

Información geoespacial y desarrollo sostenible, perspectivas globales:

# CUBO DE DATOS GEOESPACIALES DE MÉXICO



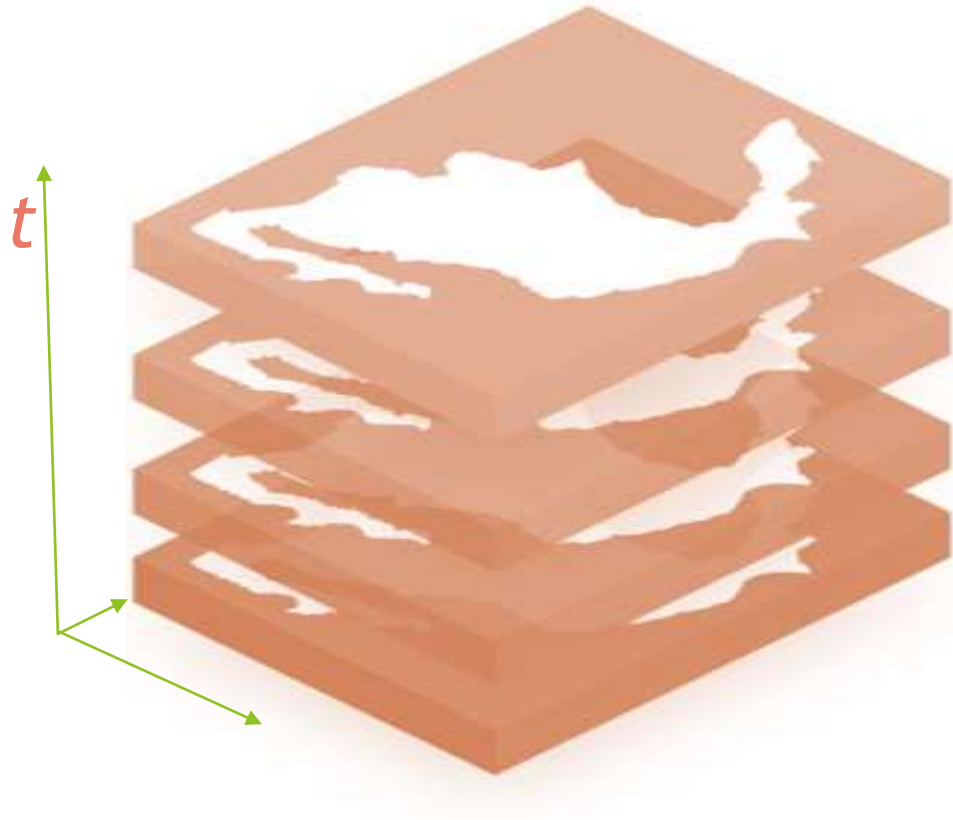
CUBO DE DATOS  
GEOESPACIALES DE MÉXICO



# Cubo de Datos

Arreglo masivo de datos  
multidimensionales  
(ráster).

# Cubo de Datos



Arreglo masivo de datos multidimensionales (ráster).

# Panorama General de la Producción Geoespacial del INEGI

El INEGI produce información geoespacial (inicialmente mapas de papel; 1970s,1980s)

Se realizan a partir de percepción remota (previamente fotografía aérea, actualmente imágenes satelitales)

Interpretación visual y métodos exhaustivos



# Panorama General de la Producción Geoespacial del INEGI

La demanda por mapas más frecuentes y detallados crece

## Limitaciones

- Resolución espacial impuesta en la producción de mapas
- La exhaustividad impide actualizaciones oportunas de la información



# Panorama General de la Producción Geoespacial del INEGI

Necesidades identificadas (aumentar la resolución/detalle y la frecuencia)

Solución:  
Big Data / Machine Learning  
(almacenamiento, manejo  
procesamiento inteligente de  
grandes y variados volúmenes  
de información)



# Panorama General de la Producción Geoespacial del INEGI

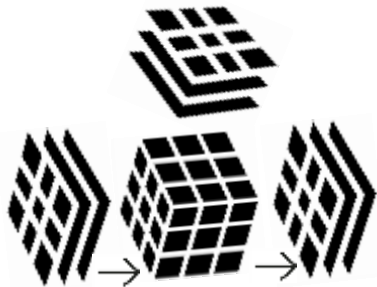
Necesidades identificadas (aumentar la resolución/detalle y la frecuencia)

Solución:  
Big Data / Machine Learning  
(almacenamiento, manejo  
procesamiento inteligente de  
grandes y variados volúmenes  
de información)



# Open Data Cube INEGI

Adquisición y  
Preprocesamiento



Análisis y  
Procesamiento



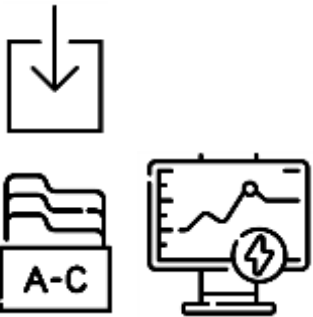
Traducir  
Resultados a  
Productos





# Ventajas


- ▶ Menos almacenamiento para el usuario (descargas específicas de área)
- ▶ Menos procesamiento (Datos Listos para el Análisis -ARD)
- ▶ Organización automática (ingesta/indexado) --acceso eficiente



# Ventajas

- ▶ Menos almacenamiento para el usuario (descargas específicas de área)
- ▶ Menos procesamiento (Datos Listos para el Análisis -ARD)
- ▶ Organización automática (ingesta/indexado) -acceso eficiente



- ▶ Homologa resultados (ARD)
- ▶ Compatibilidad (Open source)
- ▶ Al manejar tantos datos permite análisis computacional más complejo (Computer Vision)
- ▶  python™
- ▶ **x,y** (análisis a nivel pixel permite aprovechar mejor las imágenes con nubes)
- ▶ **t** (análisis con una dimensión temporal)
- ▶ Series de tiempo (monitoreo de cambio)
- ▶ ODS
  
- ▶ Ejemplo caso práctico: geomediana (resumen estadístico en 1 imagen)

# Avance actual en cuanto a:

- ❑ **INSUMOS**  
(Imágenes LANDSAT proporcionadas)
- ❑ **CAPACITACIÓN**  
(Ejercicios Piloto en 4 áreas)
- ❑ **INFRAESTRUCTURA**  
(Inicial y estimaciones futuras)
- ❑ **ESTRATEGIA**  
(Propuestas de etapas de implementación)

# Avance actual en cuanto a:

- ❑ **INSUMOS**  
(Imágenes LANDSAT proporcionadas)
- ❑ **CAPACITACIÓN**  
(Ejercicios Piloto en 4 áreas)
- ❑ **INFRAESTRUCTURA**  
(inicial y estimaciones futuras)
- ❑ **ESTRATEGIA**  
(propuestas de etapas de implementación)

# Total de Imágenes LANDSAT proporcionadas

Período ( 1985 : 2018 )  
(imágenes disponibles en estos años para ciertas áreas de Interés)



