



**Universidad de Chile**  
**Facultad de Ciencias Físicas y**  
**Matemáticas**  
**Departamento de Geología**

**Evaluación de la susceptibilidad de remociones en masa en el sector nororiente de la Cuenca de Santiago mediante métodos estadísticos multivariados.**

Memoria para optar al título de Geólogo

Paulina Victoria Schachter Valdés

**Profesor Guía:**

Sergio Sepúlveda Valenzuela.

**Profesores Comisión:**

Julián Ortiz Cabrera.

Sofía Rebolledo Lemus.

Laura Piñero Feliciangeli.

Santiago de Chile  
2008



## RESUMEN

### “EVALUACIÓN DE LA SUSCEPTIBILIDAD DE REMOCIONES EN MASA EN EL SECTOR NORORIENTE DE LA CUENCA DE SANTIAGO MEDIANTE MÉTODOS ESTADÍSTICOS MULTIVARIADOS”

El sector oriente de la Cuenca de Santiago, de acuerdo a distintos estudios realizados, es especialmente propenso a la ocurrencia de fenómenos de remoción en masa. La gran mayoría de dichos estudios se focalizan principalmente en el sector entre el Río Mapocho por el Norte y el Río Maipo por el Sur. Sin embargo, hay pocos estudios realizados respecto de este fenómeno para los sectores que se encuentran tanto al Norte como al Sur de dicha área.

El principal objetivo de este estudio consistió en determinar la susceptibilidad del sector nororiente de la Cuenca de Santiago a la generación de procesos de remoción en masa, especialmente deslizamientos, utilizándose para ello métodos estadísticos multivariados. Otro objetivo fue apreciar, mediante distintos criterios, la aplicabilidad de dos de estos métodos estadísticos a este tipo de problema.

Se aplicaron los métodos de regresión logística y de análisis discriminante, al área comprendida entre el Río San Francisco por el Este, el canal El Carmen por el Oeste, el Río Mapocho por el Sur y el Estero Colina por el Norte.

Para desarrollar este estudio, se consideraron como variables independientes diversas características de la geología, la hidrografía, la topografía y la vegetación, incorporándose además de manera especial, variables representativas de la sismicidad en la zona. Estas variables se escogieron de entre las utilizadas en otros trabajos y considerando la información disponible para dichas variables en el área de interés. Por su parte, la variable dependiente representó la evidencia de deslizamientos en el área en estudio, o sea la presencia o ausencia de escarpes de deslizamiento. Los modelos se generaron a partir de las variables ya citadas, utilizando los programas SPSS v15.0 y R 2.6.2. Para cada modelo se generó también una representación gráfica, utilizando para ello el programa ArcGIS 9.2.

Los resultados obtenidos mostraron que, para el área en estudio, las variables más determinantes para explicar los deslizamientos son: las unidades geológicas, los lineamientos y las variables sísmicas, en particular aquéllas que representan las fuentes sismogénicas corticales. Además, se obtuvo que el método de regresión logística entrega un resultado gráfico mucho más acorde con los antecedentes disponibles que el método de análisis discriminante. En general, entre los dos métodos utilizados en este estudio, el que entregó resultados más apropiados y menos sesgados, es el de regresión logística.

Se concluye que este tipo de métodos estadísticos sirven para el estudio de problemas de carácter geológico pero que, sin embargo, tienen ciertas limitaciones en lo que se refiere a las variables que representan la geología del área. En efecto, dichas variables difícilmente cumplen con los supuestos básicos de dichos métodos. No obstante, cuando la escala permite una representación adecuada y el área de trabajo es lo suficientemente grande como para constituir una visión general del problema en estudio, los métodos sí resultan aplicables.

El estudio demostró también la existencia de una importante correlación entre la sismicidad cortical y los lineamientos, así como la significancia de ambas variables en la susceptibilidad a deslizamientos. Esto permite plantear una hipótesis de causalidad entre la sismicidad cortical y los deslizamientos.

**A mis padres**

## **Agradecimientos**

Agradezco a mi familia por el apoyo incondicional que me han brindado.

Agradezco a mi profesor guía, Sergio Sepúlveda, por sus valiosas sugerencias y correcciones, las que sin duda han contribuido a que este trabajo haya podido llegar a buen término, y también por haber estado siempre dispuesto a ayudarme y a brindarme el máximo apoyo posible.

Agradezco también a los profesores integrantes de la comisión, Sofía Rebolledo y Julián Ortiz por sus acertados comentarios e importantes sugerencias y correcciones.

También a Laura Piñero por la ayuda, apoyo y guía entregada desde tan lejos.

Agradezco además a María Rosa y a Cristina, por haberme dado ánimo en las situaciones más difíciles de mi carrera y por haber estado siempre dispuestas para darme su ayuda en todo momento.

Quiero agradecer también a Rodrigo Rauld, Felipe Leyton, Jaime Campos, Raúl Aguilera y a Andrés Pavez, por sus oportunos consejos los que me fueron de mucha utilidad.

Quiero agradecer especialmente a Rodrigo Assar, por haberme apoyado y ayudado en distintos momentos de este trabajo.

Por ultimo, agradezco a mis queridos compañeros de curso: Dani, Rodolfo, Carlos Garcia, Andrea, Karina, Mito, Melu, Seba, Gotito, Paulas, Chasca, Kathy, Chicho, Romi, Chala, Cristobal, y de carrera por su amistad y desinteresada colaboración.

# ÍNDICE

	Página
<b>CAPÍTULO I - INTRODUCCION</b>	<b>1</b>
1.1 Exposición del problema	1
1.2 Ubicación zona de estudio	3
1.3 Objetivos del estudio	4
1.3.1 Objetivo general	4
1.3.2 Objetivos específicos	4
1.4 Metodología general	4
1.5 Limitaciones	5
<b>CAPÍTULO II – DESCRIPCIÓN DE LA ZONA DE ESTUDIO</b>	<b>6</b>
2.1 Marco geomorfológico regional	6
2.2 Geología regional	7
2.2.1 Cordillera de la Costa	8
2.2.2 Cordillera Principal	9
2.3 Geomorfología del área de estudio	11
2.3.1 Quebrada El Arrayán	19
2.3.2 Quebrada San Francisco	21
2.3.3 Río Mapocho	23
2.3.4 Estero Colina	23
2.4 Geología del área de estudio	23
2.4.1 Geología del Cuaternario	24
2.4.2 Intrusivos	26
2.4.3 Rocas estratificadas	27
2.5 Depósitos de remoción en masa	30
2.6 Marco tectónico	33
2.7 Descripción de la Quebrada San Ramón	38
<b>CAPÍTULO III – FORMULACION TEORICA</b>	<b>41</b>
3.1 Aplicación de métodos estadísticos a estudios de susceptibilidad a la ocurrencia de remociones en masa	42
3.2 Análisis discriminante	44
3.3 Regresión logística	45
3.4 Variables independientes	47
3.5 Variables dependientes	48
<b>CAPÍTULO IV – RESULTADOS</b>	<b>49</b>
4.1 Metodología de selección de variables	50
4.1.1 Regresión logística univariable	50
4.1.2 Regresión por pasos hacia adelante y hacia atrás	50
4.2 Regresión logística	51
4.3 Validación en la Quebrada de San Ramón (regresión logística)	56
4.4 Análisis discriminante	57
4.5 Validación en la Quebrada de San Ramón (Análisis discriminante)	60
4.6 Resultados obtenidos por Lara (2007) y comparación	62
4.7 Comparación de resultados obtenidos por los métodos análisis discriminante y regresión logística	65

<b>CAPÍTULO V – RESULTADOS INCLUYENDO VARIABLES SÍSMICAS</b>	<b>68</b>
5.1 Regresión logística	71
5.1.1 Estadísticos de colinealidad	71
5.1.2 Tests de ajuste al modelo de los datos	71
5.1.3 Tablas de clasificación	72
5.1.4 Coeficientes	73
5.2 Análisis discriminante	76
5.2.1 Estadísticos de colinealidad	76
5.2.2 Homogeneidad de las matrices de covarianza	76
5.2.3 Igualdad de medias de grupo (Lambda de Wilks)	77
5.2.4 Coeficientes estandarizados	77
5.2.5 Matriz de estructura	77
5.2.6 Correlación canónica	78
5.2.7 Tablas de clasificación	78
5.3 Resultados gráficos de los métodos utilizados	79
5.4 Comparación de métodos de análisis discriminante y regresión logística para una probabilidad de corte de 0,1	80
5.5 Métodos de regresión logística y análisis discriminante considerando sólo las variables más importantes de acuerdo a cada método	82
5.5.1 Análisis discriminante	82
5.5.1.1 Correlación entre variables utilizadas	82
5.5.1.2 Estadísticos de colinealidad	82
5.5.1.3 Tests de homogeneidad de las matrices de covarianza	82
5.5.1.4 Prueba de igualdad de medias de los grupos	82
5.5.1.5 Coeficientes estandarizados de las funciones discriminantes canónicas	83
5.5.1.6 Matriz de estructura	83
5.5.1.7 Correlación canónica	83
5.5.1.8 Resultados de la clasificación	83
5.5.1.9 Resultados gráficos del modelo	83
5.5.2 Regresión logística	84
5.5.2.1 Tabla de correlación	84
5.5.2.2 Estadísticos de colinealidad	84
5.5.2.3 Tests de ajuste del modelo a los datos	84
5.5.2.4 Tabla de clasificación	84
5.5.2.5 Variables en la ecuación	85
5.5.2.6 Resultados gráficos del modelo	85
5.5.3 Comparación de resultados obtenidos con modelos creados con variables más importantes	86
 <b>CAPÍTULO VI – DISCUSIONES</b>	 <b>87</b>
 <b>CAPÍTULO VII – CONCLUSIONES Y RECOMENDACIONES</b>	 <b>95</b>
7.1 Conclusiones	95
7.2 Recomendaciones	96
 <b>REFERENCIAS</b>	 <b>97</b>

## **ANEXOS**

**En CD**

**ANEXO 1** - PRUEBAS DE NORMALIDAD

**ANEXO 2** – DESCRIPCIÓN DE VARIABLES UTILIZADAS

**ANEXO 3** – TEORÍA MÉTODOS REGRESIÓN  
LOGÍSTICA Y ANÁLISIS DISCRIMINANTE

**ANEXO 4** – TABLAS CAPÍTULO IV

**ANEXO 5** – TABLAS CAPÍTULO V

**ANEXO 6** – RESULTADOS GRÁFICOS MODELOS CAPÍTULO V

**ANEXO 7** – TABLAS DE CORRELACIÓN ENTRE  
VARIABLES INDEPENDIENTES PARA  
REGRESIÓN LOGÍSTICA Y ANÁLISIS  
DISCRIMINANTE.

**ANEXO 8** – COMPROBACIÓN DE RELACIONES  
LINEALES ENTRE EL LOGIT Y LAS VARIABLES  
INDEPENDIENTES PARA EL CAPÍTULO IV.

# ÍNDICE DE FIGURAS

Figura	Página
<b>Figura 1.1:</b> Área de estudio encerrada con línea azul. Imagen obtenida del programa NASA World Wind. Escala 1:400.000.	3
<b>Figura 2.1:</b> Rasgos geomorfológicos principales en Chile Central. Modificado de Fock (2005).	6
<b>Figura 2.2:</b> Columnas estratigráficas de la Cordillera de la Costa y de la Cordillera Principal en Chile Central. Modificado de Fock (2005).	8
<b>Figura 2.3:</b> Modelo de evolución paleogeográfica entre los 33° y 34°S. Tomado de Fock (2005). A continuación se describe brevemente cada etapa (también tomado de Fock, 2005).	
<ul style="list-style-type: none"> <li>- (a): Exhumación general de las rocas Cretácicas generándose la proto Cordillera de la Costa.</li> <li>- (b): Depositación de la Formación Abanico, junto con una gran actividad magmática básica. Durante este periodo se produce una fuerte subsidencia regional.</li> <li>- (c): Inicio de la inversión de la cuenca subsidente donde se depositó la Formación Abanico. Se deposita el miembro inferior de la Formación Farellones de manera concordante y sintectónica sobre la Fm. Abanico.</li> <li>- (d): Continúa el proceso de inversión de la cuenca y la deformación migra hacia el Este, concentrándose en el límite oriental de la cuenca. El engrosamiento cortical, que comenzó en el episodio anterior, prosigue y el Arco migra hacia el Este.</li> <li>- (e) y (f): La deformación se concentra en el límite oriental de la cuenca. Durante el Mioceno Superior el arco continúa su migración hacia el Este. Se produce el principal evento de alzamiento de superficie y exhumación.</li> </ul>	10
<b>Figura 2.4:</b> Modelo digital de terreno mostrando ubicación del área de estudio. En rojo como referencia se encuentra el Cerro San Cristóbal. Sistema de coordenadas UTM WGS-84 (zona 19S).	11
<b>Figura 2.5:</b> Zoom del área de estudio (mostrada en la figura 2.4) en donde se aprecia que ésta comprende la Depresión Central (zonas más planas y borde) y la Precordillera o zona más occidental de la Cordillera Principal. También se aprecian algunas características del terreno, tales como variaciones topográficas, hidrográficas, y los principales rasgos topográficos y geomorfológicos.	12
<b>Figura 2.6:</b> A la izquierda arriba se muestra un ejemplo de drenaje paralelo y a la derecha arriba de uno subparalelo. Abajo a la izquierda se muestra uno de tipo subdendrítico mientras que abajo a la derecha, se muestra uno de tipo centrípeto. Las flechas indican dirección del flujo (Sesser, 2006).	14
<b>Figura 2.7:</b> (a) Mapa hipsométrico; (b) Distribución de pendientes en el área de estudio; (c) distribución de aspecto (orientación de ladera) en el área de estudio; (d) Curvatura en planta; (e) Curvatura en perfil. Para las unidades en las que se encuentra la curvatura ver Anexo 2.	15 - 17
<b>Figura 2.8:</b> Perfil Este Oeste entre los puntos: 347980 y 377630 (coordenada Este), y 6317500 (coordenada Norte). Exageración vertical aproximadamente 1H es a 4V.	17
<b>Figura 2.9</b> Perfil diagonal en dirección noreste a suroeste entre los puntos 1: 370890/6326930 y 2: 350580/6305570. Exageración vertical aproximadamente 1H a 4V.	18
<b>Figura 2.10:</b> Distribución de aspecto, tomado de la figura (2.7 (c)), con NDVI sobrepuesto. Las zonas más claras corresponden a áreas con mayor vegetación.	18
<b>Figura 2.11:</b> Perfil topográfico a lo largo del cauce del Estero El Arrayán. La figura muestra los cambios de pendiente a lo largo de la Quebrada El Arrayán el cual fue obtenido a partir de la base de datos topográfica OTAS (escala 1:50.000).	19
<b>Figura 2.12:</b> Mapa topográfico de la subcuenca del Estero el Arrayán, en donde se pueden ver los principales elementos topográficos y afluentes. Fuente: base topográfica OTAS 1:50.000.	20
<b>Figura 2.13:</b> Perfil topográfico a lo largo del cauce del Río San Francisco. La figura muestra los cambios de pendiente a lo largo del Río obtenido a partir de la base de datos topográficos OTAS (escala 1:50.000).	21
<b>Figura 2.14:</b> Mapa topográfico de la subcuenca del Río San Francisco en donde se pueden ver los principales elementos topográficos y afluentes. Fuente base topográfica OTAS 1:50.000.	22

<b>Figura 2.15:</b> Mapa geológico área en estudio. Modificado de Wall et al. (1999), mapa geológico de Chile (1:1.000.000), Thiele (1980) y trabajo de terreno. Ver simbología en mapa fuera de texto.	23
<b>Figura 2.16:</b> Depósitos fluviales antiguos Río San Francisco. Se puede ver claramente la imbricación de los clastos más grandes, además de un alto grado de redondeamiento (Vista al Este).	24
<b>Figura 2.17:</b> Vista hacia el sureste en Quebrada El Arrayán, donde se pueden observar manteos hacia el Este	28
<b>Figura 2.18:</b> Vista hacia el noreste en quebrada el Arrayán. En el fondo se ve la Formación Abanico manteando hacia el Este.	28
<b>Figura 2.19:</b> Brecha volcánica fuertemente erodada, perteneciente a la Formación, Abanico ubicada en la ladera oeste de la Quebrada San Francisco.	29
<b>Figura 2.20:</b> Vista al Sur desde cuchilla de Cerro Alto el Olivillo. Se observa pequeño deslizamiento en ladera Oeste del Río San Francisco. La ladera cae hacia el Este.	30
<b>Figura 2.21:</b> Complejo de Remociones observado en la Quebrada de San Francisco en la ladera Oeste (Fotografía tomada mirando al Oeste). Abajo el mismo complejo de remociones visto desde el camino a Farellones	30
<b>Figura 2.22:</b> (a) y (b) Corresponden a depósitos de remoción en masa observados en cortes de ladera en la Hacienda Santa Martina. (c) Imagen obtenida de Nasa World Wind, mostrando la remoción completa. A la derecha se ve el Estero El Arrayán. La flecha roja indica el Norte.	31
<b>Figura 2.23:</b> Morfología aterrazada que se puede ver en la Hacienda Santa Martina. (a) Panorámica en donde se pueden observar los tres niveles. (b) Detalle del nivel más bajo en donde se resalta la superficie ondulosa (pero muy plana a gran escala).	32
<b>Figura 2.24:</b> Perspectiva virtual al noreste. Se aprecia la estructura formada por los cerros Calán, Apoquindo y Los Rulos, y que es interpretada como la Falla de San Ramón. (Tomado de Rauld, 2002).	33
<b>Figura 2.25:</b> Imagen del área de estudio obtenida con el programa NASA World Wind, en donde se distingue el fuerte rasgo lineal que presentan el Estero El Arrayán y El Río San Francisco.	34
<b>Figura 2.26:</b> Fallas y lineamientos presentes en el área en estudio. La imagen de fondo es la banda pancromática de la imagen satelital LANDSAT 7 ETM.	35
<b>Figura 2.27:</b> Representación gráfica de la densidad de lineamientos en el área de estudio, extraído de la figura 2.26. La unidad de área es 2.500m <sup>2</sup> (píxel de 50m por 50m).	36
<b>Figura 2.28:</b> Representación gráfica de la distancia euclidiana a los lineamientos en el área en estudio, extraída de la figura 2.26. Unidad de área es 2500m <sup>2</sup> .	37
<b>Figura 2.29:</b> Geología de la quebrada San Ramón (Lara, 2007).	38 – 39
<b>Figura 4.1:</b> Mapa geológico con escarpes de remociones consideradas en el área en estudio marcadas en café. Se observa que la gran mayoría de los escarpes (en color café) se ubican en la Formación Abanico. Tomado y modificado de Wall et al. (1999), mapa geológico de Chile (1:1.000.000), Thiele (1980) y trabajo de terreno.	52
<b>Figura 4.2:</b> Escala de probabilidad que representa la susceptibilidad a la ocurrencia de remociones en masa en el área en estudio, obtenida mediante el método de regresión logística en la zona en estudio. Arriba susceptibilidad a la ocurrencia de remociones en masa sin relieve de fondo, abajo susceptibilidad a la ocurrencia de remociones en masa con relieve de fondo y escarpes.	55
<b>Figura 4.3:</b> Escala de probabilidad que representa la susceptibilidad a la ocurrencia de remociones en masa obtenidas mediante el método regresión logística en la Quebrada San Ramón. Arriba susceptibilidad a la ocurrencia de remociones en masa sin relieve de fondo, abajo susceptibilidad a la ocurrencia de remociones en masa con relieve de fondo.	56
<b>Figura 4.4:</b> Escala de probabilidad que representa la susceptibilidad a la ocurrencia de remociones en masa en el área de estudio obtenida mediante el método análisis discriminante en la zona de estudio. Arriba sin relieve de fondo, abajo con relieve de fondo y escarpes.	59
<b>Figura 4.5:</b> Escala de probabilidad que representa la susceptibilidad a la ocurrencia de remociones en masa obtenidas mediante el método análisis discriminante en la Quebrada de San Ramón. Abajo con relieve de fondo y arriba sin relieve de fondo.	60
<b>Figura 4.6:</b> a) densidad de estructuras en la Quebrada de San Ramón; b) Distancia euclidiana a las estructuras en la Quebrada de San Ramón	61
<b>Figura 4.7:</b> Índice de susceptibilidad a deslizamientos en Roca obtenido por Lara (2007) en la Quebrada de San Ramón.	63
<b>Figura 4.8:</b> Índice de susceptibilidad para deslizamientos en suelo obtenido por Lara (2007) en la Quebrada de San Ramón.	64

<b>Figura 4.9:</b> En blanco se encuentran los sectores con susceptibilidad a la ocurrencia de remociones en masa menor que 0,1, y en verde aquellas áreas con susceptibilidad mayor a 0,1. A la izquierda se encuentran los resultados obtenidos con el método de regresión logística, y a la derecha los obtenidos con el método de análisis discriminante.	65
<b>Figura 4.10:</b> En blanco se encuentran los sectores con susceptibilidad a la ocurrencia de remociones en masa menor a 0,04 en el caso de la regresión logística (a la izquierda) y menor a un 0,19 para el caso del análisis discriminante (a la derecha). En verde se muestran aquellas áreas cuya probabilidad supera la susceptibilidad definida como corte por cada método.	66
<b>Figura 4.11:</b> En blanco se encuentran los sectores con susceptibilidad a la ocurrencia de remociones en masa menor a 0,1, y en verde aquellas áreas que superan la susceptibilidad de 0,1. A la izquierda se encuentran los resultados en la Quebrada de San Ramón para el método de regresión logística y a la derecha del análisis discriminante.	67
<b>Figura 5.1:</b> Perfil con la sismicidad de la zona, a la altura de 33.5°S. Cada punto representa un sismo, independiente de su magnitud o fecha de ocurrencia. El color azulado representa la posición esquemática de la placa de Nazca y, el verde, la placa Sudamericana. Las líneas punteadas indican la posición de las principales fuentes sismogénicas: (a) interplaca tipo thrust, (b) interplaca de profundidad intermedia, (c) cortical y (d) <i>outer-rise</i> (no considerado en este trabajo).	68
<b>Figura 5.2:</b> Modelos que incluyen variable que representa la Intensidad de Mercalli Modificada para fuentes corticales (IMM3). Arriba, resultado regresión logística. Al medio, resultados regresión logística y análisis discriminante con escarpes sobrepuestos. Abajo, resultado análisis discriminante.	80
<b>Figura 5.3:</b> En blanco se encuentran los sectores con susceptibilidad a la ocurrencia de remociones en masa menor a 0,1, y en verde aquellas áreas con susceptibilidad mayor o igual a 0,1. A la izquierda se encuentran los resultados obtenidos con el método de regresión logística, y a la derecha, los obtenidos con el método de análisis discriminante para los modelos que incluyen la variable Intensidad de Mercalli Modificada para fuentes corticales (IMM3).	81
<b>Figura 5.4:</b> Escala de probabilidad que representa la susceptibilidad a la ocurrencia de remociones en masa en el área en estudio, obtenida mediante el modelo basado en el método de análisis discriminante, considerando sólo las variables más importantes.	84
<b>Figura 5.5:</b> Escala de probabilidad que representa la susceptibilidad a la ocurrencia de remociones en masa en el área en estudio, obtenidas mediante el método de regresión logística y considerando sólo las variables más importantes.	85
<b>Figura 5.6:</b> Susceptibilidad a la ocurrencia de remociones en masa. En blanco se encuentran los sectores con susceptibilidad menor a 0,1, y en verde aquellas áreas con susceptibilidad igual o mayor que 0,1. A la izquierda se encuentran los resultados obtenidos con el modelo basado en el método de regresión logística, y a la derecha los resultados obtenidos con el modelo basado en el método de análisis discriminante.	86
<b>Figura 6.1:</b> Distribución de susceptibilidades a la ocurrencia de remociones en masa para el método de regresión logística.	93
<b>Figura 6.2:</b> Distribución de susceptibilidades a la ocurrencia de remociones en masa para el método de análisis discriminante	93

# ÍNDICE DE TABLAS

Tabla	Página
<b>Tabla 4.1:</b> Variables consideradas en este trabajo y sus respectivos códigos	49
<b>Tabla 4.2:</b> $R^2$ de Cox y Snell y de Negelkerke.	53
<b>Tabla 4.3:</b> Tabla de clasificación con probabilidad de corte 0.1 que muestra el número de píxeles clasificados como casos positivos para probabilidades iguales o superiores a 0,1, o zonas altamente susceptibles, y casos negativos o baja susceptibilidad para probabilidades menores a 0,1, para el modelo mostrado en la tabla A4.7.	53
<b>Tabla 4.4:</b> Coeficientes de variables (en negritas).	54
<b>Tabla 4.5:</b> Capacidad discriminante de función canónica.	57
<b>Tabla 4.6:</b> Tabla de clasificación para el método análisis discriminante con probabilidades a priori iguales para ambos grupos (se asume una probabilidad de 0,5 para cada grupo).	58
<b>Tabla 4.7:</b> Comparación entre métodos y casos reales para una probabilidad de corte de 0,1. P se refiere a susceptibilidad a la ocurrencia de remociones en masa, RL: regresión logística, y AD: análisis discriminante.	65
<b>Tabla 4.8:</b> Comparación en terminas de área y probabilidades de corte de los resultados de la validación en este trabajo con lo obtenido por Lara (2007).	66
<b>Tabla 5.1:</b> Variables sísmicas utilizadas	70
<b>Tabla 5.2:</b> Variables sísmicas utilizadas en los casos considerados en este trabajo.	70
<b>Tabla 5.3:</b> $R^2$ de Cox y Snell y de Negelkerke.	72
<b>Tabla 5.4:</b> Resultados de los modelos en términos de porcentaje de casos correctamente clasificados por grupo y total para una probabilidad de corte de 0,1.	73
<b>Tabla 5.5:</b> Coeficientes obtenidos para todos los casos mediante el método de regresión logística.	74
<b>Tabla 5.6:</b> Correlaciones canónicas por modelo.	79
<b>Tabla 5.7:</b> Porcentaje de casos correctamente clasificados por grupo y en total para el método análisis discriminante.	79
<b>Tabla 5.8:</b> Comparación de métodos de análisis discriminante y regresión logística, para una probabilidad de corte de 0,1, en términos de número de casos o píxeles para todos los modelos vistos en este capítulo. RL: regresión logística, AD: análisis discriminante.	81
<b>Tabla 5.9:</b> Correlación canónica entre el modelo y la función discriminante.	83
<b>Tabla 5.10:</b> Tabla que muestra los resultados de la clasificación mediante el método de análisis discriminante.	83
<b>Tabla 5.11:</b> Tabla de clasificación para el modelo obtenido con el método de regresión logística para una probabilidad de corte de 0,1.	85
<b>Tabla 5.12:</b> Coeficientes de las variables y su importancia dentro del modelo.	85
<b>Tabla 5.13:</b> Comparación de resultados entre modelos y casos reales para una probabilidad de corte de 0,1. Donde P es la susceptibilidad a la ocurrencia de remociones en masa, RL: regresión logística, AD: análisis discriminante.	86

# **Capítulo I: Introducción**

## **1.1 Presentación del problema**

Las condiciones geológicas, geomorfológicas y climáticas, presentes en Chile, hacen que el país se vea constantemente afectado por remociones en masa a lo largo de su territorio. Hasta el momento, estos eventos de remoción en masa no han sido los principales causantes de pérdidas humanas y económicas para Chile, dado que, en su mayoría, ocurren en zonas cordilleranas más bien aisladas. Sin embargo, existen algunas excepciones que se destacan por su magnitud y por el daño causado, como por ejemplo: los flujos de detritos ocurridos en 1991 en Antofagasta (Hauser, 1997) y los eventos ocurridos en 1993 en el Sector Oriente de Santiago (Naranjo & Varela, 1996; Sepúlveda et al., 2006).

La Región Metropolitana presenta una gran cantidad de ejemplos de remociones en masa, especialmente en los sectores cordilleranos. Sin embargo, y a pesar de la alta tasa de ocurrencia de dichas remociones en masa, los habitantes de la Región Metropolitana no se habían visto afectados por este tipo de eventos sino hasta el explosivo crecimiento demográfico, ocurrido durante la segunda mitad del siglo XX, crecimiento que se manifestó de manera importante en las comunas más alejadas de la ciudad de Santiago centro, entre las que se encuentran Lo Barnechea, Las Condes, Vitacura, y La Florida, y que se caracterizan por o encontrarse muy cerca o en sectores cordilleranos propiamente tales (Antinao et al., 2003 (b)). Esta situación introdujo a las remociones en masa como un riesgo cierto y como factor a considerar para determinar la ubicación de futuros asentamientos humanos.

Las remociones en masa afectan de manera ocasional a las inmediaciones de la Cuenca de Santiago y de manera frecuente a los sectores cordilleranos, siendo principalmente del tipo flujo de detritos y, en menor proporción, deslizamientos. Generalmente, estos eventos se ven asociados a fuertes precipitaciones o a sismos, entre otros factores, así como también a fenómenos climáticos como el desplazamiento en altura de la isoterma cero, actuando principalmente en las partes altas de las hoyas hidrográficas y sobre acumulaciones inestables en las laderas, incluyendo los extensos abanicos aluviales que se pueden encontrar en todo el borde de la cuenca, especialmente en el borde oriental de Santiago (Antinao et al., 2003 (a)).

El sector de Santiago Oriente es especialmente propenso a la ocurrencia de este tipo de fenómenos, es por este motivo que se pueden encontrar tanto eventos recientes como antiguos, y en particular deslizamientos. Si bien hay diversos estudios acerca de la susceptibilidad o peligro de ocurrencia de remociones en masa en la Región Metropolitana como tal, la gran mayoría de éstos se refieren al sector oriente de la Cuenca de Santiago, especialmente al sector entre los ríos Mapocho y Maipo (Lara, 2007; Padilla, 2006; Sepúlveda, 1998). No se encuentran, sin embargo, este tipo de estudios para el sector nororiente o suroriente de la Cuenca de Santiago.

En función de la cantidad de ejemplos observables en el terreno, se escogió el sector nororiente de la Cuenca de Santiago como área de estudio. La cantidad y la distribución de remociones observables en dicha área permite realizar un mejor análisis estadístico, a partir del cual determinar la relación de dichos eventos con cada uno de los factores que pueden ser considerados como sus factores condicionantes, lo que a futuro permitiría poder predecir de mejor manera la ocurrencia de deslizamientos en las inmediaciones de la Cuenca de Santiago.

Los estudios geofísicos realizados por diversos investigadores, se basan principalmente en la respuesta sísmica, habiéndose descubierto que el frente cordillerano o el borde de la cuenca propiamente tal del área de estudio presenta una respuesta sísmica de buena a muy buena principalmente por su constitución rocosa. Sin embargo, dicha respuesta sísmica disminuye en calidad hacia el Oeste (von Igel et al., 2004).

También de acuerdo a von Igel et al. (2004), dentro de estas áreas con buena (o muy buena) respuesta sísmica se encuentran otras áreas que presentan una respuesta sísmica regular, en cuanto a la transmisión de la energía, las que corresponden principalmente a abanicos aluviales y coluviales (antiguos o contemporáneos) que poseen matriz arcillo-arenosa, lo que, si bien, disminuye fuertemente la respuesta sísmica, aumenta la probabilidad de ocurrencia de remociones en masa. Por lo expuesto, uno de los factores a incluir dentro de este estudio es la sismicidad, la que ha sido considerada como uno de los factores condicionantes más importantes, por ser Chile un país eminentemente sísmico. Otros factores que se consideraron fueron la topografía, la geología, la hidrografía, la vegetación y los lineamientos.

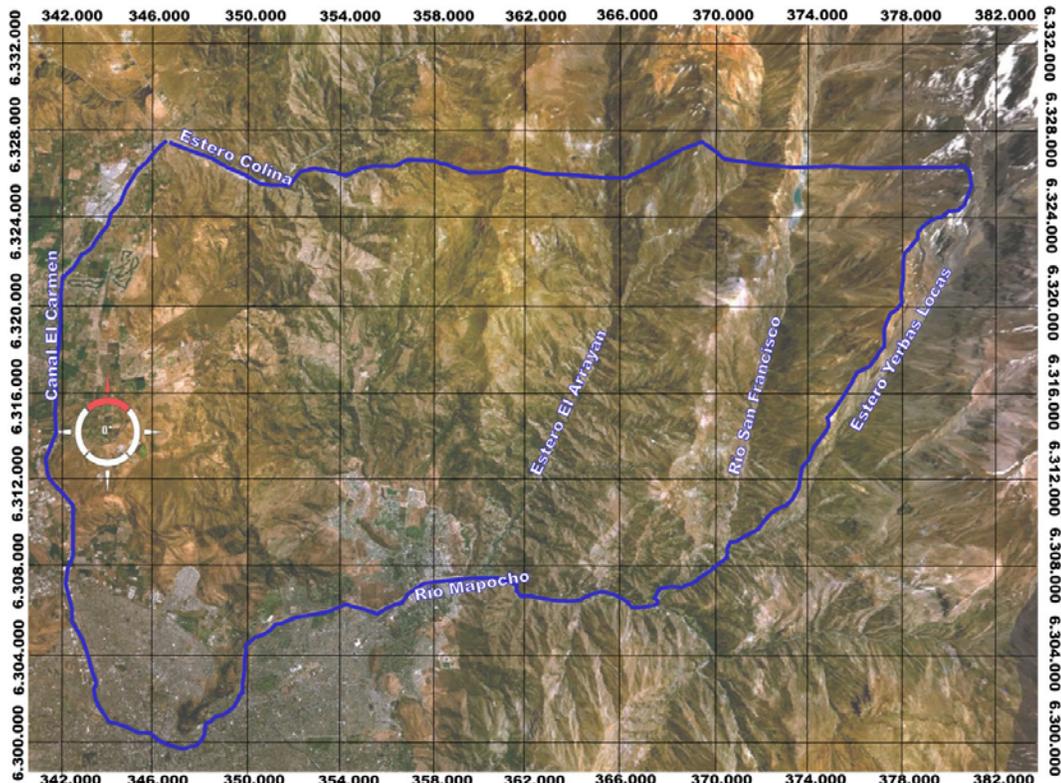
Actualmente se aplican técnicas estadísticas y de sistemas de información geográfica (GIS por sus siglas en inglés), con las que se intenta cuantificar la susceptibilidad o el peligro de ocurrencia de una remoción en masa. Además, la utilidad de la fotogeología (interpretación de fotos aéreas) y de las imágenes satelitales, es cada vez más reconocida y utilizada para la creación de bases de datos, a partir de las cuales se intenta obtener un modelo que sirva para predecir y estimar la probabilidad de ocurrencia de un evento de dicha naturaleza (Alexander, 2007), los que se basan en la importancia de la geomorfología como factor condicionante de una remoción en masa (Carrara, 2007).

Las técnicas estadísticas más aplicadas en la actualidad son la regresión logística y el análisis discriminante (Baeza, 1994; Chung et al., 1995; Gorsevski et al., 2000; Santacana et al., 2003; Lee, 2004; Ayalew & Yamagishi, 2005; Greco, et al. 2006; Nandi & Shakoor, 2006). Las que han demostrado su potencial para este tipo de estudios.

## 1.2 Ubicación del área de estudio

El área de estudio corresponde al frente precordillerano ubicado en el sector nororiente de la Cuenca de Santiago. En términos de sus coordenadas UTM (en sistema de referencia WGS – 84 zona 19°S), el área de estudio se ubica entre las coordenadas este 342000 - 378000 y norte: 6300000 - 6326000) (Figura 1.1), las que definen un área que va desde el río Mapocho hasta el Estero Colina, en sentido sur - norte, y desde el Canal El Carmen hasta el Río San Francisco en el sentido oeste - este.

Por ubicarse contigua a una importante zona urbana, el área de estudio cuenta con una buena cantidad de vías de acercamiento, principalmente al extremo occidental del área de estudio. Sin embargo, existen pocos caminos, la mayoría no pavimentados que permiten ingresar al área y que son generalmente de propiedad privada, por lo que el ingreso está restringido.



**Figura 1.1:** Área de estudio encerrada con línea azul. Imagen obtenida del programa NASA World Wind. Escala 1:400.000.

## **1.3 Objetivos del estudio**

### **1.3.1 Objetivo general**

Evaluar la susceptibilidad del sector nororiente de la Cuenca de Santiago a la generación de procesos de remoción en masa, mediante métodos estadísticos multivariados.

### **1.3.2 Objetivos específicos**

1. Levantar un catastro de remociones en masa en el área de estudio, para determinar los principales tipos de remociones en masa en la zona.
2. Identificar los factores condicionantes de las remociones en masa en la zona de estudio mediante visitas a terreno e instrumentos de percepción remota.
3. Determinar cuáles son los factores condicionantes más importantes mediante métodos estadísticos multivariados.
4. Determinar la susceptibilidad a la generación de remociones en masa en el sector nororiente de la cuenca de Santiago, informando su distribución espacial.
5. Determinar el efecto de la sismicidad en la ocurrencia de remociones en masa.
6. Determinar si un modelo generado por métodos estadísticos multivariados en un área específica para estudios de susceptibilidad de remociones en masa es aplicable en un área distinta.

## **1.4 Metodología general**

Para la realización de un catastro de remociones en masa en el área de estudio se procedió a la recopilación de antecedentes (informes previos), y al estudio de imágenes aéreas a escalas 1 a 50.000 y 1 a 70.000, para identificar las remociones existentes en la zona (antiguas o recientes), cuyo tamaño sea suficiente para ser mapeables a una escala de 1:100.000, esto es que involucre un área mayor a 2500m<sup>2</sup>, determinando la ubicación espacial de los deslizamientos existentes en el sector.

Para la verificación del catastro e identificación de los posibles factores condicionantes se realizaron 5 visitas a terreno, se realizó una investigación bibliográfica acerca del tema, y se buscó información de distintas bases de datos procedentes de sensores remotos y de la base de datos topográficos, OTAS, en escala 1 a 50.000.

Para determinar cuáles son los factores condicionantes más importantes, se procedió a la creación de una base de datos en el programa ArcGIS 9.2, para luego asignarle un peso a cada uno de dichos factores mediante métodos estadísticos multivariados, tales como la regresión logística y el análisis discriminante, utilizando los programas estadísticos SPSS v15.0, y R 2.6.2 ver anexo 3.

Aplicados dichos programas estadísticos, se procedió a la implementación de los resultados obtenidos en el programa ArcGIS 9.2, para poder obtener información gráfica acerca de la susceptibilidad espacial a la ocurrencia de remociones en masa del sector en estudio.

Las variables sísmicas utilizadas fueron la Intensidad de Mercalli Modificada (IMM) y la aceleración máxima horizontal, las que fueron obtenidas de distintas bases de datos de sismicidad para Chile central, como por ejemplo catálogo NEIC, catálogo SISRA, entre otros (Leyton et al. (2008)).

A continuación se repitieron los pasos indicados en el tercer y cuarto párrafo precedentes, para ver que efecto tenían las variables sísmicas en los modelos, y así juzgar su importancia.

Para determinar la capacidad predictiva de los modelos generados a partir de las variables del área de estudio por los métodos estadísticos multivariantes, en un área distinta a la de estudio, se escogió la Quebrada de San Ramón, por haber sido estudiada previamente por Lara (2007) pero con un método cualitativo, y por lo tanto por encontrarse los resultados disponibles para un análisis comparativo.

Por último, sobre la base de los resultados obtenidos, se concluye acerca de las bondades de los métodos utilizados para determinar la susceptibilidad del terreno a la ocurrencia de remociones en masa.

## **1.5 Limitaciones**

La principal limitación se refiere a la imposibilidad de contar con información geográfica en una escala menor que 1:50.000, es decir, con una resolución mejor que 50m, dado que esta escala suaviza el relieve del terreno.

También hay problemas asociados a la escala en la que se encuentran los datos sísmicos, ya que ésta se encuentra disponible a una escala de 1:1.000.000.

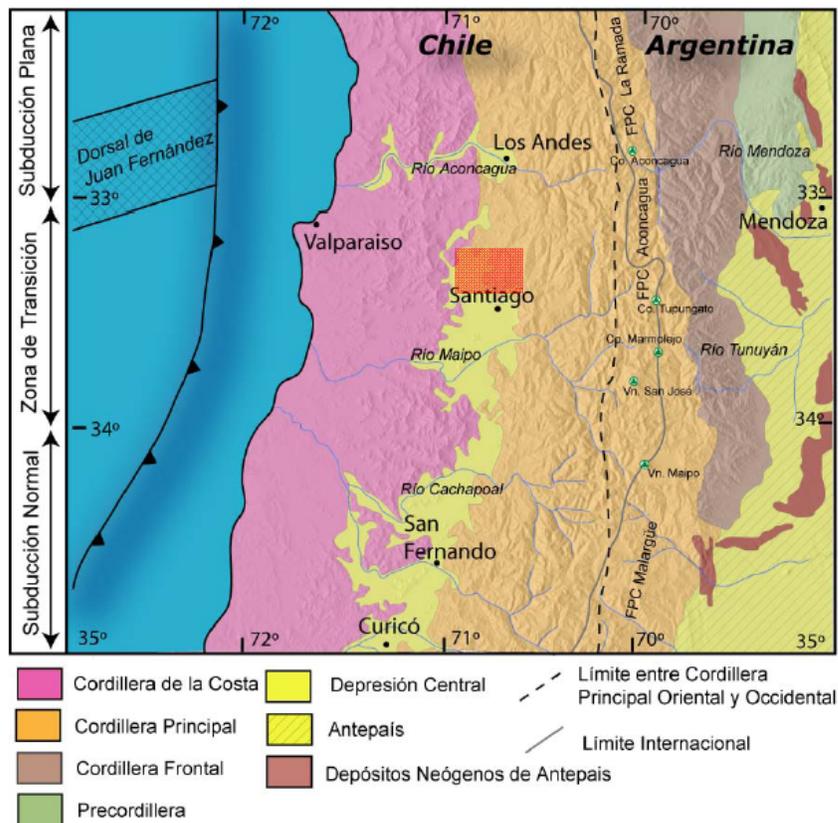
Las otras limitaciones se refieren a las dificultades de acceso al área en estudio, principalmente por:

- 1) Caminos cerrados por ser de propiedad privada
- 2) Caminos cerrados por pertenecer a propiedades mineras.
- 3) Lugares de difícil acceso (demasiada vegetación, muy escarpados, etcétera)
- 4) Caminos sin mantención después del último invierno, por lo que resulto imposible acceder o continuar por ellos.

