

Aceleración de servicios de mapas teselados mediante la aplicación de un modelo descriptivo

García, Ricardo; de Castro, Juan Pablo

Resumen

La gran proliferación en el uso de servicios de mapas a través de la Web ha motivado la necesidad de disponer de servicios cada vez más escalables en las Infraestructuras de Datos Espaciales (IDE). Los servicios de mapas basados en teselado se han perfilado como una alternativa escalable frente a los servicios de mapas tradicionales, permitiendo la actuación de mecanismos de *cache* o incluso la prestación del servicio mediante una colección de imágenes pregeneradas. Sin embargo, los requisitos de almacenamiento y tiempo de puesta en marcha de estos servicios resultan prohibitivos para muchos potenciales proveedores cuando la cartografía a servir cubre una zona geográfica extensa para un elevado número de escalas. En este trabajo se propone un modelo descriptivo, basado en el estudio de peticiones realizadas sobre diversos servicios de mapas, para anticipar qué zonas serán solicitadas con mayor probabilidad en un futuro y situar dichas teselas en *cache* para ser servidas más rápidamente.

Dada la naturaleza exponencial de la pirámide de escalas, no resulta práctico trabajar con las estadísticas de teselas individuales. Por ello, se propone un modelo simplificado para la propagación probabilística de los objetos a un nivel abordable. Se comparan las proporciones de recursos para la generación de una *cache* parcial utilizando el modelo propuesto, frente a aquellos para una *cache* conteniendo la cartografía completa. Los resultados obtenidos muestran que pueden conseguirse grandes ahorros de recursos y mantener altas tasas de acierto en *cache* mediante el uso de algoritmos de gestión de *cache* que contemplen este modelo.

PALABRAS CLAVE

Teselado, WMTS, IDEE, WMS, modelo descriptivo, *logs*, *proxy cache*.

1. INTRODUCCIÓN

Las diversas especificaciones WMS (*Web Map Service*) del OGC ofrecen una gran flexibilidad en el servicio. Los parámetros espaciales de las peticiones no están restringidos, lo que hace que cada petición de mapa deba ser atendida en tiempo real mediante un procedimiento, generalmente costoso, que implica acceso a datos de origen, aplicación de estilos, composición de capas y codificación de la imagen comprimida.

Este procedimiento se ha demostrado ineficaz para satisfacer la demanda de algunas aplicaciones de difusión masiva como se expone en [1] tras la experiencia del servidor OnEarth de la NASA. Por este motivo, generalmente los servicios comerciales más populares se prestan con servidores no OGC en los que el espacio geográfico está teselado de acuerdo a una rejilla predefinida y cuyo contenido está frecuentemente pregenerado [2].

La popularidad y necesidad de esta estrategia ha provocado la aparición de algunos mecanismos de acceso no estándar [3-5], y otras recomendaciones más abiertas como la recomendación de WMS-C (*Web Map Service - Cached*) de OSGeo [6] o el reciente estándar WMTS (*Web Map Tile Service*) del propio OGC [7].

Cuando un servicio OGC se va a utilizar en un escenario exigente con una información poco parametrizable y estacionaria, se puede utilizar el patrón *proxy web cache* para conseguir una mejora en la calidad de servicio. El *proxy* es un dispositivo que se interpone de forma preferiblemente transparente entre el cliente y el servicio final, interceptando las peticiones de los usuarios [8].

Al adoptar esta aproximación, los proveedores tienen que afrontar serias decisiones relativas al diseño y mantenimiento de sus *caches* de teselas. Una opción es pregenerar todas las posibles teselas en todas las escalas soportadas. Hoy en día, tan sólo las grandes corporaciones que ofrecen estos servicios disponen de las ingentes cantidades de espacio de almacenamiento necesarias para poder permitirse el tener pregenerados todos los objetos. No supone un problema para estas entidades (sí para otras más humildes y de menor presupuesto) el discernir qué contenido debe ser pregenerado y cuál no.

En cualquier caso, incluso contando con espacio de almacenamiento suficiente, hay capas cartográficas que deben actualizarse eventualmente y todas deben, en algún momento, partir de cero o renovar su base de teselas. En general, esta estrategia de pregenerar todos los objetos no es adecuada para servir mapas que se actualizan con mucha frecuencia, como es el caso de los mapas meteorológicos o aquellos que muestran información de tráfico rodado, por citar algunos ejemplos.

Estos escenarios requieren una nueva aproximación basada en una estrategia que equilibre cuidadosamente los procesos de población y mantenimiento de la *cache*. Estos problemas de optimización de *cache* se resuelven con frecuencia mediante heurísticas como las siguientes: (1) monitorizar la popularidad de teselas individuales, y (2)

introducir en *cache* las teselas más solicitadas. Esta aproximación presenta, sin embargo, los siguientes inconvenientes. En primer lugar, presupone que los usuarios tolerarán un peor rendimiento mientras se recogen las estadísticas de las peticiones. En segundo lugar, asume que los patrones de acceso de los usuarios son invariantes con el tiempo. En el caso de que estas suposiciones no sean válidas, pueden conducir a disminuciones en el rendimiento que podrían no ser tolerables en función de las necesidades del servicio [9].

En este artículo se presenta un modelo descriptivo, basado en el estudio de peticiones realizadas a distintos servidores de mapas, para predecir qué regiones de mapa serán solicitadas con mayor probabilidad por los usuarios y situar estos objetos en la *cache* para ser servidos más rápidamente. Para ello, se presenta en primer lugar una caracterización del espacio de teselas y se formalizan algunos conceptos. A continuación se describen brevemente los servicios de mapas analizados y la herramienta desarrollada para la extracción de los datos de las peticiones. Posteriormente se presenta el modelo simplificado propuesto y los resultados obtenidos mediante el mismo. Por último se recogen las principales conclusiones de este trabajo y las líneas futuras.

2. ESPACIO DE TESELAS

En cualquier instante se puede definir un estado para una *cache* en la que los objetos que gestiona pueden estar disponibles con una cierta probabilidad.

Como los objetos de una *cache* espacial se pueden identificar por sus coordenadas, denominaremos a las teselas como $T(i, j, n)$, con $i, j, n \in \mathbb{N}$ (donde n es el nivel en la pirámide de escalas e i, j son los índices que sitúan dicha tesela en el plano n). Así mismo, denominaremos $P_{a,b}(T(i, j, n), t) \in [0, 1]$ a la probabilidad de conseguir un acierto de *cache* al solicitar la tesela $T(i, j, n)$, en un instante de tiempo t .

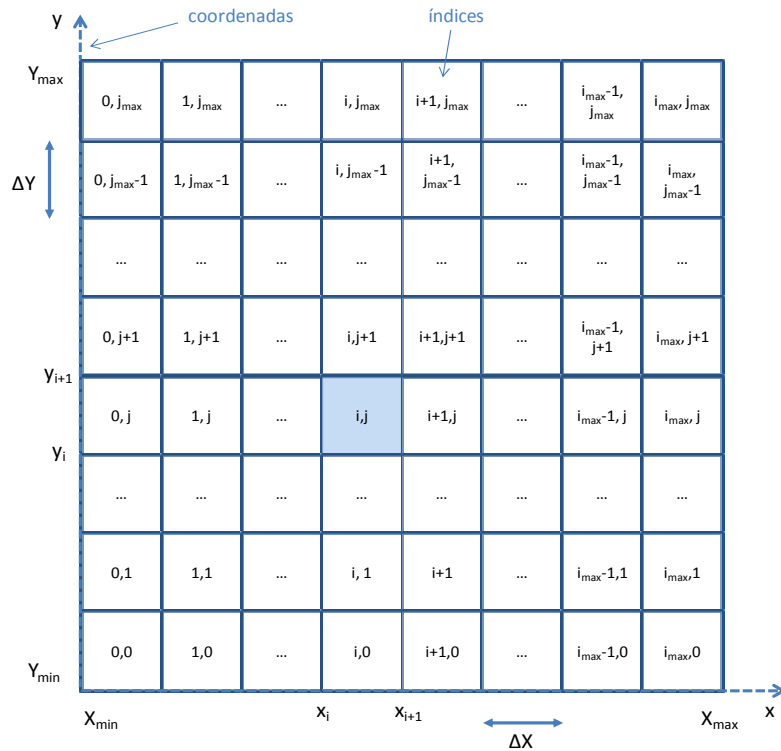


Figura 1: Espacio uniforme de teselado en un nivel concreto de la pirámide de escalas.

Denominemos C_{hit} al coste en segundos necesario para obtener un objeto de la *cache* (*hit*) y C_{miss} al coste incurrido al construir un nuevo objeto a partir de los servicios originales (*miss*).