Pathrows	Imágenes	TB
5	3,175	3.2

# Total de Imágenes LANDSAT proporcionadas

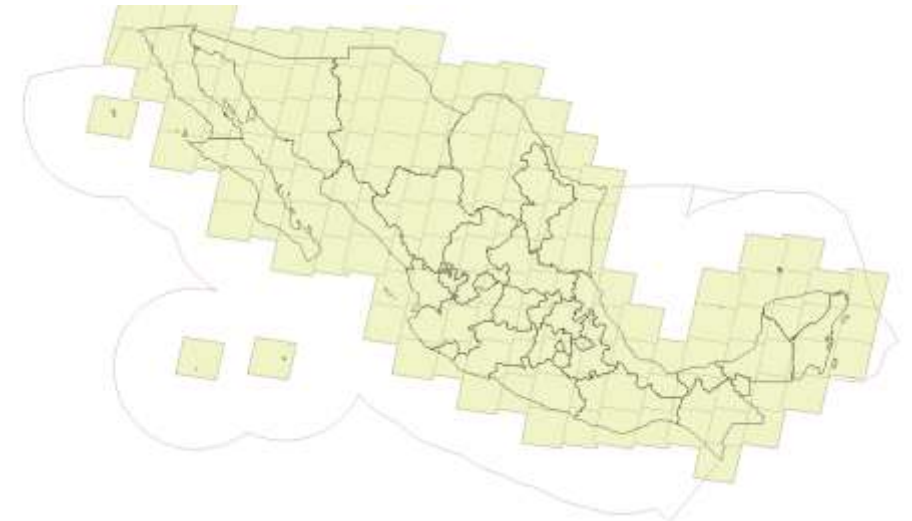
Período ( 1985 : 2018 )  
(Imágenes disponibles en estos años para ciertas áreas de interés)



Pathrows	Imágenes	TB
5	3,175	3.2



Cobertura nacional 2011 y 2015  
(imágenes disponibles en estos años)



Año	Pathrows	Imágenes	TB
2011	135	3,517	3.9
2015	138	6,074	7.5

# Total de Imágenes LANDSAT proporcionadas

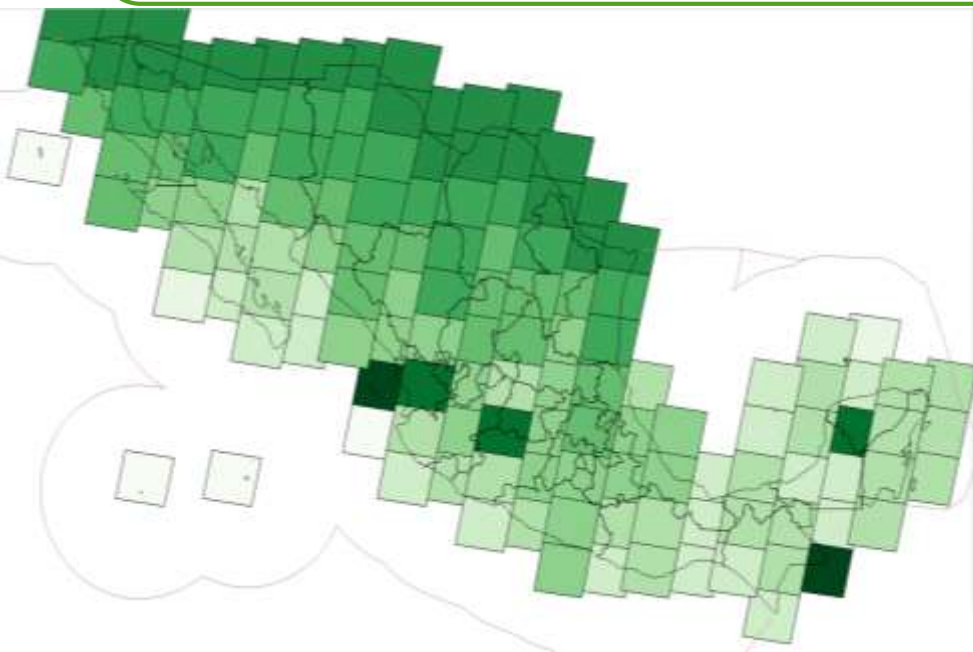
Período ( 1985 : 2018 )  
(imágenes disponibles en estos años para ciertas áreas de interés)

Pathrows	Imágenes	TB
5	3,175	3.2



Cobertura nacional 2011 y 2015  
(Imágenes disponibles en estos años)

Año	Pathrows	Imágenes	TB
2011	135	3,517	3.9
2015	138	6,074	7.5



Total de Imágenes proporcionadas por Geoscience Australia

Pathrows	Imágenes	TB
138	12,494	14.5

~ 1.2 GB/imagen

# Avance actual en cuanto a:

- ❑ INSUMOS  
(Imágenes LANDSAT proporcionadas)
- ❑ **CAPACITACIÓN**  
(Ejercicios Piloto en 4 áreas)
- ❑ INFRAESTRUCTURA  
(Inicial y estimaciones futuras)
- ❑ ESTRATEGIA  
(Propuestas de etapas de implementación)



# Algunos resultados en ejercicio piloto

- ▶ Vegetación
- ▶ Humedales y Línea Costera
- ▶ Crecimiento Urbano
- ▶ Agricultura





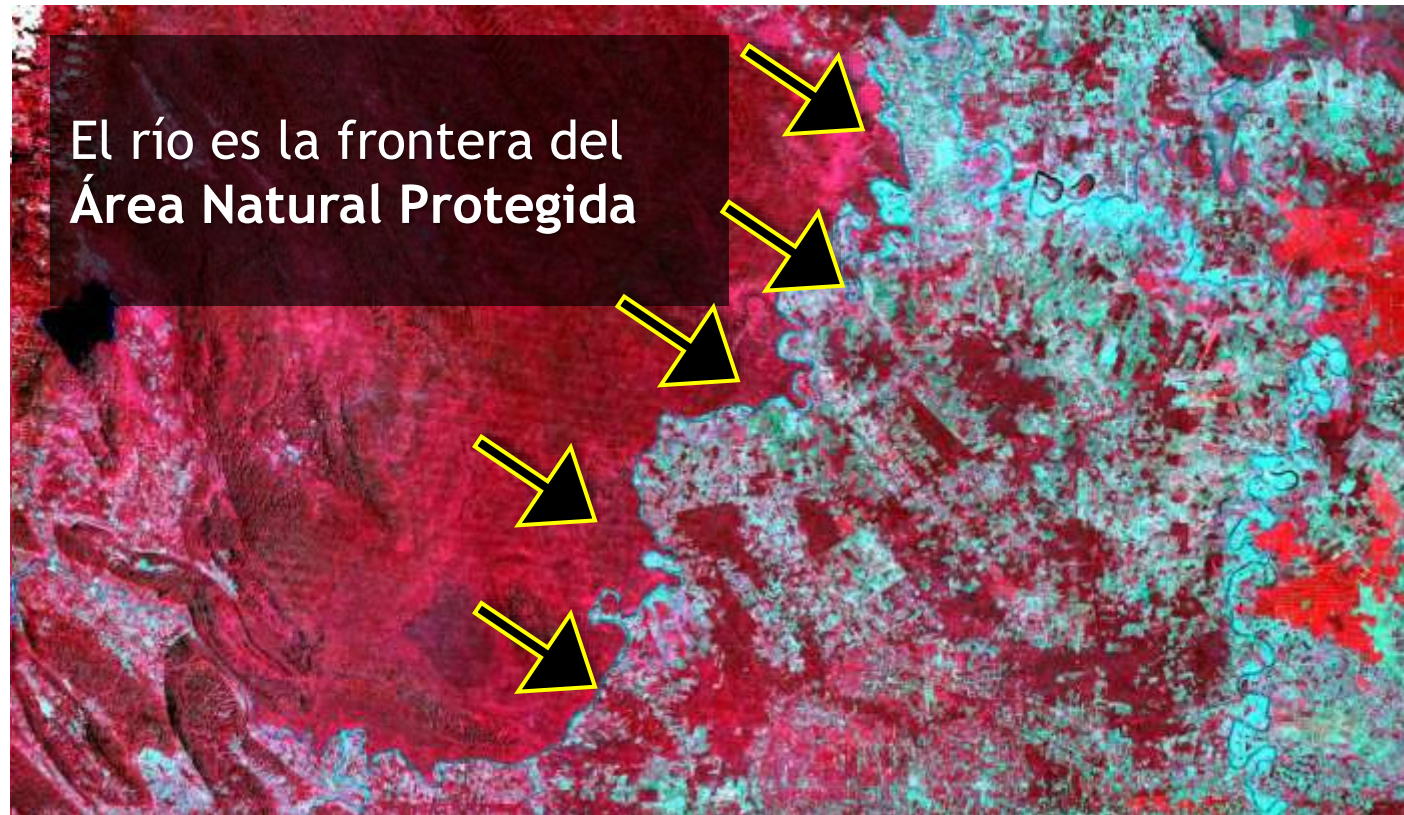
# Algunos resultados en ejercicio piloto

