



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

DISEÑO DE UN MODELO ESPACIO-TEMPORAL PARA CARACTERIZAR Y  
PREDECIR EL RIESGO CRIMINAL Y ANÁLISIS DE SU DESEMPEÑO USANDO  
TRANSFORMACIONES BASADAS EN KERNELS GAUSSIANOS PERIÓDICOS Y  
FUNCIONES WAVELET EN LA FENOMENOLOGÍA TEMPORAL.

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN CIENCIAS DE LA  
INGENIERÍA, MENCIÓN ELÉCTRICA.

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO.

PABLO FRANCISCO FUENTES POBLETE

PROFESOR GUÍA:  
MARCOS EDUARDO ORCHARD CONCHA

MIEMBROS DE LA COMISIÓN:  
CÉSAR AZURDIA-MEZA  
JUAN CARLOS AGÜERO VÁSQUEZ  
MARCOS ORCHARD CONCHA

Este trabajo ha sido parcialmente financiado por por CONICYT 4191/2015 y Carabineros  
de Chile

SANTIAGO DE CHILE.

2017



RESUMEN DE LA MEMORIA PARA OPTAR  
AL TÍTULO DE MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN ELÉCTRICA.  
E INGENIERO CIVIL ELÉCTRICO.  
POR: PABLO FRANCISCO FUENTES POBLETE  
FECHA: 2017  
PROF. GUÍA: SR. MARCOS EDUARDO ORCHARD CONCHA

DISEÑO DE UN MODELO ESPACIO-TEMPORAL PARA CARACTERIZAR Y PREDECIR  
EL RIESGO CRIMINAL Y ANÁLISIS DE SU DESEMPEÑO USANDO  
TRANSFORMACIONES BASADAS EN KERNELS GAUSSIANOS PERIÓDICOS Y  
FUNCIONES WAVELET EN LA FENOMENOLOGÍA TEMPORAL.

Una de las mayores preocupaciones a nivel nacional es velar por la seguridad ciudadana [1]. Gracias al masivo almacenamiento de datos es posible utilizar herramientas computacionales que identifiquen los patrones criminales, ayudando a optimizar los recursos policiales destinados a mitigar la delincuencia, siendo Predpol<sup>®</sup> pionero en el rubro de análisis y predicción de eventos criminales [2].

Este trabajo de tesis muestra un modelo basado en teoría de kernels y análisis de espectros de frecuencia mediante Transformada Wavelet para caracterizar la evolución espacio-temporal del riesgo criminal, permitiendo predecir la actividad delictual. El modelo utiliza información espacio-temporal de eventos criminales ocurridos anteriormente para generar una función de intensidad dinámica, la cual indica donde existe una mayor “probabilidad” de ocurrir un evento criminal en el futuro. Se separa una base de datos criminal otorgada por Carabineros de Chile en dos conjuntos: una denominada fase *priori* (captura de la tendencia criminal histórica) y otro como fase *posteriori* (captura de la tendencia criminal reciente). Usando la información espacial de los datos en cada conjunto se procede a entrenar las medias, covarianzas y pesos asociados a cada componente que define a un *Gaussian Mixture Model* (GMM). Definidos los componentes espaciales se procede a caracterizar el comportamiento temporal de los delitos asignados a cada componente por medio del entrenamiento de Kernel Periodicos Gaussianos (KPG) o por medio de análisis de Transformada Wavelet (TW). Finalmente se entrega un modelo de predicción espacio-temporal del riesgo criminal capaz de integrar la fase *priori* y *posteriori* mediante un coeficiente inter-proceso. El modelo es evaluable en diferentes comunas y su visualización es por medio de grillas espaciales de  $150 \times 150 m^2$  y muestrada cada ocho horas según los turnos policiales.

Este Trabajo de Tesis permitio diseñar e implementar un modelo probabilístico espacio-temporal basado en delitos geo-refenciados usando dos métodos de caracterización temporal: una mediante sumas de KPG y otro analizando espectros de frecuencia por medio de TW. Siendo posible comparar sus desempeños utilizando un conjunto de entrenamiento y prueba. Se desarrollaron optimizaciones paramétricas en los modelos temporales basado en minimizar el error cuadrático medio del conjunto de prueba de cada componente espacial. Los resultados arrojan que el proceso de caracterización temporal por medio de TW tiene una mejor tasa de aciertos promedio que los modelos de KPG según métricas de desempeños especialmente definidas para el problema: *High Risk Porcentage* ( $HRP_{\theta}$ ), *True Incident Percentage* ( $TIP_{\theta}$ ) y *Area Under Curve* (AUC) [3]. El modelo de caracterización espacio-temporal del riesgo criminal es ejecutado en 10 comunas de la Región Metropolitana, obteniendo una adecuada tasa de acierto con una reducida cantidad de área, siendo factible su uso para predecir en diferentes comunas. Finalmente se entrega un estudio de la importancia de la información de la fase *priori* respecto a la *posteriori* en el desempeño del modelo espacio-temporal de riesgo criminal.



*“No trato de describir el futuro, trato de prevenirlo.”*  
*Ray Bradbury. Fahrenheit 451.*



# Agradecimientos

Hago entrega de este Trabajo de Tesis agradeciendo por sobre todas las cosas a Dios. Él es el motor de mi felicidad y a Él entrego mi vida para ser instrumento de Fe en este mundo.

A mi mamá por su amor incondicional, por darme refugio en los momentos difíciles de la vida, por sus bellas manos que acarician y me llenan de paz en todo momento.

A mi papá por entregarme las herramientas indispensables para surgir como persona, por traer el pan a casa y darnos un refugio para dormir. Por su preocupación y sinceridad al momento de decirme las cosas.

A mi hermana, que es un ejemplo a seguir en lo académico y en lo personal, mostrándome que con esfuerzo todo es posible en la vida.

A mi familia entera, mis abuelos, tíos y primos, porque lo son todo para mí.

A Constanza Silva, que durante estos 5 años me ha llenado el corazón de bellas experiencias. Por tus abrazos, regalones; por comprenderme y aceptarme. Gracias por ser parte de tu vida.

A mis amigos *cartoneros*, a Matias, Lisa, Tamara, Joakin, Angelo, Seba, Kirby, Mariosky, Nelson, Aldaka y Americo, con quienes compartí los momentos más alegres durante la universidad. Por las tallas y bonitos recuerdos que tendré siempre en mi memoria.

Al Coro de Ingeniería, por darme la oportunidad de hacer lo que más amo en la vida. A sus integrantes, siempre alegres y dispuestos a acoger a más gente y hacer crecer sus sueños. A su directora Verónica que, por su entrega y dedicación, me demostró que el amor a un oficio puede llevar a la felicidad.

A la rama de natación y la profesora Mónica, con quienes aprendí a no darme por vencido y a ser perseverante para lograr mis metas.

A mis profesores de la universidad y del colegio, que permitieron que fuera posible que yo llegara tan lejos en mi vida académica. A mi profesor guía, Marcos Orchard, por recibirme siempre con un cálido abrazo y con una sonrisa en su rostro ante cualquier duda.

A Joaquin Amigo Cortes, por enseñarme que una sonrisa puede transformar el día de muchas personas y que la felicidad esta escondida en cosas pequeñas.

A CONICYT y Carabineros de Chile, por financiar y darme la oportunidad de disfrutar este trabajo de investigación.





# Tabla de Contenido

Nomenclatura	xvi
<b>Introducción</b>	<b>1</b>
<b>1. Estado del Arte y Marco Teórico</b>	<b>5</b>
1.1. Evolución de los modelos de caracterización criminal. . . . .	5
1.2. Transformaciones por medio de Kernel . . . . .	6
1.2.1. Kernel Density Estimation . . . . .	6
1.2.2. Mixture Model . . . . .	7
1.2.3. Kernel Periódico Gaussiano . . . . .	11
1.3. Análisis Wavelet . . . . .	12
1.3.1. Transformada Fourier . . . . .	12
1.3.2. Transformada Wavelet . . . . .	13
1.3.3. Funciones Wavelet . . . . .	14
1.3.4. Espectro de energía Wavelet . . . . .	16
1.3.5. Elección de escalas . . . . .	16
1.3.6. Anti-transformada Wavelet (ATW) . . . . .	17
1.4. Medidas de desempeño . . . . .	17
<b>2. Metodología</b>	<b>19</b>
2.1. Introducción . . . . .	19
2.2. Base de datos . . . . .	20
2.2.1. Datos geográficos . . . . .	20
2.2.2. Datos de eventos criminales . . . . .	20
2.3. Pre-procesamiento de datos . . . . .	21
2.3.1. Pre-procesamiento temporal . . . . .	22
2.3.2. Pre-procesamiento espacial . . . . .	24
2.4. Procesamiento espacial . . . . .	24
2.4.1. Selección de condiciones iniciales . . . . .	25
2.4.2. Entrenamiento de GMM . . . . .	26
2.5. Procesamiento temporal . . . . .	26
2.5.1. Llenado de arreglos temporales y suavizado . . . . .	26
2.5.2. Modelo temporal de Kernel Periódicos Gaussianos . . . . .	29
2.5.3. Modelo temporal Wavelet . . . . .	31
2.6. Predicción . . . . .	34
2.6.1. Evaluación del riesgo en la comuna . . . . .	35
2.6.2. Uso de métricas de desempeño . . . . .	35

<b>3. Resultados</b>	<b>38</b>
3.1. Herramientas de hardware y software utilizados . . . . .	38
3.2. Casos de estudio . . . . .	39
3.3. Resultado base . . . . .	40
3.3.1. Modelo KPG . . . . .	43
3.3.2. Modelo WV . . . . .	46
3.3.3. Predicción espacio-temporal del riesgo criminal . . . . .	50
3.4. Optimización de modelos temporales . . . . .	52
3.4.1. Optimización modelo KPG . . . . .	53
3.4.2. Optimización modelo Wavelet . . . . .	57
3.4.3. Análisis del proceso de optimización. . . . .	60
3.5. Respuesta en diferentes comunas . . . . .	64
3.6. Influencia del coeficiente inter-proceso temporal $\lambda$ . . . . .	69
<b>Conclusión</b>	<b>72</b>
<b>Bibliografía</b>	<b>76</b>
<b>A. Clasificación de los delitos utilizados en el trabajo de tesis.</b>	<b>77</b>
<b>B. Resultados optimización modelo KPG.</b>	<b>78</b>
<b>C. Resultados optimización modelo WV.</b>	<b>82</b>
<b>D. Resultados de modelos de caracterización variando el coeficiente inter-proceso <math>\lambda</math>.</b>	<b>83</b>
<b>E. Curvas promedios para distintos valores AUC.</b>	<b>86</b>

# Índice de Tablas

2.	Variación de los casos policiales acumulados a nivel nacional durante el primer semestre de 2016 . . . . .	2
1.1.	Tres funciones Wavelet expresadas en su forma temporal y frecuencial de Fourier. La función DOG es la abreviación de <i>Derivative Of a Gaussian</i> . $H(\omega)$ es la función Heaviside $H(\omega) = 1$ , si $\omega > 0$ y es 0 para cualquier otro valor. . .	15
1.2.	Relación entre la escala $s$ y el periodo bajo el concepto de Fourier para las funciones Wavelet Morlet, Paul y DOG. Tanto $\omega_0$ como $m$ son parámetros de la función. . . . .	17
3.1.	Características de los distintos hardware utilizados en el Trabajo de Tesis. . .	39
3.2.	Parámetros y datos utilizados para generar el entrenamiento espacial de los datos en la comuna de Maipú en la fecha 04-06-2016 . . . . .	41
3.3.	Periodos seleccionados utilizando el análisis Wavelet con parámetros base para la fase <i>priori</i> . . . . .	47
3.4.	Periodos seleccionados utilizando el análisis Wavelet con parámetros base para la fase <i>posteriori</i> . . . . .	48
3.5.	Arreglos escogidos para el estudio de optimización en el modelo KGP. Cada estudio muestra el tipo de arreglo utilizado, donde el primer y último valor han sido denotados entre corchetes y la diferencia entre un valor con el siguiente ha sido explicado fuera de estos. . . . .	54
3.6.	Matriz de pares combinados a analizar a partir de los arreglos de periodos agrupados ( <i>bp</i> ) y valores del parámetro $m$ . . . . .	57
3.7.	Comportamiento de la métrica AUC para cada uno de los modelos estudiados y sus variantes en el primer turno de la fecha de predicción. . . . .	61
3.8.	Comportamiento de la métrica AUC para cada uno de los modelos estudiados y sus variantes en el segundo turno de la fecha de predicción. . . . .	61
3.9.	Comportamiento de la métrica AUC para cada uno de los modelos estudiados y sus variantes en el tercer turno de la fecha de predicción. . . . .	61
3.10.	Comportamiento del tiempo de procesamiento para cada uno de los modelos estudiados al momento de generar 3 superficies de riesgos en el día de predicción. . . . .	63
3.11.	Comportamiento de los MSE de prueba del proceso <i>priori</i> para cada uno de los modelos temporales en el mes de junio 2016 . . . . .	63
3.12.	Comportamiento de los MSE de prueba del proceso <i>posteriori</i> para cada uno de los modelos temporales en el mes de junio 2016 . . . . .	63

3.13. Acumulación de todos los primeros turnos de predicción analizados durante todo el mes de junio en las diferentes comunas en base a distintos intervalos AUC. . . . .	67
3.14. Acumulación de todos los segundo turnos de predicción analizados durante todo el mes de junio en las diferentes comunas en base a distintos intervalos AUC. . . . .	67
3.15. Acumulación de todos los tercer turnos de predicción analizados durante todo el mes de junio en las diferentes comunas en base a distintos intervalos AUC. . . . .	68
3.16. Tiempo promedio en diferentes comunas al momento de generar 3 superficies de riesgo respectivas al primer, segundo y tercer turno del día siguiente a la fecha de consulta usando el modelo espacio-temporal de riesgo criminal con técnica temporal Wavelet. . . . .	69
3.17. Mejor y peor desempeño promedio del modelo espacio-temporal usando diferentes valores coeficiente inter-proceso $\lambda$ en la comuna de Maipú en el mes de junio del 2016 para cada uno de los turnos de predicción. . . . .	71
A.1. Delitos de mayor connotación social (DMCS) la base de datos. . . . .	77

# Índice de Ilustraciones

1.1.	Influencia de los parámetros asociados al KPG. Los gráficos entregados son ejecutados con parámetros $c_t = 2$ , $\sigma = 2$ , $p = 10$ y diferentes valores de $l$ explicitados en cada curva coloreada. . . . .	12
1.2.	Tres diferentes funciones Wavelet generadas a partir de la tabla 1.1. Los gráficos de la izquierda entregan el comportamiento temporal de las funciones, siendo la línea sólida la parte real y la línea punteada la parte imaginaria. Los gráficos de la derecha entrega el comportamiento en el dominio de la frecuencia. En este caso se tomó en consideración una escala igual a $s = 10\delta t$ . . . . .	15
2.1.	Esquema general de la metodología implementada para la caracterización del riesgo criminal. . . . .	19
2.2.	Esquema representativo de los procesos y resultados de la etapa de pre-procesamiento. . . . .	22
2.3.	Construcción de los arreglos temporales a <i>priori</i> , <i>posteriori</i> y predicción. La unidad mínima definida es la unidad de medida temporal uniforme ( <i>slot</i> ) que abarca un delta tiempo de $x$ horas. . . . .	23
2.4.	Sub-arreglos de entrenamiento y prueba de los arreglos $AT_{pr}$ y $AT_{po}$ (arreglo de arriba y abajo respectivamente). Las barras amarillas denotan el comienzo de cada uno de los sub-arreglos. El caso $AT_{pr}$ mostrado es idéntico al explicado anteriormente, dando énfasis a la no inclusión de la fecha de 24 de julio del 2016 en el arreglo de entrenamiento. Por otra parte, el caso $AT_{po}$ se denotan la duración del arreglo de entrenamiento y de prueba, especificando las fechas de inicio y final para un caso arbitrario. . . . .	23
2.5.	Esquema representativo de los procesos y resultados de la etapa de procesamiento espacial. . . . .	24
2.6.	Tipos de respuestas en la caracterización por medio de GMM para diferentes $N_{kG}$ . En a) representa el caso de sub-representación ( $N_{kG} = 2$ ), en b) se muestra el valor óptimo de $N_{kG}$ y c) el caso de sobre-representación usando $N_{kG}$ igual a la cantidad de delitos cometidos en el área de estudio. . . . .	25
2.7.	Esquema representativo de los procesos y resultados de la etapa de procesamiento temporal. . . . .	27
2.8.	Llenado de parte del arreglo temporal a <i>priori</i> para kernel espacial $k_3$ . Las estrellas indican los eventos criminales ocurridos en cada componente, generando un diccionario de eventos. El llenado resulta una señal discreta donde se muestran la cantidad de delitos que caen en un <i>slot</i> determinado. . . . .	28
2.9.	Proceso de suavizado de la señal temporal discreta haciendo uso del método KDE. . . . .	28

2.10.	Esquema de correlación de una señal STS con un set de $K_{pT}$ y arreglos de centros y periodos igual a $c_t^r = [0; 0,5; 1; 1,5; 2]$ y $p^r = [0,5; 1,0; 1,5; 2,0]$ respectivamente. La Figura 2.10 a) muestra el proceso de correlación entre $K_{pT}$ con centro $c_t^r = 0$ y diferentes valores de $p^r$ . La Figura 2.10 b) muestra el resultado final del proceso de correlación, usando los arreglos $p^r$ y $c_t^r$ completamente. Finalmente la Figura 2.10 c) muestra la selección de los cinco mayores valores de correlación y con ello los parámetros que hicieron posible dicha correlación	31
2.11.	Fechas de corte tomadas en cada uno de los arreglos temporales para el método Wavelet. . . . .	33
2.12.	Ejemplo de comuna con 12 grillas evaluadas con el modelo $R^{pred}$ normalizado para un tiempo arbitrario. . . . .	36
3.1.	Comunas del Gran Santiago. En color morado se muestra la comuna de Maipú. Imagen capturada gracias a herramientas de visualización dadas por QGIS®.	40
3.2.	Proceso de caracterización espacial de los eventos criminales usando técnicas de entrenamiento de GMM. Se selecciona solo la zona urbana de la comuna de Maipú. . . . .	41
3.3.	Histogramas de asignación de delitos para cada una de las etapas de la caracterización espacial. . . . .	42
3.4.	Señales de tiempo suavizadas (STS) para dos casos particulares después del proceso de entrenamiento espacial. . . . .	43
3.5.	Matrices de correlación generada usando $A_p^r$ y $A_c^r$ base. Los puntos celestes corresponden a los máximos locales encontrados usando una búsqueda basada en procesamiento de imágenes de ventana 5x5. . . . .	44
3.6.	Búsqueda del mejor candidato $N_{kG}$ usando la información de los errores asociados a cada una de las fase de evaluación. . . . .	45
3.7.	Curvas de comparación para el proceso <i>priori</i> de la STS del kernel espacial	29. 45
3.8.	Curvas de comparación para el proceso <i>posteriori</i> de la STS del kernel espacial	34. . . . . 46
3.9.	Análisis Wavelet de la señal temporal suavizada de la componente espacial	29. 47
3.10.	Análisis Wavelet de la señal temporal suavizada de la componente espacial	10. 48
3.11.	Segmentación semanal de la señal reconstruida por medio de ATW y con su factor de olvido asociado a cada una de las curvas. . . . .	49
3.12.	Señales de predicción para cada uno de los procesos usando el modelo temporal usando técnicas de Transformada Wavelet. . . . .	49
3.13.	Superficie de riesgo en formato 3D para el primer <i>slot</i> de predicción del 04-06-2016 de la comuna de Maipú. . . . .	50
3.14.	Riesgo criminal entregado para el primer <i>slot</i> de predicción del 04-06-2016 de la comuna de Maipú usando modelo Wavelet. Las grillas entregadas corresponden al 10 % del total de grillas definidas en zona urbana (202 de 2015 grillas) . .	51
3.15.	Curvas de desempeño generadas para el caso base, en los tres <i>slots</i> definidos para la fecha de predicción . . . . .	52
3.16.	Ejemplo de separación temporal del conjunto de entrenamiento y prueba en los diferentes procesos para optimización Wavelet . . . . .	53
3.18.	Ejemplos de curvas de desempeño para diferentes procesos usando el mejor par de componentes $(m, bp)$ para la señal de un componente espacial específico en el 19 de junio del 2016. . . . .	58

3.19. Cantidad de componentes espaciales que minimizan el MSE usando el par de parámetros $(m, bp)$ descritos en la matriz. . . . .	58
3.21. Comunas que entran en el proceso de caracterización del riesgo criminal durante todo el mes de junio del 2016. . . . .	65
3.22. Respuestas por medio de la métrica AUC de diferentes comunas aplicando el modelo de caracterización de riesgo criminal usando técnicas temporales WV. . . . .	66
C.1. Valores de MSE de prueba diarios para la fase <i>priori</i> y <i>posteriori</i> ocupando el método de optimización Wavelet paramétrica . . . . .	82
D.1. Promedio de valores AUC diarios del primer, segundo y tercer turno utilizando los datos del mes de junio 2016 en la comuna de Maipú utilizando modelo WV para caracterizar el comportamiento temporal con diferentes coeficientes inter-proceso $\lambda$ . En el eje de las abscisas se ordenan de mayor a menor los valores de $\lambda$ , partiendo en 0 y terminando en 1, con cada paso de avance igual a 0,1. Las líneas horizontales entregan la varianza de los datos para cada caso de estudio. . . . .	83
D.2. Promedio de valores AUC diarios del primer, segundo y tercer turno utilizando los datos del mes de junio 2016 en la comuna de Maipú utilizando modelo KPG con el caso de estudio 0 para caracterizar el comportamiento temporal con diferentes coeficientes inter-proceso $\lambda$ . En el eje de las abscisas se ordenan de mayor a menor los valores de $\lambda$ , partiendo en 0 y terminando en 1, con cada paso de avance igual a 0,1. Las líneas horizontales entregan la varianza de los datos para cada caso de estudio. . . . .	84
D.3. Promedio de valores AUC diarios del primer, segundo y tercer turno utilizando los datos del mes de junio 2016 en la comuna de Maipú utilizando modelo KPG con el caso de estudio 2 para caracterizar el comportamiento temporal con diferentes coeficientes inter-proceso $\lambda$ . En el eje de las abscisas se ordenan de mayor a menor los valores de $\lambda$ , partiendo en 0 y terminando en 1, con cada paso de avance igual a 0,1. Las líneas horizontales entregan la varianza de los datos para cada caso de estudio. . . . .	85
E.1. Comportamiento promedio de la curva HRP-TIP con diferentes intervalos de valores AUC. . . . .	86

# Nomenclatura

$AT_{po}$	Arreglo Temporal <i>posteriori</i>
$AT_{pred}$	Arreglo Temporal de predicción
$AT_{pr}$	Arreglo Temporal <i>priori</i>
$J$	largo del arreglo de escalas
$K_{pT}$	expresión matemática de la función KPG
$N_{kG}$	número de componentes presentes en la GMM
$N_{kT}$	número de funciones de Kernels Periódicos Gaussianos ( $K_{pT}$ ) asociados a una componente espacial de la GMM
$R^{pred}$	modelo de predicción del riesgo criminal
$R_M$	modelo temporal de Kernel Periódicos Gaussianos
$R_W^{post}$	modelo temporal <i>posteriori</i> Wavelet
$R_W^{prio}$	modelo temporal <i>priori</i> Wavelet
$S = \{s_j\}_{j=0, \dots, J}$	conjunto de escalas $s$ de Transformada Wavelet
$Wn$	Transformada Wavelet continua de una secuencia temporal $x_n$ discreta
$\beta_i^r$	peso asociado a la $K_{pT}$ $i$ -ésima del componente espacial $r$
$\delta j$	resolución del arreglo de escalas aplicado
$\lambda$	coeficiente inter-proceso temporal
$\psi(\eta)$	función Wavelet normalizada
$\psi(\eta)^*$	complejo conjugado de $\psi(\eta)$
$\psi_0(\eta)$	función Wavelet
$bp$	ancho de las bandas aplicada en el arreglo de Energía Acumulada por Escala
$c_i^r$	fase asociada a la $K_{pT}$ $i$ -ésima del componente espacial $r$
$f_{post}$	fecha de inicio para el estudio <i>posteriori</i>
$f_{pred}$	fecha de predicción
$m$	parámetro interno de la función Wavelet
$p_i^r$	periodo asociado a la $K_{pT}$ $i$ -ésima del componente espacial $r$
$r$	componente espacial arbitraria de la GMM entrenada
$s$	parámetro interno de la función Wavelet que indica la escala temporal asignada
$s_0$	escala mínima de resolución
$slot$	unidad de medida temporal uniforme
$A_c^r$	arreglo con candidatos a condiciones óptimas del parámetro $c_i^r$
$A_p^r$	arreglo con candidatos a condiciones óptimas del parámetro $p_i^r$
ATW	Anti-Transformada Wavelet
AUC	<i>Area Under Curve</i>



CEAMOS	Centro de Análisis y Modelamiento en Seguridad
CEP	Centro de Estudios Público
DFT	Transformada Discreta de Fourier
DMCS	Delitos de Mayor Connotación Social
EAE	Energía Acumulada por Escala
EETW	Espectro de Energía de Transformada Wavelet
EM	<i>Expectation Maximization</i>
GMM	<i>Gaussian Mixture Model</i>
$HRP_{\theta}$	<i>High Risk Percentage</i>
KDE	<i>Kernel Density Estimation</i>
KPG	Kernel Periodicos Gaussianos
MEW	Matriz de Energía Wavelet
ML	<i>Maximum Likelihood</i>
modelo KPG	modelo temporal usando técnicas de transformación mediante Kernels Periódicos Gaussianos
modelo WV	modelo temporal usando técnicas de Transformada Wavelet
MSE	<i>Mean Square Error</i>
SAIT	Sistema de Análisis de Información Territorial
$SES_q^r$	señal temporal con fuerte presencia de la escala $s_q$
$SST^r$	señal de soporte temporal
STS	señal temporal suavizada por medio de KDE
$STS^r$	señal temporal suavizada de una componente espacial $r$ arbitraria
TF	Transformada de Fourier
$TIP_{\theta}$	<i>True Incident Percentage</i>
TW	Transformada Wavelet



# Introducción

Latinoamérica es una de las regiones con mayor concentración de países de habla española a nivel mundial (15 de 21 países), rico en materias primas tales como petróleo, hierro, café, carne, frutas y verduras, entre otros. Con una gran riqueza cultural de fuertes raíces precolombinas y en donde cada país se relaciona con sus pares con tal de obtener beneficio mutuo. Cada gobierno combate con las diferentes problemáticas que acontecen en su diario vivir, tanto a nivel nacional como internacional, siendo algunas de sus prioridades erradicar el hambre [4], combatir con la desigualdad económica y social, mejorar la calidad y atención de la salud pública, luchar por una educación incluyente e igualitaria y velar por la seguridad ciudadana. El grado de importancia entregado a cada una de estas problemáticas depende de la agenda de gobierno, de sus proyectos o de las circunstancias que viven sus ciudadanos día a día.

En Chile, según el estudio realizado por el Centro de Estudios Público (CEP) entre el mes de julio y agosto del 2016, los tres problemas a los cuales el gobierno tiene que prestar mayor atención son: (i) delincuencia, (ii) salud y (iii) corrupción [1]. Lamentablemente, las actuales políticas públicas no han podido entregar una respuesta certera a los tres problemas de mayor envergadura, siendo evaluada cada una de ellas con notas bajo 3 en una escala del 1 al 7.

Abocándose específicamente en delincuencia, según datos de la fundación Paz Ciudadana, en el 2014 el índice de victimización total del país alcanzó un 43,5 %, 4,6 puntos porcentuales más que el año 2013. En el año 2015 la delincuencia es la principal preocupación de los chilenos (un 58 % de la población la menciona como prioridad según la encuesta del 2015 realizada por el CEP [5], superando incluso a temas como educación y salud) y es una de las áreas peor evaluadas en cuanto a la gestión del gobierno, obteniendo valores por sobre el 80 % de desaprobación desde el mes de noviembre del 2014 hasta junio del 2015 la encuesta Adimark [6]. Como consecuencia de las circunstancias mencionadas, la problemática en torno a la delincuencia ha pasado a tener protagonismo en las gestiones actuales de la administración pública [7].

A partir del 2010 se han creado diversos métodos para prevenir la delincuencia. Uno de los más importantes es el plan cuadrante de seguridad preventiva, el cual se define como la estrategia operacional de Carabineros de Chile, para enfrentar las demandas siempre crecientes por servicios policiales de vigilancia y seguridad que requiere la ciudadanía y, que la institución entrega en forma permanente las 24 horas, todos los días del año y en todo el territorio de la república, en el contexto urbano [8]. Otros métodos formulados por el gobierno para mitigar la delincuencia son la modernización del sistema de inteligencia nacional, generación de instrumentos para mejorar la convivencia vecinal y promover la paz social entre otros, los cuales son mostrados en Plan Nacional de Seguridad Pública y Prevención de la Violencia y el Delito publicado en agosto del 2014 [9].

Según los datos recolectados a nivel nacional por el análisis de umbral de delitos realizada por la paz ciudadana a mediados de agosto del 2016 [10] los Delitos de Mayor Connotación Social (DMCS) junto con robo en lugar habitado, robo de vehículo, hurtos y homicidios han presentado una disminución significativa desde el año 2015 al 2016, mientras que violación y lesiones presentan una disminución leve. Por otra parte, delitos clasificados como robo con violencia, robo sorpresa, robo en lugar no habitado y robo de accesorio vehicular se mantienen sin variación significativa, tal como se muestra en la Tabla 2. Sin embargo, en el mismo estudio realizado por paz ciudadana, existen 37 comunas que han aumentado significativamente la cantidad de robos con violencia, siendo San Ramón la comuna más afectadas en este caso, teniendo un aumento del 22.3% a finales de junio del 2016 respectivamente. Otro caso se da con el robo en lugar habitado, del cual 47 comunas han aumentado significativamente, siendo las más afectadas Colipulli con un aumento del 318.7% de los delitos ocurridos en esta categoría, Diego de Almagro con un 152.6%, San Ramón con un 80.3% y Quinta Normal con un 39.6%.

Delito	Acumulado a junio del 2015	Acumulado a junio del 2016	Diferencia entre 2015 y 2016	Variación relativa entre 2015 y 2016
DMCS	294.698	279.752	-14.946	-5.3 %
Robo violencia	32.409	31.732	-677	-2.1 %
Robo por sorpresa	19.265	17.867	-1.398	-7.8 %
Robo lugar habitado	34.540	31.285	-3.255	-10.4 %
Robo lugar no habitado	24.987	24.235	-752	-3.1 %
Robo de vehículo	16.093	14.745	-1.348	-9.1 %
Robo de accesorio vehículo	31.770	31.841	71	0.2 %
Hurtos	90.377	86.596	-3.781	-4.3 %
Lesiones	40.398	37.064	-3.334	-8.9 %
Homicidio	286	221	-65	-29 %
Violación	1.411	1.329	-82	-6.1 %

Tabla 2: Variación de los casos policiales acumulados a nivel nacional durante el primer semestre de 2016

El interés actual por parte de las políticas de Seguridad Ciudadanas y Carabineros de Chile es modernizar el proceso de patrullaje tanto en recursos humanos como logístico. A finales de octubre del 2016, 28 vehículos de patrullaje de última generación fueron distribuidos en diferentes comisarias. Además se ha construido una nueva sub-comisaria en Diego Portales y se ha implementado una nueva forma de analizar el crimen gracias al moderno Sistema de Análisis de Información Territorial (SAIT), que permite generar plantillas de patrullaje para diferentes horas del día. Uno de los objetivos actuales es predecir las zonas de mayor riesgo criminal con tal de poder distribuir eficientemente los recursos policiales y así disminuir la cantidad de delitos ocurridos en zonas publicas haciendo uso de un patrullaje preventivo con visión al futuro [11] [12].

# Hipótesis a demostrar

Es posible realizar una apropiada caracterización espacio-temporal del riesgo criminal en diversas comunas de la Región Metropolitana usando un modelo que mezcle el comportamiento espacial analizado por *Gaussian Mixture Model* (GMM) y modelos temporales basados en Kernel Periodicos Gaussianos (KPG) o Transformada Wavelet (TW) según alguna métrica *ad-hoc* al problema propuesto.

## Objetivos y alcances

El objetivo del trabajo de tesis es diseñar e implementar un modelo probabilístico espacio-temporal basado en delitos geo-refenciados ocurridos dentro de los últimos 6 años a partir del 2016. Usando la información espacial es posible aproximar la distribución de riesgo criminal basándose en el entrenamiento de GMM, para luego realizar un análisis temporal de cada uno de los componentes que definen dicha GMM. Este trabajo tiene como objetivo principal demostrar la efectividad de dos modelos matemáticos basados en técnicas de descomposición por medio de Transformada Wavelet y construcción de modelos periódicos usando KPG para el proceso de *forecasting* de señales temporales asociadas a cada componente de la GMM. El desempeño de la parte temporal se mide en base a *Mean Square Error* (MSE), mientras que el desempeño del modelo que integra la parte espacio-temporal se basa en la mezcla de tres conceptos definidos por *High Risk Porcentage* ( $HRP_{\theta}$ ), *True Incident Percentage* ( $TIP_{\theta}$ ) y *Area Under Curve* (AUC). Finalmente, el modelo diseñado debe ser capaz de predecir las zonas con mayor tasa de ocurrencia delictual para un determinado intervalo de tiempo, con tal de facilitar la toma de decisiones a nivel policial y así poder prever y disminuir los eventos criminales a corto o largo plazo.

Se consideran los siguientes objetivos específicos para cumplir con el objetivo principal de este trabajo de tesis:

- Encontrar una estrategia para seccionar los datos de manera que se puedan observar características históricas y recientes.
- Pre-procesar los datos con la finalidad de uniformar los intervalos de predicción.
- Implementar un algoritmo que reconozca patrones espaciales de los datos entregados y que sea capaz de discriminar los datos correspondientes a cada patrón.
- Implementar algoritmos que procesen y caractericen la información temporal de los datos mediante KPG y TW.
- Diseñar un modelo matemático que unifique la caracterización espacial y temporal, con el cual sea posible realizar predicciones espacio-temporales del riesgo criminal.
- Integrar las métricas de desempeño para el modelo construido.
- Realizar diferentes estudios de optimización del algoritmo en algunos de los parámetros que definen a los modelos temporales diseñados.
- Comparar los desempeños de cada uno de los modelos temporales generados mediante el uso de la métrica AUC para el conjunto de parámetros que optimizan su respuesta.

## Estructura de tesis

El Capítulo 1 tiene comienzo con un breve Estado del Arte del problema de caracterización del riesgo criminal a nivel mundial y finaliza con el marco teórico explicando cada una de las herramientas ocupadas para el procesamiento de datos, transformaciones por medio de diferentes *kernels*, aplicables tanto en el caso espacial como temporal y descomposición por medio de TW para el análisis de señales temporales. El Capítulo 2 muestra la metodología utilizada para generar el modelo de caracterización de riesgo criminal integrando la etapa espacial y temporal, bajo el análisis de datos ocurridos con anterioridad. El Capítulo 3 entrega los resultados generados por el modelo, mostrando de manera meticulosa cada una de las respuestas generadas en cada etapa, utilizando como caso base la comuna de Maipú para una fecha definida arbitrariamente.

Posteriormente se procede a optimizar diferentes parámetros internos que definen a los modelos temporales, utilizando siempre como caso base la comuna de Maipú y ventanas de predicción abocada al mes de junio con su respectivos datos de validación. Además se entrega un estudio del desempeño del modelo en diferentes comunas de la Región Metropolitana basándose en los estudios de optimización realizadas con anterioridad. Finalmente se entrega un análisis de los resultados obtenidos. El Capítulo 4 entrega las conclusiones del trabajo de tesis.

## Confidencialidad

Este trabajo de título está inserto en el proyecto con nombre Software de Predicción Criminal de la Subsecretaría de Prevención del Delito del Gobierno de Chile y Carabineros de Chile, el cual fue desarrollado por el Centro de Análisis y Modelamiento en Seguridad (CEAMOS) de la Universidad de Chile, dirigido por Sr. Raúl Manásevich. Debido a la presencia de acuerdos de confidencialidad entre las partes, existe información relevante referida a los algoritmos desarrollados y las bases de datos utilizadas que no será entregada detalladamente en este documento. Sin embargo, se deja constancia en este documento que el Profesor Guía está en conocimiento de toda la información no presentada en este documento.

# Capítulo 1

## Estado del Arte y Marco Teórico

### 1.1. Evolución de los modelos de caracterización criminal.

Muchas técnicas y modelos han sido desarrollados durante estas últimas décadas para alcanzar el objetivo principal de este trabajo de tesis. Uno de los más comunes es la teoría definida por *Hot Spots* explicada el año 2005, donde los incidentes criminales son proyectados sobre la superficie de análisis, formando diferentes *clusters* de interés, los cuales se asumen constantes para una ventana de predicción definida [13]. Sin embargo, una de las falencias que se presenta en dicha teoría es la incapacidad de reflejar cambios en los patrones criminales para un intervalo de tiempo futuro, siendo un modelo estático. Para solucionar este problema se han formulado modelos más sofisticados y que contemplan la evolución temporal del crimen en cierta área de interés. Por ejemplo, el trabajo de D. Brown y H. Liu [14] introduce un modelo de predicción multivariable, mezclando los conceptos *Hot Spots* y las características de la zona en donde se está realizando el análisis criminal. En este caso se implementa un modelo de densidad de transición basado en patrones puntuales para la predicción de eventos en el espacio-tiempo, entrenado según las características criminales observadas en el pasado. Otro ejemplo es el trabajo de Y. Xue y D. Brown [3] donde se combinan las teorías de Elección Discreta y *Data Mining* para el desarrollo de un modelo espacial, analizando y prediciendo el comportamiento criminal y su latencia de decisión efectiva, al igual que el trabajo realizado por M. Smith y D. Brown [15] donde se asume que los criminales eligen lugares que minimizan sus esfuerzo al momento de asaltar y que además le entrega una ganancia proporcional o mayor a dicho esfuerzo.

La desventaja de los modelos anteriormente descritos es que ninguno de ellos incorpora el componente temporal directamente al modelo, y en el caso de ser modelado, por ejemplo, usando series temporales, la interacción espacio-tiempo no queda bien definida, tal como se muestra en el trabajo de C. Ivaha, H. Al-Madfai, G. Higgs y J. Ware [16], donde la evolución criminal no depende netamente de un modelo espacio-temporal sino de las condiciones climáticas a lo largo del tiempo que condicionan la actividad delictual en una cierta región. Estudios recientes han permitido desarrollar nuevas mejoras al problema espacio-temporal aplicando *Generalized Additive Models* (GAM) que permiten combinar información espacial y temporal de forma paramétrica, considerando además diversas características de la zona (geografía y demografía, entre otras) para la predicción de nuevos incidentes criminales [17].

En el año 2008, el comisionado William Bratton del departamento de policía en Boston, Estados Unidos, en conjunto con el director de la Oficina de Asistencia Judicial, James H. Burch y de la directora del Instituto Nacional de Justicia, Kristina Rose, exploraron un nuevo concepto de predicción policial; con la ayuda de investigadores, practicantes, oficiales gubernamentales pudieron realizar diferentes simposios consecutivos que daban a conocer las distintas aristas del problema de predicción criminal. Después de dos años de constante trabajo, se pudo implementar uno de los software más nombrados en el área de predicción criminal, PredPol<sup>®</sup>, tecnología que ofrece a los oficiales la mejor oportunidad de estar en el lugar correcto, en el momento adecuado, para detener a los delincuentes antes de que cometan un acto delictual, permitiendo así reducir el crimen y aumentar la precisión de patrullaje de dichos oficiales. [2].

Gran parte de los modelos anteriormente mencionados ocupan como base datos derivados de procesos puntuales espaciales (*Spatial Point Process*). Dichos datos entregan información suficiente de las características de los delitos ocurridos, siendo los principales el lugar geo-referenciado y la fecha y hora ocurrido el evento. Al venir de una distribución de probabilidades es posible definir cada uno de los eventos como una observación dada por una variable aleatoria con distribución a estimar.

## 1.2. Transformaciones por medio de Kernel

Se define una variable aleatoria de dimensión  $L$  como  $X = (X_1, \dots, X_l, \dots, X_L)$  donde  $X_l$  corresponde a la variable  $l$ -ésima. Sea  $x_i = (x_1, \dots, x_l, \dots, x_L)$  una de las  $i$  observaciones entregadas de la variable  $X$ . Dicha variable aleatoria procede de una función de densidad  $f_X$ . Dicha función proviene de una familia de funciones definida por medio de un conjunto de parámetros desconocidos en primera instancia.

Incluso sin el conocimiento de los parámetros, es posible encontrar una o varias funciones de densidad que se ajusten de mejor manera a los datos ya sea por métodos no paramétricos o paramétricos. El primer método permite estimar la distribución de probabilidades sin tener en mente una familia de densidades candidata, tomando en consideración solo los datos entregados. El segundo método permite definir una estimación de los parámetros que optimizan un tipo de familia de funciones de densidades conocida (Gaussianas, *Student's t*, exponenciales, *Gamma*, *Beta*, entre otras) en base a la información entregada por los datos. Uno de los estimadores más recurrentes es *Maximum Likelihood* (ML) dado que resulta ser consistente y eficiente en la mayoría de los casos .

