Using Remotely Sensed Data and ArcGIS to Automate and Improve Census Bureau Geographic Update Activities

Arthur Desch
U.S. Census Bureau
Geography Division
Geographic Research and Innovation Staff
July 12, 2018

United States Census Bureau | U.S. Department of Commerce Economics and Statistics Administration U.S. CENSUS BUREAU census.gov

No Title 13 data in this presentation

Goals & Benefits

- To improve data quality for the U.S. Decennial Census and survey operations.
  - Validate existing collection of housing units and roads (MAF/TIGER).
  - Detect change (new roads/buildings) to identify areas for more intensive study.
- Reduce the amount of in-field work needed, focus on-going update work.

Block Assessment Research and Classification Application (BARCA)

- An in-office data review system using visual inspection of high resolution imagery to identify Census blocks in need of update

- Estimated that BARCA will be able to identify around 75% of the blocks needing active updating

BARCA Interface

No Title 13 data in this presentation

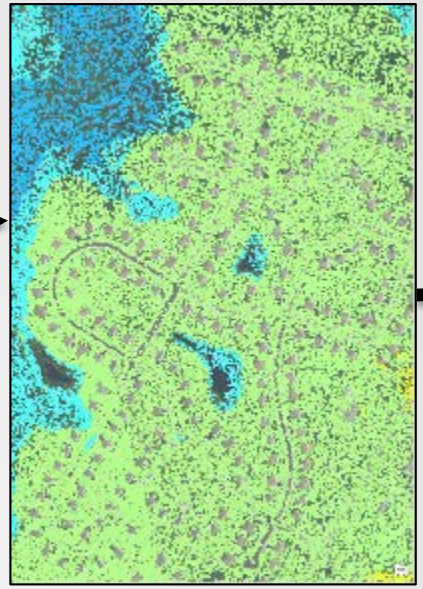
Interactive Review Automation

- Create a Geoprocessing tool that will utilize a combination of aerial imagery, LiDAR, and existing Census GIS data to identify Census blocks that have experienced change in the built environment since the last update and prioritize the blocks by amount and type of change for editing by a human analyst.

