Quantifying Analyst Bias in Mapping Flooded Areas from Landsat Images

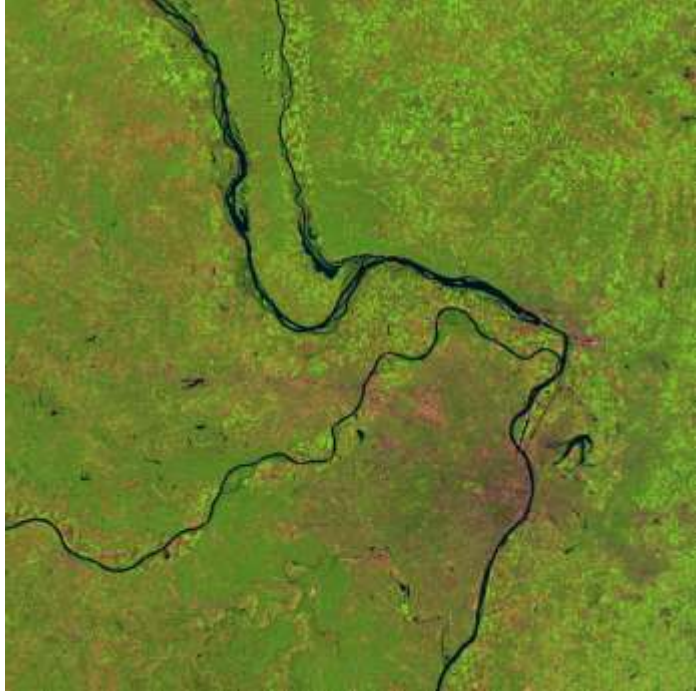
Caleb McCarragher[1], Dr. Ramesh Sivanpillai[2, 3]

