

# Taller de GEOLOCATE

---

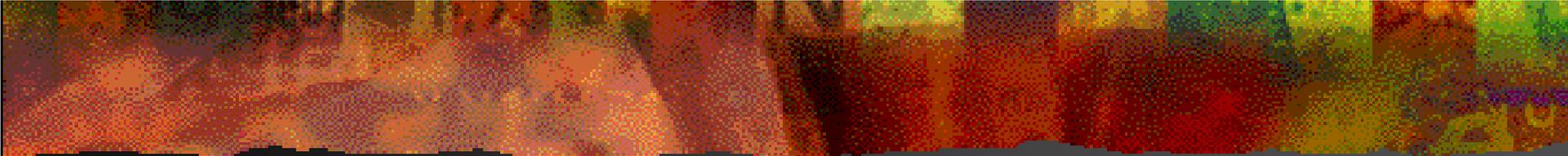
David Draper  
david.draper@upm.es

Taller de SIG y GEOLOCATE,  
Madrid 19-23 -abril 2010 GBIF

# Objetivos

---

- Presentar la herramienta y sus capacidades
- Definir una sesión básica de georreferenciación con GEOLocate
- Practicar bajo diversas situaciones
- Identificar problemas y necesidades



# Georreferenciación

---

David Draper  
david.draper@upm.es

Taller de SIG y GEOLOCATE,  
Madrid 19-23 -abril 2010 GBIF

# Punto de la situación

---

- 2.5 billones de especímenes en colecciones biológicas (Cotterill, 1995)
- Se estima que sólo el 5% se encuentran informatizados en bases de datos (Beaman & Conn, 2003)
- Una vez digitalizado, es necesario continuar trabajando para poder utilizar estos especímenes en posteriores análisis
- Un aspecto crucial de esta información es la asociación de unas coordenadas espaciales a los especímenes.

# Punto de la situación

---

- El mayor desafío de las colecciones biológicas es la asignación de las coordenadas espaciales a millones de registros antiguos (Baker & al., 1998)
- El futuro puede ser un poco más alagüeño con la confianza depositada en la tecnología GPS. Con algunas reservas...

# ¿Qué es la georreferenciación?

- La georreferenciación de localidades consiste en la asignación de coordenadas geográficas a la descripción textual de un sitio (Proctor, Blum, Chaplin, 2001; Wieczorek, 2001)
- Asignar coordenadas geográficas a las localidades contenidas en las bases de datos de las colecciones biológicas que carezcan de ellas (posprocesado)
- Georreferenciación no es únicamente asignar unas coordenadas, es necesario determinar la incertidumbre y el grado de precisión de las coordenadas

# Objetivos de la georreferenciación

- Asignar coordenadas a partir de la localidad
- Mantener o incrementar la exactitud del registro sin incrementar el error y corrigiendo localidades erróneas
- Mantener o incrementar la precisión evaluando y depurando las localizaciones disponibles para asignar coordenadas espaciales

# Fases del proceso

---

1. Normalización de las localidades
2. Asignación de coordenadas
3. Determinación de la incertidumbre
4. Validación de las coordenadas

# 1 Normalización de las localidades

- La información mínima necesaria que se debe localizar en la base de datos para poder realizar la georreferenciación es la siguiente:
  - País
  - Entidad federativa (división política)
  - Municipio (división administrativa)
  - Descripción del sitio de colecta: rasgo principal de referencia, distancia, dirección o rasgo que indique una dirección
- Identificación de duplicados:
  - **registros redundantes** que son aquellos que pertenecen a distintos especímenes pero que fueron colectados en la misma localidad (por lo tanto la descripción de localidad es la misma)
  - **falsos duplicados** que corresponden a registros de distintos colectores que trabajaron sobre una misma localidad (estos registros presentan una sintaxis distinta pero se refieren al mismo sitio)

# 1 Normalización de las localidades

CONABIO 2006

Redundantes  
Redundantes con errores de captura  
Falsos duplicados

	Localidad
{	0.25 mi E state boundary, 8 mi SW Calvillo
	0.25 mi E state boundary, 8 mi SW Calvillo
{	0.33 mi NW Ciudad de Los Ninos, 7 mi WSW Aguas Calientes
	0.33 mi NW Cuidad de Los Minos, 7 mi WSW Aguascalientes
	0.33 mi S and 1.25 E Rincon de Romos
{	0.33 mi S and 1.25 mi E Rincon de Ramos
	0.33 mi S, 1.25 mi E Rincon de Romos
{	Rincon de Romos, 5 mi W
	0.5MI W RICON DE ROMOS
	0.5 mi S La Labor
	0.5 mi N Presa Calle
	0.5 mi S LaLabor

# 1 Normalización de las localidades

Homogenización de los registros

Localidad	Descripción_homog
0.25 mi E state boundary, 8 mi SW Calvillo	0.25 mi E state boundary, 8 mi SW Calvillo
0.33 mi NW Ciudad de Los Ninos, 7 mi WSW	0.33 mi NW Ciudad de Los Ninos, 7 mi WSW
0.33 mi NW Ciudad de Los Minos, 7 mi WSW	0.33 mi NW Ciudad de Los Ninos, 7 mi WSW
0.33 mi S and 1.25 E Rincon de Romos	0.33 mi S and 1.25 mi E Rincon de Ramos
0.33 mi S and 1.25 mi E Rincon de Ramos	0.33 mi S and 1.25 mi E Rincon de Ramos
0.33 mi S, 1.25 mi E Rincon de Romos	0.33 mi S and 1.25 mi E Rincon de Ramos
Rincon de Romos, 5 mi W	Rincon de Romos, 5 mi W
0.5MI W RICON DE ROMOS	Rincon de Romos, 5 mi W
0.5 mi S La Labor	0.5 mi S La Labor
0.5 mi N Presa Calle	0.5 mi N Presa Calle
0.5 mi S LaLabor	0.5 mi S La Labor
0.5MI W RINCON DE ROMAS	0.5MI W RINCON DE ROMOS,6000FT