### 1.2.1. Kernel Density Estimation

Una de las primeras herramientas utilizadas para la estimación de la función de densidad  $f_X$  es por medio de histogramas, donde cada una de las muestras es asignada a un intervalo de valores (bin) de ancho  $h$ , siendo cada bin contabilizado respecto a los datos que se contienen en dicho intervalo y normalizado por el total de datos  $n$  observados. Lo que se consigue es estimar la función de distribución acumulativa  $F_X(x)$  definida por la ecuación 1.1.

$$\hat{F}_X(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(-\infty, x]}(x_i) \quad (1.1)$$



Es posible obtener el estimador de la pdf en base a la función de distribución acumulativa, utilizando la definición de la ecuación 1.2, donde  $h = x_{i+1} - x_i$

$$\begin{aligned}\widehat{f}_X(x) &= \lim_{h \rightarrow 0} \frac{\widehat{F}_X(x+h) - \widehat{F}_X(x)}{h} = \lim_{h \rightarrow 0} \frac{\frac{1}{n} (\sum_{i=1}^n \mathbf{1}_{(-\infty, x+h]}(x_i) - \sum_{i=1}^n \mathbf{1}_{(-\infty, x]}(x_i))}{h} \\ &= \lim_{h \rightarrow 0} \frac{\sum_{i=1}^n \mathbf{1}_{(x, x+h]}(x_i)}{nh}\end{aligned}\quad (1.2)$$

El estimador de histograma tiene la ventaja de ser fácil y simple de construir, facilitando la interpretación de los datos, sin la necesidad de recurrir a herramientas gráficas avanzadas [18]. Las propiedades del estimador dependen netamente del valor de  $h$ . Entre más pequeño sea el valor de  $h$ , menor es el sesgo del estimador pero mayor es su varianza [19]. Por otra parte, el estimador presenta serios problemas de continuidad por no ser lo suficientemente sensible a las propiedades locales de  $\widehat{f}_X$ . La ecuación 1.2 puede ser reescrita por:

$$\widehat{f}_X(x) = \lim_{h \rightarrow 0} \frac{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)}{nh}\quad (1.3)$$

Donde  $K(u) = \frac{1}{2}$  si  $u \in ]-1,1]$  y 0 en otro lugar. El parámetro  $h$  es el ancho de banda del kernel elegido.

La forma entregada en la ecuación 1.3 es denominada por *Kernel Density Estimation* (KDE) siendo los histogramas un caso particular, donde el kernel utilizado es del tipo uniforme. Gracias a esta generalización y el uso de diversos tipos de kernels es posible solucionar el problema de continuidad y sensibilidad locales, debido a que la suma ponderada de un tipo de kernel  $K$  retiene sus propiedades de continuidad y diferenciabilidad. En el siguiente apartado se entregan algunos de los kernels aplicados en la actualidad en el método KDE [20]. Dichas funciones se definen en base a una aproximación dado que la ponderación puede variar de un problema a otro.

- Gaussiano:  $K(x|h) \propto \exp\left(\frac{-x^2}{2h^2}\right)$
- Uniforme:  $K(x|h) \propto 1$ ; si  $x < h$
- Epanechnikov:  $K(x|h) \propto 1 - \frac{x^2}{h^2}$
- Exponencial:  $K(x|h) \propto \exp\left(\frac{-x}{h}\right)$
- Lineal:  $K(x|h) \propto 1 - \frac{x}{h}$ ; si  $x < h$
- Coseno:  $K(x|h) \propto \cos\left(\frac{\pi x}{2h}\right)$ ; si  $x < h$

### 1.2.2. Mixture Model

Sea un conjunto de  $M$  funciones de una misma familia de distribución de probabilidades  $f_1, f_2, \dots, f_M$ . Se dirá que  $f$  es un modelo mixto de  $M$  componentes si cumple con la forma dada por la ecuación 1.4, donde  $\lambda_m$  es el conjunto de pesos de la mezcla, restringido a que  $\lambda_m > 0$  y  $\sum_M \lambda_m = 1$

$$f(x_i) = \sum_{m=1}^M \lambda_m f_m(x_i) \quad (1.4)$$

Cada una de las funciones entregadas en el conjunto se diferencia por el valor asociado a los parámetros que la definen. Por ejemplo las funciones de la familia de las Gaussianas tendrán diferentes medias y covarianzas, las de Poisson diferentes medias y las potencias tendrán distintos exponentes. Por esta razón es necesario explicitar el vector de parámetro  $\theta_m$  que define a la función  $m$ -ésima, con el cual la ecuación 1.4 pasa a ser reescrita por 1.5.

$$f(x_i|\Theta) = \sum_{m=1}^M \lambda_m f(x_i|\theta_m) \quad (1.5)$$

Todos los parámetros que definen al modelo mixto quedan condensados en  $\Theta = (\lambda_1, \dots, \lambda_K, \theta_1, \dots, \theta_K)$ . Analizando dos extremos se tiene que [21]:

- Si  $M = 1$  entonces tendremos tan solo una distribución en el modelo, recayendo todo el proceso de cálculo en la estimación de sus parámetros por medio de un estimador ML.
- Si  $M = n$  entonces se vuelve nuevamente al caso de KDE, donde el cálculo de los parámetros queda delegado a segundo plano.

La cantidad de funciones con las cuales se trabaja en el modelo mixto tendrá una fuerte influencia en el *trade off* entre lo preciso y acertado de la respuesta del modelo.

### 1.2.2.1. Expectation Maximization

Asumiendo que las observaciones entregadas son independientes, con una función de probabilidad definida, se puede definir la función de verosimilitud del modelo mixto según la ecuación 1.6.

$$L(\Theta, X) = \prod_{i=1}^n f(x_i|\Theta) \quad (1.6)$$

Si se aplica el logaritmo a la función de verosimilitud es posible descomponer la pitatoria en una sumatoria expresada por:

$$l(\Theta, X) = \log(L(\Theta, X)) \quad (1.7)$$

$$= \sum_{i=1}^n \log(f(x_i|\Theta)) \quad (1.8)$$

$$= \sum_{i=1}^n \log\left(\sum_{m=1}^M \lambda_m f(x_i|\theta_m)\right) \quad (1.9)$$

El proceso de minimización del error asociado a la función de verosimilitud del modelo mixto recae resolver la ecuación para el conjunto  $\Theta$ :

$$\frac{\partial l(\Theta, X)}{\partial \Theta} = 0 \quad (1.10)$$

Es posible obtener una respuesta cerrada y bien definida para distribuciones de probabilidad de forma Gaussiana unidimensional, sin embargo existen casos más complejos y de mayor dimensionalidad, las cuales se debe aplicar el algoritmo *Expectation Maximization* (EM).

Se formula un conjunto de datos completos  $Z = (X, Y)$ , donde  $X$  es el vector de datos observados e  $Y$  esta asociado al vector de datos no observados. En este marco de referencia, donde cada  $x_i$  de dimensión  $n$ , es un resultado generado por una de las componentes del modelo mixto ha ajustar y  $y_{im}$  de dimensión  $M$ , toma valores 0 o 1 dependiendo de si el vector  $x_i$  pertenece o no a un componente  $m$  del modelo mixto ha ajustar, el cual se asume que distribuye con una función multinomial. Bajo este supuesto es posible generar una función de verosimilitud, definida para un conjunto de datos completos, tal como se muestra en la ecuación 1.13

$$L_c(\Theta, Z) = \prod_{i=1}^n f(x_i, y_{im} | \Theta) \quad (1.11)$$

$$= \prod_{i=1}^n (f(y_{im} | x_i, \Theta) \cdot f(x_i | \Theta)) \quad (1.12)$$

$$(1.13)$$

Aplicando logaritmo a la ecuación 1.13, es posible obtener:

$$l_c(\Theta, Z) = \sum_{i=1}^n \sum_{m=1}^M y_{im} \{ \log \lambda_m + \log f_m(x_i | \theta_m) \} \quad (1.14)$$

El algoritmo EM es un proceso iterativo que consta de dos etapas:

- 1.- *Expectation* (E): Esta etapa fue necesario calcular el valor de la esperanza condicionada de el logaritmo de  $L_c$  dado un dato observado  $x_i$ , suponiendo un conjunto de valores iniciales asociados a los parámetros  $\Theta^{(0)}$ . Dicha esperanza condicionada queda reflejada en la siguiente ecuación:

$$Q(\Theta, \Theta^{(0)}) = E_{\Theta^{(0)}} \{ \log L_c(\Theta) | x_i \} \quad (1.15)$$

Como el logaritmo de la función verosimilitud es lineal respecto a los datos no observables  $y_{im}$ , la iteración  $k + 1$  de la etapa E simplemente requiere del calculo de la actual esperanza condicionada de la variable  $Y_{im}$  dado un dato observado  $x_i$ . En este caso se tiene:

$$E_{\Theta^{(k)}} \{ Y_{im} | x_i \} = \text{pr}_{\Theta^{(k)}} \{ Y_{im} = 1 | x_i \} \quad (1.16)$$

$$= \tau_m(x_i | \Theta^{(k)}) \quad (1.17)$$

Donde  $p$  definió la probabilidad del conjunto evaluado y  $\tau_m$  fue definida como la probabilidad *posteriori* para el  $m$ -ésimo componente del modelo dado un dato observado y su definición quedo reflejada en la ecuación 1.18.

$$\tau_m(x_i|\Theta^{(k)}) = \frac{\lambda_m^{(k)} f(x_i|\Theta^{(k)})}{f(x_i|\theta_m^{(k)})} = \frac{\lambda_m^{(k)} f(x_i|\Theta^{(k)})}{\sum_{m=1}^M \lambda_m^{(k)} f(x_i|\theta_m^{(k)})} \quad (1.18)$$

Finalmente la esperanza condicionada de la etapa E queda reflejada en la ecuación 1.19

$$Q(\Theta, \Theta^{(k)}) = \sum_{i=1}^n \sum_{m=1}^M \tau_m(x_i|\Theta^{(k)}) \{ \log \lambda_m^{(k)} + \log f(x_i|\theta_m^{(k)}) \} \quad (1.19)$$

2.- *Maximization* (M): Esta etapa maximiza la información entregada por 1.19 para generar una nueva iteración respecto a los parámetros definidos en  $\Theta$ . Para el caso dado por un set de datos completos, el resultado de la maximización de los pesos  $\lambda_m$  de cada función del modelo mixto y sus parámetros internos  $\theta_m$  queda dado por las ecuaciones 1.20 y 1.21 respectivamente.

$$\lambda_m^{(k+1)} = \sum_{i=1}^n \tau_m(x_i|\Theta^{(k)})/n \quad (1.20)$$

$$\theta^{(k+1)}_m := \sum_{i=1}^n \sum_{m=1}^M \tau_m(x_i|\Theta^{(k)}) \frac{\partial \log f_m(x_i|\theta_m^{(k)})}{\partial \theta^{(k)}_m} = 0 \quad (1.21)$$

Los etapas E y M son alternadas repetidamente hasta que la diferencia entre  $L_c(\Theta^{(k+1)}) - L_c(\Theta^{(k)})$  sea menor o igual a un valor arbitrariamente pequeño [22].

Como el logaritmo de la función verosimilitud es lineal con respecto a los datos no observables  $y_{im}$ , la iteracion  $k + 1$  de la etapa E necesitara tan solo el valor condicional de  $Y$  dado un dato  $x_i$  observado, siendo en este caso igual a:

$$E_{\Theta^{(k)}}(Y_{im}|x_i) = p_{\Theta^{(k)}}(Y_{im} = 1|x_i) \quad (1.22)$$

$$= \tau_i(x_i|\Theta^{(k)}) \quad (1.23)$$

### 1.2.2.2. Gaussian Mixture Model

Se define como función de densidad de la familia Gaussiana la descrita en la ecuación 1.24.

$$f(x|\mu_m, \Sigma_m) = \frac{1}{(2\pi)^{\frac{L}{2}} |\Sigma_m|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - \vec{\mu}_m)^T \Sigma_m^{-1} (x - \vec{\mu}_m)\right) \quad (1.24)$$

Donde  $L$  es la dimensión de la variable aleatoria,  $\vec{\mu}_m$  es el vector media y  $\Sigma_m$  la matriz de covarianza [23]. Usando la ecuación 1.24 en 1.5 es posible definir un GMM.

El algoritmo EM permite calcular los parámetros establecidos en el GMM mediante las funciones de máxima verosimilitud. Para una cantidad de  $n$  observaciones de la variable aleatoria  $X$  descritas por  $x_i = \{x_1, x_2, \dots, x_L\}$ , con  $i = 1, \dots, n$ , se obtienen las siguientes funciones de verosimilitud:

- Pesos:

$$\lambda_m^{(k+1)} = \frac{\sum_{i=1}^n \tau_m(x_i | \Theta^{(k)})}{n} \quad (1.25)$$

- Medias:

$$\mu_m^{(k+1)} = \frac{\sum_{i=1}^n \tau_m(x_i | \Theta^{(k)}) x_i}{\sum_{i=1}^n \tau_m(x_i | \Theta^{(k)})} \quad (1.26)$$

- Varianza (diagonal de la covarianza):

$$\sigma_m^{2(k+1)} = \frac{\sum_{i=1}^n \tau_m(x_i | \Theta^{(k)}) x_i^2}{\sum_{i=1}^n \tau_m(x_i | \Theta^{(k)})} - \mu_m^{(k+1)} \quad (1.27)$$

El desarrollo de las ecuaciones fue obtenida según [23], pero fueron ajustadas al desarrollo de este Trabajo de Tesis, usando la anotación dada por [22].

### 1.2.3. Kernel Periódico Gaussiano

Sea una función  $u : x \rightarrow u(x)$  que permite generar un respuesta estacionaria para todo el dominio uni-dimensional de  $x$ . Un caso particular para la función  $u$  esta dada por la familia de funciones sinusoidales, dadas por:

$$u(x) = \left\{ \cos \left( 2\pi \frac{x}{p} \right); \sin \left( 2\pi \frac{x}{p} \right) \right\} \quad (1.28)$$

Tomando en consideración estas funciones base, es posible construir una transformación que considere un kernel exponencial cuadrático en el espacio generado por  $u$  tal como se muestra en la ecuación 1.29 [24] [25] [26].

$$K_{pT}(x_i | c_t, \beta, l, p) = \beta^2 \exp \left( -\frac{1}{2l^2} \sin^2 \left( \pi \left| \frac{x_i - c_t}{p} \right| \right) \right) \quad (1.29)$$

La amplitud de la expresión matemática de la función KPG ( $K_{pT}$ ) esta siendo regulada por el parámetro  $\sigma$ ,  $l$  es el ancho de banda asociada a  $K_{pT}$  y  $c_t$  es el centro o punto de inicio de  $K_{pT}$ . Con esta formulación es posible realizar inferencias arbitrarias con diferentes tipos de anchos y periodos dados por el parámetro  $p$ . Es importante notar que la escala  $l$  es relativa al periodo  $p$  utilizado. La Figura 1.1 entrega la influencia de cada uno de estos parámetros con un arreglo  $x$  de 1000 muestras entre  $[0 - 50]$ .

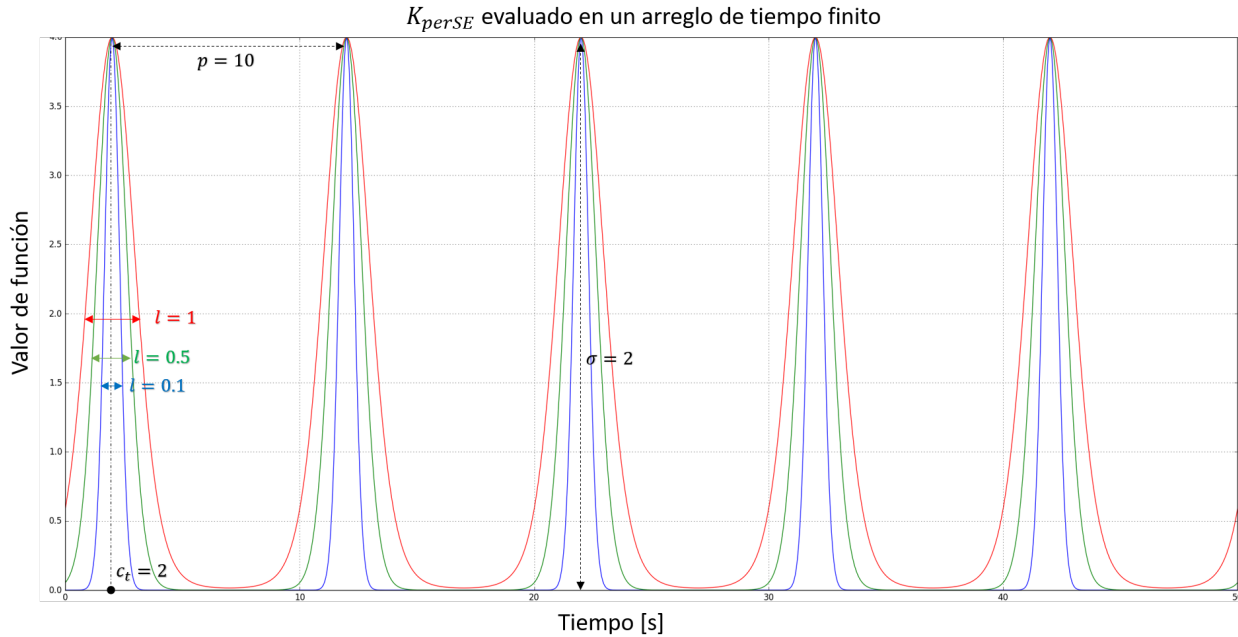


Figura 1.1: Influencia de los parámetros asociados al KPG. Los gráficos entregados son ejecutado con parámetros  $c_t = 2$ ,  $\sigma = 2$ ,  $p = 10$  y diferentes valores de  $l$  explicitados en cada curva coloreada.

### 1.3. Análisis Wavelet

El análisis de series temporales que contienen comportamientos no estacionarios usando herramientas de Transformada Wavelet (TW) se ha vuelto cada vez más común en la literatura científica. Para descomponer una serie de tiempo en el espacio de frecuencias, es necesario determinar el comportamiento de la señal, denotando la cantidad de sucesos no estacionarios que ocurren en una periodo fijo de análisis. En esta sección se entregan algunos detalles importantes a considerar para el estudio de una señal por medio de Wavelet, basando su notación matemática con el trabajo de Christopher Torrence y Gilbert P. Compo [27]. Para una mejor introducción al tema de Transformada Wavelet se sugiere consultar el trabajo de M. Lua y Y. Weng [28].

#### 1.3.1. Transformada Fourier

La Transformada de Fourier (TF) representa una de las herramientas más básicas para extraer información frecuencial de una señal temporal. La TF es aplicable a toda señal temporal con tiempo de muestreo  $\delta t$  y duración igual a  $N\delta t$  sobre una ventana deslizante de largo  $T < N\delta t$ , retornando un estudio de frecuencias entre  $T^{-1}$  y  $(2\delta t)^{-1}$ . Las ventanas deslizantes pueden ser moldeados mediante funciones como cajas cerradas (cortes abruptos tanto al inicio como al final) o por *Gaussian Windows* (segmentos suaves tanto al principio como al final).

Según Kaiser [29], la TF representa un método impreciso e ineficiente para el análisis frecuencial, ya que impone una escala o intervalo de ventana fijo igual a  $T$ . Lo impreciso surge a partir del *aliasing* de las componentes con frecuencias altas o bajas que no caen dentro del

rango establecido por el segmento. La ineficiencia recae en frecuencias  $\frac{T}{2\delta t}$ , que deben ser analizadas en cada tiempo de muestreo, independiente del tamaño del segmento entregado o las frecuencias dominantes que se presenten. Adicionalmente, se deben usar diferentes largos de ventana  $T$  con tal de entregar un análisis más apropiado de la señal temporal con diferentes frecuencias dominantes, dando a conocer una necesidad de independizar el largo de la ventana. Esta independencia es posible haciendo uso de la Transformada Wavelet.

### 1.3.2. Transformada Wavelet

La Transformada Wavelet (TW) permite analizar series temporales con comportamiento no estacionario en diferentes frecuencias. Sea una serie de tiempo  $x_n$  con un muestreo temporal de  $\delta t$  y  $n = 0, \dots, N - 1$ , con  $N$  la cantidad de puntos de la señal temporal. Se asume una función Wavelet ( $\psi_0(\eta)$ ) que depende de un parámetro no dimensional de “tiempo”. Es de esperar que  $\psi_0(\eta)$  tenga media igual a cero y que además posibilite su mapeo tanto en el espacio de tiempo y frecuencia.

El término  $\psi_0(\eta)$  es usado generalmente para hacer referencia a funciones ortogonales o no ortogonales, mientras que el término base Wavelet es aplicable solamente a un set de funciones ortogonales. Cada  $\psi_0(\eta)$  puede ser ocupada tanto para TW discretas como continuas, mientras que las bases son utilizadas solo para TW discretas. En este caso se da a conocer el desarrollo de la TW continua con series temporales discretas.

Se define la Transformada Wavelet continua de una secuencia temporal  $x_n$  discreta ( $W_n$ ) como la convolución de la señal con una versión escalada y trasladada de la función Wavelet normalizada ( $\psi(\eta)$ ), tal como se muestra en la ecuación 1.30.

$$W_n(s) = \sum_{n'=0}^{N-1} x_{n'} \psi^* \left[ \frac{(n' - n)\delta t}{s} \right] \quad (1.30)$$

Donde:

- \* es el complejo conjugado de  $\psi(\eta)$ .
- $s$  es el parámetro interno de la función Wavelet que indica la escala temporal asignada de  $\psi(\eta)$ .
- $n$  el parámetro interno de la función Wavelet que indica la traslación temporal con respecto al origen.

Con la ecuación 1.30 es posible identificar diferentes características temporales por medio de una escala fija ( $s$  fijo). Es importante notar que  $\psi_0(\eta)$  han sido normalizadas con tal de obtener  $\psi(\eta)$  (ver 1.3.2.1 Normalización). Aunque es posible calcular la TW con la ecuación 1.30, es considerablemente más rápido realizar el cálculo por medio del espacio de Fourier.

La convolución que define a TW necesita ser procesada  $N$  veces por cada escala, siendo  $N$  la cantidad de puntos que definen la señal temporal discreta. El teorema de convolución permite realizar  $N$  convoluciones simultáneamente en el espacio de Fourier, aplicando la Transformada Discreta de Fourier (DFT), la cual se define en la ecuación 1.31.

$$\hat{x}_k = \frac{1}{N} \sum_{n=0}^{N-1} x_n e^{\frac{-2\pi kn}{N}i} \quad (1.31)$$

En este caso  $k = 0, \dots, N - 1$  es el índice de frecuencia. La TF de  $\psi(\eta)$  esta dada por  $\hat{\psi}(s\omega)$ , con la cual se genera la respuesta  $Wn$  utilizando los beneficios de cálculo de la TF y de su inversa, tal como queda reflejado en la ecuación 1.32.

$$W_n(s) = \sum_{k=0}^{N-1} \hat{x}_k \hat{\psi}^*(s\omega_k) e^{i\omega_k n \delta t} \quad (1.32)$$

Donde  $\omega_k$  esta definida por:

$$f(x) = \begin{cases} \frac{2\pi k}{N\delta t} & \text{si } k \leq \frac{N}{2} \\ -\frac{2\pi k}{N\delta t} & \text{si } k > \frac{N}{2} \end{cases} \quad (1.33)$$

Usando la ecuación 1.32 y la rutina estándar de TF es posible calcular la TW (para una escala  $s$  arbitraria) para todos los  $n$  de la señal de manera simultanea y eficiente.

### 1.3.2.1. Normalización

Para asegurar que la TW sea comparable con cada una de las escalas  $s$  elegidas, es necesario que cada función Wavelet ( $\psi_0(\eta)$ ) definida en cada  $s$  sea normalizada con energía unitaria. Se define a  $\hat{\psi}_0(s\omega_k)$  como la TF de  $\psi_0(\eta)$  definida para  $s$ . Se tiene que su normalización esta dada por:

$$\hat{\psi}(s\omega_k) = \left( \frac{2\pi s}{\delta t} \right)^2 \hat{\psi}_0(s\omega_k) \quad (1.34)$$

Usando esta normalización y una escala  $s$  definida se tiene  $\sum_{k=0}^{N-1} \hat{\psi}(s\omega_k) = 1$ , donde  $N$  es la cantidad de puntos de la serie temporal.

### 1.3.3. Funciones Wavelet

En la Tabla 1.1 se dan a conocer algunas  $\psi_0(\eta)$  más usadas, descritas tanto para el espacio  $\eta$  como en el de Fourier, donde la figura 1.2 muestra gráficamente su comportamiento en cada uno de los espacios. Es importante notar que gran parte de las funciones tienen recorrido en los números complejos mientras que otras recorren solo el espacio real.

El primer paso para realizar un correcto análisis por medio TW es elegir de manera sensata la función Wavelet, la cual se condiciona con la serie temporal a procesar. Algunos de los aspectos que se deben tener en consideración son los siguientes (mayor detalle en el trabajo de Farge [30]):

- **Ortogonalidad:** El análisis Wavelet ortogonal entrega una cantidad de convoluciones proporcional al ancho de la escala de la base Wavelet. Esto permite compactar de manera simple la información de la señal, siendo una herramienta útil en el procesamiento de señales. Sin embargo, un cambio a periódico de la señal no puede ser representado por medio de este tipo de análisis. Contrario al caso anterior, un análisis no ortogonal entrega información redundante dado que la cantidad de convoluciones es fija para toda escala de la función Wavelet. Se espera que la señal a procesar para este análisis no ortogonal sea suave, con variaciones continuas en su amplitud y de tamaño limitado.



Nombre	$\psi_0(\eta)$	$\widehat{\psi}_0(s\omega)$	$e\text{-folding}$
Morlet ( $\omega_o = \text{frecuencia}$ )	$\pi^{-1/4} e^{i\omega_0\eta} e^{-\eta^2/2}$	$\pi^{-1/4} H(\omega) e^{-(s\omega - \omega_0)^2/2}$	$\sqrt{2}s$
Paul ( $m = \text{orden}$ )	$\frac{(2i)^m m!}{\sqrt{\pi(2m)!}} (1 - i\eta)^{m+1}$	$\frac{2^m}{\sqrt{\pi(2m-1)!}} H(\omega) (s\omega)^m e^{-s\omega}$	$s/\sqrt{2}$
DOG ( $m = \text{derivada}$ )	$\frac{(-1)^{m+1}}{\Gamma(m+\frac{1}{2})} \frac{d^m}{d\eta^m} (e^{-\eta^2/2})$	$\frac{(-i)^m}{\Gamma(m+\frac{1}{2})} (s\omega)^m (e^{-\eta^2/2})$	$\sqrt{2}s$

Tabla 1.1: Tres funciones Wavelet expresadas en su forma temporal y frecuencial de Fourier. La función DOG es la abreviación de *Derivative Of a Gaussian*.  $H(\omega)$  es la función Heaviside  $H(\omega) = 1$ , si  $\omega > 0$  y es 0 para cualquier otro valor.

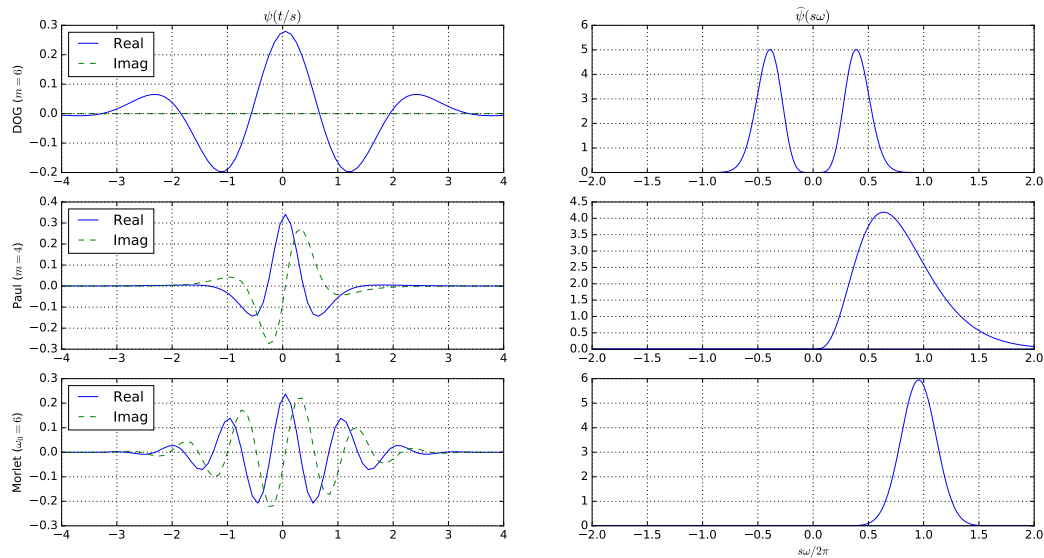


Figura 1.2: Tres diferentes funciones Wavelet generadas a partir de la tabla 1.1. Los gráficos de la izquierda entregan el comportamiento temporal de las funciones, siendo la línea sólida la parte real y la línea punteada la parte imaginaria. Los gráficos de la derecha entrega el comportamiento en el dominio de la frecuencia. En este caso se tomo en consideración una escala igual a  $s = 10\delta t$ .

- Recorrido: Funciones con recorrido en los complejos permiten adaptar de mejor manera el comportamiento oscilatorio de la serie, mientras que las funciones con recorrido solamente en los reales se acomodan de mejor manera a los *peaks* aislados o discontinuidades.
- Anchura: La anchura de la función Wavelet esta condicionado con el valor que toma su escala  $s$  y queda definida por el nombre de *e-folding* entregado en la Tabla 1.1. Este parámetro refleja el equilibrio que existe entre el espacio temporal y de Fourier. Una función corta (en el tiempo) tiene una mejor resolución temporal pero pobre en el espacio frecuencial, mientras que una función ancha tiene una pobre resolución temporal pero mejor en el espacio frecuencial.

Las funciones Wavelet entregadas en la Tabla 1.1 son todas no ortogonales. Sin embargo,

existen otros tipos de funciones tales como Haar y Daubechies con las cuales se realiza análisis Wavelet ortogonal. Más ejemplos de bases o funciones Wavelet se puede consultar el trabajo de G. Kaiser [29].

### 1.3.4. Espectro de energía Wavelet

Dado que en general  $\psi(\eta)$  es complejo se tendrá que  $W_n$  también lo es. Dicha transformada puede ser dividida en parte real  $\mathcal{R}\{W_n(s)\}$  e imaginaria  $\mathcal{I}\{W_n(s)\}$  o en amplitud  $|W_n(s)|$  y fase  $\angle W_n(s)$ . Es posible definir el Espectro de Energía de Transformada Wavelet (EETW) como  $|W_n(s)|^2$ . Para el caso de funciones Wavelet reales, la parte imaginaria es nula y su fase es indeterminada.

### 1.3.5. Elección de escalas

Una vez elegida la función Wavelet a utilizar, es necesario establecer el conjunto de escalas  $s$  de Transformada Wavelet ( $S = \{s_j\}_{j=0,\dots,J}$ ) a usar en la ecuación 1.32. Para el caso de análisis ortogonal, el set discreto de escalas esta dado por Farge [30]. Para el caso no ortogonal, el set de escalas puede ser arbitrario, tanto como la que uno desee, con tal de obtener una imagen mucho más completa de la señal por medio del EETW. Una de las convenciones más utilizadas es generar un set de escalas por medio de fracciones de la potencia de dos. Las ecuaciones 1.35 y 1.36 toman en consideración esta convención, donde la escala minima de resolución ( $s_0$ ) permite determinar el largo del arreglo de escalas ( $J$ ).

$$s_j = s_0 2^{j\delta j}, \quad j = 0, \dots, J \quad (1.35)$$

$$J = \delta j^{-1} \log_2 \left( \frac{N\delta t}{s_0} \right) \quad (1.36)$$

El parámetro  $s_0$  suele tomar como valor mínimo  $2\delta t$  donde  $\delta t$  es el tiempo de muestreo de la serie temporal.

Finalmente el parámetro  $\delta j$  define la resolución del arreglo de escalas aplicado. Entre menor sea su valor, más fino es la resolución de las escalas, pero mayor es el tiempo de cálculo de TW.

#### 1.3.5.1. Relación entre escalas y periodos de Fourier

La relación entre el periodo equivalente en Fourier y la escala  $s$  se deriva de manera analítica para todo tipo de función Wavelet [31]. En particular, la tabla 1.2 entrega los coeficientes de ponderación para las funciones hasta ahora discutidas en el análisis de Wavelet.

Nombre función Wavelet	Periodo de Fourier según la escala $s$
Morlet ( $\omega_0 = \text{frecuencia}$ )	$\frac{4\pi}{\omega_0 + \sqrt{2 + \omega_0^2}} \cdot s$
Paul ( $m = \text{orden}$ )	$\frac{4\pi}{2m+1} \cdot s$
DOG ( $m = \text{derivada}$ )	$\frac{2\pi}{\sqrt{m+1/2}} \cdot s$

Tabla 1.2: Relación entre la escala  $s$  y el periodo bajo el concepto de Fourier para las funciones Wavelet Morlet, Paul y DOG. Tanto  $\omega_0$  como  $m$  son parámetros de la función.

### 1.3.6. Anti-transformada Wavelet (ATW)

Dado que la TF es un filtro pasa bandas con función conocida, es posible realizar una reconstrucción al espacio temporal de una señal transformada aplicando el concepto de deconvolución o el de filtro inverso aplicado. La reconstrucción es sencilla para el caso de funciones ortogonales dado que no generan redundancia en el tiempo gracias al arreglo de escalas, contrariamente al caso no ortogonal. Aun con esta redundancia, es posible calcular la Anti-Transformada Wavelet (ATW) usando la reconstrucción de la señal por medio de la ecuación 1.37, siendo en este caso la suma de la parte real de la TW para todas las escalas.

$$x_n = \frac{\delta_j \delta t^{1/2}}{C_\delta \psi_0(0)} \sum_{j=0}^J \frac{\mathcal{R}\{W_n(s_j)\}}{s_j^{1/2}} \quad (1.37)$$

El factor  $\psi_0(0)$  quita la energía escalada, mientras que  $s_j^{1/2}$  convierte la TW a una densidad de energía. El factor  $C_\delta$  se calcula a partir de la TW de una señal  $\delta$  (Delta Dirac) usando la función Wavelet  $\psi_0(\eta)$ , siendo un valor constante y diferenciado para cada una de las funciones Wavelet. El proceso de reconstrucción ha sido simplificado y se entrega mayor detalles en el trabajo de Farge [30].

## 1.4. Medidas de desempeño

Dado un conjunto de datos de validación que proceden de un proceso puntual de carácter espacial y un modelo espacio-temporal cuya salida genera una superficie de riesgo, basada al cálculo de una integral bidimensional, cuyo valor es proporcional a la tendencia de ocurrir un delito para un tipo  $t'$  futuro, es posible medir su desempeño predictivo en base a las curvas definidas por X. Wang y D. Brown con el nombre de *High Risk Porcentage* ( $\text{HRP}_\theta$ ) y *True Incident Percentage* ( $\text{TIP}_\theta$ ) [17]. De su trabajo se extrae la esencia de las formulas que definen cada una de las curvas y se adaptan para el caso de estudio de este trabajo de tesis. Sea un conjunto de eventos delictuales  $\mathcal{D}$  para un tiempo  $t'$  y un conjunto de  $N$  grillas  $G = \{g^h\}$ , con  $h = 1 \dots N$ , que definen el área específica a estudiar. Si el riesgo normalizado asociado a cada grilla  $g^h$  es igual a  $R^{\text{pred}}(g^h, t')$  es posible describir la función  $\text{HRP}_\theta$  y  $\text{TIP}_\theta$  como:

$$\text{HRP}_\theta = \frac{\|g_i | \{R^{\text{pred}}(g^h, t') > \theta\}\|}{\|G\|} = \frac{\|G_\theta\|}{\|G\|} \quad (1.38)$$

$$\text{TIP}_\theta = \frac{\|\mathcal{D} \otimes G_\theta\|}{\|\mathcal{D}\|} \quad (1.39)$$

Donde  $\|\bullet\|$  es la cardinalidad del conjunto,  $\theta \in [0, 1]$  es el umbral que indica el valor de activación de las grillas y  $\otimes$  es un operador espacial que indica los delitos que caen dentro del área activada.

En este caso el  $\text{HRP}_\theta$  entrega el porcentaje de grillas que son activadas para todo el intervalo de evaluación  $\theta$  y el  $\text{TIP}_\theta$  representa la cantidad de delitos que caen efectivamente en las grillas activadas por el valor  $\theta$ . Al compartir el mismo  $\theta$  dichas funciones pueden ser relacionadas mediante una curva similar a las generadas en procesos de caracterización de operaciones (curvas ROC) [32]. En el eje  $x$  se pone el  $\text{HRP}_\theta$  y en el eje  $y$  el  $\text{TIP}_\theta$ . Si la mayoría de los delitos ocurren en las grillas de mayor riesgo criminal entonces la curva se acercará más rápidamente a la esquina superior izquierda, siendo este el mejor resultado a obtener. En caso contrario, la curva tendrá un comportamiento similar a una función lineal.

Finalmente para medir la calidad del modelo en base a las curvas ROC se toma en consideración el *Area Under Curve* (AUC), el cual, en este caso en particular, sus valores van desde 0 a 1, siendo el peor y mejor desempeño entregado por el modelo respectivamente.

# Capítulo 2

## Metodología

### 2.1. Introducción

En este trabajo de tesis la caracterización del riesgo criminal se dividió en cuatro secciones. La primera sección entrega información de la base de datos a utilizar, la cual fue separada y filtrada según una consulta específica. La segunda está relacionada en la identificación de patrones espaciales, haciendo uso de EM para generar una GMM espacial bajo ciertas condiciones iniciales. Una vez entrenados y asignados los delitos ocurridos a cada componente que define a la GMM, se utilizaron diferentes técnicas para estimar el comportamiento temporal en cada componente de la GMM, transformando el dominio temporal por medio de transformaciones por KPG o de TW. Finalmente se entrega en detalle el proceso de validación del modelo usando de las métricas descritas en la Sección 1.4. La Figura 2.1 muestra el esquema general de este Capítulo, mostrando cada una de las etapas en bloques separados.

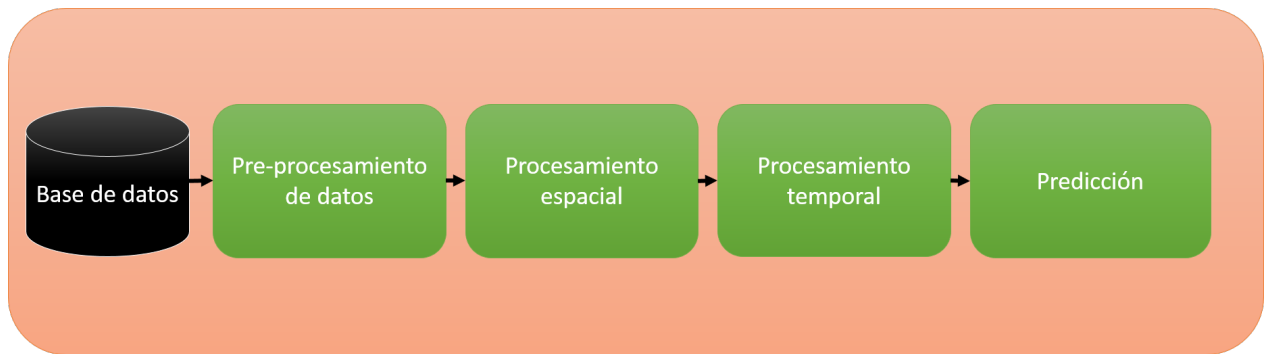


Figura 2.1: Esquema general de la metodología implementada para la caracterización del riesgo criminal.

## 2.2. Base de datos

El proyecto fue realizado gracias a Carabineros de Chile, en conjunto del Centro de Análisis y Modelamiento en Seguridad (CEAMOS), los cuales facilitaron parte de su base de datos para el desarrollo de la investigación. La base de datos contenía tanto información geográfica de las comunas de la Región Metropolitana y de eventos criminales ocurridos desde el 1 enero del 2011 hasta el 16 de agosto del 2016. La comunicación con la base de datos fue por medio de comandos PostgreSQL, los cuales filtraron los datos sujetos a diferentes características asociados a cada uno de ellos. Este proceso fue iterado en conjunto con el equipo técnico de CEAMOS, con los cuales se trabajó durante todo el año. Con tal de entender mejor la forma o estructura de los datos esta sección fue dividida en dos subsecciones: Datos geográficos y datos de incidentes criminales.

### 2.2.1. Datos geográficos

La base de datos contenía dos límites geográficos que permitieron tanto la visualización como evaluación del algoritmo de predicción en cada una de las comunas de la región Metropolitana. El primer límite de interés fue el borde asociado a la comuna, el cual ayuda en la visualización del modelo mediante software que usan sistemas de información geográfica (*Geographic Information System* GIS), siendo QGIS<sup>®</sup> el más utilizado en este trabajo de tesis. El segundo límite fue denominado con el nombre de grilla o cuadrícula siendo la unidad mínima de medida utilizada por Carabineros de Chile. Estas grillas fueron definidas por polígonos cuadrados de 22500 [ $m^2$ ] (de 150 metros cada lado) repartidos uniformemente para toda comuna a lo largo del país, donde la cantidad de grillas en una comuna esta totalmente proporcional con su área total geográfica.