## Capítulo II: Descripción del área de estudio

### 2.1 Marco Geomorfológico Regional

La zona de estudio (figuras 2.1 y 2.4) se encuentra en el sector nororiente de la Cuenca de Santiago, en el límite que separa la depresión central del frente cordillerano, dentro de la Cordillera Principal Occidental (Fock, 2005). El área está delimitada por: el Río Mapocho por el Sur, el Estero Colina por el Norte, el sector de Chicureo por el Oeste y la ribera Este del Río San Francisco (Cordón de la Yerba Loca).



**Figura 2.1:** Rasgos geomorfológicos principales en Chile Central. Modificado de Fock (2005).

Tal como se muestra en la figura 2.1 en la zona de Chile central se presentan tres unidades geomorfológicas: Cordillera de la Costa, Depresión Central y Cordillera Principal. Estas unidades se encuentran orientadas en dirección Norte-Sur y van de Oeste a Este

La Cordillera de la Costa, que rara vez sobrepasa los 2000 m s.n.m, está compuesta por granitoides del Paleozoico Superior en su sector occidental, granitoides Jurásicos en su zona central, las Formaciones volcano - sedimentarias Jurásicas y Cretácicas, y los intrusivos Cretácicos en el sector oriental. Además, destaca por su topografía suave en comparación a la de la Cordillera Principal. El relleno Cuaternario, por su parte, es principalmente de origen fluvial y coluvial, mientras que el sector costero se caracteriza por la presencia de terrazas de abrasión y de depósitos de remoción en masa.

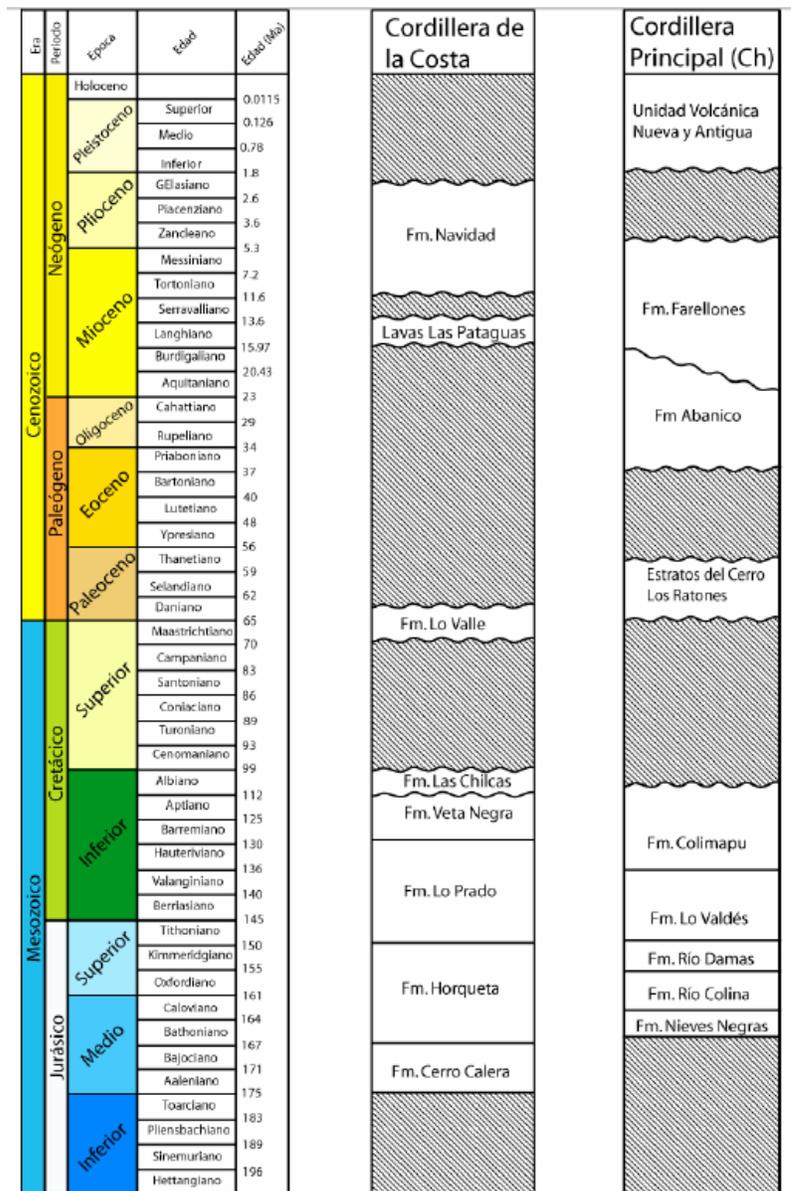
La Depresión Central se encuentra limitada al Norte por el Cordón de Chacabuco y al Sur por los cerros de Paine (este es el límite en la Región Metropolitana, pero este rasgo geomorfológico se extiende hacia el Sur). Se caracteriza por presentar una topografía extremadamente suave, que se ve interrumpida por la presencia de algunos cerros isla (en general oscila entre los 450 m s.n.m y los 550 m s.n.m, según base topográfica OTAS). Esta suave topografía se debe a que la cuenca se encuentra rellena por depósitos Cuaternarios los que, en algunos lugares, puede llegar hasta los 500m de espesor (Araneda et al., 2000), predominando los depósitos Cuaternarios de origen fluvial, coluvial y aluvial principalmente en el borde del sector oriental de la Cuenca de Santiago que corresponde al límite con el frente cordillerano.

La Cordillera Principal, por último, es la unidad geomorfológica más importante y representa un cambio abrupto, en cuanto a la topografía, ya que su altura va desde los 500 m s.n.m y superando en algunos casos los 5000 m s.n.m (Volcán San José) en dirección al Este. En la Cordillera Principal se originan los depósitos que se encuentran en la Depresión Central, los que han sido trasladados por los ríos Mapocho y Maipo. Los depósitos cuaternarios en la Cordillera Principal son de origen fluvial, coluvial o de remociones en masa.

## **2.2 Geología Regional**

El marco geológico y tectónico de Chile central se encuentra condicionado por la presencia de la subducción de la placa oceánica bajo la corteza continental, fenómeno que viene ocurriendo al menos desde el Jurásico Inferior y que ha influido principalmente en la ubicación de los arcos volcánicos y en la presencia de cuencas de trasarco, de antearco, entre otros.

La figura 2.2, a continuación, muestra dos columnas estratigráficas, una en la Cordillera de la Costa y otra en la Cordillera Principal en su lado occidental. Estas columnas corresponden a las unidades de roca estratificada cuya descripción se puede encontrar en fuentes tales como Hoja Santiago (Thiele, 1980); Hoja Til Til Santiago (Wall et al., 1999); y Mapa Geológico del área de Valparaíso – Curacavi (Gana et al., 1996).



**Figura 2.2:** Columnas estratigráficas de la Cordillera de la Costa y de la Cordillera Principal en Chile Central. Modificado de Fock (2005).

### 2.2.1 Cordillera de la Costa

Dentro de la Cordillera de la Costa, en su sector más occidental, se encuentran registros de edad Paleozoica, los que corresponden principalmente a intrusivos tonalíticos, granodioríticos, monzograníticos y graníticos, reconocidos como Unidad Mirasol por Gana (1994) (Gana et al. 1996), y que, además, forma parte del batolito de la costa pero no corresponde al arco volcánico de esta época, que se encontraba más al Este. No han sido encontrados registros correspondientes al Triásico.

A comienzos del Jurásico, se reinicia la subducción en el borde occidental de la Placa Sudamericana, por lo que en este período reaparece el arco volcánico, el cual se instala a lo largo de lo que hoy se conoce como Cordillera de la Costa. Los depósitos provenientes del arco volcánico del Jurásico son los que conforman (en su mayoría) el sector centro y oriental de la Cordillera de la Costa. Los representantes de esta etapa son: Fm. Ajiál, Fm Cerro Calera, y Fm. Horqueta, las que corresponden tanto a secuencias sedimentarias marinas como volcánicas (volcanismo no explosivo), de

carácter transicional a continental. Además, se encuentra una extensa área ocupada por los intrusivos Jurásicos de litología variable, entre gabros y leucogranitos (Gana et al, 1996), entre los que se reconocen las Unidades Lliu – Lliu, Limache, Peñuelas, Sauce y Laguna Verde.

Durante el Cretácico Inferior, el arco volcánico toma el nombre de “Arco volcánico de Lo Prado – Pelambres”, el que se encuentra al Este del arco Jurásico. El Arco de Lo Prado – Pelambres separa la cuenca de antearco de Lo Prado de la cuenca de trasarco conocida como Cuenca de Mendoza - Neuquén. Asociado a este periodo se pueden encontrar las formaciones Lo Prado y Veta Negra, las que formaban parte de la cuenca de antearco de Lo Prado (Charrier et al., 2007).

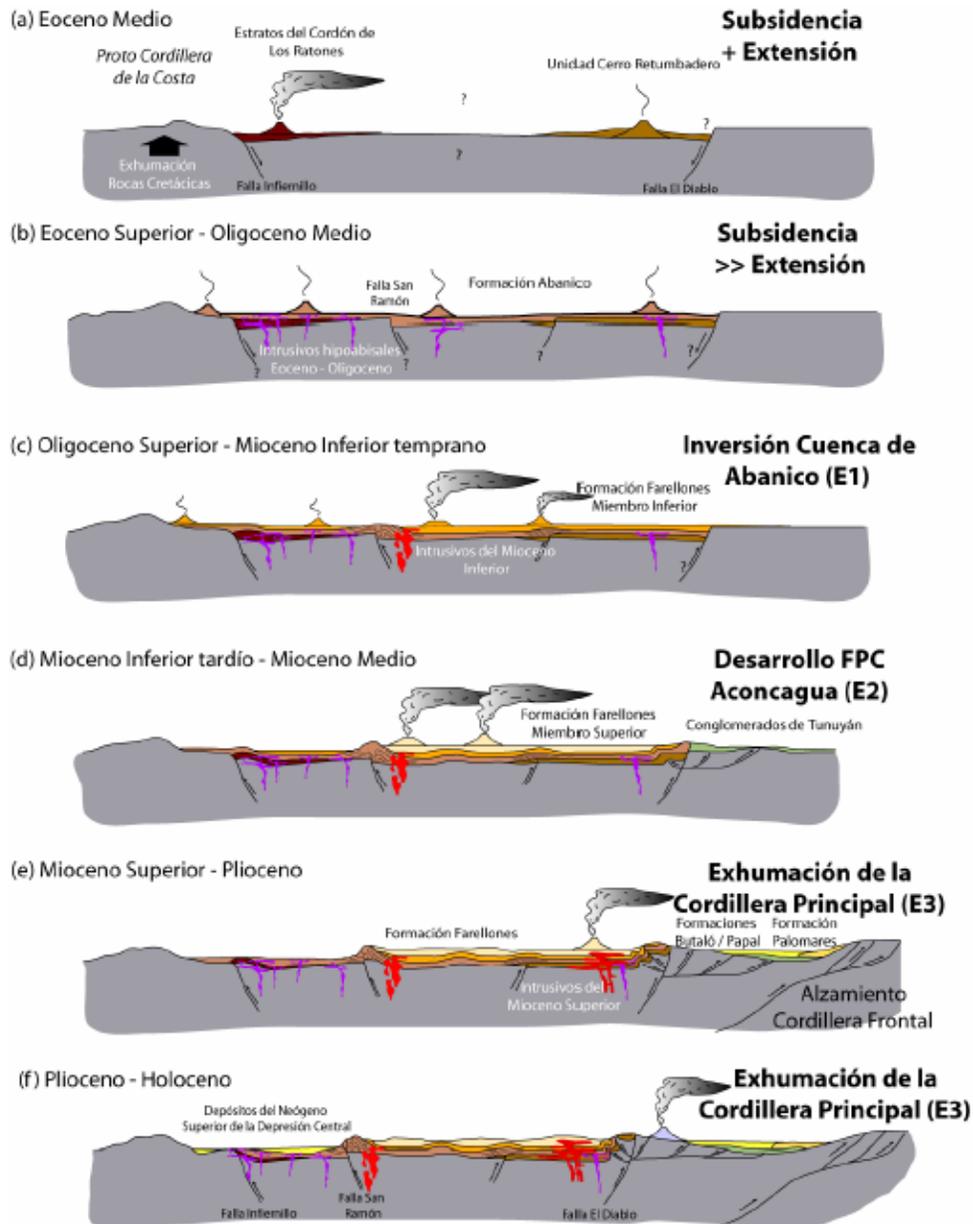
### 2.2.2 Cordillera Principal

Durante el Cretácico Tardío, hubo un cambio en la velocidad de convergencia entre las placas de Farallón y Sudamericana, cambio que habría llevado a un rejuvenecimiento de la placa oceánica subductante y, con ello, a una disminución del ángulo de subducción. Esta situación habría provocado un brusco corrimiento hacia el Este del arco volcánico, para ubicarse donde hoy se encuentra la Cordillera Principal (Charrier et al., 2007).

Durante el Jurásico y Cretácico tardío Inferior la zona al Este de la Cordillera de la Costa constituía una cuenca de trasarco conocida como Cuenca de Mendoza - Neuquén. Los registros de esta etapa se manifiestan en las Formaciones Nieves Negras, Río Colina, Río Damas, Lo Valdes y Colimapu, las cuales corresponden a depósitos sedimentarios de origen marino.

Durante el Cretácico Superior ocurre un nuevo corrimiento del arco volcánico hacia el Este y comienza la transformación de la cuenca de trasarco en una de antepaís de características extensivas (Charrier et al., 2007).

Este rasgo geomorfológico, tal cual lo conocemos hoy, comenzó a formarse recién hacia el Eoceno medio (Cretácico) (Fock, 2005) y a éste periodo corresponden las Formaciones volcano sedimentarias Abanico y Farellones entre otras. En la figura 2.3 a continuación se muestra una posible secuencia de cómo se habrían generado los principales rasgos geomorfológicos presentes en la zona de Chile Central, en particular Cordillera Principal, la Cordillera de la Costa (tal cual las conocemos hoy en día) y la Depresión Central.



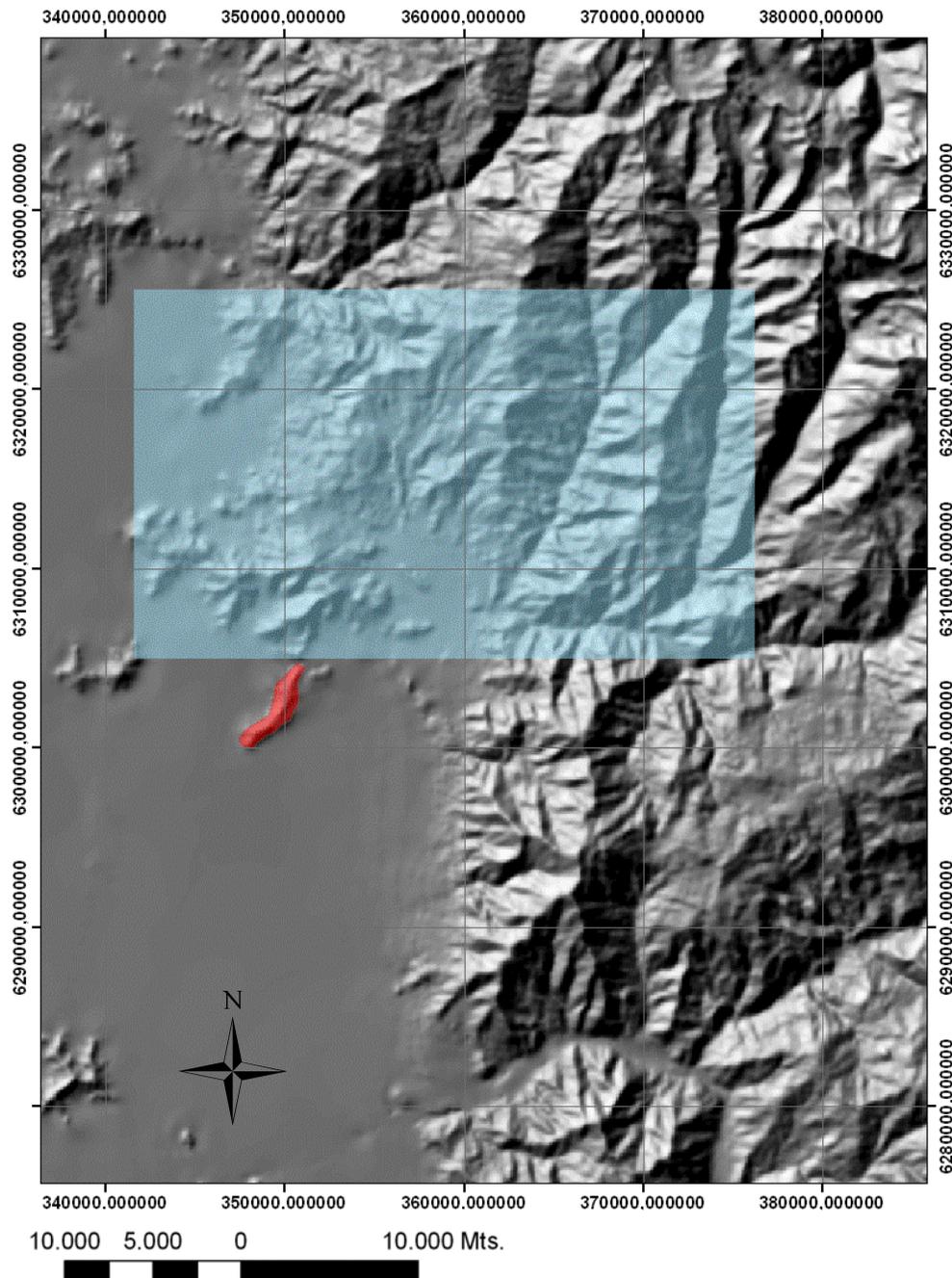
**Figura 2.3:** Modelo de evolución paleogeográfica entre los 33° y 34°S. Tomado de Fock (2005). A continuación se describe brevemente cada etapa (también tomado de Fock, 2005).

- (a): Exhumación general de las rocas Cretácicas generándose la proto Cordillera de la Costa.
- (b): Depositación de la Formación Abanico, junto con una gran actividad magmática básica. Durante este periodo se produce una fuerte subsidencia regional.
- (c): Inicio de la inversión de la cuenca subsidente donde se depositó la Formación Abanico. Se deposita el miembro inferior de la Formación Farellones de manera concordante y sintectónica sobre la Fm. Abanico.
- (d): Continúa el proceso de inversión de la cuenca y la deformación migra hacia el Este, concentrándose en el límite oriental de la cuenca. El engrosamiento cortical, que comenzó en el episodio anterior, prosigue y el Arco migra hacia el Este.
- (e) y (f): La deformación se concentra en el límite oriental de la cuenca. Durante el Mioceno Superior el arco continúa su migración hacia el Este. Se produce el principal evento de alzamiento de superficie y exhumación.

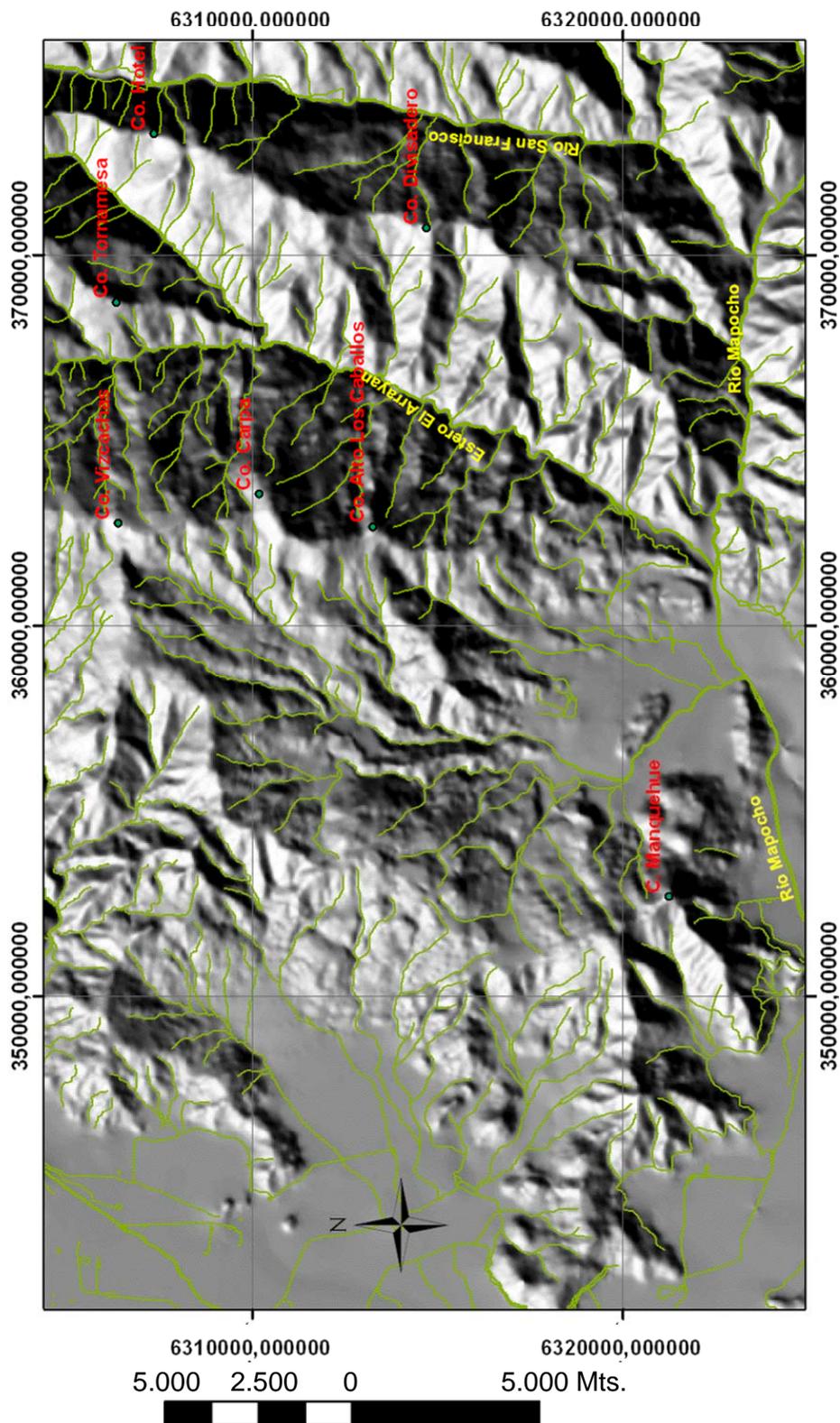
## 2.3 Geomorfología del área de estudio

El área de estudio está ubicada en el sector nororiente de la Cuenca de Santiago tal como se muestra en las figuras 2.1 y 2.4

En el área de estudio se manifiestan tanto la Depresión Central (zonas más planas) como el sector occidental de la Cordillera Principal (figura 2.5).



**Figura 2.4:** Modelo digital de terreno mostrando ubicación del área de estudio. En rojo como referencia se encuentra el Cerro San Cristóbal. Sistema de coordenadas UTM WGS-84 (zona 19S).



**Figura 2.5:** Zoom del área de estudio (mostrada en la figura 2.4) en donde se aprecia que ésta comprende la Depresión Central (zonas más planas y borde) y la Precordillera o zona más occidental de la Cordillera Principal. También se aprecian algunas características del terreno, tales como variaciones topográficas, hidrográficas, y los principales rasgos topográficos y geomorfológicos.

La topografía varía de suave a abrupta de manera muy marcada (Figuras 2.5 y 2.7(a)) a continuación, siendo bastante notorio especialmente en la dirección suroeste - noreste, entre las cotas 390 y 5405 m s.n.m., tendencia que se confirma con los perfiles Este - Oeste (figura 2.8), y con la figura 2.9 que muestra un perfil diagonal noreste – suroeste y donde se aprecia parte de la zona correspondiente a la Depresión Central que, en este caso, corresponde a la Cuenca de Santiago. Las cumbres más altas corresponden a los cerros Alto Los Caballos (2327 m s.n.m), Divisadero (2650 m s.n.m), Tornamesa (3600 m s.n.m), Hotel (3324 m s.n.m) y Vizcachas (2898 m s.n.m) (ver figura 2.5).

Las pendientes, en su mayoría, oscilan entre los 10° y los 25°, encontrándose valores menores dentro de la Cuenca de Santiago y en cursos fluviales, y superiores principalmente en cuchillas (ver figura 2.7(b)).

Por otro lado, como se muestra en las figuras 2.7(c) y 2.10, la mayoría de las laderas caen hacia el Oeste, mientras que otras caen hacia el Este, y una minoría hacia el Norte o al Sur.

Tanto la dirección de máxima pendiente (aspecto) como el valor de la pendiente, son factores importantes que determinan la cantidad de vegetación que puede encontrarse en las laderas. Esto es, a mayor pendiente menor vegetación. Por otro lado, también influye la orientación de la ladera, observándose que las laderas que miran hacia el Norte y al Oeste tienen una menor cantidad de vegetación que el resto. Este efecto es generado por que el sol pasa por el Norte y de Este a Oeste, lo que no permite que la tierra conserve una cantidad importante de humedad (ver figura 2.10).

Otra característica topográfica, no menos importante que las anteriores, es la curvatura del terreno, cuyo valor se traduce en concavidad o convexidad con respecto a una determinada referencia. Si se ubica un plano paralelo a la superficie (figura 2.7 (d)) se puede ver que la mayor parte del área de estudio no presenta curvatura, o ésta es muy débil. Las redes de drenaje, por su parte, se caracterizan por ser cóncavas, y las cuchillas convexas. Ahora, si el plano es perpendicular tanto a la superficie de la ladera como a la dirección de máxima pendiente (figura 2.7(e)), se observa que la mayoría de las laderas no presenta curvatura, pero se invierten los papeles en el caso de la red de drenaje y las cuchillas transformándose en general los primeros en convexos y los últimos en cóncavos.

Para la Quebrada El Arrayán, la Quebrada San Francisco, para el sector Este del Río Mapocho, así como para el sistema completo, el sistema de drenaje corresponde al tipo subdendrítico, paralelo o subparalelo (figura 2.5 y 2.6) más bien simétricos, lo cual es un rasgo típico en áreas donde el curso principal fluye en una zona en la que la pendiente y el control estructural son diferentes a los de la zona por donde se desarrollan los tributarios, encontrándose, además, fuertemente controlado por el tipo de sedimentos, y la litología. El resto del área de estudio se encuentra dentro de la Cuenca de Santiago, sector que presenta subcuencas con redes de drenaje de tipo centrípeta (principalmente), lo cual se corresponde con cabeceras curvadas en forma de arco, lo que también puede encontrarse relacionado con el control estructural o plegamiento erodado (figuras 2.5 y 2.6).

Por su parte, el sistema de drenaje en relación a las Formaciones Abanico y Farellones evidencia un fuerte control estructural, tanto por su linealidad observada, como por la dirección de tributarios (se incluyen quebradas sin flujo permanente) y cauces principales. En relación con los depósitos de remoción en masa, en aquellos casos en los cuales el depósito llega al cauce principal, sólo se aprecia un leve desplazamiento y una forma menos lineal o más curvilínea (levemente anastomosado). La falta de

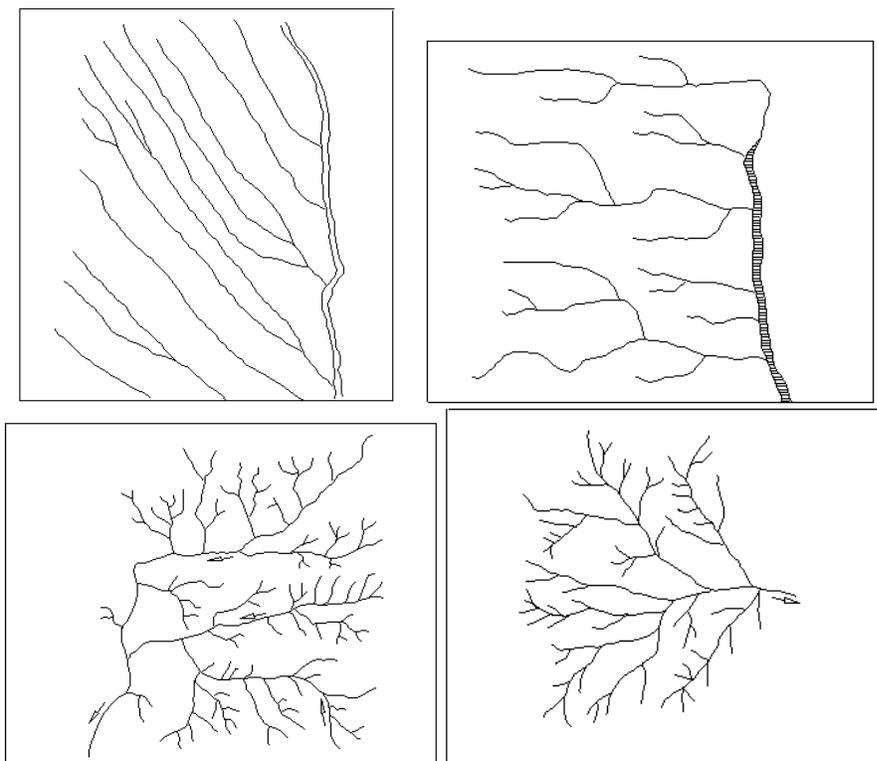
evidencias de efectos puede deberse a la antigüedad de los depósitos de remoción en masa que se encuentran en el área en estudio.

Sin perjuicio de lo anterior, existen trabajos tales como el de Ng (2006), que relacionan el nivel de evolución de una subcuenca con la densidad de deslizamientos que se pueden generar o encontrar dentro de la misma. En este caso, todas las subcuencas presentes en el área en estudio presentan un alto nivel en la escala evolutiva y se encuentran en la etapa de integración por lo que la densidad de deslizamientos por subcuenca es baja concordando con el trabajo ya citado.

Algunos de los rasgos geomorfológicos más importantes presentes en el área en estudio, de acuerdo a la figura 2.5 son:

- Río Mapocho
- Quebrada El Arrayán (Estero El Arrayán)
- Quebrada San Francisco (Río San Francisco)
- Estero Colina

Llama la atención la marcada dirección Norte Sur del Estero El Arrayán y del Río San Francisco, así como las también numerosas quebradas que se observan.



**Figura 2.6:** A la izquierda arriba se muestra un ejemplo de drenaje paralelo y a la derecha arriba de uno subparalelo. Abajo a la izquierda se muestra uno de tipo subdendrítico mientras que abajo a la derecha, se muestra uno de tipo centripeto. Las flechas indican dirección del flujo (Sesser, 2006).

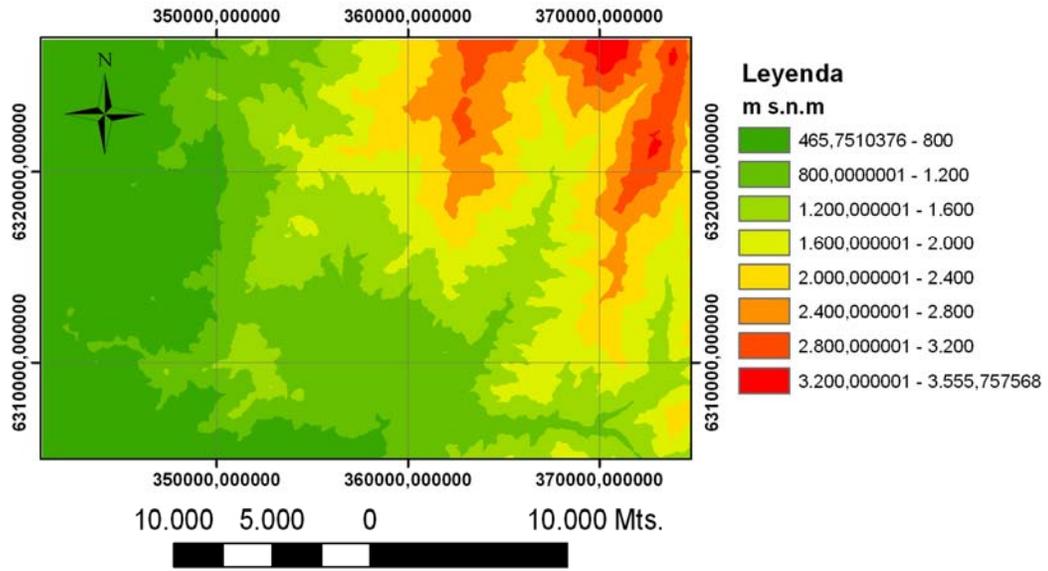


Figura 2.7 a: Mapa hipsométrico del área en estudio.

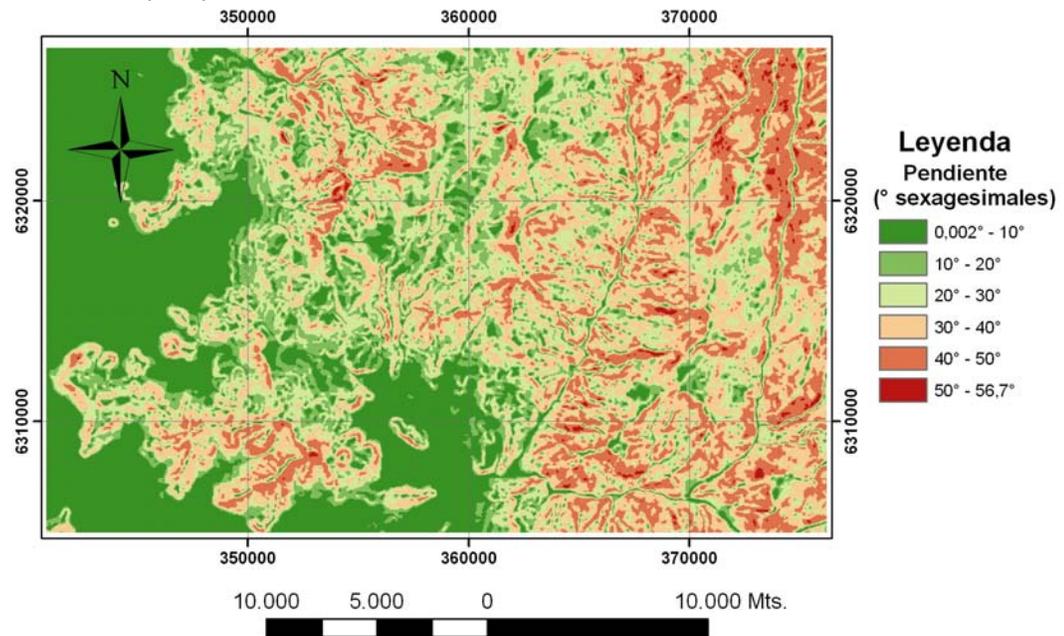


Figura 2.7 b: Distribución de pendientes de ladera en área en estudio

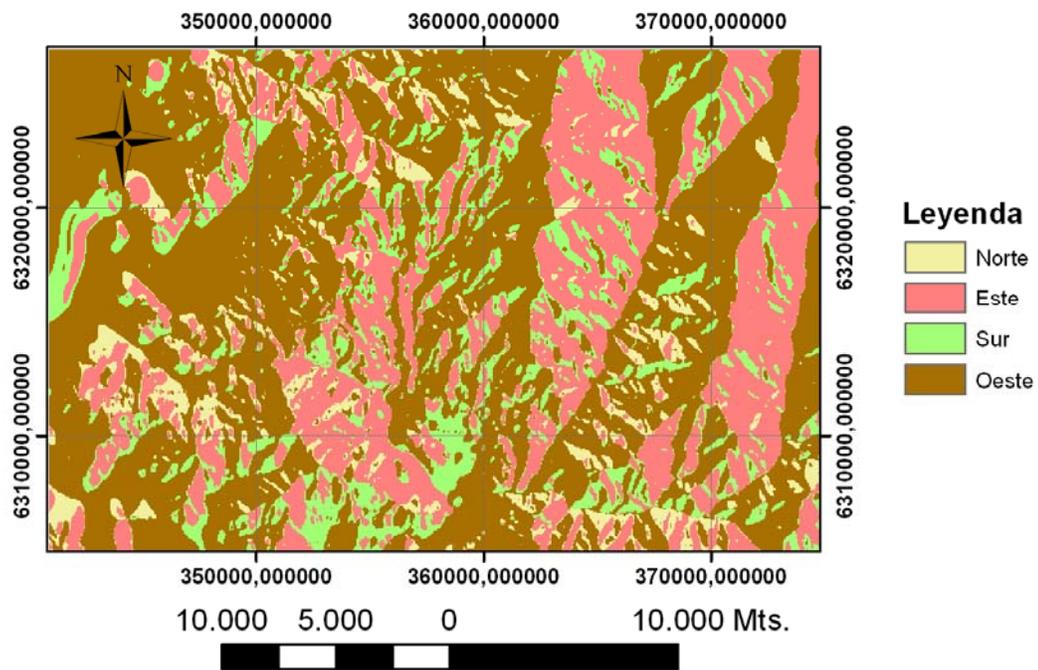


Figura 2.7 c: Distribución de aspecto (orientación de ladera) en el área en estudio

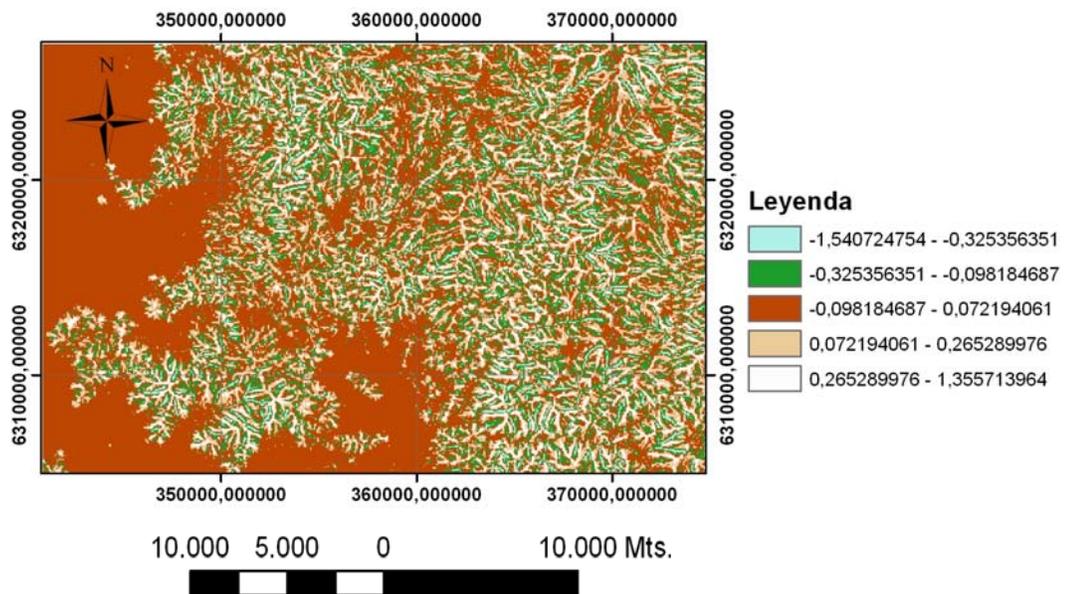


Figura 2.7 d: Curvatura en planta en el área en estudio.

