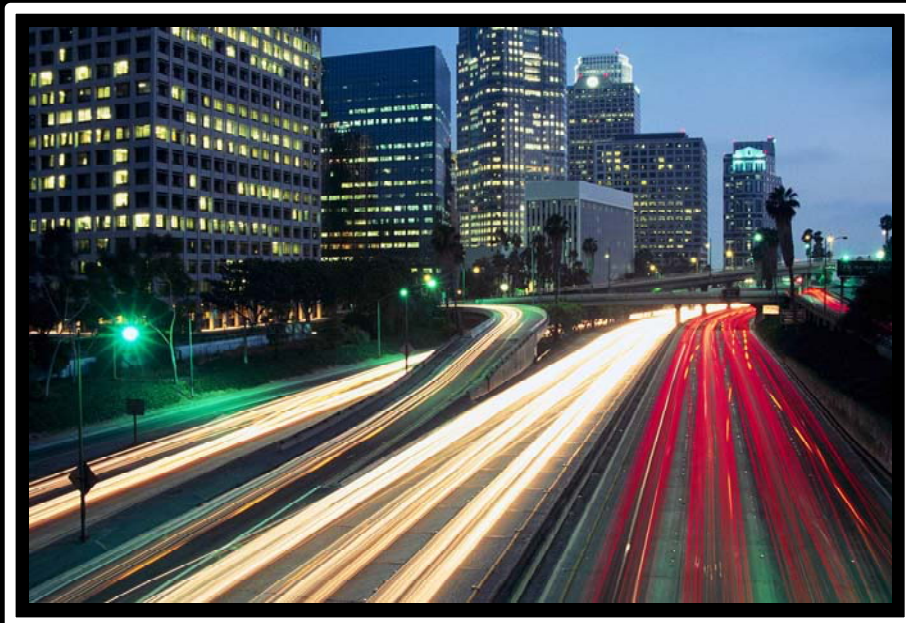


# Estimación de Matrices de Movilidad Mediante Datos de Telefonía Móvil

Noelia Cáceres Sánchez



TESIS DOCTORAL





**ESTIMACIÓN DE MATRICES DE MOVILIDAD  
MEDIANTE DATOS DE TELEFONÍA MÓVIL**

**Septiembre 2009**

**TESIS DOCTORAL**

Noelia Cáceres Sánchez



# TESIS DOCTORAL

## ESTIMACIÓN DE MATRICES DE MOVILIDAD MEDIANTE DATOS DE TELEFONÍA MÓVIL

**Presentada por:** Noelia Cáceres Sánchez  
Ingeniero de Telecomunicación

**Dirigida por:** Francisco García Benítez  
Dr. Ingeniero Industrial  
Luis M. Romero Pérez  
Dr. Ingeniero Industrial  
Johan Wideberg  
Dr. Ingeniero Industrial

**Septiembre 2009**

Departamento de Ingeniería Mecánica y de los Materiales  
Ingeniería e Infraestructura de los Transportes  
Escuela Técnica Superior de Ingenieros  
Universidad de Sevilla

Esta Tesis ha sido realizada gracias al proyecto TRA2005–09138, concedido por el *Ministerio de Ciencia y Tecnología* a través del programa “*Plan Nacional I+D*”, y a *VODAFONE* dentro del marco del “*Proyecto Minerva*” (1C-021).



*A Fran, mi compañero de viaje*

*A Pilar, nuestra ilusión*





# AGRADECIMIENTOS

Esta Tesis Doctoral, si bien ha requerido de esfuerzo y mucha dedicación por parte de la autora y sus directores, su finalización no hubiese sido posible sin la cooperación de todas y cada una de las personas y entidades que a continuación menciono. A Vodafone, especialmente a Santiago Borrero, Jesús Pérez y M<sup>a</sup> Satur Torre. Al Centro de Gestión de Tráfico del Suroeste (DGT), en particular a Ana Luz Jiménez. A Alejandro Carballar y Ana Madera del Proyecto Minerva. A Manuel García de ITSoft. A todos ellos gracias por su inestimable colaboración en el desarrollo de este trabajo.

Me gustaría agradecer sinceramente a Dr. Francisco García Benítez, Dr. Luis M. Romero Pérez y Dr. Johan Wideberg, la confianza depositada en mí, así como su dedicación y apoyo para que este trabajo llegara a buen término. Sus conocimientos, orientaciones, paciencia y motivación han sido fundamentales para mi formación como investigadora.

De igual manera agradecer a Dr. Jose María del Castillo Granados los conocimientos y consejos ofrecidos para la realización de esta Tesis, especialmente en la última etapa.

A mis compañeros de Departamento, por ofrecerme su ayuda y amistad, haciendo aún más agradable el trabajo cotidiano.

A mis padres, mi hermano y mis abuelos, por el apoyo incondicional que me han dado en todo momento y porque, a pesar de la distancia, el ánimo y alegría que me brindan me dan la fortaleza necesaria para seguir adelante.

A Fran, por ser la persona con la que comparto mi vida, por su amor, paciencia y comprensión, porque en su compañía las cosas malas se convierten en buenas, la tristeza se transforma en alegría y la soledad no existe.

Por último, y no menos importante, a mi hija Pilar, su nacimiento ha coincidido con el final de la Tesis. Ella es lo mejor que me ha pasado nunca, y ha venido a este mundo a darme el último empujón para terminar este trabajo.



# CONTENIDO

<b>CAPÍTULO 1 INTRODUCCIÓN .....</b>	<b>1</b>
1.1 MOTIVACIÓN.....	1
1.2 ÁMBITO DE LA TESIS .....	2
1.3 OBJETIVOS .....	4
1.4 ANTECEDENTES .....	6
1.5 ESTRUCTURA DE LA TESIS .....	13
<b>CAPÍTULO 2 REDES DE TELEFONÍA CELULAR: GSM .....</b>	<b>17</b>
2.1 INTRODUCCIÓN .....	17
2.2 ARQUITECTURA DE LA RED GSM .....	18
2.3 BASES DE DATOS HLR Y VLR .....	23
2.3.1 Introducción.....	23
2.3.2 Home Location Register (HLR) .....	24
2.3.3 Visitor Location Register (VLR).....	24
2.4 IDENTIDADES Y NÚMEROS DE USUARIO Y RED.....	25
2.4.1 Internacional Mobile Station Identity (IMSI).....	25
2.4.2 Temporary Mobile Subscriber Identity (TMSI) .....	26
2.4.3 Location Area Identification (LAI) .....	26
2.4.4 Cell Global Identity CGI o Cell ID .....	27
2.5 DISEÑO DE LA RED CELULAR.....	27
2.6 GESTIÓN DE MOVILIDAD .....	29
2.6.1 Actualización de la posición (Location Update) .....	29
2.6.2 Entrega de llamada (Paging) .....	30
2.6.3 Traspaso (Handover) .....	31
2.7 TRÁFICO Y DIMENSIONADO .....	32
2.7.1 Introducción.....	32
2.7.2 Modelos de tráfico telefónico .....	33

<b>CAPÍTULO 3 ESTADO DEL ARTE EN EL ÁMBITO DE LA ESTIMACIÓN DE PARÁMETROS DE TRÁFICO MEDIANTE LA TELEFONÍA MÓVIL.....</b>	<b>37</b>
3.1 INTRODUCCIÓN .....	37
3.2 MATRICES ORIGEN-DESTINO .....	37
3.3 VOLUMEN .....	42
3.4 VELOCIDAD .....	46
3.5 TIEMPOS DE VIAJE.....	50
3.6 CONGESTIÓN.....	53
3.7 DENSIDAD DE TRÁFICO.....	54
3.8 CARACTERÍSTICAS DE LOS DATOS DE TELEFONÍA MÓVIL.....	58
3.8.1 Precisión de localización.....	58
3.8.2 Tamaño de la muestra.....	60
3.8.3 Aspectos legales .....	61
3.9 CONCLUSIONES.....	62
<b>CAPÍTULO 4 INFERENCIA DE VOLÚMENES DE TRÁFICO MEDIANTE LA TELEFONÍA MÓVIL .....</b>	<b>65</b>
4.1 INTRODUCCIÓN .....	65
4.2 ANTECEDENTES Y ESTADO ACTUAL.....	65
4.3 METODOLOGÍA PARA LA INFERENCIA DE VOLÚMENES DE VEHÍCULOS .....	67
4.3.1 Planteamiento del problema .....	68
4.3.2 Conteo de paso de teléfonos entre celdas. Extracción de la muestra.....	71
4.3.3 Análisis descriptivo de los datos de llamadas .....	73
4.3.4 Tratamiento de los datos.....	76
4.3.5 Elaboración de modelos de predicción.....	82
4.3.6 Contraste de modelos .....	97
4.3.7 Conclusiones .....	112

<b>CAPÍTULO 5 APLICACIÓN A LA ESTIMACIÓN DE MATRICES ORIGEN–DESTINO.....</b>	<b>117</b>
5.1 INTRODUCCIÓN.....	117
5.2 ANTECEDENTES Y ESTADO ACTUAL.....	119
5.3 METODOLOGÍA DE AJUSTE DE MATRICES ORIGEN–DESTINO MEDIANTE VOLUMENES AGREGADOS.....	121
5.3.1 Introducción.....	121
5.3.2 Definición del problema.....	122
5.3.3 Formulación matemática.....	124
5.3.4 Algoritmo de ajuste.....	129
5.4 CASOS PRÁCTICOS.....	140
5.4.1 Experimento 1: Red Sioux–Falls.....	140
5.4.2 Experimento 2: Red Comunidad Autónoma de Madrid.....	145
<b>CAPÍTULO 6 CONCLUSIONES Y LÍNEAS FUTURAS.....</b>	<b>153</b>
6.1 CONCLUSIONES.....	153
6.1.1 Introducción.....	153
6.1.2 Evaluación del uso de teléfonos móviles para movilidad de vehículos.....	156
6.1.3 Estudio de mercado.....	159
6.2 LÍNEAS FUTURAS.....	162
<b>BIBLIOGRAFÍA.....</b>	<b>165</b>
<b>APÉNDICE I GLOSARIO.....</b>	<b>173</b>
<b>APÉNDICE II PUBLICACIONES.....</b>	<b>179</b>



# ÍNDICE DE FIGURAS

<b>Figura 2-1:</b> Arquitectura de la red GSM. ....	19
<b>Figura 2-2:</b> Estructura de IMSI (Fuente: Rec. GSM 03.03).....	26
<b>Figura 2-3:</b> Estructura de LAI (Fuente: Rec. GSM 03.03).....	26
<b>Figura 2-4:</b> Estructura de CGI (Fuente: Rec. GSM 03.03). ....	27
<b>Figura 2-5:</b> Definición de celdas y áreas de localización.....	27
<b>Figura 2-6:</b> Antena omnidireccional (izquierda) y directiva – sectores 120° (derecha). ....	28
<b>Figura 3-1:</b> Flujo (rojo) y número de llamadas de teléfonos (verde) en función del tiempo. Izquierda: northbound (desde ciudad), derecha: southbound (hacia ciudad). Todos los datos son promediados sobre intervalos de 15 minutos. (Thiessenhusen <i>et al.</i> , 2006).....	44
<b>Figura 3-2:</b> Ejemplo de ubicación de aforo virtual. ....	44
<b>Figura 3-3:</b> Comparación entre flujo real por un aforo virtual y el estimado tras el ajuste (Caceres <i>et al.</i> , 2006). ....	45
<b>Figura 3-4:</b> Representación de un aforo “virtual”. ....	46
<b>Figura 3-5:</b> Promedio velocidades: teléfonos vs. detectores. Inter-Ciudad (arriba) e Intra-Ciudad (abajo) Motorway Northbound. (Ygnace, 2001) .....	47
<b>Figura 3-6:</b> Velocidad en función del tiempo obtenidos de datos GSM (+), detectores (-), FCD (×) (Thiessenhusen <i>et al.</i> , 2006).....	48
<b>Figura 3-7:</b> Concepto de estimación de velocidad usando doble-handover (Birle <i>et al.</i> , 2006).49	
<b>Figura 3-8:</b> Zona handover (a) con solapamiento típico y (b) con solapamiento debido a edificios (Birle <i>et al.</i> , 2006). ....	50
<b>Figura 3-9:</b> Tiempos de viaje en la autovía de Ayalon, sentido Norte (Bar-Gera, 2007). ....	52
<b>Figura 3-10:</b> Estimaciones de tiempos de viajes y velocidades en función del tiempo de salida: datos GSM vs. LPR (Kummala, 2002). ....	52
<b>Figura 3-11:</b> Relación entre volumen de llamadas y nivel de incidentes en la autovía A7 (agosto y septiembre 2000). Nota: Nivel de incidentes (level of incident): 0 = ningún incidente detectado por los detectores, 14 = al menos 14 detectores (de 21) detectan un incidente en los correspondientes intervalos de 12 min; Volumen de llamadas (call volume): promedio de	

número de llamadas entrantes y salientes registradas en los correspondientes intervalos de 12 min (Ygnace, 2001).....	54
<b>Figura 3-12:</b> Intensidad de tráfico telefónico a 10:00 a.m. en el área metropolitana de Graz....	55
<b>Figura 3-13:</b> Representaciones de la densidad de llamadas de teléfonos móviles en el área metropolitana de Milán (20 km × 20 km) entre las 4 p.m. y las 6 p.m. ....	56
<b>Figura 3-14:</b> Medidas de uso de la red móvil (Erlangs) en 6 lugares de Roma (Reades <i>et al.</i> , 2007). ....	57
<b>Figura 3-15:</b> Localización por Identificación Global de Celda CGI. ....	58
<b>Figura 3-16:</b> Múltiples vías dentro de una celda. ....	59
<b>Figura 3-17:</b> Fronteras entre celdas y LAs. ....	64
<b>Figura 4-1:</b> Llamadas en movilidad: (a) handover y (b) 2 llamadas en $t_1$ y $t_2$ tal que $t_2 - t_1 \leq T = 15 \text{min}$ .....	72
<b>Figura 4-2:</b> Definición de frontera observada o aforo “virtual” en un modelo de red. ....	73
<b>Figura 4-3:</b> Tráfico telefónico (número de llamadas por hora) durante un día en una celda. ....	74
<b>Figura 4-4:</b> Promedio del número de llamadas por hora en diferentes celdas a lo largo de 18 días. ....	75
<b>Figura 4-5:</b> Número de llamadas por hora durante un día según la duración de la llamada $t_d$ ... ..	75
<b>Figura 4-6:</b> Duración media de las llamadas durante un día. ....	76
<b>Figura 4-7:</b> Ejemplos de solapamiento entre celdas.....	77
<b>Figura 4-8:</b> Puntos kilométricos en los que en sus proximidades existe una frontera entre celdas. ....	79
<b>Figura 4-9:</b> Coeficiente de Variación de la distribución de llamadas en diferentes celdas. ....	81
<b>Figura 4-10:</b> Densidad Espectral de Potencia de la distribución de llamadas en una celda. ....	81
<b>Figura 4-11:</b> Diagrama de contorno de isoprobabilidad de la cópula independencia y la empírica.....	84
<b>Figura 4-12:</b> Probabilidad de hacer una llamada a bordo de vehículo. ....	89
<b>Figura 4-13:</b> Probabilidad de handover $P_h$ según el el factor de movilidad $\alpha$ .....	90
<b>Figura 4-14:</b> Evolución de la probabilidad de handover en la celda origen de la frontera $k$ según el intervalo horario $t_i$ . ....	92



<b>Figura 4-15:</b> Evolución del nº de llamadas y de vehículos en una frontera según el intervalo horario. ....	94
<b>Figura 4-16:</b> Factor de intensidad del paso de vehículos según el intervalo horario $t_i$ .....	95
<b>Figura 4-17:</b> Factor de intensidad del paso de vehículos según el intervalo horario $t_i$ .....	96
<b>Figura 4-18:</b> Evolución horaria del error relativo en valor absoluto para cada uno de los modelos. ....	99
<b>Figura 4-19:</b> Evolución horaria del error absoluto para cada uno de los modelos. ....	99
<b>Figura 4-20:</b> Distribución acumulada del valor absoluto del error relativo en porcentajes (arriba). Zoom en torno a los valores de error relativo absoluto 15-30% (abajo). ....	100
<b>Figura 4-21:</b> Test gráfico de normalidad.....	103
<b>Figura 4-22:</b> Histograma del error relativo de cada modelo según el nº de bins definido por Knuth.....	104
<b>Figura 4-23:</b> Histograma del error relativo de cada modelo según el nº de bins definido por Stone.....	104
<b>Figura 4-24:</b> Volúmenes de vehículos observados y estimados con cada modelo que cruzan la frontera 1 y la frontera 2.....	106
<b>Figura 4-25:</b> Volúmenes de vehículos observados y estimados con cada modelo que cruzan la frontera 3 y la frontera 4.....	106
<b>Figura 4-26:</b> Volúmenes de vehículos observados y estimados con cada modelo que cruzan la frontera 5 y la frontera 6 en cada periodo horario.....	107
<b>Figura 4-27:</b> Volúmenes de vehículos observados y estimados con cada modelo que cruzan la frontera 7 y la frontera 8 en cada periodo horario.....	107
<b>Figura 4-28:</b> Volúmenes de vehículos observados y estimados con cada modelo que cruzan la frontera 9 y la frontera 10 en cada periodo horario.....	108
<b>Figura 4-29:</b> Volúmenes de vehículos observados y estimados con cada modelo que cruzan la frontera 11 y la frontera 12 en cada periodo horario.....	108
<b>Figura 4-30:</b> Volúmenes de vehículos observados y estimados con modelos 2, 3 y 6 que cruzan la frontera 1 y la frontera 2.....	109
<b>Figura 4-31:</b> Volúmenes de vehículos observados y estimados con modelos 2, 3 y 6 que cruzan la frontera 3 y la frontera 4.....	110

<b>Figura 4-32:</b> Volúmenes de vehículos observados y estimados con modelos 2, 3 y 6 que cruzan la frontera 5 y la frontera 6 en cada periodo horario. ....	110
<b>Figura 4-33:</b> Volúmenes de vehículos observados y estimados con modelos 2, 3 y 6 que cruzan la frontera 7 y la frontera 8 en cada periodo horario. ....	111
<b>Figura 4-34:</b> Volúmenes de vehículos observados y estimados con modelos 2, 3 y 6 que cruzan la frontera 9 y la frontera 10 en cada periodo horario. ....	111
<b>Figura 4-35:</b> Volúmenes de vehículos observados y estimados con modelos 2, 3 y 6 que cruzan la frontera 11 y la frontera 12 en cada periodo horario. ....	112
<b>Figura 5-1:</b> Red de Sioux Falls, incluyendo distribución ficticia de LAs y celdas. ....	141
<b>Figura 5-2:</b> Correlación entre los volúmenes observados en fronteras (grupos de arcos) y los modelados al asignar la matriz previa (a) y la matriz estimada (b). ....	144
<b>Figura 5-3:</b> Correlación entre la matriz previa y la matriz estimada. ....	145
<b>Figura 5-4:</b> Nodos de la red de Madrid agrupados en 15 LAs. ....	146
<b>Figura 5-5:</b> Correlación entre los volúmenes observados en fronteras (grupos de arcos) y los modelados al asignar la matriz previa (a) y la matriz estimada (b). ....	147
<b>Figura 5-6:</b> Correlación entre los volúmenes de arcos observados (pertenecientes a alguna frontera observada) y los modelados al asignar la matriz previa (a) y la matriz estimada (b). .	149
<b>Figura 5-7:</b> (a) Evolución de la función objetivo durante el proceso de ajuste de la matriz. (b) Zoom para los subproblemas 2–5. (c) Zoom para el término de las restricciones de la función objetivo. ....	151

# ÍNDICE DE TABLAS

<b>Tabla 3-1:</b> Comparativa de volúmenes derivados de handovers o procesos de LU. ....	64
<b>Tabla 4-1:</b> Extracto de los datos entregados por el operador. ....	70
<b>Tabla 4-2:</b> ID fronteras asociados a cada punto kilométrico y sentido. ....	79
<b>Tabla 4-3:</b> Días observados. ....	80
<b>Tabla 4-4:</b> Medidas de error para cada modelo. ....	98
<b>Tabla 4-5:</b> Percentiles del error relativo en valor absoluto para cada modelo. ....	100
<b>Tabla 4-6:</b> Correlación por rango (Spearman) y lineal (Pearson) entre $\{ y_i, \hat{y}_i \}$ para cada modelo. ....	101
<b>Tabla 4-7:</b> Valores de Logaritmo de la verosimilitud ( <i>LLF</i> ), <i>AIC</i> y <i>BIC</i> del conjunto de errores relativos de cada modelo, junto al número de parámetros de cada modelo. ....	105
<b>Tabla 5-1:</b> Fronteras observadas válidas y arcos asociados. ....	142
<b>Tabla 5-2:</b> Comparativa de ecuación de la recta de regresión y $R^2$ , antes y después de ajuste. ....	149
<b>Tabla 6-1:</b> Ejemplos de Aplicaciones según mercados de negocio. ....	161



# CAPÍTULO 1

---

## INTRODUCCIÓN

### 1.1 MOTIVACIÓN

La creciente movilidad existente en la sociedad actual, potenciada por el incremento del uso del automóvil, exige técnicas y mecanismos cada vez más complejos para una correcta gestión y planificación de la red viaria. El objetivo de dicha planificación del transporte es responder adecuadamente a las necesidades de movilidad de la población, por medio del dimensionado de la infraestructura y de la utilización de los medios de transporte, estableciendo los costes de la solución adoptada, las fuentes de los recursos necesarios, y los plazos respectivos de las etapas propuestas. Asimismo, la complejidad de los sistemas de transporte actuales ha mostrado la conveniencia de contar con herramientas analíticas que permitan al planificador disponer de un adecuado conocimiento del sistema y poder hacer predicciones. Un modo de lograrlo es caracterizar las vías en términos de una serie de parámetros como son la velocidad, tiempos de viajes, flujos, etc. La obtención de esos datos de tráfico de una forma precisa y rápida es de vital importancia para la planificación y gestión del tráfico.

Las nuevas tecnologías, cada vez más extendidas en la sociedad, juegan un papel clave para introducir mejoras en la estimación de estos parámetros y así hacer frente a los principales problemas de tráfico: movilidad, saturación y seguridad. Un modo de conseguirlo es mediante la telefonía móvil. Según los últimos estudios de mercado, las tasas de penetración de telefonía celular superan ya el 90% de la población de gran cantidad de países (CDB, 2006). Por ello, parece lógico pensar que cada vehículo puede llevar a bordo algún teléfono. La propia gestión de movilidad que poseen las redes de telefonía celular genera ciertos parámetros para su correcto funcionamiento, tal que un tratamiento apropiado de ellos, sin necesidad de revelar información confidencial de sus abonados, permitirá determinar ciertos datos relativos a la movilidad de vehículos mediante los teléfonos que se lleven a bordo.

La idea del empleo de teléfonos móviles para monitorizar el tráfico rodado no es nueva. Sobre este tema han aparecido múltiples trabajos en los últimos años, los cuales se revisan con más detalle en la Sección 1.4 y el Capítulo 3. Sobre este campo lleva varios años trabajando el Grupo de Transportes de la E.T.S.I. de Sevilla. En uno de los trabajos previos, se realizó un estudio sobre una zona piloto (Cáceres *et al.*, 2005), financiado por el Ministerio de Fomento a través de las *Ayudas a la Realización de Proyectos I+D en el Área de Transportes FOM/486/2003*, el cual demostró la viabilidad de la inferencia de datos de tráfico a través de información disponible de telefonía celular. Para ello se desarrolló una herramienta simuladora que, al tiempo que modelaba la situación de tráfico del escenario de estudio mediante un simulador de tráfico comercial, AIMSUN, generaba unos datos de localización de teléfonos similares a como se consiguen en las redes GSM. Con esos registros de localización “ficticios”, el proceso de estimación se fundamentó en un análisis para controlar el paso por las distintas celdas que cubrían el tramo estudiado, actuando los teléfonos como equipos de medida anónimos e indirectos. Los datos extraídos proporcionaron información de movilidad con la que inferir parámetros como matrices O-D y volúmenes. Además, este simulador permitió la posibilidad de evaluar los resultados inferidos puesto que, mientras generaba los registros de localización oportunos, paralelamente, almacenaba los datos de la situación de tráfico simulada a fin de evaluar la precisión de estimación de esta técnica. Los resultados obtenidos revelaron el potencial de esta nueva técnica al obtener gran precisión en la estimación de datos de tráfico. Este hecho, junto con la existencia de proyectos similares en otros países donde participan operadores de telefonía móvil, impulsó la investigación dentro de este campo de cara a la implementación sobre un escenario real.

## 1.2 ÁMBITO DE LA TESIS

Una de las principales características en la sociedad actual es el continuo incremento en la movilidad de personas y mercancías. Esto conlleva un conjunto de problemas tales como congestión y retraso. Resolver problemas derivados de la movilidad se considera hoy uno de los ejes centrales en la gestión y planificación de la red viaria. Para conocer eficientemente las características del tráfico vehicular existente en la red viaria sería necesario aforar todas las vías de manera casi permanente, lo cual implicaría disponer de una infraestructura considerable cuyo coste de instalación y mantenimiento es poco asumible. Así pues, surge la necesidad de nuevas alternativas que permitan caracterizar el tráfico de un modo rápido, preciso y continuo, sin un excesivo coste. El desarrollo de la telefonía móvil impulsa su uso como alternativa para la gestión de tráfico. Su extendido uso en las sociedades modernas hace considerar a sus abonados

como una muestra extensa de la que extraer datos, convirtiéndose sus terminales en sondas anónimas para caracterizar el tráfico de un modo transparente.

En lo que respecta a la gestión del tráfico de vehículos, una de las fuentes de información más demandadas, con vista a la gestión de la red de infraestructura terrestre y su planificación estratégica, son los datos de movilidad origen–destino. Estos datos, organizados en matrices origen–destino (O–D), proporcionan una estimación del número de personas que se desplazan entre puntos de una determinada red en un periodo de tiempo dado, definiendo el mapa de movilidad. La estimación precisa de estas matrices, para modos de transporte mecanizados, es imprescindible para que la Administración optimice el uso de esta infraestructura desde el punto de vista del usuario (en su utilización diaria) así como desde el punto de vista de las inversiones a realizar, necesarias para la adecuación de estas infraestructuras a las necesidades previstas y pronosticadas en escenarios futuros.

La estimación de una matriz O–D, desde la toma inicial de datos hasta la explotación de los primeros resultados, supone un largo proceso que puede abarcar un periodo de incluso años. Esta restricción, junto con los costes económicos y los medios necesarios para llevarla a cabo, implica que el único usuario posible de dichas matrices sea la Administración. Por ello, una matriz O–D obtenida de una forma automática y, en cierto modo, inmediata supondría una auténtica revolución en los estudios de movilidad encargados por la Administración. Junto a ello, podría dar también una respuesta inmediata a los problemas derivados de la movilidad, permitiendo abordar temas a los que hasta ahora ha dedicado pocos recursos, como la movilidad de fin de semana, movilidad asociada a ocio, movilidad a pie, etc.

Históricamente, las matrices origen–destino han sido estimadas en base a tres metodologías diferentes.

1. Estimación directa: Llevando a cabo un proceso de encuestación, domiciliario o viario (directo sobre una muestra de vehículos o por ficha proporcionada a conductores, o indirecto por captación de números de matrículas).
2. Estimación por modelos de demanda: Haciendo uso de un modelo de distribución (e.g.: modelo de gravedad).
3. Estimación mediante conteos de tráfico: Utilizando conteos de tráfico a fin de actualizar una matriz origen–destino preexistente.

De las tres aproximaciones anteriores, la primera de ellas, que utiliza resultados de la teoría de muestras, es la que proporciona los resultados más fiables. Sin embargo, llevar a cabo una encuesta amplia y detallada, que permita elaborar matrices O–D completas, resulta ser un proceso costoso, desde el punto de vista económico y social (las encuestas a vehículos sobre el viario son altamente distorsionantes, consumen mucho tiempo y dan lugar a resultados con ciertos sesgos).

La segunda de ellas aplica modelos sintéticos, de demanda de viajes, al sistema de transporte actual para obtener estimaciones de los flujos origen–destino. Estos modelos pueden ser calibrados específicamente para el área de estudio o transferidos desde áreas similares.

La tercera alternativa es la más profusamente utilizada en las últimas décadas y ha dado lugar a innumerables contribuciones desde el trabajo pionero de Low (1972). La metodología en la que se basa exige, como prerrequisito, un conjunto de aforos que recogen conteos de tráfico (volúmenes observados) y una matriz O–D previa (obtenida de algún estudio cronológicamente anterior o de una encuestación básica que predefina la estructura de la matriz final a obtener). Los aforos de tráfico es un tipo de información que se recoge de forma automática en un conjunto de puntos asociados a arcos del viario y, por consiguiente, se encuentran disponibles en áreas urbanas a muy bajo coste. De aquí que esta última alternativa presenta las ventajas económicas de eliminar encuestas, las manipulaciones de datos, los chequeos y las validaciones en gran medida. Sin embargo, presenta la desventaja de depender de la infraestructura de detectores (aforos) existente.

El problema de ajustar una matriz de viajes O–D, a partir de mediciones de aforos en una red de tráfico y de una matriz O–D desactualizada, ha suscitado un gran interés en los últimos años. Esto es debido, fundamentalmente, a que la información se puede obtener más rápida y económicamente que la obtenida tradicionalmente mediante la realización de encuestas. Es en este contexto donde se va a centrar esta Tesis, pero introduciendo una novedad respecto a los volúmenes observados. En lugar de los tradicionales conteos de tráfico realizados por los aforos, se utilizarán como volúmenes observados datos de movilidad derivados de la telefonía móvil. Esto implicará el desarrollo de modelos de inferencia de volúmenes de tráfico mediante datos de teléfonos así como algoritmos de ajuste adaptados a las características de esos volúmenes.

### **1.3 OBJETIVOS**

Como ya se ha comentado en líneas anteriores, la obtención de datos de una forma precisa y rápida es una de las principales preocupaciones para la planificación y gestión del tráfico. Entre



esta información de tráfico destacan las matrices origen–destino, las cuales son esenciales con vista a la gestión de la red y su planificación. De las metodologías existentes para la estimación de matrices O–D, la más extendida se basa en el ajuste de una matriz previa (obtenida de algún estudio cronológicamente anterior o de una encuestación básica que predefina la estructura de la matriz final a obtener) mediante conteos de tráfico medidos en un subconjunto de arcos de la red, a fin de encontrar una matriz O–D que, cuando sea asignada a la red, reproduzca los conteos de tráfico observados.

Una posible alternativa, tanto para la determinación de las matrices de movilidad O–D como para la captación de aforos en determinados puntos de la red de infraestructura terrestre (carreteras, viarios,...), consiste en la utilización de ciertos datos de posición de teléfonos móviles, los cuales son conocidos por los operadores móviles para su correcta administración de servicios. Estos terminales no tienen una posición fija sino que pueden moverse libremente. Esto obliga a los operadores a disponer de mecanismos para realizar un seguimiento de sus usuarios y así proporcionarles los servicios contratados lo más rápido posible y con menor coste de señalización.

La idea fundamental reside en que si a bordo de un vehículo hay un teléfono encendido, dicho vehículo puede ser considerado como un “sonda”. Según el informe presentado en Enero de 2009 por la CMT el número de abonados de telefonía móvil en España supera ya el 100 % de la población total. En la Unión Europea, la media se sitúa en torno al 90% (CDB, 2006), siendo esta tendencia similar en la mayoría de países del mundo desarrollado. Dicho crecimiento se refleja en nuestros hábitos ya que el teléfono móvil se convierte en un elemento indispensable en nuestra vida diaria. Este hecho hace pensar que los terminales móviles pueden convertirse en posibles “sondas” de nuestra movilidad puesto que nos acompañan a lo largo de todo el día.

No resulta extraño que cada vehículo lleve a bordo un teléfono móvil, propiedad de su conductor o del resto de sus ocupantes. La propia gestión de movilidad que poseen las redes de telefonía celular ofrece una estimación de la posición de sus abonados, referida a la zona que les proporciona cobertura. De modo que la posición de un vehículo puede ser monitorizada mediante el análisis de los registros de posición que el teléfono que lleva a bordo envía al sistema. El tratamiento de estos datos, sin la necesidad de revelar información confidencial de sus abonados, permitirá determinar ciertos datos relativos al desplazamiento de vehículos mediante los teléfonos que se lleven a bordo.

El objetivo final es presentar las nuevas tecnologías como una alternativa viable para su uso en los estudios de movilidad de transporte, para lo cual será necesario desarrollar:

- a. Una metodología para la inferencia de volúmenes (aforos) en determinadas zonas de la red utilizando información de terminales móviles anónimos a bordo de vehículos.

Para la captación de aforos en determinados puntos de la red, los datos de móviles proporcionarían flujos mediante la determinación del número de móviles que atraviesan los arcos que se desean monitorizar y serían transformados, usando métodos estadísticos de estimación oportunos, a su equivalente en número de vehículos.

- b. Un algoritmo de estimación y ajuste para la actualización de una matriz de movilidad mediante esos volúmenes inferidos de teléfonos móviles.

Conocida una matriz de movilidad ya obsoleta, denominada matriz previa, ésta se ajustaría utilizando técnicas de estimación, mediante volúmenes observados en un conjunto de arcos, adaptadas a los requerimientos que exige utilizar información derivada de teléfonos móviles. Con este nuevo algoritmo, en lugar de utilizar los tradicionales aforos, se utilizarían los volúmenes derivados de los teléfonos móviles que pasan de una celda a otra, transformados en su equivalente en vehículos. Como ya se explica en capítulos posteriores, típicamente una frontera entre celdas comprende varios arcos, de modo que el paso de terminales proporciona volúmenes agregados del grupo de arcos que cruzan la frontera. No obstante, la información de los aforos derivados de teléfonos móviles podría ser complementada con la de aforos tradicionales (espiras, cámaras de vídeo, etc.), dado que éstos son un caso especial en la que la frontera es cruzada por un único arco, dotando al algoritmo de mayor versatilidad.

Una vez que las matrices de movilidad O–D se han inferido mediante el algoritmo citado en el punto precedente, éstas se asignarían al modelo de red existente de la región bajo estudio. Los resultados proporcionados mediante este proceso de asignación serían los flujos de tráfico en cada uno de los arcos de la red.

## 1.4 ANTECEDENTES

La idea de la utilización de móviles para monitorizar las condiciones del tráfico rodado no es nueva. La aparición de legislaciones en diferentes países que exigieron precisión en la localización de llamadas de emergencias (E–112, Europa 2002/22/EC) propició una exhaustiva investigación sobre nuevas técnicas que mejoraran los datos de posición proporcionados por los operadores. Como consecuencia de estos trabajos, surgieron nuevas aplicaciones relacionadas con la información de posición (LBS), muchas de ellas enfocadas al sector del Transporte. Esta

temática está siendo estudiada en numerosos centros de investigación, universidades y departamentos relacionados con el transporte de todo el mundo. Cabe destacar la labor del Department of Civil Engineering (Univ. de Waterloo), con las publicaciones de varios trabajos de investigación (Hellings *et al.*, 2005), y la del Institute of Transportation Studies (Univ. de California, Berkeley), por las múltiples aportaciones relacionadas con la materia, como una revisión del “Estado-del-Arte” (Yim, 2003). No obstante, el creciente interés en este campo ha hecho surgir en los últimos años numerosos proyectos piloto por parte de diferentes entidades. Una revisión detallada sobre proyectos piloto y otros estudios relacionados con la materia aparece en trabajos como Caceres *et al.* (2008) o en un informe presentado por el VDOT, en referencia al proyecto NCHRP 70–01, (Smith y Fontaine, 2006). En un apéndice de este último informe se revisan diferentes proyectos (finalizados y en curso) que emplean sistemas de monitorización de tráfico basados en WLT (Wireless Location Technology). Haciendo uso de información publicada hasta entonces, se elaboró un documento que recopila datos como entidades participantes, objetivos, tecnología usada, etc., de cada uno de esos proyectos, junto con resultados obtenidos caso de haber finalizado.

Sobre este tema se han presentado en los últimos años innumerables publicaciones en revistas y libros. En 2000, se presentó un modelo de estimación de tráfico mediante un algoritmo que estimaba parámetros de tráfico a partir de datos de posición de teléfonos (Bolla *et al.*, 2000). En 2001, se elaboraron diferentes estudios sobre datos anónimos de posición de teléfonos móviles orientados a aplicaciones de transporte, como velocidades o tiempos de viaje (Lovell, 2001, Smith *et al.*, 2001). También se deben destacar estudios sobre estimación de matrices origen–destino mediante registros de teléfonos móviles (Pan *et al.*, 2006) o utilizando simulaciones de datos de telefonía móvil (Caceres *et al.*, 2007, Sohn, 2004, Akin y Sisiopiku, 2002), así como otros estudios orientados a la efectividad de la técnica, como el de Cayford y Johnson (2003), que examinaron los principales parámetros a considerar: precisión, frecuencia de las mediciones y número de localizaciones necesarias para la correcta descripción del tráfico.

De todos ellos, la precisión es el parámetro que posee mayor relevancia para alcanzar los resultados deseados. El nivel de precisión más fino proporcionado por la red GSM sin realizar ninguna modificación en ella ni en los terminales es a nivel de celda, es decir, con el área de cobertura de la estación base a la que está conectado el terminal móvil. Esta precisión depende del radio de la celda, pudiendo ser de 200 m en ciudades y oscilando entre 5 y 20 km en entornos rurales. No obstante, existen otros modelos que poseen mayor precisión pero requieren la modificación de la red (instalando elementos adicionales para realizar mediciones de señal: ángulo de llegada, nivel de señal, etc.) o implican la modificación de los terminales (GPS, relojes para medir tiempos de llegada, etc.). Dado el gasto asociado a estas modificaciones

adicionales así como que la información recopilada debe ser estadísticamente representativa, es decir, que el tamaño de la muestra sea lo mayor posible (la mayoría de los abonados no disponen de terminales modificados), esta Tesis centra su atención en la localización proporcionada por una red GSM típica sin modificar. Este nivel, a pesar de que ofrece una precisión que no es la óptima, es suficiente para una apropiada estimación de ciertos parámetros de tráfico.

En la actualidad, existen multitud de estudios piloto relacionados con esta nueva alternativa. Todos ellos han mostrado magníficos resultados para una posible implementación en la realidad. Sin embargo, pese a ser una fuente de datos fiable y real que permite monitorizar la movilidad de los habitantes de una zona, estos datos de localización agregados no se utilizan de manera sistemática para describir la planificación de sistemas urbanos o interurbanos. Principalmente, esto se debe al acceso restringido a estos datos de localización disponibles por los operadores de telefonía. Además, la carencia de una normativa detallada para el manejo de estos datos por parte de terceros junto con la creciente preocupación asociada a términos de privacidad frenan el desarrollo de esta alternativa. No obstante, numerosos estudios (Ahas y Ülar, 2005, Ratti *et al.*, 2006, y Rose, 2006) realizados en los últimos años están demostrando que la utilización de estos datos no vulnera los derechos de los abonados y que, con la colaboración del operador, se pueden obtener nuevos productos que beneficien no sólo a los operadores sino a toda la sociedad. Incluso se muestra que poseen numerosas ventajas respecto a los métodos anteriores:

- Gran tamaño de la muestra de datos, asociado con la tasa de penetración de telefonía móvil.
- Cualquier zona puede ser monitorizada, dada la extensa cobertura de telefonía.
- No se precisa de instalación adicional de elementos en el vehículo ni en la red para ser monitorizado.
- Datos generados *quasi* en tiempo real.

## ■ PROYECTOS RELACIONADOS CON DATOS DE TRÁFICO DERIVADOS DE TELEFONÍA CELULAR

Existen múltiples proyectos en los que se estudia esta técnica. Algunos de ellos utilizan la propia localización usada por los operadores GSM cuando un teléfono está simplemente encendido, otros emplean ciertos datos asociados a situaciones concretas como llamadas o datos

de facturación, incluso existen proyectos que desarrollan un software propio para captura de señalización de los teléfonos móviles. A continuación se enumeran una serie de proyectos relacionados con la materia. Todos focalizan su atención en el uso de los datos de localización asociados a teléfonos móviles para obtener información sobre desplazamientos de vehículos, respetando en todo momento la privacidad del abonado móvil.

Uno de los más recientes es el desarrollado por el SENSEable City Laboratory del MIT con estudios realizados en diferentes ciudades del mundo como Graz (Austria), Milán o Roma (Italia). En el desarrollado en Graz se consiguió mapear una ciudad en tiempo real, siguiendo a miles de personas que se desplazaban por la ciudad con sus teléfonos móviles encendidos (Ratti *et al.*, 2005). Esto fue posible gracias a la colaboración del operador austriaco A1/Mobikom que entregó datos anónimos de los teléfonos para permitir a los investigadores del MIT desarrollar el proyecto “Mobile Landscapes”. Con ello, se crearon mapas digitales del uso de los terminales en el área metropolitana de Graz, visualizando la dinámica total de la ciudad en tiempo real. Los investigadores cruzaron tres tipos de datos cedidos por el operador: la densidad de las llamadas (medida en Erlang), los orígenes y destinos de éstas (a través de “handovers”), y las posiciones de usuarios muestreados a intervalos regulares. La misma metodología se aplicó a la ciudad de Milán, obteniendo resultados similares (Ratti *et al.*, 2006).

Otro estudio desarrollado por el SENSEable City Laboratory fue en la ciudad de Roma, aunque éste estuvo enfocado a cuestiones de dinámica social de una ciudad, tales como relaciones entre hábitos telefónicos y tipos de actividad asociadas al terreno (áreas residenciales, comerciales, zonas turísticas,...). El proyecto “Real Time Rome” utilizó diferentes tipos de datos procedentes de teléfonos móviles. Uno de ellos fue una medida de uso de ancho de banda de la red, recogidos a nivel de antena (Erlangs). Los otros se correspondían con datos de posición y trayectoria de usuarios llamantes obtenidos mediante la plataforma Lochness del operador Telecom Italia. La investigación se centró principalmente en revelar la similitud entre zonas urbanas basándose en patrones del uso de la red móvil. Agrupando áreas con actividad telefónica similar se clasificaba la ciudad en zonas, de modo que dichas zonas eran comparadas con otras definidas según el tipo de actividad que se desarrollaba en el área (comercial, residencial, turística...). Para ello, se dividió la ciudad en píxeles y utilizó esos datos de uso de la red para promediar valores de Erlang que caracterizaban cada píxel (Reades *et al.*, 2007). Según los valores que tomaba dicha magnitud en cada píxel y aplicando el algoritmo de agrupación K-Means, se obtenían clusters que representaban características espacio-temporales de la ciudad en función del uso del teléfono móvil, por ejemplo, se obtenían diferentes clusters a primera hora de la mañana debido a desplazamientos asociados a la jornada laboral que a última hora de la tarde de un fin de semana, donde los cluster están asociados a actividades de ocio. Por

tanto, el proyecto reveló patrones de movilidad diarios según los hábitos telefónicos de los habitantes de una ciudad.

En Europa, se ha desarrollado un sistema de información sobre tráfico denominado MTS (Mobile Traffic Services). MTS utiliza la información de la red GSM sobre el desplazamiento de los teléfonos móviles (presentes en los vehículos) para monitorizar el tráfico y así ayudar a las autoridades a gestionar el flujo de vehículos por vías y/o posibles congestiones. El piloto se ha realizado en la provincia de Noord-Brabant en Holanda, en colaboración con LogicaCMG, Vodafone y el gobierno local (MTS, 2003). El software base fue desarrollado por Applied Generics, una empresa que proporciona soluciones avanzadas de localización de móviles e información de tráfico a los operadores, cuyo producto RoDIN24 monitoriza una red de telefonía móvil para adquirir localizaciones anónimas en tiempo real de gran cantidad de teléfonos, muchos de los cuales estarán dentro de un vehículo.

En la región de Flandes, el operador de telefonía belga PROXIMUS y la sociedad británica ITIS Holdings han desarrollado también un proyecto piloto que utiliza estas técnicas para el control de tráfico en carreteras a través de las posiciones de teléfonos móviles en los vehículos en movimiento. ITIS recibe del operador PROXIMUS el número de móviles conectados a cada estación base, sin revelar información confidencial de sus abonados, con objetivo de poder determinar, entre otros datos, la velocidad de desplazamiento de cada teléfono celular entre una antena y la siguiente (ITIS Holdings, 2004).

En 1999, TRL (Transport Research Laboratory) participó en un proyecto de actualización de matrices origen-destino con datos de telefonía móvil en el Reino Unido. La investigación encontró que era posible obtener información de matrices origen-destino mediante la localización de teléfonos móviles, junto con otras medidas de información de tráfico tales como tiempos y velocidades del viaje. TRL desarrolló un algoritmo para analizar los datos anónimos de teléfonos procedentes de facturación, proporcionados por el operador BTCellnet (actualmente O2) y convertirlos en información asociada a tráfico (White y Gower, 1999; White y Wells, 2002).

Otro proyecto europeo que ha desarrollado esta técnica es STRIP (System for Traffic Information and Positioning), el cual investigó y validó el uso de la posición de un teléfono móvil para la estimación de tiempos de viaje (Ygnace, 2001). El proyecto fue realizado bajo el programa SERTI (Southern European Road Thelematics Implementation), en colaboración con el gobierno francés y el operador de telefonía SFR (Cegetel Group). Para la obtención de datos de teléfonos se usó el método de localización conocido como “Abis Probing”, que monitoriza la

información de señalización enviada por los interfaces radio Abis y A. STRIP se dividió en dos subsistemas base: uno de captura de datos que recogía mensajes de teléfonos móviles en tiempo real y calculaba su localización; y otro de procesamiento y presentación que calculaba tiempos de viaje en secciones de vía, indicando la dirección de movimiento y mostrando la posición de las llamadas.

En 2002, Finnra (Finnish Road Administration) y el operador de telefonía móvil Radiolinja desarrollaron un sistema donde se capturaron los datos de un conjunto de teléfonos (encendidos) en un cierto tramo para proporcionar estimaciones de tiempos de viaje (Kummala, 2002, Virtanen, 2002). Teniendo en cuenta que cada terminal intercambia información periódicamente con las estaciones base más cercanas, el análisis del tiempo que llevaban los teléfonos comunicándose con estaciones bases permitió estimar tiempos de viaje. A fin de validar la técnica, estas estimaciones se compararon con otras obtenidas mediante identificación de matrícula LPR (License Plate Recognition), mostrando esta nueva alternativa resultados más exactos cuando se analizaba el tráfico en tramos largos (aproximadamente de 10 km) donde la entrada, salida o parada de vehículos es poco frecuente. Junto con ello, se observó que el nuevo sistema no se veía afectado por condiciones atmosféricas adversas, como sí ocurría con los sistemas de LPR. Respecto a la privacidad de los datos, el cifrado de las identidades de los teléfonos impedía que fuera posible conocer qué teléfono realizaba cada viaje detectado; no obstante, tampoco se almacenaba ninguno de los datos después de ser utilizados.

En Alemania, Vodafone (dentro del programa ESGI 2004) propuso ofrecer ciertos servicios de información de tráfico por carretera a sus clientes de telefonía móvil. Para ello, se planteó generar datos de flujo y velocidad de vehículos usando información de señalización que ya está disponible de las operaciones estándar de su red de telefonía móvil. En dicho proyecto se investigó la viabilidad del uso del handover para generar información de tráfico, analizando datos recogidos entre Julio y Septiembre del 2003 (previamente procesados por Vodafone) sobre una autopista en Alemania. Un handover se produce cuando un teléfono con una llamada activa se mueve de una celda a otra y la llamada es transferida desde la antena servidora a la que gestiona la nueva celda, sin interrumpirse la comunicación. Entonces, la captura de información anónima asociada a cada una de esas llamadas traspasadas (como celdas implicadas, instante y duración del handover, junto con algún otro tipo de información de señalización) permitió que fuera posible calcular datos como velocidad o flujo a través de los handover efectuados entre celdas, asociando previamente carreteras a cada una de las celdas y realizando cálculos oportunos de ubicación de fronteras de dichas celdas (Smith Institute, 2006).

Pero no sólo en Europa se están desarrollando proyectos relacionados con esta técnica, la empresa israelita Decell Ltd. tiene desarrollado un producto denominado AutoRoute-1, el cual lleva a cabo un muestreo periódico y aleatorio de teléfonos móviles que viajan a bordo de vehículos en movimiento en una determinada zona. Según estas muestras de móviles registrados en una red sobre un lugar concreto, se estiman ciertos parámetros de tráfico (como velocidad, densidad o flujos) para cada tramo de la carretera (Decell-Autoroute, 2006). Este sistema convierte datos que constantemente son generados dentro de una red móvil en información precisa y en tiempo real de tráfico por una carretera. Además provee a los conductores de rutas recomendadas que corresponden al tiempo de recorrido previsto para alcanzar un destino concreto según la ubicación estimada. Los conductores pueden también recibir alarmas instantáneas sobre incidencias relevantes de tráfico tales como accidentes, obras, etc. AutoRoute-1 genera información en tiempo real de tráfico mientras se respeta la privacidad y anonimato de sus abonados móviles. Ello se asegura teniendo en cuenta que toda la información de localización es almacenada y procesada dentro del servidor AutoRoute-1, el cual físicamente se establece dentro del entorno del operador. De forma que, al igual que pasa con información de facturación, estos datos no pueden emitirse al exterior y son inaccesibles a cualquier persona fuera de la organización del operador.

En EE.UU. existen diferentes proyectos relacionados con esta técnica. La mayoría de ellos están desarrollados por los Departamentos de Transporte de diferentes estados junto con la colaboración de empresas y operadores de telefonía. Uno de ellos es el que desarrolló VDOT (Virginia Department of Transportation) junto con AirSage para demostrar la viabilidad del empleo de datos de teléfonos para la vigilancia del tráfico (VDOT, 2005). Para ello, AirSage elaboró un software que recopilaba información de señalización asociada a localización enviada automáticamente por teléfonos que están simplemente encendidos o en llamada para controlar su movimiento (Airsage, 2007). Esa información permitía conocer la posición de teléfonos anónimos en ciertos instantes la cual, una vez procesada y representada sobre un mapa, mostraba la velocidad con la que fluctuaba el tráfico en cualquier carretera dada por la que existieran vehículos con teléfonos a bordo.

Otro proyecto americano fue el desarrollado en Maryland por Delcan Corp. junto con ITIS Holdings el cual construía un mapa en tiempo real del tráfico a lo largo de sus principales carreteras usando los movimientos de teléfonos (Delcan, 2005). Este proyecto tenía como objetivo principal el conocer el estado del tráfico en ciertas vías para enfocarlo a la detección de eventos, tales como congestión o accidentes, sin necesidad de usar sensores, cámaras u otros elementos que requieren de instalación adicional de nuevos equipos sobre el terreno. En lugar



de estos elementos se utilizaron los teléfonos móviles, propiedad de los conductores, como sondas anónimas. De esta manera se logra obtener información sobre cualquier punto de la red viaria en el que existan teléfonos, ya que los operadores conocen aproximaciones de las localizaciones de sus abonados en términos de antena (celda) a la que deben encaminar las llamadas. La información tomada de los móviles se trató anónimamente, evitando la posibilidad de rastreo de personas concretas en su movimiento hacia un destino. Dicha información pertenecía tanto a señales radio que los teléfonos envían a las estaciones base como a señales derivadas del proceso de movimiento de teléfonos con llamada activa de una estación base (antena) a otra, denominado handover. Entonces, conociendo esa información junto con los mapas de carreteras se podía conocer la ubicación aproximada de los teléfonos y cómo de rápido se estaban moviendo.

Estos son algunos de los proyectos desarrollados en los últimos años. Todos ellos centran su atención en un objetivo común, el uso de los datos de localización de teléfonos móviles, los cuales son ya conocidos por los operadores, para obtener información sobre desplazamientos de vehículos. La mayoría de ellos se han desarrollado con la colaboración de operadores móviles sin necesidad de vulnerar los derechos de intimidad de sus abonados. No obstante, la principal barrera a superar para la implantación completa de esta alternativa consistirá en el establecimiento de acuerdos con los operadores para la utilización de los sistemas de telefonía móvil en aplicaciones de valor añadido como son la gestión del tráfico y su planificación. Estos acuerdos pueden verse favorecidos cuando los operadores descubran el rendimiento económico real del uso de sus sistemas para este tipo de servicios enfocados al tráfico, sin afectar a su funcionamiento.

## **1.5 ESTRUCTURA DE LA TESIS**

El documento final de esta Tesis Doctoral se ha dividido en 6 capítulos.

En el Capítulo 1 se ha expuesto la temática en la cual se enmarca la Tesis, realizándose una introducción al problema de estimación de matrices de viajes origen–destino en redes de tráfico actuales y procediéndose a revisar los antecedentes sobre el uso de la telefonía móvil en la monitorización del tráfico.

En el Capítulo 2 se describe el funcionamiento de las redes de telefonía móvil, revisándose con especial detenimiento el tratamiento seguido para gestionar la movilidad de sus abonados a fin de comprender la utilidad de la telefonía móvil en los estudios de transporte.

En el Capítulo 3 se revisa el Estado-del-Arte, procediéndose a presentar los conocimientos actuales sobre la estimación de parámetros de tráfico mediante el uso de información procedente de los sistemas de telefonía móvil.

El Capítulo 4 está dedicado a desarrollar una metodología para la inferencia de volúmenes de tráfico de vehículos mediante el flujo de teléfonos móviles que atraviesan una determinada zona. En el capítulo se lleva a cabo una revisión de los procedimientos actuales de obtención de volúmenes en arcos, enfocando el problema al caso del nivel de precisión disponible de los sistemas de telefonía móvil. Para el desarrollo de la metodología, se propone el uso de modelos matemáticos que permita expresar en forma cuantitativa relaciones de dependencia de vehículos y teléfonos móviles de modo que puedan ser usados con fines predictivos. En particular, se pretende que los modelos planteados proporcionen medidas admisibles de volúmenes de vehículos sin necesidad de instalar una estación de aforo sobre la zona. El capítulo se completa con un ensayo numérico sobre diferentes fronteras entre celdas en el que se evalúa el rendimiento de los modelos propuestos mediante un análisis comparativo entre ellos, contrastando los volúmenes estimados con medidas proporcionadas por estaciones de aforo ubicadas en puntos próximos a dichas fronteras, y seleccionando aquellos modelos que mejor balance alcancen entre todos los criterios evaluados.

Para completar el estudio propuesto para esta Tesis, en el Capítulo 5 se aborda el problema de estimación de matrices Origen-Destino mediante un algoritmo cuya principal novedad es el uso de datos de volúmenes agregados sobre grupos de arcos como sería la información inferida mediante la telefonía móvil. En este capítulo se propone una metodología de ajuste de una matriz previa mediante un modelo de estructura binivel empleando la adaptación a nivel de grupos de arcos de un algoritmo robusto y eficiente que trabaja a nivel clásico de volumen de arco. La elección de dicho algoritmo como fundamento reside en la importancia de dos criterios tomados a la hora de emprender la resolución del problema: i) que la matriz resultante reproduzca, al ser asignada a la red, los datos de volúmenes agregados lo más fielmente posible y ii) que la información contenida en la matriz O-D previa no se vea distorsionada en exceso durante el ajuste, puesto que una distorsión excesiva de dicha información invalida cualquier esfuerzo presupuestario y humano dedicado a tareas de encuestación. Para analizar la consistencia y eficiencia del algoritmo desarrollado se completa el capítulo con un estudio numérico sobre dos redes: una de pequeña escala pero muy conocida en los estudios de transporte, la red de Sioux Falls. Y otra que se correspondería con la red de la Comunidad de Madrid, para la cual también se dispone de datos de aforos tradicionales (a nivel de arco), de

modo que es posible realizar una comparación del algoritmo propuesto frente a otros que utilizan aforos a nivel de arco.

Finalmente, en el Capítulo 6 se agrupan las conclusiones y se sugieren posibles líneas para futuras investigaciones.



# CAPÍTULO 2

---

## REDES DE TELEFONÍA CELULAR: GSM

### 2.1 INTRODUCCIÓN

Las redes telefónicas móviles surgieron como un nuevo servicio que facilitaba la comunicación entre usuarios sin requerir una conexión permanente (fija) con la red, de forma que el usuario conseguía una más amplia libertad de movimiento, pudiendo acceder, vía radio, a centrales de conmutación a través de estaciones base radioeléctricas distribuidas por la zona de servicio.

Como consecuencia, un usuario que va a realizar una llamada accede a la red a través de la estación que gestione la zona de cobertura en la que se encuentre. Por otro lado, cuando se recibe una llamada, la red debe conocer, en todo momento, dónde está el usuario destinatario de la llamada para que ésta pueda ser encaminada correctamente. Por ello debe realizar un seguimiento de forma automática que permita encontrar la estación base (celda) que da servicio a cada teléfono con el menor retraso posible. Bajo este razonamiento, un teléfono móvil que esté encendido, independientemente de que esté haciendo uso de la red (llamada, sms, etc.), puntualmente informa a la red de su posición actual en términos de la zona de cobertura en la que se encuentra. Este hecho permite que un sistema de telefonía móvil tenga continuamente localizados a sus abonados para proporcionar los servicios contratados lo más rápido posible y con el menor coste de señalización.

A esta propiedad de la red se le llama localización, la cual debe ser dinámica, ya que, por definición, el usuario puede moverse por la red, y debe actualizarse en función del desplazamiento y del tiempo. Asimismo, como consecuencia de esa movilidad, un usuario puede cambiar de zona de cobertura de las estaciones base de la red. En esos casos, debe asegurarse la continuidad de las llamadas en curso de forma automática, para lo cual existe la operación de traspaso de llamadas o handover.

El dinamismo característico de la vida moderna está modificando el estilo de vida tanto a nivel individual como colectivo de la sociedad actual. Como consecuencia, surge la necesidad de

soluciones cada vez más avanzadas que permitan la comunicación entre personas mientras éstas se están desplazando por la red. Este hecho fomenta que los sistemas de telefonía móvil estén experimentando un rápido crecimiento, especialmente impulsado por la característica de permitir establecer cualquier tipo de comunicación con la red (llamadas, sms, acceso a Internet...), independientemente de la posición y movimiento del terminal. Para garantizar esta movilidad de los abonados manteniendo la calidad de los servicios, una red de telefonía dispone de un sistema completo con funciones de acceso, transmisión, conmutación y señalización, que proporcionan servicios de telecomunicación a usuarios a través de acceso vía radio.

En los siguientes apartados se presentan conceptos básicos relacionados con los sistemas de telefonía móvil, revisando su arquitectura de red y sus principales elementos. Además se introducen nociones básicas respecto a la planificación y dimensionado de la red, describiendo las principales funcionalidades asociadas a la gestión de movilidad a fin de comprender mejor cómo puede ayudar la telefonía móvil en la inferencia de parámetros asociados al tráfico. Para su elaboración se ha utilizado la siguiente bibliografía: Gibson (1999), Hernando Rábanos (2002), Huidobro (2002), Sallent *et al.* (2003), Rappaport (2002) y Ilyas y Mahgoub (2005).

## **2.2 ARQUITECTURA DE LA RED GSM**

En lo que se refiere a la estructura básica del sistema GSM, éste se organiza como una red de estaciones base radioeléctricas que proporcionan cobertura completa al área de servicio. Cada estación base (BTS) tiene asociada una zona de cobertura, denominada celda, que opera en un conjunto de canales de radio diferentes a los usados en las celdas adyacentes y que se encuentran distribuidas según un plan celular. Un grupo de BTSs se encuentra conectado a un controlador (BSC), encargado de aspectos como el handover (traspaso de llamada de una estación base a otra) o el control de potencia de las BTSs y de los móviles. En consecuencia, el BSC se encarga del manejo de toda la red de radio y supone una auténtica novedad respecto a los anteriores sistemas celulares.

Una o varias BSCs se conectan a una central de conmutación de móviles (MSC). Éste es el corazón del GSM como responsable de la inicialización, enrutamiento, control y finalización de las llamadas, así como de la información sobre la tarificación. Es también el interfaz entre diversas redes GSM o entre una de ellas y las redes públicas de telefonía o datos. La información referente a los abonados se encuentra almacenada en dos bases de datos que se conocen como registro de posiciones base (HLR) y registro de posiciones de visitantes (VLR).

Los diferentes elementos que constituyen una red de telefonía celular GSM se clasifican en cuatro grupos fundamentales:

1. **MS (Mobile Station).**
  - ME (Mobile Equipment).
  - SIM (Subscriber Identity Module).
2. **BSS (Base Station Sub-System).**
  - BSC (Base Station Controller).
  - BTS (Base Transceiver Station).
3. **NSS (Network and Switching Subsystem).**
  - MSC (Mobile Switching Center).
  - GMSC (Gateway Mobile Switching Center).
  - VLR (Visitor Location Register).
  - HLR (Home Location Register).
  - AuC (Authentication Center).
  - EIR (Equipment Identity Register).
4. **NMC (Network Management Center).**
  - OMC (Operation and Maintenance Center).

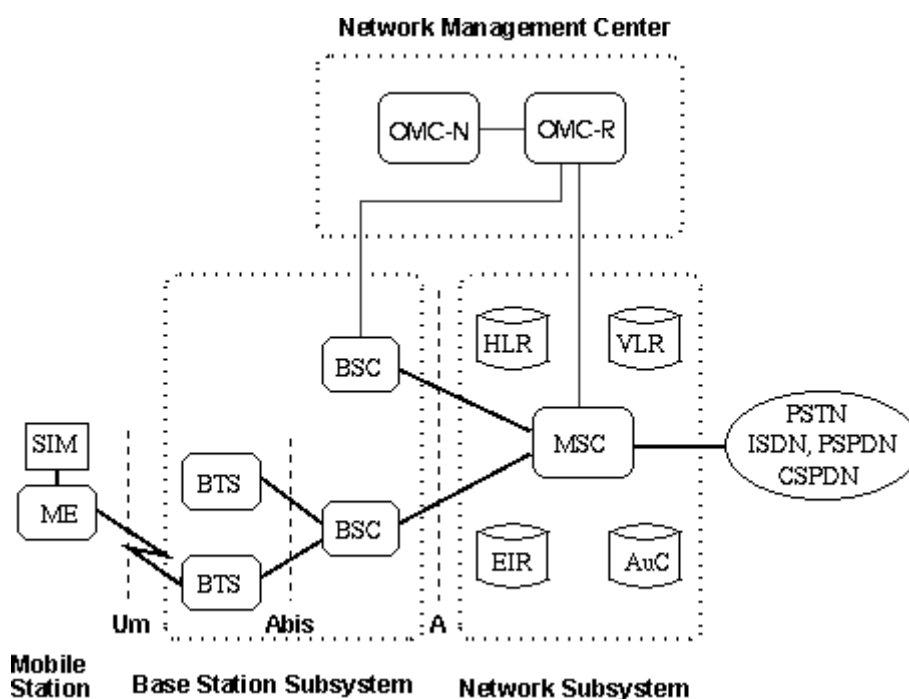


Figura 2-1: Arquitectura de la red GSM.

A continuación se describe cada elemento en base a la clasificación anterior.

**1.– Estación Móvil (Mobile Station, MS).** Consta de dos elementos básicos: el equipo móvil (ME) y la tarjeta SIM (Subscriber Identity Module).

- Mobile Equipment (ME)

Es el dispositivo móvil y se identifica dentro de cualquier red GSM por el International Mobile Equipment Identity (IMEI). El IMEI es un número de 15 cifras y tiene la siguiente estructura:

IMEI = TAC / FAC / SNR / sp donde:

TAC = Type Approval Code (6 CIFRAS).

FAC = Final Assembly Code, identifica al fabricante (2 cifras).

SNR = Serial Number (6 cifras).

sp = Cifra suplementaria de reserva (1 cifra).

Los equipos tienen la peculiaridad de variar la potencia de emisión de señal sobre la celda en la que se encuentren con el objetivo de poder mantener en todo momento el nivel de potencia de transmisión óptimo, limitando interferencias co-canal inducidas sobre celdas adyacentes y reduciendo el consumo. Esta potencia de emisión es característica del móvil, de manera que su nivel de potencia máximo es importante dado que implica un aumento de la capacidad del terminal para poder alejarse de la estación base y poder seguir disfrutando del servicio.

