



UNIVERSIDAD DE JAÉN

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

TESIS DOCTORAL

**CONTROL DE CALIDAD POSICIONAL EN
CARTOGRAFÍA: ANÁLISIS DE LOS
PRINCIPALES ESTÁNDARES Y PROPUESTA
DE MEJORA**

**PRESENTADA POR:
ALAN DAVID JAMES ATKINSON GORDO**

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

JAÉN, 18 DE JULIO DE 2005

ISBN 978-84-8439-415-0

Nombre y apellidos del autor:

ALAN DAVID JAMES ATKINSON GORDO

Título de la Tesis Doctoral:

CONTROL DE CALIDAD POSICIONAL EN CARTOGRAFÍA: ANÁLISIS DE LOS PRINCIPALES ESTÁNDARES Y PROPUESTA DE MEJORA

I.S.B.N.:

978-84-8439-415-0

Fecha de Lectura:

18 DE JULIO DE 2005

Centro y Departamento en que fue realizada la lectura:

ESCUELA POLITÉCNICA SUPERIOR DE JAÉN
Departamento de Ingeniería Cartográfica, Geodesia y Fotogrametría

Composición del Tribunal/Dirección de la Tesis:

Dirección de la Tesis	Dr. D. Francisco J. Ariza López
Presidente/a del Tribunal	Dr. D. Rafael Ferrer Torio
Vocales	Dr. D. Luis Parra Guijosa Dr. D. Sebastián Mas Mayoral Dr. D. Carlos Pinilla Ruiz
Secretario/a	Dr. D. Tomás Fernández del Castillo

Calificación Obtenida:

SOBRESALIENTE CUM LAUDE



UNIVERSIDAD DE JAÉN

tesis doctoral



Resumen

Desde siempre, la exactitud posicional se ha considerado un aspecto definitorio y primordial de todo producto cartográfico. Así, la posición afecta a la geometría, topología y a los aspectos temáticos que se reflejan en dicho modelo de la realidad. Pero la posición también tiene una importancia mayúscula en la integración o interoperabilidad entre fuentes cartográficas.

La principal herramienta empleada es la simulación, con todas sus ventajas e inconvenientes, pero que ha servido como base en el estudio y análisis comparativo del comportamiento de los diferentes estándares así como para la generación de curvas operativas.

En la tesis se analizan los principales estándares de control de calidad posicional planimétrica en cartografía: NMAS (National Map Accuracy Standard, 1947), EMAS (Engineering Map Accuracy Standard, 1991), ASPRS (Spatial Accuracy Specification for Large Scale Topographic Maps, 1983) y NSSDA (National Standard for Spatial Data Accuracy, 1999). Asimismo, se realiza un análisis entre los diferentes estimadores robustos y la estadística clásica, los estimadores circulares, la detección y corrección de sistematismos mediante ajuste por mínimos cuadrados y la aplicación de curvas operativas en procesos de aceptación de productos cartográficos.

Por último, se efectúa una propuesta metodológica de aplicación en base a los resultados obtenidos durante la investigación.

Abstract

The positional accuracy has been always considered a distinctive and essential aspect in all the cartography products.

The main tool is the simulation, with its advantages and disadvantages, but it has been used as a base in the study and comparative analysis of the different standards behavior and to the operative curves.

The main standards of the planimetric positional quality control are analyzed in the thesis: NMAS (National Map Accuracy Standard, 1947), EMAS (Engineering Map Accuracy Standard, 1991), ASPRS (Spatial Accuracy Specification for Large Scale Topographic Maps, 1983) and NSSDA (National Standard for Spatial Data Accuracy, 1999). In the same sense a analysis among the different robust estimators and the classic statistics, the circular estimators, the detection application and correction of the systematic errors by means of last squares and the application of operative curves in processes of cartographic products acceptance.

Finally, a methodological proposal of application based on the results obtained during the investigation.



UNIVERSIDAD DE JAÉN

CONTROL DE CALIDAD
POSICIONAL EN CARTOGRAFÍA:
ANÁLISIS DE LOS PRINCIPALES
ESTÁNDARES Y PROPUESTA
DE MEJORA

ALAN D.J. ATKINSON GORDO

tesis doctoral

Índice

CONTENIDO	Pág.
Memoria	
Índice	2
Agradecimientos	7
Capítulo 1: Introducción	9
Capítulo 2: Objetivos	16
Capítulo 3: Antecedentes	19
3.1 Calidad y su aplicación a la cartografía	21
3.1.1 Aseguramiento y gestión de la calidad	21
3.1.2 Los costes de la mala calidad	27
3.1.3 La calidad en cartografía	31
3.1.4 La calidad en las Agencias Nacionales de Cartografía. Encuestas	36
3.2 Estándares en el control de calidad posicional	43
3.2.1 NMAS - National Map Accuracy Standard	47
3.2.2 EMAS - Engineering Map Accuracy Standard	48
3.2.3 ASPRS – Spatial Accuracy Specification for Large Scale Topographic Maps	51
3.2.4 USGS de categorías de exactitud de MDT	53
3.2.5 Fórmula de Koppe	54
3.2.6 NSSDA - National Standard for Spatial Data Accuracy	57
3.2.7 Consideraciones sobre los estándares posicionales	60
3.3 Técnicas estadísticas	62
3.3.1 Estadística robusta	63
3.3.1.1 MVT – MLT	68
3.3.1.2 Suavizado gaussiano	69
3.3.1.3 Método danés	71
3.3.1.4 Método de Geman & Mc.Clure	72
3.3.1.5 Consideraciones sobre los métodos robustos	73
3.3.2 Estadística circular	73
3.3.2.1 Introducción	74
3.3.2.2 Test Ómnibus	78
3.3.2.3 Test de Mardia	79
3.3.2.4 Dirección predominante	81
3.3.2.5 Estudio de los módulos	81
3.3.2.6 Consideraciones sobre los estimadores circulares	82
3.4 La simulación como herramienta en el análisis de la calidad	83
3.4.1 Números aleatorios – pseudoaleatorios	84
3.4.2 Contraste de aleatoriedad	88
3.4.3 Contraste de normalidad	90
3.4.4 Tamaño muestral	93
3.4.5 Consideraciones sobre la simulación	94
3.5 El ajuste mínimo cuadrático como herramienta para la corrección de errores sistemáticos	95
3.5.1 Transformación afín 2D	96
3.5.2 Consideraciones sobre el ajuste por mínimos cuadrados	100
3.6 Empleo de curvas operativas en procesos de control de calidad	101
3.6.1 Curvas operativas	104
3.6.2 Muestreo por variables	107
3.6.3 Consideraciones sobre el empleo de curvas operativas	108

CONTENIDO	Pág.
Capítulo 4: Material y método	110
4.1 El proceso de simulación empleado	111
4.1.1 Características de los estándares para el control posicional cartográfico	112
4.1.2 Simulación en la generación de poblaciones	113
4.1.3 Simulación en la generación de muestras	115
4.1.4 Simulación en la contaminación de poblaciones	119
4.1.5 Desarrollo de una aplicación propia para el estudio teórico: Cartocal 2.0	125
4.2 Datos reales tomados sobre cartografía	128
4.2.1 La cartografía analizada	129
4.2.2 El sistema GPS	131
4.2.3 Planificación y toma de datos	132
4.3 Aplicación al control posicional	136
4.3.1 Ajuste de los estimadores robustos	137
4.3.1.1 Proceso iterativo del estimador MLT	137
4.3.1.2 Suavizado gaussiano	139
4.3.1.3 Método danés	139
4.3.1.4 Método de Geman & Mc.Clure	140
4.3.2 Estándares en el control de calidad posicional: proceso comparativo mediante simulación.	141
4.3.3 Aplicación de procesos de simulación para la obtención de curvas operativas.	145
4.3.4 Cuantificación de errores sistemáticos mediante ajuste por mínimos cuadrados	148
4.3.5 Aplicación de estimadores circulares	152
Capítulo 5: Resultados y discusión	154
5.1 Comparativa entre los diferentes estimadores robustos y la estadística clásica	156
5.2 Análisis comparativo de los diferentes estándares en el control de calidad posicional	160
5.2.1 NMAS	161
5.2.2 EMAS	165
5.2.3 ASPRS	169
5.2.4 NSSDA	174
5.3 Corrección mediante ajuste mínimo cuadrático de errores sistemáticos	183
5.4 Estimadores circulares	192
5.4.1 Test Ómnibus	192
5.4.2 Test de Mardia	194
5.4.3 Dirección predominante	194
5.4.4 Simulación de muestras	195
5.4.5 Análisis de los módulos	202
5.5 Aplicación de los estándares mediante curvas operativas	211
5.5.1 EMAS	212
5.5.2 NMAS	216
5.5.3 ASPRS	220
5.5.4 NDDSA	225
5.5.5 Estimadores circulares	234
5.6 Propuesta metodológica	244

CONTENIDO	Pág.
Capítulo 6: Conclusiones	246
6.1 Conclusiones generales	247
6.1.1 Procesos de simulación	248
6.1.2 Estándares de exactitud posicional analizados.	248
6.1.3 Empleo de estimadores robustos.	250
6.1.4 Cuantificación de sistematismos.	251
6.1.5 Estimadores circulares	252
6.1.6 Empleo de curvas operativas.	252
6.1.7 Aplicación en conjunto.	255
6.2 Perspectivas de futuro	256
6.2.1 Perspectivas del mercado	256
6.3.2 Perspectivas de investigación	257
Referencias	259

CONTENIDO	Pág.
Anexos	
A1 Encuestas sobre calidad en cartografía	267
A1.1 Carta remitida	269
A1.2 Encuesta enviada	270
A1.3 Algunas respuestas recibidas	273
A2 Procedimientos para el control de calidad posicional	275
A2.1 Selección de puntos de control	275
A2.2 Levantamiento de puntos de control	277
A2.3 Digitalización de puntos de control	278
A3 Datos tomados en campo mediante GPS	280
A3.1 Selección de los puntos de control	280
A3.2 Parámetros de transformación	283
A3.3 Toma de puntos de control y cálculo de las observaciones	293
A3.4 Reseñas de campo	294
A4 Digitalización	297
A4.1 La digitalización	297
A4.2 El tratamiento previo de los datos	299
A5 Software I+D: Cartocal 2.0	303
A5.1 Manual del usuario	303
A5.2 Manual del programador	329
A5.3 Aplicaciones Autolisp	330
A6 Propuesta metodológica: ExHor	332
A6.1 Propuesta de aplicación	332
A6.2 Anexos a la propuesta de aplicación	341
A6.2.1 Puntos de control	341
A6.2.2 Normalidad	343
A6.2.3 Aleatoriedad	346
A6.2.4 Valores atípicos	347
A6.2.5 Sistematismos	347
A6.2.6 Desviación poblacional a posteriori	349
A6.2.7 Exactitud posicional	350
A6.3 Ejemplo de aplicación	351
A7 Listados en formato digital	358
A8 Planos	359
Hoja (926) 1-1 del ICA	
Hoja (926) 1-2 del ICA	
A9 Glosario	360

Agradecimientos

En el largo, duro y, sin embargo, gratificante trabajo que se presenta en esta tesis doctoral quiero mostrar mi agradecimiento a muchos organismos e instituciones, amigos y compañeros que la han hecho posible. Especialmente, quisiera recordar en este momento a:

- ▶ Todos aquellos organismos y personas que colaboraron respondiendo amablemente a las encuestas enviadas. También quiero agradecer a José Lázaro Amaromellado su ayuda en esta tarea.
- ▶ Antonio R. Alcalá, por su ayuda en la codirección de mi PFC en IGC, y su ánimo y apoyo continuo en la utilización de los sistemas de posicionamiento.
- ▶ Elidia Beatriz Blázquez, por su apoyo y sus sugerencias.
- ▶ José Luis García Balboa, por gran amistad, apoyo y ánimo continuo, así como por todas sus sugerencias y comentarios que han contribuido a mejorar el trabajo. Gracias a Cinta.
- ▶ Ramón Martínez Carrasco y Magdalena Nuño Palacio, por su amistad y ayuda tanto durante la tesis como en mi PFC.
- ▶ M^a del Mar Pacomio Peña, especialmente por todo lo que me ha dado y ha soportado desde el comienzo de la tesis hasta casi su conclusión.
- ▶ José Rodríguez Avi, por sus sugerencias, ideas y mejoras en cuanto a las herramientas estadísticas empleadas.
- ▶ José Juan de Sanjosé Blasco, por su ayuda y la revisión del software. Especialmente por su gran amistad y apoyo mientras ha durado la realización de la tesis, al tener que llevar una mayor carga de trabajo en el Grupo de Investigación en Ingeniería Geomática y Patrimonio Urbano de la Universidad de Extremadura. Gracias a Carmen.
- ▶ Manuel A. Ureña Cámara, por sus sugerencias y ayuda en la optimización de parte del software utilizado.
- ▶ Al mis amigos y compañeros de Zarautz, Jaén, Madrid y Extremadura, por su apoyo durante estos años.

No cabe duda de que todo este trabajo jamás hubiera sido posible sin la inestimable guía, ayuda, colaboración y ... ese empuje y ánimo que, más bien como amigo que

como director, siempre me ha proporcionado el Dr. **Francisco J. Ariza López**. Muchas gracias Pachi por tu **paciencia** y **ayuda**.

Tampoco quisiera olvidarme en este momento de mi familia y que, junto a otras circunstancias, el hecho de hacer esta tesis me ha mantenido más alejado de ellos de lo que hubiera deseado. Especialmente, gracias a mi madre por el apoyo y ánimo continuo desde que comencé mis estudios allá por el año 1987 en Ingeniería Técnica en Topografía.

Y, especialmente, a **Lucila** por la nueva vida que hemos comenzado.

Alan. D.J. Atkinson.-

Capítulo 1

Introducción

INTRODUCCIÓN

En la última década del milenio, la competitividad ha dado un salto cualitativo con la apertura de los mercados internacionales y el avance imparable de las telecomunicaciones. En éste ámbito, serán los productos con mejor relación CALIDAD – PRECIO los que obtengan mayores beneficios y reconocimiento. Así, como no podía ser de otra manera, se tendrán que reducir costes y aumentar la calidad de los productos que se realicen si se quieren mantener en el mercado.

Al hablar de *calidad* en general, nos referiremos a la “propiedad o conjunto de propiedades inherentes a una cosa, que permiten apreciarla como igual, mejor o peor que las restantes de su especie”. O en sentido absoluto “buena calidad, superioridad o excelencia”, tal y como nos lo define el Diccionario de la Lengua Española. Al analizar dicha definición, se aprecia que se basa en la comparación de un determinado producto con otros similares,

Junto a este concepto se habla también de *gestión de la calidad*, entendiéndose por ello aquellas actividades, englobadas en la gestión general, que tienen por objeto implantar y determinar la política de la calidad, los objetivos y las responsabilidades mediante la planificación de la calidad, el control de la calidad, el aseguramiento de la calidad y la mejora de la calidad. El *aseguramiento de la calidad* se refiere al conjunto de acciones planificadas y sistemáticas que son necesarias para proporcionar la confianza de que un producto o servicio satisface los requisitos referentes a la calidad. Por tanto, un *sistema de calidad* es el conjunto de la estructura organizativa, de responsabilidades, de procedimientos, de procesos y de recursos que se establecen para llevar a cabo la gestión de la calidad. Las normas ISO, serie 9000, establecen un marco para el control de la calidad en cualquier producto o proceso.

La *planificación de la calidad* es una actividad que se puede considerar innata en todo trabajo cartográfico o geodésico dado que desde el desarrollo científico de estas ciencias, sus aplicaciones siempre han partido de las premisas de obtener unos errores posicionales acotados por unas tolerancias dependientes de los métodos de levantamiento.

Al hablar de *control de calidad*, se trata del proceso seguido por una empresa u organismo¹ para asegurarse de que sus productos o servicios cumplen con los requisitos mínimos de calidad, establecidos por la propia empresa. Con la política de *Gestión (o administración) de Calidad Óptima (GCO)*, toda la organización y actividad de la empresa está sometida a un estricto control de calidad, ya sea de los procesos productivos como de los productos finales. En el caso de producción de bienes, la GCO implica que tanto el diseño, como la producción y la venta, la calidad de los materiales utilizados y los procesos seguidos se ajustan a unos patrones de calidad establecidos con antelación (Encarta, 1998). El control de la calidad, tal y como se entiende en la actualidad, se ha desarrollado desde principios del siglo pasado centrado en las actividades industriales productivistas y ha generado numerosas técnicas de control estadístico de los procesos. Uno de los aspectos más importantes del control estadístico de la calidad, es lo que se denomina el *control en la aceptación de productos*.

Los procesos de control de calidad cartográficos han de poseer unas características específicas del área en el que se enmarcan. Con los medios informáticos que se disponen hoy en día, no es de extrañar que existan numerosas empresas de diferentes sectores que emplean Bases de Datos Geográficas (BDG²): empresas de telecomunicaciones, empresas de distribución y logística, organismos oficiales (ayuntamientos, hacienda, consejerías autonómicas...), y un largo etcétera. Toda esta producción de cartografía en muy diferentes formatos hace que no siempre se conozca la calidad de dichos productos.

Los actuales sistemas de procesado y tratamiento digital de la información permiten trabajar con una gran cantidad de datos georreferenciados, tal y como sucede en los Sistemas de Información Geográfica (SIG). Así, se han llegado a generar grandes sistemas de información con los pies de barro, dado que los datos soportados no son todo lo adecuados que cabría esperar. Esta situación, que puede llevar a una falta de interés o transparencia a la hora de dar a conocer la verdadera calidad de los datos, se ha unido, en cierta medida, con una especie de sentir generalizado de que el dato geográfico digital era bueno por el mero hecho de ser digital, como si esta

¹ En el texto se emplean indistintamente los términos empresa, organismo u organización (ver Anexo 9 Glosario)

² En este texto, los términos BDG y mapa son utilizados indistintamente (Anexo 9)

circunstancia de su almacenamiento lo redimiera de todos sus posibles vicios (Ariza, 2002).

Por ejemplo, en el caso de querer implementar un Sistema de Posicionamiento Global por satélite (GPS), se ha de tener muy en cuenta la cartografía a emplear en el SIG. No es de extrañar el encontrarse con casos como el analizado por el Minnesota Department of Transportation (MPLMIC, 1999): sobre una BDG en formato digital, obtenida a partir de cartografía a escala 1/100.000, con vías de comunicación, lagos, ríos, etcétera, se localizó la situación de un vehículo. Su posición se obtuvo mediante navegadores G.P.S. en tiempo real. En este caso, se puede observar en la figura 1.1 cómo la supuesta localización del vehículo lo sitúa en mitad de un lago.



Figura 1.1 – Supuesta localización de un vehículo terrestre en mitad de un lago.

Ante esta situación se plantearon la siguiente cuestión: ¿qué datos son correctos: la BDG o la posición del punto por GPS? Dado que no se dispone de información acerca de la exactitud de ninguno de ellos, en un principio podría suponerse que los puntos suministrados por el receptor GPS son bastante más precisos que la información cartográfica a 1/100.000. Sin embargo en este caso la suposición es errónea: el navegador GPS tenía una precisión de 100 metros, mientras que la cartografía empleada se asume que poseía una exactitud de 50 metros porque fue realizada en base al estándar de exactitud posicional NMAS³.

El dilema aparece al no tener suficientes datos para decidir qué tipo de información es incorrecta y poder actuar en consecuencia. Resulta de vital importancia poseer información contrastada sobre la calidad de los productos empleados.

³ El National Map Accuracy Standard (Veregin, 1999) se presenta en el capítulo 3 (3.2.1)

Dado que los actuales sistemas de procesado y tratamiento digital de la información gráfica permiten trabajar con una gran cantidad de datos georreferenciados, es de vital importancia que se asocie la información a sus parámetros de calidad intrínsecos.

La utilidad de todo el trabajo podrá estar hipotecada si se emplea una BDG georreferenciada con una exactitud inferior a la deseada. Por todo ello, son imprescindibles los metadatos⁴ adjuntos a todo documento para poder utilizarlo correctamente: una BDG no será buena o mala, sino que estará correcta o incorrectamente empleada.

La tesis doctoral se centra en uno de los aspectos de mayor importancia de una BDG: la calidad posicional. Si bien intervienen otras componentes (exactitud temática, consistencia lógica ...), es bien conocido que un producto cartográfico es, por antonomasia, la georreferenciación de la información. Si dicha georreferenciación no es adecuada, el producto final tampoco lo será.

Por otra parte, las características propias del dato geográfico hacen que exista un tratamiento claramente diferenciado entre su componente planimétrica (XY) y altimétrica (Z). Éstas podrían quedar resumidas en que:

- Los métodos empleados en de captura de la componente Z difieren sustancialmente de la componente XY .
- Generalmente, la incertidumbre en la captura de la información altimétrica es superior a la planimétrica (p.e.: fotogrametría, teledetección, topografía clásica, GPS, ...).
- Tradicionalmente, la altimetría se ha “representado” mediante simbología (curvas de nivel, sombreado, tintas hipsométricas, ...), debiendo “leerse” en vez de medirse. Así, para obtener el valor de la componente Z entre dos curvas de nivel, hay que interpolar en función de la distancia planimétrica a la que se encuentre el punto objeto de interés, suponiendo que el terreno es constante y homogéneo entre ambas curvas de nivel.

⁴ Información sobre un conjunto o conjuntos de datos (Anexo 9).

- Por el contrario, la planimetría se puede “medir” al representarse en su totalidad⁵ sobre la BDG.

Si bien gran parte de estos aspectos son directamente aplicables a la cartografía en formato analógico, también es cierto que existe un elevado número de productos en formato digital que no contemplan la componente altimétrica o que ésta aparece representada como si de cartografía analógica se tratase (p.e. mediante curvas de nivel en 2D).

Por otro lado, el tratamiento de la información altimétrica en BDGs en formato digital (MDE⁶), se realiza de forma sustancialmente diferente. Hay que tener en cuenta aspectos como: algoritmo para la generación del MDE (triangulación, malla regular, ...), tratamiento previo en la detección de valores atípicos, generalización de MDEs, representación de la altimetría en la BDG (sin ningún elemento gráfico visible en vistas de planta, mediante sólidos 3D o texturas para vistas diferentes a la planta, por curvas de nivel, tintas hipsométricas sobre el MDE, triangulación, ...), etcétera.

Todo ello hace plantear la necesidad de centrar la investigación sobre aquellos aspectos que se consideren de un mayor interés. Así, se ha optado por el análisis y el tratamiento de la componente planimétrica del dato geográfico.

Cabe destacar que la principal herramienta empleada es la simulación, con todas sus ventajas e inconvenientes, pero que ha servido como base en el estudio y análisis comparativo del comportamiento de los diferentes estándares así como para la generación de curvas operativas.

El documento se ha estructurado en 6 capítulos y un bloque de 9 anexos. En el segundo capítulo se presentan los principales objetivos a alcanzar en la tesis. En siguiente lugar se estudia el estado actual en la materia en el capítulo dedicado a los antecedentes, desglosando en seis apartados: una primera parte sobre el control de calidad a nivel genérico y su aplicación a la cartografía, el resultado de encuestas a nivel nacional e internacional, los diferentes estándares de control de calidad

⁵ El nivel de representación vendrá dado por la escala, el tipo de cartografía (base, temática...) y la utilización que se vaya a hacer de la BDG.

⁶ Modelo Digital de Elevaciones

posicional sobre entidades puntuales, las herramientas estadísticas a emplear, las herramientas basadas en procesos de simulación, el ajuste mínimo cuadrático y, por último, la utilización de curvas operativas en procesos de control de calidad.

En el cuarto capítulo, material y método, se muestran las herramientas empleadas: simulación y datos reales tomados sobre el terreno. Asimismo, se hace referencia a la principal herramienta del trabajo: el software CARTOCAL, desarrollado bajo entorno Microsoft Visual Basic 6.

En el quinto capítulo se exponen los resultados alcanzados en las diferentes partes del estudio: estimadores robustos, comparativa de los estándares más empleados en el control de calidad posicional, aplicación mediante curvas operativas, estudio geométrico, corrección y detección de sistematismos mediante ajuste mínimo cuadrático, y propuesta de un nuevo estándar.

En el capítulo sexto se muestran brevemente las conclusiones a las que llega el estudio.

Por último, los anexos del documento se estructuran en 9 apartados: encuestas sobre calidad en cartografía, procedimientos para el control de calidad, datos de campo, digitalización, software para simulación, propuesta metodológica, listados en formato digital⁷, cartografía analizada y glosario. En éste último, para evitar posibles dudas, se presenta la definición con la que se emplean numerosos términos que se utilizan a lo largo del trabajo.

⁷ Se adjunta CD-ROM con un índice en formato HTM y documentos en formato PDF.

Capítulo 2

Objetivos

2 Objetivos

Tal y como se anuncia en el título de la tesis, “*Control de calidad posicional en cartografía: análisis de los principales estándares y propuesta de mejora*”, los objetivos principales del trabajo de investigación de pueden definir en las líneas de trabajo:

2.1 Análisis de los principales estándares

Dentro del control de calidad cartográfico, se pretende ahondar en el conocimiento en el estado en la materia. Para ello se han estudiado trabajos realizados a nivel nacional e internacional sobre qué parámetros de calidad se están empleando en cartografía y sus perspectivas de futuro. Asimismo, se pretende conocer cuál es el estado del sector a nivel nacional e internacional mediante el diseño y envío de una encuesta en la materia.

En cuanto a los principales estándares que se han utilizado y actualmente se emplean para el control de la exactitud posicional, se pretende:

- Conocer su existencia, peculiaridades y funcionamiento.
- Realizar un análisis crítico de sus posibles deficiencias.
- Estudio mediante procesos de simulación del comportamiento de los estándares ante diferentes tipos de errores y distinta magnitud: outliers, sistematismos, varianza, escala a priori de la cartografía, estimadores circulares ...
- Análisis de los posibles riesgos de usuario y productor a la hora de emplear uno u otro estándar.
- Verificar si el tamaño de la muestra de control es adecuado a los resultados que, presumiblemente, proporcionan los estándares.

2.2 Propuesta de mejora

En esta segunda línea se plantean modificaciones a los procedimientos empleados actualmente. Dichos cambios van orientados a:

- Mejorar los actuales estándares en aquellos casos en los que sea posible (tamaño muestral, tolerancia, formato, escala...).
- Proponer en qué situaciones es o no recomendable emplear uno u otro estándar.
- Plantear la posibilidad de emplear algún tipo de estimador robusto con el que tratar los valores atípicos que el usuario se podría encontrar al analizar el producto. Proponer un procedimiento para su aplicación en el control de la exactitud posicional.
- Plantear la posibilidad de emplear algún estimador circular.
- Estudiar las posibilidades para que, en función del tipo de error detectado, éste pueda ser corregido o minimizado.
- Trasladar el concepto de curva operativa, empleado en otros sectores, al ámbito de la calidad posicional en cartografía, así como la posible determinación de los riesgos de usuario y productor.
- Plantear una propuesta en conjunto como estándar de exactitud posicional sobre elementos puntuales con el que proporcionar al usuario la mayor información posible acerca de la calidad del producto.

Capítulo 3

Antecedentes

3 Antecedentes

En este capítulo se presenta el estado del arte en lo que se refiere a la calidad en la cartografía, centrándose en aquellos aspectos y posibles herramientas que se podían aplicar al control posicional.

El capítulo se divide en 6 apartados:

- 1º. Calidad y su aplicación a la cartografía. Con la exposición de una serie de conceptos básicos necesarios para comprender las herramientas y los procesos empleados en el control de calidad, así como su aplicación en productos cartográficos.
- 2º. Estándares en el control de calidad posicional. Se presentan los principales estándares que se han empleado desde mediados del siglo pasado en el control de calidad cartográfico.
- 3º. Técnicas estadísticas. Se exponen diferentes técnicas que pueden aplicarse al control de calidad cartográfico, como podría ser el caso de la estadística robusta y la estadística circular.
- 4º. La simulación como herramienta en el análisis de la calidad. En este punto, se presenta una de las principales herramientas a emplear en el estudio de los estándares para el control de calidad cartográfico.
- 5º. El ajuste mínimo cuadrático como herramienta de corrección de errores sistemáticos. Si bien los mínimos cuadrados se emplean en otras áreas de la geomática (geodesia, topografía...), aquí son empleados para la cuantificación de errores posicionales sistemáticos.
- 6º. Empleo de curvas operativas en procesos de control de calidad. Por último, se muestran las curvas operativas como una herramienta más en los procesos de control de calidad.

3.1 Calidad y su aplicación a la cartografía

En este primer apartado se exponen aquellos aspectos más importantes referidos al aseguramiento y la gestión de la calidad, los costes derivados de la mala calidad, la calidad en la cartografía y su aplicación en las Agencias Nacionales de Cartografía.

3.1.1 Aseguramiento y gestión de la calidad

Al hablar de calidad, una gran mayoría lo asocia al cumplimiento de una determinada empresa con una normativa específica. La más conocida es la norma ISO 9000:2000.

Uno de los factores esenciales en el funcionamiento de la organización es la calidad de sus productos o servicios. Existe una tendencia mundial por parte de los clientes, hacia requisitos más exigentes respecto a la calidad. Al mismo tiempo se está produciendo una creciente toma de conciencia de que, para obtener de forma continuada unos buenos rendimientos económicos, es necesario, con frecuencia, mejorar la calidad de forma continua. Para que una organización sea reconocida por la calidad de sus productos o servicios, tiene que funcionar con eficacia, con método y con sistema, dentro de una norma que dé garantía al usuario de los productos o servicios.

Cuando se crea un producto o un servicio, se hace para satisfacer las necesidades y requisitos de los clientes que, generalmente, se traducen en forma de especificaciones. Sin embargo, las especificaciones técnicas no pueden, por sí solas, garantizar que se cumplirán de manera efectiva los requisitos del cliente, ya que pueden producirse deficiencias en las propias especificaciones o en el sistema organizativo. Por ello, se han desarrollado normas de sistemas de calidad y guías que contemplan los requisitos establecidos en las especificaciones técnicas del producto o servicio. La serie de Normas Internacionales ISO 9000 establece una racionalización de los numerosos y variados enfoques nacionales en este campo. Las normas nacen para que las empresas se rijan por unos principios de organización, y para que den estabilidad en el mercado y la sociedad.

Un Sistema de Gestión de la Calidad (SGC) es aquella parte del sistema de gestión de la organización que está enfocada al logro de los resultados en relación a los objetivos de la calidad, para satisfacer las necesidades, expectativas y requisitos de las partes interesadas (ISO 9000:2000). Para que exista una clientela, es necesario hacer una prospección del mercado, saber lo que las personas necesitan, y diseñar un producto o servicio acorde a esas expectativas. La satisfacción del cliente garantiza la continuidad de la organización.

En una determinada empresa, el SGC estará influenciado por los objetivos de la propia organización, sus productos o servicios y su forma de trabajar o producir, por lo que el sistema variará en función de la empresa. No obstante, los SGC quedan básicamente definidos según el modelo establecido como estándar a nivel mundial, desarrollado por la Organización Internacional de Normalización (ISO) y que se describe en las normas ISO de la serie 9000, revisadas a finales del año 2000. Estas normas han sido adoptadas por el Comité Europeo de Normalización (CEN), como Normas Europeas EN 29000 y en España, por la Asociación Española de Normalización y Certificación (AENOR), como normas UNE 66-900.

No cabe duda, que el principal objetivo de las normas ISO es el proporcionar a la empresa que las aplique una serie de herramientas que le permitirán una mejora continua al ofrecer productos de alta calidad, así como poder detectar y eliminar costes superfluos, costes de control y reparación, mejorando de forma continua sus relaciones con los clientes que son, en realidad, quienes hacen posible la supervivencia y rentabilidad de la empresa.

La familia de normas ISO 9000:2000 establece, a la hora de implementar un sistema de gestión de la calidad, el formato general de aplicación en cualquier empresa, independiente de su tamaño y tipo. Éstas son:

- ISO 9000:2000. Establece los **fundamentos** de los sistemas de gestión y la terminología a emplear para evitar ambigüedades.
- ISO 9001:2000. Especifica los **requisitos** para los **SGC**, siendo una propuesta genérica aplicable a cualquier caso. Se trata de la propuesta ISO para aquellas organizaciones con fines contractuales o de certificación, aunque también

puede emplearse a nivel interno. En esta ocasión, los requisitos de los productos quedan fuera del alcance de la normativa.

- ISO 9004:2000. Junto con la ISO 9001:2000, incluye amplias **recomendaciones** sobre los objetivos para un SGC. Éstas se orientan a la mejora de las prestaciones globales de una organización en el camino hacia la "excelencia en los negocios" (mejora del desempeño de la organización, satisfacción al cliente y demás partes implicadas). No posee carácter contractual y se recomienda su uso en conjunto con ISO 9001:2000 puesto que amplía los requerimientos mínimos establecidos por ésta.
- ISO 19011:2000. Proporciona orientación relativa a las auditorías de sistemas de gestión de la calidad y de gestión medio ambiental.

Estas normas son genéricas e independientes de cada organización. El diseño e implantación de un SGC depende de las necesidades concretas de la organización, sus objetivos particulares, productos y servicios suministrados, y de los procesos y prácticas empleadas. El objeto de estas normas internacionales no es imponer la uniformidad de los SGC, sino establecer un modelo adaptable a cualquier tipo de empresa o institución. En este sentido cabe destacar la nueva versión de ISO 9000, está más alineada que las anteriores ISO 9000:1994 con la ISO 14000 (sistemas de gestión medioambiental) y por ello facilita su integración con otros sistemas de gestión dentro de la empresa. La aplicación de estas normas les permite obtener un certificado de calidad, sometiéndose a una auditoría o inspección de su SGC.

La norma ISO 9000 establece una serie de requisitos del sistema de calidad, que se presentan en la tabla 31.1.

<p>SISTEMA DE GESTIÓN DE LA CALIDAD</p> <ul style="list-style-type: none">• Requisitos generales• Requisitos de la documentación
<p>RESPONSABILIDAD DE LA DIRECCIÓN</p> <ul style="list-style-type: none">• Compromiso• Enfoque al cliente• Política de calidad• Planificación• Responsabilidad, autoridad y comunicación• Revisión por la dirección

<p>GESTIÓN DE LOS RECURSOS</p> <ul style="list-style-type: none"> • Provisión de recursos • Recursos humanos • Infraestructura • Ambiente de trabajo
<p>REALIZACIÓN DEL PRODUCTO</p> <ul style="list-style-type: none"> • Planificación de la realización del producto • Procesos relacionados con el cliente • Diseño y desarrollo • Compras • Producción y prestación del servicio • Control de los dispositivos de seguimiento y de medición
<p>MEDICIÓN, ANÁLISIS Y MEJORA</p> <ul style="list-style-type: none"> • Generalidades • Seguimiento y edición • Control del producto no conforme • Análisis de datos • Mejora

Tabla 31.1. Requisitos del Sistema de Gestión de la Calidad.

Se podrían enumerar varios apoyos que una empresa puede utilizar para acceder al mercado y competir, pero, fundamentalmente, hay tres pilares que resultan estratégicos: *precio*, *calidad* y *plazo*.

La anticipación en el tiempo a las necesidades de los clientes era la apuesta de empresas de sofisticada tecnología en el pasado, pero esto ya no es un hecho diferenciador porque todas las empresas, sea cual sea su sector, lo están aplicando.

En la mayoría de sectores y mercados, se puede afirmar que tener precios competitivos es una condición necesaria pero no suficiente para poder permanecer en él. Por ello, la calidad se convierte en un objetivo estratégico para lograr la fidelidad del cliente y ampliar la cuota de mercado. Esto se logra a través de las mejoras en la organización y en el resultado final del producto o servicio que proporciona la implantación de un sistema de gestión de la calidad.

Partiendo de la base que no hay organización perfecta, el sistema de gestión de la calidad se fundamenta en los criterios de la mejora continua que aparecen en la figura 31.1 (Deming, 1982). De esta forma, y como nunca se consigue la perfección en una

organización (el fallo es algo intrínseco en el ser humano) y dado que cuando se alcanzan unos objetivos, se plantea la consecución de otros más exigentes.



Figura 31.1: La rueda de Deming con las etapas a seguir en un proceso de mejora continua.

Se trata de:

- 1º. Detectar problemas en el origen, evitando la multiplicidad de errores futuros.
- 2º. Permitir, a largo plazo, reducir los costes por ineficacia o por errores cometidos, ya que el suministro de un producto defectuoso, provoca coste de devolución del producto por transporte, esfuerzo comercial doble por suministrar de nuevo el producto con los consiguientes costes de envío, retrasos en la fecha de entrega, retrasos en la fecha de facturación y, por ello, del cobro, sin hablar del perjuicio que supone para la imagen de marca y empresa.

La primera toma de contacto en la implantación con el sistema y la norma que lo sustenta, es la elaboración de un manual de calidad y otro de procedimientos. Esto significa definir cada una de las funciones que realiza la empresa sobre la base de normativas y criterios explicitados de forma general en las Normas ISO. Los manuales deben ser redactados por alguien que conozca bien la empresa y su funcionamiento. Pero aún cuando se puedan encontrar ciertos problemas iniciales, la elaboración de los manuales es la parte más fácil del proceso al ser la más teórica. En la implantación del sistema se han de seguir los siguientes pasos:

1. **Fijación de objetivos** cuantitativos y cualitativos en los diferentes departamentos, encaminados a mejorar la satisfacción del cliente: acotar el número de errores, cumplimiento de los plazos de entrega y que permitan conocer a todos de qué partimos y dónde queremos llegar.
2. **Cambios en la estructura organizativa: definiendo funciones** de los diferentes departamentos y de los puestos en cada uno de ellos. Ello implicará crear unas tareas y eliminar otras, así como reasignar y realizar cambios en las que se llevan a cabo y delimitar los niveles jerárquicos.
3. **Definición de la misión de la dirección.** Los jefes departamentales tienen un papel clave en el éxito de la implantación del sistema. Ellos deben ser los primeros convencidos de las ventajas del mismo, puesto que han de transmitir la motivación necesaria para cumplir los procedimientos y exigir responsabilidades a sus subordinados.
4. **Enfoque de los RRHH.** Se ha de unir el sistema de retribuciones al alcance de los objetivos. Se debe premiar y no sancionar. Impartir la formación adecuada y reciclar al personal.
5. **Cambios en la cultura organizativa.** Potenciar el trabajo en equipo. Facilitar las nuevas opiniones de los empleados es la mejor forma de conseguir su motivación y mejorar el trabajo.
6. **Medición de resultados y análisis de desviaciones.** Realizado a partir de las reclamaciones de los clientes (recogidas por cualquier miembro de la empresa u otro instrumento para la detección de errores).
7. **Aplicación de las medidas correctoras necesarias y nuevos objetivos.** Cada 3 o 6 meses se ha de organizar una reunión con el Gerente de la empresa y todos los Jefes de departamento para analizar lo que se ha logrado y los posibles problemas que tiene la empresa, con el fin de plantear soluciones conjuntas.

La comunicación no sólo es vital en materia de calidad, sino que la falta de sistemas de información adecuados puede reducir la capacidad de crecimiento en la empresa, la obtención de beneficios y el aumento de la productividad. Si no hay una comunicación correcta, se corre el riesgo de caer en una profunda desorganización. Los sistemas de Calidad fomentan y mejoran los sistemas de comunicación en la propia empresa.

Las Normas ISO indican como se tienen que llevar a cabo las diferentes funciones a realizar, pero no qué proceso debe seguir la empresa para llevarlas a la práctica. Esto depende de la creatividad, iniciativa y experiencia que en el día a día se va adquiriendo tanto por los directivos como por los trabajadores (Rodrigo, 2002).

3.1.2 Los costes de la mala calidad

Tal y como se apuntaba anteriormente, existe una cierta sensación a tratar de esquivar la aplicación de controles y parámetros de calidad por parte de algunas empresas y organismos públicos, ante la idea de que una mejor calidad supone un mayor coste. No obstante, Feigenbaum (Harrington, 1990) demuestra, bajo experiencia empresarial, que se cumple el principio básico de que la buena calidad supone un empleo correcto de los recursos y una reducción de los costes y mayor productividad. Todo ello proporciona un aumento de los beneficios para la empresa.

Frente a la idea de los costes que supone el implementar un sistema de calidad, aparecen los costes de la mala calidad (CMC), puesto que lo que realmente cuesta es la mala calidad. El coste de mala calidad es: *el coste en el que se incurre para ayudar al empleado a que haga bien el trabajo todas las veces, el coste de determinar si la producción es aceptable, más cualquier coste en que incurre la empresa y el cliente si el producto no satisface los requisitos explícitos o implícitos* (Harrington, 1990).

En cuanto a la forma que tienen las empresas para reducirlo es mediante la implantación de sistemas de gestión de calidad (SGC). El primer problema aparece a la hora de evaluar los costes reales de dicho sistema frente a los CMC. En realidad el empresario no podrá conocer con exactitud cuál es el coste del SGC hasta que éste no se ponga en marcha: costes de burocracia, formación, mantenimiento, etc., y otros que se evitarán, como los derivados de los reprocesos, desperdicios, etc. Por lo general, los SGC deben hacer más competitivas a las empresas al reducir sus costes y ajustar sus productos a las demandas del mercado para vender más y más barato. No obstante, aunque una mayor calidad produce un aumento en las ventas, no existe una proporcionalidad entre ambos, creciendo el coste de fabricación de forma exponencial con el nivel de calidad (Orero, 1994). Por ello han de fijarse y alcanzarse

niveles de calidad realistas en función de unas determinadas necesidades que tienen los usuarios para no disparar los costes.

Puesto que la calidad consiste en hacer las cosas bien a la primera, el coste de la no calidad (Senlle y Vilar, 1996) es el coste añadido que incluye todo lo que se hace más de una vez, todo lo que se repite o se tiene que controlar. Los costes de la no calidad pueden clasificarse en:

- Coste de los **productos defectuosos** o de fallos. Un producto defectuoso es aquel que no cumple con sus especificaciones. Los productos defectuosos siguen dos vías: su identificación en la empresa (donde podrán ser desechados, reelaborados, utilizados como de 2ª categoría), y su salida al mercado (cuando ya es propiedad del cliente los costes suelen ser más altos debido a reclamaciones en garantías, etc). La tabla 31.2 presenta algunos ejemplos de costes tangibles e intangibles asociados a estas dos posibilidades.
- Coste de las **actividades de inspección** y control. Son aquellos que se producen para impedir la salida de productos defectuosos al mercado. Aquí se incluyen las inspecciones de control y análisis de los suministros internos y externos durante el proceso de fabricación. Cuanto más se pretenda asegurar que los productos que salen al mercado, no son defectuosos, más se incrementan los costes por inspección.
- Coste de **actividades preventivas**. Forman parte de la estructura de aseguramiento de calidad, intentando evitar las causas de aparición de los defectos. Aquí se incluyen los planes de calidad con todas sus actividades de formación, mantenimiento preventivo, etc.

Tabla 31.2 Costes derivados de los productos defectuosos (Otero y Domínguez, 1992).

Productos defectuosos identificados en la empresa		
Destino	Costes tangibles	Costes intangibles
Rechazo	Costes del material, mano de obra, costes indirectos, etc. menos el de la posible venta como residuo	Pérdidas en la producción Problemas de programación Incremento nivel de almacén Eventuales retrasos
Uso como producto de 2ª	Diferencia de coste entre las dos clases de productos.	Moral empresarial, etc.
Reelaborados	Coste de la reelaboración	
Reelaborados + uso	Coste de la reelaboración y de la diferencia de productos	
Productos defectuosos que salen al mercado		
Destino	Costes tangibles	Costes intangibles
Reclamación del cliente	Coste del servicio de atención	Mala fama producto, empresa, pérdida de clientes
El cliente no reclama	Ninguno	Ídem

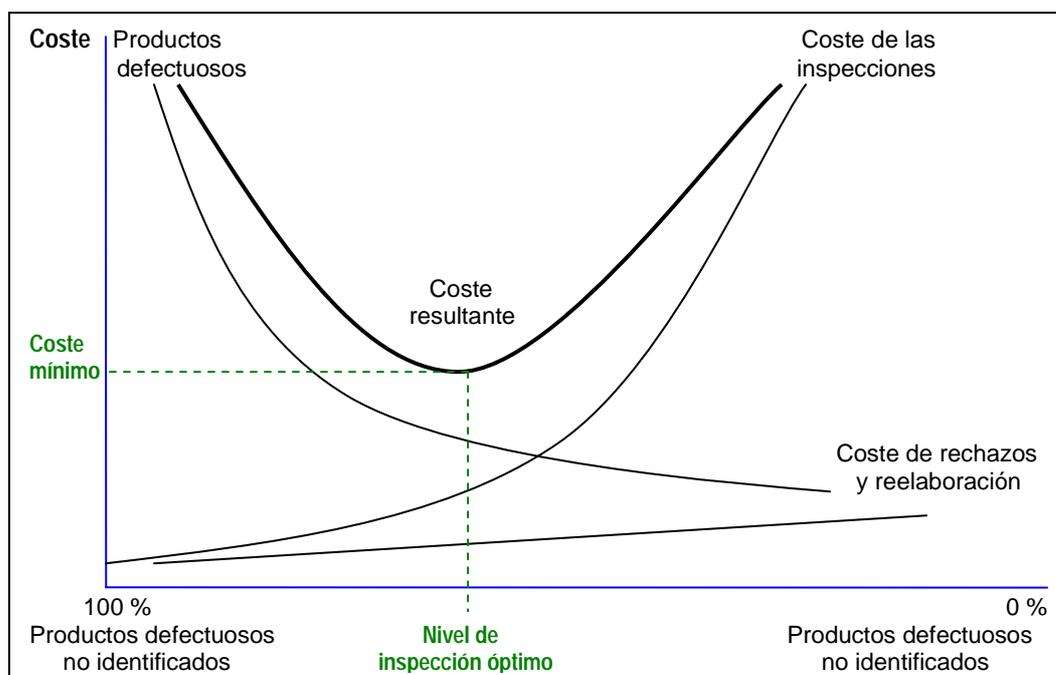


Figura 31.2: Evolución de los costes de la calidad y determinación del nivel de inspección óptimo (Ariza, 2002).

Se puede afirmar que el coste de la calidad es la suma de las tres componentes antes mencionadas. La figura 31.2 presenta una evolución típica de los costes de inspección, de rechazo y reelaboración, y de defectuosos, así como el nivel de inspección óptimo en función de los mismos. En ella se observa cómo el coste de las inspecciones se incrementa a medida que se reduce el porcentaje de defectuosos no identificados. El caso extremo supondría la revisión completa de toda la producción, lo que permitiría identificar todos los elementos defectuosos. La curva representativa de

los costes de rechazo y reelaboración se incrementa conforme se identifica un mayor número de productos defectuosos, mientras que la curva de costes de los elementos defectuosos disminuye. Si no se identifica ningún producto defectuoso, todos saldrán al mercado y generarán un elevado coste. La composición de estas tres funciones permite la obtención de una función en la que se puede determinar el nivel óptimo de inspección con el mínimo de la función (Ariza, 2002).

Un ejemplo del incremento en los costes de inspección se puede apreciar si se analiza el número de puntos a emplear para un control de exactitud posicional (Atkinson, Alcalá y Ariza, 2000), siendo directamente proporcional al número de puntos seleccionado (Figura 31.3).

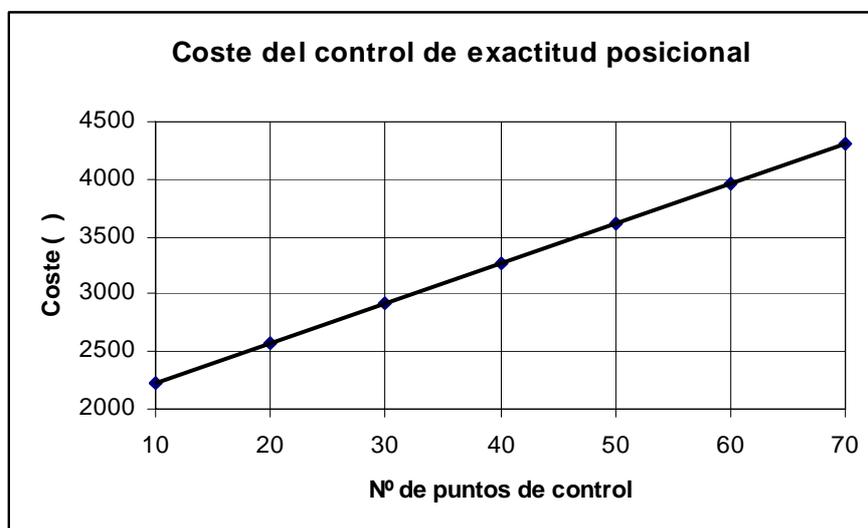


Figura 31.3: Coste total de un control de exactitud posicional en función del número de puntos de control.

Se ha de tratar de encontrar un equilibrio entre el número de puntos muestreados para el control de calidad, en función de varios factores: el nivel de confianza con el que se desee obtener los resultados finales, el coste económico que se esté dispuesto a asumir, el grado de aceptación deseado, la utilización a posteriori de la cartografía (catastral, topográfica, militar, ...), etc.

3.1.3 La calidad en cartografía

Centrándose en los productos cartográficos, resulta de vital importancia el asociar su calidad a la posterior utilización que se vaya a hacer de los mismos. La información geográfica debe adecuarse a los objetivos designados.

En la tabla 31.3, se presentan posibles no calidades que generan costes a una empresa de trabajos cartográficos y que, en algunos casos, pueden ser difíciles de evaluar.

Tabla 31.3. Posibles no calidades en los servicios de una empresa de Ingeniería cartográfica (Ariza, 2002).

No calidades	
<ul style="list-style-type: none"> ▪ Aceptar reclamaciones y dejarlas en saco roto ▪ Dedicar tiempo a arreglar las cosas mal hechas ▪ Demoras en la entrega ▪ Diseño de rutas de trabajo sin optimizar los desplazamientos ▪ El producto servido no es el pedido ▪ Fallos de facturación ▪ Falta de amabilidad ▪ Falta de elementos de demostración o de catálogos apropiados ▪ Falta de garantía o incumplimiento de ella ▪ Falta de interés para comprender al cliente ▪ Falta de profesionalidad ▪ Generar productos sin metadatos ▪ Improvisación 	<ul style="list-style-type: none"> ▪ Lentitud en el servicio ▪ No conocer las nuevas tecnologías ▪ No cumplir lo pactado ▪ No cumplir lo prometido ▪ Pasar de un teléfono a otro sin dar satisfacción a la demanda del cliente ▪ Poco conocimiento del producto ▪ Repetición de errores y fallos ▪ Repetir operaciones ▪ Salidas al campo sin prepararlas ▪ Uso de productos faltos de calidad ▪ Utilizar aparatos no calibrados ▪ Utilizar procedimientos fuera de las normas ▪ Visitas sin avisar al cliente ▪ Recepción de información en formatos no soportados ▪ Recepción de información mal codificada

Puesto que el dato geográfico se caracteriza por su posición espacial (x, y, z), sus atributos ($a_1, a_2, a_3 \dots$), y el tiempo en el que suceden ($t_1, t_2, t_3 \dots$); permite contestar a preguntas de dónde, qué, cómo, cuándo, cuánto, etc., que se pueden plantear para conocer una realidad. Como una BDG es un modelo de la realidad, interesa que éste sea lo más exacto posible; si embargo, en su proceso de generación intervienen numerosas fuentes de imprecisión. A modo de ejemplo, en una cartografía obtenida por teledetección intervienen el esquema de clasificación, los efectos atmosféricos, la calidad radiométrica de la imagen, la escala del producto de salida, el algoritmo de asignación, la determinación de los campos de entrenamiento, etc. Todo este proceso

de integración de la información está afectado de numerosos errores (Ariza, 2002) dependiendo del que se esté realizando en cada momento (tabla 31.4).

Tabla 31.4. Procesos donde pueden generarse errores en el manejo de los datos geográficos (Aronoff, 1989).

Proceso	Motivo
Modelización conceptual	<ul style="list-style-type: none"> ▪ Errores en el modelo conceptual
Recogida de datos	<ul style="list-style-type: none"> ▪ Error en los trabajos de campo. ▪ Error en las fuentes de información utilizadas.
Captura de datos	<ul style="list-style-type: none"> ▪ Inexactitud en la digitalización. ▪ Inexactitud inherente a los elementos geográficos.
Almacenamiento	<ul style="list-style-type: none"> ▪ Insuficiente precisión numérica y/o espacial. ▪ Errores de procesamiento
Manipulación	<ul style="list-style-type: none"> ▪ Intervalos de clase inapropiados. ▪ Errores de superposición. ▪ Propagación de errores.
Salidas cartográficas	<ul style="list-style-type: none"> ▪ Errores en la transformación de coordenadas ▪ Inexactitud en la escala. ▪ Inexactitud del dispositivo de salida. ▪ Deformaciones en el soporte.
Uso de los resultados	<ul style="list-style-type: none"> ▪ Entendimiento incorrecto. ▪ Uso inapropiado.

Todos estos procesos conllevan a que la calidad final de una determinada BDG vendrá definida por varios índices de calidad en función de qué componente se esté analizando: espacial, temporal o temática (Shi, 1994). Empleando una nomenclatura muy utilizada en cartografía, se pueden definir una serie de elipses de error o elipsoides (3D) de incertidumbre del dato geográfico (Veregin, 1994) tal y como se muestra en la figura 31.4.

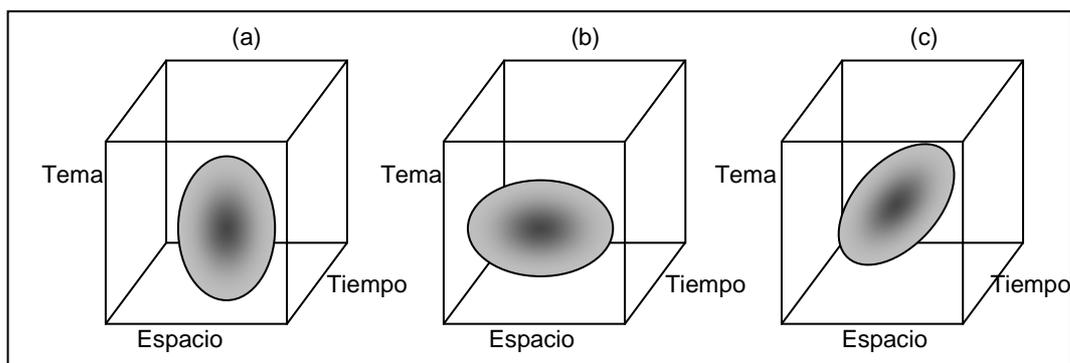


Figura 31.4: Elipsoide de inexactitud alrededor de un punto con una componente dominante: temática (a), espacial (b) y temporal (c).

Muchas veces se confunden los términos que, en un primer momento, pueden resultar similares: calidad, error, exactitud, precisión, fiabilidad ... Así, el término de "calidad"

irá íntimamente ligado a la idoneidad de uso (Ariza, 2002; Matos, 2002), mientras que el término de “error”, complementario al de calidad, hará referencia a la diferencia existente entre el modelo y la realidad. De igual manera, la “exactitud” definirá la aproximación de los resultados de la observación o medida, cálculo o estimación, respecto a los valores reales (verdaderos) o aquellos que se consideren como tales. Por otro lado, el término de “precisión” suele hacer referencia bien al número de cifras significativas que soporta un tipo de dato desde el punto de vista informático (enteros, reales cortos y largos, ...), en función de la calidad del método empleado (p.e.: nivelación de alta precisión), o bien desde un punto de vista estadístico al agrupamiento de los datos obtenidos por repetición de una misma medida (Borrough, 1996; Morant y Lerma, 1999).

Así, puede darse el caso de que se disponga de un sistema en el que la precisión sea muy elevada y, sin embargo, de exactitud muy baja (sistema poco fiable). Esto puede suceder en el caso de que el sistema tenga un sesgo que no haya sido detectado, por lo que su utilización induce a error¹.

Con estas definiciones, cabe destacar que gran parte de los datos geográficos basan sus cualidades en su resolución, con lo que se podría hablar de tantas resoluciones como dimensiones o componentes tiene el dato (Veregin, 1999):

- **Resolución espacial:** Se define como el tamaño del menor elemento que puede ser distinguido o separado de sus partes constituyentes o del conjunto al que pertenecen. En los sistemas vectoriales se refiere a la mínima unidad cartografiada, mientras que en los sistemas ráster se refiere a la dimensión de la celdilla.
- **Resolución temática:** Debe distinguirse entre datos cuantitativos y cualitativos. En el caso de los primeros es similar a la resolución espacial en tanto y cuanto se trata de una variable cuantitativa; mientras que en el caso de los datos cualitativos se refiere a la adecuación/precisión en la definición de las clases.

¹ Un claro ejemplo de ello se da en mediciones GPS cuando están afectadas por un error multicamino (Atkinson, 2003).

- **Resolución temporal:** Se refiere a la duración temporal (intervalo) del muestreo que siempre deberá ser inferior a la duración del fenómeno a representar.

Gran parte de la comunidad científica especializada (NCDCCDS, 1989; USGS, 1994; Morrison, 1995), considera que existen seis componentes básicas de la calidad en los datos geográficos:

- **Exactitud posicional.** Hace referencia a la proximidad entre las coordenadas dadas y las reales.
- **Exactitud del atributo.** Similar a la anterior pero referida al valor del atributo del elemento geográfico. Existe mayor diversidad debido a que los atributos pueden ser categóricos o numéricos, y a que el factor tiempo afecta con una mayor dinámica.
- **Exactitud temporal.** Se refiere a la adecuación de la fecha de la BDG a la del análisis concreto, en particular la acepción más usual es la actualidad de los datos. El aspecto temporal es fundamental en la representación de numerosos datos espaciales (medio ambiente, evolución de tecnoestructuras, ...).
- **Consistencia lógica.** Se refiere a las reglas lógicas que ha de cumplir la estructura de los datos, sus atributos y sus formatos físicos. Ejemplo de ello en una estructura topológica es que las líneas deben empezar y acabar en nodos, las intersecciones de líneas deben ser nodos, etc.
- **Compleción.** Mide lo completa que es la BDG. La presencia de todos los objetos del mundo real en nuestro modelo dependerá de la selección de temas y de reglas de generalización. La leyenda deberá estar completa y sin elementos extraños, debiendo ser adecuada su captura.
- **Genealogía o linaje.** Fundamentalmente referido a las fuentes, procesos de captura, métodos de análisis, sistemas de referencia, parámetros de transformación de proyección, resolución de los datos, etc.

De esta forma, a pesar de existir numerosas discrepancias en cuanto a la definición de unas características o componentes de la calidad, se están realizando esfuerzos en su definición (tabla 31.5).

Tabla 31.5. Calidad de los datos geográficos en distintas normas (Morant, 1999).

Comisión ICA	ISO/TC 211	CEN/TC 287	AENOR CTN 148
Elementos de la calidad espacial	Elementos de la calidad de los datos	Elementos de la calidad de los datos	Elementos de la calidad de los datos
Exactitud posicional Exactitud temática o de los atributos Compleción Consistencia lógica Exactitud semántica Exactitud temporal	Compleción Consistencia lógica Exactitud posicional Exactitud temporal Exactitud temática	Genealogía Utilización Parámetros de calidad	Exactitud posicional (absoluta y relativa) Exactitud semántica (textual y atributos) Coherencia Genealogía
	Otros aspectos de la calidad: Propósito Uso Genealogía	Parámetros de calidad Compleción Consistencia lógica Exactitud posicional Exactitud atributos Exactitud semántica Parámetros secundarios Fidelidad textual	Otros aspectos del intercambio: Documentación y formatos. Calidad de la organización emisora. Calidad de los propios datos

A la hora de analizar la exactitud de una determinada BDG, básicamente se podrá realizar de cinco formas diferentes (Veregin y Giordano, 1994):

- **Estimaciones de experto.** No permiten cuantificar, las conclusiones serán del tipo buena, regular o mala. Son meramente orientativas y subjetivas, dependiendo mucho del conocimiento de quien las hace.
- **Estudio de la coherencia interna.** Este tipo de metodologías tienen como objetivo la búsqueda de posibles incoherencias en el trabajo. Por ejemplo, algunos test se dirigen a establecer si existen incongruencias en los valores del conjunto de atributos temáticos, en las relaciones posibles entre objetos, etc.
- **Comparación con las fuentes.** La información se compara con los datos de los cuales ha sido obtenida. Por ejemplo, la información obtenida de la digitalización de un mapa existente se superpone al mapa original para identificar discrepancias (elementos que faltan o elementos que sobran, digitalizaciones dobles, errores en la forma o dimensiones de los elementos, en los atributos, etc.).
- **Comparación con fuentes de mayor exactitud.** Se trata de metodologías basadas en el control de los datos con fuentes más exactas o con trabajos de campo, siendo este último caso la opción más objetiva al basarse directamente en datos terreno. Consiste en realizar un contraste estadístico entre la BDG y la verdad terreno. Por razones de coste, para este tipo de análisis se emplean técnicas de muestreo tales que se tenga una baja probabilidad de aceptar una

BDG mala y una alta probabilidad de aceptar una BDG exacta, intentando minimizar el número de visitas al campo. Cuando se compara con otra fuente cartográfica, es necesario que dicha referencia sea de exactitud conocida y significativamente mejor que la BDG objeto de evaluación. Asimismo, ambas informaciones deberán haber sido compiladas con procedimientos similares, de tal forma que tengan el mismo esquema de clasificación (p.e. mapas temáticos) y las mismas unidades mínimas cartografiadas. Como puede entenderse, estos requerimientos son difíciles de cumplir.

- **Evaluación de tipo deductivo.** Se trata de técnicas aplicadas a funciones de análisis que consisten en modelizar la propagación del error para determinar valores de exactitud en los productos finales. Éstos se obtienen mediante funciones de transformación en la producción cartográfica (p.e. SIG). Las estimaciones suelen basarse en el nivel de exactitud de la fuente original.

3.1.4 La calidad en las Agencias Nacionales de Cartografía

Las Agencias Nacionales de Cartografía² de Europa están realizando un gran esfuerzo en el estudio y la aplicación de normas relacionadas con la calidad. Así, en CERCO³ se agruparon 34 de ellas⁴ con la principal misión de ayudar a sus miembros en toda Europa en sus necesidades de información georreferenciada y cartográfica. En 1993 se creó la organización MEGRIN⁵ para dirigir proyectos europeos. Esto ha incluido la armonización de las BDGs y el desarrollo de servicios de metainformación basados en Web, siendo más de la mitad de los miembros de CERCO también miembros de MEGRIN. En un esfuerzo por evitar posibles confusiones y activar las economías de ambas administraciones, se creó **EuroGeographics** como la unión de ambos en septiembre de 2000 durante la *Asamblea General de Miembros* en Malmö (Suecia). De esta forma, quedan representadas en una única entidad todas las NMAs que trabajan en pro de la Infraestructura de Información Geográfica Europea (EuroGeographics, 2002).

Actualmente, EuroGeographics posee tres grupos de trabajo:

² National Mapping Agencies, en adelante **NMA**.

³ Comité Européen des Responsables de la Cartographie Officielle, fundado hace más de 20 años.

⁴ Actualmente lo forman 44 agencias de 40 países.

⁵ Multi-purpose European Ground Related Information Network.

- Grupo de expertos en asuntos legales y comerciales
- Grupo de expertos en geodesia
- Grupo de expertos en gestión de la calidad para las especificaciones de los datos y estándares.

Este último ya fue creado bajo los auspicios de CERCO (Septiembre de 1997), siendo sus principales objetivos:

- Plantear un foro de discusión sobre la calidad de los datos y la gestión de la calidad en las NMAs.
- Promover experiencias en el campo de la calidad en la Información Geográfica (IG)
- Crear una red de expertos en calidad de los datos y la gestión de la calidad en las NMAs.

De esta forma, se han constituido varios subgrupos con tareas más específicas:

- SWG-A: **Gestión de la Calidad e ISO 9000**, encargado de analizar la oportunidad de implementar estándares ISO 9000 en las NMAs, desarrollar los argumentos que justifiquen su utilización y analizar la aplicación de los capítulos del estándar ISO 9000
- SWG-B: **Calidad de los Datos**, analizando la calidad de los datos en todo su ciclo vital: especificación de métodos, metodología para la medida de datos de calidad y control de calidad, así como las formas de asegurar los requerimientos demandados por los usuarios.
- SWG-C: **Estándares**, trabajando bajo la experiencia de la implementación de la norma CEN/TC 287 (especialmente “metadata”), que con la unión a MEGRIN fue cerrado en 1999.

En julio de 1999 CERCO realizó una encuesta sobre calidad en cartografía, recibiendo respuesta de 21 agencias de 18 países (incluido EE.UU.). El resultado sobre dicho trabajo fue presentado en el seminario del SWG-B celebrado en Estocolmo en año 2000 (Jakobsson y Vauglin, 2001)

La encuesta fue bastante extensa, cubriendo las 226 series de diferentes datos que elaboraban todas las NMAs. Como resultado, destacan cinco aspectos importantes:

1. Identificar las necesidades del usuario.
2. Especificar los productos que reúnen los requerimientos del punto 1.
3. Producir datos acordes con el punto 2.
4. Verificar que el punto 3 ha sido realizado correctamente.
5. Distribuir los datos del punto 4 a los usuarios para asegurar que los requerimientos del punto 1 han sido correctamente identificados.

Sobre los principales resultados obtenidos en dicha encuesta, cabría destacar los siguientes aspectos:

- El empleo de estándares no es muy común y ha de ser reactivado.
- Necesidad de mejorar los procesos de gestión, especialmente los procesos de evaluación de la calidad.
- Necesidad de una mejora en la gestión de los datos, fundamentalmente en sus especificaciones.
- Necesidad de una mejora en la metainformación (p.e. los resultados de la calidad).
- Formación del personal.

Existe una gran necesidad de mejorar esta situación y las NMAs europeas entienden la importancia de la calidad, pero realmente no han invertido en su evaluación. Las razones son tanto que el mercado no lo demanda, como la falta de financiación gubernamental. A pesar de ello, se reconoce el hecho de que el mercado de la información geográfica esta evolucionando y ésta será más empleada cada día, modificándose el papel del actual consumidor. Probablemente, dichos usuarios no serán profesionales expertos en el conocimiento de cómo se ha obtenido la información y qué garantías se esperan, por lo que las NMAs han de tomarse muy seriamente la calidad de sus productos. La gestión de la calidad junto con la calidad de la información, constituyen un gran avance para los productores, significando un ahorro de costes y aumento en la producción (Jakobsson y Vauglin, 2001).

No obstante y dado que, por una parte, la encuesta se realizó entre 1999 y 2000, y, que por otra, no se analizaban de forma exhaustiva las herramientas y estándares empleados en el control de calidad, se ha creído conveniente el realizar una nueva encuesta con mayor nivel de detalle en estos aspectos. Así, paralelamente a los trabajos de EuroGeographics, se ha diseñado y elaborado una encuesta. En ella, se pretende indagar acerca de los sistemas y controles de calidad que están o se van a implementar a corto o medio plazo en diferentes empresas y organismos que elaboran productos cartográficos. Junto a la encuesta, se redactó una carta de presentación y elaboró una base de datos con los resultados de una intensa búsqueda.

Partiendo de los índices de respuesta (Anexo 1), se muestran en primer lugar los resultados obtenidos en la encuesta a nivel internacional.

De las 16 agencias/empresas que contestaron, el 62.5% no tienen implementado ningún sistema de calidad, mientras que el resto sí. En aquellas que lo tienen, tres de ellas emplean el estándar ISO 9000:2000 (una de ellas con EFQM), y las otras dos restantes el ISO 9001 y 9002 respectivamente. En aquellos organismos que no tienen ningún estándar como sistema de calidad, dos de ellas emplean un sistema propio, mientras que tres más esperan implementarlo a corto o medio plazo. Resulta llamativo el hecho de que en las encuestas se reconozca que no tienen ningún sistema de calidad implementado y al final del documento “afirmen” que sus productos tienen unos controles de calidad “muy estrictos” y que, por ello, “son de muy buena calidad”.

En el apartado, herramientas de control y mejora, un elevado número de organismos emplean *Hojas de chequeo*, frente a otro tipo de herramientas. En la figura 31.5 aparecen las 7 herramientas (numeradas en la parte inferior de la 1ª a la 7ª), representando de forma conjunta cada uno de los 3 bloques del apartado correspondiente:

- Siete herramientas **básicas** (antiguas):
 1. Histogramas
 2. Diagramas causa-efecto
 3. Hojas de chequeo
 4. Diagramas de Pareto
 5. Gráficos

6. Control por tablas
 7. Diagramas de dispersión
- Siete herramientas (**nuevas**) de control:
 1. Diagramas relacionales
 2. Diagramas de afinidad (método KJ)
 3. Diagramas sistemáticos (Diagramas de Árbol)
 4. Diagramas matriciales
 5. Análisis matricial de datos
 6. Proceso de decisión por tabla de programas (PDPC)
 7. Diagramas de flechas
 - Siete herramientas de **planificación** del producto:
 1. Grupo de entrevistas
 2. Cuestionarios de inspección
 3. Análisis
 4. Comprobación por lista
 5. Tablas-tipo
 6. Análisis de conjunto
 7. Tablas de calidad

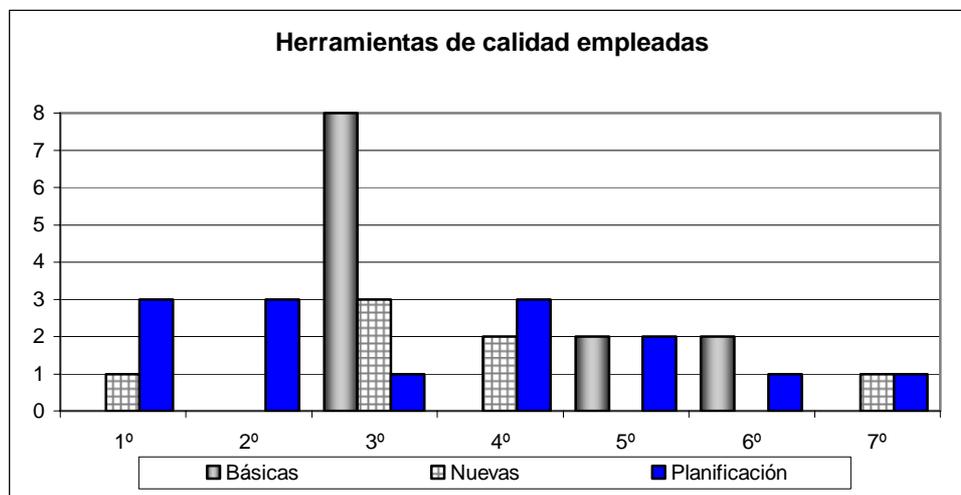


Figura 31.5: Número de empresas que emplean alguna de las herramientas para el control y mejora de la calidad.

En siguiente lugar, con gran diferencia, se encuentran los *Diagramas de árbol*, *Grupo de entrevistas*, *Cuestionarios de inspección*, y *Comprobación por lista* (estas tres últimas son consideradas como herramientas de planificación del producto).

Respecto a qué parte de la producción se hace “en casa” y qué parte se obtiene de otra empresa externa, se ha considerado que aquellas que subcontratan menos de un 5% (la mitad de ellas) es porque realizan la mayor parte del trabajo con sus propios medios, y tienen un nivel medio de subcontratación del 1%. El resto, una cuarta parte subcontratan parte de la producción (con una media del 26% de ésta), y la otra cuarta parte de las organizaciones subcontrata la mayor parte de su producción (90%).

En el apartado de la producción (figura 31.6), todas realizan mapas topográficos (al igual que sucedía en la encuesta de EuroGeographics), y la gran mayoría producen cartografía temática y ortofotos.

En cuanto al formato, prácticamente todas ofrecen productos analógicos y digitales. No obstante, llama la atención que en los digitales, a excepción de los MDE, no suelen rellenar la casilla de “resolución” y sí la de “escala” (cuando este concepto se ajusta más a la cartografía analógica).

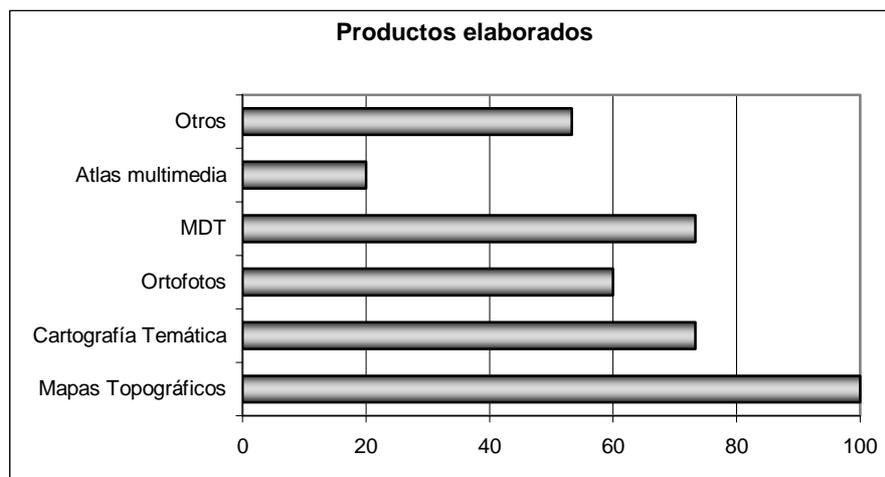


Figura 31.6: Productos cartográficos y porcentaje de empresas que los realizan.

La exactitud posicional es controlada de forma estándar tan sólo por el *Institut géographique national* de Francia., que emplea el NMAS (ver apartado 3.2.1). El *Instituto Geográfico Agustín Codazzi* (Colombia) emplea un estándar propio similar al NMAS. En cuanto al resto, dos emplean un estándar propio basado en el RMSE, y tres más utilizan otro sin especificar. Cabe destacar que más de la mitad no realizan **ningún tipo de control sobre la exactitud posicional.**

En la exactitud temática, el control es todavía inferior: tres de las empresas encuestadas afirman trabajar bajo un estándar propio (sin especificar), y tan sólo el

Agustín Codazzi emplea técnicas de matrices de confusión, aunque sin especificar los criterios estadísticos de aceptación.

Por último, en la coherencia de los datos, el 37.5% afirman realizar un control sobre la coherencia y la consistencia lógica de las BB.DD., aunque uno de ellos lo hace de forma visual.

Los resultados de la encuesta nacional se pueden resumir de forma mucho más breve en la tabla 31.6.

Tabla 31.6: Resultados de la encuesta nacional

Concepto	Número	Porcentaje
Muestra total	72	100
Respuestas recibidas	11	15.2
Reenvían a otra persona u organismo	6	8.3
En estos momentos no hacen cartografía	2	2.7
No realizan ningún control de calidad	2	2.7
Sí hacen algún tipo de control de calidad	1	1.4

A modo de anécdota, destaca la respuesta afirmativa de un organismo público (cuyo texto original se reproduce en el Anexo 1), en el que se indica que el control de calidad se realiza de forma externa por otra organización (también pública) mediante un chequeo estadístico sobre algunos puntos bien definidos sobre la cartografía, mediante observaciones desde vértices geodésicos. Asimismo, indica que “*se realizan diversos controles por parte del Director de los Trabajos, que son de carácter interno y que lógicamente no voy a detallarle*”. Resulta sorprendente este tipo de respuestas, dado que tanto el cliente como el productor debieran conocer el tipo de controles que se van a realizar sobre el producto.

Otra respuesta digna de mención, también se adjunta en el Anexo 1, es aquella en la que el Ingeniero Técnico en Topografía responsable muestra su inquietud al no recibir el apoyo necesario por parte de sus superiores en su intento de controlar aspectos mínimamente básicos en la recepción de cartografía (consistencia lógica, exactitud posicional y temática, ...). Esta última opinión parece ser la más extendida dado que, por la experiencia personal del doctorando en numerosas entrevistas y conversaciones mantenidas con los responsables de la cartografía de varios ayuntamientos y consejerías, siempre parece existir una gran descoordinación entre las empresas adjudicatarias de la cartografía y los clientes que la van a utilizar.

Como resumen de la encuesta, se podrían señalar las siguientes ideas básicas:

- No se acostumbra a emplear estándares en el control de calidad cartográfico.
- La evaluación de la calidad es prácticamente inexistente.
- Existe una gran necesidad de coordinación tanto en cuanto a adoptar sistemas de calidad entre proveedores y clientes, como entre contratistas y subcontratas.
- Falta de formación en el personal implicado.

Como se puede apreciar, y a pesar del bajo índice de respuestas recibidas, existen numerosas coincidencias entre la encuesta realizada por EuroGeographics y la realizada por el doctorando. Con todo ello, se refuerza la idea de **estandarizar los procesos** y la **evaluación de la calidad** de los productos cartográficos y sus derivados (cualquier tipo de información o BB.DD. georreferenciada). Junto a ello, cabe destacar la **necesidad de formar** en la materia al personal técnico y directivo que interviene en la producción.

3.2 Estándares en el control de calidad posicional

En el proceso cartográfico de generación de un mapa los errores pueden provenir tanto de los datos como de los procesos propiamente dichos, pero, en cualquier caso, el error se transmite y acrecienta en la mayoría de las veces a lo largo de la cadena productiva (Lunetta y col., 1991). Existen dos formas básicas de estudiar el error o exactitud final de una cartografía:

- **Física.** Consiste en estudiar todos y cada unos de los procesos de la cadena productiva, intentando comprender y acotar el error que se produce. Previamente a la realización de este tipo de estudios hace falta contar con trabajos de investigación básica en cartografía que definan adecuadamente los rangos de error de cada operación cartográfica, como los de Maling (1989).
- **Estadística.** Consiste en observar el mapa definitivo y contrastar su calidad mediante un test estadístico sobre una muestra de datos que se consideren representativos, con casi total despreocupación de lo que ocurre dentro de los procesos. Este tipo de análisis son los que se conocen como “caja negra”

(figura 32.1). Cabe destacar que los métodos estadísticos son los únicos que permiten establecer algún tipo de nivel de confianza que sirva tanto al usuario como al productor de forma concreta.

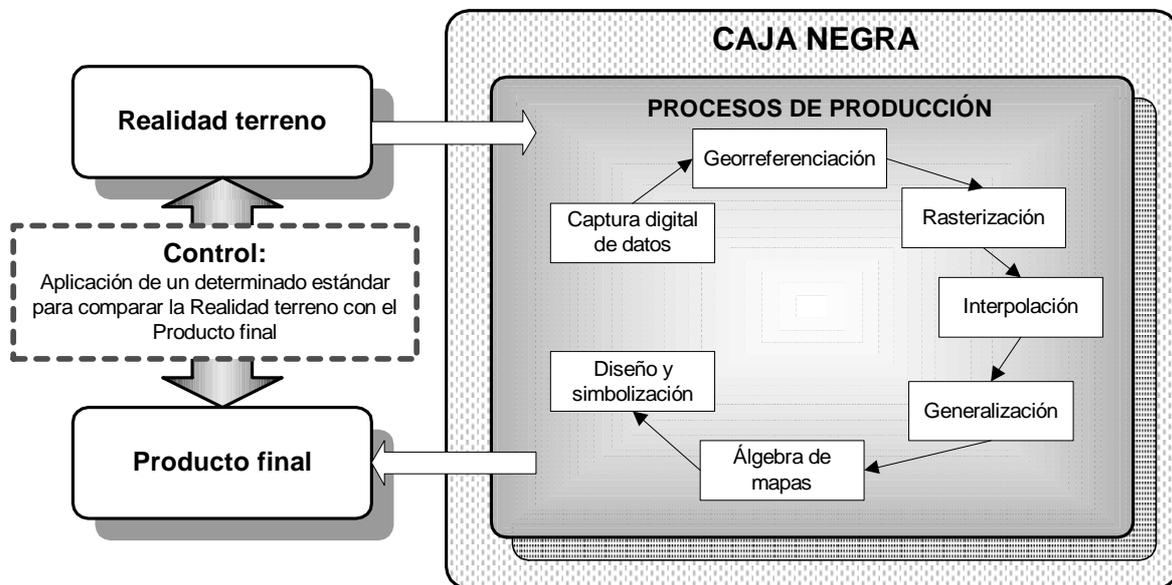


Figura 32.1: Ejemplo de control de calidad del tipo “caja negra” de un producto obtenido mediante un determinado proceso de producción.

Desde el punto de vista de la calidad posicional de los mapas topográficos, existen diversos test o métodos que permiten hacer el seguimiento de la exactitud posicional X, Y y/o Z del mapa. Si bien han de tenerse en cuenta las 3 componentes para un correcto control de la exactitud posicional, hay que reconocer el hecho diferencial de la altimetría tal y como se ha mencionado en el capítulo 1 ⁽⁶⁾.

Respecto al control posicional, al fundamentarse en procedimientos estadísticos, se pueden fijar ciertos niveles de confianza con los que se desee obtener el resultado. No obstante, numerosos autores afirman el hecho de que hay que partir de la hipótesis de normalidad en la distribución de los errores (ASPRS, 1989; Veregin, 1994; FGDC, 1998...). Este aspecto es perfectamente asumible siempre y cuando el número de datos sea lo suficientemente amplio. A pesar de ello, y en el caso de tener alguna duda al respecto, deberá acreditarse la normalidad de la muestra (Ariza, 2002).

⁶ Independientemente de la metodología empleada, generalmente la incertidumbre en Z es superior a la incertidumbre en XY : fotogrametría, teledetección...

Todos los test empleados se basan en la comparación del mapa con una fuente de mayor exactitud (FME). Ésta podrá ser bien un mapa de mayor exactitud (cartografía a escala significativamente superior y/o con una mayor calidad posicional contrastada previamente), o bien datos tomados sobre el terreno con una exactitud superior a la de la cartografía a analizar (recomendándose una fuente tres veces más exacta que el producto a contrastar). En este último aspecto, el empleo de los actuales sistemas de posicionamiento global (GPS) reducen considerablemente los costes en esta fase del proceso.

En cuanto a los errores que afectan al dato geográfico, pueden clasificarse en tres tipos:

- **Errores groseros.** Son aquellos valores considerados como valores atípicos.
- **Errores aleatorios.** Son producto de variaciones al azar producidas en las propias condiciones del trabajo, dado que al medir varias veces el mismo parámetro bajo idénticas condiciones, se obtienen valores diferentes. Si la dispersión en los valores medidos es aleatoria, los errores pueden ser tratados mediante técnicas estadísticas y obtener, a partir de un conjunto de mediciones, un valor representativo del total (Volkman, 2002). Este tipo de errores, también conocidos como accidentales, generan fallos aleatorios individualmente pequeños, y no suelen presentar una ley fija de comportamiento, por lo que generalmente se compensan parcialmente incrementando el número de datos observados (Ariza, 2002).
- **Errores sistemáticos.** Están asociados a las condiciones en que se realiza el experimento. No tienen variación estadística y el tratamiento para su posterior corrección requiere una cuidadosa revisión del montaje experimental. Las fuentes habituales de este tipo de errores se encuentran el uso de instrumentos de medición incorrectamente calibrados, o la suposición equivocada de las condiciones, como presión atmosférica, temperatura, etc. (Volkman, 2002). No obstante, centrándose en los productos cartográficos, los errores sistemáticos (constantes o variables) suceden con el mismo signo y, generalmente, con la misma magnitud en un número consecutivo de observaciones, son generados por causas permanentes de tal manera que pueden eliminarse si éstas se detectan y cuantifican (Ariza, 2002). A estas definiciones, cabría añadir el hecho de que se pueden llegar a producir sistematismos de forma local. A modo de

ejemplo, podría existir un sistematismo que produzca un desplazamiento hacia el noreste en una zona montañosa de la BDG, y un desplazamiento de diferente magnitud hacia el sur en otra zona de la BDG por errores en el apoyo fotogramétrico.

En la exposición y análisis de los diferentes test, se ha seguido a Veregin (1989, 1994) presentando los test NMAS, EMAS, Fórmula de Koppe y Test del USGS para categorías de MDT propios de la ingeniería cartográfica. Así mismo, se ha ampliado el análisis con los estándares ASPRS (ASPRS, 1989) y NSSDA del Federal Geographic Data Committee (FGDC, 1998) de reciente implantación. El esquema de trabajo es el siguiente:

1. **Aplicaciones corrientes:** entidades cartográficas que utilizan el test rutinariamente.
2. **Método de comparación:** indica la metodología utilizada para evaluar la exactitud posicional.
3. **Componente posicional:** especifica si el test se aplica a la componente vertical u horizontal.
4. **Clases de elementos:** puntos, líneas o superficies.
5. **Correspondencia con un estándar de exactitud predefinida:** indica si el estándar está construido en forma de poder evaluar la exactitud posicional de un grupo de elementos relativos a límites de error aceptables y que hayan sido predefinidos por el usuario.
6. **Descripción:** breve presentación del procedimiento.
7. **Procedimiento:** índice, paso a paso, de cómo desarrollar el test.
8. **Discusión:** breve análisis crítico de las ventajas e inconvenientes del test analizado.

En los siguientes apartados se exponen los test más empleados en el control de la calidad posicional final de productos cartográficos, independientemente de que se efectúe un control sobre su componente planimétrica y/o altimétrica.

3.2.1 NMAS - National Map Accuracy Standard

1) **Aplicaciones corrientes:** el NMAS (USBB, 1947) es el estándar empleado por el USGS desde 1947, por lo que ha sido utilizado por numerosas instituciones y organismos oficiales dedicados a la cartografía.

2) **Método de comparación:** se realiza con fuentes de mayor exactitud.

3) **Componente posicional:** analiza tanto la componente horizontal como la vertical. No existe ninguna subdivisión entre la X e Y , analizándolas de forma conjunta. Esto no sucede con la componente vertical dado que es estudiada de forma independiente.

4) **Clases de elementos:** Puntos.

5) **Correspondencia con un estándar de exactitud predefinida:** Se especifica si la cartografía cumple o no cumple con el NMAS.

6) **Descripción:** el test estima si el mapa analizado se encuentra dentro de ciertos límites de error preestablecidos tanto para la componente horizontal (XY) como para la vertical (Z). Los errores se obtienen en base a una comparación entre una muestra homogénea de puntos bien definidos, localizados sobre la cartografía a analizar, y esos mismos puntos localizados sobre una fuente de mayor exactitud.

7) **Procedimiento:**

- Seleccionar una muestra sobre el mapa de, al menos, 20 puntos de fácil determinación
- Determinar las coordenadas de los puntos sobre el mapa y sobre una fuente de mayor exactitud
- Determinar si se ha alcanzado el estándar de exactitud horizontal predefinido. En el caso del USGS, el estándar indica que:
 - Para cartografía a escala mayor que 1/20.000: como máximo el 10% de los puntos de la muestra puede tener un error horizontal mayor de 1/30 de pulgada (0.846 mm)
 - Para cartografía a escala menor que 1/20.000: como máximo el 10% de los puntos de la muestra puede tener un error horizontal mayor de 1/50 de pulgada (0.508 mm)

El error viene definido como la diferencia entre la posición de los puntos en el mapa y en la fuente de mayor exactitud

- Determinar si se ha alcanzado el estándar de exactitud vertical predefinido. En el caso del USGS, el estándar indica que, como máximo, el 10% de los puntos de la muestra pueden tener un error vertical mayor de la mitad del intervalo de las

curvas de nivel. El error en la dimensión vertical puede corregirse modificando la posición de los puntos en una cantidad igual al error horizontal aceptable.

8) **Discusión:** la ventaja del estándar NMAS es su simplicidad: los resultados se expresan EN la forma cumple / no cumple, de fácil interpretación por parte del usuario. No obstante, las desventajas del empleo del NMAS son numerosas. La principal es la falta de indicación de una medida de inexactitud presente en el mapa: existiendo sólo indicación de si cumple o no cumple, las diferencias en el nivel de exactitud pueden ser notables y no están resueltas. Por consiguiente, resulta difícil estimar cuánto puede costar obtener un mapa de mayor exactitud. Una limitación posterior del test reside en el método de muestreo utilizado. Al considerar el requisito de fácil determinación de los puntos, tanto sobre el mapa como sobre la fuente de referencia, se introduce en las mediciones un grado de dependencia de las mismas respecto al observador externo. Además, el hecho de poder corregir el error vertical de un punto actuando sobre su componente horizontal, hace alcanzable el estándar de exactitud sólo en superficies de fuerte pendiente. En definitiva, el test no suministra medida de inexactitud de los puntos tomados: el mapa pasa el test cuando no llegan al 10% los puntos que exceden de un cierto límite, sin posibilidad de conocer la superación de ése límite por los puntos.

3.2.2 EMAS - Engineering Map Accuracy Standard

1) **Aplicaciones corrientes:** El estándar original EMAS fue ideado por la Sociedad Americana de Ingenieros Civiles (ASCI, 1983). El EMAS se propone como alternativa al NMAS para grandes escalas, dando una gran flexibilidad para su uso en cualquier tipo de aplicación. La Sociedad Americana de Fotogrametría (ASP, 1985) hizo suya la idea del EMAS con ligeros cambios (clases de mapas y niveles de exactitud preestablecidos) pero tras el debate en el foro de la ASPRS (la ASP pasó a ser ASPRS) (Merchant, 1987; ASPRS, 1989) se introdujeron cambios mayores. Se mantuvieron las tres clases de mapas y las tolerancias pero se decidió simplificar el método, proponiendo el uso del error medio cuadrático en lugar del error estándar y se eliminaron las hipótesis sobre las distribuciones del error. Todo ello llevó a la posterior aparición del estándar ASPRS.

2) **Método de comparación:** emplea fuentes de mayor exactitud.

3) **Componente posicional:** analiza tanto la componente horizontal como la vertical. Permite realizar un estudio por separado de la componente X e Y , así como de la componente horizontal (XY) frente a la vertical (Z)

4) **Clases de elementos:** puntos

5) **Correspondencia con un estándar de exactitud predefinida:** el estándar nos permite definir los umbrales aceptables de error y el nivel de confianza. Está basado en hipótesis sobre los límites aceptables para los errores aleatorios y sistemáticos.

6) **Descripción:** se obtienen estadísticos para analizar si existen desplazamientos constantes (errores sistemáticos) y la variabilidad de la muestra (errores casuales o aleatorios). Se emplean de forma independiente las componentes X , Y y Z de una muestra de, al menos, 20 puntos bien definidos. La posición de los puntos sobre el mapa objeto del estudio, se compara con su posición en las fuentes de mayor exactitud.

7) **Procedimiento:**

- Seleccionar una muestra de n puntos, para $n \geq 20$
- Salcular el error para cada punto en su componente X :

$$e_{xi} = x_{mi} - x_{ti}$$

en donde:

- e_{xi} es el error en la componente X del punto i
 - x_{ti} la coordenada del punto i sobre el terreno o sobre la fuente de mayor precisión
 - x_{mi} la coordenada del punto i sobre el mapa
- Calcular el error medio en X (\bar{e}_x) para los n puntos que componen la muestra:

$$\bar{e}_x = \frac{1}{n} \sum_{i=1}^n e_{xi}$$

- Calcular la desviación en la componente X (S_x):

$$S_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (e_{xi} - \bar{e}_x)^2}$$

- Calcular el test de cumplimiento / rechazo del estándar empleando los límites de error aceptables para los errores sistemáticos (desplazamientos constantes sobre el mapa). Para ello se emplea el estadístico t_x :

$$t_x = \frac{\overline{e_x} \sqrt{n}}{S_x}$$

Si se verifica que $|t_x| \leq t_{n-1, \alpha}$, el mapa cumple; siendo:

- $n-1$: los grados de libertad
 - α : el nivel de significación
 - t : la medida teórica de una distribución t de *Student*
- Calcular el test de cumplimiento / rechazo del estándar empleando los límites de error aceptables para los errores casuales (aleatorios). Así, se determina el grado de variabilidad de los errores. Para ello se emplea el estadístico χ_x^2 :

$$\chi_x^2 = \frac{S_x^2(n-1)}{\sigma_x^2}$$

Si se verifica que $|\chi_x^2| \leq \chi_{n-1, \alpha}^2$, el mapa cumple con el estándar de calidad posicional en X , y donde:

- σ_x : límite error aceptable (varianza máxima) en la componente X (⁷)
 - $n-1$: los grados de libertad
 - α : nivel de significación
 - χ^2 : la medida teórica de la distribución *Chi-Cuadrado*
- Calcular test análogos (para errores sistemáticos y casuales) en las componentes Y y Z .
- La cartografía cumplirá con el EMAS si **todas** las componentes han superado los dos test correspondientes.

8) **Discusión:** el test EMAS presenta dos importantes diferencias con respecto al test NMAS. En EMAS las diferencias entre las localizaciones de los puntos sobre el mapa y sobre la fuente de referencia están expresadas en términos de distancias sobre el terreno, mientras que en NMAS la unidad de medida es la distancia sobre el mapa. Además, el límite de error aceptable puede ser definido por el usuario, esto tiene como consecuencia importante que los estándares de exactitud pueden definirse en base a las exigencias de aplicación que tiene el usuario de la cartografía. La desviación es

⁷ Recomendaciones de numerosos autores (Sevilla, 1991), hacen referencia al umbral de 0,2 mm de percepción visual a la escala del mapa, incrementados aproximadamente un 25 %. No obstante, se pueden aplicar ciertos "coeficientes de relajación" (Atkinson, 1998) en función del posterior uso de la cartografía.

similar al error medio cuadrático (RMSE), que es una medida utilizada con frecuencia en las valoraciones sobre la exactitud cartográfica. La desviación y el RMSE son iguales sólo cuando el error medio es igual a cero. Algunas versiones del EMAS utilizan más que el límite de error aceptable el valor de RMSE. En este caso, no es necesario desarrollar los dos pasos anteriores (tx , ty , tz), y el valor de la desviación es reemplazado por el valor de RMSE. Por ejemplo, para la componente X sería:

$$RMSE_X = \sqrt{\frac{1}{n} \sum_{i=1}^n e_{xi}^2} \quad \text{[Ecuación 32.1]}$$

en donde:

- $RMSE_X$: es el error medio cuadrático en la componente X
- e_{xi} es el error en la componente X del punto i
- n : es el número de puntos de la muestra

3.2.3 ASPRS – Spatial Accuracy Specification for Large Scale Topographic Maps

1) **Aplicaciones corrientes**: el test ASPRS especifica la exactitud de los mapas topográficos a gran escala. Esta norma, como se ha mencionado anteriormente, proviene del estándar ARPRS y fue desarrollada por la American Society of Photogrammetry and Remote Sensing (Merchant, 1987; ASPRS, 1989).

2) **Método de comparación**: con fuentes de mayor exactitud.

3) **Componente posicional**: horizontal y vertical. Componentes X e Y separadamente. El error vertical y horizontal de forma separada.

4) **Clases de elementos**: puntos.

5) **Correspondencia con un estándar de exactitud predeterminado**: en el caso de los mapas de mejor calidad (Clase **I**), se establecen límites al RMSE para las componentes X e Y según la escala del mapa (tabla 32.1). En altimetría se considera que el RMSE es de $1/3$ del intervalo entre curvas de nivel, salvo en el caso de puntos acotados que será de $1/6$ de dicho intervalo. Los límites para los mapas de las Clases **II** y **III** se obtienen multiplicando los valores dados para la Clase **I** por dos y por tres respectivamente. Se permite que un producto cumpla los requisitos de una clase en planimetría y los de otra en altimetría.

6) **Descripción**: Se calcula el RMSE para comparar los resultados con valores tabulados. Se utiliza una muestra de, al menos, 20 puntos de fácil determinación. La

localización de los puntos sobre la BDG se compara con la localización de los mismos puntos en las fuentes de mayor exactitud.

7) **Procedimiento:**

- seleccionar una muestra de, al menos, 20 puntos bien definidos sobre la cartografía y sobre la fuente de mayor exactitud.
- detectar y corregir los errores previos de inconsistencias, signos, etc.
- calcular el RMSE para cada componente. Por ejemplo, para la componente X sería empleando la Ecuación 32.1
- Comprobar el cumplimiento del estándar utilizando los límites de error marcados por la tabla 32.1. Si la BDG es de Clase **II** los valores de la tabla 32.1 se multiplican por dos, y si es de Clase **III** por tres. El error en la dimensión vertical puede ser corregido variando en cualquier dirección la posición de los puntos en una cantidad igual a dos veces el RMSE límite.
- Se incluirá una leyenda que haga referencia expresa de que el producto ha sido compilado para cumplir con el estándar o que ha sido comprobado y cumple con el estándar. Por ejemplo⁸:

“Este mapa ha sido compilado para cumplir con el estándar ASPRS para mapas de Clase I.”

O bien:

“Este mapa ha sido comprobado y cumple con el estándar ASPRS para mapas de Clase I.”

Tabla 32.1. Requerimientos de exactitud planimétrica (en coordenadas X o Y) para mapas de Clase **I**.

Escala	Límite RMSE (m)
1:50	0,0125
1:100	0,025
1:200	0,050
1:500	0,125
1:1.000	0,25
1:2.000	0,50
1:4.000	1,00
1:5.000	1,25
1:10.000	2,50
1:20.000	5,00

⁸ Texto original: “This map was compiled to meet the ASPRS standard for class (1., 2., 3.) map accuracy” o “This map was checked and found to conform to the ASPRS standard for class (1., 2., 3.) map accuracy”

3.2.4 USGS de categorías de exactitud de MDT

- 1) **Aplicaciones corrientes:** empleado por el USGS para definir la exactitud de sus modelos digitales del terreno – MDT (Veregin y Giordano, 1994).
- 2) **Método de comparación:** también emplea fuentes de mayor exactitud.
- 3) **Componente posicional:** vertical (Z).
- 4) **Clases de elementos:** al igual que en los demás test, emplea entidades puntuales.
- 5) **Correspondencia con un estándar de exactitud predefinida:** el test está basado en el empleo de umbrales para el error máximo y el error medio cuadrático (RMSE).
- 6) **Descripción:** se basa en una cuantificación de los niveles de exactitud, que son empleados posteriormente para obtener la categoría de exactitud del MDT.

7) **Procedimiento:**

- Seleccionar una muestra de, al menos, 20 puntos del MDT obteniendo su componente Z en el MDT.
- Obtener la componente Z de esos mismos puntos sobre la FME.
- Calcular el error en cada uno de los puntos antes mencionados
- Calcular el RMSE en la componente vertical según:

$$RMSE_z = \sqrt{\frac{1}{n} \sum_{i=1}^n e_{zi}^2}$$

donde:

- $RMSE_z$: Error medio cuadrático en la componente Z
 - n : el número de puntos de la muestra
 - e_{zi} : el error en la componente Z para el punto i
- Definir la categorías de exactitud para el MDT:
 - Para los MDT del Nivel 1 el valor máximo del RMSE es igual a 15 m y la muestra no debe contener puntos en los cuales el error vertical sea superior a 50 m.
 - Para los MDT de Nivel 2 el valor máximo del RMSE es igual a 7 m y la muestra no debe contener puntos en los cuales el error vertical sea superior al doble del intervalo de las curvas de nivel de la fuente de mayor exactitud.
 - Para los MDT de Nivel 3 el valor máximo del RMSE es igual a 7 m y la muestra no debe contener puntos en los cuales el error vertical sea superior al intervalo de las curvas de nivel de la fuente de mayor exactitud.
- 8) **Discusión:** Las categorías de exactitud del estándar se definen tanto en términos de nivel total de error en el MDT (mediciones del valor del RMSE) como relativas a los

puntos aislados. El test USGS de categorías de MDT es más preciso en cuanto a que en él los errores groseros, a parte de su influencia en el RMSE, contribuyen a determinar la categoría de exactitud de la BDG. Esto es importante, especialmente, cuando es notorio que las metodologías utilizadas para la recogida y creación de la BDG pueden introducir grandes errores.