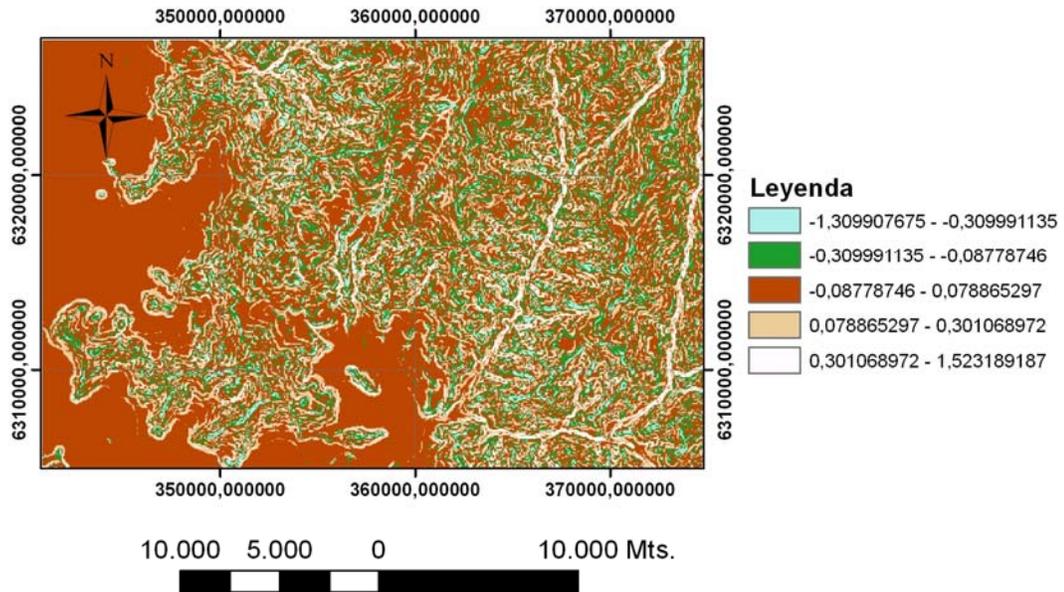
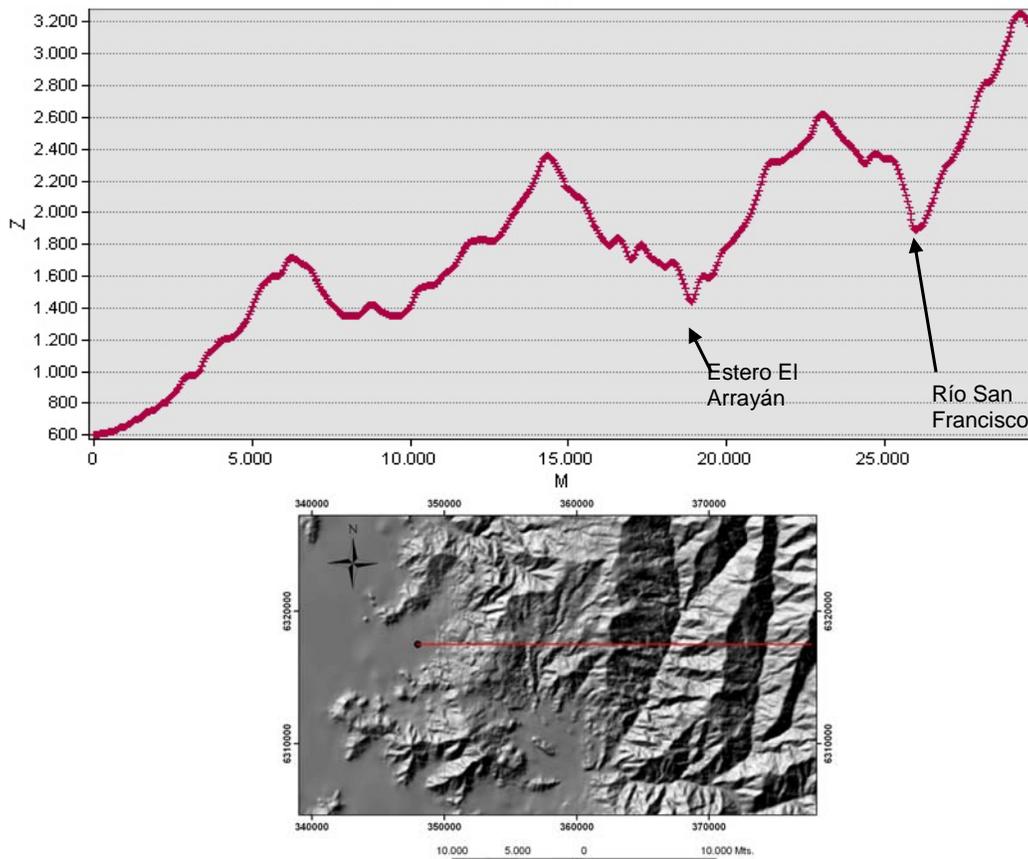
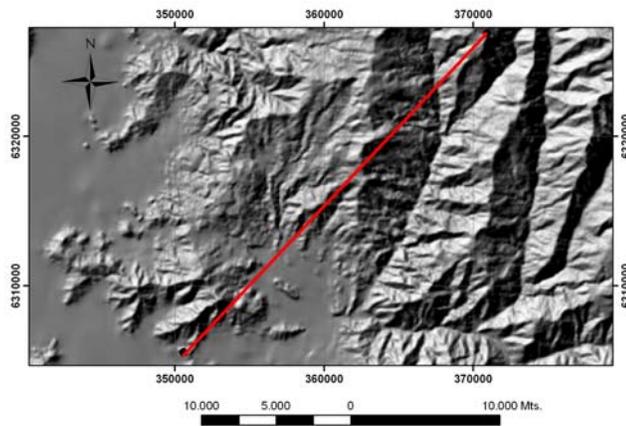
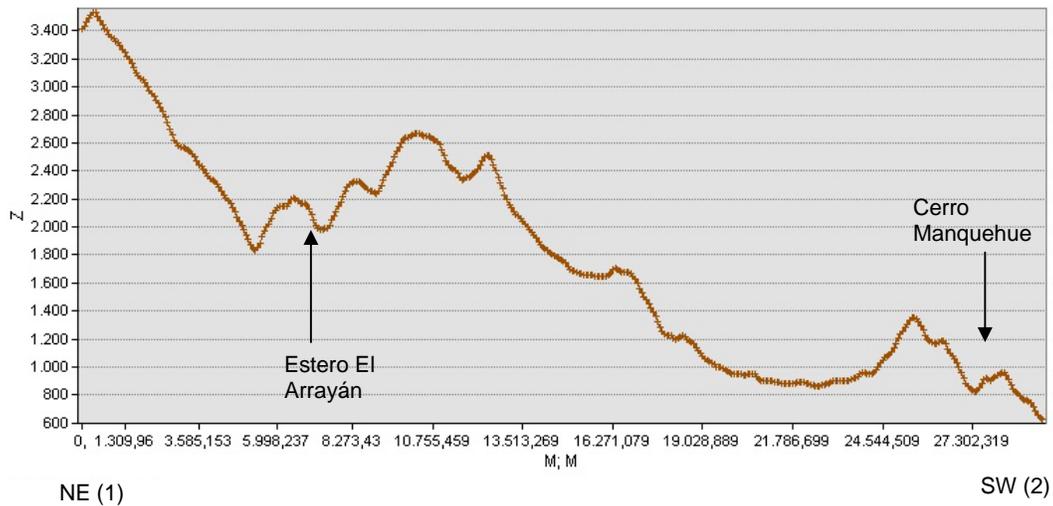


Figura 2.7 e: Curvatura en perfil en el área en estudio.

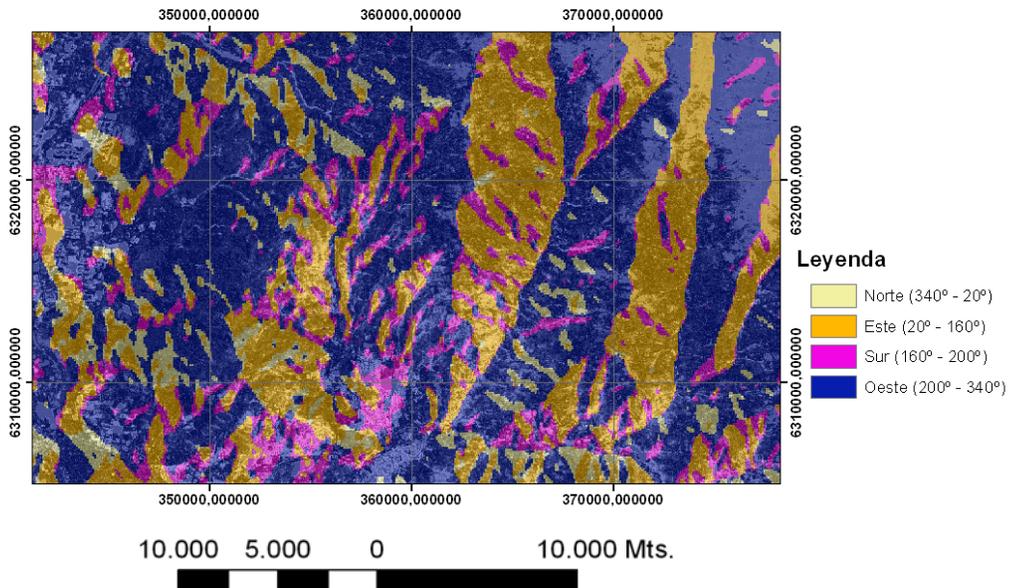
**Figura 2.7:** Para las unidades en que se encuentra la curvatura ver Anexo 2.



**Figura 2.8:** Perfil Este Oeste entre los puntos: 347980 y 377630 (coordenada Este), y 6317500 (coordenada Norte). Exageración vertical aproximadamente 1H es a 4V.



**Figura 2.9:** Perfil diagonal en dirección noreste a suroeste entre los puntos 1: 370890/6326930 y 2: 350580/6305570. Exageración vertical aproximadamente 1H a 4V.



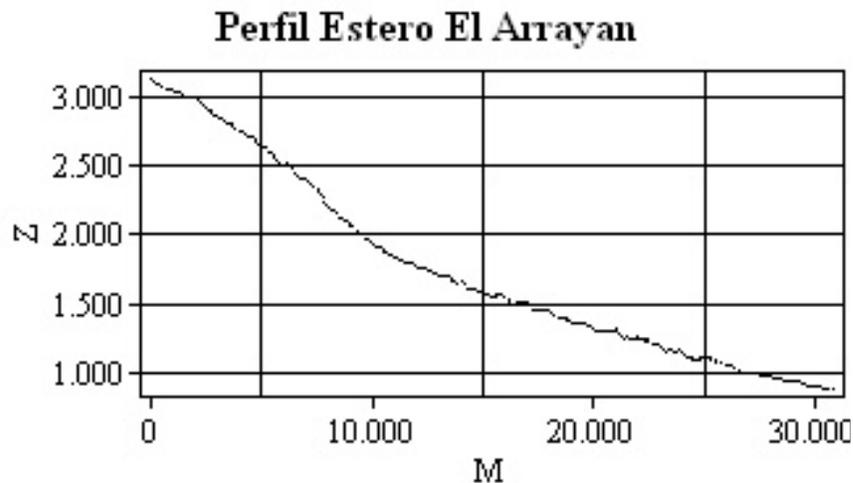
**Figura 2.10:** Distribución de aspecto, tomado de la figura (2.7 (c)), con NDVI sobrepuesto. Las zonas más claras corresponden a áreas con mayor vegetación.

### 2.3.1 Quebrada El Arrayán:

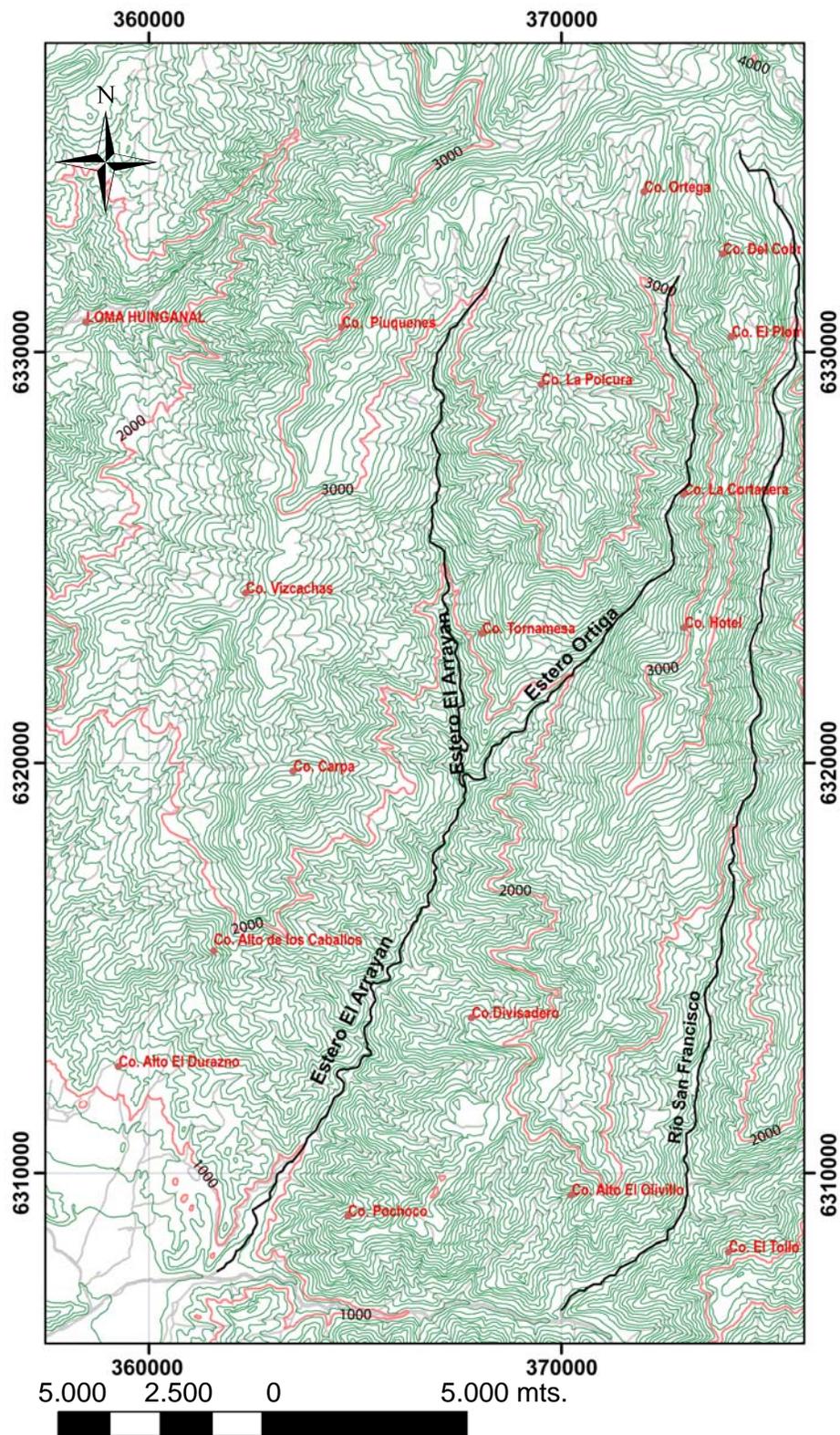
La hoya hidrográfica del Estero el Arrayán abarca una superficie de 159,356 km<sup>2</sup>. Se encuentra limitada al Norte por el Cordón de los Españoles, y al Sur por el Río Mapocho. Por el oriente limita con: el Cerro Divisadero (2650 m s.n.m), Cerro la Cortadera (3330 m s.n.m), Cerro El Plomo (3780 m s.n.m) y Hotel (3324 m s.n.m), los cuales forman parte de la Sierra de San Francisco. Al Poniente se encuentran el Cerro Alto de los Caballos (2327 m s.n.m), el Cerro Vizcachitas (2898 m s.n.m) y el Cerro Carpa, los cuales forman parte de Sierra Ortiga. El Estero el Arrayán desemboca en el Río Mapocho, poco antes de llegar a la Cuenca de Santiago, a una altura de 900 m s.n.m mientras que su nacimiento se encuentra a una altura de 3650 m s.n.m en el Cerro del Plomo (4050 m s.n.m). Uno de sus principales afluentes es el Estero Ortiga, el cual se junta con el Estero El Arrayán en la cota 1550 m s.n.m (Figura 2.12).

La red de drenaje (incluyendo quebradas y cursos fluviales activos) alcanza una longitud de 296 km de largo, por lo que la densidad de drenaje es de 1,858 (1/km). Por su parte, el Estero El Arrayán tiene un largo de 25 km y el resto corresponde a quebradas y afluentes.

Los cambios de pendiente que sufre el Estero El Arrayán se muestran en la figura 2.11. En dicha figura (la cual no se encuentra a escala) se puede ver que la pendiente en su cabecera es mucho más fuerte que al final excepto por la parte más alta, en donde la pendiente se ve influida por la presencia de nieve.



**Figura 2.11:** Perfil topográfico a lo largo del cauce del Estero El Arrayán. La figura muestra los cambios de pendiente a lo largo de la Quebrada El Arrayán el cual fue obtenido a partir de la base de datos topográfica OTAS (escala 1:50.000)



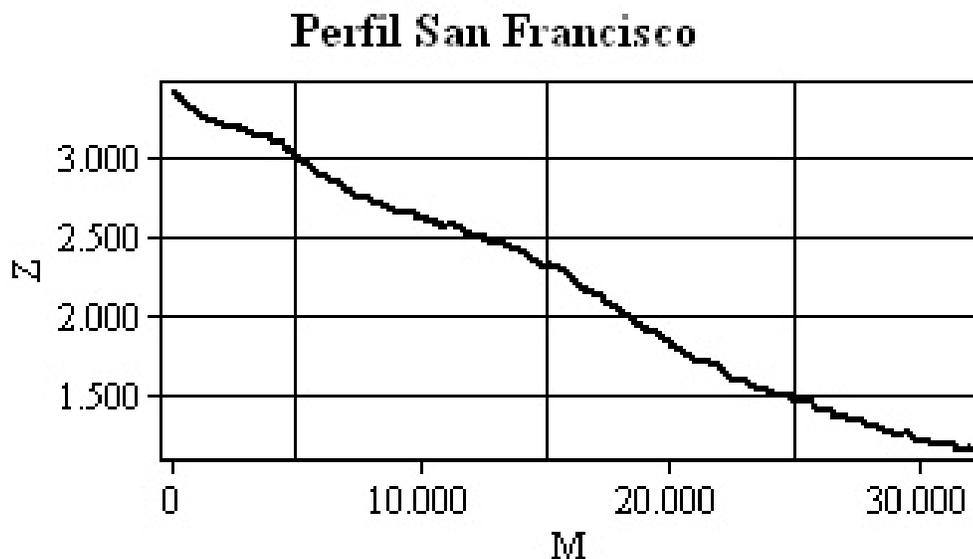
**Figura 2.12:** Mapa topográfico de la subcuenca del Estero el Arrayán, en donde se pueden ver los principales elementos topográficos y afluentes. Fuente: base topográfica OTAS 1:50.000

### 2.3.2 Quebrada San Francisco:

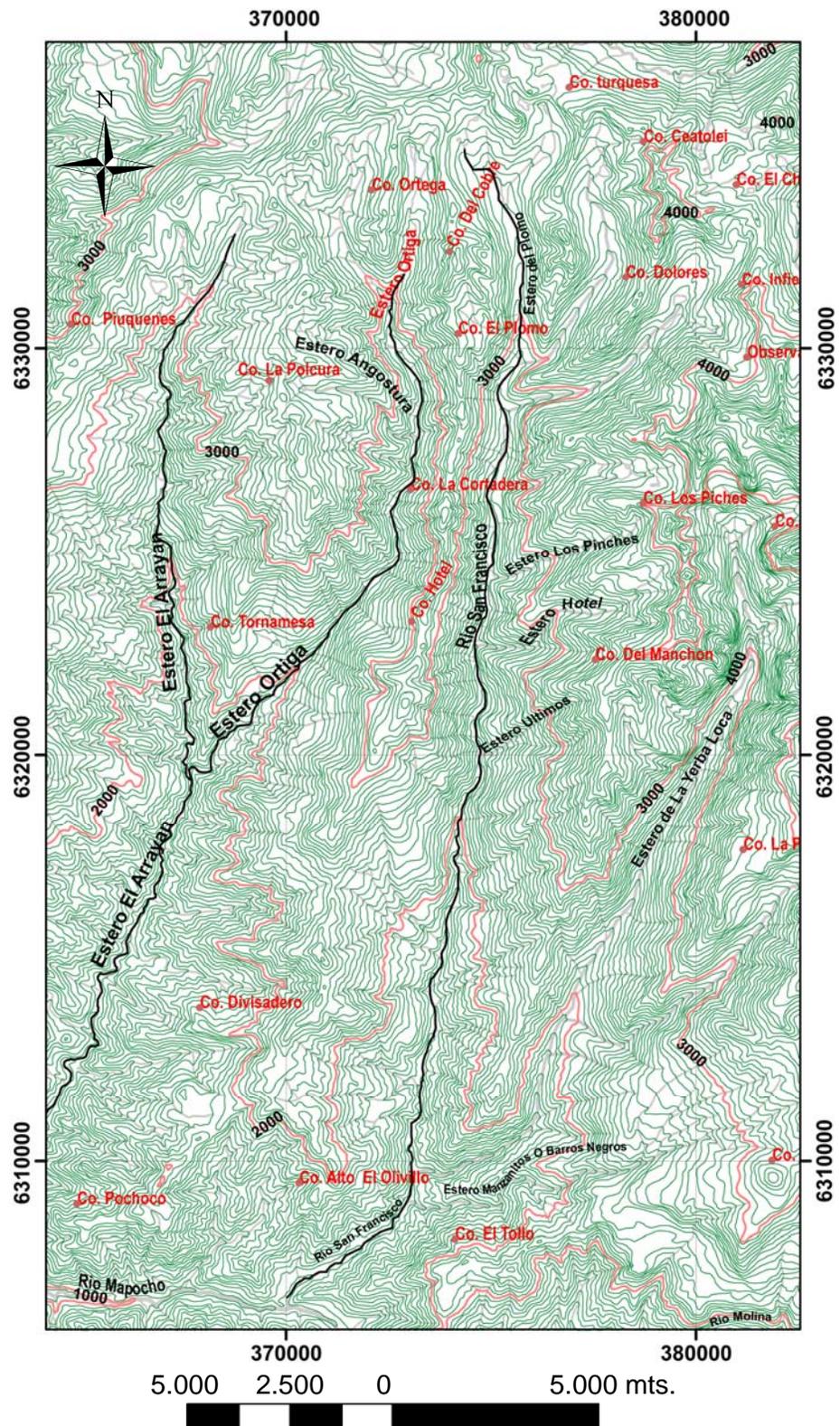
La hoya hidrográfica del Río San Francisco abarca una superficie de aproximadamente 619 km<sup>2</sup>. Se encuentra limitada al norte por el Cordón de los Españoles, al sur por el Río Mapocho, al poniente por los cerros: Cerro Divisadero (2650 m s.n.m), Cerro la Cortadera (3330 m s.n.m), Cerro El Plomo (3780 m s.n.m) y Hotel (3324 m s.n.m), los que forman parte de la Sierra de San Francisco) y por el Oriente limita con los cerros: Cerro del Manchon (3873 m s.n.m), Cerro Los Piches (4274 m s.n.m), Cerro Dolores (3760 m s.n.m) y Cerro Ceatolei (4088 m s.n.m), los que forman parte del Cordón de la Yerba Loca. El Río San Francisco desemboca en el Río Mapocho, bastante antes de llegar a la Cuenca de Santiago, a una altura de 1145 m s.n.m y nace a una altura de 3749 m s.n.m en el Cerro del Plomo (4050 m s.n.m). Uno de sus principales afluentes es el Estero El Plomo, con el que se junta en la cota 3000 m s.n.m. Otros afluentes son: Estero Hotel, Estero Los Pinches y Estero Ultimo, entre las cotas 2550 m s.n.m y 2150 m s.n.m, y el estero Manzanitos (o Barros Negros) en la cota 1300 m s.n.m (figura 2.14).

La red de drenaje (incluyendo quebradas y cursos fluviales activos) alcanza una longitud de 1206 km de largo, por lo que la densidad de drenaje es de 1,949 (1/km). El Río San Francisco tiene un largo de 51 km (el resto corresponde a quebradas y afluentes).

Los cambios de pendiente que sufre el Río San Francisco se muestran en la figura 2.13 la que no se encuentra a escala. Se aprecia que la pendiente en el inicio es mucho más fuerte que al final salvo por la parte más alta, en donde la pendiente se ve influida por la presencia de nieve. Se puede apreciar una diferencia notable con respecto al Estero El Arrayán, ya que, la pendiente que manifiesta a lo largo de su curso, el Río San Francisco, es mucho más fuerte que la del Estero El Arrayán.



**Figura 2.13:** Perfil topográfico a lo largo del cauce del Río San Francisco. La figura muestra los cambios de pendiente a lo largo del Río obtenido a partir de la base de datos topográficos OTAS (escala 1:50.000)



**Figura 2.14:** Mapa topográfico de la subcuenca del Río San Francisco en donde se pueden ver los principales elementos topográficos y afluentes. Fuente base topográfica OTAS 1:50.000.

### 2.3.3 Río Mapocho:

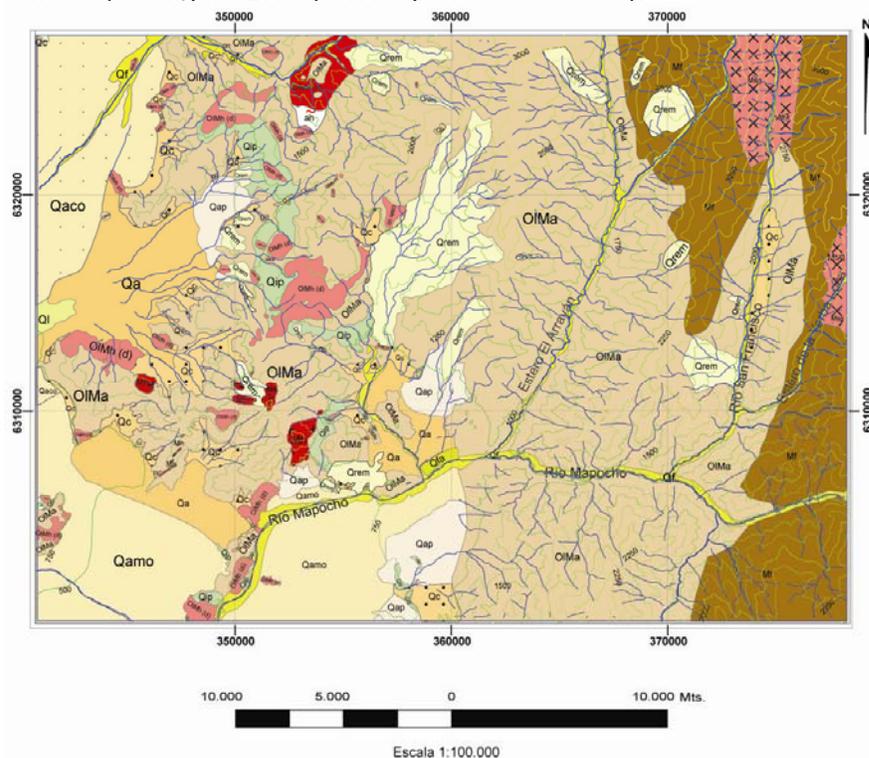
Corresponde al límite sur del área en estudio (figura 2.5) y cómo tal solo serán descritas algunas de sus características. El Principal afluente del Río Mapocho, además de los previamente mencionados (San Francisco y Arrayán), nace a 3242 m.s.n.m en el Cerro Bismark (4710 m.s.n.m) con el nombre de Río Molina. En la cota 1200 m.s.n.m pasa a llamarse Río Mapocho. El Río Mapocho cruza la Cuenca de Santiago en sentido Este Oeste, y, una vez que alcanza el borde occidental de la Cuenca de Santiago, pasa a tener sentido Norte-Sur.

### 2.3.4 Estero Colina:

Corresponde al límite norte del área de estudio (figura 2.5) y como tal solo serán descritas algunas de sus características. El Principal afluente del Estero Colina nace a 3350 m.s.n.m en el cerro Lletas (3539 m.s.n.m). El Estero Colina cruza la Cuenca de Santiago en sentido noreste a suroeste a casi este - oeste, una vez que alcanza el borde Oriental de la cuenca de Santiago pasa a tener sentido Norte Sur y no la cruza. Dentro del área de estudio ocupa un área bastante menor.

## 2.4 Geología del área de estudio

Las unidades geológicas que se pueden encontrar en la zona en estudio son principalmente de edad Cuaternaria (Pleistoceno –Holoceno) y Cenozoicas, representadas principalmente por las Formaciones Abanico y Farellones, además de los intrusivos Oligo–Miocenos. Una breve descripción de cada una de estas unidades se presenta a continuación. Tanto las unidades existentes en el área de estudio como la descripción de las mismas, fueron obtenidas de Thiele (1980), Wall et al. (1999) y trabajo de terreno (Ver figura 2.15 y en mapa fuera de texto).



**Figura 2.15:** Mapa geológico área en estudio. Modificado de Wall et al. (1999), mapa geológico de Chile (1:1.000.000), Thiele (1980) y trabajo de terreno. Ver simbología en mapa fuera de texto.

### 2.4.1 Geología del Cuaternario:

El área en estudio se caracteriza por la gran presencia de sedimentos no consolidados, de edad Cuaternaria, entre los que destacan aquellos de origen fluvial, coluvial o de remociones en masa.

#### **(Qrem) Depósitos de Remoción en Masa (Pleistoceno–Holoceno)**

Consiste en deslizamientos en suelo y/o roca, cuyo depósito cubre un área mayor a 0,1Km<sup>2</sup> (ver sección 2.5).

#### **(Qf) Depósitos Fluviales (Holoceno)**

Consiste en sedimentos no consolidados, cuyo origen está relacionado con cursos fluviales activos. Están constituidos por ripios, gravas, gravas arenosas clasto-soportadas, y con clastos redondeados y en parte imbricados (facies de relleno de canal). Incluyen lentes arenosos con estratificación plana–horizontal o cruzada, junto a escasos niveles de limos finamente laminados (depósitos de planicie de inundación).

#### **(Qfa) Depósitos fluviales antiguos (Pleistoceno – Holoceno)**

Consiste en sedimentos no consolidados, asociados a cursos fluviales abandonados, constituidos por ripios, gravas, gravas arenosas, arenas y limos. Esto incluye depósitos fluviales aterrizados, ubicados entre 3 y 20m por sobre el curso fluvial actual y que pueden presentar una cobertura de suelo moderada o bien desarrollada, y depósitos ubicados a escasa altura sobre el cauce actual, cubiertos por una delgada capa de suelo y que son esporádicamente inundados en épocas de crecida. Figura 2.16.



**Figura 2.16:** Depósitos fluviales antiguos Río San Francisco. Se puede ver claramente la imbricación de los clastos más grandes, además de un alto grado de redondeamiento (Vista al Este).

#### **(Qc) Depósitos Coluviales (Pleistoceno Superior- Holoceno)**

Consiste en sedimentos no consolidados, ubicados en las faldas y al pie de los cerros, a la salida de quebradas o en pequeñas quebradas intermontanas. Están formados principalmente por clastos y bloques mal seleccionados (escombro de falda). Constituyen mantos adosados a cerros con pendientes de más de 10°, y que presentan estratificación grosera paralela a las laderas.

Los depósitos coluviales se pueden encontrar en toda el área en estudio, pero sólo se considera su existencia en aquellos lugares donde cumplen con los criterios de escala.

#### **(Qap) Depósitos Aluviales de Piedemonte (Pleistoceno Superior – Holoceno)**

Consiste en sedimentos no consolidados, ubicados en la desembocadura de quebradas mayores. Se encuentran principalmente en el borde precordillerano, al oriente de Santiago, y forman abanicos aluviales con pendientes de entre 3 a 6°. Están constituidos por una alternancia de depósitos de remoción en masa (flujos de detritos), con depósitos fluviales y mantos de crecida, que se interdigitan y en parte progradan sobre los depósitos aluviales (Qa) de los ríos Maipo y Mapocho (Qamo).

#### **(Ql) Depósitos Lacustres (Pleistoceno – Holoceno)**

Consiste en sedimentos no consolidados, impermeables, formados por limos, arcillas, y arenas limosas, finamente laminados o macizos, ubicados en zonas deprimidas donde la napa freática se encuentra a profundidades inferiores a 5m, o que aflora a la superficie. En el área en estudio, este tipo de depósitos aparecen reducidamente en el sector occidental de la misma.

#### **(Qa) Depósitos Aluviales (Pleistoceno Medio – Pleistoceno Superior)**

Consiste en sedimentos no consolidados, que rellenan gran parte de las cuencas de Santiago. Forman parte de extensos abanicos aluviales coalescentes, con suave pendiente hacia el Oeste, provenientes de las hoyas hidrográficas principales. Hacia sus cabeceras, estos depósitos se interdigitan con depósitos coluviales (Qc) y aluviales de piedemonte (Qap), mientras que hacia el Oeste gradan hacia depósitos lacustres. Están formados por gravas y arenas con sedimentos finos en menor proporción.

Qamo corresponden a los depósitos aluviales del río Mapocho.

Qaco corresponden a depósitos aluviales del estero Colina.

#### **(Qip) Depósitos asociados a la Ignimbrita Pudahuel (Pleistoceno medio) (Wall *et al.*, 1996)**

Consiste en un depósito de flujo piroclástico de ceniza y lapilli pumíceo, de composición riolítica, de amplia distribución en la Depresión Central. Son de color blanco amarillento a pardo claro, medianamente a bien consolidado, presenta estructuras de flujo laminar matriz-soportado, constituido por una matriz de ceniza que comprende vidrio, cristales y pómez. Localmente se aprecian estructuras traccionales con lentes y niveles con acumulaciones de líticos y pómez. Se observan pómez desde 1 mm hasta 15 cm de diámetro, con escasas biotitas. La Ignimbrita Pudahuel incluye líticos accidentales, principalmente de rocas volcánicas con retrabajo fluvial e intrusivas de origen local. Esta unidad representa el producto de un evento eruptivo explosivo, proveniente de la Caldera Diamante del Complejo Volcánico Maipo, ubicado aproximadamente a 130 km hacia el sureste, en la Cordillera Andina y su edad, ha sido determinada por el método de trazas de fisión, es de  $450.000 \pm 60.000$  a A.P. (Stern *et al.*, 1984). En el área en estudio se pueden encontrar afloramientos aislados, ubicados en el cerro Apoquindo y al Este de Peñalolén intercalados con depósitos aluviales de piedemonte cordillerano (Qap).

#### 2.4.2 Intrusivos

La presencia de intrusivos va acompañada (generalmente) de altas cotas y altas pendientes (más difícilmente erodables).

##### **(OIMh) Intrusivos Hipabisales (34 -19 Ma)**

“Stocks”, filones manto, cuellos volcánicos y diques, cuya composición varía desde gabros a andesitas de piroxeno. Forman cuerpos de tamaño inferior a  $8\text{km}^2$  y afloran preferentemente en el sector Oeste del área en estudio. También en el área en estudio aparecen dos de los cuatro tipos de intrusivos reconocidos dentro de este grupo, los que poseen una edad Mioceno Inferior y se encuentran intruyendo a la Formación Abanico. Dentro de las unidades que se pueden encontrar en el área de estudio se tiene:

- OIMh (c): *Unidad de Dioritas a Dioritas Cuarcíferas de piroxeno y anfíbola*, (52 - 60)%  $\text{SiO}_2$ , los que se encuentran en stocks de hasta  $8\text{km}^2$ . Son de grano medio y de color gris oscuro, que afloran al noreste del Cerro Las Moscas y en la Quebrada La Leonera. No se han realizado dataciones radiométricas para estos intrusivos.
- OIMh (d): *Unidad de Andesitas, Andesitas – Basálticas y Dacitas de dos Piroxenos*, (52 - 67)%  $\text{SiO}_2$ . Forman “Stocks” de hasta 2km de diámetro y filones – mantos o lacolitos de hasta 150m de espesor. Son rocas porfídicas de color gris azulado, compuestas por plagioclasas (An50-35), ortopiroxeno,  $\pm$ clinopiroxeno,  $\pm$ olivino y escasos pseudomorfos de anfíbola. La masa fundamental es generalmente intergranular, con plagioclasa, opacas y granulos de piroxeno. Se conocen edades K-Ar en roca total de  $21.8\pm 0.5\text{ Ma}$  (Thiele et al., 1980),  $21.2\pm 0.2\text{ Ma}$  (Vergara y Lopez, 1980) y  $21.2\pm 1.0\text{ Ma}$  (Selles, 1999)

##### **(Mh) Intrusivos Hipabisales Andesíticos y Dacíticos (20 – 19Ma)**

Consiste en “Stocks” y diques andesíticos y dacíticos de anfíbola (57 -72)%  $\text{SiO}_2$  de hasta  $4\text{km}^2$  de superficie, que intruye en varias localidades a la Formación Abanico (Cerro Manquehue). Son rocas leucocráticas, y porfídicas con fenocristales de plagioclasa (An38 -45), hornblenda y, en menor proporción, piroxeno, parcial o totalmente transformado en anfíbola, que están inmersos en una masa fundamental compuesta por plagioclasa, cuarzo y escasa magnetita. Para estos intrusivos se han obtenido edades K- Ar en roca total de  $19.5\pm 0.5\text{ Ma}$  (Drake et al, 1976),  $20.3\pm 2.9\text{ Ma}$  y  $40\text{Ar}/39\text{Ar}$  en anfíbola de  $18.8\pm 0.8\text{ Ma}$  (edad plateau; Gana y Wall, 1997) las que se consideran como edades cercanas a la cristalización.

##### **(Msg) Intrusivos granodioríticos**

De hornblenda y biotita, y en menor proporción monzogranitos, monzonitas cuarcíferas y monzodioritas.

### 2.4.3 Rocas estratificadas

En la zona en estudio se pueden encontrar unidades litológicas de edad Eoceno a Mioceno. Estas son las Formaciones Abanico y Farellones, las cuales son de carácter volcánico.

#### **(Mf) Formación Farellones (Mioceno) (Klohn, 1960)**

Es sucesión compuesta de lavas, tobas e ignimbritas, con intercalaciones de brechas. Las lavas manifiestan un claro predominio sobre las tobas y brechas, mientras que la alternancia de rocas volcanoclásticas más finas, en tramos de 4 a 5 metros, producen una marcada estratificación en la serie (Thiele, 1980).

Esta formación sobreyace a la Formación Abanico entre los 32° S y los 35° S y también la separa en dos franjas de orientación Norte – Sur. De ellas, sólo la franja Occidental se encuentra dentro de la zona de estudio.

Aunque es motivo de discusión el tipo de contacto entre ambas formaciones, según las evidencias más recientes se puede describir como un contacto transicional y con amplias variaciones tanto Norte- Sur como Este-Oeste, reportándose en diversos sectores como discordante o pseudo-concordante, dependiendo de la reactivación de fallas asociadas a la inversión tectónica de la cuenca de la Formación Abanico (Godoy et al., 1999; Charrier et al., 2002).

El espesor reportado para la Formación Farellones es de 2.500 m, aumentado por los numerosos mantos intrusivos y lacolitos del Mioceno Superior–Plioceno que la intruyen. (Thiele, 1980; Rivano *et al*, 1990).

El techo de esta formación corresponde a la actual superficie de erosión (Thiele, 1980; Rivano *et al*, 1990).

#### **(OIMa) Formación Abanico (Eoceno Superior? – Mioceno Superior) (Aguirre, 1960)**

Es una Unidad litoestratigráfica constituida por lavas básicas a intermedias, rocas piroclásticas ácidas, e intercalaciones sedimentarias continentales (fluviales, aluviales y lacustres), formando lentes de hasta 500 m de espesor (Charrier et al., 2002; Nyström et al., 2003). Esta formación se encuentra plegada y es afectada por un pervasivo metamorfismo de muy bajo grado (Aguirre, 1960; Thiele, 1980; Levi et al., 1989). El espesor del conjunto se estima en aproximadamente 3.000 m, aún cuando se acepta que éste aparece aumentado por las numerosas intrusiones de filones-manto y lacolitos andesíticos (Thiele, 1980). Las rocas de la Formación Abanico se disponen en dos franjas Norte – Sur, separadas por los afloramientos de la Formación Farellones.

Se ha propuesto que la franja occidental de la Formación Abanico se encuentra, ya sea en contacto por falla con la Formación lo Valle, o concordante con un hiatus de 35 Ma. (Gana et al., 1997; Fuentes et al., 2000; Fuentes, 2004).

El contacto con la sobreyacente Formación Farellones es transicional y con amplias variaciones tanto Norte – Sur como Este – Oeste, reportándose en diversos sectores tanto discordante, pseudo-concordante (Charrier et al., 2002), concordante o por falla (Godoy et al., 1999)

En el área en estudio, tanto en la Quebrada del arrayán como en la de San Francisco se observan manteos hacia el Este (Figura 2.17 y 2.18) y la litología predominante son lavas de composición intermedia, andesíticas, especialmente en la quebrada El Arrayán

y brecha volcánica en ladera Este del Río San Francisco, y en ladera Oeste desde coordenada Norte 6312700 (en el sistema de referencia UTM WGS – 84, Zona 19S), esta se encuentra fuertemente erodada (figura 2.19). No se observó en ningún punto de estas quebradas estratos correspondientes a secuencias sedimentarias de origen continental.



**Figura 2.17:** Vista hacia el sureste en Quebrada El Arrayán, donde se pueden observar manteos hacia el Este



**Figura 2.18:** Vista hacia el noreste en quebrada el Arrayán. En el fondo se ve la Formación Abanico manteando hacia el Este.



**Figura 2.19:** Brecha volcánica fuertemente erodada, perteneciente a la Formación, Abanico ubicada en la ladera oeste de la Quebrada San Francisco.

## **2.5 Depósitos de remoción en masa**

La mayoría de las remociones en masa presentes en el área no cubren una superficie mayor a 100m<sup>2</sup> (figura 2.20). Estos depósitos de remoción en masa se caracterizan por ser principalmente flujos de detritos, no son considerados en este trabajo y un número de deslizamientos que, para efectos de este trabajo, son muy pequeños (la escala de trabajo es de 1:100.000). Sin embargo, existe una cantidad de deslizamientos que cubren una superficie mayor a 0.1 km<sup>2</sup> lo que es suficiente para este trabajo (figura 2.21 y 2.22 (c)).



**Figura 2.20:** Vista al Sur desde cuchilla de Cerro Alto el Olivillo. Se observa pequeño deslizamiento en ladera Oeste del Río San Francisco. La ladera cae hacia el Este.