Sea $f_{req}(x, y, n, t)$ la densidad espacial de probabilidad que caracteriza la distribución espacial de los centroides de las peticiones (*requests*) coincidentes con la escala n en un cierto instante de tiempo t . En general, para cualquier distribución de las peticiones, las teselas son solicitadas con probabilidad (1).

$$P_{req}\{T(i,j,n), t\} = \int_{y=y_{n,i}}^{y_{n,i+1}} \int_{x=x_{n,i}}^{x_{n,i+1}} f_{req}(x, y, n, t) dx dy \quad (1)$$

Aunque este resultado general será de utilidad a la hora de analizar las peticiones no teseladas dirigidas a un *proxy-cache*, en los análisis iniciales puede simplificarse suponiendo una distribución uniforme de los centroides de las peticiones dentro de cada tesela, o bien que las peticiones están restringidas a la rejilla de referencia (ver Figura 1) como en el caso de un WMS-C. En estas condiciones la probabilidad de acceso a una tesela de coordenadas $T(i,j,n)$ y de tamaño $\Delta x \times \Delta y$ es

$$P_{req}\{T(i,j,n), t\} = f_{req}(x, y, n, t) \Delta x \Delta y \quad (2)$$

Para una petición individual de una tesela situada en la coordenada (discretizada) (i, j, n) , en un instante t , la latencia observada viene determinada por la siguiente expresión:

$$\tau(i, j, n, t) = P_h\{T(i,j,n), t\} \tau_h + (1 - P_h\{T(i,j,n), t\}) \tau_m \quad (3)$$

Combinando las definiciones de (3) y (1) se obtiene una expresión probabilística conjunta de las peticiones que permite calcular una latencia media del servicio:

$$\tau(t) = \sum_{(i,j,n)} (\tau_m - P_h\{T(i,j,n), t\}(\tau_m - \tau_h)) P_{req}\{T(i,j,n), t\} \quad (4)$$

Suponiendo un tiempo de acceso constante a la *cache* y que también lo es el tiempo necesario para generar nuevos objetos (al menos en término medio), puede definirse la ganancia de uso de *cache* como el incremento porcentual de rendimiento en los aciertos de *cache*

$$G_c = \frac{\tau_m}{\tau_h} \quad (5)$$

y (4) puede escribirse como

$$\tau(t) = \tau_h \sum_{(i,j,n)} (G_c - P_h\{T(i,j,n), t\}(G_c - 1)) P_{req}\{T(i,j,n), t\} \quad (6)$$

Donde ya se pueden identificar algunos componentes que deben ser caracterizados al menos localizadamente.

En este punto no se han asumido más simplificaciones probabilísticas que, si bien podrían simplificar el diseño y funcionamiento de estos sistemas, eliminarían la información disponible para mejorar la gestión¹. Por lo tanto parece necesario estudiar cuáles son las características típicas de estas propiedades y cómo se relacionan con las métricas de calidad.

Algunas características relevantes del modelo planteado son:

- No hay independencia estadística entre $P_h\{T(i,j,n), t\}$ y $P_{req}\{T(i,j,n), t\}$ ya que resulta evidente que el estado de la *cache* está vinculado íntimamente a la historia de las peticiones del servicio.
- No hay invarianza temporal durante el régimen transitorio del sistema.
- La función densidad de probabilidad (y por ende la probabilidad expresada en (1) no es uniforme y probablemente presente vínculos directos con la estructura espacial de la información subyacente.

Otro factor a tener en cuenta es la magnitud práctica que supone gestionar estructuras de datos de dimensiones exponenciales. En una *cache* con estructura de *pirámide* de escalas, el número de elementos aumenta exponencialmente con la coordenada n por lo que la aplicación de algoritmos analíticos y predictivos se puede tornar impráctica, incluso mediante el apoyo de mecanismos heurísticos, si pretende aplicarse a todos los niveles de la *pirámide*.

Por este motivo resulta de gran interés obtener un modelo de relación estadística entre los distintos niveles de

¹ Por ejemplo, suponer una *cache* completa reduciría (6) a $\tau(t) = \tau_m G_c = \tau_h$

representación de la *pirámide* en el que se puedan extrapolar características cuantitativas de los niveles inferiores a partir de medidas recopiladas en un cierto nivel intermedio de la misma.

Vemos en (7) una estimación de la probabilidad de recibir peticiones en la zona geográfica de una tesela $T(i, j, n)$ calculada ponderadamente a partir de los niveles inferiores, hasta un cierto nivel L , según (1) (esta operación se ilustra en la Figura 2).

$$\hat{P}_{req}\{T(i, j, n), t\} = \sum_{l=L}^n \lambda_l \int_{y=y_{n,l}}^{y_{n,l}+1} \int_{x=x_{n,l}}^{x_{n,l}+1} f_{req}(x, y, l, t) dx dy \quad (7)$$

Bajo la suposición de que los centroides de las peticiones presentan una distribución uniforme dentro de cada tesela, o bien que las peticiones están restringidas a la rejilla de referencia puede hacerse uso de la expresión (2) para obtener la siguiente estimación de probabilidad:

$$\hat{P}_{req}\{T(i, j, n), t\} = \sum_{l=L}^n \lambda_l \cdot f_{req}(x, y, l, t) \Delta x \Delta y \quad (8)$$

Para incrementar la interoperabilidad entre clientes y servidores se recomienda que las capas publicadas por estos últimos utilicen un conjunto concreto de escalas en el mismo sistema de referencia de coordenadas (CRS) que la comunidad de usuarios acuerda utilizar. De esta forma, un cliente de mapas puede superponer teselas procedentes de distintos servidores sin necesidad de realizar reproyecciones ni re-escalado de teselas. Para ello se define el concepto de *Well-Known Scale Set* o "Conjunto de escalas bien conocidas", como una combinación de CRS y conjunto de escalas [7].

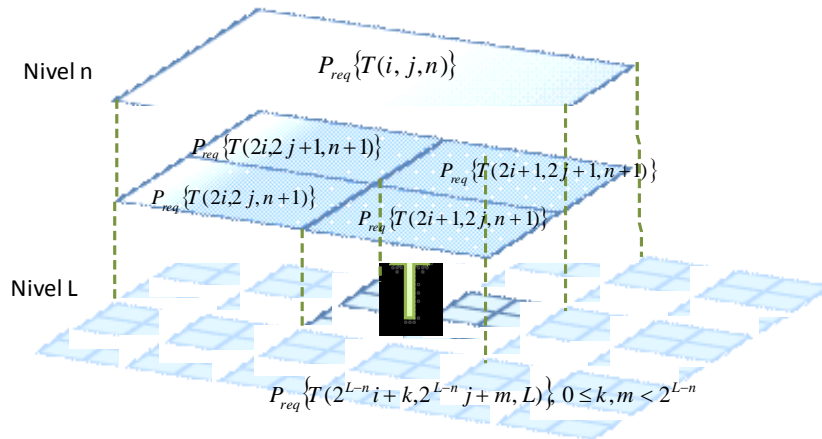


Figura 2: estimación de la probabilidad de recibir peticiones en la zona geográfica de una tesela $T(i, j, n)$ calculada a partir de los niveles inferiores, hasta un cierto nivel L .

Los esquemas de teselado definidos por varios *Well-Known Scale Set*, como el GoogleCRS84Quad o el GoogleMapsCompatible, definen una estructura piramidal en la que cada nivel de la pirámide realiza una representación de la zona geográfica utilizando el doble de teselas en cada dirección que el nivel anterior (de resolución más gruesa). En este caso se tiene que $\Delta x_n = 2^k \Delta x_{n-k}$ y $\Delta y_n = 2^k \Delta y_{n-k}$, y la expresión (8) puede reducirse a

$$\hat{P}_{req}\{T(i, j, n), t\} = \sum_{l=L}^n \lambda_l \sum_{y=y_{n,l}}^{y_{n,l}+1} \sum_{x=x_{n,l}}^{x_{n,l}+1} f_{req}(x, y, l, t) \quad (9)$$

La validez de (7)-(9) parte de la hipótesis de que la localización geográfica es una propiedad relevante para todos los niveles (o al menos para niveles próximos) y que las peticiones tienen una correlación espacial elevada. De ser así, pueden utilizarse métodos heurísticos basados en el Principio de Localidad [10] para gestionar la totalidad de la *cache* utilizando sondas estadísticas en niveles con un dominio de objetos (teselas) abordable.

3. MODELO DESCRIPTIVO

Los modelos descriptivos utilizan los registros de acceso a los servidores de mapas para determinar cuáles son las regiones de mapa más solicitadas por los usuarios. La aplicación Web *Hotmap* de Microsoft² utiliza los registros de peticiones del servicio de mapas *Bing Maps* para mostrar sobre un mapa el número de veces que se ha pedido cada tesela. *Hotmap* utiliza un *heatmap* o “mapa de calor” para representar de forma gráfica la actividad de los usuarios [11-13]. Sin embargo, no ofrece acceso a los propios datos de las peticiones, por lo que limita el posible análisis a una mera prospección visual, no pudiendo hacerse uso de algoritmos automáticos para la extracción de patrones de interés. La Figura 3 muestra una captura del uso de esta aplicación.

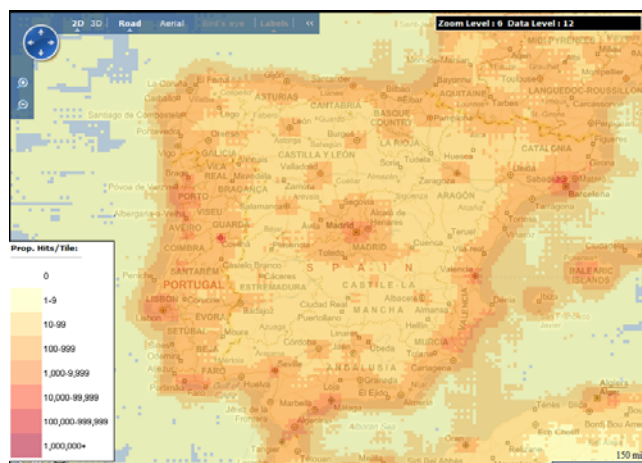


Figura 3: Visualización de patrones de acceso en Microsoft Hotmap.

