



UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN GEOGRAFÍA

APLICACIÓN DEL ANÁLISIS ESPACIAL DE VECINDAD CONTINUA AL
MODELADO DE PROBABILIDAD A DESLIZAMIENTOS EN TERRENOS
VOLCÁNICOS UTILIZANDO REGRESIÓN LOGÍSTICA MÚLTIPLE

TESIS
QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN GEOGRAFÍA

PRESENTA:
RUTILIO CASTRO MIGUEL

DIRECTOR DE TESIS
DR. GABRIEL LEGORRETA PAULÍN
INSTITUTO DE GEOGRAFÍA DE LA UNAM.

CIUDAD UNIVERSITARIA, CD.MX. AGOSTO DE 2020



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Comité tutor: Dr. Gabriel Legorreta Paulín. - Instituto de Geografía de la UNAM.

Dr. José Fernando Aceves Quesada. - Posgrado en Geografía de la UNAM.

Dr. Roberto Bonifaz Alfonzo. - Instituto de Geofísica de la UNAM.

Sínodos: Dra. Rocío Marisol Alanís Anaya. - Facultad de Filosofía y Letras de la UNAM.

Dra. Lilia Arana Salinas. - Colegio de Humanidades y Ciencias Sociales de la UACM.

Dr. Miguel Ángel Castillo Santiago. - El Colegio de la Frontera Sur.

Agradecimientos.

A la Universidad Nacional Autónoma de México por brindarme el espacio y las herramientas para poder desarrollarme personal y académicamente durante todos estos años.

Al Instituto de Geografía de la UNAM, en donde he tenido la fortuna de iniciar y continuar mi vida profesional y académica.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT), a través del Programa Nacional de Posgrados de Calidad por el apoyo económico para poder realizar los estudios de doctorado.

Al Dr. Gabriel Legorreta por sus consejos para continuar con los estudios de doctorado y por su invaluable asesoramiento para poder plantear y desarrollar este trabajo, así como el apoyo y dirección permanente durante estos años. Del mismo modo agradezco los constantes e importantes aportes a este trabajo por parte del Dr. Roberto Bonifaz y José Fernando Aceves. A todos ellos además de su asesoría profesional agradezco la amistad que me han brindado durante varios años, ha sido un gusto enorme trabajar con ustedes.

A los miembros del sínodo la Dra. Rocío Marisol Alanís, la Dra. Lilia Arana y el Dr. Miguel Ángel Castillo por su interés y disponibilidad para la revisión de la presente tesis, con sus aportes se le dio un mejor sentido y calidad a la tesis.

A Tania Fernández por compartir durante estos años la vida. Gracias a tu compañía ha sido fascinante este desarrollo personal y profesional.

A mis padres: Rutilio y Teresa, y a mis hermanos: Armando, Araceli, Cuauhtémoc, José Luis, Silvia, Graciela, Guadalupe, Roberto y Diego. Por su apoyo constante durante todo este tiempo y por todo lo que he aprendido de cada uno de ustedes.

ÍNDICE	Página
RESUMEN	1
I. INTRODUCCIÓN	2
II. PLANTEAMIENTO DEL PROBLEMA	5
III. OBJETIVOS	6
III.I. Objetivo general	6
III.II. Objetivos específicos	6
IV. HIPÓTESIS	6
CAPÍTULO 1. ANTECEDENTES Y MARCO TEÓRICO	7
1.1. La Geomática en los estudios geográficos	7
1.2. Modelado espacial de probabilidad a deslizamientos	9
1.3. Procesos gravitacionales y sus principales clasificaciones	15
1.4. Análisis Espacial de Vecindad Continua	18
1.5. Muestreo	22
1.5.1. Población y muestra	22
1.5.2. Muestreo probabilístico	23
1.5.3. Tamaño de la muestra	24
1.5.4. División de la muestra	25
1.6. Modelo estadístico de Regresión Logística Múltiple	26
1.6.1. Supuestos de la Regresión Logística	28
1.6.2. Valoración del ajuste del modelo	30
1.7. Evaluación del modelo de Regresión Logística	31
1.7.1. Curva Característica de Operación del Receptor	31
1.7.2. Matriz de clasificación	33
CAPÍTULO 2. ÁREA DE ESTUDIO.	36
CAPÍTULO 3. METODOLOGÍA.	43
3.1. Delimitación de la zona de estudio	43

	Página
3.2. Elaboración del inventario de Procesos Gravitacionales	44
3.3. Descripción y procesamiento de las variables	47
3.3.1. Variable dependiente	47
3.3.2. Variables explicativas	48
3.3.3. Estandarización de las variables espaciales	62
3.3.4. Tamaño de la muestra, muestra ampliada y tipo de muestreo	63
3.3.5. Evaluación de la colinealidad	65
3.3.6. Aplicación del Análisis Espacial de Vecindad Continua	66
3.3.6.1. A las variables cualitativas	66
3.3.6.2. A las variables cuantitativas	71
3.4. Modelos espaciales de probabilidad a deslizamientos	73
3.5. Evaluación de los modelos de probabilidad de AEVC e <i>in situ</i>	76
CAPÍTULO 4. INVENTARIO DE DESLIZAMIENTOS Y MODELOS ESPACIALES DE PROBABILIDAD A DESLIZAMIENTOS DE LA CUENCA LA CIÉNEGA	78
4.1. Inventario de deslizamientos	78
4.2. Modelos de probabilidad a deslizamientos	80
4.2.1. Colinealidad	80
4.2.2. Modelos estadísticos con base al AEVC y con información <i>in situ</i>	81
4.2.3. Modelados espaciales de probabilidad a deslizamientos con base al AEVC e <i>in situ</i>	88
4.2.4. Evaluación de los modelos espaciales	91
CAPÍTULO 5. DISCUSIÓN Y CONCLUSIONES.	93
5.1. Discusión	93
5.2. Conclusiones	106
Referencias	109

ÍNDICE DE FIGURAS	Página
Figura 1. Áreas de investigación, disciplinas y técnicas que integran la geomática.	8
Figura 2. Análisis espacial por zonas, por lo general utilizado en los modelos heurísticos.	13
Figura 3. Análisis espacial puntual, es muy utilizado en modelos estadísticos.	13
Figura 4. Toma de información in situ.	19
Figura 5. Configuración del análisis de vecindad.	20
Figura 6. Integración de la información circundante a los sitios de muestreo para el modelado con RL.	21
Figura 7. Descripción del AEVC, en la figura se representan las áreas de 0 a 10 pixeles de diámetro.	22
Figura 8. Grafica de una curva COR.	33
Figura 9. Calculo de las exactitudes y eficiencia del modelo para la clase de cabeceras (1).	34
Figura 10.- Localización del volcán Nevado de Toluca dentro del CVTM.	36
Figura 11.- Ubicación de la cuenca La Ciénega, en el estado de México.	38
Figura 12. Principales poblaciones en la cuenca La Ciénega.	39
Figura 13. a) valores altitudinales en la cuenca y b) perfil altitudinal este – oeste de la zona de estudio.	39
Figura 14. Mapa de pendientes del terreno en la zona de estudio.	40
Figura 15. Mapa geomorfológico de la cuenca La Ciénega.	41
Figura 16. Mapa de la ubicación del graben y litología de la cuenca La Ciénega.	42
Figura 17. Esquema de la generación automatizada del límite de la cuenca La Ciénega a partir del CEM v3.0 (INEGI, 2013a).	44
Figura 18. Límite final de la cuenca La Ciénega.	44
Figura 19. Distribución de los PG en la cuenca La Ciénega con base en la integración de 2 inventarios.	45
Figura 20. Forma en que fueron tomadas las medidas de la cabecera de los deslizamientos durante los recorridos en 2015 y 2016.	46

Figura 21. Digitalización del polígono de las cabeceras a partir de la ubicación y las dimensiones de los deslizamientos tomadas en campo.	46
Figura 22. Mapa de las áreas estables y de deslizamiento.	48
Figura 23. Mapa de uso de suelo y vegetación 2015.	49
Figura 24. Mapa de cambios en la cobertura de uso de suelo y vegetación durante el periodo 1983-2015.	50
Figura 25. Mapa geomorfológico de la zona de estudio.	52
Figura 26. Mapa de la litología de la zona de estudio.	53
Figura 27. Mapa edafológico de la zona de estudio.	54
Figura 28. Mapa de la textura de los suelos en la zona de estudio.	55
Figura 29. Mapa de altitudes en la cuenca La Ciénega..	56
Figura 30. Mapa de la distribución de la pendiente del terreno en la cuenca La Ciénega.	56
Figura 31. Mapa de la dirección de la pendiente del terreno en la cuenca La Ciénega.	57
Figura 32. Mapa de la curvatura del terreno en la cuenca La Ciénega.	58
Figura 33. Mapa de la rugosidad del terreno en la cuenca La Ciénega.	59
Figura 34. Mapa de la longitud de la pendiente en la cuenca La Ciénega.	59
Figura 35. Mapa de la distancia a fallas y fracturas en la cuenca La Ciénega.	60
Figura 36. Mapa de la distancia a los ríos en la cuenca La Ciénega.	61
Figura 37. Mapa de la distancia a los caminos en la cuenca La Ciénega.	61
Figura 38. Mapa del índice de Vegetación de Diferencia Normalizada de la cuenca La Ciénega.	62
Figura 39. Generación de información raster con una correspondencia a nivel de local.	63
Figura 40. Estratificación de la variable dependiente: zonas estables (0) y zonas de deslizamiento (1).	65
Figura 41. Toma de datos de las 10 variables cuantitativa para realizar el análisis de colinealidad.	66

Figura 42. División del tema cualitativo de Litología en mapas individuales, un mapa para cada clase que lo conforma.	67
Figura 43. Adquisición de datos para las áreas de vecindad correspondientes a 1, 2, 3, 4, 5, 10, 15 y 20 píxeles de radio, para un sitio estable (sitio 44) y un sitio de deslizamiento (sitio 325).	68
Figura 44. Toma de información in situ para las variables cualitativas.	70
Figura 45. Valores de -2 LL correspondientes al análisis de RL in situ y las 20 áreas de vecindad y su correspondiente gráfica. Ejemplo tema litología.	71
Figura 46. Valores de la variable cualitativa (pendiente del terreno) para los sitios de muestreo 44 y 325. Los promedios obtenidos en cada AEVC se asignaron al sitio.	72
Figura 47. Procesamiento de las tablas obtenidas de la toma de información in situ y el AEVC con la RL para identificar la distancia a la que el valor de -2 LL es menor.	73
Figura 48. Integración de la matriz con los datos seleccionados en el AEVC.	74
Figura 49. Representación de una variable espacial cualitativa en su forma original y en función de un tamaño de vecindad de 4 píxeles, en los formatos de mapa raster y matriz de datos.	75
Figura 50. Cálculo espacial del modelo de probabilidad con la fórmula de RL.	75
Figura 51. Generación de la curva COR utilizando la muestra ampliada y el modelo de probabilidad (AEVC e in situ).	76
Figura 52. Generación de la matriz de clasificación con base en la muestra ampliada y los modelos de probabilidad agrupados en los valores de 0 y 1.	76
Figura 53. a) medición de un DSE y b) vista en planta en la que se resalta la cabecera del DSE. c) canal por el que fluyen los escombros en un DSFE y d) área de desembocadura del DSFE.	78
Figura 54. DAP en el que se aprecia que el plano de deslizamiento con una altura de 9 metros está por debajo de las raíces de los árboles.	79
Figura 55. Mapa del inventario de deslizamientos 1984-2016. Por tipo de deslizamientos y su respectivo porcentaje.	79
Figura 56 Resultados del AEVC por variable explicativas.	83 y 84
Figura 57. Mapa de probabilidad a deslizamientos con valores continuos entre 0 y 1 con base al AEVC.	88

Figura 58. Mapa clasificado en zonas de probabilidad a deslizamientos y zonas estables con base en el punto de corte. Modelo con base al AEVC.	89
Figura 59. Mapa de probabilidad a deslizamientos con información in situ. En valores continuos entre 0 y 1.	90
Figura 60. Mapa clasificado en zonas de probabilidad a deslizamientos y zonas estables con base en el punto de corte. Modelo in situ.	90
Figura 61. Resultados de la curva COR para el modelo con base en el AEVC.	91
Figura 62. Resultados de la curva COR para el modelo con base en datos in situ.	92
Figura 63. Mapa de distribución del inventario de deslizamientos con relación al modelo de probabilidad con base en el AEVC.	95
Figura 64. Distribución del inventario de deslizamientos con relación al modelo de probabilidad con base en datos in situ.	96
Figura 65. Mapa de la distribución de las zonas deforestadas y su cálculo con un AEVC de 15 pixeles, y las áreas con probabilidad a deslizamiento.	97
Figura 66. Mapa de las áreas con probabilidad a deslizamientos y el AEVC para las clases de geomorfología utilizadas en el modelo de probabilidad con base al AEVC.	98
Figura 67. Mapa de las áreas con probabilidad a deslizamientos y su relación con el promedio de la curvatura del terreno a un AEVC de 6 pixeles.	99
Figura 68. Mapa de las áreas con probabilidad a deslizamientos y su relación con el promedio de la distancia a ríos dentro del AEVC de 2 pixeles de radio.	100
Figura 69. Mapa de las áreas con probabilidad a deslizamientos y su relación con la textura del suelo gruesa en un AEVC de 20 pixeles.	100
Figura 70. Mapa de las áreas con probabilidad a deslizamientos y su relación con la orientación de la pendiente tomada in situ.	101
Figura 71. Mapa de las áreas con probabilidad a deslizamientos y su relación con la distancia a ríos tomada in situ.	102
Figura 72. Mapa de las áreas con probabilidad a deslizamientos y su relación con la distancia a fallas y fracturas tomada in situ.	102

ÍNDICE DE FÓRMULAS	Página
Fórmula 1. Fórmula para el cálculo del tamaño de la muestra para proporciones en poblaciones infinitas.	24
Fórmula 2. Fórmula para el cálculo del tamaño de la muestra para proporciones en poblaciones finitas.	24
Fórmula 3. Fórmula para la regresión logística con una variable explicativa.	27
Fórmula 4. Fórmula para la regresión logística con dos o más variables explicativas.	27
Fórmula 5. Fórmula para el cálculo del Índice de Kappa.	34

ÍNDICE DE TABLAS	Página
Tabla 1 Métodos en la elaboración de mapas de probabilidad a PG.	12
Tabla 2 Clasificación de PG de acuerdo al WSDNR.	18
Tabla 3.- Principales funciones aplicables al análisis de vecindad.	20
Tabla 4. Codificación de los falsos positivos y aciertos positivos utilizados en la curva COR.	32
Tabla 5.- Clases de la variable de uso del suelo y vegetación 2015.	49
Tabla 6.- Clases de la variable de cambios de uso del suelo y vegetación (1983-2015).	50
Tabla 7.- Clases de la variable de geomorfología.	51
Tabla 8.- Clases de la variable de litología.	52
Tabla 9. Tipos de suelo en la zona de estudio.	53 y 54
Tabla 10.- Clases de la variable de textura del suelo.	54
Tabla 11. Sumatoria de pixeles de la clase con clave 1 de litología a diferentes distancias del AEVC para los sitios de muestreo 44 y 325.	69
Tabla 12. Integración de la sumatoria de las 6 clases del mapa de litología a una distancia de 1 pixel de distancia del área de vecindad. La tabla muestra solo 10 registros de los 384.	69
Tabla 13. Valores de Tolerancia y FIV.	80

Tabla 14. Resultados del Índice de Condicionamiento y la proporción de la varianza.	80
Tabla 15. Resultados para el modelo de RL con base al AEVC.	85
Tabla 16. Resultados para el modelo de RL con base a datos in situ.	87
Tabla 17. Evaluación utilizando la matriz de clasificación. a) modelo de AEVC, b) modelo in situ.	92

RESUMEN

El objetivo del trabajo es aplicar el análisis espacial de vecindad de forma continua en la adquisición de los datos para los sitios de muestreo que son utilizados en la calibración de los modelos de probabilidad al emplear la regresión logística múltiple (RL) para desarrollar un enfoque metodológico que brinde la posibilidad de elaborar un modelo espacial de probabilidad a deslizamientos de mayor precisión en comparación con uno elaborado con la toma de información *in situ* de los sitios de muestreo, siendo esta última la forma tradicional con la que son elaborados este tipo de trabajos de probabilidad de ocurrencia. El modelo de probabilidad se elaboró para la cuenca La Ciénega, localizada en la ladera este del volcán Nevado de Toluca.

Para realizar la comparación de la precisión del modelo basado en el análisis de vecindad se elaboró el modelo con la toma de información *in situ*, y así contar con información que permita conocer las ventajas y desventajas de uno y otro método. La evaluación de los modelos se realizó tanto estadística como espacialmente, en dicha evaluación se obtuvo que el modelo basado en el análisis de vecindad reporta mejores valores para la tabla de clasificación del modelo de RL con un 89.6 de correcta clasificación a nivel general, mientras el modelo de datos *in situ* produjo un valor de 85.9 %, respectivamente. En la evaluación con base a la curva característica de operación del receptor (COR) se obtuvo un valor del 96.8 % de área bajo la curva para el modelo con base al análisis de vecindad y de un 89.8 % para el modelo calibrado con datos *in situ*. Se realizó la validación espacial de los modelos mediante el cálculo del índice de Kappa, exactitud general, exactitud del productor, exactitud del usuario y la eficiencia del modelo, en todos estos el modelo basado en el área de vecindad reportó mejores resultados.

Por último, la representación cartográfica de los modelos muestra que el basado en el análisis de vecindad tiene una distribución espacial más próxima a la que presenta la información del inventario de deslizamientos que se elaboró para el presente trabajo, en comparación con el modelo basado en la toma de datos *in situ*. El modelo cartográfico de datos *in situ* reportó una sobre evaluación del 9.9 % de áreas con probabilidad a deslizamientos en comparación con el modelo basado en el análisis de vecindad.

Los resultados indican que la aplicación del enfoque metodológico propuesto produjo un modelo de probabilidad a deslizamientos para la zona de estudio de mayor precisión al que se obtiene cuando se emplea la forma tradicional de toma de datos *in situ* en la elaboración de modelos utilizando la RL.

ABSTRACT

This work aims to apply the spatial analysis of neighborhoods in a continuum way to acquire the data for sampling sites used to calibrate probability models using multiple logistic regression (LR). The above allowed to develop a methodological approach that provides the possibility of elaborating a better spatial landslides probability model compared to the model calibrated by taking information locally from the sampling sites (called in here as in-situ). The landslides probability model was developed for La Ciénega watershed on Nevado de Toluca volcano's eastern flank.

The landslides probability is calculated with two models: 1) based on the Continuum Neighborhood Spatial Analysis (CNSA) in combination whit LR and 2) taking in-situ data. The two models are compared and contrast to highlight advantages and limitations. The evaluation of the models is carried out statistically and spatially.

The study area results suggest that the model based on the CNSA has a stronger predictive capability than a model elaborated with in-situ data. The classification table of the LR model reports an 89.6 of correct classification for overall accuracy, followed by the In-situ data model whit an 85.9%. In the evaluation based on the receiver operating characteristic curve (ROC), 96.8% of the area under the curve (AUC) was obtained for the model based on the CNSA and 89.8% for the model calibrated with in-situ data. The models were also validated with the Kappa index, general accuracy, producer accuracy, user accuracy, and the model's efficiency. In all these tests, the model based on CNSA has a better coincidence with the field information.

Finally, the spatial overlay between the CNSA model and field information has a better cartographic representation and prediction of landslides areas. The results indicate that the proposed methodological approach produced a better landslide probability model for the study area than the traditional form of in-situ data collection to calibrate the RL model.

I INTRODUCCIÓN.

Los procesos gravitacionales (PG) son uno de los principales peligros geomorfológicos provocados por fenómenos naturales a nivel mundial (CENAPRED, 2001; Lugo, 2011; Nourani y Ghaffari, 2012; MatebieMeten *et al.*, 2014), lo que los hace una de las causas más frecuentes de pérdida de vidas humanas y materiales (Rodríguez *et al.*, 2006; Mirnazari *et al.*, 2014). Este trabajo se enfoca dentro de los PG a los movimientos de tierra denominados como deslizamientos de acuerdo al Departamento de Recursos Naturales del Estado de Washington (WSDNR, 2006a). En el trabajo se elabora un inventario de deslizamientos y el posterior análisis de variables asociadas a estos procesos para la generación de un modelo espacial de probabilidad de ocurrencia en el área de estudio. El modelo de probabilidad y su representación cartográfica son de gran importancia para la prevención de desastres, ya que con ello es posible planificar espacialmente acciones dirigidas a disminuir los efectos que estos pudieran tener sobre la población y sus actividades.

El trabajo se realiza en la cuenca del río La Ciénega, ubicada en la ladera este del volcán Nevado de Toluca, en el estado de México. Dentro de la cuenca se encuentra el poblado de Santa Cruz Pueblo Nuevo, el cual se asentó en el mismo sitio en el que estaba situado el poblado de Santa Cruz que en el año de 1940 fue sepultado por un flujo de escombros (Toscana y Valdez, 2014). Lo que representa un peligro latente para el poblado.

Con el propósito de alcanzar una mayor precisión en la clasificación de zonas con probabilidad a deslizamientos al aplicar la RL en el modelado espacial, se implementaron técnicas y métodos de análisis espacial y estadístico, para la evaluación de la información utilizada en la elaboración del modelo. La forma tradicional de calibrar los modelos en la RL inicia con la selección probabilística de sitios de muestreo, en la localización de cada uno de estos sitios se registra el valor de las variables explicativas y con los datos recabados se procede a calibrar el modelo estadístico para calcular la probabilidad de que una zona sea propensa a presentar deslizamientos. La recopilación de los valores de las variables explicativas únicamente en el lugar donde se posicionan los sitios de muestreo excluye una gran cantidad de información de la zona de estudio, lo que resulta en la calibración del modelo con una mínima cantidad de información. La integración de información alrededor de los sitios de muestreo al momento de adquirir los datos para la calibración del modelo de RL puede producir una mayor precisión en el modelo de probabilidad.

Para integrar la información circundante a los sitios de muestreo seleccionados se utilizó el análisis espacial de vecindad, el cual asocia un sitio con su entorno (Aronoff, 1989). Esta técnica

se aplicó aumentando el área de vecindad de manera continua, a este proceso se le llamo Análisis Espacial de Vecindad Continua (AEVC). Con la implementación del AEVC se recopilaron datos con los cuales se integró el entorno de los sitios a la elaboración del modelo de probabilidad.

En un primer análisis, las variables explicativas fueron trabajadas individualmente con el AEVC y la RL, con ello se evaluó la relación estadística existente entre el proceso de deslizamientos con una determinada variable en función del tamaño del área del área de vecindad. Esto permitió identificar para cada variable cual es el radio de vecindad que reportó estadísticamente una mayor relación con el proceso de deslizamientos. Posteriormente se elaboró una matriz de datos que integra los valores de todas las variables explicativas a la distancia seleccionada en el primer análisis. Una vez integrada la matriz con base a los resultados del AEVC se aplicó el modelo estadístico de la RL para la elaboración del modelo de probabilidad de ocurrencia a deslizamientos. Enseguida, los resultados del modelo estadístico fueron implementados en la información cartográfica de las variables explicativas seleccionadas para finalmente generar el modelo espacial de probabilidad a deslizamientos con base en el AEVC.

Con el propósito de realizar una comparación de los resultados obtenidos en el modelo elaborado con el AEVC se trabajó un modelo de probabilidad basado en la toma de información *in situ*, en el que solo se utilizaron los datos recopilados justo en la posición donde se establecieron los sitios de muestreo. Los modelos obtenidos fueron evaluados estadísticamente con la matriz de clasificación y la curva característica de operación del receptor (COR), y especialmente con el cálculo del índice de Kappa, la exactitud general, la exactitud del productor, la exactitud del usuario y la eficiencia del modelo. Lo anterior permitió identificar que el modelo con base a la información del AEVC reporto un mejor desempeño en los aspectos evaluados, obteniendo una mejor tabla de clasificación durante el proceso de calibración del modelo estadístico. En lo que respecta al modelo espacial de probabilidad con base al AEVC, este reporta un mejor nivel de clasificación en las áreas estables y de deslizamiento en comparación con el modelo con información *in situ*, siendo los valores de exactitud del usuario y la eficiencia del modelo los que tiene un mejor nivel de ajuste, el primero reportando una mejora del 4.68 %, mientras que el segundo señala un aumento en su valor de 0.08. Además de los valores más exactos del modelo basado en el AEVC, la distribución espacial de las áreas con probabilidad fue de mayor precisión en este modelo que en el modelo *in situ*. El modelo basado en el AEVC presenta una mayor coincidencia con las áreas donde se identificaron deslizamientos en el inventario realizado, mientras que el modelo con información local reporto un área con una importante probabilidad a deslizamientos en la que de acuerdo al inventario su presencia es muy escasa. Esta mayor

precisión de las áreas con probabilidad es de gran importancia ya que con ello es posible hacer una interpretación de cómo estas zonas propensas a deslizamientos podrían afectar las actividades de antrópicas dentro del área de estudio. Con esta información se puede establecer que actividades pueden ser afectadas por posibles deslizamientos, así como la localización y distribución de acciones dirigidas a minimizar o evitar el efecto negativo de los deslizamientos en las actividades antrópicas.

II. PLANTEAMIENTO DEL PROBLEMA.

En la elaboración de modelos espaciales de probabilidad a PG utilizando la RL, se hace uso de sitios de muestreo tanto para la calibración del modelo como para la validación del mismo (Hair *et al.*, 1999; Eskandarj y Chuvieco, 2015). Con respecto a la calibración del modelo, la información que se recaba es la que se encuentra justo en el sitio en donde se ubica el punto de muestreo, es decir la información *in situ* (Tayyebi *et al.*, 2010; Abdulah y Yulianti, 2015; Eskandarj y Chuvieco, 2015; Xiong y Zuo, 2018). Esta forma de tomar la información de los sitios de muestreo implica un aislamiento del punto donde es adquirida la información con relación a su entorno, ya que se registran solo los valores en donde se especifica la ubicación del mismo.

Tobler (1970, p. 236) menciona que “todo está relacionado con todo, sin embargo, las cosas cercanas están más relacionadas que las cosas lejanas”. Con base en lo anterior se supondría que tomando la información de alguna variable relacionada con los deslizamientos justo en el lugar donde se ubica el proceso se estaría adquiriendo el dato más relacionado. Pero ¿cómo afecta una distribución heterogénea de valores de una variable al sitio en que se tome la información?, ¿Y hasta que distancia, si la hay, la relación de un sitio con su entorno reporta una mayor fuerza estadística con el proceso que se analiza?

Cuando los valores de una variable alrededor de un sitio de muestreo son heterogéneos la influencia que tales datos tienen sobre el sitio está en función de la distancia a la que estos se encuentren en combinación con la diversidad e intensidad de sus valores.

La integración de la información de las zonas vecinas a los sitios de muestreo en áreas inestables y estables, y su posterior análisis estadístico utilizando la RL permitirá evaluar la relación entre las variables dependiente y explicativas, y proporcionará la información necesaria para determinar la distancia óptima a la que las variables explicativas tienen una mayor relación con los PG.

Con base en lo anterior, esta investigación tiene el propósito implementar y evaluar una metodología con base en el análisis espacial de vecindad que permita la integración del entorno, con respecto a las variables explicativas, de los sitios de muestreo utilizados en la RL, con el fin de obtener modelos de probabilidad a deslizamientos de mayor precisión en comparación con modelos generados a partir de la toma de información *in situ*.

III. OBJETIVOS.

III.I. Objetivo general.

Aplicar el análisis espacial de vecindad de manera continua y la RL para desarrollar un enfoque metodológico que tiene el propósito de generar modelos de probabilidad a deslizamientos de mayor precisión que los elaborados con información tomada de manera puntual utilizando el modelo estadístico de RL.

III.II. Objetivos específicos.

- Realizar la cartografía sobre la ubicación de deslizamientos en la zona de estudio a través de la elaboración del inventario de deslizamientos con trabajo de campo y la recopilación de información previa.
- Identificar el tamaño de radio del área de vecindad de los sitios de muestreo en el cual la relación estadística con el proceso de deslizamientos es más significativa para cada una de las variables explicativas recopiladas, utilizando el AEVC y el indicador de -2 logaritmo de la verosimilitud de la RL.
- Elaborar un modelo de probabilidad a deslizamientos de mayor precisión aplicando para cada una de las variables explicativas la distancia que reporto la mayor relación con el proceso de deslizamientos de acuerdo al AEVC
- Elaborar el modelo de RL de la probabilidad a deslizamientos con base en información *in situ* de los sitios de muestreo.
- Evaluar y comparar, estadística y espacialmente, los modelos de probabilidad generados al aplicar el modelo de RL con base en el AEVC y con base en datos *in situ*.

IV. HIPÓTESIS.

Al aplicar la función de AEVC a los sitios de muestreo para analizar las variables relacionadas con la ocurrencia de deslizamientos se estará integrando información sobre el entorno de los sitios que servirá para identificar una relación integral de las variables explicativas con el proceso de deslizamientos y en consecuencia se obtendrá un modelo espacial de probabilidad a deslizamientos de mayor precisión en comparación con el modelo elaborado con información *in situ*.

CAPÍTULO 1. ANTECEDENTES Y MARCO TEÓRICO.

1.1 La Geomática en los estudios geográficos.

La elaboración de modelos espaciales basados en el uso de herramientas informáticas se ha desarrollado a partir de los años 1990's (Berry, 2013a). Para la ciencia geográfica es de primordial importancia representar de manera espacial la distribución y localización del proceso que se estudia, además de ello es importante conocer de qué manera los elementos que componen un determinado espacio tienen relación con tal proceso (Rojas, 2005; Vargas, 2012). Para llevar a cabo esta tarea la geografía hace uso de diversas ramas del conocimiento y herramientas con las que le es posible estructurar metodologías que le permitan modelar las interacciones de las variables que dan origen a determinado proceso geográfico (Rojas, 2005). El modelado espacial en la geografía se ha venido desarrollando de una forma cada vez más especializada como resultado del avance del conocimiento geográfico y el uso de las herramientas informáticas, las cuales aumentan de manera importante las funciones posibles de realizar, así como el volumen de datos a procesar (Mas *et al.*, 2002; Álvarez *et al.*, 2003; Castro, 2011; García, 2017; Zambrano *et al.*, 2017).

Con el surgimiento de los Sistemas de Información Geográfica (SIG) a finales de los años 1960's, resultado de la implementación de programas de cómputo en los que fue posible integrar y analizar información geográfica en su parte espacial (gráficos) y descriptiva (atributos) (Gomasasca, 2009; Berry, 2013a), se pudo contar con una herramienta informática que cambió la concepción de la cartografía que hasta entonces había sido utilizada en formato análogo. Los mapas digitales contienen atributos cualitativos y/o cuantitativos que pueden ser analizados de acuerdo al tipo de dato que representan, estos pueden ser datos cualitativos de tipo nominal u ordinal o datos cuantitativos de tipo de intervalo, de razón o binarios (Berry, 2013b). De acuerdo al tipo de dato con el que sea representado un elemento del espacio geográfico será el tipo de análisis que se le puede aplicar. Por ejemplo, a los datos cualitativos se les puede aplicar un número limitado de operaciones estadísticas, en comparación a los datos cuantitativos a los cuales se les puede aplicar una amplia variedad de análisis estadísticos y matemáticos (Berry, 2013b).

Desde el surgimiento de los SIG se han desarrollado métodos para la elaboración de cartografía digital hasta la generación de modelos espaciales de probabilidad (Berry, 2013a). Esta evolución en el manejo de información cartográfica digital implicó el surgimiento e implementación de nuevas ciencias como lo es la Geomática. El término Geomática fue utilizado por primera vez a inicios de los años 1970 por el ministerio de Equipamiento y Vivienda de Francia, en esos tiempos

el termino solo hacía referencia al procesamiento automático de información geográfica (Kemp, 2008). La conceptualización de la Geomática como la conocemos hoy en día tuvo lugar en la Universidad de Laval en Canadá al comienzo de los años 80's del siglo XX (Gomasasca, 2009). Actualmente se pueden encontrar un gran número de definiciones de la geomática, entre las que se encuentran aquellas que reflejan una visión integral en las que se considera una amplia variedad de ciencias, disciplinas y técnicas que la componen y pone mayor énfasis en las funciones y flujos de información que la geomática realiza con la información geográfica (Kemp, 2008; Samuelsen, 2008; Gomasasca, 2009; DiGruttolo y Mohamed, 2011; Pinto, 2013). También se pueden encontrar otras definiciones que hace énfasis en la obtención de información georreferenciada y precisa, las cuales equiparan de manera directa las funciones de la geomática con lo que venía realizando la ingeniería topográfica (Ghilani y Wolf, 2012).

Entre las disciplinas y técnicas que conforman la geomática se encuentran: Las ciencias de la computación, geodesia, topografía, cartografía, fotogrametría, sensores remotos, sistemas de posicionamiento global, sistemas de escaneo por láser, sistemas para el soporte de decisiones, SIG, sistemas expertos, SIG web (Figura 1).

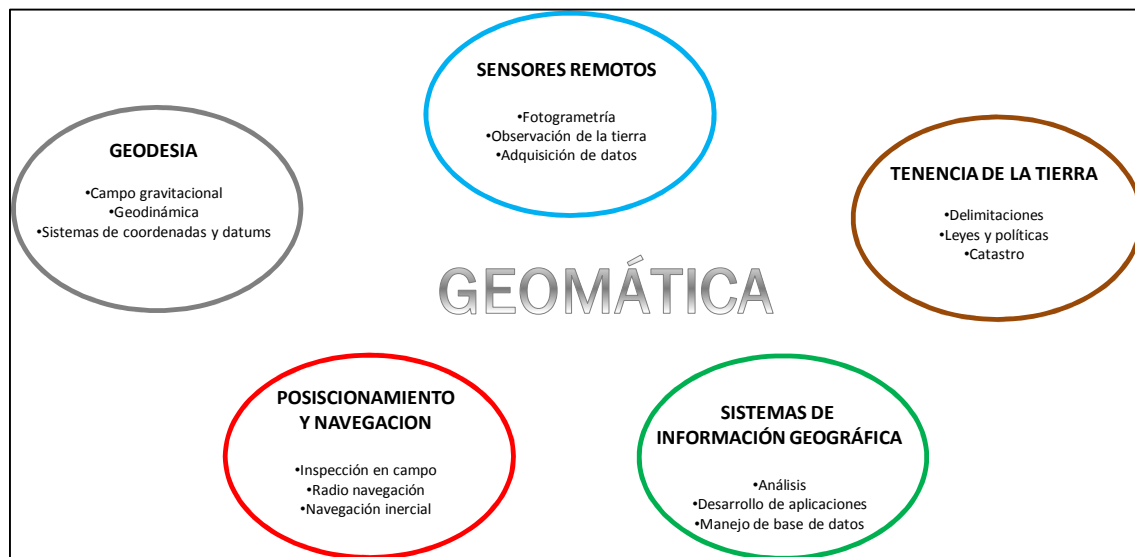


Figura 1. Áreas de investigación, disciplinas y técnicas que integran la geomática. Elaboración propia con base en Gomasasca (2009).

El desarrollo de metodologías para la elaboración de modelos espaciales ha sido un área ampliamente trabajada en estudios de carácter geográfico. Este tipo de trabajo se han realizado en una diversidad de temas como pueden ser modelos forestales (Leal-Nares *et al.*, 2012; Castro,

2014; Vilchis-Francés *et al.*, 2015), de distribución de especies animales o vegetales (Augustin *et al.*, 2001; Ibarra-Montoya *et al.*, 2012; López-Sandoval *et al.*, 2015), factores físicos como las inundaciones (Zambrano *et al.*, 2018), deslizamientos (Atkinson y Massari, 2011; Yang *et al.*, 2019) por mencionar algunos.

1.2. Modelado espacial de probabilidad a deslizamientos.

El estudio de la probabilidad de ocurrencia de PG se ha realizado desde varias perspectivas. Hasta ahora, se han desarrollado trabajos con metodologías y enfoques diversos (Jiménez-Perálvarez *et al.*, 2008; Legorreta *et al.*, 2014a; Schlögel *et al.*, 2018; Yang *et al.*, 2019). En particular, para la ubicación de las áreas susceptibles a deslizamientos se han empleado diversos métodos en diferentes condiciones geográficas (Nouran y Ghaffari, 2012; Feizizadeh y Blaschke, 2014; Legorreta *et al.*, 2014b; MatebieMeten *et al.*, 2014). Entre los métodos utilizados se encuentran la relación de frecuencias, la regresión logística (Esper, 2013; Mirnazari *et al.*, 2014; Legorreta *et al.*, 2014b), análisis multicriterio y lógica difusa (Dehban, 2013).

Ejemplo de estos estudios es el elaborado por Atkinson y Massari (2011), en el cual se realizó un modelo de la susceptibilidad a deslizamientos en los Apeninos Centrales en Italia. El estudio se basa en la aplicación del modelo de RL al que se le agrega una variable explicativa que es producto del cálculo del promedio de susceptibilidad a deslizamientos de los vecinos de una celda, esta nueva variable es agregada al modelo de RL lo que lo convierte en un modelo autologístico. Otro trabajo de este tipo es el elaborado por Yang *et al.* (2019), en la cuenca Duwen Highway, provincia de Sichuan, China, para la representación cartográfica de la susceptibilidad a deslizamientos. En este trabajo aplicaron la RL junto con el programa GeoDetector, este programa es empleado para realizar una evaluación cuantitativa de la importancia de diversas variables para el desarrollo de los PG y puede ser utilizado para la selección de variables explicativas que serán integradas en la calibración del modelo.

Los modelos realizados por Jiménez-Perálvarez *et al.* (2008) y Othman *et al.* (2018) aplican técnicas estadísticas para realizar mapas de susceptibilidad a deslizamientos; el primero en la Cordillera Bética, al sur de la Sierra Nevada, España, en este trabajo se elaboró un cruce de mapas de 4 variables explicativas (ángulo de la pendiente, altitud, dirección de la pendiente y litología) y un inventario de deslizamientos para finalmente aplicar una relación de frecuencia y así obtener un mapa de susceptibilidad a deslizamiento; el segundo en el área de Mawat, en la región Kurditan, en el noreste de Iraq, en este se aplicaron los modelos estadísticos de relación de frecuencias, de peso de la evidencia, RL y regresión Probit.

Los modelos de aprendizaje automático (MAA) son de tipo cuantitativo y han sido utilizados para la elaboración de mapas de susceptibilidad a deslizamientos, ejemplo de estos trabajos son el elaborado por Aditian *et al.* (2018). En este trabajo se compararon 3 modelos cualitativos, utilizando relación de frecuencias (RF), RL y Redes Neuronales Artificiales (RNA). Para la elaboración de mapas de susceptibilidad a deslizamientos, en la región de Ambon, Indonesia. Los resultados de este trabajo muestran que los tres métodos generan valores por arriba del 66 % en el análisis de la Curva de Operador del Receptor (COR) con respecto al área bajo la curva (ABC), siendo el modelo de RNA el que obtuvo el valor más elevado al alcanzar arriba del 71 % en los valores de ABC. Otro trabajo de este tipo es el elaborado por Zhou *et al.* (2018), en el cual se generó un modelo de susceptibilidad a deslizamientos aplicando MAA en un área de la reserva Three Gorges, en China: En este trabajo se aplican los métodos de RNA, el de Máquina de Soporte de Vectores (MSV) y la RL, los resultados muestran que el modelo elaborado con MSV da los mejores resultados con un 88 % de ABC, le sigue el de RNA con un 83.6 % de ABC, por último, está el de RL con un 69.7 % de ABC.

Otros métodos para la elaboración de este tipo de mapas son los que se basan en el conocimiento heurístico del tema, la característica de este tipo de trabajo es que las variables que son consideradas como catalizadoras de los PG son ordenadas y se les da un valor de acuerdo a la relación que estos tengan con la presencia del proceso en estudio, todo esto con base en el conocimiento de los expertos. Ejemplo de este es el mapa de susceptibilidad a deslizamiento generado por Tsangaratos y Rozos, (2013). Este mapa fue producido con un enfoque cualitativo aplicando análisis multicriterio, a través de dos técnicas que son el Sistema de Ingeniería de Rocas y el Proceso de Análisis Jerárquico (PAJ). En ambos métodos se realiza una clasificación de cada una de las variables que tienen relación con la presencia y distribución de los PG, esta clasificación le da una importancia relativa a cada una de ellas de acuerdo al conocimiento de los expertos. Una vez jerarquizadas las variables estas son trabajadas con una combinación lineal de pesos para obtener los mapas de susceptibilidad.

Otro ejemplo de este tipo de modelos es el realizado por Nourani y Ghaffari, (2012) para la elaboración de un mapa de susceptibilidad a deslizamientos aplicando el PAJ en la provincia de Zonouz al este de Azerbaijan, para ello se identificaron cuáles son los factores de mayor relación con los deslizamientos para después darle a cada uno de estos un valor de acuerdo a su importancia con relación a la presencia del proceso. Con ello se procedió a realizar matrices de comparación por pares, para producir un vector de pesos que refleja la importancia de los factores

analizados. Finalmente se agruparon los valores del vector de pesos en cuatro categorías para obtener el mapa de susceptibilidad.

Hay otro tipo de modelos en los que es necesario contar con mediciones sobre las características de las variables que serán tomadas como explicativas de los PG, un modelo de este tipo es desarrollado en el trabajo de Legorreta y Bursik, (2009), en este presentan una herramienta llamada LOGISNET que incluye 2 métodos para el análisis de estabilidad de laderas: RL y un modelo cartográfico-hidroológico conocido como Mapeo de Índices de Estabilidad (SINMAP, por sus siglas en inglés). Para poder realizar este último modelo es necesario contar con mediciones de propiedades geotectónicas de los suelos que serán integrados en la elaboración del modelo.

La amplia gama de trabajos en la elaboración de mapas de susceptibilidad a PG dan lugar a varias clasificaciones sobre las metodologías para realizar este tipo de estudios. Jiménez-Perálvarez *et al.* (2008) y Yang *et al.* (2019) las agrupan en tres tipos: empírico o heurístico, las cuales se basan en el conocimiento de los científicos responsables del estudio; cuantitativos estadísticos, estos permiten una mejor comprensión entre las relaciones de las variables explicativas y los PG a través del análisis cuantitativo de las variables; por último, los métodos físicos, estos implican un análisis de la estabilidad de laderas por medio de estudios detallados de características de la zona de estudio. Legorreta *et al.* (2012a), menciona una clasificación desde cuatro enfoques: por inventario, heurísticos, estadísticos y determinísticos. Por su parte, Aditian *et al.* (2018) los divide en métodos cuantitativos y cualitativos. En la Tabla 1 se muestra una clasificación de los diversos métodos utilizados en la elaboración de mapas de susceptibilidad a PG y sus principales características.

La subjetividad de los métodos heurísticos hace difícil mantener una consistencia en su aplicación cuando se elaboran en zonas con características diferentes. Por su parte, los modelos estadísticos tienen un enfoque objetivo, en lo que respecta a la cuantificación de la relación de las variables explicativas con los PG (Jiménez-Parálvarez *et al.*, 2008; Othman *et al.*, 2018). Lo anterior permite que los modelos estadísticos sean comparables, siempre y cuando se esté trabajando el mismo modelo, lo que permite comprobar la efectividad del modelo en diferentes áreas. En lo que respecta a los modelos físicos, la medición de parámetros de las variables involucradas para la zona que se está estudiando hace que estos modelos sean más costosos que los dos tipos de modelos anteriores. Existe la posibilidad de que los parámetros obtenidos en una determinada zona de estudio puedan ser extrapolados a otra zona con características similares.

Tabla 1 Métodos en la elaboración de mapas de probabilidad a PG.

Método	Características
Heurístico	La cuantificación de los pesos de las variables involucradas en el estudio se basa en el conocimiento de expertos. Una vez hecha la jerarquización de las variables y sus clases se procede a realizar una agregación de estos pesos para determinar las áreas con una mayor susceptibilidad al desarrollo de los procesos.
Estadístico	Se basan en la aplicación de modelos estadísticos, en ellos se busca identificar, con base en una análisis cuantitativo, la relación de hay entre las variables explicativas con los PG presentes en la zona de estudio.
Físicos	Parten de la medición detallada de las características físicas (por ejemplo: espesor del suelo, transmisividad, cohesión del terreno, entre otras) que se identifiquen con una mayor relación con el desarrollo de los PG, a partir de la toma de información se aplican modelos matemáticos con los que se llega a representar el grado de susceptibilidad a PG.

Fuente: Elaboración propia con base en información de Jiménez-Perálvarez *et al.* (2008); Legorreta *et al.* (2012); Aditian *et al.* (2018); Yang *et al.* (2019).

Los métodos estadísticos son los más utilizados en la elaboración de mapas de áreas propensas a PG, los niveles de precisión alcanzados por este tipo de modelos se encuentran entre valores que van de 70 % al 93 % (Esper, 2013; Chalkias *et al.*, 2014; Legorreta *et al.*, 2014b; Mirnazari *et al.*, 2014). La aplicación de la RL en este tipo de modelos ha mostrado su eficiencia en varios casos en los que los niveles de precisión se encuentran entre los 72 % y 84 % (Ayalew y Yamagishi, 2005; Esper, 2013; Legorreta *et al.*, 2014b).

Con respecto al análisis espacial que se emplea en la elaboración este tipo de mapas, este se puede agrupar en dos tipos. Uno de ellos es el modelado por zonas (ESRI, 2001; Berry, 2013a) (Figura 2). En este, la evaluación de la probabilidad se realiza con base a los límites de regiones, que son determinadas por el cruce de las variables involucradas en el análisis, la combinación de las clases genera nuevas áreas las cuales son caracterizadas de acuerdo al método que se esté empleando (heurístico, estadístico o físico). Por lo general los resultados reflejan límites discretos por lo que pueden ser fácilmente representados con límites lineales.

Otro tipo de análisis espacial es el local. En este, el método utilizado al momento de evaluar la zona de estudio es aplicado a nivel puntual (a nivel de pixel) (ESRI, 2001; Berry, 2013a). El análisis local permite la generación de modelos de probabilidad de tipo continuos en los que los valores tienen gradientes que permiten representar superficies sin cambios bruscos (Figura 3).

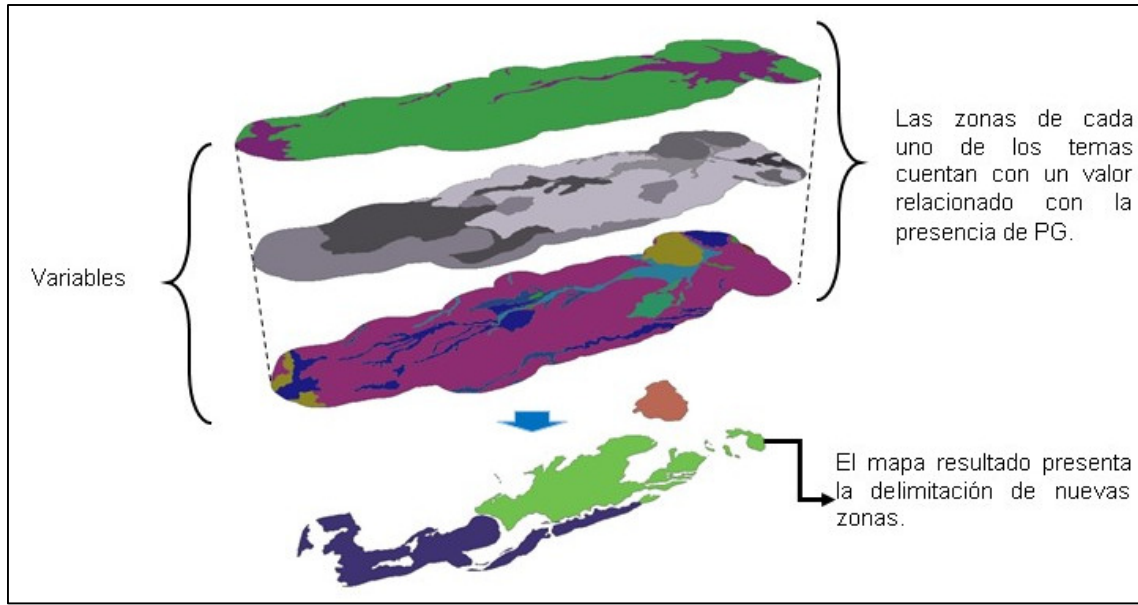


Figura 2. Análisis espacial por zonas, por lo general utilizado en los modelos heurísticos. Elaboración propia.

