

UNIVERSIDAD DE SALAMANCA



Departamento de Informática y Automática

Aportaciones al Estudio de la Anisotropía y Modelado Espacial de Información

Tesis Doctoral

M^a Dolores Muñoz Vicente

Salamanca, Julio de 2008

UNIVERSIDAD DE SALAMANCA



Aportaciones al Estudio de la Anisotropía y Modelado Espacial de Información

MEMORIA QUE PRESENTA

M. Dolores Muñoz Vicente

PARA OPTAR AL GRADO DE DOCTOR EN INFORMÁTICA

Salamanca, Julio de 2008

DIRECTOR
Dr. Andrés Molina Aguilar

Departamento de Informática y Automática

Tesis Doctoral parcialmente subvencionada por la Comisión Interministerial de Ciencia y Tecnología con el proyecto “Utilización de modelos digitales para la evaluación de la Energía Solar e integración en Sistemas de Información geográfica”



CICYT

ENE2004-07816-C03-02

La memoria titulada “Aportaciones al Estudio de la Anisotropía y Modelado Espacial de Información”, que presenta D.^a M. Dolores Muñoz Vicente para optar al grado de doctor, ha sido realizada dentro del programa de doctorado “Informática y Automática” del Departamento de Informática y Automática de la Universidad de Salamanca, bajo la dirección del doctor D. Andrés Molina Aguilar de la Universidad de Jaén.

Salamanca, Julio de 2008

La doctoranda

Fdo: M. Dolores Muñoz Vicente

El Director

Fdo: Andrés Molina Aguilar

Resumen

Cuando la visualización y observación del proceso analizado no describen de forma adecuada su comportamiento espacial, es necesario el diseño de un modelo matemático que lo explique y sintetice. Habitualmente, los fenómenos espaciales son el resultado de la suma de dos tipos de efectos: globales y locales. Los efectos globales aparecen cuando existe en el espacio variaciones significativas del valor medio del proceso. Los efectos locales se manifiestan cuando el valor medio en un punto es influenciado por el valor medio en los puntos vecinos.

El diseño de modelos espaciales sólo es posible bajo asunción de homogeneidad en los efectos globales y/o locales. Si no se asume ningún tipo de estacionalidad, sería extremadamente complicado modelizar el fenómeno debido a la gran cantidad de parámetros necesarios para la formalización matemática de su comportamiento. Además, es necesario determinar si el fenómeno presenta tendencia anisotrópica, ya que este hecho invalidaría la asunción de modelos definidos para funciones isotrópicas, y haría necesario utilizar transformaciones espaciales que permitan obtener funciones anisotrópicas a partir de modelos isotrópicos. Por este motivo, el modelado de procesos espaciales requiere tanto un análisis de la presencia de estacionalidad en las variaciones globales y locales como un estudio direccional para valorar la existencia de una componente anisotrópica.

Son muchas y muy variadas las herramientas utilizadas para el análisis de estacionalidad: triangulación de Delaunay, medias móviles espaciales, k -funciones y funciones de covarianza (covariogramas), variogramas, estadísticos de autocorrelación espacial (*Moran's I*, *Geary's C*, *Getis's G*), autocorrelogramas, tendencia espacial....

Las técnicas para el estudio de la anisotropía demuestran que, por el contrario, son escasas, y se limitan a la detección de efectos direccionales mediante la inspección de variogramas experimentales en diferentes direcciones. El examen de estos variogramas "a ojo" sólo proporciona una comprobación informal de la existencia de anisotropía y la determinación aproximada de la

dirección en que ésta se manifiesta.

Como consecuencia de esto, en esta memoria se presenta:

- Un estudio sobre el estado del arte de la anisotropía, especialmente en el aspecto relacionado con el modelado espacial de datos.
- El análisis de la problemática asociada a la determinación de la existencia y cuantificación de la componente direccional en patrones.
- La adaptación de los estadísticos circulares existentes para el análisis del comportamiento direccional de fenómenos naturales dentro del proceso general de análisis de las propiedades de primer orden (homogeneidad) y segundo orden (autocorrelación y anisotropía) en patrones.
- Un test para la existencia de direccionamiento en patrones que presentan dependencia espacial asumiendo normalidad de la población.
- Determinación de la dirección dominante en patrones, una vez establecida la existencia de anisotropía.
- Un método para el diseño de óvalos generadores en los que la superficie se distribuye de forma Normal alrededor de su dirección dominante.

Estos métodos han sido particularizados y adaptados a imágenes digitales, y más concretamente a imágenes satélites.

El contenido descrito se distribuye en los siguientes capítulos:

Capítulo I: Anisotropía. Estado del Arte

En este capítulo definiremos el concepto de anisotropía desde distintos puntos de vista, considerándola como una característica visual o como el comportamiento de ciertos minerales. Dentro de la anisotropía espacial, analizamos dos tipos: la anisotropía vertical y la horizontal, definiéndola de una manera más formal a partir de un modelo estadístico. El capítulo finaliza con el estudio del tratamiento de la anisotropía por los Sistemas de Información Geográficos más conocidos en este momento.

Capítulo II: Estadística Circular

Existen fenómenos que poseen características que cambian en función de la dirección en que son observadas. Estas variables requieren un tratamiento estadístico particular y diferenciado al de las variables lineales mediante el uso de estadísticos descriptivos, métodos de exploración y distribuciones circulares propias. En este capítulo se describen los fundamentos de la estadística circular y las herramientas de análisis estadístico direccional que utilizaremos en los capítulos posteriores.

Capítulo III: Test para determinar la existencia de direccionamiento

Las técnicas usadas para el análisis de anisotropía, se limitan a la visualización de variogramas de diferentes direcciones. En este capítulo se diseña un método basado en estadísticos circulares de segundo orden que nos permite determinar la existencia de anisotropía a partir de imágenes digitales cuantificando las direcciones en las que aparecen.

Capítulo IV: Operador Buffer

Presentamos un método para generar óvalos en los que la superficie se propaga según una distribución Normal alrededor de su dirección dominante. Para ello hemos utilizado la función de distribución de Von Misses modificada, lo que nos permitirá identificar el concepto de probabilidad con el de superficie ocupada por la propagación alrededor de un punto.

Capítulo V: Un caso de estudio: El SIG como entorno de trabajo para extracción de conocimiento

El propósito de este capítulo es aplicar las técnicas diseñadas por la doctoranda y expuestas en los capítulos precedentes a un proceso de extracción de conocimiento espacial y de técnicas de minería de datos espacial sobre colecciones heterogéneas de datos. Para ello, describiremos el desarrollo de un modelo gráfico (mapa) de coste en tiempo para el desplazamiento de un móvil, utilizando proceso de extracción de conocimiento sobre la información contenida en nuestra base de datos y sobre los datos reales (mapas temáticos) almacenados en el SIG *Idris32*.

Abstract

When the visualization and observation of the analyzed process does not adequately describe his behaviour in space, it is necessary to design a mathematical model that explains and summarize it. Usually, the space phenomena are the result of the sum of two types of effects: global and local. The overall effects appear when there are significant variations in the space of the average value of the process. The local effects are shown when the average value at one point is influenced by the average value at neighbouring points.

The design of spatial models is only possible under assumption of homogeneity in the global and/or local effects. If no kind of stationarity is assumed, it would be extremely complex to model the phenomenon due to the high number of necessary parameters for the mathematical formalization of its behaviour. Moreover, it is necessary to determine if the phenomenon presents anisotropic tendencies, since this fact would invalidate the employment of models defined for isotropic functions, and would make necessary to use space transformations to get anisotropic functions from isotropic models. For this reason, the spatial processes modelling requires not only an analysis of stationarity in global and local variations, but also a directional study to determine the existence of an anisotropic component.

Many and varied tools are used for analysing stationarity: Delaunay triangulation, spatial movable average, k-functions and covariance functions, variograms, statistics of spatial autocorrelation (*Moran's I*, *Geary's C*, *Getis's G*), autocorrelograms, spatial trend....

Techniques used for the analysis of anisotropy are not many, and they just detect directional effects by inspecting experimental variograms in different directions. The examination of these variograms "by rough estimate" only provides an informal verification of the existence of anisotropy and the approximated determination of the direction in which this one is pronounced.

As a result of this, in this report appears:

- The analysis of the problems associated with determining the existence and quantifying the component in directional patterns.
- A study on the state-of-the-art of anisotropy, especially in the aspect related to the data spatial modelling.
- The analysis of the problems associated with determining the existence and quantifying of the component in directional patterns.
- The adaptation of the existing circulars statistics to analyze the directional behaviour of natural phenomena within the overall process of analysing the first order properties (homogeneity) and second order ones (autocorrelation and anisotropy) in patterns.
- A test for the existence of addressing in patterns that present spatial dependence assuming normality of the population.
- A method for the design of generating ovals in which the surface is distributed as Normal form around its dominant direction.

These methods have been distinguished and adapted to digital images, and more specifically to satellite images.

The described content is distributed in the following chapters:

Chapter I: Anisotropy. State-of-the-Art

In this chapter we define the concept of anisotropy from different points of view, considering it as a visual characteristic or as the behaviour of certain minerals. Within the space anisotropy, we analyze two types: vertical anisotropy and horizontal, defining it in a more formal way based on a statistical model. The chapter ends with the study of the treatment of anisotropy by geographic information systems best known at this time.

Chapter II: Circular Statistic

There are phenomena that have characteristics that change depending on the direction in which they are observed. These variables require a statistical

treatment particular and differentiated to the one from the linear variables by using descriptive statistical, methods for exploration and own distributions circulars. This chapter describes the basis for circular statistical and the tools of directional statistical analysis we will use in later chapters.

Chapter III: Test to determine the address existence.

The techniques used for the anisotropy analysis, are limited to the visualization of variograms of different directions. In this chapter we design a method based on Second-Order Statistical circular that allows us to determine the existence of anisotropy from digital images quantifying the directions in which they appear.

Chapter IV: Buffer Operator

We present a method to generate ovals in which the surface spreads as a Normal distribution around its dominant direction. We have used the distribution function of Von Misses modified, which will allow us to identify the concept of probability with the one of surface occupied by the spread around a point.

Chapter V: A case study: GIS as a work environment for extracting knowledge

The purpose of this chapter is to apply the techniques designed by the doctoral student and exposed in the preceding chapters to a process of extraction of space knowledge and techniques of mining of spatial data on heterogeneous collections of data. To do so, we will describe the development of a graphic model (map) of cost in time for the displacement, using extraction process of knowledge on the information contained in our database and on real data (thematic maps) stored in GIS *Idrisi32*.

Índice de Contenidos

Capítulo I Anisotropía: Estado del Arte

1.1. Anisotropía	2
1.1.1. Anisotropía vertical	2
1.1.2. Anisotropía horizontal.....	4
1.2. La anisotropía como característica de los minerales.....	7
1.3. La Anisotropía de fisuración.....	8
1.3.1. Una descripción más formal.....	10
1.4. La anisotropía en los Sistemas de Información Geográfica	11
1.5. El Sistema de Información Geográfica <i>Idrisi</i>	18

Capítulo II Estadística Circular

2.1. Estadística direccional	24
2.2. Vector media	26
2.3. Distribución de Von Misses.....	28
2.4. Test de Rayleigh	32
2.5. Test de Cox.....	33

Capítulo III Test para determinar la existencia de direccionamiento

3.1. Test para determinar la existencia de direccionamiento	36
3.2. Test de direccionamiento	37
3.3. Elipse estándar.....	41
3.4. Elipse de confianza	44
3.5. Un ejemplo	49
3.6. Análisis global de homogeneidad	52
3.7. Análisis de dependencia espacial.....	53
3.8. Análisis de direccionamiento.....	60

Capítulo IV Operador Buffer

4.1. Operador Buffer.....	72
4.1.1. Definición.....	72
4.1.2. Creación de áreas de influencia de datos puntuales	72
4.1.3. Creación de áreas de influencia de datos lineales.....	74
4.1.4. Creación de áreas de influencia de datos poligonales.....	81
4.2. El operador buffer en SIG comerciales	85
4.2.1. <i>Arcinfo</i>	85
4.2.2. <i>ArcView</i>	86
4.2.3. <i>Idrisi</i>	89

4.2.4. Grass	89
4.3. Suma de Minkowski.....	91
4.4. Modelado del polígono generador.....	92
4.5. Visualización de la expansión	96
4.5.1. Caso 1	98
4.5.2. Caso 2	99
4.5.3. Caso 3	101
4.6. Un ejemplo ilustrativo	102
4.6.1. Determinación de las direcciones dominantes.....	102

Capítulo V Un caso de estudio: El SIG como entorno de trabajo para extracción de conocimiento

5.1. Extracción de conocimiento y minería de datos.....	110
5.2. Métodos de Minería de Datos Espaciales.....	112
5.3. Modelado de la función de coste.....	114
5.4. Información disponible y preproceso de datos.....	116
5.5. Algoritmo DM.....	118
5.6. KDSD: Asignación de fricciones.....	120
5.7. Cálculo de coste	121

Capítulo VI Conclusiones

6.1. Resultados obtenidos	130
6.2. Publicaciones asociadas al trabajo desarrollado en esta memoria	132
6.3. Trabajos futuros.....	133

Bibliografía	135
---------------------------	-----

Índice de figuras

Figura 1.1.	Anisotropía vertical número.....	3
Figura 1.2.	Efecto visual debido a la anisotropía.....	3
Figura 1.3.	“La Calle” obra de Veermer.....	5
Figura 1.4.	Leonardo da Vinci, “La Virgen de las Rocas”. Arte del Renacimiento (composición cerrada)	6
Figura 1.5.	Izquierda: Bernini, “Apolo y Dafne” (siglo XVII, estilo Barroco). Derecha: Raebun, “El reverendo Robert Walker patinando”, siglo XIX.....	7
Figura 1.6.	Ejemplo de overlay en Sistema de Análisis Espacial.....	14
Figura 1.7.	Superposición de diferentes capas de información	15
Figura 1.8.	Componentes de un SIG.....	16
Figura 1.9.	Representación de información Raster y Vectorial.....	17
Figura 1.10.	Función $f = \cos^k \alpha$	21
Figura 1.11.	Función $f = \frac{1}{\cos^k \alpha}$	21
Figura 1.12.	Comportamiento de las funciones <i>Varcost</i> y <i>Disperse</i>	22
Figura 2.1.	Representación de un vector unitario mediante un gráfico circular	24
Figura 2.2.	Representación de datos circulares como puntos sobre el círculo unidad.....	24
Figura 2.3.	Vector media	26
Figura 2.4.	Distribución circular	29
Figura 2.5.	Distribución de Von Misses para distintos valores de k	30
Figura 2.6.	Comparación de la Distribución de Von Misses frente a la distribución circular para distintos valores de k	31
Figura 3.1.	Pixelado de 17x17.....	37
Figura 3.2.	Dirección en la que se produce la anisotropía.....	38
Figura 3.3.	Elipse estándar obtenida a partir de los parámetros de la Tabla 3.3.....	43
Figura 3.4.	Elipse de confianza con un coeficiente $Q = 1 - \alpha = 95\%$	48
Figura 3.5.	Elipse de confianza con un coeficiente $\alpha = 0.1$ y $\alpha = 0.05$	48
Figura 3.6.	(a) Vista del Estrecho de Gibraltar desde el vuelo STS-58 (b) Ventana de 37x 37píxeles	49
Figura 3.7.	Límites superior e inferior para χ^2_{120}	52
Figura 3.8.	Límites superior e inferior del intervalo de confianza.....	59
Figura 3.9.	Las celdas marcadas con el símbolo = no presentan homogeneidad a nivel global pero no dependencia espacial	60
Figura 3.10.	Las celdas con valores de T^2 inferiores a $T^2(0.05)=24.75$ tienen un comportamiento isotrópico	61
Figura 3.11.	Comportamiento de las celdas de W	62
Figura 3.12.	Dirección ψ en la que se muestra la anisotropía para cada una de	

las celdas de W calculada con $d = 5$ valores	64
Figura 3.13. Dirección ψ en la que se muestra la anisotropía superpuesta a la imagen de W	65
Figura 4.1. Áreas de influencia de datos puntuales con distancia fija	72
Figura 4.2. Áreas de influencia de datos puntuales con atributos de capa	73
Figura 4.3. Cálculo de la intersección de áreas de influencia	73
Figura 4.4. Cálculo de la disolución de áreas de influencia de datos puntuales	74
Figura 4.5. Cálculo de la disolución de áreas de influencia de datos puntuales con atributos de capa	74
Figura 4.6. Creación de áreas de influencia de datos lineales	75
Figura 4.7. Segmento inicial	75
Figura 4.8. Cálculo de las coordenadas en el eje de abscisas y ordenadas	76
Figura 4.9. Representación de las coordenadas de los puntos extremos	76
Figura 4.10. Segundo segmento	77
Figura 4.11. Cálculo de las coordenadas del segundo segmento	77
Figura 4.12. Ultimo segmento	78
Figura 4.13. Creación de los extremos del área de influencia	78
Figura 4.14. Remate de los extremos del área de influencia	79
Figura 4.15. Cálculo del solapamiento entre varias áreas	79
Figura 4.16. Intersección de las áreas de influencia	80
Figura 4.17. Disolución de las áreas de influencia	80
Figura 4.18. Disolución de áreas de influencia de datos poligonales con atributos de capa ..	81
Figura 4.19. Polígono original, área de influencia exterior, área de influencia interior	81
Figura 4.20. Campus universitario	82
Figura 4.21. Creación de áreas de influencia poligonales	83
Figura 4.22. Solapamiento de áreas de influencia	84
Figura 4.23. Resultado de la generación de áreas de áreas de influencia	84
Figura 4.24. Creación del buffer con <i>ArcInfo</i>	86
Figura 4.25. Pantalla 1 para la creación de buffer con <i>ArcView</i>	87
Figura 4.26. Pantalla 2 para la creación de buffer con <i>ArcView</i>	87
Figura 4.27. Pantalla 3 para la creación de buffer con <i>ArcView</i>	89
Figura 4.28. Creación del buffer con <i>Idrisi</i>	91
Figura 4.29. Ejemplo de suma de Minkowski tomando como objeto generador un círculo (a) y un óvalo (b)	93
Figura 4.30. Comportamiento direccional de un fluido que emana desde el punto P	93
Figura 4.31. Histograma de valores	95
Figura 4.32. Representación de la función de Von Misses con el parámetro de concentración $k = 1.5$ y vector media $\bar{\Phi} = 59^\circ$	97

Figura 4.33.	Expansión del punto P	98
Figura 4.34.	Generación de áreas de influencia.....	98
Figura 4.35.	La expansión E_1 del punto P es el resultado de las expansiones de las celda1 $E_{1,1}$ y de la celda2 $E_{1,2}$: $Z1+Z2+Z3$	98
Figura 4.36.	Diferentes ejemplos de expansión en las celdas en función de los cambios en la velocidad y dirección de las fuerzas.....	101
Figura 4.37.	Imagen de control remoto de las islas de Ibiza y Formentera (Islas Baleares) tomadas desde el vuelo STS-78 de la NASA (a) y (b).....	102
Figura 4.38.	Comportamiento de las celdas que pertenecen a la ventana señalada en la Fig.4.37.(b)	103
Figura 4.39.	Direccionamiento y buffer de las celdas pertenecientes a una subventana (251x251) de la ventana señalada en la Fig.4.37.(a)	105
Figura 4.40.	Zona de influencia calculada desde el punto x de coordenadas (245, 245) en la Fig.4.37	107
Figura 4.41.	Estudio de las corrientes marinas en el Estrecho de Gibraltar	108
Figura 5.1.	Diagrama del proceso DM.....	124
Figura 5.2.	Diagrama del proceso KDSD completo.....	125
Figura 5.3	Modelo digital de elevación del terreno (Map9). El símbolo “+” representa la posición del punto_inicial.....	126
Figura 5.4.	Orientación del terreno (Map27)	126
Figura 5.5.	Pendiente del terreno (Map28)	126
Figura 5.6.	Coste total final (Map32)	127
Figura 5.7.	Coste total final (isotrópico)	127

Índice de Tablas

Tabla 2.1.	Valores de $z(\alpha)$ para las variables α y n	33
Tabla 3.1.	Matriz que almacena los valores de obtenidos a partir de la Fig.3.1	40
Tabla 3.2.	Muestras obtenidas N_1, \dots, N_6 a partir de $Z_{17,17}$ de la Fig.3.1.....	41
Tabla 3.3.	Vector media $\{\bar{m}_i, i = 1..6\}$ calculado a partir de las medias N_1, \dots, N_6	41
Tabla 3.4.	Parámetros de la elipse de confianza calculada a partir de los vectores almacenados en la Tabla 3.3.....	43
Tabla 3.5.	Valores de la luz reflejada en la escala de grises 0-255 siendo W la ventana de tamaño 27×27	51
Tabla 3.6.	Índices de Moran's $\{I(x_{i,j}, h), h = 1..5\}$ para la ventana comprendida entre los puntos $\{x_{i,j}, i = 2..17, j = 21..26\}$ de la Tabla 3.5.....	55
Tabla 3.7.	Valores de β_2 para W a partir de los índices de Moran's $\{I(x_{i,j}, h), h = 1..5\}$ y de la función lineal	57
Tabla 3.8.	Los parámetros de ajuste para los índices $\{E(I(k)), k = 1..5\} = \left\{ \frac{-1}{8}, \frac{-1}{24}, \frac{-1}{48}, \frac{-1}{80}, \frac{-1}{120} \right\}$ de la función lineal $\hat{I} = \beta_1 + \beta_2 x$	59
Tabla 3.9.	Vector media (longitud y ángulo media) de las celdas que pertenecen al rectángulo enmarcado de la Tabla 3.5	67
Tabla 3.10.	El test estadístico T^2 para W . En negrita las celdas que no verifican la condición de direccionamiento	69
Tabla 4.1.	Valores Φ_i, d_i y s_i en intervalos $\lambda = 30^\circ$	93
Tabla 4.2.	Valores de la expansión mostradas en la Fig.4.35	100
	Tabla de símbolos	144

Capítulo I

Anisotropía: Estado del Arte

En este capítulo definiremos el concepto de anisotropía desde distintos puntos de vista, considerándola como una característica visual o como el comportamiento de ciertos minerales.

Dentro de la anisotropía espacial, analizamos dos tipos: la anisotropía vertical y la horizontal, definiéndola de una manera más formal a partir de un modelo estadístico.

El capítulo finaliza con el estudio del tratamiento de la anisotropía por los Sistemas de Información Geográficos más conocidos en este momento.

1.1. Anisotropía

La palabra Anisotropía es la conjunción de tres palabras con significado propio: “an” (partícula negativa), “iso” (igual o equivalente) y “tropo” (figura) quedando por tanto definida como el “*desigual comportamiento en las diferentes direcciones del espacio*” [Mar08] [DinTur⁺01]. La anisotropía es una de las cualidades más importantes de la visión espacial del ser humano.

Se fundamenta en el hecho biológico de ser habitantes de la corteza terrestre, sin capacidad para volar de forma autónoma ni para bucear libremente en las profundidades marinas. Esta situación biológica conlleva una percepción del espacio que no es homogénea en sentido vertical y horizontal.

Nos alzamos sobre el eje vertical, y la simetría respecto a este eje nos proporciona enormes ventajas adaptativas en el consumo de energía, tanto para ejercer el movimiento como para mantener el equilibrio. Estos condicionantes nos ayudan a entender las fuerzas que actúan en el espacio en el que nos movemos, así como a comprender que no son iguales las dimensiones horizontal y vertical del espacio en el que nos desplazamos.

Medio kilómetro es una distancia que puede recorrer sin problemas cualquier individuo sano de nuestra especie. Sin embargo, subir quinientos metros en vertical es una aventura extraordinaria, incluso cuando se realiza por las escaleras de un rascacielos. Esta desigualdad entre las diferentes dimensiones del espacio, tal y como son percibidas por los humanos, es lo que se llama anisotropía espacial. Dentro de esta clasificación, se distinguen dos tipos de anisotropía: vertical y horizontal.

1.1.1. Anisotropía vertical

En la visión humana la anisotropía se manifiesta en sentido vertical debido a la acción general que ejerce sobre todos los organismos la fuerza gravitatoria, y en sentido horizontal (izquierda/derecha) a causa de la asimetría cerebral propia de los humanos.

La anisotropía es una propiedad perceptiva de nuestro campo visual que influye de manera general en todo lo que vemos. Tiene como efecto visual el

que una misma imagen no sea percibida de la misma forma cuando está situada a derecha o izquierda en el campo de visión, en la zona superior o en la zona inferior. Este efecto se puede observar al girar 180° el objeto izquierdo de la Fig.1.1.



Figura 1.1. Anisotropía vertical número

El número tres tiene un diseño característico, formado por dos lóbulos abiertos, que en la mayoría de los casos son ligeramente distintos. Al invertir el número, el lóbulo mayor, que suele ocupar la parte de abajo, cobra un protagonismo excesivo y hace que el número parezca deformado o "cabezudo". Esta desigualdad es una manifestación de anisotropía en el eje vertical.

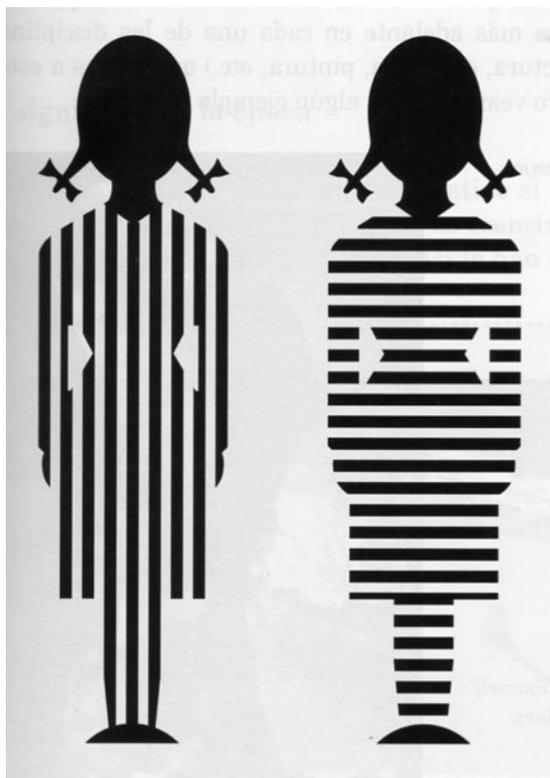


Figura 1.2. Efecto visual debido a la anisotropía

1.1.2. Anisotropía horizontal

Entre otros factores que se han considerado para justificar el fenómeno de desigualdad perceptiva entre el lado derecho e izquierdo de nuestro campo visual se han enumerado los siguientes:

- La costumbre de seguir la dirección de lectura. El problema es que los países que leen al revés o en sentido vertical, no presentan diferente anisotropía, dándose también en los analfabetos.
- Las diferencias existentes entre ambos lóbulos del cerebro.
- Las diferencias de funcionalidad e intensidad perceptiva entre ambos ojos: como sucede con las manos, siempre somos zurdos o diestros en lo que respecta a la visión. Las personas usualmente diestras (la mayoría) miran de izquierda a derecha. Por ello, la publicidad en los periódicos tiene un valor de contratación doble en las páginas pares que en las páginas impares.

La anisotropía horizontal es la responsable de que las mismas cosas, situadas a la derecha o a la izquierda del campo visual, generen sensaciones diferentes, que suelen estar polarizadas.

La obra de Vermeer "La calle", que nos muestra la Fig.1.3., con su estructura de división vertical casi al cincuenta por ciento entre el edificio y el callejón, es un buen ejemplo de cómo la percepción de una misma forma experimenta variaciones notables según esté situada en una zona u otra del espacio.

Tenemos la sensación de que el callejón se ve más o se ve menos, de que es mayor o menor el espacio que queda junto al edificio de ladrillo.

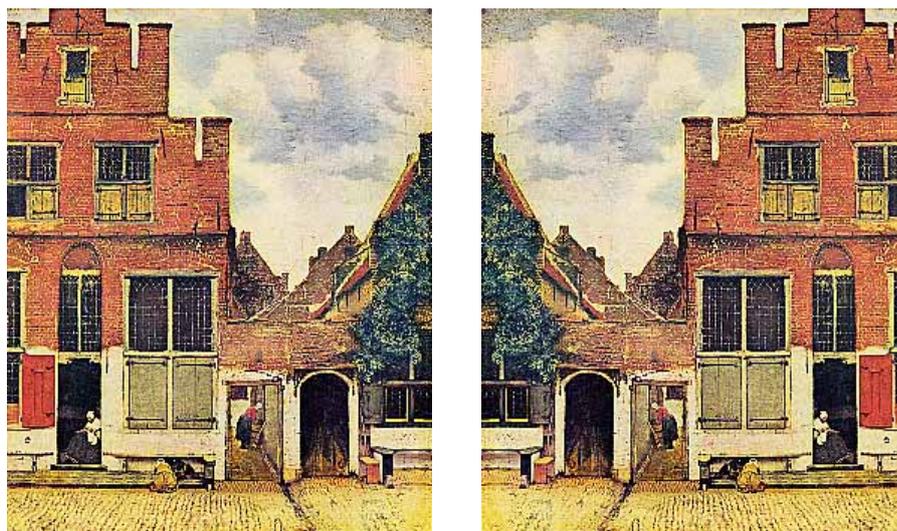


Figura 1.3. “La Calle” obra de Veermer

El direccionamiento ha sido utilizado desde antiguo en el arte a la hora de establecer la composición de la obra. La composición determina la ordenación de todos los elementos del lenguaje plástico en el conjunto de la obra de arte. No se trata por tanto de una simple ordenación de las figuras e imágenes en el espacio, sino de la articulación de todos los elementos de la expresión plástica y de la sensación que en conjunto ello nos produce [HubChe⁺08].

Las posibilidades compositivas de una pintura son casi infinitas, aunque predominan siempre dos tendencias más generales, las composiciones cerradas y las abiertas. Las composiciones cerradas o centrípetas, utilizan todos sus elementos plásticos para converger hacia el centro del conjunto, sin dirección definida, subrayando así su estabilidad y anulando cualquier indicio de movimiento en la obra. Suelen distribuir las figuras siguiendo criterios geométricos más o menos claros, como cuadrados, triángulos o círculos, transmitiendo armonía y equilibrio, lo que las convierte en las preferidas de movimientos clásicos.



Figura 1.4. Leonardo da Vinci, "La Virgen de las Rocas". Arte del Renacimiento (composición cerrada)

Las composiciones abiertas por el contrario divergen hacia los extremos de la obra, agitando toda la estructura compositiva y dinamizando la escena. En algunos casos suelen aplicarse disposiciones oblicuas o diagonales, que fomentan la sensación de movimiento, lo que explica su utilización en estilos como el Barroco o el Romanticismo. En otros casos las figuras se distribuyen en forma de zig-zag, o líneas sinuosas y curvas, en aspa, etc., consiguiendo el desequilibrio de las imágenes y empleándose cuando se quiere generar una sensación de caos y confusión.



Figura 1.5. Izquierda: Bernini, "Apolo y Dafné" (siglo XVII, estilo Barroco). Derecha: Raebun, "El reverendo Robert Walker patinando", siglo XIX (composiciones abiertas)

1.2. La anisotropía como característica de los minerales

Un mineral se define como la combinación química natural o elemento químico nativo determinado que constituye la corteza terrestre, y que presenta casi siempre una misma estructura cristalina. Una de sus propiedades fundamentales, y que por tanto definen al medio cristalino, son la homogeneidad y la anisotropía.

Cuando dos porciones iguales de una materia cristalina con la misma orientación en el espacio, no presenta ninguna diferencia en cuanto a sus cualidades fisicoquímicas, se dice que conserva una estructura homogénea.

La anisotropía se puede definir como la característica según la cual, determinadas propiedades de un cristal dependen de la orientación que se considere en cada momento. Así, la conductividad eléctrica, la calorífica, la dilatación térmica, la velocidad de propagación de la luz, etc., son muy diferentes según la dirección en que se manifieste la característica.

En el caso de la propagación de la luz en el interior de un cristal de cuarzo, por ejemplo, su velocidad dependerá de la dirección que los rayos sigan en su interior, ya que estructuralmente, la distancia entre los nudos vecinos de una red cristalina varía según la dirección, y en consecuencia, también las propiedades del cristal.

Las sustancias cristalinas cuyas propiedades físicas varían con la dirección se denominan anisotrópicas, y se clasifican dentro de las propiedades físicas vectoriales, es decir, aquellas que se representan a través de vectores que indican dirección, intensidad y sentido.

1.3. La Anisotropía de fisuración

Entre los diversos problemas que se presentan en Sismología (localización, fuente, estructura) se ha trabajado, tradicionalmente, con el supuesto de que el medio en el que tiene lugar la propagación de las ondas sísmicas, es isótropo, lo que quiere decir que sus propiedades elásticas son iguales independientemente del camino que sigue la onda en su propagación. Esta ha sido una suposición, también habitual, cuando el medio considerado es la corteza terrestre en el que, por otro lado, se ha comprobado la presencia de heterogeneidad [Oya96].

La consideración de anisotropía sísmica en la corteza terrestre viene a enriquecer la imagen que de ella disponemos, y a aportar información sobre procesos tectónicos que pueden estar en el origen de esta anisotropía.

La distinción entre un medio heterogéneo y un medio anisótropo, en consecuencia homogéneo, se establece a partir de la longitud de onda. En un medio heterogéneo, la variación de las propiedades elásticas tiene lugar para distancias mayores que la longitud de onda considerada. La hipótesis de anisotropía sísmica supone que la onda elástica encuentra diferentes propiedades en el medio dependiendo del camino que toma en su propagación. La sucesión de estas variaciones, en el camino de propagación de la onda, da lugar a un efecto observable en la velocidad y en la forma de la onda que permite detectar la presencia de anisotropía.

En los últimos años, se ha detectado anisotropía sísmica en diversas zonas de la corteza superior de la tierra. En muchos de estos casos se ha establecido que el origen de esta anisotropía era tectónico, y estaba de acuerdo con el modelo de anisotropía de fisuración (modelo E.D.A., Extensive Dilatancy Anisotropy) propuesto por Crampin en 1978. Este modelo se basa en lo siguiente: “el campo de esfuerzo actual controla la orientación de las fisuras en la corteza superior de forma tal, que tienden a alinearse vertical y paralelamente a la dirección del máximo esfuerzo compresivo horizontal”.

Imaginemos que una onda S incide verticalmente en el medio anisótropo, en su desarrollo se desdobra en dos ondas S_1 y S_2 polarizadas perpendicularmente que viajan a diferente velocidad: la primera fase S_1 (la onda S rápida) está polarizada en la dirección de las fisuras y, por tanto, del máximo esfuerzo compresivo horizontal; la segunda fase S_2 (la onda S lenta), que llega al cabo de un cierto retraso temporal, está polarizada perpendicularmente a la anterior.

Por lo tanto, los dos parámetros que permiten caracterizar la anisotropía de fisuración, y que son observables en nuestros datos son: la dirección de polarización de la onda S_1 rápida, relacionada con el régimen de esfuerzos, y el retraso temporal entre las ondas S_1 y S_2 , que está directamente ligado a la longitud del camino que recorre la onda S en el medio anisotrópico, y a la densidad de fisuración del propio medio.

Si tenemos en cuenta la anisotropía que genera la dependencia de los campos de ondas respecto a la dirección en la que éstos se registran podemos hablar de dos tipos de anisotropía: la generada por la presencia de series de estratos con espesores pequeños comparados con las longitudes de onda observadas y la presencia de diferentes configuraciones de sistemas es decir, configuraciones ortorrómbicas o monoclinicas. Las primeras reciben el nombre de anisotropías del tipo ITV (Isotropía Transversal con un eje de simetría Vertical) y la segunda ITH (Isotropía Transversal con un eje de simetría Horizontal).

1.3.1. Una descripción más formal

La especificación de los modelos estadísticos espaciales requiere en primer lugar, la visualización de datos de manera adecuada, normalmente a través de un mapa. A partir de la información obtenida, trataremos de generar un modelo estadístico para la especificación de una o más funciones de probabilidad para cada una de las variables aleatorias utilizadas en el proceso [CasMon⁺07].

En general, el comportamiento de fenómenos espaciales es a menudo el resultado de una mezcla tanto de efectos de primer como de segundo orden. Con efectos de primer orden nos referimos al valor medio de los procesos en estudio (la tendencia global o a gran escala). Los efectos de segundo orden son el resultado de la correlación espacial, es decir, de la dependencia espacial que presenta el proceso. Estos efectos están relacionados con el comportamiento de las desviaciones aleatorias de la media (efectos locales).

Las componentes de segundo orden a menudo se modelan como un proceso espacial estacionario. Un proceso espacial $\{Z(s), s \in D\}$, es estacional u homogéneo si no hay relación entre los valores estadísticos que lo describen (especialmente la media y la varianza) y su localización en el espacio. Consecuentemente, la estacionalidad implica que la media $E(Z(s))$, y la varianza $Var(Z(s))$, son constantes e independientes de su localización en D .

Estacionalidad también implica que dados dos puntos cualesquiera s_1 y s_2 , el valor de su covarianza $Cov(Z(s_1), Z(s_2))$, depende únicamente de las localizaciones relativas de los dos puntos, y no de su localización absoluta en el dominio D .

Si además de la estacionalidad, la covarianza depende únicamente de la distancia entre s_i y s_j , y no de la dirección de la recta que determinan ambas localizaciones, afirmamos que $Z(s)$ es isotrópico. Si por el contrario, la covarianza que determinan presenta dependencia de la dirección que s_i y s_j forman, $Z(s)$ es anisotrópico [BaiGat96].

Más formalmente, el proceso espacial $\{Z(s), s \in D\}$ se define como estacionario de segundo orden si:

$$E(Z(s)) = \mu \quad \text{para todo } s \in D \quad (1)$$

$$\text{Cov}(Z(s_1), Z(s_2)) = C(s_1 - s_2) \quad \text{para todos } s_1, s_2 \in D \quad (2)$$

La función $C(h)$ se denomina covariograma o función estacional de covarianza. Además, si $C(s_1 - s_2)$ depende solamente de $|s_1 - s_2|$, $Z(s)$ es isotrópico [Cre93]. Si la media, varianza o covarianza se muestran “sin rumbo” en D podemos afirmar que el proceso presenta heterogeneidad o no estacionalidad. La dependencia espacial se puede considerar como un caso especial de homogeneidad. La dependencia espacial implica que los datos referidos a una determinada unidad espacial están relacionados y son similares a los datos obtenidos en parcelas cercanas espacialmente [Get94]. La existencia de dependencia espacial es equivalente a la primera ley de la Geografía de Tobler, que establece que *“Todo está relacionado con todo pero las cosas próximas están más relacionadas que las cosas distantes”* [Tob70].

1.4. La anisotropía en los Sistemas de Información Geográfica

Para la manipulación de datos referenciados geográficamente, existen unas herramientas software de gran utilidad, denominadas Sistemas de Información Geográficos (SIG).

Una de las constantes en los manuales escritos sobre SIG es la enorme complejidad en su definición. Las causas fundamentales de esta dificultad son que integran en un mismo concepto:

- Datos y teoría sobre los datos.
- Hardware y software.
- Diversos intereses (científicos, de gestión y comerciales).
- Diversas disciplinas científicas (matemáticas, informática, cartografía, geografía, biología, ingeniería, ...) que previamente aparecían aisladas.

Existen numerosas definiciones de SIG, en función de las distintas concepciones que de él se tienen.

La primera referencia al término SIG aparece en 1967 referida a *“una aplicación informática cuyo objetivo era desarrollar un conjunto de tareas con información geográfica digitalizada”*. Se trataba del Sistema de Información Geográfica de Canadá (CGIS) [Tom67].

En 1986, Cebrián define un SIG como: *“Base de datos computerizada que contiene información espacial”* o también como: *“Una tecnología informática para gestionar y analizar información espacial”* [CebMar86].

Otra definición más amplia, y sin duda durante muchos años, la más difundida, sería la dada por Burrough en 1988: *“Un conjunto de herramientas para reunir, introducir en el ordenador, almacenar, recuperar, transformar y cartografiar datos espaciales sobre el mundo real para un conjunto particular de objetivos”* [Bur88].

En sentido similar podemos mencionar esta otra dada por Tomlinson en 1987 *“Sistema digital para el análisis y manipulación de todo tipo de datos geográficos, a fin de aportar información útil para las tomas de decisiones territoriales”* [Tom87]. O también: *“Sistema para capturar, almacenar, validar, integrar, manipular, analizar y representar datos referenciados sobre la Tierra”* [Doe87].

Otros autores insisten más en su aspecto informático: *“Sistema informático capaz de realizar las tareas para manejar datos georreferenciados: entrada, almacenamiento, recuperación, manipulación, análisis y representación”* [Ara89] o también, *“Tipo especializado de base de datos, que se caracteriza por su capacidad de manejar datos geográficos, es decir, espacialmente referenciados, los cuales se pueden representar gráficamente como imágenes”* [BraWest90].

Finalmente, es importante recordar la definición realizada en el manual sobre el tema, redactado por el NCGIA (National Center for Geographic Information and Analysis) de USA: *“Un sistema de hardware, software y procedimientos*

elaborados para facilitar la obtención, gestión, manipulación, análisis, modelado, representación y salida de datos espacialmente referenciados, para resolver problemas complejos de planificación y gestión” [Ncg90].

De manera más simple e inmediata un Sistema de Información Geográfica se puede considerar como “*un conjunto de capas de información digital con tablas de datos alfanuméricas asociadas a cada una de ellas y referidas a la misma porción del territorio, donde un lugar concreto tiene la misma localización (las mismas coordenadas) en todos los mapas incluidos en el sistema de información*”. De este modo, resulta posible realizar análisis de sus características espaciales y temáticas para obtener un mejor conocimiento de esa zona [Bos95].