Las comunas estudiadas fueron agrupadas en tres grupos diferentes. El primer grupo correspondió a las comunas predominantemente urbanas, con nada o muy pocas zonas rurales. Ejemplos de estas comunas fueron Conchalí, San Joaquín y Estación Central. Se caracterizaron por poseer superficie geográfica inferior a 60 [ $km^2$ ], siendo el conjunto de comunas con menor área estudiada y por ello con la menor cantidad de grillas de evaluación. El segundo grupo fueron las que tuvieron una proporción muy parecida entre la zona urbana y rural. Ejemplos de estas comunas fueron La Florida, Las Condes y San Bernardo. Se caracterizaron por poseer superficie geográfica entre 60 a 180 [ $km^2$ ]. En este conjunto la cantidad de grillas fluctuó entre 2700 a 8000 unidades. El último grupo correspondió a las comunas con gran presencia de zona rural, ubicadas fuera del Gran Santiago o en los bordes de la región Metropolitana. Fueron comunas tales como Lo Barnechea, Lampa o San José de Maipo. Su superficie geográfica fue superior a los 180 [ $km^2$ ] y fueron las comunas con mayor cantidad de grillas dentro de la base de datos.

### 2.2.2. Datos de eventos criminales

Fueron eventos puntuales que contenían la información detallada de los crímenes ocurridos desde enero del 2011 hasta agosto del 2016. Presentaban varias etiquetas que permitían diferenciarlos según zona, grupo, tipo de crimen entre otros. Sin embargo, para efectos del trabajo de tesis, solo las siguientes etiquetas fueron de utilidad:

- Delito: Nombre del evento producido que fue clasificado según los antecedentes descritos por la víctima. Existen 674 categorías, entregando información desde los homicidios hasta hurto de accesorios de automóviles. Un ejemplo de tipo de crimen es “Robo en lugar habitado o destinado a la habitación”
- Código del delito: Número identificador que se le asoció al delito, siendo único para cada tipo de delito. Siguiendo con el ejemplo anterior, el código 809 fue asignado a “Robo en lugar habitado o destinado a la habitación”
- Agrupado: Nombre entregado a un conjunto de delitos con similares características. Había 36 nombres diferentes para esta etiqueta. Siguiendo el hilo de los ejemplos anteriores, el agrupado Robo Fuerza agrupaba los crímenes “Robo en lugar habitado o destinado a la habitación”, “Robo en lugar no habitado”, “Robo en bienes nacionales de uso público o sitios no destinados a la habitación”, “Robo de accesorios de vehículos o especies interior vehículos” y “Robo de vehículo motorizado”
- Tipo de delito: Conjunto de agrupados asociados para el interés de investigación criminal. Existían seis categorías que definen a esta etiqueta: **Control preventivo, Delitos de mayor connotación social, Desordenes, Droga, Infracción al tránsito y Violencia intra-familiar**. Para efectos de ejemplo la etiqueta **Delitos de mayor connotación social (DMCS)** reunía los agrupados Robo Fuerza, Robo Violencia, Hurto, Lesiones, Homicidios y Violaciones.
- Comuna: Nombre de la comuna donde ocurrió el evento criminal.
- Posición: Coordenadas geo-referenciadas por latitud y longitud del evento sucedido.
- Fecha: Día y hora exacta en donde se registró el crimen por parte de la víctima a Carabineros de Chile.

En el Anexo A, la Tabla A.1 entrega información completa de la base de datos con nombres de los delitos, sus códigos y su agrupado respectivo.

## 2.3. Pre-procesamiento de datos

Definidas las características de la base de datos, se dio inicio al proceso de filtrado. Se definió a consulta aquella selección de atributos para un caso de estudio generado por una persona o algoritmo. Constó de diferentes variables, pero para efectos de este Trabajo de Tesis se dejó solamente en el nombre de la comuna a analizar, la fecha de predicción y los delitos que cumplan con estar en la categoría agrupado según lo descrito. Un ejemplo de consulta se muestra en el apartado 2.1, donde se trató de analizar la comuna de Santiago para el día 22 de abril del 2015 con todos los delitos que cumplan con el nombre agrupado Robo Violencia.

$$\{\text{Comuna} = \text{'SANTIAGO'}, F_{pred} = 2015 - 04 - 22, \text{Agrupado} = \text{'ROBO VIOLENCIA'}\} \quad (2.1)$$

Con los datos recolectados tanto geográficos como criminales, se procedió a realizar el pre-procesamiento de los datos tal como muestra la Figura 2.2. La primera etapa constó de adquirir, a partir de la base de datos geográficos, todas las grillas que definen a la comuna para su posterior evaluación. Las siguientes dos etapas fueron explicadas por sub-secciones, las cuales definen la creación de los arreglos temporales y el filtrado de los delitos a partir de la consulta realizada.

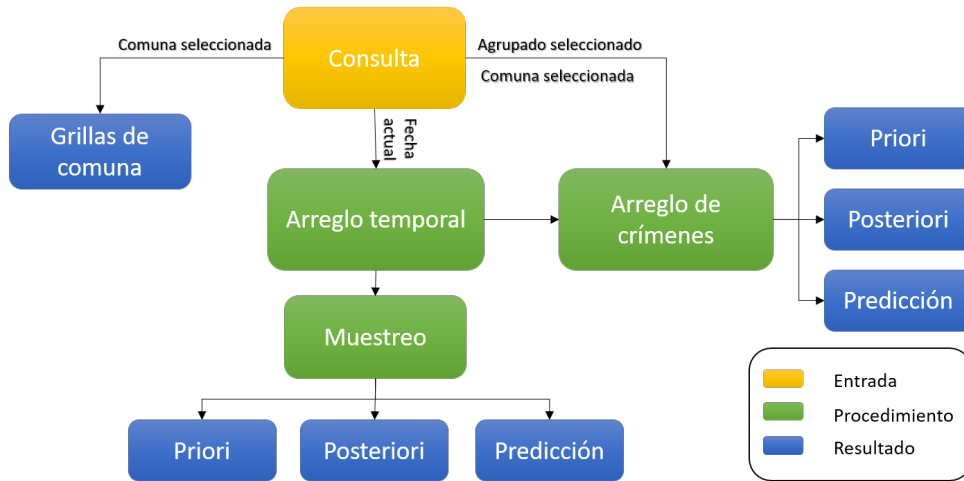


Figura 2.2: Esquema representativo de las procesos y resultados de la etapa de pre-procesamiento.

### 2.3.1. Pre-procesamiento temporal

Se usó la fecha de predicción definida en la consulta para construir tres arreglos temporales definidos con el nombre *priori*, *posteriori* y predicción. El Arreglo Temporal *priori* ( $AT_{pr}$ ) tuvo como finalidad capturar la tendencia criminal histórica ocurrida en la comuna. Su fecha de inicio parte en  $f_{init}$  y termina en  $f_{post}$ . Por otra parte el Arreglo Temporal *posteriori* ( $AT_{po}$ ) captura la tendencia criminal más reciente. Este arreglo parte desde la fecha  $f_{post}$  y termina en  $f_{pred}$ . Finalmente el Arreglo Temporal de predicción ( $AT_{pred}$ ) tuvo como objetivo entregar las fechas futuras en donde el algoritmo fue evaluado para su validación. Parte desde la fecha de predicción ( $f_{pred}$ ) entregada por la consulta y puede tener una duración de no más dos meses. Particularmente se trabajó con una duración máxima igual a una semana.

Cada uno de los arreglos temporales entró en un proceso de muestreo, el cual subdivide cada uno de los bloques respecto a una *slot*. Cada *slot* consta de una determinada cantidad de horas. Dicha cantidad debió ser divisor de 24 para evitar traslapes con los diferentes bloques. Se estableció de manera logística que el tiempo de muestreo fuera igual a ocho horas equivalente a la duración de un turno de trabajo de Carabineros de Chile.

Como resultado final se obtuvieron tres arreglos de tiempo diferenciados por distintas fechas de partida, duraciones acorde al tipo de estudio y con una cantidad determinada de *slot*. Estos arreglos fueron la base del desarrollo del procesamiento temporal dado que integraron la información de la fecha y hora de cada uno de los delitos analizados dentro de la consulta. La Figura 2.3 entrega un esquema de los arreglos temporales construidos, mostrando las diferencias entre sus largos y la cantidad de *slot* en cada uno de ellos.

Dado que en este trabajo de tesis contempla dos modelos de caracterización temporal, fue necesario estandarizar las fechas que definieron a cada uno de los arreglos temporales. Para ello se utilizaron las siguientes fechas de corte:

- $f_{init}$  : Igual a 16 meses antes de  $f_{pred}$  y con duración igual a 14 meses, terminando justo en la fecha de inicio para el estudio *posteriori* ( $f_{post}$ ) sin incluirla.
- $f_{post}$  : Parte 2 meses antes de  $f_{pred}$  y termina justo en  $f_{pred}$  sin incluirla.



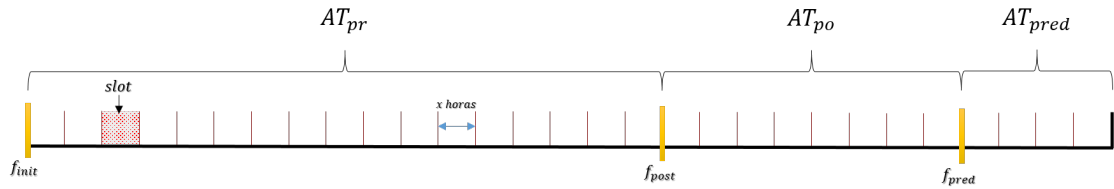


Figura 2.3: Construcción de los arreglos temporales a *priori*, *posteriori* y predicción. La unidad mínima definida es el *slot* que abarca un delta tiempo de  $x$  horas.

Además, los arreglos  $AT_{pr}$  y  $AT_{po}$  fueron separados en señales de entrenamiento y prueba, según lo descrito a continuación.

- *Priori*: La señal de entrenamiento comenzó desde  $f_{init}$  y tuvo un largo de un año completo sin contener el día asignado a  $f_{init}$ . Esto quiere decir que si  $AT_{pr}$  parte, por ejemplo, el 24 de julio del 2015, entonces la fecha final del conjunto de entrenamiento es el 23 de julio del año 2016. La señal de prueba constó del resto de las fechas entregadas por el arreglo  $AT_{pr}$  partiendo con  $f_{init}$  desplazada en un año. Siguiendo con el ejemplo, el conjunto de prueba debe contener el 24 de julio pero del año 2016 hasta el final del arreglo  $AT_{pr}$ . La Figura 2.4 muestra gráficamente el ejemplo descrito para  $AT_{pr}$ .
- *Posteriori*: Dado que  $AT_{po}$  tuvo una duración de 8 semanas, la señal de entrenamiento fue construida a partir de 7 semanas seguidas, desde el principio de  $AT_{po}$ . Finalizadas estas 7 semanas, se ocupó la última semana como señal de prueba. La Figura 2.4 muestra gráficamente el ejemplo descrito para  $AT_{po}$ .

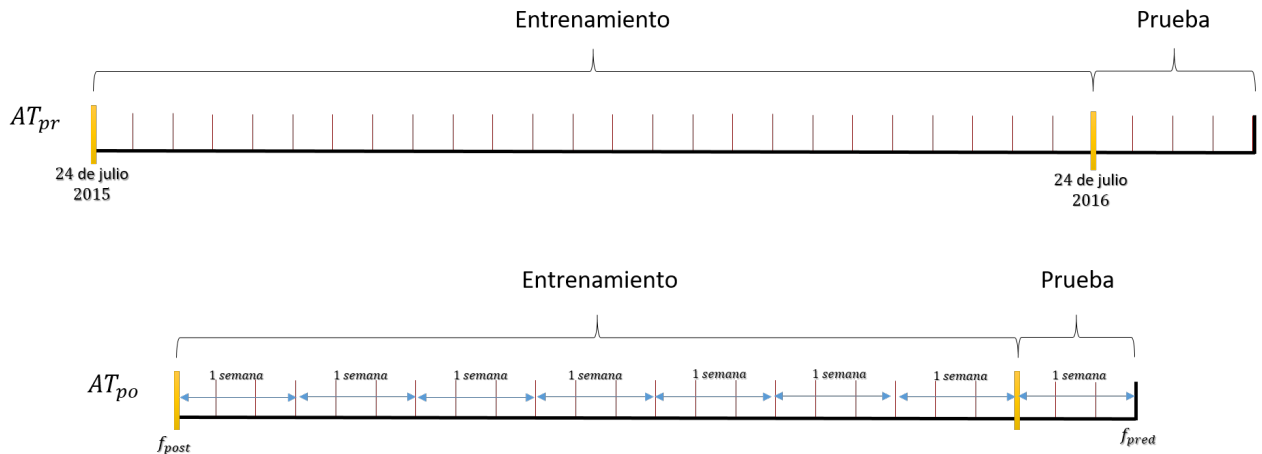


Figura 2.4: Sub-arreglos de entrenamiento y prueba de los arreglos  $AT_{pr}$  y  $AT_{po}$  (arreglo de arriba y abajo respectivamente). Las barras amarillas denotan el comienzo de cada uno de los sub-arreglos. El caso  $AT_{pr}$  mostrado es idéntico al explicado anteriormente, dando énfasis a la no inclusión de la fecha de 24 de julio del 2016 en el arreglo de entrenamiento. Por otra parte, el caso  $AT_{po}$  se denotan la duración del arreglo de entrenamiento y de prueba, especificando las fechas de inicio y final para un caso arbitrario.

### 2.3.2. Pre-procesamiento espacial

En esta etapa se trabajó con todos los delitos que cumplieron con las especificaciones entregadas por la consulta. Uno de los primeros pasos fue eliminar información irrelevante y dejar solo las características que permitían la construcción del modelo, definidas por la posición dada en latitud y longitud de los eventos criminales, la fecha y la hora exacta de ocurrencia. Ante este conjunto de datos se aplicó un nuevo filtro respecto a la fecha de ocurrencia de los delitos, separando el conjunto en tres grupos diferentes. El primer grupo definió los delitos históricos asociados al conjunto *priori*, el cual coincidió con los intervalos de tiempo entregados por el arreglo  $AT_{pr}$ . El segundo grupo reunió los delitos actuales y fueron asociado al conjunto *posteriori*, coincidiendo con el intervalo de tiempo dado por el arreglo  $AT_{po}$ . El último grupo asoció los delitos de predicción o validación y coincidió con la temporalidad del arreglo  $AT_{pred}$ , sin embargo este grupo fue valido solo cuando existían datos que cumplieran con las condiciones temporales (posteriores a la fecha predicción entregada, incluida esta), en otro caso este grupo fue considerado vacío, imposibilitando realizar la validación del modelo construido.

### 2.4. Procesamiento espacial

El procesamiento espacial permitió identificar diferentes patrones espaciales en base a las ubicaciones geo-referenciadas de los eventos criminales. Para ello se utilizó entrenamientos supervisados de parámetros que definen a una GMM bivariada (peso, vector de medias de dimensión igual a dos y su matriz de covarianza respectiva) con tal de identificar los centros de concentración criminal. La Figura 2.5 muestra el desglose de esta sección en sus respectivas sub-secciones. Cabe mencionar que este procesamiento se realizó de manera independiente para el grupo de datos *priori* y *posteriori* descritos en la sección anterior.

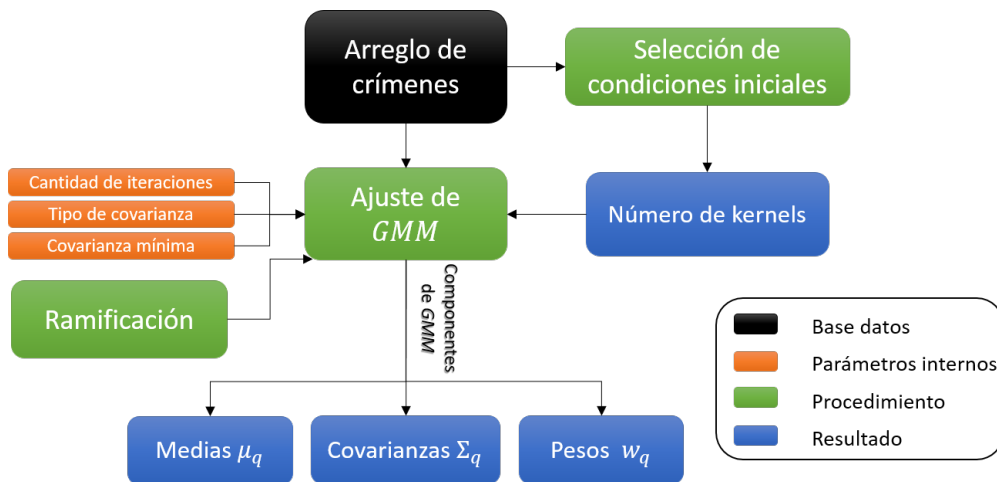


Figura 2.5: Esquema representativo de las procesos y resultados de la etapa de procesamiento espacial.

### 2.4.1. Selección de condiciones iniciales

El ajuste paramétrico de la GMM necesitó una adecuada selección de condiciones iniciales con tal de reflejar de mejor manera el comportamiento de los datos espaciales. Existieron tres condiciones iniciales que quedaron fijadas durante todo el Trabajo de Tesis y no variaron según el tipo de consulta desarrollada. Estas fueron la cantidad de iteraciones, el tipo de covarianza y la covarianza mínima, definidas también como parámetros internos del proceso de entrenamiento. Tan solo un parámetro fue variando según consulta, el cual definía el número de componentes presentes en la GMM ( $N_{kG}$ ). Este fue calculado a partir de la cardinalidad de los eventos criminales *priori* o *posteriori* según corresponda.

Al elegir un apropiado  $N_{kG}$  se pudo evitar problemas de ajustes tales como la (i) sub-representación de patrones criminales generados debido a que se entregó una cantidad insuficiente de componentes, dando como resultado pocos focos de concentración criminal y (ii) sobre-representación de patrones criminales, siendo el caso opuesto al anterior, donde la cantidad de componentes fue igual a la cantidad total de eventos en cada arreglo filtrado. Estos casos han sido representado en la Figura 2.6.

La heurística aplicada para seleccionar  $N_{kG}$  fue dada por la investigación generada por William Gosset [33], entregando que 30 muestras son suficientes para representar adecuadamente una distribución Gaussiana. Denotando a  $\|\mathcal{D}\|$  como la cantidad de delitos cometidos en el proceso *priori* o *posteriori*, se definió la cantidad de componentes según la ecuación 2.2.

$$N_{kG} = \frac{\|\mathcal{D}\|}{30} \quad (2.2)$$

Se esperó que durante el proceso de asignación de delitos a cada componente de la GMM existieran alrededor de 30 delitos. Sin embargo, no siempre se obtuvo una cantidad constante de delitos sujetos en cada componente. Para solventar este problema se diseñaron dos nuevas metodologías definidas con el nombre de Silhouette y de desviación estándar.

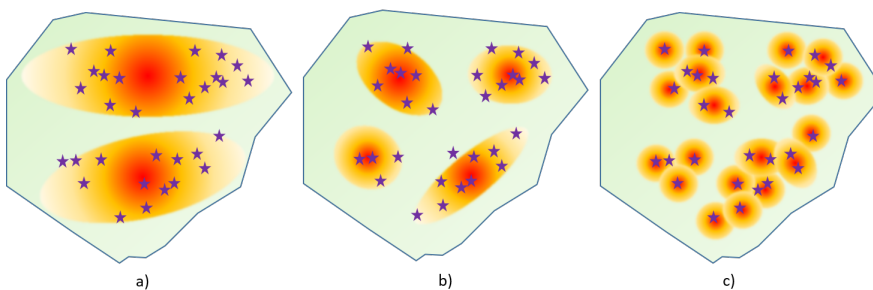


Figura 2.6: Tipos de respuestas en la caracterización por medio de GMM para diferentes  $N_{kG}$ . En a) representa el caso de sub-representación ( $N_{kG} = 2$ ), en b) se muestra el valor óptimo de  $N_{kG}$  y c) el caso de sobre-representación usando  $N_{kG}$  igual a la cantidad de delitos cometidos en el área de estudio.

- Método de Silhouette: se definió un arreglo que incluye todos los candidatos a ser el óptimo  $N_{kG}$  y se realizó un estudio de Silhouette para cada uno de estos candidatos, guardando los valores de *score* generados en dicho estudio [34]. Estos valores variaron entre  $-1$  y  $1$ , siendo el mejor candidato a representar de manera óptima una GMM el valor más cercano a  $1$ . Es por ello que se buscó la cantidad de componentes  $N_{kG}$  que maximice globalmente el arreglo de *score* generado.

- Método de desviación estándar: una vez finalizado el proceso de entrenamiento de una GMM, cada componente concentró una cantidad determinada de eventos criminales, la cual varía dependiendo del  $N_{kG}$  impuesto como condición inicial. Por ejemplo, cuando  $N_{kG} = 1$  todos los delitos se concentran en la única componente a entrenar. Con tal de uniformar la cantidad de eventos criminales asignados a cada componente, se analizó la información entregada por la desviación estándar de la cantidad de eventos criminales asignados a cada componente por GMM entrenada con diferentes valores de  $N_{kG}$ . Para ello se generó un arreglo de candidatos a ser el óptimo  $N_{kG}$  y se procedió a entrenar la GMM con cada uno de estos. El valor  $N_{kG}$  óptimo es el punto codo de la curva de desviación estándar de la cantidad de delitos en cada componente de la GMM entrenada vs el valor  $N_{kG}$ .

Estos métodos fueron implementados en el algoritmo, pero no ejecutados en este trabajo de tesis. Para mayor detalle y resultados de estas soluciones al problema de representación de patrones espaciales se sugiere consultar [35].

### 2.4.2. Entrenamiento de GMM

El ajuste de la GMM fue implementado en base al algoritmo *Expectation Maximization* (EM). El entrenamiento con los datos filtrados generó los componentes que maximizan la verosimilitud de función de probabilidades *a posteriori*. En este caso fue de interés conocer la ubicación de las medias, el valor de las covarianzas y el peso asignado que definen a la GMM. Luego se procedió a generar un diccionario de delitos cometidos en cada una de los componentes de la GMM, capturando información temporal de los delitos ocurridos, siendo posible caracterizar su comportamiento temporal. En el desarrollo de esta sub-sección se implementaron los algoritmos de optimización explicados en el Marco Teórico en Gaussian Mixture Model 1.2.2.2.

## 2.5. Procesamiento temporal

Esta etapa fue la de mayor interés en este trabajo de tesis. Esta relacionada con todo el proceso de caracterización de series temporales. Para ello se utilizaron dos métodos comparativos diferentes: modelos en base a Kernel Periodicos Gaussianos (KPG) y descomposición por medio de Transformada Wavelet (TW). El esquema de la Figura 2.7 muestra las etapas desarrolladas en esta sección. Cada componente espacial tuvo su procesamiento temporal respectivo, haciendo uso del diccionario de eventos criminales asignado y de los arreglos temporales  $AT_{pr}$  y  $AT_{po}$  pre-procesados con anterioridad.

### 2.5.1. Llenado de arreglos temporales y suavizado

La información temporal adquirida de cada componente de la GMM espacial necesito ser estandarizada con tal de uniformar el procesamiento de la señal temporal. Para ello se utilizaron los arreglos  $AT_{pr}$  y  $AT_{po}$  construidos en la etapa de pre-procesamiento temporal, partiendo con todos sus *slots* vacíos con tal de rellenarlos con la información entregada por el diccionario de eventos de cada componente.

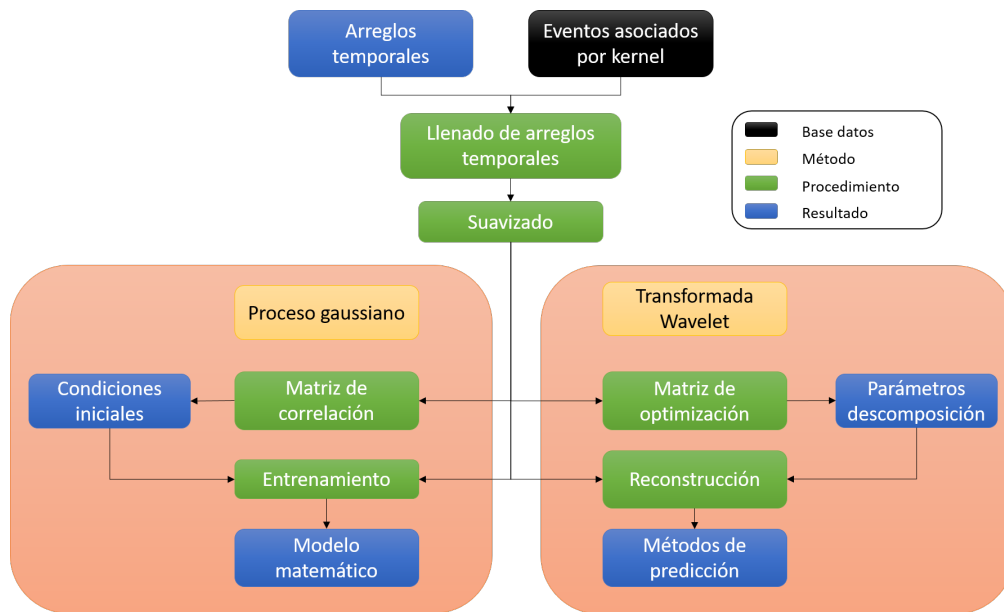


Figura 2.7: Esquema representativo de las procesos y resultados de la etapa de procesamiento temporal.

El método de llenado identificó la cantidad de delitos ocurridos dentro de cada *slot*, el cual tenía asociado una fecha con hora de inicio y fin. Con esto se tuvo, al final del proceso, un arreglo discreto, con valores  $\mathbb{N}_0$ . La Figura 2.8 esquematiza el proceso asociado a una componente específica de la GMM espacial y su resultado final. La serie temporal generada tuvo cardinalidad igual a la cantidad de *slots* del arreglo temporal llenado y un recorrido igual a la cantidad de eventos al interior de cada *slot*.

Sin embargo, las señales temporales discretas entregadas no pueden ser caracterizadas satisfactoriamente con métodos que buscan minimizar el *Mean Square Error* (MSE) durante su entrenamiento. Un ejemplo que muestra lo deficiente que podría ser este proceso es el siguiente: sea una señal temporal discreta con 17 *slots* con valor distinto a cero y otros 80 *slots* con valor igual a cero. Una vez efectuado su entrenamiento del modelo, basado en la minimización del MSE, se obtiene como respuesta una constante igual a cero para cualquier *slot* evaluado, eliminando la información entregada por los eventos criminales capturados en la serie temporal original. Esto debido a que la señal constante cero es un mínimo local dominante del problema de optimización por medio de minimización del MSE, siendo muy difícil encontrar su mínimo global.

Este problema pudo ser solucionado con técnicas que transformen las señales discretas a continuas. Una de ellas fue por medio de *Kernel Density Estimation* (KDE). Este método fue idóneo debido a que identificó dos características de suma importancia en la señal temporal discreta: (i) los *slots* donde ocurrió algún delito (soporte) y (ii) la reiteración de eventos en un mismo *slot* (intensidad). La Figura 2.9 muestra como una señal temporal discreta queda al ser suavizada por KDE.

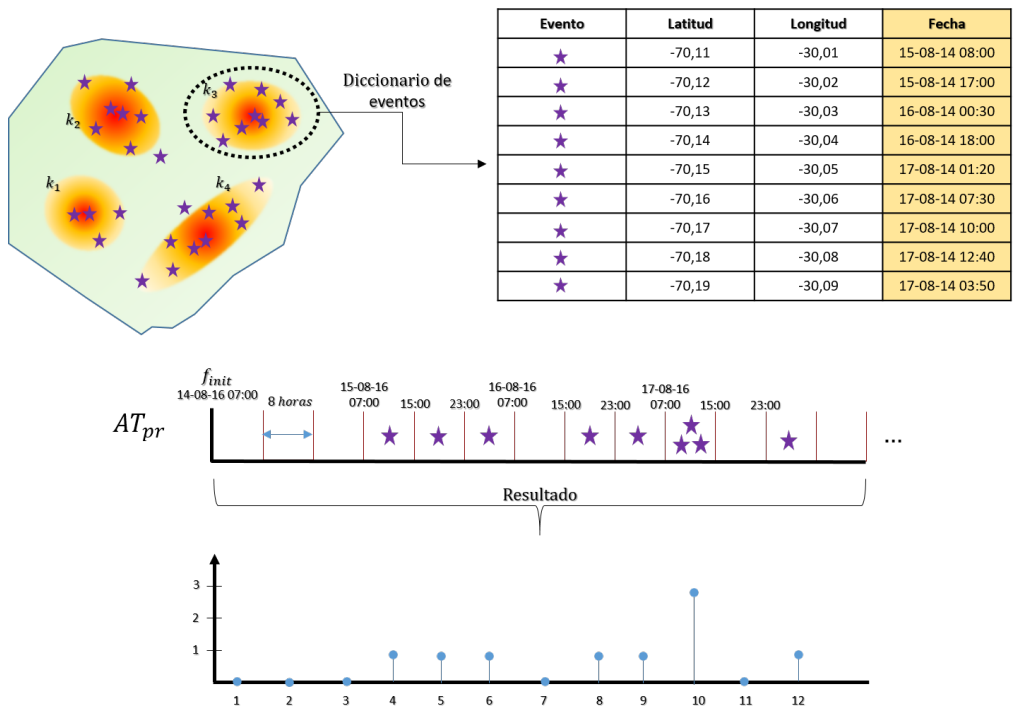


Figura 2.8: Llenado de parte del arreglo temporal a *priori* para kernel espacial  $k_3$ . Las estrellas indican los eventos criminales ocurridos en cada componente, generando un diccionario de eventos. El llenado resulta una señal discreta donde se muestran la cantidad de delitos que caen en un *slot* determinado.

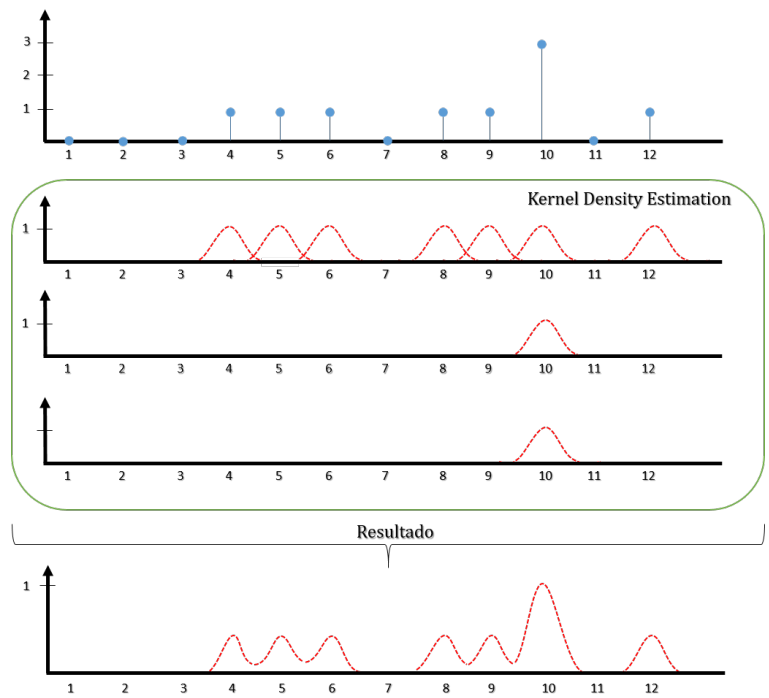


Figura 2.9: Proceso de suavizado de la señal temporal discreta haciendo uso del método KDE.

Se pudo observar que el proceso de suavizado por medio de KDE identificó tanto las posiciones temporales de los delitos cometidos (colocando un kernel del tipo Gaussiano con un ancho de banda adecuado para identificar la ocurrencia de los eventos) como la cantidad de veces ocurrido el delito en dicha ubicación temporal (agregando más kernels donde existía una mayor ocurrencia de delitos, los cuales al final del proceso fueron sumados y normalizados con valores entre 0 a 1), generando finalmente una señal temporal suavizada por medio de KDE (STS).

## 2.5.2. Modelo temporal de Kernel Periódicos Gaussianos

Se definió el modelo temporal de Kernel Periódicos Gaussianos ( $R_M$ ) como la suma ponderada funciones KPG ( $K_{pT}$ ) con diferentes parámetros internos. La ecuación 2.3 define el modelo temporal KPG de manera matemática.

$$R_M(t|r) = \sum_{i=0}^{N_{kT}} \beta_i^r \exp \left( \frac{-2 \cdot \sin \left( \frac{\pi \|t - c_i^r\|}{p_i^r} \right)^2}{l^2} \right) = \sum_{i=0}^{N_{kT}} \beta_i^r K_{pT}(t|c_i^r, p_i^r, l) \quad (2.3)$$

Donde:

- $r$  es un componente espacial arbitraria de la GMM entrenada.
- $N_{kT}$  es el número de  $K_{pT}$  asociados a un componente espacial arbitraria de la GMM entrenada.
- $\beta_i^r$  corresponde al peso asociado a la  $K_{pT}$   $i$ -ésima del componente espacial  $r$ .
- $c_i^r$  es la fase asociada a la  $K_{pT}$   $i$ -ésima del componente espacial  $r$ .
- $p_i^r$  es el periodo asociado a la  $K_{pT}$   $i$ -ésima del componente espacial  $r$ .
- $l$  es el ancho de banda de  $K_{pT}$ .
- $t$  es el argumento de la función y denota el tiempo analizado, el cual tomó valores naturales dado que se utilizó el *slot* como unidad mínima de tiempo.

Una de las razones al momento de elegir este modelo por sobre otros fue que, al ser compuesto por solamente funciones del tipo KPG ( $K_{pT}$ ), tuvo la capacidad de generar pulsos aislados con una tasa de periodo fijo e intensidad asociada al peso de la suma ponderada de diversas  $K_{pT}$ , lo cual coincide con el comportamiento general de STS. En el proceso de ajuste se buscaron los valores asociados a los parámetros que definieron los pesos, ancho de bandas, periodos y centros de cada  $K_{pT}$ . Durante el trabajo de tesis, el proceso de ajuste de una gran cantidad de parámetros fue siempre un gran limitante de cálculo y esfuerzo computacional. Es por ello que para este caso se realizó el ajuste de los pesos de cada una de las funciones usando las señales de entrenamiento y de prueba de cada arreglo temporal según corresponde, siendo los demás parámetros condiciones iniciales fijadas de manera arbitraria o con estudios previos tal como se muestra en la sub-sección 2.5.2.1.

Sea una señal temporal suavizada de una componente espacial  $r$  arbitraria ( $STS^r$ ), donde  $r \in [1 \dots N_{kG}]$  con  $N_{kG}$  la cantidad de componentes escogidos como condición inicial en el entrenamiento de la GMM espacial, la cual pudo haber sido generada en el proceso *priori* o *posteriori*.

Con el conjunto de entrenamiento se solucionó el problema de optimización definido por la ecuación 2.4, el cual minimiza el MSE entre la señal generada por el modelo  $R_M$  y la señal  $STS^r$  de entrenamiento utilizando un conjunto de pesos a optimizar  $B = \{\beta_i\}_{i=1\dots N_{kT}}$  para  $t$  definido dentro del intervalo de tiempo de entrenamiento.

$$\begin{aligned} \min_B \left( \sum_{t=1}^{t_f} \left[ STS^r(t) - \sum_{i=0}^{N_{kT}} \beta_i^r K_{pT}(t|c_t^r, p^r, l) \right] \right)^2 &= \min_B \left( \sum_{t=1}^{t_f} [STS^r(t) - R_M(t|r)] \right)^2 \\ &= \min_B e(t|r)^2 \end{aligned} \quad (2.4)$$

La optimización de 2.4 fue realiza en base al método de gradientes o Cauchy, por lo que se utilizó la información del gradiente del error de cada uno de los elementos definidos en el conjunto  $B$ , proponiendo como condición inicial  $B_0 = \{1\}_{1\dots N_{kT}}$  dado por la ecuación 2.5.

$$\nabla e(t|r) = -2 \cdot e(t|r) \cdot \sum_{t=1}^{t_f} \left[ \exp \left( \frac{-2 \cdot \sin \left( \frac{\pi \|t - c_t^r\|}{p^r} \right)^2}{l^2} \right) \right] \quad (2.5)$$

Finalmente, usando las condiciones iniciales y aplicando el entrenamiento basado por la ecuación 2.5, se obtuvo un modelo matemático evaluable en el tiempo. Sin embargo, fue necesario garantizar que el valor de número de funciones de Kernels Periódicos Gaussianos ( $K_{pT}$ ) asociados a una componente espacial de la GMM ( $N_{kT}$ ) garantice el mínimo error entre la  $STS^r$  de prueba y la evaluación del modelo en el intervalo de tiempo de prueba. Para ello se realizó una búsqueda iterativa, de la cual se necesitó de un arreglo con el  $N_{kT}$  candidatos a ser óptimos. El arreglo parte con un valor base y fue aumentando una unidad en cada iteración, realizando siempre la evaluación del MSE del modelo respecto a la  $STS^r$  de prueba. La curva generada fue analizada y procesada con tal de encontrar un mínimo local o global que denotará el mejor candidato del número de funciones de Kernels Periódicos Gaussianos ( $K_{pT}$ ) asociados a una componente espacial de la GMM.

### 2.5.2.1. Matriz de correlación y condiciones iniciales

El problema de optimización tuvo como objetivo buscar, en un componente espacial  $r$ , el mejor conjunto de pesos  $\beta_i^r$  que entregara el modelo  $R_M$  con MSE. Sin embargo, este fue tan solo uno de los cuatro conjuntos de parámetros de dicho modelo, siendo necesario elegir de manera estratégica los valores asociados a los demás con tal de propiciar la convergencia del problema de optimización. En la siguiente explicación fue asumido un componente espacial  $r$ , el cual es valido tanto para la etapa *priori* como *posteriori*.

El ancho de banda de  $K_{pT}$  ( $l$ ), criminalísticamente explicada, representaba la sensación temporal de inseguridad o de ser una nueva víctima dado un delito ocurrido. Si el valor de  $l$  aumentaba entonces el delito tendría mayor connotación durante el tiempo. Este parámetro fue fijado en cada una de los componentes espaciales, tomando un valor igual a 2, indicando que para todo delito ocurrido, su efecto temporal desaparece después de haber transcurrido dos *slots*.



Los otros parámetros dados por la  $c_i^r$  y el  $p_i^r$  fueron considerados como condiciones iniciales en el proceso de entrenamiento. Sea el arreglo con candidatos a condiciones óptimas del parámetro  $c_i^r$  ( $A_c^r$ ) y el arreglo con candidatos a condiciones óptimas del parámetro  $p_i^r$  ( $A_p^r$ ). Con cada uno de estos valores se generaron diferentes  $K_{pT}$ , los cuales fueron correlacionados con la señal STS<sup>r</sup> de entrenamiento. El resultado de cada una de las correlaciones fue condensada en una matriz, donde las filas indicaban los valores tomados por el arreglo  $A_c^r$  y las columnas los valores tomados por el arreglo  $A_p^r$ . El método de selección de condiciones iniciales para estos dos parámetros se basó en la búsqueda de los primeros  $N_{kT}$  máximos locales generados por la matriz de correlación.

La Figura 2.10 entrega un esquema de la generación de la matriz de correlación y la selección de los parámetros como condiciones iniciales.

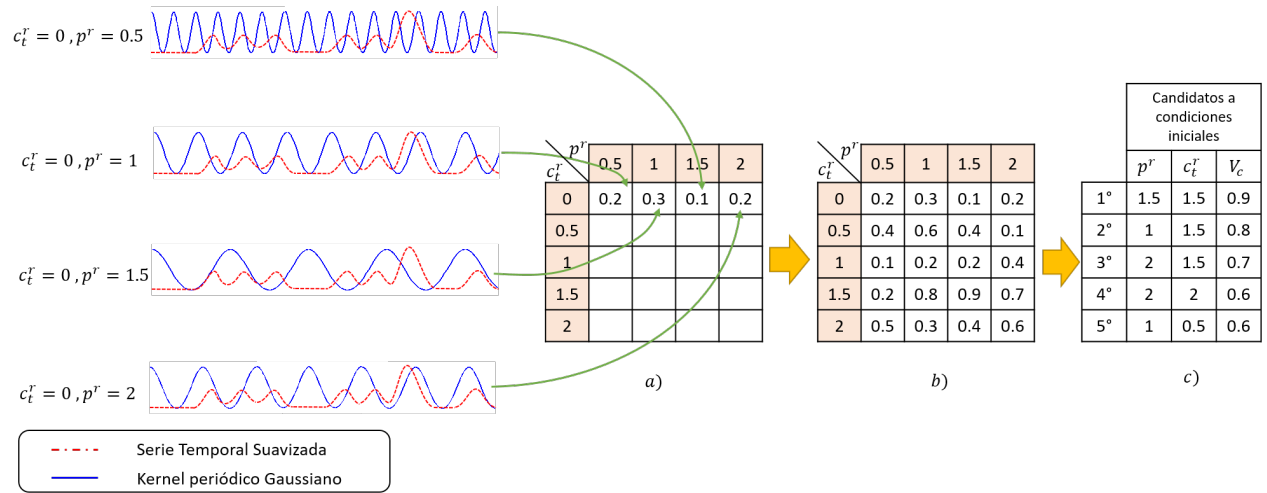


Figura 2.10: Esquema de correlación de una señal STS con un set de  $K_{pT}$  y arreglos de centros y periodos igual a  $c_t^r = [0; 0,5; 1; 1,5; 2]$  y  $p^r = [0,5; 1,0; 1,5; 2,0]$  respectivamente. La Figura 2.10 a) muestra el proceso de correlación entre  $K_{pT}$  con centro  $c_t^r = 0$  y diferentes valores de  $p^r$ . La Figura 2.10 b) muestra el resultado final del proceso de correlación, usando los arreglos  $p^r$  y  $c_t^r$  completamente. Finalmente la Figura 2.10 c) muestra la selección de los cinco mayores valores de correlación y con ello los parámetros que hicieron posible dicha correlación

### 2.5.3. Modelo temporal Wavelet

Dado que cada señal temporal suavizada por medio de KDE (STS) tuvo un comportamiento no estacionario, dado que acumuló una gran cantidad de intensidad en lapsos muy cortos de tiempo, el estudio de las periodicidades principales de cada STS por medio de la Transformada de Fourier (TF) fue imposibilitada. Ante esta situación, la Transformada Wavelet (TW) fue una solución acertada al momento de encontrar los periodos de mayor importancia de la señal dado que permitió el análisis de señales no estacionarias con gran acumulación de energía en *peaks* aislados de tiempo por medio de una descomposición espectral de energía a diferentes escalas de una función base ortogonal. Ante la posibilidad de filtrar los periodos de mayor importancia de cada STS, fue posible transformar dicha señal en una cuya base presente una importante presencia de los periodos seleccionados haciendo uso de la Anti-Transformada Wavelet (ATW).