1. Department of Geology, 2. Department of Botany, 3. Wyoming Geographic Information Science Center

2017 Undergraduate Research Day

University of Wyoming

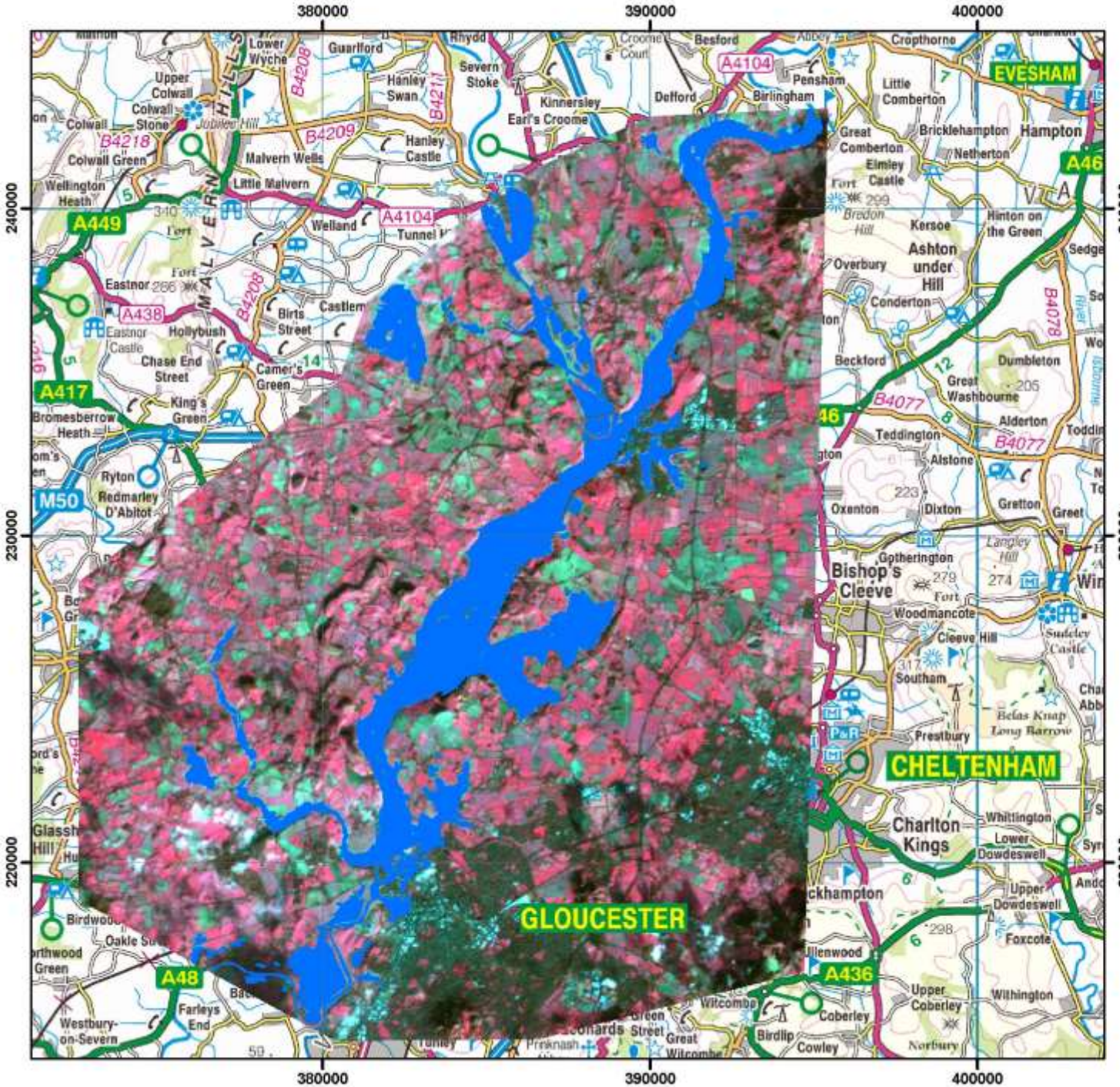
29 April 2017



https://earthobservatory.nasa.gov/IOTD/view.php?id=5422

Purpose

- Remotely sensed images are widely used for post-disaster response and recovery
  - Satellites, airplanes, and drones
- Satellites continuously collect data
  - Pre-flood and post-flood comparison
- Info derived from satellite data are combined with maps



Flood mapping of River Severn derived UK DMC-2 satellite data acquired on 07/01/2014.

Legend

Water Extent

UK DMC-2 Image

Red: NIR band
Green: Red band
