

# **MODELOS ESTADÍSTICOS SOBRE LA INTERACCIÓN ENTRE VEHÍCULOS DURANTE MANIOBRAS DE ADELANTAMIENTO**

Joaquín Espinoza Schmidt, Universidad de los Andes, Chile — [jfespinoza@miuandes.cl](mailto:jfespinoza@miuandes.cl)  
Rafael Delpiano Costabal, Universidad de los Andes, Chile — [rdelpiano@uandes.cl](mailto:rdelpiano@uandes.cl)

## **RESUMEN**

Un desafío práctico para la adopción de vehículos autónomos es la influencia de la dimensión lateral en el tráfico. El tema ha sido escasamente estudiado, y faltan descripciones cuantitativas de las interacciones entre vehículos para validar modelos.

En este trabajo, se analiza las interacciones entre variables cinemáticas en ambos ejes para pares de vehículos durante maniobras de adelantamiento, tanto con modelos lineales como no-lineales y no-paramétricos, sobre la base de datos highD.

Se encontró i) una influencia mutua entre vehículos, especialmente para distancias pequeñas, ii) que a mayor velocidad longitudinal, mayor tendencia al distanciamiento lateral, posiblemente por prevenir colisiones, y iii) aceleraciones laterales tendientes a acercarse están asociadas a aceleraciones longitudinales para alejarse. Los resultados son consistentes entre los distintos modelos y con resultados anteriores.

*Palabras clave:* modelos estadísticos, tráfico en dos dimensiones, Generalized Additive Models

## 1. INTRODUCCIÓN

La llegada de los vehículos autónomos (VV.AA.) se ha topado hasta ahora con más de algún obstáculo práctico, al punto de que el CEO de una de las principales empresas impulsoras de esta tecnología pusiera en duda la factibilidad futura de vehículos totalmente autónomos en condiciones de clima adverso. Entre dichos problemas prácticos, uno importante que sí puede ser parcialmente abordado es la escasa comprensión que se tiene de la importancia e influencia de la dimensión lateral en el tráfico, que ha contribuido a que las interacciones laterales sean una causa relevante de accidentes de VV.AA.

Trabajos previos sobre la dimensión lateral del tráfico pueden clasificarse en: i) identificación cualitativa de fenómenos, ii) identificación cuantitativa y caracterización del tráfico en dimensión lateral y iii) modelación. El primer tema es de larga data y no nos detendremos en él. Baste decir que se remonta a la descripción de fenómenos como la fricción lateral [cita]. Respecto a los otros dos, su desarrollo es todavía incipiente. En mediciones es poco lo que existe, pero destacan los trabajos de (Budhkar & Maurya, 2017) que establece una relación entre velocidad y distancia lateral entre vehículos y (Jin, 2010) que estudia el proceso de cambio de pistas. Por su parte, los autores han contribuido con un trabajo donde se individualiza un comportamiento específico de vehículos ubicados lado a lado (Delpiano, 2015), y uno donde se caracteriza la disciplina de pistas (Delpiano, 2021). Respecto a la modelación, los primeros intentos de modelar tráfico en 2D y reproducir la fricción lateral datan de (Gunay, 2003) y se han desarrollado por diversos caminos. Sólo algunos de los modelos resultantes permiten una modelación de escenarios de flujo vehicular suficientemente completos.

Adicionalmente, existe aún un gran vacío en términos de una descripción cuantitativa de las interacciones entre vehículos fuera del seguimiento vehicular directo (i.e., en direcciones distintas a la longitudinal), que permite evaluar los modelos y su capacidad de reproducir con precisión las mismas.

En este trabajo, se analiza con detención las interacciones entre variables cinemáticas (posición, velocidad, aceleración) en ambos ejes (longitudinal y lateral) para pares de vehículos durante maniobras de adelantamiento por medio tanto de regresiones lineales como de modelos no-lineales y no-paramétricos. Dentro de los modelos no-lineales se consideraron los modelos GAM (*Generalized Additive Models*) que incluyen la posibilidad de encontrar relaciones no lineales para una variable dependiente, y los modelos GAMLSS (*Generalized Additive Models for Location, Scale and Shape*), los cuales incluyen además la posibilidad de modelar todos los parámetros de la distribución de probabilidad de la variable dependiente. Los análisis se efectúan sobre la base de datos highD (Krajewski, et al., 2018), consistente en trayectorias extraídas de video aéreo grabado en la Autobahn alemana.

Se encontró y cuantificó una esperable influencia mutua entre ambos vehículos. Esta influencia no es lineal con respecto a la distancia, y es especialmente notoria para distancias pequeñas. Por otro lado, a mayor velocidad longitudinal del vehículo adelantado, más rápidamente se aleja el adelantador lateralmente, posiblemente por minimizar las probabilidades de choque, y las aceleraciones laterales tendientes a acercar a ambos vehículos están asociadas a aceleraciones longitudinales que los alejen. Los resultados son consistentes entre los distintos modelos y con resultados anteriores.

## 2. PROCESAMIENTO DATOS Y VARIABLES CONSIDERADAS

Para realizar el análisis y creación de modelos estadísticos se utilizó la base de datos *highD* (Krajewski, et al., 2018), que entrega información de trayectorias vehiculares en autopistas alemanas, cercanas a la ciudad de Köln. Entrega, entre otros, la posición, velocidad y aceleración en ambos ejes de todos los vehículos, con una frecuencia de 25 muestreos por segundo. Esta consta de 60 grabaciones de los años 2017 y 2018 en 6 diferentes segmentos de autopista de aproximadamente 400 m de longitud, identificados por números enteros del 1 al 6. Durante esta investigación se utilizó específicamente la ubicación número 1, debido a que es la que presenta la mayor cantidad de datos y para evitar correlaciones espurias debidas a distintos sitios. A continuación se procedió a encontrar todos los adelantamientos observados (i.e., cuando un auto pasa por al lado de otro en una pista adyacente, con mayor velocidad).

