

**Universidad Pública de Navarra**

**Nafarroako Unibertsitate Publikoa**

**ESCUELA TECNICA SUPERIOR  
DE INGENIEROS AGRONOMOS**

***NEKAZARITZAKO INGENIARIEN  
GOI MAILAKO ESKOLA TEKNIKOA***

**EVALUACIÓN Y COMPARACIÓN DE UN MODELO DE CLASIFICACIÓN  
GEOMORFOMÉTRICA AUTOMÁTICA A PARTIR DE DOS MDE:  
SRTM (90M) Y ALOS PALSAR (12,5M)**

presentado por

**JOSÉ GREGORIO ROA LOBO**

**MÁSTER EN SISTEMAS DE INFORMACIÓN GEOGRÁFICA Y TELEDETECCIÓN  
*MASTERRA INFORMAZIO SISTEMA GEOGRAFIKOETAN ETA TELEDETEKZIOAN***

Septiembre, 2019

### Resumen

El objetivo a desarrollar en este trabajo es la evaluación de un modelo de clasificación geomorfométrica automática ejecutado con el algoritmo KNN (vecino más cercano) y comparado con la clasificación obtenida de un algoritmo empírico (AE) manual, ambos aplicados a dos modelos digitales de elevación (MDE) ALOS PALSAR (12,5m) y SRTM (90m). El área de trabajo corresponde a una cuenca de montaña tropical localizada en los Andes venezolanos, con rangos de altitud ente los 150 y 5000 msnm, y una extensión de 3517 Km<sup>2</sup>. Las clasificaciones obtenidas fueron validadas con áreas de comprobación previamente fotointerpretadas encontrándose al método KNN y al MDE ALOS PALSAR (12,5m), como el mejor clasificador y MDE respectivamente, no evidenciándose ventaja alguna en la utilización de alguna zona de análisis en particular, sin embargo, se sugiere cautela con el uso de la matriz de confusión e índices derivados como método de validación debido a la generalización implícita de sus resultados.

**Palabras clave:** geomorfometría, modelos digitales de elevación, formas del paisaje.

### Abstract

Main objective in this work is the assessment of an automatic geomorphometric classification model executed with the KNN algorithm (*k* nearest neighbors) and compared to the classification obtained from a manual empirical algorithm (AE), both applied over two digital elevation models (DEM) ALOS PALSAR (12.5m) and SRTM (90m). The study area is a tropical mountain catchment located in the Venezuelan Andes, which ranges between 150 and 5000 m.a.s.l, and a 3517 km<sup>2</sup> area. Classifications obtained were validated with previously photointerpreted test areas, finding the KNN method and the ALOS PALSAR DEM (12.5m), as the best classifier and DEM respectively, with not evidence of any advantage in the use of some analysis area in particular, however, caution is suggested with the use of the confusion matrix and derived indices as a validation method due to the implicit generalization of its results.

**Key words:** geomorphometry, digital elevation model, landform.

## Índice

1.- INTRODUCCIÓN	5
1.1.- Definición del problema	7
1.2.- La clasificación geomorfométrica. Parámetros y clases	8
2.- OBJETIVOS DE LA INVESTIGACIÓN	9
3.- MATERIALES Y MÉTODOS	9
3.1.- El área de estudio	10
3.2.- Diseño de investigación	13
3.2.1.- Datos fuente. Los MDE Alos Palsar y SRTM	13
3.2.2.- Datos insumos. Parámetros geomorfométricos. Algoritmos de clasificación	14
3.2.3.- Datos producto. Las 13 clases geomorfométricas	15
3.2.4.- Datos validación. Las áreas de comprobación. La matriz de confusión	16
3.2.5.- Procedimiento metodológico general	19
3.3.- El algoritmo empírico (AE) de clasificación geomorfométrica	20
3.3.1.- Definición de las formas, parámetros y criterios de trabajo	20
3.3.2.- Esquema de clasificación	21
3.4.- El modelo automatizado KNN de clasificación geomorfométrica	22
3.4.1.- Definición de las formas, parámetros y criterios de trabajo	22
3.4.2.- Esquema de clasificación	23
4.- RESULTADOS Y DISCUSIÓN	24
4.1.- Expresión espacial de los criterios y formas bajo la clasificación AE	24
4.2.- Expresión espacial de los criterios y formas bajo la clasificación KNN	28
4.3.- Matriz de confusión. Medidas de rendimiento	30
4.4.- La clasificación geomorfométrica a partir del MDE SRTM 90 m	31
- La clasificación del SRTM según el AE	31
- La clasificación del SRTM según el KNN	32
4.5.- La clasificación geomorfométrica a partir del MDE Alos Palsar 12,5m	34
- La clasificación del Alos Palsar según el AE	34
- La clasificación del Alos Palsar según el KNN	35
4.6.- Análisis comparativo	37
5.- CONCLUSIONES	43
6.- BIBLIOGRAFÍA	45
7.- ANEXOS	48

**Índice de figuras**

Figura 1. Localización esquemática de los métodos de clasificación geomorfométrica	8
Figura 2. Área de estudio	10
Figura 3. Localización de los bloques de análisis	11
Figura 4. Visualización comparativa de características morfométricas MDE ALOS PALSAR (12,5m)	12
Figura 5. Visualización comparativa de características morfométricas MDE SRTM (90m)	13
Figura 6. Área de comprobación en un sector de la zona de páramo (A)	16
Figura 7. Área de comprobación en dos sectores de la zona de vertientes (B)	17
Figura 8. Flujoograma del proceso metodológico	20
Figura 9. Proceso de clasificación según el algoritmo empírico (AE)	22
Figura 10. Proceso de clasificación según el algoritmo KNN	24
Figura 11. Ejemplos comparativos de los parámetros y criterios según el AE y MDE ALOS PALSAR	25
Figura 12. Ejemplos comparativos de los parámetros y criterios según el AE y MDE SRTM	26
Figura 13. Definición de clases geomorfométricas aplicando el AE con MDE ALOS PALSAR.	27
Figura 14. Definición de clases geomorfométricas aplicando el AE con MDE SRTM	28
Figura 15. Ejemplos comparativos de la clasificación KNN con datos obtenidos del MDE ALOS PALSAR	29
Figura 16. Ejemplos comparativos de la clasificación KNN con datos obtenidos del MDE SRTM	30
Figura 17. Distribución de las medidas de rendimiento por método y áreas del MDE SRTM	33
Figura 18. Distribución de las medidas de rendimiento por método y áreas del MDE ALOS PALSAR	36
Figura 19. Distribución de los índices kappa y EG por método, áreas de aplicación y MDE	38
Figura 20. Generalización de clases en zona de páramo	41
Figura 21. Generalización de clases en zona de vertiente	42

**Índice de tablas**

Tabla 1. Métodos, clasificador y unidades de clasificación	6
Tabla 2. Descripción de algunos parámetros geomorfométricos	14
Tabla 3. Descripción de las 13 clases geomorfométricas a clasificar	15
Tabla 4. Grado de acuerdo según valores del índice Kappa	18
Tabla 5. Definición y parámetros de las medidas de rendimiento	19
Tabla 6. Formas, parámetros y criterios según el algoritmo empírico de clasificación (AE)	21
Tabla 7. Índices de las medidas de rendimiento derivadas de los parámetros del MDE SRTM	33
Tabla 8. Índices de las medidas de rendimiento derivadas de del MDE ALOS PALSAR	36
Tabla 9. Distribución comparativa de las clasificaciones geomorfométricas según su alcance	38
Tabla 10. Primeras 10 clases geomorfométricas con mejor rendimiento según según el IH	40
Tabla 11. Últimas 10 clases geomorfométricas con menor rendimiento según el IH.	40

## 1.- INTRODUCCIÓN

El presente estudio evalúa el uso y aplicabilidad de modelos digitales de elevación (MDE), tales como los obtenidos a partir de los sistemas SRTM 90m v4 (Lin., et al 2013), y ALOS PALSAR 12,5m (ASF DAAC, 2019), en la clasificación geomorfométrica. La libre disponibilidad de estos MDE, de casi cualquier parte del mundo, unido al rápido avance teórico y técnico de los sistemas de información geográfica (SIG), ha facilitado la evolución de la geomorfometría como el arte de la cuantificación topográfica, cuyo enfoque principal es la extracción automática de parámetros geomorfométricos. Por esta razón la propuesta aquí presentada busca comparar dos clasificaciones geomorfométricas de un área común, la primera desarrollada a partir de un algoritmo empírico y la segunda a partir de un proceso automatizado basado en el *K-Nearest-Neighbor* (KNN), el cual es un algoritmo basado en instancia de tipo supervisado y de aprendizaje automatizado. A su vez cada clasificación geomorfométrica basada en cada uno de los métodos descritos, se aplicará tanto a los datos extraídos del MDE ALOS PALSAR 12,5m, como del SRTM 90m, de esta manera la comparación de los productos generados no solo evaluará la eficiencia de los métodos de clasificación sino describirá el efecto de la resolución espacial de los MDE en la definición de las clases inicialmente determinadas.

En los últimos años la disponibilidad de información topográfica y de elevación a nivel mundial proveniente de plataformas satelitales ha fomentado la cuantificación de los análisis basados en las formas del terreno o del paisaje, debido a que el desarrollo de este frente de estudio de las geociencias, reduciría la complejidad de análisis de la superficie terrestre permitiendo un mejor entendimiento, manejo y modelizaje de los procesos asociados.

La cartografía de la geomorfometría del relieve reviste la subdivisión del paisaje en unidades morfométricas basado en una metodología de segmentación del terreno que evalúa su tamaño y forma, estableciendo la relación entre estas (Miliareisis G. 2006). Obviamente la segmentación del terreno según su expresión cartográfica puede tener diferente perspectiva, por ejemplo de un MDE se esperaría una segmentación por formas geométricas mientras que de una fotointerpretación de imágenes ópticas una zonificación fisiográfica. De aquí que los métodos de clasificación geomorfométrica pueden dividirse en los basados en celda (*cell-based*) y los basados en clases u objetos (*object based*) (Dragut and Blaschke, 2006; van Asselen and Seijmonsbergen, 2006; Dragut and Blaschke, 2009; Ghosh et al., 2010). Para la geomorfometría, lo más importante es la exactitud de los parámetros y clases determinadas, y no la precisión absoluta de elevaciones y otras medidas (Reuter, et al., 2009).

Jasiewicz et al. (2013), considera que los diferentes enfoques para la clasificación de las formas de paisaje dependen de la eficiencia con que utilizan la información contenida en las variables geomorfométricas, y en la selección de las clases o unidades de clasificación. Estos a su vez pueden agregarse según el tipo de clasificador a utilizar, previendo entonces los que usan un clasificador manual basado en la experticia del experto y/o la evidencia empírica (Pennock et al., 1987; Skidmore, 1990; MacMillan et al., 2000; Gallant et al., 2005; Iwahashi and Pike, 2007; Minar and Evans, 2008), y los que usan un clasificador automático generado a partir de un algoritmo de aprendizaje automático (Hengl and Rossiter, 2003; Ghosh et al., 2010).

Los algoritmos de aprendizaje vienen del mundo computacional y están diseñados para replicar el enfoque humano de inferir información a partir de un conjunto de datos y luego aplicar esa

información predictivamente. Valentine & Kalnins (2016), plantean el potencial del aprendizaje automático y sus aplicaciones en la geociencias, argumentando sobre los factores podrían determinar la conveniencia de algún algoritmo de aprendizaje en un problema geomorfológico.

La división final de estos métodos viene dada por la selección de las clases/objetos o unidades de clasificación (Minar & Evans, 2008), las cuales incluyen: formas de paisaje (*landforms*), elementos de formas de paisaje (*landform elements*) y unidades fisiográficas (Iwahashi & Pike, 2007; Stepinski & Bagaria, 2009; Dragut & Eisank, 2011).

En este trabajo se evalúan dos algoritmos de clasificación geomorfolométrica, uno basado en umbrales estadísticos que segmenta manualmente parámetros para clasificar elementos de formas de paisaje y el segundo basado en el algoritmo KNN de clasificación supervisada el cual clasifica igualmente elementos de forma de paisaje pero automáticamente. En ambos casos el método de clasificación es basado en celdas. La tabla 1, describe la ordenación de algunos trabajos de clasificación geomorfolométrica de acuerdo a los métodos de clasificación, clasificador y unidades de clasificación utilizada.

**Tabla 1.** Métodos, clasificador y unidades de clasificación utilizadas por algunos autores en investigaciones relativas a clasificación geomorfolométricas.

Métodos de clasificación geomorfolométrica	Tipo de clasificador	Unidades de clasificación	Autores y trabajos
Basados en celdas / <i>Cell based</i>	Manual	Formas de paisaje ( <i>landforms</i> )	MacMillan et al., 2009 Skidmore, 1990
		Elementos de formas de paisaje ( <i>landform elements</i> )	Olaya, 2009 MacMillan et al., 2004 Pike, 1988 Dragut & Blaschke, 2009 Minar & Evans, 2008
		Unidades fisiográficas	Pennock et al., 1987 van Asselen & Seijmonsbergen, 2006
	Aprendizaje automático	Formas de paisaje ( <i>landforms</i> )	
		Elementos de formas de paisaje ( <i>landform elements</i> )	MacMillan et al., 2000
		Unidades fisiográficas	Hengl & Rossiter, 2003
Basados en clases / <i>Objet based</i>	Manual	Formas de paisaje ( <i>landforms</i> )	Gallant et al., 2005
		Elementos de formas de paisaje ( <i>landform elements</i> )	
		Unidades fisiográficas	Argyros & Argialas, 2017 Gerçek Deniz et al., 2011
	Aprendizaje automático	Formas de paisaje ( <i>landforms</i> )	Ghosh et al., 2010
		Elementos de formas de paisaje ( <i>landform elements</i> )	Dragut & Blaschke, 2006 Dragut & Eisank, 2011 Dekavalla & Argialas, 2017 Valentine et al., 2016
		Unidades fisiográficas	Jasiewicz et al., 2013 Iwahashi & Pike, 2007 Stepinski & Bagaria, 2009

### 1.1.- Definición del problema

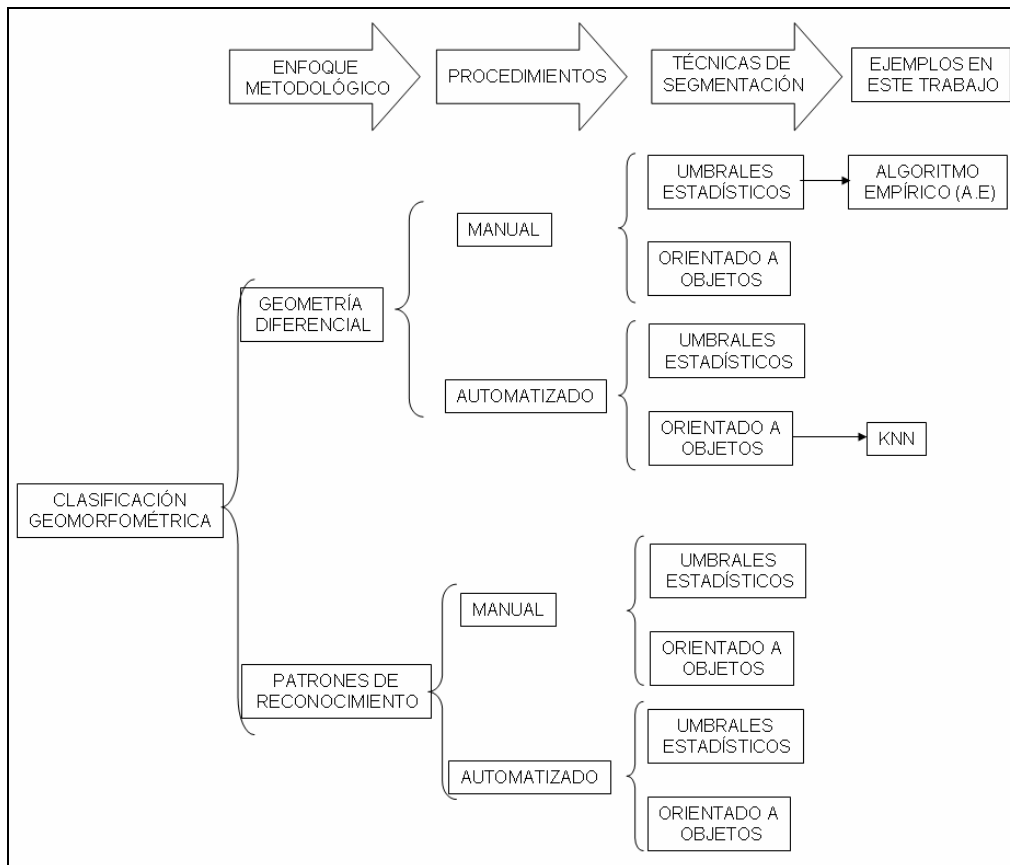
Los métodos geomorfométricos son ampliamente usados para resolver problemas multiescalares provenientes de la geomorfología, hidrología, teledetección, pedología, geología, climatología, y otras disciplinas (Clarke and Romero, 2016), es por esto que la precisión y exactitud de la clasificación geomorfométrica es una demanda importante en la comunidad científica dedicada a los estudios ambientales y de paisaje. El auge de los MDE, particularmente aquellos provenientes de sensores remotos, ha dado un reconocido impulso a las investigaciones en la morfometría del paisaje (Wood, 1996), sin embargo con estos se han revelado dos conjuntos de problemas, uno relativo a las consideraciones técnicas de los datos fuente como pueden ser la precisión y resolución horizontal / vertical de los MDE, y el otro relativo al alcance del análisis, lo cual involucra la heterogeneidad de enfoques en la definición de las clases geomorfométricas y los métodos y técnicas empleados.

La calidad del análisis geomorfométrico está determinada por la calidad de los MDE utilizados. Ni siquiera el más sofisticado algoritmo de clasificación geomorfométrica puede corregir las fallas y errores de origen contenidos en los MDE (Reuter, et al., 2009), sin embargo para reducir los problemas relativos a la precisión de los MDE, en este trabajo se utilizó el MDE SRTM v4 en su resolución original 90m, considerado la mejor versión mejorada de los datos SRTM (Lin., et al 2013) y disponible en la plataforma CGIAR (*Consultative Group on International Agricultural Research*), y en el caso del MDE proveniente del sistema ALOS PALSAR, se escogió la versión corregida (*Hi-Res Terrain Corrected*), y remuestreada a 12,5m tal como se ofrece en el portal de *Alaska Satellite Facility* (ASF DAAC 2019).

El enfoque metodológico utilizado corresponde a la geometría diferencial, siguiendo un procedimiento manual en lo que se refiere al algoritmo empírico basado en umbrales estadísticos, y un procedimiento automático en la clasificación con el algoritmo KNN basado en reconocimiento de objetos y patrones.

El problema relativo a la heterogeneidad de enfoques para la definición de clases esta ligado además de la resolución espacial de los MDE y escala de análisis, a la concepción de clases y formas de paisajes lo cual tiene un componente subjetivo. Es por ello que nuevos conceptos sobre clasificación geomorfométrica definen formas de paisaje universales que pueden reconocerse a diferentes resoluciones del MDE, tal es el caso de clasificaciones basadas en patrones geomorfológicos como *Geomorphons* (Jasiewicz et al., 2013), la cual consiste en dividir la geomorfología del paisaje en 10 clases (*flat/llano, peak/pico, ridge/cresta, shoulder/berma, spur/estribación, slope/pendiente, hollow/ahuecado, footslope/piedemonte, valley/valle, and pit/surco*), que pueden ser extraídas a partir de un MDE y con técnicas morfométricas. Esta técnica está basada en patrones de reconocimiento en vez de geometría diferencial y por lo tanto presentan una alta eficiencia computacional (Amatulli et al., 2019). Sin embargo, dado que la mayoría de los métodos de clasificación usados para determinar formas de paisaje están basados en variables geomorfométricas (Evans, 2011; Pike, 1988; MacMillan et al., 2004; Olaya, 2009), y que estas variables, expresadas en dominios de valor (*float*), son calculadas a partir de la primera y segunda derivada de un MDE, se considera que la mayoría de los métodos de clasificación automática geomorfométrica tienen sus raíces en la geometría diferencial (Jasiewicz et al., 2013).

Por lo tanto, asumiendo que la robustez de la clasificación geomorfométrica depende del ensayo y comparación de los diferentes enfoques y resoluciones del MDE, en este trabajo se definen 13 clases geomorfométricas a clasificar y comparar a partir de métodos manuales como el Algoritmo Empírico (AE) y automatizados como el algoritmo KNN, siendo el enfoque metodológico basado en la geometría diferencial. Las 13 clases geomorfométricas provienen de la fotointerpretación previa y describen las formas de paisaje encontradas. La figura 1, señala la ubicación de los casos de análisis geomorfométricos aquí presentados y su ubicación con respecto al enfoque, procedimientos y técnicas de procesamiento.



**Figura 1.** Localización esquemática de los métodos de clasificación geomorfométrica aplicados en este trabajo según sus técnicas de segmentación, procedimientos y enfoques.

### 1.2.- La clasificación geomorfométrica. Parámetros y clases.

Si bien en la figura 1, se esquematiza la ubicación de los casos de este trabajo de acuerdo al enfoque metodológico, procedimientos y técnicas empleadas, en este apartado se describen los parámetros y clases utilizados en la ejecución de los algoritmos clasificatorios.

Mientras en las clasificaciones de coberturas naturales, cultivos y uso de la tierra, se utilizan datos extraídos de los diferentes canales espectrales que poseen las imágenes satelitales, en el caso de la clasificación geomorfométrica, los datos a procesar deben ser extraídos del MDE del área de estudio, tal como lo propuso Dikau (1989) citado por Gerçek (2011), quien categorizó la



superficie en clases morfométricas del paisaje mediante el uso de la pendiente y cuatro curvaturas direccionales (curvatura del perfil, plano de curvatura, máxima curvatura y mínima curvatura), como parámetros.

Dado que la información contenida en un MDE podría ampliarse si se combinan las técnicas de análisis morfométricos junto a las de teledetección óptica (Olaya, 2009), existen trabajos sobre clasificaciones geomorfométricas que complementan los datos extraídos del MDE con datos de sensores ópticos como la integración de las imágenes Landsat en la definición de abanicos aluviales en el Valle de la Muerte, California, mostrada por Argyros & Argialas (2017), sin embargo, aún el peso metodológico recae predominantemente en los parámetros morfométricos extraídos del MDE.

En cualquier caso, la calidad de los levantamientos geomorfométricos, sus parámetros de superficie y clases a determinar, depende de ciertos factores como (Florinsky, 1998):

- La rugosidad del relieve
- La densidad de muestreo de los datos de elevación
- La resolución de la malla raster a organizar los muestreos
- El algoritmo de rasterización del MDE
- La resolución vertical del MDE
- El tipo de análisis geomorfométrico

El análisis de los parámetros geomorfométricos se realiza preferentemente a nivel local y estos pueden dividirse en dos grupos, aquellos basados en las características geométricas y los basados en umbrales estadísticos (Olaya, 2009). En este trabajo, los parámetros geomorfométricos fueron extraídos del MDE respectivo mediante las técnicas de filtrado comunes en los SIG.