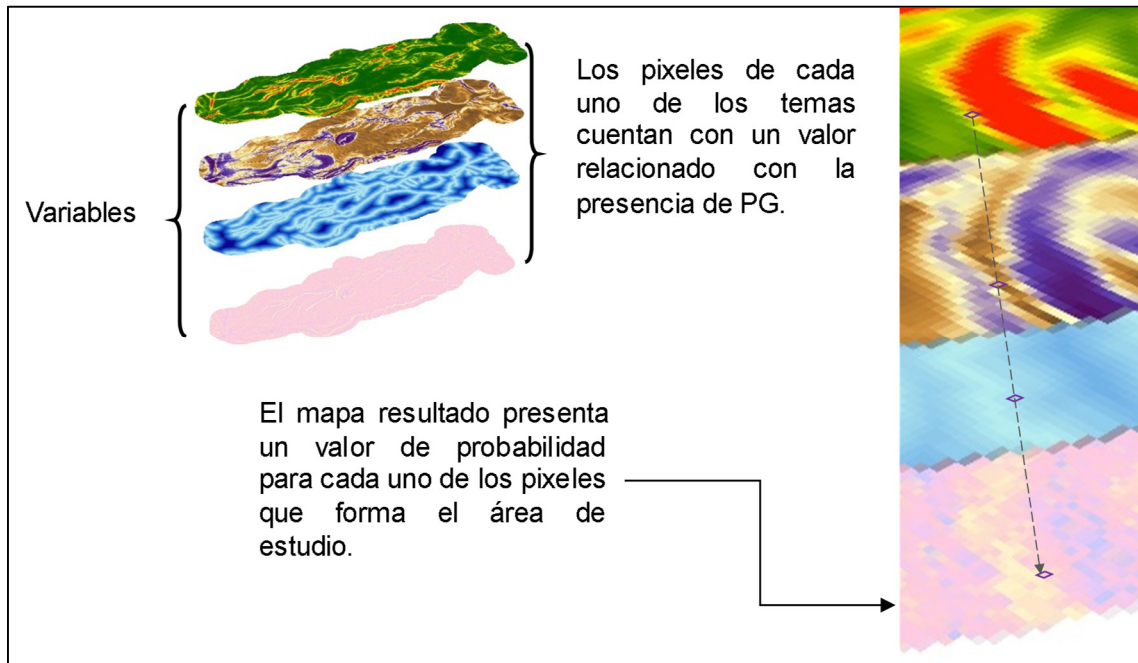


Figura 3. Análisis espacial puntual, es muy utilizado en modelos estadísticos. Elaboración propia.

En México la elaboración de modelos espaciales de probabilidad a ocurrencia de PG se comenzó a desarrollar recientemente. En 2012 Legorreta *et al.* (2012b) presenta uno de los primeros trabajos de este tipo, se trata de la elaboración de mapas de susceptibilidad a deslizamientos utilizando la aplicación LOGISNET. La zona de estudio fue la cuenca del Río Chiquito-Barranca del Muerto, en la ladera sur del volcán Pico de Orizaba, entre los estados de Puebla y Veracruz.

LOGISNET recolecta, almacena, integra, analiza y modela datos georreferenciados con el propósito de generar información sobre la distribución de zonas con susceptibilidad a deslizamientos con base en 2 métodos: la RL y un modelo índice de estabilidad. La aplicación realiza la evaluación de los modelos obtenidos comparando un inventario de deslizamientos de la zona con el mapa de susceptibilidad a deslizamientos de cada modelo. Los resultados de la evaluación reportaron que el modelo de RL obtuvo el valor más alto en la precisión del productor con un 72.33 %, mientras que para el modelo de índice fue de 43.39 %, se observó que el alto valor alcanzado por el primer método de RL se debe a que en este hubo una sobrevaluación de las áreas con susceptibilidad a deslizamientos, esto favorece a que se dé una mayor coincidencia entre las áreas que el modelo marca como zonas inestables y la información de campo.

Otro trabajo realizado es el de Peña, (2013) donde se elaboró un modelo espacial de susceptibilidad a deslizamientos en la sierra Norte de Chiapas. Este se basó en la asignación de pesos a las variables involucradas en función de su relación con la ocurrencia de deslizamientos. Se trabajó en un SIG en el que se combinaron las diferentes capas de las variables explicativas y la suma de los pesos de cada una de ellas fue la que determinó la susceptibilidad a deslizamientos en el mapa resultante. En este trabajo no se realizó una comparación entre la distribución del inventario de deslizamientos con el mapa obtenido.

El trabajo elaborado por Paz *et al.* (2017) presenta la elaboración de un mapa de susceptibilidad a procesos de remoción en masa en la ladera sur de la ciudad de Tuxtla Gutiérrez, Chiapas. En éste se aplica el método heurístico en combinación con el proceso de análisis jerárquico. Este trabajo implicó la ponderación de las clases en las variables explicativas por parte de especialistas en el tema y en la zona de estudio. Como resultado se obtuvieron 5 niveles de amenaza (muy baja, baja, media, alta y muy alta), las cuales fueron presentadas en un mapa de la zona de estudio. En este trabajo se realizó una sobreposición del inventario de deslizamientos que se elaboró en el estudio con el mapa final de susceptibilidad y se pudo observar una alta coincidencia entre las zonas clasificadas como de alta y muy alta amenaza a los PG inventariados.

En el área de estudio se han realizado trabajos sobre la evaluación de PG. En 2014 se realizó una investigación para identificar las áreas potenciales a deslizamientos y flujos de escombros utilizando un método heurístico con base en el análisis multicriterio de 7 variables: 5 de estas morfométricas (hipsometría, energía del relieve, erosión vertical, pendiente y erosión potencial), así como la geología y geomorfología. Los valores para cada variable y sus categorías se asignaron en función de la importancia de cada factor como detonante de los deslizamientos de acuerdo al conocimiento de expertos. Como resultado se obtuvo un mapa de peligros por

deslizamiento que fue categorizado 5 clases: nulo, bajo, medio, alto y muy alto. Relacionando las zonas con muy alto riesgo y la información de campo sobre la ubicación de deslizamientos se obtuvo una certeza del 75 % (Aceves *et al.*, 2014a). En 2018 García *et al.*, elaboraron un trabajo sobre la susceptibilidad geomorfológica a PG en dos arroyos secundarios de la cuenca La Ciénega. Este trabajo se desarrolló a partir de la elaboración de un inventario de deslizamientos, la elaboración de un mapa geomorfológico actualizado y el cálculo de la susceptibilidad a PG a partir de la cuantificación del rango de valoración por área y el rango de valoración por frecuencia del número y superficie de deslizamientos en relación con las formas del relieve y su superficie (García *et al.*, 2018).

1.3. Procesos gravitacionales y sus principales clasificaciones.

Los PG son estudiados por la Geomorfología, que a su vez es una especialidad de la Geografía (Siso, 2010). Estos se refieren al desplazamiento de material no consolidado de las laderas, principalmente roca y suelo, pendiente a bajo, fundamentalmente por efecto de la fuerza de gravedad y sin una influencia importante de algún otro medio como sería el agua, el aire o el hielo (Selby, 1993; Lugo, 2011). La ocurrencia de los PG incluye un gran número de variables, tanto en el tipo de materiales como en los demás elementos que determinan la disposición de los mismos. Para poder hacer una adecuada clasificación de los diferentes tipos de PG es necesario tener en cuenta los siguientes factores: la velocidad y mecanismo del movimiento, los materiales involucrados, el tipo de deformación, la geometría en el movimiento de los materiales y el contenido de agua en el material desplazado (Selby, 1993).

Por todos los elementos involucrados en los PG, además de los diversos enfoques con los que estos son abordados por diversas disciplinas existen una amplia variedad de clasificaciones (Selby, 1993; Alcántara-Ayala, 2000). De acuerdo con Selby (1993), una de las primeras clasificaciones que fue ampliamente utilizada es la elaborada por Sharpe del año de 1938. La clasificación de Sharpe se basa principalmente en la velocidad del movimiento, el cual podría ir desde un movimiento lento hasta un movimiento rápido, también se refiere a los materiales que se desplazan, a los que clasifica en rocas o suelo, otra característica importante es el contenido de hielo o agua presente en el material deslizante. En total describe 10 clases diferenciadas por tipos de movimiento (flujos y deslizamientos), por la velocidad del mismo y el contenido de agua y hielo (Selby, 1993).

Otra clasificación que tiene una amplia difusión y uso es la elaborada por Varnes (1978). En esta se diferencia a los PG con base al tipo de material involucrado, separándolos entre rocas y suelos

gruesos o finos, así como al tipo de movimiento, agrupándolos en 6 tipos: caídas, vuelcos, deslizamientos, expansiones laterales, flujos y complejos, siendo estos últimos la combinación de 2 o más tipos de los 5 primeros movimientos. La combinación de los tipos de movimiento y de los materiales da como resultado un total de 22 clases de PG.

En 1988 Hutchinson elaboró una clasificación dividida en dos secciones; la primera hace referencia a los movimientos que se presentan en la ladera, y considera los movimientos de rocas o suelos. Se basa principalmente en la morfología del material en movimiento, considerando los mecanismos, materiales y velocidad del desplazamiento. La segunda sección se basa en la geotécnica, el objetivo de esta segunda sección son los trabajos para estabilizar las áreas de desplazamiento (Selby, 1993). La primera sección describe 8 tipos de movimientos que, al considerar la morfología, los materiales y la velocidad del mismo llega a una clasificación de poco más de medio centenar de categorías de PG.

Además de las clasificaciones mencionadas, se han elaborado otras que dan respuesta a enfoques y objetivos diversos (Alcántara-Ayala, 2000). Con el propósito de hacer uso de una clasificación que sea compatible con las características de la zona de estudio, así como a los objetivos del presente trabajo se utilizó la clasificación elaborada por el Departamento de Recursos Naturales del Estado de Washington (WSDNR, por sus siglas en inglés) (WSDNR, 2006a). Esta clasificación está basada en el manual de la División de Practicas Forestales (WSDNR, 2006b), que tiene como base la clasificación de Varnes (1978). La clasificación del WSDNR divide a los deslizamientos en función de la profundidad de área del plano de deslizamiento o falla; los clasifica en superficiales o profundos: Los deslizamientos superficiales se refieren a fallas dentro de la zona de las raíces de los árboles, por lo general a una profundidad no mayor a 3 metros, y sobre el límite del lecho rocoso o sedimentos de glaciar. Entre los materiales que componen el terreno se puede encontrar suelo, regolita, coluvión, aluvión, depósitos superficiales de baja permeabilidad y otro tipo de sedimentos del manto rocoso. El deslizamiento es por lo general rápido y de corta duración. Este tipo de deslizamientos se subdivide en 4 tipos: deslizamiento de escombros, flujo de escombros, vuelcos y caídas, y avalanchas de nieve (WSDNR, 2006a).

- 1) Deslizamiento de escombros: Es un deslizamiento superficial de poca profundidad que se forma por la separación de material en una pendiente pronunciada, implica un movimiento rápido de suelo y regolita sobre el lecho rocoso. Incluye los tipos de deslizamiento conocidos como superficial rápido, deslizamientos de suelos y avalancha de escombros.

La falta de una cantidad de agua significativa es lo que diferencia a un deslizamiento de escombros de un flujo de escombros (WSDNR, 2006a).

- 2) Flujo de escombros: Es un deslizamiento superficial que fluye dentro de un canal formado ya sea por las paredes de un valle de un orden tributario bajo o por diques que el mismo flujo forma. Contiene suelo y agua con cantidades variables de escombros leñosos y se caracteriza por ser un flujo canalizado, por lo regular tiene un largo recorrido hacia su desembocadura. Puede incluir los flujos de lodo, torrentes de escombros, lodos líquidos y deslizamientos por inundaciones causados por la ruptura de presas (WSDNR, 2006a).
- 3) Vuelcos y caídas: Están constituidos por bloques individuales de suelo o roca que se desprenden de una pendiente pronunciada y caen por el aire, rebotan, o ruedan antes de descansar en pendientes más suaves. Los vuelcos y caídas de suelo tienden a desintegrarse, mientras que en los constituidos por rocas estas últimas no se desintegran. El constante vuelco y caída de bloques de suelo conducen a la formación de pie de montes convexos, mientras que los vuelcos y caídas constantes de rocas resultan en la formación de taludes (WSDNR, 2006a).
- 4) Avalancha de nieve: Se originan por la falla dentro o en la base de las capas de nieve en las áreas alpinas, dan como resultado un movimiento rápido de material compuesto por nieve, escombros leñosos y cantidades menores de sedimentos superficiales hacia la parte baja de la pendiente. El camino que deja la avalancha es un área alargada sin escombros de madera en las áreas alpinas y subalpinas, y en la base de la pendiente forman un depósito de rocas y escombros leñosos con forma de abanico (WSDNR, 2006a).

En lo que respecta a los deslizamientos de asentamientos profundo, estos corresponden a aquellos en los que la mayor área del plano de deslizamiento o falla se encuentra por debajo de la profundidad máxima de las raíces de los árboles, esta profundidad puede ir desde algunos metros hasta cientos (WSDNR, 2006a). Este tipo de deslizamiento incluye depósitos de glaciar, depósitos de regolita profunda, roca o lecho rocoso erosionado, así como suelo superficial pedogénico, en esta clasificación los deslizamientos de asentamiento profundo incluyen grandes fallas asociadas a materiales y estructuras geológicas (WSDNR, 2006a). A su vez estos asentamientos profundos pueden ser diferenciados de acuerdo al nivel de actividad y al mecanismo que los afecta. Por actividad: a) activo / reciente, b) inactivo-distinto, c) inactivo-indistinto y d) relicto. Por mecanismo: a) rotacional, b) traslacional, c) combinación y d) flujo de tierras.

Para la clasificación de los PG en la zona de estudio se utilizaron las categorías que muestra la Tabla 2.

Tabla 2 Clasificación de PG de acuerdo al WSDNR.

	Deslizamientos superficiales
1	De escombros
2	Flujo de escombros
3	Vuelco y caída de rocas o suelo
4	Avalancha de nieve
5	Deslizamientos profundos

Elaboración propia con información de WSDNR 2006a.

1.4. Análisis Espacial de Vecindad Continua.

Cuando se elaboran modelos sobre la probabilidad de ocurrencia de procesos espaciales aplicando la RL se hace uso de sitios de muestreo correspondientes a áreas donde el proceso en estudio se presenta y a otras en las que está ausente (Ayalew y Yamagishi, 2005; Castro, 2011; Esper, 2013; Legorreta *et al.*, 2014b). Al momento de adquirir los datos de las variables explicativas la toma de estos se realiza justo en la ubicación en donde el sitio es posicionado, es decir la información es registrada *in situ*, este procedimiento se realiza en la totalidad de los casos.

Cuando se está trabajando con variables explicativas espaciales que cubren la totalidad de la zona de estudio, al tomar solo los datos *in situ* de los sitios de muestreo se está desechando la gran mayoría de la información disponibles para la zona, además los registros obtenidos son aislados de su entorno en su totalidad con lo que se está perdiendo el principio geográfico de conexidad entre los datos correspondientes al proceso (variable dependiente) con relación a las posibles causales (variables explicativas) (Figura 4).

Con el propósito de integrar información de las variables explicativas circundante a los sitios de muestreo, se realiza un análisis espacial de vecindad de manera continua tomando como localidad del análisis de vecindad la ubicación del sitio de muestreo seleccionado para la elaboración del modelo de RL.

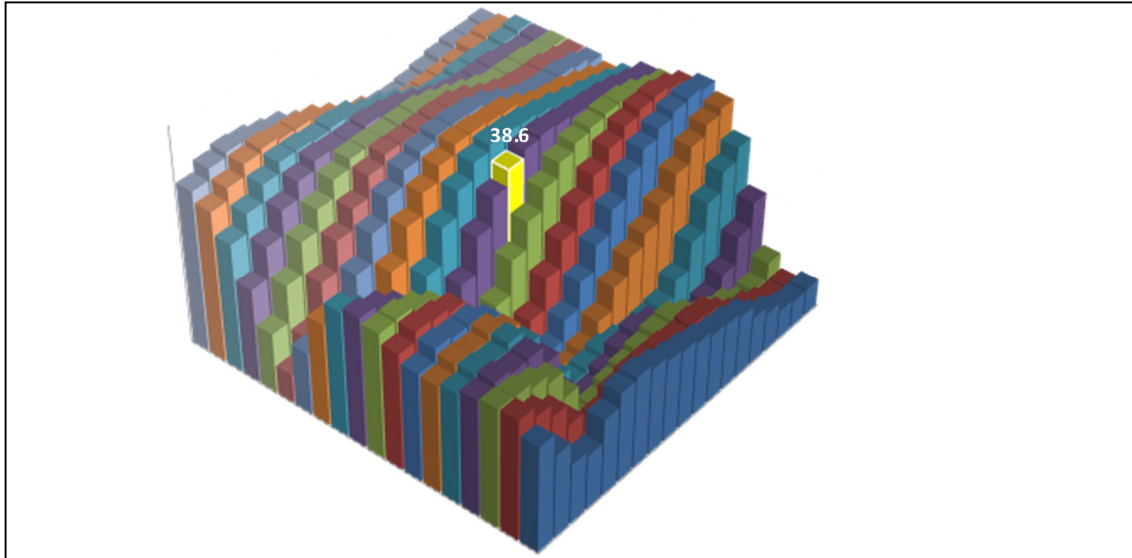


Figura 4. Toma de información *in situ* (columna amarilla). Al tratarse de una variable espacial el sitio de muestreo con un valor de 38.6 se encuentra inmerso de una gran cantidad de datos con diversos valore. Elaboración propia.

El análisis espacial de vecindad es una técnica que permite la integración de la información del entorno de un determinado sitio, este análisis se refiere a una función focal en la cual la información circundante a una localidad es cuantificada y asignada a la localidad misma (Berry, 2013b; ESRI, 2016). Para realizar un análisis de vecindad es necesario contar con: una o más localidades; definir el área alrededor de la localidad la cual será analizada; y alguna función matemática o estadística que será aplicada a los datos contenidos en el área de vecindad. La localidad hace referencia a un punto o sitio a partir del cual se especifica una zona adyacente sobre la cual se realiza una evaluación de la información contenida en esta con respecto a la característica o variable que se esté analizando (Aronoff, 1989). La zona adyacente que se analiza es llamada área de vecindad y puede tener varias formas; circular, cuadrada, rectangular, entre otras (ESRI, 2016). El análisis de los datos dentro del área de vecindad puede ser de varias formas; se puede cuantificar la variabilidad de valores, la suma de estos, el promedio en toda el área, entre los principales (Aronoff, 1989) (Figura 5). Las funciones más comunes que pueden ser aplicadas al área de vecindad se muestran en la Tabla 3.

Los datos obtenidos del área de vecindad de las localidades, las cuales en el caso de un análisis de RL estarían representadas por los sitios de muestreo, servirán para cuantificar la intensidad de la relación de las variables explicativas con el proceso en estudio más allá de la ubicación *in situ* de los sitios seleccionados, con ello se busca analizar la conexión que hay entre los sitios de muestreo con su entorno en función de cada una de las variables explicativas, partiendo del

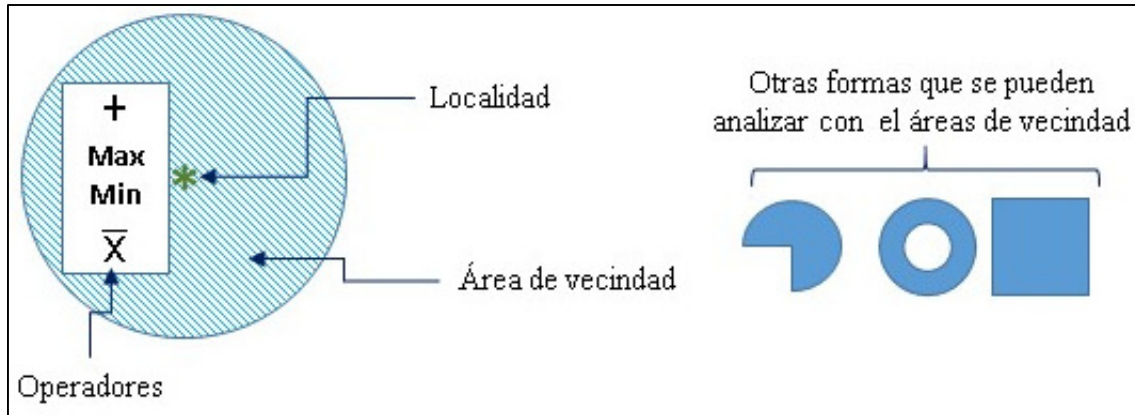


Figura 5. Configuración del análisis de vecindad. Fuente: Elaboración propia.

Tabla 3.- Principales funciones aplicables al análisis de vecindad.

Función	Descripción
Mayoría	Selecciona el valor de mayor ocurrencia en el área de vecindad.
Máximo	Reporta el valor máximo presente en el área de vecindad.
Media	Calcula el promedio del total de valores presentes en el área de análisis.
Mediana	Asigna el valor de la mediana de los valores contenidos en el área analizada.
Mínimo	Asigna el valor mínimo de los presentes en el área de vecindad.
Minoría	Reporta el valor de los datos que tiene una menor frecuencia dentro del área de análisis.
Rango	La diferencia que hay entre los valores mínimo y máximo dentro de área analizada.
Desviación	Calcula la desviación estándar de los datos del área de vecindad.
Suma	Suma los valores que presentes en el área de vecindad.
Variedad	Calcula la diversidad de valores únicos dentro del área de vecindad.

Fuente: Elaboración propia, con información de Aronoff, 1989; ESRI, 2016.

supuesto de que el área de vecindad representa una mayor integración de las características de la zona de estudio, en comparación con datos aislados como sucede cuando se considera solo el dato de la ubicación *in situ* del sitio de muestreo.

En la Figura 6a se muestran una matriz de datos de pendiente del terreno en grados en un formato *raster*, las celdas marcadas en color azul representan las localidades elegidas para realizar un análisis de vecindad, el área de vecindad 1, en color verde, y el área 2, en color rojo, muestran los datos que se toman en cuenta para hacer la evaluación estadística o matemática del área en forma circular. Entre mayor sea el tamaño del área de vecindad mayor será la cantidad de datos a considerar, lo que espacialmente representa un aumento en el área de influencia que se relaciona con el sitio. Lo anterior es un cambio importante con respecto a la toma de información para la calibración del modelo de RL, ya que si solo se toma el dato *in situ* del sitio de muestreo solo se estarían evaluando los valores 3, ignorando totalmente el entorno del sitio de muestreo y por lo tanto reflejando de manera limitada la realidad de la zona de estudio. La Figura 6b es una representación en tercera dimensión de los valores de pendiente del terreno, en esta se aprecia como a pesar de que ambos sitios contienen valores de 3, las condiciones alrededor de cada uno

de ellos son diferentes, mientras que en sitio ubicado más al sur de la zona se presenta una marcada variabilidad de valores, en el sitio al norte de la zona las condiciones son más homogéneas, estas diferencias entre las zonas son totalmente ignoradas cuando se realiza la toma de datos de manera puntual.

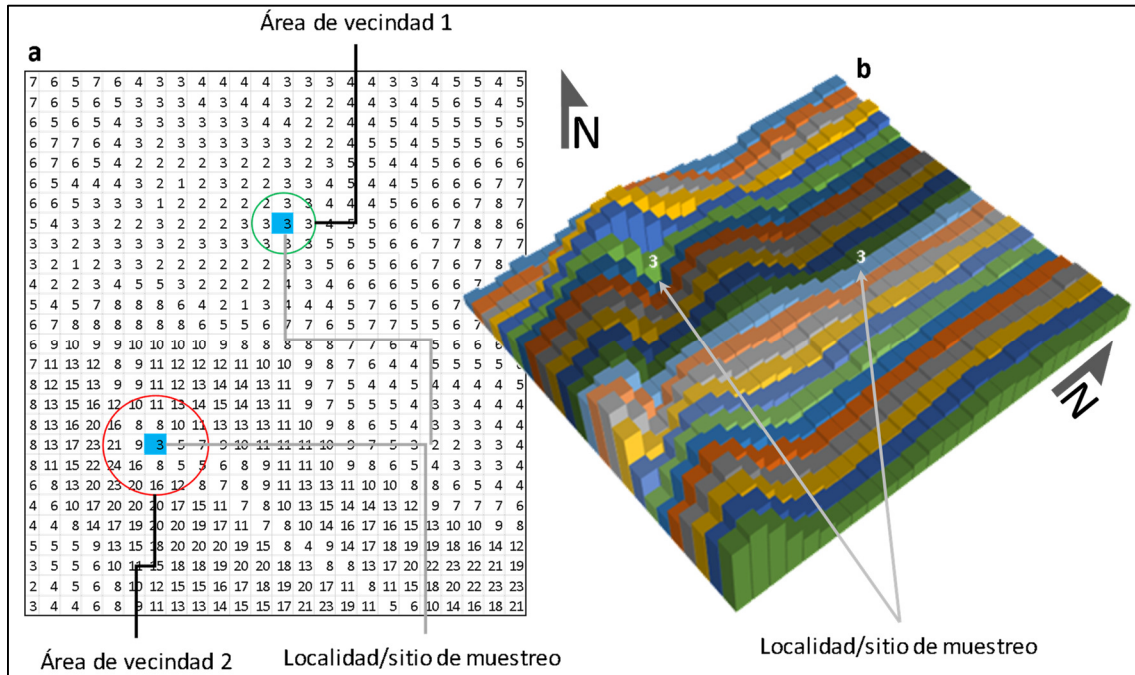


Figura 6. Integración de la información circundante a los sitios de muestreo para el modelado con RL. Elaboración propia.

Otro aspecto que se analiza es el conocer la distancia de vecindad que reporte una mayor relación con el proceso que se modela. El análisis de vecindad se elabora sobre un tamaño específico de la zona adyacente, es decir la cuantificación de los datos se realiza solo para una dimensión específica, por lo que no es posible asegurar que la distancia de vecindad analizada a un área de influencia es la que tiene una mayor relación con el proceso en estudio. Por lo tanto, se realiza el análisis de vecindad para los sitios de muestreo seleccionados de forma continua, iniciando con la lectura *in situ* del dato para posteriormente aumentar el diámetro del área de influencia de 1 en 1 celda (Figura 7).

El AEVC, se refiere a la aplicación del análisis de vecindad a partir de una distancia de un pixel de diámetro e ir aumentando de manera constante el tamaño de dicha área. Esto con el propósito de obtener los datos necesarios para realizar una evaluación estadística que cubra la vecindad al sitio de muestreo desde 1 celda hasta una distancia seleccionada de diámetro sin dejar espacio

analizado. Posteriormente los datos obtenidos son evaluados estadísticamente con la RL con el propósito de identificar cuál de las áreas cuantificadas reporta una mayor relación estadística con la probabilidad de ocurrencia del proceso que se analiza y así poder generar modelos espaciales de probabilidad de mayor precisión.

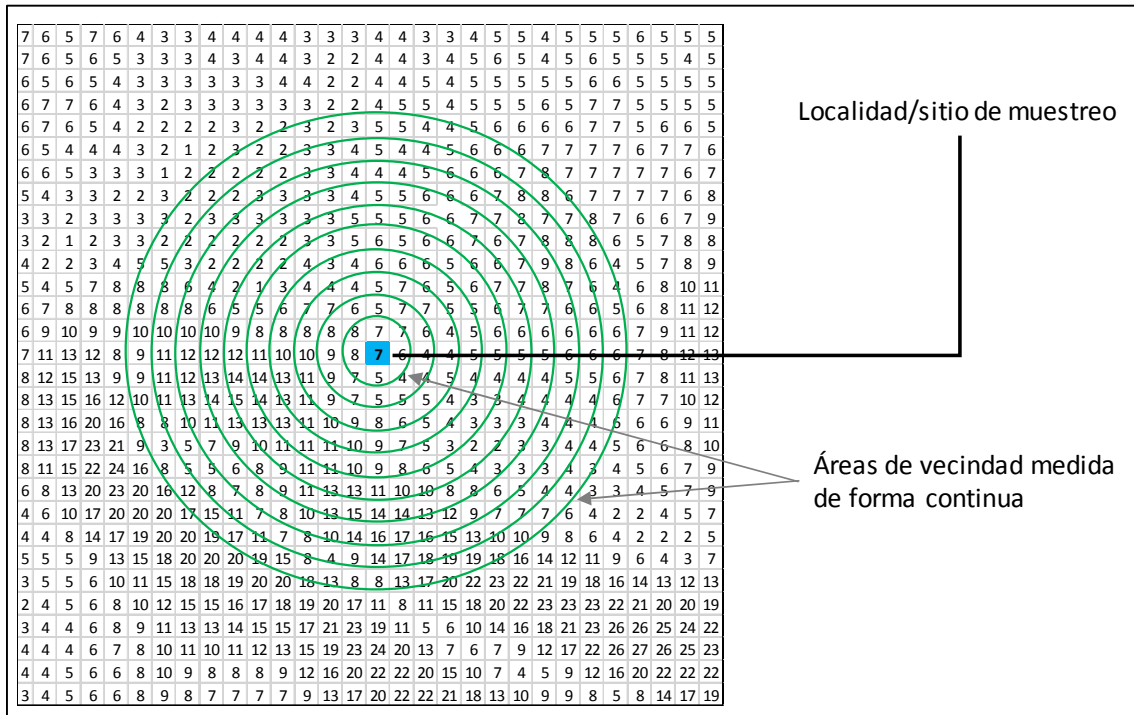


Figura 7. Descripción del AEVC, en la figura se representan las áreas de 0 a 10 píxeles de diámetro. Elaboración propia.

1.5. Muestreo.

1.5.1. Población y muestra.

La población se refiere a la totalidad de observaciones y mediciones que se pueden hacer sobre algún tema o área en particular. En estadística se habla de poblaciones finitas cuando el número de observaciones o individuos puede ser contabilizado, cuando estos no pueden ser contabilizados debido a que tienden al infinito se habla de poblaciones infinitas. Cuando se abordan problemáticas en las que el tamaño de las poblaciones sea muy grande es conveniente hacer uso de las muestras de la población (Chao, 2002; Pardo y Ruíz, 2002).

Para la especificación del muestreo en la elaboración de modelos utilizando el análisis discriminante se debe de tener en cuenta algunos aspectos importantes, entre los que se

encuentran el tipo de muestreo, el tamaño de la muestra y la división de la misma (Hair *et al.*, 1999; Pardo y Ruíz, 2002).

1.5.2. Muestreo probabilístico.

Cuando se realizan estudios sobre estadística inferencial es necesario elaborar un muestreo de tipo probabilístico, la principal cualidad de este tipo de muestreo es que todos los elementos o individuos que integran la población en estudio tienen la misma posibilidad de ser seleccionados para formar parte de la muestra, la aleatoriedad en la selección de la muestra es la que da esta cualidad en los muestreos probabilísticos (Chao, 2002; Pardo y Ruíz, 2002).

Los tipos de muestreo probabilístico más comunes son: Muestreo aleatorio simple; muestreo aleatorio sistemático; muestreo aleatorio estratificado y muestreo aleatorio por conglomerado (Chao, 2002; Pardo y Ruíz, 2002). De estos, el que mejor se ajusta a las características de este estudio y de la información utilizada es el aleatorio estratificado.

El muestreo aleatorio estratificado es utilizado cuando se tienen estratos bien definidos dentro de la población a muestrear, si alguno de los subgrupos de la población es muy pequeño con este método se puede lograr una representación mayor de este dentro de la muestra. Para ello se divide a la población en los subgrupos o estratos que la componen, a partir de los subgrupos se realiza la selección de una muestra para cada uno de estos (Chao, 2002; Pardo y Ruíz, 2002). Al generar los estratos en los que se dividirá la población lo que se busca es que los individuos dentro de un determinado estrato sean homogéneos y al mismo tiempo los estratos sean diferentes entre ellos. La suma de la selección de los subgrupos compone el total de la muestra (Chao, 2002; Pardo y Ruíz, 2002; Navarro, 2011).

En este tipo de muestreo el tamaño de la muestra seleccionada para cada estrato puede ser proporcional o no al porcentaje que este representa dentro de la población (Chao, 2002; Pardo y Ruíz, 2002). Si se conoce la varianza dentro del estrato con ello se puede decidir si el muestreo sea proporcional o no, cuando la varianza es uniforme dentro de los estratos el muestreo puede ser proporcional, pero si la varianza en alguno de los grupos es mayor se recomienda que el muestreo sea mayor a los demás en dicho grupo (Chao, 2002; Pardo y Ruíz, 2002). Otro enfoque que puede ser aplicado para decidir si la muestra se realiza de forma proporcional o no, es la importancia en el análisis que tengan los estratos o clases a muestrear, si alguno de los estratos tiene una mayor importancia dentro del estudio, utilizando el muestreo estratificado es como se le puede asignar una cantidad suficiente de sitios de muestreo para que tenga una mayor

representatividad de la que hubiera tenido si el número de sitios se hubiera determinado de manera proporcional (McGrew y Monroe, 2000).

Cuando el muestreo aleatorio estratificado se realiza utilizando información geográfica, los estratos son definidos por cada una de las clases que contenga el mapa temático que se está trabajando.

1.5.3. Tamaño de la muestra.

El tamaño de la muestra se define con base en características de la investigación y de la población que se está estudiando (Hair *et al.*, 1999; Chao, 2002; Navarro, 2011). Para calcular el tamaño necesario de la muestra se requiere de parámetros de la población a ser estudiada, entre los que se encuentran la media, el total y la proporción (McGrew y Monroe, 2000). La obtención de estos parámetros puede ser: 1) Con base a estudios anteriores; 2) Con la elaboración de un estudio piloto y 3) Para la proporción, asignando la máxima probabilidad de ocurrencia de la variable en cuestión, que en el caso de variables cualitativas binarias es del 50 %. Otro dato es el nivel de confianza que se desea obtener en los resultados, los valores más utilizados son: 99 %, 95 % y 90 %. Los valores que se utilizan en las fórmulas para cada uno de estos porcentajes corresponden al área bajo una curva normal. Por ejemplo, un 95 % tiene un valor de 1.96 (Aguilar-Barojas, 2005). La precisión absoluta es la extensión del intervalo de confianza en porcentaje, se utiliza para especificar el margen de error aceptable. Entre más pequeño sea este valor la precisión será mayor, pero será necesario una cantidad más grande de elementos en el muestreo. Los valores de la precisión absoluta corresponden a porcentajes expresados en decimales, por ejemplo, para una precisión del 5 % se utiliza el valor 0.05 (Aguilar-Barojas, 2005).

Cuando se calcula el tamaño de una muestra para variables de tipo cualitativo, en la que se trabaja la proporción del proceso sobre la población, se hace uso de las siguientes formulas. Al trabajar con poblaciones infinitas en las que no es posible conocer la totalidad del de elementos que la componen o que esta sea mayor a 10,000, se utiliza la Fórmula 1 (Aguilar-Barojas, 2005):

$$n = \frac{Z^2 pq}{d^2} \quad [1]$$

Para trabajar con poblaciones finitas, es decir se conoce el número de individuos o este es menor a 10,000 se utiliza la Fórmula 2 (Aguilar-Barojas, 2005):

$$n = \frac{N Z^2 pq}{d^2 (N-1) + Z^2 pq} \quad [2]$$

Donde:

n = tamaño de la muestra

Z = valor calculado en tablas según el nivel de confianza que se requiera

N = tamaño de la población

p = proporción aproximada del proceso en estudio en la población de referencia

q = proporción aproximada de la población de referencia que no presenta el proceso en estudio. La suma de $p + q$ siempre debe resultar en 1

d = precisión absoluta

A demás del cálculo para el tamaño de muestra anterior, cuando se aplican técnicas de análisis discriminante se recomienda utilizar 20 observaciones por cada variable explicativa analizada, si no fuera posible alcanzar este tamaño de muestra el mínimo recomendado es de 5 muestras por cada variable (Hair *et al.*, 1999). En lo que respecta al número de sitios con relación a los grupos o clases de la variable dependiente se recomienda que cuando menos sean un total de 20 observaciones por grupo (Hair *et al.*, 1999).

1.5.4. División de la muestra.

La división de la muestra tiene como objetivo el contar con dos conjuntos de sitios de muestreo que permitan realizar la calibración del modelo estadístico (muestra de análisis) y la evaluación del nivel predictivo del modelo obtenido (ampliación de la muestra), con lo que se evita una sobre evaluación del modelo, lo cual sucedería si los sitios que fueron utilizados para la elaboración del modelo discriminante fueran utilizados para la evaluación del mismo (Hair *et al.*, 1999). Algunos de los porcentajes recomendados para realizar la división de la muestra mencionan que esta puede ser un 50% de los sitios para muestra de análisis como para la ampliación de la muestra, también se manejan otros porcentajes como son el 60 % o 75 % para la muestra de análisis y el 40 % o 25 % para la muestra de validación, respectivamente. En caso que no se pudiera obtener un tamaño de muestra que permitiera contar con el número suficiente de sitios para realizar la división de la misma, es decir que solo se contara con el número mínimo para la elaboración de la función discriminante, se recomienda utilizar los sitios tanto para el cálculo de la función como para su evaluación, sin perder en cuenta que esto dará como resultado una sobre evaluación de los resultados, pero se recomienda hacer la evaluación del modelo (Hair *et al.*, 1999).

1.6. Modelo estadístico de Regresión Logística Múltiple.

La estadística se divide en dos grandes áreas: una de estas es la estadística descriptiva que se ocupa de la adquisición, organización, presentación y descripción de datos. La segunda es la estadística inferencial, a través de la cual se realizan generalizaciones utilizando solo una parte de la población que se está analizando (Chao, 2002). La estadística inferencial incluye una gran variedad de técnicas entre las que se encuentra el análisis discriminante el cual tiene como objetivo la diferenciación de clases o grupos con base en un conjunto de variables, dentro de este tipo de análisis se encuentra la técnica de la RL (Hair *et al.*, 1999).

La RL tiene el propósito de calcular la probabilidad de pertenencia de un individuo a un determinado grupo a partir de la evaluación de variables explicativas (Hair *et al.*, 1999). Con este modelo estadístico se calcula la probabilidad de ocurrencia del proceso en estudio, los resultados obtenidos son valores continuos entre 0 y 1 pero el resultado final debe de ser clasificado a solo los valores 0 o 1, que corresponden a la ausencia (0) o presencia (1) del proceso (Pardo y Ruíz, 2002). Con el propósito de realizar la clasificación de los valores continuos obtenidos de la RL a 0 o 1 se establece un punto de corte que corresponde a un número a partir del cual los valores menores a este se clasifican como 0 y los mayores como 1. Una forma de determinar un punto de corte óptimo con el que se alcancen a diferenciar al máximo ambos grupos (valores 0 y 1) es generando tablas de clasificación de la función de RL en las que se vaya variando el punto de corte hasta alcanzar el máximo de aciertos en ambos grupos (Pardo y Ruíz, 2002).

El modelo de RL es ideal para trabajar con procesos en los que la variable dependiente puede tomar dos posibles valores, que son excluyentes uno del otro (Wayne, 2004). Tal es el caso de los modelados aplicados a PG en los que se busca conocer la probabilidad de distribución de dichos procesos en una zona determina con base a la evaluación de una serie de variables explicativas en donde la variable dependiente toma los valores de áreas estables e inestables.

En el modelo de RL se pueden trabajar variables explicativas cualitativas, estas son analizadas en función de su ausencia o presencia, por lo que son codificadas con los valores 0 y 1, respectivamente. Estas son codificadas en función del número de clases que esta contiene menos 1 y como resultado se obtiene un coeficiente para cada una de las clases que contiene la variable, a excepción de la categoría que fue tomada como referencia, codificada con valor 0 (Pardo y Ruíz, 2002).

El modelo de RL puede ser aplicado con una sola variable explicativas o con 2 o más. A continuación, se presenta la ecuación de la RL para una sola variable explicativas (Fórmula 3) (Pardo y Ruíz, 2002).

$$Y = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1)}} \quad [3]$$

Cuando se tienen dos o más variables explicativas la ecuación toma la siguiente forma (Fórmula 4):

$$Y = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}} \quad [4]$$

Donde:

Y es la probabilidad de que se presente el proceso.

X_1 a X_k son las variables explicativas que forma parte del modelo.

β_0 es el coeficiente de la constante en el modelo.

β_1 a β_k son los coeficientes de las variables explicativas.

La RL utiliza la función de máxima verosimilitud transformándola con el logaritmo neperiano, esto limita los resultados a valores entre 0 y 1 con el propósito de calcular la pertenencia de la variable dependiente a alguno de los dos grupos posibles. Los valores cercanos a 0 representan poca probabilidad de ocurrencia, mientras que los valores cercanos a 1 representan una alta probabilidad (Hair *et al.*, 1999; Pardo y Ruiz, 2002).

Los coeficientes estimados en la RL corresponden a medidas de los cambios de la razón de probabilidad expresados en logaritmo. Los coeficientes con valores positivos indican un aumento en la probabilidad de ocurrencia, mientras que los coeficientes con signo negativo denotan una disminución (Pardo y Ruiz, 2002).

Una de las opciones con la que puede ser aplicada la RL es hacia delante razón de la verosimilitud. Esta forma de ejecutar la RL implica el comenzar la evaluación con un modelo vacío, sin introducir alguna de las variables explicativas, y progresivamente ir integrando de una en una las variables al modelo y evaluar si esta es significativa, si la variable recién integrada fuera significativa se mantiene en el modelo y se procede a ingresar la siguiente variable, las variables que se han mantenido en el modelo son evaluadas en conjunto para saber si la adición de la nueva variable afecta la significancia de alguna de las que ya forman parte del mismo, si alguna de las variables que ya estaban en el modelo o si la variable recién ingresada resulta no ser significativa se procede a eliminarla, así sucesivamente se van ingresando y evaluando todas

las variables. Al final se obtiene un modelo que solo está integrado por las variables que son significativas (Pardo y Ruíz, 2002).

1.6.1. Supuestos de la Regresión Logística.

Para poder llevar a cabo un análisis de regresión es necesario tener en cuenta el cumplimiento de un conjunto de supuestos o condiciones para evitar que los resultados obtenidos presenten deficiencias. Estos supuestos son: normalidad multivariante, igualdad de matrices de varianza y covarianza, y multicolinealidad.

En lo que respecta al supuesto de normalidad multivariante, este supuesto resulta nulo para la RL debido a la distribución de la variable dependiente. Esta variable tiene dos posibles resultados y por consiguiente posee una distribución binomial con lo que el supuesto de normalidad es nulo, por otra parte, la igualdad de las matrices de varianza y covarianza no se cumple ya que la distribución de la varianza de una variable dicotómica no es constante. El incumplimiento de tales supuestos no representa problema alguno al trabajar con la RL e incluso los resultados suelen ser más sólidos cuando estos no se cumplen (Hair *et al.*, 1999).

El supuesto de multicolinealidad entre las variables explicativas se refiere a la colinealidad que pudiera darse entre estas. Lo ideal en la elaboración de modelos de regresión es que las variables explicativas estén altamente relacionadas con la variable dependiente y que la relación entre las primeras sea mínima o nula. Cuando existe una relación entre variables explicativas es difícil identificar la influencia que tienen sobre el proceso en estudio variables que se vayan incorporando al análisis, ya que dicho efecto o contribución al modelo se realiza de manera combinada para las variables que presentan una alta relación (Hair *et al.*, 1999). Conforme aumenta la colinealidad entre las variables explicativas, se da una relación inversamente proporcional con el poder predictivo de las variables explicativas. Cuando existe una colinealidad perfecta entre dos variables explicativas se le llama singularidad, si esta singularidad se presenta no es posible calcular los coeficientes de la regresión, aun si la correlación entre las variables explicativas no es exacta, pero sí muy elevada la estimación de los coeficientes es incorrecta, es por ello que se debe de identificar si hay colinealidad entre las variables analizadas y si es que la hay se debe de aplicar algún método para evitarla (Hair *et al.*, 1999).

Para la identificación de colinealidad entre variables, los métodos más comunes utilizados son el cálculo del valor de Tolerancia y el Factor de Inflación de la Varianza (FIV). Estas medidas evalúan como cada variable explicativa es afectada por alguna de las otras variables del mismo

tipo. En términos simples lo que se hace con estos métodos es tomar una variable explicativa y tratarla como si fuera la variable dependiente, mientras que las demás variables son trabajadas como explicativas. Con ello se calcula como la variable tomada inicialmente como dependiente es explicada por las demás, esto se hace con cada una de las variables independientes hasta completar el análisis de todas ellas (Hair *et al.*, 1999). La Tolerancia se refiere a la cuantificación de la variabilidad de la variable explicativa seleccionada que no es explicada por las otras variables independientes, es decir que si se tiene un valor pequeño en la evaluación de la tolerancia la relación entre las variables explicativas es alta ya que la mayoría de la variabilidad en la variable independiente analizada está relacionada con las otras variables explicativas. En lo que respecta al FIV, este es el valor inverso de la Tolerancia, por lo que valores altos en este último denotan una alta colinealidad entre variables (Hair *et al.*, 1999). Los valores de la Tolerancia que podrían estar exponiendo colinealidad entre las variables explicativas serían aquellos menores a 0.10 y esto representa valores por arriba de 10 para el FIV (Hair *et al.*, 1999).

Otros métodos para identificar la existencia de colinealidad son: el Índice de Condicionamiento, el cual muestra la correlación entre variables al combinarlas en una matriz de datos; y la Matriz de Proporción de la Varianza del Coeficiente de Regresión. Estos dos indicadores se analizan en conjunto (Hair *et al.*, 1999). Se inicia con la identificación de los valores del índice de condicionamiento mayores a 30, que es valor que generalmente es utilizado, aunque también se utiliza el valor de 15. Una vez identificados los valores con índice de condicionamiento mayores al umbral seleccionado se procede a identificar valores mayores al 0.9 de la proporción de la varianza del coeficiente, si se da esta combinación en dos o más coeficientes significa que las variables asociadas a estos presentan colinealidad. Se debe de cumplir con las dos condiciones para que se pueda asumir colinealidad, si solo se sobrepasa uno de los dos umbrales (condicionamiento o proporción de la varianza) no habría colinealidad (Hair *et al.*, 1999).

En caso de que se llegue a identificar colinealidad entre variables se procede de la siguiente manera (Hair *et al.*, 1999):

- Eliminar algunas de las variables que reporten correlación, si es necesario se deben de buscar otras variables que pudieran aportar algún nivel de separación en las clases de la variable dependiente.
- Se podría utilizar el modelo resultante con las variables que reportan colinealidad, pero este solo podría ser utilizada con propósitos de predicción, sin darle alguna interpretación a los valores de los coeficientes.