Básicamente un SIG es “*un sistema para la adquisición, almacenamiento, análisis y visualización e impresión de datos geográficos*” [Mag91]. En la actualidad disponemos de muchos y variados SIG comerciales (*ArcInfo, Erdas, Grass, Idrisi...*); todos ellos tienen en común un conjunto de componentes básicos:

- **Base de datos espacial y temática.** La base de datos de un SIG, se compone de una colección de mapas digitalizados con información asociada para cada una de las entidades geográficas (parcelas, edificaciones, carreteras, ríos...). El conjunto de información que localiza espacialmente a dichas entidades forma la base de datos (BD) espacial, mientras que la información asociada a aquéllas forma la BD no espacial, temática o alfanumérica. En algunos sistemas, las BD espacial y temática se encuentran perfectamente diferenciadas tanto física como lógicamente, mientras que en otros, ambos tipos de información forman una única BD.
- **Sistema de digitalización de mapas.** Permite el paso de cartografía sobre soporte tradicional (papel) a cartografía digitalizada (soporte magnético).
- **Sistema de visualización e impresión cartográfica.** Existen diferentes componentes de software destinados a la producción de mapas (en pantalla o papel), a partir de los elementos seleccionados en la propia base de datos. El más utilizado de estos es el sistema de representación

cartográfica, que permite el uso de dispositivos de impresión, como la impresora o el trazador (plotter).

- **Sistema de gestión de Bases de Datos (SGBD).** El SGBD de un SIG no difiere significativamente de los SGBD de propósito general, ya que su función es la entrada, manejo y análisis de información no espacial o alfanumérica.
- **Sistema de análisis espacial.** El SGBD no espaciales no resuelve la mayoría de los problemas de tratamiento de información en los SIG, por lo que es necesario la adición de un módulo de análisis para el tratamiento de entidades espaciales. Un sistema de análisis espacial o territorial incluye, junto a los componentes tradicionales de consulta de bases de datos, la posibilidad de analizar datos, basándose en su localización. De hecho, esta capacidad para comparar elementos distintos o diferentes variables que se superponen en el territorio es sin duda el elemento distintivo de los SIG. Por eso, este tipo concreto de análisis se realiza mediante el proceso de superposición (overlay), que llamamos así porque tiene el mismo carácter que la tradicional superposición de mapas transparentes.

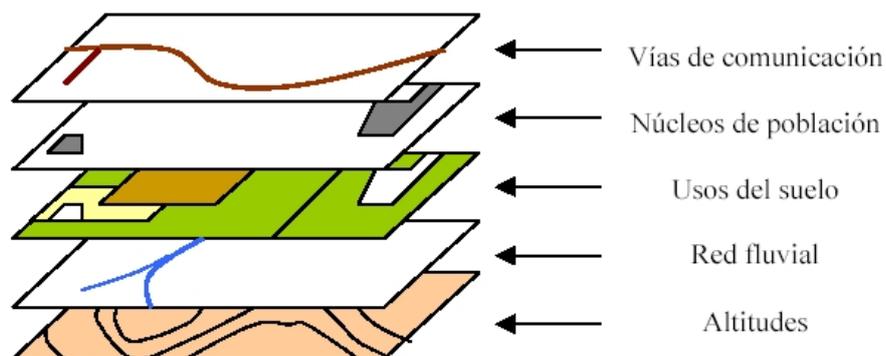


Figura 1.6. Ejemplo de overlay en Sistema de Análisis Espacial

- **Sistema de análisis estadístico.** También suele ser incluido en los SIG un conjunto de rutinas estadísticas, tanto para realizar cálculos no espaciales, como análisis de datos espaciales, que permiten descubrir patrones de comportamiento en esos datos y validar las hipótesis de trabajo.

- **Sistema de proceso de imágenes.** El más novedoso de los sistemas que incluyen los SIG es aquél que permite a éstos el análisis de imágenes captadas mediante sensores remotos. Las más usuales son imágenes de satélite LANDSAT y SPOT. El tratamiento de imágenes permite acceder a datos originales obtenidos mediante teledetección (fotos de satélite o fotografía aérea) y convertirlos en mapas temáticos gracias a procedimientos de clasificación o utilizarlos como otra fuente más de información. En la actualidad y a pesar de la importancia de esta técnica en la adquisición de datos territoriales, especialmente ambientales, no se considera como un componente esencial de un SIG, de manera que no es imprescindible la presencia de este módulo para que se pueda hablar de un verdadero SIG.

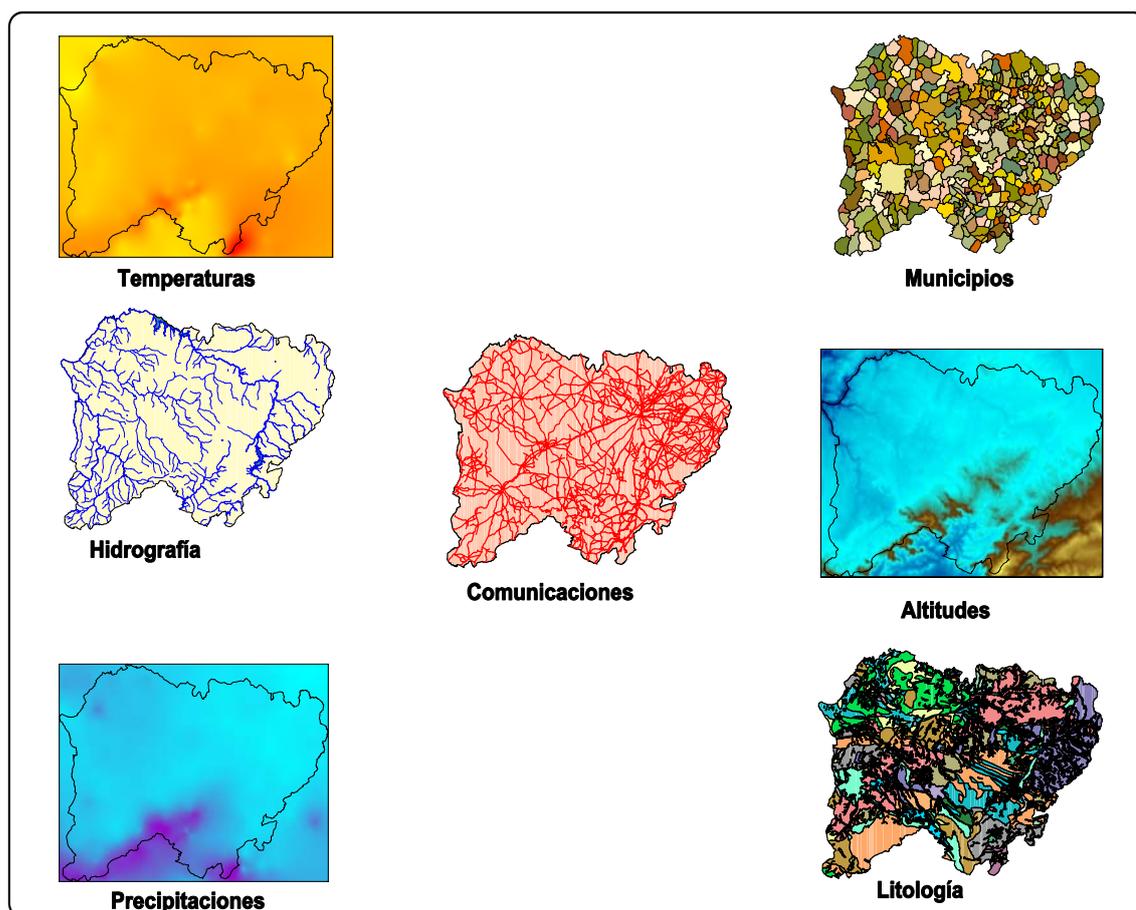


Figura 1.7. Superposición de diferentes capas de información

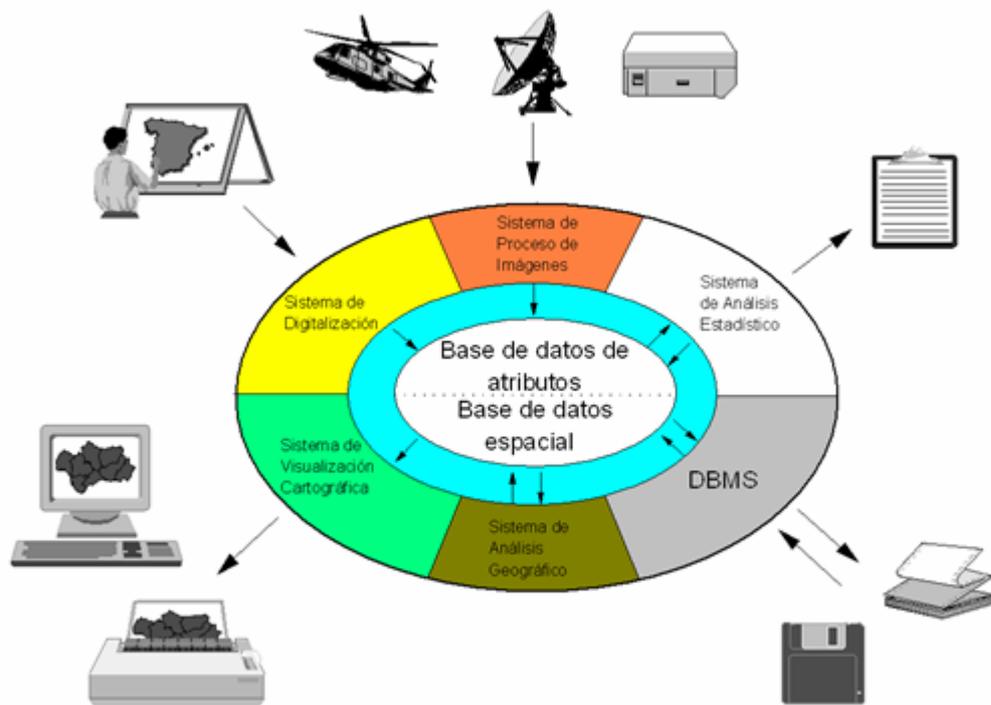


Figura 1.8. Componentes de un SIG

Uno de los problemas fundamentales en el desarrollo de un proyecto SIG es, al igual que en el caso de la cartografía tradicional, cómo representar una realidad compleja y continua de forma simplificada y en elementos discretos. El desarrollo de una base de datos espacial conlleva una simplificación de la realidad para adaptarla a un modelo de datos.

Existen dos modelos de datos básicos, el modelo vectorial y el modelo raster. En el modelo vectorial se considera que la realidad está dividida en una serie de objetos discretos (puntos, líneas, polígonos) a los que se puede asignar diversas propiedades, cualitativas o cuantitativas. Estos objetos se codifican por su posición en el espacio (puntos y líneas) o por las parcelas que limitan sus límites (polígonos). Los cambios de escala van a suponer en muchos casos que los objetos cambien de un tipo a otro.

El modelo raster o de retícula se centra en las propiedades del espacio más que en la precisión de la localización. Divide el espacio en celdas regulares o píxeles, asignándole a cada una un valor para cada una de las variables

consideradas. Los cambios de escala van a reflejarse en el tamaño de las estas celdillas.

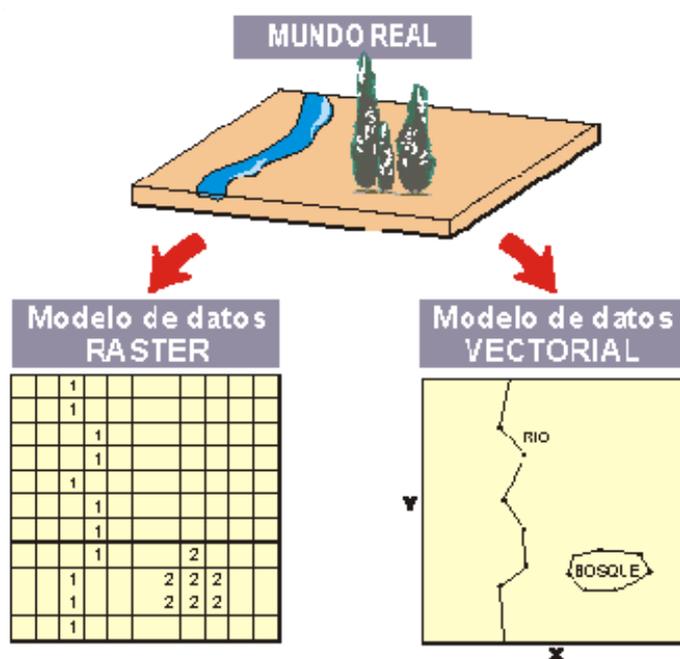


Figura 1.9. Representación de información Raster y Vectorial

Los SIG son utilizados para la resolución de múltiples problemas: análisis sobre emplazamiento óptimo de un determinado centro hospitalario y el acceso desde determinadas poblaciones, simulación y visualización de procesos medioambientales, estudio de la contaminación que producen los pesticidas y fertilizantes químicos en las aguas poco profundas de las cuencas hidrográficas, análisis de la influencia de los niveles de nitratos y sedimentos en el crecimiento de la flora, impacto ambiental producido por la implantación de nuevas industrias en las cercanías a centros urbanos...

Todas estas aplicaciones poseen un comportamiento isotrópico o no direccional. Sin embargo, existen fenómenos como la dispersión de gases contaminantes emitidos por industrias, que no puede ser considerados procesos direccionalmente uniformes, ya que la zona afectada por los gases dependerá de la dirección cambiante de los vientos. En estos casos, no es posible la utilización de las técnicas y herramientas que no consideran el aspecto direccional del problema.

Tal y como hemos comentado, los SIG comerciales poseen módulos que permiten realizar análisis espacial, sin embargo, las herramientas de análisis anisotrópico no están suficientemente desarrolladas, y se restringen a implementaciones basadas en el estudio de costes de distancias.

El SIG *Idrisi* [Idr03], producto desarrollado por los Laboratorios Clark incorpora un módulo de cálculo de coste anisotrópico llamado *Varcost* que puede ser usado para modelar los efectos de fricciones anisotrópicas asociados al desplazamiento a través de una superficie. Incluye la función anisotrópica $a(\alpha) = \cos^k(\alpha)$, siendo la fricción efectiva $d(\alpha) = \text{fricción_patrón}^{f(\alpha)}$, obteniendo diferentes comportamientos de $f(\alpha)$ para distintos valores de k .

Capacidades similares (denominadas funciones ponderadas de distancia) se han incorporado recientemente a la versión 9 de *ArcGIS* [Esr08]. *AccessMod* [BlaEbe04] [EbeMor⁺05] (desarrollado por la Organización Mundial de la Salud junto con la red de centros de excelencia) es una extensión de *ArcView* 3.x y utiliza el mismo módulo isotrópico *CostDistance* de análisis espacial que *ArcView* 3.x. y *ArcGIS* 9. Esta función también está disponible en la versión 3 de *ArcView* y en *MAPA* de *ArcGis* 9 [EbeMor⁺05] a través de una librería dinámica escrita en C++ (dll) basada en la librería *Boost* [Boo07], construida de tal modo que puede ser linkada directamente a los módulos de *AccessMod*. *MAPA* permite combinar la información isotrópica obtenida tanto de las capas de los modelos digitales del terreno, como de las capas representadas a través de vectores lineales (caminos y ríos).

El SIG *Grass* [Gra08] incluye la orden *R.spread* que permite simular expansiones anisotrópicas a través de una elipse y generar un mapa raster que acumulan valores correspondientes al tiempo invertido en generar la superficie de expansión.

1.5. El Sistema de Información Geográfica Idrisi

Idrisi [Idr03], dispone de dos módulos de cálculo de coste anisotrópico llamados *Varcost* y *Disperse*.

En los estudios de coste no solamente influye el coste generado por el

movimiento, debemos tener también en cuenta las fuerzas y fricciones que impiden o facilitan dicho desplazamiento. Cuando se usa tanto *Varcost* como *Disperse* se representan las fuerzas y las fricciones anisotrópicas a través de un vector. Para modelar el coste anisotrópico, se utilizan dos imágenes, una en la que se representa la magnitud de fuerzas y fricciones y la otra que almacena la dirección de las fuerzas y fricciones, expresadas como grados azimutales.

Debido a que los análisis pueden reflejar la actuación de diferentes fuerzas al mismo tiempo, existen otros dos módulos que nos permiten representar la combinación de fuerzas y fricciones: *Resultant* y *Decomp*. *Resultant* toma la información de una imagen que representa el par fuerza/fricción y genera otra imagen fuerza/fricción que representa el vector resultante de la combinación de acciones. *Resultant* puede ser usado sucesivamente para combinar fuerzas y fricciones produciendo una imagen de una sola magnitud (dirección) que se puede usar como entrada para *Varcost* o *Disperse*.

Decomp realiza la acción inversa a *Resultant*, dado un vector fuerza lo descompone en magnitud y dirección.

Sirvámonos de un ejemplo para explicar el coste anisotrópico. Supongamos que se quiere calcular el coste consumido al andar a una velocidad normal. Si traducimos el coste al número de calorías que quemamos a lo largo del paseo, la pendiente será un parámetro que influirá de manera directa en su consumo. Y aún más, podemos afirmar que la pendiente afectará de diferente manera dependiendo de la dirección. Ya que si subimos, la fricción producida será mayor que si nos movemos perpendicularmente a la pendiente; y en el caso de descender la pendiente actuará como una fuerza que reduce el coste. Por lo que podemos afirmar que existen fricciones anisotrópicas, es decir, elementos de fricción que producen efectos diferentes en direcciones diferentes.

Para modelar los efectos de fricciones anisotrópicas de fenómenos en movimiento producidos por una fuerza, *Idrisi* utiliza el módulo *Varcost*. Sin embargo, en aquellos casos en los que se desee modelar fenómenos de movimiento que no son producidos por ninguna fuerza externa pero que actúan

bajo fuerzas anisotrópicas, se debe utilizar la función *Disperse*.

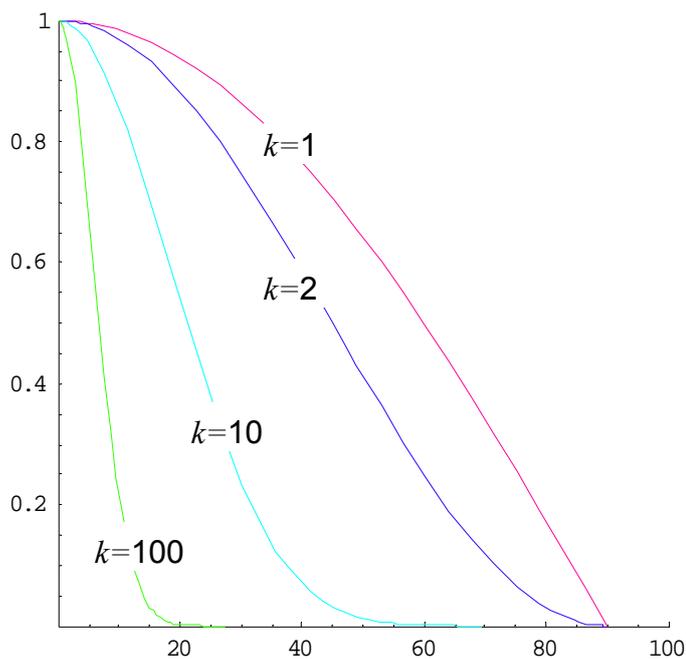
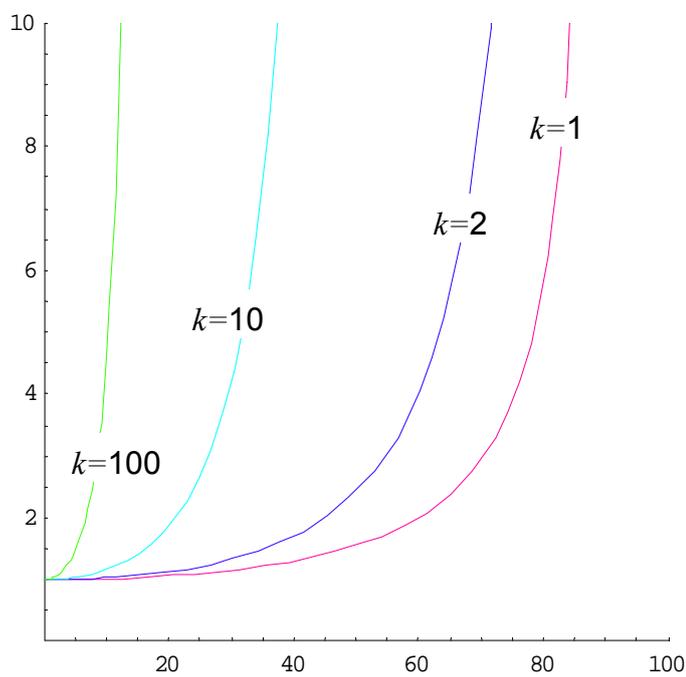
Un buen ejemplo sería el problema de la contaminación producida por los pesticidas en la tierra. Si estudiamos la absorción del pesticida por la tierra, podemos afirmar que el contaminante se moverá en las aguas subterráneas siguiendo la fuerza de gravedad según el gradiente hidráulico. El patrón de movimiento sería el siguiente: “decrece la probabilidad de que se mueva en la dirección que se ajuste al gradiente máximo (la pendiente)”.

Volviendo al ejemplo anterior en el que paseábamos en pendientes con grados diferentes, observamos que hay una conexión directa entre la pendiente y la fuerza de fricción, denominada *fricción_efectiva* o *stated friction*. Esta función cumple lo siguiente: “La *fricción_efectiva* toma el valor 1 en el punto de máxima pendiente”. De manera que si nos movemos en la dirección opuesta a la máxima fricción, la fricción podría descender progresivamente hasta alcanzar el valor de 90°.

La variación de la fricción/fuerza efectiva en función de la dirección es lo que llamamos función anisotrópica. En el caso de *Varcost* la fricción efectiva se calcula a través de la fórmula siguiente:

$$d(\alpha) = \text{fricción_efectiva} = \text{fricción_patrón}^{f(\alpha)} \quad (3)$$

donde $f(\alpha) = \cos^k \alpha$, siendo k un coeficiente definido por el usuario y α el ángulo diferencia (diferencia entre la dirección en la que el fenómeno alcanza su máximo valor de fricción y la dirección real de actuación).

Figura 1.10. Función $f = \cos^k \alpha$ Figura 1.11. Función $f = \frac{1}{\cos^k \alpha}$

Un exponente extremadamente alto haría que la fricción fuera cero. Independientemente del exponente usado, la función anisotrópica devuelve valores negativos para los ángulos comprendidos entre 90° y 270° .

Con *Disperse*, se aplica la misma lógica pero en este caso la función anisotrópica es diferente

$$d(\alpha) = \text{fricción_efectiva} = \text{fricción_patrón} * f(\alpha) \quad (4)$$

donde $f(\alpha) = 1/\cos^k \alpha$. El propósito de esta función es modificar las fricciones para alcanzar el valor máximo en el ángulo 0° con incrementos progresivos de fricciones, llegando al valor infinito para el ángulo de 90° . La Fig.1.12. muestra los valores devueltos por la función f , para el caso de las funciones *Varcost* y *Disperse*.

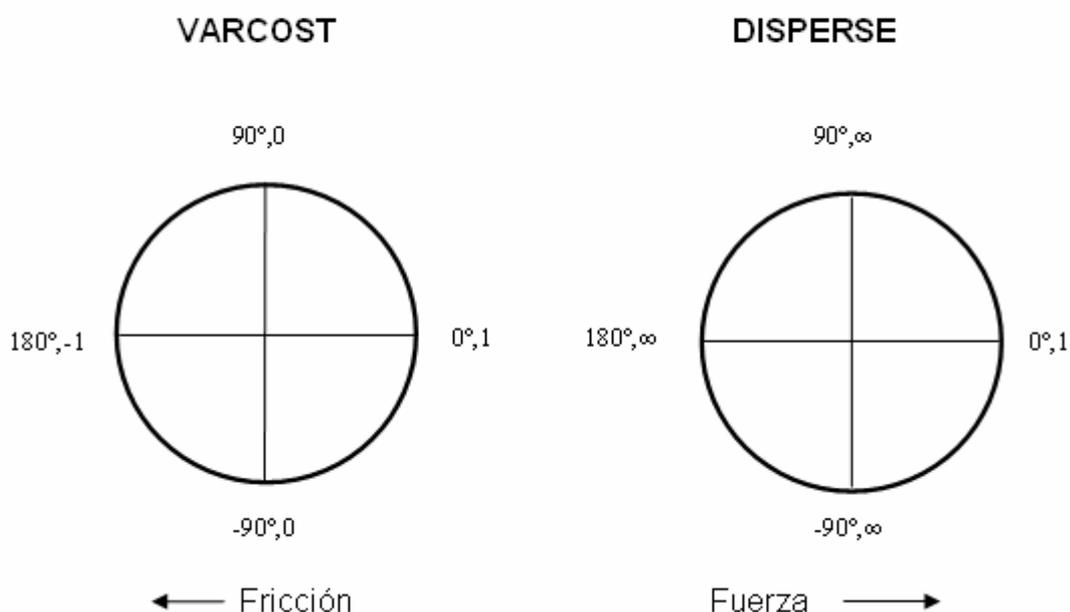


Figura 1.12. Comportamiento de las funciones *Varcost* y *Disperse* de *Idrisi*

Tanto *Varcost* como *Disperse* permiten la entrada de datos a través de una función definida por el usuario. En los dos casos el procedimiento es idéntico, se lee de un fichero los valores de la función para ángulos de diferencia de $0-360^\circ$ en incrementos de 0.05° . En este caso los valores de la función se multiplican por la fricción efectiva.

$$\text{fricción_efectiva} = \text{fricción_patrón} * f \quad (5)$$

Siendo f una función definida por el usuario.

Capítulo II

Estadística Circular

Existen fenómenos que poseen características que cambian en función de la dirección en que son observadas. Estas variables requieren un tratamiento estadístico particular y diferenciado al de las variables lineales mediante el uso de estadísticos descriptivos, métodos de exploración y distribuciones circulares propias. En este capítulo se describen los fundamentos de la estadística circular y las herramientas de análisis estadístico direccional que utilizaremos en los capítulos posteriores.

2.1. Estadística direccional

La estadística direccional se ocupa de datos que representan vectores en el plano o en el espacio 3D. En el primer caso, el espacio de muestreo es un círculo y en el segundo una esfera. Para tener en cuenta la estructura de esos espacios muestrales se necesitan métodos estadísticos especiales.

Algunos ejemplos de datos circulares son las direcciones de vientos predominantes, de vuelo de aves migratorias, es decir, cualquier dato que pueda ser medido o convertido en grados o radianes puede considerarse un dato circular.

Un dato circular se puede representar como un punto sobre un círculo unidad o como un vector unitario en el plano. Tomando una dirección como origen y eligiendo una orientación, cada dato circular se puede especificar como un ángulo θ entre la dirección inicial y la dirección correspondiente al dato (Fig.2.1.). La manera más simple de representar datos circulares es graficarlos como puntos sobre el círculo unidad y, cuando una dirección se repite, colocar los nuevos puntos fuera del círculo sobre el radio correspondiente (Fig.2.2.).

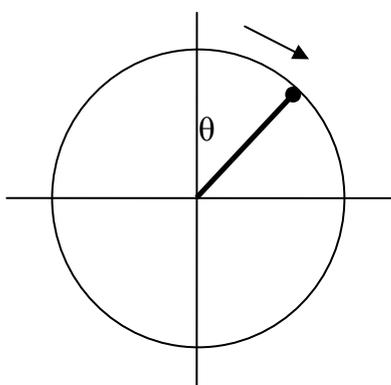


Figura 2.1. Representación de un vector unitario en el plano mediante un gráfico circular. La dirección del vector queda caracterizada por el ángulo θ

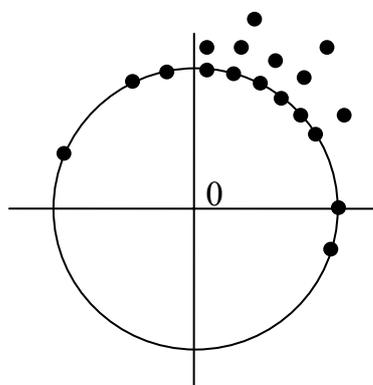


Figura 2.2. Representación de datos circulares como puntos sobre el círculo unidad

Podemos definir las variables circulares [Men04] como aquellas que representan direcciones en el plano, cuantificándose mediante ángulos que varían de 0 a 2π . Una de las diferencias más importantes respecto a las

variables lineales es que mientras que éstas pueden tomar valores de toda la recta real $(+\infty, -\infty)$, las variables circulares toman valores cíclicos y consecuentemente, la suma o diferencia de observaciones puede sobrepasar 360° o incluso dar un valor negativo, siendo posible en estos casos encontrar un valor equivalente perteneciente al intervalo $0-360^\circ$. Esta característica permite dar a las variables circulares un tratamiento distinto al de las lineales, mediante la creación de estadísticos descriptivos, métodos para análisis de correlación y distribuciones específicas para este tipo de variables.

Conceptualmente, una distribución circular puede considerarse como una distribución lineal bivariada en la que la probabilidad total (o masa total) se encuentra dispersa sobre el círculo unidad. Por lo tanto, al igual que en los estadísticos lineales bivariados, existe en estadística circular un vector media \bar{m} de módulo r y ángulo media $\bar{\Phi}$, en cuyo extremo está el centro de masas C de la distribución (Fig.2.4.).

Las referencias más remotas de análisis de datos circulares datan de finales del siglo XVIII, de la mano del geólogo y astrónomo británico Rev. John Mitchell (considerado el padre de la sismología por sus estudios sobre terremotos) quien emplea estos análisis para estudiar las separaciones angulares entre estrellas con el objeto de probar la hipótesis de que sus direcciones estaban uniformemente distribuidas [Men04] [Bat81].

Muchos de los primeros trabajos en estadística circular tuvieron su origen en intentar comprender la habilidad de las palomas mensajeras para retornar a sus hogares. Posteriormente estos análisis se extienden a los estudios del empleo que hacen las aves migratorias de su reloj interno y de la dirección del sol para orientarse.

La estadística direccional cobra un gran impulso cuando a comienzos de la década de 1950 investigadores de Geofísica y Geodesia de la Universidad de Cambridge llevan a cabo estudios de direcciones de magnetización remanente en rocas volcánicas de Islandia y encuentran que las direcciones de magnetización remanente forman dos grupos apartados 180° uno de otro. Sir Ronald Fisher, por entonces profesor de genética de esa universidad, se

interesa en el tema, retoma trabajos anteriores no publicados, diseña rápidamente su estadística en la esfera, y junto con el grupo de geofísicos aplican por primera vez estos métodos. No sería hasta 1953 cuando Fisher publica su trabajo clásico de estadística direccional “Dispersión en una esfera” [Fis53]. A partir de entonces la estadística de Fisher fue empleada continuamente por el grupo de paleomagnetistas de Cambridge, y actualmente es usada universalmente en estudios de campos magnéticos terrestres.

2.2. Vector media

El estadístico vector media \bar{m} se calcula asignando a cada uno de los n puntos del círculo de radio unidad de la Fig.2.4., una masa M , a partir de la cual se calcula el centro de masas o centro de gravedad C . Las proyecciones \bar{x} e \bar{y} de \bar{m} vienen dadas por las expresiones:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^j n_i \text{Cos } \Phi_i \qquad \bar{y} = \frac{1}{n} \sum_{i=1}^j n_i \text{Sen } \Phi_i \qquad n = \sum_{i=1}^j n_i \qquad (6)$$

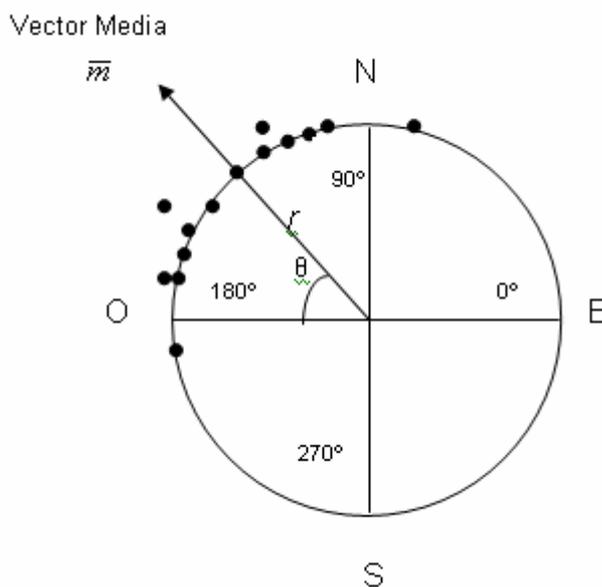


Figura 2.3. Vector media

Sea la variable aleatoria $Z(\Phi)$. Si tomamos una muestra monomodal de frecuencias n_1, n_2, \dots, n_j en las direcciones $\Phi_1, \Phi_2, \dots, \Phi_j$, se define el vector media $\bar{m}(r, \bar{\Phi})$ como

$$r = \sqrt{\bar{x}^2 + \bar{y}^2} \quad \bar{\Phi} = \begin{cases} \text{Arctan}[\bar{y}/\bar{x}] & \text{si } \bar{x} > 0 \\ 180 + \text{Arctan}[\bar{y}/\bar{x}] & \text{si } \bar{x} < 0 \end{cases} \quad (7)$$

donde \bar{x} e \bar{y} son las proyecciones de \bar{m} sobre los ejes X e Y respectivamente. La dirección del vector se encuentra en la recta que une C con las coordenadas del origen O . Cuando los datos están agrupados en arcos j con una longitud de $\lambda = \frac{2\pi}{j}$ los valores de \bar{x} e \bar{y} son los siguientes:

$$\bar{x} = \frac{1}{n} [n_1 \cos \Phi_1 + n_2 \cos \Phi_2 + \dots + n_j \cos \Phi_n]$$

$$\bar{y} = \frac{1}{n} [n_1 \text{sen} \Phi_1 + n_2 \text{sen} \Phi_2 + \dots + n_j \text{sen} \Phi_n]$$

(8)

Donde n_1, n_2, \dots, n_j son las frecuencias de los puntos medios $\Phi_1, \Phi_2, \dots, \Phi_j$ de los j arcos. El valor θ se estima de la misma forma que en muestras agrupadas. Sin embargo necesita un factor de corrección. Sin corrección, r tiende a ser un poco más pequeño. Por lo tanto hemos de multiplicar la r por un factor $c > 1$. Siendo el módulo corregido $r_c = r \cdot c$, donde el factor de corrección c es:

$$c = \frac{\lambda/2}{\text{Sen}(\lambda/2)} \quad (\lambda \text{ en radianes}) \quad (9)$$

El cálculo de \bar{m} descrito es válido para muestras monomodales. Sin embargo, la experiencia muestra que los fenómenos unidos a discontinuidades orográficas pueden ser plurimodales. El proceso de muestras v -modales (donde v es el número de modas) difiere del descrito, ya que éstas deben ser tratadas como si fueran muestras generadas por v distribuciones unimodales, por lo que podemos hablar de una mezcla de distribuciones.

Las muestras v -modales deben de ser consideradas como extraídas de una distribución generada por el solapamiento de v distribuciones monomodales. Cuando las distancias entre modas son arbitrarias no existen métodos

estándares para descomponer una muestra ν -modal en ν muestras monomodales; en la naturaleza las muestras plurimodales suelen manifestarse como bimodales y diametralmente opuestas. En este caso, es posible reducir la muestra bimodal a una muestra monomodal duplicando los ángulos. Con los nuevos ángulos, se calcula el vector media $\bar{m}_2(r_2, \bar{\Phi}_2)$ utilizando Eq.(6)(7)(8). Para obtener el ángulo modal simétrico $\bar{\Phi}_1$ de la muestra original, hemos de cancelar el efecto de la duplicación de los ángulos, siendo:

$$\bar{\Phi}_1 = \bar{\Phi}_2/2 \quad \text{o} \quad \bar{\Phi}_1 = \bar{\Phi}_2/2 + 180^\circ \quad (10)$$

2.3. Distribución de Von Misses

De entre las distribuciones circulares existentes, una de las más usadas para la modelización de variables circulares es la distribución de Von Misses. En el análisis direccional, la distribución de Von Misses juega un papel similar al de la distribución Normal en análisis lineal. Por esta razón, también es conocida como “distribución circular Normal”. La función de densidad para muestras ν -modales y simétricas es:

$$f(\Phi) = \frac{1}{2\pi I_0(k)} \text{Exp}[k \text{Cos } \nu(\Phi - \theta)] \quad (11)$$

Donde I_0 es la función de Bessel de argumento imaginario puro de orden 0, ν es el número de modas y k es el parámetro de concentración [Sch78], que indica en qué medida se concentra la distribución alrededor de la dirección dominante θ . La longitud del vector media r es un valor entre 0 y 1. Si r se aproxima a 1, la población se concentra alrededor θ . Por el contrario, valores de r próximos a 0 indica una dispersión de la población sobre el círculo.

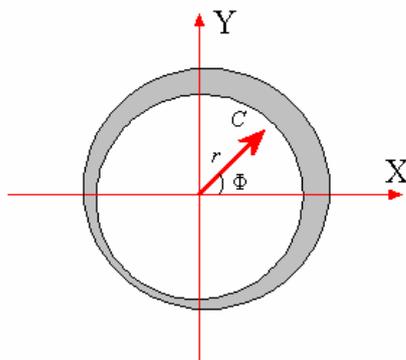


Figura 2.4. Distribución circular

Sean μ_1, ρ_1 y θ_1 respectivamente los parámetros poblacionales de m, r y $\bar{\Phi}$. Si a su vez, denotamos respectivamente $\hat{\mu}_1, \hat{\rho}_1$ y $\hat{\theta}_1$ como los estimadores de máxima verosimilitud de μ_1, ρ_1 y θ_1 , se cumple que

$$|\mu_1| = \rho_1 = A(k) \quad (12)$$

donde la función $A(x)$ se define como

$$A(x) = I_1(x)/I_0(x) \quad (x \geq 0) \quad (13)$$

Mardia demuestra [Mar75] que el estimador de máxima verosimilitud \hat{k} , es la solución de

$$A(\hat{k}) = I_1(\hat{k})/I_0(\hat{k}) = r \quad (14)$$

Usando Eq.(10) obtenemos

$$\hat{\rho}_1 = r \quad (15)$$

De modo que el estimador teórico de máxima verosimilitud de la longitud del vector media coincide con el vector media de longitud r . El parámetro k se puede calcular a través de r usando las tablas en [Bat81]. Una descripción más formal de la relación establecida entre r y k se describe en [Bur90].

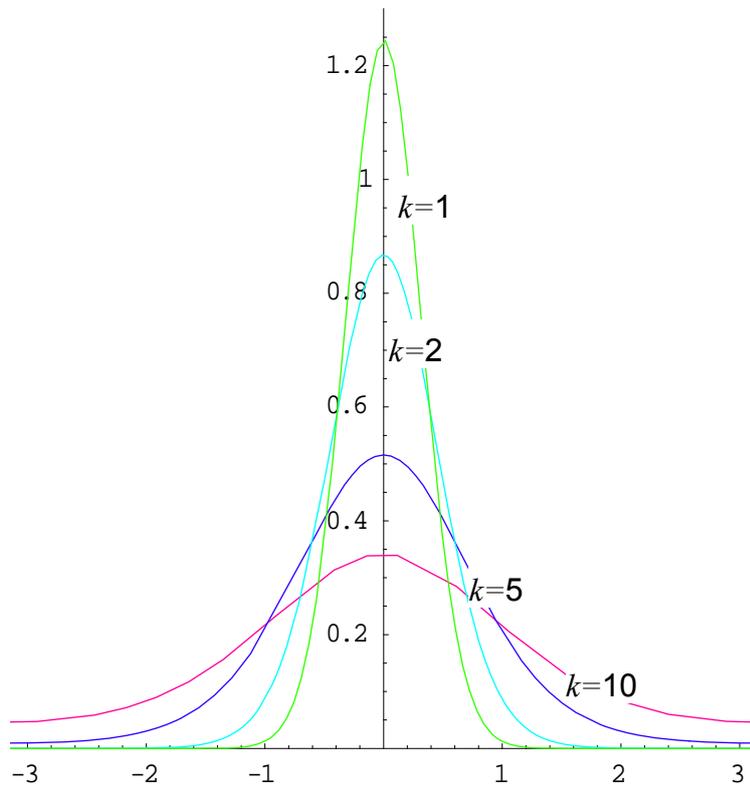


Figura 2.5. Distribución de Von Misses para distintos valores de k

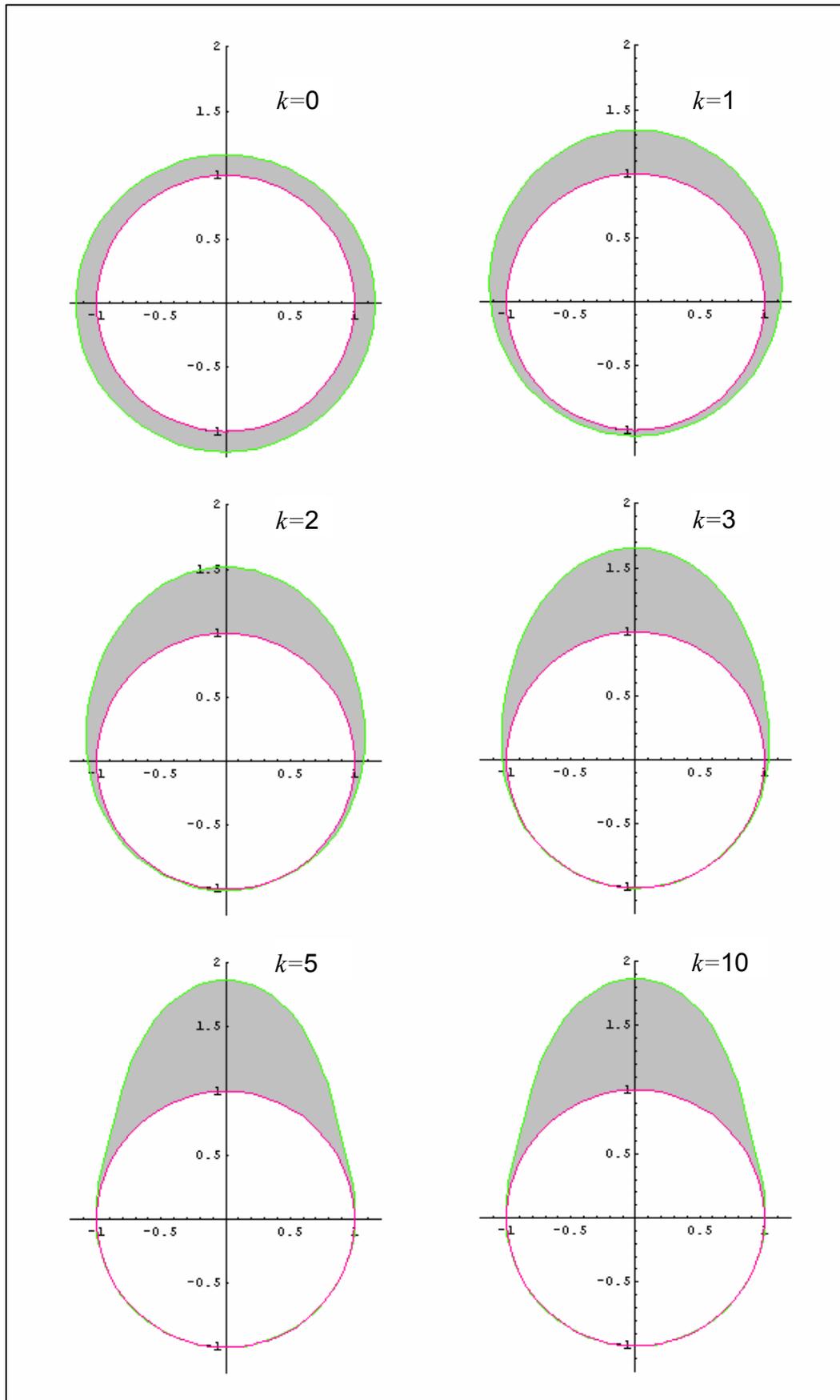


Figura 2.6. Comparación de la distribución de Von Misses frente a la distribución circular para distintos valores de K

2.4. Test de Rayleigh

Existen numerosos test de uniformidad (no direccionamiento) que permiten comprobar si la población de la que se extrae una muestra difiere significativamente de la aleatoriedad, o en otras palabras, si hay existencia de direccionamiento, o monolateralidad. Entre los numerosos test de uniformidad, el más usado por su simplicidad y generalidad es el test de Rayleigh [Bat81]. Se parte de una muestra aleatoria de n valores angulares $\Phi_1, \Phi_2, \dots, \Phi_n$ que representan direcciones en el plano.

