INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA

WAGENINGEN UNIVERSITY & RESEARCH

United Nations Statistics Division

UNSD - Wageningen University - INEGI

The INEGI Geospatial Data Cube

Background…. Big Challenges

- We have produced Natural Resources Maps – geospatial data for 50 years.
- National coverage in several scales.
- But the present needs call for:
- More spatial detail
- More often, more recent reference data

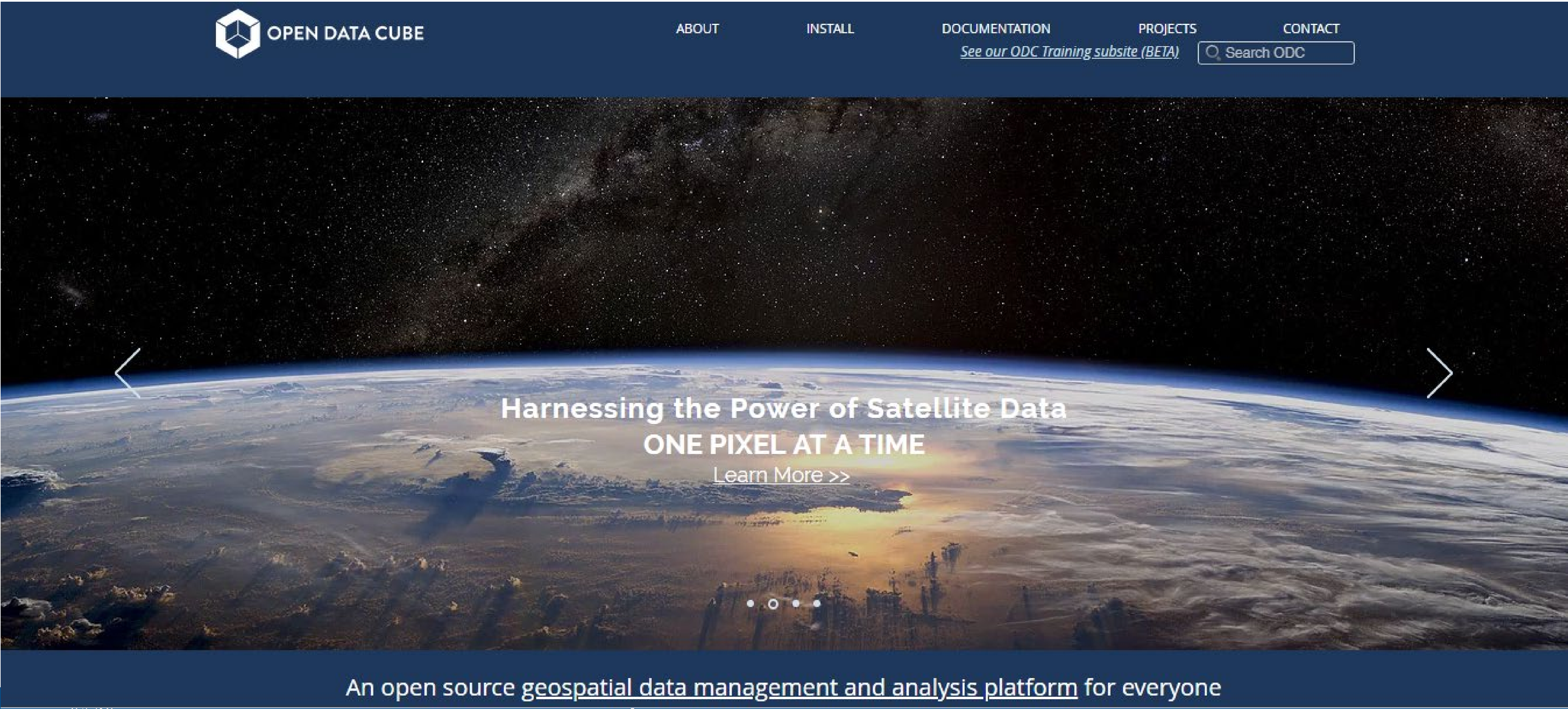
Background…. Big Challenges, big opportunities

- Massive availability of satellite images.
- Availability of many data layers – ancillary data.
- Geospatial
- Statistical (National Forest Inventory)
- New knowledge, technologies
- Big Data, Analytic tools

Background

- A **data cube** (or **datacube**) is a multidimensional array of values
- Typically, the term datacube is applied in contexts where these arrays are massively larger than the hosting computer's main memory (multi Terabyte to Petabytes)
- It fits well to massive time series of images

The Open Data Cube



Open Data Cube

- Software, tools:
- Massive storage and management of raster data
- Access to data
- Massive processing and analysys
- Products

Open Data Cube

- Open source software framework
- Geoscience Australia, CSIRO, others (GEOS)
- Massive Capacity to:
- Storage and management
- Processing and analysys at country and even continental scales
- Focus on: areas of interest – time
- On pixel at a time

ODC features

“Analysys Ready Data (ARD)”

Analysys oriented, more than data administration

Capacidad de realizar análisis masivos.

Cubo de datos geoespaciales

Pixel centered, not scene centered

.

All the images available can be used, regardless of cloud coverage.

Time series analysys:

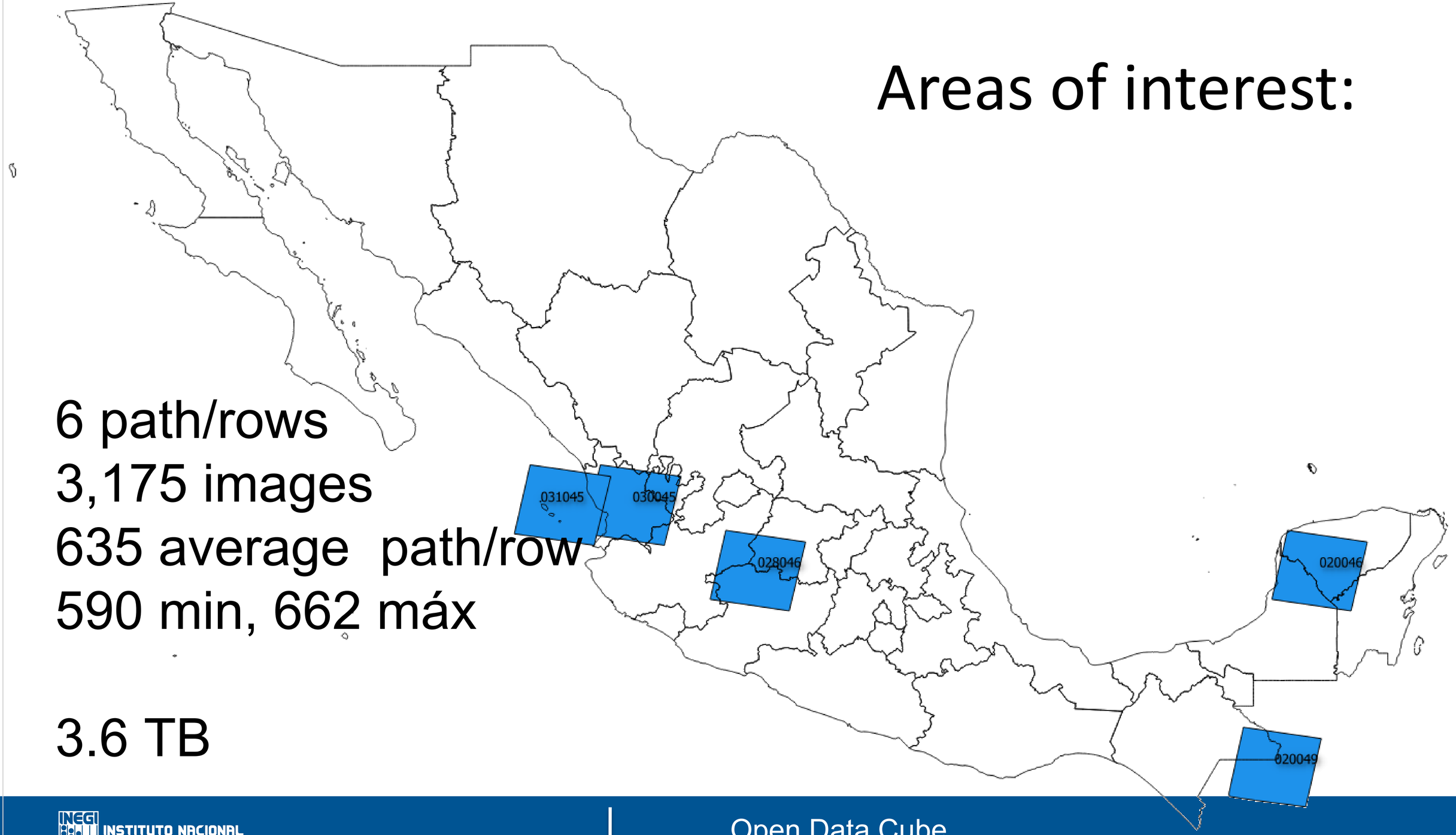
- Single day
- Weeks
- Seasons
- Years
- Decades

Training at Geoscience Australia (GA)

- ODC implementation
- Direct interaction with developers and experts at GAInteracción con especialistas y desarrolladores de GA
- LANDSAT imagery for Mexico provided by GA
- Pilot test in areas of interest

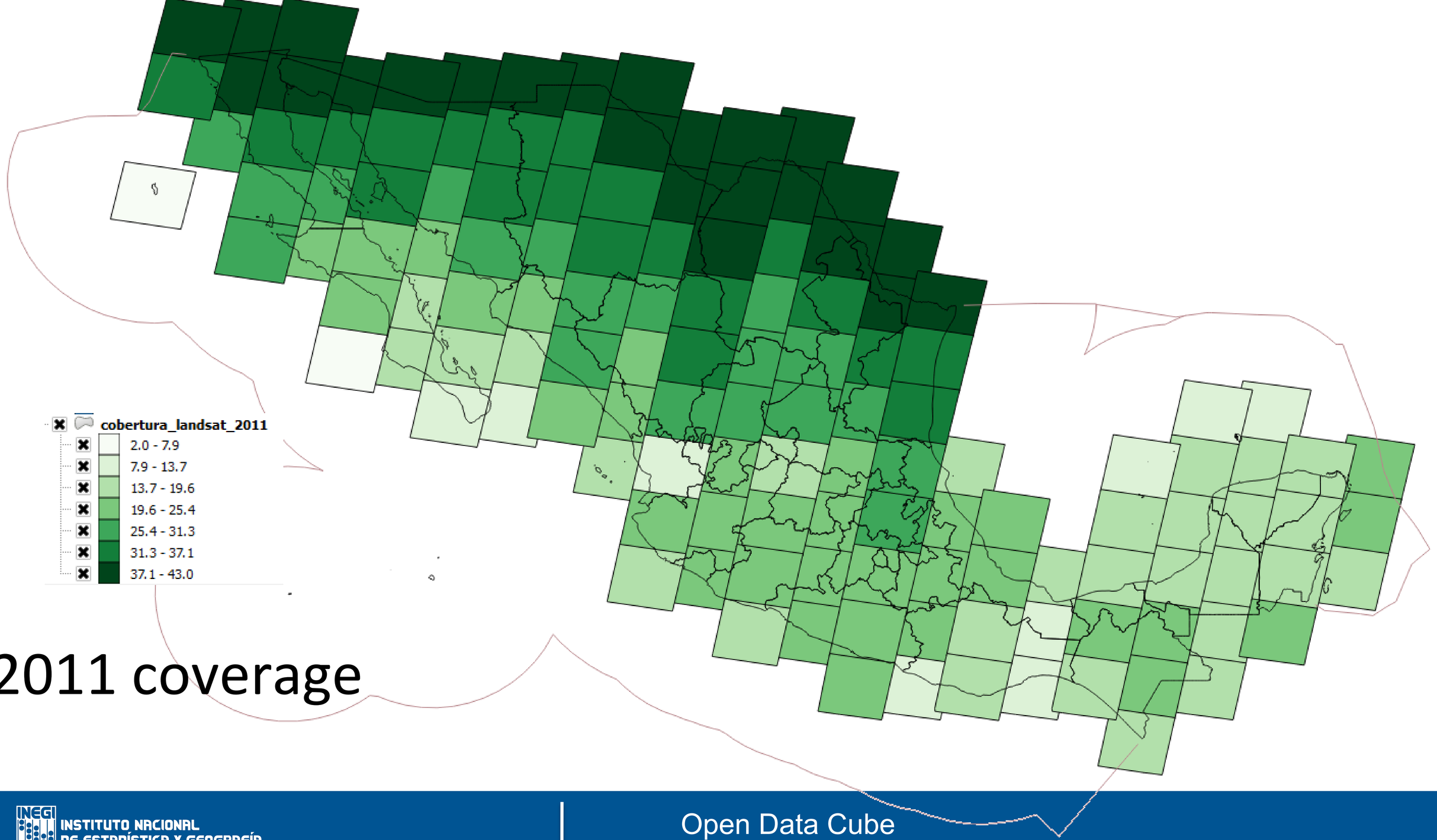
Scope: short, médium, long term

Areas of interest:

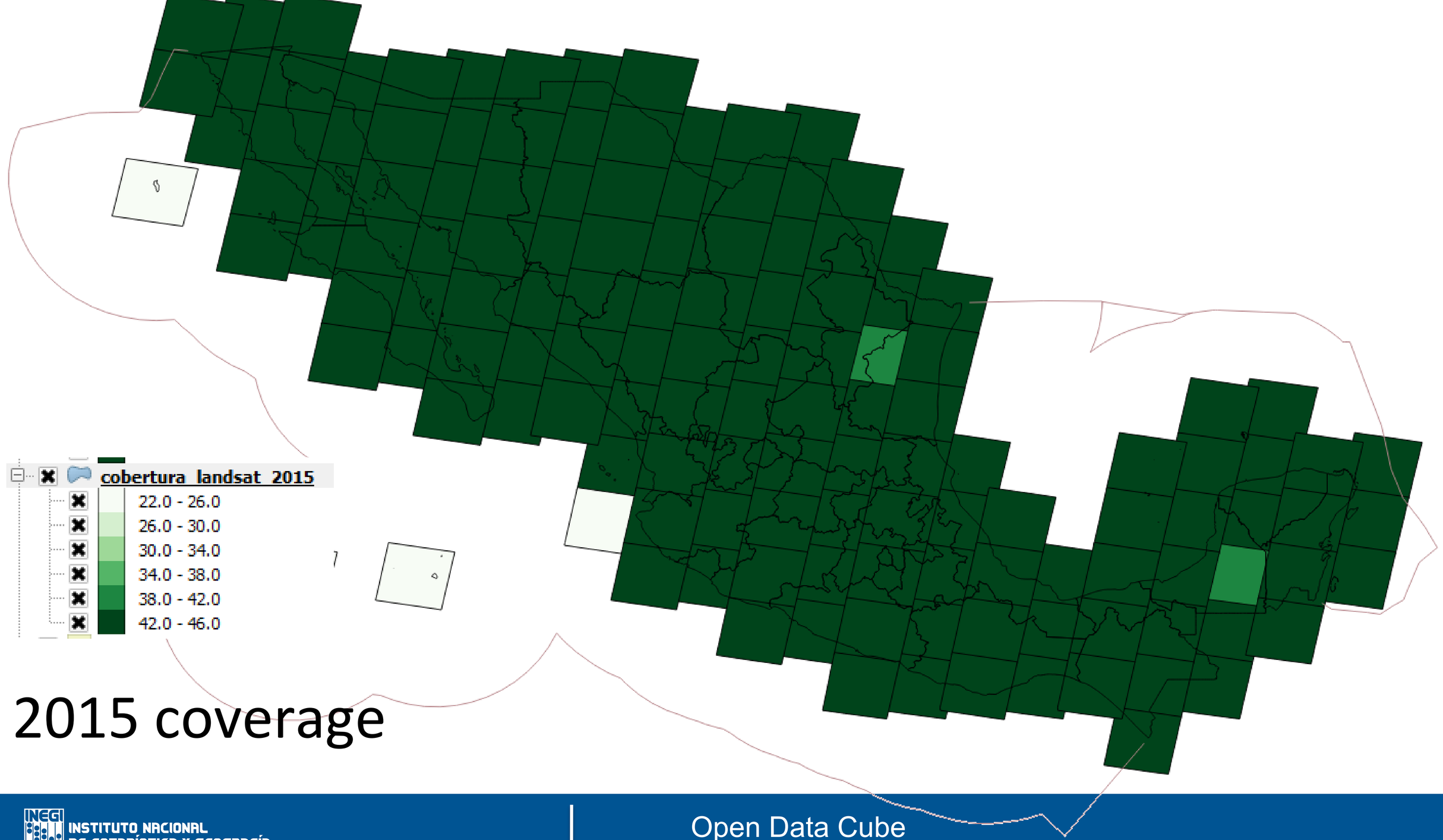


6 path/rows
3,175 images
635 average path/row
590 min, 662 máx

3.6 TB



2011 coverage



2015 coverage

Yearly coverages

2011

135 pathrows

3,517 images

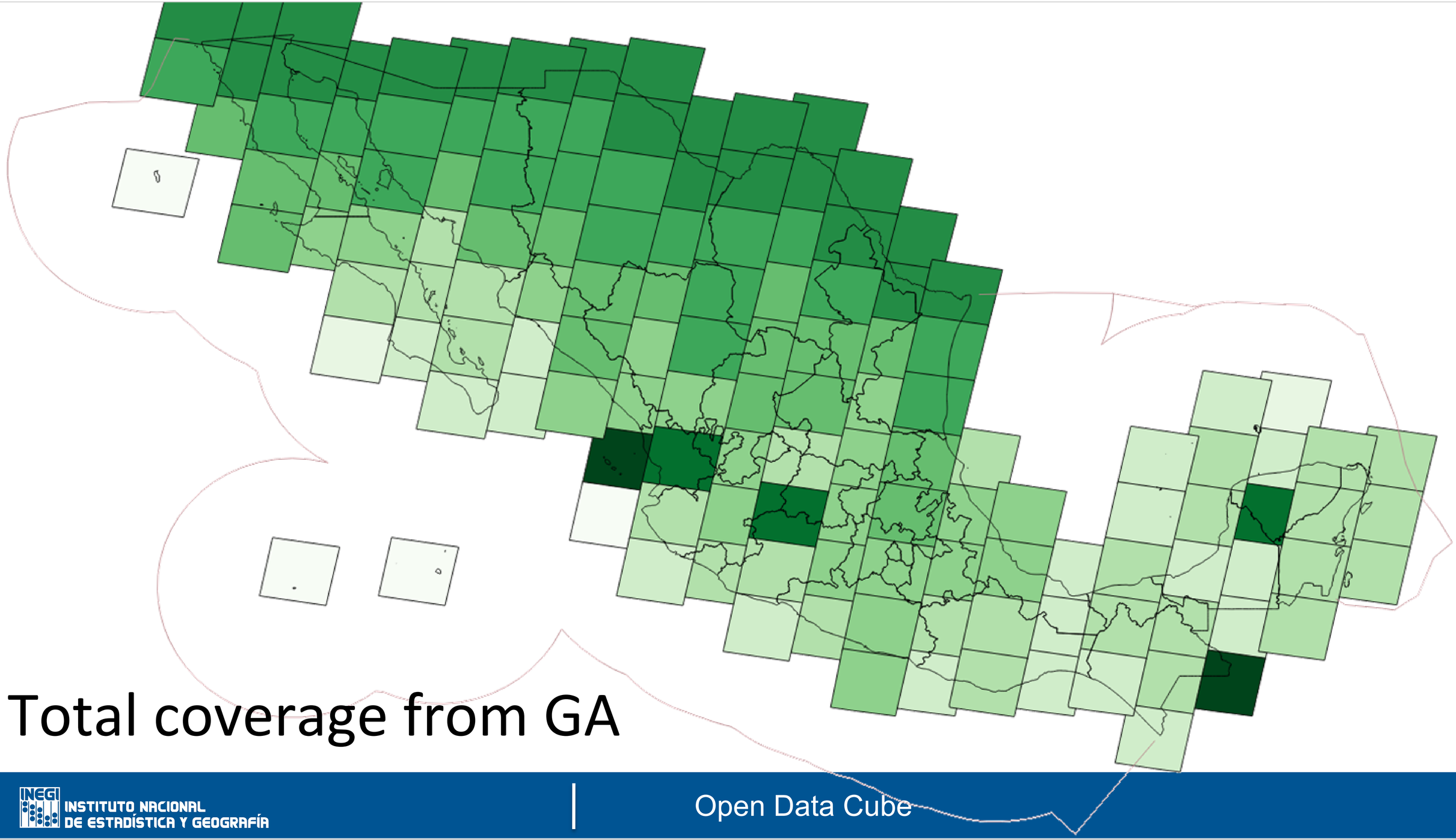
3.9 TB

2015

138 pathrows

6,074 images

7.5 TB



Total coverage from GA

Imagery from GA - USGS

138 path/rows

12,494 images

~ 1.2 GB/image

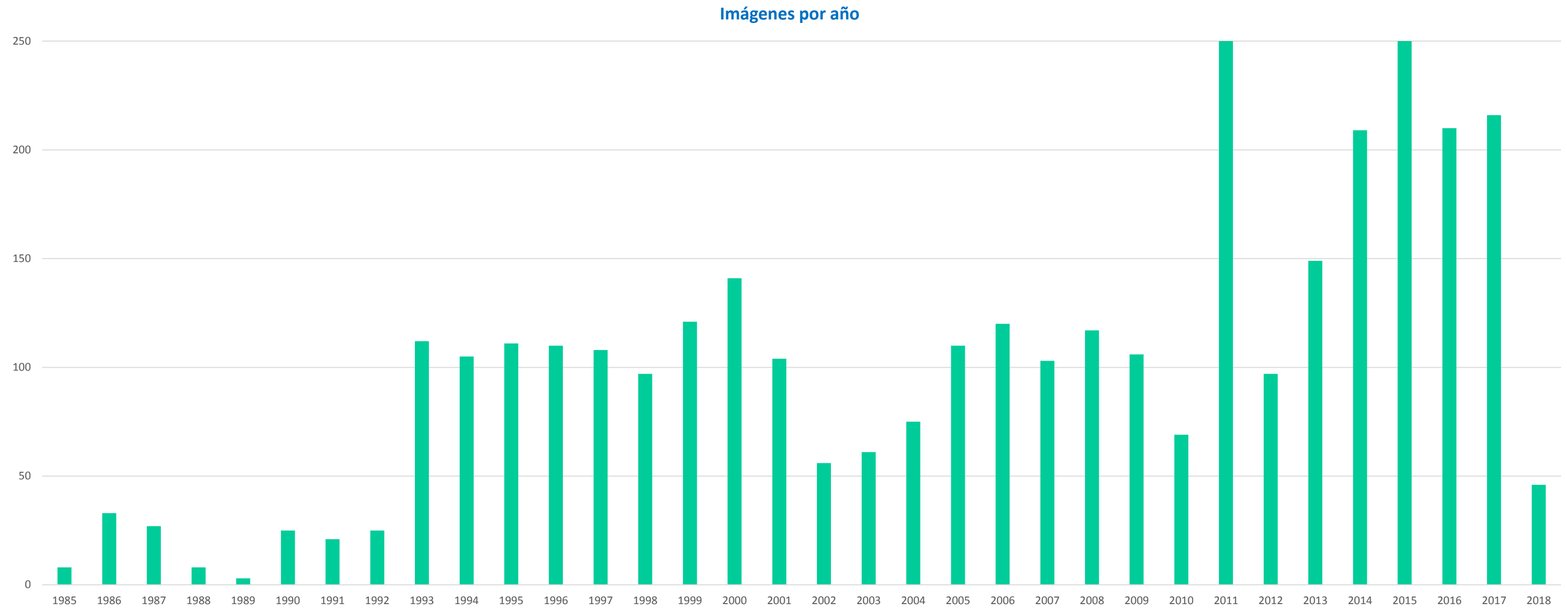
14.5 TB

Annual coverages:

- ~ 45 images path/row
- ~ 8.5 TB annual storage

Imagery per year and sensor
1985 – 2018

Año	Núm imgs	Vol Datos (MB)
1985	8	8,498,804
1986	33	34,844,028
1987	27	28,692,200
1988	8	8,618,144
1989	3	3,225,156
1990	25	26,760,040
1991	21	22,108,836
1992	25	26,598,124
1993	112	120,191,132
1994	105	112,626,092
1995	111	119,177,716
1996	110	118,753,352
1997	108	116,617,816
1998	97	104,620,596
1999	121	130,641,536
2000	141	152,510,352
2001	104	112,378,700
2002	56	60,634,312
2003	61	66,289,416
2004	75	81,778,148
2005	110	119,530,344
2006	120	130,534,620
2007	103	112,284,064
2008	117	127,778,956
2009	106	116,279,040
2010	69	75,982,440
*2011	3,517	3,861,488,257
2012	97	107,173,300
2013	149	178,250,816
2014	209	258,157,564
*2015	6,074	7,420,873,812
2016	210	259,096,484
2017	216	266,133,068
2018	46	55,628,416
TOTAL	12,494	14,544,755,681



Resuls

Vegetation

Wetlands and Coastline

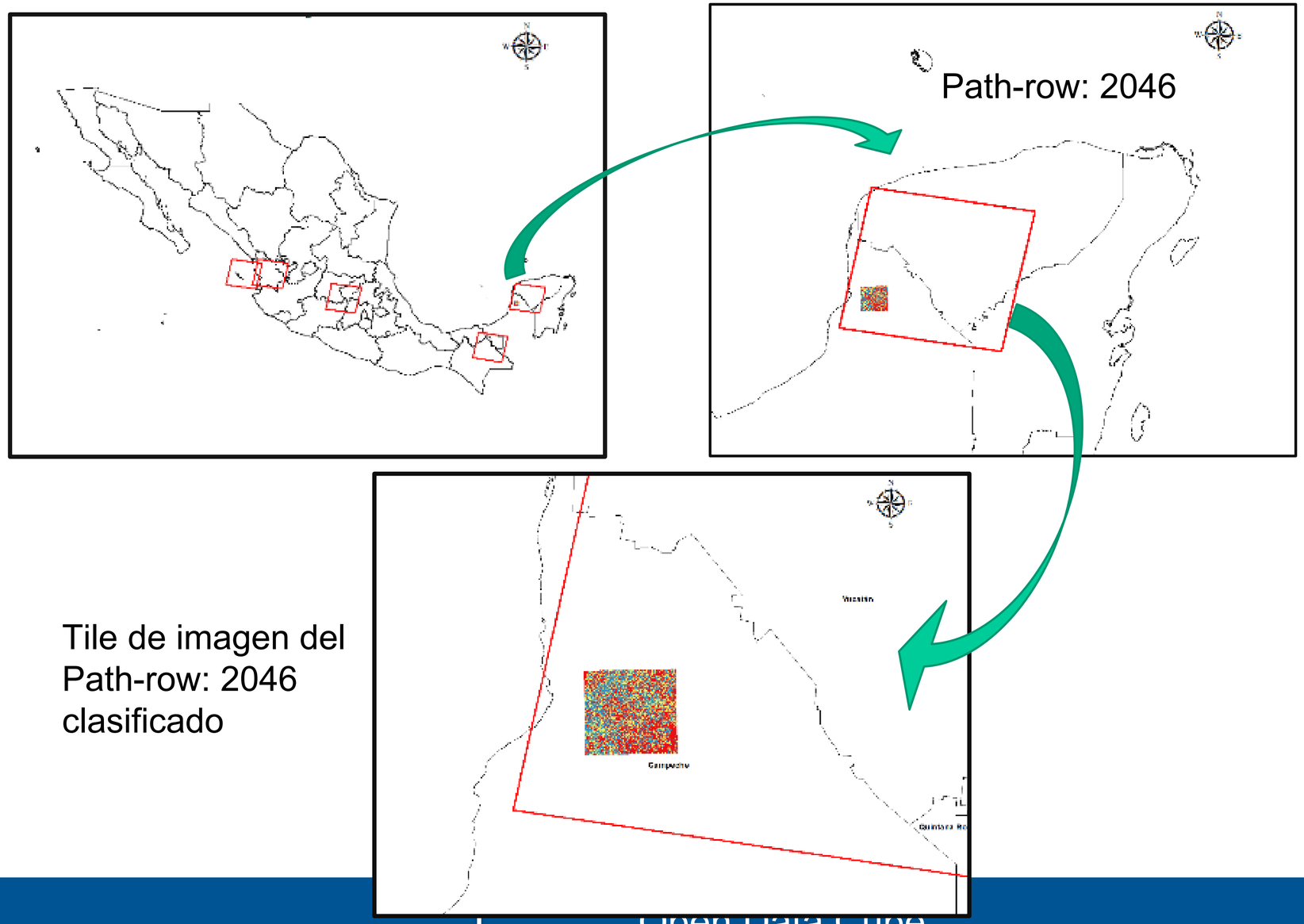
Urban Growth

Agriculture

Vegetation: Random Forest Classification

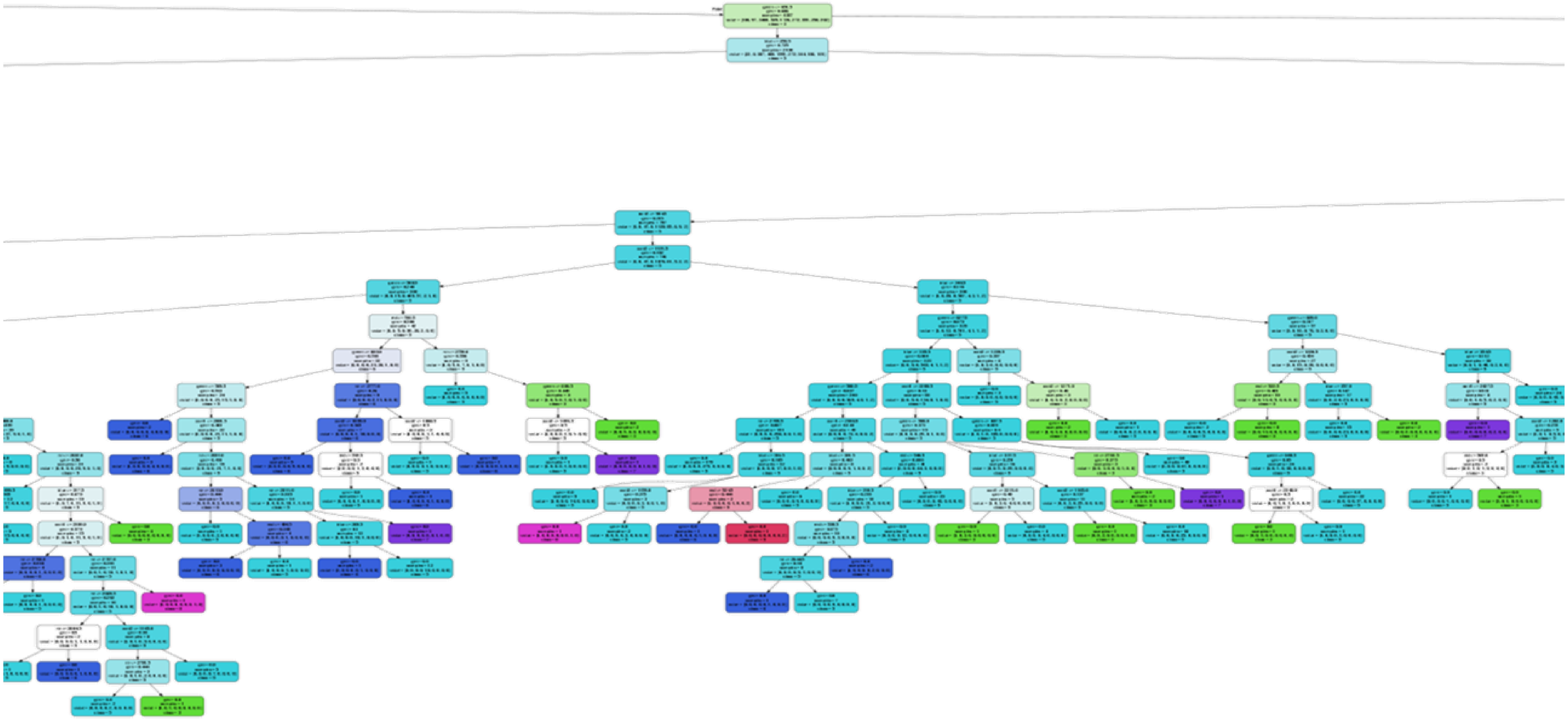
- Classification of a Geomedian mosaic in the Yucatan Peninsula
- 40 training polygons.
- Vegetation and land cover classes

Area of interest



Tile de imagen del Path-row: 2046 clasificado

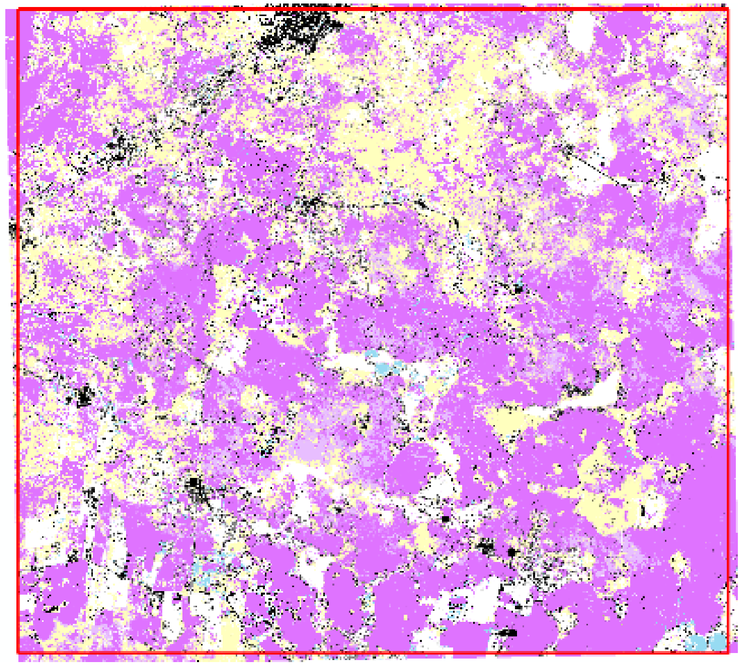
DECISION TREE

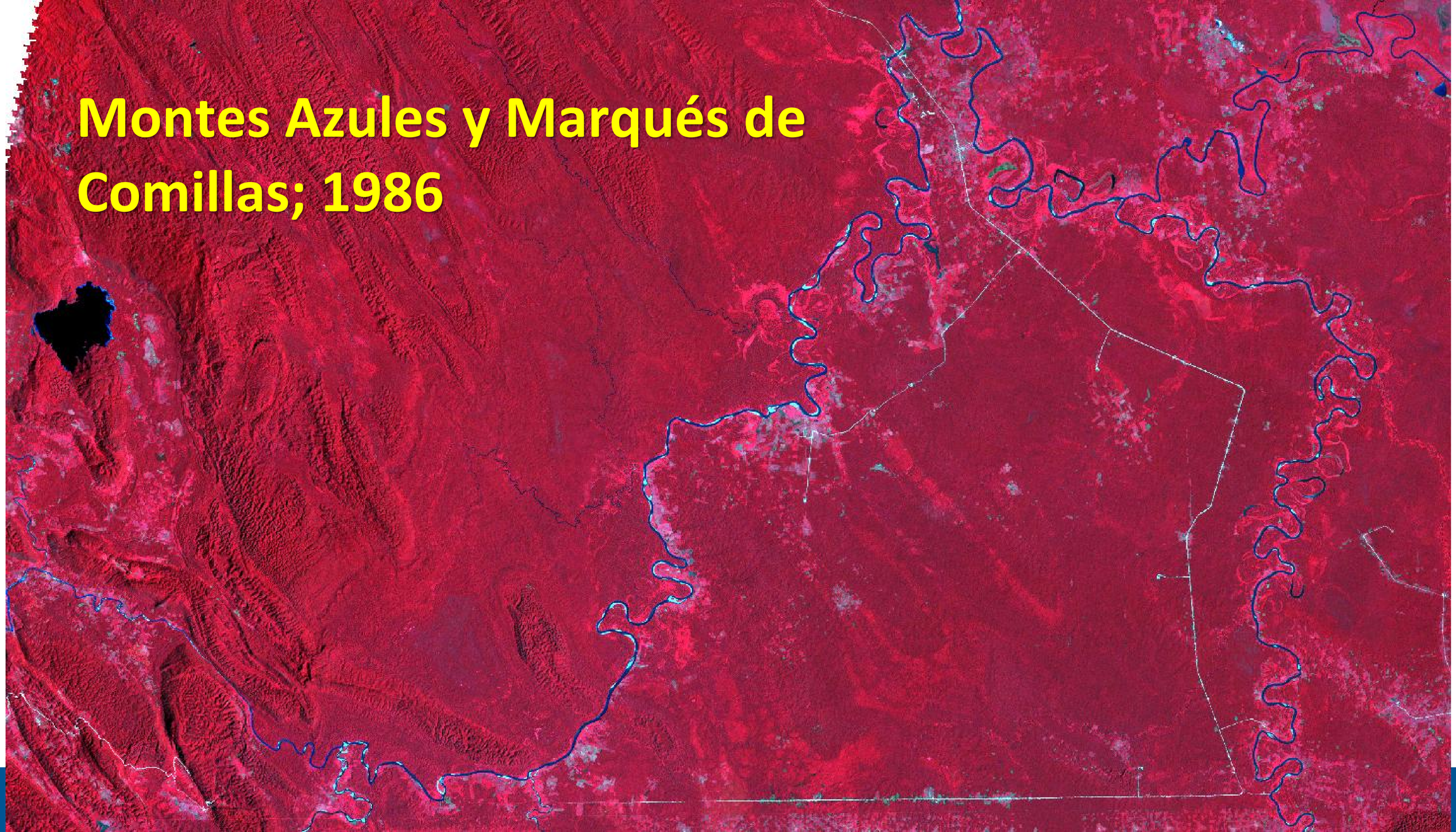


RESULTS

GEOMEDIAN

CLASSIFIED IMAGE




Montes Azules y Marqués de Comillas; 1986


Montes Azules y Marqués de Comillas; 2017

WATER, WETLANDS, COASTLINE

Pacific coastal plain in the state of Nayarit:

Marismas Nacionales, Santiago River, Tepic City, Western Sierra Madre

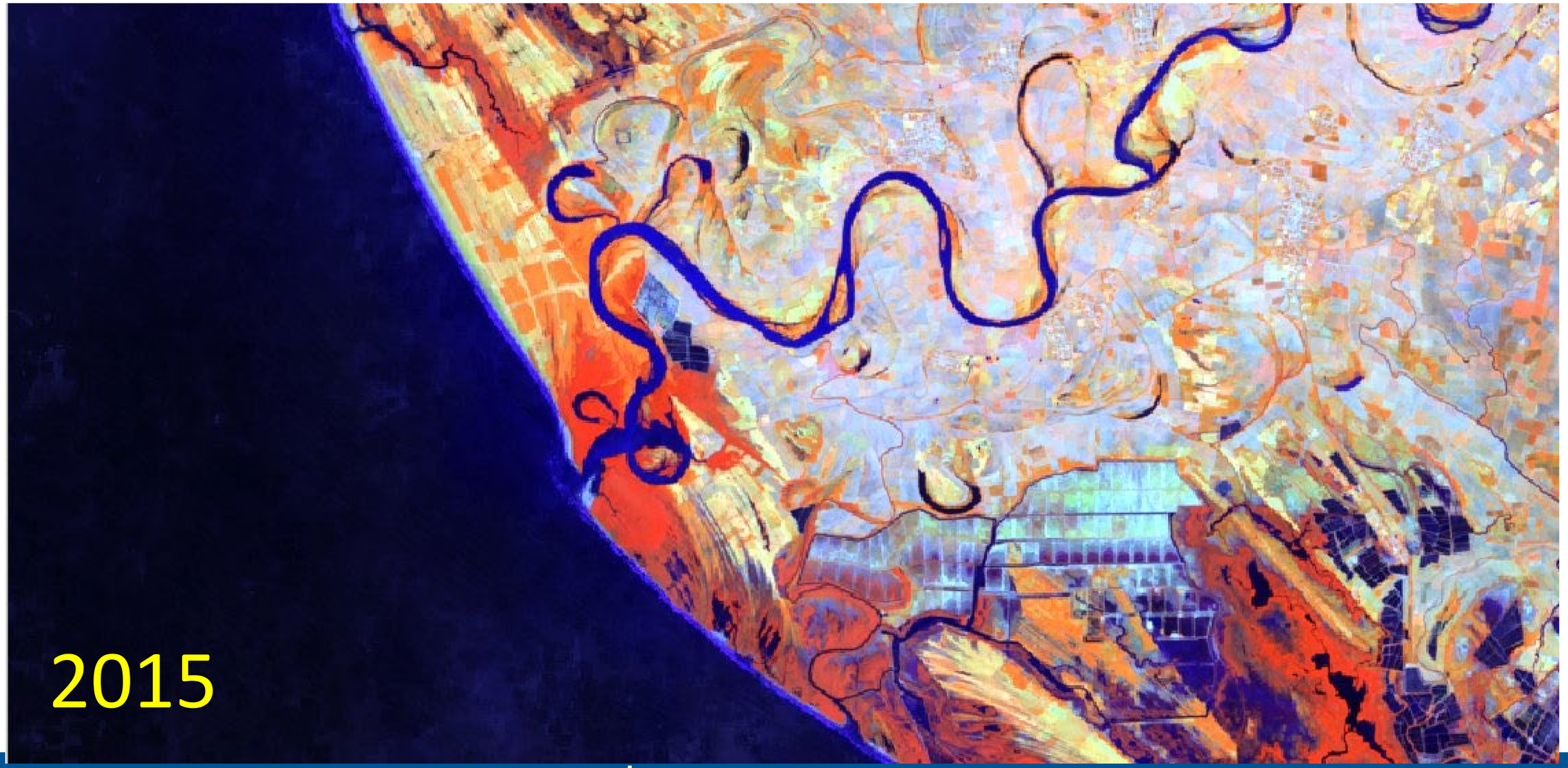
Yearly mosaics: 1986 – 2018

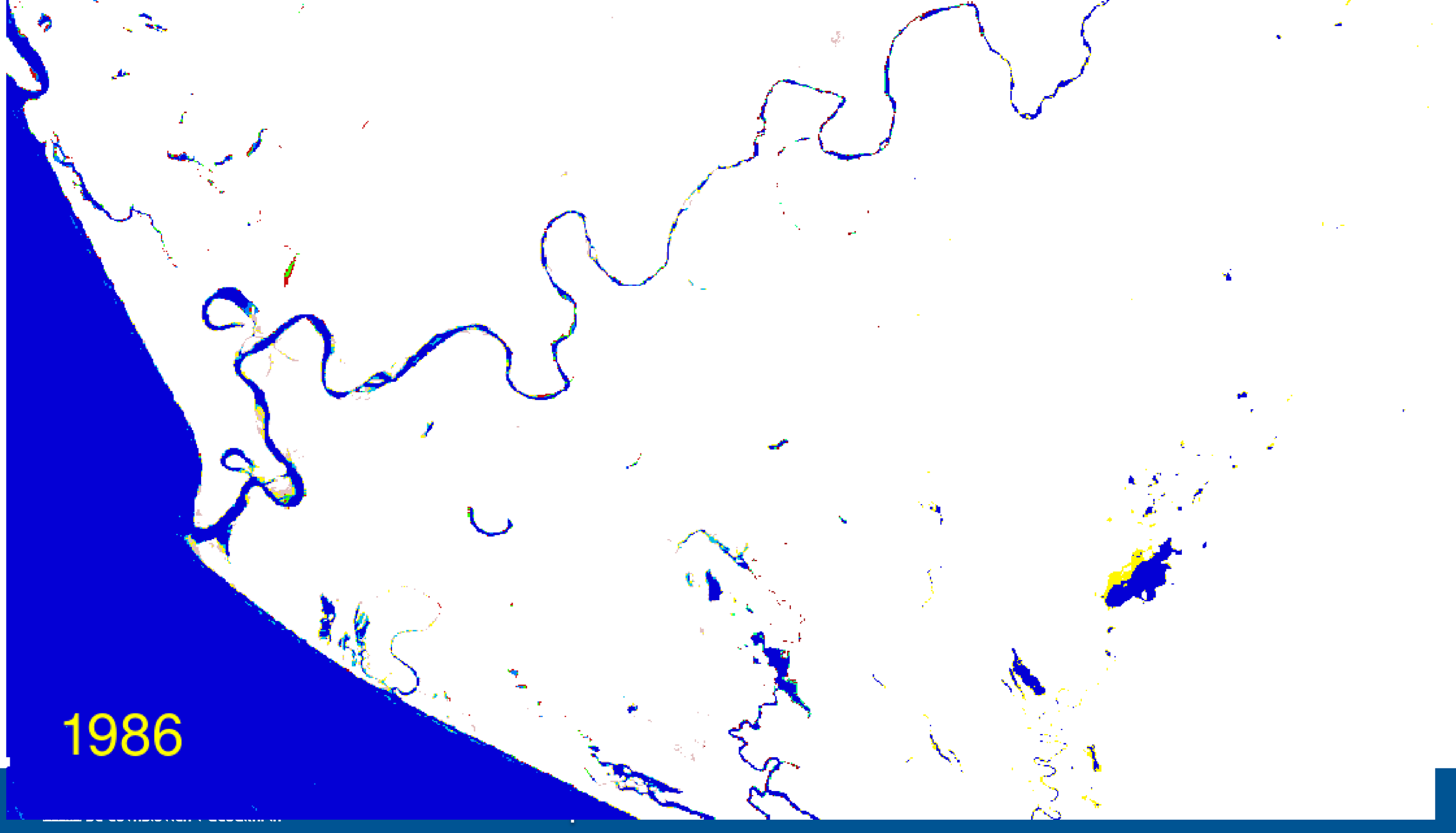
Geomediana

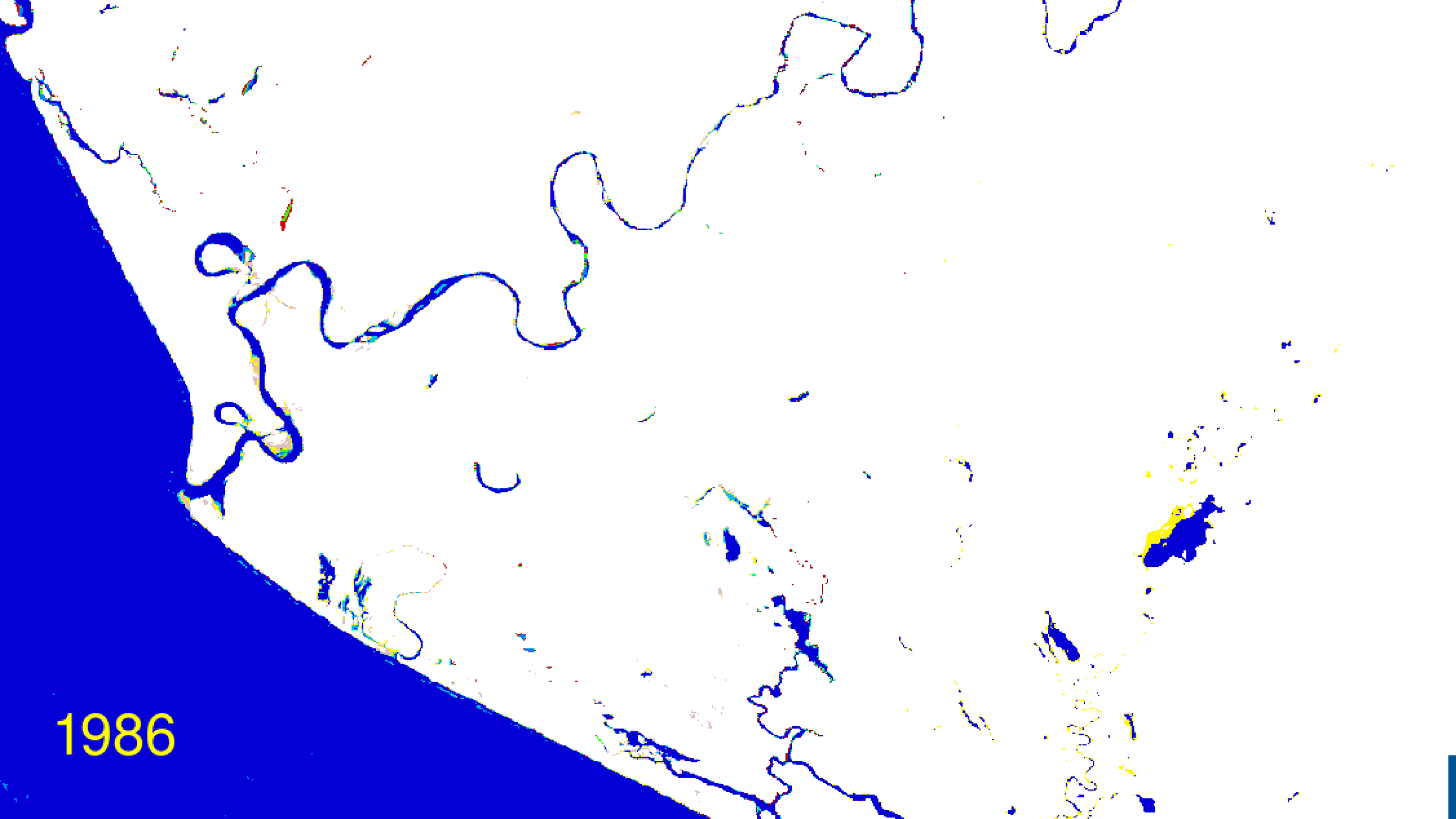
“WOFS”

Santiago River Mouth

Santiago River Mouth




1986


1986


1986


1995

Water monitoring

WOFS: Water Observations From Space.