## **2.- OBJETIVOS DE LA INVESTIGACIÓN**

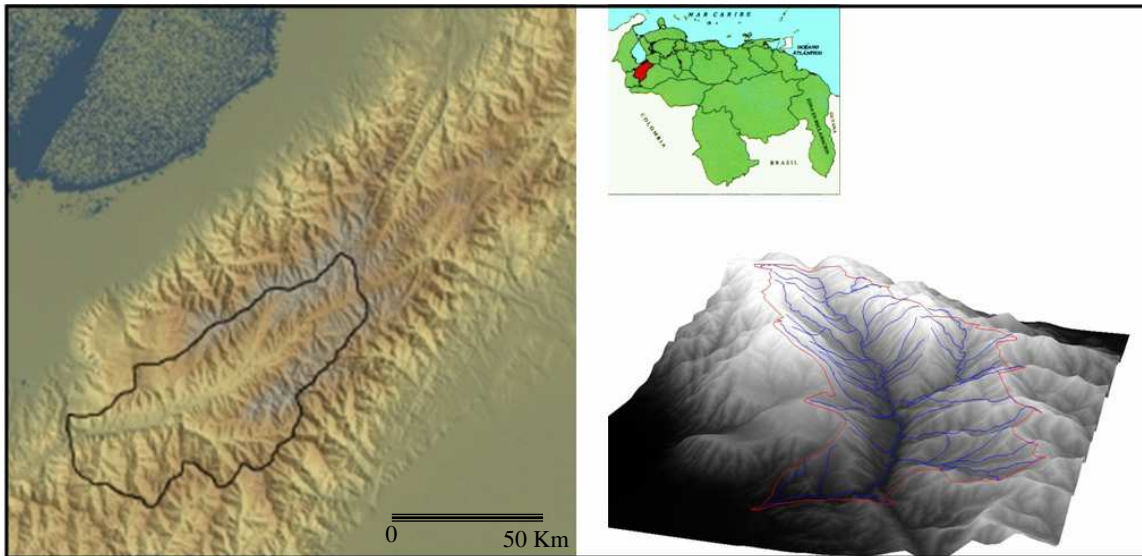
El objetivo principal a desarrollar en este trabajo es la evaluación de un modelo de clasificación geomorfométrica automática ejecutado con el algoritmo KNN y comparado con la clasificación obtenida de un algoritmo empírico manual, ambos aplicados a dos modelos digitales de elevación (MDE) ALOS PALSAR (12,5m) y SRTM (90m). Este objetivo principal es complementado por los siguientes objetivos específicos:

- Desarrollar a partir del análisis de formas un algoritmo de clasificación manual basado a partir de umbrales estadísticos y basado en geometría diferencial.
- Seleccionar el algoritmo de clasificación KNN que posea los mejores parámetros y rendimiento de ejecución.
- Clasificar 13 clases geomorfométricas en los diferentes conjuntos de datos procesar (MDE ALOS PALSAR y SRTM aplicados en zonas de páramo y vertientes).
- Comprobar las clasificaciones obtenidas a través de las medidas de exactitud y fiabilidad a partir de la matriz de confusión.
- Determinar cual MDE ofrece mejores ventajas para la extracción de los parámetros geomorfométricos y final clasificación geomorfométrica.

## **3.- MATERIALES Y METODOS**

### 3.1.- El área de estudio

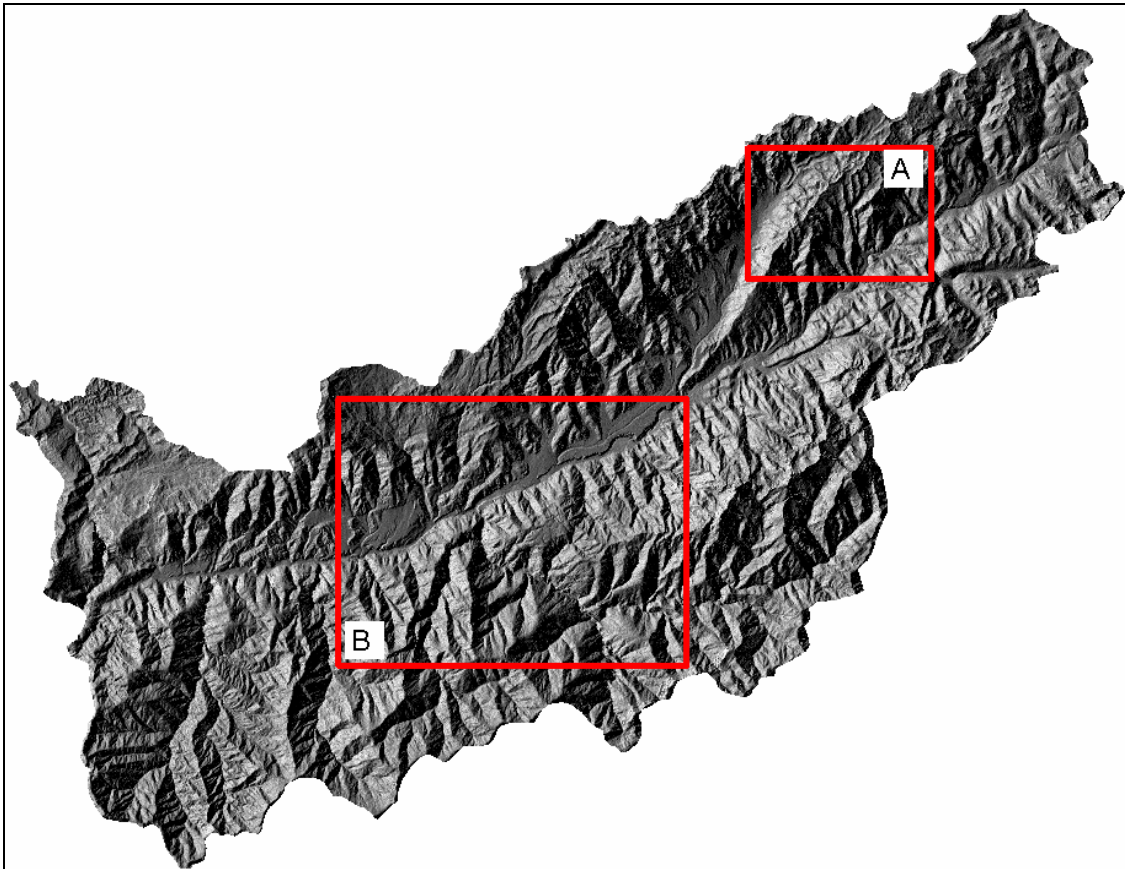
El área de trabajo corresponde a la cuenca del Río Chama, ubicada en la parte central de la Cordillera de Mérida, localizada al occidente de Venezuela (Figura 2). Si bien esta es una cuenca de montaña tropical, exhibe rangos de altitud ente los 150 y 5000 msnm, y posee un área de 3517 Km<sup>2</sup>, lo cual ofrece diversidad en cuanto a formas y procesos geomorfológicos, cuya expresión geomorfométrica es claramente visible en la superficie de la cuenca.



**Figura 2.** Área de estudio. Cuenca del Río Chama, localizada en la parte central de los andes venezolanos al occidente del país.

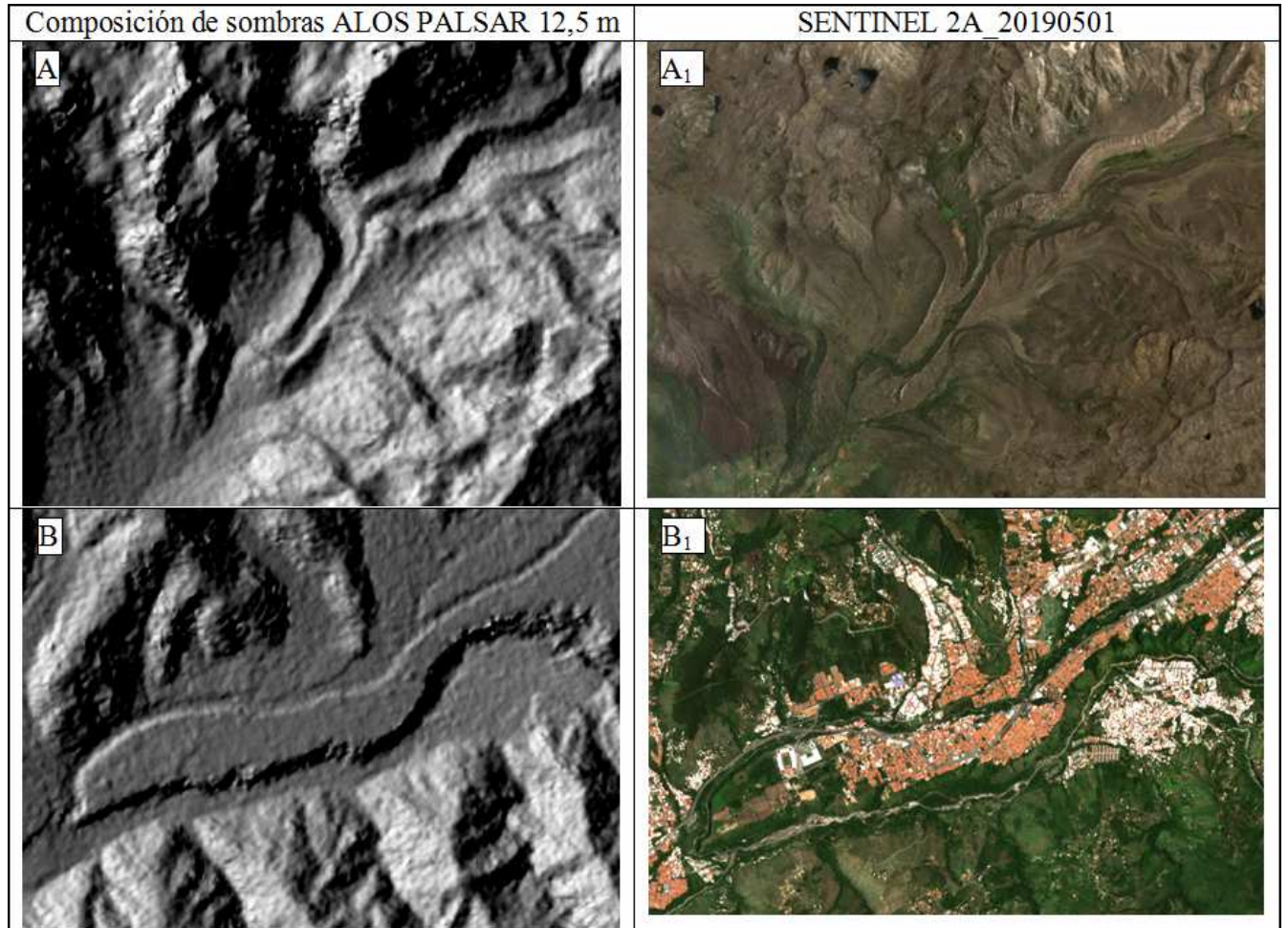
El drenaje principal de esta cuenca es el río Chama, el cual nace a una altitud de 4000 msnm y tiene un recorrido longitudinal de 180 Km. La cuenca posee desde ambientes glaciales hasta bosques secos tropicales, pasando por unidades de selva nublada lluviosa, consideradas hot spot dentro de la nomenclatura internacional como centros de alto endemismo de especies. Geomorfológicamente la cuenca está conformada por dos grandes vertientes generadas por la Sierra Nevada al SE y la Sierra de la Culata al NO, estas vertientes contentivas de formas de paisaje como afloramientos, laderas denudacionales y coluviales, confluyen en los valles de los ríos principales y secundarios donde ocurren otras expresiones geomorfométricas tales como los abanicos aluviales, las terrazas y los valles.

Se ha designado el área de estudio a toda la cuenca del río Chama pues permite asociar las diferentes morfometrías a los diferentes ambientes geomorfológicos, altitudinales y topográficos que la cuenca posee, sin embargo por el coste / tiempo computacional el análisis y clasificación pretendido en este trabajo se construyó a dos bloques dentro de la cuenca, uno representando la zona de alta montaña (A) y el segundo la zona de de vertientes y valles (B) (Figura 3). Para efectos de este trabajo la zona de alta montaña se considerará como zona de páramo y se refiere a un polígono donde por factores altitudinales y climáticos presenta una evolución geomorfológica diferente a la zona de vertientes.

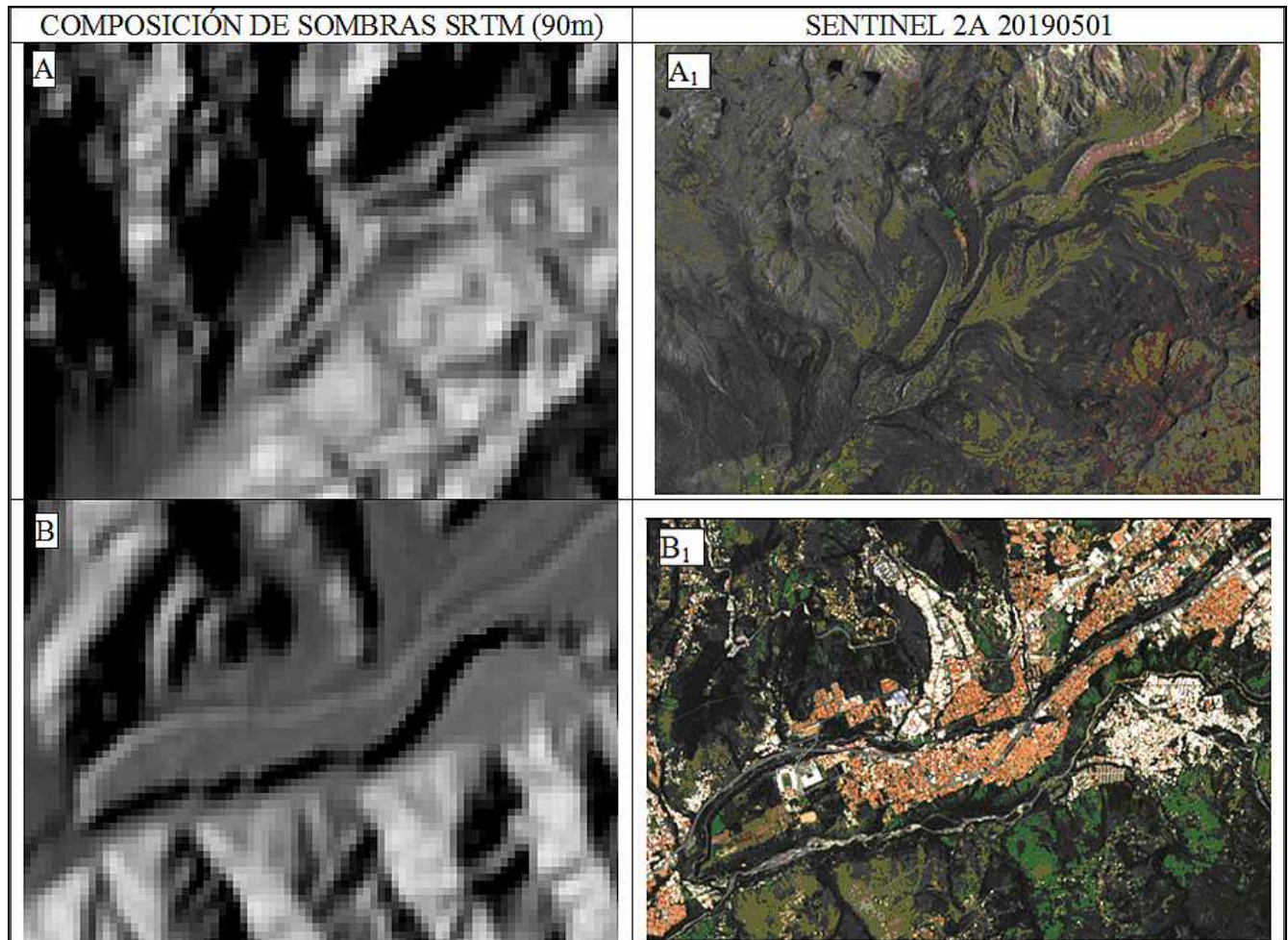


**Figura 3.** Localización de los bloques de análisis. El bloque A corresponde al polígono localizado en la zona de páramo. El bloque B es el polígono que contiene las zonas de vertientes y valle. La capa de fondo es una composición de sombras derivado del MDE ALOS PALSAR 12,5m.

La zona de páramo presenta características geomorfológicas de ambientes glaciales y periglaciales, mientras que la zona de vertientes tiene como principal característica los ambientes de acumulación representados morfométricamente por terrazas y abanicos aluviales. Elementos morfométricos del paisaje tales como morrenas, valles, terrazas, abanicos, etc, pueden ser apreciables en una imagen Sentinel 10m de resolución para un ojo entrenado, pero también pueden ser deducibles en los MDE ALOS PALSAR (12,5m) y SRTM (90m) como propuesto en este trabajo. Las figuras 4 y 5 ilustran la expresión espacial de algunas formas del paisaje en el área de estudio y su correspondiente bosquejo en imágenes ópticas y MDE.



**Figura 4.** Visualización comparativa de características morfométricas advertidas tanto en una composición de sombras basada en un MDE ALOS PALSAR (12,5m) y su correspondiente en una composición color natural de SENTINEL 2A (10m). La toma A muestra una sección de la zona de páramo donde se aprecian formas asociadas a morrenas, crestas, valles glaciales y afloramientos rocosos, estas características también son apreciables en la toma A<sub>1</sub> correspondiente a una composición color natural de SENTINEL 2A. Para el caso de la zona de vertientes B, el sombreado del MDE reporta un amplio valle de acumulación donde resalta un sistema de terrazas flanqueado por taludes, tales características se confirman en la toma B<sub>1</sub>, la cual reporta el uso urbano del valle a diferencia de las vertientes circundantes.



**Figura 5.** Visualización comparativa de características morfológicas advertidas en una composición de sombras basada en un MDE SRTM (90m) y sus correspondientes en una composición color natural de SENTINEL 2A (10m). En este caso la toma A muestra la zona de páramo con una resolución menor a la ofrecida por la composición de ALOS PALSAR pero aún es posible apreciar formas asociadas a morrenas, crestas, valles glaciales y afloramientos rocosos, las cuales se ilustran en la toma A<sub>1</sub> correspondiente a una composición color natural de SENTINEL 2A. La zona de vertientes B, el sombreado del MDE reporta igualmente el valle donde resalta las terrazas flanqueadas por taludes, características confirmadas luego en la toma B<sub>1</sub>.

### 3.2.- Diseño de investigación

#### 3.2.1.- Datos fuente. Los MDE Alos Palsar y SRTM.

Los datos fuente se refieren a los datos de elevación de los dos polígonos de estudio (páramo y vertientes) de la cuenca analizada, estos se obtuvieron de MDE elaborados a partir de datos ALOS PALSAR y de datos SRTM v4. El MDE ALOS PALSAR escogido para este análisis es el correspondiente al de alta resolución (12,5m horizontal), corregido y disponible en el portal de Alaska Satellite Facility (ASF DAAC, 2019), en el caso del SRTM se dispuso del ofrecido por la plataforma CGIAR (<http://srtm.csi.cgiar.org/>), y considerado la mejor versión mejorada (v4.1) de los datos SRTM (Lin., et al 2013), en su resolución original de 90m. Ambos MDE fueron revisados tanto en la distribución de sus datos y corrección de errores relativos a vacíos y datos fuera de rango, sin embargo no otra forma de normalización y rectificación fue aplicada.

3.2.2.- Datos insumos. Parámetros geomorfométricos. Algoritmos de clasificación.

Los datos insumos provienen del primer procesamiento de los datos fuente y representan el conjunto de datos que alimentarán los algoritmos de clasificación previstos en este trabajo. Cada MDE fue filtrado y procesado con la herramienta de modelizaje topográfico de ENVI la cual está basada en el metodo Evans-Young (Evans, 1980), para extraer los 6 parámetros geomorfométricos: Pendiente angular (Slope angle), Relieve Interno (Internal Relief), Convexidad Longitudinal (Longitudinal Convexity), Plano de Curvatura (Plan Curvature), Perfil de Convexidad (Profile Convexity), Convexidad Cruzada (Cross Convexity), a utilizar en este análisis. Estos parámetros son el resultado de la aplicación de un filtro direccional con una ventana kernel 3 x 3, donde cada MDE genera su propio conjunto de parámetros.

Los parámetros geomorfométricos son la base tanto para lograr la clasificación geomorfométrica automatizada mediante el algoritmo KNN como tambien la clasificación manual mediante el algoritmo empírico AE. Aunque existe diversidad de parámetros morfométricos que pueden ser extraídos de un MDE, en este trabajo se utilizarán sólo los 6 antes mencionados debido a:

- La integración de estos describe las condiciones geométricas locales.
- Su expresión espacial fácilmente se asocia con la existencia de formas de terreno detectadas en campo o con la fotointerpretación.
- Se utilizaron inicialmente para establecer los umbrales límite para la elaboración del algoritmo empírico AE.

Adicionalmente, el uso limitado de parámetros supone un menor gasto de tiempo computacional, permitiendo mayor frecuencia en la ejecución y calibración de los procedimientos. En el tabla 2, se describe el alcance descriptivo y espacial de algunos parámetros geomorfométricos generales usados en estudios regionales hidrográficos y los parámetros usados en este trabajo con fines de clasificación geomorfométrica.

**Tabla 2.** Descripción de algunos parámetros geomorfométricos según su alcance espacial y descriptivo.

ALGUNOS PARÁMETROS GEOMORFOMÉTRICOS GENERALES	ALCANCE ESPACIAL	ALCANCE DESCRIPTIVO
Orientación del relieve	Local	Dirección de flujos
Área de la cuenca	Regional	Magnitud de flujo
Hipsometría	Regional	Distribución del área en rangos altitudinales
Insolación	Regional / local	Intensidad radiación solar
Índice topográfico compuesto	Regional / local	Acumulación de flujos y erosividad
PARÁMETROS GEOMORFOMÉTRICOS USADOS EN ESTE TRABAJO	ALCANCE ESPACIAL	ALCANCE DESCRIPTIVO
Pendiente angular (Slope angle)	Local	Tasa de flujo / erosividad
Relieve Interno (Internal Relief)	Local	Complejidad del relieve
Convexidad Longitudinal (Longitudinal Convexity)	Local	1er mecanismo de acumulación. Forma de la vertiente paralela al flujo
Plano de Curvatura (Plan Curvature)	Local	Forma de la vertiente perpendicular al flujo
Perfil de Convexidad (Profile Convexity)	Local	2do mecanismo de acumulación
Convexidad Cruzada (Cross Convexity)	Local	Forma compuesta de la vertiente

## 3.2.3.- Datos producto. Las 13 clases geomorfométricas.

Los datos producto se refiere a las 2 clasificaciones geomorfométricas (KNN y AE) a obtener las cuales se desagregan en 4 por generarse a partir de dos conjuntos MDE (ALOS PALSAR, SRTM) y a su vez en 8 por estar aplicadas en dos zonas diferentes de la cuenca (páramo y vertiente). Estas clasificaciones vienen expresadas en un rango de 13 clases geomorfométricas, las cuales provienen de las clases fotointerpretadas previamente y que se aplican a la zona de estudio planteada, es por ello que estas surgen de un proceso semiótico entre las formas de paisajes y la percepción del investigador quien finalmente asigna el objeto, de aquí que las clasificaciones geomorfométricas son evaluadas por su capacidad de localizar y asignar los objetos previamente reconocidos por el investigador.