Para muestras procedentes de una distribución uniforme, la longitud del vector media r nos sirve como indicador de direccionamiento, r toma valores pequeños. Si r es suficientemente grande, la hipótesis de aleatoriedad puede ser rechazada asumiendo direccionamiento en la distribución de la población.

Se parte de la hipótesis nula H_0 :

$$H_0 : \text{La población se distribuye uniformemente (aleatoriedad)} \quad (16)$$

Sea r el vector media definido en Eq.(6) o bien en Eq.(8) (en el caso de que se consideren valores agrupados).

Utilizamos como test estadístico:

$$z = nr^2 \quad (17)$$

Los valores críticos $z(\alpha)$ en relación del tamaño de la muestra n y el nivel de error α , se muestran en la Tabla 2.1. Si $z \geq z(\alpha)$ el test es significativo, por lo que se rechaza la hipótesis nula H_0 a favor del direccionamiento o monolateralidad de la distribución poblacional, ya que en este caso, la longitud del vector media difiere significativamente de cero. En caso contrario, no se puede excluir la uniformidad.

n	$\alpha = 0.10$	0.05	0.02	0.01	0.001
30	$z(\alpha) = 2.30$	2.97	3.85	4.50	6.62
50	2.30	2.98	3.87	4.54	6.74
100	2.30	2.99	3.89	4.57	6.82
200	2.30	2.99	3.90	4.59	6.87
500	2.30	2.99	3.91	4.60	6.89
∞	2.30	3.00	3.91	4.61	6.91

Tabla 2.1. Valores de $z(\alpha)$ para las variables α y n

En los casos en los que se pueda asumir que la población es unimodal, significación en el Test de Rayleigh, prueba no sólo la existencia de direccionamiento, sino también la concentración de las direcciones entorno a la dirección media \bar{m} .

2.5. Test de Cox

Una distribución es un modelo que ayuda a interpretar y comprender un fenómeno estocástico. Mientras que el fenómeno es siempre complejo, el modelo ha de ser conceptualmente simple, por lo que no puede reproducir de forma exacta el comportamiento de aquél. Este hecho plantea la necesidad de validar el modelo cuantificando las desviaciones entre el comportamiento real del fenómeno y el comportamiento del modelo. Estadísticamente, esto se reduce a chequear si la distribución teórica $f(\Phi)$ se ajusta a los valores observados. Son muchos los tests utilizados para cuantificar la bondad del ajuste de una función, desde pruebas generales como la de Chi-cuadrado, hasta tests específicos para determinadas distribuciones. El test de Cox [Gsc78] está diseñado para chequear la bondad del ajuste de una muestra a la distribución de Von Misses. Dicho test se basa en los segundos momentos trigonométricos del coseno (m_c) y del seno (m_s) [Gsc78]:

$$m_c = \frac{1}{N} \sum_{i=1}^j n_i \cos(\Phi_i - \bar{\Phi}) \quad m_s = \frac{1}{N} \sum_{i=1}^j n_i \text{sen}(\Phi_i - \bar{\Phi}) \quad (18)$$

Para una población de Von Misses de parámetros $\bar{\Phi}$ y k , estos momentos

trigonométricos se distribuyen normalmente (de forma aproximada) con medias $I_2(k)/I_0(k)$ y 0 respectivamente. Se rechaza la asunción de que la población sigue una distribución de Von Misses si los segundos momentos trigonométricos m_s y m_c (Eq.(18)) se alejan demasiado de dichos valores media.

Capítulo III

Test para determinar la existencia de direccionamiento

La identificación de modelos espaciales requiere conocimientos previos sobre propiedades tales como el grado de homogeneidad en efectos locales y globales, y la existencia de una componente direccional que determine si el fenómeno es anisotrópico. Esto nos permitiría asumir algún tipo de estacionalidad, reduciendo el número de parámetros y haciendo que el modelo sea más manejable. Los métodos utilizados para estudiar la estacionalidad de forma local y global son variados y bien conocidos. Las técnicas usadas para el análisis de anisotropía, se limitan a la visualización de variogramas de diferentes direcciones. En este capítulo diseñamos un método basado en estadísticos circulares de segundo orden que nos permite determinar la existencia de anisotropía a partir de imágenes digitales cuantificando las direcciones en las que aparecen.

3.1. Test para determinar la existencia de direccionamiento

El diseño de modelos espaciales es sólo posible bajo la asunción de estacionalidad tanto a nivel local como global. Si no se asume ninguna clase de estacionalidad, sería extremadamente complejo modelar el fenómeno debido a que necesitaríamos un gran número de parámetros para conseguir formalizar matemáticamente su comportamiento. Además, es necesario determinar si el fenómeno presenta anisotropía, puesto que este hecho invalidaría el uso de los modelos definidos para las funciones isotrópicas. Por esta razón, para modelar los procesos espaciales se requiere no solamente del análisis de estacionalidad a nivel local y global sino también de un estudio direccional que nos ayude a determinar la existencia de un componente anisotrópico.

Existen muchas y variadas herramientas para estudiar la homogeneidad [Bur90]: métodos basados en la teselación (triangulación de Delanuy e interpolación de vecindad) de las muestras observadas para estimar la media de la población, medias móviles espaciales y estimaciones de densidad Kernel adaptativa para delimitar áreas homogéneas e identificar posibles modelos, K -funciones y funciones de covarianza (covariogramas) para describir respectivamente, propiedades de segundo orden de patrones y datos espaciales continuos, variogramas, estadísticos de autocorrelación espacial (Moran's I , Geary's C , Getis's G), autocorrelogramas, tendencia espacial...

Por el contrario, las técnicas de análisis de anisotropía son muy limitadas, apenas permiten detectar efectos direccionales a partir de variogramas de distintas direcciones [Wac95]. La observación de estos variogramas solamente nos permite asegurar de una manera informal que se produce direccionamiento y hacer una aproximación sobre la dirección en la que la anisotropía se muestra. A continuación se expone cómo a partir de imágenes digitales podemos determinar la existencia de anisotropía y cuantificar la dirección en la que aparece, lo que permitirá diseñar funciones anisotrópicas a partir de modelos isotrópicos [MolFei98].

3.2. Test de direccionamiento

La Fig.3.1. muestra una imagen pixelada de 17x17, en donde cada pixel almacena un valor comprendido entre 0 a 255 que representa el nivel de gris. Los píxeles que muestran intensidades de luz mayores, representan valores de la variable aleatoria $Z(s)$ (densidad de vegetación, concentración de polen en el aire...).

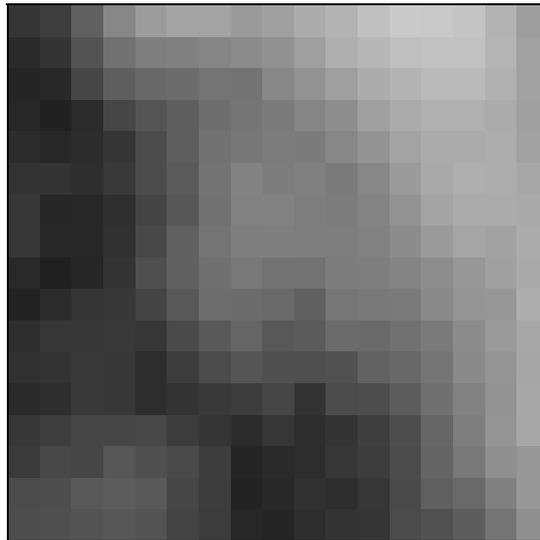


Figura 3.1. Pixelado de 17x17

Si observamos la imagen 3D mostrada en la Fig.3.2., podemos asegurar que el comportamiento del proceso en la dirección noroeste-sureste es diferente al de la dirección noreste-suroeste. Este hecho nos lleva a formular dos preguntas:

- ¿Existe realmente el comportamiento anisotrópico de $Z(s)$, o por el contrario, la tendencia direccional observada es provocada por fluctuaciones aleatorias?
- Si la anisotropía existe, ¿en qué dirección se produce?

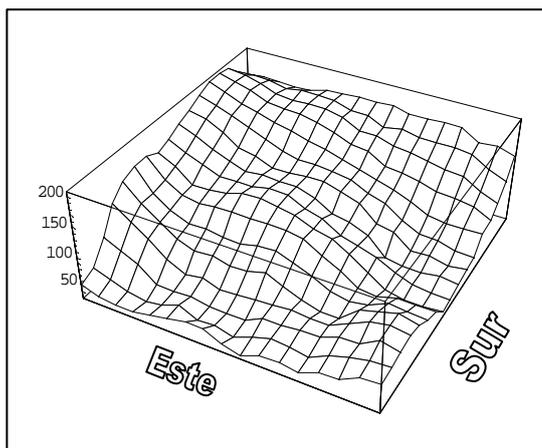


Figura 3.2. Dirección en la que se produce la anisotropía

La respuesta a estas preguntas requiere el desarrollo de métodos y el uso de las herramientas que nos permitan estudiar las variaciones espaciales del fenómeno, es decir, que nos permita establecer un patrón de anisotropía en la dirección en la que aparezca. Para esto, utilizaremos un procedimiento paramétrico basado en el cálculo de las elipses estándar y de confianza.

En la matriz $Z_{n,n}$ se almacenan los valores de la variable aleatoria $Z(s)$ obtenida a partir de la imagen:

$$Z_{l,l} = \begin{bmatrix} z_{1,l} & \dots & z_{l,l} \\ \vdots & & \vdots \\ z_{1,1} & \dots & z_{l,1} \end{bmatrix} \quad (19)$$

A partir $Z_{l,l}$, obtenemos $N = \{z_1, z_2, \dots, z_m\}$, siendo N un conjunto de m valores aleatorios. El vector $\overrightarrow{z_i z_j}$ representa una variación de $Z(s)$ con intensidad $|z_j - z_i|$ en la dirección γ , donde

$$\gamma = \left\{ \begin{array}{ll} \text{ArcTan} \left[\frac{z_{jy} - z_{iy}}{z_{jx} - z_{ix}} \right] & \text{Si } (z_j - z_i) \succ 0 \\ \text{Nulo} & \text{Si } (z_j - z_i) = 0 \\ \text{ArcTan} \left[\frac{z_{jy} - z_{iy}}{z_{jx} - z_{ix}} \right] + \pi & \text{Si } (z_j - z_i) \prec 0 \end{array} \right\} \quad (20)$$

A partir de los vectores $\overrightarrow{z_i z_j}$ [$i = 1..(l-1), j = (i+1)..l$], obtenemos

$\Gamma = \left\{ \gamma_i, i = 1..n, n = \binom{m}{2} \right\}$, que contiene las direcciones de los vectores sobre el plano.

Es evidente, que, a partir de las observaciones realizadas sobre una muestra concreta, no se puede obtener una conclusión sobre el comportamiento de los valores de la imagen para otras muestras sin tener en cuenta de cuántos datos disponemos y de la calidad del análisis. Por este motivo el análisis estadístico debe ser realizado en dos pasos, llamados análisis de primer y segundo orden o análisis de primer y segundo nivel [Bat78]:

- Para cada muestra obtenemos los ángulos γ_i mediante el cálculo de los estadísticos apropiados.
- Combinamos los estadísticos del paso anterior y testeamos su significación. Sólo entonces podemos realizar una inferencia estadística sobre el comportamiento direccional de la población a la que pertenece el conjunto de ángulos.

La medida más apropiada para nuestro propósito es el vector media \overline{m} . Dada una muestra $\Gamma = \{\gamma_i, i = 1..n\}$ su vector media \overline{m} (de longitud r y ángulo media $\overline{\Phi}$) se calcula según Eq.(6)(7).

Donde

$$n_i = \left\{ \begin{array}{ll} 0 & \text{Si } \gamma_i \text{ es nulo} \\ 1 & \text{Si } \gamma_i \text{ no es nulo} \end{array} \right\} \quad (21)$$

Para cada muestra $\Gamma_1, \dots, \Gamma_d$ calculamos el vector media de primer orden de longitud r y ángulo media $\bar{\Phi}$, obteniendo los pares $\{(r_i, \bar{\Phi}_i), i=1..d\}$. Estos valores se suponen mutuamente independientes y se calculan a partir de diferentes muestras con el mismo número de observaciones. La Tabla 3.3. muestra los valores del vector media $\{\bar{m}_i, i=1..6\}$ calculados a partir de las muestras N_1, \dots, N_6 (Tabla 3.2.) obtenidas a partir de $Z_{17,17}$ (Fig.3.1.).

54	63	97	136	157	165	164	155	161	173	180	193	203	202	197	179	159
43	53	84	114	127	129	134	140	146	160	176	183	192	194	194	181	162
38	41	71	96	105	109	117	116	136	147	160	173	180	187	187	177	162
40	34	44	71	86	95	111	118	124	134	143	161	172	177	177	172	161
45	40	45	54	77	95	114	120	125	124	136	148	164	172	172	174	162
52	52	47	58	77	91	116	131	126	129	124	136	156	170	175	174	168
56	39	40	48	69	88	114	130	131	127	125	132	147	165	172	172	171
56	40	40	50	73	97	117	127	128	127	127	130	141	155	166	163	173
42	33	39	53	80	98	113	121	115	115	126	127	133	142	152	161	171
36	44	54	56	69	89	110	109	107	97	119	122	123	139	150	151	174
47	55	56	58	55	75	90	102	89	92	108	107	114	123	141	155	168
50	52	57	56	46	61	76	86	81	80	82	99	105	118	139	149	167
43	47	58	56	46	53	58	62	71	52	76	78	94	113	131	147	169
56	63	71	71	73	61	54	45	55	47	53	64	79	103	127	149	168
59	73	71	87	81	74	62	38	43	46	56	62	76	101	122	144	151
76	79	90	93	90	71	62	36	41	50	47	55	75	97	106	129	152
78	81	84	88	84	69	61	40	38	46	53	54	76	82	94	117	144

Tabla 3.1. Matriz que almacena los valores de $Z(s)$ obtenidos a partir de la Fig.3.1.

	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8
N_1	(9,5)	(17,13)	(1,11)	(17,6)	(9,12)	(9,17)	(3,6)	(11,1)
N_2	(5,6)	(16,11)	(11,9)	(13,5)	(4,9)	(9,7)	(13,3)	(9,13)
N_3	(13,15)	(4,5)	(4,13)	(9,2)	(4,4)	(8,9)	(12,16)	(17,7)
N_4	(9,10)	(8,14)	(13,6)	(17,12)	(11,15)	(13,9)	(15,1)	(7,10)
N_5	(16,17)	(8,12)	(2,8)	(3,16)	(6,10)	(12,6)	(14,12)	(1,15)
N_6	(7,7)	(5,4)	(9,16)	(14,3)	(3,10)	(17,8)	(2,17)	(14,17)

Tabla 3.2. Muestras N_1, \dots, N_6 obtenidas a partir de $Z_{17,17}$ de la Fig.3.1.

i	1	2	3	4	5	6
r_i	.53	.60	.49	.53	.70	.59
$\bar{\Phi}_i$	30.7	36.4	50.5	65	8.5	14.5

Tabla 3.3. Vector media $\{\bar{m}_i, i = 1..6\}$ calculado a partir de las medias N_1, \dots, N_6 (Tabla 3.2.)

El vector \bar{m} se calcula a partir del ángulo $\bar{\Phi}$ y el módulo r ; en otras palabras, se tienen en cuenta el ángulo y la amplitud del módulo. Bajo estas condiciones, los pares $(r_1, \bar{\Phi}_1), (r_2, \bar{\Phi}_2), \dots, (r_d, \bar{\Phi}_d)$ se consideran una muestra bivariada de segundo orden, por lo que se puede afirmar que el tratamiento utilizado es un análisis de segundo orden.

3.3. Elipse estándar

Entre las herramientas usadas para el análisis de segundo orden se encuentra la elipse estándar. Se utiliza exclusivamente con propósitos descriptivos. Los extremos de los vectores de una muestra de segundo orden forman un diagrama de dispersión de datos con desviaciones estándar en las direcciones X e Y , con una determinada tendencia hacia arriba o hacia abajo. La elipse estándar describe este comportamiento de forma gráfica: asumiendo normalidad, aproximadamente el 40% de los puntos se encuentran en el interior dentro de la elipse y el 60% restante fuera. La población no necesita ser Normal, sin embargo es deseable que no se desvíe demasiado de la normalidad.

Sea $(x_1, y_1), (x_2, y_2), \dots, (x_d, y_d)$ una muestra aleatoria de una población bivariada. En general, x e y dependen uno de otro. Sin embargo, asumimos que (x_i, y_i) y (x_k, y_k) son independientes entre si. Definimos la elipse estándar del modo siguiente:

- Su centro es (\bar{x}, \bar{y}) .
- Construimos un rectángulo con centro (\bar{x}, \bar{y}) y lados paralelos a los ejes X e Y . Los lados tienen longitud $2s_1$ y $2s_2$ respectivamente.
- La elipse estándar se encuentra dentro del rectángulo definido anteriormente, siendo los puntos de tangencia $\bar{x} + rs_1, \bar{x} - rs_1, \bar{x} + s_1, \bar{x} - s_1, \bar{y} + rs_2, \bar{y} - rs_2, \bar{y} + s_2, \bar{y} - s_2$.

Para dibujar la elipse estándar, se necesitan dos medias, dos desviaciones típicas y un coeficiente de correlación.

$$x_i = r_i \text{ Cos } \Phi_i \qquad y_i = r_i \text{ Sen } \Phi_i \qquad (22)$$

$$\bar{x} = \frac{1}{d} \sum_{i=1}^d x_i \qquad \bar{y} = \frac{1}{d} \sum_{i=1}^d y_i \qquad (23)$$

$$s_1^2 = \frac{1}{d-1} \sum_{i=1}^d (x_i - \bar{x})^2 \qquad s_2^2 = \frac{1}{d-1} \sum_{i=1}^d (y_i - \bar{y})^2 \qquad (24)$$

$$\text{Cov}(x, y) = \frac{1}{d-1} \sum_{i=1}^d (x_i - \bar{x})(y_i - \bar{y}) \qquad r = \text{Corr}(x, y) = \frac{\text{Cov}(x, y)}{s_1 s_2} \qquad (25)$$

La ecuación de la elipse de confianza es:

$$\frac{1}{1-r^2} \left(\frac{(x-\bar{x})^2}{s_1^2} - 2r \frac{(x-\bar{x})(y-\bar{y})}{s_1 s_2} + \frac{(y-\bar{y})^2}{s_2^2} \right) = 1 \qquad (26)$$

o escrito sin fracciones

$$s_2^2 (x-\bar{x})^2 - 2rs_1s_2 (x-\bar{x})(y-\bar{y}) + s_1^2 (y-\bar{y})^2 = (1-r^2) s_1^2 s_2^2 \qquad (27)$$

La ecuación implícita de una elipse con centro en (\bar{x}, \bar{y}) es

$$A(x - \bar{x})^2 + 2B((x - \bar{x})(y - \bar{y})) + C(y - \bar{y})^2 = D \quad (28)$$

Para la elipse estándar, los coeficientes se obtienen a partir de

$$A = s_2^2 \quad B = -rs_1s_2 \quad C = s_1^2 \quad D = (1 - r^2)s_1^2s_2^2 \quad (29)$$

Los semiejes a y b siendo (a < b) son:

$$a = \left[\frac{2D}{A + C - R} \right]^{\frac{1}{2}} \quad b = \left[\frac{2D}{A + C + R} \right]^{\frac{1}{2}} \quad (30)$$

Donde

$$R = \left[(A - C)^2 + 4B^2 \right]^{\frac{1}{2}} \quad (31)$$

La muestra presenta la máxima variabilidad en la dirección del ángulo θ . Este es el ángulo formado por el eje mayor y el eje X ($-90^\circ < \theta < 90^\circ$).

$$\theta = \text{ArcTan} \left[\frac{2B}{A - C - R} \right] \quad (32)$$

Los valores de los parámetros descritos, se calculan a partir de la Tabla 3.3. y se muestran en la Tabla 3.4. En la Fig.3.3. se representa la elipse.

\bar{x}	\bar{y}	s_1	s_2	Cov	r	A	B	C	D	R	a	b	θ	ψ
.459	.291	.172	.144	-.023	-.941	.021	.023	.029	.00007	.047	.221	.038	-39.72°	32.34°

Tabla 3.4. Parámetros de la elipse de confianza calculada a partir de los vectores almacenados en la Tabla 3.3.

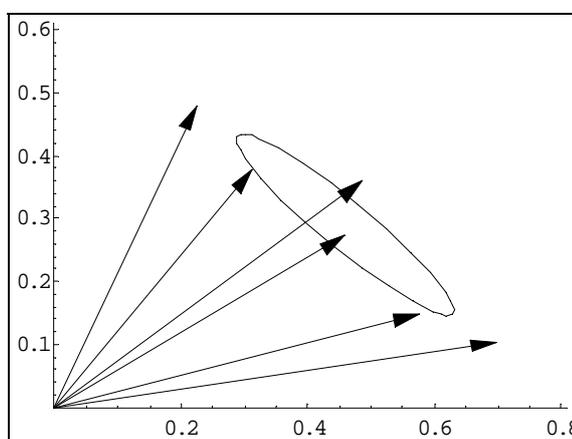


Figura 3.3. Elipse estándar obtenida a partir de los parámetros de la Tabla 3.3.

Como se ha comentado en esta sección, la elipse estándar se usa exclusivamente con una finalidad descriptiva. Sin embargo, a partir de ella podemos estimar si la asunción de normalidad en la población es razonable. Si los ángulos media se distribuyen uniformemente alrededor del origen de coordenadas, podemos considerar que la población de la que ha sido extraída la muestra no se aparta de la aleatoriedad ó monolateralidad, alejándose de la normalidad. Cuando esto ocurre, la elipse de confianza que se describe en la siguiente sección no se puede aplicar.

3.4. Elipse de confianza

Mediante la elipse estándar, hemos descrito el comportamiento espacial de los vectores media, cuantificando, mediante el cálculo de θ , la dirección de máxima variabilidad. Sin embargo, debemos determinar si la variación direccional se produce por fluctuaciones aleatorias de los vectores, o por el contrario, si la causa es la existencia de direccionamiento (anisotropía). La elipse de confianza se usa para este fin.

La elipse de confianza incluye una región definida en el plano XY que cubre el centro de la población (μ_1, μ_2) con una probabilidad preasignada $Q = 1 - \alpha$, siendo por lo tanto, una herramienta para realizar inferencia estadística. Si asumimos normalidad, la región se encuentra dentro de la elipse con el mismo centro que la elipse estándar y la misma inclinación del eje principal (θ). La determinación de la existencia de direccionamiento se resuelve generando la elipse de confianza y testeando que el origen está dentro de ella. Si no ocurre esto, el centro de la población (μ_1, μ_2) no coincide con el origen, y (\bar{x}, \bar{y}) es significativamente diferente de él (origen), concluyendo que los vectores media se orientan como grupo en la dirección ψ :

$$\psi = \text{ArcTan} \left[\frac{\bar{y}}{\bar{x}} \right] \quad (33)$$

Buscamos una región de dos dimensiones en las que se encuentra el punto (μ_1, μ_2) con probabilidad $Q = 1 - \alpha$. Para una muestra univariante, sabemos que [Hot31]

$$-t = \left(\frac{\bar{x} - \mu}{s} \right) n^{\frac{1}{2}} \quad (34)$$

sigue una distribución t con valores críticos $t(\alpha)$ y $t(-\alpha)$. Por lo tanto la desviación de μ respecto de \bar{x} está acotada por la desigualdad

$$(\mu - \bar{x})^2 \leq \frac{s^2}{n} t^2(\alpha) \quad (35)$$

El problema de los valores bivariados fue resuelto por Hotelling [Hot31] y explicado en [Hal52]. Sea

$$t_1 = \left(\frac{\bar{x} - \mu_1}{s_1} \right) n^{\frac{1}{2}} \quad t_2 = \left(\frac{\bar{y} - \mu_2}{s_2} \right) n^{\frac{1}{2}} \quad (36)$$

Hotelling descubrió, mediante métodos que caen fuera del alcance de este trabajo, que el punto (μ_1, μ_2) queda acotado por la desigualdad

$$t_1^2 - 2rt_1t_2 + t_2^2 \leq (1 - r^2)T^2 \quad (37)$$

Donde

$$T^2(\alpha) = 2 \frac{d-1}{d-2} F_{2,d-2}(\alpha) \quad (38)$$

$F_{2,d-2}(\alpha)$ representa el valor crítico F con 2 y $d-2$ grados de libertad y nivel de significación α .

Reemplazando en Eq.(36) el punto (μ_1, μ_2) por (x_1, y_1) obtenemos unas nuevas expresiones de t_1, t_2

$$t_1 = \left(\frac{\bar{x} - x_1}{s_1} \right) n^{\frac{1}{2}} \quad t_2 = \left(\frac{\bar{y} - y_2}{s_2} \right) n^{\frac{1}{2}} \quad (39)$$

Sustituyéndolas en Eq.(37)

$$\left(\frac{(x-\bar{x})^2}{s_1^2} - 2r \frac{(x-\bar{x})(y-\bar{y})}{s_1 s_2} + \frac{(y-\bar{y})^2}{s_2^2} \right) \leq \frac{T^2}{n} \quad (40)$$

y dividiendo por $1-r^2$

$$\frac{1}{1-r^2} \left(\frac{(x-\bar{x})^2}{s_1^2} - 2r \frac{(x-\bar{x})(y-\bar{y})}{s_1 s_2} + \frac{(y-\bar{y})^2}{s_2^2} \right) \leq \frac{T^2}{n} \quad (41)$$

Esta ecuación es semejante a Eq.(26) de la elipse estándar. Si eliminamos las fracciones en la Eq.(41) y la comparamos con la ecuación implícita de la elipse Eq.(28), los coeficientes A , B y C tienen los mismos valores que para la elipse estándar de la Eq.(29), siendo el coeficiente D

$$D = (1-r^2)s_1^2 s_2^2 d^{-1} T^2(\alpha) \quad (42)$$

La elipse de confianza tiene el mismo centro y el mismo valor de θ que la estándar, independiente de la variable D , como se especifica en Eq.(32). En las dos elipses, el eje principal coincide. Solamente los semiejes son variables.

Sean a_1, b_1 los semiejes cuando el parámetro $D = 1$, y obtenemos Eq.(30) para valores $D < > 0$

$$a = a_1 D^{\frac{1}{2}} \quad b = b_1 D^{\frac{1}{2}} \quad (43)$$

Puesto que a y b son proporcionales a $D^{\frac{1}{2}}$, obtenemos a partir de Eqs.(42)(43)

$$a = a_s d^{\frac{1}{2}} T(\alpha) \quad b = b_s d^{\frac{1}{2}} T(\alpha) \quad (44)$$

Donde a_s y b_s son los semiejes de la elipse estándar.

A partir de Eq.(28) podemos deducir que los puntos que se encuentran dentro de la región limitada por la elipse de confianza, satisfacen la desigualdad

$$A(x-\bar{x})^2 + 2B((x-\bar{x})(y-\bar{y})) + C(y-\bar{y})^2 < D \quad (45)$$

Si el origen cae dentro de la elipse, el centro de la población coincide con el origen, y la muestra no está direccionada. En este caso, la desigualdad Eq.

(45) se satisface siempre que $x = 0$ e $y = 0$. A partir de Eqs.(42)(45), se puede afirmar que la condición de la existencia de direccionamiento con un nivel de significación α es

$$T^2 > T^2(\alpha) \quad (46)$$

siendo el test estadístico

$$T^2 = \frac{d}{1-r^2} \left[\frac{\bar{x}^2}{s_1^2} - \frac{2r\bar{x}\bar{y}}{s_1s_2} + \frac{\bar{y}^2}{s_2^2} \right] \quad (47)$$

Al analizar la elipse estándar obtenida a partir de la (Fig.3.3.), observamos que es razonable asumir normalidad, por lo que, se puede aplicar la elipse de confianza. Para determinar la elipse de confianza con un nivel de significación de $\alpha = 5\%$ a partir de los vectores media representados en la Tabla 3.3., debemos calcular en primer lugar $T^2(0.05)$ usando Eq.(38).

De la tabla de las distribuciones F se obtiene $F_{2,4} = 6.94$, siendo por consiguiente $T^2(0.05) = 17.35$. A partir de Eqs.(42)(44) obtenemos los parámetros $D = 0.0007$, $a = 0.0221$ y $b = 0.038$ (Tabla 3.4). Se observa que la elipse de confianza (Fig.3.4.) es algo más grande que la elipse estándar (curva discontinúa).

Debido a que la elipse de confianza no contiene el origen, el centro de la población difiere significativamente de él. Por consiguiente el conjunto de vectores media está orientado en la dirección $\psi = 32.34^\circ$ (Eq.33).

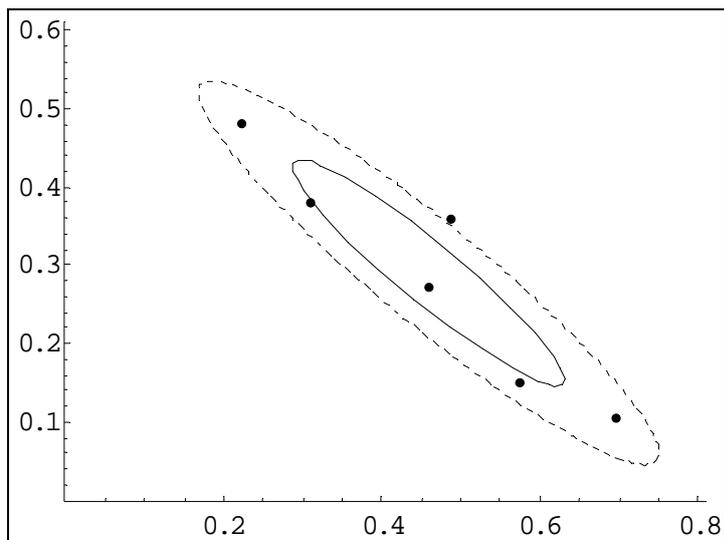


Figura 3.4. Elipse de confianza con un coeficiente $Q = 1 - \alpha = 95\%$. La curva punteada es la elipse estándar de la muestra

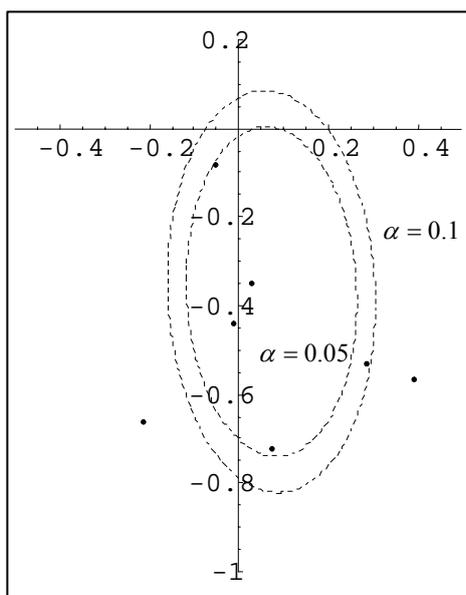


Figura 3.5. Elipses de confianza con un coeficiente $\alpha = 0.1$ y $\alpha = 0.05$

La Figura 3.5. muestra dos elipses de confianza calculadas con errores $\alpha = 0.1$ y $\alpha = 0.05$. Podemos observar que para $\alpha = 0.1$ el origen de coordenadas cae dentro de la elipse, por lo que no hay evidencia estadística de direccionamiento. Para $\alpha = 0.05$ sí podemos afirmar la existencia de direccionamiento.

3.5. Un ejemplo

La Fig.3.6. es una fotografía satélite del Estrecho de Gibraltar [JSC08], vista desde el vuelo STS-58 (STS o Sistema de Transporte Espacial es el nombre del programa global de la lanzadera espacial de la NASA). Pertenece a la colección de imágenes digitales del Centro Johnson de la NASA. El centro Johnson proporciona todas las imágenes catalogadas de las lanzaderas desde la primera misión de la lanzadera STS-1 hasta la misión ISS016 (Internacional Space Station) realizada conjuntamente con la estación rusa MIR en 2007. En la imagen se ve la bahía de Cádiz, el Estrecho de Gibraltar y la costa marroquí.

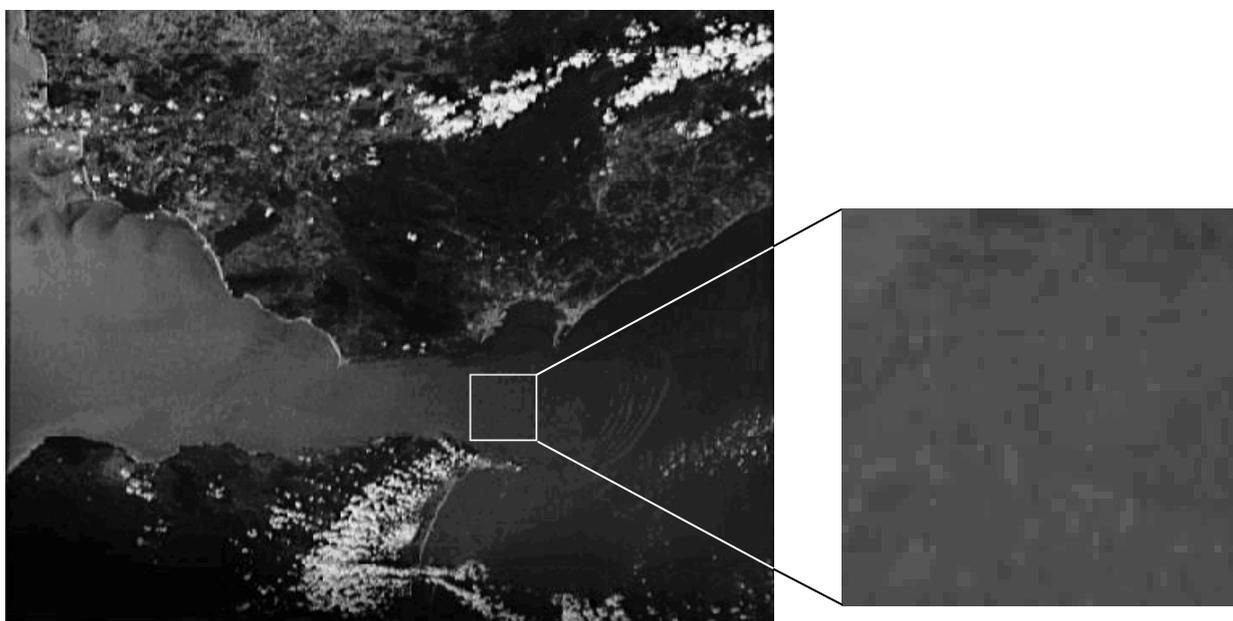


Figura 3.6. (a) Vista del Estrecho de Gibraltar desde el vuelo STS-58

(b) Ventana de 37x37 píxeles

Las corrientes del estrecho producen complicados patrones en la superficie del agua debido al reflejo del sol. Para estudiar la dirección de las corrientes, seleccionamos una ventana W de 27x27 píxeles a partir del centro de un rectángulo de 37x37 (Fig.3.6. (b)). La Tabla 3.5. muestra los valores de luz reflejados en la escala de grises 0-255, siendo W la ventana interior marcada con doble línea.

Para cada punto $x_{i,j}$ [$i=1..27, j=1..27$] de W obtenemos la submatriz $Z_{11,11}(x_{i,j})$, que contiene los píxeles de tamaño 5 alrededor de $x_{i,j}$. A partir de

$Z_{11,11}(x_{i,j})$ cogemos 5 muestras al azar N_1, \dots, N_5 y calculamos los vectores media $\bar{m}_1, \dots, \bar{m}_5$. Posteriormente, testeamos la existencia de anisotropía y cuantificamos la tendencia espacial del patrón en el caso de que exista evidencia estadística de direccionamiento.

Como ya hemos comentado, la anisotropía es una propiedad de los procesos estacionarios de segundo orden. Esto implica que el estudio solo se puede llevar a cabo bajo la asunción de estacionalidad en efectos locales. Por esta razón, antes de estudiar el direccionamiento, debemos comprobar el nivel de dependencia espacial del patrón, ya que si existe independencia espacial a nivel local, significaría que la interpretación de los resultados obtenidos no tendría ningún sentido. Para completar el análisis espacial de la imagen, examinaremos también la existencia/ausencia de homogeneidad a nivel global.

3.6. Análisis global de homogeneidad

Entre las técnicas más usadas para el estudio de la homogeneidad a nivel global, la más conocida y sencilla es el análisis de la varianza (ANOVA). Para utilizar el ANOVA es necesario asumir igualdad de varianzas (homocedasticidad). Si esto no se cumple, el ANOVA no es una prueba válida para testear la igualdad de medias en el patrón.

Asumiendo normalidad en la población de la que se han extraído los patrones, las varianzas deberían distribuirse según una distribución Chi-Cuadrado con $(11 \times 11) - 1 = 120$ grados de libertad. Los límites inferior y superior con una confianza del 95% son respectivamente 95.70 y 146.57.

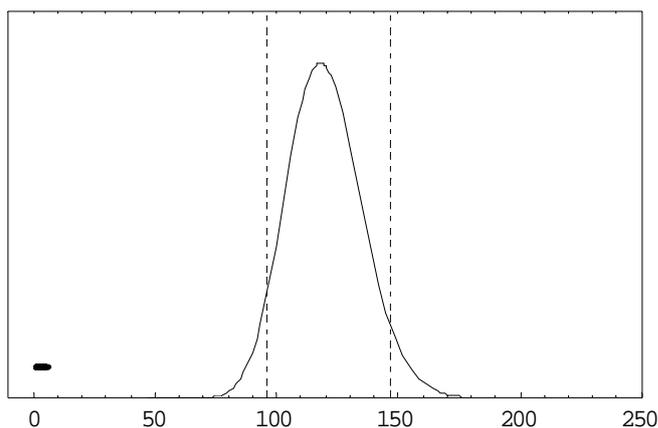


Figura 3.7. Límites superior e inferior para χ^2_{120} . Los estadísticos de prueba se encuentran cercanos a 0

Los valores de χ_{120}^2 menores que el límite inferior denotan patrones con baja varianza, y consecuentemente con una dependencia espacial fuerte, mientras que los valores por encima del límite superior indican una alta variabilidad o heterogeneidad. Los estadísticos han sido calculados estimando la varianza de la población (324.96) como la varianza de una muestra de gran tamaño de pixeles extraída de la imagen. Todos los estadísticos son inferiores a 95.70 (Fig.3.6.), por lo que podemos asumir estacionalidad a nivel global.

