

# **ESTUDIO DE LA SEVERIDAD DE ACCIDENTES DE TRÁFICO DE AUTOCARES Y AUTOBUSES EN ARGENTINA Y ESPAÑA CON ÁRBOLES DE CLASIFICACIÓN**

**Víctor Miguel Toalombo Vargas**

Profesor Escuela Superior Politécnica de Chimborazo Ecuador

**Blanca del Valle Arenas Ramírez**

Profesor Titular de Universidad. ETSII. Investigador INSIA-UPM, Universidad Politécnica de Madrid, España

**Giovanny Pablo Pillajo Quijia**

Doctorando ETSIDI. INSIA-UPM. Universidad Politécnica de Madrid, España

## **RESUMEN**

El transporte colectivo de personas en autobús y autocar es muy importante en la mayoría de los países de Iberoamérica, y los accidentes de tráfico con implicación de estos vehículos involucran pérdidas elevadas si se consideran las pérdidas humanas, materiales y de productividad de las empresas del transporte. En cualquier país del mundo en el que éstos eventos ocurren, tienen un alto impacto mediático y avivan la preocupación por la mejora de la seguridad y la protección de los ocupantes. En la región de América del Sur, la mejora de la seguridad vial necesita alcanzar un mayor nivel de desarrollo de los sistemas de información que permitan llevar a cabo una investigación detallada y en profundidad.

Para suplir la escasez de información detallada de los accidentes con implicación de autobuses y autocares del transporte colectivo de personas (AT-TCP), cinco equipos de investigadores de las universidades: UPM de España, UPS de Ecuador y UTN, UNSa y UNTDF de Argentina, participaron en un proyecto cooperativo de investigación de accidentes en transporte colectivo de pasajeros (COOPERAXVII-12) ocurridos en los tres países y desarrollaron una metodología de análisis de datos tomados de publicaciones en medios de comunicación y datos oficiales disponibles, enfocado en el desarrollo de estudios de patrones y aplicación de modelos estadísticos.

En este trabajo se evaluó la severidad de una sub-muestra de 94 accidentes AT-TCP ocurridos en Argentina y España entre los años 2017 y 2018. Los datos fueron analizados con modelos de minería de datos: árboles de clasificación y bosques aleatorios. Los modelos permiten identificar los factores de influencia en la severidad del accidente y algunos patrones según el contexto y concurrencia de diferentes factores. De los resultados, se destaca que los accidentes de mayor severidad han ocurrido en vías interurbanas y son de tipo colisión entre vehículos, vuelco sobre la calzada y salidas de la vía con desbarranco.

## 1. INTRODUCCIÓN

El transporte colectivo de personas por vía terrestre, tiene una función vital en todos los países y en especial en aquellos en los que los costos del transporte por medios más veloces son elevados para grandes sectores de la población. En los países en los que existe un sistema de transporte de personas de mayor desarrollo, los medios de alta velocidad ferroviario, ofrecen a los ciudadanos una oferta mayor en un sistema competitivo de precios del transporte. Pero en países sin un sistema de transporte con alternativas, el realizado con autobuses y autocares (AyA) cumple un papel vertebrador del territorio y de la economía. Los impactos negativos del transporte como los accidentes de tráfico en todo el mundo superan el número de 1.200.000 víctimas mortales, con las consecuencias graves desde el punto de vista humano y económico dentro de la colectividad. Distintos organismos (World Health Organization (WHO) – UN) coinciden en señalar a los accidentes de tráfico como una de las principales causas en la pérdida de vidas humanas y especialmente en los colectivos de edades más tempranas. Villar Fernández, (2012) señala que los accidentes de tráfico son la segunda causa de muerte, y la primera dentro de la población juvenil comprendida entre 15 a 29 años, siendo uno de los principales problemas de Salud Pública en España.