Para identificar los adelantamientos, se utilizaron las siguientes variables de la base de datos:

$t$ : Tiempo o cuadro (*frame*) de la grabación

$x_{it}$ : Posición lateral del vehículo  $i$  en el tiempo  $t$

$y_{it}$ : Posición longitudinal del vehículo  $i$  en el tiempo  $t$

De esta forma, se establece que el automóvil  $i = 1$  es el adelantador, e  $i = 2$  el adelantado, y se procedió a encontrar todos los instantes de cada par de autos en los que se cumplan las siguientes restricciones:

$$|x_{1t} - x_{2t}| < 5 \quad (1)$$

$$-4,5 < y_{1,x} - y_{2,x} < 1,5 \quad (2)$$

La primera condición establece que las distancias brutas (incluyendo el ancho de un auto) entre los autos sean de menos de 5 m, de manera que la distancia neta (sin contar el ancho) no supere un ancho de pista.

La segunda restricción establece que la maniobra se estudiará como un proceso prolongado en el espacio y en el tiempo, dando más importancia a la aproximación previa.

Al aplicar las restricciones, para cada instante de cada maniobra se obtienen y calculan las siguientes variables para el vehículo adelantador ( $i=1$ ) y adelantado ( $i=2$ ):

- $x_i$  : Posición lateral
- $y_i$  : Posiciones longitudinal
- $w_i$  : Ancho
- $v_{xi}$  : Velocidad lateral
- $v_{yi}$  : Velocidad longitudinal
- $a_{xi}$  : Aceleración lateral
- $a_{yi}$  : Aceleración longitudinal
- $(v_i^p)$  : Velocidades de los predecesores
- $GAP_i$  : Distancias de los predecesores

- $c_i$  : Tipos de vehículo (1 → camión, 0 → automóvil)  
 $q$  : Tipo de adelantamiento (1 → por la izquierda, 0 → por la derecha)  
 $d$  : Distancia lateral neta entre los autos,  $d = |x_1 - x_2| - \frac{w_1 + w_2}{2}$

Para las variables en el eje lateral se consideraron los ejes mostrados en la *Figura 1*, de manera que estas serán negativas si su vector apunta en el sentido donde se encuentra el otro auto, y positiva en caso contrario. Los ejes anteriores contribuyen consistentemente en el cálculo de posiciones relativas, y brindan mucha interpretabilidad, debido a que las variables serán positivas cuando hay repulsión, y negativas cuando hay atracción.

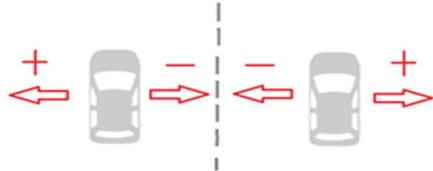


Figura 1: signos de la velocidad y aceleración lateral

### 3. MODELOS DE REGRESIÓN LINEAL PARA DISTANCIAS CORTAS

Encontradas las variables de cada adelantamiento, se procedió a establecer modelos de regresiones lineales. Al tener para cada adelantamiento una serie de instantes, para caracterizar cada uno se calculó el promedio en el tiempo de cada una de las variables. Cabe destacar que todos los modelos realizados durante esta investigación tienen como variable dependiente a velocidades y aceleraciones laterales, dando un énfasis en aquellas del adelantador. Lo anterior se debe a que cuando se realizaron modelos para el adelantado, se encontraron  $R^2$  y relaciones bajas o inexistentes entre las variables independientes y dependientes cuando se compararon con los modelos del adelantador. En consecuencia se tomó el supuesto de que es el adelantador quien reacciona de manera más directa frente a las diferentes variables cinemáticas tanto de él mismo como del adelantado. Así, el modelo base que se consideró está dado por la siguiente fórmula, en el cual la variable dependiente es la aceleración lateral del adelantador.

$$a_{x_1} = \beta_0 + \beta_1 Dist + \beta_2 a_{x_2} + \beta_3 v_{y_1} + \beta_4 v_{y_2} + \beta_5 v_{x_1} + \beta_6 v_{x_2} + \beta_7 a_{y_1} + \beta_8 a_{y_2} + \beta_9 v_1^p + \beta_{10} v_2^p + \beta_{11} GAP_1 + \beta_{12} GAP_2 + \beta_{13} c_1 + \beta_{14} c_2 \quad (3)$$

Al igual que en el trabajo de Budhkar y Maurya (2017), se utilizaron filtros de distancias laterales, es decir que se consideraron adelantamientos en los que la distancia lateral neta entre los vehículos sea menor o igual a un umbral estipulado. En adelantamientos en que la distancia lateral es muy grande, el poder explicativo de los modelos obtenidos se hace muy bajo (Espinoza, 2020). En consecuencia, se escogió un umbral de 1,2 m de distancia lateral, buscando un equilibrio entre el poder explicativo del modelo y su tamaño muestral mayor a 1000 datos de adelantamientos, obteniendo 1.091 grados de libertad (Espinoza, 2020).

### 3.1. Resultados

La Tabla 1 detalla el modelo encontrado. Se observa que son significativas al 90% de confianza las variables  $v_{x_1}$ ,  $v_{x_2}$ ,  $a_{x_2}$ ,  $v_{y_1}$ ,  $v_{y_2}$ ,  $a_{y_1}$ ,  $a_{y_2}$ ,  $v_2^p$ ,  $c_2$  y  $q$ .

Variable de Repuesta: ax1			
	Coeficiente	Error Estándar	p-value
(Intercept)	-0,04345	0,02666	0,10345
distP	0,00561	0,01911	0,76908
ax2P	-0,11105	0,04677	0,01774
vy1P	-0,00328	0,00096	0,00063
vy2P	0,004	0,00106	0,00016
vx1P	-0,28863	0,01497	< 2E-10
vx2P	-0,12638	0,02988	0,00003
ay1P	-0,03362	0,00691	1,32E-06
ay2P	0,01608	0,00899	0,07413
pvelP1	-0,000032	0,00027	0,22974
pvelP2	0,000067	0,0003	0,02621
gapM1	-0,000004	0,00007	0,52797
gapM2	-0,000008	0,00006	0,19923
class1P	-0,00596	0,00779	0,44398
class2P	0,02306	0,00533	0,00002
dummyP	0,02296	0,01169	0,04979
Res. S.E.	0,08		
R2 Aj.	0,33427		
DF	1091		

Tabla 1: Modelo de regresión lineal

Primero, se estudió la relación entre la aceleración y velocidad lateral del adelantador. La velocidad  $v_{x_1}$  establece un signo negativo en el coeficiente, indicando que cuando un adelantador tenga magnitudes altas de la velocidad lateral, este querrá centrarse en su pista disminuyendo la velocidad en base a la aceleración.