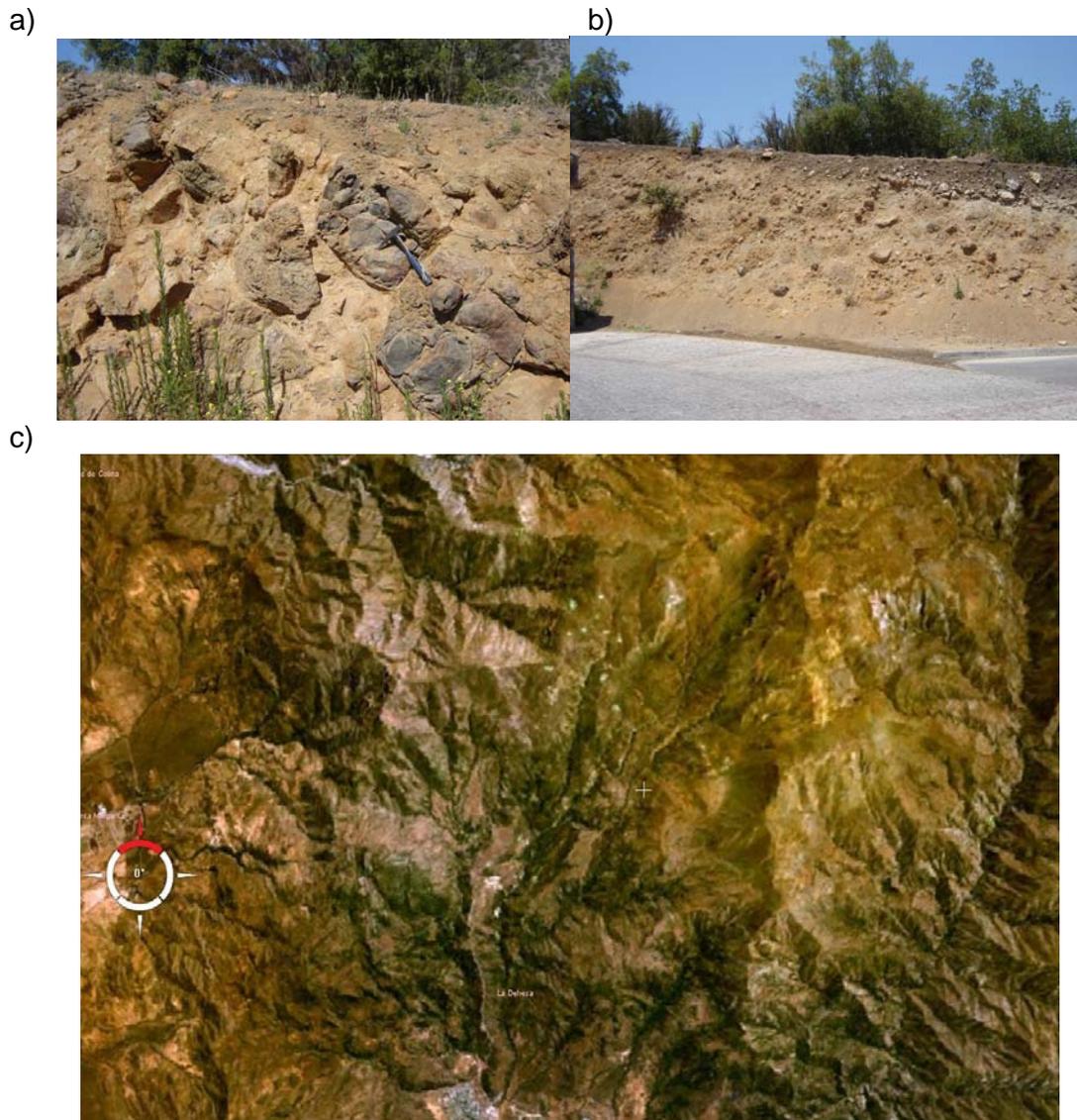


**Figura 2.21:** Complejo de Remociones observado en la Quebrada de San Francisco en la ladera Oeste (Fotografía tomada mirando al Oeste). Abajo el mismo complejo de remociones visto desde el camino a Farellones

La mayoría de los deslizamientos presentes en el área en estudio se caracteriza por presentar clastos angulosos y por ser la mayoría de ellos deslizamientos en roca, siendo muy pocos en suelo, por lo general mucho más pequeños comparados con aquellos en

roca. Dada la escala de este trabajo los deslizamientos que cubren un área mayor a 100m<sup>2</sup> son, en su mayoría, en roca (figuras: 2.21 y 2.22).

En el área en estudio se puede ver una remoción que resalta sobre el resto por su gran tamaño (Figura 2.22). Se caracteriza por presentar una morfología de lengua muy clara, aterrazada y algo lobulada (figura 2.23). El depósito se caracteriza por encontrarse bastante alterado ya que algunos clastos son muy blandos (alto grado de meteorización), con grandes clastos (de hasta cinco metros de diámetro) de brecha volcánica y andesita (principalmente). Los clastos son bastante angulosos y no se observa un alto porcentaje de matriz.

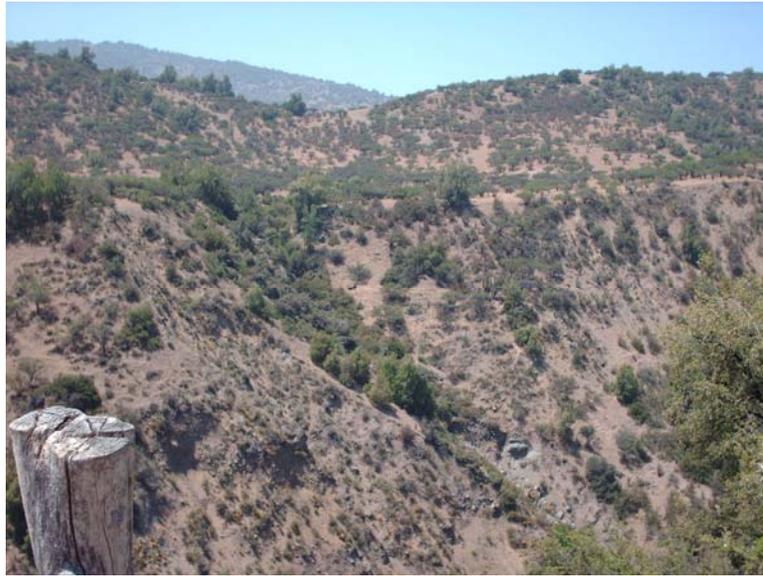


**Figura 2.22:** (a) y (b) Corresponden a depósitos de remoción en masa observados en cortes de ladera en la Hacienda Santa Martina. (c) Imagen obtenida de Nasa World Wind, mostrando la remoción completa. A la derecha se ve el Estero El Arrayán. La flecha roja indica el Norte.

a)



b)



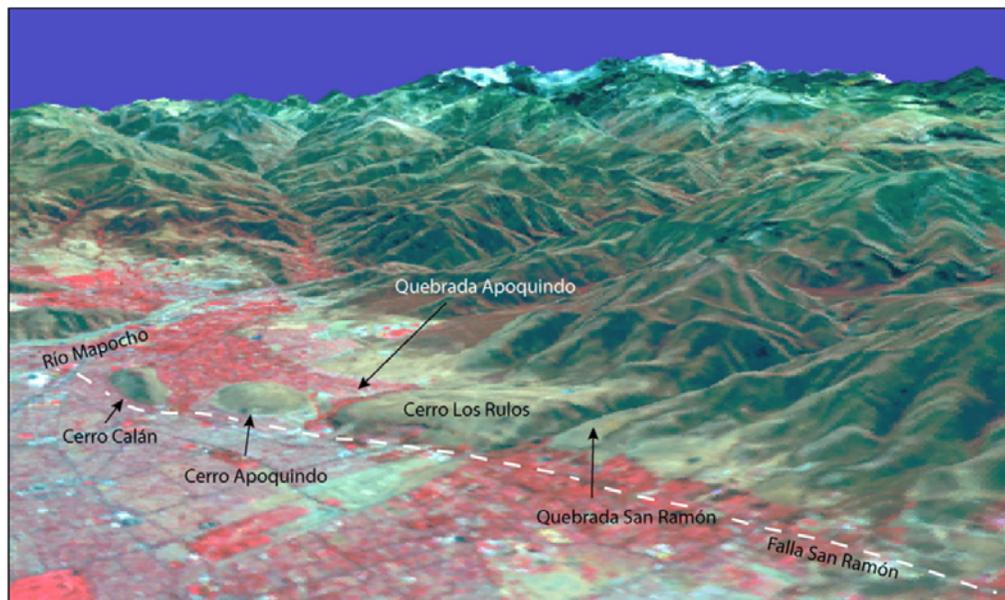
**Figura 2.23:** Morfología aterrazada que se puede ver en la Hacienda Santa Martina. (a) Panorámica en donde se pueden observar los tres niveles. (b) Detalle del nivel más bajo en donde se resalta la superficie ondulosa (pero muy plana a gran escala).

## 2.6 Marco tectónico

El área en estudio se caracteriza por no presentar un marcado rasgo tectónico a diferencia de las presentadas por la Falla de San Ramón (Rauld, 2002), misma que se encuentra al Sur del área en estudio (figura 2.24), y por la Falla Pocuro (Campbell, 2005), la que se encuentra al Norte del área en estudio. Ambas estructuras presentan una orientación aproximadamente Norte – Sur, con rasgos estructurales cuya traza se distingue por varias decenas de kilómetros.

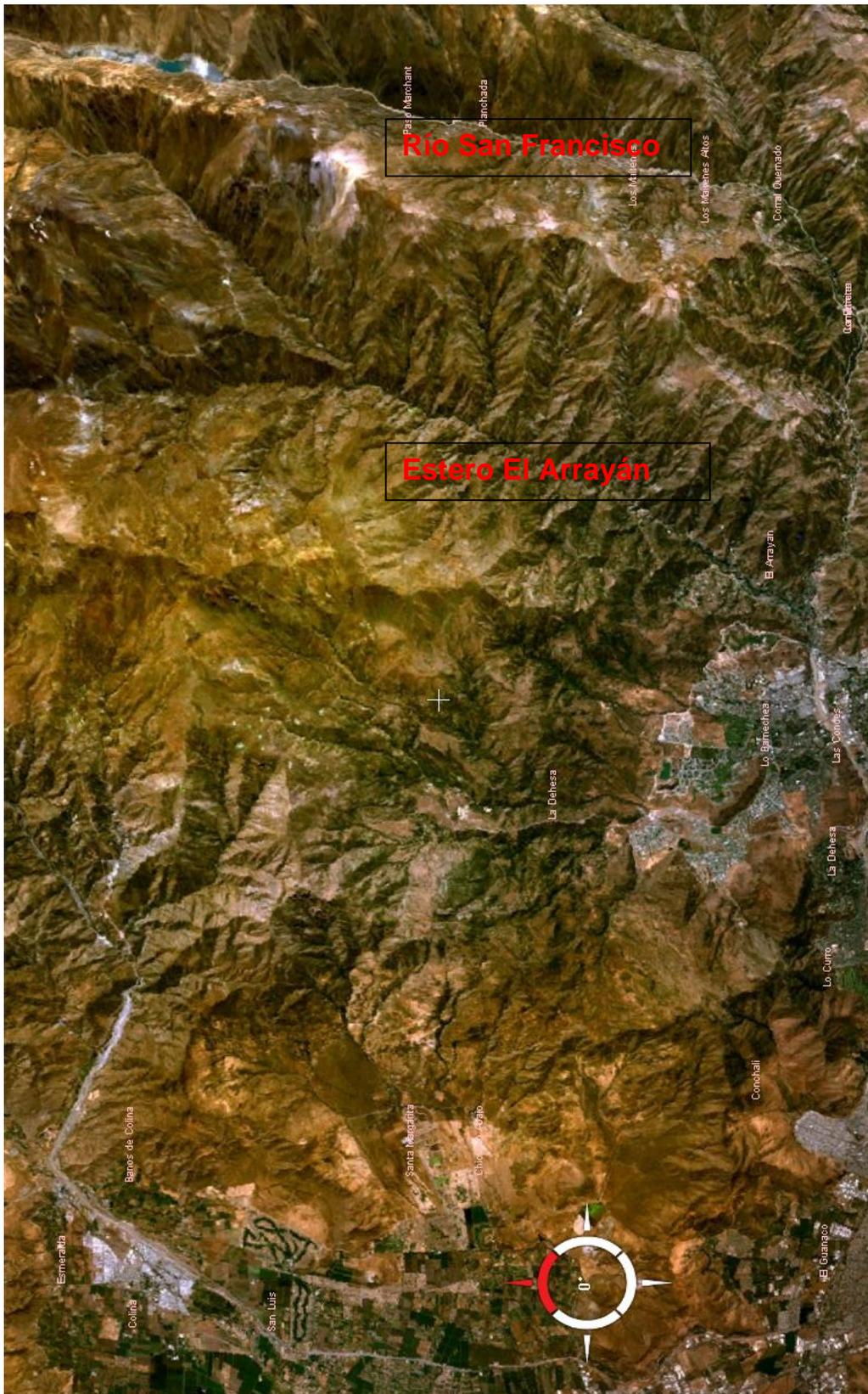
Las estructuras utilizadas en este trabajo fueron obtenidas tanto de fotogeología como del mapa geológico Til Til – Santiago.

Si bien en la figura 2.25 del área en estudio no se observa una estructura tan marcada, como el caso de la Falla de San Ramón, se puede observar numerosos rasgos lineales, a pseudo lineales con una dirección de aproximadamente N15°E (figura 2.25). Esto es muy evidente si se observa el Estero el Arrayán, y el Río San Francisco.

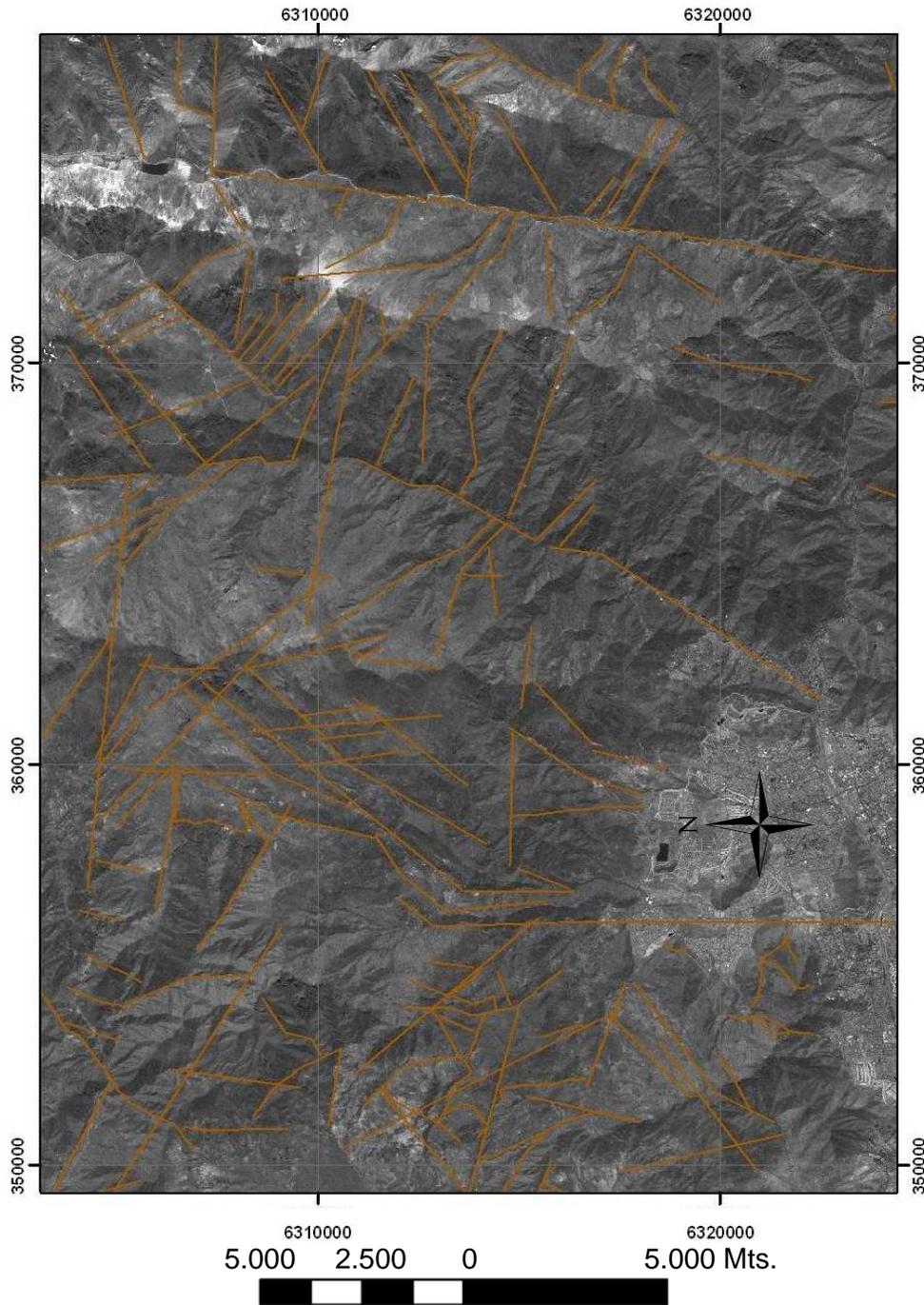


**Figura 2.24:** Perspectiva virtual al noreste. Se aprecia la estructura formada por los cerros Calán, Apoquindo y Los Rulos, y que es interpretada como la Falla de San Ramón. (Tomado de Rauld, 2002).

Además de las estructuras ya señaladas, que se pueden ver a simple vista en el área en estudio, se observan también numerosos lineamientos, tanto en las subcuencas del Estero el Arrayán y en la del Río San Francisco, como en el resto del área en estudio (figura 2.26)

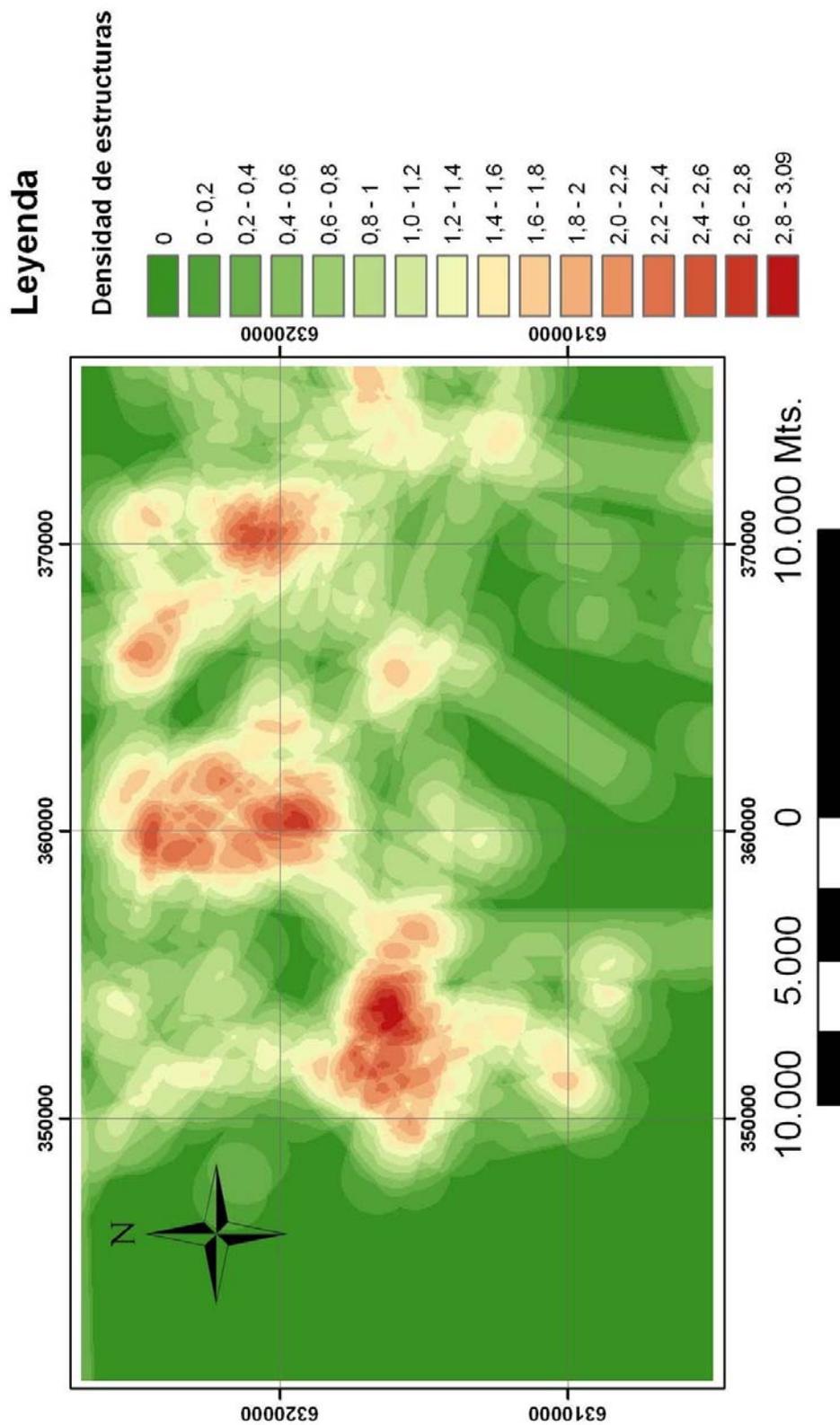


**Figura 2.25:** Imagen del área de estudio obtenida con el programa NASA World Wind, en donde se distingue el fuerte rasgo lineal que presentan el Estero El Arrayán y El Río San Francisco.

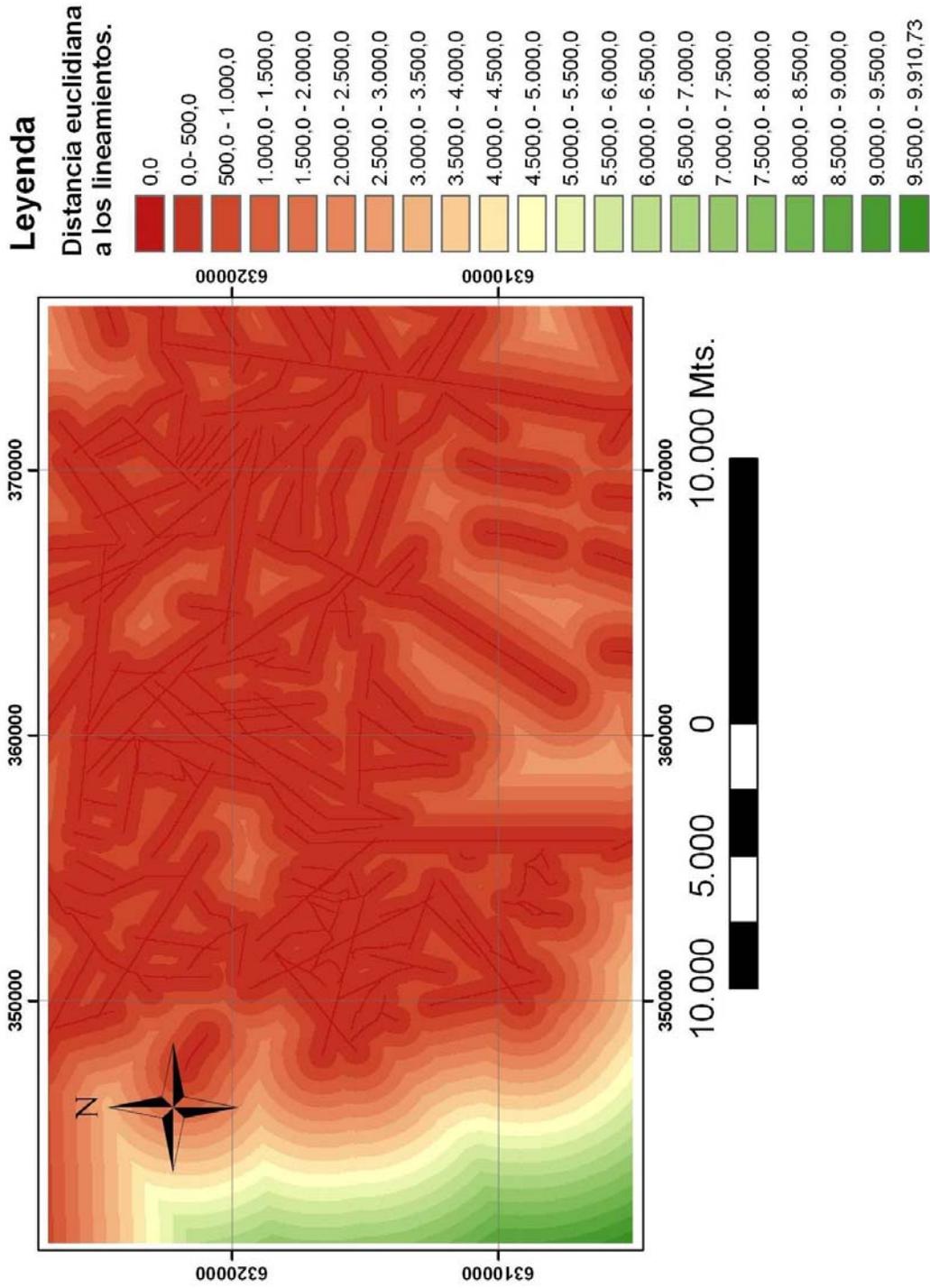


**Figura 2.26:** Fallas y lineamientos presentes en el área en estudio. La imagen de fondo es la banda pancromática de la imagen satelital LANDSAT 7 ETM.

En la figura 2.27 se observa una alta densidad de estructuras en el sector noroeste del área en estudio especialmente en aquel sector en el que no se ven estructuras claramente demarcadas. La distancia euclidiana a las estructuras (ver anexo 2) es bastante homogénea en el sector precordillerano, mientras que en la cuenca de Santiago se observan distancias mayores a las estructuras (figura 2.28).



**Figura 2.27:** Representación gráfica de la densidad de lineamientos en el área de estudio, extraído de la figura 2.26. La unidad de área es  $2.500\text{m}^2$  (píxel de 50m por 50m).



**Figura 2.28:** Representación gráfica de la distancia euclidiana a los lineamientos en el área en estudio, extraída de la figura 2.26. Unidad de área es 2500m<sup>2</sup>.



**LEYENDA**

-  **Depósitos aluviales recientes (Qar)**  
Depósitos de detritos (bloques, gravas y arenas), >3m de espesor, que rellenan los cauces actuales. Son producto de intercalaciones de flujos de detritos y flujos esporádicos de agua de diversa energía.
-  **Depósitos coluviales (Qc)**  
Acumulaciones de material grueso (bloques, grava y arena) sin o con escasa matriz, en laderas de alta pendiente o pequeñas quebradas, >1m de espesor.
-  **Escombros de falda (Qe)**  
Depósitos >1m de espesor, de arena, grava y bloques inmersos en una matriz limo arenosa que se encuentran en las laderas de los cerros.
-  **Depósitos activos de acumulación de detritos (Qd)**  
Depósitos de gravas y bloques angulosos en matriz arenosa, sueltos y sujetos a retrabajo continuo y reciente, se encuentran cubriendo laderas de baja pendiente, adyacentes a quebradas.
-  **Abanicos aluviales (Qa)**  
Depósitos aluviales de granulometría gruesa, gravo-arenosa con matriz limo-arcillosa, que se encuentran relleno el pie de las quebradas, generando un relieve suave y de poca pendiente.
-  **Depósitos aluviales antiguos (Qaa)**  
Depósitos aluviales, de granulometría media, compactos y estables. Son producto de flujos esporádicos de agua en las quebradas.
-  **Depósitos aluviales, fluviales y de remoción en masa antiguos intercalados (Qaf)**  
Intercalaciones de depósitos aluviales, fluviales y de remoción en masa, depositados en el piedemonte cordillerano, >20m de espesor, que se encuentran formando colinas.
-  **Depósitos de remoción en masa aterrazados (Qrmt)**  
Depósitos de granulometría media, con bloques en matriz limo-arcillosa, de espesor >20m, compactos y estables. Corresponden a antiguos eventos de remoción en masa. Forman un nivel aterrazado extenso.
-  **Depósitos de remoción en masa antiguos (Qrma)**  
Depósitos de antiguos eventos de remoción en masa, de espesor >20m, compuestos por bloques y grava en matriz arenosa ligeramente arcillosa. Son bastante compactos y estables. Se encuentran formando colinas.
-  **Formación Abanico (Tsa)**  
Secuencia de rocas volcánicas (lavas, brechas, y tobas) con intercalaciones sedimentarias (areniscas y lutitas), con zonas con cubierta de suelo de hasta 1,5m de espesor, principalmente en las cabeceras de la hoya hidrográfica.
-  **Intrusivo de composición basáltica (Tf)**  
Intrusivo basáltico de grano fino.
-  **Intrusivo inferido de composición basáltica (Tf)**  
Intrusivo basáltico inferido, parcialmente cubierto de suelo de espesor >1m.
-  **Intrusivo de composición granodiorítica (Mi)**  
Intrusivo granodiorítico de grano grueso.
-  Eje axial de pliegue anticlinal
-  Falla San Ramón
-  Zona de falla
-  Lineamientos
-  Escarpe de remoción en masa antigua
-  Escarpe asociado a remoción en masa reciente
-  Zonas con caídas de rocas
-  Zonas de muestreo
-  Eje de quebrada
-  Caminos
-  Curvas de nivel cada 50 m

ESCALA 1 : 40.000  
0 0,25 0,5 1 Km

Universidad de Chile  
Departamento de Geología



Marisol Lara Castillo  
2007

La hoya hidrográfica de la Quebrada de San Ramón tiene una superficie de 38.5 km<sup>2</sup>, y una extensión de 11 Km. La altura máxima en la cabecera alcanza a los 3253 m s.n.m. Presenta pendientes fuertes, del orden de 20° a 30°, como promedio, desarrollando algunos saltos rápidos de agua sobre la cota 2000 m s.n.m. Bajo dicha cota, entre los 1650 m s.n.m y los 1450 m.s.n.m, posee una pendiente de aproximadamente 10°, debido a la presencia de depósitos de remoción en masa que se encuentran en la zona. En esta zona confluyen la mayoría de los tributarios. En las laderas, las pendientes son más pronunciadas destacándose nuevamente la zona superior por presentar laderas con pendientes entre 25° y 35°, llegando incluso a 50°.

Se identifican en la zona afloramientos de rocas volcánicas (tobas, brechas y lavas) con algunas intercalaciones sedimentarias de areniscas y lutitas, asignadas a la Formación Abanico, intruidas por cuerpos de tamaño variable de basaltos, granitos y granodioritas.

Además, fueron reconocidos una serie de lineamientos que podrían asociarse a estructuras mayores, posibles causantes de la morfología que presenta la zona, principalmente de las trazas de las quebradas principales y de la configuración de los depósitos recientes.

Los depósitos más antiguos corresponden principalmente a depósitos de remoción en masa antiguos, depósitos de remoción en masa aterrazados y depósitos que conforman el relieve de los cerros La Cabreria y Los Rulos (eventos aluviales de diversa energía) (figura 2.29).

## **Capítulo III: Formulación teórica**

El presente capítulo corresponde a un resumen explicativo de la metodología aplicada para la construcción de los modelos matemáticos predictivos de las remociones en masa. Una explicación más detallada acerca del tema se puede encontrar en el anexo número tres.

El análisis discriminante y la regresión logística, son técnicas analíticas apropiadas cuando la variable dependiente es categórica o binaria y lo que se busca es determinar la probabilidad de pertenencia a un grupo determinado, dadas ciertas condiciones (variables independientes) (Hair et al., 1999)

Los dos métodos utilizados, en este trabajo permiten agrupar los casos mediante un proceso de clasificación. En el caso de la regresión logística, además permite estimar la probabilidad de ocurrencia de un evento en particular, asociado a un cierto punto, sin considerar los puntos vecinos. Con el análisis discriminante, por su parte, también se puede estimar la probabilidad de ocurrencia de un fenómeno, pero su valor va a ser estimado mediante el método de las probabilidades condicionales de Bayes, por lo que el resultado va a ser equivalente a la probabilidad de pertenencia a un grupo determinado, dadas ciertas condiciones de los datos. Por lo tanto, el método de análisis discriminante difiere del método de regresión logística ya que este último no supone un valor de probabilidad previo de pertenencia a un grupo para cada caso en función del tamaño de los grupos. Para minimizar los efectos de esta diferencia, y hacer que los resultados obtenidos sean comparables, el análisis discriminante asume una probabilidad a priori de valor 0,5, esto es, igual probabilidad de pertenencia a cualquiera de los grupos para cada caso. De esta manera se minimiza el efecto de la diferencia de tamaño de los distintos grupos.

A grandes rasgos, la metodología de trabajo utilizada con ambos métodos es la siguiente:

- 1) Comprobar que los datos cumplen con las suposiciones básicas de los métodos
- 2) Realizar cálculos y estimaciones de probabilidades por caso (o punto) para obtener un modelo con los paquetes estadísticos computacionales.
- 3) Comprobar que los modelos obtenidos se adecuan bien a los datos con los que fueron creados.
- 4) Obtención de resultados gráficos.

Es importante mencionar que al discretizar el área en estudio, ninguno de los métodos utilizados en este estudio considera la interacción entre puntos vecinos.

Para más detalles acerca de los métodos utilizados, sus pasos, pruebas y resultados consultar el anexo número tres.

### **3.1.-Aplicación de métodos estadísticos a estudios de susceptibilidad a la ocurrencia de remociones en masa.**

Los métodos estadísticos en la geología son utilizados tanto para predecir peligros como susceptibilidad, nacieron producto del progresivo aumento de la población en el mundo lo que ha presionado a la humanidad a ocupar sectores cada vez menos aptos para asentamientos humanos con el consecuente deterioro de la estabilidad de las áreas ocupadas y que aceleran los procesos de remoción en masa o que vuelven un área previamente no susceptible en susceptible (Gorsevski, 2000). La aplicación de métodos estadísticos a la predicción de remociones en masa es utilizada tanto para predecir peligro o susceptibilidad, entre otras aplicaciones.

En la actualidad, las técnicas estadísticas se han convertido en una herramienta de creciente importancia para el estudio de la susceptibilidad de peligros naturales, en particular de fenómenos de remoción en masa. Situándose como uno de los procedimientos con mejores resultados en la predicción espacial de deslizamientos (Yesilnacar et al., 2005; Greco, 2006). Esta metodología es de tipo indirecta.

La principal ventaja de los métodos estadísticos multivariados es la capacidad de relacionar distintas variables con una variable que representa la evidencia observada y, a partir de esto, predecir futuras ocurrencias (Carrara, 1991). Su objetivo principal es intentar mitigar los efectos de un posible evento de remoción en masa que no solo tiene altos costos monetarios, sino que usualmente conlleva altos costos sociales, los que generalmente se traducen en una alta cantidad de muertes. (Lee et al., 2006). Ejemplos de este tipo sobran en el mundo y algunos de los casos más importantes han ocurrido en: Malasia, Japón, Centro América, entre otras.

Uno de los problemas más grandes, que hacía difícil la aplicación de estos métodos a este tipo de problemas, era el error que producía la inadecuada tecnología de computo existente hasta hace 40 años. Sin embargo, a lo largo del tiempo se han ido desarrollando numerosas técnicas que han permitido una mejor y más confiable aplicación de metodologías estadísticas multivariadas, tales como:

- Disponibilidad de sensores remotos (fotografías aéreas y satelitales).
- SIG.

En particular, los sensores remotos han permitido obtener datos cada vez con mejor resolución, lo que tiene gran importancia para el estudio de lugares de difícil acceso. Sin embargo, acceder a este tipo de información con la calidad y en la cantidad necesaria para este tipo de estudios tiene un alto costo económico.

Por su parte, los sistemas de información geográfica, SIG (GIS por sus siglas en inglés), representan un avance importantísimo para este tipo de problemas pues permiten el manejo de datos geográficos (o capas de información) y su posterior visualización (Carrara et al., 1995).

Sin embargo y a pesar de los avances tecnológicos previamente mencionados, la solución de este tipo de problemas o, en otras palabras, la posibilidad de obtener un resultado 100% correcto con este tipo de metodologías es altamente complejo, pues va a depender de:

- Cantidad de eventos positivos.
- Tamaño de los eventos.
- Resolución (o nivel de detalle) de información que describe factores condicionantes (pérdida de resolución implica pérdida de confiabilidad en el resultado).
- Error de los datos (ajenos a los del investigador).

Si bien existe un nivel de error aceptable (asociado muchas veces a la escala de trabajo), éste debería ser minimizado, lo que no siempre es posible.

Ahora bien, con respecto a los otros problemas mencionados, el ideal es que tanto la cantidad de los eventos, como su tamaño, sean adecuados, ya que en caso contrario, los métodos estadísticos tienden a sobreestimar su valor y a ser conservadores, haciendo parecer que la susceptibilidad es mucho menor que la real. Pero si el tamaño es excesivo puede suceder lo contrario (Can et al. 2005). La gran cantidad de datos que es necesario manipular, vuelven a estos métodos relativamente engorrosos (Carrara et al., 1990).

Es necesario mencionar que, de acuerdo con Can et al. (2005), no es necesario ni deseable que un mapa de susceptibilidad de deslizamientos reproduzca completamente los datos empíricos (o teóricos), esto es, reproducir aquellos deslizamientos que ya se sabe que ocurrieron y que fueron utilizados como variables de entrenamiento. Pero hay que tener en cuenta, que mientras más alto es el porcentaje de aciertos, más específico para un área determinada se vuelve el modelo, haciendo difícil la predicción. Por lo tanto, un buen modelo es aquel que se ajusta relativamente bien a los datos entregados, pero que también es capaz de predecir aquellos que no han ocurrido aún.

Las variables que generalmente son utilizadas en este tipo de estudios son:

- Geología: factor importante en tanto exista un control geológico de los deslizamientos (esto es que la mayoría ocurra en una formación determinada).
- Topografía: es un factor condicionante muy importante, en tanto que variables como la Pendiente, Aspecto, Curvatura pueden afectar la estabilidad de una ladera frente a distintos factores gatillantes.

En general todos los autores concuerdan que los modelos geomatemáticos proveen una aproximación confiable en estudios para detectar peligros medio ambientales, y son particularmente útiles dentro del marco de proyectos multidisciplinarios para evaluación y planeamiento de uso del territorio.

### **3.2.- Análisis discriminante**

El análisis discriminante es un análisis estadístico multivariante desarrollado por Fisher en 1930. Es una técnica estadística de clasificación, que permite identificar aquellas variables que discriminan entre dos o más grupos definidos con anterioridad. Por lo tanto, permite establecer diferencias entre dichos grupos. Se parte de una muestra de N sujetos en los que se han medido p variables cuantitativas independientes, que son las que se utilizarán para tomar la decisión en cuanto al grupo en el que se clasifica cada sujeto (Hair et al., 1999).

El análisis discriminante resulta útil para construir un modelo predictivo para pronosticar el grupo de pertenencia de un caso, a partir de las características observadas de cada caso. El procedimiento genera una función discriminante o, para más de dos grupos, un conjunto de funciones discriminantes, basada en combinaciones lineales de las variables predictoras que discriminan los grupos de la mejor forma posible. Las funciones se generan a partir de una muestra de casos para los que se conoce el grupo de pertenencia. Posteriormente, las funciones pueden ser aplicadas a nuevos casos que dispongan de medidas para las variables predictoras, pero de los que se desconozca el grupo de pertenencia (fuente ayuda SPSS v 15.0; Garson, 2008 a).

La variable dependiente es la variable de criterio, la cual clasifica a los sujetos con los que se crea el modelo en determinados grupos. Es una variable categórica que tiene un mínimo de dos valores. Los grupos de la variable criterio son mutuamente excluyentes.

Las variables cuantitativas del análisis discriminante son las variables independientes o predictoras. Pueden ser variables de intervalo o de razón. Estas se utilizan para distinguir o discriminar en qué grupo se clasificara a los sujetos.

El método de análisis discriminante considera una o más variable(s) dependiente(s), categórica(s) o binaria(s), además de otras variables independientes métricas o categóricas.

Algunos de sus usos más importantes son:

- Clasificar casos en grupos utilizando una función de predicción discriminante.
- Para comprobar una teoría mediante la observación de si los casos son clasificados de acuerdo a una predicción en particular.
- Para investigar diferencias entre grupos
- Para determinar el grado de influencia en la varianza de la variable dependiente que puede ser explicado por las independientes.
- Para confirmar el grado de importancia de una determinada variable independiente en la clasificación de la variable dependiente.
- Para descartar variables que se encuentren poco relacionadas a la diferenciación de grupos

A diferencia de otros métodos, el análisis discriminante tiene requisitos que deben cumplir las variables independientes para así asegurar un buen funcionamiento del método. Estas son:

- El método supone que las variables predictoras siguen distribuciones normales.
- No existe multicolinealidad
- Homogeneidad de la varianza (homocedasticidad)
- Homogeneidad de matrices de covarianza.
- La variable dependiente debe ser categórica

- Independencia de los datos.
- El tamaño de los grupos no debe ser muy diferente.
- Las variables independientes no deben tener desviación estándar cero en uno o más grupos (Garson, 2008 b).

Para probar que las variables cumplen con estas suposiciones, y también para probar la representatividad del modelo generado por el método, existe una serie de tests y criterios estadísticos que pueden ser utilizados, algunos de los cuales se mencionan y se explican brevemente en el anexo número tres, especialmente aquellos utilizados en este trabajo.

El análisis discriminante es un método que, generalmente, ha sido poco utilizado para estudios de susceptibilidad a la ocurrencia de remociones en masa, debido a los requisitos de normalidad de las variables independientes y de homogeneidad de la varianza y de la covarianza. Estas condiciones difícilmente se cumplen cuando no hay una cantidad equivalente de casos positivos y negativos o cuando los datos no se encuentran distribuidos homogéneamente con respecto a los casos positivos (que es lo más probable en geología).

El error que puede cometer una función o modelo discriminante, en términos de la clasificación dentro de grupos, es bastante alto desde el punto de vista probabilístico, ya que, cuando se busca minimizar la probabilidad de clasificar mal en una dirección, generalmente se incrementa la probabilidad de cometer un error en la otra dirección (Gorsevski, 2000).

Sin embargo, de acuerdo a Carrara (1983) y a Gussetti et al. (2006) el método de análisis discriminante, cuando está basado en una gran cantidad de variables mapeables geológicas y geomorfológicas, es capaz de discriminar exitosamente entre áreas estables e inestables.

### **3.3.- Regresión Logística (Caballero, 2008; Garson , 2008 (c))**

La regresión logística es un modelo matemático de regresión no lineal para variables dependientes categóricas, en particular para aquellas variables dicotómicas (sí o no, éxito o fracaso, entre otras). Esta característica, es la que hace que este método sea tan utilizado en la actualidad, pues existen numerosos problemas tanto en el área científica como en otras áreas cuya variable dependiente es dicotómica.

La regresión logística es una técnica estadística multivariante que permite predecir y explicar una variable categórica en lugar de una medida dependiente métrica. También permite estimar de manera directa la probabilidad de ocurrencia de un evento en particular y clasificar individuos o casos dentro de categorías.

La regresión logística resulta útil para los casos en los que se desea predecir la presencia o ausencia de una característica o resultado, según los valores de un conjunto de predictores. Es similar a un modelo de regresión lineal, pero está adaptado para modelos en los que la variable dependiente es dicotómica. La regresión logística se puede aplicar a un rango más amplio de situaciones de investigación que el análisis discriminante (fuente ayuda SPSS v15.0).