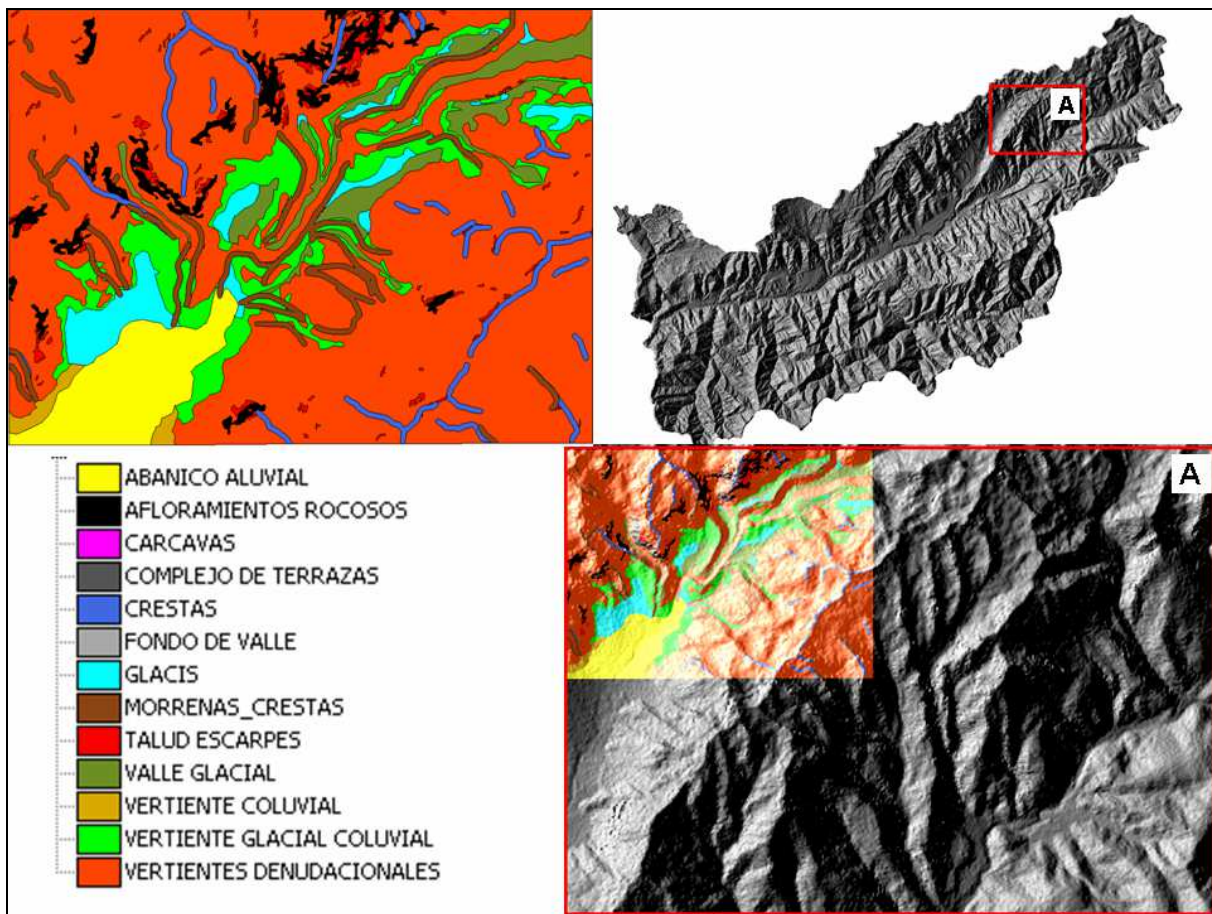
La fotointerpretación aquí presentada y elaborada por el autor, proviene de un ejercicio manual de delineación y clasificación de áreas de acuerdo a la morfometría del paisaje, basándose en fotografías aéreas e imágenes sentinel, teniendo como principal objetivo la clasificación de formas que reflejarán el paisaje predominante tanto de ambientes de alta montaña (páramo) como vertientes y valles intramontanos (vertiente). Esta fotointerpretación genera clases cuya dimensión espacial es en lo posible, asimilable a los procesos de digitalización y remuestreo a que debe someterse los efectos de este trabajo. No todas las clases estan presentes en cada zona de aplicación, por ejemplo las clases relativas a formas y procesos periglaciales predominantes en la zona de páramo no son comunes en la zona de vertiente. La tabla 3 ofrece una descripción de las 13 clases geomorfométricas a clasificar en los casos de análisis de este trabajo.

**Tabla 3.** Descripción de las 13 clases geomorfométricas a clasificar en los casos de análisis.

CLASES GEOMORFOMÉTRICAS	DESCRIPCIÓN
Abanico Aluvial	Terreno sedimentario de inclinación menor a 12° de forma cónica-triangular ubicado en transición vertiente - fondo de valle.
Afloramientos Rocosos	Rocas salientes de crestas montañosas y vertientes escarpadas.
Cárcavas	Surcos erosivos profundos de mediano y gran desarrollo longitudinal.
Complejo de Terrazas	Terrenos sedimentarios con inclinación menor a 4° delimitados por taludes y escarpes.
Crestas	Cumbreras afiladas en las divisorias de cadenas montañosas.
Fondo de Valle	Concavidad por donde discurre un drenaje y terrenos sedimentarios asociados de inclinación menor a 4°.
Glacis	Terreno sedimentario de origen periglacial de inclinación menor a 16° ubicado en transición vertiente - fondo de valle.
Morrenas - crestas	Cadenas convexas de derrubios formados a los flancos de glaciares
Taludes y Escarpes	Cambio brusco en la pendiente que genera paredes con inclinacion superiores a 70°.
Valle Glacial	Valles de alta montaña con inclinación menor a 12° acunados entre morrenas y circos glaciales.
Vertiente Coluvial	Laderas con inclinación menor a 24° formadas por acumulación de material detritico.
Vertiente Glacial Coluvial	Laderas con inclinación menor a 24° formadas por acumulación de material detritico grueso y adyacentes a rocas periglaciales.
Vertientes Denudacionales	Laderas con inclinación mayor a 24° formadas afloramientos rocosos con procesos erosivos activos.

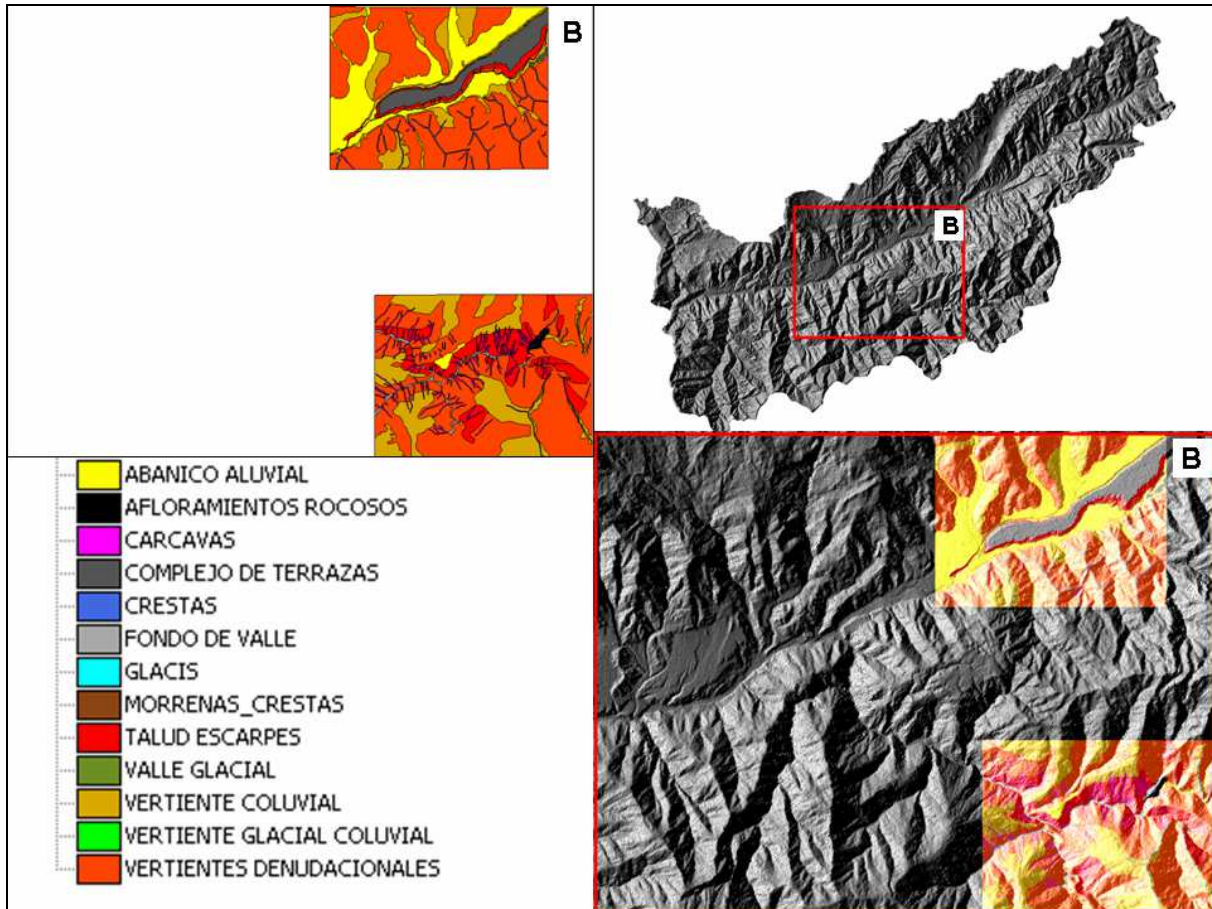
## 3.2.4.- Datos validación. Las áreas de comprobación. La matriz de confusión.

Las clasificaciones logradas con los algoritmos aquí propuestos serán validadas a través de la matriz de confusión, para esto es necesario establecer áreas de comprobación o de referencia, las cuales describen zonas cuya distribución de clases corresponden lo más preciso posible a la realidad geográfica. Las áreas de comprobación a utilizar en este trabajo provienen de la fotointerpretación previamente realizada a partir de fotografías aéreas como imagen satelital, tal como se describió en el anterior apartado, en la cual se escogieron algunos sectores de la zona páramo como en la zona de vertiente donde se delinearón las clases geomorfométricas existentes en cada zona. Las figuras 6 y 7, ilustran las áreas de comprobación a utilizar en este estudio, sus clases, su localización dentro de la zona de estudio (páramo o vertiente) y dentro de la cuenca en estudio.



**Figura 6.** Área de comprobación en un sector correspondiente a la zona de páramo (A), que describe las clases geomorfológicas previamente fotointerpretadas por el autor y que son utilizadas en este trabajo como áreas de referencia. El área restante del polígono A, será el área a clasificar según los algoritmos en este trabajo propuestos.





**Figura 7.** Área de comprobación en dos sectores correspondientes a la zona de vertientes (B), que describe las clases geomorfológicas previamente fotointerpretadas por el autor y que son utilizadas en este trabajo como áreas de referencia. El área restante no clasificada del polígono B, será el área a clasificar según los algoritmos en este trabajo propuesto.

La matriz de confusión es una técnica estadística que en el caso de la teledetección compara la fiabilidad de una clasificación espacial contra la observación de campo. Cada fila de la matriz representa el número de predicciones de cada clase, mientras que cada columna representa a las instancias en la clase real. Uno de los beneficios de la matriz de confusión es que facilita determinar si el sistema está confundiendo dos clases.

Las clasificaciones geomorfométricas obtenidas con los algoritmos propuestos (KNN y AE), son evaluadas por comparación con las áreas estipuladas como de validación en el apartado 1.3. Este análisis comparativo se realiza con la matriz de confusión y medidas de rendimiento derivadas como lo son:

- La exactitud global (*overall accuracy*), la cual es una medida porcentual global de la relación y establece un porcentaje de acuerdo entre la clasificación lograda y las áreas de validación.

- La precisión (*precision*), esta medida establece el acierto de píxeles correctamente clasificados en cada categoría por el número total de píxeles que fueron clasificados en dicha categoría.
- La confiabilidad (*reliability*), en cambio establece que porcentaje de píxeles de una clase ya definida por el área de validación que son correctamente interpretados y clasificados en la correspondiente categoría.
- El estadístico Kappa (*k*), es un índice de de concordancia propuesto por Cohen (1960), que en clasicación, mide la concordancia existente entre las frecuencias de ocurrencia de cada clase y las frecuencias obtenidas al usar el clasificador, por lo tanto su valor expresa el porcentaje de acuerdo obtenido en la clasificación una vez eliminada la parte que se debe al azar. Una interpretación del grado de acuerdo a evaluar con Kappa se describe en el tabla 4 (Landis, 1977).
- El Índice de Hellden, (IH) (Hellden, 1980), expresa la probabilidad que a su vez puede llevarse a porcentaje, de los píxeles que han sido correctamente clasificados por cada clase, tomando en cuenta los errores por omisión (1- Confiabilidad) y comisión (1- Precisión), ocurridos en la clasificación.

**Tabla 4.** Grado de acuerdo según valores del índice Kappa.

<b>Kappa</b>	<b>Grado de acuerdo</b>
< 0,00	Sin acuerdo
0,00 – 0,20	Insignificante
0,20 – 0,40	Discreto
0,41 – 0,60	Moderado
0,61 – 0,80	Substancial
0,81 – 1,00	Casi perfecto

El índice de exactitud global (EG) y el kappa (*k*), son medidas globales que interpretan el grado de acuerdo y concordancia de toda la clasificación analizada con respecto al área considerada como de validación. Las medidas de precisión ( $P_i$ ) y confiabilidad ( $C_i$ ), describen el promedio del grado de acierto que tienen las clases asignadas respecto a la validación, de estas medidas se deduce entonces el grado de error.

El índice de Hellden permite comparar las clases asignadas con respecto a las consideradas verdaderas, solo cuando se computa el índice de Hellden (IH), se considera el calculo de la precisión y confiabilidad por clases. Un resumen de las definiciones, parámetros y ecuaciones de las medidas de rendimiento derivadas de la matriz de confusión se describen en el tabla 5.

**Tabla 5.** Definición y parámetros de las medidas de rendimiento derivadas de la matriz de confusión.

DEFINICIÓN	MEDIDAS DE RENDIMIENTO / ÍNDICES	PARÁMETROS
<b>Exactitud global</b> (Overall accuracy) ( <b>EG</b> ): Porcentaje de píxeles correctamente clasificados respecto al total de píxeles verificados.	$EG = \frac{\sum_{i,j=1}^n X_{i=j}}{\sum_{i,j=1}^n X_{i,j}}$	<b>EG</b> = Exactitud global $X_{i=j}$ = Valores de la diagonal de la matriz de confusión, donde los píxeles están correctamente clasificados. $X_{i,j}$ = Valores de cada celda de la diagonal de la matriz de confusión.
<b>Precisión</b> (precision) ( <b>Pi</b> ): Probabilidad de que los píxeles clasificados como pertenecientes a una determinada clase representen realmente dicha clase en el terreno.	$P_i = \frac{X_{ii}}{X_{i+}}$	<b>Pi</b> = Precisión. $X_{ii}$ = Píxeles correctamente clasificados $X_{i+}$ = Totales de la fila de cada categoría
<b>Confiabilidad</b> (reliability) ( <b>Ci</b> ): Probabilidad de que los píxeles clasificados como pertenecientes a una determinada clase representen realmente dicha clase en la clasificación aplicada.	$C_i = \frac{X_{ii}}{X_{+i}}$	<b>Ci</b> = Confiabilidad $X_{ii}$ = Píxeles correctamente clasificados $X_{+i}$ = Totales de la columna de cada categoría
<b>Estadístico Kappa</b> ( <b>k</b> ): Grado de fiabilidad debida exclusivamente a la exactitud de la clasificación, prescindiendo de factores aleatorios.	$k = \frac{N \sum_{i=1}^n X_{ii} - \sum_{i=1}^n X_{i+} X_{+i}}{N^2 - \sum_{i=1}^n X_{i+} X_{+i}}$	<b>k</b> = Mide la diferencia entre el acuerdo observado y el esperado únicamente por azar. $N$ = Total de píxeles de verificación $X_{ii}$ = Píxeles correctamente clasificados $X_{i+} X_{+i}$ = Totales de la columna y fila de cada categoría
<b>Índice de Hellden</b> ( <b>IH</b> ): Probabilidad de que los píxeles de una determinada clase $i$ se hayan clasificado correctamente. Tiene en cuenta tanto Precisión como la Confiabilidad, y por lo tanto, los errores de omisión y comisión	$IH_i = \frac{2X_{ii}}{X_{+i} + X_{i+}}$	<b>IH</b> = Índice de Hellden $X_{ii}$ = Píxeles correctamente clasificados $X_{+i}$ = Totales de la columna de cada categoría $X_{i+}$ = Totales de la fila de cada categoría

### 3.2.5.- Procedimiento metodológico general

El procedimiento metodológico seguido en este trabajo responde a la disponibilidad y procesamiento de los datos fuente, insumos, producto y de validación. Si bien se convino que el área de estudio pertenece a una cuenca hidrográfica completa, los cálculos a efectuar con los algoritmos de clasificación son realizados en bloques representativos de la cuenca debido a la demanda de memoria y tiempo / costo computacional de estos procesos lo cual fueron limitantes técnicas. La figura 8 ofrece un flujograma de los procedimientos seguidos en este trabajo.

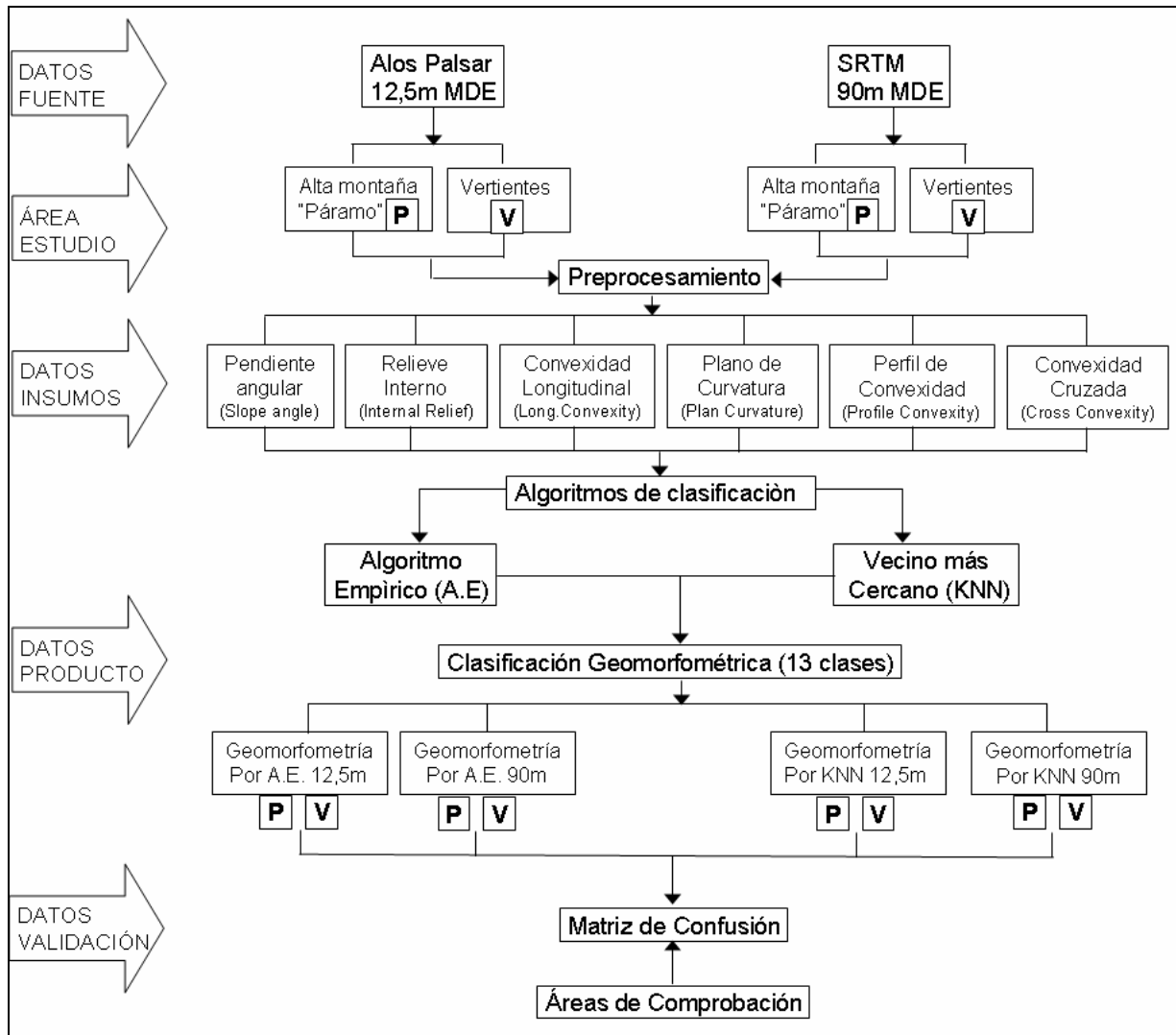


Figura 8. Flujograma del proceso metodológico seguido en este trabajo. P = páramo, V = vertiente.

### 3.3.- El algoritmo empírico (AE) de clasificación geomorfológica

El algoritmo empírico (AE) de clasificación es un protocolo que si bien puede automatizarse, para efectos de este trabajo se define como un procedimiento manual que, basado en la experiencia y las clases determinadas por la fotointerpretación, formula una serie de umbrales estadísticos a partir del histograma de los parámetros geomorfológicos. Estos umbrales son los criterios que determinan la segmentación de cada parámetro el cual luego solo o adicionado a otro criterio, definen una forma morfológica que consideraremos clase. Los criterios surgen de un proceso de optimización manual que evalúa y compara la asociación de los parámetros geomorfológicos con formas de paisaje previamente reconocidas.

#### 3.3.1.- Definición de las formas, parámetros y criterios de trabajo

El AE, es una sucesión de parámetros que son segmentados por medio de criterios y cuyo resultado se asocia a una forma de paisaje, luego estas formas son superpuestas en un orden específico y de esta manera garantizar la representación de ciertas formas menores que pudieran ser dominadas por otras de mayor distribución, por ejemplo carcavas y fondo de valle deben ir superpuesta a las vertientes coluviales o denudacionales, así como las formas glaciales y periglaciales, para asegurar su representación cartográfica. La tabla 6 describe el AE de clasificación geomorfométrica utilizado en este trabajo.

**Tabla 6.** Formas del paisaje, correspondientes parámetros representativos y criterios definitorios de la clases geomorfométricas según el algoritmo empírico de clasificación (AE).

	FORMAS	PARÁMETROS	CRITERIOS	ORDEN DE LAS CAPAS
Glacial ( $\geq 3000$ msnm)	AFLORAMIENTOS ROCOSOS	- Slope angle - Internal relief	Slope $x \geq 40$ AND Inrelief $x \geq \text{mean} + \text{STD}$	CAPA SUPERIOR  ↓  CAPA INFERIOR
	MORRENAS-CRESTAS	- Slope shape	Plan Curvature $x \leq -1(\text{mean} + \text{STD})$	
	VALLE GLACIAL	- Slope angle - Slope shape	Slope $x \leq 12$ AND Long Convex $x \leq \text{mean} + \text{STD}$	
	GLACIS	- Slope angle	Slope $12 \leq x \leq 16$	
	VERTIENTE GLACIAL COLUVIAL	- Slope angle	Slope $16 < x < 24$	
Valle y vertientes ( $< 3000$ msnm)	FONDO DE VALLE	- Slope angle - Slope shape	Slope $x \leq 12$ AND Profile Convexity $x \leq \text{mean} - \text{STD}$	
	CARCAVAS	- Slope shape	Cross Convexity $x \leq \text{mean} - 3\text{STD}$ OR Plan Convexity $x \geq \text{mean} + 3\text{STD}$ OR Profile Convexity $x \leq \text{mean} - 3\text{STD}$	
	TALUDES / ESCARPES	- Slope shape - Internal relief	Profile Convex $x \geq \text{mean} + 2\text{STD}$ OR Long Convex $x \geq \text{mean} + 2\text{STD}$ OR Internal relief $x \geq \text{mean} + 2\text{STD}$	
	COMPLEJO DE TERRAZAS	- Slope angle	Slope $x \leq 4$	
	ABANICO ALUVIAL	- Slope angle	Slope $4 < x \leq 12$	
	CRESTAS	- Slope shape	(Long Convexity $x \geq \text{mean} + \text{STD}$ AND Plan Convexity $x \leq \text{mean} - \text{STD}$ ) OR Profile Convexity $x \geq \text{mean} + 2\text{STD}$ OR Cross Convexity $x \geq \text{mean} + \text{STD}$	
	VERTIENTE COLUVIAL	- Slope angle	Slope $12 < x < 24$	
VERTIENTES DENUDACIONALES	- Slope angle	Slope $x \geq 24$		

### 3.3.2.- Esquema de clasificación

El algoritmo AE trabaja con los datos extraídos del MDE propuesto, en este análisis se utilizarán tanto el ALOS PALSAR (12,5m) y el SRTM (90m). El proceso de clasificación con el AE presenta dos bloques, el primero plantea la extracción de parámetros requeridos, que en este caso son 6. Estos parámetros son considerados parámetros geomorfométricos en cuanto se establece su asociación a determinadas formas del paisaje. Luego se procesan estadísticamente para determinar los umbrales, los cuales nuevamente en asociación con las formas de paisaje determinan un criterio que será utilizado para la asignación de la clase. La asignación de clases es el primer paso del segundo bloque o de la clasificación, una vez determinadas las clases según los criterios utilizados son superpuestas y luego filtradas (kernel 3x3), para eliminar ruido y darle mayor continuidad a las clases y de esta manera lograr la clasificación geomorfométrica final. La figura 9, describe este proceso de clasificación.