Estimation in a time period of the porcentaje of (valid) observations where a pixel was classified as water

.

0 Permanently dry areas

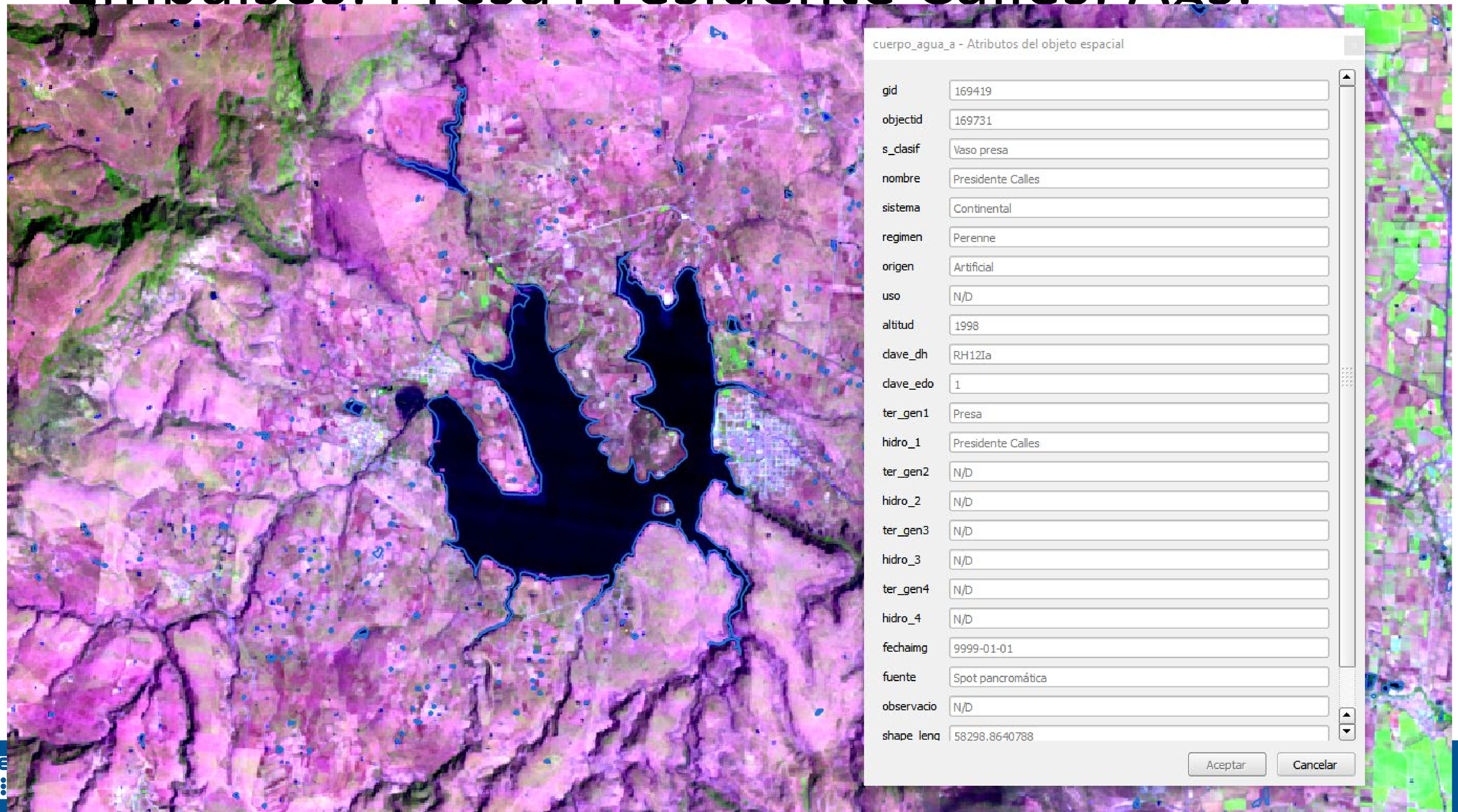
> 0 < 90 temporal flooding

>90: permanent water bodies

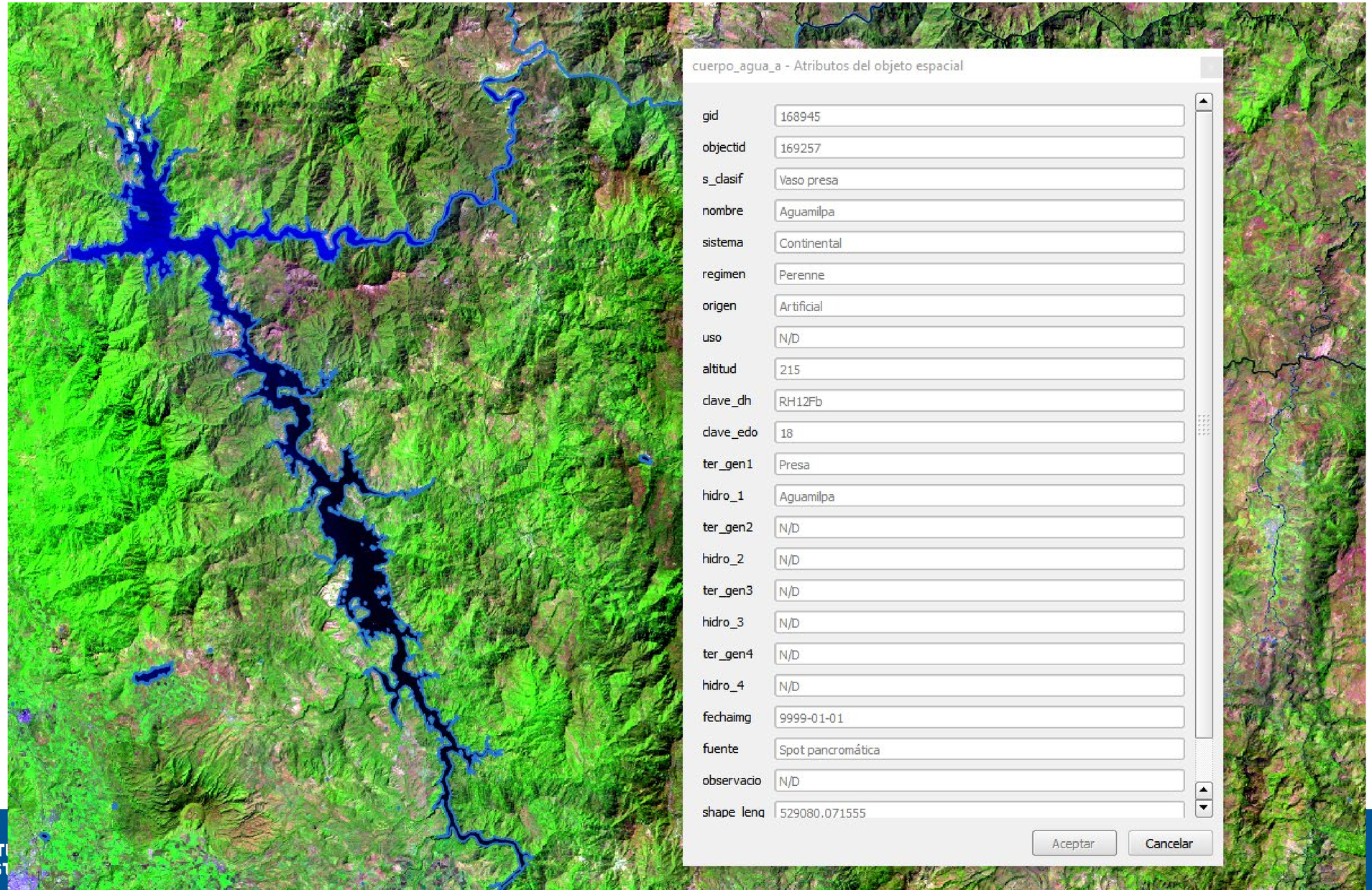
Evolution of water bodies and stream

Water body monitoring

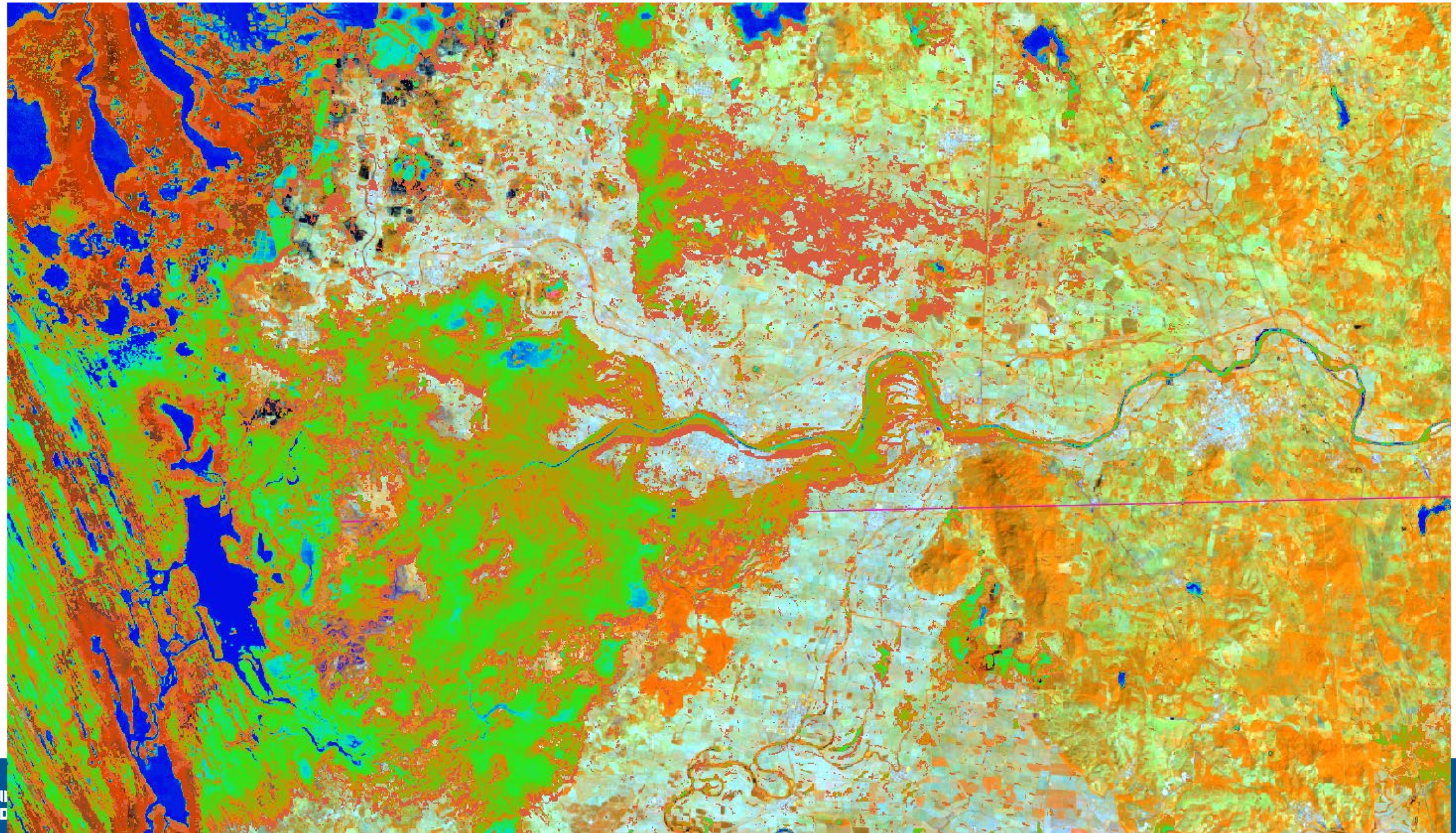
Embalses: Presa Presidente Calles, Ags.



