.91	41.26	43.77	56.58
3	66.55	56.49	55.26	62.48	52.31	56.88	66.55
4	73.42	65.69	63.66	71.83	59.57	65.12	73.42
5	77.94	70.77	69.74	77.75	65.26	70.76	77.94
6	81.34	75.08	74.20	82.18	69.43	74.96	81.34
7	83.91	78.34	77.62	84.58	72.77	78.20	83.91
8	85.75	80.86	80.21	86.48	75.54	80.81	85.75
9	87.26	82.86	82.26	88.00	77.87	82.82	87.26
10	88.58	84.49	83.98	89.17	79.86	84.43	88.58
11	89.62	85.84	85.35	90.13	81.59	85.77	89.62
12	90.56	86.98	86.57	90.93	83.11	86.89	90.56
13	91.26	87.95	87.57	91.61	84.38	87.78	91.26
14	91.87	88.79	88.44	92.27	85.48	88.60	91.87
15	92.51	90.21	89.20	92.78	86.43	89.32	92.51
16	92.96	90.80	89.86	93.22	87.25	90.01	92.96
17	93.35	91.32	90.50	93.61	87.99	90.56	93.35
18	93.71	91.78	91.02	93.95	88.66	91.06	93.71
19	93.90	92.20	91.49	94.26	89.28	91.51	93.90
20	94.18	92.57	91.91	95.00	89.82	91.91	94.18

Tabla AIV. 10. Resultados del Método MOD(Superficie Común). Base de datos BCN25

Orlado [m]	Desplazamiento promedio por zonas [m]						
	Almería	Cádiz	Co-Ja	Huelva	Málaga	Sevilla	Todas
1	1.98	2.39	2.38	2.26	2.38	2.38	1.98
2	2.73	3.57	3.58	3.21	3.68	3.53	2.73
3	3.16	4.12	4.22	3.54	4.48	4.05	3.16
4	3.34	4.33	4.57	3.54	5.07	4.37	3.34
5	3.47	4.62	4.75	3.50	5.44	4.58	3.47
6	3.52	4.73	4.86	3.36	5.74	4.70	3.52
7	3.55	4.80	4.92	3.39	5.97	4.77	3.55
8	3.58	4.85	4.98	3.40	6.13	4.80	3.58
9	3.60	4.88	5.02	3.40	6.23	4.84	3.60
10	3.58	4.92	5.03	3.41	6.30	4.87	3.58
11	3.59	4.94	5.07	3.42	6.34	4.89	3.59
12	3.55	4.96	5.07	3.42	6.34	4.92	3.55
13	3.56	4.97	5.08	3.43	6.47	4.96	3.56
14	3.57	4.99	5.09	3.41	6.36	4.98	3.57
15	3.53	4.67	5.09	3.41	6.37	5.00	3.53
16	3.54	4.69	5.10	3.42	6.38	4.99	3.54
17	3.55	4.70	5.08	3.42	6.54	5.00	3.55
18	3.55	4.72	5.08	3.43	6.38	5.02	3.55
19	3.56	4.73	5.09	3.44	6.37	5.03	3.56
20	3.57	4.74	5.09	3.15	6.37	5.04	3.57

Tabla AIV. 11. Resultados del Método MOD(Desplazamiento Promedio). Base de datos BCN25

Orlado [m]	Oscilación por zonas [oscilaciones/km]						
	Almería	Cádiz	Co-Ja	Huelva	Málaga	Sevilla	Todas
1	3.66	3.67	5.32	4.04	6.64	3.67	4.60
2	2.25	2.78	4.21	2.49	4.62	2.47	3.25
3	1.86	2.16	3.48	1.69	3.78	1.76	2.58
4	1.76	1.87	3.20	1.38	3.57	1.60	2.35
5	1.66	1.82	3.07	1.30	3.37	1.51	2.23
6	1.55	1.75	2.98	1.27	3.28	1.45	2.16
7	1.49	1.68	2.93	1.27	3.25	1.39	2.11
8	1.50	1.64	2.92	1.24	3.18	1.36	2.09
9	1.50	1.62	2.89	1.23	3.20	1.31	2.07
10	1.50	1.61	2.86	1.23	3.12	1.29	2.05
11	1.50	1.61	2.84	1.22	3.08	1.29	2.04
12	1.49	1.60	2.84	1.22	3.06	1.29	2.03
13	1.49	1.59	2.84	1.22	3.07	1.29	2.03
14	1.48	1.59	2.84	1.22	3.05	1.29	2.03
15	1.48	1.59	2.84	1.22	3.04	1.29	2.02
16	1.48	1.59	2.84	1.22	3.03	1.29	2.02
17	1.48	1.59	2.84	1.22	3.03	1.29	2.02
18	1.47	1.59	2.84	1.22	3.02	1.28	2.02
19	1.47	1.59	2.84	1.22	3.03	1.28	2.02
20	1.47	1.59	2.84	1.22	3.03	1.27	2.02