A partir de lo señalado, fue necesario explicar el proceso de descomposición por TW, su razonamiento y las técnicas de búsqueda paramétrica utilizadas para una adecuada descomposición. También se necesitó explicar el uso de la ATW para crear la nueva señal en base a los periodos seleccionados, mostrando finalmente el proceso de predicción para las STS del caso *priori* y *posteriori*.

### 2.5.3.1. Descomposición

Sea una STS de largo  $N$  y una TW con función Wavelet normalizada ( $\psi(\eta)$ ) y parámetros bien definidos. Sea un arreglo de escalas  $S = \{s_j\}_{j=0, \dots, J}$ , ordenado de menor a mayor, donde  $J$  es la dimensión del arreglo e  $j \in [1 \dots J]$ . Al aplicar la TW con un valor de escala  $s_j$  a una señal temporal cualquiera, fue posible obtener un arreglo de dimensión  $1 \times N$  con valores reales e imaginarios dependiendo de  $\psi(\eta)$ . A dicho arreglo se le aplicó el concepto de Espectro de Energía de Transformada Wavelet (EETW) con tal de obtener los componentes de energías principales del arreglo temporal transformado.

Para una  $STS^r$ , con  $r$  una componente específica de la GMM, los resultados generados por cada TW usando las escalas de  $S$  fue concatenado formando la Matriz de Energía Wavelet (MEW) de dimensión  $M \times N$ , siendo posible identificar cuales fueron las escalas de mayor concentración de energía sumando a lo largo de las filas, generando un nuevo arreglo de dimensiones  $M \times 1$  definido por el nombre de Energía Acumulada por Escala (EAE).

El siguiente paso fue escoger los periodos de mayor relevancia en la  $STS^r$  analizada. Para ello el EAE fue segmentado en bandas de escalas diferenciadas por un ancho igual a  $bp$ . Cada una de las bandas tuvo una escala representante igual al valor medio entre el valor inicial y final de su banda. El proceso de descomposición se repitió hasta que se pudo escoger las primeras diez bandas que acumularon la mayor cantidad de energía, tomando con ellas sus escalas representantes. La identificación de los periodos de mayor importancia de la  $STS^r$  analizada fue por medio de la Tabla 1.2 y el valor de las diez primeras escalas representantes del EAE.

### 2.5.3.2. Construcción de modelo temporal de predicción

El proceso de predicción por medio del análisis Wavelet se basó en la hipótesis de que los incidentes criminales cumplieron con una recurrencia cíclica (reiteración del delito en zonas donde antes habian ocurrido). En el proceso de caracterización del riesgo criminal *priori* la hipótesis de base fue que los delitos tenían una recurrencia del tipo anual, es decir que si un delito ocurrió el 24 de junio es posible que en el próximo año ocurra lo mismo en fechas cercanas. En el proceso *posteriori* se asumió que la actividad delictual semanal esta relacionada con lo observado en semanas anteriores pero con diferentes grados de importancia.

La información entregada por EAE permitió encontrar las diez primeras escalas representativas de mayor concentración de energía MEW, seleccionando las filas respectivas a cada escala  $s_q$ , con  $q \in [1, \dots, 10]$ , las cuales al aplicar la ATW sobre cada una de estas, fue posible generar una señal temporal con fuerte presencia de la escala  $s_q$  ( $SES_q^r$ ). Con dicha señal se pudo construir un modelo temporal tanto para el proceso *priori* como *posteriori* de un componente espacial  $r$  de la GMM. La ecuación 2.6 define la señal de soporte temporal ( $SST^r$ ) que reúne la información más importante extraída de MEW.

$$SST^r(t) = \sum_{q=1}^{10} SES_q^r(t) \quad (2.6)$$

Asumiendo las hipótesis entregadas, los métodos para desarrollar los modelos temporales dado un componente espacial  $r$  fueron:

- Caso *priori*: El primer paso para generar el modelo temporal *priori* Wavelet ( $R_W^{prio}$ ) fue identificar la fecha de predicción del año anterior dentro del intervalo de tiempo generado para el proceso *priori*. En la Figura 2.11 fue denotado con el nombre de  $f_{priori}$  y solo tiene validez en este estudio.  $R_W^{prio}$  consiste en replicar  $SSP^r$  entre las fechas  $f_{priori}$  y  $f_{post}$ . Con esto el soporte temporal de  $R_W^{prio}$  será igual al intervalo dado por las fechas  $f_{priori}$  y  $f_{post}$ .

La ecuación 2.7 definió matemáticamente lo descrito con anterioridad, donde  $Q_{slots}$  indica la cantidad de *slot* que se le debe sumar a  $SST^r$  con tal de eliminar los meses de diferencia entre  $f_{init}$  y  $f_{priori}$ .

$$R_W^{prio}(t|r) = \sum_{q=1}^{10} SES_q^r(t + Q_{slots}) \quad (2.7)$$

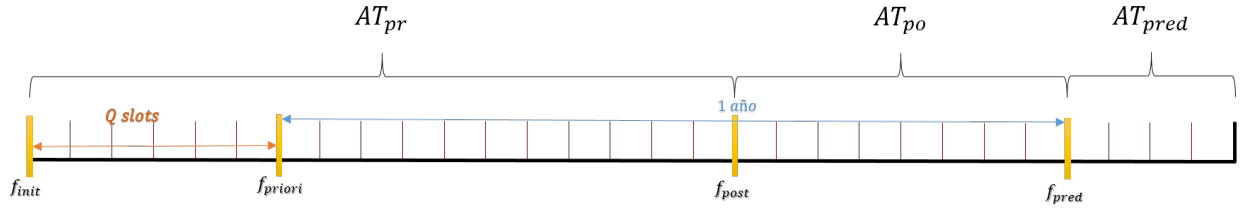


Figura 2.11: Fechas de corte tomadas en cada uno de los arreglos temporales para el método Wavelet.

- Caso *posteriori*: Se procedió a separar  $SST^r$  en diferentes señales de duración igual a una semana ( $SS_w^r$ ). Fue posible obtener 8 semanas completas dado que la señal  $SST^r$  provenía de una señal de duración igual a dos meses. A cada  $SS_w^r$ , con  $w$  la semana segmentada y perteneciente al conjunto  $[1, \dots, 8]$ , se le aplicó un factor de olvido  $\lambda_{fo}^{w-1}$ , donde  $w$  partió con valor igual a 1 y fue aumentando en una unidad por cada semana anterior a  $f_{pred}$ . Dada la construcción realizada, el modelo temporal *posteriori* Wavelet ( $R_W^{post}$ ) tuvo soporte temporal igual a una semana y quedó definido matemáticamente en la ecuación 2.8.

$$R_W^{post}(t|r) = \sum_{w=1}^8 SS_w^r(t) \cdot \lambda_{fo}^{(q-1)} \quad (2.8)$$

Los modelos  $R_W^{prio}$  y  $R_W^{post}$  fueron normalizados entre 0 a 1. Para ello se eliminó el sesgo entregado por el valor mínimo del modelo evaluado a lo largo de su soporte temporal y luego se dividió por el máximo global de la señal insesgada.

### 2.5.3.3. Matriz de optimización y parámetros iniciales

El uso de TW permitió descomponer diferentes tipos de señales. Sin embargo, fue necesario conocer los valores de los parámetros que permitieron optimizar el análisis de la TW. Dichos valores debieron estar relacionados con el comportamiento de la señal estudiada.

Durante el proceso de optimización se analizaron el parámetro interno de la función Wavelet ( $m$ ) y ancho de las bandas aplicada en el arreglo de Energía Acumulada por Escala ( $bp$ ). Para encontrar estos valores se utilizaron las señales de entrenamiento y prueba descritas en la Sección 2.3.1.

Se generaron dos arreglos con candidatos a ser óptimos de  $m$  y  $bp$ . Los valores de cada uno de los arreglos fueron mezclados entre si, con tal de formar una dupla de valores  $\{m, bp\}$ . Con cada una de las duplas se procedió a realizar, de manera independiente, la TW en cada una de las señales de entrenamiento del proceso *priori* y *posteriori*. Finalmente se calculó el *Mean Square Error* (MSE) con la señal de prueba respectiva, seleccionando la dupla que minimiza dicho error.

## 2.6. Predicción

Los modelos analizados anteriormente entregaron respuestas de forma separadas entre el proceso espacial y temporal. Sin embargo, este tipo de respuesta parcial no fue suficiente debido a que se necesita de un modelo que combine de manera satisfactoria la parte espacial con la temporal.

La caracterización espacio-temporal implementada en este Trabajo de Tesis consideró que en cada componente espacial existe un modelo temporal evaluable en un horizonte de tiempo no mayor a una semana. Dicha evaluación temporal sirve como ponderador lineal a la respuesta espacial generada por la GMM.

Sea una comuna arbitraria, la cual ha pasado por el proceso de entrenamiento espacial mediante GMM generando  $N_{k_{prio}}$  componentes espaciales *priori* ( $k_i^{prio}$ ) y  $N_{k_{post}}$  para el caso *posteriori* ( $k^{post}$ ). Cada  $k^{prio}$  se generó un modelo temporal por medio de transformaciones KPG ( $R_M^{prio}$ ) o por Transformada Wavelet ( $R_W^{prio}$ ). Lo mismo ocurrió en cada  $k^{post}$ , definiendo los modelos temporales como  $R_M^{post}$  o  $R_W^{post}$  respectivamente.

Sea, además, el coeficiente inter-proceso temporal ( $\lambda$ ) que asignó diferentes grados de importancia a cada uno de los procesos *priori* y *posteriori*. El modelo de predicción del riesgo criminal ( $R^{pred}$ ) dado un punto en el espacio  $\vec{x}$ , definido por latitud y longitud, y de  $t$ , que muestra la cantidad de pasos que se desea observar en el futuro (cuantificado por medio de *slot*) usando TW para el proceso temporal, quedó reflejado por las ecuaciones 2.9 y 2.10 en formato condensado.

$$\begin{aligned}
 R^{pred}(\vec{x}, t) = & (1 - \lambda) \sum_{r=1}^{N_{k_{prio}}} \left[ \sum_{q=1}^{10} [SES_q^r(t)] \cdot k_r^{prio}(\vec{x}) \right] \\
 & + \lambda \sum_{r=1}^{N_{k_{post}}} \left[ \sum_{q=1}^8 [SS_q^r(t) \cdot \lambda_{fo}^{(q-1)}] \cdot k_r^{post}(\vec{x}) \right] \quad (2.9)
 \end{aligned}$$

$$R^{pred}(\vec{x}, t) = (1 - \lambda) \sum_{r=1}^{N_{kprio}} R_W^{prio}(t|r) \cdot k_r^{prio}(\vec{x}) + \lambda \sum_{r=1}^{N_{kpost}} R_W^{post}(t|r) \cdot k_r^{post}(\vec{x}) \quad (2.10)$$

Por otra parte, el modelo de predicción del riesgo criminal usando el criterio KPG para el análisis temporal quedó definido para las ecuaciones 2.11 y 2.12 en formato condensado.

$$R^{pred}(\vec{x}, t) = (1 - \lambda) \sum_{r=1}^{N_{kprio}} \left[ \left[ \sum_{i=0}^{N_{kT}} \beta_1^r \exp \left( \frac{-2 \cdot \sin \left( \frac{\pi \|t - c_t^r\|}{p^r} \right)^2}{l^2} \right) \right] \cdot k_r^{prio}(\vec{x}) \right] + \lambda \sum_{r=1}^{N_{kpost}} \left[ \left[ \sum_{i=0}^{N_{kT}} \beta_1^r \exp \left( \frac{-2 \cdot \sin \left( \frac{\pi \|t - c_t^r\|}{p^r} \right)^2}{l^2} \right) \right] \cdot k_r^{post}(\vec{x}) \right] \quad (2.11)$$

$$R^{pred}(\vec{x}, t) = (1 - \lambda) \sum_{r=1}^{N_{kprio}} R_M^{prio}(t|r) \cdot k_r^{prio}(\vec{x}) + \lambda \sum_{r=1}^{N_{kpost}} R_M^{post}(t|r) \cdot k_r^{post}(\vec{x}) \quad (2.12)$$

### 2.6.1. Evaluación del riesgo en la comuna

Durante el proceso de evaluación del modelo de predicción del riesgo criminal se utilizaron las grillas entregadas por la base de datos geográficos (ver sección 2.2.1). Se definió a  $G$  como el conjunto que contenía las  $N_G$  grillas  $g^h$  de una comuna cualquiera. Sea también un tiempo de evaluación futura  $t'$ . El modelo fue evaluado en cada una de las grillas  $g^h$  calculando una integral aproximada que dio a entender el riesgo acumulado en cada una de las grillas. Esta aproximación se realizó utilizando la evaluación de 25 puntos ( $g_i^h$ ), los cuales distribuyeron de manera uniforme en cada una de las grillas del conjunto  $G$ , siendo finalmente sumados tal como muestra la ecuación 2.13.

$$R^{pred}(g^h, t') = \sum_{i=1}^{25} R^{pred}(g_i^h, t'), \forall g^h \in G \quad (2.13)$$

La evaluación finalmente aglomeró una cantidad finita de grilla evaluadas con las cuales se conformó la superficie de riesgo criminal de la comuna para un tiempo  $t'$  determinado.

### 2.6.2. Uso de métricas de desempeño

Medir el desempeño del modelo fue una tarea que tuvo cierto grado de subjetividad, debido a que depende de cuanta área es posible abarcar con los recursos humanos destinados al patrullaje preventivo de Carabineros de Chile en una cierta comuna. Sea un umbral  $\theta$  que tomó valores entre 0 y 1 de manera discreta. Sea además una superficie de riesgo criminal de una comuna arbitraria para un tiempo  $t'$  establecido. A dicha superficie generada por  $R^{pred}$  se procedió a normalizar, quedando con valores que iban desde 0 a 1. Se definió como grillas activadas a todas aquellas donde su valor de riesgo normalizado superó a un valor determinado de  $\theta$ .

El poder de acierto exigido a la respuesta del modelo de predicción varió según el valor del parámetro  $\theta$ . Existieron dos casos importantes a notar. El primero es cuando se analizaban valores pequeños del parámetro  $\theta$ , mostrando una gran cantidad de grillas activas las cuales eran susceptibles a ocurrir un delito (caso de mayor área de cobertura). Por el contrario, si  $\theta$  tomaba valores grandes, el análisis de la respuesta del modelo se volvía más exigente, dado que se activaban una baja cantidad de grillas donde podrían ocurrir delitos.

Si bien el valor de  $\theta$  pudo ser fijado arbitrariamente en este trabajo de tesis, por parte de Carabineros de Chile este factor de análisis dependió netamente del recurso humano destinado al patrullaje preventivo. Por ejemplo, si una unidad de Carabineros de Chile tuviera baja cantidad de recurso humano para el patrullaje y por ello una baja cantidad de área para ser patrullada, se esperaría por parte del modelo una alta precisión, reduciendo al máximo las grillas que muestren una tendencia a tener un delito. Es por ello que en ese caso el valor de  $\theta$  debería ser cercano a 1 (análisis más exigente)

Dado los diferentes criterios de análisis de la respuesta entregada por el modelo de predicción fue necesario establecer una métrica de desempeño que no tuviera ambigüedad al momento de decir que la respuesta del modelo fue satisfactoria o no. En este caso la métrica *Area Under Curve* (AUC) permitió analizar todos los casos denotados por el parámetro  $\theta$ , siendo posible su construcción a partir de las funciones *High Risk Percentage* ( $HRP_\theta$ ) y *True Incident Percentage* ( $TIP_\theta$ ) explicadas en la Sección 1.4.

Sea una comuna de ejemplo cuadrada de 12 grillas evaluada por  $R^{pred}$  para un tiempo determinado. La evaluación normalizada del modelo en la comuna hipotética se entrega dentro de cada grilla, la cual se le asoció colores que van desde el amarillo (menor valor de riesgo) hasta el rojo (mayor valor de riesgo). A parte se ha colocado, con una estrella color azul, los delitos que ocurrieron efectivamente dentro de cada de las grillas, tal como se muestra en la Figura 2.12 (a). El cálculo de la función  $HRP_\theta$  y  $TIP_\theta$  fue a partir de los diferentes valores de  $\theta$  y la superposición de dichos valores respecto al parámetro  $\theta$  en común permitieron generar la curva resultante con la cual se midió el AUC.

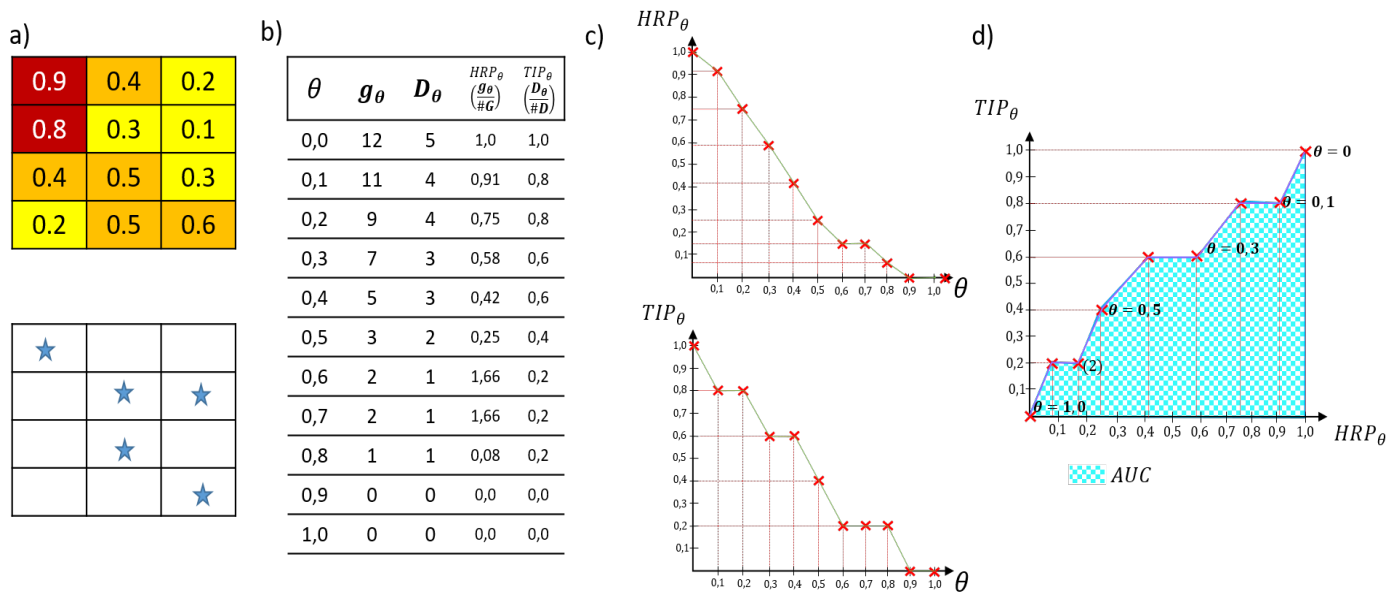


Figura 2.12: Ejemplo de comuna con 12 grillas evaluadas con el modelo  $R^{pred}$  normalizado para un tiempo arbitrario.

La Figura 2.12 a) entrega los valores de cada grilla y la ubicación de los delitos cometidos para el proceso de evaluación con estrellas. La Figura 2.12 b) muestra una tabla con el llenado de valores respecto a las grillas activadas  $g_\theta$  y los delitos que ocurren en dichas grillas  $D_\theta$  para diferentes valores de  $\theta$ . La cantidad de grillas totales esta dada por  $\#G$  y el número total de eventos dado por  $\#D$ . La figura 2.12 c) entrega las curvas  $HRP_\theta$  y  $TIP_\theta$  generadas a partir de los datos de la tabla de la Figura 2.12 b). Finalmente la curva de la Figura 2.12 d) es la superposición de las dos funciones gracias a que comparten la misma variable  $\theta$  y el área bajo la curva definida como AUC.

# Capítulo 3

## Resultados

En este capítulo se dio a conocer las respuestas tanto espaciales como temporales del modelo de caracterización del riesgo criminal mediante datos geo-referenciados, su proceso de optimización basado en diferentes parámetros internos y del desempeño en diversas comunas de la región Metropolitana. Una de las cualidades que presentó el modelo implementado fue la posibilidad de entregar una respuesta para cualquier comuna, siempre y cuando se tuviera un registro de incidentes criminales durante los últimos 2 meses. La respuesta entregada tuvo una validez igual a una semana y fue medida cada 8 horas, obteniendo una predicción de 21 pasos. Dependiendo de la fecha consultada fue posible realizar el proceso de validación usando las métrica AUC definida en la Sección 2.6.2. Además el modelo permitió realizar estudios para cualquier tipo de delito o conjunto de ellos. Todas estas características le dieron al modelo diseñado una enorme versatilidad.

### 3.1. Herramientas de hardware y software utilizados

El algoritmo del Trabajo de Tesis fue escrito en lenguaje Python (v2.7). Dicho lenguaje permitió utilizar diversas librerías disponibles en Internet y que son gratuitas y de libre instalación y administración, permitiendo posibles modificaciones con tal de ajustarse a los requerimientos del proyecto. Se utilizó como plataforma de trabajo el software Spyder y las librerías que permiten la ejecución del algoritmo fueron:

- Matplotlib (v1.5.1) = Librería que permite graficar arreglos numéricos.
- Numpy (v1.8.2) = Librería que permite el lenguaje numérico en Python.
- Pandas (v0.17.1) = Librería que facilita el uso de diccionarios y arboles, permitiendo una búsqueda eficiente de datos.
- Pyswarm (v0.6) = Librería que permite la optimización de un modelo en base a PSO.
- Scipy (v0.14.0) = Librería de Machine Learning.
- Shapely (v1.5.13) = Librería que facilita la exportación de objetos geográficos en formato SHP.
- PyQt4 (v4.11.4) = Librería que permite la conexión con base de datos usando protocolos PostgreSQL.
- Mlpy (v3.5) = Librería que entrega una sencilla manera de realizar Transformadas Wavelet.



- Pyshp (v1.2.3) = Librería que facilita el uso de objetos geográficos y de diversas operaciones sobre estos.
- Scikit-learn (v0.17.1) = Librería de Machine Learning.

Dado que el proyecto de investigación estuvo inserto en una problemática que utiliza información geo-referenciada, fue necesario integrar un software que permitiera la visualización actualizada de calles y relieves de diferentes partes del país. El uso de QGIS<sup>®</sup> entregó herramientas claves para la visualización de los resultados del modelo en diferentes comunas de la Región Metropolitana, entregando detalles importantes al momento de identificar la causa de diferentes patrones criminales presentes en la respuesta del modelo evaluado.

Por parte del hardware utilizado, CEAMOS entregó un computador portátil de alto desempeño, como se puede observar en la Tabla 3.1. Con dicho portátil se escribió parte del algoritmo y se realizaron algunas pruebas preliminares con tal de evidenciar el funcionamiento adecuado del modelo generado. Sin embargo, todos los resultados entregados en este Trabajo de Tesis fueron realizados por un computador particular con sus requerimientos descritos igualmente en la Tabla 3.1.

	Notebook	Computador
Modelo/Motherboard	ASUS ZenBook Pro UX501VW	ASROCK Z170 Extreme
CPU	Intel Core i7 6700HQ (2600 MHz - 3500 MHz)	Intel Core i7 6700K (4000 MHz)
RAM	16 Gb DDR4 2133	32 Gb DDR4 2400
GPU	NVIDIA GeForce GTX 970M	NVIDIA GeForce GTX 1060
Disco Duro	512 Gb M.2	1 Tb HDD / 256 Gb SSD
OS	Windows 10	Windows 10

Tabla 3.1: Características de los distintos hardware utilizados en el Trabajo de Tesis.

## 3.2. Casos de estudio

El modelo fue evaluado en la comuna de Maipú durante el mes de junio del 2016. La Figura 3.1 ubica a la comuna de Maipú en los límites establecidos por el Gran Santiago. La base delictual escogida consta de todos los delitos que sean del agrupado Robo Violencia o Robo Fuerza (ver tabla A.1 de Anexo A), siendo este conjunto el de mayor interés por parte de Carabineros de Chile debido a que la gran parte de estos delitos ocurren en la vía pública. La cantidad de grillas que definieron la comuna fue igual 6429, sin embargo, el modelo fue aplicado tan solo en las grillas de importancia delimitadas por Carabineros de Chile. Durante el mes seleccionado se produjeron 578 delitos, con los cuales se validó la respuesta del modelo en cada uno de los días del mes evaluado.

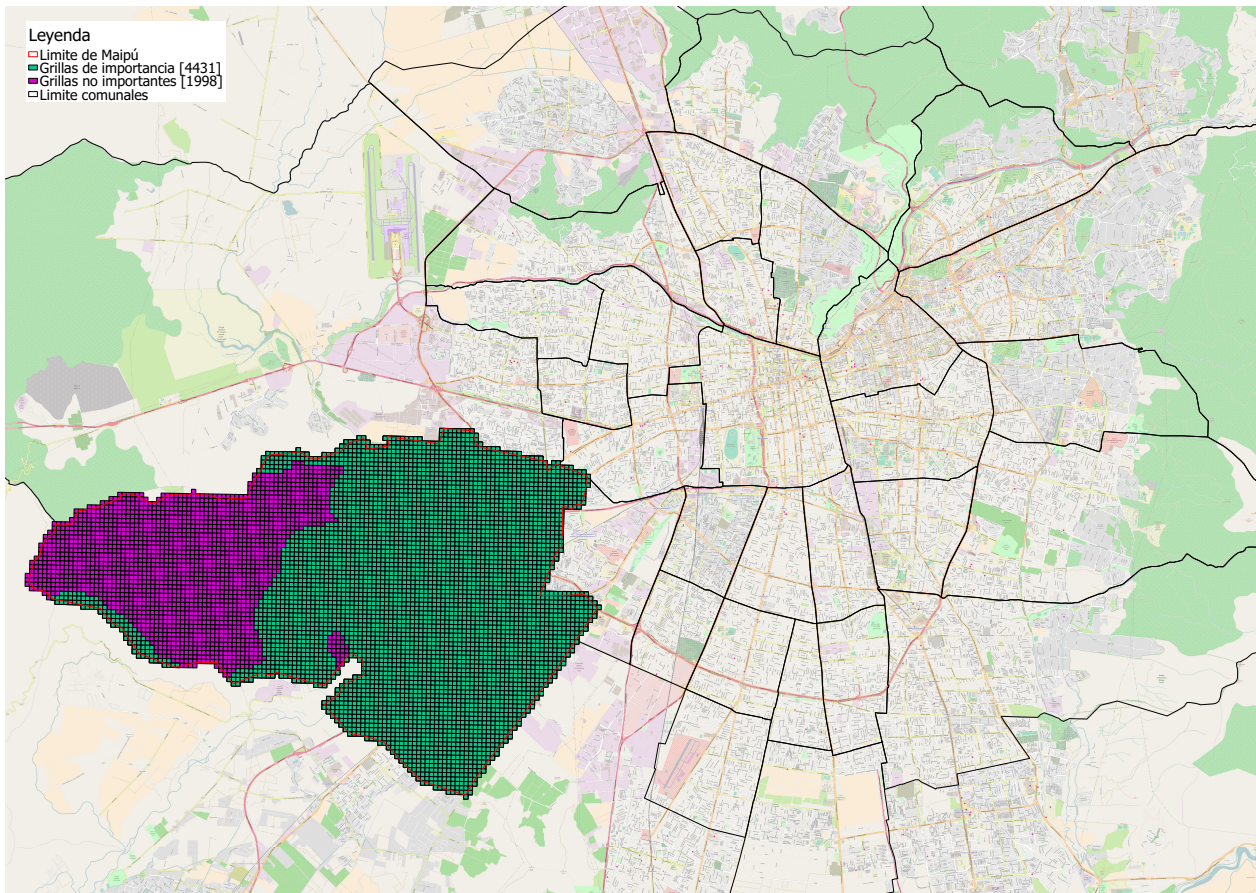


Figura 3.1: Comunas del Gran Santiago. En color morado se muestra la comuna de Maipú. Imagen capturada gracias a herramientas de visualización dadas por QGIS®.

### 3.3. Resultado base

Durante esta sección se describió cada uno de los resultados entregados por los procesos espaciales y temporales del modelo de caracterización espacio-temporal del riesgo criminal en una fecha específica y con parámetros arbitrariamente escogidos. Como primer paso se estudió la cantidad de componentes espaciales utilizados para cada caso consultado, identificando la distribución de la cantidad de delitos asignados en cada componente espacial: Esto fue de suma importancia en la creación de los arreglos temporales y su posterior predicción. Con los componentes entrenados y con sus delitos asignados, se generaron las respuestas temporales utilizando el modelo temporal de Kernel Periódicos Gaussianos ( $R_M$ ), el modelo temporal *priori* Wavelet ( $R_W^{prio}$ ) y el modelo temporal *posteriori* Wavelet ( $R_W^{post}$ ). Se aplicó en este caso base, un set de parámetros no necesariamente óptimo en cada modelo temporal, generando una señal de predicción que evoluciona en el tiempo que finalmente fue combinada con el proceso espacial tal como lo muestra la ecuación 2.9, que usó técnicas de TW y 2.11 representando el caso de transformaciones KPG.

Este Trabajo de Tesis utilizó la asignación del número de componentes presentes en la GMM ( $N_{kG}$ ) dado por la ecuación 2.2. La optimización del número de componentes espaciales para una consulta específica fue explicada en el trabajo de tesis del Sr. Mario Vergara [35]. Sin este estudio el proceso de caracterización del riesgo criminal no es óptimo según la información

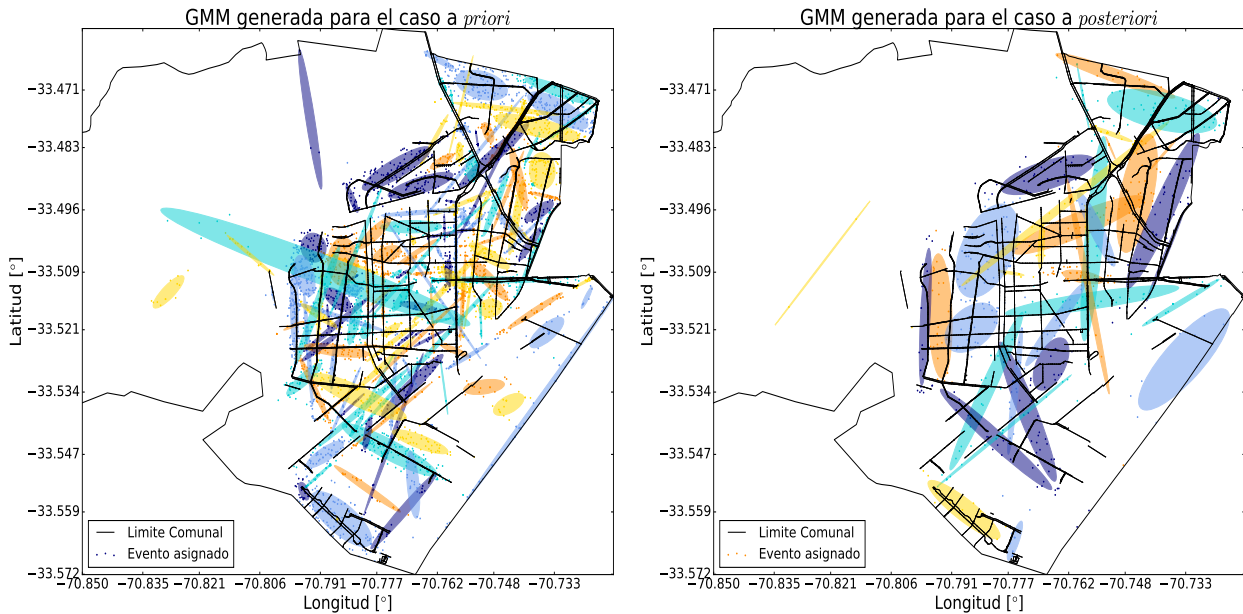
espacial. Lamentablemente el tiempo de trabajo no permitió la integración de esta mejora, pero es de esperar que en trabajos futuros estén integrados en conjunto con los estudios descritos en este Trabajo de Tesis.

La primera etapa de este caso base de estudio fue fijar una fecha dentro del mes de junio para observar su desarrollo espacial, seguido por establecer las condiciones internas de la *Gaussian Mixture Model* (GMM), siendo  $N_{kG}$  calculado en base a la cantidad de delitos para la etapa *priori* y *posteriori*. La Tabla 3.2 entrega los valores establecidos en el caso base. El coeficiente inter-proceso temporal ( $\lambda$ ) de las ecuaciones 2.10 y 2.12 quedó definido con un valor igual a 0.7, dando mayor importancia a la etapa *posteriori* de cada modelo.

Parámetro	Valor	Datos	Valor
Fecha de predicción	04-06-2016	# de eventos <i>priori</i>	7699
Cantidad de iteraciones	500	$N_{kG}$ <i>priori</i>	256
Tipo de covarianza	<i>full</i>	# de eventos <i>posteriori</i>	1107
Covarianza mínima	$10^{-8}$	$N_{kG}$ <i>posteriori</i>	36

Tabla 3.2: Parámetros y datos utilizados para generar el entrenamiento espacial de los datos en la comuna de Maipú en la fecha 04-06-2016

Las respuestas de la parte espacial, utilizando un entrenamiento por medio de *Expectation Maximization* (EM), tanto en la fase *priori* como *posteriori*, quedaron reflejadas en las Figuras 3.2a y 3.2b.



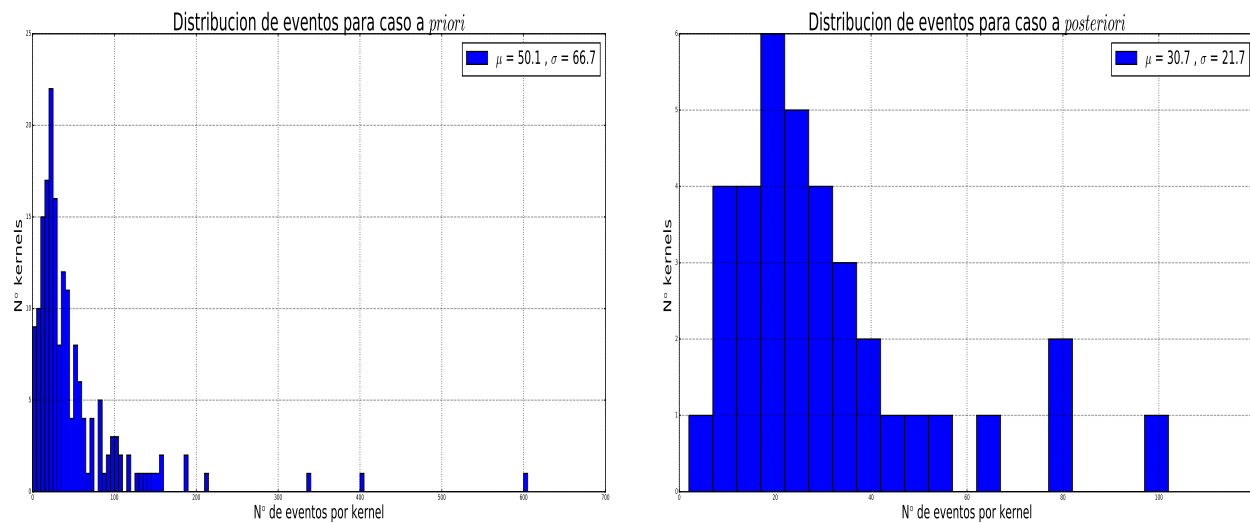
(a) Distribución de los componentes espaciales encontrados mediante el entrenamiento por EM de una GMM en la fase *priori* del caso de estudio.

(b) Distribución de los componentes espaciales encontrados mediante el entrenamiento por EM de una GMM en la fase *posteriori* del caso de estudio.

Figura 3.2: Proceso de caracterización espacial de los eventos criminales usando técnicas de entrenamiento de GMM. Se selecciona solo la zona urbana de la comuna de Maipú.

Las elipses de múltiples colores representaron la forma de las covarianzas asignadas a cada componente espacial. Las covarianzas mostradas fueron centradas en sus medias respectivas, dado el estudio de optimización mediante EM. Los eventos criminales asignados a cada componente espacial fueron superpuestos como puntos del mismo color de las elipses graficadas. El hecho de que los colores se hayan repetido no significó que las componentes espaciales hayan sido iguales y es tan solo una ayuda visual al momento de diferenciar, dentro de la GMM entrenada, un componente espacial con otro.

La Figura 3.3a y 3.3b entregaron la distribución de la cantidad de delitos asignados a cada componente espacial en la fase *priori* y *posteriori* respectivamente. Se pudo observar que el método heurístico que ajusta la cantidad de delitos en cada componente espacial, dado por la ecuación 2.2, no permitió uniformar la cantidad de delitos asignados a cada componente espacial. Era de esperar que la mayoría de los componentes tuvieran una cantidad igual a la dada por el denominador de la ecuación 2.2, sin embargo el resultado obtenido denotó una fuerte desviación en la cantidad de delitos por componentes espaciales, la cual quedó reflejada en la leyenda de los gráficos entregados. Por otro lado, la distribución presentó una media muy similar al valor denotado por el denominador de la ecuación 2.2. Ante estos problemas se dio comienzo al trabajo de optimización de este parámetro en el trabajo de tesis del Sr. Mario Vergara [35].



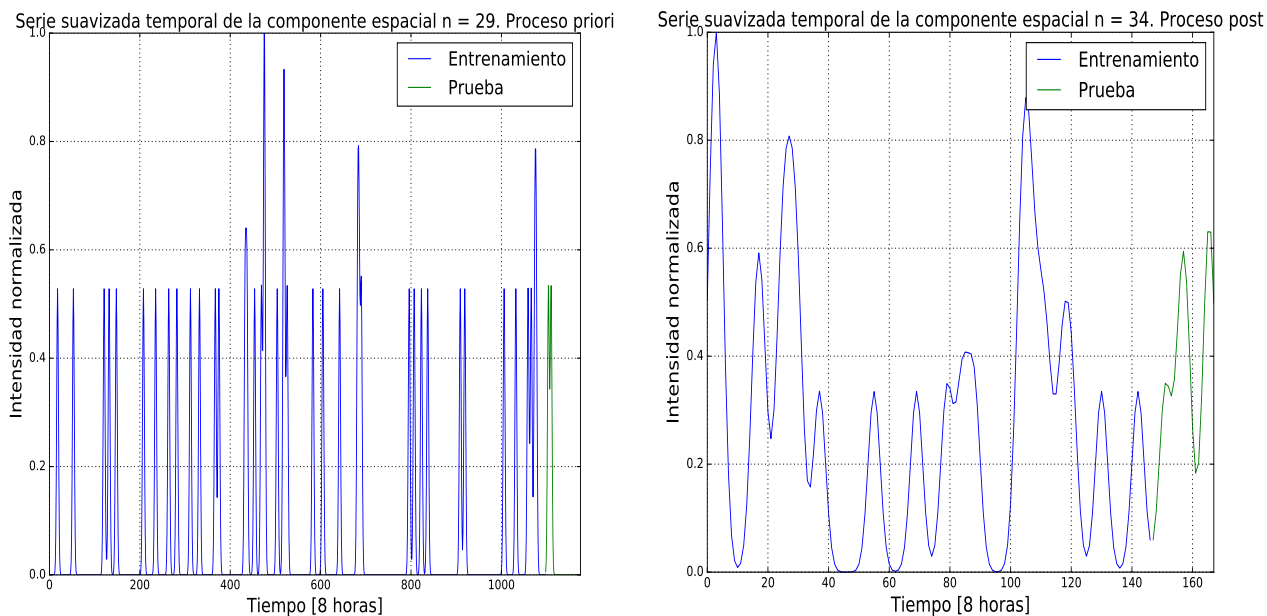
(a) Histograma de la cantidad de eventos criminales en el ajuste de GMM *priori* espacial.