- Subscriber Identity Module (SIM)

La SIM es una pequeña tarjeta inteligente dotada de memoria y microprocesador que sirve para identificar al abonado. Esta tarjeta se inserta en el interior del móvil y permite al usuario acceder a todos los servicios disponibles por parte de su operador. Sin la tarjeta SIM, el terminal no tiene utilidad porque no puede hacer uso de la red. Puesto que la tarjeta SIM permite identificar al abonado independientemente del terminal usado, se proporciona al usuario la posibilidad de continuar recibiendo y efectuando llamadas y utilizar todos los servicios que tiene contratados insertando la tarjeta SIM en otro terminal.

La SIM está protegida por un número de cuatro dígitos que recibe el nombre de PIN (Personal Identification Number). Una vez que se introduce el PIN en el terminal, éste activa el proceso de búsqueda de las redes GSM que estén disponibles para intentar validarse en ellas. Una vez que la red valida el terminal, el teléfono queda registrado en la celda que lo ha validado. Es importante destacar que la red que valida el terminal es la que se tiene contratada.



La tarjeta SIM contiene el IMSI (International Mobile Subscriber Identity), usado para identificar y validar al abonado en cualquier sistema GSM. Además, contiene los procedimientos de criptografía para la confidencialidad de la información del usuario o las memorias del teléfono (alfanuméricas o para mensajes de texto, SMS).

**2.– Subsistema de estaciones base (Base Station Subsystem, BSS):** Sirve para conectar a las estaciones móviles con los NSS (Network Sub-System), además de ser los encargados de la transmisión y recepción. Constan de dos elementos diferenciados: la estación base o BTS (Base Transceiver Station) y la BSC (Base Station Controller).

- Base Transceiver Station (BTS)

La BTS alberga todos los receptores/transmisores que dan servicio a una celda, los cuales están encargados de recibir y enviar información al canal radio, abasteciendo una interfaz física entre la estación móvil y la BSC. Las antenas usadas en cada celda de la red pueden estar situadas en el centro de la celda (omnidireccional) o directivas (sectores de 120°). Normalmente la potencia de transmisión determina el tamaño de la celda.

- Base Station Controller (BSC)

Se encarga de la gestión de varias BTSs en lo relativo a los recursos de radio: asignación, utilización y liberación de las frecuencias, handovers y funcionamiento con saltos de frecuencias, controlando la conexión entre las BTSs y las MSCs (centrales de conmutación que proporcionan la conexión con la red física y otras redes).

**3. Subsistema de conmutación y red (Network and Switching Subsystem, NSS):** Es el sistema encargado de administrar las comunicaciones que se realizan entre los diferentes elementos de la red. Para poder realizarlas, la NSS se divide en 7 sistemas diferentes, cada uno con una misión dentro de la red:

- Mobile Services Switching Center (MSC): Es el componente central del NSS y se encarga de las labores de conmutación para estaciones móviles situadas en su área de influencia, así como de proporcionar la conexión con otras redes. Además, proporciona las funciones adicionales necesarias para sustentar la movilidad y organizar la asignación de los recursos radioeléctricos. Entre esas funciones se destacan las siguientes:
  - Procedimientos para la localización y registro de abonados y su actualización.

- Gestión de llamadas y aviso a la estación móvil de llamadas entrantes.
  - Procedimientos para el traspaso de llamadas (handovers).
  - Recopilación de datos de tráfico y facturación.
- 
- Gateway Mobile Services Switching Center (GMSC): Es una MSC que además es utilizada para dirigir hacia la red móvil las llamadas originadas en la red fija. De modo que actúa como dispositivo “pasarela” que se encarga de servir de mediador entre las redes de telefonía fijas y las móviles, haciendo que los protocolos de comunicaciones que existen en ambas redes se entiendan. Esta central se encarga de interrogar al HLR adecuado para conocer la posición del móvil al que va dirigida la llamada para, posteriormente, encaminar la llamada hacia la MSC correspondiente.
  - Home Location Register (HLR): Es una base de datos que contiene información sobre los usuarios conectados a un determinado MSC. Analiza información relativa a los niveles de suscripción, servicios suplementarios y localización actual o más reciente de los móviles que pertenecen a la red local de ese MSC. El HLR funciona conjuntamente con el VLR.
  - Visitor Location Register (VLR): Contiene información sobre los niveles de suscripción, servicios suplementarios y localización para un abonado que se encuentra o al menos se encontraba recientemente visitando el área gestionada por el MSC. Esta base de datos dispone también de información relativa a si el abonado se encuentra activo o no, lo que evita el uso improductivo de la red (envío de señales a una localización que se encuentra desconectada).
  - Authentication Center (AuC): Asociado al HLR trabaja el centro de autenticación (AuC), que contiene la información por la que se comprueba la autenticidad de las llamadas con el fin de evitar los posibles fraudes, la utilización de tarjetas de abonado (SIMs) robadas o el disfrute del servicio por parte de morosos. También se encarga de soportar funciones de encriptación.
  - Equipment Identity Register (EIR): También es un elemento para la seguridad de las redes GSM pero a nivel de validez de terminales. La EIR contiene una base de datos con la identificación de todos los terminales que son válidos para ser usados en la red. De manera que se puede impedir que se realice una llamada cuando se detecte que ha sido robado, pertenece a algún modelo no homologado o sufre algún fallo susceptible de afectar negativamente a la red.

**4. Network Management Center (NMC):** Es un sistema de operación que constituye la máxima jerarquía dentro del sistema de explotación. De este centro dependen todos los demás centros de operación y mantenimiento (Operation and Maintenance Center, OMC). Estos centros tienen las siguientes funciones:

- Acceso remoto a todos los elementos que componen la red GSM (BSS, MSC, VLR, HLR, EIR y AuC).
- Gestión de las alarmas y del estado del sistema con posibilidad de efectuar varios tipos de pruebas para analizar las prestaciones y verificar el correcto funcionamiento del mismo.
- Recogida de todos los datos relativos al tráfico de los abonados necesarios para la facturación.
- Supervisión del flujo de tráfico a través de las centrales.
- Visualización de la configuración de la red con posibilidad de cambiarla por control remoto.
- Administración de los abonados y posibilidad de poder conocer su posición dentro del área de cobertura.

En redes de telefonía de grandes dimensiones pueden existir más de una OMC. En este caso existirá un OMC general desde el que es posible controlarlo todo (OMC–N) y otros OMC limitados al control de algunas zonas (OMC–R).

## **2.3 BASES DE DATOS HLR Y VLR**

### **2.3.1 Introducción**

En los sistemas de telefonía móvil, dado que los terminales pueden moverse libremente por la red, es necesario realizar un seguimiento automático de los mismos para poder encontrar la celda donde se encuentra con el menor retardo posible, lo cual es fundamental, por ejemplo, en el caso de una llamada entrante. Para lograrlo se emplea una base de datos jerárquica de dos niveles formados por el HLR y varios VLRs. Dada la importancia de estas bases de datos para las funciones de movilidad implementadas en las redes de telefonía móvil, a continuación se describe con más detalle aspectos significativos relacionados con estas bases de datos.

### 2.3.2 Home Location Register (HLR)

Cuando se inscribe un nuevo abono de usuario a la red GSM, todas las informaciones para su identificación se memorizan en el HLR de la MSC asociada al área en la que se encuentre en ese momento inicial. El HLR contiene información administrativa (nivel de subscripción, servicios suplementarios, etc.) de cada abonado asignado al mismo, junto con los datos de localización, incluyendo información sobre la posible área visitada a efectos de enrutar llamadas destinadas al mismo. Para ello, requiere trabajar conjuntamente con los VLRs. La información contenida en el HLR puede clasificarse como:

- Datos permanentes, entre los que figuran la identificación del abonado (IMSI y MSISDN), información relativa a teleservicios, servicios portadores y servicios suplementarios (con sus parámetros) contratados, junto a limitaciones de servicio.
- Datos temporales (actualizables), que son los relativos a la localización del terminal a fin de encaminar las llamadas entrantes al MSC donde está registrado. Esta información comprende el MSRN, direcciones del VLR y del MSC e identidad temporal (TMSI) que tiene asignado el terminal.

El HLR, como todas las bases de datos de la red, está implementado en una estación de trabajo cuyas prestaciones (memoria, procesadores, capacidad de los discos) son actualizables cuando crece el número de abonados. Asociado al HLR trabaja el centro de autenticación (AuC), que contiene la información por la que se comprueba la autenticidad de las llamadas con el fin de evitar los posibles fraudes, la utilización de tarjetas de abonado (SIMs) robadas o el disfrute del servicio por parte de impagados.

### 2.3.3 Visitor Location Register (VLR)

El VLR almacena, de modo temporal, información de todos los abonados que en un momento dado están registrados dentro de su área de servicio. Esta información es requerida y obtenida a partir de los datos contenidos en el HLR del que depende el usuario. En general, para simplificar las señalizaciones requeridas y la estructura del sistema, los fabricantes implementan el VLR y el MSC juntos, de modo que el área geográfica controlada por el MSC corresponde a la controlada por el VLR.

Cuando un terminal entra en la zona controlada por un MSC, esta central chequea en su base de datos asociada para determinar si el usuario está en su "Home Area" (donde se registró por primera vez), o en un "Roaming Area", es decir, que el usuario se encuentra en modo

itinerancia. En el caso de un usuario itinerante, se notifica esa situación al VLR correspondiente a ese área visitada. El terminal ejecuta entonces un proceso de inscripción o registro y recibe una dirección de visitante que se traduce en un número de ruta, MSRN (Mobile Subscriber Roaming Number), que sirve para encaminar las llamadas destinadas a él hacia el área donde se encuentra actualmente el móvil. Además, es necesario enviar esta información hacia el HLR asociado al usuario (el de su Home Area).

El VLR contiene también otras informaciones necesarias para el tratamiento de las llamadas entrantes/salientes, las cuales se obtienen mediante un diálogo entre el VLR y el HLR. La información básica contenida en el VLR, toda ella de carácter temporal, es la siguiente:

- Identidad permanente del abonado (IMSI).
- Identidad temporal del abonado (TMSI).
- Número telefónico del abonado (MSISDN).
- Número de encaminamiento (MSRN).
- Identidad de la zona de localización (LAI).
- Servicios contratados por el usuario (servicios suplementarios, restricciones...).

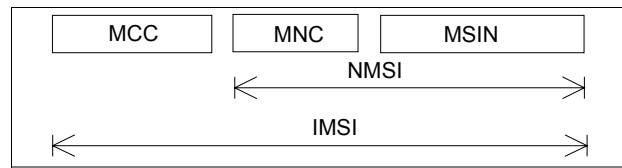
Esta base de datos dispone también de información relativa a si el abonado se encuentra activo o no, lo que evita el uso improductivo de la red (envío de señales a un terminal desconectado).

## **2.4 IDENTIDADES Y NÚMEROS DE USUARIO Y RED**

Para los procesos de autenticación de los usuarios y encaminamiento de las llamadas, se ha establecido un conjunto de números asociados a los usuarios y a los elementos de la red. A continuación se describen los más relevantes.

### **2.4.1 Internacional Mobile Station Identity (IMSI)**

El IMSI se almacena en la tarjeta SIM y se asigna al abonado en el momento de suscribirse. Es un número interno de red que identifica al abonado para el acceso a las bases de datos. Como este número se difunde por el interfaz radio y puede ser captado, para preservar la confidencialidad de la identidad, la red asigna una identidad ficticia al abonado, denominada TMSI (Temporary Mobile Subscriber Identity), una vez que éste ya está registrado. De modo que sólo se transmite por radio si es necesario. Se compone de un máximo de 15 dígitos y tiene la siguiente estructura:



**Figura 2-2: Estructura de IMSI (Fuente: Rec. GSM 03.03)**

Se aprecia que está compuesto de tres elementos:

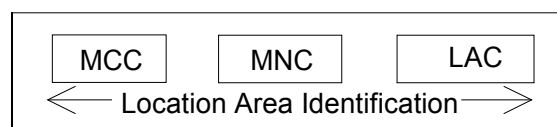
- i) Mobile Country Code (MCC): tres dígitos que identifican el país de domicilio del abonado móvil (Ej.: 214 para España).
- ii) Mobile Network Code (MNC): dos dígitos para usos GSM. EL MNC identifica al operador (GSM/PLMN) dentro del país del abonado móvil (Ej.: 01 para Vodafone).
- iii) Mobile Station Identification Number (MSIN) identifica al abonado móvil dentro de un operador (GSM/PLMN), máximo 10 cifras.

#### 2.4.2 Temporary Mobile Subscriber Identity (TMSI)

Es la identidad temporal que se le da al abonado móvil para evitar enviar continuamente el IMSI por el interfaz radio. Consiste en un identificador ficticio que tendrá el abonado durante un intervalo de tiempo. El TMSI tiene importancia únicamente local, es decir, identifica a un terminal únicamente dentro del área controlada por el VLR y, por tanto, del MSC. Su significado temporal es crear la correspondencia entre TMSI e IMSI, de modo que cuando el móvil se desplaza al área controlada por otro MSC, se le tiene que asignar un nuevo TMSI.

#### 2.4.3 Location Area Identification (LAI)

Un área de localización (LA) es una zona geográfica cubierta por una agrupación de celdas, siendo una celda la zona en la cual una estación base proporciona los servicios. Las áreas de localización se identifican mediante un número denominado LAI (Location Area Identity), que tiene la estructura mostrada en la Figura 2-3, donde MCC y MNC tienen el mismo significado que en el IMSI. El LAC es un código que identifica el LA dentro de la red.



**Figura 2-3: Estructura de LAI (Fuente: Rec. GSM 03.03)**

### 2.4.4 Cell Global Identity (CGI)

La celda dentro de un LA se identifica unívocamente mediante el parámetro CI (Cell Identity). Para identificar la celda dentro de la red se emplea el parámetro CGI o Cell-ID. Este parámetro es la identidad global de celda y se emplea para conocer la celda con la que la estación móvil tiene actualmente contacto radio o con la que tuvo el último contacto radio. El parámetro Cell ID o CGI se obtiene concatenando la identificación de LA (LAI) y la identidad de celda (CI).

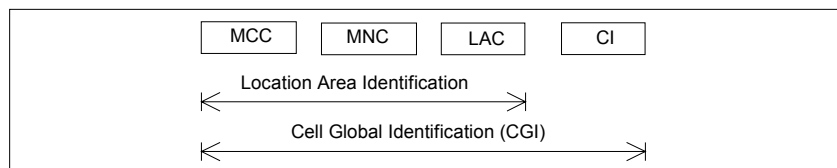


Figura 2-4: Estructura de CGI (Fuente: Rec. GSM 03.03).

## 2.5 DISEÑO DE LA RED CELULAR

Puesto que el ámbito geográfico de cobertura puede ser muy grande, se introduce el concepto de área de localización (LA), que consiste en una zona geográfica cubierta por una agrupación de celdas, siendo una celda la zona en la cual una estación base proporciona los servicios.

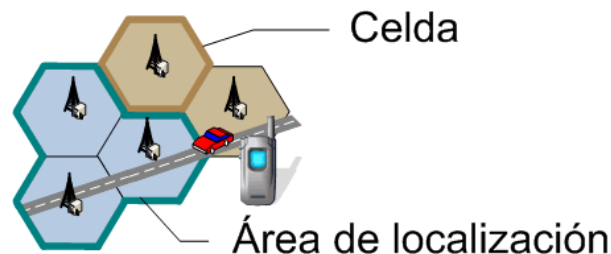
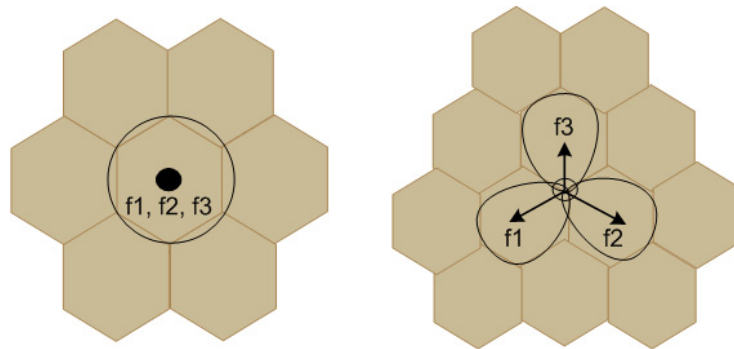


Figura 2-5: Definición de celdas y áreas de localización.

La forma de las celdas usada en la planificación radio depende del tipo de antena y de la potencia emitida por cada estación base. En los modelos de planificación, las coberturas de estaciones base se idealizan mediante hexágonos que recubren el plano sin superposición. Normalmente se usan dos tipos de antenas, las de diagrama horizontal omnidireccional y las directivas.

Si se usan antenas omnidireccionales, idealmente el área de cobertura será circular, y se utiliza un círculo para modelar dicha área. En un contexto real, si se pretende cubrir una determinada zona con antenas omnidireccionales (círculos), se producirá solapamiento entre ellos, lo cual es

poco eficiente desde el punto de vista de uso del espectro pues en la zona de superposición se atendería el tráfico con más frecuencias de las necesarias. Por otro lado, para reducir la interferencia y obtener mayor ganancia que favorezca la comunicación en entornos de cobertura difícil, como son los de naturaleza urbana e interiores de edificios, se utilizan las antenas directivas. En este caso, en cada emplazamiento se disponen tres antenas con diagramas de radiación horizontal que abarcan  $120^\circ$ , como se ve en la Figura 2-6.



**Figura 2-6: Antena omnidireccional (izquierda) y directiva – sectores  $120^\circ$  (derecha).**

Como se ha comentado en líneas anteriores, idealmente, las superficies cubiertas por cada antena se asimilan a hexágonos, y se les suele llamar sectores. Cada sector tiene sus propias frecuencias y es servido por sus propios transmisores, de forma que los sectores son celdas pero “iluminadas” desde un vértice. Esta disposición reduce la interferencia producida ya que las antenas directivas radian muy poca potencia hacia atrás y en direcciones separadas angularmente, por lo que la interferencia será pequeña en esas direcciones.

Con la incorporación de los sistemas de telefonía de generaciones más avanzadas (2.5G/3G), surgen nuevos y más completos servicios a proporcionar a los abonados. Los sistemas móviles GSM, GPRS y UMTS comparten una estructura de red para la conmutación de paquetes sustancialmente equivalente, ya que UMTS hereda la arquitectura creada para GSM y GPRS. No obstante, los nuevos servicios ofrecidos implica la incorporación de nuevos elementos en la red para lograr una optimización en la localización de los terminales en los diferentes servicios existentes. Al igual que ocurre para servicios de voz, estas redes deben mantener actualizada la información sobre la posición de cada teléfono para conocer hacia dónde enrutar los paquetes de datos que le llegan. En esos casos, se precisa conocer con mayor exactitud el conjunto de celdas sobre el que difundir el mensaje de aviso a un terminal para el envío de paquetes de datos. Con objeto de obtener una solución de compromiso entre el tráfico de señalización debido a avisos y el originado por actualizaciones de posición, adicionalmente a la jerarquía de celda y área de localización, aparecen otros niveles como son el RA (Routing Area). Un RA se define como un



subconjunto de celdas de un LA y posee una funcionalidad similar a la de las LA en GSM. El tamaño de un RA es siempre menor o igual al del LA al que pertenece.

## 2.6 GESTIÓN DE MOVILIDAD

La creciente movilidad exigida en la vida moderna está modificando los hábitos de vida tanto a nivel individual como colectivo en la sociedad actual. Como consecuencia, surge la necesidad de soluciones que permitan la comunicación entre personas mientras éstas se están desplazando por la red. Para garantizar esta movilidad de los abonados, manteniendo la calidad de los servicios, una red de telefonía celular dispone de un conjunto de procesos encargados, básicamente, del seguimiento del usuario y de la llamada. Esta sección describe los principales procesos relacionados con la gestión de la movilidad.

### 2.6.1 Actualización de la posición (Location Update)

Dado que entre los terminales móviles y la red no existe un vínculo físico continuo, se necesitan una serie de procesos para dar soporte a la movilidad de los terminales. Es decir, la red debe conocer una cierta estimación de la posición de esos terminales a fin de conocer la central con la que establecer una conexión (por ejemplo, en caso de una llamada entrante). La gestión de la movilidad se ocupa del seguimiento del usuario y, cuando proceda, del mantenimiento de la comunicación establecida. Su correcta ejecución depende del estado del terminal.

En GSM se utiliza un esquema de dos estados: *DEDICATED* (ocupado con una llamada) y *IDLE* (encendido y desocupado). En el caso de que el usuario se encuentre realizando una llamada (*DEDICATED*), como consecuencia de su movilidad puede ocurrir que se aleje de la estación base servidora y sea necesario conmutar la comunicación a otra estación base desde la que recibirá mejor calidad de señal. Esta operación se denomina Handover y se describe en el siguiente subapartado. En esta situación, la red siempre conoce la estación base a la que está conectado el terminal durante la comunicación, es decir, las celdas por las que transita, notificando la nueva posición cada vez que cambie de celda.

Cuando un terminal se encuentra encendido y desocupado (*IDLE*), el usuario debe estar localizado en todo momento para que la red pueda encaminarle las llamadas entrantes cuando éstas ocurran. El área de localización (LA) es el nivel de resolución con el que trabajará la red celular en la búsqueda de terminales. Esto es, cada terminal móvil se encuentra asociado a una única área de localización y esta información se mantiene continuamente actualizada en las base de datos. Como los terminales pueden cambiar de posición y, por tanto, el punto de acceso

a la red, debe existir un proceso automático y transparente que mantenga a la red informada de las ubicación de sus terminales a nivel de LA.

Para conseguirlo aparece el proceso Location Update (LU), el cual implica un intercambio de mensajes de señalización entre el terminal y la red, y el posterior registro en las bases de datos para actualizarlas con arreglo a esos mensajes. La frecuencia de ejecución del proceso de LU depende del estado del teléfono. Cuando el teléfono está encendido y desocupado (*IDLE*), basta que la red conozca el LA donde se encuentra para alertarle en caso de una llamada entrante. Por tanto, el teléfono puede moverse por las distintas celdas del LA sin necesidad de actualizar su posición. En este caso, los eventos que lanzan el proceso LU para actualizar la posición son los siguientes. i) Cuando el terminal móvil se enciende y se conecta a la red, se registra en el LA en el que se encuentre. ii) Por otro lado, la movilidad del terminal puede provocar un cambio de posición que suponga la entrada en un área de localización diferente. Mediante un análisis de la señalización enviada por canales de “broadcast” se conoce el LA en el que se encuentra un terminal. De ese modo, es posible detectar la entrada en un nuevo área, comparando el nuevo valor recibido con el anterior valor de LA almacenado de modo que, si ha cambiado se ejecuta el “Location Area Update”. iii) Asimismo, si el teléfono no tiene ningún tipo de actividad existe un mecanismo periódico que lanza reportes automáticos a la red con la última posición almacenada por el teléfono. Por ello, cada vez que finaliza un temporizador, se obliga a realizar un registro periódico para notificar su presencia en la red aunque se encuentre en el mismo LA. Este conjunto de procesos permite que la red conozca continuamente la ubicación del terminal en términos de área de localización.

Para gestionar la movilidad en sistemas de telefonía de generaciones más avanzadas (2.5G/3G), entre otros procedimientos, existe una actualización a nivel de RA. De este modo, cuando el terminal detecta que ha entrado en un nuevo RA, comparando el identificador de área de enrutamiento (RAI) que tiene almacenado con el RAI recibido vía radio, se inicia automáticamente un proceso para la actualización correspondiente.

### **2.6.2 Entrega de llamada (Paging)**

En caso de que haya una llamada dirigida a un terminal móvil, la red debe tener sus mecanismos para poder avisar al móvil y establecer finalmente la conexión. Este proceso se denomina “*paging*” y consiste en el envío de mensajes de aviso por canales radio destinados a ello. Cuando se presenta una llamada entrante, se consulta la base de datos HLR para determinar el LA en el que se encuentra actualmente el destinatario de la llamada. A continuación, el mensaje de búsqueda sólo se difunde en las celdas pertenecientes a ese LA, enviando un mensaje con el

identificador del terminal al que se debe entregar la llamada. Finalmente todos los terminales analizan este mensaje y responde solamente el terminal buscado. En ese momento ya se puede establecer el canal necesario para cursar la llamada puesto que se conoce la estación base (celda) que daría servicio al teléfono.

### **2.6.3 Traspaso (Handover)**

Para mantener una comunicación (voz y/o datos) con un terminal móvil, es necesario que exista una conexión “punto a punto” entre el terminal origen y el de destino, la cual debe extenderse en el tiempo mientras dure la comunicación. Por ello, la red precisa conocer continuamente el punto de acceso a la red por el que encaminar la comunicación, es decir, la estación base (celda) en la que se encuentran los terminales implicados. Esa información de posición de los abonados a nivel de celda es almacenada y actualizada en las bases de datos (HLR y VLR) oportunas.

Dado que un terminal puede moverse entre áreas de servicio de distintas estaciones base durante una conexión activa con la red (voz y/o datos), una red debe disponer de mecanismos adecuados para que ese terminal esté permanentemente conectado y controlado por la red durante la comunicación. El HandOver (HO) surge como solución al problema del mantenimiento de una comunicación cuando la calidad del enlace es insuficiente, principalmente cuando el terminal se mueve entre dos celdas de la red. Con los mecanismos adecuados, se consigue un cambio de una celda a otra de un modo transparente para el usuario y sin degradar la calidad de la comunicación. Este proceso permite que un terminal esté siempre conectado tanto a la red como a la mejor estación base servidora. Existen otros casos en los que puede iniciarse el handover tal como un elevado nivel de interferencia, el cual imposibilita la comunicación con la calidad adecuada, o bien que por razones de tráfico se deban liberar recursos de esa estación base y continuar la comunicación a través de otra con menor tráfico. En el proceso de handover intervienen mecanismos para evitar efectos indeseados, como los cambios de estación base innecesarios (efecto ping-pong), mediante el uso de parámetros que miden la calidad del enlace como el RSSI (Radio Signal Strength Indicator), BER (Bit Error Rate), relación CIR, entre otros.

Durante el diseño de la red, el operador define un conjunto de celdas para cada estación base (celda) a las que se permite realizar el handover. Típicamente, este conjunto de celdas coincide con las celdas vecinas que la rodean, aunque puede variar según la topología de la zona. Asimismo, se ordena dicha lista de celdas vecinas conforme a una prioridad en función de parámetros de diseño de la red. En el caso de que se tome la decisión de ejecutar un handover, el BSC inicia el proceso de traspaso a aquella celda de la lista con mayor prioridad.

Desde el punto de vista operativo, el handover tiene 3 etapas: medida, toma de decisión y ejecución. Por ello, durante el handover se produce un intercambio de información de señalización entre el terminal y centrales de la red (como son las BTS, BSC, MSC). Como resultado, se toma la decisión de hacer o no el handover. La decisión la toma siempre la BSC que se ocupa del abonado en ese momento. Si la decisión es negativa, se mantiene la comunicación con la estación base actual (misma celda). En caso de que se deba ejecutar el handover, se produce un intercambio de información con la red, derivándose ciertos parámetros asociados a cada una de esas llamadas traspasadas (como celdas implicadas, instante y duración del handover, junto con algún otro tipo de información de señalización).

En el caso de GSM, el handover consiste en una transferencia de recursos que implica la interrupción del enlace actual y la creación de uno nuevo. Esto se debe fundamentalmente a que las estaciones base vecinas utilizan portadoras y/o bases de tiempo distintas y los terminales no son capaces de conectarse simultáneamente a ambas. En esos casos, denominado Hard-Handover, el terminal se desconecta de su estación base origen y durante un tiempo (del orden de milisegundos) no está conectada a ninguna otra estación base. Esto provoca una interrupción de la señal vocal apenas perceptible por el abonado. Sin embargo, esta situación no es aplicable a comunicaciones de datos (2.5G/3G), donde una interrupción de ese tipo se traduce en pérdida de los datos, dado que la red no está preparada para estos cortes. Por ello existe otro tipo de handover, denominado Soft-Handover, durante el cual el terminal estará conectado a la estación base origen mediante un canal y a la estación base destino mediante otro. Así pues, la transmisión se realiza en paralelo por los dos canales, es decir, no se produce interrupción del enlace. Este es el esquema utilizado en redes más avanzadas.

## **2.7 TRÁFICO Y DIMENSIONADO**

### **2.7.1 Introducción**

El dimensionado en un sistema de telefonía móvil es una actividad compleja que requiere cuantificar, aunque sea de manera teórica e incorporando muchas simplificaciones e hipótesis, la capacidad que tendrá la red de soportar usuarios para una configuración dada y unos ciertos requisitos de calidad. Esta capacidad vendrá referida, esencialmente, al tráfico telefónico ofrecido en una celda y su distribución con respecto al tiempo, para garantizar un grado de servicio básico. Ese tráfico telefónico varía según la estación base, el mes, día e incluso instante dentro del día. La duración de las conversaciones es otra variable importante a considerar a la hora del dimensionado.

Tradicionalmente, este dimensionado se efectúa conforme a la intensidad de tráfico de la hora de mayor ocupación BH (Busy Hour) – período de 60 min. de un día para el cual la carga de tráfico es máxima –, y viene expresada en Erlangs. Un Erlang es una unidad adimensional utilizada en telecomunicaciones como una medida estadística del volumen de tráfico o intensidad; en concreto, se emplea para caracterizar el sistema en el ámbito de tráfico telefónico. Por ejemplo, si el uso total del teléfono móvil por hora en un área dada es de 180 minutos, esto representa 3 Erlangs.

El tráfico telefónico indica la cantidad de información (voz y/o datos) transportada por una vía de comunicación o canal. Este tráfico se mide en términos de tiempo (entendido como tiempo de ocupación del canal) y depende del número de comunicaciones y de la duración de las mismas. Por ejemplo, en el caso de un canal que se ha utilizado para diez llamadas en una hora, cada una con una duración de tres minutos, dado que el comportamiento del canal es en serie (las comunicaciones se suceden una detrás de la otra), se cuantifica que el canal ha estado ocupado durante un periodo de tiempo de treinta minutos: 0,5 Erlang, o bien que la ocupación del canal ha sido del 50%. Así pues, la intensidad de tráfico telefónico expresada en Erlangs mide el promedio de llamadas en progreso simultáneamente durante el periodo de una hora o la ocupación media de un canal por unidad de tiempo (hora).

### 2.7.2 Modelos de tráfico telefónico

En el despliegue y la operación de una red móvil, el dimensionado de los recursos radioeléctricos debe ser tratado como un aspecto clave. Para abordar el problema del dimensionado suelen emplearse modelos de tráfico, los cuales definen una serie de conceptos y parámetros que posteriormente serán utilizados en las metodologías desarrolladas en esta Tesis.

En redes telefónicas fijas el tráfico es caracterizado con gran precisión. En cambio, un modelado preciso de las características de tráfico en redes de telefonía móvil suele ser más complejo. En general, el tráfico telefónico generado en una zona puede describirse, principalmente, atendiendo a dos modelos de tráfico que difieren de su punto de vista de la red. i) El modelo de movilidad (*mobility model* o *traffic source model*), el cual describe el sistema desde la perspectiva del usuario móvil. En él, el escenario de tráfico telefónico se representa como un conjunto de entidades individuales (usuarios) realizando un recorrido aleatorio a través del área de servicio y con una demanda aleatoria de recursos, es decir, de canales de radio. ii) En cambio, existe un modelo de tráfico de red (*network traffic model*), para el cual el tráfico telefónico viene descrito como una observación realizada desde elementos de la red, como

estaciones base (BTS) o centrales (MSC). Este modelo caracteriza la dependencia temporal y espacial de la distribución de tráfico de llamadas mediante el tráfico ofrecido  $A^l$  en una celda.

Este último modelo es muy utilizado a la hora de determinar la ubicación de los principales componentes en una red móvil (BTS, MSC,...). Estos componentes deben estar dimensionados para soportar un tráfico próximo al previsto a fin de aumentar la eficiencia del sistema. En cambio, debido a su capacidad para describir la dependencia temporal del comportamiento de los usuarios en detalle, el modelo de movilidad es el más extendido para la caracterización del tráfico en cada celda particular.

A continuación se presenta un modelo de tráfico, simplificado pero bastante general, cuyo planteamiento está basado en la teoría de tráfico telefónico usando el citado modelo de movilidad. El modelo se encarga de estudiar el dimensionamiento y la gestión de los recursos que se ponen a disposición de un sistema de telecomunicación, en este caso, de telefonía móvil, para optimizar tanto su rentabilidad como la calidad de servicio ofrecida por celda.

En el contexto de las redes móviles 2G, como GSM, el concepto de capacidad queda fácilmente definido en la medida en que hay básicamente tráfico de voz en la red. En cambio, en el marco de los sistemas 2,5G y 3G, en los que se plantea un entorno multiservicio con requerimientos de calidad de servicio muy diversos, el concepto de capacidad de la red debe entenderse de forma más amplia y su estimación no resulta tan inmediata como en el entorno sólo de voz. En base al razonamiento anterior y dado que el estudio se centra en redes GSM, el modelo está referido únicamente al tráfico telefónico de voz, por ser el tráfico dominante en GSM. Asimismo, el modelo debe estar basado en pronósticos de demanda en función del tipo de zona, ya que el tráfico telefónico varía según las características de los abonados y de la celda en cuestión.

En este modelo se considera que la llegada de llamadas puede modelarse mediante un proceso de Poisson de tasa de llegada  $\lambda$ , de forma que la probabilidad de que en la hora de mayor ocupación (o BH) se produzcan  $n$  llamadas, se define como:

$$p_n = \frac{e^{-\lambda} (\lambda)^n}{n!} \quad (1)$$

En este modelo simplificado de tráfico de voz, si  $M$  es el número de estaciones móviles y  $L$  el número medio de intentos de llamada por móvil en la hora de mayor ocupación, BH, se tiene:

---

<sup>1</sup> El tráfico  $A$  se cuantifica mediante la intensidad del tráfico  $\lambda$  – la cual se mide como intentos de llamadas por unidad de tiempo y espacio (llamadas/seg·km<sup>2</sup>) – y teniendo en cuenta la media de duración de llamadas  $E[t_{llamadas}]$ , siendo  $A = \lambda E[t_{llamadas}]$  (en Erlang/km<sup>2</sup>). Esta medida representa la cantidad de tráfico que se ofrece en una zona determinada.

$$\lambda = \frac{M \cdot L}{3600} \quad (2)$$

En los procesos de llegada de Poisson, el tiempo  $t$  comprendido entre dos intentos consecutivos es una variable aleatoria continua que se modela mediante una distribución exponencial negativa, cuya densidad de probabilidad es:  $f(t_d) = \mu \cdot e^{-\mu t_d}, t_d \geq 0$ . El parámetro  $\mu$  de la distribución es igual a  $1/T_c$ , siendo  $T_c$  la duración media de la ocupación del recurso, es decir, duración media de la llamada. Entonces, la intensidad del tráfico  $A$  ofrecido al recurso en cuestión se define como:

$$A = \frac{\lambda}{\mu} \quad (\text{Erlang}) \quad (3)$$

Sustituyendo en (3) los valores  $\lambda$  y  $\mu$  resulta:  $A = \frac{M \cdot L \cdot T_c}{3600}$  (Erlang).

Esta expresión permite interpretar la intensidad de tráfico como el tiempo medio total de ocupación de los recursos en la hora de mayor ocupación, BH.

Con este parámetro de intensidad de tráfico quedan expuestos los conceptos básicos que más adelante serán usados en el ámbito de la Tesis. Sin embargo, es necesario indicar que una llamada generada en una celda determinada puede proceder de una llegada nueva o una llegada de traspaso ya que una llamada puede concluir en la celda donde se origina o ser objeto de traspaso a otra u otras celdas. Por tanto, en general, toda llamada puede descomponerse en dos: las que se inician y concluyen en una misma celda y las que experimentan traspasos (handovers). De modo que la tasa de llamadas entrantes  $\lambda$  del modelo de Poisson estará formada por  $\lambda_c$  y  $\lambda_h$ , siendo éstas las tasas de llegada de nuevas llamadas y las de traspasadas, respectivamente. Estas tasas  $\lambda_c$  y  $\lambda_h$  definen sendos procesos de Poisson y estarán relacionadas entre sí. Esta relación se demuestra utilizando la hipótesis de equilibrio de flujos de llamadas, en la cual la tasa de llamadas que abandonan una celda coincide con las que llegan y vale  $\lambda_h$ :

$$\lambda_h = p_h [(1 - p_b)\lambda_c + (1 - p_{bh})\lambda_h] \Rightarrow \lambda_h = \frac{p_h(1 - p_b)\lambda_c}{1 - p_h(1 - p_{bh})} \quad (4)$$

Donde  $p_h$  es la probabilidad de handover,  $p_b$  es la probabilidad de bloqueo para una nueva llamada y  $p_{bh}$  es la probabilidad de bloqueo para un intento de handover, considerando bloqueo como la ausencia de canales para asignar. De la ecuación anterior se desprende que la tasa de llamadas entrantes por traspaso en una celda  $\lambda_h$  es una fracción de la tasa de llamadas ofrecidas a esa celda  $\lambda_c$ , siendo ese parámetro de proporcionalidad  $\bar{n}_h$  el siguiente:

$$\frac{\lambda_h}{\lambda_c} = \frac{p_h(1-p_b)}{1-p_h(1-p_{bh})} = \bar{n}_h \quad (5)$$

La probabilidad de handover  $p_h$  será igual a la probabilidad de que la duración de una llamada ( $t_d$ ) supere el tiempo de permanencia en la celda ( $t_p$ ). Para el cálculo de  $p_b$  y  $p_{bh}$  se utilizan modelos de procesos estocásticos de tráfico para obtener la probabilidad de que haya  $n$  canales ocupados, generalmente mediante el modelo de Erlang-B.

En este apartado se han presentado algunos de los conceptos básicos relacionados con el tráfico de voz en el ámbito de un modelo para la evaluación y dimensionado de sistemas de telefonía móvil. Estos conceptos serán empleados posteriormente en el planteamiento de las metodologías desarrolladas en la Tesis. El modelo de tráfico expuesto puede ser aplicado a situaciones muy generales ya que sigue las principales hipótesis asumidas en redes de telefonía móvil: las llegadas de llamadas siguen un proceso de Poisson y el intervalo de tiempo entre llegadas sucesivas sigue una distribución exponencial negativa. En la literatura existen numerosos modelos de tráfico que asumen estas distribuciones (Hong y Rappaport, 1986; Steele y Nofal, 1992; Zeng *et al.*, 1997; Gavish y Sridhar, 1996), cuyos desarrollos son mucho más extensos que los expuestos en este apartado. Es necesario indicar que este apartado sólo se ha centrado en los aspectos básicos relativos al dimensionado de los recursos radio a fin de presentar unas nociones básicas al lector.



# CAPÍTULO 3

---

## ESTADO DEL ARTE EN EL ÁMBITO DE LA ESTIMACIÓN DE PARÁMETROS DE TRÁFICO MEDIANTE LA TELEFONÍA MÓVIL

### 3.1 INTRODUCCIÓN

El capítulo anterior describe los conceptos básicos y funcionalidades principales existentes en los sistemas de telefonía móvil para garantizar la movilidad de los abonados, manteniendo la calidad de los servicios. Durante su ejecución, estos procedimientos proporcionan información sobre la localización de los usuarios que, tras el tratamiento oportuno, puede ser útil para la caracterización del tráfico vehicular, con parámetros tales como velocidad, volumen o movilidad origen–destino, entre otros. En los siguientes apartados se hace una revisión exhaustiva del “Estado–del–Arte” en el ámbito de la estimación de estos parámetros a partir de la telefonía móvil, describiendo métodos usados en diferentes proyectos, tanto de simulación como pruebas piloto, existentes en el sector académico e industrial.

### 3.2 MATRICES ORIGEN–DESTINO

Las matrices origen–destino (O–D) se utilizan para cuantificar y sintetizar la movilidad asociada al desplazamiento de personas y/o mercancías sobre una región. Estas matrices aportan información sobre el número de viajes realizados entre una “zona de transporte” origen  $i$  y una destino  $j$  en un determinado intervalo de tiempo. Así pues, se trata de una estructura de información que representa la demanda de transporte sobre un área durante un período temporal específico. Tradicionalmente, las áreas que conforman las “zonas de transporte” orígenes y destinos de una matriz, es decir, las zonas donde empiezan y terminan los viajes, vienen definidas por criterios socioeconómicos o agregaciones de población que se corresponden con

zonas que pudieran tener una relación causal con los movimientos de transporte que se dan entre ellas.

Estas matrices pueden ser elaboradas con diferentes niveles de agregación, dependiendo del grado de detalle que se desee (tamaño de las zonas) o el tipo de información necesario: por modo de transporte (vehículo privado, autobús...), por bien transportado (personas, mercancías...). Las unidades empleadas para definir una matriz O–D están fuertemente relacionadas con este nivel de agregación. De este modo éstas podrían ser vehículos, personas, mercancías, o incluso teléfonos móviles.

En los sistemas de telefonía existe una clasificación de la zona de servicio en celdas y áreas de localización (LA), aunque adicionalmente existen las áreas de enrutamiento (RA) para 2.5G y 3G. Según lo visto en el Capítulo 2, la red maneja una estimación de posición de un terminal mientras esté encendido dentro de su cobertura según los diferentes estados: encendido y desocupado (*idle*), o con comunicaciones activas (*dedicated* o *active*), incluso, en 2.5G/3G existe un estado adicional en el cual el terminal se encuentra en espera de iniciar una comunicación en el contexto de paquetes una vez registrado en GPRS (*standby*). Si esto se traslada al dominio de teléfonos a bordo de vehículos en movimiento, un terminal que tenga establecida una comunicación con la red proporciona datos bastante continuos del recorrido: a nivel de celdas por las que está circulando. Incluso si el terminal no está en uso sino simplemente encendido, también ofrece datos válidos del recorrido, puesto que continuamente se conoce el LA en el que se encuentra. En este caso, aunque los datos de localización son menos precisos y frecuentes, también permiten identificar viajes. Por ello, construir una matriz O–D mediante datos de localización de teléfonos móviles exige definir el nivel de agregación que se desea conseguir, es decir, si las “zonas de transporte” serán las celdas o las LAs existentes en la zona de servicio de la red de telefonía existente en la región bajo estudio.

Según lo expuesto anteriormente, el nivel de agregación de la matriz está directamente relacionado con el tipo de datos de localización que la genere. Por ello, si se pretende obtener una matriz de viajes entre celdas origen–destino, se precisa conocer las celdas por las que los teléfonos transitan en todo momento. Esto es posible cuando se trata de datos asociados a comunicaciones (llamadas) efectuadas por los teléfonos. En estos casos, la muestra de datos asociada es más reducida que si se utilizara el conjunto de teléfonos encendidos.

Sobre esta idea de matriz entre celdas O–D se han presentado recientemente varios trabajos. Tras años de estudio, White y Wells (2002) publicaron los resultados de un proyecto desarrollado conjuntamente por TRL (Transport Research Laboratory) y HA (Highways

Agency) en el Reino Unido. Dicho proyecto consistió en la extracción de información sobre movilidad entre celdas origen–destino con datos anónimos de localización asociados a llamadas realizadas por usuarios, proporcionados por el operador BTCellnet (actualmente O2). Se sabe que, cuando un teléfono efectúa una llamada, se conoce la estación base servidora en todo momento y, por tanto, la celda en la que se encuentra. Con fines de facturación, se almacenan múltiples datos asociados a cada llamada, como la posición del teléfono al inicio y fin de la llamada o la duración de la misma, entre otros. El citado trabajo hace uso de esos datos de posiciones inicial y final de cada llamada para obtener información de movilidad sobre una región bajo estudio en términos de celdas origen y destino. Así pues, se consigue construir una matriz de viajes donde las “zonas de transporte” se corresponden con las celdas que definen la cobertura en la región analizada. El principal inconveniente de las matrices elaboradas bajo este criterio es que la información recopilada es parcial. El viaje detectado con las llamadas está limitado al recorrido que realizó el usuario con la llamada activa, el cual lógicamente no suele corresponderse con el recorrido completo realizado por el usuario; por ejemplo, en su desplazamiento de casa al trabajo. No obstante, la matriz resultante proporcionaría una matriz de partida que podría ser tratada posteriormente con algoritmos de expansión u otros procedimientos similares para inferir una matriz válida.

Lógicamente, el empleo de este tipo de información asociada a la telefonía móvil implica la colaboración de un operador que los suministre. Esos datos se recibirían y tratarían de forma agregada y anónima, según los reglamentos vigentes, de modo que no exista posibilidad de asociar datos con abonados reales. Sólo se pretende conocer información referente a la localización de un terminal, visto como una sonda anónima que se mueve a través de la red. A pesar de ello, conseguir la cooperación de un operador de telefonía no es una tarea sencilla. Por ello, es habitual recurrir a la ayuda de simuladores para conseguir datos de localización necesarios para implementar sus algoritmos de estimación ya que poseen la ventaja de que la información “real” de posición de cada vehículo puede ser extraída según un tiempo de muestreo deseado, y así validar con mayor facilidad dichos algoritmos.

Uno de esos trabajos basados en simulaciones es el desarrollado por Akin y Sisiopiku (2002), el cual también elaboró una matriz de viajes entre celdas pero mediante datos de localización generados por un conjunto específico de teléfonos sobre una cierta zona. Para desarrollar los modelos y validar sus algoritmos, los datos de localización (registros en las celdas de una región) fueron creados mediante una función generadora de números aleatorios en MS Excel. Los números generados representaban los identificadores de celdas visitadas durante un día por un conjunto de teléfonos móviles pertenecientes a un operador ficticio. Esos teléfonos

técnicamente transmitían su posición cada 5 minutos, representando una muestra de viajeros sobre una ciudad ficticia. Tras obtener una matriz mediante algoritmos oportunos, se efectuaban los correspondientes ajustes dado que los viajes se derivaban de una muestra limitada asociada a un conjunto concreto de teléfonos y había que extenderlo a la demanda total de tráfico.

Pese a los buenos resultados logrados, los anteriores estudios no aprovechan la ventaja de la alta penetración de la telefonía celular en la sociedad actual, puesto que manejan solamente una pequeña muestra del total de teléfonos existentes en la zona. En un caso utilizan sólo datos de aquellos que efectúan llamadas y en otro, datos de un conjunto exclusivo de teléfonos capaces de transmitir periódicamente su posición a un servidor central. Sería de vital importancia aumentar el tamaño de esa muestra de las que se derivan las matrices. Esto supone utilizar datos del conjunto de teléfonos que estén encendidos de un operador, de los cuales se conoce en todo momento el LA en el que se encuentran.

Este es el objetivo perseguido en el trabajo elaborado por Caceres *et al.* (2007). Su investigación planteó la obtención directa de una matriz O-D considerando como “zonas de transporte” de la matriz el conjunto de áreas de localización (LA), definidas por una red GSM, existentes sobre la región a estudiar. Debido a que la gestión de movilidad de una red GSM obliga a que el terminal notifique al sistema su posición cuando entra en un nuevo LA (proceso LU, Location Update), junto a los reportes periódicos de posición (PLU), los teléfonos móviles encendidos siempre proporcionarán información actualizada de las LAs por las que han pasado. El análisis de los registros efectuados por cada teléfono en un intervalo de tiempo dado permitió conocer el LA de origen (registro inicial) y de destino (registro final) asociado al viaje que ha seguido el vehículo que transportaba el teléfono. Además, este tipo de análisis de registros logró distinguir entre viajes válidos, pares O-D de la matriz estudiada, del resto de viajes de paso, los cuales son viajes que se iniciaron o terminaron en orígenes o destinos distintos a los que conforman la matriz. La validación de la metodología se realizó mediante un simulador de tráfico, generando bases de datos HLR y VLR ficticias con los registros de localización. Asimismo, dado que los valores estimados venían expresados en términos de movilidad de teléfonos, este trabajo implementó un algoritmo para extender datos de teléfonos encendidos de un operador concreto al conjunto global de vehículos mediante un factor de corrección. Entonces, tras comparar la matriz O-D de viajes “real” y estimada, los reducidos niveles de error mostraron la posibilidad de obtener una matriz correlada con los viajes realizados en un área extensa de estudio. Sin embargo, dado que los LA abarcan áreas de gran dimensión, lógicamente existirán desplazamientos dentro de ellas, denominados viajes intra-zonales, que no podrían ser

detectados analizando exclusivamente registros generados mediante procesos LU (cuyo resolución más fina es a nivel de LA).

Un desarrollo intermedio entre las dos ideas anteriores fue el propuesto por Pan *et al.* (2006), el cual fue validado en la ciudad de Shanghai. En el citado trabajo se combinó la alta penetración de la telefonía móvil, empleando teléfonos que simplemente están encendidos, y la precisión a nivel de celda para confeccionar una distribución de viajes. Al ejecutar el proceso de LU (off/on, cambio LA o registro periódico en ausencia de actividad telefónica), se notifica a la red el LA en el que se encuentra junto con la celda dentro de ese LA. Pero la posición de un terminal no sólo se actualiza cada vez que se ejecuta un proceso de LU, sino cuando el teléfono registra algún tipo de actividad (realice una llamada o envíe un sms). Empleando ambos conjuntos de datos se logra fusionar las ventajas de utilizar una muestra extensa (cualquier teléfono encendido) y precisión a nivel de celda. Sin embargo, se pierde cierta información de movilidad asociada al hecho de que un teléfono encendido puede moverse por las celdas de un LA sin que quede constancia de su paso por ninguna de ellas, ya que no se precisa ejecutar un proceso de LU mientras no cambie de LA o registre actividad telefónica (llamadas, sms...); no obstante el registro periódico PLU suaviza este inconveniente. Para obtener la distribución de viajes, la metodología utiliza dos escenarios de tiempo basados en las características de las zonas que conforman la matriz. Por ejemplo, los viajeros permanecen en casa desde las 10:30 p.m. hasta las 6:30 a.m. del día siguiente (escenario temporal A) y estarán en sus lugares de trabajo desde las 9:30a.m. hasta las 11:30a.m. (escenario temporal B) de ese nuevo día. Analizando los registros de localización generados por los teléfonos, una tabla de distribución de viajes es generada computando teléfonos que han permanecido en la zona *i* (e.g. casa) durante el escenario temporal A y están en la zona *j* (e.g. trabajo) durante el escenario temporal B. El almacenamiento propio de los sistemas de telefonía, tanto de la localización como los instantes de actualización, permite que los flujos O-D puedan ser obtenidos de una forma sencilla mediante un análisis de dicha información.

En cualquier caso, tanto para matrices obtenidas mediante llamadas como con procesos LU, la principal ventaja es que los datos son recopilados directamente de flujo de tráfico y no de procesos de encuestación (encuestas domiciliarias, telefónicas, reconocimiento matrículas, etc.) los cuales son procesos costosos y suelen dar lugar a resultados con ciertos sesgos. Sin embargo, ambos procesos presentan como desventaja la pobre precisión obtenida en la localización. Los datos de posición derivados de la telefonía móvil vienen expresados en términos de celdas y LA en la que se encuentran. El tamaño de esas zonas, aunque varía de entornos urbanos o rurales, suele ser grande, de modo que es normal que muchos de los viajes ocurran dentro de dichas

áreas (viajes intra-zonales). En base al razonamiento anterior, ambas metodologías pueden ser consideradas como válidas para obtener una matriz previa para, posteriormente ajustarla con otra información mediante métodos tradicionales, pero no para ser usada directamente como matriz de viajes fiable.

### 3.3 VOLUMEN

Se denomina volumen de tráfico  $Q$  al número de vehículos que pasan por un punto o sección transversal de una vía por unidad de tiempo. Las unidades más usadas son la hora y el día, definiendo el volumen horario [veh/h] y volume diario [veh/día], respectivamente. Las distribuciones temporales de los volúmenes de tráfico se deben a diferentes estilos y forma de vida, los cuales determinan patrones de viaje basados en el tiempo (ciertas épocas del año, horas específicas del día, etc.). Estos volúmenes pueden ser entendidos como la utilización de la vialidad por la demanda de transporte. De modo que los estudios de volúmenes se realizan con el propósito de obtener información sobre el movimiento de vehículos sobre puntos o secciones específicas dentro de una red viaria, proporcionando una descripción intuitiva del comportamiento dinámico del tráfico.

Para la medición de esta magnitud suelen utilizarse los aforos. Un aforo mide la cantidad de vehículos que transitan por un determinado punto o sección del viario. Tradicionalmente, estos aforos son un tipo de información que se recoge de forma automática en un conjunto de puntos asociados a tramos del viario. La medición de este parámetro de tráfico obliga por tanto a instalar elementos adicionales en las vías (espiras, cámaras, etc...). Este volumen de paso de vehículos varía según la movilidad de la población, categoría de la vía, temporalidad, etc. Para conocer de un modo preciso y continuo las características del tráfico de vehículos que circula por la red viaria sería necesario observar todas las vías de manera casi permanente, lo cual implicaría disponer de una infraestructura considerable de elementos medidores del aforo. Así pues, surge la necesidad de nuevos medios que permitan caracterizar el tráfico de un modo rápido, preciso y permanente, sin excesivo coste.

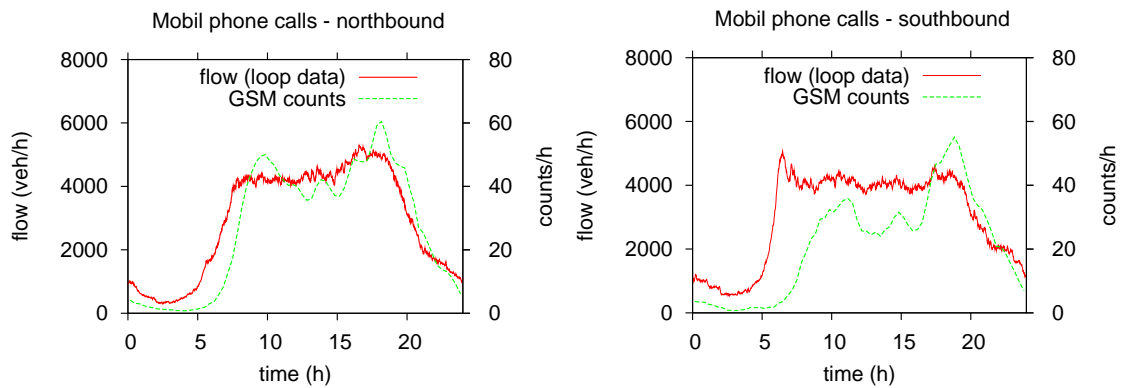
La telefonía móvil constituye una alternativa real, plausible y competitiva para conseguir esta información. En contraste con los sensores fijos, los teléfonos móviles proporcionan información respecto a comportamiento del tráfico sobre cualquier zona donde exista cobertura telefónica, sin necesidad de un despliegue adicional de infraestructura. En estos casos, el volumen estaría asociado al paso de teléfonos por un determinado punto o zona de la red, en lugar de vehículos. Para ello existen ciertos eventos y procesos en GSM relacionados con la

gestión de movilidad que permiten detectar el cambio de zonas, bien a nivel de celdas (handover) o bien a nivel de áreas de localización (proceso LU).

Analizando la topología de la red junto con la distribución de celdas y áreas de localización es posible asociar fronteras entre dichas zonas con secciones de vía en las que es posible monitorizar el paso de teléfonos. En tales fronteras se considerará que existe una especie de aforo “virtual” para la observación del paso de entidades por las vías como hacen los aforos tradicionales. Con ello, cada vez que un teléfono cambie de celda durante una llamada (handover) o cambie de LA mientras esté encendido (proceso LU), se podrá considerar que ha pasado por alguna de las vías monitorizada por un aforo “virtual” localizado en la frontera correspondiente. El número de vías que atraviesen la frontera, el cual determina la precisión de la medida del aforo, depende de la extensión de la zona fronteriza. En el caso de un aforo “virtual” que monitorice el paso de una celda a otra, el número de vías observadas suele ser menor que en caso de uno que monitorice el paso de un área de localización a otro puesto que un LA comprende una extensión de terreno mayor formado por varias celdas.

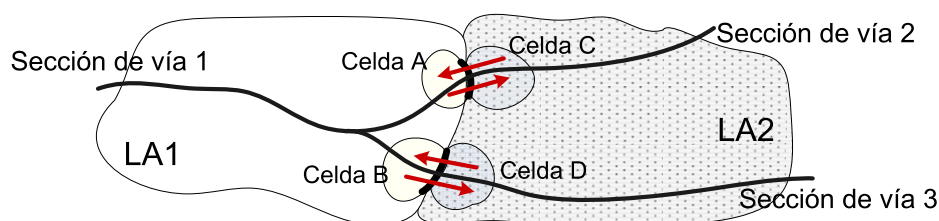
En los últimos años se han publicado diferentes estudios relacionados con esta idea. Uno de ellos es el desarrollado por Thiessenhusen *et al.* (2006) del DLR utilizando datos suministrados por Vodafone. La investigación utilizó el handover como evento para detectar el paso de teléfonos de un cierto operador por aforos “virtuales” ubicados en las fronteras entre dos celdas. Durante un handover, se almacena una serie de parámetros asociados, como celdas implicadas o instante en el que se produce. El análisis de esos parámetros permitió detectar el tránsito por la frontera correspondiente. Los resultados publicados mostraron cómo ese flujo de teléfonos (llamadas) está estrechamente relacionado con el flujo de vehículos medido por detectores instalados sobre los mismos puntos en la red. En la Figura 3-1 se comparan ambos valores a lo largo de un día. Se aprecia que existen los picos típicos asociados a horas punta tanto por la mañana como por la tarde, mostrando la influencia de los hábitos telefónicos (horarios típicos de llamadas) en los cambios de flujo de vehículos típicos.

No obstante, con el uso del handover sólo se monitoriza el paso de teléfonos que tengan una comunicación activa (p.ej., llamada). De este modo, la muestra de terminales de las que se obtienen datos es reducida e interesaría ampliar su tamaño a fin de conseguir una muestra más representativa del conjunto total de vehículos. Por ello, existe otra alternativa de monitorización de tráfico por una frontera pero, en este caso, a nivel de área de localización.



**Figura 3-1: Flujo (rojo) y número de llamadas de teléfonos (verde) en función del tiempo. Izquierda: northbound (desde ciudad), derecha: southbound (hacia ciudad). Todos los datos son promediados sobre intervalos de 15 minutos. (Thiessenhusen *et al.*, 2006)**

El trabajo presentado por Caceres *et al.* (2007) utiliza este criterio. Para validar sus algoritmos, no utilizó datos reales de un operador sino una herramienta conjunta que integraba un simulador de tráfico vehicular junto con un módulo específico para modelar la gestión de movilidad usada en una red GSM sobre un conjunto de teléfonos asociados a vehículos. Con ello se elaboraban unas bases de datos HLR y VLR ficticias con registros de localización. Como ya se ha comentado, cuando el teléfono está simplemente encendido, el teléfono notifica su nueva posición cuando entra en un nuevo LA (proceso LU), quedando reflejado el tránsito entre ese par de LA. En su investigación, se consideró la existencia de aforos “virtuales” sobre secciones de vías que coincidían con las fronteras entre áreas de localización. Aunque, dado que el paso de un LA a otro se efectúa por una celda concreta del LA destino, y esa información de celda de entrada es conocida (cuando se ejecuta el proceso de LU se indica tanto el LA como la celda dentro de ese LA), realmente se observarán fronteras entre celdas de LA contiguos por los que se va a efectuar el tránsito. En base a ello, el aforo virtual monitorizaría el paso desde un LA a otro por una celda concreta, logrando una resolución a nivel de celda para la identificación de la vía por la que circuló el teléfono (Figura 3-2).



**Figura 3-2: Ejemplo de ubicación de aforo virtual.**



Para validar esta idea se analizaron unas bases de datos generadas con simuladores buscando, para cada teléfono, los registros asociados a cambios de área de localización. Conociendo ese registro (instante de tiempo, nuevo LA y celda) junto con el LA anterior, y observando si pertenece a una frontera de interés, se conocería el instante de paso por dicha frontera, se logró computar el paso de un teléfono por el aforo virtual correspondiente. Este flujo se corresponde con teléfonos encendidos de un único operador, por lo que es extendido a un conjunto de vehículos equivalente mediante algoritmos de ajuste apropiados. La Figura 3-3 compara el flujo real y el estimado tras el ajuste que se obtuvo para un aforo virtual típico.

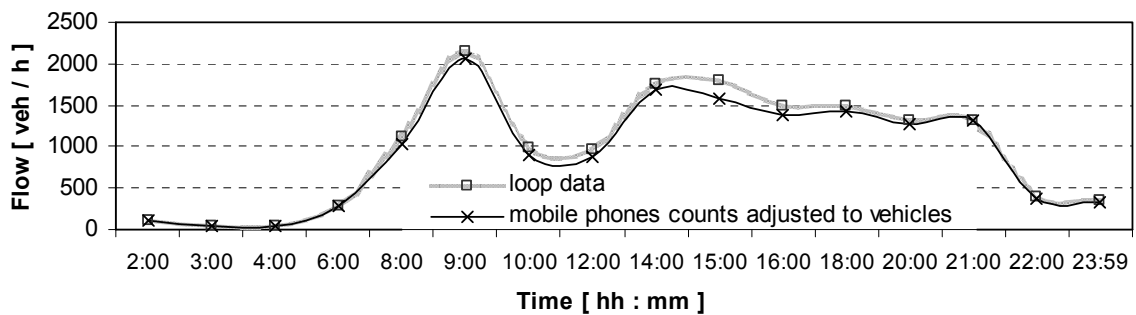


Figura 3-3: Comparación entre flujo real por un aforo virtual y el estimado tras el ajuste. (Caceres et al., 2006).

Los resultados mostraron alta correlación entre valores reales medidos por un detector y los estimados con teléfonos mediante datos procedentes de procesos de LU por cambiar de LA, debido a que la mayoría de los vehículos (teléfonos a bordo) son contabilizados. Cualquier teléfono encendido notifica a la red el cambio de LA y, por tanto, ese paso será detectado por el aforo “virtual”. Esta es una de las ventajas que esta alternativa presenta con respecto al handover, considerar como muestra el conjunto de todos los teléfonos encendidos de un operador y no exclusivamente los que están haciendo handover (llamada al pasar de una celda a otra). Asimismo, el uso de un nivel de agregación en términos de LA no supone perder precisión respecto al handover (nivel de celda). Esto se debe a que un LA es una agrupación de celdas y, por tanto, la entrada en un nuevo LA se efectúa por una de esas celdas. En la Figura 3-4, un proceso de LU ejecutado al pasar desde el LA1 al LA2 no permitiría distinguir si el paso ha sido por el punto 1 o el 3 debido a que ambos puntos pertenecen a la celda por la que se entra al LA2 (celda 3). En cambio, un proceso de LU ejecutado al pasar en el otro sentido (desde el LA2 al LA1) sí que permitiría identificar si ha ocurrido por el punto 1 o el 3, puesto que los puntos pertenecen a celdas diferentes para la entrada al LA1 (celda 1 y celda 2, respectivamente). En esos casos, se aprecia que la precisión es la misma que con el handover. No obstante, esta alternativa presenta una desventaja respecto al handover ya que las fronteras

entre LAs son espacialmente menos frecuentes que entre celdas y, por tanto, se dispondrán de menos aforos “virtuales” de ese tipo.