Tabla AIV. 12. Resultados del Método MOD(Oscilación). Base de datos BCN25

Índices

Índice de Figuras

Figura II. 1. Elipsoide de inexactitud alrededor de un punto	18
Figura II. 2. Banda Épsilon	23
Figura II. 3. Banda de error	25
Figura II. 4. Simplificación bandas no paralelas.....	25
Figura II. 5. Banda obtenida por simulación.....	26
Figura II. 6. Comportamiento de la banda de error	26
Figura II. 7. Banda épsilon en poligonales y polígonos.....	27
Figura II. 8. Función de densidad de probabilidad del segmento AB.....	27
Figura II. 9. Visualización de la Banda-g de un segmento.....	28
Figura II. 10. Región de confianza en puntos extremos	28
Figura II. 11. Banda-g en una poligonal	29
Figura II. 12. Método de banda épsilon. Puntos finales coincidentes	31
Figura II. 13. Método de banda épsilon. Puntos finales no coincidentes	32
Figura II. 14. Orlado simple.....	33
Figura II. 15. Resultados obtenidos por Goodchild y Hunter, 1997	34
Figura II. 16. Resultados método de orlado simple Johnston y col. (1999)	36
Figura II. 17. Resultados globales test orlado Hoja MTN 947. Mozas (2003)	37
Figura II. 18. Método de doble orlado.....	38
Figura II. 19. Doble orlado: Comportamiento áreas. Tveite y Langaas (1999).....	39
Figura II. 20. Doble orlado: Desplazamiento promedio. Tveite y Langaas (1999)	40
Figura II. 21. Doble orlado: Oscilación. Tveite y Langaas (1999).....	41
Figura II. 22. Doble orlado: Compleción y omisiones. Tveite y Langaas (1999)	42
Figura II. 23. Resultados método doble orlado. Goeman y col. (2005).....	44
Figura II. 24. Método basado en las distancias de Hausdorff	45
Figura II. 25. Test MPS y MDS de Veregin	48
Figura II. 26. Test MPS y MDS: Distribución frecuencias acumuladas. Veregin (2000)	48

Figura II. 27. Método de correspondencia de puntos	51
Figura II. 28. Error de comisión y omisión en búfer	53
Figura II. 29. Discrepancia del búfer	55
Figura II. 30. Métodos de selección de muestras	56
Figura II. 31. Función de densidad de probabilidad. Distribución normal	62
Figura II. 32. Función de densidad de probabilidad. Distribución uniforme.....	62
Figura II. 33. Función de densidad de probabilidad. Distribución beta	63
Figura II. 34. Función de densidad de probabilidad. Distribución gamma	64
Figura II. 35. Función de densidad de probabilidad. Distribución Weibull	64
Figura II. 36. Función de densidad de probabilidad. Distribución exponencial	65
Figura III. 1. Procedimiento General.....	69
Figura III. 2. Esquema del proceso de obtención de la BD de control	75
Figura III. 3. Determinación de ejes de líneas de control	76
Figura III. 4. Eliminación de tramos	77
Figura III. 5. Identificación de líneas homólogas. Caso 1	78
Figura III. 6. Identificación de líneas homólogas. Caso 2	79
Figura III. 7. Filtrado por distancia	80
Figura III. 8. Automatización Método MDH.....	83
Figura III. 9. Automatización Método MSK.....	85
Figura III. 10. Automatización Método MOS	87
Figura III. 11. Automatización Método MOD	88
Figura III. 12. Ejemplo del criterio base del TKS. Press y Teukolsky (1992	92
Figura III. 13. Líneas sintéticas.....	96
Figura III. 14. Zona de estudio	100
Figura III. 15. Curvas hipsométricas de la zona de estudio	103
Figura III. 16. Líneas de control GPS. Almería	105
Figura III. 17. Líneas de control GPS. Córdoba-Jaén	105
Figura III. 18. Líneas de control GPS. Cádiz.....	106
Figura III. 19. Líneas de control GPS. Huelva	106
Figura III. 20. Líneas de control GPS. Málaga	107
Figura III. 21. Líneas de control GPS. Sevilla	107