# 1 Normalización de las localidades

- Falta eliminar los **registros redundantes** que no pudieron desecharse previamente debido a que presentaban errores de captura, para esto se realiza una segunda consulta de agrupación basada en la columna duplicada de la descripción
- Para corregir los **falsos duplicados** se procede a desagregar cada descripción de localidad en sus distintos elementos (Atomización). Esta etapa permitirá la eliminación de los falsos duplicados además de estructurar de manera ordenada la información

# 2 Asignación de coordenadas

- La descripción de las localidades se pueden tipificar en 9 clases (Wieczorek & al., 2003):

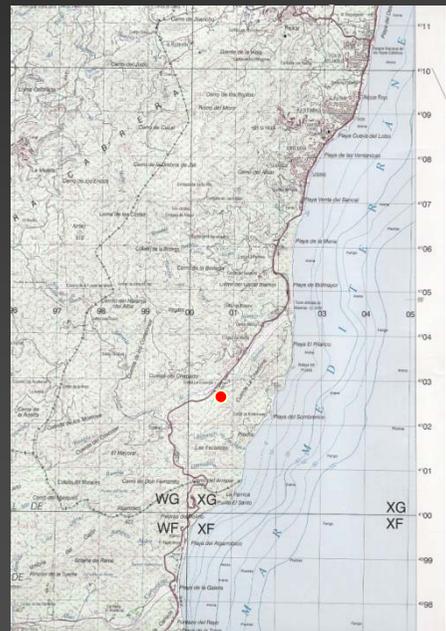
Tipo	Descripción	Ejemplos
1. Dudosa	El propio registro describe que es dudoso el nombre de la localidad o rasgo de referencia	'Tabarca?'
2. No puede ser ubicada	Los datos de la descripción se perdieron o contienen información que no corresponde con una descripción, o los nombres presentan homónimos que no pueden ser diferenciados, o no se encuentra la localidad o el rasgo en ninguna fuente disponible	'localidad no registrada', 'cumbre'
3. Comprobadamente inexacta	La descripción tiene elementos inconsistentes	'municipio de Poza Rica, Puebla'
4. Coordenadas	La descripción ya posee coordenadas	'UTM 30N 553160 4077280',
5. Localidad o rasgo	La descripción sólo posee el nombre de un elemento geográfico (una localidad, un cerro, una cueva, isla, etc.)	'unión de la carretera 70 con la 45'
6. Distancia	La descripción consiste en una distancia asociada a un elemento geográfico	'A 5 km de Sevilla'
7. Distancia a través de una ruta	Se describe una localidad con una distancia recorrida sobre una carretera, camino, río, etc.	'3.5 km N de la nacional 1'
8. Distancias en direcciones ortogonales	Se utilizan dos distancias trazadas en direcciones ortogonales a partir de un elemento geográfico	'6 km N y 4 km W de Madrid'
9. Distancia en una dirección	La descripción contiene una distancia definida en una dirección	'50 km NE de Segovia'

# 2 Asignación de coordenadas: Métodos

- Utilizando Bases de datos:

- Bases toponímicas...

- ➔ Manual (sobre mapas



- ➔ Apoyada por SIG

- Necesitamos de cartografía de base: municipios, carreteras...

# 3 Determinación de la incertidumbre

Basado en MANIS

- Fuentes potenciales de incertidumbre

Coordinate source	Source of uncertainty						
	GPS inaccuracy	locality extent	unknown datum	coordinate imprecision	distance imprecision	map scale	direction imprecision
GPS	X	X	X	X			
locality record		X	X	X			
Map		X	X	X	X	X	X
Gazetteer		X	X	X	X		X

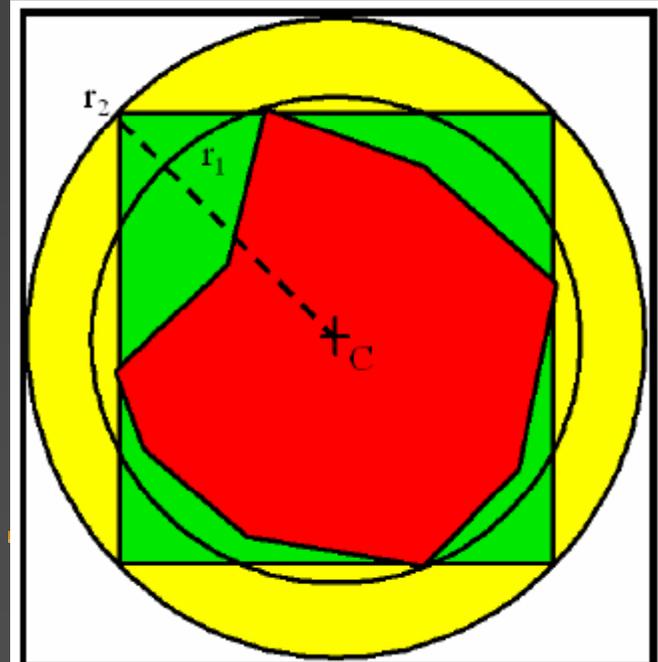
- El método de Punto-radio define una descripción de localidad con dos elementos: el punto de coordenadas y su incertidumbre, representada con una medida de longitud
- Esta distancia define una circunferencia que delimita el área en donde, con mayor probabilidad, se ubica el sitio de colecta

# 3 Determinación de la incertidumbre

- La incertidumbre está compuesta básicamente por un factor de distancia y otro de dirección que interactúan entre sí
- En la distancia se considera:
  - a) La extensión de la localidad de referencia
  - b) El desconocimiento del datum
  - c) La imprecisión en la determinación de la distancia
  - d) La imprecisión en la medición de las coordenadas
  - e) Escala del mapa utilizado
- Dentro del factor de dirección, se considera:
  - f) la imprecisión con que se definió la dirección de referencia