1.6.2. Valoración del ajuste del modelo.

Para la evaluación de los modelos de RL se utilizan indicadores con los que se cuantifica si las variables explicativas contribuyen en el comportamiento de la variable dependiente y otros con los que se mide el mejorar el ajuste que vaya reportando el modelo. El indicador que permite conocer si la introducción de alguna variable en el modelo de RL aporta información para agrupar de mejor manera las categorías de la variable dependiente es el referente a la significancia (Pardo y Ruiz, 2002). Este indicador hace referencia a la hipótesis estadística, que está compuesta de la hipótesis nula (H_0) la cual se refiere a la afirmación de que la introducción de una variable al modelo no contribuye en explicar el comportamiento de la variable dependiente, por otra parte, está la hipótesis alterna (H_1) la cual se refiere a que al introducir una variable al modelo esta contribuye en la explicación de la variable dependiente y por lo tanto en la mejora del modelo (Chao, 2002; Pardo y Ruiz, 2002; Hernández *et al.*, 2014). Cuando se trabaja con las hipótesis estadísticas se corre el riesgo de rechazar una H_0 cuando esta es verdadera y aceptar la H_1 , a este se le conoce como error de Tipo I. El riesgo de cometer este error es señalado por el valor de la significancia de las variables introducida en el modelo de RL, este representa el porcentaje del riesgo en valores decimales, el valor utilizado de manera general es del 0.05 lo que significa que se tiene un 5 % de probabilidad de cometer un error de Tipo I, es decir se tiene una confianza del 95 % de haber hecho la decisión correcta al rechazar la H_0 . Cuando se utiliza el valor de 0.05 para evaluar la significancia al introducir una variable al modelo, si el valor reportado por la variable es mayor a 0.05 se acepta la H_0 lo que significa que la variable no aporta información que permita una mejora en el modelo (Chao, 2002; Pardo y Ruiz, 2002; Spiegel y Stephens, 2009).

Para valorar el de ajuste global del modelo se analiza el indicador de -2 veces el logaritmo de la verosimilitud (-2LL). Este indicador evalúa el nivel de probabilidad de los resultados en función del modelo obtenido. Este se calcula contrastando el número de casos observados con los casos esperados al aplicar el modelo (Martín *et al.*, 2008). Cuando un modelo de RL tiene un buen ajuste reporta un valor pequeño para -2LL, siendo 0 el valor mínimo que este puede alcanzar. Una cualidad que tiene este indicador es que puede ser comparado entre dos modelos, en donde la diferencia de los valores reportados para -2LL representa el cambio en el ajuste predictivo de un modelo a otro (Hair *et al.*, 1999).

El contraste de chi-cuadrada de ómnibus para la reducción del logaritmo de la función de verosimilitud representa una medida de mejora debido a la introducción o eliminación de alguna otra variable explicativa, cuanto mayor sea el valor de chi-cuadrada refleja un mejor ajuste del

modelo conforme se va realizando la selección de las variables que estarán presentes en el modelo. El ajuste de este indicador se puede apreciar cuando para la elaboración del modelo de RL se utiliza la introducción de variables hacia delante o a hacia atrás, sin embargo, cuando se utiliza el método por bloque (todas las variables son incluidas en un solo paso) este indicador no puede ser evaluado (Hair *et al.*, 1999; Pardo y Ruiz, 2002).

La matriz de clasificación producto de la RL permite evaluar el poder predictivo del modelo en función de cada uno de los grupos de la variable dependiente. De esta matriz se puede obtener un porcentaje de correcta predicción para cada una de las clases de la variable dependiente, así como el porcentaje global de clasificación de todo el modelo (Hair *et al.*, 1999; Pardo y Ruiz, 2002).

El contraste de clasificación desarrollado por Hosmer y Lemeshow en 1980, parte de la división de los casos analizados (muestra) en 10 grupos con un número casos iguales en cada una de ellos, posteriormente se comparan el número eventos verdaderos y predichos, lo que da una medida general de la exactitud de la predicción de la variable dependiente del modelo. Para la elaboración de este contraste se requiere un tamaño de muestra que asegure un mínimo de 5 casos por grupo (Hair *et al.*, 1999).

1.7. Evaluación del modelo de Regresión Logística.

La curva Característica de Operación del Receptor y la elaboración de matrices de clasificación, también conocidas como matrices de confusión son dos métodos utilizados para la evaluación de modelos producto de funciones discriminantes como la RL.

1.7.1. Curva Característica de Operación del Receptor.

El análisis de la curva Característica de Operación del Receptor (COR) es utilizado para evaluar cómo responde una función discriminante a la clasificación de dos grupos. Esto se realiza con la generación de un gráfico y el cálculo del área bajo la curva producto del mismo (Pardo y Ruiz, 2002). Para poder realizar el análisis con la curva COR es necesario contar con la información sobre la probabilidad de pertenencia a alguna de las dos clases analizadas (modelo RL) y la pertenencia real de los sitios que serán utilizados en el análisis (información verdadera) (Pardo y Ruiz, 2002).

La información presentada en el gráfico del análisis de la curva COR, contiene los datos referentes a falsos positivos, por lo general nombrado como $1 - \text{Especificidad}$, los cuales corresponden a los errores entre el grupo real y el pronosticado de la clase codificada como 0. Así como a los aciertos positivos, generalmente etiquetado como *Sensibilidad* y que corresponden a los aciertos entre el grupo real y el pronosticado de la clase codificada con 1 (Pardo y Ruiz, 2002) (Tabla 4).

Tabla 4. Codificación de los falsos positivos y aciertos positivos utilizados en la curva COR.

		Grupo pronosticado	
		Zonas estables (0)	Cabeceras (1)
Grupo real	Zonas estables (0)	Aciertos negativos (Especificidad)	Falsos positivos ($1 - \text{Especificidad}$)
	Cabeceras (1)	Falsos negativos ($1 - \text{Sensibilidad}$)	Aciertos positivos (Sensibilidad)

Elaboración propia con base en Pardo y Ruíz 2002.

La Figura 8 muestra una gráfica de curva COR, en esta la línea azul muestra los valores de $1 - \text{Especificidad}$ y *Sensibilidad* para cada punto que contiene la información real de pertenencia a un determinado grupo y el valor obtenido en la RL (valor de probabilidad). Si la curva de la línea es más cercana a la esquina superior izquierda significa que los resultados de la RL son mejores, una curva COR ideal es aquella que tenga una *Sensibilidad* de 1, es decir que reporte el 100 % de aciertos positivos, por otro lado, si la curva resultante en la gráfica fuera una línea diagonal (color verde) se interpreta como el peor resultado ya que significa que al aplicar la RL las probabilidades obtenidas representan el 50 % de pertenencia a uno u otro grupo (Pardo y Ruíz, 2002).

El análisis de la curva COR proporciona el área debajo de la curva que representa el nivel de correcta clasificación de los valores resultado de la RL. También proporciona información sobre la significancia que resulta al contrastar la hipótesis nula, la cual dice que el área bajo la curva es de 0.5, este valor de significancia debe de ser igual o menor a 0.05 para poder rechazar la hipótesis nula y asumir que el área bajo la curva es mayor a 0.5 (Pardo y Ruíz, 2002).

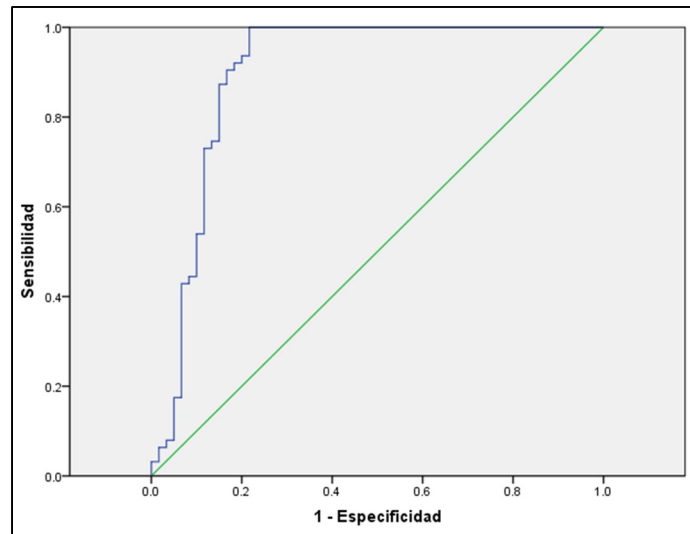


Figura 8. Grafica de una curva COR.

1.7.2. Matriz de clasificación.

Esta matriz es utilizada para evaluar la correcta capacidad de clasificación de alguna función discriminante, como es el caso de la RL (Hair *et al.*, 1999). Es generada a partir de la comparación de información verdadera con el modelo de probabilidad resultado de la RL.

Para la elaboración de la matriz de clasificación se hace uso de la muestra ampliada y su correspondencia con la información obtenida por la RL. En las filas se ubican los datos de la información real, mientras que en las columnas se ubican los datos pronosticados. Los datos que se presentan en la diagonal (color azul) corresponden a los individuos que han sido correctamente clasificados, los que se encuentran fuera de la diagonal representan datos erróneamente clasificados (Hair *et al.*, 1999) (Figura 9).

De la matriz de clasificación se obtienen indicadores como: Índice de Kappa indica el grado de acuerdo entre grupo real y grupo pronosticado. El índice se calcula tomando en cuenta precisión total del modelo junto con el número de casos de clasificación correcta que se dan por azar, estos últimos se calculan multiplicando las frecuencias marginales (total renglones y columnas), divididas entre el total de casos (tamaño de muestra) (Fórmula 5) (Pardo y Ruiz, 2002). De manera general el valor de este índice se interpreta como muy bueno si es arriba de 0.75, entre 0.40 y 0.75 como bueno y un valor menor a 0.40 como malo (Pardo y Ruiz, 2002).

Fórmula para el cálculo del Índice de Kappa:

$$k = \frac{n \sum_i n_{ii} - \sum_i (n_{i+} * n_{+i})}{n^2 \sum_i (n_{i+} * n_{+i})} \quad [5]$$

Donde:

- n = tamaño de la muestra
- n_{ii} = individuos en la diagonal
- n_{i+} = total de individuos en el renglón
- n_{+i} = total de individuos en la columna

Exactitud general, esta se refiere a la razón entre el número de elementos clasificados correctamente (diagonal) y el número total de elementos analizados en la matriz (Pardo y Ruiz, 2002; Nemmaoui *et al.*, 2013) y representa el porcentaje de elementos correctamente clasificados (Figura 9).

		Grupo pronosticado	
		Zonas estables (0)	Cabeceras (1)
Grupo real	Zonas estables (0)	(0,0)	(0,1)
	Cabeceras (1)	(1,0)	(1,1)

Exactitud general = $(0,0) + (1,1) / (0,0) + (0,1) + (1,0) + (1,1)$
Exactitud del productor = $(1,1) / (1,0) + (1,1)$
Exactitud del usuario = $(1,1) / (0,1) + (1,1)$
Eficiencia del modelo = $(1,1) - (0,1) / (1,0) + (1,1)$

Figura 9. Cálculo de las exactitudes y eficiencia del modelo para la clase de cabeceras (1). Elaboración propia.

Exactitud del productor se refiere a la razón de los individuos correctamente clasificados en una determinada clase y el total de individuos verdaderos en esa misma clase (Legorreta *et al.*, 2012a) (Figura 9). Es el porcentaje de correcta clasificación de elementos del grupo real para una determinada clase.

Exactitud del usuario este corresponde a la razón entre individuos correctamente clasificados en una determinada clase y el total de individuos clasificados por el modelo en esa misma categoría (Legorreta *et al.*, 2012a) (Figura 9). Indica el porcentaje elementos correctamente clasificados del pronóstico de una clase.

Eficiencia del modelo es la razón de individuos correctamente clasificados en una determinada clase menos los individuos incorrectamente pronosticados de la misma clase entre el número total de individuos del grupo real (Figura 9), el resultado tiene un valor de 1 cuando hay una clasificación perfecta de los individuos y reporta valores negativos cuando el número de elementos incorrectamente clasificados es mayor al correctamente pronosticados (Legorreta *et al.*, 2012a)

CAPÍTULO 2. ÁREA DE ESTUDIO.

La zona de estudio es la cuenca La Ciénega localizada en la ladera este del volcán Nevado de Toluca (VNT). El volcán se ubica en las coordenadas geográficas 99°45' W y 19°09' N, a una distancia de 23 km al suroeste de la ciudad de Toluca, capital del estado de México (Macías, 2005). Su altura es de 4,680 m s.n.m., siendo la cuarta elevación más altas de México (García-Palomo *et al.*, 2000; Norini *et al.*, 2004). El VNT es un estrato volcán andesítico-dacítico que forma parte del Cinturón Volcánico Trans-Mexicano (CVTM), el Cinturón cruza el centro de México de este a oeste y tiene una extensión de 1200 km., (Capra y Macías, 2000; García-Palomo *et al.*, 2000) (Figura 10).



Figura 10.- Localización del volcán Nevado de Toluca dentro del CVTM. Elaboración propia con base en Macías, 2005.

Sobre la geología y geomorfología del volcán Nevado de Toluca (VNT) se han realizado estudios que sirven de base para conocer las características generales de la zona de estudio. Geológicamente el VNT es un estratovolcán andesítico-dacítico de las épocas Plioceno tardío-Holoceno. Se han identificado 17 eventos volcánicos que inician hace unos 2.6 millones de años (+/- 200,000 años) con flujos de lava, flujos de escombros, flujos de depósitos lacustres, avalanchas y flujos de escombros, flujos de pómez rosa y ocre, flujo de bloques y cenizas alternados con caídas de pómez. Los 4 estratos más recientes, de hace unos 10,500 años

produjeron la caída de pómez blanca, el flujo de pómez llamado Toluca superior, el flujo de lava Tenango y por último un flujo de ceniza de hace aproximadamente 3,300 años (García-Palomo *et al.*, 2002). El VNT está constituido por una variedad de materiales los cuales han sido modificados por diversos procesos de físicos. La evolución del volcán se ha dado entre etapas de construcción y destrucción, como son erupciones, colapsos en el edificio volcánico, flujos y depósitos de material piroclástico, cenizas y pómez, principalmente (Espinosa-Rodríguez *et al.*, 2014).

En particular, la geomorfología de la zona noreste del piedemonte del VNT se desarrolla en función de sus características litológicas, de relieve, climáticas y actividades antrópicas, dando origen a una zona formada por materiales de baja compactación como son las pómez y la toba, que en conjunto con las demás variables producen condiciones que generan una dinámica que favorece los PG los cuales afectan las actividades antrópicas que se desarrollan en la zona y que se caracterizan por actividades agrícolas, pecuarias, forestales y extracción de material para la construcción (Espinosa-Rodríguez, 2003). La geomorfología de la ladera este del VNT obedece principalmente a sus características físicas como son: laderas de pómez, abanicos aluviales y pendientes del terreno (entre los 5 y 15 grados), y a actividades antrópicas como: la agricultura y los pastizales (Espinosa-Rodríguez, 2003). Está conformada por laderas bajas acumulativas producto de depósitos de pómez con una alta disección fluvial, valles y circos de origen glaciar que presentan disección fluvial de baja a media, y por laderas convexas asimétricas formadas por piroclastos con presencia de disección fluvial de moderada a intensa (Espinosa-Rodríguez *et al.*, 2014; García, 2017). Tiene una pendiente suave, pero es la ladera con mayor presencia de barrancos con algunos que llegan a medir hasta 70 m de profundidad, como es el cauce del río Arroyo Grande (Capra *et al.*, 2008).

El trabajo se realizó en la cuenca endorreica del río La Ciénega, esta se ubica en el suroeste de la subcuenca Almoloya-Otzolotepec que pertenece a la cuenca Lerma-Toluca en la región hidrográfica Lerma-Santiago (Aceves *et al.*, 2014b; INEGI, 2010) (Figura 11). La mayoría de su superficie pertenece al municipio de Tenango del Valle, la zona más occidental de La Ciénega se encuentra dentro del municipio de Toluca y solo una pequeña área cubre parte del municipio de Calimaya (INEGI, 2016).

La cuenca La Ciénega se localiza entre las coordenadas extremas: 19°7'10.7" y 19°4'21.8" latitud norte y, 99°37'28.7" y 99°44'40.7" longitud oeste. Tiene una superficie de 3,007 hectáreas y presenta una orientación de oeste a este, con una longitud de 13 km y un ancho promedio de 2.5 km (Aceves *et al.*, 2014a).

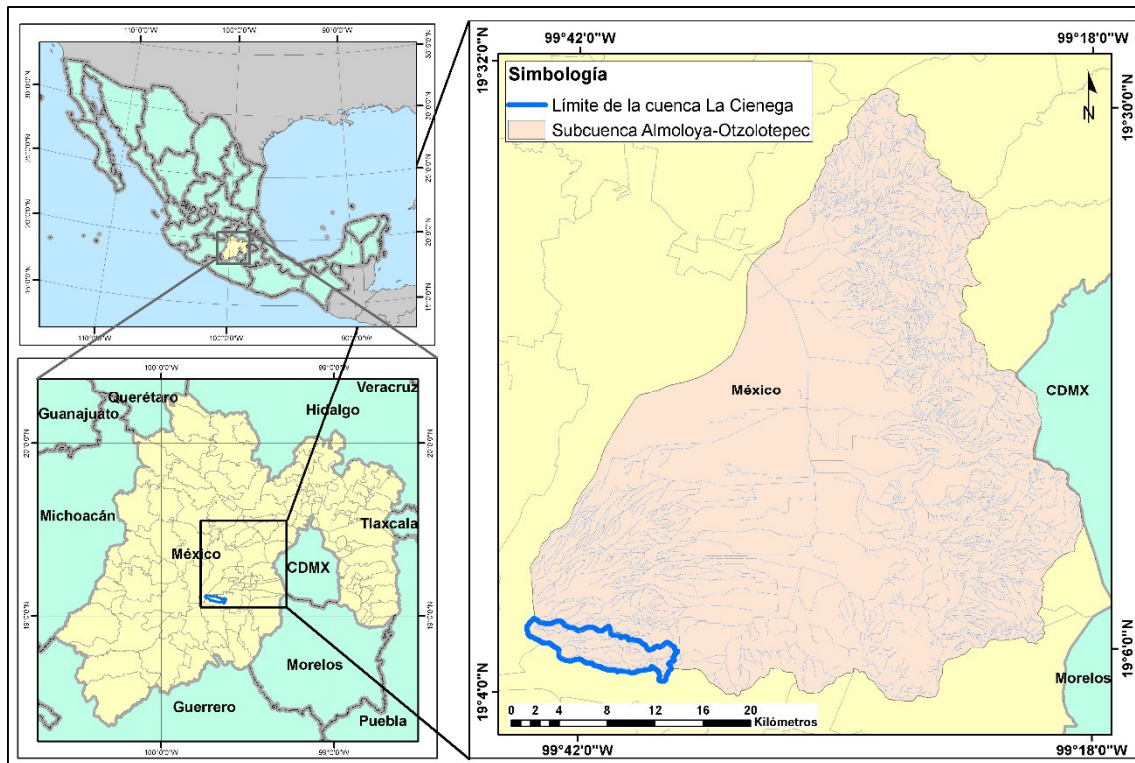


Figura 11.- Ubicación de la cuenca La Ciénega, en el estado de México. Elaboración propia con información del INEGI (2010; 2016).

Dentro de la cuenca se encuentran los poblados de San Juan Tepehuixco y Santa Cruz Pueblo Nuevo, con una población de 43 y 1,634 habitantes, respectivamente de acuerdo con el Censo de Población y Vivienda del año 2010 (INEGI, 2013b). La localidad de Santa Cruz Pueblo Nuevo, se ubicada en la parte baja del río La Ciénega, sobre un abanico aluvial activo (Figura 12). Por las características geomorfológicas de la cuenca aunado a los regímenes de precipitación en la zona el poblado se encuentra en constante peligro ya que existe la posibilidad de que ocurran nuevamente flujos de escombros como el acontecido en 1940 en el antiguo pueblo de Santa Cruz (Aceves *et al.*, 2014b; Toscana y Valdez, 2014; García *et al.*, 2018). Y el ocurrido en 1952 a causa del material producto de PG no consolidado que se encuentra principalmente en los cause. Se trata de una zona en la que los flujos de escombros ocurren cada año en la época de lluvias (Capra *et al.*, 2008).

En la cuenca se encuentran altitudes que van de los 2,807 m s.n.m. en su lado oriental, hasta los 4,423 m s.n.m. en el extremo occidental (INEGI, 2013a) dando una diferencia altitudinal de 1,616 m en sus 13 km de longitud este-oeste (Figura 13a). La disminución de la altitud se va dando de manera uniforme de oeste a este, siendo interrumpida por la presencia del cerro Tepehuixco en

la parte central de la cuenca con una altura mayor a los 3,400 m s.n.m. En la Figura 13b se muestra la ubicación del poblado de Santa Cruz Pueblo Nuevo en un perfil altitudinal.

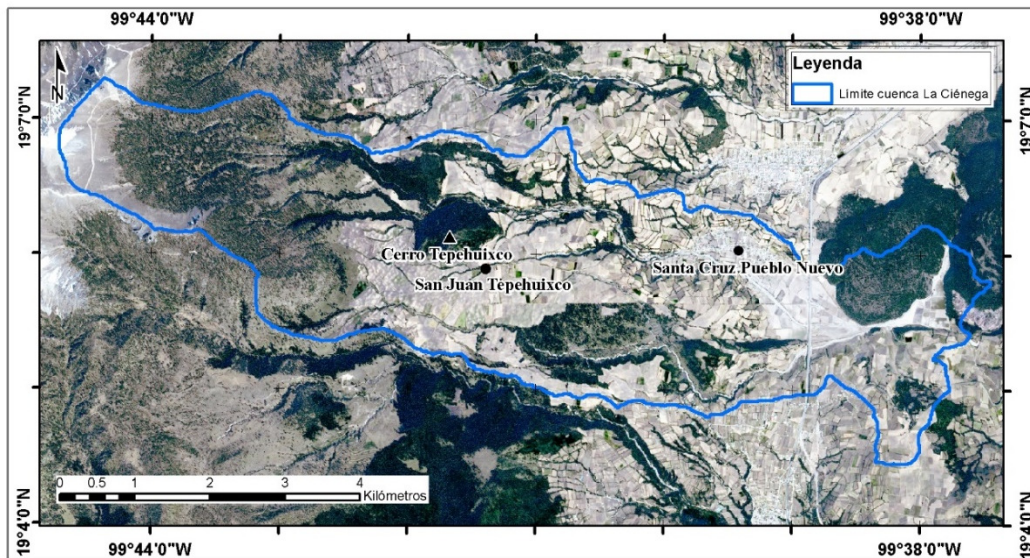


Figura 12. Principales poblaciones en la cuenca La Ciénega. Elaboración propia con información del INEGI, 2016; SIAP-SEDENA, 2016.

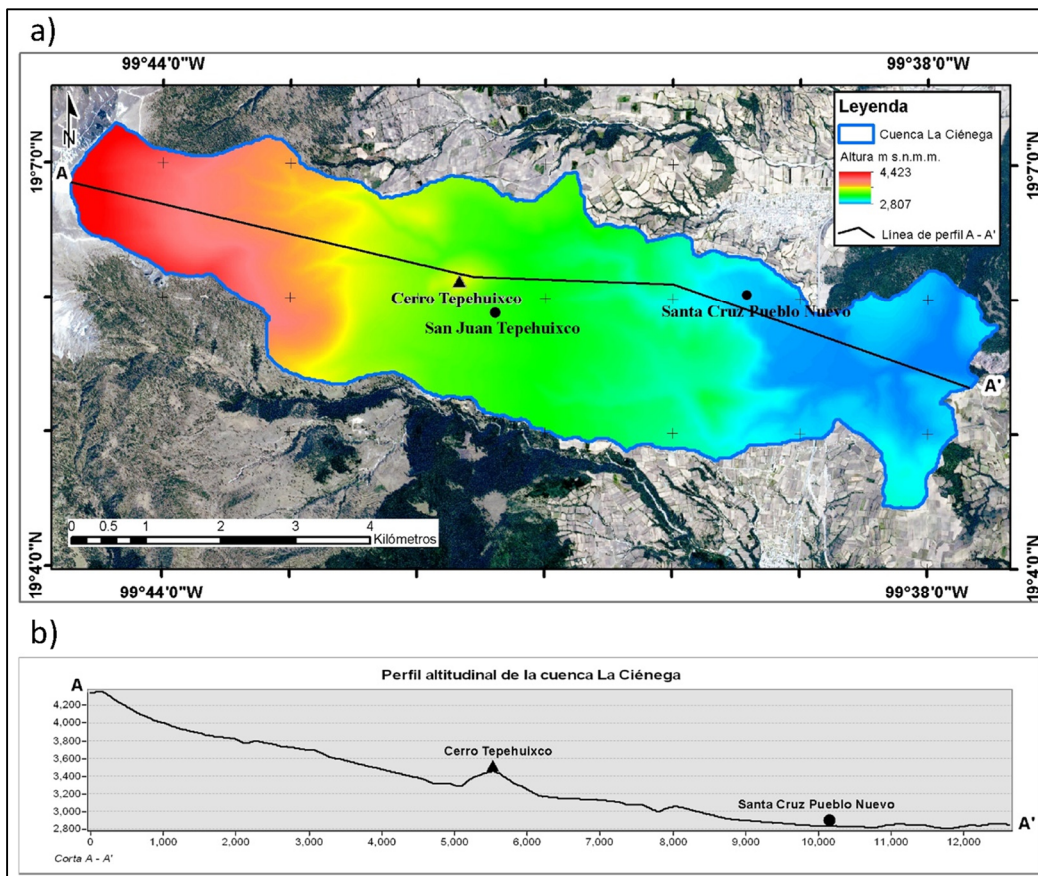


Figura 13. a) valores altitudinales en la cuenca y b) perfil altitudinal este – oeste de la zona de estudio. Elaboración propia con información de INEGI, 2016 y SIAP-SEDENA, 2016.

La pendiente del terreno en el área de estudio registra valores entre los 0° y los 48°, la mayoría de las zonas con pendientes fuertes se encuentran distribuidas en la región oeste. De manera general al este de la cuenca se encuentran distribuidas áreas con pendiente de menor gradiente, la concentración de pendientes de mayor gradiente se da en el cerro de Tepehuixco, en la zona centro. El área del abanico aluvial en la que se encuentra asentado el poblado de Santa Cruz Pueblo Nuevo registra pendientes bajas (Figura 14). En áreas con valores elevados de la pendiente del terreno la erosión hídrica y los PG suelen ser más numerosos e intensos, lo cual aunado a que el poblado se encuentra en una de las zonas de menor altitud dentro de la cuenca potencializa el peligro a la acumulación de los depósitos de los flujos de escombros en esta zona baja y plana (Aceves *et al.*, 2014b).

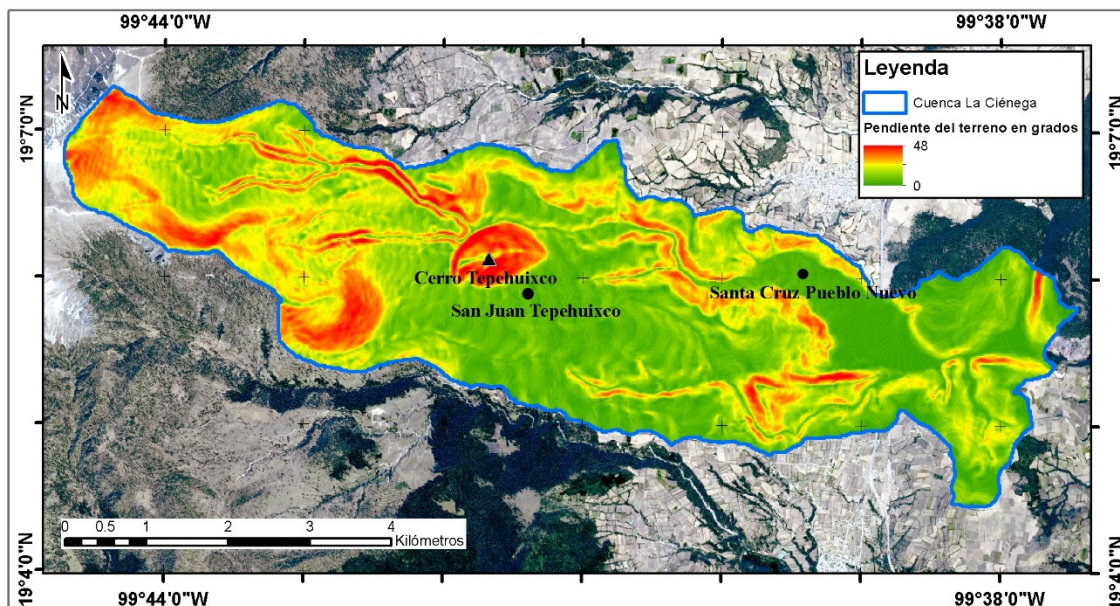


Figura 14. Mapa de pendientes del terreno en la zona de estudio. Elaboración propia con información de INEGI, 2016 y SIAP-SEDENA, 2016.

Geomorfológicamente la cuenca está limitada al oeste por el borde del cráter del VNT, en esta zona se encuentran domos de origen dacítico, al norte está limitada principalmente por laderas y lomeríos de lava cubiertos por piroclastos, al sur por coladas de lava parcialmente cubiertas por piroclastos y lomeríos de pómez y de depósitos de caída, y al este por una mesa de lava, la cual no permite la salida de la corriente fluvial, y lomeríos de lava cubiertos por pómez y ceniza (García *et al.*, 2018). La zona de estudio se encuentra sobre material piroclástico, depósitos de flujo y cenizas, avalancha de escombros, lahares, oleadas piroclásticas, flujos de pómez y depósitos de caída, elementos que favorecen el desarrollo de procesos gravitacionales debido a su poca cohesión y fácil erosión (Aceves *et al.*, 2014b; García *et al.*, 2018) (Figura 15).

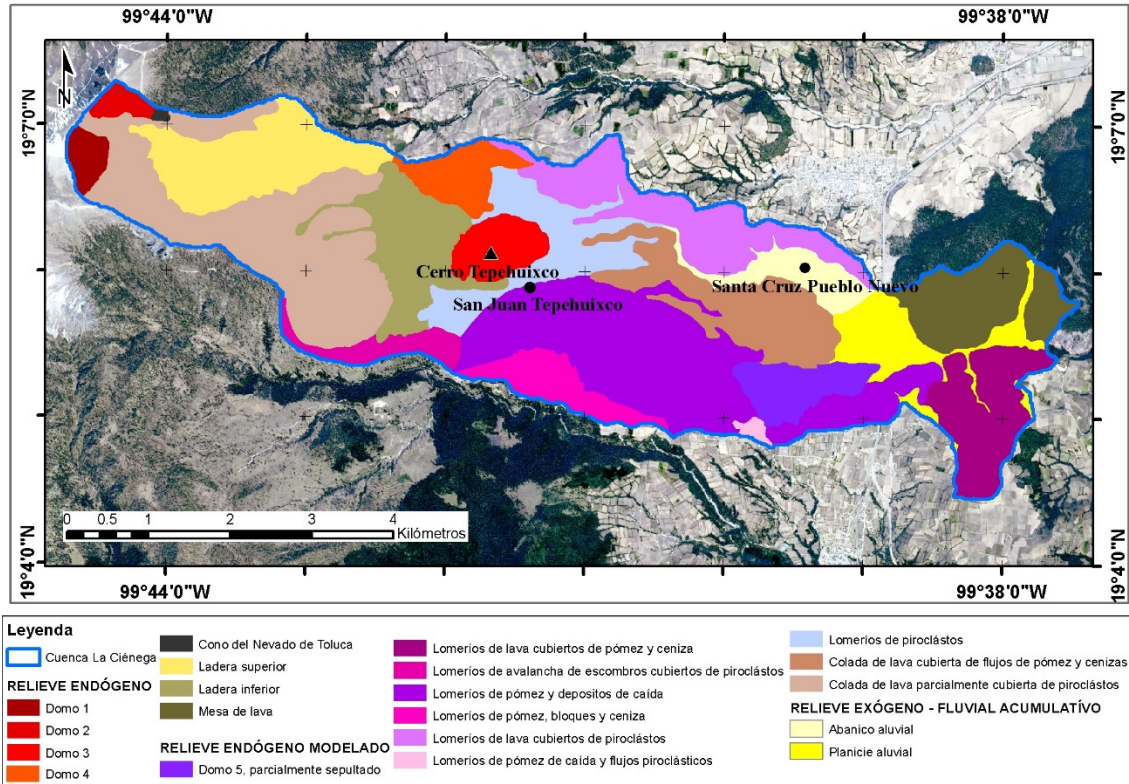


Figura 15. Mapa geomorfológico de la cuenca La Ciénega. Elaboración propia con información de García, 2017.

La cuenca es atravesada por el Sistema de Fallas Tenango (SFT) que es el sistema más reciente y activo del VNT. Dentro de la zona de estudio se formó un graben el cual ha sido rellenado por depósitos de piroclastos y epiclásticos, lo que le podría dar una inestabilidad estructural (Capra *et al.*, 2008; Norini *et al.*, 2008; Aceves *et al.*, 2014b). La litología presente en la zona está compuesta principalmente de toba y roca extrusiva, en lo que respecta al poblado de Santa Cruz Pueblo Nuevo este se encuentra sobre material aluvial. En la parte este de la cuenca se puede apreciar el derrame basáltico que da el carácter de cuenca endorreica (Figura 16).

En la cuenca se presenta tres tipos de clima. En la parte más alta, por arriba de los 4,000 m s.n.m, lo cubre un clima de tipo frío de altura E(T) que se caracteriza por presentar temperaturas menores a los 10° C durante todo el año. Por debajo de los 4,000 m el clima es de tipo templado y esta zona se encuentra dividida en dos categorías, entre los 4,000 y los 3,000 se tiene un clima templado, semifrío, subhúmedo con lluvias en verano, en el que la temperatura media anual esta entre los 4° y 12 ° C. En la zona más baja de la cuenca, con altitudes menores a los 3,000 m se presenta un clima templado subhúmedo con lluvias en verano con una temperatura media anual de 13° C. En lo que respecta a la precipitación en el área correspondiente al clima E(T) se da una

precipitación anual de 900 mm, mientras que en las zonas de clima templado la precipitación anual máxima es de 1,000 mm (García, 2006; Toscana y Valdez, 2013; García, 2017).

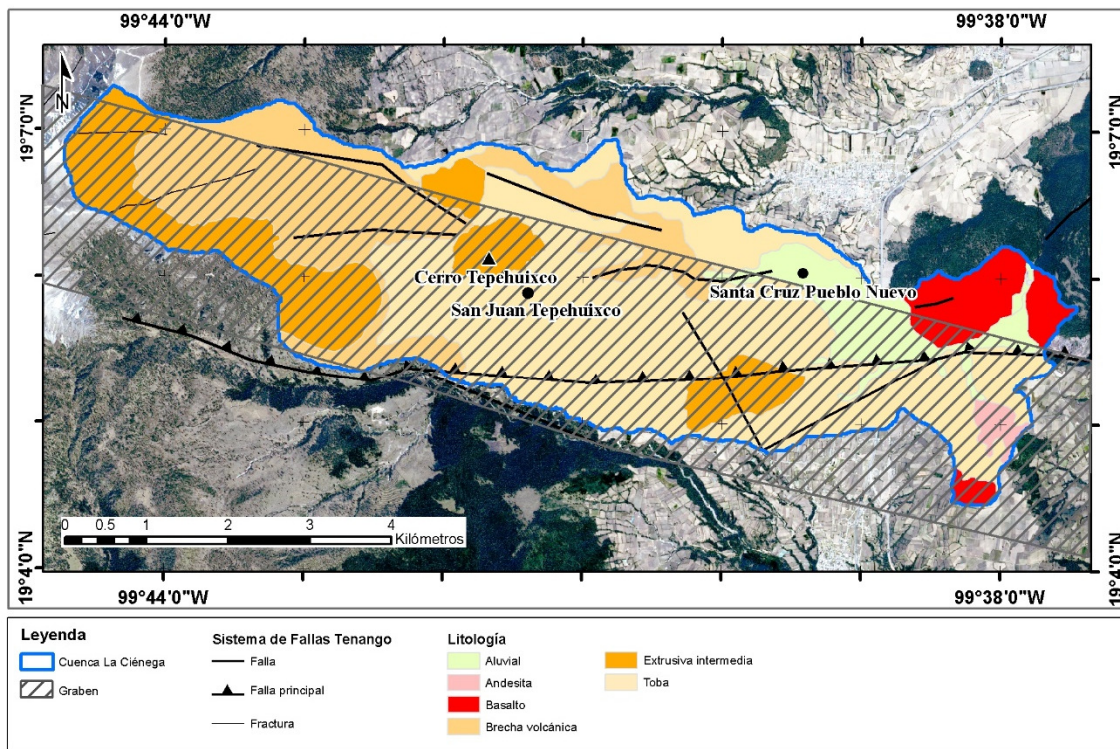


Figura 16. Mapa de la ubicación del graben y litología de la cuenca La Ciénega. Elaboración propia con información de INEGI, 2001a; Norini *et al.*, 2008 y García, 2017.

En la zona se presenta un tipo de drenaje dendrítico con rasgos paralelos, siendo el arroyo de la Ciénega un cauce de cuarto orden (García *et al.*, 2018).

CAPÍTULO 3. METODOLOGÍA.

El desarrollo del trabajo se divide en seis etapas:

- 1) Delimitación de la zona de estudio.
- 2) Elaboración del inventario de PG.
- 3) Procesamientos de las variables.
- 4) Aplicación del AEVC,
- 5) Elaboración de los modelos de probabilidad con base al AEVC y con información *in situ*.
- 6) Evaluación de los modelos del AEVC e *in situ*.

3.1. Delimitación de la zona de estudio.

La delimitación de la cuenca se realizó utilizando un método automatizado y su posterior revisión interactiva con otros límites generados en otros trabajos, así como con las curvas de nivel de la zona de estudio a escalas de 1:20,000 y 1:50,000, con ello se obtuvo un límite preciso de la cuenca en un tiempo relativamente rápido. Existen publicaciones como las de Aceves *et al.* (2014b) y García *et al.* (2018) en las que se presentan figuras con el límite de la zona de estudio, pero estas no cuentan con la información y el detalle necesario para realizar una digitalización de una manera precisa. Para delimitar la cuenca La Ciénega se utilizó como insumo principal el Continuo de Elevación Mexicano en su versión 3.0 (CEM) elaborado por el INEGI en 2013 (INEGI, 2013a). Las principales características del CEM son: proyección cartográfica Cónica conforme de Lambert; formato digital *raster* y resolución espacial de 15 metros por pixel. Se utilizaron las herramientas para el análisis hidrológico en el módulo GRID de Arc/Info® ver. 10.0 (ESRI, 2010). El procesamiento inicia con la generación del modelo de la dirección del flujo a partir del CEM, posteriormente, a partir del modelo de dirección del flujo, se calcula el modelo de depresiones. Utilizando la dirección del flujo y las depresiones se calcula el modelo que delimita la cuenca (Figura 17).

El límite de la cuenca resultado del proceso automatizado se revisó de manera interactiva con el límite presentado en los trabajos de Aceves *et al.* (2014b) y García *et al.* (2018). Así como la cartografía topográfica digital escala 1:20,000 (INEGI, 2012) y escala 1: 50,000 (INEGI, 1998; 2009; 2013c y 2014a) de la zona de estudio con el propósito de realizar una validación del límite obtenido (Figura 18).

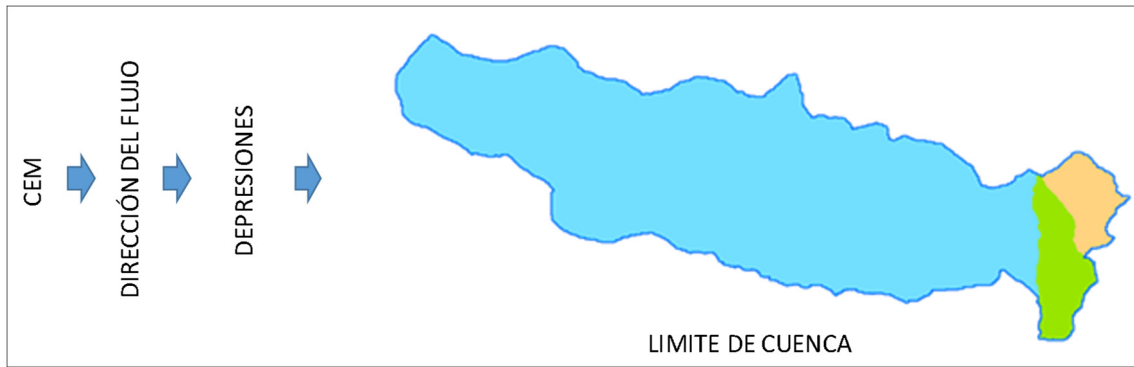


Figura 17. Esquema de la generación automatizada del límite de la cuenca La Ciénega a partir del CEM v3.0 (INEGI, 2013a). Elaboración propia.

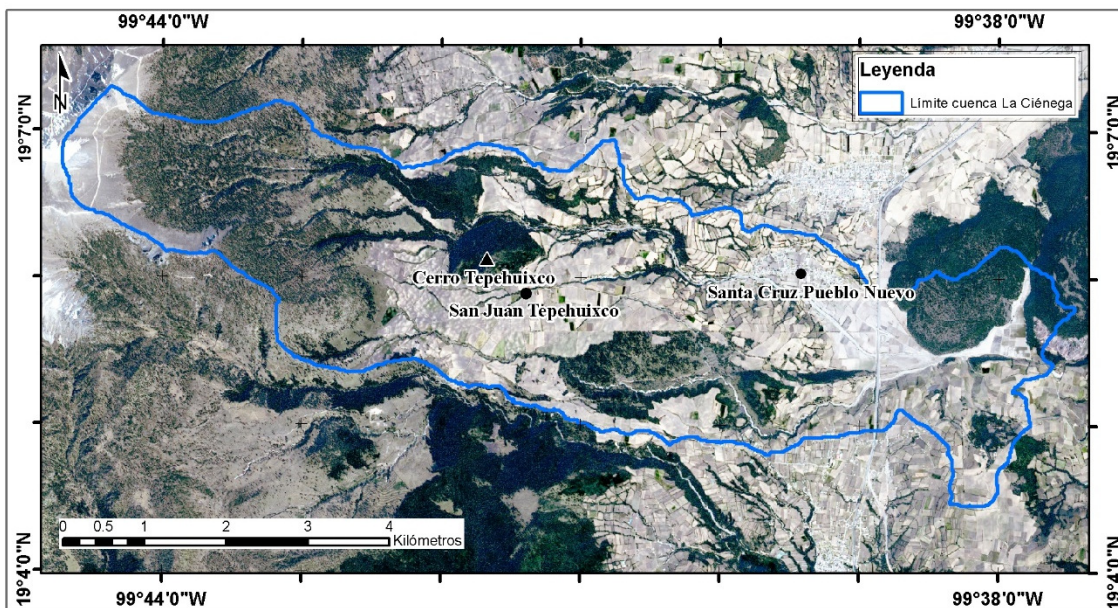


Figura 18. Límite final de la cuenca La Ciénega. Elaboración propia.

3.2. Elaboración del inventario de Procesos Gravitacionales.

La elaboración del inventario implicó la recopilación de información generada por Álvarez en 2015 la cual corresponde a un inventario de PG basado en la interpretación de fotografías aéreas para el año de 1983 en escala 1: 37,000 e imágenes del satélite Landsat 8 del año 2014, con una resolución espacial de 15 metros para la banda pancromática y de 30 metros para las bandas multiespectrales, de dicho inventario se obtuvieron 154 PG. La información de este inventario se encuentra distribuida principalmente en la zona noroeste y centro-sur de la cuenca (Figura 19). Con el propósito de contar con datos que tuvieran una distribución más completa sobre el área de la cuenca este primer inventario fue actualizado con visitas a campo durante los años 2015 y 2016. Estas vistas se enfocaron en las zonas de la cuenca en las que no se contaba con

información sobre la ubicación de los PG. Como se observa en la Figura 19, con la integración de ambos inventarios se obtuvo una mejor cobertura en la distribución de los deslizamientos inventariados en la cuenca.

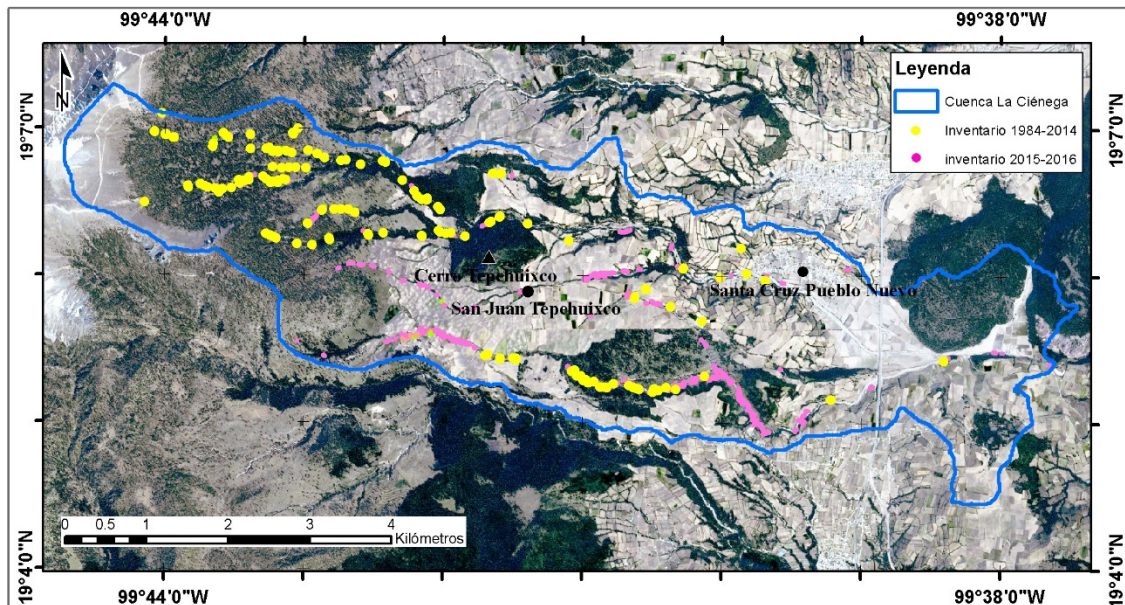


Figura 19. Distribución de los PG en la cuenca La Ciénega con base en la integración de 2 inventarios. Elaboración propia con información de Álvarez, 2015 y visitas de campo 2015 a 2016.

Durante los recorridos en campo para el levantamiento del inventario de los PG se obtuvieron los siguientes datos: Ubicación (coordenadas UTM), tipo de deslizamiento, dimensiones de la cabecera (longitud L, altura A y profundidad P), ubicación con respecto al margen del río (en relación a la dirección de la corriente) y fotografía del sitio (Figura 20). Con respecto a la ubicación de los PG esta hace referencia a dos situaciones: la primera, cuando fue posible posicionarse sobre el área del PG, las coordenadas corresponden a la localización del mismo; la segunda, cuando por condiciones de accesibilidad no fue posible ubicarse sobre el área del PG, se tomó la distancia y la dirección en la que se encuentra el PG con respecto al sitio en donde fue tomada la coordenada de referencia.

Con la información del inventario completo se elaboró una cobertura digital de puntos utilizando el programa Arc/Info® ver 10.0, a partir de esta y con el propósito de contar con la representación de la superficie de las cabeceras se realizó la digitalización de las mismas. Esto se hizo de manera interactiva ubicando cada deslizamiento sobre las imágenes Spot 6 (SIAP-SEDENA, 2016). Estas imágenes fueron procesadas para elaborar un compuesto híbrido entre las bandas multiespectrales y la banda pancromática, lo que dio como resultado una imagen multiespectral

con resolución espacial de 1.5 m por pixel (Intergraph©, 2012). Identificando cada deslizamiento sobre la imagen de satélite y con la información tomada en campo sobre las dimensiones de los PG se realizó la digitalización del área de las cabeceras de los deslizamientos (Figura 21).