Figura III. 22. Variograma para las diferencias de calidad posicional planimétrica de los datos GPS	107
Figura III. 23. Autocorrelaciones de los valores de calidad GPS	108
Figura III. 24. Histograma relativo al espaciado entre puntos	108
Figura III. 25. Histograma relativo a la velocidad entre puntos	108
Figura IV. 1. Ejemplo de aplicación del método MOS con CPLin.....	113
Figura IV. 2. Método MOS. Línea sintética nº1	117
Figura IV. 3. Método MOS. Línea sintética nº2	117
Figura IV. 4. Método MOS. Línea sintética nº3	117
Figura IV. 5. Método MOS. Línea sintética nº4	117
Figura IV. 6. Método MOS. Línea sintética nº5	118
Figura IV. 7. Método MOS. Línea sintética nº6	118
Figura IV. 8. Método MOS. Línea sintética nº7	118
Figura IV. 9. Método MOD(SC). Línea sintética nº1	120
Figura IV. 10. Método MOD(SC). Línea sintética nº2	120
Figura IV. 11. Método MOD(SC). Línea sintética nº3	120
Figura IV. 12. Método MOD(SC). Línea sintética nº4	120
Figura IV. 13. Método MOD(SC). Línea sintética nº5	121
Figura IV. 14. Método MOD(SC). Línea sintética nº6	121
Figura IV. 15. Método MOD(SC). Línea sintética nº7	121
Figura IV. 16. Método MOD(DP). Línea sintética nº1.....	123
Figura IV. 17. Método MOD(DP). Línea sintética nº2.....	123
Figura IV. 18. Método MOD(DP). Línea sintética nº3.....	123
Figura IV. 19. Método MOD(DP). Línea sintética nº4.....	123
Figura IV. 20. Método MOD(DP). Línea sintética nº5.....	123
Figura IV. 21. Método MOD(DP). Línea sintética nº6.....	123
Figura IV. 22. Método MOD(DP). Línea sintética nº7.....	124
Figura IV. 23. Método MOD(OS). Línea sintética nº1.....	125
Figura IV. 24. Método MOD(OS). Línea sintética nº2.....	125
Figura IV. 25. Método MOD(OS). Línea sintética nº3.....	126
Figura IV. 26. Método MOD(OS). Línea sintética nº4.....	126
Figura IV. 27. Método MOD(OS). Línea sintética nº5.....	126