El modelo de regresión logística permite obtener una relación de regresión multivariable entre una variable dependiente y varias independientes. Una de sus grandes ventajas consiste en que las variables independientes pueden ser continuas, discretas o

categorías y no necesariamente deben seguir una distribución normal. Donde se tenga una variable dependiente binaria, la regresión logística es aplicable (Lee et al, 2006).

Este método es el que más problemas puede presentar si la cantidad de datos positivos (Variable dependiente igual a 1), es muy distinta a la cantidad de datos negativos (variable dependiente igual a 0) (Can et al., 2005). También presenta problemas si los datos son pocos o si los eventos tienen un tamaño pequeño, no siendo recomendable su uso en tales casos. Esto último fue comprobado por Can et al (2005) para el estudio de susceptibilidad a la ocurrencia de flujos de tierra (generalmente muy pequeños) producto de lluvias.

En varios de estos estudios, la regresión logística arroja resultados muy confiables y ha sido utilizado en áreas bastante grandes, obteniéndose buenos resultados, incluso en áreas tan grandes como 300 km<sup>2</sup> (Ardizzone et al., 2001) o 850 km<sup>2</sup> (Greco et al., 2006).

Los modelos de regresión pretenden conocer la relación existente entre una variable dependiente y una o más variables independientes. El tipo de variable dependiente va a determinar el tipo de modelo de regresión, y el número de variables independientes va a determinar si el modelo es multinomial o no.

A pesar de ser equivalente al método de análisis discriminante, en cuanto a resultados entregados, se pueden observar algunas diferencias, no poco importantes, algunas de las cuales son:

- La regresión logística no descansa sobre un estricto cumplimiento de la normalidad, como sí lo hace el análisis discriminante.
- La regresión logística no descansa sobre un estricto cumplimiento de la igualdad de las matrices de varianza/covarianza intergrupo, como sí lo hace el análisis discriminante.
- Es más adecuada para variables dependientes binarias.
- Los resultados admiten explotación en términos de probabilidad.
- Las variables independientes pueden ser categóricas o continuas, a diferencia del análisis discriminante, que teóricamente, solo puede utilizar variables independientes continuas (o métricas)

Los supuestos de normalidad y de igualdad de matrices de varianza / covarianza intergrupo, que exige el análisis discriminante para que el resultado sea robusto y confiable, son condiciones que no siempre se cumplen, y una de las ventajas de la regresión logística es que puede ser robusto aún frente a la ausencia de estos supuestos. Evidentemente, si los supuestos de normalidad y de homocedasticidad (entre otros) se cumplen, entonces el resultado es aún más confiable.

Otra ventaja muy importante es que solo se necesita conocer si un suceso ocurrió o no para entonces utilizar un valor dicotómico como variable dependiente.

Pero no sólo tiene diferencias con el análisis discriminante, sino que también con otros métodos de la familia de las regresiones. Tal es el caso de la diferencia que existe entre la regresión múltiple y la logística. Para la regresión logística, los valores de probabilidad pueden ser cualesquiera entre cero y uno, y, a diferencia del método de regresión lineal no asume una relación lineal entre la variable dependiente e independiente, y puede manejar efectos no lineales incluso cuando términos exponenciales o polinómicos no han sido agregados como variables independientes.

Lo que distingue a la regresión logística de otros métodos (y en particular aquéllos de la familia de las regresiones) es que utiliza el estimador de máxima verosimilitud en lugar

del método de los mínimos cuadrados para calcular la medida de ajuste global del modelo, pues la metodología que utiliza para estimar los coeficientes del modelo es prácticamente la misma que la utilizada en el método de regresión lineal.

Al igual que el método de análisis discriminante la regresión logística posee ciertas suposiciones que idealmente deben ser cumplidas, pero, a diferencia del análisis discriminante, la regresión logística es un método mucho más robusto cuando no se cumplen fielmente estas suposiciones. Estas suposiciones son:

- La variable dependiente es dicotómica (al menos) y las variables independientes son continuas o categóricas.
- Codificación significativa: La variable independiente debe representar el factor más importante, si son más de dos, deben estar ordenados de menor a mayor importancia.
- Todas las variables independientes relevantes están incluidas en el modelo
- Todas las variables independientes irrelevantes han sido excluidas del modelo.
- El muestreo es independiente (por lo que, los errores son independientes)
- Idealmente asume bajo error de medición y no existencia de casos perdidos en las variables independientes.
- **Relación lineal entre logit y variables independientes.**
- **No existe multicolinealidad**
- No existen datos aberrantes.
- Muestras grandes (con al menos diez).
- No toma en cuenta las interacciones entre variables independientes, especialmente cuando una es función de otras.
- Cada grupo formado por la variable dependiente, debe tener una frecuencia igual o mayor que uno en las variables independientes.

Como se puede ver a simple vista, la mayoría de estas suposiciones dependen de quien crea la base de datos, ya que es en esta etapa en donde se puede asegurar si estas suposiciones fueron cumplidas.

Sin embargo existen otras suposiciones (marcadas en negritas) que necesitan ser testeadas. Para el caso de la multicolinealidad los pasos a considerar son los mismos que los utilizados en el caso del análisis discriminante para evitar este mismo problema. Los tests utilizados para comprobar las suposiciones básicas, así como también los pasos a seguir para la construcción del modelo son descritos con mayor detalle en el anexo número tres.

### **3.4 Variables independientes**

Las variables independientes utilizadas varían, dependiendo de lo que se quiera demostrar o predecir, así como también varía la metodología, o sea, la forma de definir y recolectar los datos.

Las variables independientes que generalmente son utilizadas en este tipo de procedimientos son:

- Topográficas: altura, elevación, pendiente, aspecto, curvatura (en planta y en perfil)
- Geológicas: estructuras (distancia, y densidad), geología (litología, alteraciones, tipo de suelo, espesor del suelo), entre otras

- Hidrográficas: densidad de drenaje (en subcuenca y por unidad de área), área de subcuenca, distancia a red de drenaje (puede considerarse permanente u ocasional).
- Geotécnicos: en particular resistencia, meteorización, grado de fracturamiento, entre otros.
- Vegetación: generalmente, cuando se utiliza, es la cantidad por unidad de área, y también se utiliza el tipo de vegetación.
- Uso de suelos: ciudad, plantaciones, entre otros.
- Climática: generalmente se utiliza el valor medio típico de variables climáticas de un área en particular.
- Sísmicos: en este ámbito se consideran parámetros tales como amplificación, susceptibilidad, entre otros.

Los factores condicionantes considerados como importantes varían según el criterio del investigador. Algunos sólo consideran variables topográficas (Davis et al., 2006), o topográficas más geología (Ayalew & Yamagishi, 2004), densidad de drenaje, factores climáticos, estructuras, entre otras (Concha Dimas et al., 2007; Eeckhaut et al. (2006)).

También está el caso de las variables que son desencadenantes de eventos de remoción en masa, pero que se consideran como condicionantes. Los factores gatillantes más importantes son el climático seguido por el sísmico y, es por esto que es importante ver cuan susceptible es un determinado sector a estos elementos.

### **3.5 Variables dependientes**

También varía de un autor a otro la manera de considerar las evidencias. Esto va a depender del problema que se este estudiando. Por ejemplo, en el caso de la susceptibilidad a la ocurrencia de remociones en masa, puede resultar de gran importancia el tamaño de las remociones y la escala de trabajo, entre otros. Para el caso de las remociones en masa las evidencias son consideradas de distintas formas:

- Algunos las consideran como puntos, ya sea en el escarpe, o en el depósito, o en la mitad de la remoción como un todo.
- Otros consideran líneas alrededor de los escarpes de deslizamiento.
- Otros consideran polígonos, que pueden cubrir tanto el área completa de la remoción como un sector en particular.

## **Capitulo IV: Resultados sin considerar variables sísmicas**

Las variables utilizadas en cada método varían de acuerdo con los criterios de selección de variables correspondiente a cada uno de ellos, así como también con las suposiciones. Nuevamente es importante mencionar que estos métodos no se basan en un cumplimiento estricto de las suposiciones y que, si bien es deseable que se cumplan, el no cumplimiento de estas suposiciones sólo afecta levemente los resultados. Esto es particularmente cierto para el caso del método de regresión logística.

Las variables consideradas en este trabajo, con sus respectivos códigos se muestran en la tabla 4.1. Para mayor información ver anexo dos.

**Tabla 4.1:** Variables consideradas en este trabajo y sus respectivos códigos

<b>Variables Topográficas</b>	<b>Código</b>	<b>Tipo</b>
Altura	Talt	Continua
Elevación	Telev	Continua
Pendiente	Tp	Continua
Aspecto	Tasp	Continua
Curvatura en perfil	Tper	Continua
Curvatura en planta	Tpla	Continua
<b>Variables relacionadas con Geología</b>		
Geología completa	Gt	Categórica
Fm. Abanico	Ga	Binaria
Fm. Farellones	Gf	Binaria
Intrusivos	Gi	Binaria
Depósitos coluviales	Gc	Binaria
Depósitos aluviales	Gal	Binaria
Depósitos fluviales	Gfl	Binaria
Depósitos lacustres	Gl	Binaria
Ignimbrita Pudahuel	Gip	Binaria
Depósitos de remoción en masa	Grem	Binaria
Distancia euclidiana a estructuras	Geests	Continua
Densidad de estructuras	Gdests	Continua
<b>Variables relacionadas con Hidrografía</b>		
Distancia euclidiana a la red de drenaje permanente	Hedp	Continua
Distancia euclidiana a la red de drenaje ocasional	Hedo	Continua
Área de subcuenca	Has	Continua
Densidad de drenaje	Hds	Continua
Longitud de drenaje	Hls	Continua
<b>Variables relacionadas con Vegetación</b>		
NDVI (índice de vegetación)	Hv	Discreta

## **4.1 Metodología de selección de variables**

La metodología utilizada para la selección de las variables corresponde a la propuesta por Hosmer & Lemeshow (2000), y consta de dos pasos básicos:

- 1) Regresión logística univariable.
- 2) Regresión logística por pasos hacia adelante y hacia atrás.

### 4.1.1 Regresión logística univariable

En este paso se determina si una variable puede o no ser importante para el modelo. El criterio es que el p-valor (ver anexo 3) resultante del test de Hosmer & Lemeshow debe ser menor que 0.25 pues, de lo contrario, demostraría una relación excesivamente alta con la variable dependiente. Los resultados de este test se muestran en la tabla A4.1 y de acuerdo con ella, todas las variables cumplen con este criterio menos la variable independiente Longitud de drenaje (Hls). Un punto importante es que esta prueba sólo es realizable para las variables independientes continuas o categóricas no dicotómicas y, por lo mismo en todas aquellas categóricas dicotómicas el chi cuadrado es cero (ver tabla A4.1).

### 4.1.2 Regresión por pasos hacia adelante y hacia atrás

En este paso se ve qué tan importante puede ser una variable dentro de un grupo de variables e identifica aquellas menos importantes dentro del mismo grupo de variables. Las variables van siendo incluidas una a una dentro del modelo y se observa el impacto que causa su inclusión ó exclusión.

Esta prueba se realiza sin la variable Longitud de drenaje (Hls), ya que fue rechazada en el paso anterior.

En la tabla A4.2 (variables que quedaron seleccionadas por el método por pasos hacia adelante) no se encuentra la variable curvatura en perfil. Sin embargo, en la tabla A4.3 (variables seleccionadas por el método por pasos hacia atrás) si se encuentra, por lo que se escoge dejarla en el modelo. Además, ninguna de las otras variables ingresadas fue rechazada.

Se decidió dejar en el modelo todas las variables mostradas en la tabla 4.1, con excepción de la variable Longitud de drenaje.

Si bien esta forma de selección de variables es utilizada para el método de regresión logística, se decidió utilizar las mismas variables (normalizadas) en el método de análisis discriminante. Sin embargo, como la variable geología completa (Gt) es utilizada sólo con el método de análisis discriminante por incluir todas las otras variables binarias que representan la geología (tabla 4.1) con excepción de los depósitos de remoción en masa (Qrem) las variables binarias que representan: las formaciones Abanico y Farellones, los intrusivos, los depósitos aluviales, coluviales, fluviales, e Ignimbrita Pudahuel, no son incluidas dentro del análisis por el método de análisis discriminante (ver anexo dos).

## **4.2 Regresión Logística**

El modelo fue escogido basándose en el criterio de selección de variables de Hosmer & Lemeshow, siempre cuidando que el modelo resultara suficientemente representativo del problema en cuestión.

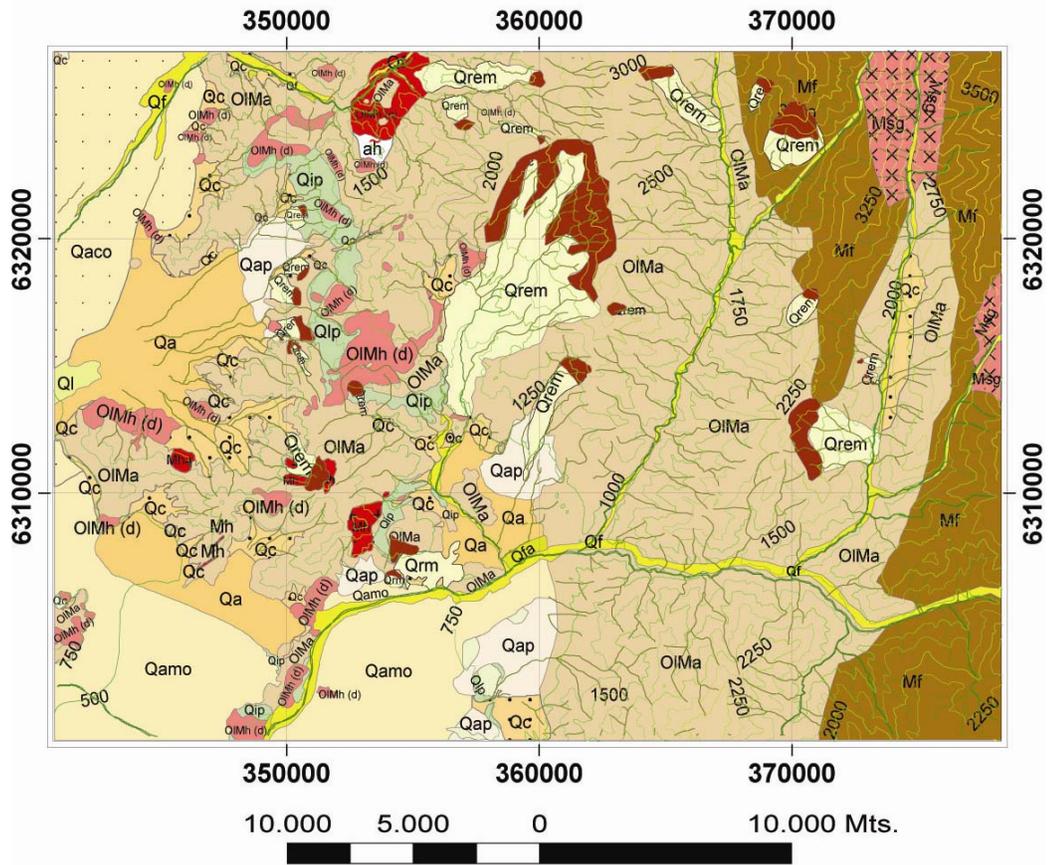
El criterio utilizado para la selección de variables, una vez utilizado el método de Hosmer & Lemeshow, fue la colinealidad dado que, el no cumplimiento de esta condición era lo que más podía perjudicar el resultado del método de regresión logística (Tabla A4.4), ya que el mejor modelo es aquel en el que sus variables se encuentran débilmente relacionadas.

La variable longitud de drenaje (Hls) no fue incluida en el modelo, por haber sido rechazada en la sección 4.1.1, probablemente por encontrarse demasiado relacionada con la variable densidad de drenaje. La variable geología completa no fue incluida en el modelo, ya que ésta corresponde a una agrupación de las otras variables geológicas presentadas, correspondientes a unidades geológicas del área en estudio (Anexo 2).

La tabla A4.4 plantea otro problema relevante, el cual consiste en decidir cuál variable debe ser escogida entre la altura (Talt) y la elevación (Telev), ya que la segunda es equivalente a la primera menos una constante (Anexo 2). Para esto, se utilizó el criterio del FIV y la tolerancia en conjunto y, dado que la variable altura (Talt) posee una menor tolerancia (tabla A4.4), lo que significa que un mayor porcentaje de la misma se encuentra explicado por otras variables, y tiene además un mayor FIV que la variable elevación (Telev) (tabla A4.4), resulto elegida la variable que representa la elevación (Telev). El modelo correspondiente a este cambio es mostrado en la tabla A4.5.

En la tabla A4.5 se aprecian cambios bastante importantes con respecto a la tabla A4.4, en particular, luego de extraer definitivamente la variable altura (Talt), dejando sólo la elevación (Telev). Sin embargo, aún persiste un problema, y éste es con la variable que representa a la Formación Abanico, y que puede deberse a la gran extensión areal que esta variable posee dentro del área de estudio, lo que la hace aparecer casi como una constante frente al resto de las variables. Al observar la tabla de correlaciones entre las variables para la regresión logística (Anexo 7), se puede ver que la correlación entre la variable que representa a la Formación Abanico con las demás variables no es tan grande como la que se observa en las variables que representan los depósitos aluviales (Gal), fluviales (Gfl), lacustres (Gl) y a la Ignimbrita Pudahuel (Gip) con las demás.

Luego, de acuerdo con lo discutido previamente, si se extraen las variables relacionadas a depósitos aluviales (Gal), fluviales (Gfl), lacustres (Gl) y a la Ignimbrita Pudahuel (Gip), se obtiene la tabla A4.6 donde el problema relacionado a la Formación Abanico disminuye de manera importante. La Formación Abanico es considerada como la más importante en este estudio debido a que la gran mayoría de los escarpes presentes en el área de estudio se encuentran en dicha Formación (figura 4.1) y es por esta razón que, pese a su aparente relación con alguna de las otras variables, es importante que este presente en el modelo, pues, de otra manera, se perdería representatividad. No ocurre lo mismo con la eliminación de las variables que corresponden a la presencia de depósitos lacustres, aluviales y fluviales y también a la Ignimbrita Pudahuel, ya que estas variables no están relacionadas con la ocurrencia de deslizamientos por corresponder a zonas de baja pendiente y por su escasa presencia en el área en estudio.



**Figura 4.1:** Mapa geológico con escarpes de remociones consideradas en el área en estudio marcadas en café. Se observa que la gran mayoría de los escarpes (en color café) se ubican en la Formación Abanico. Tomado y modificado de Wall et al. (1999), mapa geológico de Chile (1:1.000.000), Thiele (1980) y trabajo de terreno.

Para este modelo se realizaron distintos test, especialmente para medir su grado de representatividad (ver anexo 3 para más detalles).

Tests de ajuste del modelo a los datos:

- Hosmer & Lemeshow: En la tabla A4.7 se puede ver que se rechaza la hipótesis nula, correspondiente a que el modelo se ajusta a los datos (significancia es menor que 0,05). En la tabla A4.8 se puede ver una significativa diferencia entre los valores predichos y los observados para el caso positivo.
- Test Ómnibus: Este test es una alternativa un poco menos robusta al test de Hosmer & Lemeshow, en donde el resultado significativo (Sig. < 0,05) observado en la tabla A4.9 indica que el modelo se adecua bastante bien a los datos.

La tabla 4.2 muestra los estadísticos de  $R^2$  de Negelkerke y de Cox y Snell. Los valores de estos estadísticos muestran la cantidad de variabilidad que puede ser explicada por el modelo. A mayor valor de  $R^2$  mejor es el modelo, además es necesario mencionar que su uso no es recomendado para medir bondad de ajuste del modelo a los datos, sino que solo su representatividad. El  $R^2$  de Negelkerke es el de interpretación más directa, considerando que el mejor modelo es aquel con menores problemas de colinealidad. En este caso, se obtuvo un  $R^2$  de Negelkerke de 0,313, el cual no es suficientemente bueno (tabla 4.2).

**Tabla 4.2:** R<sup>2</sup> de Cox y Snell y de Negelkerke.

	R cuadrado de Cox y Snell	R cuadrado de Negelkerke
Sin variables sísmicas	0,072	0,313

Tabla de clasificación: Se asumió una probabilidad de corte de 0.1, esto es, se considera que todos aquellos casos que presenten una probabilidad de 0.1 o superior, son consideradas como zonas altamente susceptibles. En la tabla 4.3 se muestra el resultado obtenido bajo este criterio, en donde se obtuvo que un 93.4% de los casos se encuentra bien clasificado y, aún más importante, un 58,9% de los casos positivos se encuentran bien clasificados, lo que si bien no es bueno, se considera aceptable.

**Tabla 4.3:** Tabla de clasificación con probabilidad de corte 0.1 que muestra el número de píxeles clasificados como casos positivos para probabilidades iguales o superiores a 0,1, o zonas altamente susceptibles, y casos negativos o baja susceptibilidad para probabilidades menores a 0,1, para el modelo mostrado en la tabla A4.7.

Observado		Pronosticado		
		Escarpe		Porcentaje correcto
		0	1	0
Escarpe	0	285334	16737	94,5
	1	3700	5309	58,9
Porcentaje global				93,4

a El valor de corte es ,100

En este modelo no se le da importancia al hecho de que el test de bondad de ajuste de Hosmer & Lemeshow sea rechazado ya que puede ser originado por la gran cantidad de casos utilizados en este trabajo (son 311.080 casos), situación en la que una pequeña desviación puede provocar que el modelo sea rechazado.

Los signos de estos coeficientes indican qué tan importantes son estas variables, a medida que se incrementa o disminuye su valor. Un coeficiente positivo indica que a medida que aumenta el valor de la variable esta es más importante dentro del modelo y un signo negativo indica lo contrario. En el caso de la geología, el programa estadístico SPSS v 15.0 cambia la codificación de la variable, dejándola en cero si es que la característica se encuentra presente o en valor uno en el caso contrario, por lo que el signo negativo que llevan estas variables muestra (de alguna manera) qué tan importante es su ausencia dentro del modelo.

Los coeficientes de la función logística son mostrados en la tabla 4.4. En ella se puede ver que ninguna de las variables incluidas en el modelo es rechazada (Sig. < 0.05 en todos los casos).

De la tabla 4.4 se reconocen como las variables más importantes las siguientes:

- Densidad de drenaje.
- Densidad de estructuras.
- Geología.
- Curvatura en planta.
- Curvatura en perfil.

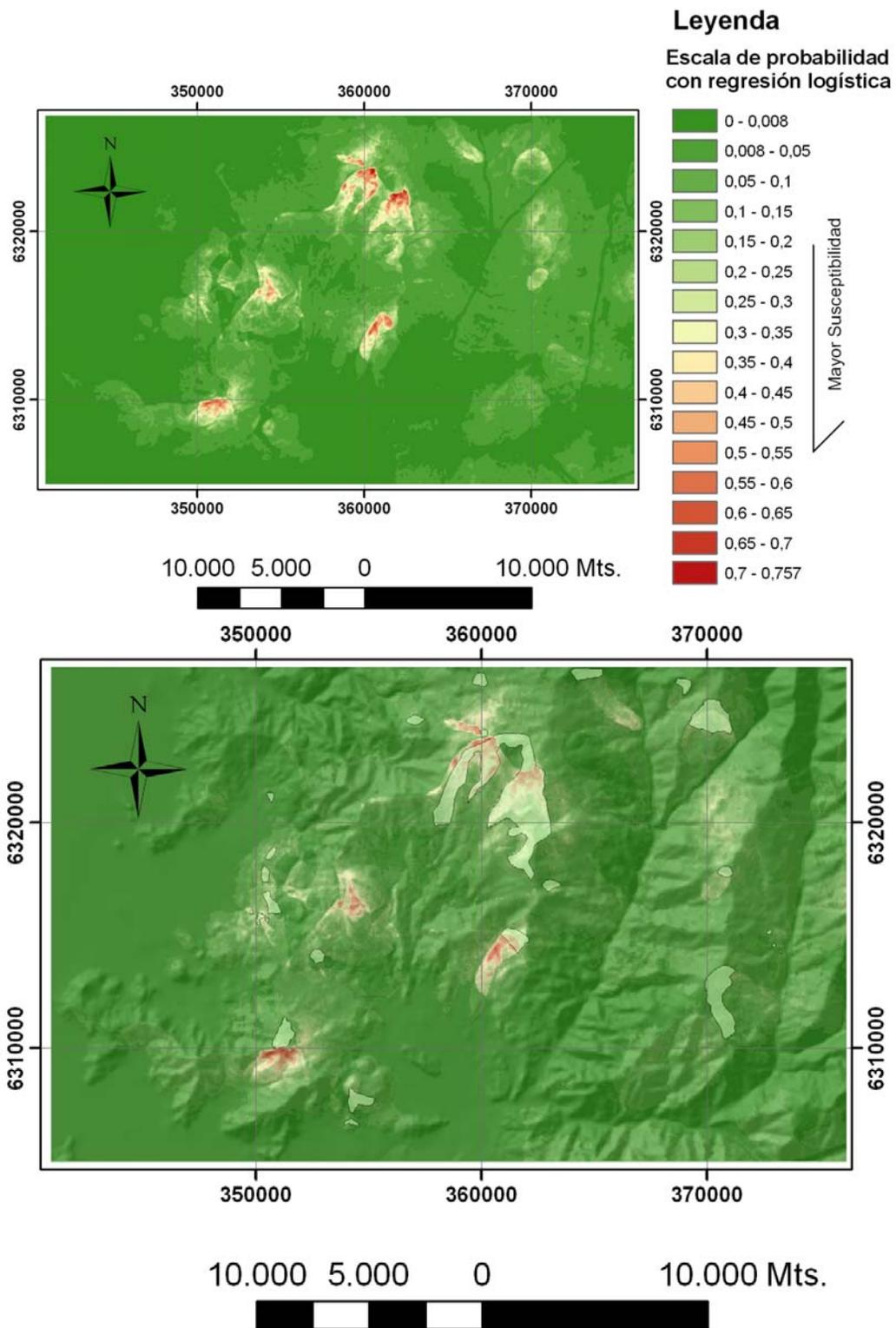
Otro índice es el Exp(B), el cual indica qué tanto cambia la importancia de una determinada variable independiente cuando el valor de ésta se incrementa en una unidad. Entonces el Odd cambia en el factor que indica Exp (B), por lo que se incrementa la probabilidad de ocurrencia del evento en ese factor, o, lo que es lo mismo,

se incrementa la susceptibilidad. La tabla 4.4 muestra que la variable densidad de estructuras es una de las más importantes dentro del modelo. Todas aquellas variables cuyo Exp(B) es uno o cercano a este indica que la variable es poco importante dentro del modelo.

**Tabla 4.4:** Coeficientes de variables (en negritas).

	<b>B</b>	Sig.	Exp(B)
Telev	<b>,001</b>	,000	1,001
Tp	<b>,037</b>	,000	1,038
Tasp	<b>,001</b>	,000	1,001
Tper	<b>-,487</b>	,000	,615
Tpla	<b>,294</b>	,000	1,342
Hv	<b>-,015</b>	,000	,985
Ga	<b>-2,680</b>	,000	,069
Gf	<b>-2,854</b>	,000	,058
Gi	<b>-1,531</b>	,000	,216
Gc	<b>-1,921</b>	,000	,146
Hedo	<b>-,001</b>	,000	,999
Hedp	<b>,001</b>	,000	1,001
Hds	<b>-5,788</b>	,000	,003
Has	<b>,001</b>	,000	1,001
Gdests	<b>1,533</b>	,000	4,630
Geests	<b>-,001</b>	,000	,999
Grem	<b>-1,559</b>	,000	,210
Constante	<b>10,724</b>	,000	45421,796

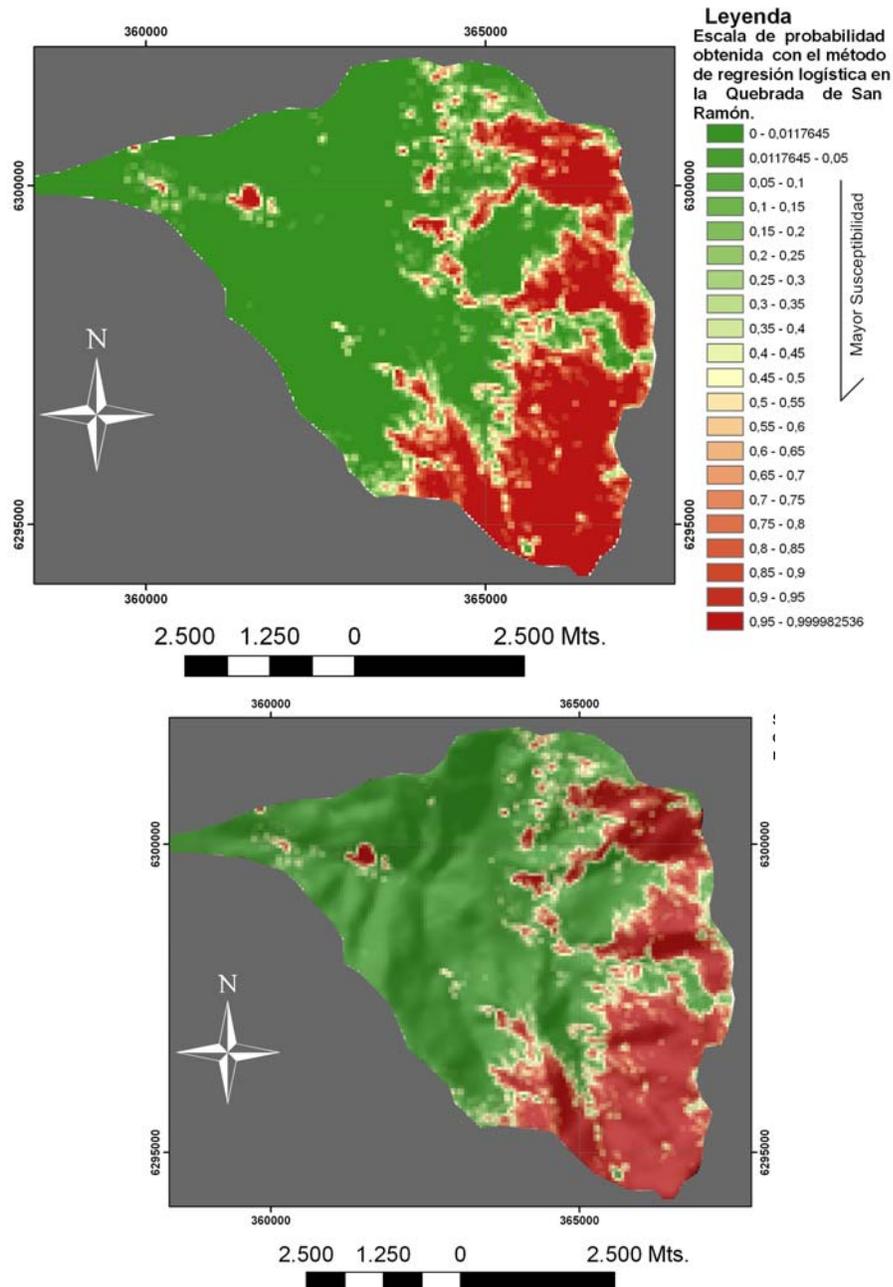
Se espera que el resultado de este modelo prediga de manera correcta, especialmente en aquellos casos donde se sabe que son casos positivos (zonas predefinidas por la variable dependiente escarpe). El resultado obtenido con este modelo es mostrado en la figura 4.2, en donde se puede ver que predice de manera bastante aceptable, especialmente en aquellas zonas donde si se sabe que es susceptible. Por esto se considera que el modelo es bueno y representativo.



**Figura 4.2:** Escala de probabilidad que representa la susceptibilidad a la ocurrencia de remociones en masa en el área en estudio, obtenida mediante el método de regresión logística en la zona en estudio. Arriba susceptibilidad a la ocurrencia de remociones en masa sin relieve de fondo, abajo susceptibilidad a la ocurrencia de remociones en masa con relieve de fondo y escarpes.

### 4.3 Validación en la Quebrada de San Ramón (regresión logística)

Con el propósito de asegurar un buen desempeño de este método, se procedió a su validación mediante una prueba en un área más pequeña, ubicada al Sur del área en estudio, correspondiente a la hoya hidrográfica de la Quebrada de San Ramón, cuyas características generales fueron descritas en el capítulo 2. Los resultados obtenidos en esta área (figura 4.3) son bastante consistentes con lo observado en la hoya hidrográfica de la Quebrada de San Ramón (figuras 4.8 y 4.9) (Lara, 2007).



**Figura 4.3:** Escala de probabilidad que representa la susceptibilidad a la ocurrencia de remociones en masa obtenidas mediante el método regresión logística en la Quebrada San Ramón. Arriba susceptibilidad a la ocurrencia de remociones en masa sin relieve de fondo, abajo susceptibilidad a la ocurrencia de remociones en masa con relieve de fondo.

## 4.4 Análisis Discriminante

Las variables utilizadas en este método fueron normalizadas mediante un método gráfico por cuantiles. Los tests de normalidad realizados a las variables normalizadas pueden encontrarse en el anexo 1. El hecho de que la condición de normalidad no se cumpliera en todas las variables utilizadas (debido a la imposibilidad de transformarlas), no se consideró como un problema y, si bien se intentó, se continuó adelante con estas variables sin normalizar para no perder representatividad en el modelo.

La condición de multicolinealidad (Tabla A4.10), a diferencia del caso de regresión logística exige que la variable que representa la geología sea “Geología completa”, la cual es una variable categórica descrita en el anexo 2. Nuevamente la variable altura es rechazada, por lo que es utilizada la elevación en su reemplazo. Las variables dentro del modelo generado por el análisis discriminante no presentan problemas de colinealidad, con excepción de la variable que representa la distancia euclidiana a las estructuras, la cual presenta correlaciones no despreciables con otras variables (tabla correlaciones análisis discriminante en Anexo 7).

El test de M de Box intenta demostrar la suposición de homogeneidad de las matrices de covarianza. Esta propiedad difícilmente se cumple, y este caso no es la excepción, dada la significancia (o  $p$ -valor  $< 0.01$ ) la cual es mostrada en la tabla A4.11

Sin embargo, de acuerdo al método de logaritmo de los determinantes (tabla A4.12), se encuentra que, si bien la variabilidad intra grupo es grande, la variabilidad intergrupos no lo es tanto, por lo que se puede asumir que existe homogeneidad en las matrices de covarianza a un nivel aceptable.

El test de igualdad de medias de grupo de Lambda de Wilks mide el potencial de cada variable independiente antes de que el modelo sea creado. Una variable independiente contribuye al modelo en la medida que sea significativa ( $p$ -valor  $< 0.01$ ), para el caso de este modelo todas las variables contribuyen (tabla A4.13). De acuerdo al Estadístico de Lambda de Wilks aquellas variables con menor valor del estadístico discriminan mejor.

Los coeficientes estandarizados (tabla A4.15) muestran, al igual que la columna con el estadístico Lambda de Wilks (tabla A4.14), cuáles son las variables con mejor capacidad discriminante. De acuerdo al criterio de los coeficientes estandarizados, las variables con mayor capacidad discriminante son aquellas con mayores valores absolutos. De acuerdo a la tabla A4.15, las variables que representan densidad de estructuras, distancia euclidiana a estructuras, hidrografía permanente, la geología, el área de la subcuenca y remociones, son las variables más importantes dentro del modelo.

La matriz de estructura muestra el grado de correlación existente entre las variables independientes y la función discriminante, en este caso, en la tabla A4.15 se ve que nuevamente la variable densidad de estructuras presenta una importancia mayor, seguida de cerca por la distancia euclidiana a las estructuras y por los depósitos de remoción en masa.

Sin embargo, de acuerdo con la tabla 4.5, la correlación canónica indica que la capacidad discriminante de esta función no es muy buena.

**Tabla 4.5:** Capacidad discriminante de función canónica.

Autovalor	% de varianza	% acumulado	Correlación canónica
0,075	100,0	100,0	0,265

La tabla de clasificación para este método, que asume una probabilidad a priori igual para ambos grupos, es mostrada en la tabla 4.6. No hay que pasar por alto que el programa estadístico SPSS v15.0 estima la probabilidad en el caso del análisis discriminante mediante la regla de Bayes pues lo que los coeficientes estandarizados permiten obtener son las puntuaciones discriminantes para cada uno de los casos. El hecho de considerar una probabilidad a priori de 0,5 permite considerar ambos grupos como iguales, lo que minimiza el efecto de la diferencia de tamaños entre grupos.

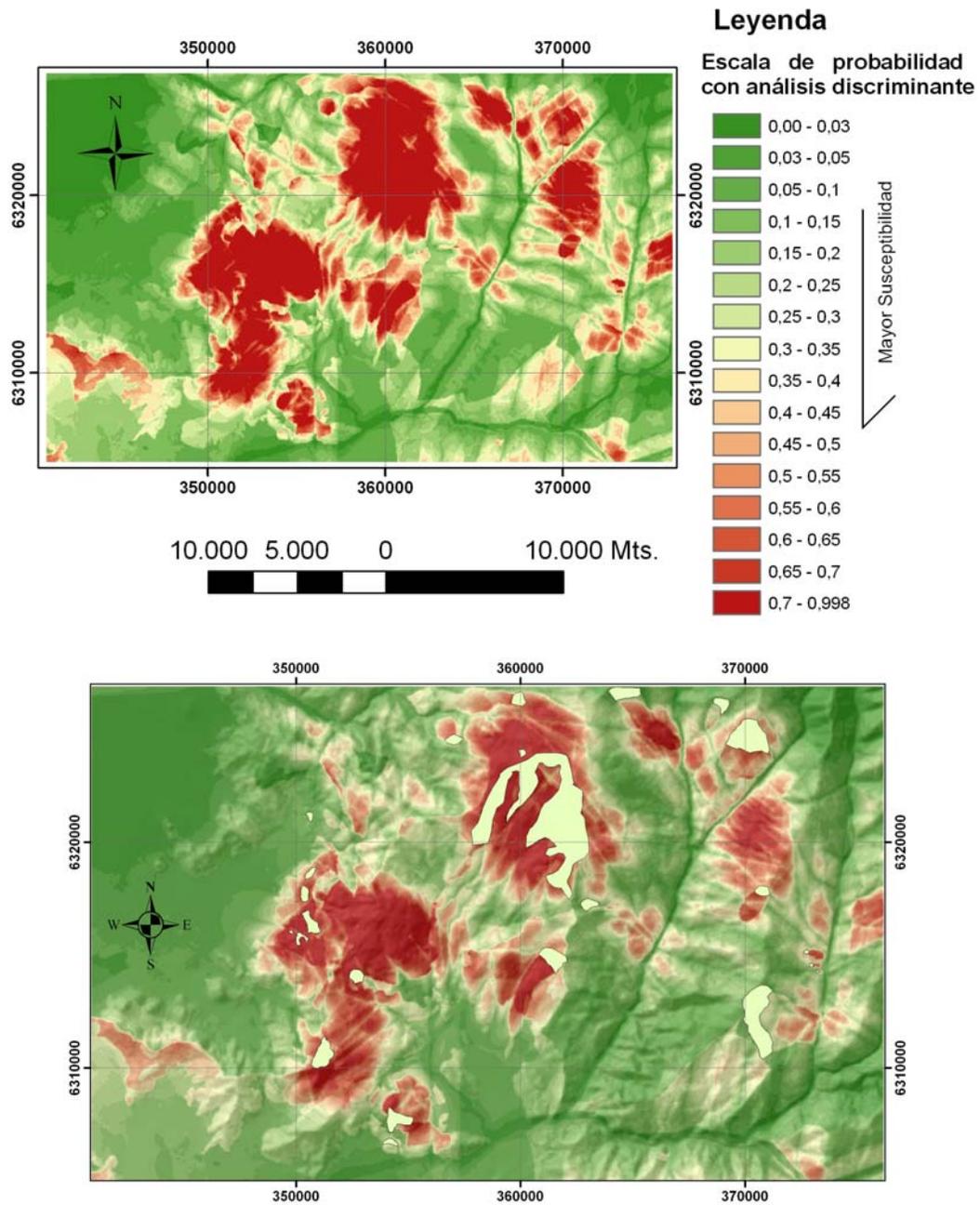
**Tabla 4.6:** Tabla de clasificación para el método análisis discriminante con probabilidades a priori iguales para ambos grupos (se asume una probabilidad de 0,5 para cada grupo)

			Grupo de pertenencia pronosticado		Total
			0	1	
Original	Recuento	0	245951	56120	302071
		1	2074	6935	9009
	%	0	81,4	18,6	100,0
		1	23,0	77,0	100,0

a Clasificados correctamente el 81,3% de los casos agrupados originales.

De acuerdo con la tabla 4.6 y pese a que, en la tabla 4.5 se obtiene un valor de correlación bajo, el porcentaje de acierto en los casos positivos es mayor que el observado con el método de regresión logística.

Efectivamente, al representar de manera gráfica la susceptibilidad a la ocurrencia de remociones en masa en el área en estudio (figura 4.4), ésta es aparentemente mucho más representativa, si bien, le asigna probabilidades superiores a 0.5 a zonas de muy baja pendiente, esto es, le asigna susceptibilidades muy altas a zonas que en realidad no la tienen, esto es, especialmente notorio en el sector suroeste del área en estudio (figura 4.4). También se puede ver una fuerte relación con las estructuras, lo que corrobora lo ya observado en las tablas. Esta situación puede deberse al hecho de haber ignorado la diferencia de tamaño entre los grupos.