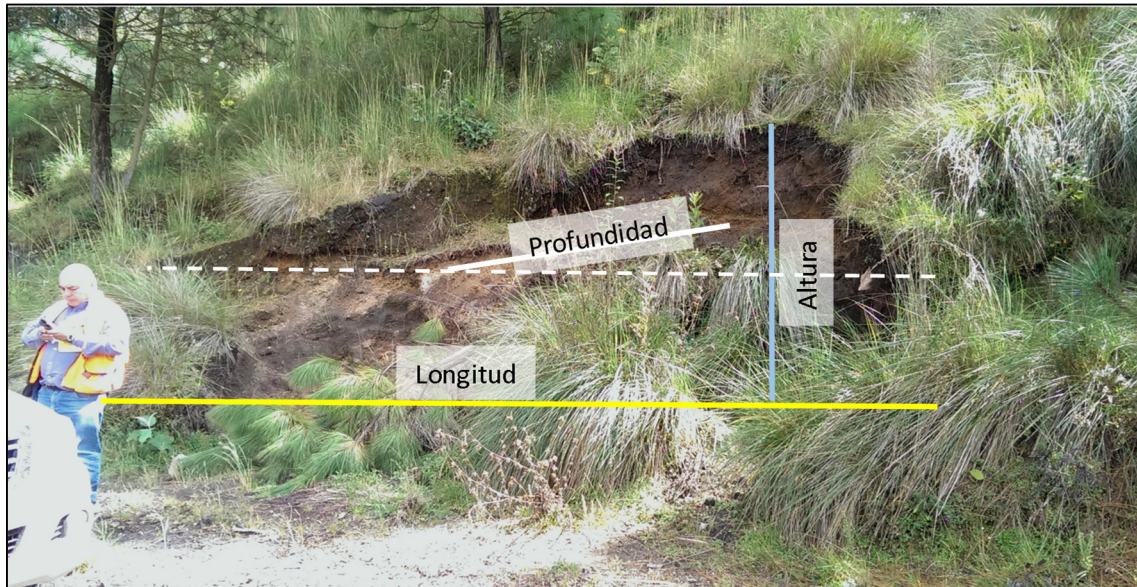


Figura 20. Forma en que fueron tomadas las medidas de la cabecera de los deslizamientos durante los recorridos en 2015 y 2016. Elaboración propia.

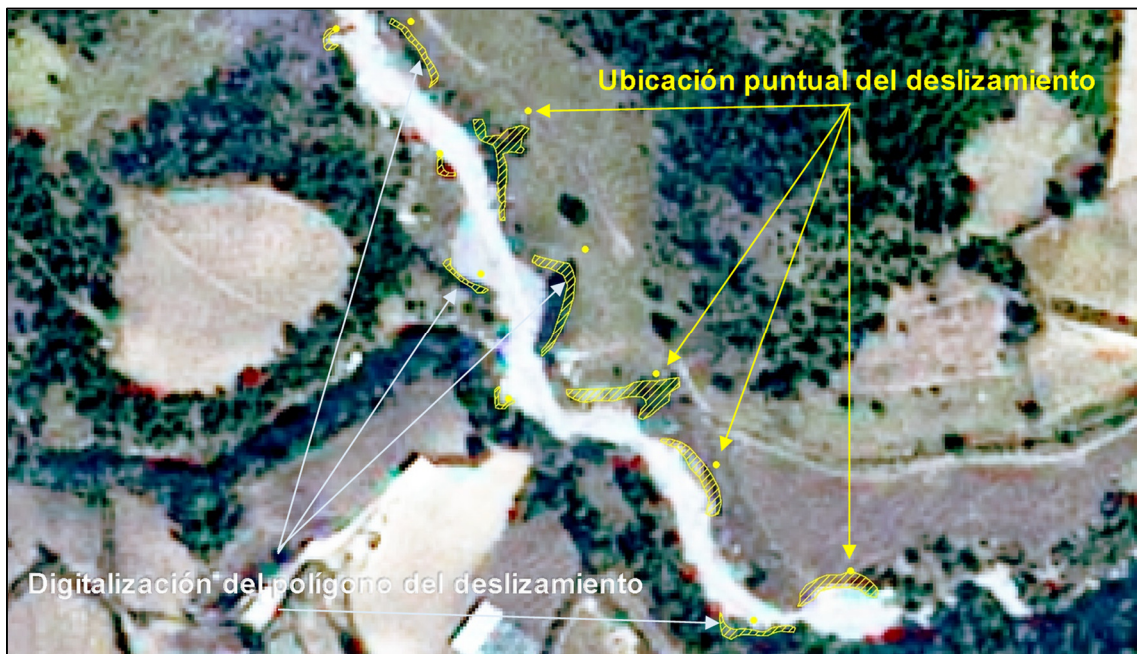


Figura 21. Digitalización del polígono de las cabeceras a partir de la ubicación y las dimensiones de los deslizamientos tomadas en campo. Fuente elaboración propia con información de campo e imágenes spot (SIAP-SEDENA, 2016).

Como resultado del procedimiento anterior se obtuvieron las cabeceras de los 290 deslizamientos que cubren una superficie del 4.8 ha. Se establecieron las zonas de cabecera como las áreas de deslizamiento, ya que estas se forman en donde tienen su origen este tipo de PG (Lugo, 2011).

3.3. Descripción y procesamiento de las variables.

A continuación, se describen las principales características de las variables dependiente y explicativas.

3.3.1. Variable dependiente.

Para la aplicación del modelo estadístico de RL es necesario que la variable dependiente sea de tipo dicotómico o binaria, es decir que solo pueda tomar dos valores, otro requisito es que los valores de esta variable sean excluyentes uno del otro (Pardo y Ruíz, 2002). En este estudio se tomó como variable dependiente la cartografía de los 290 deslizamientos de la zona de estudio, determinando a las áreas de deslizamiento y las áreas en las que no se identificó la presencia de cabeceras (zonas estables) como las dos categorías que componen dicha variable.

A partir del inventario de deslizamientos 1984-2016 se elaboró la cartografía en formato *raster* con un tamaño de 15 m por pixel con el propósito de que fuese compatible con la escala de la cartografía recabada como variables explicativas.

Con la cartografía del inventario de deslizamientos en formato *raster* se contabilizó el número de pixeles o celdas que componen el área de estudio para los deslizamientos y las áreas estables. Lo anterior dio como resultado un total de 1,299 pixel correspondientes a deslizamientos (29.2 ha) y 132,390 que representan zonas estables (2,978.8 ha), que corresponde a un 0.97 % y 99.03 %, respectivamente, del número de pixeles que integran la superficie de la cuenca (Figura 22).

El cálculo del modelo de RL se realizó a nivel local (por pixel) por ello se consideró como población total del estudio todos los pixeles que integran la zona de la cuenca.

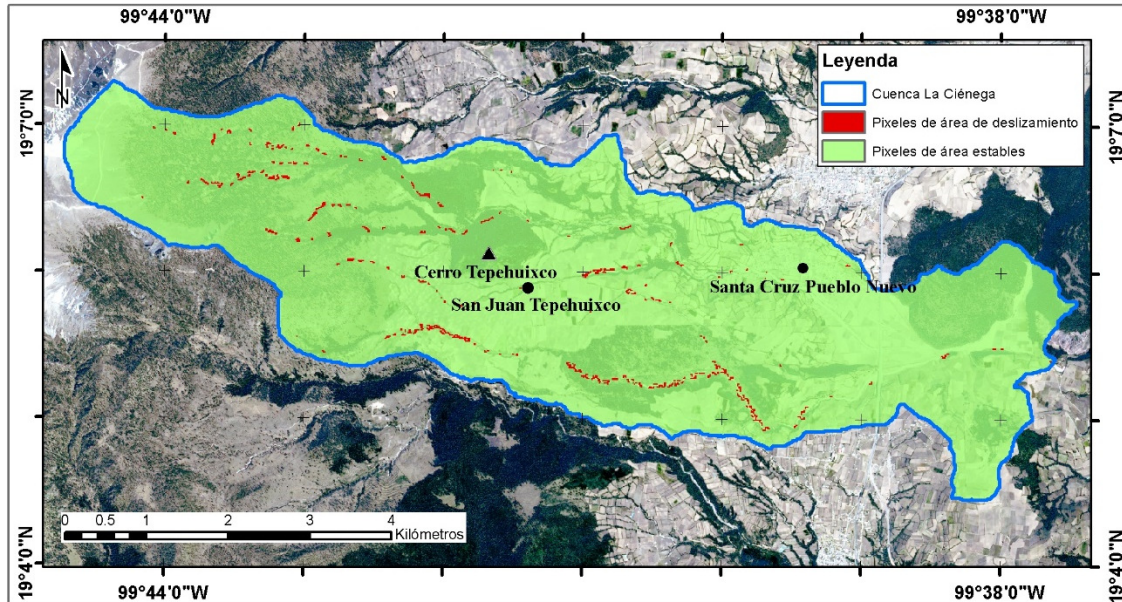


Figura 22. Mapa de las áreas estables y de deslizamiento. Elaboración propia.

3.3.2. Variables explicativas.

Con base en estudios sobre la elaboración de modelos estadísticos de probabilidad a deslizamientos (Atkinson y Massari, 2011; Von Ruetten *et al.*, 2011; Hong *et al.*, 2015; Huang *et al.*, 2015; Aceves *et al.*, 2016; Aditian *et al.*, 2018; Othman *et al.*, 2018; Pham *et al.*, 2018) se identificaron variables que constantemente son utilizadas en este tipo de trabajos. A partir de estas se realizó la búsqueda de información cartográfica disponible para la zona de estudio a una escala de 1:50,000 o mayor, teniendo como resultado la selección de las 16 variables explicativas que fueron trabajadas para la elaboración del modelo espacial de probabilidad a deslizamientos. Estas variables se separaron en dos grupos de acuerdo al tipo de datos que representan: variables cualitativas y variables cuantitativas.

Variabes Cualitativas.

En el estudio se incluyen seis variables explicativas con valores cualitativos nominales: Uso de suelo y vegetación 2015, cambios de uso de suelo y vegetación (1983-2015), geomorfología, litología, edafología y textura del suelo.

Uso del suelo y vegetación 2015.

Se obtuvo de la interpretación visual de imágenes de satélite Spot 6 (SIAP-SEDENA, 2016). La interpretación se basó en los recorridos de campo realizados en la zona de estudio. La definición de clases se hizo con base en la información de los documentos de INEGI (2015), Franco *et al.*, (2006) y Álvarez, 2015. La Tabla 5 describe las 9 clases que componen este mapa.

Tabla 5.- Clases de la variable de uso del suelo y vegetación 2015.

CLASE	CLAVE
Asentamientos humanos (AH). Poblados.	1
Agricultura (Agr). Usos del suelo de agricultura de riego y temporal.	2
Bosque fragmentado (Bf). Bosque de coníferas y latifoliadas, asociado con una mayor presencia de matorral, pastos y/o agricultura.	3
Bosque semidenso (Bs). Bosque de coníferas y latifoliadas con una presencia menor de matorral, pastos y/o agricultura.	4
Bosque denso (Bd). Bosque de coníferas y latifoliadas.	5
Pastizal natural (PN). Pastizales que, de manera natural sin la intervención de alguna actividad antrópica, se desarrollan por lo general en las áreas de transición entre los matorrales y los bosques.	6
Pastizal inducido (PI). Pastizales que se desarrollan como consecuencia de algún tipo de actividad antrópica como el desmonte, agricultura abandonada o quemas frecuentes.	7
Áreas sin vegetación (SV). Corresponde a las zonas en donde no se aparecía algún tipo de vegetación o actividad agrícola.	8
Otros tipos de vegetación (OTV). Incorpora todos aquellos tipos de vegetación que no se incluyeron en las clases anteriores.	9

Elaboración propia.

La Figura 23 muestra el mapa con la distribución de los usos de suelo y vegetación en la zona de estudio.

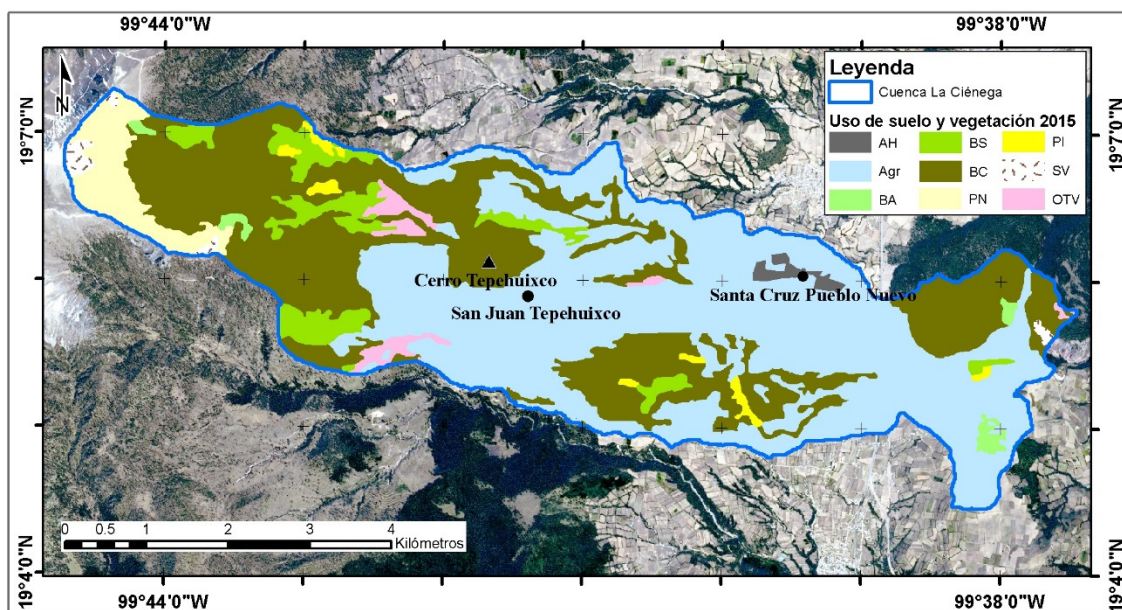


Figura 23. Mapa de uso de suelo y vegetación 2015. Elaboración propia con información del SIAP-SAGARPA, 2016.

Cambios de uso de suelo y vegetación (1983-2015).

Se obtuvo a partir de la cartografía de uso de suelo y vegetación de los años de 1983 (Álvarez, 2015) y la del año 2015. Se realizó una sobreposición de los mapas y las combinaciones de categorías se clasificaron en tres clases como se muestran en la Tabla 6. La Figura 24 muestra el mapa de cambios de uso de suelo y vegetación.

Tabla 6.- Clases de la variable de cambios de uso del suelo y vegetación (1983-2015).

CLASE	CLAVE
Permanencia. Representa aquellas áreas en las que su uso de suelo y vegetación no tuvo cambio durante el periodo.	1
Cambios. Corresponde a zonas Agr y OTV que pasaron a AH y Agr, respectivamente.	2
Deforestación. Considera el cambio de los tres tipos de bosques (Bf, Bs y Bd) a zonas Agr, PI y SV.	3

Elaboración propia.

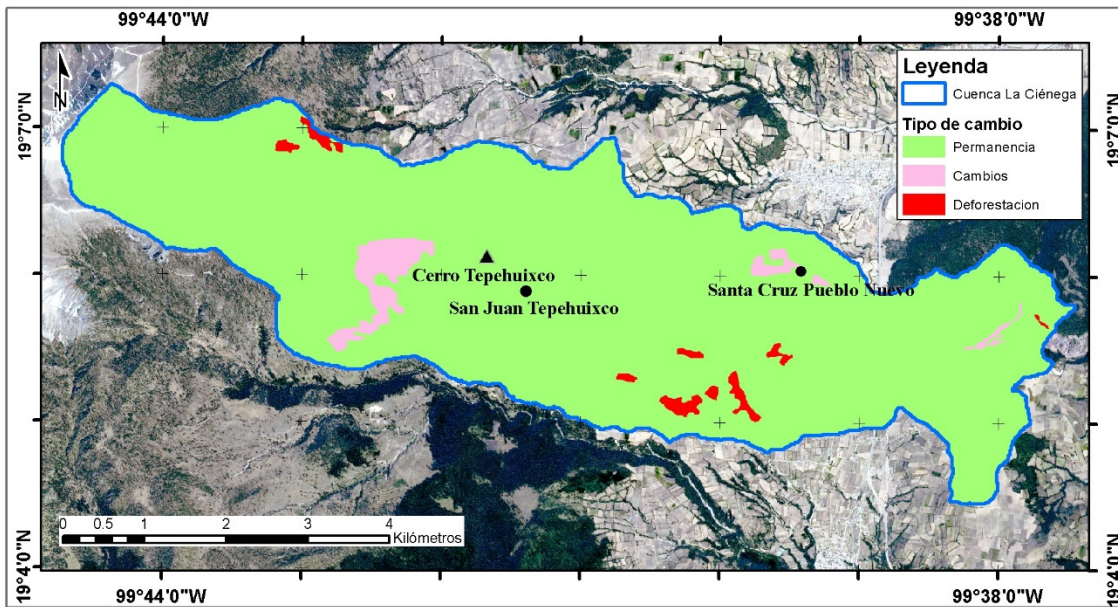


Figura 24. Mapa de cambios en la cobertura de uso de suelo y vegetación durante el periodo 1983-2015. Elaboración propia con información de Álvarez, 2015 y mapa de uso de suelo y vegetación 2015.

Geomorfología.

El mapa geomorfológico fue tomado del estudio de García (2017). En este mapa se presentan un total de 20 formas del relieve en la zona de estudio (Tabla 7). La Figura 25 muestra la distribución de las geoformas dentro de la cuenca.

Tabla 7.- Clases de la variable de geomorfología.

CLASE	CLAVE
Domo 1	1
Domo 2	2
Domo 3	3
Domo 4	4
Cono del Nevado de Toluca (NT). Originado el depósito de lava y material piroclástico. Presenta pendientes del terreno entre los 15° y 35°.	5
Ladera superior. Declive que delimita el cono del NT, de forma recta con pendientes de entre los 15° y 30°.	6
Ladera inferior. Declive que delimite a la colada de lava parcialmente cubierta por piroclastos. Con pendientes del terreno entre los 10 y 30°.	7
Mesa de lava. Estructura elevada y plana, compuesta de lava andesítica. Con pendientes de 2! a 6°.	8
Domo 5 parcialmente sepultado. Estructura de forma redondeada de composición basáltica parcialmente cubierta por material piroclástico. Con pendientes del terreno entre los 15° y 40°	9
Lomeríos de lava cubiertos de pómez y ceniza. Los lomeríos tienen su origen en la erosión de laderas rectas con una inclinación prolongada. Su constitución es de lavas cubiertas por piroclastos, pómez y ceniza.	10
Lomeríos de avalancha de escombros cubiertos de piroclastos. Lomeríos constituidos por depósitos de avalancha de escombros cubiertos por pómez y depósitos de caída.	11
Lomeríos de pómez y depósitos de caída. Lomeríos formados por depósitos de caída y flujos piroclásticos,	12
Lomeríos de pómez, bloques y cenizas. Lomeríos conformados por flujos de pómez dacítica cubiertos por depósitos de bloques y fragmentos de andesita	13
Lomeríos de lava cubiertos por piroclastos. Lomeríos compuestos por bloques cubiertos de flujos de pómez, ceniza y depósitos de caída	14
Lomeríos de pómez de caída y flujos piroclásticos. Lomeríos compuestos por depósitos de pómez y flujos piroclásticos.	15
Lomeríos de piroclastos. Lomeríos constituidos por depósitos de piroclastos.	16
Colada de lava cubierta de flujos de pómez y ceniza. Las coladas de lava corresponden a material magmático que fluye pendiente abajo. En particular esta colada de lava está cubierta por flujos de pómez, de ceniza y depósitos de caída.	17
Colada de lava parcialmente cubierta por piroclastos. Cubierta por material piroclástico de composición dacítica.	18
Abanico aluvial. Originado de la acumulación de sedimento en la desembocadura las corrientes fluviales. Con forma de medio cono con el vértice en la parte más alta.	19
Planicie aluvial. Se forma de la acumulación de material por procesos fluviales. Compuesta de material volcánico.	20

Elaboración con información de García, 2017.

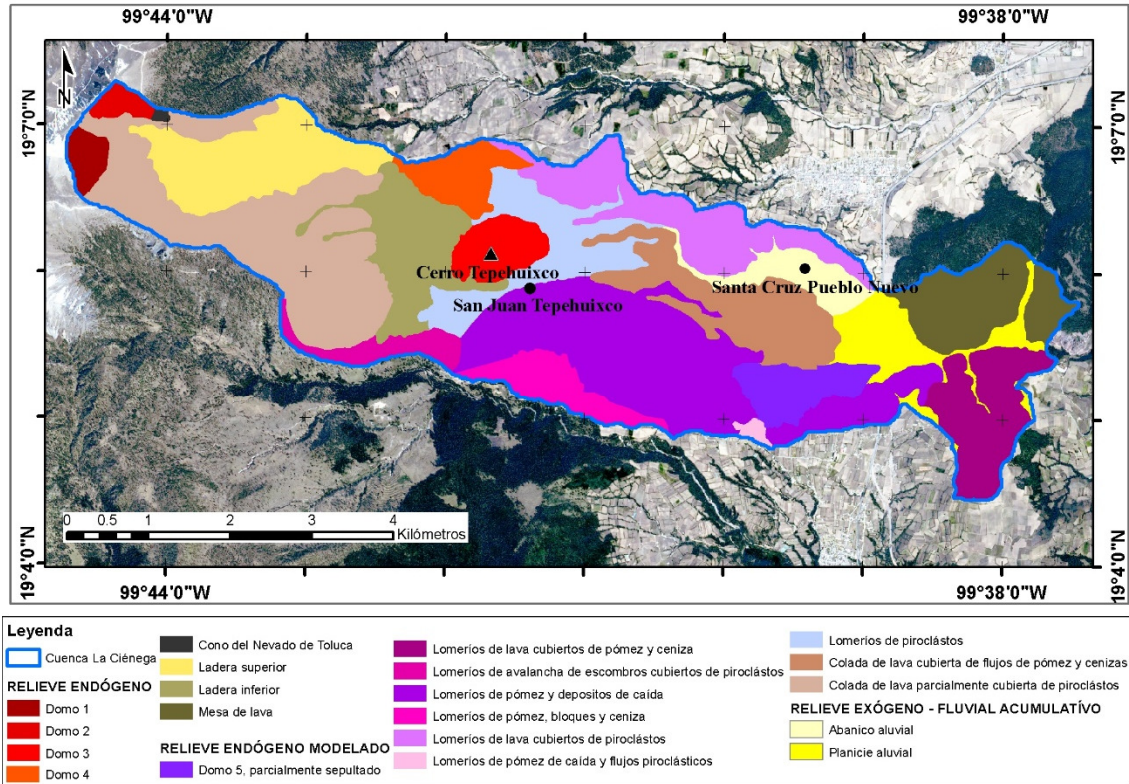


Figura 25. Mapa geomorfológico de la zona de estudio. Elaboración con información de García, 2017.

Litología.

Obtenido de la digitalización del tema de litología de las cartas geológicas en escala 1:50,000 E14-A47 (Volcán Nevado de Toluca) y E14-A48 (Tenango) de CETENAL de 1975 (INEGI, 2001a) (Figura 26). En la Tabla 8 se describen las 6 clases que cubren la zona de estudio, de acuerdo a la documentación correspondiente a la cartografía geológica en escala 1:50,000 del INEGI (2005) y a Sopena y Sánchez-Moya (2010).

Tabla 8.- Clases de la variable de litología.

CLASE	CLAVE
Ígnea extrusiva intermedia (Igei). Roca ígnea extrusiva intermedia, incluye la Traquita y Latita, son rocas de origen magmático, están compuestas por feldespatos alcalinos y plagioclasas sódicas.	1
Aluvial (al). Depósitos de material originados por un sistema fluvial.	2
Brecha volcánica (Bv). De origen magmático, la conforman materiales fragmentados que fueron proyectados por los conductos volcánicos y depositados, incluye fragmentos de diferentes tamaños y composición (incluye Riolita, Andesita y Basalto).	3
Toba (T). De origen explosivo, incluye material suelto y consolidado, integra fragmentos de diversa composición y tamaños menor a 4 mm (incluye Riolita, Andesita y Basalto).	4
Andesita (A). Roca de tipo Igei, que tiene como característica que su textura está definida por cristales de mayor tamaño sobre un fondo uniforme de cristales de menor tamaño, compuesta por plagioclasa sódica.	5
Basalto (B). Roca de origen magmático, de tipo extrusivo. Formada por plagioclasa cálcica.	6

Elaboración propia.

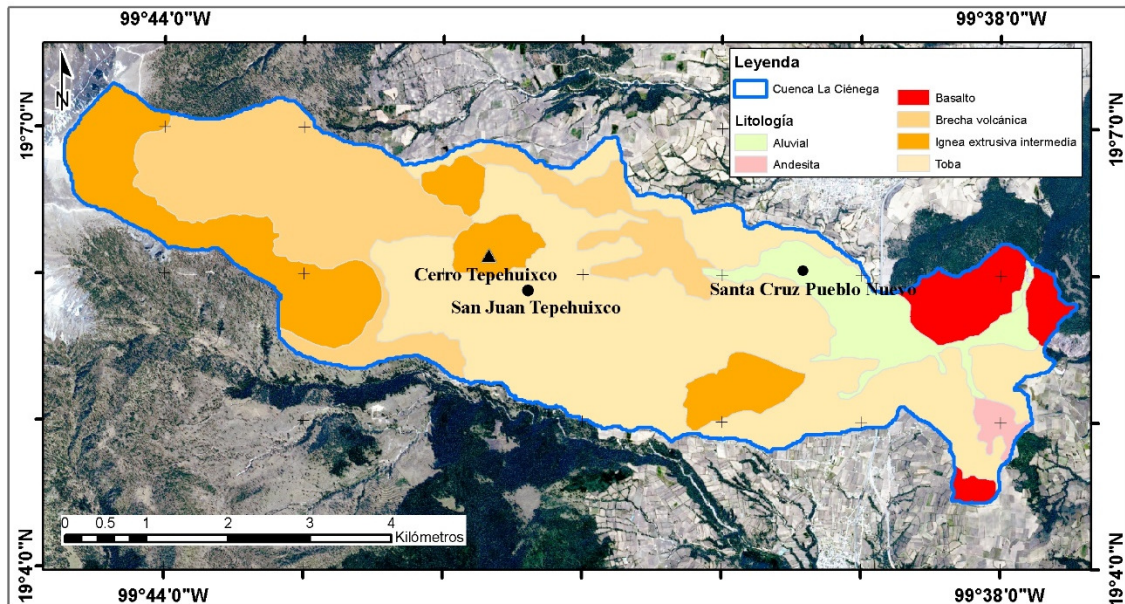


Figura 26. Mapa de la litología de la zona de estudio. Elaboración propia con información del INEGI, 2001a.

Edafología.

Se obtuvo de la digitalización del tema unidades de suelo de las cartas edafológicas en escala 1:50,000 E14-A47 (Volcán Nevado de Toluca) y E14-A48 (Tenango) de CETENAL de 1976 (INEGI, 2001b). De acuerdo al INEGI (2011; 2014b). Las principales características de los 7 tipos de suelos presentes en la zona se describen la Tabla 9. La Figura 27 presenta la distribución de los tipos de suelos de la zona de estudios.

Tabla 9. Tipos de suelo en la zona de estudio.

CLASE	CLAVE
Andosol húmico (Th). Son suelos recientes de origen volcánico, por lo general son muy aptos para cultivos. La mayoría de este tipo de suelo se encuentra sobre basaltos, andesitas, brecha volcánica básica, brecha sedimentaria y estructuras volcanoclásticas. El Andosol húmico contiene en promedio el 1% o más de carbono orgánico en los primeros 50 cm de profundidad, por lo general es de color oscuro.	1
Andosol mólico (Tm). La característica del Andosol mólico es que su horizonte superficial es oscuro, con un buen contenido de carbono orgánico y cuentan con un espesor de 20 cm o mayor.	2
Andosol ócrico (To). El Andosol ócrico se caracteriza por tener un horizonte mineral superficial, con 10 cm o más de espesor y con muy poco de carbono orgánico.	3
Feozem háplico (Hh). Son de color pardo a negro, con contenido de magnesio y potasio, con poco o sin carbonato en el subsuelo. Se desarrollan en superficies planas o ligeramente onduladas, son importantes para la agricultura de temporal. En particular los suelos háplicos no presentan desarrollo.	4
Fluvisol éútrico (Je). Contiene grandes cantidades de sedimentos fluviales, marinos o lacustres. Cuentan con una buena fertilidad y también son utilizados por asentamientos humanos. Se localizan principalmente en llanuras intermontanas y valles abiertos o ramificados. Su característica de éútrico significa que están saturados de calcio, magnesio, sodio y potasio.	5
Continúa en la siguiente página.	

Continuación de la página anterior	
CLASE	CLAVE
Leptosol-litosol- (l). Suelos con menos de 25 cm de espesor o con más del 80 % de su volumen ocupados por rocas o grava. Por lo general se localizan en áreas montañosas con pendientes mayores al 40 %. Principalmente utilizados como zonas para agostadero.	6
Regosol éútrico (Re). Suelos pedregosos de colores claros cuando son superficiales su color es muy semejante al de la roca que les da origen. El carácter éútrico determina que sus principales componentes son calcio, magnesio, sodio y potasio, lo que les da una buena fertilidad.	7

Elaboración propia.

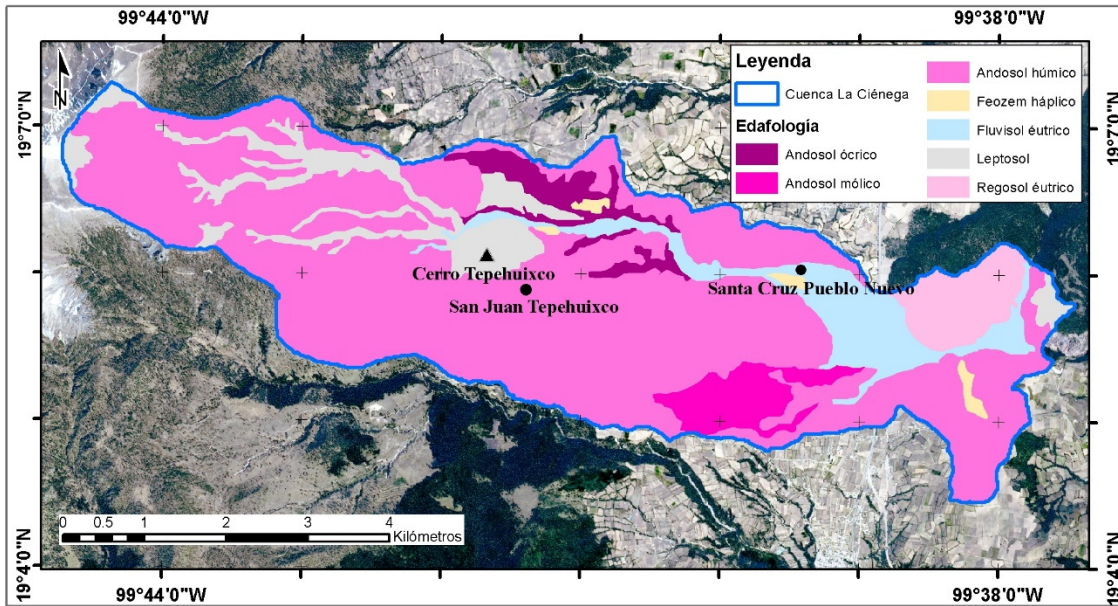


Figura 27. Mapa edafológico de la zona de estudio. Elaboración propia con información del INEGI, 2001b.

Textura del suelo.

Se obtuvo de la digitalización del tema clase textural del suelo de las cartas edafológicas en escala 1:50,000 E14-A47 (Volcán Nevado de Toluca) y E14-A48 (Tenango) de CETENAL de 1976 (INEGI, 2001b). La textura de los suelos se refiere a la proporción relativa del tamaño de promedio de las partículas de la parte fina del suelo dominante en los primeros 30 cm de profundidad. De acuerdo a la clasificación del INEGI (2011) la textura de los suelos en la zona de estudio se caracteriza por las categorías: gruesa y media (Tabla 10). En la Figura 28 se muestra la distribución de las texturas de los tipos de suelo presentes en la zona.

Tabla 10.- Clases de la variable de textura del suelo.

CLASE	CLAVE
Gruesa. Son suelos arenosos con cantidades de arena igual o mayores al 65 %, como consecuencia tienen poca capacidad de retención de agua y nutrientes para la vida vegetal.	1
Mediana. También conocidos como suelos francos. Su contenido de arena, arcilla y limo están general equilibrados.	2

Elaboración propia.

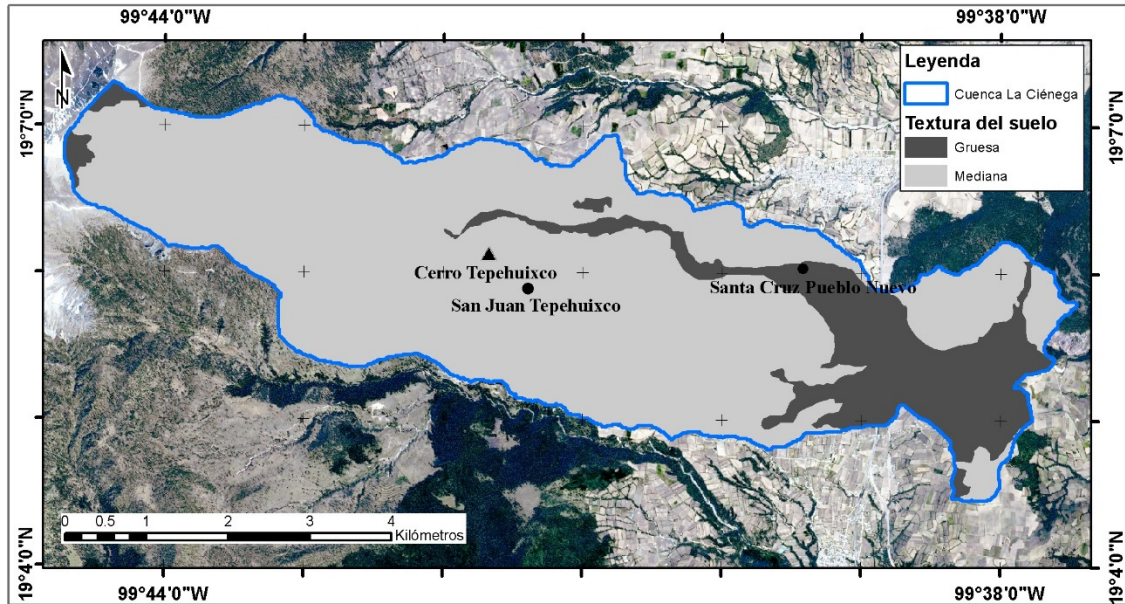


Figura 28. Mapa de la textura de los suelos en la zona de estudio. Elaboración propia con información del INEGI, 2001b.

Variables Cuantitativas.

Se trabajaron un total de 10 variables cuantitativas: Altitud en metros s.n.m., pendiente del terreno, dirección de la pendiente del terreno, curvatura de la superficie del terreno, rugosidad del terreno, longitud de la pendiente, distancia a fallas y fracturas, distancia a ríos, distancia a caminos e índice de vegetación de diferencia normalizada (NDVI).

Altitud en metros s.n.m.

Se obtuvo directamente de la información del CEM (INEGI, 2013a). Contiene los datos de la altitud del terreno en metros sobre el nivel del mar (m.s.n.m.). Este Continuo de elevación es presentado por el INEGI en formato *raster* (Figura 29), sus principales características son las siguientes: resolución espacial 15 m por pixel, tipo de datos enteros positivos y negativos, y error medio cuadrático de 4.9 m.

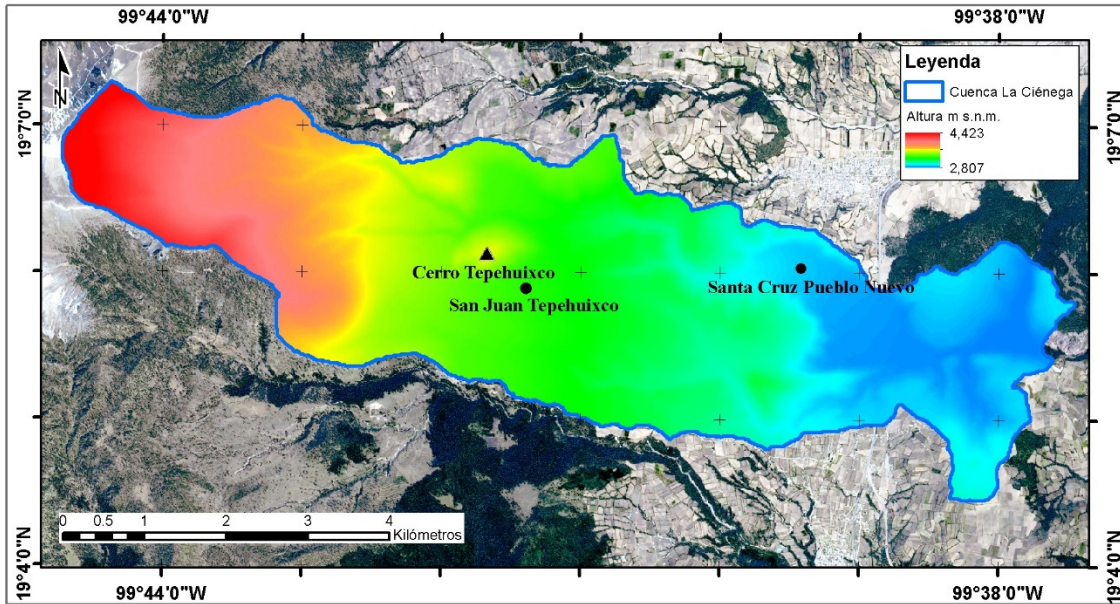


Figura 29. Mapa de altitudes en la cuenca La Ciénega. Elaboración propia con información del INEGI, 2013a.

Pendiente del terreno.

Expresa la pendiente del terreno en grados (Figura 30). Se aplicó la función *Slope* del módulo *GRID* del SIG Arc/Info 10.0 al CEM, calcula la tasa de cambio máxima (en grados).

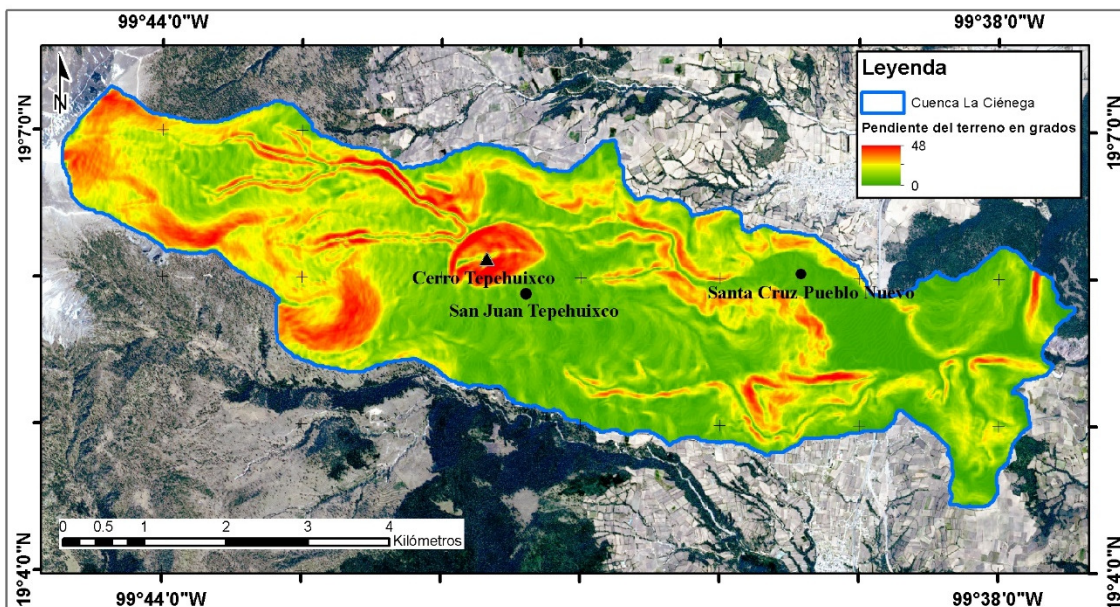


Figura 30. Mapa de la distribución de la pendiente del terreno en la cuenca La Ciénega. Elaboración propia con información del INEGI, 2013a.

Dirección de la pendiente del terreno

Se obtuvo aplicando la función de *aspect* en el programa SIG Arc/Info 10.0 al CEM, la función proporciona la dirección de orientación de la pendiente en forma descendente en una medida azimutal, con los valores de 0° cuando se tiene una dirección al norte hasta llegar a los 359°, la medida de la dirección va aumentando en el sentido de las manecillas del reloj. Las zonas planas reportan valores de -1 (Figura 31).

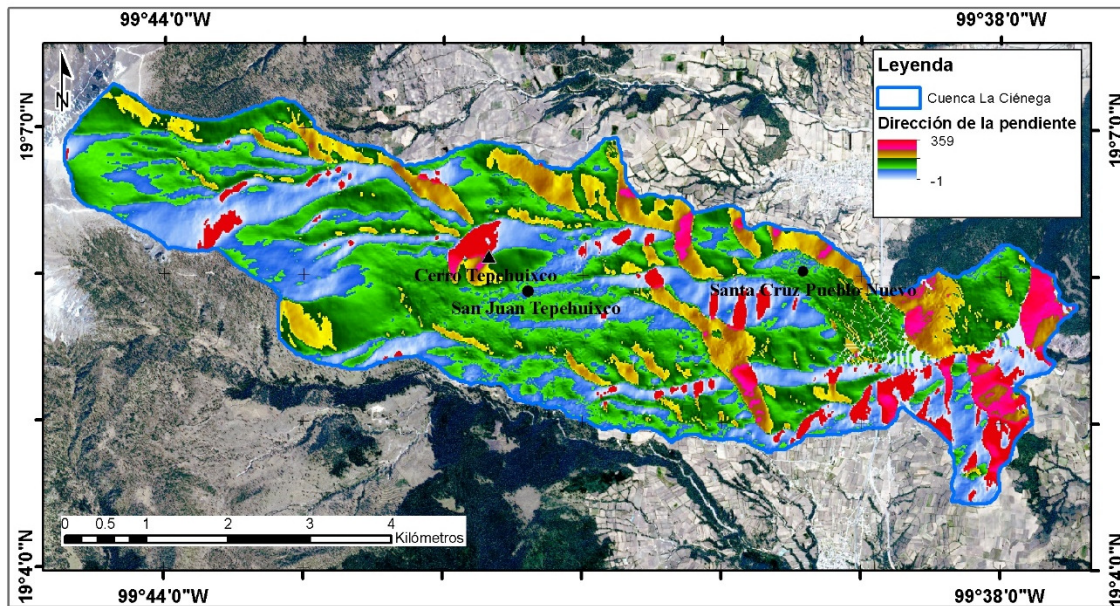


Figura 31. Mapa de la dirección de la pendiente del terreno en la cuenca La Ciénega. Elaboración propia con información del INEGI, 2013a.

Curvatura de la superficie del terreno

Se obtuvo aplicando la función *curvature* del programa SIG Arc/Info 10.0 al CEM 3.0. Los valores obtenidos de esta función son valores positivos para superficies convexas, los negativos para las superficies cóncavas, y 0 para áreas planas. La curvatura de un terreno tiene efectos sobre la aceleración y desaceleración de un flujo y como consecuencia de esto afecta directamente los procesos de erosión y sedimentación (ESRI, 2016).

De manera general los rangos de valores obtenidos de la función de curvatura de la superficie se pueden interpretar de la siguiente manera: Un área con relieve moderado dará valores entre 0.5 a -0.5, mientras que un área con relieve montañoso reportará valores entre los 4 a -4 o mayores (ESRI, 2016). En la Figura 32 se muestra la distribución de la curvatura de la superficie del terreno.

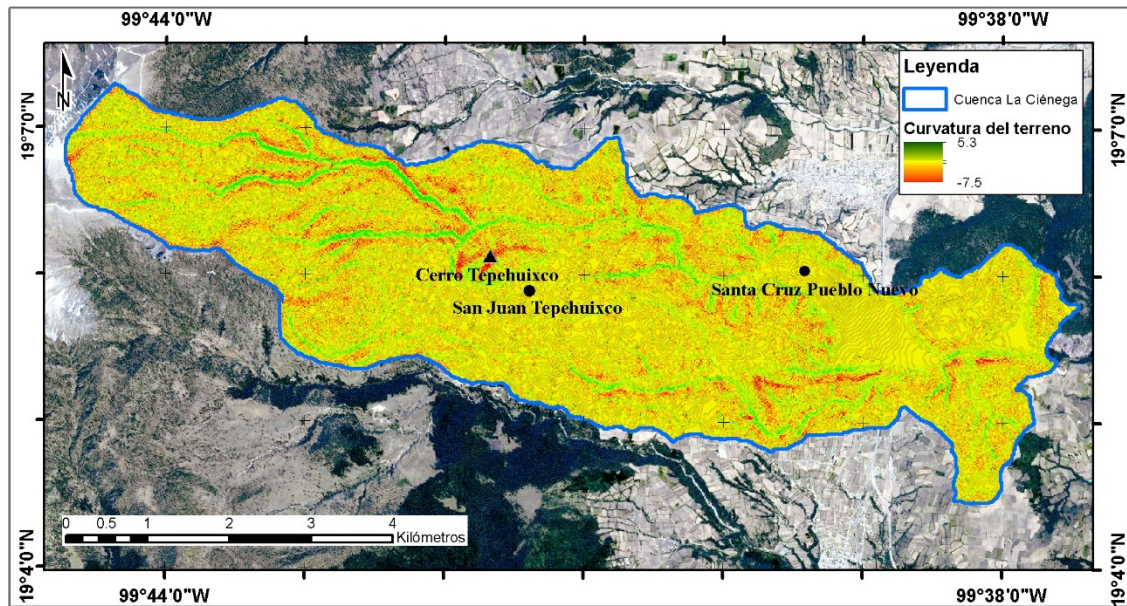


Figura 32. Mapa de la curvatura del terreno en la cuenca La Ciénega. Elaboración propia con información del INEGI, 2013a.

Rugosidad del terreno

Se calculó con la herramienta Vector Ruggedness Measure (VRM) (Sappington *et al.*, 2007). Este se basa en la medida de la heterogeneidad tanto de la pendiente como de la dirección de la pendiente por medio de la medición de la dispersión de los vectores en tres dimensiones (X, Y y Z) para un grupo de superficies (representadas por píxeles), combina ambas medidas (pendiente y dirección de la pendiente) en una sola. Los resultados son valores entre 0 y 1, los valores de 0 corresponden a zonas planas y los valores de 1 a zonas con rugosidad alta (Figura 33). Se trabajó con una ventana de 3 píxel para tener una medición de grupo de píxeles igual a la que se hace para los mapas derivados del CEM mencionados anteriormente.