En referencia a pérdidas humanas en España, durante el año 2017, los diferentes miembros policiales notificaron 102.233 accidentes con víctimas. Estos accidentes ocasionaron 1.830 fallecidos (en el momento del accidente o hasta 30 días después del mismo); además, 9.546 personas fueron ingresadas en un centro hospitalario y 129.616 resultaron heridas no hospitalizadas. Del total de víctimas en accidentes de tráfico con implicación de autobuses, se produjeron 3 muertos, 47 personas fueron ingresadas en un centro hospitalario y 1.171 resultaron heridas no hospitalizadas. La mayoría de ellos (87%) se produjeron en vías urbanas, los restantes (el 13%) en interurbanas, siendo la lesividad de los accidentes, en general, mayor en vías interurbanas. (DGT, 2017).

La Asociación Civil Luchemos por la vida de Argentina en el año 2017, reporta que 5.420 personas fallecieron (en el lapso de los 30 días posteriores al accidente), 6.626 resultaron heridos graves hospitalizados y 51.945 con heridas leves en los accidentes de tráfico ocurridos.

Otra fuente de Argentina (Anuario estadístico de siniestralidad vial) reporta para el mismo año: 5.472 víctimas fatales; 113.805 heridos, de los cuales 8.174 son graves que requerían hospitalización, en un total de 81.592 siniestros ocurridos. Las principales causas de muerte por accidentes de tránsito fueron: la velocidad y el consumo de bebidas alcohólicas. Para los peatones y los conductores, el problema principal es la distracción. De los números expuestos para Argentina, se deduce que existe dispersión entre ellos según las fuentes y no se disponen de datos desagregados para el estudio científico de los accidentes con AyA.

Contrariamente, en España, existe un sistema de información accesible a los investigadores y que posibilita la realización de estudios de investigación de colectivos de interés (vulnerables, transporte de viajeros, trabajadores in itinere, etc.).

El cuestionario aplicado en el proyecto de investigación COOPERA TCP se ha creado teniendo como base el existente en España, que la DGT aplica para la recogida de datos in situ. El cuestionario COOPERA TCP incluye además campos para las variables de infraestructura obtenidas con Google Maps, con las coordenadas del lugar del accidente en el marco del proyecto de investigación de accidentes con implicación de autobuses y autocares, en el que participaron varias universidades de 3 países participantes (Argentina, Ecuador y España).

Este trabajo analiza una muestra de accidentes de tráfico en Argentina y España, donde se han visto implicados autobuses y autocares, con el objetivo principal de identificar las variables que influyen en la severidad de los accidentes.

Para la identificación de los factores de influencia se utilizan modelos estadísticos cuyo análisis que permitirá realizar un diagnóstico más riguroso y completo dentro de los límites que proporcionan los datos disponibles. Con este estudio se pretende profundizar en el conocimiento de la situación accidentalológica de AyA en los países participantes, con limitaciones en el acceso a los datos necesarios para tal fin.

El desarrollo de estudios como éste permitirá analizar los fenómenos causales y extraer los patrones y combinaciones de variables de influencia, y así adoptar las mejores decisiones para la mejora de la seguridad del transporte colectivo de personas.

## **2. MATERIALES Y MÉTODOS**

### **2.1 Metodología de recolección de la información**

El estudio toma como base el proyecto COOPERA-TCP que estandarizó una herramienta de recolección de datos de accidentes de tráfico de autocares y autobuses (Arenas et al., 2018), y de otros vehículos TCP como microbús, combi, autocar de un piso, microbús y autocar de dos pisos en los países participantes. La recolección de datos se realizó entre 2017 – 2018, para la cual se utilizó la metodología de búsqueda, localización y análisis de medios de comunicación masivos. Esta información se contrastó con las bases de datos disponibles en el país y la información se recogió en el instrumento destinado para la recopilación de datos en formato digital.

El formulario utilizado para la ejecución del proyecto reúne los datos de accidentes, de la vía en donde ha ocurrido el accidente y del estado en el que han quedado los vehículos de transporte implicados en el accidente de tráfico, por lo cual contiene campos con las variables de interés para abordar estudios científicos en países de Latinoamérica.

En la figura 1 se presentan los factores del formulario del proyecto COOPERA-TCP, basados en la estructura del modelo MIICA.