Imagery + LiDAR = Buildings

2011 NAIP Imagery

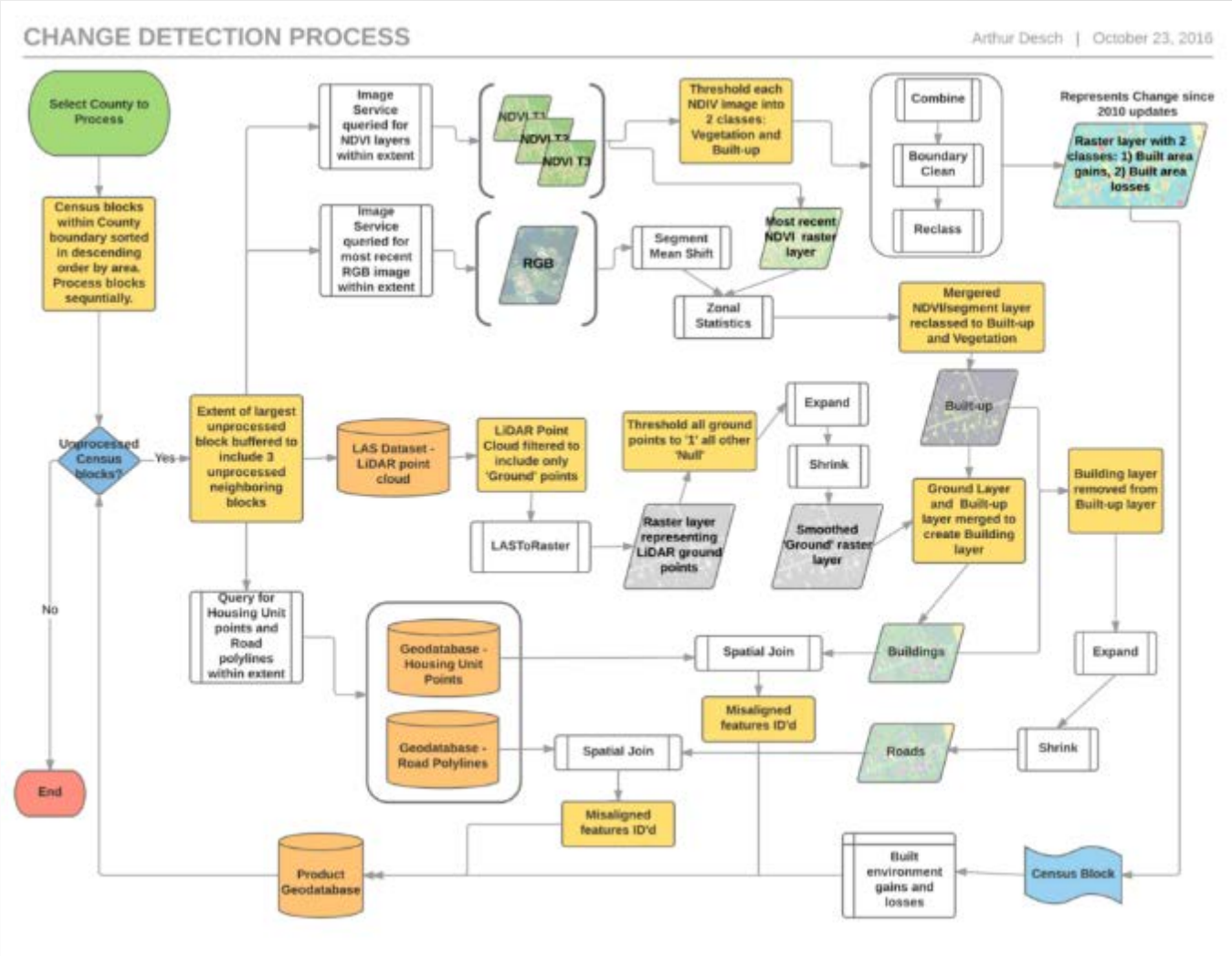
2016 NAIP Imagery



2016 LiDAR Point Cloud

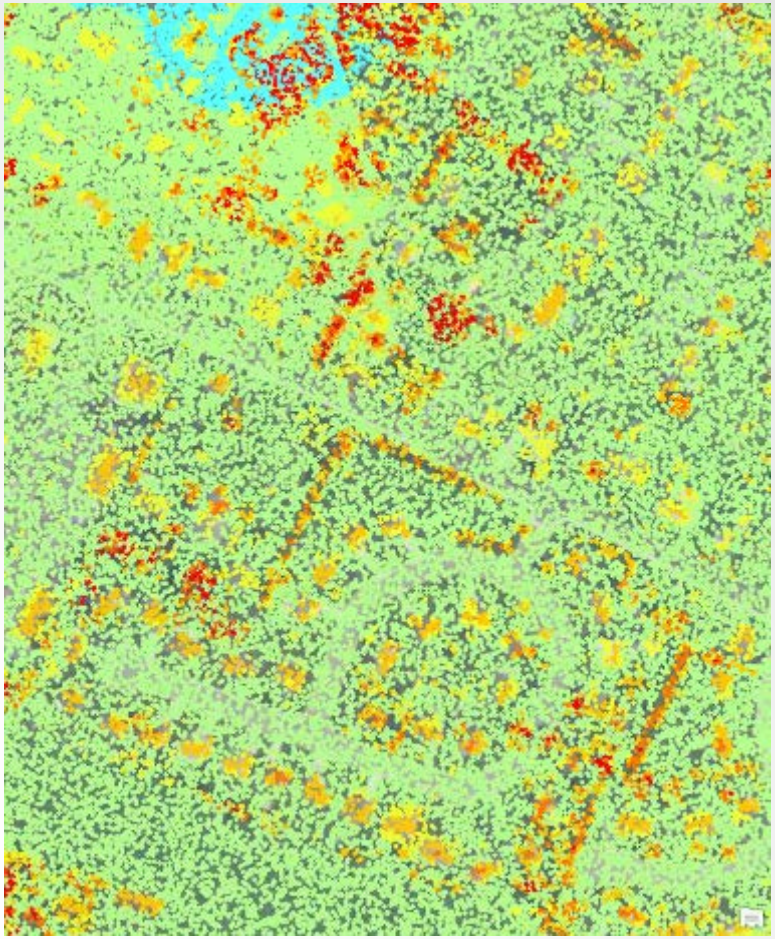
New Building Footprints

Python Processing Flow

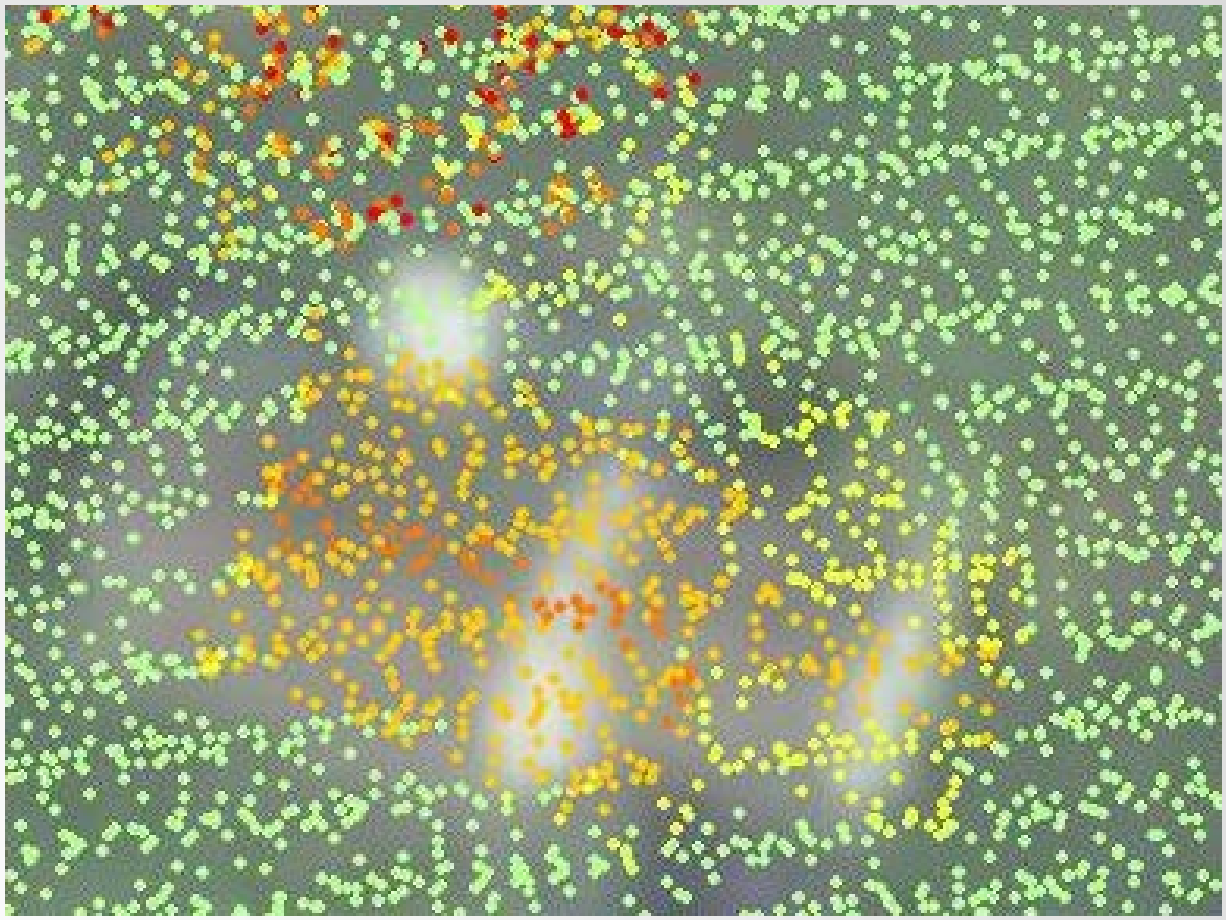


Processing Inputs

LiDAR Point Cloud



