

Obtención y análisis de aforos desagregados de transportes de mercancías peligrosas mediante sistemas ANPR

F. J. Morales¹, L.M. Romero¹, N. Cáceres², A. Reyes¹, F. G. Benítez¹

¹Ingeniería e Infraestructura de los Transportes, Universidad de Sevilla, España.

²Ingeniería de los Transportes, AICIA, España.

RESUMEN

La fuente de información disponible, a nivel nacional, sobre flujos de tráfico que involucran mercancías peligrosas se encuentra publicada anualmente en el Mapa de Tráfico editado por el Ministerio de Fomento. Esta información, de carácter informativo, es presentada de forma agregada; obtenida mediante aforos manuales que se capturan en 6 campañas anuales, en días laborables de meses alternos y en un periodo temporal limitado a seis horas. Los datos no revelan la distribución de la tipología de mercancía peligrosa transportada.

Esta distribución sí puede inferirse haciendo uso de sistemas ANPR (Automatic Number Plate Recognition), mediante la captura de datos durante las 24 horas diarias, para los 365 días del año.

El sistema y metodología que se reporta en este estudio demuestra la capacidad de estimación del flujo de vehículos y tipología de materia peligrosa transportada mediante la lectura del código ONU que identifica de forma unívoca la tipología de ésta y que se encuentra incluido dentro de la placa naranja de identificación de peligro que obligatoriamente deben llevar los vehículos que realizan el transporte de este tipo de mercancía. El sistema se encarga de la interpretación de la materia transportada y el peligro asociado, así como de la posición y dirección de dicho vehículo, registrando la información en un servidor central. El sistema está constituido por un conjunto distribuido de estaciones de lectura ubicadas en puntos de observación estratégicos.

Con la información capturada con este sistema automático, se realiza una comparativa con la fuente de información agregada, para derivar conclusiones sobre el grado de utilidad de la información publicada.

Con la inferencia de la información desagregada es posible realizar una multiplicidad de análisis tales como, clasificación de materias por su importancia en cuanto al volumen transportado, identificación de los puntos con un mayor riesgo medioambiental, y actualización de matrices de origen-destino de mercancías peligrosas, entre otros.

1. INTRODUCCIÓN

La universalización de determinadas tecnologías, que se han ido desarrollando y experimentando en proyectos pilotos, de identificación automatizada de mercancías peligrosas durante las etapas de su transporte (Yu, M. et al, 2012; Liu, Y. et al, 2012), van a facilitar conocer en un horizonte próximo, de forma exacta el volumen de MPs que circulan por las carreteras. La identificación de estos flujos permite prever las medidas más adecuadas

para proteger la población y el entorno medioambiental del riesgo asociado con este tipo de transporte. Hasta no alcanzar este horizonte, resulta fundamental obtener estimaciones cada vez más fiables de los citados flujos que circulan por las carreteras. Este trabajo ha supuesto un paso adelante en esta dirección al hacer uso de un sistema de adquisición de imágenes para identificar vehículos que transportan MPs. El resultado y objetivo práctico alcanzado ha sido el diseño de una aplicación informática eminentemente interactiva, basada en metodologías de estimaciones de máxima verosimilitud, capaz de analizar automáticamente los datos capturados por el sistema de reconocimiento automático de visión artificial, mostrando de forma gráfica e intuitiva la información más relevante para el usuario.

2. DESCRIPCIÓN DEL SISTEMA ANPR

El sistema ANPR se encarga de la interpretación de la materia transportada y el peligro asociado, así como de la posición y dirección del vehículo, registrando la información en un servidor central.

Este sistema está formado por un conjunto distribuido de estaciones de lectura ubicadas en puntos de observación estratégicos, así como de un puesto de operación, todo ello interactuando con un servidor central que almacena los datos. El criterio de selección de puntos de aforo se ha realizado en base a:

- Observación del mayor flujo posible, de vehículos de MPs caracterizado por los datos agregados, derivado de la encuesta manual publicada; o bien del flujo general de vehículos en tramos de interés específico.
- Tramos de la red RIMP (DGT, 2016), que correspondan con rutas que conecten el mayor número posible de pares origen-destino.
- Existencia de pórticos accesibles para la instalación de las cámaras ANPR.