Número de factor	Factores de influencia y otros	Tipo de factor de acuerdo a MICAA
1	Datos generales	Datos del momento y espacio. Geo-referenciación.
2	Tipo de vía	Infraestructura
3	Tipo de servicio del autobús o autocar / TCP	Vehículos
4	Lugar	Dato de tiempo y espacio
5	Tipo de accidente	
6	Condiciones en el momento del accidente	Condiciones ambientales y del tráfico
7	Vehículos	Vehículos
8	Datos de los ocupantes (conductor, pasajeros, otros) y víctimas	Actores del tráfico
9	Datos del conductor del autobús o autocar / TCP	Actores del tráfico
10	Evacuación de las personas lesionadas	Otros datos - aportes de datos para la investigación del siniestro
11	Presuntas infracciones del conductor	Actores del tráfico
12	Posible responsabilidad del accidente	Actores del tráfico
13	Factores que pudieron afectar la atención y presuntos errores del conductor	Actores del tráfico
14	Descripción del accidente	Otros datos - aportes de datos para la investigación del siniestro
15	Daño estructural en el autobús o autocar / TCP	Otros datos - aportes de datos para la investigación del siniestro
16	Esquema de daños en los vehículos	Otros datos - aportes de datos para la investigación del siniestro

**Figura 1: Factores y datos recogidos en el formulario de accidentes del Transporte colectivo de personas – Proyecto COOPERAXVII-12 (TCP). (Arenas et al., 2018).**

La base disponible para este trabajo está conformada por una muestra de 94 accidentes ocurridos en España y Argentina y contiene más de 80 variables. El análisis realizado permite resaltar la potencialidad de la metodología de COOPERA-TCP y de las herramientas propuestas para los objetivos planteados.

Se ha realizado una selección de variables con varios criterios (como evitar repetición de la información, valores nulos o vacíos, etc.) y con modelos Random Forest (RF) considerando su error de clasificación. Se seleccionan las 13 variables superiores en la clasificación con las que se elabora un modelo final. Las 13 variables explicativas se han agrupado en 4 grupos: factor infraestructura, vehículo, entorno y factor humano. Con las variables seleccionadas, se desarrollará un modelo de análisis de la severidad del accidente.

## **2.2 Metodología estadística para el análisis**

### **2.2.1 Modelos de bosques aleatorios (Random Forest (RF))**

Es una versión sofisticada del procedimiento bagging, señalada por (Breiman, 2001), que replica subconjuntos de registros y elige al azar un subconjunto de las variables de entrada (Hastie, Tibshirani, & Friedman, 2008). La herramienta indicada se enfoca en la técnica de aprendizaje automático estándar denominada “árbol de decisiones. Esta metodología también se usa para análisis de sensibilidad (Grömping, 2009).

La técnica de conjuntos de árboles tienen enfoque clásico o frecuentista, y el algoritmo utilizado es el RF (Random Forest de R Core, 2016). Este algoritmo optimiza la precisión en la clasificación mediante la incorporación de aleatoriedad en la construcción de cada clasificador individual. Para la selección y optimización se utilizan dos índices: criterio de Gini y de clasificación.

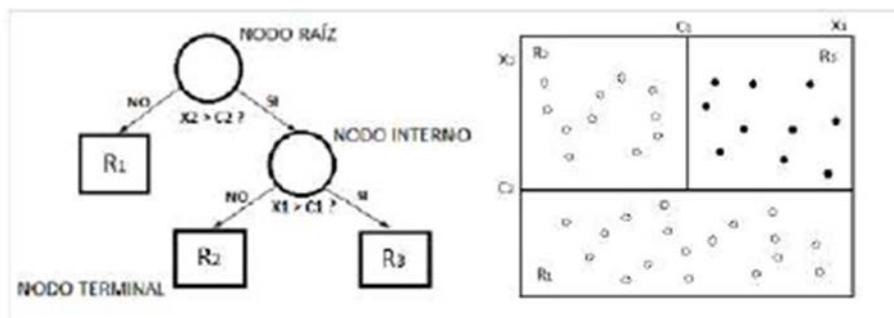
El mecanismo de construcción de RF establece un baremo que prioriza cada variable en la predicción final. Para ello se calcula el error de la muestra. Para cada variable de la muestra, se intercambia un par de elementos y se calcula el error de la muestra permutada. El resultado debería ser peor que para la muestra original. Este procedimiento se realiza para todos los valores de cada variable y se calcula el promedio. El proceso es realizado para todas las variables. Las variables de menor importancia deberían alterar menos la diferencia entre el error y el error permutado que las variables importantes.