Luego, es necesario estudiar como las velocidades y aceleraciones laterales del adelantado influyen en la variable objetivo. La velocidad  $v_{x_2}$  tiene un coeficiente negativo, y la mejor forma de interpretar dicho coeficiente es que cuando el adelantado se acerca al adelantador, este último percibe una repulsión para prevenir un choque y tiende a alejarse del primero para que haya una mayor distancia entre ambos. Cuando  $v_{x_2}$  es positiva, indica que el adelantado se estaría alejando, de manera que aumentaría la distancia entre ellos, por lo que frente a este aumento el adelantador se acercará al adelantado para mantener la distancia lateral.

Resultados e interpretaciones muy similares se encuentran cuando se analiza la aceleración lateral del adelantado. El coeficiente de la variable  $a_{x_2}$  es negativo y de magnitud considerable, indicando que a medida que el adelantado comienza a cambiar su velocidad en dirección hacia el adelantador, este último manifestará una tendencia a alejarse. La misma interpretación se obtiene cuando el adelantado tiene aceleraciones laterales positivas, pues en ese caso la intención de este sería hacer

que sus velocidades sean cada vez mayores, alejándose del adelantador, concluyendo en que el adelantador quiera acercarse a éste para mantener la distancia lateral. Los resultados anteriores, tanto para  $v_{x_2}$  como  $a_{x_2}$ , sugieren que el adelantador muestra una tendencia a mantener la distancia lateral, ya sea si el adelantado presenta intenciones de acercarse o alejarse de él.

Luego, se analizaron las influencias de las variables longitudinales. Se observa que la aceleración lateral  $a_{x_1}$  disminuirá a medida que aumente  $v_{y_1}$ , y esto sugiere que cuando un adelantador adquiera una mayor velocidad longitudinal, sus movimientos en el eje lateral serán de menor intensidad, lo que puede resultar intuitivo. La velocidad  $v_{y_2}$  presenta un coeficiente positivo, indicando que a mayores velocidades longitudinales del adelantado, más querrá el adelantador alejarse de él.

En términos de aceleraciones longitudinales, los signos también tienen la misma interpretación a la de las velocidades; es decir, un adelantador reducirá su aceleración lateral conforme aumente su aceleración longitudinal, pero la aumentará y tenderá a alejarse del adelantado conforme aumente la aceleración longitudinal de este último.

Por último, la variable binaria  $c_2$  indica que cuando los automóviles son camiones, el adelantador tenderá a tener una mayor aceleración lateral, y por ende, mayores velocidades laterales. La variable binaria  $q$ , que en este caso tiene coeficiente positivo, establece que si el adelantamiento es por la izquierda, entonces la aceleración lateral del adelantador tenderá a aumentar en promedio, y por ende, buscará alejarse del adelantado. Esto tiene como posible explicación el mayor control que tiene el conductor de lo que pasa a su izquierda porque está sentado más cerca y tiene mejor perspectiva. Esta misma hipótesis, además, explicaría simultáneamente el que la posición lateral promedio de los vehículos se ubique a la izquierda del centro de la pista (Ramírez, 2020).

Aun cuando interpretaciones anteriores entregan información relevante, es importante destacar que estas relaciones son de carácter lineal que no consideran posibles relaciones no lineales que puedan entregar información valiosa respecto a la naturaleza de estas.

## 4. MODELOS NO-LINEALES

### 4.1. Modelos Aditivos Generalizados (GAM)

La utilización de regresiones no lineales da la posibilidad de encontrar modelos con mayor poder predictivo, pero más importante, pueden entregar mayor información acerca de la relación que hay entre las variables. Desde un punto de vista estadístico, se pueden utilizar modelos paramétricos, y no paramétricos. Las diferencias entre ambos es sustancial: los primeros entregan una ecuación directa acerca de la relación entre las variables, pero a su vez también hay que entregarle al modelo la función que se espera ajustar, mientras que los no paramétricos establecen por sí mismo las curvas no lineales que hay entre las diferentes variables, y si bien estas no entregan una ecuación explícita que establezca la no linealidad en la relación de las variables, si entregan gráficos bastante interpretativos que brindan la posibilidad de observar posibles relaciones no lineales entre diferentes variables. Así, debido a la cantidad de variables a considerar, y a que para estos datos no hay funciones que se puedan distinguir clara y directamente, se utilizaron modelos no paramétricos,

específicamente Modelos Aditivos Generalizados. En consecuencia, los próximos modelos a considerar están dados por las siguientes ecuaciones:

$$a_{x_1} = X_a^T \beta_a + \sum_{i=1}^n f_{ai}(x_{ai}) \quad (4)$$

$$v_{x_1} = X_v^T \beta_v + \sum_{i=1}^n f_{vi}(x_{vi}) \quad (5)$$

Donde  $X^T \beta$ , representa la parte lineal de los modelos utilizada anteriormente, y  $\sum_{i=1}^n f_{\cdot i}(x_{\cdot i})$  son las funciones no paramétricas entre las variables, de modo que para este trabajo dichas funciones se estimaron únicamente utilizando splines de suavizado cúbico (Green & Silverman, 1993). Los modelos consideran tanto en la parte lineal como no lineal todas las variables independientes, ya que si bien esto podría no ser óptimo debido a que puede haber relaciones que sí sean lineales, permite ver las posibles mejoras de los modelos al observar cómo son los gráficos de las interpolaciones encontradas. Este tipo de modelo puede presentar sobreajuste hacia los extremos del dominio si es que existen pocos datos, y esto se ve reflejado en los intervalos de confianza.