Blue: Green band

0 2 4 8 kilometres

N

The water extent was estimated from UK DMC-2 satellite data with 22 metre resolution acquired at 10:15 UTC on 07/01/2014 . The extent of the analysis is limited to the area covered by the imagery shown.

These data were used to map the water extent within the flood analysis area. The image contained a great deal of cloud and cloud shadow and so the analysis area is restricted.

Some of these areas are water under normal conditions (e.g. rivers or lakes).

Environment Agency Geomatics
Environment Agency, National Operations

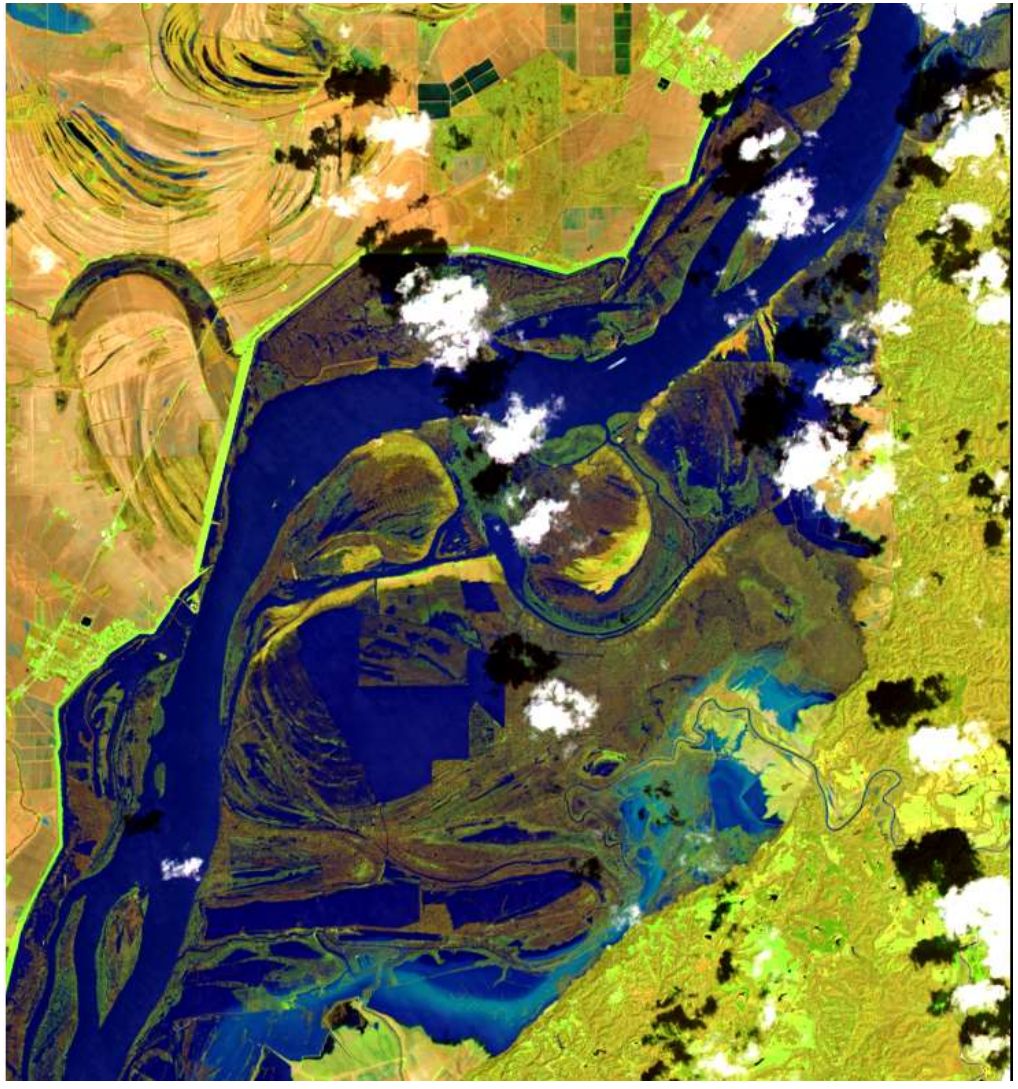
Map generated by Environment Agency, National Operations, Geomatics.