3.1 SERVICIOS DE MAPAS ANALIZADOS

En este trabajo se han utilizado registros de peticiones realizadas a los servicios WMS-C de ámbito nacional *Cartociudad*, *IDEE-BASE* y *PNOA*, facilitados por el Instituto Geográfico Nacional (IGN).

Cartociudad es el resultado de la integración y armonización de datos aportados por diferentes organismos públicos (principalmente Dirección General del Catastro, Instituto Nacional de Estadística, Sociedad Estatal de Correos y Telégrafos e Instituto Geográfico Nacional) que ha dado lugar a un sistema de información geográfica de red viaria continua, e información parcelaria, censal y postal, cuyo ámbito es todo el territorio nacional [14].

El servicio *IDEE-BASE* permite visualizar la Base Cartográfica Numérica 1:25.000 y 1:200.000 del IGN, que contiene temas de información como hidrografía, vías de comunicación, núcleos y población, toponimia, transporte, etc.

Por último, el servicio de mapas *PNOA* permite visualizar las ortofotos del Plan Nacional de Ortofotografía Aérea, cuyo objetivo es obtener una cobertura ortofotográfica bienal de todo el territorio nacional.

Las versiones teseladas de estos servicios utilizan sendas *caches* de teselas implementadas mediante *Tilecache*, de Metacarta [15]. Este sistema de *cache* sigue la especificación WMS-C de OSGeo.

La Tabla 1 muestra los parámetros principales de estos servicios extraídos del documento de metadatos del servicio (*Capabilities*).

	CARTOCIUDAD	PNOA	IDEE-BASE
layers	DivisionTerritorial_Poligono CodigoPostal, Callejero Todas, SeccionCensal	PNOA	Todas
srs	EPSG:4258	EPSG:4258	EPSG:4326
forma	image/png	image/jpeg	image/jpeg
width	256		
height	256		
bbox	minx=-180 miny=-90 maxx=180 maxy=90		

² <http://hotmap.msresearch.us/>

scales	0.70312500000000000000 0.35156250000000000000 0.17578125000000000000 0.08789062500000000000 0.04394531250000000000 0.02197265625000000000 0.01098632812500000000 0.00549316406250000000 0.00274658203125000000 0.00137329101562500000 0.00068664550781250000 0.00034332275390625000 0.00017166137695312500 0.00008583068847656250 0.00004291534423828125 0.00002145767211914063 0.00001072883605957031 0.00000536441802978516 0.00000268220901489258 0.00000134110450744629
--------	---

Tabla 1: Parámetros de interés de los servicios WMS-C analizados.

3.2 EXTRACCIÓN DE LOS DATOS DE LAS PETICIONES

Los *logs* recogidos corresponden a los registros de acceso estándar de Apache con información sobre todas las peticiones que procesa, utilizando el Formato Común de Registro (*Common Log Format* - CLF) [16]. La información extraída de estos registros es la siguiente:

- Fecha en la que el servidor recibió la petición, con precisión de segundos.
- Dirección IP o *hostname* del cliente (host remoto) que hizo la petición al servidor. Las direcciones IP que se registran no son necesariamente las direcciones de las máquinas de los usuarios finales. Si existe un servidor *proxy* entre el usuario final y el servidor, la dirección que se registra es la del *proxy*.
- Código de estado que el servidor envía de vuelta al cliente. Esta información es muy valiosa, porque revela si la petición fue respondida con éxito por el servidor (los códigos que empiezan por 2), una redirección (los códigos que empiezan por 3), un error provocado por el cliente (los códigos que empiezan por 4), o un error en el servidor (los códigos que empiezan por 5).
- Tamaño del objeto retornado por el cliente, no incluidas las cabeceras de respuesta, expresado en *bytes*.

De la línea de la petición del cliente se extraen los parámetros propios de la petición WMS-C: *service*, *version*, *request*, *layers*, *width*, *height*, *format*, *styles*, *exceptions* y *crs*.

Se ha creado un sistema para la extracción, normalización y almacenamiento de los datos procedentes de estos *logs*, para su posterior explotación. Dado el elevado volumen de registros a procesar se ha establecido como requisito primordial que la herramienta automatice en la mayor medida posible este proceso. Para ello, el sistema ha sido creado en torno a un gestor de contenidos *Open Source* Alfresco³. Alfresco ofrece múltiples protocolos para subida/bajada de ficheros, como HTTP, HTTPS, CIFS/Samba, FTP y WebDav. La subida de un fichero de *logs* a la plataforma desencadena la ejecución de un *workflow* jBPM que realiza la inserción de los datos de las peticiones en una base de datos PostgreSQL con la extensión espacial PostGIS. También se mantiene un fichero *.sql* en el repositorio para posibles necesidades futuras (ver Figura 4).

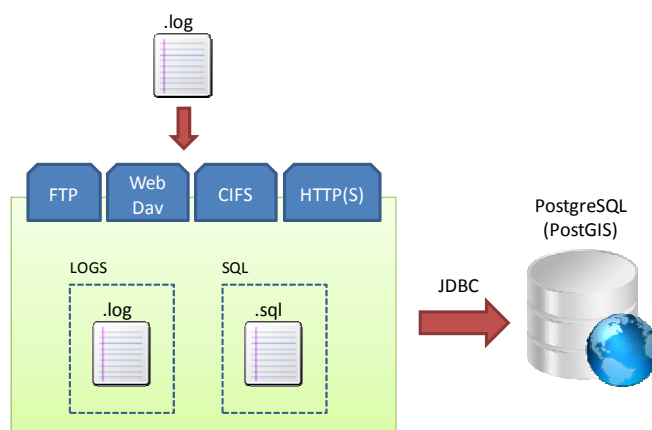


Figura 4: Escenario para la extracción, normalización y almacenamiento de los datos procedentes de los históricos de peticiones a los servidores de mapas.

En la Tabla 2 se recoge un resumen con las estadísticas globales de los registros de peticiones a los distintos servicios WMS-C considerados.

³ <http://www.alfresco.com>

	IDEE-BASE	CARTOCIUDAD	PNOA
total peticiones	16978535	3778369	9816747
válidas	16891385 (99.49%)	3362940 (89%)	9678358 (98.59%)
erróneas	87150 (0.51%)	415429 (11%)	138389 (1.41%)
GetMap	16978026 (99.997%)	3683101 (97.479%)	9816431 (99.997%)
GetCapabilities	509 (0.003%)	95260 (2.521%)	316 (0.003%)
fecha inicio	2010-03-15	2009-12-09	2010-03-19
fecha fin	2010-06-17	2010-06-20	2010-06-17
duración	93 días	193 días	89 días
frecuencia	182565 petic/día	19577 petic/día	110300 petic/día
tamaño objetos	6365.574 bytes	29181.239 bytes	12549.867 bytes

Tabla 2: Estadísticas globales de los registros de peticiones a los servicios WMS-C analizados

En las Figuras Figura 5 y Figura 6 se recoge la distribución de peticiones realizadas por los usuarios sobre las diferentes escalas. En los servicios IDEE-BASE y Cartociudad se observa un pico anómalo de peticiones realizadas sobre la escala 4. Se ha comprobado que este nivel de resolución es el más adecuado para mostrar por completo el territorio nacional en una única pantalla de mapa, resultando idóneo como pantalla inicial de mapa. Este hecho ofrece una justificación razonable sobre la existencia de estos máximos localizados en esta escala.

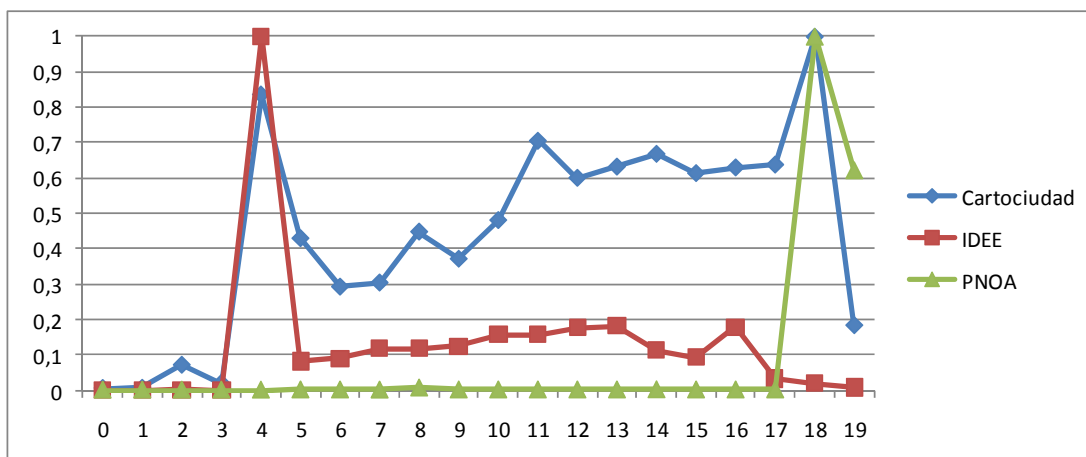


Figura 5: Distribución normalizada de peticiones a lo largo de las distintas escalas.

