



Generalitat de Catalunya
Institut d'Estadística de Catalunya

Sobre la utilización de métodos espaciales en el registro estadístico de territorio.

D. Ibáñez Vidal

dibanez@Idescat.cat

A. Bernaus Darbra

abernaus@Idescat.cat

E. Suñé Luís

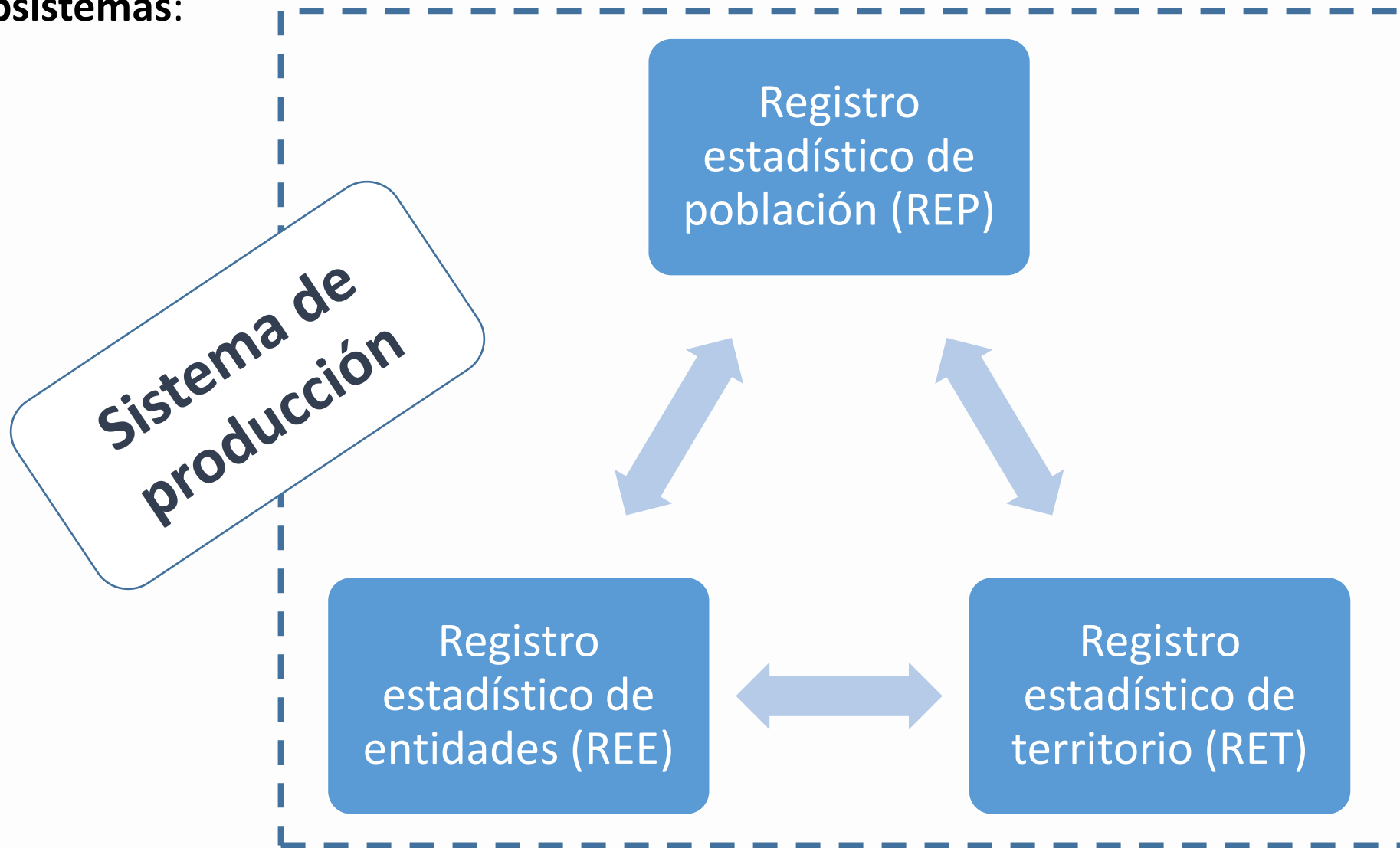
esl@Idescat.cat

M. Farré

mfarre@Idescat.cat

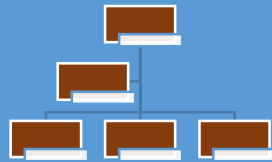
Sistema de producción basado en registros

Construcción de un **sistema de producción** basado en **registros administrativos** formado por **3 subsistemas**:



Registro estadístico de territorio

Registro estadístico de territorio (RET)



REP,REE

Objetivos

- 1 • Geocodificar microdatos
- 2 • Validar direcciones postales

Fuentes

- Direcciones postales del REP y REE. Idescat, INE...
- Portales. Base de datos municipal de direcciones de Cataluña. Instituto Cartográfico y Geológico de Cataluña (ICGC),MDAGE
- Bienes inmuebles. Dirección General del Catastro (DGC)

A2	TVIA	A2	NVIA	A2	RPOB	A2	PRO	A2	MUN	A2	DIS	A2	SEC	A2	CVIA	A2	TINUM	A2	NUMER	A2	CNUMER	A2	NUMERS	A2	CNUMERS	A2	KMT	A2	HMT	A2	BLOQ	A2	PORT	A2	ESCA	A2	PLAN	A2	PUER
PG			GRAN PASSEIG RONDA		20141 25		120		04		015		02980		1		0001		(null)		(null)		(null)		(null)		(null)		(null)		(null)		(null)		P01		0001		
PG			GRAN PASSEIG RONDA		20141 25		120		04		015		02980		1		0001		(null)		(null)		(null)		(null)		(null)		(null)		(null)		(null)		P01		0003		
PG			GRAN PASSEIG RONDA		20141 25		120		04		015		02980		1		0001		(null)		(null)		(null)		(null)		(null)		(null)		(null)		(null)		P02		0001		
PG			GRAN PASSEIG RONDA		20141 25		120		04		015		02980		1		0001		(null)		(null)		(null)		(null)		(null)		(null)		(null)		(null)		P02		0002		

vía

BDMAC, MDAGE

dirección horizontal

1

Obtener la posición {x,y,SRID}

CAT Catastro

dirección completa

2

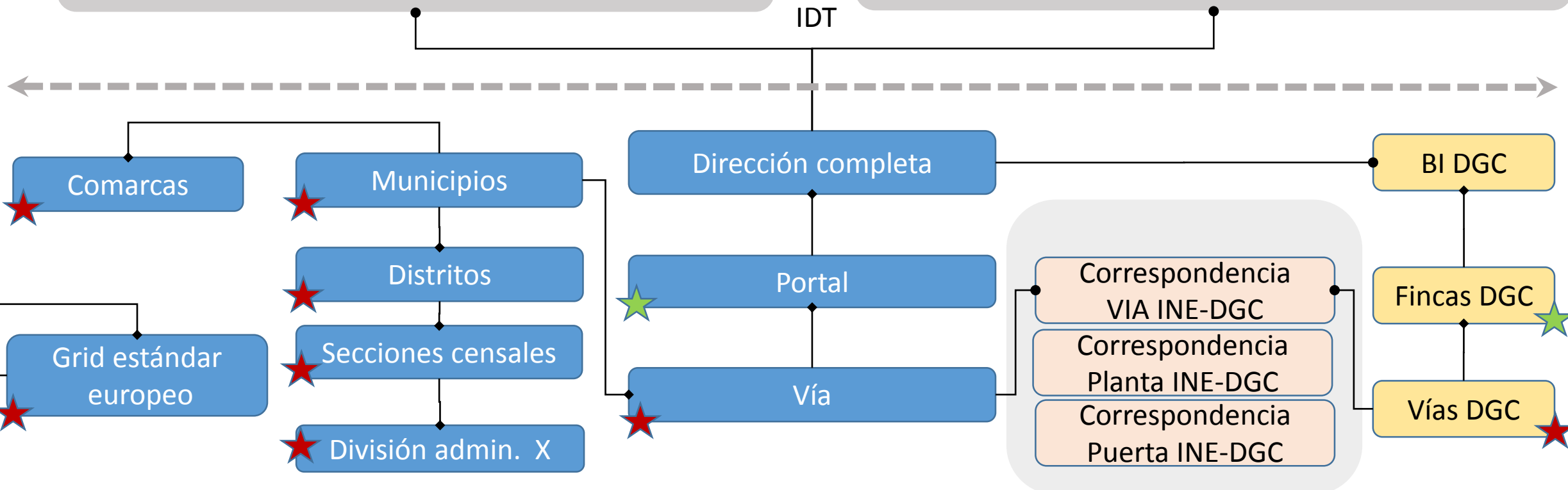
Obtener la referencia catastral

Registro estadístico de territorio

Registro estadístico de población (REP)

Registro estadístico de entidades (REE)