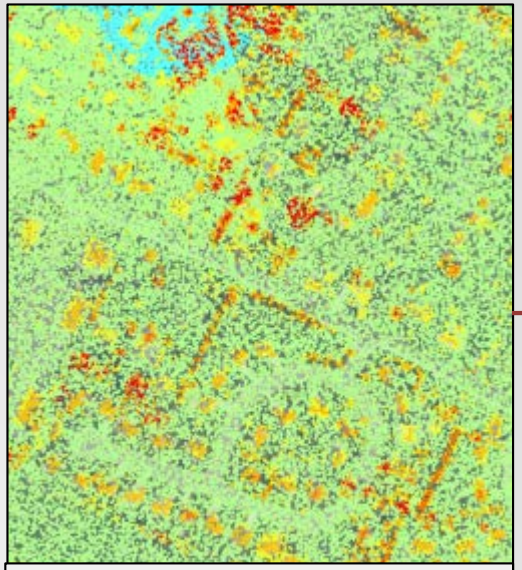
NAIP Imagery – 4 band



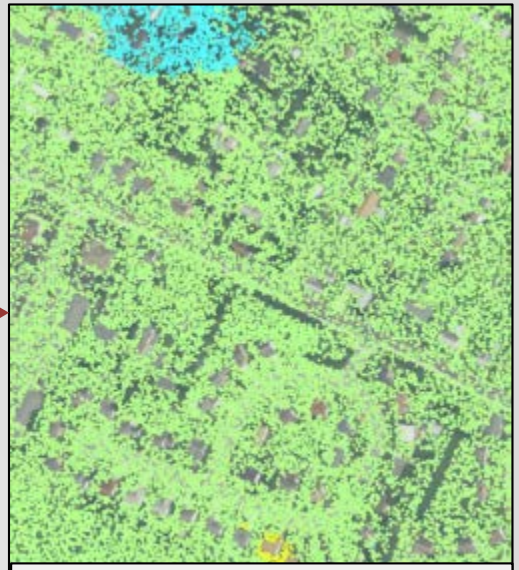
LiDAR point cloud – full sample

Processing - LiDAR

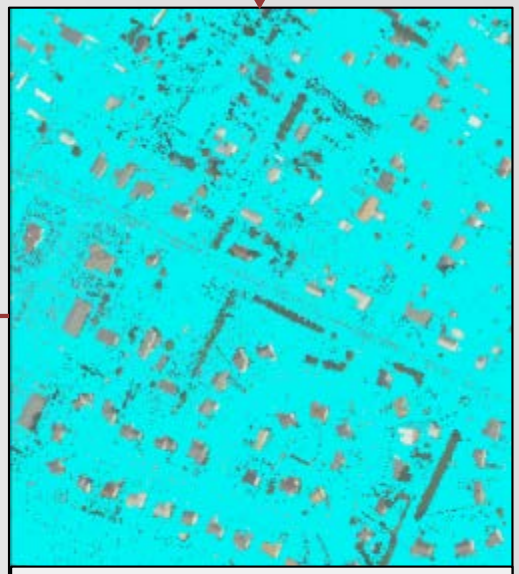
- LiDAR point cloud queried for ground points
- Points converted to raster
- Expanded to fill small gaps



