

UNIVERSIDAD DE ALCALÁ

**ESCUELA TÉCNICA SUPERIOR DE
ARQUITECTURA Y GEODESIA**

INGENIERÍA EN GEODESIA Y CARTOGRAFÍA



PROYECTO FIN DE CARRERA:

**“Metodología para el análisis de incertidumbre de un
modelo basado en técnicas de Evaluación MultiCriterio
y Sistemas de Información Geográfica”**

Andrés Calabia Aibar

2011



Universidad
de Alcalá

**ESCUELA TÉCNICA SUPERIOR DE
ARQUITECTURA Y GEODESIA
INGENIERÍA EN GEODESIA Y CARTOGRAFÍA**

PROYECTO FIN DE CARRERA:

**“Metodología para el análisis de incertidumbre de un
modelo basado en técnicas de Evaluación MultiCriterio
y Sistemas de Información Geográfica”**

AUTOR: Andrés Calabia Aibar

DIRECCIÓN: Montserrat Gómez Delgado

TRIBUNAL:

PRESIDENTE: Joaquín Bosque Sendra

VOCAL 1º: José Antonio Malpica Velasco

VOCAL 2º: Montserrat Gómez Delgado

CALIFICACIÓN:

FECHA:

AGRADECIMIENTOS

Que en el presente trabajo quede reflejado mi más sincero agradecimiento a la universidad de Alcalá de Henares y con especial atención a la directora de este proyecto, la profesora Montserrat Gómez Delgado, que ha contribuido en gran medida a la mejora de mi expresión de conocimientos en un documento escrito.

1. INTRODUCCIÓN

La validación de modelos geográficos es muy importante, pues proporciona fiabilidad y credibilidad de los resultados que estos producen.

Este proyecto pretende ofrecer de una manera práctica, un procedimiento a seguir para la validación parcial de modelos normativos obtenidos mediante procesos de Evaluación MultiCriterio (EMC) y Sistemas de Información Geográfica (SIG) mediante un análisis de incertidumbre. El modelo geográfico que se pretende validar proporciona localizaciones óptimas de suelo urbano para el año 2020 en la Comunidad de Madrid.

La validación mediante el análisis de incertidumbre se fundamenta en la transmisión de los posibles errores existentes en las variables de entrada del modelo en los resultados del mismo. Teniendo en cuenta esto, una distorsión controlada de las variables iniciales puede proporcionar un índice que valore la incertidumbre y evalúe de alguna manera la robustez de los resultados del modelo.

1.1. PLANTEAMIENTO DEL PROBLEMA

Habitualmente, la exactitud de un mapa se proporciona de una manera global, cuando en realidad puede haber variaciones espaciales y temáticas importantes. El humano no es capaz de detectar fácilmente la incertidumbre introducida en un modelo por los datos de partida y por ello se están realizando muchos proyectos y se han creado muchos métodos y herramientas, pero todavía no hay ninguna implementada en un SIG que comunique el efecto del error asociado a un modelo. Entre otros, los métodos desarrollados están basados en:

- **Reglas o procedimientos**, como los seis pasos para la gestión de riesgos en el uso de datos espaciales por Agumya y Hunter (1999), la incertidumbre en bases de datos de Bédard (1987) o el manual de instrucciones de Gervais (2003) para una estrategia de gestión de riesgos derivados de la incertidumbre.
- **Técnicas de visualización**, donde Leitner y Buttenfield (2000) representan la certeza de los datos con visualizaciones de saturación, textura y valor, Howard y MacEachren (1996) o Fisher (1994), con simbologías y estudios a cerca del esfuerzo del lector de un mapa para deducir la incertidumbre asociada y, entre otros, Buttenfield y Beard (1991), Beard y Mackaness (1993), Buttenfield (1993), McGranaghan (1993), Beard (1997), Beard y Buttenfield (1999), Drecki (2002), Devillers y Jeansoulin (2006), etcétera.

En la actualidad se están desarrollando nuevos métodos para evaluar la calidad de los datos mediante herramientas de visualización y análisis capaces de proporcionar al usuario un resumen lógico sobre la incertidumbre en los resultados de un modelo (Devillers et al. 2007).

Según Rykiel (1996), los procesos de validación han de ser necesariamente distintos, dependiendo del tipo de modelo tratado, ya sea para ciencias experimentales, ingeniería o ciencias sociales.

El propósito de estos estudios es informar de la bondad de los resultados de un modelo, pero en cualquier caso, no existe un acuerdo general en la comunidad científica sobre en qué debe consistir ese proceso de validación. Con el mismo objetivo, Qureshi et al (1999) proponen un proceso formado por tres componentes:

- Verificación de si el modelo está correctamente construido desde el punto de vista formal.
- Validación de si la estructura del modelo representa adecuadamente el sistema modelado, construyendo un modelo correcto desde un punto de vista conceptual y operacional.

- Aplicación de un análisis de sensibilidad que permita comprobar la robustez y estabilidad del modelo, examinando la variación que se produce en los resultados cuando se realizan variaciones sistemáticas en un rango de interés sobre uno o varios parámetros de entrada.

Con otras propuestas, Pagelow y Camacho Olmedo (2008), dividen el proceso en tres tareas a realizar:

- Verificación o validación interna para comprobar que el modelo está correctamente construido.
- Calibración o comprobación de si el comportamiento global del modelo se ajusta a los objetivos.
- Validación para comprobar y mejorar la robustez del modelo, punto que correspondería al análisis de incertidumbre.

Este proyecto pretende analizar la metodología empleada para la cuantificación y localización de la incertidumbre de un modelo obtenido a partir técnicas de EMC y SIG. Es sabido que las variables empleadas en los modelos están sometidas a la teoría de propagación de errores y cómo los resultados de los modelos varían cuando se modifican sus variables de entrada, estas variaciones pueden proporcionar la incertidumbre transmitida al resultado.

1.2. ESTADO DE LA CUESTIÓN: REVISIÓN TEÓRICA

La aplicación de simulaciones para estudiar la incertidumbre asociada a los datos de entrada en un SIG se ha incrementado sustancialmente en la última década. En los últimos años encontramos una cantidad importante de aplicaciones y propuestas para validar modelos. Por ejemplo, los análisis realizados por Pontius y Schneider (2001) donde se emplea el parámetro *Relative Operating Characteristic (ROC)* o los análisis realizados por Munroe et al (2002) para cuantificar la exactitud en los cambios de cubiertas vegetales en el occidente de Honduras. Acercándonos un poco más, existen los análisis de incertidumbre relacionados con los posibles errores en los datos de partida, tales como los realizados por Emmi y Horton (1996) en modelos para riesgo sísmico, las técnicas empleadas por Davis y Keller (1997) de lógica difusa y simulación de Monte Carlo en el análisis de incertidumbre de un modelo de estabilidad de taludes, los cálculos de Kiiveri (1997) para la distorsión geométrica y los análisis de incertidumbre del modelo de localización óptima para un emplazamiento de una nueva construcción o los estudios de Arbia y Haining (1998) con la propagación de errores en operaciones de superposición mediante el análisis ANOVA y el índice Kappa.

Ahora bien, en el año 2007 Burnicki y otros proponen una metodología basada en la generación de un conjunto de mapas que son perturbados con patrones específicos de error. Acto seguido comparan los mapas generados y determinan los efectos de los patrones de error en la exactitud del mapa resultante. Por otro lado, Lee (1992) también añadió una perturbación aleatoria para analizar los errores de los modelos. Así pues, se pueden generar unas variables iniciales afectadas por una distorsión controlada para obtener su modelo derivado y a continuación comparar el modelo obtenido contra el original para realizar el análisis del error transmitido.

En 2004, Gómez Delgado y Bosque Sendra proponen una metodología similar que pretende controlar el riesgo en cualquier toma de decisiones basada en los resultados de un modelo basado en técnicas de EMC y ejecutado en un entorno SIG. Esta metodología se basa en el uso combinado de un análisis de sensibilidad y un análisis de incertidumbre obteniendo unos índices que validan. Así pues, por lo que se refiere a estos análisis, este proyecto desarrolla el análisis de incertidumbre y analiza el efecto que los errores iniciales provocan en los resultados.

El método propuesto por Gómez Delgado y Bosque Sendra (2004) se basa en la introducción de perturbaciones aleatorias en cada uno de los factores por toda la superficie analizada. En cambio, otras propuestas tales como las de Emmi y Horton (1996), Davis y Keller (1997) o Crosetto y Tarantola (2001), limitan la perturbación a una franja entorno a estos factores.

Insistiendo en lo que aquí se pretende aportar, el proyecto desarrollado propone introducir un error más dirigido y realista. En otras palabras, se procura una innovación en la metodología tras haber analizado los procedimientos de distorsión disponibles. Ante las diferentes alternativas propuestas por la comunidad científica a la hora de introducir error, la elección del método de distorsión es esencial y básica para la obtención de resultados de robustez estadística.

Por otro lado, en muchos de los casos, el método resulta muy complejo y difícil de implementar, por lo que resulta necesario desarrollar propuestas que permitan ser puestas en práctica con las herramientas disponibles en cualquier SIG convencional y crear módulos de software o actualizar los existentes, para automatizar estos procesos.

1.3. OBJETIVOS

Partiendo de la situación descrita en el anterior apartado, en el presente trabajo se pretende desarrollar una metodología alternativa a la planteada por Gómez Delgado y Bosque Sendra (2004) de análisis de incertidumbre, que formaría parte de una validación parcial de un modelo de simulación de crecimiento urbano deseable para el año 2020 en la Comunidad de Madrid.

Para ello se plantean los siguientes objetivos:

- En primer lugar, se pretende identificar los posibles factores que puedan proporcionar incertidumbre en el modelo.
- Una vez seleccionado el método más adecuado para transmitir la incertidumbre de las variables iniciales en el resultado del modelo, se generarán cuatro reproducciones de cada variable, distorsionadas de manera controlada. Dichas nuevas versiones de las variables de partida, se emplearán aleatoriamente para crear treinta resultados deformados y así simular la incertidumbre de los datos iniciales.
- El siguiente punto trata de comparar los treinta resultados modificados con el resultado sin modificar, a partir del cálculo del índice de Incertidumbre Media Cuadrática propuesto por Gómez Delgado y Bosque Sendra (2004), con el fin de valorar la influencia de los errores introducidos.
- Finalmente, se realizará una comparación de la metodología empleada con otro método que realiza el mismo análisis.

2. METODOLOGÍA

En esta sección, se desarrolla el tema principal, más técnico y de mayor duración del proyecto. Este capítulo está estructurado en diferentes secciones, compuestas por subcapítulos, que explican cómo se distorsionan las variables de entrada de un modelo para luego analizar la incertidumbre del resultado. En primer lugar se procede a describir el modelo a validar, luego se desarrolla la metodología de distorsión espacial y temática y finalmente se calcula el índice de Incertidumbre Medio Cuadrático, que valida los resultados y la robustez del modelo.

En el proyecto SIMURBAN se estudian dos zonas, la Comunidad de Madrid y la aglomeración urbana de Granada. En este trabajo se valida el modelo de uso residencial de la comunidad de Madrid para el año 2020, partiendo de la situación existente en el año 2000, aunque en el proyecto también se llevó a cabo la simulación para los usos del suelo industrial y comercial. La Figura 1 muestra el resultado del modelo para los tres usos.

El escenario de simulación de innovación y sostenibilidad del que parte este trabajo se materializa a partir de factores sociales, económicos y ambientales que persiguen conformar un modelo de sostenibilidad (Tabla 1). Esta serie de factores fueron ponderados a partir del método de comparación por pares de Saaty. Una vez obtenidos los factores normalizados y sus pesos finales, se empleó una sumatoria lineal ponderada para obtener los mapas de aptitud de uso del suelo urbano.

GRUPO	Factores	Pesos
Ambiental	Usos del suelo (<i>Corine Land Cover 2000</i>)	0,18
	Vulnerabilidad a la contaminación de los mantos acuíferos	0,04
	Protección de suelos de mayor productividad.	0,09
Económico	Pendiente	0,03
	Accesibilidad a carreteras	0,16
	Accesibilidad a zonas urbanas	0,21
	Geotecnia	0,08
Social	Accesibilidad a hospitales	0,13
	Orientación del terreno	0,08

Tabla 1. Factores utilizados en el modelo SIMURBAN para la obtención de suelo residencial para el año 2020 en el escenario de innovación y sostenibilidad.

La asignación se corresponde, por un lado, con un grupo de factores que puntúan con unos valores de aptitud altos como son el uso del suelo, la pendiente, la accesibilidad a carreteras o a zonas urbanas. Por otro lado, un grupo con menor incidencia, como la geotecnia, los tipos de suelo, la vulnerabilidad a la contaminación de los mantos acuíferos y la orientación del terreno. Además, un mapa de restricciones excluye las zonas clasificadas como suelo no urbanizable de protección del año 2003, los espacios naturales protegidos y las superficies artificiales existentes en el año 2000.

Como rasgos generales, el modelo empleado presenta cierta compacidad en la distribución de los nuevos usos del suelo y ofrece un carácter sostenible al territorio para el futuro. Asimismo garantiza que el uso se asigna según los factores planteados, en las mejores condiciones ambientales, económicas y sociales y en un marco legal establecido por los entes planificadores del territorio.

2.2. DISTORSIÓN ESPACIAL Y TEMÁTICA DE LAS VARIABLES DE PARTIDA

La mayoría de modelos que se crean no tienen previsto realizar un análisis de incertidumbre y la falta de esta previsión puede causar una pérdida de tiempo a la hora de reconstruir los procesos y realizar los análisis de validación de éstos.

El análisis de incertidumbre analiza la influencia que produce la incertidumbre de las variables iniciales y su propagación en los resultados del modelo. La metodología empleada introduce un error aleatorio en las variables iniciales y posteriormente obtiene varios resultados distorsionados con variaciones que reflejan la incertidumbre de éstos. No obstante, es recomendable distorsionar las variables en su estado más inicial y obtener un número de reproducciones estadísticamente adecuado para realizar posteriormente las comparaciones con el resultado original.

El proceso de análisis comprende:

1. Insertar un error aleatorio controlado en las variables iniciales, tanto en la componente espacial como temática.
2. Ejecutar de nuevo el modelo empleando las variables a las que se les ha introducido el error.
3. Se repite el proceso del punto uno y dos un número de veces estadísticamente correcto, obteniendo así una serie de resultados sensiblemente diferentes.
4. Con todos los resultados, se puede medir el grado de variación alcanzado mediante un índice que compare los resultados modificados con el original.
5. La información resultante servirá para adoptar una decisión en función de la incertidumbre global obtenida en el resultado y del grado de confianza que nos proporcione el mismo.

La tabla 2 propone los tipos de distorsión a realizar en los factores del *Modelo de innovación y sostenibilidad de crecimiento urbano para la localización óptima de zonas residenciales en la Comunidad de Madrid para el año 2020*. Hay que hacer notar que los tipos de distorsión empleados están aplicados en función del tipo de variable que representan y de las dimensiones de sus entidades que se explicará a continuación.

En lo que se refiere a la dimensión de las entidades, el factor de vulnerabilidad a la contaminación de los mantos acuíferos, por ejemplo, contiene polígonos de muy elevada extensión y por este motivo no se aplica la distorsión temática. En otras palabras, es poco probable encontrar errores temáticos en la clasificación de polígonos tan extensos.

Por otro lado, las variables derivadas del Modelo Digital de Elevaciones (MDE) como la pendiente y la orientación del terreno, representan una variable con una estructura diferente a la de punto, línea o polígono y por tanto se les aplica otra metodología. La metodología aplicada en este tipo variable se fundamenta en la alteración del MDE del que provienen.

Factores	Tipo de distorsión
Usos del suelo (<i>Corine Land Cover 2000</i>)	Espacial y temático
Vulnerabilidad a la contaminación de los mantos acuíferos	Espacial
Tipos de suelo	Espacial y temático
Pendiente	Alteración del MDE
Accesibilidad a carreteras	Espacial y temático
Accesibilidad a zonas urbana	Espacial y temático
Geotecnia	Espacial y temático
Accesibilidad a hospitales	Espacial y temático
Orientación del terreno	Alteración del MDE

Tabla 2. Factores del modelo con los tipos de cambios realizados.

2.2.1. DEFORMACIÓN GEOMÉTRICA

La deformación geométrica se realiza mediante un proceso inverso al de georreferenciación de una imagen. En lugar de desplazar los elementos a una posición correcta, estos tomarán nuevas localizaciones dentro de un margen de incertidumbre espacial. En las siguientes secciones se pretende explicar, mediante desarrollos matemáticos y ejemplos, la mejor propuesta que se puede encontrar disponible en un SIG para deformar geoméricamente las variables de un modelo. Primero se pretende cuantificar y localizar la deformación de las variables, dando a conocer la magnitud y el ámbito de deformación. Después se exponen los métodos de deformación y se desarrolla y aplica el método más apropiado.

2.2.1.1. MARGEN DE INCERTIDUMBRE ESPACIAL

En la incertidumbre de un mapa digital, cada conjunto de coordenadas y dentro de éste, cada coordenada, tiene un grado de aproximación o dispersión que es función de un sin número de causas accidentales o sistemáticas, entre las cuales se puede citar las siguientes:

- La precisión de las observaciones en la captura, almacenamiento y tratamiento de los datos brutos que dieron lugar al conjunto de coordenadas.
- La exactitud de las correcciones aplicadas a dichas observaciones para adecuarlas a la realidad, así como la reducción de distancias y ángulos.
- Los programas informáticos y herramientas utilizadas y la forma en que se han aplicado para tratar los datos.
- La adecuación al modelo geométrico de representación, ya sea bidimensional, tridimensional y su sistema de referencia, elipsoide, geoide, proyección u otro.
- Las variaciones locales o impurezas resultado de usar un rango pequeño en la toma de muestras.

En relación con este tipo de errores, el margen de incertidumbre espacial propuesto hace referencia a la magnitud y ámbito de error que los elementos espaciales puedan tener.

Con respecto a la magnitud del error que se pretende introducir, supongamos un nodo que represente la localización de un hospital o de un almacén de residuos que tiene una incertidumbre en su posición. Esta magnitud de error en formato distancia, está aplicada en un radio de acción centrado en la posición real de la entidad, y para ello, se propone emplear una distribución aleatoria de vectores que distorsionen geoméricamente las entidades. Las magnitudes de los vectores propuestos tienen una media de 350 metros y variarán entre 0 y 700. Estos intervalos se han adecuado para un desplazamiento medio de 6 píxeles y máximo de 14, que corresponde a la resolución de 50 metros.

Por otro lado, el ámbito de deformación hace referencia a la disposición espacial de cada valor o magnitud de error en el espacio a deformar. Es decir, la frecuencia espacial en la que esa magnitud aleatoria va a estar aplicando su efecto. Para este caso, las deformaciones se van a repartir por todo el territorio de una forma aleatoria y uniforme, evitando aglomeraciones, vacíos y superposiciones.

En base a estos conceptos propuestos, se pretende aplicar una deformación empleando una transformación espacial bidimensional y restringir las deformaciones de las entidades dentro de unos límites establecidos.

Es necesario conocer algunas reglas y conceptos esenciales para el empleo de un modelo matemático que satisfaga la necesidad de transformar unas coordenadas en otras. Entre otros conceptos básicos, se pueden destacar algunos:

- El tipo de acuerdo entre las coordenadas transformadas y las conocidas en el primer sistema se emplea para deformar las entidades mediante unos vectores de identidad y una transformación.
- Se debe controlar la magnitud de los vectores de identidad así como los límites de acción de la transformación para limitar las distorsiones y conseguir coherencia en el resultado.
- La transformación será diferente si la cantidad de puntos comunes es diferente o incluso si se elige otro conjunto de puntos comunes.
- La elección de puntos comunes debe cubrir toda la zona holgadamente y con uniformidad.
- La transformación elegida y sus parámetros inherentes, se debe aplicar a todas las coordenadas de la imagen.

La figura 2 muestra el diseño de una matriz de celdas virtuales que define la distorsión del mapa mediante los vectores de identidad.

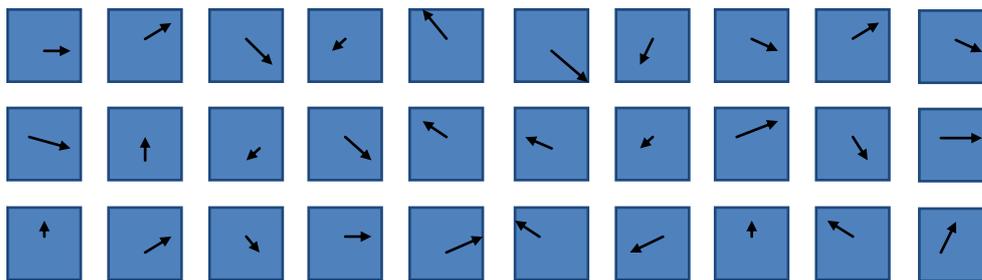


Figura 2. Ejemplo gráfico de 30 celdas virtuales con sus correspondientes vectores de identidad.

En la Figura 3 se puede observar que cada vector tiene el punto origen dentro de un ámbito de posición centrado en otro que es el ámbito de las coordenadas del punto final.

Esta distribución de celdas evita que el origen del vector se aproxime al borde de la celda exterior y salga de los límites de ésta en caso de que esté orientado hacia el exterior. Si los vectores salieran de las celdas mayores se producirían cruces entre ellos y sería una transformación poco correcta.

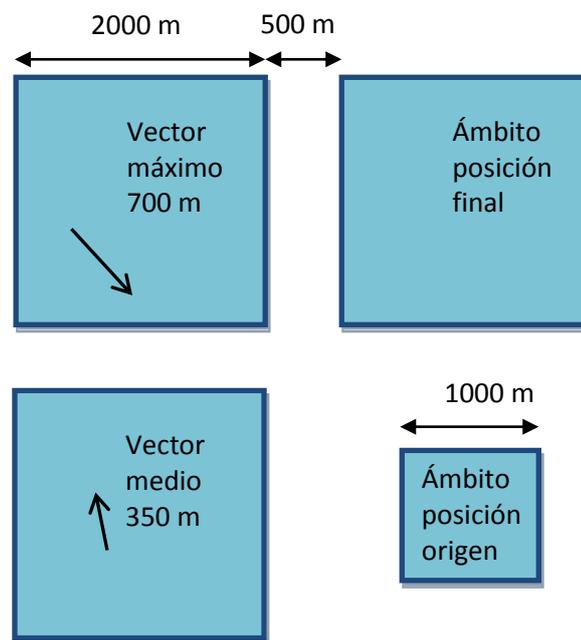


Figura 3. Disposición de los vectores identidad en la distorsión espacial.

El ámbito, la frecuencia y dimensiones de estas celdas se diseñan en función del tiempo de procesado y de la potencia del *hardware* y *software* empleado. Para ello se realizan diversas pruebas y se opta por unas dimensiones de una matriz virtual de celdas de 2 por 2 kilómetros, separadas entre ellas 500 metros. Finalmente, con estas dimensiones se han obtenido tres mil vectores para cubrir toda la imagen de estudio. En la Figura 4 se muestra el resultado en detalle.

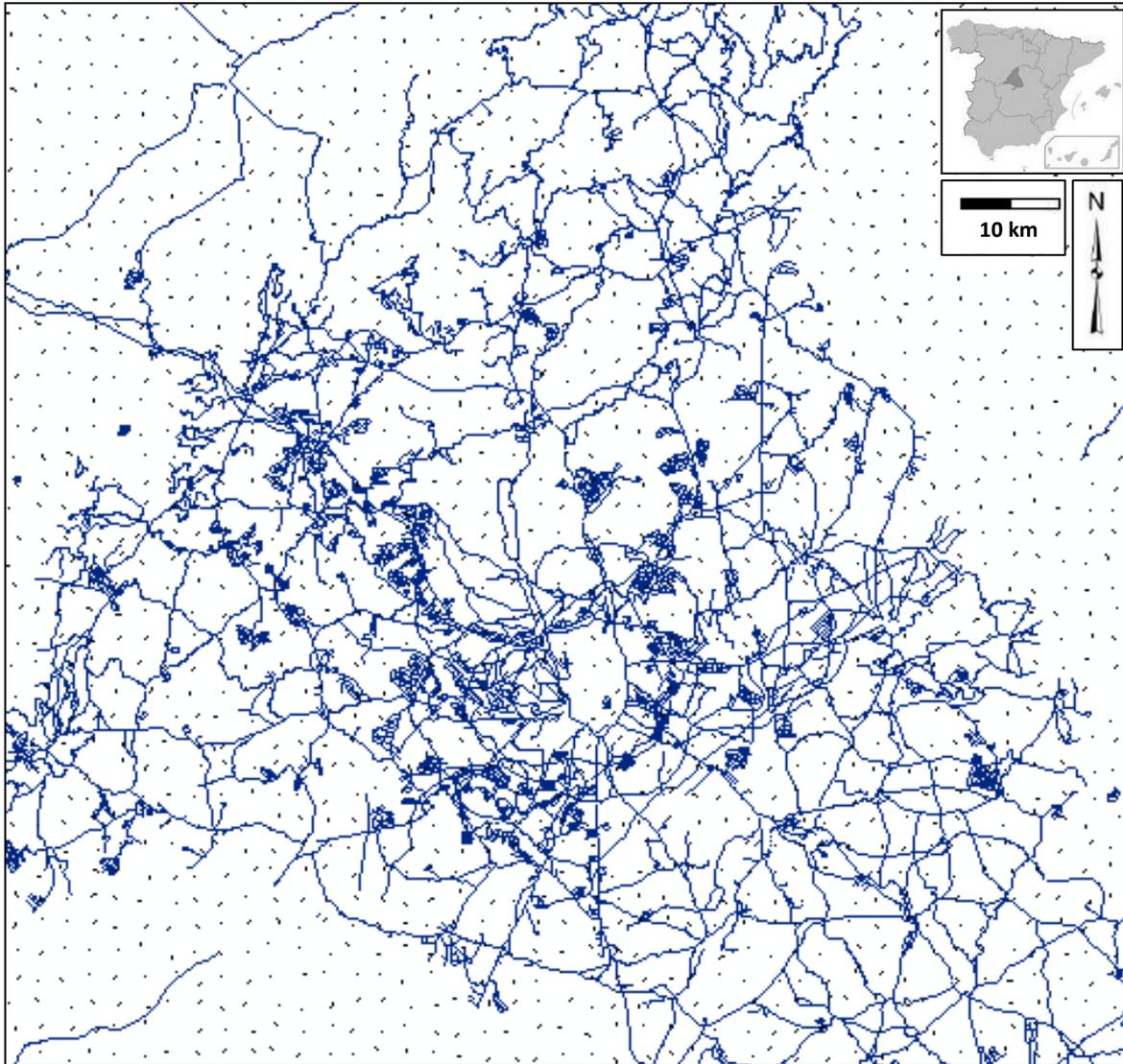


Figura 4. Tres mil vectores de identidad repartidos en el mapa de la zona de estudio.

Para crear este volumen de datos, en primera aproximación se recomienda el empleo de una hoja de cálculo tipo *Excel*, que permite realizar las operaciones simples de sumas y restas necesarias. Como segunda opción y con resultados muy potentes, se puede utilizar la programación de algoritmos en cualquier lenguaje informático.

La obtención de los vectores empieza con un conjunto de pares de coordenadas repartido por toda una superficie. El carácter aleatorio lo proporciona el empleo de números aleatorios con los que se obtienen coordenadas diferentes para cada par de puntos de cada vector.

El error a introducir en cada coordenada se puede obtener con la expresión:

$$E_{i,m,n,j} = M_i \cdot A_{i,m,n,j} \cdot (-1)^{\text{ent}(100 \cdot A_{i,m,n,j})} \quad \forall i = \{x, y\}; j = \{o, f\}$$

Donde:

M es la semimagnitud de error para el correspondiente eje de coordenadas,
m y **n** se corresponde a la fila y columna de la matriz de celdas,
i indica si es para la coordenada **x** o **y**,
j indica si es el punto origen (**o**) o punto final (**f**),
A es un número entero aleatorio entre cero y uno.

El último término de esta expresión está en forma de potencia y cambia aleatoriamente el signo de la porción aleatoria del valor **M_i**. En lo referente a la aleatoriedad del signo, se ha multiplicado por una centena el número aleatorio y se ha empleado el dígito de la posición decimal para proporcionar aleatoriamente un número entero par o impar. Como se observa en la fórmula, -1 se eleva a este número par o impar y se obtiene el cambio aleatorio de signo. Con esta fórmula obtenemos un número aleatorio que recorre el rango desde **M_i** negativo a **M_i** positivo.