IDT

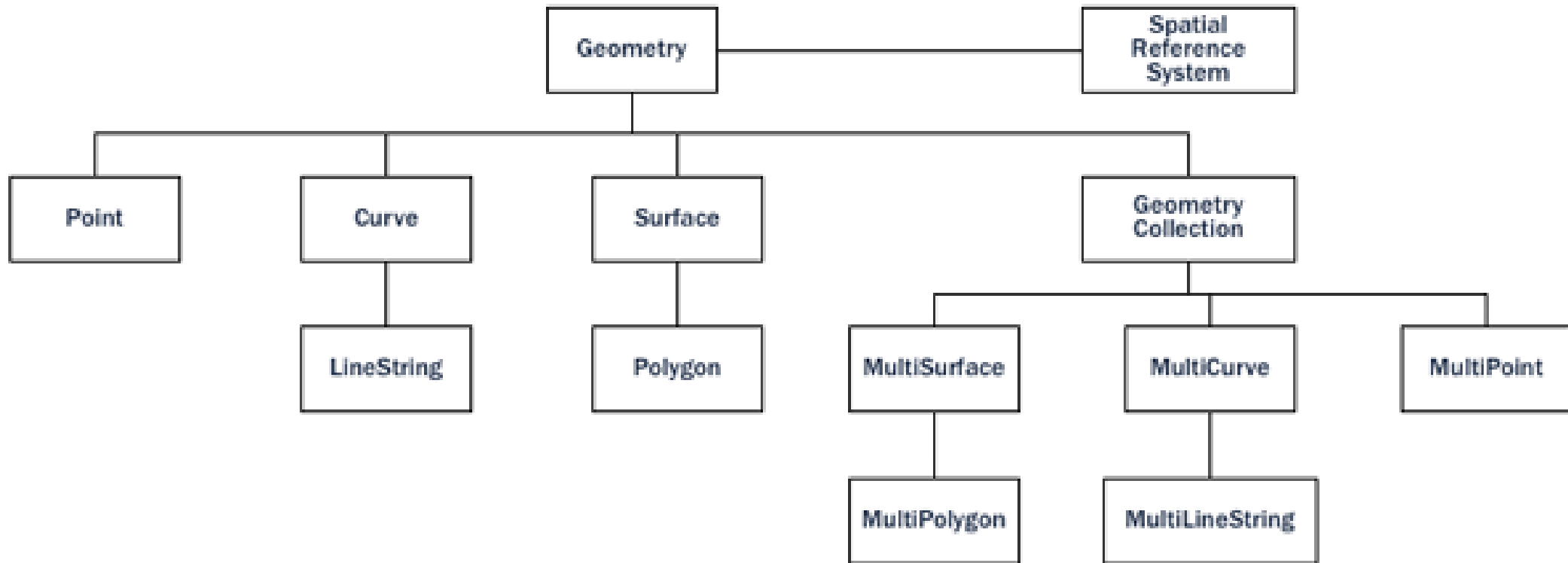


★ Tabla con columna tipo polígono/multi polígono

★ Tabla con columna tipo punto

¡ Es una BD espacial !

Tipos y funciones SFSQL (ejemplos postGIS).



- Constructores → *ST_MakePoint(x,y,SRID), ST_MakeLine(), ST_MakePolygon(), etc..*
- Funciones de acceso → *ST_IsValid(geom), ST_Envelope(geomA), etc..*
- Funciones de edición → *ST_Transform(geom,SRID), ST_SetSRID(geomA,SRID), etc..*
- Funciones de relación y medida → *ST_Within(geomA,geomB), ST_Intersects(geomA,geomB), ST_Touches(geomA,geomB), etc..*
- Funciones de procesamiento → *ST_Intersection(geomA,geomB), ST_Difference(geomA,geomB), ST_ConvexHull(), etc..*
- etc...

RET como geo database.

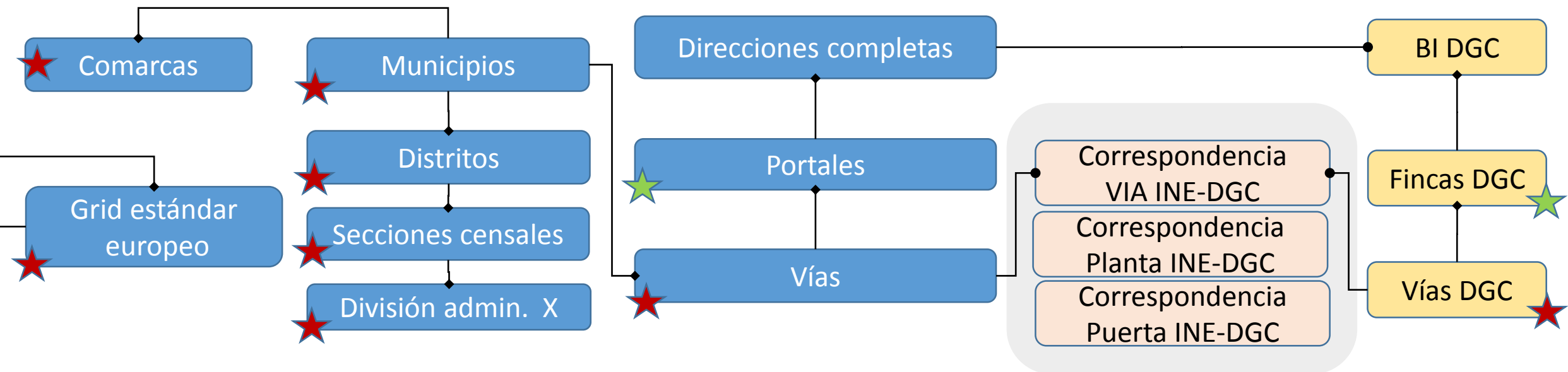
Las BD espaciales tienen meta información relativa a que columnas contienen datos espaciales, sistema de referencia etc.

Disponen de funciones de administración → AddGeometryColumn()

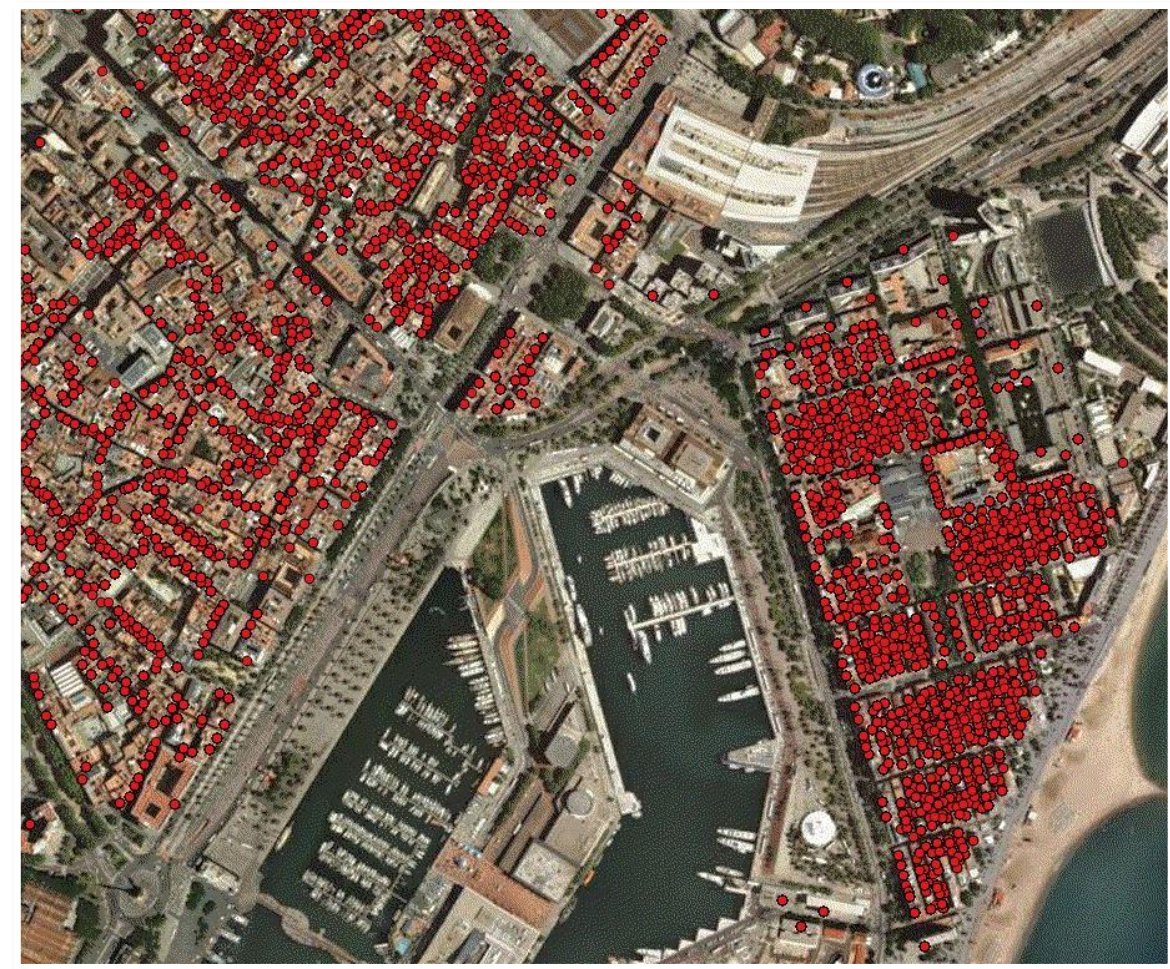
Permiten la indexación mediante las columnas geométricas (R-Trees)

Reference System
25831 UTM31N ETRS89
.....