- ▶ Vegetación
- ▶ Humedales y Línea Costera
- ▶ Crecimiento Urbano
- ▶ Agricultura



# Vegetación

Montes Azules y Marqués de Comillas; **1986** → 2017



# Algunos resultados en ejercicio piloto

- ▶ Vegetación
- ▶ Humedales y Línea Costera
- ▶ Crecimiento Urbano
- ▶ Agricultura

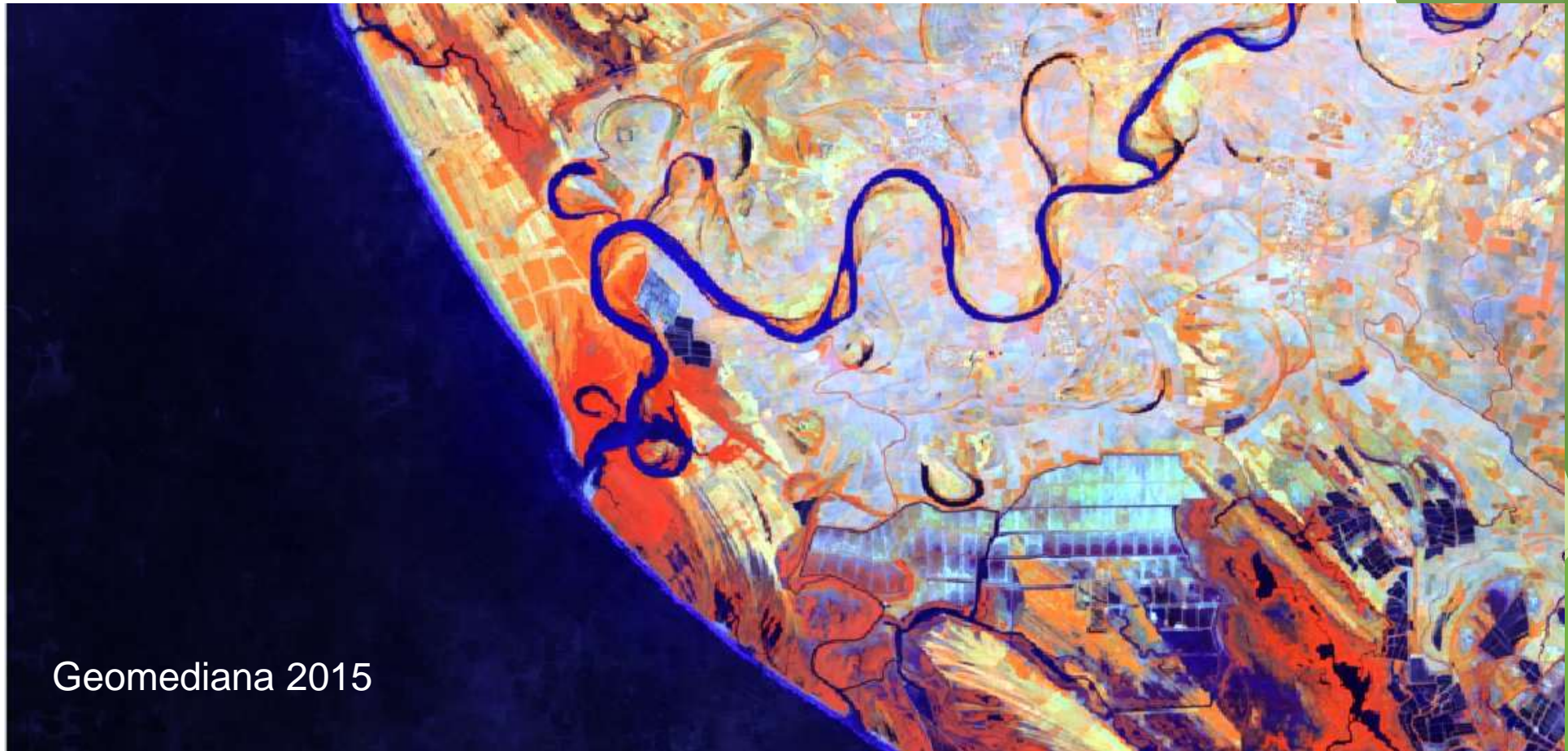
# Erosión costera en la boca del Río Santiago