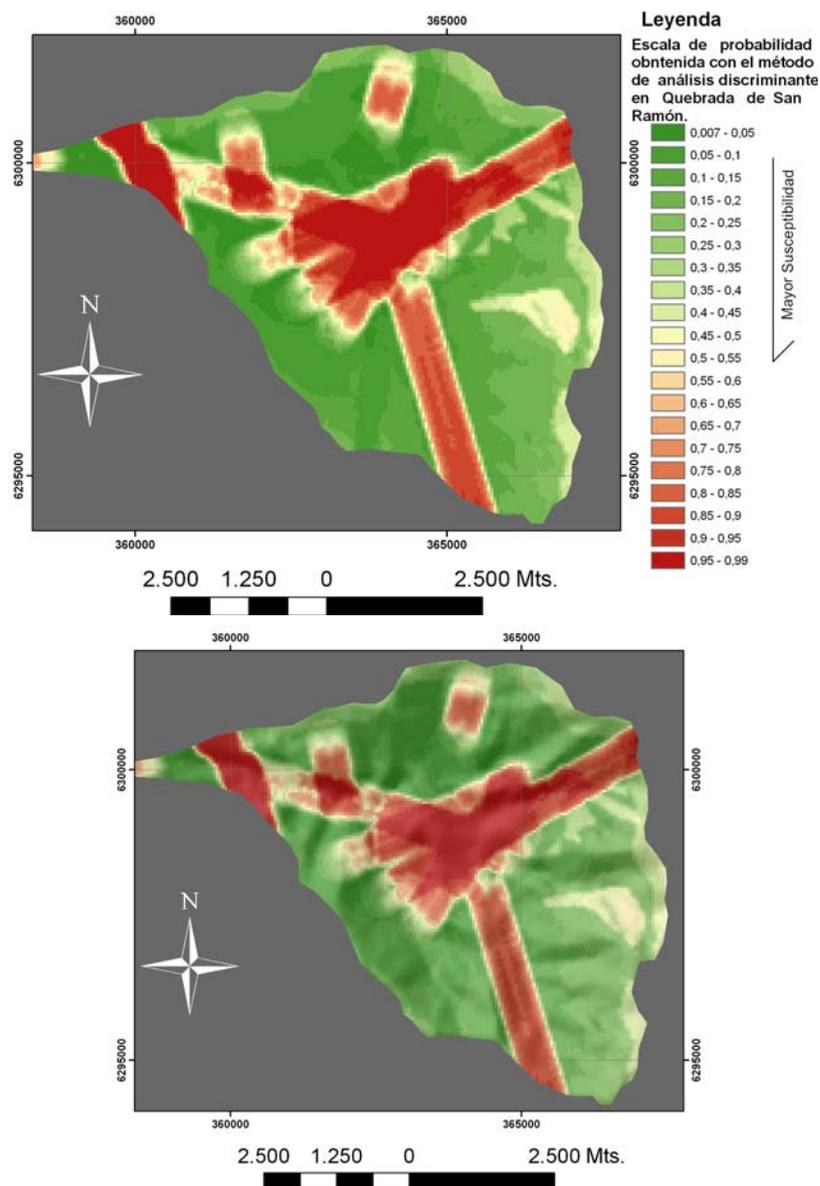


**Figura 4.4:** Escala de probabilidad que representa la susceptibilidad a la ocurrencia de remociones en masa en el área de estudio obtenida mediante el método análisis discriminante en la zona de estudio. Arriba sin relieve de fondo, abajo con relieve de fondo y escarpes.

Se aprecia una correlación bastante buena entre las zonas con escarpe y las zonas con probabilidades altas, así como también en zonas con mayor pendiente. Por otro lado no asigna probabilidades demasiado bajas a las laderas de los cerros. Es importante mencionar que en la figura 4.4 se puede apreciar una fuerte correlación del resultado obtenido con las variables densidad de estructuras y distancia euclidiana a las estructuras (ver figuras 2.27 y 2.28 en capítulo 2),

#### 4.5 Validación en Quebrada de San Ramón (análisis discriminante).

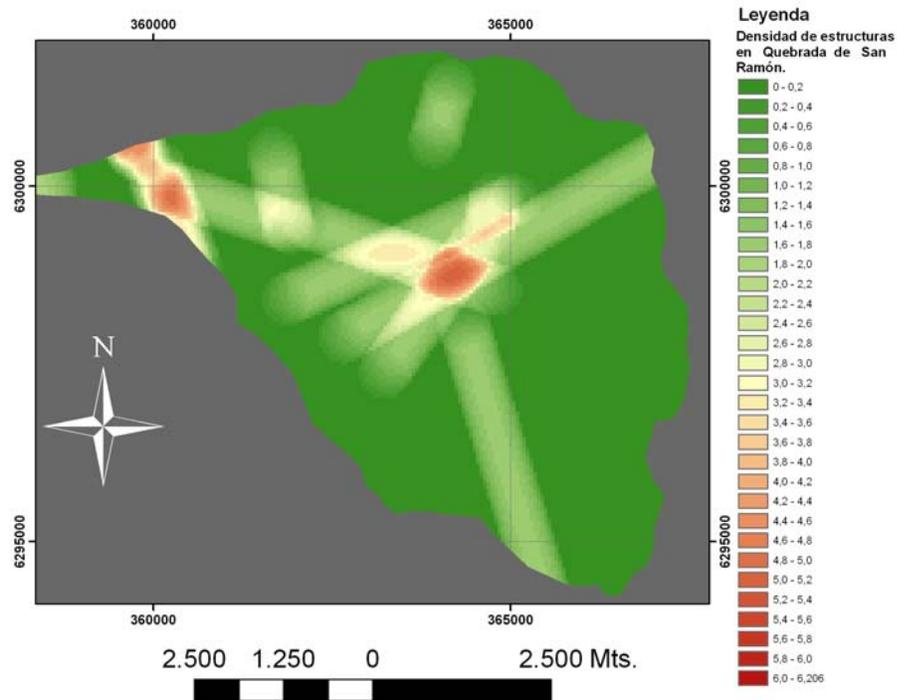
Una vez transformados los datos (normalizados) de la Quebrada de San Ramón (mediante la misma transformación aplicada en los datos del área de estudio), se pudieron estimar las probabilidades mediante el método de Bayes (esto fue realizado en el programa R2.6.2). En los resultados obtenidos en el área de prueba se puede observar una correlación no menor con las variables densidad de estructuras y distancia euclidiana a las estructuras (figura 4.5) lo que no concuerda con lo observado en esta área por Lara (2007). Es necesario mencionar que es bastante lógico el resultado obtenido en el área de prueba dada la gran importancia que le asigna el método análisis discriminante le asigna a las variables densidad de estructuras y distancia euclidiana a las estructuras.



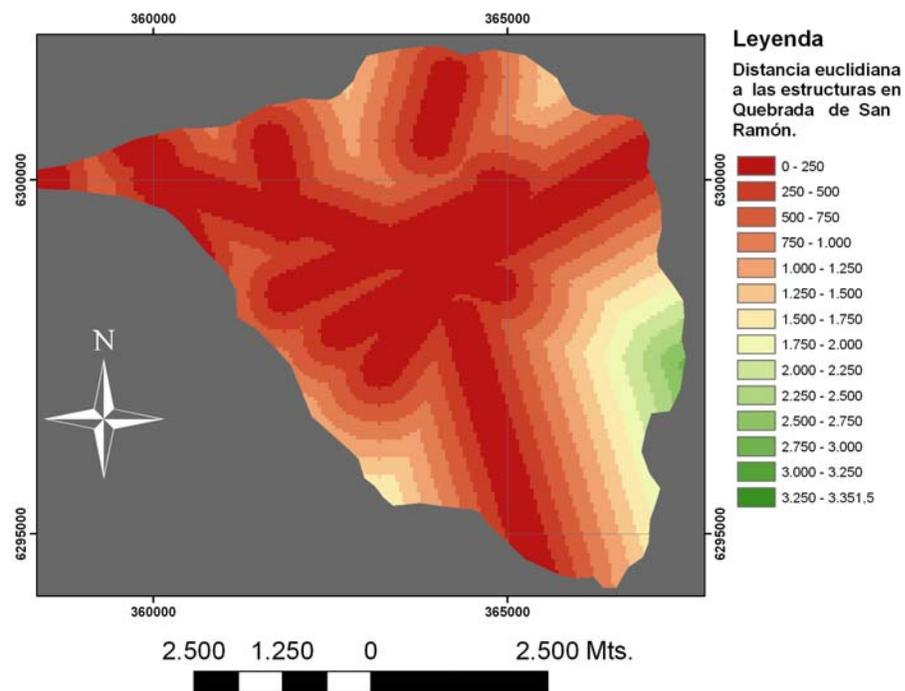
**Figura 4.5:** Escala de probabilidad que representa la susceptibilidad a la ocurrencia de remociones en masa obtenidas mediante el método análisis discriminante en la Quebrada de San Ramón. Abajo con relieve de fondo y arriba sin relieve de fondo.

Si la figura 4.5 es comparada con la figura 4.6 (densidad de estructuras y distancia euclidiana a las estructuras en la Quebrada de San Ramón), se puede ver una similitud muy marcada, lo que demuestra el grado de importancia de estas variables dentro del modelo.

a)



b)



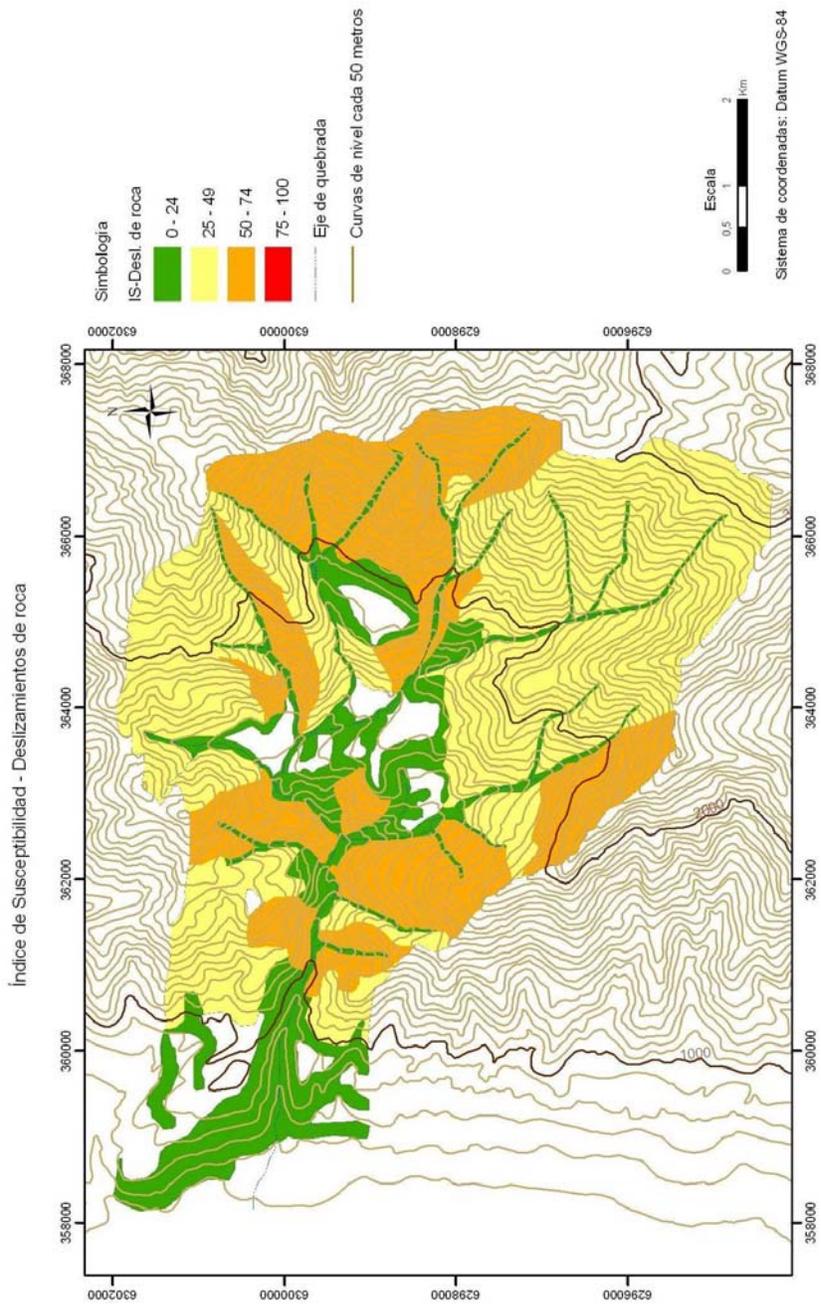
**Figura 4.6:** a) densidad de estructuras en la Quebrada de San Ramón; b) Distancia euclidiana a las estructuras en la Quebrada de San Ramón

#### **4.6.- Resultados Obtenidos por Lara (2007)**

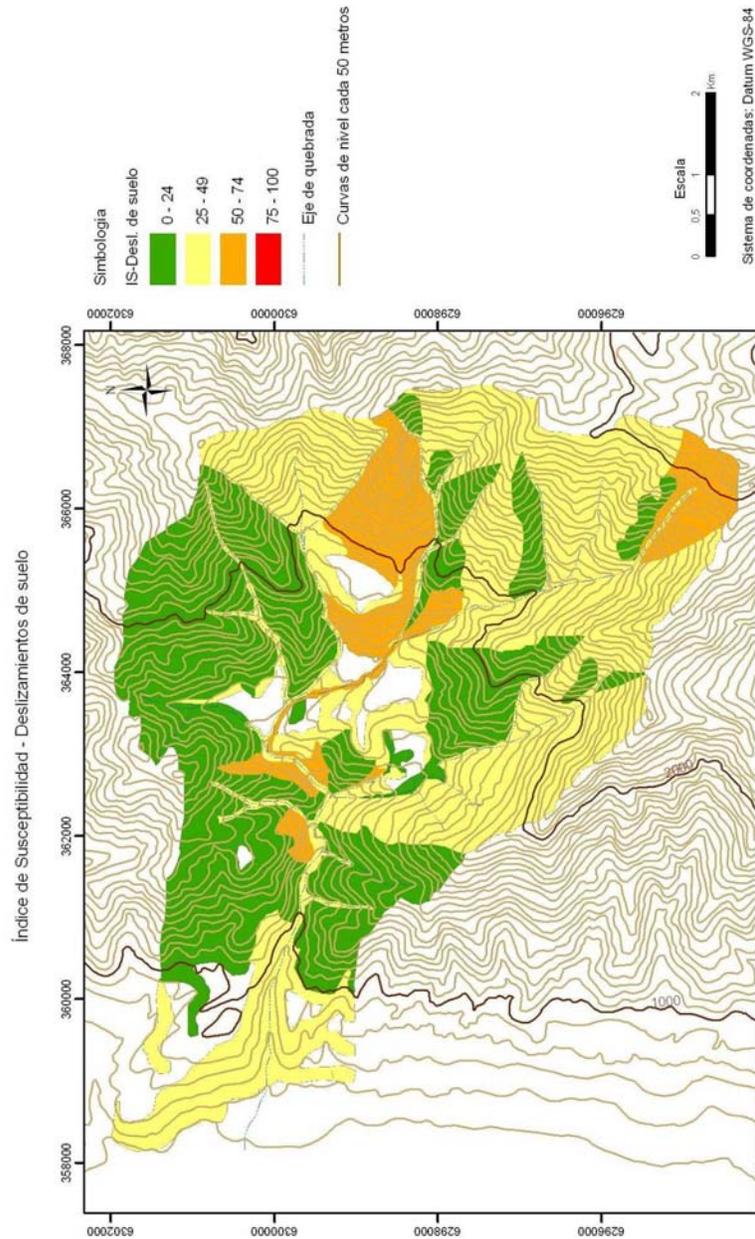
El trabajo en la Quebrada de San Ramón de Lara (2007) es de tipo cualitativo y esta constituido por las siguientes partes:

- Susceptibilidad a deslizamientos en roca
- Susceptibilidad a deslizamientos en suelo
- Susceptibilidad a caídas de roca
- Susceptibilidad a flujos

Dado que el objetivo de este trabajo es estimar la susceptibilidad a la ocurrencia de deslizamientos, los resultados obtenidos por Lara (2007) serán comparados con los resultados obtenidos en el presente trabajo. Con este fin, en la figura 4.7 se muestra el resultado de susceptibilidad a deslizamientos en roca, y en la figura 4.8 la susceptibilidad a deslizamientos en suelo obtenidos por Lara (2007).



**Figura 4.7:** Índice de susceptibilidad a deslizamientos en Roca obtenido por Lara (2007) en la Quebrada de San Ramón.



**Figura 4.8:** Índice de susceptibilidad para deslizamientos en suelo obtenido por Lara (2007) en la Quebrada de San Ramón.

Al comparar visualmente las figuras 4.7 y 4.8 con las figuras 4.3 (regresión logística testeada en la Quebrada de San Ramón) y 4.5 (análisis discriminante testeado en la quebrada de San Ramón), se observa una mucho mejor correspondencia de los resultados obtenidos por Lara (2007) con los arrojados por el método de regresión logística, no obstante que el análisis discriminante le diera importancia a otras áreas que presentan también una alta susceptibilidad pero que el método de regresión logística no considera.

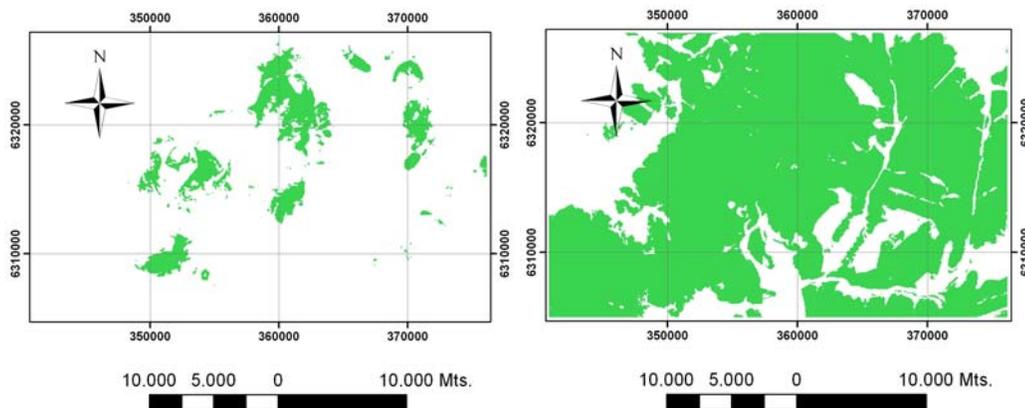
## 4.7 Comparación de resultados obtenidos por los métodos análisis discriminante y regresión logística.

En la tabla 4.7 se puede observar que los casos clasificados como positivos por ambos métodos son más que los casos reales positivos obtenidos de la información de terreno y fotogeología, lo que no es necesariamente malo, pues puede estarse cometiendo error de tipo uno, que son aquellos casos reales inicialmente clasificados como cero pero que el modelo les asigna valor uno. De acuerdo con las tablas 4.3 y 4.6 claramente no están siendo incluidos todos los casos positivos reales dentro de la clasificación, lo que corresponde a la falta de ajuste del modelo con los datos.

**Tabla 4.7:** Comparación entre métodos y casos reales para una probabilidad de corte de 0,1. P se refiere a susceptibilidad a la ocurrencia de remociones en masa, RL: regresión logística, y AD: análisis discriminante.

	Casos reales 0	Casos reales 1	Casos RL 0 (P<0,1)	Casos RL 1 (P>0,1)	Casos AD 0 (P<0,1)	Casos AD 1 (P>0,1)
<b>Sin variables sísmicas</b>	302071	9009	287796	22138	92261	217673

Al comparar gráficamente lo observado en la tabla 4.7 se obtiene la figura 4.9, en donde se ve claramente que el análisis discriminante presenta una mayor cobertura de superficie que la regresión logística en los casos considerados como positivos, sin respetar las zonas declaradas inicialmente como más susceptibles como por ejemplo los escarpes de deslizamiento.



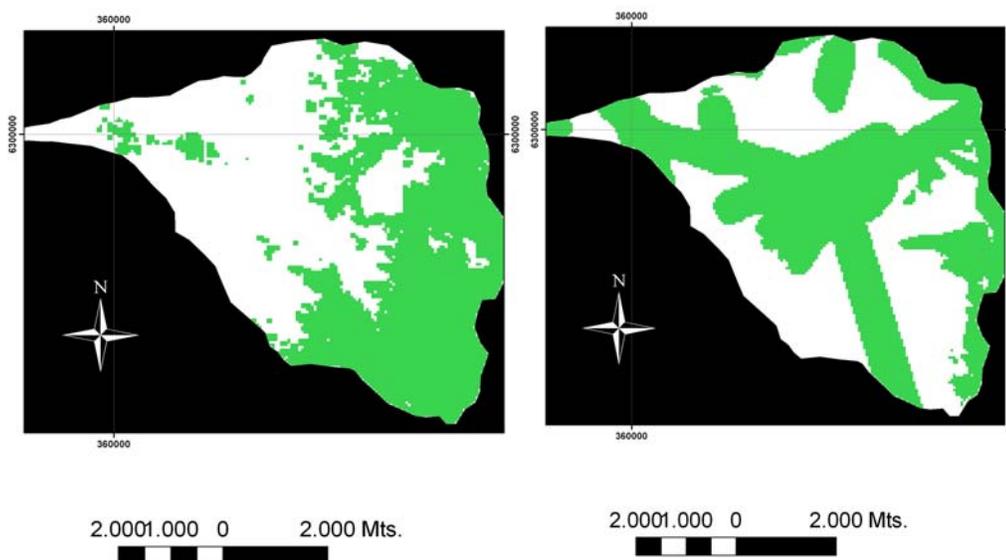
**Figura 4.9:** En blanco se encuentran los sectores con susceptibilidad a la ocurrencia de remociones en masa menor que 0,1, y en verde aquellas áreas con susceptibilidad mayor a 0,1. A la izquierda se encuentran los resultados obtenidos con el método de regresión logística, y a la derecha los obtenidos con el método de análisis discriminante.

Al comparar lo obtenido por Lara (2007) y lo obtenido por los métodos de regresión logística y análisis discriminante, en términos de área y probabilidades, se obtiene la tabla 4.8. Para esta comparación sólo se consideraron las áreas cuyo índice de susceptibilidad (obtenido por Lara (2007)) sea mayor a 49, esto es, un índice de susceptibilidad medio a alto. Debido a que las áreas de los índices de susceptibilidad para deslizamientos en roca y en suelo son distintas se considera el promedio de estas áreas (tabla 4.8).

**Tabla 4.8:** Comparación en terminas de área y probabilidades de corte de los resultados de la validación en este trabajo con lo obtenido por Lara (2007).

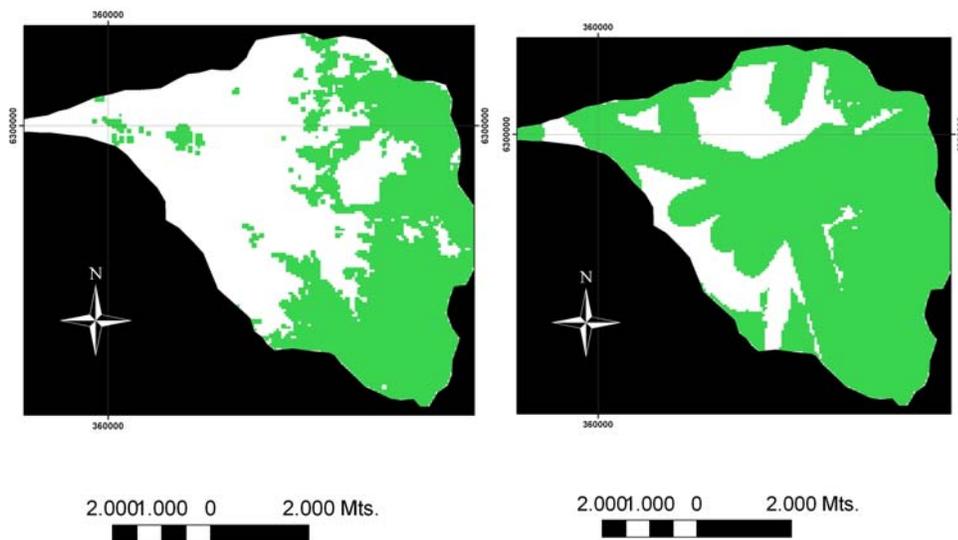
Lara (2007) (área original)	N° de Píxeles	7580
	Área (m2)	18.950.000
Regresión logística	N° de Píxeles	7534
	Área (m2)	18.835.000
	Probabilidad de corte	4%
Análisis discriminante	N° de Píxeles	7600
	Área (m2)	19.000.000
	Probabilidad de corte	19%

En la figura 4.10, se pueden ver los resultados obtenidos para la Quebrada de San Ramón para las probabilidades de corte mostradas en la tabla 4.8, esto es, las probabilidades que satisfacen la igualdad de áreas considerando como punto de comparación la áreas moderada a altamente susceptibles definidas por Lara (2007).



**Figura 4.10:** En blanco se encuentran los sectores con susceptibilidad a la ocurrencia de remociones en masa menor a 0,04 en el caso de la regresión logística (a la izquierda) y menor a un 0,19 para el caso del análisis discriminante (a la derecha). En verde se muestran aquellas áreas cuya probabilidad supera la susceptibilidad definida como corte por cada método.

En la figura 4.10 se puede apreciar que para dichas probabilidades de corte, ambos métodos cubren un área bastante similar, y más aún tienen bastantes áreas en común, lo que cambia radicalmente las observaciones previas acerca del método de análisis discriminante. En la figura 4.11 se comparan los resultados obtenidos en la Quebrada de San Ramón para la misma probabilidad (susceptibilidad) de corte.



**Figura 4.11:** En blanco se encuentran los sectores con susceptibilidad a la ocurrencia de remociones en masa menor a 0,1, y en verde aquellas áreas que superan la susceptibilidad de 0,1. A la izquierda se encuentran los resultados en la Quebrada de San Ramón para el método de regresión logística y a la derecha del análisis discriminante.

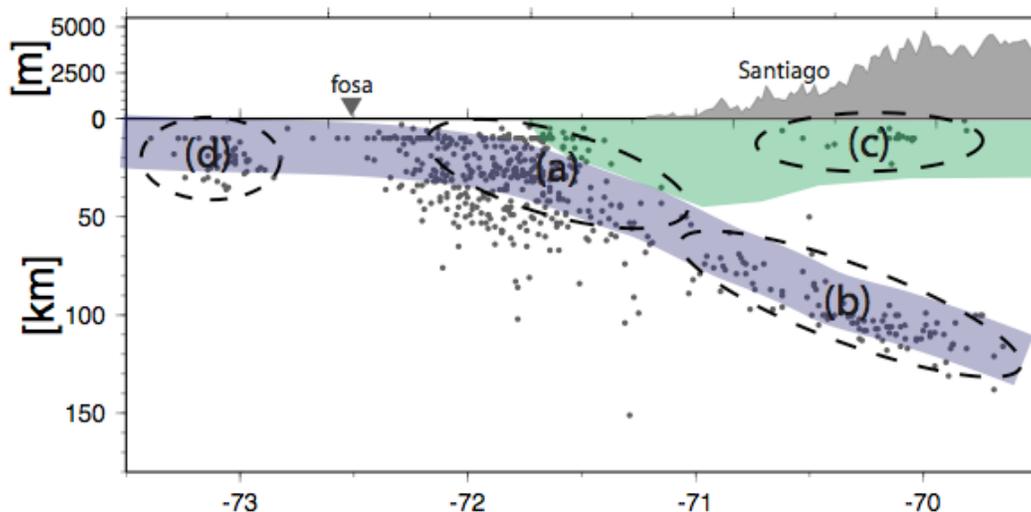
Claramente en la figura 4.11, se ve una reducción del área con susceptibilidades altas o moderadas y en el análisis discriminante una ampliación del área, lo que es obvio dado el aumento en la probabilidad de corte para el método de regresión logística y una disminución en la probabilidad de corte para el método de análisis discriminante. Si bien la regresión logística se corresponde bastante bien con lo obtenido por Lara (2007), el análisis discriminante cubre un área mucho más extensa que lo propuesto por Lara (2007), por lo que, si bien el método de análisis discriminante a primera vista no sirve, se aprecia que en esta ocasión cubre aún mejor todas las áreas cubiertas por el método de regresión logística.

## Capítulo V: Resultados incluyendo la variable sísmica

Si bien la variable sísmica es considerada, en la mayoría de los estudios realizados, como factor desencadenante de remociones en masa, en este caso, la sismicidad es considerada junto con los factores condicionantes, dada la común ocurrencia de eventos sísmicos a lo largo de Chile. Considerar variables desencadenantes como condicionantes ya ha sido utilizado por otros autores, como por ejemplo el caso de García –Rodríguez et al. (2008), quien considero el promedio de precipitaciones anuales, en una determinada zona, como factor condicionante, siendo que en la mayoría de los estudios ha sido considerada como desencadenante, la justificación es precisamente la común ocurrencia de este tipo de eventos en esta zona.

Para este estudio se consideraron tres tipos de fuente para los sismos (figura 5.1):

- Fuentes Corticales.
- Fuentes Intraplaca de profundidad intermedia.
- Fuentes Interplaca.



**Figura 5.1:** Perfil con la sismicidad de la zona, a la altura de 33.5°S. Cada punto representa un sismo, independiente de su magnitud o fecha de ocurrencia. El color azulado representa la posición esquemática de la placa de Nazca y, el verde, la placa Sudamericana. Las líneas punteadas indican la posición de las principales fuentes sismogénicas: (a) interplaca tipo thrust, (b) interplaca de profundidad intermedia, (c) cortical y (d) *outer-rise* (no considerado en este trabajo). Tomado de Leyton et al. (2008).

En particular, se consideraron las siguientes variables, obtenidas probabilísticamente mediante el método de Algermissen & Perkins (1976) para la zona de Chile Central por Leyton et al (2008):

- Intensidad de Mercalli modificada (IMM).
- Aceleración máxima horizontal (PGA).

El método probabilístico propuesto por Algermissen & Perkins (1976) fue utilizado en este trabajo por ser un estándar en Chile, debido a que lo representa bastante bien. Una descripción un poco más detallada acerca de estas variables puede ser encontrada en el anexo dos.

Por otra parte, se consideraron las siguientes fuentes sismogénicas o combinaciones de para cada una de las variables sísmicas:

- Todas las fuentes.
- Cortical.
- Intraplaca de profundidad intermedia más Interplaca.
- Intraplaca de profundidad intermedia.
- Interplaca.

Esta separación se realizó para ver tanto el efecto global de la sismicidad como el efecto de la sismicidad en la placa continental (Sudamericana), en la placa subductante y en la zona de contacto.

Los valores que toman estas variables, tanto el PGA como el IMM, son los esperados para una cierta probabilidad en un determinado período de vida útil; esto se traduce en el período de retorno, entendiendo por tal la probabilidad de exceder un determinado parámetro sísmico en un cierto periodo de tiempo (Leyton et al., 2008). En este caso, el periodo de retorno considerado es de 475 años, esto es, que posee una probabilidad de excedencia de un 10% para un periodo de vida útil de 50 años.

Para estimar las leyes de atenuación de Ms, se utilizó el catálogo NEIC, el cual posee información a partir del año 1973 (y es completa hasta un Ms ~ 7,8), por lo que se complementó con el catálogo de sismicidad histórica SISRA (Askew & Algermissen, 1985). A partir de esta información, se estimó la magnitud máxima esperada por fuente sismogénica (Leyton et al., 2008). Para la estimación de las leyes de atenuación de IMM se utilizaron tanto el catálogo SISRA como informes recientes del Servicio Sismológico Nacional de Chile (Araya, 2007), (Leyton et al., 2008). Para la estimación de las formulas de atenuación de PGA, se utilizaron los datos de acelerogramas de grandes terremotos de Chile (Leyton et al., 2008). Es importante mencionar que es necesario asumir que la sismicidad reportada (en particular aquella de origen cortical) va a seguir ocurriendo donde ha ocurrido hasta el momento (Leyton et al., 2008).

Tanto en el caso de la regresión logística como en el caso del análisis discriminante, se consideraron en el modelo las mismas variables ya señaladas en el capítulo anterior, a las que se agregaron el IMM y el PGA, de acuerdo a las modalidades previamente mencionadas. Una descripción más detallada de las variables sísmicas se puede encontrar en el anexo dos.

En la tabla 5.1 se muestra la simbología escogida para cada una de las variables sísmicas. En la tabla 5.2 se muestran los distintos modelos creados en el presente capítulo de este estudio.

**Tabla 5.1:** Variables sísmicas utilizadas

Variables sísmicas	Código
Intensidad de Mercalli Modificada para todas las fuentes	IMM
Intensidad de Mercalli Modificada para fuentes interplaca	IMM1
Intensidad de Mercalli Modificada para fuentes intraplaca oceánica de profundidad intermedia.	IMM2
Intensidad de Mercalli Modificada para fuentes corticales	IMM3
Aceleración máxima horizontal para todas las fuentes	PGA
Aceleración máxima horizontal para fuentes interplaca	PGA1
Aceleración máxima horizontal para fuentes intraplaca oceánica de profundidad intermedia	PGA2
Aceleración máxima horizontal para fuentes corticales	PGA3

**Tabla 5.2:** Variables sísmicas utilizadas en los casos considerados en este trabajo.

<b>Casos: Variables sísmicas combinadas</b>	<b>Variables consideradas</b>
Todas las fuentes	IMM, PGA
Corticales	IMM3, PGA3
Interplaca e Intraplaca de profundidad intermedia	IMM1, IMM2, PGA1, PGA2
Interplaca	IMM1, PGA1
Intraplaca de profundidad intermedia	IMM2, PGA2
<b>Casos: Variables solas</b>	<b>Variables consideradas</b>
IMM para todas las fuentes	IMM
PGA para todas las fuentes	PGA
IMM para fuentes corticales	IMM3
PGA para fuentes corticales	PGA3
IMM para fuentes intraplaca de profundidad intermedia	IMM2
PGA para fuentes intraplaca de profundidad intermedia	PGA2
IMM para fuentes interplaca	IMM1
PGA para fuentes interplaca	PGA1

## **5.1 Regresión logística**

### **5.1.1.- Estadísticos de colinealidad**

Las tablas A5.1 y A5.2 (anexo 5) muestran los estadísticos de tolerancia y FIV que miden los distintos grados de colinealidad (correlación) existentes entre las variables, para los distintos modelos considerados en este capítulo.

En la tabla A5.1 se puede apreciar que, para los modelos en los que se consideran todas las fuentes y ambas variables sísmicas (IMM y PGA), o ambas variables sísmicas (para cualquier fuente) o más de una fuente sísmica, el porcentaje de cada variable que puede ser explicado por otras es altísimo. De manera análoga, en la tabla A5.2 los valores del FIV para las variables sísmicas que se encuentran en los modelos previamente mencionados, sobrepasan con creces el valor límite para este indicador ( $FIV = 3$ ).

Para el caso en que se consideraron las variables IMM y PGA para las fuentes interplaca e intraplaca de profundidad intermedia (IMM1, IMM2, PGA1 y PGA2), la variable intensidad de Mercalli modificada para fuentes interplaca (IMM1) fue excluida por el modelo por presentar problemas de colinealidad (tablas A5.1 y A5.2). El resto de las variables no fue excluida por el modelo, pese a que también presentaban problemas de colinealidad pero en un grado menor.

Para los modelos con las variables solas (no combinadas y una fuente), el problema de colinealidad previamente observado se redujo de manera importante (tablas A5.1 y A5.3), por lo que no hubo problemas en los cálculos.

Al observar la tabla de correlaciones para la regresión logística (Anexo 7), se puede observar que la correlación existente entre las variables sísmicas con algunas de las topográficas, con las que representan las estructuras en el área de estudio y además entre las mismas variables sísmicas no es menor, lo que explica los importantes problemas de colinealidad. Es necesario considerar (sobre todo en los primeros modelos), que este problema influyó directamente en la confiabilidad de los coeficientes de los modelos (Garson, 2008 (c)).

### **5.1.2.- Tests de ajuste del modelo a los datos**

De acuerdo a la tabla A5.3, se rechazó la hipótesis nula (del test de Hosmer & Lemeshow), por lo que ninguno de los modelos expuestos en este capítulo se ajustó apropiadamente a los datos. En el anexo 5 se pueden encontrar las tablas de contingencia asociadas al test de Hosmer & Lemeshow, en las cuales se puede observar que existe una significativa diferencia entre los valores observados y los predichos.

Sin embargo, para el test Ómnibus (tabla A5.3), el que es bastante menos robusto que el test de Hosmer & Lemeshow, existe un ajuste bastante bueno del modelo a los datos.

Una de las posibles causas de que el test de Hosmer & Lemeshow rechace la suposición de que el modelo se ajusta a los datos (tabla A5.3) sería el tamaño de la muestra la cual es muy grande (más de 300.000 casos). Bajo estas circunstancias, pequeñas diferencias entre los valores observados y predichos pueden provocar que dicho test rechace tal suposición. En cambio, el test Ómnibus sólo busca probar si el modelo, en la intersección es significativamente diferente del modelo con los predictores

(Garson, 2008 (c)) y lo que indica es que al menos uno de los predictores se encuentra significativamente relacionado a la variable dependiente.

La tabla 5.3 muestra los estadísticos de  $R^2$  de Negelkerke y de Cox y Snell. Los valores de estos estadísticos muestran el porcentaje de variabilidad que puede ser explicada por el modelo. A mayor valor de  $R^2$ , mejor es el modelo. Además, es necesario mencionar que su uso no es recomendado para medir la bondad de ajuste del modelo a los datos, sino que sólo su representatividad. El  $R^2$  de Negelkerke es el indicador de interpretación más directa y, considerando el criterio de que el mejor modelo es aquel con menores problemas de colinealidad, se obtiene que el mejor modelo es el que contiene la variable Intensidad de Mercalli Modificada para fuentes corticales (tabla 5.3).

**Tabla 5.3:**  $R^2$  de Cox y Snell y de Negelkerke.

	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
<b>Todas las fuentes</b>	<b>0,08355886</b>	<b>0,36191578</b>
<b>Slab</b>	<b>0,08403233</b>	<b>0,36396654</b>
<b>Cortical</b>	<b>0,08288391</b>	<b>0,3589924</b>
<b>PGA slab</b>	<b>0,08324201</b>	<b>0,36054342</b>
<b>IMM slab</b>	0,076249367	0,330256426
<b>PGA todas las fuentes</b>	0,07366805	0,319076051
<b>IMM todas las fuentes</b>	<b>0,08271838</b>	<b>0,35827545</b>
<b>PGA1</b>	0,074858649	0,324232851
<b>IMM1</b>	0,074427723	0,322366397
<b>PGA2</b>	0,073574954	0,318672826
<b>IMM2</b>	0,074654773	0,323349809
<b>PGA3</b>	0,075655325	0,32768347
<b>IMM3</b>	<b>0,08218195</b>	<b>0,35595202</b>

### 5.1.3.- Tabla de clasificación

En la tabla 5.4 se puede apreciar el porcentaje de casos clasificados correctamente por grupo y total, para una probabilidad de corte de 0,1. Esto significa que todos aquellos casos que tengan probabilidades asociadas iguales o por sobre 0,1 pertenecen al grupo 1 y, análogamente, si la probabilidad asociada es menor que 0,1 pertenece al grupo 0. Es necesario hacer notar que los problemas de correlación pueden hacer que estos resultados se encuentren distorsionados por lo que, si bien aparentemente el modelo que considera todas las fuentes es el que mejor resultados entrega, esto puede no ser cierto. Por lo expuesto, se considera que el mejor de todos los modelos señalados en la tabla 5.4 es el que incorpora la variable Intensidad de Mercalli Modificada para fuentes corticales (IMM3).

**Tabla 5.4:** Resultados de los modelos en términos de porcentaje de casos correctamente clasificados por grupo y total para una probabilidad de corte de 0,1.

Caso	Regresion logistica		
	% de casos correctamente clasificados en grupo 0	% de casos correctamente clasificados en grupo 1	% total de casos correctamente clasificados
<b>Todas las fuentes</b>	94,7996	63,3699	93,8881
<b>Slab</b>	94,6799	62,7262	93,7532
<b>Cortical</b>	94,9047	62,6485	93,9692
<b>PGA todas</b>	94,4691	60,0844	93,4719
<b>IMM todas</b>	94,7837	61,9714	93,8321
<b>PGA slab</b>	94,5764	61,4941	93,617
<b>IMM slab</b>	94,4097	59,8735	93,4081
<b>IMM1</b>	94,2337	59,3185	93,2211
<b>PGA1</b>	94,2062	59,2408	93,1921
<b>IMM2</b>	94,2164	59,3184	93,2044
<b>PGA2</b>	94,3172	59,3296	93,3025
<b>IMM3</b>	94,7068	61,1722	93,7342
<b>PGA3</b>	94,1737	59,3739	93,1644

#### 5.1.4.- Coeficientes

En la tabla 5.5 se encuentran los coeficientes para cada una de las variables utilizadas en los distintos modelos. En la tabla 5.5 se puede apreciar que algunas variables no son importantes o que tienen muy poca importancia para el modelo, algunos ejemplos son:

- Elevación: El valor del coeficiente es de 0.001 en todos los modelos estudiados (tabla 5.5). Si bien, realiza un aporte a los modelos, éste es muy débil. Pese a que en la tabla A5.4 aparece como significativa, el  $\text{Exp}(B)$  es casi 1 (tabla A5.5).
- Aspecto: El valor del coeficiente es 0 en todos los modelos, con excepción del que considera el PGA para todas las fuentes (tabla 5.5). En las otras tablas se puede encontrar una situación parecida a la ya mencionada para la variable elevación.
- Distancia euclidiana a las estructuras: En este caso, el valor del coeficiente para algunos modelos es negativo y pequeño. Esto quiere decir que, a pesar de que la variable sea poco importante, a medida que ésta aumenta de valor es menos importante para el modelo, lo que es correcto.