Figura IV. 28. Método MOD(OS). Línea sintética nº6.....	126
Figura IV. 29. Método MOD(OS). Línea sintética nº7.....	126
Figura IV. 30. Líneas obtenidas tras el proceso de edición en el MTA	129
Figura IV. 31. Líneas obtenidas tras el proceso de edición en la BCN.....	129
Figura IV.32. Variogramas correspondientes a las diferencias de distancias entre MTA y GPS	130
Figura IV.33. Variogramas correspondientes a las diferencias de distancias entre BCN y GPS.....	130
Figura IV.34. Correlogramas correspondientes a los valores de distancias entre puntos del MTA y GPS	131
Figura IV.35. Correlogramas correspondientes a los valores de distancias entre puntos de la BCN y GPS	131
Figura IV. 36. Histogramas relativos a la distribución de casos por longitud. MTA.....	132
Figura IV. 37. Histogramas relativos a la distribución de casos por longitud. BCN	132
Figura IV. 38. Histogramas relativos a la distribución de longitudes por zona. MTA	132
Figura IV. 39. Histogramas relativos a la distribución de longitudes por zona. BCN	132
Figura IV. 40. Histogramas relativos a la distribución de longitudes por tipo de vía. MTA.....	134
Figura IV. 41. Histogramas relativos a la distribución de longitudes por tipo de vía. BCN	134
Figura IV. 42. Histograma relativo a la distribución según la orientación. MTA.....	135
Figura IV. 43. Histograma relativo a la distribución según la orientación. BCN	135
Figura IV. 44. Histograma relativo a la distribución según la orientación por zonas. MTA	136
Figura IV. 45. Histograma relativo a la distribución según la orientación por zonas. BCN.....	136
Figura IV. 46. Histograma relativo a la distribución de longitudes según sinuosidad. MTA	137
Figura IV. 47. Histograma relativo a la distribución de longitudes según sinuosidad. BCN	137
Figura IV. 48. Función de distribución MDH para distancias de Hausdorff. MTA	140
Figura IV. 49. Función de distribución MDH para distancias de Hausdorff. BCN	140

Figura IV. 50. Función de distribución MDH para distancias medias de los elementos. MTA	140
Figura IV. 51. Función de distribución MDH para distancias medias de los elementos. BCN	140
Figura IV. 52. Función de distribución MSK. MTA	142
Figura IV. 53. Función de distribución MSK. BCN.....	142
Figura IV. 54. Función de distribución del porcentaje de línea dentro del orlado. MTA	145
Figura IV. 55. Función de distribución del porcentaje de línea dentro del orlado. BCN.....	145
Figura IV. 56. Función de distribución del porcentaje vértices dentro del orlado. MTA	146
Figura IV. 57. Función de distribución del porcentaje vértices dentro del orlado. BCN.....	146
Figura IV. 58. Funciones de densidad de probabilidad de los resultados del método MOS	147
Figura IV. 59. Función de distribución de la superficie común. MTA	148
Figura IV. 60. Función de distribución de la superficie común. BCN.....	148
Figura IV. 61. Función de distribución del desplazamiento promedio. MTA.....	148
Figura IV. 62. Función de distribución del desplazamiento promedio. BCN	148
Figura IV. 63. Función de distribución la oscilación por longitud del tramo. MTA	149
Figura IV. 64. Función de distribución la oscilación por longitud del tramo. BCN	149
Figura IV. 65. Función de distribución la oscilación por kilómetro. MTA	149
Figura IV. 66. Función de distribución la oscilación por kilómetro. BCN	149
Figura IV. 67. Funciones de densidad de probabilidad de los resultados del método MOD(SC)	150
Figura IV. 68. Funciones de densidad de probabilidad de los resultados del método MOD(DP)	150
Figura IV. 69. Funciones de distribución MOS por zonas. MTA	154
Figura IV. 70. Funciones de distribución MOS por zonas. BCN	154
Figura IV. 71. Funciones de distribución MOS por tipo. MTA	156
Figura IV. 72. Funciones de distribución MOS por tipo. BCN.....	156

Figura IV. 73. Funciones de distribución MOS por longitud. MTA	157
Figura IV. 74. Funciones de distribución MOS por longitud. BCN.....	157
Figura IV. 75. Funciones de distribución MOS por orientación. MTA	160
Figura IV. 76. Funciones de distribución MOS por orientación. BCN.....	160
Figura IV. 77. Funciones de distribución MOS por sinuosidad. MTA	162
Figura IV. 78. Funciones de distribución MOS por sinuosidad. BCN.....	162
Figura IV. 79. Evolución de la estimación según MDH con el tamaño de la muestra. MTA	165
Figura IV. 80. Evolución de la estimación según MDH con el tamaño de la muestra. BCN	165
Figura IV. 81. Evolución de la estimación según MSK con el tamaño de la muestra. MTA.....	165
Figura IV. 82. Evolución de la estimación según MSK con el tamaño de la muestra. BCN	165
Figura IV. 83. Evolución de la estimación según MOS al 95% con el tamaño de la muestra. MTA.....	166
Figura IV. 84. Evolución de la estimación según MOS al 95% con el tamaño de la muestra. BCN	166
Figura IV. 85. Evolución de la estimación según MOS a la ZKS con el tamaño de la muestra. MTA.....	167
Figura IV. 86. Evolución de la estimación según MOS a la ZKS con el tamaño de la muestra. BCN	167
Figura IV. 87. Evolución de la estimación según MOS al p-valor con el tamaño de la muestra. MTA.....	167
Figura IV. 88. Evolución de la estimación según MOS al p-valor con el tamaño de la muestra. BCN	167
Figura IV. 89. Evolución de la estimación según MOD a 20m con el tamaño de la muestra. MTA.....	168
Figura IV. 90. Evolución de la estimación según MOS a 20m con el tamaño de la muestra. BCN	168
Figura IV. 91. Evolución de la estimación según MOD a la ZKS con el tamaño de la muestra. MTA.....	168

