

Procesamiento productos ráster

Automatización para INVÍAS

Introducción

Este documento relaciona los diferentes orígenes o fuentes para el proceso de automatización con el fin de dar seguimiento e identificar pasos a seguir para producir y mantener, de manera repetible un portafolio de capas ráster estandarizadas a 100 x 100 m. La solución integra fuentes oficiales y abiertas como Colombia en Mapas (IGAC/INVÍAS), SIAC-IDEAM, Global Forest Change, Copernicus DEM y Sentinel-2 en GEE— y las normaliza en una rejilla única con píxeles cuadrados alineados,

El resultado es un conjunto de productos homogéneos y comparables entre si como distancias a vías, cuerpos de agua, pérdida de cobertura arbórea y bosques; mapas categóricos de cobertura y vocación de uso; pendientes; bandas e índices.

Objetivos

Estandarizar y presentar el flujo de actividades en la ingesta de fuentes primarias y secundarias y su normalización.

1. Distancia a Vías

La red vial se obtuvo desde el portal Colombia en Mapas del IGAC (<http://www.colombiaenmapas.gov.co>) consumiendo el servicio REST de INVÍAS MapaCarreteras/RedVial/FeatureServer. El insumo corresponde a la Base de Datos Vectorial Básica a escala 1:100.000 (edición 2022; insumos más recientes 2017) y contiene la red de carreteras del país hasta nivel 6. Es un conjunto vectorial tipo polígono/línea publicado en MAGNA SIRGAS geográfico (EPSG:4686), destinado a cartografía temática y análisis de movilidad e infraestructura.

Para preparar el insumo y llevarlo a ráster de distancia se trabajó en QGIS. Primero se descargó y guardó localmente la capa de Red Vial del servicio REST, se recortó al límite nacional y se ejecutó Corregir geometrías para eliminar errores topológicos que impiden la rasterización y los cálculos de proximidad.

El dataset se reproyectó a EPSG:9377 para que todas las medidas se expresen en metros. Luego se homogenizó la grilla a 100 x 100 m exactos con warp y la opción de alineación a píxeles objetivo (equivalente a -tap), tomando como referencia la plantilla de 100 m del proyecto.

Se generaron dos resultados de ráster de entrada para la distancia.

1. La primera mantuvo la categoría de vía (ID de clasificación) convirtiendo el vector a ráster con Rasterizar (vector a ráster), resolución 100 m, remuestreo Nearest y NoData = -9.9999 para exterior; esta salida conserva el código de clase en cada celda y sirvió para la distancia “por categoría”. (Resultado en la carpeta adjunta)
2. La segunda fue un ráster binario (vías=1, fondo=0) producido con la misma herramienta pero grabando valor fijo 1 a las líneas y 0 al resto, también a 100 m.

Con ambos insumos se calculó la distancia euclidiana mediante Proximidad (distancia ráster) de GDAL, unidades “coordenadas georreferenciadas” para obtener metros reales, remuestreo Nearest y preservando el NoData acordado.

Se emitieron cuatro productos operativos según la combinación de modo (categorizado / no categorizado) y manejo de fondo:

- dist_road_cat_0_100m y dist_road_0_100m con fondo 0.
- dist_road_cat_100m y dist_road_100m con NoData = -9.9999 para preservar ausencia de dato fuera del AOI.

Todas las salidas son TIF en EPSG:9377, 100 m de resolución, píxeles cuadrados y alineados.

2. Cuerpos de agua (IGAC) — Distancia a cuerpos de agua

Los cuerpos de agua se obtuvieron desde el portal Colombia en Mapas del IGAC (<http://www.colombiainmapas.gov.co>), usando los servicios REST de la Base de Datos Vectorial Básica 1:100.000 (edición 2022; insumos más recientes 2017).

El conjunto incluye los temas de Canal doble/sencillo, Drenaje doble/sencillo, Línea costera, Otros cuerpos de agua y los polígonos de Ciénaga, Embalse, Humedal y Laguna.