**Tabla 5.5:** Coeficientes obtenidos para todos los casos mediante el método de regresión logística

	Todas las fuentes	Slab	Cortical	PGA todas	IMM todas	PGA slab	IMM slab	IMM1	PGA1	IMM2	PGA2	IMM3	PGA3
Telev	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001
Tp	0,037	0,037	0,036	0,039	0,035	0,037	0,034	0,034	0,033	0,034	0,034	0,033	0,033
Tasp	0	0	0	0,001	0	0	0	0	0	0	0	0	0
Tper	-0,326	-0,324	-0,331	-0,478	-0,335	-0,352	-0,433	-0,455	-0,449	-0,452	-0,467	-0,346	-0,44
Tpla	0,282	0,298	0,288	0,27	0,292	0,315	0,29	0,311	0,313	0,31	0,307	0,295	0,315
Hv	-0,013	-0,014	-0,014	-0,013	-0,014	-0,018	-0,01	-0,015	-0,015	-0,014	-0,015	-0,014	-0,014
Ga	-2,82	-2,864	-2,793	-2,789	-2,776	-2,852	-2,683	-2,621	-2,618	-2,621	-2,623	-2,757	-2,631
Gf	-2,78	-2,913	-2,617	-3,027	-2,641	-2,981	-2,521	-2,643	-2,619	-2,623	-2,675	-2,569	-2,659
Gi	-2,756	-2,82	-2,651	-1,703	-2,668	-2,793	-1,926	-1,736	-1,782	-1,759	-1,648	-2,664	-1,899
Gc	-2,838	-2,938	-2,741	-2,036	-2,767	-3,028	-1,951	-2,003	-2,034	-2,002	-1,946	-2,685	-2,12
Geests	-0,001	0	0	-0,001	0	0	-0,001	-0,001	-0,001	-0,001	-0,001	0	-0,001
Gdests	1,16	1,202	1,25	1,379	1,309	1,28	1,517	1,648	1,651	1,644	1,634	1,387	1,67
Hedo	0	0	0	-0,001	0	0	0	-0,001	-0,001	-0,001	-0,001	0	-0,001
Hedp	0,001	0,001	0,001	0	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001
Has	0,002	0,002	0,002	0,001	0,002	0,002	0,002	0,002	0,002	0,002	0,001	0,002	0,002
Hds	-5,485	-4,383	-5,358	-6,741	-5,02	-3,403	-5,561	-3,989	-3,778	-4,005	-4,369	-4,371	-3,59
Grem	-1,347	-1,385	-1,389	-1,519	-1,409	-1,411	-1,509	-1,541	-1,536	-1,538	-1,548	-1,433	-1,525
IMM	14,846				15,299								
IMM1							58,985	-4,227					
IMM2		-147,699					-239,6			-16,76			
IMM3			14,998									11,652	
PGA	95,979			117,481									
PGA1		-361,955				-688			-49,88				
PGA2		496,486				434,46					-23,78		
PGA3			-15,924										21,579
Cte	-186,477	990,837	-113,991	-60,276	-125,52	-43,01	1489,9	36,791	21,802	139,61	21,964	-93,795	-3,476

- Distancia euclidiana a la hidrografía ocasional: Similar al de la variable distancia euclidiana a las estructuras.
- Área de subcuenca: Similar al de la variable elevación.

El hecho de que una variable sea positiva indica que, a medida que aumenta el valor de esta variable, aumenta su importancia, y al revés en el caso negativo. Un buen ejemplo es el de la pendiente la que, si bien el valor de su coeficiente no es muy alto, es positivo por lo que se puede esperar que, a mayor pendiente, mayor sea la susceptibilidad del área a la ocurrencia de remociones en masa. Otro ejemplo es el caso del NDVI (Hv), cuyo coeficiente es negativo, lo que indica que la presencia de vegetación es inversamente proporcional a la susceptibilidad a la ocurrencia de remociones en masa, lo que también es correcto.

En el caso de las variables categóricas, tales como la geología, es necesaria una pequeña aclaración: Dado que, en este caso, las variables que representan la geología toman valores 0 ó 1, el programa estadístico SPSS v15.0 las codifica internamente de tal que si es 0 significa que es un punto en donde se puede encontrar esa característica y, si es 1, tal característica no se encuentra en ese punto. Por lo tanto, el hecho de que los coeficientes que representan a la geología sean negativos, no quiere decir que la variable sea poco importante sino que, representa en que grado disminuye la susceptibilidad, por no encontrarse la característica en ese punto. Dada esta situación, en la tabla 5.5 se puede ver que las variables más importantes dentro de la geología son las que representan a la Formación Abanico y a los depósitos coluviales, en la mayoría de los casos, y a la Formación Farellones en otros, lo que es correcto dada la alta presencia de escarpes en la Formación Abanico, en la Formación Farellones y en los depósitos coluviales.

En la tabla A5.4 se aprecian los valores del  $\text{Exp}(B)$ , los que indican, que tan importante puede ser una variable dentro del modelo:

- El aspecto y la elevación permanecen con un valor uno o muy cercano a este, en todos los casos, lo que indica que su aporte al modelo es poco importante.
- Las variables que representan las distancias euclidianas a las hidrografías permanente y ocasional, también tienen un bajo aporte al modelo, por la misma razón anterior.
- Las variables sísmicas resultan ser las más importantes en todos los modelos (en particular el IMM y el PGA para fuentes corticales), dado que el valor de  $\text{exp}(B)$  es siempre mayor que uno.

En términos de significancia de cada variable dentro del modelo, todas ellas resultan importantes, con excepción de algunas, para algunos modelos, como es el caso de la distancia euclidiana a la hidrografía ocasional cuando se considera todas las fuentes. Sin embargo, la variable aspecto prácticamente no es significativo en ninguno de los modelos (tabla A5.5).

Para los modelos obtenidos con el método de regresión logística, las variables más importantes (o más representativas del problema) resultaron ser:

- Curvatura en planta
- Formación Abanico
- Formación Farellones
- Intrusivos
- Depósitos coluviales
- Depósitos de remociones en masa
- Densidad de estructuras

- IMM3

Se incluye la variable IMM3 dado que fue con esta variable con la que se obtuvo el mejor modelo (tabla 5.3).

Lo anterior coincide bastante bien con las variables consideradas como más importantes dentro del modelo por este método sin la variable sísmica (ver capítulo cuatro).

## **5.2 Análisis discriminante**

### 5.2.1.- Estadísticos de colinealidad

Las tablas A5.6 y A5.7 muestran los estadísticos de tolerancia y FIV, que miden los distintos grados de colinealidad existentes entre las variables normalizadas, para los distintos modelos considerados en este trabajo.

En la tabla A5.6, y al igual que en el caso de la regresión logística, se puede apreciar que, para los casos en el que se consideran todas las fuentes y ambas variables (IMM y PGA) o ambas variables (para cualquier fuente) o más de una fuente, el porcentaje de estas variables que puede ser explicado por otras es elevado, lo que implica que existe un alto nivel de colinealidad entre las variables utilizadas. De manera análoga, en la tabla A5.7, los valores de FIV para las variables sísmicas que se encuentran en los casos previamente mencionados sobrepasan con creces el valor límite ( $FIV = 3$ ), aunque, en este caso, los valores de FIV obtenidos son un poco menores que para los modelos obtenidos mediante regresión logística.

Si se consideran las variables IMM y PGA, para las fuentes interplaca e intraplaca de profundidad intermedia, la variable IMM para fuentes interplaca fue excluida por el presentar problemas de colinealidad (tablas A5.6 y A5.7). Sin embargo, el modelo no excluyó a todas las variables que presentaban problemas de colinealidad, mostradas en negritas en tabla A5.6.

Para los modelos con variables solas y una fuente, el problema de colinealidad previamente observado se redujo de manera importante (tablas A5.6 y A5.7), siendo incluso menor al observado en el caso del modelo con regresión logística (tablas A5.1 y A5.2). Por lo anterior, y al igual que en el caso de la regresión logística, en los modelos con variables sísmicas solas y para una fuente, el problema de colinealidad no se consideró importante.

Al observar la tabla de correlaciones entre las variables para el análisis discriminante (Anexo 7), nuevamente se puede ver que se presentó una fuerte correlación entre las variables sísmicas con las que representan las estructuras, así como también entre las variables sísmicas propiamente tales.

### 5.2.2.- Homogeneidad de las matrices de covarianza

El test de M de Box pone a prueba la suposición de que las matrices de covarianza son las mismas entre los grupos (homogeneidad de las matrices de covarianza). La tabla A5.8 muestra que esta suposición es rechazada ya que el p-valor obtenido es 0.

Sin embargo, de acuerdo con el método del logaritmo de los determinantes (tabla A5.9), si bien la variabilidad de cada grupo es grande, la variabilidad intergrupos no lo es tanto, por lo que se consideró que la suposición se cumplió a un nivel aceptable.

Cabe hacer notar que esta es una propiedad que difícilmente se cumple y, es más difícil aún cuando los grupos son de tamaños muy diferentes, y también cuando se incluyen variables como la topografía, ya que muchos de los deslizamientos que se pueden encontrar en el área en estudio están asociados a un cierto rango de elevaciones, de pendientes, y de curvaturas, entre otros, por lo que normalmente la distribución de los valores por grupo no resulta similar.

### 5.2.3.- Igualdad de medias de grupo

El test de igualdad de medias de grupo de Lambda de Wilks mide el potencial de cada variable independiente antes de que el modelo sea creado. Una variable independiente contribuye al modelo en la medida que sea significativa ( $p$ -valor  $< 0.01$ ). Para todos los modelos las variables independiente presentes contribuyen (tabla A5.10), pero de acuerdo al estadístico Lambda de Wilks (tabla A5.10), aquellas variables con menor valor del estadístico discriminan mejor entre grupos.

En la gran mayoría de los casos observados en la tabla A5.10 las variables que mejor discriminan entre grupos son (por orden de importancia):

- Densidad de estructuras.
- Variables sísmicas.
- Distancia euclidiana a las estructuras
- Geología completa
- Depósitos de remociones en masa.
- Pendiente.

Llama fuertemente la atención la ausencia de variables topográficas que sean importantes, con excepción de la variable que representa la pendiente que tampoco es la más importante dentro del grupo de variables escogidas.

### 5.2.4.- Coeficientes estandarizados

Los coeficientes estandarizados (tabla A5.11) muestran, al igual que el estadístico Lambda de Wilks, cuáles variables poseen mejor capacidad discriminante, y, de acuerdo con este criterio son aquellas variables con mayor valor absoluto.

De acuerdo al criterio previamente mencionado, las siguientes variables son las más importantes:

- Sísmicas
- Densidad de estructuras
- Área de subcuenca
- Distancia euclidiana a hidrografía permanente
- Geología completa

Como se puede ver, nuevamente se repiten muchas de las variables previamente señaladas.

### 5.2.5.- Matriz de estructura

La matriz de estructura (tabla A5.12) muestra el grado de correlación existente entre las variables independientes y la función discriminante, y, mientras mayor sea su valor absoluto, mayor es la correlación.

En la tabla A5.13 se muestran las variables más relacionadas con la función discriminante de cada modelo. En la tabla A5.13 se puede ver que nuevamente las variables mostradas a continuación son las más importantes:

- Densidad de estructuras
- Distancia euclidiana a las estructuras
- Variables sísmicas (IMM e IMM3)
- Depósitos de remoción en masa
- Geología completa
- Pendiente.

#### 5.2.6.- Correlación canónica

Muestra cuán buena es la capacidad de la función discriminante para distinguir entre grupos. En la tabla 5.6 se puede ver que ésta capacidad para separar los grupos no es muy buena en general, pues los valores de la correlación canónica no sobrepasan el valor de 0,35, ya que un valor óptimo sería del orden de 0,5 o superior.

Al igual que en el caso de la regresión logística, los modelos que consideran las variables sísmicas combinadas entregan un mejor resultado. Pero, debido a los problemas de colinealidad que existen entre las variables, este modelo no resulta confiable y, por lo tanto, debe ser visto con cuidado.

Si se considera que los mejores modelos son aquellos con menores problemas de colinealidad, entonces sólo deberían considerarse aquellos modelos en los que se utiliza sólo una de las variables sísmicas para una fuente sismogénica. Considerándolo así, el mejor modelo es aquel que incluye la variable Intensidad de Mercalli Modificada, para fuentes corticales (tabla 5.6).

#### 5.2.7.- Tablas de clasificación

Para los modelos creados mediante el método de análisis discriminante, considerar una probabilidad a priori de 0,5 (cada caso tiene igual probabilidad de pertenencia a alguno de los grupos), el resultado que se muestra en la tabla 5.7 indica que, si bien el porcentaje de casos correctamente clasificados en el grupo 0 (o sectores no a débilmente susceptibles), disminuye fuertemente, el porcentaje de casos correctamente clasificados en el grupo 1 (o sectores moderada a altamente susceptibles), aumenta de manera importante.

**Tabla 5.6:** Correlaciones canónicas por modelo.

Modelo	Autovalor	Correlación canónica
Todas las fuentes	0,102	0,304
Slab	0,097	0,298
Cortical	0,093	0,291
IMM todas	0,091	0,289
IMM slab	0,091	0,288
IMM3	<b>0,090</b>	<b>0,287</b>
PGA todas	0,079	0,271
IMM2	0,078	0,269
PGA3	0,078	0,269
PGA slab	0,078	0,269
PGA1	0,077	0,267
IMM1	0,076	0,266
PGA2	0,076	0,266

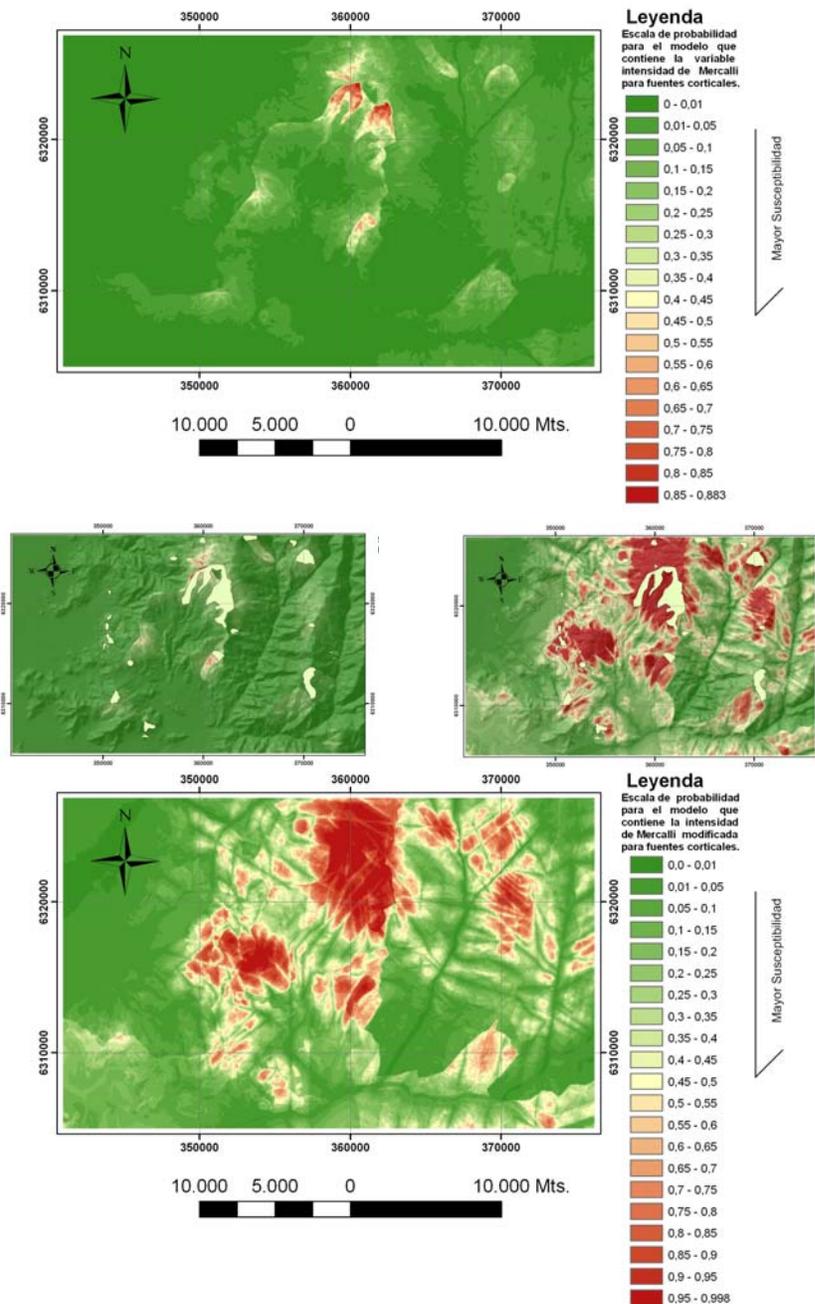
**Tabla 5.7:** Porcentaje de casos correctamente clasificados por grupo y en total para el método análisis discriminante.

Caso	Análisis discriminante		
	% de casos correctamente clasificados en grupo 0	% de casos correctamente clasificados en grupo 1	% total de casos correctamente clasificados
Todas las fuentes	83,1049	74,64757	82,9
Slab	82,4594	73,2712	82,2
Cortical	83,0471	75,5577	82,8
PGA todas	81,1196	77,0784	81
IMM todas	83,1022	71,4145	82,9
PGA slab	81,9644	76,4347	81,8
IMM slab	81,7549	74,7586	81,6
IMM1	81,6709	76,4902	81,5
PGA1	81,7147	76,4347	81,6
IMM2	81,3753	76,1128	81,2
PGA2	81,6213	76,6234	81,5
IMM3	83,2305	71,4144	83
PGA3	81,5599	76,0239	81,4

Es importante mencionar que si bien ambos métodos entregan la susceptibilidad a la ocurrencia de remociones en masa, para poder comparar, en términos de aciertos o equivocaciones, se requiere tener los mismos parámetros y, para ello, es necesaria la creación de una tabla que considere los casos correctamente clasificados, dado un parámetro fijo, que en este caso es una probabilidad de corte determinada.

### **5.3 Resultados gráficos de los métodos utilizados**

Dado que, de acuerdo a ambos métodos estadísticos utilizados la variable más importante es la Intensidad de Mercalli Modificada, para fuentes corticales (IMM3). Los resultados gráficos obtenidos con los modelos que incluyen esta variable se muestran en la figura 5.2 (el resto de las figuras es mostrada en el anexo seis). Llama fuertemente la atención la gran similitud que se observa en todos los resultados gráficos de los modelos vistos en este capítulo (anexo seis).



**Figura 5.2:** Modelos que incluyen variable que representa la Intensidad de Mercalli Modificada para fuentes corticales (IMM3). Arriba, resultado regresión logística. Al medio, resultados regresión logística y análisis discriminante con escarpes superpuestos. Abajo, resultado análisis discriminante.

## **5.4 Comparación de métodos de análisis discriminante y regresión logística para una probabilidad de corte de 0,1**

En la tabla 5.8 se observa una marcada diferencia en la forma en que los métodos de regresión logística y análisis discriminante clasifican dentro de los grupos para igual probabilidad de corte.

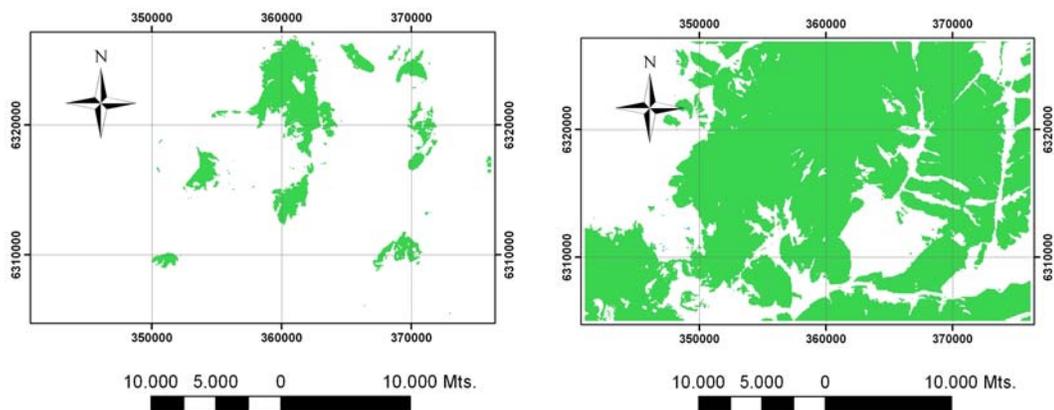
La probabilidad de corte para comparar se escogió en 0,1, dado que recién para esa probabilidad la regresión logística presentaba un nivel de acierto superior al 50% para los casos positivos, lo que es considerado aceptable.

El hecho de que los casos positivos, observados en la tabla 5.8, para los métodos de análisis discriminante y regresión logística, sean mayores que los casos positivos reales, no quiere decir que éstos estén incluidos completamente dentro de los casos positivos y, de hecho no lo están. Es por esto que los porcentajes de clasificación observados en las tablas 5.4 y 5.7 no alcanzan el 100% de casos positivos correctamente clasificados.

**Tabla 5.8:** Comparación de métodos de análisis discriminante y regresión logística, para una probabilidad de corte de 0,1, en términos de número de casos o pixeles para todos los modelos vistos en este capítulo. RL: regresión logística, AD: análisis discriminante.

	Casos reales 0	Casos reales 1	Casos RL 0 (<0,1)	Casos RL 1 (>0,1)	Casos AD 0 (<0,1)	Casos AD 1 (>0,1)
<b>Todas las fuentes</b>	300486	9009	287946	21549	137007	172488
<b>Slab</b>	300486	9009	287736	21759	130030	179465
<b>Cortical</b>	300486	9009	288377	21118	116312	193183
<b>PGA slab</b>	300486	9009	287420	22075	91722	217773
<b>IMM slab</b>	300486	9009	287163	22332	121761	187734
<b>PGA todas las fuentes</b>	300486	9009	287194	22301	100983	208512
<b>IMM todas las fuentes</b>	300486	9009	288078	21417	118099	191396
<b>PGA1</b>	300486	9009	286606	22889	94047	215448
<b>IMM1</b>	300486	9009	286674	22821	93865	215630
<b>PGA2</b>	300486	9009	286894	22601	93918	215577
<b>IMM2</b>	300486	9009	286644	22851	99069	210426
<b>PGA3</b>	300486	9009	286508	22987	96930	212565
<b>IMM3</b>	300486	9009	287902	21593	114934	194561

Al comparar gráficamente los resultados arrojados por los métodos de análisis discriminante y regresión logística, para los modelos que consideran la variable independiente Intensidad de Mercalli Modificada para fuentes corticales (IMM3), se observa nuevamente lo ya visto en el capítulo cuatro (figura 5.3), esto es, que para igual probabilidad de corte (en este caso 0,1), el análisis discriminante posee una mayor cobertura de superficie con respecto al método de regresión logística, pero no se encuentra necesariamente asociado a los escarpes de deslizamiento (variable dependiente en este trabajo). Se puede ver que, en el caso de la regresión logística, la cobertura de superficie aumenta con respecto a la observada en el capítulo cuatro (tablas 5.8 y 4.23 y figuras 5.3 y 4.10).



**Figura 5.3:** En blanco se encuentran los sectores con susceptibilidad a la ocurrencia de remociones en masa menor a 0,1, y en verde aquellas áreas con susceptibilidad mayor o igual a 0,1. A la izquierda se encuentran los resultados obtenidos con el método de regresión logística, y a la derecha, los obtenidos con el método de análisis discriminante para los modelos que incluyen la variable Intensidad de Mercalli Modificada para fuentes corticales (IMM3).

## **5.5 Métodos de regresión logística y análisis discriminante considerando sólo las variables más importantes de acuerdo a cada método.**

En este subcapítulo se consideraron las variables más importantes de acuerdo a cada método, éstas son:

### Análisis Discriminante

- Pendiente.
- Distancia euclidiana a las estructuras.
- Densidad de estructuras.
- Geología.
- Remociones.
- Intensidad de Mercalli modificada para fuentes corticales.

### Regresión logística

- Curvatura en planta
- Densidad de estructuras.
- Geología.
- Remociones.
- Intensidad de Mercalli modificada para fuentes corticales.

### 5.5.1 Análisis Discriminante

#### 5.5.1.1 Correlación entre variables utilizadas

En la tabla A5.14 se puede ver que la correlación entre las variables es lo suficientemente baja como para que no existan problemas de colinealidad.

#### 5.5.1.2 Estadísticos de colinealidad

Efectivamente, tal como fue supuesto previamente, se observa en la tabla A5.15 que los problemas de colinealidad son prácticamente inexistentes dentro de este grupo de variables.

#### 5.5.1.3 Tests de homogeneidad de las matrices de covarianza

- M de Box: al igual que en todos los casos mostrados en este trabajo, la suposición de que las matrices de covarianza poblacionales son iguales es rechazada (Tabla A5.16).
- Logaritmo de los determinantes: de acuerdo a este método existe un parecido bastante aceptable entre las matrices de covarianza poblacionales para este modelo por lo que esta suposición puede asumirse como cierta (Tabla A5.17).

#### 5.5.1.4 Pruebas de igualdad de las medias de los grupos

De acuerdo a la significancia del estadístico de Lambda de Wilks, para cada una de las variables consideradas en este modelo, todas las variables contribuyen al modelo (Sig. < 0,01) (Tabla A5.18)

### 5.5.1.5 Coeficientes estandarizados de las funciones discriminantes canónicas

Los coeficientes estandarizados (Tabla A5.19) miden el grado de importancia de cada una de las variables dentro del modelo y sirven para calcular las puntuaciones discriminantes.

### 5.5.1.6 Matriz de estructura

La matriz de estructura (tabla A5.20) muestra el nivel de correlación que existe entre la variable y la función discriminante. Para medir este nivel de correlación se considera el valor absoluto de la variable.

### 5.5.1.7 Correlación canónica

Este estadístico muestra la capacidad para distinguir los casos de la función discriminante. La correlación canónica (Tabla 5.9) muestra que la capacidad para distinguir de la función discriminante, correspondiente a este modelo, no es muy buena, y es incluso peor que las observadas en los capítulos cuatro y cinco, pero aún así se la considera mas fidedigna dado que en este caso los problemas de colinealidad son prácticamente nulos.

**Tabla 5.9:** Correlación canónica entre el modelo y la función discriminante.

Autovalor	% de varianza	% acumulado	Correlación canónica
0,061	100,0	100,0	0,239

### 5.5.1.8 Resultados de la clasificación

La tabla 5.10 muestra que un 77,8% del total de casos se encuentran correctamente clasificados. Más importante aún es que el 76,8% de los casos positivos se encuentran correctamente clasificados.

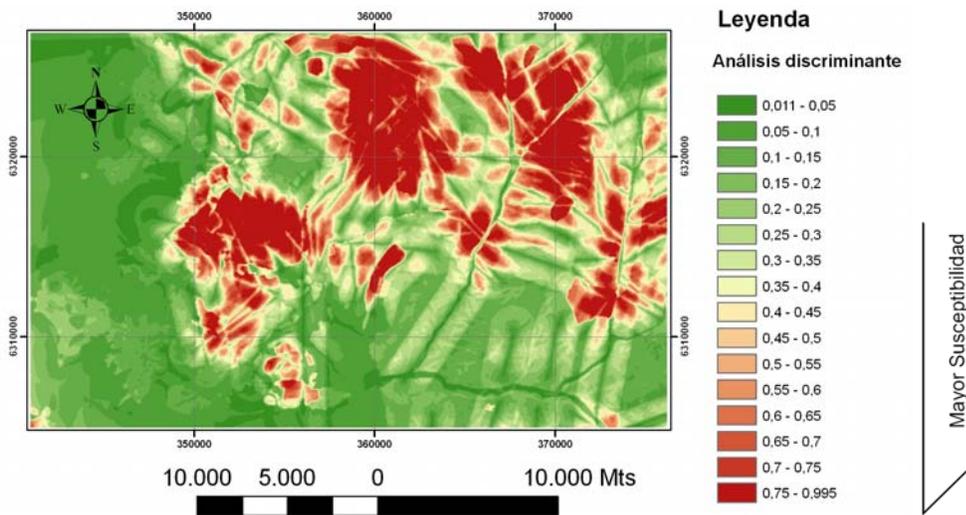
**Tabla 5.10:** Tabla que muestra los resultados de la clasificación para el modelo obtenido mediante el método de análisis discriminante.

Escarpe			Grupo de pertenencia pronosticado		Total
			0	1	
Original	Recuento	0	234799	66832	301631
		1	2088	6921	9009
	%	0	77,8	22,2	100,0
		1	23,2	76,8	100,0

Clasificados correctamente el 77,8% de los casos agrupados originales.

### 5.5.1.9 Resultados gráficos del modelo

En la figura 5.4 se ve una distribución de susceptibilidades bastante similar a la observada en los capítulos cuatro y cinco para los modelos desarrollados sobre la base del método de análisis discriminante.



**Figura 5.4:** Escala de probabilidad que representa la susceptibilidad a la ocurrencia de remociones en masa en el área en estudio, obtenida mediante el modelo basado en el método de análisis discriminante, considerando sólo las variables más importantes.

## 5.5.2 Regresión logística

### 5.5.2.1 Tabla de correlación

En la tabla A5.21 se puede ver que la correlación que existe entre las variables es lo suficientemente baja como para no ser un problema

### 5.5.2.2 Estadísticos de colinealidad

Se puede decir que los problemas de colinealidad entre estas variables son inexistentes y que el porcentaje de cada una de estas variables que puede ser explicado por otras es mínimo (tabla A5.22).

### 5.5.2.3 Tests de ajuste del modelo a los datos

Al igual que en los capítulos cuatro y cinco de este trabajo, el test de Hosmer & Lemeshow rechaza la suposición de que el modelo se ajuste a los datos, mientras que el test ómnibus la acepta (tablas A5.23 y A5.25)

- Prueba de Hosmer y Lemeshow: De acuerdo al test de Hosmer & Lemeshow el modelo no se ajusta a los datos (Sig. < 0.05) (tabla A5.23). Este hecho es corroborado por lo observado en la tabla A5.24 en donde los valores esperados difieren fuertemente de los valores observados en el caso positivo.

- Pruebas ómnibus sobre los coeficientes del modelo: De acuerdo a este test todas las variables son importantes dentro del modelo y este se ajusta lo suficientemente bien a los datos como para ser considerado un buen modelo (tabla A5.25).

### 5.5.2.4 Tabla de clasificación

El porcentaje de casos positivos correctamente clasificados alcanza a un 57% para una probabilidad de corte de 0,1, lo que es bastante bueno si se considera que el porcentaje de casos positivos correctamente clasificados, al considerar todas las variables, es menor para la misma probabilidad de corte (ver capítulos cuatro y cinco) (tabla 5.11).

**Tabla 5.11:** Tabla de clasificación para el modelo obtenido con el método de regresión logística para una probabilidad de corte de 0,1.

Observado		Pronosticado		
		escarpe		Porcentaje correcto
		0	1	0
escarpe	0	284299	17332	94,3
	1	5207	3802	42,2
Porcentaje global		92,7		

El valor de corte es ,100

### 5.5.2.5 Variables en la ecuación

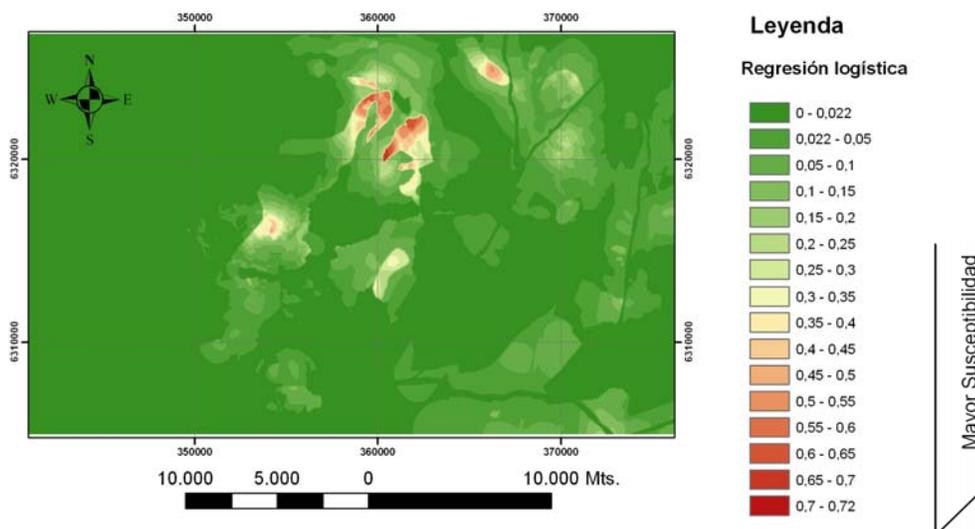
Todas las variables resultan ser importantes dentro del modelo y contribuyen en mayor o menor medida a él. Dentro de todas las variables utilizadas en este modelo, la más importante resulta ser la Intensidad de Mercalli Modificada, para fuentes corticales, seguida de la densidad de estructuras (tabla 5.12).

**Tabla 5.12:** Coeficientes de las variables y su importancia dentro del modelo.

	B	Sig.	Exp(B)
Tpla	,261	,000	1,299
Gdests	1,652	,000	5,216
Ga	-1,455	,000	,233
Gi	-,647	,000	,524
Gc	-,702	,000	,496
Grem	-,849	,000	,428
IMM3	4,706	,000	110,659
Constante	-43,980	,000	,000

### 5.5.2.6 Resultados gráficos del modelo

En la figura 5.5 se ve una distribución más generalizada que la observada en los capítulos cuatro y cinco para los resultados de los modelos basados en el método de regresión logística.



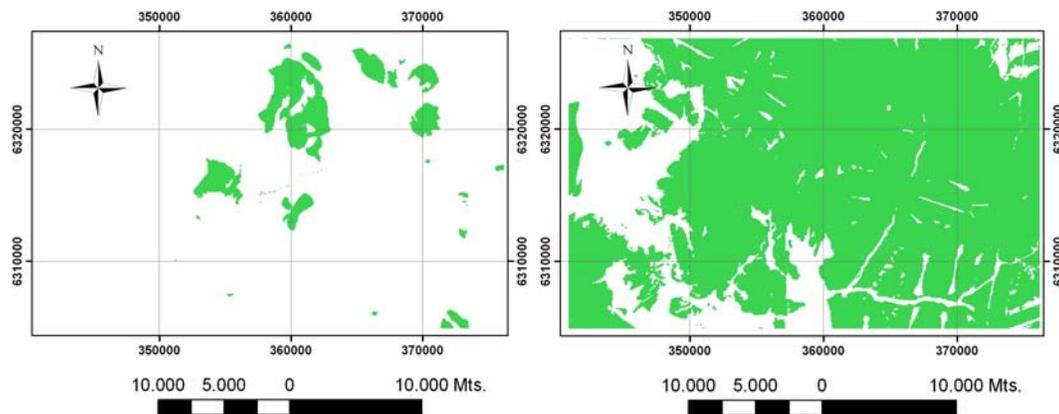
**Figura 5.5:** Escala de probabilidad que representa la susceptibilidad a la ocurrencia de remociones en masa en el área en estudio, obtenidas mediante el método de regresión logística y considerando sólo las variables más importantes.

### 5.5.3 Comparación de resultados obtenidos con modelos creados con las variables más importantes

Con respecto a los capítulos cuatro y cinco, se aprecia que en la regresión logística las susceptibilidades superiores a 0,1 disminuyen, mientras que en el análisis discriminante aumentan (tabla 5.13 y figura 5.6). Ello conlleva a un aumento del área calificada como susceptible por el análisis discriminante y a una reducción de la misma en la regresión logística.

**Tabla 5.13:** Comparación de resultados entre modelos y casos reales para una probabilidad de corte de 0,1. Donde P es la susceptibilidad a la ocurrencia de remociones en masa, RL: regresión logística, AD: análisis discriminante.

	Casos reales 0	Casos reales 1	Casos RL 0 (P<0,1)	Casos RL 1 (P>0,1)	Casos AD 0 (P<0,1)	Casos AD 1 (P>0,1)
<b>Con variables más importantes</b>	302071	9009	291096	18399	74223	235272



**Figura 5.6:** Susceptibilidad a la ocurrencia de remociones en masa. En blanco se encuentran los sectores con susceptibilidad menor a 0,1, y en verde aquellas áreas con susceptibilidad igual o mayor que 0,1. A la izquierda se encuentran los resultados obtenidos con el modelo basado en el método de regresión logística, y a la derecha los resultados obtenidos con el modelo basado en el método de análisis discriminante.

## **Capítulo VI: Discusión**

La susceptibilidad se define como la capacidad o potencialidad de una unidad geológica o geomorfológica de ser afectada o de generar un proceso geológico determinado. Esta definición corresponde a un concepto cualitativo. Sin embargo, la susceptibilidad también puede ser estimada mediante métodos estadísticos, o lo que es lo mismo, estudiarla mediante metodologías cuantitativas, pero para ello, es necesario trabajar con los hechos consumados, que en el caso de remociones en masa viene a ser la presencia de remociones declarada. Esto es debido a que los métodos que se encuentran disponibles para hacer esto, no pueden crear un modelo sólo a partir de las variables independientes, pues de esta manera, los métodos no pueden asignar un peso a cada una de las variables independientes consideradas. En muchos estudios, y este no es la excepción, los hechos consumados que se consideran en el caso de los deslizamientos, son los escarpes de estos. Esta característica es representada mediante la variable dependiente, binaria, “escarpes” (ver anexo dos), donde un valor de uno indica presencia de esta característica y cero indica su ausencia.

A lo largo de este trabajo, se desarrollaron un total de 28 modelos, de los cuales, 14 estuvieron basados en cada uno de los dos métodos utilizados. De cada grupo de estos 14 modelos, 13 consideraban la sismicidad y uno de ellos no la consideraba. De los modelos que consideraban la sismicidad, se realizó uno para cada combinación de variables y fuentes sismogénicas (en total 13 combinaciones).

De los 28 modelos estudiados, los resultados gráficos obtenidos en ellos son bastante satisfactorios, esto a pesar de que los resultados de las pruebas realizadas para cada uno de los modelos considerados indican que éstos no son lo suficientemente buenos y por lo tanto, no representan correctamente el área de estudio (de donde se obtuvieron los datos para crear el modelo). Sin embargo, es necesario considerar que un test funciona bajo suposiciones muy estrictas, que muchas veces son imposibles de cumplir, y que cuando no se cumplen, o no se cumplen de manera estricta, entonces el test es rechazado y la representatividad del modelo de acuerdo a los tests es menor que la esperada, aunque esto en la realidad no sea cierto. Uno de los ejemplos más claros de esta situación fue el que se vio con el test de Hosmer & Lemeshow, con el método de regresión logística, el cual fue rechazado en todos los modelos generados. Una explicación para este problema es entregada por Hosmer et al. (1997), en donde se explica que a medida que aumenta el número de datos el poder de los tests aumenta fuertemente para detectar desviaciones moderadas, por lo que si bien, el problema puede no ser tan importante, el test lo considera como una violación de la suposición.

Para que un modelo sea bueno, se requiere primero que nada que los datos sean representativos del área que se está estudiando. En este trabajo, se tomaron datos cada 50m dado que la base de datos de mayor resolución con la que se contaba tiene una escala de 1:50.000 (base topográfica OTAS), por lo que fue imposible realizar un trabajo con una resolución mayor a la utilizada.

La pregunta que cabe hacerse es acerca de la importancia del nivel de detalle de los datos con los que se trabaja, y para que variables es esto más importante. De acuerdo a diversos autores (Lee et al., 2002; Park et al. et al., 2008), este es un tema muy importante, en particular, si lo que se desea es un modelo que represente de la mejor manera posible la susceptibilidad a la ocurrencia de remociones en masa en un sector determinado. Dado que el área de estudio posee una topografía altamente variable, típico en la Cordillera de los Andes de Chile Central, hubiera sido ideal contar con una base de datos topográfica con resolución de al menos 20m por píxel (equivalente a una escala de 1:20.000) para así poder representar mejor la variabilidad topográfica existente en el área. Sin embargo, lamentablemente no existen bases de datos con este nivel de resolución para el área de estudio por no ser ésta una zona residencial ni de cultivo ni en la que se realice alguna actividad económica de importancia. De haberse contado con una mejor resolución, probablemente habría cambiado el nivel de importancia de algunas de las variables topográficas, como por ejemplo la pendiente (considerada débilmente importante por el método análisis discriminante), y las curvaturas (consideradas como débilmente importantes por el método de regresión logística). También podría haber resultado importante la variable que representa la densidad de vegetación, la que si bien no corresponde a una variable topográfica, esta variable juega un rol importante en la estabilidad de las laderas. Por lo expuesto, las variables indicadas resultaron con muy poca o ninguna influencia, lo que contradice la evidencia física de terreno.

Un tema muy importante a mencionar, es acerca de la transformación que se les aplica a las variables utilizadas con el método de análisis discriminante para normalizarlas, dado que ésta es una de las exigencias de este método para que el resultado a obtener sea confiable. No ocurre igual con el método de regresión logística, ya que para este método no es necesario el cumplimiento de esta condición, razón por la cual esta transformación no se aplicó. El método de transformación aplicado a las variables independientes para normalizarlas, es de tipo gráfica (cuantil a cuantil). Además, se conserva la media y varianza de los datos originales. Es importante mencionar que esta transformación fue aplicada en el programa R 2.6.2, mediante la función: "normalizeBetweenArrays", para la cual primero se creó una serie de datos de distribución normal con las medias y varianzas de cada una de las variables independientes y luego, se le pidió al programa que ambas tuvieran igual distribución. Al conservarse la media y la varianza de cada una de las variables independientes, lo único que cambia es el valor de los datos originales, los cuales son divididos en cuantiles (partes iguales), y se les asigna un valor dependiendo del rango en que se encuentren, esto es, crear una biyección entre los valores que debería tener la variable para que sea normal y su valor real. Es importante también considerar que esta transformación sólo afecta la interpretabilidad directa de los datos.

Para el estudio realizado en el capítulo cinco, sólo con las variables más importantes, en el caso de la regresión logística, entre las curvaturas fue escogida sólo la curvatura en planta por tener una correlación positiva con el modelo, esto es, a mayor valor (positivo) de esta variable mayor susceptibilidad a la ocurrencia de remociones en masa. Esto resulta bastante lógico considerando que un valor positivo de esta variable representa una topografía positiva, y un valor negativo, una negativa, mientras un valor cero representa superficies planas. La curvatura en perfil considera la topografía en sentido perpendicular a la curvatura en planta, esto es, con respecto a la dirección de máxima pendiente. El hecho de que esta variable tenga una correlación negativa con el modelo puede deberse a que es bastante raro ver un escarpe de falla asociado a un deslizamiento que sea convexo.