### **2.2.2 Modelos de árboles de clasificación (CART y Rpart)**

Los modelos de Árboles de clasificación son una técnica de aprendizaje controlado para modelar relaciones entrada – salida, no paramétrica, de segmentación binaria, esto es, las particiones se hacen de forma permanente hasta alcanzar un criterio de parada, es decir que el árbol se construye dividiendo los datos repetidamente, en cada nueva división los datos son divididos otra vez en dos grupos mutuamente excluyentes.

El enfoque algorítmico más habitual para el CART creado por (Breiman, Friedman, Stone, & Olshen, 1984) produce de manera inicial un árbol muy grande, para consecutivamente podarlo; es decir, recortar ramas que no incrementen su capacidad predictiva. El grado de poda se define en función de un compromiso entre bondad de ajuste y confusión del árbol, o incluyendo también criterios de precisión de una validación cruzada.

La figura 2 muestra un nodo raíz o madre, que se divide después en nodos hijos, y finalmente el procedimiento de partición se aplica a cada nodo hijo por separado. Estas divisiones se seleccionan, de modo que la impureza de los grupos hijos sea menor que la del grupo madre. Este procedimiento tiene por objetivo, discriminar la respuesta en grupos homogéneos, de tal manera que se pueda mantener el árbol relativamente pequeño. (Pillajo et al., 2018).



**Figura 2: Descripción de objetos de un árbol de clasificación (Zhang / Singer, 2010).**

Mediante la función Rpart del programa R, se analiza el modelo de clasificación con las 13 variables seleccionadas previamente con RF y se ejecuta la poda del árbol en función del % de error relativo y del índice de complejidad de división CP óptimo para generar el árbol de clasificación final.

### 3. RESULTADOS Y DISCUSIÓN

A continuación, se muestra el análisis descriptivo de los datos concernientes a la muestra de 94 accidentes de tráfico de autocares y autobuses ocurridos en las carreteras de Argentina y España implicados entre los años 2017– 2018.

En la Tabla 1 se indica los accidentes según tipo de vehículo de la muestra.

Argentina				Total	España				Total
Autocar/Micro		Colectivo/Autobús			Autocar/Micro		Colectivo/Autobús		
N.º	%	N.º	%		N.º	%	N.º	%	
30(0.6)	31.91	20(0.4)	21.28	50(1)	37(0.84)	39.36	7(0.16)	7.45	44(1)

**Tabla 1: Descripción accidentes de TCP en Argentina y España.**

El número de accidentes de tráfico con implicación de Autocar/Micro es mayor en España, con un total de 39.36% en comparación con el 31.91% de Argentina. Por otra parte, en lo que concierne a accidentes de tráfico con Colectivo/Autobús, se observa una mayor incidencia en Argentina con el 21.28% en relación a la de España (el 7.45%). Los porcentajes se han tomado en base a los 94 accidentes en los dos países.

Con respecto a Argentina se produjeron más accidentes de tráfico en los que están involucrados Autocar/Micro, el 60% en comparación con la implicación de Colectivo/Autobús que alcanza el 40%. Lo mismo sucede en España con una mayor implicación de Autocar/Micro en accidentes (el 84%).

Se puede identificar que en cualquier relación ya sea en función de los 94 accidentes o por país, en España existe un mayor número de accidentes con tipo de vehículos TCP Autocar/Micro.