(b) Histograma de la cantidad de eventos criminales en el ajuste de GMM *posteriori* espacial.

Figura 3.3: Histogramas de asignación de delitos para cada una de las etapas de la caracterización espacial.

Una vez asignados los delitos a cada componente espacial, se procedió a generar los arreglos temporales discretos para las fases *priori* y *posteriori*. Cada arreglo fue suavizado por medio de KDE, con los cuales se ejecutaron los entrenamientos de cada modelo temporal. La Figura 3.4 muestra las series temporales suavizadas del componente espacial número 29 de la fase *priori* y 34 de la fase *posteriori*. Cada una de las señales fueron divididas en dos conjuntos, siendo utilizadas de igual manera en el modelo temporal usando técnicas de transformación mediante Kernels Periódicos Gaussianos (modelo KPG) y el modelo temporal usando técnicas de Transformada Wavelet (modelo WV). Para la fase *priori*, la señal de entrenamiento abarcó

el primer año de datos, mientras que la señal de prueba tomó los restantes meses de datos. En la fase *posteriori* se tuvieron 7 semanas en la señal de entrenamiento y tan solo 1 semana de prueba. Las señales generadas tuvieron un tiempo de muestreo igual a 8 horas.



(a) STS de la componente espacial número 29 para la fase *priori*.

(b) STS de la componente espacial número 34 para la fase *posteriori*

Figura 3.4: Señales de tiempo suavizadas (STS) para dos casos particulares después del proceso de entrenamiento espacial.

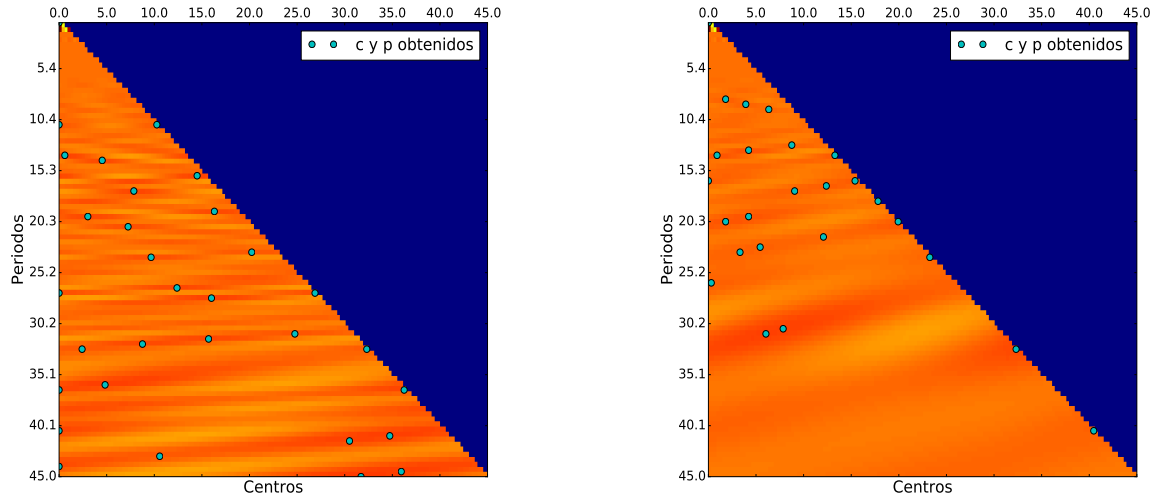
### 3.3.1. Modelo KPG

Durante el desarrollo del modelo KPG se necesitó crear la matriz de correlación, la cual fue explicada en la sección 2.5.2.1. Para ello, se establecieron el arreglo con candidatos a condiciones óptimas del parámetro  $p_1^r$  ( $A_p^r$ ) y el arreglo con candidatos a condiciones óptimas del parámetro  $c_1^r$  ( $A_c^r$ ) con la misma estructura para toda componente espacial  $r \in [1 \dots N_{kG}]$  tanto para la fase *priori* como *posteriori*.

El arreglo  $A_p^r$  permitió estudiar los periodos presentes en cada señal temporal suavizada por medio de KDE (STS), partiendo con un valor igual a 0.5, aumentando cada paso en 0.5 y terminando con un valor igual a 45, teniendo al final 90 elementos a examinar. Los valores entregados estuvieron relacionados con el tiempo de muestreo de la señal temporal, que en este caso fue de 8 horas. El primer elemento del arreglo generó un KPG con un periodo igual a 4 horas ( $0,5 \cdot 8\text{Horas} = 4\text{Horas}$ ) y el último elemento analizó un KPG con periodo igual a 360 horas, equivalente a 15 semanas.

El arreglo  $A_c^r$  entregó los centros de los KPG a correlacionar. En este caso se fijó la cantidad de elementos del arreglo a 90, partiendo en 0 y finalizando con 45. Este arreglo también fue afectado por el tiempo de muestreo de la señal. Por ejemplo, el primer elemento del arreglo denota un desfase igual a 0 horas, mientras que el último elemento permitió analizar una señal desfasada en 360 horas.

La Figura 3.5 entregó la matriz de correlación de la fase *priori* y *posteriori* de la señal de entrenamiento del componente espacial número 29 y 34 respectivamente. En cada matriz se superpusieron los máximos locales presentes, siendo estos los mejores candidatos para parametrizar el conjunto de KPG que construyeron el modelo de predicción temporal. La matriz de correlación generada tuvo una estructura triangular inferior dado que la información entregada por centros con mayor valor respecto a su periodo generaba redundancia, entregando la misma respuesta para un caso presente en alguna combinación dentro del triángulo inferior de la matriz.



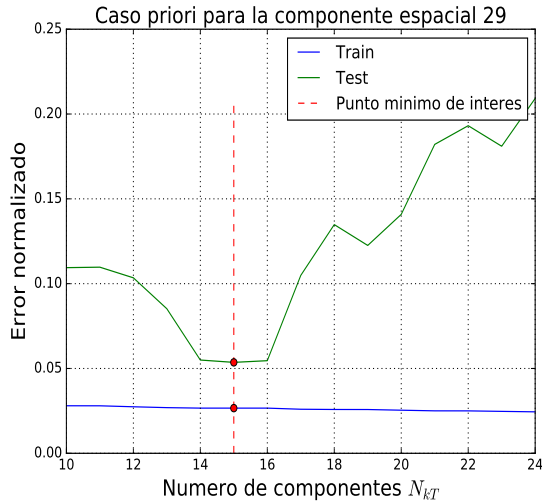
(a) Matriz de correlación para el proceso *priori* de STS de la componente espacial 29.

(b) Matriz de correlación para el proceso *posteriori* de STS de la componente espacial 34.

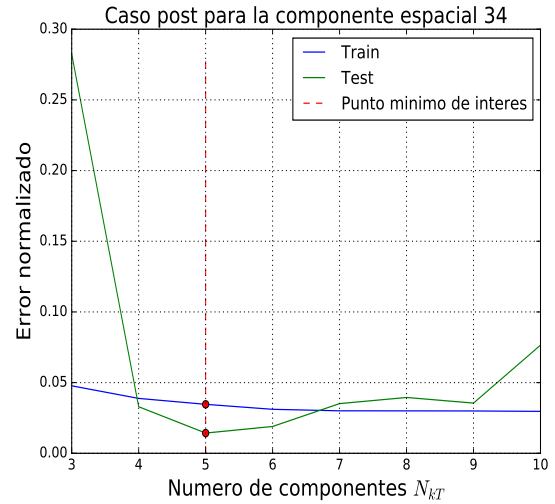
Figura 3.5: Matrices de correlación generada usando  $A_p^r$  y  $A_c^r$  base. Los puntos celestes corresponden a los máximos locales encontrados usando una búsqueda basada en procesamiento de imágenes de ventana 5x5.

El siguiente paso fue encontrar el óptimo número de funciones de Kernels Periódicos Gaussianos ( $K_{pT}$ ) asociados a una componente espacial de la GMM ( $N_{kT}$ ) a partir de la evaluación de los modelo KPG entrenados con cada uno de los máximos locales de la matriz de correlación. Los elementos fueron ordenados de manera descendente con tal de entregar mayor prioridad al máximo local de mayor valor de la matriz. La Figura 3.6 entrega las curvas de error asociado al estudio *priori* y *posteriori*. La línea punteada de color rojo marcó el candidato  $N_{kT}$  que minimizó la suma del MSE de entrenamiento y prueba de la STS en estudio.

La Figura 3.7 compara el modelo en su etapa de entrenamiento y prueba en la fase *priori*, mientras que la Figura 3.8 lo hace en la fase *posteriori*. En el caso de la fase *priori*, la Figura 3.7a entregó una baja tasa de aciertos al momento de ser comparada con la señal original de entrenamiento, perjudicando la evaluación del modelo en la etapa de prueba, dando como resultado la Figura 3.7b. Aun cuando el modelo temporal, al ser evaluado en el soporte de entrenamiento, entregó una señal que identificó gran parte de los *peaks* de mayor intensidad, los parámetros aplicados no fueron lo suficientemente buenos para entregar un buen desempeño en la etapa de prueba.



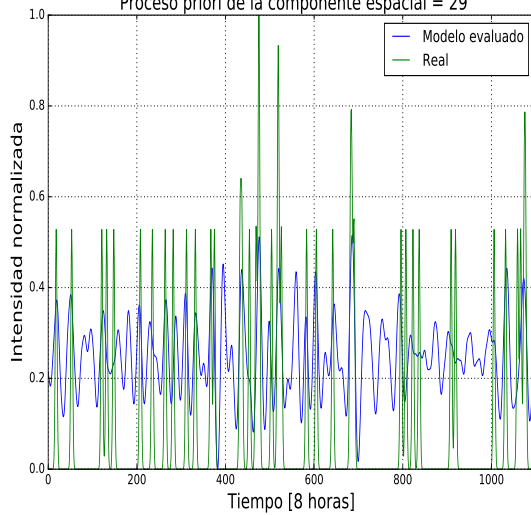
(a) Curvas de error entrenamiento y prueba para proceso *priori*.



(b) Curvas de error entrenamiento y prueba para proceso *posteriori*.

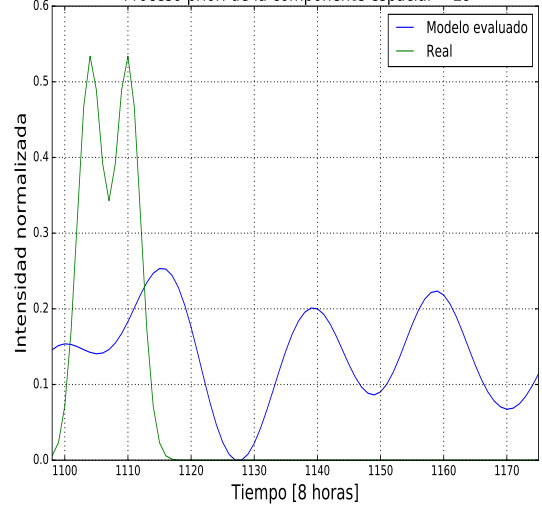
Figura 3.6: Búsqueda del mejor candidato  $N_{kG}$  usando la información de los errores asociados a cada una de las fase de evaluación.

Curvas de comparación de entrenamiento para modelo optimizado con  $N_{kT} = 15$ . Proceso *priori* de la componente espacial = 29



(a) Fase de entrenamiento.

Curvas de comparación de prueba para modelo optimizado con  $N_{kT} = 15$ . Proceso *priori* de la componente espacial = 29

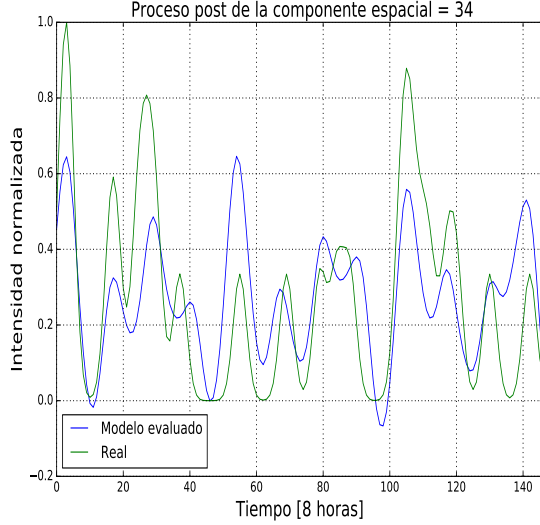


(b) Fase de prueba.

Figura 3.7: Curvas de comparación para el proceso *priori* de la STS del kernel espacial 29.

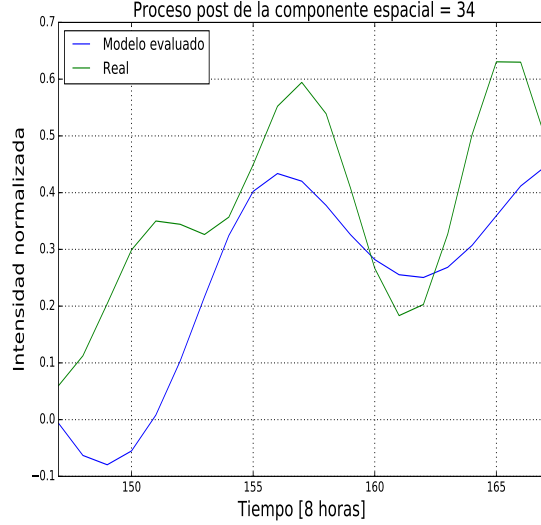
Por otra parte, la señal de entrenamiento de la fase *posteriori* presentó un mejor comportamiento tal como se observó en la Figura 3.8a. Esto permitió un mejor desempeño de la etapa de prueba, tal como se mostró en la Figura 3.8, donde el modelo evaluado sigue la tendencia de la señal original.

Curvas de comparación de entrenamiento para modelo optimizado con  $N_{kT} = 5$ .



(a) Fase de entrenamiento.

Curvas de comparación de prueba para modelo optimizado con  $N_{kT} = 5$ .



(b) Fase de prueba.

Figura 3.8: Curvas de comparación para el proceso *posteriori* de la STS del kernel espacial 34.

### 3.3.2. Modelo WV

El modelo WV entregó una respuesta heurística al momento de generar una señal de predicción, basándose en la reconstrucción de las bandas de mayor energía dadas por la Matriz de Energía Wavelet (MEW) de la señal temporal suavizada según el arreglo Energía Acumulada por Escala (EAE). Esta heurística tomó como supuesto que los delitos tenían un patrón de recurrencia que tendía a ser muy parecido al comportamiento analizado durante una ventana de tiempo mayor a una semana. Tanto la fase *priori* como *posteriori* tuvieron una señal de predicción propia, las cuales fueron descritas en la Sección 2.5.3.2.

El caso base de estudio utilizó TW de función base *DOG* debido a que su diseño estaba orientado a la descomposición de señales con pulsos repentinos o aislados. La unidad de medida temporal que definió a la función base fue el día, la cual estuvo relacionado con el tipo de estudio que se realizó, siendo de interés conocer la recurrencia diaria de los delitos en un componente espacial arbitrario. Dado que el tiempo de muestreo de STS fue de 8 horas, el valor unitario de la TW quedó definido por la expresión  $\delta t = \frac{8}{24} = \frac{1}{3}$ [día]. Otro parámetro de interés fue la resolución del arreglo de escalas aplicado ( $\delta j$ ), la cual permitió calcular el largo del arreglo de escalas ( $J$ ) (ver ecuación 1.36) y así producir el conjunto  $S = \{s_j\}_{j=0,\dots,J}$ . En este caso se dejó  $\delta j = \frac{1}{40}$ , valor que entrega una resolución apropiada para el estudio diario en base a MEW. Fue posible escoger una fracción aún más pequeña que la entregada, pero el tiempo de cálculo fue incrementando a medida que iba disminuyendo su valor, sin tener una ganancia de resolución significativa. Para este caso base se fijaron los parámetros  $m$  y  $bp$  en los valores de 6 y 3 respectivamente tanto para la fase *priori* como *posteriori*. Finalmente se analiza una ventana de 120 y 60 días para la fase *priori* y *posteriori* respectivamente, indicando que la posibilidad de repetirse un delito supera una tasa de 4 meses para el caso *priori* y de 2 meses para el caso *posteriori*.

Al aplicar la TW a la STS *priori* del componente espacial número 29 se tuvo como respuesta



la Figura 3.9 entregando información de MEW y de sus bandas de mayor acumulación de energía. Basado en la información entregada por las bandas de mayor acumulación de energía se pudo completar la Tabla 3.3, la cual entregó los diez primeros periodos que representaron de mejor manera la señal estudiada: En este caso el estudio del componente espacial número 29 indicó que los delitos tuvieron una fuerte tasa al ser repetidos cada 4 días y 12 horas. El mismo proceso se realizó en la STS *posteriori* de la componente espacial número 10, obteniendo como respuesta la Figura 3.10. Según lo mostrado por las bandas de energía acumulada fue posible generar la Tabla 3.4 entregando los periodos de mayor importancia de la STS *posteriori*. En este caso, el estudio mostró que los delitos tuvieron una fuerte presencia de repetición superiores a los 20 días, obteniendo una respuesta totalmente diferente a la fase *priori*, donde la tendencia de repetición delictual fue cercana a los 10 días.

Periodos con mayor energía ordenados de izquierda a derecha [días]									
4.5	7.5	13.6	10.5	16.6	19.7	52.9	22.6	46.8	37.8

Tabla 3.3: Periodos seleccionados utilizando el análisis Wavelet con parámetros base para la fase *priori*.

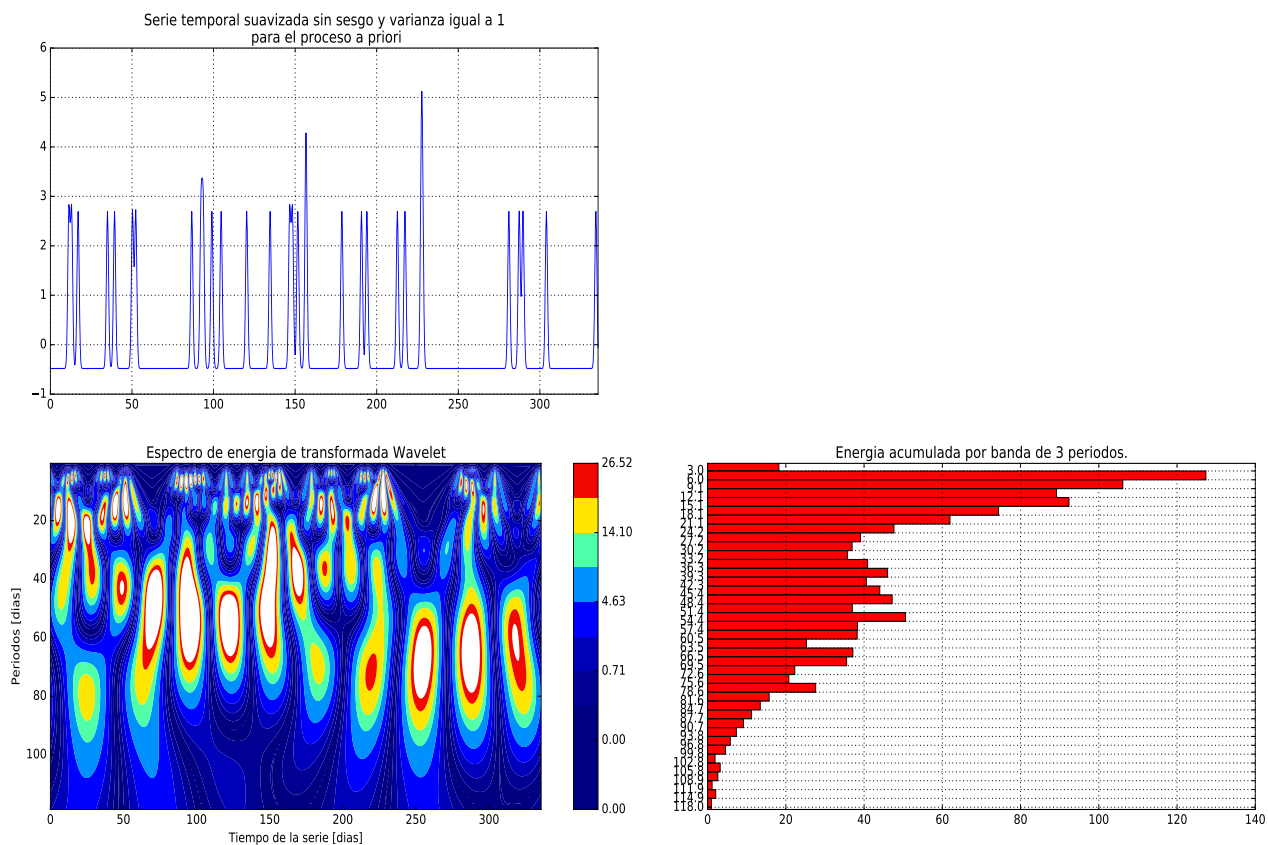


Figura 3.9: Análisis Wavelet de la señal temporal suavizada de la componente espacial 29.

Periodos con mayor energía ordenados desde izquierda a derecha [días]									
26.3	4.6	29.4	23.2	7.7	32.5	10.8	20.1	13.9	17

Tabla 3.4: Periodos seleccionados utilizando el análisis Wavelet con parámetros base para la fase *posteriori*.

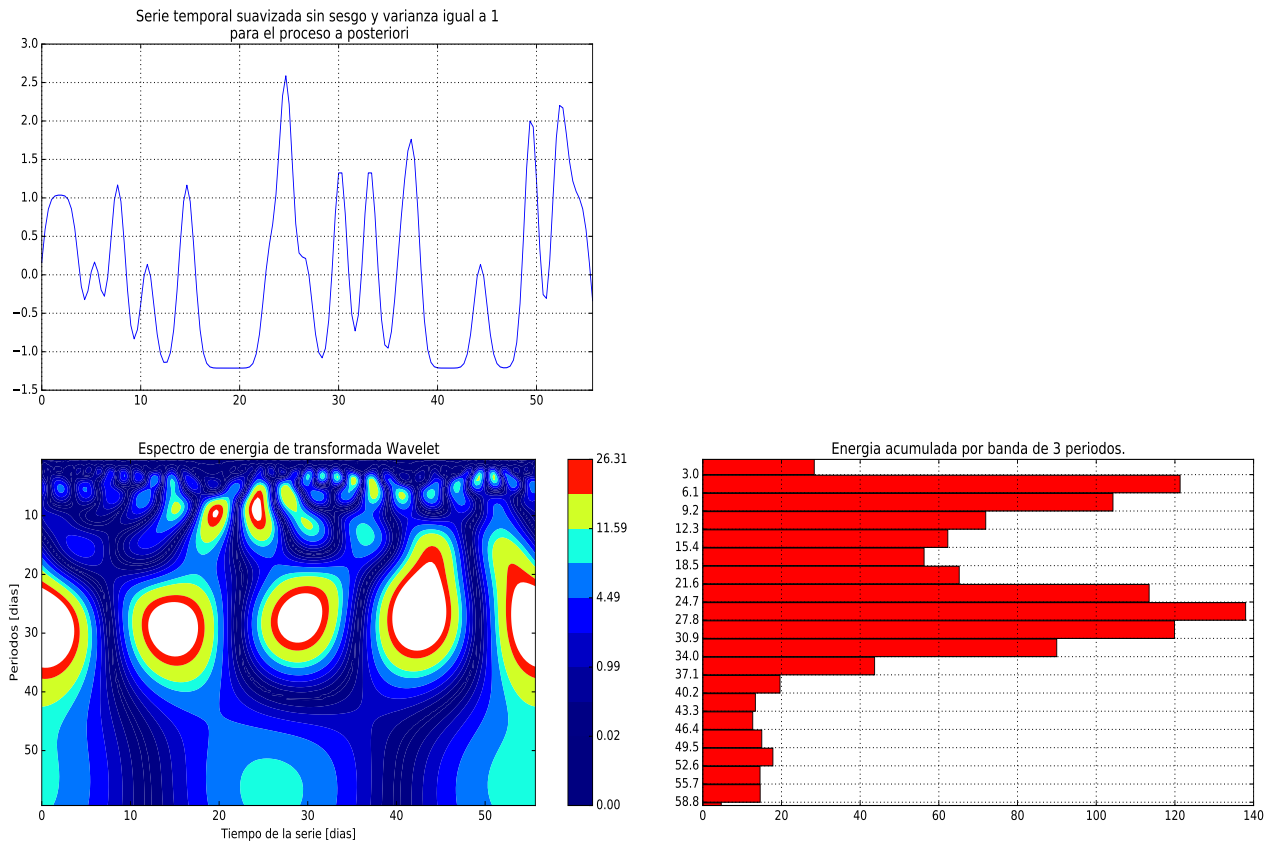


Figura 3.10: Análisis Wavelet de la señal temporal suavizada de la componente espacial 10.

La señal de predicción se generó por medio de la Anti-Transformada Wavelet (ATW) en las escalas representativas con mayor concentración energética. En este caso la Tabla 3.3 y 3.4 describieron los periodos de mayor relevancia de cada señal. Cada periodo estuvo asociado a un arreglo de energía dado por la fila correspondiente a MEW (recordando que existe una proporción asociada entre periodo-escala) y se aplicó a este set de componentes la ATW. Así para cada uno de los periodos de mayor energía, generando en este caso 10 señales diferentes.

Como se describió en el Capítulo Metodología en la Sección 2.5.3.2, cada proceso tuvo una forma diferente de generar una señal de ATW. En el caso *priori* se realizó la suma normalizada de las 10 señales generadas por la ATW de los periodos de mayor importancia. La Figura 3.12a entregó la respuesta de predicción del caso *priori*. Por otra parte la señal de predicción del caso *posteriori* fue generada a partir de la segmentación semanal de la suma normalizada de las 10 señales obtenidas por la ATW. Cada semana tuvo un factor de olvido asociado, donde su potencia fue aumentando a medida que se alejaba cada vez más de la fecha de predicción tal como se ve en la Figura 3.11 donde la semana 0 fue la más lejana a la fecha actual y la semana 8 fue justamente la anterior a dicha fecha. La suma normalizada

de todas estas señales ponderadas por su factor de olvido generó la señal de predicción para el caso *posteriori*, tal como se muestra en la Figura 3.12b.

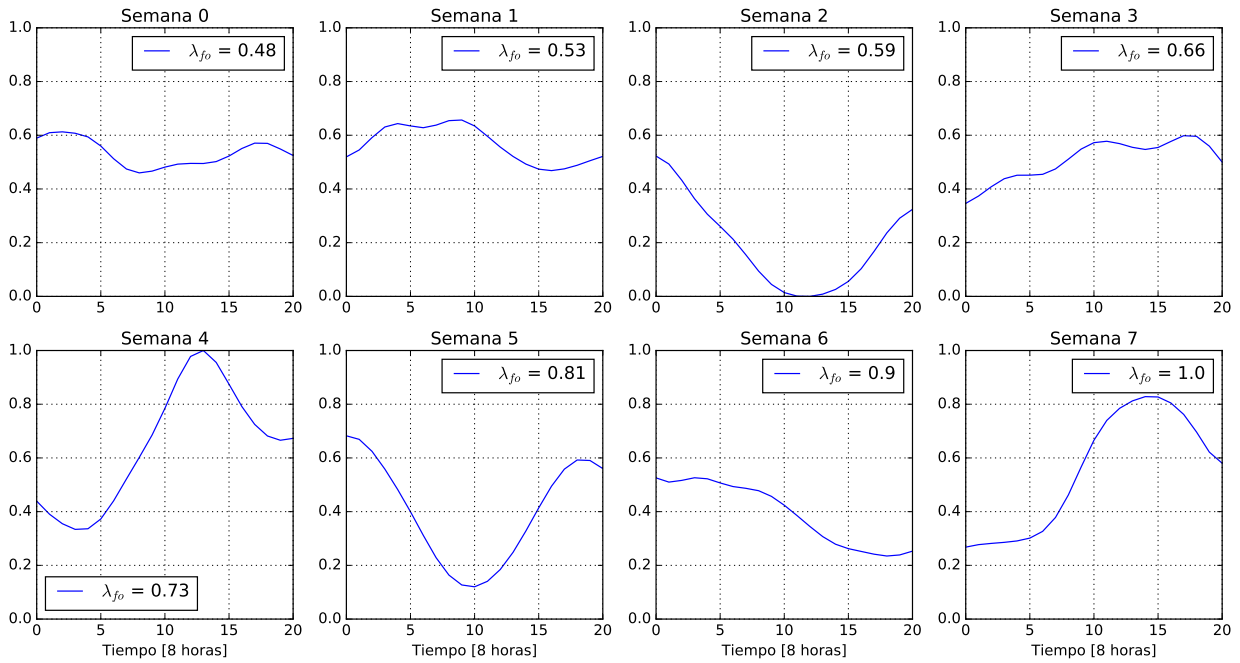
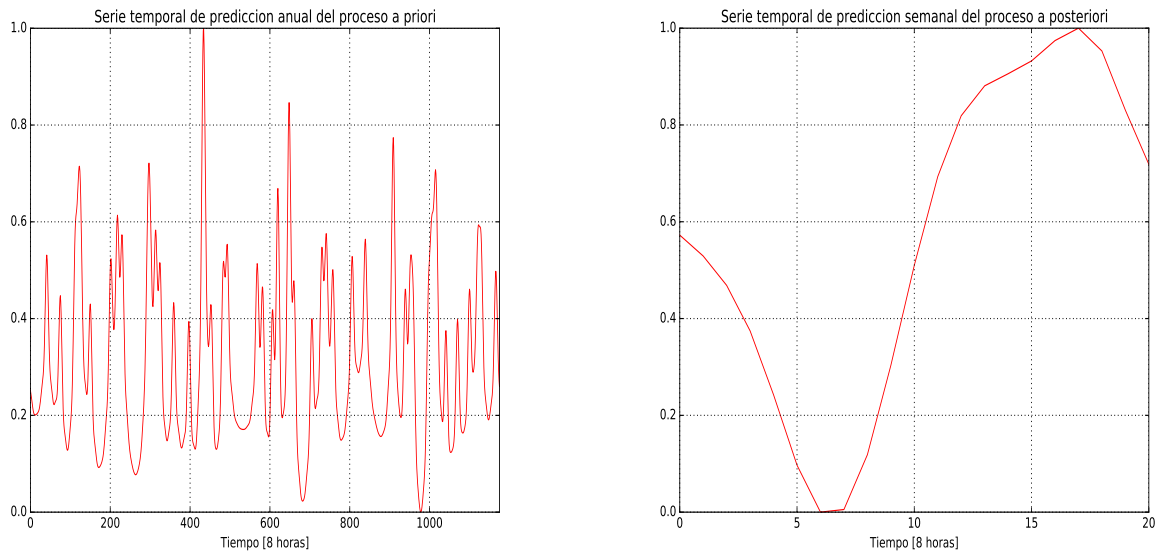


Figura 3.11: Segmentación semanal de la señal reconstruida por medio de ATW y con su factor de olvido asociado a cada una de las curvas.



(a) Señal de predicción para el proceso *priori* de la componente espacial 29. La señal tiene duración de un año.

(b) Señal de predicción para el proceso *posterior* de la componente espacial 10. La señal tiene duración de una semana.

Figura 3.12: Señales de predicción para cada uno de los procesos usando el modelo temporal usando técnicas de Transformada Wavelet.

### 3.3.3. Predicción espacio-temporal del riesgo criminal

Una vez que fueron desarrollados los modelos temporales para cada uno de las componentes espaciales, fue posible obtener una respuesta espacio temporal del riesgo criminal haciendo uso de los modelos entregados por la ecuación 2.10 o 2.12. Cabe mencionar que para este estudio base, los dos modelos presentaron las mismas componentes espaciales, con centros y valores de covarianza fijos, pero las respuestas fueron diferentes debido a que los métodos aplicados en cada uno de los modelos temporales cumplían premisas diferentes. Sin embargo fueron respuestas estandarizadas dado que ocuparon el mismo conjunto de señales de entrenamiento y prueba, lo cual permitió comparar el desempeño de cada uno de los modelos generados según la fecha de predicción y de la comuna.

Para el caso base estudiado, la predicción del modelo de caracterización espacio temporal del riesgo criminal usando el modelo KPG y modelo WV, permitió evaluar tres *slots* de 8 horas definidos desde la fecha entregada en la Tabla 3.2. El primer *slot* predijo la actividad delictual del turno de mañana entre las 07:01 horas hasta las 15:00 de la fecha de predicción. El segundo *slot* predijo la actividad delictual del turno de la tarde dado por las 15:01 hasta las 23:00 horas. Finalmente el último *slot* predijo la actividad delictual nocturna entre las 23:01 hasta las 07:00 horas del día siguiente a la fecha de predicción.

La Figura 3.13 muestra la superficie de riesgo criminal predicha a partir del modelo Wavelet dado por la ecuación 2.10 para el primer *slot*. El eje  $z$  muestra la intensidad del riesgo para un punto  $(x, y)$ .

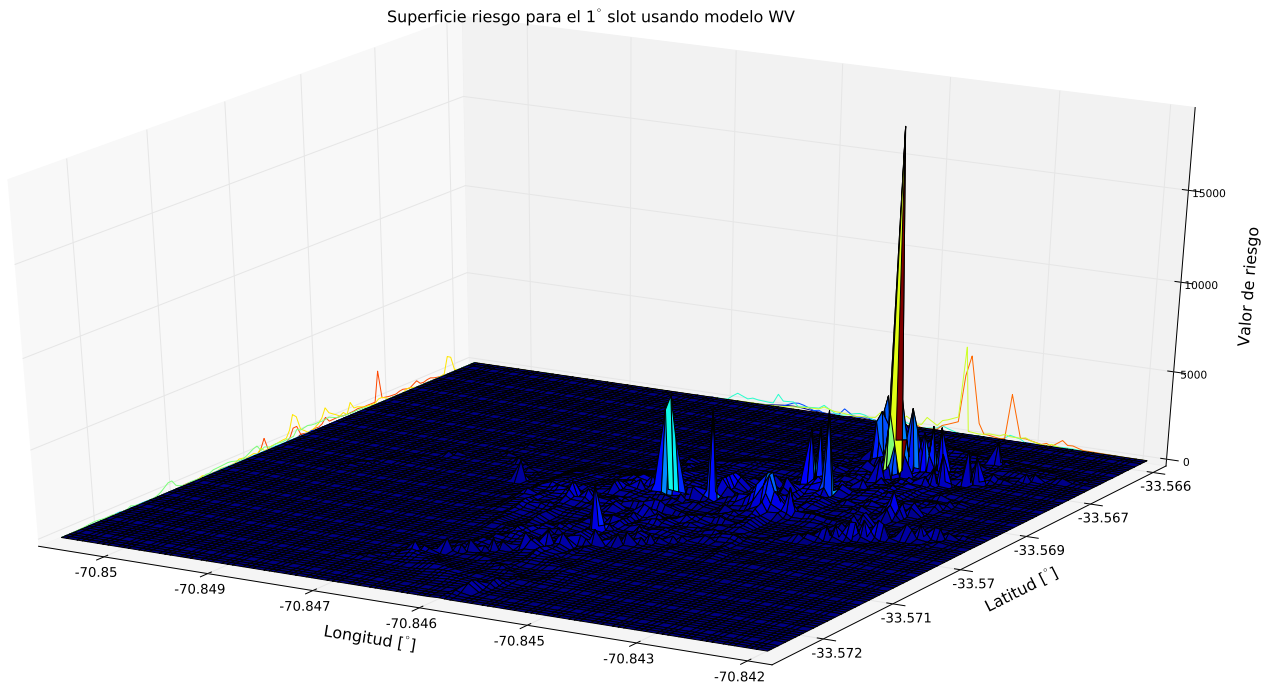


Figura 3.13: Superficie de riesgo en formato 3D para el primer *slot* de predicción del 04-06-2016 de la comuna de Maipú.

Como último paso se procedió a calcular la integral aproximada de la curva de riesgo criminal en cada uno de las grillas asociadas a Maipú, las cuales proporcionaron una medida estándar y uniforme en cada una de las comunas integradas en la base de datos. Además, permitieron generar una respuesta visual acorde a las exigencias dadas por Carabineros de

Chile. La integral aproximada constó de la suma 25 puntos uniformemente distribuidos en sus dos dimensiones definidos dentro de cada grilla, los cuales fueron evaluados con el modelo de caracterización de riesgo criminal anteriormente generado. Gracias a las herramientas de visualización entregadas por QGIS® fue posible integrar la información de calles y geografía de la zona de estudio. Además fue posible aplicar una segmentación sobre los valores entregados por el modelo en la totalidad de las grillas de la comuna, clasificando en cinco niveles diferentes la cantidad de riesgo asociado a cada grillas, asociando de manera descendente los nombres de **Muy Alto**, **Alto**, **Medio**, **Bajo** y **Muy Bajo** riesgo criminal. Dicha segmentación fue procesada por algoritmos definidos como quiebres naturales o *Jenks*, muy parecido al proceso de *K-means*, donde se maximiza la varianza entre los diferentes grupos pero se minimiza la varianza entre los datos de la misma clase.

La Figura 3.14 entregó el resultado de integrar la evaluación del modelo mostrando las grillas cuya segmentación quedó categorizada como **Muy Alto** riesgo criminal usando el motor visual de QGIS®. En este caso se pudo apreciar que 3 de los 4 delitos fueron acertados con las grillas activadas. Esto indicó un buen desempeño del modelo, debido a que con reducida cantidad de grillas (10% aproximadamente del total área entregada por Carabineros de Chile según la Figura 3.1) fue posible acertar más de la mitad de los delitos ocurridos en el primer turno.

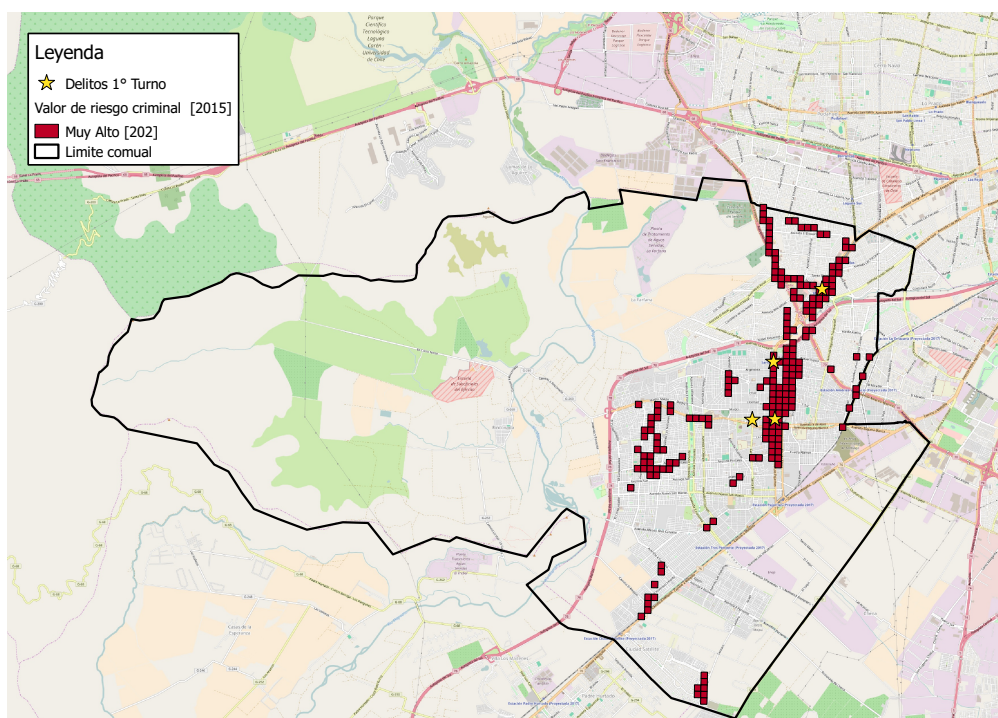


Figura 3.14: Riesgo criminal entregado para el primer *slot* de predicción del 04-06-2016 de la comuna de Maipú usando modelo Wavelet. Las grillas entregadas corresponden al 10% del total de grillas definidas en zona urbana (202 de 2015 grillas)

Sin embargo, fue importante considerar como fue el desempeño del modelo en los diferentes porcentaje de riesgo criminal. Para ello se utilizó la información entregada por las curvas  $HRP_{\theta}$  y  $TIP_{\theta}$ , mostradas en la Figura 3.15, las cuales permitieron visualizar gráficamente la tasa de aciertos dentro de dicha área ( $TIP_{\theta}$ ) según el porcentaje de área activa ( $HRP$ ). El caso analizado en la Figura 3.14 corresponde justamente a un punto de la curva de color

azul, denotando que con el 10% de grillas ( $HRP_{\theta} = 0.1$ ) fue posible acertar el 75% de los delitos ocurridos en primer *slot* ( $TIP_{\theta} = 0.75$ ). Las tres curvas indicaron el comportamiento de los tres turnos definidos en el día (3 *slots*). Además se entregó en la leyenda de cada curva el valor AUC asociado a cada *slot*, recordando que entre más cercano sea su valor a 1, mejor es el desempeño del modelo.