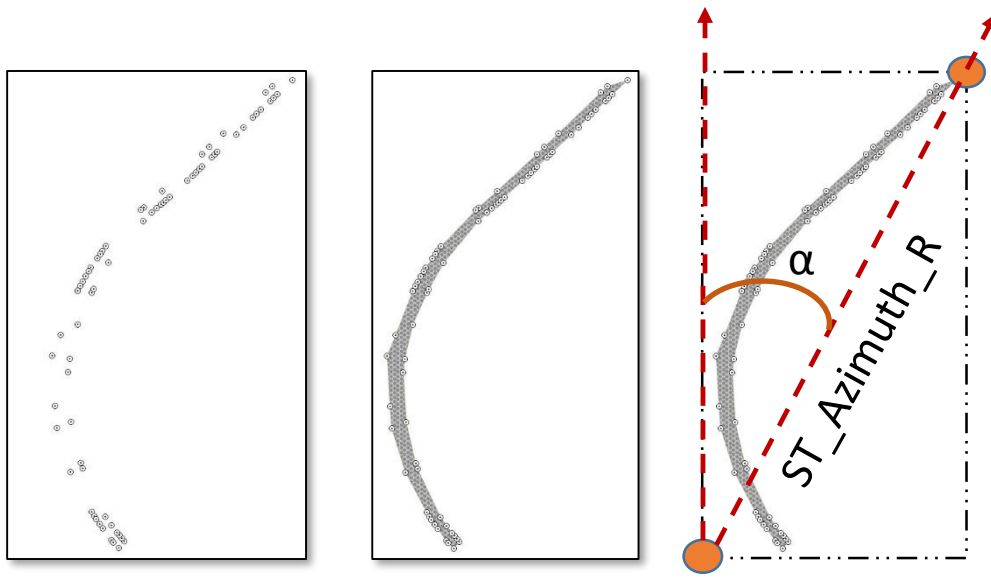
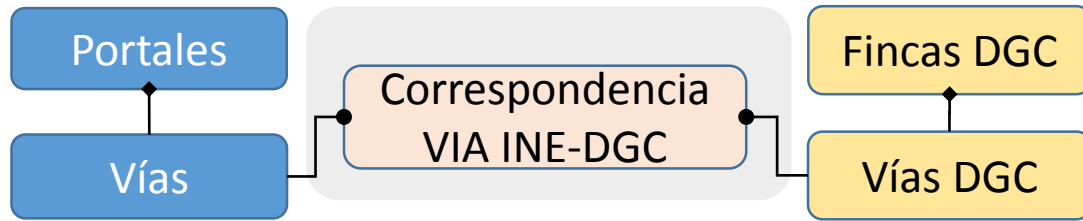
Geometry Columns
Portales, posición, POINT, 25831
Secciones censales, MULTIPOLYGON, 25831
.....



Usos de la geo database: visualización



Usos de la geo database: tabla de equivalencias vía INE- DGC



Si las dos geometrías representan la misma entidad física deben tener una elevadísima intersección y una dirección similar.

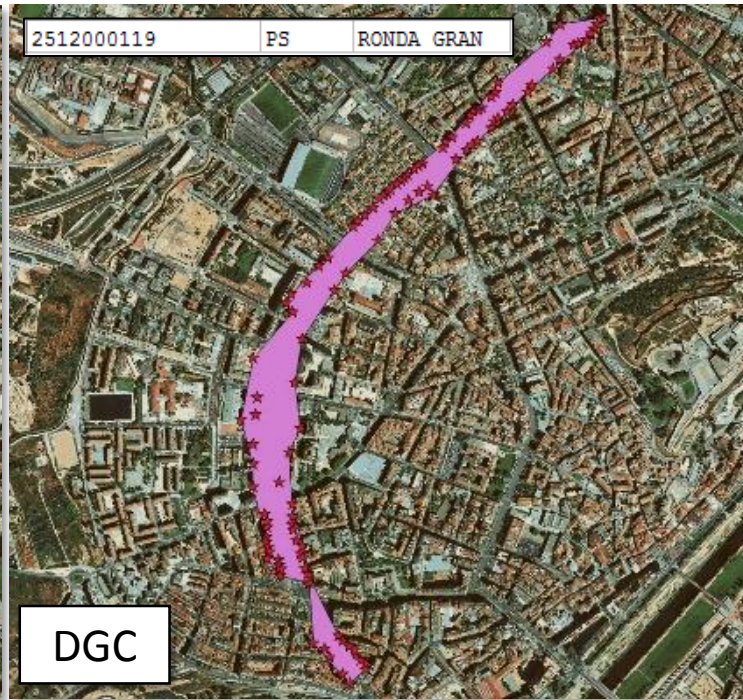
$$\frac{ST_Area(ST_Intersection(A,B))}{ST_Area(A)} \approx 1$$

$$ST_Azimuth_R(A) \approx ST_Azimuth_R(B)$$

El uso de métricas entre nombres de vías es inseguro (falsos positivos)

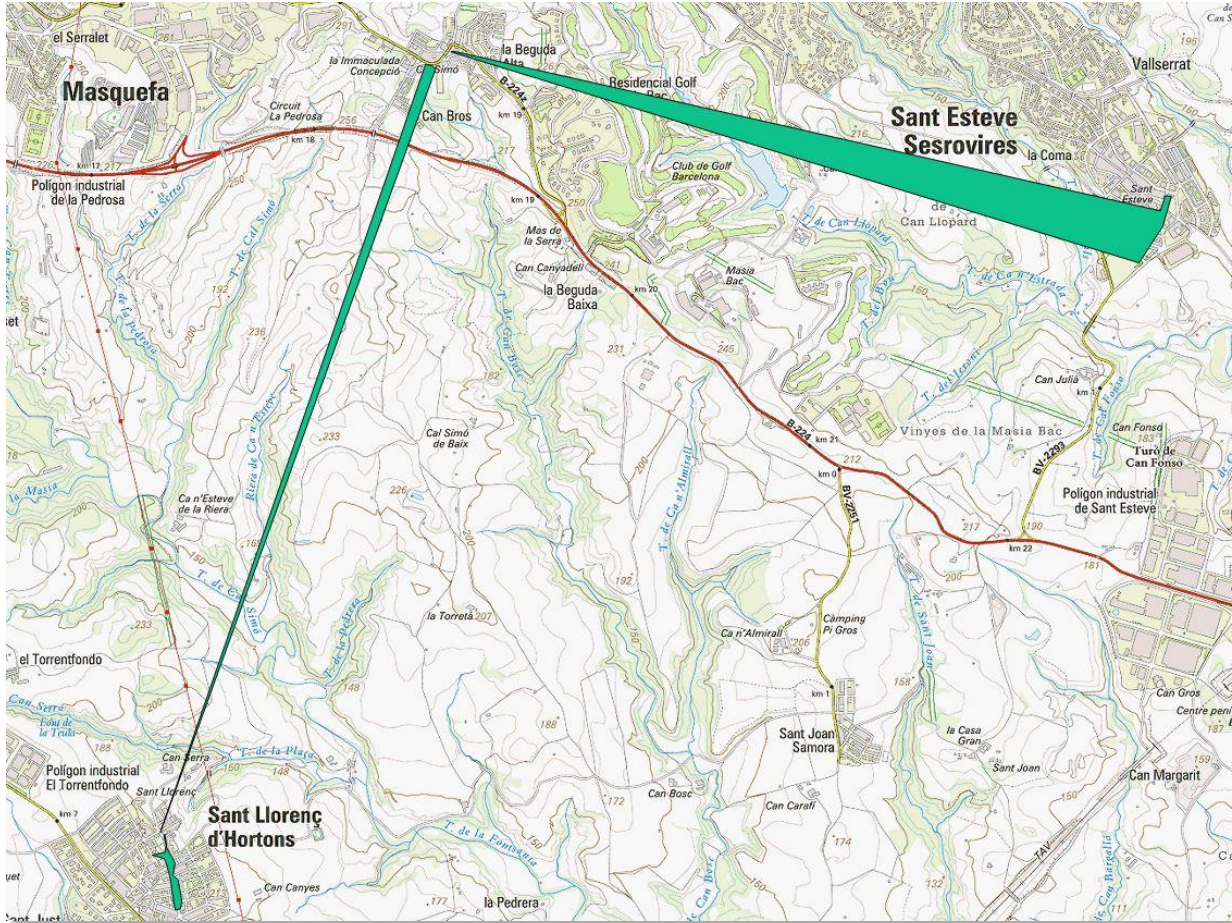
INE	DGC	JARO_S	EDIT_S
GRAN PASSEIG RONDA	SEGON DE RONDA	64	45
GRAN PASSEIG RONDA	RONDA GRAN	62	34
GRAN PASSEIG RONDA	DE LA TORRE PATRIOT	52	11
GRAN PASSEIG RONDA	RONDA	47	28