Entonces, si las coordenadas para cada vector de fila **m** y columna **n** son:

$$\bar{v}_{m,n} = (x_{m,n,f} - x_{m,n,o}, y_{m,n,f} - y_{m,n,o})$$

Siendo los valores aleatorios de la coordenada origen y final:

$$x_{m,n,o} = x_{\emptyset} + n \cdot d_x + E_{x,m,n,o}$$

$$x_{m,n,f} = x_{\emptyset} + n \cdot d_x + E_{x,m,n,o} + E_{x,m,n,f}$$

$$y_{m,n,o} = y_{\emptyset} + m \cdot d_y + E_{y,m,n,o}$$

$$y_{m,n,f} = y_{\emptyset} + m \cdot d_y + E_{y,m,n,o} + E_{y,m,n,f}$$

Donde:

x_∅ e **y_∅** son las coordenadas de la esquina del mapa,
m y **n** se corresponde a la fila y columna de la matriz de celdas,
o o **f** indica si es el origen o final de cada vector y
d_x y **d_y** son los incrementos de distancia entre cada vector, que se pueden obtener dividiendo la longitud de la imagen entre el número de vectores que se quiera introducir por cada eje **x** o **y**.

Sustituyendo el error E definido anteriormente, tendríamos:

$$x_{m,n,o} = x_{\emptyset} + n \cdot d_x + M_x \cdot A_{x,m,n,o} (-1)^{ent(100A_{x,m,n,o})}$$

$$x_{m,n,f} = x_{\emptyset} + n \cdot d_x + M_x \cdot A_{x,m,n,o} (-1)^{ent(100 \cdot A_{x,m,n,o})} + L_x \cdot A_{x,m,n,f} (-1)^{ent(100A_{x,m,n,f})}$$

$$y_{m,n,o} = y_{\emptyset} + m \cdot d_y + M_y \cdot A_{y,m,n,o} (-1)^{ent(100A_{y,m,n,o})}$$

$$y_{m,n,f} = y_{\emptyset} + m \cdot d_y + M_y \cdot A_{y,m,n,o} (-1)^{ent(100A_{y,m,n,o})} + L_y \cdot A_{y,m,n,f} (-1)^{ent(100A_{y,m,n,f})}$$

Hay que notar que para cumplir las necesidades de magnitud y ámbito de error comentados, M_i y L_i toman el valor de 500 metros. Con este valor para cada eje, la composición cuadrática es la distancia máxima de longitud del vector, es decir, $\sqrt{2} \cdot 500$, que son unos 700 metros. Por otro lado, el valor d_i debe ser de 2,5 kilómetros para que los vectores estén separados una distancia adecuada. En este caso, d_i debe ser mayor que la suma de M_i más L_i para que no se crucen entre ellos. Con respecto al empleo de celdas cuadradas, basta con que los valores $M_x + L_x = M_y + L_y$.

Finalmente, el algoritmo para obtener las coordenadas tiene la siguiente estructura:

```

INTRODUCE  $x_{\emptyset}, y_{\emptyset}, d_x, M_x, d_y, M_y, L_x$ 
DESDE 0 HASTA m
  DESDE 0 HASTA n
    A1=RANDOM
    A2=RANDOM
    A3=RANDOM
    A4=RANDOM
    IMPRIMIR  $x_{\emptyset} + n \cdot d_x + M_x \cdot A1 \cdot (-1)^{ent(100 \cdot A1)}$ 
    IMPRIMIR " "
    IMPRIMIR  $y_{\emptyset} + m \cdot d_y + M_y \cdot A2 \cdot (-1)^{ent(100 \cdot A2)}$ 
    IMPRIMIR " "
    IMPRIMIR  $x_{\emptyset} + n \cdot d_x + M_x \cdot A1 \cdot (-1)^{ent(100 \cdot A1)} + L_x \cdot A3 \cdot (-1)^{ent(100 \cdot A3)}$ 
    IMPRIMIR " "
    IMPRIMIR  $y_{\emptyset} + m \cdot d_y + M_y \cdot A2 \cdot (-1)^{ent(100 \cdot A2)} + L_y \cdot A4 \cdot (-1)^{ent(100 \cdot A4)}$ 
    IMPRIMIR (SALTO DE LINEA)
  FINDESDE
FINDESDE

```

En la tabla 3 se observa el algoritmo aplicado en una hoja de cálculo. En la columna de cabecera " $x_{\emptyset} + n \cdot d_x$ " se encuentran los valores de los orígenes sin error. A la coordenada x de la esquina del mapa se le ha sumado de celda en celda descendiente el valor d_x . Para aplicar el carácter aleatorio se suma a cada valor x e y , las columnas E_{xmno} y E_{ymno} . Cada vez se ha de sumar un número aleatorio diferente para evitar repetición o patrón posible. De esta manera se obtienen coordenadas aleatorias, una para cada par de coordenadas del origen de cada vector identidad.

Por otro lado, otros dos valores aleatorios en las columnas E_{xmnf} y E_{ymnf} son los que proporcionan las coordenadas finales de cada vector. Para ello, se suman estos valores a las coordenadas aleatorias de los orígenes.

$E_{x,m,n,o}$	$E_{y,m,n,o}$	$E_{x,m,n,f}$	$E_{y,m,n,f}$	$x_0+n \cdot d_x$	$y_0+n \cdot d_y$	$X_{m,n,o}$	$Y_{m,n,o}$	$X_{m,n,f}$	$Y_{m,n,f}$
430	209	+d _x	-13	365720	4415490	366150	4415699	366392	4415486
-109	131	+d _x	374	368220	4415490	368111	4415621	368307	4415995
-457	337	-240	-92	370720	4415490	370263	4415827	370510	4415735
148	-159	65	171	373220	4415490	373368	4415331	373432	4415502
24	-273	-297	-79	375720	4415490	375744	4415217	375447	4415138
373	-67	112	-329	378220	4415490	378593	4415423	378704	4415094
155	469	156	-101	380720	4415490	380875	4415959	381031	4415858
-180	438	-371	-465	383220	4415490	383040	4415928	382669	4415463
285	-85	61	355	385720	4415490	386005	+d _y 05	386066	4415760
282	-440	321	-134	388220	4415490	388502	4415550	388823	4414915
449	-156	326	323	365720	4415990	366169	4417834	366495	4418157
200	-85	-230	-125	368220	4415990	368420	4417905	368191	4417780
-364	239	-177	-38	370720	4415990	370356	4418229	370179	4418192
461	458	426	-270	373220	4415990	373681	4418448	374107	4418178
441	-21	236	-362	375720	4415990	376161	4417969	376397	4417607
-221	-272	-202	-102	378220	4415990	377999	4417718	377796	4417615
230	-48	406	-405	380720	4415990	380950	4417942	381356	4417536
90	-121	0	121	383220	4415990	383310	4417869	383311	4417990
498	50	433	455	385720	4415990	386218	4418040	386652	4418495
177	-407	326	-249	388220	4415990	388397	+d _y 3	388723	4417333
.
.
.

Tabla 3. Hoja de cálculo simplificada para la deformación espacial donde $M_i = L_i = 500$ y $d_i = 2500$.

2.2.1.2. MODELO MATEMÁTICO DE TRANSFORMACIÓN ESPACIAL

Para justificar la elección del modelo matemático de transformación espacial, realizaremos antes un breve repaso de las dos transformaciones disponibles en *software* del mercado. Hay que destacar que se intenta emplear las herramientas que se encuentran en estos programas para no tener que programar o utilizar módulos externos difíciles de conseguir. No obstante, se debe seleccionar la transformación que mejor se ajuste al orden de magnitud y ámbito de distorsión necesario.

Para describir mejor los efectos de cada transformación se utilizará la malla de pruebas representada en la figura 5, que muestra superpuestos los vectores de identidad que deformatán dicha malla.

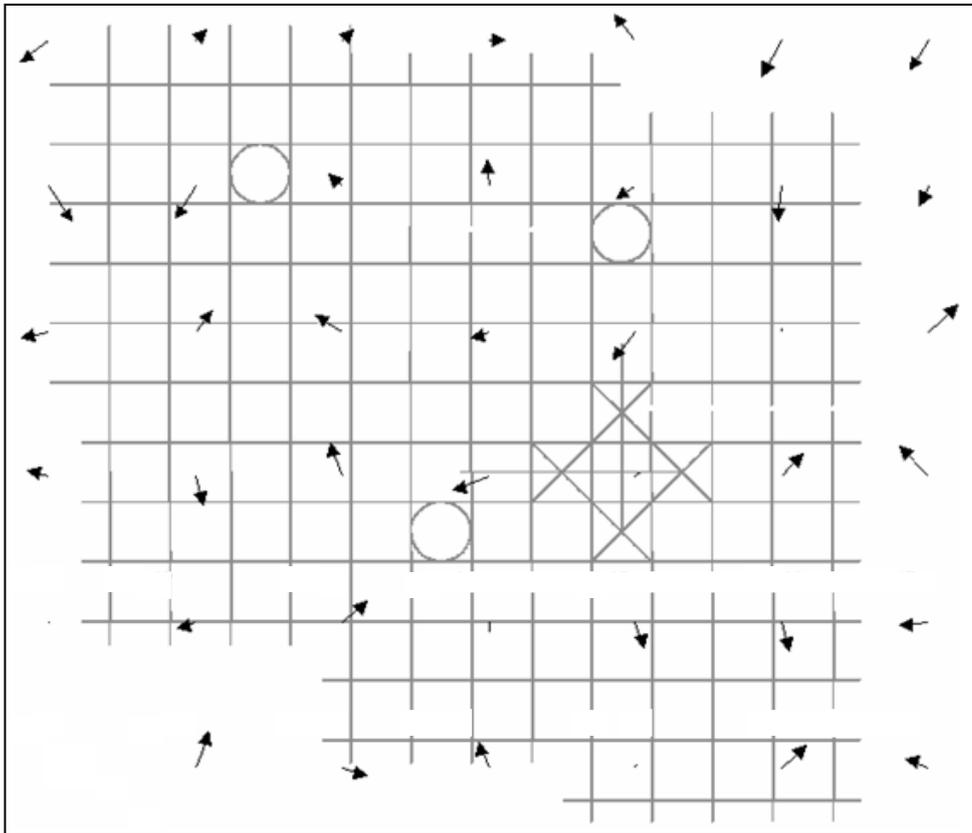


Figura 5. Vectores de identidad y malla de pruebas. Elaboración propia.

2.2.1.2.1. TRANSFORMACIÓN POLINÓMICA

La transformación polinómica está disponible en la mayoría de SIG del mercado y persigue deformar los elementos espaciales, ya sean en formato vectorial o ráster, mediante funciones de interpolación. Esta deformación consiste en hacer coincidir al máximo posible, según la función empleada, los puntos origen/final de los vectores de identidad. Asimismo, la deformación polinómica transformará toda la superficie obligando a acomodarse el resto de coordenadas a esta deformación impuesta.

En la transformación lineal, de semejanza o “Helmert” se obtienen cuatro parámetros, que son dos de translación ($\Delta X_0, \Delta Y_0$), uno de homotecia o escala μ y otro de giro κ .

Su ecuación para un vector de identidad ($X_2 - X_1, Y_2 - Y_1$) es la siguiente:

$$\begin{bmatrix} X_2 \\ Y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \end{bmatrix}_i + \begin{bmatrix} \Delta X_0 \\ \Delta Y_0 \end{bmatrix} + \mu \begin{bmatrix} \cos \kappa & -\sin \kappa \\ \sin \kappa & \cos \kappa \end{bmatrix} \begin{bmatrix} X_1 \\ Y_1 \end{bmatrix}_i \quad \forall i = \{1..n\}$$

Que es un par de ecuaciones lineales de la forma:

$$\begin{aligned} X_2 &= A + B \cdot X_1 + C \cdot Y_1 \\ Y_2 &= D + E \cdot X_1 + F \cdot Y_1 \end{aligned}$$

Donde B es igual a F y C a $-E$ para el caso de semejanza o “Helmert” y distintos en el caso de utilizar diferentes escalas en los ejes, resultando entonces la transformación afín. Estas ecuaciones plantean un primer orden del polinomio que determina la flexibilidad del ajuste y de la transformación.

Acto seguido, se crea un ajuste mediante mínimos cuadrados para reducir todos los vectores de identidad a estos cuatro parámetros (Rodríguez Jordana, 2002). Finalmente, los vectores deformadores serán simplificados a sólo dos siendo así la deformación de esta transformación muy insuficiente para el propósito planteado. La figura 6 muestra el resultado de esta transformación.

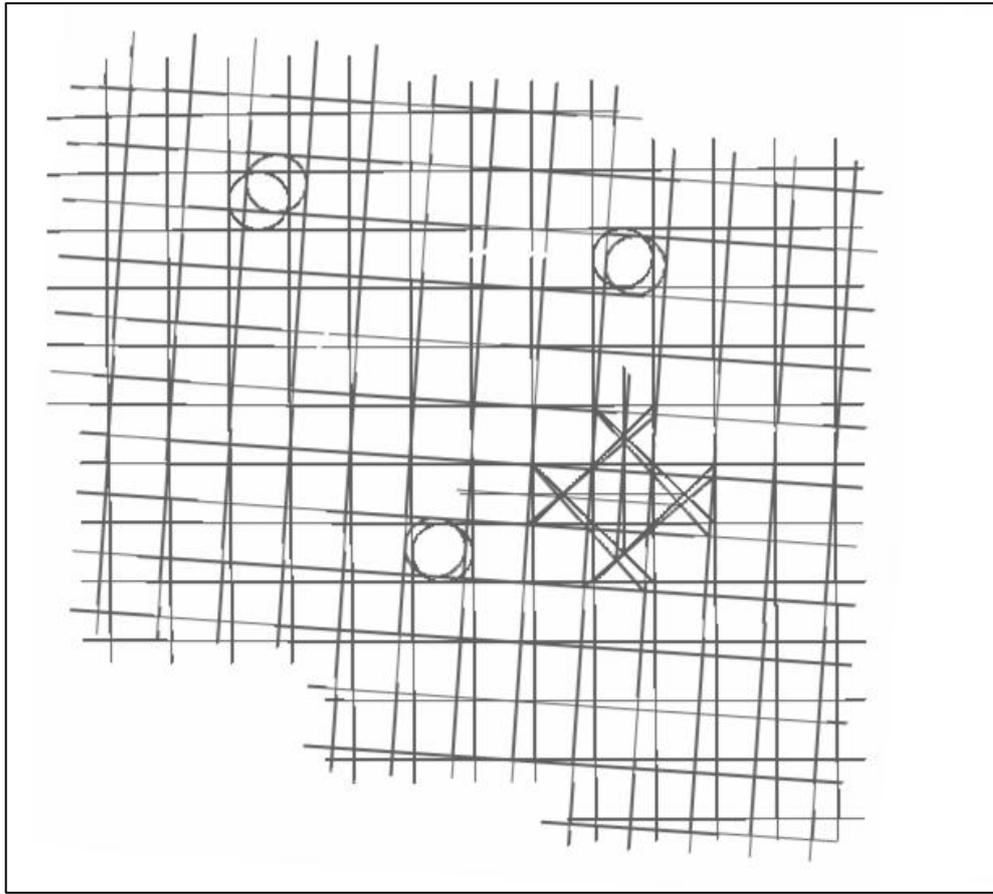


Figura 6. Resultado de la transformación *Helmert* con la malla de pruebas de la figura 5.

Por otro lado, en el polinomio de orden cuadrático o polinomio de segundo grado, se obtienen doce coeficientes para los que son necesarios como mínimo seis vectores de identidad.

Las ecuaciones que se plantean son del tipo:

$$\begin{aligned} X_2 &= A + B \cdot X_1 + C \cdot Y_1 + G \cdot X_1 \cdot Y_1 + H \cdot X_1^2 + I \cdot Y_1^2 \\ Y_2 &= D + E \cdot X_1 + F \cdot Y_1 + J \cdot X_1 \cdot Y_1 + K \cdot X_1^2 + L \cdot Y_1^2 \end{aligned}$$

De la misma manera, el polinomio de orden cúbico o polinomio de grado tres es más complejo y las ecuaciones son similares a las lineales o cuadráticas pero se incluyen términos elevados al cubo.

Por otro lado, como en el proceso se crea un sistema de ecuaciones para ajustar por el método de mínimos cuadrados los parámetros que mejor se adapten a todo el conjunto de vectores identidad, es lógico pensar que un polinomio puede absorber las heterogeneidades que se quieran introducir. Este proceso se fundamenta en hacer la suma de los residuos mínima y así obtener los coeficientes de esta transformación (Rodríguez Jordana, 2002). Por lo tanto, la transformación está representada por el número de coeficientes empleados y se crea un promedio de todos los vectores identidad. La figura 7 muestra la transformación cúbica donde se distorsiona la malla de pruebas de una forma muy global y poco aleatoria.

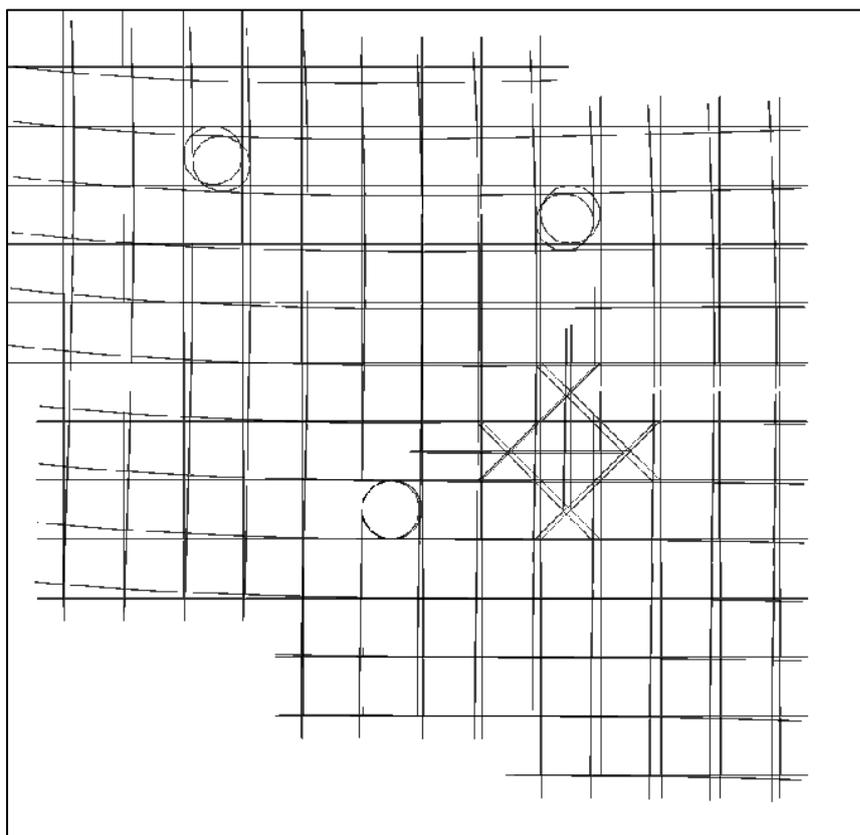


Figura 7. Transformación de polinomio de tercer grado en una malla de pruebas de la figura 5.

Con la finalidad de no distorsionar de manera tan global un mapa y de controlar los desplazamientos, la pregunta planteada en este momento es qué de otro tipo de transformación se podría emplear.

Es necesario tener en cuenta que cuanto mayor sea el grado de los polinomios implicados, más sensible será la transformación a los vectores de identidad creados, pero no se evitará la globalidad. Otros polinomios de mayor grado u otros que emplean una variable compleja (Appelbaum, 1982; Draper et al, 1981), son menos globales, pero se necesitan muchos coeficientes, mucho tiempo de cálculo y no son tan accesibles de encontrar en un SIG como IDRISI o ARCMAP.

2.2.1.2.2. TRANSFORMACIÓN DE DISTORSION

La idea de modelar una distorsión se fundamenta en la existencia de un patrón de comportamiento en la transformación. Modelar esta distorsión puede llevarse a cabo por técnicas de superficies de mínima curvatura, por mínimos cuadrados o por regresión múltiple entre otras (González Matesanz et al, 2004).

Por su complicación y para no hacer demasiado extenso este documento, la transformación por mínima curvatura y la de mínimos cuadrados se comentarán de una forma muy superficial.

La idea de superficies de mínima curvatura proviene de la ingeniería mecánica y la teoría de la elasticidad (Briggs et al, 1974; Dewhurst, 1990; Garcia de Arangoa, 1974; Love, 1994). Esta transformación minimiza la curvatura total sobre una distribución de vectores de identidad considerando una lámina en equilibrio, donde las fuerzas actúan perpendiculares a ella y no existen ni fuerzas de cizalla ni tensiones en la misma. Se describe mediante la ecuación biarmónica siguiente:

$$\frac{\delta^4 w}{\delta x^4} + \frac{2\delta^2 w}{\delta x^2 \delta y^2} + \frac{\delta^4 w}{\delta y^4} = \frac{P}{D}$$

En esta ecuación D es la rigidez a la flexión, P la fuerza y w el desplazamiento, siendo P igual a 0 para los puntos que no sean dato. Se obtiene un polinomio de tercer orden en el que las fuerzas actuantes resultan de la fuerza en el punto:

$$u(X_i, Y_i) = w_i$$

El segundo método a mencionar es la colocación mínimo cuadrática (Morizt, 1973; Lachapelle, 1982), que amplía el modelo mínimo cuadrático a:

$$A \cdot x + s + n = K$$

Donde A es la matriz de diseño, x los parámetros, K los términos independientes, s la señal o distorsión residual y n el ruido o efecto de errores de medida. Esta transformación tiene por objeto eliminar ruido mediante filtrado y cálculo de la señal en otros puntos que no intervienen en el ajuste. Dada una serie de puntos con señal conocida, es posible predecir en otro punto los valores de señal mediante algoritmos matriciales.

La transformación de distorsión más eficiente para nuestro propósito es el ajuste *rubbersheet* (Saafeld , 1985; White et al, 1985). Como se demuestra a continuación, el ajuste *rubbersheet* proporciona deformaciones más locales que los ajustes polinómicos y proporciona la posibilidad de moldear mejor el carácter aleatorio de magnitud y ámbito establecidos así como controlar los desplazamientos de las coordenadas. De esta manera, el modelo matemático de deformación geométrica se acerca mucho más a la idea de una deficiente toma de datos en campo y otras fuentes de error.

El proceso *rubbersheet* se fundamenta en encontrar una transformación entre coordenadas originales y distorsionadas mediante una triangulación de *Delaunay*, determinando un homomorfismo entre las dos partes (ver figura 8). Según esta figura, los vectores deformadores son los vínculos de desplazamiento que definen las nuevas coordenadas.

Mediante *rubbersheeting*, los vectores de identidad se reducirán a cero y los puntos origen y final de cada uno estarán solapados. El resto de coordenadas se ajustará a estas modificaciones. Las deformaciones pueden realizarse en función de la proximidad y longitud de los vínculos de desplazamiento, cumpliendo así los requisitos establecidos. Las coordenadas que menos se moverán están fuera de estos vínculos de desplazamiento y los enlaces de identidad también pueden ser utilizados para ayudar a mantener las características de ciertos lugares sin deformar.

Esta herramienta de distorsión realiza una deformación elástica utilizando dos redes temporales de triángulos irregulares (TIN): “*TIN de cambios en el eje x*” y “*TIN de cambios en el eje y*”. Estas redes se usan para interpolar los cambios en x e y de la deformación de coordenadas que definen los vectores especificados.

Cada triángulo tiene la misma estructura y los orígenes de los vectores se utilizan como esquinas de los triángulos. El nodo de cada triángulo se define por su posición x , y y un valor z . Este valor z es el que interpola la cantidad a ajustar en el eje x o y . Por ejemplo, si el cambio en x es de diez unidades, el valor z para el nodo final de la “*TIN de cambios en el eje x*” es diez.

Una vez que cada nodo tiene valor z , el correspondiente valor en z de cualquier punto que esté en ese triángulo se le interpolará. De igual manera se realiza en la “*TIN de cambios en el eje y*”. Finalmente, el valor interpolado $z_p^x = dx_p$ y $z_p^y = dy_p$ se sumará a la coordenada x e y correspondiente. La figura 8 muestra todo el proceso.

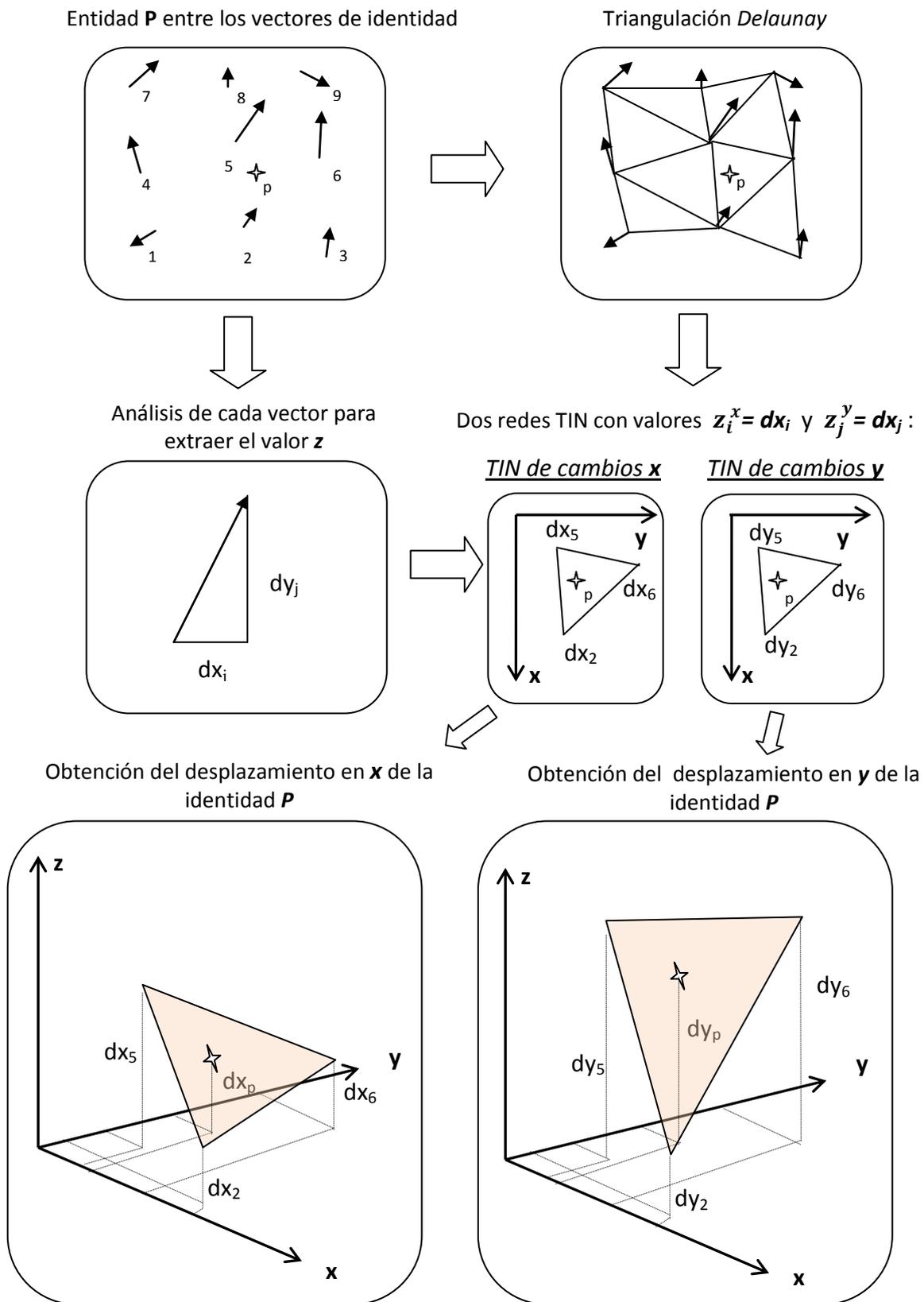


Figura 8. Transformación *rubbersheet* de interpolación. Fuente: elaboración propia.