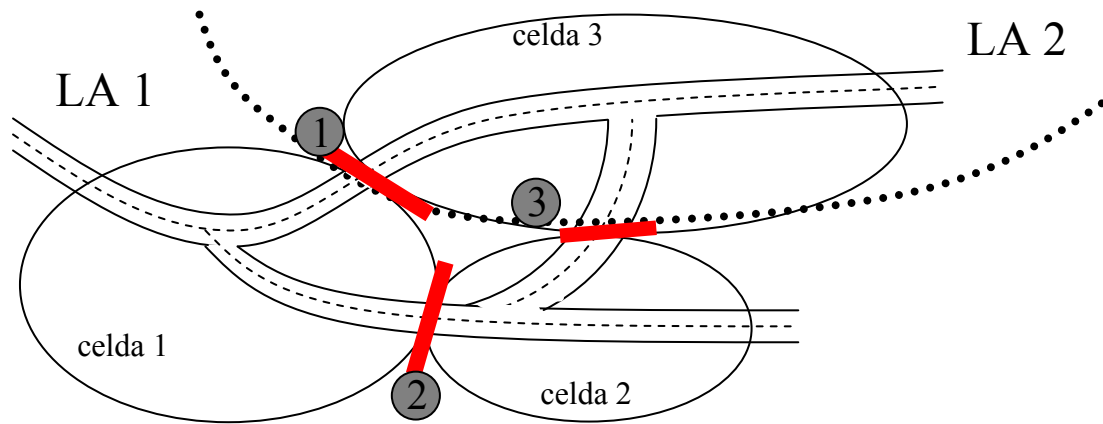


Figura 3-4: Representación de un aforo “virtual”.

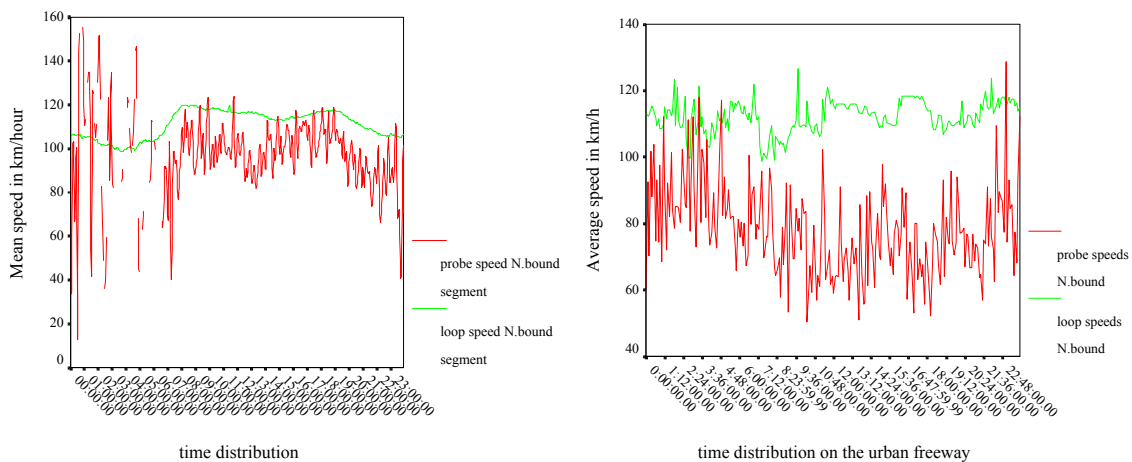
### 3.4 VELOCIDAD

La velocidad es uno de los parámetros más estudiados para evaluar la calidad del servicio de una vía según la demanda que soporta, especialmente en entornos urbanos. Muestra la relación entre el espacio recorrido por un vehículo y la unidad de tiempo (km/h o m/s). La velocidad se suele medir en un punto o tramo corto de una vía para determinar la rapidez con que pasan los vehículos (fluidez del tráfico). La velocidad puede ser definida según diferentes enfoques: velocidad instantánea, velocidad media temporal, velocidad media espacial o velocidad de recorrido, entre otros. La velocidad de cada vehículo sufre variaciones durante un viaje ya que depende de muchos factores, como la categoría de la vía (urbana, autovía...), ancho de la vía o número de carriles por sentido de circulación. Además existen otros factores variables que también influyen, como climatología, temporalidad (hora punta, día festivo...). Desde el punto de vista de tráfico, conocer la velocidad de un vehículo en un momento dado es poco representativo y resulta más eficiente trabajar con velocidades medias de recorrido. Para ello, se utilizan diferentes alternativas como el empleo de coche flotante (vehículos equipados con sensores que se sumergen en la circulación de una vía y capturan información de tráfico) o detectores en puntos fijos que realizan las mediciones mediante identificación de vehículos. Ambas alternativas suponen un gasto adicional de instalación y mantenimiento de los equipos.

Bajo el enfoque de la telefonía, Lovell (2001) comenzó e investigar sobre la viabilidad de mediciones de velocidad de vehículos usando el seguimiento anónimo de llamadas procedentes de teléfonos móviles. Para ello utilizó métodos de simulación que incluían algoritmos de

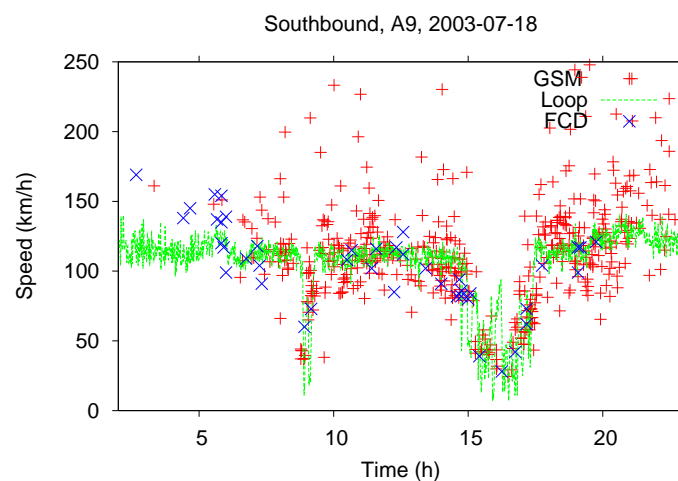
posicionamiento para diferentes técnicas de localización (AOA, TDOA), tamaños de muestra, media y varianza de velocidades de vehículos o incluso distancia. Los resultados mostraron que un sistema de este tipo podía clasificar las velocidades de los vehículos que transportaban los teléfonos, así como el grado de congestión conforme a tres niveles: bajo, medio, y alto. Este trabajo necesitó la incorporación de módulos adicionales a los que existen comúnmente en las redes de telefonía para aumentar la precisión de localización con la que efectuar los cálculos.

Algo similar presentó Ygnace (2001) en el informe del proyecto STRIP. Este proyecto instaló elementos sobre una vía a estudiar para monitorizar los interfaces radio Abis/A. Esto permitía identificar ciertos mensajes de señalización intercambiados entre un terminal encendido y la red para conseguir una estimación temporal de la posición del terminal mediante técnicas de redes neuronales (precisión: 100–150m). Esta estimación, junto con ciertos algoritmos e información histórica en bases de datos, servían para identificar correctamente la posición del vehículo sobre la vía que lo soportaba y así hacer predicciones de velocidad y tiempos de viaje. Los resultados de la estimación fueron contrastados con los medidos por detectores instalados en la vía analizada, mostrando la relación entre ambas cantidades. Estas estimaciones eran menos precisas en secciones urbanas de la vía debido al posicionamiento erróneo de usuarios que realmente estaban próximos a la vía pero no sobre ella. A pesar de ello, los resultados (Figura 3-5) mostraron que la tecnología de monitorización del interfaz Abis/A puede considerarse viable para segmentos de vía interurbanos; aunque resaltó la necesidad de una mejora de los algoritmos de posicionamiento, especialmente para aplicaciones enfocadas a su uso en tiempo real.



**Figura 3-5: Promedio velocidades: teléfonos vs. detectores Inter-Ciudad (arriba) e Intra-Ciudad (abajo) Motorway Northbound. (Ygnace, 2001)**

En general, los trabajos anteriores necesitaron la instalación de elementos adicionales en la red y/o modificación de los terminales. Sin embargo, una de las principales ventajas del uso de la telefonía celular es que no es necesario llevar a cabo dichas modificaciones para obtener parámetros de tráfico viables. El trabajo presentado por Thiessenhusen *et al.* (2006) del DLR con datos de Vodafone pone de manifiesto esta afirmación. En el citado se empleó información procedente de los handover ejecutados para medir ciertos parámetros de tráfico como la velocidad. El handover se produce cuando un teléfono pasa de una celda a otra mientras mantiene una comunicación activa (p.ej., llamada), produciéndose una conmutación de la llamada a la estación base que gestiona esa nueva celda. Ese instante en el que se produce el handover queda registrado, junto con otros parámetros como las celdas implicadas en él. Si la llamada es suficientemente larga como para que la nueva celda a la que se acaba de entrar sea atravesada completamente, se dispondrá de información temporal de entrada (handover inicial) y salida (handover final) de dicha celda. Estos tiempos, junto con las ubicaciones de las fronteras entre celdas (conocidas por el operador), permiten realizar una estimación de velocidad media para la sección de vía que atraviesa la celda comprendida entre ambas fronteras donde existió handover. Esto será viable en aquellos casos en los que la identificación de la vía dentro de la celda sea única y la duración de la llamada sea la suficiente para que se produzca un doble handover. La Figura 3-6 compara las mediciones de velocidad tomadas con tres fuentes diferentes: datos de teléfonos GSM, detectores y FDC tomados por GPS instalados en taxis.

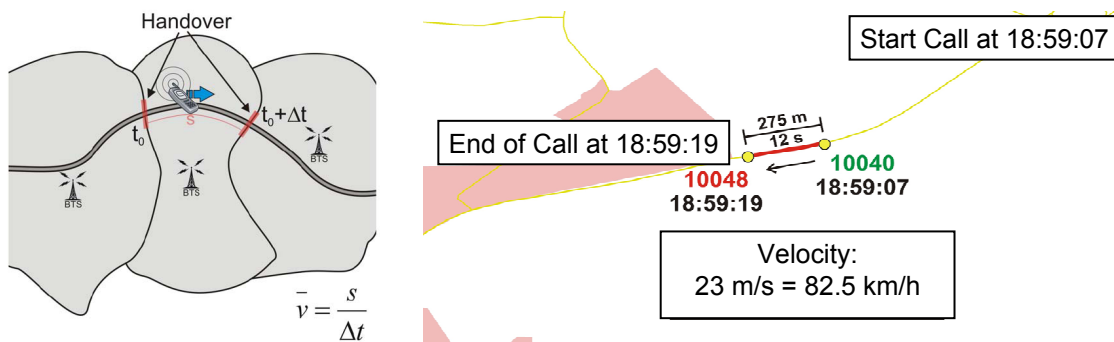


**Figura 3-6: Velocidad en función del tiempo obtenidos de datos GSM (+), detectores (-), FCD (x) (Thiessenhusen *et al.*, 2006).**

El principal problema de esta metodología es que la muestra sólo considera a aquellos teléfonos que efectúan llamadas lo suficientemente largas como para atravesar la celda completamente, lo cual es poco frecuente. Además, la identificación de la ruta seguida por el teléfono-vehículo

puede ser problemática debido a la existencia de varias posibles, especialmente en entornos urbanos. Aun así, el reducido coste de implantación de un sistema de este tipo permite considerarlo como alternativa viable para la determinación de velocidad.

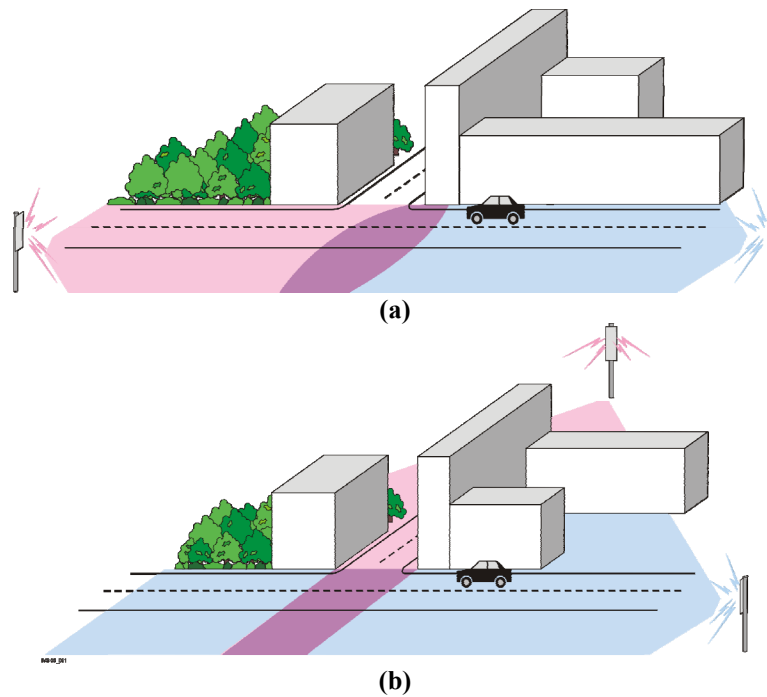
Una metodología similar presentaron Birle *et al.* (2006) en referencia al proyecto TrafficOnLine, que también consideró el handover como principio básico para estimar parámetros de tráfico como velocidad (Figura 3-7). Como se ha comentado, el método es preciso mientras sea posible identificar la sección de vía (ruta) que conecta ese par de zonas donde se producen los handovers. En general, esto ocurre en autovías o carreteras principales en las que no existen otras vías próximas con las que confundir la identificación.



**Figura 3-7: Concepto de estimación de velocidad usando doble-handover (Birle *et al.*, 2006).**

Sin embargo, en entornos urbanos suelen existir diferentes rutas para conectar zonas donde se producen los handover. Por lo que esa identificación de vía es más problemática y no es suficiente el concepto típico de handover como fuente de información. Para solventar ese problema en áreas urbanas, propusieron utilizar datos procedentes de handovers en zonas donde éstos se producen debido a un solapamiento de cobertura causado por la presencia de edificios, las cuales son más pequeñas y excepcionales dentro de una celda (Figura 3-8 b).

Con ello se consigue en áreas urbanas una identificación de localizaciones únicas dentro de una celda más viable que con el solapamiento típico (Figura 3-8 a). Sin embargo, este método requiere información adicional del nivel de potencia de señal de las celdas vecinas, es decir, informes de medida con los que identificar la ruta tomada por el terminal móvil al cruzar la celda. Los resultados conseguidos mejoraron en gran medida la precisión en la estimación de velocidad para entornos urbanos, pese a requerir un tratamiento más exhaustivo de los datos.



**Figura 3-8: Zona handover (a) con solapamiento típico y (b) con solapamiento debido a edificios (Birle *et al.*, 2006).**

### 3.5 TIEMPOS DE VIAJE

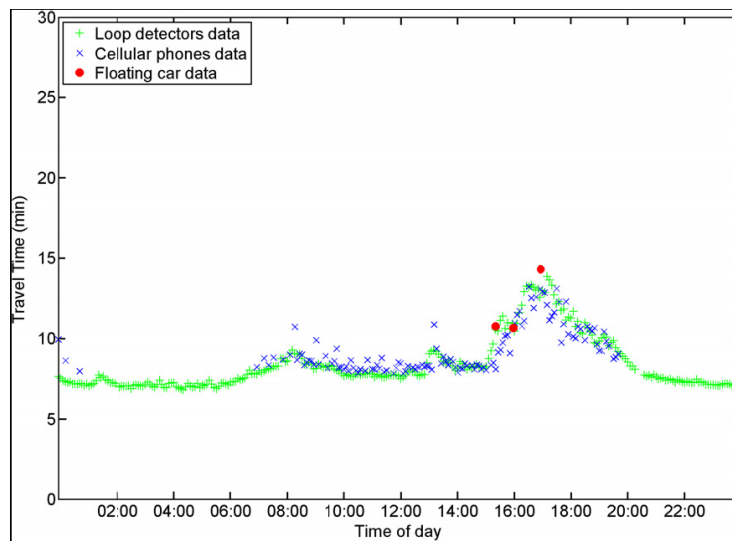
La determinación de tiempos de recorrido, junto con la de velocidad, resulta imprescindible en los estudios de planificación de una red viaria. Se denomina tiempo de recorrido al tiempo que invierte cada vehículo en desplazarse entre dos puntos fijos. Estos tiempos son medidos en tramos de vía de cierta longitud para conocer la calidad del servicio que prestan así como sus variaciones a lo largo de ellos. El modo de obtención de estos tiempos varía según el tramo observado. En tramos de longitud media y/o corta se utilizan, normalmente, equipos de observadores que anotan el instante en que entra y sale del tramo cada vehículo, distinguiéndolo mediante el número de placa, tipo de vehículo, etc. En cambio, para tramos de mayor longitud, se suele medir recorriendo varias veces dicho tramo con un vehículo y hacer mediciones de tiempo mediante la técnica de vehículo flotante. La incorporación de nuevas tecnologías está dotando de cierto automatismo a los procesos de medición con el empleo de cámaras y técnicas de reconocimiento de imágenes; aunque la estimación final continúa siendo un proceso costoso.

Lógicamente, el concepto de tiempo está estrechamente relacionado con la velocidad, por ello cualquiera de los métodos basados en telefonía móvil descritos en el apartado anterior puede emplearse para la estimación de tiempos de viaje en la sección de vía que discurre por una celda. En general, el concepto de doble-handover es el más extendido para la estimación de tiempos y velocidades. Tal es el caso del trabajo presentado por Linauer y Leihns (2003), en el

cual se identificaron las zonas donde existía el típico solapamiento de cobertura entre celdas (zonas handover). Los tiempos de viaje fueron obtenidos mediante el análisis de la señalización enviada al ejecutar el traspaso de una llamada entre dos “zonas handover” contiguas. Sin embargo, esta alternativa presenta como dificultad añadida problemas en el emparejamiento de zonas handover con la red viaria debido al tamaño que poseen. Solventar esta problemática requiere la utilización de algoritmos avanzados de “map–matching” como los desarrollados por Estimotion Ltd. (ITIS Inc.), cuya tecnología patentada se encarga de la medición y predicción de tiempos de viaje en tramos de carretera con alta calidad y en tiempo real.

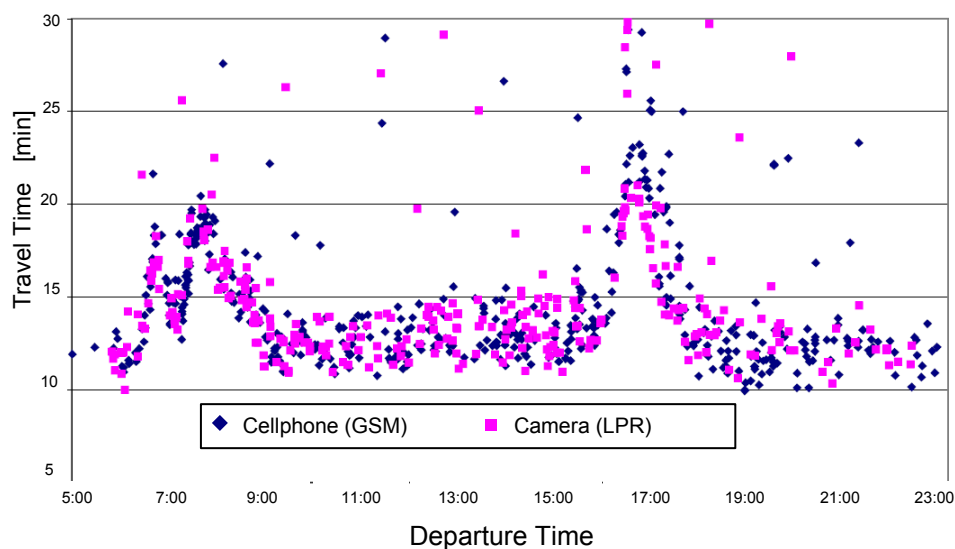
Utilizando la citada tecnología de Estimotion, Bar–Gera (2007) estudió las velocidades y tiempos proporcionados utilizando el evento de handover. El sistema recibía observaciones procedentes de handovers ejecutados por teléfonos existentes en la zona estudiada y almacenaba, para cada handover, el instante de tiempo junto con el identificador de celda asociado a las antenas de antes y después de la ejecución de dicho handover. El sistema era capaz de calcular el área donde era más probable que el terminal estuviera localizado cuando ejecutó el handover (zonas handover), con una probabilidad de acierto del 85%. La dimensión típica de esas zonas handover estaba en torno a 300–1000m, de ahí la complejidad de emparejar esas zonas con secciones de la red viaria. Esta tecnología de análisis consideraba la posibilidad de que las observaciones realmente no estuvieran relacionadas con vehículos desplazándose por la vía designada sino por otras próximas a ella, lo cual es frecuente en entornos urbanos. Finalmente, el sistema generaba estimaciones válidas de tiempos de viaje (en torno al 63%) para las secciones analizadas durante el periodo de observación (10:00–20:00), empleando una muestra cuyo tamaño está en torno al 1–3% del tráfico total. Los tiempos de viaje eran convertidos a velocidades medias de sección aplicando simplemente el ratio de longitud correspondiente a dicha sección.

Para probar la validez de las estimaciones, las medidas basadas en teléfonos móviles fueron comparadas con otras obtenidas con detectores y coches flotantes. Los resultados mostraron una buena correlación entre todas las medidas, donde sólo 4 valores atípicos fueron excluidos de la muestra al obtener medidas (con coches flotantes) considerablemente mayores que las del resto (con detectores y teléfonos móviles). La Figura 3-9 compara los tiempos de viaje obtenidos con los detectores, coches flotantes (Floating Car Data, FCD) y mediante teléfonos móviles, observándose una buena correspondencia entre los tres tipos de medidas. Según la figura, existe una leve desviación, en torno a 1 minuto, entre los tiempos de viaje estimados con datos de telefonía móvil y los procedentes de detectores, especialmente en situaciones de no-congestión.



**Figura 3-9: Tiempos de viaje en la autopista de Ayalon, sentido Norte (Bar-Gera, 2007).**

Respecto a los tiempos de viaje, cabe destacar el trabajo de investigación desarrollado por Finnra (Finnish Road Administration) en 2002 (Kummala, 2002; Virtanen, 2002). Este trabajo se basó en la idea de que los teléfonos móviles que se muevan siguiendo una cierta ruta siempre cambiarán de estación base (celda) en torno al mismo lugar. Según lo expuesto en el capítulo anterior, un teléfono móvil encendido intercambia regularmente información con las estaciones base que conforman la red. Esto permite que la red conozca, aproximadamente, dónde buscar a un teléfono, por ejemplo, a la hora de encaminar una llamada entrante.



**Figura 3-10: Estimaciones de tiempos de viajes y velocidades en función del tiempo de salida: datos GSM vs. LPR (Kummala, 2002).**



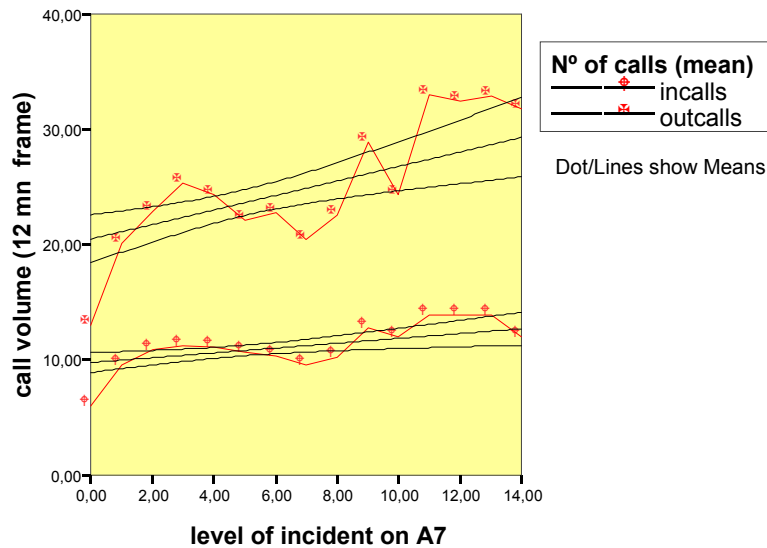
Así pues, el trabajo desarrollado por Finnra propuso una monitorización de esa señalización intercambiada entre el teléfono y la red para estudiar el tiempo que precisaba cada teléfono en cruzar una sección de vía, desde que entra en un área gestionada por una estación base determinada hasta la siguiente. Esto permite estimar la duración de tiempo de viaje asociado a esa sección. Para la prueba piloto sólo se necesitó monitorizar el 5% de los teléfonos para actualizar los patrones de tiempos cada 30 minutos. Los resultados obtenidos (Figura 3-10) fueron contrastados con los obtenidos mediante cámaras (LPR, License Plate Recognition), mostrando la alternativa basada en GSM mayor precisión debido a que no se veía afectada por problemas asociados a la visibilidad como condiciones meteorológicas adversas, frecuentes en países nórdicos.

### **3.6 CONGESTIÓN**

Un fenómeno de congestión se origina por la existencia de algún evento que motive un cambio en los valores medios típicos de velocidad y tiempos de viaje asociados a una sección. En los apartados anteriores se han descrito algunos métodos basados en telefonía que consiguen caracterizar secciones de vías en términos de velocidad y tiempos. Un análisis de variaciones bruscas en cualquiera de esos parámetros respecto a valores típicos puede ser empleado para la detección de congestiones. Análogamente, se puede hacer uso de otra magnitud relacionada con la intensidad de tráfico telefónico. El dimensionado de una red de telefonía móvil atiende a la naturaleza del tráfico telefónico cursado y su distribución con respecto al tiempo y destino. Esta magnitud varía según proceda de un tráfico continuo de llamadas (núcleos urbanos) u otro más esporádico característico de zonas de poca población. Generalmente, estas zonas se asocian a vías, como carreteras o autovías, alejadas de los núcleos urbanos. Por ello, el tráfico telefónico que soportan las estaciones base de esas vías se corresponde predominantemente con el generado por los ocupantes de los vehículos que circulen por ellas.

Para la detección de eventos como accidentes, atascos o cualquier incidente que altere el flujo normal de tráfico, es posible hacer uso de un histórico de intensidades típicas de tráfico telefónico medidas sobre esas mismas estaciones bases en diferentes horarios. Para comprender este razonamiento es necesario resaltar que los usuarios de telefonía móvil poseen ciertos hábitos en la realización de llamadas. Estos hábitos siguen ciertos patrones de modo que, cuando aparecen eventos que motivan un cambio, los valores típicos de intensidad suelen modificarse. Este es el caso de situaciones de congestión, en las cuales los usuarios tienden a realizar más llamadas por diversas razones, como avisar en el trabajo o incluso aprovechar el tiempo para hablar con algún conocido. Este motivo modifica la intensidad de tráfico típica al

incrementarse el volumen de llamadas cursado por el sistema. Una simple comparación con valores tradicionales puede detectar la existencia de alguna incidencia que motive el aumento de la actividad telefónica.



**Figura 3-11: Relación entre volumen de llamadas y nivel de incidentes en la autovía A7 (agosto y septiembre 2000). Nota: Nivel de incidentes (level of incident): 0 = ningún incidente detectado por los detectores, 14 = al menos 14 detectores (de 21) detectan un incidente en los correspondientes intervalos de 12 min; Volumen de llamadas (call volume): promedio de número de llamadas entrantes y salientes registradas en los correspondientes intervalos de 12 min (Ygnace, 2001).**

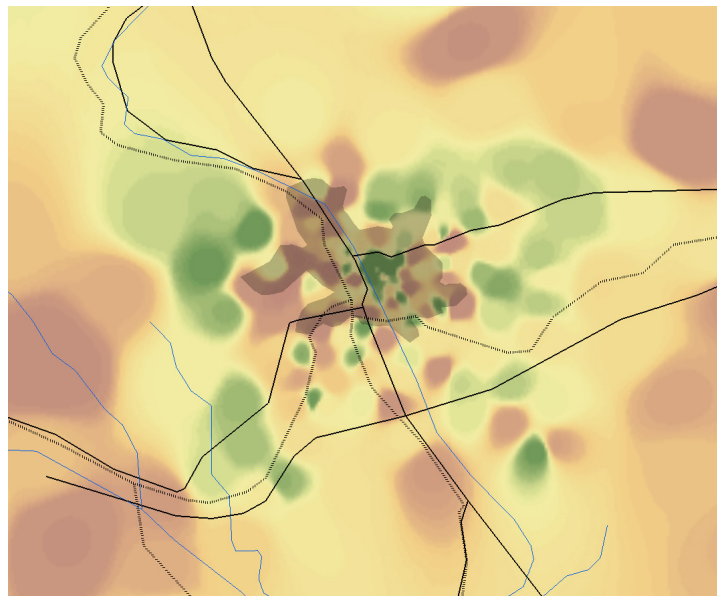
El estudio desarrollado por INRETS analizó la relación entre el volumen de llamadas (entrantes y salientes) y la congestión de una vía (Ygnace, 2001). Para ello, se consideró la hipótesis de que los conductores tienden a hacer más llamadas cuando se encuentran en alguna situación de congestión. En la Figura 3-11 se muestra la relación entre el volumen de llamadas de teléfonos móviles y el nivel de incidentes en un tramo de autovía de 80 millas. El análisis demostró que las llamadas entrantes no aportaban información relevante. Sin embargo, se detectó una relación significativa entre llamadas salientes de móviles sobre esa vía y el nivel de incidentes, la cual reflejaba el incremento de la actividad telefónica al estar los vehículos parados.

### 3.7 DENSIDAD DE TRÁFICO

En el ámbito de vehículos, la densidad es otro parámetro clave para medir la calidad de servicio que ofrece una sección de vía. La densidad de tráfico  $\rho$  representa el número de vehículos que ocupan una longitud específica de una vía o carril de ella, en un instante dado, excluyendo los vehículos estacionados. Se expresa en vehículos por kilómetro [veh/km], ya sea referido a un carril o a todos los carriles de una vía. Por su naturaleza espacial (se mide a lo largo de una

longitud), la medición directa de la densidad no es una tarea sencilla ya que precisa la observación de longitudes significativas de una sección de vía. No obstante, si se dispone de cierta información en puntos de esa sección, la densidad [veh/km] se estima como el cociente entre flujo [veh/h] y velocidad [km/h], pero esto sólo es válido bajo ciertas condiciones.

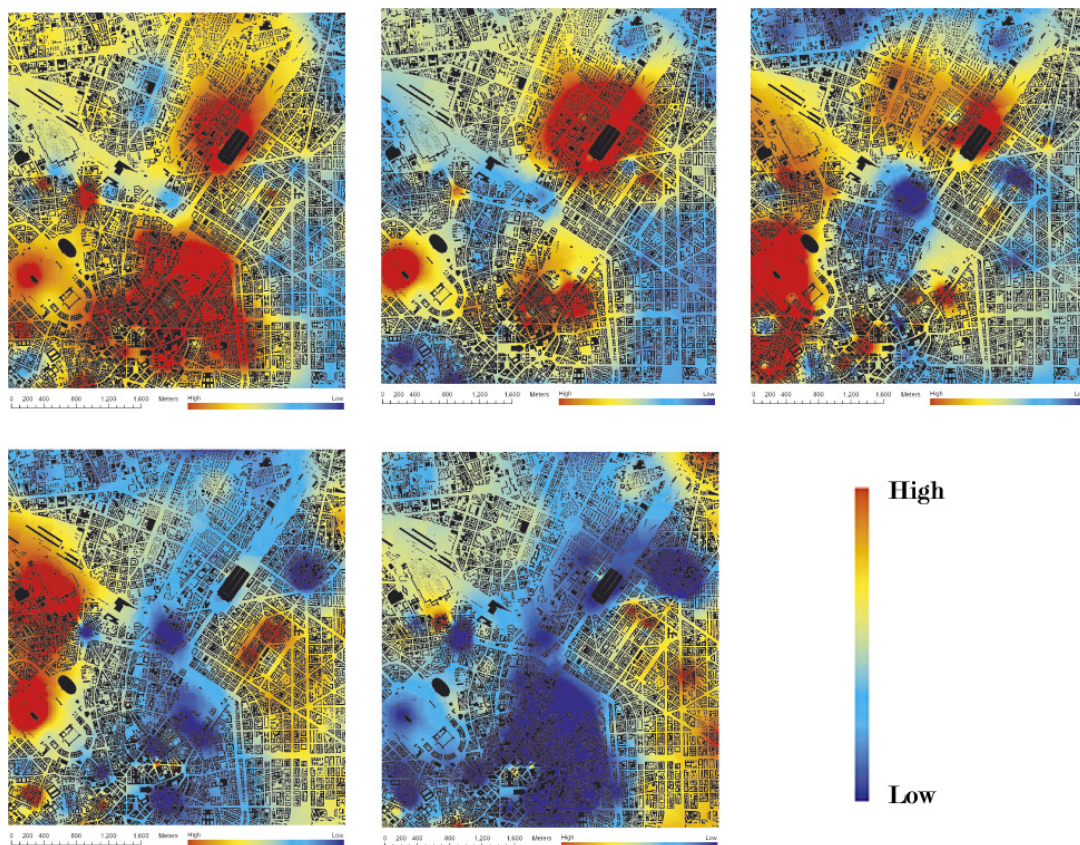
La telefonía móvil sugiere nuevas formas de obtención de esta magnitud. La unidad Erlang está relacionada con la demanda de utilización de los recursos de estaciones base que dan servicio a una zona, es decir, tráfico telefónico que se produce durante un determinado periodo de tiempo en cada celda. Como cualquier otro tipo de tráfico, el telefónico posee también un concepto ligado a la ocupación. Así pues, esta magnitud puede proporcionar una estimación de la ocupación de una celda en términos de consumo de recursos ofrecidos por la estación base asociada y, por tanto, medir el número de teléfonos a los que dicha celda está dando servicio.



**Figura 3-12: Intensidad de tráfico telefónico a 10:00 a.m. en el área metropolitana de Graz.**  
(Fuente: <http://senseable.mit.edu/projects/graz/highres/highres.zip>)

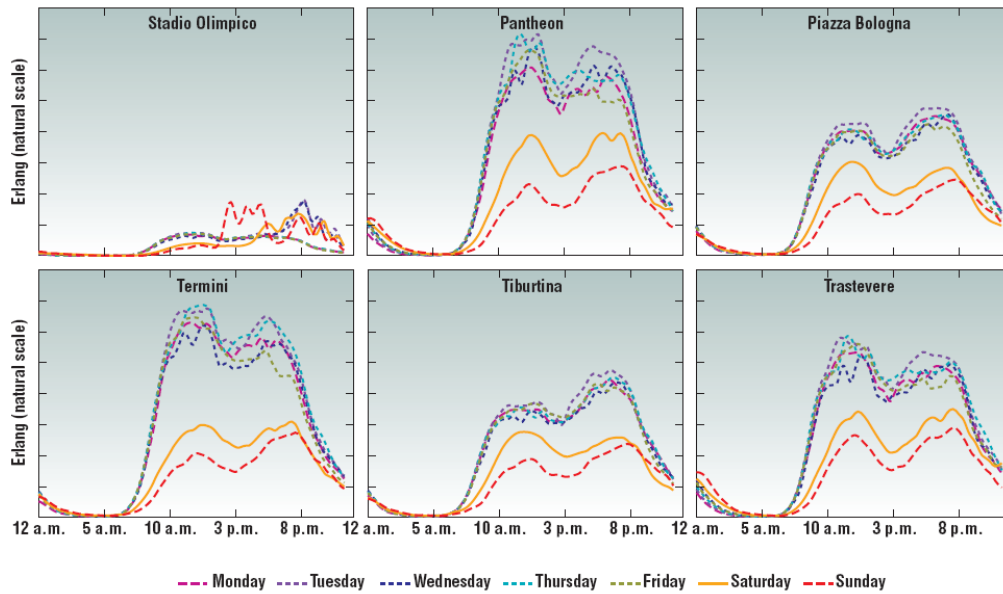
Esta línea es la que sigue Ratti *et al.* (2005) en su trabajo del SENSEable City Laboratory (MIT). En él se consigue representar sobre mapas de una ciudad ciertos parámetros de tráfico, visualizando la dinámica de sus habitantes en tiempo real. Uno de esos parámetros fue la intensidad de tráfico telefónico o densidad de las llamadas (medida en Erlang) de los abonados de un cierto operador, para lo cual contó con la cooperación del operador austriaco A1/Mobikom que proporcionó ciertos datos anónimos de teléfonos. Esta investigación también representaba sobre mapas información relacionada con la movilidad de usuarios a través de handovers. Asimismo, mediante una inscripción previa en un servicio de rastreo del operador,

se monitorizaban periódicamente las posiciones de usuarios que previamente se habían dado de alta. La Figura 3-12 muestra los valores de intensidad de tráfico gestionada por antenas de Graz (Austria) usando mapas de color, donde colores claros representan intensidades altas, y verdes y negros valores bajos o nulos de intensidad, respectivamente. En Milán (Italia) se aplicó también la misma metodología (Ratti *et al.*, 2006). La Figura 3-13 muestra gráficamente representaciones de la densidad de llamadas en diferentes horarios, observando una variación de niveles de densidad en función del tiempo.



**Figura 3-13: Representaciones de la densidad de llamadas de teléfonos móviles en el área metropolitana de Milán (20 km × 20 km) entre las 4 p.m. y las 6 p.m. (Fuente: <http://www.envplan.com/misc/b32047>).**

En ambos estudios del MIT se llegó a la conclusión de que el número de llamadas realizadas con teléfonos móviles está estrechamente relacionado con la densidad de personas y, en el caso de zonas donde no existen núcleos residenciales próximos sino autopistas o carreteras, con la densidad de vehículos mediante los teléfonos de sus ocupantes. Así pues, aplicando ciertos algoritmos de expansión, es posible convertir esa medida de intensidad de tráfico en celda [teléfonos<sub>LLAMADA</sub>/celda] a magnitudes relacionadas con la densidad en términos de vehículos/celda.



**Figura 3-14: Medidas de uso de la red móvil (Erlangs) en 6 lugares de Roma (Reades *et al.*, 2007).**

Este enfoque se comprende mejor en otro estudio desarrollado por el SENSEable City Laboratory (MIT) en Roma (Reades *et al.*, 2007). En el citado estudio, se consiguió definir patrones de actividad, según la zona y el día analizado, empleando medidas expresadas en Erlangs de uso de la red celular (llamadas) proporcionados por un operador. La Figura 3-14 muestra cómo las medidas de Erlang cambiaban con el tiempo en seis lugares seleccionados, los cuales representaban zonas residenciales (Piazza Bologna), turísticas (Panteón), ocio (Trastevere, Estadio Olímpico) y principales puntos de entrada a la ciudad por tren (Termini) y coche (Tiburtina). Como indican las gráficas, los valores de Erlangs de lunes a viernes en cada punto son muy similares, aunque el ratio varía de un lugar a otro. Por otro lado, los valores del sábado y domingo suelen tener valores inferiores respecto a los valores típicos de días laborables, reflejando un patrón de comportamiento diferente según el tipo de día. Sin embargo, a pesar de esas diferencias entre días laborables y fin de semana, la Figura 3-14 muestra que los seis lugares muestran un comportamiento similar: una rápida subida de actividad telefónica por la mañana entre las 6 y 10 a.m. en días laborables, con un ritmo más lento y algo retrasado en el tiempo durante el fin de semana, un valle al mediodía, una nueva subida por la tarde y un descenso abrupto por la noche.

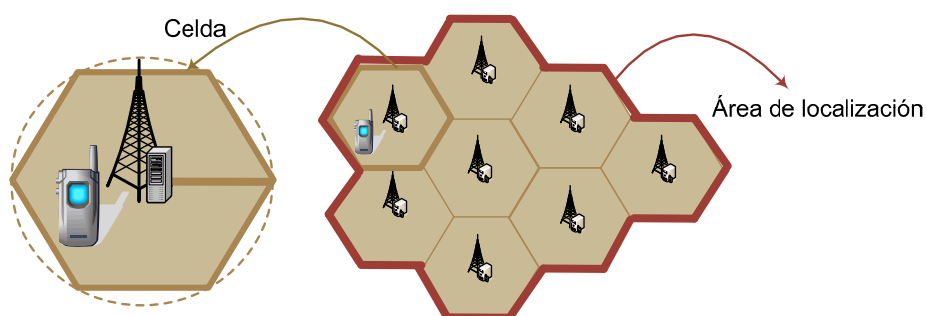
Por otro lado, los altos niveles nocturnos de uso de red, entre medianoche y las 2 a.m., los fines de semana cerca del Panteón y Trastevere, en comparación con los de lugares orientados a zonas residenciales o industriales, sugerían una actividad fuertemente orientada al ocio. Esta característica permite la posibilidad de identificar, por ejemplo, zonas de ocio y culturales

atendiendo al patrón de uso de la red celular. Por ello, este estudio también aplicó la actividad telefónica (densidad de llamadas) a otra línea como es la dinámica de una ciudad. Según esas medidas de uso de la red (en erlangs) se dividió la ciudad en zonas, de modo que se estableció una relación de dichas zonas con otras zonas definidas según la actividad desarrollada en ellas (comercial, turística, residencial,...). Ello permitió representar características espacio-temporales de la ciudad en función del uso del teléfono móvil. Se observó que la actividad telefónica de días laborables en cada zona exhibía un patrón temporal más dinámico, mientras que en fin de semana la actividad presentó mayor dispersión espacial. Una característica con implicaciones para la comprensión de la dinámica urbana de una ciudad fue el alto nivel de actividad por la mañana de días laborables en zonas asociadas a estaciones de trenes, autobuses o intercambiadores, en comparación con los niveles de zona residenciales a esa misma hora. Desde un punto de vista de planificación urbana, esta relación era obvia pero la visualización espacial y temporal de estas características a esta escala está comenzando a ser posible mediante la infraestructura ya existente en sistemas de telefonía móvil.

## 3.8 CARACTERÍSTICAS DE LOS DATOS DE TELEFONÍA MÓVIL

### 3.8.1 Precisión de localización

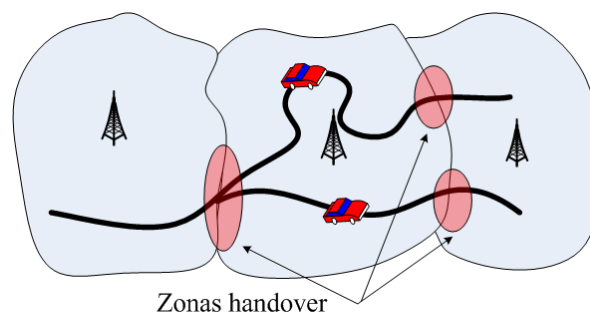
El esquema de localización en los sistemas GSM se basa en la existencia de áreas de localización y celdas (Figura 3-15). Con ello se pretende tener organizado el área de servicio (cobertura) en una serie de zonas para proporcionar eficientemente servicios a sus usuarios.



**Figura 3-15: Localización por Identificación Global de Celda, CGI.**

El nivel de precisión de posición más fino conseguido de una red de telefonía celular, sin realizar ninguna modificación en ella ni en los terminales, es a nivel de celda. Es decir, se localiza el terminal dentro del área de cobertura de la estación base a la que está conectado. Con

ello se obtiene una aproximación de la zona geográfica en la que se encuentra el teléfono, pero no la posición exacta. Por ello, existe un margen de error que depende del radio de la celda. Este radio oscila entre los 100 m y 1 km en ciudades. En entornos rurales depende de la densidad de las antenas, aunque suele oscilar entre 5 y 20 km. Ese margen de error empeora a medida que aumenta el nivel de agregación, por ejemplo, a nivel del área de localización (agrupaciones de celdas), que es el nivel de resolución con el que trabaja la red celular para gestionar terminales simplemente encendidos. El uso de resolución a nivel de LA implica que la zona geográfica en la que se puede encontrar el teléfono es más extensa, dando lugar a una mayor incertidumbre. Esta resolución afecta al mapeado de teléfonos sobre vías existentes en la zona en la que se informe que se encuentra el terminal, lo cual puede ser relevante para ciertos parámetros. Una de esas situaciones problemáticas se presenta con la existencia de múltiples vías dentro de una celda para parámetros que usen handover (véase la Figura 3-16), donde las celdas implicadas en los handovers son las mismas para las dos posibles rutas. En tales situaciones no es posible desagregar la velocidad o el tiempo de viaje estimado por ruta, ya que cada una tiene diferente longitud de sección de vía que recorre la celda. Sin embargo, esta situación no es tan relevante en otros casos como la medida de volúmenes. En ellos, la información de volúmenes de paso de teléfonos móviles entre dos celdas proporciona información agregada de paso total por las dos vías. No obstante, cada método toma el nivel de agregación apropiado según la resolución requerida para la predicción del parámetro. Incluso pueden paliarse tales efectos mediante el uso de algoritmos de “map-matching” (Pyo *et al.*, 2001, White *et al.*, 2000) o completando la información disponible con otros de los muchos parámetros que se pueden extraer de la red de telefonía móvil para determinar una posición más exacta del terminal.



**Figura 3-16: Múltiples vías dentro de una celda.**

Asimismo, hay diferencias en los requerimientos de calidad en las estimaciones de parámetros según los usos finales a los que estén destinados. Por ejemplo, datos que puedan ser usados con el fin de informar al viajero probablemente sean menos exigentes que si se utilizarán para la detección de incidentes o la gestión del tráfico. Por ello, en casos de altos niveles de calidad, los

parámetros de tráfico estimados mediante telefonía móvil deberían ser utilizados para completar otro sistema de monitorización de tráfico, pero no como sistema exclusivo. Incluso realizar modificaciones en la red o en los propios terminales también supondría una mejora en la calidad de las predicciones, por ejemplo la utilización de teléfonos equipados con A-GPS con el objetivo de obtener una mejor precisión de los datos de localización.

Otro factor que debe ser seriamente considerado es la frecuencia de renovación de esos datos de localización, la cual dependerá de si el teléfono está simplemente encendido (las actualizaciones tienen lugar tras vencer un temporizador o cambiar de área de localización) o si realiza algún tipo de uso de la red, como una llamada, en la que la actualización es continua e instantánea mientras dure la conexión.

No obstante, en términos de coste y prestaciones ofrecidas para lograr calidad en las estimaciones, la información de posición proporcionada por una red GSM sin modificaciones permite obtener datos fiables para monitorizar el movimiento de teléfonos y, por tanto, vehículos sin necesidad de desplegar una infraestructura adicional. Así mismo, se trata de un sistema bastante más económico dado que aprovecha la tecnología y módulos propios de las redes GSM para obtener la posición de cualquier terminal móvil, por muy sencillo que sea.

### **3.8.2 Tamaño de la muestra**

Como ya se ha referido previamente, una de las ventajas de la obtención de parámetros de tráfico basados en la telefonía móvil es la posibilidad de capturar información de un modo más continuo que con los métodos tradicionales. Esta continuidad en la captura depende de la frecuencia de renovación de los datos obtenidos de cada terminal. Lo deseable es conocer la posición exacta en términos de sus coordenadas espaciales y temporales ( $x, y, z; t$ ). Sin embargo, eso no es posible con redes o terminales sin modificar. Así pues, lo más preciso que se puede conseguir es conocer la localización en la que se encuentra cada terminal en términos de celda, dado que es el nivel de resolución más preciso disponible sin alterar la red ni los terminales. Pero esto sólo es posible cuando el teléfono tiene establecida una comunicación con la red. Con ello, el tamaño de la muestra de datos dependerá del número de llamadas efectuadas por los abonados así como la duración de las mismas. Sin embargo, la realización de llamadas no tiene una tasa tan elevada como la mera existencia de terminales que simplemente están encendidos. Un teléfono encendido y desocupado, estado más frecuente, ofrece datos menos precisos en términos de posición y más dispersos en el tiempo, dado que el terminal interactúa con la red menos frecuentemente. En ese estado, las notificaciones a la red de la posición son puntuales bajo ciertos eventos, como el vencimiento de un temporizador o cambio de área de



localización, por lo que la actualización de datos de posición es menos frecuente. Tampoco aportaría información a nivel de tráfico telefónico (Erlangs), dado que en ese estado no consume recursos. De modo que el empleo de información procedente de terminales en ese estado limita su aplicación a ciertos parámetros de tráfico.

En cambio, utilizar información obtenida de terminales ocupados con llamadas afectaría al tamaño de la muestra de datos disponible para las estimaciones, dado que el conjunto de terminales en ese estado se reduce considerablemente. Esta limitación restringe su uso a estimaciones de parámetros que necesiten muestras pequeñas. En algunos casos puede ser necesario obtener mayor precisión de posición y no una muestra de teléfonos de tamaño elevado, como ocurre con la velocidad o tiempos de viaje. En cambio en otros, la calidad de la estimación depende en gran medida de la alta penetración de teléfonos que actúan como “sondas” para inferir los datos de tráfico, como ocurre con las matrices O–D, densidad, etc. Por ello, las metodologías descritas anteriormente llegan a un equilibrio entre tamaño de muestra y precisión de posición deseada a la hora de su planteamiento, consiguiendo que estas limitaciones no afecten a la calidad de los resultados.

### **3.8.3 Aspectos legales**

Es importante destacar que estas técnicas de estimación de parámetros de tráfico mediante teléfonos móviles no infringen aspectos relacionados con la privacidad de los usuarios. Todos los datos analizados se reciben y se tratan en forma agregada y anónima, según las normativas europeas, de modo que no exista posibilidad de asociar datos de localización con abonados concretos. El objetivo de estas técnicas basadas en telefonía móvil es mostrar su aplicabilidad y aptitud para mejorar los métodos de cálculo de parámetros de tráfico y así lograr una eficiente gestión y planificación del tráfico.

Sin embargo, la principal preocupación asociada a estos servicios basados en localización (LBS) reside en la característica de que los movimientos que un abonado realiza sean continuamente monitorizados, especialmente cuando esos datos de localización se ponen a disposición de aplicaciones cuyos beneficiarios son terceras personas. Pero todos estos servicios se enmarcan dentro de la legalidad puesto que se rigen por normativas que se encargan de velar por la protección de la intimidad. En concreto, la directiva 2002/58/EC de la Unión Europea se encarga de regular el tratamiento de los datos personales y la protección de la intimidad en el sector de las comunicaciones electrónicas. En ella se distinguen tres grupos de datos asociados

al abonado: datos del tráfico, personales, y de localización. Los datos de mayor interés para los servicios LBS son los asociados al último grupo, los datos de localización.

En el punto 14 de la citada directiva aparece una descripción de los datos de localización, indicando que: *“Los datos de localización pueden referirse a la latitud, la longitud y la altitud del equipo terminal del usuario, a la dirección de la marcha, al nivel de precisión de la información de la localización, a la identificación de la celda de red en la que está localizado el equipo terminal en un determinado momento o a la hora en que la información de localización ha sido registrada.”* Además, esta directiva proporciona regulaciones sobre los datos de localización en el artículo 9, donde indica:

*“En caso de que puedan tratarse datos de localización [...] sólo podrán tratarse estos datos si se hacen anónimos, o previo consentimiento de los usuarios o abonados, en la medida y por el tiempo necesarios para la prestación de un servicio con valor añadido. El proveedor del servicio deberá informar a los usuarios o abonados, antes de obtener su consentimiento, del tipo de datos de localización distintos de los datos de tráfico que serán tratados, de la finalidad y duración del tratamiento y de si los datos se transmitirán a un tercero a efectos de la prestación del servicio con valor añadido.”*

Con ello, queda claro que estos servicios requieren acuerdos previos y que la protección de la privacidad es un aspecto fundamental para cualquiera de estos servicios. En particular, la inferencia de datos de tráfico mediante datos de localización procedentes de telefonía se trata de un ejemplo claro del empleo de los servicios LBS donde el beneficiario principal no sólo es la Administración sino la sociedad al completo, puesto que este servicio supone una mejora en la gestión y planificación de la red. Es necesario resaltar que el servicio no infringe la ley sobre protección de la intimidad, ya que hace uso de datos anónimos para no asociar desplazamientos a usuarios concretos. Sólo se pretende conocer información referente a la localización de un terminal, visto como un elemento anónimo de la red que se mueve a través de ella, sin importar las características del usuario u otro tipo de información asociada a él.

### **3.9 CONCLUSIONES**

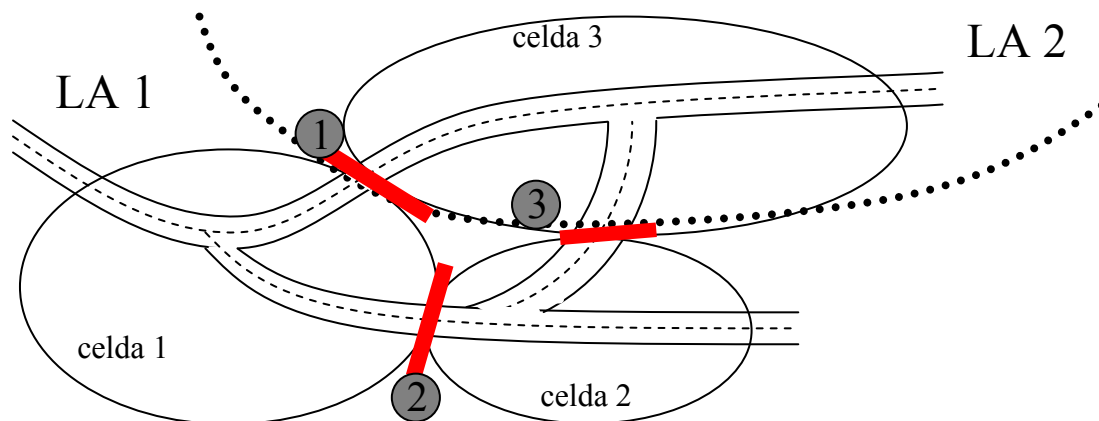
Una vez expuestos los procesos asociados a la gestión de movilidad de los sistemas de telefonía en el Capítulo 2 y revisado el “Estado del Arte” respecto a su aplicabilidad en la estimación de parámetros de tráfico, se puede concluir afirmando que son dos los procesos viables que permiten monitorizar el movimiento de terminales. Por un lado se dispone del handover, el cual permite conocer el momento en el que un terminal con una llamada activa cambia de una celda

a otra al necesitar la comunicación ser transferida de la estación base servidora a una nueva. Con ello se permite detectar el paso de terminales con llamadas activas por las fronteras entre celdas. Por otro lado existe el proceso de LU, el cual es ejecutado por cualquier terminal encendido para que la red se mantenga informada de la localización aproximada del terminal, en términos de área de localización (LA), logrando así encaminar rápidamente las llamadas entrantes en el caso que éstas ocurran. En base a este razonamiento, cada vez que un teléfono cambie de LA mientras esté encendido se considerará que ha cruzado la correspondiente frontera entre LAs.

La detección del movimiento de terminales derivados de handovers y/o procesos de LU presenta una serie de ventajas e inconvenientes comparándolos entre sí. Cuando se emplea el handover, la información extraída viene referida a frontera entre celdas, la cual, dada las dimensiones típicas de una celda, no abarca una extensión de terreno muy amplia. Con ello, el número de vías que discurren por la frontera es limitado, alcanzándose una mejor precisión cuanto menor es la zona fronteriza. En el caso de un proceso de LU para monitorizar el paso de un LA a otro, al ser un área de localización una región que comprende un grupo de celdas, la zona fronteriza observada posee un mayor tamaño que en el caso anterior. Esto provoca el aumento del número de vías que cruzan la frontera y, por tanto, del nivel de agregación de la información de trasposos proporcionada en este caso. No obstante, esta situación no ocurre siempre al utilizar el proceso de LU. Al ejecutarse este proceso no sólo se dispone de información del nuevo LA de entrada sino de la celda por la que lo ha hecho. De modo que la zona fronteriza a monitorizar realmente no comprende toda la frontera entre los LAs sino la sección de esa frontera limítrofe con una celda concreta. En base a este razonamiento, la desventaja del proceso de LU respecto al handover en términos de precisión a nivel de celda sólo ocurre en determinadas situaciones.

Esta afirmación se puede ver más claramente en la Figura 3-17. El paso desde la celda 1 a la 3 monitorizado por un aforo “virtual” que emplee el handover entre ese par de celdas observaría el movimiento de terminales por el punto 1. En cambio, ese movimiento no podría ser detectado por un aforo “virtual” que emplee el proceso de LU ejecutado cuando se cambia del LA1 al LA2, ya que éste observaría tanto el punto 1 como el 3. En este caso no se podría distinguir si el paso ha sido a través del punto 1 o del 3 debido a que ambos puntos limitan con la misma celda por la que se entra al LA2 (celda 3). Con ello, la información proporcionada por ese aforo “virtual” se correspondería con el movimiento por cualquiera de las vías asociadas a ambos puntos, es decir, información agregada de trasposo. Por el contrario, un proceso de LU ejecutado al pasar en el otro sentido (desde el LA2 al LA1) sí que permitiría identificar si ha ocurrido por

el punto 1 o el 3, puesto que los puntos limitan con celdas diferentes para la entrada al LA1 (celda 1 y celda 2, respectivamente).



**Figura 3-17: Fronteras entre celdas y LAs.**

Por otro lado, el hecho de analizar datos procedentes de handover (llamadas) conlleva que la muestra de terminales de la que se recogen los datos la forman aquellos que han tenido activa una llamada en el momento de paso por la frontera. Esto implica que el tamaño de la muestra es estadísticamente menos representativo que si se utilizaran datos de terminales que estuvieran simplemente encendidos, como ocurre en el caso de datos procedentes del proceso LU. La Tabla 3-1 resume estas características básicas.

**Tabla 3-1: Comparativa de volúmenes derivados de handovers o procesos de LU.**

	<i>Ventajas</i>	<i>Inconvenientes</i>
<b>LU</b>	Mayor tamaño de la muestra	Menor precisión (mayor número de vías por frontera)
<b>Handover</b>	Mayor precisión (menor número de vías por frontera)	Menor tamaño de la muestra

Atendiendo a estas conclusiones, una red GSM (sin modificar) puede no ofrecer prestaciones óptimas para fines de localización y monitorización de usuarios, especialmente en lo que a movilidad se refiere. No obstante, la rápida expansión de este tipo de sistemas de telefonía junto a los resultados alcanzados en proyectos similares relacionados con la obtención de parámetros de tráfico, impulsa a incorporar métodos basados en estas nuevas tecnologías en los estudios de transporte. Por ello, los Capítulos 2 y 3 han descrito y revisado conceptos y procedimientos relacionados con los sistemas de telefonía móvil para afrontar los objetivos de esta Tesis con un mayor conocimiento de la materia.

# CAPÍTULO 4

---

## INFERENCIA DE VOLÚMENES DE TRÁFICO MEDIANTE LA TELEFONÍA MÓVIL

### 4.1 INTRODUCCIÓN

Se denomina volumen de tráfico al número de vehículos que pasan por un punto o sección transversal de una vía durante un periodo de tiempo determinado. Estos volúmenes pueden ser entendidos como una medida del nivel de utilización del viario por la demanda de transporte. Los estudios de volúmenes se realizan con el propósito de obtener información del movimiento de vehículos sobre puntos o secciones específicas dentro de una red viaria, proporcionando una descripción intuitiva del comportamiento dinámico del tráfico.

Las distribuciones temporales de volúmenes de tráfico se deben a diferentes estilos y formas de vida, los cuales determinan patrones de viaje basados en el tiempo (épocas del año, horas del día, etc.). Por otro lado, los teléfonos móviles se han convertido en un elemento indispensable en nuestra rutina diaria. Éstos intercambian continuamente información con las estaciones de un sistema de telefonía móvil a fin de disponer de servicios telefónicos de calidad, con lo que una monitorización de datos generados por dichos terminales puede ofrecer un fiel reflejo de las pautas de movilidad de los usuarios. En concreto, ciertas situaciones que tienen lugar durante el funcionamiento típico de esos sistemas pueden desempeñar una tarea similar a las llevadas a cabo por las tradicionales estaciones de aforo en lo referente a conteos de tráfico. Este capítulo está dedicado a profundizar en este campo, desarrollando una metodología para inferir volúmenes de tráfico de vehículos a partir de la telefonía móvil mediante modelos matemáticos.

### 4.2 ANTECEDENTES Y ESTADO ACTUAL

Para la medición de los volúmenes de tráfico suelen utilizarse estaciones de aforo. Una estación de aforo mide la cantidad de vehículos que transitan por un determinado punto o sección del

viario; a este valor medido se le conoce con el nombre de “aforo”, término que también identifica al de “estación de aforo” o “punto de aforo”. Tradicionalmente, estos aforos son un tipo de información que se recoge de forma automática en un conjunto de puntos asociados a arcos del viario. Para ello se han de instalar elementos adicionales en las vías (espiras, cámaras de video...). Este volumen de paso de vehículos no es constante sino que oscila según la movilidad de la población, categoría de la vía, temporalidad, etc. Para conocer en detalle las características del tráfico de vehículos que circula por la red sería necesario monitorizar un número considerable de vías de modo casi permanente con algún tipo de detector, algo inviable cuando existen restricciones presupuestarias. Así pues, surge la necesidad de otras alternativas que permitan caracterizar el tráfico de un modo rápido, preciso y permanente, sin excesivo coste.

Un modo de conseguir esta información es mediante la telefonía móvil. En contraste con los sensores fijos, los teléfonos móviles proporcionan información respecto al comportamiento del tráfico sobre cualquier zona donde exista cobertura telefónica, sin necesidad de un despliegue adicional de infraestructura. En estos casos, el volumen estaría asociado al movimiento de teléfonos por una determinada zona de la red, dado que existen ciertos eventos y procesos en la telefonía móvil relacionados con la gestión de movilidad que permiten detectar el cambio de zona, bien entre celdas (mediante handover) o bien entre áreas de localización (mediante procesos LU, Location Update). Respecto a esta idea, en la sección 3.3 se presentaron los conocimientos actuales sobre el cálculo de volúmenes mediante la telefonía móvil, junto con una revisión del “Estado-del-Arte” sobre la obtención de volúmenes con datos procedentes de teléfonos.

Existen estudios en la literatura en los cuales se afirma que el flujo de teléfonos móviles derivado de llamadas con handover y el flujo de vehículos proporcionado por detectores (espiras, gomas, cámaras...) en los mismos puntos están directamente relacionados en la mayoría de los intervalos horarios de un día típico (Thiessenhusen *et al.*, 2006; Höpfner *et al.*, 2007), mostrando picos característicos similares en las horas punta de mañana y de tarde. No obstante, esos estudios manifestaron la dificultad de obtener flujos de vehículos precisos directamente de datos de llamadas debido a la naturaleza de dichos datos. Por tanto, la principal cuestión es determinar cómo el flujo de teléfonos estará relacionado con el flujo de vehículos ya que en los conteos de terminales que pasan por fronteras entre celdas no se monitoriza el conjunto completo de vehículos que la han atravesado, sino sólo una muestra de ellos.

Ciertos factores que tienen su origen en la naturaleza de los datos de teléfonos móviles afectan a la hora de cuantificar la relación entre vehículos y teléfonos. Por un lado, la telefonía móvil

aporta información asociada a movilidad de teléfonos que i) son de un cierto operador y ii) que están encendidos. Los teléfonos apagados o de otro operador son transparentes, es decir, aunque crucen la frontera no serán detectados en el proceso de conteo. En el caso del handover, existe otro condicionante adicional asociado a la naturaleza del evento, que implica que el teléfono esté realizando una llamada en el momento de cambio de celda. Adicionalmente a los factores citados, la existencia de más de un teléfono por vehículo también perturba la estimación de flujo usando datos de teléfonos.

La carencia de una relación directa e inmediata entre conteos de teléfonos que ejecutan handover y conteos de vehículos sugiere la necesidad de un tratamiento especial para relacionar ambas magnitudes, el cual puede requerir la realización de algún proceso de calibración usando conteos de vehículos obtenidos, por ejemplo, mediante detectores ubicados en la misma zona en la que se realiza el conteo de teléfonos. En la literatura se han propuesto diferentes modelos diseñados para correlacionar ambas magnitudes, como un sencillo modelo basado en funciones de transferencia generadas directamente con datos de aforos (Bekhor *et al.*, 2008) u otro más complejo basado en técnicas de asimilación de datos como la relajación Newtoniana o técnica nudging (Herrera y Bayen, 2008). En ambos casos, la utilización de modelos para relacionar ambas variables hizo posible alcanzar estimaciones más precisas de flujo de vehículos.