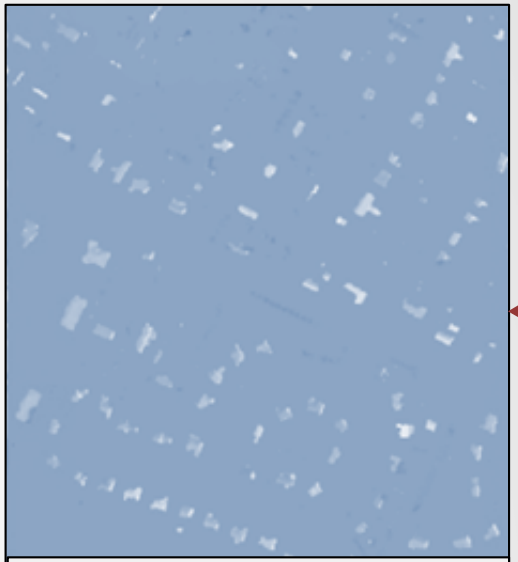
LiDAR point cloud



LiDAR 'Ground' points



'Ground' points as raster layer

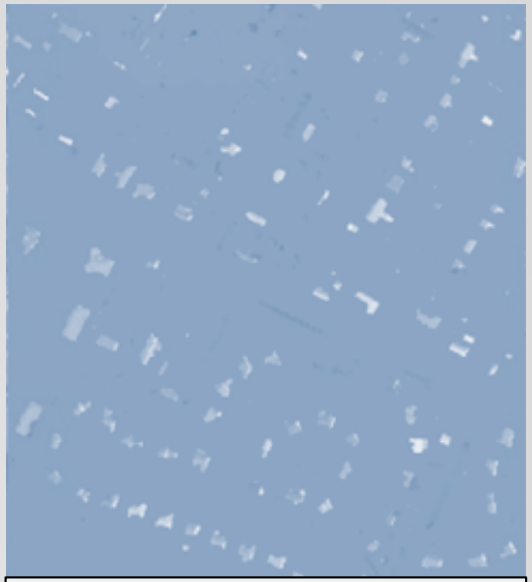


Expanded Ground raster layer

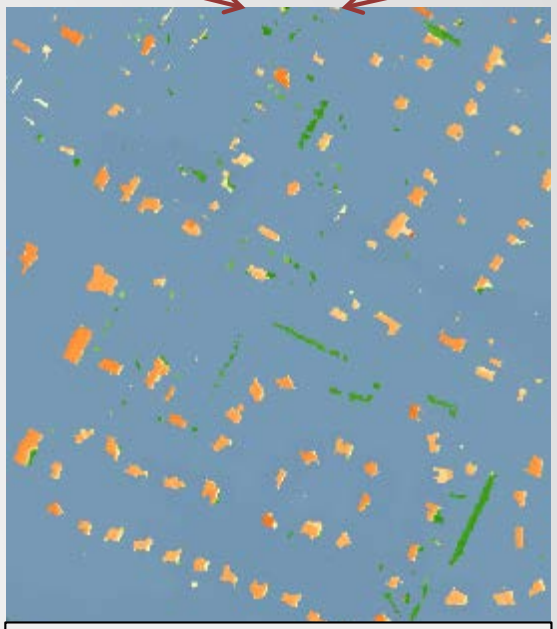
Processing – LiDAR: Buildings

- The resulting Ground Mask layer is 'transparent' where higher objects exist. Overlaying this layer on a Normalized Difference Vegetation Index (NDVI) of NAIP imagery highlights which spaces are buildings and which are trees.

NDVI layer



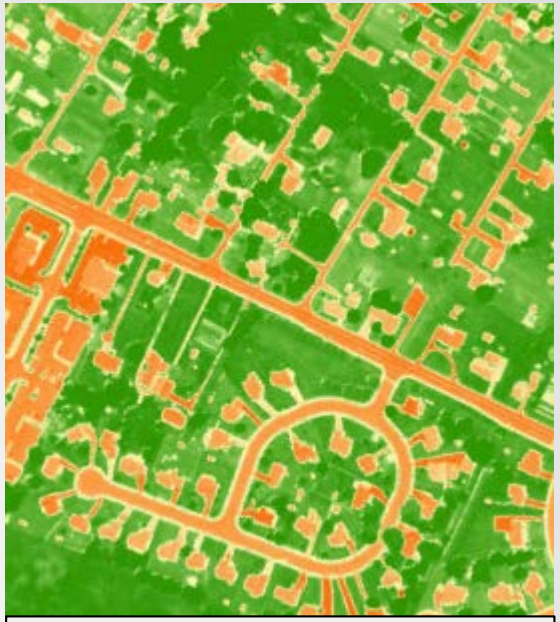
Ground Mask layer



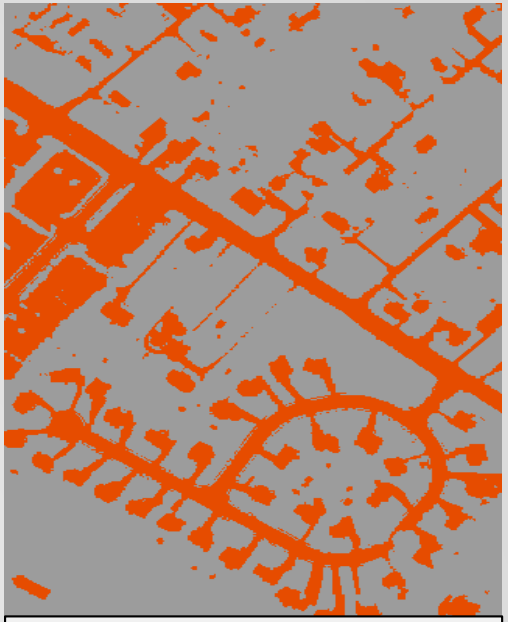
Ground Mask over NDVI

Processing – LiDAR: Buildings

- The NDVI layer is converted to a Built/not Built layer



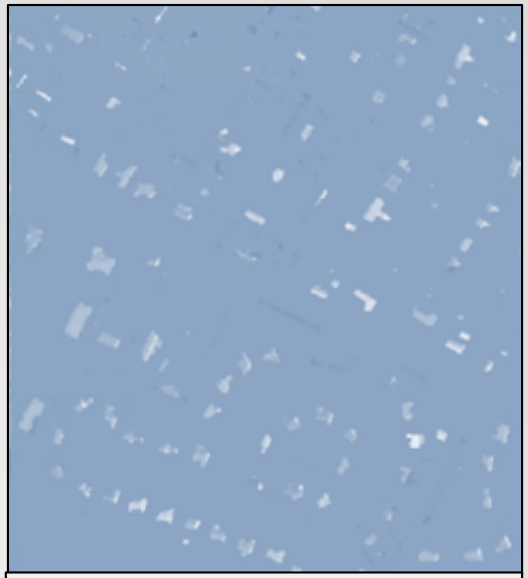
2015 NDVI layer



2015 Built layer

Processing – LiDAR: Buildings

- Expanded Ground raster used as mask to remove low areas of Built/Not Built layer
- Resultant raster layer represents primarily buildings



'Ground' points as raster layer

2015 Built layer

'Buildings' raster layer

Processing – LiDAR: Buildings

- Building raster converted to polygons
- Small polygons deleted, other noise filtered