Geomediana 1993



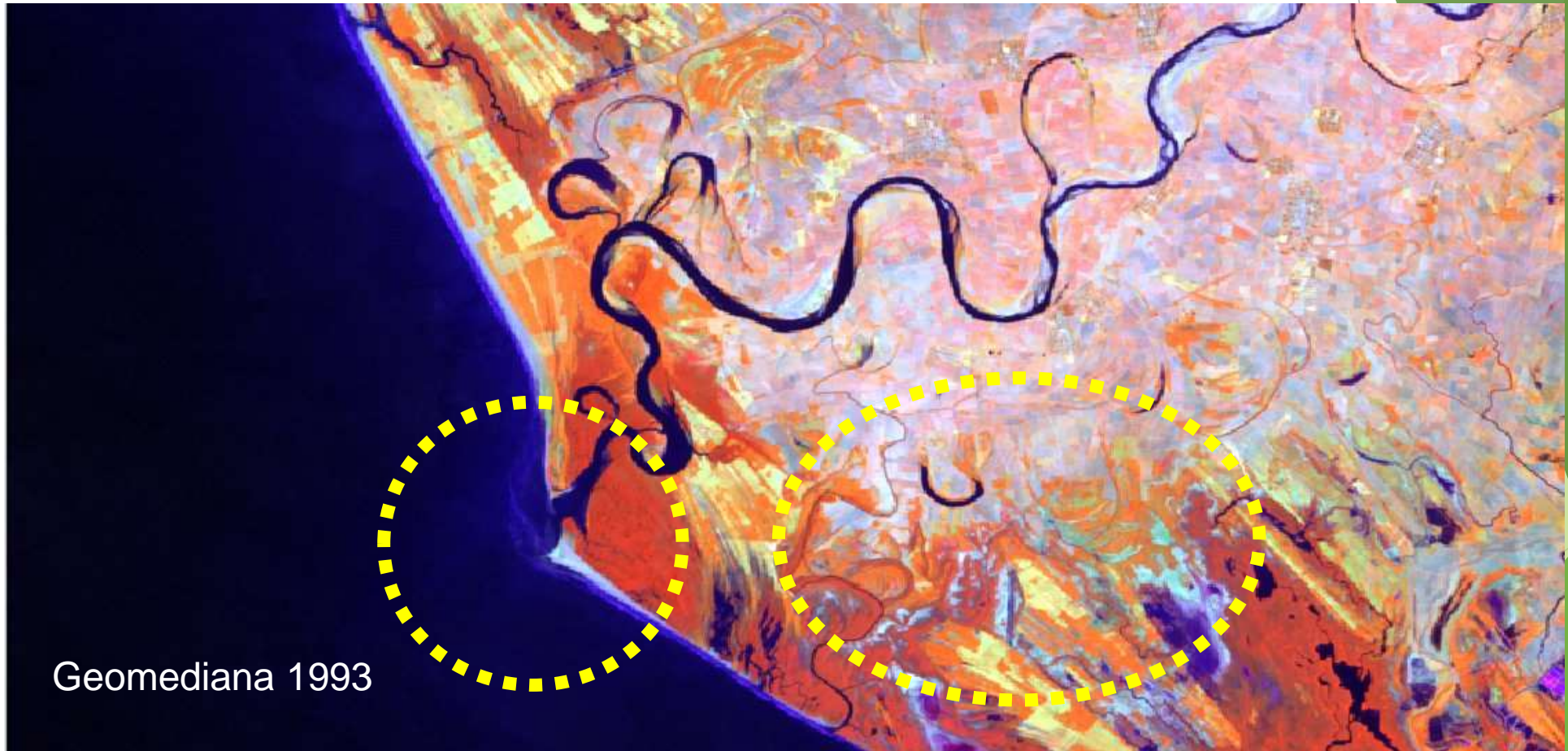
# Erosión costera en la boca del Río Santiago