Tras descargar cada capa, se añadió un campo source para trazar el origen, se recortó al límite nacional y se ejecutó Corregir geometrías para eliminar problemas topológicos.

Todo el insumo se reproyectó a EPSG:9377 y se alineó a la malla de 100 x 100 m con Warp y la alineación a píxeles; con esto las medidas de distancia quedan en metros y las salidas son compatibles con los productos anteriores.

Dado que la hidrografía combina entidades lineales y superficiales, primero se fusionaron por separado todas las líneas (canales, drenajes y línea costera) y todos

los polígonos (lagunas, cienagas, humedales, embalses), generando dos capas únicas.

Luego se rasterizó cada una a 100 m, remuestreo Nearest, creando dos variantes por convención de fondo: sin dato como 0 y sin dato como -9.9999. Las salidas intermedias fueron:

- water_l_0_100m y water_l_100m (líneas), y water_pg_0_100m y water_pg_100m (polígonos).

A continuación se combinó cada par, las de líneas y las de polígonos mediante la Calculadora ráster usando `max("water_line_bin@1","water_poly_bin@1")`, obteniendo un ráster binario único de agua (water_0_100m con fondo 0 y water_100m con NoData=-9.9999).

Sobre cada uno se calculó la distancia con Proximidad (distancia ráster) de GDAL, unidades georreferenciadas, remuestreo Nearest y conservando la misma convención de NoData; los resultados son los mapas de Distancia a cuerpos de agua en metros, alineados a 100 m, en EPSG:9377, comprimidos como TIF.

3. Pérdida de cobertura arborea

La información de pérdida de cobertura arbórea proviene del conjunto Global Forest Change 2000–2023 (Hansen/UMD), distribuido por Global Forest Watch.

Es un ráster anual de “lossyear” en formato GeoTIFF de 8 bits, organizado en teselas de 10°x10° referidas a WGS 84 geográfico (EPSG:4326); cada píxel registra el año en que un píxel pasó de condición bosque a no bosque.

Para generar el insumo corporativo de Distancia a pérdida de cobertura arbórea en 100 m, se descargaron las nueve teselas que cubren Colombia (lossyear_00N_070W a 20N_090W) desde el portal público de Google Cloud Storage.

En QGIS se construyó primero vrt en el SRC original para tener una sola capa continua; a continuación se reproyectó y alineó a EPSG:9377 con tamaño de celda 100 x 100 m.

Con el ráster ya en metros se binarizó la señal de pérdida con la Calculadora ráster, usando la siguiente expresión

`"lossyear_9377_100m@1" > 0) * 1` para marcar 1 si hubo pérdida en cualquier año y 0, no hubo.

se emitieron dos variantes por convención de fondo: thrloss_0_100m_bin_aligned (NoData=0) y thrloss_100m_bin_aligned (NoData=-9.9999), ambas TIF.

Finalmente se calculó la distancia en metros desde cada celda hasta el píxel más cercano con pérdida utilizando Proximidad de GDAL, conservando la malla, el SRC y la convención de NoData; los productos resultantes son dist_thrloss (fondo -9.9999) y dist_0_thrloss (fondo 0). Con este flujo se garantiza que el indicador quede homogéneo con el resto de capas del proyecto.

4. Cambio en la superficie cubierta por Bosque Natural 2023-2024

La capa Bosque No bosque proviene del SIAC – IDEAM, a partir del producto Cambio en la superficie cubierta por Bosque Natural 2023–2024, accesible en el visualizador institucional (sección de Datos Abiertos).

El insumo se publicó como ráster ya preprocesado por la entidad y se descargó para su homologación con el resto de capas. El objetivo fue generar el subproducto Distancia a bosques en una malla nacional de 100 × 100 m y en EPSG:9377.

En QGIS se reclasificó el ráster de origen para obtener un binario “bosque/nobosque”. Se consideró Bosque Estable (código 1) y Regeneración (código 4) como bosque = 1; todas las demás clases se asignaron a no bosque = 0 usando la Calculadora ráster.