Construir dos conjuntos de geometrías mediante la función ST_ConcaveHull, una para INE otra para DGC para cada una de las vías



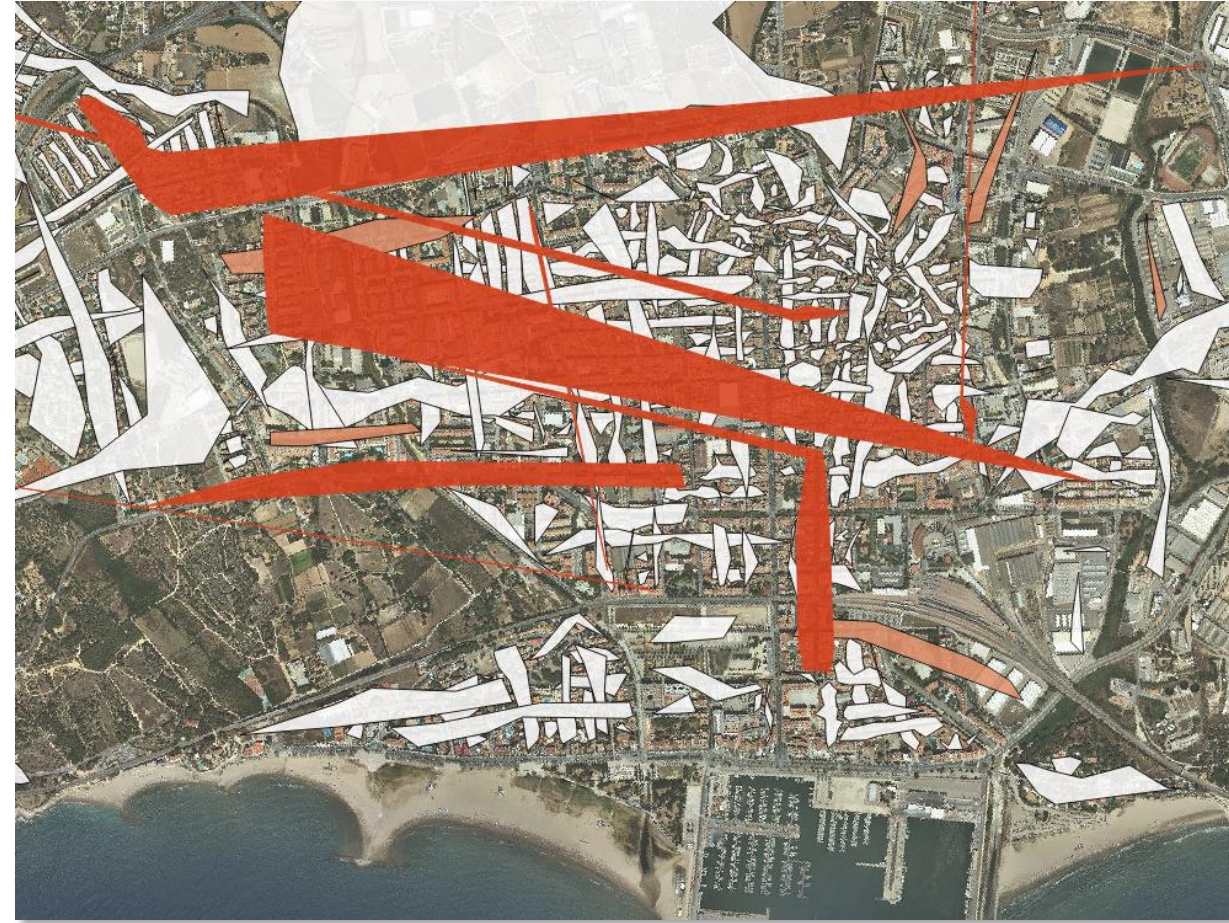
Usos de la geo database: validación coordenadas de portales (I)

La obtención de las envolventes cóncavas pone de manifiesto incoherencias entre las posiciones y los códigos de vía.



Se definió un indicador de calidad de los polígonos:

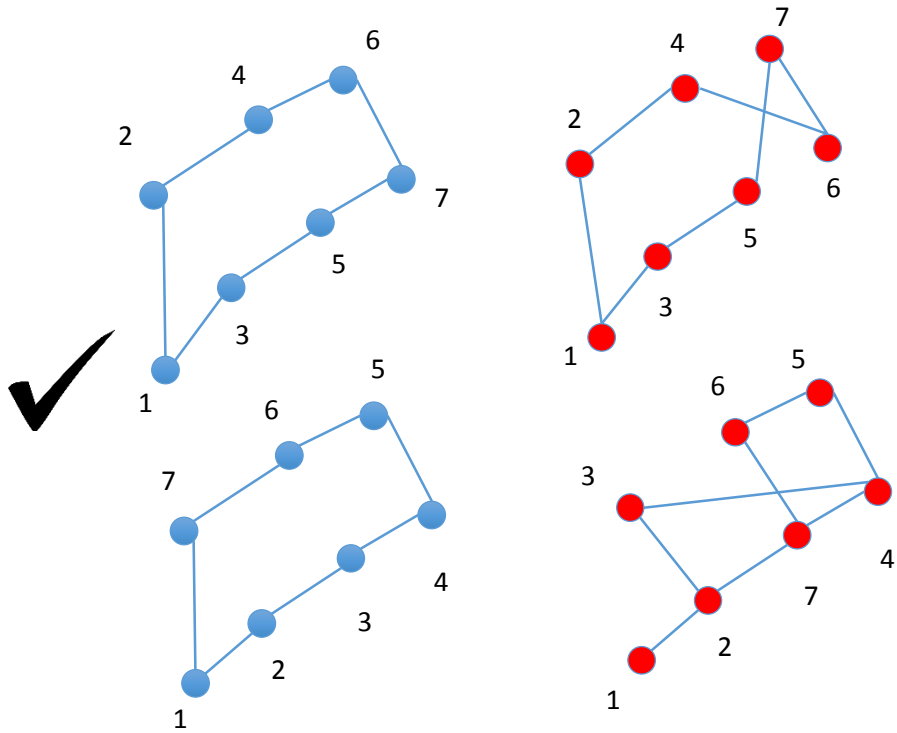
$$ICP_p = CV_p^2 \cdot \max\{longitud_i - longitud_{i+1}\}_{i=1 \dots v_p-1}$$



El grupo de polígonos con **más probabilidad de ser erróneos** es de un **2,9%** del total de la base de datos de fincas (DGC) de Catalunya, lo que conlleva que el número de puntos anómalos sea aún menor.

Usos de la geo database: validación coordenadas de portales (II)

Otra forma de validar: construir polígonos uniendo puntos según la numeración de los portales



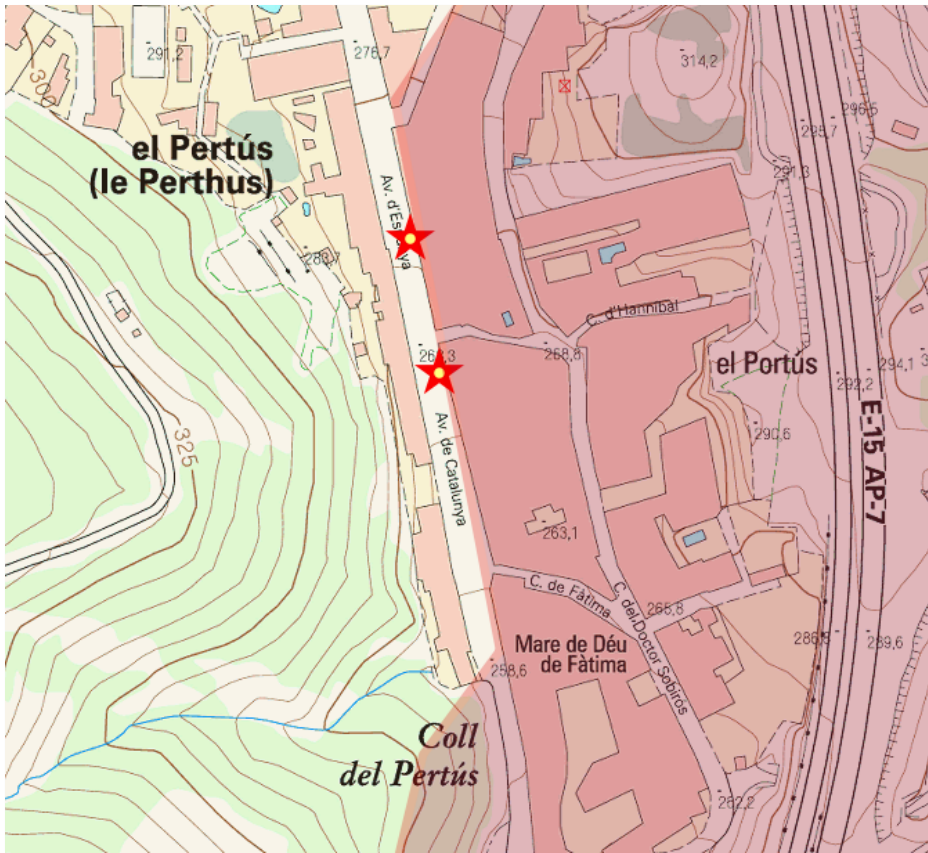
Las funciones `ST_IsValid()` y `ST_IsValidDetail()` permiten detectar auto-intersecciones y donde se producen.