Ante las conclusiones anteriores, y dada la naturaleza de la información procedente de telefonía móvil, parece lógica la necesidad de usar modelos matemáticos que se encarguen de medir esa relación con un objetivo claramente predictivo. Los modelos matemáticos son ampliamente utilizados para resolver problemas concretos de ingeniería u otras áreas científicas, y constituyen un marco teórico general para la formulación de la inferencia estadística. En los siguientes apartados se proponen una serie de modelos para predecir volúmenes de vehículos a partir de la telefonía móvil con objeto de ser usados en los estudios de transporte en sustitución o de forma complementaria a medidas de volúmenes proporcionadas por los habituales aforos.

### **4.3 METODOLOGÍA PARA LA INFERENCIA DE VOLÚMENES DE VEHÍCULOS**

La inferencia estadística comprende una serie de métodos y procedimientos para deducir propiedades (hacer inferencias) de una población, a partir de una pequeña parte de la misma (muestra). Lógicamente es imposible predecir con exactitud y es por ello que la bondad de estas deducciones se medirá atendiendo a su grado de acierto. En este trabajo se va a plantear el uso de la inferencia estadística para predecir volúmenes de vehículos utilizando información

derivada de la telefonía móvil, de modo que el estudio estadístico correspondiente comprenderá los siguientes pasos, los cuales serán abordados en los sucesivos apartados:

1. Planteamiento del problema.
2. Extracción de la muestra.
3. Análisis descriptivo de los datos.
4. Tratamiento de los datos.
5. Elaboración de modelos y estimación de los parámetros.
6. Contraste de modelos.
7. Conclusiones.

### **4.3.1 Planteamiento del problema**

En esta fase inicial del problema se definen con precisión la población, la característica a estudiar, las variables dependientes e independientes, etc. Bajo el contexto de nuestro problema, la población o universo se corresponde con el conjunto de vehículos que pasan por una zona en un periodo de tiempo determinado (por ejemplo, cada hora), de modo que la característica a estudiar será la intensidad de paso de vehículos por dicha zona.

#### **4.3.1.1 Alternativas ofrecidas por la telefonía móvil**

En esta etapa también se deben analizar los medios de los que se dispone y el procedimiento a seguir. Como se ha comentado en el apartado 3.9, la telefonía móvil ofrece dos procesos asociados a la gestión de movilidad que permiten monitorizar el paso de terminales, bien entre celdas mediante el handover, o bien entre áreas de localización (LA) mediante el proceso de Location Update (LU). A partir de la red viaria real se elabora un modelo de red compuesto, principalmente, por nodos y arcos. La mayoría de los nodos corresponden a una intersección o cruce, mientras que los arcos corresponden a los segmentos de vías de comunicación que unen los nodos. Analizando la topología de un modelo de red de transporte junto con la distribución de celdas y LAs del sistema celular existente, es posible asociar fronteras entre dichas zonas con secciones de vía (arcos). Con ello, cada vez que un teléfono cambie de celda durante una llamada (handover) o cambie de LA mientras esté encendido (proceso LU), se podrá considerar que ha pasado por alguno de los arcos que conforman la correspondiente frontera. En aquellos casos en los que la frontera sólo comprenda un arco, el paso de terminales computado se asociará íntegramente a dicho arco. Sin embargo, cuando son más de uno los arcos implicados en una frontera, el nivel de resolución proporcionado por estos datos sólo informa del paso por la frontera, pero no permite distinguir sobre qué arco del total se ha realizado el traspaso. Dada esta situación, la cual es mayoritaria, la información de traspasos proporcionada será en forma



agregada por grupos de arcos, en lugar de a nivel individual de arco (formato tradicional de las estaciones de aforo).

El procedimiento a escoger dependerá, por tanto, de las especificaciones que se necesiten para abordar el problema en cuestión. Las fronteras a monitorizar vienen ya establecidas por el propio diseño de las redes de telefonía móvil, puesto que se utilizan las mismas celdas y/o áreas de localización definidas en el dimensionado y diseño de la red celular. El tamaño de un área de localización depende en gran medida de las celdas que la forman. El radio de las celdas depende de parámetros como tráfico que tienen que soportar, topología del terreno, tipo de antena o incluso edificios alrededor de la estación base, variando su tamaño desde cientos de metros a kilómetros.

Dada una región de estudio, puesto que las áreas de localización son agrupaciones de celdas, el número de fronteras entre celdas es mayor que entre LA y, por tanto, se monitorizaría un número mayor de arcos utilizando handovers que procesos de LU. No obstante, si la zona a estudiar es extensa (p.ej. una comunidad autónoma), las fronteras entre LAs por una celda concreta pueden englobar un número suficientemente elevado de arcos para abordar un determinado estudio. En esta Tesis, el principal objetivo de los datos de teléfonos está orientado al campo de la estimación de matrices origen–destino mediante volúmenes observados, por lo que interesa disponer de información en un número de arcos de la red suficientemente representativo. Así pues, será la dimensión de la red de transporte la que condicione la elección de tomar datos procedentes de llamadas (cambios de celda) o registros de LU (cambios de LA). Por ejemplo, si el estudio se centrara en una red urbana, la monitorización del paso de un LA a otro no es admisible dado que típicamente una ciudad es cubierta con un único LA y el número de posibles fronteras a monitorizar es muy bajo; en cambio, existirán numerosas celdas. Por ello, sería más apropiado en esos casos utilizar datos de llamadas al permitir monitorizar un mayor número de fronteras.

#### **4.3.1.2 Proceso de captura de datos**

Una vez conocidos los medios de los que se dispone, y los posibles procedimientos que ofrece la telefonía móvil, es necesario elegir cuál va a ser el procedimiento a seguir. Como se ha analizado en el apartado anterior, la dimensión de la red es determinante a la hora de seleccionar el evento de telefonía del que se obtendrán los datos fuente. Sin embargo, la implementación de los modelos de captura de datos de teléfonos está sujeta en gran medida a las características técnicas de la plataforma del operador que proporcione los datos. Esa plataforma deberá tener habilitada la funcionalidad necesaria para extraer información de las bases de datos relativa a

llamadas o cambios de área de localización. En su defecto, la disposición por parte del operador para la instalación de módulos que implementen tal extracción también condiciona el éxito del proceso de captura de datos.

Para la realización de esta Tesis se ha contado con la inestimable colaboración del operador VODAFONE que se ha prestado a proporcionar datos de teléfonos debidamente encriptados. En el caso de la plataforma del citado operador, los datos de llamadas están accesibles sin coste ni impacto en los módulos del sistema puesto que son almacenados con fines de facturación y control del funcionamiento de la red. Para el caso de datos asociados a procesos de LU, aunque éstos son generados automáticamente por el sistema, sólo se almacenan temporalmente. Por ello, sería necesario realizar un volcado periódico de las bases de datos del sistema en un servidor central del que tomar los datos. Esto requiere una modificación de los módulos existentes en la plataforma del operador ya que las APIs necesarias para su almacenamiento permanente no están accesibles en la actualidad.

Tras considerar las posibilidades ofrecidas y desde la perspectiva de bajo coste e impacto en módulos de la infraestructura del operador, se prefirió que los estudios numéricos desarrollados en este Capítulo utilizaran como datos fuente información de llamadas realizadas sobre un conjunto de celdas. En concreto, el conjunto de datos proporcionado por VODAFONE se corresponde con registros encriptados de llamadas efectuadas por sus abonados en celdas bajo estudio durante periodos de 24 horas. Un extracto del formato de los datos entregados por el operador se muestra en la Tabla 4-1. En particular, estas celdas han sido previamente seleccionadas por disponer en las proximidades de sus fronteras con otras celdas de una estación de aforo con la que contrastar y/o calibrar los valores de volúmenes de tráfico inferidos con el modelo diseñado.

**Tabla 4-1: Extracto de los datos entregados por el operador.**

<b>IMSI</b>	<b>ID Número Origen</b>	<b>ID Numero Destino</b>	<b>Fecha</b>	<b>Hora (hh:mm)</b>	<b>Duración (min)</b>	<b>Cell Id (Nombre)</b>
20240	3210	508	10/04/2008	9:30:00	1,53	GSM2
20240	3210	508	10/04/2008	9:31:00	2,45	GSM2
20241	3205	17362	10/04/2008	18:10:00	0,33	GSM1
20241	3205	1055	10/04/2008	18:11:00	1,12	GSM1

Es preciso mencionar que cualquier número que aparezca en la información proporcionada por el operador respecto a la identidad de los abonados se encuentra encriptado con un identificador único y permanente durante todo el intervalo de tiempo analizado (24 horas). Con ello se

garantiza la privacidad de los abonados pero sin impedir realizar un seguimiento de la movilidad que ha tenido cada terminal en términos de celdas en las que ha efectuado llamadas. Todo ello permite, finalmente, monitorizar el paso de terminales de una celda a otra para inferir volúmenes agregados por grupos de arcos que cruzan la frontera en cuestión.

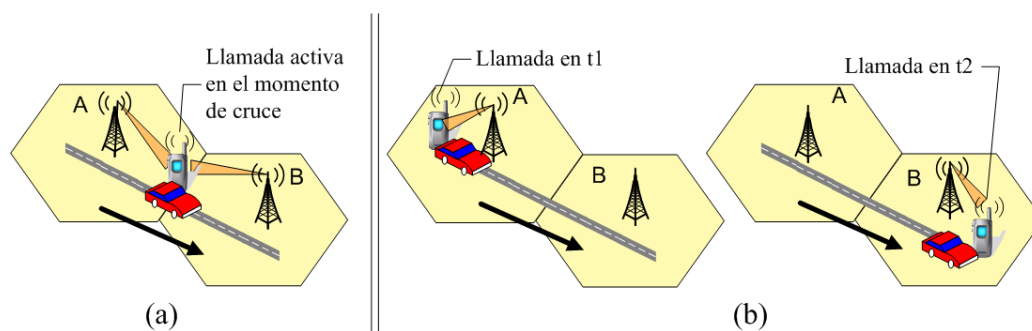
### **4.3.2 Conteo de paso de teléfonos entre celdas. Extracción de la muestra**

La inferencia estadística persigue estimar parámetros de una población a partir de una muestra de la misma. En general, la estadística lo que nos permite no es el conocimiento exacto de lo que ocurrirá, sino lo que más probablemente pueda ocurrir. Para ello dicha muestra o porción de población tiene como requisito que sea representativa de la población o universo al que pertenece, es decir, sea relativamente similar al conjunto de la que fue extraída. Dado que en nuestro problema la población la constituye el conjunto de vehículos que pasan por un tramo de vía, la selección de la muestra consistirá en tomar miembros de ese conjunto de vehículos. Según lo expuesto en apartados anteriores, la telefonía móvil está muy extendida en la sociedad actual, en especial en el caso de usuarios en movilidad dado que les proporciona un modo eficiente y rápido de comunicación. Bajo este enfoque, una posible muestra estaría relacionada con los teléfonos móviles pertenecientes a los ocupantes de los vehículos que circulan por una vía. El funcionamiento estándar de los sistemas de telefonía móvil permite conocer el flujo de teléfonos que pasan de una celda a otra mientras realizan llamadas. En entornos próximos a la red viaria, ese movimiento de teléfonos estaría directamente relacionado con los vehículos que han pasado por la zona asociada a la frontera entre celdas, en términos de la porción de vehículos que en su interior transportaban los teléfonos desde los que se han realizado tales llamadas.

Para la identificación de tal conjunto de teléfonos, lo más inmediato es tomar aquellos terminales que realizaron un handover. Se sabe que cuando un terminal con una llamada activa entra en el área de servicio de otra antena (celda), se ejecuta un handover, notificando a la red tanto su nueva posición en términos de identidad de celda como el instante de entrada. Con ello, en las bases de datos del sistema siempre se crea un registro para la entrada de un terminal con una llamada activa en una nueva celda. Un análisis de los instantes en los que se producen esos registros debido a handovers (cambios de celda) permite detectar el paso de terminales con llamadas activas por una zona asociada a la frontera de una celda a otra (Figura 4-1 a).

No obstante, emplear exclusivamente teléfonos que han generado handovers restringe en cierta medida el tamaño de la muestra dado que el handover requiere unas condiciones apropiadas para ser ejecutado. Éstas son que un usuario en movilidad (a bordo de vehículo) haga una llamada y que dure lo suficiente para que sea traspasada a otra celda. Puede darse el caso que la

llamada finalice antes de cambiar de celda, por lo que el handover no llegaría a realizarse y, a pesar de que el teléfono esté en movimiento, el teléfono no formaría parte de la muestra. Por ello, se deben considerar otras situaciones de movilidad que amplíen los criterios para la selección de la muestra. Se sabe que cuando un usuario realiza una llamada, por cuestiones de facturación, el sistema siempre almacena en una base de datos un registro conteniendo información relativa a esa llamada, como hora de inicio, duración, ID teléfono llamado, ID teléfono llamante o identificación de la celda donde se origina la llamada (LAI+CI). Por ello, también es factible detectar el paso de un terminal de una celda a otra sin necesidad de que esté realizando una llamada en el momento de cruzar la frontera, como en el caso del handover. Esto ocurre cuando desde un mismo terminal se realizan dos llamadas consecutivas en celdas diferentes y colindantes en un corto periodo de tiempo  $T$ , por ejemplo, 15 minutos (Figura 4-1 b). En ese caso, dado que hay constancia de que el terminal ha estado en dos celdas en un corto periodo de tiempo, se puede considerar que está en movimiento y que, por tanto, también ha pasado por la frontera entre celdas. Así pues, aunque el instante de paso de una celda a otra no se conozca con la misma exactitud que en el caso del handover, se supone que habrá ocurrido dentro de ese periodo  $T$ . Este tiempo ( $T=15$  minutos) es admisible dado que está dentro de los rangos de tiempo respecto de los cuales se suele proporcionar las medidas de intensidades de tráfico de las tradicionales estaciones de aforo. Esta pareja de llamadas junto a las que ejecutan handover permiten construir una muestra mayor de terminales para monitorizar su movimiento de una zona a otra, denominándose a esos tipos de llamadas como “llamadas en movilidad”.



**Figura 4-1: Llamadas en movilidad: (a) handover y (b) 2 llamadas en  $t_1$  y  $t_2$  tal que  $t_2 - t_1 \leq T = 15$  min.**

Es importante destacar que la situación mostrada en la Figura 4-1 b sólo se contabiliza como una única llamada en movilidad, pese a realizarse 2 llamadas. Esto es así dado que el concepto de movilidad se asocia al movimiento del terminal desde el cual se realizan llamadas y, en esta situación, es el mismo terminal el que realiza las dos llamadas, siendo éstas usadas para detectar su movimiento. Entonces, el análisis de registros de llamadas almacenados en las bases de datos en busca de llamadas en movilidad, realizadas en las celdas de interés, permitirá contabilizar los

terminales que se desplazan de una celda a otra como si existiera una especie de estación de aforo “virtual” en la frontera entre ellas.

Un modelo de red de transporte es una representación simplificada de la red vial y está formado, principalmente, por nodos y arcos, los cuales representan las intersecciones y los tramos viales, respectivamente, de la infraestructura vial. Analizando la distribución de las celdas y el modelo de red de transporte se identificarán las posibles fronteras observadas y se establecerá una correspondencia con los arcos del modelo. Una frontera observada o aforo “virtual” lo conforman los arcos cuyo nodo inicial y nodo final están en celdas distintas y contiguas (Figura 4-2). Con ello, cada vez que un teléfono con una llamada activa accede a la celda B procedente de otra A o el mismo teléfono realice dos llamadas consecutivas (<15 minutos) en la celda A y B, respectivamente, se podrá considerar que ha pasado por el aforo “virtual” posicionado en dicha frontera observada, es decir, por alguno de los arcos que la cruzan. La información de conteo proporcionada estará definida de forma agregada por grupos de arcos, en lugar de a nivel individual de arco (formato proporcionado por los tradicionales aforos).

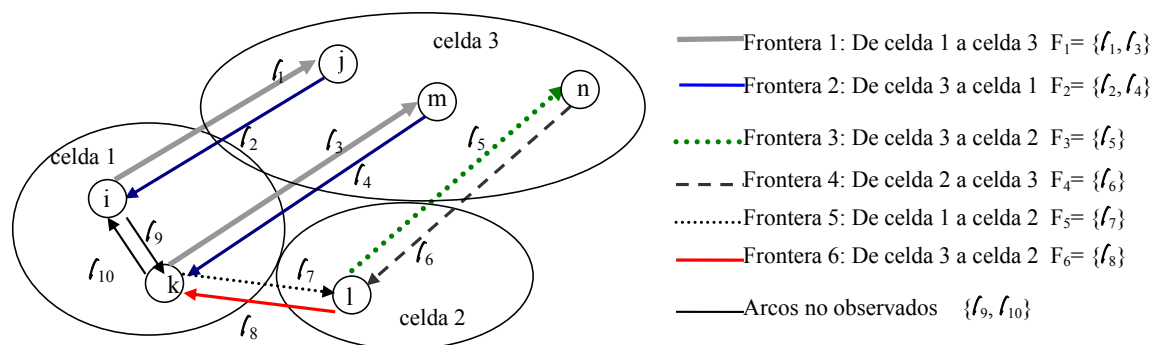


Figura 4-2: Definición de frontera observada o aforo “virtual” en un modelo de red.

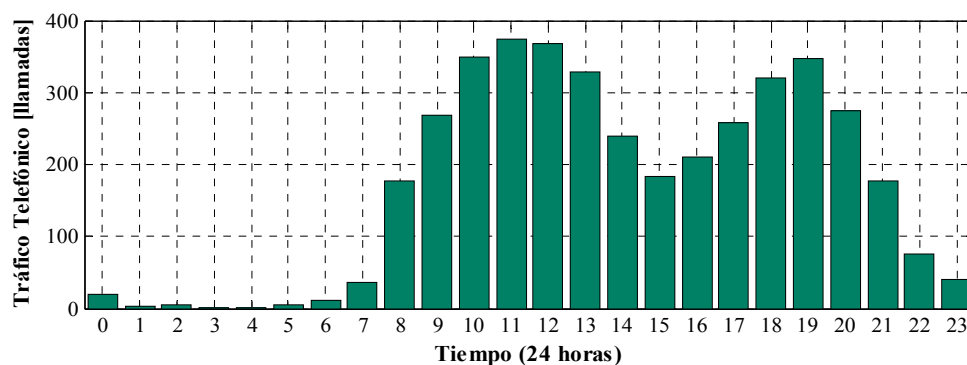
### 4.3.3 Análisis descriptivo de los datos de llamadas

Dada la naturaleza de la información usada para la inferencia de volúmenes de vehículos, asociada con llamadas realizadas por usuarios a bordo de éstos, surge la necesidad de realizar un análisis descriptivo de esos datos de llamadas. Este análisis, además de examinar el tráfico telefónico que tiene que soportar la red celular, debe atender un aspecto importante influido por la movilidad de los usuarios de teléfonos móviles, el cual viene referido a la conducta de éstos a la hora de realizar llamadas. En redes fijas, los usuarios en sus lugares de trabajo (parques industriales o empresariales) poseen diferentes hábitos de llamadas a los de áreas residenciales. Una diferenciación similar existe en los sistemas de telefonía móvil, donde existen diferentes patrones de llamadas para usuarios en movimiento y estáticos. Esto conlleva que el tráfico

telefónico en redes móviles dependa tanto del tiempo como del espacio, siendo esa dependencia espacial mucho más variable que en redes fijas.

Atendiendo a un contexto de movilidad de vehículos, un modo de abordar este análisis descriptivo es cuantificar la comodidad a la hora de efectuar llamadas. Las llamadas se pueden dividir en dos categorías: aquellas que son recibidas por el usuario, denominadas llamadas entrantes, y las que son realizadas por ese usuario, denominadas llamadas salientes. Lógicamente, un parámetro para evaluar la citada comodidad será la tasa de llamadas salientes. Destacar que cada llamada entrante tiene su origen en una saliente de otro usuario, por ello sólo se van a considerar las llamadas salientes de cada usuario al analizar el tráfico telefónico de red. La duración de una llamada también será otro parámetro característico a estudiar. Ambos parámetros están fuertemente afectados por la movilidad de los usuarios (estáticos, peatones u ocupantes de vehículos), ya que ésta condiciona la realización de llamadas. Por ejemplo, un conductor normalmente iniciará un número de llamadas menor comparado con un usuario ubicado en su trabajo; respecto a la duración, los pasajeros de vehículos harán llamadas más largas en comparación con las realizadas por conductores.

La Figura 4-3 muestra la intensidad del tráfico telefónico, representada en términos de número de llamadas realizadas, en una celda  $Z_k$  a lo largo de un día laboral típico. Para sistemas de telefonía móvil, los picos suelen producirse entre las 10:00 y 12:00 debido a la actividad comercial y laboral de los usuarios, mostrando un segundo pico entre 18:00 y 20:00 debido al fin de la jornada laboral.



**Figura 4-3: Tráfico telefónico (número de llamadas por hora) durante un día en una celda.**

Cabe señalar que la intensidad de tráfico telefónico puede variar de un día a otro, semanalmente, o incluso según el periodo del año (verano, Semana Santa,...), mostrando diferentes patrones asociados con los hábitos de los usuarios y/o las características de la zona gestionada por la celda. Aunque, normalmente, estos patrones de volumen de llamadas tienden a mostrar un

comportamiento similar en diferentes zonas. La Figura 4-4 muestra el número de llamadas realizadas a lo largo de un día laborable típico en diferentes celdas bajo estudio. En ella se aprecia que, a grandes rasgos, los patrones de llamadas son similares.

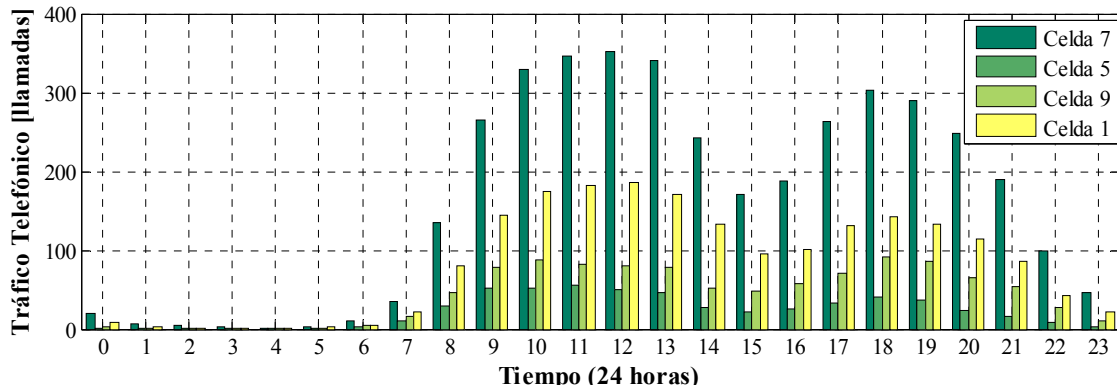


Figura 4-4: Promedio del número de llamadas por hora en diferentes celdas a lo largo de 18 días.

En la Figura 4-4 se puede observar que, esencialmente, existen tres períodos de actividad durante un día laborable típico, uno de muy poca actividad y dos claros picos donde la actividad es mayor, variando su rango de una celda a otra. El primero se corresponde con el período nocturno (00:00–7:00), donde la realización de llamadas es poco probable. El segundo periodo se extiende durante la mayor parte de la jornada laboral (8:00–16:00). El último período es el correspondiente a la tarde y el final del día (16:00–00:00). A su vez, se observa que la actividad telefónica en un día laborable aumenta bruscamente por la mañana entre las 7:00 y 10:00; en cambio, la transición hacia la noche refleja descensos más graduales.

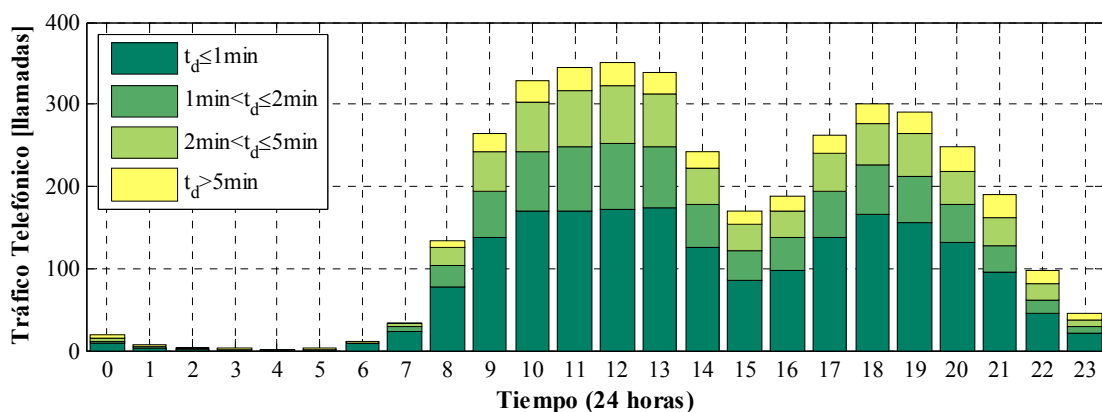


Figura 4-5: Número de llamadas por hora durante un día según la duración de la llamada  $t_d$ .

Un comportamiento similar experimenta la duración de las llamadas. En la Figura 4-5 se desglosan las llamadas realizadas según su duración, mostrando cómo varía la tendencia de duración de llamadas a lo largo del día. Analizando esa duración en términos medios (Figura

4-6), se aprecia más claramente que los usuarios tienden a hacer llamadas más largas a partir de las 19:00. La propia naturaleza del servicio condiciona esta variación en la duración, ya que puede verse influenciada por las tarifas existentes o, incluso, por la disponibilidad de los usuarios. Además, el promedio de duración de llamada puede variar según las características de la zona (p. ej., usuarios en zonas residenciales tienden a hacer llamadas más largas que usuarios en zonas comerciales durante la misma hora).

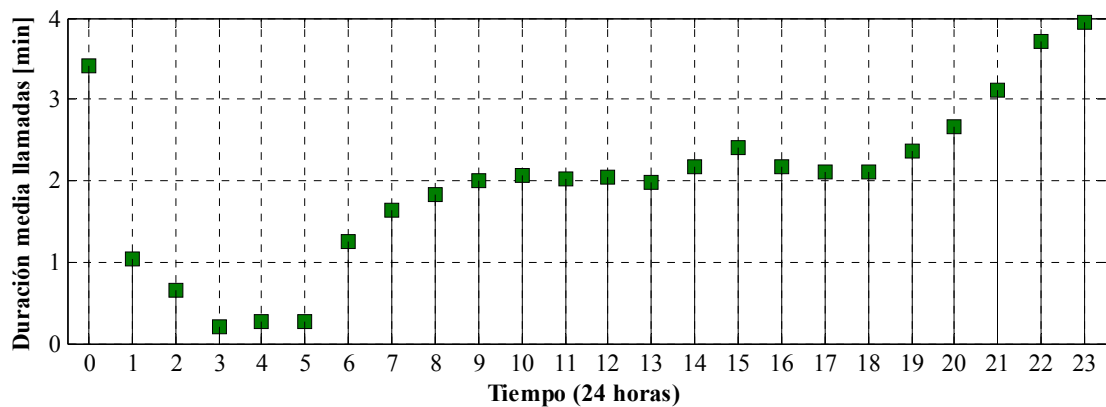


Figura 4-6: Duración media de las llamadas durante un día.

#### 4.3.4 Tratamiento de los datos

En el apartado anterior se han analizado los rasgos característicos de los datos en términos de número de llamadas realizadas y su duración. Esta sección se encarga de revisar en detalle otros aspectos significativos relacionados con los datos de llamadas, así como el tratamiento que se ha realizado sobre ellos para eliminar posibles errores.

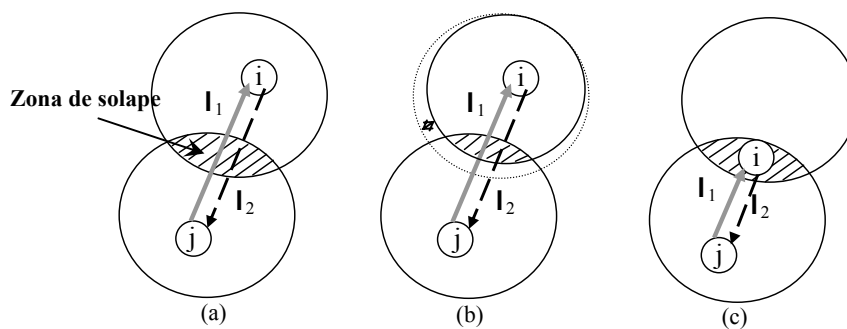
##### ■ Radio de cobertura de las celdas

Cada antena transmite a teléfonos móviles residentes en el espacio radioeléctrico de su área de cobertura denominada celda. Ese área de cobertura viene caracterizado por dos radios: i) el radio total de la celda y ii) el radio efectivo. El radio de servicio efectivo, definido mediante niveles potencia de señal, es aquel en el cual un teléfono se comunica con una única estación base servidora. El resto del área de servicio se diseña de modo que sea gestionado por más estaciones base, denominándose zona de solape. Es necesario tener en cuenta que el radio total de las celdas no es estático, sino que puede variar dinámicamente en función de ciertos factores aleatorios (condiciones meteorológicas, carga de llamadas...). No obstante, esta incertidumbre en el radio no afecta a la detección de paso de terminales de una celda a otra debido al propio concepto de frontera observada (Figura 4-2). En ella, se identifica la celda donde se ubican el nodo inicial y final de un arco, para conocer los arcos que cruzan una frontera observada. De



modo que la fluctuación del radio no suele tener gran impacto en aquellos nodos ubicados dentro del radio efectivo de una celda, como se aprecia en la Figura 4-7 (a) y (b).

Por otro lado, se ha comentado ya que el área de servicio de una celda suele superponerse parcialmente con el de una o más de sus celdas vecinas. Este solapamiento tiene lugar para evitar “agujeros” en la cobertura debido a una disminución del nivel de potencia de señal lejos de la antena servidora. El tamaño de la zona de solape se determina mediante el establecimiento de un umbral mínimo de señal recibida. Éste es un parámetro de diseño importante en las redes celulares puesto que garantiza que un abonado situado cerca del borde de una celda pueda recibir servicio de múltiples direcciones en caso que la servidora falle. Más información respecto al solapamiento entre celdas puede ser encontrada en Thornton y Grace (2005).



**Figura 4-7: Ejemplos de solapamiento entre celdas.**

Un valor típico de la zona de solape en una celda es el 30–40% de la cobertura completa de la celda (EIA/TIA IS-95, 1993). Este valor puede variar dependiendo, entre otros factores, de la densidad de población y/o de la interferencia causada por grandes edificios. En general, una celda tiene 6 celdas vecinas, de modo que su área total de solape es la suma del solapamiento con cada una de esas vecinas, situándose entonces la zona de solape de una celda con alguna de sus vecinas en torno al 5–7% del total de cobertura. En entornos no-urbanos, el radio típico de una celda se sitúa alrededor de 1 km, por tanto la zona de solape entre dos celdas puede llegar a ser mayor que el área definida para ser representada por un nodo. Cuando ocurre esta situación, el concepto de frontera observada no se cumple puesto que no se satisface la condición de estar el nodo inicial y final en celdas diferentes. En esta situación (Figura 4-7 c), el nodo (inicial o final) de un arco se encuentra en una zona que pertenece a dos celdas, pudiendo ser una de ellas la misma en la que se encuentra el otro nodo (final o inicial). En ese caso, esa frontera no puede ser considerada como frontera observada al no cumplir la definición. Todo ello sugiere la necesidad de prestar especial atención a la hora de identificar el conjunto de fronteras observadas dado que no todas las fronteras son susceptibles de ser monitorizadas.

### ■ *Tráfico peatonal*

La presencia de peatones es un aspecto importante a considerar en la inferencia de volúmenes de vehículos mediante la realización de llamadas en movilidad. Estas llamadas están asociadas a terminales desplazándose de una celda a otra, sin conocer la velocidad a la que lo hacen. Esto implica que las llamadas en movilidad observadas pueden ser efectuadas por teléfonos en movimiento tanto a bordo de vehículos como de peatones en el caso de que la frontera entre celdas esté próxima a entornos urbanos. En base a ese razonamiento, no es posible distinguir únicamente con datos de llamadas si un terminal que se mueve de una celda a otra pertenece a un usuario de vehículo o a un peatón. Introduciendo modificaciones en la red, la velocidad de traspaso podría ser detectada, resolviéndose así este problema. No obstante, dado que se ha optado por emplear datos de telefonía sin introducir modificaciones, se considerará que las vías que soporten tráfico de peatones no son válidas para ser monitorizadas. Es decir, si existiera una frontera entre celdas asociada a ese tipo de vías, las llamadas observadas procederían de usuarios de diferente tipo de movilidad. Como el interés se centra en el estudio de vehículos, cuando una frontera entre celdas esté formada por arcos con un tráfico de peatones significativo, esa frontera es desechada del conjunto de fronteras observadas, aunque cumpla la definición de frontera observada. Sólo fronteras cuyos arcos no poseen un tráfico de peatones significativo son consideradas como válidas para formar parte del conjunto de fronteras observadas. Esto implica la necesidad de un procesado previo para evaluar las fronteras entre celdas y calificarlas como aptas para el proceso de inferencia.

### ■ *Puntos kilométricos en estudio*

Se pretende determinar el volumen de vehículos asociado a un número conocido de llamadas en movilidad generadas en una frontera entre celdas. Para alcanzar este objetivo, se van a plantear una serie de modelos que, con fines predictivos, se encarguen de proporcionar esa magnitud de un modo similar a las clásicas estaciones de aforo. La elaboración de esos modelos no es una tarea sencilla ya que, como se expone más adelante, no existe una dependencia directa entre ambas magnitudes. Por ello, una vez definidos los modelos, se va a emplear un histórico de datos, tanto de llamadas en movilidad como de vehículos que han pasado por una zona, no sólo para llevar a cabo el ajuste de parámetros de esos modelos sino para la validación de las predicciones realizadas. Así pues, se exige que las fronteras entre celdas dispongan de una estación de aforo en las proximidades encargada de proporcionar los volúmenes de vehículos que realmente atraviesan la frontera entre celdas para constituir el citado histórico de datos.

Asimismo, bajo las condiciones expuestas en el punto anterior, la selección de los puntos kilométricos en los que se localizan las fronteras entre celdas en las que se realizará la

monitorización estará restringida a zonas en las que sea poco probable que los usuarios en movilidad existentes se correspondan con peatones. Seleccionando fronteras entre celdas cruzadas por vías donde el tráfico sea esencialmente de vehículos, alejadas de entornos urbanos, se elimina el efecto negativo que los peatones introducen. Esta será una condición necesaria para considerar como válida la información que se pueda derivar de las llamadas.

Finalmente, para identificar tales fronteras se realizó un estudio de campo sobre vías pertenecientes a anillos alrededor del área metropolitana de Sevilla. Una vez conocidos los puntos de la red viaria en los que existen estaciones de aforo, se seleccionaron los más adecuados para realizar el estudio. La Figura 4-8 muestra los 6 puntos kilométricos escogidos pertenecientes a las vías sobre las que se realiza la monitorización. Estos 6 puntos dan lugar a 12 fronteras observadas, dos para cada punto kilométrico, estando cada una de ellas asociada a un sentido de paso. Sobre esos puntos, el Centro de Gestión de Tráfico de la DGT facilitó las mediciones de intensidades horarias de vehículos y VODAFONE los registros de llamadas generados en las celdas implicadas en las fronteras durante los días seleccionados para ser observados. Con ellos se elaboró el histórico de datos que relacionaba número de vehículos y de llamadas en movilidad  $\{(x_1, y_1), \dots, (x_n, y_n)\}$ . La Tabla 4-2 muestra los IDs de cada una de las fronteras observadas en este estudio.



Figura 4-8: Puntos kilométricos en los que en sus proximidades existe una frontera entre celdas.

Tabla 4-2: ID fronteras asociados a cada punto kilométrico y sentido.

<b>ID Frontera</b>	1	2	3	4	5	6
<b>PK</b>	A 49 PK 46 S.HUELVA	A 4 PK 518 S.MADRID	A 4 PK 488 S.MADRID	A 4 PK 527 S.MADRID	A 92 PK 32 S.MALAGA	A 92 PK 43 S.MALAGA
<b>ID Frontera</b>	7	8	9	10	11	12
<b>PK</b>	A 4 PK 518 S.SEV	A 49 PK 46 S.SEV	A 92 PK 32 S.SEV	A 92 PK 43 S.SEV	A 4 PK 488 S.SEV	A 4 PK 527 S.MADRID

### ■ *Días de observación*

Los datos con los que se debe trabajar proceden de medidas u observaciones tomadas a partir de la muestra. Dado que se pretende que la información extraída sea lo más representativa posible, se seleccionarán para monitorizar días que, a priori, presentan pocas variaciones respecto a un día “típico”. Esto implica que las mediciones se deben llevar a cabo en días centrales de la semana: Martes, Miércoles y Jueves. Además, deben estar sujetos a las siguientes condiciones: no ser festivo, que el día anterior y posterior tampoco sea festivo, y que no pertenezca a una semana con eventos especiales (Semana Santa, Feria de Abril,...). En base a esos criterios, los días seleccionados para tomar las observaciones de los datos de llamadas son 18 (martes, miércoles y jueves de 6 semanas), los cuales se muestran en la Tabla 4-3.

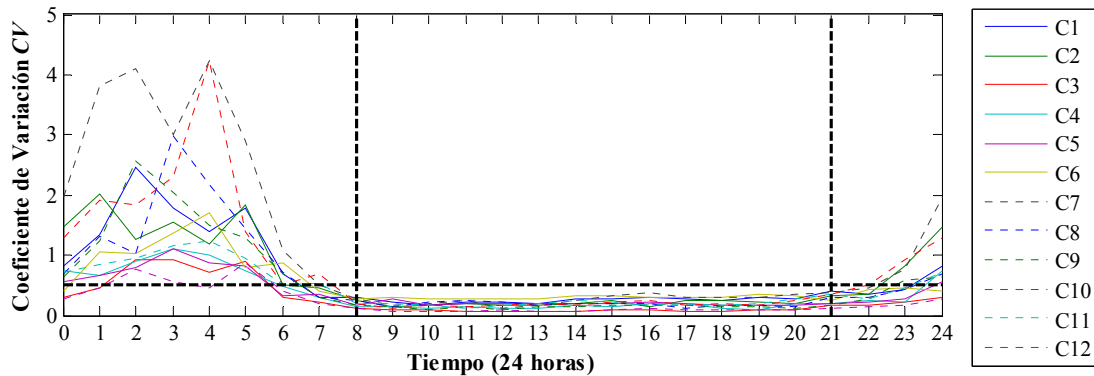
**Tabla 4-3: Días observados.**

MARZO 2008							ABRIL 2008							CÓDIGO DE COLORES	
L	M	X	J	V	S	D	L	M	X	J	V	S	D		
					1	2		1	2	3	4	5	6		
3	4	5	6	7	8	9	7	8	9	10	11	12	13		
10	11	12	13	14	15	16	14	15	16	17	18	19	20		
17	18	19	20	21	22	23	21	22	23	24	25	26	27		
24	25	26	27	28	29	30	28	29	30						
31															

### ■ *Horario de observación*

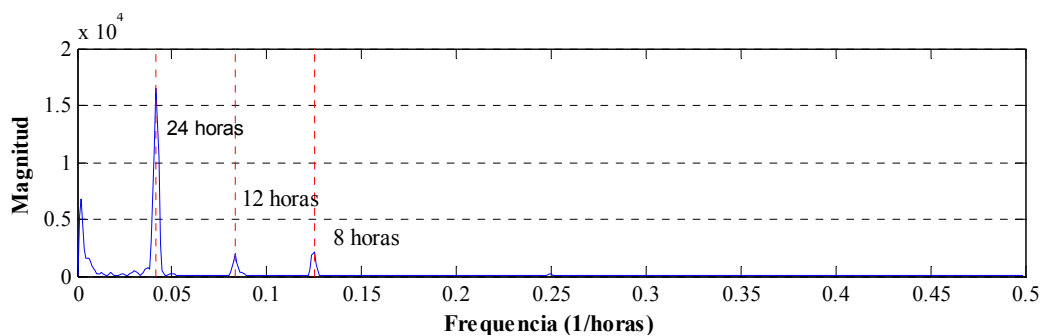
El comportamiento de los usuarios a la hora de realizar llamadas también está condicionado por la hora, bien por las tarifas existentes, la jornada laboral o porque en ciertos periodos puede ser socialmente inadecuado realizar una llamada. Incluso, éste puede verse influido por la celda según las características de la zona a la que da servicio (área residencial, centro comercial...). Prueba de ello es que la actividad telefónica durante la franja horaria entre la 1:00 a.m. y 6:00 a.m. es muy reducida, siendo poco probable hacer una llamada; en cambio, a partir de las 18:00 aumenta la tendencia a realizar llamadas por el inicio de tarifas reducidas. En este sentido, es conveniente emplear medidas de dispersión para determinar el rango horario donde el número de llamadas realizadas no presenta mucha dispersión respecto de la media. Una de estas medidas es el coeficiente de variación ( $CV$ ), definido como el cociente entre la desviación típica y la media,  $CV = \frac{\sigma}{\mu}$ , el cual cuantifica la separación (dispersión o variabilidad) de los valores de la distribución respecto a un valor central. El valor del  $CV$  será igual a 0 cuando no existan diferencias entre los datos, resultando entonces una distribución totalmente homogénea. La

Figura 4-9 muestra el *CV* conseguido para las distribuciones de llamadas realizadas en cada una de las celdas bajo estudio, procedentes de datos cedidos por VODAFONE.



**Figura 4-9: Coeficiente de Variación de la distribución de llamadas en diferentes celdas.**

Desde el punto de vista práctico, interesa centrar el análisis en un rango horario donde las curvas presenten valores bajos de *CV*. Esto ocurre dentro del intervalo 8:00–21:00 observando la Figura 4-9, en la cual el valor de *CV* permanece en todo momento por debajo de 0.5. Fuera de ese intervalo, los valores de *CV* alcanzados son bastante mayores que la unidad, indicando en ese caso que la realización de llamadas posee una variabilidad mayor respecto de la media. Entonces, se tomará como intervalo de observación la franja horaria entre las 8:00 y 21:00, en la cual la realización de llamadas responde a un comportamiento, en cierto grado, homogéneo para todas las celdas. Con ello se consigue que los modelos de predicción diseñados sean más estables para cualquier periodo de tiempo comprendido en dicha franja.



**Figura 4-10: Densidad Espectral de Potencia de la distribución de llamadas en una celda.**

Por otro lado, realizando un análisis de la actividad telefónica generada por usuarios de teléfonos móviles, se aprecian ciertos ciclos de comportamiento relacionados con los patrones de movilidad de los usuarios. Analizando la Densidad Espectral de Potencia de las llamadas (Figura 4-10), para realizar un estudio en frecuencia, se observan 3 picos en torno a periodos de

24 horas, 12 horas y 8 horas, los cuales son representativos de la periodicidad del conjunto de datos de llamadas analizados. Lógicamente, el pico de las 24 horas está asociado a la frecuencia con que los datos de llamadas son almacenados. El pico en torno a las 12 horas representa ciclos repetitivos de comportamiento cada 12 horas correspondientes a pautas día/noche. Asimismo, el pico de periodicidad en torno a las 8 horas responde a comportamientos asociados a la jornada laboral.

### 4.3.5 Elaboración de modelos de predicción

#### 4.3.5.1 Introducción

La inferencia estadística es una técnica mediante la cual se deducen propiedades o se hacen predicciones en base a una información parcial o completa obtenida mediante técnicas descriptivas sobre individuos de una muestra. En nuestro problema en cuestión, se pretende predecir volúmenes de tráfico de vehículos a partir de información de llamadas generadas por una muestra constituida por ciertos teléfonos móviles que van a bordo de esos vehículos. Por ello, inicialmente, se buscó analizar la relación de dependencia entre vehículos y las denominadas llamadas en movilidad (Figura 4-1). Para abordar el problema, se ha dispuesto de un histórico dado por una muestra bidimensional que relaciona ambas variables  $\{(x_1, y_1), \dots, (x_n, y_n)\}$ , siendo  $x_i$  e  $y_i$  observaciones de llamadas en movilidad y de vehículos, respectivamente, que han pasado por alguna frontera observada entre celdas en un intervalo de tiempo  $t_i$ .

Es en este marco de cuantificación de dependencia entre variables en el que aparecen unas funciones denominadas cópulas. Las cópulas son una poderosa herramienta para representar las relaciones de dependencia entre distintas variables aleatorias a través de la distribución de probabilidad conjunta (Nelsen, 2006). En base a ellas, la función de distribución conjunta de una serie de variables puede expresarse como la función cópula aplicada sobre las distribuciones marginales consideradas individualmente, esto es,  $F(x_1, x_2, \dots, x_m) = C(F_1(x_1), F_2(x_2), \dots, F_m(x_m))$ . Así pues, una cópula,  $C$ , es una función de distribución multivariante cuyas leyes marginales se distribuyen uniformemente entre  $[0,1]$ . En el caso bivalente,  $C(u, v) = p[U \leq u, V \leq v]$  es una función de distribución definida en  $[0,1]^2 \rightarrow [0,1]$ , donde  $U$  y  $V$  son dos variables aleatorias uniformemente distribuidas. Entonces, dadas dos variables aleatorias  $X$  e  $Y$  con distribuciones marginales  $F_X(x)$  y  $F_Y(y)$  respectivamente, existe una cópula  $C$ , tal que  $F_{XY}(x,y) = C(F_X(x), F_Y(y))$ .

Para la obtención de la cópula asociada a las variables de interés  $X$  (número de llamadas en movilidad) e  $Y$  (número de vehículos), se ha optado por el cálculo de la cópula empírica. Las

cóputas empíricas fueron estudiadas originalmente por Deheuvels (1979). La idea consiste en construir una función cópula a partir de valores muestrales  $\{(x_1, y_1), \dots, (x_n, y_n)\}$  recogidos para las variables univariantes sin establecer dependencia de ningún parámetro. De esta forma, la cópula es no paramétrica y queda definida únicamente a partir de la muestra de datos disponible.

Sea  $(x, y) \equiv \{(x_1, y_1), \dots, (x_n, y_n)\}$  una muestra de tamaño  $n$  obtenida a partir de una distribución bivalente  $(X, Y)$ , la definición de cópula empírica (Nelsen, 2006) responde a la expresión:

$$C_n\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{n^\circ \text{ de pares } (x, y) \text{ de la muestra tales que } x \leq x_{(i)} \text{ e } y \leq y_{(j)}}{n} = \frac{1}{n} \sum_{k=1}^n I(x_k \leq x_{(i)}, y_k \leq y_{(j)}) \quad i, j = 1, \dots, n \quad (6)$$

con  $x_{(i)}$  e  $y_{(j)}$   $1 \leq i, j \leq n$ , los estadísticos de orden definidos a partir de la muestra. Deheuvels (1979) demostró que cuando el tamaño muestral crece, la cópula empírica converge a la cópula verdadera. Es por ello que la cópula empírica constituye un instrumento adecuado para estudiar, a grandes rasgos, la relación de dependencia entre  $U$  y  $V$ , y por tanto, la relación entre  $X$  e  $Y$ .

En este sentido, existe una cópula que representa la condición de independencia entre variables, definida como  $\Pi(u, v) = u \cdot v$ . Se dice que dos variables  $X$  e  $Y$  no son dependientes si y sólo si  $C(u, v) = \Pi(u, v)$ . Cuando  $C(u, v) \neq \Pi(u, v)$ , las variables  $X$  e  $Y$  son dependientes y la dependencia depende de las variables a través de transformaciones crecientes no lineales de las variables. Esta es una propiedad destacable de las funciones cópula, la posibilidad para detectar independencia entre variables frente a otras medidas clásicas que no lo permiten como el coeficiente de correlación lineal (que en caso de ser cero, únicamente permite confirmar la ausencia de dependencia lineal entre las variables).

Una manera sencilla de presentar y comparar las cópulas es a través de los diagramas de contorno. El diagrama de contorno de una cópula es la representación gráfica de sus curvas de nivel, las cuales son conjuntos en  $[0, 1]^n$  dados por  $C(u, v) = k$ , con  $k$  constante. Empleando estos diagramas, se pueden comparar la cópula de independencia y la cópula empírica obtenida para el conjunto de datos  $\{x_i, y_i\}$  en los puntos kilométricos o fronteras estudiadas (Figura 4-11). Se aprecia que ambas cópulas muestran un comportamiento similar. A la vista de la similitud entre ambas cópulas, y dado que la cópula empírica converge hacia la verdadera conforme el tamaño muestral crece (Deheuvels, 1979), es posible afirmar que las variables  $X$  e  $Y$  no poseen una relación de dependencia directa.

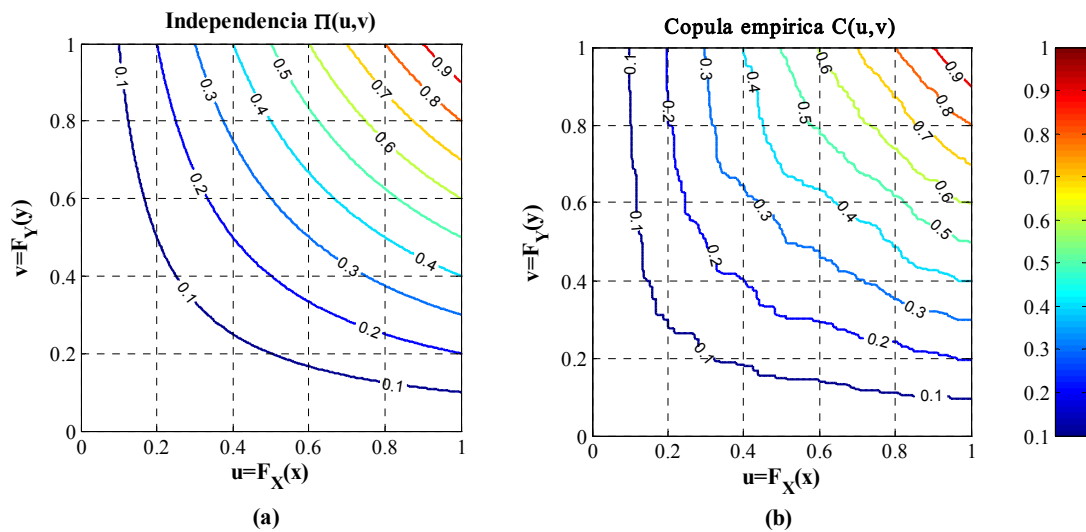


Figura 4-11: Diagrama de contorno de isoprobabilidad de la cópula independencia y la empírica.

En estas condiciones, se hace necesario introducir en los modelos otras dependencias que se encarguen de medir esa relación desconocida entre  $X$  e  $Y$  que no tiene por qué ser cuantificable por un único parámetro. A la vista del análisis descriptivo de los datos de llamadas, definir modelos en los que, además de las llamadas en movilidad, intervenga una dependencia temporal parece lo más intuitivo. Desde el enfoque de datos discretos, existe una amplia gama de modelos estadísticos con capacidad de predicción como modelos lineales generalizados, modelos paramétricos, etc. En los siguientes apartados se van a plantear una serie de modelos empleando ecuaciones paramétricas para modelar la relación entre las variables  $Y$  (número de vehículos) y  $X$  (número de llamadas en movilidad), junto a otros términos dependientes, de cara a que los modelos puedan ser utilizados con un objetivo predictivo.

Para la elección del modelo más adecuado parece conveniente decidirse por aquél que mejores resultados proporcione. Como se ha comentado, para abordar el problema se dispone de un histórico dado por una muestra bidimensional que relaciona ambas variables. En este sentido, para cada uno de los días observados se dispone de una muestra formada por 13 parejas de datos: número de llamadas y número de vehículos registrados cada hora desde las 8:00 a las 21:00. Con objeto de disponer de un tamaño de muestra suficiente para ajustar un modelo, se obtuvo para cada punto kilométrico una muestra mayor mediante la agregación de los  $D$  días observados ( $D=18$ ) en una sola muestra. Finalmente, el tamaño de la muestra resultante,  $\{(x_1, y_1), \dots, (x_n, y_n)\}$ , ha sido de  $n=234$  registros ( $18 \times 13$ ) para cada uno de los  $K=12$  puntos kilométricos estudiados. De modo que el estudio predictivo hará uso de esas variables para ajustar los parámetros de los modelos. Esos parámetros se encargarán de cuantificar adecuadamente la relación entre las variables de cara a que el modelo proporcione medidas admisibles de volúmenes de vehículos sin necesidad de instalar una estación de aforo en la zona.



Lógicamente es imposible predecir con total exactitud, por lo que los resultados obtenidos por los modelos contendrán ciertos errores de predicción. En este sentido, es importante indicar que no todo el histórico será utilizado para el proceso de estimación de parámetros, ya que una parte del histórico, denominada conjunto de validación, estará reservada para ser utilizada con posterioridad en la evaluación de las predicciones realizadas, no participando en el ajuste. Por lo general, se reserva una tercera o cuarta parte de los datos disponibles para validar los resultados. El resto del histórico se utilizará para realizar los ajustes y recibe el nombre de conjunto de entrenamiento.

En la notación de los modelos que a continuación se van a formular se empleará el índice  $i$  para designar el intervalo horario,  $d$  para el día observado, y  $k$  para el punto kilométrico o frontera.

#### **4.3.5.2 Modelo basado en la probabilidad de llamada y probabilidad de handover**

Se va a plantear un modelo de predicción basado en el fenómeno físico asociado a las llamadas en movilidad (Figura 4-1). Para definirlo se ha considerado que un teléfono móvil ha efectuado una llamada en movilidad cuando el usuario o bien ha realizado una llamada en cada una de las dos celdas que componen la frontera en un pequeño intervalo de tiempo (suceso M), o bien realiza una llamada en su celda origen con la suficiente duración como para que se efectúe el handover (suceso L).

Es importante destacar que, aunque el suceso M exige la realización de dos llamadas, sólo se contabiliza como una única llamada en movilidad dado que el concepto de movilidad se asocia únicamente al terminal desde el cual se han realizado las dos llamadas, necesarias para la detección de movimiento.

A la vista de estas situaciones, parece lógico que las observaciones de llamadas en movilidad estén afectadas por el hecho de que las llamadas se realicen en periodos de tiempo donde sea más o menos probable realizarlas. Además, en el caso de handovers, la duración de las llamadas también afectará a esas observaciones ya que, cuanto mayor sea la duración de una llamada, más probable será que se inicie un handover si el teléfono está en movimiento. Por ello, la dependencia temporal del comportamiento de los usuarios al realizar llamadas, puesta de manifiesto al comprobar como el número total de llamadas y su duración media varían a lo largo del día (apartado 4.3.3), debe ser considerada a la hora de diseñar un modelo que permita estimar el número de vehículos conocido un número de llamadas en movilidad.

Para la formulación de un modelo teórico que represente con fidelidad la relación entre el número de llamadas en movilidad y el de vehículos, se plantea una expresión que modele cada una de las situaciones anteriormente citadas haciendo uso de parámetros característicos de las

llamadas, como probabilidad de hacer una llamada o su duración, según el intervalo horario  $t_i$  y la frontera observada  $k$ . Cabe resaltar que la naturaleza de las celdas estudiadas, alejadas de entornos urbanos, permite suponer que prácticamente la totalidad de las llamadas en movilidad observadas el intervalo horario  $t_i$  y la frontera  $k$ ,  $n_{ll\ mov}(t_i, k)$ , serán realizadas por usuarios en movimiento asociados al tráfico vehicular sobre las vías existentes en esa frontera. Estos usuarios constituyen, por tanto, una muestra de la población de vehículos que han cruzado la frontera  $k$  en cada intervalo  $t_i$ ,  $n_{veh}(t_i, k)$ .

Respecto al primer tipo de llamadas en movilidad (suceso M), se hace necesario introducir un término que refleje la importancia de la probabilidad de que un teléfono del operador monitorizado, y a bordo de un vehículo, realice dos llamadas en el intervalo horario  $t_i$ . Teniendo en cuenta que el evento “hacer dos llamadas a bordo de vehículo con el operador de interés” es una superposición de dos realizaciones independientes del evento  $A \equiv$  “hacer una llamada a bordo de vehículo con el operador de interés”, y considerando que este evento  $A$  tiene una probabilidad de éxito que varía con el tiempo,  $P_A \equiv P_{ll\ veh}(t_i)$ , entonces la probabilidad de éxito del suceso M se puede aproximar por  $P_{ll\ veh}(t_i) \cdot P_{ll\ veh}(t_i)$ . Dado que se asume que la población que genera las llamadas en movilidad son el conjunto de vehículos que han cruzado la frontera,  $n_{veh}(t_i, k)$ , el número de llamadas en movilidad originadas por el suceso M vendrá dado por:

$$n_M(t_i, k) = n_{veh}(t_i, k) \cdot P_M = n_{veh}(t_i, k) \cdot P_{ll\ veh}(t_i) \cdot P_{ll\ veh}(t_i) = n_{veh}(t_i, k) \cdot P_{ll\ veh}^2(t_i) \quad (7)$$

Para el segundo tipo de llamada en movilidad (suceso L), el planteamiento es similar pero introduciendo un término relacionado con la probabilidad de que la llamada ejecute handover. Así pues, la probabilidad de éxito de este suceso L precisa de dos terminos; por un lado, la antes mencionada probabilidad de  $A \equiv$  “hacer una llamada a bordo de vehículo con el operador de interés”, y por otro, la probabilidad de que dicha llamada requiera handover,  $P_h$ . Respecto a ello, la probabilidad de que sobre una llamada se ejecute un handover viene dada por el hecho de que la duración de la misma,  $t_d$ , supere el tiempo de permanencia en la celda,  $t_p$ . Asumiendo que la población que genera las llamadas en movilidad es el conjunto de vehículos que han cruzado la frontera,  $n_{veh}(t_i, k)$ , el número de llamadas en movilidad generadas en las condiciones asociadas al suceso L vendría definido como:

$$n_L(t_i, k) = n_{veh}(t_i, k) \cdot P_L = n_{veh}(t_i, k) \cdot P_{ll\ veh}(t_i) \cdot P_h(t_i, k) \quad (8)$$

El total de llamadas en movilidad observadas en el intervalo horario  $t_i$  se corresponde con la suma de ambos conjuntos de llamadas  $\{n_M$  y  $n_L\}$ , obteniéndose una expresión en la que el valor observado de las llamadas en movilidad se corresponde con la media de vehículos que realizan una llamada en movilidad:

$$\begin{aligned} n_{ll\ mov}(t_i, k) &= n_{veh}(t_i, k) \cdot P_{ll\ veh}^2(t_i) + n_{veh}(t_i, k) \cdot P_{ll\ veh}(t_i) \cdot P_h(t_i, k) \\ &= n_{veh}(t_i, k) \cdot [P_{ll\ veh}^2(t_i) + P_{ll\ veh}(t_i) \cdot P_h(t_i, k)] \end{aligned} \quad (9)$$

Esta expresión es obtenida considerando despreciable la probabilidad de sucesos tales como que se realicen dos llamadas desde el mismo vehículo en la misma celda, o que se llame en cada celda efectuando adicionalmente el proceso de handover. Dichos sucesos, posibles en la realidad, no llegaron a presentarse en ninguna ocasión sobre la muestra tomada en los puntos monitorizados, por lo que no fueron considerados como situaciones en movilidad factibles. Atendiendo a la expresión anterior, la probabilidad de que un terminal efectúe una llamada en movilidad, denominada  $P(t_i, k)$ , puede aproximarse por:

$$P(t_i, k) = P_{ll\ veh}^2(t_i) + P_{ll\ veh}(t_i) \cdot P_h(t_i, k) \quad (10)$$

Entonces, el número de vehículos que cruzan una frontera  $k$  en el intervalo horario  $t_i$  vendrá dado por:

$$n_{ll\ mov}(t_i, k) = P(t_i, k) \cdot n_{veh}(t_i, k) \Rightarrow n_{veh}(t_i, k) = \frac{n_{ll\ mov}(t_i, k)}{P(t_i, k)} = \frac{n_{ll\ mov}(t_i, k)}{P_{ll\ veh}^2(t_i) + P_{ll\ veh}(t_i) \cdot P_h(t_i, k)} \quad (11)$$

Esta expresión proporciona un valor para el volumen de vehículos que pasan por frontera entre celdas en función de las llamadas en movilidad detectadas y otras variables relacionadas con las llamadas. El número de llamadas en movilidad asociadas al cruce de una frontera en el intervalo horario dado,  $n_{ll\ mov}(t_i, k)$ , se obtiene mediante un análisis de los registros de llamadas realizadas sobre las celdas implicadas en la frontera  $k$ , proporcionados por el operador. Para los otros términos: probabilidad de que un vehículo haga una llamada y probabilidad de handover,  $P_{ll\ veh}$  y  $P_h$  respectivamente, se precisa de un tratamiento estadístico adicional atendiendo a la dependencia temporal de las características de las llamadas. Los siguientes puntos exponen en detalle el método desarrollado para la obtención de los citados términos.

■ *Probabilidad de hacer una llamada a bordo de vehículo con el operador de interés,  $P_{ll\ veh}(t_i)$*

Es habitual que un conjunto de datos siga algún tipo de distribución estadística. Definir adecuadamente la probabilidad de éxito de algún suceso puede resultar tedioso si no se conoce esa distribución. La propia representación de los valores nos puede sugerir el empleo de alguna de las muchas distribuciones conocidas (Poisson, Uniforme, Normal, etc.), determinando así la que mejor se ajuste a los datos. Sin embargo, hay algunas formas de acercarse a tal término de probabilidad sin necesidad del conocimiento previo de la distribución; tal es el caso de un procedimiento empírico en el cual la asignación de las probabilidades de los sucesos de interés se basa en la información observada. En esos casos, se requiere realizar un gran número de experimentos para hallar la probabilidad de éxito del suceso, en nuestro caso la probabilidad de

hacer una llamada a bordo de un vehículo con el operador monitorizado (Vodafone). Bajo este enfoque empírico, la probabilidad se determina sobre la base de la proporción de veces que ocurre un evento favorable o éxitos respecto al total de resultados posibles.

$$\frac{\text{Nº veces que ocurre el evento de interés}}{\text{Nº total de resultados posibles}}$$

Como ya se ha comentado anteriormente, la naturaleza de las celdas estudiadas, alejadas de entornos urbanos y de zonas residenciales, permite suponer que prácticamente la totalidad de las llamadas en movilidad observadas en cada una de las fronteras serán realizadas por usuarios inmersos en el tráfico vehicular. Así pues, la probabilidad de “hacer una llamada a bordo de vehículo con el operador de interés”,  $P_{ll\ veh}(t_i)$ , será igual al número de llamadas de dicho operador realizadas a bordo de vehículos en el intervalo horario  $t_i$  dividido por el número de total de vehículos que han pasado en dicho intervalo:

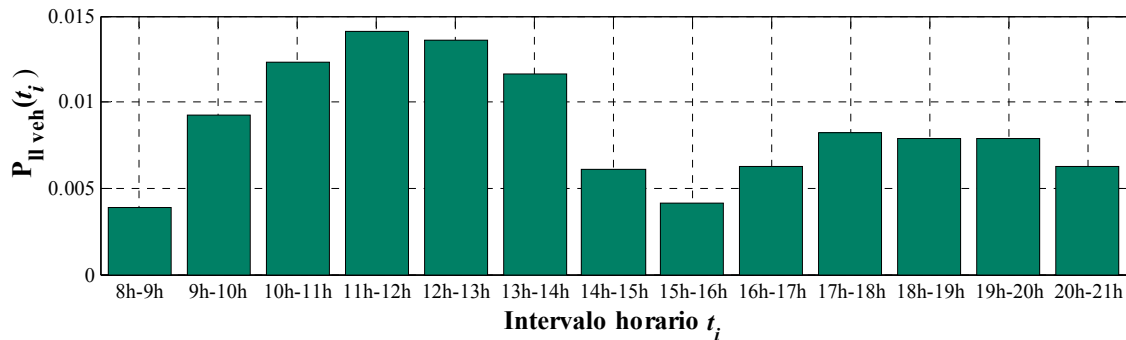
$$\frac{\text{Nº veces que ocurre el evento de interés}}{\text{Nº total de resultados posibles}} \Bigg|_{t_i} = \frac{\text{Nº llamadas realizadas por vehículos}}{\text{Nº total de vehículos cruzan fronteras}} \Bigg|_{t_i}$$

En esta expresión, el número total de vehículos que cruzan las fronteras se conoce a partir de las estaciones de aforo existentes en las proximidades de cada frontera. Por otro lado, el término de llamadas del operador de interés realizadas desde vehículos estará relacionado directamente con las llamadas en movilidad observadas. Teniendo en cuenta que una llamada en movilidad originada por el suceso M implica la realización de dos llamadas por parte del terminal que viaja a bordo del vehículo, la probabilidad de “hacer una llamada a bordo de vehículo con el operador de interés” en el intervalo horario  $t_i$  se define como:

$$P_{ll\ veh}(t_i) = \frac{\sum_{d=1}^D \sum_{k=1}^K [n_L(t_{i,d}, k) + 2 \cdot n_M(t_{i,d}, k)]}{\sum_{d=1}^D \sum_{k=1}^K n_{veh}(t_{i,d}, k)} \quad (12)$$

siendo  $K$  el número total de fronteras estudiadas y  $D$  el número total de días observados. Como se puede notar, se ha establecido una dependencia según el intervalo horario  $t_i$  considerado. A la vista de las conclusiones alcanzadas en la sección 4.3.3, la fuerte dependencia temporal existente en el comportamiento de los usuarios para la realización de llamadas también influye en esta probabilidad. La Figura 4-12, que representa la probabilidad  $P_{ll\ veh}(t_i)$  obtenida empíricamente, refleja tal variación con el tiempo. Cabe destacar que esta probabilidad  $P_{ll\ veh}(t_i)$  es válida para cualquiera de las fronteras monitorizadas ya que el evento “hacer llamada a bordo de vehículo con el operador de interés” es independiente de la celda en la que se haga, salvo quizá zonas agrestes de difícil conducción en las que la incomodidad del viaje hace poco deseable la realización de una llamada. No obstante, sería interesante recalibrar dicha

probabilidad cada cierto tiempo para atender posibles modificaciones en la red viaria, como creación de nuevas vías cruzando una frontera, o cambios en la conducta de los usuarios al realizar llamadas.