3.2.5 Fórmula de Koppe

- 1) **Aplicaciones corrientes:** se emplea como alternativa al estándar NMAS.
- 2) **Método de comparación:** también emplea fuentes de mayor exactitud.
- 3) **Componente posicional:** tan sólo analiza la componente vertical (Z).
- 4) **Clases de elementos:** puntos
- 5) **Correspondencia con un estándar de exactitud predefinida:** no está previsto.
- 6) **Descripción:** En este caso, se emplean coeficientes derivados empíricamente para estudiar los efectos de las pendientes del terreno sobre el error medio en la componente vertical, siguiendo las relaciones establecidas por Koppe.
- 7) **Procedimiento:**
 - Seleccionar una muestra de puntos
 - Obtener la altitud Z de los puntos en la cartografía y sobre una fuente de mayor exactitud
 - Calcular el error vertical de cada punto
 - Calcular el valor de la pendiente en cada punto
 - Calcular el análisis de regresión empleando la tangente del valor de la pendiente como variable independiente y el error vertical como variable dependiente
 - Calcular el error vertical medio según:

$$\bar{e}_V = \pm(A + B \cdot \tan \alpha)$$

donde:

- \bar{e}_V : error vertical medio
 - A : el término independiente de la regresión
 - B : el coeficiente de la regresión
 - α : la pendiente del terreno
- Calcular el error horizontal medio (\bar{e}_H) según:

$$\bar{e}_H = \pm(B + A \cdot \cot \alpha)$$

Tradicionalmente la altitud siempre se ha “leído” sobre el mapa, en lugar de medirse (Maling, 1989), por lo que siempre ha supuesto una interpretación más o menos precisa de la realidad que el mapa estaba representando. Junto a este aspecto, también influye el hecho de que las técnicas empleadas para la toma de las alturas en el terreno han sido diferentes a las empleadas para la captura de los elementos planimétricamente. Con la aplicación de las nuevas tecnologías digitales de representación se han solventado la mayor parte de estos problemas, pero como todavía se utiliza asiduamente cartografía en formato analógico, sigue siendo un elemento diferenciador a tener en consideración.

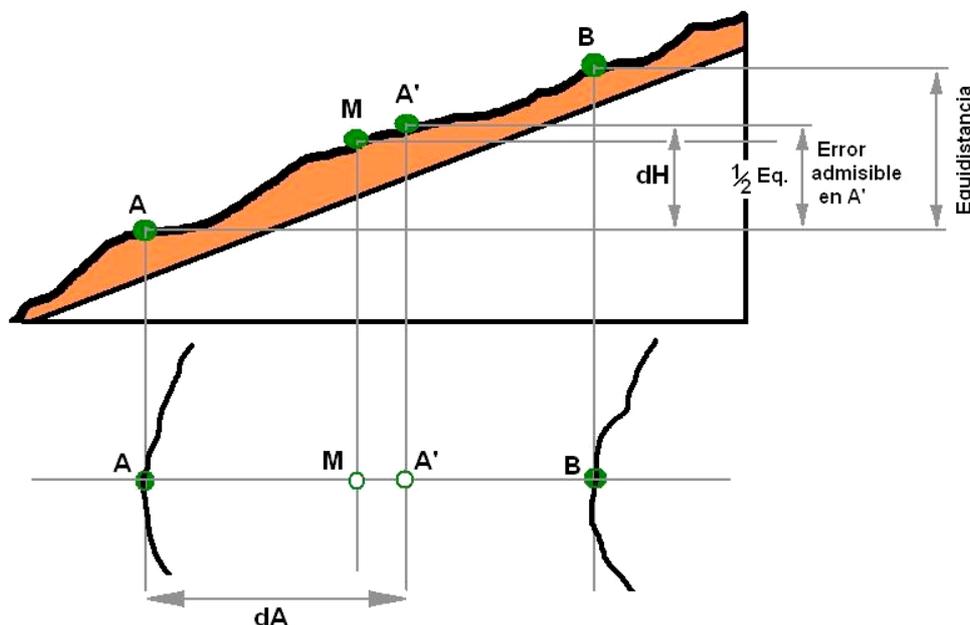


Figura 32.2: Relación entre la componente vertical y horizontal.

Dado que la representación de la tercera dimensión cartográfica se acostumbra a realizar mediante curvas de nivel, existirá una relación $Z = F(x,y)$ (ver la figura 32.2), tal que para dos ubicaciones de una misma curva (A, correcta; y A* incorrecta), en donde se podría considerar un desplazamiento en planimetría dA y un desplazamiento altimétrico dH en función de la pendiente del terreno (α), tendremos que dicha relación será:

$$\tan \alpha = \frac{dH}{dA}$$

Koppe halló una relación empírica entre el RMSE horizontal y vertical, tal que:

$$RMSE_v = \pm(A + B \cdot \tan \alpha)$$

Así, el primer sumando del segundo término de dicha ecuación es una componente independiente de la pendiente del terreno, mientras que el segundo sumando supone la contribución dependiente de dicha pendiente. Los valores A y B se calculan de forma empírica para un mapa o una serie y dependen de los métodos levantamiento, por lo que, como demuestran Gustafson y Loon (1982), existe una tendencia histórica a la disminución de ambos valores. Esto se debe a que los métodos actuales de levantamiento son cada vez más independientes del propio terreno, lo que induce a pensar en una menor aplicación práctica de esta fórmula.

Esta relación permite modificar la posición horizontal de un punto para ajustar su posición vertical y viceversa. La forma de obtener dichas expresiones para una cartografía en concreto consiste en seleccionar una muestra, determinar las elevaciones en la BDG a contrastar y en otra BDG de mayor exactitud o por otro método más exacto (por ejemplo, trabajo de campo). Con los dos valores de elevación para cada punto se determinan los errores existentes en los mismos y las pendientes para cada punto. A continuación se realiza un análisis de regresión entre error y la pendiente, determinando los valores de A y B de la ecuación anterior. Autores como Thompson y Rosenfeld (1971) y Gustafson y Loon (1982) sugieren que estos ajustes se hagan para zonas de pendientes homogéneas, por lo que deben considerarse varios ajustes para cartografías que cubren áreas de gran variabilidad. El error horizontal medio puede definirse transformando la ecuación anterior como sigue:

$$RMSE_H = \pm(B + A \cdot \cot \alpha)$$

En la tabla 32.2 aparecen las relaciones entre los errores $RMSE_{verticales}$ y el desplazamiento planimétrico según la fórmula de Koppe para las escalas más comunes (Maling, 1989).

Tabla 32.2. Valor de las constantes estimadas por Koppe para su ecuación, en función de la escala, el intervalo entre curvas de nivel y la pendiente del terreno.

Escala	Equidistancia de las curvas de nivel	$RMSE_{vertical}$ (m)	$RMSE_{horizontal}$ (m)
1/1.000	1	$\pm(0.1 + 0.3 \tan \alpha)$	$\pm(0.1 \cot \alpha + 0.3)$
1/5.000	5	$\pm(0.4 + 3 \tan \alpha)$	$\pm(0.4 \cot \alpha + 3)$
1/10.000	10	$\pm(1 + 5 \tan \alpha)$	$\pm(\cot \alpha + 5)$
1/25.000	10	$\pm(1 + 7 \tan \alpha)$	$\pm(\cot \alpha + 7)$
1/50.000	20	$\pm(1.5 + 10 \tan \alpha)$	$\pm(1.5 \cot \alpha + 10)$

3.2.6 NSSDA - National Standard for Spatial Data Accuracy

1) **Aplicaciones corrientes:** las agencias federales de los Estados Unidos⁹ 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 (FGDC, 1998)

2) **Método de comparación:** emplea fuentes de mayor exactitud.

3) **Componente posicional:** analiza tanto la componente horizontal (XY de forma conjunta), como la componente vertical (Z).

4) **Clases de elementos:** puntos

5) **Correspondencia con un estándar de exactitud predefinida:** basado en el RMSE de la muestra. El propio usuario, o agencia que solicita la cartografía, es quien decide los umbrales de aceptación, aunque el estándar ofrece unas tablas en las que se presentan unas exactitudes recomendadas. Este estándar viene a sustituir al NMAS y al ASPRS.

6) **Descripción:** con la obtención del RMSE para XY por un lado, y para 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 %). Este test muestra un índice de calidad de la cartografía en unidades reales sobre el terreno.

7) **Procedimiento:**

- Seleccionar una muestra de, al menos, 20 puntos bien definidos sobre la cartografía y sobre la fuente de mayor exactitud.
- Detectar los errores previos de inconsistencia, signos, ...
- Calcular el RMSE_X para la componente X , y de forma análoga para componente Y , según:

$$RMSE_X = \sqrt{\frac{1}{n} \sum_{i=1}^n e_{xi}^2}$$

donde:

- $RMSE_X$: error medio cuadrático en la componente X
- n : el número de puntos de la muestra

⁹ El Federal Geographic Data Committee está formado por numerosas agencias estadounidenses: Department of Agriculture, Commerce, Defense, Energy, Housing and Urban Development, Interior, State, Transportation, Env. Protection Agency, Fed. Emergency Management Agency, Library of Congress, Nat. Aeronautics and Space Adm., Nat. Archives and Records Adm., Tennessee Valley Authority.

- e_{xi} : el error en la componente X para el punto i
- Calcular el $RMSE_r$ para la componente posicional (XY) según:

$$RMSE_r = \sqrt{RMSE_x^2 + RMSE_y^2}$$

- Calcular el coeficiente de exactitud posicional a un 95% de nivel de confianza dependiendo de dos opciones:
- a) En caso de que $RMSE_x = RMSE_y$, se utilizará la ecuación:

$$RMSE_r = \sqrt{2 \cdot RMSE_x^2} = \sqrt{2 \cdot RMSE_y^2}$$

$$RMSE_r = \sqrt{2} \cdot RMSE_x = 1.4142 \cdot RMSE_x = 1.4142 \cdot RMSE_y$$

Se asume que los errores sistemáticos han sido eliminados de la mejor forma posible y que los errores se distribuyen de forma Normal e independiente entre la X e Y . Así, el factor de 2.4477 es el que deberá emplearse para calcular la exactitud horizontal al 95 % de nivel de confianza (Greenwalt y Schultz, 1968). Cuando se dan estas condiciones, el valor de la exactitud de acuerdo al NSSDA se obtiene mediante la ecuación:

$$Exactitud_r = 2.4477 \cdot RMSE_x = 2.4477 \cdot RMSE_y$$

$$Exactitud_r = \frac{2.4477 \cdot RMSE_r}{\sqrt{2}} = \frac{2.4477 \cdot RMSE_r}{1.4142}$$

$$\boxed{Exactitud_r = 1.7308 \cdot RMSE_r}$$

- b) En caso de que $RMSE_x \neq RMSE_y$, se utilizará la ecuación:

Si la relación entre $RMSE_{min}/RMSE_{max}$ ⁽¹⁰⁾ está entre 0,6 y 1,0, el *Error Circular Estándar* al 39.35% de nivel de confianza puede ser aproximado por (Greenwalt y Schultz, 1968):

$$ErrorCircularEstándar \approx 0.5 \cdot (RMSE_x + RMSE_y)$$

Si los errores se distribuyen de forma Normal e independiente entre la X e Y , el valor de la exactitud de acuerdo al NSSDA puede ser aproximado mediante la siguiente fórmula:

$$Exactitud_r \approx 2.4477 \cdot 0.5 \cdot (RMSE_x + RMSE_y)$$

$$\boxed{Exactitud_r = 1.22385 \cdot (RMSE_x + RMSE_y)}$$

- Calcular el $RMSE_z$ para la componente Z de forma análoga

¹⁰ $RMSE_{min}$ es el valor mínimo entre el $RMSE_x$ y el $RMSE_y$. Análogamente, $RMSE_{max}$ es el valor máximo entre el $RMSE_x$ y el $RMSE_y$.

- Calcular el coeficiente de exactitud posicional a un 95 % de confianza para la componente Z:

$$\boxed{Exactitud_z = 1.96 \cdot RMSE_z}$$

- Junto a la leyenda del mapa deberá aparecer el siguiente texto¹¹:

*“Comprobado para ____ metros de exactitud horizontal al 95% de nivel de confianza.”
y / o “Comprobado para ____ metros de exactitud vertical al 95% de nivel de confianza.”*

O bien:

*“Compilado para ____ metros de exactitud horizontal al 95% de nivel de confianza.” y /
o “Compilado para ____ metros de exactitud vertical al 95% de nivel de confianza.”*

8) **Discusión:** este es el estándar empleado actualmente por las agencias federales de los Estados Unidos para analizar BDG. La mayor ventaja que ofrece es que no nos dice si el mapa es aceptado o denegado, tal y como hacían los test anteriores, sino que nos ofrece un índice de calidad de la cartografía en unidades reales sobre el terreno. En este caso, le corresponde al usuario de la cartografía el establecer los límites de aceptación en función de sus necesidades. El dato ofrecido por el NSSDA deberá aparecer junto con la metainformación del mapa. No obstante, el propio estándar ofrece una serie de recomendaciones del error admisible en función del tipo de cartografía y su escala que, a modo de ejemplo, se presentan en la tabla 32.2.

¹¹ Texto original: “Using the National Standard for Spatial Data Accuracy, the data set tested ____ meters (or feet) horizontal (or vertical) accuracy at 95 % confidence level.” y “Using the National Standard for Spatial Data Accuracy, the data set compiled to meet ____ meters (or feet) horizontal (or vertical) accuracy at 95 % confidence level.”

Tabla 32.2. Algunos ejemplos de las recomendaciones de exactitud aconsejables en función del tipo de mapa (FGDC, 1998; FGDC 2002).

Tipo de mapa	Escala	Exactitud horizontal	Exactitud vertical	Eq. curvas de nivel
Construcción en general (planos topográficos de detalle)	1/500	100 mm	50 mm	250 mm
Planos de situación de zonas recreativas (campos de golf, campos de atletismo ..)	1/1.000	500 mm	100 mm	500 mm
Planos de Administración (diseño de interiores/exteriores)	1/250	50 mm	N/P	N/P
Planeamiento general de mapas de localización (antenas AM/FM, SIG)	1/5.000	1.000 mm	1.000 mm	1.000 mm
Planos de movimiento de tierras (carreteras, drenajes, bordillos, ...)	1/500	250 mm	100 mm	500 mm
Mapas de los servicios de emergencias/SIG	1/10.000	25.000 mm	N/P	N/P
Mapas de suelos contaminados / Modelización	1/500	1.000 mm	500 mm	500 mm

3.2.7 Consideraciones sobre los estándares posicionales

Como se puede observar, todos los estándares analizados se basan en entidades puntuales. Es sobre este tipo de elementos sobre el que se realizan diferentes pruebas estadísticas en sus componentes tanto horizontal (X e Y) como vertical (Z). En todo momento se emplea una fuente de mayor exactitud con la que realizar la comparación de las coordenadas de los puntos que forman la muestra. Se considera que dicha fuente ha de tener una exactitud contrastada de, al menos, tres veces mejor que la que se desee analizar.

Según se indica, esta muestra ha de estar formada en todos los casos por *“al menos 20 puntos bien definidos”*. A la hora de realizar su elección, los puntos no deberán entrañar ninguna posible confusión en cuanto a su localización tanto sobre el mapa como sobre la fuente de mayor exactitud. Así mismo, deberán estar homogéneamente distribuidos. Siguiendo las recomendaciones del FGDC (1998), los puntos se deberán distribuir de tal forma que, en cada uno de los cuadrantes del mapa, se sitúen al menos el 20% de los mismos. Así mismo, la separación entre los puntos no será inferior a la décima parte de la diagonal del mapa a controlar (figura 32.2).

La homogeneidad de la muestra no solamente se refiere a su localización espacial sobre el mapa. Algunos autores mencionan el hecho de que también se deberá tener

en cuenta el aspecto temático: por ejemplo, si el 90% de los elementos representados son caminos carreteros, la muestra seleccionada para realizar el control de calidad debería estar acorde con esta proporción. De igual forma, en aquellas zonas donde exista una mayor densidad de elementos cartografiados, se tomarán los puntos de control con una mayor densidad.

Otro de los aspectos más controvertidos es el hecho de tomar “al menos 20 puntos”. Siempre se ha de tomar un número lo suficientemente grande como para que se cumplan las hipótesis de normalidad, lo cual viene determinado por el teorema de los grandes números en estadística. Por ello, las recomendaciones siempre hablan de “al menos 20 puntos” independientemente del área a controlar o del tamaño de la base de datos (MPLMIC, 1999). Por supuesto, siempre se deberá tomar un cierto número de datos superior, dado que habrá que eliminar aquellos que sean considerados como errores groseros. Este número de puntos debería ser suficiente para asegurar, con un nivel de confianza dado, que el usuario no adquiere una cartografía de calidad no aceptable y, además, ser lo más pequeña posible para minimizar los costes del control de calidad.

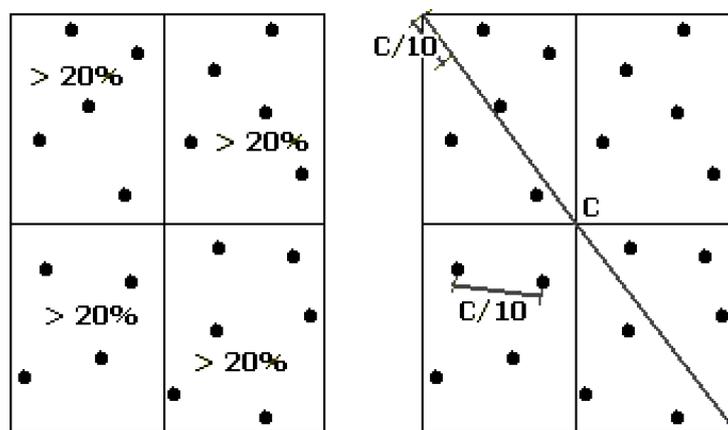


Figura 32.2: Distribución homogénea de la muestra (MPLMIC, 1999).

No cabe duda que el control de calidad posicional en cartografía es uno de los aspectos de mayor relevancia y, a su vez, quizás uno de los más descuidados. Si se observan las normas de calidad que se exigen a la hora de recepcionar los productos cartográficos por parte de una empresa u organismo, en numerosas ocasiones tan sólo se presta atención a determinados aspectos de la calidad: consistencia lógica, compleción... De esta forma, se está pasando por alto una componente básica y fundamental: la posición de los elementos representados.

El hecho de realizar un control de calidad posicional a posteriori, ofrece una garantía final del producto que se está adquiriendo / ofreciendo.

3.3 Técnicas estadísticas

La mayor parte de los controles de calidad se basan en un estudio estadístico diseñando muestreos que permiten determinar si un producto cumple o no con unas determinadas especificaciones. Dicho control se puede realizar desde diferentes perspectivas:

- Desde un punto de vista de la estadística clásica, tal y como hacen la práctica totalidad de los estándares presentados en el apartado 3.2, desechando errores groseros y considerando los valores como componentes unidimensionales e independientes entre sí.
- Desde un punto de vista de la estadística robusta, permitiendo que aquellos valores que puedan ser considerados como outliers o valores atípicos puedan intervenir en los cálculos y el análisis final del producto, sin que dicha influencia sea determinante¹².
- Desde un punto de vista de la estadística circular, no considerando la información geográfica como la componente cuadrática de dos componentes unidimensionales e independientes entre sí (X, Y) , si no como un único valor bidimensional que aparecerá representado como un vector en el plano¹³.

A continuación se exponen algunas posibilidades para el empleo de estadísticos robustos y circulares que, a priori, se podrían aplicar en al estudio de la exactitud posicional.

¹² En este texto, el término de valor atípico, outlier o error grosero, se han utilizado bajo un mismo significado, por lo que se emplean indistintamente (Anexo 9).

¹³ En el caso de trabajar con información tridimensional (X, Y, Z) en vez de herramientas de estadística circular, se emplearían herramientas de estadística esférica (considera la posición de los puntos como un vector tridimensional en el espacio).

3.3.1 Estadística robusta

En el análisis de la exactitud posicional, muchos usuarios desconocen que los métodos empleados no se comportan correctamente si los datos no satisfacen estrictamente las suposiciones establecidas a priori. Hampel (1973) expone diversas justificaciones para la utilización de estimadores robustos en su artículo (he aquí un extracto):

“Pero, ¿cuáles son las razones para usar procedimientos robustos? Hay principalmente dos observaciones que combinadas dan una respuesta. A menudo en estadística se usa un modelo paramétrico, lo que implica un conjunto muy limitado de distribuciones de probabilidad como posibles, como el modelo común de errores distribuidos normalmente. Los estadísticos clásicos (paramétricos) proporcionan resultados bajo la suposición de que estos modelos son estrictamente verdaderos. Pero sin embargo estos modelos no son casi nunca estrictamente verdaderos. Podemos distinguir tres principales razones para las desviaciones: (1) redondeo y agrupamiento y otras inexactitudes locales, (2) la existencia de errores groseros en la observación, puntos decimales equivocados, errores de copia, etc., (3) el modelo ha sido concebido sólo como una aproximación, por ejemplo, en virtud del teorema central del límite.”

En el caso de existir alguna duda respecto al modelo estadístico, la calidad de la muestra o incertidumbre acerca de las decisiones adoptadas, se hace necesario el emplear este tipo de métodos. Así, determinar qué es lo estimado cuando la suposición del modelo no es correcta, puede no estar demasiado claro y cuestionar si este tipo de procedimientos son necesarios. Podría ser suficiente un análisis en dos pasos:

- 1º. Limpieza de los datos aplicando algún tipo de método para rechazar observaciones atípicas.
- 2º. Utilización de la estimación clásica y procedimientos de comprobación en los datos restantes.

Si bien, parece sencillo a priori, el problema resulta algo más complicado ya que:

1. Normalmente no es posible separar claramente los dos pasos. Por ejemplo, en problemas de regresión multiparamétrica, estos errores son difíciles de reconocer a no ser que se utilicen estimadores robustos.
2. Aunque el conjunto original de observaciones esté formado por datos normales entremezclados con valores atípicos, posiblemente los datos ya limpios no seguirán una distribución normal. Por ello, la teoría de la normalidad no es aplicable en muestras limpias, y este procedimiento en dos pasos puede ser más complicado de aplicar que un procedimiento basado en teoría robusta.
3. Es un hecho comprobado que los mejores procedimientos de rechazo no alcanzan las funciones de los mejores procedimientos robustos. Éstos, aparentemente, son mejores porque pueden realizar una transición más suave entre la aceptación total y el rechazo de una observación.

A todas estas consideraciones planteadas por Domingo (2000), cabría destacar la importancia que tiene el utilizar todos y cada uno de los datos al trabajar con muestras relativamente pequeñas (*"... al menos 20 puntos de control..."* FGDC, 1998).

Los estándares mencionados en el apartado 3.3 (NMAS, EMAS, ASPRS ...), trabajan siempre bajo la perspectiva de la estadística clásica. Esto es, no tienen en cuenta aquellos valores ("outliers"), que se encuentran alejados del valor mediano una determinada magnitud. Estos valores se consideran atípicos y procedentes de errores groseros o equivocaciones. Entre ellos, se pueden establecer dos categorías detectables mediante diagramas Box-Whiskers:

- 1 Atípicos o atípicos medios: son aquellos que se distancian más de $1.5 \cdot RQ$ ⁽¹⁴⁾.
- 2 Atípicos extremos: son aquellos que se distancian más de $3.0 \cdot RQ$.

No obstante, otra forma relativamente sencilla y muy empleada para detectar "outliers" es localizando aquellos valores que se distancian más de $3 \cdot \sigma$ veces del valor medio. De esta forma, dichos valores son eliminados al considerarlos como posibles equivocaciones en su medición o localización tanto sobre la BDG, como sobre la fuente de mayor exactitud.

⁽¹⁴⁾ Se entiende por RQ el recorrido o rango intercuartílico: la variación alrededor del centro de una distribución de categorías ordenadas. Este valor mide la variabilidad en el espacio que va entre Q1 (el primer cuartil) y Q3 (el tercer cuartil). Como se aprecia, RQ no es otra cosa que una medida de variación alrededor de la mediana.

Si se observa el ejemplo mostrado en la tabla 33.1, se puede apreciar cómo los puntos nº 1 y 10 podrían considerarse como valores atípicos procedentes de algún tipo de equivocación en su medida. Si se siguen las recomendaciones de la estadística clásica, dichos valores serían suprimidos, quedando la muestra reducida a 8 elementos con una considerable reducción de la desviación típica y una aproximación del valor medio a la zona central de la muestra. Por el contrario, si no se realiza el filtrado previo de los datos, los resultados finales de los estimadores estadísticos difieren considerablemente:

- El valor promedio se ve muy afectado por el atípico extremo (punto nº 10), desplazándose hacia el lado derecho de la muestra.
- La desviación típica triplica su valor.

Si, a la hora de realizar el control de calidad posicional, se toma una reducida muestra de puntos de control (p.e. de 30 ó 40) para minimizar los costes de su captura sobre el terreno, cabe la posibilidad de que una pequeña parte de ellos podría considerarse como valores atípicos. Desechándolos, se perdería una parte importante de la información, y por tanto, no podría ser transmitida al usuario que va a adquirir el producto cartográfico.

Llegados a este punto, ha de plantearse la siguiente cuestión: ¿qué representa realmente un atípico?, ¿no se puede encontrar el usuario del mapa con valores atípicos?, ¿en qué medida y magnitud? Muy probablemente la mayoría de estos aspectos podrán solventarse empleando estimadores robustos.

Tabla 33.1: Ejemplo de aplicación de estimadores clásicos en el cálculo del error medio y la desviación típica muestral.

Ejemplo:										
Sobre una BDG a escala 1/1.000 ($\sigma = 0.25m$) se efectúa un control de calidad posicional tomando una muestra de 10 puntos. La diferencia entre las coordenadas tomadas sobre el mapa y las tomadas sobre el terreno, presentan una serie de valores para la componente X (en metros sobre el terreno):										
Punto nº:	1	2	3	4	5	6	7	8	9	10
Error en X :	-1.11	-0.37	-0.34	-0.14	-0.05	0.09	0.10	0.21	0.32	1.84
Estimadores clásicos:										
Detectando y eliminando valores atípicos:										
Eliminación de "outliers": Mediante <i>Box-Whiskers</i> $Q1 = -0.29$; $Q3 = 0.18$; $RQ = 0.47$ $Q1 - 1.5 \cdot RQ = -1.00m$; $Q1 - 3.0 \cdot RQ = -1.71m$ $Q3 + 1.5 \cdot RQ = 0.90m$; $Q3 + 3.0 \cdot RQ = 1.61m$ <ul style="list-style-type: none"> Por lo que el punto nº 1 sería un atípico medio, y el punto nº 10 sería un atípico extremo. O también: $3 \cdot \sigma = 3 \cdot 0.25 = \pm 0.75m$ <ul style="list-style-type: none"> Por lo que los puntos nºs 1 y 10 serían valores atípicos. 					Cálculo de estimadores: Se han suprimido del cálculo los puntos nºs 1 y 10: <ul style="list-style-type: none"> Error medio de la muestra $\bar{e}_X = -0.02m$ Desviación típica de la muestra $S_X = 0.25m$ 					
Sin detectar ni eliminar valores atípicos:										
Se emplean todos los datos.					Cálculo de estimadores: El cálculo se realiza con todos los puntos: <ul style="list-style-type: none"> Error medio de la muestra $\bar{e}_X = 0.06m$ Desviación típica de la muestra $S_X = 0.75m$ 					

Numerosos autores comentan que se pueden esperar en torno a un 10% de observaciones con errores de medida o grabación (Rencher, 1998). Otra hipótesis común es que los datos son casi normalmente distribuidos¹⁵ pero tienden a aumentar en las colas, de modo que unos cuantos valores extremos pueden alterar sustancialmente la media e incrementar la varianza. Para evitarlo se pueden utilizar estimadores robustos, que son menos sensibles a las observaciones extremas que los estimadores clásicos (p.e.: la media). Un estimador robusto se dice eficiente si proporciona una menor estimación de la varianza que S^2 para distribuciones con

¹⁵ La mayoría de los estándares de control de calidad posicional cartográfico se basan en la hipótesis de que los errores de distribuyen de forma normal, aleatoria e independiente en cada una de sus componentes (X , Y , Z).

grandes colas. Así, Mosteller y Tuckey (1977) comparan la desviación media con S en una distribución normal contaminada, de modo que el 1% de los datos proviene de una población con desviación típica 3 veces mayor. Para este caso, encontraron que la desviación media es un 144% más eficiente que S . Un posible esquema para obtener estimadores robustos puede ser el M de Huber, que obtiene una solución para $\tilde{\mu}$ en:

$$\sum_{i=1}^n \psi\left(\frac{y_i - \tilde{\mu}}{S}\right) = 0$$

para lo que hay propuestas múltiples funciones ψ .

Otra manera de encontrar estimadores robustos multivariantes para μ consistentes en vectores de estimadores robustos para cada variable univariante, donde:

- 1 \mathbf{y}^*_M es el vector de medianas.
- 2 \mathbf{y}^*_{HL} es el vector de *Hodges-Lehmann*, siendo cada uno la mediana de las medias que se obtienen entre cada par de observaciones
- 3 $\mathbf{y}^*_{T(\alpha)}$ es el vector de medias α -recortado, cada una de las cuales es la media resultante tras descartar una proporción α de las menores y mayores observaciones (por ejemplo, $\alpha = 0.1$)
- 4 \mathbf{y}^*_{sen} es el vector de estimadores M para cada variable, en donde cada $\tilde{\mu}_j$ en α es la solución de:

$$\sum_{i=1}^n \psi\left(\frac{y_i - \tilde{\mu}}{S}\right) = 0$$

$$\text{en donde } \psi \text{ es: } \psi(z) = \begin{cases} \text{sen}(z/2.1) & \text{si } |z| < 2.1\pi \\ 0 & \text{en otro caso} \end{cases}$$

No obstante, dado que los estimadores presentados suprimen datos, sería deseable emplear algún estimador similar al recortado de máxima verosimilitud (**MLT**¹⁶), pero basándose en una ponderación de la muestra. Esto es, aquellos datos que se encuentren más alejados de los valores más probables, tendrán un menor peso en el cálculo del valor medio y desviación típica.

A continuación se presentan algunos posibles estimadores robustos que podrían aplicarse en procesos de control de calidad posicional. Así, se expone una breve descripción del mismo junto con el **procedimiento de aplicación**. En todos ellos se considera como un paso intermedio a aplicar en los estándares de control de calidad,

¹⁶ Maximum Likelihood Trimming

después de seleccionar los puntos de control y **previamente a la realización de las pruebas estadísticas** propias de cada estándar (ASPRS, NSSDA, ...).

3.3.1.1 MVT – MLT

Se trata de dos estimadores que podrían considerarse como aceptables: **MVT** (*Multivariate Trimming*) y **MLT** (*Maximum Likelihood Trimming*) (Devlin, Gnanadesikan y Kettenring, 1981). El **MVT** o *recorte multivariante* es un proceso iterativo, basado en la distancia de Mahalanobis (d_i^2), eliminándose aquellos valores que se encuentren más alejados. Por el contrario, el estimador *recortado de máxima verosimilitud* (**MLT**) es un caso especial de un estimador M en el que, en vez de recortes, se utilizan pesos basados en la distancia de Mahalanobis. El no recortar la muestra y emplearla en su totalidad, es lo que le hace ser un estimador potencialmente adecuado en el control de calidad posicional cartográfica.

El **procedimiento** de aplicación sería el siguiente:

- Obtención, mediante estadística clásica, de los valores iniciales para la media ($\tilde{\mu}$) y la desviación (S).
- Comienzo del proceso de iteración:
 - Cálculo de la distancia de Mahalanobis para el punto i :

$$d_i^2 = (y_i - \tilde{\mu})' \cdot \tilde{S}^{-1} \cdot (y_i - \tilde{\mu})$$

en donde:

- d_i^2 : es la distancia de Mahalanobis en el punto i
- y_i : es la matriz de errores en el punto i . Para el análisis de la componente horizontal (X e Y), se trata de una matriz de dos dimensiones (2×2).
- $\tilde{\mu}$: es la matriz con los errores medios en la iteración $(i-1)$.
- \tilde{S} : es la matriz varianza-covarianza de los errores en la iteración $(i-1)$.
- Detección de valores atípicos en función de las distancias de Mahalanobis frente al coeficiente de una distribución χ^2 con p grados de libertad (siendo p el número de componentes analizadas, en este caso $p = 2$):
 - **Atípicos medios**: son aquellos cuya $d_i^2 > \chi_{2,0.05}^2 (= 5.9915)$. El contraste se realiza al 95% de nivel de confianza.

- **Atípicos extremos:** son aquellos cuya $d_i^2 > \chi_{2,0.01}^2$ (= 9.2103). El contraste se realiza al 99% de nivel de confianza¹⁷.
- Cálculo de los estadísticos ponderados a emplear en el control de calidad:

$$\bar{\mu} = \frac{\sum_{i=1}^n \omega_i \cdot y_i}{\sum_{i=1}^n \omega_i} \quad \tilde{S} = \frac{1}{n} \sum_{i=1}^n \omega_i (y_i - \bar{\mu})(y_i - \bar{\mu})$$

en donde:

$$\omega_i = \frac{p+1}{d_i^2 + 1}$$

siendo:

- ω_i : matriz de pesos
- p : número de variables (para un control planimétrico en X e Y , $p = 2$)
- d_i^2 : distancia de Mahalanobis en el punto i

Los estimadores se denominan **S□MVT** y **R□MVT**, para el caso del recorte, ó **S□MLT** y **R□MLT** para la ponderación.

- Determinar si se cumple el criterio de convergencia. Devlin, Gnanadesikan y Kettenring (1981) sugieren empezar con $\bar{\mu}$ y S , y terminar con 25 iteraciones o cuando la transformada de Fisher de los r_{ij} cambie menos de 0.001. Ésta viene dada por:

$$z_{jk} = \frac{1}{2} \ln \left(\frac{1 + \tilde{r}_{jk}}{1 - \tilde{r}_{jk}} \right) = \tanh^{-1} \tilde{r}_{jk}$$

de forma que los r_{jk} se pueden obtener a partir de: $\tilde{R} = D^{-1} \cdot \tilde{S} \cdot D^{-1}$; en donde **D** es la matriz diagonal cuyos elementos son las raíces cuadradas de los elementos de la diagonal de **S□**.

- Si no se cumple el criterio de convergencia, realizar la siguiente iteración (en el cálculo de la distancia de Mahalanobis).

3.3.1.2 Suavizado gaussiano

Partiendo de la idea propuesta por el MLT, se emplea un suavizado gaussiano sobre aquellos valores que se encuentren a una determinada distancia con respecto del valor medio (García, 1998). Así, los errores suavizados de un punto i , se obtienen a

¹⁷ El valor de χ^2 es la medida teórica de la distribución *Chi-Cuadrado*. Para un análisis tridimensional (X, Y, Z), los valores límite son: 7.8147 ($p = 3$; $1-\alpha = 0.05$) para los atípicos medios, y 11.3449 ($p = 3$; $1-\alpha = 0.01$) para los atípicos extremos.

través de una ponderación que sigue una distribución normal, en forma de campana de Gauss (figura 33.1), en la que intervienen tanto el error del punto a suavizar como la distancia a la que se encuentre del valor medio muestral ($e_i - \bar{e}$). En el centro de la campana se sitúa dicho valor, y en los extremos los factores de ponderación van disminuyendo simétricamente para aquellos puntos que se encuentren a mayor distancia (e_i). La función gaussiana viene dada por la expresión:

$$g_{\sigma}(e_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\bar{e}-e_i)^2}{2\sigma^2}}$$

siendo:

- $g_{\sigma}(e_i)$: función gaussiana
- σ : desviación típica
- e_i : error para el punto de orden i
- \bar{e} : error medio

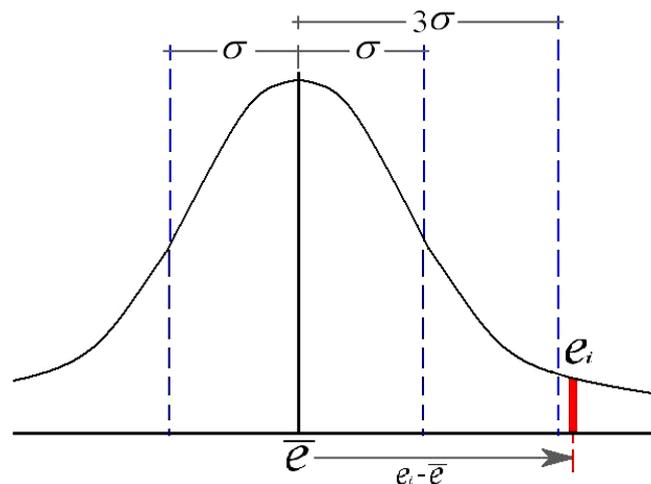


Figura 33.1: Campana de Gauss y valor atípico a ponderar (e_i) mediante la función gaussiana $g_{\sigma}(e_i)$.

El **procedimiento** de aplicación sería el siguiente:

- Detección de valores atípicos, considerando como tales aquellos que, por ejemplo, superen en 3 veces la desviación poblacional. Otra posibilidad, es mediante la diferenciación en atípicos medios y extremos en función de a qué distancia se sitúen del valor promedio y el recorrido intercuartílico.
- Ponderación mediante la función gaussiana de aquellos puntos en los que alguna de sus componentes pueda considerarse como un valor atípico. Éste vendrá ponderado según la ecuación:

$$e_{X_i}(w) = e_{X_i} \cdot g_{\sigma}(e_{X_i})$$

Dicho error ponderado servirá para realizar cualquier cálculo posterior (errores medios, desviaciones típicas, ...). De esta forma, se minimiza el efecto producido por el valor atípico, y a un mismo tiempo, permite trabajar con todos los datos sin eliminar ninguno de ellos.

3.3.1.3 Método danés

El método fue desarrollado a partir de las ideas de Krarup, Kubik y Juhl (1967). Si bien no se emplea en procesos de control de calidad posicional, sí que se ha utilizado por el Instituto Geodésico de Dinamarca para detectar errores groseros en soluciones por mínimos cuadrados¹⁸. El punto de partida es un ajuste convencional por mínimos cuadrados y, a partir de los residuos del primer ajuste, se calculan nuevos pesos para cada medida de forma individual. Éstos se basan en la siguiente función de pesos:

$$p = \begin{cases} 1 & \text{para } |v| \leq 2 \cdot \sigma \\ \text{proporcional a } \exp(-cv^2) & \text{para } |v| > 2 \cdot \sigma \end{cases}$$

Con estos nuevos pesos, se realiza un nuevo ajuste por mínimos cuadrados, repitiéndose el proceso de reponderación y ajuste hasta que se logre la convergencia (generalmente después de cinco o diez iteraciones). Las observaciones afectadas por este tipo de errores tienen peso cero y sus residuos son una medida de la magnitud del error correspondiente.

Este método ha sido suficientemente probado y resulta muy efectivo con datos erróneos en ajustes por mínimos cuadrados. Empleando técnicas de simulación, este método es más efectivo que otros en ajustes por mínimos cuadrados. Dados los buenos resultados que describe Domingo (2000) en fotogrametría, se ha realizado una adaptación del mismo para su aplicación en la detección y ponderación de atípicos en el control de calidad posicional cartográfico.

Así, el **procedimiento** de aplicación sería el siguiente:

¹⁸ A principios de los años setenta, utilizaron una rutina automática de búsqueda de errores que fue aplicada en el cálculo de problemas geodésicos de gran envergadura.

- Detección de valores atípicos, considerando como tales aquellos que, por ejemplo, superen en 3 veces la desviación poblacional (v.g. $[\bar{e}-e_{Xi}] > 3 \cdot \sigma_{mapa}$).
- Cálculo de los pesos a aplicar, empleando la ecuación:

$$W_i = e^{-k \cdot e_{Xi}^2}$$

siendo:

- W_i : el peso a aplicar en el punto i
- e_{Xi} : el valor del error en la componente X para el punto i
- k : un valor constante a determinar empíricamente
- Ponderación de los valores atípicos, siendo el error asociado al punto i en la componente X :

$$e'_{Xi} = e_{Xi} \cdot W_i$$

- En caso de que fuera necesario, repetir el proceso de ponderación.

3.3.1.4 Método de Geman y Mc.Clure

Otra de las posibilidades propuestas por Domingo (2000), es el empleo del método de Geman & Mc.Clure. En este caso, la función de pesos viene dada por una constante así como el valor asociado como parte del divisor.

El **procedimiento** de aplicación sería el siguiente:

- Detección de valores atípicos, considerando como tales aquellos que, por ejemplo, superen en 3 veces la desviación poblacional (v.g. $[\bar{e}-e_{Xi}] > 3 \cdot \sigma_{mapa}$).
- Cálculo de los pesos a aplicar, empleando la ecuación:

$$W_i = k \cdot \frac{1}{1 + e_{Xi}^2}$$

siendo, al igual que en el caso anterior:

- W_i : el peso a aplicar en el punto i
- e_{Xi} : el valor del error en la componente X para el punto i
- k : un valor constante a determinar empíricamente

- Ponderación de los valores atípicos, siendo el error asociado al punto i en la componente X :

$$e'_{Xi} = e_{Xi} \cdot W_i$$

- En caso de que fuera necesario, repetir el proceso de ponderación.

3.3.1.5 Consideraciones sobre los métodos robustos

A la hora de implementar cualquier nueva metodología, para que sea efectiva y se utilice, ha de ser relativamente fácil de aplicar. Así, a priori, el primer método mostrado (MVT – MLT), si bien es muy recomendado en otros campos de actuación (Rodríguez, 2003), podría presentar alguna complicación a la hora de llevarlo a cabo dada su dificultad de cálculo.

En cuanto al suavizado gaussiano, como todos los estándares empleados en el control posicional se basan en la hipótesis de distribución normal de los errores, cabe esperar que ofrezca unos resultados más que aceptables en su ponderación. No obstante, la ecuación a aplicar para la obtención de los pesos (la función gaussiana), también resulta algo complicada.

Por el contrario, la ventaja a priori de los dos últimos estimadores robustos (método de Geman y McClure), es que resultan de muy fácil implementación a la hora de realizar los cálculos. Sin embargo, se fundamentan en una constante k que ha de ser ajustada empíricamente mediante un intenso estudio en función del tipo de error, la desviación poblacional, el nivel de contaminación, la magnitud de los valores atípicos...

3.3.2 Estadística circular

Para el estudio de las posibles herramientas de estadística circular, se presenta una breve introducción sobre ésta al tratarse de una técnica novedosa aplicada al control de la exactitud posicional cartográfica. Seguidamente, se muestran algunos de los posibles estimadores a emplear en la detección de direcciones predominantes en la distribución de errores, Ómnibus y Mardia, así como la herramienta para calcular el

valor de dicha dirección y la posibilidad de realizar algún tipo de análisis sobre los módulos de los vectores de error.

3.3.2.1 Introducción

Los datos direccionales pueden llegar a ser una información muy valiosa a la hora de analizar cualquier tipo de información geográfica. Así, por ejemplo, se puede emplear en geología para detectar planos inclinados, fallas o puntos de unión caracterizados por sus afloramientos o hundimientos sobre el terreno (Fisher, 1995). De igual forma sucede con los desplazamientos sobre glaciares (Sanjosé, 2003), en los que pueden llegar a realizarse procesos automatizados de predicción sobre posibles cambios medioambientales en función de análisis temporales de distintos glaciares.

En el caso de fotografías aéreas o imágenes de teledetección, es fácil que se detecten posibles patrones de orientación lineal sobre fenómenos geográficos, por lo que es necesario disponer de herramientas de análisis específico que reflejen la naturaleza circular (en 2 dimensiones) o esférica (en 3 dimensiones) de los datos direccionales. Si se aplicasen técnicas de estadística circular sobre los vectores de error posicional de una BDG, podrían detectarse comportamientos sistemáticos en cuanto a su dirección.

Para representar datos circulares se emplean vectores unitarios sobre gráficos circulares de radio la unidad. En la figura 33.2 puede observarse un ejemplo de datos circulares con afloramientos de estriaciones glaciares al sur de Finlandia. Su representación por vectores unitarios se muestra en la figura 33.3a. Si el círculo se subdivide en segmentos (p.e. cada 10°) y se contabiliza el número de vectores incluido en cada uno de ellos, se obtiene un diagrama de rosa o histograma circular (figura 33.3b). Los valores direccionales de vectores unitarios, también pueden representarse de forma gráfica tal y como aparecen en la figura 33.4.

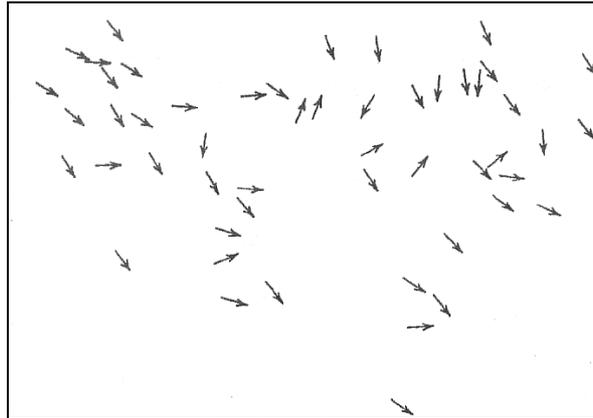


Figura 33.2: Vectores de localización y dirección de 51 estricciones en un glaciar al sur de Finlandia.

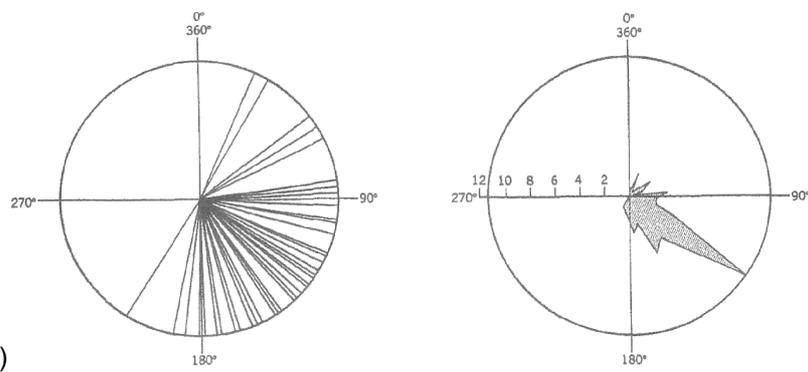


Figura 33.3: Diagramas con las direcciones de las estricciones del glaciar representadas por vectores unitarios (a) y en diagrama de rosa agrupados en intervalos de 10° (b).

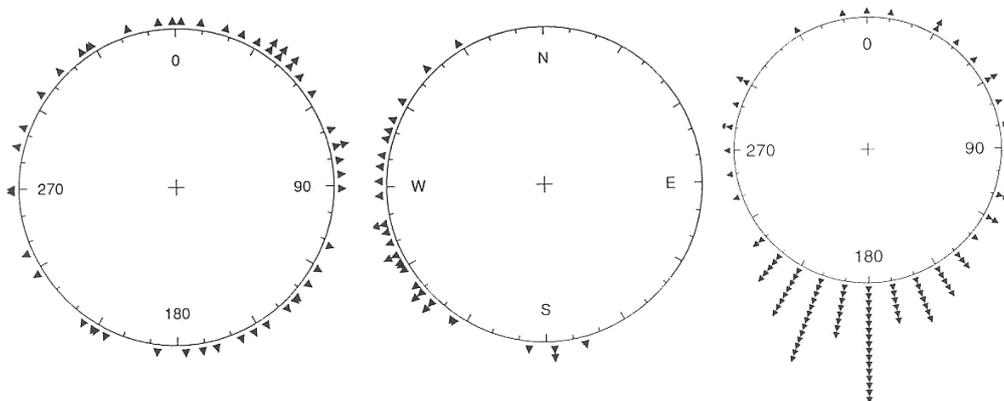


Figura 33.4: Ejemplos de representación de distintos datos circulares por vectores unitarios con comportamientos claramente diferenciados.

En una serie de vectores, su dirección dominante puede detectarse por el cálculo del vector resultante. Las coordenadas X, Y y el punto final de un vector unitario, cuya dirección viene dada por su acimut (θ), son:

$$X_i = \text{Cos}\theta \quad ; \quad Y_i = \text{Sin}\theta$$

Así, para 3 vectores unitarios (figura 33.5a), su vector resultante R se obtiene por la suma de los senos y cosenos de los vectores unitarios. Esto es análogo al valor medio de una serie de medidas escalares, siendo la dirección resultante:

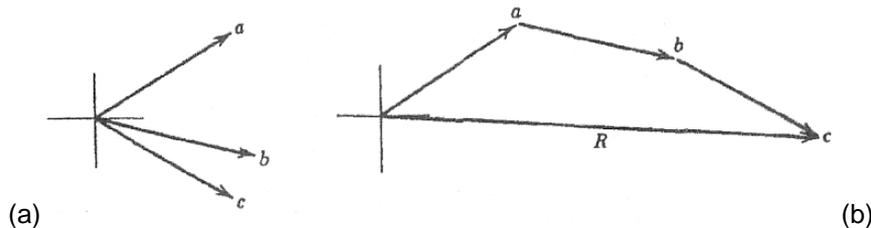
$$\bar{\theta} = A \tan\left(\frac{Y_r}{X_r}\right) = A \tan\left(\frac{\sum_{i=1}^n \text{Sin}\theta_i}{\sum_{i=1}^n \text{Cos}\theta_i}\right)$$


Figura 33.5: Dirección media de 3 vectores unitarios (a) y su vector resultante R (b).

Lógicamente, la magnitud del vector resultante depende tanto de la dispersión muestral como del número de vectores. Para poder comparar muestras de diferente tamaño, éstas han de estandarizarse dividiendo las coordenadas del vector resultante por el número de observaciones que posea la muestra:

$$\bar{C} = \frac{X_r}{n} = \frac{1}{n} \sum_{i=1}^n \text{Cos}\theta_i \quad ; \quad \bar{S} = \frac{Y_r}{n} = \frac{1}{n} \sum_{i=1}^n \text{Sin}\theta_i$$

El vector resultante informa sobre el valor de la dirección media de la muestra y la distribución de los vectores sobre éste. Por ejemplo, en la figura 33.6a se observa cómo los 3 vectores desvían ligeramente la dirección media, mientras que en la en figura 33.6b producen un vector resultante muy pequeño al estar la muestra muy dispersa. De esta forma, el vector resultante (o longitud media resultante, \bar{R}) estará comprendido en el intervalo $[0,1]$ y vendrá dado por la componente cuadrática de:

$$\bar{R} = \frac{R}{n} = \sqrt{\bar{C}^2 + \bar{S}^2}$$

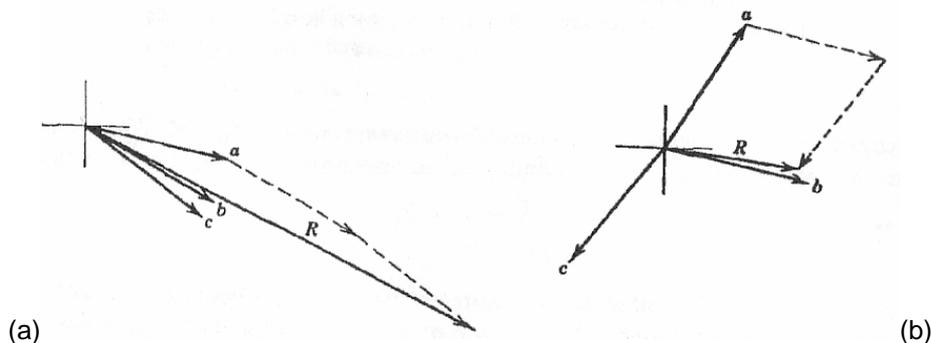


Figura 33.6: Estudio de la dispersión a partir del vector resultante: (a) vectores muy agrupados en una dirección, y (b) vectores con direcciones muy dispersas.

Se trata de una medida de dispersión similar a la varianza de la estadística clásica, pero expresada en sentido opuesto. Esto es, para valores grandes de \bar{R} , se corresponden observaciones muy agrupadas con una pequeña dispersión; mientras que los valores pequeños representan observaciones muy dispersas (ver figuras 33.4 y 33.6).

Tal y como sucede en estadística clásica, para poder realizar el análisis sobre una distribución circular, hay que disponer de un modelo de probabilidad de características conocidas frente a los datos que se deseen estudiar. Existen analogías circulares a las distribuciones univariantes, siendo la más empleada la distribución de Von Mises. Se trata de la equivalente a la distribución normal y, de forma análoga, posee 2 parámetros que indican sus características: la dirección media (θ) y el parámetro de concentración (k).

La distribución de Von Mises es unimodal y simétrica con respecto a la dirección media. Cuando aumenta el parámetro de concentración, aumenta la probabilidad de que se observe una dirección muy próxima a la dirección media; mientras que si $k = 0$, significa que todas las direcciones son igualmente probables y la distribución se transforma en una circular uniforme. La figura 33.6a muestra la forma de la distribución de Von Mises para diferentes valores de k . Asimismo, la distribución también puede representarse de forma convencional, a modo de histograma de frecuencias, como se aprecia en la figura 33.6b.

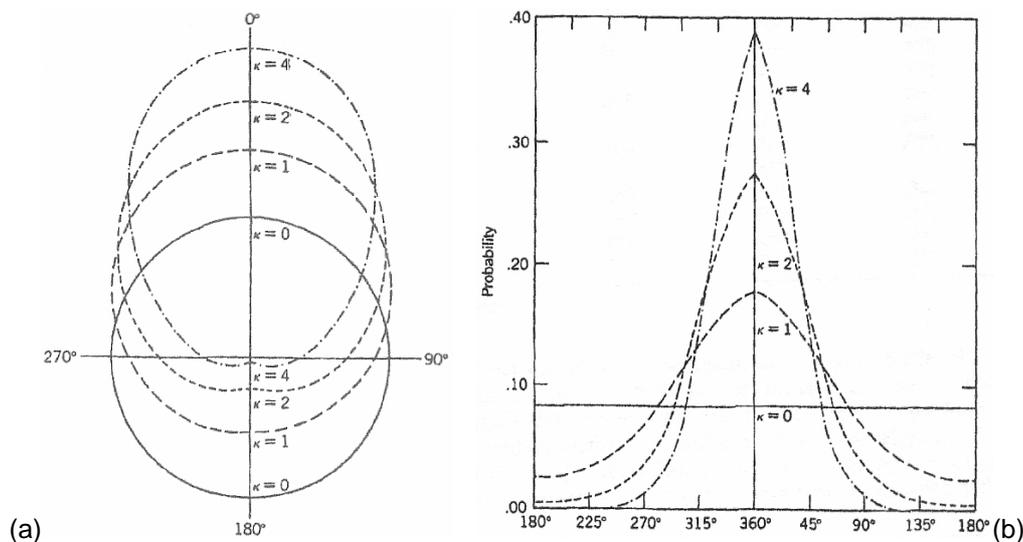


Figura 33.6: Distribuciones de Von Mises con distintos parámetros de concentración: (a) gráfico circular, y (b) gráfico representado a modo de distribución de probabilidad convencional (Gumbel, 1954).

Tabla 33.2: Umbrales máximos estimados del parámetro de concentración k para valores calculados de \bar{R} (Gumbel, 1954)

\bar{R}	k	\bar{R}	k	\bar{R}	k
0.00	0.00000	0.35	0.74783	0.70	2.01363
.01	.02000	.36	.77241	.71	2.07685
.02	.04001	.37	.79730	.72	2.14359
.03	.06003	.38	.82253	.73	2.21425
.04	.08006	.39	.84812	.74	2.28930
.05	.10013	.40	.87408	.75	2.36930
.06	.12022	.41	.90043	.76	2.45490
.07	.14034	.42	.92720	.77	2.54686
.08	.16051	.43	.95440	.78	2.64613
.09	.18073	.44	.98207	.79	2.75382
.10	.20101	.45	1.01022	.80	2.87129
.11	.22134	.46	1.03889	.81	3.00020
.12	.24175	.47	1.06810	.82	3.14262
.13	.26223	.48	1.09788	.83	3.30114
.14	.28279	.49	1.12828	.84	3.47901
.15	.30344	.50	1.15932	.85	3.68041
.16	.32419	.51	1.19105	.86	3.91072
.17	.34503	.52	1.22350	.87	4.17703
.18	.36599	.53	1.25672	.88	4.48876
.19	.38707	.54	1.29077	.89	4.85871
.20	.40828	.55	1.32570	.90	5.3047
.21	.42962	.56	1.36156	.91	5.8522
.22	.45110	.57	1.39842	.92	6.5394
.23	.47273	.58	1.43635	.93	7.4257
.24	.49453	.59	1.47543	.94	8.6104
.25	.51649	.60	1.51574	.95	10.2716
.26	.53863	.61	1.55738	.96	12.7661
.27	.56097	.62	1.60044	.97	16.9266
.28	.58350	.63	1.64506	.98	25.2522
.29	.60625	.64	1.69134	.99	50.2421
.30	.62922	.65	1.73945	1.00	∞
.31	.65242	.66	1.78953		
.32	.67587	.67	1.84177		
.33	.69958	.68	1.89637		
.34	.72356	.69	1.95357		

Resulta difícil de determinar directamente el valor de k , aunque puede obtenerse a partir de \bar{R} si se asume que la muestra procede de una población que sigue una distribución de Von Mises. En la tabla 33.2 se muestran los umbrales máximos de probabilidad para el cálculo del coeficiente k a partir de \bar{R} .

3.3.2.2 Test Ómnibus

La uniformidad de una muestra de datos circulares, puede ser una propiedad difícil de ver a simple vista, dependiendo del modelo de uniformidad escogido. Si únicamente se trata de detectar cualquier tipo de uniformidad frente a cualquier modelo de distribución (unimodal, bimodal, trimodal...), ha de seleccionarse un test Ómnibus o global. No obstante, dado que el contraste se realiza frente a cualquier alternativa, es

posible que no sea muy efectivo a la hora de realizar el contraste frente a una alternativa concreta (p.e.: unimodal).

Partiendo de que se trabaja con datos continuos (desagrupados), se calculan los valores X_i :

$$X_1 = \frac{\theta_1}{2\pi}; X_2 = \frac{\theta_2}{2\pi} \dots X_n = \frac{\theta_n}{2\pi}$$

siendo θ_i las direcciones o acimutes de cada vector unitario de la muestra, para los datos ordenados en función de θ : $\theta_1 \leq \theta_2 \leq \theta_3 \leq \dots \leq \theta_n$

Asimismo, han de calcularse los siguientes estadísticos:

$$D_n^+ = \text{Máximo de } : \frac{1}{n} - x_1; \frac{2}{n} - x_2 \dots \frac{i}{n} - x_i \dots 1 - x_n$$

$$D_n^- = \text{Máximo de } : x_1; x_2 - \frac{1}{n}; x_3 - \frac{2}{n} \dots x_i - \frac{i-1}{n} \dots x_n - \frac{n-1}{n}$$

$$V_n = D_n^+ + D_n^-$$

$$V = V_n \left(\sqrt{n} + 0,155 + \frac{0,24}{\sqrt{n}} \right)$$

El test Ómnibus se basa en la hipótesis de contraste:

$$\begin{cases} H_0 : V < V_{Kuiper} \xrightarrow{\text{entonces}} \text{Distribución Circular Uniforme} \\ H_1 : V \geq V_{Kuiper} \xrightarrow{\text{entonces}} \text{Distribución con una dirección dominante} \end{cases}$$

siendo los valores críticos de decisión para V_{Kuiper} los mostrados en la tabla 33.3.

Tabla 33.3: Umbrales máximos de decisión estimados para el estadístico (V) modificado de Kuiper en función del nivel de significación (Stephens, 1974).

	Punto porcentual 100 α %			
α	0,15	0,10	0,05	0,01
V_{Kuiper}	1,537	1,620	1,747	2,001

3.3.2.3 Test de Mardia

Asumiendo que las observaciones proceden de una distribución estadística de Von Mises, la hipótesis equivalente a contrastar es que si el parámetro $k \approx 0$, es porque se trata de una distribución circular uniforme. En términos formales, tendríamos las hipótesis:

$$\begin{cases} H_0 : k \approx 0 \xrightarrow{\text{entonces}} \text{Distribución Circular Uniforme} \\ H_1 : k > 0 \xrightarrow{\text{entonces}} \text{Distribución con una dirección dominante} \end{cases}$$

Para realizar un contraste de la aleatoriedad en la distribución de las direcciones, tan sólo habrá que calcular el valor medio de \bar{R} , comparando el estadístico con el valor crítico de \bar{R} para el nivel de significación deseado. Si las observaciones siguen una distribución circular uniforme, puede esperarse que \bar{R} sea muy próximo a 0; mientras que si se supera el valor crítico (valores muy grandes), se rechaza H_0 asumiendo que existe una dirección dominante.

Este test fue originariamente desarrollado por Lord Rayleigh a finales del s. IXX, aunque a finales del s. XX Mardia (1972) diseñó una versión actualizada con un mejor comportamiento al analizar la aleatoriedad en las distribuciones. En la tabla 33.4 se muestran los valores críticos de \bar{R} en función del nivel de confianza y el tamaño muestral ($1-\alpha, n$).

Tabla 33.4: Valores críticos de \bar{R} para el test de Rayleigh en función del nivel de confianza y el tamaño muestral (Mardia, 1972).

n	Nivel de significación α (%)			
	10 %	5 %	2,5 %	1 %
4	0.768	0.847	0.905	0.960
5	.677	.754	.816	.879
6	.618	.690	.753	.825
7	.572	.642	.702	.771
8	.535	.602	.660	.725
9	.504	.569	.624	.687
10	.478	.540	.594	.655
11	.456	.516	.567	.627
12	.437	.494	.544	.602
13	.420	.475	.524	.580
14	.405	.458	.505	.560
15	.391	.443	.489	.542
16	.379	.429	.474	.525
17	.367	.417	.460	.510
18	.357	.405	.447	.496
19	.348	.394	.436	.484
20	.339	.385	.425	.472
21	.331	.375	.415	.461
22	.323	.367	.405	.451
23	.316	.359	.397	.441
24	.309	.351	.389	.432
25	.303	.344	.381	.423
30	.277	.315	.348	.387
35	.256	.292	.323	.359
40	.240	.273	.302	.336
45	.226	.257	.285	.318
50	.214	.244	.270	.301

3.3.2.4 Dirección predominante

Cuando se ha detectado la existencia de una dirección dominante, bien mediante el test Ómnibus o el test de Mardia, puede efectuarse una prueba para definir el valor de la dirección predominante.

La forma más sencilla de determinar dicho valor es mediante el cálculo de la dirección o acimut medio ($\bar{\theta}$) y su intervalo de confianza (por ejemplo para $1-\alpha = 0,95$). El intervalo de confianza se basa en el error estándar en la determinación de la dirección media, en función del tamaño muestral y su dispersión. El error estándar aproximado de la dirección media viene dado por:

$$S_e = \frac{1}{\sqrt{n \cdot \bar{R} \cdot k}}$$

Asumiendo que los errores se distribuyen de forma normal, el intervalo de confianza para la dirección media vendrá dado por:

$$\bar{\theta} \pm Z_{\alpha} \cdot S_e \rightarrow \text{siendo para } \alpha = 0,05 \rightarrow \bar{\theta} \pm 1,96 \cdot S_e$$

Si el intervalo de confianza es muy amplio, el valor de la dirección media detectada carecerá de sentido. En este caso, es el propio usuario del test quien ha de decidir si el intervalo es o no excesivamente amplio.

3.3.2.5 Estudio de los módulos

Por último, y dado que los estimadores circulares tan sólo realizan un análisis sobre vectores unitarios, se ha tratado de desarrollar algún tipo de estudio que tenga en cuenta la magnitud y la distribución de los módulos de error en función de la dirección en la que éstos se encuentren.

Para ello, se ha optado por trabajar con los datos presentados en forma de gráficos Acimut / Módulo (A/M). Esto es, sobre el eje horizontal se distribuyen los acimutes de 0 a 2π , mientras que sobre el eje vertical se representa la magnitud del módulo del vector de error (ver ejemplo en la figura 33.7).

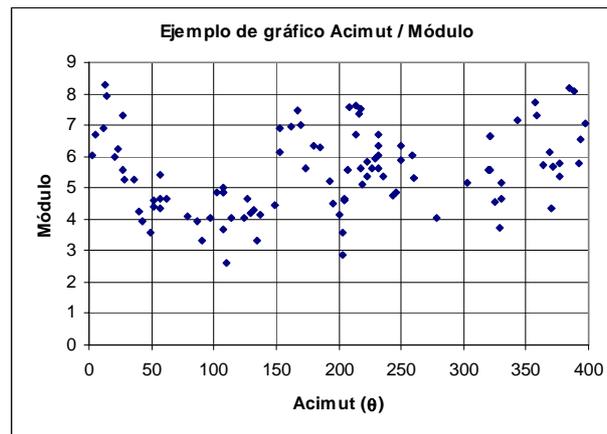


Figura 33.7: Ejemplo de gráfico de relación Acimut / Módulo con la distribución de los vectores de error.

Sobre las series de datos desarrolladas de esta forma, se pueden realizar diferentes análisis empleando estadística clásica, que analicen la distribución de los módulos con respecto a su dirección, como por ejemplo: estudio de los principales estadísticos (media, mediana, desviación ...), test de contraste (test de rachas para el estudio de la aleatoriedad...), etcétera.

3.3.2.6 Consideraciones sobre los estimadores circulares

Cabe destacar que no se han encontrado experiencias previas de aplicación de estimadores circulares en el análisis de los errores posicionales sobre BDGs. Dado que los errores que se puedan producir sobre BDGs es posible que estén afectados por algún tipo de sistematismo (p.e. por una incorrecta georreferenciación en una imagen de teledetección o por errores en el sensor de captura de los datos), los estimadores circulares pueden ser una herramienta eficaz a la hora de detectarlos. Como toda información georreferenciada posee una componente en posición horizontal (XY), el estudio de este tipo de desplazamientos de forma conjunta, empleando el acimut y el módulo del vector de error, puede proporcionar a usuarios y productores información muy valiosa para la mejora continua de los productos.

Asimismo, tampoco se han encontrado trabajos basados en estimadores circulares sobre el análisis de los módulos con respecto a su dirección. Es de destacar, que en estadística circular, siempre se ha trabajado sobre vectores unitarios, por lo que el

tratar de plantear algún tipo de análisis sobre esta componente podría proporcionar una mayor información de la BDG analizada.

3.4 La simulación como herramienta en el análisis de la calidad

No cabe duda que la simulación proporciona una gran herramienta de trabajo a los investigadores. En ella se basan, entre otros, muchos de los modelos de predicción: medicina, biología, economía, medio ambiente, hidrología, riesgos naturales,.... Así, en cartografía, la simulación se puede orientar tanto hacia aspectos de la calidad temática como posicional (Ariza, 2000; Ariza, 2002).

El Método de Montecarlo es uno de los más empleados. Se podría afirmar que la simulación es una simple programación en un computador de un modelo que describe el comportamiento de un determinado sistema. De esta forma, se pueden obtener una serie de conclusiones con las que apoyar la toma de decisiones final. La simulación no es otra cosa que una metodología de análisis basada en un soporte informático y en la teoría de sistemas. Independientemente del tipo de simulación que se realice, y del tipo de sistema que se esté modelizando, la simulación se puede definir según el diagrama de la figura 34.1.

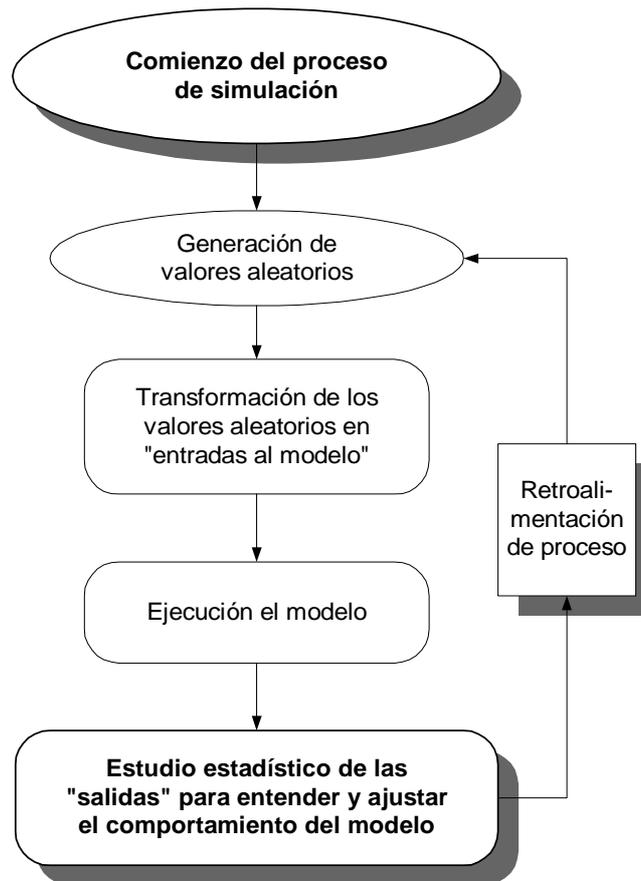


Figura 34.1: Diagrama de flujo para un proceso de simulación.

En todo proceso de simulación, resulta de vital importancia el hecho de que la generación de números aleatorios (seudoaleatorios) sea correcta, así como el poder contrastar adecuadamente dicha aleatoriedad. Por otra parte, el tamaño muestral es otro de los aspectos importantes a tener en cuenta.

3.4.1 Números aleatorios – pseudoaleatorios

Dado que la simulación se basa en procesos aleatorios, si los datos aleatorios no se obtienen correctamente, los resultados de la investigación tampoco serán correctos. Así, los números aleatorios se podrían obtener tanto desde procedimientos al azar, no predecibles, como desde números pseudoaleatorios.

En el caso de los procedimientos al azar, se ha de recurrir a procedimientos más costosos y lentos (p.e.: lanzar una moneda o un dado, escoger una bola dentro de una

caja negra...), en comparación con los datos pseudoaleatorios generados artificialmente.

Los números pseudoaleatorios, tal y como su propio nombre indica, son falsos números aleatorios. Es decir, aparentemente son números aleatorios, pero en realidad son predecibles si se conoce el algoritmo que los ha generado. Para ello, se pueden emplear “semillas” o valores iniciales con los que alimentar el algoritmo, y que éstos sean completamente aleatorios. Dichos valores iniciales se pueden obtener, por ejemplo, de distintas tablas.

Los números pseudoaleatorios se suelen obtener en forma de distribución estadística uniforme $U(0,1)$ debido a que son las más fáciles de generar y se pueden convertir a cualquier otro tipo de distribución estadística mediante un proceso de transformación inversa (Ariza, 2002). Así, las principales características en la generación de números aleatorios podrían resumirse como las que se muestran en la tabla 34.1.

Tabla 34.1. Principales características de los algoritmos de generación de números pseudoaleatorios.

Algoritmo	<ul style="list-style-type: none"> - Distribución estadística uniforme $U(0,1)$ - Longitud de la serie - Rapidez de ejecución - Poco consumo de memoria - Portabilidad - Sencillez de implementación - Reproducibilidad - Mutabilidad - Periodo largo
-----------	---

En cuanto a los métodos para la generación de números pseudoaleatorios, existe una extensa bibliografía (ver Ríos, 1997), presentándose a continuación algunos de ellos.

- Método congruencial

Este es uno de los más empleados en la actualidad y consiste en la generación de un valor R_{i+1} como resto de una operación de división que se convierte a un valor $r_i \rightarrow U(0,1)$. Los pasos a seguir son:

$$1) R_{i+1} = \text{RESTO} \left\{ \frac{a \cdot R_i + b}{T} \right\}$$

$$2) r_{i+1} = \frac{R_{i+1}}{T}$$

Para entender mejor el proceso, la tabla 34.2 presenta un ejemplo numérico de cálculo de siete números generados por este método. Las constantes R_0, a, b, T , se eligen de tal forma que el generador tenga el mayor número de cualidades positivas. El parámetro T determina el número de cifras de R_i , y se suelen tomar potencias de 2 ($T=2^K$). Los valores de a y b ($a, b \in [0, 1, \dots, T-1]$) deben cumplir ciertas relaciones con T para reducir la autocorrelación y retrasar la aparición de ciclos. Cuando $b=0$, el generador se denomina multiplicativo.

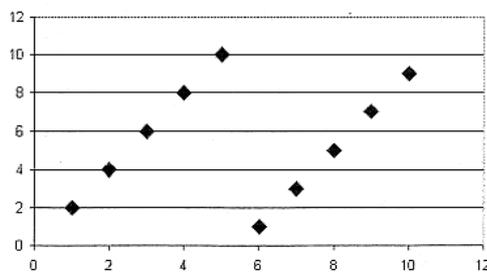
Este tipo de métodos es más rápido y, aunque se reduce un poco la longitud del ciclo, la aparición de éstos no es un problema si T se elige suficientemente grande. Valores comúnmente usados son (Ríos, 1997) $T = 2147493647, a = 7^5 = 16807$, donde a es raíz de $2^{31}-1$, por lo que el generador es de periodo máximo. La tabla 34.3 muestra distintas series en función de los parámetros iniciales utilizados y cuyo comportamiento puede visualizarse en la figura 34.2 al representar un valor frente al anterior.

Tabla 34.2: Ejemplo de la generación de números pseudoaleatorios por el método del congruencial (Ríos, 1997).

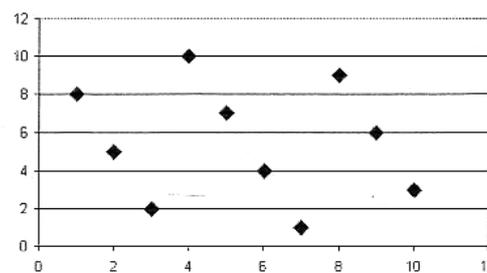
Parámetros: $a=5, b=0, T=11, Semilla = 1$			
$R_1 = \text{Resto } \{(aR_0+b)/T\} = \text{Resto } \{(5 \times 1 + 0) / 11\} = 6$	$r_1 = R_1/T = 6 / 11$		
$R_2 = \text{Resto } \{(aR_1+b)/T\} = \text{Resto } \{(5 \times 6 + 0) / 11\} = 8$	$r_2 = R_2/T = 8 / 11$		
$R_3 = \text{Resto } \{(aR_2+b)/T\} = \text{Resto } \{(5 \times 8 + 0) / 11\} = 7$	$r_3 = R_3/T = 8 / 11$		
$R_4 = \text{Resto } \{(aR_3+b)/T\} = \text{Resto } \{(5 \times 7 + 0) / 11\} = 2$	$r_4 = R_4/T = 2 / 11$		
$R_5 = \text{Resto } \{(aR_4+b)/T\} = \text{Resto } \{(5 \times 2 + 0) / 11\} = 10$	$r_5 = R_5/T = 10 / 11$		
$R_6 = \text{Resto } \{(aR_5+b)/T\} = \text{Resto } \{(5 \times 10 + 0) / 11\} = 6$	$r_6 = R_6/T = 6 / 11$		
$R_7 = \text{Resto } \{(aR_6+b)/T\} = \text{Resto } \{(5 \times 6 + 0) / 11\} = 8$	$r_7 = R_7/T = 8 / 11$		

Tabla 34.3: Ejemplo de series generadas por el método del congruencial (Ríos, 1997).

Caso	a	b	T	Semilla	Resultados															
1	5	0	11	1	6	8	7	2	10	6	8	7	2	10	6	8	7	2		
2	6	0	11	3	7	9	10	5	8	4	2	1	6	3	7	9	10	5		
3	7	0	11	5	2	3	10	4	6	9	8	1	7	5	2	3	10	4		
4	5	0	13	5	12	8	1	5	12	8	1	5	12	8	1	5	12	8		
5	6	0	13	1	6	10	8	9	2	12	7	3	5	4	11	1	6	10		
6	7	0	13	10	5	9	11	12	6	3	8	4	2	1	7	10	5	9		



(a) Caso 3 de la tabla 34.3



(b) Caso 2 de la tabla 34.3

Figura 34.2: Ejemplo de comportamiento de un generador congruencial (Ríos, 1997).

El principal problema del método es la autocorrelación, ya que los valores de R_j , R_{j+1} están relacionados por una expresión lineal. Para evitarlo, algunos programas generan series de datos que luego son extraídos de forma aleatoria. El coeficiente de correlación viene dado por:

$$\rho \approx \frac{1}{a} - \frac{6b}{aT} \left(1 - \frac{b}{T}\right) \pm \frac{a}{T}$$

- Método de registro de desplazamiento

Este método supone una generalización de los generadores congruenciales con recursiones de mayor orden. Para $K \geq 1$, T primo, se define:

$$1) \quad R_n = \text{RESTO} \left\{ \frac{a_1 \cdot R_n + \dots + a_k \cdot R_{n-k}}{T} \right\}$$

$$2) \quad r_n = \frac{R_n}{T}$$

- Método de Fibonacci retardado

Este método es otra generalización de los generadores congruenciales. En este caso se parte de una semilla inicial formada por una serie de números Fibonacci (1,1,2,3,5,8,13,...), en la que se utiliza la recursividad.

$$R_j = R_{j-r} \oplus R_{j-s}$$

donde: r y s son retardos ($r > s$); y \oplus es una operación +, -, XOR

- Método específico para generar una distribución Normal: **Box-Müller**

Existen algunos métodos específicos para generar series de números aleatorios que se distribuyan normalmente (a partir de una uniforme). Uno de los más empleados es el de Box-Müller (Box y Müller, 1958). Las ecuaciones a emplear son:

$$x_1 = \sqrt{-2 \cdot \text{Ln}(r_1)} \cdot \text{Cos}(2 \cdot \pi \cdot r_2)$$

$$x_2 = \sqrt{-2 \cdot \text{Ln}(r_1)} \cdot \text{Sen}(2 \cdot \pi \cdot r_2)$$

$$x_3 = \dots \dots \dots \dots$$

siendo x_1 , x_2 , ... los valores generados aleatoriamente y distribuidos según una normal; r_1 y r_2 los números aleatorios de una distribución uniforme.

3.4.2 Contraste de aleatoriedad

Dado que los estándares de control de calidad posicional se basan, entre otros aspectos, en la presumible aleatoriedad en la distribución de los errores, un buen generador de números aleatorios parece el más adecuado para trabajar en procesos de simulación. No obstante, en aquellos casos en los que no se disponga de información contrastada sobre el generador de datos aleatorios o sobre la muestra, es necesario el hacer un contraste de dicha aleatoriedad. Éste se podrá realizar mediante un contraste de rachas, de huecos... (Spiegel, 1970; Caridad, 1985; Earickson y Harlin, 1994).

En el caso de trabajar sobre datos reales, el orden de los puntos de control, generalmente viene dado por el orden seguido en la toma de datos sobre la FME. Así, aquellos puntos correlativos serán los que se encuentren en una misma zona geográfica. Si existiera algún tipo de comportamiento diferenciado en un área determinada, éste sería detectado como una racha continua en el contraste de aleatoriedad.

- Contraste de rachas

Este método se puede aplicar tanto sobre datos nominales como sobre secuencias binarias en función de que los datos se encuentren o no ordenados de forma creciente, dos a dos. De esta forma, "racha creciente" (o decreciente) será aquel dato o conjunto de datos en el que todos los elementos de la muestra estén formados por un mismo valor binario ("1" o "0"), y este estudio se realice en el mismo orden en el que han sido extraídos los datos. Una vez contabilizado el número de rachas, se estudia su aleatoriedad mediante el establecimiento de una hipótesis nula en la que se supone un comportamiento normal asintótico según una:

$$N(\mu, \sigma^2) = N\left(\frac{2 \cdot n - 1}{3}, \frac{16 \cdot n - 29}{90}\right)$$

Observando el comportamiento de las rachas, se tendrá una idea aproximada del comportamiento de los datos: si las rachas son muy extensas indican un comportamiento no aleatorio, mientras que las rachas muy cortas pueden dar una idea de comportamiento aleatorio (figura 34.4).

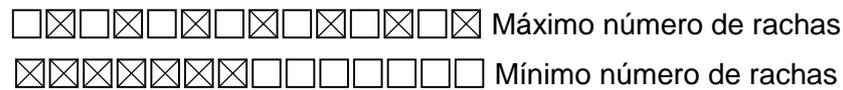


Figura 34.3: Ejemplo del comportamiento de dos rachas extremas para una serie de 14 datos.

Asimismo, un exceso o un defecto en el número y longitud de las rachas, también puede indicar la falta de aleatoriedad de la serie. Para comprobarlo, se pueden contrastar las frecuencias observadas con una distribución $U(0,1)$, empleando un estadístico similar al χ^2 entre ambas.

▪ Contraste de rachas de Wald–Wolfowitz

Se trata de un test con el que poder determinar si una muestra ordinal se distribuye de forma aleatoria. Así, se parte de dos hipótesis:

- H_0 : la muestra es aleatoria simple.
- H_1 : la muestra no es aleatoria.

Para el cálculo de dicho test, en primer lugar se ha de obtener la mediana (m_e) de la muestra en cuestión, ordenada de menor a mayor. Las observaciones z_i se clasifican en dos grupos (p.e.: "w" y "s"): si $z_i \leq m_e$ se asigna z_i al grupo "s", y si $z_i > m_e$ se asigna z_i al grupo "w". Con estos dos conjuntos, refundidos en el mismo orden en el que se obtuvieron, se denomina "racha" al conjunto de datos de la misma muestra (" R_w " o " R_s ") que aparezcan contiguos. El estadístico del test es el número de rachas: $R = R_w + R_s$.

La regla de decisión es que se acepta la hipótesis H_0 al nivel α si, en función del número de rachas obtenidas, se cumple que:

$$R \in C_0 = (r_{1-\alpha}; n + m)$$

siendo $r_{1-\alpha}$ valores tabulados. En el caso de que $n, m \geq 20$, se puede tomar la distribución como asintótica a la normal, sin necesidad de disponer de las tablas específicas de Wald–Wolfowitz.

$$R \rightarrow N(\mu_R; \sigma_R^2)$$

Así, los valores de media y varianza vienen determinados por:

$$\mu_R = \frac{2n \cdot m}{n + m} + 1 \quad \sigma^2 = \frac{2n \cdot m \cdot (2n \cdot m - n - m)}{(n + m)^2 \cdot (n + m - 1)}$$

$$z = \frac{R - \mu_R \pm 0.5}{\sigma}$$

siendo:

- $r_{1-\alpha}$: estadístico con el que comparar el número de rachas obtenidas.
- α : el nivel de significación
- n, m : el número de elementos que componen cada uno de los subconjuntos ("w" y "s")
- \pm : el signo de la ecuación será positivo (+) si $R < \mu_R$; y negativo (-) si $R \geq \mu_R$

La regla de decisión es que se acepta H_0 si: $Z \in (z_{\alpha/2}; z_{1-\alpha/2})$

▪ Contraste de huecos

Fijados dos valores α, β tales que $0 < \alpha < \beta < 1$, se considera que una sucesión presenta un hueco de longitud $m+1$ si $U_j, U_{j+m} \in [\alpha, \beta]$, pero tal que $U_{j+1}, \dots, U_{j+m-1} \in [\alpha, \beta]$. Bajo la hipótesis de aleatoriedad de la serie, la distribución de los huecos debe seguir una distribución geométrica de parámetro $P(\alpha \leq X_j \leq \beta) = \beta - \alpha = q$. Esto es, que la probabilidad del hueco de longitud m [$P(\text{hueco longitud} = m) = q^m (1-q)$].

3.4.3 Contraste de normalidad

Otro de los aspectos en los que se fundamenta la mayoría de los estándares para el control de calidad posicional, es en la presumible distribución estadística de los datos según una normal $N(\mu, \sigma)$. Por ello, el usuario del estándar debería realizar un contraste previo para comprobar que los datos se asemejan a una normal. Dos de los posibles chequeos son mediante un contraste bien del tipo Kolmogorov-Smirnov o bien del tipo D'Agostino.

▪ Contraste de normalidad según el test de Kolmogorov-Smirnov

Este último contraste, previo al del propio control de calidad, es uno de los más adecuados a la hora de estudiar la semejanza de una determinada distribución a una función (en este caso la función normal).

Al igual que en el estudio de la aleatoriedad, se parte de dos hipótesis:

- H_0 : la muestra se asemeja a una distribución normal.
- H_1 : la muestra no sigue una distribución normal.

El test analiza las diferencias existentes entre la distribución de los errores (para cada una de las componentes X e Y), con la distribución normal. Para ello se basa en los diagramas de frecuencias acumuladas: en primer lugar se ordenan todos los valores obtenidos para cada una de las componentes, de menor a mayor, obteniendo así los valores de frecuencias acumuladas observadas entre 0 y 1 . En siguiente lugar, se tipifican los valores según:

$$Ztip_i = \frac{e_i - \bar{e}}{S}$$

De esta forma, se puede obtener en cada punto (i), el valor de la función de distribución normal estándar acumulativa:

$$f(z;0,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Por último, se calculan las diferencias (D_i) entre el valor de la frecuencia acumulada observada [$S(x)$] y el de la función de distribución normal estándar [$F^*(x)$].

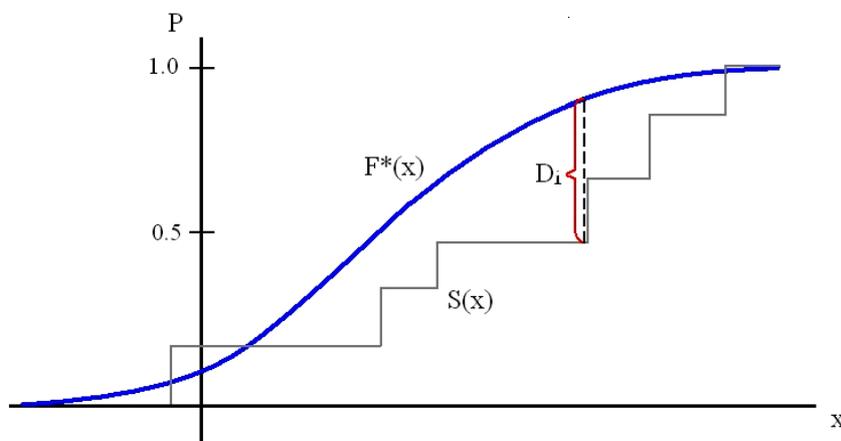


Figura 34.4: Distribución de frecuencias acumuladas observadas [$S(x)$] y de la función normal [$F^*(x)$]. ⁽¹⁹⁾

El test de Kolmogorov-Smirnov proporciona determinados valores críticos tabulados que indican si la distribución analizada se asemeja a una normal con un 95 % de nivel

¹⁹ Siendo D_i la distancia máxima entre ambas.

de confianza. A modo de ejemplo, en la tabla 34.4 se presentan algunos de éstos valores críticos en función del número de puntos de control.

Tabla 34.4: Valores límite de Kolmogorov-Smirnov.

Tamaño muestral (n)	Valor límite de Kolmogorov-Smirnov ($\alpha=0.05$)
15	0.338
20	0.294
25	0.27
30	0.24
35	0.23
> 35	$1.36/\sqrt{n}$

De esta forma, se aceptará la hipótesis nula H_0 , si la distancia máxima ($D_{i\ max}$) para cada una de las componentes analizadas (p.e. X e Y) es inferior al estadístico proporcionado por Kolmogorov-Smirnov: $D_{i\ max} < Kolmogorov$.

- Contraste de normalidad según la prueba de D'Agostino

Al igual que en el caso anterior, se parte de dos hipótesis:

- H_0 : la muestra se asemeja a una distribución normal.
- H_1 : la muestra no sigue una distribución normal.

El procedimiento consiste en ordenar la serie de n datos ($x_1, x_2, x_3, \dots, x_n$) de menor a mayor, obteniendo los valores ($x_1^*, x_2^*, x_3^*, \dots, x_n^*$). El estadístico de verificación es:

$$D(w) = \frac{\sum_{k=1}^n x_k^* \left(k - \frac{n+1}{2} \right)}{\sqrt{n^3 S_{xx}}} \quad \text{siendo: } S_{xx} = \sum_{k=1}^n (x_k - \bar{x})^2$$

La región crítica es el exterior (cerrado) del intervalo ($k1, k2$), donde $k1$ y $k2$ se obtienen para cada nivel de confianza de los valores determinados por D'Agostino. Se aceptará la hipótesis nula H_0 , si el estadístico $D(w)$ pertenece a dicho intervalo.

3.4.4 Tamaño muestral

Otro de los aspectos a tener en cuenta en los procesos de simulación, es el tamaño de la muestra que hay que generar para contrastar correctamente las hipótesis. Si bien en los estándares objeto de estudio tan solo se menciona “...al menos veinte puntos de control...”, es conveniente realizar un análisis algo más exhaustivo al respecto.

Aunque en los procesos de simulación tan sólo se encuentran referencias a la secuencia de generación de aleatorios, el tamaño muestral tiene una gran importancia porque influye en la precisión de la estimación (Ariza, 2002). Al ser los estimadores de la varianza inversamente proporcionales al número de datos, se deberá trabajar con una muestra lo suficientemente grande como para que garantice una precisión estimada a priori que satisfaga los resultados que se quieran obtener.

Así, la muestra ha de ser lo mayor posible porque, utilizando un estimador consistente, al aumentar el tamaño de la muestra el estimador converge en probabilidad hacia su valor real, disminuyendo el error. De esta forma, se denomina fracción muestral f ($f = n/N$, tal que $0 < f \leq 1$) a la parte de la población investigada, dependiendo la precisión de una muestra tanto de n como de f . Por ejemplo, para el caso de la estimación del valor medio, el error estándar es:

$$SE = \sqrt{\frac{N-n}{N-1}} \cdot \frac{S}{\sqrt{n}}$$

En las poblaciones infinitas ($N \rightarrow \infty$), el primer cociente de la expresión anterior tiende a la unidad, por lo que al aumentar n disminuye SE y aumenta la precisión en la estimación. Al aumentar el tamaño muestral n , se puede conseguir reducir el error en la estimación o aumentar el intervalo de confianza.

Otro factor importante es la variabilidad de la muestra. Ésta afecta a la precisión de tal forma que, si la variabilidad es grande, la muestra también tendrá que serlo para alcanzar un nivel de precisión predeterminado.

Este razonamiento lleva a la paradoja de que para establecer el tamaño muestral es necesario disponer de cierta información sobre la propia muestra (con estudios previos, estudios de tipo piloto o por un planteamiento analítico de las hipótesis más desfavorables). La tabla 34.5 presenta la formulación de los tamaños necesarios en

tres casos concretos (estimación de proporciones, valores medios y sumas para una precisión y nivel de confianza determinados), pudiendo recurrirse a cualquier manual especializado en muestreos para determinar la formulación apropiada (Azorín y Sánchez-Crespo, 1986).

Tabla 34.5: Tamaños muestrales²⁰ (Ariza, 2002).

Estimación de	Fórmula	Simplificaciones
Una proporción en un muestreo con reemplazamiento	$n = \frac{N \cdot P \cdot Q}{P \cdot Q + \frac{(N-1) \cdot e^2}{z_{\alpha/2}^2}}$	$\frac{z_{\alpha/2}^2}{4 \cdot e^2}$
Una media	$n = \frac{N \cdot \sigma^2}{\sigma^2 + \frac{N \cdot e^2}{z_{\alpha/2}^2}}$	Si se puede considerar la precisión como proporcional a la varianza: $e = k \cdot \sigma_X$ entonces: $n = \frac{N}{1 + \frac{N \cdot k^2}{z_{\alpha/2}^2}}$
Total o suma de varias variables aleatorias	$n = \frac{N \cdot \sigma^2}{\sigma^2 + (N-1) \left(\frac{e}{N \cdot z_{\alpha/2}} \right)^2}$	$n = \left(\frac{N \cdot \sigma \cdot z_{\alpha/2}}{e} \right)^2$

3.4.5 Consideraciones sobre la simulación

Si bien estos estudios teóricos pueden resultar de cierta ayuda a la hora de seleccionar el tamaño muestral, el empleo de técnicas de simulación es un factor determinante. De esta forma, el estudio se puede realizar analizando la variabilidad obtenida en resultados de laboratorio bajo condiciones controladas.

En este tipo de trabajos, el investigador posee un control de los datos poblacionales, pudiendo observar cómo afecta el diferente tamaño muestral a los resultados. Estudiando la relación población–muestra, se puede estimar el tamaño adecuado de ésta en función del nivel de confianza con el que se pretendan obtener los resultados.

No cabe duda de que la problemática de escoger un tamaño muestral adecuado de acuerdo para el propósito con el que se desea realizar el análisis, ha sido, es y será siempre un aspecto muy controvertido. Todo ello cobra una mayor importancia al tener

²⁰ Tamaños muestrales o número de simulaciones para la estimación de proporciones, valores medios y sumas para la precisión y nivel de confianza determinados. Nota: se ha empleado la definición de e como la precisión esperada a un nivel de confianza $1-\alpha$. Si la varianza σ^2 se estima mediante S^2 , se deberán emplear los cuantiles de la $t_{\alpha/2, (n-1)}$ en lugar de $z_{\alpha/2}$.

que reducir al máximo los costes derivados de la toma de datos cuando ésta se produce sobre el terreno (en el caso de una muestra de puntos sobre una fuente de mayor exactitud que el producto).

3.5 El ajuste mínimo cuadrático como herramienta para la corrección de errores sistemáticos

Dado que todo el estudio de errores presentado por el doctorando se basa en un análisis final del tipo “caja negra”, en este caso se pretende abordar el problema de igual forma. Así, las preguntas a plantear para cuantificar errores sistemáticos podrían ser:

- ¿Cómo se puede obtener una información más exhaustiva sobre un error sistemático detectado previamente?
- ¿Qué magnitud tiene dicho error?
- ¿Qué tipo de sistematismo es?
- ¿Se trata de una traslación, un factor de escala, un giro o una composición de todos o parte de ellos?

Cartográficamente, un error posicional sistemático podría ser una traslación (en cualquier componente X, Y, Z), un factor de escala (en X, Y y/o Z), o un giro (en X, Y y/o Z), con una combinación total o parcial de ellos. Si se trabaja únicamente sobre la componente planimétrica, el tipo de sistematismo se reduce a los mencionados en la tabla 35.1.

Tabla 35.1: Posibles errores sistemáticos en planimetría.

Sistematismo	Magnitud del error
Traslación en X	X_0
Traslación en Y	Y_0
Factor de escala en X	λ_x
Factor de escala en Y	λ_y
Giro	α

De esta forma, se podrían llegar a detectar todos estos errores mediante un ajuste mínimo cuadrático de una transformación afín en dos dimensiones. Dicha técnica es muy empleada en procesos geodésicos y fotogramétricos, pero no así en el control de calidad.

La transformación de puntos de un sistema de coordenadas a otro es un problema común en la ingeniería cartográfica. Por ejemplo, si inicialmente se realiza un proyecto en un sistema local de coordenadas, es posible que necesite transferir dicho proyecto a un sistema nacional de coordenadas. En las mediciones GPS y en el campo de la fotogrametría, las transformaciones de coordenadas se emplean exhaustivamente. Desde el comienzo del *North American Datum* de 1983 (NAD 83), numerosos ingenieros, agencias estatales, de transportes, y de todo tipo, han tenido que realizar transformaciones de coordenadas entre el antiguo datum (NAD 27) y el nuevo, desarrollándose varios modelos matemáticos para realizar este tipo de conversión (Wolf y Ghilani, 1997). No obstante, los ajustes mínimo cuadráticos en la transformación de coordenadas provienen de años anteriores (Rainsford, 1957).

3.5.1 Transformación afín 2D

La transformación de coordenadas bidimensional conforme, también conocida como transformación afín 2D, tiene la característica de que las formas se mantienen después de la transformación. Generalmente es empleada en ingeniería para la conversión de dos mediciones separadas a un único sistema de referencia. Esta transformación conlleva un proceso en tres pasos:

- 1º. Escalado: para crear dos sistemas de coordenadas con iguales dimensiones
- 2º. Rotación: para conseguir que los ejes de abcisas y ordenadas de ambos sistemas sean paralelos.
- 3º. Traslación: para obtener un mismo origen entre los dos sistemas de coordenadas.

Si bien el escalado y la rotación se pueden definir por dos únicos parámetros (Wolf y Ghilani, 1997), podría darse el caso de que el eje X se encuentre afectado por un factor de escala diferente al eje Y (error en el apoyo, digitalización, impresión...), por lo que para la escala se emplearán dos coeficientes diferentes. En cuanto a la traslación, se emplearán otros dos parámetros diferentes. De esta forma, se trabajará con un total de cinco parámetros (tabla 35.1)

Para poder resolver el sistema, serán necesarios un mínimo de 3 puntos de control con sus coordenadas planimétricas conocidas en ambos sistemas. Así, se dispondrá de 6 ecuaciones con 5 incógnitas que permitirán la aplicación de los mínimos cuadrados. Lógicamente, cuanto mayor sea el número de puntos de control, mayor redundancia tendrá el ajuste.

En la figura 35.1 aparece representado un punto (P) con sus coordenadas en ambos sistemas: coordenadas BDG (X, Y) y coordenadas terreno o FME (X', Y').

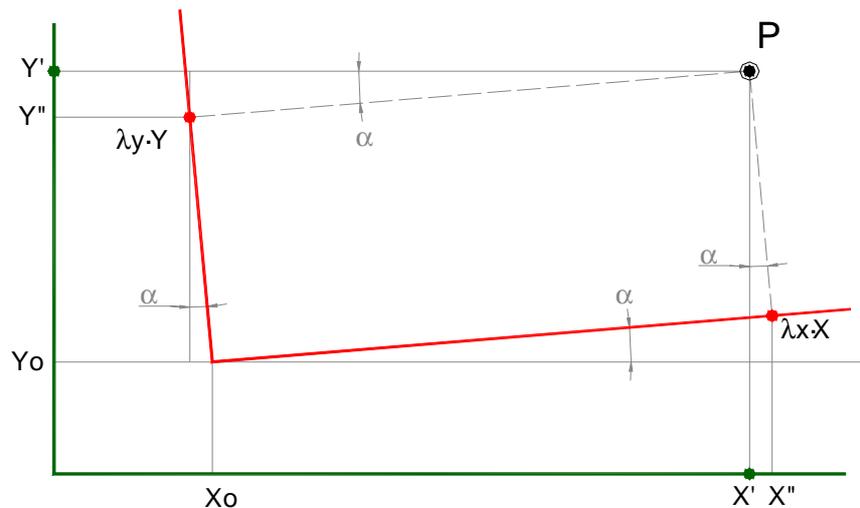


Figura 35.1: Representación de un punto (P) en ambos sistemas de coordenadas²¹.

Sobre las coordenadas de la BDG que definen la posición del punto P , deberán aplicarse las siguientes transformaciones o correcciones para referirlas al sistema de coordenadas terreno (FME):

- 2 factores de escala (λ_x, λ_y). Para hacer que las longitudes de los ejes en ambos sistemas de coordenadas sean coincidentes, será necesario multiplicar cada una de las coordenadas por dicho factor de escala:

$$\begin{aligned} X &= \lambda_x \cdot X_{mapa} \\ Y &= \lambda_y \cdot Y_{mapa} \end{aligned} \quad (22)$$

- 1 giro (α). El ángulo entre el eje de coordenadas X de la BDG y el del terreno, que ha de girarse para alcanzar el paralelismo entre ambos sistemas.

²¹ X e Y de la BDG en color rojo, y X' e Y' del terreno o FME en color verde.

²² Para simplificar la terminología, en la figura 35.1 aparece únicamente $\lambda_x \cdot X$.

- 2 traslaciones (X_0 , Y_0). Para hacer coincidir el origen de coordenadas en ambos sistemas.

Para cada una de las componentes, el cálculo de las ecuaciones que definen la transformación será el que aparece en la tabla 35.2. No se ha tenido en consideración la inclusión de pesos porque el error a cuantificar en cualquiera de las componentes tiene una misma importancia y, por tanto, ponderación.

Tabla 35.2 Cálculo de los parámetros de la transformación afín entre el sistema de coordenadas BDG y terreno (FME).

Componente X	Componente Y
$X' = X_0 + \overline{X_0 X''} - \overline{X' X''}$ $\text{Cos} \alpha = \frac{\overline{X_0 X''}}{X \cdot \lambda x} \rightarrow \overline{X_0 X''} = X \cdot \lambda x \cdot \text{Cos} \alpha$ $\text{Sen} \alpha = \frac{\overline{X' X''}}{Y \cdot \lambda y} \rightarrow \overline{X' X''} = Y \cdot \lambda y \cdot \text{Sen} \alpha$ $X' = X_0 + X \cdot \lambda x \cdot \text{Cos} \alpha - Y \cdot \lambda y \cdot \text{Sen} \alpha$	$Y' = Y_0 + \overline{Y_0 Y''} - \overline{Y' Y''}$ $\text{Cos} \alpha = \frac{\overline{Y_0 Y''}}{Y \cdot \lambda y} \rightarrow \overline{Y_0 Y''} = Y \cdot \lambda y \cdot \text{Cos} \alpha$ $\text{Sen} \alpha = \frac{\overline{Y' Y''}}{X \cdot \lambda x} \rightarrow \overline{Y' Y''} = X \cdot \lambda x \cdot \text{Sen} \alpha$ $Y' = Y_0 + X \cdot \lambda x \cdot \text{Sen} \alpha + Y \cdot \lambda y \cdot \text{Cos} \alpha$
Cambio de variable:	
$X' = a \cdot X + b \cdot Y + X_0$	$Y' = c \cdot X + d \cdot Y + Y_0$
Coeficientes:	
$a = \lambda x \cdot \text{Cos} \alpha ; \quad b = -\lambda y \cdot \text{Sen} \alpha$ $c = \lambda x \cdot \text{Sen} \alpha ; \quad d = \lambda y \cdot \text{Cos} \alpha$	
Parámetros de la transformación:	
$\lambda x = \sqrt{a^2 + c^2} \quad ; \quad \lambda y = \sqrt{b^2 + d^2} \quad ; \quad \alpha = \text{Arctg} \left(\frac{c}{a} \right) = \text{Arctg} \left(\frac{-b}{d} \right)$	

De esta forma, quedarán definidos los 5 parámetros de transformación en función de los coeficientes del ajuste. Para poder resolverlos, se deberá plantear el sistema de, al menos, 6 ecuaciones y 5 incógnitas que se podrá resolver empleando un ajuste por mínimos cuadrados.