Purpose

- Remotely sensed images are widely used for post-disaster response and recovery
  - Satellites, airplanes, and drones

- Raw satellite images cannot be used effectively
  - Even using false color images, water can be challenging to differentiate

Satellite data in different band combinations

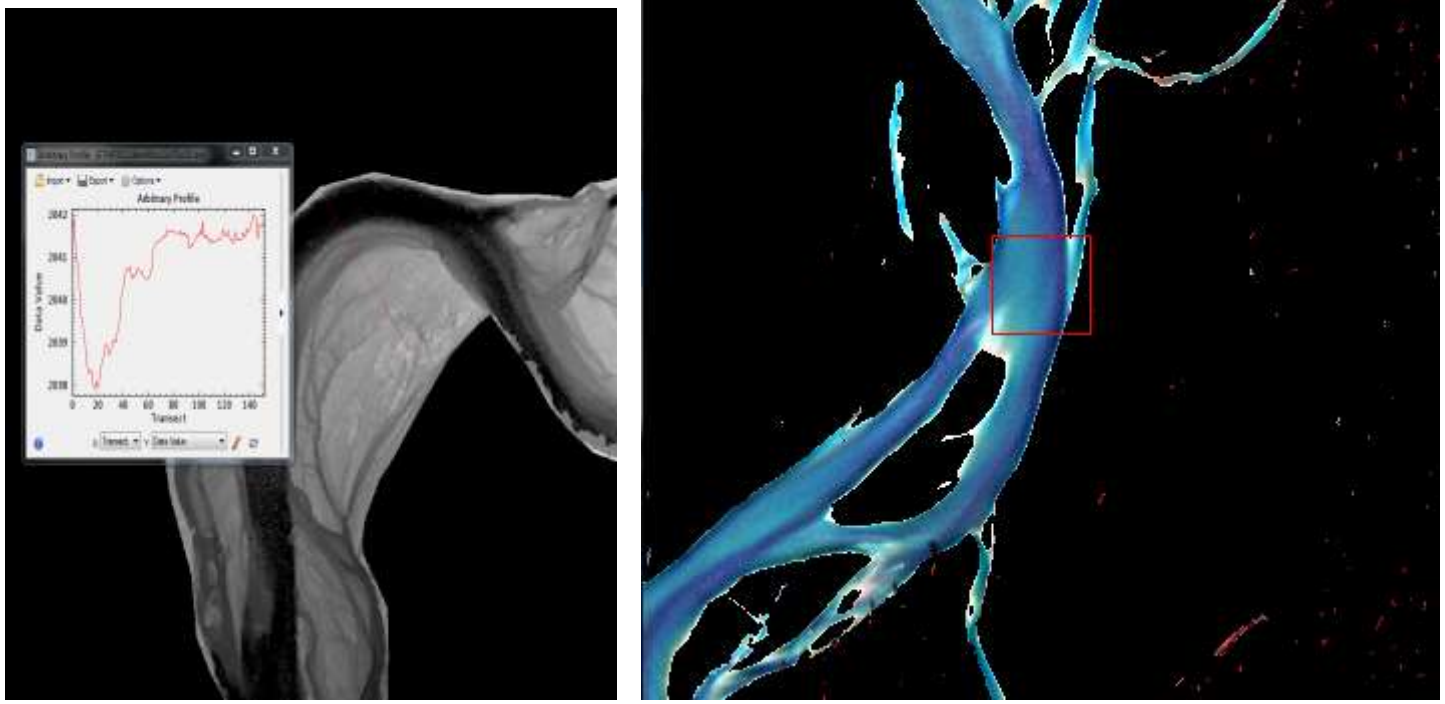


Purpose

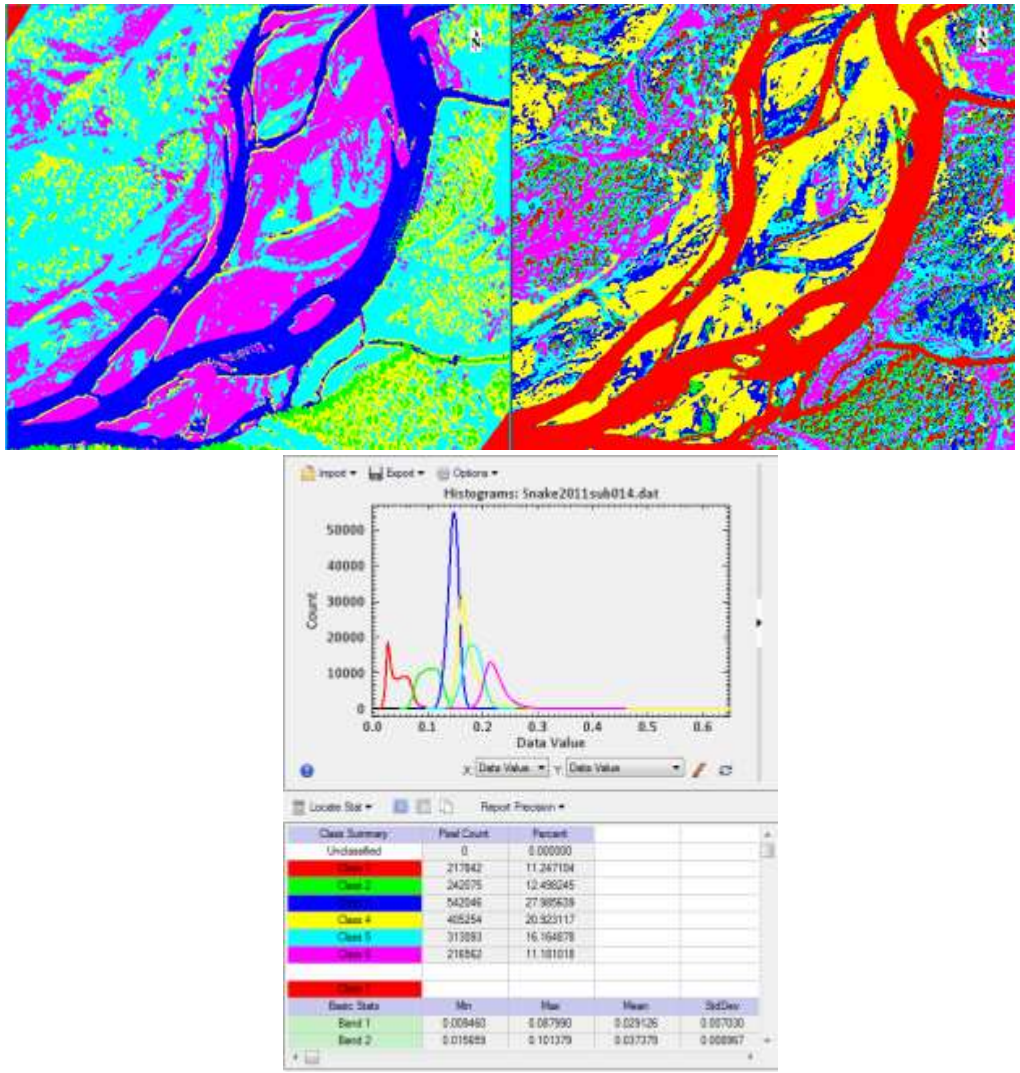
- Remotely sensed images are widely used for post-disaster response and recovery
  - Satellites, airplanes, and drones
- Raw satellite images cannot be used effectively
  - Even using false color images, water can be challenging to differentiate
- Processing ranges from simple to advanced
  - However processing time is critical (lives are at stake)

Mapping water extent

Water extent in Snake River, Wyoming



Extract waterbodies from imagery (method 1)