**Figura 4-12: Probabilidad de hacer una llamada a bordo de vehículo.**

Se aprecia que esta probabilidad de hacer una llamada a bordo de un vehículo por parte de usuarios en movilidad sigue en gran medida los patrones de la actividad telefónica diaria reflejados en la Figura 4-3 (cualquier usuario). El pico de mañana se centra en torno al mismo rango horario en ambos casos; mientras que el pico de la tarde está más suavizado en el caso de llamadas realizadas por usuarios en movilidad. Esta característica puede tener su origen en que, durante la jornada laboral, el hecho de estar en movimiento no impide que los usuarios realicen llamadas cuando así lo requieran. Mientras que, fuera de ese horario laboral, el usuario es más reacio a llamar mientras esté en movimiento. De ahí la menor tendencia a realizar llamadas fuera de la jornada laboral en el caso de usuarios en movimiento.

#### ■ Probabilidad de handover $P_h(t_i, k)$

Se considera a continuación un escenario de handover simplificado (Sallent *et al.*, 2003), suponiendo un sistema celular regular con celdas por las cuales discurren una serie de vías. Se modela la distancia que recorre un móvil dentro de una celda antes de cruzar la frontera de esa celda como una variable aleatoria con distribución uniforme en el intervalo  $[0, L]$  metros. El tiempo de permanencia en la celda,  $t_p$ , si el móvil se desplaza a velocidad uniforme de  $V$  m/s será también una variable aleatoria con distribución uniforme en el intervalo  $t_p \in [0, L/V]$  segundos, de manera que su función de densidad de probabilidad vendrá dada por:

$$f(t_p) = \frac{1}{L/V} \quad t_p \in \left[0, \frac{L}{V}\right] \quad (13)$$

Si la duración de la llamada,  $t_d$ , puede modelarse como una variable aleatoria exponencial de media  $T_c$  segundos, su correspondiente función de densidad de probabilidad se expresa como:

$$f(t_d) = \frac{1}{T_c} e^{-\frac{t_d}{T_c}} \quad t_d \in [0, +\infty) \quad (14)$$

La probabilidad de que dicho usuario deba efectuar un handover a lo largo de la llamada viene dada por el hecho de que la duración de la misma supere el tiempo de permanencia en la celda. Así, la probabilidad de handover puede calcularse a partir de un ejercicio de probabilidad condicionada:

$$P_h = \int_0^{\frac{L}{V}} P(t_d > t_p | t_p) f(t_p) dt_p \quad (15)$$

A su vez, la probabilidad condicionada de que la duración de la llamada sea superior al tiempo de permanencia en la celda vale:

$$P(t_d > t_p | t_p) = \int_{t_p}^{\infty} f(t_d) dt_d = e^{-\frac{t_p}{T_c}} \quad (16)$$

Definiendo el factor de movilidad como  $\alpha = \frac{L}{V \cdot T_c}$ , finalmente se tiene que:

$$P_h = \int_0^{\frac{L}{V}} e^{-\frac{t_p}{T_c}} f(t_p) dt_p = \int_0^{\frac{L}{V}} e^{-\frac{t_p}{T_c}} \frac{1}{L/V} dt_p = \frac{1}{\alpha} [1 - e^{-\alpha}] \quad \text{con} \quad \alpha = \frac{L}{V \cdot T_c} \quad (17)$$

La Figura 4-13 representa la probabilidad de handover en función del factor de movilidad  $\alpha$ . De la lectura de la ecuación (17) se puede concluir afirmando que será más fácil que la llamada deba ejecutar un handover:

1. Cuanto más pequeñas sean las distancias a recorrer (menor  $L$ ).
2. Cuanto mayor sea la velocidad del móvil (mayor  $V$ ).
3. Cuanto más larga sea la duración de las llamadas (mayor  $T_c$ ).

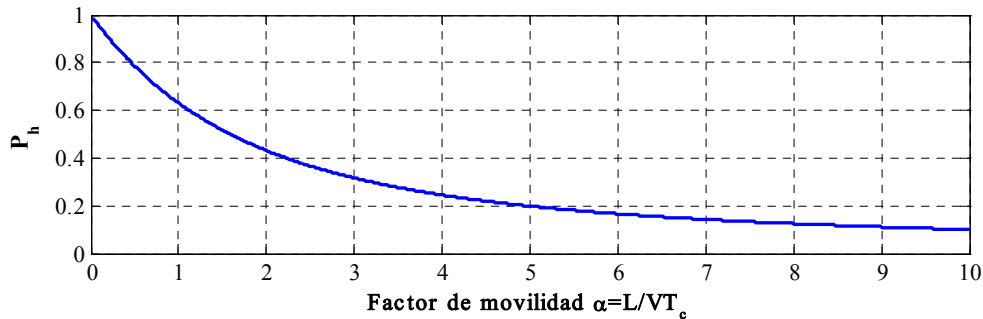


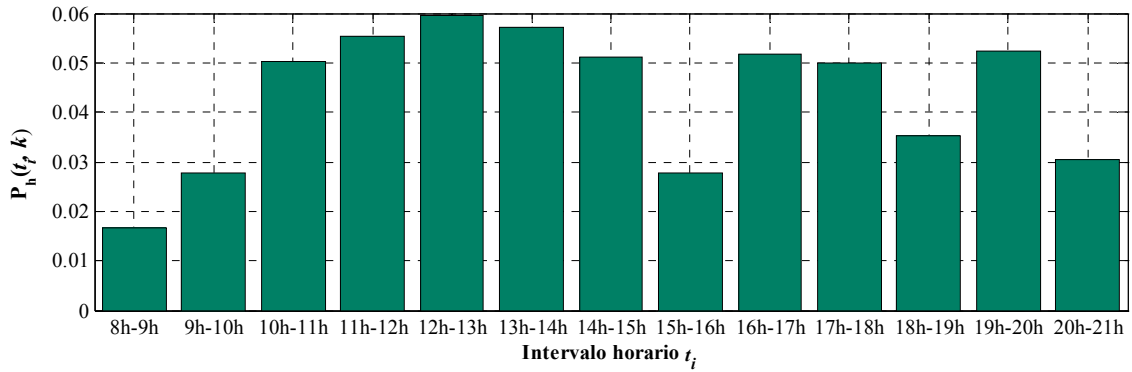
Figura 4-13: Probabilidad de handover  $P_h$  según el el factor de movilidad  $\alpha$ .

Cabe destacar que  $\alpha$  depende de 3 parámetros: longitud  $L$ , velocidad  $V$  y duración media de la llamada  $T_c$ . Respecto al parámetro  $L$ , al tratarse de la distancia que debe recorrer un móvil

dentro de una celda hasta entrar en otra celda (cruzar la frontera), esta longitud depende de las vías existentes dentro de la celda origen de la frontera. Algo similar ocurre con la velocidad,  $V$ , donde su valor depende del tipo de vía que tenga asociada (autovía, carretera nacional,...). Aunque la velocidad podría oscilar para la misma vía según el intervalo horario debido, por ejemplo, al nivel de saturación que experimente, se considerará como uniforme dentro de la celda origen de la frontera. Por ello, el valor de  $\alpha$  dependerá muy estrechamente de la frontera  $k$  observada en términos de longitud y velocidad asociada a la vía que discurre por la celda origen de esa frontera. De igual modo, teniendo en cuenta la dependencia temporal de la duración de las llamadas (apartado 4.3.3),  $\alpha$  también dependerá del intervalo horario analizado. Así pues, el valor de la probabilidad de handover será función del intervalo horario y de la frontera,  $P_h(t_i, k)$ .

Para el cálculo de la longitud  $L$  se podría realizar un estudio previo que utilizara información de cobertura proporcionada por el operador, permitiendo definir, con ayuda de un GIS, la longitud del tramo que debe recorrer un móvil dentro de una celda hasta entrar en otra. No obstante, existen funcionalidades básicas en la red que permiten conocer el identificador de la celda (antena) a la que está conectado un teléfono. De manera que, con ayuda de un GPS, es posible establecer una correspondencia entre punto kilométrico de una vía y celda que le da servicio. En base a esto, se puede conocer no sólo la longitud del tramo que recorre un móvil dentro de una celda, sino la velocidad media de viaje por dicho tramo en cualquier momento del día, todo ello sin necesitar que el operador revele información confidencial de cobertura. Con esta sencilla aplicación, implementada sobre dispositivos móviles (BlackBerry, Windows Mobile,...), se recorrerían varias veces los tramos de vía mediante la técnica de coche flotante, a fin de poder caracterizar la celda origen de cada frontera  $k$  en términos de longitud  $L$  y velocidad  $V$ .

Por último, faltaría estimar la duración media de la llamada  $T_c$  para determinar completamente el valor de  $\alpha$  y, por consiguiente, de la probabilidad de handover. Esta duración es posible obtenerla empíricamente mediante el análisis de los datos de llamadas proporcionados por el operador. Considerando la dependencia temporal puesta de manifiesto en el apartado 4.3.3, se determina un valor de duración media de llamadas según el intervalo horario considerado,  $T_c(t_i)$ . Esa duración dependiente del intervalo  $t_i$ , junto con los valores de velocidad y longitud correspondientes al tramo de vía que discurre por la celda origen de la frontera  $k$ , permiten definir la probabilidad de handover como una función variable con el tiempo y la frontera,  $P_h(t_i, k)$ . La Figura 4-14 muestra la probabilidad de handover alcanzada en la celda origen de una frontera, centrándose en el intervalo 8:00–21:00 según lo expuesto en el apartado 4.3.4.



**Figura 4-14: Evolución de la probabilidad de handover en la celda origen de la frontera  $k$  según el intervalo horario  $t_i$ .**

#### ■ Definición del modelo

Muchos procedimientos estadísticos suponen que los datos siguen algún tipo de modelo matemático que se define mediante una ecuación, en la que se desconoce alguno de sus parámetros, siendo éstos calculados o estimados a partir de la información obtenida en un estudio diseñado para tal fin. En este caso, se va a definir un modelo a partir del fenómeno físico de llamadas en movilidad empleando la expresión teórica del volumen de vehículos que pasan por una frontera  $k$  en un intervalo horario  $t_i$  definida en (11). Tras sustituir las expresiones anteriores de  $P_{ll\ veh}(t_i)$  y  $P_h(t_i, k)$ , esa ecuación quedaría expresada como:

$$n_{veh}(t_i, k) = \frac{n_{ll\ mov}(t_i, k)}{P_{ll\ veh}^2(t_i) + P_{ll\ veh}(t_i) \cdot \frac{1}{\alpha(t_i, k)} [1 - e^{-\alpha(t_i, k)}]}, \quad \text{siendo } \alpha(t_i, k) = \frac{V_k}{L_k \cdot T_c(t_i)} \quad (18)$$

Como ya se comentó, la expresión (18) es obtenida asumiendo una serie de hipótesis y considerando despreciables sucesos tales como que se llame en cada celda efectuando adicionalmente el proceso de handover, o, incluso, obviando ciertas situaciones en movilidad no detectadas, como llamadas que terminan antes de que el vehículo abandone la celda (no handover) y, al no realizarse una llamada en la nueva celda, no es detectado. Estos sucesos inducen pérdidas deben ser consideradas en la definición del modelo. En este sentido, se introducen una serie de parámetros en la expresión (18) para corregir las indeterminaciones introducidas con las hipótesis necesarias para generar el modelo, quedando formulado como:

$$\hat{y}(t_i, k) = \frac{a \cdot x(t_i, k)}{P_{ll\ veh}^2(t_i) + P_{ll\ veh}(t_i) \cdot \frac{b_1}{\alpha(t_i, k)} [1 - e^{-b_2 \cdot \alpha(t_i, k)}] + c} + d, \quad \text{con } \alpha(t_i, k) = \frac{V_k}{L_k \cdot T_c(t_i)} \quad (19)$$

En ese modelo, las variables  $x$  e  $y$  se corresponden, respectivamente, con el número de llamadas en movilidad realizadas entre dos celdas,  $n_{ll\ mov}(t_i, k)$ , y el número de vehículos que traspasan



una frontera entre esas celdas,  $n_{veh}(t_i, k)$ . La naturaleza de los términos y variables implicadas en el modelo condiciona que la variable de salida sea función del tiempo y el espacio.

Para estimar los parámetros  $\{a, b_1, b_2, c, d\}$  del modelo existen diferentes procedimientos de ajuste, los cuales utilizarían la parte del histórico que relaciona ambas variables denominada conjunto de entrenamiento para fijar los parámetros. En este caso, se va a emplear una técnica de ajuste de parámetros tratando de satisfacer algún criterio de ajuste. El criterio empleado atiende al error relativo, a la vista de que no posee la misma importancia que una estimación exceda en 50 sobre un valor real de 100 que sobre uno de 1000, siendo más grave la primera situación que la última. Así mismo, dado que en este caso tiene la misma importancia errar por exceso o por defecto, el valor absoluto también debe ser considerado en el criterio. Así pues, el ajuste buscará la minimización de la suma del valor absoluto del error relativo entre valores observados y modelados, resolviendo el siguiente problema de optimización:

$$\text{Minimizar}_{a, b_1, b_2, c, d} \sum_{j=1}^N \left| \frac{y_j - \hat{y}_j}{y_j} \right|, \text{ siendo } y_j \text{ los valores observados e } \hat{y}_j \text{ los valores modelados}$$

La resolución de este problema se ha implementado sobre Matlab, resultando los siguientes valores de los parámetros:  $\{a, b_1, b_2, c, d\} \equiv \{2.96417, -43.40174, -0.167134, -0.01205, 707.4249\}$ .

#### 4.3.5.3 Modelos adicionales

El objetivo final de esta investigación es correlacionar el volumen de vehículos ( $Y$ ) en una determinada sección del viario con el número de llamadas en movilidad ( $X$ ) generadas en la frontera asociada entre celdas. El anterior modelo estimaba unos valores para ese volumen de vehículos a partir del número de llamadas en movilidad observado pero, adicionalmente, introducía una dependencia temporal y espacial que hacía que las predicciones variaran según el periodo horario y la frontera en cuestión. Este hecho implica realizar un estudio previo para caracterizar la frontera en términos de longitud y velocidad de las vías existentes en la celda origen correspondiente. En general, lo ideal sería emplear modelos que sólo tuvieran dependencia con una variable, en este caso, llamadas en movilidad, pero las conclusiones extraídas de las cópulas revelan la imposibilidad de emplear ese tipo de modelos. Eliminar la dependencia espacial evita realizar el estudio para conocer las características geométricas de cada frontera monitorizada; en cambio, a la vista del análisis descriptivo de datos de llamadas, mantener en la formulación de los modelos la dependencia temporal parece lo más apropiado.

La dependencia asociada al periodo horario en cuestión es importante puesto que las dos variables implicadas en el proceso, llamadas realizadas y vehículos circulando, varían en

función del tiempo de una forma diferenciada. Ya se comentó anteriormente que los usuarios poseen diferentes hábitos en la realización de llamadas según la hora considerada debido a criterios como tarifas u horario admisible de realización de llamada. Mientras que, en relación con el flujo de vehículos, éste fluctúa con el tiempo atendiendo a criterios diferentes, como hora punta por inicio de jornada laboral, hora valle, etc.

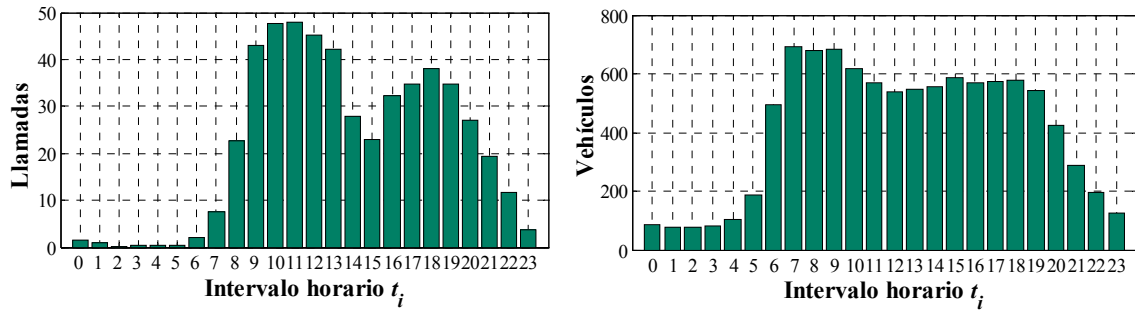


Figura 4-15: Evolución del nº de llamadas y de vehículos en una frontera según el intervalo horario.

La Figura 4-15 muestra la variación temporal de la intensidad del número de llamadas realizadas y del número de vehículos que pasan por una frontera determinada en un día concreto. Se observa que en ambos casos los periodos horarios de mayor intensidad están próximos en el tiempo pero con cierto desfase; mientras que el pico de vehículos aparece en torno las 7:00 (inicio jornada laboral), a esa hora el número de llamadas es bajo ya que es poco usual realizar llamadas en ese intervalo horario. Este comportamiento es extensible a otras fronteras. En este sentido, las predicciones deben incluir algún término asociado a esa variabilidad horaria tanto de la intensidad de llamadas como de la intensidad de vehículos. Los modelos que a continuación se presentan introducen unos factores asociados al tiempo, que modelan la dependencia temporal de esas intensidades. El cálculo de esos factores de intensidad se realizará de modo que sean independientes de las características de la frontera, eliminando la dependencia espacial y dando lugar a modelos más versátiles.

#### ■ Factor de intensidad de vehículos

El factor de intensidad,  $f_v$ , asociado a la variación horaria del flujo de vehículos, se define como:

$$f_v(t_i) = \frac{\bar{n}_{veh}(t_i)}{\bar{n}_{veh}}, \text{ con } \left\{ \begin{array}{l} \bar{n}_{veh}(t_i) \text{ es el nº medio de vehículos que pasan en el intervalo } t_i \text{ de cada uno de los} \\ D \text{ días observados por las } K \text{ fronteras, tal que } \bar{n}_{veh}(t_i) = \frac{1}{D \cdot K} \cdot \sum_{d=1}^D \sum_{k=1}^K n_{veh}(t_i, d, k) \\ \bar{n}_{veh} \text{ es el nº medio de vehículos que han pasado desde las 8h hasta las 21h,} \\ \text{tal que } \bar{n}_{veh} = \frac{1}{H} \cdot \sum_{i=8h}^{21h} \bar{n}_{veh}(t_i), \text{ siendo } H \text{ el nº de intervalos horarios entre 8h-21h} \end{array} \right.$$

El factor de intensidad es, por tanto, la relación entre el número medio de vehículos contados en un periodo horario  $t_i$  y el número medio del total de vehículos contados desde las 8:00 hasta las 21:00. Así pues, este factor se encarga de regular las observaciones en cada periodo horario según su importancia relativa dentro de los periodos horarios considerados. Por ejemplo, en intervalos horarios donde el flujo de tráfico tiende a ser mayor debido a hora punta, el factor de intensidad  $f_v$  hará que las observaciones destaquen sobre las de otros intervalos. La Figura 4-16 muestra la variación con el tiempo del factor de intensidad  $f_v(t_i)$ .

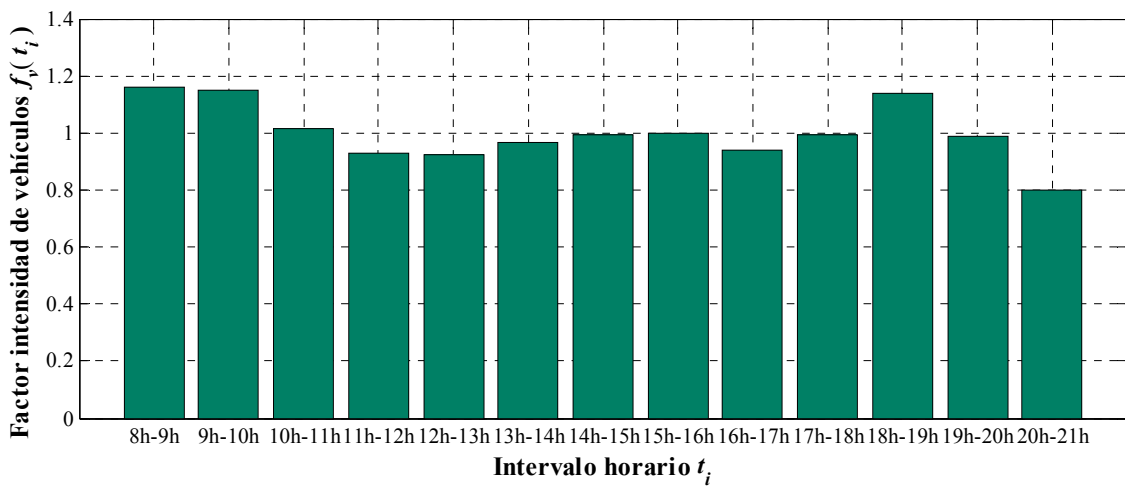


Figura 4-16: Factor de intensidad del paso de vehículos según el intervalo horario  $t_i$ .

■ *Factor de intensidad de llamadas*

Análogamente, el factor de intensidad de llamadas,  $g_{ll}$ , se encarga de reflejar la variación temporal del número de llamadas realizadas. Este factor  $g_{ll}$  se define como la relación entre la probabilidad de que un vehículo realice una llamada con el operador de interés en el intervalo horario  $t_i$ , definida en (12), y la probabilidad media en todo el rango horario estudiado, atendiendo a la siguiente expresión:

$$g_{ll}(t_i) = \frac{P_{ll\ veh}(t_i)}{\bar{P}_{ll\ veh}}, \text{ donde } \begin{cases} P_{ll\ veh}(t_i), \text{ es la probabilidad de hacer una llamada con el operador de interés} \\ \text{a bordo de vehículo en el intervalo horario } t_i, \text{ definida en (12)} \\ \bar{P}_{ll\ veh} = \frac{1}{H} \cdot \sum_{i=8h}^{21h} P_{ll\ veh}(t_i), \text{ siendo } H \text{ el n}^\circ \text{ de intervalos horarios entre 8h-21h} \end{cases}$$

Con ello, se busca regular las observaciones de llamadas realizadas en función de su importancia dentro de los intervalos horarios considerados. La Figura 4-17 refleja la variación con el tiempo del factor de intensidad  $g_{ll}(t_i)$ , mostrando que en periodos horarios donde es más probable hacer llamadas, este factor de intensidad realzará las observaciones.

Cabe destacar que ambos factores de intensidad  $f_v(t_i)$  y  $g_{ll}(t_i)$ , obtenidos empíricamente con datos observados, se pueden considerar válidos durante un cierto tiempo (p.ej: 5 meses), debido a que las características de la muestra son más o menos estables. No obstante, sería aconsejable realizar un calibrado de ambos factores periódicamente, a fin de actualizarlos con posibles cambios de tendencia asociados a los usuarios de la vía.

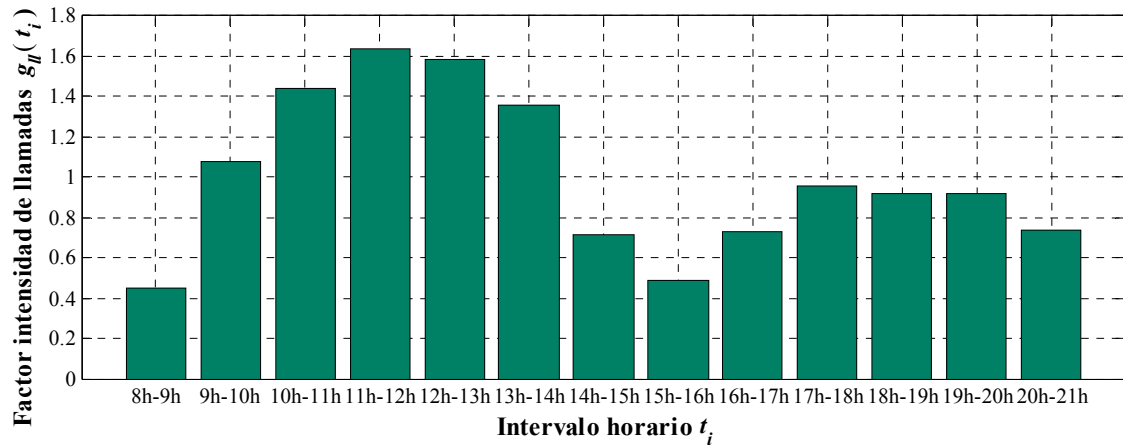


Figura 4-17: Factor de intensidad del paso de vehículos según el intervalo horario  $t_i$ .

#### ■ Definición del modelo

Los anteriores factores de intensidad  $f_v(t_i)$  y  $g_{ll}(t_i)$  constituyen la base principal en la definición de los nuevos modelos para proporcionar una predicción del número de vehículos que atraviesan una frontera  $k$  en función del número de llamadas en movilidad observadas  $n_{ll\ mov}(t_i, k)$ . Es importante resaltar que, a pesar de que esas llamadas en movilidad observadas están asociadas a un punto kilométrico o frontera  $k$ , la forma funcional de los modelos no tiene en cuenta esta dependencia y tratará las observaciones de cada frontera del mismo modo.

Así pues, se proponen varios modelos, con diferente número de parámetros, para encontrar la forma funcional que mejor se ajuste al fenómeno. En todos ellos, la variable dependiente es  $y(t_i)$ , el número de vehículos que cruzan una frontera en un intervalo horario  $t_i$ .

- Modelo 1:  $\hat{y}_1(t_i) = a \cdot f_v^\alpha(t_i) \cdot g_{ll}^\beta(t_i)$
- Modelo 2:  $\hat{y}_2(t_i) = (a + b \cdot x(t_i)) \cdot f_v^\alpha(t_i) \cdot g_{ll}^\beta(t_i)$ , donde  $x(t_i) = n_{ll\ mov}(t_i, k)$
- Modelo 3:  $\hat{y}_3(t_i) = (a + b \cdot x(t_i) + c \cdot x^2(t_i)) \cdot f_v^\alpha(t_i) \cdot g_{ll}^\beta(t_i)$ , donde  $x(t_i) = n_{ll\ mov}(t_i, k)$

El modelo 1, con 3 parámetros  $\{a, \alpha, \beta\}$ , carece de información sobre el número de llamadas en movilidad observadas. Se ha propuesto para estudiar la importancia del uso de las llamadas en movilidad en la predicción del volumen de paso de vehículos. El modelo 2, con 4 parámetros

$\{a, b, \alpha, \beta\}$ , establece una dependencia de primer orden con el número de llamadas en movilidad observadas. Por último, el modelo 3, con 5 parámetros  $\{a, b, c, \alpha, \beta\}$ , es similar al modelo 2 pero estableciendo una dependencia de segundo orden con el número de llamadas en movilidad observadas.

Así mismo, para estudiar la importancia del uso de los factores de intensidad  $f_v(t_i)$  y  $g_{ll}(t_i)$  en la formulación del modelo, se plantean dos modelos adicionales en los cuales  $\alpha$  y  $\beta$  son fijados a cero, definiendo dos sencillos modelos que únicamente emplean las llamadas en movilidad observadas. Éstos son los siguientes:

- Modelo 4:  $\hat{y}_4(t_i) = a + b \cdot x(t_i)$ , donde  $x(t_i) = n_{ll\ mov}(t_i, k)$ , 2 parámetros  $\{a, b\}$ .
- Modelo 5:  $\hat{y}_5(t_i) = a + b \cdot x(t_i) + c \cdot x^2(t_i)$ , donde  $x(t_i) = n_{ll\ mov}(t_i, k)$ , 3 parámetros  $\{a, b, c\}$

Para la estimación de los parámetros de cada uno de estos cinco modelos, se emplea el mismo procedimiento presentado en el apartado 4.3.5.2 donde, utilizando el conjunto de entrenamiento definido a partir del histórico de datos observados, los parámetros son ajustados buscando la minimización de la suma del valor absoluto del error relativo entre valores observados y modelados:

$$\underset{\text{parámetros}}{\text{Minimizar}} \sum_{j=1}^N \left| \frac{y_j - \hat{y}_j}{y_j} \right|, \text{ siendo } y_j \text{ los valores observados e } \hat{y}_j \text{ los valores modelados}$$

La resolución de este problema para cada uno de los modelos, implementada sobre Matlab, ha permitido inferir los parámetros que caracterizan los modelos de predicción. Estos cinco modelos, junto al definido en (19), enumerado como modelo 6  $-\hat{y}_6 = f(x, t_i, k)-$ , serán evaluados y comparados entre sí para seleccionar el que proporcione mejores predicciones del número de vehículos que atraviesan una frontera asociado a llamadas en movilidad en ella.

## 4.3.6 Contraste de modelos

### 4.3.6.1 Introducción

Para abordar la evaluación y comparación entre los seis modelos anteriormente planteados se ha utilizado el conjunto de validación, formado por una parte del histórico de valores observados reservada para este fin. Dada la variedad de los modelos, esta comparación se ha realizado en base a criterios tales como las medidas de error clásicas (error medio, error absoluto y relativo, etc.) o la evaluación de la distribución acumulada del error relativo absoluto en términos de percentiles. Incluso, se han empleado otros criterios más minuciosos, como la correlación por

rango (o de Spearman) y el coeficiente de correlación lineal (o de Pearson), para comparar, uno a uno, los valores  $\hat{y}_i$  estimados por cada modelo frente a los observados  $y_i$ . Adicionalmente, se han estudiado otras medidas de contraste de modelos en las que también se tiene en cuenta la eficiencia y simplicidad de los modelos en términos del número de parámetros. Las más extendidas son el criterio de información de Akaike, *AIC*, y criterio de información bayesiano, *BIC*. Por último, lógicamente, se han comparado los modelos gráficamente, representando en la misma figura los valores observados y las predicciones hechas por cada modelo en cada una de las fronteras. Los modelos más adecuados han sido los que presentan mejor balance entre todos los criterios considerados.

#### 4.3.6.2 Medidas de error

Cuando se utilizan modelos de predicción, es inevitable que el resultado esté afectado por errores debidos a los propios modelos, puesto que éstos siempre llevan implícita alguna indeterminación. Dado que se dispone de un histórico de valores observados  $y_i$  con los que comparar las predicciones  $\hat{y}_i$  hechas por cada uno de los modelos, es posible efectuar un análisis directo de las medidas de error. Los estadísticos más comunes a la hora de cuantificar el error son el error medio, el error absoluto medio y el error relativo medio. A la vista de que presenta la misma gravedad errar en una predicción tanto en exceso como por defecto, el error relativo se ha estudiado en valor absoluto. La Tabla 4-4 muestra esas medidas de error para cada uno de los modelos empleando los valores observados y las predicciones. También se han incluido otras medidas relacionadas con los valores máximos del error absoluto y del error relativo en valor absoluto que se han encontrado entre los pares de datos.

**Tabla 4-4: Medidas de error para cada modelo.**

	<b>Modelo 1</b>	<b>Modelo 2</b>	<b>Modelo 3</b>	<b>Modelo 4</b>	<b>Modelo 5</b>	<b>Modelo 6</b>
<b><i>ME</i></b>	118.5430	103.8163	106.6021	112.8573	113.5008	95.1117
<b><i>MAE</i></b>	237.2021	210.6609	210.0152	223.9775	223.2356	203.5965
<b><i>MARE</i></b>	0.2318	0.2039	0.2032	0.2169	0.2157	0.2000
<b><i>Error absoluto máximo</i></b>	1488.0287	1399.7732	1406.7949	1562.8313	1552.9211	1449.5156
<b><i>Error relativo absoluto máximo</i></b>	0.9761	0.8302	0.8529	1.0103	1.0723	0.9985

(ME: error medio; MAE: Error absoluto medio; MARE: Error relativo medio en valor absoluto)

En base a estas medidas, los mejores modelos son el 6, 2 y 3. Conviene destacar que no es posible establecer un orden claro de prioridad entre esos tres modelos dado que sus medidas de

error son muy similares entre ellas. Las Figura 4-18 y Figura 4-19 muestran, respectivamente, el error relativo medio en valor absoluto y el error absoluto medio para cada modelo atendiendo al periodo horario, apreciándose que los modelos 2, 3 y 6 son los que presentan mejor comportamiento.

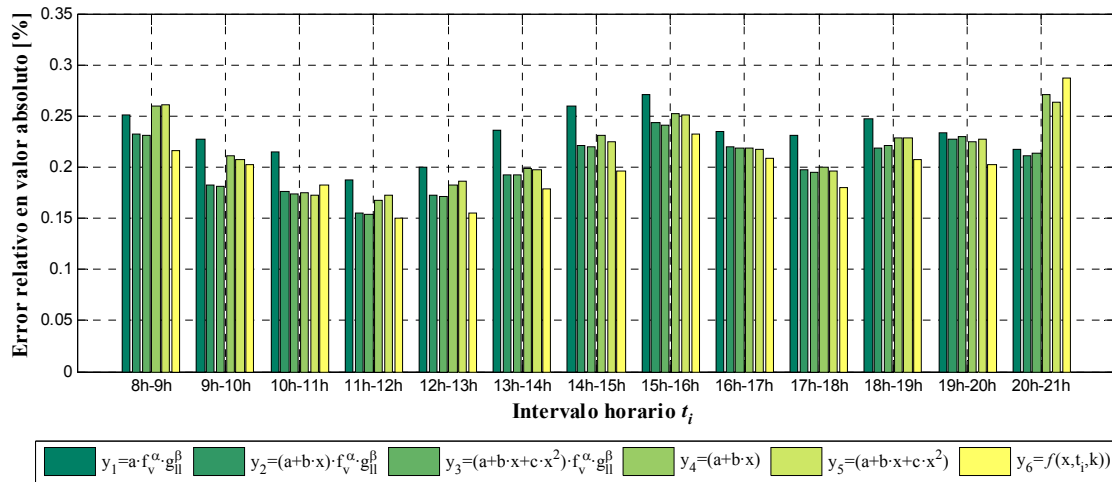


Figura 4-18: Evolución horaria del error relativo en valor absoluto para cada uno de los modelos.

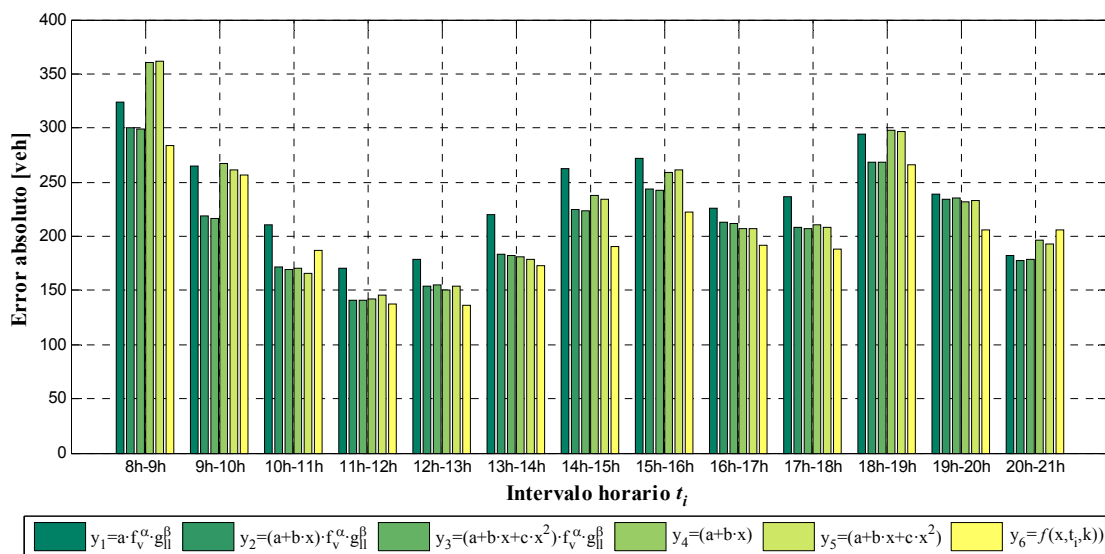


Figura 4-19: Evolución horaria del error absoluto para cada uno de los modelos.

#### 4.3.6.3 Distribución acumulada de error relativo en valor absoluto. Percentiles

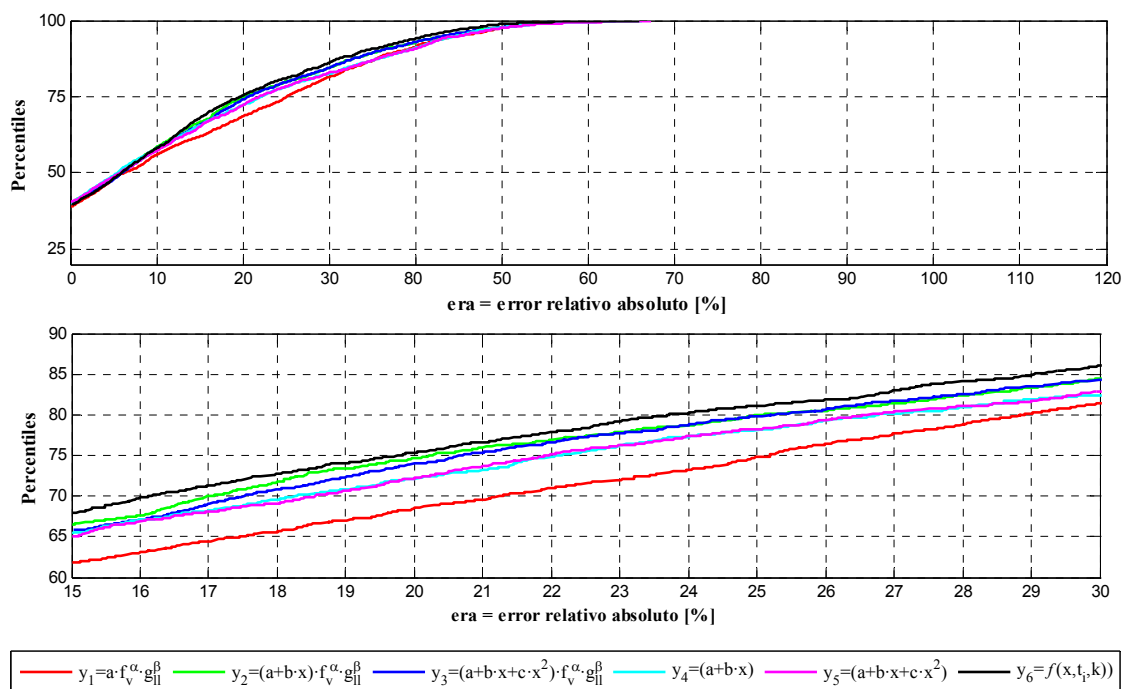
Los percentiles pertenecen al grupo de estadísticos asociados a medidas de posición. Se emplean para proporcionar información sobre el conjunto de datos que se está analizando, indicando qué valores de la variable se presentan en la realidad con una frecuencia por debajo de un

porcentaje. Analizando la distribución de frecuencias del error relativo en valor absoluto para cada uno de los modelos, los percentiles alcanzados se muestran en la siguiente tabla:

**Tabla 4-5: Percentiles del error relativo en valor absoluto para cada modelo.**

Percentiles	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
25	0,0799	0,0769	0,0763	0,0777	0,0799	0,0772
30	0,0982	0,0933	0,0933	0,096	0,0967	0,0938
50	0,2039	0,1694	0,1694	0,1763	0,1783	0,1646
60	0,26	0,2141	0,2121	0,2286	0,2246	0,212
75	0,3596	0,3076	0,3034	0,3351	0,3311	0,2986
85	0,4298	0,379	0,3791	0,4117	0,4113	0,3703
100	0,9765	0,83	0,8492	1,0723	1,0103	0,9985

Los mejores modelos serán aquellos que posean valores más pequeños de error relativo en valor absoluto para cada uno de los percentiles. Numéricamente, se observa que los modelos 6, 3 y 2 destacan en estos términos respecto a los demás. En ellos, por ejemplo, el percentil 60 refleja que el 60% de las predicciones alcanzaron errores relativos en valor absoluto inferiores a 0.21. Esto también se aprecia más claramente a nivel gráfico. En este sentido, la Figura 4-20 muestra la distribución acumulada del error relativo en valor absoluto conseguido por cada modelo. El eje vertical está graduado en percentiles, por ello la media se alcanza en el percentil 50. Se observa que los mejores modelos son, por orden de preferencia, el modelo 6, el 3 y el 2, dado que la curva para el resto de modelos es visiblemente inferior a la de esos tres modelos.



**Figura 4-20: Distribución acumulada del valor absoluto del error relativo en porcentajes (arriba). Zoom en torno a los valores de error relativo absoluto 15-30% (abajo).**



#### 4.3.6.4 Correlación por rango (de Spearman) y correlación lineal (de Pearson)

Otro criterio a tener en cuenta dentro de la comparativa de modelos es la correlación entre las predicciones y los valores reales. Se considera que dos variables están correlacionadas cuando los valores de una de ellas varían sistemáticamente con respecto a los valores de la otra. En este sentido, si tenemos dos variables (A y B) existe correlación si al aumentar los valores de A lo hacen también los de B y viceversa. Aplicando esto al conjunto de valores observados  $\{y_i\}$  y estimados  $\{\hat{y}_i\}$ , se puede evaluar la calidad de las predicciones examinando si existe correlación. El coeficiente de correlación lineal, o coeficiente de Pearson, es el instrumento tradicionalmente utilizado para medir la dependencia entre variables aleatorias. Si  $r = 0$ , no existe relación lineal. Pero esto no implica necesariamente una independencia total entre las dos variables. Puede ocurrir que, aún con  $r$  próximo a 0, la variación de una de ellas influya en el valor que pueda tomar la otra caso que existan relaciones no lineales entre las dos variables.

Estas limitaciones no existen si se utiliza la correlación entre rangos, la cual considera la concordancia de los datos. El concepto de concordancia está relacionado con la probabilidad de encontrar valores altos (o bajos) de una variable asociados a valores altos (o bajos) de la otra variable. Si esto no ocurre, las variables son discordantes. De modo que la concordancia detecta relaciones no lineales que la correlación no puede detectar. En consecuencia, las medidas de correlación que trabajan con rangos no presentan las limitaciones de la correlación lineal. Dentro de estas medidas destaca el denominado coeficiente de correlación de Spearman.

Se procede a analizar la calidad de las predicciones  $\hat{y}_i$  de cada modelo analizando la dependencia en términos de correlación por rango de Spearman y correlación lineal o de Pearson. La siguiente tabla muestra estas correlaciones para cada uno de los modelos.

**Tabla 4-6: Correlación por rango (Spearman) y lineal (Pearson) entre  $\{y_i, \hat{y}_i\}$  para cada modelo.**

	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
<i>Correlación por rango</i>	0.25102	0.51398	0.51375	0.39904	0.39904	0.56623
<i>Correlación lineal</i>	0.28436	0.4698	0.48134	0.33953	0.34893	0.53333

En términos de correlación por rango, destacan claramente sobre los demás los modelos 6, 2 y 3; especialmente el modelo 6, alcanzando valores elevados para el tamaño de la muestra considerada. Respecto al coeficiente de correlación lineal, ocurre algo similar. Por ello, se establece que los mejores modelos bajo este enfoque de correlación son el modelo 6, seguido de los modelos 2 y 3.

#### 4.3.6.5 Criterio de información de Akaike y criterio de información bayesiano: AIC BIC

El modelo perfecto no existe, puesto que todos constituyen simplificaciones de la realidad a través de una serie de parámetros. Siempre son preferibles modelos con menos parámetros debido a que, además de ser más sencillos de estimar, son más estables y están sometidos a menos sesgo. Por ello existen unas medidas de contraste entre modelos que penalizan en cierto modo que éstos tengan muchos parámetros.

Las más conocidas son el criterio de información de Akaike, *AIC*, y criterio de información bayesiano, *BIC*. Ambos estadísticos se utilizan en la selección de modelos para tomar el mejor de entre un conjunto de varios admisibles. Así pues, un modelo es mejor que otro si tiene un valor *AIC* (o *BIC*) menor. Las fórmulas matemáticas del *AIC* y *BIC* se enuncian a continuación:

$$\begin{aligned} AIC &= -2 \cdot LLF + 2 \cdot P \\ BIC &= -2 \cdot LLF + P \cdot \ln(N) \end{aligned} \quad (20)$$

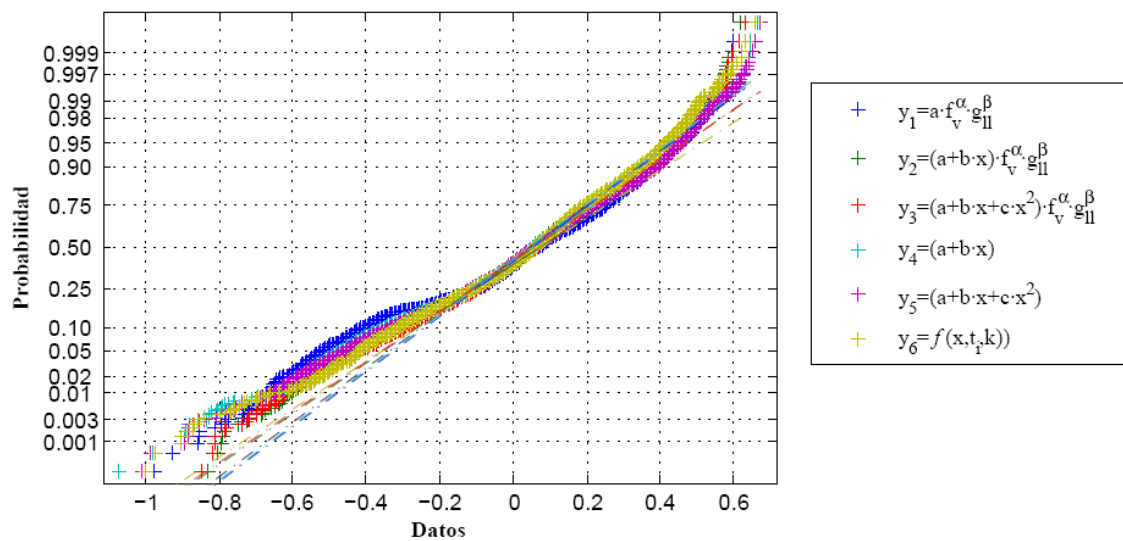
En ellas intervienen el logaritmo de la función verosimilitud de cada modelo (*LLF*), el número de parámetros ajustados (*P*) y el número total de observaciones o tamaño de la muestra (*N*). En el cálculo de la función de verosimilitud interviene un producto de probabilidades de observaciones individuales, por lo que habitualmente interesa tomar logaritmos, ya que éstos transforman los productos en sumas y los cocientes en restas. De ahí que el cálculo de *AIC* y *BIC* emplee el logaritmo de la función de verosimilitud.

La función de verosimilitud es un indicador para señalar la capacidad de ajuste de un modelo, permitiendo con ello su empleo en la comparación de modelos. Su valor será mayor cuanto mejor sea el ajuste, mientras que tendrá un valor más cercano a 0 cuando el ajuste sea malo. Por ello, junto a los estadísticos *AIC* y *BIC* de cada uno de los modelos, también se emplea el logaritmo de la función verosimilitud (*LLF*) para el contraste de modelos. El cálculo de estos tres indicadores se realiza sobre un conjunto de datos que, en este caso, se tratará de medidas del error relativo entre valores reales y predicciones,  $y_i$  e  $\hat{y}_i$  respectivamente, para cada modelo,

$$er_i = \frac{y_i - \hat{y}_i}{y_i}.$$

Para determinar el *LLF* se precisa conocer la función de verosimilitud del conjunto de medidas del error relativo, en la cual interviene la función de densidad de probabilidad seguida. La construcción de esta función se simplifica, en gran medida, si se asume que la distribución del conjunto de datos es normal. Sin embargo, realizando un test de normalidad, se demuestra que los datos analizados, en este caso errores relativos entre los valores observados y estimados por

cada uno de los modelos, no siguen una distribución normal. Esta afirmación se puede comprobar gráficamente mediante el test gráfico de normalidad mostrado en la Figura 4-21. En ella se representa la probabilidad acumulada empírica de cada conjunto de datos. Si los puntos se ajustasen muy bien a la línea recta sería señal de que se acepta la bondad de ajuste y los datos provienen de una distribución normal. Otras funciones de densidad de probabilidad introducirían curvatura en el gráfico, siendo esto lo que ocurre para las medidas consideradas de error relativo.



**Figura 4-21: Test gráfico de normalidad.**

En consecuencia, la hipótesis de distribución normal no es válida, de manera que se precisa otro procedimiento para determinar la función de verosimilitud de cada conjunto de errores relativos de los modelos. La conveniencia del uso de histogramas como estimadores no-paramétricos de la función de densidad de una variable sugiere que un modo de lograrlo consiste en discretizar la variable aleatoria continua, asociada a cada conjunto de datos, empleando un número apropiado de bins. Así pues, las funciones de densidad de probabilidad empleadas estarán basadas en histogramas cuyo número de bins será fijado por algún criterio. En particular, se han usado el criterio de Knuth y el criterio de Stone para determinar dicho número de bins de un modo óptimo (Knuth, 2006). Aplicando este procedimiento al conjunto de medidas de error relativo de cada modelo, se obtienen los siguientes histogramas (Figura 4-22 y Figura 4-23), representando la distribución estadística experimental (histograma) de los errores relativos de cada modelo junto con una curva de densidad normal superpuesta sobre cada histograma. Se observa que la discretización alcanzada con el número de bins definido por Knuth es más armoniosa que la resultante de Stone. Es decir, en el histograma generado con el número de bins definido por Stone (Figura 4-23), la altura de las barras no sigue una tendencia estable de subida o bajada, sino que se aprecia una oscilación en la altura de las barras. En cambio, en el histograma

generado por Knuth (Figura 4-22), la altura de las barras mantiene la tendencia de subida o bajada respecto del máximo. Por ello, se toma el criterio de Knuth para la discretización del error relativo y la obtención de las funciones de densidad de probabilidad que permitan calcular la verosimilitud junto a los estadísticos *AIC* y *BIC*.

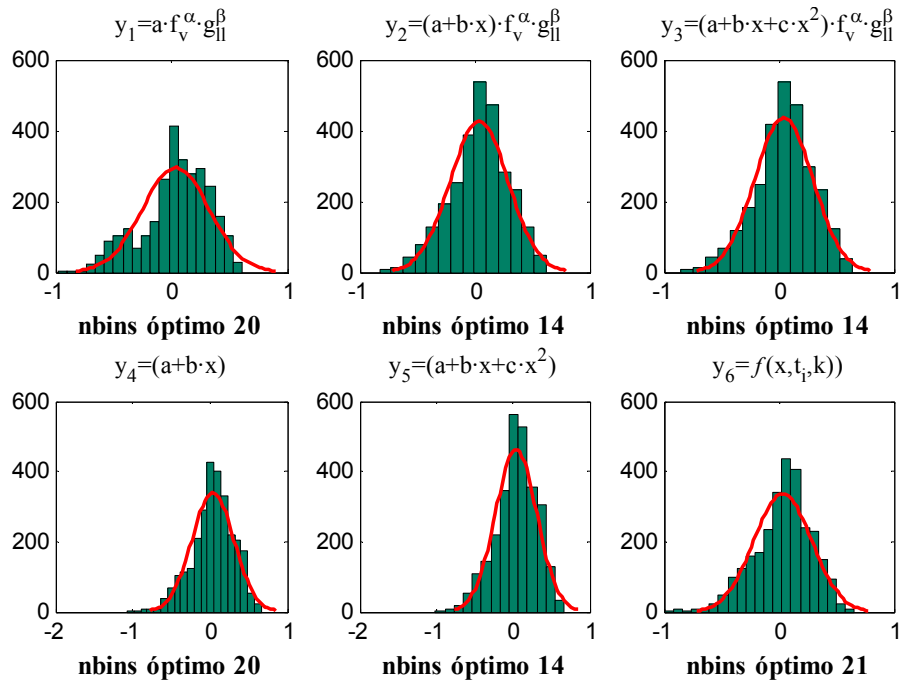


Figura 4-22: Histograma del error relativo de cada modelo según el n° de bins definido por Knuth.

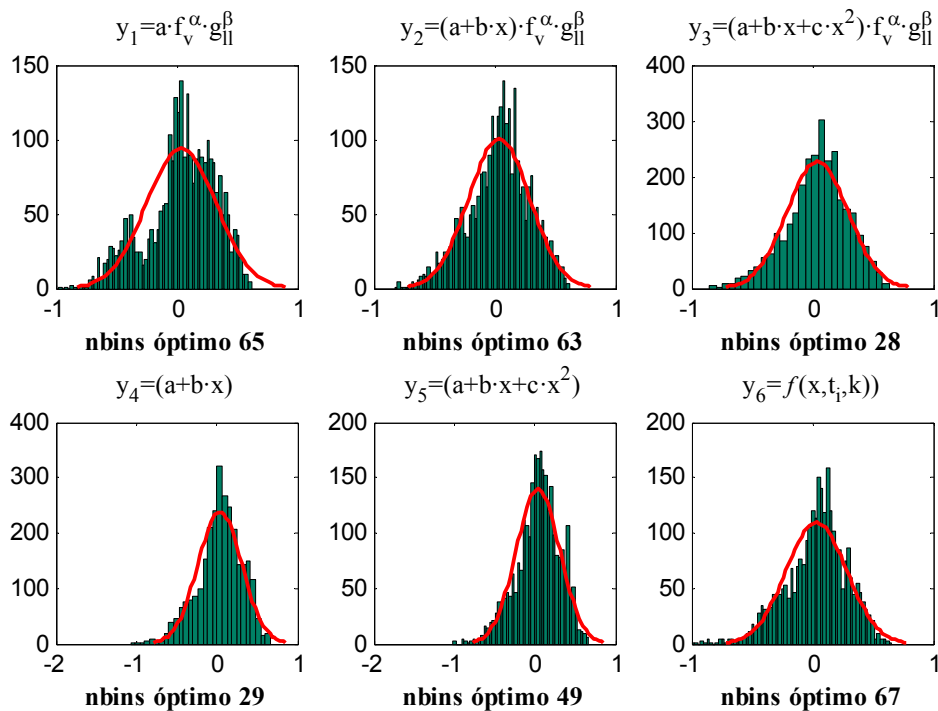


Figura 4-23: Histograma del error relativo de cada modelo según el n° de bins definido por Stone.

Así, tras realizar el proceso y obtener, posteriormente, los estadísticos *AIC* y *BIC*, se obtienen los valores mostrados en la Tabla 4-7.

**Tabla 4-7: Valores de Logaritmo de la verosimilitud (*LLF*), *AIC* y *BIC* del conjunto de errores relativos de cada modelo, junto al número de parámetros de cada modelo.**

	<i>LLF</i>	<i>AIC</i>	<i>BIC</i>	Nº param
<i>Modelo 1</i>	-7345	14697	14714	3
<i>Modelo 2</i>	-6423	12854	12877	4
<i>Modelo 3</i>	-6356	12721	12751	5
<i>Modelo 4</i>	-7082	14167	14179	2
<i>Modelo 5</i>	-6175	12357	12375	3
<i>Modelo 6</i>	-7113	14237	14267	5

Dado que los mejores modelos son los que consiguen el mayor valor de *LLF*, y menor valor de *AIC* y *BIC*, el ranking de modelos según cada criterio es:

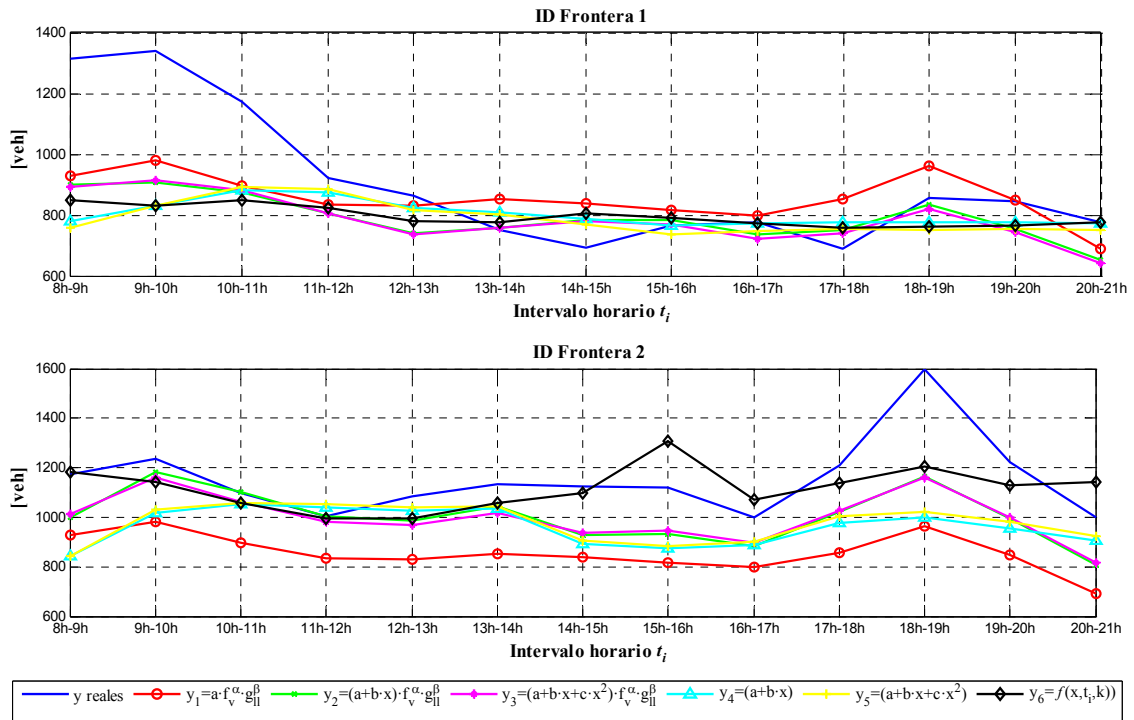
- Ranking de mejor modelo por Log-Likelihood: 5 3 2 4 6 1
- Ranking de mejor modelo por *AIC*: 5 3 2 4 6 1
- Ranking de mejor modelo por *BIC*: 5 3 2 4 6 1

Es necesario comentar que estos criterios de selección de modelo también consideran el número de parámetros existente, penalizando aquellos que tengan muchos parámetros. De ahí que, en el orden de prioridad anterior, los modelos 5 y 4 destaquen en la selección de modelos. Estos modelos 4 y 5, junto al modelo 1, son los que menos parámetros poseen. Sin embargo, atendiendo a los criterios estudiados en los anteriores apartados (medidas directas de error, percentiles y correlación por rango y lineal), parece conveniente que los modelos 1, 4 y 5 sean descartados del conjunto de modelos admisibles. En consecuencia, la selección de modelos se establecería entre los modelos 2, 3 y 6, siendo el nuevo orden de prioridad:

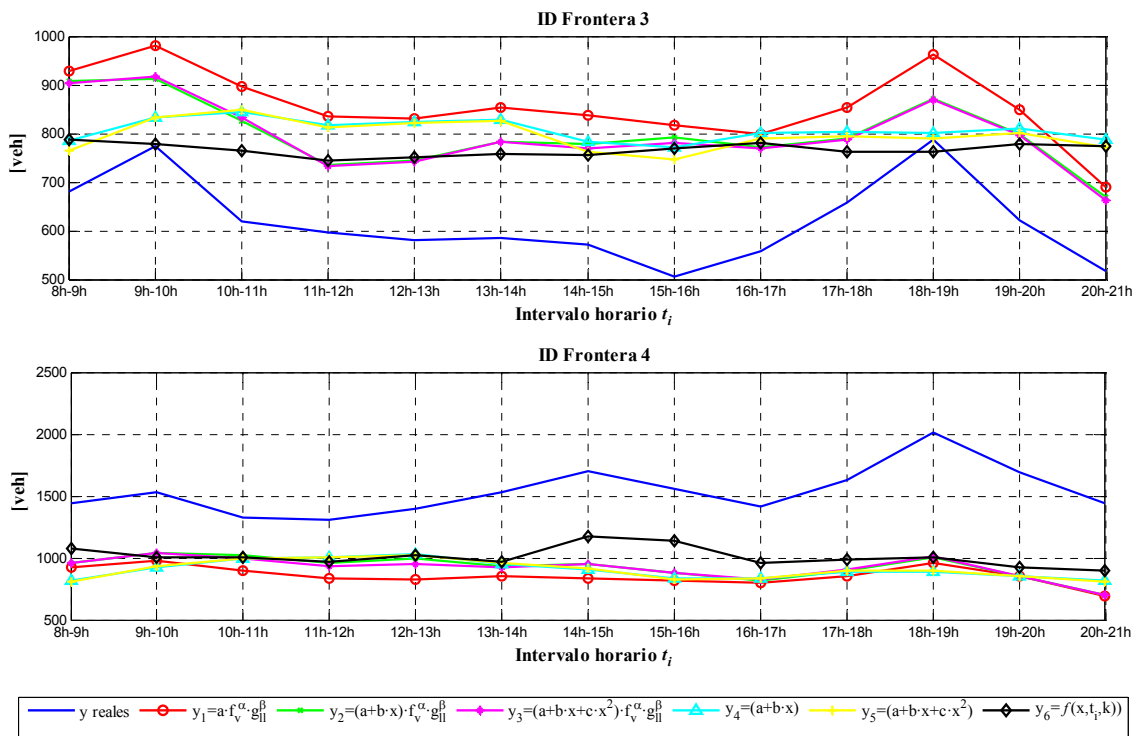
- Ranking de mejor modelo por Log-Likelihood: 3 2 6
- Ranking de mejor modelo por *AIC*: 3 2 6
- Ranking de mejor modelo por *BIC*: 3 2 6

#### 4.3.6.6 Análisis gráfico

Por último, la comparación de modelos se realizará gráficamente utilizando el conjunto de validación. Para ello, se representan las predicciones de vehículos hechas por cada uno de los modelos para cada frontera según el número medio de llamadas en movilidad observadas en cada periodo horario. Junto a ellas, también se representan los valores medios de vehículos observados en cada periodo horario y frontera, estableciendo una comparación visual de la precisión alcanzada por cada modelo así como el orden de magnitud del error de predicción.