3.7. Análisis de dependencia espacial

El análisis de la dependencia espacial, se realiza examinando la estructura de la covarianza espacial del patrón. La técnica más usada comúnmente es el test de autocorrelación basado en algún estadístico diseñado para ello, tal como los multiplicadores de Lagrange, *Moran's I*, *Geary's C* o *Getis's G* [Cho95]. *Moran's I* es el más usado. Para una matriz $Z_{n,n}$, este estadístico a una distancia h se estima como

$$I(h) = \frac{n \sum_{i=1}^n \sum_{j=1}^n \delta_{i,j}^{(h)} (z_i - \bar{z})(z_j - \bar{z})}{\sum_{i=1}^n (z_i - \bar{z})^2 \sum_{i \neq j} \delta_{i,j}^{(h)}} \quad (48)$$

El valor esperado $E(I)$ y la varianza $Var(I)$ son

$$E(I) = -\frac{1}{n-1} \quad Var(I) = \frac{n^2 S_1 - n S_2 + 3 S_0^2}{S_0^2 (n^2 - 1)} \quad (49)$$

Donde

$$S_0 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\delta_{i,j} + \delta_{j,i})^2 \quad S_1 = \sum_{i=1}^n \left(\sum_{i=1}^n \delta_{i,j} + \sum_{j=1}^n \delta_{j,i} \right)^2 \quad (50)$$

En Eq.(48), $\delta_{i,j}^{(h)}$ representa la correlación espacial entre las celdas i -ésima y j -ésima. Aunque los coeficientes de correlación se restringen normalmente al rango (-1,1), utilizando el estadístico de *Moran's I* no se consigue. Para lograr

que se ajuste al rango (-1,1), debemos dividirlo por un factor de corrección $c^{(h)}$:

$$c^{(h)} = \frac{n}{\sum_{i \neq j} \sum \delta_{i,j}^{(h)}} \left(\frac{\sum_{i=1}^n \left(\sum_{j=1}^n \delta_{i,j}^{(h)} (z_i - \bar{z}) \right)^2}{\sum_{i=1}^n (z_i - \bar{z})^2} \right)^{\frac{1}{2}} \quad (51)$$

Los patrones agrupados se asocian a valores altos y positivos. Por el contrario un valor negativo alejado de cero implica un patrón disperso. Los valores cercanos a cero representan ausencia de dependencia espacial por lo que se asocian a patrones espacialmente aleatorios. Si observamos la Eq.(48), podemos comprobar que I no es una buena medida para patrones en los que todas las celdas toman el mismo valor, ya que resultaría una indeterminación del tipo $\frac{0}{0}$.

Una práctica común utilizada para explorar la dependencia espacial es estimar la correlación espacial a diferentes distancias para la construcción de un correlograma en donde $I^{(h)}$ se representa contra varios valores de h .

La observación de este correlograma, y especialmente sus picos, permiten obtener la información sobre el nivel y la dirección de la autocorrelación, así como de las distancias en las que se alcanzan sus valores máximos y mínimos.

Otro método para analizar el nivel de dependencia espacial del patrón es calcular $I^{(1)}$ a partir de las matrices $\{Z_{2h+1,2h+1}(x_{i,j}, h), h = 1..5\}$, donde $Z_{2h+1,2h+1}(x_{i,j}, h)$ es la matriz que contiene los píxeles de distancia h alrededor de $x_{i,j}$. Para cada punto $x_{i,j}$, obtenemos 5 índices $\{I^{(1)}(x_{i,j}, h), h = 1..5\}$ de autocorrelación.

La Tabla 3.6. muestra estos índices para la ventana $\{x_{i,j}, i = 2..17, j = 21..26\}$ (rectángulo discontinuo de la Tabla 3.5.). Cada uno de estos grupos de cinco

índices se consigue después de eliminar las indeterminaciones obtenidas de la función lineal $\hat{I} = \beta_1 + \beta_2 x$. Los valores máximos β_2 para cada punto de W se representan en la Tabla 3.7.

-.0130	-.0004	.0016	.0036	-.0033	-.0058	-.0096	-.0038	.0020	-.0055	.0100	-.0110	.0100	.0097	-.0029	.0031
.0024	.0066	.0110	.0120	.0140	.0120	.0120	.0070	-.0009	.0110	.0150	.0110	.0110	.0120	.0016	.0038
.0060	.0087	.0110	.0130	.0100	.0090	.0085	.0076	.0065	.0080	.0091	.0079	.0053	.0037	.0074	.0068
.0074	.0078	.0093	.0087	.0085	.0067	.0061	.0062	.0076	.0071	.0068	.0070	.0063	.0073	.0080	.0110
.0082	.0087	.0079	.0080	.0081	.0081	.0059	.0062	.0073	.0077	.0077	.0073	.0083	.0086	.0095	.0110
-.0074	.0024	Indet.	Indet.	-.0009	.0016	-.0038	.0038	-.0073	-.0029	.0120	-.0069	.0100	.0130	-.0150	-.0034
.0008	.0023	.0066	.0002	-.0003	-.0001	-.0009	-.0018	-.0009	.0062	.0120	.0097	.0095	.0110	.0009	.0023
.0034	.0062	.0084	.0098	.0074	.0075	.0068	.0055	.0058	.0087	.0099	.0090	.0064	.0027	.0042	.0024
.0062	.0068	.0084	.0076	.0073	.0062	.0059	.0058	.0073	.0063	.0057	.0054	.0040	.0049	.0055	.0100
.0068	.0071	.0060	.0061	.0060	.0063	.0050	.0056	.0057	.0061	.0052	.0044	.0055	.0059	.0075	.0095
-.0230	.0017	Indet.	Indet.	Indet.	Indet.	.0016	-.0073	-.0007	-.0008	.0030	-.0034	0	.0093	-.0072	-.0150
.0021	.0021	.0047	0	.0018	.0034	.0047	.0028	-.0009	.0033	.0095	.0077	.0110	.0110	.0029	.0021
.0031	.0044	.0036	.0036	.0014	.0008	-.0001	.0006	.0030	.0072	.0065	.0071	.0053	.0032	.0050	.0019
.0056	.0055	.0066	.0061	.0061	.0050	.0058	.0054	.0072	.0064	.0063	.0055	.0037	.0034	.0025	.0069
.0059	.0061	.0049	.0049	.0053	.0060	.0050	.0063	.0058	.0061	.0048	.0033	.0043	.0044	.0066	.0092
-.0110	.0047	-.0031	-.0008	Indet.	Indet.	.0016	-.0067	-.0029	.0016	-.0009	-.0007	-.0008	-.0017	-.006	-.0058
.0069	.0092	.0097	.0064	.0051	.0009	.0016	.0001	.0028	.0009	.0050	.0047	.0042	.0052	.0017	-.0005
.0064	.0063	.0062	.0072	.0055	.0046	.0021	.0050	.0050	.0059	.0044	.0069	.0056	.0033	.0044	.0016
.0042	.0026	.0026	.0028	.0032	.0027	.0044	.0040	.0067	.0058	.0063	.0050	.0035	.0039	.0022	.0064
.0052	.0050	.0039	.0040	.0049	.0059	.0049	.0064	.0061	.0061	.0056	.0041	.0042	.0041	.0058	.0078
0	.0087	.0031	.0058	-.0008	Indet.	-.0008	-.0031	-.0029	.0038	-.0009	-.003	.0057	-.0008	Indet.	-.0017
.0070	.0110	.0100	.0091	.0095	.0018	-.0002	.0003	.0044	.0023	.0062	.0073	.0035	.0035	-.0006	-.0030
.0066	.0065	.0066	.0073	.0044	.0040	.0044	.0086	.0049	.0073	.0064	.0073	.0063	.0009	.0026	-.0013
.0042	.0044	.0045	.0043	.0047	.0046	.0065	.0055	.0076	.0069	.0067	.0059	.0048	.0048	.0036	.0067
.0050	.0044	.0035	.0035	.0045	.0057	.0045	.0058	.0057	.0058	.0055	.0047	.0045	.0045	.0058	.0074
-.0059	.0031	-.0050	.0052	.0016	Indet.	-.0008	-.0031	-.0012	-.0034	.0060	.0026	.0026	.0025	-.0017	-.0060
.0021	.0062	.0077	.0069	.0086	.0022	.0027	.0071	.0091	.0055	.0130	.0130	.0100	.0110	-.0027	-.0039
.0053	.0055	.0061	.0066	.0040	.0031	.0053	.0081	.0060	.0086	.0079	.0088	.0092	.0029	.0041	.0022
.0062	.0074	.0073	.0068	.0063	.0060	.0076	.0061	.0072	.0062	.0067	.0074	.0060	.0054	.0029	.0053
.0056	.0059	.0056	.0054	.0062	.0073	.0060	.0070	.0067	.0068	.0059	.0044	.0041	.0043	.0056	.0070

Tabla 3.6. Índices de Moran's $\{I(x_{i,j}, h), h = 1..5\}$ para la ventana comprendida entre los puntos $\{x_{i,j}, i = 2..17, j = 21..26\}$ del rectángulo de la Tabla 3.5.

AQUI VA LA TABLA 3.7 DE A3 EN LA PÁGINA 57 DEL DOCUMENTO

AQUI VA LA TABLA 3.7 DE A3 EN LA PÁGINA 57 DEL DOCUMENTO

A partir de Eq.(49) obtenemos los índices esperados $\{E(I(h)), h = 1..5\} = \left\{\frac{-1}{8}, \frac{-1}{24}, \frac{-1}{48}, \frac{-1}{80}, \frac{-1}{120}\right\}$. La Tabla 3.8. muestra los parámetros del ajuste lineal de estos valores y de sus intervalos de confianza calculados con una significación de $\alpha = 5\%$.

De las 729 celdas (Tabla 3.5.), solamente 24 de ellas (marcadas como el símbolo “=” en Fig.3.9.) se pueden considerar espacialmente independientes, ya que presentan pendientes con valores fuera del intervalo de confianza.

Parámetro	Estimador	Error Estándar	Intervalo de Confianza
1	-.1204	.0300	(-.0216, -.0250)
x	.0262	.0090	(-.0025, .0550)

Tabla 3.8. Los parámetros de ajuste para los índices $\{E(I(k)), k = 1..5\} = \left\{\frac{-1}{8}, \frac{-1}{24}, \frac{-1}{48}, \frac{-1}{80}, \frac{-1}{120}\right\}$ de la función lineal $\hat{I} = \beta_1 + \beta_2 x$

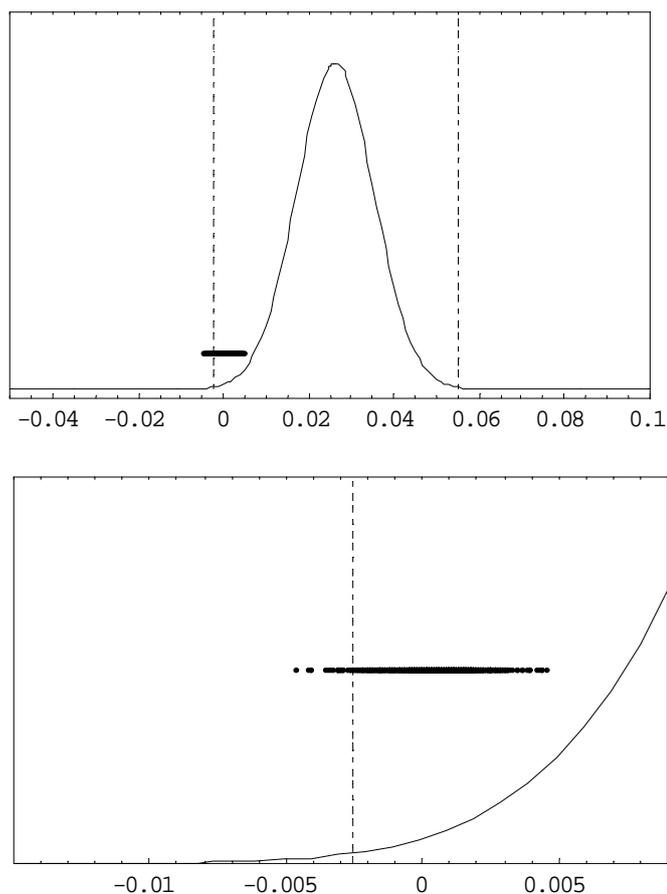


Figura 3.8. Límites superior e inferior del intervalo de confianza. El estadístico de prueba de las celdas espacialmente independientes se encuentra por debajo del límite inferior

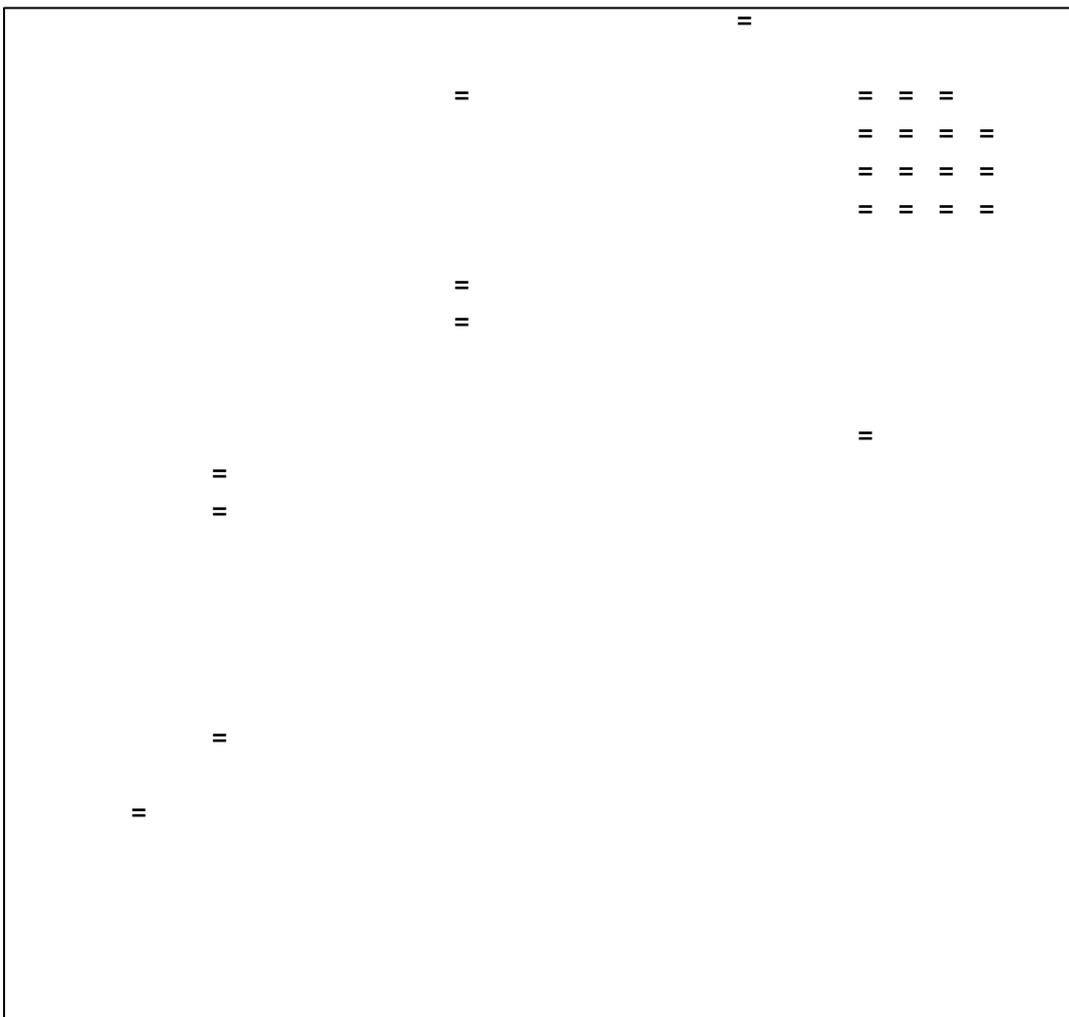


Figura 3.9. Las celdas marcadas con el símbolo “ = ” presentan homogeneidad a nivel global pero no dependencia espacial. El resto sí son celdas espacialmente dependientes

3.8. Análisis de direccionamiento

Como se comentó en puntos anteriores, el primer paso para analizar el direccionamiento es comprobar si hay evidencia estadística de un comportamiento anisotrópico. Para cada $Z_{11,11}(x_{i,j})$ de W extraemos 5 muestras aleatorias N_1, \dots, N_5 y calculamos los vectores media $\{\overline{m}_i, i = 1, \dots, 5\}$ aplicando Eqs.(6-7). En la Tabla 3.9. se muestra la longitud de r y el ángulo media $\overline{\Phi}$ de las celdas que pertenecen al rectángulo remarcado de la Tabla 3.5. Una vez obtenido los vectores media de cada una de las celdas de la ventana W , determinamos los parámetros de la elipse de confianza que nos permite obtener el estadístico T^2 (Eq.(46)) y aplicamos la condición Eq.(46) para testear la anisotropía.

Todas las elipses de confianza han sido calculadas con $d = 5$ vectores, por lo que, el valor crítico $T^2(0.05)$ (Eq.(38)) para todas las celdas es 25.47. La Tabla 3.10. muestra los valores de los estadísticos T^2 para W . La dirección en la que la anisotropía aparece para cada punto viene dada por ψ (Eq.(33)), y se muestra gráficamente en la Fig.3.11.

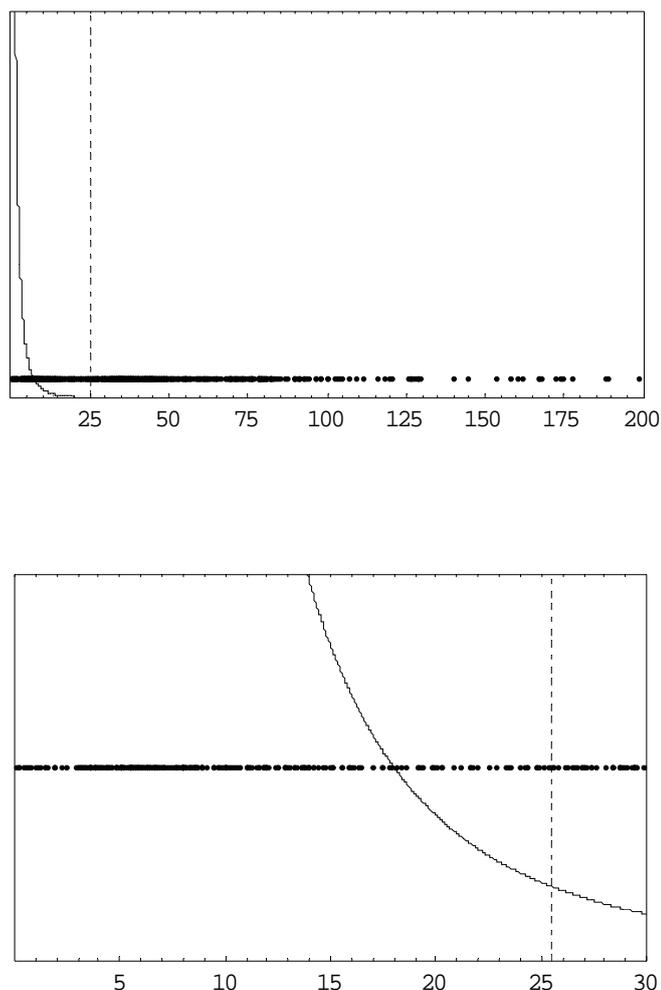


Figura 3.10. Los celdas con valores de T^2 inferiores a $T^2(0.05) = 24.75$ tienen un comportamiento isotrópico. El resto son celdas con un comportamiento direccional

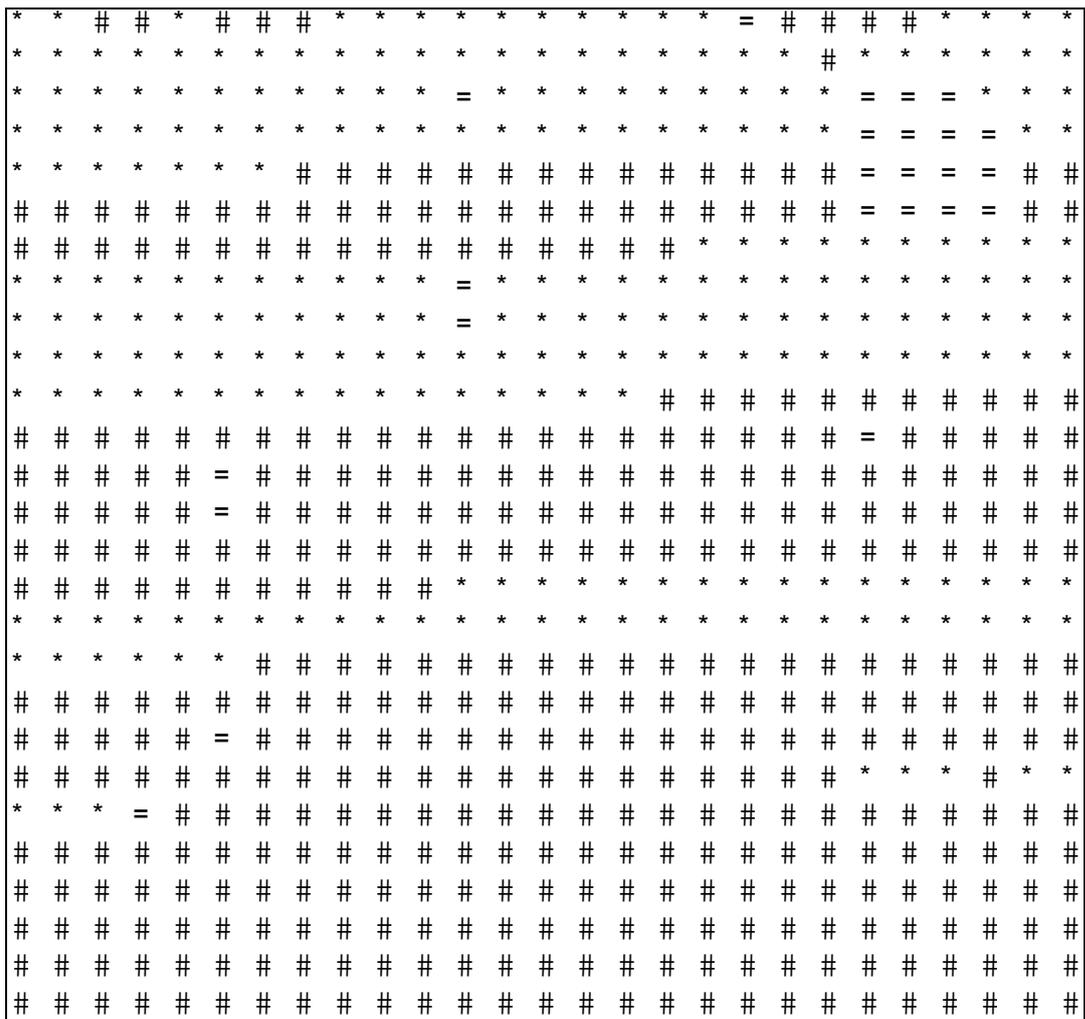


Figura 3.11. Comportamiento de las celdas de W : celdas espacialmente independientes (“=”), celdas isotrópicas (“*”) y celdas anisotrópicas (“#”)

En la Fig.3.11. se ha representado la existencia de homogeneidad a nivel global, la dependencia espacial local y la anisotropía. Analizándola, podemos observar que solamente 24 de las 729 celdas (marcadas con “=”) presentan homogeneidad a nivel global y ausencia de dependencia espacial a nivel local. El resto de las celdas presenta homogeneidad global y dependencia espacial local. En cuanto al direccionamiento, observamos la existencia de un grupo de celdas que muestran anisotropía (marcadas con “#”), nos estamos refiriendo a los dos tercios inferiores de la figura. La parte superior representa tres grupos bien definidos, dos de ellos presentan isotropía (marcadas con “*”) y el tercero direccionamiento.

Si examinamos las direcciones ψ representadas en la Fig.3.12., observamos variaciones en cuanto a la dirección, habiendo una tendencia claramente anisotrópica en la dirección oeste-este. Esto es debido a las corrientes marinas que produce el Atlántico al atravesar el Estrecho de Gibraltar hacia el mar Mediterráneo. Estas corrientes generan olas internas representadas a través de diferentes intensidades de luz (Tabla 3.5.) [MolAgu00].

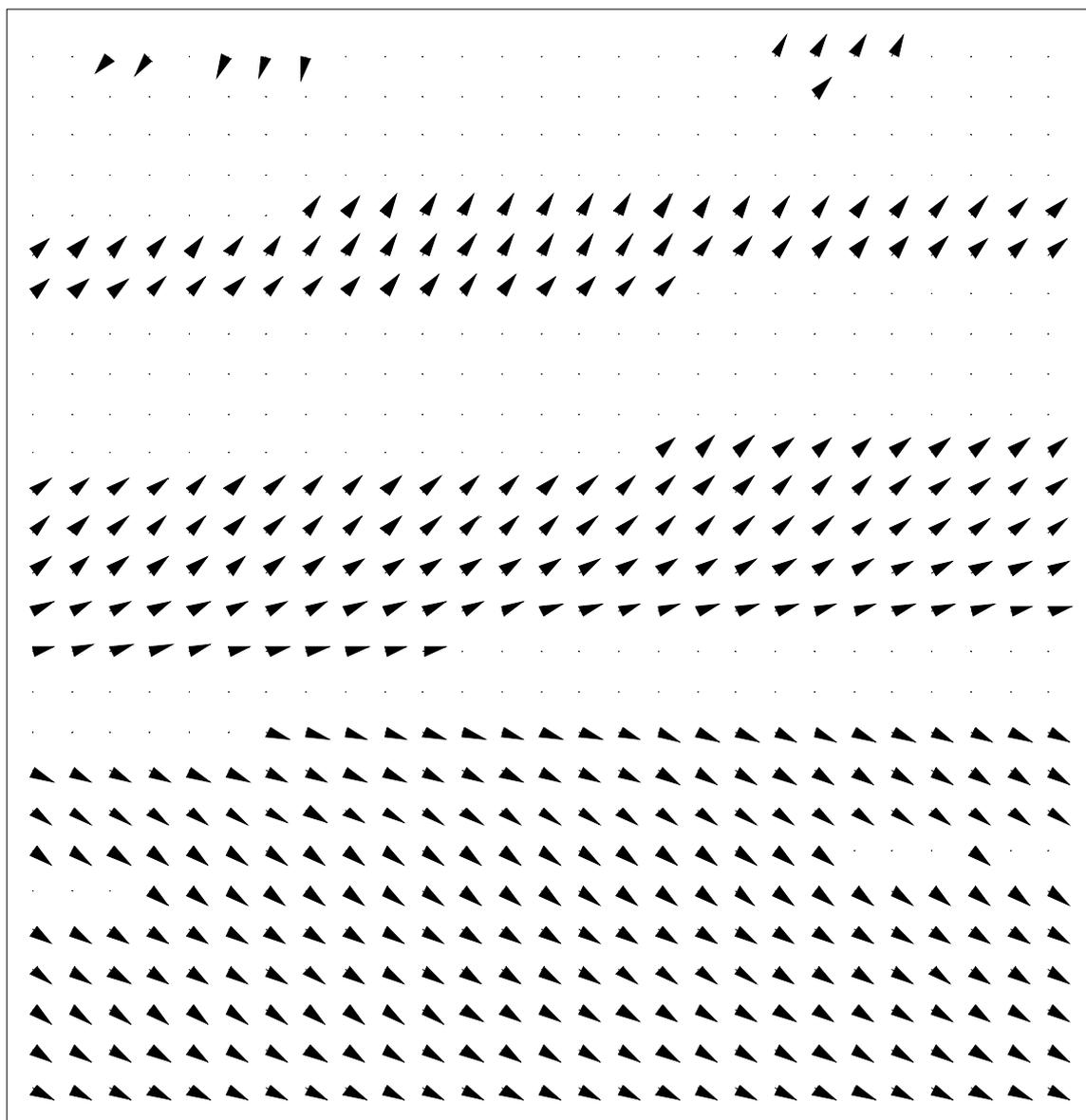


Figura 3.12. Dirección ψ en la que se muestra la anisotropía para cada una de las celdas de W calculada con $d = 5$ vectores

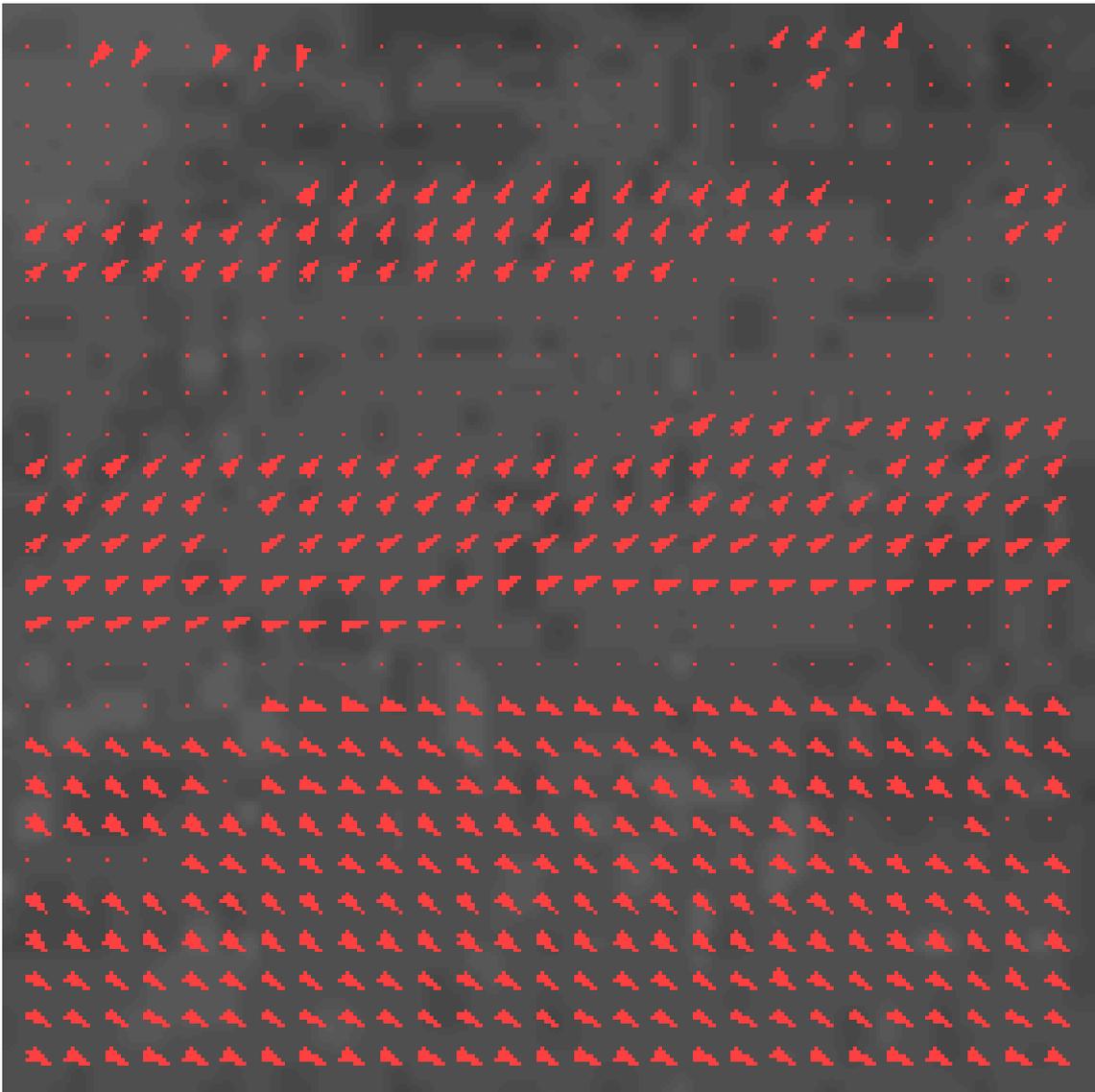


Figura 3.13. Dirección ψ en la que se muestra la anisotropía superpuesta a la imagen de W

AQUÍ VA LA TABLA 3.9 DE A3 PAGINA 67 DEL DOCUMENTO

AQUÍ VA LA TABLA 3.10 DE A3 PAGINA 69 DEL DOCUMENTO

Capítulo IV

Operador Buffer

La mayoría de los SIG conocidos disponen de uno o varios módulos que generan áreas de influencia isotrópica, pero muchos de los fenómenos observados son direccionales, por lo que se requieren operadores capaces de modelar comportamientos anisotrópicos. En este capítulo presentamos un método para generar óvalos en los que la superficie se propaga según una distribución Normal alrededor de su dirección dominante. Para ello hemos utilizado la función de distribución de Von Misses modificada, lo que nos permitirá identificar el concepto de probabilidad con el de superficie ocupada por la propagación alrededor de un punto.

4.1. Operador buffer

4.1.1. Definición

El término buffer, también denominado buffering o zona de influencia, se define de forma clásica como el lugar geométrico de los puntos que se encuentran a una distancia menor o igual a un objeto (punto, polilínea o polígono) dado. En el estudio del buffer se distinguen los siguientes casos:

- Creación de áreas de influencia de datos puntuales.
- Creación de áreas de influencia de datos lineales.
- Creación de áreas de influencia de datos poligonales.

4.1.2. Creación de áreas de influencia de datos puntuales

Con frecuencia los SIG requieren la generación de áreas de influencia en determinadas operaciones de análisis espacial de datos. La forma más simple de área de influencia es la que se genera en torno a datos puntuales, ya que el proceso implica tan solo la creación de un polígono “circular” alrededor de cada punto, de radio equivalente a la distancia del buffer [Buf08]. Hay dos formas de asignar la anchura del buffer, la primera aplica una distancia de buffer fija (especificada por el usuario) para todos los puntos de una capa.

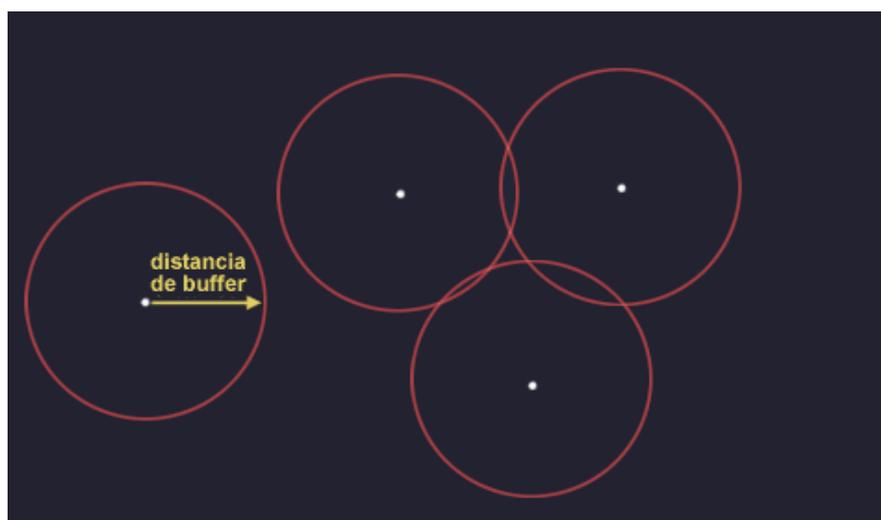


Figura 4.1. Áreas de influencia de datos puntuales con distancia fija

La segunda consiste en asignar a cada punto un valor de distancia individual basado en los atributos de otra capa del sistema (la ponderación o peso). Los atributos de anchura del área de influencia en este caso quedaran almacenados en una tabla.

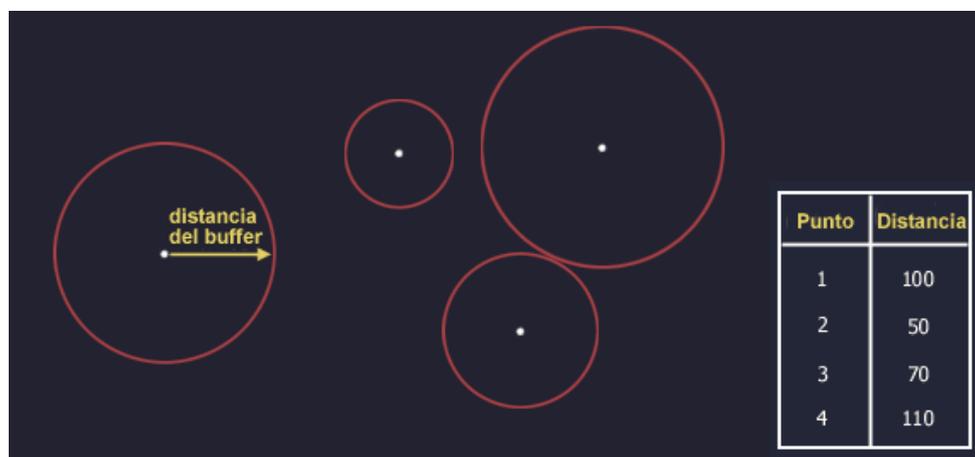


Figura 4.2. Áreas de influencia de datos puntuales con atributos de capa

En el caso de existir múltiples puntos en la capa a analizar, el sistema debe comprobar la existencia de solapamientos entre las áreas de cada punto. Estos solapamientos deben ser borrados de forma que el resultado sea una capa poligonal que represente la zona cubierta por la unión de todas las áreas. Por lo que este procedimiento, implica la aplicación de dos operaciones adicionales: la intersección y la disolución (*dissolve*).

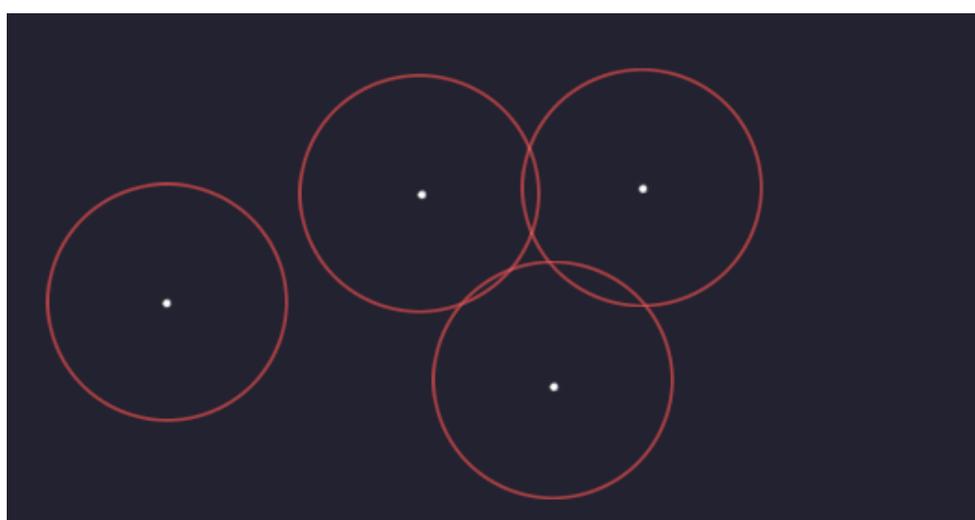


Figura 4.3. Cálculo de la intersección de áreas de influencia

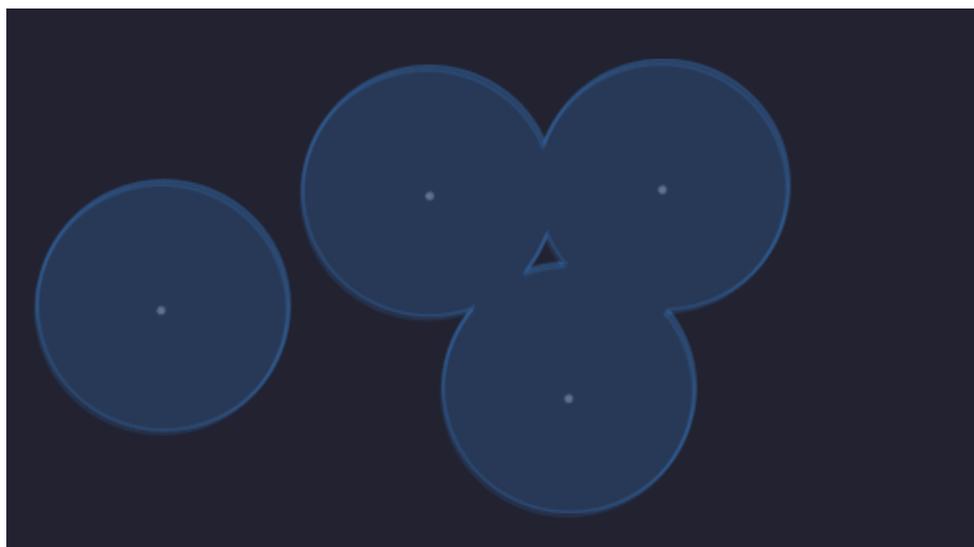


Figura 4.4. Cálculo de la disolución de áreas de influencia de datos puntuales

La creación de buffers da como resultado una nueva capa de tipo poligonal en el sistema, que representa las zonas de influencia generadas a partir de valores de distancia tanto fijos como ponderados. La tabla de polígonos resultante contendrá los identificadores de los polígonos creados en el procedimiento, y un nuevo atributo que indica si el polígono se encuentra dentro o fuera del área de influencia.

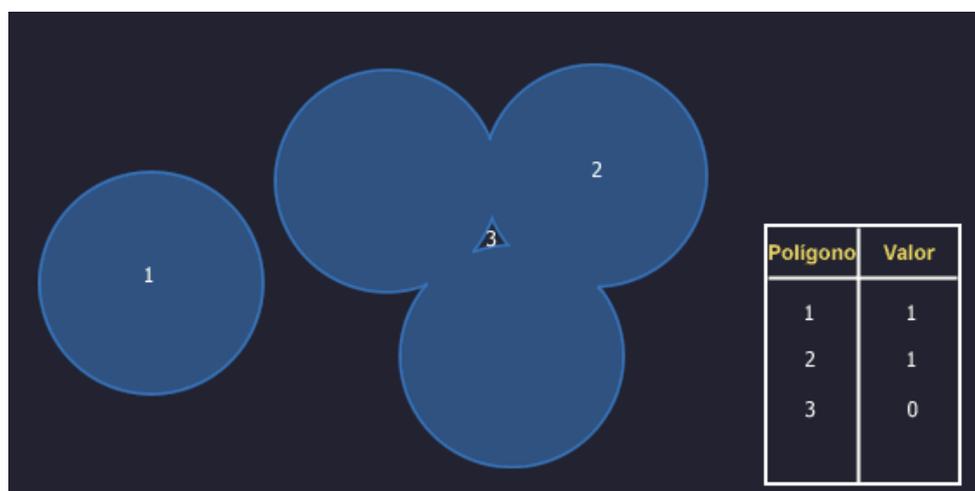


Figura 4.5. Cálculo de la disolución de áreas de influencia de datos puntuales con atributos de capa

4.1.3. Creación de áreas de influencia de datos lineales

El algoritmo en el caso de datos de tipo lineal es más complejo que el de datos puntuales, ya que las líneas se pueden componer de múltiples segmentos.