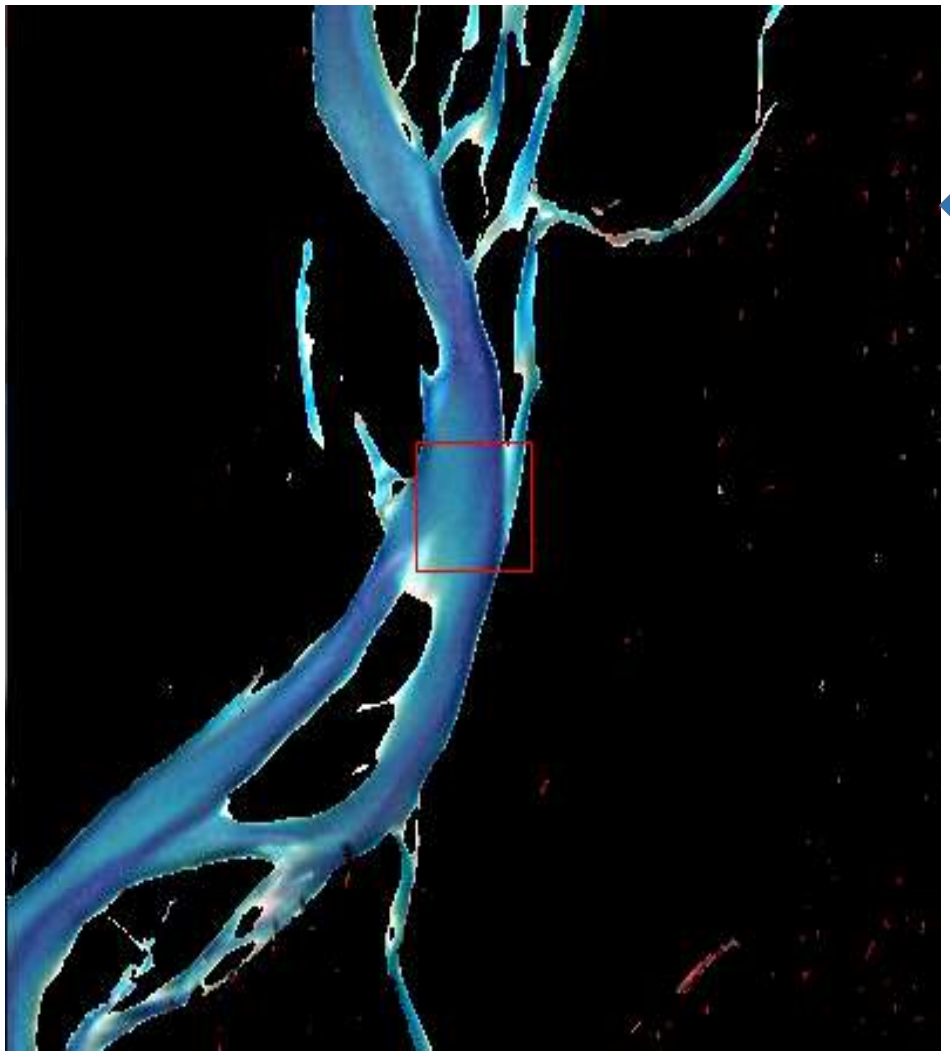


Extract waterbodies from imagery (method 2)

Purpose

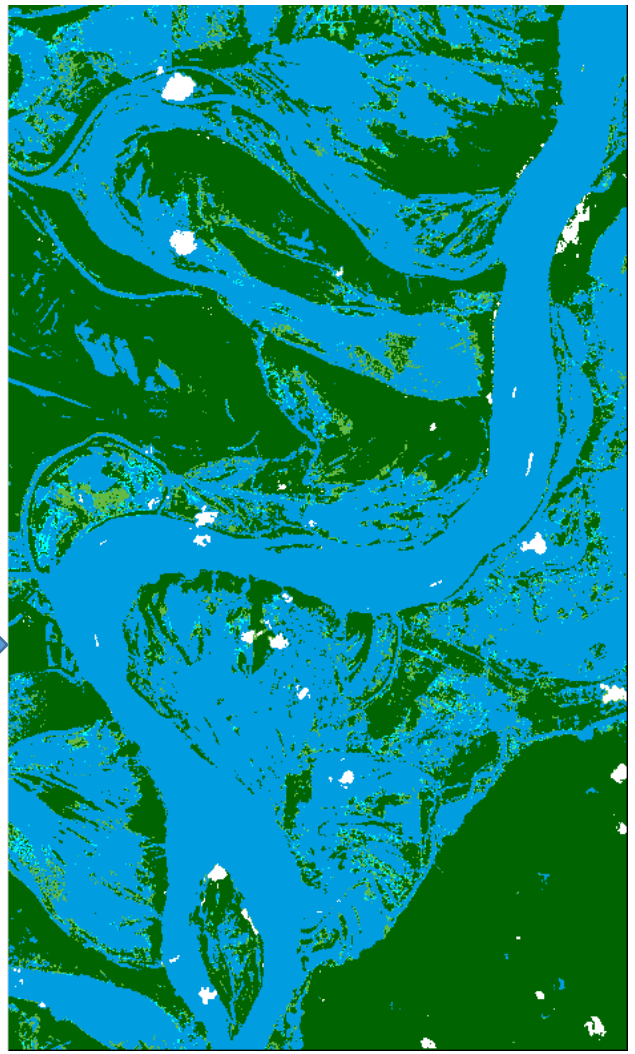
- Remotely sensed images are widely used for post-disaster response and recovery
- Raw satellite images cannot be used effectively
- Processing ranges from simple to advanced
- In the event of a disaster, volunteers are solicited for processing imagery
  - We cannot expect everyone to have similar expertise and experience
  - Classification output will be different
  - Some methods will need reference data

Classification output



Water depth

Standing water



Satellite data for disaster mapping

- Goal: Process them rapidly and deliver to responders

- Challenges
  - Raw images have to processed to extract useful information
  - Numerous methods and expertise of volunteers' varies

  - For large disasters, several volunteer analysts will classify images
  - Decision making could be problematic

Objective

- Quantify the variability among analysts
  - If the same image is processed by different analysts, how much variability will be there?

- Based on the findings from this study:
  - What level of instruction must be provided to minimize variability?