Para el desarrollo matemático, a los vínculos de identidad anteriormente desarrollados, se hace un cambio de variable $\mathbf{z}_j^x = d\mathbf{x}_j$, $\mathbf{z}_j^y = d\mathbf{y}_j$ y $\bar{\mathbf{v}}_{m,n} = \bar{\mathbf{v}}_j$ quedando:

$$\bar{\mathbf{v}}_{m,n} = (x_{m,n,f} - x_{m,n,o}, y_{m,n,f} - y_{m,n,o}) = (z_j^x, z_j^y)$$

Entonces, dada una entidad \mathbf{P} y las coordenadas del triángulo definido por los puntos **1**, **2** y **3** al que pertenece:

$$(x_p, y_p, z_p^i)$$

$$(x_1, y_1, z_1^i)$$

$$(x_2, y_2, z_2^i)$$

$$(x_3, y_3, z_3^i)$$

La ecuación de la superficie que contiene el triángulo se puede obtener de la ecuación general del plano:

$$A \cdot x + B \cdot y + C \cdot z = D$$

Donde \mathbf{D} es una constante a determinar y \mathbf{A} , \mathbf{B} y \mathbf{C} son las coordenadas del vector perpendicular al plano que se calcula mediante el producto vectorial de dos vectores contenidos en el mismo.

$$\overrightarrow{12} = (x_2 - x_1, y_2 - y_1, z_2^i - z_1^i)$$

$$\overrightarrow{32} = (x_2 - x_3, y_2 - y_3, z_2^i - z_3^i)$$

$$\overrightarrow{12} \wedge \overrightarrow{32} = \vec{i} \cdot A + \vec{j} \cdot B + \vec{k} \cdot C = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_2 - x_1 & y_2 - y_1 & z_2^i - z_1^i \\ x_2 - x_3 & y_2 - y_3 & z_2^i - z_3^i \end{vmatrix} =$$

$$= \vec{i} [(y_2 - y_1)(z_2^i - z_3^i) - (z_2^i - z_1^i)(y_2 - y_3)] - \vec{j} [(x_2 - x_1)(z_2^i - z_3^i) - (z_2^i - z_1^i)(x_2 - x_3)] + \vec{k} [(x_2 - x_1)(y_2 - y_3) - (y_2 - y_1)(x_2 - x_3)]$$

Las coordenadas del vector perpendicular al plano son **(A,B,C)**:

$$A = [(y_2 - y_1)(z_2^i - z_3^i) - (z_2^i - z_1^i)(y_2 - y_3)]$$

$$B = [(x_2 - x_1)(z_2^i - z_3^i) - (z_2^i - z_1^i)(x_2 - x_3)]$$

$$C = [(x_2 - x_1)(y_2 - y_3) - (y_2 - y_1)(x_2 - x_3)]$$

El valor **D** se obtiene despejando la ecuación general del plano tras substituir un punto conocido (por ejemplo, la coordenada del punto **1**):

$$D = x_1[(y_2 - y_1)(z_2^i - z_3^i) - (z_2^i - z_1^i)(y_2 - y_3)] - y_1[(x_2 - x_1)(z_2^i - z_3^i) - (z_2^i - z_1^i)(x_2 - x_3)] + z_1 [(x_2 - x_1)(y_2 - y_3) - (y_2 - y_1)(x_2 - x_3)]$$

Finalmente, los valores $z_p^x = dx_p$ y $z_p^y = dy_p$ se obtienen substituyendo las coordenadas del punto **p** en la ecuación del plano obtenido en cada triángulo.

Que despejando queda:

$$z_p^i = \frac{-x_p \cdot A - y_p \cdot B + D}{C} \quad \forall i = \{x, y\}$$

Estas coordenadas proporcionan los pares de valores de desplazamiento de toda coordenada **P** en el mapa a distorsionar. La figura 9 muestra el proceso gráficamente.

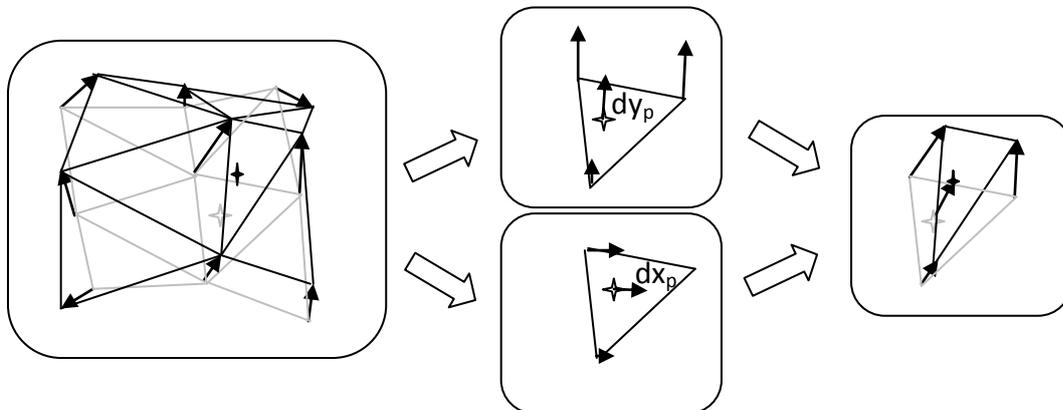


Figura 9. Desplazamiento para un punto **P** cualquiera en un triángulo de distorsión. Elaboración propia.

El algoritmo de cálculo para cada coordenada del mapa a distorsionar empieza identificando cada píxel con el triángulo al que pertenece y luego sólo queda calcular el valor dx_p y dy_p con las fórmulas anteriores. Asimismo, en el software ARCGIS 9.2 se puede encontrar esta herramienta que sólo exige el archivo con las coordenadas de los vínculos de identidad en formato txt.

La figura 10 muestra la transformación *rubbersheet* mediante los mismos vectores de identidad usados en las otras pruebas. Se puede observar que dicha transformación formaliza una nueva disposición espacial que distorsiona de una forma local y aleatoria en función de los vectores deformadores.

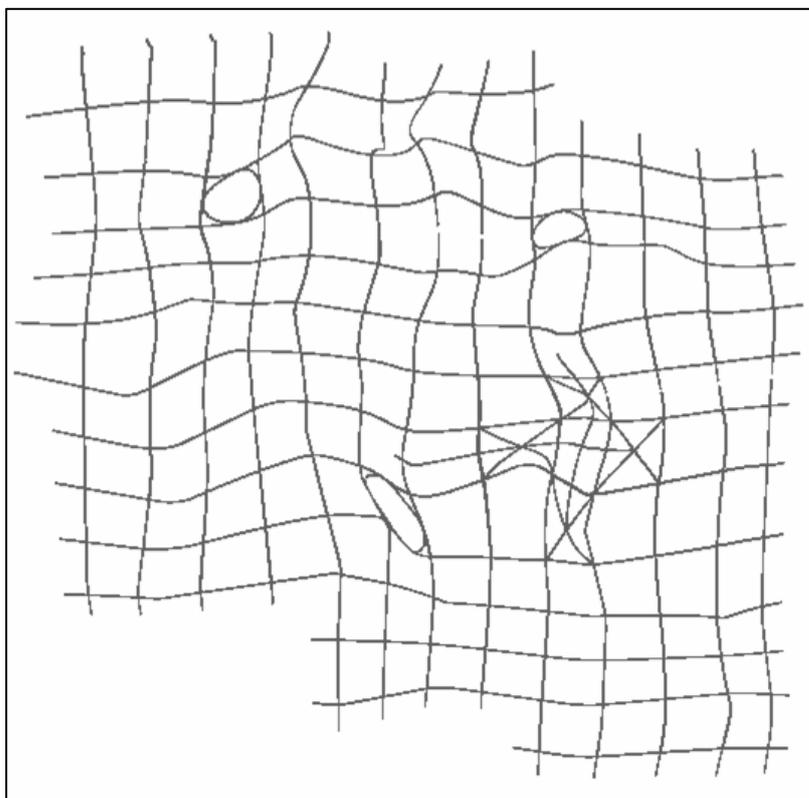


Figura 10. Transformación *rubbersheet* por interpolación en la malla de pruebas de la figura 5.

En la figura 11 podemos observar los efectos de la transformación *rubbersheet* sobre las entidades del mapa de tipo de suelos. Para poder apreciar las deformaciones que esta transformación proporciona, se han superpuesto al mapa deformado los contornos de las entidades del mapa original.

En esta transformación se conservan los límites de magnitud establecidos y ningún elemento se desplaza más que la distancia que los vectores han deformado una identidad. Por tanto, si los vectores deformadores no sobrepasan los 700 metros, ningún elemento se desplazará más de 700 metros.

Se puede concluir que la deformación por *rubbersheet* es la más adecuada en cuanto a deformación geométrica y, por tanto, la que se va a aplicar en este estudio.

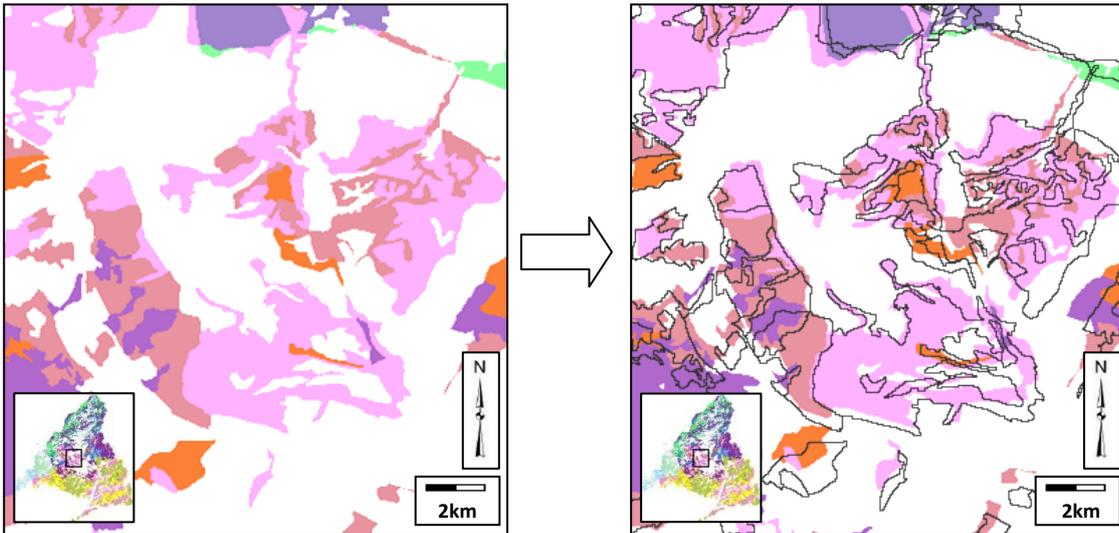


Figura 11. Efecto de la distorsión geométrica en una zona del mapa de tipo de suelos.

Para las entidades lineales o puntuales se realiza el mismo proceso. La figura 12 muestra en niveles de gris, la diferencia de valores que proporcionan el factor distancias a carreteras deformado y sin deformar.

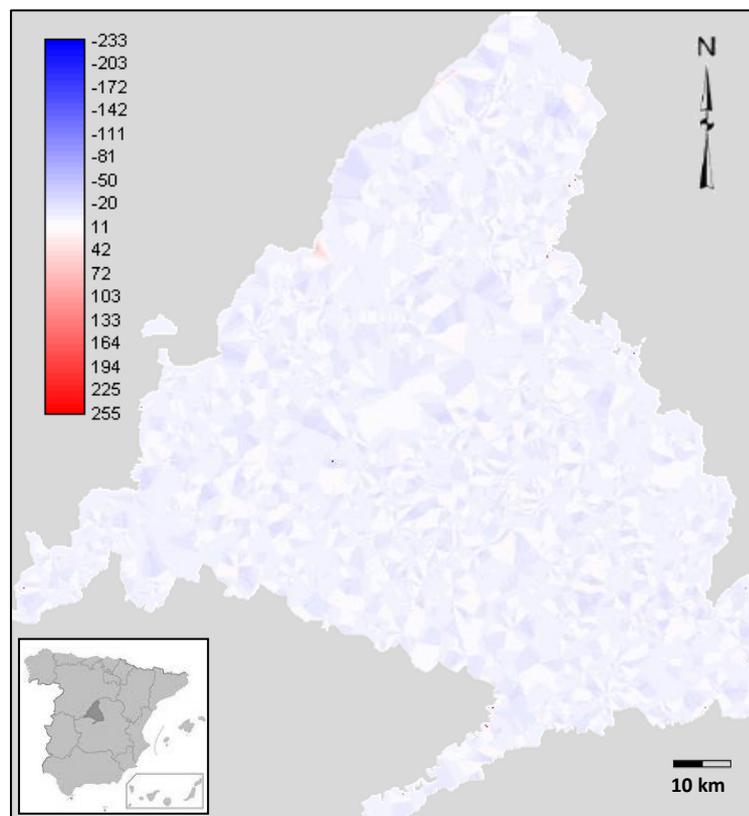


Figura 12. Diferencia entre factor distancia a carreteras original y del distorsionado.

2.2.2. DEFORMACIÓN DE MODELOS DIGITALES DEL TERRENO

En el modelo analizado se emplea un Modelo Digital del Terreno que representa las elevaciones del terreno. A partir de este MDE se derivan dos Modelos Digitales Derivados para crear el factor de pendiente y de orientación del terreno. Es razonable pensar que si se modifica el MDE original, los Modelos Digitales Derivados también serán modificados. Con respecto a la deformación de estas variables, la incertidumbre no es del mismo tipo que la de las variables de tipo polígono, línea o punto y no es necesario distinguir entre distorsión geométrica y temática. Es por esto que esta variable se presenta en una sección individual y con un método de distorsión diferente.

Algunos autores añaden una perturbación aleatoria de distribución conocida para analizar los errores de modelos derivados de un MDE. Por ejemplo, Lee (1992) emplea una malla obtenida a partir de clústeres. A continuación, Lee genera un MDE afectado por esta distorsión y obtiene su modelo derivado para comparar el modelo obtenido con el original y realizar así el análisis de error.

Así pues, se propone distorsionar el MDE para obtener las variables derivadas pendiente y orientación deformadas. El MDE a distorsionar es una matriz regular a la que se puede introducir una perturbación de error mediante una simulación de errores. Se trata de sumar un modelo digital de superficie de error creado a partir de puntos que varíen entre dos cantidades.

Lee et al (1992) simulan la magnitud de los errores mediante un generador de números aleatorios. Los números aleatorios generados se vuelven a escalar para estar dentro de una magnitud especificada y luego son cambiados y controlados para tener varios grados de clústeres espaciales. Esta autocorrelación espacial se realiza mediante el índice de Moran (Goodchild, 1986).

El método propuesto en este proyecto simplifica la superficie de error perturbadora a una red de triángulos irregulares de magnitud y distribución controlada.

2.2.2.1. MARGEN DE INCERTIDUMBRE VERTICAL

Para poder abordar la distorsión del MDE es necesario conocer la estructura de la capa representada. El MDE es una estructura numérica de datos que representa la distribución espacial de la altitud de la superficie del terreno de la forma:

$$Z = \delta(X, Y)$$

Donde Z es la altitud y δ es una función que relaciona la variable con su localización geográfica.

Los errores en el MDE pueden provenir de diversas fuentes y se pueden obtener diferentes precisiones en función de la forma del terreno. Mukai et al. (1989) obtienen un Error Medio Cuadrático (EMC) de 26 metros en los Alpes centrales de Japón en restituciones a partir de imágenes SPOT-HRV. En cambio, Kubik y Wu (1995) o Priebbenow y Clerici (1988) en Australia, obtienen alrededor de 6 metros de EMC. Por otro lado, Sasowsky et al (1992) en Alaska obtienen un error de 19 metros de EMC.

Actualmente, muchos trabajos tratan de estudiar las mejores técnicas de explotación de los datos estereoscópicos espaciales para la generación de MDE mediante técnicas automáticas de correlación, Cuartero Sáez (2003). En interferometría radar, los MDE obtenidos por TOPSAR tienen 5 metros de EMC en la componente horizontal y hasta 3 metros en la vertical. Estos valores se pueden emplear en la distorsión cuando no se dispone de los metadatos que informan de la precisión del MDE.

Las fuentes secundarias de error pueden provenir de la propia digitalización, de errores producidos por deformación, referenciación, equivocaciones topológicas, de imprecisión, de generalización, de procesos fotogramétricos o propagación del error planimétrico a la altimetría. Tahiri y Donnai (1995) señalan un error de 5 a 10 metros para mapas impresos de escala 1:50.000, con 10 metros de equidistancia en curvas de nivel.

Por lo tanto, tras los ejemplos de error altimétrico mostrados y atendiendo a un tamaño de píxel de 50 metros, la distorsión que se propone aplicar en el MDE de la Comunidad de Madrid tiene un rango de variación en altura de 10 metros y un máximo de variación de pendiente del 10 %.

2.2.2.2. DISTORSIÓN A PARTIR DE UNA RED DE TRIÁNGULOS IRREGULARES

Como ya se ha comentado anteriormente, esta distorsión pretende modificar el MDE original para que los Modelos Digitales Derivados también varíen. Para ello, se procede a sumar al MDE un Modelo Digital de Incertidumbre, creado a partir de puntos que varíen entre dos cantidades.

Este método propuesto crea la superficie de error a partir de una red de triángulos irregulares de magnitud y distribución controlada. Para ello, se propone una matriz de celdas cuadradas de 400 metros de lado, separadas cada 100 metros, donde en cada celda se introduce un punto aleatorio en planimetría y altimetría. Siguiendo el límite de incertidumbre establecido, el rango de variación en altimetría propone valores entre -5 y +5. Por otro lado y en caso extremo, los puntos van a estar a un mínimo de 100 metros de proximidad y un máximo de unos 1000 metros. Se obtendrá por tanto, un valor máximo en pendiente del 10 %. Las figuras 13 y 14 muestran la geometría diseñada para la magnitud y distribución de los puntos de error de cota, donde un punto está colocado aleatoriamente en cota y en planimetría dentro de cada región.

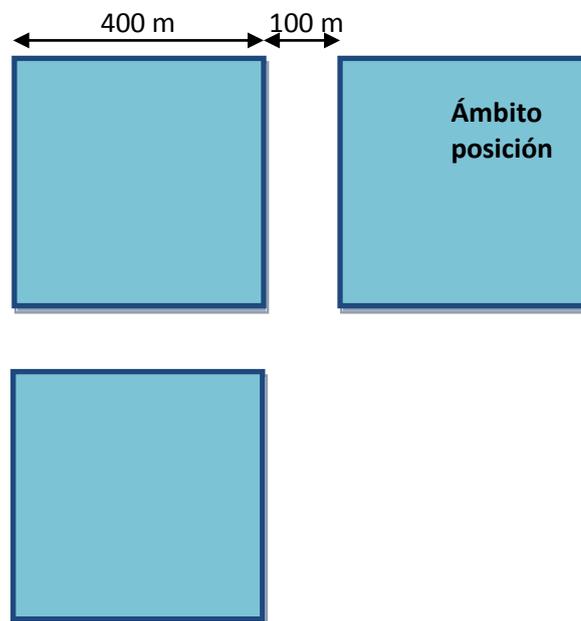


Figura 13. Distribución de los puntos acotados en la deformación del Modelo Digital de Incertidumbre.

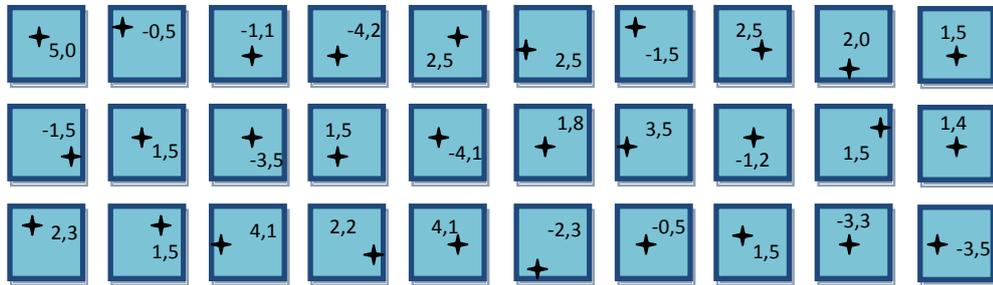


Figura 14. Ejemplo gráfico de treinta puntos acotados en sus correspondientes celdas.

Es un caso similar al de obtención de los vectores de deformación pero los centros de las celdas están separados quinientos metros en lugar de 2,5 kilómetros.

Las coordenadas para cada punto de fila m y columna n son:

$$\bar{v}_{m,n} = (x_{m,n}, y_{m,n})$$

Que son los valores para cualquier punto aleatorio con coordenadas del tipo:

$$x_{m,n} = x_{\emptyset} + n \cdot d_x + E_{x,m,n}$$

$$y_{m,n} = y_{\emptyset} + m \cdot d_y + E_{y,m,n}$$

Donde:

x_{\emptyset} e y_{\emptyset} son las coordenadas de la esquina del mapa y d_x y d_y son los incrementos de distancia entre cada celda.

Los valores d_x y d_y se pueden obtener dividiendo la longitud de la imagen por el número de puntos que se quiera introducir en cada eje x o y .

Para la aleatoriedad, se usa el mismo error definido en el capítulo de vectores de deformación (ver capítulo 2.2.1.1.):

$$E_{i,m,n} = M_i \cdot A_{i,m,n} \cdot (-1)^{\text{ent}(100 \cdot A_{i,m,n})} \quad \forall i = \{x, y\}$$

Recordando que \mathbf{A} es un número aleatorio entre cero y uno y \mathbf{M} la semimagnitud máxima de error, sustituyendo queda:

$$x_{m,n} = x_{\emptyset} + n \cdot d_x + M_x \cdot A_{x,m,n} (-1)^{ent(100A_{x,m,n})}$$

$$y_{m,n} = y_{\emptyset} + m \cdot d_y + M_y \cdot A_{y,m,n} (-1)^{ent(100A_{y,m,n})}$$

Asimismo, se propone usar el valor $\mathbf{M}_x = \mathbf{M}_y$ e igual a 200 metros para crear las celdas cuadradas de 400 metros de lado.

Finalmente, la coordenada de elevación distorsionada es:

$$z_{m,n} = M_z \cdot A_{m,n} \cdot (-1)^{ent(100 \cdot A_{m,n})}$$

Con estas dimensiones se consigue para toda la imagen un centenar de miles de puntos, repartidos aleatoriamente y con una cota aleatoria entre $-\mathbf{M}_z$ y $+\mathbf{M}_z$. Como ya se ha comentado, para \mathbf{M}_z se usa el valor de 5 metros de incertidumbre.

Para calcular todos los puntos de la imagen, sus coordenadas se pueden calcular fácilmente programando en cualquier lenguaje de programación con bucles del tipo *for* o *while*. El algoritmo para crear las coordenadas tiene la estructura siguiente:

```

INTRODUCE  $x_{\emptyset}, y_{\emptyset}, d_x, M_x, d_y, M_y, L_x$ 
DESDE 0 HASTA m
  DESDE 0 HASTA n
    A1=RANDOM
    A2=RANDOM
    A3=RANDOM
    IMPRIMIR  $x_{\emptyset} + n \cdot d_x + M_x \cdot A1 \cdot (-1)^{ent(100 \cdot A1)}$ 
    IMPRIMIR " "
    IMPRIMIR  $y_{\emptyset} + m \cdot d_y + M_y \cdot A2 \cdot (-1)^{ent(100 \cdot A2)}$ 
    IMPRIMIR " "
    IMPRIMIR  $M_z \cdot A3 \cdot (-1)^{ent(100 \cdot A3)}$ 
    IMPRIMIR (SALTO DE LINEA)
  FINDESDE
FINDESDE

```

El fichero de texto obtenido se puede cargar en uno de los muchos programas existentes que contienen la herramienta de triangulación de puntos acotados. De esta manera se obtiene el Modelo Digital de Incertidumbre que se suma al MDE para distorsionarlo.

La creación de estos puntos también se puede realizar fácilmente aplicando el algoritmo en una hoja de cálculo. En la tabla 4, la creación de los puntos se desarrolla en *excel* de manera muy similar a la empleada en la creación de los vectores de deformación espacial.

$E_{x,m,n}$	$E_{y,m,n}$	$x_o + n \cdot d_x$	$y_o + n \cdot d_y$	$x_{m,n}$	$y_{m,n}$	$z_{m,n}$
+d _x	63	365720	4415490	365790	4415427	-1
+d _x	45	366220	4415490	366061	4415535	4
	-25	366720	4415490	366754	4415465	1
152	149	367220	4415490	367372	4415639	1
168	-76	367720	4415490	367888	4415414	-4
-124	138	368220	4415490	368096	4415628	-3
-169	173	368720	4415490	368551	4415663	3
-177	-120	369220	4415490	369043	4415370	1
41	195	369720	4415490	369761	441585	-2
127	-132	370220	4415490	370347	441558	4
-63	-176	365720	4415990	365657	4415814	1
4	-23	366220	4415990	366224	4415967	5
133	-80	366720	4415990	366853	4415910	-3
-176	123	367220	4415990	367044	4416113	-2
100	-15	367720	4415990	367820	4415975	-3
174	126	368220	4415990	368394	4416116	-1
111	-11	368720	4415990	368831	4415979	-3
25	76	369220	4415990	369245	4416066	-3
-154	-84	369720	4415990	369566	4415906	-4
-101	119	370220	4415990	370119	4415109	-2
.
.
.

Tabla 4. Hoja de cálculo simplificada para obtener los puntos acotados para la distorsión del MDE. $M_z = 5$, $M_x = M_y = 200$, $d_x = d_y = 500$.

Las tres últimas columnas de la tabla 4 se importan en formato *txt* al *software* informático que vectoriza los puntos y triangula la superficie. El paso de elementos vectoriales a una superficie continua se debe obtener mediante una función de interpolación que proporcione la información no existente entre cada punto acotado. La interpolación se debe realizar mediante una estimación del valor en cada punto en función de los puntos del entorno.

La materialización de una TIN en una capa ráster se realiza estimando valores mediante interpolación entre los puntos acotados. Existen varios métodos a la hora de interpolar redes:

- Estimación en función de la distancia, Kitanidis (1997).
- Interpolación lineal: Donde la altitud se estima directamente a partir de la ecuación del plano definido por los tres vértices de cada triángulo de *Delaunay*. Es muy rápida pero provoca discontinuidades entre triángulos con cambios bruscos de orientación y pendiente.

- Interpolación quíntica: En la que se considera la superficie como un continuo, suavizando las zonas de los vértices y lados gracias al uso de un polinomio de quinto grado que genera el MDT sin facetas.

La figura 15 muestra el resultado del método de interpolación lineal mediante el comando *tinsurf* de IDRISI. Resumiendo muy brevemente, este resultado corresponde a un Modelo Digital de Incertidumbre que ha sido generado por interpolación lineal de una TIN obtenida a partir de puntos acotados aleatorios.

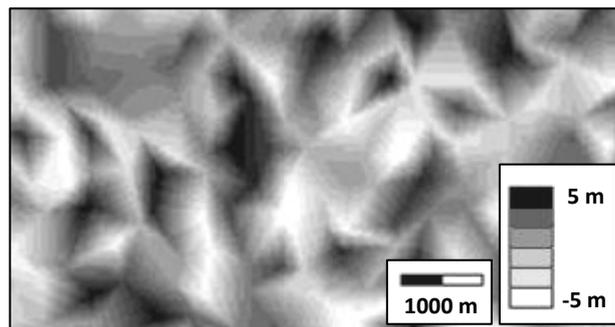


Figura 15. Sección del Modelo Digital de Incertidumbre. Elaboración propia.

Una vez obtenida la superficie de error para cada distorsión, sólo queda sumar los valores al MDE y obtener los Modelos Digitales Derivados distorsionados.

2.2.2.2.1. DISTORSIÓN DEL FACTOR PENDIENTE

En primer lugar, para distorsionar el factor pendiente se debe modificar el MDE primitivo. Para ello, en IDRISI se suma con la función **overlay** la superficie de error aleatorio al MDE. Como se muestra en la figura 16, los procedimientos para la creación de la superficie de error son, primero **xyzidrisi** para vectorizar el archivo de texto con el listado de coordenadas, a continuación el comando **tin** para crear la red de triángulos y finalmente **tinsurf** para rasterizar la capa vectorial. El proceso de obtención del factor pendiente es tan simple como aplicar **slope** y normalizar con **fuzzy**. Finalmente se enmascara el contorno de la comunidad multiplicando con **overlay** para limitar la zona de estudio. La Figura 17 muestra la diferencia del valor de adecuación para ver los cambios producidos respecto al factor sin modificar.