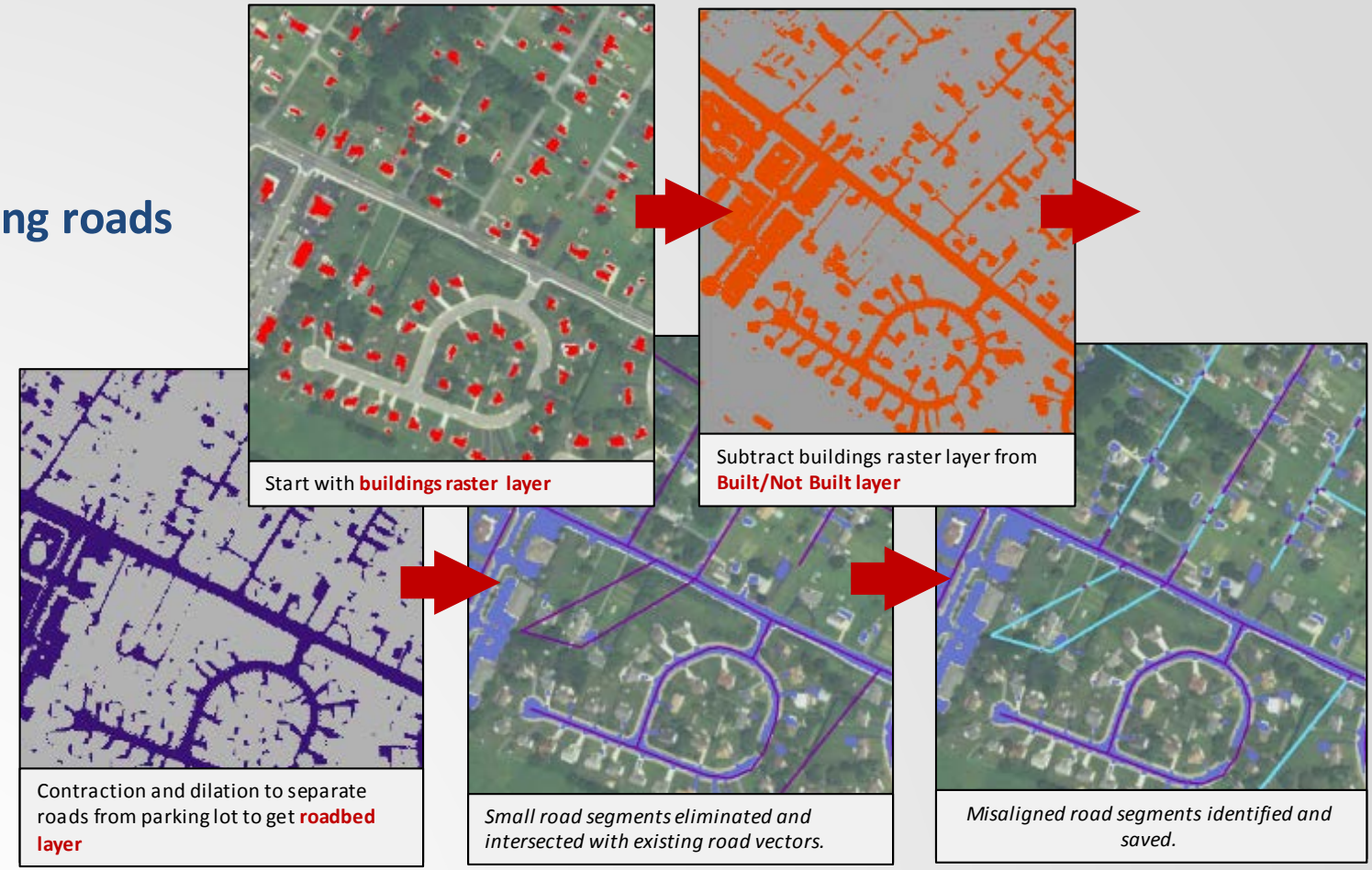
'Buildings' Polygons

Processing – LiDAR: Buildings

- Building polygons spatially joined to housing unit points to identify points that do not intersect building footprints
- Misaligned points written to output geodatabase
- Building footprints and approximate heights written to output geodatabase

'Buildings' Polygons

Validating roads



Block Product

- Number of misaligned features, total features found by tool, and total from MAF/TIGER for each Census Block within the extent are tallied and written to output geodatabase
- Total and relative amount of area change per Census Block written to output geodatabase
- Total and relative amount of change can be used to prioritize assignment to human editor

Interactive Review Comparison

- 95% agreement for blocks identified as ‘Active’
  - Most disagreements result from older LiDAR
- 65% agreement on ‘Passive’ blocks
  - Tool IDs more buildings
  - Tool can’t differentiate residential vs commercial
  - Multi-unit buildings difficult to flag correctly
- Human – ~90 sec/block Program – ~90 sec/block

Data Currency is Problematic

Similar Datasets Available

'Building' Polygons – Census Tool

'Building' Polygons – ORNL

Developing Approaches to Identifying Change

- Needs to utilize data that is current enough to be relevant to Census needs
- Needs to be at fine enough resolution to identify building level change
- Need to be able to process data for entire country quickly
- Needs to be cost effective