El proceso de generación es el siguiente:

El primer paso consiste en asignar a cada segmento de línea la anchura correspondiente del área de influencia (la cual podría ser fija para todas las líneas o ponderada). A esta distancia la denominaremos b .

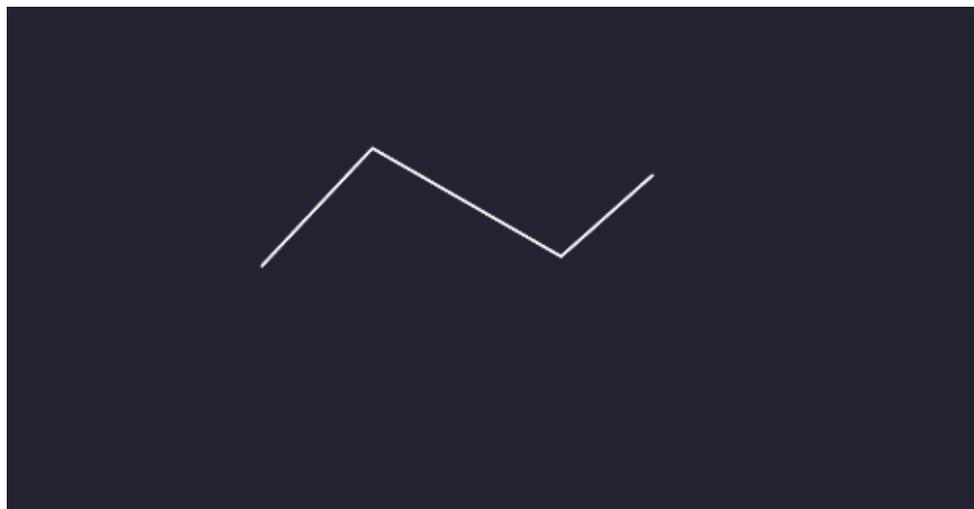


Figura 4.6. Creación de áreas de influencia de datos lineales

Los segmentos de la línea se procesan de forma individual. Cada segmento de la línea tiene un nodo inicial (E_1, N_1) y un nodo final (E_2, N_2) .

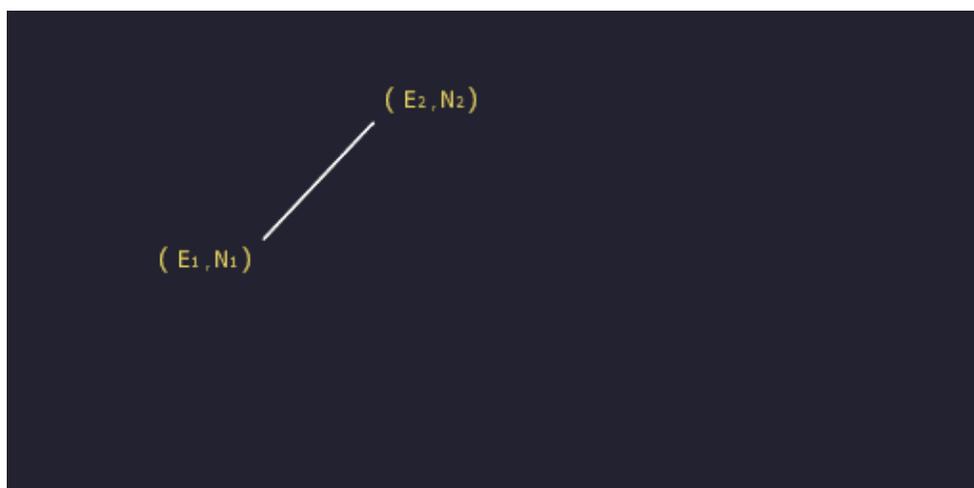


Figura 4.7. Segmento inicial

A partir de estas coordenadas se calculan los valores Δx y Δy entre los dos extremos.

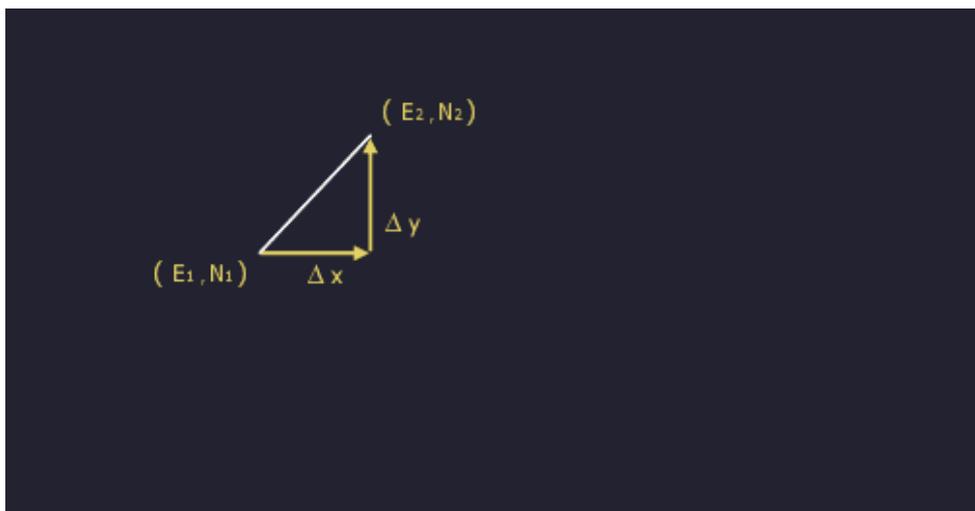


Figura 4.8. Cálculo de las coordenadas en el eje de abscisas y ordenadas

Las coordenadas de los puntos extremos de las líneas paralelas a crear (a una distancia perpendicular b), se calculan a partir de las fórmulas que se muestran a continuación:

$$\begin{aligned}
 E2 \pm b \cdot \text{Sen}[\text{Tan}^{-1}(\Delta x / \Delta y)] \\
 N2 \pm b \cdot \text{Cos}[\text{Tan}^{-1}(\Delta x / \Delta y)]
 \end{aligned}
 \tag{52}$$

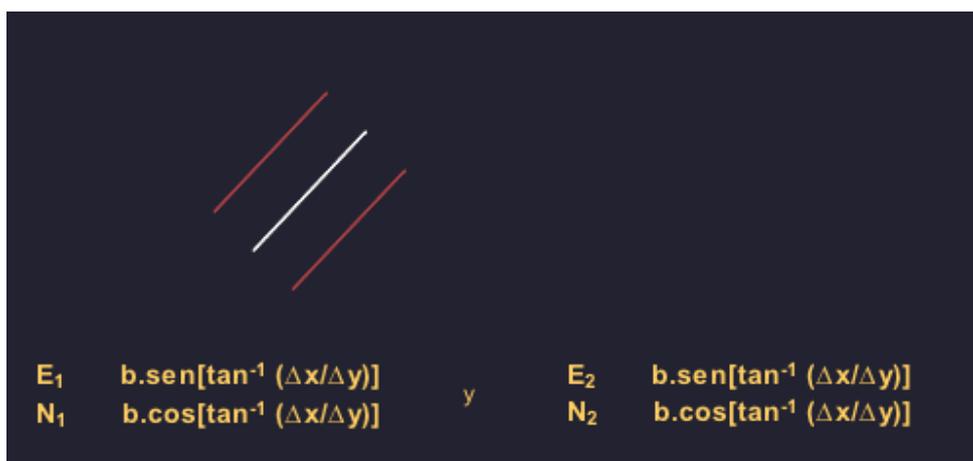


Figura 4.9. Representación de las coordenadas de los puntos extremos

En el caso de que la línea sea horizontal o vertical respecto al sistema de ejes, las fórmulas se simplifican notablemente. Llegado a este punto, el sistema examina el valor de la tangente al segmento de línea para determinar si se

producen incrementos crecientes o decrecientes en el eje de abscisas u ordenadas en función de la dirección de la línea.

Una vez que se han identificado las dos nuevas líneas paralelas de la futura área de influencia, se puede procesar de la misma manera el siguiente segmento.

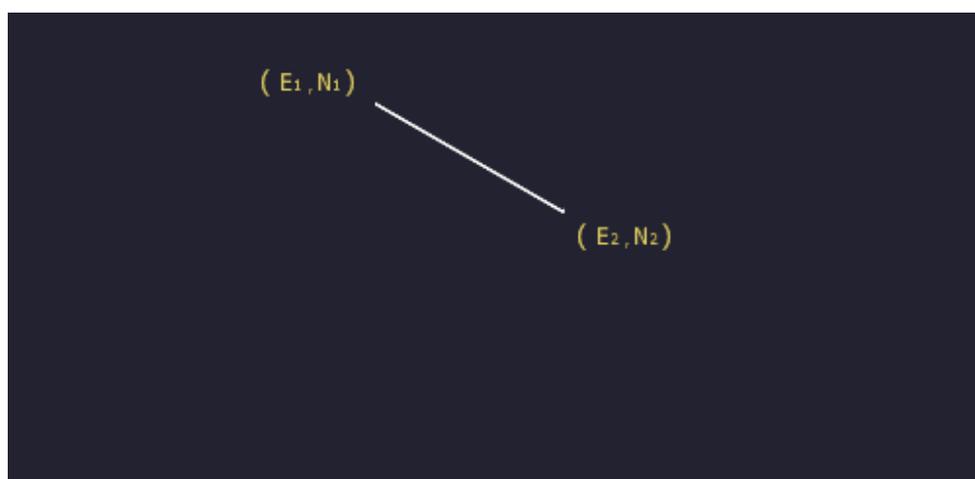


Figura 4.10. Segundo segmento

Determinadas las líneas paralelas al siguiente segmento, se calculan las intersecciones entre las líneas paralelas trazadas para cada segmento y se asignan nuevas coordenadas a los vértices comunes.

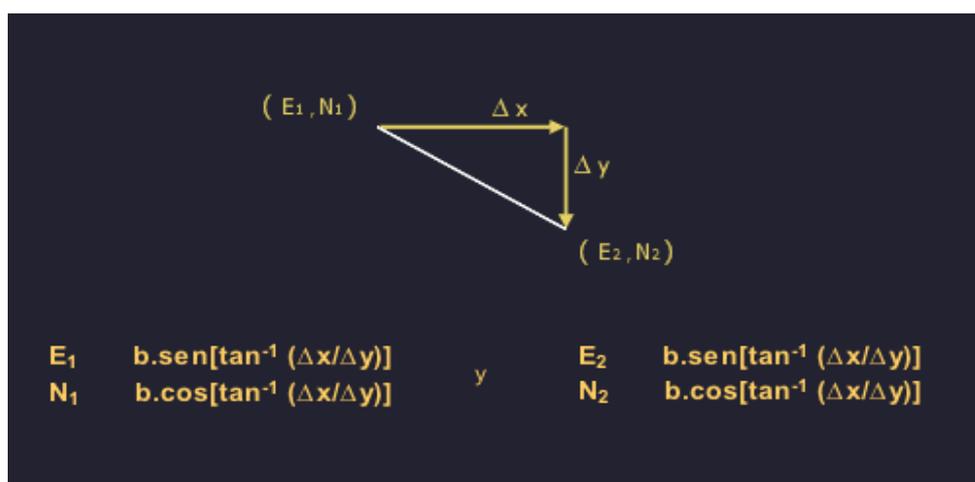


Figura 4.11. Cálculo de las coordenadas del segundo segmento

Estos pasos se repiten hasta que se procesa el último de los segmentos de la línea inicial.

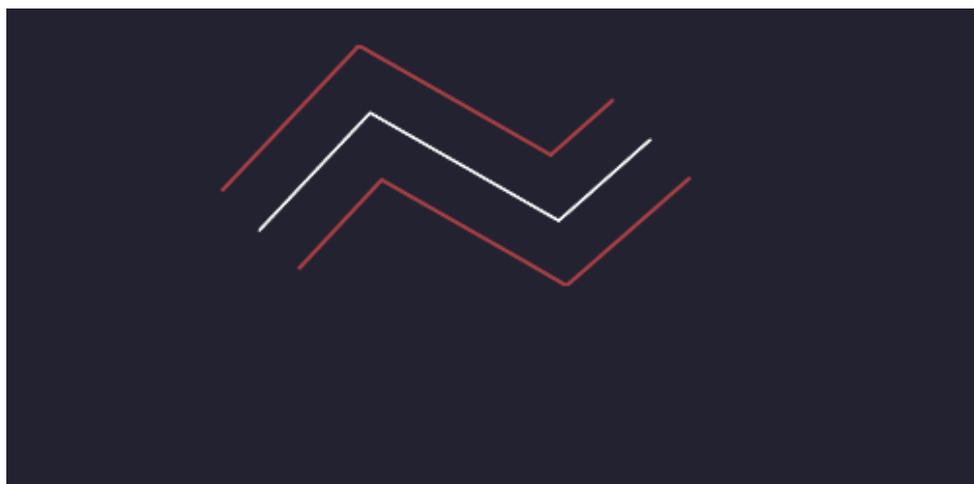


Figura 4.12. Último segmento

La última fase consiste en la definición de la forma de los extremos del área de influencia. Los diferentes SIG definen los extremos finales de un área de influencia de forma distinta. Para ello se aplican algunos de los siguientes criterios:

- Un sencillo truncamiento de los extremos finales de las líneas paralelas.
- Rematar en punta los extremos de las líneas paralelas para hacerlos coincidir.
- Rematar el punto de inicio y fin de las líneas paralelas con un polígono en forma de semicírculo de radio b .

En Fig.4.14. se ha utilizado el último método.

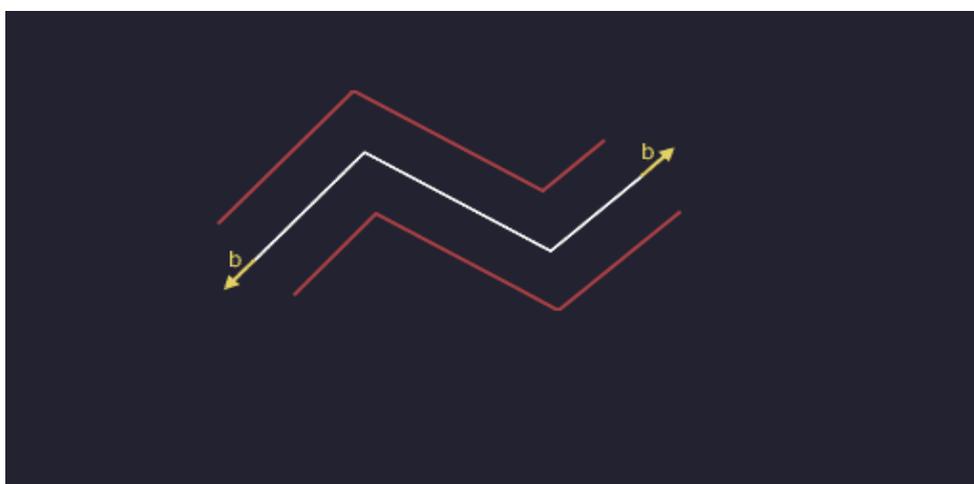


Figura 4.13. Creación de los extremos del área de influencia

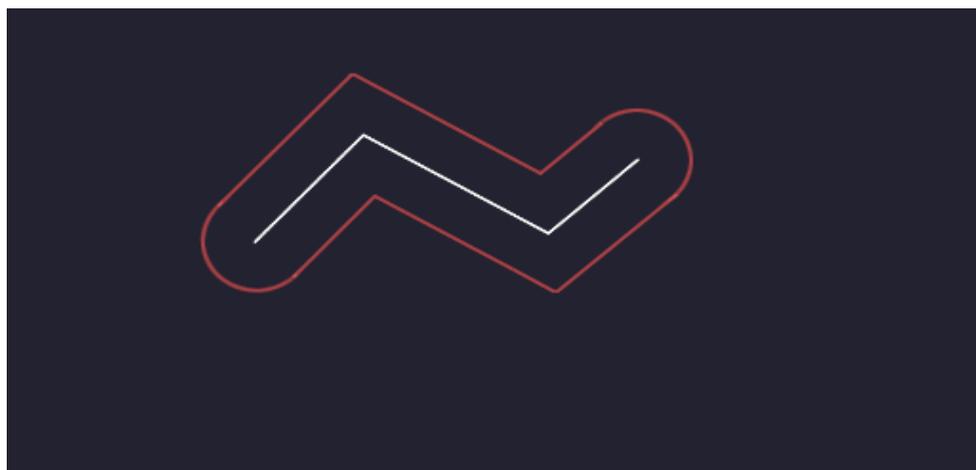


Figura 4.14. Remate de los extremos del área de influencia

Si existen varias líneas en la capa en la que se van a crear las áreas de influencia, el sistema debe comprobar la existencia de solapamientos entre los buffers de cada línea.

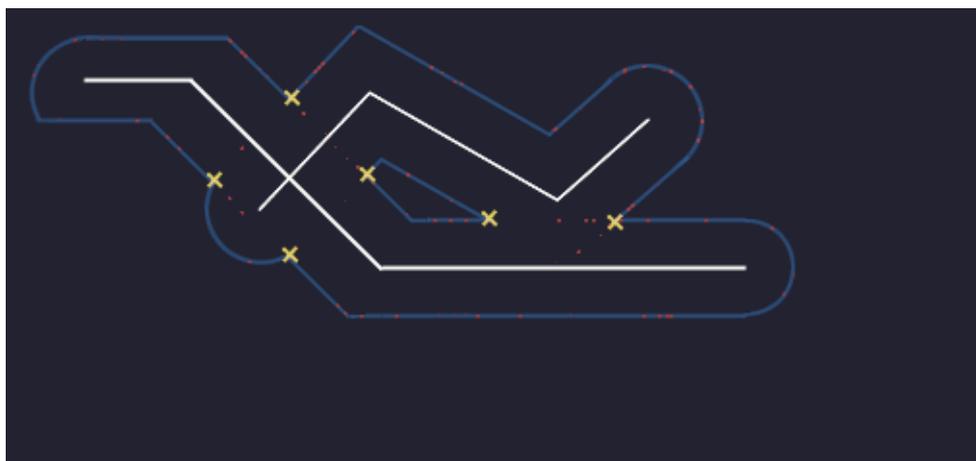


Figura 4.15. Cálculo del solapamiento entre varias áreas de influencia

Se deben borrar todas las regiones que se solapan, de forma que el resultado sea una capa poligonal que represente la zona cubierta por la unión de todas las áreas de influencia. En este procedimiento también se tienen que aplicar las operaciones de intersección y la disolución.

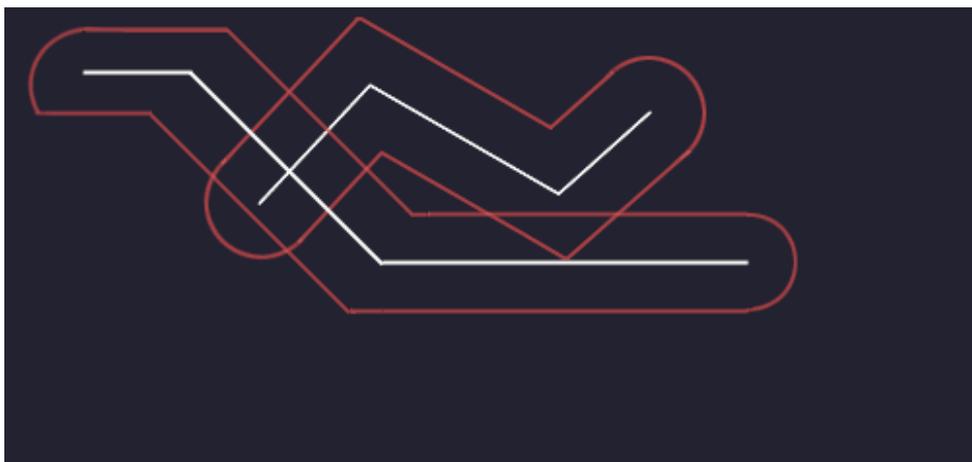


Figura 4.16. Intersección de las áreas de influencia



Figura 4.17. Disolución de las áreas de influencia

La creación de áreas de influencia da como resultado una nueva capa de tipo poligonal en el sistema, que representa las zonas de influencia generadas a partir de valores de distancia tanto fijos como ponderados. La tabla de polígonos resultante tendrá los identificadores de los polígonos creados en el procedimiento, y un nuevo atributo que indica si el polígono se encuentra dentro o fuera del área de influencia. En este ejemplo, el 1 indica que el polígono está dentro de dicha área y el 0 que está fuera de la misma.

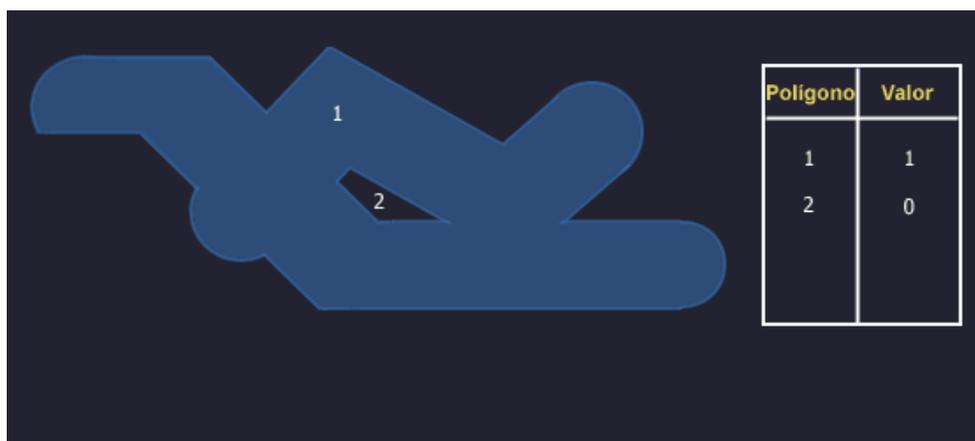


Figura 4.18. Disolución de áreas de influencia de datos poligonales con atributos de capa

4.1.4. Creación de áreas de influencia de datos poligonales

El algoritmo para la creación de áreas de influencia en torno a polígonos utiliza el mismo método que en el caso de datos lineales, con una pequeña diferencia: el polígono buffer resultante se crea solo hacia uno de los lados que definen el polígono original (hacia fuera o hacia dentro).

Por defecto se genera el área de influencia alrededor de los límites que definen el polígono, algunos SIG ofrecen también la posibilidad de crear el buffer hacia el interior del polígono problema. La Fig.4.19. muestra un ejemplo de las 2 posibles áreas de influencia de un polígono.



Figura 4.19. Polígono original, área de influencia exterior, área de influencia interior

Para explicar este último método partimos del siguiente supuesto:

A partir del plano de un campus universitario, se necesita determinar la localización de zonas de seguridad para reunir al personal evacuado en caso de incendio. La imagen siguiente representa las capas de terreno del campus y la ubicación actual de los edificios (representadas en verde).



Figura 4.20. Campus universitario

Las zonas de seguridad deben estar situadas al menos a 20 metros de los edificios del campus. Con el fin de encontrar los lugares donde se encuentra la zona de seguridad contra incendios, se crea una capa con las áreas de influencia generadas a partir de la situación de los edificios del campus a una distancia fija de 20 metros (Fig.4.21. (a)).



Figura 4.21. Creación de áreas de influencia poligonales

Como se ve en la figura esto produce algunos solapamientos. Por lo que posteriormente es necesario aplicar las operaciones de intersección y disolución para fusionar estas zonas de intersección.

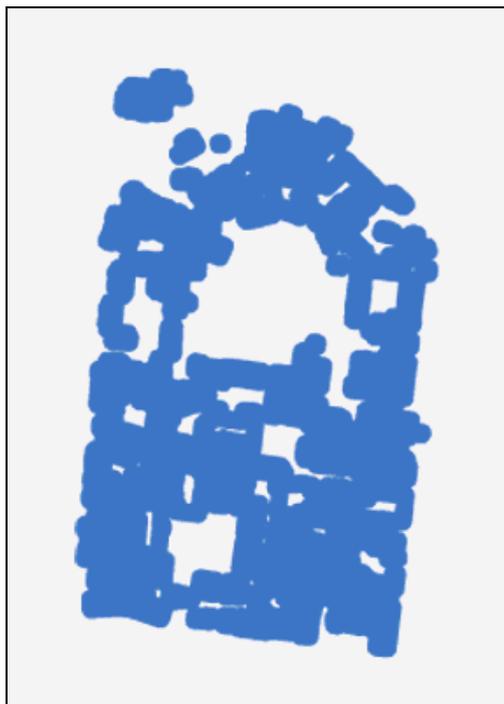


Figura 4.22. Solapamiento de áreas de influencia

El resultado muestra que las zonas de seguridad contra incendios se encuentran fuera de las áreas de influencia.



Figura 4.23. Resultado de la generación de áreas de influencia

Para determinar de forma apropiada estas zonas, se necesitará realizar operaciones de superposición adicionales.

En todos los casos mostrados la operación buffer es isotrópico o direccionalmente uniforme, ya que la distancia del objeto al borde del buffer es constante en cualquier dirección del plano.

4.2. El operador buffer en SIG comerciales

A continuación analizamos las peculiaridades más significativas del operador buffer en algunos de los SIG comerciales más conocidos: *Arcinfo*, *ArcView*, *Idrisi* y *Grass*. En todos ellos, este operador utiliza al menos dos parámetros de entrada (la imagen a procesar y la distancia o amplitud del buffer) y uno de salida (la imagen resultante o un fichero/tabla que contiene los atributos referentes a la zona de influencia calculada).

4.2.1. Arcinfo

El análisis de la información que hace *Arcinfo* a través del operador buffer es de tipo vectorial. Supongamos que partimos de una imagen denominada *carreteras* a la que se desea realizar un buffer de 5 km de distancia, el resultado de la operación se almacena en la imagen denominada *circunvalación-5000*.

La operación se expresaría mediante la orden:

```
Buffer carreteras circunvalación-5000 5000
```

Observando la sintaxis del operador vemos que está formada por la palabra reservada `buffer` seguida de tres parámetros que hacen referencia a la siguiente información: el primero de ellos es el nombre de la imagen a procesar (`carreteras`) a continuación se introduce el nombre de la imagen donde se volcará el resultado (`circunvalación-5000`) seguido del último parámetro indicador de la amplitud del buffer, que en el caso que nos ocupa, es de 5000 metros.

Arcinfo posee además un módulo de visualización denominado *Arcplot*, a partir del cual también se puede calcular la zona de influencia. La única diferencia

respecto al caso anterior estriba en la presentación del resultado. En esta ocasión se utilizan diferentes colores para representar tanto la imagen inicial como la imagen resultado; para ello es necesario incluir en la sintaxis del operador dos nuevos parámetros relativos a los colores utilizados. Con lo que, haciendo uso del módulo *Arcplot*, la operación anterior se trasladaría a la orden:

```
Mapex circunvalación-5000  
Linecolor 2  
Arcs carreteras  
Linecolor 5  
Arcplot arcs circunvalación-5000
```

El resultado de la operación anterior se muestra en la imagen que figura a continuación, en la que dos colores diferentes (negro y gris) sirven para representar la imagen inicial y la zona de influencia.

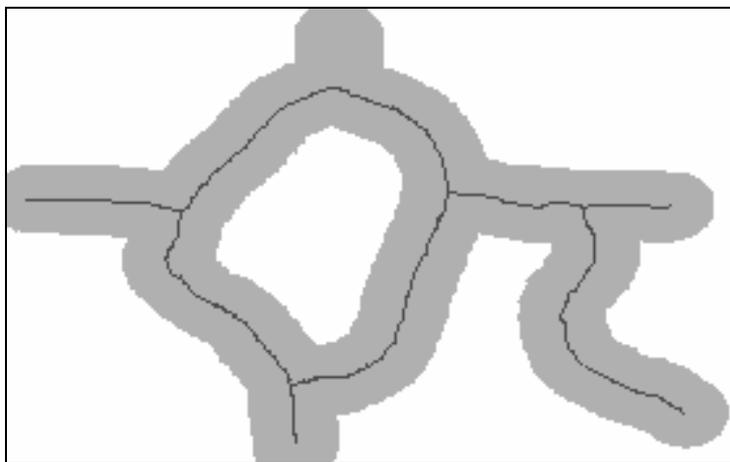


Figura 4.24. Creación del buffer con *ArcInfo*

4.2.2. ArcView

ArcView utiliza información proveniente de *Arcinfo*, *Dbase* o de ficheros tipo texto. En este caso, se parte de un objeto geométrico y una distancia dada y se genera un segundo objeto que circunda al inicial. Las imágenes siguientes

muestran los pasos a seguir para generar un buffer en la citada aplicación.

En el ejemplo de la Fig.4.26. el operador buffer se ejecuta sobre la información almacenada en el fichero *Quake99.dbf*. En este caso se desea conocer si en el terreno comprendido en un radio de 500 km hay riesgo de que se produzca un terremoto de magnitud superior o igual a cinco grados en la escala Richter.

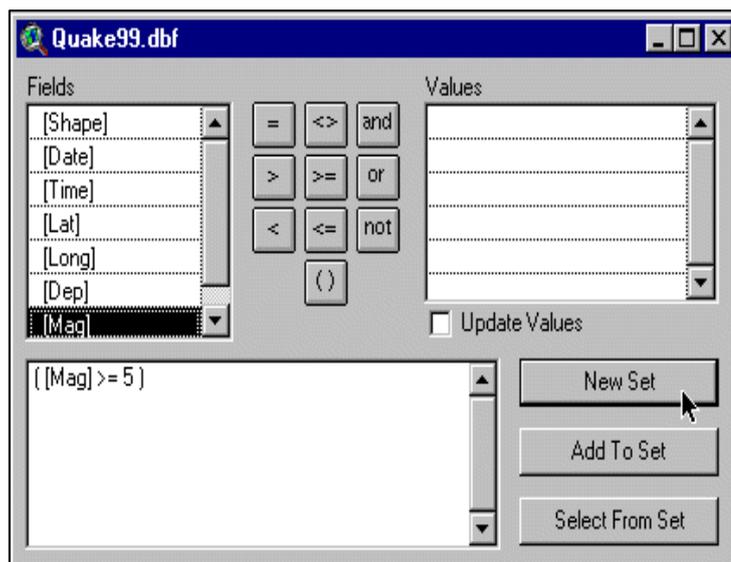


Figura 4.25. Pantalla 1 para creación de buffer con ArcView

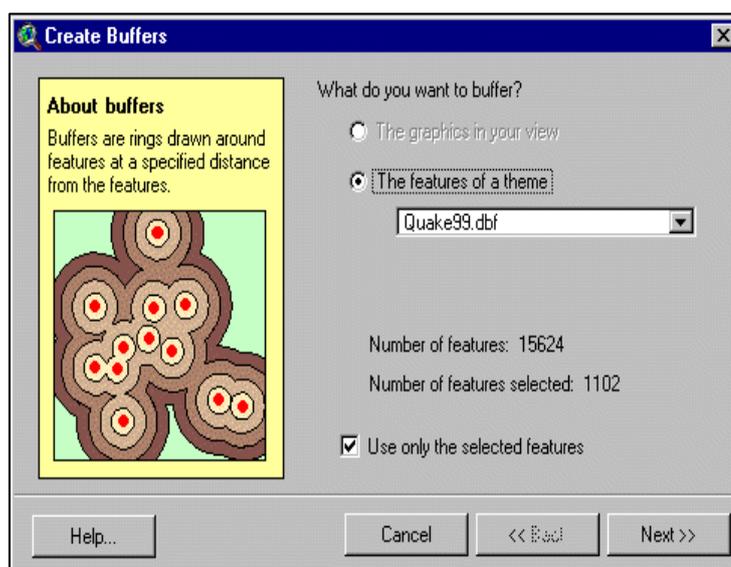


Figura 4.26. Pantalla 2 para creación de buffer con ArcView

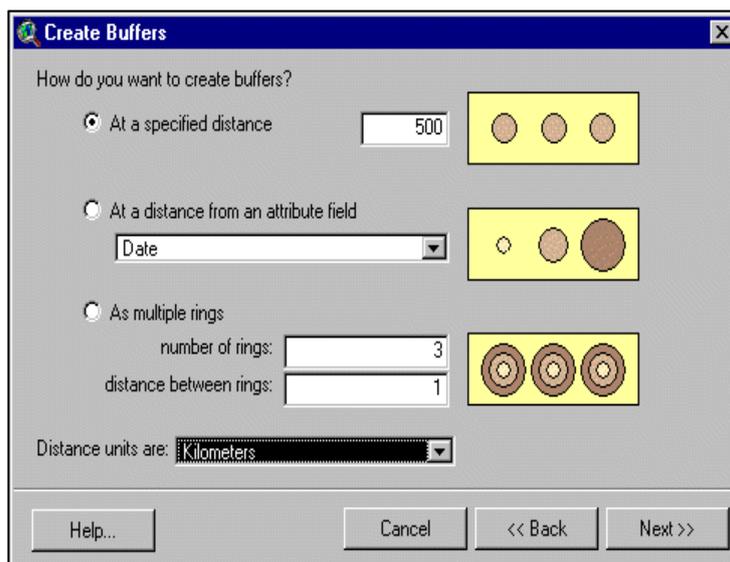


Figura 4.27. Pantalla 3 para creación de buffer con *ArcView*

En Fig.4.27. se marca la primera opción ya que, como se explicó anteriormente, se pretende calcular el buffer a una distancia de 500 km. Esta aplicación permite calcular el buffer de tres formas diferentes:

- Especificando la amplitud del buffer como un valor constante.
- Especificando la amplitud del buffer en función de un determinado atributo de la tabla.
- Múltiples buffers concéntricos.

Si se elige esta última opción, se debe indicar el número de buffer y la distancia existente entre los diferentes anillos. En todos los casos el usuario deberá indicar las unidades utilizadas para la operación.

ArcView también dispone de otro procedimiento para realizar el buffer, a través de la orden `SE_shape_generate_buffer`. Partimos de una imagen inicial, se especifica una determinada longitud y el resultado se vuelca a otra imagen. El usuario puede elegir el número de puntos que tendrá el polígono resultante de la operación. La sintaxis del operador respondería a:

```
SE_shape_generate_buffer (img_inicial,distancia, max_pts,img_final)
```

`Img_inicial` es la imagen sobre la que se le realiza el buffer, el parámetro `distancia` indica la amplitud del buffer, `max_pts` es el número de puntos que tendrá el polígono generado a partir de la operación e `img_final` como su nombre indica, guardará la imagen producida a través del buffer.

4.2.3. Idrisi

En Fig.4.28. se muestran los parámetros que se deben introducir para la generación de zonas de influencia en *Idrisi*.

Figura 4.28. Creación del buffer con *Idrisi*

Una vez introducido el nombre de la imagen a procesar y el nombre del fichero donde se almacena el resultado de la operación, se debe indicar el tamaño del buffer. A continuación se especifican tres valores numéricos que servirán para representar en la imagen resultado la zona objeto de estudio, el área incluida en la zona buffer y el área ajena a las dos anteriores. Para terminar se debe especificar tanto las unidades de medida como el título de la imagen resultante.

4.2.4. Grass

Grass genera buffers a partir de imágenes raster. La imagen resultante se encuentra dividida en celdas etiquetadas con un valor numérico que identifica el buffer al que pertenece dicha celda, de forma que todo terreno perteneciente a un buffer concreto viene representado por un determinado número. En *Grass* el buffer se realiza a través de dos operadores diferentes: *r.buffer* y *s.buffer* que comparten la misma sintaxis.

El operador *r.buffer* utiliza distancias euclidianas para medir la separación entre los centros de celdas colindantes y distancias geodésicas elipsoidales para expresar tanto la longitud como la latitud. A través de *r.buffer* se pueden realizar a una imagen hasta 60 buffers diferentes. La distancia entre los diferentes buffer se mide en metros, kilómetros, pies o millas.

Para explicar el funcionamiento de los dos módulos simulamos la creación de cinco buffers con una distancia entre ellos de 10 metros.

La sintaxis del operador *r.buffer* es la siguiente:

```
r.buffer input= map.in output=map.out distances=10,20,30,
40,50 units=meters
```

Si realizamos el buffer a partir del operador *s.buffer* la sintaxis del operador pasaría a ser:

```
s.buffer input=map.in output=map.out radius =10
```

Las unidades referentes al radio indican el número de píxeles que distan un buffer de otro.

Uno de los métodos utilizados para generar zonas de influencia es la Suma de Minkowski. En la Suma de Minkowski se define un polígono secundario o polígono generador, que al situarse sobre un punto o desplazarse sobre una polilínea o polígono genera una superficie formada por los puntos que dicho polígono generador encuentra a su paso.

El buffer isotrópico se aplica en la simulación de procesos medioambientales, en este caso el polígono generador es un círculo, lo que implica una distancia constante del borde del buffer al objeto. Tal y como se expuso en el Capítulo 1, existen fenómenos que no pueden ser considerados procesos direccionalmente uniformes, ya que dependen de la dirección. En estos casos, no es posible la utilización del operador buffer definido, por lo que es necesario un nuevo operador buffer anisotrópico, capaz de calcular zonas de influencia direccionalmente no uniformes.

4.3. Suma de Minkowski

Dadas dos imágenes A y B en R^2 , la Suma de Minkowski se define como [GooO'Ro04].

$$A \oplus B := \bigcup_{b \in B} A + b \tag{53}$$

donde A es el polígono generador, y B el esqueleto o elemento (punto, polilínea o polígono) primario. $A \oplus B$ se genera trasladando A a través de cada elemento $b \in B$, y sumando posteriormente el resultado de todas las traslaciones. La traslación del polígono generador A a través del elemento $b \in B$ se define como

$$A + b := \{a + b, a \in A\} \tag{54}$$

Si tomamos como polígono generador A un círculo y como elemento primario el conjunto de puntos $B = \{(2,3), (3,4), (2,5), (1,5)\}$:

$$A \oplus B = [(A + (2,3)) \cup (A + (3,4)) \cup (A + (2,5)) \cup (A + (1,5))] \tag{55}$$

La Fig.4.29. muestra el resultado de $A \oplus B$, $A \oplus L$ y $A \oplus P$, ambas operaciones tomando como elemento primario un punto B , una polilínea L y el polígono P respectivamente.

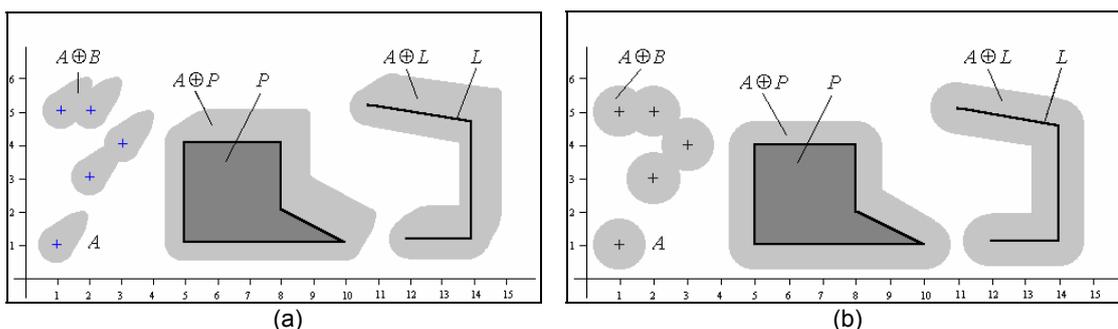


Figura 4.29. Ejemplo de Suma de Minkowski tomando como objeto generador un círculo (a) y un óvalo (b)

Conceptualmente, la Suma de Minkowski es una dilatación o expansión de la imagen primaria B , cuya forma viene determinada por el polígono generador A . En Fig.4.29.(a) hemos escogido como imagen generadora un círculo. Al ser ésta una imagen simétrica respecto a ambos ejes, la expansión de la imagen primaria es direccionalmente uniforme o isotrópica. Si por el contrario tomamos

como polígono generador un óvalo, la expansión resultante será direccionalmente no uniforme o anisotrópica, teniendo al frente su mayor avance en la dirección α en la que el óvalo presenta su máximo radio polar. La Fig.4.29.(b) muestra las sumas $A \oplus B$, $A \oplus L$ y $A \oplus P$ siendo A un óvalo con $\alpha = 45^\circ$.

4.4. Modelado del polígono generador

La Fig.4.30. representa la “mancha” que dibuja la expansión de un fluido que emana del punto P . Considerando que las fuerzas externas (viento, corrientes marinas, ...) que han distribuido el fluido han actuado sobre ella en una dirección constante, los pasos a seguir para el diseño de una función que simule su comportamiento y nos permita visualizar el proceso de expansión son los siguientes:

a) Muestreo de la mancha. Fijada la longitud del arco λ , hemos de muestrear la superficie que cubre la mancha en los $j = 2\pi/\lambda$ intervalos $\{[\Phi_i - \lambda/2, \Phi_i + \lambda/2), i = 1..j\}$ de marca $\Phi_i = \lambda(i - 1)$. Una buena aproximación a la superficie del sector i es $s_i = (\lambda/2) \cdot d_i^2$, en donde d_i es la distancia del punto P al frente de la mancha en la dirección Φ_i . La Tabla 4.1. muestra los valores Φ_i , d_i y s_i para $\lambda = 30^\circ$, siendo su histograma el que muestra la Fig.4.31.

La observación del histograma proporciona una primera aproximación de la ausencia/presencia de monomodalidad y simetría. En el caso de asimetría, la distribución de Von Misses no es válida, lo que haría necesario el ajuste mediante alguna otra función circular sesgada.