# Método de punto-rádío

- Una vez definida la coordenada se define el grado de incertidumbre en forma de círculo que circunscribe dicha área
- Es el método más elemental
- Las coordenadas y el radio definen la posición y su incertidumbre



# Ejemplo punto-radio

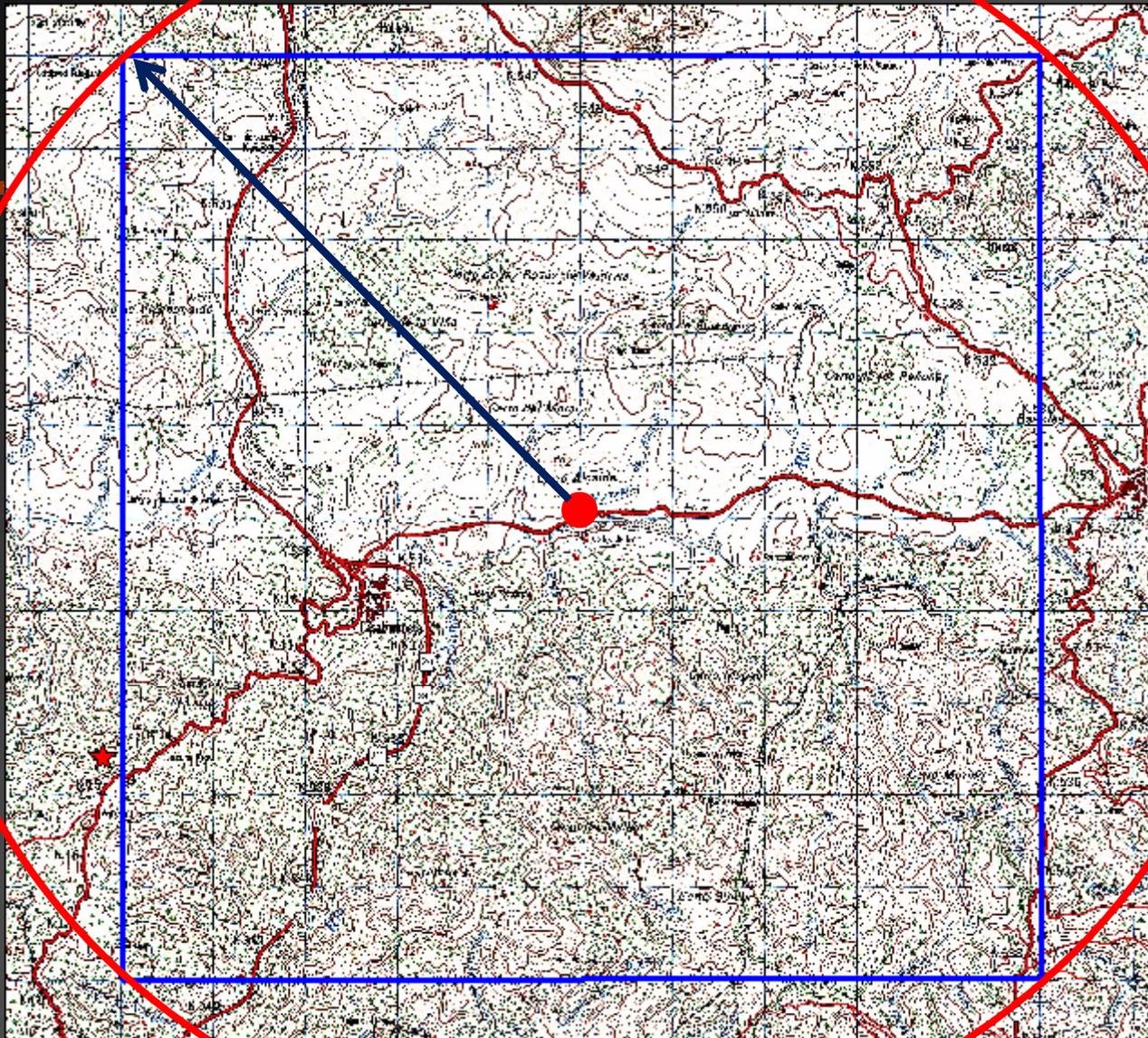
---

- Málaga: Antequera. Sierra de Camarolos: entre las poblaciones de Colmenar y Villanueva de Cauche (pedanía de Antequera), 30SUF78
  - Provincia: Málaga
  - Municipio: Antequera
  - Localidad: Sierra de Camarolos: entre las poblaciones de Colmenar y Villanueva de Cauche

# Ejemplo punto-radio

---

- Málaga: Antequera. Sierra de Camarolos: entre las poblaciones de Colmenar y Villanueva de Cauche (pedanía de Antequera), 30SUF78
  - Información accesoria
    - UTM: 30SUF78



Taller de SIG y GEOLOCATE,  
Madrid 19-23 -abril 2010 GBIF

# Otros métodos

- Shapefile method (Murphey & al., 2004)  
(<http://www.calacademy.org/research/informatics/georef/>)

Se define un polígono que representa el error exacto del registro.

- Ventajas:
  - Se digitaliza la representación real del error
- Desventajas:
  - Incremento en el tiempo de procesamiento
  - Dificultad de comparación entre errores de localización por la complejidad de las siluetas
  - Necesidad de personal especializado