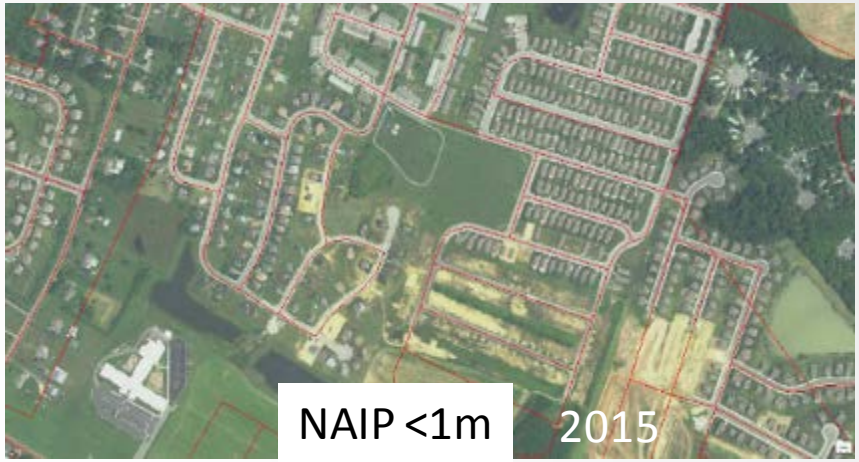
Landsat – Not enough spatial resolution

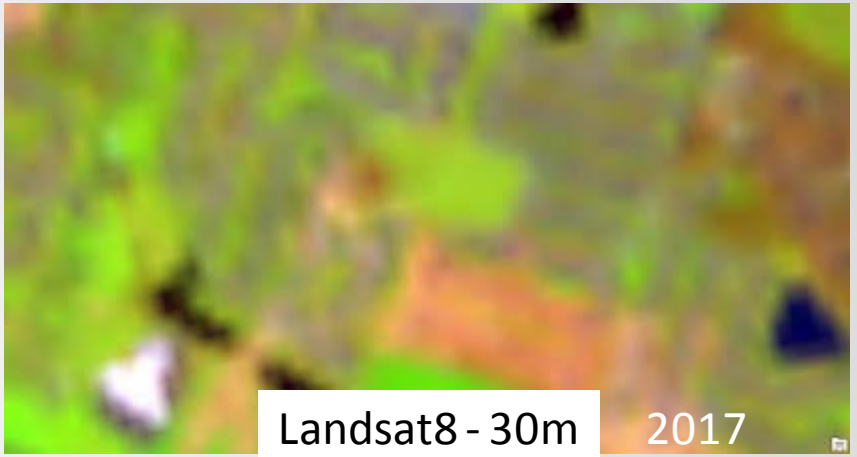
Sentinel 2

- 13 spectral bands
- Four bands at 10 m, six bands at 20 m and three bands at 60 m spatial resolution
- Orbital swath of 290 km, product provided as 100 km^2 tiles
- Level-1C tiles are orthorectified and corrected to TOA reflectance
- Tools available to correct to BOA
- 5 day repeat
- FREE - Available from ESA and USGS

Sentinel 2 – 10m Change Signal?

Sentinel 2 - 10m





Sentinel 2 – 10m Change Signal!