Para obtener los valores de los coeficientes, se tienen que determinar los coeficientes a , b , c , d , X_0 e Y_0 de forma que V será el residuo del ajuste:

$$a \cdot X_1 + b \cdot Y_1 + X_{01} = X'_1 + Vx_1$$

$$c \cdot Y_1 + d \cdot Y_1 + Y_{01} = Y'_1 + Vy_1$$

$$a \cdot X_2 + b \cdot Y_2 + X_{02} = X'_2 + Vx_2$$

$$c \cdot Y_2 + d \cdot Y_2 + Y_{02} = Y'_2 + Vy_2$$

$$a \cdot X_3 + b \cdot Y_3 + X_{03} = X'_3 + Vx_3$$

... ..

Expresándolo de forma matricial, quedaría:

$$A \cdot X = T + V$$

en donde:

- A : es la matriz de las observaciones, de dimensiones $[2 \cdot n \times 6]$ (siendo n el número de puntos empleados en el ajuste)
- X : es la matriz de incógnitas, de dimensión $[6 \times 1]$
- T : es la matriz del término independiente, con los datos sobre la fuente de mayor exactitud, y de dimensión $[2 \cdot n \times 1]$
- V : es la matriz de residuos del ajuste, de dimensión $[2 \cdot n \times 1]$

En el ajuste por mínimos cuadrados, se pretende que la suma de los cuadrados de los residuos sea mínima:

$$M = \sum_{i=1}^n (V_i)^2$$

siendo la resolución del sistema:

$$A \cdot X = T + V$$

$$X = (A^t \cdot A)^{-1} \cdot A^t \cdot T$$

$$V = A \cdot X - T$$

y la bondad del ajuste vendrá dada por:

$$\sigma_0^2 = \frac{V^t \cdot V}{2 \cdot n - 6} \quad ; \quad \sum \hat{x}\hat{x} = \sigma_0^2 \cdot (A^t \cdot A)^{-1} \quad ; \quad \sigma = \sqrt{\sum \hat{x}\hat{x}}$$

Una vez calculados los coeficientes del ajuste y verificada la bondad del mismo, se podrán obtener las coordenadas en uno u otro sistema de forma indiferente. En la tabla 35.3 se muestran las ecuaciones para referir las coordenadas del terreno a la BDG.

De esta forma, si se pudiese detectar y cuantificar un error sistemático con la suficiente precisión, se podría tener un conocimiento del mismo mucho más

exhaustivo que con las técnicas empleadas actualmente. Asimismo, al conocer la magnitud de dicho error, podría permitir al usuario efectuar la corrección de dicho sistematismo mejorando la bondad del producto a un coste relativamente bajo.

Tabla 35.3 Cálculo de las ecuaciones para referir el sistema de coordenadas terreno (FME) al sistema de la BDG.

Componente X	Componente Y
$\begin{cases} X' = a \cdot X + b \cdot Y + X_0 \\ Y' = c \cdot X + d \cdot Y + Y_0 \end{cases} \Rightarrow \begin{cases} Y = \frac{X'}{b} - \frac{a \cdot X}{b} - \frac{X_0}{b} \\ Y = \frac{Y'}{d} - \frac{c \cdot X}{d} - \frac{Y_0}{d} \end{cases}$	$\begin{cases} X' = a \cdot X + b \cdot Y + X_0 \\ Y' = c \cdot X + d \cdot Y + Y_0 \end{cases} \Rightarrow \begin{cases} X = \frac{X'}{a} - \frac{b \cdot Y}{a} - \frac{X_0}{a} \\ X = \frac{Y'}{c} - \frac{d \cdot Y}{c} - \frac{Y_0}{c} \end{cases}$
$d \cdot X' - d \cdot a \cdot X - d \cdot X_0 = b \cdot Y' - b \cdot c \cdot X - b \cdot Y_0$	$c \cdot X' - c \cdot b \cdot Y - c \cdot X_0 = a \cdot Y' - a \cdot d \cdot Y - a \cdot Y_0$
$X \cdot (b \cdot c - a \cdot d) = b \cdot Y' - d \cdot X' + d \cdot X_0 - b \cdot Y_0$	$Y \cdot (b \cdot c - a \cdot d) = -a \cdot Y' + a \cdot Y_0 + c \cdot X' - c \cdot X_0$
$X = \frac{b \cdot (Y' - Y_0) + d \cdot (X_0 - X')}{(b \cdot c - a \cdot d)}$	$Y = \frac{a \cdot (Y_0 - Y') + c \cdot (X' - X_0)}{(b \cdot c - a \cdot d)}$

3.5.2 Consideraciones sobre el ajuste por mínimos cuadrados.

Este tipo de técnicas, tal y como se ha mencionado anteriormente, han sido profusamente empleadas en otras ramas de la geomática (geodesia, fotogrametría, topografía...), pero en ningún caso se han encontrado referencias al control de calidad cartográfico, y tampoco a la detección y eliminación de errores sistemáticos sobre BDGs.

No obstante, si bien a priori se presenta como una herramienta adecuada para detectar y corregir sistematismos, también es cierto que, dadas las características del dato geográfico, éste viene afectado de una variabilidad propia de la escala a la que se representa (incertidumbre posicional). Ésta afecta directamente a la desviación típica del elemento geográfico, por lo que la diferencia existente entre la posición real y la posición en la que se representa tendrá dos claras componentes²³:

- Componente sistemática, que podrá detectarse y corregirse mediante una transformación afín ajustada por mínimos cuadrados.

²³ Se supone que, previamente, se han eliminado o ponderado aquellos valores considerados como atípicos.

- Componente aleatoria, intrínseca del propio dato geográfico en función de la escala de representación.

Esta doble vertiente de la incertidumbre, hace que el ajuste por mínimos cuadrados se “contamine” por la componente aleatoria del dato. Así, se absorberá parte del error aleatorio, incrementando los valores residuales del ajuste. Esta razón obliga a tratar de diferenciar ambos tipos de error y determinar entre qué parámetros se puede aplicar la técnica.

3.6 Empleo de curvas operativas en procesos de control de calidad

Tal y como se emplean los estándares de control de calidad posicional, los resultados finales se fundamentan en una muestra de puntos de control. Así, una BDG (o lote de BDGs) podrá ser aceptada o rechazada mediante la inspección de todos los elementos que la componen (inspección 100 %), o mediante la inspección de una muestra o porción de ella. No cabe duda que los aspectos económicos son la principal ventaja de un muestreo en comparación con la inspección 100 %, dado que el coste de dicha inspección se incrementa notablemente (figuras 31.2 y 31.3).

En este apartado se presentan una serie de conceptos básicos necesarios para poder utilizar y comprender el funcionamiento de las curvas de aceptación y los muestreos por variables.

El muestreo tiene ciertas ventajas e inconvenientes, tal y como se muestra en la tabla 36.1

Tabla 36.1: Ventajas e inconvenientes de las inspecciones por muestras o fracción de lotes.

Ventajas	Inconvenientes
<ul style="list-style-type: none"> ▪ Reducción de costes por una inspección de sólo parte del producto ▪ Menores daños de manipulación durante la inspección ▪ Menos inspectores, simplificando los problemas de contratación y formación ▪ Mejora del trabajo de inspección, pasando de las monótonas decisiones pieza a pieza a las decisiones lote por lote. ▪ Aplicabilidad al ensayo destructivo, con un cuantificable nivel de aseguramiento de la calidad del lote. ▪ Rechazos de lotes enteros y no simples retornos de los defectuosos, lo que da una mayor motivación para la mejora. 	<ul style="list-style-type: none"> ▪ Existe el riesgo de aceptar lotes «malos» y rechazar lotes «buenos». ▪ Hay que agregar planificación y documentación. ▪ La muestra, generalmente, proporciona menos información acerca del producto que la inspección 100 %.

Mediante una inspección muestral, se trata de: proteger al productor de rechazos en lotes cuando los productos se han obtenido bajo control, y son satisfactorios; proteger al consumidor de la aceptación de lotes malos; proteger al consumidor durante largo tiempo; animar al productor a mantener el proceso bajo control; minimizar los costes de muestreo, inspección y administración; y, por último, proporcionar información acerca de la calidad del producto (Cowden, 1957).

Cualquier aplicación del muestreo de aceptación, debe distinguir si el objetivo es reunir información sobre los productos analizados o sobre el proceso que ha producido el lote. Así, deberán distinguirse dos tipos de muestreo:

Tipo A. Muestreo para aceptar o rechazar el lote de producto.

Tipo B. Muestreo para determinar si el proceso que ha elaborado el producto está dentro de los límites aceptables.

Otro aspecto de gran importancia en el muestreo de aceptación, es el tipo de datos a analizar:

- **Atributos:** obteniendo información sobre si son correctos o no. Éstos pueden ser:

Defectuosos: habitualmente medidos en porcentaje de defectuosos. Indica la aceptabilidad de unidades de producto para una amplia gama de características.

Defectos: habitualmente medidos por recuento o como ratio de defectos por unidad. Indica el número de defectos hallados en las unidades inspeccionadas que, por tanto, puede ser mayor que el número de éstas.

- **Variables:** obteniendo información sobre la medición. Habitualmente medidas mediante la media y la desviación típica. Indica la distribución de una característica medible del producto inspeccionado.

De esta forma, los planes de muestreo pueden clasificarse en dos tipos: planes por atributos y planes por variables.

Planes por atributos. En estos planes, se toma una muestra del lote y se clasifica cada unidad como *conforme / no conforme*. El número de disconformidades es comparado con el número de aceptación establecido en el plan, y se toma la decisión de aceptar o rechazar el lote. A su vez, los planes por atributos pueden clasificarse en:

1. Planes que, admitiendo determinados riesgos de muestreo, proporcionan protección lote por lote. Estos riesgos son:
 - a. En cada lote, un determinado nivel de calidad (en términos de porcentaje defectuoso) tiene un cierto riesgo (p.e. 10%) de ser aceptado por el comprador. Ese nivel de calidad es conocido como porcentaje de unidades defectuosas toleradas en el lote (p_2); el riesgo elegido es conocido como riesgo del comprador (β).
 - b. El plan de muestreo aceptará un cierto porcentaje (p.e. 95%) de los lotes inspeccionados que tengan un determinado nivel de calidad. Este nivel de calidad es llamado *Nivel de Calidad Aceptable* (NCA). El riesgo de rechazar un lote de NCA (p_1) es conocido como riesgo del productor (α).
2. Planes que proporcionan un limitado porcentaje medio de artículos defectuosos, a lo largo de la fabricación. Este valor se conoce como *Limite de Calidad Media de Salida* (LCMS).

Planes por variables. En estos planes, se toma una muestra y se mide una determinada característica de su calidad. Estas mediciones se resumen en un estadístico sencillo (por ejemplo, la media de la muestra) y los valores observados se comparan con un valor predeterminado que se define en el plan; y a continuación, se toma la decisión de aceptar o rechazar el lote.

Los planes por atributos se aplican, generalmente, en función del porcentaje defectuoso (para controlar el porcentaje de producto defectuoso). Los planes por variables tienen mayor sensibilidad, pero es necesario que la forma de la distribución de las medidas individuales sea conocida y estable. Ésta se utiliza para “traducir” el porcentaje defectuoso a valores específicos de los parámetros del proceso que son controlados (media, desviación típica). Los planes por variables también se utilizan para estudiar el proceso de producción cuando las especificaciones se refieren a su media o variabilidad, y no al porcentaje defectuoso.

Por último, en cuanto al **control de aceptación y muestreo de aceptación**, se denomina muestreo de aceptación a la aplicación de un plan de muestreo específico a un determinado lote o secuencia de lotes. Sin embargo, los procedimientos del muestreo de aceptación pueden ser utilizados en un programa de control de aceptación para alcanzar una mejor calidad a menor coste, mejorar el control e incrementar la productividad. De esta manera, los planes de muestreo pueden utilizarse para complementar cualquier otro plan utilizado en un programa continuo de control de aceptación para la mejora de la calidad con una inspección reducida. Este es el objetivo del control de aceptación que, posteriormente, podría ser reemplazado por una certificación del proveedor o un control de proceso.

3.6.1 Curvas operativas

Para poder emplear curvas operativas, es necesario conocer el significado del **riesgo** que corren tanto el productor como el usuario al utilizar muestreos para conocer la calidad de los productos. Cuando se lleva a cabo un muestreo de aceptación, las partes de mayor interés son el vendedor (suministrador, o departamento de Producción) y el comprador, es decir, el cliente o empresa que adquiere el producto y que ha de utilizarlo. Ya que el muestreo corre el riesgo de rechazar lotes «buenos» y

de aceptar lotes «malos», con las correspondientes graves consecuencias, vendedores y compradores han tenido que estandarizar los conceptos de lo que constituyen buenos y malos lotes, y estandarizar también los riesgos asociados al muestreo. Estos riesgos dependen de varios parámetros (tabla 36.2).

Tabla 36.2: Riesgo del vendedor y riesgo del comprador (Juran y Gryna, 1997).

Riesgo del vendedor.	Riesgo del comprador.
<p>El riesgo del vendedor (riesgo α) es la probabilidad de que un “buen” lote sea rechazado por el plan de muestreo. En algunos planes, este riesgo se fija en 0,05; en otros, varía del 0,01 al 0,10. El riesgo se establece en conjunción con la definición numérica de un máximo nivel de calidad que puede alcanzarse con el plan, habitualmente llamado <i>Nivel de Calidad Aceptable (NCA)</i>.</p> <p>Así, el NCA es el máximo porcentaje defectuoso (o máxima cantidad de defectos por cada cien unidades) que, a los fines de la inspección, puede considerarse satisfactorio como media del proceso. Un plan de muestreo debe tener un bajo riesgo del vendedor para una calidad igual o mejor que el NCA.</p>	<p>El riesgo del comprador (riesgo β) es la probabilidad de que un lote “malo” sea aceptado por el plan de muestreo. El riesgo se establece en conjunción con una definición numérica de la calidad rechazable, tal como el <i>Porcentaje de unidades Defectuosas Toleradas en el Lote (PDTL)</i>. El PDTL es el nivel de calidad que es insatisfactorio y que debe ser rechazado por el plan de muestreo. Comúnmente se asume un riesgo del comprador del 0,10 y el PDTL ha sido definido como la calidad del lote cuya probabilidad de aceptación es del 0,10; es decir, sólo el 10 % de tales lotes será aceptado. El PDTL es un caso especial del concepto de <i>calidad límite (CL)</i> o de nivel de calidad rechazable. Estos últimos términos son utilizados en tablas que proporcionan planes para varios valores del riesgo del comprador, en contraste con el valor del 0,10 para el PDTL.</p>

Curva característica. La curva característica (curva CC) es el gráfico que, para un determinado plan de muestreo, relaciona la probabilidad de aceptación con el porcentaje de defectuoso del lote (Ariza, 2002).

La figura 36.1a muestra una curva CC para el caso en que se desee aceptar todos los lotes hasta con un 3 % de defectuosos, y rechazar todos los lotes que tengan un nivel de calidad mayor del % defectuoso (Juran y Gryna, 1997). Todos los lotes con menos del 3 % defectuoso tienen una probabilidad de aceptación del 1,0 (certeza de aceptación), mientras que todos los lotes de más del 3 % defectuoso tienen una probabilidad de aceptación igual a cero. Sin embargo, no existe ningún plan de muestreo que pueda hacer esta discriminación perfectamente, dado que siempre hay algún riesgo de que algún lote “bueno” sea rechazado o de que algún lote “malo” sea aceptado. Por ello, se han de controlar estos riesgos.

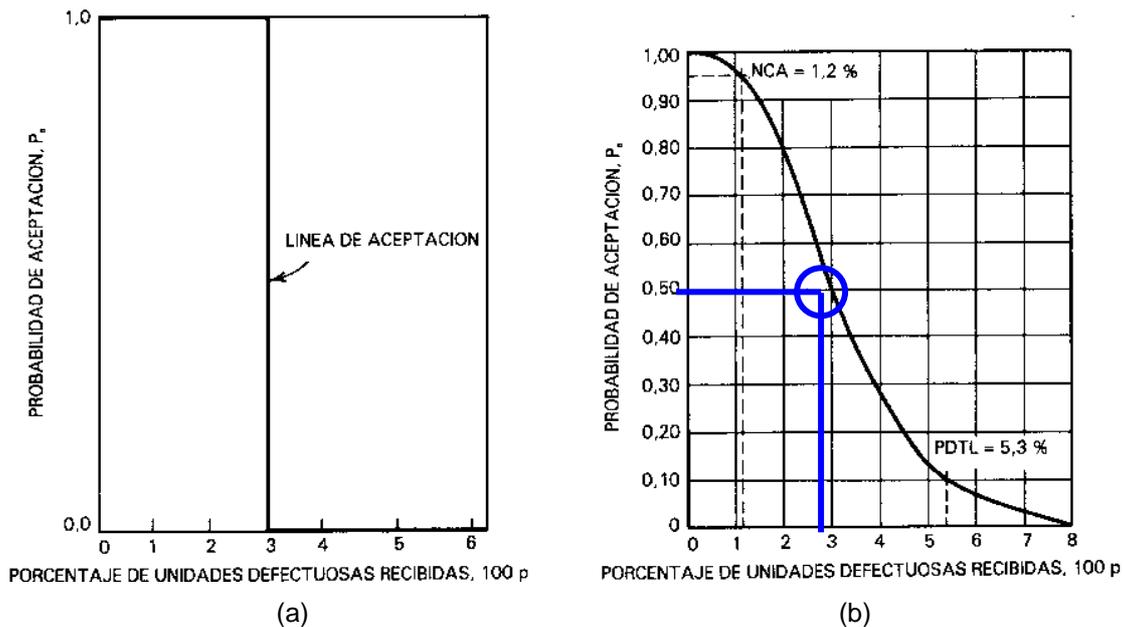


Figura 36.1: a) Funcionamiento de un plan de muestreo ideal; b) Funcionamiento de un plan de muestreo real.

La figura 36.1b muestra la curva de comportamiento que se obtendría si se le dijera a un inspector que tomara una muestra de 150 piezas como parte de un lote, y lo aceptara si no encuentra más de cuatro piezas defectuosas. Estas curvas se pueden construir a partir de distribuciones de probabilidad.

En la curva se puede apreciar que un lote 3 % defectuoso tiene una probabilidad de cada dos de ser aceptado. Sin embargo, un lote 3,5 % defectuoso, técnicamente un lote "malo" tiene 39 % de probabilidad de ser aceptado. De igual manera, un lote 2,5 % defectuoso, técnicamente un lote "bueno", tiene un 34 % de ser rechazado.

Límite de calidad media de salida. El nivel de calidad aceptable (NCA) y el porcentaje defectuoso tolerado en el lote (PDTL) son dos índices de calidad habituales en los planes de muestreo. Un tercer índice es el "Límite de Calidad Media de Salida" (LCMS) que corresponde al límite superior de la calidad media del producto de salida, incluidos los lotes aceptados y los rechazados que han sido cribados y mejorados.

3.6.2 Muestreo por variables

Al utilizar los planes por variables, se toma una muestra y se mide una determinada característica de su calidad en cada unidad. Los planes por variables proporcionan el mismo grado de protección al comprador que los planes por atributos utilizando muestras más pequeñas.

Cuando el interés se centra en conocer la proporción del producto que no cumple con las especificaciones de la medida, se pueden usar los planes de variables para determinar el porcentaje de unidades defectuosas. Estos planes relacionan la proporción de unidades defectuosas con la media de la población mediante la teoría de probabilidades. La media muestral se utiliza para probar la situación de la media en la población.

Suponiendo que se conoce que la distribución de las mediciones individuales es normal y que se desea un plan tal que la curva CC pase por los puntos $(p_1, 1-\alpha)$ y (p_2, β) donde:

- p_1 : nivel de calidad aceptable
- $1-\alpha$: probabilidad de aceptación al nivel p_1
- p_2 : nivel de calidad rechazable
- β : probabilidad de aceptación al nivel p_2

La curva CC debe aparecer tal como se indica en la figura 36.2.

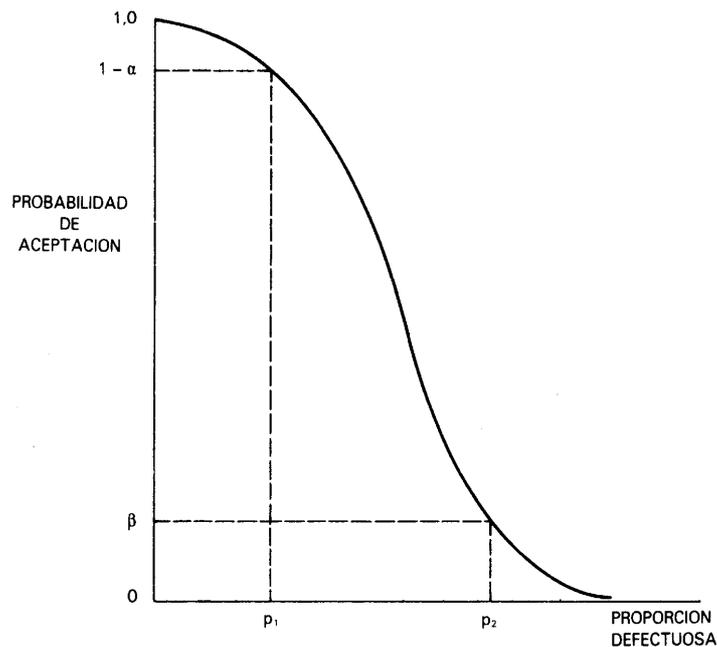


Figura 36.2: Curva característica $(p_1, 1-\alpha)$; (p_2, β)

La teoría de probabilidades puede ser útil para obtener las curvas CC y otras propiedades de esos procedimientos, pudiendo deducirse fórmulas que se ajusten a determinadas condiciones (Juran y Gryna, 1997). Las curvas CC de los planes por variables se consideran, por lo general, del tipo B, es decir, que se trata de un muestreo del proceso de fabricación de las piezas examinadas, más que del lote de material que se está analizando.

3.6.3 Consideraciones sobre el empleo de curvas operativas

No cabe duda que la determinación de curvas operativas supone un importante reto en el área de la cartografía. Así, en el caso del control de la exactitud posicional sobre una BDG, cabría mencionar que:

- Se trata de un muestreo por variables, en el que se analiza el error existente en alguna o varias de sus componentes (X, Y, Z) .
- Las muestras han de cumplir ciertos condicionantes previos (USBG, 1947, ASPRS, 1989, FGDC, 1998): un número de, al menos, 20 puntos de control, distribución homogénea, perfectamente identificables sobre la BDG y la FME, normalidad, aleatoriedad en la distribución de errores, ...

- Se han de establecer a priori, los riesgos de productor y usuario asumidos por ambos.
- Ha de definirse el estándar a aplicar.
- En función del estándar y los riesgos a asumir, se han de establecer las curvas operativas basadas en el tamaño muestral (número de puntos de control con el que realizar el control).

No obstante, se presentan numerosas dificultades que hacen realmente complicado obtener dichas curvas operativas:

- Elevado coste en la realización de los experimentos, si es que éstos se realizan sobre datos reales tomados en campo.
- Necesidad de realizar un estudio para observar el comportamiento de las BDGs en función de su escala de representación (simbología, generalización ...).
- Análisis por separado de la componente planimétrica (XY) de la altimétrica (Z), dadas las peculiaridades que las diferencian en cuanto a la metodología empleada tanto en su captura como en su representación.
- Necesidad de conocer el comportamiento poblacional (inspección 100 %) para poder interpretar el comportamiento muestral. Este aspecto se hace prácticamente imposible de determinar sobre datos reales.

Ante este reto, el empleo de técnicas basadas en la simulación ofrece al investigador una herramienta fundamental para la generación de posibles curvas operativas.

Capítulo 4

Material y método

4. Material y método

A continuación se presenta el material y la metodología empleada para el desarrollo de la investigación por parte del doctorando. El capítulo se ha distribuido en 3 grandes apartados:

- 1º. El proceso de simulación empleado. En él se exponen las principales características de la simulación aplicada al control de la exactitud posicional en cartografía. Asimismo, se presentan los pasos seguidos para la generación en laboratorio de poblaciones y la selección de muestras aleatorias de puntos de control, afectados tanto por los errores propios asociados a la cartografía analizada, como por la posible contaminación por otro tipo de errores. En último lugar, se hace una breve descripción de la aplicación propia para la realización de los procesos de simulación mencionados: Cartocal 2.0.
- 2º. Datos reales tomados sobre cartografía a diferente escala. Para la comprobación de los resultados obtenidos tras la simulación, se han tomado datos reales sobre la cartografía y en el campo, los cuales se muestran en este apartado.
- 3º. Aplicación al control posicional. Por último lugar, se expone la aplicación de los materiales mencionados a los estimadores robustos, los principales estándares de control posicional, las curvas operativas, el posible estudio geométrico, el ajuste mínimo cuadrático y los estimadores circulares.

4.1 El proceso de simulación empleado

Tal y como se ha indicado en el apartado 3.6, la simulación se presenta como una herramienta de gran valor para el investigador a la hora de tratar de buscar soluciones y respuestas a los problemas que se presentan en innumerables áreas de la ciencia. A continuación se exponen los procesos de simulación desarrollados.

4.1.1 Características de los estándares para el control posicional cartográfico

Los principales estándares presentados en el apartado 3.2 coinciden en los siguientes aspectos:

1. Toman una muestra de, al menos, 20 puntos de control.
2. La muestra ha de estar homogéneamente distribuida.
3. No ha de existir duda razonable en la identificación correcta de los puntos de control.
4. Los errores en sus componentes han de distribuirse:
 - a. De forma aleatoria.
 - b. De forma independiente
 - c. Asemejándose a una distribución estadística Normal
5. Una BDG correcta debería:
 - a. No tener errores groseros (valores atípicos)
 - b. No tener errores sistemáticos.
 - c. Los errores aleatorios deberían estar dentro de unos límites o tolerancias en función de las características establecidas a priori para la BDG.

Algunos estándares suponen como condición previa ciertas características que debieran poseer las BDG buenas: sin errores sistemáticos, normalidad, aleatoriedad... Por todo ello, sería necesario el comprobar previamente estos aspectos antes de aplicar el estándar en cuestión.

Así, el estándar puede ofrecer un resultado de distinto tipo:

- La aceptación o rechazo en función de unas tolerancias o límites de error máximo.
- La clasificación de la cartografía de acuerdo al estándar.
- La aportación al usuario del nivel de exactitud de la BDG.

En base a todo ello, los procesos de simulación se han orientado hacia la investigación sobre los principales aspectos que tienen en común los estándares.

4.1.2 Simulación en la generación de poblaciones

Como una de las principales características de los estándares es el hecho de que los errores se distribuyen de forma independiente, aleatoria y siguiendo una distribución estadística normal, el trabajo está fundamentado en poblaciones de errores generadas artificialmente de tal forma que cumplan dichas características. Para ello, se ha empleado el algoritmo de Box-Müller (ver apartado 3.4.1).

Una vez obtenidas las series de datos, se han tipificado de acuerdo con las necesidades del proceso de simulación a aplicar. Con la población de errores tipificados, a modo de datos supuestamente obtenidos empleando el total de puntos posibles en una hoja cartográfica, se ha trabajado aplicando tanto los diferentes estándares sobre el total de la población, como la simulación en la elección de muestras aleatorias (ver apartado 4.1.3).

En el diagrama de la figura 41.1 se muestra el esquema de trabajo en el proceso de simulación de poblaciones aleatorias.

En base a experimentos previos, el proceso de trabajo con poblaciones simuladas se ha repetido en 100 ocasiones para darle una mayor robustez y estabilidad a los resultados finales. Esta decisión se fundamenta en que la variabilidad obtenida entre dicho número de poblaciones es mínima y se ajusta a los requerimientos de precisión exigidos por el trabajo.

Respecto al tamaño poblacional, éste puede variar en función de las limitaciones estadísticas en cuanto a su relación con el tamaño muestral. No obstante, y dado que se podría considerar que una BDG está formada por infinitos puntos, se parte de que una población de 1000 puntos ($M = 1000$) tiene la suficiente significación estadística como para asemejarla al total de puntos que podrían definir un mapa (Rodríguez, 2003; Atkinson, García y Ariza. 2001).

Por todo ello, el proceso de simulación se fundamenta en considerar que un mapa está formado por 1000 puntos homogéneamente distribuidos, de tal forma que sus errores son independientes, aleatorios y se asemejen a una distribución estadística normal con una desviación típica preestablecida.

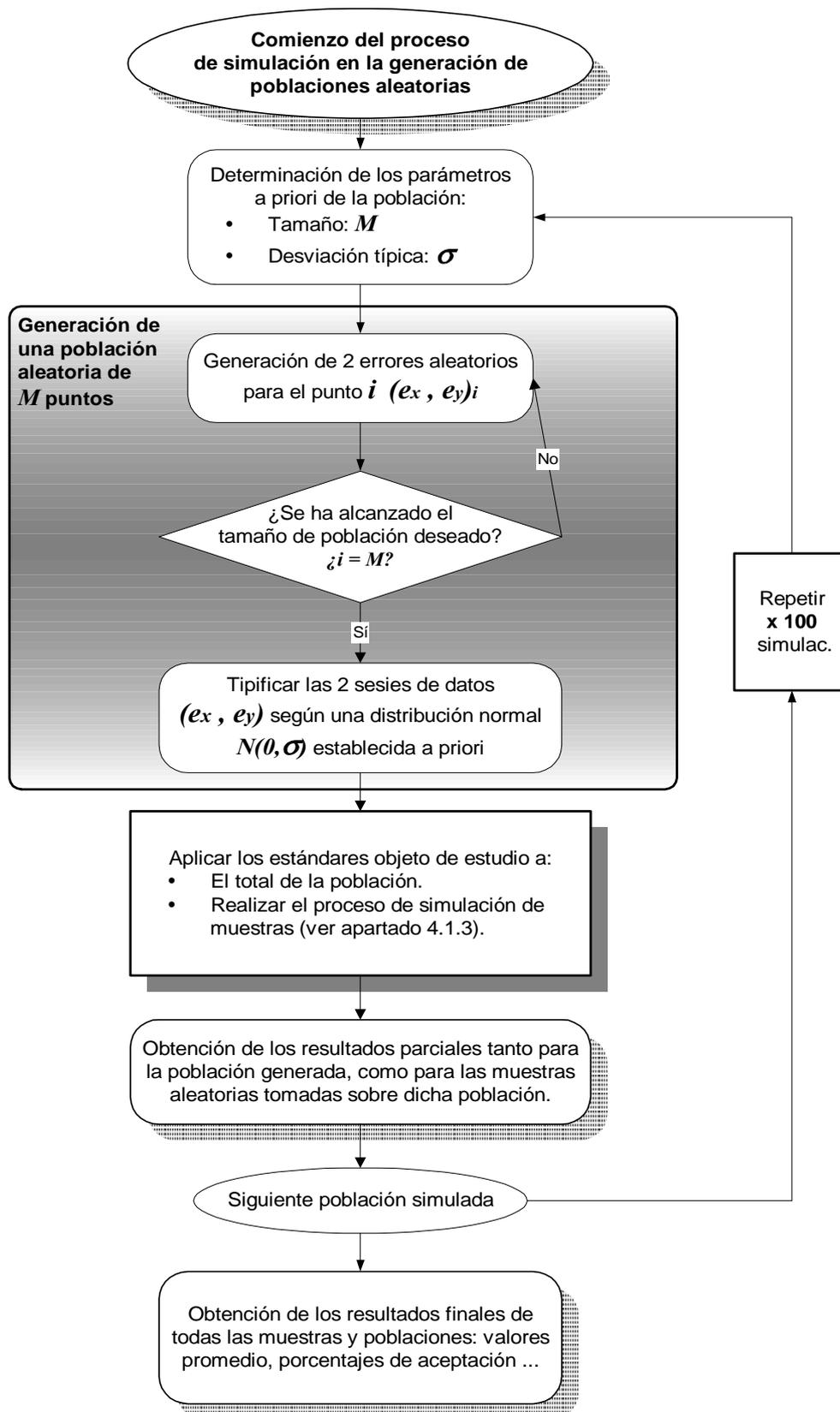


Figura 41.1: Diagrama de flujo seguido en el proceso de simulación para la generación de poblaciones aleatorias.

4.1.3 Simulación en la generación de muestras

Uno de los aspectos más interesantes de la investigación, se ha basado en el estudio del comportamiento de los diferentes estándares en función de las variaciones producidas sobre el tamaño muestral.

Para ello, se ha trabajado bajo las hipótesis previas que los estándares analizados reconocen sobre la distribución de los errores: independencia, aleatoriedad y normalidad. Dado que uno de los principales objetivos es el analizar el tamaño muestral, también se ha dado por supuesto la homogeneidad espacial y temática en la distribución de los puntos.

Si bien es cierto que sobre datos reales se pueden presentar comportamientos zonales diferenciados (p.e. sistematismos en zonas montañosas), éstos deberán detectarse por el estándar.

Una forma de simplificar el tratamiento de la información, ha sido emplear los valores de error en cada una de las componentes (e_x , e_y). De esta forma, obviando los valores en cuanto a la posición absoluta de los puntos, coordenadas X e Y , se presupone que los puntos seleccionados como muestra de control se distribuyen de acuerdo a las exigencias que los estándares indican en cuanto a su homogeneidad.

Sobre una determinada población de puntos P , con un tamaño poblacional M , se selecciona una muestra aleatoria de tamaño n . Repitiendo dicho proceso un determinado número de veces, pueden obtenerse resultados en laboratorio que serían muy difíciles y costosos de obtener en campo. Asimismo, si se varía el tamaño muestral, los resultados serían aplicables para el conocimiento del tamaño muestral adecuado al estándar objeto del análisis.

En cuanto al número de muestras aleatorias que pueden seleccionarse sobre una determinada población, dicho valor vendrá dado por las combinaciones lineales de M elementos tomados de n en n . Así, para el caso de una población de 100 puntos de control, y muestras de 30 puntos, se podrán escoger hasta $2,937 \cdot 10^{25}$ muestras diferentes:

$$\binom{M}{n} = \frac{P_{M,n}}{n!} = \frac{M!}{n!(M-n)!} \quad \text{en donde: } P_{M,n} = \frac{M!}{(M-n)!}$$

siendo: M : el tamaño de la población

n : el tamaño de la muestra

$$\binom{100}{30} = \frac{100!}{30!(100-30)!} = 2,937 \cdot 10^{25}$$

En el caso de la elección aleatoria de muestras, el procedimiento es aplicable tanto sobre poblaciones simuladas como sobre datos reales tomados en campo que posean una sobreabundancia tal que permitan el proceso de simulación.

No obstante, al seleccionar las muestras aleatorias cabría el plantearse la siguiente pregunta:

¿Podría repetirse la misma muestra de puntos de control?

La respuesta ha de ser afirmativa dado que dos operarios diferentes que analicen un mismo mapa podrían coincidir al escoger los mismos puntos de control puesto que éstos deben cumplir con unas características concretas (homogeneidad, fácil identificación, accesibilidad para la toma de datos en campo, ...). Así, a la hora de diseñar el algoritmo para la elección aleatoria de muestras, se ha contemplado la posibilidad de admitir dos o más muestras con los mismos puntos de control. El proceso básico aparece en la figura 41.2

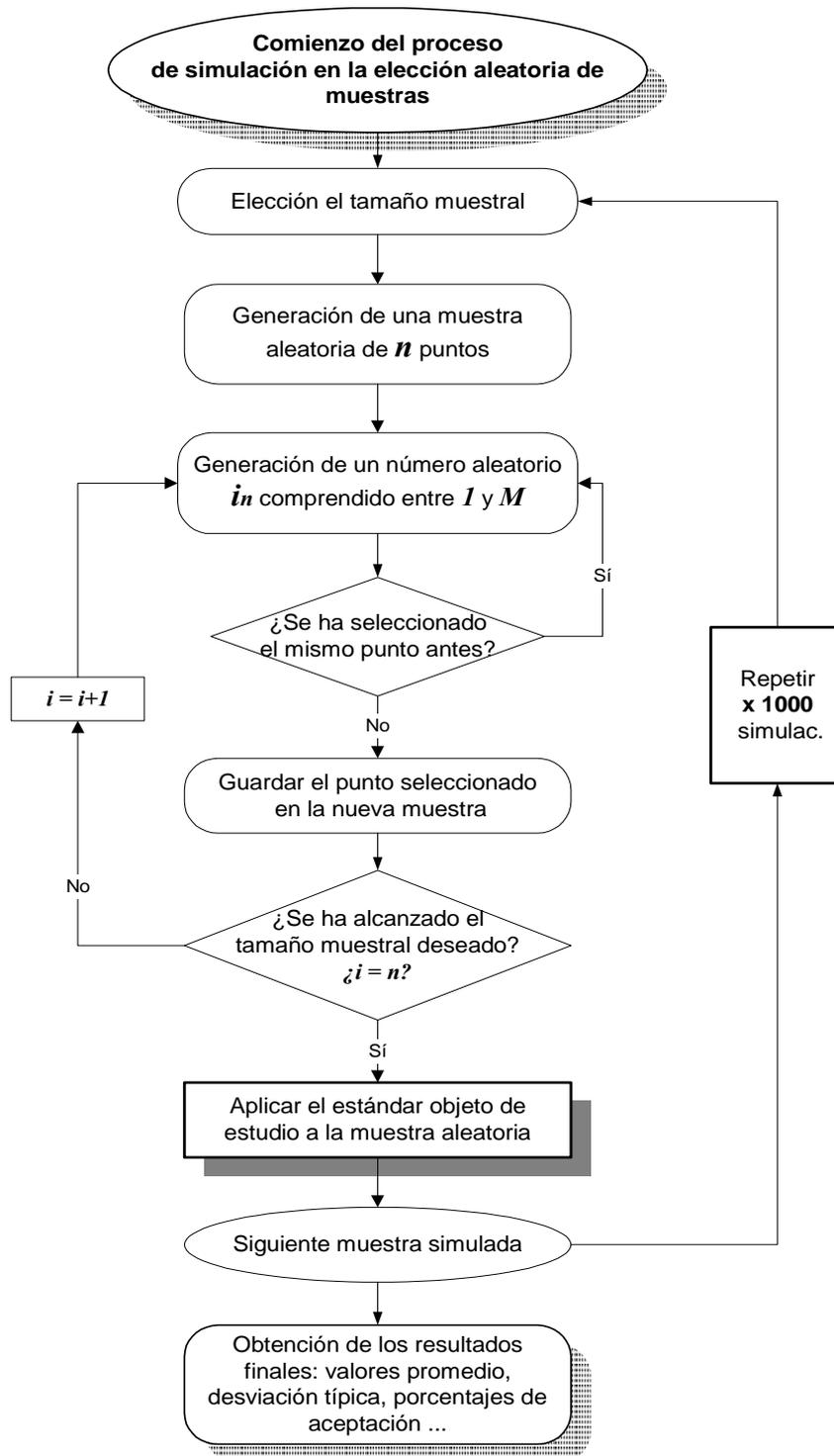


Figura 41.2: Diagrama de flujo seguido en el proceso de simulación para la elección de muestras aleatorias.

En este caso, en base a experimentos previos, se ha tomado la decisión de repetir la simulación en la elección de muestras aleatorias hasta 1000 veces para darle mayor robustez y estabilidad a los resultados. Esta decisión se fundamenta en la variabilidad obtenida entre las muestras.

Toda esta serie de procesos de simulación, tanto en poblaciones como en muestras, tienen como finalidad el tratar de generar posibles curvas operativas para cada uno de los estándares, de tal forma que puedan determinarse los riesgos reales de usuario y productor (ver apartado 3.6).

Al efectuar el control de calidad posicional de acuerdo con un estándar, la mayoría de ellos¹ determinará si la BDG analizada es adecuada (supera el estándar) o inadecuada de acuerdo a unas condiciones previas (escala, sistematismos, varianza muestral...). Así, el productor correrá un determinado riesgo (α) si el estándar rechaza una BDG que sí cumple con dichas condiciones. De igual forma, el usuario correrá un determinado riesgo (β) si el estándar acepta una BDG que no cumple con las condiciones exigidas (figura 41.3).

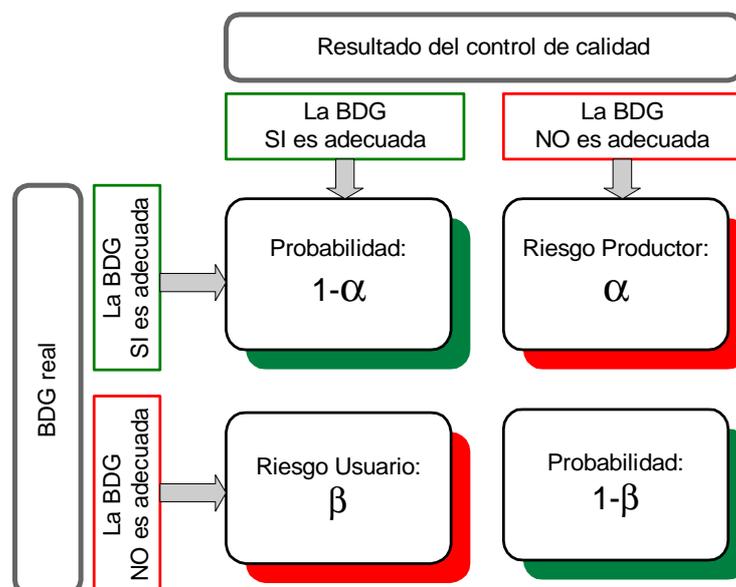


Figura 41.3: Niveles de riesgo genéricos para el productor y el usuario en el control de calidad posicional.

¹ Tan sólo el estándar NSSDA no proporciona al usuario un resultado de Aceptación / Rechazo de la BDG, aunque podría considerarse dicha opción si la *Exactitud_r* < Exactitud exigida a priori.

4.1.4 La simulación en la contaminación de poblaciones

Al analizar una muestra de puntos de control tomados sobre una determinada cartografía, básicamente ésta puede estar contaminada por dos tipos de errores:

- Errores groseros o valores atípicos (outliers).
- Errores sistemáticos.

A continuación se estudia cómo han sido tratados ambos tipos de errores en los procesos de simulación.

a) Contaminación de poblaciones con valores atípicos

Tal y como se ha mencionado en el apartado 3.3.1, los valores atípicos son considerados como outliers, procedentes de errores groseros o equivocaciones. Así en el caso de una población que se asemeje a una distribución estadística normal $N(0,1)$, podría estar contaminada en torno a un 10 % con valores atípicos. Dichos valores pueden no seguir ninguna distribución estadística conocida o seguir una distribución diferente: $N(0,3)$, $N(2,5)$, $U(2,1.5)$, $N(0,8)$...

Basándose en las hipótesis sobre las que se fundamentan los estándares en control de calidad posicional, la contaminación de las poblaciones se ha efectuado en diferentes porcentajes y con muestras normales de características distintas a la de la población principal.

Así, se ha generado aleatoriamente una muestra de K puntos con una distribución estadística $N_K(\mu_k, \sigma_k)$ diferente a la distribución estadística poblacional $N(0,\sigma)$, tal que:

$$N_K(\mu_k, \sigma_k) \neq N(0,\sigma)$$

siendo:

- μ_k : media de la muestra contaminada. En el caso de que coincida con la poblacional ($\mu_k = 0$), la desviación $\sigma_k \neq \sigma$
- σ_k : desviación de la muestra contaminada. En el caso de que coincida con la poblacional ($\sigma_k = \sigma$), la media $\mu_k \neq 0$
- $N(0,\sigma)$: distribución estadística de la población principal, según una normal de media 0 y desviación típica σ .

En la figura 41.4 se presenta, a modo de resumen, el proceso general de alteración de la población y tratamiento de los datos contaminados. Como se puede apreciar, la contaminación de la población se realiza sustituyendo un número determinado de puntos, en función del porcentaje de contaminación que se desee realizar, por nuevos datos provenientes de la muestra generada mediante una distribución estadística diferente (muestra contaminada). La inserción de estos datos sobre la población total se realiza de forma aleatoria tanto para la componente X como para la componente Y .

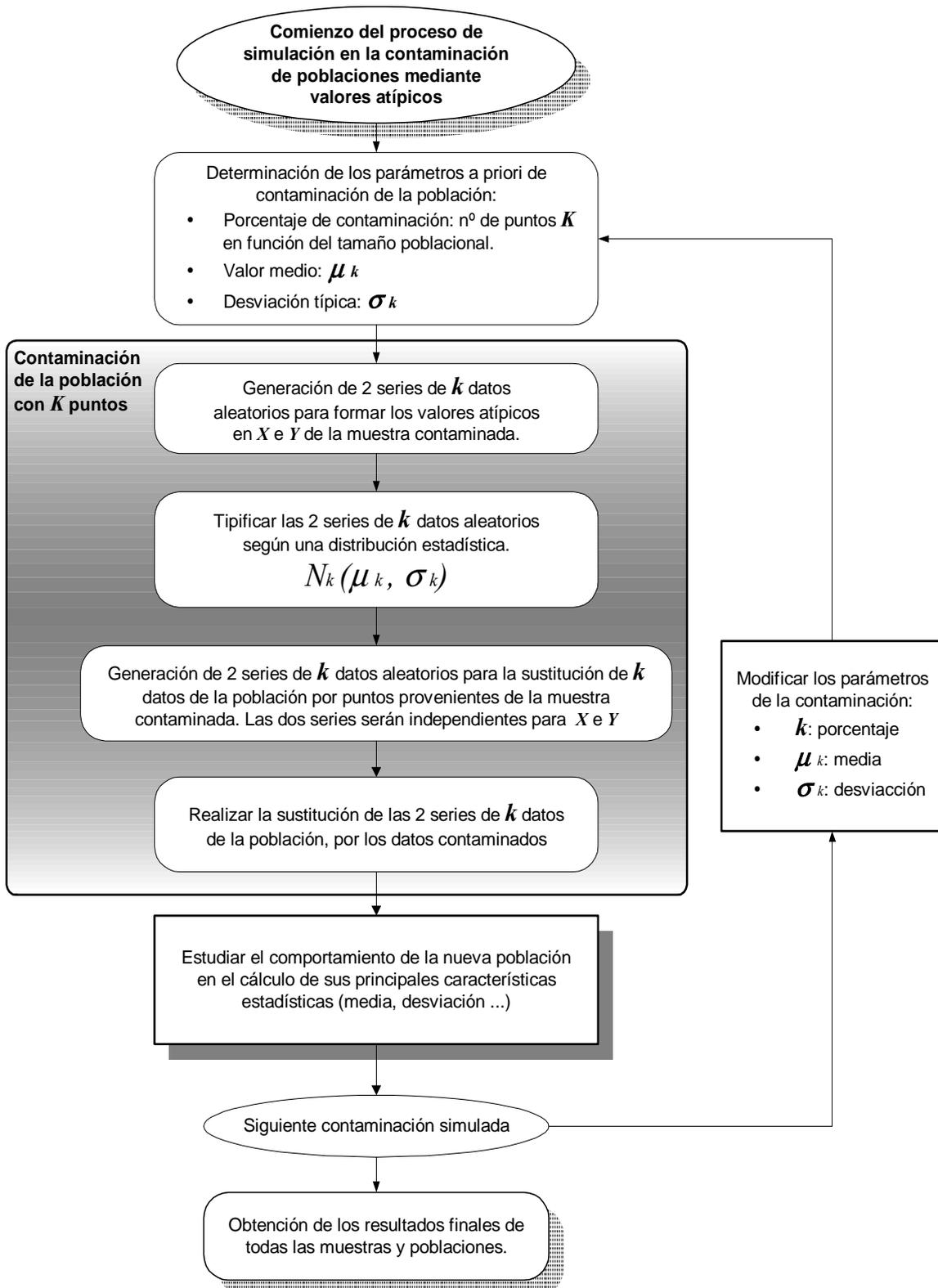


Figura 41.4: Diagrama de flujo seguido en el proceso de contaminación de poblaciones mediante valores atípicos con una distribución estadística diferente.

b) Contaminación de poblaciones con errores sistemáticos

Si bien se puede entender que la contaminación de la población con una muestra que se distribuye con un error medio $\mu_k \neq 0$ podría considerarse como un error sistemático en traslación, la investigación se ha desarrollado bajo la pauta de que un error sistemático es aquel que afecta a toda la población (de ahí su denominación: “*sistemático*”).

De entre los posibles errores sistemáticos (tabla 35.1), los más fáciles de detectar son aquellos que suponen algún tipo de traslación constante en cualquier eje (X, Y), observándose cómo el error medio se aleja del valor cero ($\mu \neq 0$). Por el contrario, los factores de escala y los giros son más difíciles de identificar por su aparente comportamiento no sistemático en función de dónde se sitúe el origen del factor de escala y del giro (que pueden no coincidir en el mismo punto si se dan ambas distorsiones a un mismo tiempo).

En este caso, se ha empleado un procedimiento de contaminación diferente al utilizado en los apartados anteriores. Así, sobre la población de datos generados aleatoriamente se ha contaminado sistemáticamente en función de unos parámetros de traslación, factor de escala y giro establecidos previamente por el investigador. De esta forma, conociendo los errores sistemáticos que afectan a la población, se han tratado de detectar y cuantificar empleando técnicas de ajuste por mínimos cuadrados. La contaminación sistemática se realiza aplicando una transformación afín 2D a la población original $N(0, \sigma)$.

Al igual que en los apartados anteriores, en la figura 41.5 se muestra el diagrama de flujo seguido en el proceso de simulación para la contaminación de poblaciones con errores sistemáticos.

Otro aspecto de gran relevancia es el tipo de población sobre la que se va a trabajar. Esto es, la población inicialmente “buena” (sin errores sistemáticos) se ajustará a una distribución estadística de datos normales $N(0, \sigma)$ y su desviación típica estará acorde con la desviación típica estimada a priori para el mapa (en función de su escala). Al trabajar con cartografía a diferente escala, dicha desviación variará, y este aspecto podría alterar los resultados obtenidos en la detección y cuantificación de los errores

sistemáticos. Por todo ello, durante el proceso de simulación, el investigador deberá trabajar con poblaciones simuladas de diferente desviación típica en función de la escala a la que se podría encontrar el mapa.

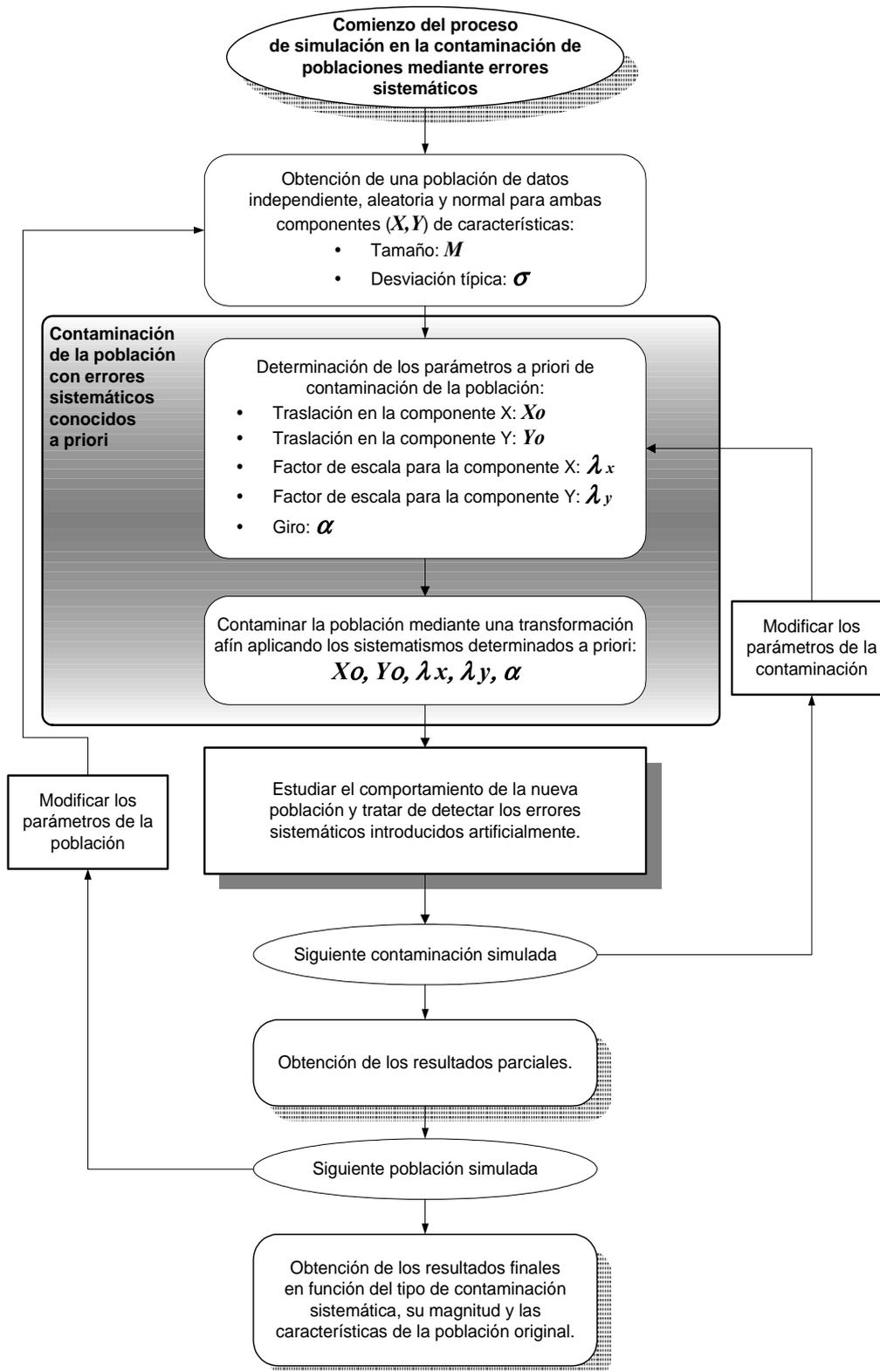


Figura 41.5: Diagrama de flujo seguido en el proceso de contaminación de poblaciones mediante errores sistemáticos.

4.1.5 Desarrollo de una aplicación propia para el estudio teórico: Cartocal 2.0

Para poder realizar la mayoría de los procesos de simulación descritos, se ha diseñado una aplicación propia, denominada Cartocal 2.0, bajo Microsoft Visual Basic 6. En el Anexo 5 se expone con mayor detalle el manual del usuario y del programador de la aplicación. No obstante, a modo de resumen, la tabla 41.1 presenta la estructura básica de menús.

Tabla 41.1: Estructura de menús de Cartocal. 2.0.

Menú principal	Submenú
Fichero	Abrir Guardar fichero de errores Salir
Tipo de mapa	Seleccionar tipo ...
Test	Parámetros estadísticos EMAS Test EMAS Test NMAS Test ASPRS Test NSSDA
Simulación	Simulación aleatoria de muestras Simulación aleatoria de errores
Afín	Transformación AFIN
Ayuda	Ayuda de CARTOCAL Acerca de ...

El programa se puede “alimentar” tanto de datos reales tomados sobre BDGs y fuentes de mayor exactitud, como de procesos simulación. Así, puede efectuar las siguientes operaciones:

- Intercambio de datos. Permite abrir, cargar datos y guardar resultados en archivos de formato estándar de intercambio ASCII.
- Aplicación de estándares. Calcula los estándares EMAS, NMAS, ASPRS y NSSDA bajo distintas posibilidades en cuanto a la definición personalizada de tolerancias, niveles de confianza, Bonferroni, ...
- Simulación. Se trata de una de las partes más significativas del programa. Permite aplicar uno o varios de los estándares mencionados sobre muestras de

distinto tamaño, seleccionadas aleatoriamente, de una única población, o repetir el proceso sobre diferentes poblaciones generadas aleatoriamente.

- Transformación afín. Aplica y muestra los resultados de una transformación afín ajustada por mínimos cuadrados para la detección y cuantificación de errores sistemáticos.
- Ayuda. Contiene una breve ayuda para la correcta utilización del programa (en formato PDF).

En cuanto a las principales funciones y algoritmos empleados, éstos quedan resumidos en la tabla 41.2.

Tabla 41.2: Principales algoritmos y funciones de cálculo en Cartocal. 2.0.

- **CalculaBarraProgreso:** calcula el número total de simulaciones que ha de realizar para dimensionar las barras de progreso y actualizarlas correctamente.
- **Calcular_Constante_NSSDA:** calcula el valor de la constante para el estándar NSSDA en función del nivel de confianza.
- **Desv_Media:** calcula la desviación media de la muestra
- **Error_Medio:** calcula el error medio de la muestra
- **ErroresAleatorios:** genera poblaciones de errores aleatorios distribuidos $N(0, \sigma_p)$
- **EscribirCabecera:** escribe la cabecera en el archivo ASCII de resultados en función de las opciones seleccionadas al verificar los estándares.
- **EscribirDatos:** escribe los resultados de las muestras y las poblaciones que han superado o no los diferentes estándares.
- **EscribirDesviaciónFINAL:** escribe en el informe compacto de resultados la desviación de todos los datos.
- **EscribirMediaFINAL:** escribe en el informe compacto de resultados la media de todos los datos.
- **GeneraFicheroErroresAleatorios:** genera una única población de errores aleatorios según una $N(0, \sigma_p)$
- **Interpola:** función auxiliar de interpolación para las tablas *t-Student* y *Chi-Cuadrado* cuando no se trabaja enlazado con Microsoft Excel ©
- **InterpolarEstadisticoC:** interpolación en la tabla *Chi-Cuadrado*
- **InterpolarEstadisticoS:** interpolación en la tabla *t-Student*
- **IteraciónDeMuestras:** selección de muestras aleatorias sobre la población
- **PuntoAleatorioDeMuestra:** selecciona un punto de forma aleatoria sobre la población para ser introducido en la muestra
- **Test:** función complementaria a la función "Interpola"
- **Test_C:** calcula el test para la detección de errores casuales del estándar EMAS
- **Test_S:** calcula el test para la detección de errores sistemáticos del estándar EMAS
- **TestASPRS:** calcula el estándar ARPRS con los puntos de la muestra
- **TestNMAS:** calcula el estándar NMAS con los puntos de la muestra
- **TestNSSDA:** calcula el estándar NSSDA con los puntos de la muestra

Por último, el programa ha sido verificado, incluyéndose en el código del programa los procesos de validación interna y externa de los algoritmos empleados y su programación en Visual Basic, mediante anotaciones similares a las que aparecen en la figura 41.6.

4.2.1 La cartografía analizada

Los trabajos de control de calidad posicional se han realizado sobre el Mapa Topográfico de Andalucía del ICA³ en formato analógico a escala 1/10.000.

El Mapa Topográfico de Andalucía (MTA) 1/10.000 ha sido realizado por la Junta de Andalucía, a través de la Consejería de Obras Públicas y Transportes, y actualizado por el Instituto de Cartografía de Andalucía. Se realiza por fotogrametría aérea mediante vuelos fotogramétricos a escala 1/25.000 y 1/20.000 para su actualización. El apoyo topográfico determina la posición planimétrica y altimétrica de un mínimo de cuatro puntos por par estereoscópico y la restitución de la minuta se realiza a escala 1/10.000. El sistema de referencia utiliza:

- Elipsoide Internacional de Hayford.
- Datum Potsdam.
- Orígenes de Coordenadas Geográficas referidas al Ecuador y al Meridiano de Greenwich.
- Origen de altitudes referidas al nivel medio del mar en Alicante.

El sistema geodésico es el definido por el Instituto Geográfico Nacional y el de Representación es la Proyección Universal Transversa de Mercator (UTM), resultando el territorio de Andalucía incluido entre los husos 29, 30 y 31.

La información, obtenida mediante restitución analógica, refleja los detalles en los fotogramas, recogida y completada en la revisión de campo, representándose en su posición exacta y forma real, con dimensión mínima a escala de un milímetro. El relieve se representa mediante curvas de nivel con equidistancia de 10 metros, curvas maestras cada 50 metros y puntos acotados de los detalles planimétricos más significativos (vértices geodésicos, señales de nivelación, vértices topográficos, puntos de apoyo, cumbres, collados, cruces de vías, estaciones de ferrocarril, puentes, cambios de pendiente, ...).

³ Instituto Cartográfico de Andalucía.

El Mapa Topográfico de Andalucía 1/10.000 contiene información planimétrica y altimétrica detallada de todas las vías de comunicación y elementos relacionados (autopistas, carreteras, caminos, sendas, ferrocarriles, líneas eléctricas de alta tensión, torres de tendido eléctrico, ...), hipsografía (sierras, cumbres, collados, barrancos, ...), todos los vértices geodésicos y topográficos, hidrografía (ríos, afluentes, arroyos, canales, acequias, ...), planimetría urbana (manzanas, calles, edificios aislados, parcelario urbano, espacios libres, identificación de edificios públicos, ...) y rústica (definición de parcelas, límites de cultivos, usos de suelo rústicos, ...), así como límites administrativos (nacional, provincial y municipal), recogiendo las últimas modificaciones debidas a segregaciones o anexiones de términos municipales.

En el aspecto toponímico se rotulan los accidentes principales del relieve, la hidrografía, todos los niveles de población, edificios habitados, construcciones aisladas de importancia, así como todos los edificios destinados a usos públicos, señalados en las normas oficiales. Los caminos y cañadas sólo cuando poseen nombres propios, y los vértices geodésicos y topográficos con el nombre que figuran en el Instituto Geográfico Nacional, los primeros, y con el nombre del paraje, los segundos. Para la recogida de estos topónimos se utiliza el Catastro Topográfico Parcelario y fuentes orales, hasta proporcionar la densidad correspondiente a este tipo de escala.

Se reproduce en soporte indeformable (formato DIN A-1), representándose en cada hoja una superficie aproximada de 3.400 hectáreas. Contiene, además del ámbito territorial en el que se dibujan las cuadrículas UTM cada un kilómetro, identificación numérica, términos municipales representados, leyenda con signos convencionales y usos de suelo, así como los datos técnicos y gráficos de situación. Al ser el 1/10.000 resultado de dividir el Mapa Topográfico Nacional 1/50.000 en dieciséis partes (4x4), para su identificación numérica se toma como referencia la numeración del 1/50.000, seguido de la columna y la fila en donde se encuentra la hoja correspondiente.

Cabe destacar que el propio ICA no posee toda la infraestructura necesaria para la realización de esta cartografía, teniendo que subcontratar algunas fases, como puede ser la adquisición de los vuelos al Instituto Cartográfico de Cataluña (ICC).

Asimismo, es importante mencionar que la restitución fotogramétrica ha sido analógica, con lo que los errores en ésta fase del trabajo son superiores a lo que serían con el empleo de otras técnicas (analítica o digital).

4.2.2 El sistema GPS

La tecnología del Sistema de Posicionamiento Global (GPS) está disponible desde hace más de una década (Ashkenazi y Moore, 1986; Alcalá, 1996). Aunque fue desarrollada por el Department of Defense de los Estados Unidos de América, como un sistema de navegación puramente militar, en la actualidad también se emplea en aplicaciones civiles. Estas incluyen: levantamientos hidrográficos, control y seguimiento de la dinámica de la corteza terrestre, levantamientos de triangulación geodésica, navegación marítima, aérea y terrestre, control y seguimiento de mareógrafos, instalación de plataformas marítimas, agricultura, topografía de ingeniería, Sistemas de Información Geográficos...

Existen dos tipos diferentes de observaciones que se pueden efectuar con GPS. El primero es el de las *pseudodistancias*, que se emplean para el posicionamiento instantáneo en navegación. Se puede efectuar con una precisión de unos 10 – 15 metros en posición horizontal, y de unos 30 metros en la componente vertical. No obstante, la mayor parte del error puede eliminarse con la utilización de dos receptores que se encuentren operando de forma simultánea en la zona de trabajo (GPS diferencial – DGPS).

Esta forma de operar (por pseudodistancias), se basa en el cálculo de la distancia entre el satélite y el receptor en función del tiempo que tarda la señal en viajar del satélite al receptor. Éste se compara con los dos osciladores y su código. El satélite genera un código que se recibe retrasado en el tiempo en el receptor. Éste lo duplica y correla, calculando la distancia absoluta. El cálculo de pseudodistancias supone el cálculo de la distancia geométrica más el error provocado por los osciladores, y multiplicado todo ello por la velocidad de propagación de la luz. Si los osciladores estuviesen perfectamente sincronizados, la intersección de las esferas que nos marcan las distancias centradas en cada satélite, nos darían el punto exacto en el que se encuentra el receptor. Al no suceder así, se hace necesario un cuarto satélite para

poder resolver el problema: 4 incógnitas $(X, Y, Z)_{\text{receptor}}$ y ξ_t . Dependiendo del tiempo de estacionamiento, el equipo receptor utilizado, la distancia entre los receptores, el número y la geometría de los satélites, las precisiones alcanzadas con dos receptores suelen ser inferiores al metro.

El segundo tipo de observaciones que se puede realizar con GPS es mediante la *medida de fase* de la portadora. En este caso se utiliza la portadora sin modular, sin medir el tiempo, y calculando el número de ciclos entre el satélite y el receptor, que multiplicado por la longitud de la onda proporciona la distancia. En este tipo de observaciones aparece un nuevo factor: la ambigüedad (el número entero de ciclos en el momento en que se comienza a medir, el cual se desconoce). Habrá una ambigüedad diferente para cada satélite que se observe. Dependiendo del tiempo de estacionamiento, el equipo receptor utilizado, la distancia entre los dos receptores, el número y geometría de los satélites, las precisiones alcanzadas con dos receptores suelen ser del orden de $5\text{mm} \pm 1\text{ppm}$ de la longitud de la línea base medida (vector formado entre los receptores).

En el caso de levantamientos de vías de comunicación, el modo de posicionamiento es el cinemático relativo por medida de fase (en aquellos casos donde sea posible la resolución de ambigüedades) y por pseudodistancias (cuando éstas no se puedan resolver).

Con las características del sistema, el trabajo en modo de código diferencial (o pseudodistancias), garantiza la toma del punto de una forma rápida y con errores inferiores al metro. Dicha precisión es adecuada para trabajos a escala 1/10.000.

4.2.3 Planificación y toma de datos

La primera tarea a realizar fue seleccionar las hojas a controlar. Tomando los trabajos realizados en 1998 (Atkinson, 1998), se han empleado las hojas 926 1-1 (en la que aparece el término municipal y el casco urbano de Mengibar) y 926 1-2 (en la que apenas hay carreteras, cruces de caminos y edificaciones) del MTA como ejemplo representativo de dos mapas: uno con una gran densidad de puntos fácilmente identificables y otro con pocos puntos identificables a priori (ver Anexo 8: Planos).

Para poder aplicar procesos de simulación en la elección de muestras sobre ambas hojas, se seleccionaron algo más de 100 puntos de control sobre cada una de ellas. Dichos puntos han sido escogidos de forma aleatoria y homogénea (ver Anexo 3, figuras A31.1 y A31.2). Al decir homogénea, se refiere tanto a su distribución espacial como al tipo de punto seleccionado (cruces de caminos, canales de riego, edificios aislados, postes, ...). Así mismo, se seleccionaron más puntos en aquellas zonas donde existía un mayor número de detalles planimétricos, y menor en las zonas menos densas⁴.

Una vez escogidos, se realizó una visita previa de campo a la zona de trabajo. Ésta consta de dos partes :

1. Ayuntamiento/s de la zona. Contactando con el personal cualificado de las respectivas Oficinas Técnicas para conocer, sobre los planos, las posibles modificaciones efectuadas desde la realización del vuelo (Septiembre de 1991) hasta la fecha de efectuar el control de calidad (Septiembre de 1997). Realizar un análisis tanto de las carreteras y caminos, como de las edificaciones, líneas eléctricas, canales, etc. Se desecharán los puntos previos seleccionados sobre las zonas modificadas. Por ejemplo, en el caso del presente trabajo, el Sr. D. Arturo de Urda (Aparejador en el Excmo. Ayuntamiento de Mengibar), fue quien primero advirtió de la autovía Bailén - Jaén que discurre de Norte a Sur atravesando las dos hojas por su zona más oriental. Así mismo advirtió de las nuevas edificaciones realizadas en el casco del municipio y las ya derruidas o reconstruidas.
2. Reconocimiento de la zona en campo. Consultando a los lugareños sobre las posibles modificaciones realizadas en edificaciones aisladas, caminos, sendas, etc. Verificación, por parte del técnico que realiza el control de calidad, de la actual existencia de los puntos preseleccionados, así como la inexistencia de posibles confusiones en su localización en campo. Así mismo, se tendrá en cuenta la accesibilidad a los diferentes puntos y zonas del mapa para ahorrar tiempo y optimizar costes en la posterior toma de datos.

⁴ Debido a esta mayor densidad, existen puntos que se encuentran más próximos, de tal forma que al efectuar los procesos de simulación únicamente se selecciona uno de estos puntos por muestra, sin que existan dos puntos muy próximos en la misma muestra.

Al tratarse de un trabajo sobre cartografía a escala 1/10.000, se diferenciaron distintos tipos de puntos basándose en la leyenda de signos convencionales del mapa. Tanto a la hora de seleccionarlos y capturarlos en la cartografía y el terreno, se establecieron diferentes procesos a seguir para:

- La elección de puntos de control (ver Anexo 2, apartado A2.1).
- El levantamiento de puntos de control (ver Anexo 2, apartado A2.2).
- La digitalización de puntos de control (ver Anexo 2, apartado A2.3).

En la fase de planificación, se establecieron los itinerarios a seguir en campo para optimizar los desplazamientos (ver Anexo 3, figuras A31.1 y A31.2). Al tratarse de dos hojas contiguas, la planificación de éstos se hizo de forma coordinada (hay itinerarios que transcurren por ambas hojas).

En cuanto a la campaña de campo y procesamiento de las observaciones GPS con los puntos de control, en la figura 42.1 se puede observar el diagrama de flujo seguido. En el Anexo 3 se muestran detalladamente los pasos seguidos y los cálculos realizados para obtener las coordenadas finales de los puntos de control sobre la fuente de mayor exactitud (terreno). De igual forma, en el Anexo 4 se adjunta el proceso de digitalización manual seguido para la obtención de dichos puntos sobre la cartografía analógica.

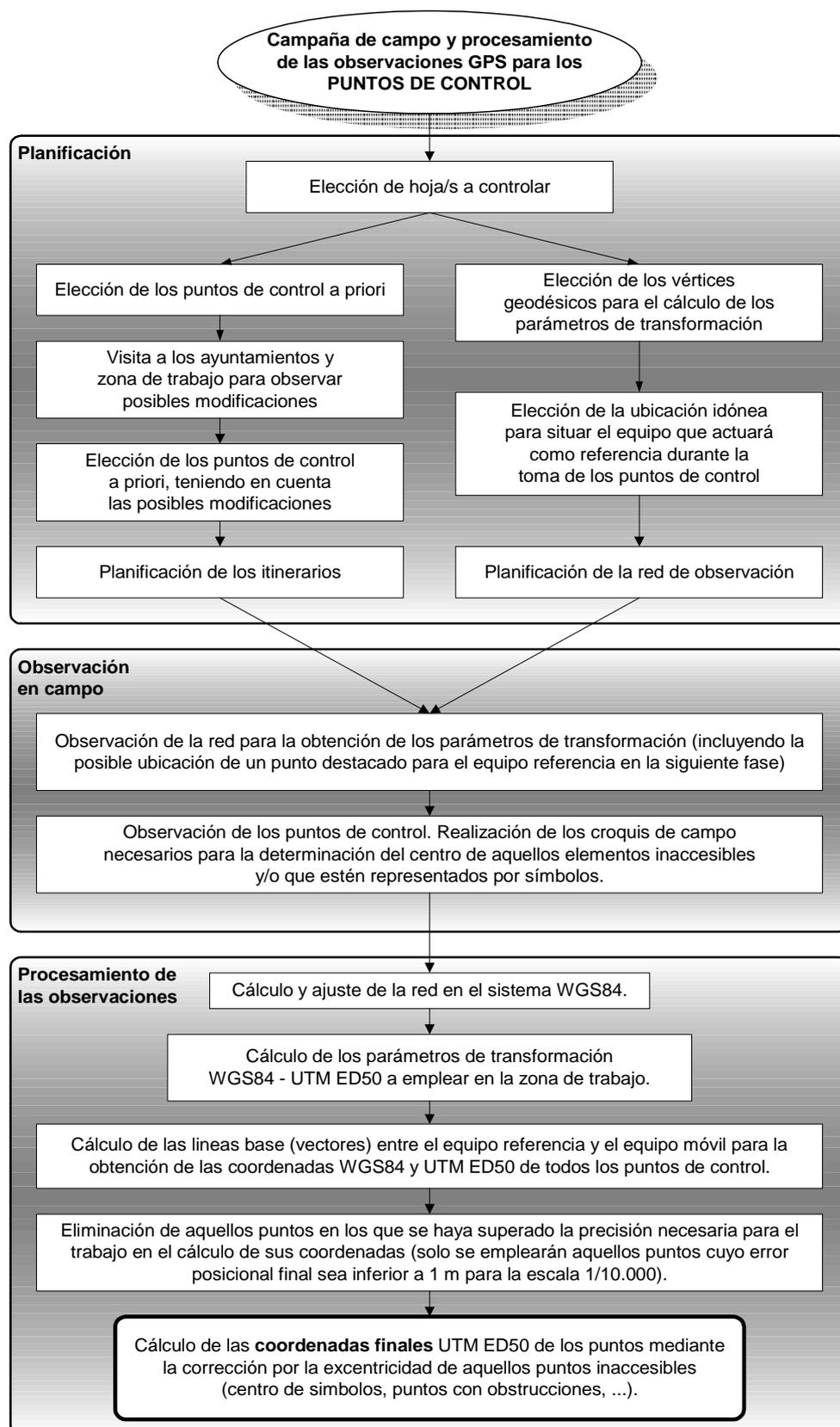


Figura 42.1: Diagrama de flujo seguido en la planificación, observación y cálculo de los datos GPS para la toma de puntos de control.

Por último, en la tabla 42.1 aparecen las precisiones obtenidas en cada una de las fases de observación y cálculo, así como las tolerancias máximas admisibles para la cartografía a escala 1/10.000.

Tabla 42.1: Errores en las diferentes fases de observación y cálculo, y tolerancias máximas admisibles para la cartografía analizada⁵.

Fase de observación – cálculo		Error máximo observado (en metros)
		1/10.000
F_1	Observación y ajuste de la red	0,0087
F_2	Parámetros de transformación	0,048
F_3	Observación y cálculo de los puntos de control	0,79
F_4	Observación y cálculo de puntos inaccesibles	0.25
F_5	Falta de homogeneidad y paralelismo en el levantamiento de los márgenes de las carreteras y errores por la interpolación de los ejes	-
Total (componente cuadrática):		0,83
Tolerancia máxima admisible en función de la escala		0,83
Ecuación:		$Tolerancia \geq \sqrt{F_1^2 + F_2^2 + F_3^2 + F_4^2 + F_5^2}$

4.3 Aplicación al control posicional

En este último apartado se presenta la aplicación tanto de los procesos de simulación como sobre datos reales, para:

- La detección y ponderación de valores atípicos.
- El estudio del comportamiento de los principales estándares de control de exactitud posicional.
- La obtención de curvas operativas.
- La corrección mediante ajuste por mínimos cuadrados de errores sistemáticos.
- La aplicación de estimadores circulares.

⁵ La tolerancia se ha obtenido como la tercera parte de la desviación típica estimada a priori para la cartografía en función de su escala $\sigma_{BDG}/3$ (ver capítulo 3, apartado 3.2).

4.3.1 Ajuste de los estimadores robustos

Trabajando bajo la hipótesis de que sólo una pequeña parte de los puntos de control pueden ser objeto de un error atípico, se ha realizado el estudio de los estimadores robustos MLT, gaussiano, método danés y Geman y Mc. Clure, frente a los estimadores clásicos. Dado que los principales estándares (v.g. EMAS) se fundamentan en estadísticos como la media y la desviación, sus resultados serán muy sensibles a cualquier variación que se pueda producir sobre éstas. En la medida en que sea posible minimizar los efectos de valores atípicos, los resultados serán más fiables.

Para realizar el experimento, se ha generado una población aleatoria de 100 puntos con errores en dos componentes (X e Y), tipificados según una $N(0,1)$ ⁶, otros tantos según una $N(0,2)$ y una $N(0,10)$. Seguidamente, se ha procedido a contaminar dicha población con otras, de menor número de elementos (5%, 10% y 20% del total de la muestra principal), distribuidas de diferente forma (v.g. $N(0, 1.1)$, $N(0, 1.5)$, $N(1.5, 1.5)$, $N(0, 6)$...). Partiendo de estos datos, a modo de supuestos puntos tomados para un control cartográfico, se ha realizado el estudio del comportamiento de media y desviación para aplicaciones iterativas del estimador MLT, y de forma comparativa con el resto de estimadores: gaussiano, método danés y Geman y Mc.Clure. A continuación se muestran los procesos realizados para lograr el ajuste de los estimadores en su aplicación al control de calidad posicional.

4.3.1.1 Proceso iterativo del estimador MLT

Tal y como se menciona en el capítulo 3 (ver apartado 3.3.1.1), el estimador MLT se puede aplicar de forma iterativa en la ponderación de la media y la desviación típica muestral (μ y S respectivamente) hasta 25 veces, o hasta que el estimador de Fisher varíe menos de 0.001 (Rodríguez, 2003). No obstante, esto no es factible sobre las muestras de puntos de control, dado que se alteran excesivamente los valores μ y se infravaloran los de σ .

⁶ Con la nomenclatura $N(0,1)$, se presenta una población con distribución Normal de media 0 ($\mu = 0$) y desviación 1 ($\sigma = 1$). De igual forma se presentan otras distribuciones normales con diferentes valores de media y desviación: $N(\mu, \sigma)$.

A modo de ejemplo, en la figura 43.1 se representa el valor del error (μ y σ) frente al número de iteraciones del estimador MLT. Como se puede apreciar, para una muestra $N(0,1)$ contaminada en un 10% con una distribución $N(1.5,1.5)$, se observa cómo los valores de μ se reducen ligeramente. No obstante, la mayor variación se produce en la σ para ambas componentes (figura 43.2): la reducción es de tal envergadura que, para una supuesta muestra original de $\sigma = 1$, en la iteración nº 12 ésta queda reducida a 0.76.

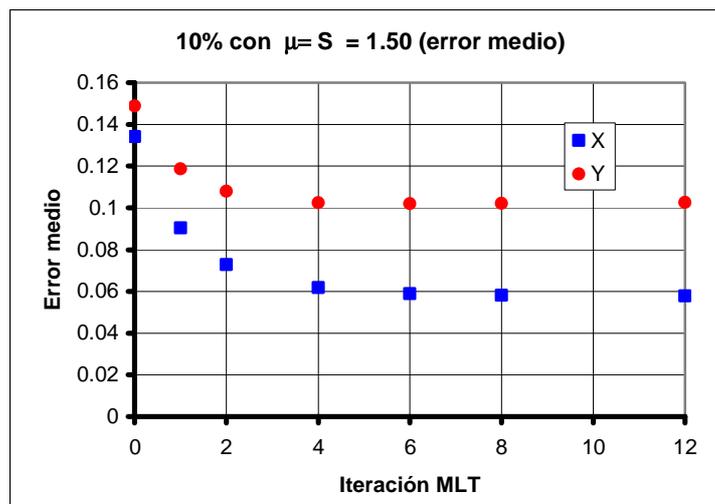


Figura 43.1: Valores del error medio resultante (μ) en función del número de iteraciones del estimador MLT.

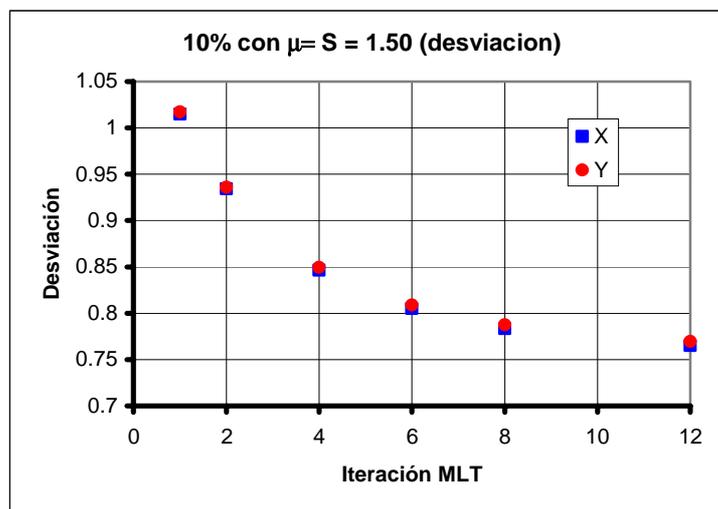


Figura 43.2: Valores de la desviación resultante (σ) en función del número de iteraciones del estimador MLT.

Sin embargo, los resultados en la primera iteración no son totalmente incorrectos tal y como podrá observarse más adelante.

4.3.1.2 Suavizado gaussiano

A la hora de aplicar el suavizado gaussiano, los valores finales varían sustancialmente en función de dónde se sitúe el grado de corte o aplicación. Así, deberá establecerse un criterio de aplicación de la ponderación del error por la función gaussiana, que podría ser:

- A partir de los atípicos medios.
- A partir de los atípicos extremos⁷.
- Aquellos que se encuentren a más de $3\cdot\sigma_p$; siendo σ_p la desviación típica estimada a priori para la población (1, 2 y 10 para cada una de las poblaciones empleadas en la simulación)
- Otros (v.g. a más de $2.5\cdot\sigma_p$)

Dado que la práctica totalidad de los estándares aplicados al control de calidad posicional cartográfico (NMAS, EMAS, ASPRS, NSSDA...) se basan en la hipótesis de la distribución normal de los errores, se ha realizado la simulación empleando dicha función sobre una población normal.

4.3.1.3 Método danés

En el caso de la aplicación del método danés, se ha realizado un estudio empírico basado en las muestras simuladas con diferente grado de contaminación, y de esta forma se ha ajustado el coeficiente constante k para optimizar los resultados finales.

A la hora de obtener dicho coeficiente, no solamente se ha trabajado con las muestras mencionadas al comienzo de este apartado, si no que se ha trabajado con muestras de diferente tamaño y distinto grado de contaminación, tanto en magnitud, como tipo (sobre la media y la desviación) y porcentaje. Así, partiendo de la ecuación general planteada en el capítulo 3 (ver apartado 3.3.1.3), se obtendría la ecuación ajustada para la reducción de la influencia de los valores atípicos en el control de la exactitud posicional:

⁷ Esta opción no se ha considerado en el estudio al ser relativamente pocos los valores considerados como atípicos extremos que intervienen en el cálculo.

$$W_i = e^{-0.05 \cdot e_{Xi}^2}$$

siendo:

- W_i : el peso a aplicar en el punto i
- e_{Xi} : el valor del error en la componente X para el punto i
- 0.05 : el valor constante (k) ajustado empíricamente

De esta forma, el error asociado al punto i en la componente X (de forma análoga se actuaría con la componente Y), vendría dado por:

$$e'_{Xi} = e_{Xi} \cdot W_i$$

En este caso, no se realiza ninguna iteración dado que los pesos propuestos modificarán sustancialmente el resultado final del error medio y la desviación típica de la muestra, minimizando los atípicos.

4.3.1.4 Método de Geman y Mc. Clure

De forma similar al punto anterior, Geman y Mc. Clure (Domingo, 2000) proponen una ponderación basada en una constante, a determinar empíricamente, y un valor inversamente proporcional al cuadrado del error para el dato a analizar. Al igual que sucedía con el método danés, se ha realizado un estudio exhaustivo para llegar a determinar empíricamente el valor de la citada constante. Así, la ecuación ajustada para la reducción de la influencia de los valores atípicos en el control de la exactitud posicional cartográfica quedará definida por:

$$W_i = 5 \cdot \frac{1}{1 + e_{Xi}^2}$$

siendo, al igual que en el caso anterior:

- W_i : el peso a aplicar en el punto i
- e_{Xi} : el valor del error en la componente X para el punto i
- 5 : un valor constante (k) ajustado empíricamente

El error asociado al punto i en la componente X (de forma análoga se actuaría con la componente Y), viene dado por:

$$e'_{Xi} = e_{Xi} \cdot W_i$$

Tampoco se realiza ninguna iteración dado que los pesos propuestos modificarán sustancialmente el resultado final del error medio y la desviación típica de la muestra, minimizando los atípicos.

Respecto al proceso de ajuste de los valores constantes, tanto en el método danés como en el de Geman y Mc. Clure, se ha pretendido en todo momento que los resultados en cuanto a la media y la desviación de la población contaminada se aproximasen lo más posible a los datos previos al proceso de contaminación. De esta forma, los valores contaminados quedarían minimizados frente a la población original pero sin tener que eliminar ningún punto de control.

4.3.2 Estándares en el control de calidad posicional: proceso comparativo mediante simulación.

Partiendo de las consideraciones preliminares enunciadas sobre los estándares (capítulo 3), se ha ideado un proceso de simulación iterativo (apartado 4.1) en el que el investigador ha de introducir una serie de parámetros preliminares y el software desarrollado realiza los procesos de simulación ofreciendo una serie de resultados que deberán ser analizados.

Así, el proceso general de simulación se podría resumir en varios bucles anidados iterativos:

- El primero de ellos genera poblaciones de errores que se distribuyen de forma aleatoria e independiente, ajustándose a una distribución Normal. Tanto el error medio como su desviación son tipificados y controlados por el investigador. Este bucle, al no centrarse exclusivamente en una única población aislada y obtenerse una serie de valores promediados, ofrecerá una idea de la robustez de la metodología en función de la variabilidad de los resultados finales.
- El segundo de ellos realiza una selección de muestras aleatorias sobre las poblaciones generadas en el bucle anterior. Esta selección es variable en función del rango de tamaño muestral sobre el que se desee trabajar. Este es

uno de los bucles más importantes dado que se desea conocer cómo afecta el tamaño muestral al resultado final ofrecido por el estándar.

- Por último lugar, sobre cada una de las muestras seleccionadas se realizan las pruebas estadísticas necesarias determinadas por los estándares (NMAS, EMAS, ASPRS y NSSDA).

Cabe destacar que cada uno de ellos posee unas características propias, por lo que el último de los bucles deberá ser adaptado a dichas peculiaridades con el fin de conocer sus puntos fuertes y débiles, así como poder efectuar el proceso de comparación.

Para poder trabajar con datos normalizados y extrapolables a otro tipo de cartografía, se han empleado poblaciones aleatorias distribuidas según una $N(0,1)$. Así, y de acuerdo con las recomendaciones internacionales de expertos y organizaciones cartográficas (USBB, 1947; ASCI, 1983; ASPRS, 1989, FGDC, 1998...), una cartografía con una desviación típica en sus errores posicionales de 1 m, sería aconsejable realizarla a una escala de 1/4.000 ($[0.2\text{mm} + 25\%] \times 4000 = 1 \text{ m}$). De esta forma, se podría considerar que las poblaciones generadas serían acordes con los parámetros de calidad exigibles para una cartografía a escala 1/4.000. Todo el proceso de simulación se presenta esquematizado en el diagrama de flujo de la figura 43.3.

a simulación consiste en generar, para cada componente horizontal de la BDG, poblaciones normales aleatorias de $N_{p,x}$ y $N_{p,y}$ con valores según el método de Box-Muller (Box, 1958) tal que se distribuyan según una Normal de parámetros teóricos $N(0,1)$. Asimismo, se verifica el cumplimiento de las hipótesis mediante test de adherencia a la normal y test de aleatoriedad.

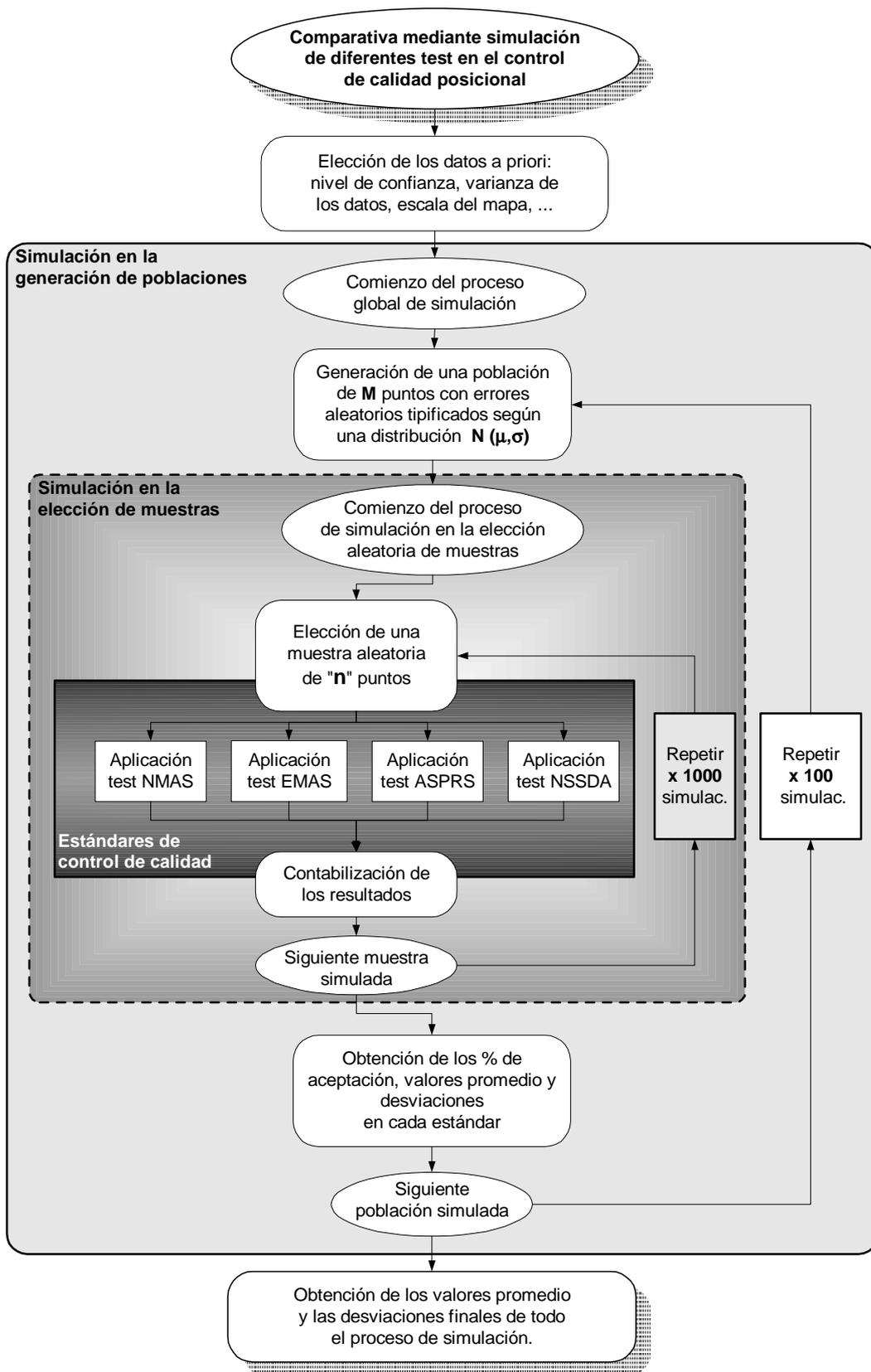


Figura 43.3: Diagrama de flujo del proceso general de simulación para comparar los diferentes estándares.

Estas dos poblaciones N_{px} y N_{py} , donde $N_{px} = N_{py} = 1000$, se conforman de esta manera como una base de datos de errores planimétricos simulados para las componentes X e Y de la BDG. Sobre ellas, aleatoriamente, se extraerán I diferentes muestras m_{nx} , m_{ny} , cada una de ellas con n número de elementos ($n = 5, 10, 15, 20, 25, 30, \dots$), empleándose como si fueran los errores encontrados en I diferentes muestras de 5, 10, 15, 20,.... puntos de control. Todas ellas se utilizarán para determinar nivel de aceptación o rechazo de la calidad posicional de una BDG.

Dados los tamaños de población y muestras utilizadas, el proceso de extracción aleatoria de las I muestras no llega a agotar las posibles combinaciones de elementos que quedan acotadas por los valores de C_n^{Np} , pero además, en el caso de que este valor quedara sobrepasado ($I > C_n^{Np}$), dado el mecanismo aleatorio de extracción, se generaría de manera automática un proceso de *bootstrapping*.

Otro factor importante a tener en cuenta, es el hecho de que cuando el tamaño muestral se aproxima al tamaño poblacional, los resultados pierden validez y ha de tenerse en cuenta el *Factor de Corrección de Poblaciones Finitas* (FCPF):

$$FCPF = \frac{M_{población} - n_{muestra}}{M_{población}}$$

en donde:

- $M_{población}$: es el tamaño de la población
- $n_{muestra}$: es el tamaño de la muestra

Cuanto más se aleje el FCPF de la unidad, menor fiabilidad tendrán los datos obtenidos. Por ejemplo, si se va a trabajar con poblaciones de 1000 puntos y muestras de 25, el FCPF = 0,975. Si se aumenta el tamaño de la muestra a 200, el FCPF = 0,8. Así, se ha trabajado con poblaciones de 1000 puntos para tamaños muestrales de hasta 125 puntos (FCPF \approx 0.88).

4.3.3 Aplicación de procesos de simulación para la obtención de curvas operativas.

Trabajando prácticamente con el mismo motor de simulación que el descrito en el diagrama de flujo de la figura 43.3, surge la siguiente idea:

¿Qué sucedería al aplicarse el estándar bajo condiciones de aceptación más/menos estrictas en cuanto a varianza estimada a priori para el mapa?

Si se considera, por ejemplo en el **EMAS**, un nivel de exigencia superior al que realmente tiene la población ($\sigma_p^2 > \sigma_{EMAS-Mapa}^2$)⁸. Para una población distribuida con una $\sigma_p = 1$ y el EMAS con un nivel de exigencia $\sigma_{EMAS-Mapa} = 0,8$, la probabilidad de aceptación disminuye adoptando comportamientos muy claros de rechazo. En este caso se han generado 100 poblaciones aleatorias de 1000 puntos cada una ($N_{p1} = N_{p2} = \dots = N_{p100} = 1000$), para ambas componentes (X e Y), con errores aleatorios según una distribución Normal tipificada $N(0;1)$. En cada una de estas poblaciones p_i , se ha desarrollado el anterior proceso de simulación con la elección de 1000 muestras aleatorias para diferentes tamaños muestrales ($n = 20, 30, 40 \dots$). Todo ello se ha repetido aplicando niveles de exigencia superiores al de la varianza poblacional ($\sigma_{EMAS-Mapa} = 0,95; 0,925; 0,95 \dots 0,70$). El proceso de simulación aparece en el diagrama de la figura 43.4.

⁸ La denominación $\sigma_{EMAS-Mapa}$ (o σ_{EMAS} , o σ_{Mapa}), se refiere a la desviación estimada a priori para la BDG. La σ_{EMAS} es la empleada en el estándar EMAS para el estudio de la varianza o prueba χ^2 (ver capítulo 3, apartado 3.2.2).

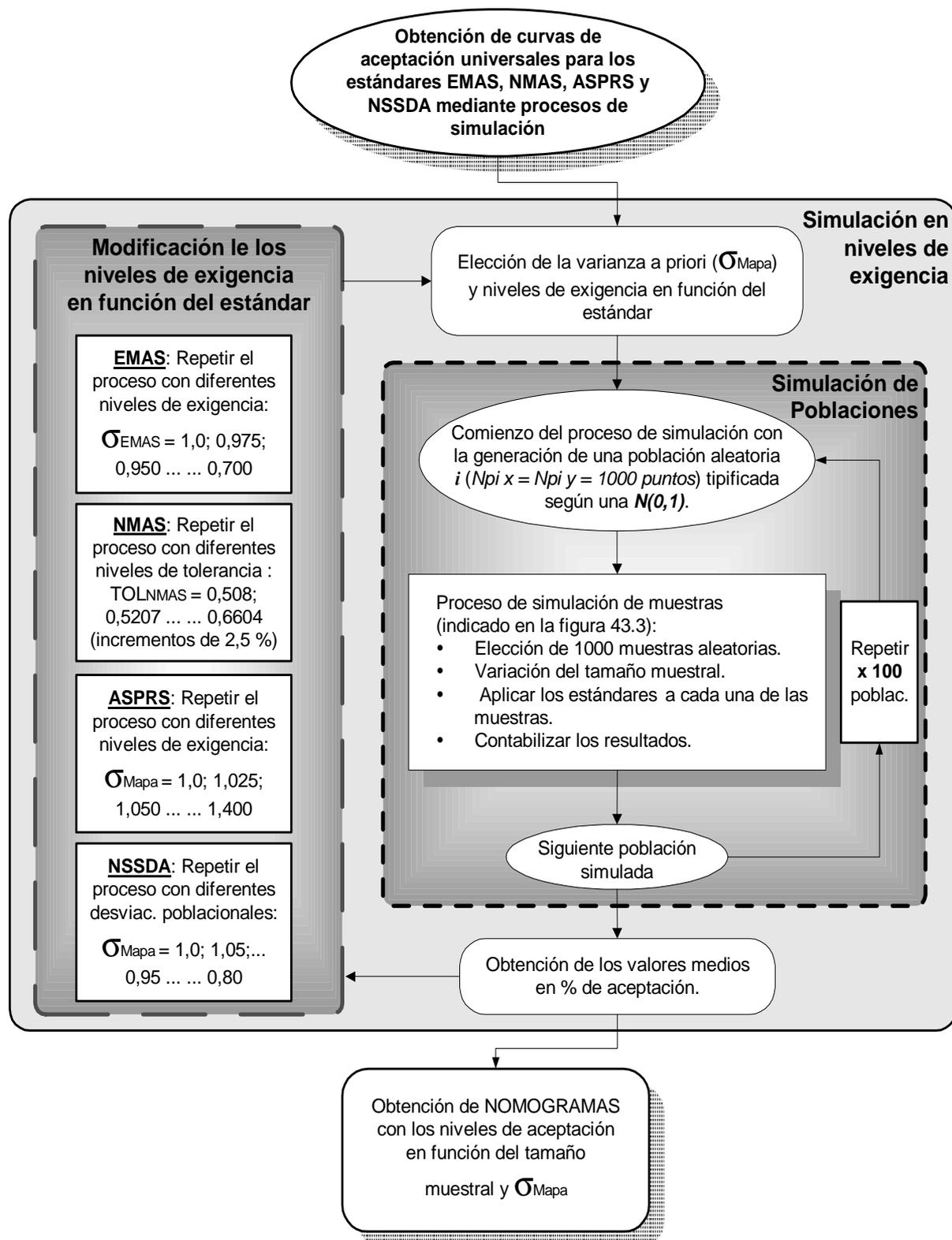


Figura 43.4: Diagrama de flujo del proceso de simulación para la obtención de curvas de aceptación en los estándares EMAS, NMAS y ASPRS.

De forma similar se podría trabajar con los estándares NMAS y ASPRS:

- **NMAS:** Los niveles de aceptación, en función del tamaño muestral, podrían variar sustancialmente dependiendo de la distribución que siga la población (σ_p) y la estimada a priori para la cartografía (σ_{Mapa}). La forma de simular esta situación es aplicando, en una población distribuida según una $N(0,1)$, distintos niveles de exigencia del estándar (TOL_{NMAS})⁹. El proceso de simulación sería el indicado en la figura 43.4.
- **ASPRS:** La aceptación o inclusión de una muestra en una determinada categoría del estándar, vendrá dada por la variabilidad de la población (σ_p) y la estimada a priori para la cartografía de Clase I (σ_{Mapa}). Si se varía el nivel de exigencia para dicha categoría o la desviación típica de la muestra, también cambiarán los niveles de aceptación. El proceso general de simulación es muy similar al anterior (figura 43.4).