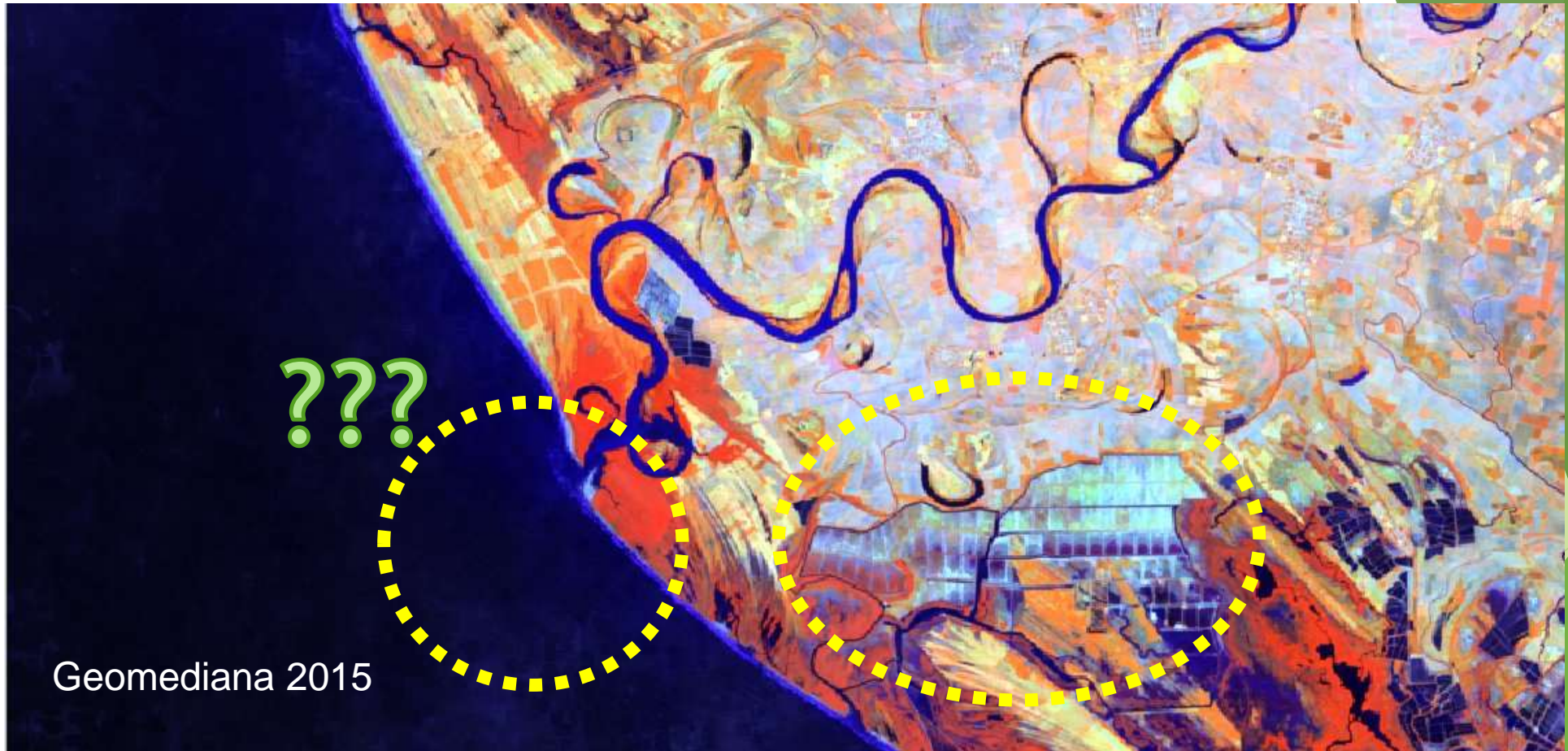
Geomediana 2015



# Erosión costera en la boca del Río Santiago

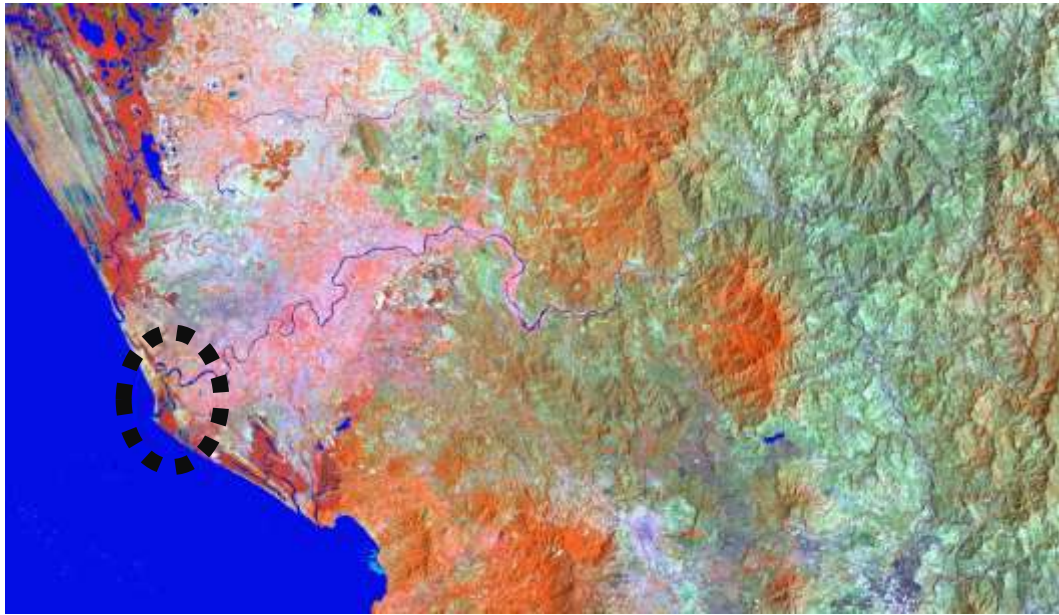


# Erosión costera en la boca del Río Santiago





# Erosión costera en la boca del Río Santiago



Geomédiana 1986



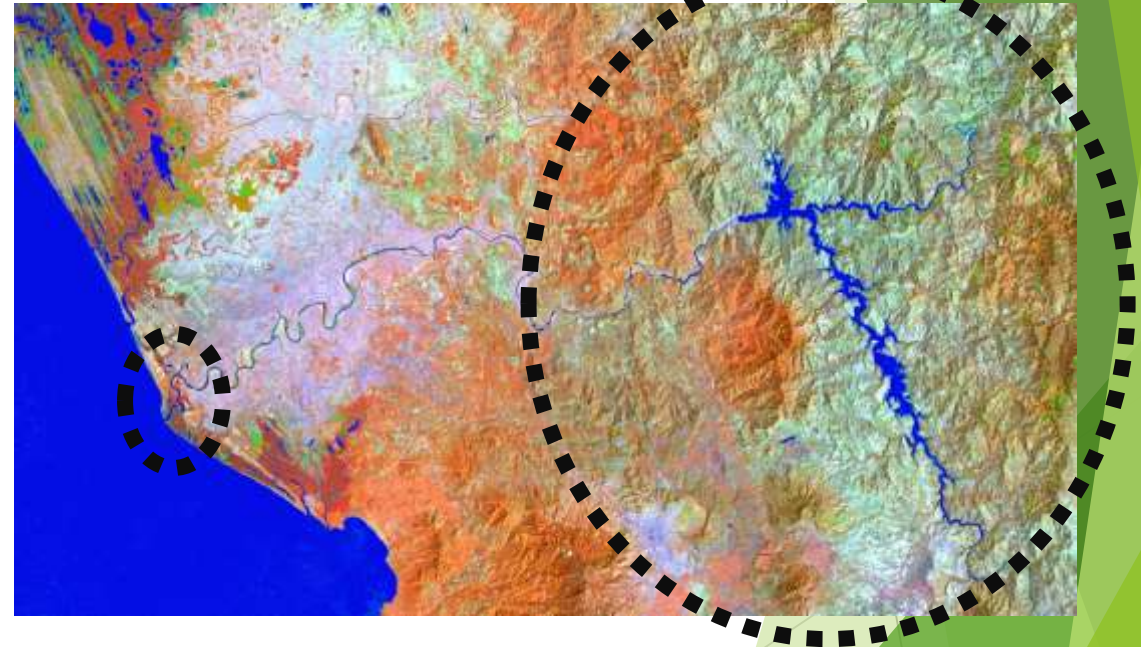
Geomédiana 1995

# Erosión costera en la boca del Río Santiago

Al identificarse el gran cambio, se buscó una explicación



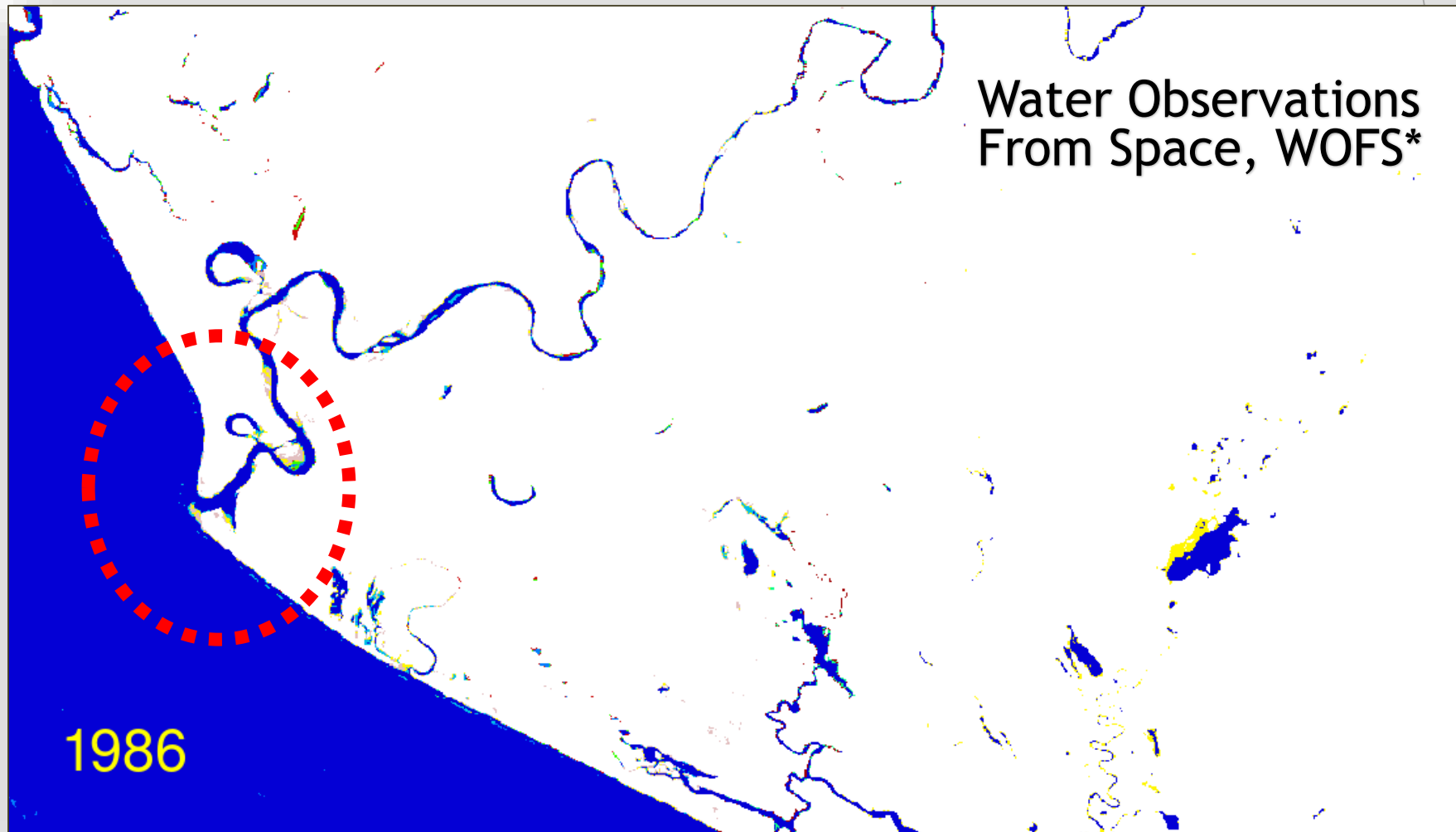
Geomedia 1986



Geomedia 1995



# Erosión costera en la boca del Río Santiago

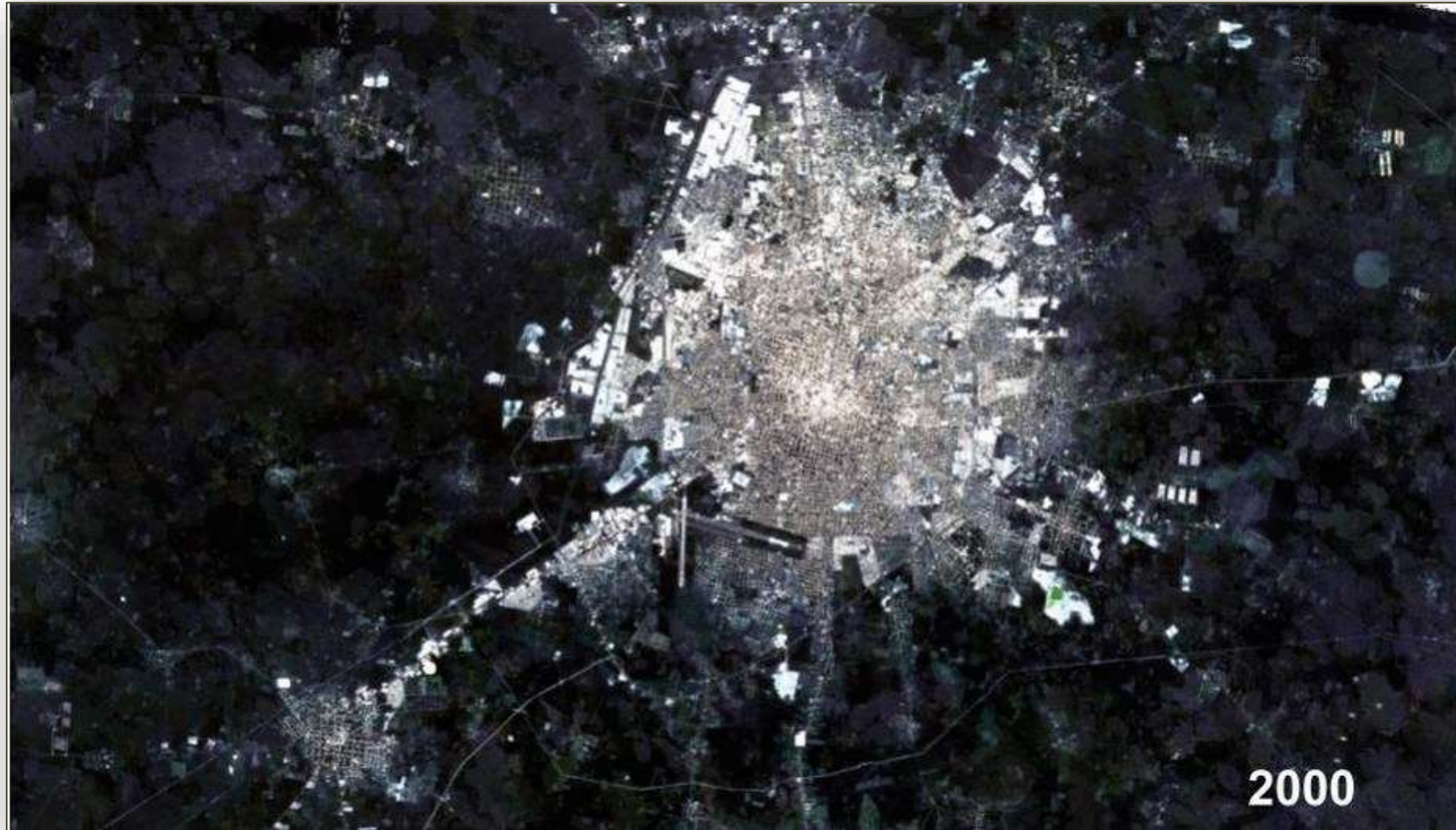


\*WOFS Algorithm calculates the proportional presence of water at each pixel during certain amount of time

# Algunos resultados en ejercicio piloto

- ▶ Vegetación
- ▶ Humedales y Línea Costera
- ▶ Crecimiento Urbano
- ▶ Agricultura

# Geomediana anual en Mérida 2000-2017



Y  
:  
eomedian

# Algunos resultados en ejercicio piloto

- ▶ Vegetación
- ▶ Humedales y Línea Costera
- ▶ Crecimiento Urbano
- ▶ Agricultura



# Identificación de cultivos

Mosaico Geomediano Landsat  
junio a agosto de 2015, Guanajuato



Clasificación supervisada  
Random Forest Classifier