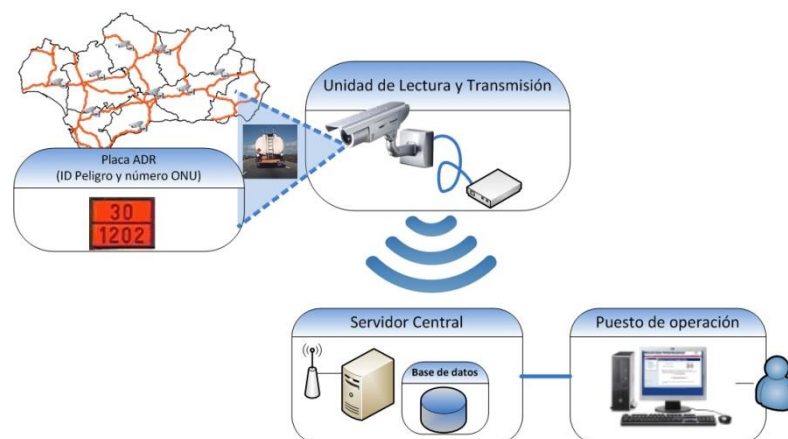


Figura 1: Arquitectura del sistema

La arquitectura del sistema (Figura 1) es modular, con los siguientes elementos físicos y lógicos:

- **Unidad de Lectura y Procesado (ULP).** Cada unidad cuenta con una cámara con capacidad de obtener el código de peligro y número ONU de la placa ADR en diferentes condiciones de iluminación, y una sub-unidad de transmisión integrada. Lleva a cabo el reconocimiento e interpretación de la información de la placa, para posteriormente transmitir dicha información a un Servidor Central. Esta unidad se encarga de la generación de los conteos de flujo desagregado, por clase de mercancía, en base a la información capturada por la cámara de visión artificial.
La ULP reconoce automáticamente las placas de identificación de peligro, de color naranja, retroreflectantes y con un borde limitador negro (Figura 2). Las placas tienen unas dimensiones normalizadas (30x40 cm), permitiendo la existencia de variaciones no estándares que el sistema es capaz de reconocer; contienen en su interior dos grupos de números de identificación que son reconocidos por el sistema.
La ubicación habitual de estas placas es en la parte frontal y posterior del vehículo, aunque en ocasiones se sitúa también en los laterales. En determinados casos la placa puede no contener los números de identificación, siendo ésta una salida posible del algoritmo (panel de MP sin identificación de peligro ni materia). Tras completar el reconocimiento de los números estandarizados, la UPL ejecuta un análisis sintáctico de sus valores para validar la relación peligro/materia mediante el acceso a una base de datos con las correspondencias posibles, que permite contrastar los resultados del proceso de reconocimiento de caracteres. En cualquier caso, la UPL es capaz de reconocer vehículos de MPs y extraer números de placa ADR para posteriormente almacenar los contenidos en un repositorio donde estos datos están accesibles desde el puesto de operación para la consulta y/o explotación.
- **Servidor Central (SC).** Se trata de un ordenador de gran capacidad que aloja las bases de datos con la información transmitida por la red de ULPs, así como otra información complementaria: riesgos medioambientales geográficos, información de flujos anuales de mercancías, conteos de este tipo de vehículos, relaciones posibles peligro/materia, ubicación de las unidades de intervención en caso de accidente, etc. Así mismo, actualiza las matrices origen-destino oportunas mediante la contrastación de las diferentes fuentes de información (Romero et al., 2014).
- **Puesto de operación (PO).** Este elemento del sistema está destinado a la resolución manual de las incidencias por parte de un operador. La operativa permite conocer el estado del sistema y facilita el acceso para consulta de las bases de datos.

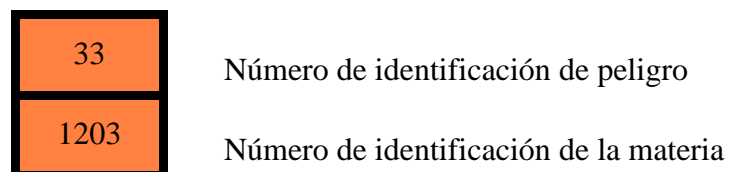


Figura 2: Placa ADR

El sistema proporciona, entre otros, los principales datos para cada registro:

- Para todos los vehículos:
 - instante de tiempo en el que se ha detectado un vehículo;
 - carril de circulación.

- Para vehículos con mercancías peligrosas:
 - instante de tiempo en el que se ha detectado un vehículo;
 - carril de circulación;
 - grupo de peligro;
 - número de identificación de peligro;
 - número de identificación de la materia.

Hay que hacer constar que la identificación del vehículo se encuentra anonimizada, limitándose la detección al caso de la tipología del vehículo (ligero, pesado) y, para aquellos que dispongan de placa ADR, sus identificadores ADR. El sistema permite la opción, no implementada, de instalación de un módulo de almacenamiento de imagen y lectura de placa de matrícula.