Por el contrario, en el caso del estándar **NSSDA**, al no trabajar con el concepto de aceptación/rechazo si no con un valor discreto de exactitud posicional, el procedimiento no es aplicable en un primer momento. Tan sólo se podrá obtener la curva en la que se exprese el valor de dicha exactitud y su variabilidad en función del tamaño muestral. Cuanto más se aproxime dicho valor muestral al poblacional, mejor será el resultado. No obstante, y a pesar de todo ello, se ha estudiado el comportamiento del estándar en función de la σ_{Mapa} y el tamaño muestral empleando poblaciones con desviación inferior y superior a 1: $N(0, 0.8)$, $N(0, 0.85)$, $N(0, 0.9)$, $N(0, 0.95)$, $N(0, 1)$, $N(0, 1.05)$, $N(0, 1.1)$, $N(0, 1.15)$ y $N(0, 1.2)$. Asimismo, se han establecido posibles reglas de decisión para aceptar o rechazar una BDG en función de la *Exactitud*, obtenida por el estándar.

En todos los estándares se trata de determinar los riesgos tanto del usuario como del productor (figura 41.3) en función del tamaño muestral. De esta forma, a la hora de emplear alguno de los estándares analizados, ambos podrán conocer en qué márgenes de riesgo están trabajando.

⁹ El estándar original establece distintas tolerancias en función de la escala de la cartografía (mayor o menor que 1/20.000). La población se asemeja a la de una cartografía 1/4.000 ($\sigma_{Mapa} = 1$ m), y se han modificado los niveles de tolerancia del estándar: 0,508 mm (= 1/50"), 0,5207 (=1/50"+2.5%), 0,5334 (=1/50"+ 5%)... Éstas se basan en el nivel de tolerancia del NMAS para cartografía a escala <1/20.000, con incrementos de un 2,5 %: 1,000; 1,025; 1,050...

En el caso de los estimadores circulares (**Ombibus** y **Mardia**), el proceso de simulación es similar al descrito para los casos anteriores, salvo que se ha introducido artificialmente un determinado nivel de contaminación como error sistemático en una dirección concreta.

4.3.4 Cuantificación de errores sistemáticos mediante ajuste por mínimos cuadrados.

Cuando se detecta un error sistemático (tabla 35.1), cabe la posibilidad de tratar de identificarlo y cuantificarlo. Si bien, los errores producidos por traslación en ambos ejes son bastante fáciles de observar mediante una comprobación visual o un estudio de los errores medios en sus componentes, el trabajo se complica si se combinan otros sistematismos como variaciones de escala y giros.

En este apartado no se pretende detectar el sistematismo, sino averiguar qué tipo de error es (traslación, escalado y/o giro) y **cuantificarlo**. Para ello se propone la utilización de una transformación afín en dos dimensiones (para errores planimétricos), ajustada mediante mínimos cuadrados (mm.cc.).

Para poder aplicar esta técnica, se deberán tener en cuenta tres aspectos fundamentales:

1. Los puntos se encuentran afectados de errores aleatorios
2. Los puntos se encuentran afectados por uno o varios errores sistemáticos
3. El número de puntos que interviene en el ajuste

Si bien el segundo de los aspectos se puede cuantificar mediante un ajuste mínimo cuadrático, el primero de ellos supondrá una distorsión que afectará tanto a los parámetros de la transformación afín como a los residuos del ajuste mínimo cuadrático. El error aleatorio estará íntimamente ligado al número de puntos que se empleen en el ajuste, dado que, a mayor error aleatorio, mayor número de puntos necesarios. Esta teoría quedará demostrada en base a los procesos de simulación.

En primer lugar, se ha generado una población distribuida en forma de malla regular uniforme de 1000 puntos (figura 43.5) y sobre ella se ha generado una contaminación

en base a una distribución normal aleatoria de media cero y desviación uno: $N(0,1)$. En siguiente lugar, se ha contaminado con dos traslaciones de magnitud conocida (X_0, Y_0) , dos factores de escala conocidos (λ_x, λ_y) , y un giro (α) también conocido (figura 43.6). Por último, se ha aplicado el estándar EMAS¹⁰ para comprobar su comportamiento en la detección de errores sistemáticos, y se ha procedido a un ajuste por mínimos cuadrados de la transformación afín en dos dimensiones para cuantificar la magnitud de cada uno de los errores sistemáticos.

Dada la gran influencia existente entre el tamaño muestral y el nivel de contaminación aleatoria de la población, se ha trabajado con diferentes tamaños de población a diferentes niveles de contaminación normal. El diagrama de flujo seguido en el proceso de simulación se muestra en la figura 43.7.

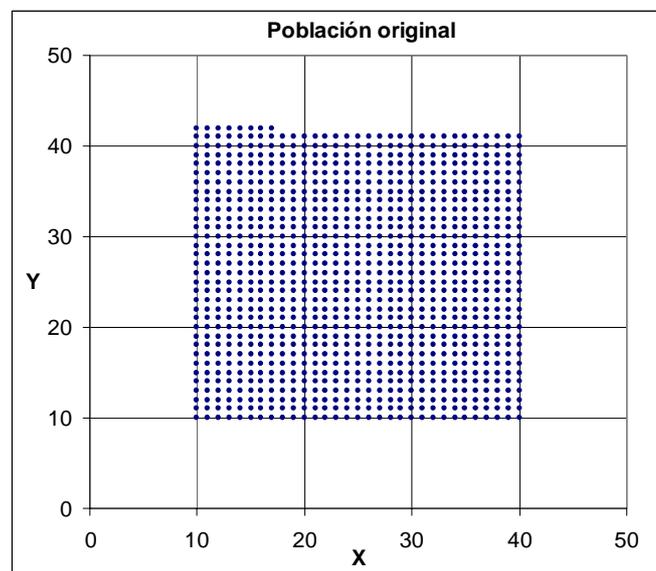


Figura 43.5: Distribución de la población original de 1000 puntos simulada de forma homogénea.

¹⁰ Se ha decidido utilizar el estándar EMAS por su capacidad para detectar errores sistemáticos.

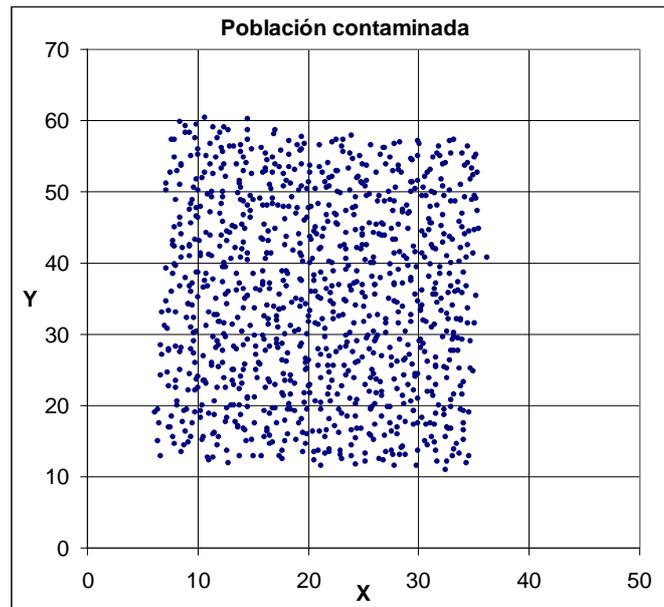


Figura 43.6: Distribución de la población contaminada por dos traslaciones, dos factores de escala, un giro y errores aleatorios distribuidos de forma normal.

El decidir contaminar en primer lugar con una distribución aleatoria $N(0,\sigma)$, se debe a que, intrínsecamente, en condiciones habituales todos los datos tendrán un determinado error aleatorio. Asimismo, su variabilidad vendrá fijada por la exactitud a la que haya sido realizada la BDG objeto del análisis. En segundo lugar, se ha procedido a la contaminación de los datos con uno o varios errores sistemáticos, porque se entiende que éstos se producirán como causa del tratamiento y manipulación de los datos en los procesos de producción cartográfica. Aunque pudiera parecer un hecho sin relevancia, éste resulta de vital importancia puesto que los resultados pueden variar sustancialmente.

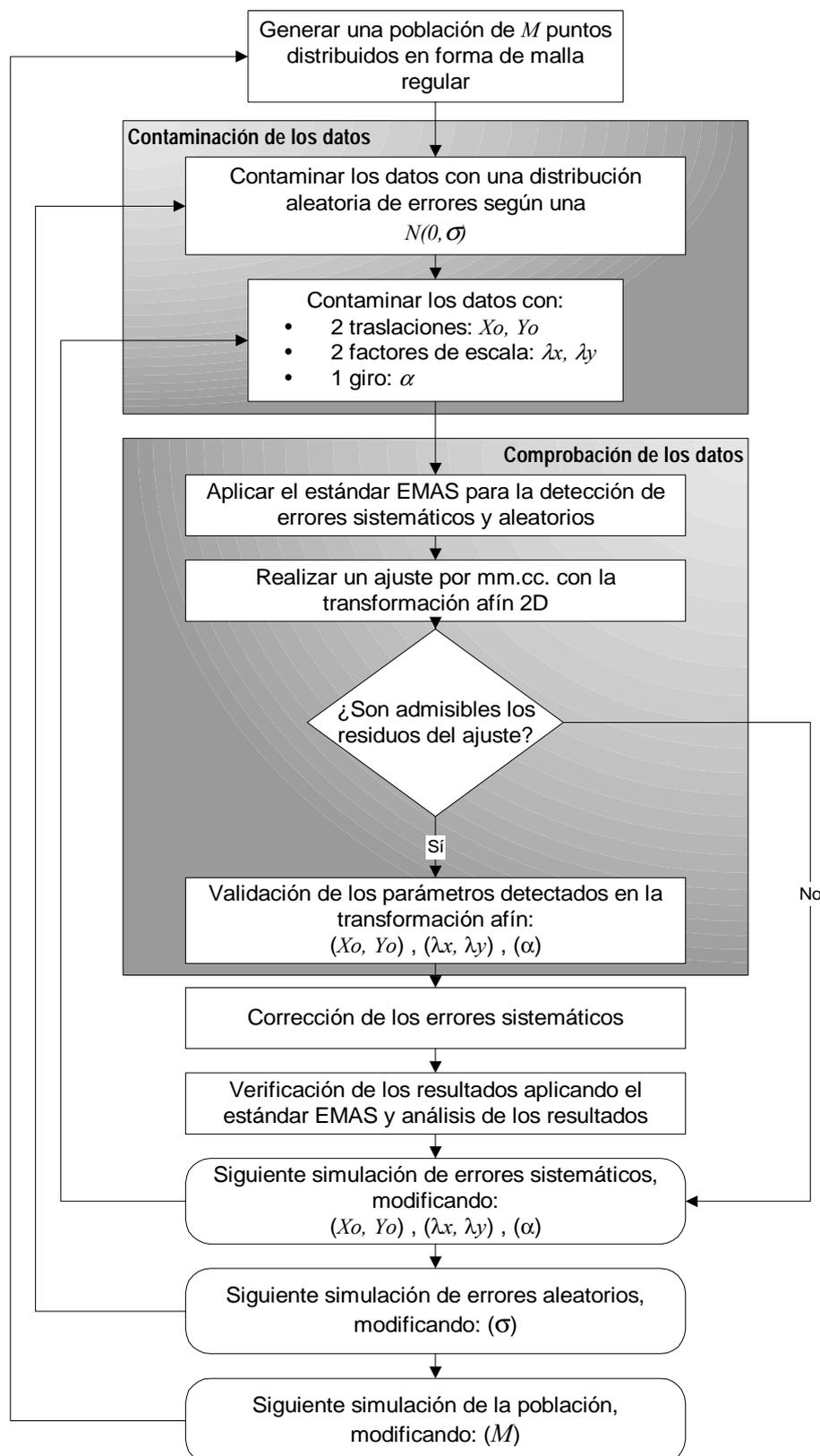


Figura 43.7: Diagrama de flujo seguido en el proceso de simulación para el estudio de la aplicación del ajuste por mm.cc.

Tal y como se ha mencionado anteriormente, los primeros experimentos se desarrollaron sobre una población de 1000 puntos con una contaminación aleatoria según una distribución $N(0,1)$. De esta forma, se pretende alcanzar una gran robustez en los resultados del ajuste por mm.cc.

4.3.5 Aplicación de estimadores circulares.

Para analizar el comportamiento de los estimadores circulares, se ha optado por trabajar con diferentes poblaciones de 1000 puntos de control distribuidos según una $N(0,1)$.

Dadas las características de los estimadores circulares y, al basarse en el análisis de las direcciones, se han contaminado las poblaciones artificialmente con errores de traslación sistemática sobre una o las dos componentes posicionales (XY), estudiando el grado de sensibilidad en su detección y la capacidad de obtener dicha dirección.

Al igual que en la cuantificación de sistematismos del apartado anterior (4.3.4), varios de los experimentos se han realizado sobre las mismas poblaciones contaminadas. Asimismo, se ha procedido a contaminar de forma progresiva y en idéntica magnitud cada componente posicional de la distribución $N(0,1)$.

Los estándares analizados han sido:

- Test Ómnibus.
- Test de Mardia.
- Test para la determinación de la dirección predominante y su intervalo de confianza.

Todos ellos se han contrastado al 95 % de nivel de confianza.

Junto con las poblaciones generadas artificialmente, los test también se han aplicado sobre datos reales tomados en la Hoja 926 1-1 del ICA.

Dado que la estadística circular se basa en el análisis de las direcciones, se ha tratado de estudiar el comportamiento de los módulos del error en función de la dirección en la que se producen. Para ello, se ha optado por desarrollar en forma de gráfico Acimut / Módulo los puntos de control utilizados en el control de la exactitud posicional.

Capítulo 5

Resultados y discusión

5 Resultados y discusión

Basándose en el trabajo realizado anteriormente, en este capítulo se presentan los resultados del mismo así como su análisis e interpretación. El capítulo 5 se divide en 5 apartados:

- 1º. Comparativa entre los diferentes estimadores robustos y la estadística clásica. Se exponen los resultados de aplicar estimadores robustos y técnicas de estadística clásica mediante la simulación de una serie de poblaciones con diferente grado de contaminación. Asimismo, se hace un estudio comparativo para analizar cuál o cuales son los más apropiados para el control de calidad posicional.
- 2º. Análisis comparativo de los distintos estándares en el control de calidad posicional. Con los procesos de simulación realizados en el capítulo 4, se exponen los resultados alcanzados en cuanto a niveles de exigencia, número de puntos de control, características propias de cada estándar, etcétera.
- 3º. Corrección mediante ajuste mínimo cuadrático de errores sistemáticos. Se plantea la utilización de la herramienta del ajuste por mm.cc. para cuantificar y corregir sistematismos en BDGs.
- 4º. Aplicación de estimadores circulares para la detección de posibles sistematismos en forma de direcciones predominantes en la distribución de errores, así como la determinación de dicha dirección y el análisis de los módulos de error.
- 5º. Aplicación de los distintos estándares mediante curvas operativas. Para cada estándar se desarrollan nomogramas de aceptación en función del tamaño muestral y las características de la distribución de los errores de la BDG con respecto a los niveles de exigencia aplicables en cada caso. De esta forma, se trata de determinar los riesgos de usuario y productor al aplicar cada uno de los estándares.
- 6º. Propuesta metodológica de aplicación en conjunto de aquellos aspectos que ofrecen mejores resultados en un posible “*Estándar de exactitud posicional en información georreferenciada (EPIG)*”, centrado sobre la exactitud horizontal de datos puntuales en el “*Estándar de Exactitud Horizontal en información puntual georreferenciada (ExHor)*”.

Al final de cada estudio, se exponen una serie de consideraciones a modo de resumen en función de los resultados obtenidos.

5.1 Comparativa entre los diferentes estimadores robustos y la estadística clásica

Con el análisis previo y ajuste de los parámetros constantes (ver capítulo 4, apartado 4.3.1), se ha procedido a realizar un estudio comparativo de los distintos estimadores robustos. Cabe destacar que todos los resultados se basan en la hipótesis de la distribución normal de los errores, al igual que consideran la mayoría de los estándares de control posicional cartográfico (NMA, EMAS, ASPRS, NSSDA...).

Se ha analizado el comportamiento de los estimadores robustos calculando los valores de μ y σ con la totalidad de los datos y eliminando aquellos valores que superan $3\cdot\sigma$ tal y como se suele actuar habitualmente.

Los resultados que aparecen en la tabla 51.1, se muestran de la siguiente forma:

- En las dos primeras columnas las características de la muestra original, así como el porcentaje y el tipo de contaminación aplicada.
- Las siguientes columnas presentan los valores de μ y σ para:
 - o Estimación clásica empleando todos los puntos y eliminando aquellos que superen $3\cdot\sigma_p$ (¹).
 - o Estimación por MLT con una única iteración.
 - o Estimación gaussiana por atípicos medios (en función del RQ), estimación gaussiana a más de $3\cdot\sigma_p$, y estimación gaussiana a más de $2.5\cdot\sigma_p$.
 - o Estimación aplicado el método danés con una ponderación sobre aquellos datos que superen $3\cdot\sigma_p$ y sobre aquellos datos que superen $2.5\cdot\sigma_p$.

¹ Por σ_p se entiende la desviación típica poblacional estimada a priori.

- o Estimación aplicado el método de Geman y Mc.Clure con una ponderación sobre aquellos datos que superen $3 \cdot \sigma_p$ y sobre aquellos datos que superen $2.5 \cdot \sigma_p$.
- Las filas muestran los resultados obtenidos para cada una de las componentes (X, Y) y en cada uno de los experimentos simulados.

Para facilitar el análisis de los resultados, se ha marcado para cada población en color verde aquel estimador con el que se obtiene el resultado más favorable (la media se aproxima a 0 y la desviación a 1, 2 o 10, dependiendo del tipo de muestra principal empleada), en color amarillo el segundo mejor resultado y en naranja el tercero. No obstante, en algunos casos las diferencias son mínimas, por lo que es interesante detenerse a observar las magnitudes de los valores de la tabla (51.1).

Tabla 51.1 Resultados obtenidos tras aplicar diferentes estadísticos para calcular μ y σ . Comparativa con diferentes poblaciones y niveles de contaminación.

Pobla.Orig.	Experimento	Componente	Contaminación (%)		CLÁSICOS				MLT		Función GAUSIANA				Método DANÉS				Geman & Mc-Clure					
					TODOS		SOLO < 3 σ		μ	σ	CUARTILES		3· σ		2.5· σ		3· σ		2.5· σ		3· σ		2.5· σ	
					μ	σ	μ	σ			μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
$\mu = 0$ $\sigma = 1$	A	X	0	1.1	-0.007	1.010	0.007	1.010	0.024	0.923	0.007	1.010	0.007	1.010	-0.020	0.973	-0.007	1.010	0.002	0.991	-0.007	1.010	0.004	0.986
		Y	0	1.1	0.023	0.988	-0.023	0.988	0.015	0.865	-0.048	0.954	-0.023	0.988	-0.048	0.954	0.023	0.988	0.030	0.973	0.023	0.988	0.031	0.970
	B	X	0	1.5	-0.006	1.076	-0.038	1.032	-0.004	0.977	-0.025	1.029	-0.025	1.029	-0.026	0.958	0.006	1.048	0.007	1.013	0.010	1.040	0.012	0.998
		Y	0	1.5	0.044	1.047	-0.055	0.986	-0.017	0.906	-0.077	0.870	-0.075	0.997	-0.100	0.962	0.056	1.016	0.063	1.001	0.061	1.008	0.069	0.990
	C	X	1.5	1	-0.006	1.076	-0.038	1.032	-0.004	0.977	-0.025	1.029	-0.025	1.029	-0.026	0.958	0.006	1.048	0.007	1.013	0.010	1.040	0.012	0.998
		Y	1.5	1	0.044	1.047	-0.055	0.986	-0.017	0.906	-0.077	0.870	-0.075	0.997	-0.100	0.962	0.056	1.016	0.063	1.001	0.061	1.008	0.069	0.990
	D	X	1.5	1.5	0.133	1.146	-0.093	1.102	-0.099	1.027	-0.067	1.050	-0.098	1.093	-0.067	0.977	0.105	1.082	0.105	1.045	0.095	1.067	0.094	1.021
		Y	1.5	1.5	0.164	1.161	-0.090	1.048	-0.066	0.981	-0.089	0.911	-0.088	1.037	-0.137	0.971	0.125	1.066	0.125	1.037	0.113	1.050	0.112	1.015
	E	X	0	2	-0.017	1.124	-0.043	1.053	-0.001	1.002	-0.024	1.046	-0.024	1.046	0.002	0.935	0.007	1.061	-0.002	1.006	0.013	1.052	0.002	0.984
		Y	0	2	0.014	1.144	0.000	1.038	0.030	0.979	-0.036	0.996	-0.011	1.028	-0.036	0.996	0.011	1.063	0.018	1.049	0.011	1.045	0.019	1.029
	F	X	0	3	0.014	1.335	-0.007	1.080	-0.007	1.141	-0.014	1.055	-0.014	1.055	-0.041	1.019	-0.016	1.091	-0.008	1.074	-0.006	1.076	0.005	1.054
		Y	0	3	0.011	1.289	0.027	1.027	0.028	1.066	-0.017	0.967	0.039	1.046	-0.042	0.932	-0.025	1.087	0.000	1.031	-0.029	1.065	0.004	0.995
	G	X	0	4	-0.017	1.531	-0.003	1.075	0.013	1.187	-0.015	0.982	-0.045	1.025	-0.042	0.942	0.033	1.046	0.031	1.002	0.035	1.038	0.031	0.981
		Y	0	4	-0.031	1.533	-0.011	1.095	0.036	1.245	-0.027	1.053	-0.027	1.053	-0.001	1.021	0.015	1.108	0.014	1.079	0.018	1.083	0.017	1.048
H	X	4	4	0.383	1.953	-0.011	0.948	-0.109	1.421	0.057	0.984	0.003	0.909	0.003	0.909	0.009	1.029	0.026	0.992	0.001	1.012	0.023	0.965	
	Y	4	4	0.369	1.957	0.038	1.023	-0.086	1.412	0.000	0.974	-0.031	1.023	-0.048	0.907	0.056	1.027	0.063	1.012	0.051	1.006	0.059	0.988	
I	X	0	2	-0.058	1.108	0.055	0.979	0.059	1.109	0.054	1.000	0.054	1.000	0.001	0.928	-0.053	1.036	-0.038	1.002	-0.052	1.019	-0.032	0.976	
	Y	0	5	-0.016	1.795	0.039	1.041	0.086	1.336	0.014	1.027	0.014	1.027	0.014	0.961	-0.043	1.093	-0.043	1.063	-0.030	1.059	-0.031	1.024	
J	X	0	1.2	-0.034	0.980	0.034	0.980	0.039	0.897	0.034	0.980	0.034	0.980	0.034	0.980	-0.034	0.980	-0.034	0.980	-0.034	0.980	-0.034	0.980	
	Y	0	1.2	-0.074	1.065	0.074	1.065	0.106	0.923	0.074	1.065	0.074	1.065	0.074	1.065	-0.074	1.065	-0.009	1.009	-0.074	1.065	-0.057	1.031	
K	X	1.2	1.2	0.245	1.046	-0.248	1.051	-0.236	0.967	-0.245	1.046	-0.245	1.046	-0.188	0.975	0.245	1.046	0.201	0.977	0.245	1.046	0.219	0.993	
	Y	1.2	1.2	0.268	1.127	-0.235	1.083	-0.222	0.995	-0.199	0.973	-0.233	1.078	-0.199	0.973	0.214	1.041	0.193	1.010	0.231	1.053	0.219	1.031	
L	X	0	3	-0.052	1.548	-0.026	0.979	0.018	1.242	-0.076	1.017	-0.047	0.975	-0.023	0.944	0.106	1.021	0.083	0.982	0.041	1.065	0.026	1.035	
	Y	3	2	0.595	1.779	-0.182	1.169	-0.289	1.457	-0.269	1.280	-0.256	1.197	-0.284	1.076	0.172	1.111	0.158	1.007	0.244	1.147	0.223	1.078	
$\mu = 0$ $\sigma = 2$	M	X	0	4	-0.015	2.180	-0.010	1.918	0.044	1.863	-0.054	1.844	-0.039	1.898	-0.009	1.898	0.026	1.901	0.016	1.931	0.022	1.901	0.012	1.931
		Y	0	4	-0.112	2.189	0.060	2.096	0.020	1.999	0.043	2.081	0.043	2.081	0.050	1.929	-0.050	2.082	-0.059	1.978	-0.050	2.082	-0.058	1.975
	N	X	4	4	0.331	2.536	-0.011	1.966	-0.134	2.173	-0.024	2.035	-0.054	2.035	-0.024	2.035	0.052	2.047	0.005	1.996	0.060	2.048	0.009	1.986
		Y	4	4	0.388	2.530	-0.127	2.072	-0.308	2.138	-0.125	1.988	-0.186	2.078	-0.176	1.916	0.159	1.960	0.117	1.944	0.153	1.960	0.125	1.930
	O	X	0	6	-0.069	2.628	0.138	2.056	0.145	2.156	0.125	1.896	0.135	2.042	0.077	1.959	-0.127	2.034	-0.059	1.973	-0.128	2.045	-0.067	1.973
Y		0	6	0.046	2.633	-0.020	2.085	-0.100	2.225	-0.009	1.995	-0.059	1.951	-0.004	1.842	0.042	1.987	-0.005	1.933	0.026	1.989	0.021	1.906	
P	X	0	8	-0.033	3.066	0.048	1.932	0.103	2.401	0.113	1.846	-0.029	2.026	-0.029	2.026	0.027	2.028	0.027	2.028	0.023	2.030	0.023	2.030	
	Y	0	8	-0.024	3.101	0.052	2.041	0.015	2.471	0.045	2.008	0.045	2.038	0.049	1.856	-0.055	2.030	-0.042	1.907	-0.038	2.032	-0.033	1.895	
$\mu = 0$ $\sigma = 10$	Q	X	0	20	-0.577	11.173	0.841	9.992	0.534	9.715	0.571	10.160	0.571	10.160	0.556	9.505	-0.571	10.160	-0.311	9.830	-0.571	10.160	-0.313	9.830
		Y	0	20	-0.634	11.133	0.170	10.384	0.745	9.591	0.256	10.478	0.256	10.478	0.502	10.173	-0.256	10.478	-0.256	10.478	-0.257	10.478	-0.257	10.478
	R	X	25	30	2.084	15.216	0.484	10.365	-0.814	11.940	0.063	10.523	0.063	10.523	-0.135	9.486	-0.063	10.523	-0.094	9.763	-0.059	10.523	-0.091	9.763
		Y	25	30	2.435	15.227	0.442	10.189	-0.774	12.316	0.415	9.876	0.415	9.876	0.415	9.876	-0.415	9.876	-0.415	9.876	-0.409	9.876	-0.409	9.876
	S	X	0	40	-0.134	15.416	0.149	10.127	0.157	12.420	0.382	9.804	0.382	9.804	0.121	9.455	-0.382	9.804	-0.121	9.455	-0.382	9.804	-0.124	9.455
		Y	0	40	0.337	15.482	-0.374	10.237	-0.283	12.485	-0.084	10.207	-0.084	10.207	-0.339	9.876	0.084	10.207	0.339	9.876	0.085	10.207	0.338	9.876
T	X	15	50	1.836	18.492	-0.434	10.620	-0.976	13.547	-0.476	10.261	-0.476	10.261	-0.737	9.906	0.476	10.261	0.472	9.561	0.478	10.261	0.474	9.560	
	Y	25	30	2.274	15.275	-0.164	11.050	-0.716	12.670	-0.157	11.061	-0.157	11.061	0.171	10.014	0.157	11.061	-0.410	10.292	0.161	11.061	-0.403	10.292	

Leyenda	
	1º mejor resultado
	2º mejor resultado
	3º mejor resultado

Por ejemplo, en el caso de las dos primeras filas se ha trabajado con la población sin contaminar, distribuyéndose los errores en X e Y según una $N(0,1)$. Lógicamente, los mejores resultados se obtienen empleando la totalidad de los datos, mediante estadística clásica. También se obtienen resultados muy similares tanto con el método danés como con el de Geman y Mc.Clure al no existir ponderación.

En el caso del experimento *A*, los mejores resultados se obtienen mediante la ponderación gaussiana (con un “corte” a partir de $3\cdot\sigma_p$), así como con el estimador clásico (eliminando aquellos datos que superan $3\cdot\sigma_p$).

Se ha trabajado de igual forma en el resto de los experimentos. Así, en el experimento *O*, los mejores resultados se obtienen con la ponderación por el método danés (también con un “corte” a partir de $3\cdot\sigma_p$), así como con la ponderación de Geman y Mc.Clure (tanto con un “corte” de $3\cdot\sigma$ como de $2.5\cdot\sigma_p$). En el caso del experimento *R*, el mejor resultado se obtiene con la ponderación a partir de $2.5\cdot\sigma_p$ con los pesos de Geman & Mc.Clure, seguido por la ponderación del método danés (también a partir de $2.5\cdot\sigma_p$) y con la ponderación a partir de $3\cdot\sigma_p$ de Geman y Mc.Clure.

Observando los resultados de forma global, el método danés resulta el más idóneo si se aplica con un corte de ponderación sobre aquellos datos que superen $2.5\cdot\sigma_p$. El siguiente estimador más apropiado resulta ser el de Geman y Mc.Clure cuando también se aplica con un corte de ponderación sobre aquellos datos que superen $2.5\cdot\sigma_p$ (en este caso es el 2º mejor estimador), así como si se aplica con un corte sobre aquellos datos que superen $3\cdot\sigma_p$ (siendo el 3º mejor estimador).

El estimador MLT no resulta aconsejable en este tipo de trabajos dado que, incluso con una única iteración, los resultados no se aproximan sustancialmente a los valores deseados. Tan sólo en el caso de los experimentos *D* y *E* obtiene los mejores valores finales.

En cuanto a la ponderación gaussiana con diferentes niveles de “corte”, si bien en un principio se obtuvieron resultados muy prometedores (Atkinson, Ariza y Rodríguez, 2002), al realizar la ponderación mediante el método danés y mediante el suavizado de Geman y Mc.Clure, los resultados mejoran ligeramente.

Si, por el contrario, se atiende a las columnas de estimadores clásicos, prácticamente resulta desaconsejable el emplear la totalidad de los datos sin realizar un estudio previo de posibles atípicos. Asimismo, el hecho de eliminar aquellos valores que superan $3\cdot\sigma_p$, aparte de desechar posible información útil, sólo ofrece resultados óptimos con muy bajos índices de contaminación muestral.

Consideraciones sobre los estimadores robustos:

Actualmente, el procedimiento comúnmente aceptado es el de la supresión de aquellos valores considerados como atípicos al encontrarse a una distancia determinada del valor medio. No obstante, y dado que el usuario de una BDG puede encontrarse con estos valores, se le ha de presentar algún informe en el que queden especificados y cuantificados.

Asimismo, teniendo en cuenta el coste que supone la captura de datos en campo (en el caso de que ésta sea la fuente de mayor exactitud), es importante el tratar de emplear toda la información disponible sin eliminar ningún tipo de dato.

A modo de resumen, el mejor resultado de todos los experimentos efectuados lo ofrece el método danés con un valor de “corte” en la ponderación sobre aquellos valores que superan $2.5\cdot\sigma_p$. De esta forma se reduce considerablemente la influencia de valores atípicos procedentes de un proceso de contaminación sobre la muestra principal, aproximándose los valores de media y desviación globales a los datos originales.

5.2 Análisis comparativo de los diferentes estándares en el control de calidad posicional

Tal y como se ha mencionado anteriormente, a la hora de realizar el proceso comparativo entre los diferentes estándares, se ha optado por la simulación como herramienta básica de investigación (capítulo 4, apartados 4.1 y 4.3.2).

El proceso de simulación se ha centrado sobre los cuatro estándares más empleados: NMAS (*National Map Accuracy Standard, 1947*), EMAS (*Engineering Map Accuracy Standard,*

1991), ASPRS (*Spatial Accuracy Specification for Large Scale Topographic Maps, 1983*) y NSSDA (*National Standard for Spatial Data Accuracy, 1999*) (USBB, 1947; ASCI, 1983; ASPRS, 1989, FGDC, 1998).

A continuación se enuncian brevemente sus principales características y los resultados obtenidos con la simulación.

5.2.1 NMAS

En el caso del NMAS, sobre cada una de las muestras escogidas al azar, se ha obtenido el porcentaje medio de aceptación del mapa en función del número de puntos seleccionado al formar la muestra. Este valor medio viene dado por las 1000 muestras aleatorias tomadas sobre cada población. Asimismo, se ha obtenido la desviación que sufre dicha media en las 100 poblaciones simuladas. Ésta ofrecerá un intervalo en la variabilidad de los resultados, proporcionando las desviaciones finales de todo el proceso de simulación. Dichas desviaciones ofrecen al investigador una cuantificación en la estabilidad y robustez de los resultados.

Tal y como se menciona en el capítulo 3, aplicando el NMAS a una población con desviación de 1 m (escala 1/4.000), el límite máximo para los errores sería de 0,846 mm (1/30 de pulgada) al tratarse de una cartografía a escala mayor que 1/20.000.

Si se aplica el estándar al total de las poblaciones, en todos los casos la cartografía resulta aceptada al no superarse más de un 10 % de los puntos con dicho error. Empleando la simulación, se obtuvieron los resultados referentes al estándar NMAS que figuran en la tabla 52.1. En la primera columna (experimento A), aparecen los porcentajes medios de aceptación de las muestras aleatorias (sobre mapas simulados). La estabilidad de dichos resultados (desviación) no supera el 3 %. Tal y como se puede apreciar, el nivel de aceptación de los mapas es muy elevado a pesar de trabajar con muestras muy pequeñas: se obtiene el 99,95 % de aceptación con tan sólo 10 puntos de control.

Por el contrario, si se tratase de una cartografía a escala menor que 1/20.000, debería emplearse como límite de aceptación el que no exista más de un 10 % de puntos de

control con un error superior a 1/50 de pulgada (0,508 mm). Dado que se pretenden extrapolar los resultados del experimento a cualquier tipo de cartografía, se ha modificado el estándar original para que éste sea el límite de aceptación. Si se aplica dicho límite al total de la población, se obtiene como resultado el rechazo del mapa. En la siguiente columna (experimento *B*) de la tabla 52.1 aparecen los porcentajes promedio de aceptación para dichas condiciones. Se observa que, conforme aumenta el tamaño muestral, se reducen los porcentajes de aceptación con una tendencia hacia el 100 % de rechazo.

Ahora bien, ¿a qué se debe este elevado porcentaje de rechazos? El estándar NMAS resulta muy permisivo a la hora de aceptar cartografía de escala mayor que 1/20.000, pero es mucho más estricto cuando se trabaja con escalas inferiores a dicho límite. Este aspecto resulta chocante dado que es en este tipo de cartografía donde los efectos de la generalización y la simbología empleada en la representación son mucho más acusados.

Así, para que este tipo de cartografía resulte aceptada en un porcentaje en torno al 95 %, ha de ser elaborada con una calidad propia de una escala un 15 % mayor que la escala final. Esto es, para que no exista más de un 10% de puntos con un error superior a 1/50 de pulgada (0,508 mm) como promedio de las muestras seleccionadas, la desviación típica poblacional establecida a priori ha de ser un 15 % inferior que la de la escala del mapa. Por ejemplo, para una cartografía a escala 1/50.000, la desviación típica poblacional a priori sería de 12,5 m ($(0,2 \text{ mm} + 25 \%) \times 50.000 = 12,5 \text{ m}$), mas si se desea cumplir con el estándar NMAS, deberá ser producida para una desviación un 15% inferior (10,625 m). Los resultados de este proceso de simulación se aprecian en la columna de resultados de la tabla 52.1 correspondiente al experimento *C*, en donde se alcanza el 95 % de aceptación cuando el tamaño muestral oscila en torno a los 120 puntos de control.

Tabla 52.1: Porcentajes promedio de aceptación en el estándar NMAS².

Experimento →	A	B	C	D
Tamaño muestral <i>n</i>	Límite: 1/30" (0,846 mm)	Límite: 1/50" (0,508 mm)	Límite: 1/50" (0,508 mm) Cartografía un 15 % mejor	Límite: 1/40" (0,677 mm)
10	99.52	63.01	86.47	94.38
15	99.93	57.10	86.30	96.37
20	99.97	52.31	86.28	97.42
25	99.99	49.40	86.50	98.42
30	99.99	46.56	87.32	98.97
35	100.00	43.80	87.60	99.38
40	100.00	41.50	88.00	99.56
45	100.00	39.20	88.60	99.68
50	100.00	37.63	90.06	99.80
55	100	36.10	90.20	99.89
60		34.94	90.48	99.93
65		33.40	91.20	99.98
70		31.87	91.80	99.99
75		30.70	92.10	99.99
80	100	29.85	92.33	99.99
85		28.80	93.00	99.99
90		27.77	93.66	99.99
95		26.40	94.20	100.00
100		25.54	94.28	100.00
105	100	24.90	94.60	100
110		24.43	94.85	
115		23.40	94.96	
120		22.56	95.23	
125		21.90	95.57	

Para tratar de evaluar un posible caso intermedio entre las propuestas realizadas por el USGS, se ha repetido el experimento empleando como límite de aceptación el que no exista más de un 10 % de los puntos que superen un error de 1/40 de pulgada (0,677 mm). En este caso, la supuesta cartografía habrá sido elaborada de forma acorde con la desviación típica a priori: para una cartografía a escala 1/50.000 será de 12,5 m. Si se realiza el test NMAS al total de la población, se obtiene como resultado la aceptación del mapa. Los resultados del proceso de simulación se muestran en la última columna de la tabla 52.1 (experimento **D**). Lógicamente, los porcentajes de aceptación son superiores a los del experimento anterior para el mismo tamaño muestral, alcanzándose más del 97 % de aceptación para, tan sólo, 10 puntos de control.

En la figura 52.2 se presentan las gráficas representativas de los niveles de aceptación en función del tamaño muestral para cada uno de los 4 experimentos.

² La simulación se ha realizado sobre poblaciones de tamaño $m = 1000$ puntos con errores en la componente X e Y distribuidos según una $N(0,1)$

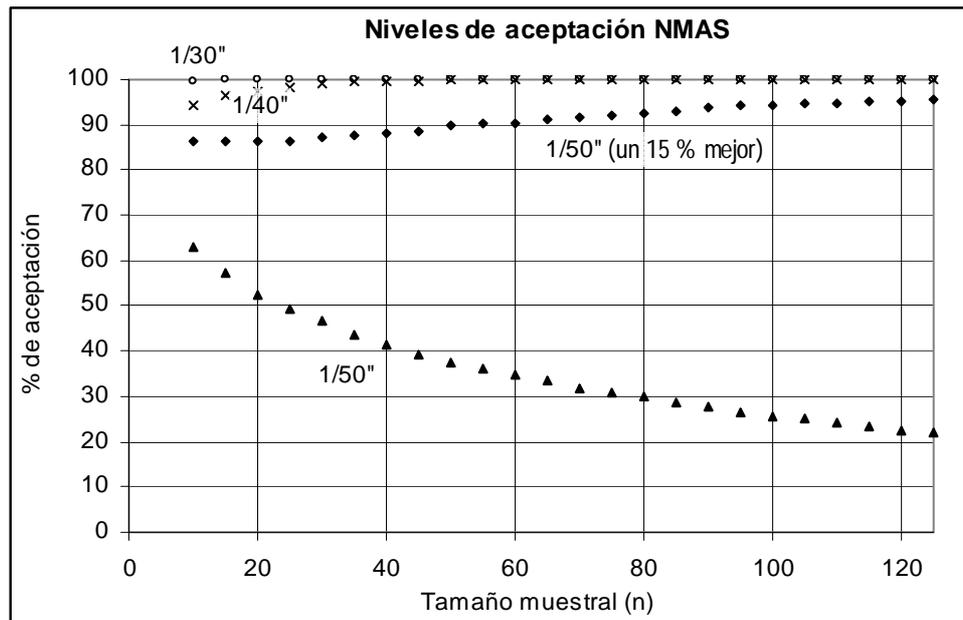


Figura 52.2: Niveles de aceptación del estándar NMAS en función del tamaño muestral³.

Consideraciones sobre el NMAS:

Aunque anticuado, es un estándar que ha sido profusamente empleado por organismos cartográficos y que todavía sigue apareciendo en algunos de los actuales Pliegos de Prescripciones Técnicas para la elaboración de cartografía. En cuanto a las características que muestra, cabría destacar que:

- El estándar es muy permisivo con cartografía a escala $> 1/20.000$ y, por el contrario, es muy restrictivo con escalas $< 1/20.000$.
- Para aceptar una BDG a escala $< 1/20.000$, ésta deberá ser un 15 % más exacta que lo estrictamente recomendable para su escala y tendrán que emplearse un mínimo de 120 puntos de control.
- Tan solo realiza un estudio sobre el porcentaje de valores que se encuentran por encima de una determinada tolerancia dada en función de la escala.
- En el caso de que no se supere, no se puede estimar cuánto costaría hacer una cartografía que cumpliera con el estándar. Análogamente, si se supera, no

³ El porcentaje de aceptación de muestra para cada uno de los experimentos de la tabla 52.1: **A:** límite 1/30" (0,846 mm); **B:** límite: 1/50" (0,508mm); **C:** límite: 1/50" (0,508 mm) con cartografía un 15 % mejor; y **D:** límite: 1/40" (0,677mm)

se puede estimar cuánto mejor es la cartografía y saber si se está produciendo más exacta de lo estrictamente necesario (con el correspondiente incremento de costes para el productor).

5.2.2 EMAS

El EMAS es uno de los test más exhaustivos, dado que primero estudia la existencia de algún tipo de comportamiento sistemático, y después analiza el comportamiento de la varianza para la detección de errores casuales o accidentales. La cartografía será aceptada solamente si cumple con todas las hipótesis: no posee errores sistemáticos ni errores casuales en ninguna de sus componentes (X e Y).

5.2.2.1 Nivel de confianza para el estándar EMAS

Tras un estudio del test, se observa que la distribución *t-Student* posee 2 colas, con lo que en vez de emplear α deberá utilizarse $\alpha/2$ para mantener el nivel de significación final (empleando $|t_x| \leq t_{n-1, \alpha/2}$ en vez de $|t_x| \leq t_{n-1, \alpha}$). Tal y como aparece en la referencia bibliográfica (ASP, 1985), la distribución *t-Student* empleada era de un nivel de significación del 5 % en cada cola, por lo que, en realidad, se estaba trabajando al 90 % de nivel de confianza para la mencionada prueba.

Una útil práctica para comprobar varias hipótesis, se basa en utilizar la corrección de Bonferroni (García y Martín, 1997). Si se desea comprobar “ n ” hipótesis independientes con un nivel de confianza final de 0,95, deberá emplearse un nivel de significación de $(0.05/n)$ para cada una de ellas. Es decir, si se realizan dos contrastes de hipótesis, se declarará haber obtenido un resultado estadísticamente significativo sólo si $p < 0,025$ (Tobías, 2001).

Por ello, se ha de aplicar Bonferroni (Rodríguez, 2000) para evitar que la acumulación de dicho nivel aumente progresivamente. De esta forma, se presentan dos situaciones: el nivel de confianza real tal y como se ha aplicado el test hasta ahora (ver *test original* en la tabla 52.2), en el que el nivel de confianza final sería $0.95^2 \times 0.95^2 \times 0.95 \times 0.95 \approx 0.73$; y el nivel de confianza real que se ha empleado en el proceso de simulación

aplicando Bonferroni (*test corregido*) para analizar la componente de la calidad planimétrica (X e Y), siendo éste de $0.9875 \times 0.9875 \times 0.9875 \times 0.9875 \approx 0.95$.

Este aspecto es de vital importancia y ha sido validado en el proceso de simulación, tal y como se podrá comprobar al final del apartado.

Tabla 52.2 – Relación entre los niveles de significación parciales y el nivel de confianza global para un control de calidad posicional en planimetría (X e Y) aplicando Bonferroni⁴.

Test →	Niveles de significación parciales: α %				Nivel de significación acumulado α %	Nivel de confianza final $(1-\alpha)$ %
	t-Student (2 colas)		Chi-Cuadrado (1 cola)			
Componente →	X	Y	X	Y		
Test original →	5 (x2)	5 (x2)	5	5	27	73
Valores límite para $n = 25 \rightarrow$	$t_{n-1, \alpha(x2)} = 1.71$		$\chi^2_{n-1, \alpha} = 36.42$			
Bonferroni →	$\alpha/4$	$\alpha/4$	$\alpha/4$	$\alpha/4$	5	95
Test corregido →	1.25	1.25	1.25	1.25		
Valores límite para $n = 25 \rightarrow$	$t_{n-1, \alpha/4} = 3.00$		$\chi^2_{n-1, \alpha/4} = 42.12$			

5.2.2.2 Análisis del estándar EMAS

Tras la obtención de las muestras, cada una de ellas es sometida al test EMAS bajo el supuesto de media $\mu = 0$, y de varianza $\sigma^2 = 1$. Para su estudio y representación gráfica (figura 52.4), en función del número n de puntos en la muestra, se contabilizan los porcentajes de paso de la componente sistemática (test t_x y t_y de forma conjunta), y la componente de variabilidad (test χ_x^2 , χ_y^2 en conjunto). Junto a ellos, en la figura 52.4 se representan los porcentajes de aceptación globales para el EMAS. Como puede observarse, el porcentaje de rechazo del EMAS viene determinado prácticamente por la suma directa de los rechazos de las componentes de sistematismo y varianza lo que supone la independencia en su comportamiento. Además, la componente del sistematismo es la que presenta un comportamiento de aceptación más restrictivo. La figura 52.4 podría considerarse como la curva de aceptación que genera el Test EMAS bajo los criterios anteriormente expuestos ($\mu = 0$, $\sigma^2 = 1$).

⁴ Ejemplo de aplicación para $n = 25$ puntos de control a un nivel de confianza del 95 %.

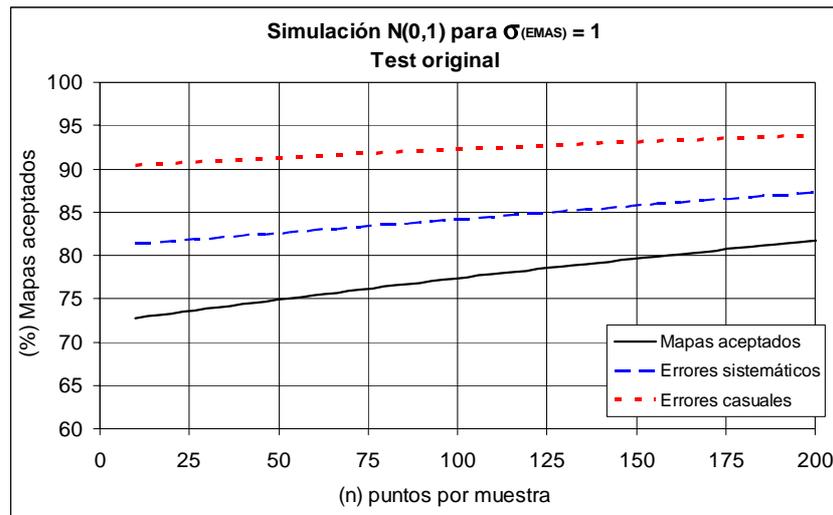


Figura 52.4: Porcentajes de aceptación del estándar EMAS original⁵.

Otro aspecto de interés es el comportamiento en el origen. La curva arranca con valores próximos al 90%, 80% y 70% de aceptación en varianza, sistematismo y global, respectivamente. Estos valores tienen su justificación en la propia construcción del estándar: se trata de una batería de 4 test que hay que superar para considerar la aceptación de la BDG.

El comportamiento señalado demuestra que el test EMAS es demasiado restrictivo según el planteamiento propuesto originariamente, lo cual plantea serios problemas tanto al productor como al usuario. Para el primero se rechazan un gran número de mapas que son correctos, mientras que el segundo tiene problemas en aceptarlos, por lo que deberá pagar más para obtener la calidad que desea.

La racionalización del comportamiento descrito en el EMAS pasa por la aplicación de niveles de significación en la batería de test, que tengan en cuenta una significación global $\alpha = 5\%$ (según los criterios de Bonferroni descritos en la tabla 52.2).

En la forma que adopta la nueva curva (figura 52.5), los comportamientos respecto a la aceptación según los sistematismos, la varianza y el Test global, alcanzan valores más adecuados con puntos de arranque en el 95 % para el caso global (tal y como se pretendía). Cabe destacar que con los 20 – 25 puntos de control como tamaño

⁵ Las curvas representan el porcentaje de muestras sin errores sistemáticos (en X e Y), sin errores casuales o aleatorios (X e Y), y valores en su conjunto (muestras aceptadas sin errores sistemáticos ni aleatorios).

muestral recomendado por el estándar, se obtienen los niveles de aceptación deseados (95 %).

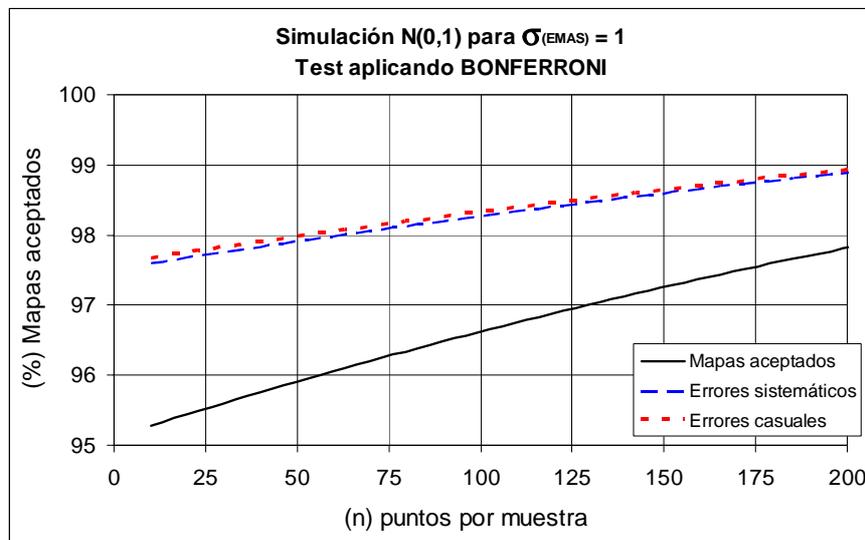


Figura 52.5: Test con corrección en el nivel de significación global (Bonferroni).

Todo ello muestra la validez de:

- el empleo de la simulación como herramienta apropiada y necesaria en el análisis de los estándares, y
- la necesidad de aplicar la corrección de Bonferroni a la hora de determinar los niveles de significación parciales y globales de una batería de test, como es el caso del estándar EMAS.

Consideraciones sobre el EMAS:

A pesar de no haber sido empleado profusamente, es uno de los estándares que, a priori, ofrecen una información más exhaustiva sobre los errores de una BDG. A pesar de ello, el estándar original posee graves deficiencias:

- El nivel de significación empleado en el contraste para la detección de sistematismos (*t-Student*) es del 10 % para cada componente. Esto hace que el estándar sea muy restrictivo.

- El nivel de significación para cada componente en el contraste para el estudio de la varianza (*Chi-Cuadrado*) es del 5 %. Con ello, el estándar se hace más restrictivo.
- Si no se efectúa la corrección por Bonferroni, el nivel de confianza final del estándar es del 73 % en vez del 95 % (tal y como se entiende de la lectura del estándar). De esta forma, el riesgo del productor es excesivamente elevado (el 27 % de los mapas “buenos” son rechazados), mientras que prácticamente se ha eliminado el riesgo del usuario.

5.2.3 ASPRS

Tal y como se mencionaba en el capítulo 3, el estándar se basa en el establecimiento de categorías de exactitud para los diferentes mapas en función del RMSE. En el caso de los mapas de mejor calidad (Clase I), se establecen límites al RMSE para las componentes X e Y en función de su escala, siendo los límites para las clases II y III, el producto los valores dados para la clase I (x_2) y (x_3) respectivamente.

Empleando el mismo proceso de simulación que en casos anteriores (ver capítulo 4, figura 43.3), se ha trabajado con una serie de 100 poblaciones de 1000 puntos con errores en X e Y distribuidos según una $N(0,1)$. En este caso, el total de la población se correspondería a un mapa con desviación típica a priori de 1 m para la Clase I. Mediante simulación, el programa ofrece como resultado el porcentaje de pertenencia a cada una de las clases antes mencionadas. En cuanto a la estabilidad y robustez de los resultados (desviación típica entre poblaciones), ésta se sitúa entre un 2 y un 4,5 % (tabla 52.3).

Tabla 52.3 – Porcentaje de mapas pertenecientes a la clase *I* del estándar ASPRS en función del tamaño muestral y estabilidad en los resultados obtenidos (desviación del proceso de simulación).

Muestra de tamaño <i>n</i>	% ∈ Clase <i>I</i>	Desviación (%)
10	31,68	1.23
20	29,30	1.05
30	29,28	0.91
40	28,51	0.73
50	27,73	0.57
60	27,43	0.50
70	27,38	0.49
80	27,08	0.46
90	27,20	0.38
100	27,59	0.32
150	27,05	0,21

Analizando estos valores, si bien el total de la población se ajustaría a un mapa de la Clase *I*, se observa que:

- En el tramo entre 10 y 90 puntos de tamaño muestral, el porcentaje de pertenencia a la Clase *I* disminuye desde un 31 a un 27 % (con una variabilidad de $\pm 2,5$ %).
- En el tramo entre 100 y 150 puntos de tamaño muestral, el porcentaje de pertenencia a la Clase *I* se mantiene en torno al 27 % (con una variabilidad de ± 4 %).

Los resultados muestran la tendencia del estándar a clasificar el mapa en una clase inferior a la que realmente pertenece (en el 75 % de los casos), por lo que se trata de un estándar muy restrictivo.

Para estudiar esta hipótesis de forma más exhaustiva, se han realizado nuevos experimentos: manteniendo las características de las poblaciones generadas aleatoriamente (distribución $N(0,1)$), se han establecido diferentes desviaciones típicas como las estimadas a priori para el mapa. Los resultados se muestran en la tabla 52.4 ⁽⁶⁾, así como su representación gráfica en las figuras 52.6 y 52.7.

⁶ En la columna referida al experimento **F**, se han incluido los resultados de la tabla 52.3 para facilitar su análisis.

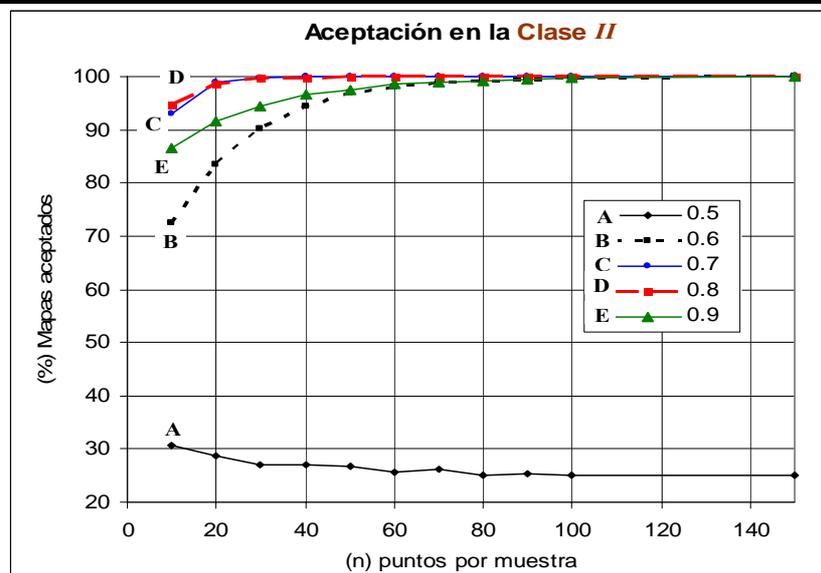
Entre las columnas **A** y **E** se ha aplicado un criterio de exigencia superior al de la población total, reduciendo la desviación típica estimada para la Clase *I* a la mitad de la que posee dicha población. Entre 0,5 y 0,7 veces dicha desviación (experimentos **A** – **C**), el resto de los mapas se clasifican como pertenecientes a la Clase *III*, mientras que entre 0,8 y 0,9 veces (experimentos **D** y **E**), lo hacen como pertenecientes a la Clase *I*. Observando la figura 52.6, se aprecia cómo el comportamiento del experimento **A** es muy similar al del experimento **F** (figura 5.7), dado que se ha trasladado el nivel de exigencia en un orden de categoría, por lo que los porcentajes de aceptación también se trasladan en dicho orden de categoría. Entre 0,6 y 0,7 desviaciones típicas, el porcentaje de pertenencia a la Clase *II* va en aumento de forma progresiva con el tamaño muestral, mientras que el resto de las muestras que no aparecen representadas por la gráfica, son clasificadas como pertenecientes a una clase de categoría inferior (Clase *III*).

Este último aspecto cambia al trabajar entre 0,8 y 0,9 veces la desviación típica (experimentos **D** y **E**), en donde el resto de las muestras pasa a pertenecer a una clase de categoría superior (Clase *I*). Esto es, si por ejemplo un cliente analiza un mapa con un nivel de exigencia de una décima parte mejor que la población real ($\sigma_{\text{mapa}} = 1 \text{ m}$ y $\sigma_{\text{clase } I} = 0,9 \text{ m}$), con 40 puntos de control tendrá más de un 95 % de posibilidades de que se clasifique como perteneciente a la Clase *II* (porcentaje de muestras que pertenece a la clase *II* = 96,6 %).

Observando el resto de los experimentos (**G**, **H** e **I**), se aprecia cómo se comporta el estándar cuando el nivel de exigencia de pertenencia a la Clase *I*, es inferior al de la desviación poblacional. Cuando ésta es un 10 % menos exigente, (experimento **G**, Clase *I* = 1,1), el porcentaje de pertenencia se sitúa en valores próximos al 95 % con tamaños muestrales de 150 puntos. Dicho tamaño disminuye a medida que el nivel de exigencia se reduce: 50 puntos para un 20 % menos estricto y 20 puntos para un 30 % menos estricto.

Tabla 52.4 – Porcentajes de mapas pertenecientes a una determinada clase del estándar ASPRS en función del tamaño muestral y la desviación típica estimada a priori⁷.

Experimento	A	B	C	D	E	F	G	H	I
$\sigma_{\text{clase I}} = (m)$	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3
Muestra de tamaño n	% \in Clase II	% \in Clase I	% \in Clase I	% \in Clase I	% \in Clase I				
10	30,78	72,62	93,08	94,55	86,47	31,68	52,54	71,68	85,27
20	28,83	83,35	98,75	98,60	91,49	29,30	59,54	82,73	94,98
30	27,94	90,29	99,77	99,65	94,44	29,28	64,74	89,28	98,24
40	27,91	94,37	99,95	99,85	96,60	28,51	69,82	93,98	99,39
50	27,51	96,81	99,99	99,96	97,41	27,73	73,70	96,07	99,79
60	27,13	98,09	100	99,97	98,52	27,43	77,21	97,48	99,91
70	27,01	98,89	100	99,98	98,81	27,38	80,70	98,49	99,96
80	26,97	99,25	100	99,99	99,06	27,08	83,75	99,13	100
90	26,99	99,57	100	100	99,42	27,20	85,82	99,48	100
100	27,47	99,75	100	100	99,60	27,59	87,24	99,72	100
150	26,70	100	100	100	100	27,05	94,53	99,99	100
200	26,62					26,60	98,15		
250	26,79					26,69	99,2		
300	26,70					26,71	99,91		
350	25,94					26,74	100		
400	26,26					26,69	100		
450	26,58	100	100	100	100	26,62	100	100	100
500	27,03					26,97	100		
550	26,39					26,64	100		
600	26,99					26,97	100		
650	26,61					26,73	100		
700	26,25					26,23	100		
Resto muestras \in a la Clase:	III	III	III	I	I	II	II	II	II

Figura 52.6: ASPRS: porcentajes de aceptación en la Clase II en función del tamaño muestral y la desviación estimada a priori para dicha clase⁸.

⁷ La población original se posee una desviación de 1 m, mientras que el nivel de exigencia del estándar (desviación típica estimada a priori, $\sigma_{\text{clase I}}$) varía entre 0,5 y 1,3 m. En la última fila se indica la categoría en la que han sido clasificadas el resto de muestras.

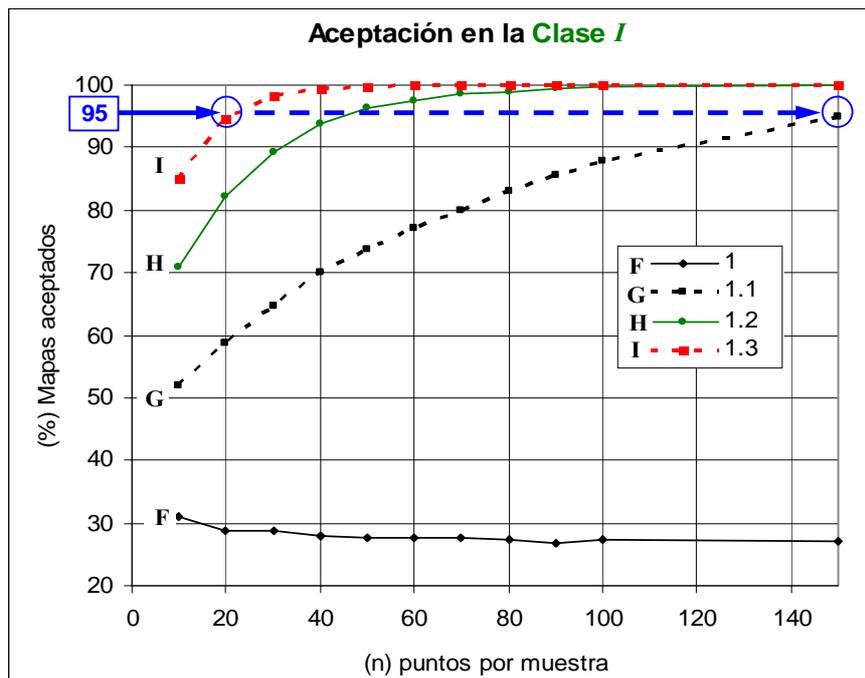


Figura 52.7: ASPRS: porcentajes de aceptación en la **Clase I** en función del tamaño muestral y la desviación estimada a priori para dicha clase⁹.

Con todo ello, se aprecia en el comportamiento del ASPRS, un nivel de exigencia de hasta un 30 % más que el de la población cuando se trabaja con tamaños muestrales del orden de los 20 puntos de control, mientras que si se aumenta a 150 puntos, dicho nivel de exigencia se reduce al 10%.

Realizando una interpolación entre los experimentos **H** y **G** para que el 95 % de las muestras pertenezcan a la Clase *I*, con un tamaño muestral de 100 puntos de control, la cartografía debería ser aproximadamente un 15 % más exigente.

Consideraciones sobre el ASPRS:

En este caso, el estándar establece categorías de mapas en función del *RMSE* detectado sobre una muestra de, al menos, 20 puntos de control. No obstante, el estándar exige al productor una cartografía de calidad muy superior a la inicialmente estimada:

⁸ El resto de muestras correspondientes a los experimentos **A**, **B** y **C**, \in **Clase III**, y en el caso de las series **D** y **E**, \in **Clase I**.

⁹ El resto de muestras de todas las series, **F**, **G**, **H** e **I**, \in **Clase II**. La línea en color azul indica las intersecciones con el 95 % de los mapas aceptados.

- Para un tamaño muestral de 20 puntos de control, con un 95 % de nivel de confianza, se aceptan los mapas como pertenecientes a la *Clase I* cuando éstos son un 30 % más exactos que la exigencia deseada para dicha clase.
- En el caso de emplear 40 puntos de control, la cartografía deberá ser un 20 % más exacta.
- Si se empleasen, al menos, 150 puntos de control, la cartografía debería ser un 10 % más exacta.

5.2.4 NSSDA

En este estándar se cambia completamente la filosofía con respecto a los anteriores: ya no se establece un valor que afirme si la cartografía ha sido aceptada o no, sino que se ofrece al usuario un valor estimativo del error posicional y será éste quien decida si debe o no emplear el mapa para determinada finalidad.

Partiendo de la opción de que $RMSE_x = RMSE_y$ (capítulo 3), se ha trabajado con el motor de simulación desarrollado (capítulo 4, apartado 4.1) teniendo en consideración las peculiaridades del estándar. Así, y puesto que en este caso el resultado de pasar el test no es la aceptación o rechazo de la muestra, sino que se ofrece un valor estimativo de la *Exactitud_r*, el resultado con 1000 muestras aleatorias sobre una determinada población ofrecerá, como estadísticos:

- valor medio del error, y
- desviación típica asociada al mismo.

Repitiendo el proceso sobre 100 poblaciones aleatorias, se obtiene la desviación típica entre poblaciones. Este factor proporciona al investigador una magnitud acerca de la estabilidad y fiabilidad de los resultados finales.

Al contrario de lo que sucedía en los procesos de simulación aplicados en estándares anteriores (NMA, EMAS...), en este caso la variabilidad detectada entre poblaciones es superior al 5 % para tamaños muestrales altos. Por ello, en base a los aspectos relacionados con el FCPF mencionados en el apartado 4.3.2, se ha optado por variar

el tamaño poblacional en función del tamaño muestral para aproximar el FCPF a la unidad. De esta forma:

- para un tamaño muestral entre 0 y 150, se han empleado poblaciones de 30.000 puntos,
- para un tamaño muestral entre 150 y 225, se han empleado poblaciones de 40.000 puntos, y
- para un tamaño muestral entre 225 y 800, se han empleado poblaciones de 100.000 puntos¹⁰.

Las variables introducidas en este caso han sido el número de puntos tomado en cada una de las muestras y el nivel de confianza exigido para el estándar. Partiendo de las poblaciones mencionadas, y con un 95 % de nivel de confianza, se han seleccionado muestras aleatorias de $n = 10, 15, 20... 800$ puntos. Los resultados obtenidos se presentan a modo de resumen en la tabla 52.5 y en las figuras 52.8 y 52.9.

Dado el elevado número de iteraciones, se ha alcanzado una gran robustez en los resultados finales, estabilizándose los porcentajes de variación obtenidos entre el 0.2 y 0.7 % (ver columna "e", tabla 52.5).

En primer lugar, y como aclaración previa, ha de mencionarse que siempre se trabaja con una BDG supuestamente buena: si se verifica la población total simulada, el valor obtenido para el estándar NSSDA es de 2,447 m (figura 52.8).

Tal y como se aprecia en la tabla 52.5, los valores se sitúan en $2,416 \pm 0,382$ m para muestras de 10 puntos. Esta variabilidad supone el 15,8 % del valor medio, por lo que los valores obtenidos no son aceptables. La horquilla se reduce considerablemente al aumentar el tamaño muestral, alcanzando valores de $\pm 1\%$ para muestras de tamaño próximo a 700 puntos. De esta forma, tomando 95 puntos de control, el valor obtenido por el NSSDA es de $2,443 \pm 0,121$ m. Dicha horquilla supone aproximadamente el $\pm 5\%$ ($n=93$) del valor medio mostrado por el test, por lo que si se desea trabajar a un 95 % de nivel de confianza, NO será aconsejable utilizar menos de 95-100 puntos de control por muestra. La estabilidad en los resultados obtenidos por simulación es del

¹⁰ Se ha efectuado un contraste de rachas, del tipo Wald-Wolfowitz, para comprobar que el generador de aleatorios no se agota.

0,4% (variabilidad alcanzada en 100 poblaciones), por lo que la horquilla final se situaría entre 4,6 y 5,4 %. Asimismo, para el tamaño muestral recomendado por el estándar, el valor medio de *Exactitud_r*, obtenido es de 2,434 m (lo que supone un 0,6 % inferior al valor poblacional de 2,447 m). Otro factor a tener en cuenta es el hecho de que el NSSDA presupone que se han eliminado los errores sistemáticos, hecho que no es posible contrastar al estudiar una cartografía real si previamente no se comprueba este tipo de errores (p.e. mediante un contraste del tipo *t-Student*).

Tabla 52.5 – NSSDA: valores promedio de *Exactitud_r*, y variabilidad en la simulación de muestras y poblaciones¹¹.

n	NSSDA	Desv.	Variación	Estab.	FCPF	n	NSSDA	Desv.	Variación	Estab.	FCPF
(a)	m	±m	±%	%	(f)	(a)	m	±m	±%	%	(f)
10	2,416	0,382	15,8	0,7	1.00	200	2,445	0,075	3,1	0,4	1.00
15	2,426	0,312	12,9	0,6	1.00	225	2,446	0,070	2,9	0,4	1.00
20	2,432	0,270	11,1	0,5	1.00	250	2,445	0,065	2,6	0,4	1.00
25	2,434	0,241	9,9	0,5	1.00	275	2,446	0,061	2,5	0,3	1.00
30	2,437	0,219	9,0	0,5	1.00	300	2,446	0,058	2,4	0,3	1.00
35	2,438	0,203	8,3	0,4	1.00	325	2,445	0,055	2,3	0,3	1.00
40	2,439	0,189	7,8	0,4	1.00	350	2,446	0,052	2,1	0,3	1.00
45	2,440	0,178	7,3	0,4	1.00	375	2,446	0,048	2,0	0,3	1.00
50	2,441	0,168	6,9	0,4	1.00	400	2,446	0,046	1,9	0,3	1.00
55	2,441	0,160	6,6	0,4	1.00	425	2,446	0,043	1,8	0,3	1.00
60	2,442	0,153	6,3	0,4	1.00	450	2,446	0,042	1,7	0,3	1.00
65	2,442	0,147	6,0	0,4	1.00	475	2,446	0,040	1,6	0,3	1.00
70	2,442	0,141	5,8	0,4	1.00	500	2,446	0,039	1,6	0,3	1.00
75	2,443	0,136	5,5	0,4	1.00	525	2,447	0,037	1,5	0,3	0.99
80	2,443	0,131	5,4	0,4	1.00	550	2,446	0,034	1,4	0,2	0.99
85	2,443	0,127	5,2	0,4	1.00	575	2,446	0,033	1,3	0,3	0.99
90	2,443	0,124	5,1	0,4	1.00	600	2,446	0,031	1,3	0,2	0.99
95	2,444	0,121	5,0	0,4	1.00	625	2,446	0,030	1,2	0,2	0.99
100	2,444	0,115	4,7	0,4	1.00	650	2,446	0,028	1,1	0,2	0.99
110	2,444	0,109	4,5	0,4	1.00	675	2,446	0,026	1,0	0,2	0.99
125	2,444	0,099	4,1	0,4	1.00	700	2,446	0,025	1,0	0,2	0.99
150	2,446	0,088	3,6	0,3	1.00	800	2,447	0,024	1,0	0,2	0.99
175	2,446	0,083	3,4	0,4	1.00						

¹¹ Las columnas representan: (a) tamaño de las 1000 muestras aleatorias; (b) valor medio de exactitud horizontal (*Exactitud_r*) tras aplicar el estándar NSSDA al 95% de nivel de confianza; (c) desviación del valor medio obtenido en el cálculo; (d) porcentaje que representa la desviación con respecto al valor medio (se presenta redondeado a 1 decimal aunque en la simulación el valor se ha ajustado con 3 decimales); (e) estabilidad alcanzada en la obtención de los resultados repitiendo el proceso con 100 poblaciones aleatorias $N(0,1)$, en base a la desviación de los valores medios y sus desviaciones; y (f) FCPF para $M = 30.000$ (cuando $10 \leq n < 150$), $M = 40.000$ ($150 \leq n < 225$) y $M = 100.000$ ($225 \leq n \leq 800$). En **negrita** se presentan los valores más significativos: resultados con 25 puntos de control y porcentajes de variación del $\pm 5\%$, $\pm 2,5\%$ y $\pm 1\%$.

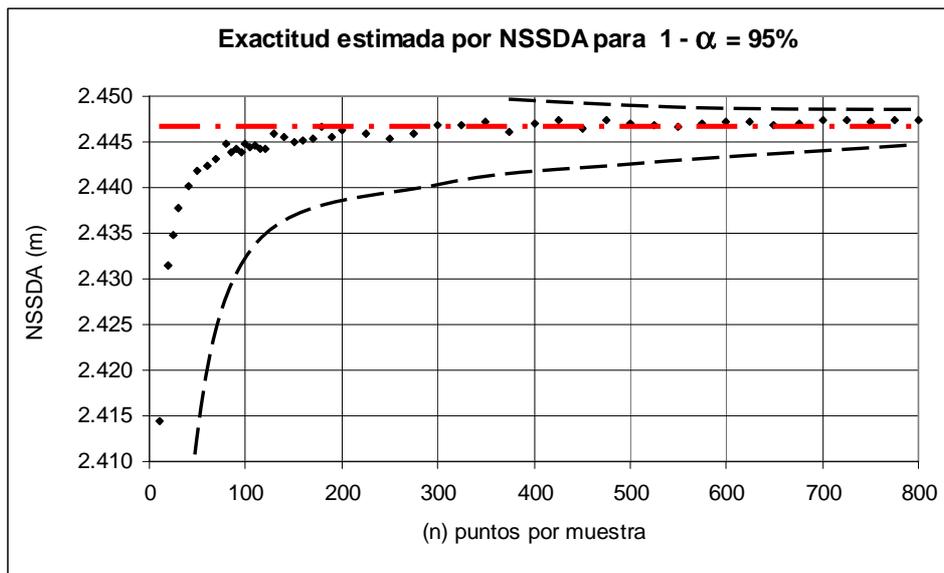


Figura 52.8: Valores medios obtenidos en el cálculo del estándar NSSDA en función del tamaño muestral¹².

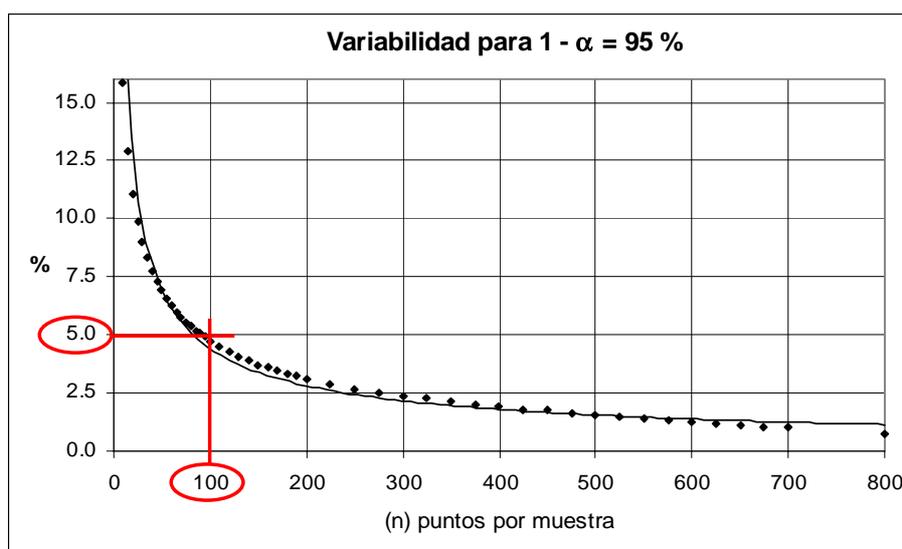


Figura 52.9: Variabilidad obtenida en el cálculo del estándar NSSDA en función del tamaño muestral¹³.

Siguiendo la misma metodología, se obtienen las tablas 52.6 a 52.8 en las que se presenta el tamaño muestral aconsejable para diferentes niveles de confianza en la

¹² Los valores medios aparecen representados por puntos, la línea horizontal discontinua (en color rojo) representa el valor al que ha de tender (*Exactitud*, poblacional = 2,447 m), y el trazo discontinuo muestra su variabilidad en metros escalada respecto a la *Exactitud*, poblacional $\times (1/10)$.

¹³ Los datos poblacionales y muestrales son los mismos que en la figura 52.8. Junto con los datos discretos, se ha ajustado la curva de regresión $\% = 88,466 \cdot n^{-0,6508}$ (siendo $r^2 = 0,9779$).

aplicación del estándar NSSDA. El propio estadístico¹⁴ varía dependiendo del nivel de confianza aplicable, por lo que ha sido modificado en función del índice de variabilidad buscado. Así, se obtiene el tamaño muestral para un determinado nivel de probabilidad con el que verificar la exactitud horizontal.

Tabla 52.6 – Valor promedio de $Exactitud_r$ para un 90 % de nivel de confianza en el NSSDA ($Exactitud_r$ poblacional = 2,415).

n	NSSDA	Desv.	Variación	Estab.	FCPF
(a)	m	$\pm m$	$\pm\%$	%	(f)
(a)	(b)	(c)	(d)	(e)	(f)
10	2.118	0.336	15.9	0.4	1.00
15	2.128	0.275	12.9	0.4	1.00
20	2.132	0.238	11.2	0.3	1.00
24	2.135	0.218	10.2	0.3	1.00
25	2.135	0.214	10.0	0.3	1.00
26	2.136	0.210	9.8	0.3	1.00
30	2.137	0.187	8.8	0.3	1.00
35	2.138	0.180	8.4	0.3	1.00

Tabla 52.7 – Valor promedio de $Exactitud_r$ para un 97,5 % de nivel de confianza en el NSSDA ($Exactitud_r$ poblacional = 2,715).

NSSDA al 97,5% = 2,715 m					
n	NSSDA	Desv.	Variación	Estab.	FCPF
(a)	m	$\pm m$	$\pm\%$	%	(f)
(a)	(b)	(c)	(d)	(e)	(f)
350	2.716	0.072	2.6	0.4	1.00
360	2.716	0.072	2.6	0.4	1.00
370	2.719	0.072	2.6	0.4	1.00
380	2.713	0.068	2.5	0.4	1.00
390	2.720	0.068	2.5	0.4	1.00
400	2.712	0.067	2.5	0.3	1.00
410	2.715	0.067	2.5	0.3	1.00
420	2.714	0.066	2.4	0.3	1.00

Tabla 52.8 – Valor promedio de $Exactitud_r$ para un 99 % de nivel de confianza en el NSSDA ($Exactitud_r$ poblacional = 3,03).

n	NSSDA	Desv.	Variación	Estab.	FCPF
(a)	m	$\pm m$	$\pm\%$	%	(f)
(a)	(b)	(c)	(d)	(e)	(f)
2000	3.039	0.035	1.1	0.2	0.98
2100	3.037	0.032	1.1	0.2	0.98
2200	3.035	0.033	1.1	0.2	0.98
2300	3.035	0.032	1.0	0.2	0.98
2400	3.035	0.030	1,0	0.2	0.98
2500	3.034	0.029	0.9	0.2	0.98
2600	3.034	0.031	0.9	0.2	0.97
2700	3.039	0.027	0.9	0.2	0.97

Si se desea aplicar el estándar NSSDA para el control de calidad posicional cartográfico, deberán tenerse en cuenta los siguientes aspectos:

¹⁴ El estadístico aplicado en el NSSDA (ver apartado 3.2.6) se obtiene a partir de la función de densidad de la distribución circular normal.

- El estándar muestra un valor estimativo de la exactitud posicional del mapa al 95 % de nivel de confianza, con un tamaño muestral mínimo de 20 puntos. Sin embargo, dicho tamaño muestral supone una variabilidad en su aplicación superior al 11 %, así como una subestimación de la *Exactitud*, poblacional del 0,6 %.
- Al afirmar que el mapa analizado posee una determinada exactitud al 95 % de nivel de confianza, deberán emplearse, al menos, 95 puntos de control (NSSDA = 2,444 ±0,121 m).
- Con un tamaño muestral inferior, no es posible afirmar que la exactitud posicional horizontal alcanzada está verificada al 95 % de nivel de confianza, si no que vendrá dada en función de las tablas 52.5 a 52.8, y que pueden resumirse en:
 - 25 puntos para un 90 % de nivel de confianza.
 - 100 puntos para un 95 % de nivel de confianza.
 - 380 puntos para un 97.5 % de nivel de confianza.
 - 2400 puntos para un 99 % de nivel de confianza.

Con todo ello, se observan ciertas peculiaridades que no aparecen mencionadas en la documentación del propio estándar, como es el tamaño muestral frente al nivel de confianza del análisis.

Otro aspecto relevante es que, dado que el valor medio proporcionado por el NSSDA para una muestra es, en general, inferior al valor poblacional, el estándar subestima el error de la BDG. Así, cuando un usuario utilice un número determinado de puntos de control ($n_1, n_2, n_3...n_m$), correrá un determinado riesgo de subestimar el error real del mapa (figura 52.10). Por encima del valor estimado para la población (NSSDA_{Pobl}) se representa el riesgo del productor de que el resultado del estándar proporcione un valor superior al poblacional, mientras que por debajo es el riesgo del usuario a que dicho valor sea inferior.

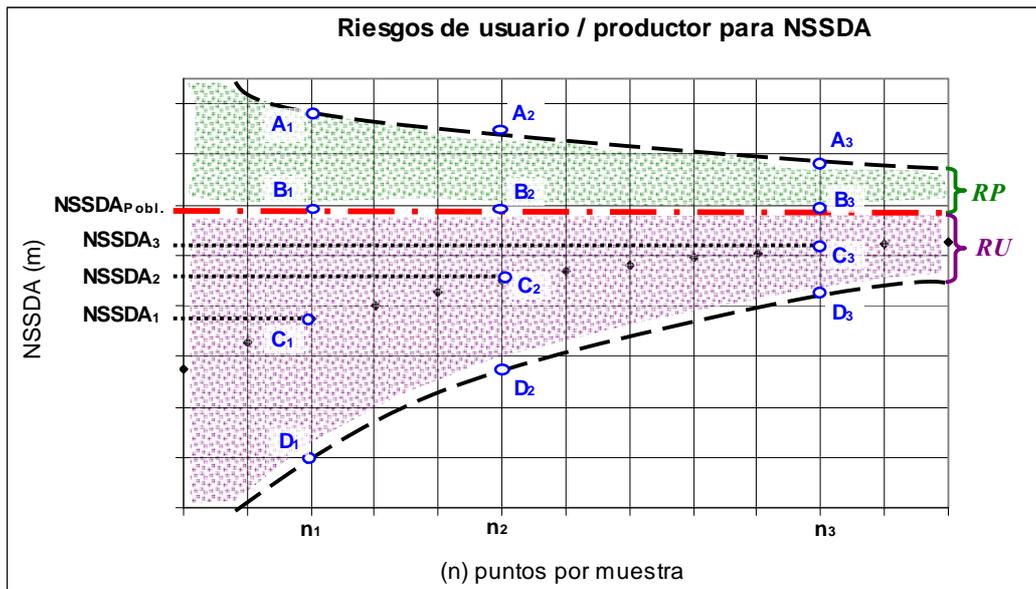


Figura 52.10: Riesgo de usuario y productor al determinar la *Exactitud_r*, mediante el estándar NSSDA¹⁵.

Para obtener los riesgos máximos, la desviación obtenida en la columna “c” de las tablas 52.5 a 52.8, deberá incrementarse en $\pm k \cdot \sigma = \pm 1,96 \cdot \sigma$ (probabilidad al 95 % de nivel de confianza). Si se deseara trabajar bajo estas condiciones, para que el índice de variabilidad fuese de ± 5 % del valor medio, se debería trabajar con 275 puntos de control:

$$Exactitud_r \pm k \cdot \sigma \rightarrow 2,446 \text{ m} \pm 1,96 \cdot 0,061 \text{ m} = 2,446 \text{ m} \pm 0,120 \text{ m}$$

Con un incremento tan considerable en el número de puntos de control, se ha pasado de los 20 recomendados por el estándar a 275 (un incremento del 1375%), en contraste de la exactitud posicional se hace difícil de realizar en la práctica al aumentar los costes de forma exponencial.

En cuanto a los riesgos que deberán asumir tanto productor como usuario, éstos serán:

$$RU(n_1) = \frac{B_1 D_1}{A_1 D_1}; RU(n_2) = \frac{B_2 D_2}{A_2 D_2} \dots = RU(n_m) = \frac{B_m D_m}{A_m D_m} \xrightarrow{\text{Aprox.}} RU^*(n_m) = \frac{C_m D_m}{A_m D_m}$$

$$RP(n_1) = \frac{A_1 B_1}{A_1 D_1}; RP(n_2) = \frac{A_2 B_2}{A_2 D_2} \dots = RP(n_m) = \frac{A_m B_m}{A_m D_m} \xrightarrow{\text{Aprox.}} RP^*(n_m) = \frac{A_m C_m}{A_m D_m}$$

¹⁵ Los puntos A₁, A₂, A₃ representan el valor superior de la variabilidad para un determinado tamaño muestral; B₁, B₂, B₃ son el valor constante de NSSDA_{poblacional}; C₁, C₂, C₃ muestran mediante simulación del NSSDA el valor promedio estimado en función de n; y D₁, D₂, D₃ representan el valor inferior de la variabilidad.

siendo $RU(n_m)$ y $RP(n_m)$ los riesgos del usuario y productor para un número n_m de puntos de control. Cuando el valor $NSSDA_{Pobl.}(valor B_m)$ es desconocido, la utilización como aproximación del valor medio obtenido en la simulación de muestras (C_m), aumenta la probabilidad de infravalorar el error real de la BGD. Dado que siempre el valor $B_m > C_m$, los riesgos aproximados de usuario (RU^*) y productor (RP^*) son:

$$\begin{aligned} RU^*(n_m) &< RU(n_m) \\ RP^*(n_m) &> RP(n_m) \end{aligned}$$

Como los puntos A_m y D_m son simétricos respecto al valor medio estimado por simulación para el NSSDA (C_m), siempre se da la circunstancia de que:

$$RU^*(n_m) = RP^*(n_m) = 0,50$$

Para conocer en qué magnitud se infravaloran o sobreestiman los riesgos, a partir de la tabla 52.5 e incrementando la desviación obtenida en $\pm 1,96 \cdot \sigma$, se obtiene la tabla 52.10 con los riesgos reales de usuario y productor en función del tamaño muestral. La diferencia existente con respecto a los valores estimados ($RU^* = RP^* = 0,50$), será lo que se infravaloren o sobreestimen los riesgos a la hora de obtener la *Exactitud*.

Como puede observarse, el estándar NSSDA representa un riesgo, tanto para el usuario como para el productor, muy similar y próximo al 50 % (independientemente del tamaño muestral). Se aprecia una muy ligera tendencia de un mayor riesgo para los usuarios y menor para los productores, aunque, lógicamente, ambos tienden a equipararse al aumentar el número de puntos de control. Tanto los riesgos de tipo α como β son del 50 % con una probabilidad del 95 % de nivel de confianza ($\pm k \cdot \sigma = \pm 1,96 \cdot \sigma$).

Tabla 52.10 – NSSDA: riesgo del usuario (*RU*) y del productor (*RP*), en función del tamaño muestral, obtenidos a partir de $NSSDA_{Pobl}$

<i>n</i>	Riesgos		<i>n</i>	Riesgos		<i>n</i>	Riesgos	
	<i>RU</i>	<i>RP</i>		<i>RU</i>	<i>RP</i>		<i>RU</i>	<i>RP</i>
10	0.521	0.479	90	0.508	0.492	400	0.506	0.494
15	0.517	0.483	95	0.506	0.494	425	0.506	0.494
20	0.514	0.486	100	0.507	0.493	450	0.506	0.494
25	0.514	0.486	110	0.507	0.493	475	0.506	0.494
30	0.512	0.488	125	0.508	0.492	500	0.507	0.493
35	0.511	0.489	150	0.503	0.497	525	0.507	0.493
40	0.511	0.489	175	0.503	0.497	550	0.508	0.492
45	0.510	0.490	200	0.507	0.493	575	0.508	0.492
50	0.509	0.491	225	0.504	0.496	600	0.508	0.492
55	0.510	0.490	250	0.508	0.492	625	0.509	0.491
60	0.508	0.492	275	0.506	0.494	650	0.509	0.491
65	0.509	0.491	300	0.504	0.496	675	0.510	0.490
70	0.509	0.491	325	0.509	0.491	700	0.510	0.490
75	0.508	0.492	350	0.505	0.495	800	0.500	0.500
80	0.508	0.492	375	0.505	0.495			

Consideraciones sobre el NSSDA:

El estándar tan solo se limita a mostrar un valor que cuantifica la exactitud al 95% de nivel de confianza. Este estándar, que es el que actualmente han de emplear la Agencias Federales de los EE.UU. que trabajen con información georreferenciada (FGDC, 1998), a pesar de su actualidad tiene algunas deficiencias:

- Se presupone que se han eliminado los sistematismos y que los errores se distribuyen de forma normal y aleatoria. Para ello no muestra al usuario una herramienta de cómo poder realizar este tipo de contrastes.
- El estándar ofrece un valor medio en el que subestima la exactitud horizontal real de la población analizada.
- El tamaño muestral recomendado por el estándar para presentar el valor de exactitud al 95 % de nivel de confianza es de, al menos, 20 puntos de control. No obstante, con dicho tamaño muestral, la variabilidad que presenta es $\pm 11\%$, por lo que en realidad se obtiene un valor de exactitud al 89 % de nivel de confianza. Si se desea que dicha variabilidad se reduzca a $\pm 5\%$, han de utilizarse al menos 100 puntos de control. Asimismo, para que la probabilidad de riesgo al 95 % de nivel de confianza sea máxima ($\pm k \cdot \sigma = \pm 1,96 \cdot \sigma$), el tamaño muestral no ha de ser inferior a 275 puntos.

- Los riesgos de usuario / productor se sitúan en valores próximos al 50 %, independientemente del tamaño muestral, equiparándose a medida que éste aumenta. Lógicamente, tan sólo existe un leve mayor riesgo para el usuario que para el productor ($\beta > \alpha$). Todo ello hace que la utilización del NSSDA se sitúe en el Punto de Riesgo Compartido (ver capítulo 3, apartado 3.6), al tener ambos riesgos muy similares, por lo que resulta ser un importante aspecto de valor añadido en el estándar.

5.3 Corrección mediante ajuste mínimo cuadrático de errores sistemáticos

En este caso, se emplearon los procesos de simulación descritos en el apartado 43.4 (ver figura 43.7). Para ello, se ha analizado una población sin ningún tipo de sistematismo, únicamente con la contaminación aleatoria intrínseca de una cartografía a escala 1/4.000 ($\sigma = 1\text{m}$).

Tabla 53.1: Cuantificación de errores sistemáticos mediante ajuste por mm.cc.¹⁶

¹⁶ Simulación realizada sobre una población de 1000 puntos con errores aleatorios según una $N(0,1)$. Sobre cada experimento (filas A – N) se ha introducido una determinada contaminación (primera fila en color rojo de cada experimento, columnas T, U, V)

Población original: distribución N(0,1)																
Experimento	Valores obtenidos sobre la población contaminada sin corrección por mm.cc.				Detección errores aplicando EMAS				Contaminación y resultados en la cuantificación de los errores sistemáticos	Contaminación por errores sistemáticos y cuantificación mediante ajuste obtenido por m.cc.					Desviación obtenida en el ajuste	
	Media		Desv.		Sistem.		Aleat.			Traslación		Factor escala		Giro	V_{x_0, y_0}	$V\lambda\alpha$
	μ_x	μ_y	σ_x	σ_y	S_x	S_y	C_x	C_y		X_0	Y_0	λ_x	λ_y	α		
A	0	0	1	1					Contaminación →	0	0	1	1	0	0,127	0,003
	Tras el ajuste →									0,137	0,478	0,991	0,988	-0,296		
B	2	0,5	1	1	X	X			Contaminación →	2	0,5	1	1	0	0,121	0,003
	Tras el ajuste →									2,120	0,958	0,991	0,988	-0,296		
C	1	3,5	1	1	X	X			Contaminación →	1	3,5	1	1	0	0,117	0,003
	Tras el ajuste →									1,141	3,928	0,991	0,988	-0,296		
D	-2,75	2,33	1,47	1,24	X	X	X	X	Contaminación →	0	0	0,9	1,1	0	0,127	0,003
	Tras el ajuste →									0,137	0,478	0,892	1,086	-0,296		
E	4,15	7,32	1,73	2,75	X	X	X	X	Contaminación →	0	0	1,2	1,4	0	0,127	0,004
	Tras el ajuste →									0,137	0,478	1,189	1,383	-0,296		
F	-0,33	0,33	1	1	X	X			Contaminación →	0	0	1	1	0,75	0,127	0,003
	Tras el ajuste →									0,137	0,478	0,991	0,988	0,454		
G	-0,88	0,88	1,04	1,03	X	X			Contaminación →	0	0	1	1	2	0,127	0,003
	Tras el ajuste →									0,137	0,478	0,991	0,988	1,704		
H	2,15	1,84	1,26	1,55	X	X	X	X	Contaminación →	5	-3	0,95	1,15	3,5	0,124	0,004
	Tras el ajuste →									5,078	-2,517	0,941	1,136	3,204		
I	1,83	-1,25	1,02	1,18	X	X		X	Contaminación →	2,5	-0,5	1	0,95	1,5	0,123	0,003
	Tras el ajuste →									2,611	-0,032	0,991	0,938	1,204		
J	2,34	0,58	1,02	1,01	X	X			Contaminación →	3	0	1	1	1,5	0,120	0,003
	Tras el ajuste →									3,109	0,458	0,991	0,988	1,204		
K	3,63	0	1,24	1	X		X		Contaminación →	1,5	0	1,1	1	0	0,124	0,004
	Tras el ajuste →									1,623	0,468	1,090	0,988	-0,296		
L	1	1	1	1	X	X			Contaminación →	1	1	1	1	0	0,123	0,003
	Tras el ajuste →									1,131	1,459	0,991	0,988	-0,296		
M	0,51	0,27	1,03	1,05	X	X			Contaminación →	1	-1	0,98	1,05	0	0,127	0,003
	Tras el ajuste →									1,124	-0,517	0,971	1,037	-0,296		
N	0,28	2,38	1,04	1,05	X	X			Contaminación →	1,2	0,8	0,98	1,05	1	0,123	0,003
	Tras el ajuste →									1,329	1,260	0,971	1,037	0,704		
	P	Q	R	S						T	U	V	W	X		

En este caso, el estándar EMAS no detecta ningún error sistemático ni aleatorio. Para estudiar el comportamiento de la transformación afín, se procedió a su ajuste por mm.cc., obteniendo los resultados que aparecen en la tabla 53.1 (experimento A). Para cada experimento (filas de la A – N) se indica en la margen izquierda de la tabla los valores obtenidos antes del ajuste para la media y la desviación típica (columnas P,

y se ha estudiado la detección de sistematismos y aleatoriedad mediante el estándar EMAS (columnas R y S). Los resultados tras cuantificar los sistematismos por mm.cc. aparecen en azul (segunda fila de cada experimento, columnas T, U, V), y la desviación obtenida en el ajuste se presenta en las dos últimas columnas (W y X).

Q), así como si se ha detectado o no algún tipo de error sistemático (S_x, S_y) o aleatorio (C_x, C_y), marcándolo con una "X" (columnas R, S).

En la margen derecha, en la primera fila de cada experimento, aparece en color rojo la magnitud de la contaminación sistemática ("Contaminación \rightarrow ") que se ha aplicado tanto en la traslación (X_0, Y_0), los factores de escala (λ_x, λ_y) como el giro (α) (columnas T, U, V). Bajo dichos valores, en la segunda fila de cada experimento (en color azul), aparece el resultado del ajuste por mm.cc. con la cuantificación que el sistema ha conseguido detectar. Por último, en las dos últimas columnas (W, X), se muestra la desviación alcanzada para la traslación ($V_{x_0 y_0}$) y el escalado y giro ($V\lambda\alpha$)⁽¹⁷⁾.

El hecho de presentar la contaminación en la primera fila de cada experimento y en la segunda el resultado obtenido por mm.cc. (columnas T, U, V), facilita la visualización de cuánto se ha aproximado el ajuste a la contaminación real.

En toda la simulación, tras validar y realizar el ajuste por mm.cc., siempre se han obtenido valores muy similares a los buscados en la población original (para la media y desviación típica).

Volviendo al experimento inicial, fila "A" en la tabla 53.1, a pesar de no existir ningún sistematismo, el ajuste detecta una traslación en Y de 0,47m y un giro de $-0,3^\circ$. Ello se debe a los propios errores aleatorios de una $N(0,1)$. Como se puede apreciar, más claramente en la traslación, ésta es inferior a la desviación propia del mapa (1m), por lo que sería despreciable. No obstante, estos falsos errores sistemáticos, se mantendrán constantes a lo largo del proceso de simulación, alterando el resultado final.

Para conocer la influencia de dicho error sobre los valores del ajuste final, se realizó un proceso de simulación con más de 100 poblaciones contaminadas con una $N(0,1)$, obteniéndose unos valores máximos en la distorsión del ajuste de 0,65m en traslación, 0,97 como factor de escala y $0,3^\circ$ en el giro. En cuanto a las desviaciones del ajuste

¹⁷ Dado que los valores de desviación para los dos factores de escala y el giro son muy similares, en la tabla aparecen promediados, al igual que sucede con la traslación en X e Y.

por mm.cc., éstas no han superado el 20% de la desviación del mapa en el caso de las traslaciones y 0,0065 en el caso de los factores de escala y giros.

Así, se puede afirmar que las precisiones alcanzadas en la cuantificación y tipificación de los errores sistemáticos es:

- En las **traslaciones**, son detectadas con una precisión mejor que $\pm\sigma$ m (generalmente en torno a $\pm 0,65\cdot\sigma$).
- En los **factores de escala**, al ser adimensionales, se puede afirmar que alcanzan una precisión aceptable hasta el redondeo al primer o segundo decimal.
- Y en cuanto al **giro**, éste podría ser válido redondeando el valor obtenido al grado sexagesimal.

Tras esta primera parte del experimento, en la tabla 53.1 se aprecian los buenos resultados obtenidos con diferentes niveles de contaminación por errores sistemáticos. En los experimentos B y C se ha alterado únicamente mediante una traslación, obteniéndose buenos resultados siempre que los sistematismos sean inferiores a σ . En los experimentos D y E se ha alterado sólo el factor de escala, obteniendo también resultados muy favorables. En los experimentos F y G, se aprecia cómo la metodología es correcta en giros de 1° o superiores. A continuación, se ha procedido a contaminar con varios tipos de errores sistemáticos de forma conjunta y en diferente magnitud, con lo que se viene a confirmar los valores de precisión mencionados anteriormente. Si se atiende a las desviaciones del ajuste, éstas se encuentran dentro de los márgenes mencionados.

En siguiente lugar, se ha trabajado con distintos niveles de contaminación aleatoria (siempre bajo el supuesto de que este tipo de errores se distribuye de forma normal). Los resultados, a diferente nivel de contaminación sistemática, aparecen reflejados en las tablas 53.2, 53.3 y 53.4. En ellas se ha trabajado con contaminación aleatoria en desviación de 0,125m, 0,50m y 2,00m (que se corresponderían a las escalas 1/500, 1/2.000 y 1/8.000 respectivamente).

Tabla 53.2: Cuantificación de errores sistemáticos mediante simulación según una distribución $N(0,0.125)$.