Embalses: Presa Aguamilpa



Floodings, San Pedro River – Tuxpan, Nay. 2008



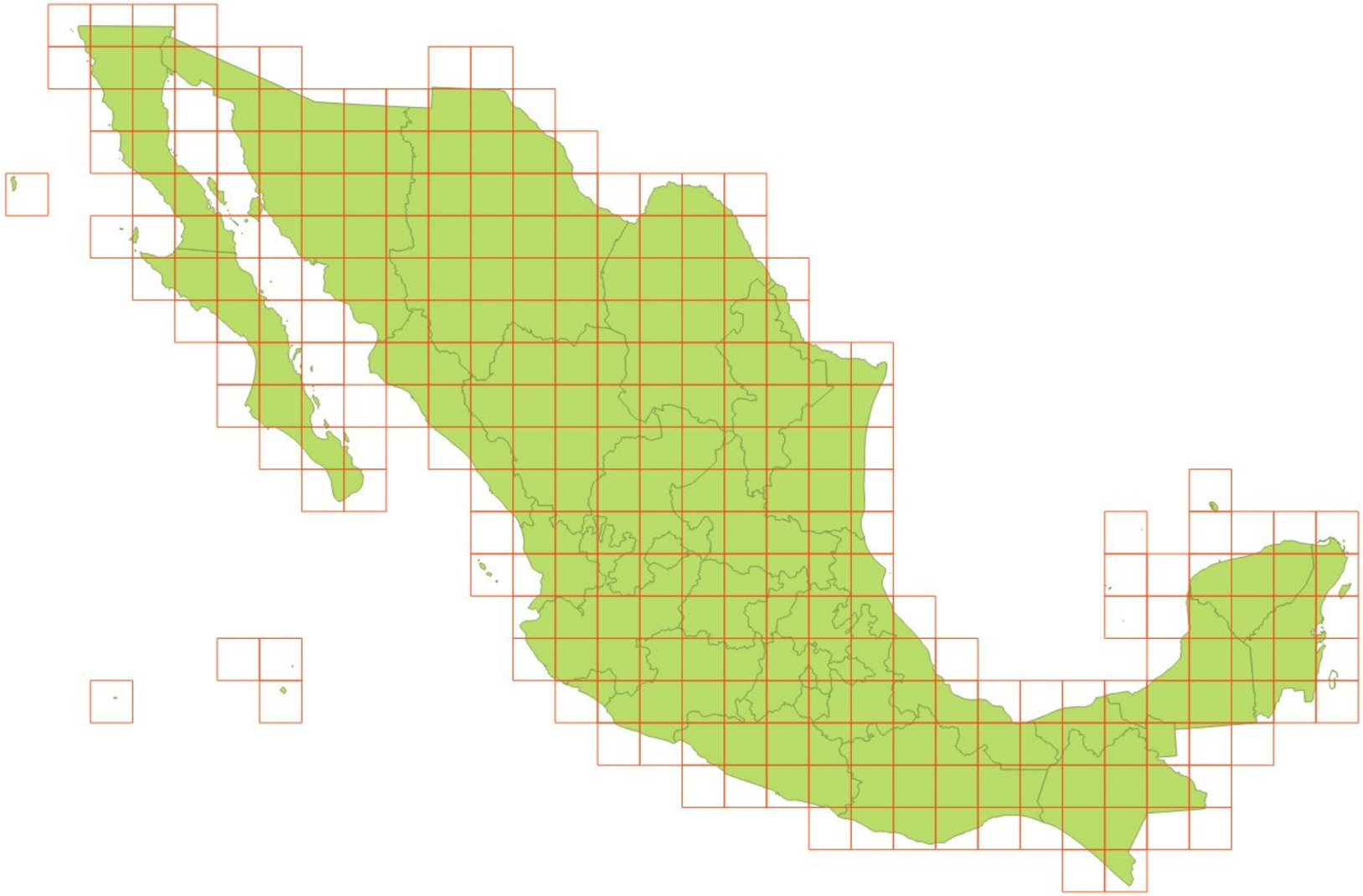
Proposed GRID

AEA Projection

100x100Km tiles

25 m Pixel

283 tiles

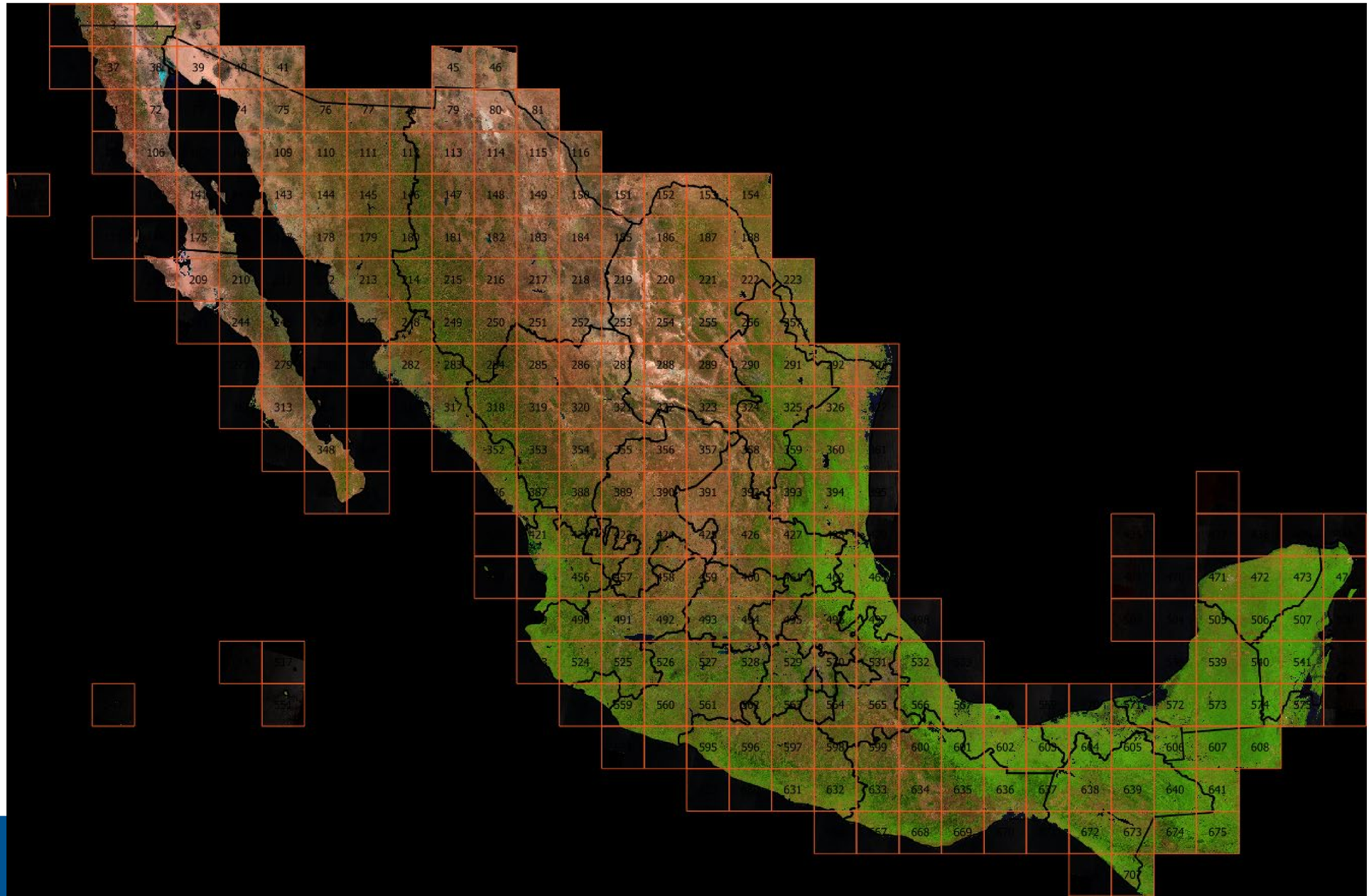