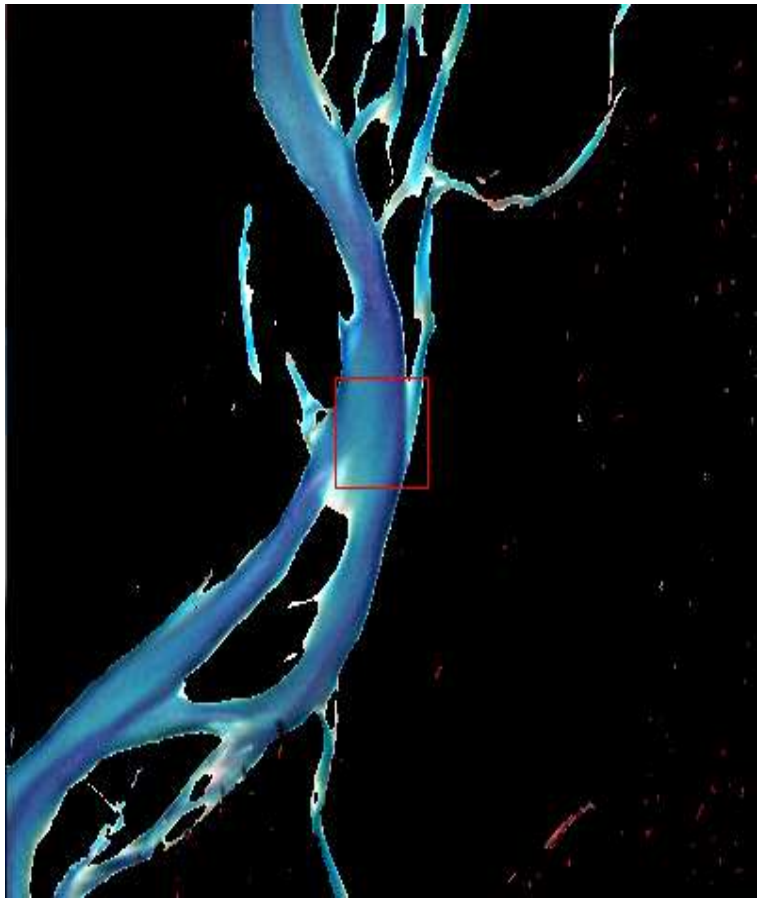
Methods

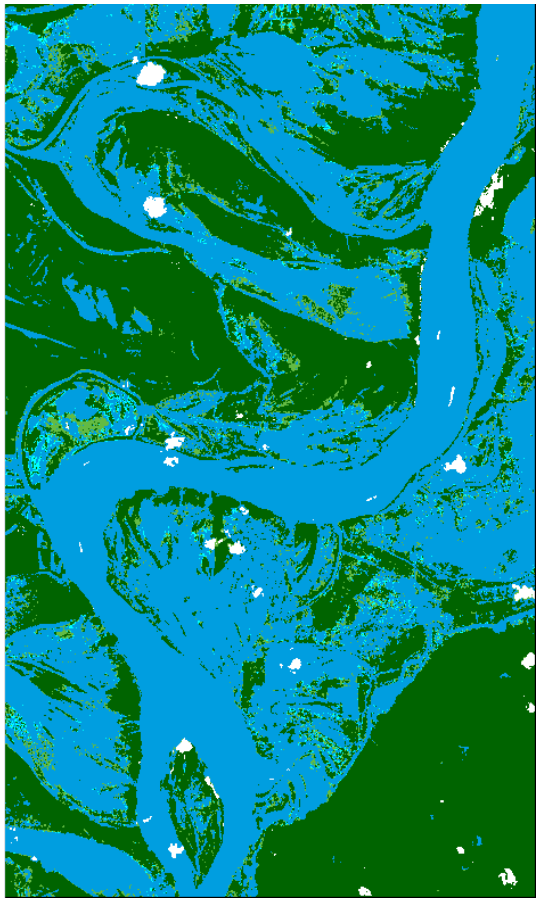
- 4 post-flood satellite images
  - 2, 2011 US Midwest floods, and 2, 2016 US Southeast floods
  - 2011 – Landsat 5 (both images were cloud free)
  - 2016 – Landsat 8 images (one had few clouds, the other more clouds)

- 6 analysts with similar level of experience and expertise
  - Goal was to identify areas inundated by water
  - Processed these images in ERDAS Imagine (Unsupervised classification)
  - No specific directions were provided about the parameters