Población original: distribución $N(0, 0.125)$																
Experimento	Valores obtenidos sobre la población contaminada sin corrección por mm.cc				Detección errores aplicando EMAS				Contaminación y resultados en la cuantificación de los errores sistemáticos	Contaminación por errores sistemáticos y cuantificación mediante ajuste obtenido por m.cc.					Desviación obtenida en el ajuste	
	Media		Desv		Sistem.		Aleat.			Traslación		Factor escala		Giro	$V_{x_0y_0}$	$V_{\lambda\alpha}$
	μ_x	μ_y	σ_x	σ_y	S_x	S_y	C_x	C_y		X_0	Y_0	λ_x	λ_y	α		
A	0	0	0,125	0,125					Contaminación →	0	0	1	1	0	0,116	0,005
									Tras el ajuste →	0,092	0,250	0,999	0,983	0,097		
B	0,25	0,06	0,125	0,125	X	X			Contaminación →	0,25	0,06	1	1	0	0,116	0,005
									Tras el ajuste →	0,342	0,308	0,999	0,983	0,097		
C	0,125	0,4	0,125	0,125	X	X			Contaminación →	0,125	0,4	1	1	0	0,113	0,005
									Tras el ajuste →	0,215	0,643	0,999	0,983	0,097		
D	-2,76	1,04	1,02	0,15	X	X	X	X	Contaminación →	0	0	0,9	1,1	0	0,116	0,006
									Tras el ajuste →	0,092	0,250	0,899	1,081	0,097		
E	4,15	3,29	1,5	0,33	X	X	X	X	Contaminación →	0	0	1,2	1,4	0	0,116	0,008
									Tras el ajuste →	0,092	0,250	1,199	1,376	0,097		
F	-0,15	0,33	0,125	0,16	X	X		X	Contaminación →	0	0	1	1	0,75	0,116	0,005
									Tras el ajuste →	0,092	0,250	0,999	0,983	0,847		
G	-0,39	0,87	0,13	0,32	X	X		X	Contaminación →	0	0	1	1	2	0,116	0,005
									Tras el ajuste →	0,092	0,250	0,999	0,983	2,097		
H	-1,25	2,37	0,49	0,49	X	X	X	X	Contaminación →	0,75	-0,5	0,95	1,15	3,5	0,121	0,006
									Tras el ajuste →	0,844	-0,244	0,948	1,131	3,597		
I	-0,59	0,35	0,13	0,27	X	X		X	Contaminación →	-0,3	0,25	1	0,95	1,5	0,114	0,005
									Tras el ajuste →	-0,209	0,496	0,999	0,934	1,597		
J	0,8	-0,81	0,127	0,25	X	X		X	Contaminación →	1	0	1	0,9	1	0,116	0,005
									Tras el ajuste →	1,091	0,247	0,999	0,885	1,097		
	P	Q	R	S						T	U	V	W	X		

Tabla 55.3: Cuantificación de errores sistemáticos mediante simulación de una N(0,0.50).

Población original: distribución N (0 , 0.50)																
Experimento	Valores obtenidos sobre la población contaminada sin corrección por mm.cc				Detección errores aplicando EMAS				Contaminación y resultados en la cuantificación de los errores sistemáticos	Contaminación por errores sistemáticos y cuantificación mediante ajuste obtenido por m.cc.					Desviación obtenida en el ajuste	
	Media		Desv		Sistem.		Aleat.			Traslación		Factor escala		Giro	V_{x_0, y_0}	$V\lambda\alpha$
	μ_x	μ_y	σ_x	σ_y	S_x	S_y	C_x	C_y		X_0	Y_0	λ_x	λ_y	α		
A	0	0	0,5	0,5					Contaminación → Tras el ajuste →	0 0,001	0 0,065	1 0,998	1 0,998	0 -0,073	0,064	0,002
B	1	0	0,5	0,5	X				Contaminación → Tras el ajuste →	1 0,999	0 0,064	1 0,998	1 0,998	0 -0,073	0,063	0,002
C	0,5	3	0,5	0,5	X	X			Contaminación → Tras el ajuste →	0,5 0,505	3 3,059	1 0,998	1 0,998	0 -0,073	0,060	0,002
D	2,7	2,3	1,1	1	X	X	X	X	Contaminación → Tras el ajuste →	0 0,001	0 0,065	0,9 0,898	1,1 1,098	0 -0,073	0,064	0,002
E	0	0	1,5	2,7	X	X	X	X	Contaminación → Tras el ajuste →	0 0,001	0 0,065	1,2 1,198	1,4 1,397	0 -0,073	0,064	0,002
F	0,33	0,33	0,5	0,5	X	X			Contaminación → Tras el ajuste →	0 0,001	0 0,065	1 0,998	1 0,998	0,75 0,677	0,064	0,002
G	0,8	0,8	0,6	0,6	X	X	X	X	Contaminación → Tras el ajuste →	0 0,001	0 0,065	1 0,998	1 0,998	2 1,927	0,064	0,002
H	0,35	1,38	0,8	1,4	X	X	X	X	Contaminación → Tras el ajuste →	2,5 2,494	-1,2 -1,135	0,95 0,948	1,15 1,148	3,5 3,427	0,063	0,002
	P	Q	R	S						T	U	V	W	X		

Tabla 55.4: Cuantificación de errores sistemáticos mediante simulación de una N(0,2).

Población original: distribución N (0 , 2)																
Experimento	Valores obtenidos sobre la población contaminada sin corrección por mm.cc				Detección errores aplicando EMAS				Contaminación y resultados en la cuantificación de los errores sistemáticos	Contaminación por errores sistemáticos y cuantificación mediante ajuste obtenido por m.cc.					Desviación obtenida en el ajuste	
	Media		Desv		Sistem.		Aleat.			Traslación		Factor escala		Giro	V_{x_0, y_0}	$V\lambda\alpha$
	μ_x	μ_y	σ_x	σ_y	S_x	S_y	C_x	C_y		X_0	Y_0	λ_x	λ_y	α		
A	0	0	2	2					Contaminación → Tras el ajuste →	0 1,079	0 1,194	1 0,958	1 0,957	0 -0,087	0,248	0,007
B	4	0,5	2	2	X	X			Contaminación → Tras el ajuste →	4 4,910	0,5 1,656	1 0,958	1 0,957	0 -0,087	0,227	0,007
C	1	6	2	2	X	X			Contaminación → Tras el ajuste →	1 2,029	6 6,934	1 0,958	1 0,957	0 -0,087	0,217	0,007
D	2,7	2,3	2,4	2	X	X	X		Contaminación → Tras el ajuste →	0 1,079	0 1,194	0,9 0,862	1,1 1,053	0 -0,087	0,248	0,007
E	4,1	7,3	2,2	3	X	X	X	X	Contaminación → Tras el ajuste →	0 1,079	0 1,194	1,2 1,149	1,4 1,340	0 -0,087	0,248	0,009
F	0,3	0,3	2	2	X	X			Contaminación → Tras el ajuste →	0 1,079	0 1,194	1 0,958	1 0,957	0,75 0,663	0,248	0,007
G	0,9	0,9	2	2	X	X			Contaminación → Tras el ajuste →	0 1,079	0 1,194	1 0,958	1 0,958	2 1,913	0,248	0,007
H	2,1	1,8	2,2	2,2	X	X	X	X	Contaminación → Tras el ajuste →	5 5,872	-3 -1,699	0,95 0,910	1,15 1,101	3,5 3,413	0,240	0,007
	P	Q	R	S						T	U	V	W	X		

Como se puede apreciar, la metodología sirve para clasificar y detectar correctamente los sistematismos. Es más, cabe destacar la equivocación que sufre el estándar EMAS al confundir errores sistemáticos con aleatorios. Esto sucede cuando el sistematismo afecta a los factores de escala o giro (experimentos D, E y H en las tablas 55.2, 3 y 4). Ello es normal debido a que este tipo de sistematismo afecta tanto a los valores promedio como a las desviaciones típicas de la población.

Respecto al nivel de contaminación aleatoria y el tamaño de la muestra con el que realizar el ajuste por mm.cc., al aumentar la desviación típica de los errores aleatorios, el parámetro encargado de cuantificar la bondad del ajuste (los residuos) que debía mantenerse próximo a cero, aumenta considerablemente. Esto conlleva la invalidación del resultado en los parámetros de la transformación afín.

Por el contrario, si se disminuye la desviación típica de los errores aleatorios, también se puede reducir el número de puntos de control con el que realizar el ajuste, obteniéndose valores admisibles para los residuos.

Después de numerosos ensayos, se definió la gráfica de la figura 53.1, en la que se muestra la relación entre el tamaño de la muestra y la desviación típica estimada para el mapa. Para aquellos puntos que se encuentren en la parte inferior de la gráfica, los resultados obtenidos en el ajuste por mm.cc. no serán admisibles dado que sus residuos son elevados. Por el contrario, aquellos puntos que se encuentren en la parte superior de la gráfica ofrecerán resultados óptimos en el ajuste por mm.cc.

Como aproximación, se podría afirmar que deberán emplearse, al menos, un número de puntos de control no inferior al proporcionado por la curva de ajuste de la figura 53.1, y que queda definida mediante la ecuación:

$$n = -150,02 \cdot \sigma^2 + 764,81 \cdot \sigma + 30,794$$

siendo:

- n : número mínimo de puntos de control recomendado
- σ : valor estimado a priori como la desviación típica del mapa (en metros)

Así, para poder cuantificar la magnitud de uno o varios errores sistemáticos que hayan sido detectados previamente (por ejemplo mediante un contraste del tipo *t-Student*), se

deberá tener en cuenta la escala del mapa (o exactitud intrínseca del mismo) y el número admisible de puntos de control.

Por ejemplo, si se ha detectado un sistematismo en una cartografía a escala 1/500 ($\sigma = 0,125\text{m}$), deberán tomarse al menos 125 puntos de control perfectamente definidos y homogéneamente distribuidos. Si, por el contrario, se trata de una cartografía a escala 1/1000 ($\sigma = 0,25\text{m}$), deberán ser ≈ 225 los puntos a emplear. En el caso de que la escala sea menor, conviene plantearse la viabilidad económica a la hora de capturar un mayor número de puntos de control, dado que la relación entre el número de puntos de control y la escala sería:

- Escala 1/2.000 ($\sigma = 0,50\text{m}$) \rightarrow 380 puntos de control
- Escala 1/5.000 ($\sigma = 1,25\text{m}$) \rightarrow 760 puntos de control
- Escala 1/10.000 ($\sigma = 2,50\text{m}$) \rightarrow 1.010 puntos de control
-

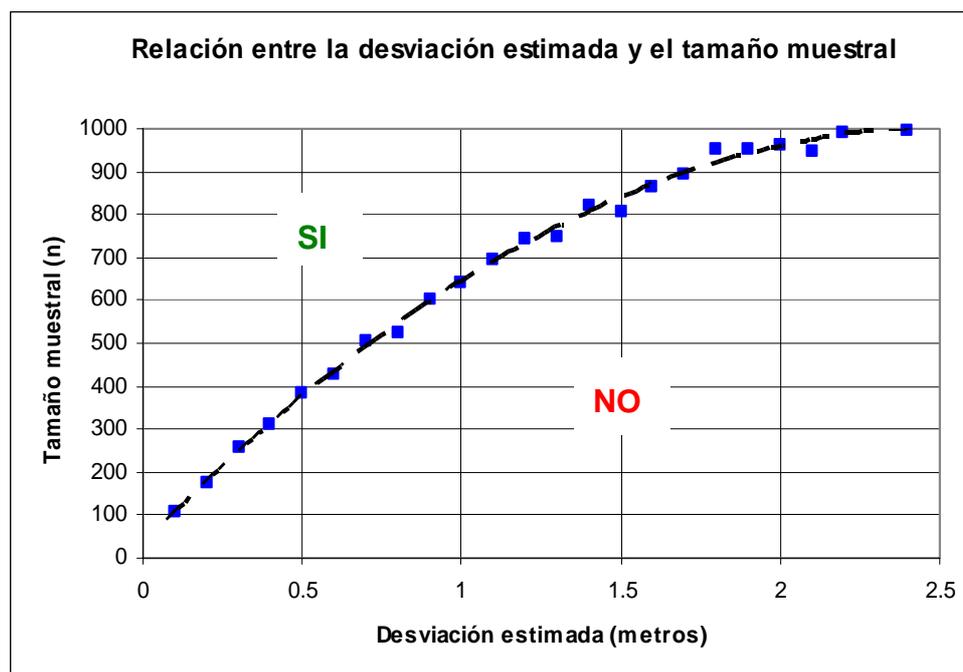


Figura 53.1: Relación entre la desviación estimada y el tamaño muestral para la cuantificación de sistematismos por mm.cc.¹⁸

¹⁸ En la curva se determina de la zona de validez (por encima de ella) para la aplicación del ajuste por mm.cc. de una transformación afín en la cuantificación del error sistemático. El gráfico se define en función del tamaño muestral y la desviación típica estimada para el mapa.

Para aquellos casos en los que el número de puntos de control haga económicamente inviable el empleo de los mm.cc., tratarán de corregirse los sistematismos en traslación mediante la corrección por el error medio sobre las coordenadas de cada componente:

$$e_i' = e_i - \bar{e}$$

siendo:

- e_i' : error para el punto i corregido de traslación.
- e_i : error para el punto i sin corregir la traslación.
- \bar{e} : error medio y traslación detectada.

Consideraciones sobre el ajuste por mm.cc.:

Ha de tenerse en cuenta que tanto en la gráfica anterior como en su curva de ajuste, tan sólo se han tenido en cuenta los errores aleatorios (distribuidos normalmente). En el caso de que existan valores atípicos, deberán eliminarse y/o ponderarse, de tal forma que queden atenuados. También es conveniente aumentar de forma significativa el tamaño de la muestra de puntos de control para obtener unos residuos adecuados tras el ajuste.

No obstante, y a pesar de los graves inconvenientes que se presentan al trabajar con escalas inferiores a 1/1.000, cabe destacar que la información suministrada es de una importancia más que considerable, dado que:

- Detecta y cuantifica la magnitud de cualquier traslación producida sobre el mapa en cualquiera de sus ejes (X_0 , Y_0). La precisión ronda la desviación típica estimada a priori para el mapa.
- Detecta y cuantifica la magnitud de cualquier cambio en la escala de sus ejes (λ_x , λ_y). La precisión oscila en torno a un primer decimal en el factor de escala.
- Detecta y cuantifica la magnitud de cualquier giro o rotación producida sobre el mapa (α). La precisión alcanzada en este tipo de error sistemático es del grado sexagesimal.

Asimismo, permite la posibilidad de corregir y eliminar todos estos errores sistemáticos de una forma relativamente rápida y sencilla: si se dispone de la BDG en formato

digital, tan sólo habrá que aplicar la transformación afín detectada a las coordenadas que la definen.

Se ha de tener en cuenta que, hasta la fecha, no se ha encontrado ningún estándar que proporcione una información tan exhaustiva y fidedigna sobre los errores sistemáticos que afectan a una determinada cartografía. Resulta de vital importancia la detección y eliminación de posibles sistematismos, dado que los nuevos estándares (FGDC, 1998) parten de la hipótesis de que éstos han sido corregidos previamente.

Aunque el ajuste por mm.cc. es una herramienta antigua (Rainsford, 1957), ésta es empleada de forma innovadora en el campo de la calidad en cartografía mediante la cuantificación y eliminación de errores sistemáticos en el control posicional. Ello obligará al usuario a habituarse a los resultados que proporciona, y a tratar de discernir cuándo los residuos proporcionados en el ajuste son o no los óptimos en función de la escala y del tamaño muestral. Para ello se cuenta con una primera aproximación mostrada en la figura 53.1.

5.4 Estimadores circulares

Los estimadores aplicados han sido: test Ómnibus, test Mardia y determinación de la dirección predominante junto con su intervalo de confianza. Asimismo, se han realizado diversas pruebas estadísticas mediante procesos de simulación en la elección de sub-muestras y sobre los puntos de control desarrollados en función de la dirección y el módulo del error.

5.4.1 Test Ómnibus

El test Ómnibus frente a cualquier alternativa (apartado 3.3.2.2), realiza un contraste para determinar si una serie de datos angulares ($\theta_1, \theta_2 \dots \theta_n$) poseen alguna dirección específica.

Los datos sobre los que se han realizado los experimentos han sido, al igual que en el apartado anterior (5.3), poblaciones de 1000 puntos con errores distribuidos según una

$N(0,1)$ (experimento A). En la tabla 54.1, se muestran los resultados sobre el total de la población para los distintos experimentos en función del contraste realizado. Para poder efectuar un análisis comparativo frente a los estimadores clásicos (no circulares), se ha aplicado un contraste del tipo *t-Student* al 95 % de nivel de confianza, teniendo en cuenta la corrección por Bonferroni: $Sudent\alpha_x = Student\alpha_y = 0,025$. Aquellas casillas en las que aparece marcada una "X" (columnas Q, R, S), indican que el test ha detectado una dirección predominante.

Tabla 54.1: Detección de dirección predominante mediante estimadores circulares¹⁹.

Población original: distribución $N(0,1)$																	
Experimento	Valores obtenidos sobre la población contaminada				Detección de una dirección predominante				Contaminación por errores sistemáticos introducidos artificialmente						Dirección detectada por estimadores circulares		
	Media		Desviación		Análisis t-Student		Estimadores circulares		Traslación		Factor Escala		Giro	Direcc. introdc.	Dirección detectada por estimadores circulares		
	μ_x	μ_y	σ_x	σ_y	S_x	S_y	Test Omnibus	Test Mardia	X_0	Y_0	λ_x	λ_y	α	θ^g	θ^g	Intervalo (g)	
A	0	0	1	1					0	0	1	1	0	-	-	-	
B	2	0.500	1	1	X	X	X	X	2	0.5	1	1	0	284,404	283,315	2,225	
C	1	3.500	1	1	X	X	X	X	1	3.5	1	1	0	217,717	218,073	1,136	
D	-2.750	2.330	1.47	1.24	X	X	X	X	0	0	0.9	1.1	0	144,749	145,871	1,619	
E	-0.330	0.330	1	1	X	X	X	X	0	0	1	1	0.75	150,000	150,319	9,173	
F	1	1	1	1	X	X	X	X	1	1	1	1	0	250,000	250,736	1,158	
G	0.500	0.000	1	1	X	X	X	X	0.500	0.000	1	1	0	100,000	100,155	8,409	
H	0.353	0.353	1	1	X	X	X	X	0.353	0.353	1	1	0	50,000	49,406	8,160	
I	0.200	0.200	1	1	X	X	X	X	0.200	0.200	1	1	0	50,000	49,329	15,693	
J	0.150	0.150	1	1	X	X	X	X	0.150	0.150	1	1	0	50,000	49,224	21,293	
K	0.100	0.100	1	1	X	X	X	X	0.100	0.100	1	1	0	50,000	49,597	33,259	
L	0.075	0.075	1	1			X	X	0.075	0.075	1	1	0	50,000	49,436	45,488	
M	0.050	0.050	1	1					0.050	0.050	1	1	0	50,000	-	-	
	N		P		Q		R	S	T		U		V	W	Z		

Al trabajar con estimadores circulares, la contaminación sobre cada componente (X e Y), ha de analizarse de forma conjunta. Esto es, para una contaminación de, por ejemplo, $0,5 m$ en X y $0,0 m$ en Y (experimento G):

$$Contaminación_{XY} = \sqrt{Cont_x^2 + Cont_y^2} = \sqrt{0,5^2 + 0,0^2} = 0,5 m$$

De forma análoga, en el experimento H se produce la misma contaminación global:

$$Contaminación_{XY} = \sqrt{Cont_x^2 + Cont_y^2} = \sqrt{0,353^2 + 0,353^2} \approx 0,5 m$$

¹⁹ Simulación realizada sobre una población artificial de 1000 puntos con errores aleatorios según una $N(0,1)$. En cada experimento (filas A – M) se ha introducido una determinada contaminación (columnas T-U), calculándose la dirección predominante que dicha contaminación produce sobre la población (columna V). La existencia de alguna dirección determinante se ha contrastado mediante una prueba del tipo t-Student (columnas Q), así como por la aplicación de los test de estadística circular Ómnibus y Mardia (columnas R y S respectivamente). En los casos en que ésta ha sido detectada, se indica la dirección predominante junto con su intervalo de confianza (columnas W-Z).

Tal y como puede apreciarse en el test Ómnibus, la detección de algún tipo de error sistemático que produzca una dirección sensiblemente predominante en la distribución de los errores, hace que la hipótesis nula no se cumpla, detectándose el sistematismo. Es más, para valores muy reducidos (experimento L), es ligeramente más sensible que el clásico contraste mediante una prueba del tipo t-Student.

5.4.2 Test de Mardia

El test de Mardia, como una mejora al test de Rayleigh y al igual que el Ómnibus, detecta la existencia de una dirección predominante en la distribución de los acimutes. En este caso, el análisis se realiza en función del valor de la longitud media resultante \bar{R} (apartado 3.3.2.3).

Los resultados de la tabla 54.1 (columna S), muestran un comportamiento muy similar al test Ómnibus: en todos los casos se detecta el sistematismo a excepción de cuando éste se reduce considerablemente con respecto a la σ_{mapa} :

$$Relación \Rightarrow \frac{Contaminación_{XY}}{\sigma_{\text{mapa}}} \left\{ \begin{array}{l} Relación_L \Rightarrow \frac{\sqrt{0.075^2 + 0.075^2}}{1} = \frac{0.106}{1} = 0.1 \rightarrow Si\ se\ detecta \\ Relación_M \Rightarrow \frac{\sqrt{0.050^2 + 0.050^2}}{1} = \frac{0.071}{1} = 0.07 \rightarrow No\ se\ detecta \end{array} \right.$$

5.4.3 Dirección predominante

Una vez que ha sido detectada la existencia de una dirección predominante (mediante el test de Mardia), se procede al cálculo de ésta junto con su intervalo de confianza. Para ello se emplea el test para una dirección específica (apartado 3.3.2.4), y cuyos resultados aparecen en las comunas W y Z de la tabla 54.1.

Puede observarse que en todos los experimentos la dirección detectada es correcta (se ajusta a la dirección introducida), siendo ésta ligeramente inferior a la real (columna V). En cuanto al intervalo de confianza al 95 % de probabilidad (columna Z), como era de esperar, éste aumenta conforme se reduce la relación entre

Contaminación_{XY} y σ_{mapa} . Así, la dirección predominante obtenida pudiera dejar de tener sentido cuando el intervalo fuese, por ejemplo, superior a 20° o 25° (experimento J).

5.4.4 Simulación de muestras

Bajo la misma perspectiva de trabajo que la empleada en el análisis de otros estándares, se ha utilizado la simulación en la elección de muestras aleatorias como herramienta para el estudio del comportamiento de los estimadores circulares.

Sobre las poblaciones de la tabla 54.1, se ha realizado un proceso de simulación seleccionando 1000 muestras aleatorias. En el experimento A (no existe ninguna dirección predominante ni sistematismo alguno), los porcentajes de aceptación de muestras son superiores al 95 % para cualquier tamaño muestral. En aquellos en los que la submuestra ha determinado la posible existencia de una dirección predominante (menos del 5 %), su intervalo de confianza es del orden de $\pm 40^{\circ}$ (independientemente del tamaño muestral); siendo la desviación típica tras el proceso de simulación muy elevada, por lo que dicha dirección queda completamente invalidada (ver tabla 54.2).

Tabla 54.2: Experimento A: porcentaje de submuestras aceptadas sin dirección predominante y acimut medio e intervalo de confianza en los casos rechazados.

Experimento A: población original N(0,1) sin errores sistemáticos							
Tamaño muestral (n)	Porcentaje de muestras aceptadas			Determinación de la dirección media y su intervalo de confianza en aquellas muestras que no superan el test			
	t-Student	Test Ómnibus	Test Mardia	Dirección media ($^{\circ}$)	Desviación (dirección) ($^{\circ}$)	Intervalo de confianza +/- ($^{\circ}$)	Desviación (intervalo) ($^{\circ}$)
20	97,7	95,5	95,4	220,6511	112,8764	32,1312	6,7468
30	97,2	94,7	94,6	238,259	147,7138	38,8326	10,9256
40	97,5	95,5	95,6	215,9037	185,209	42,3995	15,5092
50	97,9	96,0	95,7	201,175	225,4204	44,0216	18,0842
60	97,3	96,0	96,1	202,8341	255,136	42,2739	17,7645
70	98,7	95,9	96,1	218,7297	271,2245	44,0294	19,2301
80	98,0	96,1	97,3	219,9177	341,2944	44,3237	23,6415
90	98,1	96,3	96,5	231,3983	311,5583	44,0868	20,9046
100	98,0	97,0	96,8	213,7662	334,1685	43,9699	22,3155
110	98,4	97,4	97,4	192,4051	392,677	45,3468	26,4398
120	98,8	97,4	97,7	202,4569	421,9285	44,6566	27,1012
130	98,6	97,1	97,7	235,6848	434,9681	44,0276	26,4612
140	99,1	97,2	98,5	222,3672	547,6213	45,2138	35,3557
150	99,3	96,9	97,8	219,3079	456,4449	46,1075	30,5932
	Q	R	S	T	U	V	W

Lógicamente, al no existir ninguna dirección específica, las desviaciones típicas son muy elevadas (en algunos casos superiores a 400°). Hay que tener en cuenta que, por ejemplo, para un tamaño muestral de 140 puntos de control, tan sólo 15 de las 1000 muestras han sido rechazadas, siendo éstas las empleadas para la obtención de la dirección media en la distribución de los errores.

Para el resto de los experimentos (B – M) se ha seguido el mismo procedimiento, mostrándose en la tabla 54.3 los resultados de algunos de los casos más relevantes. En la margen izquierda aparecen los valores numéricos de cada proceso de simulación y en la derecha su representación gráfica.

Tabla 54.3: Experimentos de simulación: porcentaje de submuestras aceptadas sin dirección predominante y acimut medio e intervalo de confianza en los casos rechazados.

Experimento	n	Porcentaje de muestras aceptadas			Determinación de la dirección media e intervalo		Representación gráfica de los porcentajes de aceptación: — t-Student - - Test Ómnibus - - - Test de Mardia
		t-Student	Test Ómnib.	Test Mardia	Dirección media	Interv. conf. +/-	
B	20	0	0	0	283.70	13.04	
	30	0	0	0	283.11	11.10	
	40	0	0	0	283.42	9.82	
	50	0	0	0	283.19	9.09	
	60	0	0	0	283.39	8.39	
	70	0	0	0	283.31	7.79	
	80	0	0	0	283.26	7.34	
	90	0	0	0	283.25	6.97	
	100	0	0	0	283.33	6.68	
	110	0	0	0	283.19	6.39	
	120	0	0	0	283.25	6.13	
	130	0	0	0	283.30	5.95	
	140	0	0	0	283.30	5.72	
	150	0	0	0	283.34	5.52	
	G	20	65.2	64.7	60.6	99.90	
30		46.2	48.5	44.2	100.95	31.60	
40		29.3	34.5	30	100.81	31.87	
50		19	24.8	20.9	99.91	31.83	
60		9.4	17.4	14.8	100.42	31.31	
70		3.4	10.6	8.1	100.48	30.21	
80		1.9	6.6	3.9	100.49	29.24	
90		0.5	3.1	1.6	99.85	27.79	
100		0.1	1.8	1.5	99.98	26.58	
110		0.2	1.1	0.1	100.31	25.42	
120		0	0.6	0.2	100.09	24.55	
130		0	0.1	0	100.03	23.56	
140		0	0	0	100.00	22.71	
150		0	0.1	0	100.05	21.85	
H		20	71.8	63.4	57.7	64.16	28.94
	30	56.3	48.2	42.9	55.86	31.55	
	40	38.5	30.5	23.7	56.21	31.96	
	50	26.1	21.8	15.7	54.02	32.09	
	60	15.9	11.9	9.2	51.23	30.70	
	70	9	6.8	3.5	51.23	29.63	
	80	4.8	4.3	2.4	49.71	28.20	
	90	2.3	2	1	49.52	26.88	
	100	0.7	0.9	0.3	49.48	25.54	
	110	0.4	0.7	0.4	49.72	24.72	
	120	0.3	0.4	0.1	49.29	23.54	
	130	0	0	0	49.10	22.46	
	140	0	0	0	49.35	21.78	
	150	0	0	0	49.63	20.95	

... continúa

Tabla 54.3: ... continuación ...

Experimento	n	Porcentaje de muestras aceptadas			Determinación de la dirección media e intervalo		Representación gráfica de los porcentajes de aceptación: — : t-Student - - : Test Ómnibus - - - : Test de Mardia
		t-Student	Test Ómb.	Test Mardia	Dirección media	Interv. conf. +/-	
J	20	94.6	93.3	92.1	110.30	33.13	
	30	91.7	89.6	88.2	125.17	37.96	
	40	88.6	86.6	85	90.04	39.14	
	50	86.5	84	82.8	93.15	40.37	
	60	83.9	81.1	79.3	71.90	41.02	
	70	81.3	78.5	76	73.56	41.35	
	80	79.6	76.3	73.3	66.68	41.83	
	90	77.1	73.3	69.9	77.58	41.63	
	100	74	70.2	66.1	72.43	41.28	
	110	70.6	67.5	63.7	62.67	41.20	
	120	67.4	64.2	60.7	65.75	41.04	
	130	64	61	56.9	62.84	41.21	
	140	61.3	58.5	54.3	61.17	40.95	
	150	57	54.4	50.3	58.72	40.85	
K	20	95.13	92.1	92.69	160.07	33.39	
	30	95.37	92.87	92.22	159.66	38.77	
	40	95.5	93.1	92.46	126.20	41.94	
	50	94.7	92.63	91.39	142.49	41.01	
	60	94.6	91.15	90.3	102.89	42.71	
	70	93.3	90.4	91.3	124.66	42.17	
	80	91.2	87.6	89.1	96.36	43.16	
	90	91.1	87.8	89.1	86.47	42.76	
	100	89.4	86.7	86.4	97.19	43.24	
	110	89.3	85.03	86	88.17	42.50	
	120	88.36	84.85	86.75	83.83	42.83	
	130	88	85.39	86.63	78.38	42.94	
	140	85.6	84.3	83.8	75.81	43.68	
	150	83.8	82.2	82.3	91.39	43.84	
L	20	96.6	94.1	93.8	177.92	32.77	
	30	96.7	95.6	95	159.41	38.09	
	40	96.5	94.2	93.4	148.70	40.49	
	50	95.6	92.6	94.06	113.33	39.91	
	60	95.1	92.9	92.4	109.51	42.98	
	70	95.3	93.5	93.93	117.84	42.88	
	80	94.63	93.4	94.2	120.52	43.14	
	90	95.1	92.1	92.6	113.67	43.44	
	100	94.3	93.1	93.3	86.42	43.18	
	110	95	90.6	92.2	107.24	44.25	
	120	94.1	91.8	91.1	101.88	44.36	
	130	92.7	90.3	91	91.30	43.74	
	140	93.4	90.9	92.8	115.39	44.16	
	150	93.4	90.1	91.5	98.87	44.10	

Tanto la prueba t-Student como los estimadores circulares, ofrecen resultados óptimos en la detección de una dirección predominante en los experimentos B, C, D y F ⁽²⁰⁾: en todos los casos, independientemente del tamaño muestral, se detecta y determina correctamente dicha dirección. No obstante, esto no sucede así cuando se reduce el coeficiente de contaminación con respecto a la desviación poblacional. En cuanto a los experimentos mostrados en la tabla 54.3, cabe destacar los siguientes aspectos:

- **Experimento B.** Independientemente del tamaño muestral, no hay ninguna submuestra que supere ninguno de los test, por lo que en todos los casos se detecta la existencia de una dirección predominante. Dicha dirección ($284,4^\circ$), se ajusta a la poblacional con un intervalo de confianza que, lógicamente, se reduce al aumentar el tamaño muestral. Para el caso de 100 puntos de control, éste es de $6,68^\circ$.
- **Experimento G.** Si se reduce considerablemente el sistematismo (tan solo se da en la dirección del eje $X \rightarrow Ex = 0,5m, \theta = 100^\circ$), los test siguen detectando la existencia de una dirección significativa, aunque el intervalo de confianza aumenta considerablemente. No obstante, el proceso de simulación proporciona valores medios en la determinación de la dirección que se ajustan al valor poblacional. Con tamaños muestrales reducidos ($n < 80$), existe cerca de un 5% de probabilidad de que no sea detectada por los test Ómnibus y Mardia.
- **Experimento H.** En este caso, la componente cuadrática del error en X e Y es la misma que en el experimento anterior ($0,5m$), pero repartido de igual forma entre ambas. Los niveles de aceptación son similares al experimento G, sin apreciarse un comportamiento significativamente diferenciado. Lo mismo sucede con la determinación de la dirección predominante.
- **Experimento J.** Al reducirse la relación entre la magnitud del error y la desviación poblacional ($Relación_J = 0,28$), existe una mayor incertidumbre en su detección por cualquiera de los test: con 100 puntos de control cerca de un 33% de las muestras lo detecta. Lo mismo sucede con la dirección media del sistematismo: el valor medio determinado por el test mediante simulación ($72^\circ \pm 41^\circ$) difiere considerablemente de la real (50°). No obstante, el propio test

²⁰ En la tabla 54.3 tan sólo aparece el experimento B dado que los demás ofrecen idénticos resultados.

advierte de esta situación al proporcionar un valor muy amplio para el intervalo de confianza al 95 %.

- **Experimento K.** Conforme se va reduciendo la magnitud del error direccional introducido, los niveles de aceptación aumentan, detectándose un menor número de muestras con una dirección específica.
- **Experimento L.** Por último, tal y como se aprecia para el total de la población en la tabla 54.1, existe una ligera mayor sensibilidad entre los test Ómnibus y Mardia con respecto a la prueba t-Student a la hora de detectar un error direccional (los porcentajes de aceptación son algo superiores en este último caso). No obstante, una vez detectado, la dirección posee un intervalo de confianza excesivo como para ser considerado un error significativo. Asimismo, la variabilidad en las curvas es mucho mayor que en los experimentos anteriores (presentan un comportamiento más aleatorio e irregular).

En general, y para todos los experimentos, la curva de aceptación del test t-Student se sitúa por encima de los estimadores circulares. En siguiente lugar aparece la curva del test Ómnibus y, por último, la del test de Mardia. Todo ello, a priori, muestra un comportamiento ligeramente más restrictivo de los estimadores circulares con respecto al clásico t-Student.

También se han aplicado estos estimadores sobre datos reales (hoja 926 1-1 del ICA²¹), mostrándose los resultados en la tabla 54.4. Dado que dicha cartografía posee errores sistemáticos²², y que la dirección dominante en la que se encuentra dicho error es de $216,82^\circ$, se ha estudiado la sensibilidad de los diferentes estimadores en su detección mediante procesos de simulación de muestras.

²¹ Previamente se han eliminado aquellos valores considerados como atípicos.

²² Error medio en X = -0,59 m ; Error medio en Y = -2,21 m

Tabla 54.4: Porcentaje de sub-muestras aceptadas sin dirección predominante y acimut medio e intervalo de confianza en los casos rechazados para la hoja 926 1-1 del ICA con y sin errores sistemáticos.

Hoja 926 1-1 con errores sistemáticos poblacionales: $\overline{E}_x = -0,59 m$; $\overline{E}_y = -2,21 m$						
n	Porcentaje de muestras aceptadas			Determinación de la dirección media e intervalo		Representación gráfica de los porcentajes de aceptación: — : t-Student - - : Test Ómnibus - - - : Test de Mardia
	t-Student	Test Ómnib.	Test Mardia	Dirección media	Interv. conf. +/-	
30	57.99	38.53	44.15	209.94	21.15	
40	39.94	21.77	28.3	209.30	17.73	
50	24.31	10.93	17.41	209.39	14.63	
60	9.59	2.79	5.52	209.41	11.10	
70	1.80	0.28	0.82	209.51	8.59	
Hoja 926 1-1 sin errores sistemáticos poblacionales: $\overline{E}_x = 0 m$; $\overline{E}_y = 0 m$						
n	Porcentaje de muestras aceptadas			Determinación de la dirección media e intervalo		Representación gráfica de los porcentajes de aceptación: — : t-Student - - : Test Ómnibus - - - : Test de Mardia
	t-Student	Test Ómnib.	Test Mardia	Dirección media	Interv. conf. +/-	
30	97.06	93.88	94.57	125.20	49.27	
40	97.96	94.11	95.31	127.47	53.50	
50	98.55	94.60	97.13	121.43	64.60	
60	99.05	95.47	97.86	108.79	56.43	
70	99.79	97.39	98.95	28.11	78.32	

Si bien no aparecen indicadas en la tabla 54.4 las desviaciones de los valores medios de dirección e intervalo de confianza detectados (tal y como sucedía en el caso de la tabla 54.2), éstos son muy reducidos para el caso en el que la hoja tiene errores sistemáticos y aumentan considerablemente para la hoja en la que han sido corregidos (segunda parte de la tabla 54.4). Como puede apreciarse, se confirma el comportamiento de los estimadores circulares es correcto para una relación entre el sistematismo y la $\sigma_{\text{mapa}}^{(23)}$ de:

$$\text{Relación}_{\text{Hoja 1-1}} \Rightarrow \frac{\text{Sistematismo}_{XY}}{\sigma_{\text{mapa}}} = \frac{\sqrt{0.59^2 + 2.21^2}}{4.5} = \frac{2.29}{4.5} = 0.51$$

²³ Mediante el estándar EMAS, se ha estimado que $\sigma_{\text{mapa}} = 4,50 m$

Asimismo, el ángulo determinado como dirección predominante se ajusta al valor poblacional, reduciéndose el intervalo de confianza conforme aumenta el tamaño muestral.

En cuanto a la segunda mitad de la tabla 54.4, se trata de la Hoja 926 1-1 con los errores sistemáticos por traslación corregidos. Los porcentajes de aceptación han aumentado hasta cerca del 95 % de los casos. Ello confirma que, independientemente de la σ_{mapa} , al trabajar los estimadores circulares con vectores unitarios, su análisis es correcto en cuanto a la distribución de los errores: no hay sistematismos. Asimismo, los intervalos de confianza calculados para aquellos casos en los que sí se ha podido detectar algún tipo de dirección dominante (menos del 5 % de las submuestras), son excesivamente amplios como para que puedan tenerse en consideración.

5.4.5 Análisis de los módulos

Dado que los estimadores circulares comúnmente empleados no analizan los módulos (solamente estudian direcciones sobre vectores unitarios), se ha procedido a desarrollar los datos en forma de gráficos Acimut / Módulo (A/M). Lo que se pretende en este apartado, es obtener algún tipo de indicador que relacione los valores del módulo con respecto a la dirección, de tal forma que pueda complementarse con los estimadores circulares e incluirse como análisis de la magnitud del error.

En este caso, los experimentos se han realizado sobre los datos reales de la hoja del ICA 926 1-1 empleados al final del apartado anterior (tanto con errores sistemáticos como sin ellos). Así, se dispone de 100 puntos de control ordenados de menor a mayor en función de su acimut (listado XXXXXX), los cuales se representan en una gráfica A/M (figura 54.1).

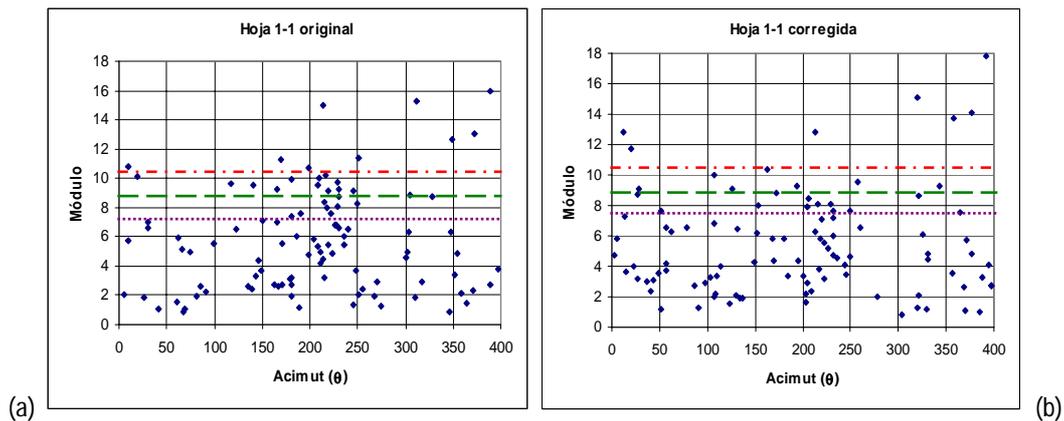


Figura 54.1: Gráfica A/M de la Hoja 926 1-1: (a) datos originales con errores sistemáticos y (b) datos corregidos de sistematismos.

Las líneas horizontales en trazado discontinuo de la figura 54.1, representan (en valor absoluto al tratarse de módulos):

- La probabilidad al 90 %: $Z_{\alpha/2} \times \sigma_{mapa} = 1,64 \times 4,50 = 7,40 \text{ m}$ (línea inferior en color magenta punteado).
- La probabilidad al 95 %: $Z_{\alpha/2} \times \sigma_{mapa} = 1,96 \times 4,50 = 8,82 \text{ m}$ (línea intermedia en color verde con trazos prolongados).
- La probabilidad al 97.5 %: $Z_{\alpha/2} \times \sigma_{mapa} = 2,24 \times 4,50 = 10,09 \text{ m}$ (línea superior en color rojo con trazos cortos intercalados por puntos).

Como todas las medidas tomadas sobre datos reales medidos tanto en campo como sobre la BDG, vienen afectadas de un cierto ruido en la medición, sobre ambas series de datos, se han efectuado distintos procesos de suavizado de los módulos distribuidos respecto a su dirección acimutal²⁴ para amortiguar dicho ruido. Asimismo, se ha procedido a su agrupamiento por intervalos de 10° y 20° :

- **Suavizado 1:** se han utilizado los valores inmediatamente anterior y posterior al valor suavizado, sin ninguna ponderación. La ecuación utilizada es:

$$Mod_Suav_i = \frac{Mod_{i-1} + Mod_i + Mod_{i+1}}{3}$$

- **Suavizado 2:** se han utilizado los valores inmediatamente anterior y posterior al valor suavizado, ponderando al doble del valor central. La ecuación es:

²⁴ Los datos próximos a 0° y a 400° han sido suavizados con una réplica de la serie original $\pm 2\pi$.

$$Mod_Suav_i = \frac{Mod_{i-1} + 2 \times Mod_i + Mod_{i+1}}{4}$$

- **Suavizado 3:** se han utilizado los valores inmediatamente anterior y posterior al valor suavizado, realizando una ponderación de seis veces el valor central. La ecuación empleada es:

$$Mod_Suav_i = \frac{Mod_{i-1} + 6 \times Mod_i + Mod_{i+1}}{8}$$

- **Suavizado 4:** se han utilizado los dos valores inmediatamente anterior y posterior al valor suavizado, sin ponderación. La ecuación es:

$$Mod_Suav_i = \frac{Mod_{i-2} + Mod_{i-1} + Mod_i + Mod_{i+1} + Mod_{i+2}}{5}$$

- **Suavizado 5:** se han utilizado los dos valores inmediatamente anterior y posterior al valor suavizado, ponderando al triple del valor central y al doble del inmediatamente anterior y posterior. La ecuación utilizada es:

$$Mod_Suav_i = \frac{Mod_{i-2} + 2 \times Mod_{i-1} + 3 \times Mod_i + 2 \times Mod_{i+1} + Mod_{i+2}}{9}$$

Una vez transformados y suavizados todos los puntos de control, se ha efectuado una serie de cálculos estadísticos cuyos resultados se muestran en las tablas 54.5 (para la Hoja 1-1 con errores sistemáticos) y 54.6 (para la Hoja 1-1 corregida de dichos errores). En cuanto a los datos agrupados por intervalos, éstos han sido tratados en la tabla 54.7.

Tabla 54.5: Estadísticos obtenidos sobre las gráficas A/M de la Hoja 926 1-1 original.

Estadísticos		Gráfica A/M
Experimento:	Hoja 1-1 original	
Media	5.784	
Desv. Típica	3.508	
Mediana	5.435	
Nº de rachas	51	
Estadístico Z	0.001	
Test W.W.:	Distribución aleatoria	
Experimento:	Suavizado 1	
Media	5.784	
Desv. Típica	1.981	
Mediana	5.821	
Nº de rachas	27	
Estadístico Z	-4.8	
Test W.W.:	Distribución no aleatoria	
Experimento:	Suavizado 2	
Media	5.784	
Desv. Típica	2.094	
Mediana	5.879	
Nº de rachas	27	
Estadístico Z	-4.8	
Test W.W.:	Distribución no aleatoria	
Experimento:	Suavizado 3	
Media	5.784	
Desv. Típica	2.668	
Mediana	5.645	
Nº de rachas	41	
Estadístico Z	-1.999	
Test W.W.:	Distribución no aleatoria	
Experimento:	Suavizado 4	
Media	5.784	
Desv. Típica	1.510	
Mediana	5.813	
Nº de rachas	18	
Estadístico Z	-6.6	
Test W.W.:	Distribución no aleatoria	
Experimento:	Suavizado 5	
Media	5.784	
Desv. Típica	1.641	
Mediana	5.935	
Nº de rachas	17	
Estadístico Z	-6.8	
Test W.W.:	Distribución no aleatoria	

Tabla 54.6: Estadísticos A/M de la Hoja 926 1-1 corregida de sistematismos.

Estadísticos		Gráfica A/M
Experimento:	Hoja 1-1 corregida	
Media	5.495	
Desv. Típica	3.416	
Mediana	4.657	
Nº de rachas	48	
Estadístico Z	-0.503	
Test W.W.:	Distribución aleatoria	
Experimento:	Suavizado 1	
Media	5.495	
Desv. Típica	1.831	
Mediana	5.659	
Nº de rachas	23	
Estadístico Z	-5.528	
Test W.W.:	Distribución no aleatoria	
Experimento:	Suavizado 2	
Media	5.495	
Desv. Típica	2.022	
Mediana	5.427	
Nº de rachas	29	
Estadístico Z	-4.322	
Test W.W.:	Distribución no aleatoria	
Experimento:	Suavizado 3	
Media	5.495	
Desv. Típica	2.622	
Mediana	5.105	
Nº de rachas	41	
Estadístico Z	-1.910	
Test W.W.:	Distribución aleatoria	
Experimento:	Suavizado 4	
Media	5.495	
Desv. Típica	1.288	
Mediana	5.507	
Nº de rachas	17	
Estadístico Z	-6.734	
Test W.W.:	Distribución no aleatoria	
Experimento:	Suavizado 5	
Media	5.495	
Desv. Típica	1.449	
Mediana	5.523	
Nº de rachas	19	
Estadístico Z	-6.332	
Test W.W.:	Distribución no aleatoria	

Junto con los más comunes estadísticos (media, mediana y desviación típica), se ha realizado un estudio de aleatoriedad sobre la distribución de los errores en función de su dirección (gráficas A/M), mediante un contraste de rachas aplicando el test de Wald–Wolfowitz (ver apartado 3.4.2).

La existencia de una mayor o menor concentración de vectores en una dirección, se ve representado en las gráficas A/M por un mayor número de puntos en dicha dirección. Si estos puntos se encuentran en la parte superior o inferior, indica que su magnitud (módulo) es de mayor o menor envergadura.

Sobre los datos de la hoja original (tabla 54.5), se observa una mayor concentración de puntos en el intervalo $[200^{\circ} - 250^{\circ}]$. Ello es debido a que el sistematismo direccional detectado para los 100 puntos de control se produce a $209,4^{\circ} \pm 25,4^{\circ}$. Si se examinan las gráficas de la hoja corregida (tabla 54.6), dicha concentración desaparece y no se observa aglomeración alguna de puntos.

Dependiendo del suavizado que se aplique, los resultados tanto gráficos como analíticos varían sustancialmente. La concentración de puntos en el comienzo del tercer cuadrante se hace mucho más visible en las gráficas suavizadas 5 y 6. Asimismo, éstas presentan una mayor regularidad gráfica, con varios “picos” en diferentes direcciones. En la figura 54.2 puede observarse una ampliación del suavizado 5 sobre la Hoja 926 1-1 sin corregir de errores sistemáticos, así como la representación de los módulos sobre una gráfica radial.

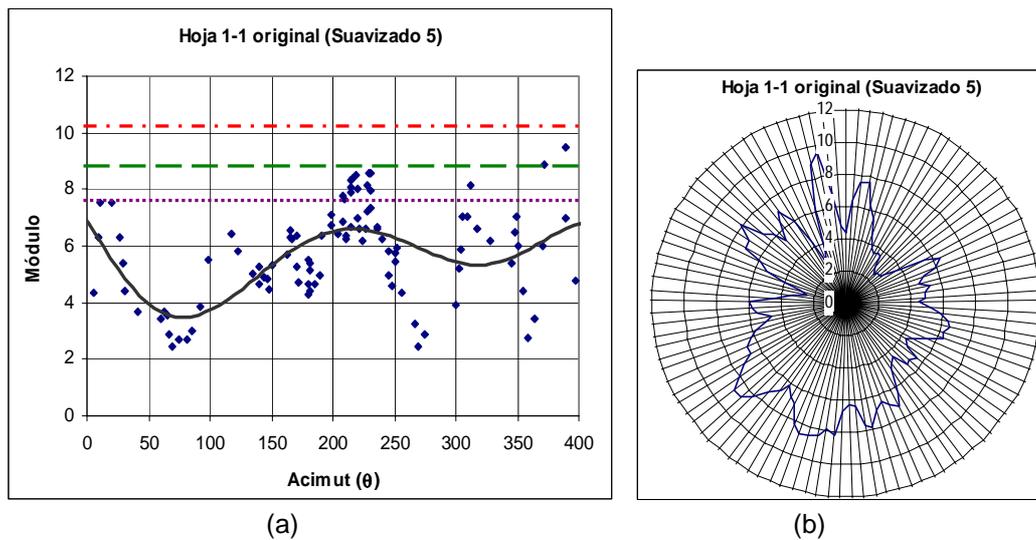


Figura 54.2: Gráfica A/M de la Hoja 926 1-1 original y línea de tendencia en la magnitud de los módulos (a), y dicha gráfica representada de forma radial (b)

La línea continua sinusoidal, muestra más claramente la tendencia entre la dimensión de los módulos y la dirección que siguen. Se continúa apreciando la concentración de vectores de mayor magnitud en el tercer cuadrante (dirección del sistematismo), así como dos “picos” con vectores mínimos en las direcciones de $\sim 70^\circ$ y $\sim 270^\circ$. En el resto de las direcciones la dispersión es mayor. Estas direcciones se aprecian más claramente en las gráficas de la hoja con sistematismos que en la corregida, donde no se puede afirmar que exista ninguna tendencia clara.

Estudiando los resultados analíticos para ambas hojas, también existe un cambio radical si se trabaja con los datos originales o los suavizados. Un claro ejemplo de ello es el número de rachas detectadas sobre los datos sin suavizar: se aproximan a 50⁽²⁵⁾ tanto para la hoja con sistematismos como sin ellos. Sin embargo, cuando se procede al suavizado, éstas disminuyen sustancialmente y tan solo en el caso del suavizado 3, se aprecia un comportamiento diferenciado entre ambas hojas.

El test de rachas de Wald–Wolfowitz para el estudio de la aleatoriedad en la distribución de los módulos, muestra que en la práctica totalidad de los casos suavizados (a excepción del suavizado 3 para la hoja corregida), los vectores no se distribuyen de forma aleatoria. Esto es, el test presenta una gran sensibilidad al estudio de la aleatoriedad de los módulos. Es más, podría afirmarse que es extremadamente sensible a este tipo de comportamientos en el control de la exactitud

²⁵ Hay que recordar que el número total de puntos de control es de $N = 100$.

posicional, por lo que, a priori, no sería aconsejable su utilización en estas condiciones.

Lógicamente, la diferencia existente entre la mediana y la media de los módulos, es mayor entre ambas hojas cuando los datos no han sido suavizados. Si bien el número de rachas es similar (51 para la hoja con sistematismos y 48 para la corregida), sí que existe una diferencia sustancial entre el estadístico de Wald–Wolfowitz para uno y otro caso (0.001 y -0.503 respectivamente).

Para que el test pudiera detectar un comportamiento diferenciado, el nivel de confianza con el que habría que realizar la prueba no sería admisible. Así, por ejemplo, en el caso de los experimentos realizados:

$$\left. \begin{array}{l} \text{Probabilidad al } 30\% \\ \text{de nivel de confianza} \end{array} \right\} \Rightarrow 1 - \alpha = 0,30 \rightarrow \alpha = 0,70 \rightarrow P_{\frac{\alpha}{2}; 1 - \frac{\alpha}{2}} = [-0,39; +0,39]$$

En cuanto a los datos agrupados por tramos para reducir los niveles de ruido en la medición (tabla 54.7), cabría destacar que:

- Al agruparse en intervalos se pierde la visión sobre una mayor o menor concentración de vectores en una determinada dirección, siendo más acusado para tramos de 20^º. Para poder detectar este factor, el análisis ha de acompañarse con gráficas en las que se indique el número de vectores por intervalo (figura 54.3) para poder detectar direcciones predominantes.

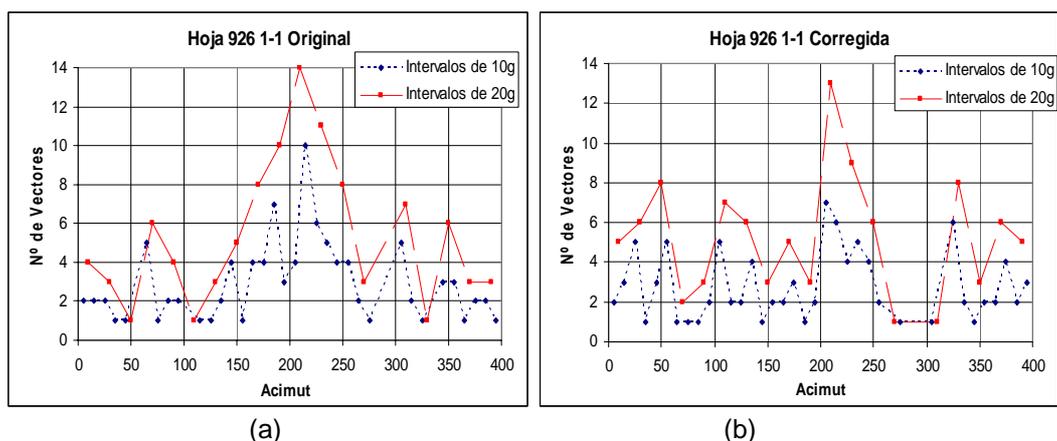


Figura 54.3: Histograma con el número de vectores (puntos de control) incluidos en cada intervalo de 10^º y 20^º. para la hoja 926 1-1 original (a) y corregida de sistematismos (b).

Tabla 54.7: Estadísticos A/M sobre datos agrupados de la Hoja 926 1-1 original y corregida de sistematismos.

Estadísticos		Gráfica A/M
Experimento:	Hoja 1-1 original Intervalo agrup.: 10 ^g	
Media	5,578	
Desv. Típica	2,515	
Mediana	5,384	
Nº de rachas	16	
Estadístico Z	-0,68	
Test W.W.:	Distribución aleatoria	
Experimento:	Hoja 1-1 original Intervalo agrup.: 20 ^g	
Media	5,560	
Desv. Típica	2,206	
Mediana	5,587	
Nº de rachas	11	
Estadístico Z	0,01	
Test W.W.:	Distribución aleatoria	
Experimento:	Hoja 1-1 corregida Intervalo agrup.: 10 ^g	
Media	5,292	
Desv. Típica	2,144	
Mediana	5,310	
Nº de rachas	22	
Estadístico Z	0,85	
Test W.W.:	Distribución aleatoria	
Experimento:	Hoja 1-1 corregida Intervalo agrup.: 20 ^g	
Media	5,285	
Desv. Típica	1,988	
Mediana	5,781	
Nº de rachas	10	
Estadístico Z	-0,23	
Test W.W.:	Distribución aleatoria	

- Empleando este tipo de gráficas (figura 54.3), se detecta la existencia de una dirección predominante por la concentración de vectores en el intervalo de 210^g (vectores agrupados en tramos de 20^g). Así, para la hoja original, entre 170^g y 250^g se encuentra más del 50% de los vectores de error de los puntos de control, mientras que en la hoja corregida es tan sólo del 36% (en un arco de $\pi/4$).

- Respecto a las pruebas estadísticas realizadas sobre los módulos en la tabla 54.7, en todos los casos el resultado del test de rachas acepta la hipótesis nula: los módulos medios por intervalo se distribuyen de forma aleatoria.
- Los valores de media, desviación y mediana, son similares en todos los casos, por lo que no existe un comportamiento claramente diferenciado entre los datos agrupados en tramos de 10^9 ó 20^9 para la hoja original o la hoja corregida de errores sistemáticos.

Tras los experimentos realizados, y a modo de conclusión sobre el empleo de los módulos, cabría mencionar que los gráficos A/M proporcionan al investigador una valiosa información sobre la distribución y magnitud de los mismos con respecto a su dirección. Se observa una mayor concentración de vectores en aquellos casos en los que existe una dirección predominante.

No obstante, las pruebas estadísticas aplicadas no poseen una sensibilidad acorde con el tipo de datos analizados: para la práctica totalidad de los suavizados se advierte de un comportamiento no aleatorio, mientras que sobre los datos originales no se detecta el sistematismo. Si se agrupan por intervalos de 10^9 o 20^9 , en ningún caso se han detectado los sistematismos (a excepción de la distribución del porcentaje de vectores por cuadrante figura 54.3 – a).

5.5 Aplicación de los estándares mediante curvas operativas

Partiendo de los experimentos realizados en el apartado anterior, se pueden obtener nomogramas a modo de curvas operativas (apartados 3.6 y 4.3.3), de tal forma que determinen la probabilidad de aceptación para un producto en función de diferentes parámetros:

- Número de puntos empleado como tamaño muestral.
- Nivel de confianza con el que se deseen obtener los resultados finales.
- Relación existente entre la distribución estimada y la distribución real de los errores en la BDG (Ratio: $\sigma_{\text{Estimado}} / \sigma_{\text{Mapa}}$).

De esta forma se han desarrollado diferentes nomogramas de aceptación para la aplicación de los principales estándares en los que esta técnica podría ser viable: EMAS, NMAS y ASPRS. El cambiar el orden de exposición de los estándares, se debe a que los resultados sobre EMAS ofrecen una idea más clara y robusta del funcionamiento de las curvas operativas aplicadas al NMAS y ASPRS.

En el caso del estándar NSSDA, dado que no se trata de un test en el que se acepte o rechace la BDG sino que ofrece un valor estimado de su exactitud, el tratamiento ha sido diferente para adaptarlo a dicha peculiaridad.

Por último, se han empleado técnicas de simulación sobre los estimadores circulares para observar su comportamiento y posible aplicación en la detección y determinación de direcciones dominantes en la distribución de errores.

5.5.1 EMAS

Tras realizar la corrección por Bonferroni (apartado 5.2.2) y calcular la curva con los niveles de aceptación del estándar (figura 52.5), se obtuvieron dos ideas básicas:

- La aplicación de la batería de test que conforman el EMAS proporciona una información muy exhaustiva a cerca del comportamiento, el tipo y la magnitud de los errores del mapa para cada una de sus componentes.
- Empleando tamaños muestrales reducidos, los niveles de aceptación son óptimos al 95 % de nivel de confianza.

De esta forma, mediante procesos de simulación descritos en el apartado 4.3.3 (figura 43.4), se puede llegar a construir un nomograma de curvas aceptación de carácter universal que muestren la capacidad²⁶ del estándar. Dichos resultados se presentan en la figura 55.2 de forma gráfica y en la tabla 55.1 de forma numérica. Sobre poblaciones generadas artificialmente y distribuidas según una $N(0,1)$, se ha aplicado

²⁶ Ver Anexo 9.

el estándar EMAS con la corrección por Bonferroni bajo distintos niveles o ratios de exigencia²⁷, siendo $\sigma_{EMAS} < \sigma_{Mapa}$ ($\sigma_{Mapa} = 1,00$; $\sigma_{EMAS} = 0,975; 0,950... 0,700$).

Su estabilidad y robustez es muy elevada por el propio proceso de simulación, obteniéndose una variabilidad máxima en las curvas representadas de $\pm 0,75\%$. Esto es, se han llegado a conseguir curvas de aceptación, genéricas y universales en función de la relación existente entre la varianza poblacional, la varianza muestral en la aplicación del EMAS y el tamaño de la muestra utilizada.

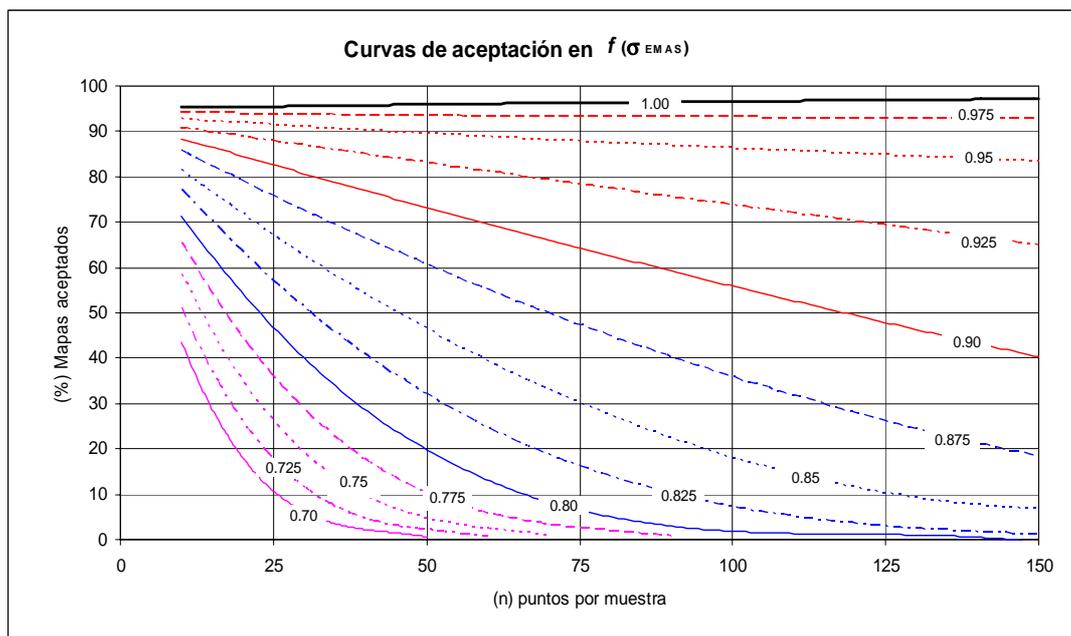


Figura 55.2: Nomograma gráfico del estándar EMAS con las curvas de aceptación en función del tamaño muestral y el ratio de exigencia $\sigma_{EMAS} / \sigma_{Mapas}$.

²⁷ Se emplea la denominación σ_{EMAS} para referirse al nivel de exigencia ($\sigma_{Estimado}$) del estándar EMAS en el análisis de la varianza.

Tabla 55.1: Nomograma analítico del estándar EMAS con los (%) de aceptación en función del tamaño muestral (n) y el ratio de exigencia $\sigma_{EMAS} / \sigma_{Mapa}$.

n	Ratio aplicado: $\sigma_{EMAS} / \sigma_{Mapa}$												
	0,700	0,725	0,750	0,775	0,800	0,825	0,850	0,875	0,900	0,925	0,950	0,975	1,00
20	17,9	25,7	34,7	44,3	54,3	63,4	71,8	78,8	84,6	88,8	91,8	94,0	95,4
30	6,4	11,6	19,2	28,5	39,6	51,7	62,7	72,8	80,6	86,6	90,8	93,7	95,6
40	2,1	4,8	9,7	17,9	28,3	40,6	54,2	66,7	76,8	85,0	90,3	93,6	95,7
50	0,6	1,9	4,9	10,4	19,6	32,0	46,3	60,9	73,2	83,1	89,6	93,4	95,9
60	0,2	0,7	2,2	5,9	13,1	24,6	39,2	55,5	69,5	80,9	88,8	93,3	96,0
70	0,0	0,2	1,0	3,2	8,6	18,5	32,9	49,8	66,2	79,2	88,2	93,2	96,3
80	0,0	0,1	0,4	1,8	5,6	13,9	27,4	44,8	62,6	77,6	87,4	93,2	96,5
90	0,0	0,0	0,2	0,9	3,5	10,1	22,3	40,0	59,2	75,7	86,9	93,1	96,4
100	0,0	0,0	0,1	0,4	2,1	7,3	18,4	35,8	56,0	73,9	86,3	93,0	96,7
110	0,0	0,0	0,0	0,2	1,3	5,2	14,6	31,6	52,6	71,9	85,8	93,1	96,7
120	0,0	0,0	0,0	0,1	0,7	3,6	11,8	27,7	49,3	70,2	84,9	92,9	96,8
130	0,0	0,0	0,0	0,0	0,4	2,5	9,3	24,5	46,5	68,2	84,5	93,1	97,0
140	0,0	0,0	0,0	0,0	0,2	1,7	7,2	21,3	43,2	66,7	83,9	92,9	97,1
150	0,0	0,0	0,0	0,0	0,1	1,1	5,5	18,6	40,3	64,9	83,4	92,9	97,2

La mejor forma para comprender la utilización del nomograma es mediante un ejemplo práctico. Sea el caso de una hoja cartográfica a escala 1/10.000 donde se ha tomado una muestra de 98 puntos de control. Para cada punto se han obtenido sus coordenadas en el terreno y en la BDG (ver apartado 4.2), con un comportamiento en las desviaciones de $\sigma_{px} = 3,76$ m y $\sigma_{py} = 4,80$ m. El cartógrafo que ha de contrastar la calidad del producto, se encuentra con que no puede tomar el total de la población (ya que un mapa tiene infinitos puntos), y ha de trabajar con una reducida muestra de puntos de control. Este problema se puede solventar empleando procesos de simulación.

Así, de los 98 puntos de la muestra inicial se toman 1000 sub-muestras aleatorias de tamaño $n_m = 20, 30$ y 40 puntos cada una de ellas. A continuación, se aplica el estándar a cada sub-muestra con un valor de $\sigma_{EMAS} = 4,00$ m, ligeramente inferior al de aceptación para los 98 puntos. Para cada proceso de simulación en la elección de muestras, se ha obtenido el porcentaje de aceptación del estándar EMAS (tabla 55.2). Al entrar en el nomograma (ver detalle en la figura 55.3), se observa que los porcentajes resultantes de aceptación indican una situación intermedia entre las curvas de $\sigma_{EMAS} = 0,90$ y $0,875$. Dado que el nomograma está tipificado para una población $\sigma_p = 1$, realizando una sencilla interpolación se puede obtener que la varianza del mapa analizado está en torno al valor de $\sigma_{EMAS}^2 = (4,50 \text{ m})^2$. Si se repite el proceso de simulación aplicando dicho valor como criterio de varianza en el EMAS, los

porcentajes de aceptación están en torno al 95 %, independientemente del tamaño muestral. Esto indica que la desviación típica del mapa analizado es de 4,50 m (0,45 mm a la escala del mapa) con un 95 % de nivel de confianza. El valor obtenido es intermedio entre los calculados para el total de la muestra ($\sigma_{px} = 3,76$ m y $\sigma_{py} = 4,80$ m), lo que indica la existencia de una cierta correlación en el comportamiento de ambas componentes.

Tabla 55.2: Porcentajes de aceptación para la cartografía analizada en función del tamaño muestral y de la σ_{EMAS}

n_m	Porcentajes de aceptación	
	$\sigma_{EMAS} = 4,0$	$\sigma_{EMAS} = 4,5$
20	81,2 %	95,1 %
30	75,2 %	96,1 %
40	67,5 %	96,7 %

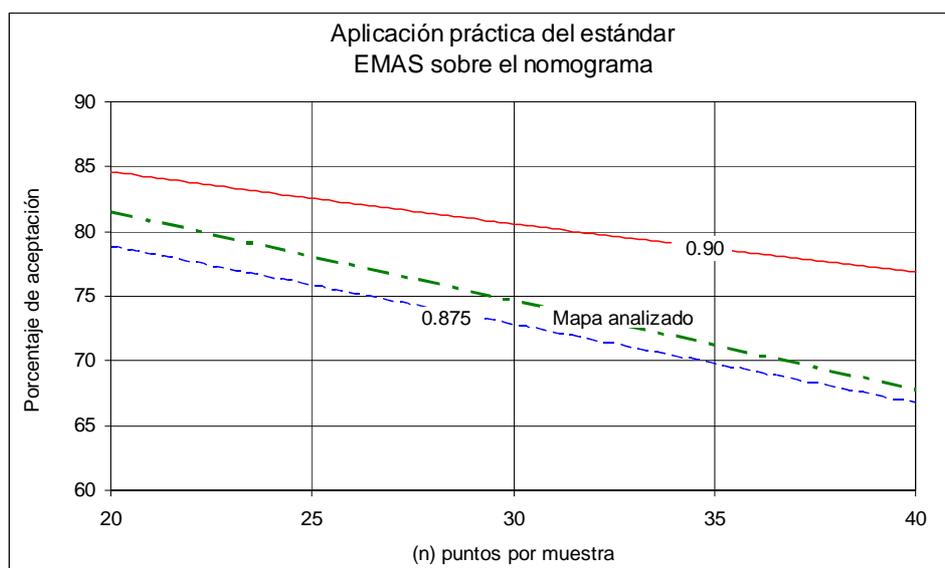


Figura 55.3: Detalle del nomograma y su aplicación práctica a un caso real con el estándar EMAS.

Asimismo, el nomograma permite conocer los riesgos tanto del usuario como del productor a la hora de aceptar o rechazar una determinada BDG en función del tamaño muestral y la capacidad del estándar o ratio de exigencia entre la exactitud posicional demandada (σ_{EMAS}) y la real (σ_{Mapa}):

$$Capacidad \Rightarrow \text{Ratio de Exigencia} = \frac{\text{Exactitud demandada}}{\text{Exactitud real}} = \frac{\sigma_{EMAS}}{\sigma_{Mapa}}$$

Si, por ejemplo, se desea conocer el riesgo del productor de que sea rechazada una BDG buena (Ratio-Riesgo =1); en la tabla 55.1 se obtiene que:

$$n = 20 \rightarrow RP = 4,6\% [100 - 95,4 = 4,6]$$

$$n = 50 \rightarrow RP = 4,1\% [100 - 95,9 = 4,1]$$

$$n = 100 \rightarrow RP = 3,3\% [100 - 96,7 = 3,3]$$

Por el contrario, si se desea conocer el riesgo del usuario de aceptar una BDG con una exactitud real menor que la exigida (p.e.: Ratio-Riesgo = 0,8):

$$n = 20 \rightarrow RU = 54,3\%$$

$$n = 50 \rightarrow RU = 19,6\%$$

$$n = 100 \rightarrow RU = 2,1\%$$

De esta forma, con el nomograma propuesto se puede conocer en todo momento los riesgos que corren tanto usuarios como productores a la hora de contrastar la exactitud posicional aplicando el estándar EMAS.

5.5.2 NMAS

En el estándar NMAS influyen diferentes factores: escala a priori de la BDG²⁸, tolerancia aplicable del estándar y tamaño muestral. En este caso se ha trabajado partiendo de la tolerancia aplicable a cartografía a escala <1/20.000 (0,508 mm) porque, en caso contrario, el estándar tiene un comportamiento altamente permisivo (ver apartado 52.1).

Empleando el proceso de simulación descrito (figura 43.3), se han variado los niveles de tolerancia del NMAS sobre poblaciones distribuidas según $N(0,1)$. Dichos niveles de exigencia se han reducido progresivamente en intervalos de 2,5 % (tabla 55.3). De esta forma se determina la capacidad del estándar en función del tamaño muestral y del ratio de exigencia ($Tol_{NMAS} / \sigma_{Mapa}$).

²⁸ Su correspondencia sería la desviación estimada a priori (σ_{Mapa}).

Tabla 55.3: Correspondencia entre los ratios de exigencia del NMAS, la Tol_{NMAS} y el coeficiente para el nomograma.

Tolerancia NMAS (pulgadas)	Tolerancia NMAS (mm)	Ratio de exigencia NMAS	
		Mayor exactitud para la BDG (%) a escala <1/20.000	Coeficiente en el nomograma
$(1/50)''$	0,508	=	0,750
$(1,025/50)''$	0,5207	2,5 %	0,775
$(1,050/50)''$	0,5334	5,0 %	0,800
$(1,075/50)''$	0,5461	7,5 %	0,825
$(1,100/50)''$	0,5588	10,0 %	0,850
$(1,125/50)''$	0,5715	12,5 %	0,875
$(1,150/50)''$	0,5842	15,0 %	0,900
$(1,175/50)''$	0,5969	17,5 %	0,925
$(1,200/50)''$	0,6096	20,0 %	0,950
$(1,225/50)''$	0,6223	22,5 %	0,975
$(1,250/50)''$	0,6350	25,0 %	1,00
$(1,300/50)''$	0,6604	30,0 %	1,05

Mediante el proceso de simulación descrito (figura 43.4), se ha obtenido el nomograma para el estándar NMAS (figura 55.4 y tabla 55.4). Como coeficiente, se ha tomado la unidad para cartografía un 25 % más exacta que la exigida por el NMAS (a escala < 1/20.000), porque para el mínimo recomendado como tamaño muestral ($n = 20$) se obtiene un porcentaje de aceptación ≈ 95 %.

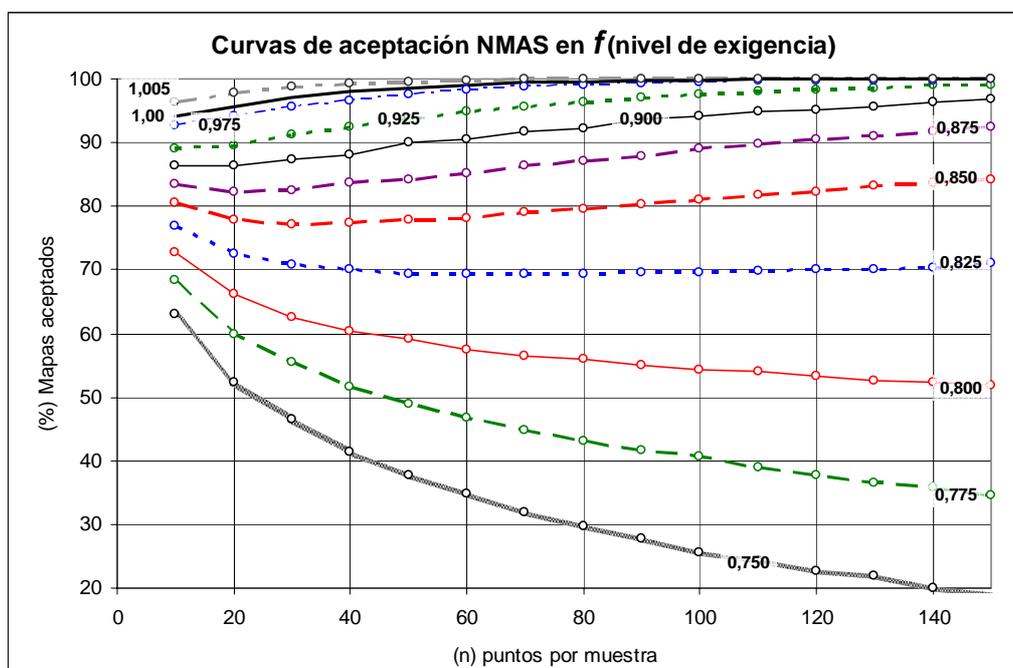


Figura 55.4: Nomograma gráfico del estándar NMAS con las curvas de aceptación en función del tamaño muestral y el ratio de exigencia.

Tabla 55.5: Nomograma analítico del estándar NMAS con los (%) de aceptación en función del tamaño muestral (n) y el ratio de exigencia.

n	Nivel de exigencia											
	=	2,5 %	5 %	7,5 %	10 %	12,5 %	15 %	17,5 %	20 %	22,5 %	25 %	30 %
Coef. NMAS	0,750	0,775	0,800	0,825	0,850	0,875	0,900	0,925	0,950	0,975	1,00	1,05
10	63,01	68,51	72,85	76,96	80,45	83,37	86,47	88,94	90,98	92,71	94,28	96,44
20	52,31	59,89	66,15	72,43	77,77	82,29	86,28	89,60	92,13	94,07	95,69	97,87
30	46,56	55,45	62,66	70,87	77,23	82,55	87,32	91,18	93,45	95,60	97,01	98,83
40	41,50	51,57	60,43	70,04	77,39	83,60	88,00	92,46	94,83	96,69	97,97	99,30
50	37,63	48,96	59,13	69,47	77,94	84,27	90,06	93,72	95,98	97,68	98,57	99,56
60	34,94	46,78	57,54	69,46	78,21	85,27	90,48	94,88	96,82	98,32	99,10	99,81
70	31,87	44,86	56,49	69,30	78,98	86,38	91,80	95,64	97,45	98,81	99,41	99,93
80	29,85	43,14	55,94	69,42	79,48	87,16	92,33	96,41	97,99	99,11	99,55	99,95
90	27,77	41,54	54,95	69,56	80,42	87,94	93,66	97,04	98,56	99,36	99,70	99,95
100	25,54	40,59	54,26	69,54	81,03	89,07	94,28	97,57	98,73	99,53	99,78	99,98
110	24,43	39,05	53,96	69,79	81,82	89,86	94,85	98,02	98,99	99,66	99,91	100,00
120	22,56	37,85	53,20	69,99	82,27	90,42	95,23	98,41	99,19	99,75	99,91	100,00
130	21,96	36,43	52,70	70,17	83,23	91,04	95,68	98,65	99,41	99,80	99,94	100,00
140	20,03	35,85	52,28	70,40	83,59	91,80	96,43	99,00	99,56	99,88	99,97	100,00
150	19,02	34,54	51,75	70,98	84,20	92,40	96,75	99,14	99,69	99,91	99,97	100,00

Sobre el caso práctico del apartado anterior (5.5.1), se ha realizado un proceso similar simulación adaptado al NMAS. Se han generado 1000 sub-muestras aleatorias de tamaño $n_m = 30, 40, 50, 60$ y 70 puntos. Partiendo de que la escala a priori del mapa es obtenida en función de la $\sigma_{EMAS-mapa}$ (4,50 m), se aplica el estándar a cada una de ellas con un valor de $Tol_{NMAS} = 0,508$ mm. con incrementos de un 2,5 %. Para que los resultados en la simulación de sub-muestras sean estables, se ha repetido el proceso 30 veces y se han obtenido los valores promedio de aceptación (tabla 55.6). Gráficamente se ha representado en el nomograma del estándar sobre la figura 55.5.

Tabla 55.6: Porcentajes de aceptación y ratio de exigencia NMAS interpolado para la cartografía analizada²⁹.

n_m	Porcentaje de mapas aceptados	Ratio de exigencia NMAS (interpolado)
30	81,82 %	1,122
40	84,33 %	1,129
50	89,41 %	1,147
60	91,77 %	1,157
70	95,33 %	1,173

²⁹ Se ha empleado: Tol_{NMAS} (0,508 mm) y escala 1/18.000 ($\sigma_{Mapa} = 4,50$ m).

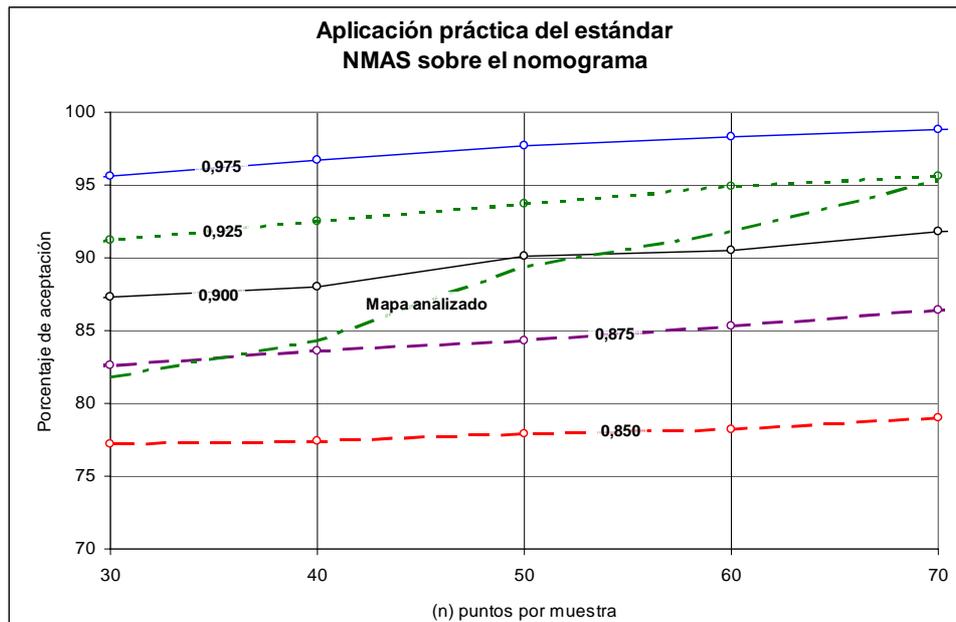


Figura 55.5: Detalle del nomograma y su aplicación práctica a un caso real con el estándar NMAS.

Se puede observar que el comportamiento de la curva obtenida empleando las submuestras del ejemplo, no sigue un mismo parámetro al cruzarse con las curvas del nomograma (a pesar de haber repetido el proceso de simulación y estabilizarse los resultados³⁰). No obstante, el valor medio obtenido es un coeficiente de 1,146, lo que reafirma la teoría de que la cartografía ha de ser, al menos, un 15 % mejor que la exigida a priori por el estándar.

El hecho de que la curva obtenida se cruce y que para que se estabilicen los resultados se deban repetir las 1000 iteraciones en 30 ocasiones (30.000 iteraciones) indica que el método no es todo lo adecuado que cabría esperar.

En cuanto a la aplicación del nomograma para conocer los riesgos de usuario y productor, en el caso del estándar NMAS también es posible. Si, por ejemplo, el productor quiere conocer qué riesgo tiene de que sea rechazada una BDG con una exactitud un 25 % mayor que la estrictamente necesaria (según el NMAS para cartografía a escala < 1/20.000):

$$\text{Capacidad} \Rightarrow \text{Ratio de Exigencia} = \frac{\text{Exigencia demandada}}{\text{Exactitud real}} = \frac{\text{Tot}_{\text{NMAS}}}{\sigma_{\text{Mapa}}} = \frac{\left(\frac{1}{50}\right)^n}{\left(\frac{1,25}{50}\right)^n} = \frac{0,508\text{mm}}{0,635\text{mm}} = 0,8$$

Entrando en la tabla 55.5 ($\text{Coef}_{\text{NMAS}} = 1$), se obtiene que:

³⁰ La variabilidad resultante del proceso de simulación ha sido de $1,146 \pm 0,021$.

$$n = 20 \rightarrow RP = 4,31\% [100 - 95,69 = 4,31]$$

$$n = 50 \rightarrow RP = 1,43\% [100 - 98,57 = 1,43]$$

$$n = 100 \rightarrow RP = 0,22\% [100 - 99,78 = 0,22]$$

Por el contrario, para conocer el riesgo del usuario de aceptar una BDG con una exactitud real igual a la exigida por el NMAS:

$$\text{Capacidad} \Rightarrow \text{Ratio de Exigencia} = \frac{\left(\frac{1}{50}\right)^n}{\left(\frac{1}{50}\right)^n} = \frac{0,508 \text{ mm}}{0,508 \text{ mm}} = 1,0 \rightarrow \text{Coef}_{\text{NMAS}} = 0,750$$

$$n = 20 \rightarrow RU = 52,31\%$$

$$n = 50 \rightarrow RU = 37,63\%$$

$$n = 100 \rightarrow RU = 25,54\%$$

Así, se puede conocer el riesgo del usuario y del productor al contrastar la exactitud posicional aplicando el estándar NMAS para cartografía a escala $<1/20.000$.

5.5.3 ASPRS

En la generación del nomograma para el ASPRS, se ha trabajado bajo la hipótesis de los experimentos *F*, *G*, *H* e *I* de la tabla 52.4. Dado que el nivel de exigencia del estándar es muy riguroso (ver apartado 5.2.3), las curvas de aceptación representan aquellas muestras que, reuniendo las características propias de una BDG Clase *I*, son clasificadas como pertenecientes a dicha clase. Así, para un tamaño muestral de 20 puntos de control, tan sólo un 29 % de las muestras se consideran de Clase *I*. Si, por el contrario, la calidad real de la BDG es un 30 % superior a la exigida, con 20 puntos de control se clasificarán el 95 % de las muestras en la Clase *I*. Así, se ha obtenido porcentaje de muestras clasificadas como pertenecientes a la Clase *I* en función del tamaño muestral y del ratio de exigencia con respecto a la cartografía real ($\sigma_{\text{ASPRS_Clase I}} \Leftrightarrow \sigma_{\text{Mapa}}$). Como resultado, se ha obtenido la figura 55.6 (y su representación analítica en la tabla 55.7).

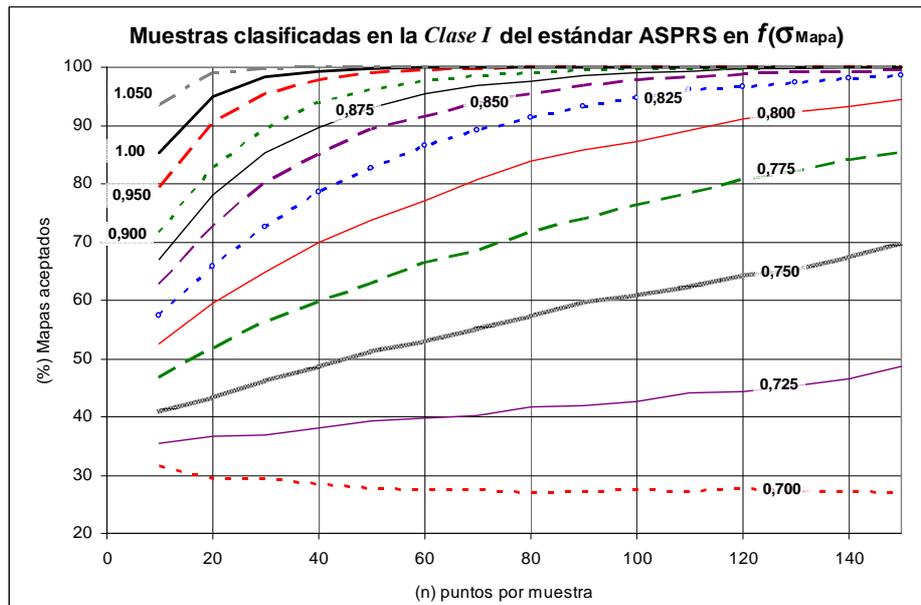


Figura 55.6: Gráfico del estándar ASPRS con las curvas del porcentaje de muestras clasificadas como pertenecientes a la Clase *I* en función del tamaño muestral y el ratio de exigencia.

Al igual que en el estándar NMAS, los coeficientes del gráfico se han establecido a partir de la determinación del porcentaje de aceptación $\approx 95\%$ (para $n = 20$).

Conforme aumenta el tamaño muestral, aumentan los porcentajes de aceptación dado que, en este caso, las curvas representan el porcentaje de muestras que, habiendo sido producidas con un nivel de exigencia igual o superior al requerido por el estándar ASPRS para la Clase *I*, son clasificadas en dicha categoría (ver la curva con el coeficiente 0,700).

Dado lo restrictivo del estándar, tras el proceso de simulación se observa que el mayor riesgo lo asume el productor. Esto es, para el ejemplo de una cartografía producida con un nivel de exigencia un 5% mejor que el estrictamente necesario de la Clase *I*, utilizando 30 puntos de control, tan sólo un 46,6% de las muestras son clasificadas en dicha categoría, mientras que un 53,4% de ellas se asignan a una categoría inferior: la Clase *II* (ver tabla 55.7).

Tabla 55.7: Porcentaje de muestras clasificadas como pertenecientes a la Clase *I* en función del tamaño muestral y el ratio de exigencia ASPRS.

<i>n</i>	Ratio de exigencia ASPRS														
	1,000	1,025	1,050	1,075	1,100	1,125	1,150	1,175	1,200	1,225	1,250	1,275	1,300	1,325	1,350
Ratio (%)	0,700	0,725	0,750	0,775	0,800	0,825	0,850	0,875	0,900	0,925	0,950	0,975	1,00	1,025	1,050
Coef. ASPRS															
10	31,68	35,41	40,95	46,77	52,54	57,39	62,56	67,08	71,68	75,65	79,16	82,52	85,27	89,56	93,43
20	29,30	36,51	43,41	51,83	59,54	65,82	72,49	78,03	82,73	87,80	90,40	92,72	94,98	97,50	98,92
30	29,28	36,77	46,30	56,39	64,74	72,58	80,31	85,26	89,28	93,36	95,43	97,19	98,24	99,28	99,80
40	28,51	38,02	48,73	59,85	69,82	78,63	85,12	89,59	93,98	96,30	97,75	98,78	99,39	99,86	99,93
50	27,73	39,30	51,34	62,98	73,70	82,62	89,41	93,07	96,07	97,98	99,01	99,58	99,79	99,93	100
60	27,43	39,80	53,05	66,48	77,21	86,56	91,60	95,40	97,48	98,97	99,54	99,86	99,91	99,97	100
70	27,38	40,13	55,17	68,44	80,70	89,23	93,92	96,96	98,49	99,52	99,78	99,91	99,96	100	100
80	27,08	41,66	57,33	71,88	83,75	91,27	95,54	97,59	99,13	99,77	99,88	99,96	100	100	100
90	27,20	41,90	59,86	74,09	85,82	93,21	96,77	98,64	99,48	99,89	99,97	99,99	100	100	100
100	27,59	42,70	60,99	76,34	87,24	94,71	97,83	99,06	99,72	99,95	99,98	100	100	100	100
110	27,26	44,17	62,38	78,41	89,15	96,04	98,34	99,33	99,78	99,96	100	100	100	100	100
120	27,63	44,43	64,26	80,72	91,03	96,73	98,68	99,64	99,87	99,99	100	100	100	100	100
130	27,28	45,36	65,29	82,19	92,25	97,46	99,20	99,79	99,92	99,99	100	100	100	100	100
140	27,28	46,51	67,49	84,02	93,26	98,10	99,28	99,91	99,95	99,99	100	100	100	100	100
150	27,05	48,56	69,91	85,31	94,53	98,61	99,60	99,90	99,99	100	100	100	100	100	100

Para facilitar la comprensión del comportamiento ascendente de las curvas, se ha realizado un planteamiento diferente: partiendo de las curvas complementarias, se han obtenido las curvas con los porcentajes de muestras que **no se asignan** como pertenecientes a la Clase *I* (ver figura 55.7 y tabla 55.8). En este caso muestran los riesgos del productor: probabilidad de que una BDG realizada con una exactitud igual o mejor que la demandada por el estándar ASPRS para la Clase *I*, sea rechazada.

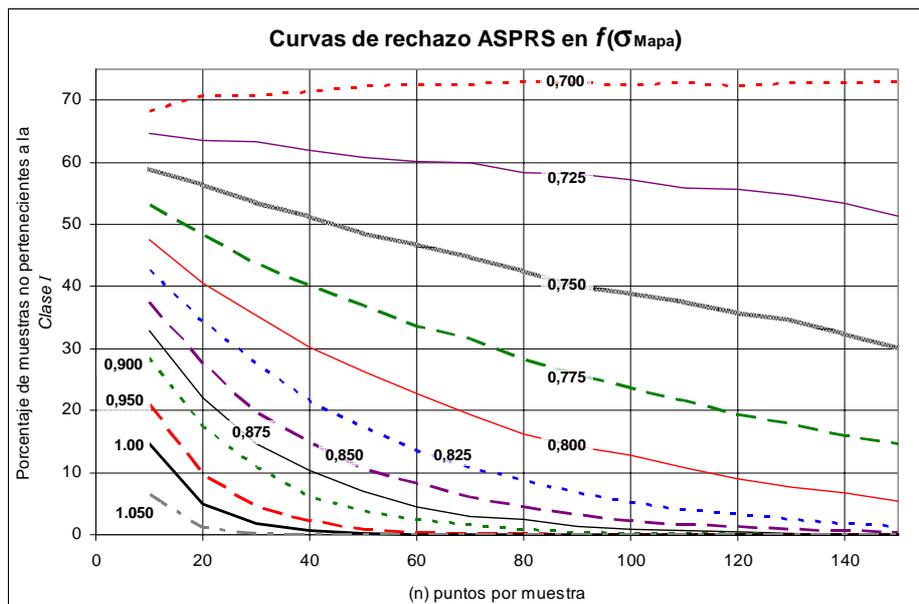


Figura 55.7: Nomograma gráfico del estándar ASPRS con las curvas de rechazo: porcentaje de muestras clasificadas como no pertenecientes a la Clase I en función del tamaño muestral y el ratio de exigencia.

Tabla 55.8: Nomograma analítico de muestras clasificadas como no pertenecientes a la Clase I en función del tamaño muestral y el ratio de exigencia ASPRS.

n	Ratio de exigencia ASPRS															
	1,000	1,025	1,050	1,075	1,100	1,125	1,150	1,175	1,200	1,225	1,250	1,275	1,300	1,325	1,350	
Coef, ASPRS	0,700	0,725	0,750	0,775	0,800	0,825	0,850	0,875	0,900	0,925	0,950	0,975	1,00	1,025	1,050	
10	68,32	64,59	59,05	53,23	47,46	42,61	37,44	32,92	28,32	24,35	20,84	17,48	14,73	10,44	6,57	
20	70,70	63,49	56,59	48,17	40,46	34,18	27,51	21,97	17,27	12,20	9,60	7,28	5,02	2,50	1,08	
30	70,72	63,23	53,70	43,61	35,26	27,42	19,69	14,74	10,72	6,64	4,57	2,81	1,76	0,72	0,20	
40	71,49	61,98	51,27	40,15	30,18	21,37	14,88	10,41	6,02	3,70	2,25	1,22	0,61	0,14	0,07	
50	72,27	60,70	48,66	37,02	26,30	17,38	10,59	6,93	3,93	2,02	0,99	0,42	0,21	0,07	0	
60	72,57	60,20	46,95	33,52	22,79	13,44	8,40	4,60	2,52	1,03	0,46	0,14	0,09	0,03	0	
70	72,62	59,87	44,83	31,56	19,30	10,77	6,08	3,04	1,51	0,48	0,22	0,09	0,04	0	0	
80	72,92	58,34	42,67	28,12	16,25	8,73	4,46	2,41	0,87	0,23	0,12	0,04	0	0	0	
90	72,8	58,10	40,14	25,91	14,18	6,79	3,23	1,36	0,52	0,11	0,03	0,01	0	0	0	
100	72,41	57,30	39,01	23,66	12,76	5,29	2,17	0,94	0,28	0,05	0,02	0	0	0	0	
110	72,74	55,83	37,62	21,59	10,85	3,96	1,66	0,67	0,22	0,04	0	0	0	0	0	
120	72,37	55,57	35,74	19,28	8,97	3,27	1,32	0,36	0,13	0,01	0	0	0	0	0	
130	72,72	54,64	34,71	17,81	7,75	2,54	0,80	0,21	0,08	0,01	0	0	0	0	0	
140	72,72	53,49	32,51	15,98	6,74	1,90	0,72	0,09	0,05	0,01	0	0	0	0	0	
150	72,95	51,44	30,09	14,69	5,47	1,39	0,4	0,10	0,01	0	0	0	0	0	0	

Para facilitar la comprensión del comportamiento descendente de la nueva familia de curvas, se muestra su comportamiento sobre el caso práctico mencionado en apartados anteriores (Hoja 926 1-1 del ICA). La desviación de partida para la Clase I

del ASPRS ha sido de $\sigma_{Mapa} = 5$ m, ligeramente superior que la mayor desviación de sus componentes ($\sigma_x = 3,4$ m; $\sigma_y = 4,7$ m). Bajo dicha hipótesis se ha efectuado la simulación en la elección de sub-muestras, obteniéndose los porcentajes de muestras que no pertenecen a la Clase I (tabla 55.9) y su interpolación sobre el nomograma ASPRS (figura 55.8). Asimismo, se ha repetido todo el proceso en 30 ocasiones para estabilizar el resultado final³¹.

Tabla 55.9: Porcentajes de aceptación / rechazo y ratio de exigencia ASPRS interpolado para la cartografía analizada ($\sigma_{Mapa} = 5$ m).

n_m	Porcentaje de muestras \in Clase I	Porcentaje de muestras \notin Clase I	Ratio de exigencia NMAS interpolado
30	73,2 %	26,8 %	1,127
40	76,1 %	23,9 %	1,118
50	80,6 %	19,4 %	1,119
60	85,9 %	14,1 %	1,123
70	91,6 %	8,4 %	1,138

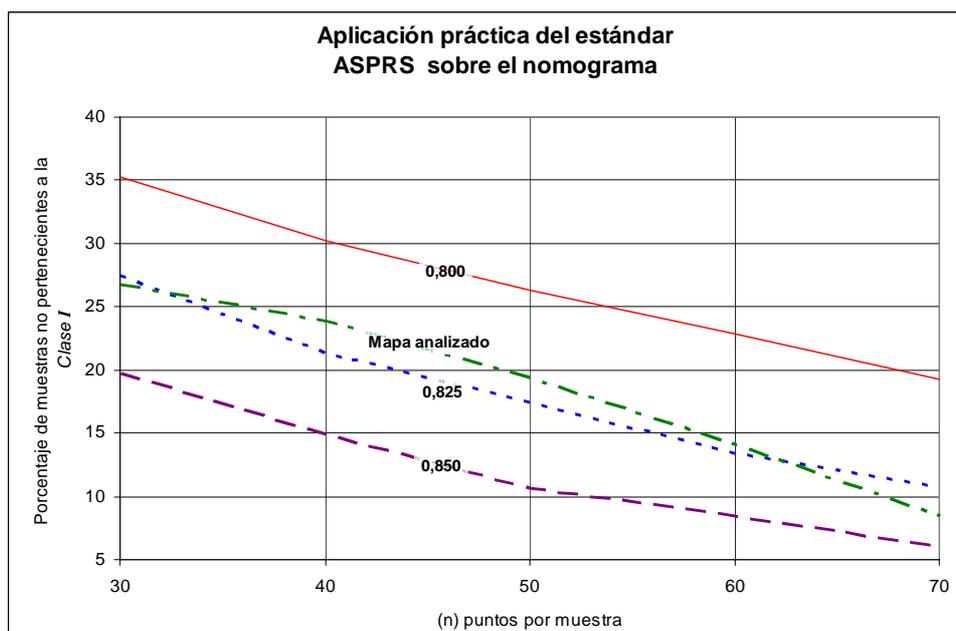


Figura 55.8: Detalle del nomograma y su aplicación práctica a un caso real con el estándar ASPRS.

En la interpolación la curva del ejemplo se cruza con las del nomograma, aunque en menor cuantía que en el caso anterior (NMAS). Los resultados promediados tras la interpolación, muestran un valor de 1,133 como ratio de exigencia del ASPRS. Esto es, la cartografía analizada es un 13 % más exacta que la $\sigma_{Mapa-estimada}$, por lo que ésta

³¹ La variabilidad resultante del proceso de simulación ha sido de $1,133 \pm 0,022$.

será de $\sigma_{\text{Mapa}} = 4,40$ m (valor que coincide sensiblemente con el obtenido sobre el nomograma del estándar EMAS).

En cuanto a la aplicación de este nuevo nomograma para determinar los riesgos de usuario y productor (ASPRS), su utilización es muy similar a los casos anteriores, teniendo en cuenta que el nomograma de la figura 55.7 (tabla 55.8) representa el porcentaje de muestras que no son asignadas a la Clase I. Para un productor que desea conocer el riesgo que tiene de rechazo de una BDG con una exactitud, por ejemplo, un 30 % más exacta que la estrictamente necesaria (según la *Clase I* ASPRS), será:

$$\text{Capacidad} \Rightarrow \text{Ratio de Exigencia} = \frac{\text{Exigencia demandada}}{\text{Exactitud real}} = \frac{\sigma_{\text{ASPRS}}}{\sigma_{\text{Mapa}}} = \frac{1,30 \text{ m}}{1,00 \text{ m}} = 1,30$$

Entrando en la tabla 55.8 ($\text{Coef}_{\text{ASPRS}} = 1$):

$$n = 20 \rightarrow RP = 5,02\%$$

$$n = 50 \rightarrow RP = 0,21\%$$

$$n = 100 \rightarrow RP = 0,00\%$$

Por el contrario, para conocer el riesgo del usuario de aceptar una BDG con una exactitud real igual o inferior a la exigida por la *Clase I* ASPRS:

$$\text{Capacidad} \Rightarrow \text{Ratio de Exigencia} = 1,0 \rightarrow \text{Coef}_{\text{ASPRS}} = 0,700$$

$$n = 20 \rightarrow RU = 29,30\% \quad [100 - 70,70 = 29,30]$$

$$n = 50 \rightarrow RU = 27,73\% \quad [100 - 72,27 = 27,73]$$

$$n = 100 \rightarrow RU = 27,59\% \quad [100 - 72,41 = 27,59]$$

5.5.4 NSSDA

Dadas las características propias del estándar NSSDA, a priori no pueden generarse nomogramas con los porcentajes de aceptación en función del tamaño muestral, dado que el resultado proporcionado el estándar es un valor de exactitud horizontal a un determinado nivel de confianza.