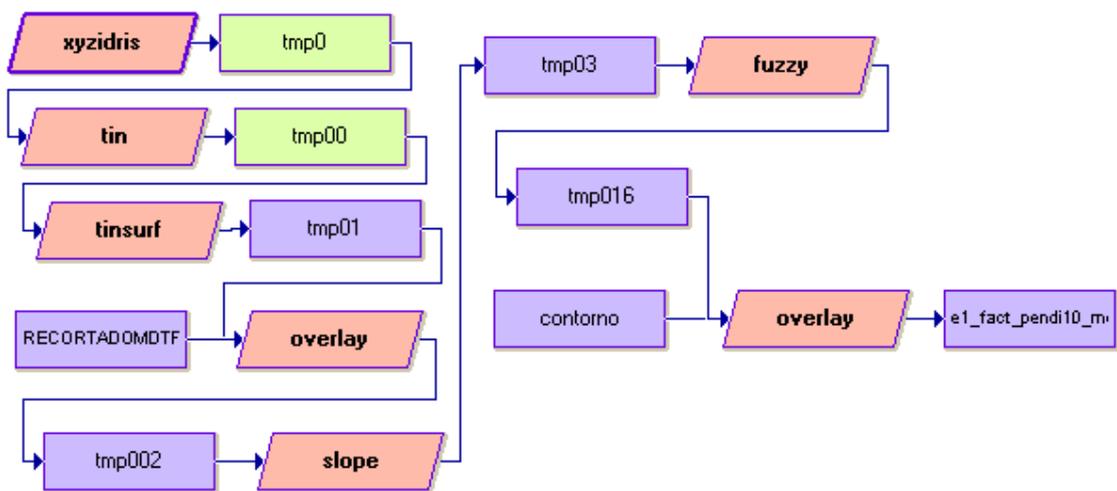


Figura 16. Reproducción mediante la macro del factor de pendiente.

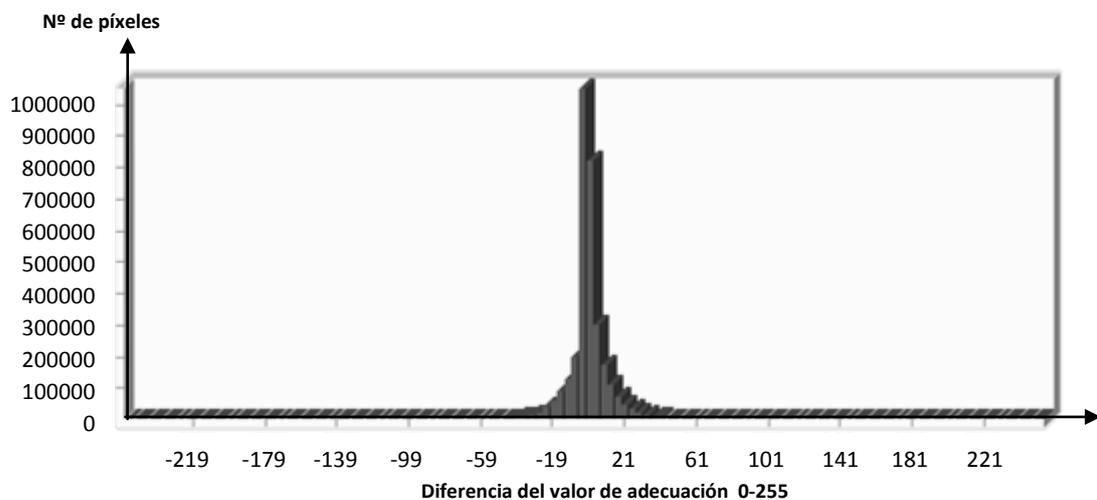


Figura 17. Histograma de la diferencia de valores de adecuación entre el factor pendiente modificado y sin modificar.

2.2.2.2.2. DISTORSIÓN DEL FACTOR ORIENTACION

El factor de orientación se genera del mismo modo que el factor pendiente. La diferencia es que no se usa el comando *slope* sino *aspect*. La diferencia del valor de adecuación se muestra en la figura18 y no se aprecian valores fuera de lo normal.

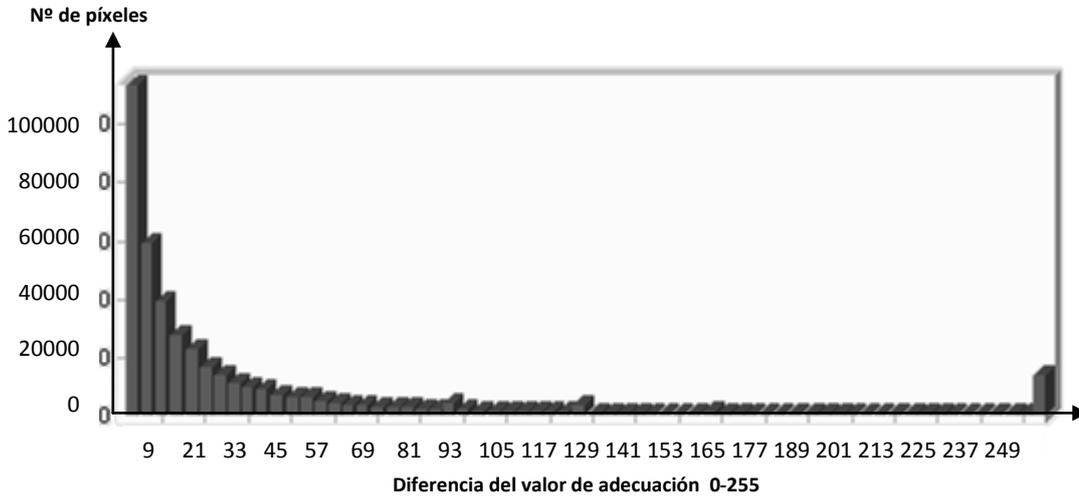


Figura 18. Valores positivos del histograma de la diferencia de valores de adecuación entre el factor orientación modificado y sin modificar.

2.2.3. DEFORMACIÓN TEMÁTICA

La deformación temática pretende simular la incertidumbre de la clasificación de las entidades de un mapa. En las siguientes secciones se muestran dos métodos empleados para la perturbación de datos temáticos. El primer método emplea una técnica que introduce ruido temático (sección 2.2.3.1.) y el segundo, realiza cambios en los atributos de las entidades (sección 2.2.3.2.). Estos métodos se exponen mediante algoritmos y hojas de cálculo. Finalmente, se empleará la mejor propuesta para la perturbación temática de las variables iniciales del modelo a analizar.

Los fallos introducen errores obvios y fáciles de chequear, pero existen errores de procedencia más fina que sólo pueden ser detectados cuando se trabaja muy íntimamente con los datos. Otros errores de origen más difícil de averiguar son los obtenidos como resultado de ciertos tipos de procesos. Su detección requiere un conocimiento íntimo no solo de los datos, sino también de los modelos, la estructura de datos y algoritmos utilizados. Como consecuencia, estos errores pasan desapercibidos para la mayor parte de los usuarios.

El origen de los errores de la componente temática de los datos espaciales puede provenir de diversos factores:

- La medición en radiómetros, termómetros, gravímetros, contadores u otros, son datos de baja calidad y/o los observadores son imprecisos.
- Los datos de campo afectados por el factor humano es uno de los errores más importantes en la recolección de datos basados en la intuición, puesto que algunos observadores son más perceptivos que otros a la hora de clasificar.
- La calidad de los resultados en el laboratorio se supone superior a la realizada en campo y los que se llevan a cabo en distintos laboratorios no son tan reproducibles.
- El área de estudio se puede detallar o generalizar en función de la uniformidad del muestreo de datos.
- En mapas continuos obtenidos a partir de interpolación, la densidad de información es un factor importante.
- Los datos pueden estar tomados en diferentes épocas y por tanto incluyen cambios no previstos.
- Muchos datos no son relevantes para el propósito que tiene su utilización y sustituyen a otros más caros o inexistentes.
- Muchos datos no son accesibles por ser secreto de estado, problemas militares, políticos u otros.

Del mismo modo que en la deformación espacial, es necesario conocer algunas reglas y conceptos esenciales de tal manera que se utilice un modelo de distorsión adecuado.

Entre los conceptos básicos se pueden destacar algunos como:

- Los cambios deben seguir unas reglas probabilísticas de incertidumbre en función de la probabilidad de cambio entre sus clases temáticas.
- Hay diferentes valores de probabilidad de cambio entre clases. Por ejemplo, es menos probable que una superficie de agua sea otro tipo de cubierta que un suelo descubierto sea un cultivo de secano.
- Los elementos pueden tener menor error en su clasificación si su tamaño es muy grande. Por tanto, el error en la clasificación temática puede depender del tamaño o de otro atributo de cada entidad.

2.2.3.1. DISTORSIÓN TEMÁTICA A NIVEL PÍXEL

El proceso empleado por Gómez Delgado y Bosque Sendra (2004) empieza estableciendo el volumen de píxeles a distorsionar de un dos a cuatro por ciento del área de estudio.

Para variables de tipo cualitativo, se lleva a cabo un cambio de categoría aleatorio, asignando nuevos identificadores a cada uno de los píxeles. Esto se realiza mediante una distribución de valores de tipo rectilínea para que todos los valores tengan la misma probabilidad de ser seleccionados y con valor máximo el máximo número de categorías existente en la imagen original. La asignación se realiza a nivel de píxel y no se aplica la probabilidad de asignación de clase ni restricciones por atributos (área, longitud, etcétera).

Esta distorsión también se puede aplicar en mapas de variable cuantitativa, donde se modifica un porcentaje en función del valor máximo registrado en cada factor original con valores extraídos de una tabla de números aleatorios.

Para introducir el error en cada variable, ya sea esta cualitativa o cuantitativa, se debe saber la cantidad de píxeles a distorsionar siendo, según un escenario “optimista”, el cuatro por ciento del área de estudio. Estos píxeles de “ruido” crean un efecto de “sal y pimienta”, que se seleccionan de manera aleatoria con el comando *sample* del software IDRISI (Ver figura 19).

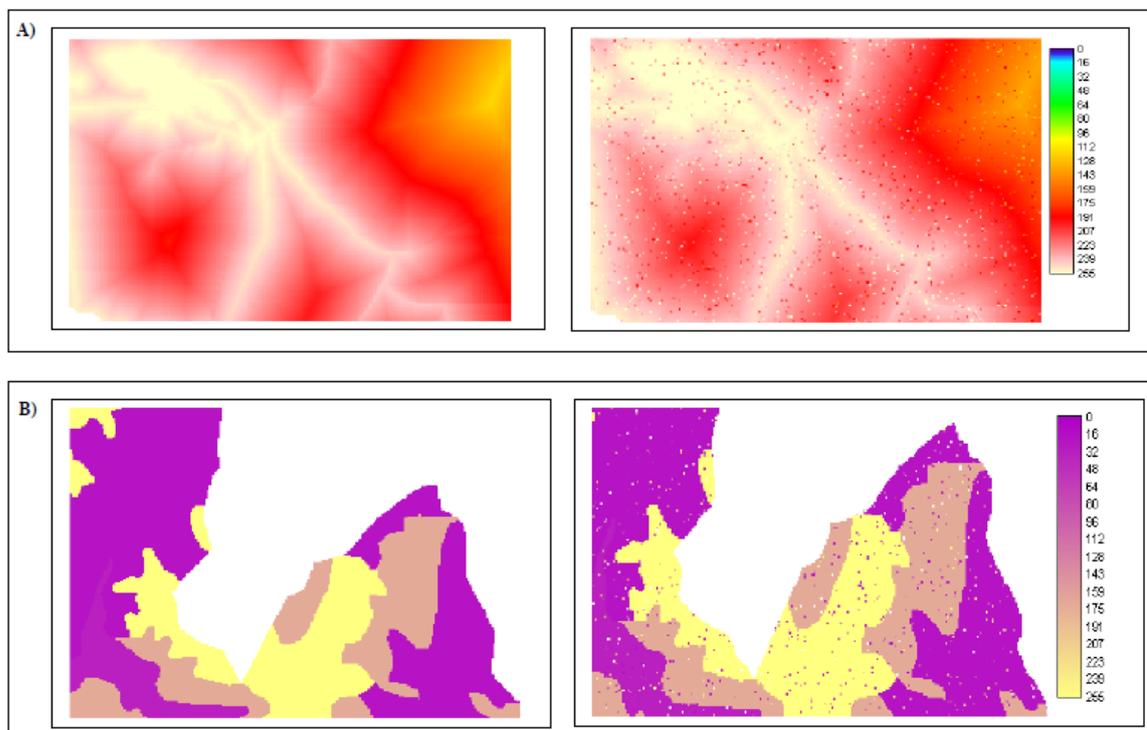


Figura 19. Detalle del efecto de introducción de “ruido” temático. En la parte de la izquierda aparece el factor original y a la derecha el modificado. Fuente: Gómez y Bosque (2004).

2.2.3.2. DISTORSIÓN TEMÁTICA A NIVEL DE ENTIDAD

El objetivo principal de este capítulo es describir una metodología alternativa a la explicada en la sección anterior. En lugar de realizar variaciones a nivel píxel, se propone una perturbación dirigida, que varíe los atributos de cada identidad, con la posibilidad de restringir el cambio en función de las características de cada elemento.

Esta perturbación dirigida utiliza como base para las variables de tipo cualitativo una matriz de confusión, empleada habitualmente para verificar la fiabilidad de la clasificación de las entidades. Es sabido que la matriz de confusión se obtiene tomando como verdad terreno lo que se entiende como la asignación temática y como clasificación, la realizada en el mapa, Chuvieco (2008). Por lo tanto, este estadístico puede informar de la probabilidad de cambio entre clases y ver cómo de buenas son las clasificaciones realizadas. Las filas de esta matriz reflejan la cartografía clasificada y las columnas la verdad terreno. La diagonal principal recoge los polígonos que se han cartografiado correctamente y el resto de valores, reflejan los errores cometidos (ver tablas 5, 7, 9, u 11). Así pues, se pretende introducir un error temático en función de la fiabilidad de la clasificación de entidades que nos proporciona la matriz de confusión de cada mapa.

Otra de las diferencias con la metodología de Gómez Delgado y Bosque Sendra (2004) es que se pretende incorporar restricciones a determinados cambios en función de características de cada entidad. De esta manera se evita cambiar el atributo a polígonos que superen un área, pues es poco probable que un polígono urbano, tal y como es el de Madrid, cambie erróneamente a otra clasificación, ya que dicho error hubiese sido detectado y corregido. De la misma forma, se evitará realizar cambios cuantitativos poco probables. Así y como se explicará a continuación, la cantidad de distorsión en el volumen de población de cada núcleo urbano se realizará de manera controlada y no superando unas cantidades que hubiesen sido fácilmente detectadas y corregidas. Volviendo al ejemplo anterior, es poco probable que si el núcleo urbano de Madrid apareciese con un volumen de población de cinco mil habitantes hubiese pasado desapercibido.

Así pues, en función de la incertidumbre temática, el proceso de introducción de error se fundamenta en la aplicación de cambios de clase asignada en cada entidad. Para implementar esta distorsión se pueden emplear hojas de cálculo o programas en algún lenguaje de programación, implementando los algoritmos desarrollados más adelante. El empleo de números aleatorios es de gran utilidad a la hora de introducir probabilidades de cambio temático, pues el cambio se realiza en un algoritmo condicional que asigna otro atributo en función del valor aleatorio obtenido. Por ejemplo, si la probabilidad de cambiar de clase es de un 20 %, un valor aleatorio que esté entre 0 y 80 % no permite cambiar el atributo, pero si está entre 80 y 100 %, el atributo cambia a otra clase.

En las siguientes secciones se describe la distorsión realizada a cada uno de los factores incluidos en el modelo original, de entre los cuales, el factor de vulnerabilidad a la contaminación de los mantos acuíferos no se distorsiona porque todos polígonos tienen una extensión muy grande y, por tanto, cualquier error hubiera sido detectado con facilidad.

2.2.3.2.1. DISTORSIÓN DEL FACTOR TIPO DE SUELOS

Con este factor se intenta que las nuevas zonas urbanas no sean asignadas, hasta donde sea posible, en los suelos de mayor productividad agrícola.

La tabla 5 muestra la matriz de confusión empleada para la distorsión temática de este mapa. Esta matriz se ha creado empleando documentación del *Reference Soil Group WRB* (Fiedler) y consultas a especialistas. El estudio realizado informa de la posibilidad de encontrar más error en la clasificación de Anthrosoles y Regosoles.

Para aplicar la asignación a otras clases, los valores de la tabla 5 proponen valorar un 70 % de confianza en Anthrosoles y Regosoles y al resto, aplicarle un 82 %. Por otro lado, la probabilidad de ser asignado a otra clase, se reparte con un valor del 5 % para Anthrosoles y Regosoles y un 2 % para el resto.

	Luvisoles	Fluvisoles	Cambi y Calcisoles	Rego y leptosoles	Anthro y Alisoles	Gypsi y Gleysoles	Urb. , parques y Embalses
Luvisoles (6)	82	2	2	5	5	2	2
Fluvisoles (96)	2	82	2	5	5	2	2
Cambri y calcisoles (153)	2	2	82	5	5	2	2
Rego y leptosoles (185)	5	5	5	70	5	5	5
Anthro y Alisoles (217)	5	5	5	5	70	5	5
Gypsi y Gleysoles (236)	2	2	2	5	5	82	2
Urb. Parques y Emb. (255)	2	2	2	5	5	2	82

Tabla 5. Matriz de confusión empleada para distorsionar el mapa tipo de suelos. Valores en %

En la tabla 5, los tipos de suelos están acompañados por el valor de adecuación final de las entidades (valores entre paréntesis). Estos valores provienen del identificador del tipo de suelo, que cambia en función de los valores de la matriz de confusión.

El siguiente algoritmo podría ser utilizado para introducir estos cambios en las entidades, con sentencias condicionales que asignan el atributo en función del área y de un valor aleatorio. La variable V_O es el valor original y CODE es el nuevo valor asignado:

SI(Área>1000)	SINOSI (aleatorio <0,85)
CODE= V_O	CODE=153
SINO	SINOSI (aleatorio <0,90)
aleatorio=RANDOM	CODE=217
SI(V_O=6)	SINOSI (aleatorio <0,95)
SI(aleatorio <0,82)	CODE=236
CODE=6	SINO CODE=255
SINOSI(aleatorio <0,84)	FINSI
CODE=96	SINOSI(V_O=217)
SINOSI (aleatorio <0,86)	SI(aleatorio <0,70)
CODE=153	CODE=185
SINOSI (aleatorio <0,91)	SINOSI(aleatorio <0,75)
CODE=185	CODE=6
SINOSI (aleatorio <0,96)	SINOSI (aleatorio <0,80)
CODE=217	CODE=96
SINOSI (aleatorio <0,98)	SINOSI (aleatorio <0,85)
CODE=236	CODE=153
SINO CODE=255	SINOSI (aleatorio <0,90)
FINSI	CODE=185
SINOSI(V_O=96)	SINOSI (aleatorio <0,95)
SI(aleatorio <0,82)	CODE=236
CODE=96	SINO CODE=255
SINOSI(aleatorio <0,84)	FINSI
CODE=6	SINOSI(V_O=236)
SINOSI (aleatorio <0,86)	SI(aleatorio <0,82)
CODE=153	CODE=236
SINOSI (aleatorio <0,91)	SINOSI(aleatorio <0,84)
CODE=185	CODE=6
SINOSI (aleatorio <0,96)	SINOSI (aleatorio <0,86)
CODE=217	CODE=96
SINOSI (aleatorio <0,98)	SINOSI (aleatorio <0,91)
CODE=236	CODE=185
SINO CODE=255	SINOSI (aleatorio <0,96)
FINSI	CODE=217
SINOSI(V_O=153)	SINOSI (aleatorio <0,98)
SI(aleatorio <0,82)	CODE=153
CODE=153	SINO CODE=255
SINOSI(aleatorio <0,84)	FINSI
CODE=6	SINOSI(V_O=255)
SINOSI (aleatorio <0,86)	SI(aleatorio <0,82)
CODE=96	CODE=255
SINOSI (aleatorio <0,91)	SINOSI(aleatorio <0,84)
CODE=185	CODE=6
SINOSI (aleatorio <0,96)	SINOSI (aleatorio <0,86)
CODE=217	CODE=96
SINOSI (aleatorio <0,98)	SINOSI (aleatorio <0,91)
CODE=236	CODE=185
SINO CODE=255	SINOSI (aleatorio <0,96)
FINSI	CODE=217
SINOSI(V_O=185)	SINOSI (aleatorio <0,98)
SI(aleatorio <0,70)	CODE=236
CODE=185	SINO CODE=153
SINOSI(aleatorio <0,75)	FINSI
CODE=6	FINSI
SINOSI (aleatorio <0,80)	FINSI
CODE=96	

La tabla 6 muestra la hoja de cálculo en la que se ha programado el algoritmo de asignación. En ella se obtiene el archivo que asigna los nuevos valores a los polígonos del mapa tipo de suelo. Las dos primeras columnas de la tabla 6 se emplean para crear este fichero y luego se aplica en IDRISI con el comando **assign**. Cada polígono debe tener un identificador diferente para poder cambiar su código con el fichero de asignación.

=SI(Área>1000; V_O; SI(V_O= 6;SI(aleatorio <0,82;6;SI(aleatorio <0,84;96;SI(aleatorio <0,86;153;SI(aleatorio <0,91;185;SI(aleatorio <0,96;217;SI(aleatorio <0,98;236;255)); etc...

ID	CODE	V_O	Tipo	Clase	Texto	aleatorio	Área
104	6	6	LV15	LV15	Luvisoles	0,45	1598
126	6	6	LV13	LV13	Luvisoles	0,92	1191
127	6	6	LV3	LV3	Luvisoles	0,03	6707
129	6	6	LV12	LV12	Luvisoles	0,18	3110
138	6	6	LV8	LV8	Luvisoles	1,00	4498
150	6	6	LV17	LV17	Luvisoles	0,96	3822
162	6	6	LV15_1	LV15	Luvisoles	0,48	2640
176	6	6	LV11	LV11	Luvisoles	0,05	3191
178	6	6	LV2_1	LV2	Luvisoles	0,65	3895
185	6	6	LV17_1	LV17	Luvisoles	0,41	880
189	217	6	LV2_2	LV2	Luvisoles	0,91	464
190	6	6	LV2	LV2	Luvisoles	0,11	116
194	6	6	LV1	LV1	Luvisoles	0,82	3741
.
.
.

Tabla 6. Distorsión temática para los polígonos del mapa tipo de suelos.

En la tabla 6 se puede comprobar que el polígono número 189 cambia de Luvisol, con valor de adecuación 6, a Antrohosol, con adecuación 217. Este cambio está dirigido tal y como se indica en la matriz de confusión. El cambio no está restringido porque el valor aleatorio es mayor a 82%.

El paso final es crear un fichero de asignación e introducirlo mediante el comando **assign** en IDRISI. El fichero de asignación contiene dos columnas, la de código de identidad de cada elemento y la de valor temático.

La figura 20 muestra el proceso para la obtención de este factor. Esta figura se inicia con el comando **shapeidr**, que importa el mapa vectorial “tipo de suelo” con formato **shape**. El comando **polyras** convierte a raster este mapa y con **assign** asignamos el fichero de códigos distorsionados. Finalmente sólo queda enmascarar el mapa con el contorno de la comunidad de Madrid multiplicando mediante **overlay**.

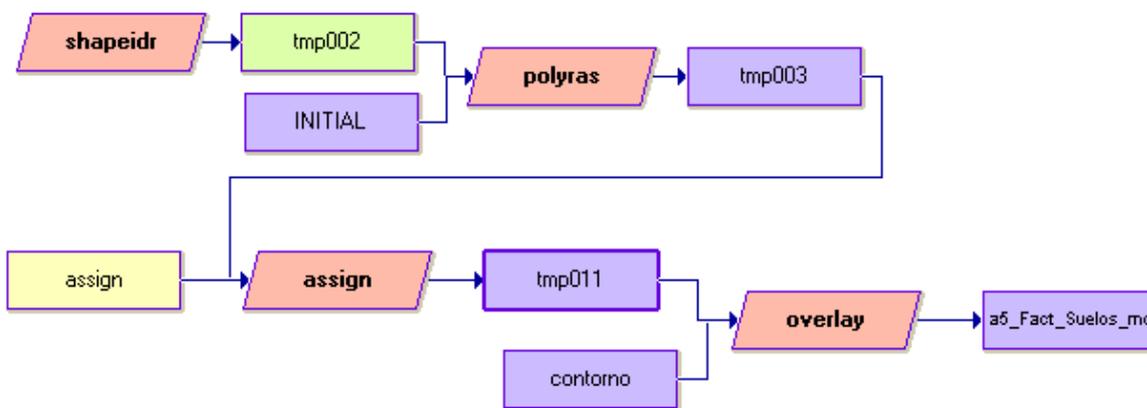


Figura 20. Reproducción del factor tipo de suelos con distorsión temática.

La figura 21 muestra el paso del mapa de tipo de suelo al factor original de aptitud desarrollado en la figura 20 (factor sin distorsionar).

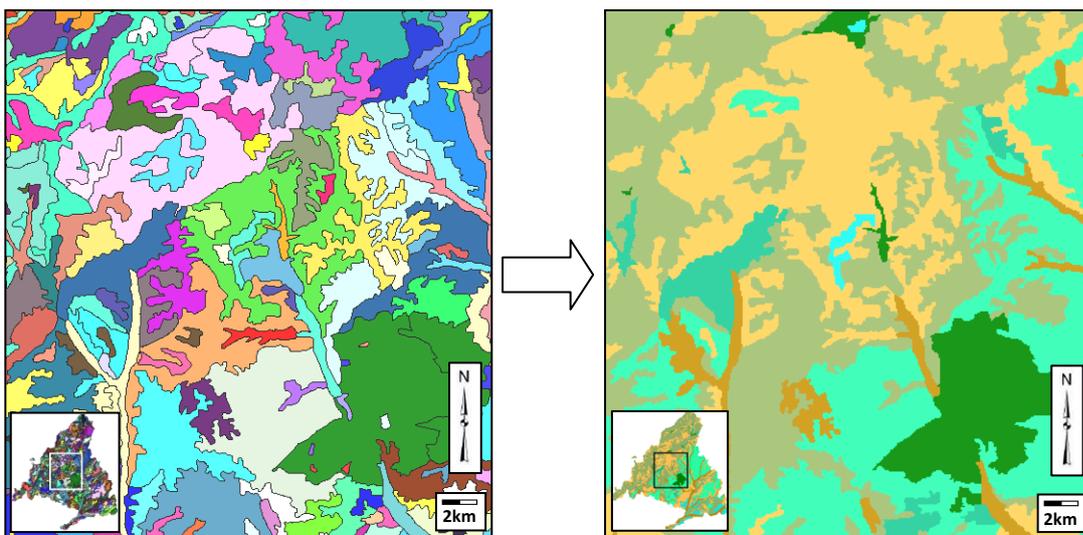


Figura 21. Mapa de tipo de suelos (a) y su valor de aptitud sin distorsionar (b).

La Figura 22 muestra el resultado final de la distorsión explicada. Estos dos mapas provienen del mapa vectorial de la Figura 21 (a) y se puede observar que pequeños polígonos toman diferentes valores respecto del original mostrado en la Figura 21 (b).

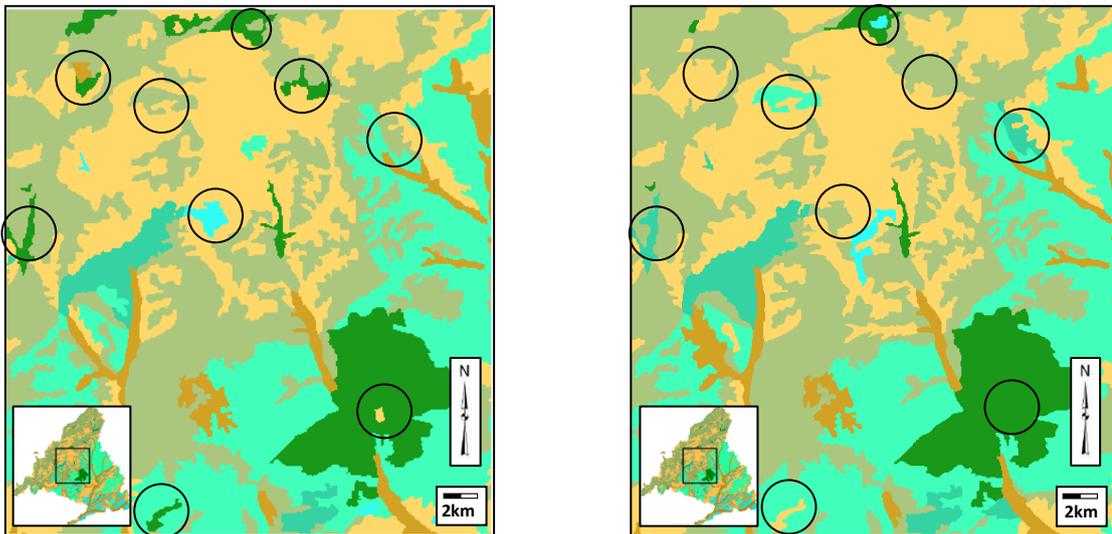


Figura 22. Mapas de aptitud según tipo de suelos tras la distorsión temática. Los círculos indican algunos de los cambios.