Este caso es correcto.

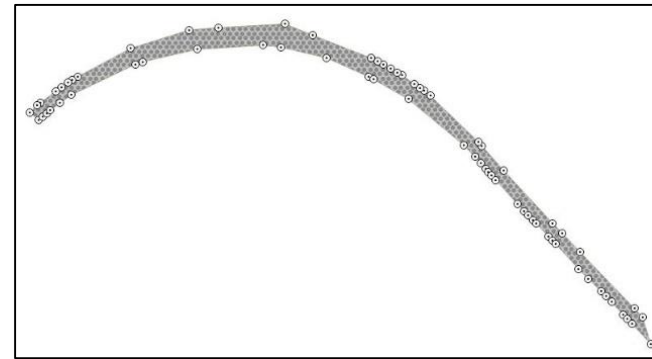


Usos de la geo database: validación coordenadas de portales (III)

PRO	MUN	DIS	SEC	CVIA	TVIA	NVIA	NUMER	CNUMER	NUMERS	CNUMERS
17	086	01	001	00076	AV	CATALUNYA	0005	(null)	(null)	(null)
17	086	01	001	00076	AV	CATALUNYA	0008	(null)	(null)	(null)
17	086	01	001	00076	AV	CATALUNYA	0015	(null)	(null)	(null)
17	086	01	001	00076	AV	CATALUNYA	0021	(null)	(null)	(null)

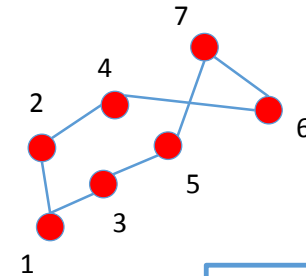


ST_Within(A,B)



ST_ConcaveHull({p})

$$ICP_p = CV_p^2 \cdot \max\{longitud_i - longitud_{i+1}\}_{i=1 \dots v_p-1}$$



ST_IsValid(P)

(x,y,SRID)

...
...

Usos de la geo database: creación de nuevas geometrías

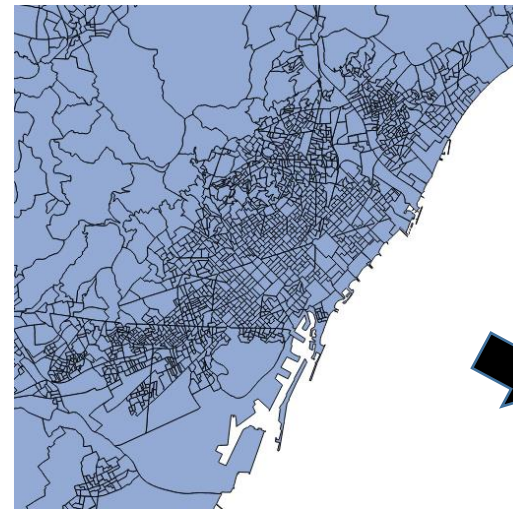
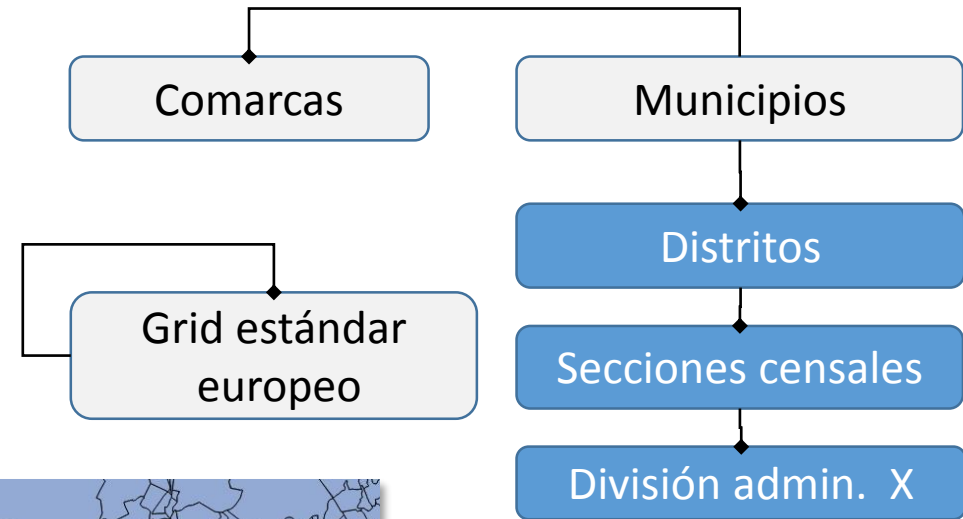
municipi character varying(6)	districte character varying(2)	seccio character varying(3)	mundissec character varying(11)	geom geometry(MultiPolygon,25831)
080193	01	001	08019301001	0106000020E76400000100000001030000000
080193	01	002	08019301002	0106000020E76400000100000001030000000
080193	01	003	08019301003	0106000020E76400000100000001030000000
080193	01	004	08019301004	0106000020E76400000100000001030000000
080193	01	005	08019301005	0106000020E76400000100000001030000000
080193	01	006	08019301006	0106000020E76400000100000001030000000
080193	01	007	08019301007	0106000020E76400000100000001030000000
080193	01	008	08019301008	0106000020E76400000100000001030000000
080193	01	009	08019301009	0106000020E76400000100000001030000000

```
select municipi,districte,ST_Union(geom) from seccions2016 where
municipi='080193' group by municipi,districte
```



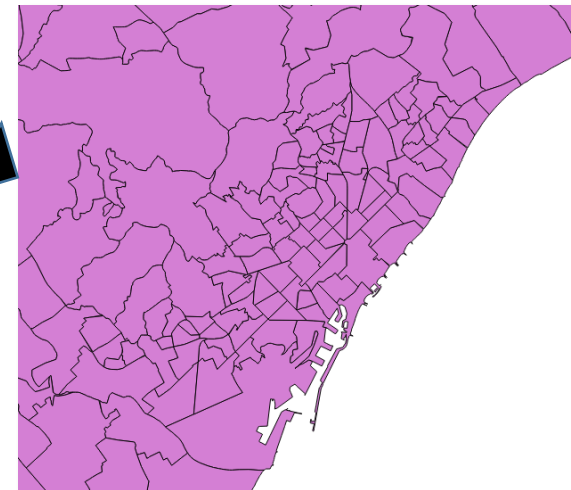
municipi character varying(6)	districte character varying(2)	st_union geometry
080193	09	0103000020E7640000010000000F700000040C6DC35275A1A4144696F68017E5141906588239F581A4
080193	04	0103000020E76400000100000057010000E04F8D17920B1A4114AE474988795141E09387855D0A1A4
080193	05	0106000020E7640000030000000103000000010000002700000000BC055266CB19418CB96BB9F97D!
080193	03	0106000020E76400000200000001030000000100000073000000060984CD5F3581A41C64B37C11278!
080193	01	0103000020E764000001000000B2010000E06800EF5B661A41F831E69ED179514170711B0DF4651A4
080193	08	0103000020E764000001000000700100008036AB7E4D5B1A419E80262E2A7F51418036AB7E4F5B1A4
080193	02	0103000020E764000001000000B50000003033333313421A41E2C798035D7951418095438BCC411A4
080193	06	0103000020E7640000010000005901000003D9BD5B63A1A412CF697C5B77B5141A01A2F9D643A1A4
080193	07	0103000020E7640000010000004602000020C9E53F78481A4126C286B7ED7C5141205A647BC7461A4
080193	10	0106000020E7640000040000000103000000010000002900000040795868E9941A41FE65F714EC7C!

```
select b.área_s,ST_Union(a.geom) from seccions2016 a,sección_área b
where a.sección=b.sección group by b.area
```