No obstante, para estudiar su comportamiento ante distintos tipos de población, se han variado las condiciones en la distribución de errores: se han generado poblaciones aleatorias distribuidas según una $N(0, 0.8)$, $N(0, 0.85)$, $N(0, 0.9)$, $N(0, 0.95)$, $N(0, 1)$, $N(0, 1.05)$, $N(0, 1.1)$, $N(0, 1.15)$ y $N(0, 1.2)$. Para cada serie de poblaciones y sus correspondientes sub-muestras, se han obtenido los valores promedio de *Exactitud_r*,

NSSDA en función de σ_{Mapa} y el tamaño muestral. Los resultados se muestran en las figuras 55.9 y 55.10, así como en la tabla 55.10.

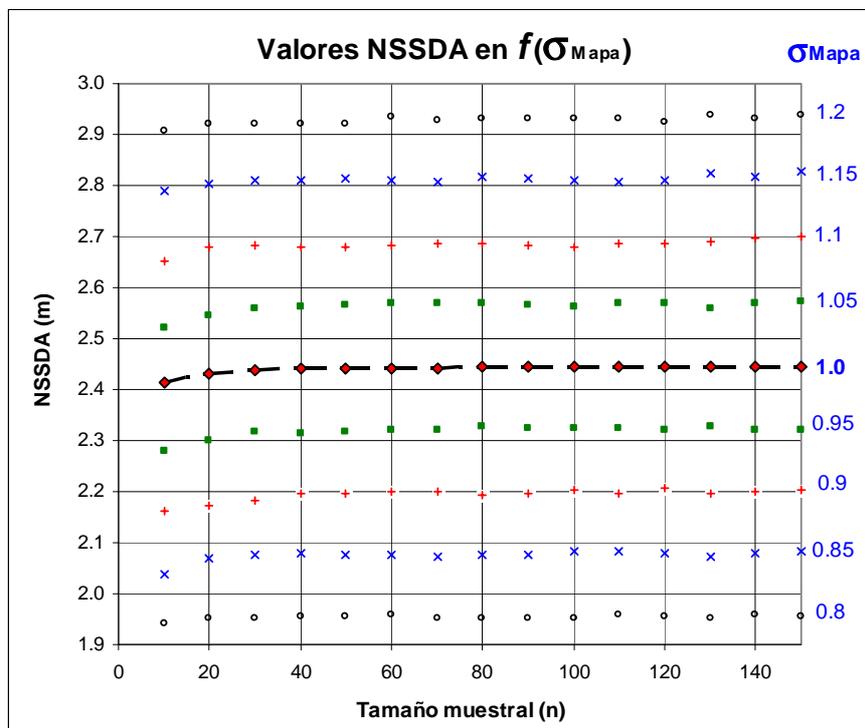


Figura 55.9: Promedios de *Exactitud*, NSSDA en función de σ_{Mapa} y el tamaño muestral³².

Se observa que los valores medios de exactitud posicional para el NSSDA, al modificar la σ_{Mapa} , se repiten de forma sensiblemente paralela al caso de $N(0,1)$ (ver apartado 5.2.4), con ligeras alteraciones en función de la propia variabilidad de los resultados. Ésta se mantiene constante respecto a los valores medios de exactitud con los mismos resultados que los obtenidos en la figura 52.9. Asimismo, para todos los casos la variabilidad del NSSDA se aproxima al 5 % del valor de exactitud poblacional cuando se emplean 100 puntos de control.

³² En la margen derecha aparece el valor de la desviación poblacional (σ_{Mapa}), y en trazado discontinuo muestran los valores obtenidos para una población $N(0,1)$ (ver figura 52.8).

Tabla 55.10: Promedios de $Exactitud_r$ NSSDA en función de la σ_{Mapa} y el tamaño muestral

n	σ_{Mapa}								
	0,80	0,85	0,90	0,95	1,00	1,05	1,10	1,15	1,20
10	1,941	2,038	2,163	2,280	2,414	2,520	2,652	2,788	2,906
20	1,951	2,068	2,172	2,300	2,432	2,546	2,678	2,804	2,921
30	1,950	2,075	2,184	2,317	2,438	2,558	2,683	2,811	2,920
40	1,955	2,078	2,195	2,315	2,440	2,562	2,678	2,809	2,920
50	1,956	2,077	2,198	2,317	2,442	2,565	2,680	2,814	2,921
60	1,958	2,077	2,199	2,319	2,443	2,568	2,684	2,811	2,934
70	1,952	2,073	2,198	2,320	2,443	2,570	2,686	2,808	2,928
80	1,953	2,075	2,194	2,326	2,445	2,568	2,686	2,818	2,932
90	1,952	2,077	2,198	2,324	2,444	2,566	2,683	2,814	2,932
100	1,952	2,083	2,203	2,324	2,445	2,564	2,678	2,811	2,933
110	1,958	2,082	2,196	2,324	2,445	2,570	2,686	2,807	2,930
120	1,955	2,079	2,207	2,322	2,444	2,568	2,686	2,810	2,924
130	1,953	2,073	2,198	2,328	2,446	2,560	2,691	2,823	2,937
140	1,959	2,079	2,199	2,321	2,446	2,567	2,696	2,819	2,932
150	1,957	2,083	2,203	2,321	2,445	2,573	2,701	2,826	2,938

La distancia existente entre las diferentes curvas aparece representada en la figura 55.10. En ella se puede observar el valor de la $Exactitud_r$ NSSDA poblacional en función de σ_{Mapa} . Lógicamente, siguen una progresión directamente proporcional a la σ_{Mapa} , ajustada a la recta:

$$Exactitud_r \text{ NSSDA} = 2,447 \cdot \sigma_{Pobl.} - 0,00056 \quad (R^2 = 1)$$

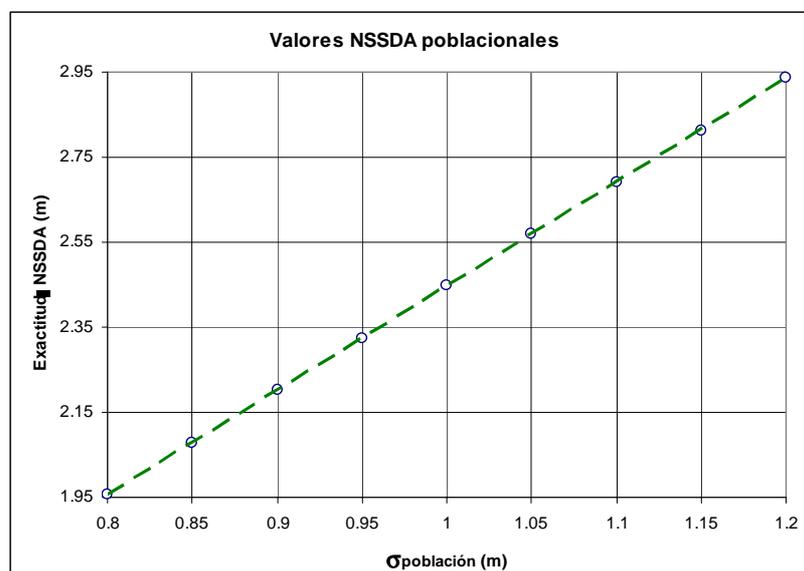


Figura 55.10: Recta de ajuste y valores del estándar NSSDA para la población total en función de σ_{Mapa} .

Respecto a los riesgos de usuario y productor, se mantienen constantes independientemente de las variaciones que se puedan producir sobre σ_{Mapa} . Tan sólo podría hablarse de modificaciones al emplear como referencia la relación porcentual entre el valor medio de la $Exactitud_r$ NSSDA y las curvas de variabilidad por encima y

por debajo de σ_{Mapa} con distintos niveles de confianza (ver figura 52.10: zona RP entre A y B , y zona RU entre B y D). Así, modificando los valores de k (nivel de confianza: $k \cdot \sigma$), se obtienen los coeficientes:

$$Coef_{RP} = \frac{RP}{Exactitud_{r,NSSDA}} (\%) = \left(\frac{A_m B_m}{Exactitud_{r,Poblac.}} \right) \cdot 100 (\%)$$

$$Coef_{RU} = \frac{RU}{Exactitud_{r,NSSDA}} (\%) = \left(\frac{B_m D_m}{Exactitud_{r,Poblac.}} \right) \cdot 100 (\%)$$

Basándose en dichos coeficientes, pueden obtenerse curvas de variabilidad para los valores medios del estándar NSSDA con distintos niveles de confianza. Los coeficientes obtenidos se muestran en la tabla 55.11 y la figura 55.11.

Tabla 55.11: Coeficientes $Coef_{RP}$ y $Coef_{RU}$ en función del tamaño muestral y del nivel de confianza en variabilidad de $Exactitud_{r,NSSDA}$.

n	$Coef_{RP}$ (%)								$Coef_{RU}$ (%)								
	α	0,990	0,975	0,950	0,900	0,800	0,750	0,683	0,500	0,990	0,975	0,950	0,900	0,800	0,750	0,683	0,500
Z		2,576	2,241	1,960	1,645	1,282	1,150	1,000	0,674	2,576	2,241	1,960	1,645	1,282	1,150	1,000	0,674
10		38,94	33,72	29,33	24,41	18,74	16,69	14,34	9,26	41,48	36,26	31,86	26,94	21,27	19,22	16,87	11,80
15		31,98	27,72	24,13	20,11	15,48	13,81	11,89	7,74	33,70	29,44	25,85	21,83	17,20	15,53	13,60	9,46
20		27,81	24,12	21,01	17,54	13,53	12,08	10,42	6,83	29,03	25,34	22,24	18,76	14,75	13,31	11,64	8,06
25		24,84	21,54	18,77	15,67	12,09	10,80	9,31	6,11	25,90	22,61	19,83	16,73	13,15	11,86	10,38	7,17
30		22,64	19,65	17,13	14,31	11,06	9,89	8,54	5,63	23,46	20,47	17,95	15,13	11,88	10,70	9,35	6,45
35		21,00	18,23	15,89	13,28	10,26	9,18	7,92	5,23	21,74	18,96	16,63	14,01	11,00	9,91	8,66	5,96
40		19,57	16,99	14,81	12,38	9,57	8,56	7,39	4,88	20,22	17,64	15,47	13,03	10,23	9,21	8,05	5,54
45		18,45	16,02	13,97	11,68	9,04	8,08	6,99	4,62	19,02	16,59	14,54	12,25	9,61	8,65	7,56	5,19
50		17,44	15,14	13,21	11,05	8,55	7,65	6,62	4,39	17,93	15,63	13,70	11,54	9,04	8,14	7,11	4,88
55		16,60	14,41	12,57	10,51	8,13	7,28	6,29	4,17	17,09	14,90	13,06	11,00	8,62	7,77	6,78	4,66
60		15,90	13,81	12,05	10,08	7,81	6,99	6,05	4,01	16,31	14,22	12,46	10,49	8,22	7,40	6,45	4,42
65		15,27	13,26	11,57	9,68	7,49	6,71	5,80	3,85	15,68	13,67	11,98	10,09	7,90	7,11	6,21	4,26
70		14,64	12,71	11,09	9,27	7,18	6,42	5,56	3,68	15,05	13,12	11,50	9,68	7,59	6,83	5,96	4,09
75		14,15	12,29	10,73	8,98	6,96	6,23	5,39	3,59	14,48	12,62	11,06	9,31	7,29	6,56	5,72	3,91
80		13,63	11,84	10,33	8,64	6,70	5,99	5,19	3,45	13,95	12,16	10,66	8,97	7,02	6,32	5,51	3,77
85		13,21	11,47	10,01	8,37	6,49	5,81	5,02	3,34	13,53	11,80	10,34	8,70	6,81	6,13	5,35	3,66
90		12,80	11,19	9,77	8,17	6,33	5,67	4,90	3,25	12,90	11,52	10,10	8,50	6,66	5,99	5,23	3,58
95		12,30	10,80	9,57	8,01	6,21	5,57	4,82	3,21	12,60	11,21	9,81	8,26	6,46	5,81	5,07	3,46
100		11,98	10,41	9,09	7,61	5,90	5,28	4,58	3,05	12,23	10,66	9,33	7,85	6,15	5,53	4,82	3,29
110		11,35	9,86	8,61	7,20	5,59	5,00	4,33	2,88	11,60	10,11	8,85	7,45	5,83	5,25	4,58	3,13
125		10,30	8,95	7,81	6,53	5,06	4,53	3,92	2,61	10,54	9,19	8,05	6,78	5,31	4,78	4,17	2,85
150		9,22	8,02	7,01	5,87	4,57	4,10	3,55	2,38	9,30	8,10	7,09	5,96	4,65	4,18	3,64	2,47

Se observa que siempre se mantiene $Coef_{RP} > Coef_{RU}$ puesto que el valor medio de la $Exactitud_{r,NSSDA} < Exactitud_{r,Poblacional}$. Así, para un 95 % de nivel de confianza y con 100 puntos de control, el riesgo del productor a que el valor medio de exactitud proporcionado por el estándar NSSDA sea superior al de la BDG es del 49,3% (ver tabla 52.10), mientras que el riesgo del usuario de que dicho valor sea inferior al de la BDG es del 50,7%. Si ahora se entra en la tabla 55.11, puede observarse que la relación existente entre el valor medio proporcionado por el NSSDA con respecto al

riesgo del productor es de un 9,09%, mientras que para el usuario es del 9,33 %. Si se emplean 20 puntos de control (el mínimo recomendado por el estándar), la relación entre ambos coeficientes aumenta al 21,01 % y 22,24 % respectivamente.

Aunque tanto en la tabla como en la figura 55.11 se han representado valores máximos de $n = 150$, los cálculos se han efectuado para tamaños muestrales de hasta $n = 700$ puntos de control, obteniéndose coeficientes de relación del 5 % al 95 % de nivel de confianza con tamaños muestrales de $n = 275$ puntos de control ($Coef_{RP} \approx Coef_{RU} \approx 5\%$).

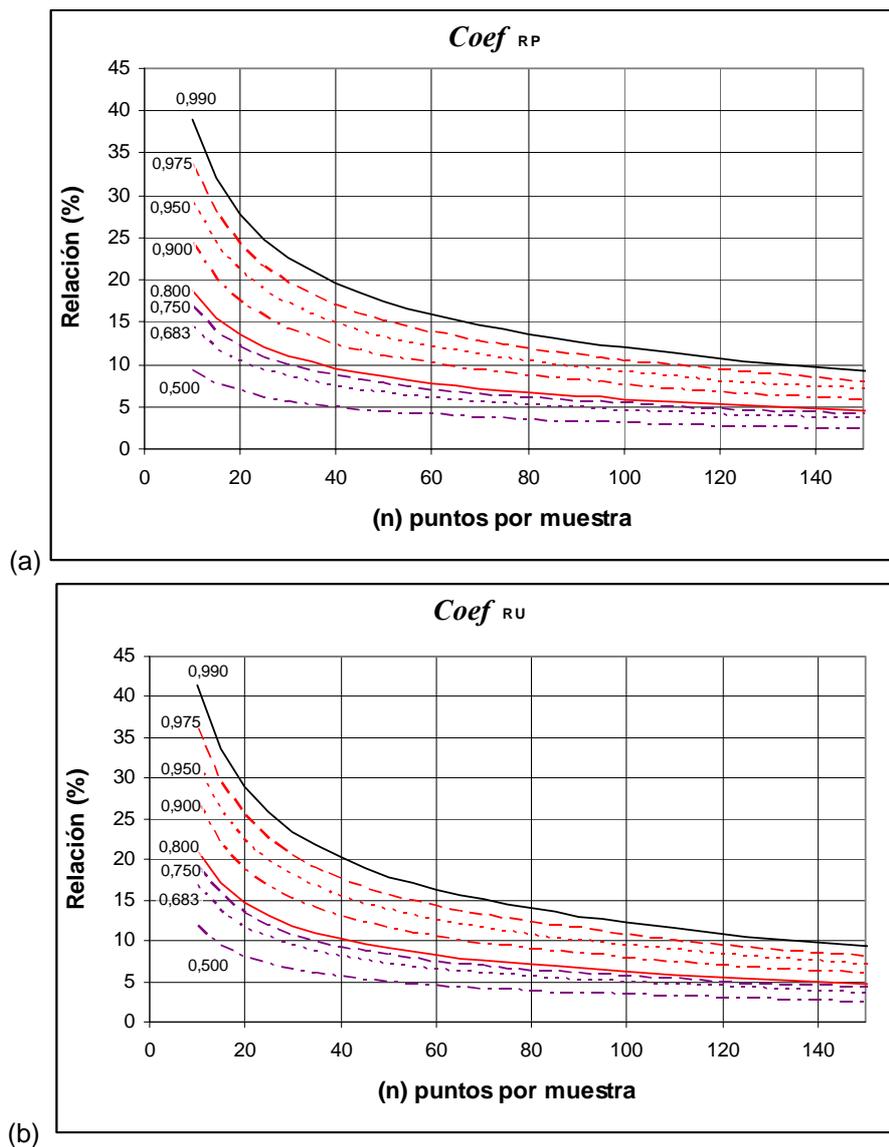


Figura 55.11: Coeficientes $Coef_{RP}$ (a) y $Coef_{RU}$ (b) en función del tamaño muestral y del nivel de confianza para $k \cdot \sigma = 0,990; 0,975; 0,950; 0,900; 0,800; 0,750; 0,683; y 0,500$ ⁽³³⁾

³³ Se incluye el valor de probabilidad de 0,683 porque $Z_{0,683} = 1$. En el eje vertical se muestra la relación porcentual existente entre el valor medio proporcionado por el NSSDA y el riesgo del productor.

Para poder generar un nomograma similar a los presentados en anteriores estándares (nomograma de aceptación / rechazo), éste debería incluir una regla de decisión para la aceptación o no de las muestras. En este caso, se han establecido las siguientes reglas de decisión:

$$R_1: \text{Exactitud}_r \text{ Exigida} \geq \text{Exactitud}_r \text{ NSSDA} \xrightarrow{\text{Entonces}} \text{Mapa Aceptado}$$

$$R_2: \text{Exactitud}_r \text{ Exigida} < \text{Exactitud}_r \text{ NSSDA} \xrightarrow{\text{Entonces}} \text{Mapa Rechazado}$$

Bajo ellas, serán rechazadas aquellas muestras cuya exactitud sea inferior a la exigida, siendo ésta establecida en función de la σ_{Mapa} estimada o exigida a priori. La simulación se ha basado en la generación de poblaciones aleatorias $N(0,1)$ y, tras aplicar el estándar NSSDA, verificar la aceptación / rechazo de 1000 sub-muestras aleatorias con distintos ratios de exigencia (ver tabla 55.12).

Tabla 55.12: Ratios de exigencia R_I aplicados en el estándar NSSDA (valores en m).

σ_{Mapa}	$\text{Exactitud}_r \text{ Mapa}$	Ratio de exigencia R_I	Umbral límite de $\text{Exactitud}_r \text{ NSSDA}$ para R_I
1,0	2,446	0,80	1,957
		0,85	2,079
		0,90	2,201
		0,95	2,324
		1,00	2,446
		1,05	2,568
		1,10	2,691
		1,15	2,813
		1,20	2,935

Dichos niveles de exigencia se han generado a partir de la desviación poblacional ($\sigma_{\text{Mapa}} = 1$), con incrementos de $\pm 0,5$ % hasta alcanzar el 20 %. Como puede observarse en los resultados obtenidos (ver nomograma de la figura 55.12 y tabla 55.13), para una población con el mismo ratio de exigencia para la regla R_I del estándar NSSDA ($\sigma_{\text{Mapa}} = 1$), la probabilidad de aceptación es ligeramente superior al 50 %. Ello viene confirmado porque el RU y el RP son muy similares y próximos a dicho valor (ver apartado 5.2, tabla 52.10).

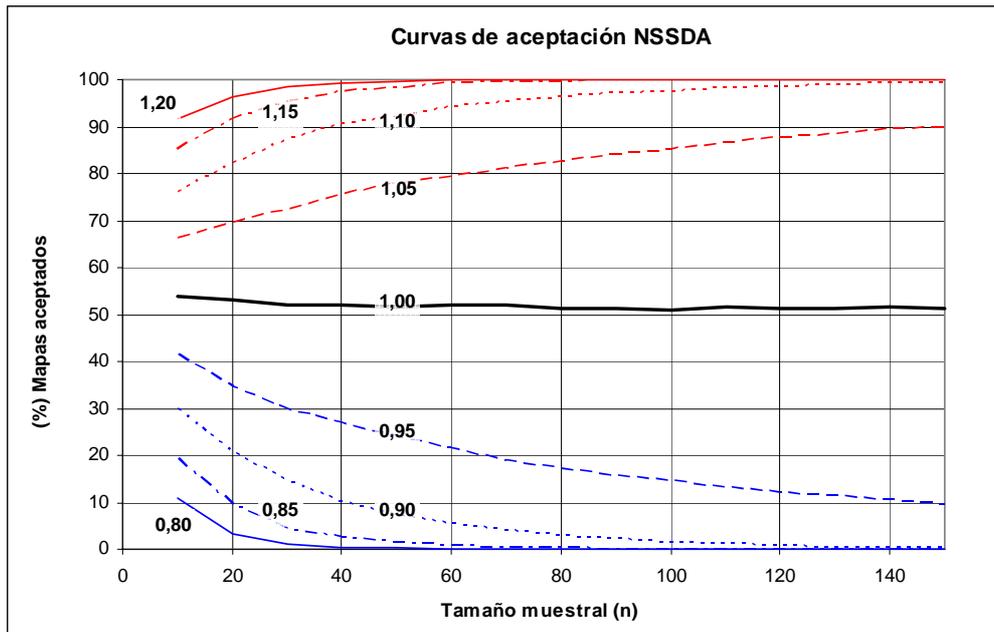


Figura 55.12: Nomograma con el porcentaje medio de mapas aceptados en función del ratio de exigencia NSSDA y el tamaño muestral.

Sin embargo, cuando se produce una mínima modificación sobre las condiciones de aceptación R_i , los porcentajes de aceptación cambian sustancialmente y, para el tamaño muestral recomendado de 100 puntos de control, tan sólo se acepta un 14 % de las muestras que tengan una *Exactitud*, NSSDA un 5 % menor que la exigida (y en un 85 % para las que son un 5 % más exactas).

Tabla 55.13: Porcentaje medio de mapas aceptados en función del ratio de exigencia NSSDA y el tamaño muestral

<i>n</i>	Ratio de exigencia para R_1 NSSDA								
	0.80	0.85	0.90	0.95	1.00	1.05	1.10	1.15	1.20
10	10.8	19.3	29.8	41.4	53.9	66.3	76.1	85.1	91.5
20	3.3	9.5	20.6	34.4	53.0	69.4	82.4	91.5	96.4
30	1.1	4.5	14.5	29.8	52.2	72.5	87.3	95.4	98.7
40	0.3	2.7	10.3	26.9	52.0	75.6	90.4	97.4	99.3
50	0.2	1.4	7.3	24.1	51.7	77.8	92.2	98.2	99.8
60	0.1	0.8	5.4	21.4	51.9	79.3	94.2	99.2	100.0
70	0.0	0.4	4.0	19.0	51.8	81.0	95.2	99.6	100.0
80	0.0	0.3	2.9	17.2	51.4	82.6	96.4	99.8	100.0
90	0.0	0.1	2.2	15.7	51.3	84.0	97.1	99.9	100.0
100	0.0	0.1	1.4	14.5	50.9	85.2	97.6	100.0	100.0
110	0.0	0.0	1.1	13.0	51.7	86.4	98.2	100.0	100.0
120	0.0	0.0	0.8	12.1	51.3	87.7	98.6	100.0	100.0
130	0.0	0.1	0.5	11.2	51.1	88.4	99.0	100.0	100.0
140	0.0	0.0	0.3	10.5	51.5	89.3	99.3	100.0	100.0
150	0.0	0.0	0.2	9.6	51.1	89.9	99.4	100.0	100.0

Dado que se ha constatado que uno de los principales inconvenientes del estándar es el hecho de que infravalora el valor de la $Exactitud_r$, especialmente para tamaños de muestra inferiores a 100 puntos de control, cabría efectuarse la siguiente pregunta:

¿Sería posible estimar el valor de la $Exactitud_r$ poblacional y, por tanto la $\sigma_{poblacional}$, con muestras de menos de 100 puntos de control?

Para poder realizar este tipo de estimaciones, a partir de la tabla 52.5 se ha obtenido la tabla 55.14 con los coeficientes de corrección, en función del tamaño muestral, para de relación entre el valor de $Exactitud_r$ NSSDA calculado y la $Exactitud_r$ Poblacional:

$$Coef_Corrección_{NSSDA}(n) = \frac{Exactitud_r_{NSSDA}}{Exactitud_r_{Poblacional}}(n)$$

Si se analizan los datos de los ejemplos anteriores (Hoja 926 1-1 del ICA), se obtiene que, tras realizar un proceso de simulación en la elección de 1000 sub-muestras de distinto tamaño muestral, los valores promedio del estándar NSSDA y sus desviaciones típicas son los que aparecen en la tabla 55.15.

Para obtener el valor de la *Exactitud_r Poblacional*, tan sólo habrá que aplicar el coeficiente de corrección a los valores medios estimados para cada tamaño muestral. Asimismo, si se desea obtener la $\sigma_{poblacional}$, ha de dividirse el valor de la *Exactitud_r NSSDA* estimada en el nomograma (tabla 55.14) por el valor medio estimado en la simulación de las sub-muestras (tabla 55.15). Los resultados (mostrados en la tabla 55.16), ofrecen una gran robustez en la obtención de ambos valores: *Exactitud_r Poblacional* y σ_{Mapa} .

Tabla 55.14 – Coeficiente de corrección para el valor medio de *Exactitud_r NSSDA*.

<i>n</i>	<i>Exactitud_r NSSDA</i> (m)	<i>Exactitud_r Poblacional</i> (m)	<i>Coef_Corrección NSSDA</i>
10	2,416	2,447	1,01296
15	2,426	2,447	1,00857
20	2,432	2,447	1,00629
25	2,434	2,447	1,00526
30	2,437	2,447	1,00431
35	2,438	2,447	1,00377
40	2,439	2,447	1,00336
45	2,440	2,447	1,00283
50	2,441	2,447	1,00262
55	2,441	2,447	1,00238
60	2,442	2,447	1,00221
65	2,442	2,447	1,00205
70	2,442	2,447	1,00192
75	2,443	2,447	1,00184
80	2,443	2,447	1,00180
85	2,443	2,447	1,00160
90	2,443	2,447	1,00156
95	2,444	2,447	1,00180
100	2,444	2,447	1,00143

Tabla 55.15: Valores promedio de *Exactitud_r NSSDA* y su desviación típica tras el proceso de simulación.

<i>n_m</i>	<i>Exactitud_r NSSDA</i> (m)	Desviación típica (m)
20	11,3643	1,199
30	11,3853	0,936
40	11,4063	0,766
50	11,4156	0,626
60	11,4230	0,491
70	11,4091	0,394
80	11,4058	0,318

Hoja 926 1-1 del ICA.

n_m	<i>Exactitud_r Poblacional</i> estimada (m)	Desviación típica poblacional estimada σ_{Mapa} (m)
20	11,436	4,673
30	11,434	4,673
40	11,445	4,677
50	11,446	4,677
60	11,448	4,679
70	11,431	4,671
80	11,426	4,670

Como prueba de la robustez de los resultados, las desviaciones típicas durante el proceso de simulación de la *Exactitud_r Poblacional* y σ_{Mapa} estimadas con respecto al tamaño muestral, son mínimas en ambos casos (0,008 y 0,003 respectivamente). Asimismo, el valor calculado σ_{Mapa} se corresponde sensiblemente con el proporcionado por otros estándares (p.e.: EMAS).

Los procesos de simulación en la elección de sub-muestras, junto con los coeficientes de la tabla 55.14, permiten obtener un valor fiable de la *Exactitud_r Poblacional* con tamaños muestrales inferiores a 100 puntos de control. No obstante, y para que la metodología propuesta funcione correctamente, se obtienen mejores resultados cuanto mayor es el número de puntos de control.

5.5.5 Estimadores circulares

Por último lugar, se ha estudiado la posibilidad de aplicar curvas operativas similares a las obtenidas en los estándares de estadística clásica (NMAS, EMAS...) para el caso de los estimadores circulares de Ómnibus, Mardia o la determinación de la dirección dominante. Después de los óptimos resultados obtenidos en el apartado 5.4, se han efectuado los procesos de simulación de forma conjunta con un contraste del tipo t-Student al 95% de nivel de confianza para comparar los estimadores circulares con los de estadística clásica.

El proceso de simulación se ha basado en la generación de poblaciones de errores artificiales del 1000 puntos distribuidos según una $N(0,1)$. Éstas han sido contaminadas con errores sistemáticos por traslación en X e Y incrementados en

tramos de 0,025 hasta llegar a 0,375 ⁽³⁴⁾. La contaminación es similar a la realizada en los apartados 5.3 y 5.4: se efectúa una traslación sobre el error posicional en sus componentes (X e Y), de tal forma que se introduce un sistematismo direccional que es conocido y controlado previamente.

Una vez contaminados los datos, han sido seleccionadas 1000 sub-muestras aleatorias de diferente tamaño aplicando en cada una de ellas los contrastes t-Student, Ómnibus y Mardia. Los resultados obtenidos se muestran en:

- Contraste t-Student para X e Y al 95 % de nivel global de confianza: figura 55.13 y tabla 55.17.
- Contraste Ómnibus al 95 % de nivel de confianza: figura 55.14 y tabla 55.18.
- Contraste de Mardia al 95 % de nivel de confianza: figura 55.15 y tabla 55.19.

El coeficiente que aparece sobre cada curva, tanto en los nomogramas gráficos como en la identificación de las curvas de los nomogramas analíticos para todos los contrastes, representa la relación existente entre la desviación poblacional ($\sigma_{\text{mapa}} = 1$) y la magnitud del sistematismo introducido artificialmente:

$$\text{Error sistemático} : E_S = E_{SX} = E_{SY} \rightarrow \text{Coeficiente} : \text{Coef}_{SIST} = 1 - E_S$$

$$\text{Ejemplos} : \begin{cases} \text{Coef}_{SIST} = 0,9m = 1 - E_S \rightarrow E_S = 0,1m \\ \text{Coef}_{SIST} = 0,825m = 1 - E_S \rightarrow E_S = 0,175m \end{cases}$$

³⁴ El valor de 0,375 (Sistemismo_x = Sistemismo_y = 0,375) supone un sistematismo por traslación de $(0,375^2 + 0,375^2)^{1/2} = 0,53$ en la dirección de 50° sobre una población $\sigma_{\text{mapa}} = 1 \rightarrow (1^2 + 1^2)^{1/2} = 1,41$ (contaminación superior al 37 % de la σ_{mapa}).

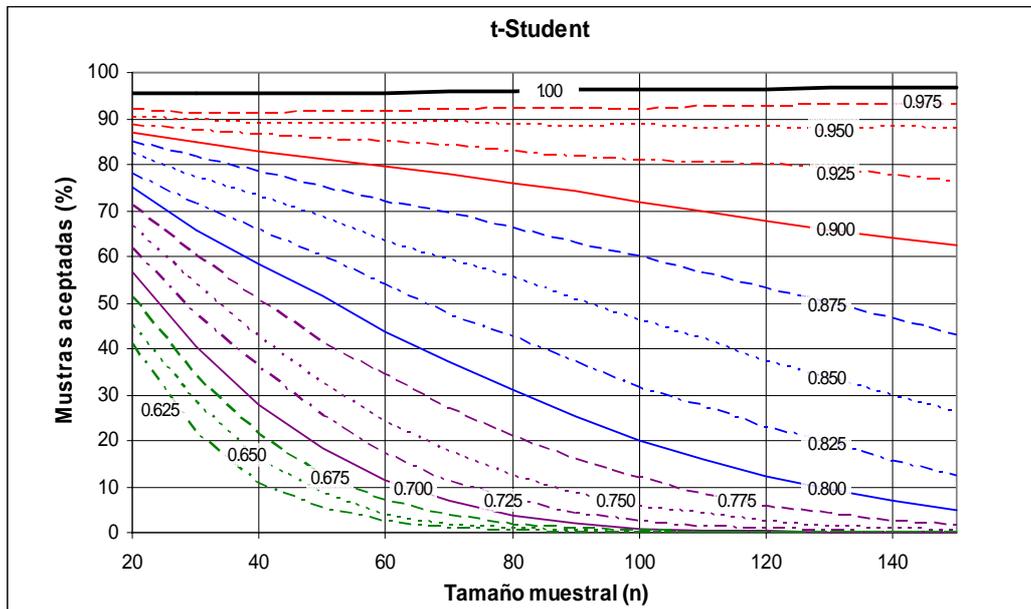


Figura 55.13: Nomograma gráfico de aceptación para una prueba t-Student en función del tamaño muestral y el coeficiente de contaminación por errores sistemáticos de traslación.

Tabla 55.17: Nomograma analítico con porcentajes de aceptación para un contraste del tipo t-Student.

n	t-Student: Coeficiente de contaminación sistemática por traslación															
	1,000	0,975	0,950	0,925	0,900	0,875	0,850	0,825	0,800	0,775	0,750	0,725	0,700	0,675	0,650	0,625
20	95,3	91,8	90,0	88,7	87,1	85,0	82,3	77,8	75,0	71,2	66,4	61,6	56,7	50,9	45,1	40,9
30	95,4	91,2	89,6	87,4	84,9	81,7	77,2	71,5	65,9	60,0	53,7	47,2	40,6	34,0	28,2	21,7
40	95,7	91,2	89,1	86,6	82,9	78,4	73,2	65,7	58,5	50,7	42,3	35,7	27,8	21,1	15,7	10,8
50	95,5	91,5	89,1	85,8	81,3	75,0	68,5	60,0	51,3	41,4	32,4	25,3	18,3	12,1	8,6	5,4
60	95,6	91,6	89,1	85,1	79,6	72,0	63,4	53,7	43,9	34,3	24,2	17,2	11,4	6,9	3,8	2,5
70	95,9	91,9	89,2	84,1	78,0	69,3	59,3	47,5	37,1	26,8	17,7	11,2	6,9	3,5	1,8	0,8
80	96,1	92,2	88,6	82,9	76,1	66,0	55,4	42,4	31,0	20,8	12,4	7,2	3,6	1,7	0,9	0,3
90	96,2	92,2	88,2	81,8	74,2	62,9	50,4	37,3	25,5	15,9	8,6	4,2	2,0	0,8	0,3	0,1
100	96,5	91,8	88,4	81,0	72,0	60,0	46,1	31,5	20,1	11,9	5,8	2,5	1,0	0,3	0,1	0,0
110	96,2	92,6	87,9	80,4	70,0	56,2	41,9	27,3	15,7	8,0	3,9	1,3	0,4	0,1	0,0	0,0
120	96,4	92,8	88,0	79,8	67,8	53,2	37,1	22,9	12,3	5,8	2,4	0,7	0,3	0,0	0,0	0,0
130	96,9	92,9	87,6	79,0	65,9	49,8	33,6	19,4	9,3	4,0	1,2	0,4	0,0	0,0	0,0	0,0
140	96,6	93,1	88,1	77,6	64,3	46,6	29,9	15,6	7,1	2,4	0,7	0,1	0,0	0,0	0,0	0,0
150	96,9	92,9	87,9	76,2	62,5	42,9	26,0	12,3	5,0	1,5	0,4	0,1	0,0	0,0	0,0	0,0

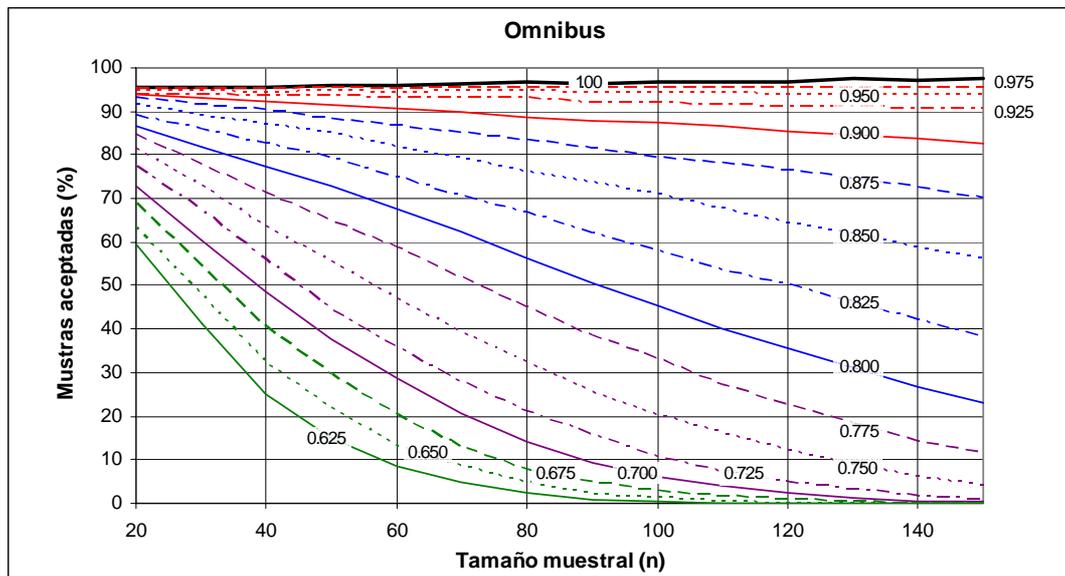


Figura 55.14: Nomograma gráfico de aceptación para una prueba Ómnibus en función del tamaño muestral y el coeficiente de contaminación por errores sistemáticos de traslación.

Tabla 55.18: Nomograma analítico con porcentajes de aceptación para un contraste Ómnibus.

n	Ómnibus: Coeficiente de contaminación sistemática por traslación															
	1,000	0,975	0,950	0,925	0,900	0,875	0,850	0,825	0,800	0,775	0,750	0,725	0,700	0,675	0,650	0,625
20	95,6	95,1	94,6	94,0	93,8	93,0	91,4	89,1	86,6	84,6	81,3	77,4	72,8	68,7	63,0	59,7
30	95,7	95,2	94,7	93,9	93,1	91,3	89,1	85,9	81,8	77,9	73,0	67,4	60,4	54,1	47,8	41,4
40	95,7	95,1	94,5	93,7	92,3	90,2	86,9	82,8	77,5	71,4	63,7	55,8	48,7	40,6	32,6	25,3
50	96,0	95,4	94,7	93,4	91,3	88,4	84,9	79,2	72,8	64,7	55,3	44,7	37,8	29,5	21,7	14,9
60	95,9	95,2	94,4	93,3	90,7	86,7	81,7	75,0	67,5	58,7	47,0	35,8	28,6	20,3	13,0	8,7
70	96,3	95,5	94,6	93,3	89,7	85,1	79,3	70,4	62,2	51,8	39,1	28,0	20,7	12,8	8,5	5,0
80	96,7	95,6	94,6	93,0	88,7	83,5	76,1	66,7	56,2	44,9	32,4	21,2	14,0	7,9	4,8	2,3
90	96,4	95,4	94,4	92,0	88,0	81,5	73,9	62,0	50,6	38,6	25,6	15,6	9,2	5,0	2,2	0,9
100	96,8	95,6	94,4	92,0	87,3	79,4	70,7	57,8	45,2	33,1	20,4	10,6	6,0	2,8	1,3	0,5
110	96,8	95,5	94,3	91,5	86,4	78,0	67,8	53,4	40,2	27,2	16,0	7,4	4,0	1,4	0,4	0,1
120	96,9	95,4	94,0	91,2	85,3	76,7	64,5	50,1	35,4	22,7	12,1	5,1	2,3	0,8	0,2	0,0
130	97,5	95,7	94,0	91,2	84,5	74,5	61,9	46,1	31,1	18,3	8,7	3,1	1,2	0,3	0,0	0,0
140	97,3	95,6	93,8	90,7	83,7	72,6	58,9	42,1	26,7	14,0	5,9	1,7	0,5	0,2	0,0	0,0
150	97,7	95,7	93,8	90,5	82,7	70,2	56,3	38,0	22,9	11,5	4,1	1,0	0,3	0,1	0,0	0,0

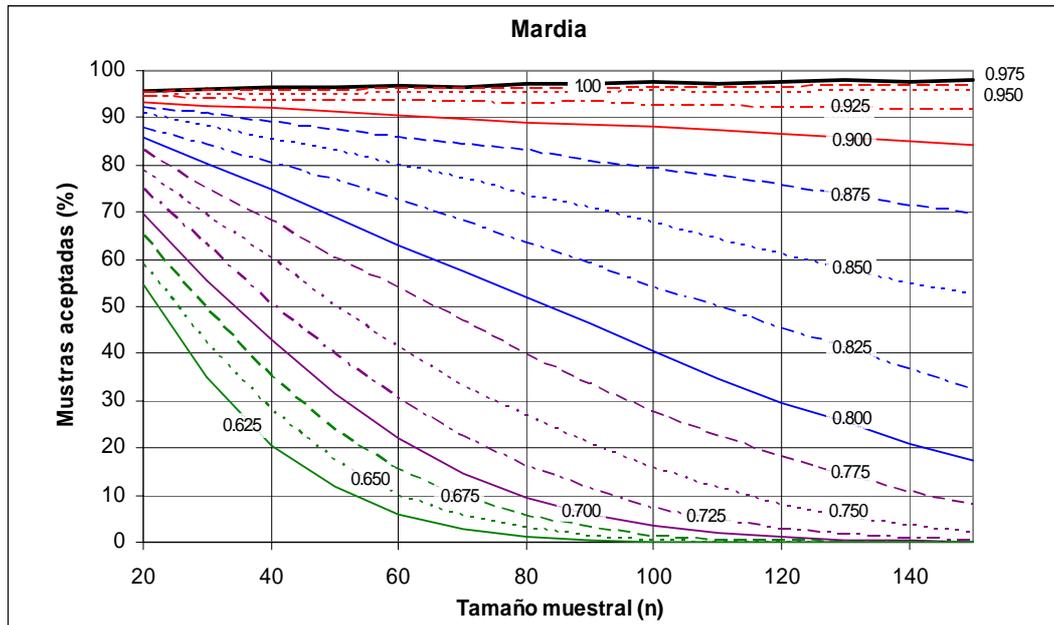


Figura 55.15: Nomograma gráfico de aceptación para una prueba de Mardia en función del tamaño muestral y el coeficiente de contaminación por errores sistemáticos de traslación.

Tabla 55.19: Nomograma analítico con porcentajes de aceptación para un contraste de Mardia.

n	Mardia: Coeficiente de contaminación sistemática por traslación															
	1,000	0,975	0,950	0,925	0,900	0,875	0,850	0,825	0,800	0,775	0,750	0,725	0,700	0,675	0,650	0,625
20	95,7	95,1	94,6	94,4	93,4	92,2	90,8	88,0	85,6	82,9	78,8	74,8	69,6	64,8	59,1	54,7
30	96,2	95,6	94,9	94,0	92,6	90,9	88,3	84,1	80,2	74,9	69,1	62,8	55,4	49,4	42,0	35,0
40	96,4	95,5	94,7	93,9	92,0	89,1	85,3	80,3	74,8	68,1	60,2	50,3	43,0	34,9	28,0	20,5
50	96,4	95,7	95,0	93,6	91,3	87,4	82,9	76,7	68,8	60,4	50,1	39,7	31,6	23,8	17,4	11,8
60	96,7	95,9	95,1	93,6	90,4	86,0	80,0	72,4	63,1	53,8	41,2	30,3	22,1	15,3	10,0	6,1
70	96,6	96,0	95,3	93,2	89,8	84,4	77,1	68,0	57,6	46,7	33,2	22,6	14,7	9,3	5,6	2,8
80	97,2	96,2	95,3	93,1	89,1	82,9	73,8	63,2	51,9	39,6	26,9	16,2	9,3	5,6	3,1	1,3
90	97,2	96,3	95,3	93,1	88,7	80,9	70,7	58,9	46,5	33,4	20,7	11,3	5,8	3,3	1,2	0,4
100	97,6	96,6	95,6	92,7	88,0	79,0	67,5	54,0	40,4	27,6	15,9	7,2	3,7	1,3	0,6	0,2
110	97,4	96,4	95,5	92,5	87,2	77,4	64,2	50,0	34,6	22,3	11,6	4,7	2,0	0,5	0,2	0,1
120	97,5	96,4	95,3	92,2	86,7	75,7	61,0	45,2	29,6	18,1	8,1	2,9	1,1	0,3	0,1	0,0
130	98,1	96,9	95,6	92,3	85,8	73,5	58,0	41,3	25,7	14,0	5,3	1,7	0,5	0,1	0,0	0,0
140	97,8	96,8	95,8	91,9	85,2	71,4	54,7	36,6	21,0	10,7	3,5	0,7	0,2	0,1	0,0	0,0
150	98,2	97,0	95,8	91,6	84,2	69,6	52,2	32,2	17,4	8,0	2,1	0,4	0,1	0,0	0,0	0,0

Sobre el caso práctico de los apartados anteriores, se ha realizado el proceso de simulación en la elección de sub-muestras. La desviación poblacional estimada para el mapa es la calculada por el estándar EMAS ($\sigma_{\text{mapa}} = 4,50$ m). Trabajando sobre los datos originales de la Hoja 926 1-1 ⁽³⁵⁾, se ha efectuado la simulación en la elección de sub-muestras, obteniéndose los resultados de la tabla 55.20 y su interpolación en los nomogramas t-Student, Ómnibus y Mardia (figura 55.16). Asimismo, se ha repetido todo el proceso en 30 ocasiones para estabilizar el resultado final.

Tabla 55.20: Porcentajes de aceptación para cada contraste en función del tamaño muestral.

N	Porcentaje de mapas aceptados		
	Contraste t-Student	Contraste Ómnibus	Contraste de Mardia
30	43,91 %	39,05 %	44,32 %
40	27,58 %	22,17 %	28,36 %
50	13,80 %	10,41 %	16,98 %
60	3,94 %	2,86 %	5,47 %
70	0,41 %	0,33 %	0,81 %

Puede observarse en las gráficas con las curvas de aceptación sobre los nomogramas, que:

- *t-Student*: el porcentaje de muestras aceptadas cruza cortando las curvas del nomograma, por lo que, a priori, no ofrecerá un resultado estable en cuanto a la magnitud de la contaminación sistemática existente sobre la población.
- *Ómnibus*: el porcentaje de muestras aceptadas es inferior a la curva con mayor error sistemático, por lo que se presupone que el sistematismo direccional será de mayor magnitud que la recogida por el nomograma.
- *Mardia*: si bien la curva del mapa analizado intersecta con las del nomograma (coeficientes de 0,650 y 0,625), su comportamiento es sensiblemente más estable que en el caso t-Student.

Realizando la interpolación sobre los nomogramas, y trasladando los valores obtenidos a la "escala" ($\sigma_{\text{mapa}} = 4,50$ m) de la cartografía analizada, se obtienen los valores de la tabla 55.21 para los errores sistemáticos detectados.

³⁵ Una vez eliminados los valores atípicos, los 100 puntos de control tienen como error medio para cada componente: $E_x = -0,598$ m ; $E_y = -2,21$ m.

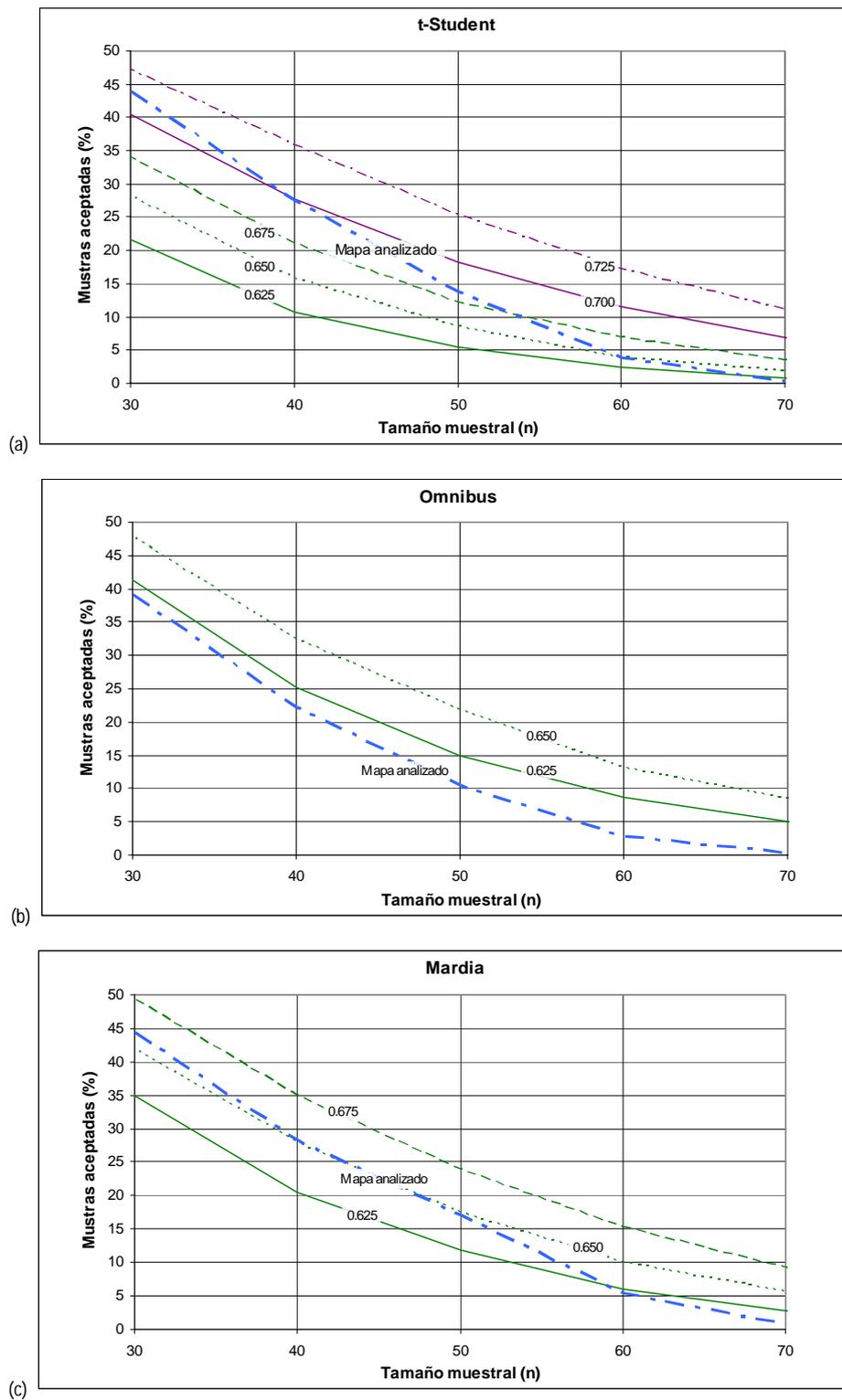


Figura 55.16: Detalle de los nomogramas y su aplicación práctica sobre la Hoja 926 1-1 para los contrastes: (a) t-Student; (b) Ómnibus y (c) Mardia.

Tabla 55.21: Interpolación del coeficiente de contaminación y magnitud del posible sistematismo para la Hoja 926 1-1 ⁽³⁶⁾

Interpolación sobre los nomogramas						
<i>n</i>	Contraste t-Student		Contraste Ómnibus		Contraste de Mardia	
	Coeficiente interpolado	Sistematismo detectado	Coeficiente interpolado	Sistematismo detectado	Coeficiente interpolado	Sistematismo detectado
30	0,713	1,29 m	< 0,625	> 1,69 m	0,658	1,54 m
40	0,699	1,35 m	< 0,625	> 1,69 m	0,651	1,57 m
50	0,682	1,43 m	< 0,625	> 1,69 m	0,648	1,58 m
60	0,651	1,57 m	< 0,625	> 1,69 m	< 0,625	> 1,69 m
70	< 0,625	> 1,69 m	< 0,625	> 1,69 m	< 0,625	> 1,69 m

Salvo en el caso del contraste de Mardia, los procesos de simulación no ofrecen resultados estables:

- *t-Student*: los valores, completamente inestables y variables en función del tamaño muestral, muestran la existencia de errores sistemáticos no inferiores a 1,29 m y, posiblemente, superiores a 1,69 m.
- *Ómnibus*: tan sólo se advierte que los sistematismos son superiores a 1,69 m.
- *Mardia*: en este último caso, los resultados algo más estables, muestran la existencia de sistematismos de 1,54 m y, posiblemente, superiores a 1,69 m.

Si se trabaja con el total de los datos obtenidos en campo, el posible error sistemático existente en la BDG sea de:

$$\text{Errores medios} \begin{cases} \overline{E_{SX}} = -0,598 \text{ m} \\ \overline{E_{SY}} = -2,211 \text{ m} \end{cases}$$

Al emplear estimadores circulares, pierde su sentido el disociar los datos en dos componentes. No obstante, el sistematismo en la componente *Y* es considerablemente superior al obtenido tras la interpolación sobre los nomogramas, mientras que el de la componente *X* es inferior. Dado que los coeficientes de los nomogramas representan un error sistemático y de igual magnitud sobre cada componente, ¿qué sería más adecuado: obtener el coeficiente del sistematismo para todas la BDG a partir de su componente cuadrática o de su valor promediado? El sistematismo y su coeficiente para el total de la muestra serían:

³⁶ Los valores en *cursiva* son datos que no se han podido interpolar al no encontrarse el mapa analizado entre dos de las curvas del nomograma respectivo.

$$\left\{ \begin{array}{l} \text{Componente cuadrática} \rightarrow \overline{E_{SXY}} = \sqrt{\overline{E_{SX}} + \overline{E_{SY}}} = 2,290 \text{ m} \rightarrow \text{Coeficiente} = 1 - \left(\frac{2,209}{4,5}\right) = 0,491 \\ \text{Valor promedio} \rightarrow |\overline{E_{SXY}}| = \left| \frac{\overline{E_{SX}} + \overline{E_{SY}}}{2} \right| = 1,404 \text{ m} \rightarrow \text{Coeficiente} = 1 - \left(\frac{1,404}{4,5}\right) = 0,688 \end{array} \right.$$

Si bien, a priori, el valor promedio es el más adecuado por la propia forma de construcción de los nomogramas, la interpolación no se ajusta ni estabiliza en el coeficiente de 0,688. En el caso de su componente cuadrática, los nomogramas no son aplicables al encontrarse fuera del intervalo. Por todo ello, se ha repetido el experimento bajo dos nuevas condiciones de trabajo:

- a) Corrigiendo completamente los errores sistemáticos de tal forma que los errores medios para todos los puntos de control son:

$$\overline{E_x} = \overline{E_y} = 0 \text{ m}$$

Tras la simulación de sub-muestras y la interpolación en los nomogramas, los porcentajes de aceptación para los tres contrastes han sido del 95 % o superiores independientemente del tamaño muestral.

- b) Corrigiendo completamente el error sistemático en la componente X y aproximadamente en la mitad de su magnitud para la componente Y , los errores medios resultantes son:

$$\overline{E_x} = 0 \text{ m} ; \overline{E_y} = -1,21 \text{ m}$$

Tras realizar los procesos de simulación en la elección de sub-muestras, los resultados no se estabilizan en ninguno de los contrastes (para todos los casos se cruzan las curvas de los nomogramas) y las magnitudes obtenidas no se ajustan al valor para el total de los datos de la BDG.

Por todo ello, se observa que los procesos de simulación seguidos en otros estándares, en cuanto a los porcentajes de aceptación de sub-muestras, no son aplicables para detectar la relación entre la magnitud de los errores sistemáticos y la desviación poblacional estimada por estadísticos circulares.

Consideraciones sobre los nomogramas y las curvas operativas

Cabe destacar el buen resultado obtenido con el nomograma para el estándar **EMAS** al ser totalmente consistente; es decir, independientemente de qué valor porcentual de aceptación se introduzca en el nomograma (σ_{Mapa}), el resultado obtenido es muy

similar. De igual forma sucede con cualquier otro mapa en el que se emplee esta técnica, independientemente de su escala y exactitud posicional.

Respecto al nomograma para el estándar **NMAS**, si bien al obtenerse mediante simulación de forma genérica no hubo problemas en cuanto a la estabilidad de sus resultados, esto no sucede así utilizando la técnica de elección de sub-muestras aleatorias sobre casos prácticos. Así, a pesar de que los resultados finales son correctos en la determinación de σ_{mapa} (coincide sensiblemente con la obtenida por otros estándares), se desaconseja su utilización.

En cuanto al nomograma para el estándar **ASPRS**, ha resultado ser bastante más estable que el caso anterior cuando es aplicado sobre datos prácticos. Los valores finales estimados para σ_{mapa} , corroboran los resultados del nomograma EMAS.

Tal y como se ha mencionado al comienzo del apartado, por las características del estándar **NSSDA**, su tratamiento y resultados son sustancialmente diferentes del resto. Así, se observa que la curva para la población distribuida según una $N(0,1)$ (figura 52.8) se repite por encima y por debajo de ésta cuando se varían las condiciones de la distribución poblacional (figura 55.9), manteniéndose constante su variabilidad relativa en función del tamaño muestral. Para generar nomogramas de aceptación similares a los estándares anteriores, pueden establecerse reglas de decisión en función de la *Exactitud_r NSSDA* (figura 55.12), observándose un comportamiento estable en los resultados. Asimismo, el estándar permite la determinación de la *Exactitud_r poblacional* con una alta fiabilidad en función del tamaño muestral, gracias a los procesos de simulación en la elección de sub-muestras y la utilización de los coeficientes de corrección mostrados en la tabla 55.14.

Analizando las curvas de todos los nomogramas, se aprecia que éstas se encuentran tanto más próximas cuanto menor es el tamaño de la muestra, lo que indica una mayor incertidumbre. Obviamente, si se desea un mayor nivel de seguridad en los resultados a la hora de obtener la σ_{Mapa} , ha de aumentarse el tamaño muestral durante el proceso de simulación.

Respecto a los estadísticos circulares **Ómnibus** y **Mardia** frente a la prueba t-Student, la aplicación de técnicas de simulación en la elección de sub-muestras no obtiene resultados óptimos en la cuantificación de un posible error sistemático.

Por último, en cuanto a los riesgos de usuario y productor, los nomogramas se han mostrado como una herramienta fiable y eficaz en su cuantificación.-

5.6 Propuesta metodológica

Por último, en función de los resultados obtenidos mediante el análisis de los principales estándares, así como otras posibles herramientas (estadística robusta, estimadores circulares...), a continuación se presenta una propuesta metodológica para el análisis de la exactitud posicional planimétrica en BDGs. Para ello se ha diseñado un posible “*Estándar de exactitud posicional en información georreferenciada (EPiG)*”, centrado sobre la exactitud horizontal de datos puntuales en el “*Estándar de Exactitud Horizontal en información puntual georreferenciada (ExHor)*”. Dicha propuesta se encuentra ampliamente desarrollada en el Anexo 6.

En el *ExHor* se han tratado de rescatar aquellos aspectos que han proporcionado los mejores resultados en los actuales estándares. De forma esquemática, el procedimiento de aplicación quedaría estructurado de la siguiente forma:

- 1º. Selección de la muestra de control. El número adecuado es de, al menos, 100 de puntos de control para garantizar unos resultados fiables. Éstos, deberán estar distribuidos homogéneamente y estar bien definidos sobre la BDG y la FME. Ésta será, al menos, tres veces más exacta que la BDG.
- 2º. Estudio de las hipótesis previas: normalidad, aleatoriedad y valores atípicos. Se obtendrán los errores entre la BDG y la FME. Dichos valores han de cumplir una serie de condiciones previas:
 - a. Normalidad: se verifica si siguen una distribución estadística Normal mediante un contraste del tipo Kolmogorov-Smirnov o D’Agostino.
 - b. Aleatoriedad: se verifica si los errores se distribuyen de forma aleatoria mediante un contraste de rachas del tipo Wald–Wolfowitz.

- c. Valores atípicos: se detectan y ponderan mediante el método danés aquellos valores considerados como atípicos (a más de $2,5 \cdot \sigma_{priori}$).
- 3º. Detección y corrección de sistematismos. La detección de éstos se realizará mediante un contraste del tipo t-Student. Si existe algún sistematismo y, en función de la variabilidad de la BDG (σ_{priori}), es viable emplear el número de puntos de control necesarios para su cuantificación, se realizará un ajuste por mm.cc. y se procederá a su corrección. Asimismo, si es detectada alguna dirección dominante, ésta deberá determinarse e indicarse.
- 4º. Estudio de la varianza. El valor de la varianza de la BDG se obtiene en dos pasos:
 - a. Estimación de la desviación poblacional a posteriori (empleando procesos de simulación en la elección de sub-muestras y el nomograma de la figura 55.2).
 - b. Cálculo del valor de exactitud posicional al 95 % de nivel de confianza (mediante la ecuación utilizada en el estándar NSSDA)
- 5º. Informe final. Por último, junto a la BDG, deberá aparecer toda la información e incidencias relacionadas con el resultado de cada una de las pruebas efectuadas en los pasos anteriores: hipótesis previas, sistematismos y varianza.

Capítulo 6

Conclusiones

6 Conclusiones

Por último lugar, en este capítulo se presentan las principales conclusiones a las que se llega tras el trabajo realizado por el doctorando. El capítulo se estructura en dos apartados:

- 1º. Conclusiones generales. Se exponen las principales conclusiones alcanzadas tras el análisis de los estándares, el empleo de estimadores robustos, la cuantificación de sistematismos y el empleo de curvas operativas. Por último lugar, se presenta la posibilidad de aplicar aquellas partes de los estándares analizados que mejores resultados ofrecen para el conocimiento de la exactitud posicional en BDGs.
- 2º. Perspectivas de futuro. Se exponen las líneas de investigación en las que esta trabajando el doctorando, así como otras posibles alternativas.

6.1 Conclusiones generales

Como principal aspecto, cabe destacar la revisión crítica realizada de los principales estándares aplicados al control posicional, analizando su comportamiento y peculiaridades gracias al uso de la simulación. Asimismo, se ha realizado un análisis sobre la posible aplicación en este campo de dos nuevas alternativas estadísticas, robusta y circular, con la idea de conocer si podrían mejorar los planteamientos actuales sobre el control de la exactitud posicional. Otro aspecto relevante, es la posibilidad que se presenta con la aplicación del ajuste por mm.cc. para la detección, cuantificación y corrección de BDGs. que posean errores sistemáticos. Finalmente, se ha incorporado de una manera explícita en el control de la exactitud posicional, la consideración de los riesgos de usuario y productor por medio de curvas operativas.

A continuación se exponen, de forma más detallada, las principales conclusiones a las que se ha llegado en cada uno de los apartados analizados.

6.1.1 Procesos de simulación

Se ha demostrado la eficacia de la simulación como herramienta de laboratorio válida en estudios sobre el comportamiento de los estándares de control posicional cartográfico. Gracias a la simulación se han podido comprobar los niveles de aceptación en función del tamaño muestral y el nivel de exigencia de cada uno de ellos.

Para el desarrollo de la simulación, ha sido necesario diseñar y crear una aplicación propia bajo entorno Visual Basic. De esta forma se han podido generar nomogramas y curvas de aceptación aplicables sobre los estándares EMAS, NMAS y ASPRS, así como la determinación del número mínimo de puntos de control recomendable para el NSSDA y el comportamiento de los estimadores circulares Ómnibus y Mardia.

Gracias a la simulación, se han podido determinar los riesgos de usuario y productor para cada uno de los estándares.

6.1.2 Estándares de exactitud posicional analizados

Basándose en el análisis sobre los estándares para el control de calidad posicional cartográfico (capítulo 5), se exponen a continuación las siguientes propuestas de mejora:

NMAS

Respecto a este estándar, ha de destacarse que:

- Es muy permisivo con cartografía a escala $> 1/20.000$. 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.
- Por el contrario, es muy restrictivo con escalas $< 1/20.000$ (en las que influye mucho más aspectos como la generalización o la simbología). Así, para aceptar

una BDG 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 BDG no cumple con el estándar, es difícil estimar cuánto costaría producir una cartografía que sí cumpliera con él y viceversa.

Dadas las peculiaridades del estándar, no se recomienda su utilización.

EMAS

Se trata de uno de los estándares que realiza un análisis más exhaustivo y detallado. No obstante, han de tenerse en cuenta los siguientes aspectos:

- 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.
- Si no se efectúa la corrección por Bonferroni, el nivel de confianza final del estándar original es del 73 % en vez del 95 % (tal y como se desprende de la lectura del estándar). Así, el riesgo del productor es excesivamente elevado (el 27 % de los mapas “buenos” son rechazados), mientras que prácticamente se ha eliminado el riesgo del usuario.
- Para solventar el anterior problema, se recomienda aplicar Bonferroni a los niveles de significación parciales, de tal forma que éste quede repartido de forma proporcional a cada componente y a cada prueba estadística: X e Y; sistematismos y estudio de la varianza.

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. No obstante, dado el nivel de exigencia del estándar, se recomienda:

- Trabajar con, al menos, 150 puntos de control y una cartografía un 10 % más exacta que la exigida para la *Clase I*.
- Si se desea producir cartografía con el mismo nivel de exactitud que la exigida, deberá aplicarse un factor de relajación a los límites (*RMSE*) de pertenencia a la *Clase I* en función del tamaño muestral: para 150 puntos de control → un factor de relajación de + 10%; para 40 puntos de control → un factor de relajación del + 20 %. No obstante, es preferible emplear 150 puntos de control.

NSSDA

Si bien se trata del estándar más actual de todos los analizados y sitúa al usuario como responsable final en la utilización de la BDG, cabría resaltar que

- Si se emplea el número mínimo recomendado de puntos de control (20), el nivel de confianza en la obtención del valor de la exactitud horizontal no es del 95% (tal y como nos indica el estándar), si no del 89 %.
- Para alcanzar un nivel de confianza del 95 %, el tamaño muestral ha de estar comprendido entre 95 y 100 puntos de control.
- No ha de descuidarse el análisis previo en cuanto a la existencia de posibles sistematismos mediante un contraste del tipo *t-Student*.

Todos los estándares analizados, dado que parten de una serie de hipótesis comunes, deberían proporcionar al usuario las herramientas necesarias para contrastar las hipótesis de aleatoriedad y normalidad en la distribución de errores.

6.1.3 Empleo de estimadores robustos.

Dado que el usuario de una BDG puede encontrarse con valores atípicos, se recomienda que:

- Se empleen todos los datos, independientemente de que éstos sean atípicos o no.
- Empleando el método danés, se ponderen aquellos valores que estén a una distancia $> 2,5 \cdot \sigma_{\text{priori}}$ del valor medio. Cuando un punto sea considerado como

atípico para cualquiera de sus componentes (X o Y), deberá ser ponderado en ambas.

- En el caso de que el número de valores atípicos sea considerablemente superior al 10 % de la muestra, debería estudiarse la posibilidad de que la σ_{priori} de la BDG no sea correcta o esté infravalorada.
- Ha de presentarse un informe al usuario en el que se indique el porcentaje de puntos que han sido considerados como valores atípicos.

6.1.4 Cuantificación de sistematismos.

En el caso de que se detecte la presencia de algún error sistemático, dependiendo de la σ_{priori} de la BDG y del número de puntos de control, podrán cuantificarse y corregirse mediante un ajuste por mm.cc. El proceso a seguir sería el siguiente:

- Detección de errores sistemáticos mediante un contraste del tipo *t-Student*.¹
- Si existen:
 - o Determinar la dirección predominante y su intervalo de confianza.
 - o Determinar si es posible cuantificar los sistematismos mediante un ajuste por mm.cc. en función de la desviación poblacional estimada a priori (σ_{priori}) y del número de puntos de control necesario. Si es así, efectuar el ajuste y su correspondiente corrección.
 - o Si el número de puntos es económicamente inviable, se eliminarán los posibles sistematismos en traslación (X,Y) aplicando la corrección de los errores medios obtenidos.

Una vez realizada la corrección, se recomienda realizar un nuevo contraste *t-Student* para verificar la eliminación de los sistematismos.

En aquellos casos en los que se hayan detectado errores sistemáticos, en la metainformación de la BDG deberá incluirse una nota alusiva al tipo de error, su dirección predominante e intervalo de confianza, así como la cuantificación y eliminación del sistematismo que se haya o no efectuado.

¹ También podrá emplearse un contraste por estimadores circulares (Ómnibus o Mardia).

6.1.5 Estimadores circulares.

La utilización de estimadores circulares ofrece resultados óptimos en la detección de comportamientos sistemáticos en la distribución de los errores posicionales. Respecto a los estimadores analizados, cabe destacar:

- Test Ómnibus: Analiza, frente a cualquier modelo de distribución (unimodal, bimodal, trimodal...), el comportamiento uniforme o la existencia de algún tipo de sistematismo en las direcciones de los vectores de error. Los resultados son óptimos en todos los casos.
- Test de Mardia: Basándose en el modelo de distribución de Von Mises, analiza existencia de comportamientos sistemáticos en las direcciones de los vectores de error posicional. Su sensibilidad es óptima y similar a la que presentan estimadores clásicos, como un contraste del tipo *t-Student*, u otros circulares como el Ómnibus.
- Dirección predominante: En aquellos casos en los que se ha detectado algún sistematismo direccional, se puede determinar dicha dirección empleando este tipo de contraste. Asimismo, la precisión en su determinación viene dada por el intervalo angular de confianza al 95%.

Como los estimadores circulares no realizan ningún tipo de estudio sobre los módulos de los vectores de error, se han realizado pruebas basadas en la relación existente entre el acimut y el módulo (gráficas A/M). Los resultados de las pruebas estadísticas realizadas no son concluyentes con la sensibilidad de los errores en BDGs.

6.1.6 Empleo de curvas operativas.

Basándose en las técnicas de simulación empleadas, han podido obtenerse una serie de curvas operativas para los distintos estándares analizados que permiten determinar los riesgos de usuario y productor. Éstos varían en función del estándar aplicado, el número de puntos de control y la relación existente entre la exactitud real de la BDG y la exactitud exigida.

Entre todas las metodologías analizadas, la fundamentada en el estándar EMAS es la que ofrece los mejores resultados al tener una más exhaustiva base estadística:

realiza un contraste de posibles sistematismos y analiza la varianza en la distribución de los errores.

Respecto al proceso recomendado a seguir para la obtención de este tipo de curvas en cualquier método que se considere, podría definirse como el presentado en la figura 61.1.

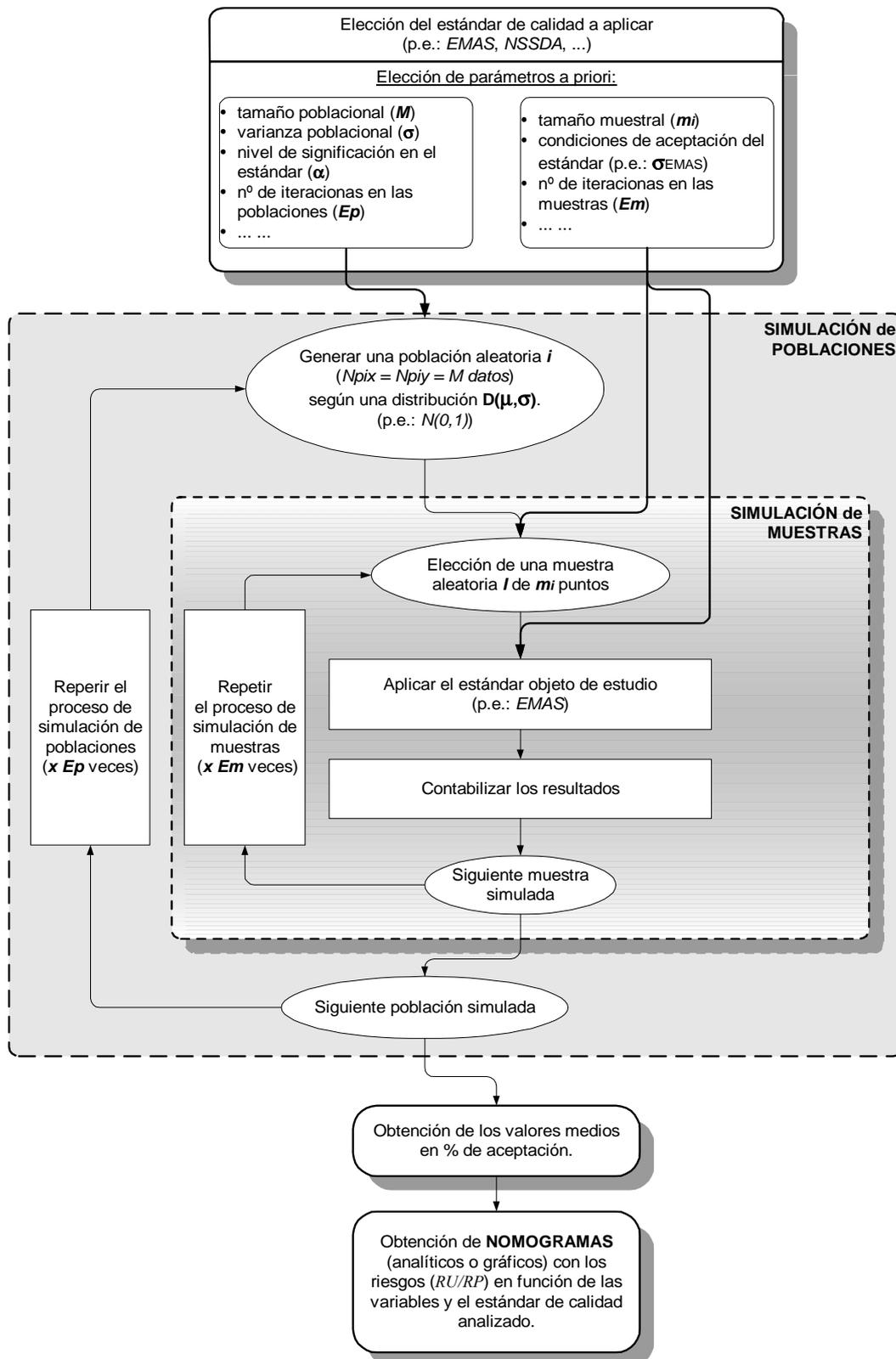


Figura 61.1: Diagrama de flujo para la obtención de nomogramas de curvas operativas.

6.1.7 Aplicación en conjunto.

Una vez estudiadas con detenimiento las diferentes características de los estándares analizados, se propone una metodología basada en los mejores aspectos de cada uno de ellos para poder efectuar un análisis exhaustivo de la exactitud posicional en BDGs. La propuesta metodológica es válida para cualquier formato de cartografía, analógica o digital, y aplicable tanto sobre productos discretos (distribuidos en hojas) como continuos.

Asimismo, se propone como denominación para el mismo, la de “Estándar de Exactitud Horizontal en información puntual georreferenciada (ExHor)”². En cuanto a los procesos de cálculo, se prevé continuar en la actual línea de investigación, proporcionando a la comunidad científica el software³ que realice de forma automatizada los cálculos matemáticos necesarios.

La metodología propuesta en el *ExHor* se resume en cinco pasos:

- 1º. Selección de la muestra de control⁴.
- 2º. Estudio previo:
 - a. Normalidad.
 - b. Aleatoriedad.
 - c. Valores atípicos.
- 3º. Detección, cuantificación y corrección de sistematismos.
- 4º. Estudio de la varianza.
- 5º. Informe final adjunto a la BDG.

Esta metodología propuesta se encuentra desarrollada en el Anexo 6. Asimismo, se incluye un ejemplo de aplicación para facilitar su comprensión.

² Dicho estándar puede considerarse como una parte de un estándar de más amplia aplicación con datos en tres dimensiones (X,Y,Z) puntuales y lineales, que podría denominarse: “Estándar de exactitud posicional en información georreferenciada (EPiG)”,

³ El software será distribuido, en primer lugar, mediante una versión de evaluación (Beta) para todos aquellos organismos y empresas nacionales e internacionales que estuvieran interesados en el mismo. Tanto el idioma del entorno gráfico, como el del manual de usuario, estará disponible en español e inglés.

⁴ Los criterios que se exponen para la selección de los puntos de control son genéricos. Es conveniente, mediante un estudio pormenorizado, elaborar una serie de “Métodos de procedimiento para la elección de puntos de control” (similares al expuesto en el Anexo 2.1) para diferentes tipos de BDG y escalas de trabajo.

6.2 Perspectivas de futuro

Para finalizar, cabría destacar como principales perspectivas de futuro dos líneas fundamentales: perspectivas del mercado y perspectivas de investigación.

6.2.1 Perspectivas del mercado

En el sector de la Geomática se está detectando un gran interés para tratar de implantar Sistemas de Gestión de la Calidad en empresas y organismos que trabajan con información georreferenciada. Un claro ejemplo de ello son los pasos dados por EuroGeographics (en Europa) y el FGDC (en EE.UU.) en la búsqueda y el establecimiento de estándares de calidad. Cada vez son más numerosos los organismos, instituciones y empresas que demandan productos cartográficos con una calidad adecuada a sus necesidades.

No obstante, a pesar de la gran difusión que ha tenido la información cartográfica en la última década con el abaratamiento de los costes de producción mediante el empleo de técnicas digitales, existe un profundo desconocimiento sobre las peculiaridades de las BDGs. Así, la difusión masiva de información georreferenciada, ha hecho que la práctica totalidad de la población disponga de ella sin que tenga una formación mínima sobre su utilización. Si bien por desconocimiento del usuario o por una incorrecta elaboración de la BDG por parte del productor, su utilización es inadecuada, conlleva a una calidad inadecuada de los productos y el asumir riesgos que no han sido tenidos en cuenta (especialmente por los usuarios).

A pesar de este último inconveniente, también es cierto que éstos han comenzado a exigir identificadores de calidad (p.e. ISO 9000:2000) en los productos que adquieren; por lo que, de forma paralela a la sociedad, los organismos de normalización están trabajando en el desarrollo de estándares para el sector. Es de destacar que en el área de la geomática, a pesar de tener una gran tradición en cuanto al cuidado de la exactitud y el control del error durante todos los procesos de producción, hasta ahora no se había trasladado ésta inquietud al desarrollo de estándares de calidad universales.

Asimismo, con la gran difusión de los sistemas de navegación y posicionamiento por satélite, unido al despegue europeo con los sistemas EGNOS y GALILEO (este último se prevé que esté operativo para el año 2008), se están creando nuevas necesidades comerciales de localización que, hasta hace unos años, eran inimaginables: control de flotas, sistemas de emergencias y rescate, aplicaciones militares, aplicaciones de seguridad nacional, productos de valor añadido a la telefonía móvil (localización de servicios en tiempo real sobre BDGs actualizadas...), optimización de trayectos y control del tráfico (gestión de la saturación del tráfico, cumplimiento de la legislación vigente...), localización de personas (reclusos, enfermos, sanitarios...) y animales (sementales o especímenes valiosos, especies en peligro de extinción...), localización de vehículos (sistemas anti-robo)... De hecho, todavía no se conocen las innumerables aplicaciones que tendrán dichos sistemas empleados de forma combinada con redes terrestres de localización (GSM, UMTS...). Todo ello ya está produciendo la necesidad de gran cantidad de productos cartográficos de calidad contrastada.

En este punto, aparece un abanico de necesidades en información georreferenciada que no está disponible o cuya calidad no ha sido contrastada. Por ello, se pueden dar situaciones de riesgo (principalmente económico) al emplear productos inadecuados.

6.2.2 Perspectivas de investigación

En cuanto a las líneas de investigación a seguir, el doctorando se propone continuar con:

- Divulgación de los resultados obtenidos:
 - o Mediante la presentación de comunicaciones en congresos nacionales e internacionales.
 - o La publicación de artículos específicos (estimadores robustos, características y mejoras de los principales estándares, empleo de curvas operativas en el control posicional...) en revistas de reconocido prestigio.
 - o Difusión de la propuesta del *ExHor* entre investigadores y empresas especializadas en productos cartográficos. Para ello se pretende desarrollar, sobre la base de Cartocal 2.0, una aplicación informática de

libre acceso que permita implementar el nuevo estándar con los procesos de simulación en la elección de sub-muestras.

- Formación en calidad:
 - o Ampliar los conocimientos relacionados con la calidad y la normativa aplicable a otros sectores.
 - o Estudiar su posible implementación en productos cartográficos.
- Investigación sobre entidades puntuales:
 - o Implementar el análisis de la componente altimétrica de forma conjunta en los estudios de control de calidad posicional.
 - o Estudiar las características propias de esta nueva componente y su comportamiento diferenciado en función del estándar a aplicar. Previsiblemente, el tratamiento será similar al aplicado para la componente planimétrica pero ajustado a la tercera dimensión.
- Investigación sobre elementos lineales:
 - o Continuar en colaboración con el Grupo de Investigación de Ingeniería Cartográfica mediante la participación en el Proyecto de Investigación Nacional *Control de calidad en cartografía por elementos lineales – ConPoCar* (⁵), dirigido por el también director de esta tesis el Dr. D. Francisco Javier Ariza López de la Universidad de Jaén.
 - o Estudio de un posible estándar de fácil aplicación para el control geométrico. Dicho estándar debería poder analizar: valores atípicos, sistematismos y varianza.
- Ampliar las relaciones con otros grupos de investigación que se encuentren trabajando en la materia.

⁵ Ref.: BIA2003-02234

Referencias

Referencias

- Alcalá, A.R. (1996). *Apuntes de Instrumentos y Sistemas de Posicionamiento*. Universidad de Jaén.
- Ariza, F.J. (2000). *Modelización, simulación y toma de decisiones con sistemas de información geográfica*. Curso de Doctorado de I.G.C. y F. en la E.P.S., Universidad de Jaén, Jaén.
- Ariza, F.J. (2002). *Calidad en la Producción Cartográfica*. Ra-MA, Madrid.
- Aronoff, (1989). *Geographic Information Systems: A Management Perspective*. WDL Publications, Ottawa, Canada.
- 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, Nueva York.
- Ashkenazi, V., Moore, T. (1986). *The Navigation of Navigation Satellites*. En Journal of Navigation, vol. 39 (3).
- ASP, (1985). *Accuracy specification for large-scale Line Maps*. En PE&RS, vol. 51, nº 2.
- ASPRS, (1989). *Accuracy standards for large scale maps*. En PE&RS, vol. 56, nº 7.
- Atkinson, A. (1998). *Control de calidad posicional por G.P.S. de las hojas 926 1-1 y 1-2 del Mapa Topográfico de Andalucía a escala 1:10.000*. PFC en Ingeniería en Geodesia y Cartografía, EPS; Universidad de Jaén, Jaén.
- Atkinson, A.; Alcalá, A.R. y Ariza, F.J. (2000). *Control de calidad posicional en cartografía mediante simulación y técnicas G.P.S.* En XII Congreso Internacional de Ingeniería Gráfica. Valladolid.
- Atkinson, A., García, J.L. y Ariza, F.J. (2001). *Comparativa mediante procesos de simulación de los test mas empleados en el control de calidad posicional cartográfico*. En XIII Congreso Internacional de Ingeniería Gráfica. Badajoz.
- Atkinson, A. y Ariza, F.J. (2002). *Nuevo enfoque para el análisis de la calidad posicional en cartografía mediante estudios basados en la geometría lineal*. En XIV Congreso Internacional de Ingeniería Gráfica. Santander.
- Atkinson, A.; Ariza, F.J. y Rodríguez, J. (2002). *Empleo de estimadores robustos en control posicional cartográfico*. En III Conferência Nacional de Cartografía e Geodesia. Aveiro.
- Atkinson, A. (2003). *Los sistemas de posicionamiento global por satélite: fundamento y aplicaciones*. Conferencia invitada en I Ciclo Ibérico de Conferências – Geomática. Beja.

- Azorín, F. y Sánchez Crespo, J.L. (1986). *Métodos y aplicaciones de muestreo*. Alianza Universidad.
- Blakemore, M. (1984). *Generalisation and Error in Spatial Data Bases*. En *Cartographica*, vol. 21 nº 2/3.
- Borrough, P.A. (1996). *Principles of Geographical Information Systems for Land Resources Assessment*. Clarendon Press, Oxford.
- Box, G.; Muller, M. (1958). *A note on the generation of random normal derivates*. En *Ann. Math. Sttistics*, vol. 29.
- Caridad, J.M. (1985). *Cálculo de probabilidades y análisis de datos I y II*. Universidad de Córdoba. Córdoba.
- Carstensen, L. (1990). *Angularity and Capture of the cartographic line during digital data entry*. En *Cartography and Geographic Information Systems*, vol. 17, nº 3.
- Caspary, W. (1992). *Error Bands and measures of geometrical accuracy*. En *Proceedings of the 3rd European Conference and Exhibition on Geographical Information Systems. EGIS'92*. Munich.
- Chen, Z. y Finn, J. (1994). *The estimation of digitizing error and its propagation with possible application to habitat mapping*. En *International symposium on the spatial accuracy of natural resource data bases* (Congalton, Ed.). Virginia.
- Cowden, D.J. (1957). *Statistical Methods in Quality Control*. Prentice-Hall, Englewood Cliffs, NJ (EE.UU.).
- Deming, W.E. (1982). *Quality Productivity and Competitive Position*. MIT Center for Advanced Engineering Study. Cambridge. MA.
- Devlin, S. J.; Gnanadesikan, R. y Kettenring, J. R. (1981) *Robust estimations of dispersion matrices and principal components*. En *Journal of the American Statistical Association* 76.
- Dodge, H.F. y Roming, H.G. (1959). *Sampling Inspection Tables*. 2ª ed. John Wiley & Son. Nueva York
- Domingo, A. (2000). *Investigación sobre los métodos de estimación robusta aplicados a la resolución de los problemas fundamentales de la fotogrametría*. Dep. IGTEG, Universidad de Cantabria, Santander.
- Earickson, R. y Harlin, J. (1994). *Geographic Measurement and Quantitative analysis*. Macmillan College Publishing Company, Inc. Nueva York
- Encarta (1998). *Control de calidad*. Enciclopedia Microsoft Encarta, Microsoft Corporation, EE.UU.
- Enell, (1954). *What Sampling Plan Shall I Chose?*. En *Industrial Quality Control*, vol.10, nº 6.

- Eurogeographics (2004). Información y documentación electrónica. [online] Disponible en Internet vía WWW. URL: <http://www.eurogeographics.org>, actualizado en junio de 2004.
- Ferrer, R., (2000). *Principales herramientas de inferencia estadística empleadas en metrología*. En Curso 9.3 Topografía aplicada, XVI Cursos de verano de Laredo, Universidad de Cantabria, Laredo.
- FGDC (1998). *Geospatial Positioning Accuracy Standards*. Ninnnesota Planning, 1999. FGDC-STD-007-1998. Virginia (EE.UU.).
- FGDC (2002). *Geospatial Positioning Accuracy Standards. PART 4: Standards for Architecture, Engineering, Construction (A/E/C) and Facility Management*. Ninnnesota Planning, FGDC-STD-007.4-2002. Virginia (EE.UU.).
- Fisher, N.I. (1995), *Statistical analysis of circular data*. Cambridge University Press, Cambridge, N.Y. (EE.UU.).
- García, M. y Martín, M. (1997). *La multiplicidad de p-valores: un riesgo a considerar*. Ed. Jano.
- García, J.L. (1998). *Generalización de elementos lineales mediante algoritmos en el dominio espacial y de la frecuencia*. PFC en Ingeniería en Geodesia y Cartografía, EPS; Universidad de Jaén, Jaén.
- Geman S. y McClure, D.E. (1992). *A nonlinear filter for film restoration and other problems in image processing*. CVGIP: Graphical Models and Image Processing. Vol 54. Nº 4.
- Goodchild, M.F. y Hunter, G.J. (1997). *A simple positional accuracy for linear features*. En Int. Journal Geographical Information Science, Vol. 11, Nº 3.
- Greenwalt, C.R. y Schultz, M.E. (1968). *Principles and Error Theory and Cartographic Applications*. ACIC Thechnical Report nº 96, St. Louis, Mo., Aeronautical Chart and Information Cernter, U.S. Air Force.
- Guftanson G. y Loon, J. (1982). *Contour Accuracy and the National Map Accuracy Standards*. En Surveying and Mapping, vol. 42, nº 4.
- Gumbel, E.J. (1954). *Applications of the circular normal distribution*. Journal Am. Statistic. Assoc. vol. 49, 267-97 (13).
- Harrington, H. (1990). *El coste de la mala calidad*. Ed. Díaz Santos. Madrid.
- Hofmann-Wellenhof, B.; Linchtenegger, H. y Collins, J. (1994). *GPS. Theory and Practice*. Springer-Verlag Wien. Nueva York.
- Jakobsson, A. y Vauglin, F. (2001). *Status of Data Quality in European National Mapping. Agencies*, En Proceeding of the 20th International Cartographic Conference, vol. 4, 2875-2883.

- Juran, M.J. y Gryna, F.M. (1997). *Manual de control de calidad*. Vol. I y II. 4ª ed. McGraw-Hill, Madrid.
- Krarup, K., Kubik, K y Juhl, J. (1967). *Götterdämmerung over least squares adjustment*. En proceedings XIV Congress of ISP. Hamburgo.
- Lacy, C.; Rodríguez-Caderot, G.; Gil, A.J.; Barzaghi, R. y Sansó, F. (1998). *Quasigeoid computation in Jaen*, En International Geoid Service Bulletin, Vol. 7, Milán.
- Leung, Y (1998). *A locational error model for spatial features*. En International Journal of Geographical Information Science, vol12, nº 16.
- Lunetta, R. y col. (1991). *Remote sensing and geographic information systems data integration: error sources and research issues*. En PE&RS, vol. 57, Nº 6.
- Maling, D. (1989). *Measurements from maps*. Pergamon Press. Oxford.
- Mardia, K.V. (1972). *Statistics of Directional Data*. Academic Press, London.
- Matos, J. (2002). *Proyecto de información geográfica*. En Datum XXI, año II, nº 3.
- Merchant, D. (1987). *A Spatial Accuracy Specification for Large Scale Topographic Maps*. En PE&RS, vol. 53.
- Morant, T. y Lerma, J.L. (1999). *La calidad en las bases de datos espaciales*. Universidad de Las Palmas de Gran Canaria. Gran Canaria.
- Morrison, J. (1995). *Spatial data quality*. En Elements of spatial data quality. Ed. Guptil & Morrison. ICA. Elsevier, Oxford.
- Mosteller, F. y Tuckey, J. W. (1977). *Data analysis and regression*. Reading, MA: Addison –Wesley.
- Mozas, A.T. (2000). *Establecimiento de una estación de referencia GPS en la Universidad de Jaén*. PFC en Ingeniería en Geodesia y Cartografía, EPS; Universidad de Jaén, Jaén.
- MPLMIC (1999). *Positional Accuracy Handbook*. Minnesota Planning, Land Management Information Center, Minnesota (EE.UU.).
- NCDCCDS (1989). *Spatial data transfer Standard (SDTS)*. National Committee for Digital Cartographic Data Standards.
- Nugent, J. (1995). *Quality control techniques for a GIS data base development project*. En PE&RS, vol. 61, nº 5.
- Otero, A. y Domínguez, E. (1992). *Introducción a la Gestión de Producción*. UPM, Madrid.

- Perkal, J. (1966). *On the length of empirical curves. Discussion paper nº 10.* Michigan interuniversity Community of Mathematical Geographers.
- Rencher, A.C. (1998). *Multivariate Statistical Inference and applications.* Wiley.
- Richardson, H.F. (1957). *Survey Adjustments and Last Squares.* Constable & Co. Londres.
- Ríos, D.; Ríos, S. (1997). *Simulación. Métodos y aplicaciones.* Ed. Ra-Ma. Madrid.
- Rodríguez, J. (2000). *Apuntes de Análisis Multivariante I y II.* Diplomatura en Estadística, Facultad de Ciencias Experimentales, Universidad de Jaén. Jaén.
- Rodríguez, J. (2003). *Comunicación personal.* Departamento de Estadística e Investigación Operativa, Universidad de Jaén. Jaén.
- Sanjosé, J.J. (2003). *Estimación de la dinámica de los glaciares rocosos mediante modelización ambiental y técnicas fotogramétricas automáticas.* Universidad Politécnica de Valencia, Valencia.
- Senlle, A. y Vilar, J. (1996). *ISO 9000 En empresas de servicios.* Ed. Gestión 2000. Barcelona.
- Sevilla, (1991) *Criterios de precisión cartográfica.* En Catastro; Año III, No 8.
- Shi, W. (1994). *Modelling and visualizing uncertainty in multi data based spatial analysis.* En European Conference on Geographical Information Systems. EGIS'92. Munich
- Spiegel, M. (1970). *Estadística.* McGrawHill.
- Stephens, M.A. (1974). EDF Test for goodness of fit and some comparisons. *Journal Amer. Statist. Assoc.*, vol. 69, 730-7 (225).
- Thapa, K. (1992). *Accuracy of spatial data used in Geographic Information Systems.* En PE&RS, vol. 58, nº 6.
- Thompson, M. y Rosenfield, G. (1971). *On Map Accuracy Specifications.* En *Surveying and Mapping* vol. 31, nº 1.
- Tobias, A. (2001). *Elección de la prueba de estabilidad.* En Institut Municipal d'Investigació Mèdica (IMIM), Barcelona
- USBB (1947). *United States National Map Accuracy Standards.* U.S. Bureau of the Budget.
- USGS (1994). *Standards for Digital Line Graphs.* National Mapping Program. Technical Instructions. USGS.

- Veregin, H. (1989). *A Taxonomy of Errors in Spatial Data Bases*. NCGIA. Technical Paper 89-12, NCGIS, Santa Barbara, California.
- Veregin, H. (1989). *Error Modelling for the Map Overlay Operation*. En M. Goodchild, S. Gopal (Eds.), *The Accuracy of Spatial Databases*. Londres, Taylor & Francis, Londres.
- Veregin, H. (1994). *Integration of Simulation Modelling and Error Propagation for the Buffer Operation in GIS*. En. PE&RS, vol.60, nº 4.
- Veregin, H. (1999). *Data quality measurement and assessment*. NCGIA Core Curriculum.
- Veregin, H. y Giordano, A. (1994). *Il controllo di qualità nei sistemi informativi territoriali*. Ed. El Cardo, Venecia.
- Volkman, U. (2002). *Elementos básicos para procesamiento de datos experimentales*. Fac. de Física, Pontificia Univ. Católica de Chile. Santiago.
- Wolf, P.R. y Ghilani, C.D. (1997). *Adjustment computations. Statistics and Least Squares*. En Surveying and GIS. John Wiley & sons, Inc. Nueva York.

Anexos

ANEXO 1

Encuestas sobre Calidad en Cartografía

Para la encuesta diseñada y enviada durante la investigación, se ha formado una base de datos con las empresas y organismos a los que ha sido enviada. Dicha base de datos se divide en dos grandes grupos: por un lado la parte nacional, con 72 envíos, y por otro la internacional, con 120. En cuanto al apartado internacional, cabe destacar que se hizo especial hincapié en Europa y el continente americano.

Dicha base de datos se ha obtenido tanto en internet, como sobre la bibliografía existente en la materia. Cabe destacar que los mejores resultados fueron obtenidos con la búsqueda en la red dado que las direcciones postales varían con el paso del tiempo. La base de datos se hizo a finales del año 2000 y principios del 2001, realizando el primer envío en la primavera de dicho año.

En cuanto al nivel de respuestas recibidas en la encuesta internacional, fue bastante desalentador a pesar de haber realizado dos envíos: el primero de ellos por correo electrónico y correo postal, y el segundo de ellos únicamente por correo electrónico. En la figura A.1, se muestran los porcentajes de empresas y organismos oficiales a los que se les envió la encuesta, así como el nivel de respuestas obtenido.

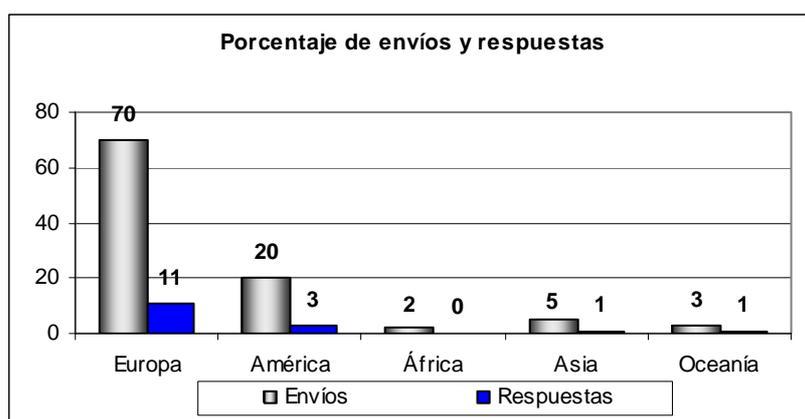


Figura A.1: Porcentaje de encuestas enviadas y respuestas recibidas en la encuesta internacional.

En cuanto a la encuesta a nivel nacional, de un total de 72 envíos, tan sólo se obtuvo algún tipo de respuesta en 11 casos (15%), siendo 6 de ellos (8%) para comunicar que la reenviaban a otro organismo o responsable.

Tanto los nombres de las empresas y organismos, como la provincia o el país al que pertenecen, aparecen en el listado 25 del Anexo 7.

La encuesta fue estructurada en 10 apartados:

- **Institución:** con los datos básicos sobre el organismo o empresa: nombre, dirección ...
- **Sistema de calidad:** si se tiene implementado alguno (ISO, EFQM, ...) o si está en proyecto implementarlo
- **Control de calidad y herramientas de mejora continua de la calidad:** con cuatro sub-apartados indicando las posibles herramientas a emplear para controlar y/o mejorar la calidad: herramientas básicas (histogramas, hojas de chequeo, ...), nuevas herramientas (diagramas racionales, de afinidad, ...), herramientas de planificación (entrevistas, cuestionarios, ...) y otros.
- **Producción:** indicando si la producción es propia o apoyada en la subcontratación.
- **Tipo de productos y servicios:** topográficos, temáticos, ortofotos, ...
- **Control de la exactitud posicional:** estándares planimétricos y altimétricos.
- **Control de la exactitud temática:** toponimia, simbología, clasificación,...
- **Coherencia lógica:** indicando el tipo de control y análisis estadístico
- **Otros:** indicando cualquier otro aspecto que se considere de interés en la materia.

Junto a la encuesta, se remitió una carta de presentación del doctorando en la que se solicitaba la colaboración del organismo para cumplimentar la encuesta y remitirla por correo electrónico o postal (apartado A1.1). A continuación se presenta la encuesta enviada (apartado A1. 2).

Por último, en el apartado A1. 3 se muestran algunas de las respuestas recibidas, que constituyen una visión realista de cuál es el estado en cuanto a calidad en cartografía en la gran mayoría de organismos que trabajan con ella en España.

A1. 1 Carta remitida

Muy Sr. mío:

Actualmente estoy realizando mi Tesis Doctoral sobre la calidad en cartografía. Uno de los aspectos que debo conocer es el estado de esta cuestión en las instituciones que, como la suya, realizan sistemáticamente trabajos de control y producción cartográfica. Por este motivo, me he atrevido a adjuntarle un cuestionario al que, si es posible, agradecería mucho le dedicara unos minutos. En él incluyo las cuestiones básicas que necesito para continuar el desarrollo de mi trabajo.

En el caso de que no sea usted la persona encargada del control de calidad en su institución, además de disculparme, me sería muy útil que me comunicara el nombre o dirección más adecuados para dirigirme directamente a ella; mi dirección de correo electrónico es atkinson@unex.es

Agradeciéndole de antemano toda su colaboración y ayuda, se despide atentamente:

A1.2 Encuesta enviada

INSTITUCION

Nombre:
 Dirección:
 E-mail:
 Web:

SISTEMA DE CALIDAD

¿Tienen algún sistema de calidad? :

- Si
- ISO 9000:1994 ISO9001 ISO9002
 ISO 9000:2000
 EFQM
 Otro (POR FAVOR ADJUNTE EXPLICACIÓN en documento adjunto):

Fecha de implementación: / /
 Responsable de Calidad:
 Persona encargada:
 E-mail:

- No
- ¿Tienen pensado establecer algún sistema de calidad?

Posible fecha de implementación : / /

Incluso si no tienen ningún Sistema de Calidad, ¿realizan algún tipo de control a sus proveedores o en alguna fase del proceso de producción?

CONTROL DE CALIDAD Y HERRAMIENTAS DE MEJORA CONTINUA DE LA CALIDAD

¿Emplean alguna de estas herramientas para el control de calidad, mejora, planeamiento, ...?