# Ejemplo Shapefile method

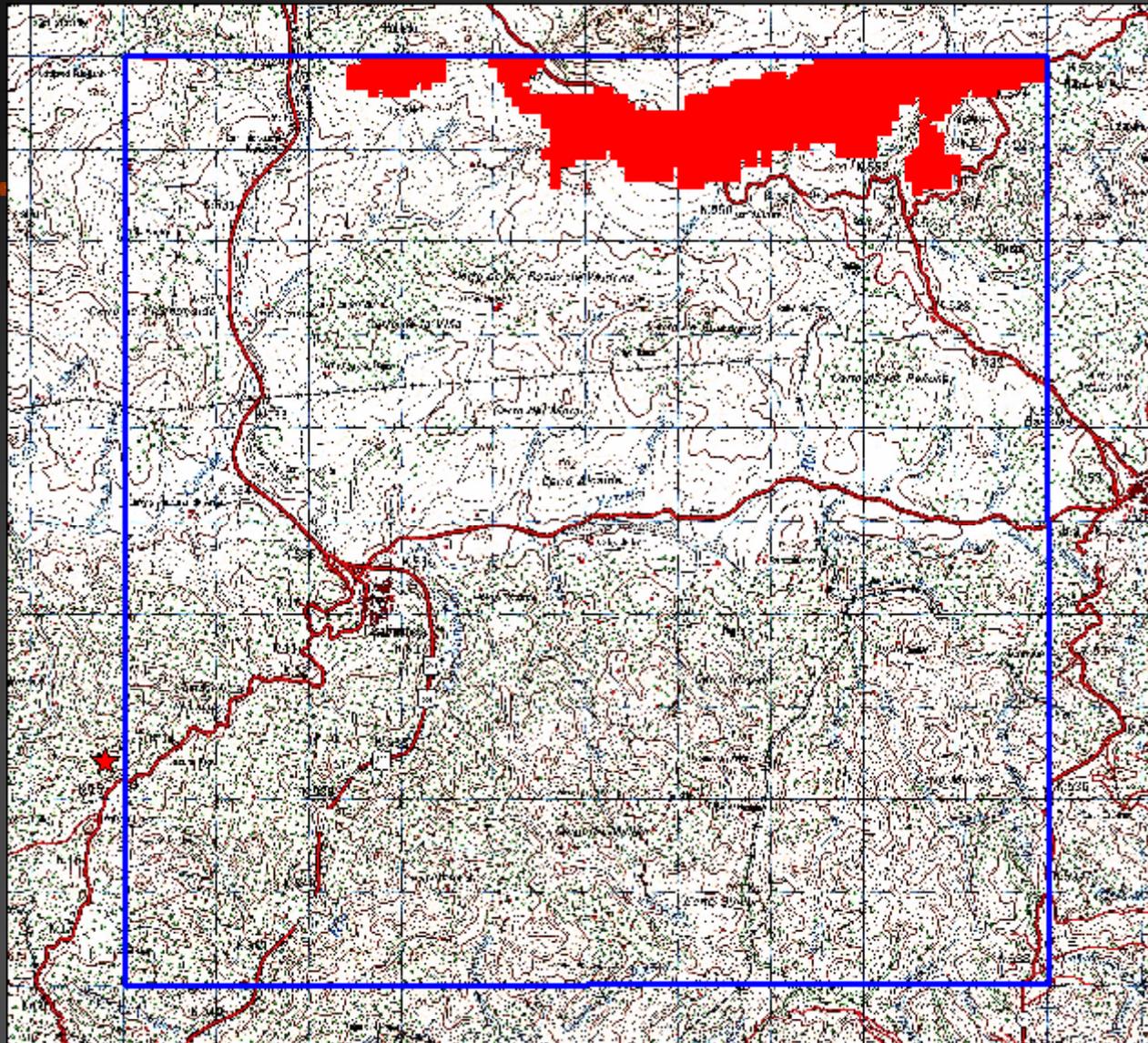
---

- Málaga: Antequera. Sierra de Camarolos: entre las poblaciones de Colmenar y Villanueva de Cauche (pedanía de Antequera), 30SUF78, 860-980 m
  - Provincia: Málaga
  - Municipio: Antequera
  - Localidad: Sierra de Camarolos: entre las poblaciones de Colmenar y Villanueva de Cauche

# Ejemplo Shapefile method

---

- Málaga: Antequera. Sierra de Camarolos: entre las poblaciones de Colmenar y Villanueva de Cauche (pedanía de Antequera), 30SUF78, 860-980 m
  - Información accesoria
    - UTM: 30SUF78
    - Rango altitudinal: 860-980 m



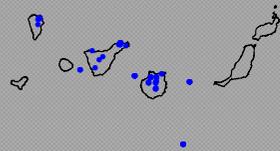
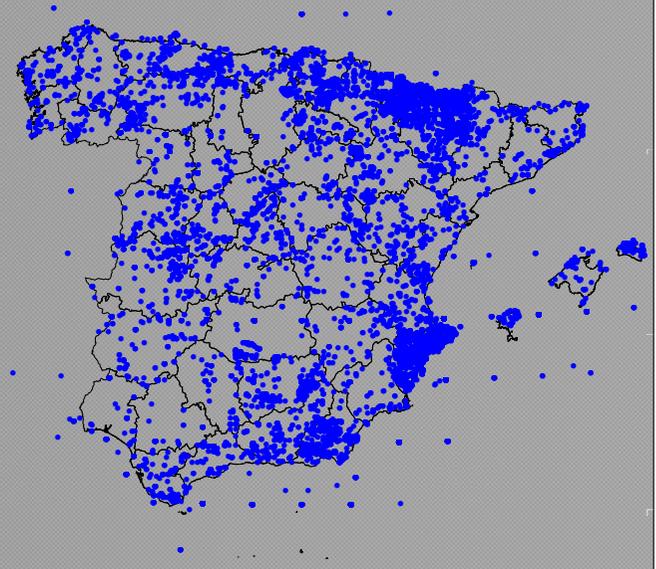
# 4 Validación de las coordenadas

- Validación con el límites administrativos (Hijmans & al., 1999)