Figura IV. 92. Evolución de la estimación según MOD a la ZKS con el tamaño de la muestra. BCN	168
Figura IV. 93. Evolución de la estimación según MOD al p-valor con el tamaño de la muestra. MTA.....	169
Figura IV. 94. Evolución de la estimación según MOD al p-valor con el tamaño de la muestra. BCN	169
Figura IV. 95. Muestra Almería-Málaga-Sevilla	170
Figura IV. 96. Muestra Córdoba_Jaén-Cádiz-Huelva.....	170
Figura IV. 97. Evolución valores medios MDH con eliminación de vértices. MTA.....	173
Figura IV. 98. Evolución valores medios MDH con eliminación de vértices. BCN	173
Figura IV. 99. Evolución valores medios MSK con eliminación de vértices. MTA.....	174
Figura IV. 100. Evolución valores medios MSK con eliminación de vértices. BCN	174
Figura IV. 101. Funciones de distribución MOS con eliminación vértices de control. MTA	175
Figura IV. 102. Funciones de distribución MOS con eliminación vértices de control. BCN.....	175
Figura IV. 103. Funciones de distribución MOD(SC) con eliminación vértices de control. MTA	176
Figura IV. 104. Funciones de distribución MOD(SC) con eliminación vértices de control. BCN	176
Figura AIII. 1. Ajuste a distribución exponencial - MOS. MTA.....	216
Figura AIII. 2. Ajuste a distribución exponencial - MOS. BCN	216
Figura AIII. 3. Ajuste a distribución Gamma - MOS. MTA	216
Figura AIII. 4. Ajuste a distribución Gamma - MOS. BCN	216
Figura AIII. 5. Ajuste a distribución exponencial - superficie común. MTA.....	217
Figura AIII. 6. Ajuste a distribución exponencial - superficie común. BCN	217
Figura AIII. 7. Ajuste a distribución Gamma - superficie común. MTA	217
Figura AIII. 8. Ajuste a distribución Gamma - superficie común. BCN	217
Figura AIII. 9. Ajuste a distribución exponencial - desplazamiento promedio. MTA	218
Figura AIII. 10. Ajuste a distribución exponencial - desplazamiento promedio. BCN.....	218
Figura AIII. 11. Ajuste a distribución Gamma - desplazamiento promedio. MTA	218
Figura AIII. 12. Ajuste a distribución Gamma - desplazamiento promedio. BCN.....	218