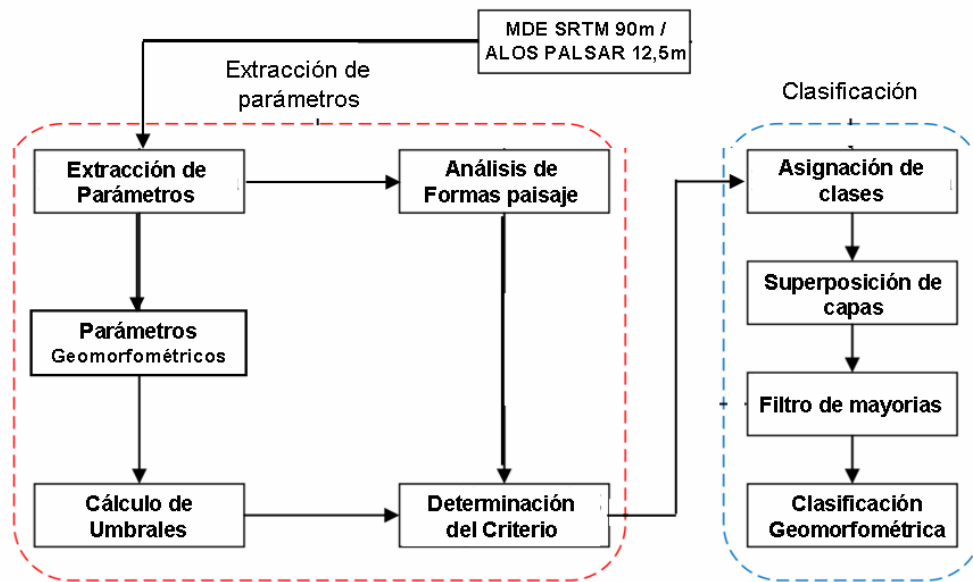


Figura 9. Proceso de clasificación según el algoritmo empírico (AE).

### 3.4.- El modelo automatizado KNN de clasificación geomorfométrica

El algoritmo *K-Nearest-Neighbor* (KNN), es un sistema de instrucciones supervisado de aprendizaje automático basado en instancias, usado en la clasificación de nuevas clases (valores discretos) o para predecir valores continuos. En el caso de la clasificación de clases de un conjunto de capas o imágenes, KNN “*aprende*” las combinaciones de datos más cercanos y similares que definen una clase y luego busca en el conjunto asociaciones similares y de esta manera clasifica el conjunto restante y otro nuevo dato que se quiera incorporar. El algoritmo es supervisado porque se entrena en un conjunto de datos, capas o imágenes, que ya han sido clasificados en clases por la relación de atributos que poseen, y se basa en instancia porque repite las instrucciones de entrenamiento y no un modelo matemático.

#### 3.4.1.- Definición de las formas, parámetros y criterios de trabajo

Debido a la naturaleza comparativa de este trabajo, los parámetros y clases a ingresar en el algoritmo KNN son los mismos que se utilizaron en el AE. Los parámetros se refieren a las 6 capas parámetros extraídas tanto del MDE ALOS PALSAR como del SRTM (pendiente angular, relieve interno, convexidad longitudinal, plano de curvatura, perfil de convexidad, convexidad cruzada), mientras que las clases están definidas en las áreas de entrenamiento que utiliza el algoritmo para “*aprender*” la instancia y posteriormente usarla para clasificar los datos restantes. Estas clases se describen en la tabla 3 del apartado 3.2.3 y su distribución espacial se ilustra en las figuras 6 y 7 del apartado 3.2.4.

El algoritmo KNN aquí implementado fue ejecutado en código python en el *notebook Jupyter* de la suite Anaconda (<https://www.anaconda.com/>). Las ventajas de esta plataforma es la

importación de librerías, modelos y procesadores como la *scikit-learn*, el *neighbors.KneighborsClassifier* y el *NumPy* respectivamente, los cuales facilitan la funcionalidad del clasificador y su automatización en lo referente a la evaluación e instrumentación de sus parámetros internos, la selección de áreas de entrenamiento y test independientes, la validación de los resultados, y finalmente el procesador *NumPy* el cual permite el manejo y producción de datos matriciales como lo son los parámetros y la clasificación geomorfométrica final.

En la utilización del *notebook Jupyter* se describe las partes del código aplicado (anexo 1), en este caso las facilidades y requerimientos del script son:

- Importación de las librerías y procesadores que contienen los protocolos a aplicar como el KNN y el manejo matricial de la información (*NumPy*).
- Lectura del conjunto de datos, se ingresan al sistema los datos de trabajo, en este caso corresponden a las 6 capas de los parámetros geomorfométricos.
- Lectura de las áreas de entrenamiento, conviene ingresar en acto aparte las áreas de entrenamiento ya que por lo general las áreas de entrenamiento vienen en datos enteros (integer) mientras que los datos a procesar pueden ser como en este caso expresados en reales (float).
- División del conjunto de datos entre los datos de ingreso y los datos objetivo
- División del conjunto de datos de ingreso entre datos de entrenamiento y datos de validación. Es de señalar acá que la librería *scikit-learn* posee una función (*train\_test\_split*), que permite separar de las áreas de entrenamiento las que propiamente entrenarán el algoritmo y otras que servirán de validación.
- Construcción del modelo KNN bajo diferentes parámetros y configuraciones. Se define el grid de parámetros a utilizar, estos nos darán todas las posibles configuraciones del clasificador KNN puesto que cada combinación de parámetros es una configuración diferente. En este caso se optó por definir una selección de vecinos (*n\_neighbors*) de 1, 3, 5, 9, con pesos (*weights*) de tipo uniforme (*uniform*) y distancia métrica (*minkowski*) y un *p* de 1, 2, 1.5, 3.
- Medida de rendimiento para la selección de los mejores parámetros. Esta grid de parámetros fueron evaluados y seleccionado el de mejor rendimiento que en nuestro caso correspondió al de vecinos 5 y peso *uniform*, distancia métrica *minkowski* y *p*:2.
- Ejecución del modelo
- Clasificación del área total (datos de ingreso y datos objetivo)

#### 3.4.2.- Esquema de clasificación

Al igual que en el anterior algoritmo, el KNN trabaja con los datos extraídos del MDE propuesto, el cual puede ser tanto el ALOS PALSAR (12,5m) como el SRTM (90m). El KNN presenta dos bloques de procesamiento definidos, el primero correspondiente al entrenamiento y el segundo a la clasificación. En el bloque de entrenamiento se ingresan los datos y de estos se extraen y definen sus características, así mismo se analizan los patrones de distribución de estos para asignarle las clases correspondientes. Estos datos ya clasificados como clases se asocian con las características de que se extrajeron del conjunto de datos, este proceso genera la instancia, parte esencial del aprendizaje automático. Paralelo a este proceso, en el bloque de clasificación se

extraen los datos que luego serán utilizados para la validación y se procede a extraer las características del conjunto total de datos los cuales posteriormente han de ser clasificados por la instancia de clasificación (aprendizaje automático) que viene del bloque de entrenamiento y validada con los datos antes extraídos. Esta etapa genera la clasificación geomorfométrica final. La figura 10 describe este proceso de clasificación.

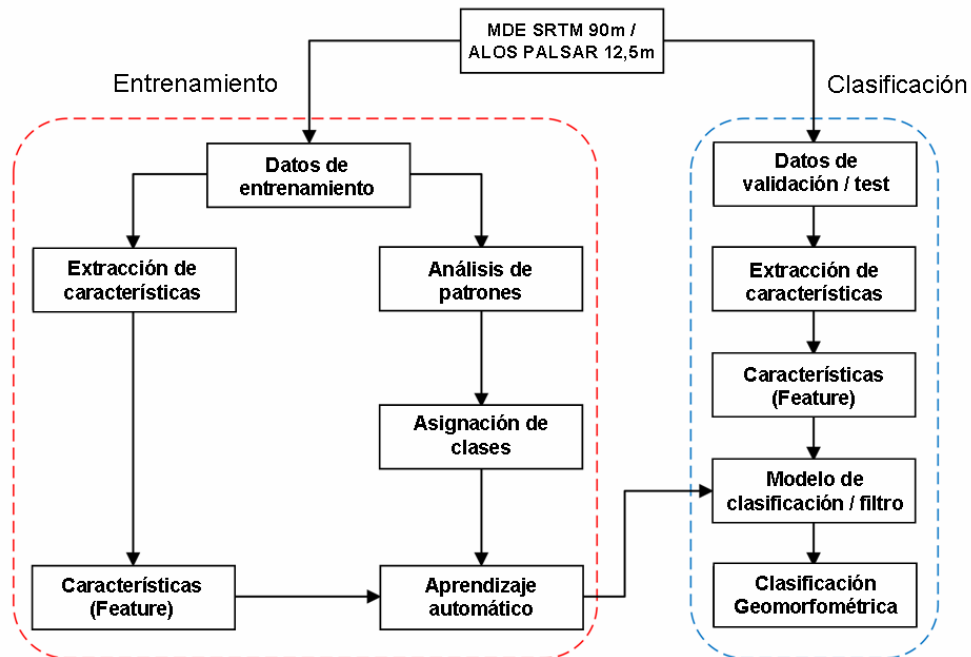


Figura 10. Proceso de clasificación según el algoritmo KNN.

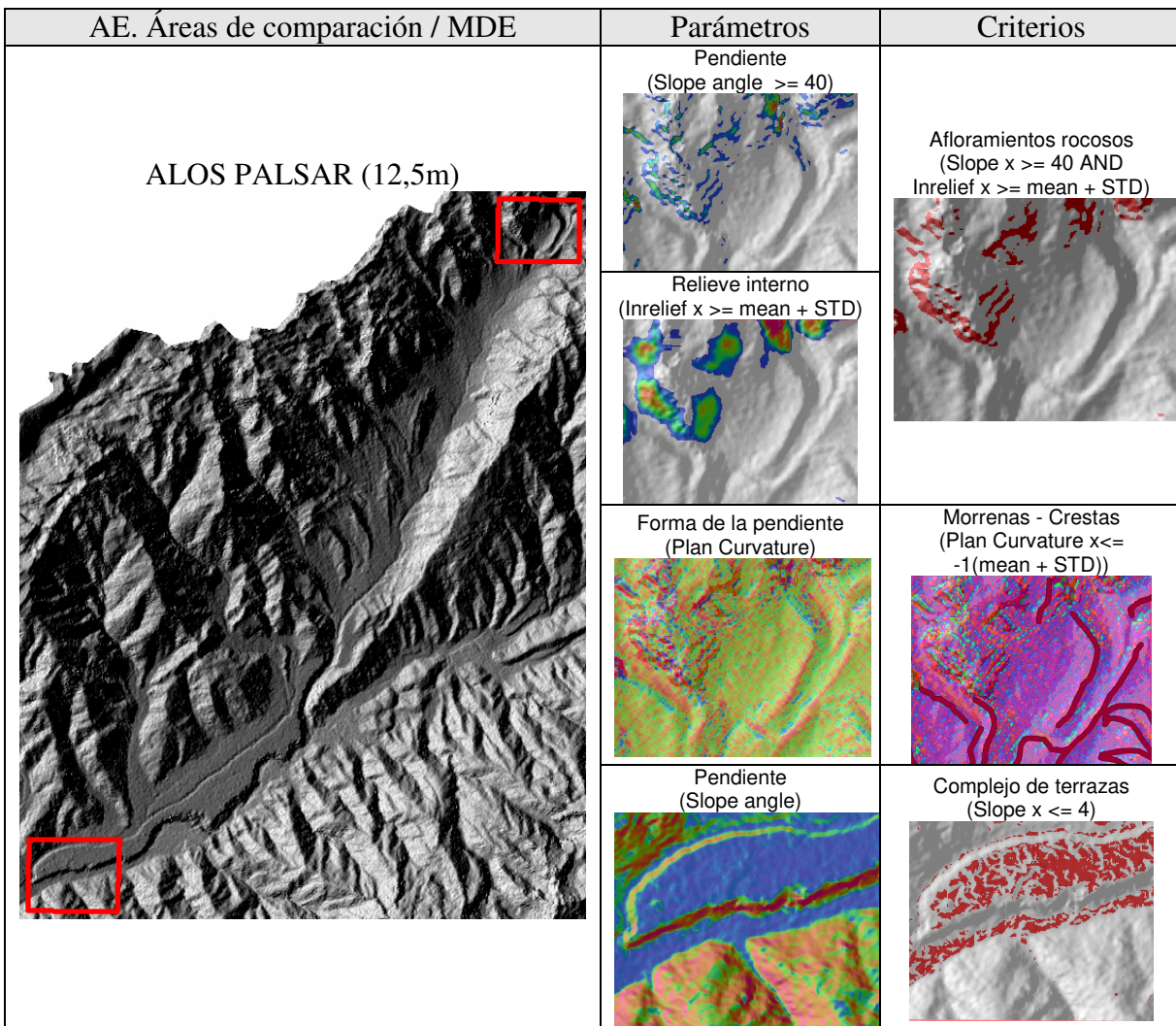
## 4.- RESULTADOS Y DISCUSIÓN

### 4.1.- Expresión espacial de los criterios y formas bajo la clasificación AE.

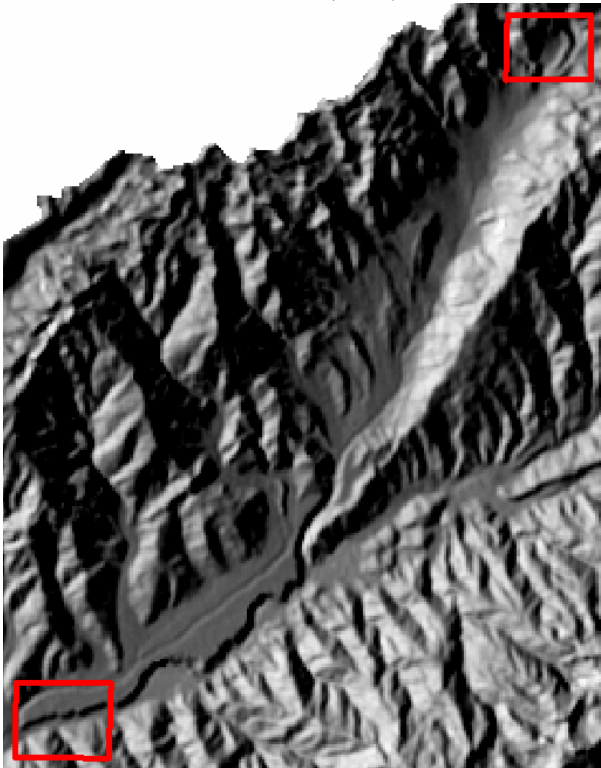
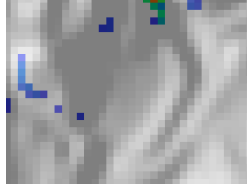
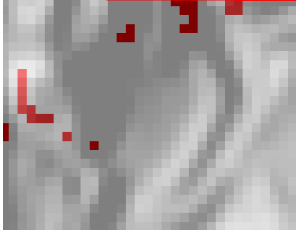
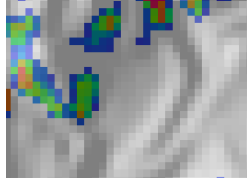
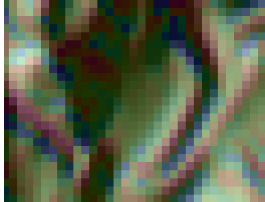

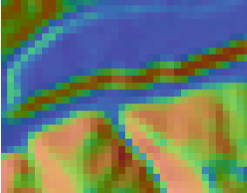
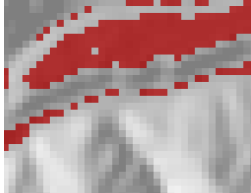
En el AE, Los parámetros geomorficos son una suerte de segmentación del MDE los cuales son luego nuevamente segmentados mediante los criterios de determinación de formas. Tal como se había advertido en el apartado 1.1, las formas o clases a obtener dependen en su expresión espacial de la resolución del MDE empleado. Cuando los parámetros proceden del MDE ALOS PALSAR (12,5m), la información que se aporta tiene mayor continuidad espacial del parámetro por lo que la forma o clase puede ser mejor discriminada, pero también se incluye un mayor ruido o dispersión. En el caso de los parámetros provenientes del SRTM (90m), estos presentan mayor generalización visual pero en algunos casos facilitan la continuidad espacial de la clase determinada.



Tal es el caso de la clase afloramientos rocosos para un sector del páramo a partir del MDE ALOS PALSAR, tanto en los parámetros como en el criterio definitorio se observa una distribución de los afloramientos más acorde al relieve dado por la composición de sombras del MDE (Figura 11), mientras que en el elaborado a partir del MDE SRTM, tal clase se construye a la limitada información que ofrece el parámetro pendientes (Figura 12). Sin embargo en el caso de la clasificación del complejo de terrazas, los datos del ALOS PALSAR generan una clase dispersa mientras que en los datos basados en el SRTM se muestra continua.

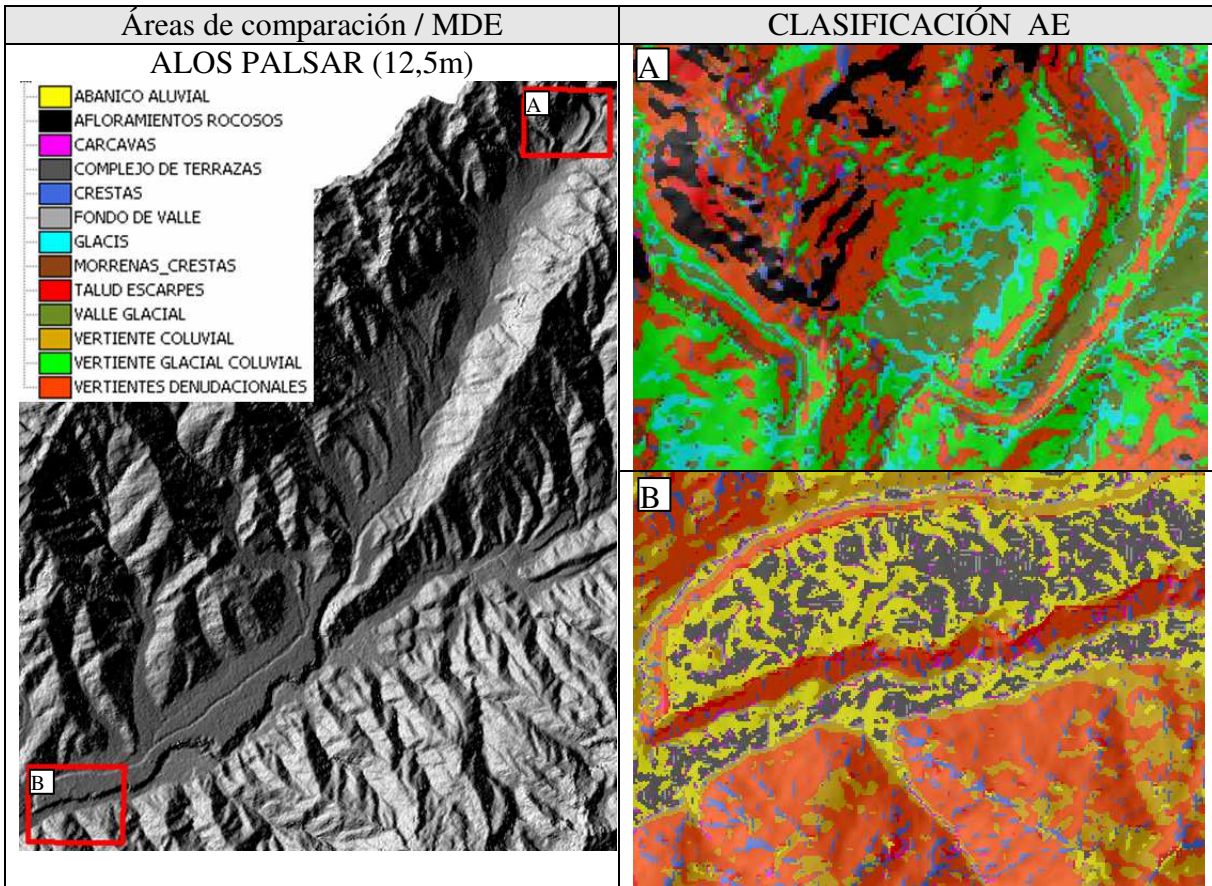


**Figura 11.** Ejemplos comparativos de los parámetros y criterios aplicados según el AE para la determinación de clases geomorfométricas, en dos ambientes del área de estudio y a partir del MDE ALOS PALSAR (12,5m).

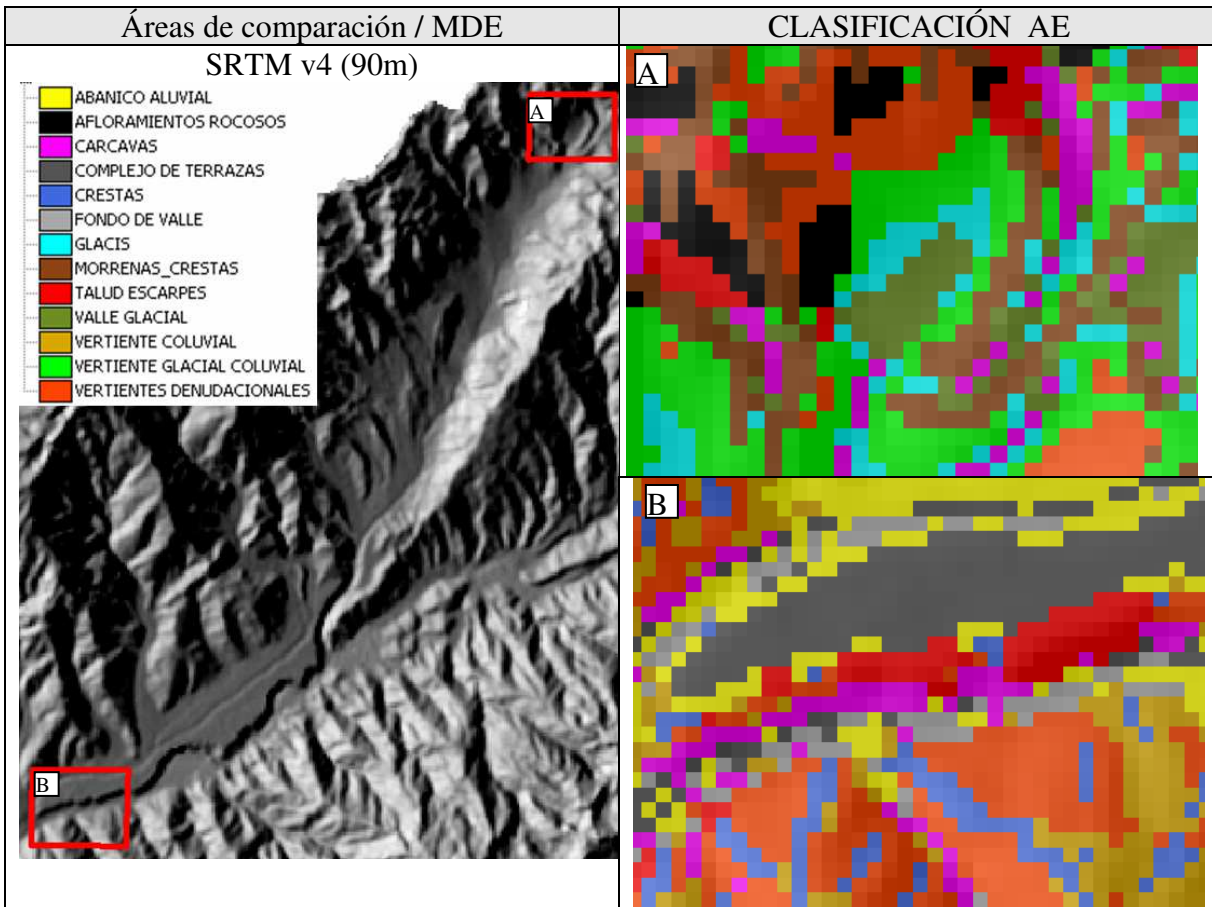
AE. Áreas de comparación / MDE	Parámetros	Criterios
<p data-bbox="402 338 618 373">SRTM v4 (90m)</p> 	<p data-bbox="862 241 1055 296">Pendiente (Slope angle <math>\geq 40</math>)</p> 	<p data-bbox="1149 327 1382 401">Afloramientos rocosos (Slope <math>x \geq 40</math> AND Inrelief <math>x \geq \text{mean} + \text{STD}</math>)</p> 
	<p data-bbox="836 478 1081 533">Relieve interno (Inrelief <math>x \geq \text{mean} + \text{STD}</math>)</p> 	
	<p data-bbox="850 720 1066 774">Forma de la pendiente (Plan Curvature)</p> 	<p data-bbox="1166 720 1356 795">Morrenas - Crestas (Plan Curvature <math>x \leq -1(\text{mean} + \text{STD})</math>)</p> 
	<p data-bbox="894 984 1019 1039">Pendiente (Slope angle)</p> 	<p data-bbox="1154 984 1365 1039">Complejo de terrazas (Slope <math>x \leq 4</math>)</p> 

**Figura 12.** Ejemplos comparativos de los parámetros y criterios aplicados según el AE para la determinación de clases geomorfométricas, en dos ambientes del área de estudio y a partir del MDE SRTM (90m).