- Ejemplo: Familia Leguminosae

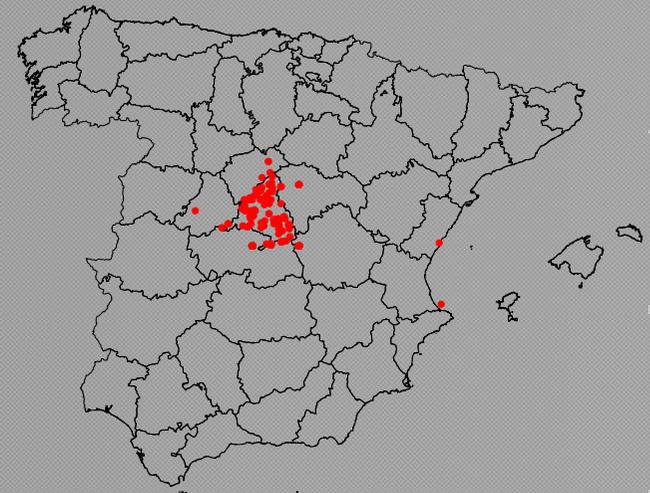
- Fuente: ABH, ARAN, BCN, BIO, COA, FCO, HGI, HUAL, IPE-CSIC, MA, MAF, SALA, SANT y VAL
- Total de registros: 43.613
- Registros georreferenciados: 21.263 (48%)
- Registros fuera del territorio nacional: 1050!!!



# 4 Validación de las coordenadas

⇒ Ejemplo: Provincia de Madrid

- Total de registros: 1353
- Registros georreferenciados: 1062 (48%)
- Registros fuera de la provincia: 147!!!



# Evaluación de métodos de georreferenciación

- Murphey & al., 2004 compararon 4 herramientas:
  - Biogeomancer (Beaman and Conn, 2003), University of Kansas  
(<http://www.biogeomancer.org> )
  - MANIS Georeferencing Calculator, University of California at Berkeley  
(<http://elib.cs.berkeley.edu/manis/gc.html> )
  - GEOLocate, Tulane University  
(<http://www.museum.tulane.edu/geolocate/demo.aspx> )
  - ArcView Georeferencing Extension, California Academy of Sciences  
(<http://www.calacademy.org/research/informatics/georef/> )

# Evaluación de métodos de georreferenciación

Feature/Capability	Biogeomancer	Georeferencing Calculator	GEOLocate	ArcView Georeferencing Extension
Format	Online	Online	Software download	Software download
Georeferencing Tool	Y	N	Y	Y
Error Estimate	N	Y	Y	Y
Error Expression	N	Point-Radius Method: Numerical error in miles	low, medium, or high precision	polygon (visual), and "span" value
Batch Processing	Y (small batches)	N	Y	N
Single Offsets	Y	Y	Y	Y
Multiple Offsets	N	N	Y	N
Abbreviations	N	N	some	N
PLSS	N	N	Y (no_sec)	N
Place Names Database	Y	N	Y	Y (user supplied)
Street Addresses	N	N	Y	N
Highways	N	N	N	N
Elevation	N	N	N	N
Multiple Bearings (SSE, NNW, etc.)	Y	N	N	N
Text Descriptions	N	N	Y	N
Parse/Ignore Capability	N	Y	N	Y
Uses Map Interface	N	N	Y	Y
Average correct on initial georeferencing:	34%	N/A	48%	25%
Total average time per record:	6.25	5 (including manual georeferencing)	4.1	7
Time: Place Name/Coverages	0	0	0	1
Time: Database Preparation	1 (due to smaller batches) 0.25	0.5	0.5	0.5
Time: Georeferencing (coordinates)		3 (manual georeferencing)	0.1	2.5 (essentially manual georeferencing)
Time: Georeferencing (error)	2 (necessary to regraph coordinates)	0.5	2 (necessary to regraph coordinates on topo map in order to find extents)	1
Time: Quality Checking	3 (many incorrect records returned)	1	1.5 (some incorrect records returned)	2 (very difficult to check because of opaque error shapefiles)

# Localidad

---

- Específica
- Sucinta
- Inequívoca
- Completa
- Exacta

# Localidad

---

- Evitar la incertidumbre debido a la imprecisión de las definiciones mediante la descripción de un lugar en términos de una distancia a lo largo de un camino, o por dos distancias desde el lugar
- Utilice sólo un punto de referencia, debe ser de tamaño pequeño, estable en su posición, en su tamaño, en el tiempo, y fácil de encontrar en los mapas o en nomenclátories
- Evitar el uso de términos vagos como "cerca de..."

# Elevación

---

- Proporcionar un valor de elevación
- Cuidado de la elevación del GPS, es menos precisa que altímetro barométrico (cuando calibrado)
- Indicar la fuente de elevación (mapa, altímetro, ...)

# Coordenadas

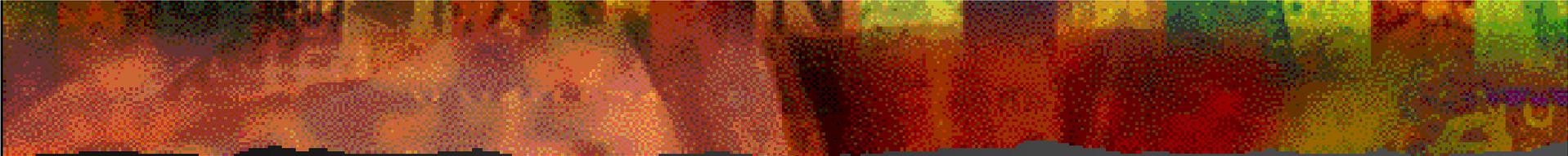
---

- No son suficientes para describir una localidad
- Las coordenadas originales del GPS son mejores que conversiones posteriores
- Los grados decimales pueden ser más precisos que los GG:MM:SS.