El sistema desarrollado permite el auto-aprendizaje, generando soluciones más fidedignas conforme las bases de datos, alimentadas por las cámaras de visión artificial ADR, van ampliándose a lo largo del tiempo con la información capturada. Ésta información puede hacerse más completa a través de una ampliación de la red cubierta por las cámaras, dando lugar a una mayor precisión y confianza de las predicciones. Al ser el sistema modular, es ampliable y adaptable. Este proyecto, piloto, ha limitado su frontera al uso de un número limitado de cámaras distribuidas entre las provincias de Cádiz y Sevilla, operando en un periodo de captación de datos limitado en el tiempo (2014-actualidad) con la finalidad de demostrar las capacidades del sistema desarrollado.

Los equipos instalados se encuentran enumerados en la Tabla 1 y marcados en el mapa de la Figura 3.

Ctra.	PK	Sentido	Prov.	Tipo	Nº Carriles	Nº Carriles Monitorizados
A-49	4,200	Decreciente	Sevilla	Pórtico	3	2
A-4	525,200	Creciente	Sevilla	Pórtico	2	2
A-66	783,100	Decreciente	Sevilla	Pórtico	2	2
A-4	551	Decreciente	Sevilla	Banderola	2	2
A-382	1,200	Creciente	Cádiz	Banderola	2	2

Tabla 1: Equipos instalados.



Figura 3: Ubicación de los equipos

2. METODOLOGÍA

El análisis de los datos es un paso previo al desarrollo de cualquier modelo de predicción o de cualquier metodología de monitorización. Actualmente, un enfoque común es la interpretación gráfica de los patrones temporales de tráfico recogidos a través de detectores (Chung, 2003; Song y Miller, 2012). Dentro de esta interpretación gráfica, uno de los análisis más recurrente es el obtener patrones diarios y su posterior clusterización (Chung, 2003; Guardiola et al., 2014).

La obtención fidedigna de clusterizaciones exige la detección y eliminación previa de datos atípicos (*outliers*); en la literatura publicada existen multitud de técnicas propuestas para llevar a cabo la depuración de *outliers* (Febrero-Bande et al., 2007), entre las que se encuentra el Análisis de Componentes Principales (PCA: Principal Component Analysis). El PCA (Pearson, K., 1901) es un procedimiento estadístico que usa una transformación ortogonal para convertir una serie de observaciones de variables posiblemente correladas en otro conjunto de valores linealmente incorrelados llamados componentes principales. El número de componentes principales es menor o igual al número de variables originales. Esta transformación se define de forma que la primera componente principal tiene la mayor varianza posible (es decir, es la componente que explica en mayor medida la varianza de los datos originales), y las sucesivas componentes a su vez tiene la mayor varianza posible bajo la restricción de que sean ortogonales a las componentes anteriores. De esta forma, los vectores resultantes (componentes principales) forman una base ortogonal, ya que son los autovectores de la matriz de covarianza, la cual es simétrica. De esta manera se pueden proyectar los datos originales en un nuevo subespacio formado por las componentes principales de forma que los datos transformados están referidos a nuevas coordenadas llamadas comúnmente “*scores*”. Normalmente, los primeros dos o tres “*scores*” sirven para

detectar grupos, *outliers* y patrones de comportamiento presentes en los datos.

La metodología PCA ha sido una de las utilizadas en este trabajo, y la que ha permitido alcanzar un resultado de mayor fiabilidad en la detección de *outliers* y la posterior clusterización de los datos capturados por el sistema desarrollado.

Los datos capturados por el sistema se almacenan en registros, individuales para cada uno de los vehículos captados en el punto de detección (cualquier tipo de vehículo, incluyendo los que portan placas ADR). Los registros se procesan para obtener agregaciones en periodos temporales (e.g. 5, 15, 30 y 60 minutos) para inferir los perfiles de tráfico diario. Se han utilizado procedimientos de ventanas deslizantes, con aperturas fijas cuantificadas a los periodos temporales indicados, para detectar la granularidad de mayor fiabilidad.

A continuación se presenta una esquematización de la analítica de datos llevada a cabo, a posteriori del procesado de los datos capturados por las cámaras. En primer lugar se realiza un filtrado de los datos para descartar errores en las mediciones. Con esto se consigue descartar días con falta de suministro eléctrico, fallo de la ULP, medidas erróneas por deslumbramiento, etc. Tras este primer filtrado se procede a graficar los datos para comprobar que los perfiles proporcionados por el sistema son adecuados (Figura 4).