La definición espacial de las formas y ajuste de estas a la información topográfica base es el principal efecto de la resolución espacial del MDE sobre la clasificación geomorfométrica final. Sin embargo esta consecuencia no anula la presencia sistemática de las clases geomorfológicas en cada clasificación, por ejemplo en la clasificación aplicada a los datos del MDE ALOS PALSAR (12,5m) con el AE, se generan un número de clases, que luego están presentes también en la clasificación hecha con los datos del MDE SRTM (90m). Esta característica de preservar las clases aún en deterioro de la resolución espacial es propia del AE, una mejor ilustración se puede observar en las figuras 13 y 14, donde se ejemplifica la capacidad del AE para la discriminación de clases geomorfológicas, tanto en áreas de alta montaña como en vertientes y tanto a partir de datos ALOS PALSAR como del SRTM.



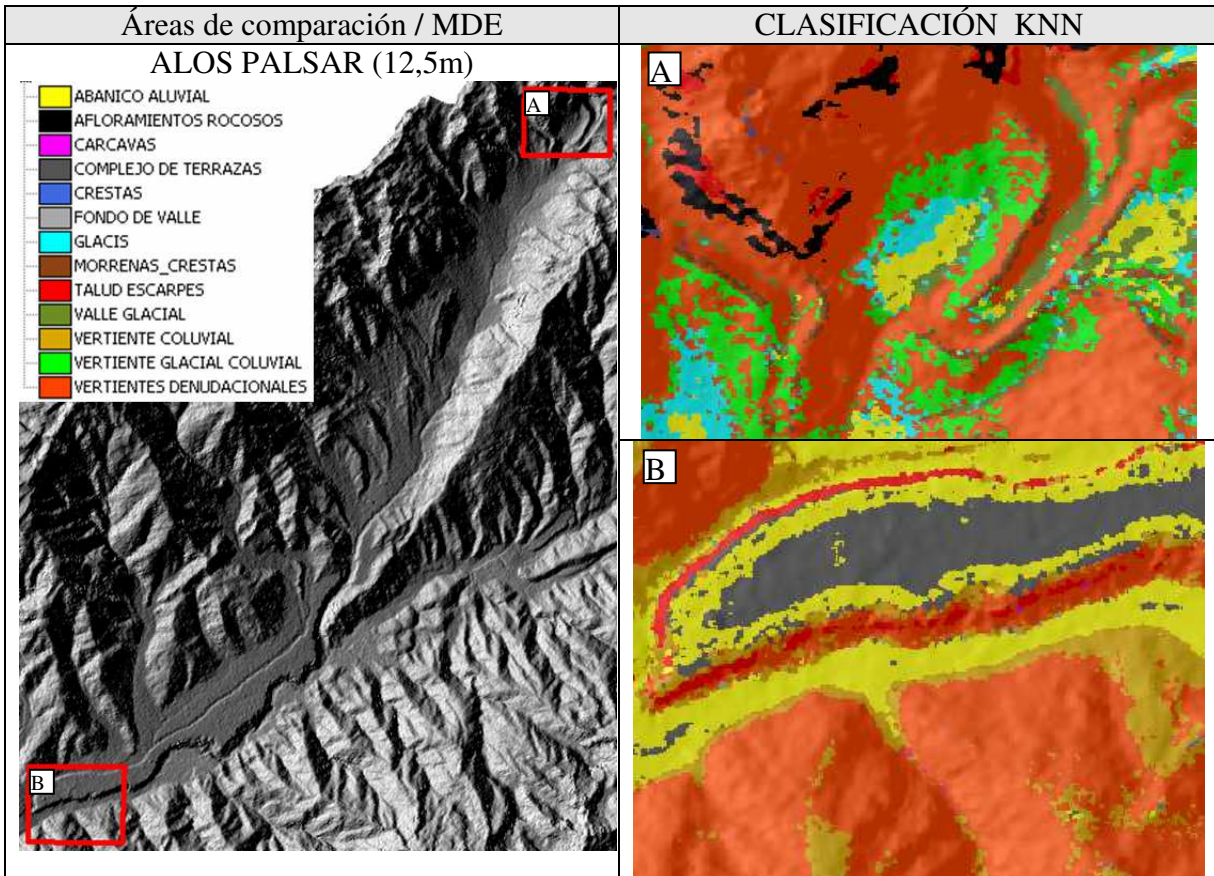
**Figura 13.** Definición espacial y discriminación de clases geomorfométricas a partir de una clasificación aplicando el AE con los parámetros geomorficos extraídos del MDE ALOS PALSAR.



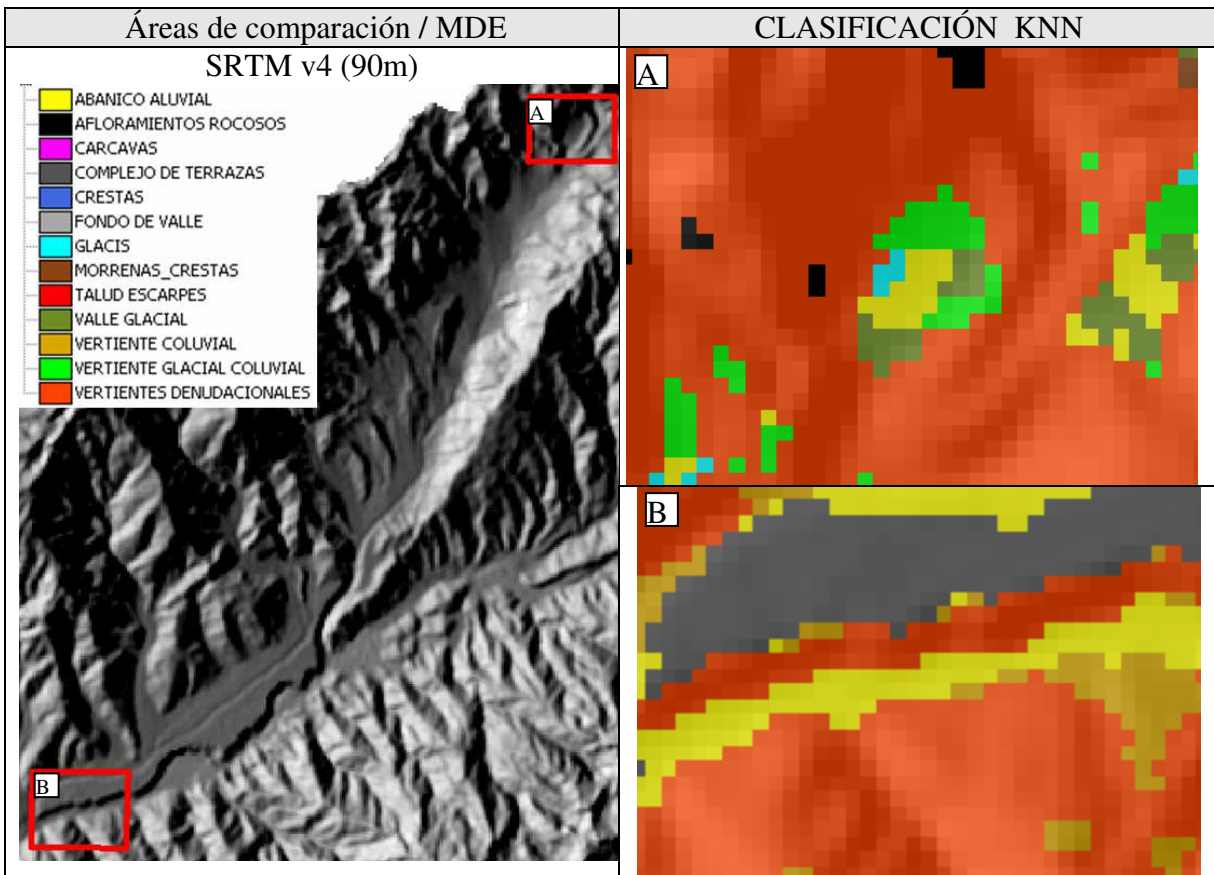
**Figura 14.** Definición espacial y discriminación de clases geomorfológicas a partir de una clasificación aplicando el AE con los parámetros geomorfológicos extraídos del MDE SRTM (90m).

#### 4.2.- Expresión espacial de los criterios y formas bajo la clasificación KNN

En el algoritmo KNN, la clasificación de las clases depende de la instancia o nivel de asociación que presenten los parámetros geomorfológicos que están catalogados como una determinada clase, debido a que esta instancia establece el aprendizaje automático. Mientras mayor sea la información extraída de los parámetros, mayor será el nivel de asociación. Es por eso que al ofrecer el MDE ALOS PALSAR (12,5m) mayor resolución espacial, las instancias a generar poseen mayor nivel de detalle y mayor margen de maniobrabilidad al momento de clasificar los nuevos datos. Caso contrario podría suceder con la clasificación realizada con los parámetros extraídos del SRTM (90m), donde las instancias están limitadas en el nivel de detalle, por lo que si bien ciertas formas de la zona de páramo como las morrenas y valles glaciales que son representadas generosamente en los datos provenientes del MDE ALOS PALSAR (Figura 15), son en cambio subrepresentadas en términos de tamaño y forma en la clasificación hecha a partir del MDE SRTM. No obstante, si evaluamos la clasificación hecha en el área de vertientes, podemos advertir que aún afectadas por el problema de generalización, las clases indicadas en la clasificación del ALOS PALSAR están presentes en la clasificación del SRTM (Figura 16).



**Figura 15.** Ejemplos comparativos de la expresión espacial de la clasificación obtenida con el algoritmo KNN en un ambiente de páramo clasificando valles glaciales, glacis, morrenas y afloramientos rocosos entre otros (A), y en un ambiente de vertiente clasificando complejo de terrazas, taludes, abanicos aluviales, vertientes coluviales y denudacionales (B), con datos obtenidos del MDE ALOS PALSAR (12,5m).



**Figura 16.** Ejemplos comparativos de la expresión espacial de la clasificación obtenida con el algoritmo KNN en un ambiente de páramo clasificando valles glaciales, glacis, morrenas y afloramientos rocosos entre otros (A), y en un ambiente de vertiente clasificando complejo de terrazas, taludes, abanicos aluviales, vertientes coluviales y denudacionales (B), con datos obtenidos del MDE SRTM (90m).

#### 4.3.- Matriz de confusión. Medidas de rendimiento

La matriz de confusión mide el grado de concordancia entre las clases geomorfológicas asignadas por cualquiera de los algoritmos clasificadores propuestos y las clases consideradas de referencia y que provienen de la fotointerpretación de algunos sectores de las zonas de análisis. Existen índices que derivados de una matriz de confusión, pueden comparar el nivel de asociación de dos clasificaciones, por ejemplo la exactitud global (EG) (Chuvieco & Huete, 2010), el coeficiente Kappa ( $k$ ) (Hudson et. al, 1987), y el índice de Hellden (IH) (Hellden,

1980). El primero es la proporción entre el total las clases correctamente clasificadas (celdas de la diagonal principal), y el total de elementos en la matriz. El coeficiente Kappa es una medida basada en la diferencia entre el porcentaje de concordancia indicado por los valores de la diagonal principal y el acuerdo aleatorio a posteriori, estimado a partir de los valores marginales (totales de las filas y columnas), y el índice de Hellden evalúa el rendimiento de la clasificación a nivel de clase.

#### 4.4.- La clasificación geomorfométrica a partir del MDE SRTM 90m.

Una particularidad del procesamiento de los datos obtenidos a través del MDE SRTM es el carácter homogéneo de las categorías extraídas, siendo esto válido tanto para los parámetros geomorfométricos como para las clases geomorfométricas. Esta condición tiene su razón en la resolución espacial del SRTM (90m), lo cual confiere a las categorías espaciales límites más toscos pero mayor uniformidad en los conjuntos clasificados. Tal condición es apreciable tanto en la clasificación realizada con AE como con KNN.

Como se viene explicando, las clasificaciones geomorfológicas obtenidas con AE y KNN se aplicaron en dos zonas definidas del área de estudio (páramo y vertiente), por lo que se consiguen 4 clasificaciones las cuales fueron comparadas con sus respectivas áreas de validación en los términos e índices ofrecidos por la matriz de confusión. A manera de ilustración la tabla 7 describe los índices, áreas de validación, y clasificación obtenida a partir de los métodos propuestos y consecuente área de aplicación, basados en los datos aportados por el MDE SRTM 90m.

##### *- La clasificación del SRTM según el AE*

Si bien el algoritmo empírico basa su protocolo de segmentación en umbrales estadísticos, su constitución inicial es heurística y se diseña con los datos que aporta este análisis es decir con los parámetros geomorfométricos y su asociación con las formas de paisaje, por esta razón se asume que una clasificación geomorfométrica hecha a partir de este algoritmo podría expresar con mayor propiedad las formas del paisaje vigentes en las zonas de análisis. Sin embargo en la comparación realizada con la matriz de confusión, esta clasificación luce insignificante en cuanto al índice EG y apenas se reconoce como discreta en el área de vertiente según el estadístico kappa.

El contraste de la segmentación que presenta la clasificación AE se acentúa con respecto al carácter poligonal simple, de clases continuas en las áreas de validación, debido precisamente los límites estáticos que imponen los umbrales estadísticos que emplea el algoritmo (anexo 2). Clases de tipo lineal como crestas, morrenas, fondo de valle y carcavas que en mucho de los casos su expresión en las áreas de validación no son más anchas que la magnitud del pixel, son ampliamente delineadas en la clasificación AE, siendo este contraste la razón por la cual su acierto se subestima en la matriz de confusión. Caso contrario ocurre con las clases abanicos aluviales y complejo de terrazas, las cuales son mejor delineadas tanto en las áreas de validación como en la clasificación AE, donde se mantiene el carácter continuo y contiguo de esta forma de paisaje (anexo 5).

- *La clasificación del SRTM según el KNN*

A diferencia del algoritmo AE, que supone un modelo estricto de umbrales estadísticos, la clasificación basada en el KNN esta basada en el aprendizaje que sucede en el mismo momento en el que se prueban los datos de test. El algoritmo clasifica cada dato nuevo en el grupo que corresponda, según tenga  $k$  vecinos más cerca de un grupo o de otro, es decir, calcula la distancia del elemento nuevo a cada uno de los existentes, y ordena dichas distancias de menor a mayor para ir seleccionando la clase a pertenecer. Esta clase será, por lo tanto, la de mayor frecuencia con menores distancias.

Precisamente las clases de mayor frecuencia y consecuentemente de menores distancias por su ubicuidad en las áreas de validación son predominantemente las clases: vertientes denudacionales y abanicos aluviales. Probablemente el KNN establece alrededor de los núcleos de estas clases los pivotes de su clasificación y de esta manera acierta un índice EG mayor al 70% y con el estadístico kappa calificado como moderado.

Una ilustración de este argumento lo encontramos en la clasificación de la zona páramo (Tabla 7), en donde la clasificación KNN con solo generalizar las clases vertientes denudacionales y abanicos aluviales, acierta por azar la distribución de estas mismas clases en las áreas de validación, aún a costa de ignorar la ocurrencia de otras clases de menor distribución (anexo 4).

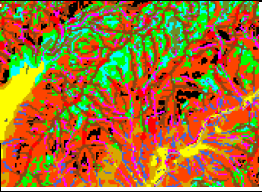
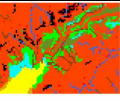
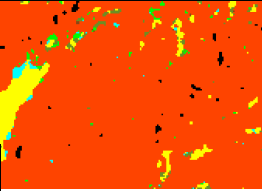
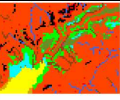
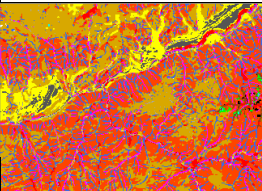
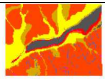
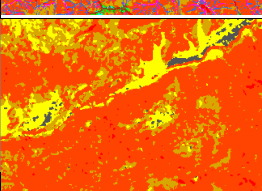
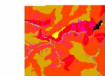
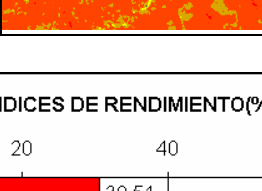
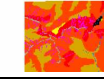
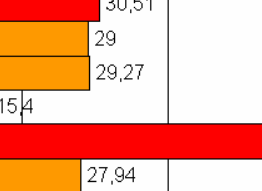
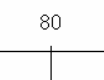
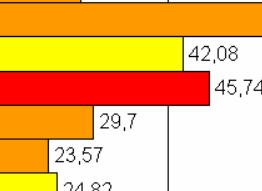
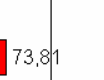
Efectivamente, el EG sitúa esta comparación en 73,81% en la zona de páramo y 75,74% en la zona de vertiente, lo cual no se refleja de la misma manera en el estadístico kappa (páramo 42,08% y vertiente 57,08%), lo que comprueba el carácter aleatorio en parte importante de la asignación de clases.

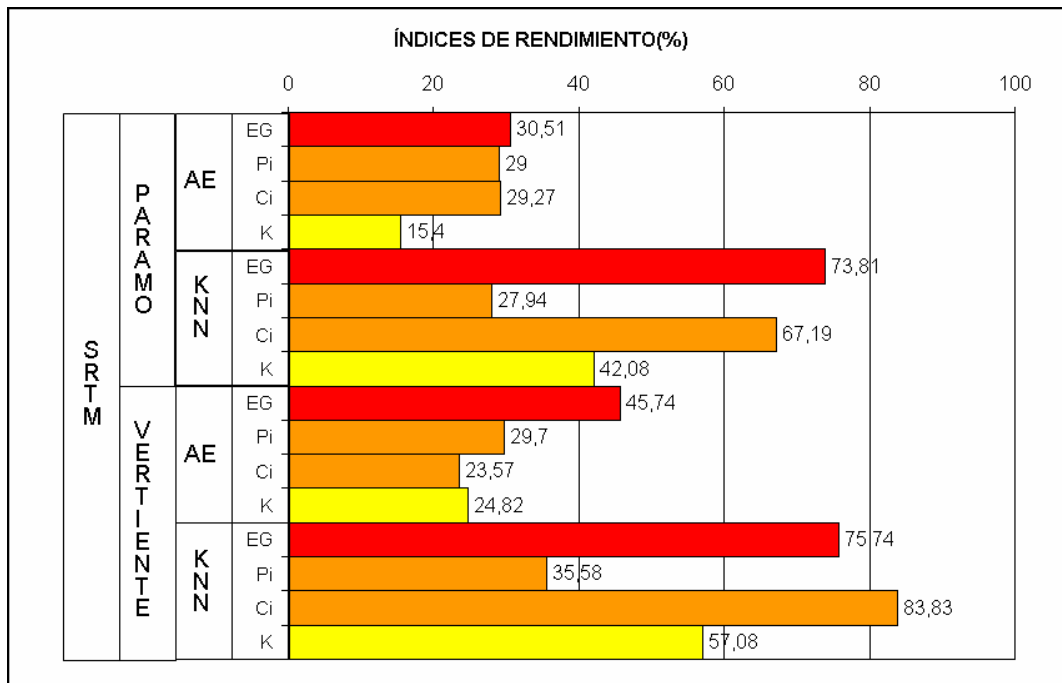
A grandes rasgos la clasificación geomorfométrica del SRTM según los índices derivados de la matriz de confusión aplicada (Figura 17) muestra:

- Un menor rendimiento del AE en todos los índices, llegando a considerarse insignificante la concordancia.
- Un mayor rendimiento del KNN, particularmente en EG y confiabilidad.
- La zona de análisis donde los índices muestran mejor desempeño es la zona de vertiente.
- El conjunto de índices medidos del AE, se muestra más compacto en sus apreciaciones, en contraste con el conjunto de los medidos del KNN, los cuales por ejemplo registran gran acierto en la confiabilidad (KNN vertiente, 83,83%), pero probablemente en demérito de la precisión (35,58%), lo cual deduce un nivel de error por comisión substancial.



**Tabla 7.** Índices de las medidas de rendimiento derivadas de la matriz de confusión, por método y áreas de aplicación, realizadas a partir de los parámetros extraídos del MDE SRTM (90m).

MDE	ÁREAS	MÉTODO	CLASIFICACIÓN	ÁREAS DE VALIDACIÓN	ÍNDICES
SRTM (90m)	PÁRAMO	AE			EG = 30,51%
					Pi = 29,00%
		KNN			EG = 73,81%
					Pi = 27,94%
	VERTIENTES	AE			Ci = 29,27%
					K = 15,40%
		KNN			EG = 45,74%
					Pi = 29,70%
VERTIENTES	AE			Ci = 23,57%	
				K = 24,82%	
	KNN			EG = 75,74%	
				Pi = 35,58%	
VERTIENTES	KNN			Ci = 83,83%	
				K = 57,08%	



**Figura 17.** Distribución de las medidas de rendimiento por método y áreas de aplicación, realizadas a partir de los parámetros extraídos del MDE SRTM (90m).

#### 4.5.- La clasificación geomorfométrica a partir del MDE Alos Palsar 12,5m

La mayor ventaja que ofrece un MDE de alta resolución espacial como el ALOS PALSAR a cualquier protocolo de clasificación geomorfométrica es su capacidad de reflejar con mayor veracidad las condiciones topográficas superficiales de un área determinada. Comparando las capacidades del MDE ALOS PALSAR con las mostradas por el MDE SRTM en términos de resolución espacial, puede decirse que las potencialidades para la determinación de parámetros geomorfométricos aumentan en más del 80%.

Sin embargo, una particularidad del procesamiento de los datos obtenidos a través del MDE ALOS PALSAR en comparación con el MDE SRTM, es el carácter heterogéneo de las categorías extraídas, siendo esto válido tanto para los parámetros geomorfométricos como para las clases geomorfométricas, debido a la inclusión de mayor información / pixel por unidad de área.

Si bien la resolución espacial del ALOS PALSAR confiere a las categorías espaciales, límites más suavizados, existe menor contigüidad en los conjuntos clasificados. Tal condición es apreciable tanto en la clasificación realizada con AE como con KNN. A manera de ilustración la tabla 8, describe los índices, áreas de validación, y clasificación obtenida a partir de los métodos propuestos y consecuente área de aplicación, basados en los datos aportados por el MDE ALOS PALSAR (12,5m).

##### *- La clasificación del Alos Palsar según el AE*

El algoritmo empírico (AE) basa su protocolo de segmentación en umbrales estadísticos definidos en el MDE SRTM (90m), por lo que su aplicación y evaluación en el MDE ALOS PALSAR (12,5m), revela que tan reproducible puede ser el AE y que ventajas o desventajas se obtienen al trabajar con MDE de mayor resolución espacial.

En una primera aproximación visual, se denota en la clasificación AE, una mayor presencia de formas de paisaje que en los correspondientes hechos con el MDE SRTM, además de un mayor suavizado dado por el tamaño del pixel, por ejemplo, si comparamos la clasificación geomorfométrica AE del SRTM aplicada a la zona páramo con la misma lograda con el MDE ALOS PALSAR (Tabla 8), se observa una mayor concordancia de esta misma con el área de validación, lo cual confirma un aspecto ventajoso de trabajar con mayor resolución espacial.

Para el caso de la zona de vertientes, el clasificador AE logra al menos visualmente configurar unidades espaciales similares a los polígonos provistos por las áreas de validación, pero con mayor dispersión de las clases, efecto visual que es menos percibido en la clasificación AE de vertiente hecha con el SRTM. Este resultado está ligado a la mayor resolución espacial del MDE lo cual introduce mayor ruido aún siendo el resultado aquí mostrado el producido por un filtro de mayoría 3x3.

En la comparación realizada con la matriz de confusión, la clasificación basada en AE obtenida con el MDE ALOS PALSAR describe valores discretos en los índices globales EG (páramo 47,77% y vertiente 53,79%) y kappa (páramo 31,79% y vertiente 32,15%). A simple vista, se

denota el carácter poligonal simple de las áreas de validación (anexo 2), en contraste con la copiosa segmentación que presenta la clasificación con AE, la cual es aún mayor que la generada con los datos del SRTM, no solo por el carácter estático de los límites que imponen los umbrales estadísticos, sino por la mayor resolución del MDE ALOS PALSAR (anexo 5).