Secciones censales

sección-área sanitaria



Áreas sanitarias

Usos de la geo database: validación y variación temporal del seccionado

Validación

- No pueden existir intersecciones entre secciones

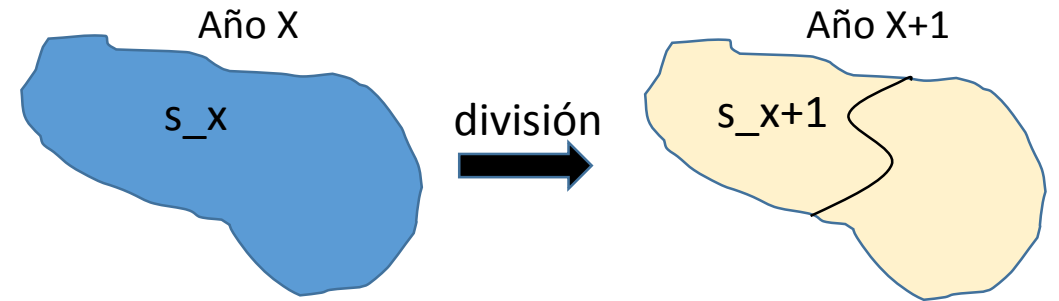
```
select sum(ST_Area(ST_Intersection(a.geom,b.geom))) from secciones2016 a,secciones2016 b where a.mundisec <> b.mundisec = 0
```

- La unión de las secciones tiene que ser todo el territorio

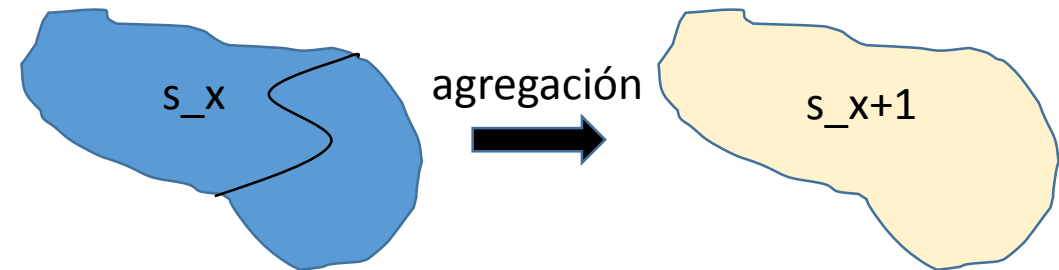
```
select ST_Union(geom) from secciones = Todo_El_territorio
```

¿Cual es la historia de cambios en el seccionado?

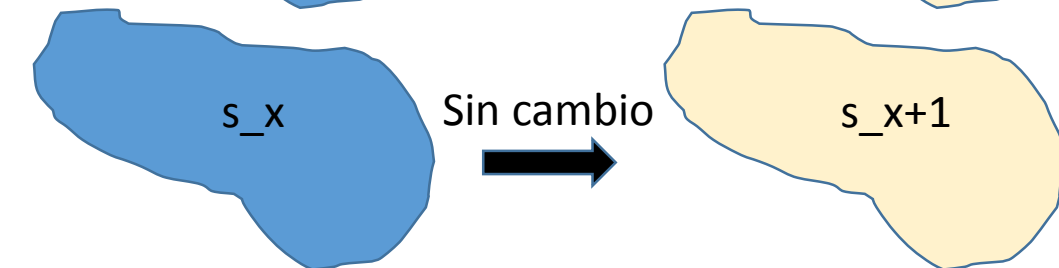
$ST_Intersection(s_x, s_{x+1}) = s_{x+1}$ and $ST_Difference(s_x, s_{x+1}) \neq null$



$ST_Intersection(s_x, s_{x+1}) = s_x$ and $ST_Difference(s_x, s_{x+1}) \neq null$



$ST_Equals(s_x, s_{x+1}) = true$



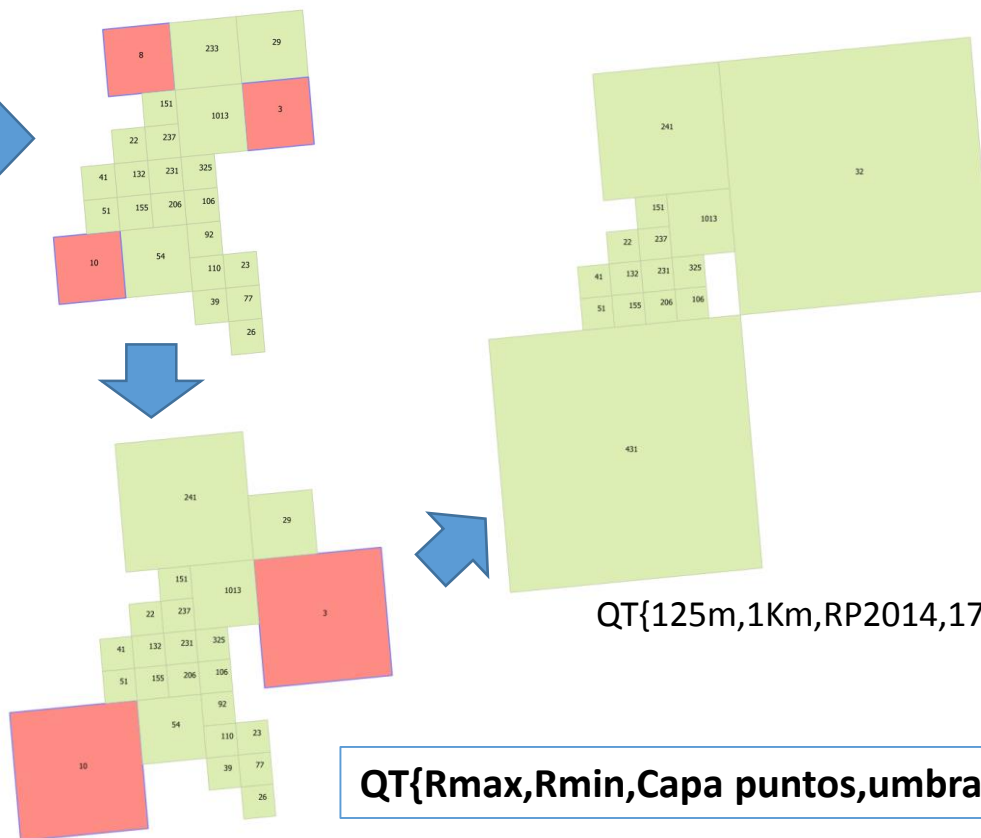
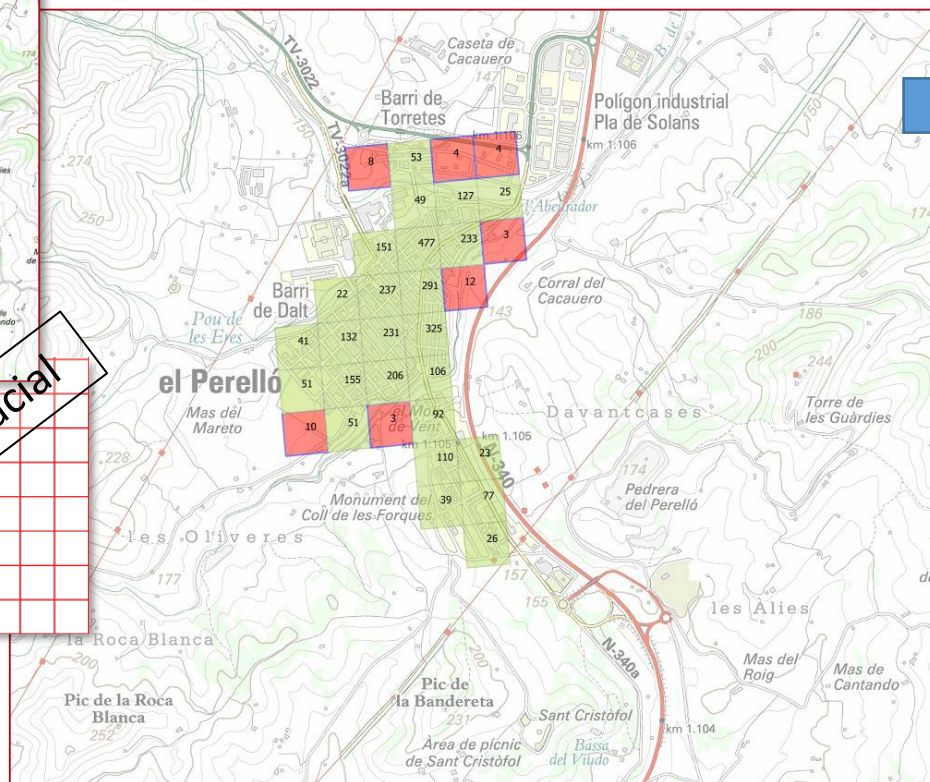
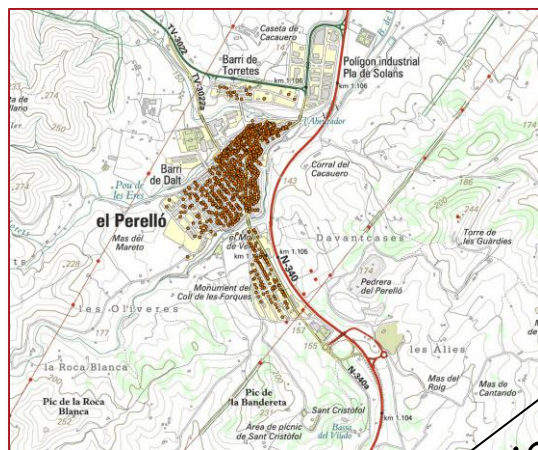
Usos de la geo database: difusión mediante quadrees

Grid estándar europeo

1 Km.

500 m.