<p>Siete herramientas básicas (antiguas):</p> <p> <input type="checkbox"/> Histogramas <input type="checkbox"/> Diagramas causa-efecto <input type="checkbox"/> Hojas de chequeo <input type="checkbox"/> Diagramas de Pareto <input type="checkbox"/> Gráficos <input type="checkbox"/> Control por tablas <input type="checkbox"/> Diagramas de dispersión </p>	<p>Siete herramientas (nuevas) de control:</p> <p> <input type="checkbox"/> Diagramas relacionales <input type="checkbox"/> Diagramas de afinidad (método KJ) <input type="checkbox"/> Diagramas sistemáticos (Diagramas de Árbol) <input type="checkbox"/> Diagramas matriciales <input type="checkbox"/> Análisis matricial de datos <input type="checkbox"/> Proceso de decisión por tabla de programas (PDPC) <input type="checkbox"/> Diagramas de flechas </p>
<p>Siete herramientas de planificación del producto:</p> <p> <input type="checkbox"/> Grupo de entrevistas <input type="checkbox"/> Cuestionarios de inspección <input type="checkbox"/> Análisis <input type="checkbox"/> Comprobación por lista <input type="checkbox"/> Tablas-tipo <input type="checkbox"/> Análisis de conjunto </p>	<p> <input type="checkbox"/> Despliegue de funciones de calidad <input type="checkbox"/> Métodos Taguchi </p>

<input type="checkbox"/> Tablas de calidad	
--	--

PRODUCCION

Sus productos son realizados:

- Totalmente por ustedes
 Subcontratan algunos trabajos
 La mayor parte del proceso es subcontratado

¿Qué partes del proceso son subcontratadas?:

¿Qué porcentaje de su producción es subcontratada? % (Aprox.)

TIPO DE PRODUCTOS Y SERVICIOS

¿Qué tipo de productos y servicios generan? :

- Mapas topográficos y bases de datos
 Cartografía temática
 Ortofotos
 Modelos digitales del terreno
 Atlas multimedia
 Otros (especificar) :

Por favor, indique sus productos y principales características:

Nombre del producto	Tipo (Topo, Temático,)	Soporte	Escala Resolución
		<input type="checkbox"/> Analógico <input type="checkbox"/> Digital	<input type="checkbox"/> 1/ <input type="checkbox"/> (m)
		<input type="checkbox"/> Analógico <input type="checkbox"/> Digital	<input type="checkbox"/> 1/ <input type="checkbox"/> (m)
		<input type="checkbox"/> Analógico <input type="checkbox"/> Digital	<input type="checkbox"/> 1/ <input type="checkbox"/> (m)
		<input type="checkbox"/> Analógico <input type="checkbox"/> Digital	<input type="checkbox"/> 1/ <input type="checkbox"/> (m)
		<input type="checkbox"/> Analógico <input type="checkbox"/> Digital	<input type="checkbox"/> 1/ <input type="checkbox"/> (m)

Si disponen de algún listado con la metainformación de cada producto, por favor adjunte una copia.

CONTROL DE LA EXACTITUD POSICIONAL

¿Emplean alguno de estos test en el control posicional planimétrico? (XY)

- NMAS
 EMAS (ASPRS)
 NSSDA
 Otros (explicar y, si está publicado, indicar/incluir referencia bibliográfica):

¿Emplean alguno de estos test en el control posicional altimétrico? (Z)

- NMAS
 EMAS (ASPRS)
 NSSDA
 USGS (para MDT)

Otros (explicar y, si está publicado, indicar/incluir referencia bibliográfica):

CONTROL DE EXACTITUD TEMÁTICA

Si realizan algún tipo de control de la exactitud temática (en la toponimia, símbolos, cobertura del terreno, etc.) por favor expliquen el proceso y el grado de aceptación estadística (si está publicado, indicar/incluir referencia bibliográfica):

COHERENCIA LÓGICA

Si realizan algún tipo de control en la coherencia lógica de los datos (en la información gráfica y alfanumérica, la estructura intermedia y/o final de los datos o ficheros, etc.), por favor expliquen el proceso y el grado de aceptación estadística (si está publicado, indicar/incluir referencia bibliográfica):

OTROS

Si desea realizar algún tipo de comentario sobre algo, por favor indíquelo a continuación:

A1.3 Algunas respuestas recibidas

Para mantener el anonimato de las respuestas recibidas, se ha omitido la provincia de procedencia.

D.G. de Admón. Local y Política Territorial Área Cartografía y S.I.G.:

Atendiendo su solicitud de información respecto a control de calidad debo responderle que respecto a la calidad de nuestras producciones cartográficas (e 1/1.000 y 1/5.000), entre otras se realiza el chequeo estadístico, por parte de un Organismo externo (IGN), de algunos puntos bien definidos en la cartografía, con observación desde Vértices Geodésicos habiendo obtenido hasta la fecha resultados aceptables.

*Previamente a la recepción definitiva de los trabajos, se realizan diversos controles por parte del Director de los Trabajos, que son de carácter interno y **que lógicamente no voy a detallarle**, en los que sí se detectan diversos errores de tipo accidental (no sistemáticos ni procedimentales), y que en cumplimiento de las especificaciones del Pliego de Prescripciones Técnicas, son requeridos por el Director al Contratista para que se resuelvan antes de la recepción definitiva.*

Delegación Provincial de Agricultura y Medio Ambiente:

Recibí hace unas semanas el cuestionario sobre Calidad en Cartografía, y lamento no haber podido responder antes al mismo, pero me pareció sumamente interesante en cuanto a despertarme numerosas reflexiones.

Soy Ingeniero Técnico en Topografía desde hace 16 años, y mi área de trabajo es las Obras Públicas y más concretamente el Servicio de Carreteras, en donde desempeño labores de Control de Obra, Gestión de Expropiaciones, y Almacenamiento, gestión y mantenimiento de Cartografía de la red de carreteras adscritas a este Servicio. Como dato te puedo indicar que soy el único I.T.Top de toda la consejería y solo somos dos en toda la Junta de Comunidades de (el otro está en Agricultura).

Esta comunidad no posee centro cartográfico alguno, ni se produce Cartografía en series (básicas o superiores). Por el contrario, se produce numerosa cartografía en

*servicios como el que me ocupa, si bien esta va incluida en proyectos cerrados, en los que el apartado Cartografía (una de las principales partidas en un proyecto) **no se somete a ningún tipo de control.***

*A mí me toca el tratamiento posterior de esta información sin homogeneidad, con **enormes deficiencias**, con escasas referencias a su sistema de producción y en muchos casos con enormes errores posicionales. Con estos trabajos se proyecta, se cubican movimientos de tierras, se miden superficies a expropiar, se valoran usos de suelo, y sin posibilidad de reclamación, pues cuando llega a mis manos el proyecto, ya se abonó íntegramente hace x años y en muchos casos la cartografía, mucho antes.*

Por describirte la cartografía que aquí se produce, se trata de series de planos a escala 1000 hechos por fotogrametría a partir de vuelos 5000 o 6000, con apoyo continuo o aerotriangulación (no se especifica nada en los pliegos). La banda restituida suele ser una franja entre 100 y 150 m a ambos lados de la vía con ampliaciones en zonas de especial interés (variantes de población, cruces con zonas críticas, enlaces o intersecciones u otras causas). Se genera en UTM ed50 (pero no siempre, aunque eso sí lo especifican los pliegos). No existe ninguna normativa en cuanto a formatos de salida ni niveles de información por lo que queda al arbitrio de la empresa de cartografía, lo que supone un problema de homogenización de la información.

En conclusión, vistos estos antecedentes, como comprenderás difícilmente puedo aplicar ningún sistema de calidad. Mi lucha es más en la de conseguir controlar los procesos de creación y entrega de todo lo relacionado con la cartografía, al menos de mi servicio.

Te puedo contar que en otras direcciones generales (Urbanismo, vivienda o transportes) la situación es aun peor, puesto que la información generada yace en planeros o cajones y, por poner un ejemplo, ni siquiera se reclama en formato digital. Espero aclararte de alguna manera la lamentable situación con que contamos en la región.

Te adjunto el formulario solicitado, pero creo que no te aclarará mucho más.

Un saludo

ANEXO 2

Procedimientos para el control de calidad posicional

Si bien los métodos que se presentan están diseñados para el MTA a escala 1/10.000, con pequeñas variaciones son aplicables a cualquier otra cartografía (p.e. 1/25.000).

A2.1 Selección de puntos de control

En primer lugar, deberá realizarse una preselección de puntos sobre el mapa. Asimismo, a todos los puntos se les ha asignado un código de tres caracteres en función del tipo de elemento al que representan para facilitar su posterior identificación¹.

1. Puntos especiales. Por “puntos especiales” se entienden los Vértices Geodésicos y Puntos de Apoyo, Es aconsejable que se seleccionen todos aquellos que sean posibles dada su importancia cartográfica.
2. Obras lineales.

Como norma general, en el caso de que las intersecciones, se tratará de seleccionar aquellas cuyo ángulo de intersección sea próximo a 90° para que el punto sea perfectamente identificable sobre la BDG y la FME.

2.1.- Carreteras de 1^{er} orden.

- a) Intersecciones: se tomará el centro del cruce de la carretera de menor orden, con el borde de la carretera de mayor orden. En el caso de que ambas sean de igual importancia, se tomará la de mayor intensidad de tráfico como la de mayor orden.
- b) Comienzos y finales: en este caso no tiene sentido dado que las carreteras de 1^{er} orden siempre desembocan en otra de orden inferior o en intersecciones con otras carreteras.

2.1.- Carreteras de 2^o y 3^o orden, pistas, caminos carreteros, vías pecuarias y sendas.

¹ Ver Anexo 4, tabla A4.1

- a) Intersecciones: se tomará el centro de la intersección de ambas vías de comunicación, dado que se trata de símbolos cuyas dimensiones no se ajustan al tamaño real del elemento representado.
- b) Comienzos y finales: se tomará el centro del símbolo.

2.2.- Ferrocarril.

- a) Intersecciones: se tomará el centro de la intersección del ferrocarril con la otra vía de comunicación. No se tomarán intersecciones de ferrocarril entre sí, a menos que éstas sean perfectamente identificables sobre el mapa (poco común dado que la mayoría de ellas poseen ángulos próximos a los 180°).
- b) Comienzos y finales: se tomará, también, el centro de la vía de ferrocarril.

2.3.- Canales de riego.

- a) Intersecciones: se tomará el centro de la intersección.
- b) Comienzos y finales: se tomará en el centro del canal.

3. Obras puntuales.

3.1.- Edificios aislados.

- a) Si se trata de un símbolo: se tomará el centro del símbolo.
- b) Si se trata de un edificio con dimensiones reales: se tomará el punto más próximo a la vía de comunicación y que sea perfectamente identificable en el mapa (ángulo próximo a 90° o inferior).

3.2.- Edificios en núcleos de población. Se tratará de seleccionar edificios grandes en la periferia de la población, que estén separados a una distancia aproximada de 600 a 700 m. entre sí, aplicando los mismos criterios que en el apartado 3.1-b. En cuanto al interior de la población, se tratará de localizar edificios, a ser posible históricos (con lo que se evitará que hayan sido modificados), que se encuentren cerca de parques y espacios abiertos para evitar, en lo posible, futuros problemas en la recepción y posicionamiento con el sistema GPS. La malla de edificios seleccionados será aproximadamente de unos 500 a 700 m. de lado, dependiendo siempre de las zonas abiertas de la localidad en cuestión.

3.3.- Depósitos y estanques. Se aplicará el mismo criterio que en los edificios aislados (apartado 3.1. a y b).

3.4.- Pozos. Se seleccionará el centro del mismo.

3.5.- Torres metálicas y postes. Se tomará el centro del símbolo.

3.6.- Muros, tapias y alambradas. Se tomarán los quiebros más próximos a los 90° y que disten más de 50 m. entre sí. Los puntos seleccionados serán los más cercanos al

camino o carretera más próxima para facilitar la accesibilidad y rapidez en el trabajo de campo.

3.7.- Obras de fábrica.

- a) Si se trata de un símbolo: se tomará el centro del símbolo (p.e. : una obra de drenaje que salva un regato en el cruce con un camino).
- b) Si se trata de una obra de fábrica con dimensiones reales: se tomará el/los puntos identificables correctamente sobre el mapa.

A2.2 Levantamiento de puntos de control

Trabajo en campo.

1. Visita previa.

Visita a los ayuntamientos y reconocimiento de la zona detectando posibles cambios o diferencias entre el terreno y la cartografía a analizar, y que hayan tenido lugar entre la fecha del vuelo y la del control de calidad.

2. Itinerarios.

Diseño de los itinerarios que optimicen la toma de datos en campo para que ésta se pueda llevar a cabo en el menor tiempo posible (dado que es la parte más costosa de todo el proceso: personal especializado, dietas, transporte, equipo GPS, ...). Los itinerarios se deberán establecer de forma que se aprovechen al máximo las vías de comunicación existentes (carreteras y caminos), para optimizar el tiempo en la toma de datos, teniendo en cuenta la homogeneidad en su distribución y los criterios de selección antes expuestos.

3. Toma de datos por GPS

3.1.- Condiciones técnicas y de la Red.

Elaboración de una red previa para la obtención de los parámetros de transformación WGS84 – UTM (datum ED-50). Para evitar la acumulación de errores, dicha red se realizará en modo estático con medida de fase para obtener precisiones centimétricas. En cuanto a la toma de los puntos a controlar, se realizará en código diferencial (dado que se alcanza la precisión requerida en un mapa a escala 1/10.000).

3.2.- Condiciones en la toma de los puntos de control.

Se tomará el punto en campo siempre y cuando no exista una duda razonable de que se trate del mismo que ha sido preseleccionado sobre el mapa. En el caso de que

existiera dicha duda, se tomarán los puntos susceptibles de serlo, haciendo el correspondiente croquis y tomando las anotaciones que se estimen oportunas para aclararla.

Si existiera algún obstáculo o impedimento en la toma de algún punto (árboles que ocultan los satélites, etc.), y en el caso de que sea posible tomar otro punto próximo identificable en el mapa, se harán las pertinentes modificaciones y anotaciones sobre los croquis de campo.

Para tomar símbolos (edificios aislados, pozos, cruces de carreteras y caminos...), cuando no sea posible levantar el centro del mismo, se tomará un punto del que esté referenciado el centro del símbolo representado en el mapa. Dicha referenciación se realizará mediante algún procedimiento expedito. Por ejemplo, se podría emplear: brújula de precisión aceptable para señalar la dirección y un flexómetro para obtener la distancia a la que se encuentra el punto (evitando distancias superiores a los 10 metros para no aumentar excesivamente los errores). De esta forma, queda asegurado el estar por debajo de la tolerancia para la escala 1:10.000.

A.2.3 Digitalización de puntos de control

En función del producto a controlar, éste podrá encontrarse en formato digital o analógico. En el caso de ser necesaria la realización de procesos de digitalización, se recomienda la consulta de bibliografía específica (véase: Burrough, 1986; Carstensen, 1990; Thapa, 1992; Chen y Finn, 1994; Nugent, 1995; Ariza, 2002). Para estudiar la calidad posicional sobre cartografía analógica mediante puntos de control, se aconseja seguir las pautas que se indican a continuación.

Se realizará sobre el soporte del mapa que se desee controlar su calidad (en este caso sobre el poliéster distribuido por la empresa LTC²). Para evitar los errores en la digitalización y reducirlos al máximo, se efectuará de la siguiente forma:

- Cada digitalización se realizará de una sola vez sobre una mesa digitalizadora en la que se pueda trabajar sobre todo el mapa a controlar.

² LTC es la empresa autorizada por el Instituto Cartográfico de Andalucía para la distribución en exclusiva de sus productos. LTC, C/ Dr. Pedro de Castro, 2 - Portal 1, 41.004 Sevilla.

- Se tomarán todas las cruces de coordenadas que se aprecien correctamente para proceder a la calibración de la hoja sobre la tableta.
- Se digitalizarán los puntos siguiendo un estricto orden de la cuadrícula del mapa para evitar posibles confusiones y errores en la digitalización (puntos digitalizados por defecto o por exceso).
- Se digitalizará cada hoja un mínimo de seis veces (Giordano y Veregin, 1994), para obtener un valor medio en cada uno de los puntos.

ANEXO 3

Datos tomados en campo mediante GPS

Una vez diseñados los procedimientos para realizar el control de calidad posicional (Anexo 2), han de seleccionarse los puntos de control, efectuar la toma de datos en campo mediante GPS, y su posterior procesado.

A3.1 Selección de los puntos de control

Previamente a la toma de datos, se han de definir los itinerarios que optimicen la campaña de campo para que se pueda realizar en el menor tiempo posible. En las figuras A3.1 y A3.2 se presentan a modo de croquis los puntos seleccionados e itinerarios seguidos para la toma de datos. Dado que se estudiaron dos hojas contiguas, 926 1-1 y 926 1-2, el diseño de los itinerarios también se ha realizado de forma conjunta.

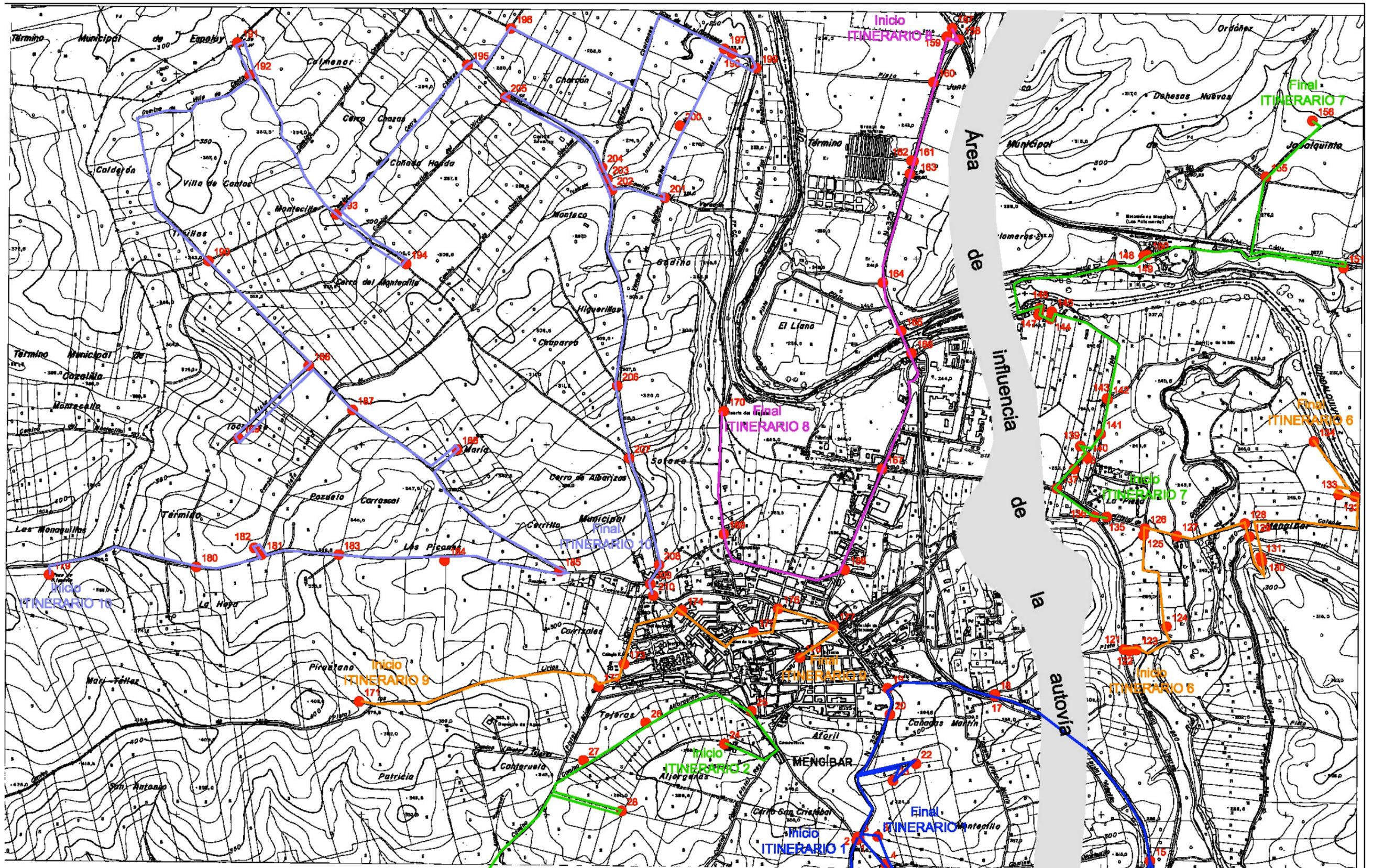


Figura A3.1: Croquis con la situación de los puntos de control y los itinerarios seguidos en campo en la hoja 926 1-1 del MTA.



Figura A3.2: Croquis con la situación de los puntos de control y los itinerarios seguidos en campo en la hoja 926 1-2 del MTA.

A continuación se muestran los datos observados en campo y los cálculos realizados para la obtención de las coordenadas terreno de los puntos de control.

Dichos trabajos han sido realizados por el doctorando en el año 1997 (Atkinson, 1998).

A3.2 Parámetros de transformación

Siguiendo el procedimiento indicado en la figura 42.1 (Capítulo 4), tras escoger las hojas 926 1-1 y 926 1-2, se seleccionaron tres vértices geodésicos que circundan la zona para la obtención de los parámetros de transformación y la situación de un punto destacado.

En primer lugar, estudiando la zona se observa que dentro de las dos hojas a controlar no existe ningún Vértice Geodésico, por lo que se han de apoyar en Vértices circundantes:

- Viñas (en la hoja 926); con coordenadas UTM ED-50 (434949.182, 4199838.312, 393.4)
- Coronilla (en la hoja 925) ; con coordenadas UTM ED-50 (424373.379, 4204663.175, 394.9)
- Atalaya de Mengibar* (en la hoja 925) ; con coordenadas UTM ED-50 (424005.122, 4199896.714, 470.4)

Se midieron todas las líneas base posible entre ellos y un “Vértice Principal” o punto destacado (denominado MENGIBAR), que se ubicó aproximadamente en el centro de la zona de estudio.

Dado que se trata de dos hojas contiguas, se sitúa en la parte inferior de la hoja 926 1-1, de forma que sea fácilmente accesible, no tenga ningún impedimento cercano que dificulte la recepción de la señal de los satélites, y se pueda establecer una buena materialización del Vértice de forma que perdure a lo largo de un cierto tiempo para posibles comprobaciones o su utilización para otros fines (ver la reseña en el apartado A3.4).

El equipo utilizado en la toma de datos ha sido :

- Dos receptores GPS Sensor WILD SR299 y sus respectivas controladoras.
- Software GPS SKI de Leica para el procesado de los datos.
- Dos Walkie Talkies YAETSU.
- Dos trípodes de topografía y un trípode de pinzas para jalón.
- Un jalón para antena GPS
- Dos basadas.
- Cuatro baterías GPS de 6 horas de autonomía.
- Brújula de mano SUNTO MC.1
- Material complementario: flexómetro, Fixolid, cinta de embalaje para sujeción de los receptores GPS a los V.G., ...

El equipo GPS de Leica posee las siguientes características técnicas para posicionamientos en modo estático relativo:

- Error en medida de fase: $5\text{mm} \pm 1\text{ppm}$
- Error en código: 50 cm
- Tiempo para posicionamiento en código: entre 0.5 y 3 minutos dependiendo del GDOP y el número de satélites.

La red ha sido tomada en modo medida de fase con estacionamientos de un mínimo de 25 minutos por Vértice. De esta forma se asegura la resolución de la ambigüedad en líneas base de hasta 11 kilómetros. En el caso de la toma de datos para la muestra de puntos de control a cotejar con el mapa, se ha realizado en código diferencial con tiempos entre 2 y 4 minutos, dependiendo del número de satélites y el GDOP. El orden en la toma de datos para la red ha sido el que aparece en la tabla A3.1.

Tabla A3.1: Orden seguido en la observación de la red.

Observación	Equipo 1		Equipo 2	
	1	Viñas	Fijo	Mengibar
2	Viñas	Fijo	Coronilla	Móvil
3	Viñas	Fijo	Atalaya	Móvil
4	Mengibar	Móvil	Atalaya	Fijo
5	Coronilla	Móvil	Atalaya	Fijo
6	Coronilla	Fijo	Mengibar	Móvil

Los itinerarios se han planificado para ahorrar el mayor tiempo posible en los desplazamientos. De esta forma se contribuye a reducir, en la medida de lo posible, la

duración de la campaña de campo con GPS. En las figuras A3.3 y A3.4 se muestran tanto los itinerarios como las líneas base medidas para el cálculo de la red y la obtención de los parámetros de transformación.

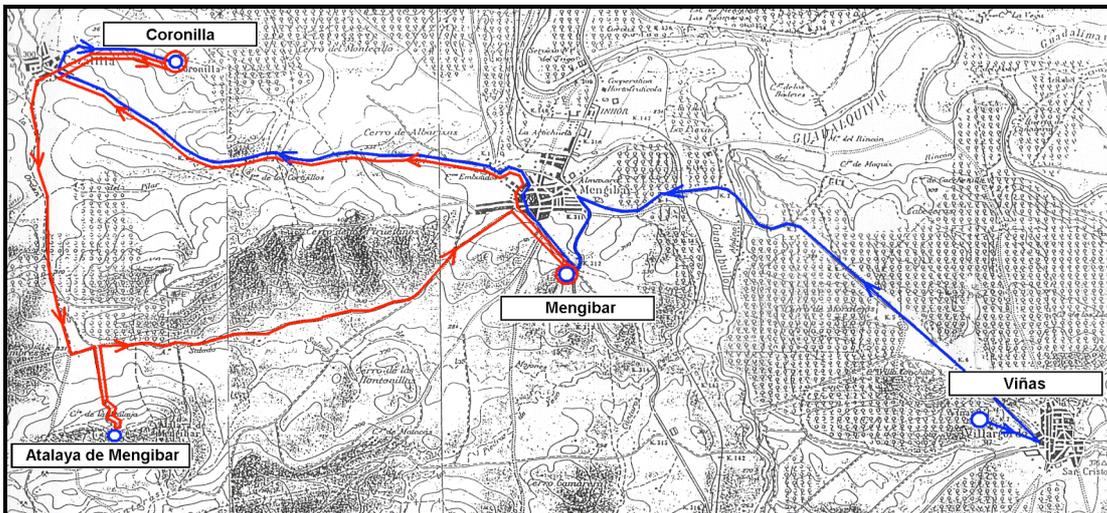


Figura A3.3: Croquis de los itinerarios seguidos para la observación de la red de vértices geodésicos y el punto destacado "Mengibar"¹.

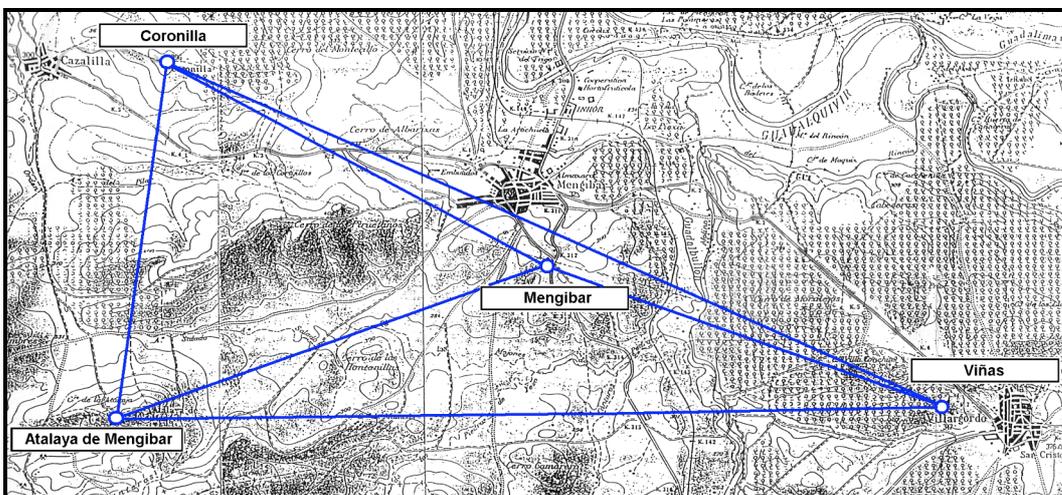


Figura A3.4: Croquis de las líneas base (vectores) observados en la red de vértices geodésicos y el punto destacado "Mengibar".

En cuanto a las características para la toma de datos, fueron las siguientes :

- Tiempo mínimo de observación en cada Vértice: 25 minutos
- Máscara de elevación: 15°

¹ En azul aparece el "Equipo 1" y en rojo el "Equipo 2".

- Intervalo de toma en modo agrupado: 30 segundos

Una vez realizada la observación, los datos se procesaron con el Software GPS SKI de Leica, obteniendo el ajuste de la red y los parámetros de transformación para la zona.

En primer lugar, se seleccionaron todas las combinaciones posibles, procesando los datos de la red y calculando las seis líneas base. Los datos de campo procesados así como las matrices varianza-covarianza de cada una de las líneas base son:

Total no of baselines computed : 6

BL id	CH id	Rover	Reference	First common ep Mod	Amb	Frq	Obs
1	1	atalaya	coronilla	08/09/97 16:17:45	Y	1+5	STS
2	2	atalaya	mengibar	08/09/97 15:26:30	Y	1+5	STS
3	3	atalaya	vinas	08/09/97 14:12:30	Y	1+5	STS
4	4	coronilla	vinas	08/09/97 13:12:00	Y	1+5	STS
5	5	coronilla	mengibar	08/09/97 17:24:30	Y	1+5	STS
6	6	mengibar	vinas	08/09/97 12:09:30	Y	1+5	STS

COORD. INICIALES:

Rover :

Point id : atalaya

X 5025375.1240 m Y -339634.5421 m Z 3900679.2930 m
Lat 37 56 30.18981 N Lon 3 51 59.01747 W h 544.7097 m

Point id : coronilla

X 5022506.5999 m Y -339092.8218 m Z 3904408.3499 m
Lat 37 59 3.36227 N Lon 3 51 44.79665 W h 553.4375 m

Point id : mengibar

X 5024418.3410 m Y -333899.5370 m Z 3902180.6200 m
Lat 37 57 35.27260 N Lon 3 48 7.27682 W h 412.8261 m

Point id : vinas

X 5025998.6018 m Y -328681.3312 m Z 3900628.5473 m
Lat 37 56 30.97712 N Lon 3 44 29.75947 W h 431.0403 m

```
#####
#####
# BL.1 atalaya coronilla 08/09/97 16:17:45 #
#####
#####
```

Rov:atalaya Ref:coronilla Amb:Y 08/09/97 16:17:45

Cartesian :

X 5025472.5191 m Y -339618.1247 m Z 3900692.6992 m
dX 2965.9192 m dY -525.3029 m dZ -3715.6507 m
sX 0.0008 m sY 0.0003 m sZ 0.0006 m

Geodetic :

Lat 37 56 28.61709 N Lon 3 51 58.07777 W h 628.7144 m
dLat 2 34.74518 dLon 13.28112 dh 75.2769 m
sLat 0.0004 m sLon 0.0003 m sh 0.0009 m

Distance :

Slope 4783.1663 m sSlope 0.0004 m

BL_VC.1 VARIANCE-COVARIANCE MATRIX

```

-----
a posteriori rms :    0.0037 m
Co-factor matrix (upper triangle) :
      qx      qy      qz
qx +4.4802699E-002 -2.6446392E-003 +2.5349434E-002
qy          +7.5381417E-003 -1.5721496E-004
qz          +2.6510045E-002

#####
#####
# BL.2 atalaya      mengibar      08/09/97 15:26:30 #
#####
#####

```

```

Rov:atalaya Ref:mengibar Amb:Y 08/09/97 15:26:30
Cartesian :
  X 5025386.9054 m   Y -339620.2961 m   Z 3900674.9884 m
 dX  968.5644 m   dY -5720.7591 m   dZ -1505.6316 m
 sX   0.0007 m   sY   0.0004 m   sZ   0.0005 m
Geodetic :
Lat 37 56 29.86448 N Lon 3 51 58.40288 W h 550.5755 m
dLat  1 5.40813 dLon  3 51.12606 dh 137.7495 m
sLat  0.0004 m sLon  0.0004 m sh 0.0007 m
Distance :
Slope 5994.3413 m sSlope 0.0004 m

```

BL_VC.2 VARIANCE-COVARIANCE MATRIX

```

-----
a posteriori rms :    0.0037 m
Co-factor matrix (upper triangle) :
      qx      qy      qz
qx +3.7408244E-002 -2.0095785E-003 +1.1496767E-002
qy          +9.5871948E-003 +4.9148442E-004
qz          +1.6006183E-002

#####
#####
# BL.3 atalaya      vinas      08/09/97 14:12:30 #
#####
#####

```

```

Rov:atalaya Ref:vinas Amb:Y 08/09/97 14:12:30
Cartesian :
  X 5025354.6551 m   Y -339611.5284 m   Z 3900647.3511 m
 dX -643.9467 m   dY -10930.1972 m   dZ  18.8038 m
 sX   0.0008 m   sY   0.0004 m   sZ   0.0007 m
Geodetic :
Lat 37 56 29.81103 N Lon 3 51 58.13370 W h 507.7403 m
dLat  1.16610 dLon  7 28.37423 dh 76.7000 m
sLat  0.0004 m sLon  0.0004 m sh 0.0010 m
Distance :
Slope 10949.1658 m sSlope 0.0004 m

```

BL_VC.3 VARIANCE-COVARIANCE MATRIX

```

-----
a posteriori rms :    0.0044 m
Co-factor matrix (upper triangle) :
      qx      qy      qz
qx +3.4602817E-002 +2.3799893E-003 +2.1160051E-002
qy          +6.9622160E-003 -6.1216339E-004
qz          +2.6109690E-002

```

```
#####
#####
# BL.4 coronilla vinas 08/09/97 13:12:00 #
#####
#####
```

```
Rov:coronilla Ref:vinas Amb:Y 08/09/97 13:12:00
Cartesian :
  X 5022388.7793 m Y -339086.2086 m Z 3904363.0283 m
  dX -3609.8225 m dY -10404.8774 m dZ 3734.4810 m
  sX 0.0028 m sY 0.0006 m sZ 0.0014 m
Geodetic :
  Lat 37 59 4.55899 N Lon 3 51 44.85149 W h 432.5406 m
  dLat 2 33.58187 dLon 7 15.09202 dh 1.5003 m
  sLat 0.0011 m sLon 0.0006 m sh 0.0029 m
Distance :
Slope 11629.2150 m sSlope 0.0008 m
```

BL_VC.4 VARIANCE-COVARIANCE MATRIX

```
-----
a posteriori rms : 0.0068 m
Co-factor matrix (upper triangle) :
      qx      qy      qz
qx +1.6937925E-001 +1.1554672E-003 +6.4902306E-002
qy          +7.5452663E-003 +4.0594878E-004
qz          +4.4190805E-002
```

```
#####
#####
# BL.5 coronilla mengibar 08/09/97 17:24:30 #
#####
#####
```

```
Rov:coronilla Ref:mengibar Amb:Y 08/09/97 17:24:30
Cartesian :
  X 5022421.0243 m Y -339094.9879 m Z 3904390.6878 m
  dX -1997.3167 m dY -5195.4509 m dZ 2210.0678 m
  sX 0.0007 m sY 0.0003 m sZ 0.0007 m
Geodetic :
  Lat 37 59 4.61206 N Lon 3 51 45.12141 W h 475.3868 m
  dLat 1 29.33946 dLon 3 37.84459 dh 62.5607 m
  sLat 0.0003 m sLon 0.0003 m sh 0.0009 m
Distance :
Slope 5988.8550 m sSlope 0.0003 m
```

BL_VC.5 VARIANCE-COVARIANCE MATRIX

```
-----
a posteriori rms : 0.0038 m
Co-factor matrix (upper triangle) :
      qx      qy      qz
qx +3.6673885E-002 +4.7836971E-004 +2.6077375E-002
qy          +5.2422257E-003 -1.1937732E-003
qz          +2.9251583E-002
```

```
#####
#####
# BL.6 mengibar vinas 08/09/97 12:09:30 #
#####
#####
```

```
Rov:mengibar Ref:vinas Amb:Y 08/09/97 12:09:30
Cartesian :
  X 5024386.1560 m Y -333890.7524 m Z 3902153.0066 m
  dX -1612.4458 m dY -5209.4212 m dZ 1524.4593 m
  sX 0.0006 m sY 0.0003 m sZ 0.0005 m
Geodetic :
  Lat 37 57 35.21878 N Lon 3 48 7.00520 W h 370.0614 m
```

dLat 1 4.24166 dLon 3 37.24573 dh -60.9789 m
 sLat 0.0004 m sLon 0.0003 m sh 0.0007 m
 Distance :
 Slope 5662.3340 m sSlope 0.0003 m

BL_VC.6 VARIANCE-COVARIANCE MATRIX

 a posteriori rms : 0.0031 m
 Co-factor matrix (upper triangle) :
 qx qx qy qz
 qx +4.0036947E-002 -5.3115012E-003 +1.2249589E-002
 qy +6.9342358E-003 -3.5458449E-003
 qz +2.1013716E-002

Seguidamente he se procedió al ajuste de la red en el sistema WGS-84. Se trata de un ajuste libre con 18 observaciones y 12 incógnitas (9 grados de libertad). Partiendo de los vectores de la tabla A3.2, se calcularon las coordenadas cartesianas de los vértices y sus consiguientes errores medios cuadráticos (tabla A3.3). Las elipses de error para cada vértice se muestran en la tabla A3.4.

Tabla A3.2: Vectores – ajuste.

Del punto	Al punto	Obs. nº	DX	DY	DZ
Coronilla	Atalaya	01_001	2965.919	-525.303	-3715.651
Mengibar	Atalaya	01_002	968.566	-5720.763	-1505.634
Viñas	Atalaya	01_003	-643.947	-10930.197	18.804
Viñas	Coronilla	01_004	-3609.883	-10404.874	3734.431
Mengibar	Coronilla	01_005	-1997.367	-5195.448	2210.030
Viñas	Mengibar	01_006	-1612.468	-5209.420	1524.445

Tabla A3.3: Coordenadas cartesianas en WGS-84 y errores medios cuadráticos.

Vértice		Coordenada	e.m.c.
Atalaya	X:	5025386.8906	± 0.00733
	Y:	-339620.2987	± 0.00339
	Z:	3900674.9874	± 0.00560
Coronilla	X:	5022420.9683	± 0.00792
	Y:	-339094.9878	± 0.00292
	Z:	3904390.6420	± 0.00641
Mengibar	X:	5024418.3414	± 0.00580
	Y:	-333899.5369	± 0.00252
	Z:	3902180.6198	± 0.00458
Viñas	X:	5026030.8179	± 0.00827
	Y:	-328690.1132	± 0.00335
	Z:	3900656.1784	± 0.00643

Tabla A3.4: Elipses de error para los Vértices.

Vértice	Semieje mayor	Acimut	Semieje menor	Altura
Atalaya	0.0057	0.6	0.0051	0.0084
Coronilla	0.0053	176.5	0.0044	0.0095
Mengibar	0.0049	177.8	0.0038	0.0066
Viñas	0.0072	170.3	0.0049	0.0093

Se alcanzaron las precisiones deseadas e incluso mejores: no hay ningún Vértice en el que se supere el centímetro de error. También se obtuvo la coordenada Z por si pudiera ser de utilidad para posteriores trabajos, aunque en este caso no fue necesario dado que tan sólo se estudió la calidad posicional en X e Y.

Seguidamente se calcularon los parámetros de transformación entre el sistema WGS-84 y el de la proyección de los mapas a controlar: Elipsoide Internacional de Hayford en la proyección UTM del Huso 30. Efectuando primero un Helmert 2D (coordenadas X e Y) y después una transformación unidimensional (coordenada Z), no fue necesario un conocimiento detallado del geoide (Hofmann-Wellenhof, 1994). Esta técnica se pudo aplicar dado que se trataba de una zona poco montañosa y, fundamentalmente, por ser un área de trabajo de reducido tamaño en el que las variaciones se pueden despreciar para el tipo de trabajo que se necesita (precisiones suficientes para obtener los puntos de control del mapa ≈ 1 m.).

La transformación Helmert 2D consta de una traslación (componentes c_1 y c_2), un factor de escala (μ) y un ángulo de rotación (α):

$$\begin{aligned}x_t &= c_1 + \mu \cdot x \cdot \cos\alpha - \mu \cdot y \cdot \sin\alpha \\y_t &= c_2 + \mu \cdot x \cdot \sin\alpha - \mu \cdot y \cdot \cos\alpha\end{aligned}\quad (\text{Ec. A3-1})$$

siendo:

- x_t, y_t : las coordenadas transformadas (2º sistema de referencia)
- x, y : las coordenadas originarias (1º sistema de referencia)

Dichos parámetros son desconocidos y pueden ser determinados utilizando puntos comunes en ambos sistemas de coordenadas. En este caso se utilizan los tres vértices geodésicos con las coordenadas WGS-84 antes obtenidas (1º sistema de referencia), y las coordenadas en el Elipsoide Internacional de Hayford en la proyección UTM (2º sistema de referencia) de la tabla A3.5.

Tabla A3.5: Coordenadas UTM de los Vértices Geodésicos

Vértice	X	Y	Z
Viñas	434949,182	4199838,312	393,4
Coronilla	424373,379	4204663,175	394,9
Atalaya	424005,122	4199896,714	470,4

De las ecuaciones (Ec. A3-1) se pueden obtener los parámetros utilizando las variables auxiliares:

$$\begin{aligned} p &= \mu \cdot \cos \alpha \\ q &= \mu \cdot \sin \alpha \end{aligned} \quad (\text{Ec. A3-2})$$

$$\begin{aligned} x_i &= c_1 + p \cdot x - q \cdot y \\ y_i &= c_2 + q \cdot x + p \cdot y \end{aligned} \quad (\text{Ec. A3-3})$$

Disponiendo de, al menos, dos puntos comunes, se puede resolver el sistema de ecuaciones. En el caso de disponer de más, se aplica un ajuste por mínimos cuadrados, con lo que el factor de escala y el ángulo de rotación se obtienen al conocer las variables auxiliares:

$$\begin{aligned} \mu &= \sqrt{p^2 + q^2} \\ \tan \alpha &= \frac{q}{p} \end{aligned} \quad (\text{Ec. A3-4})$$

Para todo ello, se introducen en el programa SKI las coordenadas UTM de los Vértices Geodésicos y se procede al cálculo:

Transformación Helmert 2D:

Nº de puntos comunes: 3

Origen de rotación: X0: 427775.894 m
Y0: 4201466.067 m

Parámetros de transformación:

Nº	Parámetro	Valor	R.m.s.	Dim.
1	Desplazamiento dX (c_1)	0.000	0.0178	m
2	Desplazamiento dY (c_2)	0.000	0.0178	m
3	Rotation al rededor del eje Z	-1.036496	0.6611	["]
4	Factor de escala (μ)	-15.224580	3.2050	[ppm]

Residuos del ajuste por mínimos cuadrados [m]

Id. del punto	dx	dy
Atalaya	0.0414	-0.0245
Coronilla	-0.0261	0.0370
Viñas	-0.0153	-0.0125

Por último, la transformación unidimensional (o transformación para alturas). Ésta consiste en un desplazamiento a lo largo del eje vertical (c_3), una rotación del eje Norte-Sur (α_1) y otro giro sobre el eje Este-Oeste (α_2). Para ello son necesarios, al menos, 3 puntos comunes en ambos sistemas de referencia (WGS-84 y UTM ED-50). Conocidas las alturas ortométricas y las alturas sobre el elipsoide, el modelo matemático para la transformación unidimensional viene dado por:

$$H_i - h_i = \Delta h - y_i \cdot d\alpha_1 + x_i \cdot d\alpha_2 \quad (\text{Ec. A3-5})$$

siendo:.

- H_i : altura ortométrica
- h_i : altura sobre el elipsoide
- Δh : desplazamiento vertical (c_3)
- $d\alpha_1$ y $d\alpha_2$: ángulos de rotación sobre los ejes X e Y

Geoméricamente, puede ser interpretada como la ecuación de un plano que permite la interpolación de las alturas del geode $N = h-H$ en otras por los puntos comunes. De esta forma, los puntos de elevación en un área reducida de trabajo (10 km. \times 10 km.) pueden ser determinados con una precisión en torno a los 3 cm. Además, y dadas las características del trabajo, el cálculo de la coordenada Z no es necesario. No obstante:

Transformación para alturas:

Nº de puntos comunes: 3

Parámetros de alturas: -0.00005095 -0.00002455 -81.394 m

Id. del punto	GPS [m]	orig. [m]	transf. [m]	res.[m]
Atalaya	550.563	470.400	470.400	0.000
Coronilla	475.315	394.900	394.900	0.000
Viñas	473.834	393.400	393.400	0.000
Media de la precisión en el ajuste: dHm:		0.000 m.		

Obviamente, y puesto que solamente se parte de tres puntos con cota conocida en los dos sistemas -WGS84 y UTM ED50- el ajuste en Z no tiene comprobación (por tres puntos solamente pasa un único plano).

Una vez calculados los parámetros de transformación entre ambos sistemas de coordenadas, se obtuvieron las coordenadas UTM del Vértice de referencia MENGIBAR, que sirvió como base fija para la posterior medición y cálculo de las

coordenadas de todos los puntos de control en las dos hojas del MTA. Las coordenadas definitivas en UTM son: Mengibar (429663.76, 4201862.16, 332.4).

A3.3 Toma de puntos de control y cálculo de las observaciones

En observables de código, con los datos del Vértice Principal Mengibar, se obtuvieron las precisiones para cada uno de los puntos de control (ver listado 18 en el Anexo 7) y con los parámetros de transformación del apartado anterior, se calcularon las coordenadas UTM ED50 de todos los puntos de control. Así, se obtiene el listado nº 1 de coordenadas (ver listado 1, Anexo 7), en el cual aparecen conjuntamente las hojas 926 1-1 y 1-2.

Seguidamente se corrigió la excentricidad en aquellos puntos en los que no fue posible tomar el centro del símbolo: torretas metálicas y postes, cruces de carreteras con tráfico, edificios aislados, etc. ...

Para corregir las coordenadas de cada punto en la dirección y magnitud correspondiente, se empleó el programa AutoCAD. El proceso de transferencia del fichero ASCII con las coordenadas UTM obtenidas en campo ha sido realizado mediante la aplicación en Autolisp: "Insertap.lsp" (el código se adjunta en el Anexo 5, apartado A5.3). No obstante, hoy en día existen numerosas aplicaciones comerciales que permiten realizar este tipo de procesos (p.e.: Autodesk Land Desktop, TCP-IT, ...). Después de introducir los puntos tomados en campo, se corrigieron las coordenadas de cada uno de ellos por excentricidad según se indica en las reseñas de campo (A3.4 y listado 49 del Anexo 7). Con ello se generó el nuevo fichero de puntos: listados nº 2 y 3 (Anexo 7). Para extraerlos desde AutoCAD se utilizó la aplicación de Autolisp "contap.lsp" (el código de ésta se adjunta en el Anexo 5, apartado A5.3).

Todos los puntos tomados en campo tienen una codificación en función del tipo de elemento al que representan, tal y como se recoge en el Anexo 2 (apartado A2.2)

A3.4 Reseñas de campo

A continuación se reseña, en primer lugar, el Vértice Principal MENGIBAR (estación fija en la toma de la nube de puntos de control).

En el Anexo 7 (listado 49) aparecen las reseñas de cada uno de los puntos tomados para comprobar la calidad posicional de las dos hojas del MTA.

En dichas reseñas aparece en la parte superior izquierda, el número identificativo del punto. En la superior derecha el código de la cuadrícula en la que se encuentra (ver la codificación al comienzo de las reseñas – figura A.3.7). A la izquierda, bajo el número, aparece el código identificativo del tipo de punto del que se trata (cruce de camino, muro, punto de apoyo,...²). Junto a él, está anotada la corrección por la excentricidad (dirección y distancia) existente entre el símbolo que aparece en el mapa y el punto real tomado en campo (en casos en los que no hay que aplicar ninguna corrección está marcado como “OK”). La nomenclatura empleada para indicar hacia qué dirección hay que aplicar la corrección es la que se indica en la figura A3.5.

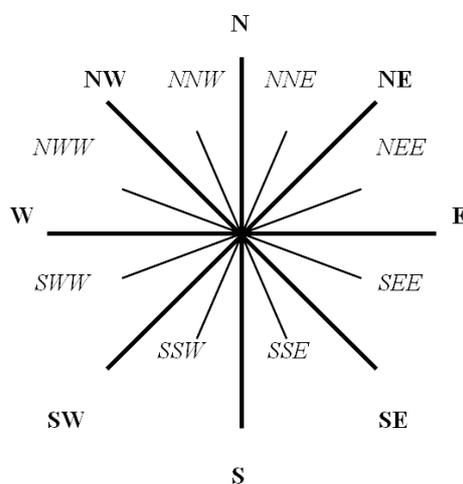


Figura A3.5: Croquis de direcciones seguidas para la corrección por excentricidad de los puntos durante la toma de datos en campo.

En la cuadrícula “Obs. :” se anota cualquier observación aclaratoria del punto o sus condiciones en la toma.

² Ver la codificación empleada para la cartografía 1/10.000 del MTA en el Anexo 4 (Digitalización).

Reseña Vértice Principal MENGIBAR

- SEÑAL : Punto de Fixolid negro sobre pivote de hormigón de Telefónica, con el anagrama "Mengibar" y la numeración de telefónica:

1
143

- LOCALIZACIÓN: Se encuentra a 1.4 km. del cruce de la estación de autobuses de Mengibar, en dirección Sur - Jaén, pasando frente al cuartel de la Guardia Civil. El Vértice se sitúa en el cruce con la carretera que accede al cementerio y al sur del casco urbano de Mengibar. Según se accede al cruce desde la población, a mano derecha hay dos bancos de hormigón. Entre éstos se encuentra el Vértice.
- COORDENADAS UTM : X = 429663.76 ; Y = 4201862.16
- CROQUIS: (ver figura A3.6)

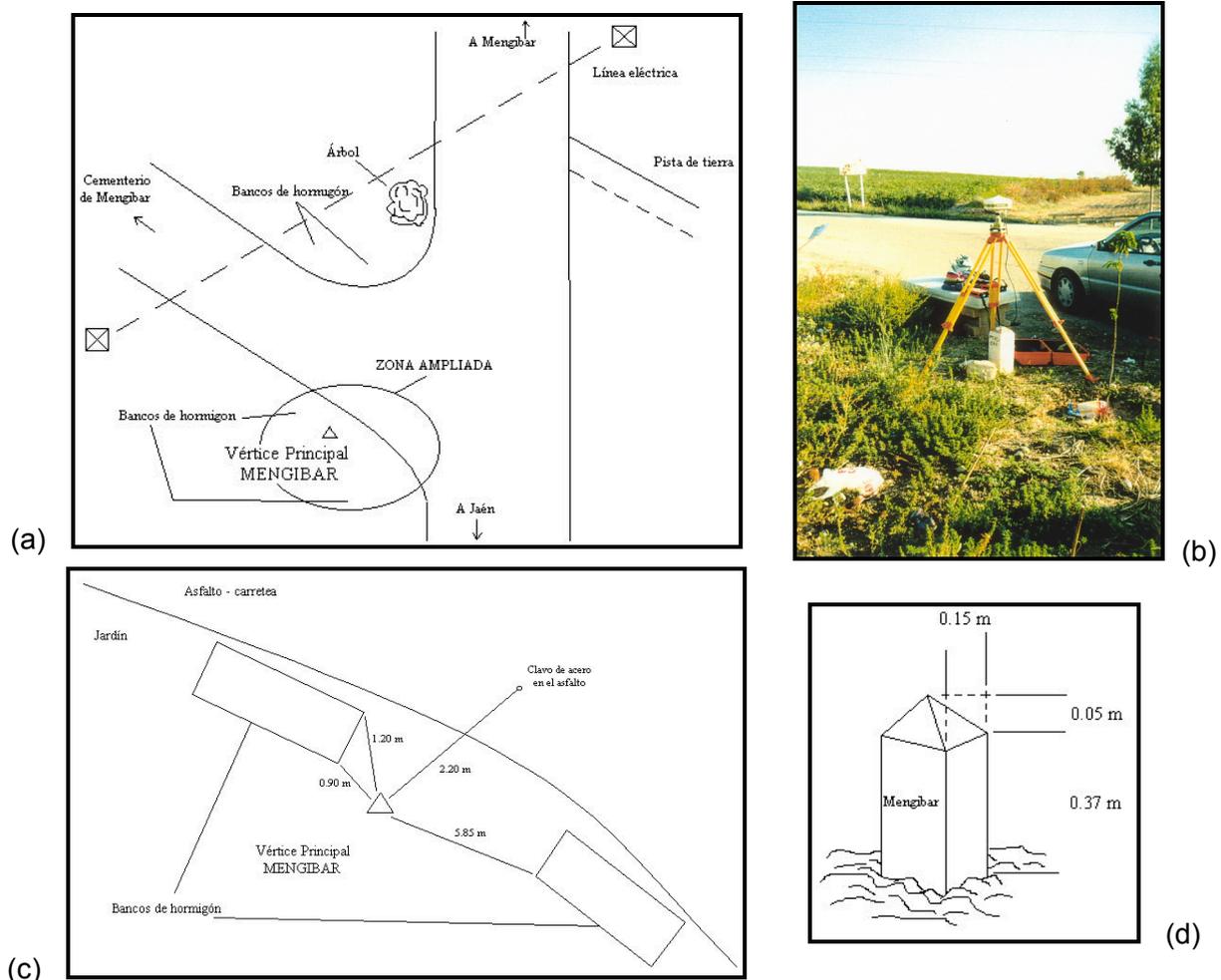


Figura A3.6: (a) Croquis de la situación del punto, (b) fotografía del punto, (c) croquis ampliado y acotado, y (d) detalle de la señal.

(a)

A-1	A-2	A-3	A-4	A-5	A-6	A-7	A-8
B-1	B-2	B-3	B-4	B-5	B-6	B-7	B-8
C-1	C-2	C-3	C-4	C-5	C-6	C-7	C-8
D-1	D-2	D-3	D-4	D-5	D-6	D-7	D-8
E-1	E-2	E-3	E-4	E-5	E-6	E-7	E-8
F-1	F-2	F-3	F-4	F-5	F-6	F-7	F-8

(b)

A-1	A-2	A-3	A-4	A-5	A-6	A-7	A-8
B-1	B-2	B-3	B-4	B-5	B-6	B-7	B-8
C-1	C-2	C-3	C-4	C-5	C-6	C-7	C-8
D-1	D-2	D-3	D-4	D-5	D-6	D-7	D-8
E-1	E-2	E-3	E-4	E-5	E-6	E-7	E-8

Figura A3.7: Croquis de situación de los cuadrantes para la ubicación de los puntos de control sobre las hojas 926 1-1 (a) y 926 1-2 (b)

ANEXO 4

Digitalización

Todo el proceso de digitalización se ha realizado de acuerdo con las indicaciones del Anexo 2, apartado A2.3. Se ha trabajado sobre el soporte de mayor precisión y estabilidad referido a las hojas 926 1-1 y 1-2 del Instituto Cartográfico de Andalucía: reproducciones en poliéster suministradas por la empresa concesionaria. Asimismo, se ha efectuado un procesado de los datos para detectar y depurar posibles errores en esta fase del proceso.

A4.1 La digitalización

Para reducir al máximo los errores en la digitalización, se ha trabajado sobre una tableta digitalizadora formato A0 capturando todos los puntos de cada hoja de una sola vez. Se ha digitalizado cada hoja un total de siete veces (Giordano y Veregin, 1994), con lo que se han obtenido los valores promedio para cada punto.

En cuanto al modelo de tableta digitalizadora empleado, éste ha sido la Calcomp 95480, con las siguientes características técnicas:

- Resolución: 10,160 líneas por pulgada (400 lpmm)
- Precisión: ± 0.005 pulgadas
- Cursor: 16 botones
- Dimensiones: 44 × 60 pulgadas (1219 × 1524 mm)

El fabricante la acompaña de su correspondiente Ficha de Calibración (ver figuras A4.1, A4.2 y A4.3 al final de este anexo).

En lo que respecta a la calibración de las hojas sobre la tableta en cada sesión de digitalización, han sido tomadas todas las cruces de coordenadas que aparecen en cada hoja, siempre y cuando éstas se aprecien correctamente y no suponga una fuente de error, sino de corrección y ajuste. Así, para la hoja 926 1-1 han sido tomadas un total de 25 cruces para cada una de las siete digitalizaciones, mientras que para la

hoja 926 1-2 han sido 24 cruces. La calibración de cada hoja fue siempre mediante una transformación afín en la que los errores eran inferiores a la tolerancia del mapa. El proceso de digitalización ha sido realizado bajo el software AutoCAD. En la primera digitalización de cada hoja se ha asignado a cada punto, además del nº identificativo, un determinado código. Con éste, se señala la categoría de punto a la que pertenece (cruce de camino, edificio aislado, punto de apoyo, ...). La codificación empleada es la que aparece en la tabla A4.1.

Tabla A4.1: Codificación de los puntos según tipo de elemento que representan

GRUPO 1: Puntos especiales	Código	
Vértice Geodésico	vgd	
Punto de Apoyo	pap	
GRUPO 2: Obras lineales	Intersección	Comienzo / Fin
Carreteras de 1 ^{er} orden: Autopistas y Nacionales	c1x	c1f
Carreteras de 2 ^o orden: Comarcales	c2x	c2f
Carreteras de 3 ^{er} orden: Locales	c3x	c3f
Caminos carreteros y vías pecuarias	cmx	cmf
Sendas	snx	snf
Canales de riego	crx	crf
Pasos del ferrocarril	frx	frf
GRUPO 3: Obras puntuales	Código	
Edificios aislados	eda	
Edificios en núcleos de población	edp	
Depósitos y estanques	dep	
Pozos	poz	
Torres metálicas y postes	tor	
Muros, tapias y alambradas	mur	
Obras de fábrica	odf	

En los elementos del Grupo 2 (Obras lineales) se diferencian de si se trata de un cruce o del comienzo/final de uno de los elementos representados. El criterio de asignación es el de la vía de comunicación de mayor importancia. Así, por ejemplo, en un cruce de un camino carretero con una carretera comarcal (de 2^o orden), el código asignado al punto es "c2x".

En el resto de las digitalizaciones tan sólo ha sido asignado el nº del punto para su posterior identificación.

Para evitar posibles errores a la hora de transcribir los datos de la digitalización a un fichero ASCII, y buscando una forma de agilizar el proceso, se ha empleado la

aplicación en AutoLisp que, al igual que en el Anexo 3, extrae los puntos automáticamente. La aplicación en concreto se denomina “contap.lsp” y su código aparece en el Anexo 5 (A5.3).

En cuanto a la digitalización propiamente dicha, se han capturado los puntos siguiendo un estricto orden por cuadrículas UTM. Esta numeración de cuadrículas sirve también para localizar fácilmente sobre el mapa los puntos que aparecen en las reseñas (ver Anexo 3). Al digitalizar cuadrícula por cuadrícula, se reduce notablemente el número de puntos a los que el operador ha de prestar atención, evitándose errores tanto por defecto como por exceso, así como confusiones entre unos y otros puntos. El orden seguido en la digitalización ha sido, comenzando por la margen superior izquierda, de Oeste a Este y de Norte a Sur.

Los listados con los puntos de cada una de las hojas y la digitalización a la que pertenecen, aparecen en el Anexo 7 (A7 – listados 4 a 17). En ambas hojas (926 1-1 y 1-2), los puntos han sido digitalizados hasta un total de 7 veces.

A4.2 El tratamiento previo de los datos

Antes de calcular el promedio de todas las digitalizaciones para cada punto y comparar sus coordenadas con las tomadas en campo por GPS, se han tratado los datos de la siguiente forma.

En primer lugar se ha pasado un test para la detección de errores groseros en cada serie de siete digitalizaciones. Se supone la hipótesis nula $H_0 \Rightarrow$ *la muestra pertenece a la población con una desviación típica de 2.5 m* (la estimada a priori para un mapa a escala 1/10.000). El nivel de significación es de $\alpha = 5 \%$ como probabilidad de rechazar una hipótesis nula, siendo verdadera. La serie será de n elementos¹. El test “ t ”, consiste en calcular todos los residuos e_{xi} con respecto al valor medio para cada componente (\bar{x}), y aplicar la siguiente formulación²:

¹ En este caso $n = 7$ digitalizaciones como punto de partida

² Como ejemplo, se presentan las ecuaciones para la componente X.

$$e_{xi} = x_i - \bar{x} \quad m = \sqrt{\frac{\sum_{i=1}^n e_{xi}^2}{n-1}} \quad \tau = \frac{m}{\sigma}$$

Con ello, se obtiene para cada punto un valor “ τ ”. Éste se compara con el valor “ τ_α ”, y si $\tau > \tau_\alpha$ con el riesgo de estar equivocado en un 5 %, se afirma que la hipótesis nula es rechazada. Esto implica que existen errores groseros, rechazándose aquella coordenada que más se aleje de la media, y volviendo a calcular nuevamente el test. Al rechazar alguna de las componentes, es eliminado el punto por completo, anulando también la otra componente, pues se considera que la digitalización ha sido errónea.

Si, por el contrario, $\tau < \tau_\alpha$ se acepta la hipótesis de que no hay errores groseros en la digitalización de dicho punto, con una fiabilidad del 95 %. Los valores de τ_α son: para $n = 7 \rightarrow \tau_\alpha = 1.455$; y para $n = 6 \rightarrow \tau_\alpha = 1.49$.

En los listados 19 y 20 del Anexo 7 aparecen los valores de τ obtenidos para cada punto. Como se puede apreciar, tan sólo es en la hoja 1-2 en la que se ha producido algún error grosero en la digitalización (valores marcados en rojo en las tablas). En el caso del punto nº 7, el error estaba en la 1ª digitalización, componente Y . En el punto nº 56, el error estaba en la 3ª digitalización en la componente Y . Por último, en el punto nº 118, el error se encontraba en la 3ª digitalización y también en la componente Y .

Una vez eliminados esos puntos en dichas digitalizaciones, se ha pasado nuevamente el test obteniendo los valores de la tabla A4.2.

Tabla A4.2: Test Tau sin errores groseros

Punto	Código	Tau X	Tau Y
7	c2x	0,4164	0,7671
56	mur	0,6722	1,3060
118	eda	0,6576	0,4043

Una vez suprimidos los errores groseros, y para hacerse una idea de las precisiones reales que se pueden llegar a alcanzar en la digitalización sobre este tipo de mapas a escala 1/10.000, se han calculado las desviaciones típicas para cada punto (ver listados 21 y 22 del Anexo 7).

Por último, se exponen en los listados 23 y 24 el valor medio calculado para las coordenadas X e Y de cada punto digitalizado. Éste no contiene errores groseros y es el que será utilizado para el control de calidad.

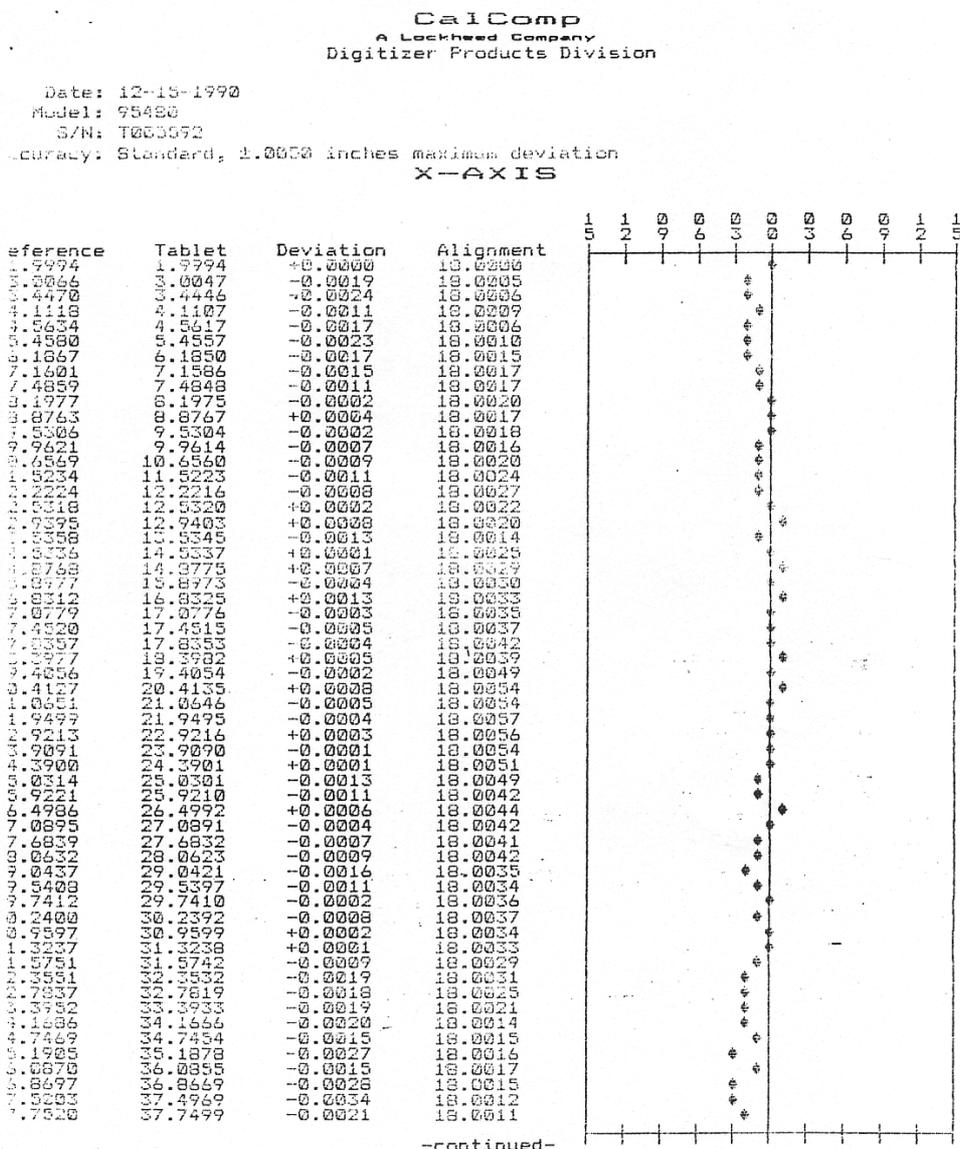


Figura A4.1: Certificado de calibración de la tableta: Eje X.

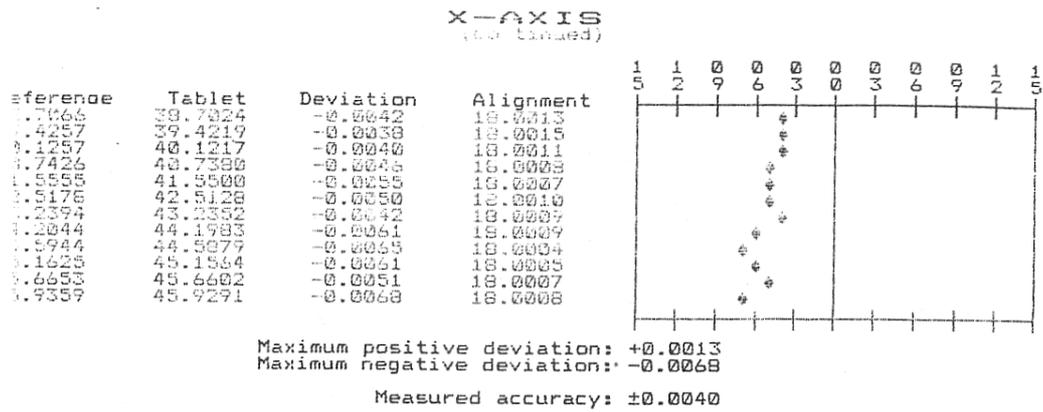


Figura A4.2: Certificado de calibración de la tableta: Eje X (cont.).

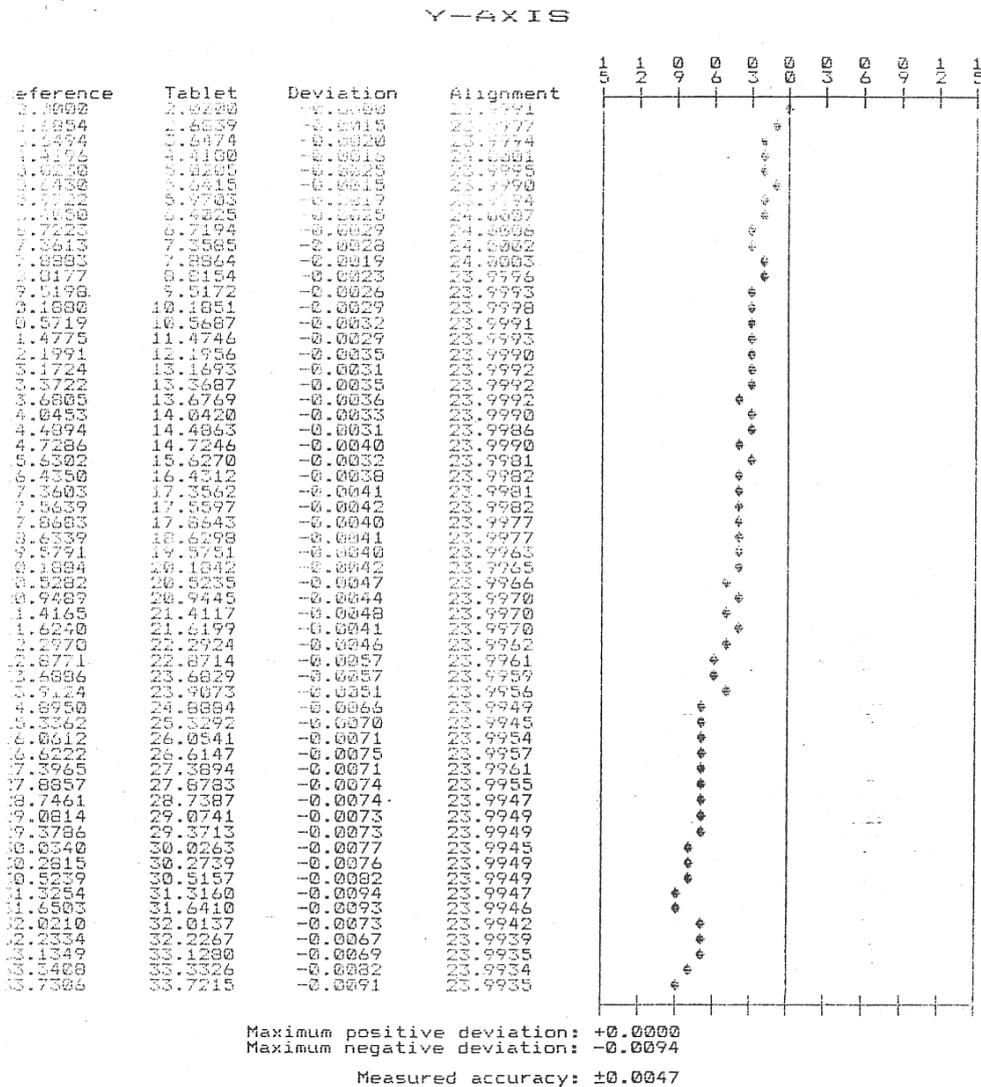


Figura A4.3: Certificado de calibración de la tableta: Eje Y.

ANEXO 5

Software I+D: Cartocal 2.0

Para poder realizar todos los procesos de simulación descritos en el capítulo 4, se ha diseñado una aplicación propia bajo entorno Microsoft Visual Basic 6.

La ideal original de la versión 2.0 de Cartocal, se fundamenta en el software y los algoritmos desarrollados por el doctorando en su Proyecto Final de Carrera (Atkinson, 1998).

A5.1 Manual del usuario

El presente Manual de usuario se encuentra en formato PDF¹ como un acceso desde la propia ayuda del programa para facilitar su utilización.

Presentación:

Cartocal es una herramienta desarrollada por el doctorando en el seno de los Grupos de Investigación de Ingeniería Cartográfica (Universidad de Jaén), e Ingeniería en Geomática y Patrimonio Urbano (Universidad de Extremadura).

En las líneas de investigación seguidas, la aplicación ha sido desarrollada dentro de las labores de producción cartográfica y dirigida hacia la investigación en el control de calidad posicional.

Condiciones de utilización

La aplicación Cartocal se ha desarrollado con finalidad de investigación, siendo de uso restringido y quedando totalmente prohibida su difusión o copia. El autor del programa no se hace responsable de cualquier empleo del mismo, quedando exento de cualquier daño o perjuicio que pudiera causar. Los riesgos de utilización del presente programa corren a cargo del usuario del mismo, aceptando las condiciones mencionadas.

¹ Se adjunta en formato digital junto al Anexo 7.

Instalación de Cartocal 2.0

Los requisitos mínimos para instalar y poder trabajar con el programa son:

- Sistema operativo: Microsoft Windows XP (o compatible)
- Procesador con velocidad ≥ 300 Mhz (²)
- Memoria ≥ 64 Mb RAM
- Espacio libre en disco duro ≥ 40 Mb
- Unidad lectora de CD

Para una correcta instalación de Cartocal 2.0, ejecutar el archivo “*setup.exe*” ( *setup.exe*) que se encuentra en el CD de instalación. Automáticamente se iniciará un tutorial que le guiará en la instalación del programa.

Por defecto, quedará instalado en “*C:\Archivos de programa\Cartocal*”, generándose en el menú de “*Inicio → Programas ...*” una carpeta con la aplicación Cartocal. Para desinstalar el programa, en dicha carpeta se encuentra la aplicación “*Uninstall Cartocal*” ( Uninstall).

Introducción a Cartocal 2.0

Cartocal implementa los cálculos necesarios a realizar en la aplicación de 4 estándares diferentes para el control de la calidad posicional planimétrica y la aplicación de diferentes estimadores circulares:

- NMAS - National Map Accuracy Standard
- EMAS - Engineering Map Accuracy Standard
- ASPRS - Spatial Accuracy Specification for Large Scale Topographic Maps
- NSSDA - National Standard for Spatial Data Accuracy
- Estimadores circulares: Ómnibus, Mardia y determinación de la dirección dominante.

² Dado que durante la simulación la mayor parte del trabajo recae sobre el microprocesador, es muy recomendable emplear equipos 1 GHz o superior.

Para ello, el usuario deberá cumplir con las exigencias demandadas por cada uno de los estándares en cuanto al tipo de puntos a emplear en el control, el número mínimo de datos, su distribución, ... (ver capítulo 3, apartado 3). El programa trabaja con ficheros ASCII para que su manipulación con cualquier editor resulte sencilla y rápida. Así, bastará seguir unas sencillas indicaciones a la hora de realizar la cabecera y el propio archivo, para que Cartocal lo reconozca y cargue satisfactoriamente.

La estructura de los menús, tal y como se ha mencionado en el capítulo 4, es la que se indica a continuación:

- Fichero
 - Abrir
 - Guardar fichero de errores
 - Salir
- Tipo de mapa
 - Seleccionar tipo ...
- Test
 - Parámetros estadísticos EMAS
 - Test EMAS
 - Test NMAS
 - Test ASPRS
 - Test NSSDA
 - Estimadores circulares
- Simulación
 - Simulación aleatoria de muestras
 - Simulación aleatoria de errores
- Afín
 - Transformación AFIN
- Ayuda
 - Ayuda de CARTOCAL
 - Acerca de ...

A continuación, se expone cómo trabaja el programa y las indicaciones a seguir para su correcta utilización.

Fichero → Abrir

El primero de los menús desplegables de Cartocal es el de “*Fichero*”. Sobre él aparece el submenú “*Abrir*” (figura A5.1)

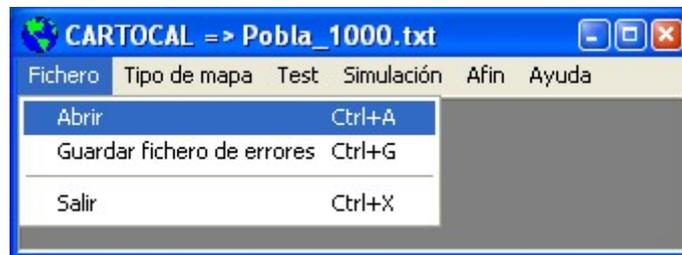


Figura A5.1: Menú Fichero → Abrir

Seleccionando dicho submenú, se abre el cuadro de diálogo de Microsoft Windows “*Abrir*”, con la extensión por defecto “*.TXT*” (figura A5.2). El programa también permite abrir archivos con extensión “*.CAL*”⁽³⁾ o cualquier otra “**.**”

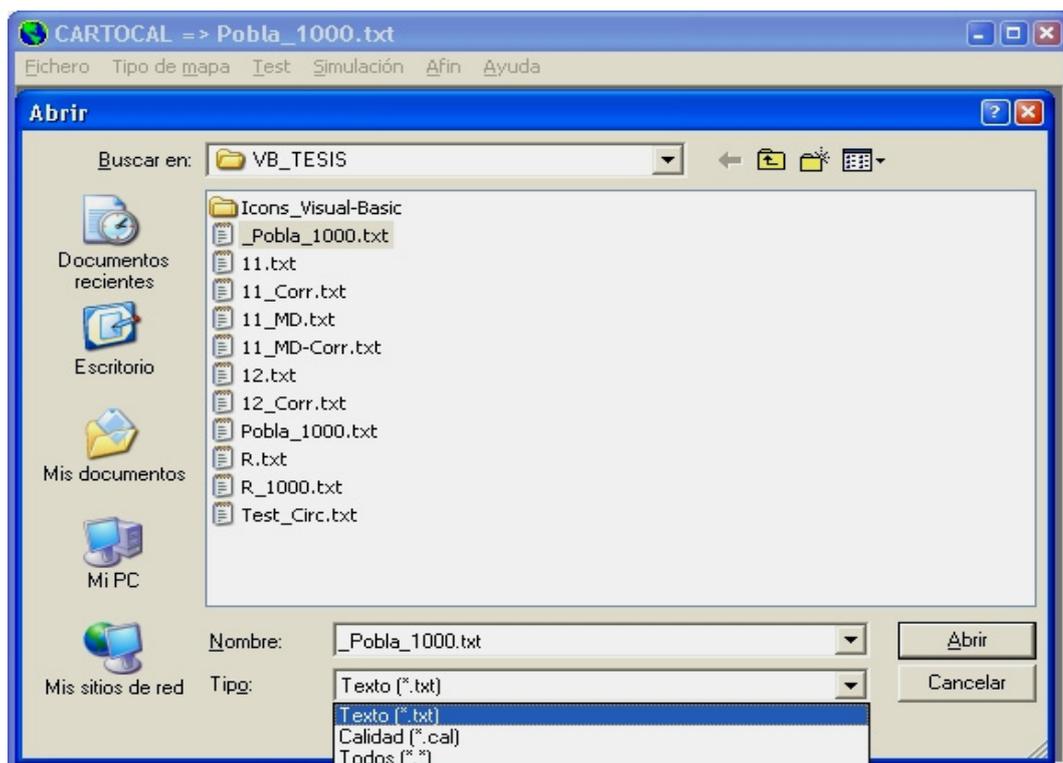
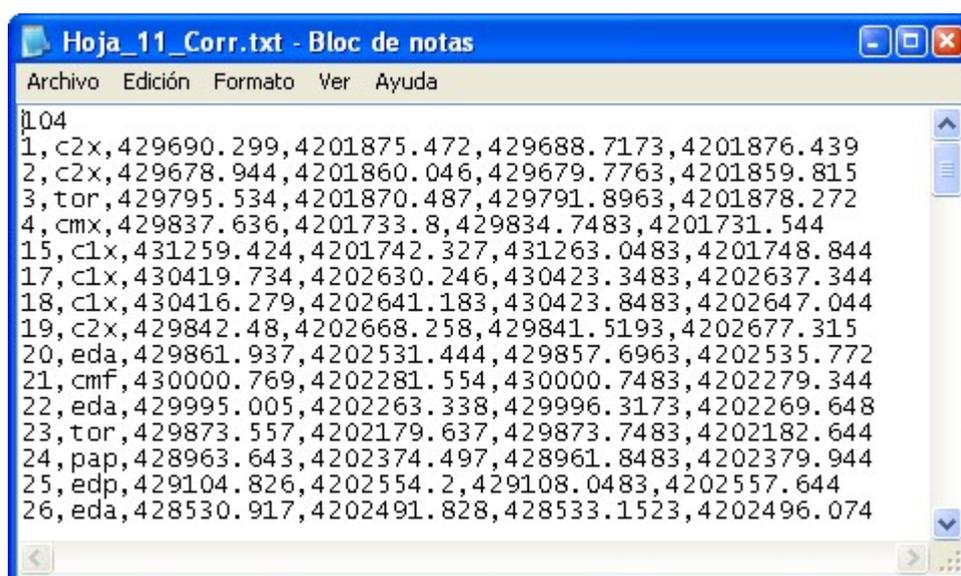


Figura A5.2: Cuadro de diálogo “Abrir”

³ La versión 1 de Cartocal trabaja con archivos ASCII de extensión “*.CAL*”. La versión 2 permite abrir correctamente este tipo de archivos.

El formato ASCII en el que deberán encontrarse los datos es el siguiente:

- En la primera fila aparecerá el número total de puntos de control que componen el archivo
- En las siguientes líneas aparecerán, separados por comas (“,”), los datos de cada uno de los puntos de control (un punto por cada línea) de la siguiente forma (figura A5.3):
 - o Número de punto
 - o Código identificador del tipo de punto de control entre comillas (“”)
 - o Coordenada X del punto de control medida sobre el mapa.
 - o Coordenada Y del punto de control medida sobre el mapa.
 - o Coordenada X del punto de control medida sobre la fuente de mayor exactitud.
 - o Coordenada Y del punto de control medida sobre la fuente de mayor exactitud.



```
104
1, c2x, 429690.299, 4201875.472, 429688.7173, 4201876.439
2, c2x, 429678.944, 4201860.046, 429679.7763, 4201859.815
3, tor, 429795.534, 4201870.487, 429791.8963, 4201878.272
4, cmx, 429837.636, 4201733.8, 429834.7483, 4201731.544
15, c1x, 431259.424, 4201742.327, 431263.0483, 4201748.844
17, c1x, 430419.734, 4202630.246, 430423.3483, 4202637.344
18, c1x, 430416.279, 4202641.183, 430423.8483, 4202647.044
19, c2x, 429842.48, 4202668.258, 429841.5193, 4202677.315
20, eda, 429861.937, 4202531.444, 429857.6963, 4202535.772
21, cmf, 430000.769, 4202281.554, 430000.7483, 4202279.344
22, eda, 429995.005, 4202263.338, 429996.3173, 4202269.648
23, tor, 429873.557, 4202179.637, 429873.7483, 4202182.644
24, pap, 428963.643, 4202374.497, 428961.8483, 4202379.944
25, edp, 429104.826, 4202554.2, 429108.0483, 4202557.644
26, eda, 428530.917, 4202491.828, 428533.1523, 4202496.074
```

Figura A5.3: Archivo ejemplo con los datos de la Hoja (926) 1-1 del ICA corregida de errores sistemáticos. El archivo ASCII esta formado por 104 puntos de control.

Con la instalación del programa se adjuntan varios archivos de ejemplo.

La generación de este tipo de archivos se puede realizar fácilmente empleando una hoja de cálculo (p.e. Microsoft Excel ©, ver apartado “Visualizar resultados” al final de este documento).

El nombre del fichero abierto aparece indicado en la barra principal del programa: CARTOCAL => "nombre del archivo". Por defecto, al iniciar el programa se abre automáticamente el archivo "Pobla_1000.txt" (4).

Fichero → Guardar como

Sobre el menú "Fichero" aparece el submenú "Guardar como" (figura A5.4). Seleccionando esta opción, se abre el cuadro de diálogo de Microsoft Windows "Guardar como", con la extensión por defecto "*.ERR" (archivos de errores posicionales) o cualquier otra "*.*". El programa genera un archivo ASCII en el que, tras una cabecera y separando los datos por comas, se presenta la siguiente información (figura A5.5):

- Número del punto de control
- Error en la componente X
- Error en la componente Y

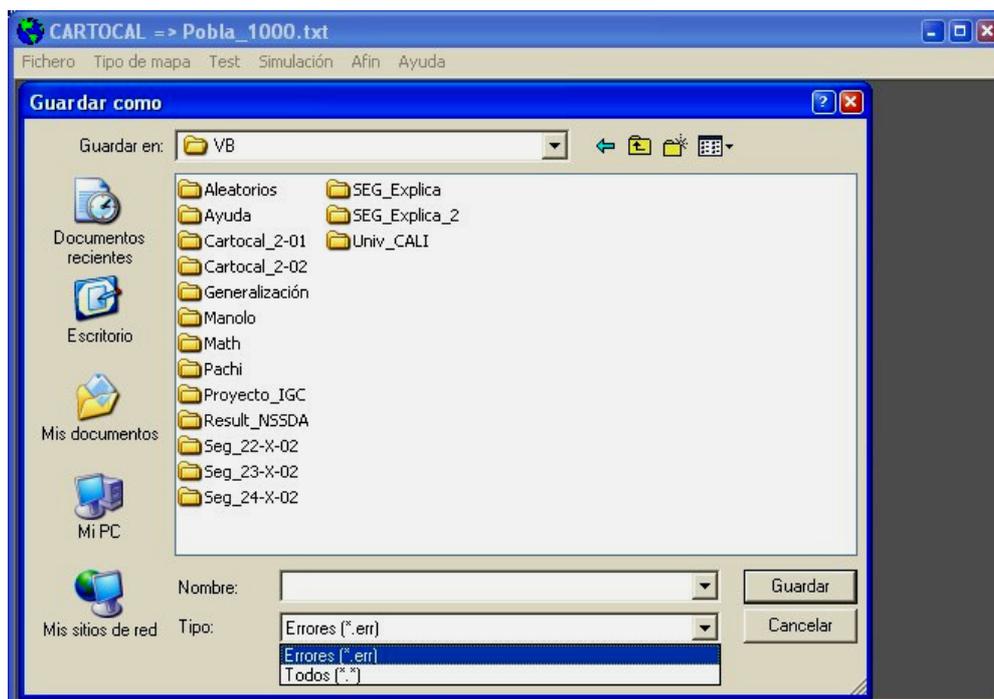
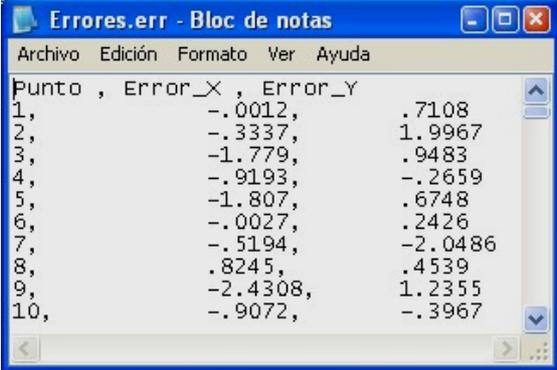


Figura A5.4: Cuadro de diálogo "Guardar como"

⁴ El archivo de ejemplo "Pobla_1000.txt" contiene 1000 puntos con errores aleatorios generados según una distribución $N(0,1)$



The screenshot shows a Notepad window with the following content:

Punto	Error_X	Error_Y
1,	-.0012,	.7108
2,	-.3337,	1.9967
3,	-1.779,	.9483
4,	-.9193,	-.2659
5,	-1.807,	.6748
6,	-.0027,	.2426
7,	-.5194,	-2.0486
8,	.8245,	.4539
9,	-2.4308,	1.2355
10,	-.9072,	-.3967

Figura A5.5: Ejemplo de presentación de resultados en el archivo con los errores detectados en el archivo "Errores.err"

La visualización de este tipo de archivos se puede realizar fácilmente empleando una hoja de cálculo (p.e. Microsoft Excel ©, ver apartado "Visualizar resultados" al final de este documento).

Fichero → Salir

Finaliza y cierra el programa Cartocal.

Tipo de mapa → Seleccionar tipo ...

Sobre el segundo menú desplegable ("*Tipo de mapa*") se abre la opción para la selección del tipo de mapa que se va a analizar. Despliega un formulario en el que se muestra la información sobre la desviación típica estimada a priori para un mapa en función de su escala (0.2mm + 25 %). El usuario debe introducir dicho dato de forma manual. Por defecto toma el valor de $\sigma_{\text{mapa}} = 1 \text{ m}$ (figura A5.6).

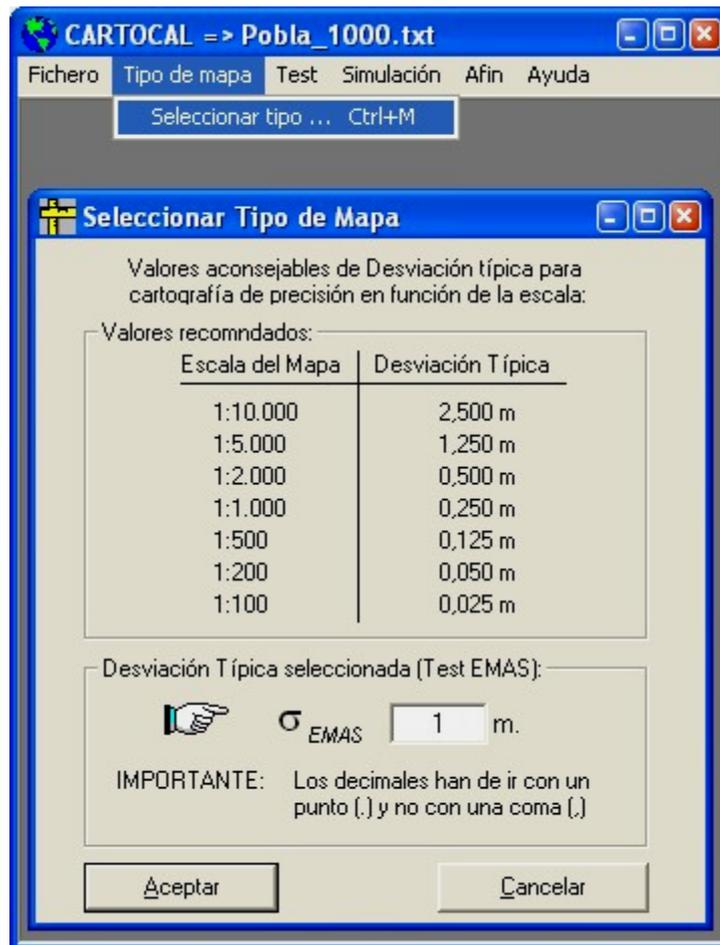


Figura A5.6: Formulario con la desviación típica a priori estimada para una cartografía a escala 1/4.000.

Test → Parámetros estadísticos

El tercer menú desplegable muestra al usuario los cuatro estándares con los que trabaja el programa (EMAS, NMAS, ASPRS y NSSDA), así como los estimadores circulares. Para el primero de ellos permite la opción de seleccionar el nivel de confianza parcial y final con el que trabajar en el submenú “*Parámetros estadísticos*”.

Al seleccionar esta opción, se presenta un formulario (figura A5.7) en el que el usuario puede escoger el nivel de confianza con el que realizar cada una de las pruebas parciales⁵. Por defecto, el programa calcula los estadísticos *t-Student* y *Chi-Cuadrado* empleando las librerías de Microsoft Excel ©. Si se desactiva la casilla superior

⁵ Detección de errores sistemáticos (*t-Student*) en X e Y; y estudio de la varianza (*Chi-Cuadrado*) en X e Y.

izquierda, el programa realiza los cálculos mediante interpolación sobre tablas estadísticas.

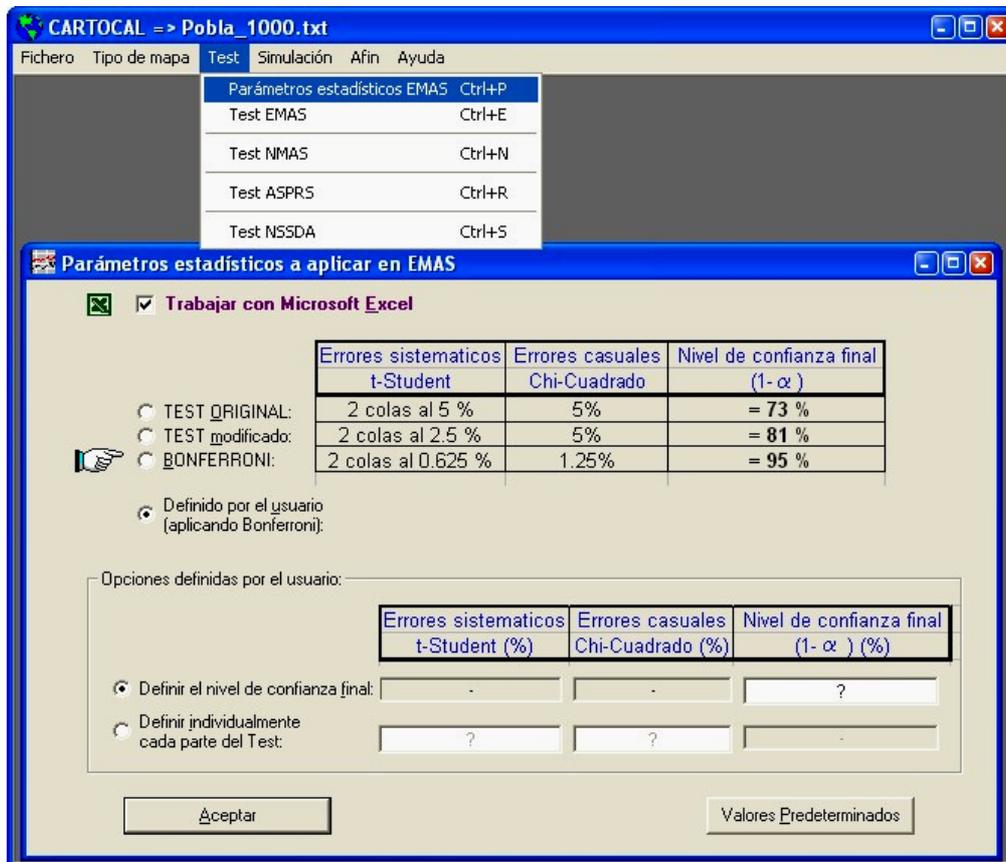


Figura A5.7: Formulario con los parámetros para el establecimiento del nivel de confianza final⁶. Asimismo, por defecto trabaja al 95 % de nivel de confianza aplicando Bonferroni, aunque también permite:

- Aplicar el test original, para cada componente (X e Y) con 2 colas al 5% de nivel de significación para la *t-Student* y 1 cola del 5 % para la *Chi-Cuadrado*. El nivel de confianza total es del 73 %.
- Aplicar el test modificando para cada componente (X e Y) con 2 colas para la *t-Student* (5 % en total) y con 1 cola del 5 % para la *Chi-Cuadrado*. El nivel de confianza total es del 71 %.
- Aplicar el test corregido por Bonferroni: para cada componente (X e Y) 2 colas para la *t-Student* (1.25% en total) y con 1 cola de 1.25 % para la *Chi-Cuadrado*. El nivel de confianza total es del 95 %.

⁶ Se establece bajo 5 opciones: estándar original (1- α =73 %), estándar modificado (81%), estándar aplicando Bonferroni (95%) o definiendo por el usuario⁶ el nivel de confianza final o parcial.

- Definir por el usuario el nivel de confianza total, distribuyendo los niveles de confianza parciales de forma proporcional aplicando Bonferroni.
- Definir por el usuario el nivel de confianza parcial para una de las pruebas (*t-Student* o *Chi-Cuadrado*), calculando el resto de forma proporcional aplicando Bonferroni.

Estas dos últimas opciones tan solo están disponibles si se trabaja empleando las librerías de Microsoft Excel © dado que, de no ser así, los estadísticos *t-Student* y *Chi-Cuadrado* se obtienen mediante interpolación en tablas estadísticas.

Test → Test EMAS

El siguiente submenú desplegable muestra un formulario con los resultados tras el cálculo del estándar EMAS (figura A5.8).

The screenshot shows the 'CARTOCAL' application window with the 'Test' menu open and 'Test EMAS' selected. The 'Test EMAS' dialog box is displayed, showing the following data:

Errores:	
Errores máximos	Pto nº
Ex= 18.7166	43
Ey= -32.773	42
Errores medios:	
Emx= 0	
Emy= 0.0119	
Desviaciones medias:	
σ_{mx} = 4.2903	
σ_{my} = 5.6526	

Test EMAS parameters and results:

- Archivo: C:_Tesis\WB\Hoja_12_Corr.txt
- Nº puntos: 100
- Desviación estimada σ_{EMAS} : 4.5
- Errores sistemáticos: Sx: Verdadero, Sy: Verdadero
- Errores casuales: Cy: Verdadero, Cy: Falso
- El mapa ha sido: **DENEGADO**

Buttons at the bottom: Recalcular, Recalcular y salir, Cancelar.

Figura A5.8: Formulario con el resultado de aplicar el estándar EMAS al archivo de puntos de control "Hoja_12_Corr.txt" bajo una desviación a priori de 4.5m.

En el formulario aparece la siguiente información:

- Nombre del archivo. Éste se sustituye por “ \rightarrow ÚLTIMA SIMULACIÓN \leftarrow ” en el caso de que el archivo abierto haya sido sustituido por un proceso de simulación de poblaciones. Esta información también se muestra en todo momento sobre la barra principal del programa.
- Número de puntos de control: el número total de puntos que contiene el archivo.
- Desviación estándar a priori para la población (σ_{EMAS}). Este valor puede ser modificado para efectuar nuevos cálculos.
- Errores:
 - o Error máximo en valor absoluto para la componente X y número del punto en el que se da dicho error.
 - o Error máximo en valor absoluto para la componente Y, así como el número del punto en el que se da dicho error.
 - o Error medio para la componente X.
 - o Error medio para la componente Y.
 - o Desviación media para la componente X.
 - o Desviación media para la componente Y.
- Test EMAS:
 - o Errores sistemáticos (S_x, S_y): superación o rechazo del estándar en función de si se han detectado errores sistemáticos en la componente X o Y. El programa muestra el rótulo “Verdadero” cuando se supera el estándar (no contiene errores sistemáticos, se acepta H_0) o “Falso” en caso contrario.
 - o Errores casuales (C_x, C_y): superación o rechazo del estándar en función de si se han detectado errores casuales en la componente X o Y. El programa muestra el rótulo “Verdadero” cuando se supera el estándar (los errores casuales se encuentran por debajo de los límites tolerables, se acepta H_0) o “Falso” en caso contrario.
 - o “El mapa ha sido DENEGADO” cuando en alguna de las pruebas anteriores no se ha superado el estándar, o “El mapa ha sido ACEPTADO” cuando se han superado todas ellas.

Para el caso de efectuar alguna modificación sobre la desviación a priori, el programa permite al usuario emplear la opción de “Recalcular” el test EMAS (mostrando los nuevos resultados), emplear la opción de “Recalcular y salir” cerrando el formulario, o “Cancelar” cerrando el formulario sin efectuar ningún cálculo nuevo.

Test → Test NMAS

El siguiente submenú desplegable muestra un formulario con los parámetros y resultados tras el cálculo del estándar NMAS (figura A5.9). El programa indica:

- El nivel de exigencia del test NMAS:
 - Aplicar el límite de tolerancia del estándar NMAS con 1/30” o 1/50” en función de si la escala es mayor o menor que 1/20.000.
 - Aplicar un límite de tolerancia definido por el usuario.
- Permite definir la escala del mapa. Por defecto trabaja con la escala 1/4.000 dado que ésta es la que se corresponde a un mapa con $\sigma_{\text{priori}}=1\text{m}$.
- Muestra el resultado de aplicar el estándar NMAS, indicando si “*Mapa ha sido ACEPTADO / RECHAZADO*” de acuerdo al estándar EMAS.

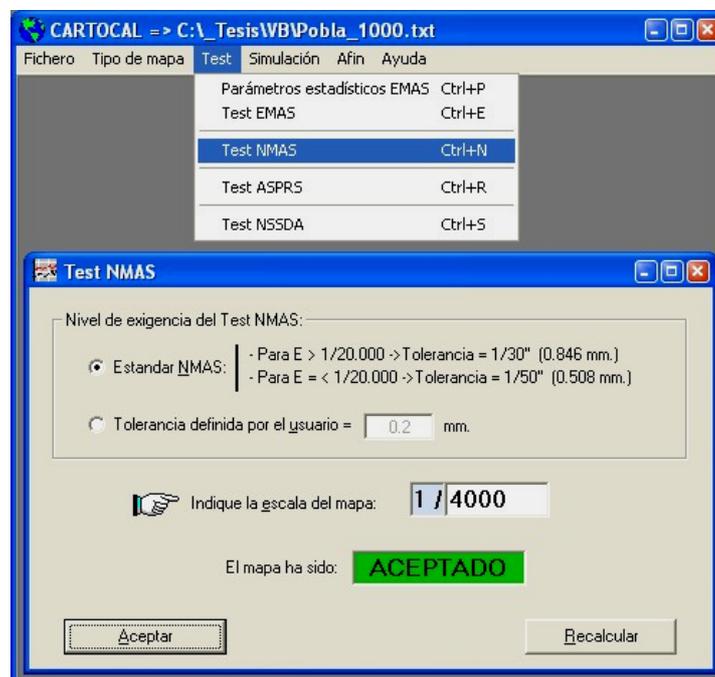


Figura A5.9: Formulario con el resultado de aplicar el estándar original NMAS para una cartografía a escala 1/4.000 (archivo “Pobla_1000.txt”).