Al igual que en la clasificación AE SRTM, clases lineales como crestas, morrenas, fondo de valle y carcavas, tienen en esta clasificación mayor ocurrencia, pero contrariamente pierden contigüidad como ocurre con las clases abanicos aluviales y complejo de terrazas, las cuales son mejor delineadas pero presentan mayor pixelización y por lo tanto no mantienen el carácter continuo y contiguo de esta forma de paisaje.

*- La clasificación del Alos Palsar según el KNN*

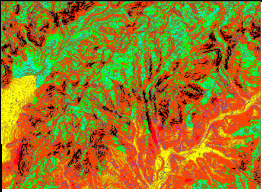
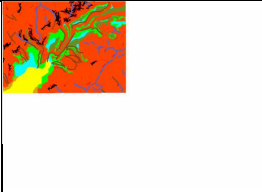
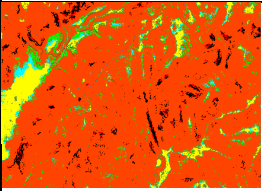
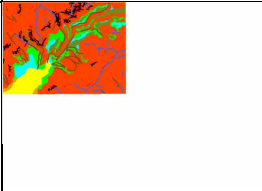
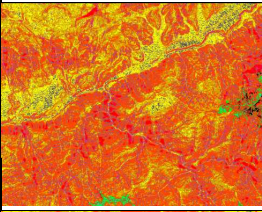
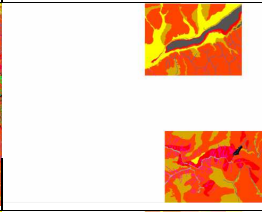
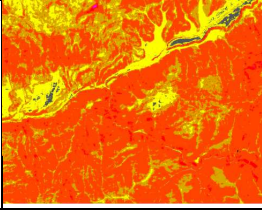
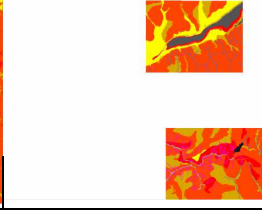
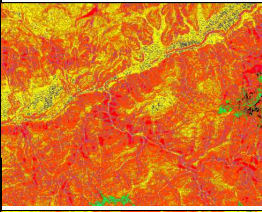
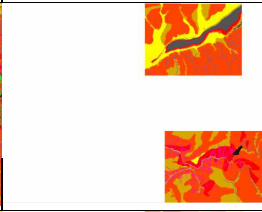
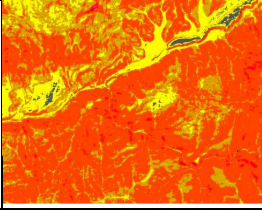
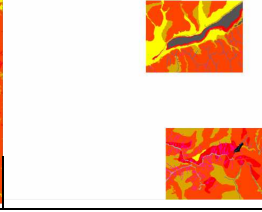
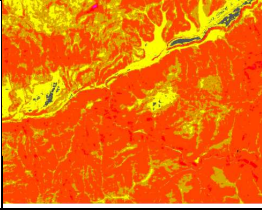
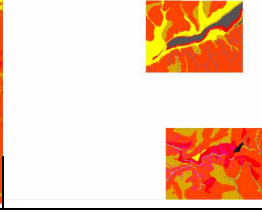
Dado que el algoritmo KNN clasifica cada dato nuevo en la clase de mayor frecuencia y con menores distancias, el efecto de pixeles dispersos en las zonas de análisis con el uso de parámetros derivados del MDE ALOS PALSAR se multiplican debido a la resolución del mismo (anexo 6). Este problema es común cuando se utilizan MDE de alta resolución espacial (Drăguț and Blaschke 2006).

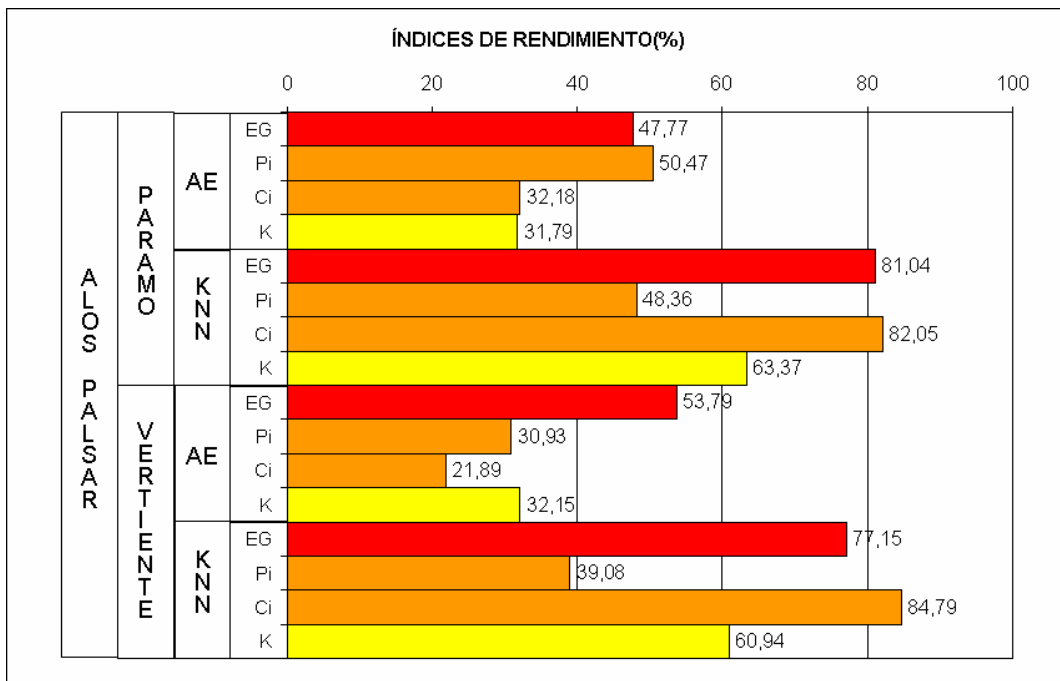
A pesar que las clases de mayor frecuencia y consecuentemente de menores distancias por su ubicuidad en las áreas de validación son predominantemente las clases: vertientes denudacionales y abanicos aluviales, la mayor resolución del ALOS PALSAR permite ser incluyente con otras clases menores tal como sucede en la clasificación hecha en la zona de páramo, donde el KNN logra reproducir con bastante acierto ( $EG = 81,04\%$ ), además de la vertientes, otras formas del paisaje como los valles glaciales, glacis y afloramientos rocosos.

Para el caso de la zona de vertiente el protocolo de aprendizaje del KNN permite establecer el carácter contiguo de las terrazas y abanicos aluviales en mayor definición que el reproducido con los parámetros del SRTM. La clasificación geomorfométrica generada con KNN a partir del ALOS PALSAR evaluada a través de los índices derivados de la matriz de confusión (Figura 18) muestra:

- Una mejoría en el rendimiento del AE en todos los índices comparado con los logros alcanzados con el MDE SRTM.
- Un definitivo mayor rendimiento del KNN, particularmente en EG (páramo 81,04% y vertiente 77,15%), y estadístico kappa (páramo 63,37% y vertiente 60,94%).
- A diferencia del resultado logrado con AE, en este caso la zona de análisis donde los índices muestran mejor desempeño es la zona de páramo.
- En todos los conjuntos de índices medidos tanto del AE como el KNN, se registra gran acierto en la confiabilidad, lo cual es probablemente en demérito de la precisión.

**Tabla 8.** Índices de las medidas de rendimiento derivadas de la matriz de confusión, por método y áreas de aplicación, realizadas a partir de los parámetros extraídos del MDE ALOS PALSAR (12,5m).

MDE	ÁREAS	MÉTODO	CLASIFICACIÓN	ÁREAS DE VALIDACIÓN	ÍNDICES
ALOS PALSAR	PÁRAMO	AE			EG = 47,77%
					Pi = 50,47%
		KNN			EG = 81,04%
					Pi = 48,36%
	VERTIENTES	AE			Ci = 32,18%
					K = 31,79%
		KNN			EG = 53,79%
					Pi = 30,93%
ALOS PALSAR	VERTIENTES	AE			Ci = 21,89%
					K = 32,15%
		KNN			EG = 77,15%
					Pi = 39,08%
ALOS PALSAR	VERTIENTES	KNN			Ci = 84,79%
					K = 60,94%



**Figura 18.** Distribución de las medidas de rendimiento por método y áreas de aplicación, realizadas a partir de los parámetros extraídos del MDE ALOS PALSAR (12,5m).

#### 4.6.- Análisis comparativo

Si bien por motivos de costo computacional el área de estudio referida en principio a la cuenca del Río Chama, Mérida – Venezuela, se concentró en dos zonas de análisis, una representando ambientes de alta montaña y el segundo mayormente enfocando en vertientes y valles. Sin embargo, el objetivo de este trabajo continuó siendo la evaluación del modelo de clasificación geomorfométrica automática ejecutado con el algoritmo KNN, comparandola con la clasificación obtenida de un algoritmo empírico manual, ambos aplicados a dos modelos digitales de elevación (MDE) ALOS PALSAR (12,5m) y SRTM (90m).

En una primera aproximación, se podría establecer como elementos determinantes en la evaluación a: el método de clasificación, el MDE y la zona de análisis. En este caso y de acuerdo a los datos ilustrados en la figura 19, se observa que:

- El AE se muestra menos eficiente que el KNN según el rendimiento descrito en EG y kappa.
- El primer determinante es el método KNN el cual se establece como el método de las cuatro primeras clasificaciones.
- El segundo determinante es el MDE, el cual se establece al MDE ALOS PALSAR como la mejor base para clasificar con cualquiera de los dos metodos comparados.
- La zona de análisis no es determinante en la eficiencia de los clasificadores.
- Existe una brecha entre el EG y kappa lo cual se explica por la sobreclasificación en ciertas clases en detrimento de otras lo cual aumenta la confiabilidad pero disminuye la precisión, manifestandose ampliamente en el caso de las clasificaciones: AE SRTM PÁRAMO y KNN SRTM PARAMO.
- Para efectos de este trabajo, se consideran buenas a las clasificaciones cuyos EG y kappa sean mayores a 70% y 50% respectivamente.
- Por su parte las clasificaciones de alcance deficiente son aquellas con EG menor al 50%, el rango restante se considera regular.

En el tabla 9 se describe la distribución comparativa de las clasificaciones geomorfométricas según su alcance, método, MDE y zona de análisis.

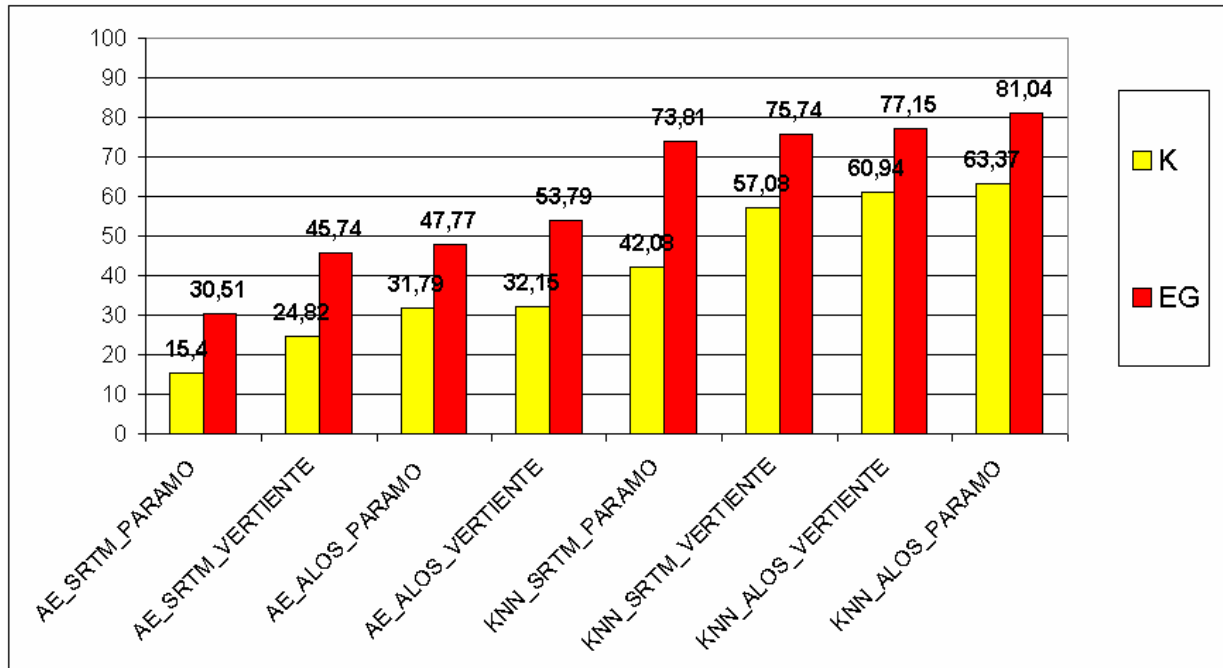


Figura 19. Distribución de los índices globales kappa y EG por método, áreas de aplicación y MDE base.

Tabla 9. Distribución comparativa de las clasificaciones geomorfométricas según su alcance, método, MDE y zona de análisis.

ALCANCE	MÈTODO	MDE	ZONA DE ANÁLISIS	CLASIFICACIÓN GEOMORFOMÉTRICA
Clasificación buena	KNN	ALOS PALSAR (12,5m)	PÁRAMO	KNN ALOS PÁRAMO
			VERTIENTE	KNN ALOS VERTIENTE
		SRTM (90m)	VERTIENTE	KNN SRTM VERTIENTE
Clasificación regular	KNN	SRTM (90m)	PARAMO	KNN SRTM PARAMO
	AE	ALOS PALSAR (12,5m)	VERTIENTE	AE ALOS VERTIENTE
Clasificación deficiente	AE	ALOS PALSAR (12,5m)	PÁRAMO	AE ALOS PÁRAMO
			VERTIENTE	AE SRTM VERTIENTE
		SRTM (90m)	PÁRAMO	AE SRTM PÁRAMO

El anterior análisis resume el rendimiento y alcance según la eficiencia de las clasificaciones geomorfométrica obtenidas, pero basado en medidas globales como lo son el EG y kappa. Se hace necesario entonces describir la eficiencia de la clasificación a nivel de clases. Para esto se utilizará el índice de Hellden (IH), el cual evalúa el acierto dado por un clasificador a nivel de clases tomando en cuenta los errores de omisión y comisión generados en la clasificación.

Al igual que en el anterior análisis, en esta evaluación también se considerarán: el método de clasificación, el MDE, la zona de análisis y las clases geomorfométricas, como elementos determinantes. De la lista de 104 clases que generó la aplicación del IH, se escogieron las primeras 10 clases geomorfométricas con mejor rendimiento (Tabla 10), y las últimas 10 clases con el menor (Tabla 11). De la interpretación de estos resultados se observa que:

- El AE se muestra menos eficiente que el KNN según el rendimiento descrito en EG y kappa.
- El primer determinante es el método KNN el cual se establece como el método de las cuatro primeras clasificaciones.
- El segundo determinante es el MDE, el cual se establece al MDE ALOS PALSAR como la mejor base para clasificar con cualquiera de los dos metodos comparados.
- La zona de análisis no es determinante en la eficiencia de los clasificadores.
- Existe una brecha entre el EG y kappa lo cual se explica por la sobreclasificación en ciertas clases en detrimento de otras lo cual aumenta la confiabilidad pero disminuye la precisión, manifestandose ampliamente en el caso de las clasificaciones: AE SRTM PÁRAMO y KNN SRTM PARAMO.
- Para efectos de este trabajo, se consideran buenas a las clasificaciones cuyos EG y kappa sean mayores a 70% y 50% respectivamente.
- Por su parte las clasificaciones de alcance deficiente son aquellas con EG menor al 50%, el rango restante se considera regular.
- El método con mejor rendimiento según el IH es el KNN.
- Las clases clasificadas que muestran mayor rendimiento en cuanto a su concordancia con las áreas de validación son las vertientes denudacionales y los abanicos aluviales en cualquier MDE empleado y con el método KNN.
- Las clases que muestran menor rendimiento en cuanto a su concordancia con las áreas de validación son las clases cuya expresión espacial es de menor extensión, lineales no ubicuas, tales como las crestas, morrenas-crestas, cárcavas y fondo de valles.
- Las clases de menor rendimiento no estan determinadas ni por el MDE empleado ni por el área de análisis, ni por el método de clasificación.

**Tabla 10.** Primeras 10 clases geomorfométricas con mejor rendimiento según el IH.

CLASES	MDE	AREA	METODO	HELLDEN (%)	PRECISIÓN (%)	CONFIDAD. (%)
Vertiente denudacional	ALOS P (12,5m)	Páramo	KNN	89,37	97,00	83,00
Vertiente denudacional	ALOS P (12,5m)	Vertiente	KNN	85,56	93,00	76,00
Vertiente denudacional	SRTM (90m)	Páramo	KNN	84,72	99,00	74,00
Abanico aluvial	SRTM (90m)	Páramo	KNN	84,30	98,00	74,00
Vertiente denudacional	SRTM (90m)	Vertiente	KNN	83,84	95,00	75,00
Afloramientos rocosos	ALOS P (12,5m)	Páramo	KNN	83,18	80,00	87,00
Abanico aluvial	SRTM (90m)	Páramo	AE	82,51	82,00	83,00
Abanico aluvial	ALOS P (12,5m)	Vertiente	KNN	81,84	92,00	74,00
Abanico aluvial	SRTM (90m)	Vertiente	KNN	81,39	87,00	76,00
Abanico aluvial	ALOS P (12,5m)	Páramo	KNN	80,38	95,00	69,00


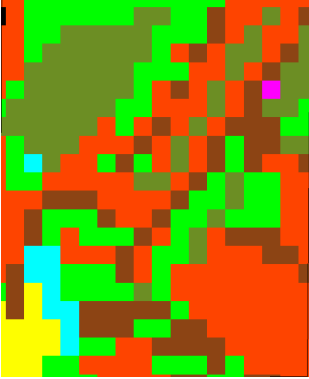


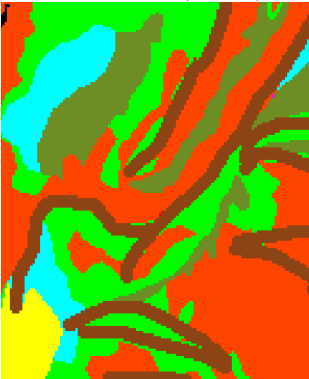


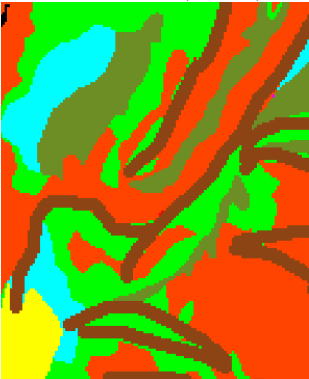
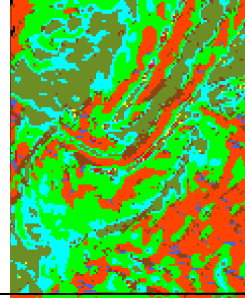
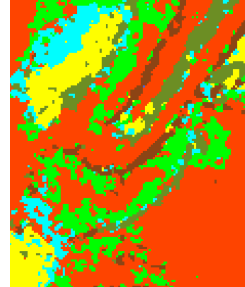
**Tabla 11.** Últimas 10 clases geomorfométricas con menor rendimiento según el IH.

CLASES	MDE	AREA	METODO	HELLDEN (%)	PRECISIÓN (%)	CONFIDAD. (%)
Crestas	ALOS P (12,5m)	Vertiente	AE	5,82	22,00	3,00
Morrenas y crestas	SRTM (90m)	Páramo	AE	5,53	16,00	3,00
Vertiente coluvial	ALOS P (12,5m)	Páramo	KNN	5,01	3,00	100,00
Cárcavas	ALOS P (12,5m)	Vertiente	AE	4,92	51,00	40,00
Crestas	ALOS P (12,5m)	Vertiente	KNN	4,63	2,00	96,00
Cárcavas	SRTM (90m)	Vertiente	AE	4,12	8,00	3,00
Cárcavas	ALOS P (12,5m)	Vertiente	KNN	3,32	2,00	92,00
Crestas	SRTM (90m)	Vertiente	AE	2,67	16,00	1,00
Morrenas y crestas	SRTM (90m)	Páramo	KNN	1,79	1,00	33,00
Fondo de valle	SRTM (90m)	Vertiente	KNN	1,57	0,00	100,00


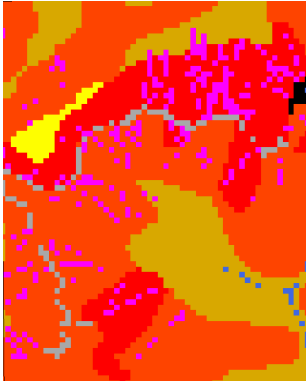
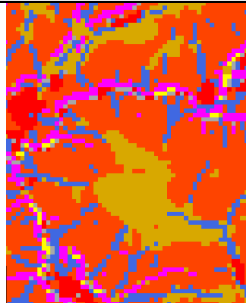
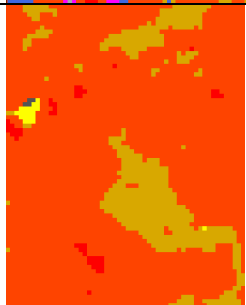
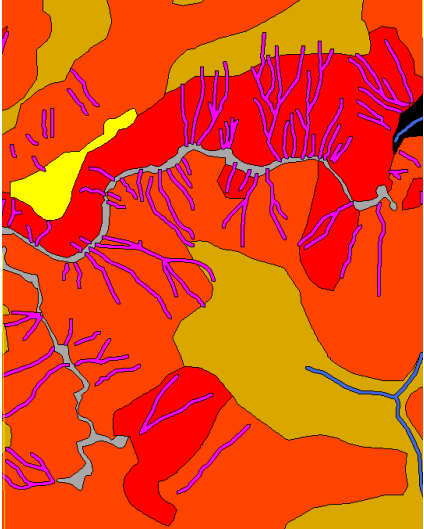
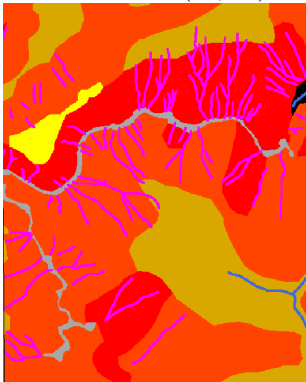
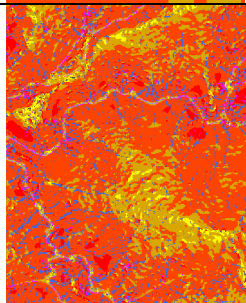
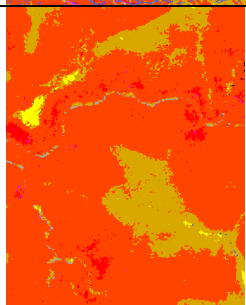
Las medidas provenientes de la matriz de confusión favorecen la clasificación que acierta a la clase dominante aún cuando esta puede estar generalizando la distribución de clases con reducida expresión espacial. Por ejemplo en el caso de la clase vertiente denudacionales, la cual es ampliamente distribuida en las zonas de análisis, cualquier método de clasificación aún por azar puede acertar y producir un nivel de concordancia que luego es registrado como eficaz por el EG e inclusive el IH.

Sin embargo, esta clase puede haberse asignado a clases vecinas que por su carácter lineal como morrenas, crestas, cárcavas y fondo de valle, son luego remuestreadas a la resolución del MDE respectivo para su procesamiento, por lo que no mantienen un mínimo de área efectiva para lograr ser interpretadas en los procesos de clasificación y por lo tanto presentan muy bajo rendimiento. En las figuras 20 y 21 se ilustra el efecto del remuestreo y método de clasificación en la asignación de clases geomorfométricas en dos limitadas porciones tanto de la zona de páramo como la de vertiente.