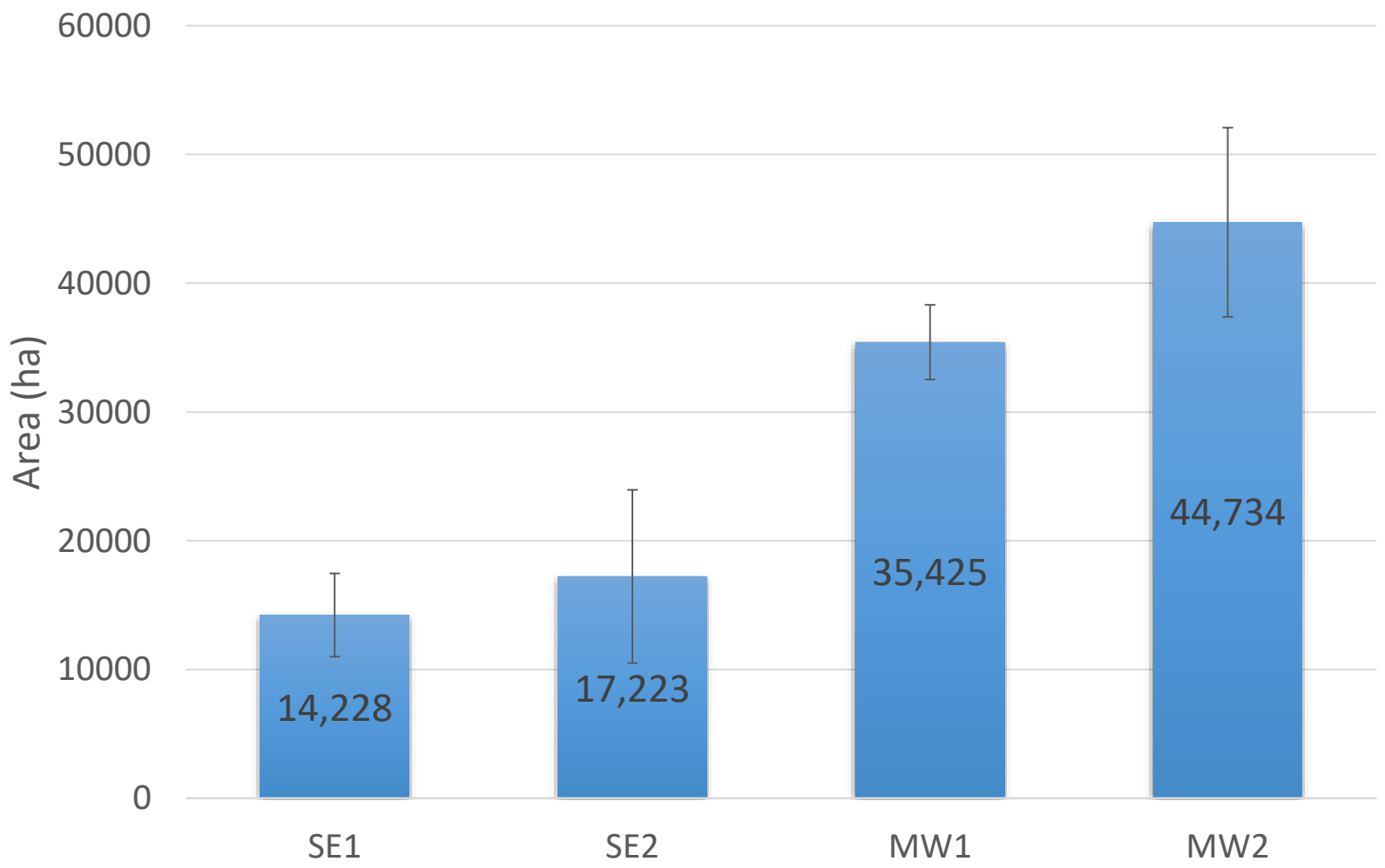
Methods

- Classification is sorting pixels into different clusters based on their spectral values
  - Several classification techniques
  - Some are simpler than others