Índice de Tablas

Tabla II. 1. Definiciones de errores circulares. Probabilidades	22
Tabla II. 2. Precisiones planimétricas usuales (Sevilla, 1991.....	23
Tabla II. 3. Resumen de los modelos de error posicional para elementos lineales	30
Tabla III. 1. Comparativa estudios con elementos lineales	82
Tabla III. 2. Resumen estadístico de distancia media entre vértices de los elementos lineales en el MTA	98
Tabla III. 3. Resumen estadístico de distancia media entre vértices de los elementos lineales en la BCN	99
Tabla III. 4. Listado de hojas del MTN50 que cubren la zona de estudio	102
Tabla III. 5. Resumen de características descriptivas de la zona de estudio	103
Tabla IV. 1. Resultados Método MDH líneas sintéticas	114
Tabla IV. 2. Resultados Método MSK líneas sintéticas	115
Tabla IV. 3. Resultados Método MOS líneas sintéticas	116
Tabla IV. 4. Resultados Método MOD (superficie común) líneas sintéticas	120
Tabla IV. 5. Resultados Método MOD (desplazamiento promedio) líneas sintéticas.....	122
Tabla IV. 6. Resultados Método MOD (oscilación) líneas sintéticas	125
Tabla IV. 7. Resultados del proceso de edición y depuración por tipo	128
Tabla IV. 8. Resultados del proceso de edición por zona	128
Tabla IV. 9. Distribución relativa de elementos por longitud.....	131
Tabla IV. 10. Distribución de los elementos por zonas. MTA	133
Tabla IV. 11. Distribución de los elementos por zonas. BCN.....	133
Tabla IV. 12. Distribución de los elementos por tipo de vía. MTA.....	134
Tabla IV. 13. Distribución de los elementos por tipo de vía. BCN	134
Tabla IV. 14. Distribución de los elementos por orientación. MTA	135
Tabla IV. 15. Distribución de los elementos por orientación. BCN	136
Tabla IV. 16. Distribución de los elementos por sinuosidad. MTA	137

Tabla IV. 17. Distribución de los elementos por sinuosidad. BCN	137
Tabla IV. 18. Resultados MDH para el MTA.....	139
Tabla IV. 19. Resultados MDH para la BCN	139
Tabla IV. 20. Resultados para distintos niveles de confianza MDH	141
Tabla IV. 21. Resultados MSK para el MTA.....	141
Tabla IV. 22. Resultados MSK para el BCN	142
Tabla IV. 23. Resultados para distintos niveles de confianza MSK	142
Tabla IV. 24. Resultados del número de cruces. MTA	143
Tabla IV. 25. Resultados del número de cruces. BCN.....	143
Tabla IV. 26. Resultados MOS(PL) para el MTA	144
Tabla IV. 27. Resultados MOS(PL) para la BCN.....	144
Tabla IV. 28. Resultados MOS(PV) para el MTA	145
Tabla IV. 29. Resultados MOS(PV) para la BCN	145
Tabla IV. 30. Resultados Kolmogorov-Smirnov de resultados de longitud y número de vértices	146
Tabla IV. 31. Resultados MOD Superficie común y desplazamiento promedio. MTA	147
Tabla IV. 32. Resultados MOD Superficie común y desplazamiento promedio. BCN.....	147
Tabla IV. 33. Resumen de resultados numéricos para cada método de control	152
Tabla IV. 34. Resultados similitud por zonas para el MTA	155
Tabla IV. 35. Resultados similitud por zonas para la BCN	155
Tabla IV. 36. Resultados similitud por tipo de vía para el MTA	156
Tabla IV. 37. Resultados similitud por tipo de vías para la BCN	157
Tabla IV. 38. Resultados similitud por longitud del tramo para el MTA	158
Tabla IV. 39. Resultados similitud por longitud del tramo para la BCN	159
Tabla IV. 40. Resultados similitud por orientación para el MTA.....	160
Tabla IV. 41. Resultados similitud por orientación para la BCN	161
Tabla IV. 42. Resultados similitud por sinuosidad para el MTA	162
Tabla IV. 43. Resultados similitud por sinuosidad para la BCN	162
Tabla IV. 44. Muestras para comprobación del tamaño de la muestra.....	170
Tabla IV. 45. Resultados muestras para comprobación del tamaño de muestras: MDH Y MSK.....	171