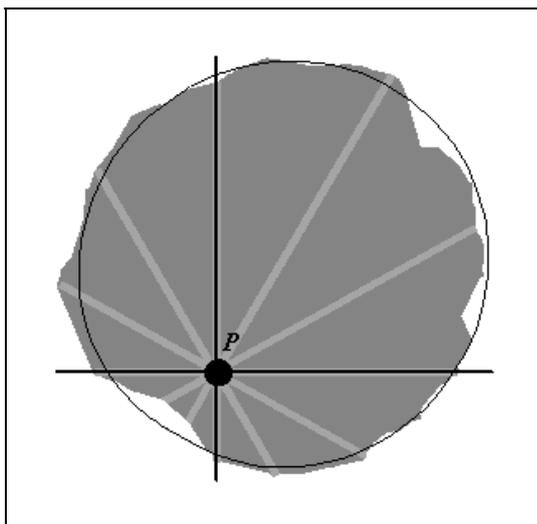


Figura 4.30. Comportamiento direccional de un fluido que emana desde el punto P

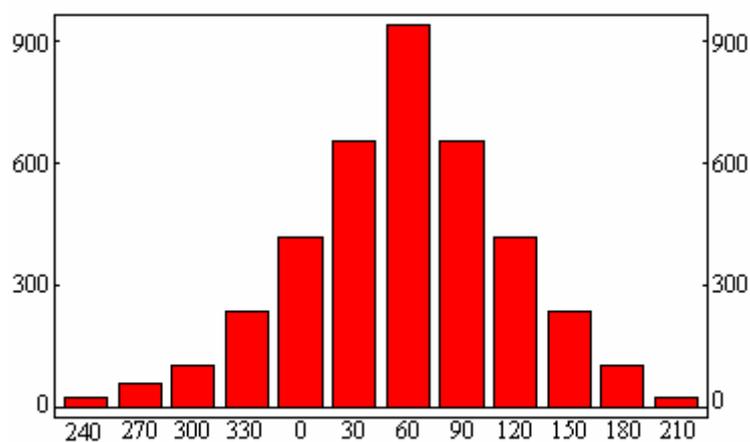


Figura 4.31. Histograma de valores. Este histograma permite comprobar la ausencia/presencia de monomodalidad y simetría

Intervalo	Φ_i	d_i	s_i
[345, 15)	0°	40	418.879
[15, 45)	30°	50	654.498
[45, 75)	60°	60	942.478
[75, 105)	90°	50	654.498
[105, 135)	120°	40	418.879
[135, 165)	150°	30	235.619
[165, 195)	180°	20	104.720
[195, 225)	210°	10	26.1799
[225, 255)	240°	10	26.1799
[255, 285)	270°	15	58.9049
[285, 315)	310°	20	104.720
[315, 345)	340°	30	235.619

Tabla 4.1. Valores Φ_i , d_i y s_i en intervalos $\lambda = 30^\circ$

b) Cálculo del vector media \bar{m} . Aplicando las Eqs.(6-9) a los datos de la Tabla 4.1., obtenemos $r = 0.597$, $\bar{\Phi} = 59^\circ$, $c = 1.01115$, $r_c = 0.597$. Una vez calculado \bar{m} , ya es posible testear si las variaciones de la muestra obedecen a la existencia de direccionamiento, o bien a desviaciones aleatorias de una expansión uniforme. Aplicando el test de Rayleigh (Sección 2.4.), obtenemos que el estadístico de prueba $z = N \cdot r^2 = 135$. Al ser el valor crítico $z(0.05) = 2.98$, encontramos sobrada significación, por lo que rechazamos H_0 a favor de la hipótesis alternativa de direccionamiento.

c) Obtención y validación de la función de densidad $f(\Phi)$. Una vez obtenido r_c , calculamos el parámetro de concentración k ($k = 0$ si asumimos uniformidad). Observamos que con el valor $k = 1.5$ se cumple Eq.(14), siendo la función densidad de probabilidad:

$$f(\Phi) = 0.097 \cdot \text{Exp}[k \text{Cos}(\Phi - 59^\circ)] \quad (56)$$

Una distribución es un modelo que ayuda a interpretar y comprender un fenómeno estocástico. Mientras que el fenómeno es siempre complejo, el modelo ha de ser conceptualmente simple, por lo que no puede reproducir de forma exacta el comportamiento de aquel. Este hecho plantea la necesidad de validar el modelo mediante la cuantificación de las desviaciones entre el comportamiento real del fenómeno y el comportamiento del modelo. Estadísticamente, esto se reduce a chequear si la distribución teórica $f(\Phi)$ se ajusta a los valores observados.

Aplicando Eq.(18) para realizar el Test de Cox (Sección 2.5), obtenemos $m_c = 0.172$ y $m_s = 0.013$. Puesto que ambos momentos se aproximan a sus respectivas medias ($I_2(1.5)/I_0(1.5) = 0.205$ y 0), podemos asumir que el ajuste es bueno.

A partir de la función densidad de probabilidad $f(\Phi)$, obtenemos la función de distribución $F(\alpha, \beta, \Phi)$:

$$F(\alpha, \beta, \Phi) = \int_{\alpha}^{\beta} f(\Phi) d\Phi \quad (58)$$

que devuelve la probabilidad de que la variable aleatoria esté comprendida entre los ángulos α y β , y cuyo valor coincide con el área de la campana comprendida entre ambos ángulos (Fig.4.32.(a)).

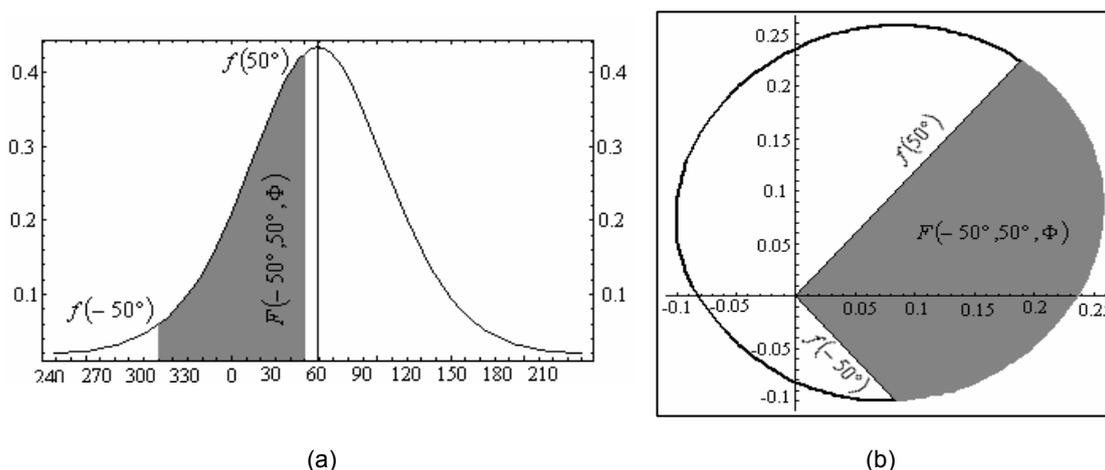


Figura 4.32. Representación de la función de Von Mises con el parámetro de concentración $k = 1.5$ y vector media $\bar{\Phi} = 59^\circ$. La superficie de la campana comprendida entre los dos ángulos (a) debe ser igual a la superficie del sector curvilíneo del óvalo generador limitado por dichos ángulos (b)

d) Modelado del polígono generador. La imagen generadora ha de ser un polígono que distribuya la superficie alrededor de su centro según la función de densidad $f(\Phi)$. Esto significa que la superficie de la campana entre dos ángulos debe de ser igual a la superficie del sector curvilíneo del polígono generador limitado por dichos ángulos. Si el borde del polígono posee como coordenadas polares $\rho = g(\Phi)$, lo anterior se traduce a la expresión

$$\int_{\alpha}^{\beta} f(\Phi) d\Phi = \frac{1}{2} \int_{\alpha}^{\beta} [g(\Phi)]^2 d\Phi \quad (59)$$

de lo que se deduce que el polígono generador es el óvalo de radio polar

$$\rho = g(\Phi) = \sqrt{2f(\Phi)} = \frac{1}{\sqrt{\pi I_0(k)}} \text{Exp} \left[\frac{k}{2} \text{Cos}(\Phi - \bar{\Phi}) \right] \quad (60)$$

Podemos observar que para $k = 0$ la expansión es uniforme, quedando

reducido el ovalo a un circulo de radio $1/\sqrt{\pi}$. La Fig.4.32.(b) muestra el óvalo generador de la función de densidad Eq.(56).

Durante la expansión producida en el periodo de tiempo t_i , las fuerzas externas tendrán una dirección dominante θ_i y una velocidad v_i , a la que viajará el frente de la mancha. En consecuencia, al final del periodo t_i , el frente de la mancha avanzará una distancia $l_i = v_i \cdot t_i$. Ello hace necesario el modelado de óvalos generadores $G(l, \theta)$ cuyo frente se encuentre a una distancia l del centro. De Eq.(60) se deduce que el frente del ovalo se encuentra en la dirección $\bar{\Phi}$, ya que es el valor de Φ el que hace máximo ρ , por lo que el radio polar de $G(l, \theta)$ vendrá dado por la expresión

$$g_l(\Phi) = l \cdot \text{Exp} \left[\frac{k}{2} [\text{Cos}(\Phi - \bar{\Phi}) - 1] \right] \quad (61)$$

4.5. Visualización de la expansión

El proceso de expansión final E, es el resultado de la acumulación de las expansiones parciales E_i que se realizan en los periodos de tiempo t_1, t_2, \dots, t_n :

$$E = \bigcup_{i=1}^n E_i(l_i, \theta_i) = \bigcup_{i=1}^n E_i(v_i \cdot t_i, \theta_i) \quad (62)$$

Partiendo de una imagen primaria P, la expresión anterior es equivalente a

$$E = [\dots[[P \oplus G(l_1, \theta_1)] \oplus G(l_2, \theta_2)] \oplus \dots \oplus G(l_n, \theta_n)] \quad (63)$$

Por la propiedad asociativa de la suma de Minkowski, la Eq.(63) puede escribirse como

$$E = P \oplus [G(l_1, \theta_1) \oplus G(l_2, \theta_2) \oplus \dots \oplus G(l_n, \theta_n)] = P \oplus G_t \quad (64)$$

por lo que la suma de n expansiones parciales puede ser interpretada como una única expansión en la que el polígono generador es G_t de radio polar

$$g_t(\Phi) = \sum_{i=1}^n l_i \cdot \text{Exp} \left[\frac{k}{2} [\text{Cos}(\Phi - \theta_i) - 1] \right] \quad (65)$$

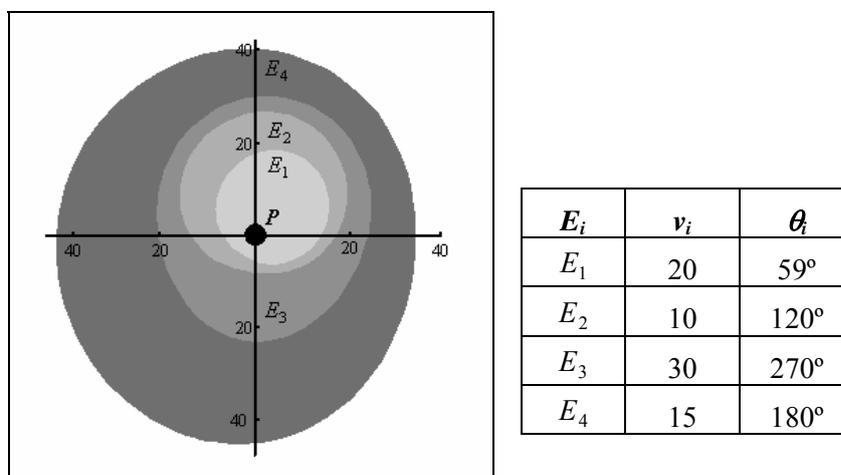


Figura 4.33. Expansión del punto P . La tabla muestra los valores v_i y θ_i utilizados en cada E_i . El polígono generador utilizado (Eq (60)) posee valor $k=1.5$

Las consideraciones anteriores han sido expuestas asumiendo que la dirección dominante y la velocidad de las fuerzas externas son constantes en cada periodo i . Esta asunción choca con la realidad, ya que a medida que aumenta el tamaño de la mancha, es más probable que ésta incluya zonas con distintas fuerzas externas en un mismo periodo de tiempo. La posible casuística cuando en un periodo i varía v_i y/o θ_i , es la siguiente:

Caso 1: La expansión se inicia con distintas fuerzas que se mantienen constantes durante todo el periodo (la imagen primaria comprende zonas con distintos valores de v_i y/o θ_i que no se modifican durante el periodo, por lo que los polígonos generadores presentan valores constantes de v_i y/o θ_i).

Caso 2: La expansión se inicia con una única fuerza que varía durante el periodo (la imagen primaria posee valores fijos de v_i y θ_i , y el polígono generador encuentra a su paso zonas con distintos valores de v_i y/o θ_i).

Caso 3: La expansión se inicia con distintas fuerzas que varían durante el periodo (la imagen primaria comprende zonas con distintos valores de v_i y/o θ_i).

y el polígono generador encuentra a su paso zonas con distintos valores de v_i y/o θ_i). Este caso supone la ocurrencia simultánea de los casos 1 y 2.

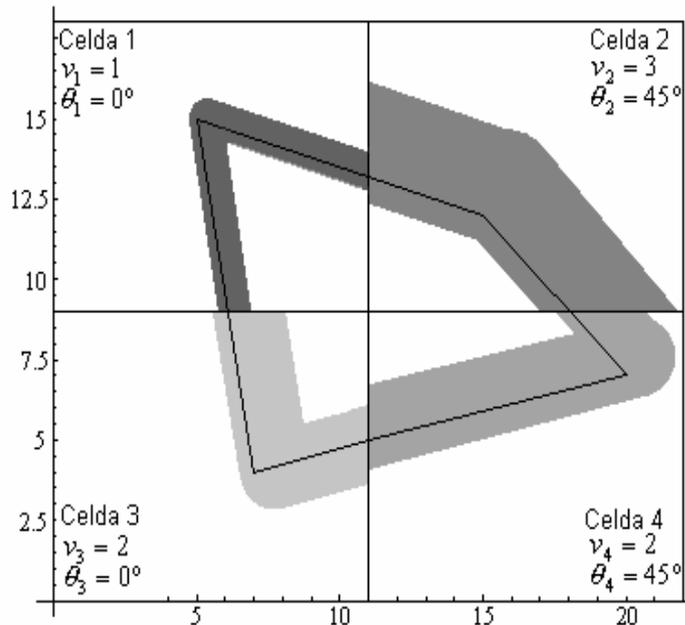


Figura 4.34. Generación de áreas de influencia. Cada celda contiene diferentes valores en función de la velocidad y la dirección dominante de las fuerzas externas que actúan sobre ella

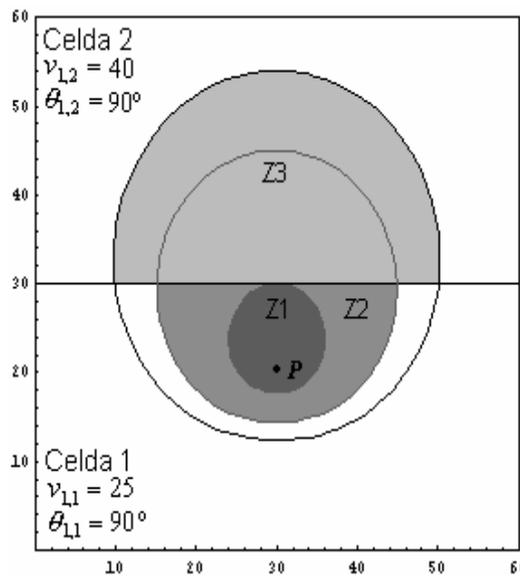


Figura 4.35. La expansión E del punto P es el resultado de las expansiones de la celda1 $E_{1,1}$ (Z1+Z2) y de la celda2 $E_{1,2}$ (Z3)

4.5.1. Caso 1

La Fig.4.34. muestra el polígono primario de vértices $P = \{(5,15), (15,12), (20,7), (7,4)\}$. La superficie en la que éste se inscribe se

encuentra dividida en celdas, cada una de las cuales posee un valor de dirección dominante y velocidad de las fuerzas externas que actúan en ella. Al existir al menos un lado de P incluido en celdas de distintos valores, el buffer no puede ser calculado aplicando la suma de Minkowski tal y como se hizo en los apartados anteriores.

La generación de la zona de influencia requiere:

a) Segmentación de cada lado de la imagen en tantos segmentos como celdas ocupe aquél, de forma que tras la división, cada uno de los segmentos resultantes posea un único valor de v y θ . Tras esta operación el polígono P queda como $P = \{(5,15), (11,13.2), (15,12), (18,9), (20,7), (11,5), (7,4), (6.09,9)\}$, habiendo añadido los vértices $(11,13.2), (18,9), (11,5)$ y $(6.09,9)$.

b) Suma de Minkowski sobre cada uno de los segmentos s_j de la imagen, utilizando como polígonos generadores en cada segmento $G(l_j, \theta_j)$, en donde l_j y θ_j son los valores de velocidad por unidad de tiempo y dirección dominante en su celda continente c , debiendo la suma circunscribirse a dicha celda.

$$A \oplus B = \{X + Y : X \in A, Y \in B\} \quad (66)$$

4.5.2. Caso 2

Es frecuente que durante el transcurso de un periodo, se produzcan cambios en la velocidad y/o dirección de las fuerzas externas que determinan la forma de la mancha. A nivel de la aplicación del operador suma de Minkowski, este hecho se traduce en que el polígono generador encuentra a su paso zonas con distintos valores de v_i y/o θ_i , lo que hace necesario la utilización de tantos polígonos generadores como pares (v_i, θ_i) distintos encuentre la expansión. El caso más simple se muestra en la Fig.4.35. En ella observamos que en el periodo $i = 1$, la expansión del punto P es el resultado de las expansiones en la celda 1 y en la celda 2:

$$E_1 = E_{1,1} + E_{1,2} = Z1 + Z2 + Z3 \quad (67)$$

La expansión exclusiva en la celda 1 (zona Z1) se produce desde el momento inicial $t=0$ hasta el momento $t=t_c$, a partir del cual se produce simultáneamente la expansión en ambas celdas. La expansión en la celda 1 sigue el comportamiento habitual durante todo el periodo, por lo que asumiendo $t_1 =$ la unidad de tiempo:

$$E_{1,1} = (P \oplus G(l_{1,1}, \theta_{1,1})) \cap \text{celda 1} = Z1 + Z2 \quad (68)$$

En la celda 2, la expansión dependerá del momento t_c :

$$E_{1,2} = (P \oplus G(v_{1,1}, v_{1,2}, \theta_{1,1}, \theta_{1,2}, t_c)) \cap \text{celda 2} = Z3 \quad (69)$$

Siendo $P \oplus G(v_{1,1}, v_{1,2}, \theta_{1,1}, \theta_{1,2}, t_c)$ el polígono generador de radio polar

$$g(\Phi) = v_{1,1} \cdot t_c \cdot \text{Exp}\left[\frac{k}{2} [\text{Cos}(\Phi - \theta_{1,1}) - 1]\right] + v_{1,2} \cdot (1 - t_c) \cdot \text{Exp}\left[\frac{k}{2} [\text{Cos}(\Phi - \theta_{1,2}) - 1]\right] \quad (70)$$

Los valores de la Fig.4.35. se muestran en la Tabla 4.2.

	$0 < t \leq t_c$	$t_c < t \leq 1$
$E_{1,1}$	$P \oplus G(25, 90^\circ) = Z1$	$(P \oplus G(25, 90^\circ)) \cap \text{celda 1} = Z2$
$E_{1,2}$	-----	$(P \oplus G(25, 40, 90^\circ, 90^\circ, 0.4)) \cap \text{celda 2} = Z3$

Tabla 4.2. Valores de la expansión mostradas en la Fig.4.35.

La Fig.4.36. muestra distintos ejemplos de variaciones en la velocidad y/o dirección de las fuerzas. Las expansiones en cada celda son las dadas por Eqs.(68-70). Aunque el procedimiento ha sido descrito por motivos didácticos tomando como imagen primaria un punto, éste es extensible a imágenes primarias de tipo polilínea y polígono.

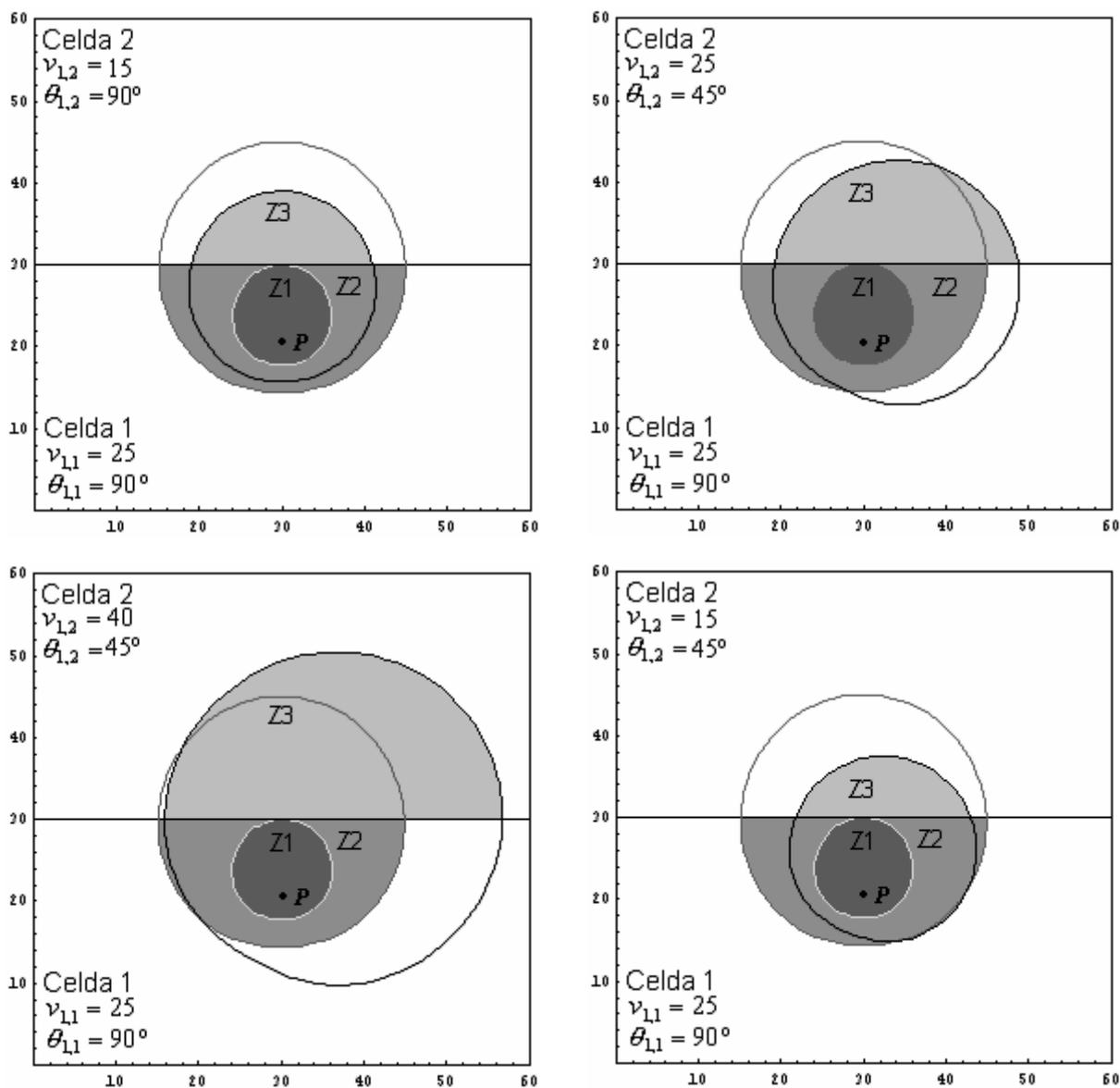


Figura 4.36. Diferentes ejemplos de expansión en las celdas en función de los cambios en la velocidad y dirección de las fuerzas

4.5.3. Caso 3

Como ya se apuntó anteriormente, éste es el caso en el que se producen simultáneamente los dos anteriores. Para su resolución basta con realizar la segmentación descrita en el caso 1 y sobre el nuevo conjunto de puntos, aplicar el caso 2.

4.6. Un ejemplo ilustrativo

La Fig.4.37.(a) es una fotografía [Gat08] de Ibiza y Formentera (tomadas desde el vuelo STS-78) perteneciente a la colección de imágenes digitales del Centro Johnson de la NASA [Jsc08]. Para simular y visualizar la expansión de los fenómenos anisotrópicos, seleccionamos la ventana W de 501 x 501 píxeles (Fig.4.37.(b)). Antes de la visualización del fenómeno de expansión anisotrópica, se debe determinar si existe direccionamiento en cada uno de los píxeles de W y en caso de que exista, debe calcularse su dirección dominante.

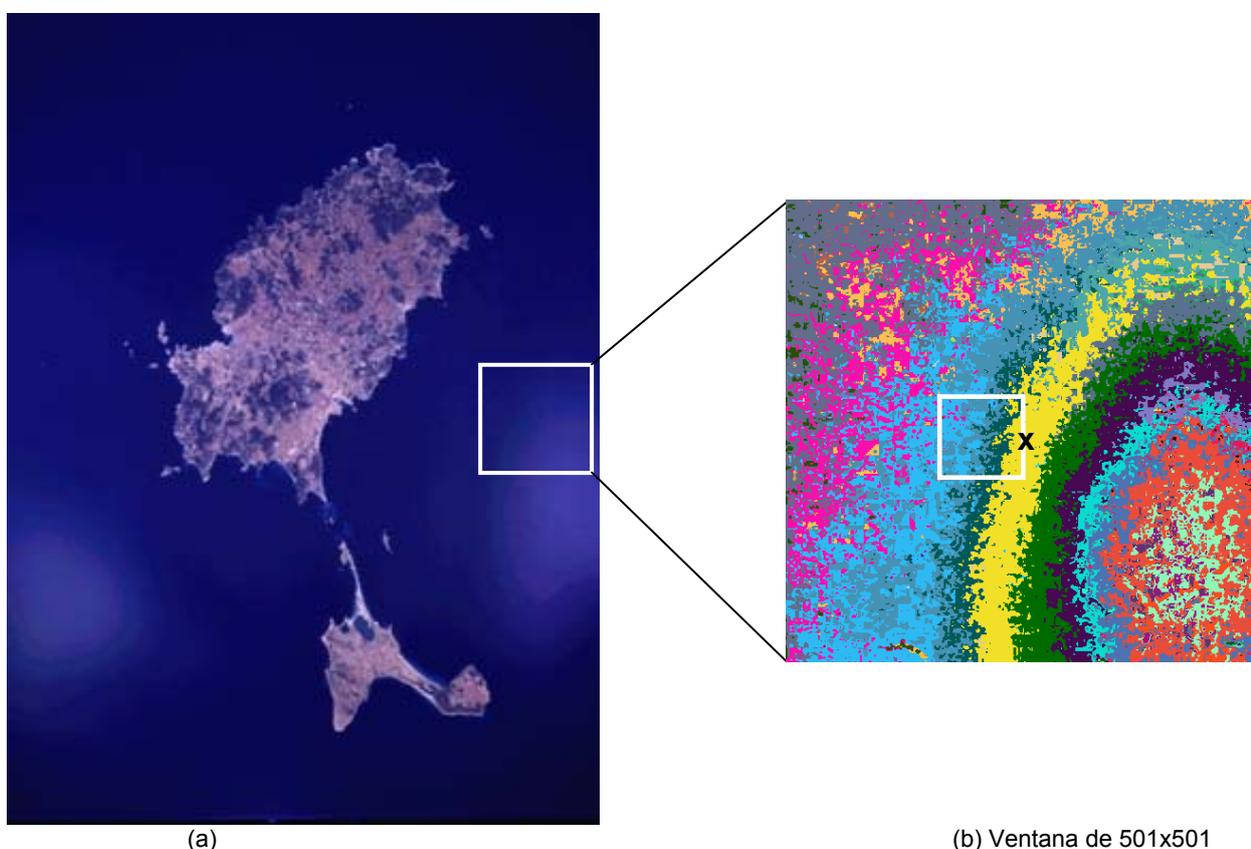


Figura 4.37. Imagen de control remoto de las islas de Ibiza y Formentera (Islas Baleares) tomadas desde el vuelo STS-78 de la NASA(a). El flujo de las corrientes marinas produce patrones complicados en la superficie del agua debido al reflejo de la luz solar (b)

4.6.1. Determinación de las direcciones dominantes

Aplicando a la Fig.4.37.(b) de 501x501 celdas, el método de análisis espacial descrito en la Sección 3.4. para la determinación de las direcciones dominantes, obtenemos que de los $(501-5) \times (501-5)=246016$ celdas,

solamente 2182 se pueden considerar espacialmente independientes, ya que presentan valores dependientes fuera del intervalo de confianza $\chi^2_{246015} = (244863-246016)$.

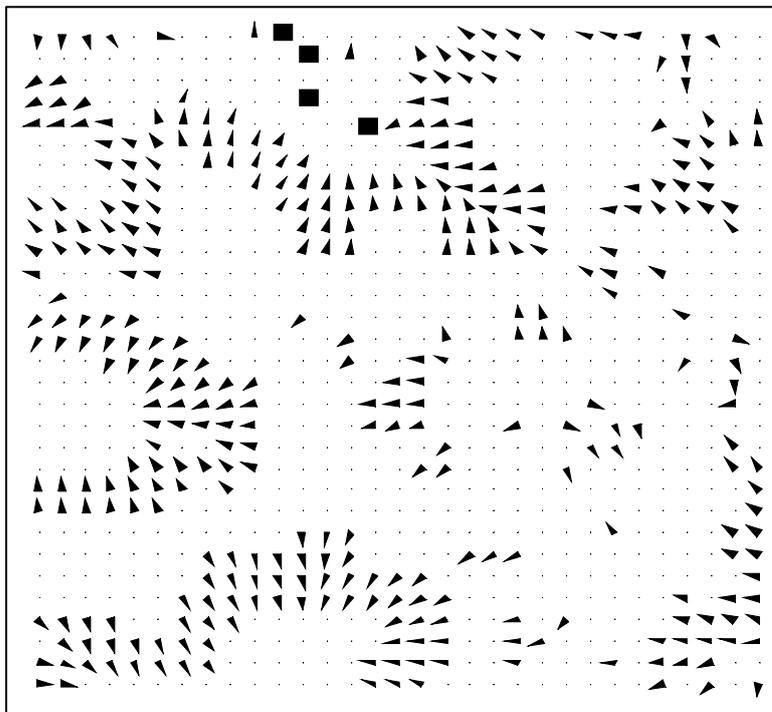


Figura 4.38. El comportamiento de las celdas que pertenecen a la ventana señalada en la Fig.4.37.(b)

El comportamiento de las celdas que pertenecen a la ventana señalada en la Fig.4.37.(b) se muestra de forma gráfica en la Fig.4.38.

Los símbolos se interpretan de la siguiente forma:

“■” Homogeneidad en los efectos globales y la ausencia de dependencia espacial en los efectos locales.

“▲” Homogeneidad en los efectos globales y la dependencia espacial en los efectos locales con comportamiento isotrópico. Las flechas representan la homogeneidad en los efectos globales y la dependencia espacial en los efectos locales, con comportamiento anisotrópico en la dirección ψ calculada con los vectores media $d = 5$.

AQUÍ VA LA FIGURA 4.39 EN A3 PAGINA 105 DEL DOCUMENTO

La Fig.4.40.(a) muestra la zona de influencia calculada usando el método descrito, partiendo del punto x de coordenadas (245, 245) de la ventana W (Fig.4.37.(b)).

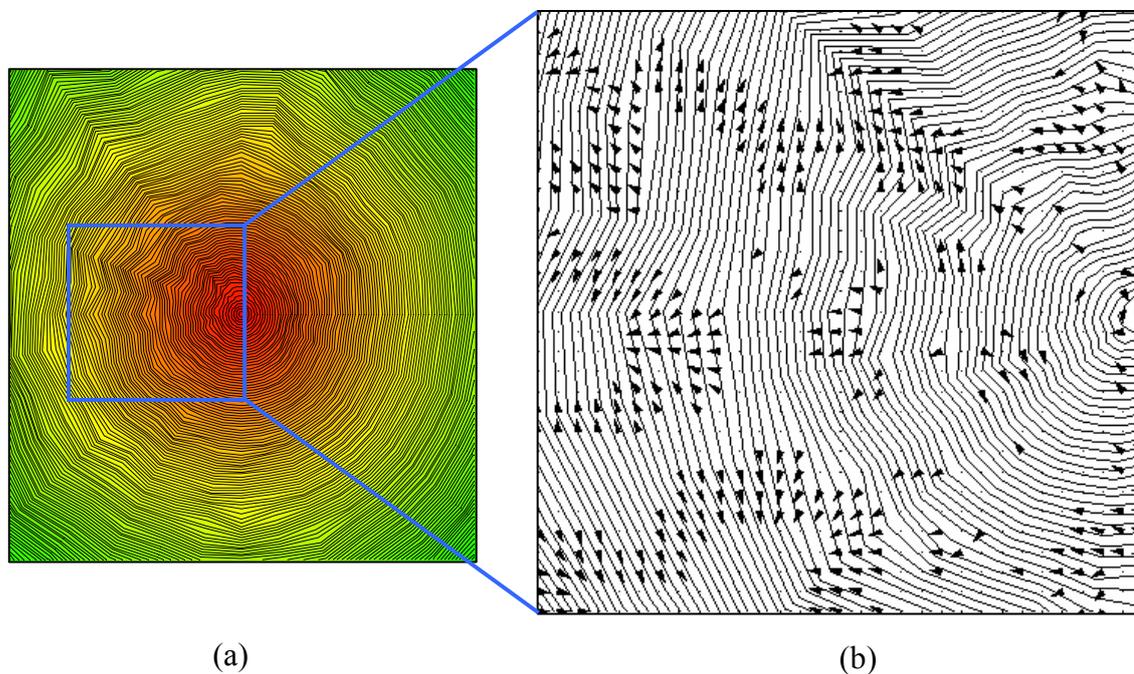


Figura 4.40. Zona de influencia calculada desde el punto x de coordenadas (245, 245) en la Fig.4.37

En la Sección 3.5. se calcularon las direcciones de las corrientes marinas en el Estrecho de Gibraltar a partir de una imagen satélite. La Fig.4.41.(b) muestra los resultados de las direcciones ψ obtenidas (Eq.33). Se observa que las variaciones del proceso en dirección horizontal son diferentes de las de la dirección vertical, ya que hay una clara tendencia anisotrópica en la dirección este-oeste. Las zonas de influencia (Fig.4.41.(c)) se han generado usando un óvalo $g_l(\Phi)$ (Eq.61) con parámetros $l = 0.8$ y $k = 0.4$.

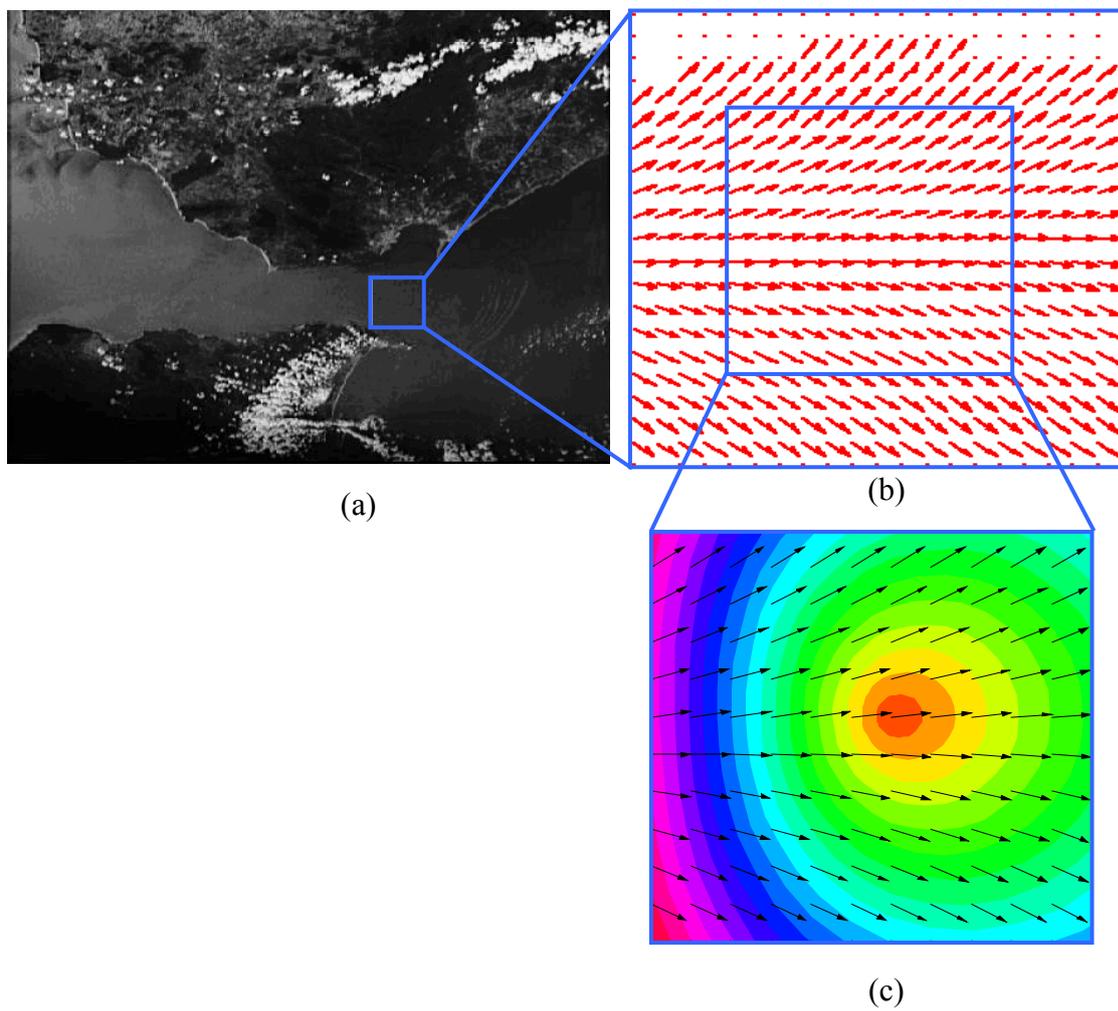


Figura 4.41. Estudio de las corrientes marinas en el Estrecho de Gibraltar

Capítulo V

Un caso de estudio: El SIG como entorno de trabajo para extracción de conocimiento

El propósito de este capítulo es aplicar las técnicas diseñadas por el doctorando y expuestas en los capítulos precedentes a un proceso de extracción de conocimiento espacial y de técnicas de minería de datos espacial sobre colecciones heterogéneas de datos. Para ello, describiremos el desarrollo de un modelo gráfico (mapa) de coste en tiempo para el desplazamiento de un móvil, utilizando proceso de extracción de conocimiento sobre la información contenida en nuestra base de datos y sobre los datos reales (mapas temáticos) almacenados en el SIG *Idris32*.

5.1. Extracción de conocimiento y minería de datos

Los avances tecnológicos en la recogida de información espacial (teledetección, telemetría vía satélite, fotografía aérea ...) han permitido la creación de sistemas de información que gestionan grandes volúmenes de datos organizados en forma de cartografía temática y almacenados en grandes bases de datos espaciales (Spatial Data Bases o SDB) referenciadas geográficamente.

Esta creciente cantidad de información ha hecho que se cuestionen los métodos tradicionales de análisis de datos y de reconocimiento de patrones, planteándose la necesidad de desarrollar enfoques y técnicas que permitan extraer conocimiento e información original no evidente de grandes SDB. Para ello nació la minería de datos espaciales (Spatial Data Mining o SDM) y el denominado KDSD (Knowledge Discovery in Spatial Databases), disciplinas que integran diversos campos como aprendizaje automático, sistemas de bases de datos, estadísticos espaciales y no espaciales, visualización de datos, etc. SDM y KDSD aluden a la extracción de conocimiento implícito, relaciones espaciales u otros patrones (modelos) no almacenados explícitamente en las SDB [Est95] .

La creación de modelos a partir de datos es conocida por diferentes nombres (incluido DM) en diferentes comunidades (extracción de conocimiento, arqueología de datos, recolección de información, procesado de patrones de datos, etc.). Para evitar confusiones, en esta memoria utilizaremos el término KDSD para referirnos al proceso total de descubrimiento de información útil a partir de datos espaciales, siendo DM un paso particular de este proceso: la aplicación de algoritmos específicos para la extracción de patrones a partir de datos ¹.

Las experiencias en el campo KSDS no han sido numerosas. Entre ellas podemos destacar el proyecto (POSS-II) [Fay96] de clasificación de galaxias y estrellas desarrollado por el observatorio Monte Palomar en Estados Unidos y

¹ Hemos optado por mantener en el texto los acrónimos en inglés, ya que están admitidos de forma generalizada por la comunidad científica.

Magellan Study [Fay93], sistema de detección y clasificación de volcanes a partir de imágenes satélite de alta resolución. La cuestión inmediata que se desprende de este hecho es la siguiente: pareciendo demostrada su utilidad ¿Porqué investigadores y usuarios de información espacial se muestran remisos a la aplicación de KDSD y DM en sus procesos de análisis y modelado?

Debido a la diversa procedencia de la información espacial, los almacenes de datos suelen estar compuestos por cartografía digital e imágenes raster con distintos formatos de almacenamiento, utilizándose para su gestión los SIG. La implementación de las técnicas KDSD, tal y como se ha realizado hasta ahora, requiere arquitecturas específicas distintas a los SIG, habitualmente caras, lo que produce una fuerte inercia al cambio.

El propósito de este capítulo es mostrar, mediante la descripción de un caso, que las capacidades analíticas y de transformación de cartografía de los SIG's proporcionan un marco adecuado para la aplicación de KDSD y de técnicas DM sobre colecciones heterogéneas de datos. Para ello, describiremos el desarrollo de un modelo gráfico (mapa) de coste en tiempo para el desplazamiento de un móvil, utilizando KDSD sobre la información contenida en nuestra base de conocimiento y sobre los datos reales (mapas temáticos) almacenados en nuestro SIG *Idrisi32* [Idr03].