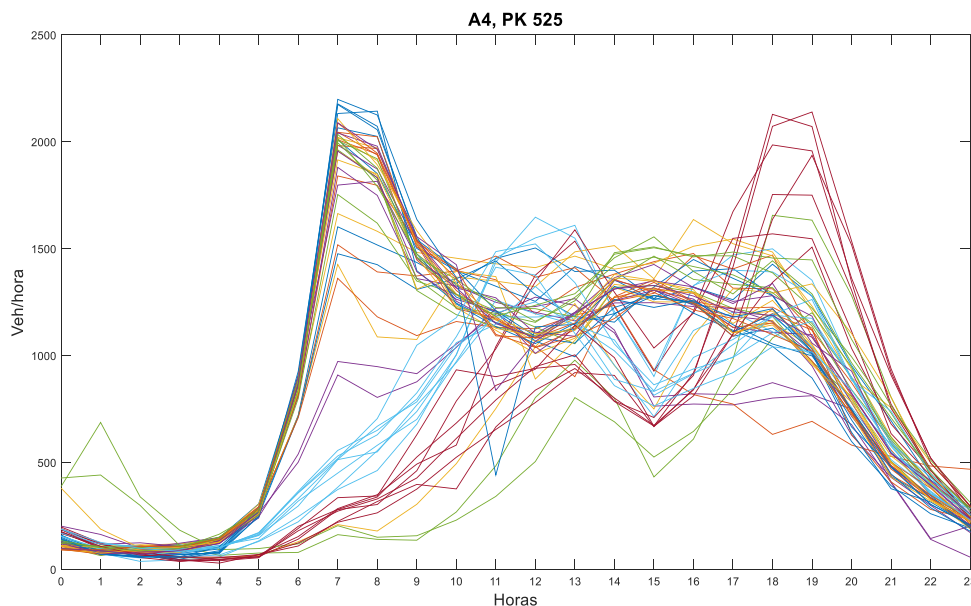


Figura 4: Datos agregados horarios de la cámara A4,PK 525

Los datos considerados para este ejemplo son los comprendidos entre el 18 de Diciembre de 2015 y el 13 de Febrero de 2016. Se ha considerado este periodo para comprobar que el proceso automático es capaz de detectar los días festivos y tratarlos separadamente del resto de días.

La Figura 5 muestra los resultados obtenidos después de la aplicación de la metodología PCA, en la que se aprecia claramente el agrupamiento de los datos de forma que pueden distinguirse los días entre semana (marcados como L, M, X, J, V), los sábados (S), los Domingos (D) y los días especiales (F: festivos, -F: previos a festivo, y +F: posterior a

festivo).

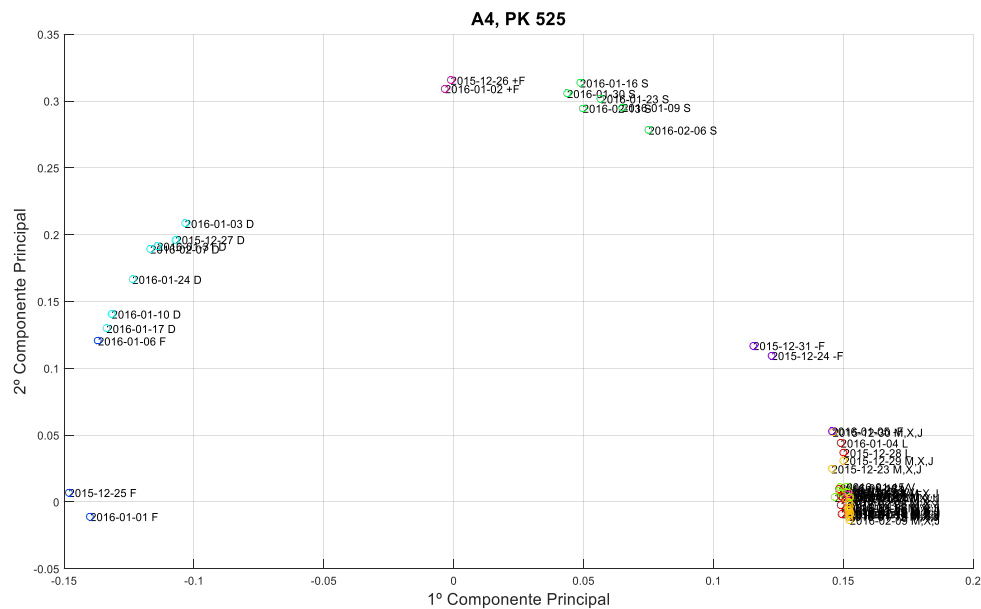


Figura 5: 1ª y 2ª Componente Principal

Centrándose en el cluster asociado a los días entre semana (Figura 6) se puede apreciar la diferencia entre estos días dentro y fuera del periodo vacacional de Navidad.