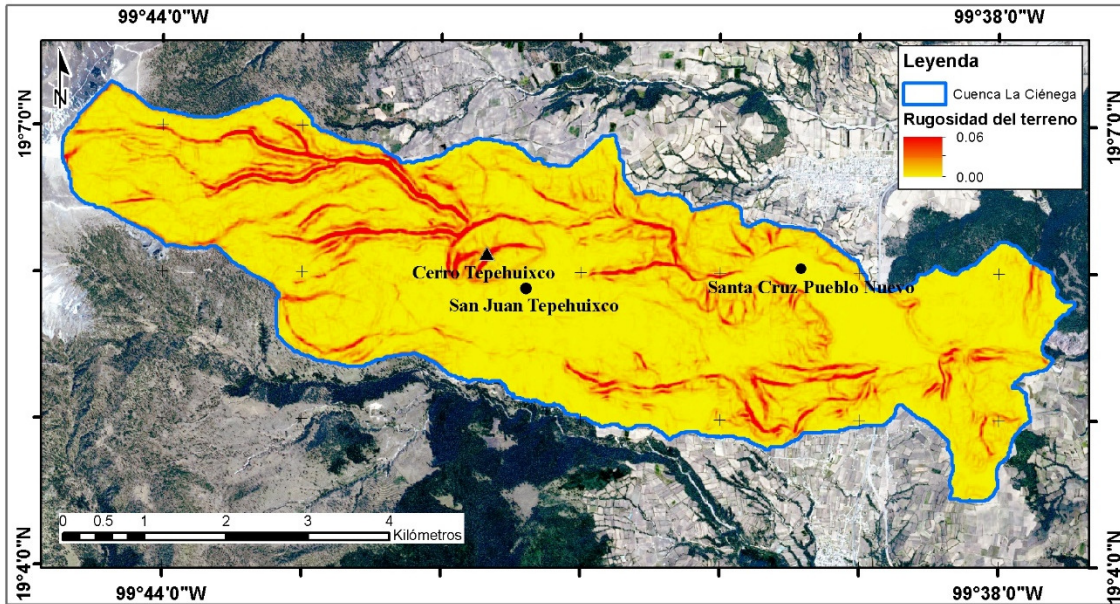


Figura 33. Mapa de la rugosidad del terreno en la cuenca La Ciénega. Elaboración propia con información del INEGI, 2013a.

Longitud de la pendiente

Este mapa se generó a partir del CEM aplicando la función *slope length* del SIG QGIS® versión 3.6.3.A. (FSFI, 2016). La Figura 34 representa la longitud de la pendiente en metros, iniciando en la parte más alta de la zona hacia la parte baja. La medición de la longitud de la pendiente se reinicia cuando se encuentra un cambio ascendente en la pendiente.

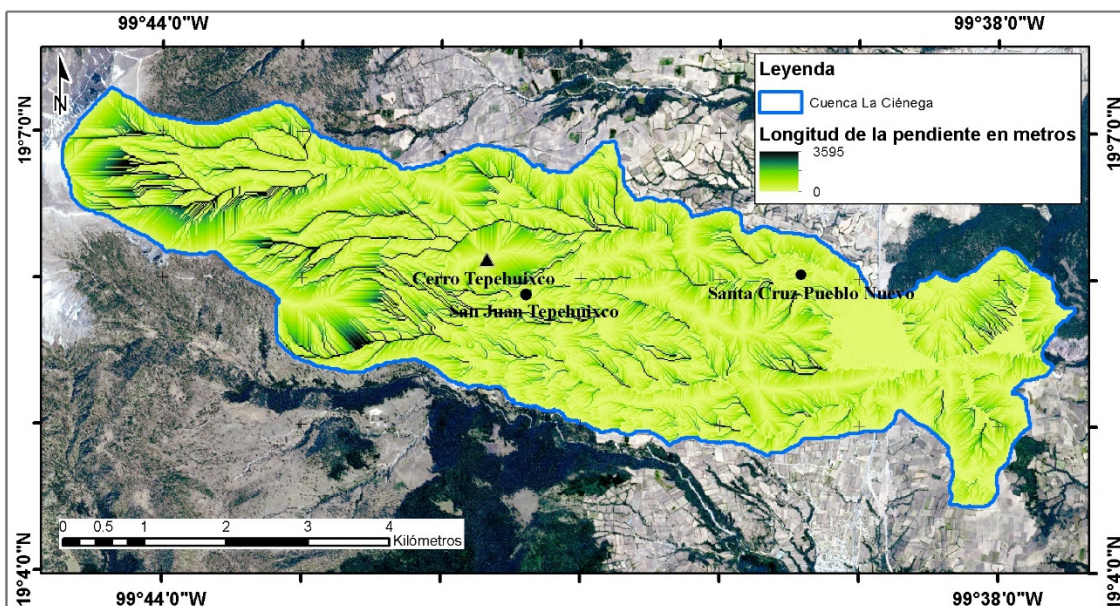


Figura 34. Mapa de la longitud de la pendiente en la cuenca La Ciénega. Elaboración propia con información del INEGI, 2013a.

Distancia a fallas y fracturas.

Se obtuvo de la digitalización de fallas y fracturas de las cartas geológicas en escala 1:50,000 E14-A47 (Volcán Nevado de Toluca) y E14-A48 (Tenango) de CETENAL de 1975 (INEGI, 2001a) y la información generada por García (2017). Ambos rasgos se agruparon en uno solo, posteriormente se aplicó la función *euclidean* en el programa SIG Arc/Info 10.0 para el cálculo de distancia euclidiana (ESRI, 2016). Con lo que se obtuvo el mapa continuo con la distancia a fallas y fracturas de la zona de estudio (Figura 35).

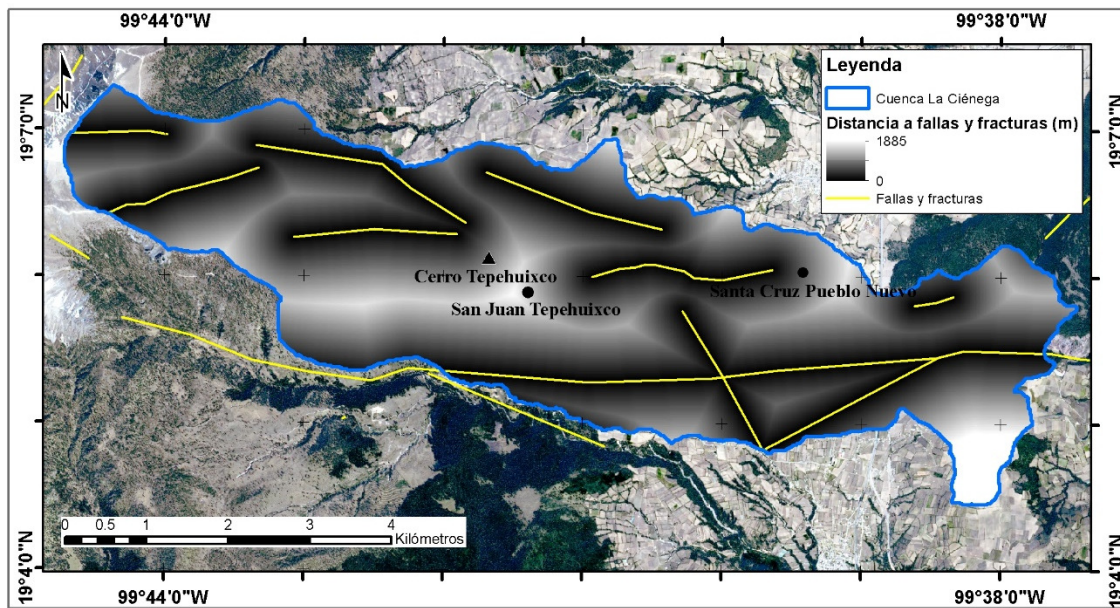


Figura 35. Mapa de la distancia a fallas y fracturas en la cuenca La Ciénega. Elaboración propia con información de García, 2017 e INEGI, 2001a.

Distancia a ríos.

Los ríos se obtuvieron a partir de la información digital de la red hidrográfica elaborada por el INEGI en 2010, esta fue complementada con la interpretación de imágenes de satélite Spot 6 del año 2015. Por último, se aplicó la función *euclidean* en el programa SIG Arc/Info 10.0 para el cálculo de distancia euclidiana en metros a los ríos (Figura 36).

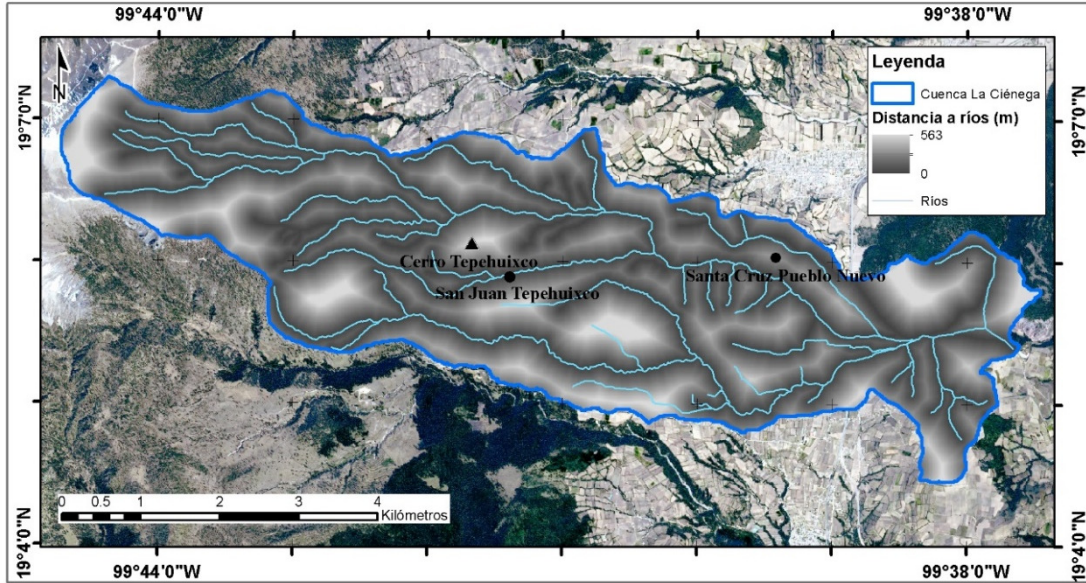


Figura 36. Mapa de la distancia a los ríos en la cuenca La Ciénega. Elaboración propia con información de INEGI, 2010.

Distancia a caminos.

Se obtuvieron de las coberturas topográficas en escala 1:50,000 E14-A47 (Volcán Nevado de Toluca) y E14-A48 (Tenango) (INEGI, 2013c; 2014a). Los tipos de vías de comunicación seleccionados fueron caminos y carreteras. La información fue complementada con la interpretación de imágenes Spot 6 del año 2015. Para obtener la distancia en metros a las carreteras y caminos se aplicó la función euclidean en el programa SIG Arc/Info 10.0 (Figura 37).

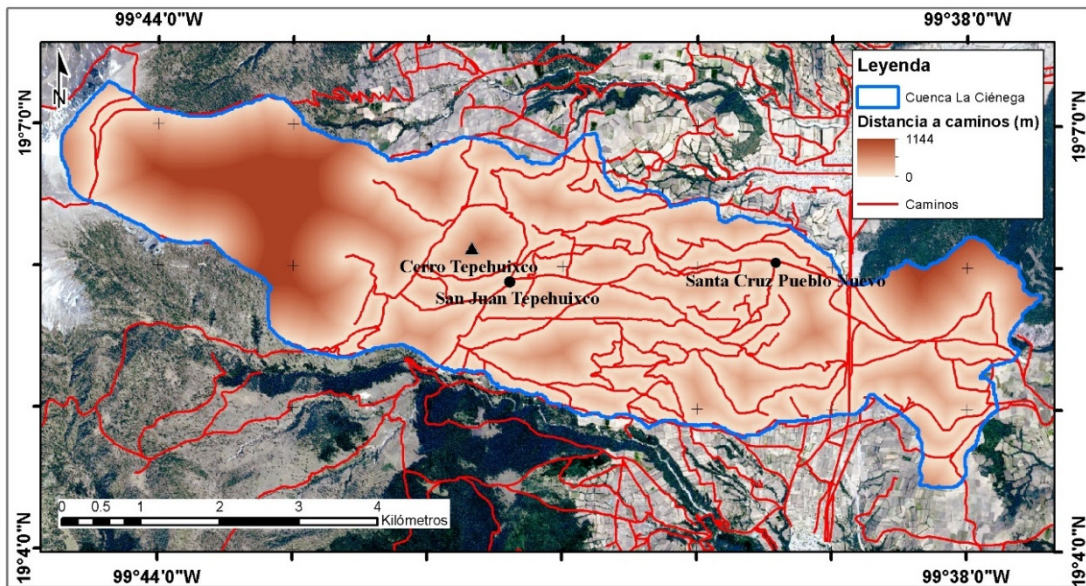


Figura 37. Mapa de la distancia a los caminos en la cuenca La Ciénega. Elaboración propia con información de INEGI, 2013c; 2014a.

Índice de vegetación de diferencia normalizada (NDVI).

Fue generado a partir de imágenes Spot 6 correspondientes al año 2015 (SIAP-SEDENA, 2016) utilizando las bandas rojo e infrarrojo cercano. El resultado se encuentra entre los valores -1 a 1. De manera general los valores cercanos a -1 corresponden a zonas cubiertas por agua o nieve, mientras que los valores cercanos a 1 corresponden a vegetación con una elevada presencia de clorofila, cuando se trata de vegetación natural esta podría interpretarse como una vegetación abundante o sana (Bannari *et al.*, 1995) (Figura 38).

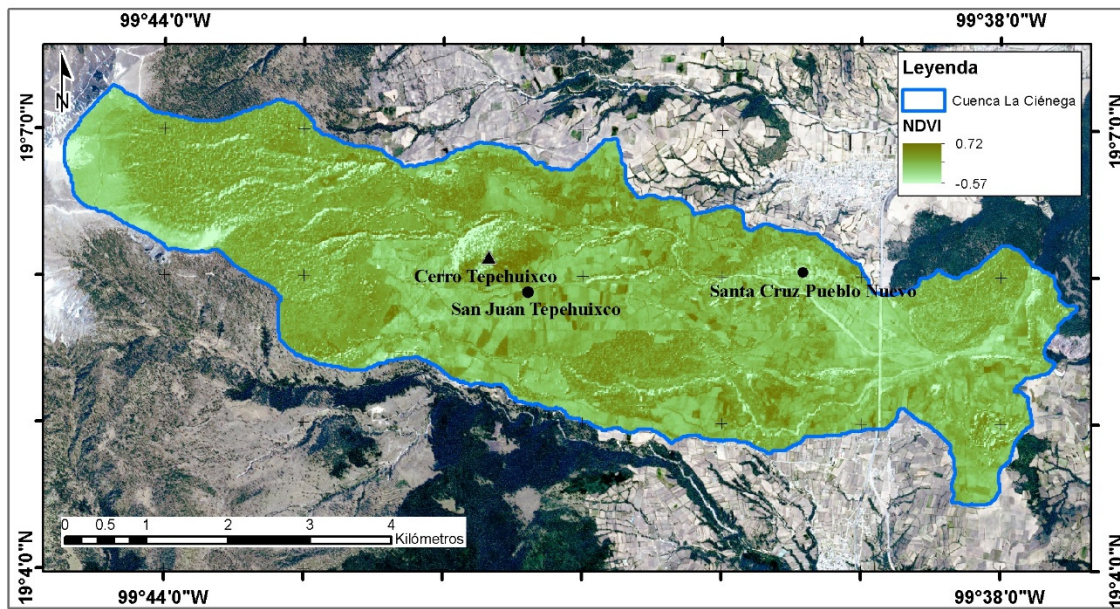


Figura 38. Mapa del índice de Vegetación de Diferencia Normalizada de la cuenca La Ciénega. Elaboración propia con información de SIAP-SEDENA, 2016.

3.3.3. Estandarización de las variables espaciales.

La totalidad de las variables dependiente y explicativas fueron estandarizadas a un formato *raster* con las siguientes características: Resolución espacial de 15 m, proyección cartográfica Universal Transversa de Mercator (UTM), unidades en metros, datum WGS 1984 y esferoide WGS 1984. Lo anterior con el propósito de hacerlas comparables. La resolución espacial se definió con un tamaño adecuado a la escala de origen de la mayoría de la información cartográfica utilizada como fuente de las variables a trabajar, siendo esta la escala 1:50,000.

Debido a que el análisis espacial que se realizó en la elaboración de los modelos de probabilidad es a nivel local (pixel), durante el proceso de rasterización de la información se cuidó que existiera una correspondencia precisa entre la ubicación de los pixeles de un tema a otro, con el propósito de asegurar que el cálculo entre mapas se realice de celda a celda (Figura 39), para ello se

especificó la generación de los mapas *raster* en función de un mapa base creado a partir de la información de límite de la cuenca.

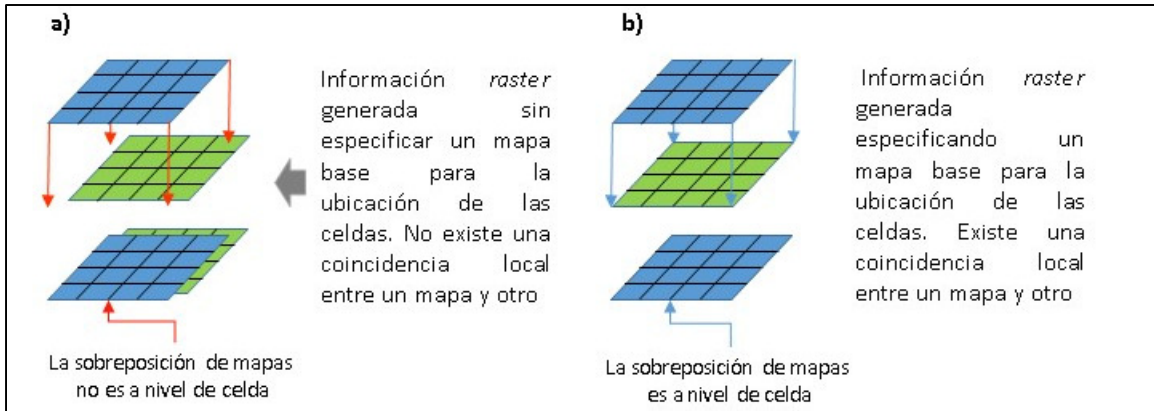


Figura 39. Generación de información *raster* con una correspondencia a nivel de local. Elaboración propia.

3.3.4. Tamaño de la muestra, muestra ampliada y tipo de muestreo.

Un aspecto importante para la elaboración de los modelos estadísticos de RL es lo referente al cálculo del tamaño de la muestra, el tipo de muestreo que se utilizará y la división de la muestra.

Tamaño de la muestra.

Los cálculos y la elaboración del modelo de probabilidad se realizaron a nivel de pixel por lo que la totalidad de pixeles que componen la cuenca La Ciénega fue considerada como la población en estudio, siendo esta mayor a 100,000 individuos. Por lo tanto, para el cálculo del tamaño de la muestra se utilizó la fórmula 1 correspondiente a una población infinita.

Los parámetros utilizados fueron: nivel de confianza del 95 %, margen de error del 5 % y para la proporción de ocurrencia se aplicó un valor del 50% para cada categoría de la variable dependiente, lo anterior debido a que la proporción de ocurrencia de los PG en la zona de estudio no se conoce con exactitud ya que el inventario realizado no cubre el 100% de los deslizamientos, por lo que se decidió asignar el porcentaje de proporción de ocurrencia máximo que puede presentar una variable dicotómica para cada categoría. A continuación, se presentan los parámetros aplicados a la Fórmula 1:

$Z = 1.96$ (coeficiente del nivel de confianza)

$p = 0.5$ (proporción de ocurrencia del valor 0 de la variable dependiente)

$q = 0.5$ (proporción de ocurrencia del valor 1 de la variable dependiente)

$d = 0.05$ (margen de error)

Quedando la siguiente fórmula y su correspondiente cálculo:

$$n = \frac{1.96^2 (0.5 * 0.5)}{0.05^2} = 384.16$$

El resultado del tamaño de la muestra es de 384 individuos.

Para el tamaño de la muestra, con respecto las 16 variables explicativas que son analizadas en el modelo de regresión, lo recomendable son 20 sitio por variable lo que da como resultado un total de 320 sitio. Por lo anterior se considera que con los 384 sitios calculados con base a la población se cubre el tamaño de la muestra necesaria para la calibración del modelo de RL en la cuenca La Ciénega.

Para el cálculo del tamaño de la muestra ampliada se considerará una proporción de 75 % de sitios para la calibración del modelo y un 25 % para la validación del mismo (muestra ampliada). Tomando los 384 sitios calculados en el tamaño de la muestra como el 75 % de sitios que serán utilizados para la calibración del modelo es necesario agregar el 25 % de sitios que corresponderán a la muestra ampliada. Lo anterior da un total de 512 sitios, de los cuales 384 (75 %) serán utilizados para la calibración y 128 sitios (25 %) servirán para la validación.

Con base a la proporción obtenida entre las áreas de las cabeceras y las áreas estables dentro de la zona de estudio, la cual es de 1 : 99, y con el propósito de darle una adecuada representatividad en la toma de la muestra a las áreas correspondientes a las cabeceras de los PG, que es la categoría de mayor importancia para la investigación, el levantamiento de los sitios de muestreo se realizó utilizando el 50 % para cada una de las categorías de la variable dependiente. Por lo tanto, el 100 % de la muestra (512 sitios) estará compuesta por 256 sitios ubicados en las áreas de las cabeceras y 256 sitios en las áreas estables. De tal forma que de los 384 sitios utilizados para la calibración del modelo 192 corresponden a zonas de deslizamientos y 192 a zonas estables y para la muestra de validación 64 se ubican en zonas de deslizamiento y 64 en zonas sin deslizamientos.

Tipo de muestreo

La selección de los 512 sitios de la muestra es realizada aplicando un muestreo aleatorio estratificado. La población fue dividida en dos estratos que corresponden a los valores de la variable dependiente, áreas de las cabeceras de los deslizamientos y áreas donde no se tiene registro de cabeceras y que son consideradas como áreas estables.

Espacialmente para la estratificación de la muestra, la cartografía digital correspondiente al inventario de deslizamientos se dividió en dos mapas, áreas estables y áreas de cabeceras (Figura 40).

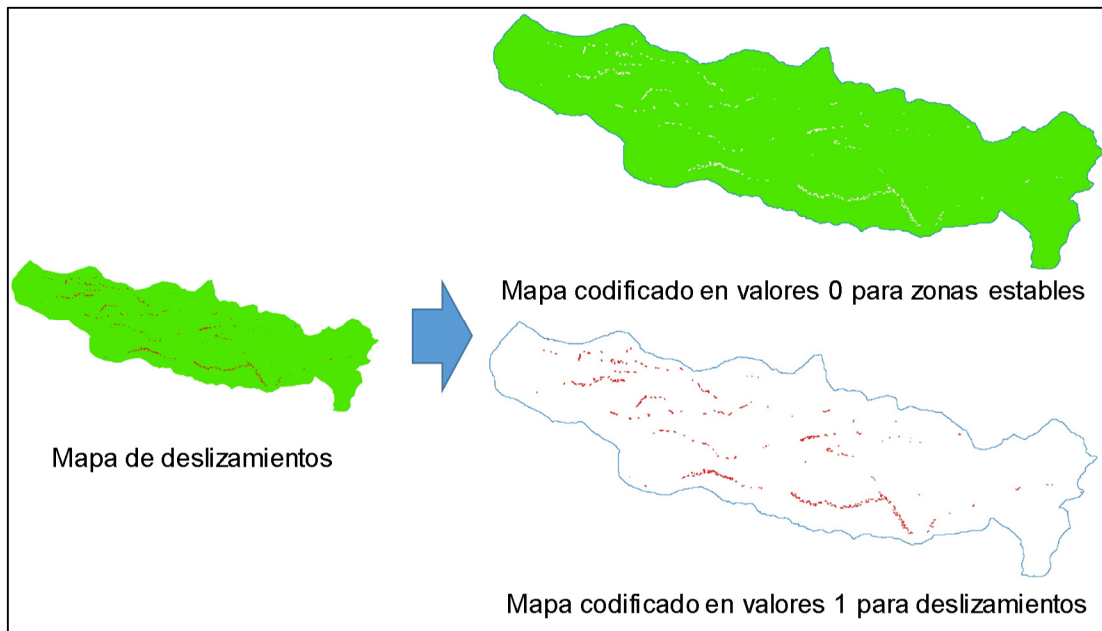


Figura 40. Estratificación de la variable dependiente: zonas estables (0) y zonas de deslizamiento (1). Elaboración propia.

Con los mapas correspondientes a cada uno de los valores de la variable dependiente se realizó la selección aleatoria del número de sitios de muestreo para cada estrato.

3.3.5. Evaluación de la colinealidad.

Esta evaluación se realizó a las 10 variables explicativas de tipo cuantitativo, para ello se adquirió de información de cada una de ellas a partir de la ubicación de los 384 sitios que son utilizados para la calibración del modelo de RL (Figura 41).

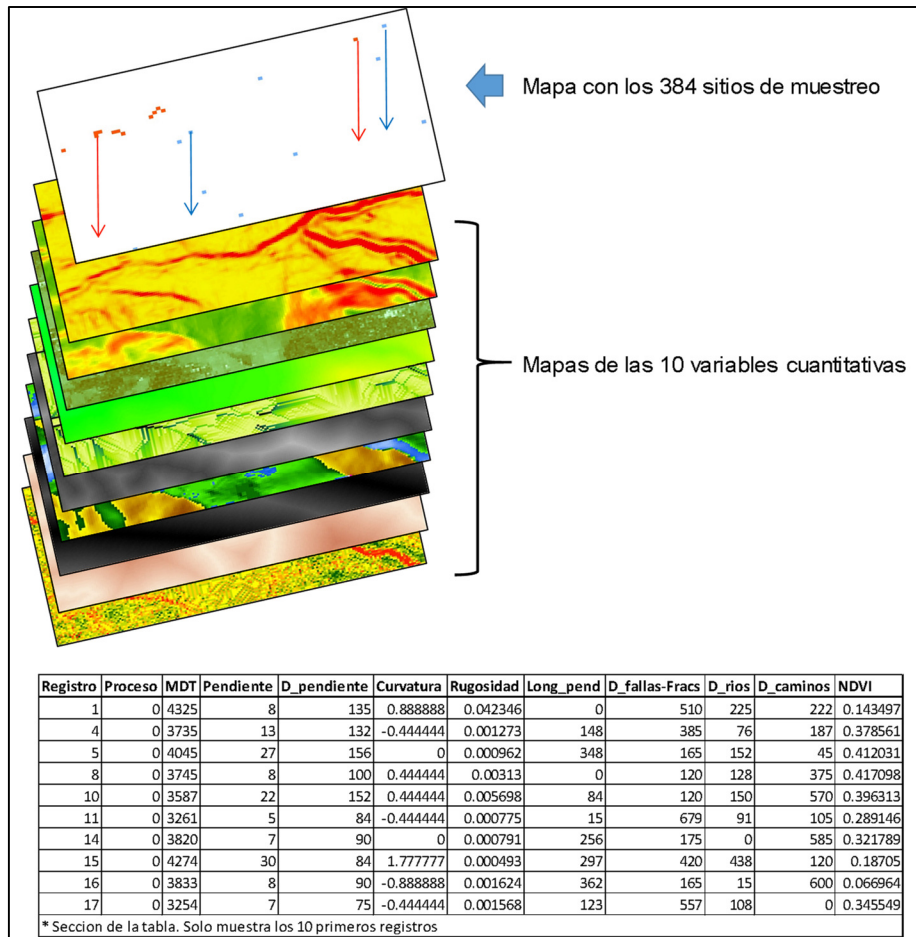


Figura 41. Toma de datos de las 10 variables cuantitativa para realizar el análisis de colinealidad. Elaboración propia.

A la matriz de datos generada del cruce de los sitios de muestreo con las variables cuantitativas se aplicaron los análisis para la evaluación de la Tolerancia, el FIV y el Índice de Condicionamiento, utilizando el programa SPSS® ver 18.0.

3.3.6. Aplicación del Análisis Espacial de Vecindad Continua.

3.3.6.1. A las variables cualitativas.

Para las 6 variables cualitativas (uso de suelo y vegetación 2015, cambios de uso de suelo y vegetación 1983-2015, geomorfología, litología, edafología y textura del suelo) el AEVC se implementó mediante el siguiente procesamiento con el propósito de que la información fuese compatible con el tipo de análisis espacial y estadístico que se le aplicó.

Estas variables se analizaron con base en su presencia o ausencia alrededor de los sitios de muestreo. Para ello se generó un mapa para cada una de las clases que componen cada uno de los mapas. Los mapas que representan cada una de las clases fueron clasificada con valores de 1 para las áreas que representan la clase en turno y con un valor de 0 para las áreas que están cubiertas por las otras clases. Por ejemplo, la Figura 42 muestra la separación del mapa de litología en 6 mapas individuales, iniciando con la clase de extrusiva intermedia donde las áreas correspondientes a esta categoría se codifican con el valor de 1 y el resto de la superficie se codifica con el valor 0, así sucesivamente hasta generar los mapas individuales para las clases que componen el mapa de litología.

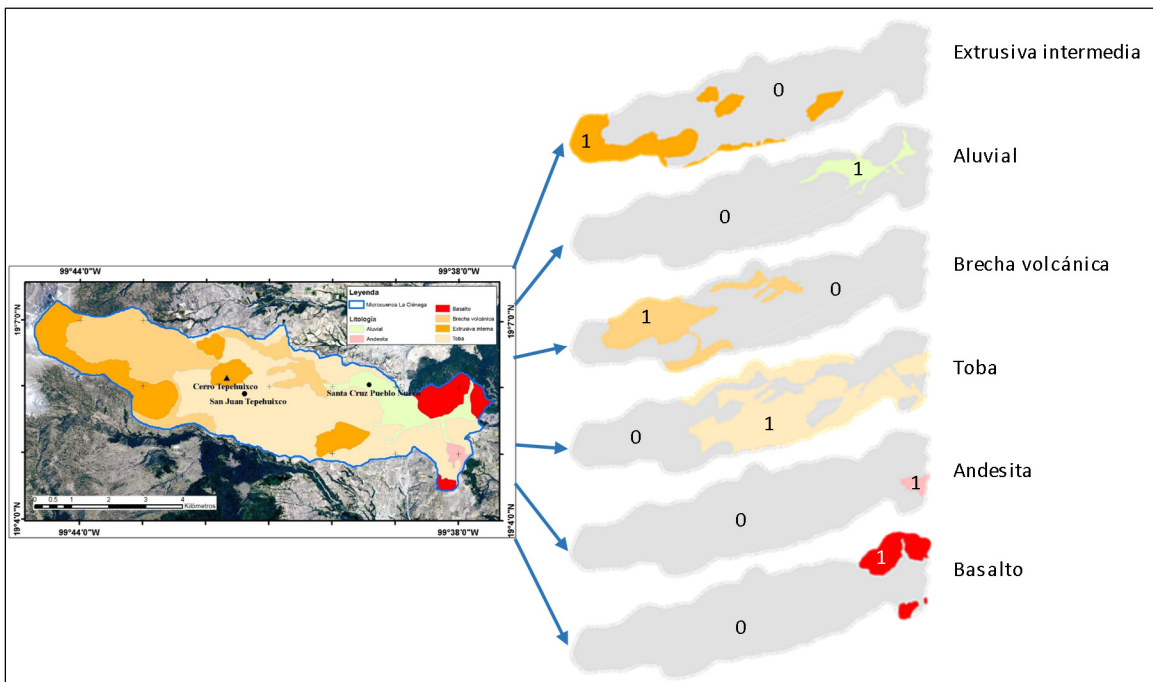


Figura 42. División del tema cualitativo de Litología en mapas individuales, un mapa para cada clase que lo conforma. Elaboración propia.

Mediante este proceso se obtuvieron un total 47 mapas cualitativos de acuerdo al número de clases que componen cada mapa: 9 de uso de suelo y vegetación 2015, 3 de cambios de uso de suelo y vegetación 1983-2015, 20 de geomorfología, 6 de litología, 7 de edafología y 2 de textura del suelo. Estos 47 mapas fueron procesados posteriormente con el AEVC.

Los mapas que representan una categoría de un determinado tema se trabajaron de manera individual haciendo el cruce de los 384 sitios de la muestra para la calibración del modelo con cada uno de estos. Posteriormente, a cada cruce le fue aplicado el análisis de vecindad, iniciando

a un radio de 1 pixel y aumentando de 1 en 1 pixel hasta llegar a los 20 pixeles, lo que da como resultado un AEVC que va de 1 pixel hasta los 20 pixeles (Figura 43).

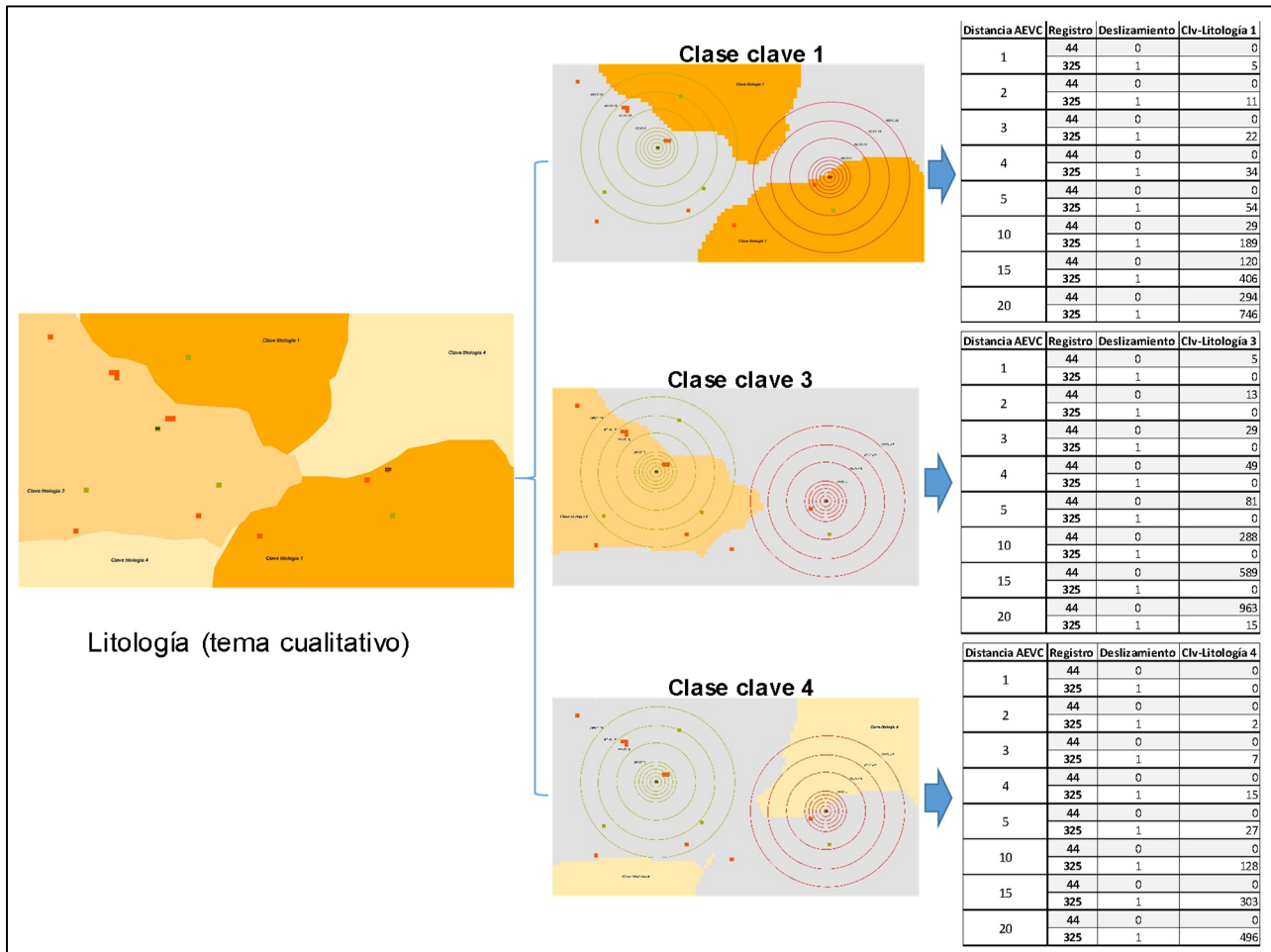


Figura 43. Adquisición de datos para las áreas de vecindad correspondientes a 1, 2, 3, 4, 5, 10, 15 y 20 pixeles de radio, para un sitio estable (sitio 44) y un sitio de deslizamiento (sitio 325). Elaboración propia.

Para el cálculo de información dentro del área de vecindad se aplicó la función de suma para cuantificar los pixeles que se encontraron dentro del área analizada, el resultado fue asignado como valor del sitio de muestreo del que se trate (Tabla 11). Con ello se evaluó la presencia del número de pixeles de una determinada clase de litología alrededor de cada uno de los sitios de muestreo.

Una vez que se obtuvieron las 20 tablas, correspondientes a los tamaños del área de vecindad por clase de litología se realizó la integración de una matriz única que contiene los datos de las 6 clases correspondientes (continuando con el tema de litología) a una sola distancia del AEVC. Ejemplo de esta matriz es la Tabla 12 que representa la integración de datos del tema litología a

una distancia de 1 pixel del AEVC. En esta matriz se presenta la cuantificación del número de pixeles por clase temática que rodean a cada sitio en el área correspondiente a un pixel de radio.

Tabla 11. Sumatoria de pixeles de la clase con clave 1 de litología a diferentes distancias del AEVC para los sitios de muestreo 44 y 325.

Distancia AEVC	Registro	Proceso	Litología-Clave1
1	44	0	0
	325	1	5
2	44	0	0
	325	1	11
3	44	0	0
	325	1	22
4	44	0	0
	325	1	34
5	44	0	0
	325	1	54
10	44	0	29
	325	1	189
15	44	0	120
	325	1	406
20	44	0	294
	325	1	746

Elaboración propia.

Tabla 12. Integración de la sumatoria de las 6 clases del mapa de litología a una distancia de 1 pixel de distancia del área de vecindad. La tabla muestra solo 10 registros de los 384.

Registro	Proceso	Litología-Clave1	Litología-Clave2	Litología-Clave3	Litología-Clave4	Litología-Clave5	Litología-Clave6
1	0	5	0	0	0	0	0
44	0	0	0	5	0	0	0
5	0	5	0	0	0	0	0
257	1	0	0	5	0	0	0
259	1	0	0	5	0	0	0
260	1	0	0	5	0	0	0
8	0	0	0	5	0	0	0
262	1	0	0	5	0	0	0
325	1	5	0	0	0	0	0
264	1	0	0	5	0	0	0

Elaboración propia.

El concentrado de información por mapa temático correspondiente a una distancia del AEVC (Tabla 12) fue procesado en el programa SPSS® 18.0 aplicando el modelo de RL. Como los valores en la tabla representan cantidad de pixeles de una determinada clase alrededor de los sitios de muestreo estos datos fueron procesados como variables cuantitativas para evaluar la relación de la variable en cuestión para cada una de las distancias trabajadas (1 a 20 pixeles). La aplicación de la RL reportó para cada una de las distancias analizadas el valor correspondiente a -2LL.

Con el propósito de evaluar la relación que existe entre el proceso de deslizamientos y la información de las variables cualitativas a 0 pixeles de distancia (datos *in situ*) se realizó el cruce de los sitios de la muestra para calibración del modelo (384 sitios) con el mapa cualitativo que integra la totalidad de las clases que lo componen. En la Figura 44 se muestra el ejemplo de este procesamiento aplicado al mapa de litología.

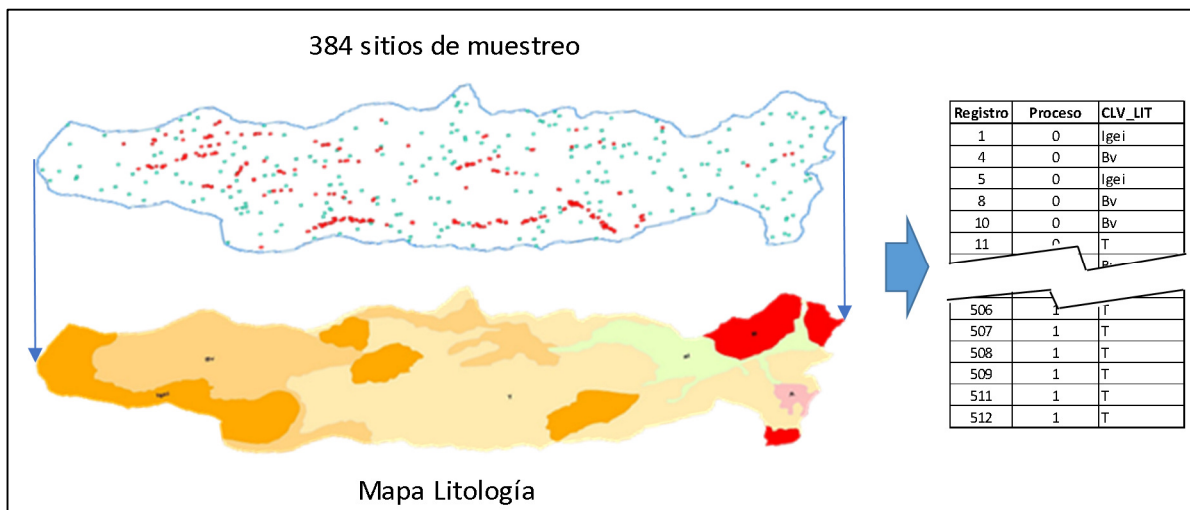


Figura 44. Toma de información *in situ* para las variables cualitativas. Elaboración propia.

Con ello se obtuvieron datos de tipo cualitativo correspondientes a las clases que integran la variable de litología, los cuales fueron analizados con la RL convirtiendo estas clases categóricas a variables ficticias, del resultado obtenido se tomó el valor reportado por el indicador -2 LL.

Por último, se realizó la integración de los valores obtenidos para el indicador de -2 LL correspondientes al AEVC para las 20 áreas analizadas, así como el obtenido del análisis *in situ*, con ello se obtuvo la información correspondiente a 21 valores del indicador de -2 LL que cubren desde una distancia de 0 (*in situ*) hasta el radio de 20 pixeles del análisis de vecindad (Figura 45).

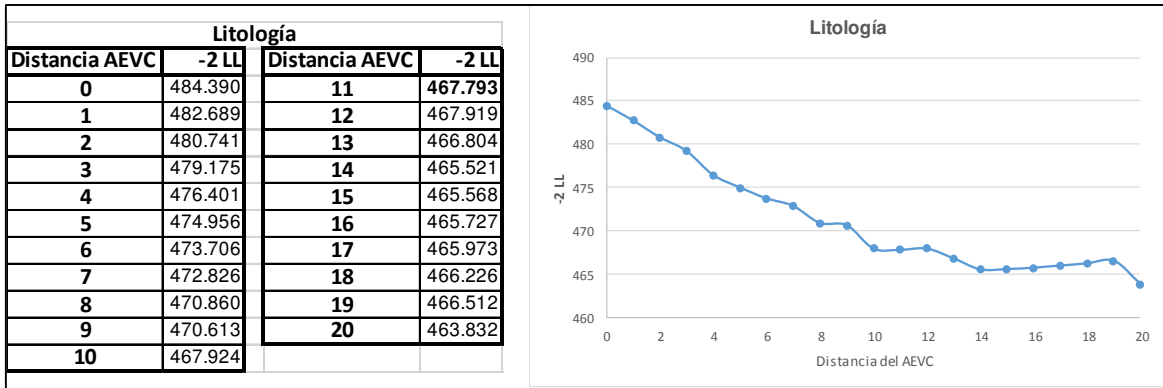


Figura 45. Valores de -2 LL correspondientes al análisis de RL *in situ* y las 20 áreas de vecindad y su correspondiente gráfica. Ejemplo tema litología. Elaboración propia.

Con los valores de -2 LL se determinó la distancia en la que se tiene una mayor relación del proceso de deslizamientos con la variable, con base a que a menor valor de -2 LL se tiene un mejor ajuste predictivo del modelo de RL con el proceso en estudio.

La aplicación del AEVC y su análisis estadístico se realizó en las 6 variables cualitativas. Se generaron un total de 946 tablas individuales correspondientes a cada una de las clases que integran los temas cualitativos y sus respectivos análisis de vecindad.

3.3.6.2. A las variables cuantitativas.

La implementación del AEVC a las variables cuantitativas se realizó de la siguiente manera:

Para todos los mapas temáticos cuantitativos se trabajaron los datos con el nivel de precisión con el que fueron generados por las funciones utilizadas para su cálculo, a fin de contar con la información de mayor detallada disponible.

Para las 10 variables de tipo cuantitativos la función aplicada para el cálculo de la información dentro de las áreas de vecindad fue la de promedios de los valores que se encontraron dentro de cada una de las 20 distancias analizadas con el AEVC (Figura 46).

El AEVC se realizó de la misma manera que con las variables cualitativas, se inició con la intersección de los 384 sitios de muestreo con cada uno de los mapas de las variables cuantitativas, posteriormente a cada intersección de sitios con una determinada variable se aplicó el análisis de vecindad para las distancias de 1 a 20 pixeles de diámetro. El AEVC dio como resultado un total de 20 tablas de datos para cada mapa temático, una para cada distancia analizada. Se obtuvieron un total de 200 tablas de datos producto de la aplicación del AEVC a

las 10 variables cuantitativas las cuales fueron analizadas aplicando la RL, posteriormente se extrajo el valor correspondiente a -2 LL de cada tabla procesada.

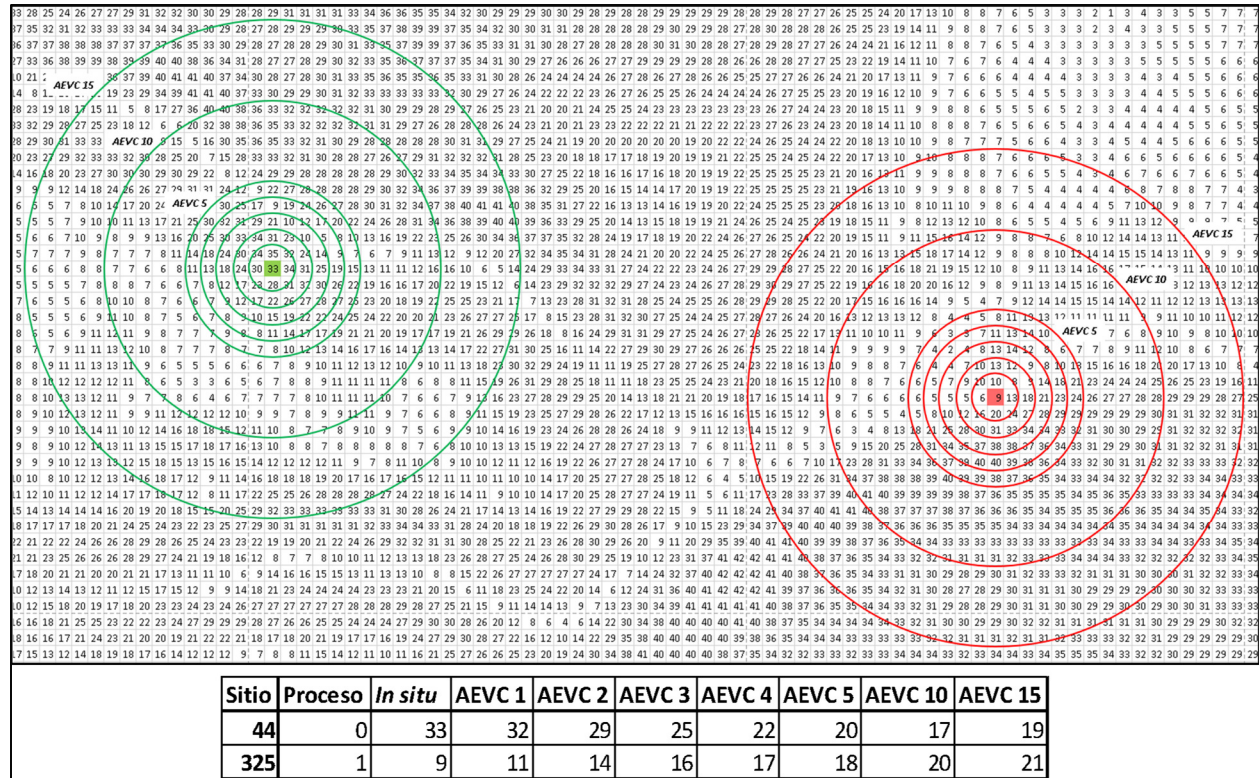


Figura 46. Valores de la variable cualitativa (pendiente del terreno) para los sitios de muestreo 44 y 325. Los promedios obtenidos en cada AEVC se asignaron al sitio. Elaboración propia.