Este proceso requiere, en primer lugar, establecer los parámetros que determinan el coste mediante la identificación de reglas de asociación espacial que descubran dependencias entre características. Asumiremos que dichos parámetros son la densidad del ecosistema forestal y las discontinuidades orográficas encontradas en el camino. Cada uno de estos factores generará una fricción que será en último término, la determinante de la función de coste. En el caso del ecosistema forestal, la fricción no depende de la dirección que tome el móvil en su desplazamiento (función isotrópica), siendo calculada mediante operaciones escalares sobre los atributos contenidos en el mapa de ecosistemas forestales actuales.

Las discontinuidades del terreno, por el contrario, presentan fricciones

anisotrópicas o direccionales. La asignación de fricciones a fenómenos que presentan un comportamiento direccional requiere el uso de funciones anisotrópicas, ya que en cada localización la fricción dependerá del ángulo diferencia δ entre la dirección en la que el fenómeno alcanza su valor máximo y la dirección real de actuación.

5.2. Métodos de Minería de Datos Espaciales.

Son muchos los criterios utilizados para clasificar los métodos usados en DM: tipo de base de datos en la que se ejecuta la DM (bases de datos relacionales, orientadas a objetos, espaciales, deductivas, data warehouses ...), el tipo de conocimiento buscado (reglas de asociación, reglas de características, reglas de clasificación, clustering ...) y el tipo de técnicas de DM utilizadas (minería basada en generalización, orientada a patrones, mediante análisis estadístico, redes neuronales ...).

No es nuestra intención realizar una descripción exhaustiva de estos métodos, por lo que nos limitaremos a describir algunos de ellos particularizados para información contenida en grandes bases de datos geográficas: análisis espacial de datos, métodos basados en generalización, cluster de datos espaciales y reglas de asociación espacial.

El análisis espacial de datos [BaiGat96] [FotRog94] [Goo87] es hoy en día la forma más común de explorar y modelar datos espaciales, y aunque el termino DM ha tenido connotaciones negativas en la literatura estadística [Chat95] [Gli97] son dos campos “condenados a entenderse”. La DM está situada entre la informática y la estadística, y sólo mediante la simbiosis de ambas disciplinas es posible extraer información útil y no almacenada de forma explícita en grandes bases de datos. Las herramientas más usadas en análisis espacial de datos [Cre91] [Bur90] son las expuestas en la Sección 3.1. (métodos gráficos y numéricos de visualización y descripción, métodos del vecino más próximo y k-funciones para la exploración de patrones, métodos de suavizado bayesiano, autocorrelación espacial, variogramas, correlogramas, kriging, co-kriging, regresión espacial...).

Algunas técnicas como la de aprendizaje mediante ejemplos, necesitan de un

preproceso de datos que elimine inconsistencias que podrían generar reglas equivocadas, para lo que son utilizados los métodos basados en generalización. Estos métodos describen el conocimiento en forma de relaciones entre atributos espaciales y no espaciales a altos niveles de generalización, en función de las jerarquías espaciales y no espaciales establecidas previamente.

Son dos los algoritmos de generalización utilizados: generalización de predominio no espacial y generalización de predominio espacial [Lu93]. El primero de ellos realiza la generalización sobre los atributos no espaciales mediante un cambio de nivel en la jerarquía (normalmente mediante reclasificación de valores) y la posterior unión de las tuplas idénticas. Esta operación implica que zonas adyacentes con el mismo valor en el atributo generalizado se fusionen en una. La generalización de predominio espacial por el contrario, realiza en primer lugar una generalización mediante la selección de los objetos espaciales que delimitan las regiones sobre las que posteriormente se generalizarán los atributos no espaciales.

El análisis de cluster es una herramienta estadística bien conocida utilizada frecuentemente en diversas disciplinas. Para su adaptación a DM espacial, se han desarrollado dos algoritmos similares a los descritos en los métodos basados en generalización: un enfoque de predominio espacial y un enfoque de predominio no espacial [NgHan94].

En el enfoque espacial, los objetos espaciales son clasificados en clusters usando un algoritmo de clasificación basado en muestras, para después realizar una inducción sobre los atributos no espaciales de los objetos incluidos en cada cluster. En el enfoque no espacial, por el contrario, se aplica en primer lugar una generalización sobre los atributos no espaciales, agrupando los objetos espaciales generalizados. Los clusters superpuestos y sus correspondientes atributos no espaciales generalizados deben ser fusionados.

Las reglas de asociación espaciales [KopHan95] aparecen como una extensión de las reglas de asociación de propósito general propuestas por Agrawal en [Agra93]. Mediante estas reglas es posible asociar uno o más objetos

espaciales con otro u otros objetos espaciales. Una regla de asociación espacial simple es una expresión de la forma:

$$\text{Predicado_A}(x, \text{valor_atri_A}) \rightarrow \text{Predicado_B}(x, \text{valor_atri_B}) \quad (c\%)$$

en donde:

Predicado_A y Predicado_B deben ser predicados espaciales que establezcan una relación topológica (intersección, superposición, está_dentro, está_fuera, cubre, cubierto_por, contiene, cerca_de, lejos_de ...) entre el objeto x y valor_atri , y $c\%$ es el nivel de confianza de la regla, e indica el porcentaje de objetos x que, satisfaciendo el predicado antecedente de la regla, también satisfacen el predicado consecuente.

Para cuantificar la fortaleza de una regla, además del nivel de confianza $c\%$, hemos de calcular el valor $s\%$ apoyo de la regla, que se define como el porcentaje de objetos x que satisfacen el predicado antecedente sobre total de objetos existentes. Para seleccionar una regla, se fijan unos niveles mínimos de apoyo y confianza, siendo considerada una regla como fuerte (y consecuentemente aceptada como válida) si tiene valores de apoyo y confianza no menores que los niveles preestablecidos.

5.3. Modelado de la función de coste.

El primer paso es determinar mediante un modelo el propósito del proceso KDSD. En nuestro caso, deseamos determinar el coste en tiempo del desplazamiento de un móvil. Esto incluye conocer de qué parámetros depende el coste, el modelo usado para describirlo y el tipo de representación a utilizar. La experiencia muestra que los parámetros que determinan el coste son las continuidades orográficas del terreno: pendiente y orientación.

Tal y como vimos en la Sección 1.5., a cada clase observada de pendiente y orientación, se le asignará un nivel de fricción o *stated friction* relativa a un valor patrón de fricción. En consecuencia, el coste para desplazarse entre dos puntos en términos de tiempo, energía, dinero, etc., dependerá tanto de la distancia entre ambos como de la fricción encontrada en el camino.

Tomemos la pendiente del terreno como un ejemplo de entidad geográfica. Si deseamos calcular el coste económico de la instalación de una canalización de agua desde la fuente hasta un hipotético punto de destino, asignaríamos fricciones a cada punto en función del tipo de pendiente, teniendo en cuenta que las fricciones en zonas abruptas son mayores que en zonas de pendiente nula o moderada debido a las dificultades que implica la construcción en terrenos con discontinuidades orográficas. Sin embargo, la fricción estimada para un cierto tipo de pendiente será la misma con independencia de la dirección que toma la canalización, por lo que podemos afirmar que el coste es isotrópico o independiente de la dirección.

El coste en tiempo del desplazamiento del móvil también depende de la pendiente, pero éste será diferente si el camino tomado es pendiente arriba, pendiente abajo o perpendicular a la pendiente. En este caso, el coste no depende sólo de la pendiente asociada al camino, sino que también dependerá de la dirección que tome el móvil. En consecuencia, el coste es anisotrópico, siendo el modelo que utilizaremos el siguiente:

$$c = d(s, a) + m \quad (71)$$

una función lineal en la que c representa el coste total, d el coste debido a las discontinuidades orográficas (pendiente s y orientación a del terreno) y m el coste debido al resto de los parámetros desconocidos que contribuirían al coste total.

La representación usada para este modelo será un mapa de costes almacenado en formato *Idrisi32*: un fichero (.RST) y su fichero de documentación asociado (.RDC). El problema es ahora descubrir cuáles serían, de existir, estos parámetros desconocidos y cómo contribuyen al modelo. Para ello, buscaremos reglas de asociación espacial entre clases mediante DM.

5.4. Información disponible y preproceso de datos.

El primer paso a realizar cuando se afronta un proceso de DM, tanto para bases de datos de propósito general como para SDB, es determinar qué información hay disponible en el sistema de información. La cartografía inicial

disponible en nuestro SIG incluye:

Mapas temáticos en formato vectorial. Los mapas temáticos muestran la distribución espacial de una realidad cualitativa o cuantitativa, que al estar representada mediante un sistema georeferenciado, adquiere propiedades espaciales. Los mapas temáticos vectoriales incluidos en nuestro sistema son: Régimen de Humedad y Régimen Térmico (índices 1º y 2º de la clasificación climática de Papadakis), Índice de Evapotranspiración Potencial (índice de Penman), Recursos Hídricos Subterráneos (incluyendo tipo de acuífero), Usos del Suelo (urbano, agricultura, ...), Unidad Edafológica (nivel 1 de clasificación de la F.A.O. [FAO94]), Ecosistema Forestal Actual (pinar, acebuchal, encinar, ...) con valores de densidad (número de especies por unidad de superficie) y tipo de aprovechamiento (madera, corcho, piña, ...) y Caminos GPS con valores de dirección y tiempo en cada tramo. Para describir el proceso, hemos etiquetado estos mapas como `map1`, `map2`, ... `map8`. El mapa de caminos ha sido generado a partir de los ficheros de puntos obtenidos mediante muestreo GPS cinemático.

- Mapas temáticos en formato raster: Modelo Digital de Elevación (MDE) del terreno (`map9`).
- Base de Conocimiento. Contiene información no espacial (y consecuentemente no asociada a ningún mapa) sobre experiencias y aspectos relacionados con la cartografía temática. En ella se encuentran los ficheros GPS de muestras de puntos.

Cuando hablamos de “información disponible”, no sólo consideramos los mapas descritos anteriormente. También aludimos a mapas no accesibles de forma inmediata, y obtenidos mediante operaciones escalares, algoritmos simples o mediante operaciones de álgebra relacional (principalmente proyecciones) sobre la tabla o tablas que almacenan datos almacenados en un mapa. A la cartografía obtenida de este modo se le denomina “cartografía derivada”. Los mapas de pendiente y orientación del terreno, así como los mapas de concavidad y convexidad son cartografía derivada del MDE.

Aunque el gestor de base de datos relacionales que incorpora *Idrisi32* (Database Workshop) ejecuta gran cantidad de operaciones, no permite la automatización de consultas sobre mapas. Por ello, las consultas que cuentan el número de ocurrencias de cada regla serán reemplazadas por una tabla de frecuencias obtenida mediante llamadas al módulo *Crosstab* empotradas en un script *.IML* (*Idrisi* Macro Lenguaje). Los archivos *.IML* se usan para automatizar la ejecución de proceso que requieren múltiples pasos analíticos simples. El fichero contiene los comandos y todos los parámetros necesarios para llevar a cabo las llamadas a los módulos *Idrisi* que ejecutan dichos pasos. *Crosstab* compara las clases de dos mapas, generando una tabla de frecuencias de las combinaciones de sus categorías.

Resulta evidente que para iniciar el proceso de DM, será necesario realizar un preproceso consistente en cambiar la estructura de almacenamiento de los mapas almacenados en formato vectorial a formato raster. Una vez realizada la transformación, para cada celda/localización de coordenadas *X* e *Y*, tendremos una observación de la coordenada *Z* (*map9*) y un valor/clase de Régimen de Humedad (*map10*), Régimen Térmico (*map11*), Índice de Evapotranspiración Potencial (*map12*), Recursos Hídricos Subterráneos, (*map13*), Uso de Suelos (*map14*), Unidad Edafológica (*map15*), Ecosistema Forestal Actual (*map16*) y Tiempos en Tramos de Caminos (*map17*)(Fig.5.1.).

Para esta conversión, se han utilizado los módulos *Polyras* y *Lineras*, seleccionando el atributo apropiado del fichero de valores de atributo polígono/línea (*.MDB*) asociado al mapa.

```
map10 := map1 a formato raster
map11 := map2 a formato raster
      :           :
map17 := map8 a formato raster
```

Map9, *map12* y *map17* almacenan valores numéricos que han de ser convertidos a categorías para poder ejecutar el algoritmo DM. Para ello,

realizamos una generalización de dominio no espacial mediante la cual, asignamos a cada intervalo numérico de la clase, un valor nominal de una escala ordinal previamente definida (módulo *Reclass*).

```
map18 := generalización de predominio no espacial sobre map9  
map19 := generalización de predominio no espacial sobre map12  
map20 := generalización de predominio no espacial sobre map17
```

5.5. Algoritmo DM.

Puesto que el propósito del proceso de DM es identificar reglas espaciales del tipo “Si una parcela tiene un suelo de tipo Cambisol, su uso es agricultura en un 60% de los casos”, la sintaxis de la regla será

```
Predicado_1(zona, categoria_1) → Predicado_2(zona, categoria_2) (c%)
```

cuya interpretación dependerá de los mapas/clases sobre los que esta regla sea aplicada. En nuestro algoritmo, imponemos dos restricciones: las categorías consideradas en la regla no pueden pertenecer a la misma clase y la zona de ocurrencia de la regla se restringe a una celda del grid.

Con estas restricciones, sólo hemos de definir el predicado *Es_del_tipo*(*l*, *categoria*). Si el predicado alude al ecosistema forestal con categoría “pino”, éste debe ser interpretado como “En la celda *l* hay pinos” o “El ecosistema forestal presente en la celda *l* es pinar”. La regla de asociación espacial es por consiguiente

```
Es_del_tipo(l, categoria_1) → Es_del_tipo(l, categoria_2) (c%)
```

El valor de confianza *c%* de una regla es el valor de la tabulación cruzada entre ambas categorías. El apoyo *s%* viene dado por la relación entre el número de celdas que satisfacen el antecedente de la regla y el número total de celdas que contiene el mapa (n_t). Los niveles mínimos preestablecidos para considerar que una regla es fuerte son $c\% = 30\%$ y $s\% = [0.5(n_t/N)*100]\%$, en donde *N* es el número de categorías de la clase.

En la sección anterior hemos especificado que cada regla de asociación espacial debe contener, al menos, un predicado espacial. Sin embargo, el predicado `Es_del_tipo(l, categoría)` no tiene ninguna connotación espacial. ¿Es esto una contradicción?, el método de tabulación cruzada utilizado para calcular la confianza y el apoyo, implica que la regla

$$\text{Es_del_tipo}(l, \text{categoría}_1) \rightarrow \text{Es_del_tipo}(l, \text{categoría}_2) \quad (c\%)$$

es equivalente a la regla:

$$[\text{Es_del_tipo}(l_1, \text{categoría}_1) \wedge \text{dentro_de}(l_1, l_2)] \rightarrow \text{Es_del_tipo}(l_2, \text{categoría}_2) \quad (c\%)$$

ya que l_1 es igual a l_2 para todas las ocurrencias de la tabulación.

Una vez establecido el conjunto de mapas y categorías, definida la regla de asociación espacial y el predicado espacial contenido en ella, y fijados los niveles de confianza y apoyo, podemos describir el algoritmo DM:

Cjto_Clasas := clases C_1, C_2, \dots, C_9 definidas en $\text{map}C_1, \text{map}C_2, \dots, \text{map}C_9$

Cjto_Categorías_ C_i := t_1, t_2, \dots categorías de la clase C_i

Para cada par $(C_i, C_j) \{C_i, C_j \in \text{Conjunto_de_Clases}, i \neq j\}$

Tabla := "CROSSTAB x $\text{map}C_i \text{map}C_j$ 2 n"

Para cada par

$(t_m, t_n) \{t_m \in \text{Conjunto_Categorías_}C_i, t_n \in \text{Conjunto_Categorías_}C_j\}$

$$\text{Si } \frac{\text{Tabla}[t_m, t_n]}{\text{N}^\circ \text{ total de celdas } t_m} \geq c\% \wedge \frac{\text{N}^\circ \text{ total de celdas } t_m}{\text{N}^\circ \text{ total de celdas}} \geq s\%$$

Entonces

$$\text{Es_del_tipo}(l, \text{categoría}_1) \rightarrow \text{Es_del_tipo}(l, \text{categoría}_2)$$

es una regla fuerte

Las reglas obtenidas después de llevar a cabo el proceso de DM confirman algunas asociaciones entre categorías ya establecidas en trabajos anteriores [Mol97] en los que fueron modeladas relaciones entre variables climáticas mediante un ajuste log-lineal de frecuencias. Concretamente, encontramos

asociaciones muy fuertes entre los regímenes térmicos “subtropical cálido” ($c = 87\%$) y “templado cálido” ($c = 70\%$) con el régimen de humedad “mediterráneo húmedo”, y valores de ETP (800-1000 litros/m²) asociados al ecosistema forestal “encinar” ($c = 60\%$) y “chopera” ($c = 36\%$).

El proceso ha descubierto también una fuerte relación entre los suelos de tipo “fluvisol”, “arenosol” y “luvisol”, con los ecosistemas forestales “Matorral Mediterráneo” ($c = 31\%$), “Eucaliptal” ($c = 31\%$) y “Encinar” ($c = 35\%$) respectivamente.

Entre las reglas identificadas como fuertes, ninguna de ellas indica una relación entre el tiempo necesario en un desplazamiento y el resto de las clases involucradas en el proceso. Si éste se repite con un nivel $c\% = 20\%$, observamos que los ecosistemas forestales “pinar” y “chopera” están asociados a caminos con valores de tiempo medio-altos con una confianza del 21%. Estos ecosistemas suelen presentar altas concentraciones de individuos por unidad de superficie, lo que induce a pensar que los tiempos altos podrían estar asociados no a determinadas especies forestales, sino a valores de altas densidades. Esta idea se confirma si analizamos la tabulación cruzada entre el mapa de densidad generalizado (map_{21}) obtenido del mapa de ecosistemas forestales y el mapa de tiempos (map_{20}), ya que densidades altas se asocian a tiempos elevados con una confianza del 33%.

```
map21 := map7 a formato raster
```

```
map22 := generalización de predominio no espacial sobre map21
```

El conocimiento adquirido a través del descubrimiento de reglas permite establecer que el parámetro m del modelo de coste es el nivel de densidad de especie forestal que el móvil encuentra en su desplazamiento. La Fig.5.1. esquematiza los pasos del proceso de minería.

5.6. KDSD: Asignación de fricciones.

Una vez modelada la función de coste, generamos los mapas de pendientes y de orientación del terreno usando el módulo *Surface*.

```
vecinos(X,Y) := {(X,Y+1), (X-1,Y), (X,Y-1), (X+1,Y)}
map26(X,Y) := pendiente máxima de map9(X,Y) y map9(vecinos(X,Y))
```

El mapa de coste será generado sólo para el rectángulo de coordenadas UTM (440950,4188150) y (477810,4206170), por lo que hemos de seleccionar el área de estudio mediante el módulo *Window*.

```
Área:= rectángulo [(440950,4188150), (477810,4206170)]
map27(X,Y) := map24(X,Y) si (X,Y) en área
map28(X,Y) := map26(X,Y) si (X,Y) en área
map29(X,Y) := map22(X,Y) si (X,Y) en área
```

El mapa de fricción debida a la densidad forestal, será obtenido mediante generalización no espacial, asignando un valor de coste asociado a cada intervalo de densidad. Para ello utilizaremos el módulo *Reclass*.

```
map31(X,Y) := generalización predominio no esp. sobre map29(X,Y)
```

De la misma manera, los valores de pendiente comprendidos en el intervalo (0%,3%] son generalizados al concepto de “zona plana”, valores en (3%,7%] a “pendiente suave”, valores en (7%,15%] a “pendiente moderada”, valores en (15%,30%] a “pendiente fuerte”, valores en (15%,45%] a “pendientes muy fuertes” y valores con pendiente $\geq 45\%$ a “zona abrupta”. La siguiente jerarquía de generalización asigna valores de fricción (*stated_friction*) a cada tipo de pendiente: fricción 1 a zonas planas y zonas con pendiente suave, fricción 1.5 a pendientes moderadas, fricción 2 a pendientes fuertes y fricción 3 a pendientes muy fuertes y zonas abruptas.

```
map30(X,Y) := generalización predominio no esp.sobre map28(X,Y)
```

5.7. KDSD: Cálculo de Coste.

El mapa de coste debido a la componente anisotrópica será generado a partir del mapa de fricciones *map30*, el mapa de orientaciones y la función $a(\delta)$ mediante el módulo *Varcost*. Como vimos en Sección 1.5., este módulo,

además de incluir la función anisotrópica $a(\delta) = \text{Cos}^k(\delta)$, permite el uso de funciones definidas por el usuario. Para ello, debemos crear un fichero ASCII o binario de valores reales que describan los 7200 valores que toma la función desde 0 a 360° en incrementos de 0.05°. El diseño de la función anisotrópica que genera dicho fichero se ha realizado con la metodología descrita en Capítulo 2. Los costes son calculados desde el punto inicial hasta el destino, generándose una superficie en donde el coste es medido como el mínimo esfuerzo para moverse a través de la superficie. Dada una celda, los movimientos se realizan en las 8 direcciones principales o “movimiento del rey” (en alusión a las posibles jugadas que puede realizar el rey de ajedrez), lo que genera un coste de valor $stated_friction \cdot a(\delta)$, siendo $\delta(X,Y) := dirección(X,Y) - map27(X,Y)$.

```
punto_inicial:= punto (459380, 4199578)
vecinos(X,Y) := {(X,Y+1), (X-1,Y+1), (X-1,Y), (X-1,Y-1), (X,Y-1),
                (X+1,Y-1), (X+1,Y), (X+1,Y+1)}
map32(X,Y) := mínimo {map30(X,Y) * a(dirección(vecinos(X,Y)) -
                        map27(vecinos(X,Y))}
```

Por último, el coste total ($map33$) será la suma del coste debido al ecosistema forestal y del coste debido a la orografía del terreno (Fig.5.3.).

```
map33(X,Y) := map31(X,Y) + map32(X,Y)
```

La Fig.5.2. describe la totalidad del proceso KDSD. La Fig.5.4. muestra el mapa de coste total considerando las discontinuidades del terreno como un factor no direccional. Comparando las Fig.5.6. y Fig.5.7. observamos los distintos comportamientos de ambos modelos .

La cuestión final que se plantea al concluir un proceso KDSD es la misma que se planteaba inicialmente ¿Cuál es la información disponible? Las operaciones realizadas han generado un conjunto de mapas (pendiente, orientación, fricciones, costes ...) e información no referenciada (reglas). De toda esta cartografía, sólo el mapa de costes y el mapa de fricciones debido a las discontinuidades orográficas serán incluidos en nuestro SIG, ya que el resto de

los mapas pueden ser obtenidos como cartografía derivada mediante operaciones simples. Asimismo, las reglas asumidas como válidas serán incluidas en nuestra base de conocimiento, ya que podrían resultar útiles en futuros procesos KDSD.

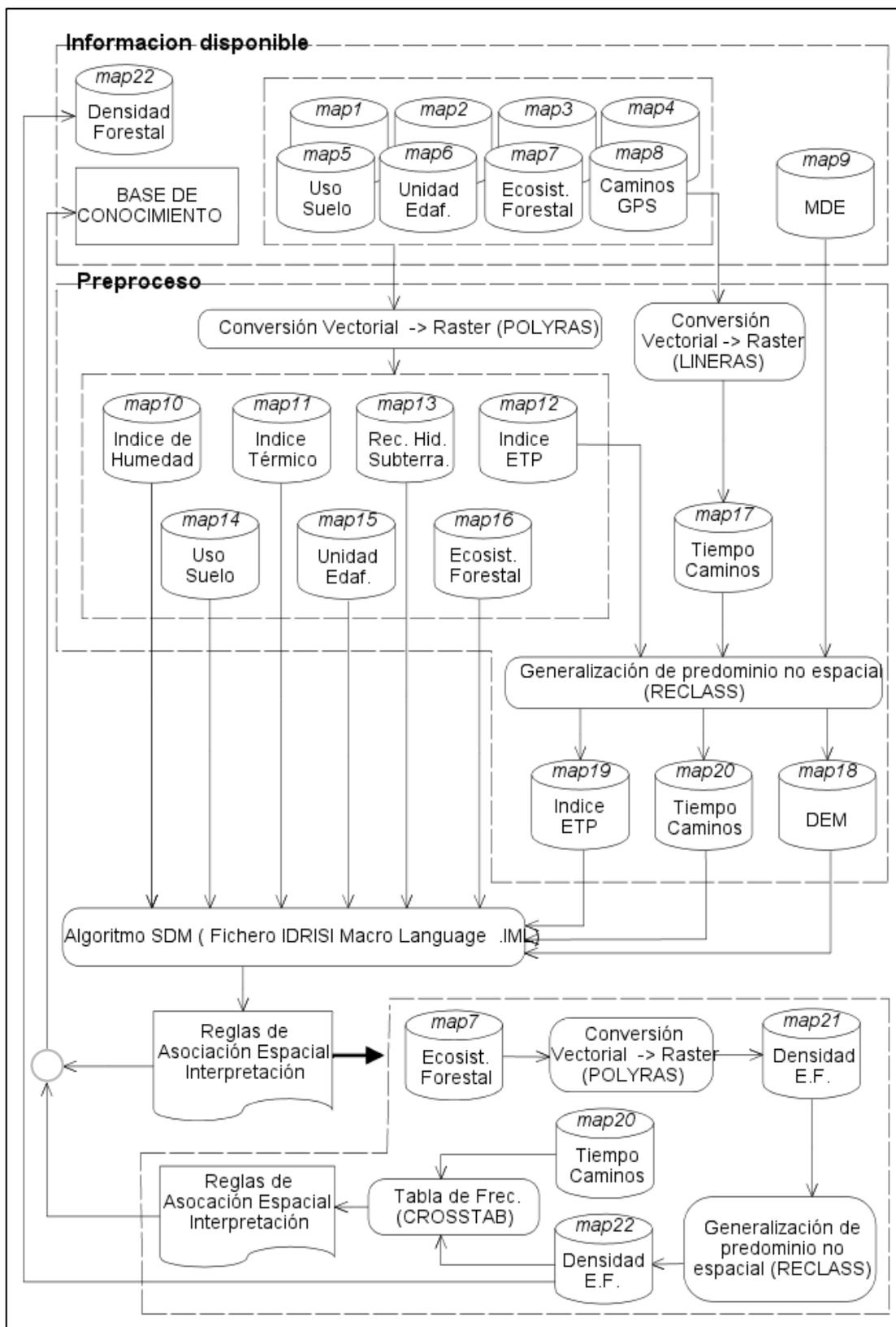


Figura 5.1. Diagrama del proceso DM

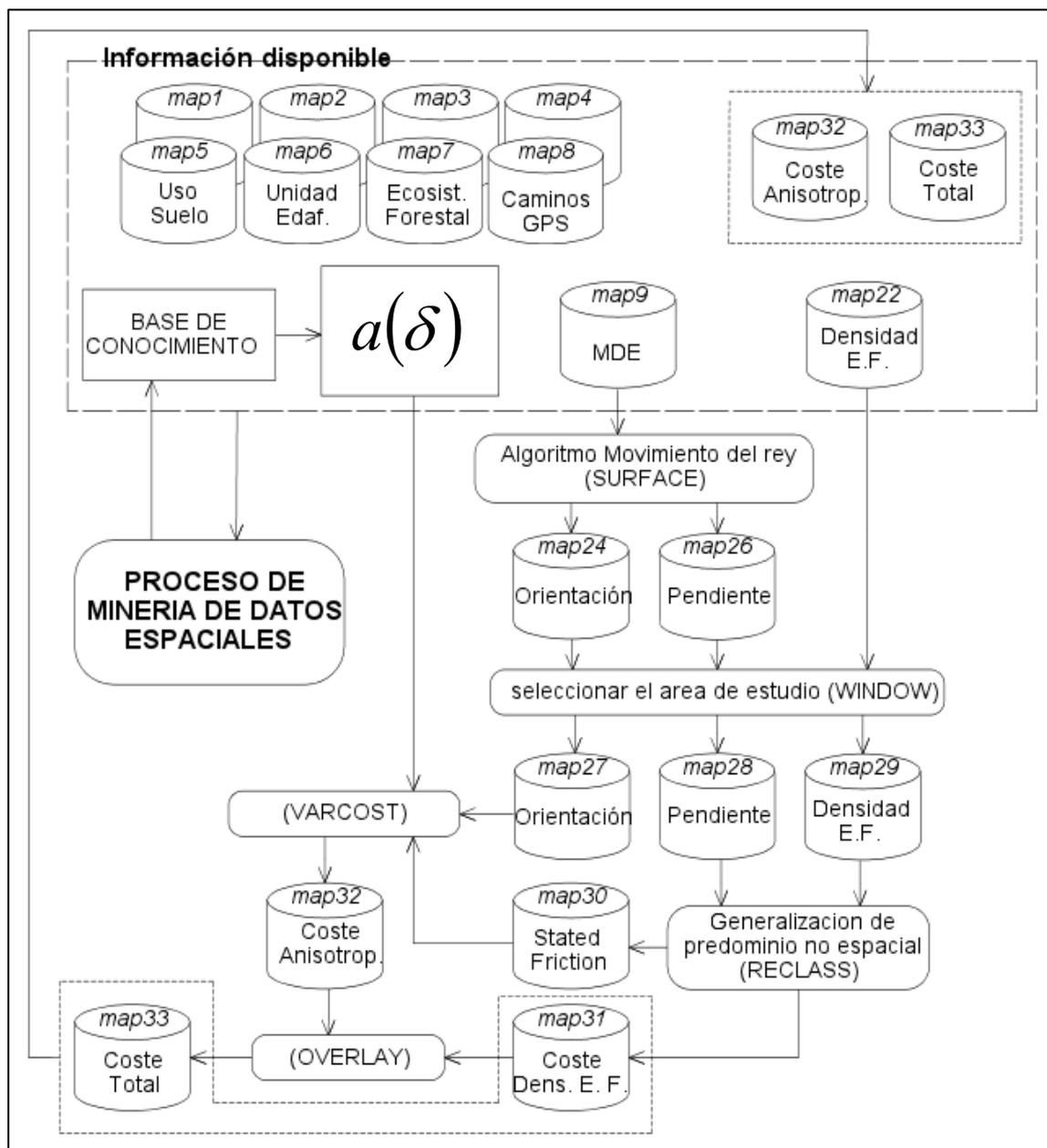


Figura 5.2. Diagrama del proceso KDSD completo

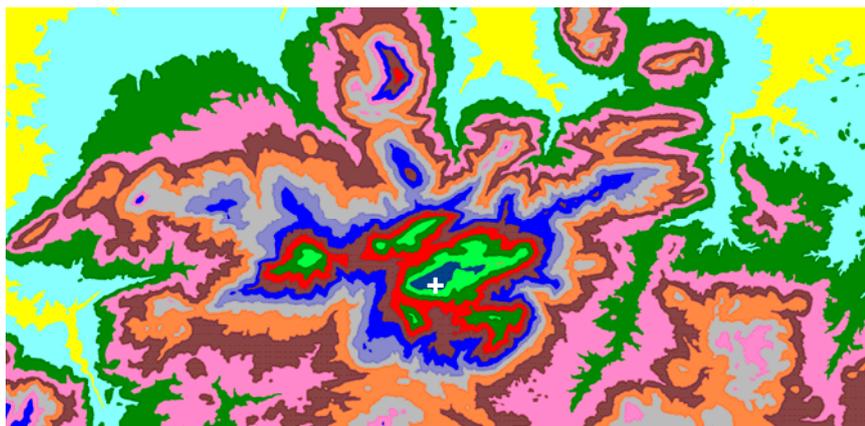


Figura 5.3. Modelo digital de elevación del terreno (Map9). El símbolo “+” representa la posición del punto_inicial

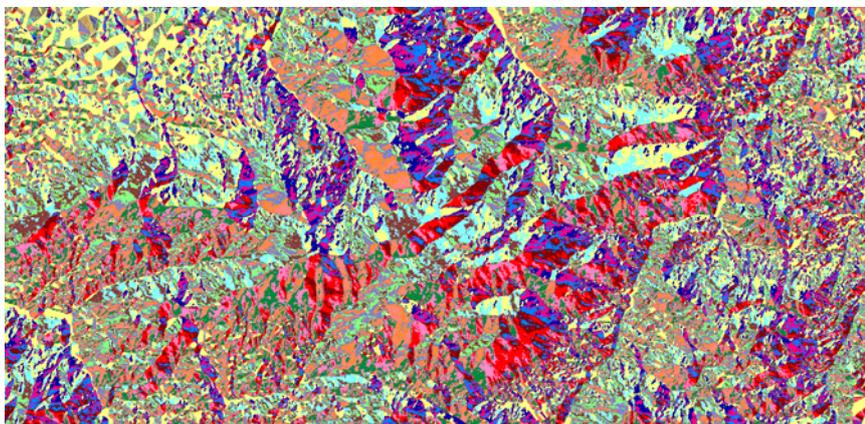


Figura 5.4. Orientación del terreno (Map 27)

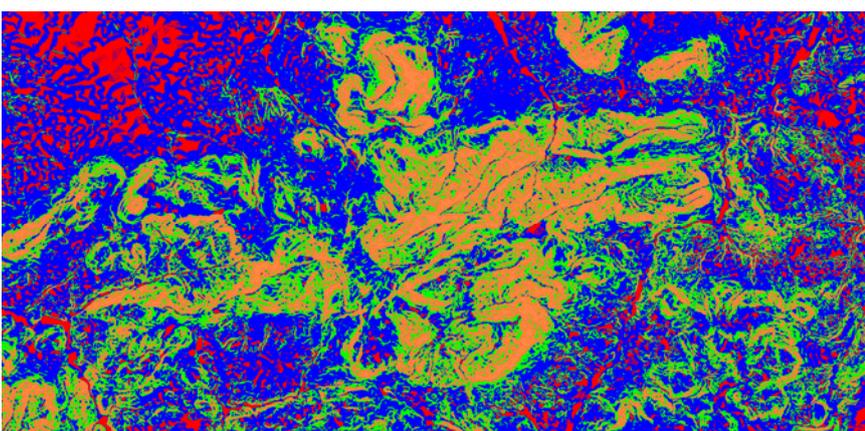


Figura 5.5. Pendiente del terreno (Map 28)

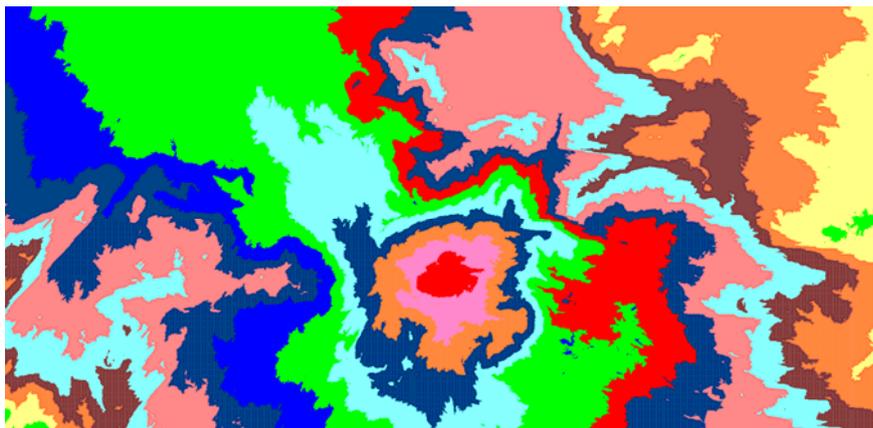


Figura 5.6. Coste total final (Map32)



Figura 5.7. Coste total final (isotrópico)

Capítulo VI

Conclusiones

Finalizamos la memoria con un resumen de las conclusiones obtenidas y de las principales aportaciones desarrolladas en ella. Asimismo, se presentan las publicaciones relacionadas y se comentan las líneas de investigación futuras que se derivan.

6.1. Resultados obtenidos

Cuando la visualización y observación del proceso analizado no describen de forma adecuada su comportamiento espacial, es necesario el diseño de un modelo matemático que lo explique y sintetice. Habitualmente, los fenómenos espaciales son el resultado de la suma de dos tipos de efectos: globales y locales. Los efectos globales aparecen cuando existe en el espacio variaciones significativas del valor medio del proceso. Los efectos locales se manifiestan cuando el valor medio en un punto es influenciado por el valor media en los puntos vecinos.

El diseño de modelos espaciales sólo es posible bajo asunción de homogeneidad en los efectos globales y/o locales. Si no se asume ningún tipo de estacionalidad, sería extremadamente complicado modelizar el fenómeno debido a la gran cantidad de parámetros necesarios para la formalización matemática de su comportamiento. Además, es necesario determinar si el fenómeno presenta tendencias anisotrópicas, ya que este hecho invalidaría la asunción de modelos definidos para funciones isotrópicas, y haría necesario utilizar transformaciones espaciales que permitan obtener funciones anisotrópicas a partir de modelos isotrópicos. Por este motivo, el modelado de procesos espaciales requiere tanto un análisis de la presencia de estacionalidad en las variaciones globales y locales como un estudio direccional para valorar la existencia de una componente anisotrópica.

Son muchas y muy variadas las herramientas utilizadas para el análisis de estacionalidad: triangulación de Delaunay, medias móviles espaciales, k-funciones y funciones de covarianza (covariogramas), variogramas, estadísticos de autocorrelación espacial (*Moran's I*, *Geary's C*, *Geti's G*), autocorrelogramas, tendencia espacial....

Las técnicas para el estudio de la anisotropía demuestra que, por el contrario, son escasas, y se limitan a la detección de efectos direccionales mediante la inspección de variogramas experimentales en diferentes direcciones. El examen de estos variogramas "a ojo" sólo proporciona una comprobación informal de la existencia de anisotropía y la determinación aproximada de la

dirección en que ésta se manifiesta.

Como consecuencia de esto, en esta memoria hemos presentado:

- Un estudio sobre el estado del arte de la anisotropía, especialmente en el aspecto relacionado con el modelado espacial de datos.
- El análisis de la problemática asociada a la determinación de la existencia y cuantificación de la componente direccional en patrones.
- La adaptación de los estadísticos circulares existentes para el análisis del comportamiento direccional de fenómenos naturales dentro del proceso general de análisis de las propiedades de primer orden (homogeneidad) y segundo orden (autocorrelación y anisotropía) en patrones.
- Un test para la existencia de direccionamiento en patrones que presentan dependencia espacial asumiendo normalidad de la población.
- Determinación de la dirección dominante en patrones, una vez establecida la existencia de anisotropía.
- Un método para el diseño de óvalos generadores en los que la superficie se distribuye de forma Normal alrededor de su dirección dominante

Estos métodos han sido particularizados y adaptados a imágenes digitales, y más concretamente a imágenes satélite.

6.2. Publicaciones asociadas al trabajo desarrollado en esta memoria

A continuación se presenta un listado de las publicaciones asociadas a esta memoria.

Revistas internacionales:

- Molina A., Muñoz M.D., Tovar J.; “Design of Generating Ovals for Anisotropic Buffer Operator”, *Computer & Geosciences*, Elsevier Science (Reino Unido) (en fase de revisión).

Capítulos de libro:

- Molina A., Muñoz M.D. “Análisis 2D. Aplicaciones. En: Plataforma avanzada de modelado paramétrico en CAD”, Universidad de Jaén, pp. 333-346, 2004.

Publicaciones en congresos internacionales:

- Muñoz M.D., Cuartero A., “Aplicaciones y tendencias de los sistemas de información geográficos” en Congreso Internacional de Ingeniería Gráfica, INGEGRAF´2000, Valladolid, 2000.
- Muñoz M.D., Molina A., “Simulation and visualization of anisotropic expansion phenomena” en actas AGILE Conference on Geographic Information Science, Palma de Mallorca, pp. 557-568, 2002.
- Molina A., Muñoz M.D., “Visualization of environmental phenomena using GIS and directional buffer operator” en I International Conference on Multidisciplinary Information Sciences and Technologies, Mérida (Badajoz), pp. 467-471, 2006.

Publicaciones en congresos nacionales:

- Muñoz M. D., Molina A., Feito F., “ALES: Una aproximación a los Sistemas Expertos Espaciales” en actas IX Congreso Español de Informática Gráfica CEIG´99, Jornadas de Sistemas de Información Geográficos JSIG´2002, pp. 371-372, 1999.
- Molina A., Muñoz M. D., “Obtención de un modelo cartográfico de coste mediante minería de datos” en actas II Jornadas de Sistemas de Información Geográficos JSIG´2002, pp. 49-59, 2002.
- Muñoz M.D., Molina A., “Diseño de funciones anisotrópicas mediante análisis de datos” en actas III Jornadas Regionales de Informática Gráfica, Jaén, pp. 63-74, 2002.
- Muñoz M.D., Molina A.,”Implementación del operador buffer anisotrópico” en actas I Jornada de Sistemas de Información Geográficos JISIG´01, Almagro (Ciudad Real), pp. 85-94, 2001.

6.3. Trabajos futuros

En relación con los procesos de simulación y visualización, los autores tienen abiertas dos líneas de investigación:

- Particularización de los fenómenos de expansión a fenómenos de transporte.
- Análisis de interacciones (continuación, creación, disipación, división y fusión) entre diferentes objetos.

Bibliografía

- [Agra93] Agrawal R., "Mining association rules between sets of items in large databases" en actas International Conference on Management of Data ACM-SIGMOD, pp. 207-216, 1993.
- [Ara89] Aranoff S., "Geographic information systems: a management perspective", WDL Publications, Ottawa, 1989.
- [Bat78] Batschelet E., "Second-order statistical analysis of directions" en Schmidt-Koenig K, Keeton WT, editores Animal, Migration, Navigation and Homing, Springer, pp. 3-24, Berlin, 1978.
- [Bat81] Batschelet E., "Circular statistics in biology", Academic Press, Londres, 1981.
- [BaiGat96] Bailey T.C., Gatrell A.C., "Interactive spatial data analysis", Longman, London, 1996.
- [BlaEbe+04] Black M., Ebener S., Najera P., Vidaurre M. y El Morjani Z., "Using GIS to measure physical accessibility to health care", en actas International Health Users Conference '04, 2004.
- [Ber00] Bergougnoux P., "Geographic information systems and intelligent transportation systems" *Geoinfo*, 4(2), pp. 123-125, 2000.
- [BerKre⁺00] Berg M., Kreveld M., Overmars M., Schwarzkopf O., "Computational geometry algorithms an applications second edition", Springer Verlag, 2000.
- [Boo07] Boost C++ Libraries. <http://www.boost.org>
- [Bos95] Bosque Sendra J., "Sistemas de información geográfica", Rialp, 2ª Edición, Madrid, 1995.
- [BosEsc⁺94] Bosque J., Escobar F.J., García E., Salado M.J., "Sistemas de información geográfica; practicas con Pc *ArcInfo* e *Idrisi*", Addison_Wesley Iberoamericana, Rama, 1994.
- [BraWes90] Bracken y Wester, "Information technology in geography and planning" en *Principles of geographic information systems*, Routledge, Londres, 1990.