El aspecto es una variable topográfica, que en muchos estudios parece no tener importancia (Ohlmacher et al., 2003), y este trabajo no fue la excepción. Una explicación

para esto es, primero la orientación norte sur de las quebradas más importantes del área de estudio, lo que reduce drásticamente la orientación de la gran mayoría de las laderas a Este u Oeste, y segundo la ubicación de los escarpes, los que no tenían una orientación preferencial, pues había tantos ubicados en laderas Este como Oeste, lo que no permitía a los métodos asignarle una importancia mayor a esta variable.

Es curioso, aunque si bastante aceptable, que la variable elevación no haya resultado fuertemente importante para ninguno de los dos métodos utilizados en este trabajo, dado que existe una correlación no menor entre esta variable y la ubicación de los escarpes. Pero nuevamente, la gran dispersión de alturas existente con respecto a la ubicación de los escarpes la hacen una variable poco importante para los modelos.

Lo previamente discutido influye directamente sobre la discusión acerca de los criterios de escala. Ya que, probablemente, las variables elevación y aspecto, por los motivos previamente expuestos, no resultarían importantes al mejorar la escala dada la situación topográfica del área de estudio, y la gran variabilidad que presenta la variable elevación frente a la ubicación de los escarpes de deslizamiento. Esto no invalida lo previamente expuesto acerca de los criterios de escala, ya que probablemente, al mejorar la escala hubiera aumentado la importancia de las variables que representan las curvaturas y la pendiente.

Las variables que representan la geología mostraron tener importancia en todos los modelos, lo que indica que esta variable es relevante en la ocurrencia de los deslizamientos en el área de estudio. Esto es en particular significativo para las Formaciones Abanico y Farellones, así como también para los depósitos coluviales. Pero también se observa una importancia no menor, que le asignan los métodos estadísticos multivariados a los depósitos de remoción en masa. Esta importancia puede deberse a diversos factores, entre ellos, los depósitos de remoción en masa se encuentran fuertemente asociados a los escarpes de deslizamiento, ya que existe una zona en la que el depósito de remoción en masa y el escarpe se traslapan; además los escarpes de deslizamiento son considerados en este estudio como evidencias.

Un lineamiento es una característica topográfica o tonal, lineal en la superficie, que podría indicar zonas de debilidad los que pueden ser fallas, fracturamientos, entre otros. Las variables que consideran los lineamientos son variables consideradas como relevantes por ambos métodos estadísticos, pero el método de análisis discriminante considera ambas variables que relacionan a los lineamientos con el área de estudio (Densidad de lineamientos (Gdests) y distancia euclidiana a los lineamientos (Geests)) como importantes, y de hecho resultaron ser las de mayor peso para todos los modelos generados con el método de análisis discriminante junto con las variables sísmicas. El método de regresión logística en cambio, le da mucho más importancia a la variable densidad de lineamientos (Gdests). Efectivamente, al observar la ubicación de los escarpes en el área en estudio y las zonas con mayor densidad de lineamientos, existe una fuerte correlación entre estas variables.

Las variables que representan la hidrografía y la densidad de vegetación (NDVI) del área de estudio en ambos métodos, son consideradas como poco importantes, lo que no significa que lo sean en la realidad. De hecho, al observar la ubicación de los escarpes se puede ver que la gran mayoría de estos está ubicado en zonas con baja densidad de vegetación, lo que muestra que existe una relación entre los deslizamientos y esta variable. Esta relación es diferente a la que se puede observar entre la densidad de vegetación y la altura.

Una pregunta no menos importante, es qué tan relevante es la cantidad de datos para este tipo de problemas, y cómo puede afectar el resultado. Evidentemente que a medida

que aumenta el número de datos, algunos tests son cada vez más inestables y otros más exigentes, por lo que una suposición puede ser rechazada por pequeñas desviaciones que en la práctica no son importantes. Sin embargo, si es un problema la fuerte diferencia que puede existir entre grupos en la variable dependiente, pues es ideal que la proporción entre grupos sea a lo menos 10:90 (Garson, 2008 (c)). A medida que la diferencia entre grupos aumenta, también aumenta con ella la dificultad para discriminar, esto es especialmente cierto en el caso de la regresión logística (Herrera et al., 2007). Es importante mencionar que, la diferencia entre grupos en este trabajo es bastante grande, es de 9009:302071  $\Leftrightarrow$  10:335 (aproximadamente 2,5:90), por lo que el no cumplimiento de la condición citada más arriba los resultados obtenidos se ven afectados.

La capacidad predictiva de estos métodos en un área externa a la de estudio, va a ser función de, las similitudes topográficas, geológicas, hidrográficas, hidrogeológicas, de vegetación, estructurales, entre otras. Evidentemente, un modelo generado con datos obtenidos en la cordillera de Chile Central, no es de ninguna manera, aplicable en la cordillera del Norte de Chile, debido a las diferencias existentes entre los parámetros utilizados para generar el modelo, lo que es una limitación bastante fuerte para este tipo de métodos, pues los resultados no son directamente extrapolables.

Si bien se puede ver que la capacidad predictiva en el área testada (la cual cumple con la condición de similitud) es bastante aceptable en el caso de la regresión logística, por lo que se considera que los resultados obtenidos en este trabajo son buenos, en el caso del análisis discriminante los resultados obtenidos en la Quebrada de San Ramón muestran la excesiva relación que existe entre las variables densidad de estructuras, distancia euclidiana a las estructuras y la función discriminante, lo que hace que el resultado obtenido con este método no sea adecuado para la Quebrada San Ramón. Por lo tanto el método de análisis discriminante no es adecuado debido a la linealidad propia del método, lo que hace que aquellas variables con las que la variable dependiente se encuentre más correlacionada, sean las que tengan más peso dentro del modelo. Justamente, esto es lo que ocurre en este caso, pues los escarpes generalmente se encuentran cerca de los lineamientos y asociados a zonas de alta densidad de lineamientos, lo que las hace estar altamente correlacionadas con la variable dependiente utilizada en este trabajo, teniendo por ello más peso dentro de los modelos obtenidos con el método de análisis discriminante.

Es importante mencionar que si bien, la regresión logística demuestra una mejor relación con lo obtenido por Lara (2007), tampoco le da importancia a otras áreas que si le da el análisis discriminante. Esto simplemente marca la diferencia entre los métodos y las variables a las que le da importancia cada método, así como a la manera de cada uno de éstos para enfrentar el problema, esto es, la regresión logística lo enfrenta desde un punto de vista no lineal y el análisis discriminante desde un punto de vista lineal, lo que explica las fuertes diferencias obtenidas.

Una desventaja de los métodos utilizados en este trabajo, es la incapacidad que éstos poseen para considerar la relación entre los casos (datos), ya que en este tipo de problemas es evidente que existe una relación (no despreciable) entre casos, lo que afecta considerablemente los resultados obtenidos, ya que, una de las suposiciones más importantes de estos métodos, es justamente la independencia entre los casos (datos) con los que se genera el modelo. Es por este motivo, que sería más apropiado utilizar métodos que permitan trabajar con casos no independientes.

Cabe hacer notar que no se consideraron todas las variables que hubiera sido deseable, como por ejemplo, variables que representen la hidrogeología, los muestreos de las capas litológicas (distinto de aspecto), entre otras, ya que, no fue posible por razones técnicas,

de escala y de tiempo. Es dable suponer que la no consideración de todas las variables posibles podría afectar la representatividad de los modelos generados.

Cuán buenos son estos métodos para este tipo de problemas, va a depender principalmente del nivel de exactitud que el investigador desee, pues una de sus características es que si bien, el investigador puede escoger las variables que considere importantes, en definitiva es el método el que le va asignar los pesos. Además, los métodos estadísticos, tampoco consideran el problema desde el punto de vista de su naturaleza, sino que simplemente desde el punto de vista numérico. Por ello, resultan adecuados sólo como primera aproximación al problema que se estudia. Sin embargo, si resultan apropiados para definir cuáles variables son más importantes dentro de un grupo.

Luego, se puede decir que para aplicar un modelo obtenido con este tipo de metodologías en otra zona, se requiere primero que nada, que la nueva área sea similar, luego, al generar el modelo se debe cumplir con los criterios de escala, esto es, la escala de trabajo debe ser tal que permita representar el área que se estudia lo mejor posible, así, al aplicarlo en otra zona, se puede esperar obtener mejores resultados.

La introducción del parámetro sísmico mejora notablemente los resultados obtenidos en el capítulo cuatro, tanto desde el punto de vista estadístico como gráfico (aunque en este último sea levemente notorio), por lo que se puede suponer que estas variables explican mucho mejor lo que ocurre en el área de estudio. Sin embargo, asociados a estas variables sísmicas es necesario considerar otros factores que no han sido mencionados previamente, estos son:

- Efecto de amplificación.
- Mediciones de IMM y PGA con las que se trabajo son para suelo y no para roca, con excepción del caso cortical.
- Efectos de los distintos tipos de fuentes considerados en este trabajo.

La gran mayoría de las ciudades a lo largo de Chile están construidas sobre suelo, y sólo unas pocas sobre roca. Es por este motivo, que las mediciones sísmicas son realizadas en suelo. Considerando este punto, es también necesario tomar en cuenta que el efecto de sitio (amplificación) es mucho más grande en suelo seguido de rocas blandas (rocas sedimentarias principalmente), por lo que el hecho de que se presenten aceleraciones mucho más altas en suelo (zonas planas en el área de estudio) no es contradictorio, especialmente en aquellos casos en los que la onda sísmica ha viajado una cierta distancia, como es el caso de las fuentes interplaca e intraplaca de profundidad intermedia.

Al observar la distribución de las aceleraciones máximas horizontales y de las intensidades máximas de acuerdo al tipo de sismos (ver anexo dos), se puede ver lo siguiente:

- En el caso de los sismos interplaca tanto las altas intensidades como las aceleraciones máximas están concentradas en la zona de la cuenca de Santiago en el área de estudio, y en la zona precordillerana se ven intensidades menores a 7 en la escala de Mercalli y aceleraciones bajas que se encuentran bajo los 0,35g.
- En el caso de los sismos intraplaca de profundidad intermedia, se observa lo mismo que en el caso de los sismos interplaca pero con aceleraciones más altas (menores a 0,6g) e intensidades menores a 7,92 en la escala de Mercalli.
- En el caso de los sismos corticales se observan bajas aceleraciones e intensidades en el sector de la cuenca de Santiago, y en el sector precordillerano

se ven aceleraciones sobre 0,45g e intensidades sobre 8,7 en la escala de Mercalli.

Las rocas que conforman las formaciones Abanico y Farellones son principalmente de origen volcánico (con algunas unidades sedimentarias dentro de la Formación Abanico que no se encuentran dentro del área de estudio), y como consecuencia de ello su dureza es más bien alta, con excepción de algunas zonas de la Formación Abanico que se encuentran bastante deformadas y fracturadas lo que afecta su dureza.

Dado que las formulas de atenuación utilizadas son para mediciones obtenidas en suelo (con excepción del caso cortical) (Leyton et al., 2008), las aceleraciones utilizadas para los casos interplaca e intraplaca de profundidad intermedia funcionan bien en el sector de la cuenca de Santiago y en los valles, pero de no representan bien lo que ocurre en el sector cordillerano del área de estudio.

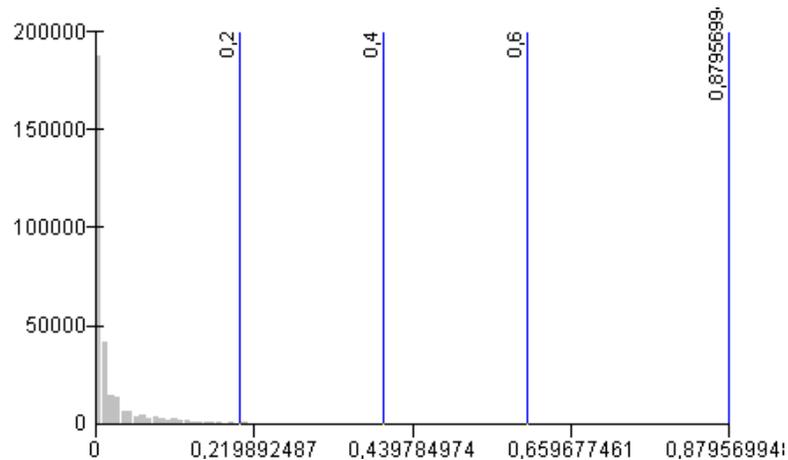
El caso cortical, efectivamente, representa bastante bien el hecho de que en roca dura las aceleraciones son menores. Sin embargo, las aceleraciones más altas en el caso cortical, se observan en el sector cordillerano. Además, asociado a esto, se encuentran los valores de IMM para fuentes corticales, en los que se puede ver que la intensidad de Mercalli modificada es marcadamente mayor en el sector cordillerano con respecto a lo observado con las otras fuentes. Esta situación es explicada por la ubicación de las fuentes consideradas en el caso cortical, y es que, las fallas o las fuentes sismogénicas se encuentran en el mismo lugar, por lo que afectan directamente la zona cordillerana del área de estudio, y la velocidad de las ondas sísmicas es mayor.

Los efectos de la amplificación en el caso de la sismicidad cortical se tornan especialmente importantes en aquellas zonas con depósitos coluviales, de remoción en masa (Bozzano et al., 2007) o simplemente en aquellas zonas con cobertura de suelo (ya sea este residual o transportado).

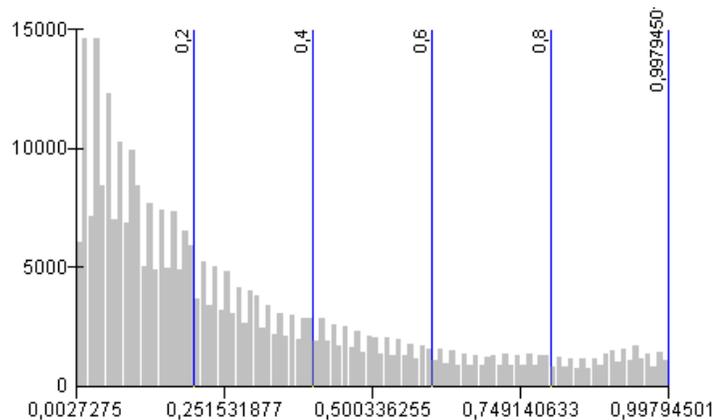
Los resultados obtenidos al aplicar el método de regresión logística con menos variables son muy similares a los obtenidos con todas las variables (luego de la selección de variables), sólo que con un detalle marcadamente menor, por lo que no se puede decir que el resto de las variables que no fueron utilizadas en los modelos con menos variables no sean importantes, sino que más bien su aporte sería menor. En cambio, en el caso del análisis discriminante, en ambos casos se obtuvo prácticamente el mismo resultado independientemente del número de variables.

Las probabilidades obtenidas con el método de regresión logística son mucho más bajas que las obtenidas con el método de análisis discriminante, y la cobertura areal de altas susceptibilidades obtenidas con el método de análisis discriminante son mucho mayores que las obtenidas con el método de regresión logística (ver capítulo cuatro y cinco). Sin embargo, esta no es una situación nueva. Ayalew et al. (2005) definió zonas medianamente susceptibles para probabilidades superiores a 0,14 y la cobertura areal de ésta es de aproximadamente un 5%. Ésta situación puede tener que ver con la distribución areal de los deslizamientos, los cuales son pocos (por la extensión considerada) frente al área de estudio. Por otro lado de acuerdo a Malamud et al. (2003) cuando el área es pequeña estos siguen una distribución exponencial, por lo tanto su comportamiento será el de una exponencial. En este caso la distribución de probabilidades observada para el método de regresión logística es la de una exponencial negativa muy marcada (figura 6.1). Tampoco hay que olvidar que las probabilidades en el método de regresión logística se estiman mediante una exponencial, por lo que se esperaba una distribución de este estilo. El hecho de la distribución de los deslizamientos se puede ver también en el análisis discriminante, el cual es un método lineal por lo que se espera que la distribución de probabilidades sea

lineal y lo que se observa es una curva que describe una forma levemente curva (figura 6.2), es importante acá también recordar que las probabilidades para el método de análisis discriminante son estimadas mediante el método de Bayes, el cual no es lineal.



**Figura 6.1:** Distribución de susceptibilidades a la ocurrencia de remociones en masa para el método de regresión logística.



**Figura 6.2:** Distribución de susceptibilidades a la ocurrencia de remociones en masa para el método de análisis discriminante.

Desde el punto de vista geológico, por el contrario, los resultados obtenidos no son despreciables, pues se observa una correlación más bien alta entre las zonas donde han ocurrido deslizamientos y las variables que representan los lineamientos, la sismicidad cortical, y las unidades geológicas. La importancia de la sismicidad cortical, y de los lineamientos en los modelos permitiría suponer la existencia de una relación de causa efecto, apreciable los deslizamientos y la sismicidad, especialmente la cortical, sin restarle importancia a los factores climáticos.

Si se toman en cuenta los siguientes factores presentes en el área de estudio:

- La Formación Abanico se encuentra fuertemente fracturada y plegada.
- La presencia de las fallas de San Ramón y de Pocuro ubicadas en la vecindad del área.
- La gran cantidad de lineamientos sin orientación preferente.

Dada la marcada orientación Norte – Sur del Río San Francisco, del Estero El Arrayán, y de los lineamientos observados en la hacienda Santa Martina, habría sido lógico esperar que los resultados del estudio realizado arrojaran luz o indicaran acerca de la presencia de una falla en el área en estudio que uniera las dos ya citadas. Sin embargo,

los resultados no son concluyentes ya que los lineamientos en el área de estudio poseen una imagen difusa, por lo que es imposible sin la realización de más estudios inferir si existe una continuación entre las fallas de San Ramón y de Pocuro.

En el contexto del área en estudio, es bastante notoria la importancia que adquieren las variables que representan los lineamientos, la geología y la sismicidad.

Queda aún por discutir cuál de los dos métodos utilizados resulta ser el más adecuado para este tipo de estudios. De acuerdo con los resultados, el método de análisis discriminante sería aparentemente un método poco robusto. Sin embargo, más que un método poco robusto, el análisis discriminante es un método lineal y demasiado restrictivo, esto es, deben cumplirse demasiados supuestos para obtener un resultado confiable, como por ejemplo, las suposiciones de homogeneidad de las matrices de covarianza, la normalidad de la ninguna relación entre las variables independientes, lo que en el caso de las variables que representan la geología, o la hidrografía, o la topografía es prácticamente imposible de cumplir. Por su parte, el método de regresión logística se muestra como un mejor, más robusto, más confiable y de más fácil aplicación que el anterior, ya que es un método no lineal que no impone tantas restricciones, salvo la condición de no existencia de multicolinealidad, y además, facilita enormemente la interpretación y la definición la importancia relativa de las variables dentro de un grupo.

# **Capítulo VII: Conclusiones y recomendaciones**

## **7.1 Conclusiones**

- Los métodos estadísticos multivariados son aplicables para la evaluación de la susceptibilidad a la ocurrencia de deslizamientos pero con limitaciones, impuestas por las exigencias de los métodos estadísticos multivariados.
- Los resultados obtenidos con los métodos estadísticos multivariados de regresión logística y de análisis discriminante son en general buenos, dependiendo de la escala a la que se realiza el estudio. Esto es, para un área como la estudiada, o de mayor tamaño, este tipo de métodos es aplicable a una escala de 1:50.000 o menor.
- El método estadístico multivariado de regresión logística, debido a la no linealidad y menor cantidad de restricciones que posee, es mucho más aplicable a este tipo de problemas que el método de análisis discriminante.
- La capacidad predictiva de los modelos obtenidos con estos métodos, para otros lugares, está condicionada a la similitud en las propiedades y características de las áreas en estudio. Además, tampoco sería posible generar un modelo utilizable a escala de la Región Metropolitana, a partir de una zona de menor extensión, como lo es el área estudiada en este trabajo.
- El área de estudio es un área susceptible a la ocurrencia de remociones en masa, como lo han demostrado otros estudios realizados en el sector oriente de la Cuenca de Santiago, teniendo cuidado de considerar este estudio sólo como una primera aproximación al problema.
- Las variables independientes consideradas como más importantes por cada método son aquellas que representan los lineamientos, la geología, la sismicidad cortical, y algunas variables topográficas, en particular, las curvaturas (por el método de regresión logística) y la pendiente (por el método de análisis discriminante).
- Los modelos indican una correlación importante entre las variables que representan los lineamientos (o estructuras), las variables sísmicas y los escarpes de deslizamiento, resultando ser la más importante la fuente sismogénica cortical, lo que permite plantear la hipótesis de causalidad entre la sismicidad cortical y los deslizamientos.
- La incapacidad de estos métodos para considerar la correlación entre los datos y entre las variables, que es típica de los problemas geológicos, es una limitación si lo que se quiere es un modelo que represente correctamente el problema en estudio.
- La diferencia de tamaño entre los grupos definidos por la variable dependiente es, en este trabajo, un problema, por lo que esta condición es una limitación seria de estos métodos.
- Una resolución (grilla) de 50m para la topografía no resulta adecuada para este tipo de métodos, toda vez que minimiza el efecto de variables que, desde el punto de vista geológico, son importantes.

- Las diferencias metodológicas de estimación de las probabilidades entre ambos métodos, hace que el método de regresión logística entregue como resultado una mayor cantidad de bajas probabilidades asociadas a la susceptibilidad a la ocurrencia de deslizamientos que la que entrega el método de análisis discriminante.

## **7.2 Recomendaciones**

Dado lo observado a lo largo de este trabajo, no se recomienda el uso del método de análisis discriminante para problemas geológicos, dado su carácter lineal, y excesivamente restrictivo, esto es, muchas condiciones a cumplir para que el método entregue un resultado confiable, que en la gran mayoría de los casos son imposibles de cumplir para las variables geológicas.

Dado lo observado a lo largo de este trabajo se recomienda el uso del método de regresión logística para problemas geológicos, dado que, demostró entregar resultados gráficos bastante acordes con los observados tanto en el área de estudio como en otros sectores similares. Sin embargo, la interpretación de sus resultados sólo entregan una primera aproximación al problema, ya que este método, al igual que el de análisis discriminante, no considera la interacción entre los datos.

Ambos métodos utilizados permiten definir cuáles variables independientes son las más importantes dentro de un modelo determinado. Sin embargo, es necesario considerar, que el método de análisis discriminante lo hace desde un punto de vista lineal, por lo que es mucho mejor el resultado entregado por el método de regresión logística.

La escala de trabajo depende de la variabilidad que presenten los datos considerados dentro del estudio. Esto es, mejorar el nivel de detalle sólo es importante en aquellas variables en la que se pierde información que puede ser valiosa para la creación del modelo.

## Referencias

- Aguirre, L., 1960. Geología de los Andes de Chile Central. Boletín N° 9, Instituto de Investigaciones Geológicas, Santiago, 70 p.
- (a) Antinao R., Jose L; Fernandez M., Juan Cristóbal; Naranjo Jose Antonio; Villarroel A., Patricio. 2003. Peligro de remociones en masa e inundaciones de la Cuenca de Santiago, Región Metropolitana. Escala 1:100000. SERNAGEOMIN, 1 Mapa, 1 mapa pleg col, Santiago.
- (b) Antinao R., Jose L; Fernandez M., Juan Cristóbal; Iriarte D., Sergio. 2003. Geología para el Ordenamiento Territorial, Cuenca de Santiago, Región Metropolitana, de Santiago. Escala 1:100000. SERNAGEOMIN, Carta Geológica de Chile, Serie Ambiental, N°5, il; fot.col; 1 Mapa pleg en bolsillo, Santiago.
- Alexander, D. E. A brief Survey of GIS in mass – movements studies, with reflections on theory and methods. *Geomorphology* (2007), doi: 10.1016/j.geomorph.2006.09.022.
- Algermissen, S.T; Perkins, D.M. 1976. A probabilistic estimate of maximum acceleration in rock in the contiguous United States. U.S. Geological Survey Open – File Report 76 – 416.
- Araya, P. 2007. Catalogo actualizado de intensidades sísmicas para Chile. Memoria para optar al título de Ingeniero Civil, Universidad de Chile.
- Ardizzone, F. Cardinali, M. Carrara, A. Guzzetti, F. Reichenbach, P. 2002. Uncertainty and errors in landslide mapping and natural hazards assessment. *Natural Hazards and Herat Systems Sciences*, Vol. 2, 1 – 14.
- Askew, B; Algermissen, S.T. 1985. Proyecto SISRA, programa para la mitigación de los efectos de los terremotos en la región.
- Ayalew, L. Yamagishi, H. 2004. The application of GIS- based logistic regression for landslide susceptibility mapping in the Kakuda – Yahiko Mountains, Central Japan. *Geomorphology* 65, 15 – 31.
- Ayalew, L; Yamagishi, H., 2005. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology*, v. 65, iss. 1-2, p. 15-31.
- Baeza, C., 1994. Evaluación de las condiciones de ruptura y la movilidad de los deslizamientos superficiales mediante el uso de técnicas de análisis multivariante. Tesis doctoral. Departamento de Ingeniería y Cartografía.

- Bozzano, F. Lenti, L. Martino, S. Paciello, A. Scarascia Mugnozza, G. 2007. Possible role of local seismic amplification in reactivation of large landslides: insights from Italian case histories. GNGTS. Sesión 2.1.
- Caballero, F. J. 2008. Modelos de regresión logística incondicional. Sociedad Andaluza de Enfermedades Infecciosas (SAEI). <http://saei.org>
- Campbell, D. 2005. Termocronología del sector comprendido entre los ríos Rocin y aconcagua: implicancias en la evolución meso – Cenozoica de la zona. Memoria para optar al Título de Geólogo. Departamento de Geología, Universidad de Chile
- Can, T. Nefeslioglu, H. Gokceoglu, C. Sonmez, H. Duman, T. 2005. Susceptibility assessments of shallow earthflows triggered by heavy rainfall at three catchments by logistic regression analyses. *Geomorphology* 72, 250 – 271.
- Carrara, A. Multivariate models for landslide hazard evaluation. *Mathematical Geology*, Volumen 15 N°3, Junio de 1983. Doi: 10.1007/BF01031290. P 403 – 426.
- Carrara A., Cardinali M., Detti R., Guzzetti F., Pasqui V. & Reichenbach P. (1990) Geographical Information Systems and Multivariate Models in Landslide Hazard Evaluation. Proceedings VI ICFL - ALPS 90, Milan, 12 September 1990, 17-28.
- Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V., Reichenbach, P., 1991. GIS techniques and statistical models in evaluating landslide hazard. *Earth Surface Processes and Landforms* 16, 427–445.
- Carrara, A. Cardinali, M. Guzzetti, F. Reichenbach, P. 1995. GIS technology in mapping landslide hazard. *Geographical information systems in assessing natural hazards*, pp. 135 – 175. © 1995 Kluwer Academic Publishers. Printed in the Netherlands.
- Carrara. GIS technology and models for assessing landslide hazard and risk. *Geomorphology* (2007) Editorial.
- Charrier, R. Baeza, O. Elgueta, S. Flynn, J.J. Gans, P. Kay, S.M. Muñoz, N. Wyss, A.R. Zurita, E. 2002. Evidence for Cenozoic extensional basin development and tectonic inversion south of the flat – slab segment, Southern Central Andes, Chile (33° - 36° S.L). *Journal of South American Earth sciences*. Vol. 15, Issue 1, pp. 117 – 135. Doi: 10.1016/S0895 – 9811(02)00009-3
- Charrier, R., Pinto, L., y Rodríguez, M.P., 2007. Tectono – Stratigraphic evolution of the Andean origin in Chile, in: *Geology of Chile*, Chapter 3 (Gibbons, W. and Moreno, T., editors), The Geological Society, London, Special Publication, p. 21 – 116.
- Chung, C. F.; Fabbri, A. G and Van Westen, C. J., 1995. Multivariate regression analysis for landslide zonation. *Geographical information Systems in assessing natural hazards*. Kluwer Academic Publishers, 107-133.
- Concha Dimas, A. Campos – Vargas, M. Lopez – Miguel, C. 2007. Comparing heuristic and bivariate GIS Based methods for refining landslide susceptibility maps in Northern Mexico City. 2004. *Environmental and engineering Geosciences*, Vol XIII, N° 4 , pp. 277 - 287.

- Davis, J. Chung, C. J. Ohlmacher, G. 2006. Two models for evaluating landslide hazards. *Computer & Geosciences* 32, 1120 – 1127.
- Drake, E.R.; Curtis, G.; Vergara, M. 1976. Potassium-argon dating of igneous activity in the central Chilean Andes. *Journal of Volcanology and Geothermal Research*, Vol. 1, p. 285-295.
- Eeckhaut, M Van Den. Vanwalleghem, T. Poesen, J. Govers, G. Verstraeten, G. Vandekerckhove, L. 2006. Prediction of landslide susceptibility using rare events logistic regression. A case – study in the Flemish Ardennes (Belgium). *Geomorphology* 76, 392 – 410.
- Fock, A. 2005. Cronología y tectónica de la exhumación en el Neógeno de los Andes de Chile Central entre los 33° y los 34°S. Memoria para optar al Título de Geólogo. Departamento de Geología. Universidad de Chile.
- Fuentes, F., Vergara, M., Aguirre, L., Féraud, G. 2002. Relaciones de contacto de unidades volcánicas terciarias de los Andes de Chile central (33° S): una reinterpretación sobre la base de dataciones 40Ar/39Ar. *Revista Geológica de Chile*, Vol. 29, Nº 2, p. 207- 225.
- Fuentes, F. 2004. Petrología y metamorfismo de muy bajo grado de unidades volcánicas Oligoceno – Miocenas en la ladera occidental de los Andes de Chile Central (33°S). Ph.D Thesis. (unpublished). Universidad de Chile, Departamento de Geología: 407p.
- Gana, P.; Wall, R.; Gutiérrez, A.; Yáñez, G. 1994. Geología de la cordillera de la costa, Regiones de Valparaíso y Metropolitana, al oeste de los 71°00`W. In *Congreso Geológico Chileno*, Nº. 7, Actas, Vol. 1, pp. 43 - 46. Concepción.
- Gana, P. and Wall, R., 1997. Evidencias geocronológicas 40Ar/39Ar y K-Ar de un hiatus Cretácico Superior-Eoceno en Chile Central (33°-33° 30' S). *Revista Geológica de Chile*, Vol. 24, Nº 2, p. 145-163.
- García –Rodríguez, M.J. Malpica, J.A. Benito, B. Díaz, M. 2008. Susceptibility assessment of earthquake triggered landslides in El Salvador using logistic regression. *Geomorphology* 95, pp. 172 – 191.
- (a) Garson, G. Discriminant function analysis. North Carolina state University. 2008. <<http://www2.chass.ncsu.edu>>.
- (b) Garson, G. Testing of assumptions. North Carolina state University. 2008. <http://www2.chass.ncsu.edu>.
- (c) Garson, G. Logistic regression. North Carolina state University. 2008. <http://www2.chass.ncsu.edu>.
- Godoy, E. Yáñez, G. Vera, E. 1999. Inversion of an Oligocene volcanic tectonic basin and uplifting of its superimposed Miocene magmatic arc in the Chilean central Andes: first seismic and gravity evidences. *Tectonophysics*, 306: 217 – 236.
- Gorsevski, P.V., Gessler, P., Foltz, R.B., 2000. Spatial prediction of landslide hazard using discriminant analysis and GIS. *GIS in the Rockies 2000 Conferences*, Denver, Colorado.

- Greco, R.; Sorriso-Valvo, M.; Catalano, E., 2006. Logistic regression analysis in the evaluation of mass movements susceptibility: The Aspromonte case study, Calabria, Italy.
- Guzzetti, F. Reichenbach, F.A. Cardinali, M. Galli, M. 2006. Estimating the quality of landslide susceptibility models. *Geomorphology*, Vol. 81, pp. 166 – 184.
- Hair, J. F.; Anderson, R. E.; Tatham, R. L.; Black, W. C. Análisis multivariante, 5a ed. Prentice Hall Iberia, Madrid, 1999.
- Hauser, A., 1997. Los aluviones del 18 de junio de 1991 en Antofagasta: Un análisis crítico a cinco años del desastre. *Boletín N° 49*. Servicio Nacional de Geología y Minería.
- Herrera, A.N. Gomez, J. 2007. Influence of equal or unequal comparison group simple sizes on the detection of differential item functioning using the Mantel – Haenszel and logistic regression techniques. *Qual Quant*, DOI: 10.1007/s11135-006-9065-z.
- Hosmer, D.W. Hosmer, T. Cessie. S. Le. Lemeshow, S. 1997. A comparison of Goodness of fit tests for the logistic regression models. *Statistics in Medicine*, Vol 16, pp. 965 – 980.
- Hosmer, D.W. & Lemeshow, S. *Applied logistic regression*. New York: John Wiley & Sons, INC, 2000.
- Huberty, C.J. 1989. Problems with stepwise methods: Better alternatives. In B. Thompson (Ed.). *Advances in social science methodology* (Vol. 1). Pp. 43 - 70. Greenwich, CY: JAI Press.
- Innes, P. 2008. *Testing & Fixing for Normality*. UQ Business School. [www.business.uq.edu.au](http://www.business.uq.edu.au)
- Klohn, C., 1960. *Geología de la Cordillera de los Andes de Chile Central, Provincia de Santiago, Colchagua y Curicó*. Instituto de Investigaciones Geológicas, Santiago, *Boletín N° 8*, 95 p.
- Lacourly, N. 2002. *Apuntes de estadística*. Universidad de Chile. Departamento de ingeniería civil matemática.
- Lara, M. 2007. *Metodología para la evaluación y zonificación de peligro de remociones en masa con aplicación en Quebrada San Ramón, Santiago Oriente, Región Metropolitana*. Tesis para optar al grado de Magíster en Ciencias mención Geología. Memoria para optar al título de geólogo. Universidad de Chile.
- Lee S., 2004. Application of Likelihood Ratio and Logistic Regression Models to Landslide Susceptibility Mapping Using GIS. *Environmental Management*. 24, 2, 223 – 232.
- Lee, S. Chwae, U. Min, K. 2002. Landslide susceptibility mapping by correlation between topography and geological structure: the Janghung area, Korea. *Geomorphology*, Vol. 46, Issues 3 -4, pp. 149 – 162.

- Lee, S. Pradhan, B. 2007. Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides* Vol 4, 33 – 41. DOI: 10.1007/s10346-006-0047-y.
- Levene, H. 1960. Robust test for equality of variances. *Contributions to probability and statistics: Essays in honor of Harold Hotelling*, I. Olkin et al., Stanford University Press: 278 – 292.
- Levi, P. Aguirre, L. Nyström, J.O. Padilla, H. Vergara, M. 1989. Low – grade regional metamorphism in the Mesozoic - Cenozoic volcanic sequences of the Central Andes. *Journal of metamorphic geology*, 7: 487 – 495.
- Leyton, F; Ruiz, S; Sepúlveda, S. 2008. Reevaluación del peligro sísmico en Chile Central. *Revista Geológica de Chile*, enviado 23 de Mayo 2008
- Lilliefors. H. W. 1967. On the Kolmogorov Smirnov test for normality with mean and variance unknown. *Journal Amer. Stat. Assoc.*, 62, 399 – 402.
- Mahía, R. 2002. Introducción al análisis discriminante. Universidad Autónoma de Madrid (UAM), Departamento de economía aplicada. [www.uam.es](http://www.uam.es).
- Nandi, A., Shakoor, A. Preparation of a landslide susceptibility map of Summit County, Ohio, USA, using numerical models. *Proceedings, 10th Congress of the International Association of Engineering Geology and the Environment*, Nottingham, UK, paper No. 660.
- Naranjo, J.A. & Varela, J., 1996. Flujos de detritos y barro que afectaron el sector oriente de Santiago el 3 de mayo de 1993. *Servicio Nacional de Geología y Minería, Boletín No.47*.
- NASA World Wind, 2008. Base mundial online de registros de fotografías aéreas.
- NEIC, National Earthquake Information Center, [http:// earthquake.usgs.gov](http://earthquake.usgs.gov)
- Ng, K. 2006. Landslides locations and drainage network development: A case study of Hong Kong. *Geomorphology*, Vol 76, pp. 229 – 239.
- Nyström, J.O. Vergara, M. Morata, D. Levi, B. 2003. Tertiary volcanism during extension in the Andean foothills of central Chile (33°15' - 33°45' S). *Geological Society of America, Bulletin* 115: 1523 – 1537.
- Ohlmacher, G.C., Davis. J.C. 2003. Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. *Engineering Geology*, Vol. 69, pp. 331 – 343.
- Padilla, C. 2006. Análisis de factores meteorológicos desencadenantes de remociones en masa en el sector oriente de Santiago. Memoria para optara al Título de Geólogo. Universidad de Chile.
- Park, N.W., Chi, K.H. 2008. Quantitative assessment of landslide susceptibility using high resolution remote sensing data and a generalized additive model. *International journal of remote sensing*, Vol. 24, N° 1-2, pp. 247-264.

- Rauld, R., 2002. Análisis morfoestructural del frente cordillerano de Santiago Oriente, entre el río Mapocho y la quebrada de Macul. Memoria para optar al Título de Geólogo. Departamento de Geología, Universidad de Chile.
- Rivano, S., Godoy, E., Vergara, M., Villarroel, R., 1990. Redefinición de la Formación Farellones en la Cordillera de Los Andes de Chile central (32° - 34° S). Revista geológica de Chile. V. 17, pp. 205 – 214.
- Rodríguez, M. 2005. Caracterización de la respuesta sísmica de los suelos. Aplicación a la ciudad de Barcelona. Universitat Politècnica de Catalunya. Minor Thesis. <https://upcommons.upc.edu>.
- Santacana, N., Baeza, B., Corominas, J., De Paz, A., Marturiá, J., 2003. A GIS-based statistical analysis for shallow landslide susceptibility mapping in La Pobla de Lillet area (Eastern Pyrenees, Spain). Natural Hazards 30, 281 – 295.
- Selles, D. 1999. La Formación abanico en el cuadrángulo Santiago (33°15' - 33°30' S; 70°30' - 70°45' O), Chile Central. Memoria de título (Inédito). Universidad de Chile.
- Sesser, M.A. 2006. Manual de Fotointerpretación. [www.fing.edu.uy](http://www.fing.edu.uy).
- Sepulveda, S. 1998. Metodología para evaluar el peligro de flujo de detritos en ambientes montañosos: aplicación en la Quebrada Lo Cañas, Región Metropolitana. Memoria para optar al Título de Geólogo. Universidad de Chile.
- Sepúlveda, S.A., Rebolledo, S., Lara, M., Padilla, C., 2006. Landslide hazards in Santiago, Chile: an overview. 10º Congreso International association of Engineering Geology and the Environment, Nottingham, Reino Unido. Paper No. 105 (CD-Rom).
- SPSS v15.0.1 Para Windows. 2006.
- Statsoft. 2008. Discriminant function analysis. StatSoft.Inc. [www.statsoft.com](http://www.statsoft.com)
- Thiele, R. 1980. Geología de la Hoja Santiago, Región Metropolitana. Carta Geológica de Chile N° 39, Instituto de investigaciones geológicas. 51p, 1 Mapa pleg.
- Stern, C.; Amini, H.; Charrier, R.; Godoy, E.; Hervé, F.; Varela, J. 1984. Petrochemistry and age of rhyolitic pyroclastic flows which occur along the drainage valleys of the Río Maipo and Río Cachapoal (Chile) and the Río Yaucha and Río Papagayos (Argentina). Revista Geológica de Chile, No. 23, p. 39-52.
- UITS. 2008. Testing Normality. UITS centerfor Statistical and Mathematical computing. [www.indiana.edu/~statmath](http://www.indiana.edu/~statmath).
- Vahdani, S., Wikstrom, S. 2002. Response of the Tarzana strong motion site during the 1994 Northridge earthquake. Soil Dynamics and Earthquake Engineering 22, 837 – 848.
- Vergara, M.; López, L. 1980. Geología y petrología de los intrusivos subvolcanicos de la precordillera andina entre Santiago y Colina. Universidad de Chile, Departamento de Geología, Comunicaciones, N°. 29, pp. 1 - 21, Santiago.

- Wall, R.; Gana, P.; Gutiérrez, A. 1996. Mapa Geológico del área de San Antonio-Melipilla, regiones de Valparaíso, Metropolitana y del Libertador General Bernardo O'Higgins. Servicio Nacional de Geología y Minería, Mapas Geológicos, No. 2, 1 mapa (1:100.000).
- Wall, R.; Sellés, D.; Gana, P. 1999. Mapa geológico del área de Tiltil-Santiago, Región Metropolitana. Servicio Nacional de Geología y Minería, Mapas Geológicos, No. 11, 1 mapa escala 1:100.000. Santiago.
- Wikipedia. 2008. [www.wikipedia.com](http://www.wikipedia.com)
- Yesilnacar, E. Topal, T. 2005. Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). *Engineering Geology*, Vol. 79, pp. 251 – 266.

# **ANEXOS**

**Anexo 1:**  
**Pruebas de normalidad**

**En CD fuera de texto**

**Anexo 2:**  
**Descripción de variables utilizadas**  
**En CD fuera de texto**

**Anexo 3:**  
**Teoría métodos regresión logística y  
análisis discriminante**

**En CD fuera de texto**

**Anexo 4:**  
**Tablas capitulo 4**

**En CD fuera de texto**

**Anexo 5:**  
**Tablas capitulo 5**

**En CD fuera de texto**

**Anexo 6:**  
**Resultados gráficos modelos capítulo 5**

**En CD fuera de texto**

**Anexo 7:**  
**Tablas de correlación entre variables  
independientes para regresión logística  
y análisis discriminante**

**En CD fuera de texto**

**Anexo 8:**  
**Comprobación de relaciones lineales  
entre el logit y las variables  
independientes para el capítulo 4**

**En CD fuera de texto**