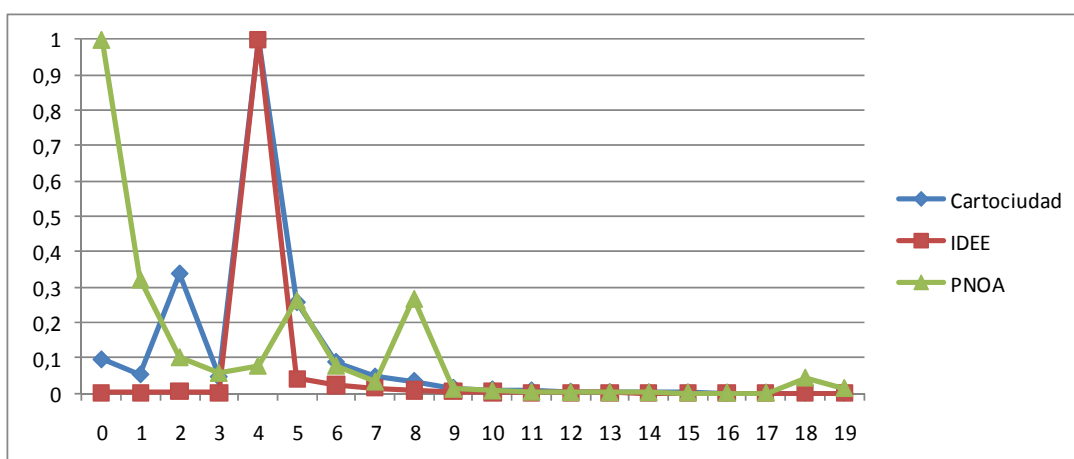


Figura 6: Distribución normalizada de densidad de peticiones a lo largo de las distintas escalas.

A pesar de que, por lo general, se recibe un mayor número de peticiones en los niveles de resolución más altos (ver Figura 5), son las teselas pertenecientes a los niveles de resolución más bajos (de resolución más gruesa) las que reciben, en término medio, una mayor densidad de peticiones por tesela (ver Figura 6). Este resultado respalda la práctica habitual de generar primero las teselas para los niveles de resolución más bajos, y generar posteriormente las teselas pertenecientes a escalas más altas o bien dejar que éstas vayan poblando la *cache* a medida que son pedidas.

Se ha comprobado que la mayor parte de las peticiones de mapas registradas se realizan sobre el territorio nacional, como era de esperar dado el ámbito de estos servicios. Por ello, se ha limitado la zona geográfica de estudio a esta región, concretamente al *bounding box* [-11.9971, 32.8711, 5.5371, 46.0107] (ver Figura 7). Este *bounding box* está alineado con la rejilla definida por el servicio teselado en el nivel de resolución 12, abarcando 400 y 300 teselas en las direcciones horizontal y vertical, respectivamente, en este nivel.



Figura 7: Región geográfica considerada (área sombreada) para el estudio.

3.3 MODELO SIMPLIFICADO

Dada la naturaleza exponencial de la pirámide de escalas y la imposibilidad de trabajar con las estadísticas de teselas individuales, se ha empleado un modelo simplificado que aproxima la probabilidad de petición de una tesela a partir de las estadísticas recogidas en otro nivel (digamos l) de resolución en la región geográfica cubierta por esa tesela, como se formaliza en la expresión (10). Concretamente, se ha simplificado el modelo a la rejilla de tamaño 400x300 definida por la región antes mencionada en la escala de resolución 12 ($l = 12$). La estructura piramidal de escalas se transforma, en cierto modo, en una estructura tipo prisma con el mismo número de elementos (400x300) en todas las escalas (ver Figura 8).

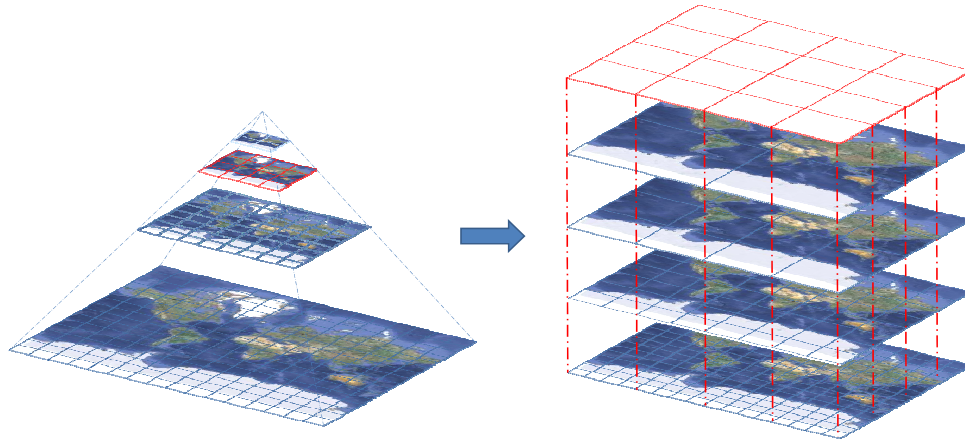


Figura 8: Paso “conceptual” de pirámide de escalas a prisma mediante el modelo simplificado.

$$\hat{F}_{reg,i}[T(l,j,n)] = \int_{y=y_{n,i}}^{y_{n,i+1}} \int_{x=x_{n,i}}^{x_{n,i+1}} f_{reg}(x,y,l) dx dy \quad (10)$$

donde

$$f_{reg}(x,y,l) = \frac{n_{reg}(l,j,n)}{\Delta x \Delta y N_{Total,i}}, \text{ para } \begin{cases} x_{n,i} \leq x < x_{n,i+1} \\ y_{n,j} \leq y < y_{n,j+1} \end{cases} \quad (11)$$

con

$$N_{Total,i} = \int_{y=y_{min,i}}^{y_{max,i}} \int_{x=x_{min,i}}^{x_{max,i}} f_{req}(x,y,i) dx dy \quad (12)$$

Dada la naturaleza *booleana* del *cacheado* (una tesela, o bien se introduce en *cache*, o por el contrario no lo hace) la salida del modelo viene dada por una función $\theta_i\{T(i,j,n)\}$ de esta misma naturaleza que se aplica sobre la probabilidad de petición de las teselas.

Se define un umbral de probabilidad μ de forma que las teselas precargadas ($\theta_i\{T(i,j,n)\} = 1$) son aquellas cuya estimación de probabilidad de ser pedida supera este umbral:

$$\theta_i\{T(i,j,n)\} = \begin{cases} 1, & \text{si } \hat{P}_{req,i}\{T(i,j,n)\} \geq \mu \\ 0, & \text{si } \hat{P}_{req,i}\{T(i,j,n)\} < \mu \end{cases} \quad (13)$$

La elección de este umbral $\mu \in (0,1)$ determinará el número de teselas *cacheadas* y los aciertos de *cache* obtenidos. Evidentemente, si $\mu = 0$, se conseguirá un 100% de aciertos de *cache*, a costa de introducir en *cache* todas las teselas. En el lado opuesto, si $\mu = 1$ no se introducirá en *cache* ningún objeto por lo que no se producirán aciertos de *cache*. La elección de este umbral debe hacerse de tal forma que se maximice la probabilidad de acierto manteniendo el consumo de recursos por debajo de un nivel dado.

Sea γ una función *booleana* que determina si se ha pedido ($\gamma = 1$) o no ($\gamma = 0$) una tesela en el histórico de peticiones analizadas

$$\gamma\{T(i,j,n)\} = \begin{cases} 1, & \text{si } n_{req}\{T(i,j,n)\} > 0 \\ 0, & \text{si } n_{req}\{T(i,j,n)\} = 0 \end{cases} \quad (14)$$

$\Psi_1(i,j,n) \in (0,1)$ representa la fracción de teselas de la celda (i,j,n) del prisma que han sido solicitadas alguna vez, para el modelo simplificado a nivel i :

$$\Psi_1(i,j,n) = \frac{\int_{y=y_{min,i}}^{y_{max,i}} \int_{x=x_{min,i}}^{x_{max,i}} \gamma(x,y,n) dx dy}{n_{req,i,j,n}} \quad (15)$$

Se introduce el concepto de “heterogeneidad” de una celda, formalizado según la expresión (16). Su interpretación es sencilla, cuando se recoge un elevado número de peticiones en una celda, pero éstas se efectúan sobre un reducido número de teselas pertenecientes a esa celda, se considera que ésta es “heterogénea”. Estas zonas requieren un análisis a más bajo nivel.

$$heterogeneidad(i,j,n)_i = \hat{P}_{req,i}\{T(i,j,n)\}(1 - \Psi_1(i,j,n)) \quad (16)$$

3.4 EXPERIMENTO Y RESULTADOS

Con el objetivo de experimentar con el modelo propuesto, los registros de peticiones se han dividido en dos rangos de tiempo: el primero de ellos se ha utilizado como fuente para la realización de las predicciones y el segundo para la validación de las mismas. Se ha utilizado el modelo simplificado a la rejilla definida por el nivel de resolución 12 ($i = 12$).