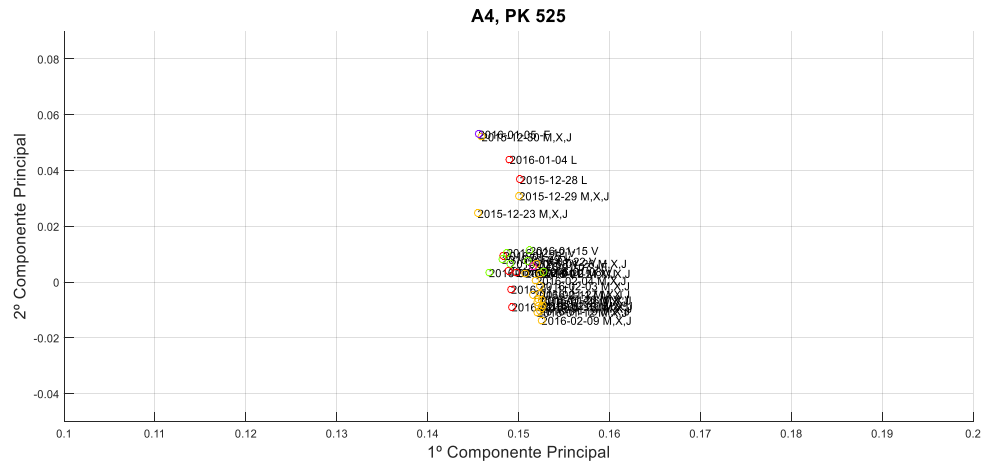


Figura 6: 1ª y 2ª Zoom Componentes Principales en días entre semana

Aunque visualmente, Figura 5, pueden apreciarse los distintos días que deben pertenecer a cada grupo, el proceso exige ser automatizado computacionalmente; exigiendo llegar a definir una operativa lo suficientemente fiable para la determinación del número óptimo de grupos presentes en el conjunto de datos. La literatura publicada recoge una gran variedad de índices y criterios que facilitan la obtención del número de grupos, definiendo la clusterización óptima (Milligan y Cooper, 1985). De entre ellos, el índice más efectivo es el Calinski-Harabasz's (CH), que evalúa la validez del grupo basándose en el error cuadrático medio entre grupos, de forma que un valor elevado del índice CH indica una buena

clusterización. En la Figura 7 se muestra un ejemplo de la variación del índice CH para los datos capturados por una de las cámaras.

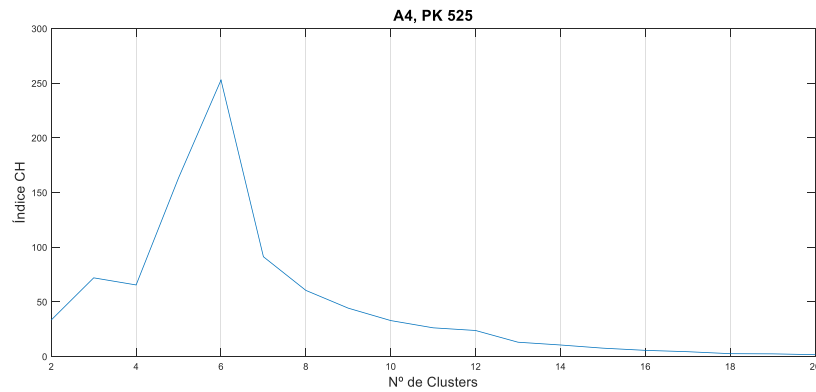


Figura 7: Índice CH

Finalmente, definido el número óptimo de grupos, se procede a clusterizar los datos para comprobar que la agrupación alcanzada es correcta. En la Figura 8 se aprecia como realmente cada uno de los grupos obtenidos se corresponde con un tipo distinto de día.