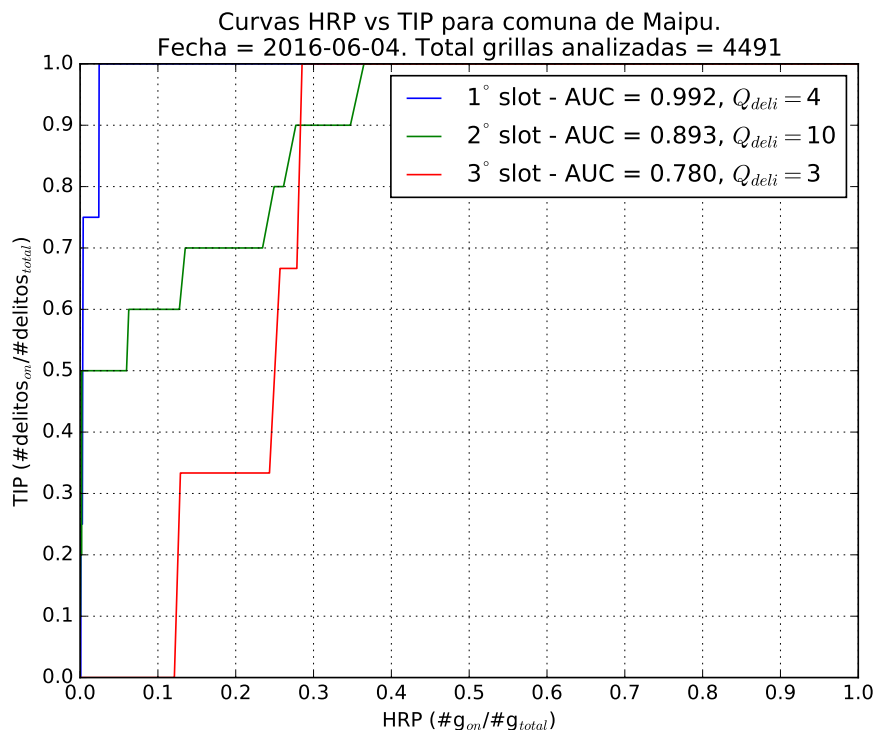


Figura 3.15: Curvas de desempeño generadas para el caso base, en los tres *slots* definidos para la fecha de predicción

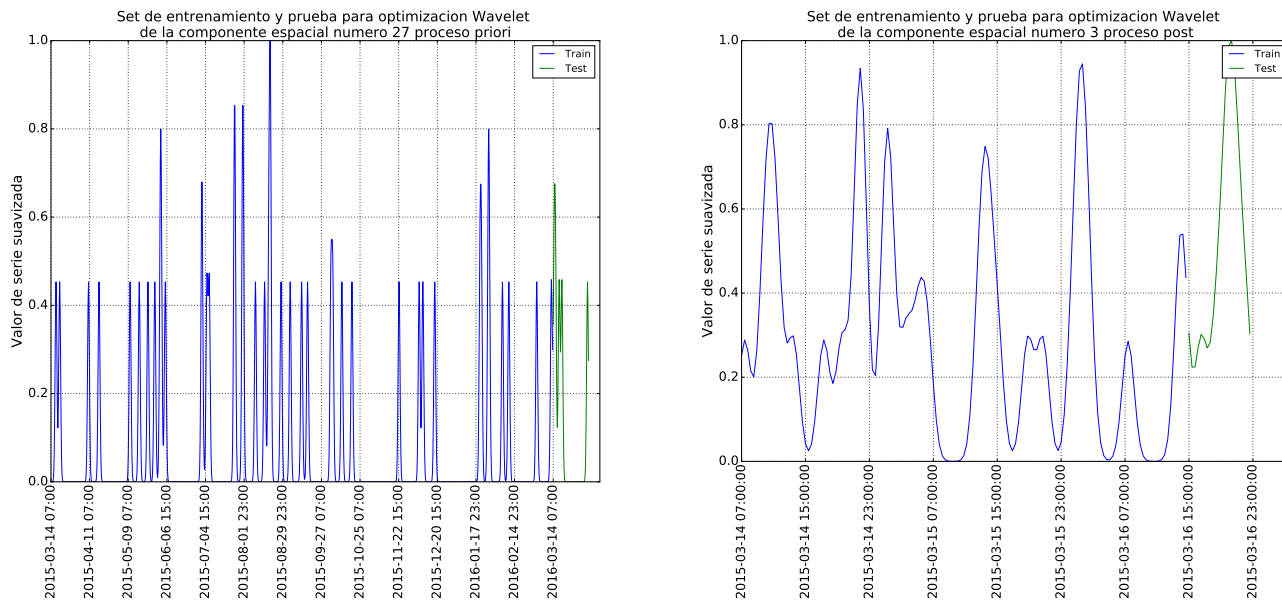
### 3.4. Optimización de modelos temporales

En esta Sección se dieron a conocer diferentes técnicas que permitieron optimizar los parámetros internos de los modelos temporales descritos en este trabajo de tesis con el objetivo de mejorar el desempeño del modelo de caracterización espacio-temporal del riesgo criminal usando como caso de estudio los delitos ocurridos durante el mes de junio del 2016 en la comuna de Maipú.

Dado que se tuvieron dos modelos temporales diferentes, se buscó que cada una de las respuestas entregadas por el modelo espacio-temporal del riesgo criminal tuviera una salida comparable. Por ello se utilizó en cada día del mes de junio un entrenamiento espacial previo, donde se pudieron definir los distintos componentes espaciales de la GMM, para luego asignar los delitos que caen dentro de cada uno de ellos y así generar la STS de cada componente, creando de esta manera una base común para el modelo KPG y el modelo WV. Esto permitió que el desempeño de los modelos espacio-temporal del riesgo criminal fueran comparable, incluso ocupando diferentes técnicas de predicción temporal. Se generó diariamente

una cantidad aproximada de 210 componentes espaciales en la fase *priori*, de los cuales, en promedio, 92 cumplieron con tener 30 delitos asociados. Lo mismo ocurrió con la fase *posteriori*, de la cual se obtuvo en promedio una cantidad de 15 componentes con al menos 30 delitos asociados por día.

La Figura 3.16 entregó un ejemplo de las series de entrenamiento y prueba tanto para la fase *priori* y *posteriori* del día 17 de junio del 2016. La señal de prueba del caso *priori* tuvo la misma fecha de inicio que la señal de entrenamiento pero con un año de desfase. En el caso *posteriori* se denotaron por líneas verticales segmentadas las divisiones semanales, siendo la señal de prueba la última semana.



(a) Señales de entrenamiento y prueba de la fase *priori*.

(b) Señales de entrenamiento y prueba de la fase *posteriori*.

Figura 3.16: Ejemplo de separación temporal del conjunto de entrenamiento y prueba en los diferentes procesos para optimización Wavelet

### 3.4.1. Optimización modelo KPG

El objetivo de esta sección fue entregar el desempeño diferenciado del modelo temporal KPG usando diferentes arreglos  $A_c^r$  y  $A_p^r$  en cada componente espacial de la GMM entrenada. Sea un conjunto de arreglos de estudio definidos por  $\mathcal{A} = [\{A_{p_i}^r, A_{c_i}^r\}]$ , con  $i \in [0, 1, 2, 3]$ . A partir de cada par generado por  $\mathcal{A}$  fue posible construir una matriz de correlación de diferentes dimensiones y características, permitiendo explorar distintas condiciones iniciales para el entrenamiento del modelo KPG en cada una de las componentes espaciales, tal como se mostró en la Sección 3.3.1

Basándose en los datos recolectados durante todo el mes de junio del año 2016 y de las STS definidas en cada componente espacial, se procedió a medir el desempeño del modelo temporal usando los arreglos denotados en la Tabla 3.5.

Sea un ejemplo hipotético donde el entrenamiento espacial del día 17 de junio del 2016 arrojó 200 y 20 componentes espaciales para la fase *priori* y *posteriori* respectivamente. Para

Estudio 0	Estudio 1	Estudio 2	Estudio 3
$A_{p_0}^r$ [0.5 - 45] con paso de 0.5	$A_{p_1}^r$ [0.5 - 45] con paso de 0.1	$A_{p_2}^r$ [0.5 - 90] con paso de 0.5	$A_{p_3}^r$ [0.5 - 90] con paso de 0.1
$A_{c_0}^r$ [0 - 45] con paso de 0.5	$A_{c_1}^r$ [0.5 - 45] con paso de 0.1	$A_{c_2}^r$ [0 - 90] con paso de 0.5	$A_{c_3}^r$ [0 - 90] con paso de 0.1

Tabla 3.5: Arreglos escogidos para el estudio de optimización en el modelo KGP. Cada estudio muestra el tipo de arreglo utilizado, donde el primer y último valor han sido denotados entre corchetes y la diferencia entre un valor con el siguiente ha sido explicado fuera de estos.

cada fase se tendrían dos MSE asociados al proceso de entrenamiento y prueba. Todo esto aplicado en cada componente espacial. La forma de comparar el desempeño de cada modelo generado por un tipo de estudio (ver Tabla 3.5) sería buscando el mejor comportamiento de los MSE generados por el proceso de prueba en cada uno de los componentes espaciales. En este caso se tendrían 200 MSE *priori* y 20 MSE *posteriori* del proceso prueba. Dada la gran cantidad de valores entregados, se entregaría dicha información por medio de gráficos que darían a conocer la media, desviación, máximos y mínimos de cada una de las fechas de estudio.

Los arreglos de la Tabla 3.5 caracterizaron, por distintos medios, las posibles condiciones iniciales aplicables a los periodos y centros de cada KPG del modelo temporal a entrenar. El caso de estudio 0 permitió buscar periodos y centros espaciados con un tiempo equivalente a 4 horas, teniendo como cota inferior de búsqueda en el arreglo de periodos ( $A_{p_0}^r$ ) un valor igual a 4 horas y un máximo valor dado por 360 horas, equivalente a 15 días, mientras que en el arreglo de centros ( $A_{c_0}^r$ ) se tuvo una búsqueda acotada inferiormente por un valor igual a 0, indicando que KPG no tiene desplazamiento alguno respecto al origen, y una cota superior igual a 360 horas. El caso de estudio 0 fue aplicado anteriormente en el caso base de la Sección 3.3.1. Los demás casos de estudios tuvieron variaciones tanto en sus límites como en su espaciado, los cuales fueron explicados en los siguientes puntos:

- Caso estudio 1: Mantuvo los límites inferior y superior establecidos en el caso de estudio 0, pero su espaciado fue más fino tanto para el arreglo de periodos como de centros, tomando valores equivalentes a 48 minutos en cada paso, permitiendo que el espectro temporal de búsqueda fuera más amplio y considere valores anteriormente no accesibles por el caso de estudio 0.
- Caso estudio 2: Mantuvo el espaciado establecido en el caso de estudio 0, pero el límite superior del arreglo de periodos y centros fue ampliado al doble de su valor original (30 días).
- Caso estudio 3: Mezcló las modificaciones dadas por el caso estudio 1 y 2, obteniendo los arreglos de gran longitud y con resolución temporal mucho más fina comparados con los arreglos del caso de estudio 0, 1 y 2.

Las Figuras B.1, B.2, B.3 y B.4, ubicadas en el Anexo B, dieron a conocer el valor promedio, desviación estándar, mínimo y máximo de los MSE del conjunto de componentes espaciales usando las STS del proceso de prueba en las distintas fechas del mes de junio, tanto en la fase *priori* como *posteriori*. El objetivo de dichas Figuras fue mostrar, de alguna manera, los momentos que podrían definir las distribuciones de los MSE de prueba de cada uno de los casos de estudios realizados. Es importante notar que para cada fecha del mes de junio se generó un modelo diferente, obteniendo un total de 30 modelos para todo el mes en cada uno de los casos de estudio.

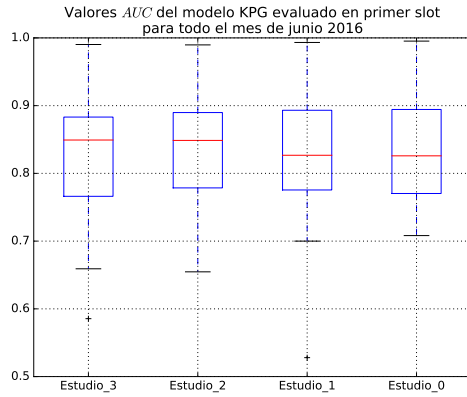


Según lo entregado por las Figuras B.1, B.2, B.3 y B.4 se pudo observar principalmente:

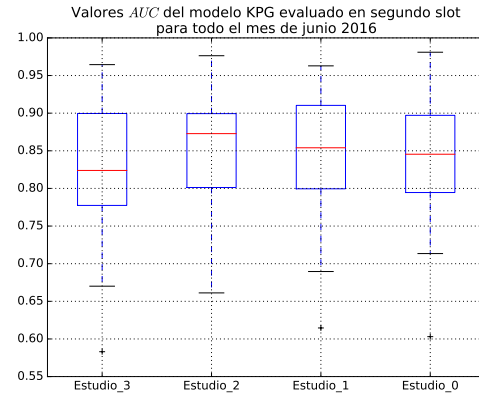
1. La peor respuesta al momento de observar la curva de los máximos MSE de la fase *posteriori* fue entregada por el caso de estudio 1, la cual presentó cambios abruptos entre un día a otro, mostrando poca estabilidad entre el valor promedio y algún componente espacial. Lo sigue el caso de estudio 0, presentando también fuertes perturbaciones, pero con valores más cercanos al origen, seguido por el caso de estudio 3 y dejando como mejor comportamiento para este estadístico al caso de estudio 2, donde la mayoría de los máximos entregados no sobrepasan la unidad a los valores promedio MSE diarios.
2. En la fase *priori* se observó que la señal generada por los MSE máximos con peor comportamiento fue en el caso de estudio 0, seguido de igual manera por el caso de estudio 1, pero con una menor amplitud entre sus máximos y la media diaria generada. También se pudo observar que el caso de estudio 2 presentó nuevamente el mejor comportamiento, obteniendo valores máximos bastante acotados y con una baja tasa de cambio de un día a otro.
3. Respecto al comportamiento de los promedios MSE en la fase *priori* se pudo observar una baja variación en cada uno de los casos de estudio. Por otra parte, los promedios de la fase *posteriori* se presentó en todos los casos de estudio una leve variación en alguno de los días del mes de junio.
4. Respecto al comportamiento de la desviación estándar de la fase *posteriori* se apreció que el caso de estudio 0 y 1 tuvo valores que aumentaron a más del doble de un día para el otro. Esto también fue presente en el caso de estudio 3, pero en menos días. El caso de estudio 2 tuvo el comportamiento más estable para este estadístico

Finalmente, en cada modelo espacio-temporal de riesgo criminal generado diariamente en el mes de junio, se procedió a evaluar su desempeño con los delitos de validación del primer, segundo y tercer turno del día siguiente. Esto quiere decir que si un modelo espacio-temporal fue entrenado el 17 de junio entonces su desempeño se validará con los delitos ocurridos entre las 07:00 y 15:00 (primer turno), entre las 15:00 hasta las 23:00 (segundo turno) y entre las 23:00 hasta las 07:00 del día subsiguiente (tercer turno). La evaluación consistió en ver cuantos delitos fueron acertados con diferentes porcentajes de cobertura, reflejados en las curvas  $HRP_{\theta}$  y  $TIP_{\theta}$ , con las cuales se aplicó el valor AUC a cada una de las curvas generadas por turno. La Figura 3.17 entregó el desempeño de cada modelo espacio-temporal generados durante todo el mes de junio, aplicando la métrica AUC en cada caso de estudio.

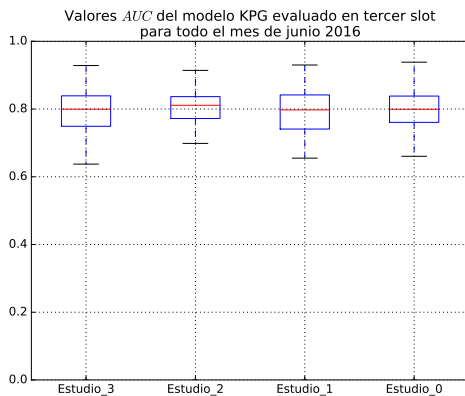
Los gráficos generados permitieron analizar la concentración de valores AUC mensual de cada caso de estudio en cada turno explicado anteriormente usando tres elementos visuales. El primero fue el rectángulo que aglomera el 80 % de los datos generados, el segundo fue dado por la línea horizontal roja al interior del rectángulo entregando información de la mediana de los datos analizados y el último elemento fue las líneas verticales punteadas que marcaron el máximo y mínimo valor entregado por cada estudio. Las marcas en forma de cruz fueron datos que no representaron el comportamiento general observado (*outlayer*). El gráfico de la Figura 3.17d entregó los tiempos de procesamiento de entrenamiento y evaluación de cada modelo generado diariamente.



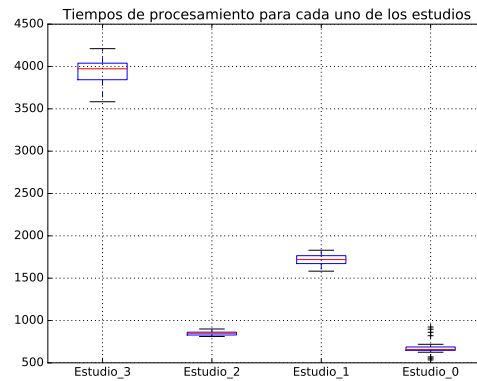
(a) Aglomeración de la métrica AUC para todo el mes de junio del 2016 en el primer *slot* de predicción según el estudio realizado.



(b) Aglomeración de la métrica AUC para todo el mes de junio del 2016 en el segundo *slot* de predicción según el estudio realizado.



(c) Aglomeración de la métrica AUC para todo el mes de junio del 2016 en el tercer *slot* de predicción según el estudio realizado.



(d) Aglomeración de los tiempos arrojados por el modelo KPG al momento de generar 3 *slots* de predicción en todo el mes de junio del 2016 según el estudio realizado.

Figura 3.17: Gráficos de desempeño del modelo basándose en los arreglos de estudio entregados por la tabla 3.5.

De la Figura 3.17 se pudo destacar lo siguiente:

1. El caso de estudio 2 presentó el rectángulo de menor dimensión en cada uno de los turnos estudiados, denotando una baja tasa de desviación en el 80 % de los datos acumulados.
2. Los mejores valores de mínimos y máximos fueron para el caso de estudio 0 en el primer y segundo turno (*slot*), mientras que en el tercer turno, el mejor mínimo fue adjudicado por el caso de estudio 2 y el mejor máximo por el caso de estudio 0.
3. En todos los casos entregados se tiene un valor por sobre 0.6, siendo el tercer turno el con menor desempeño para todos los casos de estudio, donde el caso de estudio 3 tuvo un AUC con valor mínimo de 0.62.
4. Se pudo observar que los mejores desempeños, independiente del caso de estudio, fueron entregados por el primer turno, seguido por el segundo turno.

- Respecto a los tiempos de cálculo, el caso de estudio 3 sobrepasa por lejos a los demás casos de estudio, tardando casi 8 veces más que el caso de estudio 0, siendo este el más rápido en tiempos de procesamiento computacional.

### 3.4.2. Optimización modelo Wavelet

Durante el proceso de optimización del modelo basado en TW se buscó minimizar el MSE de la señal de prueba respecto al modelo generado, variando dos variables de interés: (i) el parámetro interno  $m$  de la función Wavelet escogida para la transformación de la señal y (ii) la banda de periodos  $bp$  que permitió segmentar la información temporal de la Matriz de Energía Wavelet (MEW). Según lo visto en la Sección 2.5.3.3, se realizó el proceso de optimización de TW en base a dos arreglos que contienen los valores de los parámetros  $m$  y  $bp$  candidatos a ser óptimos en este problema. La Tabla 3.6 entregó todos los pares combinados a usar en cada una de las señales temporales de entrenamiento tanto para la fase *priori* como *posteriori*.

		Cantidad de periodos agrupados ( $bp$ )								
		1	2	3	4	5	6	7	8	9
Valor parámetro $m$	4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)	(7,4)	(8,4)	(9,4)
	6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)	(7,6)	(8,6)	(9,6)
	8	(1,8)	(2,8)	(3,8)	(4,8)	(5,8)	(6,8)	(7,8)	(8,8)	(9,8)
	10	(1,10)	(2,10)	(3,10)	(4,10)	(5,10)	(6,10)	(7,10)	(8,10)	(9,10)
	12	(1,12)	(2,12)	(3,12)	(4,12)	(5,12)	(6,12)	(7,12)	(8,12)	(9,12)

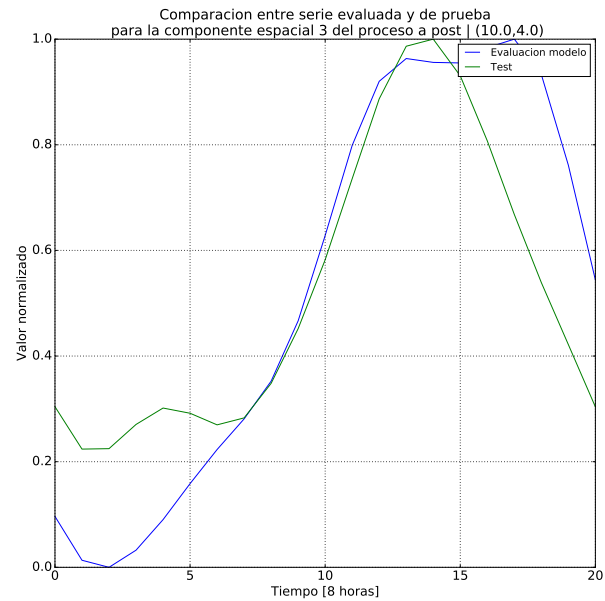
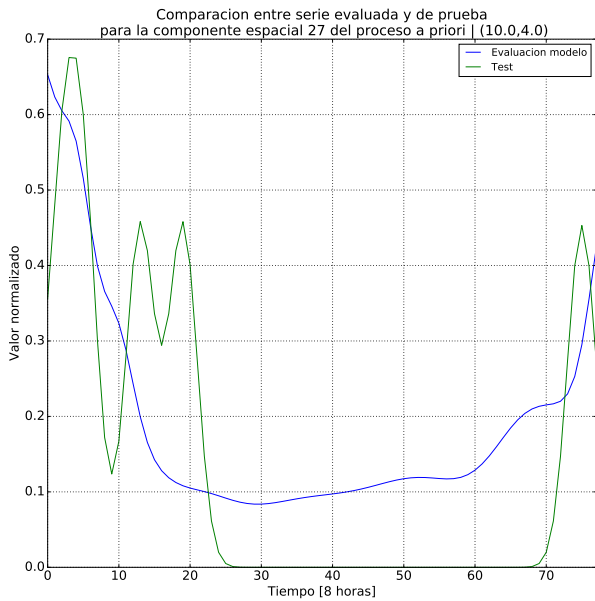
Tabla 3.6: Matriz de pares combinados a analizar a partir de los arreglos de periodos agrupados ( $bp$ ) y valores del parámetro  $m$ .

Los valores del parámetro  $m$  fueron siempre números pares y superiores a 4, mientras que los valores  $bp$  fueron acumulando los periodos de MEW de forma diaria, partiendo con un día de acumulación como banda de periodo, incrementando su valor en una unidad hasta llegar a una banda con una acumulación de nueve días de periodo.

El estudio utilizó el mismo conjunto de GMM entrenado para cada día del mes de junio aplicado en el modelo KPG y explicado al comienzo de la Sección 3.4. En esta etapa de optimización se dejaron fijos los siguientes parámetros: **a)** Función Wavelet base: *DOG*, **b)**  $\delta t = \frac{1}{3}$  [día], **c)**  $\delta j = \frac{1}{40}$ , **d)**  $s_0 = \frac{1}{6}$ , **e)**  $N\delta t_{priori} = 120$  [días] y **f)**  $N\delta t_{posteriori} = 90$  [días].

Realizando todas las Transformadas Wavelet con cada par de parámetros definidos por en la Tabla 3.6 en cada una de las componentes espaciales, se buscó el par  $(m, bp)$  que minimice el MSE de prueba tanto para la fase *priori* como *posteriori*. Este par es único y permitió construir un modelo cuya señal de evaluación en el proceso de prueba fuera mínimo en todo el espacio de búsqueda entregado. Siguiendo con el ejemplo entregado en la Figura 3.16 se generó la Figura 3.18, mostrando el modelo WV en el soporte temporal de prueba en conjunto con la señal de prueba original. En el título se escribió entre paréntesis el par de componentes  $(m, bp)$  que hizo posible el menor MSE según las combinaciones de la Tabla 3.6.

Con tal de observar como se comportó la elección de cada par  $(m, bp)$  se utilizaron las matrices de la Figura 3.19 indicando la cantidad de veces que fue elegida cada par  $(m, bp)$  durante todo el mes de estudio.



(a) Desempeño del modelo generado para la componente espacial 27 respecto la señal prueba de la fase *priori*. El par óptimo encontrado fue  $(m, bp) = (10, 4)$

(b) Desempeño del modelo generado para la componente espacial 3 respecto la señal prueba de la fase *posteriori*. El par óptimo encontrado fue  $(m, bp) = (10, 4)$

Figura 3.18: Ejemplos de curvas de desempeño para diferentes procesos usando el mejor par de componentes  $(m, bp)$  para la señal de un componente espacial específico en el 19 de junio del 2016.



Figura 3.19: Cantidad de componentes espaciales que minimizan el MSE usando el par de parámetros  $(m, bp)$  descritos en la matriz.

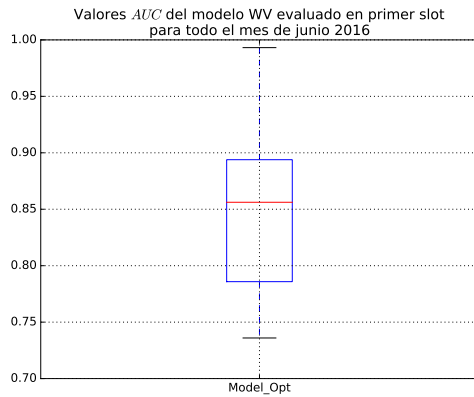
Una de las características que se presentó en la fase *priori* fue la fuerte tendencia a encontrar los mínimos MSE en los pares  $(m, bp)$  que tuvieran a  $m$  con valores iguales a 6 y 10. Al contrario, los pares  $(m, bp)$  que tuvieron como parámetro valores de  $m$  igual a 4, 10 y 12 desarrollaron una menor participación en el proceso de selección, indicando que al momento de utilizar estos valores gran parte de las componentes espaciales arrojaron MSE de prueba elevados. En la fase *posteriori* se presentó un comportamiento totalmente diferente. En este caso se vio una mayor presencia en los pares  $(m, bp)$  con valores  $m$  igual a 10 y 12, y valores  $bp$  superiores a 4 principalmente.

Descrito el proceso de elección y el comportamiento de los pares  $(m, bp)$  que optimizaron la formación del modelo WV para cada una de las STS de prueba de la fase *priori* y *posteriori*, se realizó el estudio de la pseudo-distribución del MSE de prueba de cada día en todo el mes de junio, generando finalmente la Figura C.1 ubicada en el Anexo C. De la Figura C.1 se pudo observar:

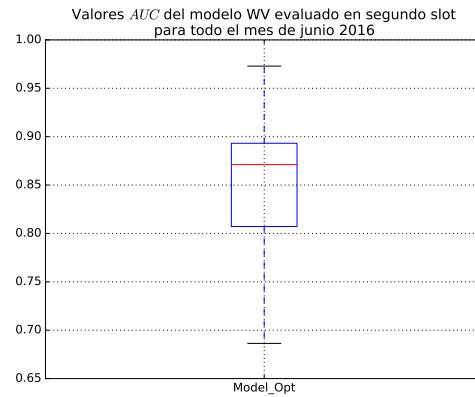
- Que los valores MSE de la fase *priori* no presentaron cambios significativos en gran parte de sus estadísticos, manteniendo una tasa casi constante del valor promedio y desviación estándar de todos los MSE procesados diariamente. En este caso los máximos presentaron un leve cambio entre el 15 y 20 de junio, pero todos muy bien acotados, con variaciones que no mayores a 0.2 unidades.
- La fase *posteriori* presentó diversas fluctuaciones tanto en el estadístico promedio, desviación estándar, mínimo y máximo, siendo este último el que presentó mayor variación comparado con los otros tres nombrados. Se pudo observar una correlación en el comportamiento de los estadísticos con respecto a los valores máximos entregados (si aumenta, los demás estadísticos también aumentan). La desviación estándar fue uno de los casos más estables en la fase *posteriori*, tomando valores cercanos a  $\pm 1,0$  en la mayoría de los días.
- El MSE tendió a entregar una respuesta más uniformada ante señales de prueba de mayor largo, tal como se mostró en el caso *priori*, donde se analizó una señal con 2 meses de prueba, siendo la fase *posteriori* mucho más castigada con estos *peaks* de error dado que se analizó una semana de prueba.

Finalmente, en la Figura 3.20 se entregó los resultados de los modelos espacio-temporales generados en todo el mes de junio con el ajuste paramétrico Wavelet utilizando los datos de validación del primer, segundo y tercer turno del día siguiente a la fecha del modelo entrenado.

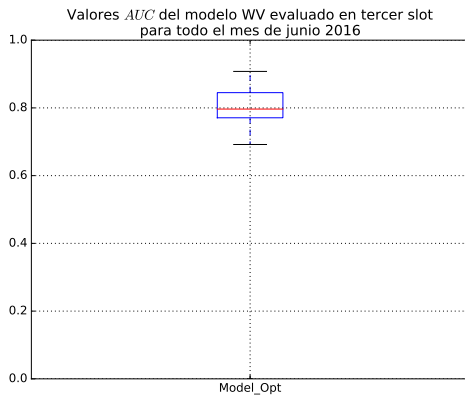
Se pudo observar que el primer y segundo turno tuvieron un comportamiento bastante similar, tanto en la mediana como en el largo del rectángulo que condensa el 80% de los datos, diferenciándose tan solo en los valores AUC máximos y mínimos, obteniendo el primer turno el mejor comportamiento en estos estadísticos. El tercer turno nuevamente presentó el peor desempeño comparado con los turnos anteriores. Por otra parte, el modelo presentó una ventaja significativa en los tiempos de procesamiento, demorando en promedio 11 minutos en entrenar y evaluar cada modelo espacio-temporal de riesgo criminal diariamente.



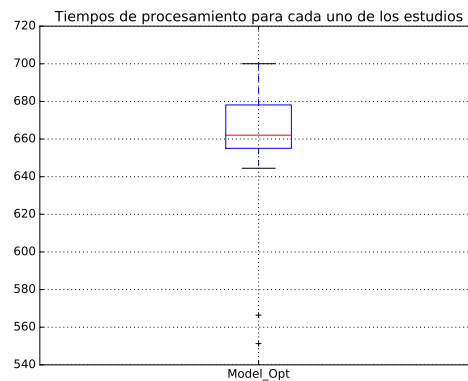
(a) Aglomeración de la métrica AUC para todo el mes de junio del 2016 en el primer *slot* de predicción.



(b) Aglomeración de la métrica AUC para todo el mes de junio del 2016 en el segundo *slot* de predicción.



(c) Aglomeración de la métrica AUC para todo el mes de junio del 2016 en el tercer *slot* de predicción.



(d) Aglomeración de los tiempos arrojados por el modelo WV al momento de generar 3 *slots* de predicción.

Figura 3.20: Gráficos de desempeño del modelo WV basándose el proceso de optimización de parámetros en todo el mes de junio del 2016.

### 3.4.3. Análisis del proceso de optimización.

La finalidad de las Secciones 3.4.1 y 3.4.2 fue poder comparar el desempeño de los dos modelos temporales bajo las mismas condiciones de entrenamiento y prueba en todo el mes de junio del 2016. A cada proceso de optimización se consideró el MSE de prueba como métrica a minimizar y el valor AUC de la etapa de validación como una métrica a maximizar. La primera métrica entregó información de la parte temporal, mientras que la segunda métrica definió la evaluación del modelo espacio-temporal de riesgo criminal. El Anexo E entregó un estudio del comportamiento promedio de curvas  $HRP_{\theta}$  y  $TIP_{\theta}$  con distintos valores de AUC con tal de facilitar el entendimiento del concepto de desempeño que entrega la métrica AUC durante este trabajo de tesis. Por ejemplo, de la Figura E.1 fue posible analizar la curva promedio de 10736 casos con AUC entre 0.90 y 0.95 observando que con un 10% de área encendida (dado por el valor 0.1 en el eje  $HRP_{\theta}$ ) se acertó el 90% de los delitos ocurridos

(dado por el valor 0.9 en el eje  $TIP_\theta$ ). Con esta información se volvió más comprensible el significado de los valores entregados por la métrica AUC en los distintos casos de estudios.

Basándose en los estadísticos de promedio, desviación estándar, mínimo y máximo de los valores AUC generados diariamente por los modelos espacio-temporales de riesgo criminal usando la caracterización temporal dadas por el modelo KPG y modelo WV durante el mes de junio del 2016 se pudo entregar las Tablas 3.7, 3.8 y 3.9 que definieron el desempeño general de dichos métodos en el primer, segundo y tercer turno respectivamente, marcando de color rojo el mejor valor obtenido según el estadístico y el turno analizado.

	Modelo KPG				Modelo WV
	Estudio 0	Estudio 1	Estudio 2	Estudio 3	Optimizado
<i>mean</i>	0.830	0.824	0.835	0.823	<b>0.853</b>
<i>std</i>	0.075	0.092	0.084	0.092	<b>0.070</b>
<i>min</i>	0.708	0.527	0.654	0.585	<b>0.735</b>
<i>max</i>	<b>0.995</b>	0.993	0.989	0.990	0.993

Tabla 3.7: Comportamiento de la métrica AUC para cada uno de los modelos estudiados y sus variantes en el primer turno de la fecha de predicción.

	Modelo KPG				Modelo WV
	Estudio 0	Estudio 1	Estudio 2	Estudio 3	Optimizado
<i>mean</i>	0.841	0.842	<b>0.849</b>	0.828	0.848
<i>std</i>	0.081	0.081	0.078	0.088	<b>0.072</b>
<i>min</i>	0.603	0.614	0.661	0.583	<b>0.686</b>
<i>max</i>	<b>0.980</b>	0.962	0.976	0.964	0.972

Tabla 3.8: Comportamiento de la métrica AUC para cada uno de los modelos estudiados y sus variantes en el segundo turno de la fecha de predicción.

	Modelo KPG				Modelo WV
	Estudio 0	Estudio 1	Estudio 2	Estudio 3	Optimizado
<i>mean</i>	0.772	0.767	<b>0.780</b>	0.766	<b>0.780</b>
<i>std</i>	0.158	0.159	<b>0.156</b>	0.159	<b>0.156</b>
<i>min</i>	0.0	0.0	0.0	0.0	0.0
<i>max</i>	<b>0.938</b>	0.929	0.914	0.928	0.907

Tabla 3.9: Comportamiento de la métrica AUC para cada uno de los modelos estudiados y sus variantes en el tercer turno de la fecha de predicción.

Se procedió a comparar y definir el método de caracterización temporal que obtuvo las mejores características al momento de predecir el riesgo criminal en la comuna de Maipú en todo el mes de estudio. En base a la información entregadas por las Tablas 3.7, 3.8 y 3.9 se tuvo que:

- El modelo KPG con el caso de estudio 2 presentó el mejor desempeño promedio en el segundo y tercer turno, mientras que el modelo WV presentó el mejor desempeño promedio en el primer y tercer turno. En el tercer turno existió una igualdad en el valor promedio considerando tres cifras significativas entre el modelo KPG con caso de estudio 2 y el modelo WV. Los mejores valores promedio entregados en el primer y

segundo turno indicaron que con un 20 % de área cubierta fue posible acertar aproximadamente el 70 % de los delitos ocurridos, mientras que para el tercer turno, utilizando el mismo porcentaje de área, tan solo se pudo acertar el 40 % de los delitos ocurridos aproximadamente.

- El modelo WV presentó los mejores valores promedio de desviación estándar en todos los turnos analizados, siendo el menor de todos ellos el entregado en el primer turno. Por otra parte, el modelo KPG presentó solamente en el tercer turno el menor valor promedio de desviación estándar usando el caso de estudio 2. Esto permitió definir al modelo de riesgo criminal usando el método de caracterización del modelo WV como aquel cuyas predicciones espacio-temporales estuvieron más centradas al valor promedio de aciertos delictuales para cada uno de los turnos del día de predicción
- Se observó que el modelo WV presentó los mejores valores mínimos en todos los turnos analizados. En este caso se buscó el mínimo global presente en todo el mes de junio, esperando que su valor tendiera a ser lo más cercano posible a 1. Con esto se pudo denotar que el modelo WV entregó un piso de desempeño superior al dado por la metodología entregada por el modelo KPG, entregando en el peor de los casos (segundo turno) un 40 % de aciertos usando el 20 % del área encendida. En el tercer turno se observó un valor mínimo global igual a 0 en todas las metodologías. Esto fue producido dado que un día del mes de junio no tuvo delitos de validación, obteniendo por ello un valor AUC igual a 0. Sin embargo, utilizando la información gráfica de las Figuras 3.17 y 3.20, se pudo estimar el comportamiento representativo de este estadístico, siendo el modelo WV y modelo KPG con caso de estudio 0 los que obtuvieron los mínimos globales más cercanos a 1 para el tercer turno.
- Los mejores valores máximos fueron adjudicados por el modelo espacio-temporal de riesgo criminal que usó la metodología temporal definida en el modelo KPG con el caso de estudio 0, siendo el mejor caso dado en el primer turno, donde alcanzó un AUC igual a 0.995, denotando un acierto del 100 % de los delitos de validación con menos del 10 % del área encendida.

Según lo descrito, el modelo KPG usando el caso de estudio 2 y modelo WV caracterizaron de mejor manera el riesgo criminal. El modelo WV permitió entregar una respuesta en la fase de validación del modelo espacio-temporal del riesgo criminal con una menor tasa de dispersión respecto a la media AUC generada y una cota inferior de desempeño mucho mayor que los otros modelos para el mes de junio en cada uno de los turnos analizados, reflejando que la respuesta del modelo WV fue mucho más precisa, más acotada y que trabajó con un valor base más alto respecto a los casos de estudio del modelo KPG. Por otra parte, el modelo KPG presentó mejores comportamientos usando arreglos  $A_c^r$  y  $A_p^r$  con espaciados igual a 0.5, indicando que el proceso de caracterización temporal fue mucho más eficiente para KPG con periodos y centros seleccionados cada 4 horas de diferencia.

El tiempo de procesamiento para entrenar los modelos espacio-temporales de riesgo criminal y evaluar las superficies de riesgo en la comuna de Maipú durante el mes de junio 2016 fue considerada como otra medida de desempeño. Esta métrica asoció la capacidad de respuesta que podría tener Carabineros de Chile al momento de distribuir recursos policiales preventivos a diferentes focos de concentración criminal. Se buscó minimizar el tiempo de procesamiento computacional sin perjudicar el desempeño de la validación. La Tabla 3.10 entregó el valor promedio, desviación estándar, mínimos y máximos de los tiempos de procesamiento en cada uno de los casos de estudios para el mes de junio del 2016, marcando con color rojo los



mejores valores asociados a cada estadístico.

	Modelo KPG				Modelo WV
	Estudio 0	Estudio 1	Estudio 2	Estudio 3	Optimizado
<i>mean</i>	680.067	1719.344	848.479	3946.710	<b>660.07</b>
<i>std</i>	88.303	67.248	<b>23.964</b>	147.308	31.239
<i>min</i>	<b>534.474</b>	1582.789	811.444	3583.771	551.27
<i>max</i>	923.307	1830.039	899.893	4210.662	<b>700.03</b>

Tabla 3.10: Comportamiento del tiempo de procesamiento para cada uno de los modelos estudiados al momento de generar 3 superficies de riesgos en el día de predicción.

Se obtuvo que el mejor valor promedio fue entregado por el modelo WV, demorando aproximadamente 11 minutos en generar 3 superficies de riesgo criminal. Por otro lado, el modelo KPG entregó una menor tasa de dispersión, respecto la media, al momento de usar el caso de estudio 2 y el mínimo tiempo de procesamiento para el caso de estudio 0 con un valor aproximado de 9 minutos. Por último, el mínimo valor máximo de tiempo se produjo en el modelo WV, siendo 2,5 veces menor al mayor mínimo entregado por el caso de estudio 3. Todo dio a entender que las mejores características en los tiempos de procesamiento se presentaron en el modelo WV optimizado, el cual obtuvo valores cercanos en desviación estándar y mínimo comparados con el modelo KPG con el caso de estudio 0 y 2.

Finalmente se justificaron los resultados expuestos en las Tablas 3.7, 3.8 y 3.9 usando el comportamiento del MSE de prueba en la fase *priori* y *posteriori*. Las Tablas 3.11 y 3.12 entregaron la información condensada de las Figuras B.1, B.2, B.3, B.4 y C.1, calculando la media del estadístico promedio, desviación estándar, mínimo y máximo diario para todo el mes de junio del 2016.