Geomedian 2015

283 tiles

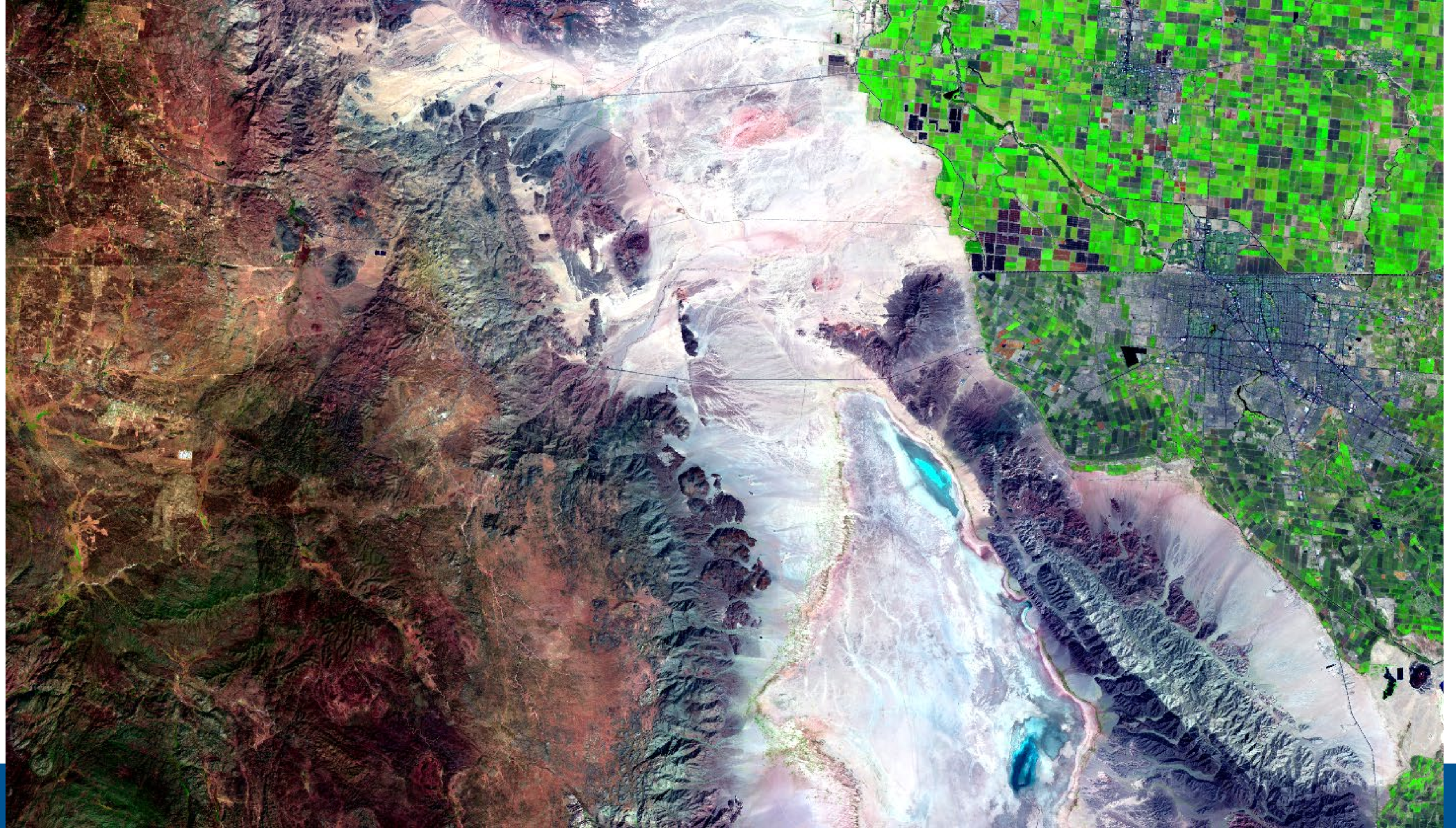
- Continental land
- Islands

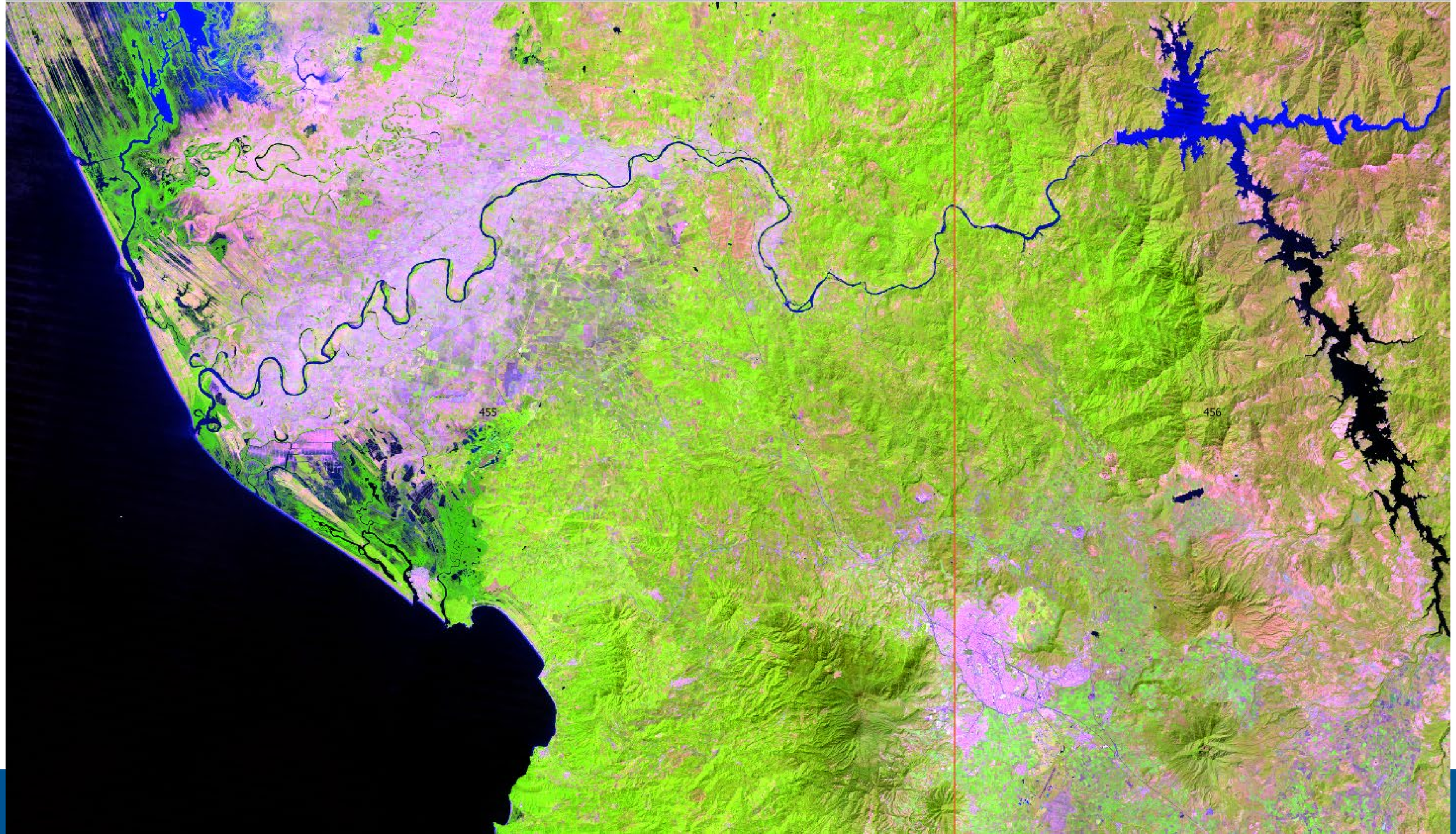
Web Map Service working

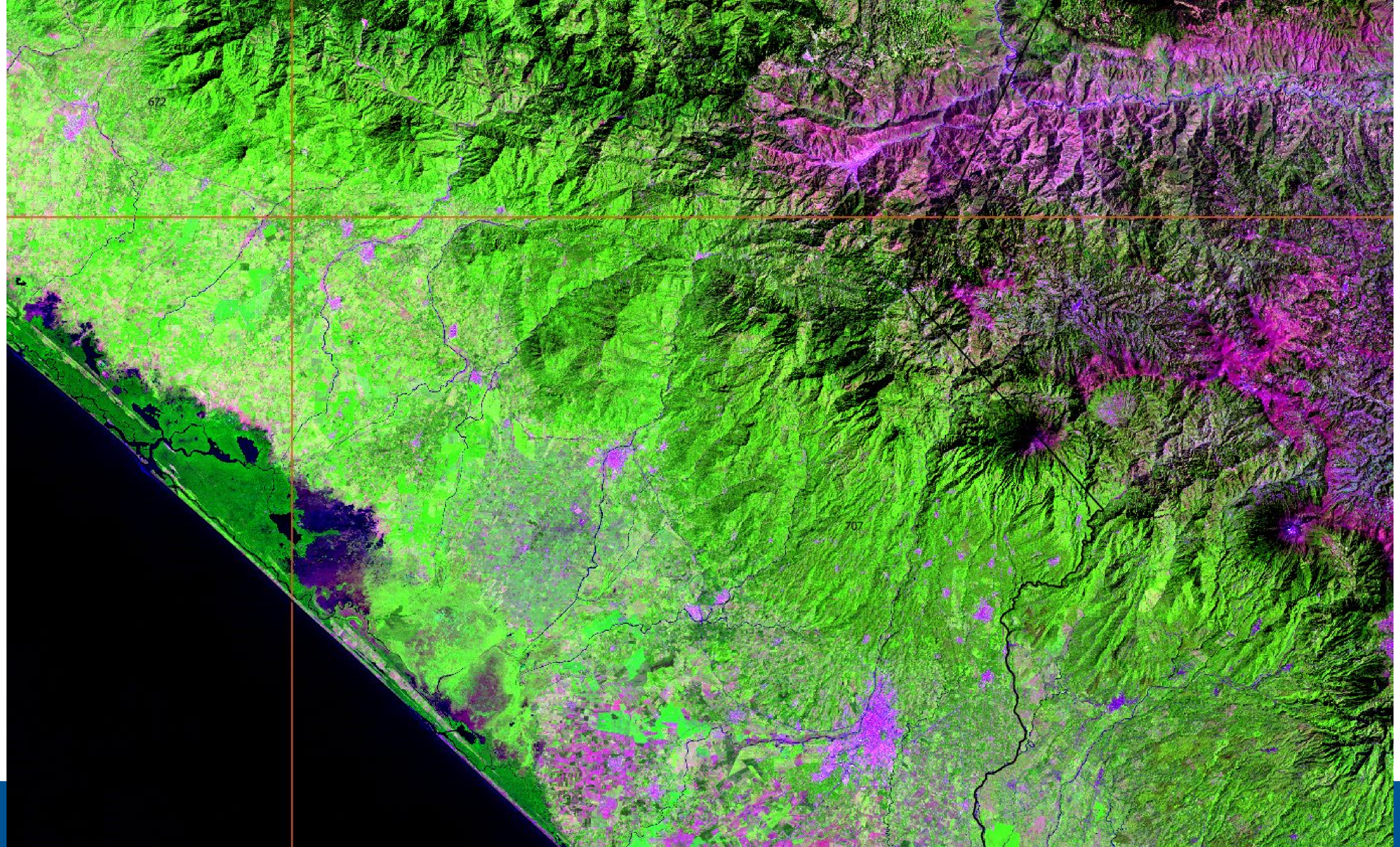


False color (infrared, red, green)