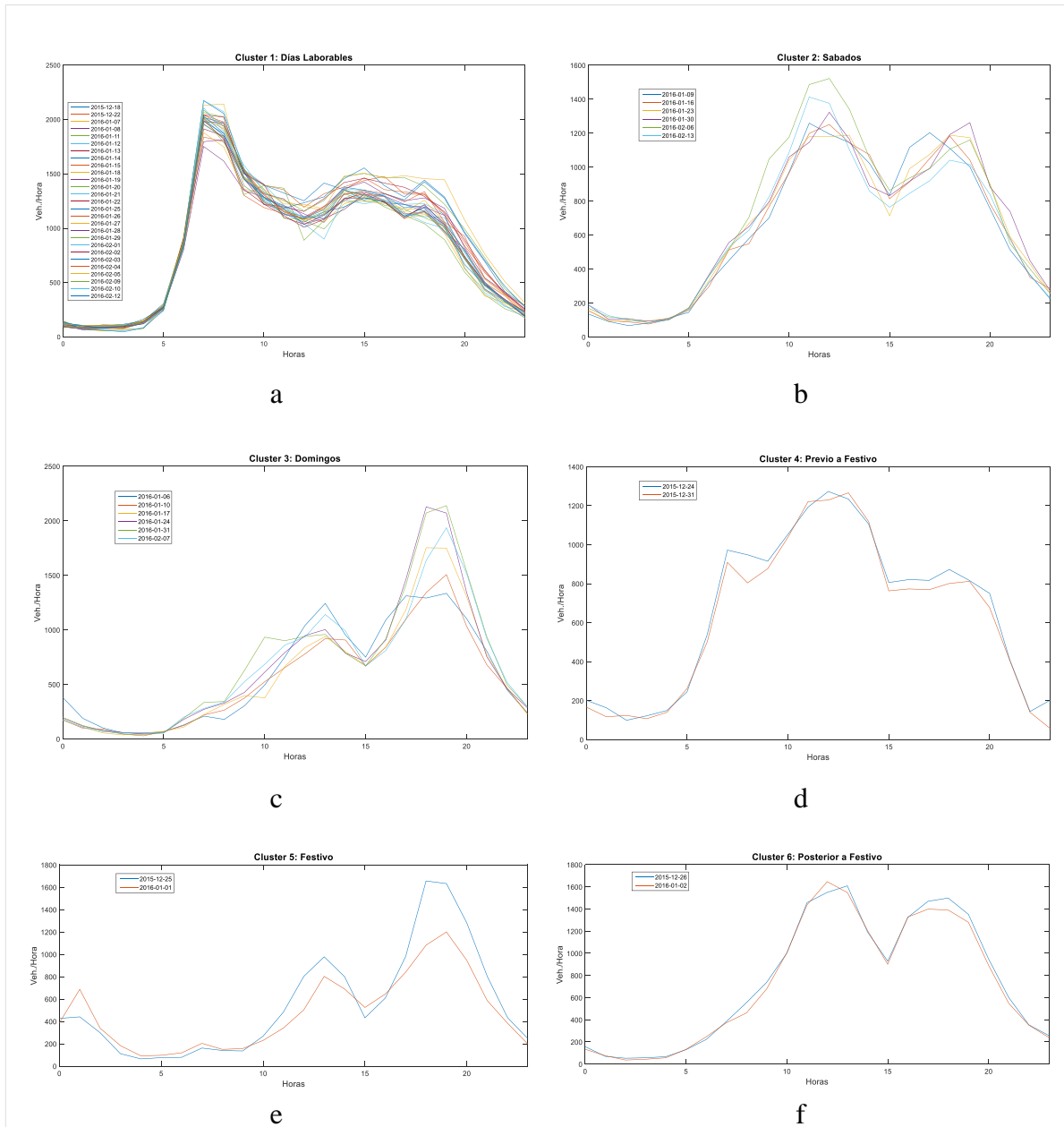


Figura 8: Clusters. a) Laborables, b) Sábados, c) Domingos, d) Previos a Festivo, e) Festivos, f) Posteriores a Festivo

Este mismo proceso se lleva a cabo teniendo en cuenta únicamente los vehículos de mercancías peligrosas. La clusterización asociada a la muestra completa de vehículos (ADR y no-ADR) permite una mayor exactitud en la obtención óptima de la distribución de la tipología de días, debido a que los vehículos de mercancía peligrosa tienen restringida por ley la circulación los domingos y festivos.

La utilización de la clusterización óptima a los registros asociados a los vehículos de mercancía peligrosa y su posterior procesado, permite la obtención de los perfiles de días tipo. La Figura 9 recoge los perfiles para días laborables y para sábados (no existe el perfil para el domingo por la prohibición de circulación ese día de la semana para este tipo de vehículos).