	Modelo KPG				Modelo WV
	Caso base	Estudio 1	Estudio 2	Estudio 3	Optimizado
<i>mean</i>	0.047	<b>0.043</b>	0.047	0.044	0.083
<i>std</i>	0.038	<b>0.029</b>	<b>0.029</b>	<b>0.029</b>	0.038
<i>min</i>	0.003	<b>0.002</b>	0.003	0.003	0.015
<i>max</i>	0.272	0.172	<b>0.163</b>	0.174	0.196

Tabla 3.11: Comportamiento de los MSE de prueba del proceso *priori* para cada uno de los modelos temporales en el mes de junio 2016

	Modelo KPG				Modelo WV
	Caso base	Estudio 1	Estudio 2	Estudio 3	Optimizado
<i>mean</i>	0.276	0.297	0.232	0.282	<b>0.100</b>
<i>std</i>	0.376	0.427	0.311	0.446	<b>0.042</b>
<i>min</i>	0.026	0.024	<b>0.023</b>	0.026	0.033
<i>max</i>	1.433	1.643	1.235	1.698	<b>0.182</b>

Tabla 3.12: Comportamiento de los MSE de prueba del proceso *posteriori* para cada uno de los modelos temporales en el mes de junio 2016

Según lo entregado en la Tabla 3.11, el modelo KPG, usando el caso de estudio 1, entregó los mejores valores promedio tanto para media, desviación estándar y mínimo valor de MSE. Sin embargo, este caso de estudio no figuró con los mejores desempeños en la etapa de validación, dando a entender que los errores asociados a la etapa a *priori* no presentaron una fuerte correlación con la etapa de validación para este caso de estudio. Por otra parte, la Tabla 3.12 indicó que el modelo WV obtuvo el mejor comportamiento respecto al valor promedio de los estadísticos asociados al MSE de prueba en la fase *posteriori*, entregando una respuesta muy bien acotada y centrada en el valor media mensual calculado. El modelo KPG, con el caso de estudio 2, presentó el siguiente mejor comportamiento para la fase *posteriori*.

El buen desempeño entregado por el modelo WV en la fase *posteriori* indicó que el comportamiento temporal de los delitos ocurridos durante una semana fue bien descrita por medio de la suma ponderada de arreglos temporales de soporte semanal anterior a la semana a predecir, aplicando un factor de olvido igual a  $0,9^x$ , donde  $x$  fue aumentando en una unidad a medida que se iba alejando semanalmente de la fecha de predicción, tal como se explicó en la Sección 2.5.3.2.

Si bien el modelo espacio-temporal para predecir el riesgo criminal tuvo uno de sus mejores estadísticos de desempeño al momento de utilizar el modelo WV (comparado con los modelos KPG) y este, a su vez, presentó un buen comportamiento promedio respecto al MSE de la STS de prueba de la fase *posteriori*, no se pudo concretizar si existió una correlación directa entre este desempeño temporal con la respuesta generada por el modelo de caracterización espacio-temporal.

Dado que los modelos espacio-temporales generados para predecir el riesgo criminal en los días del mes de junio del 2016 utilizaron siempre un coeficiente inter-proceso temporal ( $\lambda$ ) fijo e igual a 0,7, el cual favoreció la información más reciente de los datos caracterizados (fase *posteriori*). Esto significó que un modelo temporal pudo haber tenido un buen desempeño en la caracterización de la data histórica (fase *priori*), pero debido a que el valor de  $\lambda$  utilizado no presentó mejora alguna en el desempeño de predicción de eventos criminales. Siguiendo con la idea, el modelo WV obtuvo el mejor desempeño promedio en la caracterización temporal de la fase *posteriori* y bajo en el caso *priori*, fue posible obtener una predicción mucho más acertada comparada con los otros modelos temporales analizados.

Para confirmar la importancia de  $\lambda$  en el desempeño de caracterización espacio-temporal, se entregó la Sección 3.6 con tal obtener una idea de cual fase tuvo mayor importancia al momento de usar el modelo KPG y WV en la predicción de los eventos criminales.

### 3.5. Respuesta en diferentes comunas

Esta sección demostró la versatilidad que tuvo el modelo predictivo espacio-temporal del riesgo criminal al momento de ser ejecutado en diversas comunas de la región Metropolitana. Las pruebas fueron desarrolladas para la comuna de Conchalí, La Florida, Las Condes, Recoleta, San Bernardo, Estación Central, Quinta Normal, Pudahuel y Puente Alto. La Figura 3.21 entregó los límites y ubicación de cada comuna en la región Metropolitana.

Las pruebas ejecutadas utilizaron como modelo temporal el WV optimizado, debido al mejor desempeño denotado respecto al modelo KPG, usando los mismos parámetros descritos en la Sección 3.4.2 y un coeficiente inter-procesos  $\lambda = 0,7$ .

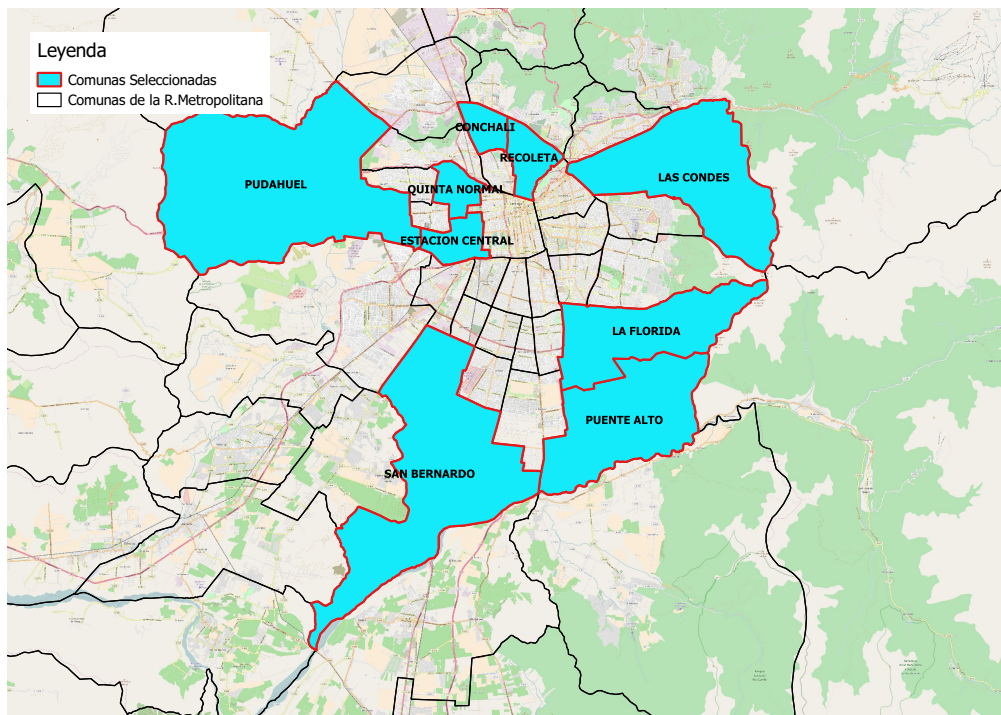


Figura 3.21: Comunas que entran en el proceso de caracterización del riesgo criminal durante todo el mes de junio del 2016.

Los días de análisis contemplaron todo el mes de junio del 2016, generando en cada día un modelo diferente, siendo evaluado para los tres turnos definidos del día siguiente de su entrenamiento. Por ejemplo, si se entrenó un modelo el 15 de junio del 2016 para la comuna de Conchalí, entonces se procede a evaluar con el primer, segundo y tercer turno del día 16 de junio dicho modelo generado. Así para cada día del mes y para cada comuna seleccionada. En el proceso de validación del modelo espacio-temporal se utilizó la métrica AUC para comparar el desempeño de cada modelo generado. El resultado fue mostrado en la Figura 3.22, condensando el comportamiento de cada una de las comunas en sus respectivos turnos de predicción.

Se distinguió diferentes comportamientos de desempeño del modelo espacio-temporal generado según la comuna elegida en la etapa de validación. En comunas de gran tamaño, tales como San Bernardo o Pudahuel, los valores AUC se concentraron por sobre los 0.8, al contrario que comunas pequeñas como Conchalí y Quinta Normal donde no se observó una fuerte concentración de dichos valores. Sin embargo, las comunas como Recoleta o Estación Central, que entran en la categoría de comunas pequeñas, sí presentaron una concentración de valores AUC por sobre 0.7 para el primer y segundo turno. Por otro lado, comunas como La Florida, Puente Alto y Las Condes presentaron concentraciones por sobre los 0.7 pero menores a 0.9 en casi todos los turnos analizados.

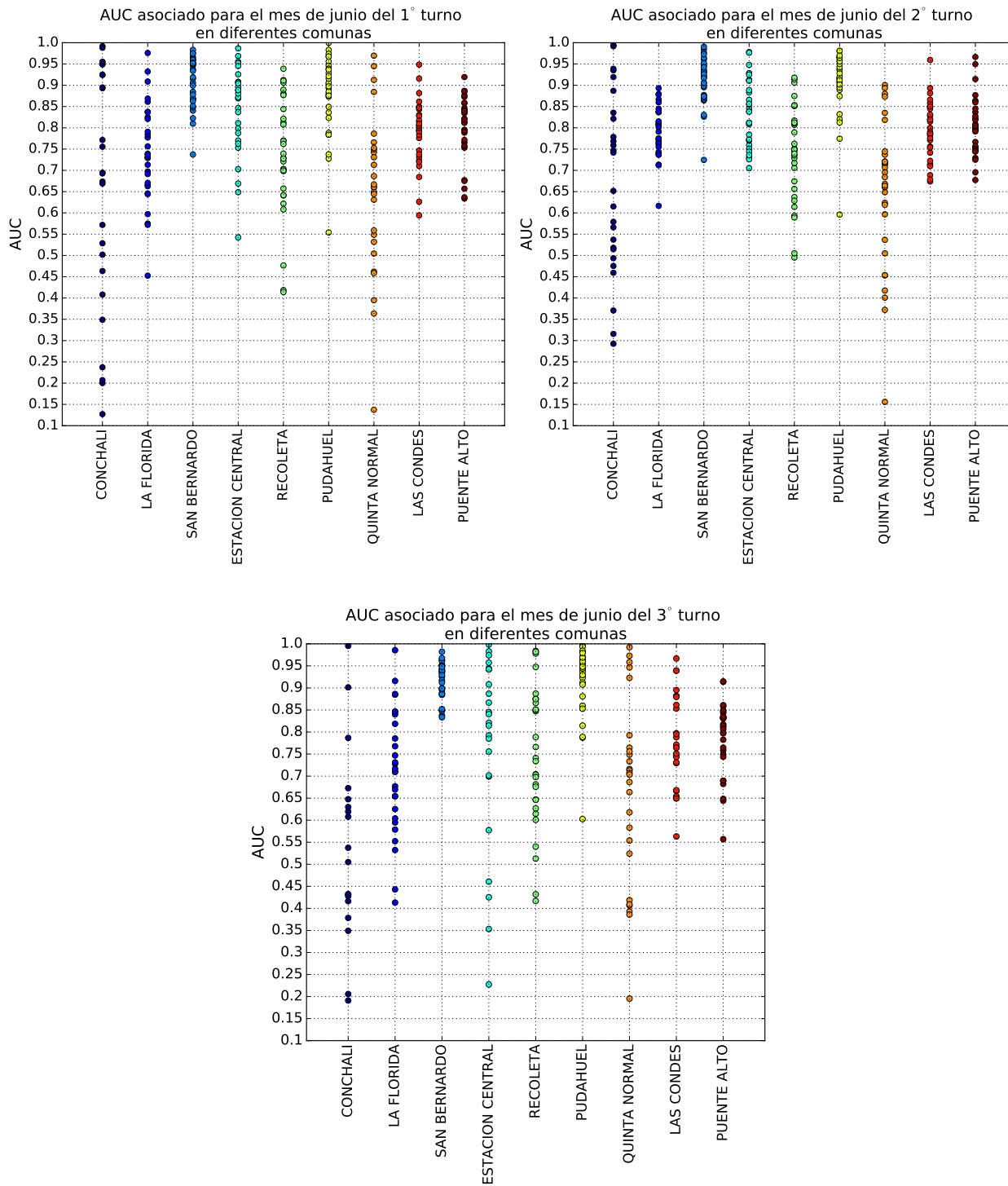


Figura 3.22: Respuestas por medio de la métrica AUC de diferentes comunas aplicando el modelo de caracterización de riesgo criminal usando técnicas temporales WV.

Dado el comportamiento variado de los valores AUC obtenidos en cada una de las comunas de estudio, se realizó una segmentación de las respuestas con tal de clasificar de mejor manera el comportamiento general de cada una de las comunas. Para ello se crearon las Tablas 3.13, 3.14 y 3.15 que resumen la información entregada por la Figura 3.22, segmentando en nueve

intervalos los valores AUC. Cada uno de ellos representó diferentes zonas de desempeño, (i) el segmento escrito con letras de color verde fueron aquellos casos que superaron o igualaron un 35 % de aciertos con un 10 % de área fija cubierta, (ii) el segmento escrito con color naranja fueron las que obtuvieron una tasa de aciertos entre el 15 % y 35 % con un 10 % de área fija cubierta y (iii) finalmente el segmento escrito con color rojo marcó los casos de peor desempeño, obteniendo respuestas bajo el 15 % de aciertos para el 10 % de área fija cubierta. Al final de cada una de las Tablas 3.13, 3.14 y 3.15 se entregó una columna con la suma de casos que cumplen con cada uno de los intervalos AUC estudiados por comuna y una fila al final de dichas Tablas, indicando la cantidad de casos que caen dentro del segmento analizado independiente de la comuna estudiada. La suma acumulada de la última fila no necesariamente fue igual a 30, debido a que en ciertos días no se registraron delitos de validación, dando valores AUC igual a cero. Estos casos no son considerados y quedan filtrados al momento de tomar valores AUC sobre 0,1.

	AUC									# de casos por comuna
	1.00-0.95	0.95-0.90	0.90-0.85	0.85-0.80	0.80-0.75	0.75-0.70	0.70-0.65	0.65-0.60	0.60-0.10	
Conchalí	4	3	2	0	2	0	4	0	10	25
La Florida	1	2	2	4	4	6	5	2	4	30
San Bernardo	10	7	8	4	0	1	0	0	0	30
Estación Central	4	5	6	3	5	1	1	1	1	27
Recoleta	0	4	3	4	3	6	3	4	3	30
Pudahuel	6	8	7	3	3	2	0	0	1	30
Quinta Normal	1	2	1	0	3	4	4	3	9	27
Las Condes	0	2	2	11	6	6	1	1	1	30
Puente Alto	0	1	7	10	7	0	3	2	0	30
# de casos por intervalo	26	34	38	39	33	26	21	13	29	259

Tabla 3.13: Acumulación de todos los primeros turnos de predicción analizados durante todo el mes de junio en las diferentes comunas en base a distintos intervalos AUC.

	AUC									# de casos por comuna
	1.00-0.95	0.95-0.90	0.90-0.85	0.85-0.80	0.80-0.75	0.75-0.70	0.70-0.65	0.65-0.60	0.60-0.10	
Conchalí	3	3	1	2	3	2	1	1	11	27
La Florida	0	0	5	10	8	6	0	1	0	30
San Bernardo	7	11	8	2	0	1	0	0	0	29
Estación Central	2	6	5	6	6	5	0	0	0	30
Recoleta	0	3	4	4	3	7	2	3	4	30
Pudahuel	10	11	4	3	1	0	0	0	1	30
Quinta Normal	0	1	3	2	0	7	5	3	9	30
Las Condes	1	0	6	7	9	4	3	0	0	30
Puente Alto	1	2	3	8	8	6	2	0	0	30
# de casos por intervalo	24	37	39	44	38	38	13	8	25	266

Tabla 3.14: Acumulación de todos los segundo turnos de predicción analizados durante todo el mes de junio en las diferentes comunas en base a distintos intervalos AUC.

	AUC									# de casos por comuna
	1.00-0.95	0.95-0.90	0.90-0.85	0.85-0.80	0.80-0.75	0.75-0.70	0.70-0.65	0.65-0.60	0.60-0.10	
Conchalí	1	1	0	0	1	0	1	4	10	18
La Florida	1	1	2	4	2	6	4	2	6	28
San Bernardo	6	11	7	4	0	0	0	0	0	28
Estación Central	4	4	2	4	3	1	1	0	5	24
Recoleta	3	1	6	1	2	4	3	5	4	29
Pudahuel	10	10	3	1	2	0	0	1	0	27
Quinta Normal	3	2	0	0	3	6	2	1	9	26
Las Condes	1	2	6	0	7	4	5	2	1	28
Puente Alto	0	2	2	13	5	2	3	2	1	30
# de casos por intervalo	29	34	28	27	25	23	19	17	36	238

Tabla 3.15: Acumulación de todos los tercer turnos de predicción analizados durante todo el mes de junio en las diferentes comunas en base a distintos intervalos AUC.

Se observó que tanto Conchalí como Quinta Normal concentraron un gran porcentaje de casos definidos en el peor intervalo de desempeño en todos los turnos, indicando que estas comunas tuvieron complicaciones al momento de tratar de caracterizar el riesgo criminal y predecir su comportamiento. Una razón por la cual estas comunas tendieron a tener mal desempeño fue la baja cantidad de delitos que ocurrieron durante el proceso de validación, siendo en muchos casos tan solo 1 o 2 delitos por turno, influyendo en la interpretación de la métrica AUC. Cuando existe tan solo un delito de validación, la curva AUC pasa de 0% a 100% de aciertos instantáneamente para un determinado porcentaje de cobertura. En este caso el valor AUC sería igual a  $1 - x$ , donde  $x$  es el área mínima con la cual fue posible capturar el evento ocurrido. Además, si el comportamiento del delito no coincidiera con lo predicho, el valor AUC sería fuertemente afectado, entregando valores bajo los 0,5. Otra razón que no favoreció al proceso de validación del modelo espacio-temporal en estas comunas fue la baja cantidad de grillas asociadas, donde cada grilla tiene asociada un alto porcentaje de cobertura respecto al área total de la comuna.

Caso muy contrario fue denotado por San Bernardo, el cual no obtuvo ningún caso con valores de AUC bajo 0,6. Esto principalmente debido a que, al ser una comuna grande, la cantidad de grillas que estuvieron asociadas al análisis criminal permitió diferenciar con mayor resolución el porcentaje de grillas encendidas, siendo una grilla encendida un porcentaje mucho menor al caso de comunas chicas, donde existían pocas grilla de análisis. Además, San Bernardo concentro en casi todos los turnos una cantidad superior a 3 delitos, permitiendo una diferenciación mucho mejor del valor AUC. Otro factor de importancia fue que en comunas grandes con prominente centralización y gran cantidad de zona rural se necesitó de un bajo porcentaje de grillas encendidas para acertar a la mayoría de los delitos, dado que gran parte de estos ocurrieron justamente en zonas urbanizadas.

Una comuna que no presentó una concentración prominente en alguno de los tres segmentos de las Tablas 3.13, 3.14 y 3.15 fue Recoleta. En esta comuna la mayor parte de los casos estuvo en el segmento que representó una tasa de aciertos entre 15% y 35% con un 10% de área fija cubierta, seguido por el segmento definido con letras verdes y finalmente por el caso segmentado con letras rojas. Como se pudo observar en la Figura 3.21, la comuna de Recoleta tiene dimensiones geográficas muy semejantes a las comunas de Quinta Normal y Conchalí, siendo esta última su colindante a ella. Las comunas descritas presentan similitudes en el

tamaño y en la ubicación dentro de la Región Metropolitana, lo cual permite entregar una correlación respecto a los resultados entregados en las comunas de Quinta Normal, Conchalí y Recoleta. La diferencia sustancial que existe entre las comunas recién nombradas fue la cantidad de delitos ocurridos durante el proceso de validación, en donde la comuna de Quinta Normal presenta una tasa de ocurrencia muy por sobre a la comuna de Conchalí o Recoleta.

Abocándose al comportamiento general de las comunas de estudio, se observó que el 52,8 % de los casos totales analizados para el primer turno estuvieron en el rango deseable de desempeño (color verde), mientras que en el segundo turno alcanzo un 54,1 % y para el tercer turno solo un 49,5 %. Esto denota que aproximadamente la mitad de la totalidad de los casos analizados tuvieron una tasa de acierto por sobre el 35 % con tan solo un 10 % de área cubierta independiente del turno escogido para la predicción. En la zona de desempeño denotada por las letras de color naranja se obtuvo un 35,9 %, 36,4 % y 35,2 % para el primer, segundo y tercer turno respectivamente. Con lo observado fue posible aseverar que el modelo de caracterización del riesgo criminal funcionó con muy buenos rendimientos para diferentes comunas de la región metropolitana basándose en la métrica de desempeño AUC, siendo en algunas comunas un poco más difícil el proceso de caracterización que en otras.

Finalmente, la Tabla 3.16 entregó información respecto a los tiempos promedio de generación y evaluación de cada uno de los modelos en los diferentes días asociados al mes de junio 2016.

	Conchalí	La Florida	San Bernardo	Estación Central	Recoleta	Pudahuel	Quinta Normal	Las Condes	Puente Alto
Tiempos promedio por turnos [segundos]	164	690	473	247	229	120	251	459	571

Tabla 3.16: Tiempo promedio en diferentes comunas al momento de generar 3 superficies de riesgo respectivas al primer, segundo y tercer turno del día siguiente a la fecha de consulta usando el modelo espacio-temporal de riesgo criminal con técnica temporal Wavelet.

Fue posible apreciar que existió una relación directamente proporcional entre el tiempo de ejecución del modelo y su evaluación para los tres turnos versus la dimensión geográfica de la comuna. Esta directa proporcionalidad fue justificada por el método para calcular el valor de riesgo , siendo necesario pasar por una suma de 25 puntos por cada grilla.

### 3.6. Influencia del coeficiente inter-proceso temporal $\lambda$

Tal como se mostró en la Sección 2.6, el modelo espacio-temporal de caracterización del riesgo criminal presentó un parámetro denominado por el nombre de coeficiente inter-proceso  $\lambda$  que permitió mezclar información de la fase *priori* y *posteriori*. Durante el desarrollo de las Secciones 3.3, 3.4 y 3.5 este parámetro fue fijado en 0,7. En la Sección 3.4 se observó que la caracterización del riesgo criminal usando el modelo WV tuvo mejor desempeño de predicción en el mes de junio del 2016 en la comuna de Maipú respecto a los modelos temporales KPG utilizando la métrica de evaluación AUC para los tres turnos del día siguiente a la fecha de ejecución de la consulta.

Con tal de identificar si el desempeño temporal en la fase *poseriori* del modelo WV tuvo

relación con el mejor desempeño en la caracterización espacio-temporal de riesgo criminal, se propuso cambiar los valores del coeficiente inter-proceso, variando la importancia entregada por la información de la fase *priori* y *posteriori*. En este caso se utilizaron los datos del mes de junio del 2016 en la comuna de Maipú, siendo anteriormente generadas las GMM espaciales, permitiendo consistencia al momento de comparar estos resultados con los generados en la Sección 3.4. Los valores de  $\lambda$  fueron variando entre 0 y 1 con un espaciado igual a 0,1. Un valor de  $\lambda = 0$  indicó una total importancia de la fase *priori* del modelo, mientras que  $\lambda = 1$  generó un modelo que solo ocupó información de la fase *posteriori*. El desempeño de la etapa de evaluación siguió siendo por medio de la métrica AUC. Las Figuras D.1, D.2 y D.3 mostraron la respuesta diaria promedio del mes de junio del 2016 en la comuna de Maipú para los modelos WV y KPG con caso de estudio 0 y 2 a diferentes valores  $\lambda$  variando de 0 a 1 cada 0,1, con las cuales se pudo entregar las siguientes observaciones.

### Modelo WV

- Los peores desempeños promedios fueron obtenidos por modelos que utilizan un valor de  $\lambda = 1$ , denotando que sin información histórica o fase *priori* el modelo no entrega un valor óptimo de desempeño. Por otra parte, los mejores desempeños promedio fueron con  $\lambda = 0,2$  en el primer y segundo turno y  $\lambda = 0,3$  en el tercer turno.
- El valor de desempeño promedio usando el valor  $\lambda = 0,7$  tomó el cuarto lugar para el primer turno y octavo lugar para los otros dos turnos. En particular se pudo apreciar que en el segundo y tercer turno existió una notable disminución de desempeño a medida que se fue quitando importancia a la información histórica en el modelo espacio-temporal de riesgo criminal.
- Los valores entregados en el segundo turno presentaron menor dispersión respecto a la media comparado con el primer y tercer turno, siendo este último el que presentó la mayor tasa de dispersión.

### Modelo KPG

- Tanto en el caso de estudio 0 y 2, el peor desempeño promedio fue para los modelos con coeficientes inter-proceso  $\lambda = 1$ , indicando una baja tasa de aciertos para modelos contruidos con solo información delictual reciente o fase *posteriori*.
- El tercer turno presentó un valor promedio de AUC relativamente constante entre valores  $\lambda = 0,1$  y  $\lambda = 0,6$ , dando a conocer que la mezcla entre la información de la fase *priori* y *posteriori* permitió un desempeño bastante uniformado siempre y cuando se le entregue mayor importancia a la fase *priori*.
- Para el primer y segundo turno de predicción los valores extremos de  $\lambda$  entregaron los peores desempeños promedios. Por otra parte, los mejores valores de desempeño fueron con  $\lambda$  igual a 0,1, 0,3, 0,4 y 0,5, predominando la importancia de la fase *priori* en mejorar el desempeño del modelo espacio-temporal.

Según lo descrito en los puntos anteriores se pudo identificar que el buen desempeño de la caracterización temporal de la fase *posteriori* no tuvo incidencia importante en el desempeño promedio del modelo espacio-temporal utilizando técnicas de descomposición Wavelet en las señales temporales. Sin embargo, realizando una mezcla entre la información histórica y reciente de los datos al momento de generar los modelos espacio-temporales de caracterización



del riesgo criminal, fue posible mejorar el desempeño promedio de predicción para los tres primeros turnos del día siguiente. Dicha mezcla permitió aumentar casi un 10% la tasa de aciertos en el segundo y tercer turno utilizando un 10% de área de cobertura fijo, tanto por las técnicas temporales entregadas por el modelo WV como por el modelo KPG.

Finalmente se ha generado la Tabla 3.17 con información condensada de los mejores y peores casos de estudio dados por el modelo WV y el modelo KPG en el mes de junio del 2016.

		AUC promedio	$\lambda$	AUC promedio	$\lambda$	AUC promedio	$\lambda$
Primer turno	Mejor	0.855	0.2	0.841	0.4	0.845	0.3
	Peor	0.838	0.0	0.824	0.9	0.821	1.0
Segundo turno	Mejor	0.869	0.2	0.851	0.1	0.856	0.3
	Peor	0.821	1.0	0.814	1.0	0.83	1.0
Tercer turno	Mejor	0.804	0.3	0.785	0.5	0.784	0.3
	Peor	0.75	1.0	0.761	1.0	0.759	0.0
		WV		KPG con caso estudio 0		KPG con caso estudio 2	

Tabla 3.17: Mejor y peor desempeño promedio del modelo espacio-temporal usando diferentes valores coeficiente inter-proceso  $\lambda$  en la comuna de Maipú en el mes de junio del 2016 para cada uno de los turnos de predicción.

De la Tabla 3.17 se destacó el mejor desempeño del modelo WV usando mayor ponderación en la fase *priori* en cada uno de los turnos analizados por sobre los modelos KPG con sus casos de estudio. Fue importante notar que en ningún caso se presentó el valor de  $\lambda = 0,7$  como mejor candidato, siendo los modelos entrenados en las secciones anteriores sub-óptimos del caso actualmente analizado. Esto implicó que el modelo WV entregó un mejor desempeño promedio no por haber obtenido mejores estadísticos en la caracterización temporal de la fase *posteriori*, sino porque el modelo en si pudo caracterizar de mejor manera la tendencia criminal usando tanto la información histórica y reciente respecto al modelo KPG.

# Conclusión

Basándose en los objetivos y resultados entregados en este trabajo de tesis se puede concluir que:

- Se ha diseñado y analizado el desempeño de un modelo de predicción espacio-temporal de riesgo criminal en diferentes comunas de la región Metropolitana utilizando dos técnicas de caracterización temporal basados en transformaciones de kernels periódicos Gaussianos y funciones Wavelet, el cual permite reconocer los lugares con mayor tendencia a ocurrir un delito dentro de la categoría Robo Violencia y Robo Fuerza para el primer, segundo y tercer turno del día siguiente de la consulta. Más de la mitad de los casos analizados entregan una precisión por sobre el 35 % de aciertos de delitos ocurridos en la fase de validación con solo el 10 % de área efectiva de la comuna.
- La optimización de los modelos temporales en base a la minimización del error cuadrático medio (MSE) de las señales temporales suavizadas en la etapa de prueba y un conjunto de valores paramétricos definidos según el comportamiento empírico de las señales temporales construidas permite encontrar mejoras en la caracterización temporal de la señal y disminuir el valor MSE promedio de la totalidad de componentes. El modelo KPG del caso de estudio 0 (kernels Gaussianos con periodos espaciados cada 4 horas) obtiene el menor valor MSE promedio en el conjunto de prueba de la fase *priori*, mientras que el modelo WV, optimizado bajo un conjunto de duplas ( $bp, m$ ) empíricamente entregadas, entrega el menor valor MSE promedio en el conjunto de prueba de la fase *posteriori*.
- En la etapa de validación, las curvas HRP y TIP ayudaron a comprender el desempeño del modelo ante diferentes escenarios criminales, permitiendo saber la efectividad del modelo espacio-temporal de riesgo criminal a diferentes porcentajes de área activada en las comunas de estudio. Por otra parte, la métrica de desempeño AUC permite simplificar el estudio gráfico entregado por las curvas HRP y TIP en tan solo un valor, facilitando la comprensión del desempeño de cada modelo generado de manera rápida y concisa gracias al Anexo E.
- Durante la comparación de las dos técnicas temporales aplicadas en el mes de junio del 2016 para la comuna de Maipú, el modelo espacio-temporal de riesgo criminal obtiene su mejor rendimiento usando la estrategia de caracterización temporal basada en Transformada Wavelet por sobre el modelo generado por el entrenamiento de la suma ponderada de kernels periódicos Gaussianos. Todo esto en base a los valores promedio generados por los estadísticos media, desviación estándar, mínimo y máximos de la métrica de desempeño AUC.
- Según la Sección 3.5, el modelo espacio-temporal de caracterización del riesgo criminal permite la ejecución en diferentes comunas de la región Metropolitana, entregando

respuestas certeras y efectivas en gran parte de las comunas analizadas. Los casos de estudios demostraron que la superficie de la comuna influye al momento de calcular el desempeño basado en la métrica AUC, entregando una tasa de acierto superior al 35 % usando solo el 10 % del área efectiva en comunas de gran tamaño, tales como Pudahuel o Las Condes. Por otra parte, comunas como Conchalí o Quinta normal presentan mayor dificultad al momento de acertar por sobre el 35 % de los delitos ocurridos con un 10 % de área efectiva encendida. Esto debido a que las comunas tienen menor dimensión y regularmente en cada turno ocurren 1 o 2 delitos aproximadamente afectando los valores entregados por la métrica AUC. El modelo también entrega una respuesta eficiente disponiendo, en el peor de los casos, de 12 minutos en promedio para la generación y evaluación del modelo en tres turnos siguientes a la fecha de consulta.

- La mezcla de la información delictual histórica y reciente permite una mejora significativa en el desempeño de predicción, donde la fase *priori* contribuye con mayor importancia a una mejora del modelo comparada con la fase *posteriori*, la cual sin apoyo de la información histórica entrega los peores valores promedios de desempeño basados en la métrica AUC para la comuna de Maipú con los datos de junio del 2016.

## Trabajo Futuro

### Búsqueda paramétrica del modelo temporal KPG

Este modelo presenta uno de los mayores potenciales al momento de caracterizar las series temporales suavizadas de los eventos criminales debido a que su kernel periódico emula de muy buena manera el comportamiento delictual en las comunas de la región Metropolitana. Sin embargo, es un modelo bastante complejo y que depende de muchos parámetros. En este Trabajo de Tesis se ha considerado como condiciones iniciales el conjunto de periodos y centros que definen la sumatoria de kernels periódicos Gaussianos, siendo entrenado tan solo el peso asociado a cada función. Una mejor estrategia que permita entregar un entrenamiento acotado de los parámetros centros, periodos y pesos puede ser conseguido por medio de optimización por enjambre de partículas (*particle swarm optimization*, PSO) o algoritmos evolutivos, siempre y cuando se tenga la libertad de tiempo de ejecución, la cual lamentablemente en este trabajo de tesis no pudo ser concedida por las exigencias de operación de Carabineros de Chile.

### Mejoras en el modelo Wavelet

El modelo temporal Wavelet entrega dos comportamientos totalmente diferentes. En la fase *posteriori* se observa un modelo que permite un buen ajuste de la señal temporal de prueba, obteniendo los mejores valores MSE promedio para todos los casos de estudios generados en la comuna de Maipú en junio del 2016. Por otra parte, la fase *priori* presenta problemas al momento de predecir la señal de prueba, entregando los peores valores promedios MSE comparado con los modelos KPG. Ante esta situación es pertinente modificar el modelo de predicción entregado para esta fase, la cual puede ser modificada por un modelo ARMAX con la información filtrada con el análisis Wavelet.

## Tiempos de procesamiento

Existieron dos acciones claves al momento de calcular los tiempos de procesamiento del modelo espacio-temporal de riesgo criminal, la primera fue entrenar y la segunda fue evaluar el modelo en la comuna consultada, siendo la de mayor demanda computacional la etapa de entrenar el modelo, dada la cantidad de puntos evaluados en cada grilla de la comuna. Un camino para reducir los tiempos de evaluación es utilizar un estimador de riesgo en base a un muestreo del modelo, tal y como se realiza en filtro de partículas o método de Montecarlo. Dichas partículas estarán concentradas en la zonas de mayor riesgo criminal, siendo posible contar y tener un valor de riesgo proporcional al calculado por medio de los 25 puntos de cada grilla. La mejora consta en utilizar librerías especialmente diseñadas en generar partículas de un modelo bi-dimensional y contabilizar en base a diferentes restricciones definidas por las grillas de la comuna.

# Bibliografía

- [1] R. G. T., “Estudio nacional de opinión pública número 77,” *Centro de Estudios Públicos.*, 2016.
- [2] W. L. Perry, B. McInnis, D. Brown, C. C. Price, D. Brown, S. Smith, and J. Hollywood, *Predictive Policing, The Role of Crime Forecasting in Law Enforcement Operations.* RAND, 2013.
- [3] Y. Xue and D. Brown, “Spatial analysis with preference specification of latent decision makers for criminal event prediction,” *Decision Support Systems*, vol. 41, no. 3, pp. 560–573, 2006.
- [4] S. E. L. y del Caribe., “Combatir el hambre, prioridad en latinoamérica.” *Prensa Latina*, 2016.
- [5] R. G. T., “Estudio nacional de opinión pública número 75,” *Centro de Estudios Públicos.*, 2015.
- [6] G. Adimark, “Encuesta:evaluación gestión de gobierno junio 2015,” *Informe mensual.*, 2015.
- [7] F. P. Ciudadana, *Memoria Institucional 2014.* Chile: Fundación Paz Ciudadana. Políticas públicas en seguridad y justicia, 2014.
- [8] C. de Chile, “Nuevo manual operativo del plan cuadrante de seguridad preventiva,” *Boletín Oficial Carabineros de Chile.*, 2010.
- [9] S. de Prevención del Delito Ministerio del Interior y Seguridad Pública, “Plan nacional de seguridad pública y prevención de la violencia y el delito,” *Seguridad para Todos.*, 2015.
- [10] F. P. Ciudadana, “Análisis de umbral de delitos acumulado al primer semestre de 2016.” *Políticas públicas en seguridad y justicia*, 2016.
- [11] C. de Chile, “Prefectura siat: Investigación integral.” *Revista Carabineros de Chile*, vol. 738, 2016.
- [12] C. O. F. Ulloa, “Primer seminario nacional de geociencias y geotecnologías.” *Sistema de Análisis de Información Territorial SAIT 2.0*, 2015.
- [13] J. Eck, S. Chainey, J. Cameron, and R. Wilson, *Mapping crime: Understanding hotspots.* National Institute of Justice, 2005.
- [14] H. Liu and D. Brown, “Criminal incident prediction using a point-pattern-based density model,” *International Journal of Forecasting*, vol. 19, no. 4, pp. 603–622, Oct. 2003.
- [15] M. A. Smith and D. E. Brown, “Discrete choice analysis of spatial attack sites,” *Information Systems and e-Business Management*, vol. 5, no. 3, pp. 255–274, 2007.
- [16] C. Ivaha, H. Al-Madfai, G. Higgs, and J. Ware, “The Dynamic Spatial Disaggregation Approach: A Spatio-Temporal Modelling of Crime.” in *World Congress on Engineering*, 2007, pp. 961–966.

- [17] X. Wang and D. Brown, “The spatio-temporal modeling for criminal incidents,” *Security Informatics*, vol. 1, no. 1, pp. 1–17, 2012.
- [18] L. Wasserman, *All of Statistics: A Concise Course in Statistical Inference*. Berlin: Springer-Verlag, 2006.
- [19] Q. Li and J. S. Racine, *Nonparametric Econometrics: Theory and Practice*. New Jersey: Princeton University Press, 2007.
- [20] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [21] C. Rohilla S., “Advanced Data Analysis from an Elementary Point of View,” Cambridge University Press, Pittsburgh, Tech. Rep., 2016.
- [22] G. McLachlan and D. Peel, *Finite Mixture Models*. John Wiley and Sons, 2000.
- [23] D. A. Reynolds, *A Gaussian Mixture Modeling Approach to Text-Independent Speaker Identification*. Georgia Institute of Technology, 1992.
- [24] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.
- [25] D. J. C. MacKay, “Introduction to gaussian processes.” *Neural Networks and Machine Learning*, p. 84–92, 1998.
- [26] S. Roberts, M. Osborne, M. Ebdon, S. Reece, N. Gibson, and S. Aigrain, “Gaussian processes for timeseries modelling.” *Signal processing and inference for the Physical Sciences*, vol. 371, p. 84–92, 2013.
- [27] C. Torrence and G. P. Compo, “A practical guide to wavelet analysis.” University of Colorado, Tech. Rep., 1997.
- [28] K. M. Lua and W. H. Y., “Climate signal detection using wavelet transform: How to make a time series sing.” *Bulletin of the American Meteorological Society*, vol. 76, p. 2391–2402, 1995.
- [29] G. Kaiser, *A Friendly Guide to Wavelets*. Birkhäuser, 1994.
- [30] M. Farye, “Wavelet transforms and their applications to tubulence.” *Annu. Rev. Fluid Mech*, vol. 24, pp. 395–457, 1992.
- [31] S. D. Meyers, B. G. Kelly, and J. J. O’Brien, “An introduction to wavelet analysis in oceanography and meteorology: With application to the dispersion of yanai waves.” *American Meteorological Society*, vol. 121, p. 2858–2866, 1993.
- [32] T. Fawcett, “An introduction to roc analysis,” *Pattern recognition letters*, vol. 27, no. 8, pp. 861–874, 2006.
- [33] W. S. S. Gosse, “The probable error of a mean,” *Biometrika*, vol. 6, no. 1, pp. 1–25, 1908.
- [34] K. L. and P. J. Rousseeuw., *Finding Groups in Data: An Introduction to Cluster Analysis*. NJ: John Wiley and Sons, 1990.
- [35] M. Vergara, “Caracterización y predicción de riesgo criminal basado en kernels espacio-temporales incluyendo fenomenología espacial.” Master’s thesis, Universidad de Chile, 2017.

# Capítulo A

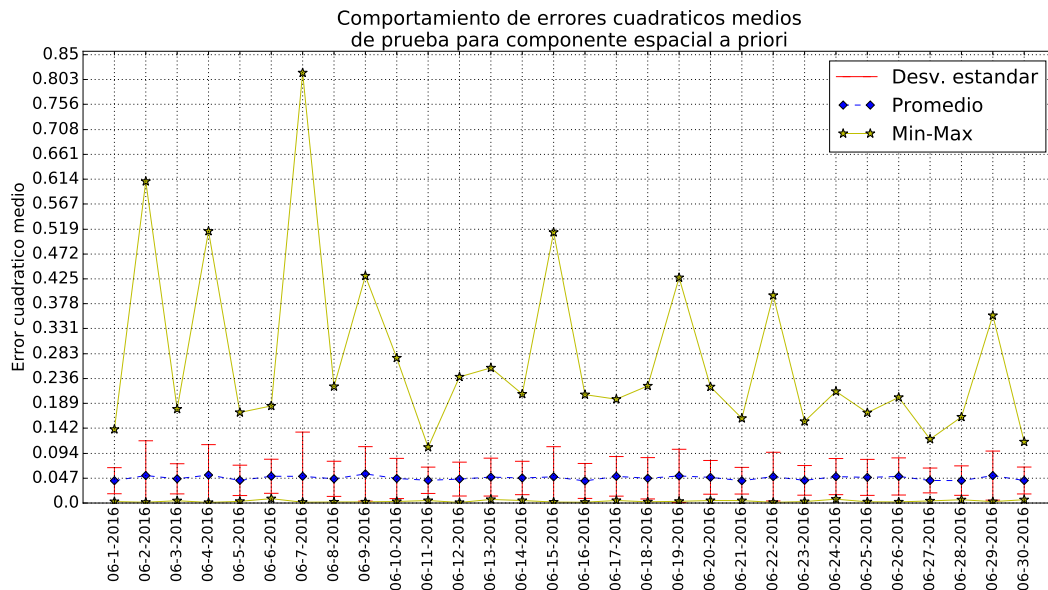
## Clasificación de los delitos utilizados en el trabajo de tesis.