ÁREA DE VALIDACIÓN PÁRAMO, MORRENAS-CRESTAS ▶	REMUESTREO SEGÚN MDE ▶	CLASIFICACIÓN SEGÚN MÉTODO	
cc452 Sentinel 2A 	SRTM(90m) 	AE	
Fotointerpretación geomorfológica 	ALOS PALSAR(12,5m) 	KNN	
Fotointerpretación geomorfológica 	ALOS PALSAR(12,5m) 	AE	
		KNN	

**Figura 20.** Clases de mayor distribución y ubicuidad como las vertientes denudacionales pueden avasallar a clases de carácter lineal como morrenas y valles glaciales las cuales pueden en un principio ser generalizados por el proceso de remuestreo a la resolución de cada MDE empleado y luego subestimados por el método de clasificación. En el caso del AE, este logra clasificar la morrena señalada en el área de validación pero reduce su eficacia cuando se aplica al MDE ALOS PALSAR. Caso contrario el KNN no logra definir la morrena con los datos provenientes del MDE SRTM pero si la clasifica cuando se aplica al MDE ALOS PALSAR.

ÁREA DE VALIDACIÓN VERTIENTE CARCAVAS, FONDO DE VALLE ▶	REMUESTREO SEGÚN MDE ▶	CLASIFICACIÓN SEGÚN MÉTODO	
cc452 Sentinel 2A 	SRTM(90m) 	AE	
		KNN	
Fotointerpretación geomorfológica 	ALOS PALSAR(12,5m) 	AE	
		KNN	

**Figura 21.** Al tener mayor distribución y ubicuidad clases como vertientes denudacionales se sobreponen a clases como crestas y carcavas, cuyo carácter lineal puede ser generalizado por el proceso de remuestreo a la resolución de cada MDE empleado y luego subestimados por el método de clasificación. Para el caso del AE, este método logra clasificar crestas y carcavas similares a las mismas señaladas en el área de validación, manteniendo su eficacia cuando se aplica al MDE ALOS PALSAR. Por el contrario el KNN no logra definir crestas y carcavas ni con los datos provenientes del MDE SRTM o ALOS PALSAR al generalizar el área como vertiente denudacional.

Según los resultados aquí obtenidos la clasificación geomorfométrica de mayor acierto y concordancia con respecto a las áreas de validación es la obtenida a partir de la aplicación del

KNN a los parámetros geomorfométricos extraídos del MDE ALOS PALSAR (12,5m). Sin embargo, aunque en algunos casos es posible evaluar y clasificar la localización espacial de cauces y cuencas hidrográficas (Wise, 2000), otras clases y parámetros geomorfométricos tales como la pendiente dependen de la escala de análisis (Florinsky, 1998), por lo tanto, no existen guías consistentes que evalúen la exactitud de parámetros y clases geomorfométricas en el campo. La mayoría de las evaluaciones sobre los MDE y sus parámetros de superficie son aún de carácter visual (Reuter, et al., 2009).

Tal como lo señala Gercek et al. (2011), la clasificación de las formas de paisaje posee características muy diferentes a la clasificación del uso de la tierra o de coberturas de superficie debido a:

- Las formas del paisaje son características basadas en la percepción y semántica y no en características “reales” como puede ser un cultivo o un bosque.
- La superficie es continua naturalmente y las formas de paisaje no tienen límites determinados.
- La clasificación de formas de paisaje depende de la ventana de percepción es decir es escala dependiente.

La evaluación de la clasificación geomorfométrica es un caso especial dentro de los procedimientos estándar de la evaluación de clasificaciones de imágenes, por lo que se recomienda una medida de similaridad ponderada entre clases como alternativa a las medidas de rendimiento tradicionales (Gerçek et al., 2011).

## **5.- CONCLUSIONES**

En este trabajo se ha probado la eficiencia de una clasificación geomorfométrica utilizando tres variables claves como son la fuente del MDE, la zona de análisis y el clasificador aplicado. Cada una de estos determinantes se empleo para generar una clasificación geomorfométrica final que luego fue evaluada en terminos de su nivel de acierto con los índices derivados de la matriz de confusión.

En cuanto a los MDE, se emplearon el MDE SRTM v4 (90m) y el ALOS PALSAR (12,5m), de los cuales el que sustentó la clasificación mejor evaluada fue el ALOS PALSAR, cuyo mayor detalle espacial permitió enriquecer la información aportada por los parámetros geomorfométricos extraídos y de aquí mejorar la expresión espacial de las formas de paisaje clasificadas. Si bien el MDE ALOS PALSAR dada su alta resolución espacial aporta ruido y mayor segmentación en la integración final de los datos, es muy conveniente para estudios a nivel de cuenca hidrográfica y mayor detalle. Las desventajas antes mencionadas podrían corregirse con una mayor calibración de los parámetros y un filtro de mayoría al resultado final. El MDE SRTM v4, es idóneo para desarrollar análisis cuantitativos de formas de paisaje a nivel regional debido a la consistencia de sus datos y donde se precisa una leyenda más generalizada de formas de paisaje.

No se evidenció ventaja alguna en los resultados obtenidos por utilizar alguna zona de análisis en particular, tanto la zona de páramo como la de vertiente reportaron logros y fallas similares. La única ventaja que poseía la zona de páramo fue la de un menor costo/tiempo computacional, comparado con la zona de vertiente, particularmente cuando se ejecutaba el KNN, lo cual no significó una saturación del proceso que pudiese alterar resultados. A pesar que para efectos de aplicación de los métodos no se utilizó la totalidad de la cuenca, se decidió mantener el concepto de cuenca como área de estudio para no perder la noción de relieve y conjunto de procesos que generados en una cuenca determinan la dimensión de las formas existentes en esta.

El clasificador que logró la clasificación mejor evaluada fue el KNN con los siguientes parámetros:  $n\_neighbors = 5$ ,  $weight = uniform$  distancia métrica = *minkowski* y  $p = 1$ . Sin embargo se advierte que:

- La generalización es considerablemente mayor en las clasificaciones obtenidas con el KNN, tanto en cualquier zona de análisis como bajo cualquier MDE.
- La segmentación es mayor en las clasificaciones obtenidas con la aplicación del AE.
- Mientras el KNN generaliza clases en detrimento de otras, en el AE la rigidez de los umbrales permite la representación certera o fallida de las clases esperadas.

Puede decirse que bajo los objetivos, materiales y métodos aplicados, la mejor clasificación geomorfométrica lograda fue la generada a partir del MDE ALOS PALSAR y aplicando el KNN como método clasificador. Dado que las clases geomorfométricas delineadas en las zonas de validación provienen de un ejercicio manual de fotointerpretación geomorfológica, estas clases se perfilan más como polígonos en cuya área probablemente exista cierto nivel de generalización de la diversidad geomorfológica existente, por lo que la correlación entre las áreas de validación y clasificaciones con sustanciales niveles de generalización son esperados.

En cuanto al trabajo futuro es importante experimentar con otros métodos de aprendizaje automático tales como los métodos de árboles aleatorios (random forest), los cuales pueden aportar nuevas ideas sobre la clasificación de formas de paisajes y niveles de eficiencia. Se debe mejorar las áreas de entrenamiento mediante la formulación de formas más basadas en la geometría que en los procesos geomorfológicos así como también que su dimensión espacial sea suficiente como para sobreponerse a la generalización implícita en todo el proceso de clasificación. Es evidente que la matriz de confusión e índices derivados no son métodos adecuados para evaluar el desempeño tanto del AE como del KNN, puesto que generalizan clases predominantes sobre clases de menor distribución y extensión.

## 6.- BIBLIOGRAFÍA

- Amatulli G, McInerney D, Sethi T, Strobl P, Domisch S. 2019. Geomorpho90m - Global high-resolution geomorphometry layers: empirical evaluation and accuracy assessment. *PeerJ Preprints* 7:e27595v1 <https://doi.org/10.7287/peerj.preprints.27595v1>
- Argyros Argyridis & Demetre Argialas. 2017. GEOBIA-based identification of alluvial fans and bajadas through geomorphometry, image análisis and fuzzy ontology, *International Journal of Image and Data Fusion*, 8:4, 309-331, DOI: 10.1080/19479832.2017.1342707
- ASF DAAC. 2019. ALOS PALSAR\_Radiometric\_Terrain\_Corrected\_high\_res. En línea: <https://earthdata.nasa.gov/about/daacs/daac-asf> Última visita 11 Mayo 2019. DOI: 10.5067/JBYK3J6HFSVF
- Chuvieco, E., & Huete, A. (2010). *Fundamentals of satellite remote sensing*. CRC Press.
- Clarke, K.C., & B.E. Romero. 2016. On the topology of topography: A review. *Cartography & Geographic Information Science*, 44, 271-282.
- Cohen, J. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37-46. <http://doi.org/10.1177/001316446002000104>
- Dekavalla Maria & Demetre Argialas. 2017. Evaluation of a spatially adaptive approach for land surface classification from digital elevation models, *International Journal of Geographical Information Science*, 31:10, 1978-2000, DOI: 10.1080/13658816.2017.1344984
- Dikau, R. 1989. The Application of a Digital Relief Model to Landform Analysis in Geomorphology. In *Three Dimensional Applications in Geographical Information Systems*, edited by J. F. Raper, 51-77. London: Taylor & Francis.
- Drăguț, L., and C. Eisank. 2011. Object representations at multiple scales from digital elevation models. *Geomorphology* 129, no. 3-4: 183-189.
- Drăguț L., Eisank C., Strasser T., & T. Blaschke. 2009. A comparison of methods to incorporate scale in geomorphometry. In: Purves R., Gruber S., Straumann R., Hengl T., editors. *Proceedings Geomorphometry 2009*. University of Zurich; Zurich: pp. 133-139
- Drăguț, L. and T., Blaschke. 2006. Automated classification of landform elements using objectbased image analysis. *Geomorphology*, 81, 330-344. doi:10.1016/j.geomorph.2006.04.013
- Evans, I., and J. Minár. 2011. A classification of geomorphometric variables. In *Geomorphometry 2011*, edited by T. Hengl, I. S. Evans, J. P. Wilson and M. Gould, 105-108. Redlands, CA.
- Evans, I. 1980. An Integrated System of Terrain Analysis and Slope Mapping. *Zeitschrift für Geomorphologie N.F., Supplement-Band 36 (1980): 274-295.*
- Florinsky, I.V. 1998. Combined analysis of digital terrain models and remotely sensed data in landscape investigations. *Prog. Phys. Geogr.* 22, 33-60.
- Gallant, A. L., T. R. Loveland, T. L. Sohl, and D. Napton. 2005. Using an ecoregion framework to analyze land cover and land use dynamics. *Environmental Management* 34 (): S89-S110.

Gerçek Deniz., Vedat Toprak & Josef Strobl. 2011. Object-based classification of landforms based on their local geometry and geomorphometric context, *International Journal of Geographical Information Science*, 25:6, 1011-1023, DOI: 10.1080/13658816.2011.558845

Ghosh S., Stepinski T., & R. Vilalta. 2010. Automatic annotation of planetary surfaces with geomorphic labels. *IEEE Transactions on Geoscience and Remote Sensing*; 48:175–185.

Hengl, T., and D. G. Rossiter. 2003. Supervised landform classification to enhance and replace photo-interpretation in semi-detailed soil survey. *Soil Science Society America Journal* 67: 1810-1822.

Hellden, U. 1980. A test of Landsat-2 imagery and digital data for thematic mapping illustrated by an environmental study in northern Kenya. Sweden, Lund University, Natural Geography Institute Report No. 47.

Hudson, W. D., & Ramm, C. W. 1987. Correct formulation of the Kappa coefficient of agreement. *Photogrammetric Engineering & Remote Sensing*, 53(4), 421–422.

Iwahashi, J., and R. J. Pike. 2007. Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. *Geomorphology* 86: 409-440.

Jasiewicz Jarosław & Tomasz F. Stepinski. 2013. Geomorphons — a pattern recognition approach to classification and mapping of landforms, *Geomorphology*, Volume 182, 147-156.

Lin, S., Jing, C., Coles, N. A., Chaplot, V., Moore, N. J., & Wu, J. 2013. Evaluating DEM source and resolution uncertainties in the Soil and Water Assessment Tool. *Stochastic Environmental Research and Risk Assessment*, 27(1), 209–221. <https://doi.org/10.1007/s00477-012-0577-x>

MacMillan, R., Shary, P., 2009. Landforms and landform elements in geomorphometry. In: Hengl, T., Reuter, H. (Eds.), *Geomorphometry, Concepts, Software, Application*. Elsevier, pp. 227–275 (Ch. 9).

MacMillan, R. A., R. K. Jones, and D. H. McNabb. 2004. Defining a hierarchy of spatial entities for environmental analysis and modeling using digital elevation models (DEMs). *Computers, Environment and Urban Systems* 28 (): 175-200.

MacMillan, R., Pettapiece, W., Nolan, S., Goddard, T. 2000. A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. *Fuzzy Sets and Systems* 113, 81–109.

Miliaresis G. 2006. Geomorphometric mapping of asia minor from globe digital elevation model. *Geografiska Annaler: Series A, Physical Geography*, 88:3, 209-221.

Minar J., Evans I.S. 2008. Elementary forms for land surface segmentation: the theoretical basis of terrain analysis and geomorphological mapping. *Geomorphology*.;95:236–259

Olaya V. 2009. Basic land-surface parameters. In: Hengl T., Reuter H.I., editors. *Geomorphometry — Concepts, Software, Applications*. vol. 33. Elsevier; Amsterdam:. pp. 141–169.

Pennock, D. J., B. J. Zebarth, and E. de Jong. 1987. Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada. *Geoderma* 40: 297-315.

Pike, R. J. 1988. The geometric signature: quantifying landslide-terrain types from Digital Elevation Models. *Mathematical Geology* 20 (1988): 491-511

Reuter, H.I., Hengl, T., Gessler, P. & P. Soille. 2009. Chapter 4 Preparation of DEMs for Geomorphometric Analysis, Editor(s): Tomislav Hengl, Hannes I. Reuter, *Developments in Soil Science*, Elsevier, Volume 33, Pages 87-120, ISSN 0166-2481, ISBN 9780123743459.

Skidmore, A. K. 1990. Terrain position as mapped from a gridded digital elevation model. *International Journal of Geographical Information Systems* 4: 33-49

Stepinski T.F., Bagaria C. 2009. Segmentation-based unsupervised terrain classification for generation of physiographic maps. *IEEE Geoscience and Remote Sensing Letters*; 6:733–737.

Valentine, A. & L. Kalnins. 2016. An introduction to learning algorithms and potential applications in geomorphometry and Earth surface dynamics, *Earth Surf. Dynam.*, 4, 445–460, doi:10.5194/esurf-4-445-2016.

van Asselen, S., A.C. Seijmonsbergen. 2006. Expert-driven semi-automated geomorphological mapping for a mountainous area using a laser DTM, *Geomorphology*, Volume 78, Issues 3–4, Pages 309-320, ISSN 0169-555X, <https://doi.org/10.1016/j.geomorph.2006.01.037>.

Williams, K., & M. Zuber. 1998. Measurement and analysis of Lunar basin depths from Clementine altimetry. *Icarus* 131, 107-122.

Wood, J. 1996. *The Geomorphological Characterisation of Digital Elevation Models*. University of Leicester. Thesis PhD.

## 7.- ANEXOS

### Anexo 1. Copia del script en python utilizado para ejecutar la clasificación del KNN

El siguiente código tiene como objetivo realizar una clasificación supervisada a partir del método

1-Knn, con 3, 5 y 9 vecinos más cercanos

El conjunto de datos a procesar son los parámetros geomorfológicos extraídos de modelos de elevación digital (MDE), tales como SRTM (90m) y ALOS PALSAR (12,5m). Los parámetros son 6 a saber:

1.- Pendiente angular (Slope angle) 2.- Relieve Interno (Internal Relief) 3.- Convexidad Longitudinal (Longitudinal Convexity) 4.- Plano de Curvatura (Plan Curvature) 5.- Perfil de Convexidad (Profile Convexity) 6.- Convexidad Cruzada (Cross Convexity)

se aplica este algoritmo a un conjunto de datos por cada MDE.

Se incorporará también una capa de datos sobre áreas de entrenamiento / validación.

Para realizar operaciones con imágenes se emplea la librería imageio. Todas las funciones de dicha librería se pueden consultar en la siguiente URL:

<http://imageio.readthedocs.io/en/latest/userapi.html>

In [ ]: *#Cargar la librería que permite el tratamiento de las imágenes*

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

pd.set_option('display.notebook_repr_html', False)
pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', 150)
pd.set_option('display.max_seq_items', None)

##config InlineBackend.figure_formats = {'pdf',}
%matplotlib inline
```



```
In [30]: #http://imageio.readthedocs.io/en/latest/userapi.html
# Importar la clase imageio para realizar la lectura de las imágenes con su función volread

import imageio

bandas= imageio.volread('MATRIZ_ESTUDIO.tif')

bandas=np.array(bandas)

#En el Array deseamos tener es una matriz tridimensional
# [NÚMERO DE FILAS, NÚMERO DE COLUMNAS, NÚMERO DE BANDAS]

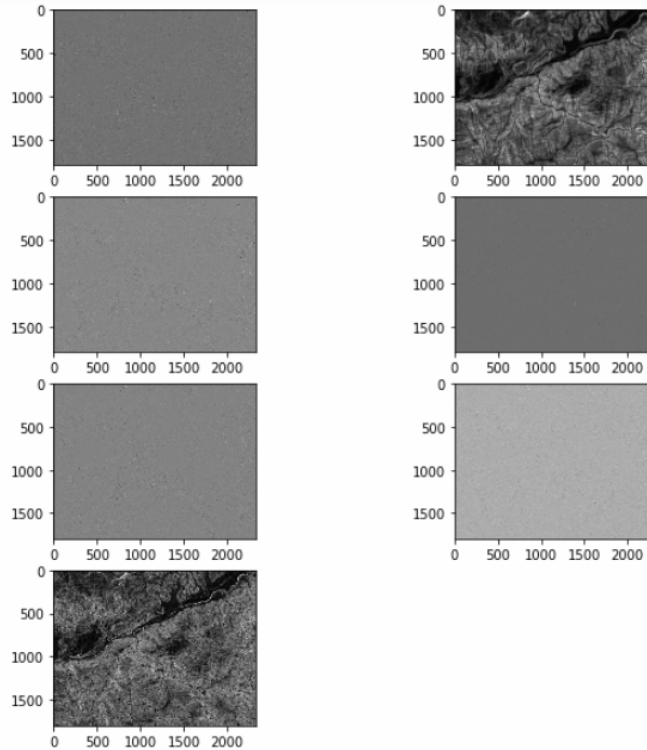
bandas = bandas.swapaxes(0,2)
bandas = bandas.swapaxes(0,1)

print(type(bandas))
print(bandas.dtype)
print("Dimensiones tras la lectura: \n", bandas.shape)

<class 'numpy.ndarray'>
float64
Dimensiones tras la lectura:
(1798, 2335, 7)
```

```
In [33]: # Creamos una figura
fig = plt.figure(1, figsize=(10,10))
# Vamos a mostrar la imagen de cada banda en blanco y negro
for i in range(0,7): # Todas las matrices menos la última porque es la de las
    # Hacemos una gráfica compuesta por 2 filas y 2 columnas
    # La función add_subplot tiene 3 argumentos
    # El primero indica el número de filas de la figura
    # El segundo indica el número de columnas de la figura
    # El último indica la posición que vamos a utilizar en ese momento
    ax = fig.add_subplot(4, 2, i+1)
    # Mostramos la imagen en escala de grises (cmap=plt.cm.gray)
    plt.imshow(bandas[:, :, i], cmap=plt.cm.gray)

plt.show()
```



Documentación de la librería ndimage para realizar procesamiento de imagen en la URL:  
<https://docs.scipy.org/doc/scipy/reference/tutorial/ndimage.html>

Matplotlib para realizar gráficas en Python en la URL:  
[https://matplotlib.org/users/plot\\_tutorial.html](https://matplotlib.org/users/plot_tutorial.html)

In [69]: *#Cargar áreas de entrenamiento*

```
areas_entrenamiento=np.array(imageio.imread('zonas_entrenamiento.tif'))
print(areas_entrenamiento[5,:])
print(type(areas_entrenamiento))
print(areas_entrenamiento.dtype)
print("Dimensiones tras la lectura: \n", areas_entrenamiento.shape)
```

```
[0 0 0 ... 0 0 0]
<class 'numpy.ndarray'>
uint8
Dimensiones tras la lectura:
(1798, 2335)
```

Ahora visualizar la imagen con las clases de todos los píxeles (los que servirán para entrenar). Los píxeles no clasificados están etiquetados con el valor 0 (serán mostrados en blanco).

```

import matplotlib.patches as mpatches
from matplotlib.colors import ListedColormap, BoundaryNorm

# Se define un array con los nombres de las clases (tipos de cultivo)
clases = ['Sin clasificar', 'a', 'b', 'c', 'd', 'e', 'f', 'h', 'i', 'j']

# Creamos un mapa de colores con colores prefijados para cada tipo de cultivo
cmap = ListedColormap(['w', 'c', 'b', 'y', 'g', 'k', 'r', 'm',])

# Se delimita el rango de cada color para que en todas las imágenes
#salgan los mismos colores para cada tipo de cultivo
bounds = np.arange(10)
norm = BoundaryNorm(bounds, cmap.N)

# Se obtienen los valores únicos de las clases (0-9)
values = np.unique(areas_entrenamiento.ravel())
print(values)
# Se crea una figura
plt.figure(1, figsize=(10,10))

# mostrar la imagen con la información de las clases: se presenta en color para ver diferencias
im = plt.imshow(areas_entrenamiento[:,:], cmap)

# Se crea un vector de tantos colores como valores únicos predichos
#de acuerdo al mapa de color utilizado en imshow
colors = [im.cmap(im.norm(value)) for value in values]

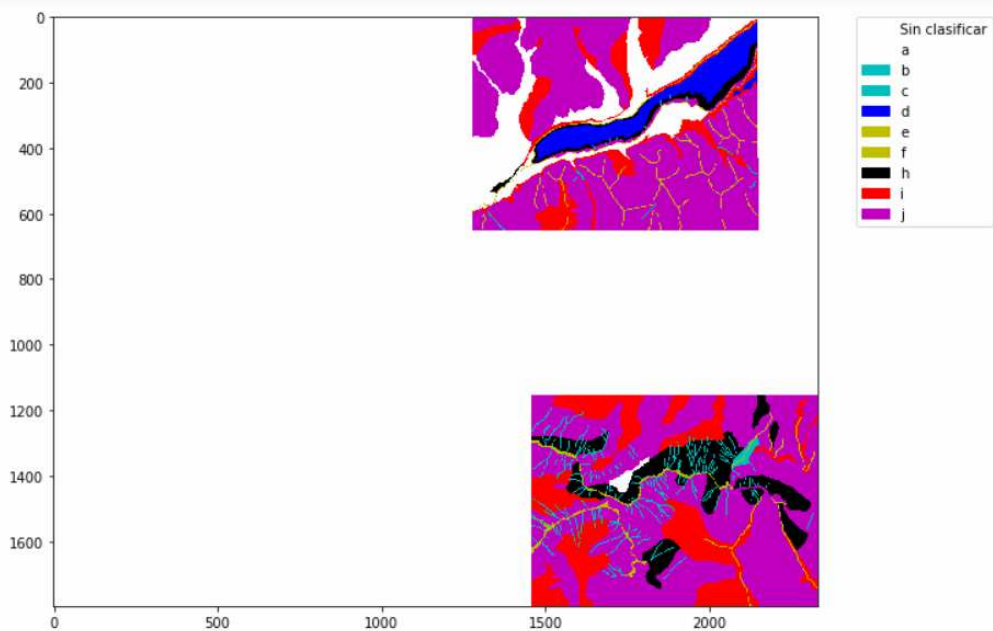
# Se crea un patch para cada color: relaciona cada color con cada etiqueta (clase)
patches = [mpatches.Patch(color=colors[i], label=clases[i]) for i in range(len(values))]

# Se ponen los patch creados como leyenda de la figura y se define la posición de la leyenda
#(a la derecha y fuera de la imagen)
plt.legend(handles=patches, bbox_to_anchor=(1.05, 1), loc=2, borderaxespad=0.)