# Coordenadas

Precision	0 degrees Latitude	30 degrees Latitude	60 degrees Latitude	85 degrees Latitude
1.0 degree	156,904 m	146,962 m	124,605 m	112,109 m
0.1 degree	15,691 m	14,697 m	12,461 m	11,211 m
0.01 degree	1,570 m	1,470 m	1,246 m	1,121 m
0.001 degree	157 m	147 m	125 m	112 m
0.0001 degree	16 m	15 m	13 m	12 m
0.00001 degree	2 m	2 m	2 m	2 m
1.0 minute	2,615 m	2,450 m	2,077 m	1,869 m
0.1 minute	262 m	245 m	208 m	187 m
0.01 minute	27 m	25 m	21 m	19 m
0.001 minute	3 m	3 m	3 m	2 m
1.0 second	44 m	41 m	35 m	32 m
0.1 second	5 m	5 m	4 m	4 m
0.01 second	1 m	1 m	1 m	1 m

**Table 4.** Table showing metric uncertainty due to precision of coordinates based on the WGS84 datum at varying latitudes. Uncertainty values have been round up in all cases. From [Wieczorek \(2001\)](#).



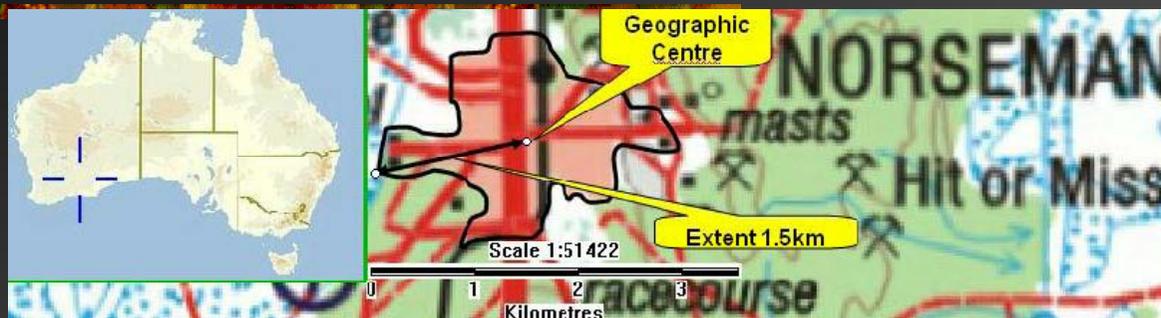
# Guía rápida de Georreferenciación

---

Cristina González Aguado  
Madrid, Réplica del II Taller sobre  
Georreferenciación  
28-30 noviembre 2007

# Lugar citado A) entidad o área urbana

Ejemplo: “Las Vegas”



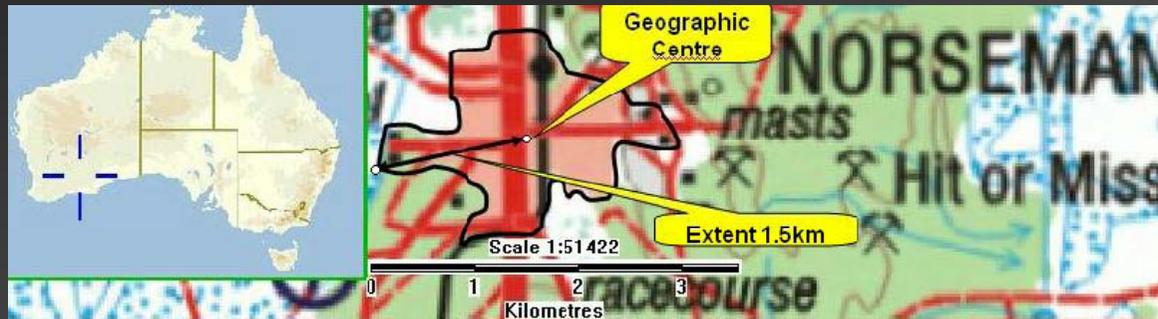
- **Procedimiento:**

Estimaremos el centro geográfico (punto medio entre los valores extremos de longitud y latitud). Si no se puede localizar fácilmente, dibuje un rectángulo que incluya totalmente el lugar citado y trace las diagonales para determinar el centro.

Si éste no está dentro del área urbana, seleccionamos el punto más cercano al centro geográfico que esté dentro del área de la figura.

## Lugar citado B) Lugar remoto

Ejemplo: “*FontLlonga, cerca del Pantano de Camarasa*”

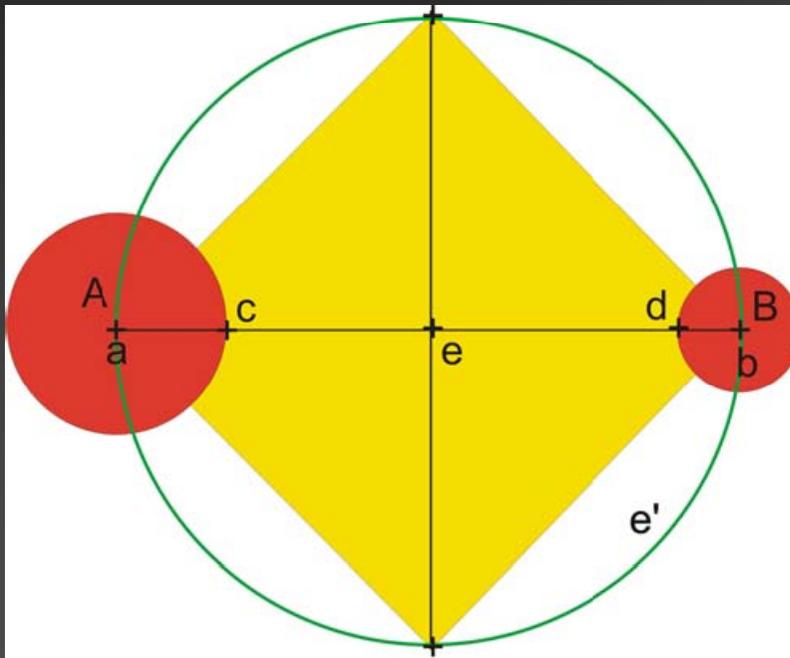


### ■ Procedimiento:

Documéntese y ponga las coordenadas en un punto que juzgue adecuado sobre el mapa.

# Entre dos lugares citados

Ejemplo: “*Entre Llafranc y Calella*”



- Procedimiento:  
Georreferenciar el punto medio entre los dos centros de los dos lugares citados.