## Leyenda

### Cultivos

- Sorgo
- Maíz
- Alfalfa
- Otros cultivos

### Otros elementos

- Agua
- Urbano
- Suelo
- Bosque, matorral y otro tipo de vegetación



Datos de campo  
(2015)

# Identificación de cultivos

Mosaico Geomediano Landsat  
junio a agosto de 2015, Guanajuato



Clasificación supervisada  
Random Forest Classifier



## Leyenda

### Cultivos

- Sorgo
- Maíz
- Alfalfa
- Otros cultivos

### Otros elementos

- Agua
- Urbano
- Suelo
- Bosque, matorral y otro tipo de vegetación



Datos de campo  
(2015)

- Buena definición en terrenos grandes.
- Se necesitan tests con mejor resolución espacial - Sentinel.

# Avance actual en cuanto a:

- ❑ INSUMOS  
(Imágenes LANDSAT proporcionadas)
- ❑ CAPACITACIÓN  
(Ejercicios Piloto en 4 áreas)
- ❑ **INFRAESTRUCTURA**  
(Inicial y estimaciones futuras)
- ❑ ESTRATEGIA  
(Propuestas de etapas de implementación)

# Necesidades de Infraestructura

Landsat



sentinel

ALMACENAMIENTO

1 año ~ 8.5 TB (2015)  
19 años ~ 145 TB (2000:2018)  
34 años ~ 290 TB (1985:2018)

1 año ~ 100 TB (2018)  
3 años ~ 300 TB (2016:2018)



ANCHO DE BANDA  
100 Mbps

1 año de Imágenes cada 7 días

1 año de Imágenes cada 8 días  
(formato JPEG2000)