## 4.2. Resultados

Para partir a evaluar los resultados de los modelos se calculó primero el índice  $R^2$ , pero ya no es tan adecuado utilizarlo debido a que no se calcula de la misma forma que para las regresiones lineales, pues se utilizan otros tipos de este índice que tratan de ser generalizados para interacciones no lineales. De esta forma, es necesario complementar con los criterios de información Akaike (Akaike, 1974) y Bayesiano (Kass & Raftery, 1995) para analizar cómo evolucionan los modelos.

Variable	Modelo	$R^2$	AIC	BIC
$a_{x_1}$	Regresión Lineal	0,3343	-2433	-2348
	GAM	0,521	-2765	-2519
$v_{x_1}$	Regresión Lineal	0,3277	-1199	-1114
	GAM	0,457	-1401	-1141

Tabla 2: comparación de modelos

De la Tabla 2 se desprende que los modelos GAM que incorporan las funciones no lineales, tienen mejor poder explicativo que los lineales. Esto indica que los modelos no lineales presentan una predicción con mejores resultados para las variables dependientes, pero no que necesariamente todas las variables presenten relaciones no lineales.

Para poder establecer si es que los modelos GAM encontraron relaciones no lineales significativas es necesario estudiar tanto su p-value, como los EDF (Effective Degrees of Freedom). Estos últimos representan la complejidad de la curva estimada, donde un EDF igual 1 es equivalente a una línea recta, si es igual a 2 es equivalente a una curva cuadrática y así, de manera que cuando se observan valores más altos, la curva será más ondulatoria.

Variable de Repuesta: ax1		
	EFD	p-value
s(distP)	6,92035	0,00447
s(ax2P)	0,82579	0,00229
s(vy1P)	5,27547	0,08758
s(vy2P)	3,1618	0,13853
s(vx1P)	5,57193	< 2E-10
s(vx2P)	2,66857	0,06109
s(ay1P)	3,31636	0,03128
s(ay2P)	3,61291	0,22322
s(pvelP1)	1,02027	0,23253
s(pvelP2)	0,07592	0,0021
s(gapM1)	2,81728	0,00858
s(gapM2)	0,00444	0,0021
R2 Aj.	0,52064	

Tabla 3: significancia de variables explicativas

Se puede observar en la Tabla 3 que las variables con menores p-value son  $d, v_{x_1}, v_{x_2}, a_{x_2}, v_{y_1}, a_{y_1}, v_2^p, GAP_1$  y  $GAP_2$ , pero es importante notar que si bien presentan significancia, no necesariamente deberían ser consideradas para estudiar las relaciones, pues algunas pueden llegar a ser curvas complejas y difíciles de interpretar debido a posibles sesgos y sobreajustes, o también podrían presentar comportamientos aparentemente lineales que no otorgan información adicional a la obtenida con los modelos anteriores. Así, a continuación se presentan los gráficos de algunas de las relaciones no lineales más relevantes para este artículo:

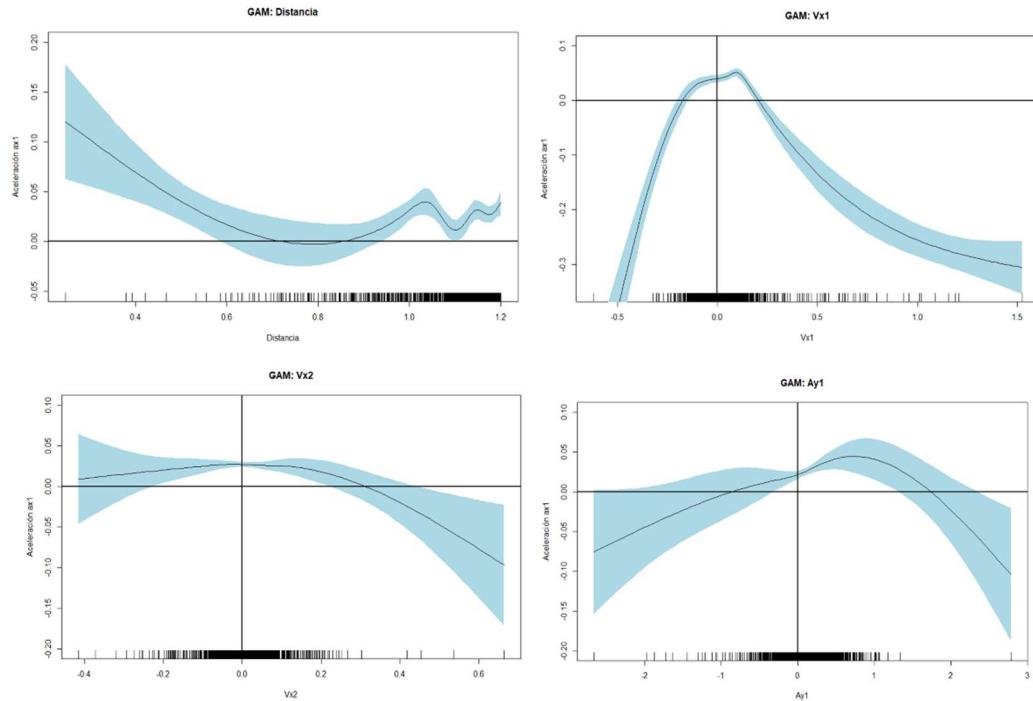


Figura 2: Relaciones no-lineales de la aceleración lateral sugeridas por modelo GAM

En la Figura 2 se puede observar la forma de cada una de las curvas que resultaron ser significativas para el modelo, junto a sus respectivas bandas de probabilidad al 95.

El primer gráfico (2a) muestra que a menor distancia lateral, mayor tenderá a ser la aceleración lateral del adelantador, tendiente a alejarse del vehículo adelantado. Aunque la relación es aparentemente lineal, las bandas de probabilidad no permiten determinarlo concluyentemente. La parte derecha de la curva, específicamente después valores a 1 metro donde se presenta una curva ondulatoria, se puede deber a posibles sesgos de los datos y sobreajuste de la curva.

