

# 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 ejecutado 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 tomo 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 las 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 el 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 las 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 las 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 <sub>θ</sub>	<i>High Risk Porcentage</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 <sub>q</sub> <sup>r</sup>	señal temporal con fuerte presencia de la escala $s_q$
SST <sup>r</sup>	señal de soporte temporal
STS	señal temporal suavizada por medio de KDE
STS <sup>r</sup>	señal temporal suavizada de una componente espacial $r$ arbitraria
TF	Transformada de Fourier
TIP <sub>θ</sub>	<i>True Incident Percentage</i>
TW	Transformada Wavelet