**Figura 4-24: Volúmenes de vehículos observados y estimados con cada modelo que cruzan la frontera 1 y la frontera 2.**



**Figura 4-25: Volúmenes de vehículos observados y estimados con cada modelo que cruzan la frontera 3 y la frontera 4.**

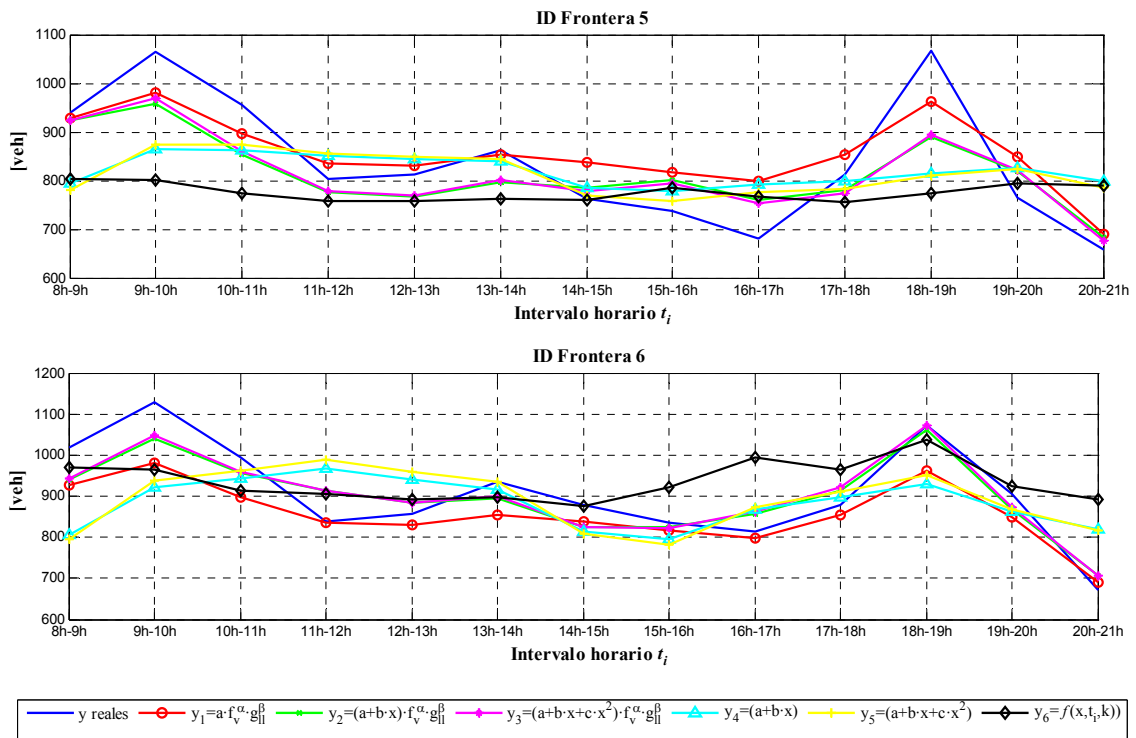


Figura 4-26: Volúmenes de vehículos observados y estimados con cada modelo que cruzan la frontera 5 y la frontera 6 en cada periodo horario.

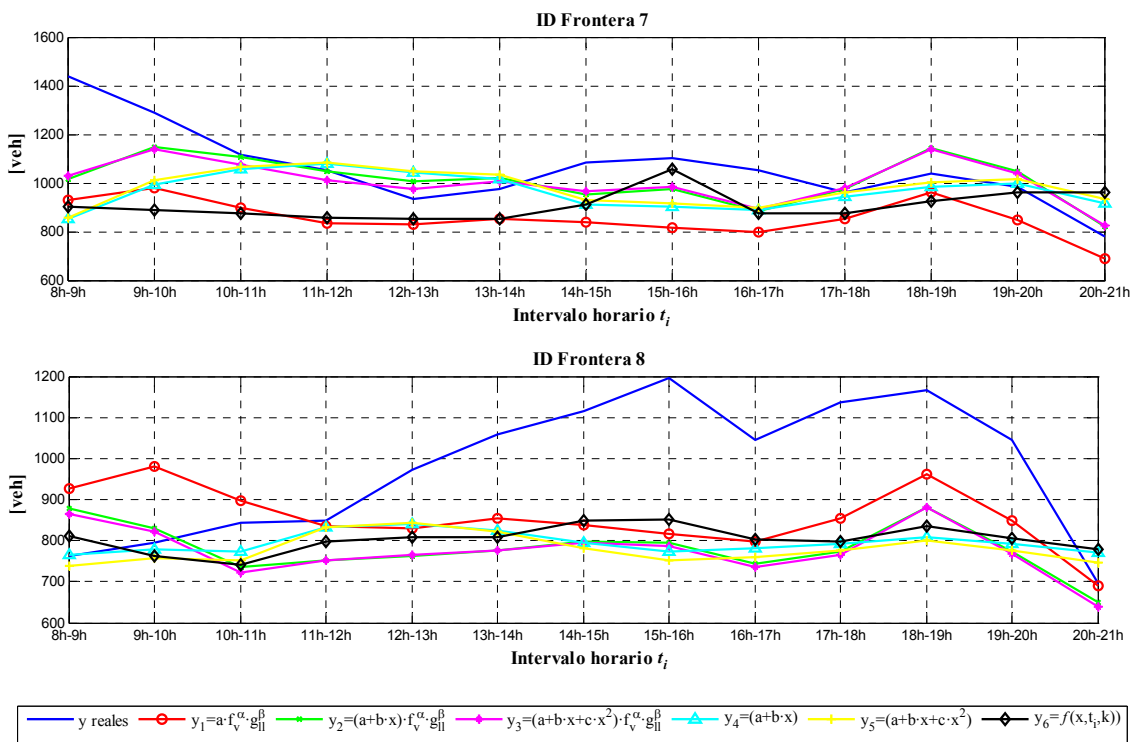
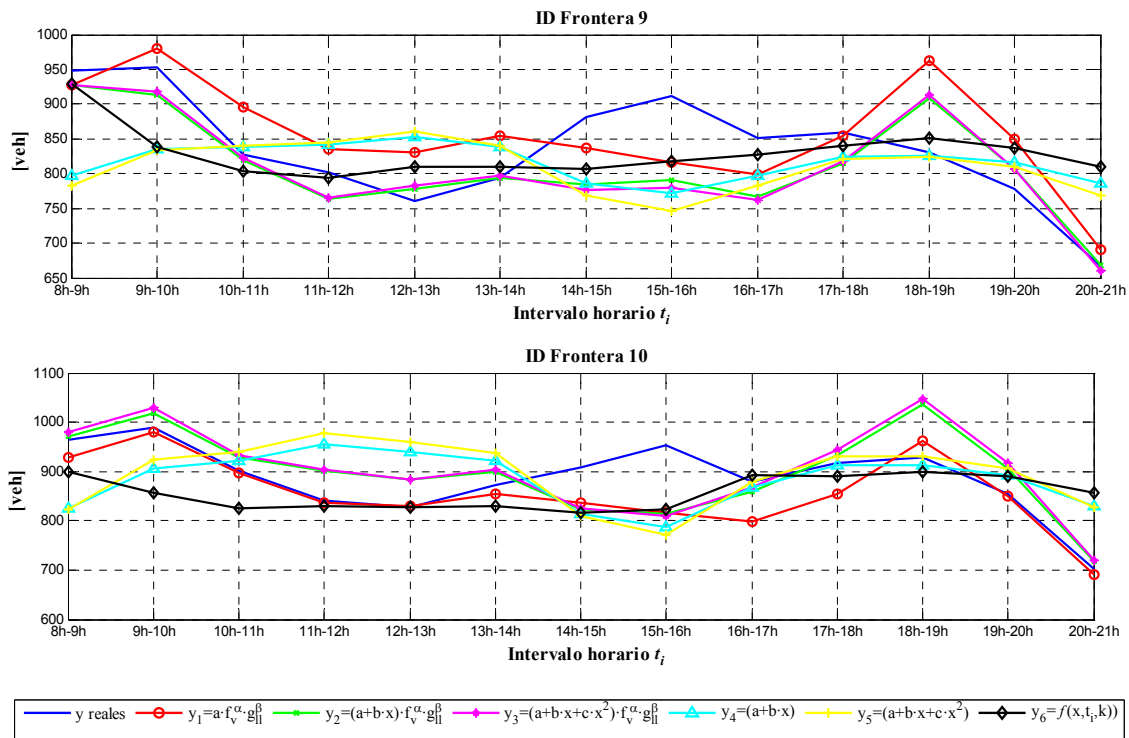
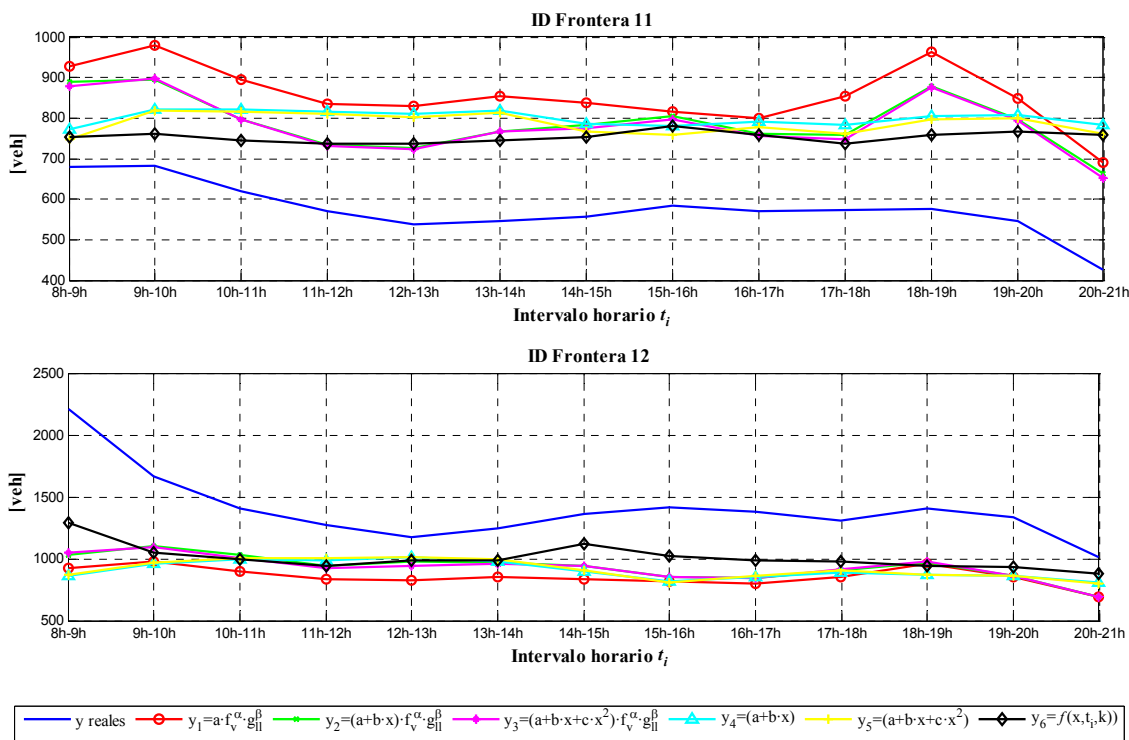


Figura 4-27: Volúmenes de vehículos observados y estimados con cada modelo que cruzan la frontera 7 y la frontera 8 en cada periodo horario.



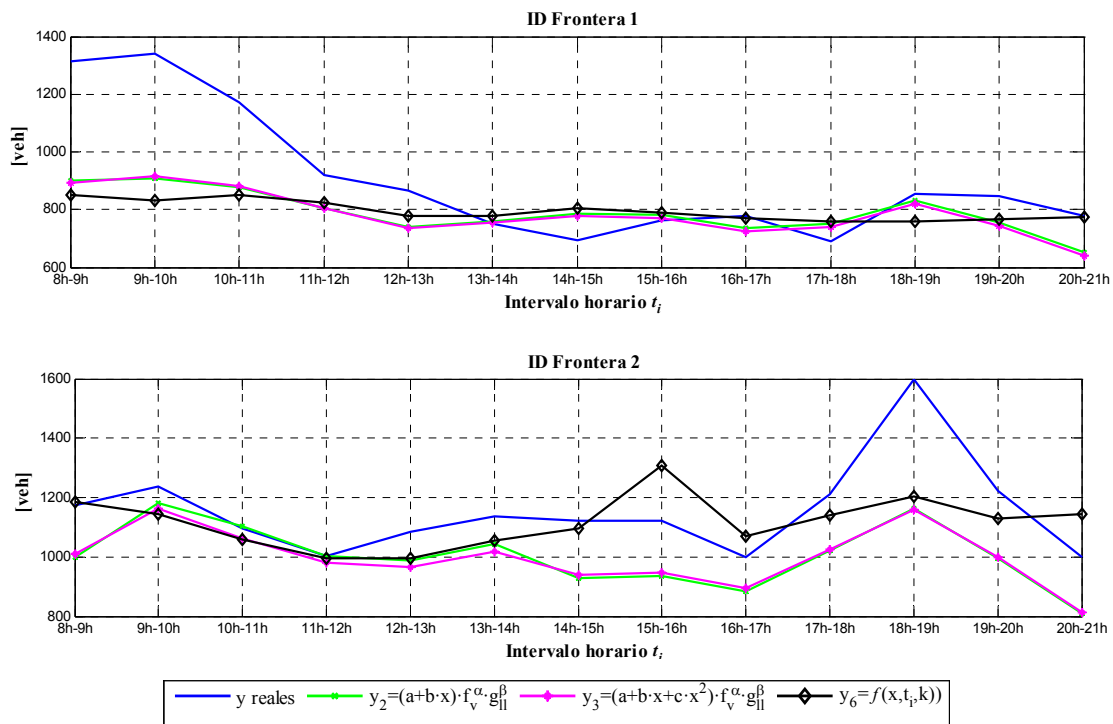
**Figura 4-28:** Volúmenes de vehículos observados y estimados con cada modelo que cruzan la frontera 9 y la frontera 10 en cada periodo horario.



**Figura 4-29:** Volúmenes de vehículos observados y estimados con cada modelo que cruzan la frontera 11 y la frontera 12 en cada periodo horario.



Gráficamente, se aprecia que los modelos 2 y 3 consiguen valores muy similares de estimaciones de volúmenes de vehículos dado un número de llamadas en movilidad. Esto tiene su origen en la forma funcional de los modelos, dado que ambas se diferencian únicamente en el término cuadrático que multiplica al número de llamadas en movilidad ( $c \cdot x^2$ ). El orden de magnitud de ese parámetro  $c$  no introduce un cambio significativo en la predicción, mientras que el resto de parámetros se mantienen del mismo orden en ambos modelos. En consecuencia, las diferencias entre las predicciones de ambos modelos apenas son apreciables en la escala gráfica utilizada, no ocurriendo lo mismo atendiendo a otros criterios como error absoluto, error relativo o correlación, entre otros, como ya se demostró con anterioridad. Algo similar ocurre con las predicciones de los modelos 4 y 5. Por ello, resulta complejo establecer una clasificación clara atendiendo a un análisis visual. No obstante, centrándose en los modelos seleccionados según los criterios desarrollados en los apartados anteriores, modelo 2, 3 y 6, se observa que sus predicciones son las que mejor seguimiento realizan de los picos y valles de los valores observados de volúmenes de vehículos. Este hecho encaja con las conclusiones alcanzadas anteriormente en términos de error, percentiles y correlación. Por ello, parece lógico pensar que esos tres modelos – modelo 2, 3 y 6 – son los más adecuados para la predicción de volúmenes de vehículos a partir de un número de llamadas en movilidad realizadas en un periodo de tiempo. Las siguientes figuras se centran en las predicciones de volúmenes de vehículos hechas por esos tres modelos y los valores reales observados por estaciones de aforo.



**Figura 4-30: Volúmenes de vehículos observados y estimados con modelos 2, 3 y 6 que cruzan la frontera 1 y la frontera 2.**

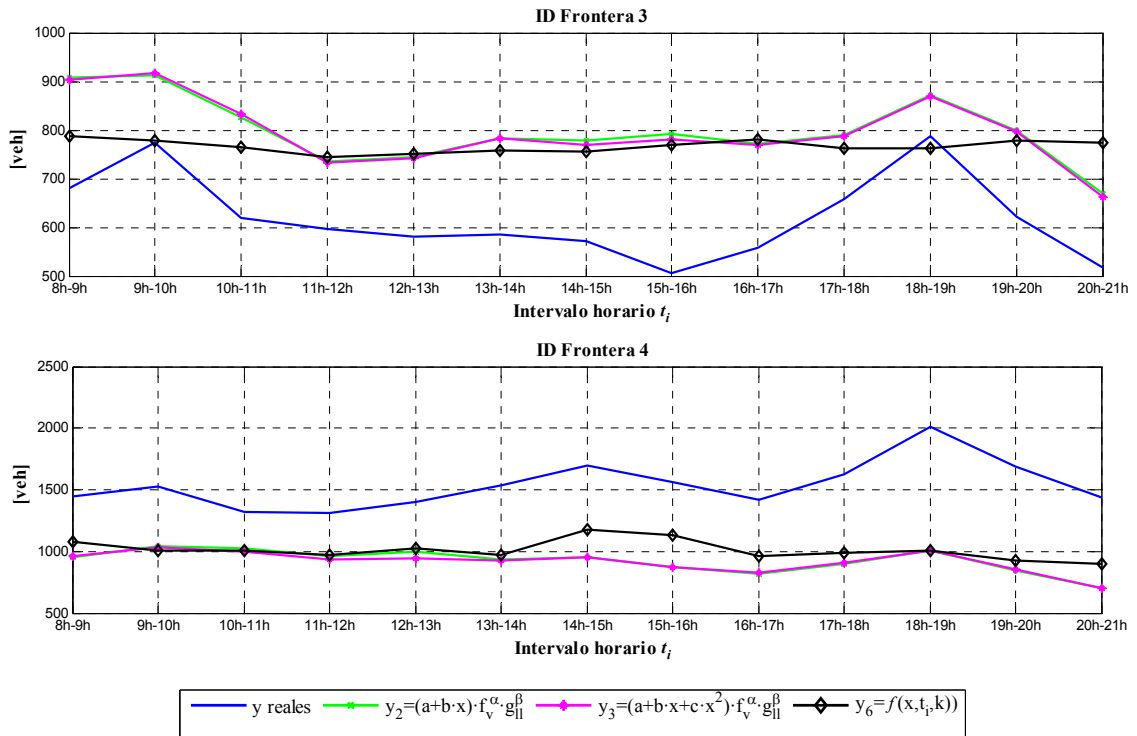


Figura 4-31: Volúmenes de vehículos observados y estimados con modelos 2, 3 y 6 que cruzan la frontera 3 y la frontera 4.

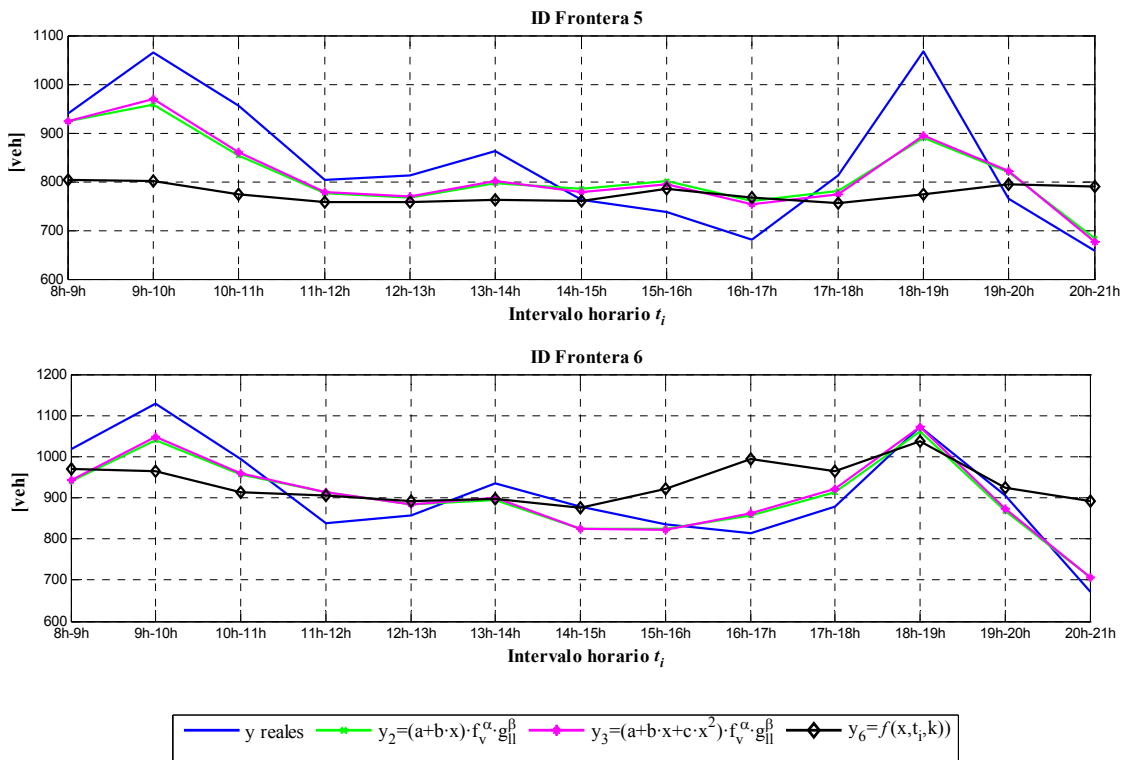


Figura 4-32: Volúmenes de vehículos observados y estimados con modelos 2, 3 y 6 que cruzan la frontera 5 y la frontera 6 en cada periodo horario.

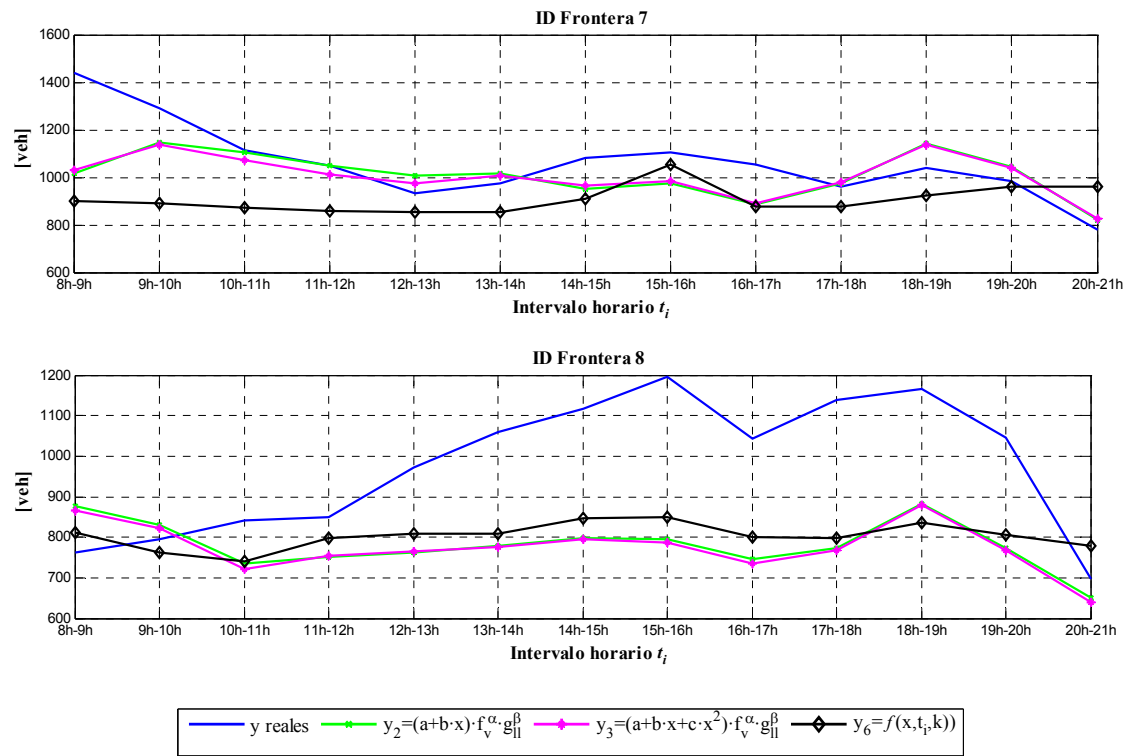


Figura 4-33: Volúmenes de vehículos observados y estimados con modelos 2, 3 y 6 que cruzan la frontera 7 y la frontera 8 en cada periodo horario.

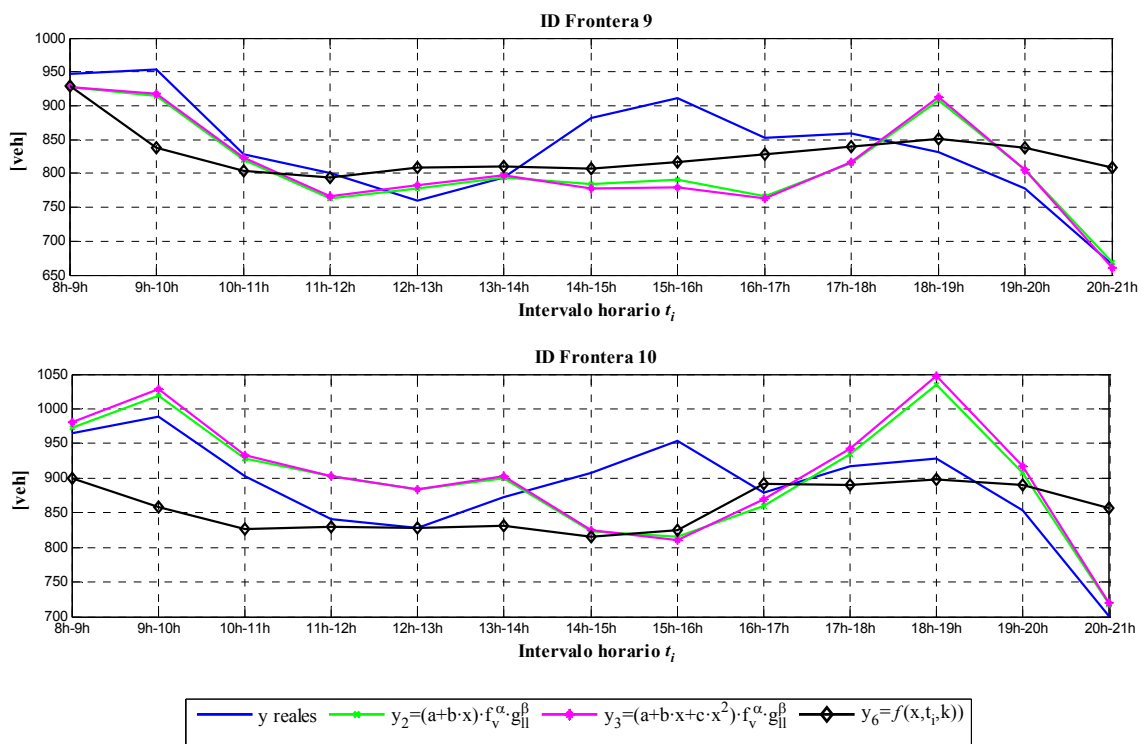
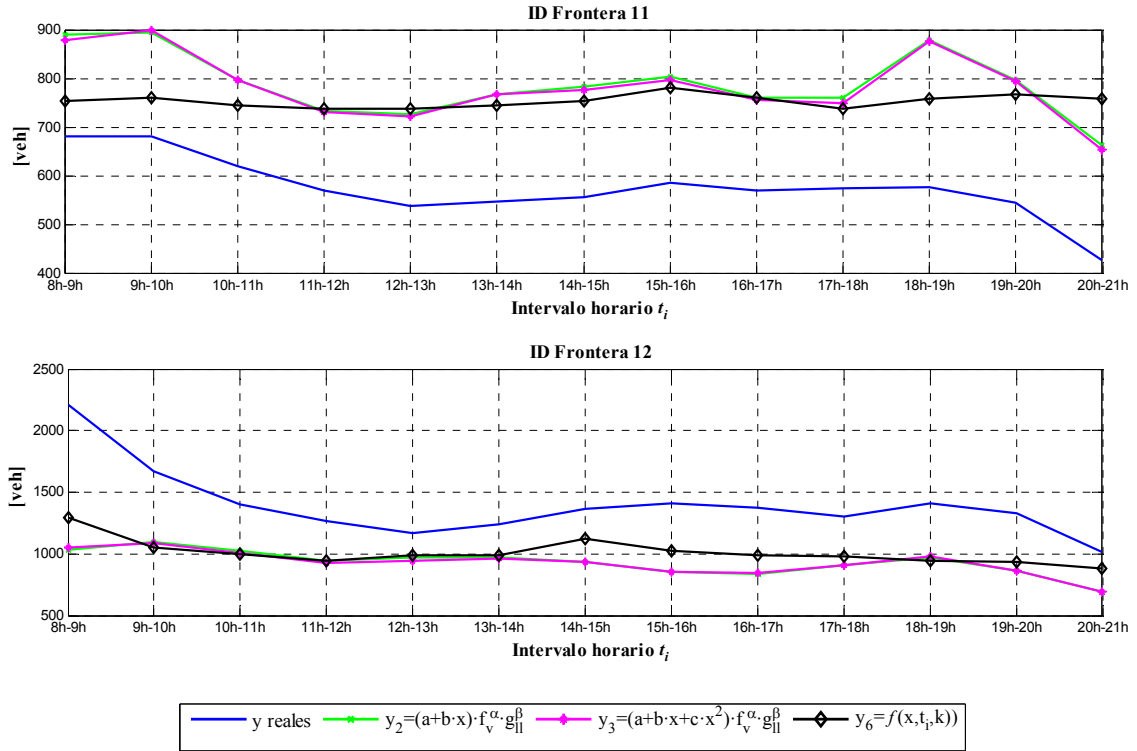


Figura 4-34: Volúmenes de vehículos observados y estimados con modelos 2, 3 y 6 que cruzan la frontera 9 y la frontera 10 en cada periodo horario.



**Figura 4-35: Volúmenes de vehículos observados y estimados con modelos 2, 3 y 6 que cruzan la frontera 11 y la frontera 12 en cada periodo horario.**

### 4.3.7 Conclusiones

En este Capítulo se ha presentado una metodología para estimar volúmenes de vehículos que cruzan una frontera entre celdas en función de datos de una red de telefonía móvil. Tras discutir la información necesaria para este objetivo, se han desarrollado seis modelos para inferir dichos volúmenes empleando las llamadas en movilidad generadas en cada periodo horario, junto con otra información asociada a las características de esas llamadas, tales como intensidad horaria de llamadas, duración, o incluso las características de las vías que cruzan la frontera. Los modelos formulados han sido los siguientes:

- Modelo 1:  $\hat{y}_1(t_i) = a \cdot f_v^\alpha(t_i) \cdot g_{ll}^\beta(t_i)$ .
- Modelo 2:  $\hat{y}_2(t_i) = [a + b \cdot x(t_i)] \cdot f_v^\alpha(t_i) \cdot g_{ll}^\beta(t_i)$ , donde  $x(t_i) = n_{ll\ mov}(t_i, k)$ .
- Modelo 3:  $\hat{y}_3(t_i) = [a + b \cdot x(t_i) + c \cdot x^2(t_i)] \cdot f_v^\alpha(t_i) \cdot g_{ll}^\beta(t_i)$ , donde  $x(t_i) = n_{ll\ mov}(t_i, k)$ .
- Modelo 4:  $\hat{y}_4(t_i) = a + b \cdot x(t_i)$ , donde  $x(t_i) = n_{ll\ mov}(t_i, k)$ .
- Modelo 5:  $\hat{y}_5(t_i) = a + b \cdot x(t_i) + c \cdot x^2(t_i)$ , donde  $x(t_i) = n_{ll\ mov}(t_i, k)$ .

- Modelo 6: 
$$\hat{y}_6(t_i, k) = \frac{a \cdot x(t_i, k)}{P_{ll\ veh}^2(t_i) + P_{ll\ veh}(t_i) \cdot \frac{b_1}{\alpha(t_i, k)} [1 - e^{-b_2 \cdot \alpha(t_i, k)}]} + d, \text{ con } \alpha(t_i, k) = \frac{V_k}{L_k \cdot T_c(t_i)}$$

Los modelos 1, 2, 3, 4 y 5 no establecen diferenciación en sus predicciones según la frontera monitorizada, de modo que la principal variable de entrada al modelo,  $x(t_i)$ , es el número de llamadas en movilidad detectadas en una frontera  $k$  cualquiera,  $x(t_i) = n_{ll\ mov}(t_i, k)$ . En cambio, el modelo 6 sí que posee una dependencia espacial, estableciendo una diferenciación según la frontera monitorizada. Así pues, las variables de entrada al modelo son las llamadas en movilidad en una determinada frontera,  $x(t_i, k) = n_{ll\ mov}(t_i, k)$ , junto a otro tipo de información relacionada con características de las llamadas y de la propia frontera.

El ajuste de los parámetros de los modelos definidos se ha llevado a cabo buscando la minimización de la suma del valor absoluto del error relativo entre valores observados y modelados. El Capítulo se ha completado con un estudio numérico en el que se contrastan los modelos y se comparan entre sí en base a una serie de criterios. Dada la variedad de modelos, estos criterios de comparación contemplan indicadores tales como las clásicas medidas de error (medio, absoluto y relativo), percentiles del error relativo absoluto, o correlaciones (lineal y por rango) entre volúmenes reales y predicciones, además de un análisis gráfico de resultados. También se han empleado el criterio de información de Akaike y el criterio de información bayesiano como medidas de contraste para evaluar la forma funcional en la selección de los mejores modelos.

Tras el estudio de comparación y contraste se extraen una serie de conclusiones, destacando que aquellos modelos que incluyen en su forma funcional tanto información de llamadas en movilidad como sobre variabilidad horaria de llamadas realizadas y/o de vehículos circulando, alcanzan predicciones más exactas que el resto. Tal es el caso de los modelos 2, 3 y 6.

En base a los criterios estudiados, se puede afirmar que los modelos 1, 4 y 5 son los que peores prestaciones consiguen. En lo que respecta al modelo 1, esto se refleja en el hecho de que el modelo depende exclusivamente de los factores de variabilidad horaria de intensidad de llamadas y de vehículos circulando, independientes de la frontera. La información del número de llamadas en movilidad realizadas no interviene en el modelo, perdiéndose entonces la información respecto a un cierto orden de magnitud del paso por la frontera en cuestión. Respecto a los modelos 4 y 5 ocurre algo similar pero en este caso, aunque su forma funcional sí que depende de las llamadas en movilidad realizadas en cada frontera en cuestión, carece de información sobre variabilidad horaria. Ya se razonó en apartados anteriores que el intervalo horario es un aspecto importante a la hora de realizar las predicciones. De ahí que sus

prestaciones, aunque en menor medida que las del modelo 1, sigan proporcionando resultados limitados. En cambio, los modelos 2 y 3, los cuales sí contienen información tanto de las llamadas en movilidad como de variabilidad horaria, consiguen resultados eficientes, demostrando una capacidad predictiva admisible para cualquier frontera entre celdas.

Por otro lado, el modelo 6 alcanza predicciones razonables de volúmenes de vehículos aunque exige realizar un estudio previo para caracterizar la frontera en términos de velocidad y longitud de las vías que discurren por la celda origen de cada frontera. Su forma funcional establece una dependencia espacial con la frontera monitorizada, por lo que las predicciones no sólo dependen del número de llamadas en movilidad sino sobre qué frontera han sido observadas. Este modelo es menos flexible que los modelos 2 y 3, dado que requiere un estudio previo de la frontera, pero su capacidad predictiva, evaluada en anteriores apartados, también lo habilita como un modelo apto para estimar volúmenes de vehículos. En consecuencia, los modelos 2, 3 y 6 serán los seleccionados para predecir volúmenes de vehículos, por ser los que mejores prestaciones alcanzan; mientras que los modelos 1, 4 y 5 son descartados como modelos viables para la inferencia.

Observando los resultados respecto a la comparación de los criterios estudiados sobre esos tres modelos {2, 3 y 6}, resulta difícil establecer un orden de prioridad claro dado que:

- i) respecto a las medidas de error medio, error absoluto y error relativo, así como de percentiles del error relativo en valor absoluto, aunque el modelo 6 presenta una ligera ventaja respecto a 2 y 3, el orden de magnitud es muy similar en los tres modelos;
- ii) gráficamente, las predicciones de los tres modelos reflejan un comportamiento semejante a la hora de alcanzar con exactitud los picos y valles en las curvas de los volúmenes reales;
- iii) respecto a correlación por rango y correlación lineal, el modelo 6 refleja una cierta mejora respecto a los modelos 3 y 2;
- iv) aunque, respecto a los estadísticos *AIC*, *BIC* y log-verosimilitud, los modelos 3 y 2, son mejores que el modelo 6.

En general, puede decirse que el modelo 6 presenta una capacidad predictiva ligeramente superior al considerar dependencia espacial, función de características geométricas que habría que medir específicamente para cada punto kilométrico monitorizado. Sin embargo dicha mejora puede no ser suficiente para justificar la pérdida de flexibilidad del modelo frente a los modelos 2 y 3, que son independientes de la frontera. A la vista de estas conclusiones,

finalmente se ha optado por considerar como igualmente viables los modelos 2, 3 y 6 para estimar volúmenes de vehículos, poniéndose de manifiesto:

- i) lo insuficiente de usar exclusivamente información de llamadas en movilidad,
- ii) la necesidad de incorporar en los modelos información sobre la variabilidad horaria de la conducta de los usuarios que se desplazan, en medios mecanizados, a la hora de realizar llamadas, u otras características asociadas a la circulación de vehículos, y
- iii) la aplicabilidad de la metodología a cualquier frontera dentro de unos rangos admisibles de error en comparación con medidas proporcionadas por las tradicionales estaciones de aforo.

Respecto a este último punto, es necesario destacar lo siguiente. Los teléfonos móviles se han convertido hoy en día en un elemento indispensable en nuestra vida diaria, especialmente en lo que respecta a nuestra movilidad. Los datos que los sistemas de telefonía móvil generan para su propio funcionamiento pueden arrojar cierta información relacionada con nuestros movimientos. De ahí que existan en la literatura diferentes estudios en los que se afirma que la telefonía móvil puede convertirse en una herramienta viable para la monitorización del tráfico. En este Capítulo se han propuesto una serie de modelos para la estimación de volúmenes de vehículos en función de, entre otras variables, llamadas realizadas por teléfonos móviles bajo ciertas situaciones expuestas en el apartado 4.3.3, denominadas llamadas en movilidad. La naturaleza de la principal variable de entrada al modelo, llamadas en movilidad, posee una fuerte componente de aleatoriedad. Este tipo de datos no representa el conjunto total de teléfonos que han pasado por una zona sino sólo aquellos pertenecientes al operador de interés que, a su vez, han realizado llamadas bajo ciertas condiciones. Ello condiciona en gran medida la representatividad de la muestra y, por tanto, de los resultados. Pese a estas dificultades, los valores de predicción de los modelos finalmente seleccionados han alcanzado resultados razonables en comparación con valores observados por estaciones de aforo. En este sentido, la validación experimental de los modelos propuestos ha manifestado evidencias de que la telefonía móvil, mediante llamadas en movilidad, constituye una alternativa válida para la determinación de volúmenes de vehículos pasando por una zona frente a los procedimientos tradicionales basados en aforos.

Una aplicación inmediata de esos volúmenes inferidos de la telefonía móvil se puede enfocar al campo de las metodologías de ajuste de matrices de movilidad, tradicionalmente realizadas a partir de aforos de la red, aunque realizando una serie de modificaciones que contemplen el origen de los volúmenes utilizados. El siguiente capítulo aborda esta cuestión.





# CAPÍTULO 5

---

## APLICACIÓN A LA ESTIMACIÓN DE MATRICES ORIGEN–DESTINO

### 5.1 INTRODUCCIÓN

Las matrices de viajes Origen–Destino (O–D) son una fuente esencial de información correspondiente a la demanda de transporte en todos los estudios de planificación, gestión y control del sistema de transporte. Estas matrices aportan información sobre el número de viajes realizados entre una zona origen  $i$  y una destino  $j$  en un determinado intervalo de tiempo. Por ello, para caracterizar el funcionamiento de un sistema de transporte es básico conocer, entre otros parámetros, la matriz origen–destino. Históricamente, las matrices origen–destino han sido estimadas en base a tres metodologías diferentes:

1. Estimación directa: Llevando a cabo un proceso de encuestación, domiciliario o viario (directo sobre una muestra de vehículos o por ficha proporcionada a conductores, o indirecto por captación de números de matrículas).
2. Estimación por modelos de demanda: Haciendo uso de un modelo de distribución (e.g.: modelo de gravedad).
3. Estimación mediante conteos de tráfico: Utilizando conteos de tráfico a fin de actualizar una matriz origen–destino preexistente.

De las tres aproximaciones anteriores, la tercera alternativa es la más profusamente utilizada en las últimas décadas y ha dado lugar a innumerables contribuciones desde el trabajo pionero de Low (1972). La metodología en la que se basa exige, como prerrequisito, un conjunto de conteos de flujos de tráfico (volúmenes observados) y una matriz O–D previa, la cual puede ser obtenida de algún estudio cronológicamente anterior o de una encuestación básica que predefina la estructura de la matriz final a obtener. Bajo el enfoque de esta metodología, suelen utilizarse modelos basados en optimización matemática para intentar determinar la matriz de viajes más

idónea. Dicha matriz debe ser consistente con la información contenida en los volúmenes de tráfico en los arcos observados mediante el uso de una función objetivo a minimizar basada en información con respecto a la matriz de viajes previa. En estos modelos basados, la matriz previa se supone que es una matriz O-D desactualizada a la que hay que realizar ajustes/cambios para adecuarse a los volúmenes observados cuando sea asignada a la red.

Recientemente, los sistemas de telefonía móvil se están considerando como una tecnología prometedora para la captura de datos de tráfico debido a sus ventajas con respecto a las técnicas tradicionales. Respecto a ello, en el anterior capítulo se presentó una metodología que permitía inferir volúmenes de tráfico a partir de conteos de paso de teléfonos por fronteras entre celdas. Dichas fronteras pueden englobar uno o varios arcos de la red de transporte, proporcionando en ese caso información agregada de volúmenes para grupos de arcos en lugar de volúmenes a nivel de arco individual. Parece lógico que las metodologías de estimación clásicas deban adaptarse para incorporar este tipo de información a su formulación de modo que pueda ser usada en lugar de (o además de) medidas proporcionadas por los tradicionales aforos.

Este capítulo presenta una metodología para la estimación o ajuste de una matriz previa cuya principal novedad es el uso de información agregada de volúmenes observados en agrupaciones de arcos inferidos de la telefonía móvil. En particular, se ha optado por adaptar la metodología propuesta por Doblas y Benitez (2005), que ya demostró su eficiencia en la estimación de matrices con aforos tradicionales, para que pueda ser empleada con la información inferida de la telefonía, es decir, a nivel de volumen por grupos de arcos. La elección como base teórica de esa metodología reside en que, a diferencia de otras, ésta implementa un control de la distorsión experimentada por la matriz previa durante el ajuste. En general, este tipo de metodologías busca que la matriz resultante reproduzca, al ser asignada a la red, los volúmenes observados lo más fielmente posible, y derive de una matriz previa que proporciona información útil al proceso dado que se supone como una observación de la “verdadera” matriz O-D a estimar. Llevar a cabo una encuesta amplia y detallada que permita elaborar una matriz previa resulta ser un proceso costoso, por lo que una distorsión excesiva de dicha información invalida cualquier esfuerzo presupuestario y humano dedicado a su elaboración. En base a este razonamiento, en la definición de la función objetivo se buscará un compromiso entre la distorsión de la solución con relación a la matriz previa y a los volúmenes observados a nivel agregado de arcos.

Se ha empleado como punto de partida la metodología de Doblas y Benitez (2005) por haber sido ampliamente estudiada y desarrollada por nuestro grupo de investigación, así como por haber mostrado magníficos resultados en el campo del ajuste de matrices con volúmenes observados. No obstante, la adaptación implementada en la formulación a nivel de volumen por

grupos de arcos también sería válida para cualquier otra metodología de ajuste de matrices basada en volúmenes observados (por ejemplo, Spiess, 1990).

## 5.2 ANTECEDENTES Y ESTADO ACTUAL

Los métodos de estimación de matrices origen–destino basados en la utilización de datos de aforos automáticos para actualizar una matriz previa han experimentado un gran desarrollo en los últimos 30 años. Esto lo ha motivado, fundamentalmente, el hecho de que este tipo de información se puede conseguir más rápida y económicamente que la obtenida tradicionalmente mediante la realización de encuestas.

Estimar una matriz O–D puede ser visto como el problema de encontrar una matriz que, cuando sea asignada a la red, reproduzca los conteos de tráfico observados. En general, hay un gran número de matrices que pueden reproducir los conteos de tráfico observados, de modo que el sistema de ecuaciones resultante es indeterminado y puede tener muchas soluciones posibles. El número de pares O–D (elementos de la matriz) excede en gran medida el número de arcos para los cuales se han recogido los conteos de tráfico. Por tanto, el problema pasaría a ser la elección de la mejor matriz entre todas las posibles que reproducen los conteos de tráfico observado. Algunos modelos resuelven este problema postulando un modelo general de distribución de viajes, por ejemplo un modelo de gravedad, mientras que otros adoptan técnicas de inferencia estadística. Son innumerables las contribuciones en este campo con algoritmos desarrollados para la resolución del problema de estimación de la matriz O–D usando conteos de tráfico. Los métodos más comunes incluyen maximización de la entropía (Van Zuylen y Willumsen, 1980), máxima verosimilitud (Spiess, 1987), mínimos cuadrados generalizados (Cascetta, 1984; Bell, 1991) o técnicas de inferencia bayesiana (Maher, 1983). Una revisión de esos modelos se puede consultar en Cascetta y Nguyen (1988), Chen y Florian (1996), Abrahamsson (1998).

Respecto a la obtención de una matriz O–D, diferentes estudios revisados en el Apartado 3.2 relacionados con uso de teléfonos móviles (Pan *et al.*, 2006; White y Wells, 2002; Akin y Sisipiku, 2002), sugieren la posibilidad de obtener directamente matrices O–D mediante registros de localización generados por usuarios de teléfonos móviles. Una matriz O–D elaborada mediante datos de teléfonos puede clasificarse según el nivel de agregación que se logra alcanzar, es decir, según las “zonas de transporte” sean las celdas o las áreas de localización existentes en la zona de servicio de red existente en la región bajo estudio. Esta diferenciación es necesaria puesto que está estrechamente relacionada con el tamaño de la muestra que generará la matriz. Cuando las “zonas de transporte” son las celdas, se precisa que los registros de localización estén actualizados a nivel de celda, lo cual sólo ocurre en el caso de

teléfonos que realicen llamadas. En el caso de que se trabaje a nivel de áreas de localización, los registros deben de estar actualizados a nivel de LA, lo cual ocurre en los sistemas de telefonía con independencia del estado del teléfono, siendo por tanto la muestra más extensa.

Independientemente del nivel de agregación seleccionado, la elaboración de este tipo de matrices implicaría monitorizar los registros iniciales y finales que cada teléfono efectúa sobre la región analizada en el intervalo de tiempo de interés, los cuales determinan la zona origen y destino del viaje asociado al usuario portador del teléfono. Este análisis permite distinguir entre los viajes de interés (pares O–D de la matriz) del resto de viajes de paso, los cuales son viajes que se iniciaron o terminaron en otros orígenes o destinos distintos a los estudiados.

En general, un proceso de encuestación tradicional requiere trabajar con una muestra representativa de la población. Debido a ello, las metodologías de estimación basadas en matrices procedentes de encuestas resultan o poco fiables o muy costosas desde el punto de vista económico y/o social. Sin embargo, el empleo de datos de teléfonos móviles supone utilizar una muestra muy extensa de la población, puesto que las tasas de penetración de telefonía se sitúan ya en torno al 90% en muchos países europeos (CDB, 2006), siendo en España el 110.8% (CMT, Enero 2009). Junto a ello destaca su bajo coste de implantación dado que utiliza la misma infraestructura existente en los sistemas de telefonía, sin necesidad de instalar módulos adicionales. Sin embargo, en lo que respecta a precisión de localización, este tipo de matrices derivadas directamente de la telefonía móvil posee ciertos inconvenientes que las invalidan como matrices de viajes finales. La desventaja principal se encuentra fuertemente relacionada con el tamaño de las “zonas de transporte” derivadas de una zonificación que emplee la distribución celular de una red de telefonía móvil.

Desde el punto de vista de dimensión de las zonas de transporte que definen una matriz O–D, la zonificación tradicional se realiza minuciosamente mediante criterios socioeconómicos o agregaciones de población seleccionando zonas que pudieran tener una relación causal con los movimientos que se dan entre ellas. El tamaño de estas zonas debe ser tal que el error de agregación causado por el supuesto que todas las actividades se concentran en ella no sea muy grande, es decir, bajo impacto de viajes intra–zonales. Además, estas zonas deben ser lo más homogéneas posible en cuanto al uso del suelo y/o la composición de la población. En cambio, en matrices elaboradas directamente con registros telefónicos, las “zonas de transporte” vienen ya fijadas por el propio diseño del sistema de telefonía, puesto que se utilizan las mismas celdas y/o áreas de localización definidas en el dimensionado de la red celular. El tamaño de las celdas respecto al ámbito geográfico de su cobertura depende en gran medida de parámetros tales como tráfico que tiene soportar, topología del terreno, tipo de antena o incluso edificios alrededor de

la estación base, variando el radio de una celda desde cientos de metros (entornos urbanos) a kilómetros (entornos rurales). Considerando a su vez que las áreas de localización se definen como agrupaciones de celdas, este tamaño será considerablemente mayor.

Atendiendo al enfoque de uso de esas áreas (celdas y/o LA) como “zonas de transporte”, no será válida la hipótesis de bajo impacto de viajes intra-zonales dado que es lógico que comprendan varias de las clásicas zonas de transporte entre las que sí se dan viajes. Con las posibilidades que ofrece la telefonía móvil, esos viajes intra-zonales serán generalmente imposibles de detectar aunque, si se atiende a matrices en términos de macro zonas (agrupación de varias “zonas de transporte” clásicas según un determinado criterio), se permitiría elaborar una matriz de viajes. Por tanto, el uso de datos de telefonía es válido para obtener una matriz final a nivel de macro zona o una matriz exclusivamente de viajes inter-zonales que pueda ser utilizada como previa para, posteriormente, ajustarla con otra información (aforos, volúmenes agregados derivados de teléfonos,...) mediante métodos tradicionales, pero no para usarla directamente como matriz de viajes final. Además, la obtención de una matriz de movilidad monitorizando el origen y destino de una muestra de teléfonos móviles presenta como desventaja que las matrices O-D inferidas no estarían desagregadas ni por motivo ni por modo de transporte, ni aportarían información sobre características de los individuos, como así ocurre en el caso de las procedentes de encuestas. En cambio presentarían como ventaja destacable que la obtención de la matriz sería de forma automática y sobre una muestra de tamaño elevado.

No obstante, los métodos más extendidos para la resolución del problema de estimación de una matriz de viajes O-D son aquellos basados en modelos matemáticos que utilizan las mediciones de aforos en un conjunto de arcos de la red de tráfico junto con otra información previa. Son en ellos en los que se ha centrado esta Tesis para la estimación de matrices, fusionando conceptos tradicionales de estimación mediante modelos matemáticos junto con una nueva fuente de información de volúmenes de tráfico derivados de la telefonía móvil.

## **5.3 METODOLOGÍA DE AJUSTE DE MATRICES ORIGEN-DESTINO MEDIANTE VOLUMENES AGREGADOS**

### **5.3.1 Introducción**

Una matriz O-D de viajes,  $T=[T_{ij}]$ , es una tabla de doble entrada que, en una celda genérica  $ij$ , contiene el número de viajes que tienen lugar entre la “zona de transporte” origen  $i$  y la de destino  $j$  durante un cierto periodo de tiempo. Los aforos o conteos de tráfico tradicionales

proporcionan el volumen total de vehículos que circula por cada uno de los arcos observados de la red de transporte en dicho periodo de tiempo.

La mayoría de los modelos de actualización de matrices emplean conteos de tráfico obtenidos de aforos (detectores) instalados sobre arcos del viario. Ciertos sistemas que realizan tales conteos suelen ser descartados para su uso en métodos de estimación de matrices debido a que no son capaces de desagregar la información medida sobre grupos de arcos en información individual de cada arco (e.g. las IMD no desagregan por sentido de circulación). Este trabajo propone una metodología para el uso de la información agregada proporcionada por sistemas de estas características. En particular, la metodología se va a centrar en datos de volúmenes derivados de la telefonía móvil por considerarse ésta como una tecnología prometedora para la captura de datos de tráfico. Para ello es necesario introducir novedades metodológicas para la explotación de la información agregada de movilidad entre zonas derivada de tales sistemas de telefonía, es decir, volúmenes observados sobre grupos de arcos de la red en lugar de sobre arcos individuales (tradicionales aforos). Se ha diseñado una metodología fácil de aplicar, adecuada para abordar problemas de estimación de matrices en aplicaciones reales de gran envergadura con un gran número de zonas de transporte involucradas.

### **5.3.2 Definición del problema**

La necesidad de obtener matrices de viajes O–D de forma rápida y económica es esencial en la planificación del transporte. Una matriz O–D es difícil y, en ocasiones, costosa de obtener por mediciones directas o encuestas; en cambio, se puede obtener una estimación razonable de ella usando conteos de tráfico en un conjunto de arcos de la red de transporte y otra información adicional. Esta información puede estar disponible en términos de una matriz O–D previa,  $\bar{T}$ , y/o en el número de viajeros que se atrajeron o se originaron en diversas zonas.

Los métodos de estimación de matrices más extendidos solamente necesitan de conteos de tráfico en un conjunto de arcos y una información a priori de la estructura de la matriz de viajes. En general, la mayoría de los modelos para la estimación de una matriz O–D utilizan información a priori sobre ella en términos de una matriz O–D previa. Ésta puede ser obtenida por una encuesta sobre una muestra o a partir de una matriz antigua, probablemente desactualizada considerada como una observación próxima a la “verdadera” matriz que se desea estimar. Por otro lado, los datos del conjunto de conteos de tráfico también se consideran como una observación directa de la realidad que el modelo matemático debe reproducir. Por tanto, en el proceso de ajuste, la matriz previa se considera que es una matriz O–D obsoleta a la que hay

que realizar cambios para obtener las mínimas desviaciones posibles entre los volúmenes estimados en los arcos y los observados.

En general, se puede concluir que los distintos métodos para el problema de estimación de matrices de viajes O–D, desarrollados en la literatura a partir de volúmenes de tráfico observados, presentan la forma genérica siguiente (Yang *et al.*, 1992):

$$\begin{array}{ll} \text{Minimizar} & \alpha F_1(\mathbf{T}, \bar{\mathbf{T}}) + \beta F_2(\mathbf{v}, \bar{\mathbf{v}}) \\ \mathbf{v}, \mathbf{T} & \\ \text{s.a.} & \mathbf{v} = M(\mathbf{T}) \end{array} \quad (21)$$

Las funciones  $F_1$  y  $F_2$  son dos métricas que miden la distancia entre la matriz O–D estimada,  $\mathbf{T}$ , y la matriz previa dada,  $\bar{\mathbf{T}}$ , y entre los volúmenes estimados y los observados en los arcos aforados,  $\mathbf{v}$  y  $\bar{\mathbf{v}}$  respectivamente. Los procedimientos más habituales para las funciones  $F_1$  y  $F_2$  son el de máxima verosimilitud, mínimos cuadrados generalizados y derivaciones del principio de máxima entropía. Los parámetros  $\alpha$  y  $\beta$  son los correspondientes factores de peso que reflejan la confianza relativa en los datos disponibles  $\bar{\mathbf{T}}$  y  $\bar{\mathbf{v}}$ . Finalmente, la expresión  $\mathbf{v} = M(\mathbf{T})$  representa el proceso de asignación considerado para modelar los volúmenes en los arcos a partir de la matriz estimada.

Típicamente, el problema de estimación puede ser interpretado como la búsqueda de una matriz O–D que, una vez asignada a la red, reproduzca los conteos de tráfico observados y, además, sea semejante a la matriz previa. La nueva contribución respecto a los métodos existentes en la literatura es el uso de información agregada de volúmenes observados sobre grupos de arcos de la red en lugar de sobre arcos individuales. Este tipo de información puede ser usada conjuntamente con los conteos de tráfico obtenidos de las tradicionales estaciones de aforo, dado que los aforos son un caso especial en el que el número de arcos del grupo es uno. El procedimiento de estimación presentado en esta Tesis se basa en los dos criterios siguientes, que permiten traducirlo a expresiones matemáticas con vistas a su implementación práctica:

1. Que la matriz resultante reproduzca, al ser asignada a la red, los datos de los volúmenes agregados lo más fielmente posible.
2. Que la información contenida en la matriz O–D previa, obtenida habitualmente por medio de complejos y costosos procesos de encuestación domiciliaria, no se vea distorsionada en exceso. Una distorsión excesiva de dicha información invalida cualquier esfuerzo presupuestario y humano dedicado a tareas de encuestación.

Esta metodología implementa un algoritmo de ajuste de matrices a partir de información agregada de volúmenes observados, junto con una matriz previa. La matriz previa a utilizar

procederá de una generada mediante un proceso de encuestación o de una matriz antigua (probablemente desactualizada). Para los volúmenes, pese a que la información puede proceder de cualquier sistema que proporcione conteos de tráfico en ese formato agregado sobre grupos de arcos, la metodología se centrará en el uso de conteos derivados de sistemas de telefonía móvil. Recientemente, estos sistemas están siendo empleados con éxito para la monitorización del tráfico; por ello el estudio se centra en el contexto de volúmenes agregados derivados de ellos. No obstante, la metodología es aplicable a cualquier otro sistema que proporcione volúmenes de forma agregada, incluso puede ser combinada con los tradicionales aforos. El algoritmo de ajuste fue diseñado para ser implementado fácilmente en códigos comerciales de asignación – EMME/2 (1990), MINUTP (1993), TRANSCAD (2006), TRIPS (1998), etc.

### 5.3.3 Formulación matemática

La idea más intuitiva y sencilla para traducir los criterios anteriormente expuestos a una formulación matemática adecuada es plantear el problema como uno de optimización matemática en el que se trata de minimizar una función cuadrática que mida las discrepancias entre volúmenes observados (ya sean de aforos automáticos, derivados de teléfonos móviles o de cualquier otro sistema) y volúmenes teóricos o modelados (los que se obtienen por asignación), así como la “distancia” entre la matriz O–D de viajes previa y los sucesivos estimados que se obtienen como resultado del proceso de ajuste. En la formulación que se propone se respeta esta idea en lo que respecta a los volúmenes observados, sin embargo, para controlar la distorsión de la matriz previa se opta por un conjunto de restricciones de variables acotadas (para cada celda de la matriz O–D) y de restricciones funcionales (para la información contenida en la matriz O–D en distintos niveles de agregación: viajes generados al agregar según destinos, viajes atraídos al agregar según orígenes y viajes totales al agregar en orígenes y destinos) que mantengan la variación de la información contenida en la matriz estimada respecto de la previa dentro de unos rangos considerados admisibles.

Se considera que el área objeto de nuestro estudio se encuentra dividida en  $n_o+n_d$  zonas de transporte ( $n_o$  orígenes y  $n_d$  destinos) con viajes desde cada una de las zonas origen hasta todos los destinos. Cada zona es representada por un nodo denominado *centroide*, donde los viajes se inician o terminan. La matriz de viajes se denotará por  $\mathbf{T} = [T_{ij}]$ , siendo su elemento  $(i, j)$  el número de viajes desde el origen  $i$  al destino  $j$  durante un cierto periodo de tiempo. La red de transporte correspondiente al área estudiada se representa mediante un grafo dirigido  $G=(N,A)$  donde  $N$  y  $A$  son el conjunto de *nodos* y *arcos* (dirigidos) respectivamente. En los arcos del modelo de la red de transporte pueden existir congestión y esto se suele modelar por funciones de coste  $s_a(v_a)$ . Las funciones de coste describen la relación entre el volumen de tráfico en el



arco y el coste (o el tiempo) para recorrerlo. Se establece la hipótesis de que las funciones de coste son separables, el tiempo en recorrer un arco no depende del volumen de otros arcos. El tratamiento de los efectos de la congestión es una característica importante para distinguir entre los modelos para la evaluación de la matriz O–D. Cada modelo supone que la congestión se puede tratar exógenamente (por asignación proporcional) o endógenamente (por asignación en equilibrio). Finalmente, la asignación de la matriz O–D al modelo de red para obtener los volúmenes y tiempos de viaje modelados sobre cada arco se basa en el equilibrio de usuario (user–equilibrium) determinista o estocástico, cumpliendo los requisitos de asignación en equilibrio según el primer principio de Wardrop (Patriksson, 1994).

A continuación se presenta la notación matemática necesaria para la formulación del modelo.

### Índices:

$i \in I$	zonas origen ( $n_o$ elementos)
$j \in J$	zonas destino ( $n_d$ elementos)
$a \in A$	arcos genéricos de la red
$b \in B$	fronteras (cada frontera comprende un conjunto de arcos observados)
$k \in K_{ij}$	rutas o caminos desde origen $i$ a destino $j$

### Constantes:

$\delta_{ak}$	1 si el arco $a$ pertenece a la ruta $k$ , 0 en otro caso
$\delta_{a b}$	1 si el arco $a$ pertenece a la frontera $b$ , 0 en otro caso
$u_{ij}, l_{ij}$	cotas superiores e inferiores para el par O–D ( $i, j$ )
$u_i^O, l_i^O$	cotas superiores e inferiores para viajes generados por la zona $i$
$u_j^D, l_j^D$	cotas superiores e inferiores para viajes atraídos por la zona $j$
$u, l$	cotas superiores e inferiores para los viajes totales
$\bar{v} _b$	volumen observado en la frontera $b$
$\alpha _b$	factor de peso (opcional) otorgado a la frontera $b$

### Funciones:

$s_a(v_a)$	funciones de coste del arco $a$
------------	---------------------------------

**Variables:**

$v_a$  volumen en el arco  $a$

$h_k$  flujo en el camino  $k$

$p_k = \frac{h_k}{T_{ij}}$  proporción de viajes (del total  $T_{ij}$ ) que circula por el camino  $k$

$P_{ij,a} = \sum_{k \in K_{ij}} \delta_{ak} p_k$  proporción de viajes desde el nodo  $i$  al nodo  $j$  a través del arco  $a$ .

$v|_b$  volumen en la frontera  $b$

$T_{ij}$  demanda (viajes) desde el origen  $i$  al destino  $j$  (Destacar que es una variable para el proceso global de ajuste, pero constante en cada etapa de asignación)

Con esto, la formulación matemática que se propone es la siguiente.

## 1. Nivel Superior

$$\text{Minimizar}_{T_{ij}} f(T_{ij}) = \frac{1}{2} \sum_{b \in B} \alpha|_b (v|_b - \bar{v}|_b)^2 \quad (\text{a})$$

$$s.a. \quad l_{ij} \leq T_{ij} \leq u_{ij} \quad \forall i \in I, \forall j \in J \quad (\text{b})$$

$$l_i^O \leq \sum_{j \in J} T_{ij} \leq u_i^O \quad \forall i \in I \quad (\text{c}) \quad (22)$$

$$l_j^D \leq \sum_{i \in I} T_{ij} \leq u_j^D \quad \forall j \in J \quad (\text{d})$$

$$l \leq \sum_{i \in I} \sum_{j \in J} T_{ij} \leq u \quad (\text{e})$$

## 2. Nivel Inferior

$$\text{Minimizar}_{v_a} g(v_a) = \sum_{a \in A} \int_0^{v_a} s_a(v) dv$$

$$s.t. \quad v_a = \sum_{i \in I} \sum_{j \in J} \sum_{k \in K_{ij}} \delta_{ak} h_k, \quad \forall a \in A \quad (23)$$

$$\sum_{k \in K_{ij}} h_k = T_{ij}, \quad \forall i \in I, j \in J$$

$$h_k \geq 0 \quad \forall k \in K_{ij}, i \in I, j \in J$$

El modelo propuesto tiene una estructura binivel, donde el problema de nivel superior estima la matriz de viajes  $T_{ij}$  en función de un volumen  $v_a$  en los arcos dado. El problema de nivel inferior es el conocido como TAP (Traffic Assignment Problem), un problema de asignación que determina los volúmenes  $v_a$  en los arcos, según las condiciones de equilibrio de usuario, para una matriz de viajes  $T_{ij}$  dada (Beckmann *et al.*, 1956).