PROCESAMIENTO

32 núcleos  
1 año de imágenes  
2 millones de km<sup>2</sup>

72 horas



cobertura nacional de un indicador por 1 año de información

27 días

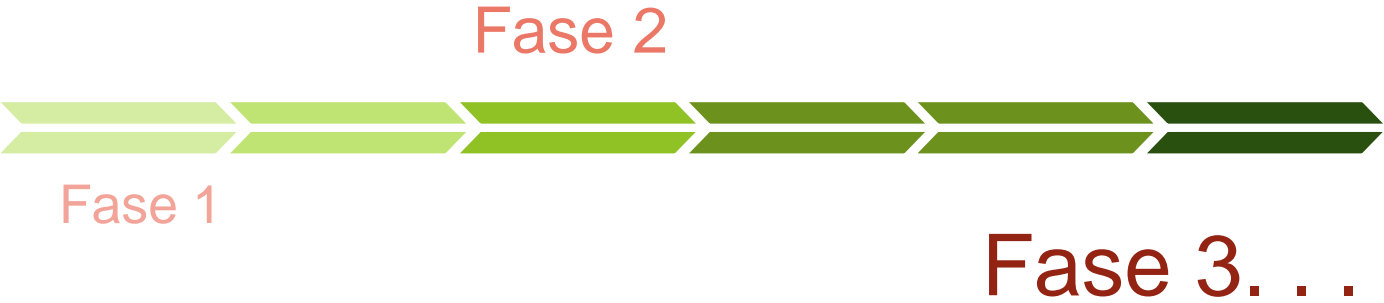




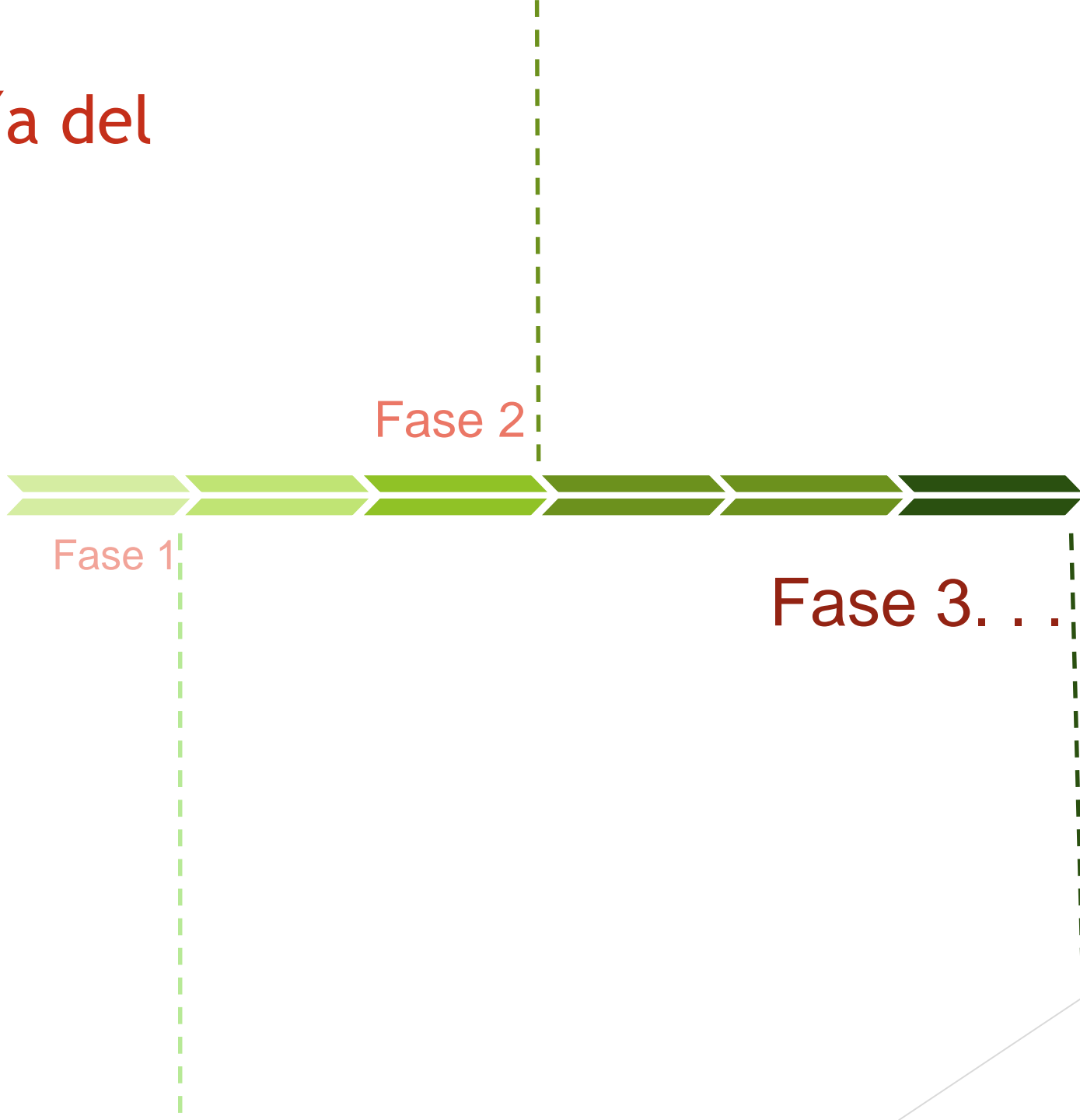
# Avance actual en cuanto a:

- ❑ INSUMOS  
(Imágenes LANDSAT proporcionadas)
- ❑ CAPACITACIÓN  
(Ejercicios Piloto en 4 áreas)
- ❑ INFRAESTRUCTURA  
(Inicial y estimaciones futuras)
- ❑ **ESTRATEGIA**  
(Propuestas de etapas de implementación)