El segundo gráfico (2b) muestra velocidad vs aceleración lateral del adelantador, y evidencia varias ideas interesantes sobre el comportamiento promedio. En primer lugar, a velocidades laterales bajas, la aceleración es siempre positiva, es decir, tendiente a alejarse del vehículo adelantado. En segundo lugar, la aceleración de vuelve negativa para velocidades altas de cualquier signo. En ambos casos, esto es consistente con que el adelantador dé por terminada la maniobra de adelantamiento y se prepara para lo que viene. En el caso de las velocidades positivas, los adelantadores mayoritariamente frenan para dejar de alejarse del adelantado, ya adquirida la distancia que consideraban prudente. En el caso de las velocidades negativas los adelantadores en promedio aceleran en dirección al adelantado, lo que es consistente con la ejecución en curso de una maniobra de cambio de pista que puede incluso haber motivado el adelantamiento. Nótese que en este caso el rango de velocidades posible es mucho más acotado.

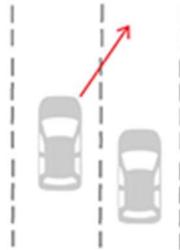


Figura 3: adelantamiento con cambio de pista

En el tercer gráfico (2c), muestra la relación de la aceleración lateral del adelantador respecto a la velocidad del adelantado en el mismo eje. Se observa una aceleración pequeña pero tendiendo significativamente a alejarse del adelantado cuando este último mantiene su posición lateral. Dicha tendencia se anula o incluso revierte cuando el adelantado se está alejando, lo que también es consistente con la prevención de choque.

El cuarto gráfico (2d), por su parte, muestra la relación entre aceleración lateral y longitudinal del adelantador. Nuevamente se observa una ligera tendencia general a alejarse del adelantado, además de una correlación positiva entre ambas aceleraciones cuando el adelantador desacelera o no acelera mucho, lo que es consistente con una tendencia general del volante a una dirección que se aleje del vehículo adelantado. Esta tendencia se revierte para aceleraciones longitudinales grandes, lo que a su vez sería consistente con la preparación de una maniobra de cambio hacia la pista del adelantado (Figura 3).

Cabe destacar que en los modelos GAM utilizados, se introdujeron en los mismos modelos una parte que representa la regresión lineal y la otra que representa los comportamientos no lineales, de manera que ambas partes se complementan para definir la variable dependiente. Además, el

análisis anterior consideró aquellas curvas que presentan interpretabilidad, y no aquellas que se asemejen a relaciones lineales, o curvas con varias oscilaciones que no puedan ser interpretadas de una forma directa e intuitiva. El modelo GAM para la velocidad  $v_{x_1}$  entregó también algunas relaciones no lineales dignas de analizar, las cuales se pueden observar en la Figura 4.

En el primer gráfico (Figura 4a) está asociado a la distancia lateral. En él se puede observar que al igual que para la aceleración  $a_{x_1}$ , existen valores de velocidades más grandes y positivas para distancias pequeñas, es decir, el adelantador tenderá a alejarse. Cabe destacar que esta curva al igual que aquella de la aceleración tiende a estabilizarse conforme las distancias aumentan.

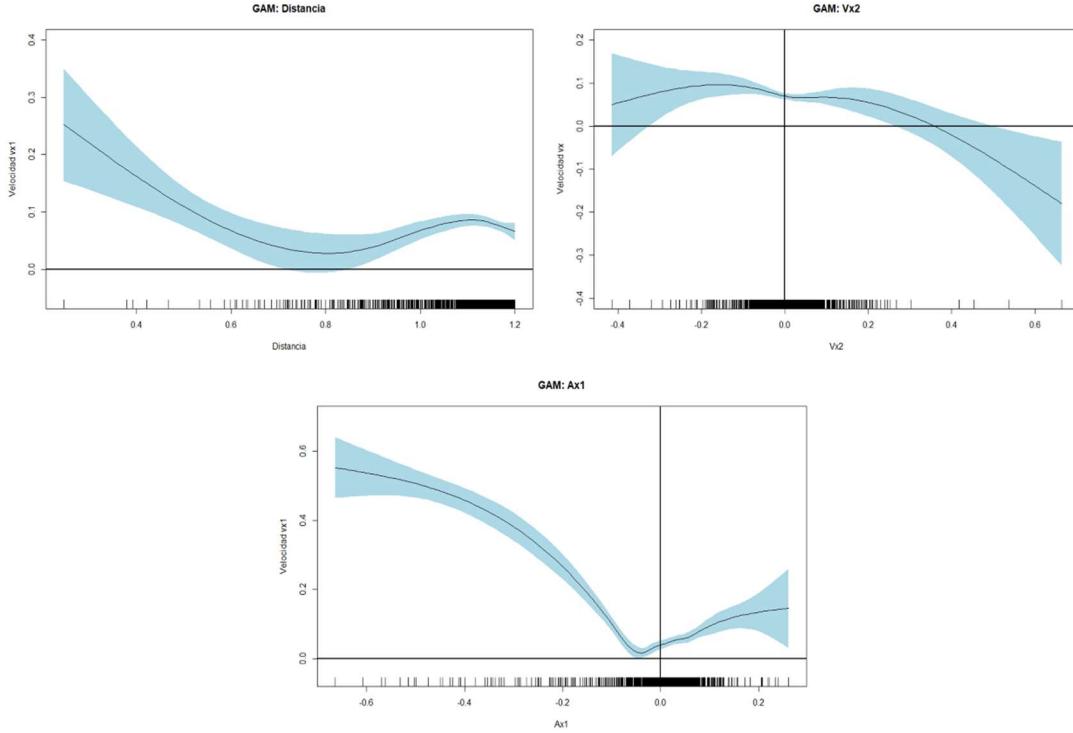


Figura 4: relaciones no lineales de la velocidad lateral

El segundo gráfico (Figura 4b) asociado a  $v_{x_2}$  muestra una posible relación negativa entre ambas velocidades laterales cuando  $v_{x_2}$  es positiva, es decir que el adelantador tenderá a acercarse cuando el adelantado se esté alejando a él. Para valores negativos de la velocidad  $v_{x_2}$  no se puede establecer una interpretación concreta, porque si bien para valores en el intervalo  $[-0,2;0)$  se nota también una relación negativa (el adelantador tiene a alejarse si el otro se acerca), después el intervalo de confianza se vuelve muy grande para valores más negativos.