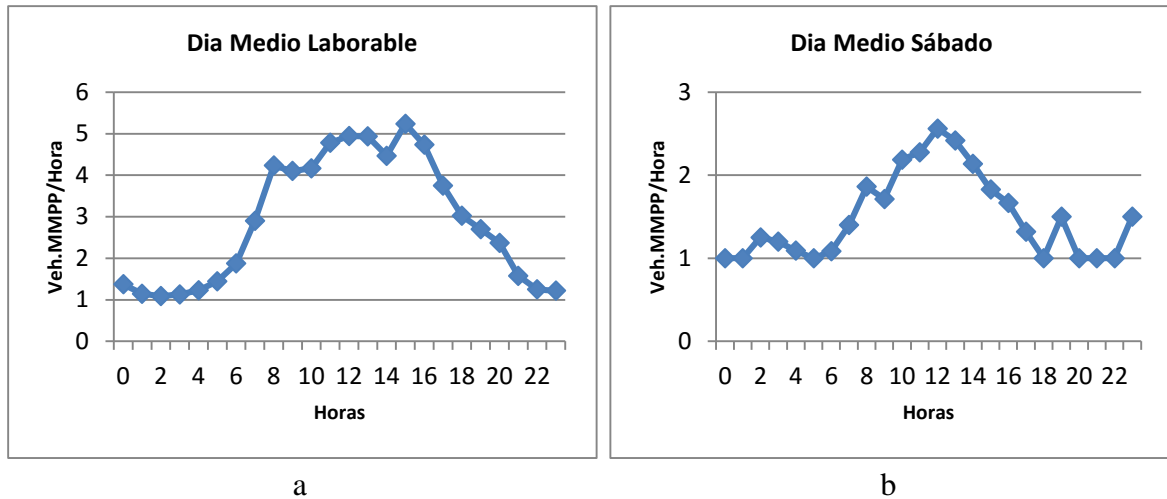


Figura 9: a) Día medio tipo laborable de vehículos de Mercancías Peligrosas, b) Día medio tipo sábado de vehículos de Mercancías Peligrosas

3. ANÁLISIS COMPARATIVO ENTRE DISTINTAS FUENTES DE DATOS

En esta sección, como ejemplo demostrativo, se realiza una comparativa entre los datos publicados por el Mapa de Tráfico y una semana tipo derivada de los datos capturados por la ULP situada en la carretera A-49 punto kilométrico 4,200 y sentido descendente. La estación de aforo asociada es la SE-244-2, situada en el punto kilométrico 5 de esa misma carretera (Tabla 2). Los datos de ambas fuentes deben ser similares ya que entre el kilómetro 5 y el 4,200 no existen cruces ni bifurcaciones.

Ha de tenerse en cuenta que los datos proporcionados por el Mapa de Tráfico incluyen los tres carriles de circulación disponibles en el punto analizado mientras que el sistema de visión artificial solo monitoriza dos de ellos (el derecho y el central). No obstante, un elevado porcentaje de la mercancía peligrosa transportada se realiza mediante vehículos pesados que no hacen uso del tercer carril.

IDAFORO	CTRA	PKP	CALZADA	PROV	MMPP
SE-244-2	A-49	5	Ascendente	SEVILLA	177
SE-244-2	A-49	5	Total	SEVILLA	287
SE-244-2	A-49	5	Descendente	SEVILLA	110

Tabla 2: Datos proporcionados por el Mapa de Tráfico.

IDAFORO	CALZADA	PROV	DIASEM	MMPP
A49, PK 4.200	Descendente	SEVILLA	1	112
A49, PK 4.200	Descendente	SEVILLA	2	140
A49, PK 4.200	Descendente	SEVILLA	3	132
A49, PK 4.200	Descendente	SEVILLA	4	143
A49, PK 4.200	Descendente	SEVILLA	5	133
A49, PK 4.200	Descendente	SEVILLA	6	41

Tabla 3: Datos proporcionados por el sistema ANPR.

En la Tabla 3, la columna DIASEM indica el tipo de día de la semana (1:Lunes, ..., 6: Sábado). El disponer de información de forma ininterrumpida, durante todos los días del año, permite detectar picos y valles en el flujo de mercancías peligrosas que transitan por la sección monitorizada. Puede observarse que, en el punto objeto del caso mostrado, el volumen de mercancía peligrosa transportada se encuentra subestimado; existen días en los que transitan hasta 140 vehículos, mientras que el dato estimado es de 110, alcanzándose errores superiores al 20% (sin tener en cuenta que el sistema únicamente monitoriza dos de los tres carriles). Adicionalmente, el dato proporcionado por el Mapa de Tráfico presenta un único valor frente a la segregación disponible mediante este sistema. Comparando ambas fuentes se evidencia lo siguiente (datos referidos a un miércoles tipo):

- Información proporcionada por el Mapa de Tráfico: 110 vehículos.
- Información proporcionada por el sistema: 132 vehículos.