Se alineó la grilla y se reproyectó al límite oficial de Colombia, preservando píxel cuadrado de 100 m. Se emitieron dos variantes por convención de fondo: forest_bin_0_100m.tif con NoData = 0 y forest_bin_100m.tif con NoData = -9.9999, ambas en formato TIF.

Sobre el binario alineado se calculó la distancia euclidiana en metros mediante la herramienta Proximidad de GDAL, manteniendo SRC, tamaño de celda y extensión del recorte.

5. Coberturas vegetales

La capa Coberturas vegetales (mapa de ecosistemas continentales) del IDEAM/SIAC se obtuvo como shapefile temático en EPSG:4686 desde el portal de ecosistemas (enlace de descarga oficial).

Es un inventario poligonal a escala 1:100.000 que delimita biomas y ecosistemas continentales y trae atributos descriptivos como ecos_sinte, ecos_gener, gran_bioma, clima, relieve, suelos, entre otros. Para integrarla al modelo raster nacional se generó el subproducto lulc_landcover con un identificador numérico por clase y malla de 100 m.

En QGIS primero se escogió el campo de clasificación y se disolvieron los polígonos por ese atributo para eliminar fronteras internas; luego la capa se reproyectó a EPSG:9377 para trabajar en metros

Finalmente se rasterizó con GDAL a 100 × 100 m, alineando a la plantilla del proyecto, usando Nearest neighbour para preservar categorías.

6. Clasificación de las Tierras por su Vocación de Uso

El Mapa de Clasificación de Tierras por Vocación de Uso 1:100.000 del IGAC se descargó desde el geoportal de datos abiertos de agrología en formato Shapefile. La capa describe, a escala 1:100.000, la aptitud de uso del suelo del país en dos niveles jerárquicos: una vocación general y, como subdivisión, los usos principales recomendados.

Para integrarla al esquema raster nacional se tomó el shapefile oficial y, dentro de QGIS, se seleccionó el atributo de clasificación de vocación y uso.

Se eliminaron fronteras internas disolviendo por esa clase y se creó un identificador entero (lulc_id) con la calculadora de campos para garantizar un código único por categoría.

La capa se reproyectó al sistema métrico EPSG:9377 y se verificó la coincidencia espacial con el recorte nacional.

Seguidamente se rasterizó con GDAL empleando lulc_id como valor de celda, remuestreo Nearest para preservar categorías, malla objetivo 100 × 100 m y alineación estricta a la plantilla del límite Colombia con las mismas características. El resultado se guardó como TIF bajo el nombre lulc_landuse.tif.

7. Pendientes

El modelo de pendientes proviene del Copernicus DEM GLO-30 (colección COPERNICUS/DEM/GLO30), un MDE global con resolución nominal de ~30 m derivado del programa Copernicus y distribuido como mosaicos continuos.

Para garantizar compatibilidad con el resto de insumos, el procesamiento se ejecutó en GEE sobre el AOI de Colombia (Plantilla ya usada con los parámetros anteriores) cargado desde el asset projects/ee-jasminprez1/assets/limitecolombia, centrando el mapa y recortando todos los cálculos a ese límite.

Se configuraron los parámetros de exportación para obtener un GeoTIFF a 100 m en EPSG:4326 y con NoData = -9.9999. La colección del DEM se filtró por el AOI, se seleccionó la banda DEM, se mosaicó y recortó al país; sobre ese mosaico se calculó la pendiente en grados mediante ee.Terrain.slope, renombrando la banda a sm_slope.

Antes de exportar se aplicó `unmask(-9.9999)` para asegurar que cualquier área fuera de cobertura del DEM o del AOI se codifique como NoData con ese mismo valor.

El resultado se exportó con nombre `sm_slope_100m.tif`, quedando un raster de pendientes continuo, alineado a 100 m.