En las Figuras Figura 10, Figura 11 y Figura 12 se muestran los *heatmaps* o mapas de calor para los registros de peticiones registradas en los servicios de *Cartociudad*, *PNOA* e *IDEE-BASE*, respectivamente, según el modelo. En estas figuras pueden identificarse elementos que reciben un elevado número de peticiones, como las zonas costeras, núcleos urbanos o las principales vías de transporte. Estos elementos pueden utilizarse en forma de entidades vectoriales para su uso en un modelo predictivo, como se expone en [9].

En las Tablas Tabla 3-Tabla 5 se recogen los porcentajes de acierto para los servicios considerados, utilizando el modelo simplificado, utilizando como umbral la media ($\mu = E\{\hat{P}\}$). Estas tablas se interpretan como el porcentaje de aciertos conseguidos en el nivel de resolución identificado por la columna a partir de las estadísticas recogidas en aquel identificado por la fila, según el modelo simplificado. La última columna muestra el consumo de recursos, medido como porcentaje de teselas *cacheadas* por el modelo. Así, la celda sombreada en la Tabla 3 indica que, utilizando como fuente de predicción las estadísticas recogidas para el nivel de resolución

13, se consigue un 91.95% de aciertos de *cache* en el nivel de resolución 18, para lo cuál se requiere *cachear* el 8.83% de los objetos.

	12	13	14	15	16	17	18	19	recursos
12	87.72	90.45	91.50	90.27	94.44	95.40	96.07	92.30	18.08
13	69.67	78.86	82.33	83.01	88.67	80.88	91.95	85.09	8.83
14	56.62	69.88	76.49	78.88	84.64	78.87	89.04	80.25	5.40
15	41.40	54.97	65.82	74.59	78.40	74.59	85.51	73.37	3.00
16	33.20	45.61	57.69	66.35	73.81	71.24	84.10	69.05	2.04
17	29.32	42.41	53.36	63.56	54.82	78.21	81.57	66.49	1.61
18	26.39	37.71	50.47	61.57	52.82	67.84	81.79	63.44	1.35
19	15.26	23.44	33.68	44.71	37.99	48.68	72.75	44.52	0.76
prop	67.75	76.35	80.60	84.42	87.49	90.81	92.74	83.80	6.68

Tabla 3: Porcentaje (%) de aciertos de *cache* mediante el modelo simplificado obtenido con los registros de Cartociudad, con $\mu = media$.

	12	13	14	15	16	17	18	19	recursos
12	88.06	85.72	93.76	94.52	96.09	93.75	93.57	79.31	21.02
13	74.86	85.96	90.71	92.18	94.09	89.75	82.49	54.88	15.28
14	48.11	57.73	84.37	86.60	89.16	78.54	59.86	38.77	7.63
15	34.62	45.64	76.00	82.75	81.70	66.90	50.63	31.59	4.58
16	41.62	53.14	80.94	84.07	89.80	74.78	56.24	37.58	6.10
17	30.20	37.63	57.01	60.13	62.28	69.51	55.51	23.46	3.26
18	23.57	25.59	41.75	46.15	45.86	41.45	61.97	33.37	2.33
19	8.86	8.68	12.45	13.13	14.33	12.27	13.69	44.11	1.23
prop	67.18	78.03	87.70	90.15	92.39	86.83	83.95	71.60	12.50

Tabla 4: Porcentaje (%) de aciertos de *cache* mediante el modelo simplificado obtenido con los registros de IDEE-BASE, con $\mu = media$.

	12	13	14	15	16	17	18	19	recursos
12	21.10	25.17	27.47	18.88	20.30	32.07	11.36	12.74	0.64
13	11.20	22.82	21.83	17.17	30.66	29.32	11.58	13.18	0.25
14	6.60	15.92	21.45	17.88	28.14	23.29	10.59	12.22	0.17
15	3.45	9.93	15.67	17.85	29.45	18.81	8.43	9.70	0.11
16	2.40	6.53	11.17	12.45	28.00	13.46	6.16	7.10	0.06
17	0.95	4.57	7.52	9.63	12.64	5.94	4.58	5.20	0.03
18	45.25	64.98	73.61	75.58	86.04	89.34	86.49	88.61	9.06
19	39.40	61.72	69.89	67.10	72.86	69.04	82.54	85.22	7.71
prop	40.05	55.97	64.05	61.22	72.28	64.53	76.77	80.26	8.93

Tabla 5: Porcentaje (%) de aciertos de *cache* mediante el modelo simplificado obtenido con los registros de PNOA, con $\mu = media$.

Destacar, sin embargo, que el beneficio del uso de una *cache* parcial no reside en el ahorro en el número de teselas *cacheadas*, sino en el espacio de almacenamiento y el tiempo de generación ahorrados. Como se expone en [9] y se ha verificado en el presente estudio, el ahorro en el número de teselas *cacheadas* supera al conseguido en relación al espacio de almacenamiento ahorrado. Estos resultados revelan que las teselas de mayor “interés” (las más solicitadas) incurrir un mayor coste de almacenamiento que aquellas omitidas. Generalmente, las zonas “populares” presentan una mayor variación cromática y de patrones, por lo que estas zonas que se introducen en *cache* almacenadas como imágenes comprimidas en formatos como JPEG o PNG, usan una mayor proporción de espacio en disco que las teselas no-*cacheadas* que presentan pocos contrastes en la imagen. A modo de ejemplo, se recoge en la Tabla 6 la gran diferencia entre el espacio de almacenamiento en disco ocupado por una tesela que recibe un elevado número de peticiones (una zona urbana) y otra que no se ha solicitado (una zona de mar). Este efecto aplica incluso en mayor proporción en el tiempo de creación de las teselas.

CARTOCIUDAD	PNOA	IDEE-BASE
-------------	------	-----------

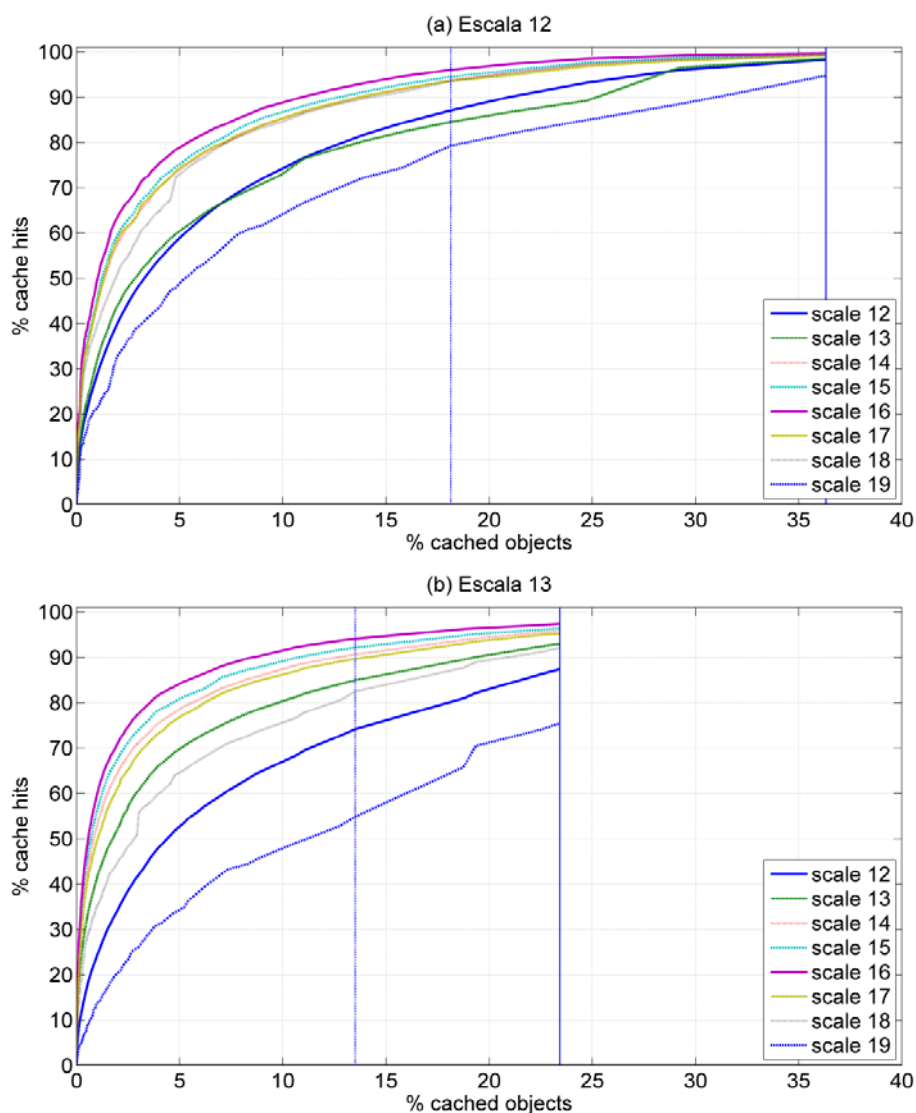
Núcleo urbano	101,08 KB	29,95 KB	12,56 KB
Zona de mar interior	3,57 KB	5,49 KB	1,73 KB