La distribución de clases de mercancías peligrosas para los 132 vehículos se indica en la Tabla 4; **Error! No se encuentra el origen de la referencia..**

Número de vehículos	Materia transportada
14	223/1972
13	30/1202
11	80/1830
9	33/1203
5	99/3257
4	60/2312
3	90/3082
3	80/3264
3	23/1077
2	40/1350
2	80/1971
2	88/1809
2	22/2187
2	225/1073
57	Otras

Tabla 4. Distribución de vehículos de MP y materia transportada

La disponibilidad de datos, de naturaleza desagregada, de flujos de vehículos de mercancía

peligrosas permite alcanzar una serie de beneficios, entre los que se enfatizan:

- Contraste o calibración de los datos ofrecidos en los Mapas de Tráfico de los Planes de Aforo anuales.
- Identificar zonas especialmente sensibles frente a un riesgo determinado.
- Clasificación de materias por su importancia en cuanto al volumen transportado.
- Potenciar la creación de planes de autoprotección por zonas enfocados a peligros concretos.
- Establecer un determinado control del tráfico de una específica materia peligrosa en función de la locación y hora del día.
- Actualización de matrices de origen-destino de mercancías peligrosas.

4. CONCLUSIONES

En este trabajo se ha presentado una metodología que permite, de forma automática, clusterizar los datos provenientes de un sistema de visión artificial de captura de placas de transporte de mercancías peligrosas. El procedimiento es capaz de detectar todos los tipos de días sin necesidad de definir a priori los tipos a considerar y de proporcionar patrones medios para cada uno de los grupos caracterizados.

En las secciones previas se han descrito el funcionamiento básico del sistema de captura de datos, la comparativa entre la calidad de los datos disponibles (datos agregados capturados manualmente y publicados en el Mapa de Tráfico) y los capturados de forma automática por el sistema de visión artificial, y las conclusiones generales derivadas de este estudio.

AGRADECIMIENTOS

Este trabajo se ha desarrollado gracias a la financiación del proyecto FEDER de la Unión Europea con título “Estimación y actualización del tráfico de vehículos de mercancías peligrosas en la red de carreteras de la Comunidad Autónoma de Andalucía y pronóstico de riesgos medioambientales”, Programa Operativo FEDER de Andalucía 2007-2013, (Agencia de Obra Pública, Junta de Andalucía). Los resultados de este estudio se han podido alcanzar gracias a la colaboración de la Dirección General de Tráfico (DGT) a través del Centro de Gestión del Tráfico del Suroeste y de la Subdirección General de Gestión de la Movilidad.

Por último, uno de los autores (N. Cáceres) también agradece al Ministerio de Economía y Competitividad el soporte proporcionado a través del Programa de Ayudas Torres-Quevedo (PTQ-13-06428).

REFERENCIAS

Chung, E., (2003). Classification of traffic pattern. In: *Proc 10th World Congress on Intelligent Transport Systems (CD)*, Madrid, Spain.

DGT (2016). *RIMP: Red de Itinerarios de Mercancías Peligrosas*. Dirección General de Tráfico, Ministerio del Interior, España.

Febrero-Bande, M., Galeano, P., Gonzalez-Manteiga, W., (2007). A functional analysis of NOx levels: location and scale estimation and outlier detection. *Computational Statistics* 22 (3), pp. 411–427.

Guardiola, I.G., Leon, T., Mallor, F.I.G. (2014). A functional approach to monitor and recognize patterns of daily traffic profiles. *Transportation Research Part B* 65 (2014), pp. 119–136

Liu, Y., Li, C., Pang L. B., (2012). The Application of RFID Technology in the Dangerous Goods Logistics. *Applied Mechanics and Materials*, Vols. 195-196, pp. 904-907.

Milligan, G.W., Cooper, M.C., (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika* 50 (2), pp. 159–179.

Pearson, K. (1901). On Lines and Planes of Closest Fit to Systems of Points in Space. *Philosophical Magazine* 2(11): pp. 559–572.

Romero, L. M., Caceres, N., Reyes, A., Benitez, F.G. (2014). Modelos de demanda de transporte de mercancías peligrosas por carretera en la Comunidad Autónoma de Andalucía. XVIII Congreso Pan-Americano de Ingeniería de Tránsito, Transporte y Logística. XVIII PANAM USB, Paper: 234, ISBN: 978-84-617-0085-1.

Song, Y., Miller, H.J., (2012). Exploring traffic flow databases using space-time plots and data cubes. *Transportation* 39 (2), pp. 215–234.

Yu, M., Deng T., Fu J., (2012). Application of RFID and GPS Technology in Transportation Vehicles Monitoring System for Dangerous Goods. *2nd International Conference on Remote Sensing, Environment and Transportation Engineering (RSETE)*. Nanjing, China.