A continuación se relaciona el script

```
var AOI_ASSET = 'projects/ee-jasminprez1/assets/limitecolombia';
var EXPORT_CRS = 'EPSG:4326';
var EXPORT_SCL = 100;
var FOLDER = 'GEE_DEM';
var NODATA = -9.9999;

var aoi = ee.FeatureCollection(AOI_ASSET).geometry();
Map.centerObject(aoi, 6);
Map.addLayer(aoi, {color:'red'}, 'AOI');

var dem = ee.ImageCollection('COPERNICUS/DEM/GLO30')
  .filterBounds(aoi)
  .select('DEM')
  .mosaic()
  .clip(aoi);
var slope_deg = ee.Terrain.slope(dem).rename('sm_slope');
var slope_pct =
slope_deg.tan().multiply(Math.PI/180).multiply(100).rename('sm_slope_pct'
);
var slope_out = slope_deg.unmask(NODATA);
Map.addLayer(slope_deg, {min:0, max:60,
palette:['white','tan','brown','black']}, 'Slope (deg)');
Export.image.toDrive({
  image: slope_out,
  description: 'sm_slope_100m',
  fileNamePrefix: 'sm_slope_100m',
  folder: FOLDER,
  region: aoi,
  crs: EXPORT_CRS,
  scale: EXPORT_SCL,
  maxPixels: 1e13
});
```

8. Áreas protegidas

La capa de Áreas protegidas de Colombia se consumió desde Colombia en Mapas como servicio REST (URL:

mapas.parquesnacionales.gov.co/arcgis/rest/services/pnn/runap/MapServer), escala de referencia 1:500.000 y fecha de elaboración 01/01/2020.

El insumo delimita polígonos de áreas protegidas registrados en el RUNAP y fue estandarizado para análisis raster a 100 m.

En QGIS se descargó el servicio, se filtró la subcapa de Zonas Núcleo y se reproyectó a metros (EPSG:9377). Sobre el vector reproyectado se creó un identificador entero de clase (campo runap_id) y se rasterizó con GDAL, así como en los pasos anteriores, bajo los mismos parámetros. A partir de ese raster categórico se generó un binario interno de presencia (runap_id > 0) y se calculó la distancia con Proximidad, unidades georreferenciadas y valor NoData = -9.9999 para el exterior; el resultado se guardó como dist_RUNAP.tif,

9. Imágenes sentinel 2

El insumo se generó directamente en Google Earth Engine a partir de la colección COPERNICUS/S2_SR_HARMONIZED. Se delimitó el AOI con el asset projects/ee-jasminprez1/assets/limitecolombia y el periodo de análisis comprendido entre 01-ene-2024 y 31-dic-2024.

Para asegurar calidad radiométrica se filtró por CLOUDY_PIXEL_PERCENTAGE ≤ 10 y se aplicó una máscara de nubes usando los bits 10 y 11 del QA60 iguales a cero.

Sobre cada escena enmascarada se estandarizaron las bandas ópticas nativas (B1...B12) y se renombraron a la convención B01, B02, ..., B12.. A partir de ese mosaico se calcularon los índices espectrales requeridos sobre las bandas Sentinel-2 ya escaladas (entre otros: AVI, BSI, EVI, GCI, GNDVI, MSI, NBRI, NDMI, NDVI, SAVI, ARVI, NPCRI, SI, SIPI, NDGI, NDSI, NDWI-McFeeters y NDWI-Gao).

Las salidas se exportaron desde GEE como GeoTIFF individuales, una por banda y una por índice, con CRS = EPSG:3857 (metros) y resolución de 100 m (scale: 100).

Ya en QGIS, todas las capas exportadas se reproyectaron y alinearon a la malla operativa EPSG:9377 mediante GDAL Warp forzando -tr 100 100 y -tap para que el origen de píxel coincida con la plantilla del límite de Colombia.