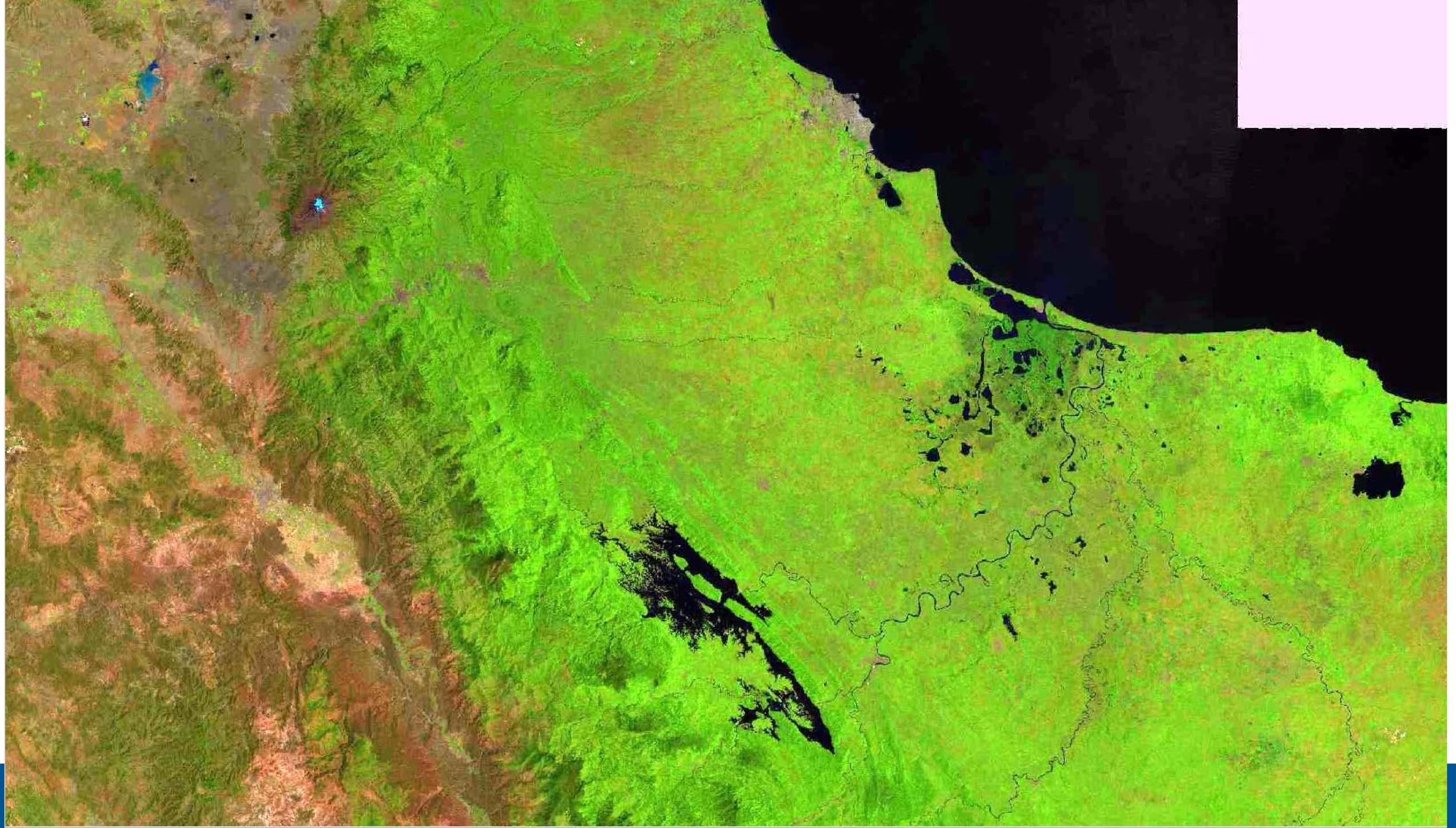
State boundaries

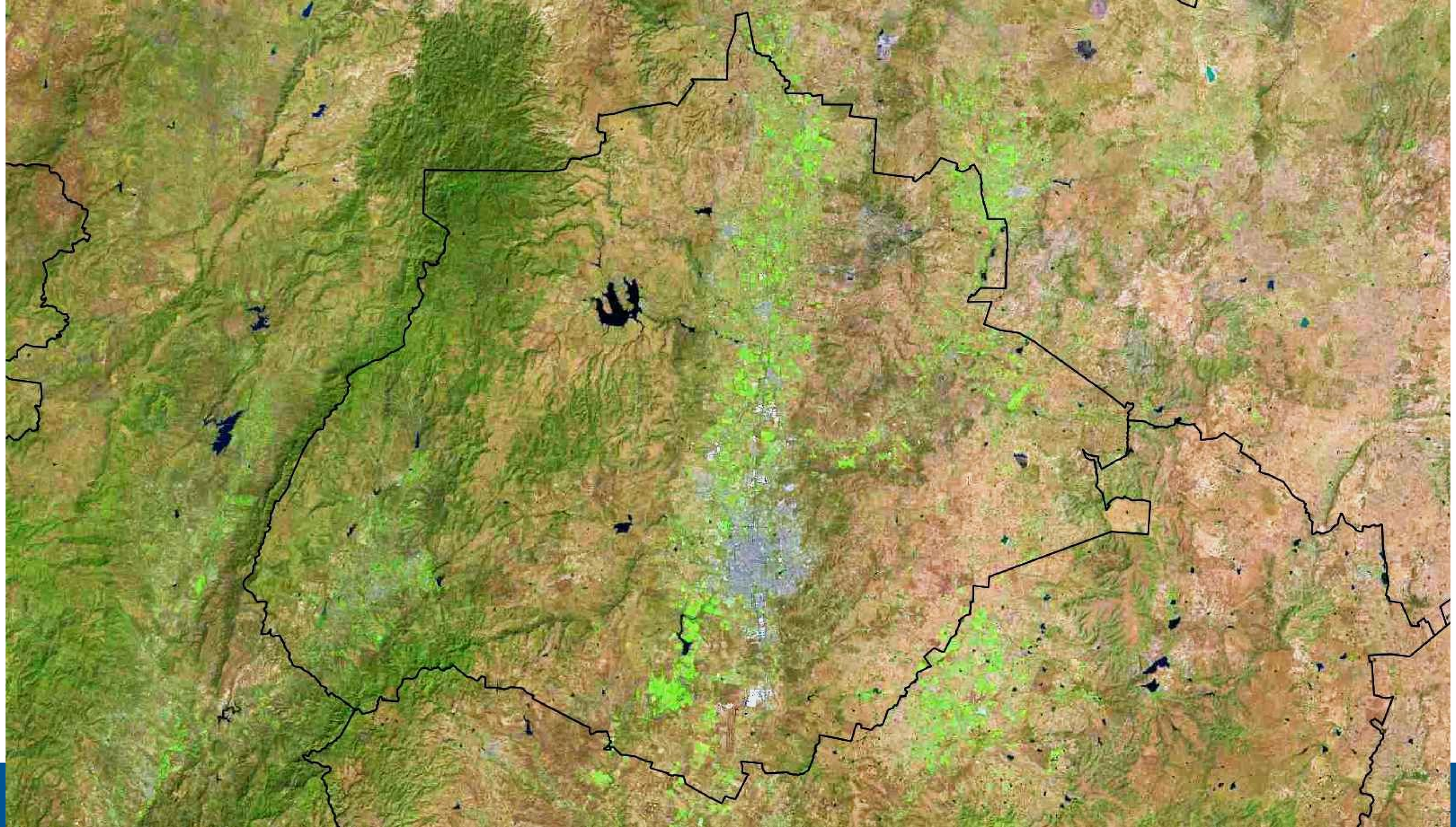


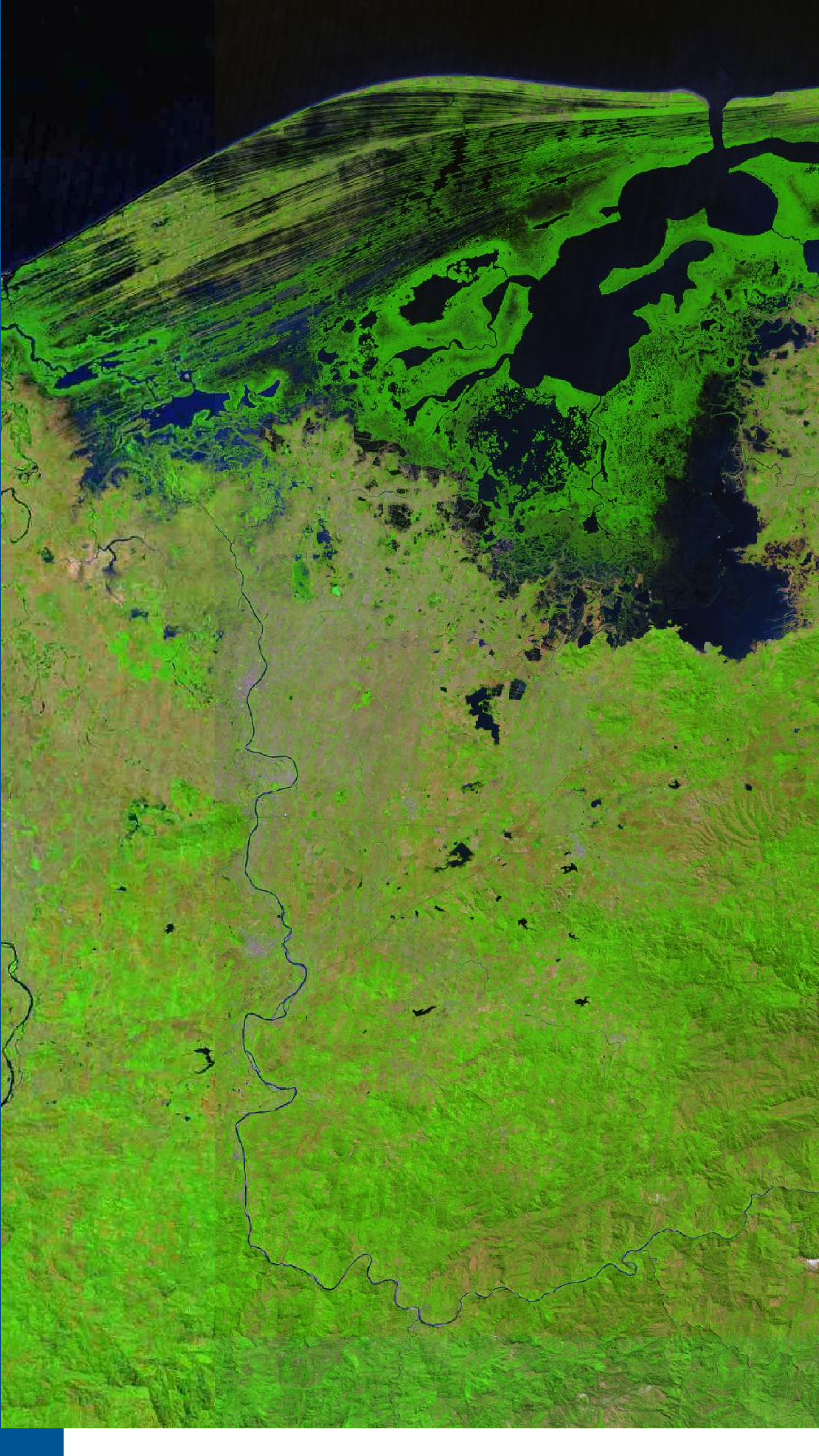

455
456


672
707


707







Prospective

- Oportunity to produce answers to new data requirements – need (public policies, knowledge)

- Improve existant products: vegetation, land cover, water.
  - Spatial detail, quality, time, pertinence.
  - Adapting, updating methodologies
  - Leverage technology and expertise and knowledge
- New data products
  - More often
  - Simple, useful

Prospective: SDG, SEEA

From improved products

New products and variables

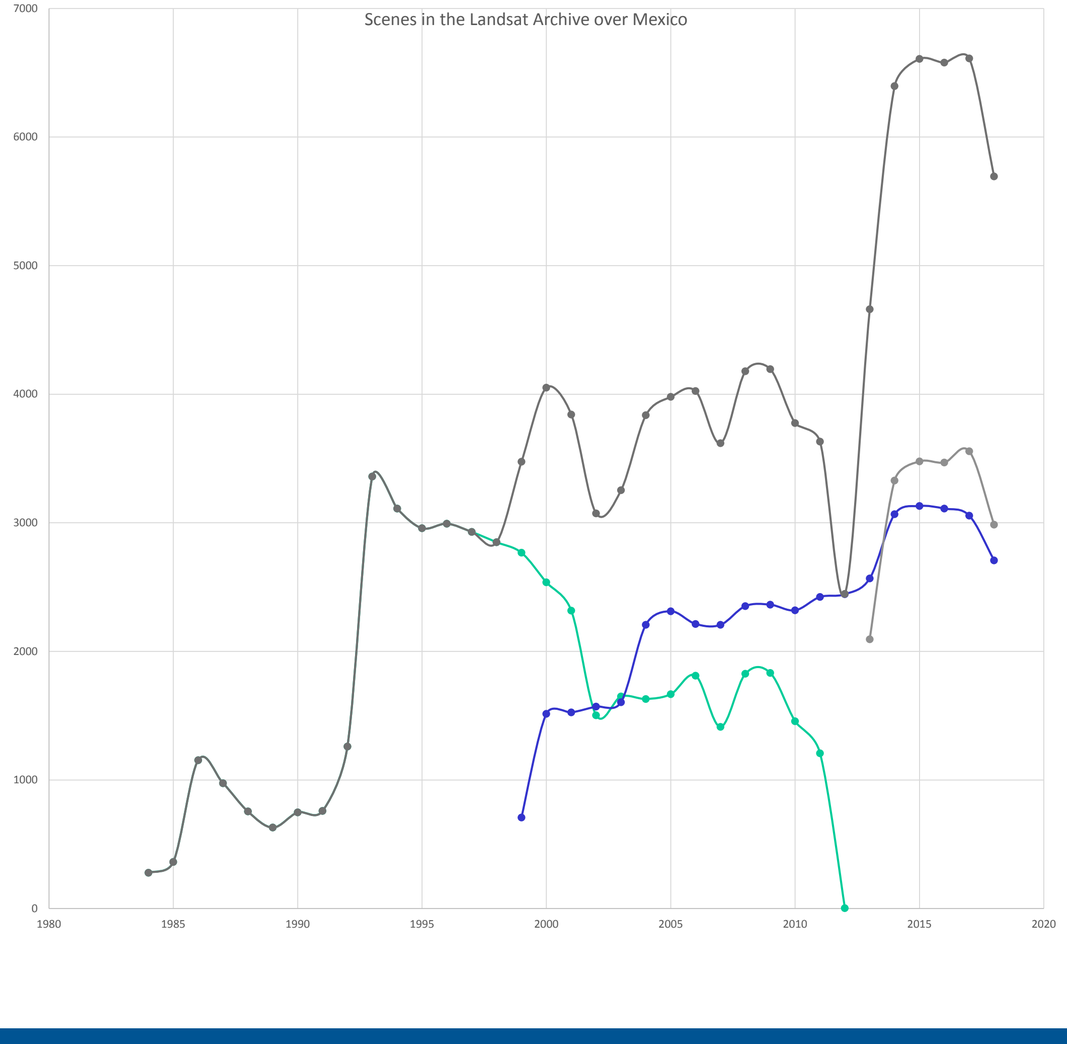
- NDVI
- Other indexes

Other applications for modelling and analysys using ODC tools.

Historical changes: 1980´s to present

LANDSAT ARCHIVE FOR MEXICO

Landsat Archive Summary for Mexico				
	Landsat 5	Landsat 7	Landsat 8	
Year	TM	ETM+	OLI	TOTAL
1984	278			278
1985	362			362
1986	1154			1154
1987	975			975
1988	755			755
1989	630			630
1990	748			748
1991	759			759
1992	1260			1260
1993	3360			3360
1994	3110			3110
1995	2958			2958
1996	2993			2993
1997	2929			2929
1998	2849			2849
1999	2767	708		3475
2000	2537	1514		4051
2001	2317	1525		3842
2002	1503	1570		3073
2003	1649	1604		3253
2004	1630	2207		3837
2005	1667	2312		3979
2006	1811	2213		4024
2007	1413	2206		3619
2008	1826	2352		4178
2009	1832	2363		4195
2010	1456	2319		3775
2011	1208	2423		3631
2012	3	2444		2447
2013		2567	2094	4661
2014		3067	3329	6396
2015		3130	3477	6607
2016		3110	3469	6579
2017		3055	3556	6611
2018		2708	2985	5693
TOTAL (all years)	48739	45397	18910	113046
	Landsat 5	Landsat 7	Landsat 8	TOTAL
Compressed Volume per Scene (MB)	150	235	919	
Total Compressed Volume (TB)	7.3	10.7	17.4	35.4
Uncompressed Volume per Scene (MB)	500	785	1610	
Total Uncompressed Volume (TB)	24.4	35.6	30.4	90.5



INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA

Thank you!