# Granja, Masia

Ejemplo: “*Can Pipirimosca*”

---

- **Procedimiento:**

Utilice las coordenadas del edificio que aparezca al mapa, si es que no se ha podido delimitar exactamente su extensión.

---

# Cruce

Ejemplo: *"cruce de las calles Rossellón con Dos de Mayo"*

---

- **Procedimiento:**

**Localice en el mapa los dos ríos o carreteras y obtenga las coordenadas del punto de intersección.**

---

# Lago, montaña o entidad geográfica

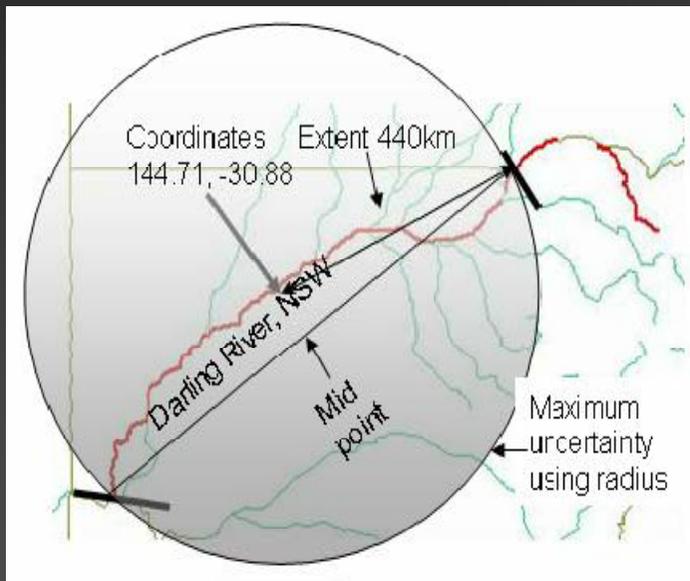
Ejemplo: “*Montserrat*”

- **Procedimiento:**

Sitúe las coordenadas en el centro geográfico. Si no se puede localizar fácilmente, dibuje un rectángulo que incluya totalmente el lugar citado y trace las diagonales para determinar el centro.

# Ríos, calles o caminos

## Ejemplo: “Río Copotaza”

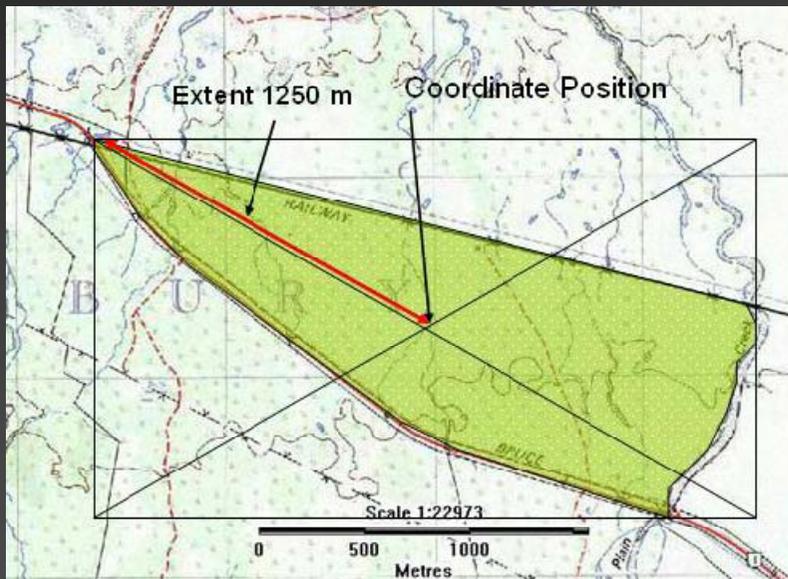


- Procedimiento:  
Dibuje una línea recta desde la desembocadura del río hasta su cabecera. Buscaremos el centro de esta recta y situaremos las coordenadas sobre el río, en el punto más cercano posible a este centro.

# Entre dos caminos, ríos o calles

Ejemplo: “entre el Ebro y el Gállego”

- Procedimiento:  
Realice un polígono entre el punto de unión de los dos ríos o carreteras y sus extremos. Dibuje un rectángulo que incluya totalmente el lugar citado y trace las diagonales para determinar el centro.



## Desembocadura o Cabecera de un río

Ejemplo: “*nacimiento del río Llobregat*”

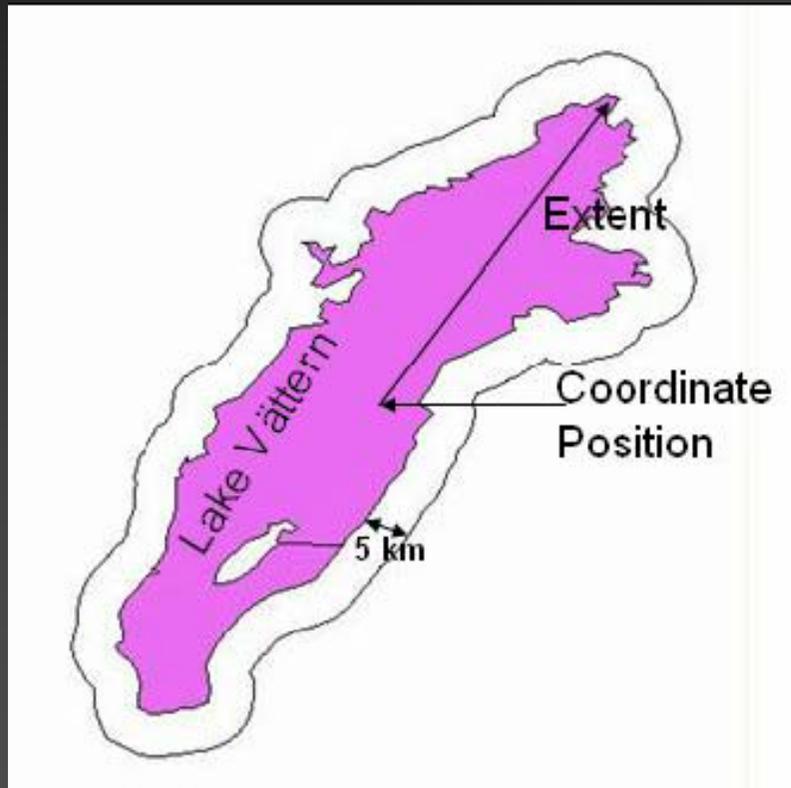
- **Procedimiento:**

*Desembocadura:* georreferenciar donde el cuerpo entra en contacto con un cuerpo de agua más grande.