La tercera figura asociada a  $a_{x_1}$  muestra que la parte negativa de esta variable es más larga que la parte positiva, y esto confirma lo evaluado en el modelo GAM para  $a_{x_1}$ ; es decir que para adelantamientos en los que la distancia lateral es pequeña se esperan una mayor cantidad de combinaciones entre las dos variables donde la aceleración es negativa, y la velocidad es positiva. Esto significa que el conductor se está alejando del adelantado, pero tiende a centralizarse y estabilizarse en su pista. Luego, la curva establece que, para valores positivos de la aceleración, hay valores positivos de la velocidad, y esto puede deberse a posibles cambios de pista del

adelantador a pistas más lejanas, porque las combinaciones en las que las dos variables son positivas indican que cada vez el vehículo tenderá a alejarse más.

### 4.3. Modelos Aditivos Generalizados para Ubicación, Escala y Forma (GAMLSS)

Los modelos lineales y no lineales que se utilizaron entregaron información importante acerca de las relaciones entre las variables, pero aún es posible profundizar en su estudio. Para esto, se utilizaron los modelos GAMLSS (Stasinopoulos & Rigby, 2005), los cuales además de definir una ecuación para el valor esperado de las diferentes variables, permiten definir ecuaciones que definen parámetros de escala, asimetría y curtosis. Lo que hacen estos modelos es primero definir la distribución de probabilidad que sigue la variable dependiente, y luego realizar un estudio sobre si es necesario establecer ecuaciones para cada parámetro de dicha distribución. Dicho lo anterior, se encontró que la distribución Johnson SU se ajusta bien a las distribuciones de  $a_{x_1}$  y  $v_{x_1}$  tanto para cuando hay no hay filtros de distancias, como cuando no los hay.

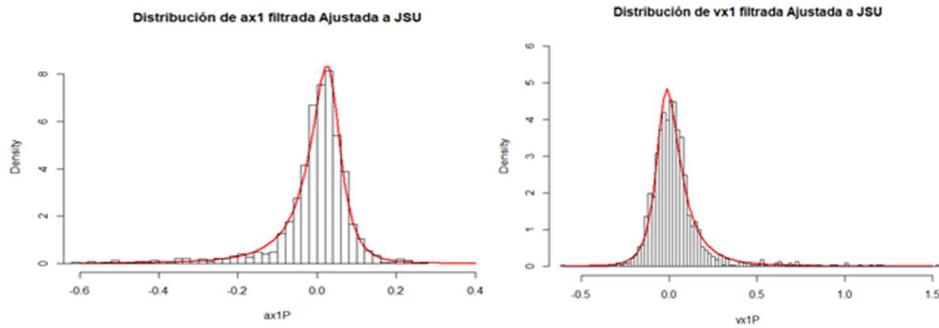


Figura 5: Distribución de aceleraciones y velocidades laterales

Una vez escogida la distribución de probabilidad, es importante decidir qué parámetros tienen que ser modelados, pues la distribución JSU tiene específicamente 4: uno de posición, otro de escala, un tercero de asimetría y el cuarto de curtosis. Debido a que los modelos GAM ya consideran el parámetro de posición (en estos se asumió normalidad en la distribución de las variables dependientes, por lo que este modelaba el valor esperado), es muy útil realizar modelos para el parámetro de escala, pues entregarían información importante respecto a cómo se comporta la dispersión de la variable dependiente para diferentes valores de las independientes. Además, tal como se puede observar en la Figura 5, las variables laterales filtradas parecen tener cierta asimetría: para las dos, una de las colas parece más grande que la otra, por lo que puede ser útil determinar si es que el parámetro de asimetría entrega información relevante. Así, se formulan las siguientes ecuaciones para la variable  $a_{x_1}$ :

$$\begin{aligned}
 a_{x_1} &\sim JSU(\xi, \lambda, \gamma, \delta), \\
 \xi &= X_a^T \beta_a + f_{Dist}^\xi(Dist) + f_{v_{x_1}}^\xi(v_{x_1}) + f_{v_{x_2}}^\xi(v_{x_2}) + f_{a_{y_1}}^\xi(a_{y_1}) \\
 \ln(\lambda) &= f_{Dist}^\lambda(Dist) + \sum_{i=1}^3 f_{i1}^\lambda(x_{i1}) + \sum_{i=1}^4 f_{i2}^\lambda(x_{i2}) \\
 \gamma &= f_{Dist}^\gamma(Dist) + \sum_{i=1}^3 f_{i1}^\gamma(x_{i1}) + \sum_{i=1}^4 f_{i2}^\gamma(x_{i2})
 \end{aligned} \tag{6}$$

Donde  $\sum_{i=1}^3 f_{i1}(x_{i1})$  representan las funciones no paramétricas de las velocidades y aceleraciones del adelantador (que no consideran  $a_{x_1}$ ) y  $\sum_{i=1}^4 f_{i2}(x_{i2})$  son las funciones correspondientes a las aceleraciones y velocidades del adelantado. Por simplicidad, en los parámetros de escala y de asimetría se consideraron variables de velocidades y aceleración de los vehículos partícipes del adelantamiento, y también la variable de su distancia, sin considerar vehículos predecesores ni variables binarias. Cabe notar que inicialmente no se están utilizando las relaciones lineales en los parámetros de escala y asimetría por simplicidad, ya que en caso de que haya un comportamiento de este tipo que sea significativo, podrá ser observado en las relaciones no lineales.

#### 4.4. Resultados

Teniendo en cuenta las ecuaciones para los diferentes parámetros, se establecen 3 modelos: el primer modelo incluye la especificación de que la variable dependiente sigue la distribución Johnson SU y una ecuación para el parámetro de posición, el segundo añade a este último la ecuación para el parámetro de escala, y el último añade la ecuación del parámetro de asimetría.