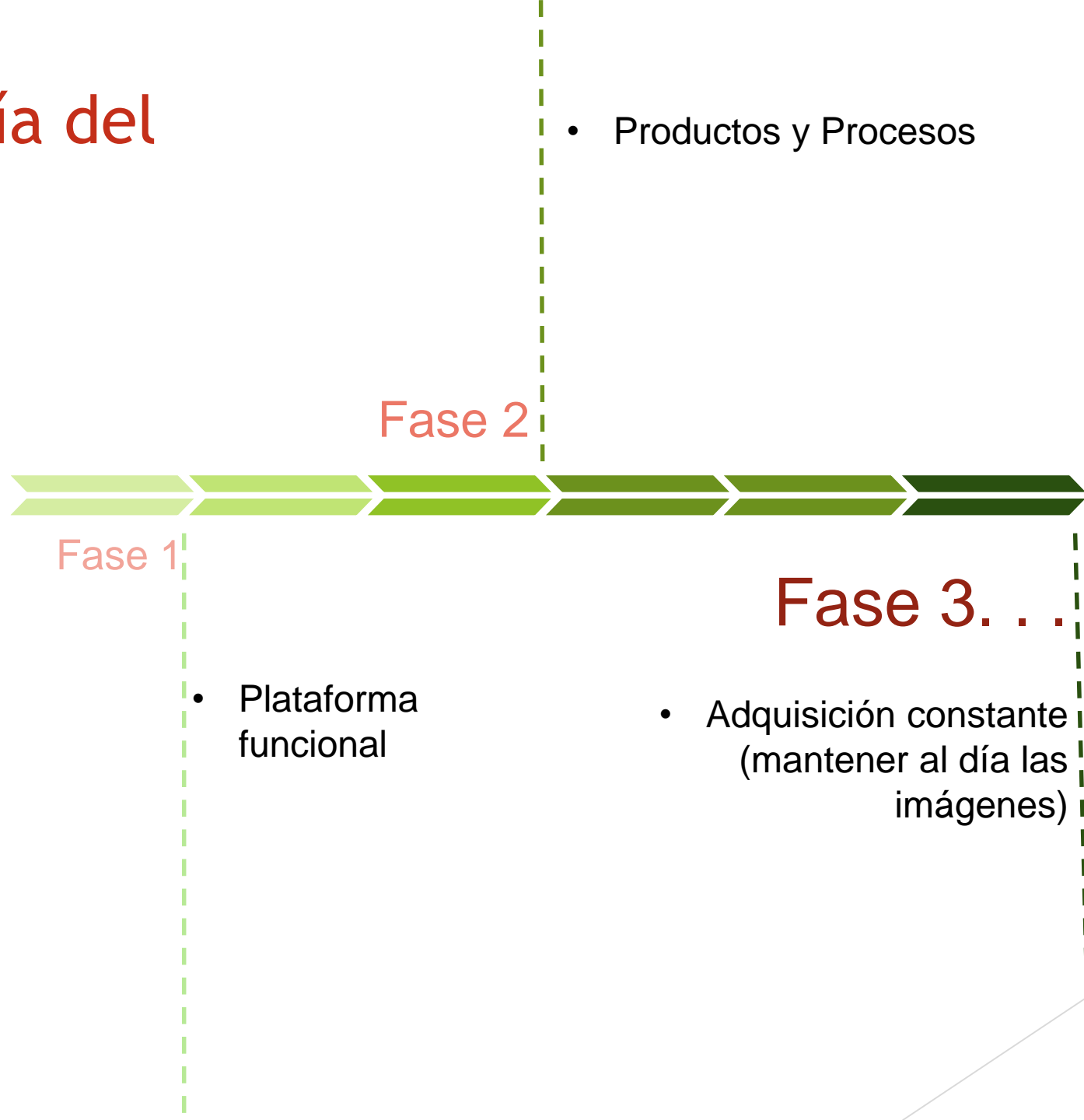
# Cronología del proyecto



# Cronología del proyecto



# Cronología del proyecto





# Cronología del proyecto

## + dimensionamiento



• **20 TB**  
(previamente en discos duros externos)

### Fase 1

- Plataforma funcional

### Fase 2

- Productos y Procesos
- 30 años - Landsat  
**300 TB**



### Fase 3...

- Adquisición constante (mantener al día las imágenes)
- 3 años - Sentinel  
**300 TB**



Cronología  
del proyecto  
+ dimensionamiento  
**+ diseño**



- Definir primeros productos
- Planear la adaptación de productos y procesos y su difusión
- Diseñar la implementación de la infraestructura



CUBO DE DATOS  
GEOESPACIALES DE MÉXICO

## 15.4.2 Mountain Green cover



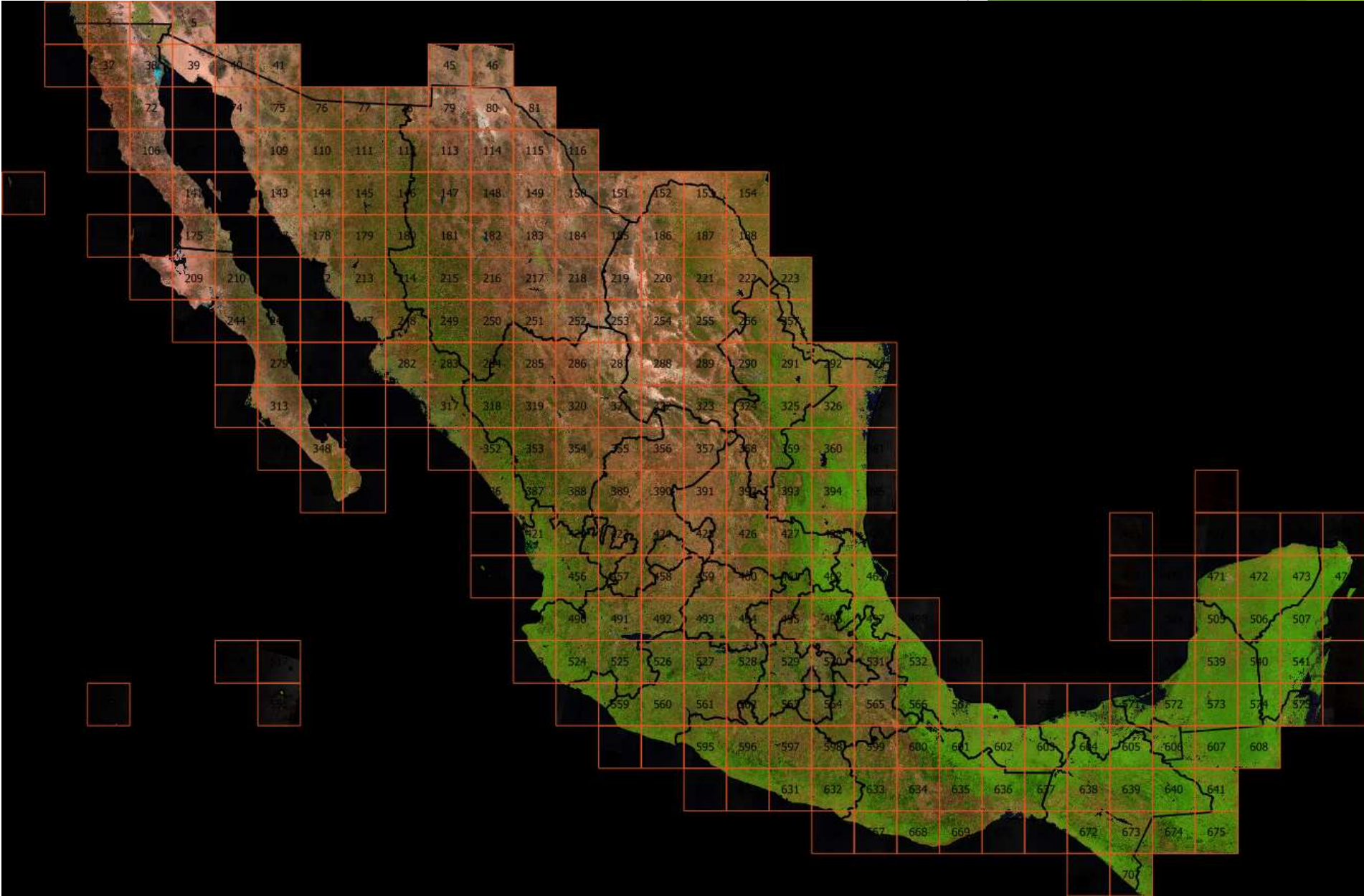
La primera clasificación es una conversión del Mapa de USV a 6 clases

Los procesos de ODC permiten actualizaciones constantes a la clasificación nacional porque se generan automáticamente

Se facilitaría una verificación de campo para evaluar la calidad de los datos

Actividad	Sin ODC		Con ODC	
	Progreso	Progreso	Progreso	Progreso
Usar definiciones de Intergovernmental Panel on Climate Change (6 clases)	✓	✓	✓	✓
Usar Mapa de USV	✓	✓	✓	✓
Convertir la clasificación (original -> 6 clases)	✓	✓	✓	✓
Obtener muestra de datos convertidos			✓	diseño
Usar muestra y 6 otros indicadores (ODC) como db de entrenamiento para la clasificación			✓	indicadores ODC (geomedian)
<b><u>Realizar una clasificación nacional con with Machine Learning</u></b>			✓	
Cruzar el ráster resultante con el Modelo Digital de Elevación (áreas montañosas)	✓	✓	✓	
Calcular el índice Green Cover en el MDE	✓	✓	✓	
<b><u>Posible validación en campo para aseguramiento de la calidad en la db de entrenamiento</u></b>			✓	
Retroalimentación a FAO	✓	✓	✓	

# Geomediana 2015



► 283 celdas









Gracias  
por su atención.