May 2016

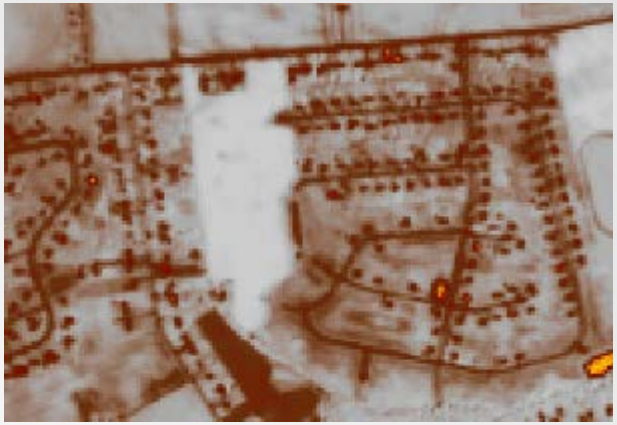
September 2017

NDVI May 2016

NDVI September 2017

New housing

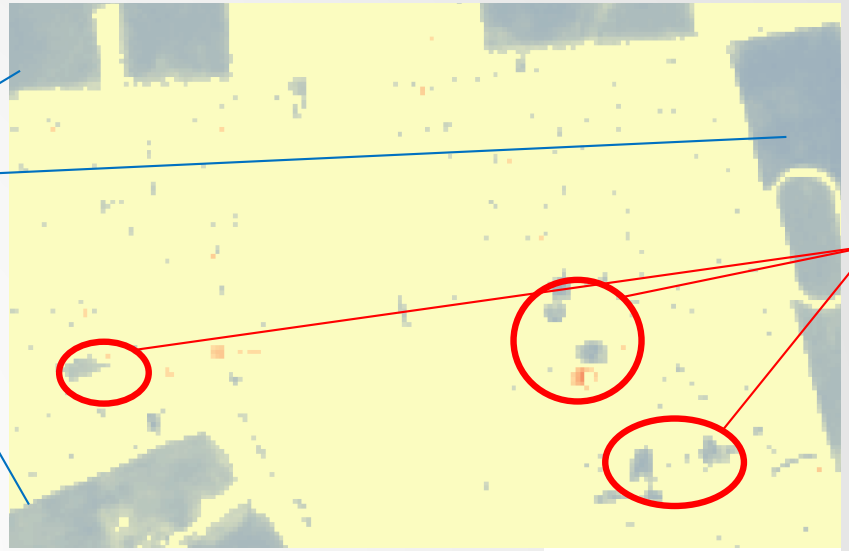
Sentinel 2 – 10m Change Signal



NDVI May 2016

NDVI September 2017

Agricultural fields

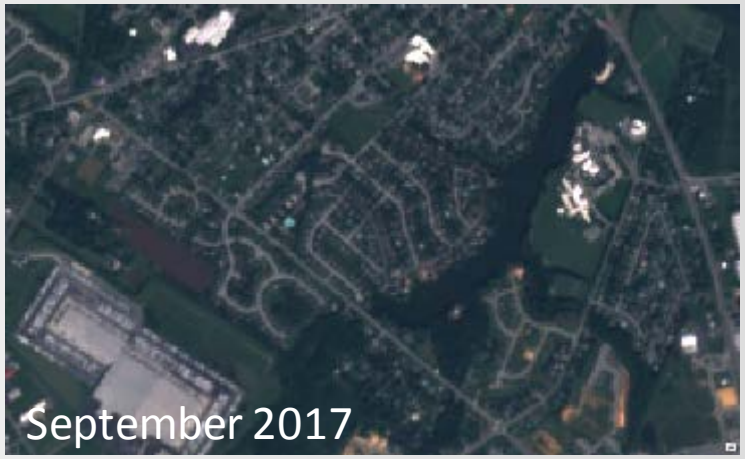


New housing

NDVI Change

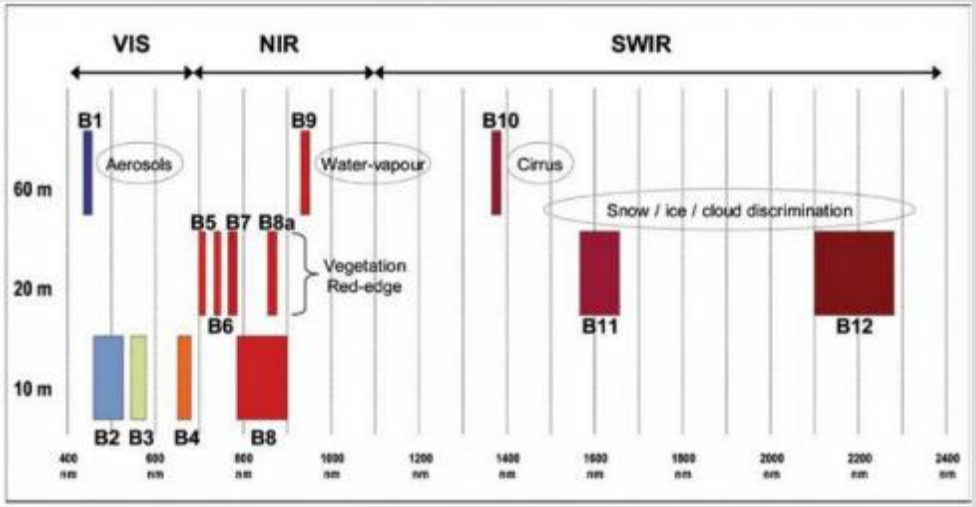
Sentinel 2 – 10m Change Signal

Detectable at time and space scales relevant to Census needs?



Challenge: Image Data Are Huge!

- Multiple dimensions to consider:
  - Spatial resolution
  - Spectral resolution
  - Temporal resolution
- Example numbers using Sentinel-2 (10m^2 spatial resolution):
  - ~80 trillion pixels of entire planet.
  - ~98 billion pixels of the U.S.
- Each pixel has 13 bands
  - 10-60m^2 resolution
- And that's just for a single pass!
  - Updated every 5 days.