A continuación se detalla las formulas con sus respectivas formulas

Índice	Qué mide	Fórmula (Landsat)
AVI	Vegetación avanzada (vigor)	$AVI = (B5 * (1 - B4) * (B5 - B4))^{\frac{1}{3}}$
BSI	Suelo desnudo	$BSI = \frac{((B6 + B4) - (B5 + B2))}{((B6 + B4) + (B5 + B2))}$
EVI	Vegetación mejorada (corrige atmósfera/suelo)	$EVI = \frac{2.5 * (B5 - B4)}{(B5 + 6 * B4 - 7.5 * B2 + 1)}$

Índice	Qué mide	Fórmula (Landsat)
GCI	Clorofila (índice verde)	$GCI = \frac{B5}{B3} - 1$
GNDVI	Vegetación (verde) normalizada	$GNDVI = \frac{(B5 - B3)}{(B5 + B3)}$
MSI	Estrés hídrico	$MSI = \frac{B6}{B5}$
NBRI	Severidad de quema (NBR)	$NBRI = \frac{(B5 - B7)}{(B5 + B7)}$
NDMI	Humedad normalizada	$NDMI = \frac{(B5 - B6)}{(B5 + B6)}$
NDVI	Vegetación normalizada	$NDVI = \frac{(B5 - B4)}{(B5 + B4)}$
SAVI	Vegetación ajustada al suelo (L=0.5)	$SAVI = \frac{1.5 * (B5 - B4)}{(B5 + B4 + 0.5)}$
ARVI	Vegetación resistente a aerosoles	$ARVI = \frac{(B5 - 2 * B4 + B2)}{(B5 + 2 * B4 + B2)}$
NPCRI	Pigmento/clorofila normalizado	$NPCRI = \frac{(B4 - B2)}{(B4 + B2)}$
SI	Índice de sombra	$SI = ((1 - B2)(1 - B3)(1 - B4))^{\{\frac{1}{3}\}}$
SIPI	Carotenoides vs clorofila	$SIPI = \frac{(B5 - B2)}{(B5 - B4)}$
NDGI	Glaciar/nieve (verde vs rojo)	$NDGI = \frac{(B3 - B4)}{(B3 + B4)}$
NDSI	Nieve (verde vs SWIR1)	$NDSI = \frac{(B3 - B6)}{(B3 + B6)}$

Índice	Qué mide	Fórmula (Landsat)
NDWI (McFeeters)	Agua superficial (verde vs NIR)	$NDWI = \frac{(B3 - B5)}{(B3 + B5)}$
NDWI (Gao) (≡ NDMI)	Humedad (NIR vs SWIR1)	$NDWI = \frac{(B5 - B6)}{(B5 + B6)}$

Script relacionado:

```

var AOI_ASSET = 'projects/ee-jasminprez1/assets/limitecolombia';
var START    = '2024-01-01';
var END      = '2024-12-31';
var CLOUD_MAX = 10;
var EXPORT_CRS = 'EPSG:3857';
var EXPORT_SCL = 100;
var EPS       = 1e-6;

var aoi = ee.FeatureCollection(AOI_ASSET).geometry();
Map.centerObject(aoi, 6);
Map.addLayer(aoi, {color:'red'}, 'AOI');

var nativeNames =
['B1','B2','B3','B4','B5','B6','B7','B8','B8A','B9','B11','B12'];
var stdNames    =
['B01','B02','B03','B04','B05','B06','B07','B08','B8A','B09','B11','B12']
;

function maskS2sr(img){
  var qa = img.select('QA60');
  var mask = qa.bitwiseAnd(1<<10).eq(0)
              .and(qa.bitwiseAnd(1<<11).eq(0));

  var scaledRenamed = img.updateMask(mask)
    .select(nativeNames)
    .multiply(0.0001)
    .rename(stdNames);

  return scaledRenamed.copyProperties(img, img.propertyNames());
}

var s2 = ee.ImageCollection('COPERNICUS/S2_SR_HARMONIZED')
  .filterBounds(aoi)
  .filterDate(START, END)
  .filter(ee.Filter.lte('CLOUDY_PIXEL_PERCENTAGE', CLOUD_MAX))
  .map(maskS2sr);

print('Nº imágenes usadas:', s2.size());