La figura 23 muestra un máximo en 16.000 píxeles para una diferencia del valor de adecuación de 69 entre el factor tipo de suelo modificado y sin modificar. Sabiendo que la imagen entera es de 7 millones de píxeles, este máximo no sobrepasa ni el 3‰ de todos los píxeles de la imagen.

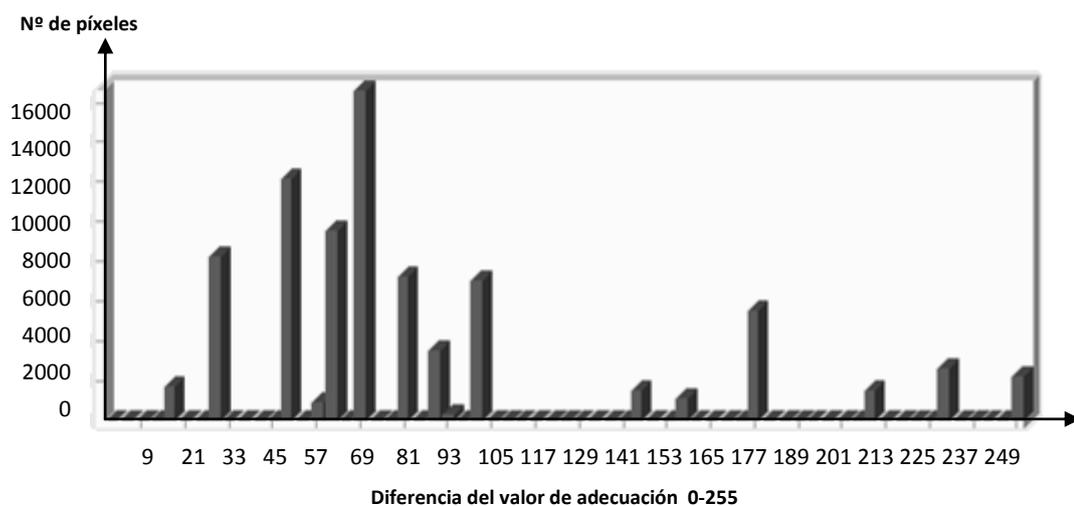


Figura 23. Valores positivos del histograma que muestran la diferencia del valor de adecuación entre el factor tipo de suelo modificado y sin modificar.

2.2.3.2.2. DISTORSION DEL FACTOR USOS DEL SUELO

Este factor considera la protección de las zonas agrícolas, los pastizales y otros usos del suelo con un valor ambiental y económico. Empleando los datos del censo agrario a nivel de unidades territoriales, la cobertura CORINE se subdivide en cuatro zonas para diferenciar las zonas de mayor productividad agraria. De esta forma, estos usos tendrán un valor de adecuación diferente en función de la zona de la región en la que se encuentren (ver zonas Z1, Z2, Z3 y Z4 en la cabecera de la tabla 7). Por ejemplo, una zona agrícola tendrá mayor valor en la parte sur de la región que en la zona norte, donde la productividad y la presencia de este uso son menos significativas.

La matriz de confusión de la tabla 7 está obtenida a partir del estudio “Metodología aplicada a la detección de errores temáticos en el CORINE Land Cover a través del estudio de cambios: Comunidad de Madrid (2000-2006)”, Barreira (2011).

Z1	0	238	0	0	255	68	68	0	119	0
Z2	0	243	0	0	255	58	116	232	255	0
Z3	0	242	128	128	255	51	115	0	153	0
Z4	0	240	0	0	255	30	0	120	255	0
	Tejido Urbano Continuo									
	Tejido Urbano Discontinuo									
	Industriales o comerciales									
	Redes viarias y asociados									
	Aeropuertos									
	Zonas de extracción minera									
	Escombreras y vertederos									
	Zonas en construcción									
	Zonas verdes urbanas									
	Inst deportivas y recreativas									
	Tierras de labor en secano									
	regados permanentemente									
	Vitales									
	Frutales									
	Olivares									
	Praderas									
	Mosaico de cultivos									
	agrícolas con vegetación natural									
	Sistemas agroforestales									
	Bosques de frondosas									
	Bosque de coníferas									
	Bosque mixto									
	Pantizales naturales									
	Vegetación Esclerofila									
	Matorral boscoso de transición									
	Playas dunas y arenales									
	Roquedo									
	Espacios con vegetación escasa									
	Humedales y zonas pantanosas									
	Cursos de agua									
	Laminas de agua									
111	80									
112	50									
121		10								
122		10								
124										
131										
132										
133										
141										
142										
211										
212										
221										
222										
223										
231										
242										
243										
244										
311										
312										
313										
321										
323										
324										
332										
333										
334										
411										
511										
512										

Tabla 7. Matriz de confusión a utilizar en la distorsión del mapa de usos del suelo. Valores en %.

El estudio de Barreira (2011) aplica la matriz de confusión para la determinación de la calidad temática en el *CORINE Land Cover*. De entre los estudios realizados, para verificar la fiabilidad de la clasificación de los polígonos, Barreira (2011) toma como verdad terreno lo que se entiende como la asignación temática y como clasificación, la realizada por el mapa Corine para el año 2000. A continuación obtiene una matriz para ver cómo de buenas son las clasificaciones realizadas. En esta matriz, las filas reflejan la cartografía y las columnas lo que debería haberse cartografiado.

Al aplicar el modelo, las clases se agrupan. Pero no todas las clases de un mismo grupo tienen la misma probabilidad de cambio entre ellas. Además, la distorsión realiza cambios entre los nueve grupos y como consecuencia se obtiene el mismo valor de adecuación para cambiar a suelo urbano discontinuo o a urbano continuo (en los dos casos se adquiere el mismo valor).

El siguiente algoritmo aplica la matriz de confusión a los atributos de los polígonos. En la variable PRECODE se asignan los nuevos valores y C_O es el código original. Para diferenciar las zonas se ha modificado el identificador (ID) de cada identidad en función de la zona a la que pertenezca.

```

aleatorio = RANDOM
SI(O(C_O = 411; C_O = 511; C_O = 512))
    PRECODE = 0
FINSI
SI( ID<100000) //polígonos zona 1
    SI(C_O =111)
        SI(aleatorio<0.8)
            PRECODE=0
        SINOSI(aleatorio <0.9)
            PRECODE=68
        SINO
            PRECODE=119
        FINSI
    SINOSI(C_O =112)
        SI(aleatorio<0.7)
            PRECODE=0
        SINOSI(aleatorio <0.8)
            PRECODE=238
        SINOSI(aleatorio<0.9)
            PRECODE=68
        SINO
            PRECODE=119
        FINSI
    SINOSI(C_O =121)
        SI(aleatorio<0.9)
            PRECODE=0
        SINO
            PRECODE=238
        FINSI
    SINOSI(O(C_O =122; C_O =124; C_O =131;
        C_O =132))
        PRECODE=0
    SINOSI(C_O =133)
        SI(aleatorio<0.6)
            PRECODE=0
        SINOSI(aleatorio<0.8)
            PRECODE=238
        FINSI
        SINOSI(C_O =141)
            SI(aleatorio<0.8)
                PRECODE=0
            SINO
                PRECODE=68
            FINSI
        SINOSI(C_O =142)
            SI(aleatorio<0.9)
                PRECODE=0
            SINO
                PRECODE=68
            FINSI
        SINOSI(C_O =211)
            SI(aleatorio<0.6)
                PRECODE=238
            SINO
                PRECODE=0
            FINSI
        SINOSI(C_O =212)
            SI(aleatorio<0.6)
                PRECODE=0
            SINOSI(aleatorio<0.8)
                PRECODE=68
            SINO
                PRECODE=119
            FINSI
        SINOSI(O(C_O =221; C_O =222;
            C_O =223))
            SI(aleatorio<0.9)
                PRECODE=0
            SINO
                PRECODE=68
            FINSI

```

```

FINSI
SINOSI( C_O =231)
SI(aleatorio <0.6)
  PRECODE=226
SINOSI(aleatorio <0.7)
  PRECODE=68
SINO
  PRECODE=0
FINSI
SINOSI( C_O =242)
SI(aleatorio <0.7)
  PRECODE=68
SINOSI(aleatorio <0.8)
  PRECODE=238
SINO
  PRECODE=0
FINSI
SINOSI( C_O =243)
SI(aleatorio <0.7)
  PRECODE=68
SINOSI(aleatorio <0.8)
  PRECODE=238
SINOSI(aleatorio <0.9)
  PRECODE=119
SINO
  PRECODE=0
FINSI
SINOSI( C_O =244)
SI(aleatorio <0.7)
  PRECODE=68
SINO
  PRECODE=119
FINSI
SINOSI( C_O =311)
SI(aleatorio <0.6)
  PRECODE=68
SINOSI(aleatorio <0.8)
  PRECODE=119
SINO
  PRECODE=0
FINSI
SINOSI( O(C_O =312; C_O =313))
SI(aleatorio <0.8)
  PRECODE=68
SINO
  PRECODE=119
FINSI
SINOSI( C_O =321)
SI(aleatorio <0.7)
  PRECODE=0
SINOSI(aleatorio <0.8)
  PRECODE=255
SINO
  PRECODE=238
FINSI
SINOSI( C_O =323)
SI(aleatorio <0.6)
  PRECODE=119
SINOSI(aleatorio <0.8)
  PRECODE=238
SINO
  PRECODE=0
FINSI
SINOSI( C_O =324)
SI(aleatorio <0.7)
  PRECODE=119
SINOSI(aleatorio <0.8)
  PRECODE=68
SINO
  PRECODE=0
FINSI
SINOSI(O( C_O =332; C_O =333;
  C_O =334)
  PRECODE=0
FINSI
FINSI
SINOSI( ID<200000) //polígonos zona 2
// SE EMPLEA EL MISMO CODIGO QUE LA
ZONA 1 PERO CON LOS VALORES DE LA
ZONA 2
SINOSI( ID<300000) //polígonos zona 3
// SE EMPLEA EL MISMO CODIGO QUE LA
ZONA 1 PERO CON LOS VALORES DE LA
ZONA 3
SINOSI( ID<400000) //polígonos zona 3
// SE EMPLEA EL MISMO CODIGO QUE LA
ZONA 1 PERO CON LOS VALORES DE LA
ZONA 4
FINSI

```

Finalmente, el algoritmo de restricción de área aplica el código distorsionado. En este caso se realizan cambios para polígonos menores a 1.000 hectáreas. Este límite corresponde a evitar cambios en un 4% de todos los polígonos (52% del área de la Comunidad de Madrid).

```

SI(área>1000)
  CODE= V_O
SINO
  CODE= PRECODE
FINSI

```

En la tabla 8 se aplica en una hoja de cálculo el algoritmo diseñado. Aplicando la restricción de área de 1.000 hectáreas, podrán cambiar el 96 % de los polígonos, valor que corresponde a un 45 % del área total. En definitiva, tras añadir la distorsión dirigida por matriz de confusión, este valor queda reducido al 15 % del área total.

```
=SI(área>1000; V_O; PRECODE)
```

```
=SI(C_O = 411; 0; SI(C_O =111; SI(Numero aleatorio <0.8; 0; SI(Numero aleatorio <0.9; 68; 119)); etc...
```

ID	CODE	PRECODE	V_O	C_O	Texto	área	aleatorio
1	0	0	0	411	Humedales y zonas pantanosas	60	0,219
2	119	119	0	111	Tejido Urbano Continuo	197	0,913
3	119	119	0	111	Tejido Urbano Continuo	20	0,903
4	0	0	0	111	Tejido Urbano Continuo	39	0,437
5	0	0	0	111	Tejido Urbano Continuo	48	0,464
6	68	68	0	111	Tejido Urbano Continuo	7	0,862
7	0	0	0	111	Tejido Urbano Continuo	44	0,768
8	0	0	0	111	Tejido Urbano Continuo	20	0,579
9	68	68	0	111	Tejido Urbano Continuo	24	0,850
10	0	0	0	111	Tejido Urbano Continuo	51	0,578
11	0	0	0	111	Tejido Urbano Continuo	6	0,632
.
.
.

Tabla 8. Distorsión para los once primeros polígonos de la Z1 del mapa de usos del suelo.

Las dos primeras columnas de la tabla 8 se emplean en el fichero de asignación que cambia los códigos de los polígonos.

El procedimiento del diagrama de la figura 24 muestra el proceso para la obtención del factor de usos del suelo. Éste se inicia con el comando *shapeidr*, que importa el mapa vectorial en formato *shape*. Acto seguido, la capa vectorial se convierte a raster con el comando *polyras*. A continuación se superpone el mapa de zonificación de mayor productividad agraria mediante *overlay* y se asigna un valor a cada polígono en función de la zona en la que esté.

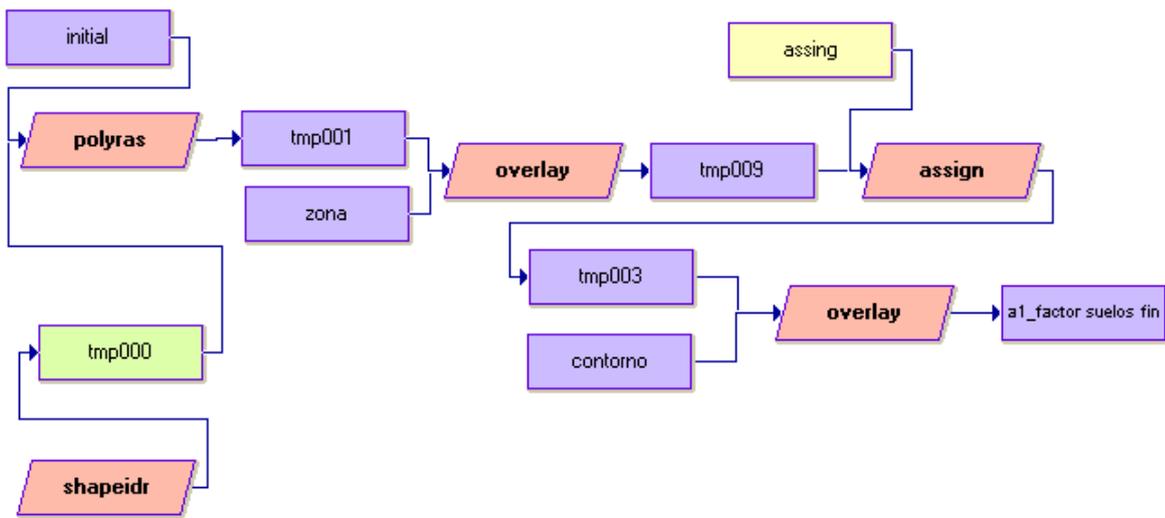


Figura 24. Reproducción del factor usos del suelo con distorsión temática.

El histograma que aparece en la figura 25 muestra la diferencia del valor de adecuación entre el factor usos de suelo modificado y sin modificar temáticamente. Recordando que la imagen es de 7 millones de píxeles, el pico de 30.000 píxeles no supera el 5 % del total.

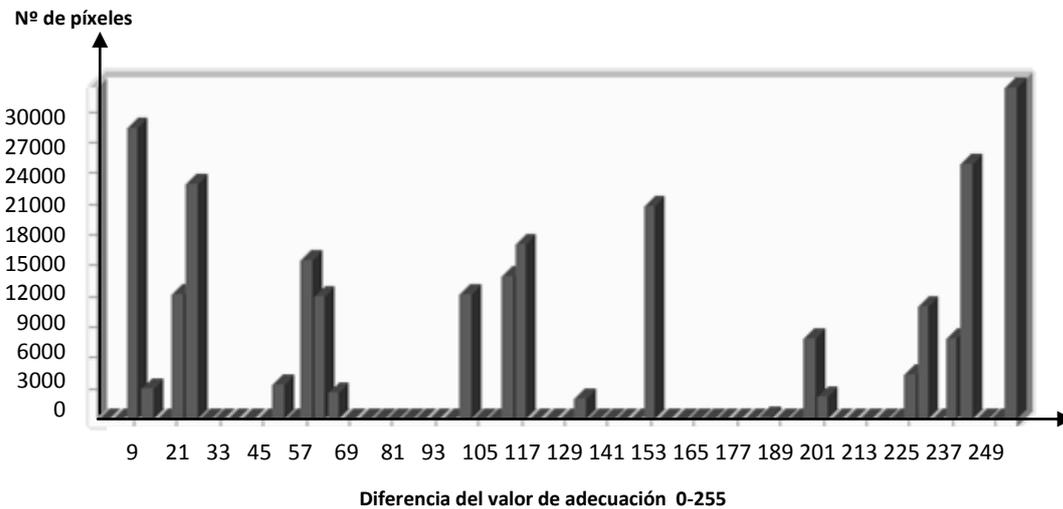


Figura 25. Histograma de la diferencia del valor de adecuación entre el factor usos de suelo modificado y sin modificar.

La figura 26 muestra en binario, el resultado de los polígonos que han cambiado tras realizar la distorsión temática dirigida. Es el resultado de restar el mapa distorsionado temáticamente con el mapa original.

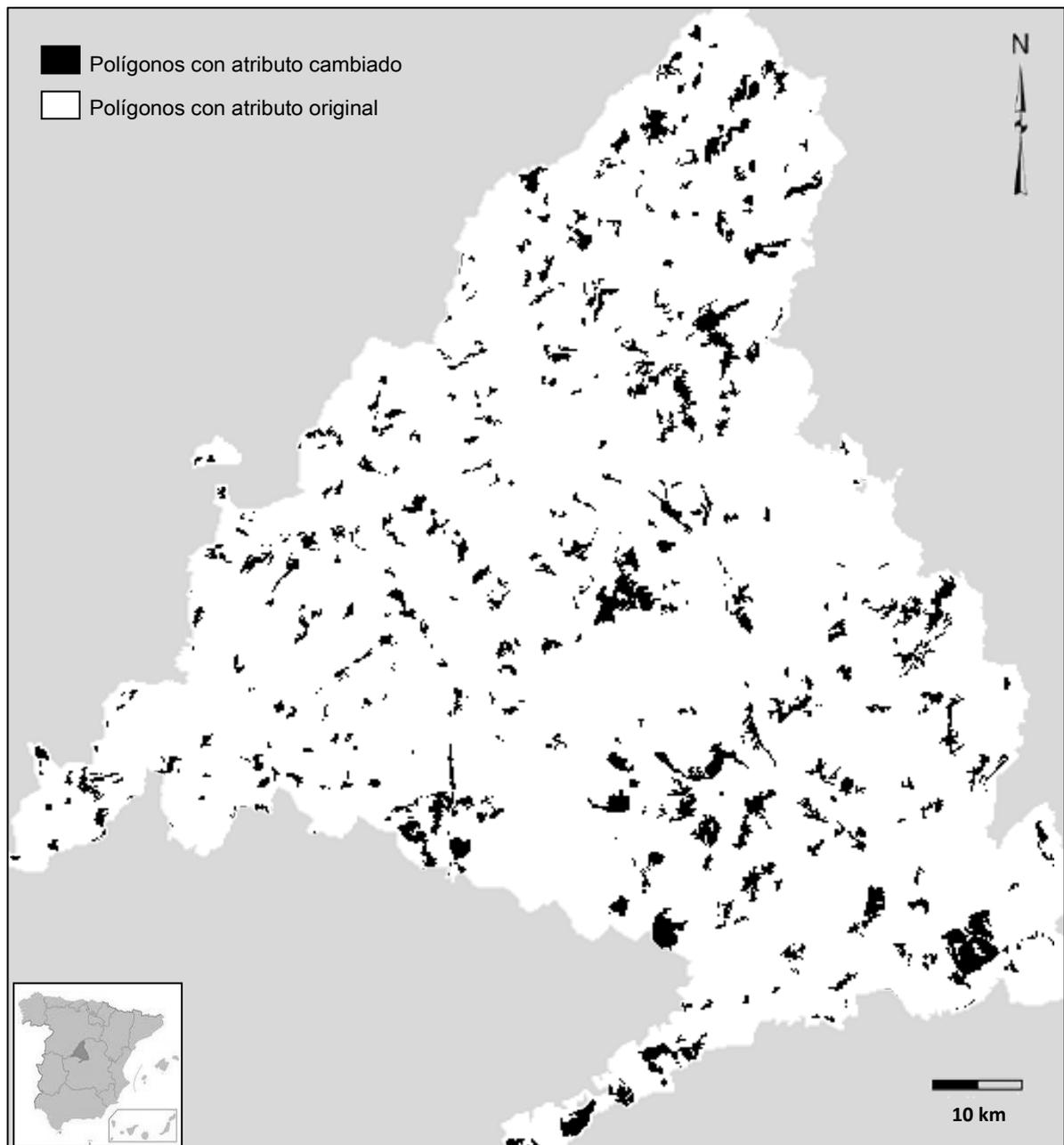


Figura 26. Mapa de los polígonos que han cambiado en la distorsión temática del mapa de usos del suelo.

2.2.3.2.3. DISTORSIÓN DEL FACTOR DISTANCIA A CARRETERAS

Este factor se incluye en el modelo objeto de estudio con la finalidad de que las nuevas zonas urbanas tengan buena accesibilidad.

En el modelo, las carreteras se agrupan por tipos de la siguiente forma:

- Tipo 1 Autovías Nacionales
- Tipo 2 Carreteras Nacionales
- Tipo 3 Carretera Principal-Doble Calzada
- Tipo 4 Carreteras Principales
- Tipo 5 Carretera Secundaria M-407
- Tipo 6 Carreteras Secundarias
- Tipo 7 Carreteras Locales
- Tipo 8 Caminos
- Tipo 9 Travesías
- Tipo 10 Calles, Urbanizaciones

Para obtener el factor distancia a carreteras, el modelo elimina las carreteras menos importantes (tipo 9 y 10). La distorsión temática intercambiará los identificadores de las carreteras más probables de confundir y las carreteras del tipo 1 a 6, como no tienen probabilidad de cambiar al tipo 9 o 10, no se eliminarán. En cambio, las del tipo 7 y 8 podrán cambiar a tipo 9 o 10 con su consecuente eliminación.

En la tabla 9, la matriz de confusión elaborada posibilita que el 20% de carreteras del tipo 7 y el 80% del tipo 8 se eliminen. Los tipos 9 y 10 se eliminan según el modelado original.

	tipo 8	tipo 7	tipo 1 a 6	resto tipos	nulo
tipo 8	20	0	0	0	80
tipo 7	0	80	0	0	20
tipo 1 a 6	0	0	100	0	0
resto tipos	0	0	0	0	100
nulo	0	0	0	0	100

Tabla 9. Matriz de confusión para la distorsión del factor carreteras. Valores en %.

El algoritmo empleado para la creación del archivo de asignación es el siguiente:

```

SI(Tipo=7)
  SI(ALEATORIO())<0.2)
    CODIGO=0
  SINO
    CODIGO=1
  FINSI
FINSI
SI(Tipo =8)
  SI(ALEATORIO())<0.2
    CODIGO=1
  SINO
    CODIGO=0
  FINSI
FINSI
SI(O(Tipo =1; Tipo =2; Tipo =3; Tipo =4; Tipo =5; Tipo =6)
  CODIGO=1
SINO
  CODIGO=0
FINSI
    
```

La aplicación del algoritmo en una hoja de cálculo se muestra en la tabla 10. El valor que se emplea en la variable CODIGO ha eliminado un 20 % de polilíneas de tipo 7 y un 80 % de tipo 8, tal y como se indica en la matriz de confusión. Las carreteras de tipo 1 a 6 se mantienen y las de tipo 9 y 10 se eliminan.

=SI(Tipo=7;SI(ALEATORIO())<0.2;0;1);SI(Tipo =8;SI(ALEATORIO())<0.2;1;0);SI(O(Tipo =1; Tipo =2; Tipo =3; Tipo =4; Tipo =5; Tipo =6);1;0)))

ID polilínea	CODIGO	Tipo	LENGTH	NOMBRE	CARRET_ID
1	1	2	260.717	N-601	1
2	1	1	635.236	N-1	2
3	1	2	212.819	N-110	3
4	1	2	4428.566	N-110	4
5	1	2	4149.501	N-601	5
6	1	1	3549.621	N-1	6
7	1	2	4884.242	N-110	7
8	1	8	2717.064	camino	8
9	0	8	809.916	camino	25
10	0	8	1631.406	camino	9
11	0	8	133.271	camino	10
.
.

Tabla 10. Distorsión temática para las polilíneas del factor distancia a carreteras.

Como en los anteriores procedimientos, el diagrama de la Figura 27 comienza importando el archivo *shape* con el comando **shapeidr**. Acto seguido se rasteriza con **lineras** y cada identificador se modifica con el comando **assign**. El procedimiento **distance** calcula las distancias desde cada píxel a la entidad más cercana.

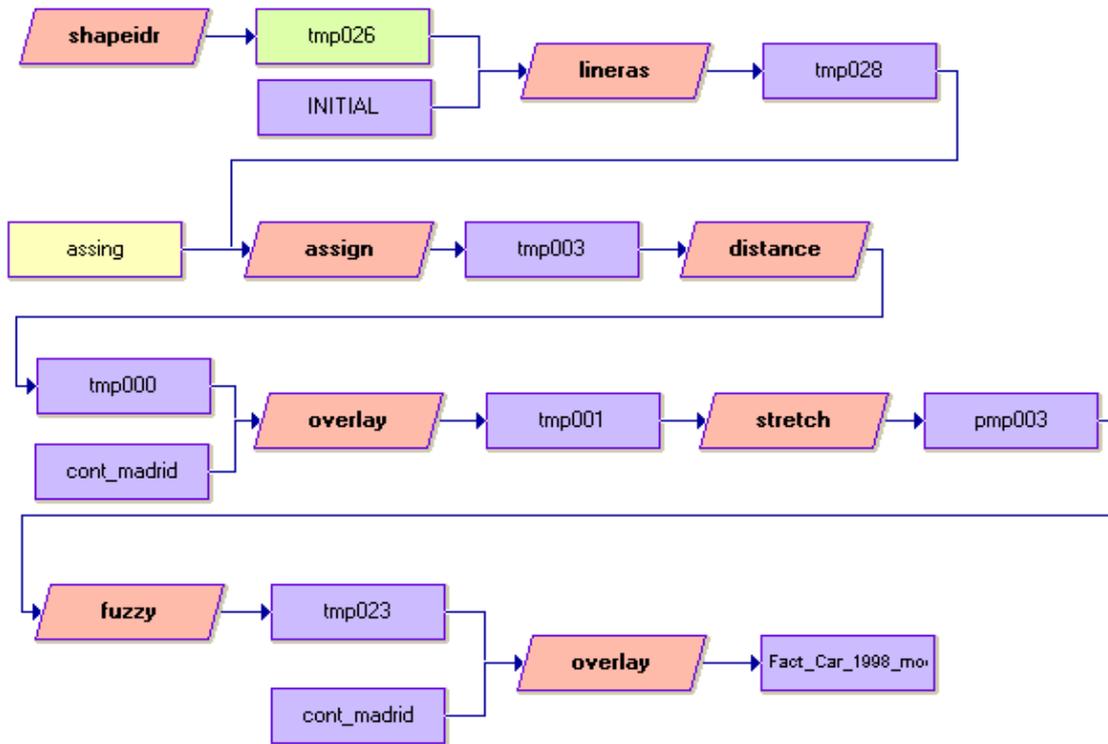


Figura 27. Reproducción del factor de distancia a carreteras.

El histograma de la figura 28 muestra la diferencia del valor de adecuación entre el factor modificado y sin modificar con un máximo en 30.000 píxeles para valores cercanos a 0. Sabiendo que la imagen entera es de 7 millones de píxeles, este máximo no sobrepasa ni el 5‰ de todos los píxeles de la imagen.