# Se muestra la figura
plt.show()

```

[ 0 1 2 3 4 5 6 9 11 13]



```
In [74]: # Se obtienen las posiciones de los píxeles de las áreas de entrenamiento
# Esta condición se debe poner en la condición de la función argwhere de numpy
indicesTrain = np.argwhere(areas_entrenamiento[:,:]>0)

# índices de los 5 primeros píxeles de zonas de entrenamiento
print(indicesTrain[:5])

# dimensiones de los índices obtenidos
print(indicesTrain.shape)
```

```
[[ 3 1282]
 [ 3 1283]
 [ 3 1284]
 [ 3 1285]
 [ 3 1286]]
(1126927, 2)
```

Ahora se forma el conjunto de entrenamiento. Para ello, se crean dos matrices:

1-Xtrain (entradas)  
 2-Ytrain (salidas)

Para crear la matriz de entradas (Xtrain) se deben tomar los valores de todas las matrices (excepto banda 1) de todos los píxeles que ocupen áreas de entrenamiento (indicesTrain). Es decir, las variables de entrada van a ser las bandas del sensor.

Como la matriz es tridimensional, en la primera posición se determinan las filas de todos los elementos que se desean seleccionar, en la segunda posición se deben determinar las columnas y en la tercera las capas de la matriz.

Para crear la matriz de salidas (Ytrain) se debe realizar el mismo proceso pero tomando solamente los valores de la última capa de la matriz (banda).

```
In [75]: #Matriz de entrada con variables
Xtrain = bandas[indicesTrain[:,0],indicesTrain[:,1],0:7]

#Matriz de predicción
Ytrain = areas_entrenamiento[indicesTrain[:,0],indicesTrain[:,1]]

print(type(Xtrain),type(Ytrain))
print(Xtrain.dtype, Ytrain.dtype )
print("Dimensiones tras la lectura: \n", Xtrain.shape, Ytrain.shape)
```

```
<class 'numpy.ndarray'> <class 'numpy.ndarray'>
float64 uint8
Dimensiones tras la lectura:
(1126927, 7) (1126927,)
```

Mostrar el número de píxeles de cada clase. Para ello se utiliza la función `bincount` de NumPy pasando como argumento de entrada la variable que contiene las salidas (Ytrain). Esta función cuenta el número de elementos con cada valor.

```
In [77]: PixelesPorClase= np.bincount((Ytrain))
print(PixelesPorClase)

[ 0 115231 2828 26656 44975 14473 13057 0 0 86933
 0 191670 0 631104]
```

Generado el conjunto de entrenamiento es posible aplicar cualquier método de aprendizaje. En este, el algoritmo de los k vecinos más cercanos.

```
In [78]: # Importar la librería neighbors
from sklearn import neighbors

# Llamamos al constructor de KNN
clasificador = neighbors.KNeighborsClassifier()
```

```
In [79]: # Entrenar KNN
clasificador = clasificador.fit(Xtrain,Ytrain)

#imprimir el modelo
print(clasificador)

KNeighborsClassifier(algorithm='auto', leaf_size=30, metric='minkowski',
                    metric_params=None, n_jobs=None, n_neighbors=5, p=2,
                    weights='uniform')
```

```
In [80]: # Realizar las predicciones para los píxeles de entrenamiento
predicciones = clasificador.predict(Xtrain)
```

```
In [81]: # Importamos la librería metrics
from sklearn import metrics

# Calculamos el porcentaje de acierto
acc = (metrics.accuracy_score(Ytrain,predicciones)*100)
print(acc)
```

73.59411922866343

Matriz de confusión para observar en qué clases se confunde más el clasificador. Utilizar la función `confusion_matrix` de la librería `metrics`.

```
In [82]: # Mostramos la matriz de confusión
print(metrics.confusion_matrix(predicciones, Ytrain))

[[ 98996   12    66  16868   615   3143   1108   9322   5462]
 [     1   349    21     8    21    20    73    57    175]
 [     7    22  2763    24    62   106  1021   442  1927]
 [  6872    21    20  23537   178   530   167  1202  1016]
 [    64    12    33    49  1378    49   137   562   830]
 [   289    29   114   159    33  2842   196   652  1089]
 [   283   183  2552   113   309   134 24968   1806 12185]
 [  5982   273  2378  2604  4061  2641  5995 117615  51516]
 [   2737  1927 18709  1613   7816  3592  53268  60012 556904]]
```

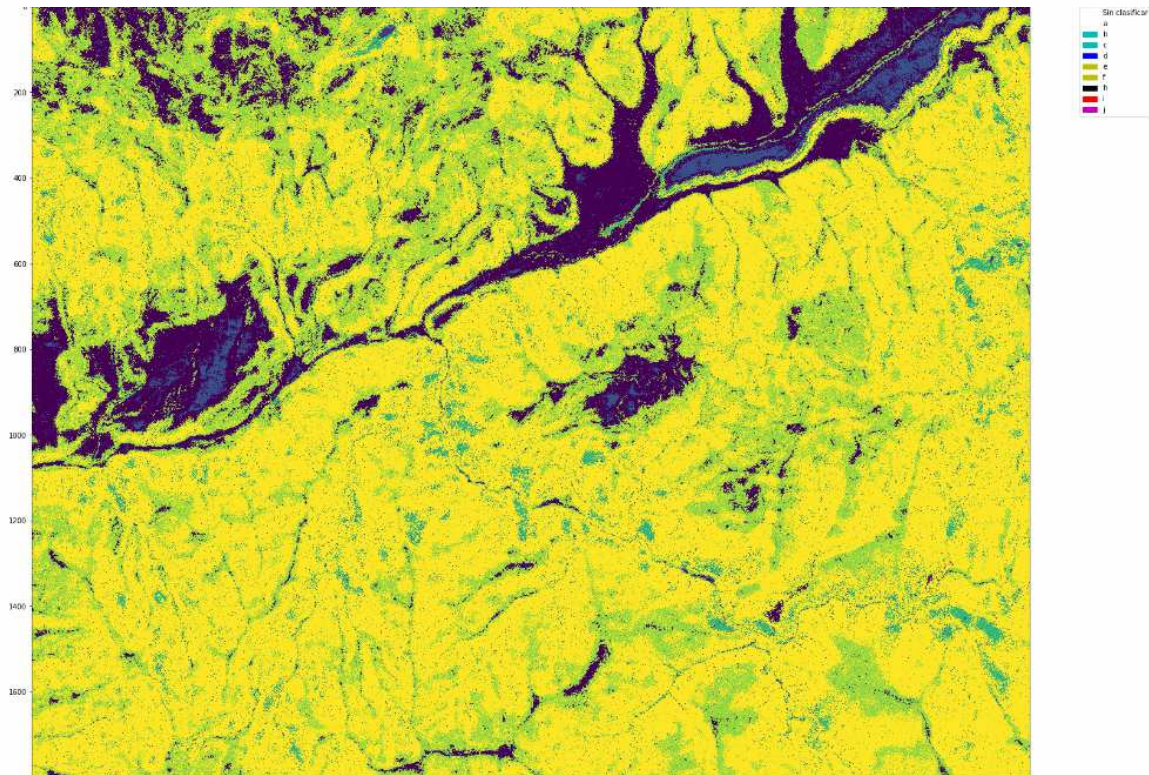
```
In [83]: # Obtenemos los índices de todos los píxeles
indicesTodo = np.argwhere(areas_entrenamiento[:,:]>=0)

# Creamos la matriz de entrada de todos los píxeles
XTodo = bandas[indicesTodo[:,0],indicesTodo[:,1],0:7]
```

```
In [85]: # Predicciones de todos Los píxeles
prediccionesTodo = clasificador.predict(XTodo)
```

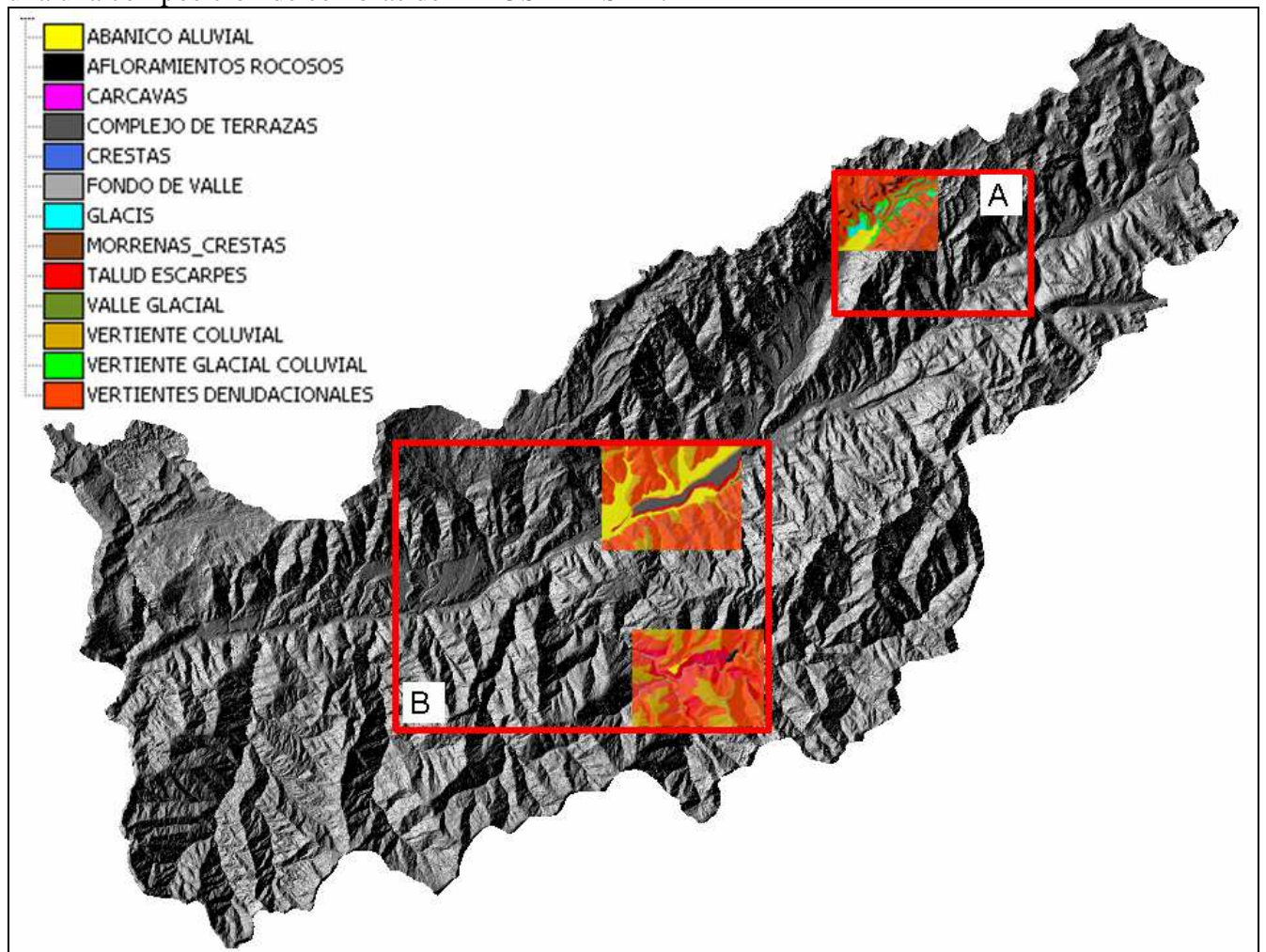
Para ver gráficamente la clasificación geomorfométrica en toda el área de análisis se genera una figura.

```
In [86]: # Se cambia el tamaño de las predicciones para crear una matriz
# de las dimensiones de la imagen de partida
prediccionesTodo = np.resize(prediccionesTodo, bandas[:, :, 0].shape)
# Se crea una figura
plt.figure(1, figsize=(25,25))
# Se muestran las predicciones
im = plt.imshow(prediccionesTodo)
# Se ponen los patch creados al principio de la práctica como leyenda de la figura
plt.legend(handles=patches, bbox_to_anchor=(1.05, 1), loc=2, borderaxespad=0.)
# Se muestra la figura
plt.show()
```

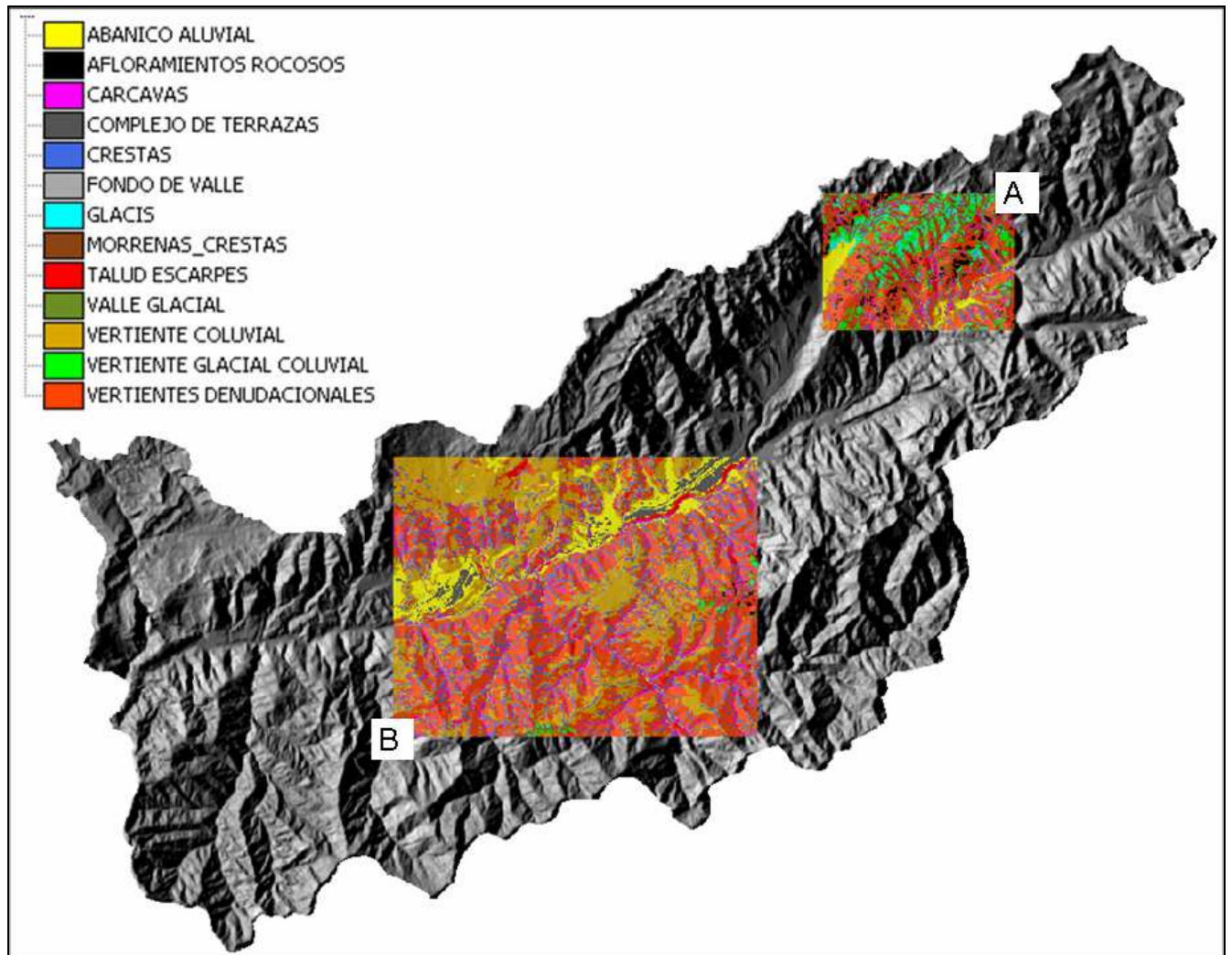


```
In [87]: #Guardar la imagen.
imageio.imwrite('Prediccion_geomorfometria_vertiente.tif', prediccionesTodo)
```

**Anexo 2.** Áreas de validación y entrenamiento creadas a partir interpretación de fotografías aéreas e imágenes sentinel. Las zonas de análisis páramo (A) y vertiente (B) se proyectan sobre una una composición de sombras del ALOS PALSAR.

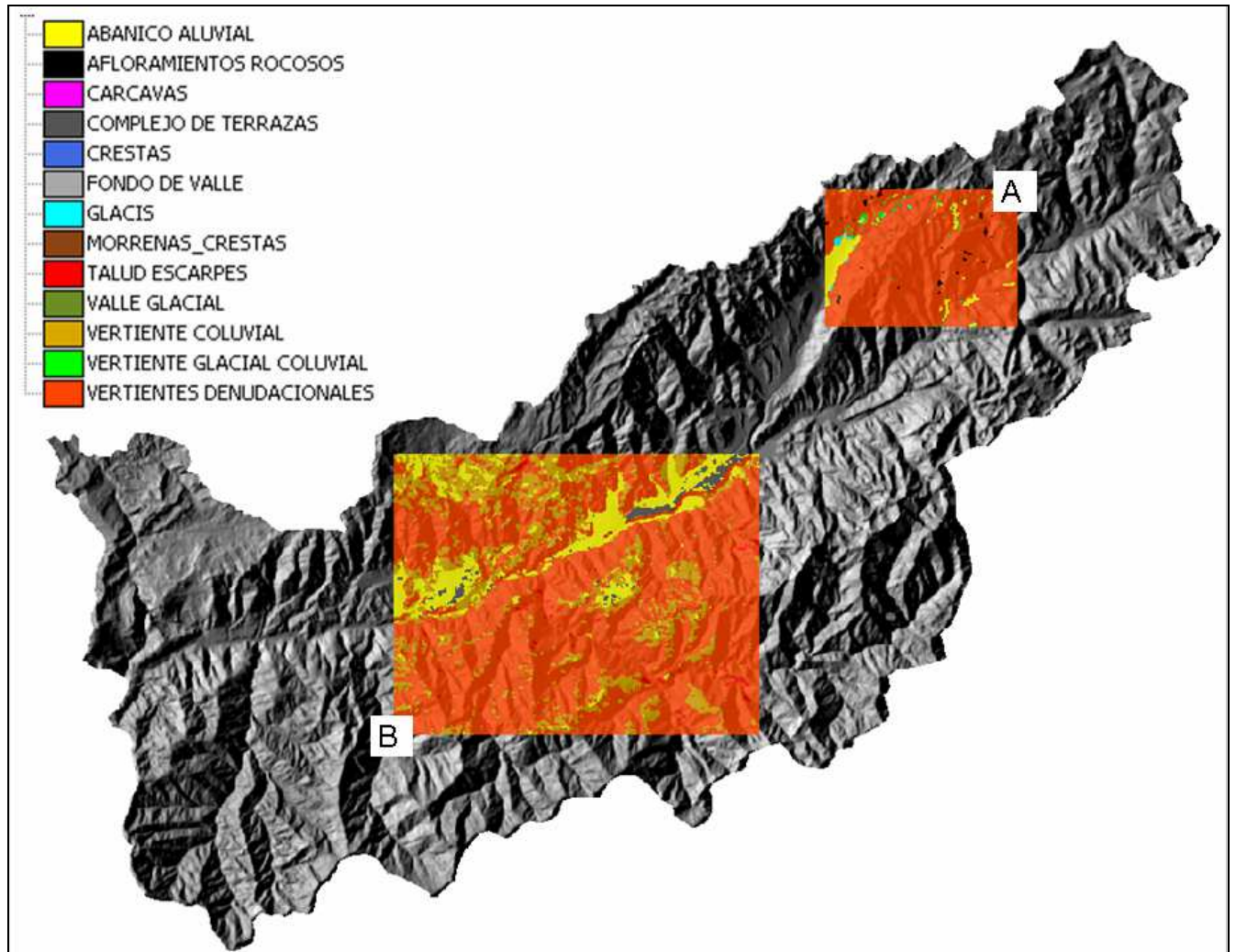


**Anexo 3.** Clasificación geomorfométrica a partir del MDE SRTM (90m), aplicando el AE. Las zonas de análisis páramo (A) y vertiente (B) se proyectan sobre una composición de sombras del SRTM.

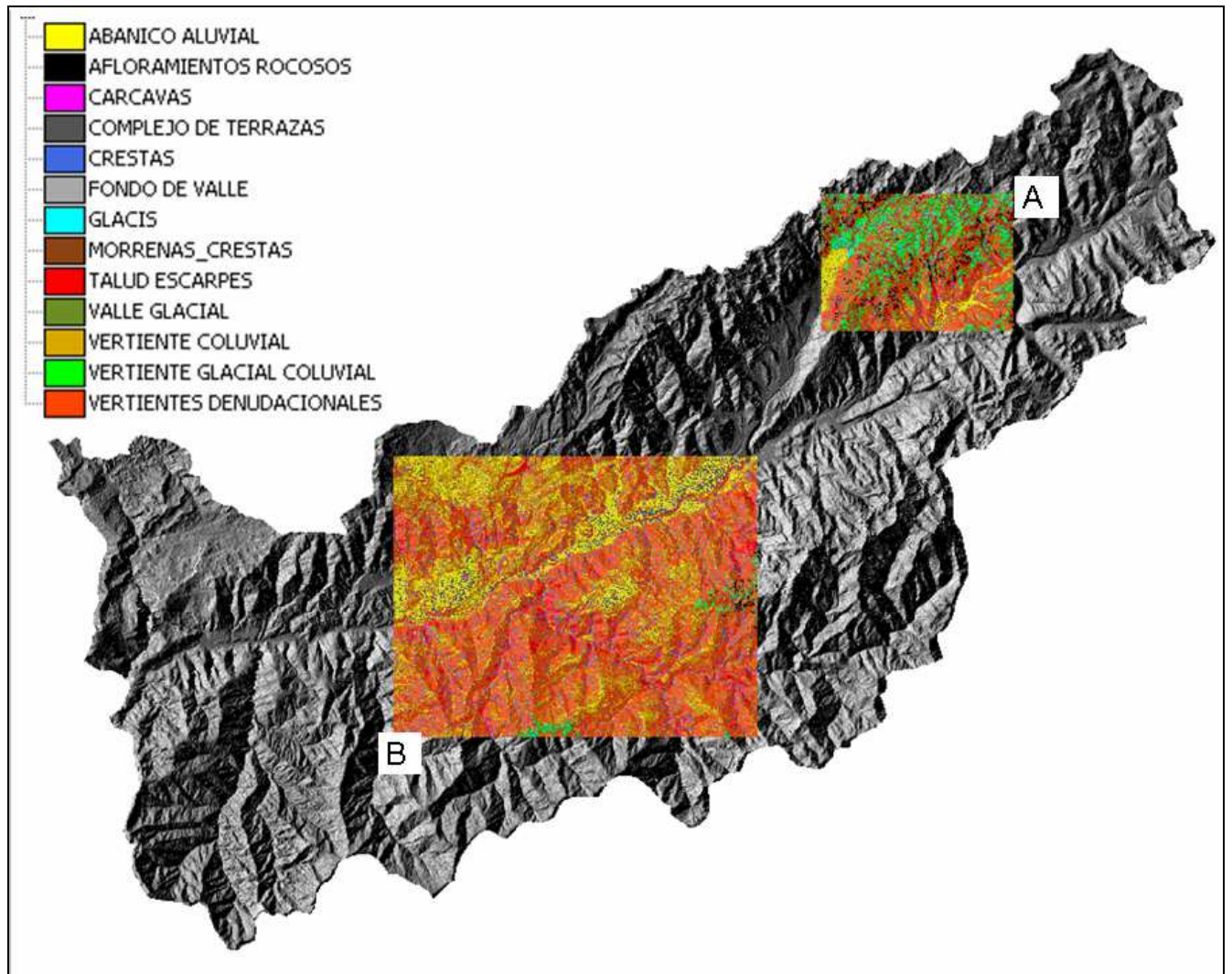




**Anexo 4.** Clasificación geomorfológica a partir del MDE SRTM (90m), aplicando el KNN. Las zonas de análisis páramo (A) y vertiente (B) se proyectan sobre una composición de sombras del SRTM.



**Anexo 5.** Clasificación geomorfométrica a partir del MDE ALOS PALSAR (12,5m), aplicando el AE. Las zonas de análisis páramo (A) y vertiente (B) se proyectan sobre una composición de sombras del ALOS PALSAR.



**Anexo 6.** Clasificación geomorfométrica a partir del MDE ALOS PALSAR (12,5m), aplicando el KNN. Las zonas de análisis páramo (A) y vertiente (B) se proyectan sobre una composición de sombras del MDE ALOS PALSAR.