```

```

var mosaic = s2.median().clip(aoi);
Map.addLayer(mosaic.select(['B04','B03','B02']), {min:0, max:0.3}, 'RGB
median');

var bandNames =
['B01','B02','B03','B04','B05','B06','B07','B08','B8A','B09','B11','B12']
;

function exportSingle(image, name, isBinary){
  var img = isBinary ? image.resample('nearest') :
image.resample('bilinear');
  Export.image.toDrive({
    image: img,
    description: name,
    fileNamePrefix: name,
    region: aoi,
    crs: EXPORT_CRS,
    scale: EXPORT_SCL,
    maxPixels: 1e13
  });
}

bandNames.forEach(function(bn){
  exportSingle(mosaic.select(bn).rename(bn), bn, false);
});

var B2  = mosaic.select('B02'); // Blue
var B3  = mosaic.select('B03'); // Green
var B4  = mosaic.select('B04'); // Red
var B8  = mosaic.select('B08'); // NIR
var B11 = mosaic.select('B11'); // SWIR1
var B12 = mosaic.select('B12'); // SWIR2

function safeDiv(num, den){ return num.divide(den.max(EPS)); }

var AVI  =
(B8.multiply(ee.Image(1).subtract(B4)).multiply(B8.subtract(B4)))
  .pow(ee.Image(1/3)).rename('i_avi');
var BSI  = safeDiv(B11.add(B4).subtract(B8.add(B2)),
  B11.add(B4).add(B8).add(B2)).rename('i_bsi');
var EVI  = safeDiv(B8.subtract(B4).multiply(2.5),

B8.add(B4.multiply(6)).subtract(B2.multiply(7.5)).add(1)).rename('i_evi')
;
var GCI  = B8.divide(B3.max(EPS)).subtract(1).rename('i_gci');
var GNDVI = safeDiv(B8.subtract(B3), B8.add(B3)).rename('i_gndvi');
var MSI  = B11.divide(B8.max(EPS)).rename('i_msi');
var NBRI  = safeDiv(B8.subtract(B12), B8.add(B12)).rename('i_nbri');
var NDMI  = safeDiv(B8.subtract(B11), B8.add(B11)).rename('i_ndmi');

```

```

var NDVI = safeDiv(B8.subtract(B4), B8.add(B4)).rename('i_ndvi');
var SAVI = safeDiv(B8.subtract(B4).multiply(1.5),
    B8.add(B4).add(0.5)).rename('i_savi');
var ARVI = safeDiv(B8.subtract(B4.multiply(2)).add(B2),
    B8.add(B4.multiply(2)).add(B2)).rename('i_arvi');
var NPCRI = safeDiv(B4.subtract(B2), B4.add(B2)).rename('i_npcri');
var SI = (ee.Image(1).subtract(B2))
    .multiply(ee.Image(1).subtract(B3))
    .multiply(ee.Image(1).subtract(B4))
    .pow(ee.Image(1/3)).rename('i_si');
var SIPI = safeDiv(B8.subtract(B2),
    B8.subtract(B4).max(EPS)).rename('i_sipi');
var NDGI = safeDiv(B3.subtract(B4), B3.add(B4)).rename('i_ndgi');
var NDSI = safeDiv(B3.subtract(B11), B3.add(B11)).rename('i_ndsi');
var NDWI_McF = safeDiv(B3.subtract(B8),
    B3.add(B8)).rename('i_ndwi_mcfeeters');
var NDWI_Gao = safeDiv(B8.subtract(B11),
    B8.add(B11)).rename('i_ndwi_gao');

AVI, BSI, EVI, GCI, GNDVI, MSI, NBRI, NDMI, NDVI, SAVI,
    ARVI, NPCRI, SI, SIPI, NDGI, NDSI, NDWI_McF, NDWI_Gao
].forEach(function(img){
    var nm = img.bandNames().get(0);
    exportSingle(img, ee.String(nm).getInfo(), false);
});

```