Modelo	AIC	BIC
GAM	-2860	-2568
GAMLSS (1)	-2959	-2788
GAMLSS (2)	-3147	-2817
GAMLSS (3)	-3165	-2674

Tabla 4: comparación de modelos

Se puede notar entonces que el AIC disminuyó conforme se agregan variables a los modelos, donde alcanzó un valor mínimo de -3165 cuando se define la ubicación, escala y asimetría de la distribución de los datos, por lo que, según este criterio, el mejor modelo para definir  $a_{x_1}$  sería el modelo GAMLSS (3). Analizando los valores del BIC, se aprecia que su menor valor es en el GAMLSS (2), ya que para el GAMLSS(3) aumentó considerablemente. Lo anterior establece que dicho criterio realizó una gran penalización por el ingreso al modelo de la ecuación de asimetría. Debido a que se busca la mayor simplicidad posible para mejor interpretaciones de los resultados, se establece que el mejor modelo es el GAMLSS(2), pues es aquel que tiene mejor BIC.

Si bien el modelo escogido es el mejor en base a los criterios, no necesariamente es óptimo. Lo anterior se debe a que en la ecuación de escala hay 8 variables consideradas, y todas definidas por funciones no lineales, por lo que puede haber relaciones que no aporten mucha información al modelo, y también pueden haber algunas que se comporten de forma lineal, y que sea mejor formularlas de esa forma. Para analizar esto, se hizo un análisis gráfico e iterativo, de manera que las curvas más claras son de las variables  $v_{x_1}$ ,  $v_{y_1}$  y  $v_{y_2}$ , pues las regiones de confianza no son tan amplias, y aquellas asociadas a las últimas dos parecen tener un comportamiento lineal.

Si se quitan las variables no significativas, y se iteran diferentes modelos, se encuentra que hay un modelo cuyas métricas son mejores que las demás, y que se confirma en base a los gráficos anteriores, pues su ecuación de escala está definida de forma no lineal por  $v_{x_1}$ , pero de forma lineal por las velocidades  $v_{y_1}$  y  $v_{y_2}$ . Dicho modelo se define por las siguientes ecuaciones

$$\begin{aligned} a_{x_1} &\sim JSU(\xi, \lambda, \gamma, \delta) \\ \xi &= X_a^T \beta_a + f_{Dist}^\xi(Dist) + f_{v_{x_1}}^\xi(v_{x_1}) + f_{v_{x_2}}^\xi(v_{x_2}) + f_{a_{y_1}}^\xi(a_{y_1}) \\ \ln(\lambda) &= f_{v_{x_1}}^\lambda(v_{x_1}) + \beta_1 v_{y_1} + \beta_2 v_{y_2} \end{aligned} \quad (7)$$

Modelo	AIC	BIC
GAMLSS	-3148	-2947

Tabla 5: Modelo GAMLSS definitivo

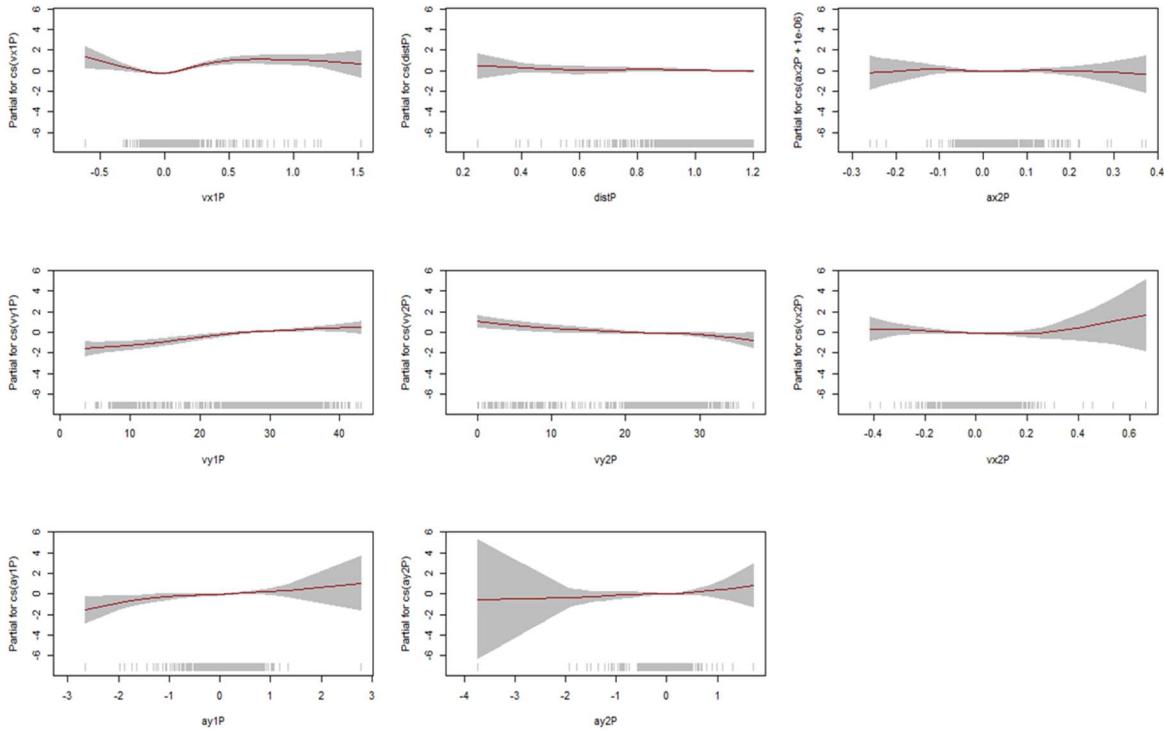


Figura 6: Relaciones del parámetro de escala sugeridas por modelo GAMLSS

La Tabla 5 confirma la elección del modelo, pues presenta mejores criterios que los previamente construidos. El modelo establece una dependencia lineal positiva con  $v_{y_1}$ , y negativa con  $v_{y_2}$ , de modo que a medida que el adelantado adquiera mayor velocidad, la aceleración lateral del adelantador tenderá a alejarse menos de la media, pero cuando el adelantador aumente su velocidad, la aceleración será más dispersa.

El resultado del modelo GAMLSS para la aceleración  $a_{x_1}$  entrega información relevante acerca de su naturaleza y relación con las demás variables, y por lo mismo se realizó durante la investigación

el mismo procedimiento, pero para la variable  $v_{x_1}$ , de manera que los resultados de los modelos creados para dicha velocidad son los mismos que para la aceleración  $a_{x_1}$  (Espinoza, 2020).