DELITO	CODIGO DEL DELITO	AGRUPADO	TIPO DEL DELITO
FEMICIDIO ART.390 INC.2	720	HOMICIDIO	DMCS
HOMICIDIO ART. 391 NRO.2	702	HOMICIDIO	DMCS
HOMICIDIO CALIFICADO ART. 391 NRO. 1	703	HOMICIDIO	DMCS
HOMICIDIO DE GENDARME EN EL DESEMPEÑO DE SUS FUNCIONES	11101	HOMICIDIO	DMCS
HOMICIDIO EN RIÑA O PELEA ART. 392	705	HOMICIDIO	DMCS
INFANTICIDIO ART. 394	707	HOMICIDIO	DMCS
PARRICIDIO ART. 390 INC.1	701	HOMICIDIO	DMCS
APROPIACION DE CABLES DE TENDIDO ELECTRICO O DE COMUNICACIONES	849	HURTO	DMCS
HURTO AGRAVADO (ART. 447 CODIGO PENAL)	826	HURTO	DMCS
HURTO DE BIENES PERTENECIENTES A REDES DE SUMINISTRO PUBLICO	853	HURTO	DMCS
HURTO DE HALLAZGO ART. 448	821	HURTO	DMCS
HURTO FALTA 494 BIS CODIGO PENAL	13028	HURTO	DMCS
HURTO SIMPLE	801	HURTO	DMCS
HURTO SIMPLE POR UN VALOR DE 4 A 40 UTM. ART. 432 Y 446 NRO. 2	847	HURTO	DMCS
HURTO SIMPLE POR UN VALOR DE MEDIA A MENOS DE 4 UTM. A.432 Y 446 NRO.3	848	HURTO	DMCS
HURTO SIMPLE POR UN VALOR SOBRE 40 UTM. ART. 432 Y 446 NRO. 1	846	HURTO	DMCS
HURTO SUPERMERCADO O CENTROS COMERCIALES	843	HURTO	DMCS
CASTRACION Y MUTILACION ART. 395 y 396	718	LESIONES	DMCS
LESIONES CON ARMA BLANCA	712	LESIONES	DMCS
LESIONES CON ARMA DE FUEGO	714	LESIONES	DMCS
LESIONES CON FUEGOS ARTIFICIALES	713	LESIONES	DMCS
LESIONES GRAVES ART. 397 NRO. 2	709	LESIONES	DMCS
LESIONES GRAVES GRAVISIMAS ART. 397 NRO. 1	717	LESIONES	DMCS
LESIONES LEVES ART. 494 NRO. 5	13001	LESIONES	DMCS
LESIONES MENOS GRAVES ART. 399	710	LESIONES	DMCS
MUTILACION	719	LESIONES	DMCS
RIÑA PUBLICA 496 NRO. 10 CODIGO PENAL	13033	LESIONES	DMCS
ROBO FRUSTRADO	844	ROBO FRUSTRADO	DMCS
ROBO DE ACCESORIOS DE VEHICULOS O ESPECIES INTERIOR VEHICULOS	842	ROBO FUERZA	DMCS
ROBO DE VEHICULO MOTORIZADO ART. 443 INC. 2	831	ROBO FUERZA	DMCS
ROBO EN BIENES NACIONALES DE USO PUBLICO O SITIOS NO DEST.HAB. ART.443	808	ROBO FUERZA	DMCS
ROBO EN LUGAR HABITADO O DESTINADO A LA HABITACION ART. 440	809	ROBO FUERZA	DMCS
ROBO EN LUGAR NO HABITADO ART. 442	810	ROBO FUERZA	DMCS
ROBO CON CASTRACION, MUTILACION O LESIONES GRAVES GRAVISIMAS A.433 N.1	829	ROBO VIOLENCIA	DMCS
ROBO CON HOMICIDIO ART. 433 NRO. 1	827	ROBO VIOLENCIA	DMCS
ROBO CON INTIMIDACION ART. 433, 436 INC. 1 438	802	ROBO VIOLENCIA	DMCS
ROBO CON RETENCION DE VICTIMAS O CON LESIONES GRAVES ART. 433 NRO. 2	830	ROBO VIOLENCIA	DMCS
ROBO CON VIOLACION ART. 433 NRO. 1	828	ROBO VIOLENCIA	DMCS
ROBO CON VIOLENCIA ART. 436 INC. 1 433, 438, 439	803	ROBO VIOLENCIA	DMCS
ROBO POR SORPRESA ART. 436 INC. 2	804	ROBO VIOLENCIA	DMCS
VIOLACION	607	VIOLACION	DMCS
VIOLACION CON HOMICIDIO ART. 372 BIS.	628	VIOLACION	DMCS
VIOLACION DE MAYOR DE 14 AÑOS ART. 361	637	VIOLACION	DMCS
VIOLACION DE MENOR DE 14 AÑOS ART. 362	621	VIOLACION	DMCS

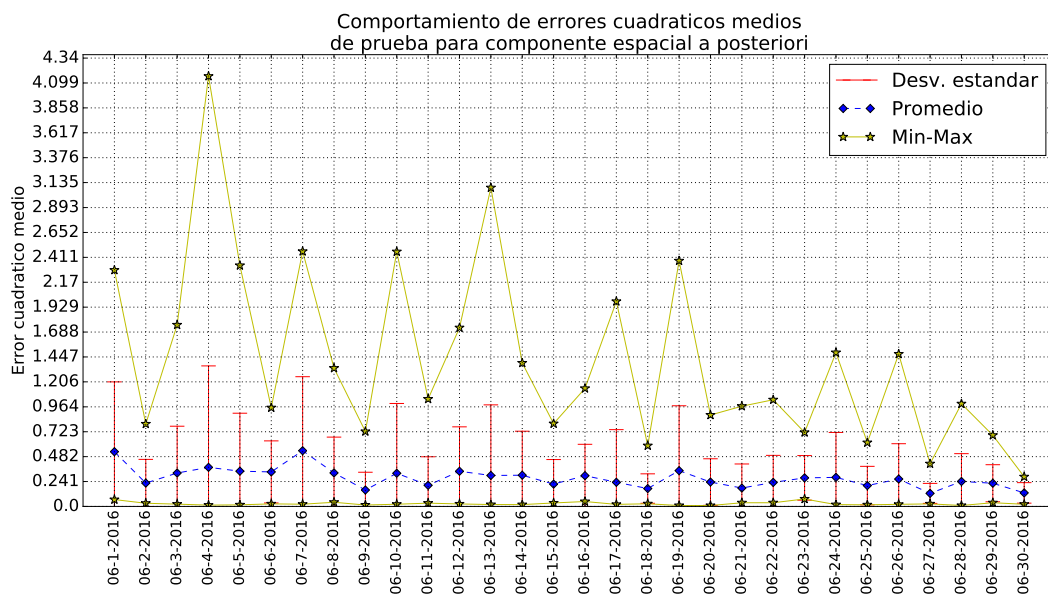
Tabla A.1: Delitos de mayor connotación social (DMCS) la base de datos.

# Capítulo B

## Resultados optimización modelo KPG.



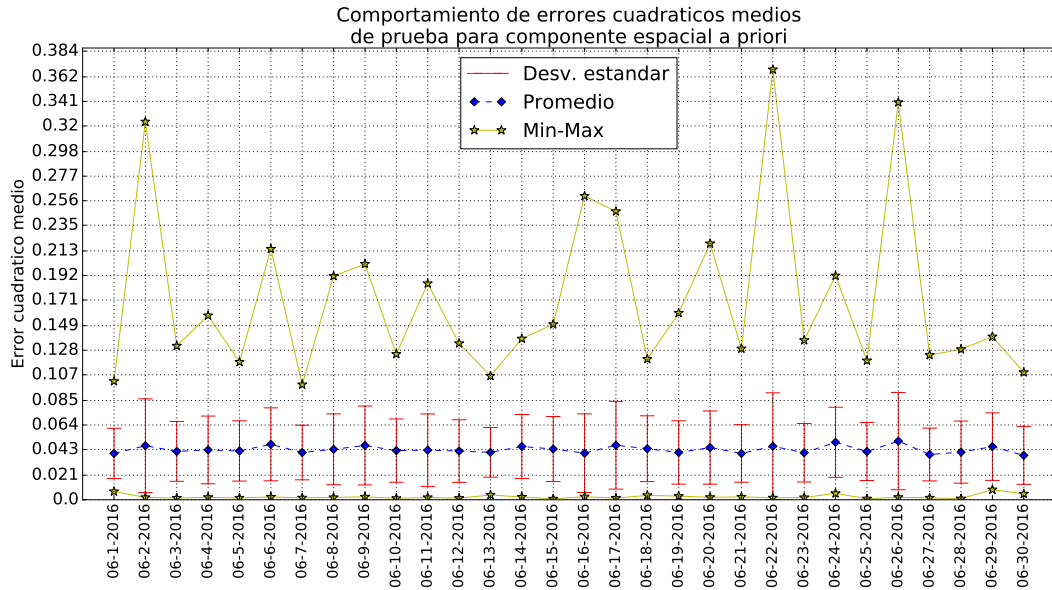
(a) MSE diarios del conjunto de prueba de la fase *priori* en todas las componentes espaciales entrenadas de la comuna de Maipú para el mes de junio del 2016



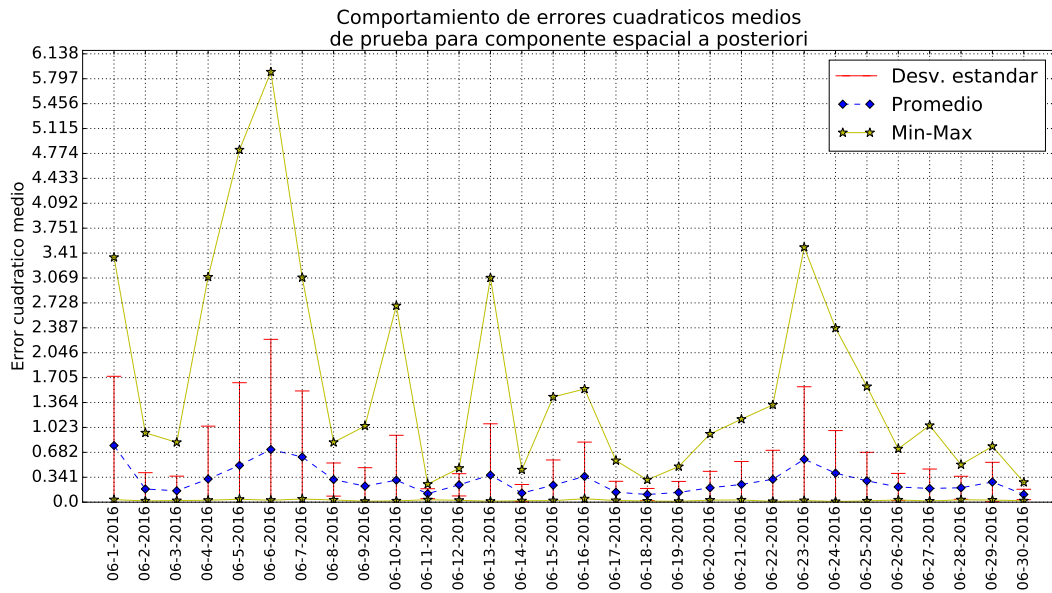
(b) MSE diarios del conjunto de prueba de la fase *posteriori* en todas las componentes espaciales entrenadas de la comuna de Maipú para el mes de junio del 2016

Figura B.1: Curvas de MSE diarios para el conjunto de prueba en la fase *priori* y *posteriori*, usando los arreglos  $A_{p_0}^r$  y  $A_{c_0}^r$  del caso de estudio 0



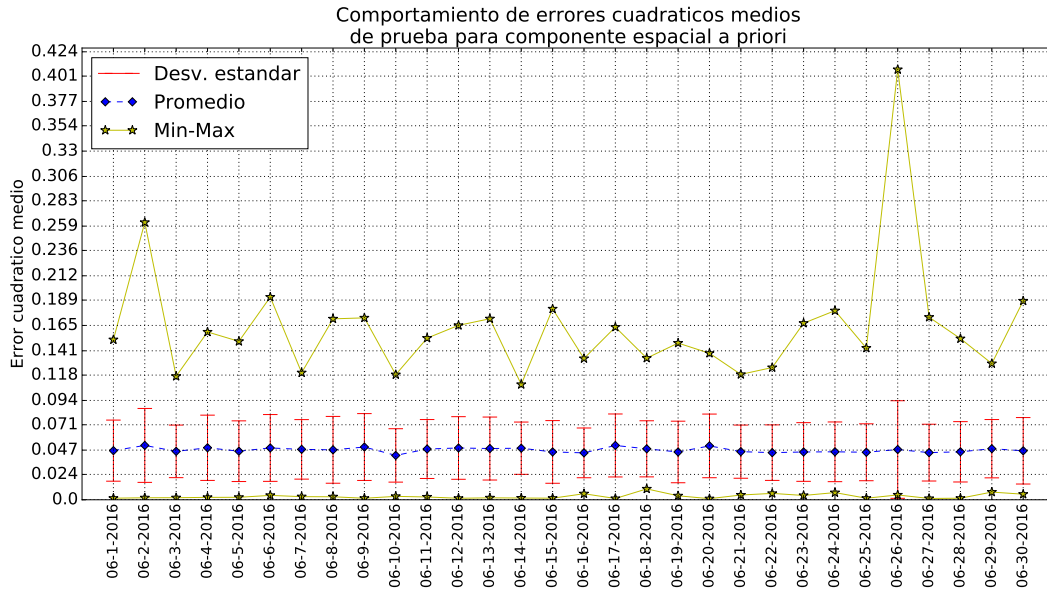


(a) MSE diarios del conjunto de prueba de la fase *priori* en todas las componentes espaciales entrenadas de la comuna de Maipú para el mes de junio del 2016

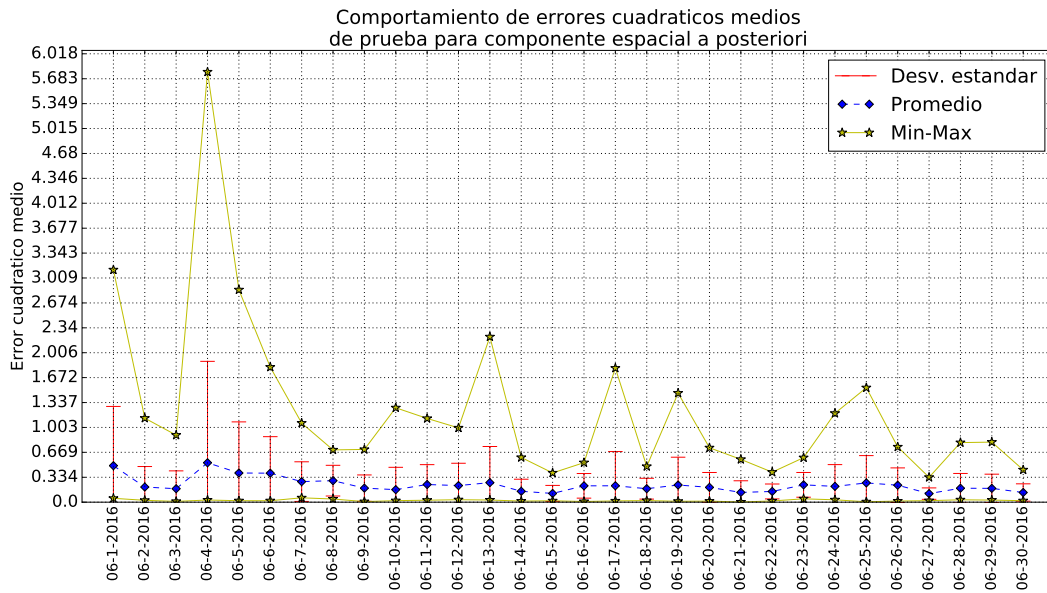


(b) MSE diarios del conjunto de prueba de la fase *posteriori* en todas las componentes espaciales entrenadas de la comuna de Maipú para el mes de junio del 2016

Figura B.2: Curvas de MSE diarios para el conjunto de prueba en la fase *priori* y *posteriori*, usando los arreglos  $A_{p_1}^r$  y  $A_{c_1}^r$  del caso de estudio 1

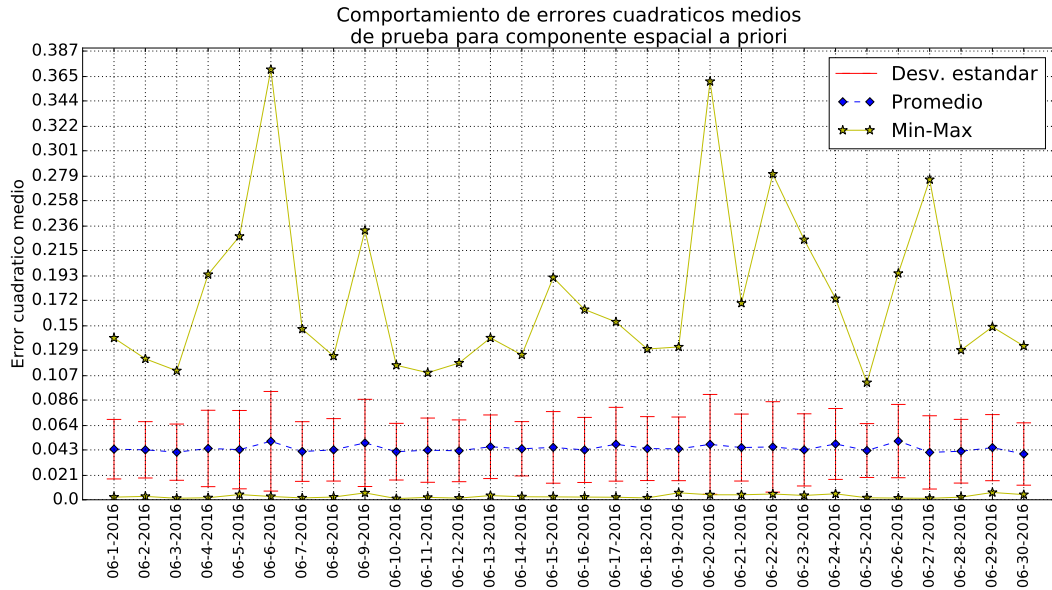


(a) MSE diarios del conjunto de prueba de la fase *priori* en todas las componentes espaciales entrenadas de la comuna de Maipú para el mes de junio del 2016

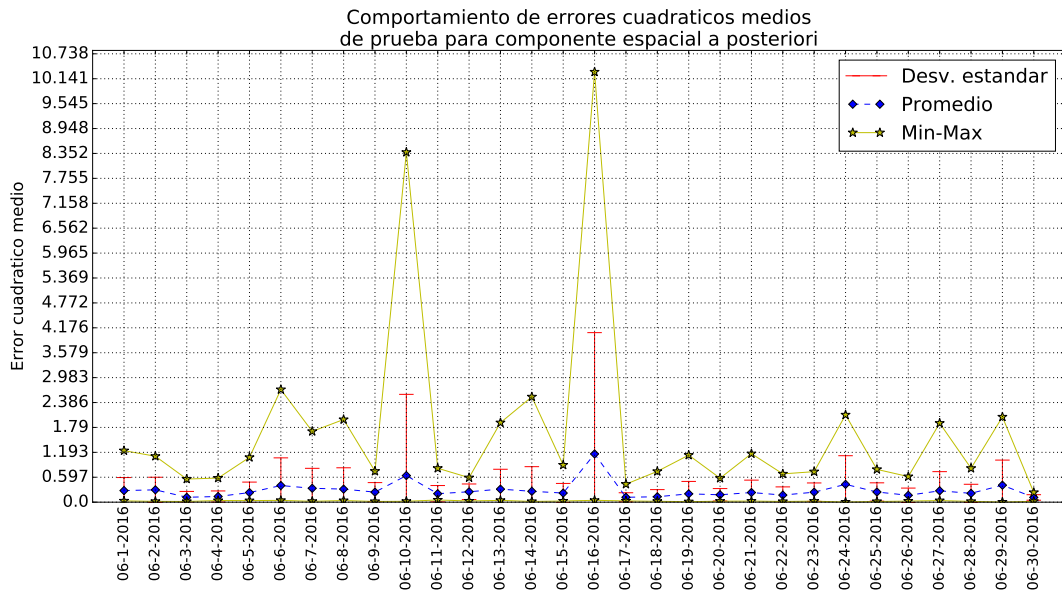


(b) MSE diarios del conjunto de prueba de la fase *posteriori* en todas las componentes espaciales entrenadas de la comuna de Maipú para el mes de junio del 2016

Figura B.3: Curvas de MSE diarios para el conjunto de prueba en la fase *priori* y *posteriori*, usando los arreglos  $A_{p_2}^r$  y  $A_{c_2}^r$  del caso de estudio 2



(a) MSE diarios del conjunto de prueba de la fase *priori* en todas las componentes espaciales entrenadas de la comuna de Maipú para el mes de junio del 2016

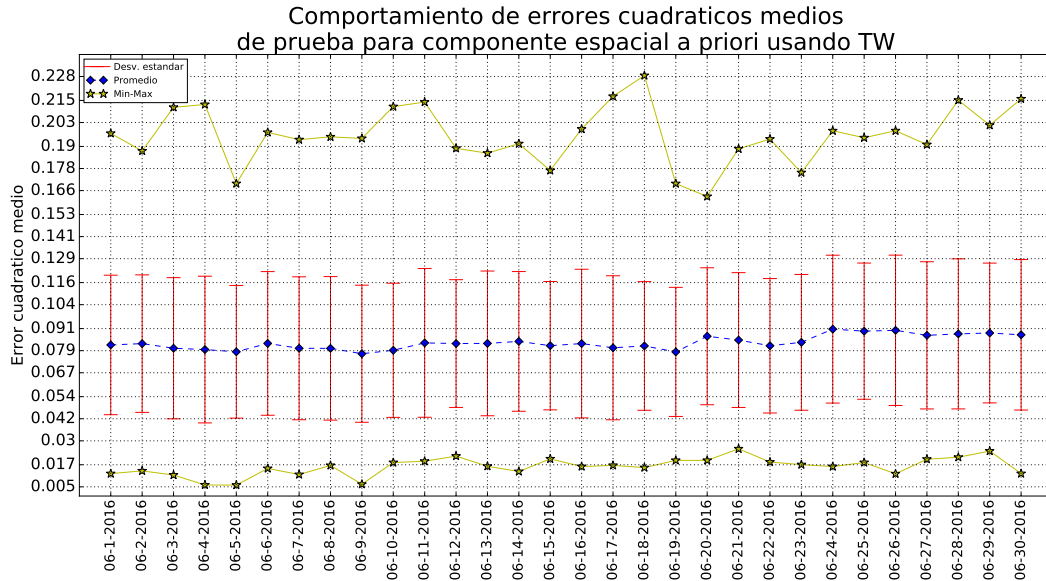


(b) MSE diarios del conjunto de prueba de la fase *posteriori* en todas las componentes espaciales entrenadas de la comuna de Maipú para el mes de junio del 2016

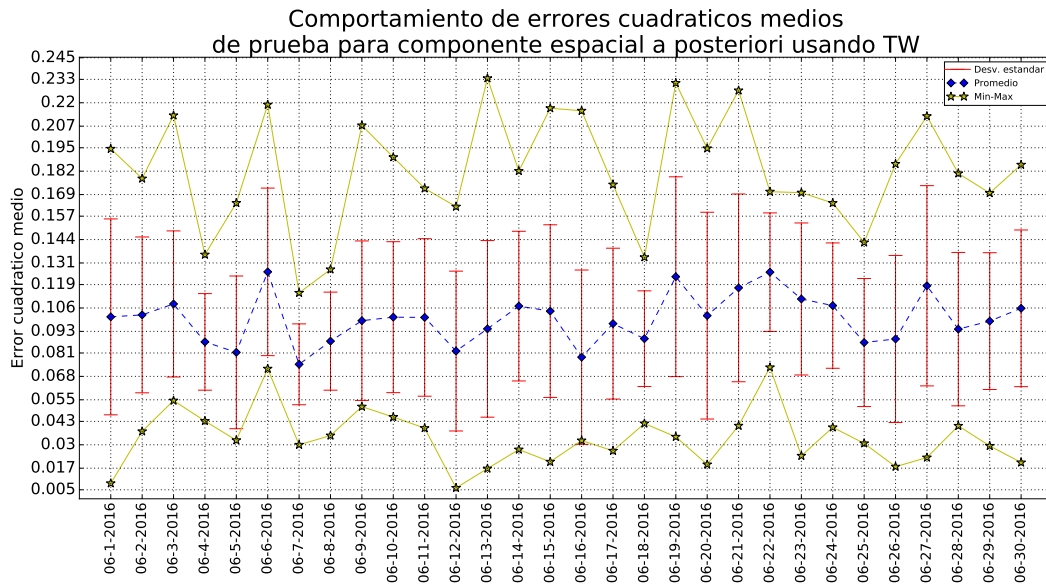
Figura B.4: Curvas de MSE diarios para el conjunto de prueba en la fase *priori* y *posteriori*, usando los arreglos  $A_{p_3}^r$  y  $A_{c_3}^r$  del caso de estudio 3

# Capítulo C

## Resultados optimización modelo WV.



(a) MSE diarios del conjunto de prueba de la fase *priori* en todas las componentes espaciales entrenadas de la comuna de Maipú para el mes de junio del 2016



(b) MSE diarios del conjunto de prueba de la fase *posteriori* en todas las componentes espaciales entrenadas de la comuna de Maipú para el mes de junio del 2016

Figura C.1: Valores de MSE de prueba diarios para la fase *priori* y *posteriori* ocupando el método de optimización Wavelet paramétrica

# Capítulo D

## Resultados de modelos de caracterización variando el coeficiente inter-proceso $\lambda$ .

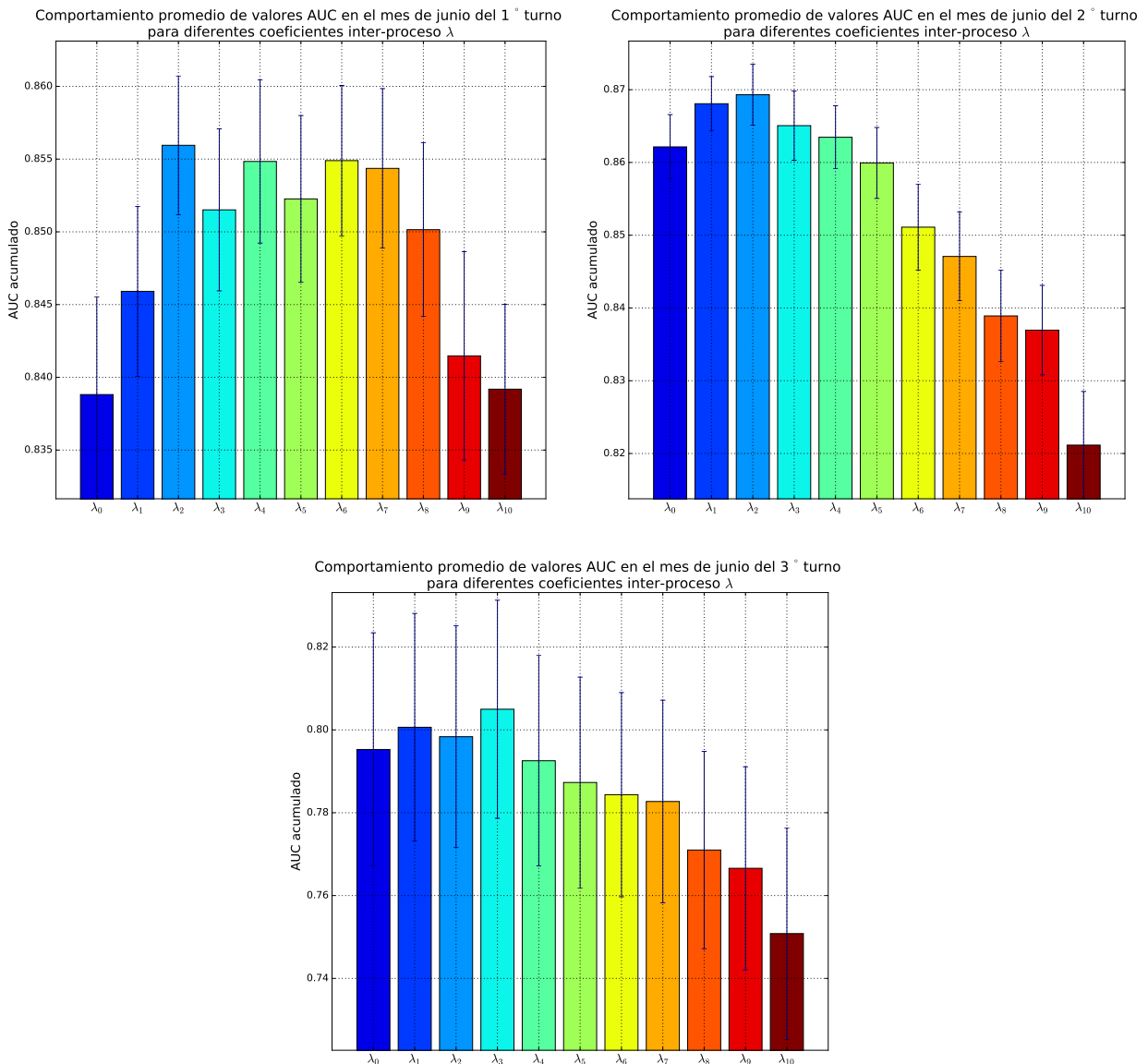


Figura D.1: Promedio de valores AUC diarios del primer, segundo y tercer turno utilizando los datos del mes de junio 2016 en la comuna de Maipú utilizando modelo WV para caracterizar el comportamiento temporal con diferentes coeficientes inter-proceso  $\lambda$ . En el eje de las abscisas se ordenan de mayor a menor los valores de  $\lambda$ , partiendo en 0 y terminando en 1, con cada paso de avance igual a 0,1. Las líneas horizontales entregan la varianza de los datos para cada caso de estudio.

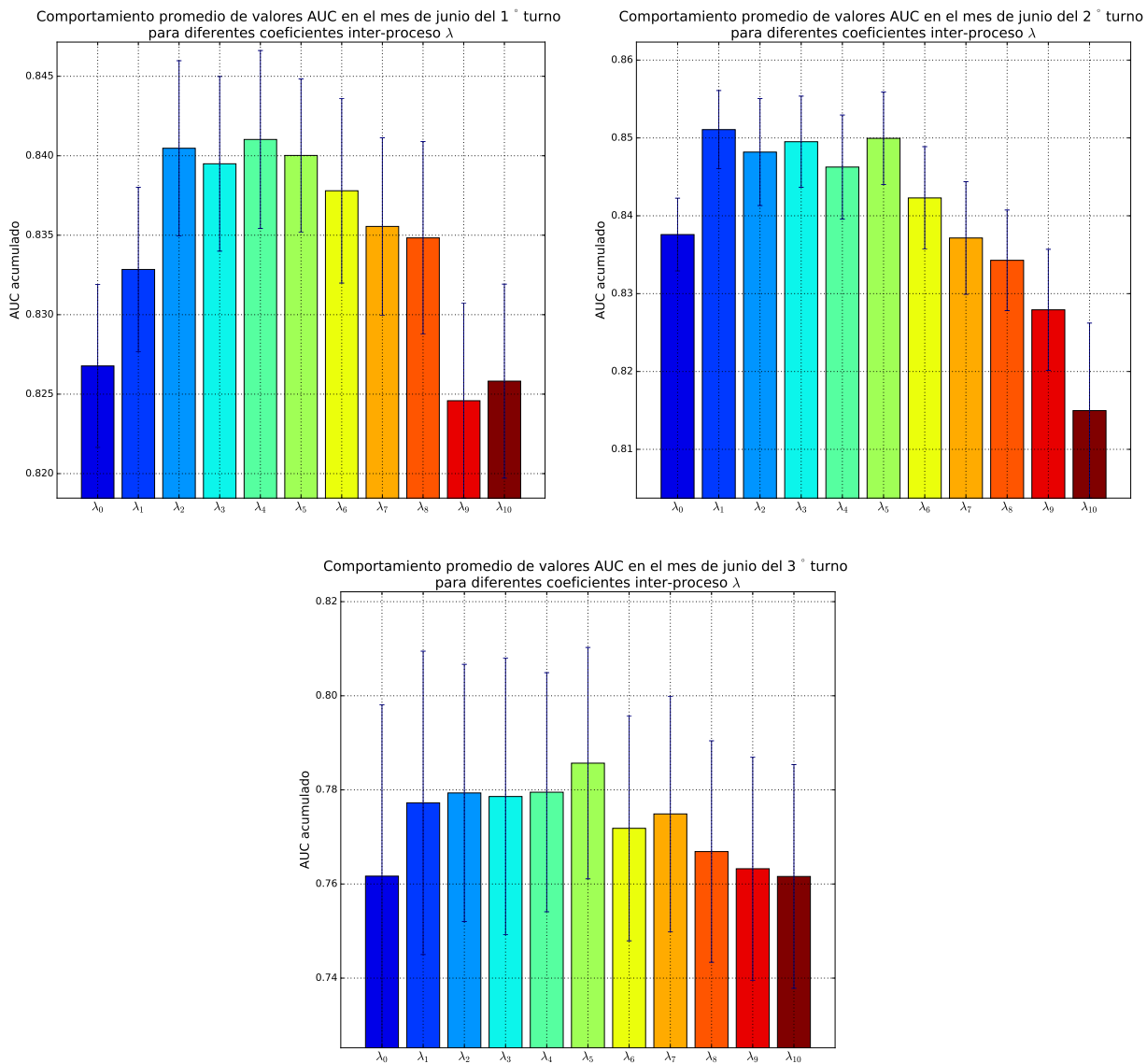


Figura D.2: Promedio de valores AUC diarios del primer, segundo y tercer turno utilizando los datos del mes de junio 2016 en la comuna de Maipú utilizando modelo KPG con el caso de estudio 0 para caracterizar el comportamiento temporal con diferentes coeficientes inter-proceso  $\lambda$ . En el eje de las abscisas se ordenan de mayor a menor los valores de  $\lambda$ , partiendo en 0 y terminando en 1, con cada paso de avance igual a 0,1. Las líneas horizontales entregan la varianza de los datos para cada caso de estudio.

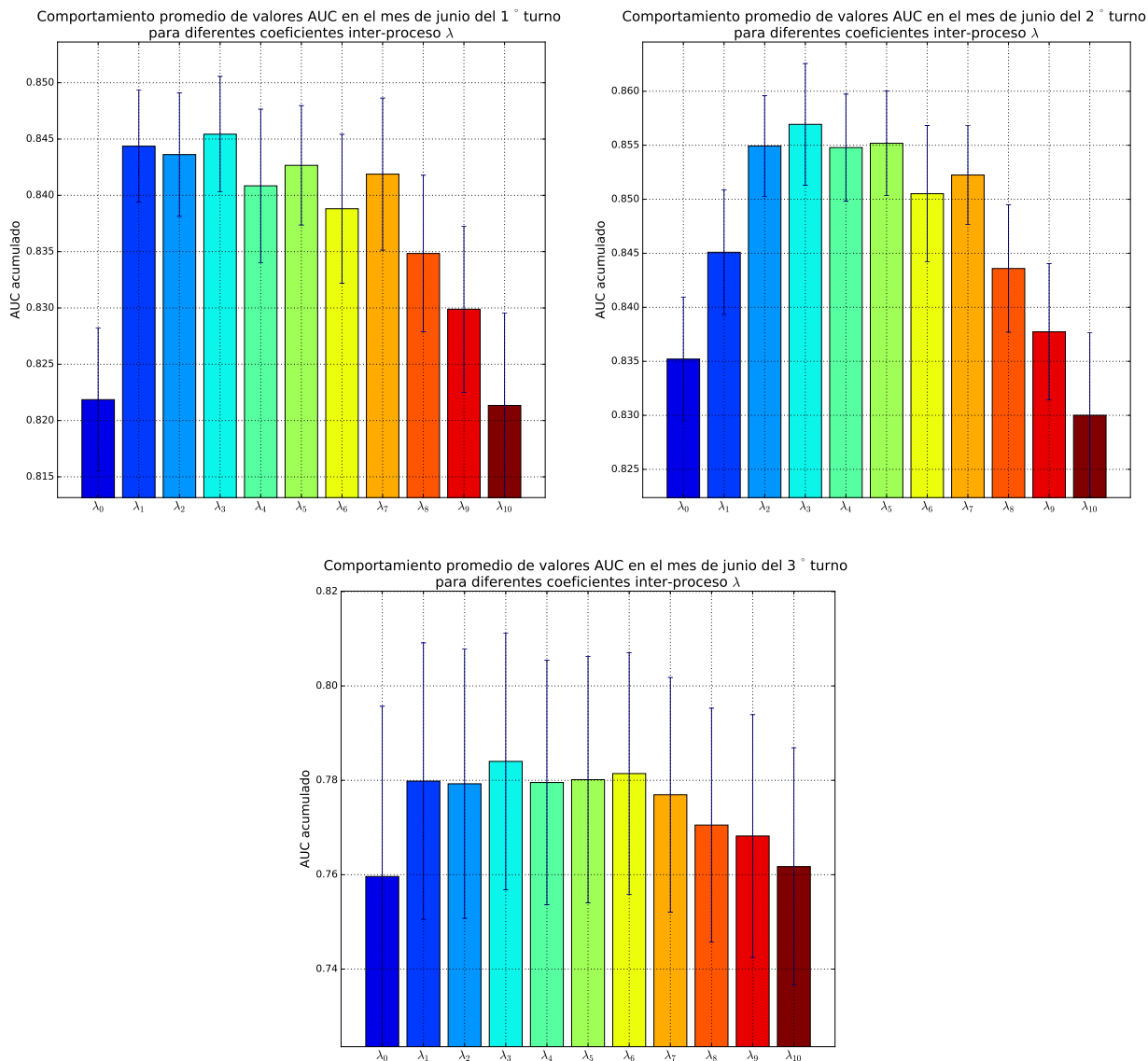


Figura D.3: Promedio de valores AUC diarios del primer, segundo y tercer turno utilizando los datos del mes de junio 2016 en la comuna de Maipú utilizando modelo KPG con el caso de estudio 2 para caracterizar el comportamiento temporal con diferentes coeficientes inter-proceso  $\lambda$ . En el eje de las abscisas se ordenan de mayor a menor los valores de  $\lambda$ , partiendo en 0 y terminando en 1, con cada paso de avance igual a 0,1. Las líneas horizontales entregan la varianza de los datos para cada caso de estudio.

# Capítulo E

## Curvas promedios para distintos valores AUC.

En este anexo se muestran las curvas promedio de diferentes casos HRP y TIP con valores AUC dentro de un intervalo determinado. Se generaron, de manera artificial, 10000 curvas HRP y TIP diferentes, cada una de ellas con 15 eventos. A cada una de las curvas se le aplica la métrica AUC para posteriormente ser filtradas según un rango de valores definido. Finalmente se calcula el comportamiento promedio en cada uno de los conjuntos filtrados, entregando como resultado de interés para este trabajo de tesis la Figura E.1.

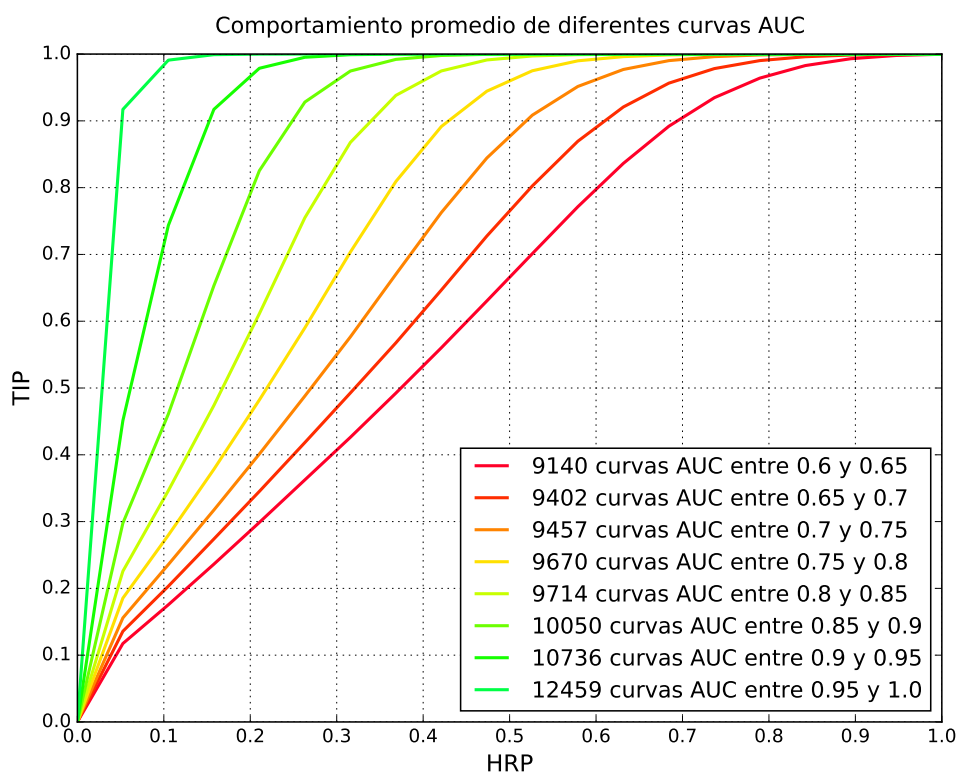


Figura E.1: Comportamiento promedio de la curva HRP-TIP con diferentes intervalos de valores AUC.