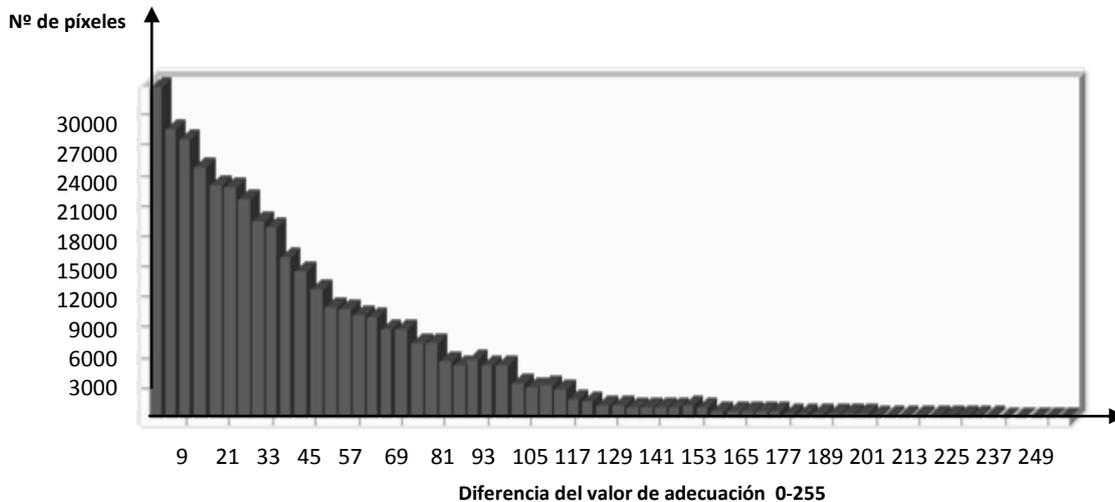


Figura 28. Histograma de diferencia de adecuación entre el factor modificado y sin modificar.

La figura 29 muestra la diferencia entre el mapa deformado y sin deformar del factor distancias a carreteras. Esta distorsión temática refleja variaciones locales producidas por la eliminación de carreteras.

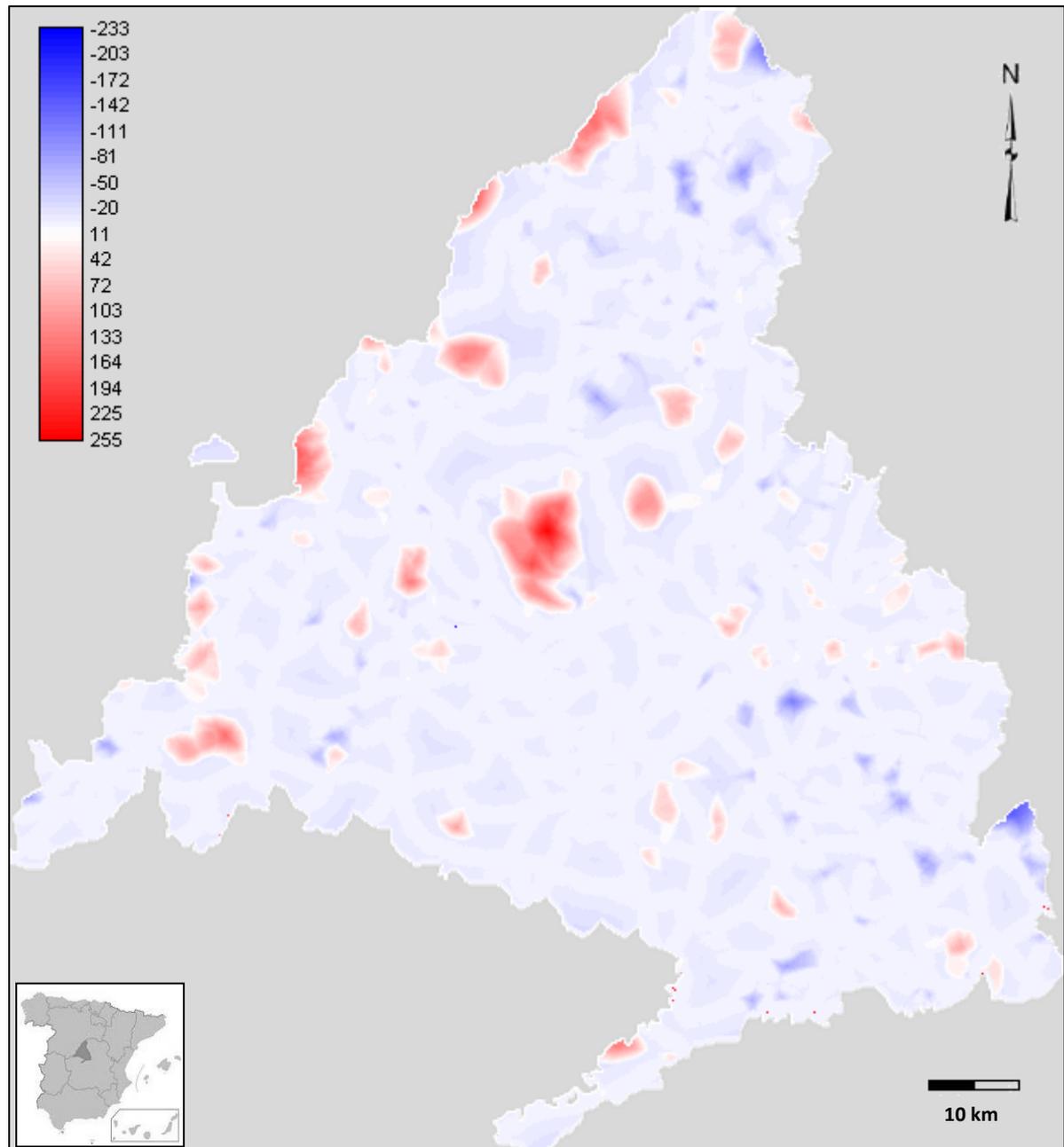


Figura 29. Diferencia del valor de adecuación entre el del factor distancia a carreteras original y el distorsionado temáticamente.

2.2.3.2.4 DISTORSIÓN DEL FACTOR GEOTECNIA

Este factor representa el riesgo de fallas activas, sísmico o sitios de baja estabilidad geomorfológica. Para el diseño de este factor se aprovechó la información que aporta el mapa de geotecnia de la Comunidad de Madrid sobre problemas de compactación, cimentación y fisuración de las rocas, poniendo especial atención en todas aquellas zonas que tuviesen alguno de los siguientes problemas o una combinación de varios de ellos:

- riesgo de deslizamiento,
- oquedades subterráneas y
- nivel freático a escasa profundidad.

Para reducir la complejidad del mapa geotécnico en el modelo original, las entidades se agruparon en función de los problemas anteriores, otorgándoles un peso a cada uno de ellos como aparece a continuación (entre paréntesis se indica el valor de adecuación final, una vez normalizada la variable entre 0 y 255):

- Alto riesgo (5): Combinación de los tres problemas
 Elevado riesgo (14): Combinación de dos problemas
 Bajo riesgo (186): Contiene uno de los problemas
 Nulo riesgo (255): No tiene ninguno de los problemas

La distorsión temática se efectúa en función de estos grupos y se propone dar una probabilidad de cambio según la tabla 11.

	alto riesgo	elevado riesgo	bajo riesgo	nulo riesgo	
alto riesgo	85	15	0	0	100
elevado riesgo	15	85	0	0	100
bajo riesgo	0	0	85	15	100
nulo riesgo	0	0	15	85	100
	100	100	100	100	

Tabla 11. Matriz de confusión empleada para distorsionar el factor geotecnia. Valores en %.

Los nuevos identificadores de los polígonos cambiarán en función de la matriz de confusión de la tabla 11 y del área de cada polígono. Como en la distorsión de los usos del suelo (sección 2.2.3.2.2.), para polígonos mayores de 10.000 hectáreas no se efectúa distorsión. En este caso los polígonos son de mayor tamaño y resulta un valor de 22 % para los polígonos que no se van a distorsionar (93% del área de la Comunidad de Madrid).

El siguiente algoritmo condicional muestra la introducción de distorsión temática en cada entidad. Este algoritmo comienza restringiendo las identidades inferiores al área indicada. Luego guarda el valor aleatorio en la variable ALEATORIO. A continuación se ejecutan una serie de condiciones para comprobar el identificador de la entidad. Posteriormente se cambia el valor de la variable CODIGO en función del valor aleatorio obtenido. Este algoritmo de condición se debe aplicar a cada entidad, y para ello, se puede emplear un algoritmo que modifique todos los identificadores (se puede emplear un bucle del tipo *for* o *while*).

```

SI(área<10000000)
ALEATORIO=RANDOM
  SI(Valor original=5)
    SI(ALEATORIO())<0.15)
      CODIGO=14
    SINO
      CODIGO=Valor original
  FINSI
  SI(valor original=14)
    SI(ALEATORIO())<0.15)
      CODIGO=5
    SINO
      CODIGO=Valor original
  FINSI
  SI(valor original=186)
    SI(ALEATORIO())<0.15)
      CODIGO=255
    SINO
      CODIGO=Valor original
  FINSI
  SI(valor original=255)
    SI(ALEATORIO())<0.15)
      CODIGO=186
    SINO
      CODIGO=Valor original
  FINSI
  SINO
    CODIGO="error"
  SINO
    CODIGO=Valor original
  FINSI

```

La tabla 12 muestra la hoja de cálculo en la que se ha programado el algoritmo de introducción de error. Con ella se obtiene el archivo que asigna los nuevos valores a los polígonos del mapa original de geotecnia. Para ello, cada polígono debe tener un identificador diferente y luego cambiar su código con el fichero de asignación. Las dos primeras columnas de la tabla 12 se emplean para crea este fichero y luego se aplica al mapa con el comando *assign* de IDRISI.

```
=SI(área<10000000; SI(Valor original=5; SI(ALEATORIO())<0.15;14; Valor original);SI(valor original=14; SI(ALEATORIO())<0.15;5; Valor original); SI(valor original=186; SI(ALEATORIO())<0.15;255; Valor original); SI(valor original=255; SI(ALEATORIO())<0.15;186; Valor original );"error")); Valor original)
```

ID polígono	CODIGO	Valor original	área
1	186	186	1212484
2	186	186	814264
3	255	255	2561859
4	255	186	522037
5	186	186	1895923
6	186	186	635467
7	186	186	1448374
8	186	186	655946
9	186	186	607771
10	255	186	3156087
11	186	186	518970
12	186	186	251853
13	186	186	229404
.	.	.	.
.	.	.	.

Tabla 12. Aplicación del algoritmo de distorsión temática para los polígonos del factor geotecnia.

El diagrama de la figura 30 muestra el proceso para la obtención del factor geotécnico. Esta figura se inicia con el comando **shapeidr**, que importa el mapa vectorial geotécnico en formato **shape**. El comando **polyras** convierte a raster este mapa y con **assign** se asigna el fichero de códigos distorsionados. Finalmente sólo queda enmascarar el mapa con el contorno de la comunidad de Madrid multiplicando mediante **overlay**.

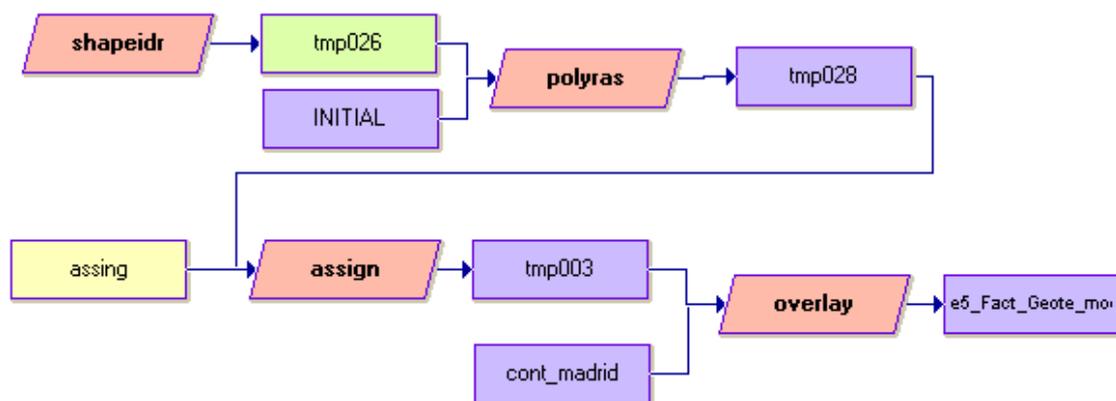


Figura 30. Reproducción del factor de geotecnia con distorsión.

El histograma de la figura 31 muestra un máximo de 18.000 píxeles para la diferencia de valor de adecuación de 69. Esta diferencia es el paso de bajo riesgo a nulo riesgo. Sabiendo que la imagen entera es de 7 millones de píxeles este máximo no sobrepasa ni el 3% de todos los píxeles de la imagen.

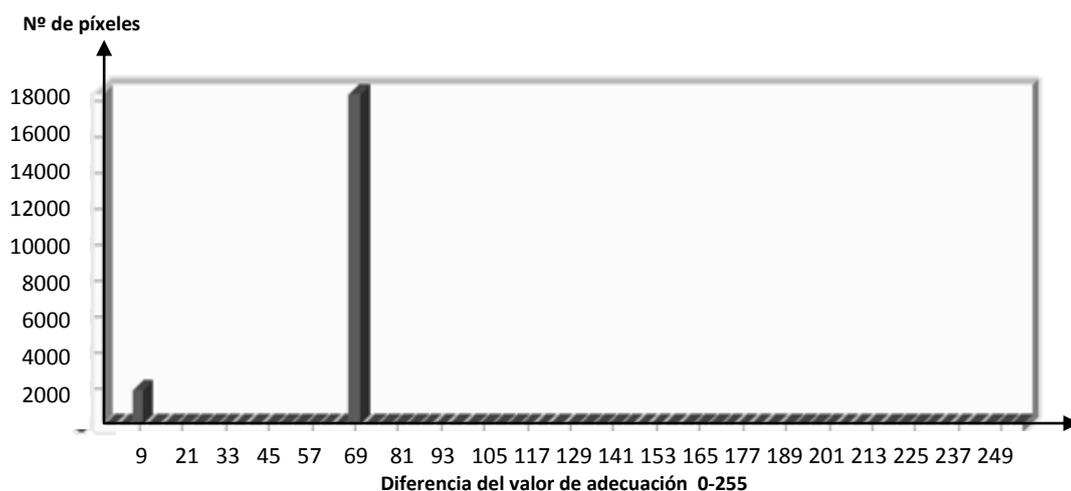


Figura 31. Histograma de la diferencia del valor de adecuación entre el factor de geotecnia modificado y sin modificar.

2.2.3.2.5. DISTORSIÓN DEL FACTOR DE DISTANCIA A ZONAS URBANAS

En este factor se pretende que las nuevas zonas urbanas estén cerca de otros núcleos de población de tamaño intermedio. Para ello se crean tres grupos en función de los habitantes de cada núcleo urbano y posteriormente se hace una suma lineal ponderada con los mapas de distancia obtenidos de dichos grupos.

En este caso, el error temático se realiza mediante varios algoritmos que introducen cambios contemplando el tamaño de la población. Este método cambia de clase las ciudades que están al borde de pertenecer a otra clase.

La siguiente fórmula modifica los valores de población para que posteriormente los núcleos urbanos cambien de grupo asignado:

$$NUEVO\ TOTAL = TOTAL + S * A * (-1)^{entero(100 * A)}$$

Donde,

- A** es un número aleatorio entre cero a uno y
- S** toma valores en función del tamaño de la población.

La variable **S** crea variaciones de un 5 % en las ciudades para provocar que los núcleos cambien de grupo.

Los posibles valores de la variable **S** son los siguientes:

- **S** = 10 000 si población < 10 000 habitantes
- **S** = 20 000 si 10 000 < población < 50 000 habitantes
- **S** = 40 000 para el resto.

El algoritmo condicional empleado para la creación del archivo de asignación comienza guardando valores en la variable CODIGO en función del TOTAL, que es la población de cada entidad:

```
SI(TOTAL<10000)
  CODIGO=1
SINOSI(TOTAL<50000)
  CODIGO=2
SINOSI(TOTAL <100000)
  CODIGO=3
SINOSI(TOTAL <200000)
  CODIGO=4
SINO
  CODIGO=5
FINSI
```

El algoritmo condicional empleado para la variable NUEVO refleja los nuevos valores asignados a cada núcleo urbano:

```

SI(CODIGO=1)
    NUEVO=TOTAL+10000*ALEATORIO()*(-1)^ENTERO(100*ALEATORIO())
SINOSI(CODIGO=2)
    NUEVO= TOTAL+20000*ALEATORIO()* (-1)^ENTERO(100*ALEATORIO())
SINO
    NUEVO=TOTAL +40000*ALEATORIO()*(-1)^ENTERO(100*ALEATORIO())
FINSI

```

Finalmente, el algoritmo condicional asigna los nuevos identificadores en la variable COD* :

```

SI(NUEVO<10000)
    COD1 =1
    COD2 =0
    COD3 =0
    COD4 =0
    COD5 =0
SINOSI(NUEVO <50000)
    COD1 =0
    COD2 =1
    COD3 =0
    COD4 =0
    COD5 =0
SINO SI(NUEVO<100000)
    COD1 =0
    COD2 =0
    COD3 =1
    COD4 =0
    COD5 =0
SINOSI(NUEVO<200000)
    COD1 =0
    COD2 =0
    COD3 =0
    COD4 =1
    COD5 =0
SINO
    COD1 =0
    COD2 =0
    COD3 =0
    COD4 =0
    COD5 =1
FINSI

```

La tabla 13 muestra la aplicación del algoritmo en una hoja de cálculo. En la primera columna encontramos el número de habitantes original y, en la última, el número de habitantes modificado (los valores negativos son resultados simbólicos). Estos nuevos valores pasan a la columna CÓDIGO, en función de los límites de población establecidos para cada grupo, tomando valores de uno a cinco. Los cinco pares de columnas corresponden a los cinco grupos donde cada núcleo se ha asignado binariamente en función de la columna CÓDIGO.

$$=SI(CODIGO=1;TOTAL+10000*ALEATORIO()*(-1)^ENTERO(100*ALEATORIO()));SI(CODIGO=2; TOTAL+20000*ALEATORIO()*(-1)^ENTERO(100*ALEATORIO()));TOTAL +40000*ALEATORIO()*(-1)^ENTERO(100*ALEATORIO()))$$

$$=SI(TOTAL<10000;1;SI(TOTAL <50000;2;SI(TOTAL <100000;3;SI(TOTAL <200000;4;5))))$$

$$=SI(NUEVO<10000;1;SI(NUEVO <50000;0; SI(NUEVO<100000;0;SI(NUEVO<200000;0;0))))$$

$$=+SI(NUEVO<10000;0;SI(NUEVO <50000;1;SI(NUEVO <100000;0;SI(NUEVO <200000;0;0))))$$

$$=SI(NUEVO <10000;0;SI(NUEVO <50000;0;SI(NUEVO <100000;1;SI(NUEVO <200000;0;0))))$$

$$=SI(NUEVO<10000;0;SI(NUEVO <50000;0;SI(NUEVO <100000;0;SI(NUEVO <200000;1;0))))$$

$$=SI(NUEVO <10000;0;SI(L2<50000;0;SI(NUEVO <100000;0;SI(NUEVO <200000;0;1))))$$

TOTAL	<POB <10000		<POB <50000		<POB <100000		<POB <200000		<POB >200000		CODIGO	NUEVO
	ID	COD1	ID	COD2	ID	COD3	ID	COD4	ID	COD5		
2306	1	1	1	0	1	0	1	0	1	0	1	3893
4573	2	1	2	0	2	0	2	0	2	0	1	5959
175237	3	0	3	0	3	0	3	1	3	0	4	163218
91964	4	0	4	0	4	1	4	0	4	0	3	93446
1209	5	1	5	0	5	0	5	0	5	0	1	1778
150489	6	0	6	0	6	0	6	1	6	0	4	139733
2193	7	0	7	0	7	0	7	0	7	0	1	-6845
2193	8	0	8	0	8	0	8	0	8	0	1	923
2193	9	1	9	0	9	0	9	0	9	0	1	4720
2193	10	1	10	0	10	0	10	0	10	0	1	1247
1307	11	0	11	1	11	0	11	0	11	0	2	10107
1176	12	1	12	0	12	0	12	0	12	0	1	8617
1176	13	1	13	0	13	0	13	0	13	0	1	4254
12994	14	1	14	0	14	0	14	0	14	0	1	5355
4862	15	1	15	0	15	0	15	0	15	0	1	2089
1422	16	0	16	1	16	0	16	0	16	0	2	11368
.
.
.

Tabla 13. Distorsión temática para los polígonos del factor de distancia a zonas urbanas.

El diagrama de la figura 32 crea diferentes mapas en función del número de habitantes. Empleando el mismo mapa inicial y el fichero de asignación de cada caso, se asignan valores binarios en función de la dimensión de las ciudades.

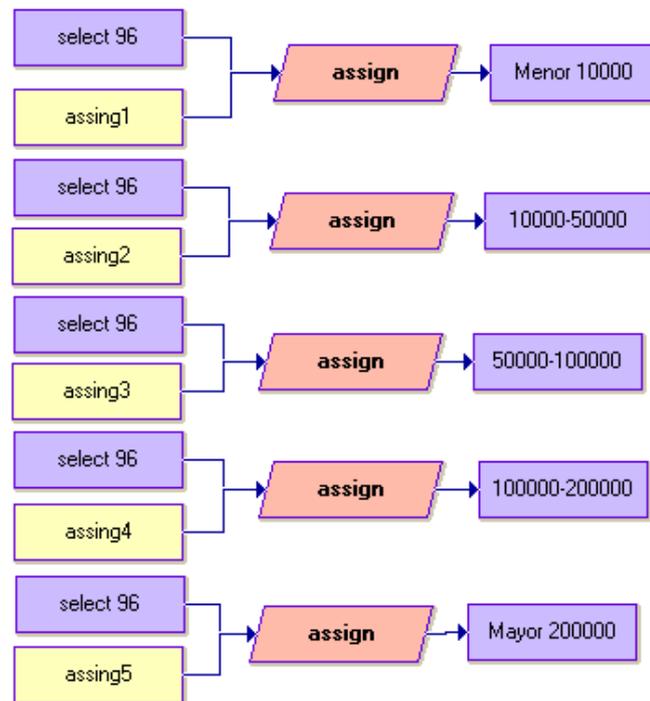


Figura 32. Asignación de nuevos identificadores en el mapa de población para crear grupos en función del número de habitantes.

En el diagrama de la figura 33 se muestra el proceso para obtener los mapas de distancia para cada grupo de ciudades. El procedimiento **distance** calcula las distancias desde cada píxel a la entidad más cercana.

El mapa de restricciones *tmp014* contiene todas las ciudades y se empleará más adelante, en el paso final, para obtener el factor distancia a zonas urbanas. Como *tmp014* depende de la distorsión temática, éste se obtiene tras realizar la asignación sumando los distintos mapas de ciudades con el comando **overlay** y luego se reclasifica a binario con **reclass**.

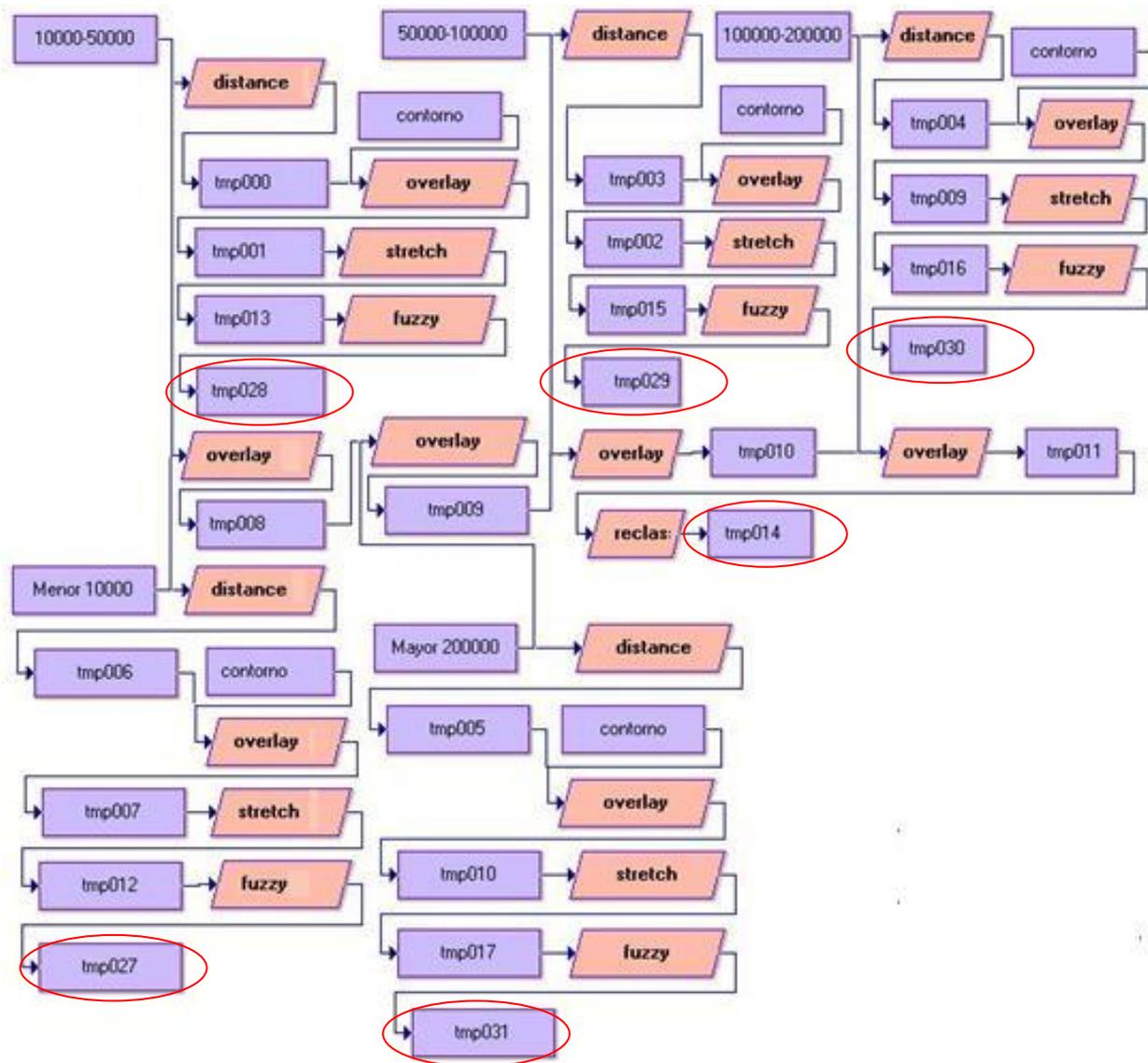


Figura 33. Proceso para realizarla distorsión temática del factor distancia zonas urbanas.

La última fase para la reproducción del factor distancias a zonas urbanas es la suma lineal ponderada de todos los mapas de distancias. En este último proceso se le aplica el mapa de restricciones *tmp014*. La figura 34 muestra las diferencias del mapa original respecto al mapa distorsionado temática y geoméricamente.