Con respecto a la evaluación de la toma de información de datos de manera *in situ* para estas variables cualitativas se realizó la intersección del mapa correspondiente a los 384 sitios utilizados para la calibración del modelo con los mapas de cada variable. Una vez obtenida la tabla con información del sitio esta fue procesada con la RL para obtener el valor del indicador de -2 LL.

Los valores de -2 LL producto del análisis *in situ* y de la aplicación del AEVC, distancias de 1 a 20 pixeles de radio, fueron integrados a una tabla con cual se pudo determinar la distancia con una mayor relación entre el proceso de deslizamientos con base al menor valor de -2 LL (Figura 47).

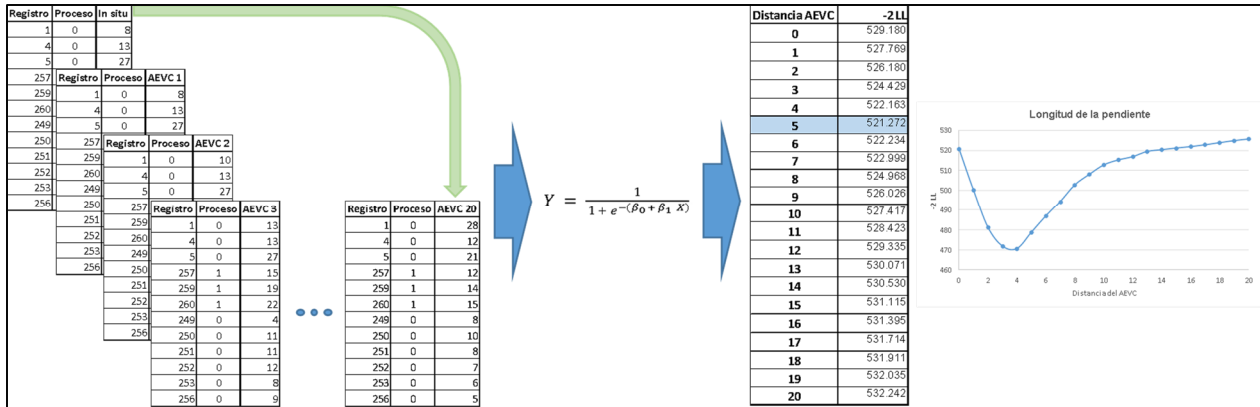


Figura 47. Procesamiento de las tablas obtenidas de la toma de información *in situ* y el AEVC con la RL para identificar la distancia a la que el valor de -2 LL es menor. Elaboración propia.

3.4. Modelos espaciales de probabilidad a deslizamientos.

Se elaboraron dos modelos de probabilidad, el primero se realizó con base al AEVC, el segundo se realizó con la toma de información *in situ*, la elaboración de este segundo modelo fue para poder evaluar y comparar los resultados de ambos.

La elaboración de los modelos se dividió en tres etapas: La primera fue la integración de una matriz con los datos que contiene la totalidad de la información de las 16 variables explicativas. Para el caso del modelo con base al AEVC la selección de la información de las 16 variables explicativas se realizó con base a la distancia que haya dado el menor valor de -2 LL en los resultados del AEVC (Figura 48). Mientras que para el caso del modelo *in situ*, la matriz con la información de las 16 variables explicativas fue integrada con las tablas obtenidas directamente del cruce de los 384 sitios de muestreo con cada uno de los mapas correspondientes a las variables explicativas.

La segunda etapa fue la aplicación del modelo de RL a las tablas que integra la información de las 16 variables. Ambas tablas (con base en AEVC e *in situ*) fueron trabajadas con el programa estadístico SPSS® 18.0 aplicando la función de Regresión Logística Binaria con el método adelante razón de verosimilitud. Para el análisis de las variables en la RL estas fueron definidas de acuerdo al tipo de dato que representan, es decir, para las variables cualitativas en el modelo con base a los datos *in situ* estas fueron trabajadas como variables categóricas, mientras que para el modelo con base en el AEVC las variables cualitativas fueron trabajadas como cuantitativas, pues como se mencionó anteriormente las variables categóricas en el análisis de vecindad fueron procesadas de tal manera que se cuantificó la cantidad de pixeles de cada una de las clases alrededor de cada sitio.

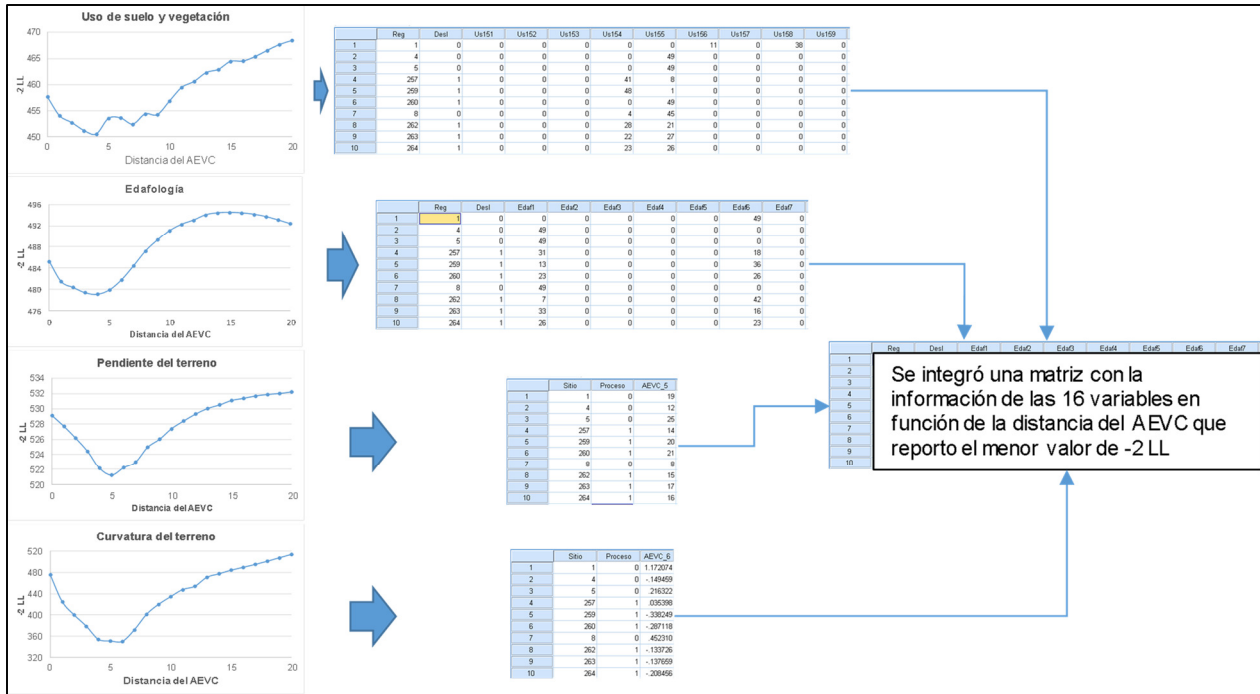


Figura 48. Integración de la matriz con los datos seleccionados en el AEVC. Elaboración propia.

En lo que respecta a las variables de tipo cuantitativo todas estas fueron trabajadas como variables métricas dentro del análisis de RL para ambos modelos (*in situ* y AEVC).

La tercera etapa fue la elaboración del mapa de probabilidad a deslizamientos. Con base en los resultados de los modelos estadísticos, se seleccionaron las variables que son significativas para formar parte del modelo. Para aplicar la información del modelo estadístico en la elaboración del modelo espacial los mapas de las variables explicativas se trabajaron de la siguiente manera: Para el modelo con base al AEVC cuando se trató un mapa temático cualitativo en función de una distancia de vecindad se identificó cuál o cuáles de las clases que componen el mapa quedaron como significativas, para las clases seleccionadas se retomaron los mapas generados para la aplicación del AEVC a las variables cualitativas (Figura 42), posteriormente, se aplicó el análisis de vecindad con la función de sumatoria a la distancia del AEVC que fue seleccionada, este cálculo se realizó sobre toda la superficie de la cuenca (Figura 49).

Para las variables cuantitativas que quedaron como significativas para formar parte del modelo con base en el AEVC, a sus respectivos mapas temáticos les fue aplicado el análisis de vecindad con la función de promedio a la distancia seleccionada durante el análisis del AEVC.

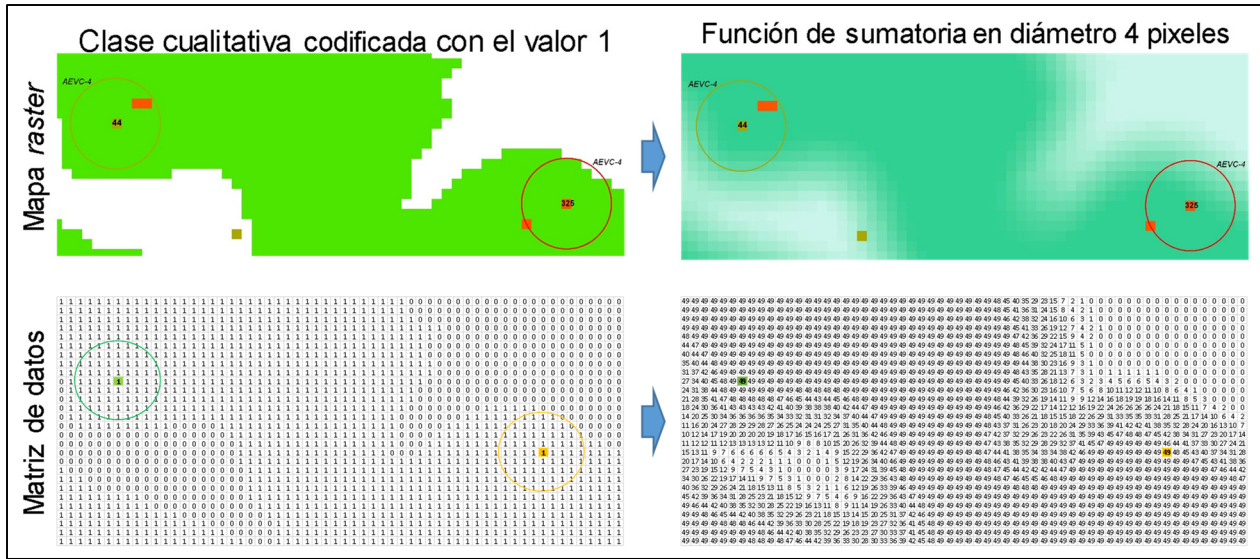


Figura 49. Representación de una variable espacial cualitativa en su forma original y en función de un tamaño de vecindad de 4 pixeles, en los formatosde mapa *raster* y matriz de datos. Elaboración propia.

Para el modelo espacial con base en datos *in situ* se utilizaron como variables explicativas los mapas temáticos que resultaron ser significativos en el modelo estadístico correspondiente. Estos mapas corresponden a los descritos en la sección 3.3 descripción y procesamiento de las variables.

Cada modelo de probabilidad a deslizamientos (AEOV e *in situ*) se calculó utilizando los mapas de las variables seleccionadas en el modelo estadístico junto con sus coeficientes aplicando la fórmula de RL con más de una variable explicativas (Fórmula 4) (Figura 50).

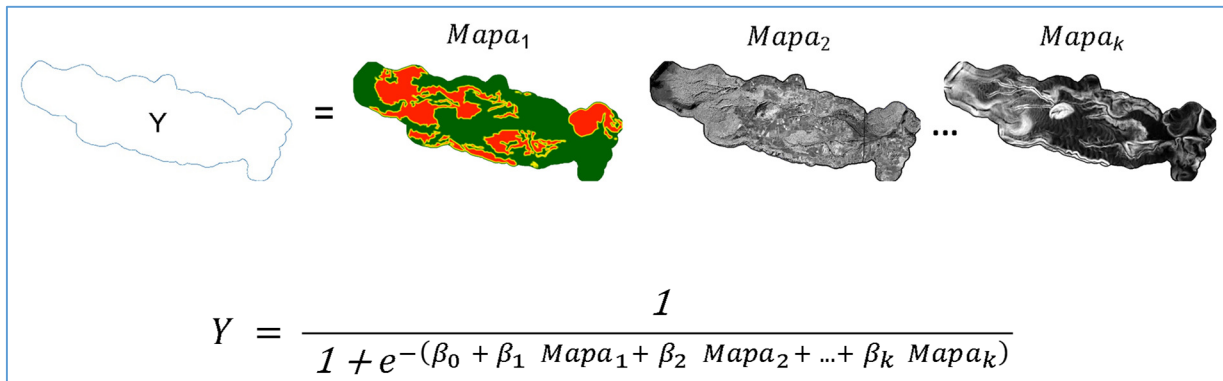


Figura 50. Calculo espacial del modelo de probabilidad con la fórmula de RL. Elaboración propia.

3.5. Evaluación de los modelos de probabilidad de AEVC e *in situ*.

Una vez obtenidos los dos modelos espaciales (AEVC e *in situ*) se realizó la evaluación de cada uno de ellos utilizando la curva COR y la matriz de clasificación.

Para la evaluación en función de la curva COR se usaron los 128 sitios de la muestra extendida, se realizó el cruce de estos con cada uno los modelos espaciales de probabilidad (AEVC e *in situ*). Del cruce se obtuvo una tabla que contiene los datos de los sitios con la información real, si se trata de un sitio estable (valor 0) o un sitio de deslizamiento (valor 1) y el valor de probabilidad del modelo (valores entre 0 y 1). Con esta información fue posible generar el análisis de la curva COR para cada uno de los modelos (Figura 51).

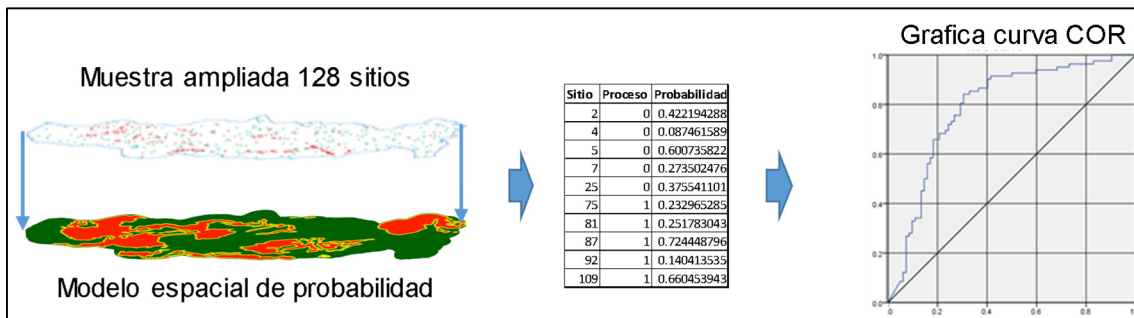


Figura 51. Generación de la curva COR utilizando la muestra ampliada y el modelo de probabilidad (AEVC e *in situ*). Elaboración propia.

Evaluación de los modelos espaciales de probabilidad utilizando la matriz de clasificación. Para la generación de la matriz de clasificación se utilizó la muestra ampliada y los mapas de probabilidad obtenidos, estos últimos ya reclasificados en valores únicos de 0 y 1 correspondientes a zonas estables y zonas de deslizamiento, respectivamente (Figura 52).

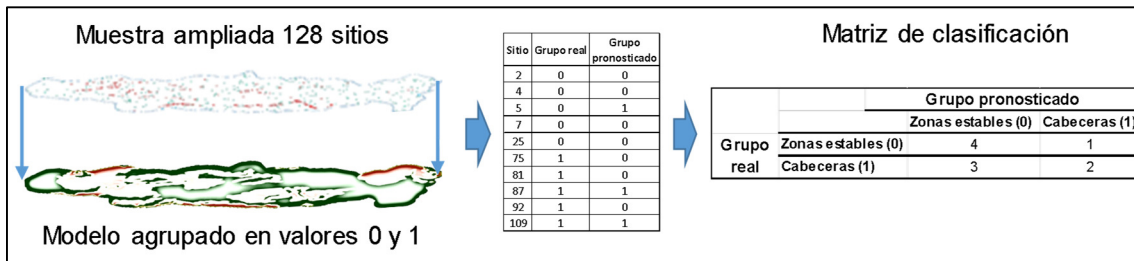


Figura 52. Generación de la matriz de clasificación con base en la muestra ampliada y los modelos de probabilidad agrupados en los valores de 0 y 1. Elaboración propia.

Con el cruce de la información espacial (muestra ampliada y modelo de probabilidad clasificado en 0 y 1) se generaron las matrices de clasificación correspondientes para los modelos de AEVC

e *in situ*. A partir de las matrices se realizaron el cálculo del índice de Kappa y las evaluaciones de exactitud general, exactitud del productor, exactitud del usuario y eficiencia del modelo.

CAPÍTULO 4. INVENTARIOS DE DESLIZAMIENTOS Y MODELOS ESPACIALES DE PROBABILIDAD A DESLIZAMIENTOS DE LA CUENCA LA CIÉNEGA.

A continuación, se presentan los principales resultados en los que se incluye el mapa del inventario de deslizamientos, así como los correspondientes al tema principal de la investigación sobre los modelos de probabilidad a deslizamientos.

4.1. Inventario de deslizamientos.

La información de campo fue trabajada con base en la clasificación para PG del Departamento de Recursos Naturales del Estado de Washington (WSDNR, 2006a). En la zona se encontraron los siguientes tipos de deslizamientos: Deslizamientos superficiales de escombros (DSE), Deslizamientos superficiales de flujo de escombros (DSFE) y deslizamientos de asentamiento profundo (DAP). En la Figura 53 se presentan ejemplos de deslizamientos superficiales y en la Figura 54 de deslizamiento de asentamiento profundo.

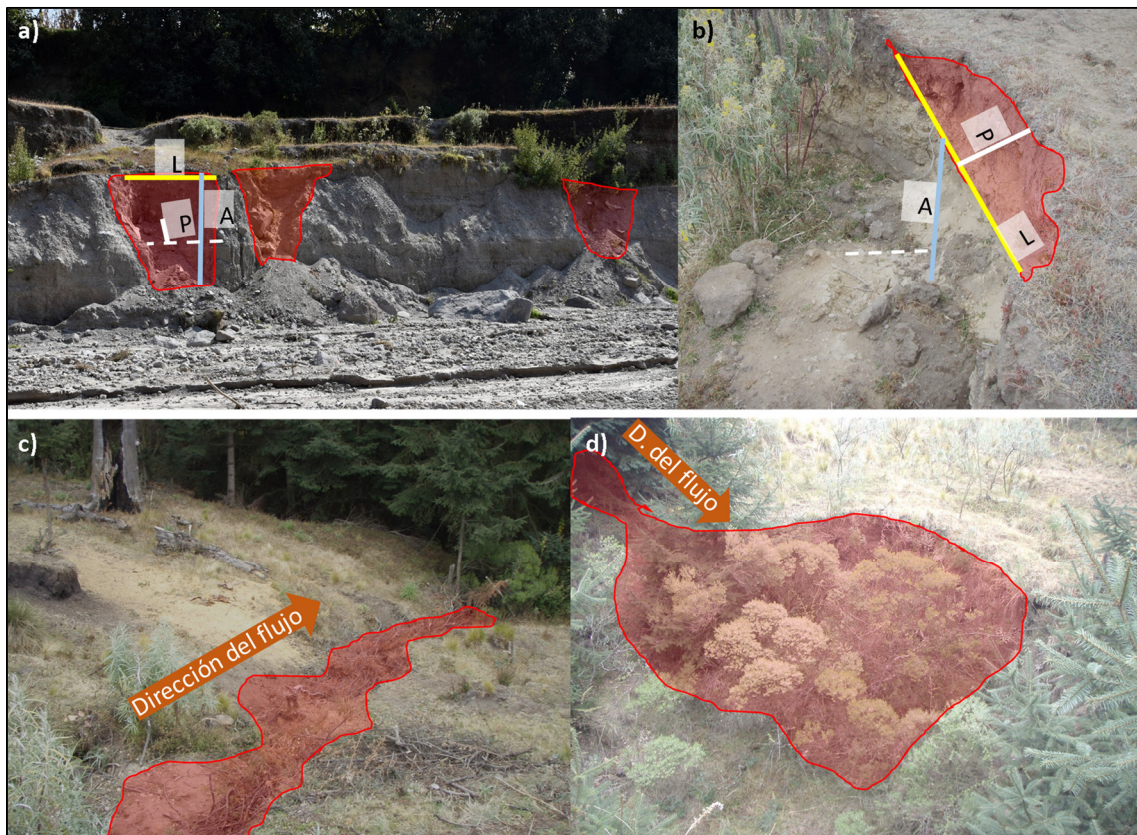


Figura 53. a) medición de un DSE y b) vista en planta en la que se resalta la cabecera del DSE. c) canal por el que fluyen los escombros en un DSFE y d) área de desembocadura del DSFE. Elaboración propia con información de campo.



Figura 54. DAP en el que se aprecia que el plano de deslizamiento con una altura de 9 metros está por debajo de las raíces de los árboles. Elaboración propia con información de campo.

Como resultado de la integración de los inventarios 1984-2014 y 2015-2016 se cartografió un total de 290 deslizamientos de los cuales 250 corresponden a DSE, 33 a DSFE y 7 a DAP. En la Figura 55 se muestra la distribución y el porcentaje de los deslizamientos inventariados en sus diferentes tipos.

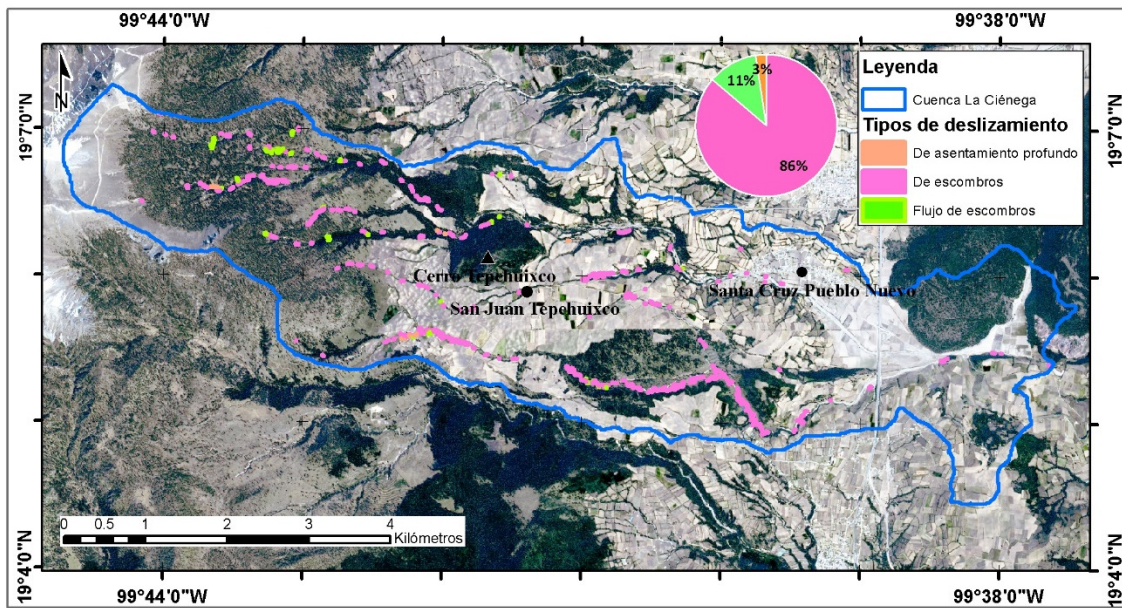


Figura 55. Mapa del inventario de deslizamientos 1984-2016. Por tipo de deslizamientos y su respectivo porcentaje. Elaboración propia con información de campo.

4.2. Modelos de probabilidad a deslizamientos.

Los resultados obtenidos en la elaboración de los modelos de probabilidad se presentan en las 4 secciones siguientes.

4.2.1. Colinealidad.

En la tabla 13 se presentan los resultados de Tolerancia y el FIV producto del análisis de colinealidad entre las 10 variables cuantitativas.

Tabla 13. Valores de Tolerancia y FIV.

Modelo	Estadísticas de colinealidad	
	Tolerancia	VIF
Altitud en metros s.n.m.	.474	2.112
Pendiente del terreno	.693	1.444
Dirección de la pendiente del terreno	.874	1.145
Curvatura de la superficie del terreno	.682	1.467
Rugosidad del terreno	.609	1.643
Longitud de la pendiente	.840	1.190
Distancia a fallas y fracturas	.804	1.243
Distancia a ríos	.775	1.291
Distancia a caminos	.525	1.904
NDVI	.794	1.260

Elaboración propia.

Los valores obtenidos para la tolerancia están por arriba de 0.10, por lo tanto, no existe colinealidad entre las variables. Lo mismo ocurre usando el FIV donde los valores calculados están por debajo de 10.

El otro método utilizado para evaluar la colinealidad fue el Índice de Condicionamiento. La Tabla 14 presenta los valores obtenidos para este índice.

Tabla 14. Resultados del Índice de Condicionamiento y la proporción de la varianza.

Índice de condición	Prporción de la varianza										
	(Constant)	Altitud m.s.n.m.	Pendiente del terreno	Dirección de la pendiente del terreno	Curvatura de la superficie del terreno	Rugosidad del terreno	Longitud de la pendiente	Distancia a fallas y fracturas	Distancia a ríos	Distancia a caminos	NDVI
1.000	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
2.053	.00	.00	.00	.00	.15	.06	.04	.02	.06	.00	.00
3.009	.00	.00	.02	.00	.00	.05	.63	.03	.01	.01	.00
3.158	.00	.00	.01	.00	.31	.00	.10	.05	.06	.16	.00
4.015	.00	.00	.00	.01	.12	.24	.05	.17	.39	.05	.02
4.074	.00	.00	.00	.13	.11	.00	.02	.02	.34	.23	.03
4.741	.00	.00	.01	.03	.26	.49	.00	.51	.08	.00	.02
5.683	.00	.00	.75	.20	.00	.01	.01	.00	.00	.16	.00
6.253	.00	.00	.11	.51	.01	.05	.00	.11	.01	.04	.33
8.391	.02	.01	.01	.04	.03	.07	.13	.08	.04	.00	.59
50.372	.98	.98	.09	.06	.00	.00	.02	.00	.00	.33	.00

Elaboración propia.

Para determinar si existe correlación entre variables, primero, se identificaron valores del índice de condicionamiento mayores a 15. En este estudio se obtuvo un registro con un valor de 50.3

(último renglón tabla 14), enseguida se buscó dentro de este renglón si más de una variable presenta valores mayores a 0.9 en la proporción de la varianza. Como se puede observar solo la variable de altitud (Altit_d0) reporta un valor de 0.98, ninguna otra variable explicativa presenta un valor cercano o mayor a 0.9, por lo anterior se concluye que no se presenta colinealidad entre variables explicativas.

Dado que no existe colinealidad entre las 10 variables cuantitativas todas estas fueron analizadas con el AEVC y utilizadas para la calibración del modelo de probabilidad a deslizamiento con información *In situ*.

4.2.2. Modelos estadísticos con base en el AEVC y con información *in situ*.

Con base en el Análisis Espacial de Vecindad Continua.

En la figura 56 se muestran los resultados del AEVC para las 16 variables explicativas con respecto al indicador -2 LL (Figuras 56a a 56o). Con base en estos se realizó la identificación del valor mínimo de -2 LL y su distancia correspondiente en el AEVC. De esta manera se seleccionó la distancia óptima del AEVC para cada variable, la cual fue integrada en la calibración del modelo de probabilidad a deslizamientos.

Las distancias seleccionadas para cada variable explicativas fueron las siguientes: uso de suelo y vegetación a la distancia de 4 píxeles (Figura 56a), cambio de uso de suelo y vegetación a la distancia de 15 píxeles (Figura 56b), geomorfología a la distancia de 20 píxeles (Figura 56c), litología a la distancia de 20 píxeles (Figura 56d), edafología a la distancia de 4 píxeles (Figura 56e), textura del suelo a la distancia de 20 píxeles (Figura 56f), altitud a la distancia de 20 píxeles (Figura 56g), pendiente del terreno a la distancia de 5 píxeles (Figura 56h), dirección de la pendiente a la distancia de 20 píxeles (Figura 56i), curvatura del terreno a la distancia de 6 píxeles (Figura 56j), rugosidad del terreno a la distancia de 2 píxeles (Figura 56k), longitud de la pendiente a la distancia de 4 píxeles (Figura 56l), distancia a fallas y fracturas a la distancia de 1 píxeles (Figura 56m), distancia a ríos a la distancia de 2 píxeles (Figura 56n), distancia a caminos a la distancia de 9 píxeles (Figura 56ñ) y NDVI a la distancia de 20 píxel (Figura 56o).

Las matrices de los datos seleccionadas fueron integradas en una sola tabla, esta última fue procesada estadísticamente con la función de RL para obtener los resultados del modelo de probabilidad. Con el propósito de generar un modelo estadístico que solo contenga las variables significativas, el modelo logístico se calculó con la opción hacia adelante razón de la verosimilitud.

Las Tablas 15a a la 15e muestran los resultados obtenidos. En un primer procesamiento de la información con la RL, la tabla 15a muestra que la variable de uso de suelo y vegetación 2015 correspondiente a la clase bosque semidenso (Us154) reporta un valor de significancia mayor a 0.05 que es límite máximo considerado para que una variable sea mantenida en la ecuación (marcada en color naranja). Por lo tanto, se realizó un segundo modelo de RL en que solo se integran las variables que reportan una significancia menor o igual a 0.05, siendo estas variables las correspondientes a la clase deforestación del tema de Cambios de uso de suelo y vegetación (Cam3), las clases del tema de geomorfología de ladera superior (Geom6), lomeríos de avalancha de escombros cubiertos de piroclastos (Geom11) y lomeríos de pómez y depósitos de caída, (Geon12), la clase de textura del suelo gruesa (text1), la curvatura del terreno (curv_d6) y la distancia a ríos (drio_d2). Los resultados de este segundo procesamiento se muestran en las Tablas 15b a 15e donde se incluye información sobre el resumen del modelo en el que se reporta el valor mínimo alcanzado por -2 LL (Tabla 15b). La tabla de clasificación de la variable dependiente en la que se presentan los porcentajes de correcta clasificación alcanzada por el modelo siendo esta de 89.6 % para las clases de estable y deslizamiento, así como para el porcentaje global (Tabla 15c). La tabla de contingencia de la prueba de Hosmer y Lemeshow con la información del muestreo en 10 grupos con número de muestras muy cercano entre todos ellos, de los cuales se presentan el número de valores observados y esperados para las clases de la variable dependiente (estables 0 y deslizamientos 1) (Tabla 15d). Por último, la información acerca de las variables en la ecuación, en la que se puede observar que todas las variables reportan un valor de significancia menor a 0.05, estas variables y sus correspondientes coeficientes (β) son las que formaran parte del modelo final del AEVC (Tabla 15e).

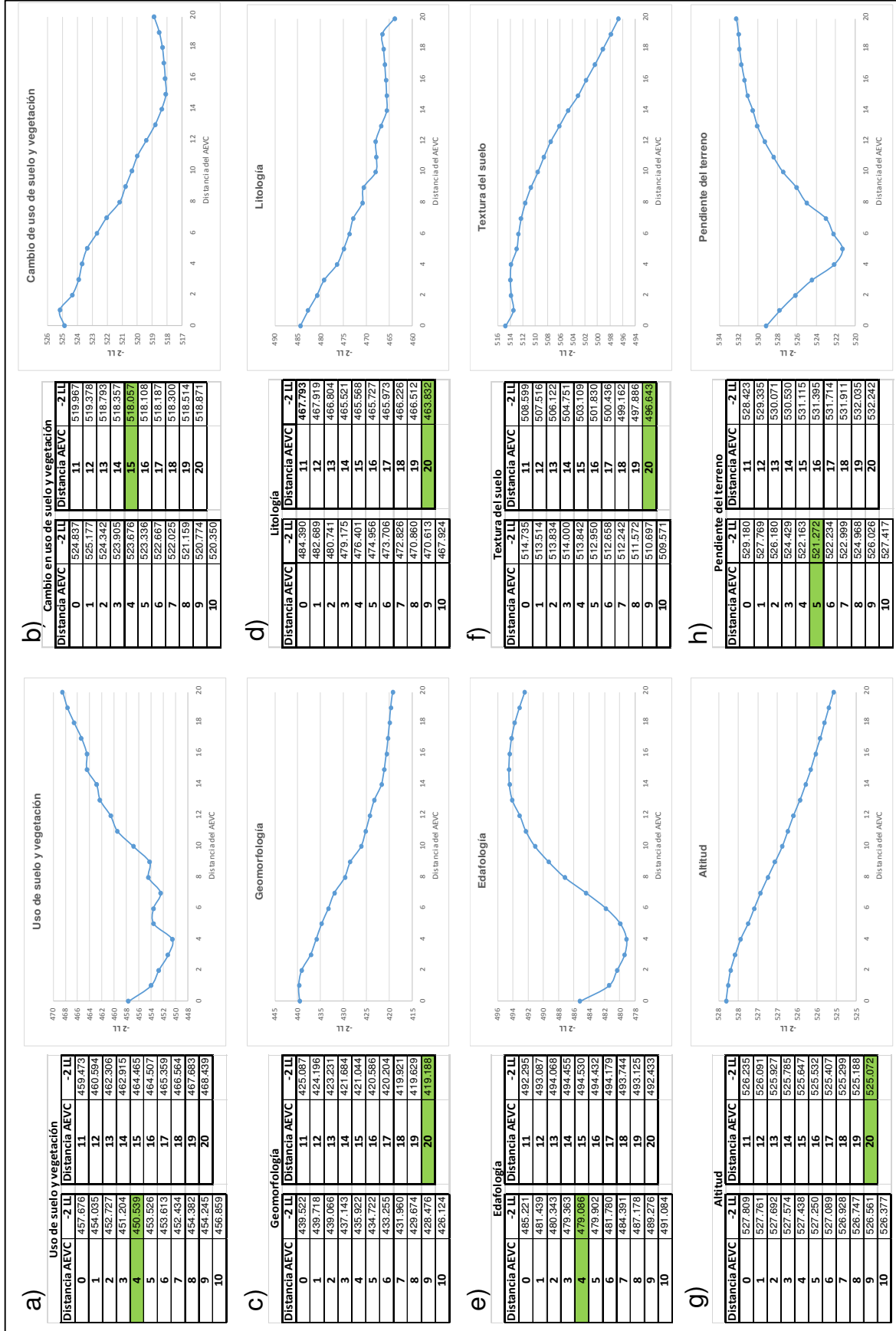


Figura 56 Resultados del AEC por variable explicativas (continúa en la siguiente página). Elaboración propia.

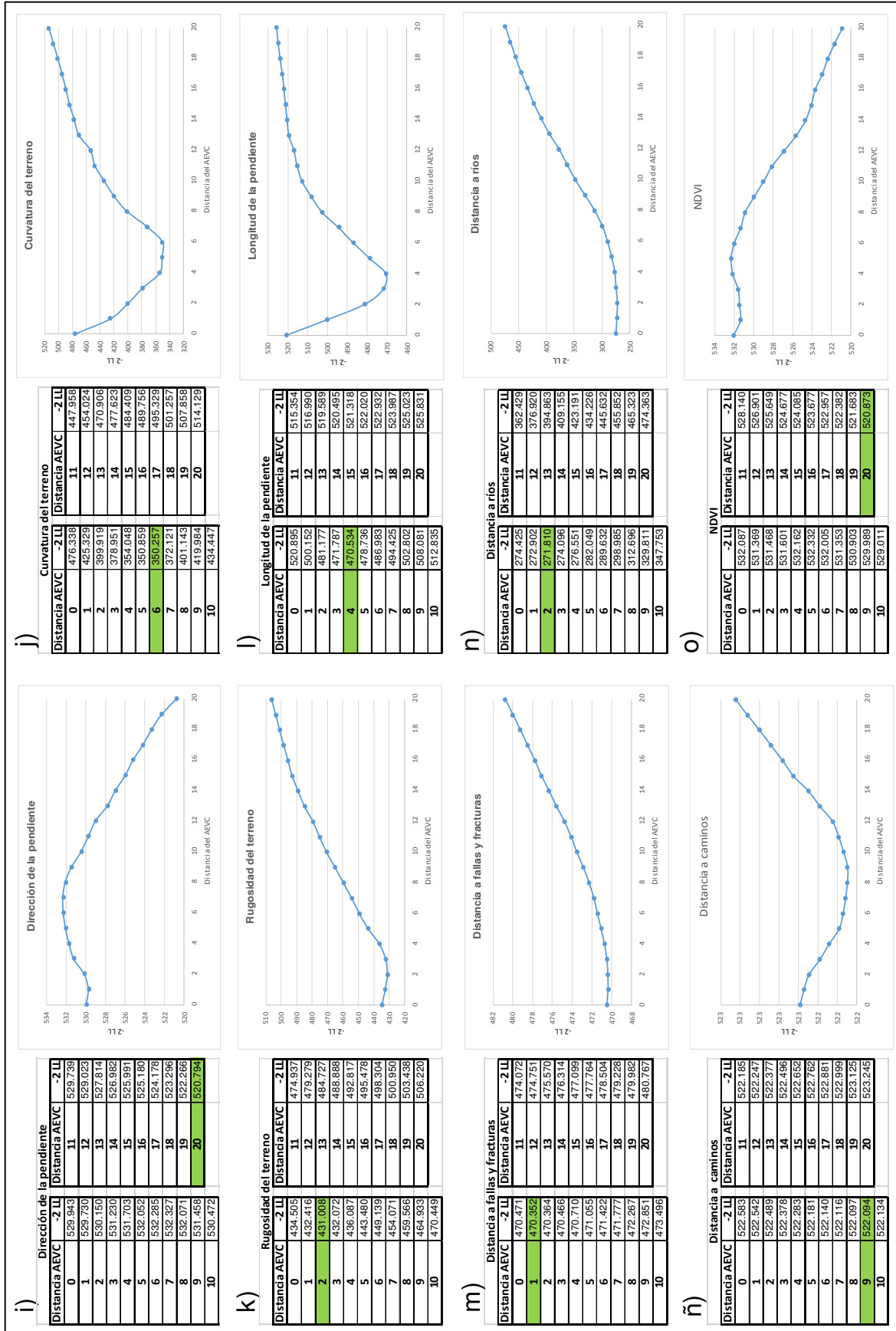


Figura 56 Resultados del AECV por variable explicativas (continuación de la página anterior. Elaboración propia.

Tabla 15. Resultados para el modelo de RL con base al AEVC.

a) Variables en la ecuación

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. para EXP(B)	
							Bajo	Superior
Us154	.049	.025	3.695	1	.055	1.050	.999	1.104
Camb3	.017	.007	5.384	1	.020	1.017	1.003	1.031
Geom6	.002	.001	11.350	1	.001	1.002	1.001	1.004
Geom11	.005	.001	11.268	1	.001	1.005	1.002	1.008
Geom12	.001	.001	5.378	1	.020	1.001	1.000	1.002
text1	-.002	.001	6.883	1	.009	.998	.997	1.000
curv_d6	-4.846	1.308	13.734	1	.000	.008	.001	.102
drío_d2	-.055	.010	31.915	1	.000	.947	.929	.965
Constant	.542	.433	1.571	1	.210	1.720		

b) Resumen del modelo

-2 Log likelihood	Cox & Snell R ²	Nagelkerke R ²
191.469	.588	.785

c) Tabla de clasificación

Observado	Pronosticado		
	Deslizamiento		Percentage Correct
	0	1	
Deslizamiento 0	172	20	89.6
Deslizamiento 1	20	172	89.6
Porcentaje general			89.6

Punto de corte 0.585

d) Tabla de contingencia prueba de Hosmer y Lemeshow

	Estables= 0		Deslizamiento= 1		Total
	Observado	Esperado	Observado	Esperado	
1	38	37.816	0	.184	38
2	37	36.073	1	1.927	38
3	34	32.987	4	5.013	38
4	29	29.216	9	8.784	38
5	23	23.429	15	14.571	38
6	12	15.319	26	22.681	38
7	10	9.653	29	29.347	39
8	4	5.081	34	32.919	38
9	4	2.043	34	35.957	38
10	1	.384	40	40.616	41

e) Variables en la ecuación

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. para EXP(B)	
							Bajo	Superior
Camb3	.016	.007	5.493	1	.019	1.016	1.003	1.029
Geom6	.002	.001	12.462	1	.000	1.002	1.001	1.004
Geom11	.005	.001	10.792	1	.001	1.005	1.002	1.008
Geom12	.001	.001	6.350	1	.012	1.001	1.000	1.003
text1	-.002	.001	6.882	1	.009	.998	.997	1.000
curv_d6	-4.962	1.317	14.189	1	.000	.007	.001	.093
drío_d2	-.051	.009	32.050	1	.000	.951	.934	.967
Constant	.512	.430	1.419	1	.234	1.668		

Elaboración propia.

Con información *in situ*.

Para la elaboración de este modelo se utilizó la información las 16 variables explicativas tomada *in situ* e integrada en una sola matriz, sobre la cual se procedió a realizar el análisis de RL con la opción hacia adelante razón de la verosimilitud.

Los resultados en la elaboración de este modelo tuvieron un comportamiento similar al de los obtenidos en el modelo con base en el AEVC. En un primer procesamiento con la RL se integraron al modelo algunas variables que sobrepasaban el límite de significancia de 0.05. En la Tabla 16a se presentan los valores para las variables en la ecuación de un primer procesamiento, como se observa las clases del tema de uso del suelo y vegetación (usv_2015) reportan valores de significancia mayores a 0.05 (marcados en color naranja), con excepción de la clase de referencia (1^{er} renglón) la cual no tiene calculado un valor para su coeficiente. Por lo anterior se realizó un segundo procesamiento con las variables que reportaron valores de significancia aceptables, siendo estas la dirección de la pendiente (dirp_d0), distancia a fallas y fracturas (dflr_d0) y la distancia a ríos (drio_d0). Los resultados de este segundo modelo se muestran en la Tabla 16b, en los datos de resumen del modelo se aprecia el valor de 239.15 alcanzado por el indicador -2 LL, como se observa este valor es mayor al obtenido en el modelo con base al AEVC que fue de 191.47 (Tabla 15b), como se ha mencionado esta disminución en el valor de -2 LL indica un mejor ajuste de un modelo a otro. La tabla de clasificación (Tabla 16c) reporta el porcentaje de correcta clasificación alcanzado por el modelo que es de 85.9 % para las tres categorías, estos valores son menores a los alcanzados por el modelo de AEVC (Tabla 15c). La información de la tabla de contingencia de la prueba de Hosmer y Lemeshow (Tabla 16d) reporta una mayor diferencia entre los valores observados y esperados en comparación con los del modelo del AEVC (Tabla 15d). Por último, las variables en la ecuación (Tabla 16e) muestran que sus niveles de significancia están por debajo del valor máximo permitido de 0.05. Los resultados de este segundo modelo serán utilizados en la elaboración del modelo espacial *in situ*.

Tabla 16. Resultados para el modelo de RL con base a datos *in situ*.

a) Variables en la ecuación

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. para EXP(B)	
							Bajo	Superior
USV_2015			19.794	8	.011			
USV_2015(1)	21.195	19296.655	.000	1	.999	1603234780.123	.000	.
USV_2015(2)	21.186	19296.655	.000	1	.999	1588641811.708	.000	.
USV_2015(3)	1.708	30122.702	.000	1	1.000	5.518	.000	.
USV_2015(4)	22.553	19296.655	.000	1	.999	6232402316.856	.000	.
USV_2015(5)	23.741	19296.655	.000	1	.999	20442609221.607	.000	.
USV_2015(6)	24.007	19296.655	.000	1	.999	26675866533.288	.000	.
USV_2015(7)	24.614	19296.655	.000	1	.999	48923785787.876	.000	.
USV_2015(8)	7.182	21847.481	.000	1	1.000	1316.002	.000	.
dirp_d0	.012	.003	16.204	1	.000	1.012	1.006	1.018
dfir_d0	-.002	.001	6.394	1	.011	.998	.996	1.000
drío_d0	-.065	.009	49.139	1	.000	.937	.920	.954
Constant	-20.439	19296.655	.000	1	.999	.000		

b) Resumen del modelo

-2 Log likelihood	Cox & Snell R ²	Nagelkerke R ²
239.152	.534	.712

c) Tabla de clasificación

Observado	Pronosticado		
	Deslizamiento		Percentage Correct
	0	1	
Deslizamiento 0	165	27	85.9
Deslizamiento 1	27	165	85.9
Porcentaje general			85.9

The cut value is .652

d) Tabla de contingencia prueba de Hosmer y Lemeshow

	Estables= 0		Deslizamiento= 1		Total
	Observado	Esperado	Observado	Esperado	
1	33	33.275	5	4.725	38
2	30	27.713	8	10.287	38
3	17	24.833	21	13.167	38
4	19	22.243	19	15.757	38
5	26	18.988	12	19.012	38
6	24	16.530	14	21.470	38
7	16	14.109	22	23.891	38
8	12	12.424	26	25.576	38
9	8	11.137	30	26.863	38
10	7	10.747	35	31.253	42

e) Variables en la ecuación

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. para EXP(B)	
							Bajo	Superior
dirp_d0	.010	.003	15.213	1	.000	1.011	1.005	1.016
dfir_d0	-.003	.001	17.713	1	.000	.997	.996	.998
drío_d0	-.057	.008	57.746	1	.000	.944	.931	.958
Constant	1.871	.376	24.784	1	.000	6.492		

Elaboración propia.

4.2.3. Modelos espaciales de probabilidad a deslizamientos con base al AEVC e *in situ*.

Con la información de las variables y sus coeficientes obtenida en los modelos estadísticos finales (Tablas 15e y 16e) se elaboraron los modelos espaciales de probabilidad a deslizamientos.

Para el modelo con base en el AEVC fueron 7 las variables utilizadas: deforestación del tema de cambio de uso de suelo y vegetación con un AEVC de 15 pixeles; las clases de ladera superior, lomeríos de avalancha de escombros cubiertos de piroclastos y, lomeríos de pómez y depósitos de caída, estas tres del tema de geomorfología con un AEVC de 20 pixeles; la clase de textura del suelo gruesa con un AEVC de 20 pixeles; la curvatura del terreno con un AEVC de 6 pixeles; y la distancia a ríos con un AEVC de 2 pixeles. Para cada una de las variables se generó la información espacial correspondiente al AEVC seleccionado y de acuerdo al tipo de variable, cualitativa o cuantitativa se utilizó la función de sumatoria o promedio, respectivamente, para calcular las áreas de vecindad.