Tabla IV. 46. Resultados muestras para comprobación del tamaño de muestras:	
MOS Y MOD	171
Tabla IV. 47. Resultados similitud con MOS en eliminación de vértices de control	
para el MTA.....	175
Tabla IV. 48. Resultados similitud con MOS en eliminación de vértices de control	
para la BCN	176
Tabla IV. 49. Resultados similitud con MOD en eliminación de vértices de control	
para el MTA.....	177
Tabla IV. 50. Resultados similitud con MOD en eliminación de vértices de control	
para la BCN	177
Tabla AI. 1. Requerimientos de exactitud planimétrica. Mapa Clase I.....	202
Tabla AI. 2. Categorías de Exactitud para MDE. USGS	206
Tabla AIV. 1. Resultados del Método MDH. Base de datos MTA10.....	220
Tabla AIV. 2. Resultados del Método MDH. Base de datos BCN25	220
Tabla AIV. 3. Resultados del Método MSK. Base de datos MTA.....	221
Tabla AIV. 4. Resultados del Método MSK. Base de datos BCN25	221
Tabla AIV. 5. Resultados del Método MOS. Base de datos MTA10	223
Tabla AIV. 6. Resultados del Método MOS. Base de datos BCN25	225
Tabla AIV. 7. Resultados del Método MOD(Superficie Común).Base datos MTA10	226
Tabla AIV. 8. Resultados del Método MOD(Desplazamiento Promedio).	
Base de datos MTA10	227
Tabla AIV. 9. Resultados del Método MOD(Oscilación). Base de datos MTA10.....	228
Tabla AIV. 10. Resultados del Método MOD(Superficie Común). Base de	
datos BCN25	229
Tabla AIV. 11. Resultados del Método MOD(Desplazamiento Promedio).	
Base de datos BCN25	230
Tabla AIV. 12. Resultados del Método MOD(Oscilación). Base de datos BCN25	231

Acrónimos

Acrónimos

ASCI	<i>American Society of Civil Engineers</i>
ASPRS	<i>American Society of Photogrammetry and Remote Sensing</i>
BCN	<i>Base Cartográfica Numérica</i>
BDG	<i>Base de Datos Geográfica</i>
BOS	<i>Buffer Overlay Statistics Method</i>
CMSPE	<i>Error circular medio</i>
CMAS	<i>Estándar de exactitud circular del mapa</i>
ConPoCar	<i>Proyecto de investigación: "Control de calidad posicional de cartografía por elementos lineales"</i>
CPE	<i>Error circular probable</i>
CPLin	<i>Control posicional de líneas</i>
CSE	<i>Error circular estándar</i>
CSG	<i>Consejo Superior Geográfico</i>
DCW	<i>Digital Chart of the World</i>
DDU	<i>Distancia de distorsión uniforme</i>
DMA	<i>Defence Mapping Agency. Estados Unidos</i>
DOD	<i>Department of Defense. Estados Unidos</i>
EMAS	<i>Engineering Map Accuracy Standard</i>
FGDC	<i>Federal Geographic Data Committee</i>
GPS	<i>Sistema de Posicionamiento Global</i>
ICA	<i>Instituto de Cartografía de Andalucía</i>
IGN	<i>Instituto Cartográfico Nacional. España</i>
IGNF	<i>Institut Géographique National. Francia</i>
ITAM	<i>Integrated Training Area Management. Estados Unidos</i>
MDE	<i>Modelo digital de elevaciones</i>
MDS	<i>Estándar de máxima distorsión</i>
MDH	<i>Método de las Distancias de Hausdorff</i>
MOD	<i>Método del Orlado Doble</i>
MOD(SC)	<i>Método del Orlado Doble: Superficie común</i>
MOD(DP)	<i>Método del Orlado Doble: Desplazamiento promedio</i>
MOD(OS)	<i>Método del Orlado Doble: Oscilación</i>
MOS	<i>Método del Orlado Simple</i>
MPS	<i>Estándar de máxima proporción</i>
MSK	<i>Método de Skidmore o Banda Épsilon</i>
MTA	<i>Mapa Topográfico de Andalucía</i>
MTN	<i>Mapa Topográfico Nacional</i>
NCDCCDS	<i>National Committee for Digital Cartographic Data Standards</i>
NMA	<i>Norwegian Mapping Authority</i>

NMCA	<i>National Mapping Council of Australia</i>
NMAS	<i>National Map Accuracy Standard</i>
NSSDA	<i>National Standard for Spatial Data Accuracy</i>
RMSE	<i>Raíz del error cuadrático medio</i>
SIG	<i>Sistema de información geográfica</i>
USBB	<i>United States Bureau of the Budget</i>
USGS	<i>United States Geological Survey</i>
USGS DLG	<i>Digital Line Graphs del USGS</i>
USGS DRG	<i>Data Raster Graphics del USGS</i>