Se han generado dos graficas de árbol de clasificación (Fig. 3 y 4) para Argentina y España respectivamente por medio de modelos CART de severidad de los accidentes de tráfico de TCP para las que se ha utilizado el Software R Studio. Dentro de la misma la librería llamada “rpart” que se encarga de realizar los cortes para las ramificaciones, que se han utilizado para el cálculo de Cp (Complexity parameter) y el Error Relativo. La elección de las variables más importante se realizó con RandomForest (RF) de R Studio para la variable respuesta “severidad”. Con respecto a Argentina indica que la variable más importante es el Tipo de arcén, en segundo lugar el tipo de accidente, tercer lugar el tipo de vía, mientras que para España la variable más importante es el Tipo de vía, seguido de Tipo de accidente y en tercer lugar Responsable; con lo cual se aprecia que dentro de los tres primeros lugares están presentes en los dos países las variables Tipo de accidente y Tipo de vía, lo cual indica que estos dos son la factores que más inciden en la severidad de accidentes de tráfico con autocares y autobuses.



**Figura 3: Árbol de severidad de Argentina.**

En síntesis, en Argentina cuando el tipo de arcén es natural, o cuando está en mal estado, pavimentado u otros, el tipo de accidente es un alcance, y el accidente se ha producido en una intersección resultaría una alta lesividad.



**Figura 4: Árbol de severidad de España**

En España, cuando el tipo de vía es autopista o carretera de doble sentido y el tipo de accidente es por alcance al autobús, salida de la vía con desbarranco, salida de la vía u vuelco la severidad es alta. De la misma manera, si el tipo de accidentes es por alcance del autobús, choque o vuelco sobre la calzada, y el estado de la vía es mojado y ambos vehículos implicados son responsables, el accidente tiene una alta severidad.

#### 4. CONCLUSIONES

Al existir un incremento de la velocidad en carretera para el tipo de vehículo 'autocar', la colisión tiende a ser más grave, por lo que se manifiesta una mortalidad mayor, y no hay diferencias significativas en Argentina con respecto a España.

El factor infraestructura (referente a tipo de vía) es el que resulta con mayor connotación; así como también el factor humano (responsabilidad) en algunos tipos de accidentes.

Los vuelcos de autocares son los accidentes más severos tanto en Argentina como en España, en comparación con autobuses (que se mueven más en ámbitos urbanos o suburbanos).

El método descriptivo estadístico utilizado con BBDD-AyA, es adecuado para identificar y comprender los factores y características que influyen en los accidentes de tráfico con implicación de autocares y autobuses tanto en España como en Argentina. La metodología aplicada en este estudio ha permitido realizar e interpretar clasificaciones claras de variables y combinaciones que determinan la severidad de los accidentes con implicación de AyA.

#### AGRADECIMIENTOS

Agradezco a la Universidad Politécnica de Madrid por haberme permitido formar parte de este proceso académico, asimismo a los Docentes que han contribuido con sus conocimientos para culminar esta etapa. Los autores agradecen además, a la Comunidad de Madrid por su financiación parcial que, a través del programa SEGVAUTO 4.0-CM (P2018 / EMT-4362) ha contribuido a su desarrollo y difusión.

## REFERENCIAS

ARENAS, A., ARENAS, B., ESPINOZA, F., TARCAYA, H., APARICIO, F. (2018). Investigación de accidentes en unidades del transporte colectivo de personas COOPERAXVII-12(TCP).

BREIMAN, L.; FRIEDMAN, J.; STONE, CJ; OLSHEN, árboles de clasificación y regresión de RA; CRC Press: Boca Raton, FL, EE. UU., 1984.

DGT. (2016). Las principales cifras de la siniestralidad vial en España 2016. Dirección general de tráfico.

DGT. (2016). Seguridad vial, tablas estadísticas.

DIRECCIÓN GENERAL DE TRÁFICO (DGT) (2017). Las Principales Cifras de la Siniestralidad Vial. España. 2017. Disponible online: <http://www.dgt.es/es/seguridad-vial/estadisticas-e-indicadores/publicaciones/principales-cifras-siniestralidad/> (consultado el 18 de noviembre de 2019).

DIRECCIÓN GENERAL DE TRÁFICO (DGT) (2018). Las Principales Cifras de la Siniestralidad Vial. España. 2018. Disponible online: <http://www.dgt.es/es/seguridad-vial/estadisticas-e-indicadores/publicaciones/principales-cifras-siniestralidad/> (consultado el 18 de noviembre de 2019).