Respecto al problema de estimación de una matriz de viajes O–D expuesto en la ecuación (22), los volúmenes derivados de la telefonía móvil con los que realizar el ajuste,  $\bar{v}|_b$ , son obtenidos sobre fronteras  $b$  entre celdas de la red celular, cada una de las cuales comprende un grupo de arcos del modelo de red. La formulación que se propone en este trabajo incluye un término cuadrático para controlar la distancia entre los volúmenes observados y estimados (modelados). Sin embargo, para controlar la distorsión de la matriz estimada (o ajustada) respecto de la previa se opta por un conjunto de restricciones de variables acotadas (22) (b) y de restricciones funcionales (22) (c, d, e). Dichas restricciones deben definir unos rangos admisibles de variación para la información contenida en la matriz O–D en distintos niveles de agregación (para cada celda de la matriz O–D, viajes generados al agregar según destinos, viajes atraídos al agregar según orígenes, y viajes totales al agregar en orígenes y destinos, respectivamente). A diferencia de la formulación empleada en (21), la medida de la distorsión entre las matrices estimada y previa se ha modelado mediante dicho conjunto de restricciones en lugar de mediante una función  $F_l$ . Para obtener estrictamente la notación equivalente a (21) bastaría con definir la lagrangiana del problema y así elevar las restricciones (22) (b, c, d y e) a la función objetivo.

Respecto a los volúmenes  $v|_b$  de todas las fronteras observadas  $b \in B$ , es necesario realizar un cálculo adicional sobre los volúmenes  $v_a$  obtenidos del nivel inferior (23). De modo que el volumen en la frontera  $b$  puede ser expresado, en términos de proporciones de viajes, como la suma de los volúmenes  $v_a$  de cada uno de los arcos que cruzan la frontera  $b$ , obteniendo la siguiente expresión:

$$v|_b = \sum_{a \in b} v_a = \sum_{a \in A} \delta_{a|b} v_a = \sum_{a \in A} \delta_{a|b} \left( \sum_{i \in I} \sum_{j \in J} T_{ij} \sum_{k \in K_j} \delta_{ak} p_k \right) = \sum_{a \in A} \delta_{a|b} \left( \sum_{i,j} T_{ij} P_{ij,a} \right) \quad \forall b \in B \quad (24)$$

Si las funciones de coste son tales que  $\frac{ds_a(v_a)}{dv_a} > 0$  para cada arco  $a \in A$ , es decir, el tiempo necesario para recorrer un arco aumenta conforme aumenta el tráfico en él, está garantizada la

unicidad de solución con respecto a los volúmenes en arcos  $v_a$  (Sheffi, 1985); pudiéndose derivar una solución factible para cada  $v|_b$  conforme a la expresión (24).

El problema de la estimación de matrices O–D mediante volúmenes observados consiste en encontrar la distribución de viajes entre  $n_o \times n_d$  celdas ( $T_{ij}$ ) a partir de los aforos disponibles en la red, en este caso, fronteras observadas. En la mayor parte de los casos el número de incógnitas ( $n_o \times n_d$ ) es muy superior al de fronteras observadas, por lo que es imposible determinar una matriz O–D única, considerando el problema indeterminado. Esto significa que existen muchas matrices que satisfacen las ecuaciones del problema y el objetivo pasa a ser el encontrar la solución que mejor se adapte a determinados criterios pre–especificados. Para la presente metodología, dichos criterios son establecidos y modelados matemáticamente en el problema de nivel superior (22) mediante las restricciones (22) (b, c, d y e).

La exactitud de las matrices obtenidas mediante esta formulación se encuentra afectada por la precisión de los resultados de partida: volúmenes observados, viajes generados y atraídos por cada centroide, etc. En una red real, la falta de exactitud de estos datos hace que, en ocasiones, no se verifiquen las condiciones de equilibrio en la red modelada. Bajo esta hipótesis, es razonable permitir que las soluciones obtenidas en el proceso de la estimación no reproduzcan los datos exactamente; consecuentemente, el objetivo a perseguir será obtener la matriz O–D estimada que mejor reproduzca los datos observados tanto a nivel de volúmenes como de matriz previa.

Es necesario realizar algunos comentarios sobre el sistema de inecuaciones lineales mostrados en (22) (b, c, d y e). Dado que el objetivo de dichas restricciones es controlar la distorsión de la información contenida en la matriz estimada  $\mathbf{T}$  respecto de la matriz previa  $\bar{\mathbf{T}}$ , es evidente que las correspondientes cotas deberían ser definidas, bien en términos absolutos o relativos, conforme a valores derivados de la estimación inicial  $T_{ij}^{(0)}$ , lo cual garantizaría una región factible no–vacía. Ello se describe con más detalle en el apartado 5.3.4.3. Asimismo, merece la pena destacar que las cotas superiores e inferiores empleadas en (22) para cada restricción dependen del diseñador, es decir, no hay un valor universal para todos los casos sino valores específicos para cada problema dado. Estas consideraciones, junto a algunas orientaciones para la definición de dichas cotas, son parte de la práctica usual seguida en aplicaciones desarrolladas y testeadas en el campo de estimaciones de matrices O–D (Doblas y Benitez, 2005).

### 5.3.4 Algoritmo de ajuste

Respecto al problema de estimación de matrices, dadas las dimensiones de las redes reales y la capacidad de los ordenadores actuales, puede resultar inviable acometer el problema mediante un método que trabaje directamente con las restricciones funcionales y de variables acotadas propuestas en (22). Esto se debe, principalmente, al cálculo y manipulación de matrices de grandes dimensiones; por ejemplo, una red de transporte real de tamaño medio puede tener del orden de unos 500 centroides, cantidad que implica una matriz O–D de 250000 celdas.

Se pretende diseñar una metodología fácil de aplicar y adecuada para abordar problemas de estimación de matrices en aplicaciones reales de gran envergadura (elevado número de zonas y centroides). Por ello, la limitación antes mencionada ha motivado la utilización de un método de transformación (penalización) en el cual las restricciones funcionales se integran en una función objetivo equivalente, pudiéndose tratar directamente las restricciones de variables acotadas.

El planteamiento seguido en esta Tesis está basado en el método propuesto por Doblas y Benítez (2005) que combina el método de transformación denominado método del Lagrangiano aumentado (o de los multiplicadores) con el método de aproximación lineal de Frank–Wolfe. Éstos centraron el problema de estimación de matrices en el uso de conteos de tráfico en arcos preservando la estructura de una matriz previa conocida. Sin embargo, el planteamiento desarrollado en esta Tesis introduce una modificación para utilizar en el proceso de ajuste volúmenes agregados en grupos de arcos en lugar de volúmenes en arcos individuales (aforos tradicionales). Esta modificación es esencial para la utilización de información derivada de la telefonía móvil como volúmenes observados, debido a que estos sistemas proporcionan información agregada a nivel de fronteras entre celdas (grupos de arcos que la cruzan) y no a nivel individual de arco. No obstante, el modelo formulado puede ser usado conjuntamente con los tradicionales conteos de tráfico dado que los aforos son un caso especial en el que el número de arcos del grupo es uno. Por ello, es necesario resaltar que la metodología propuesta no excluye el uso de información procedente tanto de aforos tradicionales como de otras fuentes, aunque por razones de claridad se ha dirigido el problema de estimación exclusivamente a información de volúmenes observados derivada de teléfonos móviles.

Esta sección se dedica a introducir la metodología propuesta por Doblas y Benitez (2005), así como a presentar y discutir las modificaciones necesarias que afectan al algoritmo de resolución del problema de optimización propuesto para utilizar los datos derivados de teléfonos móviles.

### 5.3.4.1 Introducción

A continuación se presenta el método combinado de la Función Lagrangiana Aumentada con el algoritmo de Frank–Wolfe. Una descripción más detallada sobre el proceso de optimización matemática puede ser encontrada en Doblas y Benitez (2005).

#### ■ Método de la Función Lagrangiana Aumentada o de los Multiplicadores

El problema (22) puede formularse como un problema genérico de minimización con restricciones lineales  $g_l(\mathbf{x})$  y variables acotadas  $[\mathbf{l}, \mathbf{u}]$ :

$$\begin{aligned} & \underset{\mathbf{x}}{\text{Minimizar}} && f(\mathbf{x}) \\ & \text{s.a.} && g_l(\mathbf{x}) \geq 0, \quad \forall l=1,2,\dots,L \\ & && \mathbf{l} \leq \mathbf{x} \leq \mathbf{u} \\ & && \mathbf{x} \in \mathbb{R}^N \end{aligned} \tag{25}$$

donde  $g_l(\mathbf{x})$  serían las restricciones funcionales definidas en (22) (c, d y e), y  $\mathbf{l} \leq \mathbf{x} \leq \mathbf{u}$  las restricciones de variables acotadas definidas en (22) (b). Para tal problema, se sabe que las condiciones de Kuhn–Tucker son necesarias y suficientes para caracterizar el óptimo de dicho problema. El Método de la Función Lagrangiana Aumentada o Método de los Multiplicadores (Reklaitis *et al.*, 1983) consiste, básicamente, en añadir una serie de términos cuadráticos a la función clásica de Lagrange  $f(\mathbf{x})$ , formando una función sin restricciones cuyo mínimo es un punto que cumple las condiciones de Kuhn–Tucker del problema original. Sin embargo, es necesario indicar que la función Lagrangiana Aumentada conseguida en nuestro caso no incluirá las restricciones de variables acotadas. De modo que la función lagrangiana aumentada queda expresada como:

$$L_{aug}(\mathbf{x}, \boldsymbol{\sigma}) = f(\mathbf{x}) + \sum_{l=1}^L \omega_l \left( \langle g_l(\mathbf{x}) + \sigma_l \rangle^2 - \sigma_l^2 \right) \tag{26}$$

donde  $\boldsymbol{\sigma} = (\sigma_l) \in \mathbb{R}^L$ , siendo  $\sigma_l$  y  $\omega_l$  el “multiplicador” y el factor de peso (o escala) asociado a la restricción  $l$ -ésima, respectivamente. Cabe destacar que el empleo de la función Lagrangiana Aumentada para resolver el problema puede ser visto, en general, como una optimización sin restricciones pero, debido a la estructura de este problema, se consigue conservar las  $n_o \times n_d$  restricciones de las variables acotadas del problema original. El término correspondiente a las restricciones  $\sigma_l$  se encarga de penalizar las violaciones de las restricciones que se cometan durante el proceso de optimización. La aplicación de este método a problemas del tipo (25) garantiza la convergencia hacia un punto que cumple las condiciones de optimalidad, siendo los

términos de penalización y los multiplicadores (actualizados según una sencilla regla) los que fuerzan la convergencia hacia una *solución factible óptima*.

El procedimiento consiste en optimizar (minimizar) en  $x$  la función lagrangiana aumentada (26), siendo el término  $\omega_l$  constante de una etapa a otra; mientras que los multiplicadores  $\sigma_l$  sólo permanecen constantes durante cada etapa o subproblema. En cada etapa se hallan los  $x$  óptimos para esos multiplicadores de modo que, una vez obtenido el óptimo para esa etapa del proceso, se procede a actualizar los multiplicadores conforme a la regla  $\sigma_j^{(n+1)} = \langle g_j(x^{(n)}) + \sigma_j^{(n)} \rangle$  para acometer la siguiente etapa ( $n+1$ ) del proceso de optimización. Esto se repite hasta alcanzar un grado de convergencia razonable. El operador  $\langle \cdot \rangle$  representa:

$$\langle \alpha \rangle = \begin{cases} \alpha & \text{si } \alpha < 0 \\ 0 & \text{si } \alpha \geq 0 \end{cases} \quad (27)$$

En cada etapa o subproblema, es decir, en cada resolución aproximada de las sucesivas lagrangianas aumentadas, el método empleado es el de Frank–Wolfe. Este método consiste en hallar una dirección auxiliar resolviendo el problema (26) linealizado sin salirse de la región factible. Esta dirección auxiliar determina una dirección de descenso que permite hallar un nuevo estimado que mejore el valor de la función de lagrange aumentada.

■ *Método de Frank–Wolfe*

Centrando la atención en el algoritmo de Frank–Wolfe (F–W) presentado en términos de la función Lagrangiana Aumentada genérica sujeta a restricciones de variables acotadas, el problema que pretende ser optimizado mediante el método de aproximación lineal de F–W es:

$$\begin{aligned} &\underset{\mathbf{x}}{\text{Minimizar}} && L_{aug}(\mathbf{x}) \\ &\text{s.a.} && \mathbf{l} \leq \mathbf{x} \leq \mathbf{u} \\ &&& \mathbf{x} \in \mathbb{R}^N \end{aligned} \quad (28)$$

para un vector constante de multiplicadores  $\sigma$  asociados a esta etapa del proceso. Se denota  $m$  al contador para identificar las sucesivas estimaciones obtenidas por el algoritmo. El método de F–W se basa en aproximar linealmente la función objetivo en torno al punto actual  $\mathbf{x}^{(m)}$ . A continuación, se busca el punto  $\mathbf{y}$  de la región factible que minimice la función objetivo linealizada, que se traduce en el problema de optimización lineal definido como:

$$\begin{aligned} &\underset{\mathbf{y}}{\text{Minimizar}} && \nabla L_{aug}(\mathbf{x}^{(m)})^T \mathbf{y} && \text{(a)} \\ &\text{s.a.} && \mathbf{l} \leq \mathbf{y} \leq \mathbf{u} && \text{(b)} \\ &&& \mathbf{y} \in \mathbb{R}^N && \text{(c)} \end{aligned} \quad (29)$$

Entonces, con la solución en la  $m$ -ésima iteración del problema linealizado,  $\mathbf{y}_{ij}^{(m)}$ , se obtiene un nuevo estimado de (28) mediante la siguiente regla de actualización:

$$\mathbf{x}_{ij}^{(m+1)} = \mathbf{x}_{ij}^{(m)} + \lambda^{(m)} (\mathbf{y}_{ij}^{(m)} - \mathbf{x}_{ij}^{(m)}) = \mathbf{x}_{ij}^{(m)} + \lambda^{(m)} \mathbf{d}_{ij}^{(m)} \quad (30)$$

Esta expresión está escrita en términos de la dirección de descenso en el estimado actual,  $\mathbf{d}^{(m)}$ , la cual es la diferencia entre el punto extremo auxiliar  $\mathbf{y}^{(m)}$  y el estimado actual  $\mathbf{x}^{(m)}$ . El paso  $\lambda^{(m)}$  se obtiene como solución del siguiente problema de optimización unidimensional:

$$\begin{aligned} \text{Minimizar}_{\lambda} \quad & L_{aug} (\mathbf{x}^{(m+1)} = \mathbf{x}^{(m)} + \lambda \mathbf{d}^{(m)}) \\ \text{s.a.} \quad & 0 \leq \lambda \leq 1 \end{aligned} \quad (31)$$

donde  $\lambda$  pertenece al intervalo  $[0, 1]$  puesto que el punto auxiliar  $\mathbf{y}$  está localizado en un vértice de la región factible definida por el conjunto poliédrico definido por (29).

Por tanto, según el método de la Función Lagrangiana Aumentada, las restricciones funcionales (22) (c, d y e) se integran en una función objetivo equivalente, mientras que las restricciones de variables acotadas (22) (b) son tratadas directamente con el método de direcciones factibles (Frank–Wolfe). Con ello se consigue que las restricciones de variables acotadas se cumplan en todo momento, mientras que el cumplimiento de las restricciones funcionales es producto de un proceso iterativo que va penalizando las violaciones cometidas por medio de términos que aumentan la función de Lagrange clásica.

#### 5.3.4.2 Implementación

A continuación, se introduce una notación adicional a la presentada en la sección 5.3.3 para reescribir el método combinado de la función Lagrangiana Aumentada con el algoritmo de Frank–Wolfe en términos de magnitudes asociadas al problema de estimación de matrices O–D.

#### Constantes:

$\omega_i^O$  factor de peso asociado a las restricciones de viajes generados por la zona  $i$

$\omega_j^D$  factor de peso asociado a las restricciones de viajes atraídos por la zona  $j$

$\omega$  factor de peso asociado a las restricciones de viajes totales



**Variables:**

- $\sigma_i^O$     multiplicador asociado a la cota superior de la restricción de viajes generados por la zona  $i$   
 $\tau_i^O$     multiplicador asociado a la cota inferior de la restricción de viajes generados por la zona  $i$   
 $\sigma_j^D$     multiplicador asociado a la cota superior de la restricción de viajes atraídos por la zona  $j$   
 $\tau_j^D$     multiplicador asociado a la cota inferior de la restricción de viajes atraídos por la zona  $j$   
 $\sigma$     multiplicador asociado a la cota superior de la restricción de viajes totales  
 $\tau$     multiplicador asociado a la cota inferior de la restricción de viajes totales  
 $y_{ij}$     punto extremo auxiliar del algoritmo de Frank–Wolfe  
 $d_{ij}$     dirección de descenso del algoritmo de Frank–Wolfe  
<sup>(m)</sup>     $m$ -ésima iteración

Con la anterior notación, la función Lagrangiana Aumentada (26) puede ser escrita como:

$$\begin{aligned}
 L_{aug}(\mathbf{T}, \boldsymbol{\sigma}, \boldsymbol{\tau}) = & \frac{1}{2} \sum_{b \in B} \sum_{i \in I} \sum_{j \in J} \alpha|_b \left( \sum_{a \in b} \left( T_{ij} \sum_{k \in K_{ij}} \delta_{ak} p_k \right) - \bar{v}|_b \right)^2 \\
 & + \frac{1}{2} \left( \sum_{i=1}^{n_o} \omega_i^O \left( \left\langle u_i^O - \sum_{j=1}^{n_d} T_{ij} + \sigma_i^O \right\rangle^2 - (\sigma_i^O)^2 \right) + \sum_{i=1}^{n_o} \omega_i^O \left( \left\langle \sum_{j=1}^{n_d} T_{ij} - l_i^O + \tau_i^O \right\rangle^2 - (\tau_i^O)^2 \right) \right) \\
 & + \frac{1}{2} \left( \sum_{j=1}^{n_d} \omega_j^D \left( \left\langle u_j^D - \sum_{i=1}^{n_o} T_{ij} + \sigma_j^D \right\rangle^2 - (\sigma_j^D)^2 \right) + \sum_{j=1}^{n_d} \omega_j^D \left( \left\langle \sum_{i=1}^{n_o} T_{ij} - l_j^D + \tau_j^D \right\rangle^2 - (\tau_j^D)^2 \right) \right) \\
 & + \frac{1}{2} \omega \left( \left( \left\langle u - \sum_{i=1}^{n_o} \sum_{j=1}^{n_d} T_{ij} + \sigma \right\rangle^2 - \sigma^2 \right) + \left( \left\langle \sum_{i=1}^{n_o} \sum_{j=1}^{n_d} T_{ij} - l + \tau \right\rangle^2 - \tau^2 \right) \right)
 \end{aligned} \tag{32}$$

donde  $\mathbf{T} = [T_{ij}]$ , siendo  $\boldsymbol{\sigma} = \{\sigma^O, \sigma^D\}$ , y  $\boldsymbol{\tau} = \{\tau^O, \tau^D\}$  vectores de multiplicadores asociados a cotas superiores e inferiores de las restricciones, respectivamente. En cada etapa o subproblema  $(m)$  los multiplicadores permanecen constantes, la variable de decisión es la matriz  $T_{ij}$ . Por lo que el problema de optimización se define como:

$$\begin{aligned}
\text{Min}_{T_{ij}} L_{aug}(T_{ij}) &= \frac{1}{2} \sum_{b \in B} \sum_{i \in I} \sum_{j \in J} \alpha|_b \left( \sum_{a \in b} \left( T_{ij} \sum_{k \in K_{ij}} \delta_{ak} p_k \right) - \bar{v}|_b \right)^2 \\
&+ \frac{1}{2} \left( \sum_{i=1}^{n_o} \omega_i^O \left( \left\langle u_i^O - \sum_{j=1}^{n_d} T_{ij} + \sigma_i^O \right\rangle^2 - (\sigma_i^O)^2 \right) + \sum_{i=1}^{n_o} \omega_i^O \left( \left\langle \sum_{j=1}^{n_d} T_{ij} - l_i^O + \tau_i^O \right\rangle^2 - (\tau_i^O)^2 \right) \right) \\
&+ \frac{1}{2} \left( \sum_{j=1}^{n_d} \omega_j^D \left( \left\langle u_j^D - \sum_{i=1}^{n_o} T_{ij} + \sigma_j^D \right\rangle^2 - (\sigma_j^D)^2 \right) + \sum_{j=1}^{n_d} \omega_j^D \left( \left\langle \sum_{i=1}^{n_o} T_{ij} - l_j^D + \tau_j^D \right\rangle^2 - (\tau_j^D)^2 \right) \right) \quad (33) \\
&+ \frac{1}{2} \omega \left( \left( \left\langle u - \sum_{i=1}^{n_o} \sum_{j=1}^{n_d} T_{ij} + \sigma \right\rangle^2 - \sigma^2 \right) + \left( \left\langle \sum_{i=1}^{n_o} \sum_{j=1}^{n_d} T_{ij} - l + \tau \right\rangle^2 - \tau^2 \right) \right)
\end{aligned}$$

$$\text{s.a. } l_{ij} \leq T_{ij} \leq u_{ij}$$

Al finalizar cada etapa ( $m$ ) de optimización, los multiplicadores para el nuevo subproblema ( $m+1$ ) se actualizan en función del estimado óptimo  $T_{ij}^{(m)}$  y de los multiplicadores  $\sigma$  y  $\tau$  correspondientes a la etapa ( $m$ ), de acuerdo con la regla (34). El proceso se da por concluido cuando se alcanza un grado de convergencia que se considere razonable.

$$\begin{aligned}
\sigma_i^{O(m+1)} &= \left\langle u_i^O - \sum_{j=1}^{n_d} T_{ij}^{(m)} + \sigma_i^{O(m)} \right\rangle & \tau_i^{O(m+1)} &= \left\langle \sum_{j=1}^{n_d} T_{ij}^{(m)} - l_i^O + \tau_i^{O(m)} \right\rangle, & \forall i \in I \\
\sigma_j^{D(m+1)} &= \left\langle u_j^D - \sum_{i=1}^{n_o} T_{ij}^{(m)} + \sigma_j^{D(m)} \right\rangle & \tau_j^{D(m+1)} &= \left\langle \sum_{i=1}^{n_o} T_{ij}^{(m)} - l_j^D + \tau_j^{D(m)} \right\rangle, & \forall j \in J \\
\sigma^{(m+1)} &= \left\langle u - \sum_{i=1}^{n_o} \sum_{j=1}^{n_d} T_{ij}^{(m)} + \sigma^{(m)} \right\rangle & \tau^{(m+1)} &= \left\langle \sum_{i=1}^{n_o} \sum_{j=1}^{n_d} T_{ij}^{(m)} - l + \tau^{(m)} \right\rangle
\end{aligned} \quad (34)$$

Para minimizar la función Lagrangiana Aumentada mediante el algoritmo de Frank–Wolfe es necesario calcular su gradiente, resolver el problema linealizado para hallar el punto extremo auxiliar y la dirección de descenso, y obtener el paso óptimo. Teniendo en cuenta la ecuación (24) de los volúmenes en fronteras  $v|_b$  expresada en términos de  $T_{ij}$  y de proporciones  $p_k$ , de caminos o rutas usadas por los usuarios, es posible introducir una hipótesis que permita obtener una aproximación de la derivada analítica y así obtener  $\nabla L_{aug}(\mathbf{T})$  fácilmente. Dicha hipótesis consiste en considerar constantes las proporciones de viaje  $p_k$  correspondientes a la matriz O–D actual,  $T^{(m)}$ , al producirse un cambio infinitesimal en los viajes de una única casilla de la matriz, conservando el resto constantes. De modo que, las derivadas parciales de los volúmenes de frontera  $v|_b$  con respecto a las variables  $T_{ij}$  se definen como:

$$\begin{aligned} \frac{\partial v_a}{\partial T_{ij}} &= \sum_{k \in K_{ij}} \delta_{ak} p_k \\ v|_b &= \sum_{a \in A} \delta_{a|b} v_a \end{aligned} \Rightarrow \frac{\partial v|_b}{\partial T_{ij}} = \sum_{a \in A} \delta_{a|b} \frac{\partial v_a}{\partial T_{ij}} = \sum_{a \in A} \delta_{a|b} \sum_{k \in K_{ij}} \delta_{ak} p_k \quad (35)$$

Con la expresión anterior se puede obtener fácilmente la expresión del gradiente de la Lagrangiana Aumentada, que viene dada por:

$$\begin{aligned} \frac{\partial L_{aug}(\mathbf{T})}{\partial T_{ij}} &= \sum_{b \in B} \alpha|_b \left[ \left( \sum_{a \in b} \left( T_{ij} \sum_{k \in K_{ij}} \delta_{ak} p_k \right) - \bar{v}|_b \right) \cdot \sum_{a \in b} \sum_{k \in K_{ij}} \delta_{ak} p_k \right] \\ &+ \omega_i^O \left( \left\langle \sum_{j=1}^{n_d} T_{ij} - l_i^O + \tau_i^O \right\rangle - \left\langle u_i^O - \sum_{j=1}^{n_d} T_{ij} + \sigma_i^O \right\rangle \right) \\ &+ \omega_j^D \left( \left\langle \sum_{i=1}^{n_o} T_{ij} - l_j^D + \tau_j^D \right\rangle - \left\langle u_j^D - \sum_{i=1}^{n_o} T_{ij} + \sigma_j^D \right\rangle \right) \\ &+ \omega \left( \left\langle \sum_{i=1}^{n_o} \sum_{j=1}^{n_d} T_{ij} - l + \tau \right\rangle - \left\langle u - \sum_{i=1}^{n_o} \sum_{j=1}^{n_d} T_{ij} + \sigma \right\rangle \right) \quad \forall i \in I, j \in J \end{aligned} \quad (36)$$

Una vez calculado el gradiente, se procede a obtener el punto extremo  $y_{ij}$  del problema linealizado (29) que, adaptado al problema de ajuste de la matriz  $T_{ij}$ , se convierte en:

$$\begin{aligned} \text{Minimizar}_{y_{ij}} \quad & \sum_{i=1}^{n_o} \sum_{j=1}^{n_d} \frac{\partial L_{aug}}{\partial T_{ij}} \Big|_{T_{ij}^m} \cdot y_{ij} & (a) \\ \text{s.a.} \quad & l_{ij} \leq y_{ij} \leq u_{ij} & (b) \\ & y_{ij} \in \mathbb{R}^N & (c) \end{aligned} \quad (37)$$

Debido a la estructura especial del problema, este punto extremo  $y_{ij}^{(m)}$  se puede hallar de forma sencilla, sin necesidad de utilizar el método simplex, mediante la siguiente expresión lógica:

$$y_{ij}^{(m)} = \left( \frac{\partial L_{aug}}{\partial T_{ij}} \Big|_{T_{ij}^{(m)}} > 0 \right) \cdot l_{ij} + \left( \frac{\partial L_{aug}}{\partial T_{ij}} \Big|_{T_{ij}^{(m)}} = 0 \right) \cdot T_{ij}^{(m)} + \left( \frac{\partial L_{aug}}{\partial T_{ij}} \Big|_{T_{ij}^{(m)}} < 0 \right) \cdot u_{ij} \quad \forall i \in I, j \in J \quad (38)$$

Las sucesivas estimaciones del algoritmo de F-W son computadas con la siguiente adaptación de la expresión (30) conforme a la notación de matrices O-D:

$$T_{ij}^{(m+1)} = T_{ij}^{(m)} + \lambda^{(m)} \left( y_{ij}^{(m)} - T_{ij}^{(m)} \right) = T_{ij}^{(m)} + \lambda^{(m)} d_{ij}^{(m)}, \quad \text{donde: } \lambda^{(m)} \in \mathbb{R} \quad (39)$$

para cada  $i \in I, j \in J$ , donde  $d_{ij}^{(m)} = y_{ij}^{(m)} - T_{ij}^{(m)}$  es la dirección de descenso F-W.

Finalmente, el cálculo del paso óptimo  $\lambda$  será la solución del problema unidimensional de optimización (31) adaptado a nuestro problema en notación O–D. Para ello es necesario formular la función Lagrangiana aumentada (32) para el nuevo estimado  $(m+1)$  expresándola en función del paso  $\lambda$ , es decir, evaluarla para  $T_{ij}^{(m+1)} = T_{ij}^{(m)} + \lambda^{(m)} d_{ij}^{(m)}$ , donde  $T_{ij}^{(m)}$  y  $d_{ij}^{(m)}$  son conocidos; consiguiendo una función unidimensional  $L_{aug}(\lambda)$ :

$$\begin{aligned}
L_{aug}(\lambda) = & \frac{1}{2} \sum_{b \in B} \sum_{i \in I} \sum_{j \in J} \alpha_b \left[ \sum_{a \in b} \left( T_{ij}^{(m)} \sum_{k \in K_{ij}} \delta_{ak} p_k \right) + \lambda \sum_{a \in b} \left( d_{ij}^{(m)} \sum_{k \in K_{ij}} \delta_{ak} p_k \right) - \bar{v}_b \right]^2 \\
& + \frac{1}{2} \sum_{i \in I} \omega_i^O \left( \left\langle u_i^O - \sum_{j \in J} T_{ij}^{(m)} - \lambda \sum_{j \in J} d_{ij}^{(m)} + \sigma_i^O \right\rangle^2 - (\sigma_i^O)^2 + \left\langle \sum_{j \in J} T_{ij}^{(m)} + \lambda \sum_{j \in J} d_{ij}^{(m)} - l_i^O + \tau_i^O \right\rangle^2 - (\tau_i^O)^2 \right) \\
& + \frac{1}{2} \sum_{j \in J} \omega_j^D \left( \left\langle u_j^D - \sum_{i \in I} T_{ij}^{(m)} - \lambda \sum_{i \in I} d_{ij}^{(m)} + \sigma_j^D \right\rangle^2 - (\sigma_j^D)^2 + \left\langle \sum_{i \in I} T_{ij}^{(m)} + \lambda \sum_{i \in I} d_{ij}^{(m)} - l_j^D + \tau_j^D \right\rangle^2 - (\tau_j^D)^2 \right) \\
& + \frac{1}{2} \omega \left( \left\langle u - \sum_{i \in I} \sum_{j \in J} T_{ij}^{(m)} - \lambda \sum_{i \in I} \sum_{j \in J} d_{ij}^{(m)} + \sigma \right\rangle^2 - \sigma^2 + \left\langle \sum_{i \in I} \sum_{j \in J} T_{ij}^{(m)} + \lambda \sum_{i \in I} \sum_{j \in J} d_{ij}^{(m)} - l + \tau \right\rangle^2 - \tau^2 \right)
\end{aligned} \tag{40}$$

El paso óptimo será la solución de un problema unidimensional de minimización de la función  $L_{aug}(\lambda)$  donde  $\lambda \in [0, 1]$  (31). Cabe destacar que la magnitud del desplazamiento permitido de  $T_{ij}^{(m)}$  hasta  $y_{ij}^{(m)}$  impide utilizar expresiones analíticas, como la usada en Spiess (1990), para el cálculo del paso ya que el cambio no es infinitesimal y no puede admitirse la hipótesis de que las proporciones permanecen constantes. Junto a ello, la existencia del operador  $\langle \cdot \rangle$ , definido en (27), condiciona la elección de un algoritmo robusto de búsqueda que, sin hacer uso del cálculo de derivadas, se encargue de minimizar una función unidimensional sobre un intervalo  $[a, b]$ , tales como búsqueda dicotómica, exhaustiva, sección dorada, etc. En el algoritmo de Doblas y Benítez (2005) se utilizó el método de bisección pero, tras considerar la dificultad de usar derivadas y expresiones analíticas, se ha optado por el algoritmo denominado búsqueda por la “Sección Dorada”. Además de no usar información de la derivada de la función, este eficiente algoritmo tiene como ventaja el ser óptimo con respecto al número de evaluaciones de la función objetivo, conocido el intervalo de incertidumbre, el de definición de la función y sin necesidad de predefinir el número de iteraciones (i.e.: Fibonacci). Este hecho, teniendo en cuenta las dimensiones de las matrices manejadas, simplifica en gran medida los cálculos necesarios; de ahí su elección como método de resolución. El algoritmo se basa en la búsqueda del mínimo de una función unimodal, en este caso la función definida en (40), mediante reducciones iterativas del rango de búsqueda de valores de  $\lambda$  dentro del cual el óptimo se sabe que existe (Bazaraa y Shetty, 1979). El algoritmo garantiza que cada nueva evaluación de la función acotará el mínimo, tras aplicar los ratios correspondientes, dentro de un intervalo

0.61803 (razón áurea) veces más pequeño que el intervalo de búsqueda de la anterior iteración. En nuestro problema, el valor óptimo de  $\lambda$  se encuentra aplicando el algoritmo iterativamente dentro del intervalo inicial  $[0, 1]$  en la dirección  $d_{ij}^{(m)}$ . De modo que una vez encontrado el paso óptimo, se procede a la actualización de la matriz O–D para la siguiente iteración.

Para la finalización del algoritmo de optimización, la aproximación relativa a las proporciones de rutas  $p_k$  implica cierto zigzagueo en los valores de la función objetivo entre estimaciones sucesivas, lo cual dificulta el poder incorporar un criterio de finalización basado en valores de las derivadas, en valores de la función objetivo en dos pasos consecutivos, o en la diferencia relativa entre dos estimados sucesivos. Por ello, el criterio de terminación del algoritmo se basa en i) un número máximo de subproblemas (o etapas Lagrangianas), y ii) un número máximo de iteraciones por subproblema. Es necesario especificar que no hay reglas específicas para elegir este número máximo de subproblemas e iteraciones, al igual que ocurría con las cotas superiores e inferiores definidas en (22), sino que estos valores dependen del criterio del planificador y su conocimiento práctico. Por último, indicar que una característica de este algoritmo es que la matriz O–D, para la cual se obtiene el mejor valor de la función objetivo dentro las diferentes iteraciones de una etapa o subproblema concreto, es almacenada y usada como punto de partida para el siguiente subproblema.

### 5.3.4.3 Especificaciones del modelo

Por razones de simplificación de las ecuaciones, los valores de los factores de peso opcionales  $\alpha|_b$  serán fijados a 1 para todas las fronteras de los casos prácticos que en los siguientes apartados se evalúan. Asimismo, se ha puesto especial interés en la elección de los factores de peso  $(\omega_i^O, \omega_j^D, \omega)$  asociados con las restricciones funcionales integradas dentro de la función Lagrangiana aumentada (32), y sobre el comportamiento del algoritmo desarrollado para niveles de congestión de la red (cuantificada por el número de iteraciones necesarias para conseguir una solución en equilibrio).

El extenso rango o intervalo de valores en el cual se encuentran las magnitudes implicadas en el problema de estimación de matrices O–D, las cuales son viajes generados, atraídos y totales, insta hacia un proceso natural de escalado de las restricciones basado en los valores iniciales de dichas magnitudes, es decir, respecto al estimado inicial  $T_{ij}^{(0)}$  a diferentes niveles de agregación. Por tanto, cada restricción funcional (22) (c, d y e) es normalizada (o escalada) por su correspondiente valor inicial, resultando un conjunto de restricciones normalizadas que dan lugar a penalizaciones de violaciones de las restricciones proporcionales a valores relativos de las correspondientes magnitudes. En otras palabras, no posee la misma importancia exceder en

un viaje una cota superior de 10 que una cota superior de  $10^4$ , estando más penalizada la violación de la cota superior en la primera situación que en la última.

El proceso de escalado de restricciones se presenta esquemáticamente para una pareja genérica de restricciones de viajes generados, lo cual obliga a modificar levemente la notación matemática para mostrar claramente la relación directa en la expresión (32). En primer lugar, se formulan las restricciones y la pareja correspondiente de multiplicadores como se indica:

$$\left. \begin{aligned} \frac{l_i^O}{O_i^{(0)}} \leq \frac{O_i = \sum_{j \in J} T_{ij}}{O_i^{(0)}} \leq \frac{u_i^O}{O_i^{(0)}} \\ O_i^{(0)} = \sum_{j \in J} T_{ij}^{(0)} \end{aligned} \right\} \Rightarrow \begin{aligned} \frac{u_i^O - \sum_{j \in J} T_{ij}}{O_i^{(0)}} \geq 0 \rightarrow \sigma_i'^O \\ \frac{\sum_{j \in J} T_{ij} - l_i^O}{O_i^{(0)}} \geq 0 \rightarrow \tau_i'^O \end{aligned} \quad (41)$$

para cada  $i \in I$  con  $O_i^{(0)} \neq 0$ , donde  $T_{ij}^{(0)}$  representa la estimación actual, que en nuestro caso coincide con la matriz de partida o matriz O-D previa. Las restricciones impuestas sobre los viajes atraídos y los viajes totales deberán ser modificada de forma análoga usando las magnitudes  $D_j^{(0)} = \sum_{i \in I} T_{ij}^{(0)}$  para cada  $j \in J$  con  $D_j^{(0)} \neq 0$  y  $T^{(0)} = \sum_{i \in I} \sum_{j \in J} T_{ij}^{(0)}$ .

$$\begin{aligned} \frac{l_i^O}{O_i^{(0)}} \leq \frac{\sum_j T_{ij}}{O_i^{(0)}} \leq \frac{u_i^O}{O_i^{(0)}} \\ \frac{l_j^D}{D_j^{(0)}} \leq \frac{\sum_i T_{ij}}{D_j^{(0)}} \leq \frac{u_j^D}{D_j^{(0)}} \\ \frac{l}{T^{(0)}} \leq \frac{\sum_i \sum_j T_{ij}}{T^{(0)}} \leq \frac{u}{T^{(0)}} \end{aligned} \quad (42)$$

En base al razonamiento anterior, dado que cada restricción está escalada (normalizada) respecto a su magnitud correspondiente –  $O_i^{(0)}$ ,  $D_j^{(0)}$  o  $T^{(0)}$  –, el término cuadrático de penalización de la función lagrangiana aumentada es re-formulado introduciendo un factor de peso global  $W$  común a todas ellas ( $\omega_i^O$ ,  $\omega_j^D$ ,  $\omega$ ). Este factor de peso  $W$  se refiere a la ponderación del término de restricciones de la Lagrangiana Aumentada (32) con respecto a los términos de la función objetivo original:

$$W \sum_{i \in I} \left( \left\langle \frac{u_i^O - \sum_{j \in J} T_{ij}}{O_i^{(0)}} + \sigma_i'^O \right\rangle^2 - (\sigma_i'^O)^2 + \left\langle \frac{\sum_{j \in J} T_{ij} - l_i^O}{O_i^{(0)}} + \tau_i'^O \right\rangle^2 - (\tau_i'^O)^2 \right) \quad (43)$$

donde las siguientes manipulaciones llevan hacia la definición original dada en (32):

$$\begin{aligned}
 &= W \sum_{i \in I} \left( \frac{1}{O_i^{(0)}} \right)^2 \left( \left\langle u_i^O - \sum_{j \in J} T_{ij} + \sigma_i'^O O_i^{(0)} \right\rangle^2 - (\sigma_i'^O O_i^{(0)})^2 + \left\langle \sum_{j \in J} T_{ij} - l_i^O + \tau_i'^O O_i^{(0)} \right\rangle^2 - (\tau_i'^O O_i^{(0)})^2 \right) \\
 &= \left\{ \begin{array}{l} \sigma_i^O = \sigma_i'^O O_i^{(0)} \\ \tau_i^O = \tau_i'^O O_i^{(0)} \\ \omega_i^O = \frac{2W}{(O_i^{(0)})^2} \end{array} \right\} = \frac{1}{2} \sum_{i \in I} \omega_i^O \left( \left\langle u_i^O - \sum_{j \in J} T_{ij} + \sigma_i^O \right\rangle^2 - (\sigma_i^O)^2 + \left\langle \sum_{j \in J} T_{ij} - l_i^O + \tau_i^O \right\rangle^2 - (\tau_i^O)^2 \right) \quad (44)
 \end{aligned}$$

Aquellos casos en los cuales  $O_i^{(0)} = 0$  o  $D_j^{(0)} = 0$ , serán tratados individualmente conforme a sus cotas superiores. La tarea reside ahora en la elección de un valor apropiado para  $W$ . Es conveniente usar valores suficientemente grandes a fin de no partir de una solución inicial con múltiples restricciones violadas; de lo contrario el proceso reajustado a través de la penalización del término cuadrático de la función Lagrangiana aumentada y las sucesivas actualizaciones de los multiplicadores puede no comportarse correctamente. Por otro lado, el mencionado parámetro  $W$ , aunque suficientemente grande, no debería ser mayor que un valor capaz de bloquear el proceso de optimización, ya que valores “elevados” de  $W$  pueden impedir la evolución del algoritmo hacia estimados que mejoren el valor de la función objetivo. Sin embargo, valores “pequeños” pueden dar lugar a una relajación excesiva en el cumplimiento de las restricciones que se pueda hacer irreversible. Por ello, Doblas y Benítez (2005) proporcionan una regla simple avalada por un buen comportamiento en trabajos reales de ajuste de matrices, la cual consiste en la elección de valores de  $W$  del mismo orden de magnitud que el valor inicial de la función objetivo definida en (22) (a).

De forma análoga a lo sucedido con las restricciones, el término de la función Lagrangiana aumentada (32) asociado con los volúmenes de frontera debe ser modificado mediante una normalización respecto del volumen observado en cada frontera,  $\bar{v}|_b$ , quedando:

$$L_{aug}(\mathbf{T}, \boldsymbol{\sigma}, \boldsymbol{\tau}) = \frac{1}{2} \sum_{b \in B} \sum_{i \in I} \sum_{j \in J} \alpha|_b \left( \frac{\sum_{a \in b} \left( T_{ij} \sum_{k \in K_j} \delta_{ak} p_k \right) - \bar{v}|_b}{\bar{v}|_b} \right)^2 + \dots \quad (45)$$

No obstante, por cuestiones de sencillez en la formulación matemática así como de tamaño de las ecuaciones, esta normalización no es introducida en la formulación original. El hecho de presentar un modelo que incluya esta normalización puede complicar la comprensión del modelo por lo que se ha optado por no reformular las expresiones en función de dichas normalizaciones.

## 5.4 CASOS PRÁCTICOS

### 5.4.1 Experimento 1: Red Sioux–Falls

#### 5.4.1.1 Configuración de la red

Se ha aplicado el procedimiento de estimación descrito en la sección anterior a la red de Sioux–Falls. Esta red consta de 24 nodos, los cuales son todos centroides, 528 pares O–D y 76 arcos dirigidos. Esta red de transporte originalmente fue descrita por LeBlanc (1975) y posteriormente usada por múltiples autores, como Suwansirikul *et al.* (1987), del cual se han tomado los atributos de arcos para este ensayo. Es necesario resaltar que, para este modelo de red, los atributos son idénticos para el arco dirigido desde el nodo A al nodo B y el asociado al sentido contrario, aunque esto no suele ocurrir en el caso de redes reales.

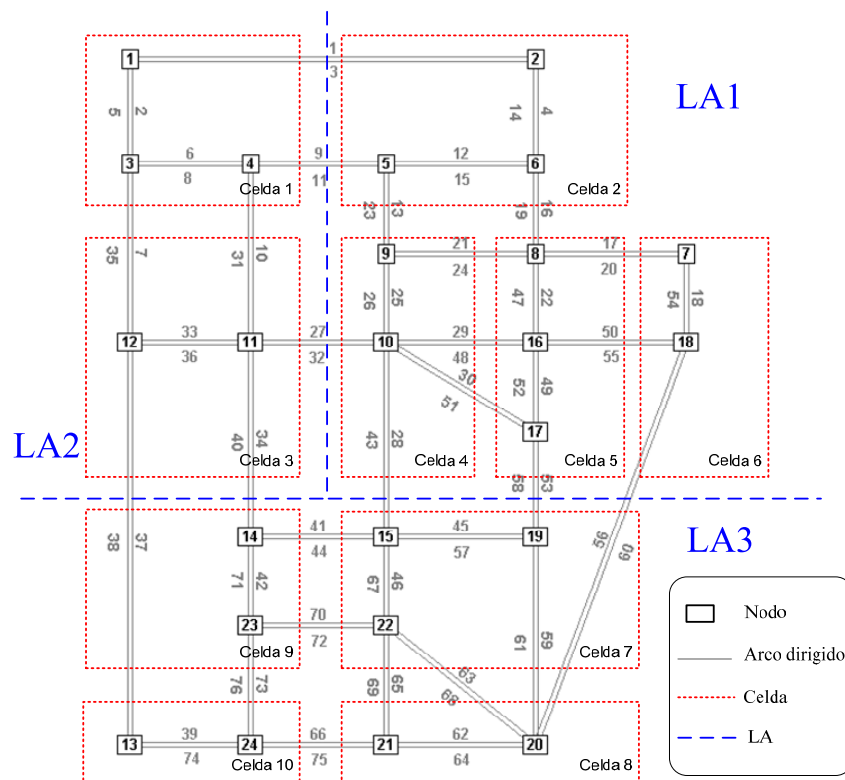
La información topológica para el diseño de red se ha tomado de la web de Bar–Gera (2007b). Las funciones de coste o demora responden a la expresión general de las BPR (Bureau of Public

Roads),  $s_a(v_a) = T_a \left[ 1 + 0.15 \cdot \left( \frac{v_a}{c_a} \right)^4 \right]$ , donde  $v_a$  es el volumen del arco  $a$ ,  $T_a$  representa el tiempo

de viaje a velocidad de flujo libre y  $c_a$  es la capacidad del arco  $a$ .

Para completar la construcción de la red para este ensayo es necesario definir la distribución celular de la red de telefonía asociada a dicha red de transporte, es decir, las celdas y áreas de localización que cubren los nodos de la red de transporte. Esto permite identificar el conjunto de fronteras observadas, y sus correspondientes grupos de arcos observados, que constituirán el conjunto de volúmenes de tráfico inferidos de la telefonía móvil. Típicamente, el radio de una celda varía desde 200 metros en entornos urbanos a dimensiones entre los 5 y 20 km en entornos rurales. Debido al carácter de este ensayo junto a la sencillez topológica de esta red, no se han utilizado criterios de dimensión o tamaño para el diseño celular, de modo que los nodos se han agrupado en celdas de un modo arbitrario atendiendo a criterios de vecindad. Posteriormente, dichos conjuntos de celdas se han agrupado en áreas de localización, obteniendo la red mostrada en la Figura 5-1, la cual contiene 22 fronteras observadas válidas con 34 arcos implicados.





**Figura 5-1: Red de Sioux Falls, incluyendo distribución ficticia de LAs y celdas.**

Inicialmente existían un conjunto mayor de fronteras entre celdas pero, tras realizar un procesado previo, finalmente sólo 22 fronteras fueron calificadas como válidas para ser utilizadas en el ensayo. La razón principal residió en criterios topográficos dado que había arcos que cruzaban fronteras entre celdas pero cuyo nodo origen o destino no estaba incluido en alguna de las dos celdas asociadas a la frontera. Ese fue el caso de los arcos 37 y 38, los cuales cruzaban las fronteras entre la celda 3 y 9 y entre la 9 y 10, pero su nodo origen o su nodo destino estaba en la celda 3 o en la 10. Por tanto, esos arcos introducían distorsión en la monitorización de esas fronteras (entre 3 y 9 y entre 9 y 10), dado que pese a cruzarlas, no cumplían la definición de frontera observada indicada en el apartado 4.3.2. Una frontera observada o aforo “virtual” lo conforman los arcos cuyo nodo inicial y nodo final están en celdas distintas y contiguas. Ante esta situación, las 4 fronteras correspondientes (una frontera entre celdas son 2 fronteras observadas, 1 por cada sentido) no podían pertenecer al conjunto de fronteras observadas válidas. Idéntica situación experimentaban los arcos 56 y 60, anulando las dos fronteras observadas asociadas a la frontera entre la celda 7 y 8. La Tabla 5-1 muestra las fronteras observadas consideradas definitivamente válidas y sus correspondientes arcos.

**Tabla 5-1: Fronteras observadas válidas y arcos asociados.**

<b>Id frontera</b>	<b>Celda Origen</b>	<b>Celda Destino</b>	<b>Id Arcos</b>	<b>Id frontera</b>	<b>Celda Origen</b>	<b>Celda Destino</b>	<b>Id Arcos</b>
1	1	2	1, 9	12	5	2	19
2	1	3	7, 10	13	5	4	21, 48, 51
3	2	1	3, 11	14	5	6	20, 50
4	2	4	13	15	5	7	53
5	2	5	16	16	6	5	17, 55
6	3	1	31, 35	17	7	4	43
7	3	4	32	18	7	5	58
8	4	2	23	19	7	9	44, 70
9	4	3	27	20	8	10	66
10	4	5	24, 29, 30	21	9	7	41, 72
11	4	7	28	22	10	8	75

La metodología propuesta exige, como prerrequisito, un conjunto de volúmenes en fronteras (o volúmenes agregados sobre grupos de arcos) y una matriz O-D previa. En general, esta matriz previa es obtenida de algún estudio cronológicamente anterior o de una encuestación básica que predefine la estructura de la matriz final a obtener. Para este ensayo, no se dispone de tal matriz de partida aunque sí de una matriz O-D factible, disponible en la web de Bar-Gera (2007b). De modo que dicha matriz será considerada como matriz auxiliar para obtener una matriz previa mediante un proceso de distorsión. Este proceso de distorsión consiste en multiplicar la matriz auxiliar por una matriz de coeficientes aleatorios entre 0.5 y 1.5, según posibles incrementos/decrementos del número de viajes en cada par O-D; obteniendo finalmente una matriz de partida (o previa) para ajustar, la cual estará formada también por 528 pares O-D definidos por los 24 centroides de la red.

Por otro lado, para el proceso de ajuste también se necesitan datos de volumen en cada frontera observada, es decir, volumen agregado por grupos de arcos que cruzan cada frontera. Obviamente, no se dispone de datos derivados de una red de telefonía real para este ensayo. Por ello, la obtención de unos volúmenes para ser utilizados en el proceso implica algún proceso de simulación que permita obtener tales datos. De modo que, una vez que se tienen identificados los arcos que pertenecen a cada frontera observada válida, se obtiene el volumen de cada frontera sumando los volúmenes del grupo de arcos que conforman tal frontera. Estos volúmenes de arco son tomados de una solución factible con respecto a flujos en arcos, proporcionada por Bar-Gera (2007b). Como se explicó en la sección 4.3 la inferencia de volúmenes mediante datos de teléfonos no es un proceso exacto sino que ofrece pequeñas fluctuaciones o errores de estimación respecto a la medida proporcionada por un aforo instalado

en la misma frontera. En este sentido, el error relativo absoluto medio alcanzado por los modelos de predicción seleccionados en el Capítulo 4 durante la validación numérica empleando volúmenes observados por estaciones de aforo se situó en torno al 20%. Bajo esta condición, finalmente, los volúmenes observados de frontera son obtenidos multiplicando el volumen agregado de cada frontera anteriormente explicado por un coeficiente aleatorio entre 0.8 y 1.2. Este rango es usado para introducir una distorsión en el volumen de frontera observado a fin de simular la discrepancia entre volúmenes inferidos de datos de teléfonos móviles y volúmenes reales de vehículos que realmente circularon.

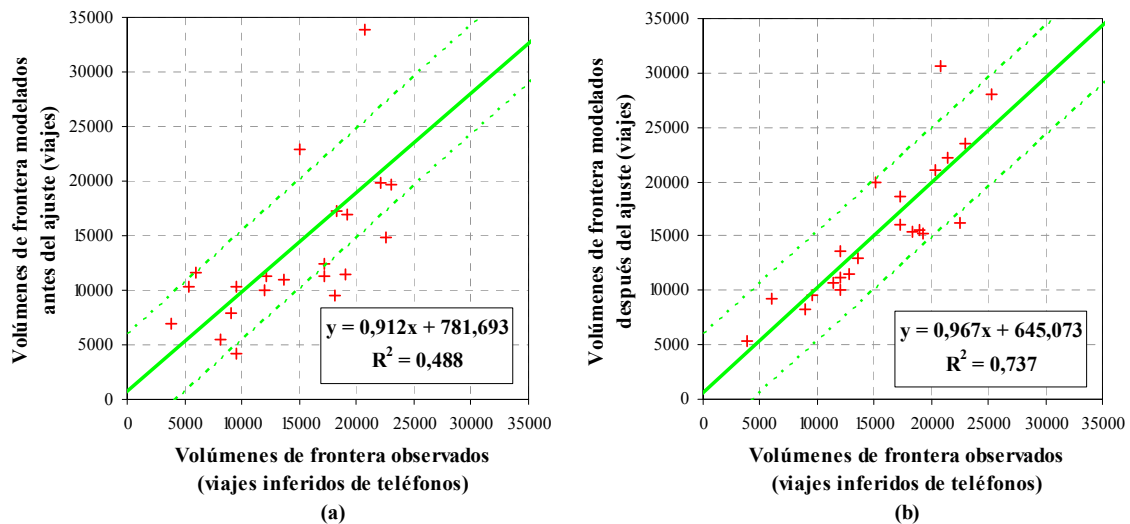
#### 5.4.1.2 Resultados

El objetivo de este apartado es presentar los resultados de este primer ensayo realizado para verificar la consistencia y eficiencia del algoritmo desarrollado. Para este caso se ha tomado un valor de  $W = 10$  junto con unos criterios de terminación de asignaciones para el código comercial usado (EMME/2) basados en un número medio de iteraciones por asignación para alcanzar las condiciones de equilibrio de usuario –  $N^{iter}=20$ ,  $relative\ gap\ (\%)=0.5$ ,  $normalized\ gap\ (min)=0.5$ . Para más detalles sobre estos criterios de terminación de la etapa de asignación (número de iteraciones,  $rgap$ ,  $ngap$ ) y otros, consultar las referencias EMME/2 (1990), MINUTP (1993), TRANSCAD (2006), y TRIPS (1998).

Para el problema de nivel superior (22), los valores de las cotas superiores e inferiores,  $\{u_{ij}, l_{ij}\}$  en el caso de cada viaje o par O–D, serán +25% y –25% de  $T_{ij}^{(0)}$  respectivamente; para viajes generados  $\{u_i^O, l_i^O\}$ , +15% y –15% de  $O_i^{(0)} = \sum_{j \in J} T_{ij}^{(0)}$ ; para viajes atraídos  $\{u_j^D, l_j^D\}$ , +15% y –15% de  $D_j^{(0)} = \sum_{i \in I} T_{ij}^{(0)}$ ; y para viajes totales  $\{u, l\}$ , +10% y 0% de  $T^{(0)} = \sum_{i \in I} \sum_{j \in J} T_{ij}^{(0)}$ . Se aprecia que estos valores de cotas están expresados en términos del estimado inicial para que, tras normalizar, se obtengan cotas relativas. Estos valores limitarán la distorsión experimentada por la matriz O–D previa.

Para presentar los resultados más significativos se utiliza tanto formato numérico como gráfico. La Figura 5-2a muestra los volúmenes de frontera (o agregados por grupos de arcos que cruzan cada frontera) obtenidos al asignar la matriz O–D previa a la red de Sioux Falls frente a los volúmenes observados en dichas fronteras. Estos últimos procederían de datos de movilidad derivados de conteos de teléfonos cruzando fronteras entre celdas colindantes. La Figura 5-2b corresponde a la representación de los volúmenes de frontera obtenidos por asignación de la matriz estimada usando el método propuesto frente a los observados derivados de paso de teléfonos por el mismo conjunto de fronteras. Esta figura muestra una buena correlación entre

los volúmenes observados y los modelados al asignar la matriz final estimada ( $R^2=0.737$ ), teniendo en cuenta el interés de preservar la información contenida en la matriz O-D de partida, alcanzando una razonable mejora respecto al  $R^2$  de partida ( $R^2=0.488$ ). Respecto a la pendiente de la recta ( $y=a+bx$ ), también se logró una mejora ( $b=0.967$ ) respecto de la previa ( $b=0.912$ ).



**Figura 5-2: Correlación entre los volúmenes observados en fronteras (grupos de arcos) y los modelados al asignar la matriz previa (a) y la matriz estimada (b).**

Lo anterior evalúa los resultados del ajuste en términos de volúmenes, no obstante, la metodología también atiende a criterios de control de la distorsión que experimenta la información contenida en la matriz estimada, a diferentes niveles de agregación (viajes generados, atraídos y totales), durante el proceso de ajuste. En este sentido, para evaluar la distorsión experimentada por la matriz estimada respecto a la previa durante el ajuste, se presenta la Figura 5-3, la cual compara el número de viajes de cada celda de la matriz O-D previa frente al número de viajes de los mismos pares O-D de la matriz final estimada usando la metodología. El gráfico revela los reducidos niveles de desviación del número de viajes de cada par O-D de la matriz estimada respecto a los de la matriz previa ( $R^2=0.934$ ). Se puede ver claramente que la nube de puntos queda confinada en una estrecha banda delimitada por los incrementos y decrementos máximos establecidos como parámetros del proceso de ajuste y que, en este caso, son de un 25%. Una gráfica similar se puede obtener para los viajes generados (agregación en destinos) y para los viajes atraídos (agregación en orígenes), pero las bandas en este caso situarán el número de viajes dentro de unos incrementos/decrementos del 15%.

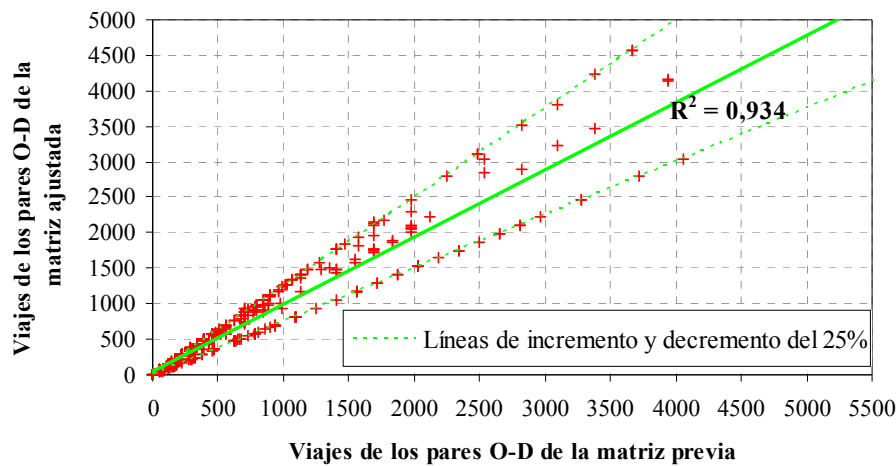


Figura 5-3: Correlación entre la matriz previa y la matriz estimada.

## 5.4.2 Experimento 2: Red Comunidad Autónoma de Madrid

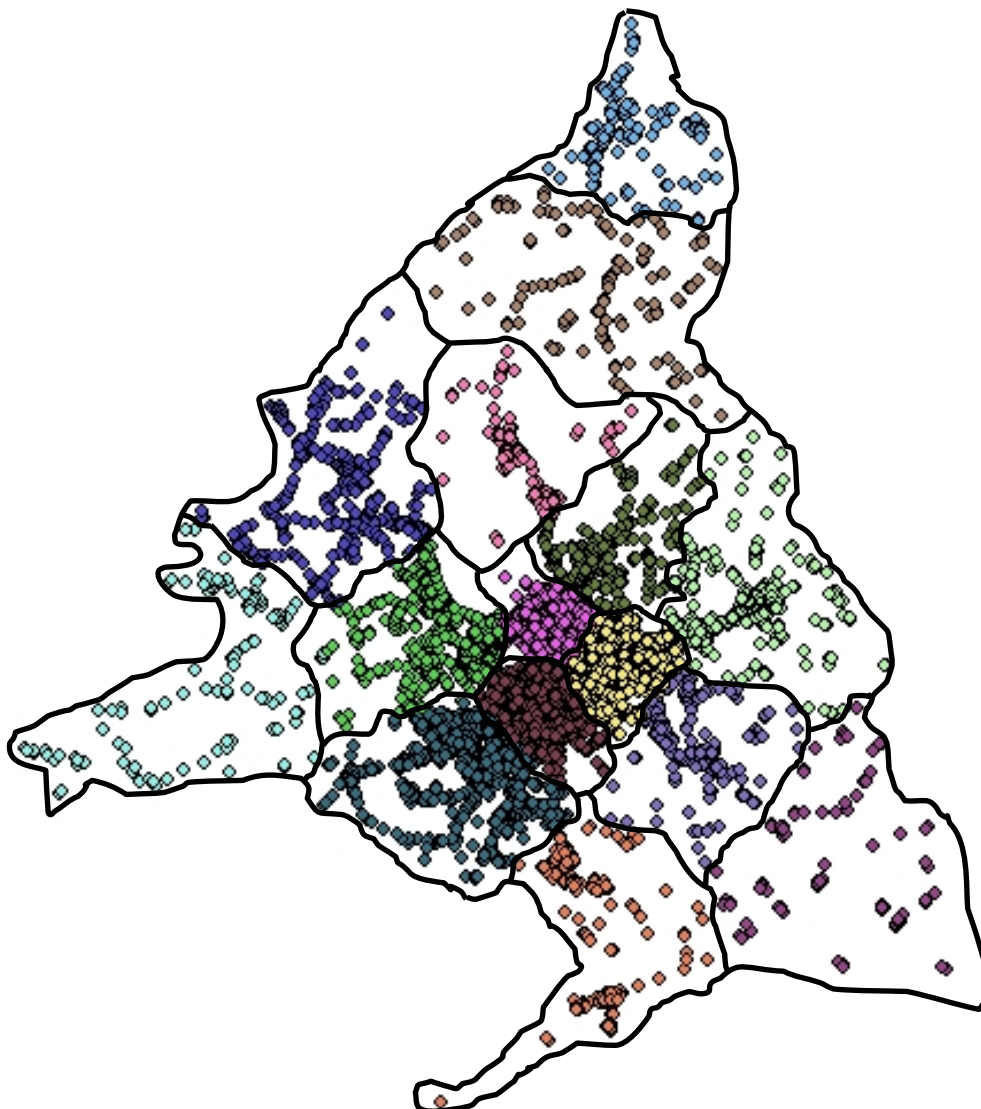
### 5.4.2.1 Configuración de la red

En este segundo ensayo, se ha tomado una red de gran tamaño correspondiente a la red de la Comunidad Autónoma de Madrid (para el año 2004), cuyas dimensiones topológicas son 1079 centroides, 9557 nodos regulares, 1079×1079 pares O–D y 21713 arcos dirigidos (CRT, 2004). Por cuestiones de nitidez gráfica, la representación del modelo de red ha sido omitida ya que la visualización de los arcos y nodos resulta poco nítida debido a sus dimensiones.

Como se ha mencionado anteriormente, se precisa un conjunto de volúmenes en fronteras (agregados sobre grupos de arcos) junto con una matriz O–D previa para realizar el proceso de ajuste. La matriz O–D previa se toma de un estudio anterior desarrollado sobre esta red, el cual dispone de una matriz en período de hora punta de la mañana (8:00–9:00 a.m.), correspondiente a la Encuesta Domiciliaria de Movilidad 2004, con 1179×1179 pares O–D.