250 m.



Join espacial

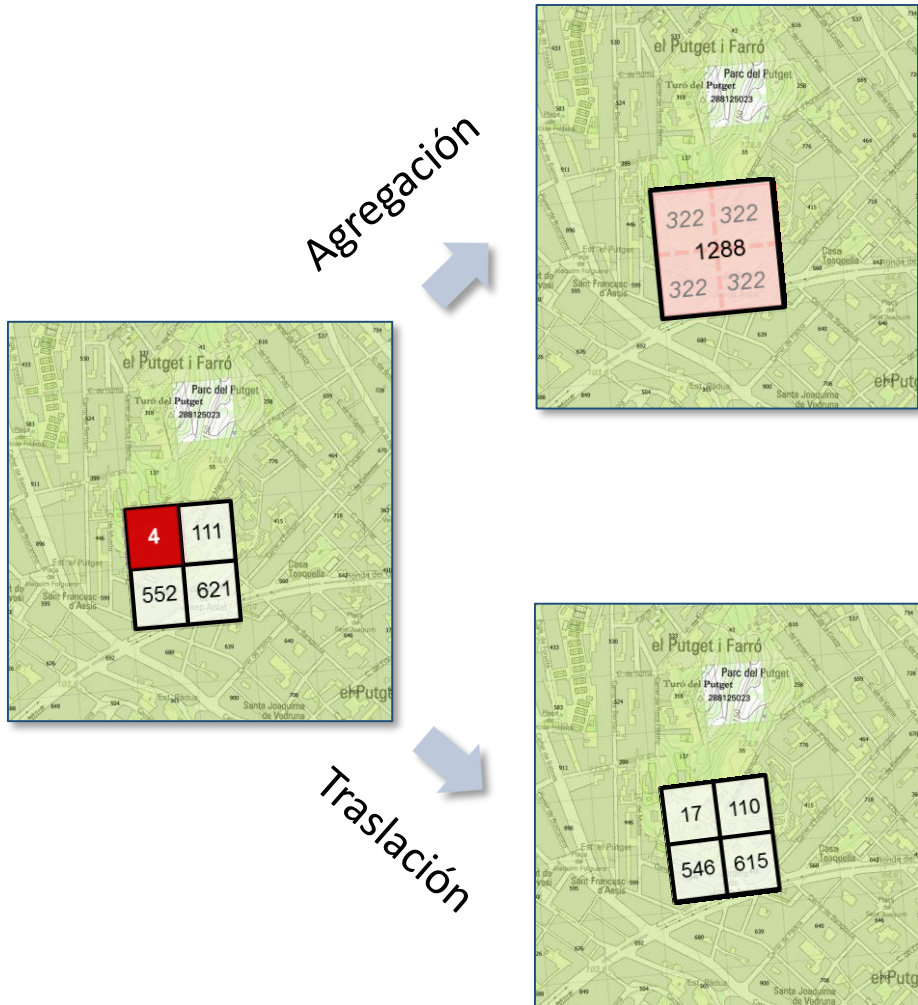
ST_Within(a,b)

QT{125m,1Km,RP2014,17}

QT{Rmax,Rmin,Capa puntos,umbral}

Quadrees. Efecto frontera

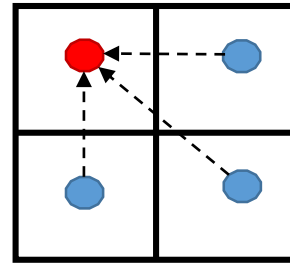
Efecto frontera: agregaciones producidas por altos valores de la varianza



Error absoluto cuando agregamos

$$\epsilon = \sum |n_i - \mu|$$

Pueden evitarse cuando el error absoluto en la traslación es menor que en la agregación



Error absoluto cuando trasladamos

$$\epsilon' = \sum |n_i - n_i'|$$

Cuadrados de	% población	
	Agregación	Traslación
250 m	15.97	3.52
125 m	84.03	96.48
Total	100	100

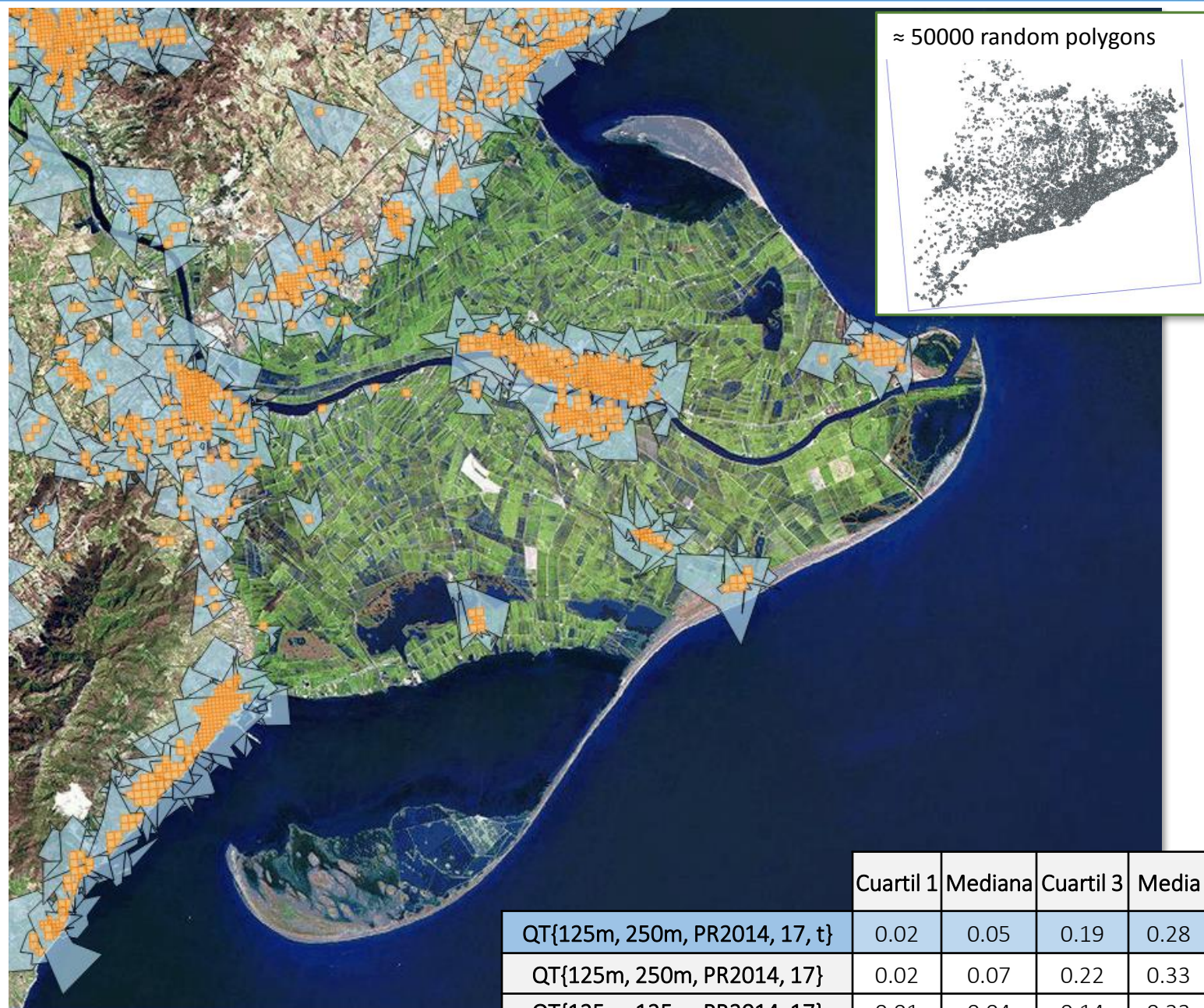
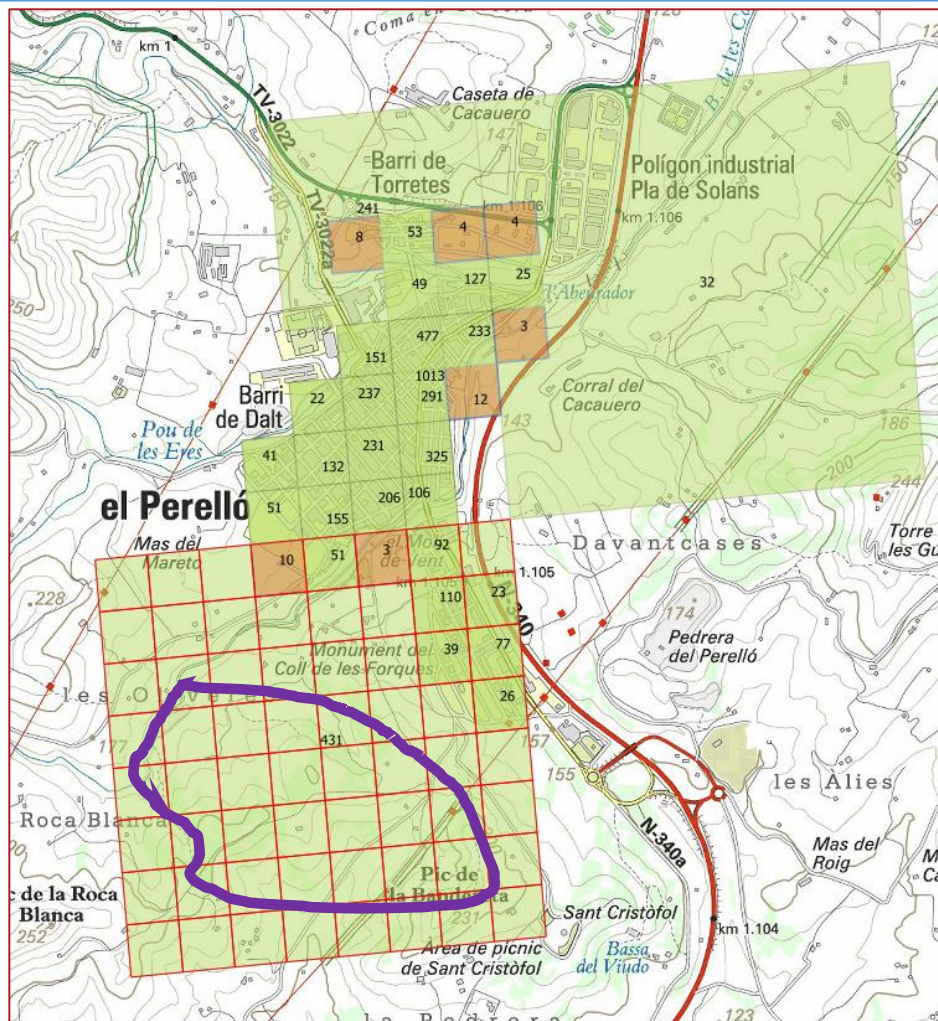
Solo el **0.85%** del total de la población ha sido trasladada