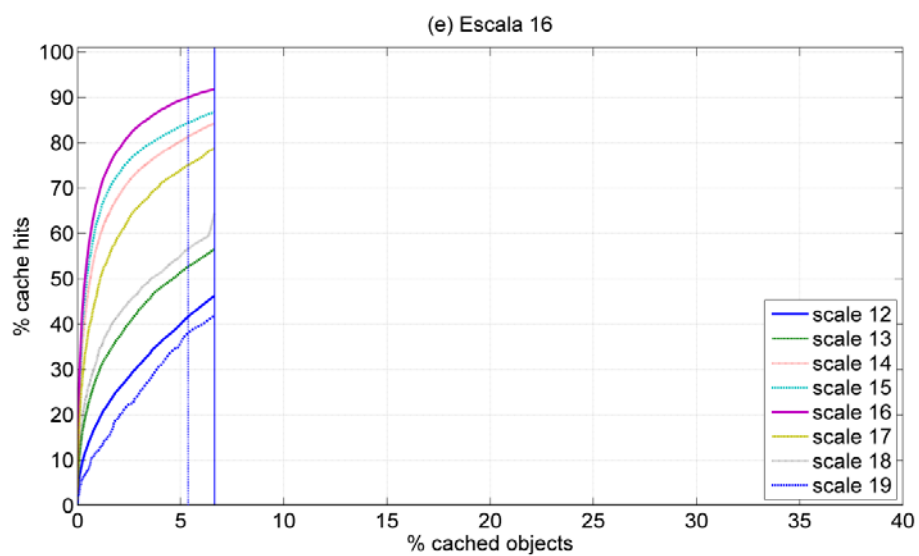
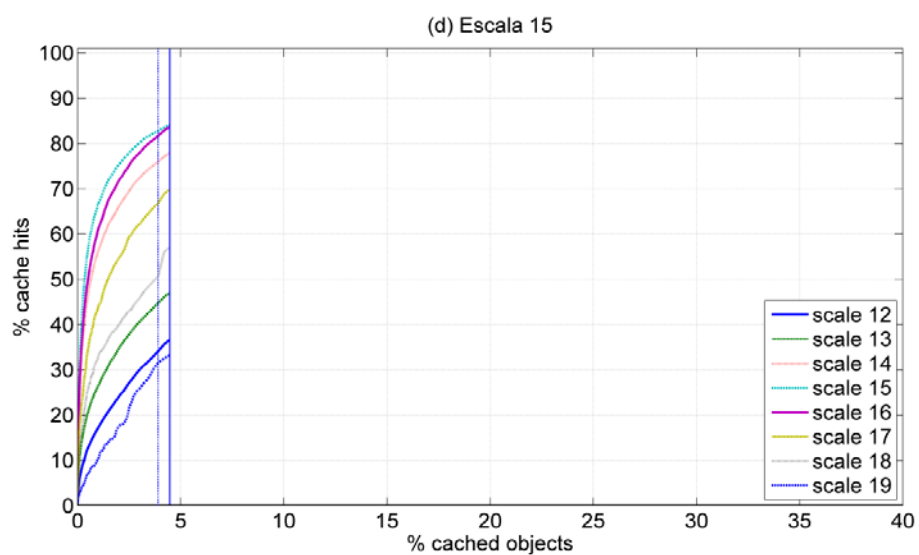
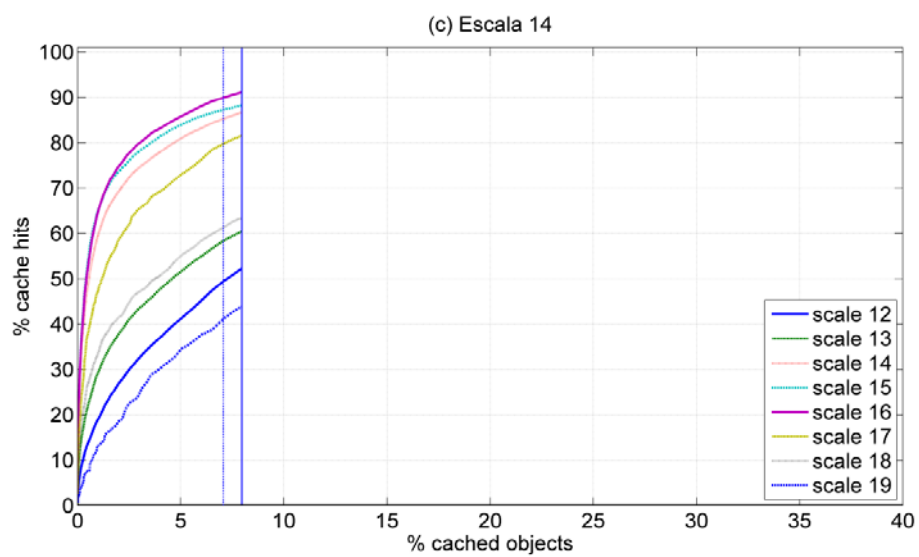
Tabla 6: Espacio de almacenamiento de sendas teselas: una perteneciente a una zona de mar y otra urbana.

Las bajas tasas de acierto conseguidas por el modelo para el servicio *PNOA* se deben a que, como se observa en la Figura 5, este servicio presenta una distribución poco uniforme de las peticiones por las distintas escalas (el 99% de las peticiones se concentran en las dos escalas más altas). Para estos servicios en los que los patrones de acceso son muy heterogéneos entre distintos niveles, el modelo no es capaz de realizar buenas predicciones a partir del conocimiento extraído de otros niveles.

En la Figura 9 se representa el porcentaje de aciertos conseguidos por el modelo frente al porcentaje de objetos *cacheados* para el servicio *IDEA-BASE*. La línea vertical punteada identifica el umbral de probabilidad ($\mu = E\{F\}$). A partir de un determinado porcentaje de objetos *cacheados*, identificado por una línea vertical continua, el modelo no es capaz de realizar ninguna predicción. Los objetos restantes (fracción de objetos a la derecha de dicha línea) corresponden a teselas que no han sido solicitadas y por tanto no figuran peticiones de las mismas en los *logs*.

Como se puede observar, por lo general el modelo simplificado ofrece mejores resultados cuando, como fuente de la predicción para un determinado nivel de resolución, se utiliza otro nivel de resolución cercano, empeorando cuando éstos son distantes. Al descender en la pirámide de escalas, hacia niveles de resolución más fina, el porcentaje de objetos solicitados es cada vez menor, por lo que el rango en el que el modelo es capaz de realizar predicciones también lo es.





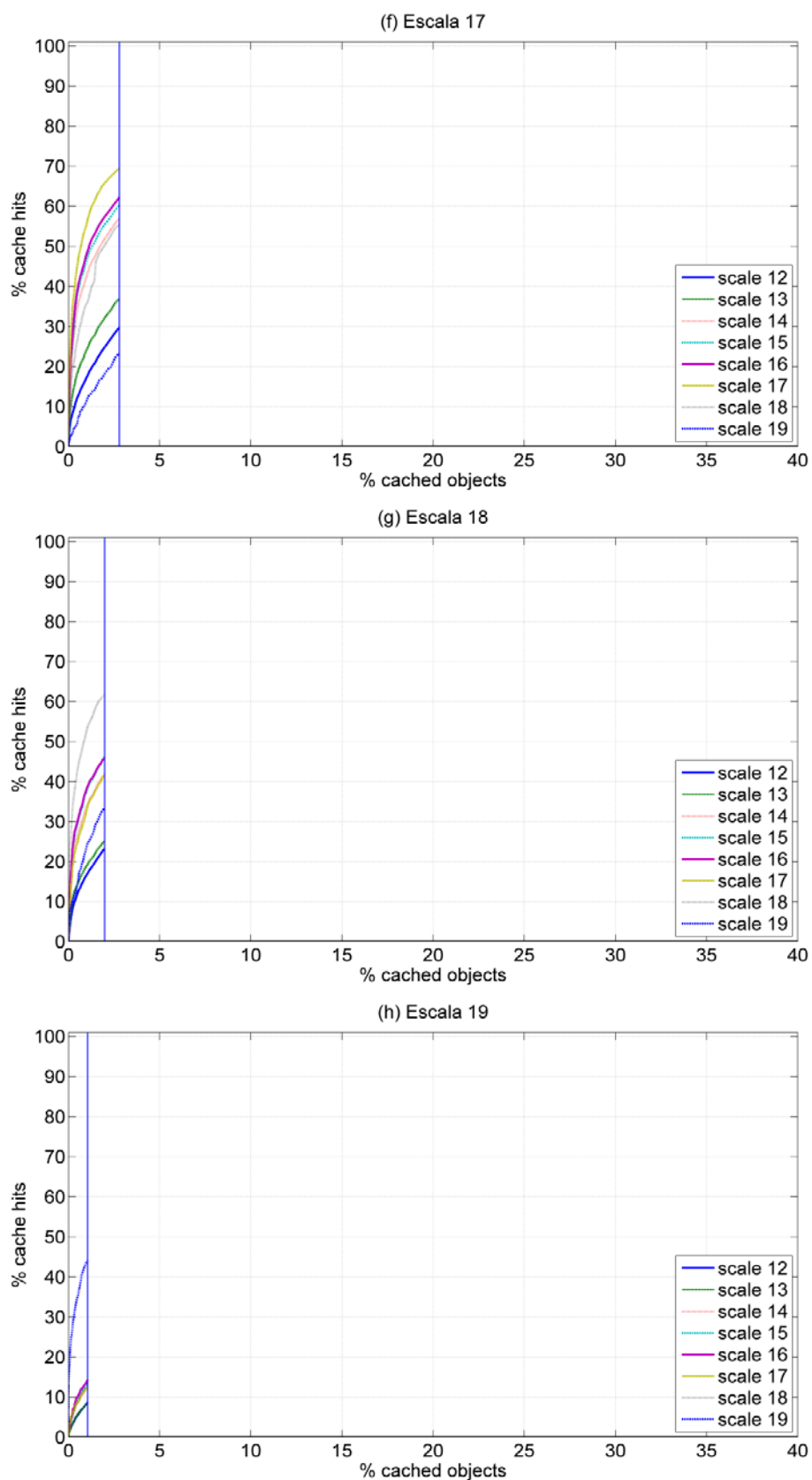


Figura 9: % Aciertos de *cache* Vs % Objetos *cacheados* para el servicio *IDEE-BASE*, mediante el modelo simplificado a la rejilla definida por el nivel de resolución 12.