DIRECCIÓN NACIONAL DE OBSERVATORIO VIAL. Anuario Estadístico de siniestralidad vial Año 2018 Dirección Nacional de Observatorio Vial (2018). Disponible en: [https://www.argentina.gob.ar/sites/default/files/ansv\\_denov\\_anuario\\_2018.pdf](https://www.argentina.gob.ar/sites/default/files/ansv_denov_anuario_2018.pdf).

DONNET, et al., (2018). Investigación de accidentes en el ámbito del transporte público de personas (TPP). XIII Congreso de Ingeniería del Transporte. España.

ESPINOZA, F., NARVÁEZ, P., ARENAS, B., APARICIO, F., CARAGUAY, J., YANZA, B., (2018). Análisis de accidentes de tránsito de autobuses en Ecuador. XIII Congreso de Ingeniería del Transporte. Guijón, España; del 06 al 08 de Junio; pp 1964 – 1974.

FERNÁNDEZ-DELGADO, M.; CERNADAS, E .; BARRO, S .; AMORIM, D. ¿Necesitamos cientos de clasificadores para resolver problemas de clasificación del mundo real? J. Mach. Aprender. Res. 2014 , 15 , 3133–3181.

GUIDICI P. (2003). Applied Data Mining. Statistical Methods for Business and Industry. John Willey& Sons Ltd, The Atrium, Southern Gate. Chichester. West Sussex PO19 8SQ. England.

MAYORGA S., VALDEZ O., ARENAS A., ROBIN J., TUERO I., LUGONES L. (2014). Evaluación de la Situación Accidentalógica en la ciudad de Salta. XI Congreso de Ingeniería del Transporte CIT. ISBN 978-84-697-0359-5. Santander, España. Ed.

EUG. 9 al 11 de junio de 2014. 13 pág.

OMS (2010). Sistema de datos. Manual de seguridad vial para decisores y profesionales.

OMS (2018). Lesiones por accidentes de tránsito. Disponible en línea: <http://www.who.int/mediacentre/factsheets/fs358/en/> (consultado el 17 de noviembre de 2019).

OMS (2015). Informe sobre la situación mundial de la seguridad vial 2015. Organización Mundial de la Salud . Disponible en línea: [http://www.who.int/violence\\_injury\\_prevention/road\\_safety\\_status/2015/en/](http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/) (consultado el 18 de noviembre de 2019).

PALACIOS, B. D. (2016). Análisis mediante el MEF de la protección de los conductores en autobuses.

PILLAJO, G., ARENAS, B., APARICIO, F. (2018). Análisis de severidad de accidentes con implicación de una furgoneta con modelos de clasificación y bosques aleatorios.

PILLAJO, G., ARENAS, B., GONZÁLEZ, C., & APARICIO, F. (2020). Influential Factors on Injury Severity for Drivers of Light Trucks and Vans with Machine Learning Methods. *Sustainability*, 12(4), 1324.

RAMÍREZ, B. A., IZQUIERDO, F. A., & FERNÁNDEZ, C. G. (2007) Factores de influencia sobre la frecuencia de accidentes en la red de carreteras del estado, mediante árboles de regresión y clasificación.

REGISTRO NACIONAL DE ACCIDENTES DE TRÁNSITO DE ARGENTINA. (2017). Mapa interactivo de la seguridad vial, víctimas fatales. Disponible en: <https://www.argentina.gob.ar/seguridadvial/mapa-interactivo-seguridad-vial#>

RINCÓN, D., ALCALÁ, E., BLÁZQUEZ, V., & MARTÍN, Á. (2018). Estudio experimental de la reparación de estructuras de autobuses tras accidente. *Revista Iberoamericana de Ingeniería Mecánica*, 22(2).

VILLAR, L. (2012). Actividad asistencial de las unidades de soporte vital avanzado a accidentes de tráfico en Asturias en el año 2010.

ZHANG, H., & SINGER, B. (2010). *Recursive Partitioning and Applications*. New York: Springer.