Una vez obtenidos los mapas temáticos de las variables explicativas, a la distancia del análisis de vecindad seleccionado, junto con sus correspondientes coeficientes (Tabla 15e), se utilizó la fórmula de RL (Fórmula 4) para obtener el mapa de probabilidad a deslizamientos con valores entre 0 y 1 (Figura 57).

$$Y = \frac{1}{1 + e^{-(0.512 + (0.16 * Camb3) + (0.002 * Geom6) + (0.005 * Geom11) + (0.001 * Geom12) + (-0.002 * text1) + (-4.962 * curv_d6) + (-0.051 * drio_d2))}}$$

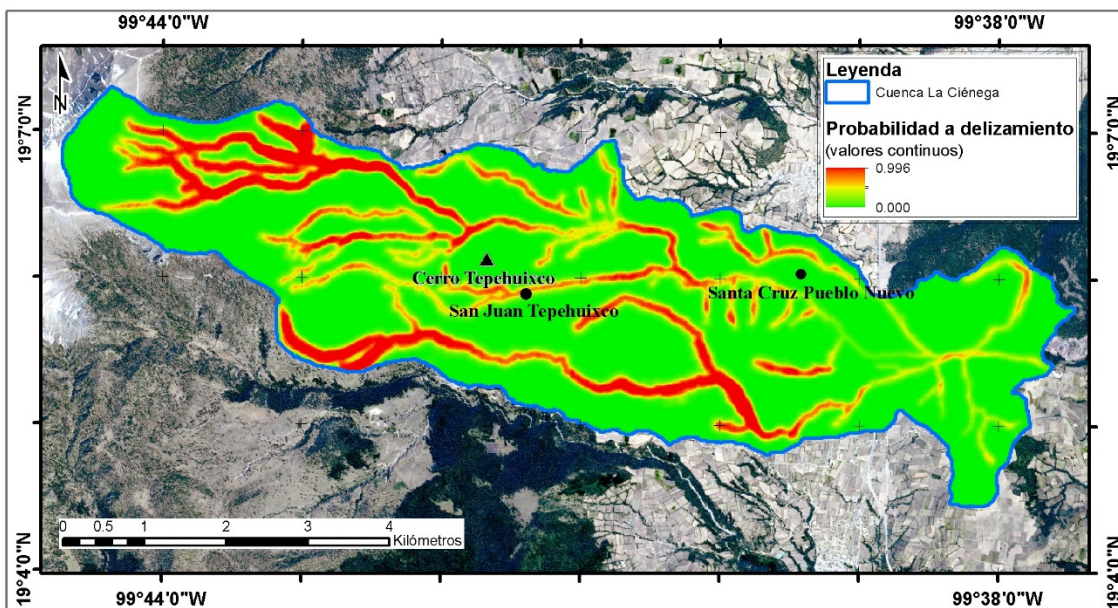


Figura 57. Mapa de probabilidad a deslizamientos con valores continuos entre 0 y 1 con base al AEVC. Elaboración propia.

Posteriormente este mapa fue reclasificado en las clases de 0 y 1 a partir del punto de corte óptimo con un valor de 0.585 (Tabla 15e), este valor fue seleccionado ya que con él se obtiene una igualdad en cuanto al nivel de correcta clasificación entre las dos clases de la variable dependiente (Figura 58).

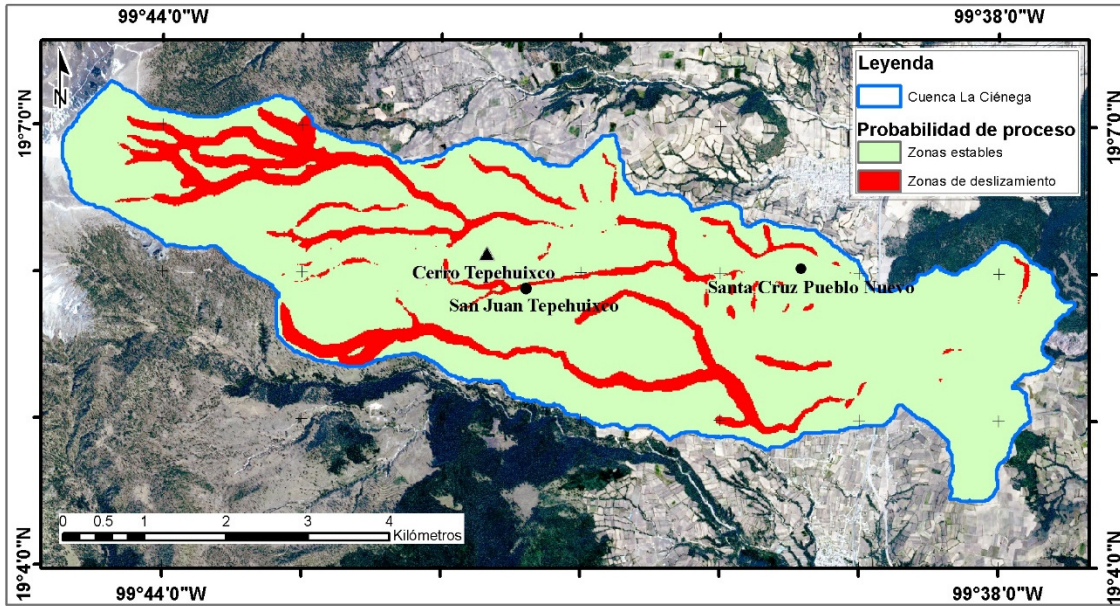


Figura 58. Mapa clasificado en zonas de probabilidad a deslizamientos y zonas estables con base en el punto de corte. Modelo con base al AEVC. Elaboración propia.

Para el mapa ya clasificado en zonas de probabilidad estables y de deslizamiento se calculó el porcentaje de pixeles que fueron clasificados como áreas con probabilidad a deslizamiento en función del área total de la cuenca. De un total de 133,654 pixeles que cubren la cuenca, 18,768 corresponden a zonas con probabilidad a deslizamientos, esto equivale al 14.04 % de la zona de estudio.

Para el modelo con base en información *in situ* se utilizaron 3 variables: dirección de la pendiente; distancia a fallas y fracturas; y distancia a ríos (Tabla 16e). Tomando la información correspondiente a los mapas temáticos de cada variable se les aplicó la fórmula de la RL (Fórmula 4). Como primer resultado se tuvo el mapa con los valores de probabilidad del proceso de deslizamientos entre 0 y 1 (Figura 59).

$$Y = \frac{1}{1 + e^{-(1.871 + (0.010 * dirp_0) + (0.003 * dflr_d0) + (0.057 * drio_d0))}}$$

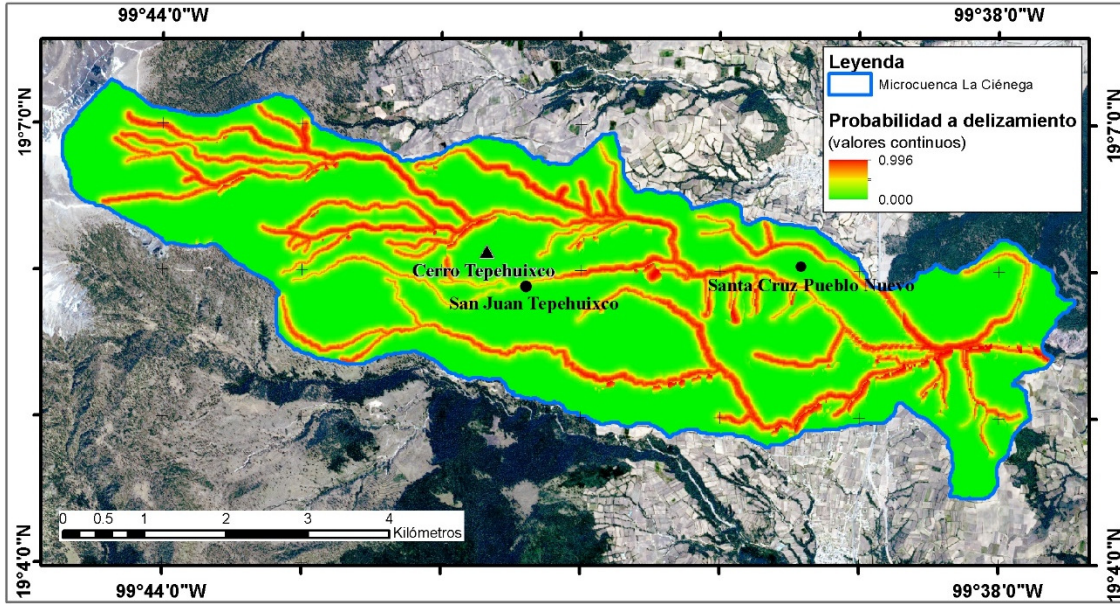


Figura 59. Mapa de probabilidad a deslizamientos con información *in situ*. En valores continuos entre 0 y 1. Elaboración propia.

El mapa de valores continuos fue clasificado en las clases 0 y 1 correspondientes la probabilidad de zonas estables y zonas de deslizamientos, respectivamente. Esta clasificación se hizo utilizando un punto de corte de 0.652, pues con este valor con que se alcanza un equilibrio entre los pixeles correctamente clasificados para ambas clases (Figura 60).

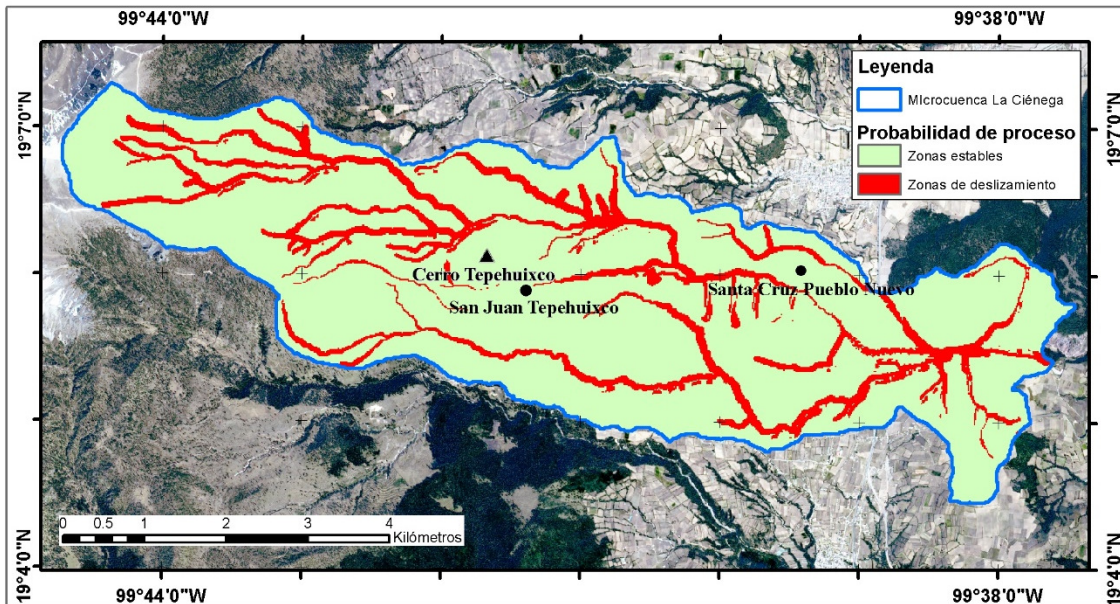


Figura 60. Mapa clasificado en zonas de probabilidad a deslizamientos y zonas estables con base en el punto de corte. Modelo *in situ*. Elaboración propia.

Para el modelo con base en los datos *in situ* se obtuvieron un total de 20,639 pixeles con probabilidad a deslizamiento lo que equivale al 15.44 % de la superficie de la cuenca.

4.2.4. Evaluación de los modelos espaciales.

Con base en la curva COR.

En la Figura 61 se muestran los resultados para el modelo a partir de la información del AEVC. El área bajo la curva representa que 96.8 % de los sitios que componen la muestra extendida fueron clasificados correctamente por el modelo, el error estándar reporta el valor de 0.014, una significancia de 0.00 y los límites del intervalo de confianza entre 0.941 y 0.955 (Figura 61a). La Figura 61b corresponde al gráfico de la curva COR, en este se aprecia como la línea en color azul inicia con un valor en la Sensibilidad mayor a 0.6 y una 1-Especificidad igual a 0 lo que da como resultado el alto porcentaje de área bajo la curva.

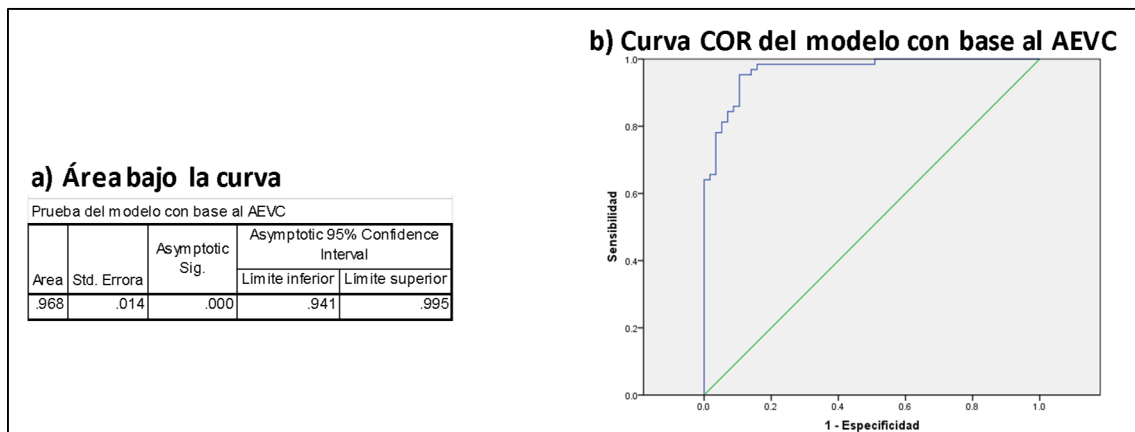


Figura 61. Resultados de la curva COR para el modelo con base en el AEVC. Elaboración propia.

En la Figura 62 se muestra los resultados de la curva COR para el modelo *in situ*. Para este modelo el área bajo la curva reporto un valor del 89.8 %, en cuanto al error estándar este fue de 0.033, para la significancia se mantuvo en 0.00 y el intervalo de confianza fue de 0.833 a 0.963 (Figura 62a), a excepción de la significancia los demás valores reportados en este modelo fueron inferiores al reportados por el modelo con base al AEVC. En la Figura 62b se aprecia como la relación entre 1-Especificidad y la Sensibilidad comienza a alejarse rápidamente del eje de las Y, lo que hace que la curva cubra del 89.9 %, un 7 % menor a la cubierta por el modelo resultado del AEVC.

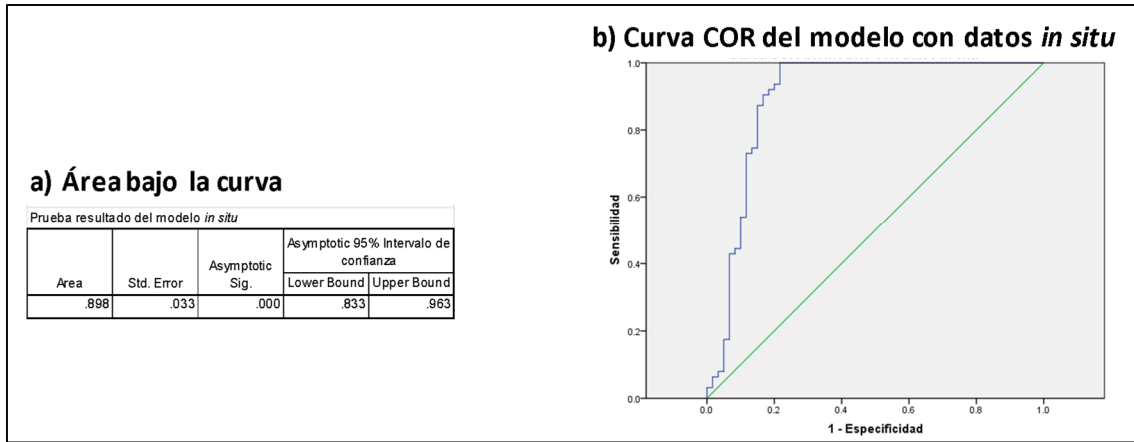


Figura 62. Resultados de la curva COR para el modelo con base en datos *in situ*. Elaboración propia.

Con base en la matriz de clasificación.

En la Tabla 17 se presentan los resultados obtenidos para cada modelo.

Tabla 17. Evaluación utilizando la matriz de clasificación. a) modelo de AEVC, b) modelo *in situ*.

a)					b)						
Matriz de clasificación del modelo con base al AEVC					Matriz de clasificación del modelo con datos <i>in situ</i>						
		Grupo pronosticado					Grupo pronosticado				
		Zonas establ	Cabeceras				Zonas estables	Cabeceras			
				0	1			0	1		
Grupo real	Zonas establ	0	58	6			Zonas estables	0	55	9	
	Cabeceras	1	9	55			Cabeceras	1	11	53	
		Índice de Kappa			0.77			Índice de Kappa			0.69
		Exactitud general			88.28%			Exactitud general			84.38%
		Exactitud del productor			85.94%			Exactitud del productor			82.81%
		Exactitud del usuario			90.16%			Exactitud del usuario			85.48%
		Eficiencia del modelo			0.77			Eficiencia del modelo			0.69

Elaboración propia.

Como se observa en las Tablas 17a y 17b los valores obtenidos por el modelo de AEVC fueron mejores en todos los aspectos evaluados en comparación con los obtenidos por el modelo con información *in situ*, siendo la eficiencia del modelo la que reporta una mayor diferencia entre los modelos, con 0.08 mayor en el modelo con base en el AEVC.

CAPÍTULO 5. DISCUSIÓN Y CONCLUSIONES.

5.1. Discusión.

Es importante resaltar que, aun cuando el procesamiento del modelo de RL en su variante hacia adelante teóricamente debe de eliminar las variables que estén fuera del rango de significancia mayor a 0.05, en el modelo resultado se mantienen algunas variables que no cumplen con esta condición (Tablas 15a y 16a), por ello se elaboró un segundo modelo de RL eliminando las variables que no cumplieron con esta condición en el primero (Tablas 15e y 16e). Con lo anterior se pudo rechazar para todas las variables en el segundo modelo la H_0 de la significancia estadística y por la tanto se acepta la H_1 de que al introducir una variable al modelo esta contribuye en la explicación de la variable dependiente.

Para el modelo elaborado con datos del AEVC, de las 7 variables que lo componen, 4 tienen un signo positivo en sus coeficientes lo que significa que con forme aumenta el valor de estas variables se incentiva la ocurrencia de deslizamientos. Las variables con coeficientes positivos fueron: la deforestación del tema de cambio de uso del suelo (Camb3) y para el tema de geomorfología las clases de ladera superior (Geom6), lomeríos de avalancha de escombros cubiertos de piroclastos (Geom11), y lomeríos de pómez y depósitos de caída (Geom12) (Tabla 15e). En lo que respecta al valor del coeficiente este señala que la variable con la cifra más elevada es la que tiene un mayor peso en el cálculo de la probabilidad de ocurrencia de los deslizamientos, con base en los datos de la Tabla 15e se observa que la deforestación (Camb3) es la que más influye con un valor en su coeficiente de 0.016, le siguen la presencia de lomeríos de avalancha de escombros cubiertos por piroclastos (Goem11) con un valor de 0.005, después la ladera superior (Geom6) con un valor de 0.002, y por último los lomeríos de pómez y depósitos de caída (Geom12) con un valor en su coeficiente de 0.001. Con respecto a las variables con coeficiente negativo estas representan características en las que conforme aumenta el valor en la variable explicativa la probabilidad a ocurrencia de deslizamientos disminuye, para el modelo con base al AEVC los resultados obtenidos (Tabla 15e) muestran que las variables con esta característica son la textura del suelo de tipo gruesa (text1), la curvatura del suelo (curv_d6) y la distancia a ríos (drio_d2). De estas la que más podría limitar probabilidad a deslizamientos, con base en el valor que tome la variable es la curvatura del terreno (curv_d6) con un valor en su coeficiente de -4.462, les sigue la distancia a ríos (drio_d2) con valor de -0.051 y por último la textura del suelo (text1) con un valor de -0.002.

Los variables y sus respectivos coeficientes que quedaron en el modelo final generado con información *in situ* (Tabla 16e) muestran que de las 3 variables la correspondiente a dirección de

la pendiente tiene una influencia positiva con respecto a la probabilidad de ocurrencia de los deslizamientos (valor del coeficiente de 0.010), mientras que las variables de distancia a ríos y distancia a fallas y fracturas, al tener un valor de coeficientes negativo (-0.057 y -0.003, respectivamente), son variables que conforme aumenta su valor limitan la probabilidad de ocurrencia, siendo la distancia a ríos la que tiene el mayor efecto de disminución de la probabilidad, ya que a esta le corresponde el mayor valor absoluto en su coeficiente.

Con respecto al desempeño de los modelos, la matriz de clasificación producto de la calibración del modelo estadístico para el AEVC reporta una correcta clasificación de 89.6 % para ambos grupos (zonas estables y de deslizamiento) (Tabla 15c), mientras que el modelo con información *in situ* reportó un valor de 85.9 (Tabla 16c), es decir un 3.7 % menos. El modelo espacial basado en el AEVC muestra una menor superficie con probabilidad a deslizamientos, siendo del 14.04 % del área de la cuenca, mientras que en el modelo *in situ* el porcentaje de estas áreas es del 15.44 %, que equivalen a 422.28 ha y 464.38 ha respectivamente, es decir, hay un aumento del 9.9 % del área susceptible a deslizamientos en el modelo *in situ* con respecto al modelo del AEVC. Los resultados de la curva COR reportan un área bajo la curva de 96.8 % para el modelo de AEVC y un área de 89.8 % para el modelo *in situ*, que equivale a un 7 % de mejor clasificación en el modelo con base al AEVC.

Un aspecto de mucha importancia y directamente relacionado con el análisis geográfico es la precisión del modelo producto del AEVC el cual muestra un mayor ajuste en la ubicación de las áreas con probabilidad en comparación con el modelo *in situ*. Como se observa en la Figura 63, el modelo basado en el AEVC, además de reportar una mayor exactitud en la clasificación de las zonas estables y con probabilidad a deslizamientos (Tabla 17a), también reporta una mayor precisión en la distribución de tales zonas con respecto a la información del inventario de deslizamientos.

En la Figura 63 se aprecia como en la zona noroeste de la cuenca se presenta una concentración importante de deslizamientos, al mismo tiempo la zona reporta una alta probabilidad a deslizamientos. Otra zona importante con probabilidad a deslizamientos se presenta en forma longitudinal de este a oeste en la parte centro sur de la cuenca, de igual modo la presencia de deslizamientos registrados en la zona durante la elaboración del inventario es alta (círculos rojos). Con respecto a las áreas estables, la zona este de la cuenca presenta poca superficie con probabilidad a deslizamiento (círculo verde) y en esta área se registró muy baja presencia de deslizamientos. De manera general para este modelo con base en el AEVC las áreas con probabilidad a deslizamientos tienen una mayor presencia en la parte oeste y en la zona central

de la cuenca, mientras que al este de la zona de estudio la presencia de área con probabilidad es casi nula.

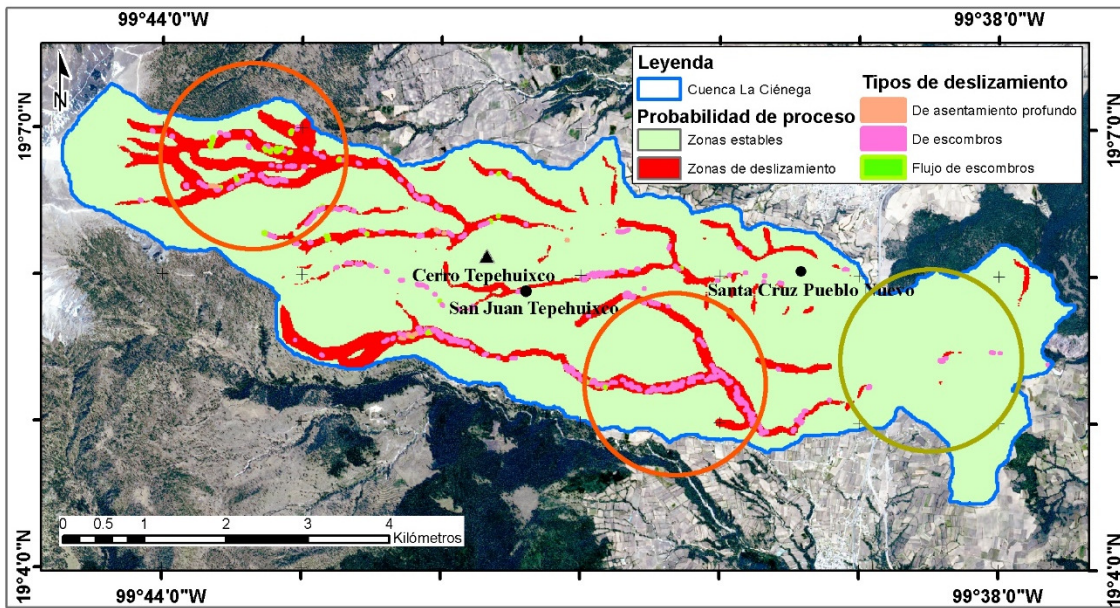


Figura 63. Mapa de distribución del inventario de deslizamientos con relación al modelo de probabilidad con base en el AEVC. Elaboración propia.

El modelo espacial con base a información *in situ* (Figura 64), muestra una sobreestimación en algunas áreas. Por ejemplo, se identifican dos zonas que presentan alta probabilidad a deslizamientos según el modelo, una se presenta al noreste del cerro Tepehuixco, la segunda se ubica en la zona este de la cuenca (círculos azules). Sin embargo, en estas zonas el inventario reporta una presencia escasa de deslizamientos.

La importancia de contar con una mayor precisión en el modelo espacial de probabilidad deslizamientos radica en que con ello se tiene una mayor certeza en la ubicación y distribución de las áreas con probabilidad, y por lo tanto brinda la posibilidad de realizar un mejor análisis espacial de las posibles causas, consecuencias y medidas de mitigación a implementar con miras a disminuir los riesgos potenciales a determinadas actividades antrópicas dentro de la zona de estudio.

La relación de los modelos de probabilidad a deslizamientos y las variables espaciales que fueron utilizadas para su elaboración se describe a continuación. Esta se realiza individualmente con cada uno de los mapas temáticos correspondientes a las variables explicativas, sin perder de vista que el análisis de RL implica la evaluación de los efectos combinados de todas las variables involucradas en cada uno de los modelos finales.

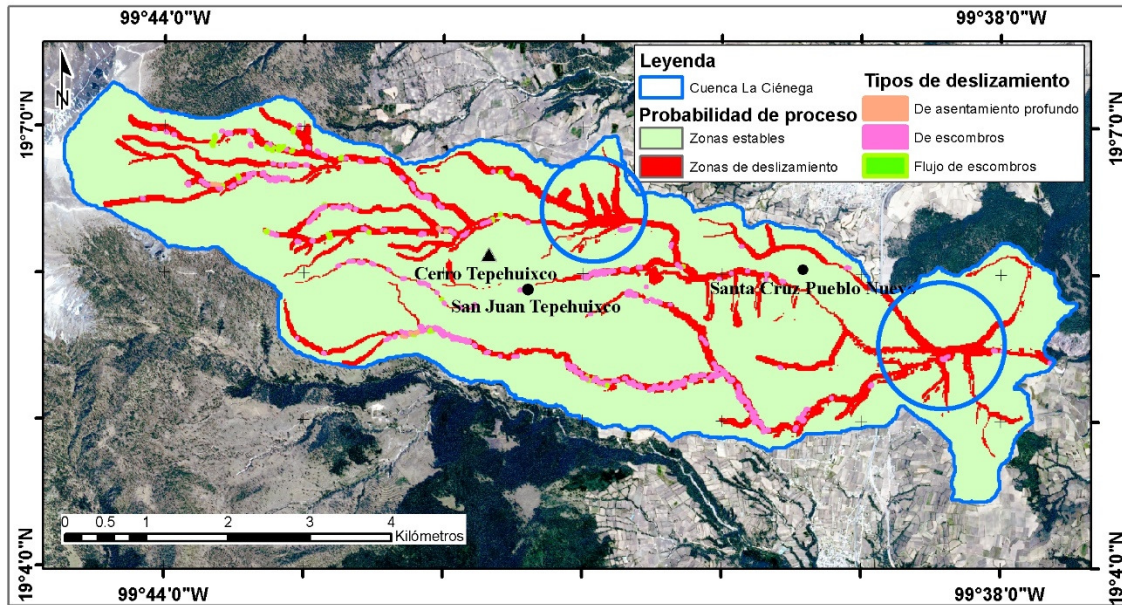


Figura 64. Distribución del inventario de deslizamientos con relación al modelo de probabilidad con base en datos *in situ*. Elaboración propia.

En la Figura 65 se muestra el mapa de la distribución de las zonas con probabilidad a deslizamientos producto del AEVC y su relación con el cambio de uso de suelo correspondiente a deforestación, esta variable resulto ser la que tiene un mayor impacto para la presencia de zonas con probabilidad a deslizamientos ya que reporta el valor más alto en los coeficientes positivos (0.016) (Tabla 15e). Lo anterior resalta la importancia de mantener la cobertura forestal en la zona de estudio (Figura 23), poniendo especial atención en las zonas que resultaron con probabilidad de deslizamientos al oeste de la cuenca y al sur del poblado de Santa Cruz Pueblo Nuevo (círculos verdes. Figura 65). Sin dejar de lado las áreas forestales en las que la presencia de la deforestación es escasa, ya que si se elimina la cobertura forestal esto resultaría en la presencia de nuevas áreas con probabilidad a deslizamientos. La deforestación implica la pérdida de las raíces de la vegetación forestal y como consecuencia la pérdida del refuerzo que estas aportan al sustrato, lo que aumenta la posibilidad de generación de fallas en el terreno y como consecuencia una mayor probabilidad a la ocurrencia de deslizamientos.

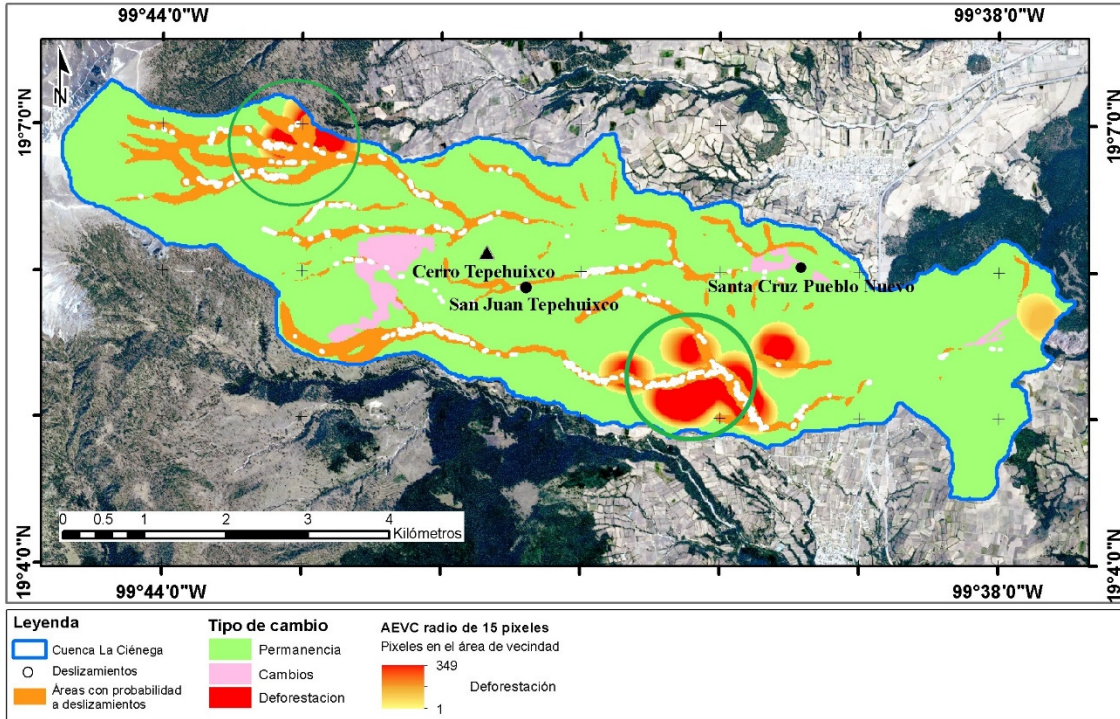


Figura 65. Mapa de la distribución de las zonas deforestadas y su cálculo con un AEVC de 15 píxeles, y las áreas con probabilidad a deslizamiento. Elaboración propia.

Para las variables de geomorfología que fueron incluidas en el modelo final de AEVC (Tabla 15e), se considera que debido a que estas geoformas se componen por depósitos de diversos materiales piroclásticos poco consolidados y fáciles de erosionar, entre más inmerso este un sitio en alguna de estas clases, mayor será su probabilidad a deslizamientos (Figura 66). La clase de lomeríos de avalancha de escombros cubiertos de piroclastos es la que reporta el mayor valor en su coeficiente. Esta geoforma, tiene un desnivel altitudinal de 576 m con pendientes entre los 6° y los 30° y tiene una densidad de red fluvial que va de 1 km /km² a 2 km / km² con una profundidad en sus valles entre los 35 m y 50 m. Es de destacar la mayor profundidad de sus valles lo que junto con la fuerte inclinación vertical de las paredes, la poca consolidación y facilidad de erosión en sus materiales hacen que las paredes de estas depresiones sean muy propensas al desarrollo de deslizamientos. Le sigue la categoría de ladera superior la cual está constituida por roca ígnea dacítica con depósitos de bloque y ceniza, depósitos de caída y depósitos de flujos piroclásticos, presenta una diferencia altitudinal de 699 m con pendientes del terreno entre los 15° y 30°, una alta densidad en su red fluvial mayor a los 3 km / km² y una profundidad en sus valles de 15 m a 25 m En esta geoforma la predisposición a deslizamientos superficiales es potencialmente menor porqué entre los materiales que la constituyen se encuentra roca ígnea y depósitos de bloque los cuales le aportan una mayor consolidación. Por

último, los lomeríos de pómez y depósitos de caída constituidos por grandes depósitos de caída y flujos piroclásticos, con una diferencia altitudinal de 453 m y pendientes del terreno entre los 6° y 20°, con densidad fluvial menor a 4km /km2 y la profundidad de sus valles es menor a los 50 m, para esta geoforma sobresale la menor diferencia altitudinal y los valores más bajos en la pendiente del terreno como posibles causas para que sean las áreas de menor predisposición al desarrollo de deslizamientos. Como se mencionó la RL evalúa la interacción de las variables en conjunto por lo que se debe de tener especial cuidado en no desarrollar actividades que promuevan la deforestación dentro de las áreas que están cubiertas por alguna de estas 3 categorías de geomorfología ya que el efecto de ambas variables (deforestación y geomorfología) se combinaría y aumentaría la posibilidad a deslizamientos.

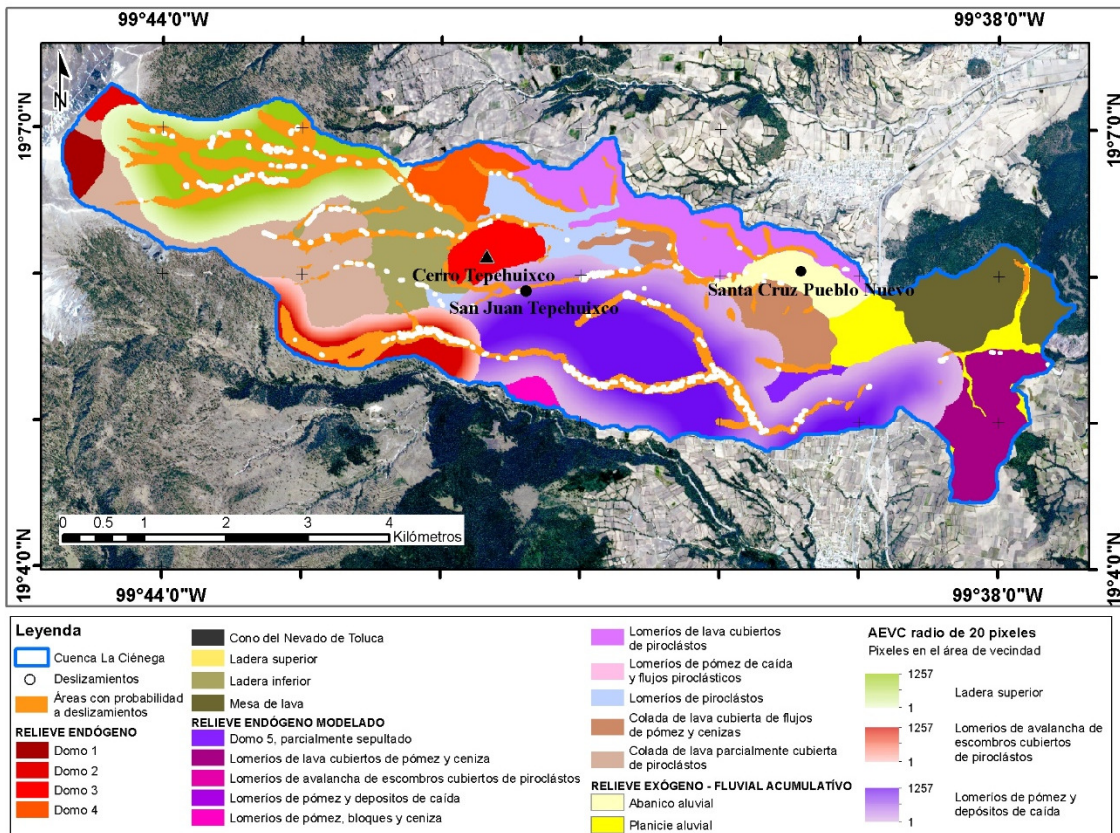


Figura 66. Mapa de las áreas con probabilidad a deslizamientos y el AEVC para las clases de geomorfología utilizadas en el modelo de probabilidad con base al AEVC. Elaboración propia.

En la Figura 67 se muestra la relación entre la curvatura del terreno y las áreas con probabilidad a deslizamientos. Como se observa las áreas con probabilidad presentan una alta coincidencia con valores bajos en la curvatura del terreno (-1.5), significa que se trata de superficies cóncavas del terreno, las cuales tienen una alta coincidencia con la ubicación de los cauces de los ríos, y

es donde tiene lugar una intensa socavación en las paredes de los lechos lo que acentúa la probabilidad a deslizamientos. Al ir aumentando los valores de la curvatura se comienzan a presentar formas convexas, es decir áreas fuera de los cauces y por lo tanto el efecto de socavación es prácticamente nulo limitando la presencia de deslizamientos.

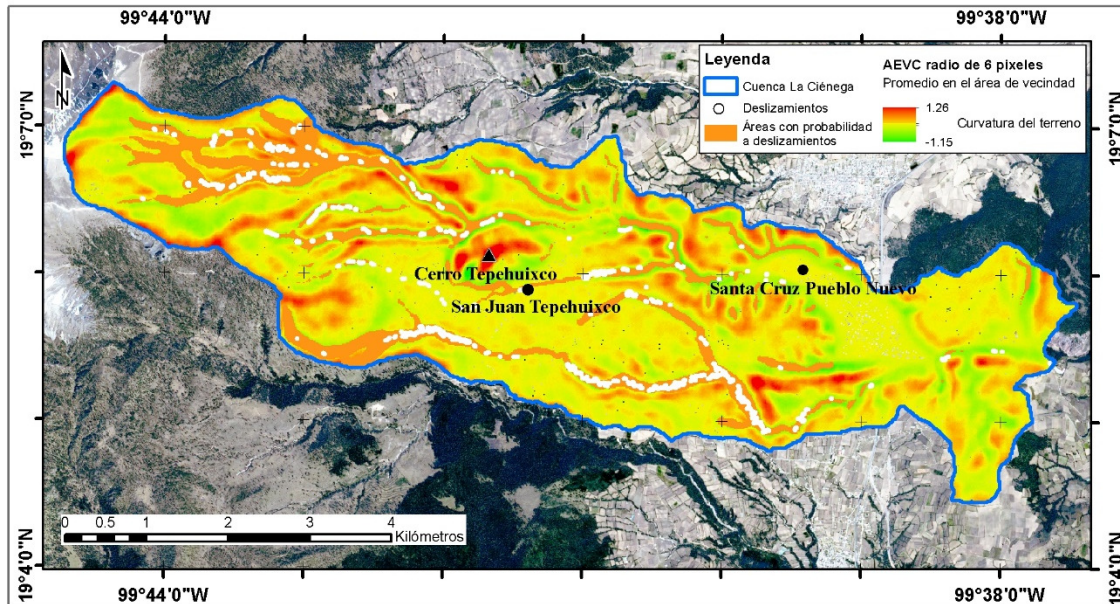


Figura 67. Mapa de las áreas con probabilidad a deslizamientos y su relación con el promedio de la curvatura del terreno a un AEVC de 6 píxeles. Elaboración propia.

Para la variable de distancia a ríos en la Figura 68 se observa que conforme se está más alejado de las corrientes de agua se va presentando una menor probabilidad a deslizamientos. Los deslizamientos presentes en la cuenca se desarrollan principalmente en las paredes de los cauces y por lo tanto estas áreas son las que tendrían una mayor probabilidad a presentar deslizamientos. Esto se debe a que la saturación de agua en el terreno es una de las principales causas de los deslizamientos y dicha saturación tiene una importante presencia a la orilla de los ríos. Es de resaltar que en la parte este de la cuenca (circulo amarillo), a pesar de que existe la presencia de ríos, en esta zona no aparecen áreas con probabilidad a deslizamientos esto es por el efecto de alguna otra variable, que en este caso podría ser la textura del suelo la cual se explica a continuación.

La textura del suelo gruesa es la variable que reporto el coeficiente de menor valor (-0.002) con signo negativo (Tabla 15e). El resultado muestra que entre más inmerso se encuentra un sitio en esta clase menor será su probabilidad a presentar deslizamientos. La causa de este efecto es la disminución en el deterioro de los suelos debido a la erosión la cual es uno de los factores que dan origen a los deslizamientos, ya que los suelos de textura gruesa (arenas) son menos

susceptibles a la erosión que los de texturas finas (limos y arcillas) pues permiten el libre paso del agua disminuyendo así el efecto de escorrentía. Como se muestra en la Figura 69 la presencia de este tipo de textura del suelo se ubica en gran parte en la zona este de la cuenca, siendo esta la posible causa que limita la presencia de zonas con potencial a producir deslizamientos.

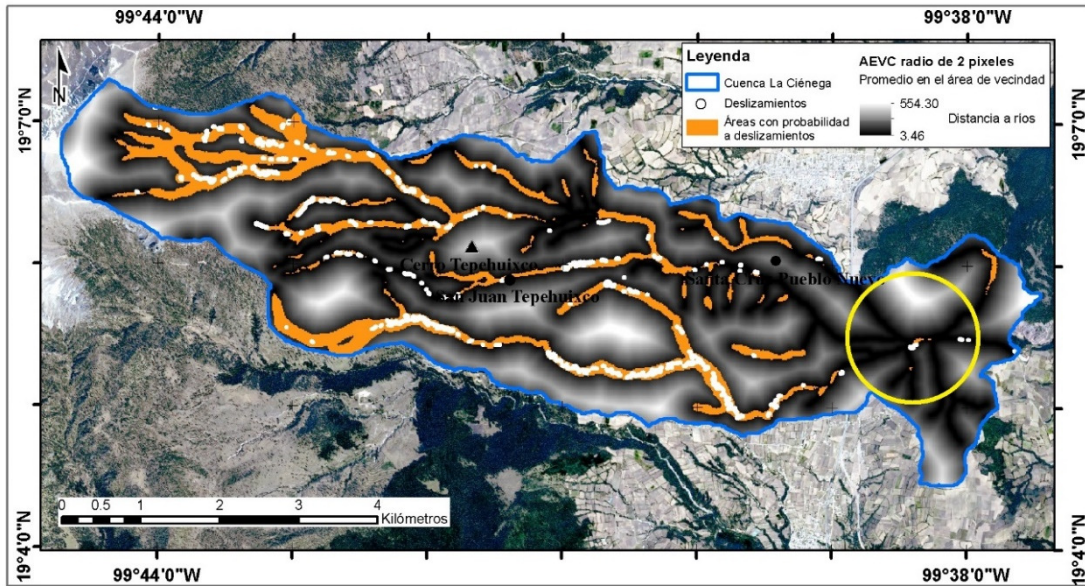


Figura 68. Mapa de las áreas con probabilidad a deslizamientos y su relación con el promedio de la distancia a ríos dentro del AEVC de 2 pixeles de radio. Elaboración propia.

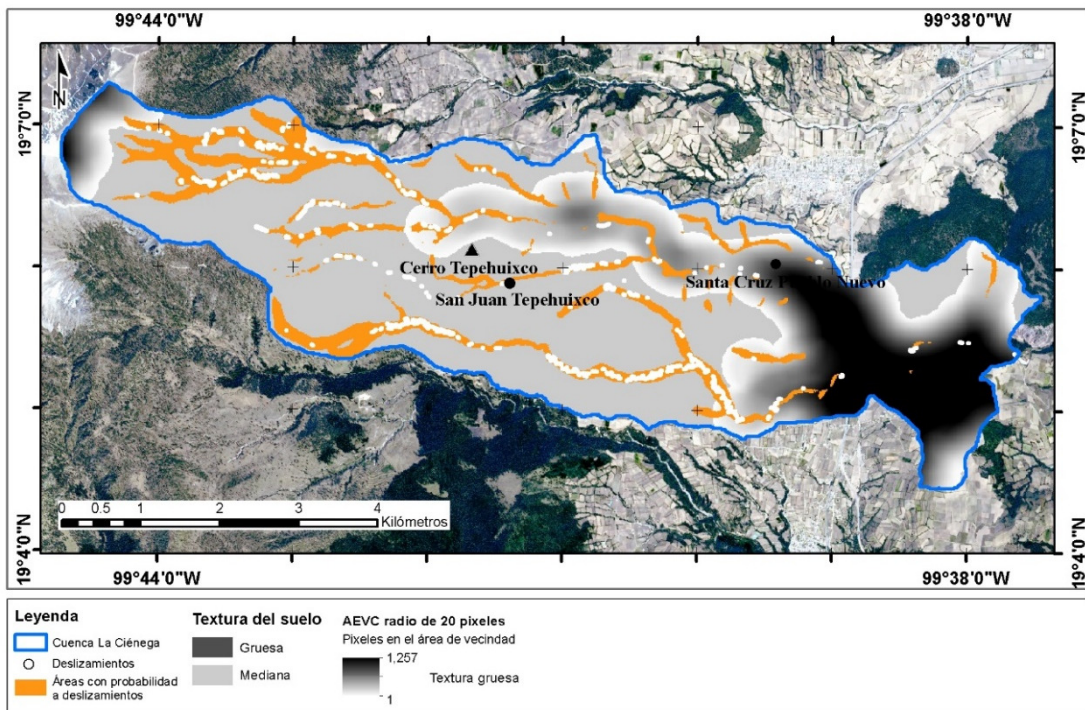


Figura 69. Mapa de las áreas con probabilidad a deslizamientos y su relación con la textura del suelo gruesa en un AEVC de 20 pixeles. Elaboración propia.

Para el modelo de probabilidad a deslizamientos con base en la toma de información *in situ* se utilizaron 3 variables para su elaboración (Tabla 16e). De estas, la variable de dirección de la pendiente es la única que reporto un coeficiente positivo, es decir que el aumento en su valor está directamente relacionado con el aumento en la probabilidad de deslizamiento de una determinada zona, con base en lo anterior se puede decir que cuando la pendiente presenta una orientación hacia el noroeste (valores más altos en la dirección de la pendiente) estas zonas tendrán una mayor probabilidad a deslizamientos que las pendientes que presenten una orientación noreste o sureste (Figura 70). Este aumento de la probabilidad a deslizamientos se debe a que las laderas orientadas al nornoroeste tienen más horas a la sombra y como resultado una mayor cantidad de humedad, lo cual se relacionan principalmente con el desarrollo de deslizamientos de tipo superficial.