- [Bur88] Burrough P., "Principles of geographical information systems for land resources assessment", Oxford University Press, Oxford, 1988.
- [Bur90] Burrough P.A., "Methods of spatial analysis in GIS", International Journal of Geographical Information Systems, 4, pp. 221-238, 1990.
- [BuzDur97] Buzai G. y Durán D., "Enseñar e investigar con sistemas de información geográfica", Troquel, Buenos Aires, Argentina, 1997.
- [CasMon⁺07] Casciola G., Montefusco L.B. y Morigi S., "The regularizing properties of anisotropic radial basis functions applied mathematics and computation", 190(2), pp.1050-1062, 2007
- [CebMar86] Cebrián de Miguel J.A. y MARK D., "Modelos topográficos digitales en métodos cuantitativos en geografía: enseñanza, investigación y planeamiento", AGE, Madrid, 1986.
- [Chat95] Chatfield, C., "Model uncertainty, data mining and statistical inference (with discussion)", Journal of the Royal Statistical Society, 158, pp. 945-957, 1995.
- [CheZhe06] Chen J. y Zheng Y., "Redesign of a conformal boundary recovery algorithm for 3D Delaunay triangulation", Zhejiang University Press, 7(12), pp.2031-2042, 2006.
- [Cre91] Cressie N.A., "Statistics for Spatial Data". John Wiley & Sons, New York, 1991.
- [Cre93] Cressie N., "Statistics for spatial data", Chichester: John Wiley, 1993.
- [CoxHin74] Cox D. y Hinkley D.V., "Theoretical statistics", Chapman and Hall, Londres, 1974.
- [DinTur⁺01] Dinh H. Q., Turk G. y Slabaugh G., "Reconstructing surfaces using anisotropic basis functions", en actas International Conference on Computer Vision, ICCV, Vancouver, Canada, pp. 606-613, 2001.
- [Doe87] DoE (Department of Environment), "Handling geographic information", Informe de la Comisión Inverstigadora dirigida por Lord Chorley, Oficina de documentación de Su Majestad, Londres, 1987.
- [EbeMor⁺05] Ebener S., El Morjani Z., Ray N., Black M., "Physical accessibility to health care: from isotropy to anisotropy" 2005. <http://www.gisdevelopment.net/magazine/years/2005/jun/care.htm>

- [Esr08] ESRI - The GIS software leader.
<http://www.esri.com>
- [Est95] Ester M., "Knowledge discovery in large spatial databases: focusing techniques for efficient class identification", en actas Symposium on Large Spatial Databases, SSD'95, pp. 67-82, 1995.
- [FAO94] F.A.O.: The digital soil map of the world, 1994.
<http://edcwww.cr.usgs.gov/glis/hyper/guide/fao#fao91994>
- [Fay93] Fayyad U., "Image database exploration: progress and challenges", en actas Knowledge Discovery in Databases Workshop, pp. 14-27, 1993.
- [Fay96] Fayyad U., "From data mining to knowledge discovery: an overview" en Advances in knowledge discovery and data mining, Cambridge, MA: MIT Press, 1996.
- [Fis53] Fisher R.A., "Dispersion on a sphere", en actas Royal Society, A, V (217), pp. 295-305, Londres, 1953.
- [FotRog94] Fotheringham S. y Rogerson, P., "Spatial analysis and GIS", Taylor&Francis, Londres, 1994.
- [FayPia⁺96] Fayyad U., Piatetsky-Shapiro G. y Smyth P., "The KDD process for extracting useful knowledges from volumes of data", Communication of the ACM 39, pp. 27-34, 1996.
- [FayPia⁺96b] Fayyad U., Piatetsky-Shapiro G. y Smyth P., "From data mining to knowledge discovery: A overview", en "Advances in knowledge discovery and data mining" Fayad et al. editores, AAAI Press, pp. 1-30, 1996.
- [Gat08] The Gateway to Astronaut Photography of Earth
<http://eol.jsc.nasa.gov/scripts/sseop/photo.pl?mission=STS078&roll=729&frame=4&QueryResultsFile=1058183394569.tsv>
- [Get94] Getis A., "Homogeneity and proximal databases" en "Spatial analysis and GIS", Fotheringham S. y Rogerson P. editores, Taylor&Francis, Londres, 1994.
- [Gli97] Glimour C., "Statistical themes and lessons for data mining. Data mining and knowledge discovery", 1, pp.11-28, 1997.
- [Goo87] Goodchild M.F., "A spatial analytical perspective on GIS", International Journal of Geographical Information Systems. 1(4), pp. 335-354, 1987.
- [GooO'Ro04] Goodman J.E. y O'Rourke J., "Handbook of discrete and computational geometry", CRC Press, Boca Raton, 2004.

- [Gra08] Manual de *Grass*.
http://grass.itc.it/grass63/manuals/html63_user
- [GutGou08] Gutierrez J. y Gould M., "Sistemas de información geográfica", Sintesis, 2008.
- [Hal52] Hald A., "Statistical theory with engineering applications", Chapman and Hall, Londres, 1952.
- [Hot31] Hotelling H., "The generalization of student's ratio", *Annals of Mathematical Statistics*, 2, pp. 370-378, 1931.
- [HubChe⁺08] Hubball D., Chen M. y Grant P.W., "Image-based aging using evolutionary computing", *Computer Graphics Forum*, 27(2), pp. 607-616, 2008.
- [Ins08] Instituto mejicano del petróleo, 2008.
http://www.insituh.com/bajadas/Yacimientos_naturalmente_fracturados_IMP.pdf
- [Idr03] *Idrisi* Kilimanjaro. Guide to GIS and image processing.
<http://www.clarklabs.org>
- [JSC08] JSC Digital Image Collection
<http://images.jsc.nasa.gov/search/search.cgi?photoidsearch=Go&hitsperpage=5&photoid=STS058-73-009&submit.x=4&submit.y=7>
- [KopHan95] Kopersky J. y Han J., "Discovery of spatial association rules in geographic information databases", en *actas Symposium on Large Spatial Databases, SSD'95*, pp.47-66, 1995.
- [Lan02] Lantada N., "Sistemas de información geográfica. Prácticas con *ArcView*", Universidad Politécnica de Cataluña, 2002.
- [Lop00] López, C., "On the improving of elevation accuracy of Digital Elevation Models: a comparison of some error detection procedures", 4(1), *Transactions on GIS*, pp. 43-64, 2000.
- [Lu93] Lu, W., "Discovery of general knowledge in large spatial databases. Proceedings far east" en *actas Workshop in GIS*, pp. 275-289, 1993.
- [Mag91] Maguire D.J., "An overview and definition of GIS" en *Geographical Information Systems: Principles and Applications*, Maguire D.J., Goolchid M.F.y Rhind D.W., editores, Logman, Londres 1991.
- [Mar75] Mardia K.V., "Statistics of directional data", *Journal of Royal Statistical Society*, 37, pp. 349-393, 1975.

- [Mar08] Martínez Val, J., "Introducción a la lógica virtual", 2008.
http://www.imageandart.com/tutoriales/morfologia/logica_visual/inex5.htm
- [Men04] Mena M., "Aplicaciones de estadística circular a problemas de ciencias naturales", Akadia, 2004.
- [Mol97] Molina, A., "Aportaciones al estudio de los Sistemas de Información Geográfica y técnicas estadísticas asociadas" Tesis Doctoral, 1997.
- [Mol00] Molina A., "A vector approach to the analysis of (patterns with) spatial dependence" en Advances in pattern recognition, Lectures Notes in Computer Science ,Springer Verlag 2000.
- [Mol02] Molina A., "Integrating GIS, circular statistics and KSDS for modelling spatial data", Geographical and Environmental Modelling, 6(1), Taylor&Francis (U.S.A.), 2002.
- [Mol02] Molina A., "A method for testing anisotropy and quantifying its direction in digital images", Computer & Graphics, Elsevier Science ,Reino Unido, 2002.
- [MolAgu00] Molina A., Aguilera J.J., "Modelo gráfico de transporte de agua a través del Estrecho de Gibraltar a partir de imágenes satélite" en actas II Jornadas de análisis y simulación numérica de intercambio de masas de agua a través del Estrecho de Gibraltar, 2000.
- [MolFei98] Molina A., Feito F.R., "Design of anisotropic functions within analysis of data" en actas Computer Graphics International CGI'98, pp.726-729, 1998.
- [MolMuñ03] Molina A., Muñoz M.D., "Simulación y visualización de fenómenos de expansión anisotrópicos" en actas IX Conferencia Iberoamericana de SIG, SIBSIG, 2003.
- [MolMuñ04] Molina A., Muñoz M.D., "Análisis 2D. Aplicaciones" en Plataforma avanzada de modelado paramétrico en CAD, 2004.
- [MolMuñ06] Molina A., Muñoz M.D., "Visualization of environmental phenomena using GIS and directional buffer operator " en actas I International Conference on Multidisciplinary Information Sciences and Technologies, 2006
- [MuñCua00] Muñoz M.D., Cuartero A., "Aplicaciones y tendencias de los sistemas de información geográficos" en actas Congreso Internacional de Ingeniería Gráfica, 2000.

- [MuñMol01] Muñoz M.D., Molina A., "Implementación del operador buffer anisotrópico" en actas I Jornada de Sistemas de Información Geográficos, 2001.
- [MuñMol02] Muñoz M.D., Molina A., "Diseño de funciones anisotrópicas mediante análisis de datos" en actas III Jornadas Regionales de Informática Gráfica, 2002.
- [MuñMol02] Muñoz M.D., Molina A., "Simulation and visualization of anisotropic expansion phenomena" en actas 5th AGILE Conference on Geographic Information Science, 2002.
- [MolMuñ02] Molina A., Muñoz M.D., "Obtención de un modelo cartográfico de coste mediante minería de datos" en actas II Jornadas de Sistemas de Información Geográficos JSIG'2002, 2002.
- [Ncg90] NCGIA Core Curriculum, Universidad de California, Santa Bárbara, 1990.
- [NgHan94] Ng R. y Han J., "Efficient an effective clustering method for spatial data mining", en actas International Conference on Very Large Data Bases, pp. 144-155,1994.
- [OkaBoo+00] Okabe A., Boots B. y Sugihara K., "Spatial tessellations-concepts and applications of Voronoi diagrams", Wiley, New York, 2000.
- [Oya96] Oyamburu Fernández, I., "Estudio de la anisotropía sísmica cortical en el sur de la Península Ibérica a partir de la polarización de las ondas S", 1996.
<http://www.minas.upm.es/fundacion/jgs/trabajos/99c02.htm>
- [Pia 91] Piatetsky-Shapiro, G., "Knowledge discovery in real databases", en actas IJCAI-89 Workshop, AI Magazine, 11(5), pp. 68-70, 1991.
- [SerUbe+00] Servigne S., Ubeda T., Puricelli A. y Laurini R., "A methodology for spatial consistency improvement of geographic databases", Geoinfo 4(1), pp. 7-34, 2000.
- [Sch78] Schow G., "Estimation of the concentrator parameter in Von Misses-Fisher distributions", Biometrika, 65, pp. 369-375, 1978.

-
- [Str89] Strang G., "Introduction to applied mathematics", Wellesley-Cambridge Press, pp. 510, 1989.
- [Tob70] Tobler W.R., "A computer movie simulating urban growth in the Detroit region", *Economic Geography*, 1970.
- [Tom67] Tomlinson R.F., "An introduction to the geo-information system of the canadian land inventory", Departamento canadiense de desarrollo rural, Ottawa, Canada, 1967.
- [Tom87] Tomblinson R.F., "Review of North American experience of current and potencial uses of Geographic Information Systems", en DoE (Departamento de Medioambiente), Handling Geographic Information, Oficina de documentación de su Majestad, Londres, 1987.
- [TouChe01] Toutin T. y Cheng P., "DEM generation with ASTER stereo data, 10(6), *Earth Observation Magazine*, pp. 10-13, 2001.
- [Wac95] Wackernagel H., "Multivariate geostatistics", Springer-Verlag, Berlin, 1995.
- [WadYou⁺94] Wadge G., Young P.A.V. y Mckendrick I.J., "Mapping lava flow hazards using computer simulation", *J. Geophys. Res.* 99, pp. 489-504, 1994.
- [WatWill56] Watson G. y Williams E., "On the construction of significance test on the circle and the sphere", *Biometrika*, 43, pp. 344-352, 1956.

Tabla de símbolos

Símbolo	Descripción
f	Función anisotrópica <i>Varcost</i>
k	Coefficiente definido por el usuario
α	Ángulo diferencia entre la dirección en la que el fenómeno alcanza su máximo valor de fricción y la dirección real de actuación
f	Función anisotrópica <i>Disperse</i>
$A(x)$	Función relativa a la función de Bessel
c	Factor de corrección
$E_{i,j}$	Expansión durante el período t_i en la celda j
$f(\Phi)$	Función de densidad de Von Misses
$F(\alpha, \beta, \Phi)$	Función de distribución de Von Misses
$g(\Phi)$	Radio polar del óvalo generador
$g_l(\Phi)$	Radio polar del óvalo generador de longitud máxima l
$G(l, \theta)$	Óvalo generador
$I_n(x)$	Función de Bessel modificada de orden n
\hat{I}	Estimador de máxima verosimilitud de I
$I^{(\tau)}$	Estadístico de Moran's I a una distancia τ
k	Parámetro de concentración
\hat{k}	Estimador de máxima verosimilitud de k
l	Longitud máxima del óvalo generador
\bar{m}	Vector media de muestra monomodal
\bar{m}_2	Vector media de muestras bimodales

Tabla de símbolos

Símbolo	Descripción
m_c	Segundo momento trigonométrico del coseno
m_s	Segundo momento trigonométrico del seno
r	Longitud del vector media
r_c	Longitud del vector media corregido
r_2	Longitud del vector media de muestra bimodal
T^2	Test estadístico de direccionalidad
$T^2(\alpha)$	Valor crítico de T^2 con un nivel de significación α
v	Número de modas
v	Velocidad de las fuerzas externas
$\delta_{i,j}^{(\tau)}$	Conexión espacial entre las i -ésimas y j -ésimas celdas a una distancia τ
$\bar{\Phi}$	Ángulo media de la muestra
$\bar{\Phi}_2$	Ángulo media de la muestra bimodal
$\bar{\Phi}_1$	Ángulo media de la muestra bimodal simétrica
μ_1	Vector media de la población
θ	Dirección dominante del ovalo generador /fuerzas externas
ρ_1	Longitud del vector media de la población
$\hat{\rho}_1$	Estimador de máxima verosimilitud de ρ_1
τ	Distancia del análisis espacial
ψ	Dirección de la anisotropía
\oplus	Operador Suma de Minkowski

Tabla de símbolos

.0041	.0040	.0009	.0007	.0010	.0011	0	-.0019	.0032	.0046	.0036	-.0004	.0023	.0022	.0018	.0031	.0033	-.0001	-.06	-.0022	.0023	.0039	.0028	.0034	.0055	.0043	.0048
.0043	.0047	.0020	.0011	.0005	.0017	.0023	.0026	.0019	.0019	.0023	-.0013	.0033	-.0008	-.0007	.0031	.0022	.0033	.0035	-.0012	.0010	.002	.0003	-.0005	.0040	.0029	.0049
.0013	.0034	.0014	-.0002	.0015	.0021	.0016	.0009	.0011	.0034	.0018	-.0020	.0018	-.0015	-.0020	.0050	.0033	.0032	.0017	-.0018	.0022	-.0018	-.0023	-.0032	.0006	.0034	.0058
.0012	.0060	.0012	.0004	.0017	.0015	.0012	.0008	.0030	.0021	.0017	0	.0011	.0001	-.0017	.0027	.0052	.0055	.0047	-.0006	.0030	-.0034	-.0011	-.0012	-.0028	.0041	.0047
.0008	.0030	-.0006	.0007	.0006	-.0003	.0013	.0009	.0030	.0022	.0014	.0014	.0010	.0009	.0010	.0024	.0034	.0054	.0049	-.0007	.0022	-.0036	-.0014	-.0016	-.0045	.0037	.0049
.0013	.0007	-.0015	-.0005	-.0009	.0006	.0012	.0017	.0023	.0020	.0009	.0014	.0014	-.0001	.0012	.0020	.0028	.0037	.0022	-.0021	.0013	-.0024	-.0009	-.0028	-.0033	.0036	.0050
.0021	.0027	.0007	.0021	0	.0007	.0018	.0019	.0019	.0014	.0021	-.0007	-.0002	-.0001	-.0002	.0020	.0035	.0027	.0014	-.0024	-.0002	-.0004	.0027	-.0009	.0012	.0028	.0050
.0036	.0030	-.0001	.0004	.0020	.0008	.0012	.0017	.0015	.0008	.0029	-.0026	-.0023	.0006	-.0010	.0006	.0052	.0030	.0013	-.0018	.0004	.0009	.0037	0	.0037	.0034	.0063
.0033	.0026	.0019	.0004	.0023	.0007	.0008	.0025	.0028	.0022	.0046	-.0027	-.0015	.0021	0	-.0001	.0055	.0024	.0014	.0001	.0026	.0018	.0032	.0013	.0034	.0040	.0058
.0029	.0036	.0027	.0010	-.0002	.0020	-.0006	.0019	.0024	.0013	.0033	-.0001	-.0015	.0013	.0024	.0024	.0044	.0062	.0062	.0033	.0012	.002	.0009	.0002	.0011	.0028	.0034
.0012	.0024	.0027	.0016	-.0016	.0026	-.0002	-.0009	.0005	.0013	.0018	.0015	-.0006	-.0012	-.0018	-.0010	-.0007	-.0006	.0002	.0007	.0011	-.001	.0004	.0005	.0008	.0034	.0047
-.0002	.0016	.0004	.0020	-.0007	-.0009	.0013	-.0019	0	.0019	.0024	.0014	.0011	-.0002	-.0009	-.0011	-.0003	.0010	-.0003	.0010	.0005	-.0026	0	.0008	.0025	.0024	.0034
-.0010	.0034	-.0008	.0015	0	-.0031	.0019	.0002	.0003	.0030	.0030	.0027	.0013	.0003	-.0001	.0001	.0011	-.0002	-.0007	-.0002	-.0001	-.0015	-.0002	.0021	.0030	.0028	.0027
-.0008	.0040	-.0007	.0015	.0001	-.0030	.0021	.0002	-.0006	.0041	.0028	.0028	.0022	.0017	.0024	.0011	.0004	-.0007	-.0007	-.0006	.0015	.0019	.0008	.0017	.0039	.0036	.0015
.0010	.0020	.0002	.0017	-.0008	-.0007	.0016	-.0017	-.0002	.0048	.0040	.0029	.0016	.0017	.0027	.0022	.0007	.0027	.0038	.0038	.0048	.0034	.0038	.0057	.0060	.0054	.0019
.0043	.0009	.0020	.0014	-.0016	.0028	.0005	-.0001	.0019	.0038	.0061	.0026	.0017	.0013	.0015	.0036	.0023	.0027	.0013	.0019	.0035	.0028	.0026	.0037	.0062	.0061	.0058
.0034	.0006	.0024	.0016	-.0007	.0020	.0004	.0034	.0061	.0056	.0056	.0027	.0021	.0016	.0009	.0032	.0043	.0050	.0051	.0037	.0025	.0008	.0022	.0050	.0066	.0068	.0058
.0042	.0022	.0013	.0027	.0017	.0020	.0025	.0052	.0066	.0068	.0045	.0025	.0027	.0026	.0011	.002	.0037	.0006	-.0004	0	.0006	.0021	.0042	.0042	.0053	.0033	.0033
.0007	.0021	.0016	.0008	.0009	.0013	.0030	.0044	.0050	.0047	.0042	.0027	.0027	.0016	.0019	.0004	.0015	-.0008	.0017	.0033	.0037	.004	.0040	.0048	.0036	.0032	.0035
.0014	.0006	.0009	-.0010	-.0021	-.0026	-.0001	.0035	.0039	.0039	.0031	.0025	.0036	.0029	.0029	.0006	.0019	.0016	.0008	.0028	.0036	.0019	.0047	.0025	.0023	.0022	.0044
.0006	.0020	-.0006	-.0016	.0027	-.0003	-.0008	.0030	.0031	.0038	.0035	.0028	.0032	.0036	.0012	.0031	.0035	.0030	.0026	.0020	.0046	.0018	.0038	.0046	.0013	.0024	.0030
.0022	.0008	.0002	-.0032	-.0005	-.0009	-.0003	.0029	.0028	.0024	.0023	.0032	.0032	.0017	.0027	.0026	.0026	.0028	.0019	.0015	.0015	.001	.0033	.0049	.0014	.0033	.0018
.0021	.0026	.0018	-.0010	-.0007	-.0011	.0017	.0038	.0028	.0021	.0021	.0029	.0025	.0020	.0008	.0013	.0021	.0016	.0008	.0018	.0017	.0024	.0039	.0042	.0008	.0018	.0029
.0033	.0017	.0014	.0030	.0034	.0031	.0013	.0027	.0029	.0020	.0023	.0023	.0020	.0012	.0013	.0003	.0033	.0006	.0019	.0019	.0010	.0013	.0027	.0028	.0024	.0018	.0036
.0047	.0014	.0028	.0054	.0052	.0047	.0051	.0033	.0028	.0023	.0032	.0028	.0027	.0020	.0006	.0015	.0019	.0017	.0011	.0029	.0025	.0017	.0015	.0025	.0050	.0044	.0044
.0034	.0014	.0013	.0038	.0053	.0037	.0025	.0031	.0017	.0030	.0021	.0021	.0022	.0016	.0006	.0016	.0026	.0017	.0013	.0033	.0038	.0028	.0035	.0025	.0045	.0074	.0033
-.0002	-.0003	0	.0033	.0025	.0001	.0015	.0004	.0018	.0016	.0005	.0001	.0002	.0001	.0019	.0006	.0058	.0010	.0024	.0022	.0033	.0008	.0013	.0046	.0035	.0071	.0034

Tabla 3.7. Valores de β_2 para W a partir de los índices de Moran's $\{I(x_{i,j}, h), h = 1..5\}$ y de la función lineal $\hat{I} = \beta_1 + \beta_2 h$

(.011,226)	(.007,241)	(.004,236)	(.002,158)	(.007,109)	(.012,107)	(.015,111)	(.012,107)	(.007,119)	(.004,140)	(.004,7)	(.006,348)	(.002,289)	(.004,158)	(.005,108)	(.007,68)
(.006,36)	(.011,51)	(.018,49)	(.022,55)	(.026,62)	(.026,68)	(.026,72)	(.026,78)	(.025,78)	(.024,64)	(.024,58)	(.025,53)	(.025,57)	(.024,62)	(.024,67)	(.024,67)
(.032,91)	(.034,88)	(.037,85)	(.041,81)	(.041,83)	(.042,83)	(.042,85)	(.043,86)	(.044,83)	(.044,81)	(.043,79)	(.042,80)	(.040,77)	(.037,77)	(.033,80)	(.029,85)
(.040,244)	(.035,247)	(.032,246)	(.030,243)	(.025,244)	(.022,240)	(.020,238)	(.017,246)	(.015,248)	(.014,244)	(.015,239)	(.016,247)	(.018,246)	(.021,252)	(.023,251)	(.026,245)
(.090,244)	(.084,245)	(.081,245)	(.077,245)	(.073,245)	(.068,245)	(.065,246)	(.063,246)	(.062,247)	(.061,246)	(.061,246)	(.061,245)	(.062,247)	(.064,246)	(.066,245)	(.068,244)
(.007,174)	(.005,145)	(.005,145)	(.005,145)	(.005,155)	(.005,175)	(.004,205)	(.003,240)	(.003,227)	(.003,199)	(.001,30)	(.003,43)	(.002,108)	(.005,161)	(.003,162)	(.002,105)
(.005,292)	(.004,304)	(.005,342)	(.005,1)	(.006,23)	(.006,34)	(.006,40)	(.005,42)	(.005,29)	(.007,8)	(.008,3)	(.010,0)	(.009,1)	(.007,3)	(.006,11)	(.007,16)
(.010,112)	(.011,106)	(.012,101)	(.013,91)	(.014,92)	(.015,94)	(.016,97)	(.016,98)	(.016,92)	(.016,85)	(.015,80)	(.014,80)	(.014,73)	(.014,73)	(.014,76)	(.012,84)
(.023,217)	(.020,219)	(.019,218)	(.019,217)	(.017,217)	(.017,214)	(.017,212)	(.015,212)	(.013,211)	(.012,208)	(.012,207)	(.011,213)	(.012,215)	(.012,223)	(.013,225)	(.016,221)
(.075,239)	(.072,239)	(.071,240)	(.069,239)	(.067,239)	(.064,240)	(.062,241)	(.061,241)	(.060,241)	(.060,241)	(.059,241)	(.059,240)	(.059,241)	(.060,241)	(.061,240)	(.061,240)
(.006,202)	(.003,220)	(.003,220)	(.003,220)	(.003,220)	(.003,220)	(.003,240)	(.003,249)	(.003,231)	(.003,220)	(.002,253)	(.001,323)	(.001,270)	(.003,225)	(.004,246)	(.004,246)
(.005,278)	(.005,277)	(.006,291)	(.006,286)	(.007,280)	(.007,282)	(.008,286)	(.008,290)	(.008,294)	(.008,303)	(.008,311)	(.009,314)	(.008,308)	(.008,304)	(.008,296)	(.009,300)
(.007,204)	(.006,208)	(.006,209)	(.004,213)	(.004,207)	(.004,197)	(.005,194)	(.005,196)	(.004,202)	(.003,224)	(.003,246)	(.003,251)	(.003,271)	(.003,273)	(.002,279)	(.002,254)
(.024,221)	(.023,222)	(.022,221)	(.021,221)	(.020,221)	(.020,221)	(.020,221)	(.019,222)	(.019,223)	(.018,224)	(.018,225)	(.017,228)	(.017,227)	(.016,230)	(.017,229)	(.018,223)
(.053,226)	(.052,227)	(.051,227)	(.050,227)	(.049,228)	(.047,229)	(.046,229)	(.045,229)	(.044,229)	(.044,229)	(.043,229)	(.043,229)	(.043,230)	(.043,230)	(.044,229)	(.045,228)
(.006,232)	(.006,252)	(.007,254)	(.007,252)	(.007,252)	(.007,252)	(.006,257)	(.006,256)	(.006,252)	(.007,257)	(.007,261)	(.007,267)	(.006,270)	(.007,265)	(.008,266)	(.009,261)
(.009,273)	(.010,275)	(.011,280)	(.011,278)	(.012,275)	(.012,276)	(.011,277)	(.011,280)	(.012,282)	(.011,284)	(.011,288)	(.011,290)	(.011,288)	(.011,289)	(.011,285)	(.012,286)
(.010,215)	(.010,216)	(.010,217)	(.010,222)	(.011,225)	(.011,227)	(.011,231)	(.011,237)	(.011,243)	(.010,246)	(.010,247)	(.010,246)	(.010,245)	(.010,243)	(.009,243)	(.009,241)
(.027,224)	(.026,225)	(.026,225)	(.025,225)	(.024,226)	(.024,227)	(.024,228)	(.024,228)	(.023,229)	(.023,230)	(.023,232)	(.023,233)	(.022,234)	(.022,235)	(.022,235)	(.022,230)
(.048,223)	(.047,223)	(.046,224)	(.045,224)	(.045,225)	(.044,226)	(.043,226)	(.043,225)	(.042,225)	(.042,225)	(.041,224)	(.040,224)	(.04,224)	(.039,223)	(.040,222)	(.040,220)
(.007,248)	(.007,258)	(.008,259)	(.009,256)	(.009,254)	(.009,254)	(.009,256)	(.008,255)	(.008,257)	(.008,262)	(.009,260)	(.008,260)	(.008,264)	(.007,265)	(.007,265)	(.008,262)
(.011,269)	(.012,272)	(.012,275)	(.013,273)	(.013,271)	(.013,271)	(.013,272)	(.013,275)	(.014,277)	(.014,276)	(.014,277)	(.014,277)	(.014,278)	(.013,279)	(.013,278)	(.013,278)
(.013,222)	(.013,223)	(.013,224)	(.012,227)	(.012,229)	(.013,229)	(.013,233)	(.013,237)	(.012,239)	(.012,239)	(.012,240)	(.012,240)	(.012,238)	(.012,238)	(.011,239)	(.011,240)
(.027,220)	(.027,221)	(.027,221)	(.027,222)	(.027,224)	(.027,226)	(.027,227)	(.027,228)	(.026,229)	(.026,230)	(.025,230)	(.025,228)	(.025,227)	(.024,226)	(.024,224)	(.023,220)
(.043,217)	(.042,218)	(.042,219)	(.041,219)	(.041,220)	(.040,221)	(.040,221)	(.040,221)	(.040,221)	(.040,222)	(.039,222)	(.039,222)	(.038,222)	(.038,222)	(.038,221)	(.039,220)
(.007,249)	(.007,255)	(.007,257)	(.007,252)	(.008,249)	(.008,249)	(.008,251)	(.008,251)	(.008,258)	(.008,263)	(.009,257)	(.009,252)	(.008,255)	(.007,260)	(.007,256)	(.006,255)
(.011,266)	(.011,268)	(.011,270)	(.011,269)	(.011,266)	(.012,266)	(.012,267)	(.012,272)	(.013,274)	(.013,273)	(.013,272)	(.013,270)	(.012,271)	(.011,273)	(.011,275)	(.011,275)
(.013,222)	(.013,225)	(.014,228)	(.014,231)	(.014,233)	(.014,233)	(.014,235)	(.014,237)	(.014,237)	(.014,237)	(.014,238)	(.013,238)	(.014,234)	(.013,232)	(.013,231)	(.012,229)
(.025,211)	(.025,212)	(.025,213)	(.024,214)	(.024,216)	(.024,218)	(.024,219)	(.024,220)	(.024,220)	(.024,220)	(.023,221)	(.023,219)	(.023,219)	(.023,218)	(.023,218)	(.023,217)
(.042,212)	(.042,212)	(.041,213)	(.041,213)	(.041,214)	(.040,216)	(.040,217)	(.040,217)	(.040,218)	(.040,219)	(.040,219)	(.039,220)	(.039,220)	(.039,220)	(.040,220)	(.040,219)

Tabla 3.9. Vector media (longitud r y ángulo media $\bar{\Phi}$) de las celdas que pertenecen al rectángulo enmarcado de la Tabla 3.5.

2.90	3.97	33.62	51.94	24.14	46.54	44.67	28.39	19.77	14.57	1.55	19.17	23.34	25.29	11.56	7.99	.37		1.85	119.18	45.89	119.91	115.6	11.91	6.75	7.19	4.91
4.53	4.42	2.40	1.13	1.50	3.06	8.55	11.90	13.90	13.72	3.11	.20	.02	.11	1.23	3.07	2.95	2.23	4.41	12.52	43.65	22.84	16.31	15.92	18.07	7.90	4.77
5.52	5.87	4.83	3.20	3.26	4.39	5.03	4.58	3.97	3.60	1.83	.73	.55	.64	.99	1.42	1.41	1.90	5.21	15.09	12.79	8.01	3.49		4.81	4.16	3.95
4.25	3.86	3.55	3.66	3.84	3.57	3.55	3.81	4.14	4.30	4.77	6.47	8.01	7.24	7.66	9.96	6.73	4.52	3.57	3.31	3.31		5.18	5.67		6.10	5.82
7.37	9.71	14.13	14.39	17.40	18.60	19.92	25.81	31.48	32.83	44.91	59.19	62.02	56.74	56.50	57.58	56.93	52.40	54.75	50.24	53.21		39.21	43.52		37.34	34.03
30.62	33.65	37.26	37.98	46.41	48.70	47.72	59.35	58.70	55.82	60.11	62.67	58.96	52.02	52.57	49.68	46.01	39.65	42.23	41.01	38.23	37.08	33.95			34.93	32.48
39.67	42.03	50.12	56.43	68.37	78.63	72.02	79.83	81.67	80.36	73.05	78.87	54.59	49.35	45.8	36.69	30.02	23.50	21.89	2.20	19.31	16.93	17.38	15.46	14.67	13.40	11.92
12.36	12.06	13.66	13.94	13.58	13.98	13.30	15.51	17.80	21.54	22.50		19.09	16.09	15	12.45	11.22	13.21	11.14	8.90	7.92	8.11	7.10	6.59	6.07	5.78	5.62
5.54	5.45	5.20	5.03	5.16	5.51	5.89	6.14	6.76	7.93	7.48		6.58	6.28	5.74	5.38	5.03	5.05	6.27	8.58	9.92	8.85	8.27	1.03	13.87	12.90	11.74
12.81	13.55	13.69	1.22	6.83	6.08	6.27	6.49	5.45	5	5.02	5.57	6.58	6.87	6.98	6.42	5.90	5.52	5.64	5.93	6.23	6.40	6.36	7.05	8.25	1.05	11.74
9.96	8.81	8.40	6.55	5.51	5.41	5.69	6.14	6.99	8.12	8.66	8.85	9.38	11	13.05	19.16	26.98	37.93	37.27	48.89	40.02	39.16	31.17	29.40	37.68	35.64	34.23
36.98	40.14	42.70	54.73	52.61	64.26	71.85	75.81	79.37	81.85	80.40	89.55	92.56	79.93	99.75	102.06	96.13	92.92	80.48	78.53	69		71.36	63.01	65.63	75.71	89.11
75.82	91.20	102.91	108.72	126.71		126.28	125.85	111.56	106.90	127.51	128.68	128.61	129.65	139.84	144.18	166.57	188.69	206.73	242.90	258.74	262.97	262.85	248.55	222.22	222.5	216.02
231.12	225.38	187.57	172.43	167.73		174.29	166.58	160.39	157.95	161.88	177.61	198.40	218.77	221.63	240.66	203.82	153.65	97.56	75.91	60.85	64.76	64.79	83.91	86.12	78.82	73.51
70.43	60.72	50.15	51.01	60.83	70.16	64.57	51.30	46.71	48.49	51.24	48.47	44.17	43.65	40.77	45.12	42.07	48.38	6.54	76.26	80.98	90.05	81.8	84.87	78.34	81.70	71.60
74.25	67.96	62.82	71.56	90.74	104.85	90.68	75.54	44.63	36.30	30.61	17.85	15.81	1.70	9.68	7.96	8.87	8.65	9.05	1.27	9.64	1.49	9.43	8.56	8.51	7.92	7.58
7.72	8.03	8.14	8	7.70	7.65	7.94	8.34	8.44	8.48	8.91	9.42	9.98	11.01	11.61	11.76	11.44	11.09	1.95	11.87	12.85	14.20	14.92	15.80	16.44	18.31	2.23
2.83	21.14	19.81	2	21.58	23.24	26.51	27.33	26.38	25.50	25.52	26.12	28.86	29.35	31.16	32.33	34.83	37.36	46.42	55.03	66.11	78.48	72.64	78.34	73.25	77.88	83.17
96.16	118.17	104.08	92.34	78.60	68.60	69.06	77.89	87.14	80.22	82.62	73.25	67.82	69.28	71.01	78.14	89.70	94.02	100.35	97.81	79.18	82.44	79.28	81.77	75.18	73.35	79.45
70.34	74.94	76.92	70.66	59.01		41.65	40.24	38.79	40.85	43.92	44.52	41.37	38.90	43.41	42.90	46.61	43.80	45.90	43.46	37.88	35.89	33.33	33.69	33.11	30.58	30.31
31.25	33.37	34.68	36.22	35.61	32.55	30.65	28.85	28.01	29.38	28.63	28.70	28.64	26.97	27.29	29.08	26.88	27.10	26.92	27.60	26.66	25.03	24.24	24.24	26.20	24.81	25.46
23.94	24.69	24.78		33.52	36.78	35.21	37.62	35.96	34.90	34.79	34.57	34.46	34.37	34.38	34.39	37.98	35.48	32.52	34.99	36.87	37.70	42.07	43.98	37.71	38.19	38.89
36.03	32.03	31.98	37.61	42.72	47.71	48.81	48.90	49.19	55.31	54.94	48.89	54.26	54.19	54.50	55.56	61.56	61.65	63.66	69.40	69.53	61.19	70.22	69.05	50.14	47.16	45.08
42.01	39.96	40.03	39.82	44.42	44.58	44.21	47.75	44.55	44.98	43.35	42.85	43.25	43.35	43.45	44.35	44.70	48.13	47.96	47.85	47.65	45.29	44.96	44.46	45.35	42.04	42.08
41.52	42.79	41.09	42.47	42.20	38.63	38.28	36.44	34.45	34.69	40.35	38.11	38.10	38.35	36.70	34.34	31.99	31.61	31.52	31.50	31.72	32.07	35.91	38.25	41.19	39.10	39.33
35.80	34.33	34.36	35.74	33.74	33.57	33.46	33.58	32.43	32.47	32.19	31.49	31.62	31.83	32.79	31.85	29.85	28.34	28.35	29.33	29.50	29.88	34.97	35.16	37.77	39.13	40.42
36.21	34.21	34.22	33.96	33.81	33.78	33.93	34.28	34.44	34.15	33.77	31.48	31.67	33.26	33.23	35.61	37.05	37.79	41.63	41.62	35.36	32.11	33.41	33.76	35.25	35.46	35.62

Tabla 3.10. El test estadístico T^2 para W . En negrita las celdas que no verifican la condición de direccionamiento, y consecuentemente, no es posible asumir anisotropía.

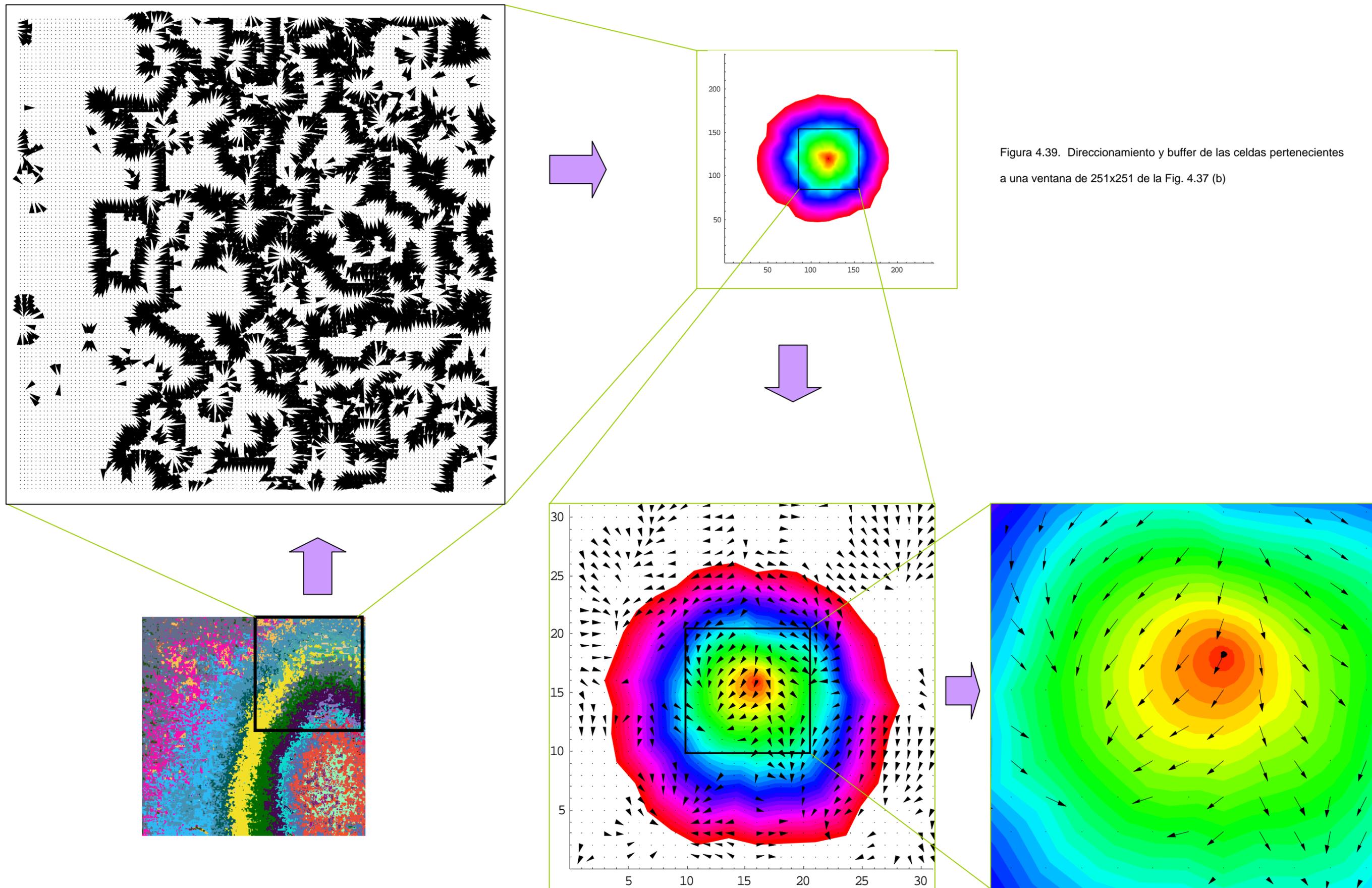


Figura 4.39. Direccionamiento y buffer de las celdas pertenecientes a una ventana de 251x251 de la Fig. 4.37 (b)