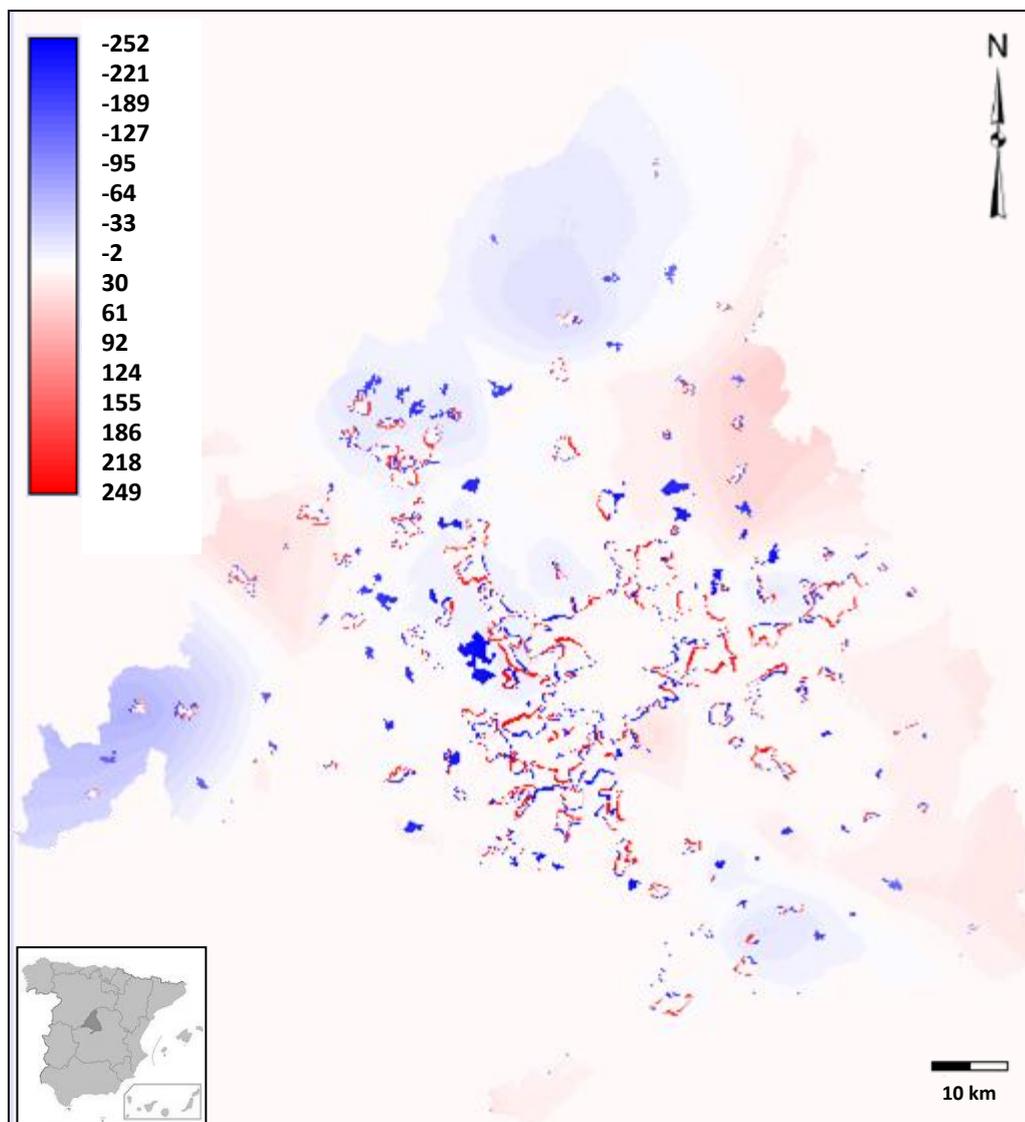


Figura 34. Diferencias del valor de adecuación final entre el factor original de distancia a zonas urbanas y el modificado temáticamente y geoméricamente.

El histograma de la figura 35 muestra las diferencias del valor de adecuación final entre el factor zonas urbanas modificado y sin modificar con un máximo de 450.000 píxeles (diferencia del valor de adecuación cercano a cero). Sabiendo que la imagen entera es de 7 millones de píxeles, este máximo no sobrepasa ni el 7% de todos los píxeles de la imagen.

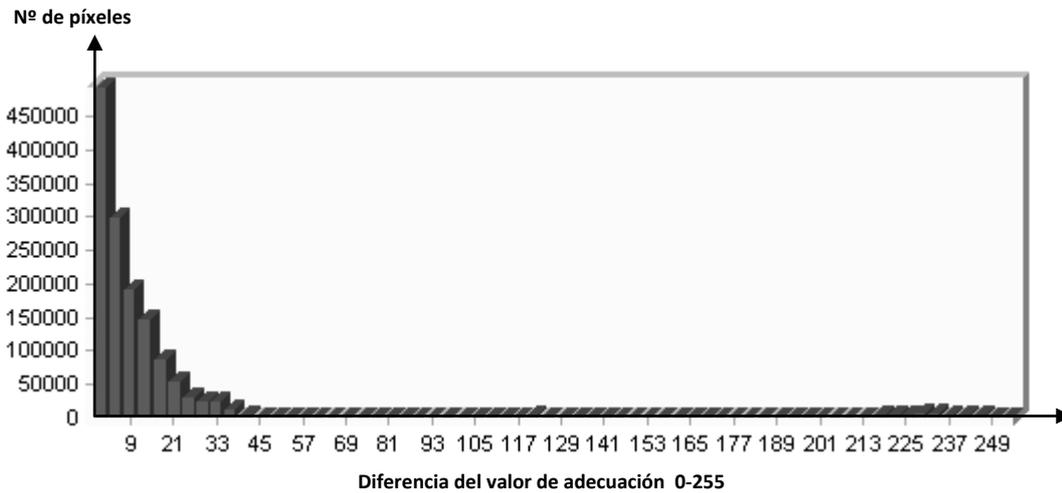


Figura 35. Histograma de las diferencias del valor de adecuación final entre el factor zonas urbanas modificado y sin modificar temática y geoméricamente.

2.2.3.2.6. DISTORSIÓN DEL FACTOR DISTANCIA A HOSPITALES

Este factor pretende aportar cierta sostenibilidad desde el punto de vista social, facilitando el acceso de la población a los servicios básicos y minimizando la distancia a centros sanitarios. Para su distorsión temática se propone cambiar los identificadores para eliminar aleatoriamente el 15 % de los hospitales. El algoritmo consiste entonces en una prueba condicional del tipo:

```
SI(ALEATORIO())<0.15)
  CODE =0
SINO
  CODE = 1
FINSI
```

La tabla 14 muestra la aplicación del algoritmo en una hoja de cálculo en lenguaje de programación excel:

ID	CODE	MUNICIPIO	CENTRO
1	1	Madrid	429
2	1	Majadahonda	319
3	1	Madrid	313
4	1	Madrid	099
5	0	Ciempozuelos	532
6	1	Madrid	407
7	1	Madrid	148
8	1	Madrid	398
.	.	.	.
.	.	.	.
.	.	.	.

Tabla 14. Distorsión temática para los polígonos del factor distancia a hospitales.

El diagrama que aparece en la figura 36 reproduce el factor distancia a hospitales. Como en los anteriores procedimientos, se comienza importando el archivo *shape* con el comando *shapeidr* y a continuación se rasteriza con *pointras*. El identificador temático se modifica con el comando *assign*. El procedimiento *distance* calcula las distancias desde cada píxel a la entidad más cercana y este mapa obtenido se limita espacialmente con la máscara obtenida a partir del *búffer* de cada entidad con 10 km de distancia, mediante el comando *overlay*.

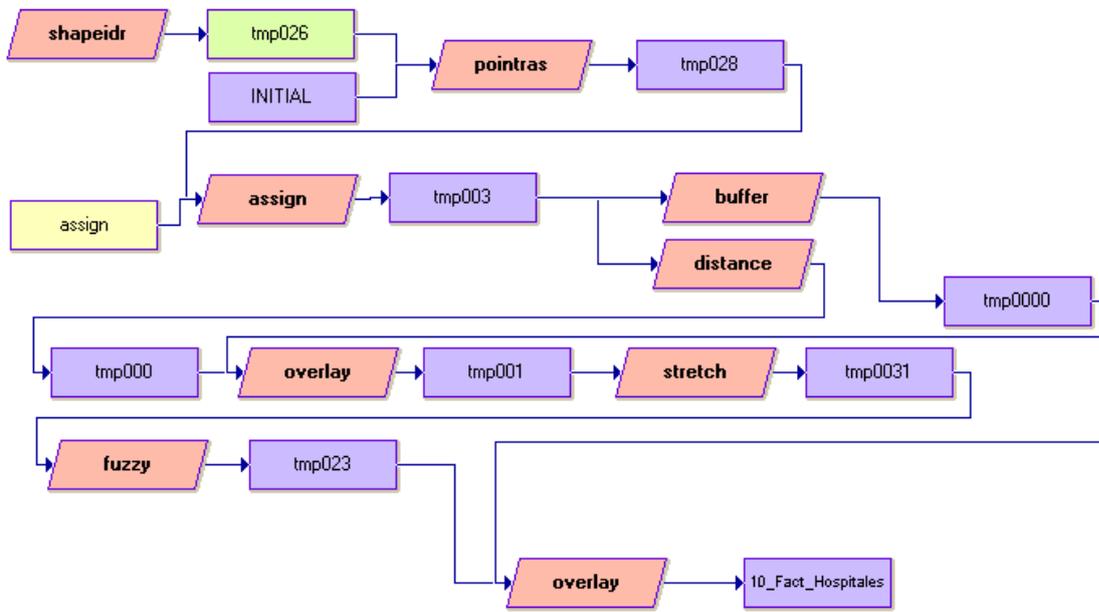


Figura 36. Proceso de distorsión temática del factor distancia a hospitales.

El histograma de la figura 37 muestra las diferencias del valor de adecuación final entre el factor modificado y sin modificar con un máximo de 120.000 píxeles (diferencia del valor de adecuación cercano a cero). Sabiendo que la imagen entera es de 7 millones de píxeles, este máximo no sobrepasa ni el 2 % de todos los píxeles de la imagen.

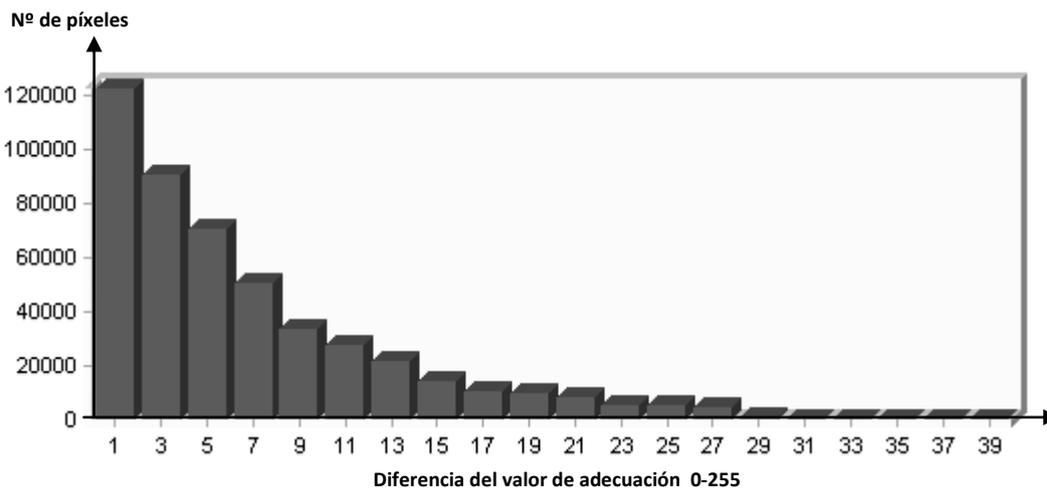


Figura 37. Histograma de diferencias del valor de adecuación final entre el factor distancia a hospitales modificado y sin modificar temáticamente.

2.3. EJECUCIÓN DEL MODELO DISTORSIONADO Y EVALUACIÓN DE LOS RESULTADOS A PARTIR DEL CÁLCULO DE LA INCERTIDUMBRE MEDIA CUADRÁTICA

Este capítulo muestra los últimos pasos en el análisis de incertidumbre de un modelo basado en técnicas de EMC y SIG.

Una vez distorsionados temática y geoméricamente cuatro veces todos los factores, se procedió a normalizar de nuevo todos y cada uno de ellos. Con todas las versiones distorsionadas se ejecutó el modelo treinta veces, escogiendo aleatoriamente los factores de entre los cinco disponibles (los cuatro distorsionados más el original).

Para el cálculo del índice de Incertidumbre Medio Cuadrático (IMC) se propone emplear las treinta ejecuciones del modelo, con los diferentes factores distorsionados. Los factores empleados en cada ejecución se escogen aleatoriamente de entre las réplicas de los factores distorsionados. El algoritmo de aleatoriedad empleado para crear la macro que ejecuta los diversos resultados cambia las terminaciones de los comandos que llaman a las cinco variables:

```
C=ENTERO(ALEATORIO()*5)
SI(O(C=0;C=5)
D=""
SINO
D=C
FINSI
CONCATENAR("1_FACT_USO_DEL_SUELO_mod";D)
```

Acto seguido, con la serie de resultados obtenidos se obtiene el IMC para cada píxel. Este índice ha sido desarrollado por M. Gómez Delgado y J. Bosque Sendra (2004). El IMC muestra para cada píxel el cálculo del error medio cuadrático con un valor que expresa la diferencia media de los diferentes resultados distorsionados respecto al resultado original. Su fórmula es la siguiente:

$$IMC = \sqrt{\frac{\sum_0^n (X_i - X)^2}{n-1}}$$

Donde:

X_i es el resultado de cada versión distorsionada

X es el resultado original

n el número de ejecuciones realizadas.

Este cálculo es posible realizarlo de manera sencilla a partir de la combinación de los comandos **Overlay** y **Scalar** de IDRISI. Para agilizar el proceso se utiliza una macro creada en una hoja Excel y un cálculo final en la calculadora de imágenes de IDRISI que realiza la raíz cuadrada de las diferencias.

3. RESULTADOS

En el estudio realizado se ha calculado el índice de IMC del Modelo de innovación y sostenibilidad de crecimiento urbano para la localización óptima de zonas residenciales en la Comunidad de Madrid para el año 2020. Este índice se ha obtenido a partir de treinta ejecuciones con diferentes distorsiones de los nueve factores de dicho modelo. Los factores empleados en cada una de las treinta ejecuciones se han escogido aleatoriamente de entre cinco replicas distorsionadas.

Todos los factores se han distorsionado geoméricamente mediante la transformación *rubbersheet* tras definir y calcular sus vínculos de deformación. En la distorsión temática se ha empleado una recodificación de polígonos mediante matrices de confusión y los cambios temáticos se han restringido aplicando unos límites de cambio según ciertas características o atributos de las entidades. Los Modelos Digitales Derivados se han distorsionado a partir de su MDE mediante varios Modelos Digitales de Incertidumbre obtenidos a partir de Redes de Triángulos Irregulares Aleatorias previamente definidas y calculadas.

Para analizar la distorsión que se ha introducido en las variables, se ha calculado el coeficiente de correlación entre variables deformadas y sin deformar para los diferentes métodos de distorsión empleados. La tabla 15 muestra los coeficientes de cada tipo de distorsión y la combinación de ambas, para cada factor.

	Distorsión geométrica por <i>rubbersheet</i>	Distorsión temática a nivel de entidad	Combinación de ambas distorsiones
Usos del suelo (CLC1990)	0,86	0,88	0,78
Vulnerabilidad mantos acuíferos	0,97	np	0,97
Tipo de suelo	0,95	0,96	0,93
Pendiente	np	np	0,99
Accesibilidad a carreteras	0,99	0,99	0,99
Distancia a zonas urbanas	np	np	0,95
Geotecnia	0,97	0,99	0,97
Accesibilidad a hospitales	np	np	0,94
Orientación del terreno	np	np	0,83

Tabla 15. Coeficientes de correlación entre los factores originales y distorsionados. (np = no practicado).

Teniendo en cuenta que se han aplicado los mismos criterios de distorsión, el factor que muestra más incertidumbre es el de usos del suelo, con un coeficiente de correlación del 78%. Es muy probable que este resultado venga afectado por emplear una matriz de confusión con observaciones en campo y por el reducido tamaño de las entidades del mapa CORINE. Por otro lado, el factor que menos incertidumbre produce es el factor de accesibilidad a carreteras, resultado poco esperado pero razonable, puesto que es el único factor con elementos lineales y es muy probable que este tipo de entidad proporcione poca incertidumbre.

En las regresiones para los Modelos Digitales Derivados obtenidos a partir de la distorsión del MDE, el coeficiente de correlación muestra valores muy elevados para el factor de pendiente y cercano al 85% para el factor de orientación.

Como ya se ha comentado anteriormente, una vez diseñadas y aplicadas las metodologías de deformación, se han obtenido treinta resultados del modelo con los factores distorsionados para evaluar su robustez mediante un análisis de incertidumbre. Posteriormente, la validación del modelo ha finalizado con la obtención del índice de IMC y ahora se procede a su evaluación numérica y su contraste con los resultados del modelo.

El índice de IMC indica los valores de incertidumbre para cada píxel, pero los píxeles con valor de adecuación nula puede que no pertenezcan a la solución. Por este motivo, en la figura 38 se ha superpuesto los polígonos de asignación óptima para zonas residenciales. Por su parte, las localizaciones más óptimas corresponden a las que contienen bajos niveles de incertidumbre.

En la figura 39 se muestra la distribución de la IMC para todo el área de estudio. El valor de máxima incertidumbre está en 34%, pero se puede observar que la mayoría de los valores están por debajo del 15%. Por tanto y atendiendo al proceso de validación, el modelo presenta valores de incertidumbre moderados, asegurando robustez en los procesos de obtención de resultados.

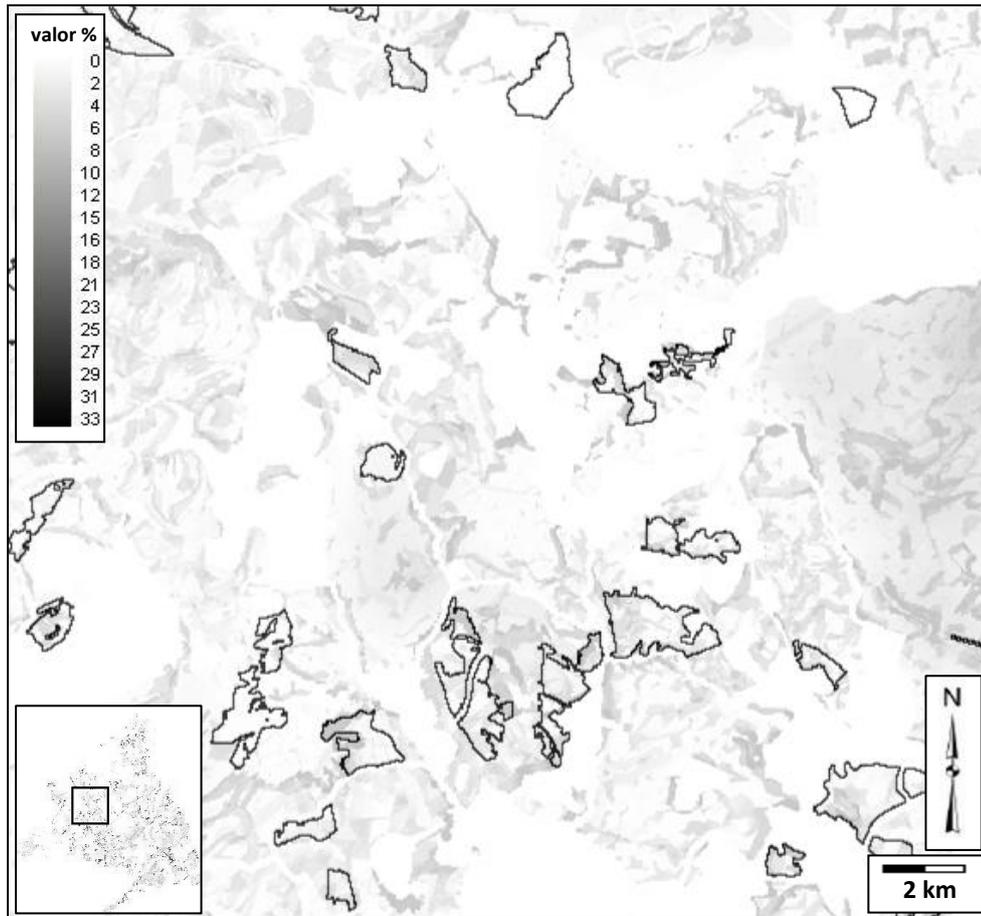


Figura 38. Resultado del cálculo de la IMC, introduciendo distorsión en los 9 factores y ejecutando el modelo 30 veces. Las zonas residenciales óptimas obtenidas en el modelo original aparecen superpuestas.

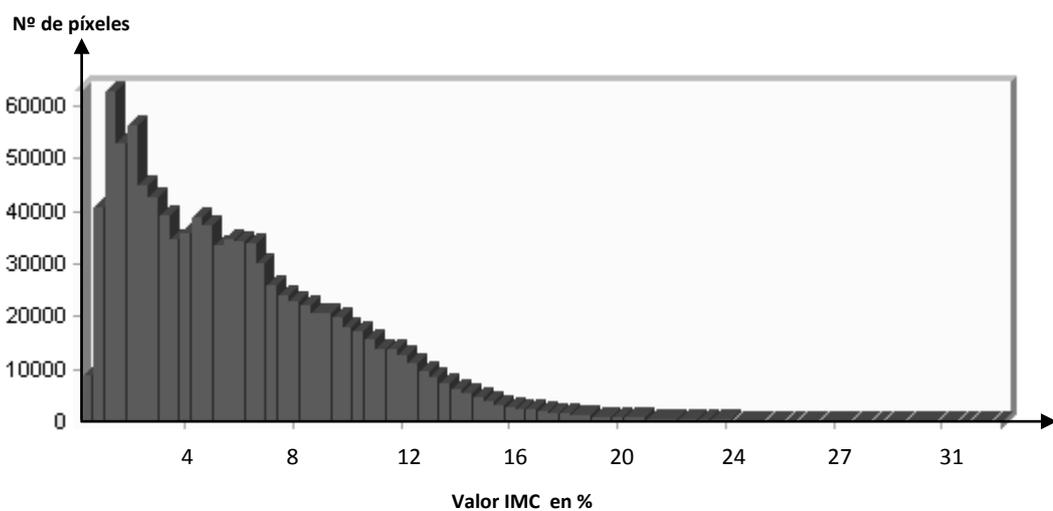


Figura 39. Distribución de los valores de IMC del modelo analizado.

Para reflejar el efecto producido por la cantidad de distorsión introducida, se ha calculado una serie de escenarios con distorsión temática en sólo uno de sus factores. La figura 40 muestra el mapa de IMC empleando sólo diez distorsiones temáticas del factor usos de suelo. Este mapa resultante muestra que el efecto de distorsión temática en un solo factor incide en el índice de IMC en función del número de distorsiones empleadas.

Por tanto, cuando se utilizan muchas variables iniciales distorsionadas, los resultados de la IMC muestran una mejor propagación en el modelo de la magnitud de error de las variables iniciales, pues la gama de posibilidades aumenta y se pueden diferenciar más niveles de fiabilidad en los resultados.

En otras palabras, se refleja mejor la incertidumbre espacial y temática de los mapas iniciales con el empleo de un número elevado de variables iniciales distorsionadas para el cálculo del índice de IMC. Los valores del IMC quedan mejor definidos si los errores se propagan con una amplia gama de valores aleatorios.

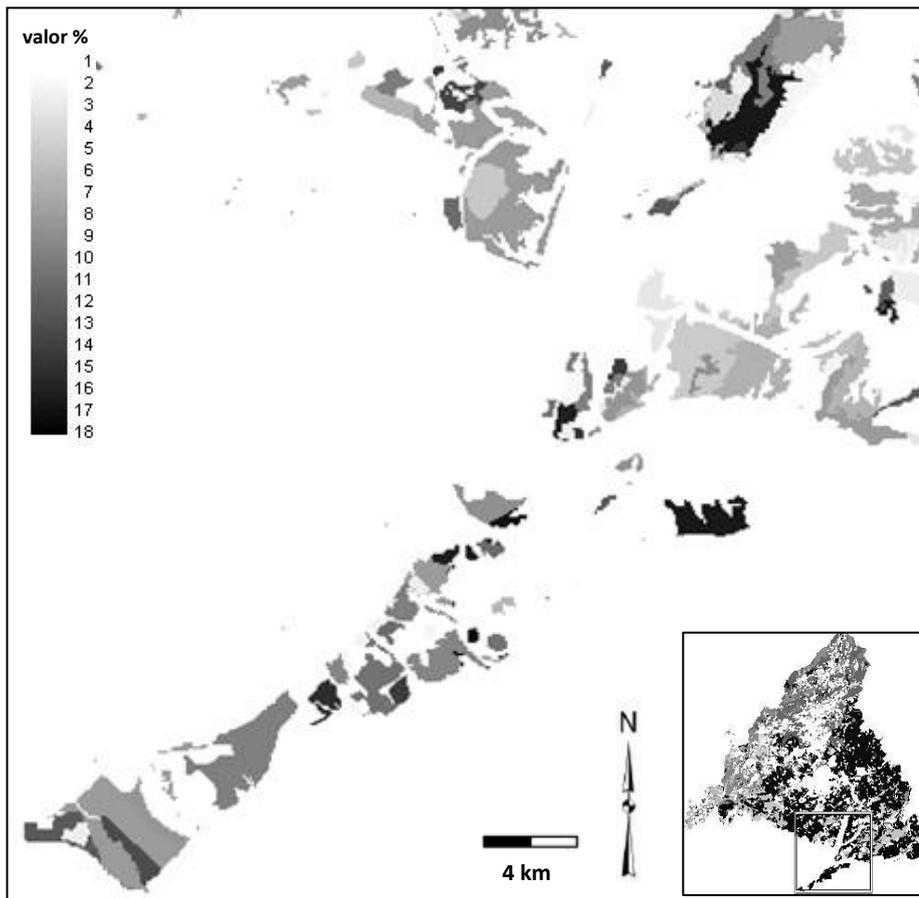


Figura 40. Resultado de IMC del modelo empleando diez distorsiones temáticas del factor usos de suelo. Se observa una amplia gama de tonos que corresponden a los valores de fiabilidad de cada clase, obtenidos de su matriz de confusión correspondiente.

No obstante, en el proyecto de investigación SIMURBAN se realiza un paso más para la metodología de asignación espacial de parcelas. En este paso entran a competir por el espacio los usos residencial, comercial e industrial. En la figura 41 se muestra el mapa de aptitud de las zonas residenciales sin distorsionar. Ante la asignación multiobjetivo, los valores del índice de IMC se pueden aplicar a cada mapa de aptitud y así provocar un descenso del valor en los píxeles que tienen una incertidumbre alta.

Para aplicar el índice de IMC al mapa de aptitud, se debe normalizar e invertir el índice de IMC y luego multiplicarlo pixel a pixel con su mapa de aptitud. De esta manera, los píxeles que tengan más incertidumbre verán disminuido su valor de aptitud. La figura 42 muestra la combinación del índice de IMC con el resultado del modelo.

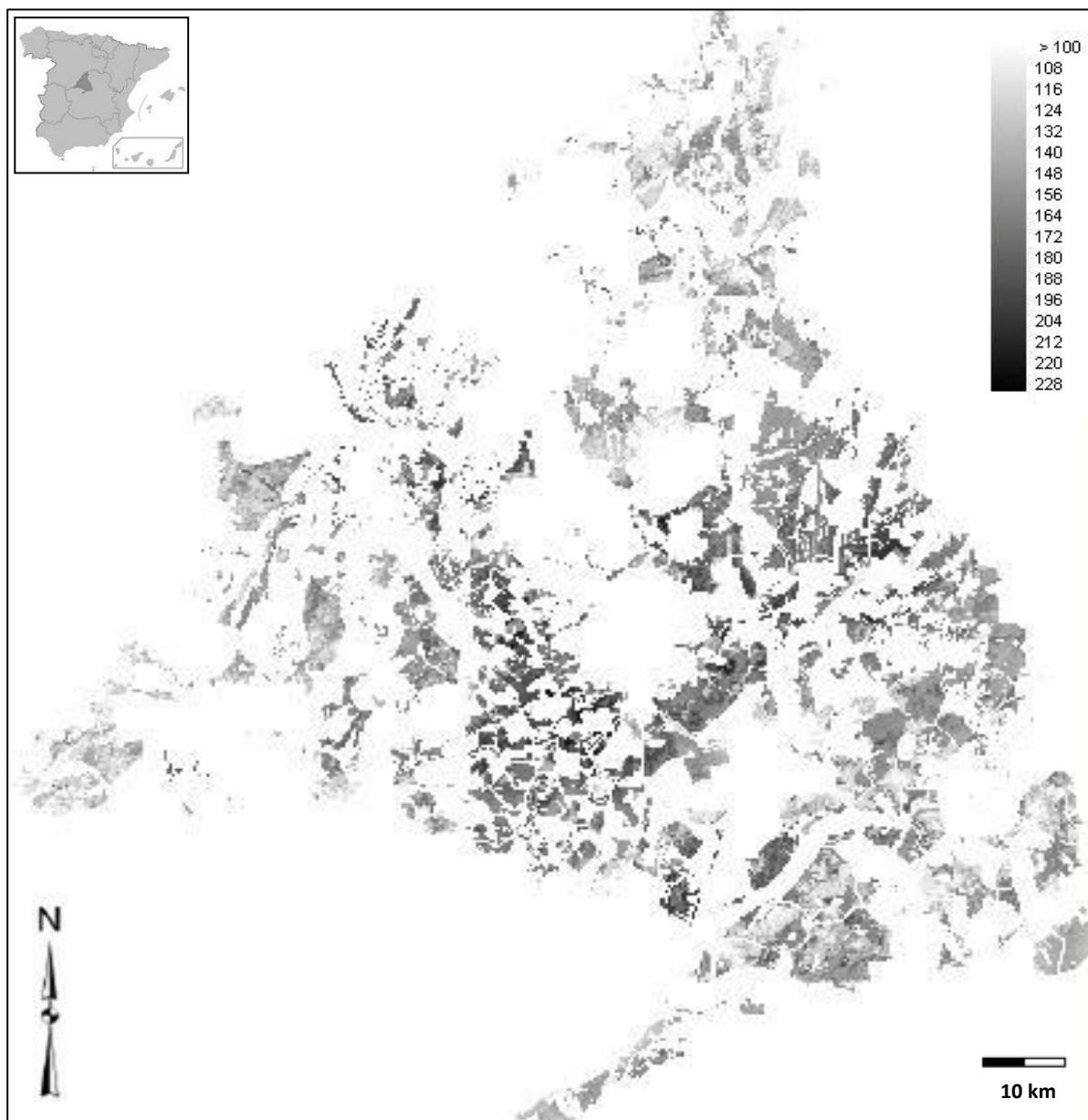


Figura 41. Mapa de aptitud de las zonas residenciales sin distorsión.