Para los volúmenes observados se ha utilizado un procedimiento basado en el citado estudio previo desarrollado sobre esta red. De él, se tomó una matriz O–D resultado de un proceso de ajuste con aforos tradicionales. Esta matriz, tras asignarla a la red, permitió disponer de unos volúmenes de todos los arcos que podían considerarse, teóricamente, como “reales”. Estos volúmenes de arco serían empleados más adelante para obtener los volúmenes de frontera (grupos de arcos). Basándose en los resultados del ensayo desarrollado en el Capítulo 4 sobre volúmenes inferidos con datos de telefonía móvil, éstos presentan pequeñas diferencias respecto a la medida proporcionada por una estación de aforo instalada en la misma frontera. De modo que esos volúmenes de arco son escalados por un coeficiente aleatorio entre 0.8 y 1.2 para

introducir una distorsión en el volumen final observado que permita simular el valor de error relativo en valor absoluto, alcanzado en el citado ensayo, entre volúmenes inferidos de datos de teléfonos móviles y volúmenes de vehículos que realmente circularon. Pero la metodología emplea volúmenes sobre grupos de arcos correspondientes a fronteras entre celdas. Por ello, se requería disponer de una distribución celular, la cual se trazó de un modo ficticio usando 15 LAs y 20 celdas por LA (Figura 5-4). Esta distribución permitió identificar los grupos de arcos que conforman cada frontera, agrupando arcos usando la definición de frontera observada: arcos cuyos nodos de inicio y fin están en celdas diferentes y vecinas.



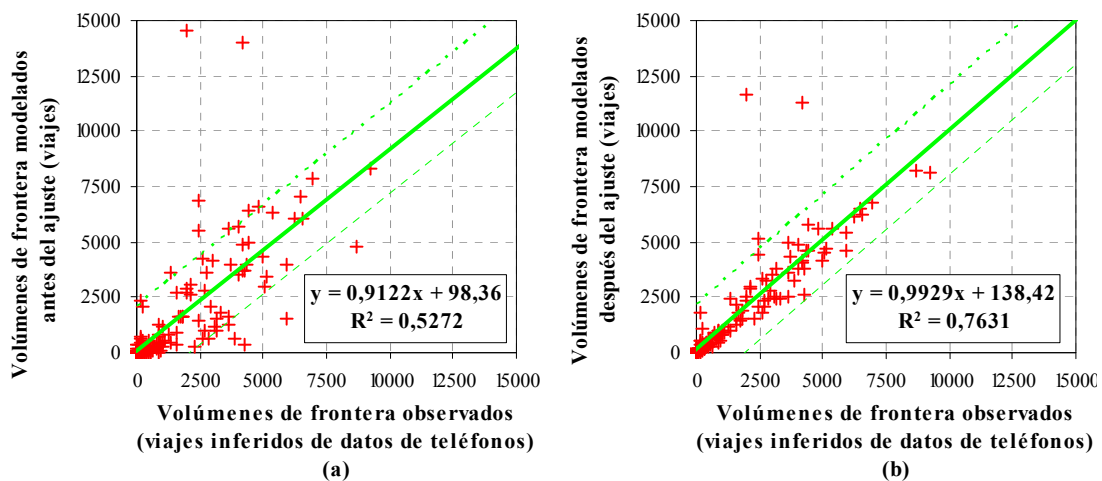
**Figura 5-4: Nodos de la red de Madrid agrupados en 15 LAs.**

Una vez conocidos los grupos de arcos para cada frontera existente, el volumen de una frontera observada se generó sumando los volúmenes de cada uno de los arcos identificados como parte

de cada frontera. Sin embargo, no todas las fronteras existentes sobre una red pueden ser usadas como válidas debido a sus características. Esto implica la realización de un pre-procesado sobre el conjunto inicial de fronteras entre celdas para obtener el conjunto final de fronteras “válidas” usadas como observadas para el proceso de ajuste. En esta red, existen 412 fronteras observadas válidas con 1265 arcos implicados, donde el principal criterio seguido para su selección ha sido evitar posibles problemas derivados de la existencia de peatones, descartando fronteras en las que existiera algún arco con tráfico de peatones asociado.

### 5.4.2.2 Resultados

En este caso, se ha tomado un valor de  $W=10^3$  – en torno al orden del valor inicial de la función objetivo (32) evaluada para este problema – junto con los criterios por defecto de terminación de las asignaciones, que emplean un número alto de iteraciones por asignación (primer criterio a ser cumplido en todos los casos) –  $N^{iter}=200$ ,  $relative\ gap(\%)=1$ ,  $normalized\ gap(\min)=1$ . Estos valores difieren de los tomados en la red de Sioux Falls, en especial en lo que respecta al número de iteraciones por asignación, debido a que los niveles de congestión sobre la red de Madrid son ciertamente más altos, de modo que las condiciones de equilibrio de usuario son más difíciles de obtener. En el problema de nivel superior (22), los valores de las cotas superiores e inferiores de las restricciones son iguales a las del ensayo en la red de Sioux Falls.



**Figura 5-5: Correlación entre los volúmenes observados en fronteras (grupos de arcos) y los modelados al asignar la matriz previa (a) y la matriz estimada (b).**

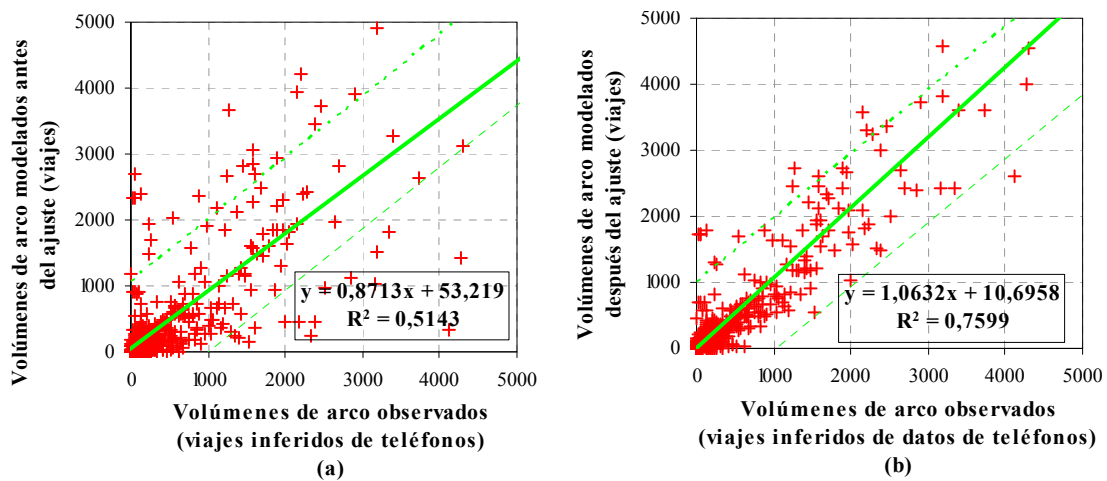
Como se ha comentado previamente, la metodología de ajuste atiende a dos criterios: i) que la matriz resultante reproduzca, al ser asignada a la red, los datos de los volúmenes agregados lo más fielmente posible; y ii) que la información contenida en la matriz O–D previa no se vea distorsionada en exceso durante el proceso de ajuste. Para evaluar la calidad del ajuste en

términos de volúmenes se procede a comparar, numérica y gráficamente, los coeficientes  $R^2$  antes y después del ajuste a nivel de volúmenes de frontera. La Figura 5-5a muestra los volúmenes agregados por grupos de arcos que cruzan cada frontera observada obtenidos tras asignar la matriz O-D previa a la frente a los volúmenes observados en dichas fronteras derivados de datos de teléfonos. La Figura 5-5b corresponde a la representación de los volúmenes de frontera obtenidos al asignar la matriz ajustada usando la metodología frente a los volúmenes observados de frontera derivados de los teléfonos. Se aprecia que el proceso de ajuste consigue valores aceptables ( $R^2=0.7631$ ) para el número de arcos implicados (1265 arcos, un 5.83% del total de arcos de la red). Respecto a la valoración de la metodología en términos de distorsión de la información contenida en la matriz O-D estimada respecto a la de partida, la matriz O-D previa y la ajustada alcanzan un indudable alto valor de correlación entre ellas ( $R^2=0.9893$ ) debido a las restricciones impuestas.

Cabe destacar que observando la Figura 5-5 se aprecian dos outliers que potencialmente afectan a los valores de  $R^2$ . Esos outliers pueden ser debidos a que la frontera o grupo de arcos contiene, por ejemplo, algún arco con línea de transporte público asociada o incluso, que tenga próximo un núcleo residencial que introduzca pérdidas. Estas situaciones inducen alteraciones respecto a la media, de modo que los modelos de estimación de volúmenes no pronostican con la precisión esperada. Se podría estudiar en detalle el origen de esos outliers para decidir si se consideran válidas o no las fronteras asociadas; sin embargo, dado que en el planteamiento inicial se consideraron válidas, se ha optado por mantenerlas en la representación de los resultados.

En este ensayo, al disponer de volúmenes de arcos para generar los de frontera surge la posibilidad de evaluar las comparaciones de volúmenes pero a nivel de arco, es decir, desagregando los volúmenes observados de frontera (grupos de arcos) en volumen a nivel individual de arco. La Figura 5-6 a y b muestra la representación de los volúmenes observados desagregados a nivel de arco frente a los volúmenes obtenidos por asignación de la matriz O-D previa y la ajustada, respectivamente, sobre el conjunto de arcos implicados en las fronteras. Observando la Figura 5-5b y la Figura 5-6b se hace patente el comportamiento similar en términos de correlación entre los volúmenes observados y estimados, representándolos tanto a nivel de frontera (agregado por arcos) como a nivel de arco.





**Figura 5-6: Correlación entre los volúmenes de arcos observados (pertenecientes a alguna frontera observada) y los modelados al asignar la matriz previa (a) y la matriz estimada (b).**

La Tabla 5-2 compara los resultados ( $R^2$  y ecuación de recta regresión) evaluados a nivel individual de arco y de grupos de arcos. Lógicamente, el  $R^2$  alcanzado tras el ajuste evaluado a nivel de volumen agregado (grupos de arcos) será mejor que el evaluado a nivel de volumen de arco, dado que la información observada en la que se basa el proceso de ajuste viene expresada a esa escala agregada por grupos de arcos. Bajo este mismo razonamiento, el offset y la pendiente de la recta evaluada a nivel de volumen de arco no alcanzarán una mejora significativa frente a los de la previa. Mientras que en el caso de la recta evaluada sobre volúmenes agregados tras el ajuste sí se consigue que la pendiente sea muy próxima a la unidad ( $b=0.9929$ ), idealmente  $45^\circ$ , manteniendo un valor de offset bajo pese a trabajar con escala agregada de volúmenes.

**Tabla 5-2: Comparativa de ecuación de la recta de regresión y  $R^2$ , antes y después de ajuste.**

	Antes del ajuste	Después del ajuste
Comparando volúmenes en grupos de arcos	$R^2=0.5272$ $y=98.36+0.9122x$	$R^2=0.7631$ $y=138.42+0.9929x$
Comparando volúmenes en arcos	$R^2=0.5143$ $y=53.219+0.8713x$	$R^2=0.7599$ $y=10.696+1.0632 x$

Otro criterio para evaluar este planteamiento de ajuste con volúmenes agregados sobre grupos de arcos en lugar de mediante volúmenes en arcos es comparar los resultados que se obtendrían empleando ambas metodologías: con volúmenes a nivel de arco y de frontera (grupos de arcos). De cara a la elección de una metodología de estimación que trabaje con volúmenes de arcos para ajustar una matriz previa, la más adecuada, a efectos de comparación con la planteada en

esta Tesis, es la desarrollada por Doblaz y Benitez (2005). Se trata de valorar el uso de volúmenes agregados sobre grupos de arcos en lugar de los clásicos conteos de tráfico, por ello esa metodología es la que mejor comparativa de prestaciones ofrecería dado que, en esencia, el algoritmo de ajuste seguido en esta Tesis está basado en él, pero con las modificaciones oportunas para trabajar a nivel de grupos de arcos.

En base al razonamiento anterior, se realizó un proceso de ajuste de la misma matriz O-D de partida pero utilizando como volúmenes observados medidas proporcionadas por aforos instalados sobre 1420 arcos de la red, disponibles de un estudio anterior desarrollado sobre la misma red. Es preciso hacer notar que en el ensayo empleando volúmenes de frontera se monitorizaron 1265 arcos asociados a las 412 fronteras observadas. De manera que, aplicando la metodología desarrollada por Doblaz y Benitez (2005), la cual realizaba el proceso de ajuste con volúmenes observados a nivel de arcos, el coeficiente  $R^2$  entre los volúmenes observados y los estimados ( $R^2 \approx 0.80$ ) no reflejó una mejora muy significativa comparada con la metodología propuesta a nivel de grupos de arcos ( $R^2 \approx 0.76$ ). Sin embargo, utilizar datos procedentes de estaciones de aforo implica unos costes asociados a la infraestructura de detectores que con la telefonía móvil no existirían. En este último caso, los volúmenes se basarían en datos de teléfonos ya disponibles por los operadores para su propio funcionamiento, sin necesidad de instalación de módulos adicionales ni reducir significativamente la exactitud obtenida del ajuste, como se ha demostrado en este ensayo.

Un último aspecto a comentar es la convergencia del algoritmo de ajuste. La Figura 5-7a muestra la evolución de la función objetivo de la Lagrangiana Aumentada (32) durante el proceso de ajuste, mostrando la rápida convergencia lograda, dadas las dimensiones del problema, ejecutándolo con 5 subproblemas de 10 iteraciones cada una. La evolución típica de la Lagrangiana aumentada, respecto al término más significativo, se aprecia con más detalle en la Figura 5-7b. El término correspondiente a las restricciones funcionales (Figura 5-7c) muestra un orden de magnitud muy inferior al del término cuadrático asociado a los volúmenes de frontera, reflejando el nivel de control que sobre la información contenida en la matriz O-D ejerce el término correspondiente a las restricciones. Este efecto se debe a que las restricciones de viajes generados, atraídos y totales no suelen violar las cotas superiores e inferiores que tienen impuestas, por lo que no son tan restrictivas como las de variables acotadas. La rápida convergencia alcanzada con las etapas del algoritmo también puede ser observada en las gráficas demostrando que, a efectos prácticos, el proceso iterativo se podría haber detenido antes de que los criterios de terminación impuestos sean alcanzados (número de subproblemas y número de iteraciones por subproblema), consiguiendo ajustes más rápidos.

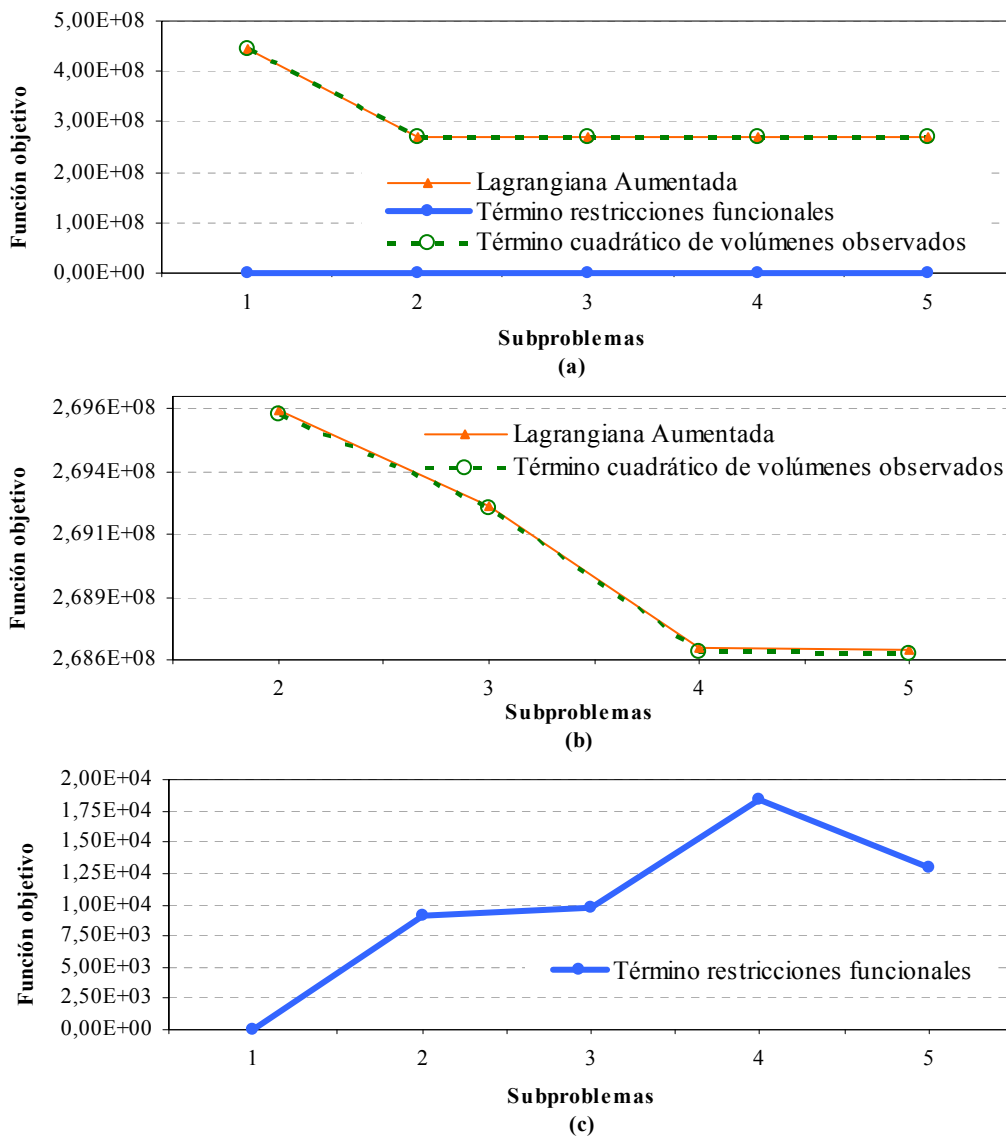


Figura 5-7: (a) Evolución de la función objetivo durante el proceso de ajuste de la matriz. (b) Zoom para los subproblemas 2–5. (c) Zoom para el término de las restricciones de la función objetivo.

Con ello se ha realizado una validación experimental en la que se revela que emplear volúmenes agregados sobre grupos de arcos, en lugar de desagregados a nivel individual de arco, en los procedimientos para el ajuste de matrices constituye una alternativa viable que no penaliza significativamente la exactitud obtenida. En especial, si estos volúmenes agregados se infirieran de la telefonía móvil se lograría reducir el coste económico y temporal del estudio frente a los procedimientos tradicionales basados en aforos.



# CAPÍTULO 6

---

## CONCLUSIONES Y LÍNEAS FUTURAS

### 6.1 CONCLUSIONES

#### 6.1.1 Introducción

Esta Tesis está enfocada a la aplicación de los sistemas de telefonía móvil en los estudios de movilidad de transporte. El objetivo principal ha sido la realización de aportaciones importantes en el campo de la inferencia de volúmenes de tráfico y en la estimación de matrices O-D, orientándolas a datos tomados de los sistemas de telefonía móvil. La telefonía móvil se está convirtiendo en un sector de gran utilidad en el ámbito del tráfico. En este contexto, el Capítulo 3 revisó en profundidad el “Estado-del-Arte” presentando los conocimientos actuales sobre la estimación de parámetros de tráfico mediante el uso de información procedente de sistemas de telefonía móvil, así como las prestaciones y beneficios que ofrecen este tipo de estimaciones.

En el Capítulo 4 se formularon un conjunto de modelos para la inferencia de volúmenes de tráfico de vehículos mediante teléfonos móviles que pasan por una determinada zona (frontera de una celda a otra). Tras realizar un minucioso estudio comparativo entre ellos, atendiendo a diferentes criterios, y contrastar los resultados, se seleccionaron aquellos que presentaron mejor balance entre todos los criterios junto a una cierta exactitud de las predicciones. La forma funcional de los modelos seleccionados puso de manifiesto:

- i) lo insuficiente de usar exclusivamente información de llamadas en movilidad,
- ii) la necesidad de incorporar en los modelos información sobre la variabilidad horaria de la conducta de los usuarios que se desplazan, en medios mecanizados, a la hora de realizar llamadas, y
- iii) la aplicabilidad de la metodología a ciertas fronteras entre celdas dentro de unos rangos admisibles de error.

En el estudio numérico, los valores de predicción de los modelos seleccionados mostraron resultados razonables en comparación con volúmenes observados por estaciones de aforo instaladas sobre las mismas vías, pese a la naturaleza de la principal variable de entrada al modelo: llamadas en movilidad. Este tipo de datos no representa el conjunto total de teléfonos que han pasado por una zona sino sólo aquellos pertenecientes al operador de interés que, a su vez, han realizado llamadas bajo ciertas condiciones analizadas en el apartado 4.3.2. Ante esta situación, la representatividad de la muestra queda condicionada en cierta medida. Lo ideal sería emplear datos generados por teléfonos que simplemente estuvieran encendidos, sin condicionar la muestra a la realización de llamadas. Aunque estos datos son generados internamente por los sistemas de telefonía en su funcionamiento, como ya se explicó en el Capítulo 2, recopilarlos implica una modificación de los módulos existentes en la plataforma del operador ya que las APIs necesarias para su almacenamiento permanente no están accesibles. En consecuencia, y buscando un bajo impacto en los módulos del operador de telefonía, se optó por emplear datos de llamadas; más concretamente, llamadas en movilidad, las cuales se asocian a un conjunto menos representativo de la población. La validación experimental realizada de los modelos propuestos, los cuales consideraban esas dificultades, mostró evidencias de que la telefonía móvil constituye una alternativa válida para la determinación de volúmenes de vehículos pasando por una zona frente a los procedimientos tradicionales con aforos, dado que no se reducía en gran medida la precisión lograda.

Una aplicación inmediata de esos volúmenes inferidos de la telefonía móvil se enfocó al ámbito de la obtención de matrices de movilidad. Bajo este planteamiento, el Capítulo 5 abordó el problema de estimación de matrices O-D presentando un algoritmo cuya principal novedad es el uso de información agregada de volúmenes sobre grupos de arcos, la cual podría proceder, entre otros, de los sistemas de telefonía móvil. Para ello, se propuso una metodología de ajuste de una matriz previa mediante un modelo de estructura binivel empleando la adaptación a escala de volúmenes por grupos de arcos de un algoritmo robusto y eficiente presentado por Doblas y Benitez (2005), el cual trabajaba a nivel clásico de volumen de arco. Respecto al algoritmo de resolución propuesto se realizó un estudio numérico que permitió efectuar una comparativa frente a ajustes de matrices mediante datos de aforos, destacando i) la convergencia del algoritmo propuesto y ii) la aplicabilidad de la metodología a redes reales de mediano o gran tamaño. En este sentido se puso de manifiesto que emplear volúmenes agregados sobre grupos de arcos en los procedimientos para el ajuste de matrices constituye una alternativa viable, sin disminuir significativamente la exactitud alcanzada, frente a los procedimientos basados en volúmenes de arco. Adicionalmente, si los volúmenes agregados se infirieran de la telefonía

móvil se lograría reducir a su vez el coste económico y temporal del estudio respecto del uso tradicional de estaciones de aforo.

Finalmente, se desea destacar que los mayores beneficios que puede reportar un desarrollo integrado de los métodos planteados en esta Tesis, tanto de inferencia de volúmenes como de ajuste de matrices, vienen expresados en términos de ventajas para la sociedad obtenidos desde el punto de vista tecnológico. El hecho de que cada uno de esos métodos pueda ser abordado mediante la misma tecnología usada por los sistemas de telefonía convencionales y que exista la posibilidad de desarrollar una herramienta capaz de integrarlos es de sumo interés para los planificadores de transporte. Una herramienta de ajuste de matrices mediante volúmenes derivados de teléfonos móviles confiere un gran potencial práctico al proceso, dado que resuelve la problemática debida a, fundamentalmente:

- Coste. El desarrollo de modelos de estimación que tomen datos de sistemas de telefonía móvil y terminales convencionales usados hoy en día por cualquier usuario supone un ahorro económico importante asociado a gastos de instalación y mantenimiento que conllevarían las estaciones de aforo. Desde el punto de vista de cobertura, la red de telefonía también ofrece un ahorro considerable dado que su zona de servicio abarca toda la cobertura nacional, incluyendo la red viaria, mientras que la infraestructura de los clásicos aforos es limitada, condicionando por tanto la captura de conteos de tráfico.
- Tiempo. La realización de grandes encuestas de movilidad supone un coste económico y humano muy elevado, por lo que su ejecución suele abordarse aproximadamente cada 5 años. Es evidente que durante ese tiempo pueda producirse un incremento/decremento de la movilidad en una ciudad o región debido a la variación en la población o del parque automotor, entre otros. Con lo cual, la percepción de movilidad reflejada por esas encuestas suele quedarse obsoleta al poco tiempo de su elaboración. El disponer de datos de movilidad generados por teléfonos de usuarios en los desplazamientos que habitualmente realizan durante un día, permite actualizar de un modo más continuo esos datos, incluso pueden servir para elaborar otro tipo de matrices de movilidad a niveles más extensos de zonificación (macrozonas).
- Fiabilidad. Aunque la realización de encuestas está sujeta a fuertes controles de calidad para regular su fiabilidad, en ocasiones, suelen introducir cierto sesgo debido al diseño, selección o determinación del tamaño de la muestra que la genera. Incluso, puede darse el caso que los individuos más móviles sean los que presenten menor

probabilidad de ser entrevistados, pues son los que pasan menos tiempo en casa. Por otro lado, la realización de encuestas directamente a vehículos sobre el viario es ciertamente incómoda para los usuarios y consumen mucho tiempo, con la consiguiente distorsión que se deriva de estos hechos. En cambio, utilizar como muestra ciertos teléfonos móviles de individuos que los llevan consigo en todo momento (i.e.: encuesta panel), lo cual es algo habitual en nuestros estilos de vida, permite obtener datos razonablemente representativos de las pautas de movilidad de los usuarios sin provocar ningún tipo de perturbación.

No obstante, el mejor criterio para evaluar el uso de la telefonía móvil en el ámbito de la movilidad de vehículos es determinar el impacto de los datos estimados a partir de ella en términos económicos y técnicos. Los siguientes apartados analizan las prestaciones de estimaciones derivadas de sistemas de telefonía móvil, enfocándolas a un nivel generalizado entre los posibles parámetros de tráfico que se pueden derivar de estos sistemas.

### **6.1.2 Evaluación del uso de teléfonos móviles para movilidad de vehículos**

Este apartado revisa las principales ventajas e inconvenientes a nivel técnico que se han encontrado tanto en el desarrollo de este trabajo, como en otros realizados en diferentes proyectos pilotos.

#### **6.1.2.1 Ventajas**

Un modo de estudiar la efectividad de estas técnicas basadas en la telefonía móvil es compararlas con los resultados obtenidos con metodologías tradicionales, respecto a las cuales presenta ciertas ventajas. Principalmente, el tratamiento de esos datos supone manejar una muestra muy extensa de la población, debido a las altas cuotas de penetración en gran parte de las ciudades del mundo. Junto a ello, destaca su bajo coste de implantación debido a que cualquier terminal de los que son utilizados hoy en día por la mayoría de usuarios puede ser válido como sonda de tráfico.

Para conocer detalladamente las características del tráfico vehicular existente en la red viaria sería necesario aforar un gran número de vías de manera casi permanente, lo cual implicaría disponer de una infraestructura considerable. Por ello, en contraste con las redes de sensores fijos (detectores, cámaras, etc.), la telefonía móvil proporciona información para analizar y cuantificar las variables de circulación sobre cualquier zona donde exista cobertura telefónica (GSM, GPRS y/o UMTS), sin necesidad de un costoso despliegue de infraestructura adicional. Algunos de los trabajos presentados hasta el momento dejan constancia de que este tipo de



técnicas consiguen magníficos resultados sin necesidad de gastos excesivos de mantenimiento, debido a que no sufre los deterioros de los equipos que están instalados a lo largo de las vías. Asimismo, su funcionamiento es independiente de las condiciones meteorológicas, como pasaría bajo situaciones de poca visibilidad con otro tipo de sistemas como el de lectura del número de matrícula.

Por otro lado, las soluciones basadas en telefonía móvil permiten derivar resultados de la información procesada de un modo más rápido que con las técnicas tradicionales. Esto ocurre en el caso de la estimación tradicional de matrices O-D mediante encuestas que, desde la toma inicial de datos hasta la explotación de los primeros resultados, supone un largo proceso que puede abarcar un periodo de incluso años. Las metodologías presentadas en el “Estado-del-Arte” muestran la rapidez conseguida a la hora de confeccionar información de las características del tráfico (tiempos de viaje, velocidad, etc.), llegando algunas de ellas a proporcionar información en tiempo real. Esta capacidad mejora de modo considerable los sistemas de información de usuarios, como consulta WEB/WAP/SMS o mediante señalización VMS (Variable Message Signs), dado que facilita la toma de decisiones sobre la ruta a tomar, hora de salida e incluso el modo de transporte a emplear. En consecuencia, se produce un efecto de “realimentación” que favorece el flujo de circulación al conseguirse reducir los tiempos de viaje empleados por los viajeros, debido a que éstos escogen las rutas más rápidas y menos congestionadas.

No obstante, la principal ventaja se refleja en la mejora de la gestión y planificación del tráfico debido al reducido tiempo de procesado de los datos. Esta mejora no se puede cuantificar en términos económicos, pero sí en calidad del servicio desde la perspectiva de uso diario de las vías por parte del usuario. Según lo expuesto en el apartado anterior, estas soluciones basadas en telefonía móvil conseguirían mejorar los sistemas de monitorización de tráfico vehicular desde el punto de vista de tiempo, coste (utilizaría la infraestructura existente de telefonía celular), cobertura y representatividad estadística de la muestra (alta tasa de penetración telefónica).

#### **6.1.2.2 Inconvenientes**

Las metodologías basadas en la telefonía celular poseen ciertas limitaciones de naturaleza técnica que podrían afectar a la capacidad de la estimación de datos de tráfico. Estas limitaciones tienen diferentes impactos según el parámetro de tráfico estudiado y, aunque ya han sido comentadas a lo largo de la Tesis, a continuación se describen las más importantes.

### ■ *Precisión en la información de posición*

La precisión de localización obtenida a partir de datos de teléfonos móviles varía según el estado del teléfono (encendido, llamada...). Lo ideal sería conocer la posición exacta en términos de  $(x, y, z, t)$ . Sin embargo, esto no es posible sin modificaciones en la red o en los terminales. El nivel de precisión más fino proporcionado por un sistema de telefonía móvil es a nivel de celda durante una llamada. De modo que se localiza al teléfono dentro del área de cobertura de la estación base a la que está conectado. Esta ubicación es una aproximación de la zona geográfica en la que un teléfono se encuentra, pero no su posición exacta. Hay un margen de error que depende del radio de la celda. Este radio oscila entre los 100 m y 1 km en ciudades. En entornos rurales depende de la densidad de las antenas, aunque suele oscilar entre 5 y 20 km. Ese margen de error empeora para los teléfonos que no tienen una conexión activa a la red (no-llamada), donde la precisión de la información de posición es a nivel de área de localización (LA). Por ejemplo, si se trabaja a nivel de LA (agrupaciones de celdas) en lugar de a nivel de celda, la zona geográfica donde el teléfono se puede encontrar es más amplia y, por tanto, hay una mayor incertidumbre.

Con independencia del inconveniente asociado al tamaño de las celdas, y más aún de las áreas de localización, existe otra dificultad que afecta al mapeado o “matching” de teléfonos sobre vías existentes en el área de servicio de estas zonas. Este efecto se puede suavizar mediante un estudio exhaustivo entre la topología de la red viaria y las celdas de la red, lo cual es una tarea difícil, o mediante algoritmos de map-matching. El nivel de error medio en el mapeado puede ser relevante para ciertos parámetros, por lo que es un aspecto a tener en cuenta a la hora de proporcionar resultados.

Adicionalmente, es necesario indicar que el uso del cambio de área de localización quizá no es el evento más apropiado para el estudio de movilidad de áreas urbanas. Debido a la propia naturaleza de la distribución celular de una red, las áreas de localización (grupos de celdas) pueden abarcar extensiones de terreno de grandes dimensiones. Esto implica que las fronteras entre LA sean espacialmente poco frecuentes y, por tanto, el conjunto de áreas de las que se podrá inferir los volúmenes de paso será pequeño.

### ■ *Validez de los datos*

La existencia de un único teléfono activo (de un pasajero o del propio conductor) dentro del vehículo asegura dicho vehículo como una sonda válida. Sin embargo, esta suposición no es suficiente en zonas con densidad de población alta donde los peatones, usuarios de teléfonos móviles que están en edificios o cualquier tipo de usuarios estáticos, pueden afectar a la

estimación de parámetros de tráfico. Una solución sencilla es considerar válidas para extraer datos de tráfico mediante telefonía móvil únicamente aquellas zonas alejadas de núcleos urbanos o de áreas residenciales dado que en esos casos los usuarios serán los ocupantes de los vehículos que circulan por vías como autovías, vías interurbanas o similares

Si se quisiera extraer información de tráfico mediante la telefonía en entornos urbanos, considerando la grave influencia de la presencia de peatones o usuarios estáticos, se hace indispensable una diferenciación entre éstos y los usuarios en movimiento. Los procedimientos encargados de la identificación y filtrado de esos teléfonos no-válidos resultan ser una tarea compleja pero necesaria en la estimación de la mayoría de parámetros de tráfico. Otros parámetros, tales como velocidad o tiempos de viaje, no se ven tan afectados por estas circunstancias puesto que los teléfonos estáticos son directamente filtrados debido al propio carácter del evento usado: el handover.

Por tanto, la estimación de datos de tráfico basados en telefonía móvil debe combinarse con otros procedimientos basados en velocidad u otros métodos capaces de implementar tal diferenciación. En ese sentido, el IAGB (Institute for Applications of Geodesy to Engineering) dentro del proyecto Do-iT desarrolló algoritmos para determinar y describir patrones de movilidad mediante datos anónimos de telefonía móvil para la identificación del modo de transporte que está utilizando el teléfono. (Wiltschko, 2006). En ellos se desarrollaron unas funciones de análisis y modelos de predicción para estimar la cantidad de miembros en cada clase o modo de transporte, logrando así distinguir entre clases “activas” o “no activas” (válidas o no válidas). En consecuencia, se consiguió mayor precisión que utilizando modelos de estimación de verosimilitud, requiriendo la integración de horarios y rutas de líneas de transporte público (autobuses, trenes, etc.) en los modelos. Por otra parte, es necesario resaltar que la presencia de varios teléfonos móviles en un vehículo también perturba la estimación de parámetros de tráfico. Por todo ello, es necesario indicar que los datos obtenidos de teléfonos móviles deben ser calibrados con otros métodos de monitorización, como detectores o cámaras, antes de usarlos como valor definitivo.

### **6.1.3 Estudio de mercado**

Según lo expuesto anteriormente, los datos de movilidad origen-destino son una de las fuentes de información más importantes requeridas con vista a la gestión de la red de infraestructura terrestre y su planificación estratégica. En el capítulo anterior se ha planteado una nueva alternativa para la obtención de esos datos de movilidad de una forma más rápida y precisa, haciendo uso de volúmenes de tráfico inferidos de teléfonos móviles.

Es obvio que las nuevas tecnologías deben ser aprovechadas para introducir mejoras en la sociedad actual. Y qué mejor beneficio que hacer frente a los dos problemas fundamentales de tráfico: la saturación y la seguridad. La saturación en las carreteras, especialmente aguda en las zonas urbanas y en las grandes ciudades, provocan molestias a los conductores, con los consiguientes efectos negativos sobre el medio ambiente. Por otro lado, para los conductores y otros usuarios de la carretera, unos sistemas que permitan reforzar la seguridad en términos de tiempo de respuesta constituyen la primera preocupación. De manera que unos sistemas de información de tráfico que permitan hacer predicciones de movilidad de usuarios según el historial de movimientos de fechas anteriores darían un paso adelante para solventar estos problemas. Este apartado analiza la aplicabilidad y rentabilidad de herramientas basadas en telefonía móvil, haciendo énfasis en los campos de interés de esta Tesis: volúmenes de tráfico y matrices de movilidad.

#### **6.1.3.1 Aplicabilidad**

Una herramienta de ajuste de matrices mediante volúmenes derivados de teléfonos móviles confiere un gran potencial práctico al proceso. No obstante, la información derivada de la telefonía permite obtener un amplio abanico de productos relacionados con datos de tráfico, los cuales pueden ser proporcionados a administraciones públicas, consultoras de transporte, otras compañías de telecomunicaciones, compañías de logística u otras organizaciones relacionadas con vehículos.

Principalmente, la metodología expuesta para estimación de matrices de movilidad está enfocada a un uso por parte de la Administración. Es ella la encargada de optimizar el uso de la infraestructura terrestre desde el punto de vista del usuario (en su utilización diaria) así como desde el punto de vista de las inversiones a realizar para la adecuación de estas infraestructuras a las necesidades previstas y pronosticadas en escenarios futuros. Con técnicas tradicionales, los datos de movilidad obtenidos resultan en ocasiones poco fiables. Por ello, las cuantiosas inversiones que realiza para afrontar sus objetivos pueden verse justificadas al obtener unas matrices O-D ajustadas con datos reales y actuales procedentes de telefonía celular. Estos datos de telefonía, dado su extendido uso en la sociedad española actual con una cuota de penetración de 110.8% (CMT, Enero 2009) y teniendo Vodafone una cuota de mercado por clientes del 31.2% (CMT, Enero 2009), representan a una muestra significativa de la población, contribuyendo con ello a la mejora de los estudios de movilidad.

No obstante, sus resultados beneficiarán no sólo a la Administración, sino a usuarios del sistema de transporte (tanto privado como público) y a la sociedad en su conjunto. La Tabla 6-1 muestra

otras aplicaciones adicionales que se pueden derivar de un sistema de información de este tipo, algunas de las cuales ya fueron comentadas en el capítulo del “Estado-del-Arte”.

**Tabla 6-1: Ejemplos de Aplicaciones según mercados de negocio.**

<b>Sector</b>	<b>Aplicaciones</b>
Administración y otras autoridades públicas	Matrices Origen Destino Volúmenes y flujos Vigilancia Congestión e Incidentes Estudios de tiempos de viaje y/o velocidades Estudios de planificación Estudios de contaminación
Compañías de logística y flotas	Gestión y planificación de rutas
Proveedores de servicios LBS	Direccionamiento guiado
Consultoras de Transporte	Vigilancia de congestión Estudios de tiempos de viaje Planificación Contaminación atmosférica Estudios de transporte
Empresas de Telecomunicaciones	Servicios de información en tiempo real (sms, WAP)
Marketing	Campañas, ubicaciones estratégicas, etc

### 6.1.3.2 Rentabilidad

En 2006 la Unión Europea (UE) aprobó definitivamente la directiva 2006/24/CE sobre retención de datos telefónicos y de comunicaciones electrónicas. Dicha normativa obliga a los operadores a retener ciertos datos durante un periodo de entre 6 y 24 meses para la lucha contra el terrorismo y otros delitos graves. Todas las llamadas telefónicas deberán ser registradas, tanto las recibidas como las perdidas o no contestadas. La citada directiva también afectará a la retención de los datos correspondientes a la localización del lugar desde el que se realizan las llamadas desde teléfonos móviles.

Según los cálculos de las autoridades europeas, la creación de una base de datos con todas las llamadas efectuadas y recibidas por cada usuario, junto con la infraestructura necesaria, le supondrá a cada compañía una inversión que superará los 100 millones de euros, sin incluir el gasto anual asociado a su mantenimiento. Esta infraestructura para almacenar los datos puede utilizarse adicionalmente para otro tipo de aplicaciones. En concreto, según lo comentado en párrafos anteriores, sería de gran utilidad en aplicaciones relacionadas con los sistemas inteligentes de transporte (ITS), los cuales requieren datos de localización anónimos para efectuar sus estimaciones y ajustes. Con ello, el gasto para desplegar esta nueva infraestructura impuesta por la Unión Europea podría verse compensado con la distribución de este nuevo

sistema de información de tráfico, el cual resultaría muy llamativo para la Administración, que se encarga de realizar las inversiones para mejorar la gestión y planificación de la red.

El coste medio de una encuesta (encuesta+procesado) para estimar matrices O-D oscila alrededor de los 30 €. Para ajustar las matrices O-D de modo apropiado, estas encuestas deben efectuarse sobre una muestra representativa de la población; por ello, su coste resulta muy elevado. A ello se le une el coste temporal, debido a que el tratamiento de los datos de las encuestas supone un largo proceso. Por ello estas encuestas suelen actualizarse con los datos recogidos por los tradicionales aforos. El emplear información derivada de la telefonía para actualizar o, incluso, generar estas matrices sería una alternativa de especial interés para la Administración puesto que se conseguirían revolucionar los estudios de movilidad desde el punto de vista de coste (utilizaría la infraestructura existente de telefonía celular), fiabilidad (son datos de movilidad reales) y temporal, pudiendo dar respuesta inmediata a los problemas derivados de ella.

### **6.1.3.3 Valor Añadido para la Sociedad de la Información**

Como ya se ha comentado, una matriz O-D obtenida de una forma automática y, en cierto modo, inmediata supondría una auténtica revolución en los estudios de movilidad encargados por la Administración. Junto a ello, podría dar también una respuesta inmediata a los problemas derivados de la movilidad, permitiendo abordar temas a los que hasta ahora ha dedicado pocos recursos, como la movilidad en fin de semana, movilidad motivo ocio, movilidad a pie, etc.

Asimismo, este sistema de información puede solventar otros problemas derivados del incremento de tráfico experimentado en la mayoría de las áreas metropolitanas de los países desarrollados como la saturación, la seguridad y la escasez de nuevos servicios. Por ello, sector público y privado deben trabajar de manera conjunta para introducir estas herramientas de automatización, fomentando la integración de las nuevas tecnologías en nuestra vida cotidiana.

Adicionalmente, este tipo de desarrollo en I+D+i aportaría un retorno proveniente de la imagen social que se generaría en la colaboración operador-organismo público para la incorporación de nuevas tecnologías en la mejora de la planificación de la movilidad mecanizada.

## **6.2 LÍNEAS FUTURAS**

Tras la realización de esta Tesis, y teniendo en cuenta el “Estado-del-Arte” y el “Estado-de-la-Práctica” en el momento de su terminación, así como las conclusiones anteriormente expuestas, el siguiente párrafo presenta unas sugerencias para posibles futuras líneas de investigación:

- Perfeccionamiento de los modelos de predicción de volúmenes de vehículos propuestos.

En primer lugar, resulta prometedora la continuación directa de la línea de investigación de la Tesis mediante el diseño de modelos más robustos para estimar los volúmenes debido a que las llamadas en movilidad presentan una fuerte componente de aleatoriedad y un reducido tamaño de la muestra respecto a la población. Se sugiere un estudio más completo que tenga en cuenta patrones de movilidad asociados a las vías que conforman la frontera entre celdas así como sus características topológicas (cercanía a una ciudad, polígonos en sus alrededores...).

A continuación se proponen nuevas líneas directamente relacionadas con la Tesis.

- Investigación y desarrollo de modelos de estimación de volúmenes de tráfico basados en datos del cambio de área de localización generados por teléfonos que simplemente se encuentran encendidos.
- Estudio de otros parámetros de tráfico (velocidades, tiempos de viaje, densidad...).
- Desarrollo de una metodología para la elaboración de matrices a nivel de macrozonas.
- Desarrollo de modelos de identificación de peatones para el filtrado y depuración de la muestra con terminales válidos.
- Análisis de clusterización entre curvas de intensidades de vehículos proporcionados por estaciones de aforo e intensidades inferidas de la telefonía móvil.





# BIBLIOGRAFÍA

---

Abrahamsson T. Estimation of origin-destination matrices using traffic counts – a literature survey. Interim report, *International Institute for Applied Systems Analysis*, A-2361. Laxenburg, Austria, 1998.

Ahas R., Ülar M. Location based services–new challenges for planning and public administration?. *Futures*, vol. 37(6), 547-561, 2005

AirSage. *Using Cellular Technology to Provide an Intelligent Traffic Roadmap*. Disponible on-line (Acceso Junio, 2007): [http://www.airsage.com/pdf/trafficwise\\_release.pdf](http://www.airsage.com/pdf/trafficwise_release.pdf)

Akin D., Sisiopiku V. P. Estimating origin-destination matrices using location information from cell phones. *Proc. of 49th Annual North American Meetings of the Regional Science Association International*, Puerto Rico, 2002.

Bar-Gera H. Evaluation of a cellular phone-based system for measurement of traffic speeds and travel times: A case study from Israel. *Transportation Research Part C*, vol. 15(6), 380–391, 2007.

Bar-Gera H. *Transport Network Test Problems*. Ben-Gurion University of the Negev. Disponible online (Acceso Mayo, 2007b): <http://www.bgu.ac.il/~bargera/tntp/>

Bazaraa M. S., Shetty C. M. *Nonlinear Programming: Theory and Algorithms*. John Wiley & Sons, Inc, 1979.

Beckmann M., McGuire C., Winsten C. *Studies in the Economics of Transportation*. Yale University Press, New Haven, Connecticut, 1956.

Bekhor S., Hirsh M., Nimre S., Feldman I. Identifying Spatial and Temporal Congestion Characteristics using Passive Mobile Phone Data. *Transportation Research Board 87th Annual Meeting*, CD-ROM paper no. 1534, 2008.

Bell M.G.H. The estimation of an origin-destination matrix by constrained generalized least squares. *Transportation Research* 25, 13–22, 1991.

Birle C., Wermuth M. The Traffic Online Project. Special Session: Cellular-based traffic data collection. SS45 (EU). *13th World Congress on ITS*, London, United Kingdom, 2006.

Bolla R., Davoli F., Giordano F. Estimating road traffic parameters from mobile communications. *7th World Congress on ITS*, Torino, Italy, 2000.

Caceres N., Benitez F.G., Romero L.M., Vázquez J., Wideberg J. *Herramienta para la inferencia y actualización de matrices O-D de movilidad a través de datos de telefonía celular*. Informe Final FOM/486/2003, 2005.

Caceres N., Wideberg J. P., Benitez F. G. Deriving origin-destination data from a mobile phone network. *IET Proc. Intelligent Transport Systems*, vol. 1(1), pp. 15–26, 2007.

Caceres N., Wideberg J., Benitez F.G. Review of traffic data estimations extracted from cellular networks, *IET Proc. Intelligent Transport Systems*, vol. 2(3), pp179–192, 2008.

Cascetta E. Estimation of trip matrices from traffic counts and survey data: A generalized least squares estimator, *Transportation Research*, 18B, 1984.

Cascetta E., Nguyen S. A unified framework for estimating or updating origin/destination matrices from traffic counts. *Transportation Research*, 22B. 437-455, 1988.

Cayford R., Johnson T. Operational Parameters Affecting Use of Anonymous Cell Phone Tracking for Generating Traffic Information. *Transportation Research Board 82th Annual Meeting*, vol. 1(3):03–3865, 2003.

CMT (Comisión del Mercado de las Telecomunicaciones). *Informe Telefonía Móvil*. Enero, 2009.

CDB, Communities Dominate Brands. *A mobile phone for every living person in Western Europe: penetration hits 100%*. Disponible on-line (Acceso Octubre, 2007): [http://communities-dominate.blogs.com/brands/2006/03/phone\\_for\\_every.html](http://communities-dominate.blogs.com/brands/2006/03/phone_for_every.html)

CRT, Consorcio Regional de Transportes de la Comunidad de Madrid, *Modelización de la Movilidad en Día Laborable de 2004 en la Comunidad de Madrid*, 2004.

Chen Y., Florian M. O–D demand adjustment problem with congestion: Part I. model analysis and optimality conditions. In L. Bianco and P. Toth (Eds.), *Advanced Methods in Transportation Analysis*, pp. 1–22. Berlin: Springer-Verlag, 1996.

Decell AutoRoute1 – *Real-Time Route Guidance*. Disponible on–line (Acceso Junio, 2006): [http://www.decell.com/autoroute1/real\\_time.htm](http://www.decell.com/autoroute1/real_time.htm)

Deheuvels P. La Fonction de Dependence Empirique et ses Proprietés. Un test non parametrique d'independance. *Academie Royale de Belgique Bulletin de la Classe des Sciences-5e*, 65: 274–292, 1979.

Delcan. Press Release. *Delcan.NET and ITIS Holdings in Partnership*. Disponible on–line (Acceso Abril, 2005): <http://www.itisholdings.com/pdf/delcanitis290405.pdf>

Doblas F.J., Benitez F.G., 2005. An Approach to Estimating and Updating Origin-Destination Matrices Based Upon Traffic Counts Preserving the Prior Structure of a Survey Matrix. *Transportation Research Part B–Methodological*, vol. 39, no. 7, pp. 565–591.

EIA/TIA IS-95. *Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System*, 1993.

EMME/2. *User's Manual, Software Release 9*. INRO Consultants Inc, 1998.

Gavish B., Sridhar S. The impact of Mobility on Cellular Network Configuration, *4th Intern. Conf. on Telecommunication Systems: Modelling and Analysis*, pp. 197–211, ATSSMA, 1996.

Gibson J.D. (ed). *The mobile communications handbook, Second Edition*. CRC Press & IEEE Press, 1999.

Hellinga B., Fu L., Takada H. Traffic network condition monitoring via mobile phone location referencing – an emerging opportunity. *Proc. of the 6th Transportation Specialty Conference held in conjunction with the Annual Conference of the Canadian Society of Civil Engineers*, Toronto, Canada, 2005.

Hernando Rábanos J.M. *Sistemas de Comunicaciones Móviles de Tercera Generación IMT–2000 (UMTS)*. Ed. Fundación Airtel-Vodafone, 2002.

Herrera J., Bayen A. Traffic flow reconstruction using mobile sensors and loop detector data. *Transportation Research Board 87th Annual Meeting*, CD-ROM paper no. 1868, 2008.

Hong D., Rappaport S.S. Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures, *IEEE Trans. Veh. Technol.*, vol. VT–35, no. 3, pp. 448–461, 1986.

Höpfner M., Lemmer K., Ehrenpfordt I. Cellular Data for Traffic Management – First Results of a Field Test. *ITS Europe Conference*, CD-ROM paper no. 2407, 2007.

Huidobro J.M. *Comunicaciones Móviles*. Ed. Paraninfo, 2002.

Ilyas M., Mahgoub I. (Editors). *Mobile Computing Handbook*. CRC Press, 2005.

ITIS Holdings – Press Releases. *Estimation Technology chosen for traffic collection in Antwerp*. Disponible online (Acceso Septiembre 2004): <http://www.itisholdings.com/pdf/itisproximus.pdf>

Kummala J. *Travel time service utilising mobile phones*. Finnish Road Administration, Helsinki. Finnra Report 55/2002. 67pp, 2002.

LeBlanc L.J. An Algorithm for the Discrete Network Design Problem. *Transport Science*, 9(3), pp. 183–199, 1975.

Linauer M., Leihs D. Generating floating car data by using GSM–network. *Proc. of the 10th World Congress and Exhibition on Intelligent Transport Systems and Services*. Madrid, 2003.

Lovell D. J. Accuracy of speed measurements from cellular phone vehicle location systems”. *ITS Journal*, vol. 6, no. 4, pp. 303–325, 2001.

Low D.E. A new approach to transport system modelling. *Traffic Quarterly*, vol. 26, no. 3, pp. 391–404, 1972.

Maher M.J. Inferences on trip matrices from observations on link volumes: A Bayesian statistical approach. *Transportation Research*, 17B. 435–447, 1983.

MINUTP. *Technical User Manual*. Hague Consulting Group, HCG. V.93A, 1993.

MTS (Mobile Traffic Service), LogicaCMG. *LogicaCMG provides reliable and comprehensive traffic information based on GSM network data*. Disponible on–line (Acceso Noviembre, 2003): <http://www.logica.com/logicacmg+provides+reliable+and+comprehensive+traffic+information+based+on+gsm+network+data/400002401>

Nelsen R. B. *An Introduction to Copulas (Second Edition)*. Springer Series in Statistics. Springer, New York, 2006.

Pan C., Lu J., Di S., Ran B. Cellular-Based Data-Extracting Method for Trip Distribution. *Transportation Research Record: Journal of the Transportation Research Board* No. 1945, pp. 33–39, 2006.

Patriksson M. *The Traffic Assignment Problem: Models and Methods*. VSP BV, 1994.

Pyo J.S., Shin D.H., Sung T.K. Development of a Map Matching Method Using the Multiple Hypothesis Technique. *IEEE Intelligent Transportation Systems Conference Proceedings*. Oakland, Calif., pp. 23–27, 2001.

Rappaport T.S. *Wireless Communications: Principles and Practice, Second Edition*. Prentice Hall, New Jersey, 2002.

Ratti C., Pulselli R.M., Williams S., Frenchman D. Mobile Landscapes: using location data from cell-phones for urban analysis. *Environment and Planning B: Planning and Design*, vol. 33(5), pp 727-748, 2006.

Ratti C., Sevtsuk A., Huang S., Pailer R. Mobile Landscapes: Graz in Real Time. *Proceedings of the 3rd Symposium on LBS & TeleCartography*. Vienna, Austria, 2005.

Reades J., Calabrese F., Sevtsuk A., Ratti C. Cellular Census: Explorations in Urban Data Collection. *IEEE Pervasive Computing*, vol. 6, no. 3, pp. 30–38, 2007.

Reklaitis G.V., Ravindran A., Ragsdell K.M. *Engineering Optimization: Methods and Applications*. John Wiley & Sons, Inc, 1983.

Rose G. Mobile phones as traffic probes: practices, prospects and issues. *Transport Reviews*, vol. 26, no. 3, 275–291, 2006.

Sallent O., Valenzuela J.L., Agustí R. *Principios de Comunicaciones Móviles*. Ediciones UPC, Barcelona, España, 2003.

Sheffi Y. *Urban Transport Networks: Equilibrium Analysis with Mathematical Programming Methods*. Prentice-Hall, Inc., 1985.

Smith B.L., Fontaine M.D. *Private-Sector Provision of Congestion Data*. NCHRP 70-01, University of Virginia, 2006.

Smith B.L., Pack M.L., Lovell D.J., Sermons M.W. Transportation management applications of anonymous mobile call sampling. *Proceedings of the 11th Annual Meeting of ITS America*, Miami, Florida, 2001.

Smith Institute. *Real-time traffic monitoring using mobile phone data*. Disponible on-line (Acceso Mayo, 2006): [www.smithinst.ac.uk/Projects/ESGI49/ESGI49-Vodafone/Report/TrafficMonitoring.pdf](http://www.smithinst.ac.uk/Projects/ESGI49/ESGI49-Vodafone/Report/TrafficMonitoring.pdf)

Sohn K. *Dynamic estimation of origin-destination flows using cell phones as probes*. SDI 2004-R-04, Dept. of Urban Transportation, Seoul Development Institute, Korea, 2004.

Spieß H. A maximum likelihood model for estimating origin-destination matrices. *Transportation Research*, 21B. 395–412, 1987.

Spieß H. *A Gradient Approach for the O-D Matrix Adjustment Problem*. Publication no. 693, Centre for Research on Transportation, University of Montreal, 1990.

Steele R., Nofal M. Teletraffic performance of microcellular personal communication networks, *IEE Proceedings-I*, vol. 139, no. 4, pp. 448–461, 1992.

Suwansirikul C., Friesz T.L., Tobin R.L. Equilibrium decomposed optimization: A heuristic for the continuous equilibrium network design problem. *Transport Science*, 21, pp. 254–263, 1987.

Thiessenhusen K.-U., Schäfer R.-P., Lang T. Traffic data from cell phones: a comparison with loops and probe vehicle data. *13th World Congress on ITS*, CD-ROM, paper no. 1550. London, United Kingdom, 2006.

Thornton J., Grace D. Effect of Lateral Displacement of a High Altitude Platform on Cellular Interference and Handoff. *IEEE Trans. Wireless Communications*, vol. 4(4), pp. 1483–1490, 2005.

TRANSCAD. *User's Guide*, Version 4.8. Newton, MA, Caliper Corporation, 2006.

TRIPS. *MVESTM program*. INSMVA Systematica, 1990.

Van Zuylen H.J., Willumsen L.G. The most likely trip matrix estimated from traffic counts. *Transportation Research*, 14B, 281–293, 1980.

VDOT. *Cell phones could keep you out of traffic jams*. Disponible on-line (Acceso Septiembre, 2005): <http://www.airsage.com/pdf/news-4-27-05.pdf>

Virtanen J. *Mobile phones as probes in travel time monitoring*. Finnish Road Administration, Helsinki, 2002.

White C.E., Bernstein D., Kornhauser A.L. Some Map Matching Algorithms for Personal Navigation Assistants. *Transportation Research Part C*, vol. 8, pp. 91–108, 2000.

White J.K., Gower P. *Collecting origin-destination information from cellular phone location data*. TRL Unpublished Report, 1999.

White J., Wells I. Extracting origin destination information from mobile phone data. *11th International Conference on Road Transport Information and Control* (Conf. Publ. No. 486). pp. 30–34, 2002.

Wiltshko T., Schwieger V, Möhlenbrink W. Generating Floating Phone Data for Traffic Flow Optimization. *Proceedings of the 3rd International Symposium Networks for Mobility*, Stuttgart, Germany, 2006.

Yang H., Sasaki T., Iida Y., Asakura Y. Estimation of origin-destination matrices from link traffic counts on congested networks. *Transport Research* 26B, 6, pp. 417–434, 1992.

Ygnace J.-L. *Travel time/speed estimates on the French Rhone corridor network using cellular phones as probes*. Final report of the SERTI V program, INRETS, Lyon, France, 2001.

Yim Y. *The state of cellular probes*. California PATH Research Report: UCB-ITS-PRR-2003-25, 2003.

Zeng Q-A., Mukumoto K., Fukuda K. Performance analysis of mobile cellular radio systems with two-level priority reservation handoff procedure, *IEICE Trans. Commun.*, vol. E80-B, no. 4, pp. 598–604, 1997.





# APÉNDICE I

## GLOSARIO

### A

AIC: Akaike Information Criterion (Criterio de Información de Akaike).

AMPS: Advanced Mobile Phone System.

AOA: Angel Of Arrival.

AuC: Authentication Center.

### B

BER: Bit Error Rate.

BH: Busy Hour (hora de mayor ocupación).

BIC: Bayesian Information Criterion (Criterio de Información Bayesiano).

BPR: Bureau of Public Roads.

BSC: Base Station Controller.

BSS: Base Station Sub-System.

BTS: Base Transceiver Station.

## C

CC: Country Code.

CGI: Cell Global Identity.  $CGI = LAI + CI$ .

CI: Cell Identity.

CIR: Carrier to Interference Ratio.

CMT: Comisión Mercado de las Telecomunicaciones.

CSPDN: Circuit Switched Public Data Network.

CV: Coeficiente de Variación.

## E

EIR: Equipment Identity Register.

## F

FCD: Floating Car Data.

F-W: Frank-Wolfe.

## G

GMSC: Gateway Mobile Services Switching Center.

GPRS: General Packed Radio Service.

GSM: Global System for Mobile Communications.

## H

HLR: Home Location Register.

HO: Handover.

## I

IMD: Intensidad Media Diaria.

IMEI: International Mobile Equipment Identity.

IMSI: International Mobile Subscriber Identity.  $IMSI = MCC + MNC + MSIN = MCC + NMSI$ .

ISDN: Integrated Services Digital Network.

ITS: Intelligent Traffic Systems.

## L

LA: Location Area (área de localización).

LAC: Location Area Code.

LAI: Location Area Identification.  $LAI = MCC + MNC + LAC$ .

LLF: Log-Likelihood Function.

LMSI: Local Mobile Station Identity.

LPR: License Plate Recognition.

LU: Location Update.

## M

MCC: Mobile Country Code.

ME: Mobile Equipment.

MNC: Mobile Network Code.

MS: Mobile Station.

MSC: Mobile services Switching Center.

MSIN: Mobile Station Identification Number.

MSISDN: Mobile Subscriber ISDN Number.  $MSISDN = CC + NDC + SN$

MSRN: Mobile Station Roaming Number.  $MSRN = CC + NDC + SN$

MT: Mobile Termination.

## N

NDC: National Destination Code.

NMC: Network management Center.

NMSI: National Mobile Station Identity.

NSS: Network Sub-System

## O

O-D: Origen-Destino.

OMC: Operation and Maintenance Center.

OMC-N: Operation and Maintenance Center National.

OMC-R: Operation and Maintenance Center Regional.

## P

PDU: Protocol Data Unit.

PIN: Personal Identification Number.

PLMN: Public Land Mobile Network.

PLU: Periodic Location Update.

PSPDN: Packet Switched Public Data Network.

PSTN: Public Switched Telephone Network.

## R

RA: Routing Area.

RAI: Routing Area Identification.

RLC: Radio Link Control.

RLP: Radio Link Protocol.

RSSI: Radio Signal Strenght Indicator.

RTB: Red Telefónica Básica.

## S

SIM: Subscriber Identity Module.

SN: Subscriber Number.

## T

TAP: Traffic Assignment Problem.

TAZ: Traffic Analysis Zone.

TDOA: Time Difference Of Arrival.

TMSI: Temporary Mobile Subscriber Identity.

## U

UMTS: Universal Mobile Telecommunications System.

## V

VLR: Visitor Location Register.

VMS: Variable Message Signs.

# APÉNDICE II

## PUBLICACIONES

Algunos de los resultados obtenidos en esta Tesis se encuentran recogidos en las publicaciones que se listan a continuación, junto con otras aportaciones también relacionadas con ella.

### PUBLICACIONES EN REVISTAS

- Caceres N., Wideberg J.P., Benitez F.G. Deriving origin–destination data from a mobile phone network. *IET Proc. Intelligent Transport Systems*, vol. 1(1), pp. 15–26, 2007.
- Caceres N., Wideberg J., Benitez F.G. Review of traffic data estimations extracted from cellular networks, *IET Proc. Intelligent Transport Systems*, vol. 2(3), pp. 179–192, 2008.
- A Bi–Level Approach to Estimating Origin-Destination Trip Matrices from Cellular Phone Data preserving the Structure of Prior Matrices' Information. (En revisión 2009).
- Forecasting Traffic Flow Using Anonymous Mobile Phone Data. (En revisión 2009).

### CONGRESOS INTERNACIONALES

- Caceres N., Wideberg J. P., Benitez F. G. Deriving Traffic Data from a Cellular Network. *13th World Congress on Intelligent Transport System and Services*. London, United Kingdom. 2006.

## CONGRESOS NACIONALES

- Cáceres N., Romero L. M., Vázquez J., Wideberg J. P., Benítez F. G. Inferencia y Actualización de Matrices O-D de Movilidad a Través de Datos de Telefonía Celular. *VII Congreso de Ingeniería del Transporte*. Ciudad Real, España, 2006. CIT 2006.
- Cáceres N., Wideberg J. P., Benítez F. G. Monitorización de la Movilidad y el Tráfico Mediante Datos de Telefonía Móvil. *VII Congreso Español Sobre Sistemas Inteligentes de Transporte*. Valencia. ITS España. 2007.
- Cáceres N., Romero L.M., Benítez F. G., Wideberg J. P. Metodología Bi-nivel para la Estimación de Matrices de Movilidad con Datos Agregados de Telefonía Móvil. *VIII Congreso de Ingeniería de Transporte*. A Coruña, España, 2008. CIT 2008.

## OTRAS PUBLICACIONES

- Cáceres N., Benítez F.G., Romero L.M., Vázquez J., Wideberg J. *Herramienta para la inferencia y actualización de matrices O-D de movilidad a través de datos de telefonía celular*. Informe Final FOM/486/2003. ISBN: 84-88783-77-9, 2005.

## OTROS MÉRITOS

*I Premio Vodafone al Proyecto de Desarrollo e Innovación Móvil en el ámbito de la Investigación*, entregado por la **Fundación Vodafone España** en 2007 al equipo formado por Noelia Cáceres Sánchez, Francisco García Benítez y Johan P. Wideberg del grupo de Ingeniería e Infraestructura de los Transportes (Escuela Técnica Superior de Ingenieros, Universidad de Sevilla), por el proyecto “**Uso de datos de una Red GSM para la inferencia de datos de movilidad interurbana de vehículos**”.