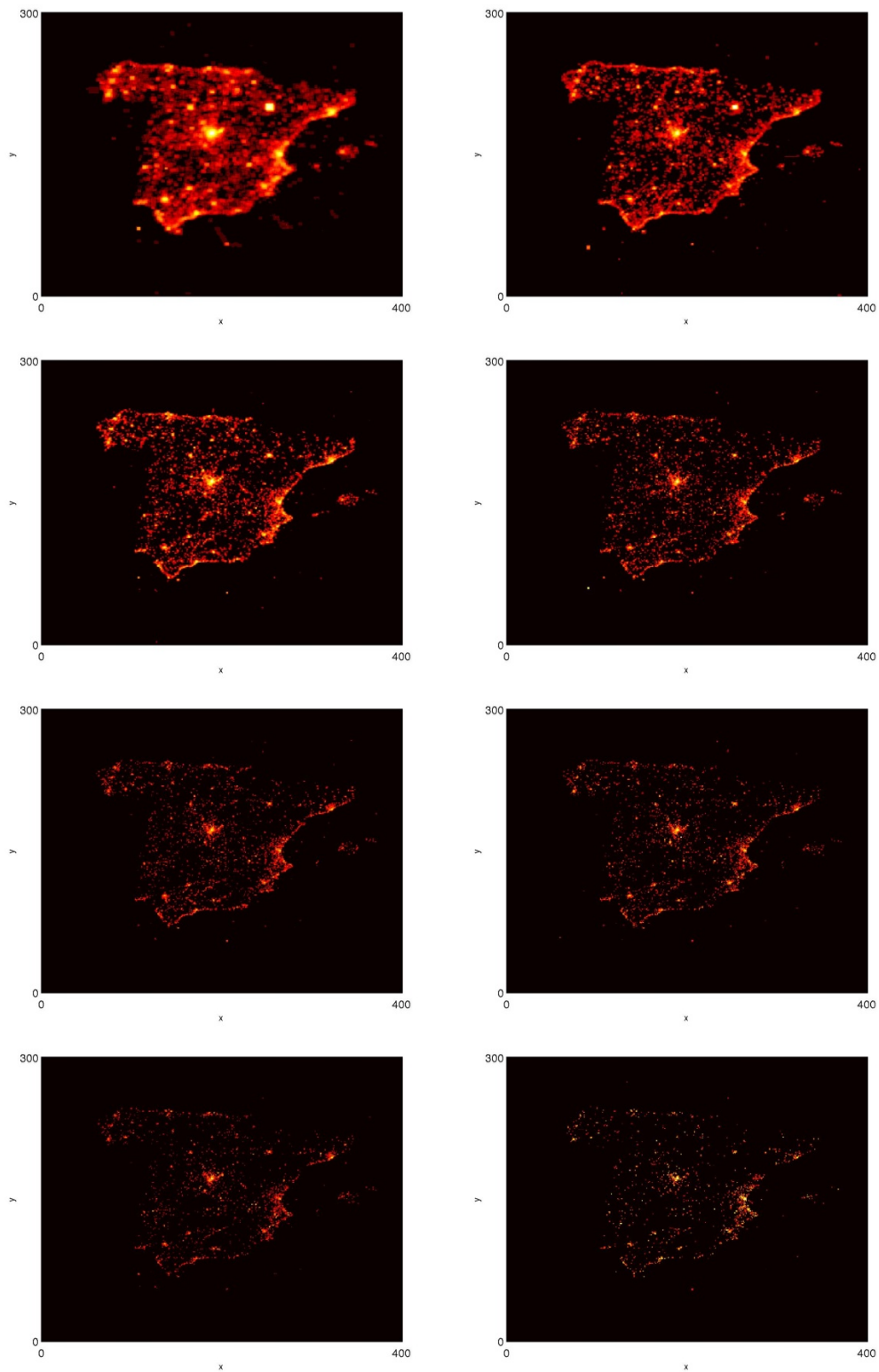


Figura 10: Registros de peticiones al servicio WMS-C de *Cartociudad* propagadas al nivel de resolución 12, para las escalas 12 a 19 (ordenadas de izda. a dcha. y de arriba a abajo).

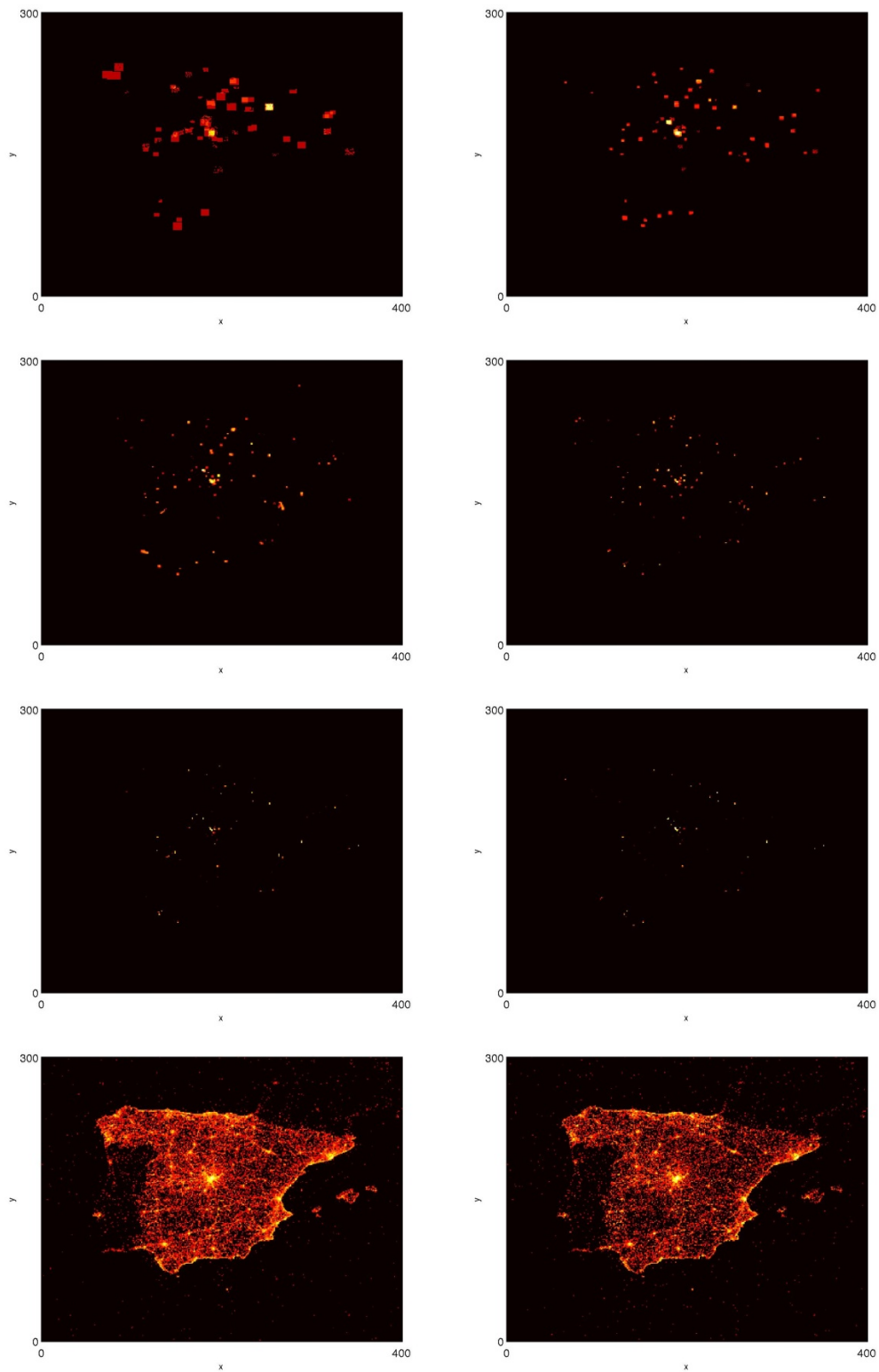


Figura 11: Registros de peticiones al servicio WMS-C de *PNOA* propagadas al nivel de resolución 12, para las escalas 12 a 19 (ordenadas de izda. a dcha. y de arriba a abajo).