Si $\epsilon' < \epsilon$ es mejor trasladar

Nuevo quadtree con traslaciones:

QT{125m,250m,PR2014,17,t}

Usos de la geo database: estimación de errores



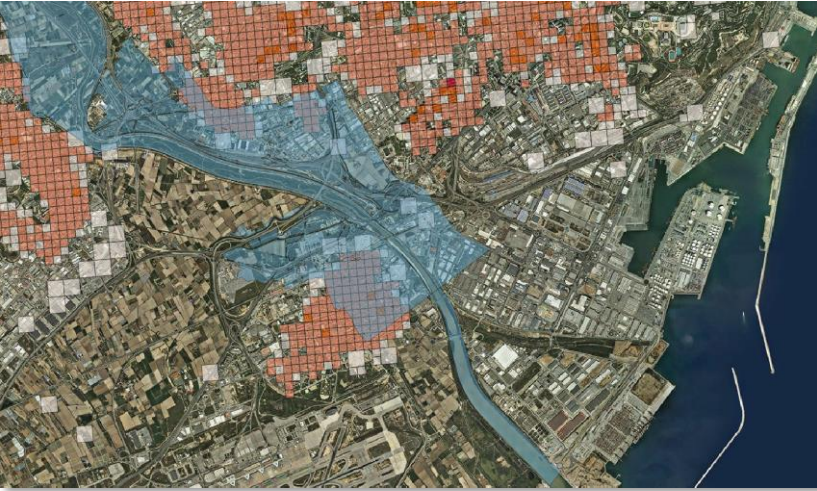
$$\epsilon_i = \frac{|p'_i - p_i|}{p_i}$$

- ① p_x = efectivos dentro de la geometría S_x
- ② $p'_x = \sum p_i * ST_AREA(Q_i \cap S_x) / ST_AREA(Q_i)$

	Cuartil 1	Mediana	Cuartil 3	Media
QT{125m, 250m, PR2014, 17, t}	0.02	0.05	0.19	0.28
QT{125m, 250m, PR2014, 17}	0.02	0.07	0.22	0.33
QT{125m, 125m, PR2014, 17}	0.01	0.04	0.14	0.23

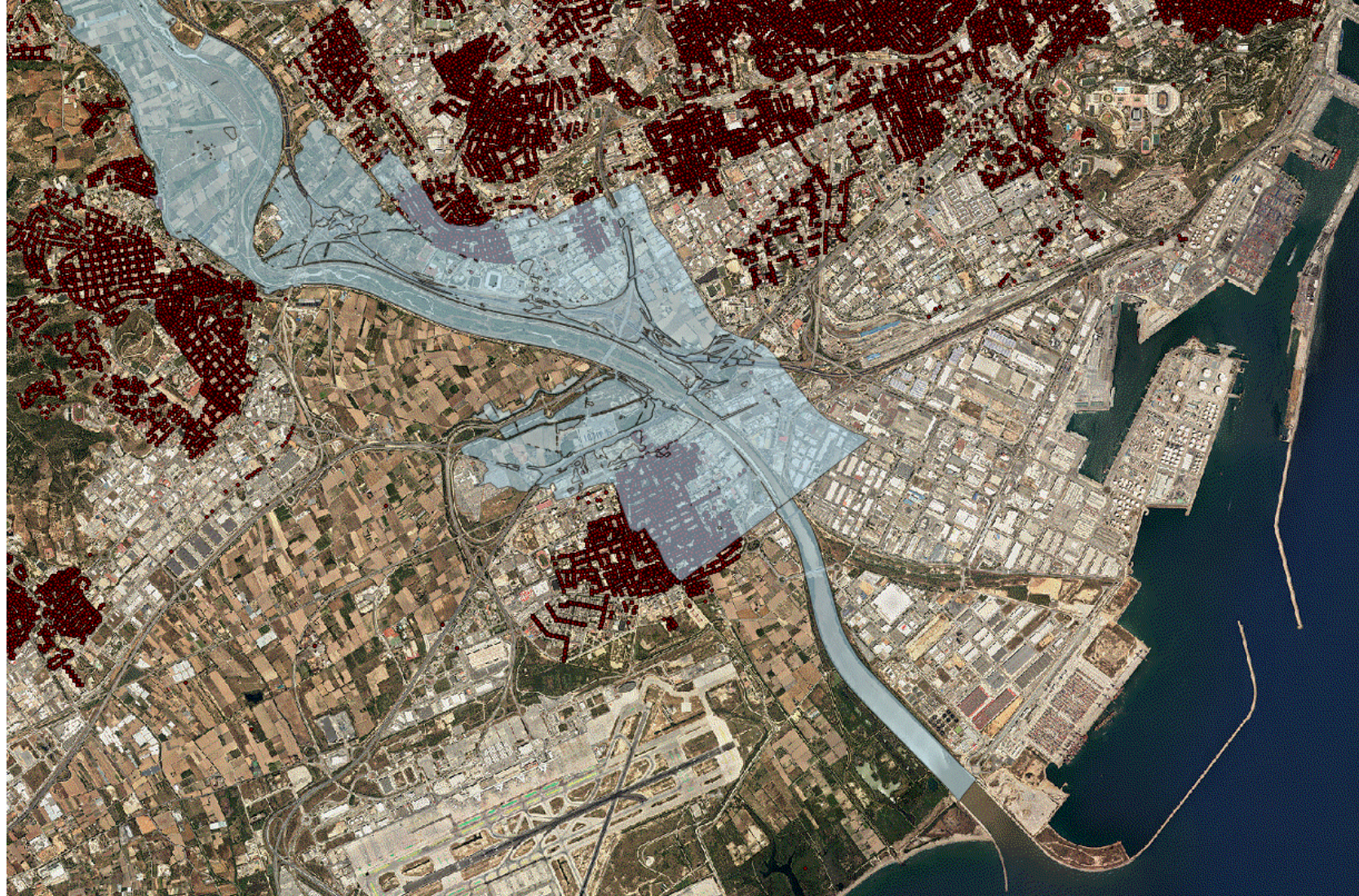
Usos de la geo database: cálculos ad-hoc

Si los cálculos son críticos y no es asumible un 5% de error relativo, el usuario puede proporcionarnos su geometría y Idescat realizará los cálculos sobre la capa de puntos y no sobre el Quadtree.



Ejemplos:

- ❖ Población en zonas de riesgo de inundación.
- ❖ Transporte de mercancías peligrosas.
- ❖ Posicionamiento de alarmas en zonas de riesgo químico.
- ❖ Estudios de contaminación atmosférica





Generalitat de Catalunya
Institut d'Estadística de Catalunya

iGracias!