## 5. CONCLUSIONES

Dentro de la investigación, se encontraron varios resultados concluyentes acerca de la interacción entre vehículos durante un adelantamiento. La primera, es que hay una influencia mutua entre ambos vehículos durante la maniobra, y se intensifica cuando la distancia es pequeña. Este resultado es muy relevante, pues indica que efectivamente existe una Fricción Lateral, pero más importante, su magnitud estará determinada por la distancia que hay entre los vehículos.

Hay una directa influencia entre las variables de velocidades y aceleraciones laterales del adelantador y del adelantado. El adelantador siente una repulsión y se alejará en caso de que el adelantado tenga intenciones de acercarse a él, pero disminuirá la distancia lateral si el adelantado tiende a alejarse. Esto define que efectivamente el adelantado tiene una tendencia a mantener la distancia lateral que hay entre los vehículos. Los resultados indican también que a medida que hay una mayor velocidad y aceleración longitudinal del adelantado, más querrá alejarse el adelantador. Esto se puede asociar a una constante prevención del choque de parte del adelantador, y que para adelantamientos con velocidades altas, se requerirá un mayor alejamiento.

Existe una relación muy directa e interesante entre la aceleración y velocidad lateral del adelantador. En esta, se concluye que aceleraciones laterales tendientes a acercar ambos vehículos (cuando tanto aceleración como velocidad lateral tienen signos negativos), están asociadas a un cambio de pista al final de la maniobra, y se respalda debido a que lo anterior sucede sólo para velocidades negativas, pero de poca magnitud. Estas aceleraciones que tienden a acercar ambos vehículos pueden estar asociadas a una mayor aceleración longitudinal del adelantador (cambio de pista agresivo, o reacción agresiva no-cooperativa a un intento de cambio de pista del adelantado).

Además, la velocidad longitudinal afecta no sólo a la velocidad y aceleración lateral promedio, sino también a su dispersión. Específicamente, las desviaciones estándar dependen linealmente de la velocidad longitudinal, lo que es consistente con resultados anteriores (Delpiano, 2021).

Por último, es muy importante destacar que los diferentes modelos utilizados, ya sea lineales y no lineales, presentan resultados consistentes entre ellos, de manera que los últimos fueron capaces de mostrar situaciones específicas durante adelantamientos que los lineales no pudieron.

## 6. INVESTIGACIÓN FUTURA

Es importante destacar que si bien correlación entre variables no implica causalidad, se pueden plantear diferentes hipótesis para la explicación de algunas de las relaciones encontradas a lo largo de toda la investigación, cuya demostración es materia idónea de investigación futura. Primero, si bien no se presentaron modelos para variables del adelantado debido a que los modelos del adelantador presentaban métricas considerablemente más precisas, es necesario investigar más a fondo la posibilidad de modelar las variables del primero en situaciones más específicas donde pueda haber una mayor explicación de sus decisiones en base a otras variables cinemáticas.

A lo largo de la investigación se dedujo que las variables interaccionan de mejor forma cuando existen distancias laterales pequeñas, y se escogió un umbral de 1.2 metros para esta investigación, pero esto naturalmente puede variar dependiendo de factores inherentes a la pista, periodo y ubicación, por lo que es idóneo estudiar este mismo fenómeno para otras ubicaciones y situaciones.

Durante el análisis de los modelos GAM, se propuso que comportamientos dentro de algunas curvas se debían a que el adelantador tenía intenciones de cambiarse a la pista del adelantado, pero esto requiere un mayor análisis, es decir estudiar los adelantamientos cuando efectivamente haya un cambio de pista, analizar sus resultados y comparar con los encontrados de esta investigación.

Por último, queda propuesta la posibilidad de definir cada adelantamiento completo de la forma más precisa posible. Lo anterior debido a que, si bien en este trabajo se encontraron resultados precisos al caracterizar cada variable por su promedio durante cada adelantamiento, podrían utilizarse métodos que puedan caracterizar de forma más específica y detallada cada maniobra.

## REFERENCIAS

- Akaike, H., 1974. A new look at the statistical model identification. En: *IEEE Transactions on Automatic Control*, 19(6). s.l.:s.n., pp. 716 - 723.
- Budhkar, A. K. & Maurya, A. K., 2017. Characteristics of lateral vehicular interactions in heterogeneous traffic with weak lane discipline. En: *Journal of Modern Transportation*, 25(2). s.l.:s.n., pp. 74-89.
- Delpiano, R., 2015. *Modelo Microscópico de Tráfico en Dos Dimensiones Basado en Fuerzas Sociales (PhD)*, Santiago de Chile.
- Delpiano, R., 2021. Understanding the lateral dimension of traffic: measuring and modeling lane discipline. *Transportation Research Record*. Forthcoming.
- Delpiano, R., Herrera, J. C., Laval, J. & Coeymans, J. E., 2020. A two-dimensional car-following model for two-dimensional traffic flow problems. En: *Transportation Research Part C: Emerging Technologies*, 114. s.l.:s.n., pp. 504-516.
- Espinoza, J., 2020. Modelos Estadísticos de Tráfico en Dos Dimensiones (Memoria de título). Universidad de los Andes.
- Green, P. J. & Silverman, B. W., 1994. *Nonparametric regression and generalized linear models: a roughness penalty approach*. London: Chapman and Hall.
- Kass, R. E. & Raftery, A. E., 1995. Bayes Factors. En: *Journal of the American Statistical Association*, 90 (430). s.l.:s.n., pp. 773-795.
- R. Krajewski, J. B. L. K. & Eckstein, L., s.f. The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems. En: *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2118-2125.
- Ramirez, D., 2020. *Caracterización de la Disciplina de Pistas en Autopistas Urbanas de Santiago mediante Técnicas de Visión Computacional*. (Memoria de título). Universidad de los Andes.
- Stasinopoulos, M. & Rigby, R. A., 2005. Generalized additive models for location, scale and shape. En: *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(3). s.l.:s.n., pp. 507-554.
- Stasinopoulos, M. y otros, 2017. *Flexible Regression and Smoothing Using GAMLSS in R*. 1 ed.:Chapman and Hall.