*Cabecera:* georreferenciar donde nace el río.

# Sólo desplazamiento

Ejemplo: “A 5 Km. de Girona”



- Procedimiento:  
Las coordenadas serán las mismas del centro del lugar citado.

## Desplazamientos a lo largo de un río o carretera en una dirección

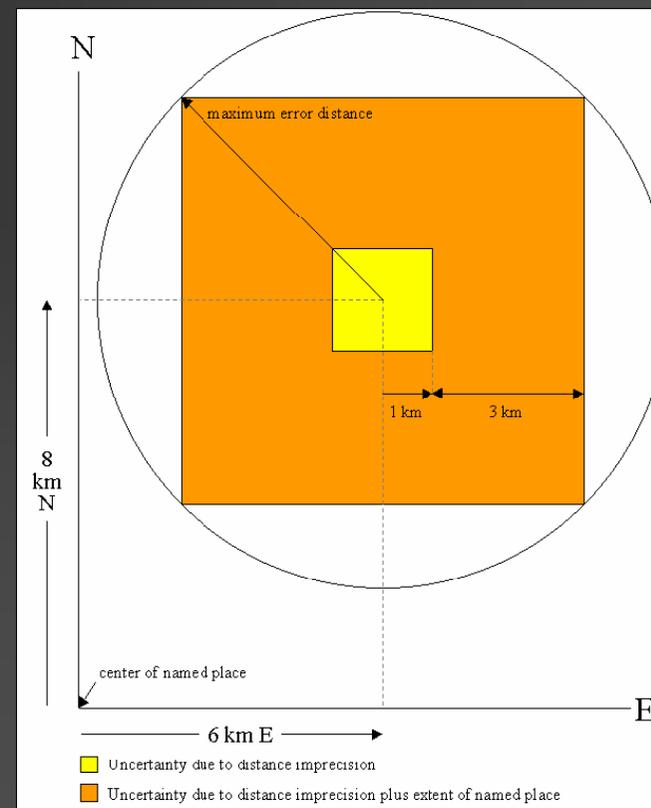
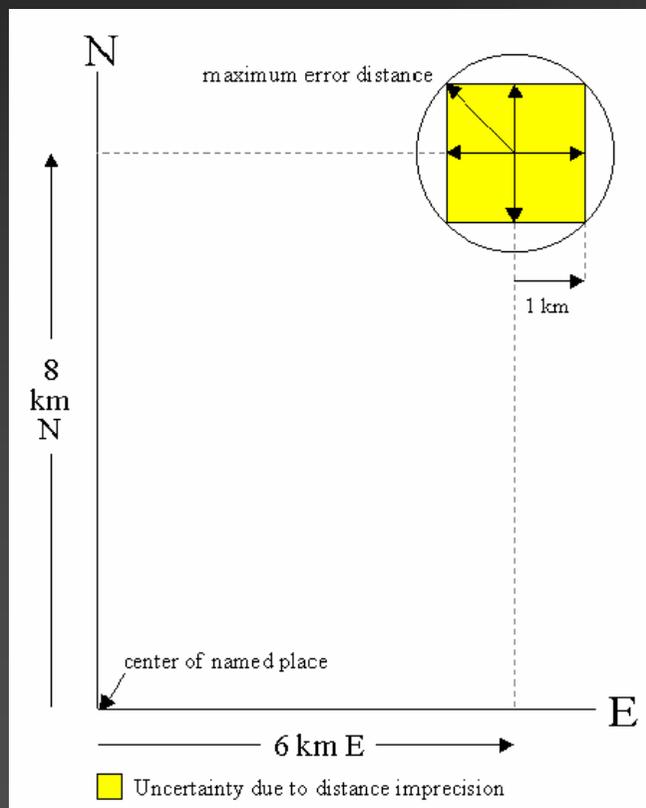
*Ejemplo: “15 Km. al Oeste de Girona por el río Onyar”*

- Procedimiento:

Comience en el centro del punto de partida y utilice la aplicación que resulte más apropiada para reseguir la carretera o río hasta la distancia citada.

# Desplazamientos en direcciones ortogonales

Ejemplo: “2 Km. al Sur y 4 al Este de Tortosa”



# Coordenadas Latitud y Longitud

Ejemplo:  $36^{\circ} 31' 21,4''$  N;  $114^{\circ} 09' 50,6''$

## ■ Procedimiento:

Si la localidad lleva asociadas unas coordenadas coherentes con el resto de la descripción, generalmente se toman como buenas.

# Coordenadas UTM

Ejemplo: “4291492 N; 456156 E”

---

- Procedimiento:

Si no nos dicen el huso, buscaremos la zona mediante información de la descripción (país, provincia..). Nos podemos ayudar utilizando la web:

<http://www.dmap.co.uk/utmworld.htm>

---

# No se puede georreferenciar

- Desplazamiento sólo en una dirección y sin ninguna distancia dada (ex: O de Vic)
- Dudoso (ex: Probablemente Chile central, Isla boca brava?)
- No se puede localizar o es ambiguo (ex: San José, Méjico; Lugar no anotado)

# Casos particulares

---

- Todos los registros no georreferenciados conviene que sean marcados como tales y indicar las causas de la no georreferenciación