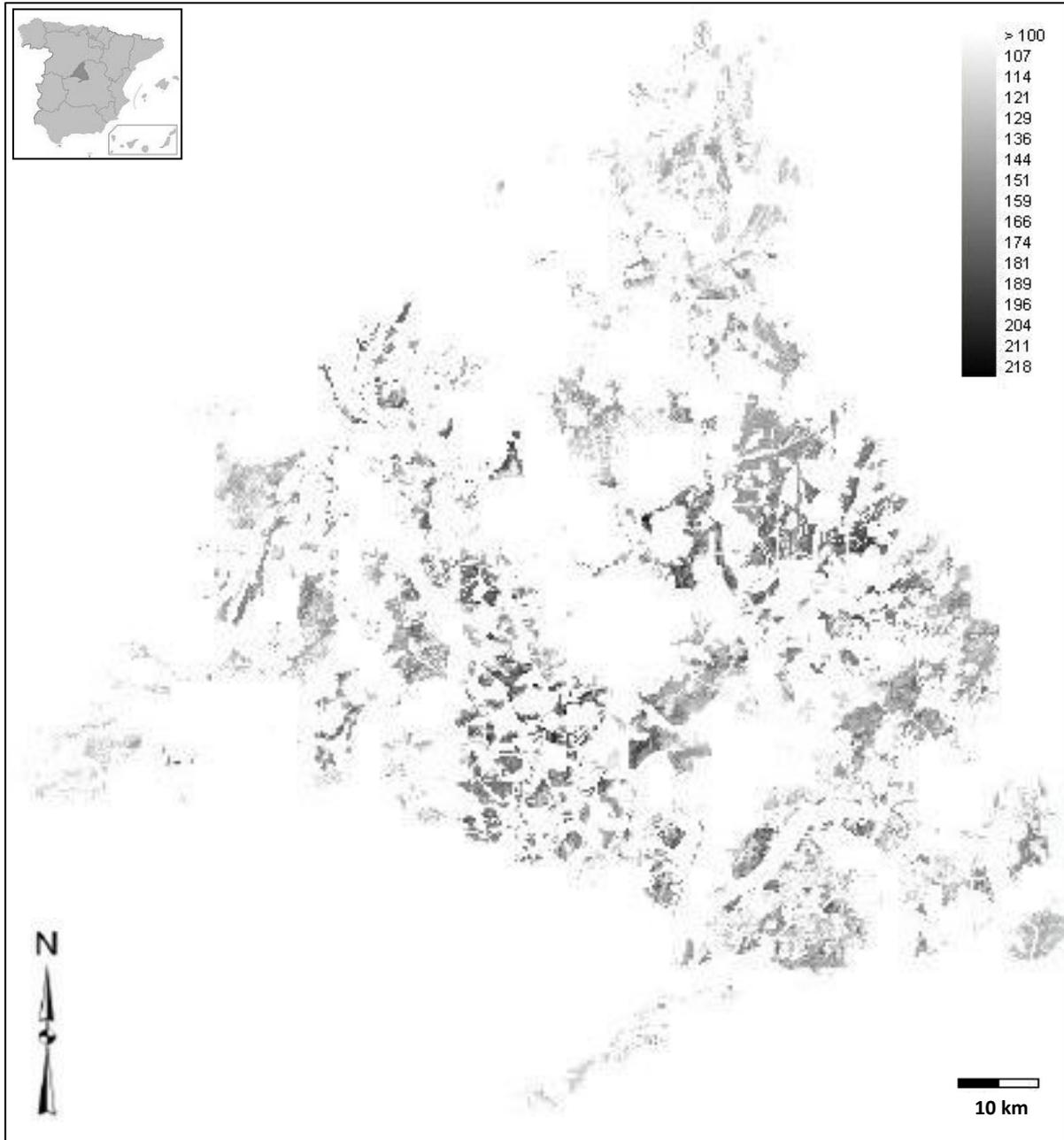


Figura 42. Mapa de aptitud de las zonas residenciales tras aplicar su índice de Incertidumbre Media Cuadrática.

4. CONTRASTE DE LA METODOLOGÍA

El contraste con otras metodologías favorece la validación de éstas a la vez que desarrolla nuevas áreas de investigación.

Para ello, se pretende contrastar la metodología empleada con la propuesta por Gómez Delgado y Bosque Sendra (2004). El trabajo a comparar ha sido desarrollado por Grace Benavides como parte del proyecto fin de máster “Análisis de incertidumbre de un modelo basado en técnicas SIG y EMC” (2011) donde ha analizado el mismo modelo. En la figura 43 se compara gráficamente el índice de IMC obtenido en ambos proyectos una zona del área de estudio.

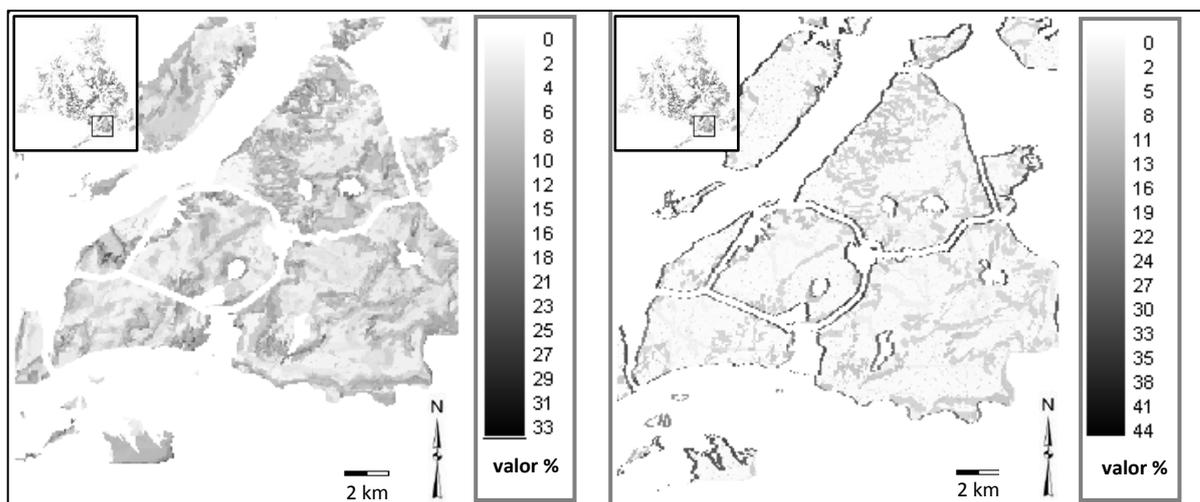


Figura 43. A la izquierda, índice de IMC obtenido en este estudio. A la derecha, aparece el índice de IMC obtenido por distorsión con polinomios cúbicos y ruido temático. Fuente: Benavides, 2011.

Tras la comparación gráfica entre los dos métodos, la metodología empleada en el presente proyecto arroja resultados más uniformes. En cambio, la metodología contrastada muestra valores altos en los contornos, producto en gran medida de la distorsión geométrica.

Por otro lado, las figuras 44 y 45 muestran los valores del índice de IMC obtenidos en ambas metodologías.

Realizando una comparación entre los dos histogramas, se puede observar una distribución más dispersa en la metodología empleada en el presente proyecto. Además, en la figura 45 parece poder diferenciarse los dos tipos de distorsión, pues presenta una distribución bimodal. Una de las distribuciones está centrada en el valor 75 y la otra en los valores cercanos a 0. Estas distribuciones deben corresponder respectivamente a la distorsión geométrica y a la temática, puesto que los valores más elevados se muestran en los contornos ya mencionados de la figura 43. Aunque no sea tan evidente, estos datos podrían indicar que se le da más importancia a la incertidumbre geométrica que a la temática, campo de estudio que se deja como propuesta para otros proyectos.

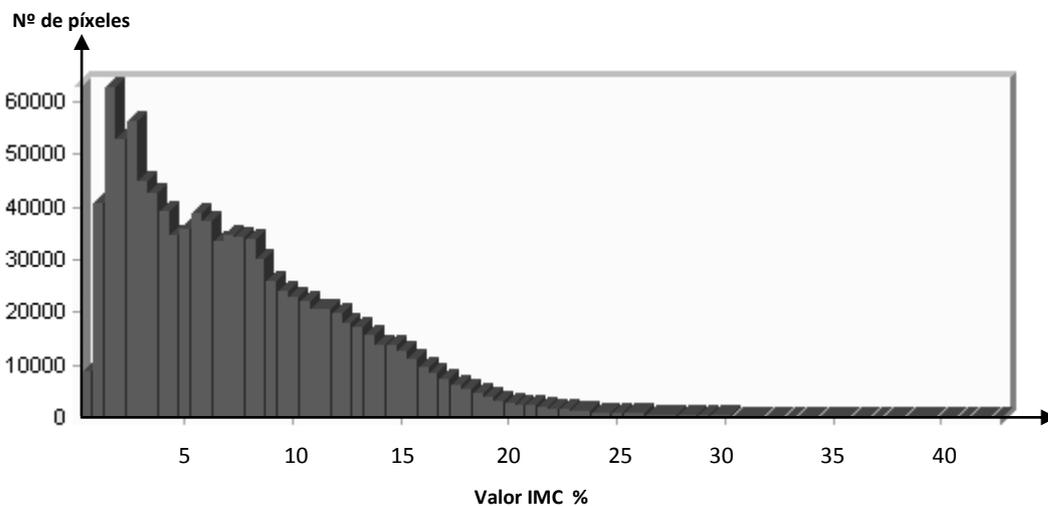


Figura 44. Distribución de la IMC obtenida mediante la metodología empleada en este proyecto.

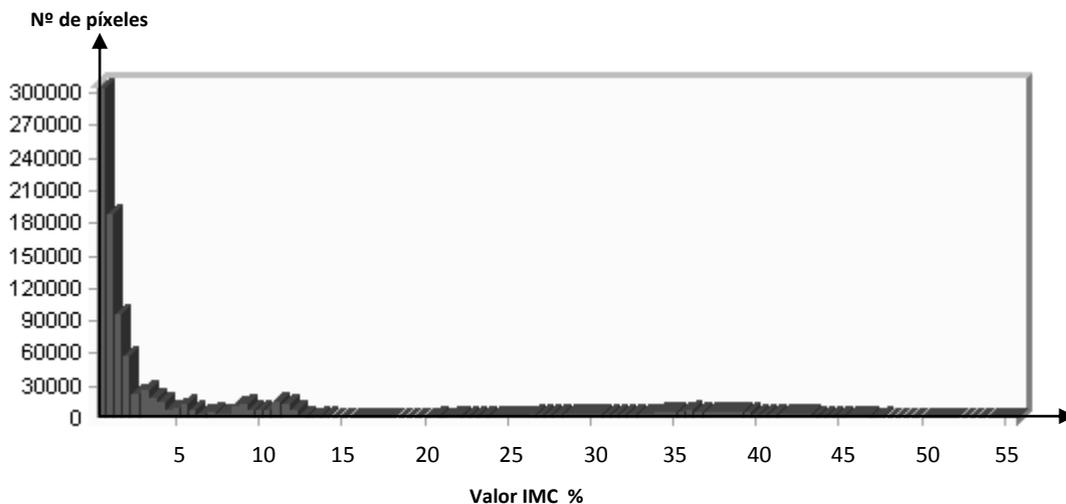


Figura 45. Distribución de la IMC obtenida en el trabajo de Benavides (2011).

5. CONCLUSIONES

En el presente trabajo fin de carrera se ha desarrollado una metodología alternativa para el análisis de la incertidumbre que en los resultados de un modelo basado en técnicas EMC y SIG pudieran introducir los posibles errores en los datos de partida.

Ante los resultados obtenidos, queda demostrado que el empleo de este tipo de análisis es una herramienta indispensable para la validación de estos modelos y para confirmar la robustez de los resultados que se obtienen. No obstante, la metodología empleada es una técnica con un cierto nivel computacional pues repite numerosos procesos, calculando treinta resultados del modelo y posteriormente analizando las variaciones que éstos sufren.

Como se ha demostrado, la metodología propuesta dirige y controla las perturbaciones a la hora de distorsionar los factores empleados en los modelos basados en técnicas de EMC y SIG. Para poder realizar este análisis de incertidumbre se precisa una serie de parámetros que indique los niveles de distorsión espacial y temática a realizar.

En cuanto a la metodología relacionada con la distorsión espacial, se debe plantear el método matemático que proporcione la mejor deformación. No obstante, la lógica geométrica induce a que unas simples traslaciones pueden ser suficientes para introducir la incertidumbre en el resultado del modelo, siempre y cuando se apliquen en las diversas direcciones de los puntos cardinales. Asimismo, a pesar de que la transformación por distorsión sea laboriosa y complicada, es la opción más recomendada para la introducción de error, pues queda suficientemente demostrado que el empleo de las transformaciones polinómicas no controla la magnitud ni el ámbito de las deformaciones aplicadas. Por otro lado, los índices de correlación muestran que el tamaño de las entidades es inversamente proporcional a la incertidumbre transmitida y por tanto, ante la misma magnitud de error empleada, polígonos de reducida área aumentarán la incertidumbre del resultado.

Por parte de la metodología empleada en la distorsión de los Modelos Digitales Derivados, los resultados obtenidos muestran que los diferentes factores pueden responder de manera diferente ante la misma distorsión del MDE original. Sólo queda entonces por denotar que ésta podría ser una nueva línea de investigación a desarrollar. No obstante, la creación de la superficie de error para distorsionar los MDE se puede crear fácilmente a partir de una Red de Triángulos Irregulares Aleatorios, cosa que simplifica mucho los procesos de distorsión.

Con respecto a la metodología empleada en la distorsión temática, los análisis realizados muestran que aunque la matriz de confusión no se obtenga realizando un trabajo de muestreo en campo, ésta se puede obtener realizando apreciaciones sobre la incertidumbre entre clases. Además, los mapas distorsionados proporcionan la misma matriz de confusión y esto es una característica muy importante en la distorsión.

Con este trabajo queda suficientemente demostrado el empleo de la matriz de confusión para la introducción de error puede ser un método óptimo para la transmisión de la incertidumbre temática.

Como ya se ha comentado anteriormente, el error de las variables iniciales presenta más aleatoriedad estadística cuanto más elevado es el número de reproducciones distorsionadas. Por lo tanto, es necesario tener un espacio muestral inicial elevado, que acerque el error introducido a una normalidad estadística. En este contexto, queda demostrado que no es suficiente emplear sólo un par de modificaciones para cada factor, pues no refleja la matriz de confusión en toda su integridad y no se reflejan todas las direcciones posibles de la incertidumbre espacial. Una muestra no tiene que ser muy grande para que la distribución de muestreo de la media se acerque a la normal. Los estadísticos utilizan la distribución normal como una aproximación a la distribución de muestreo, siempre que el tamaño de la muestra sea al menos de treinta, pero la distribución de muestreo de la media puede ser casi normal con muestras incluso de la mitad de ese tamaño. En este caso, empleando unas diez alteraciones para diez variables de entrada puede ser más que suficiente.

Finalmente y como ya se ha comentado anteriormente, se puede concluir que aplicar el índice de IMC en los mapas de aptitud para la asignación final del uso el suelo modelado favorece la seguridad de los resultados y reduce la posibilidad de encontrar errores producidos por la incertidumbre de las variables iniciales.

6. REFERENCIAS

- AGUMYA, A. y HUNTER, G.J. (1999): "A risk-based approach to assessing the 'fitness for use' of spatial data". *URISA Journal*, 11, pp. 33-44.
- APPELBAUM L.T. (1982): "Geodetic datum transformation by multiple regression equations" New Mexico State University. p. 207-203.
- ARBIA, G. Y HAINING, R. (1998): "Error propagation modelling in raster GIS: overlay operations", *International Journal of Geographical Information Science*, 12, 2, pp. 145-167.
- BARREIRA, P. (2011): "Metodología aplicada a la detección de errores temáticos en el CORINE Land Cover a través del estudio de cambios: Comunidad de Madrid (2000-2006)", Proyecto fin de Carrera. Universidad de Alcalá.
- BEARD, K. (1997): "Representations of data quality". *Geographic Information Research: Bridging the Atlantic*, M. Craglia y H. Couclelis (Eds), pp. 280-294 (London: Taylor & Francis).
- BEARD, K. y BUTTENFIELD, B., (1999): "Detecting and evaluating errors by graphical methods". *Geographical Information Systems*, P.A. Longley, M.F. Goodchild, D.J. Maguire y D.W. Rhind (Eds), pp. 219-233 (Chichester, UK: Wiley).
- BEARD, K. y MACKANESS W. (1993): "Visual access to data quality in geographic information systems". *Cartographica*, 30, pp. 37-45.
- BÉDARD, Y. (1987): "Uncertainties in land information systems databases". *Proceedings of Eighth International Symposium on Computer-Assisted Cartography*. Baltimore, pp. 175-184.
- BENAVIDES, G. (2011): "Análisis de incertidumbre de un modelo basado en técnicas SIG y EMC" Proyecto fin de Máster. Universidad de Alcalá.
- BURNICKI, A.C., BROWN, D.G., GOOVAERTS, P. (2007): "Simulating error propagation in land-cover change analysis: the implications of temporal dependence". *Computers, Environment and Urban Systems*, Vol.31 (3), pp. 282-302.
- BUTTENFIELD, B.P. y BEARD, K. (1991): "Visualizing the quality of spatial information". *Proceedings of AUTO-CARTO 10*, pp. 423-427.
- BUTTENFIELD, B.P. (1993): "Representing data quality". *Cartographica*, 30, pp. 1-7.
- BRIGGS, I.C. (1984): "Machine contouring using minimum curvature". *Geophysics*. p. 39-48.

- CHUVIECO, E., (2008): "Teledetección Ambiental: La observación de la Tierra desde el espacio" p. 501.
- CROSETTO, M. Y TARANTOLA, S. (2001): "Uncertainty and sensitivity analysis: tools for GIS-based model implementation", *International Journal of Geographical Information Science*, 15, 5, pp.415-437.
- CUARTERO SÁEZ A. (2003): "MDE generados con imágenes SPOT-HRV y TERRA-ASTER". Tesis doctoral codirigida por Dr. D. F. J. Ariza López, de la Universidad de Jaén y Dr. D. A. M. Felicísimo Pérez, de la Universidad de Extremadura Cáceres.
- DAVIS, T.J. Y KELLER, C.P. (1997): "Modelling uncertainty in natural resource analysis using fuzzy sets and Monte Carlo simulation: slope stability prediction". *International Journal of Geographical Information Science*, 11, 5, pp. 409-434.
- DEVILLERS, R., y JEANSOULIN, R. (Eds) (2006): "Fundamentals of Spatial Data Quality" (London: ISTE).
- DEVILLERS, R.; BÉDARD, Y.; JEANSOULIN, R. y MOULIN, B. (2007): "Towards spatial data quality information analysis tools for experts assessing the fitness for use of spatial data". *International Journal of Geographical Information Science*, Vol.21, No.3-4, pp. 261-282.
- DEWHURST, W.T. (1990): "The aplicaton of minimum curvature-derived surfaces in the transform of positional data from the North America datum of 1927 to the North America datum of 1983". *NOAA Technical Memorandum* 30pp.
- DRAPER, N.R., SMIDTH, H. (1981): "Aplied regression analysis" *Wiley*.
- DRECKI, I. (2002): "Visualisation of uncertainty in geographic data. *Spatial Data Quality*". W. Shi, P.F. Fisher and M.F. Goodchild (Eds), pp. 140–159 (London: Taylor & Francis).
- EMMI, P.C. Y HORTON, C.A. (1996): "Seismic risk assessment, accuracy requirements, and GIS-based sensitivity analysis". En Goodchild, M.F, Steyaert, L.T. Y Parks, B.O. (Eds.): *GIS and Environmental modeling: progress and research issues*. Fort Collins, CO: GIS World Books, pp. 191-195
- FIEDLER, S.: "Simplified key of the REFERENCE SOIL GROUP (WRB)". Universität Hohenheim. Institute of Soil Science and Land Evaluation. World Reference Base. FAO Soil Map of the World. *World Reference Base the Revised Legend of FAO*. p. 19.
- FISHER, P.F., (1994): "Animation and sound for the visualization of uncertain spatial information". En *Visualization in Geographic Information Systems*, H.M. Hearnshaw y D.J. Unwin (Eds), pp. 181–185 (Chichester, UK: Wiley).

- GERVAIS, M. (2003): "*Pertinence d'un manuel d'instructions au sein d'une stratégie de gestion du risque juridique découlant de la fourniture de données géographiques numériques*". PhD thesis, Sciences Géomatiques, Université Laval, Québec.
- GERVAIS, M. (2006): "On the Importance of External Data Quality in Civil Law". En *Fundamentals of Spatial Data Quality*, R. Devillers y R. Jeansoulin (Eds), pp. 283–300 (London: ISTE).
- GARCIA DE ARANGO, A. (1974): "Elasticidad teórica". *Instituto Geográfico Nacional*.
- GÓMEZ DELGADO, M. Y BOSQUE SENDRA, J. (2004): "Aplicación de análisis de incertidumbre como método de validación y control del riesgo en la toma de decisiones", *GeoFocus (Artículos)*, nº 4, p. 179-208. ISSN: 1578-5157
- GÓMEZ DELGADO, M. (2003): "Sistemas de Información Geográfica y toma de decisiones: control del riesgo a partir de análisis de sensibilidad y análisis de incertidumbre". Tesis doctoral. Texto policopiado. Universidad de Alcalá.
- GONZALEZ-MATESANZ, F.J., QUIROS R., CANO, M. , SÁNCHEZ SOBRINO, J. Y DALDA, A. (2004): "Cambio de datum ED50 – ETRS89. Métodos y resultados". *Congreso Nacional de Topografía y Cartografía TOPCART*.
- HOWARD, D. y MACEACHREN, A.M. (1996): "Interface design for geographic visualization: Tools for representing reliability". *Cartography and Geographic Information Systems*, 23, pp. 59–77.
- HUNTER, G.J. (1999): "Managing uncertainty", en Longley, P.A., Goodchild, M.F, Maguire, D.J. y Rhind, D.W. (Eds.): *Geographical information systems*. Volume II, New York, John Wiley & Sons, Inc., pp. 633-641.
- HUNTER, G.J. y REINKE, K.J. (2000): "Adapting Spatial Databases to Reduce Information Misuse Through Illogical Operations". En *Proceedings of 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences (Accuracy 2000)*, Amsterdam, pp. 313–319.
- KENNEDY, W y GENTLE, J. (1980): "Statistical Computing". Marcel Dekker Press, et al (1992) *Numerical Recipes in C*. Cambridge University Press Ross, S. (1990). *A course in Simulation*. Macmillan Publishing Company.
- KIIVERI, H.T. (1997): "Assessing, representing and transmitting positional uncertainty in maps". *International Journal of Geographical Information Science*, 11, 1, pp. 33-52.
- KITANIDIS P.K. (1997): "Introduction to geostatistics: Applications to hydrogeology" *Cambridge University Press*, p. 65

- KUBIK, K. y WU X. (1995): "Mapping from SPOT images using digital photogrammetric workstation". *17th International Cartographic Conference Proceedings*. Barcelona. 1, 266-274.
- LACHAPELLE, G. (1982): "Least Squares predictions of horizontal coordinate distortion in Canada" *Bulletin Geodesique* 56 242-257
- LEE, J., P. K. SNYDER, Y P. F. FISHER (1992): "Modeling the effect of data errors on feature extraction from digital elevation models" *Photogrammetric Engineering and Remote Sensing*, 58(10): 1461-1467
- LEITNER, M. y BUTTENFIELD, B.P. (2000): "Guidelines for the Display of Attribute Certainty". *Cartography and Geographic Information Science*, 27, pp. 3-14.
- LÓPEZ DE ULLIBARRI GALPARSORO I, PITA FERNÁNDEZ, S. (1999): "Medidas de concordancia: el índice de Kappa". Unidad de Epidemiología Clínica y Bioestadística. Complejo Hospitalario-Universitario Juan Canalejo. A Coruña (España) *Cad Aten Primaria*; 6: 169-171.
- LOVE, A .E.H. (1994): "A treatise on the mathematical theory of elasticity" Dover p.
- MCGRANAGHAN, M. (1993): "A cartographic view of spatial data quality". *Cartographica*, 30, pp. 8-19.
- MORITZ H. (1973): "Least-squares collocation". München, Bayerische Akademie der Wissenschaften, Beck in Kommission.
- MUKAI, Y., SUGIMURA, T. & ARAI, K. (1990): "Automated generation of digital elevation model using system corrected SPOT data". *Proceedings 23th International Symposium on Remote Sensing of Environment*. Bangkok. 1107-1115.
- MUKAI, Y., SUGIMURA, T., IJIMA, T., AYABE, K., TACHI, K. & NAKAZAWA, T.(1989): "Automatic computation of elevation data of side-lap area using system corrected Landsat TM data". *Technical Report. Remote Sensing Technology Center of Japan*. Tokyo.
- MUNROE, D., J. SOUTHWORTH y C. M. TUCKER. (2002): "The dynamics of land-cover change in western Honduras: Exploring spatial and temporal complexity". *Agricultural Economics*.
- PLATA ROCHA, W., GÓMEZ DELGADO, M. Y BOSQUE SENDRA, J. (2010): "Desarrollo de modelos de crecimiento urbano óptimo para la Comunidad de Madrid", *GeoFocus* (Artículos), nº 10, p. 103-134. ISSN: 1578-5157.

- PLATA ROCHA, W., GÓMEZ DELGADO, M., BOSQUE SENDRA, J. (2011): "Simulating urban growth scenarios using GIS and Multicriteria Evaluation techniques". Case study: Madrid Region, Spain" en *Environment and planning B: Planning and Design*. Aceptado y en prensa.
- PONTIUS, R. G., y SCHNEIDER, L. C. (2001): "Land cover change model validation by an ROC method for the Ipswich watershed". Massachusetts, USA. *Agriculture, Ecosystems and Environment*, 85, pp. 239–248.
- PRIEBBENOW, R. y CLERICI, E. (1988): "Cartographic applications of SPOT imagery". *International Archives of Photogrammetry*. 37 (Part B4). 289-297.
- REINKE, K.J. y HUNTER, G.J. (2002): "A theory for communicating uncertainty in spatial databases". *Spatial Data Quality*, W. Shi, P.F. Fisher y M.F. Goodchild (Eds) (London: Taylor & Francis), pp. 77–101.
- RODRIGUEZ JORDANA, J. J. (2002): "Ajuste de observaciones: el método de los mínimos cuadrados con aplicaciones a la topografía". Edicions UPC, p. 68.
- RYKIEL, E.J. JR. (1996): "Testing ecological models: the meaning of validation". *Ecological modelling*, Vol.90, pp. 229-244.
- SAAFELD, A. (1985): "A fast Rubbersheeting transformation using simplicial coordinates" *The American Cartographer*. Vol 12. p.169-173.
- SASOWSKY, K.C., G.W. PETERSEN Y B.M. EVANS, (1992): "Accuracy of SPOT digital elevation model and derivatives: Utility for Alaskas's North Slope". *Photogrammetric Eng. Remote Sens.*, 58: 815-824.
- TAHIRI D. y DONNAY J-P. (1995): "Modélisation de la précision des modèles numériques de terrain dérivés des cartes topographiques". *Proceedings of the 17th International Cartographic Conference of the ICA*, Barcelona, 1, pp.469-473
- WHITE, M.S., GRIFFIN, P. (1985): "Piecewise Linear Rubber-Sheer Map Transformation" *The American Cartographer*. Vol 12. p. 123-131