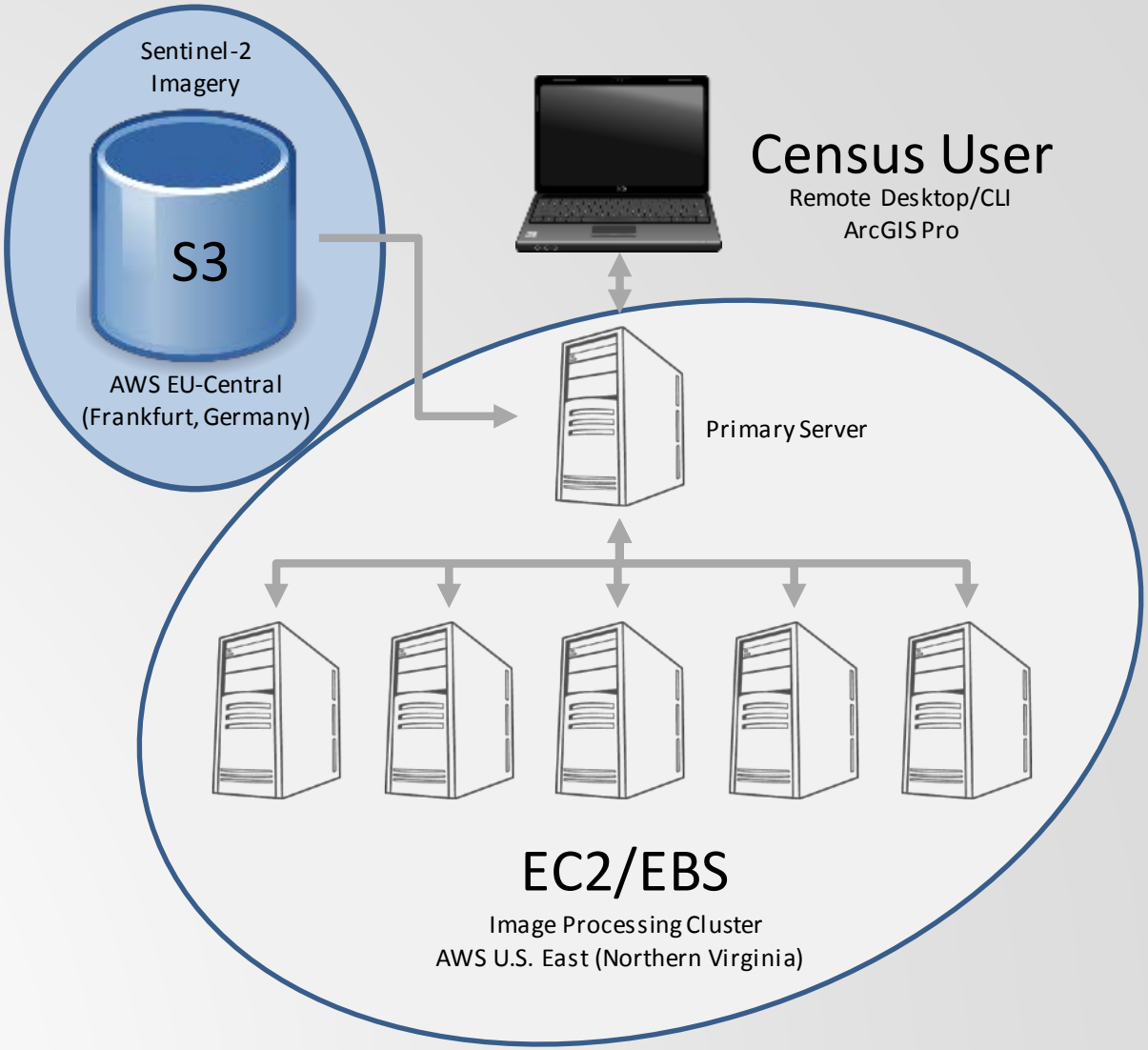
Prototype Architecture

Cloud environment:

- Amazon Web Services (AWS)
  - Elastic Compute Cloud (EC2)
  - Simple Storage Service (S3)
  - Elastic Block Store (EBS)
- Commercial Cloud vs. GovCloud

Image processing tools:

- ArcGIS Pro
- ArcGIS Enterprise
  - ArcGIS Server
  - ArcGIS Image Server



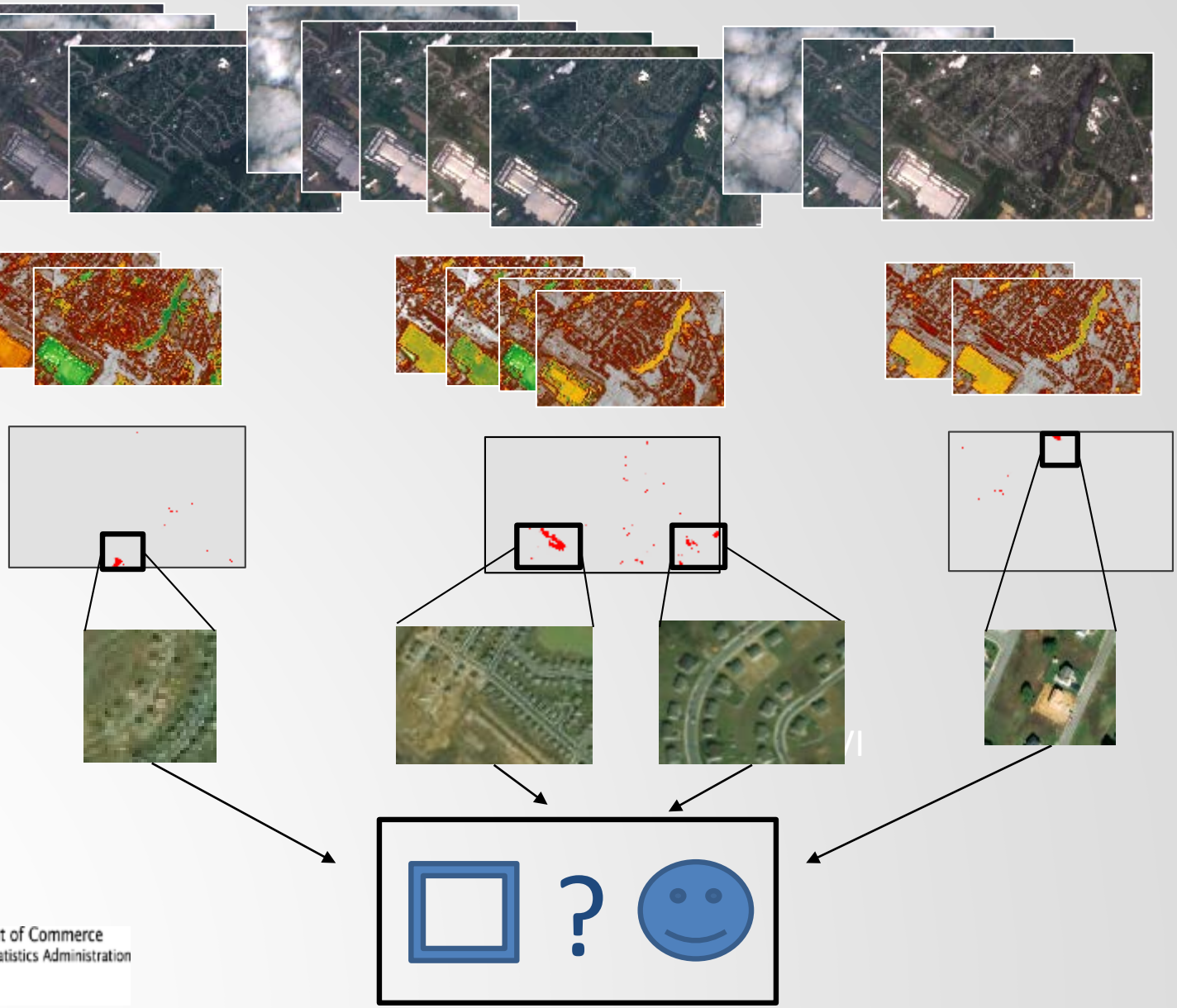
Processing Pyramid

Sentinel Imagery

Data Reduction/ Analysis

Persistent Change Detection

Hi-Res Data Analysis



Thank you!

Kip Desch

arthur.desch@census.gov