Test → Test ASPRS

El siguiente submenú desplegable muestra un formulario con los resultados tras el cálculo del estándar ASPRS (figura A5.10). El programa indica:

- La desviación estimada para los mapas de Clase *I*, Clase *II* y Clase *III*. Dichos campos se calculan automáticamente en función de la σ_{priori} introducida en los formularios previos. Por defecto, dichos valores son 1, 2 y 3 m respectivamente.
- La categoría de mapa con la que cumple la muestra seleccionada, de acuerdo al estándar ASPRS

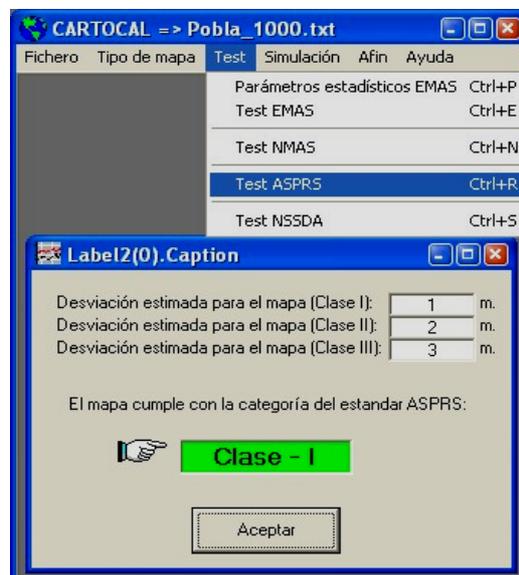


Figura A5.10: Formulario con el resultado de aplicar el estándar ASPRS para una cartografía a con desviación estándar a priori de 1 m (E: 1/4.000).

Test → Test NSSDA

El siguiente submenú muestra un formulario con los parámetros y resultados tras aplicar el estándar NSSDA (figura A5.11a). El programa indica:

- El nivel de confianza con el que se ejecuta el estándar. Por defecto, el programa trabaja al 95 %, pero permite al usuario modificarlo.
- La selección de la hipótesis a priori: $RMSE_X=RMSE_Y$ ó $RMSE_X \neq RMSE_Y$
- El resultado del cálculo de la exactitud horizontal para el estándar NSSDA al nivel de confianza indicado. Dicha información deberá aparecer indicada en la metainformación de la cartografía.

El programa permite “Recalcular” el estándar en el caso de efectuar algún cambio (p.e. el nivel de confianza), o “Aceptar” y cerrar el formulario.

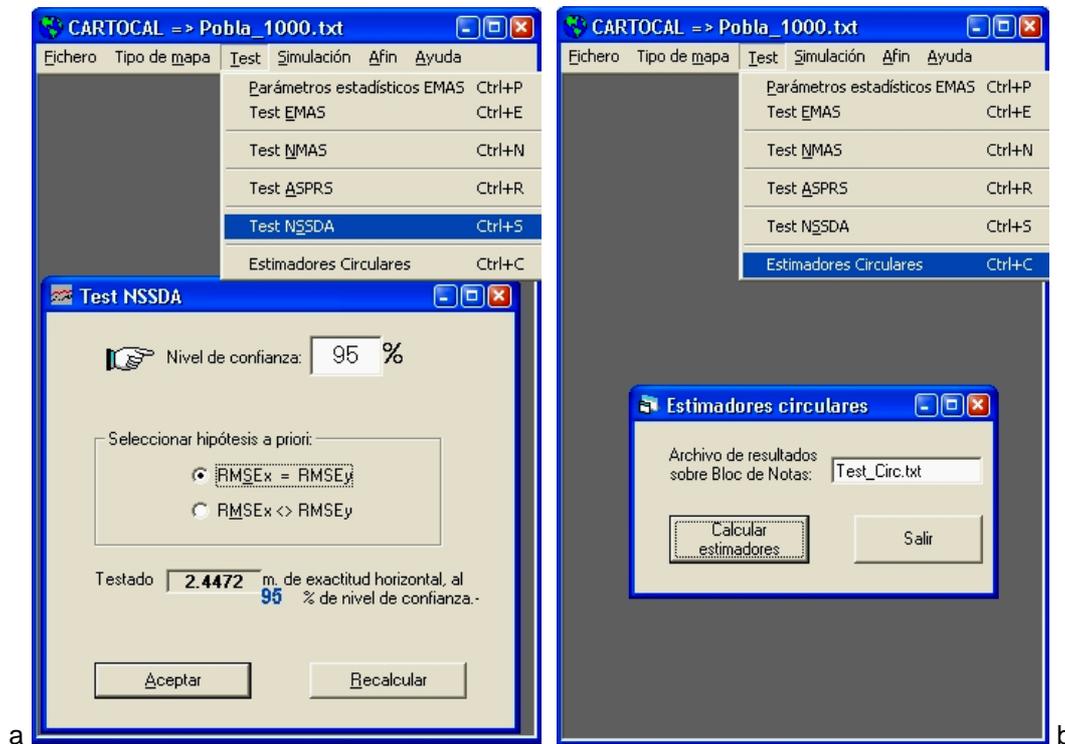


Figura A5.11: (a) Formulario con el resultado de aplicar el estándar NSSDA al 95% de nivel de confianza. Los errores se distribuyen de forma similar para cada componente ($RMSE_x = RMSE_y$); y cálculo de los estimadores circulares (b).

Test → Estimadores circulares

Realiza el cálculo de los estimadores circulares Ómnibus, Mardia y determinación de la dirección dominante, presentado los resultados sobre el Bloc de Notas (figura A5.11b).

Simulación → Simulación aleatoria de muestras

Tras la aplicación de los estándares a una determinada muestra de puntos de control, el siguiente menú desplegable presenta las opciones para el proceso de simulación. El programa despliega un formulario en el que permite introducir los parámetros con los que realizar la simulación en la elección de muestras aleatorias (figura A5.12):

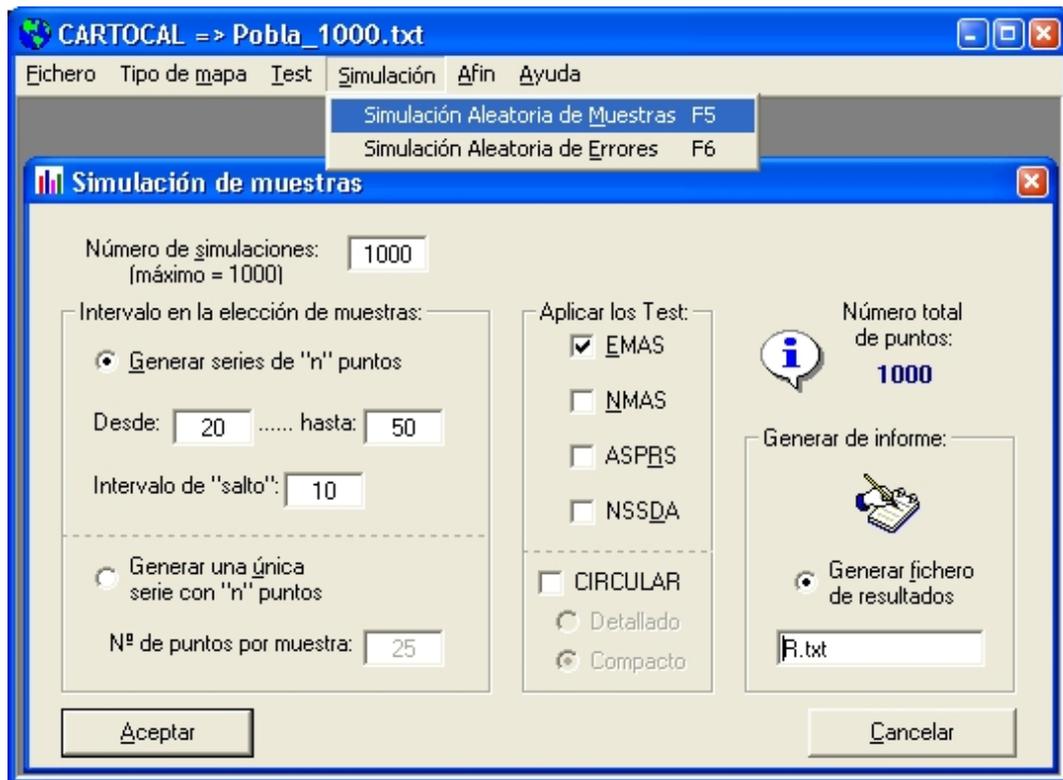


Figura A5.12: Formulario para realizar el proceso de simulación de muestras aleatorias⁷.

- En primer lugar se solicita el número de muestras⁸ que se escogerán aleatoriamente. El programa permite la repetición de una misma muestra dado que, esta situación es factible en casos reales al poder escoger dos operarios los mismos puntos de control por su fácil identificación, distribución y acceso.
- Intervalo en la elección de muestras:
 - Generando series de “*n*” puntos. Dicho valor se puede programar automáticamente indicando el valor inicial y final del intervalo, así como la cadencia o “salto” (“*Steep*”).
 - Generando una única serie de muestras aleatorias con “*n*” puntos de control.
- Generación de informe:
 - Se crea un fichero ASCII de resultados con el nombre que se indique.
- Aplicar los test:

⁷ En la figura aparece el ejemplo con el archivo poblacional “*Pobla_1000.txt*”: se escogerán 1000 muestras de 20 puntos de control y se aplicará el estándar EMAS; se escogerán otras 1000 muestras de 30 puntos y se realizará el mismo cálculo ... y así hasta alcanzar un tamaño muestral de 50 puntos de control. Los resultados se escribirán en el archivo ASCII “*R.txt*”.

⁸ Este se encuentra limitado a un máximo de 1000 muestras (valor por defecto).

- Permite seleccionar uno o varios test a aplicar a un mismo tiempo sobre cada una de las muestras: NMAS, EMAS, ASPRS, NSSDA y los estimadores circulares (Ómnibus, Mardia y dirección dominante).

El programa permite “*Aceptar*” y comenzar el proceso de simulación o “*Cancelar*” la operación y cerrar el formulario sin efectuar ningún cálculo. Durante el proceso de simulación, el programa muestra una barra de progreso en la que se indica el porcentaje de los cálculos que se han realizado (figura A5.13).

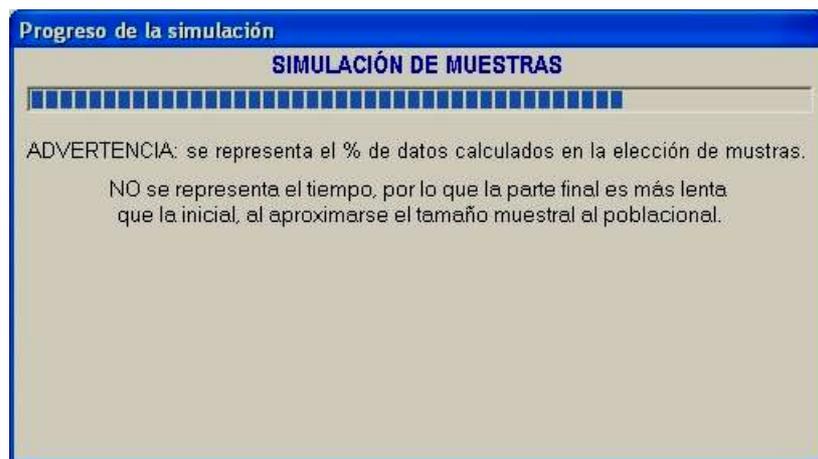


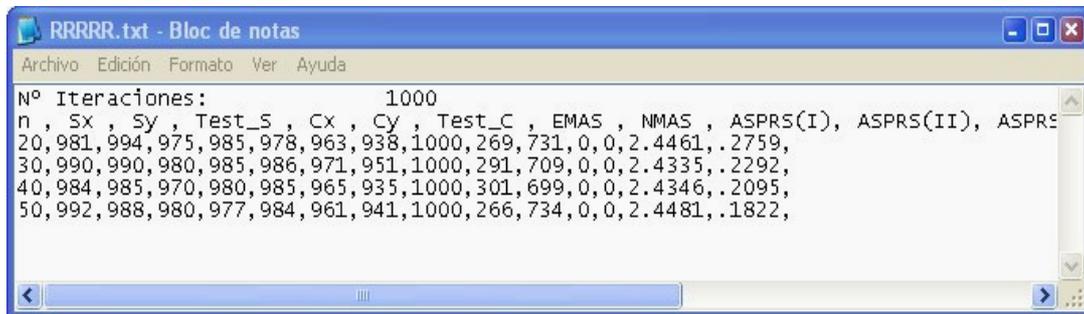
Figura A5.13: Barra de progreso durante la simulación en la elección de muestras.

Tras la simulación el programa cierra el formulario y retorna a la ventana principal de Cartocal. Dado que la muestra principal no ha cambiado, el archivo de datos cargado sigue siendo el mismo (en el ejemplo: “*Pobla_1000.txt*”). Para visualizar los resultados⁹, deberá abrirse el fichero ASCII (figura A5.14). En él aparece:

- El número de muestras que han superado cada una de las pruebas parciales del estándar EMAS (sitematismos en X e Y, aleatoriedad en X e Y), así como los valores totales.
- El número de muestras aceptadas para el NMAS.
- Número de muestras perteneciente a cada una de las clases del ASPRS.
- Valor promedio y desviación típica después de aplicar el estándar NSSDA a las 1000 muestras aleatorias.

⁹ Los resultados se pueden observar más claramente importándolos con una hoja de cálculo (ver el apartado “Visualizar resultados”).

- Número de muestras que no poseen sistematismos al aplicar los estimadores circulares (Ómnibus y Mardia), así como el valor medio de la dirección dominante junto con su intervalo de confianza.



```

RRRRR.txt - Bloc de notas
Archivo Edición Formato Ver Ayuda
Nº Iteraciones: 1000
n , Sx , Sy , Test_S , Cx , Cy , Test_C , EMAS , NMAS , ASPRS(I), ASPRS(II), ASPRS
20,981,994,975,985,978,963,938,1000,269,731,0,0,2.4461,.2759,
30,990,990,980,985,986,971,951,1000,291,709,0,0,2.4335,.2292,
40,984,985,970,980,985,965,935,1000,301,699,0,0,2.4346,.2095,
50,992,988,980,977,984,961,941,1000,266,734,0,0,2.4481,.1822,

```

Figura A5.14: Resultado de aplicar los estándares EMAS, NMAS, ASPRS y NSSDA a 1000 muestras aleatorias de tamaño muestral 20, 30, 40 y 50.

Simulación → Simulación aleatoria de errores

El siguiente submenú presenta el formulario de “*Simulación aleatoria de errores*” (figura A5.15). En este caso se deja de trabajar con un archivo de datos concreto que actuará como población ya que las poblaciones son generadas automáticamente. Así, se presenta un formulario en el que introducir los parámetros con los que se desea realizar la simulación de errores en las poblaciones:

- Simulación en la elección de muestras:
 - Al igual que en el formulario anterior, permite introducir el número de muestras simuladas, así como la/s serie/s de diferente tamaño muestral a generar.
- Generación de poblaciones aleatorias $N(0, \sigma_p)$:
 - Permite seleccionar el número de poblaciones a generar (“*Generar:*”), el tamaño de éstas (“ N_p ”), así como su desviación (“ σ_p ”).
 - También permite generar series de poblaciones, con una desviación inicial estimada para el mapa σ_{EMAS} . Por defecto aparece el valor de 1. Si en el apartado “*Serie de:*” se indica “0”, durante todo el proceso de simulación se trabajará con una única desviación estimada a priori para la cartografía. Si, por el contrario, se escribe algún valor¹⁰, todo el proceso de simulación se repetirá para diferentes niveles de

¹⁰ El valor ha de ser un número entero. En caso contrario se anula la ejecución de la simulación y se muestra un mensaje de alerta para subsanar el error.

exigencia del estándar EMAS. El valor de la desviación a priori (σ_{EMAS}) se irá reduciendo en “*incrementos de:*” durante tantas veces como el valor introducido en la casilla “*Serie de:*”.

- Los resultados se grabarán en uno o varios archivos ASCII (figura A5.17), pudiendo presentar un informe exhaustivo, con los resultados para cada población, o compacto, con los valores promedio y desviaciones para el total de las poblaciones. En el caso de que se apliquen los estándares para diferentes niveles de exigencia de la desviación a priori, se generarán múltiples archivos cuyo nombre variará en función de dicho valor (figura A5.18).
- Aplicar los test:
 - Permite seleccionar uno o varios test a aplicar a un mismo tiempo sobre cada una de las muestras: NMAS, EMAS, ASPRS NSSDA y estimadores circulares.

El programa permite “*Aceptar*” y comenzar el proceso de simulación cerrando el formulario, o “*Cancelar*” la operación y cerrar el formulario sin realizar ningún tipo de simulación. Durante el proceso de simulación, el programa muestra una doble barra de progreso, en la que se indica el porcentaje de los cálculos que se han realizado en la simulación de muestras y poblaciones (figura A5.16).

Figura A5.15: Formulario para realizar un proceso de simulación de muestras y poblaciones de forma conjunta¹¹.

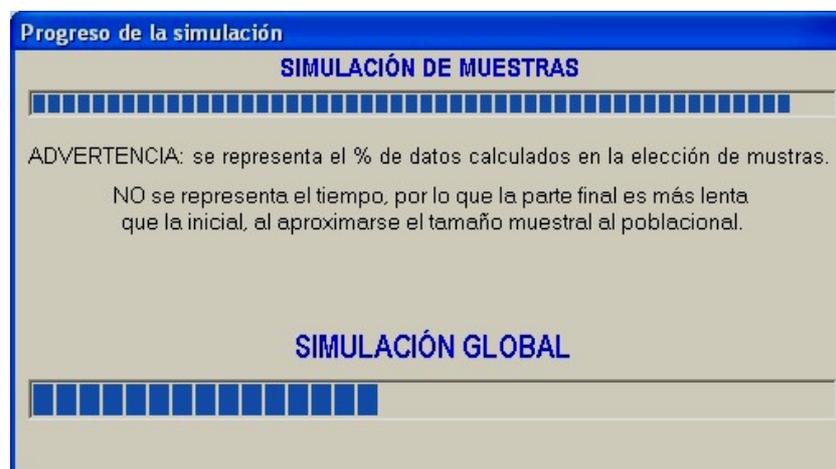


Figura A5.16: Barra de progreso para la simulación en la elección de muestras (simulación de muestras) y poblaciones (simulación global).

¹¹ En el ejemplo, la simulación toma 1000 muestras sobre 100 poblaciones generadas aleatoriamente. Cada población aleatoria estará compuesta por 1000 puntos con errores distribuidos según una función $N(0,1)$ para cada componente (X,Y) . Sobre ésta se escogerá una única serie de 1000 muestras de 25 puntos de control y se aplicarán los estándares EMAS, NMAS y NSSDA. El proceso se repetirá para diferentes niveles de exigencia del estándar EMAS (1, 0.975, 0.95, ... 0.75). Los resultados se escribirán en diferentes archivos "*Resultados_XXX.txt*", siendo "XXX" el valor de la desviación σ_{EMAS} aplicado en cada caso

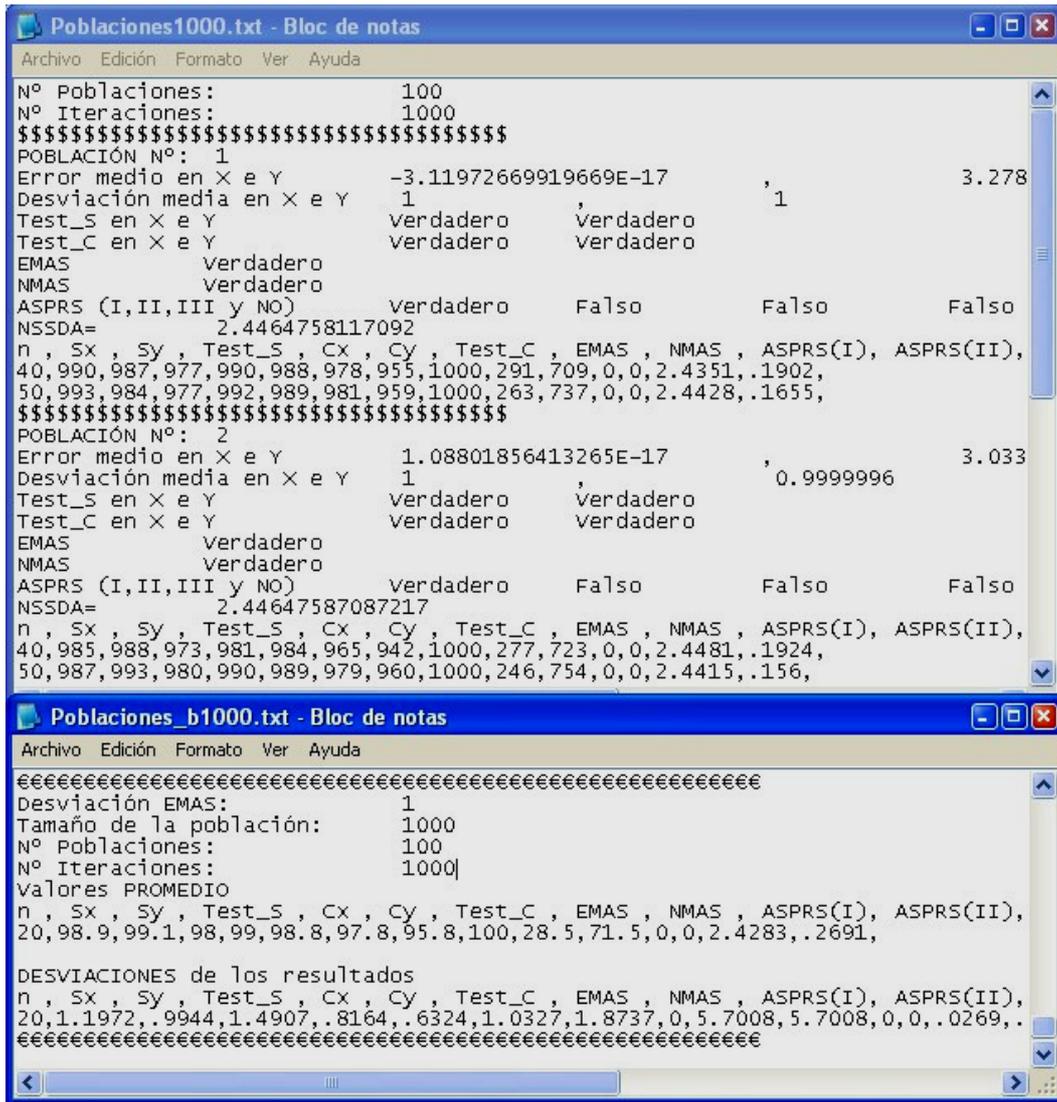


Figura A5.17: Informe de resultados tras el proceso de simulación¹².

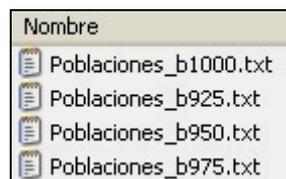


Figura A5.18: Nombres de los archivos con los resultados generados para 4 series de nivel de exigencia del estándar EMAS¹³.

¹² En el ejemplo, se han aplicado los estándares EMAS, NMAS, ASPRS y NSSDA a 1000 muestras (iteraciones) aleatorias de tamaño muestral 40 y 50 (informe exhaustivo “Poblaciones1000.txt”) o sobre un tamaño muestral de 20 puntos (informe compacto “Poblaciones_b1000.txt”). Al final del informe compacto aparecen las desviaciones obtenidas sobre las 100 poblaciones de 1000 puntos cada una.

¹³ En el ejemplo, se han generado los archivos para $\sigma_{EMAS} = 1.000$ (Poblaciones_b1000.txt), 0.975 (Poblaciones_b975.txt), 0.950 (Poblaciones_b950.txt) y 0.925 (Poblaciones_b925.txt).

Afín → Transformación afín

El siguiente menú desplegable muestra la opción para cuantificar y eliminar los errores sistemáticos mediante el ajuste por mínimos cuadrados (mm.cc.) de una “Transformación afín” (figura A5.19).

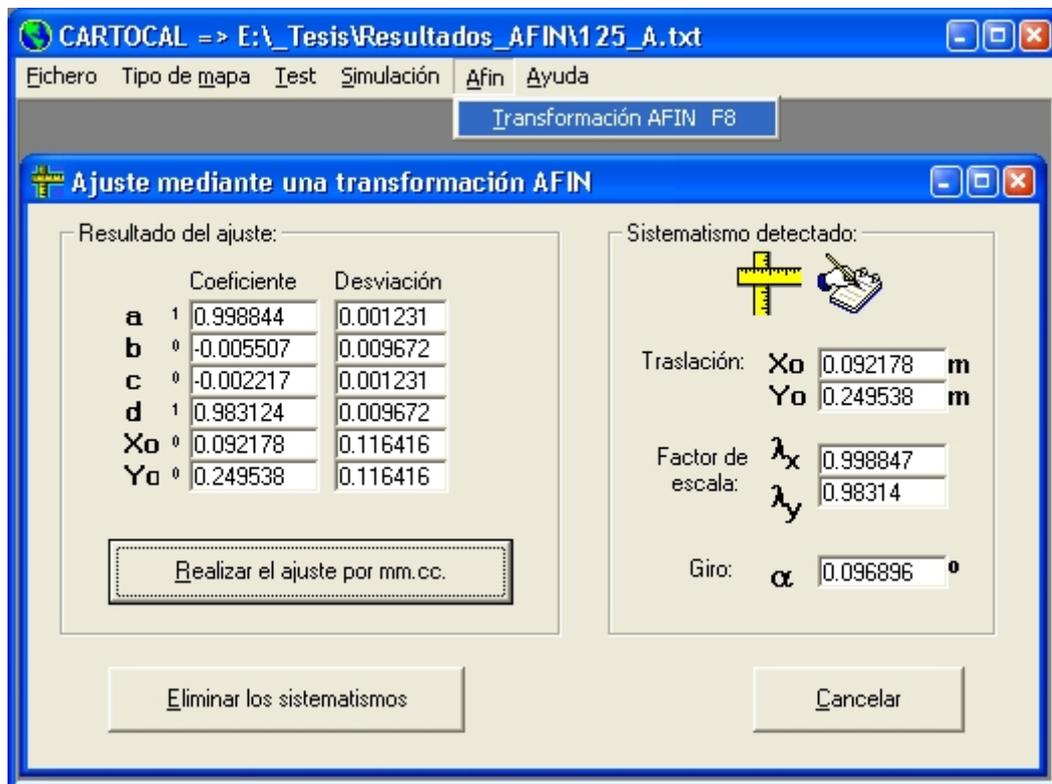


Figura A5.19: Resultados tras el ajuste por mm.cc. de una transformación afín con dos traslaciones (X_0 , Y_0), dos factores de escala (λ_x , λ_y) y un giro (α).

El formulario permite al usuario realizar el ajuste por mm.cc., visualizar los resultados y salir del mismo eliminando los sistematismos detectados o cancelando la operación.

Se despliega la siguiente información:

- En “Resultados del ajuste” se muestran los coeficientes de los 6 parámetros del ajuste (a , b , c , d , X_0 , Y_0), así como sus desviaciones. Cuanto menores sean las desviaciones, mejor será el ajuste y los sistematismos se habrán detectado con una mayor precisión.
- En “Sistematismo detectado” se presentan los valores detectados, tras despejar en función de los coeficientes anteriores, de traslación (X_0 , Y_0), factores de escala (λ_x , λ_y) y giro (α). La obtención del valor del giro es redundante pues se puede hacer partiendo de los parámetros a - c ó b - d .

Permite “*Eliminar los sistematismos*” aplicando los parámetros detectados y corrigiendo las coordenadas de la BDG, o “*Cancelar*” la operación y cerrar el formulario sin efectuar ningún cálculo.

Ayuda → Ayuda de CARTOCAL

Abre el archivo de ayuda “*Manual_Usuario_CARTOCAL.pdf*” (figura A5.20). Para poderlo visualizar es necesario tener instalado el programa Acrobat Reader ©¹⁴.



Figura A5.20: Manual de usuario de CARTOCAL 2.0 en formato PDF.

¹⁴ Se ha de disponer de la versión 5 o superior. El programa se puede descargar de forma gratuita en www.adobe.com o instalarlo desde el CD de instalación con el programa CARTOCAL 2.0.

Ayuda → Acerca de ...

El último de los submenús de Cartocal muestra el formulario “Acerca de ... CARTOCAL” (figura A5.21). En él se muestran los datos acerca de la fecha y versión del programa, la Universidad de Jaén, los grupos de investigación participantes, el autor del programa e información de contacto.



Figura A5.21: Formulario con información Acerca de ... CARTOCAL 2.0.

Para obtener más información, pueden dirigirse a:

- Director de la Tesis: **Francisco J. Ariza López** (fjariza@ujaen.es) Catedrático en la Universidad de Jaén, Departamento y Área de Ingeniería Cartográfica, Geodésica y Fotogrametría.
- Autor del programa y de la Tesis: **Alan D.J. Atkinson Gordo** (atkinson@unex.es, www.unex.es/igpu/Alan), Profesor en la Universidad de Extremadura (www.unex.es), Departamento de Expresión Gráfica, Área de Ingeniería Cartográfica, Geodésica y Fotogrametría.

- **Grupo de Investigación en Ingeniería Cartográfica** (www.ujaen.es/dep/ingcar/Investigacion/Grupo_inv_Ing_Cartografica) de la Universidad de Jaén (www.ujaen.es).
- **Grupo de Investigación en Ingeniería Geomática y Patrimonio Urbano** (www.unex.es/igpu) de la Universidad de Extremadura (www.unex.es).

Visualizar resultados

Los archivos ASCII empleados por Cartocal 2.0, tanto para obtener información sobre los puntos de un determinado control de exactitud posicional, como para la presentación de resultados, son generados de forma sencilla para que puedan ser tratados mediante una hoja de cálculo. Un claro ejemplo sería la utilización de Microsoft Excel ©.

En la figura A5.22 se aprecia cómo se podrían introducir o tratar los datos procedentes de un control posicional. Para exportarlos a un archivo ASCII que pueda capturar Cartocal, deberá seleccionarse “*Guardar como*” con la opción de texto separado por tabulador. Seguidamente se podrán reemplazar semiautomáticamente los tabuladores por comas para que lo reconozca Cartocal.

	A	B	C	D	E	F	G
1	Nº Total de puntos	104	Código	Xmapa	Ymapa	Xterreno	Yterreno
2	Nº de punto	1	c2x	429690.299	4201875.472	429688.717	4201876.439
3		2	c2x	429678.944	4201860.046	429679.776	4201859.815
4		3	tor	429795.534	4201870.487	429791.896	4201878.272
5		4	cmx	429837.636	4201733.800	429834.748	4201731.544
6		15	c1x	431259.424	4201742.327	431263.048	4201748.844

Figura A5.22: Ejemplo de archivo con los datos de un control de exactitud posicional.

Para visualizar ficheros de resultados, Cartocal genera archivos ASCII separando los datos por comas. Para visualizarlos correctamente, se puede emplear una hoja de cálculo. En el caso de la figura A5.23 se ha empleado Microsoft Excel © con las opciones:

- Abrir archivo en formato texto (“*.TXT”).
- Campos delimitados por caracteres o comas (no de ancho fijo).

- Emplear separadores de espacio y coma (considerando más de un separador junto como uno único).
- Emplear el separador decimal el punto.

	A	B	C
1	Punto	Error_X	Error_Y
2	1	-0.0012	0.7108
3	2	-0.3337	1.9967
4	3	-1.7790	0.9483
5	4	-0.9193	-0.2659

Figura A5.23: Ejemplo visualización de un archivo con los errores posicionales tratados con Cartocal.

Empleando este mismo procedimiento, se pueden visualizar correctamente otros archivos de resultados. En el caso de la figura A5.24 se muestran los resultados tras un proceso de simulación con 1000 muestras aleatorias sobre una misma población. Se ha realizado un proceso de simulación aleatoria en el elección de muestras ($n = 20, 30, 40, y 50$) para los test: EMAS, NMAS, ASPRS y NSSDA.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Nº Iteraciones:			1000											
2	n	Sx	Sy	Test_S	Cx	Cy	Test_C	EMAS	NMAS	ASPRS(I)	ASPRS(II)	ASPRS(III)	ASPRS(IV)	Media NSSDA	Desv. NSSDA
3	20	990	992	982	987	988	976	958	1000	297	703	0	0	2.4225	0.2671
4	30	988	990	978	991	984	976	954	1000	287	713	0	0	2.4325	0.2238
5	40	991	987	978	985	980	965	944	1000	282	718	0	0	2.4362	0.1964
6	50	990	993	983	981	987	968	951	1000	293	707	0	0	2.4375	0.1737

Figura A5.24: Ejemplo visualización de resultados del archivo ASCII de simulación.

La figura A5.25 muestra los resultados tras un proceso de simulación con 100 poblaciones y 1000 muestras aleatorias sobre cada una de ellas, aplicando los estándares EMAS y NMAS. En el informe exhaustivo aparecen los resultados parciales para cada una de las 100 poblaciones.

	A	B	C	D	E	F	G	H	I
1	Nº Poblaciones:	100							
2	Nº Iteraciones:	1000							
3	\$								
4	POBLACIÓN Nº: 1								
5	Error medio en X e Y	-3.1197	3.28E-09						
6	Desviación media en X e Y	1	1						
7	Test_S en X e Y	Verdadero	Verdadero						
8	Test_C en X e Y	Verdadero	Verdadero						
9	EMAS	Verdadero							
10	NMAS	Verdadero							
11	ASPRS (I	II	III y NO)	Verdadero	Falso	Falso	Falso		
12	NSSDA=	2.4464758117092							
13	n	Sx	Sy	Test_S	Cx	Cy	Test_C	EMAS	NMAS
14	40	990	987	977	990	988	978	955	1000
15	50	993	984	977	992	989	981	959	1000
16	\$								
17	POBLACIÓN Nº: 2								
18	Error medio en X e Y	1.0880	3.03E-08						

Figura A5.25: Ejemplo visualización de resultados del informe exhaustivo tras un proceso de simulación de poblaciones.

En el informe compacto (figura A5.26) se muestran los resultados promedio y desviaciones finales para las 100 poblaciones mencionadas (test EMAS).

123	\$								
124	Desviación EMAS:	1							
125	Tamaño de la población:	1000							
126	Nº Poblaciones:	100							
127	Nº Iteraciones:	1000							
128	Valores PROMEDIO								
129	n	Sx	Sy	Test_S	Cx	Cy	Test_C	EMAS	NMAS
130	20	98.9	99.1	98	99	98.8	97.8	95.8	1000
131									
132	DESVIACIONES de los resultados								
133	n	Sx	Sy	Test_S	Cx	Cy	Test_C	EMAS	NMAS
134	20	1.1972	0.9944	1.4907	0.8164	0.6324	1.0327	1.873	1000
135	\$								

Figura A5.26: Ejemplo visualización de resultados del informe compacto tras un proceso de simulación de poblaciones.

A5.2 Manual del programador

El programa se ha desarrollado sobre diferentes formularios son contenedores para controles. El formulario de interfaz de múltiples documentos (MDIForm1.frm) actúa como un contenedor para los 12 formularios secundarios (Form1.frm, Form2.frm ...). Asimismo, existen dos módulos (estándar y de clase). En la figura A52.1 se puede apreciar la estructura de formularios y módulos del programa.

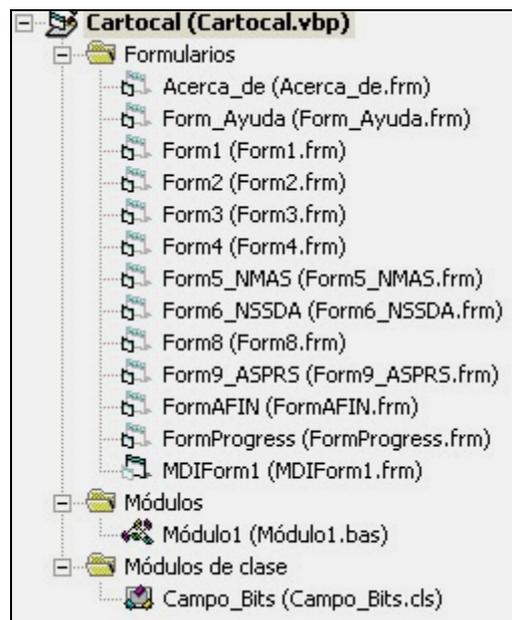


Figura A52.1: Estructura de formularios y módulos de Cartocal.

Bajo el entorno de los formularios indicados en el Manual de Usuario, se ha implementado la aplicación con las funciones y algoritmos que permiten su cálculo.

Para cada módulo y cada formulario, se ha editado un listado con el código del programa explicando detalladamente cómo se realizan los cálculos y las revisiones que ha tenido el programa. Básicamente se ha partido de una programación inicial, así como dos o tres validaciones internas realizadas por el doctorando. Por otra parte, el Dr. José Juan de Sanjosé y, fundamentalmente, el Dr. Francisco J. Ariza, han realizado diferentes validaciones externas del código planteando modificaciones y mejoras que han contribuido a su depuración.

Asimismo, el Dr. Manuel A. Ureña contribuyó en la optimización de parte del código que permite la ordenación y comparación de los datos de forma eficaz.

Toda esta información se puede visualizar en el Anexo 7, listados en formato digital:

- Listado 26: MDIForm1
- Listado 27: Acerca_de
- Listado 28: Form_Ayuda
- Listado 29: Form1
- Listado 30: Form2
- Listado 31: Form3
- Listado 32: Form4
- Listado 33: Form5_NMAS
- Listado 34: Form6_NSSDA
- Listado 35: Form8
- Listado 36: Form9_ASPRS
- Listado 37: FormAFIN
- Listado 38: FormProgress
- Listado 39: Módulo1
- Listado 40: Campo_Bits

A5.3 Aplicaciones AutoLisp

Para el proceso de intercambio de datos entre las coordenadas de los puntos de control sobre la cartografía digitalizada y el programa Cartocal (en formato ASCII), se han empleado dos pequeñas aplicaciones en Auto Lisp:

- Función INSERTAP.LSP: permite la inserción de puntos desde un fichero ASCII a AutoCAD. Dicha función inserta bloques con su código de punto en un dibujo.
- Función CONTAP.LSP: permite la extracción de puntos desde AutoCAD a un fichero ASCII. Dichos puntos estarán definidos por el bloque "P" con los atributos "numeroPunto" y "codigoPunto". Como resultado, devuelve un fichero de texto (*.TXT) con "NumeroPunto, CodigoPunto, X, Y".

El código de ambas aplicaciones aparece en los listados del Anexo 7:

- Listado 41: Función INSERTAP.LSP
- Listado 42: Función CONTAP.LSP

ANEXO 6

Propuesta metodológica: *ExHor*

Dicha propuesta se encuentra ampliamente desarrollada en el Anexo 6.

En este Anexo se presenta una propuesta metodológica para el análisis de la exactitud posicional planimétrica en BDGs. Para ello se ha diseñado un posible “*Estándar de exactitud posicional en información georreferenciada (EPIG)*”, centrado sobre la exactitud horizontal de datos puntuales en el “*Estándar de Exactitud Horizontal en información puntual georreferenciada (ExHor)*” (apartado A6.1). En el apartado A6.2 se adjuntan una serie de anexos para su correcta aplicación. Por último, en el apartado A6.3, se plantea un ejemplo práctico de aplicación del estándar *ExHor*.

A6.1 Propuesta de aplicación

OBJETIVO

El *ExHor* implementa una metodología basada en el análisis estadístico para estimar la exactitud posicional horizontal de elementos puntuales en BDGs, respecto a posiciones georreferenciadas sobre el terreno o sobre una FME.

ALCANCE

El *ExHor* es aplicable a todo tipo de BDGs independientemente de su formato (analógico / digital, vectorial / raster). La BDG puede haber sido obtenida tanto a partir de datos fotogramétricos, imágenes de satélite o levantamientos sobre el terreno.

El estándar se presenta en un lenguaje común para expresar la exactitud y el tipo de errores que afectan a la BDG, para facilitar la identificación de datos espaciales para aplicaciones geográficas.

El estándar no impone los límites o umbrales de exactitud posicional. Si bien se propone una serie de recomendaciones, ha de ser el usuario quien los establezca en función de las necesidades de aplicación de la BDG.

RESPONSABILIDADES

Se establece que el responsable final de la BDG es el director de producción de dicho proyecto. Es a él a quien se recomienda comprobar mediante auditorías internas tanto el proceso global de producción de la BDG, como los resultados previsibles ante el *ExHor*.

Respecto a los responsables del control de calidad, éstos son:

- Responsable del control de calidad. Es el director del proyecto de control de calidad empleando el *ExHor*.
- Responsable del análisis. Es el responsable de los cálculos y procesado de la información.
- Responsable de campo. Es el responsable de la toma de datos (puntos de control) sobre la BDG y la FME.

En la tabla A6.1 se exponen los responsables en cada una de las fases del de control de calidad mediante el *ExHor*.

Tabla A6.1 Responsables en el control de calidad mediante el *ExHor*.

Fase	Responsable
Planificación: selección de la muestra de control	- Responsable del análisis - Responsable del control de calidad
Toma de datos: obtención de las coordenadas de los puntos de control	- Responsable de campo
Estudio previo: verificación de hipótesis previas	- Responsable del análisis
Sistematismos: detección y corrección	- Responsable del análisis
Variabilidad: estudio de la varianza	- Responsable del análisis
Informe final	- Responsable del análisis - Responsable del control de calidad
Auditorías internas al control de calidad mediante el <i>ExHor</i>	- Responsable del control de calidad

DEFINICIONES

Ver Anexo 9: Glosario

EJECUCIÓN

El procedimiento para la aplicación del estándar *ExHor* es el siguiente:

- 1º. Selección de la muestra de control
- 2º. Estudio previo: normalidad, aleatoriedad y valores atípicos
- 3º. Detección y corrección de sistematismos
- 4º. Estudio de la varianza
- 5º. Informe final

A modo de esquema, el diagrama de flujo de la figura A6.1 muestra los pasos a seguir.

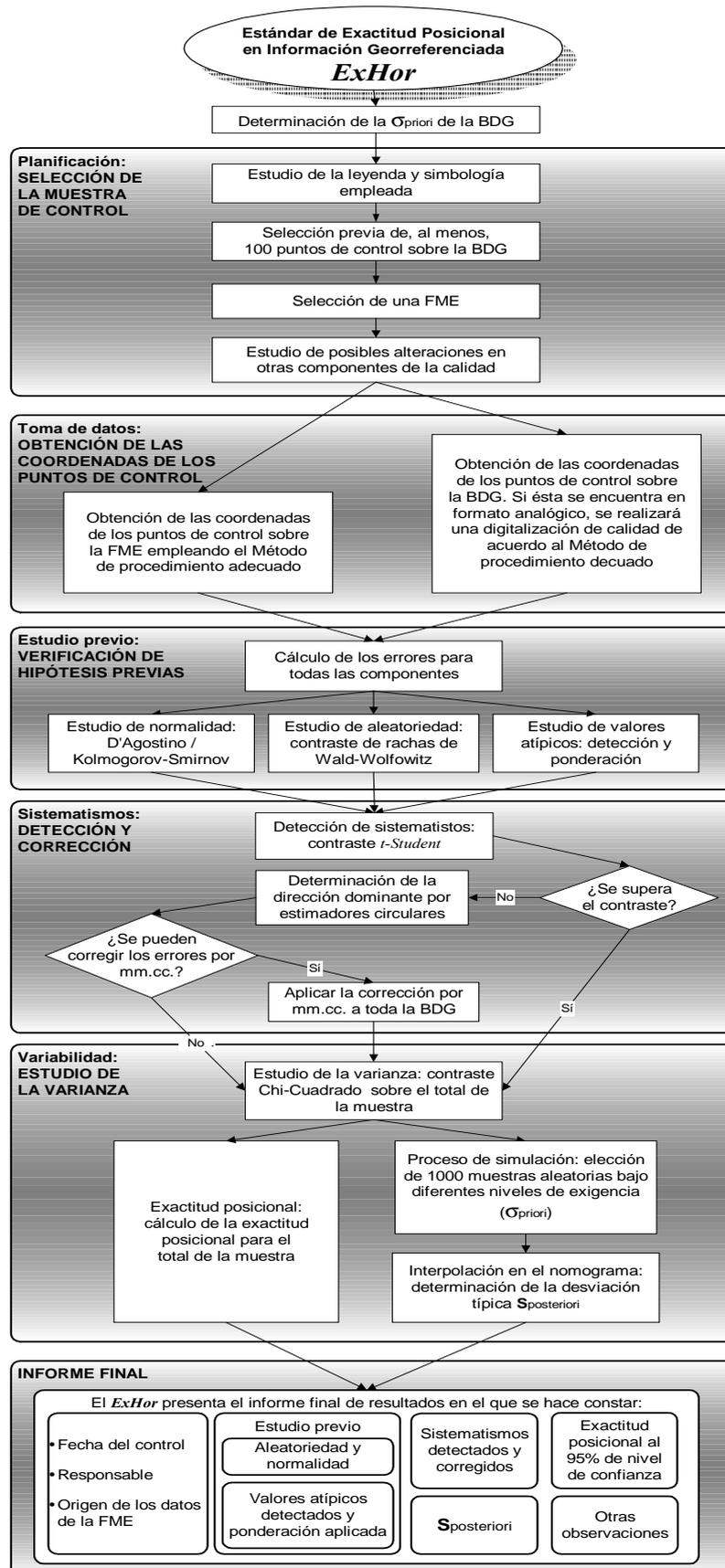


Figura A6.1: Diagrama de flujo de aplicación del estándar ExHor.

1º. Selección de la muestra de control

Se realizará una selección de, al menos, 100 puntos de control sobre la BDG y sobre una fuente, al menos, tres veces más exacta¹.

Para la selección de los puntos de control, deberá realizarse un estudio previo de la leyenda y la simbología empleada. Se realizará una preclasificación del tipo de elementos que aparecen representados por su pertenencia a las diferentes clases que aparecen en la leyenda, así como si están representados en sus dimensiones reales o empleando algún tipo de signo convencional (p.e. edificaciones aisladas de reducido tamaño que no serían visibles sobre la BDG). Se hará una selección previa de los posibles puntos de control, la cual deberá depurarse cuando se realiza la toma de datos.

Asimismo, se realizará un estudio preliminar de la zona en cuanto a posibles alteraciones de otras componentes de la calidad (exactitud temporal, compleción, consistencia lógica ...). Cabe la posibilidad de que, por ejemplo, desde que se ha producido la BDG al momento en el que se realiza el control de calidad, algunos elementos cartografiados hayan sufrido modificaciones.

Los puntos de control deberán reunir las características indicadas en el apartado A62.1. de igual forma, si la FME es la toma de datos sobre el terreno, también deberán tenerse en cuenta los aspectos indicados en dicho apartado

De igual forma, en el caso de que la BDG se encuentre en formato analógico, se deberá proceder a la digitalización² de los puntos de control. Este proceso podrá efectuarse mediante tableta digitalizadora o escaneado y posterior digitalización manual o asistida. Dicha digitalización deberá ser de calidad, empleando un método de procedimiento estandarizado que garantice los resultados finales.

¹ Preferiblemente, la fuente de mayor exactitud serán datos obtenidos directamente sobre el terreno. Para ello se empleará un sistema adecuado que proporcione el nivel de exactitud requerido.

² La digitalización deberá realizarse empleando equipos calibrados geoméricamente siendo, al menos, tres veces más exacta que la BDG.

2º. Estudio previo: normalidad, aleatoriedad y valores atípicos

Una vez se hayan obtenido las coordenadas de los puntos sobre la BDG y la FME, se calculará el error para cada punto de control³, así como los valores promedio mediante las ecuaciones:

$$e_{xi} = x_{mi} - x_{ti} \quad ; \quad \bar{e}_x = \frac{\sum_{i=1}^n e_{xi}}{n}$$

en donde:

- e_{xi} : es el error o incertidumbre en la componente X del punto i
- x_{ti} : la coordenada del punto i sobre el terreno o FME
- x_{mi} : la coordenada del punto i sobre la BDG
- \bar{e}_x : es el error medio en la componente X
- n : es el número total de puntos de control

De forma similar se actuará sobre la otra componente (Y).

En siguiente lugar, se realizará un estudio sobre la normalidad, aleatoriedad y los posibles valores atípicos.

a) Normalidad

La distribución de los errores deberá asemejarse a una función estadística normal. Para ello se efectuará un contraste de normalidad del tipo Kolmogorov-Smirnov o D'Agostino (ver apartado A62.2).

En el caso de que la distribución de los errores no se asemeje a una función Normal, dicho dato deberá quedar reflejado en el informe final.⁴

³ Como ejemplo, se presentan las ecuaciones para la componente X .

⁴ Al no distribirse los errores según una función normal, las consideraciones sobre las que se fundamentan todos los estándares analizados (incluida la propuesta del *ExHor*), no serían aplicables. No obstante, se estima conveniente el continuar con el análisis del resto de los parámetros propuestos para un mejor conocimiento de los errores de la BDG.

b) Aleatoriedad

En este caso, se trata de verificar la hipótesis previa en cuanto a la aleatoriedad en la distribución de los errores en ambas componentes. Para ello, se ha de efectuar un contraste de rachas similar al de Wald–Wolfowitz (ver apartado A62.3).

En el caso de que la distribución de los errores no sea aleatoria, dicho dato también quedaría reflejado en el informe final.

c) Valores atípicos

Se considerarán como valores atípicos aquellos puntos cuyo error en alguna de sus componentes sea superior a $(2,5 \cdot \sigma_{priori})$. La σ_{priori} será la misma que la considerada por el estándar ASPRS (ASPRS, 1989) en la exactitud planimétrica de los mapas de Clase I (tabla A6.2).

Tabla A6.2. Desviación típica a priori estimada para una BDG en función de su escala de representación.

Escala	σ_{priori} (m)
1:50	0,0125
1:100	0,025
1:200	0,050
1:500	0,125
1:1.000	0,25
1:2.000	0,50
1:4.000	1,00
1:5.000	1,25
1:10.000	2,50
1:20.000	5,00

Se detectarán y ponderarán aquellos valores atípicos que se encuentren en la muestra. El procedimiento de ponderación será mediante el método danés indicado en el apartado A62.4. A partir de este momento, todos los cálculos se realizarán con los valores ponderados de los atípicos, sin eliminar ninguno de ellos.

3°. Detección y corrección de sistematismos

La detección de posibles errores sistemáticos se realizará mediante un contraste de hipótesis del tipo *t-Student* similar al empleado en el estándar EMAS (ASPRS, 1989). El procedimiento para cada componente se explica en el apartado A62.5. En el caso de que se detecte algún sistematismo en cualquiera de las componentes analizadas, éste se tratará de cuantificar y corregir tal y como se indica en dicho apartado. Asimismo, deberá determinarse la dirección dominante (junto con su intervalo de confianza) en la que se produce dicho sistematismo.

4°. Estudio de la varianza

Para el estudio de la varianza, el estándar se aplica en 2 fases:

a) Estimación de la desviación poblacional a posteriori.

Mediante un proceso de simulación en la elección aleatoria de muestras de diferente tamaño, se determinará la desviación a posteriori para las posibles muestras escogidas ($S_{\text{posteriori}}$). Éste dato será representativo de la desviación poblacional a posteriori ($\sigma_{\text{posteriori}}$). El procedimiento para su cálculo se indica en el apartado A62.6.

b) Cálculo de la exactitud posicional al 95 % de nivel de confianza.

Para la determinación de la exactitud posicional, se seguirá un proceso similar a la última fase del NSSDA (FGDC, 1998), empleando el total de los puntos de control. El cálculo de dicho valor se expone en el apartado A62.7.

5°. Informe final

Siempre que se aplique el estándar *ExHor*, deberá adjuntarse a la BDG un informe final en el que aparezcan los datos que se indican a continuación en cursiva:

A) Para todo tipo de BDG, deberá indicarse

- Fecha en la que se ha sido realizado el control de calidad mediante el *ExHor*.
- Empresa y/o responsable/s del control de calidad.
- Origen de los datos de la FME y método de captura (p.e. GPS).

B) Se podrá presentar un informe detallado o resumido del estándar:

B.1) Informe detallado:

“De acuerdo con el Estándar de Exactitud Horizontal en información puntual georreferenciada, esta base de datos geográfica posee las siguientes características posicionales en planimetría:”

- *“Los errores se distribuyen de forma normal y aleatoria”.* En el caso de que no se cumpla alguna de estas condiciones, se indicará: *“Los errores no se distribuyen de forma normal y/o aleatoria, por lo que el resto de consideraciones que se presentan son meramente informativas y no deberán considerarse de forma rigurosa.”*
- *“El porcentaje de valores atípicos es de ... %.”* Indicar el número de puntos que han sido ponderados por el método danés (en tanto por ciento).
- *“No posee errores sistemáticos.”* Si, por el contrario, tuviera este tipo de errores, se indicaría la dirección dominante (*“Se han detectado errores sistemáticos en la dirección de ...⁵”*), y no se hubieran podido corregir, se indicaría: *“Dichos errores sistemáticos no han podido ser cuantificados en su totalidad, por lo que no han podido ser corregidos.”* En el caso de que los errores se hubieran podido cuantificar y corregir, se indicaría: *“Los errores sistemáticos detectados [traslación (X_0 , Y_0), factor de escala (λ_x , λ_y) y giro (α)] han sido corregidos y eliminados.”⁵*
- *“La desviación estimada es de $S_{\text{posteriori}}$. Dicho valor es $R(\sigma/S)$ veces más/menos exacto que su valor estimado a priori (σ_{priori}).”⁶* Se indicará que es *“más exacto”* cuando el valor de $R(\sigma/S)$ sea > 1 , y *“menos exacto”* cuando sea < 1 . En el caso de que $R(\sigma/S) = 1$, se indicará: *“Dicho valor de corresponde con su estimación a priori.”*
- *“La exactitud horizontal al 95 % de nivel de confianza es de Exactitud_r”⁷*

⁵ Deberán indicarse de forma cuantitativa, en unidades del a BDG, los valores ajustados por mm.cc.

⁶ Indicar los valores obtenidos por el estándar.

⁷ Idéntico al aplicado para la detección de sistematismos.

- Otras observaciones que se deseen hacer constar (p.e.: referencias a la exactitud temática, consistencia lógica, exactitud temporal ...).

B.1) Informe resumido:

- *“De acuerdo con el estándar **ExHor**; existe un ... % de valores atípicos; no/si posee sistematismos⁸; la desviación estimada es de $S_{\text{posteriori}}$; y la exactitud horizontal al 95 % de nivel de confianza es de Exactitud_r .”*

REFERENCIAS

Ver “Referencias” bibliográficas de la tesis.

A6.2 Anexos a la propuesta de aplicación

A6.2.1 Puntos de control

Los puntos de control deberán reunir las siguientes características:

- Su localización sobre la BDG y la FME se realizará empleando el mismo sistema de referencia que el de la BDG.
- Deberán ser perfectamente identificables tanto sobre la BDG como sobre la FME.
- En aquellos casos en los que la representación de los elementos geográficos se realice mediante simbología, se trabajará con las coordenadas del centro del símbolo sobre la BDG y su localización correspondiente en la FME.
- Los puntos deberán distribuirse de forma homogénea tanto temática como geográficamente. Esto es:
 - El número de puntos de control que representen un mismo tipo de elemento en la leyenda, deberá ser sensiblemente proporcional al número de elementos que aparezcan sobre la BDG.

⁸ En el caso de existir sistematismos, ha de indicarse si se han eliminado o no.

- La distribución espacial aleatoria para el control de una hoja, deberá ser de tal forma que, al menos, exista más de un 20 % de los puntos de control sobre cada cuadrante. La distancia mínima aconsejable de separación entre los puntos será de $(^{1/50}) \cdot D$, (⁹) siendo D la diagonal mayor que define la zona (figura A6.2).
- No obstante, en el caso de que el estudio posicional solamente se realizase sobre uno o varios tipos de elementos o en una zona determinada, se podrá afirmar que la BDG ha sido analizada de acuerdo al *ExHor* siempre y cuando quede indicado y delimitado en la metainformación.

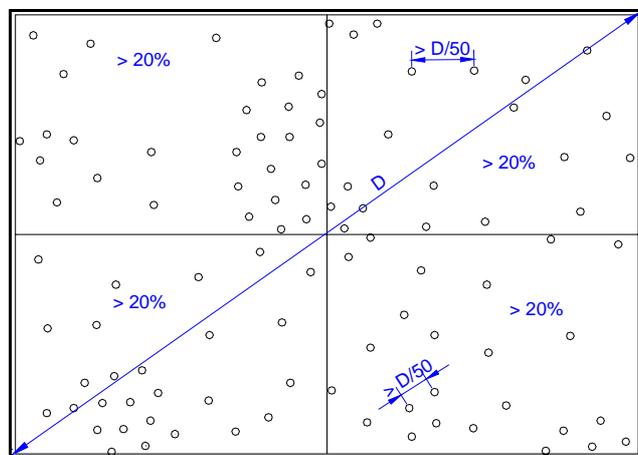


Figura A6.2: Ejemplo de distribución espacial homogénea de 100 puntos de control¹⁰.

- Cuando los puntos representen intersecciones de elementos lineales (p.e. cruces de carreteras, esquinas de edificios ...), se tomarán aquellos cuyo ángulo de intersección permita definir el punto con la menor ambigüedad posible. Para ello, los ángulos de intersección serán próximos a 90° y nunca inferiores a 25° o superiores a 155°.
- Cuando exista alguna duda sobre la identificación de un punto sobre alguna de las fuentes, preferiblemente deberá tomarse otro punto de similares características en la misma zona de trabajo. Cuando esto no sea posible, deberá desecharse el punto o asegurarse de que el elemento capturado en la

⁹ Basándose en las recomendaciones del FGDC (1998) que mencionan $^{1/10} \cdot D$ para más de 20 puntos de control, se ha estimado que, proporcionalmente, dicha distancia mínima debería reducirse a $^{1/50} \cdot D$ para el caso de más de 100 puntos de control.

¹⁰ Las zonas en las que aparece una mayor densidad (centro superior y esquina inferior izquierda), son debidas a una mayor concentración de elementos geográficos.

FME se corresponde al de la BDG empleando otras fuentes para contrastar dicha información.

En el caso de que la FME sea la toma de datos sobre el terreno, también deberán tenerse en cuenta los siguientes aspectos:

- Es recomendable que los puntos de control se ubiquen de forma que sean accesibles desde vías de comunicación próximas, para facilitar el acceso y optimizar la productividad en la toma de datos.
- En el caso de que el sistema empleado para la toma de datos en campo sea el GPS, preferiblemente los puntos de control deberán estar alejados de obstrucciones que dificulten su observación. Cuando esto no fuera posible, aquellos puntos que no sean accesibles serán capturados mediante el empleo de técnicas topográficas combinadas que aseguren su exactitud final (cinta métrica, brújula, excéntricas, estación total, ...). El método de procedimiento podría ser similar al indicado en el Anexo 2.
- Durante el estudio previo, se diseñarán los posibles itinerarios a seguir para optimizar el tiempo en la captura de datos.

A6.2.2 Normalidad

Existen numerosos procedimientos para el ajuste de una muestra de datos a una distribución normal. A continuación se exponen dos de ellos: el test de Kolmogorov-Smirnov y el de D'Agostino.

A6.2.1.1 Prueba de Kolmogorov-Smirnov

La prueba de Kolmogorov-Smirnov parte de dos hipótesis:

- H_0 : la muestra se asemeja a una distribución normal.
- H_1 : la muestra no sigue una distribución normal.

Su procedimiento de aplicación es el siguiente:

- En primer lugar se ordenan todos los valores obtenidos para cada una de las componentes, de menor a mayor, obteniendo así los valores de frecuencias acumuladas observadas entre 0 y 1.
- En siguiente lugar, se tipifican los valores según:

$$Ztip_i = \frac{e_i - \bar{e}}{S}$$

- De esta forma, se puede obtener en dicho punto (i) el valor de la función de distribución normal estándar acumulativa:

$$f(z;0,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

- Por último, se calculan las diferencias (D_i) (figura A6.3) entre el valor de la frecuencia acumulada observada [$S(x)$] y el de la función de distribución normal estándar [$F^*(x)$].

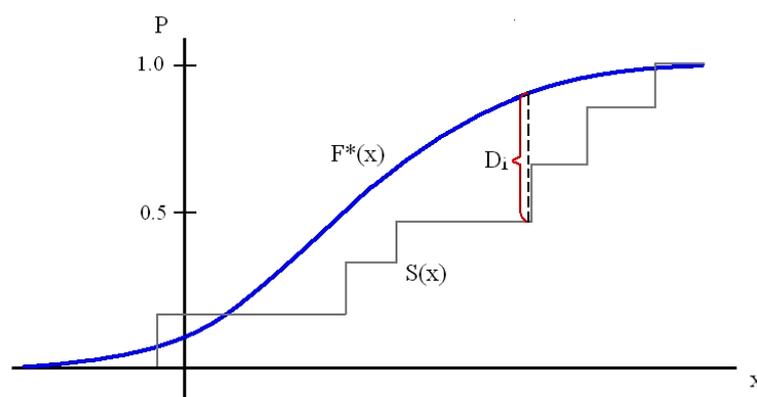


Figura A6.3: Distribución de frecuencias acumuladas observadas [$S(x)$] y de la función normal [$F^*(x)$]. Distancia máxima D_i entre ambas.

La prueba de Kolmogorov-Smirnov proporciona los valores críticos tabulados (tabla A6.3) que indican si la distribución analizada se puede asemejar a una normal con un 95 % de nivel de confianza.

Tabla A6.3: Valores límite de Kolmogorov-Smirnov.

Tamaño muestral (n)	Valor límite de Kolmogorov-Smirnov ($\alpha=0.05$)
15	0.338
20	0.294
25	0.27
30	0.24
35	0.23
> 35	$1.36/\sqrt{n}$

- Se aceptará la hipótesis nula H_0 , si la distancia máxima ($D_{i_{max}}$) para cada una de las componentes analizadas (X e Y) es inferior al estadístico proporcionado por Kolmogorov-Smirnov: $D_{i_{max}} < Kolmogorov$.

A6.2.1.2 Prueba de D'Agostino

Al igual que en el caso anterior, la prueba de D'Agostino parte de dos hipótesis:

- H_0 : la muestra se asemeja a una distribución normal.
- H_1 : la muestra no sigue una distribución normal.

Analiza las diferencias existentes entre la distribución de los errores (para cada una de las componentes X e Y), con la distribución normal. Procedimiento:

- Para cada una de las componentes, se ordenan los n valores obtenidos ($x_1, x_2, x_3, \dots, x_n$) de menor a mayor, obteniendo los valores ($x^*_1, x^*_2, x^*_3, \dots, x^*_n$).
- El estadístico de comprobación es:

$$D(w) = \frac{\sum_{k=1}^n x_k^* \left(k - \frac{n+1}{2} \right)}{\sqrt{n^3 S_{xx}}} \text{ siendo: } S_{xx} = \sum_{k=1}^n (x_k - \bar{x})^2$$

- La región crítica es el exterior (cerrado) del intervalo (k_1, k_2) donde k_1 y k_2 se obtienen para cada nivel de confianza de las tablas D'Agostino
- Se aceptará la hipótesis nula H_0 , si para cada una de las componentes analizadas (X e Y), el estadístico $D(w)$ pertenece al intervalo proporcionado por D'Agostino.

A6.2.3 Aleatoriedad

Se trata de una prueba con la que determinar si una muestra se distribuye de forma aleatoria. La verificación de esta hipótesis se realizará mediante un contraste de rachas de Wald–Wolfowitz. Para ello se parte de dos hipótesis:

- H_0 : la muestra es aleatoria simple.
 - H_1 : la muestra no es aleatoria.
- Para el cálculo de dicho test, en primer lugar se ha de calcular la mediana (m_e) de la muestra, ordenada de menor a mayor. Las observaciones z_i se clasifican en dos grupos (p.e.: "w" y "s"): si $z_i \leq m_e$ se asigna z_i al grupo "s", y si $z_i > m_e$ se asigna z_i al grupo "w". Con estos dos conjuntos, refundidos en el mismo orden en el que se obtuvieron, se denomina "racha" a un conjunto de datos de la misma muestra ("RW" o "Rs") que aparezcan contiguos. El estadístico del test es el número de rachas $R = RW + Rs$.
- La regla de decisión es que se acepta la hipótesis H_0 al nivel α si, en función del número de rachas obtenidas:

$$R \in C_0 = (r_{1-\alpha}; RW + Rs)$$

siendo $r_{1-\alpha}$ valores tabulados. En el caso de que $n, m \geq 20$, se puede tomar la distribución como asintótica a la normal, sin necesidad de disponer de las tablas específicas de Wald–Wolfowitz.

$$R \rightarrow N(\mu_R; \sigma_R^2)$$

En este caso, los valores de media y varianza vienen determinados por:

$$\mu_R = \frac{2 \cdot RW \cdot Rs}{RW + Rs} + 1 \quad \sigma^2 = \frac{2 \cdot RW \cdot Rs \cdot (2RW \cdot Rs - RW - Rs)}{(RW + Rs)^2 \cdot (RW + Rs - 1)}$$

$$z = \frac{R - \mu_R \pm 0.5}{\sigma}$$

siendo:

- $r_{1-\alpha}$: estadístico con el que comparar el número de rachas obtenidas.
- α : el nivel de significación
- RW, Rs : el número de elementos que componen cada uno de los subconjuntos ("w" y "s")
- R : el número de rachas.

- \pm : el signo de la ecuación será positivo (+) si $R < \mu_R$; y negativo (-) si $R \geq \mu_R$
- La regla de decisión es que se acepta H_0 , la muestra se distribuye de forma aleatoria, si $z \in (z_{\alpha/2}; z_{1-\alpha/2})$

A6.2.4 Valores atípicos

La detección y ponderación de valores atípicos que se encuentren en la muestra, se hará mediante el procedimiento siguiente:

- Detección de aquellos puntos cuyo error en alguna de sus componentes sea superior a $(2,5 \cdot \sigma_{priori})$. Si el número de atípicos es muy superior al 10% de la muestra, es posible que la σ_{priori} no sea correcta, por lo que habría que regresar a este punto y efectuar una nueva ponderación si así se determina en el estudio de la varianza ($S_{posteriori}$).
- Ponderación en base al método danés para cada componente¹¹ de dichos puntos:

$$W_i = e^{-0.05 \cdot e_{Xi}^2} \quad ; \quad e'_{Xi} = e_{Xi} \cdot W_i$$

siendo:

- W_i : el peso a aplicar en el punto i
- e_{Xi} : el valor del error en la componente X para el punto i
- e'_{Xi} : valor ponderado del error en la componente X para el punto i

A6.2.5 Sistematismos

El proceso para la detección de posibles errores sistemáticos es el siguiente:

- calcular el error medio en cada componente¹² (\bar{e}_x) para los n puntos que componen la muestra:

¹¹ Como ejemplo, aparece la componente X

$$\bar{e}_x = \frac{1}{n} \sum_{i=1}^n e_{xi}$$

- calcular la desviación en cada (S_x):

$$S_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (e_{xi} - \bar{e}_x)^2}$$

- calcular el test de existencia / ausencia de sistematismos empleando los límites de error aceptables para errores sistemáticos. Para ello se emplea el estadístico t_x :

$$t_x = \frac{\bar{e}_x \sqrt{n}}{S_x}$$

Si se verifica que $|t_x| \leq t_{n-1, \alpha}$, la muestra no contiene sistematismos,

siendo:

- $n-1$: los grados de libertad
- α : nivel de significación¹³
- t : medida teórica de una distribución *t-Student* de 2 colas $\alpha/2$

En el caso de que se detecte algún sistematismo en cualquiera de las componentes analizadas, deberá actuarse de la siguiente forma:

- Determinación de la posible dirección dominante empleando estimadores circulares. Para ello se obtendrá la dirección media y su intervalo al 95 % de nivel de confianza (ver apartado 3.3.2 y 3.3.2.4).
- Determinación del número aproximado de puntos de control necesario para poder corregir los errores sistemáticos, definido por la ecuación:

$$n = -150.02 \cdot \sigma_{priori}^2 + 764.81 \cdot \sigma_{priori} + 30.794$$

siendo:

- n : número mínimo de puntos de control recomendado
- σ_{priori} : valor estimado a priori como la desviación típica del mapa
- Si el número de puntos es admisible en función de los costes para la corrección de la BDG, se tomarían dichos puntos y se realizaría el ajuste por mm.cc.

¹² Como ejemplo, aparece la componente X

¹³ Para el nivel de significación se tendrá en cuenta Bonferroni (p.e. en un análisis de la exactitud horizontal, para cada componente $\alpha = (0.95)^{(1/4)} = 0.98726$).

- Si no resultase viable económicamente su corrección por mm.cc., se corregirán los errores medios (traslación) en cada componente y se verificará nuevamente la existencia / ausencia de sistematismos. Si persisten los sistematismos, esto se hará constar en el informe final.

A6.2.6 Desviación poblacional a posteriori

El proceso para el cálculo de la desviación poblacional a posteriori es el siguiente:

- Sobre la muestra principal, de al menos 100 puntos de control, se seleccionarán 1000 sub-muestras aleatorias de diferente tamaño (30, 40, 50, 60 ...).
- A cada una de las sub-muestras aleatorias, se le aplicará un contraste para la detección de errores aleatorios (contraste *Chi-Cuadrado*) con diferentes niveles de exigencia sensiblemente más estrictos que la desviación poblacional a priori (σ_{priori}). El contraste de errores aleatorios se efectuará mediante el cálculo, para cada componente¹⁴, del estadístico:

$$\chi_x^2 = \frac{S_x^2(n-1)}{\sigma_{\text{priori}}^2}$$

Si se verifica que $|\chi_x^2| \leq \chi_{n-1, \alpha}^2$, el mapa cumple con el estándar de calidad posicional en X, y donde:

- σ_{priori} : límite de error aceptable (desviación máxima) para la BDG.
- S : desviación típica
- $n-1$: grados de libertad
- α : nivel de significación¹⁵
- χ^2 : medida teórica de la distribución *Chi-Cuadrado*
- Contabilización del número de muestras que han superado el contraste e interpolación en el nomograma propuesto (figura 55.2).
- Cálculo y verificación de la desviación a posteriori ($S_{\text{posteriori}}$).

¹⁴ Como ejemplo, aparece la componente X.

¹⁵ Idéntico al aplicado para la detección de sistematismos.

- En el caso de que $S_{posteriori}$ sea diferente a σ_{priori} , debería retornarse al apartado de detección de valores atípicos para efectuar una nueva ponderación sobre aquellos puntos que superen ($2,5 \cdot S_{posteriori}$).
- Calcular la relación entre ambas desviaciones:

$$R(\sigma/S) = \frac{\sigma_{priori}}{S_{posteriori}}$$

A6.2.7 Exactitud posicional

Para la determinación de la exactitud posicional al 95 % de nivel de confianza, se seguirán los siguientes pasos:

- calcular el $RMSE_X$ para la componente X , y de forma análoga para componente Y , según:

$$RMSE_X = \sqrt{\frac{1}{n} \sum_{i=1}^n e_{xi}^2}$$

donde:

- $RMSE_X$: error medio cuadrático en la componente X
 - n : el número de puntos de la muestra
 - e_{xi} : el error en la componente X para el punto i
- calcular el $RMSE_r$ para la componente horizontal (XY) según:

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

- calcular el coeficiente de exactitud posicional horizontal al 95% de nivel de confianza dependiendo de dos posibilidades:

- a) en caso de que $RMSE_x = RMSE_y$, se utilizará la ecuación:

$$Exactitud_r = 1.7308 \cdot RMSE_r$$

- b) y en caso de que $RMSE_x \neq RMSE_y$, (la relación entre $RMSE_{min}/RMSE_{max}$) se utilizará la ecuación:

$$Exactitud_r = 1.22385 \cdot (RMSE_x + RMSE_y)$$

A6.3 Ejemplo de aplicación

Sobre los datos originales tomados en cartografía a escala 1/10.000 del MTA en la hoja 926 1-1 (Atkinson, 1998), se ha aplicado la propuesta de estándar presentada. En este apartado se exponen los cálculos realizados sobre dichos datos y los resultados que se ofrecerían tras aplicar el *ExHor*. De esta forma se pretende facilitar su comprensión y aplicación siguiendo las indicaciones del apartado A6.1.

1º. Selección de la muestra de control

Siguiendo las indicaciones del estándar, se han seleccionado y tomado sobre la FME y sobre la BDG un total de 104 puntos de control (ver Anexos 3, figuras A3.1 y A3.2; y Anexo 7, listados A7.1 y A7.23).

2º. Estudio previo: normalidad, aleatoriedad y valores atípicos

En siguiente lugar se han calculado los errores para cada punto (ver Anexo 7, listado A7.43) y se ha procedido al estudio previo. Los errores medios calculados para cada componente han sido de -0.65 y -1.86 m para X e Y respectivamente.

a) Normalidad

Se han verificado la distribución normal de los datos, en este caso empleando la prueba de Kolmogorov-Smirnov con los resultados de la tabla A6.4 (ver anexo A7.44 con los cálculos).

Tabla A6.4. Resultados de las pruebas de normalidad para ambas componentes.

Componente	Distancia máxima observada (m)	Valor límite de Kolmogorov – Smirnov para $n = 104$ puntos
X	0.1327	0.1334
Y	0.0698	

Como se puede apreciar, en ningún caso se supera el valor límite de Kolmogorov-Smirnov, por lo que la muestra se distribuye de forma normal.

b) Aleatoriedad

Se han verificado la aleatoriedad de la muestra mediante la prueba de rachas de Wald-Wolfowitz, obteniéndose los resultados de la tabla A6.5 (ver anexo A7.45 con los cálculos).

Tabla A6.5. Resultados de las pruebas de aleatoriedad para ambas componentes.

Componente	w	s	Número de Rachas	μ_R	σ_R	Estadístico z	Valor límite ($z_{\alpha/2}; z_{1-\alpha/2}$)
X	52	52	44	73	25,748	-0,330	(-1,96 ; 1,96)
Y	52	52	38	73	25,748	-0,563	

Se puede observar que no se supera el valor límite establecido al 95 %, por lo que la muestra se distribuye de forma aleatoria.

c) Valores atípicos

La σ_{priori} para una cartografía a escala 1/10.000 es, en principio, de 2,50 m (ver tabla A6.2).

Siguiendo dicho criterio, se han detectado 11 atípicos en la componente X y 30 en la componente Y (ver Anexo A7.46). En total hay 36 puntos que se consideran como valores atípicos en alguna de sus componentes. Esta cifra supone un 34,6 % del total de la muestra¹⁶ por lo que es muy probable que la σ_{priori} en función de la escala ha quedado subestimada.

Para obtener un valor de σ_{priori} más aproximado a la realidad, sin efectuar la ponderación mencionada se han eliminado los sistematismos (traslación en X e Y) y calculado la $S_{\text{posteriori}}$ para ver la relación existente entre ambas. Dicho estudio ofrece un valor de $S_{\text{posteriori}} = 5,00$ m, por lo que se ha procedido a una nueva detección de valores atípicos teniendo en cuenta dicho criterio.

¹⁶ Se trata de una cifra muy elevada teniendo en cuenta que lo normal ronda no más del 10% (Rencher, 1998).

Así, si $\sigma_{\text{priori}} = 5,00$ m, serán atípicos aquellos valores que se encuentren a más de $2,5 \times 5,00 = 12,50$ m (ver Anexo A7.47). En total hay 6 puntos que se consideran como valores atípicos (3 en X y otros tantos en Y). Esta cifra supone un 5,8 % del total de la muestra por lo que en este caso la σ_{priori} ha quedado mejor estimada.

3º. Detección y corrección de sistematismos

En siguiente lugar se han calculado los errores medios, desviaciones típicas y estadísticos de contraste para la detección de sistematismos aplicando Bonferroni (tabla A6.6).

Tabla A6.6. Cálculos efectuados para la detección de sistematismos ambas componentes.

Componente	Error medio	Desviación típica	Estadístico t	Valor límite <i>t-Student</i>
X	-0,44	3,40	1,31	2,54
Y	-2,02	4,71	4,38	

Se aprecia que la componente Y posee un error sistemático ($4,38 > 2,54$), por lo que se realiza un contraste mediante el estimador circular para la obtención de la posible dirección predominante de los errores, obteniéndose como valor medio para el total de la muestra: $209,44^g \pm 25,47^g$ (estimación realizada al 95 % de nivel de confianza).

El siguiente paso sería proceder a su corrección sobre la BDG. Para ello, se determina el número de puntos necesario para realizar un ajuste por mm.cc. en función de σ_{priori} . Así, empleando la ecuación propuesta por el estándar, se obtiene:

$$n = -150,02 \cdot \sigma_{\text{priori}}^2 + 764,81 \cdot \sigma_{\text{priori}} + 30,794$$

$$n = -150,02 \cdot 5^2 + 764,81 \cdot 5 + 30,794 = 104,344$$

Dado que la ecuación muestra un valor que sobrepasa el límite (algo más de 104 puntos), se hace muy aconsejable analizar los residuos del ajuste para ver si éste es fiable. Así, tras efectuar el ajuste por mm.cc., y analizar los resultados del mismo (tabla A6.7), se observa que los residuos para las traslaciones en X e Y no son admisibles, por lo que dichos sistematismos no se podrán cuantificar empleando este sistema. Por el contrario, los residuos que se refieren a los factores de escala y el giro son muy

bajos, por lo que los resultados en estos parámetros sí son validos: no hay sistematismos en factores de escala ni en giro.

Tabla A6.7. Resultados del ajuste por mm.cc.

Coefficiente	Valor	Residuo
a	0,999589	0,000228
b	-0,000244	0,000307
c	-0.000352	0,000228
d	0.998663	0,000307
X_0	1199,522	1302,700
Y_0	5771,906	1302,700
Sistematismo detectado		
	X	Y
Traslación	-	-
Factor de escala	0,999589	0,998663
Giro	-0,003087°	

Ello obliga a desechar el ajuste en cuanto a traslaciones se refiere, y efectuar la corrección de sistematismos por una traslación de coordenadas de igual magnitud y signo contrario a los errores medios detectados para cada componente. Así, se desplazarán todas las coordenadas de la BDG las magnitudes indicadas en la tabla A6.8.

Tabla A6.8. Corrección de errores sistemáticos por traslación.

Componente	Magnitud
X	+ 0,44
Y	+ 2,02

Como resultado se obtienen los datos de coordenadas corregidas de la BDG indicadas en el Anexo A7.48.

4°. Estudio de la varianza

a) Estimación de la desviación poblacional a posteriori

Sobre la muestra de 104 puntos se han seleccionado de forma aleatoria 1000 submuestras de 30, 40 y 50 puntos. A cada una de ellas se le ha efectuado un contraste

del tipo Chi-Cuadrado, verificando el número (en tanto por ciento) de muestras que cumplen con el estándar. Los resultados se muestran en la tabla A6.9.

Tabla A6.9: Porcentajes de cumplimiento con el estándar

n_m	$\sigma_{priori} = 4.0$
30	77,7 %
40	76,4 %
50	67,7 %

Interpolando los resultados sobre el nomograma de la figura 55.2, se calcula el valor estimado (ver figura A6.4) para la BDG. De la interpolación se obtiene que el valor estimado para $S_{posteriori}$ es de ~ 4,50 m. La relación con la σ_{priori} original es de:

$$R\left(\frac{\sigma}{S}\right) = \frac{\sigma_{priori}}{S_{posteriori}} = \frac{2,5}{4,5} = 0,55$$

De esta forma la BDG es 0,55 veces menos exacta que la estimada a priori en función de su escala.

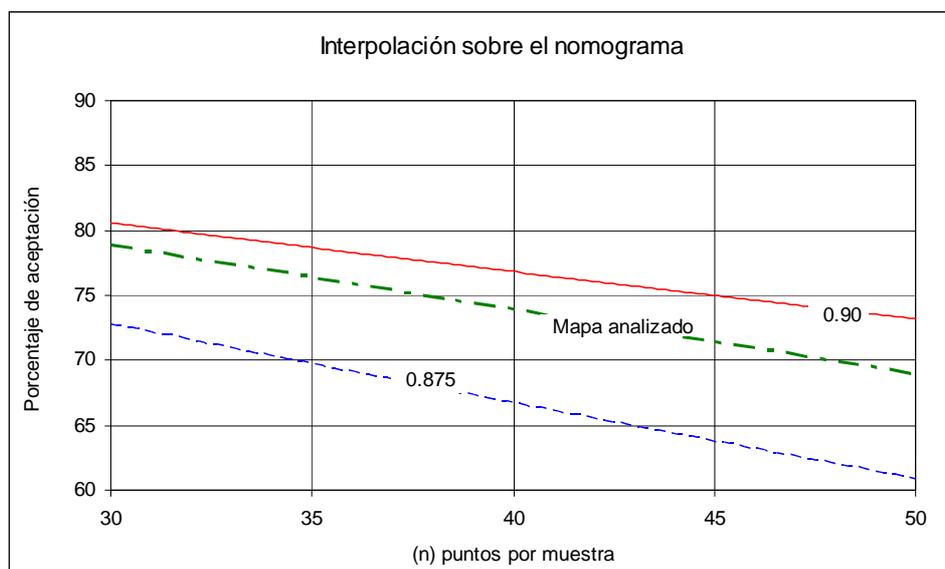


Figura A6.4: Detalle del nomograma e interpolación de los resultados de la simulación de muestras (BDG de la tabla A6.9).

b) Cálculo de la exactitud posicional al 95 % de nivel de confianza.

Se calcula el RMSE para cada componente, obteniendo los valores de:

$$RMSE_x = 3,38 \text{ m}$$

$$RMSE_y = 4,69 \text{ m}$$

Como $RMSE_x \neq RMSE_y$, se obtiene que:

$$Exactitud_r = 1.22385 \cdot (RMSE_x + RMSE_y)$$

$$Exactitud_r = 1.22385 \cdot (3,38 + 4,69) = 9,88 \text{ m}$$

5°. Informe final

Si se presenta el informe detallado, éste sería:

- Control de calidad posicional realizado el 12 de septiembre de 1997.
- Responsable del control: Alan D.J. Atkinson
- Origen de los datos de la FME: datos tomados sobre el terreno mediante DGPS en código.
- De acuerdo con el Estándar de Exactitud Horizontal en información puntual georreferenciada, esta base de datos geográfica posee las siguientes características posicionales en planimetría:
 - Los errores se distribuyen de forma normal y aleatoria.
 - El porcentaje de valores atípicos es del 5,8 %.
 - Se han detectado errores sistemáticos en la dirección de $209,44^{\circ} \pm 25,47^{\circ}$. Los errores sistemáticos detectados en traslación ($X_0 = -0,44 \text{ m}$, $Y_0 = -2,02 \text{ m}$), han sido corregidos y eliminados.
 - La desviación poblacional estimada es de 4,50 m. Dicho valor es 0,55 veces menos exacto que su valor estimado a priori (2,50 m).
 - La exactitud horizontal al 95 % de nivel de confianza es de 9,88m.
 - Otras observaciones:
 - Se han detectado errores en calidad temática (confusión, omisión y comisión).
 - La fecha del vuelo es de, al menos, 4 años antes al control de calidad, por lo que no aparece reflejada en la BDG la autovía Bailén – Granada. Ésta zona de influencia no se ha tenido en cuenta para el control de calidad.
 - De acuerdo con la desviación estimada a priori en función de la escala (2,50 m), se ha detectado un 34,6 % de valores atípicos, por lo que ésta ha sido reconsiderada en 4,50 m.

Por el contrario, si se presenta el informe resumido:

- Control de calidad posicional realizado el 1 de julio de 2004
- Responsable del control: Alan D.J. Atkinson
- Origen de los datos de la FME: datos tomados en el terreno mediante DGPS en código.
- De acuerdo con el estándar *ExHor*; existe un 5,8 % de valores atípicos; si posee sistematismos y éstos han sido corregidos; la desviación estimada es de 4,50 m; y la exactitud horizontal al 95 % de nivel de confianza es de 9,88 m.

ANEXO 7

Listados en formato digital

En el Anexo 7 se adjuntan aquellos listados y datos que, por su extensión, se aconsejaba su presentación en formato digital. Por ello, en este trabajo se adjunta un CD-ROM denominado “Tesis ATKINSON (Anexo 7)”.

Dependiendo de la configuración del ordenador en el que se visualice, el CD se ejecutará automáticamente¹. Para la correcta visualización del índice del anexo (ANEXO_7.htm), es necesario disponer de un navegador de internet que abra archivos HTM (p.e. Internet Explorer, Netscape ...).

Los listados con los documentos se encuentran en formato PDF. Para visualizar este tipo de archivos es necesario disponer de la versión 4 o superior de Adobe Acrobat Reader. Dicho programa se puede obtener de forma gratuita en la dirección:

<http://www.adobe.com/products/acrobat/readstep2.html>

Asimismo, se adjunta el archivo con el Manual de Usuario del programa CARTOCAL.



¹ De no iniciarse automáticamente, abrir el archivo “ANEXO_7.htm”

ANEXO 8

Planos

En el Anexo 8 se adjunta la cartografía empleada en el control de calidad:

- Hoja (926) 1-1 del ICA
- Hoja (926) 1-2 del ICA

ANEXO 9

Glosario

Por último, en el presente anexo se incluye una relación de términos que aparecen a lo largo del documento y que, por su especial significado, se ha hecho recomendable el realizar una breve descripción de los mismos. El glosario está basado en el propuesto por Ariza (2003)¹ con aquellas modificaciones, aclaraciones e inclusión de términos que se han considerado oportunas.

Acreditación: Reconocimiento formal, por una tercera parte autorizada, de la competencia técnica de una entidad (de certificación, de inspección, etc.) para la realización de una actividad determinada y perfectamente definida. Este mecanismo proporciona la confianza necesaria en los certificados, informes de inspección, actas de ensayo, certificados de calibración y validaciones medioambientales emitidos por las entidades de certificación de diferentes países.

Actualización: Operación para corregir o reemplazar datos en un conjunto de datos.

Aleatorio: Valor o suceso fortuito.

Aseguramiento de la calidad: Conjunto demostrable de acciones planificadas, implantadas dentro del sistema de la calidad para proporcionar la confianza adecuada de que una entidad cumplirá con los requisitos de la Calidad.

Atípico: Ver error grosero.

Atributo: Representación de una propiedad, cualidad o rasgo esencial de un objeto o entidad.

Auditoria de la calidad: Proceso sistemático, independiente y documentado para obtener evidencias y determinar si las actividades y los resultados relativos a la calidad cumplen las disposiciones establecidas, si éstas disposiciones están implantadas de forma efectiva y son adecuadas para alcanzar los objetivos.

Base de Datos Geográficos (BDG): Conjunto o colección identificable de datos geográficos.

Calidad: Conjunto de características de una entidad que le confieren su aptitud para satisfacer las necesidades explícitas e implícitas.

Capacidad: Se refiere a la probabilidad de que el proceso analizado satisfaga las tolerancias marcadas.

Certificación: Proceso llevado a cabo por una entidad reconocida como independiente de las partes interesadas, en la que se manifiesta por escrito que se dispone de la confianza

¹ Se ha optado por dicho glosario por ser el que más se ajusta al significado con el que se han empleado dichos términos en el texto.

adecuada de que un producto, proceso o servicio debidamente identificado, es conforme con una norma específica.

Cliente: Organización o persona que recibe un producto. Cliente interno se refiere a clientes de la propia organización.

Codificación: Procedimiento de conversión de la información en un formato especificado para su procesado.

Coherencia lógica: Ver consistencia lógica.

Compleción: Hecho de estar completo o pleno. Parámetro de calidad que describe la presencia y ausencia de instancias de entidades, instancias de relaciones e instancias de atributos.

Confianza: Elemento de metacalidad que describe la corrección de la información de calidad.

Conformidad: Cumplimiento de requisitos especificados.

Consistencia lógica: Grado de corrección o conformidad de un conjunto de datos geográficos con respecto a la estructura interna descrita por sus especificaciones.

Control de calidad: Las técnicas y actividades de carácter operativo utilizadas para cumplir los requisitos para la calidad.

Coste de la Mala Calidad: Cualquier coste en el que incurren la organización (reprocesado, desecho, control de la producción, etc.) y el cliente, cuando existen problemas de calidad tales que el producto no satisface los requisitos explícitos o implícitos.

Curva Característica Operativa: Es la función que, para un plan de muestreo determinado, relaciona la probabilidad de aceptación con el porcentaje de defectuoso del lote.

Datos geográficos: Tipo de información que concierne a fenómenos directa o indirectamente asociados con una ubicación relativa a la Tierra, tratable por un ordenador.

Datos terreno: Información geográfica capturada directamente sobre el terreno empleando técnicas directas (p.e.: GPS, estación total, visita a la zona ...).

Digitalizar: Convertir unos datos o informaciones analógicas en información codificada digitalmente para ser tratada por soportes físicos y lógicos.

Edición: procedimiento para: a) eliminación de errores en BDG, o para b) la codificación de los atributos de los datos geográficos.

Eficacia: Cumplimiento de los objetivos marcados.

Eficiencia: Relación entre los resultados y los recursos utilizados.

Elemento de calidad: Ítem de información que describe la calidad de un conjunto de datos geográficos.

Entidad: Fenómeno real localizado sobre la superficie terrestre que no puede ser dividido en fenómenos del mismo tipo. Cada clase de información definida por unas propiedades comunes.

Empresa: Término empleado de forma indistinta a organismo u organización (ver organización).

Error: Separación respecto a un valor teórico o real de una medida o atributo.

Error accidental: También denominado aleatorio dado que no responde a una ley fija. Se corresponden con defectos que se compensan cuando se realiza un número elevado de mediciones.

Error grosero: Equivocación que se puede cometer en una medida o transcripción de un dato.

Error sistemático: El debido a la existencia de una causa permanente que introduce un sesgo o desviación que puede eliminarse si se conoce.

Especificación Técnica: Documento que define las características requeridas en un producto (por ejemplo, niveles de calidad, uso específico, seguridad, dimensiones ...).

Estratificación: Proceso por el que un conjunto mayor se divide, según cierta propiedad, en subconjuntos menores para su análisis.

Exactitud: Grado de concordancia entre el resultado de un ensayo y el valor de referencia aceptado. La exactitud se refiere a la calidad de un resultado y se diferencia del concepto de precisión en que es relativa a la calidad de la operación más que al propio resultado.

Exactitud posicional: Grado de exactitud/fiabilidad del procedimiento de adquisición de los atributos posicionales del dato geográfico. Parámetro de calidad que describe la exactitud (X, Y, Z) de las posiciones geográficas en un conjunto de datos geográficos.

Exactitud temática (exactitud semántica): Grado de exactitud/fiabilidad del procedimiento de adquisición de los atributos temáticos del dato geográfico. Parámetro de la calidad que describe la exactitud de los aspectos temáticos/semánticos de un conjunto de datos geográficos.

Exactitud temporal: Parámetro de calidad que describe la exactitud de los aspectos temporales de los datos geográficos.

Fiabilidad: Elementos de la metacalidad que describen la probabilidad de que una muestra de un conjunto de datos geográficos usada para la evaluación de la calidad, sea representativa del conjunto total de los datos.

Fuente de Mayor Exactitud (FME): Se trata de una fuente de datos que ha de ser, preferiblemente, al menos tres veces más exacta que la BDG a controlar. Así, podrá ser bien un mapa de mayor exactitud (cartografía a escala significativamente superior y/o con una mayor calidad posicional contrastada previamente), o bien datos tomados sobre el terreno empleando algún sistema de captura que garantice dicha exactitud superior.

Genealogía: Ver linaje.

Geomática: El autor emplea el término como la ciencia o el ámbito de las ciencias que estudian y analizan la Tierra y los elementos geográficos desde una perspectiva ingenieril. En ella se agrupan conocimientos en: Cartografía, Geodesia, Fotogrametría, Teledetección, GPS, Topografía ...

Gestión de la Calidad: Actividades coordinadas para dirigir y controlar una organización en lo relativo a la calidad.

- Incertidumbre:** Falta de certeza, nivel de error.
- Indicador de calidad:** Para un conjunto de datos geográficos dado, es el conjunto de medidas que describe un parámetro de calidad.
- Información geográfica:** Información que concierne a fenómenos directa o indirectamente asociados con una ubicación relativa a la Tierra.
- Inspección:** Actividades tales como medir, examinar, ensayar o comparar con un patrón una o varias características y confrontar los resultados con requisitos especificados para así establecer si se logra la conformidad para cada una de estas características.
- Linaje:** Registro de las características, de las fuentes de datos, de las operaciones, etc. que han creado el producto final, con el fin de poder reconstruir la historia de un conjunto de datos geográficos.
- Lote:** Conjunto de productos obtenidos de un proceso homogéneo.
- Magnitud:** Atributo de un fenómeno u objeto que puede ser distinguido cualitativamente y determinado cuantitativamente.
- Manual de calidad:** Documento que describe el sistema de calidad de la organización y en el que se define la política de calidad, la estructura organizativa y las políticas sobre aspectos específicos del sistema de calidad.
- Manual de procedimientos:** Documento complementario al manual de calidad que describe las actividades que se realizan en la organización, especialmente aquellas que afectan al sistema de calidad.
- Mapa:** En este texto se emplea con el mismo significado que BDG (ver Base de Datos Geográficos).
- Medición:** Operación cuya finalidad es determinar el valor de una magnitud.
- Mejora de la calidad:** Parte de la gestión de la calidad orientada a aumentar la capacidad de cumplir con los requisitos de la calidad
- Metacalidad:** Datos que informan sobre los datos de calidad de un conjunto de datos geográficos.
- Metadatos:** Datos sobre uno o varios conjuntos de datos geográficos.
- Modelo Digital de Elevaciones (MDE, DEM):** Es un MDT en el que el valor de Z se refiere a las elevaciones o cotas topográficas del terreno.
- Modelo Digital del Terreno (MDT, DTM):** Modelo cuantitativo digital que para cada posición (X,Y) del espacio almacena un valor Z correspondiente a cualquier variable de interés (por ejemplo, elevaciones, temperaturas, etc.).
- Modelo ráster:** Modelo de datos para la codificación de información geográfica que utiliza como elemento base una rejilla de celdas de forma regular, generalmente cuadrada.
- Modelo vectorial:** Modelo de datos para la codificación de la información geográfica que utiliza como elemento base los segmentos orientados o vectores.
- Mylar:** Material de tipo poliéster muy utilizado como soporte de las minutas y separaciones de color en la producción cartográfica basada en procesos analógicos y fotomecánicos.

Nivel de calidad Aceptable (NCA, AQL): En la aceptación, es el máximo número de defectuosos que se admite como promedio de los porcentajes de defectuoso que aparecen en los lotes.

No conformidad: Fallo en el cumplimiento de un requisito especificado en una norma o en los procedimientos de calidad.

Norma: Documento público, no obligatorio, consensuado entre las partes interesadas y aprobada por un organismo de Normalización. Recoge conocimientos y recomendaciones científicotécnicas surgidas de la experiencia, orientado a la simplificación unificación y clasificación de los procesos y productos dentro de un ámbito territorial determinado.

Norma Europea: Especificación técnica que se desarrolla en base al consenso de los organismos de normalización europea, miembros del CEN, CENELEC, ó ETSI que adopta el estatus de norma nacional por la publicación del texto o su ratificación en un periodo no superior a seis meses.

Normalización: Conjunto de actividades consistentes en la elaboración, difusión y aplicación de normas como forma de establecer soluciones productivas a situaciones repetitivas.

Objeto: Fenómeno individual que existe en el mundo real.

Organización: Compañía, corporación, firma, empresa o institución, o parte de las mismas constituidas como sociedad o no, pública o privada, que tiene sus propias funciones y su propia administración.

Outlier: Ver error grosero.

Parámetro de calidad: Elemento de calidad cuantificable que describe la concordancia de un conjunto de datos geográficos con su terreno nominal.

Plan de calidad: Documento que enuncia las prácticas, los recursos y la secuencia de las actividades relacionados con la calidad, que son específicas a un producto, un proyecto o un contrato en particular.

Plan de muestreo: Conjunto de características que definen una aceptación por muestreos: tamaño del lote, calidad considerada, tamaño de la muestra, número de aceptación, número de rechazo, etc.

Planificación de la calidad: Actividades que establecen los objetivos de calidad, así como los requisitos para la aplicación de los elementos del sistema de calidad.

Política de calidad: Directrices y objetivos generales de una organización con respecto a la calidad, expresadas de manera formal, y que es aprobada y revisada periódicamente por la alta dirección de la organización.

Precisión: Calidad asociada al grado de refinamiento de los instrumentos de medida al que determina el grado de uniformidad o coherencia de sucesivas mediciones del mismo fenómeno. La precisión también se refiere al número de cifras decimales con las que se aproxima una medición.

Procedimiento: Forma especificada de desarrollar una actividad o un proceso.

Proceso: Conjunto de recursos y actividades interrelacionados que transforma entradas en salidas o productos.

Producto: Resultado de las actividades o procesos.

Punto de Riesgo Compartido (PRC): En una curva característica operativa, es el punto en el que los riesgos del productor y usuario valen lo mismo (50%).

Racha: En una serie, el conjunto de elementos que comparten una característica.

Registro: Documento que suministra evidencia objetiva de las actividades efectuadas o de los resultados alcanzados.

Relación: Asociación de dos o más entidades.

Repetitividad: Grado de concordancia entre los resultados de mediciones sucesivas sobre un mismo mensurando.

Representatividad: Índice que expresa si la información contenida en una BDG es efectivamente suficiente para describir una cierta área geográfica (equivalente a la compleción).

Reproducibilidad: Grado de concordancia entre los resultados de las mediciones de un mismo mensurando realizadas bajo diferentes condiciones de medida.

Requisito: Necesidad o expectativa establecida, implícita u obligatoria.

Resolución: Dimensión del más pequeño elemento representado en un mapa. Dimensión de la celda ráster, el lapso temporal en el paso de un satélite de observación de la Tierra.

Riesgo del Productor: En aceptación, que se le rechace un trabajo bien hecho.

Riesgo del Usuario: En aceptación, que acepte por bueno un trabajo cuya calidad es inferior a la establecida.

Seudoaleatorio: Números con apariencia aleatoria generados mediante una computadora y algoritmo.

Simulación: Programa de ordenador que incorpora un modelo de un sistema tal que permite estudiarlo y obtener conclusiones para la toma de decisiones.

Sistema de Gestión de la Calidad (SGC): Parte del sistema de gestión de la organización enfocada al logro de resultados en relación con los objetivos de la calidad.

Sistema de Información Geográfica (SIG, GIS): Sistema de soporte a la decisión integrado por datos georreferenciados, soportes físicos, lógicos y organizativos y demás recursos materiales, humanos y de conocimientos que mantienen las informaciones actualizadas y aptas para su uso.

Sistema de Posicionamiento Global (GPS): Se trata de un sistema de posicionamiento y navegación basado en 24 satélites, que proporcionan posiciones en tres dimensiones, velocidad y tiempo, las 24 horas del día, en cualquier parte del mundo y en todas las condiciones climáticas.

Sistema: Conjunto de elementos mutuamente relacionados o que interactúan.

Sistematismo: Presencia de una perturbación constante en un conjunto de datos.

Tableta digitalizadora: Dispositivo, generalmente electromagnético, para la captura digital vectorial de coordenadas desde soportes analógicos.

Trazabilidad: Capacidad para rastrear la historia, la aplicación o la localización de una entidad de interés mediante identificaciones registradas.

Valor atípico: Ver error grosero.

Vectorización: Procedimiento para la conversión de informaciones en modelo ráster a informaciones constituidas por líneas y áreas (modelo vectorial).

Verificación: Confirmación mediante examen y aporte de evidencia objetiva de que se han cumplido los requisitos especificados.