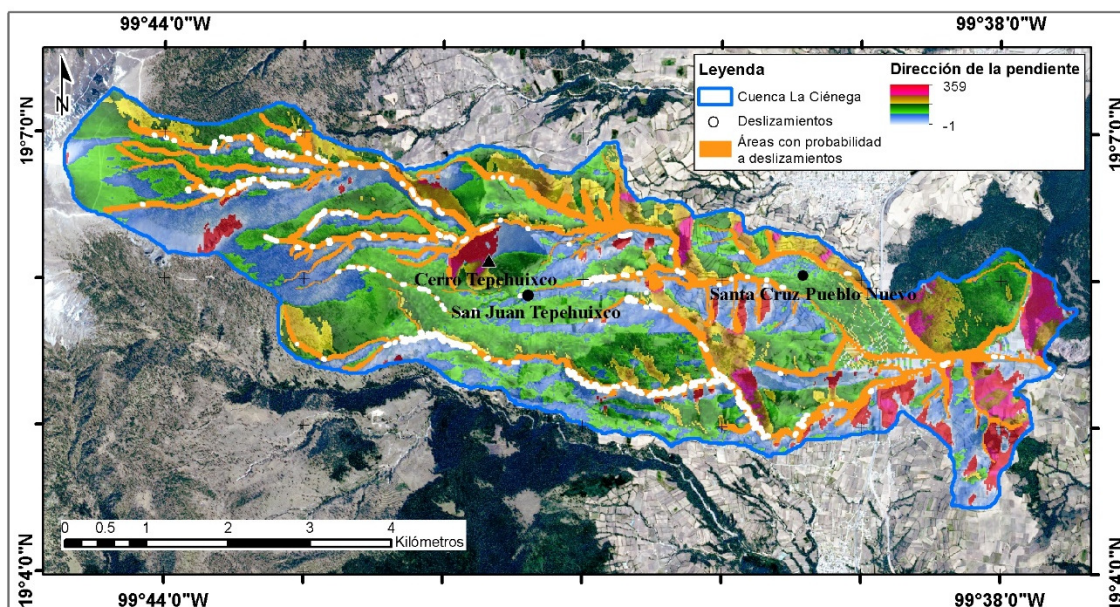


Figura 70. Mapa de las áreas con probabilidad a deslizamientos y su relación con la orientación de la pendiente tomada *in situ*. Elaboración propia.

Las otras dos variables trabajadas en el modelo *in situ* reportaron signo negativo en sus coeficientes por lo que el aumento en su valor tiene un efecto atenuante para la presencia de zonas con probabilidad a deslizamientos (Tabla 16e). De estas variables la distancia a ríos fue la que dio el mayor valor con coeficiente negativo, es decir, en cuando más alejada se encuentre un sitio de un río menor será su probabilidad a deslizamiento (Figura 71). La saturación de agua en las laderas, así como el corte lateral que provocan las corrientes de los ríos en las paredes de los cauces es una de las principales causas de los deslizamientos, por lo tanto, en cuanto más alejada este una zona del cauce de un río menor será su probabilidad a deslizarse.

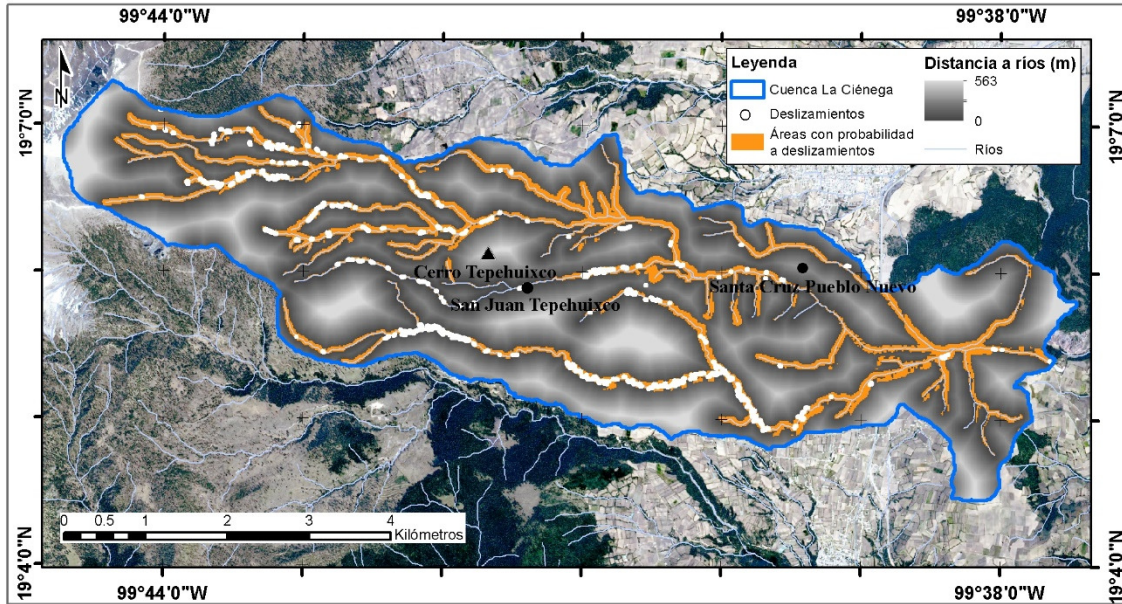


Figura 71. Mapa de las áreas con probabilidad a deslizamientos y su relación con la distancia a ríos tomada *in situ*. Elaboración propia.

La tercera variable en el modelo *in situ* es la distancia a fallas y fracturas, esta variable reporta un valor negativo en su coeficiente (Tabla 16e). El efecto de esta variable se podría deber a que por lo general las fallas y fracturas debilitan la resistencia en las rocas por pérdida de cohesión, además de generar inestabilidad en la zona lo que resultaría en una mayor probabilidad de deslizamientos en áreas cercanas a estas (Figura 72).

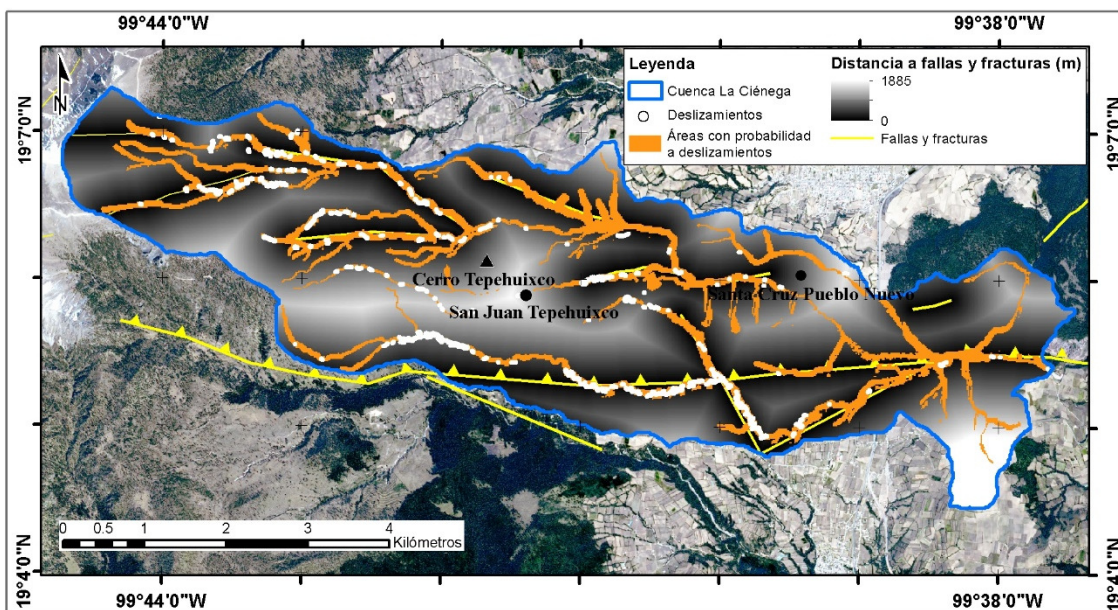


Figura 72. Mapa de las áreas con probabilidad a deslizamientos y su relación con la distancia a fallas y fracturas tomada *in situ*. Elaboración propia.

Con base en la descripción de la relación de las variables explicativas con los modelos de probabilidad a deslizamientos se interpreta que en el modelo elaborado con AEVC permitió integrar el entorno de los sitios, con respecto a las variables explicativas al cálculo espacial del mapa de probabilidad. Como resultado de esta integración se pudo evaluar un efecto de mayor alcance que las variables tienen sobre el cálculo de la probabilidad a deslizamientos (Figuras de la 65 a la 69), es decir que el efecto de las variables explicativas tiene un área de influencia, tanto para las variables que estimulan como las que limitan los deslizamientos. Ejemplo del área de influencia de una variable que incentiva los deslizamientos se puede observar en la zona noroeste de la cuenca (Figura 66) donde las áreas con probabilidad a deslizamientos ubicadas sobre la clase de ladera superior muestran una mayor amplitud en comparación con el modelo *in situ* (Figura 60), en lo que respecta a una variable limitante un ejemplo es la textura del suelo gruesa la cual su área de influencia provoca que en la parte este de la cuenca la presencia de área con probabilidad sea escasa (Figura 69).

Para el modelo espacial de probabilidad a deslizamientos con base en datos *in situ* (Figura 60) la toma de información local y por lo tanto el aislamiento de los sitios de su entorno dio como resultado un modelo espacial que se ajuste a los límites espaciales de las variables involucradas en el modelado, ejemplo de esto es la zona en la parte este del modelo *in situ* (Figura 71), en la que se presentan varias zonas con probabilidad a deslizamientos muy ajustadas a lo largo de la distribución de cauces sin importar que en la información del inventario no se identificaron deslizamientos en dicha área. Otro ejemplo se da en la zona noreste de la cuenca donde el modelo limita la presencia de zonas con probabilidad a la forma y presencia de los ríos. En el modelo *in situ* la probabilidad a deslizamientos se distribuye a lo largo de la gran mayoría de los ríos sin importar la presencia y densidad de deslizamientos identificados en el inventario de PG.

Debido a que el AEVC es un análisis espacial y estadístico de las condiciones que presentan las variables explicativas dentro de la zona de estudio con respecto a un área de vecindad a los sitios de muestreo, es de suponerse que los resultados de este análisis van a cambiar de una zona a otra, por lo tanto, las distancias reportadas para cada variable, así como los valores de los coeficientes que componen el modelo de probabilidad no serán los mismos. Incluso es de esperarse que las variables que reporten una relación con el proceso de deslizamientos no sean las mismas de una zona de estudio a otra. Por lo tanto, cuando se quiera aplicar esta metodología a la elaboración de modelos de probabilidad en diferentes zonas o con distribuciones y tamaños de muestra distintos es necesario realizar todo el proceso del AEVC y posteriormente la calibración del modelo con las distancias seleccionadas en el análisis de vecindad. Es decir, lo

que se puede compartir de una zona a otro es la metodología para aplicar el AEVC y no el modelo final de probabilidad (variables y sus coeficientes).

El motivo principal para el desarrollo del trabajo fue la implementación y evaluación de los resultados de un método de análisis espacial de vecindad que permita asociar las áreas cercanas a un sitio que vaya a ser utilizado para la generación de modelos de probabilidad de ocurrencia de PG utilizando la RL y con ello obtener modelos de probabilidad de mayor precisión, teniendo como base el principio geográfico referente a la causalidad o conexidad. Con ello se tuvo la posibilidad de identificar con un mayor éxito las relaciones que existen entre los elementos que componen un determinado espacio. Por lo tanto, la investigación se centró en la cuantificación de la relación estadística de un conjunto de sitios con su entorno con base en un análisis espacial y estadístico, en cómo esta integración espacial de los datos aporta información para la calibración de modelos de probabilidad utilizando la RL. Se buscó establecer con la mayor precisión posible una distancia en la que la relación entre los sitios y su entorno reporte un mejor ajuste estadístico para la elaboración del modelo de probabilidad. Con el AEVC se pudo conocer la relación que guarda el entorno de las variables explicativas con la probabilidad a ocurrencia de los deslizamientos.

El trabajo demuestra que las características de un sitio no dependen únicamente de los valores donde este se ubica, ya que este es influenciado por su zona vecina, y por lo tanto el realizar un análisis de vecindad aporta información con la cual se pueden generar modelos de RL de mayor exactitud. Lo anterior comprueba que el análisis de vecindad espacial tiene efectos positivos en la elaboración de los modelos de probabilidad a deslizamientos.

El aumento en la precisión del modelo con datos del AEVC con respecto al basado en datos *in situ* no solo se da en la parte estadística si no también la parte espacial, esto se constata por un lado con la menor cantidad de píxeles clasificados con riesgo de probabilidad, lo que implica una menor sobre evaluación de dichas áreas y se traduce en un mejor ajuste espacial en la localización y distribución de las zonas con riesgo de probabilidad a deslizamientos. Esta mayor precisión en el modelo espacial con base al AEVC también se ve reflejada en la evaluación de las matrices de clasificación que reportan mayores niveles de exactitud, así como de la eficiencia del mismo. Al aplicar el AEVC se obtuvo un modelo que representa de mejor manera la distribución de las zonas con probabilidad a deslizamientos.

Comparando los resultados con otras investigaciones que utilizan la RL, el valor de 96.8 % para la curva COR del modelo con datos del AEVC es mayor a los alcanzados por Zhou *et al.* (2018) con un valor de 69.7 % para modelo basado en la RL y en el de Aditian *et al.* (2018) con un valor

de 71 % en su modelo basado en RNA. Con respecto a la exactitud del productor alcanzado por el modelo basado en el AEVC es mayor a la reportada por Legorreta et al. (2012b), siendo para el primero de 85.94 % y para el segundo de 72.33 %.

5.2. Conclusiones

Con base en los resultados de los modelos estadístico y espacial basados en el AEVC e información *in situ*, se concluye que al aplicar el AEVC se obtiene un modelo de probabilidad a deslizamientos con mayor precisión al que se obtuvo utilizando la información tomada de manera puntual. Lo anterior se da tanto en el modelado estadístico como en el espacial.

Una parte importante en el modelado de deslizamientos es la elaboración de un inventario a detalle de las áreas que presenten dicho proceso. Se debe de tener una cartografía lo más completa posible de toda la zona de estudio, ya que con este inventario se calibraron y validaron los modelos estadístico y espacial.

La selección y procesamiento de las variables que serán utilizadas en el modelo debe de realizarse con base en el conocimiento de expertos en el tema y de ser posible con base a trabajos previos en la zona de estudio o en zonas con características similares. Las variables cartográficas que se trabajen en este tipo de modelos deben de cumplir características espaciales, temporales, temáticas y geográficas (por ejemplo: escala, resolución y proyección) para que sean compatibles con el proceso en estudio y con la metodología que se utiliza.

Con los datos reportados por el AEVC sobre las variables explicativas se pudo realizar un seguimiento puntual del comportamiento de la relación de las variables con el proceso de deslizamientos y así poder determinar con precisión cuál de los radios de vecindad analizado es el que tiene una mayor afectación sobre el conjunto de sitios para que estos desarrollen el proceso de deslizamientos. Al aplicar el análisis de vecindad a la distancia seleccionada en el AEVC para cada variable explicativas se derivaron los mapas temáticos necesarios para la elaboración de un modelo espacial de probabilidad a deslizamientos de mayor precisión que el obtenido con información *in situ*.

La elaboración de un segundo modelo de probabilidad con base a la toma de información *in situ* fue la mejor manera de evaluar la efectividad del modelo producto del AEVC ya que ambos fueron realizados con los mismos sitios de muestreo, los mismos parámetros del modelo de RL y la selección del punto de corte para los dos modelos se realizó en función de la mejor clasificación para ambas clases de la variable dependiente. Al aplicar el mismo procesamiento en la RL para la selección de las variables, se aseguró que el modelo final estuviera conformado solo por aquellas variables que reportaron una significancia estadística aceptable.

Al momento de realizar las evaluaciones también se trabajó con la misma muestra ampliada de sitios de muestreo en ambos modelos con el propósito de realizar una evaluación equivalente y así la comparación de resultados reflejara de mejor manera la ventajas o desventajas entre estos. Lo anterior garantizó la consistencia en la evaluación. Como se aprecia en los resultados de la evaluación el modelo basado en el AEVC tuvo un mejor desempeño en todos los aspectos evaluados. Desde el punto de vista geográfico es importante destacar que el modelo espacial de probabilidad a deslizamientos basado en el AEVC presento una mejor ubicación y distribución de las áreas con probabilidad lo que le da una ventaja sobre el modelo *in situ* al proporcionar información de mayor precisión espacial, dicha ventaja se verá reflejada en una mejor localización y distribución de acciones que tengan como propósito el contener o disminuir los efectos de los deslizamientos sobre las actividades antrópicas dentro de la cuenca.

Con los resultados obtenidos del AEVC se puede concluir que la evaluación de las áreas de vecindad a los sitios de muestreo permite identificar de mejor manera la relación de las variables explicativas con el proceso de deslizamientos, en comparación con la toma de datos *in situ*, con esto se puede confirmar estadísticamente el principio geográfico de conexidad. Como se pudo observar durante el análisis de las variables con el AEVC estas reportaron diferentes distancias en las que se registró una mejor relación de los datos con la evaluación del modelo, esto puede ser producto de varias características de las variables espaciales, como la distribución e intensidad de los valores que cada una de estas tiene dentro de la zona de estudio. Hasta ahora prácticamente la totalidad de los modelos espaciales basados en el modelo estadístico de RL han sido elaborados con datos tomados puntualmente lo cual como se ha señalado aíslan y limitan de manera importante la interacción un sitio con su entorno, lo que resta capacidad de evaluación al modelo estadístico que se ve reflejado en un modelo espacial de menor precisión (modelo *in situ*).

Con respecto a la aplicación del AEVC a las variables explicativas queda la posibilidad de realizar estudios en los que se evalué como sería el comportamiento del indicador -2 LL si se hace uso de la evaluación de rangos en lugar de promedios dentro de las áreas de vecindad en las variables explicativas de tipo cuantitativo, incluso el establecer estadísticamente estos rangos puede ser un campo de análisis a desarrollar. Podría esperarse que el resultado de un modelo generado con promedios sea diferente a uno generado con rangos por lo que una evaluación estadística y espacial entre ambos podrían señalar posibles ventajas y desventajas de uno u otro.

Con respecto a la forma del área de vecindad utilizada esta podría cambiarse y emplear una forma rectangular, anular o de alguna otra geometría, con el propósito de evaluar cómo afecta la forma del área de vecindad en un análisis de este tipo.

Cabe mencionar que existe un aumento importante en tiempo y cantidad de datos a procesar al aplicar el AEVC a los sitios de muestreo, el cual se ve reflejado principalmente durante su desarrollo e implementación, sin embargo, una vez que se automatiza el proceso del AEVC a los sitios de muestreo los costos de tiempo y procesamiento de datos bajo de manera muy considerable, lo que hace que la aplicación de esta metodología sea rápida y viable.

Con el apoyo de la ciencia Geomática fue posible implementar en el estudio geográfico metodologías de análisis de la información espacial que integren una gran cantidad de datos. La disponibilidad de equipo de cómputo con un alto desempeño, el uso de la programación, programas estadísticos, de SIG, programas de percepción remota, programas para tablas de datos, entre otros permitió desarrollar esta metodología de manera eficiente.

Por último, es importante mencionar que la posibilidad de utilizar el AEVC en modelos estadísticos no se limita solo a la aplicación del modelo de RL, esta metodología puede ser utilizada en otros casos en los que se emplee información de sitios de muestreo para la calibración de modelos, como son los modelos de regresión lineal. Así mismo, los temas en los que se puede aplicar pueden ser variados: modelos de deforestación, modelos de riesgo de inundaciones, de distribución de especies, etc.

De manera general y no solo en el estudio de los PG, la elaboración de modelos espaciales de probabilidad de mayor exactitud y precisión es de gran importancia en el campo de la geografía ya que a partir de una mejor ubicación y distribución de la probabilidad de ocurrencia de un determinado proceso es posible plantear tareas más precisas para tratar de minimizar o maximizar los efectos que estos procesos pudieran tener sobre alguna actividad de interés.

Referencias.

- Abdulah, L. y Yulianti, M. (2015). A Spatial Modeling Approach with Logistics Binary Regression for Bamboo Forest Distribution Mapping. *10th World Bamboo Congress*, Korea.
- Aceves, J. F., López-Blanco, J. y Martín del Pozzo, A. L. (2006). Determinación de peligros volcánicos aplicando técnicas de evaluación multicriterio y SIG en el área del Nevado de Toluca, centro de México. *Revista Mexicana de Ciencias Geológicas*, 23 (2), 113-124.
- Aceves, J. F., Legorreta, G. y Álvarez-Ruíz, Y. (2014a). Gravitational processes on the Eastern flank of the Nevado de Toluca, México. *Zeitschrift für Geomorphologie*, 58 (2), 185–200.
- Aceves, J. F., Legorreta, G. y Álvarez-Ruíz, Y. (2014b). Cartografía geomorfológica para el inventario de procesos gravitacionales en la cuenca endorreica del arroyo La Ciénega, flanco oriental del volcán Nevado de Toluca. *Boletín de la Sociedad Geológica Mexicana*, 66 (2), 329-342.
- Aceves, J. F., Legorreta, G., Lugo, J., Umaña, J., y Legorreta, H. A. (2016). Sistemas de información geográfica y cartografía geomorfológica aplicados al inventario de deslizamientos y cartografía de susceptibilidad en la cuenca del río El Estado, Pico de Orizaba, México. *Investigaciones Geográficas*, 2016 (91), 43–55. <https://doi.org/10.14350/rig.46503>
- Aditian, A., Kubota, T. y Shinohara, Y. (2018). Comparison of GIS-based landslide susceptibility model using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia. *Geomorphology*, 318, 101-111. <https://doi.org/10.1016/j.geomorph.2018.06.006>
- Aguilar-Barojas, S. (2005). Fórmulas para el cálculo de la muestra en investigaciones de salud. (2005). *Salud en Tabasco*, 11 (1–2), 333–338.
- Alcántara-Ayala, I. (2000). Landslides: ¿deslizamientos o movimientos del terreno? Definición, clasificaciones y terminología. *Investigaciones Geográficas*, 41, 7–25. https://doi.org/http://www.scielo.org.mx/scielo.php?pid=S0188-46112000000100002&script=sci_arttext
- Álvarez, R., Bonifaz, R., Lunetta, R.S., García, G., Gómez, G., Castro, R. ... Cabrera, A. L. (2003). Multitemporal Land-cover classification of México using Landsat MSS imagery. *International Journal Remote Sensing*, 24 (12), 2501-2514
- Álvarez, Y. (2015). *Relación entre los deslizamientos y cambios en la cubierta de uso de suelo y vegetación entre los años 1983 y 2014 en el arroyo la Ciénega, volcán Nevado de Toluca*,

- México. (Tesis para obtener el grado de bióloga). Facultad de Ciencia. Universidad Nacional Autónoma de México, México.
- Aronoff, S. (1989). *Geographic Information System: A Management Perspective*. Ottawa, Canada. WDL Publications.
- Atkinson, P. M. y Massari, R. (2011). Autologistic modelling of susceptibility to landsliding in the Central Apennines, Italy. *Geomorphology*, 130, 55-64. DOI: 10.1016/j.geomorph.2011.02.001
- Augustin, N., Cummins, R. y French, D. (2001). Exploring spatial vegetation dynamics using logistic regression and a multinomial logit model. *Journal of Applied Ecology*, (38), 991-1006.
- Ayalew, L. y Yamagishi, H. (2005). The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko mountains, central Japan. *Geomorphology* 65, 15 - 31 pp. doi:10.1016/j.geomorph.2004.06.010.
- Bannari, A., Morin, D., Bonn, F., y Huete, A. R. (1995). A review of vegetation indices. *Remote sensing reviews*, 13(1-2), 95-120.
- Berry, J. K. (2013a). *Beyond Mapping II – Spatial Reasoning*. Compilation of Beyond Mapping columns appearing in GeoWorld magazine 1993 to 1996. Basis Press. Recuperado de http://www.innovativegis.com/basis/BeyondMappingSeries/BeyondMapping_II/Default.htm
- Berry, J. K. (2013b). *Beyond Mapping III – Map Analysis*. Compilation of Beyond Mapping columns appearing in GeoWorld magazine 1996 to 2007. Basis Press. Recuperado de http://www.innovativegis.com/basis/BeyondMappingSeries/BeyondMapping_III/Default.htm
- Capra, L., y Macías, J. L. (2000). Pleistocene cohesive debris flows at Nevado de Toluca Volcano, central Mexico. *Journal of Volcanology and Geothermal Research*, 102 (1–2), 149–167. [https://doi.org/10.1016/S0377-0273\(00\)00186-4](https://doi.org/10.1016/S0377-0273(00)00186-4)
- Capra, L., Norini, G., Groppelli, G., Macías, J. L., y Arce, J. L. (2008). Volcanic hazard zonation of the Nevado de Toluca volcano, México. *Journal of Volcanology and Geothermal Research*, 176 (4), 469–484. <https://doi.org/10.1016/j.jvolgeores.2008.04.016>
- Castro, R. (2011). *Elaboración de un modelo espacial de riesgo de deforestación en bosques*. (Tesis para obtener el grado de Licenciado en Geografía). Facultad de Filosofía y Letras, Universidad Nacional Autónoma de México. Distrito Federal, México.

- Castro, R. (2014): *Modelo especial de probabilidad a la deforestación en bosques para el estado de Oaxaca*. (Tesis para obtener el grado de maestro en Geografía). Posgrado en Geografía, Universidad Nacional Autónoma de México. Distrito Federal, México.
- CENAPRED (Centro Nacional de Prevención de Desastres) (2001). *Diagnóstico de peligros e identificación de riesgos de desastres en México*. Distrito Federal, México: Secretaría de Gobernación.
- Chao, L. (2002). *Introducción a la estadística*. Distrito Federal, México: Continental.
- Chalkias, C., Ferentinou, M. y Polykretis, C. (2014). GIS-Based Landslides Susceptibility Mapping on the Peloponnese peninsula, Greece. *Geosciences*, 4, 176-190.
- Dehban, A. S. (2013): Landslide Susceptibility Mapping for Poulrood Earth Fill Dam Reservoir (The Comparison of Two Methods) Iran, Guilan Province. *Universal Journal of Geoscience*, 1 (2), 69-79.
- DiGruttolo, N., y Mohamed, A. H. (2011). *Geomatics core competencies*. Recuperado de https://www.researchgate.net/publication/292665600_Geomatics_core_competencies
- Environmental Systems Research Institute (ESRI) (2001). *ArcGis™ Spatial Analyst: Advanced GIS Spatial Analysis Using Raster and Vector Data*. California: ESRI.
- Environmental Systems Research Institute (ESRI) (2010). Arc/Info Help. *Surface hydrologic analysis: Watersheds*. California: ESRI.
- Environmental Systems Research Institute (ESRI) (2016). *ArcGis Desktop*. Recuperado de <http://desktop.arcgis.com/en/arcmap/>
- Eskandari, S., y Chuvieco, E. (2015). Fire danger assessment in Iran based on geospatial information. *International Journal of Applied Earth Observation and Geoinformation*, 42, 57–64. <https://doi.org/10.1016/j.jag.2015.05.006>
- Esper, A. (2013). Debris flow susceptibility mapping in a portion of the Andes and Preandes of San Juan, Argentina using frequency ratio and logistic regression models. *Earth Sciences Research Journal*, 17 (2), 159-167.
- Espinosa-Rodríguez, L. M. (2003). Elementos geomorfológico-cuantitativos del piedemonte nororiental del volcán Nevado de Toluca. *Ciencia Ergo Sum*, 10 (2), 187-195. Recuperado de <https://www.redalyc.org/articulo.oa?id=10410208>
- Espinosa-Rodríguez, L. M., Balderas-Plata, M. A. y Cabañas-Báez, H. V. (2014). Caracterización geomorfológica del área natural protegida Nevado de Toluca: complejo de volcanes Nevado de Toluca y San Antonio. *CienciaUAT*, 9 (1), 6-14: ISSN 2007-7521. Recuperado de http://www.scielo.org.mx/scielo.php?script=sci_abstract&pid=S2007-78582014000100006&lng=es&nrm=iso&tlng=es

- Feizizadeh, B. y Blaschke, T. (2014). An uncertainty and sensitivity analysis approach for GIS-based multicriteria landslide susceptibility mapping. *International journal of geographical information science*, 28 (3), 610-638. <http://dx.doi.org/10.1080/13658816.2013.869821>.
- Franco, S., Regil, H. H., González, C. y Nava G. (2006). Cambio de uso del suelo y vegetación en el Parque Nacional Nevado de Toluca, México, en el periodo 1972-2000. *Investigaciones Geográficas* (61). 38-57. ISSN 0188-4611.
- Free Software Foundation Inc. (FSFI) (2016). QGIS 3.6.3-Noosa. Recuperado de <https://qgis.org/en/site/>
- García, E. (2006). *Modificaciones al sistema de clasificación climática de Köppen*. Ciudad de México, México: Instituto de Geografía-UNAM.
- García, S. (2017). *Susceptibilidad geomorfológica a deslizamientos en arroyos secundarios del sur de la cuenca la Ciénega, volcán Nevado de Toluca*. (Tesis para obtener el grado de Licenciada en Geografía). Facultad de Filosofía y Letras, Universidad Nacional Autónoma de México. Distrito Federal, México.
- García, S., Legorreta, G., y Aceves, F. (2018). Susceptibilidad geomorfológica a procesos gravitacionales en arroyos secundarios del sur de la cuenca La Ciénega, volcán Nevado de Toluca. *Investigaciones Geográficas*, (96).1-14. <https://doi.org/10.14350/rig.59613>
- García-Palomo, A., Macías, J. L., y Garduño, V. H. (2000). Miocene to recent structural evolution of the Nevado de Toluca volcano region, Central Mexico. *Tectonophysics*, 318, 281–302. [https://doi.org/10.1016/S0040-1951\(99\)00316-9](https://doi.org/10.1016/S0040-1951(99)00316-9)
- García-Palomo, A., Macías, J. L., Arce, J. L., Capra, L., Garduño, V. H., y Espíndola, J. M. (2002). Geology of Nevado de Toluca Volcano and surrounding areas, central Mexico. *Geological Society of America Map and Chart Series*, MCH089, 26.
- Ghilani, D. and Wolf, P. (2012). *Elementary Surveying. An Introduction to Geomatics*, (13). New Jersey: Prentice Hall, Pearson Education, Inc.
- Gomasasca, M. A. (2009). Basics of geomatics. *Basics of Geomatics*, (September), 1–656. <https://doi.org/10.1007/978-1-4020-9014-1>
- Hair, J., Anderson, E., Tatham, R. y Black W. (1999). *Análisis Multivariante*. Madrid: Pearson.
- Hernández, R., Fernández, C. y Baptista, P. (2014). *Metodología de la investigación*. México: McGraw-Hill / Interamericana Editores, S.A. de C.V.
- Hong, H., Pradhan, B., Xu, C., y Tien Bui, D. (2015). Spatial prediction of landslide hazard at the Yihuang area (China) using two-class kernel logistic regression, alternating decision tree and support vector machines. *Catena*, 133, 266–281. <https://doi.org/10.1016/j.catena.2015.05.019>

- Huang, J., Zhou, Q., y Wang, F. (2015). Mapping the landslide susceptibility in Lantau Island, Hong Kong, by frequency ratio and logistic regression model. *Annals of GIS*, 21(3), 191–208. <https://doi.org/10.1080/19475683.2014.992373>
- Ibarra-Montoya, J. L., Rangel-Peraza, G., González-Farías, F. A., Anda, J., Martínez-Meyer, E. y Macías- Cuellar, H. (2012). Uso del modelado de nicho ecológico como una herramienta para predecir la distribución potencial de *Microcystis sp* (cianobacteria) en la Presa Hidroeléctrica de Aguamilpa, Nayarit, México. *Ambi-Agua*, 7 (1), 218-234. DOI: 10.4136/1980-993X
- Instituto Nacional de Estadística y Geografía (1998). *Carta topográfica en escala 1:50,00*. Tenango de Arista E14A48. Recuperado de <http://www.inegi.org.mx>
- Instituto Nacional de Estadística y Geografía (2001a). *Carta geológica en escala 1:50,00*. Recuperado de <http://www.inegi.org.mx>
- Instituto Nacional de Estadística y Geografía (2001b). *Carta edafológica en escala 1:50,00*. Recuperado de <http://www.inegi.org.mx>
- Instituto Nacional de Estadística y Geografía (2005). *Guía para la interpretación de la carta geológica*. Recuperado de <http://www.inegi.org.mx>
- Instituto Nacional de Estadística y Geografía (2009). *Carta topográfica en escala 1:50,00*. Volcán Nevado de Toluca E14A47. Recuperado de <http://www.inegi.org.mx>
- Instituto Nacional de Estadística y Geografía (2010). *Red hidrográfica escala 1:50,000*. Documento técnico descriptivo de la red hidrográfica escala 1:50,000. Edición 2.0. Dirección general de geografía y medio ambiente. Recuperado de <http://www.inegi.org.mx>
- Instituto Nacional de Estadística y Geografía (2011). *Guía para la interpretación de cartografía edafología: escala 1:250,000: serie II*. Recuperado de <http://www.inegi.org.mx>
- Instituto Nacional de Estadística y Geografía (2012). *Carta topográfica en escala 1:20,00*. E14A48d. Recuperado de <http://www.inegi.org.mx>
- Instituto Nacional de Estadística y Geografía (2013a). *Continuo de Elevación Mexicano 3.0 (CEM 3.0)*. Recuperado de <http://www.inegi.org.mx>
- Instituto Nacional de Estadística y Geografía (2013b). *Censo de Población y Vivienda 2010*. Recuperado de <http://www.inegi.org.mx>
- Instituto Nacional de Estadística y Geografía (2013c). *Conjunto de datos vectoriales de información topográfica en escala 1:50,00*. Serie III. Tenango de Arista E14A48. Recuperado de <http://www.inegi.org.mx>

- Instituto Nacional de Estadística y Geografía (20014a). *Conjunto de datos vectoriales de información topográfica en escala 1:50,00*. Serie III. Volcán Nevado de Toluca E14A47. Recuperado de <http://www.inegi.org.mx>
- Instituto Nacional de Estadística y Geografía (2014b). *Guía para la interpretación de cartografía edafología: escala 1:250,000: serie III*. Recuperado de <http://www.inegi.org.mx>
- Instituto Nacional de Estadística y Geografía (2015). *Guía para la interpretación de cartografía. Uso del suelo y vegetación. Escala 1:250,000, serie V*. Aguascalientes, México. Recuperado de <http://www.inegi.org.mx>
- Instituto Nacional de Estadística y Geografía (2016). *Marco Geoestadístico versión junio de 2016*. Recuperado de <http://www.inegi.org.mx>
- Intergraph Corporation (2012). *ERDAS IMAGINE 2013, versión 13.0. Resolution Merge. Desktop help*. Recuperado de C:\Intergraph\ERDASIMAGINE2013\help\html\ii_pan_sharpen\resolution_merge.html
- Jiménez-Perálvarez, J., Irigaray, C., El Hamdouni, R. y Chacón, J. (2008). Building models for automatic landslide-susceptibility analysis, mapping and validation in ArcGis. *Nat Hazards*. DOI 10.1007/s11069-008-9305-8
- Kemp, K. (2008). *Encyclopedia of Geographic Information Science*. California, EE.UU. SAGE Publications, Inc.
- Leal-Nares, O., Mendoza, M., Pérez-Salicrup, D., Geneletti, D., Lopéz-Granados, E. y Carranza, E. (2012). Distribución potencial del *Pinus martinezii*: un modelo espacial basado en conocimiento ecológico y análisis multicriterio, *Revista Mexicana de Biodiversidad*, 83, 1152-1170. DOI: 10.7550/rmb.27199
- Legorreta, G. y Bursik, M. (2009). Logisnet: A tool multimethod, multiple soil layers slope stability analysis. *Computers & Geosciences*, 35 (5), 1007-1016. doi:10.1016/j.cageo.2008.04.003
- Legorreta, G., Bursik, M., Aceves, F., Paredes, M., Lugo, J., Alcántara, I. y López, J. (2012a). Evaluación práctica y teórica de los modelos de regresión logística múltiple y cartográfica-hidrológica para la predicción de deslizamientos de tierra. En José Antonio Quintero Pérez (Coord.), *Temas selectos de Geomática* (pp. 47-76). Ciudad Universitaria, México: Instituto de Geografía.
- Legorreta, G., Bursik, M., Ramírez-Herrera, M. T., Contreras, T., Polenz, M., Hubp, J. L., ... Arana, L. (2012b). Landslide inventory mapping and landslide susceptibility modeling assessment on the SW flank of Pico de Orizaba volcano, Puebla-Veracruz, Mexico. *Zeitschrift Für Geomorphologie*, 57 (3), 371–385. <https://doi.org/10.1127/0372-8854/2012/0097>

- Legorreta, G., Bursik, M., Solene, P., Hubp, J. L., Paredes, L. M. y Aceves, F. (2014a). Inventario multitemporal, análisis de susceptibilidad y estimación de volumen de deslizamiento en el flanco SW del volcán Pico de Orizaba, Puebla-Veracruz. *Boletín de la Sociedad Geológica Mexicana*, 66 (2), 343-354.
- Legorreta, G., Bursik, M., Aceves, F., Paredes, L. M., Hubp, J. L., Ayala, I. y López, J. (2014b). *Temas selectos de geomática: Métodos y aplicaciones. Evaluación práctica y teórica de los modelos de regresión logística múltiple y cartográfica-hidrográfica para la predicción de deslizamientos de tierra*. Distrito federal, México: Instituto de Geografía, UNAM.
- López-Sandoval, J. A., López-Mata, L., Cruz-Cárdenas, G., Vibrans, H., Vargas, O. y Martínez, M. (2015). Modelado de los factores ambientales que determinan la distribución de especies sinantrópicas de *Physalis*. *Botanical Sciences*, 93 (4), 755-764. DOI: 10.17129/botsci.192
- Lugo, J. (2011). *Diccionario geomorfológico*. Distrito Federal, México. Instituto de Geografía, UNAM.
- Macías, J. L. (2005). Geología e historia eruptiva de algunos de los grandes volcanes activos de México. *Boletín de La Sociedad Geológica Mexicana*, 57(3), 379–424.
- Martín, Q., Cabero, M. T. y Paz, Y. R. (2008). *Tratamiento estadístico de datos con SPSS*. Madrid, España. Thomson.
- MatebieMeten, PrakashBhandary, N. y RyuichiYatabe (2014). *The application of weights of evidence modeling for landslide susceptibility mapping of Dejan-GohaTsiyon transect in the Blue Nile Gorge, central Ethiopia*. International symposium. Goehazard: Science, engineering and management. November 20 - 21. Kathmandu, Nepal.
- Mas, J. F., Velásquez, A., Castro, R. y Schmitt, A. (2002). Una evaluación de los efectos del aislamiento, la topografía, los suelos y el estatus de protección sobre las tasas de deforestación en México. *RA'EGA o espaço geográfico em análise*, (6), 61-73
- McGrew, J. C. y Monroe, C. B. (2000). *An Introduction to statistical problem solving in geography*. United States of America: McGraw Hill.
- Mirnazari, J., Ahmad, B., Mojaradi, B. y Sattari, F. (2014). Using frequency ratio for spatial landslide prediction. *Research Journal of Applied Sciences, Engineering and Technology*, 7 (15). 3174 - 3180.
- Navarro, J. A. (2011). Introducción al diseño y análisis del muestreo de poblaciones finitas. En F. Bautista (Ed), *Técnicas de muestreo para manejadores de recursos naturales*. (pp. 3-58). Ciudad de México, México: UNAM.

- Nemmaoui, A., y García-Lorca, A. M., Aguilar, F. J., y Aguilar, M. A. (2013). Series temporales de vegetación para un modelo forestal destruido: El caso de Tadla Azilal (Morocco). *Agrociencia*, 47 (3), 267-280. ISSN: 1405-3195. <https://www.redalyc.org/articulo.oa?id=302/30226978006>
- Norini, G., Groppelli, G., Capra, L., y De Beni, E. (2004). Morphological analysis of Nevado de Toluca volcano (Mexico): New insights into the structure and evolution of an andesitic to dacitic stratovolcano. *Geomorphology*, 62, 47–61. <https://doi.org/10.1016/j.geomorph.2004.02.010>
- Norini, G., Capra, L., Groppelli, G., y Lagmay, A. M. F. (2008). Quaternary sector collapses of Nevado de Toluca volcano (Mexico) governed by regional tectonics and volcanic evolution. *Geosphere*, 4 (5), 854-871. <https://doi.org/10.1130/GES00165.1>
- Nourani, V. y Ghaffari, H. (2012). *Landslide susceptibility mapping with analytical hierarchical process and GIS techniques (case study: Zonouz, East Azerbaijan, IRAN)*. 10th International congress on advances in civil engineering. 17 - 19 october. Middle east technical university, Turkey.
- Othman, A. A., Gloaguen, R., Andreani, L. y Rahnama, M. (2018). Improving landslide susceptibility mapping using morphometric features in the Mawat area, Kurdistan Region. NE Iraq: Comparison of different statistical models. *Geomorphology*, 319, 147-160. doi.org/10.1016/j.geomorph.2018.07.018
- Pardo, A. y Ruiz, M. A. (2002). *SPSS 11*. Madrid: Mc Graw Hill.
- Paz, J. A., González, R., Gómez, M. y Velasco, J. A. (2017). Metodología para elaborar mapas de susceptibilidad a procesos de remoción en masa, análisis del caso ladera sur de Tuxtla Gutiérrez, Chiapas. *Investigaciones Geograficas*, (92), 128-143. <https://doi.org/10.14350/rig.52822>
- Peña, I. C. (2013). *Modelo cuantitativo para inestabilidad de laderas en un ambiente SIG, aplicado en la Sierra Norte de Chiapas*. (Tesis para obtener el grado de Licenciado en Geografía). Facultad de Filosofía y Letras, Universidad Nacional Autónoma de México. Distrito Federal, México.
- Pham, B. T., Prakash, I., y Tien Bui, D. (2018). Spatial prediction of landslides using a hybrid machine learning approach based on Random Subspace and Classification and Regression Trees. *Geomorphology*, 303, 256–270. <https://doi.org/10.1016/j.geomorph.2017.12.008>

- Pinto, F. (2013). *Geomática Tecnologías de punta*. Recuperado de https://kupdf.net/download/libro-geomatica-tecnologias-de-punta-pdf_5913f9b6dc0d60ea6be5e554_pdf.
- Rojas, T. (2005). Epistemología de la geografía...una aproximación para entender esta disciplina. *Terra Nueva Etapa*. XXI (30), 141-162. ISBN: 1012-7089
- Rodríguez, S., Mora-González, I. y Murrieta, H. (2006). Flujos de baja concentración asociados con lluvias de intensidad extraordinaria en el flanco sur del volcán Pico de Orizaba (Citlaltepetl), México. *Boletín de la Sociedad Geológica Mexicana*, LVIII (2), 223-236.
- Samuelsen, R. (2008). The field of Geomatics: Scientific, Technological, Business and Infrastructural Aspects. *Nordlandsforskning*. (1027), pp. 1-30.
- Sappington, J.M., K.M. Longshore, and D.B. Thomson. 2007. Quantifying Landscape Ruggedness for Animal Habitat Anaysis: A case Study Using Bighorn Sheep in the Mojave Desert. *Journal of Wildlife Management*. 71(5): 1419 -1426
- Schlögel, R., Marchesini, I., Alvioli, M., Reichenbach, P., Rossi, M., y Malet, J. (2018). Optimizing landslide susceptibility zonation: Effects of DEM spatial resolution and slope unit delineation on logistic regression models. *Geomorphology*, 301, 10–20. <https://doi.org/10.1016/j.geomorph.2017.10.018>
- Selby, M. J. (1993). *Hillslope materials and processes*. New York: Oxford University Press.
- Servicio de Información Agroalimentaria y Pesquera – Secretaria de la Defensa Nacional (SIAP-SEDENA) (2016). Spot6_MS y Spot6_P. Ladera este volcán Nevado de Toluca.
- Siso, G. (2010). ¿Qué es la Geografía? *Terra Nueva Etapa*, XXVI (39), 147-182.
- Sopeña, A. y Sánchez-Moya, Y. (2010). IV Los sistemas aluviales. En A. Arche. (Ed.), *Sedimentología del proceso físico a la Cuenca sedimentaria* (pp. 73-83). Madrid, España: Consejo Superior de Investigación Científica.
- Spiegel, M. R. y Stephens, L. J. (2009). *Estadística*. México: McGraw-Hill / Interamericana Editores, S.A. de C.V.
- Tayyebi, A., Reza, M., Yazdanpanah, M., Pijanowsky, B., Saeedi, S. y Hossein, A. (2010). A spatial logistic regression model for simulating land use patterns: A case study of the Shiraz Metropolitan Area of Iran. En Chuvieco, E., Jonathan, L. y Xiaojun, Y. (Ed.), *Advances in Earth Observation of Global Change* (pp. 27.42). Springer.
- Tobler, R., W. (1970). A Computer Movie Simulation Urban Growth in the Detroit Region. *Economic Geography*, 46, 234-240. <https://doi.org/10.1126/science.11.277.620>.

- Toscana, A. y Valdez, V. (2014). Representaciones sociales del desastre de 1940 en Santa Cruz Pueblo Nuevo, Estado de México. *Investigaciones Geográficas*, 83, 88-101. doi: 10.14350/rig.36918.
- Tsangaratos, P. y Rozos, D. (2013). Producing landslide susceptibility maps by applying expert knowledge in a GIS – based environment. *Bulletin of the Geological Society of Greece*, vol. XLVII. Proceeding of the 13th International Congress, Chania.
- Vargas, G. (2012). Espacio y territorio en el análisis geográfico. *Reflexiones*, 91 (1): 313-326. ISBN: 1021-1209/2012
- Varnes, D. J. (1978). Slope movement types and processes. R. L. Schuster and R. J. Krizec. (Eds.), *Special report 176. Landslides: Analysis and control* (pp. 11-33). Washington D. C., EEUU: TRB, National Research Council.
- Vilchis-Francés, A., Díaz-Delgado, C., Magaña-Lona, D., Bâ, K. y Gómez-Albores, M. A. (2015). Modelo espacial para peligro de incendios forestales con predicción diaria en la cuenca del río Balsas. *Agrociencias [online]*, 49 (7), 803-820. ISSN 2521-9766
- Von Ruetten, J., Papritz, A., Lehmann, P., Rickli, C., y Or, D. (2011). Spatial statistical modeling of shallow Landslides-Validating predictions for different landslide inventories and rainfall events. *Geomorphology*, 133(1–2), 11–22. <https://doi.org/10.1016/j.geomorph.2011.06.010>
- Washington State Department of Natural Resources (2006a). *Landslide Hazard Zonation Project Protocol version 2.1*. Recuperado de https://www.dnr.wa.gov/publication/fp_lhz_protocol_v2_1_final.pdf?ee5umyf
- Washington State Department of Natural Resources (2006b). *Forest Practices Board Manual*. Recuperado de http://file.dnr.wa.gov/publications/fp_board_manual.pdf
- Wayne, W. D. (2004): *Bioestadística: Base para el análisis de las ciencias de la salud*. Distrito federal, México: LIMUSA S.A. de C.V. ISBN 968-18-61647.
- Xiong, Y., y Zuo, R. (2018). GIS-based rare events logistic regression for mineral prospectivity mapping. *Computers and Geosciences*, 111 (October 2017), 18–25. <https://doi.org/10.1016/j.cageo.2017.10.005>
- Yang, J., Song, C., Yang, Y., Xu, C., Guo, F., y Xie, L. (2019). New method for landslide susceptibility mapping supported by spatial logistic regression and GeoDetector : A case study of Duwen Highway Basin, Sichuan Province, China. *Geomorphology*, 324, 62–71. <https://doi.org/10.1016/j.geomorph.2018.09.019>
- Zambrano, L., Pacheco-Muñoz, R., y Fernández, T. (2017). A spatial model for evaluating the vulnerability of water management in Mexico City, Sao Paulo and Buenos Aires

considering climate change. *Anthropocene*, 17, 1–12.
<https://doi.org/10.1016/j.ancene.2016.12.001>

Zambrano, L., Pacheco-Muñoz, R., y Fernández, T. (2018). Influence of solid waste and topography on urban floods: The case of Mexico City. *Ambio*, 47 (7), 771-780.
<https://doi.org/10.1007/s13280-018-1023-1>

Zhou, C., Yin, K., Cao, Y., Ahmed, B., Li, Y., Catani, F., y Pourghasemi, H. R. (2018). Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir area, China. *Computers and Geosciences*, 112(April 2017), 23–37. <https://doi.org/10.1016/j.cageo.2017.11.019>