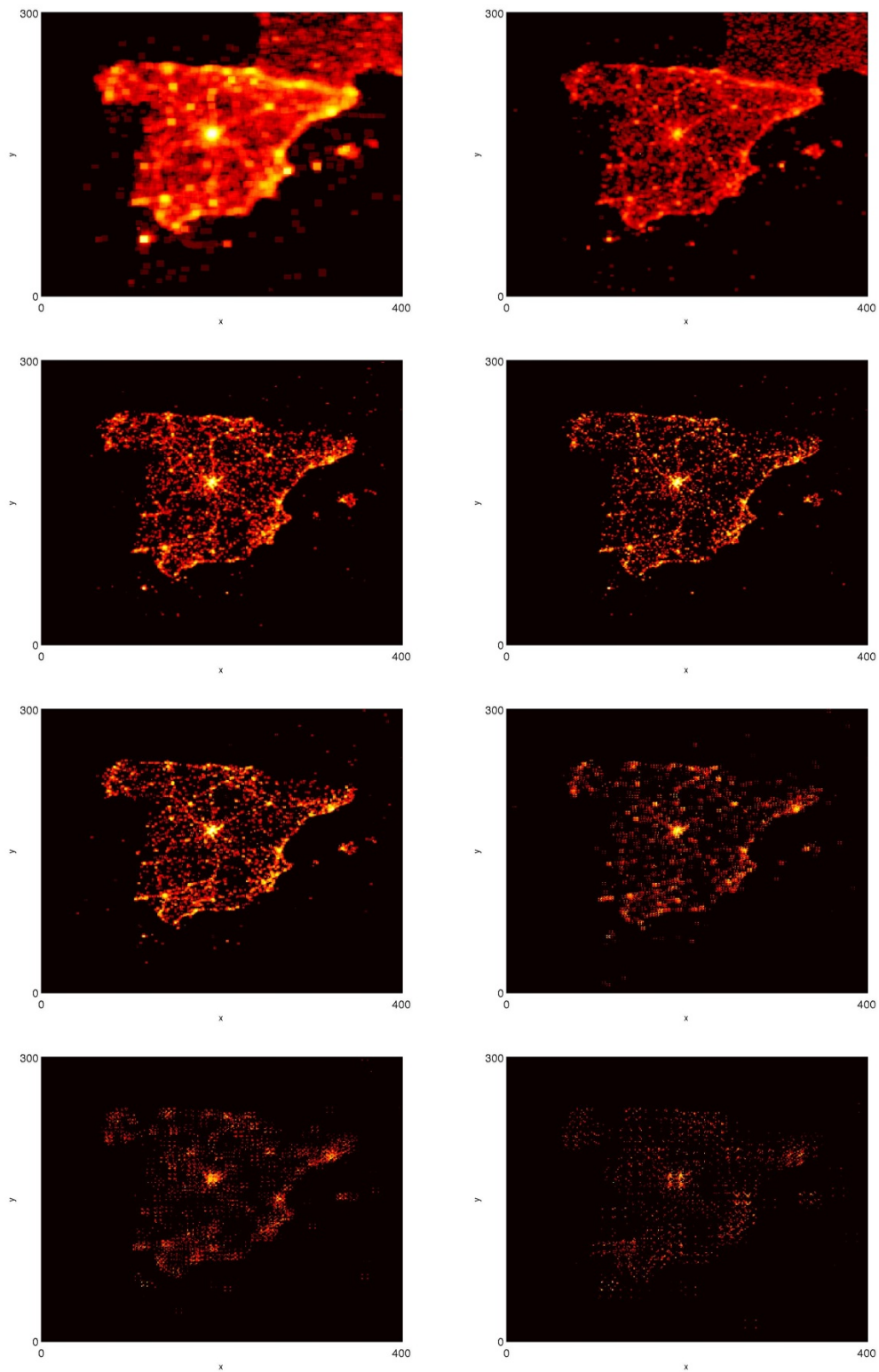


Figura 12: Registros de peticiones al servicio WMS-C de *IDEE-BASE* propagadas al nivel de resolución 12, para las escalas 12 a 19 (ordenadas de izda. a dcha. y de arriba a abajo).

4. CONCLUSIONES

Los servicios de mapas basados en teselado han adquirido una gran popularidad en los últimos, mejorando los tiempos de respuesta y la escalabilidad respecto a los servicios tradicionales, al servir imágenes pregeneradas a partir de una *cache*. En escenarios con recursos de almacenamiento escasos o en los que la cartografía se actualiza con mucha frecuencia no resulta adecuado el pregenerar todo el contenido cartográfico y hay que recurrir al uso de *caches* incompletas. La solución propuesta en este trabajo se basa en la definición de zonas prioritarias para la actuación de mecanismos de precarga y de reemplazo, de forma que se maximice la QoS experimentada por los usuarios manteniendo un consumo de recursos por debajo de un nivel dado. Se trata de mantener en *cache* los objetos con mayor probabilidad de ser pedidos. Para la determinación de las zonas prioritarias se ha propuesto un modelo descriptivo basado en el estudio de los registros de peticiones realizadas a diversos servicios de mapas de ámbito nacional, como *Cartociudad*, *PNOA* e *IDEE-BASE*. Las altas tasas de acierto conseguidas validan la hipótesis de que puede realizarse una predicción del acceso futuro a las teselas atendiendo exclusivamente a la información disponible en accesos pasados, cuando los patrones de uso del servicio presentan una cierta invarianza temporal. El análisis multi-resolución realizado respalda la utilización de estadísticas recogidas en un cierto nivel de la pirámide de escalas para predecir el comportamiento en otros niveles cercanos.

A modo de líneas futuras, podrían conseguirse mayores tasas de acierto en *cache* combinando de forma ponderada la información recopilada en distintos niveles de la pirámide de escalas, como se recoge en la expresión (9), y no solo a partir de un nivel individual como se realiza en el presente trabajo.

A partir de los registros de uso de los servidores de mapas analizados, se han identificado visualmente elementos que reciben un elevado número de peticiones, como las zonas costeras, núcleos urbanos o las principales vías de transporte. Estos elementos pueden utilizarse en forma de entidades vectoriales para su uso en un modelo predictivo para la identificación de objetos prioritarios a tener en cuenta durante las tareas de mantenimiento de la *cache*.

5. AGRADECIMIENTOS

El desarrollo de este trabajo ha sido posible gracias a la financiación por parte del Instituto Geográfico Nacional en el marco del Proyecto Conjunto al amparo del convenio de colaboración entre la dirección general del Instituto Geográfico Nacional y la Universidad de Valladolid en su edición de 2009-2010.

REFERENCIAS

- [1] Lucian Plesea, "The Design, Implementation and operation of the JPL OnEarth WMS Server," Geospatial Services and Applications for the Internet, Sample, J.T., Shaw, K., Tu, S., y Abdelguerfi, M., eds., Berlin: Springer, 2008, págs. 93-109.
- [2] Matt Mills, "NASA World Wind Tile Structure" Available: <http://www.ceteranet.com/nww-tile-struct.pdf>.
- [3] K.P. Přidal, "GDAL2Tiles - Publish Maps by Google Earth (KML SuperOverlay), Google Maps or OpenLayers as online static files (pyramid tile structure generator)" Available: <http://www.klokan.cz/projects/gdal2tiles/>.
- [4] OSGeo, "Tile Map Service Specification," Tile Map Service Specification - OSGeo Wiki Available: http://wiki.osgeo.org/wiki/Tile_Map_Service_Specification.
- [5] J. Schwartz, "Bing Maps Tile System," Microsoft Developer network Available: <http://msdn.microsoft.com/en-us/library/bb259689.aspx>.
- [6] OSGeo, "WMS Tiling Client Recommendation - OSGeo Wiki" Available: http://wiki.osgeo.org/wiki/WMS_Tiling_Client_Recommendation.
- [7] Joan Masó, Keith Pomakis, y Núria Julià, "OpenGIS Web Map Tile Service Implementation Standard | OGC®" Available: <http://www.opengeospatial.org/standards/wmts>.
- [8] Kai Cheng, Y. Kambayashi, y M. Mohania, "Efficient management of data in proxy cache," Database and Expert Systems Applications, 2001. Proceedings. 12th International Workshop on, 2001, págs. 479-483.
- [9] S. Quinn y M. Gahegan, "A Predictive Model for Frequently Viewed Tiles in a Web Map,"

Transactions in GIS, vol. 14, 2010, págs. 193-216.

- [10] P.J. Denning, "The locality principle," Commun. ACM, vol. 48, 2005, págs. 19-24.
- [11] D. Fisher, "Hotmap: Looking at geographic attention," Visualization and Computer Graphics, IEEE Transactions on, vol. 13, 2007, págs. 1184-1191.
- [12] D. Fisher, "How we watch the city: Popularity and online maps," Workshop on Imaging the City, ACM CHI 2007 Conference, 2007.
- [13] D. Fisher, "The Impact of Hotmap," 2009.
- [14] "Portal de CartoCiudad" Available: <http://www.cartociudad.es/portal/1024/index.htm>.
- [15] MetaCarta, "TileCache, from MetaCarta Labs," TileCache Available: <http://tilecache.org/>.
- [16] "Archivos de Registro (Log Files) - Servidor HTTP Apache" Available: <http://httpd.apache.org/docs/2.0/es/logs.html>.

2.3 CONTACTOS

Ricardo García
ricgar@tel.uva.es
Universidad de Valladolid
Laboratorio de Infraestructuras de
Datos Espaciales (IDELab)

Juan Pablo de Castro
jpdecastro@tel.uva.es
Universidad de Valladolid
Laboratorio de Infraestructuras de
Datos Espaciales (IDELab)