Results – variability among analysts

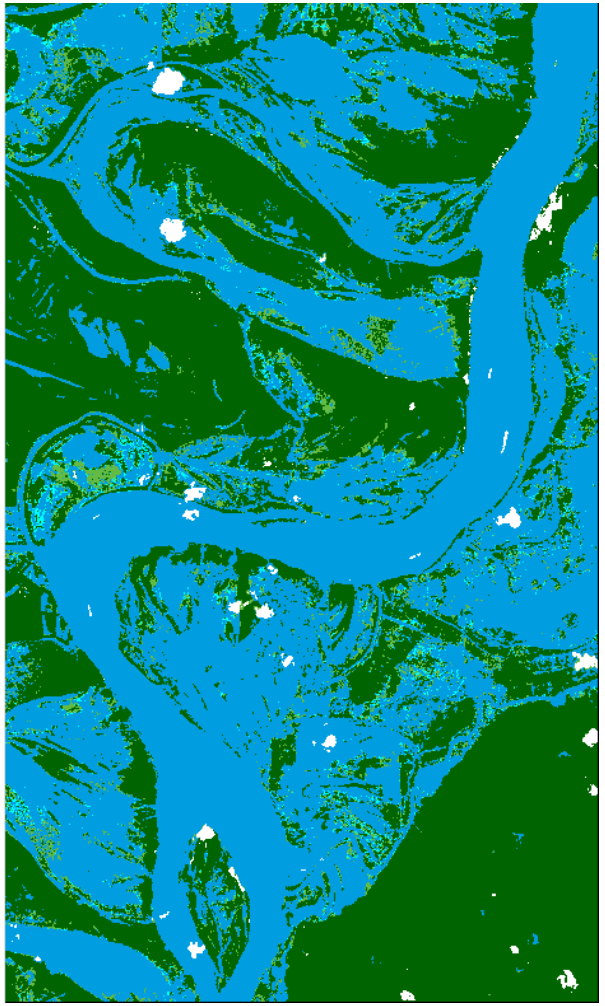


Southeast 1

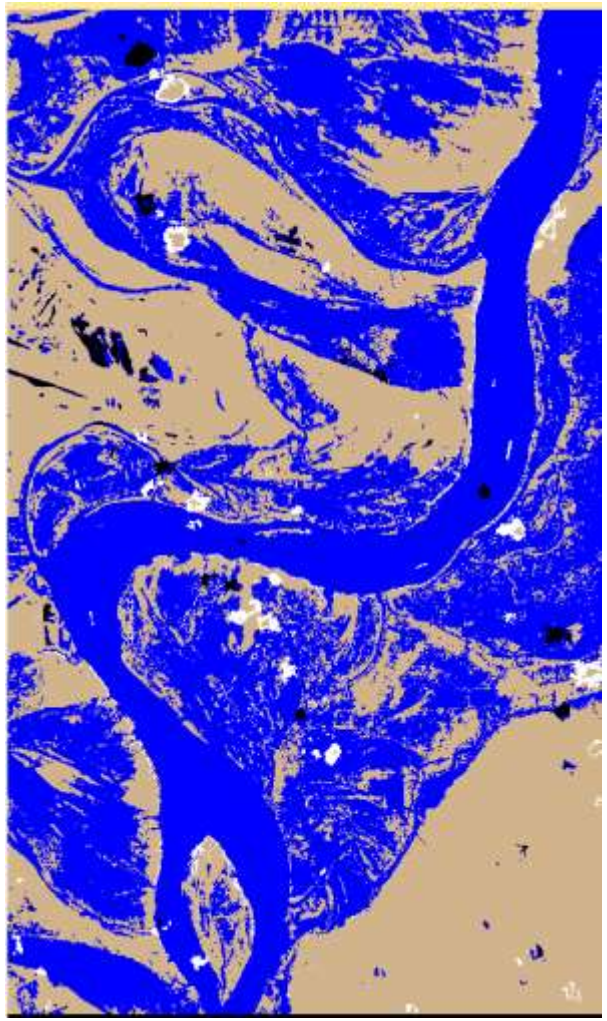
Landsat 8 image

SR – 200 clusters and assigned them to 3 thematic classes

CM – 100 clusters and assigned them to 3 thematic class



SR



CM

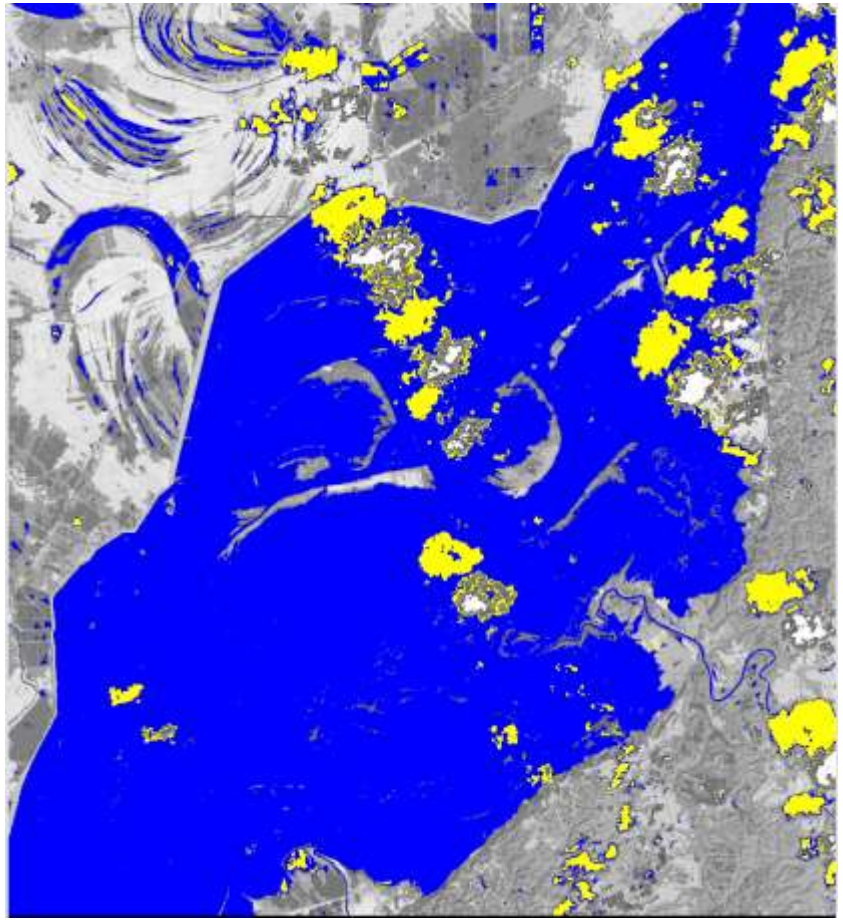
Minimum difference

Southeast 2

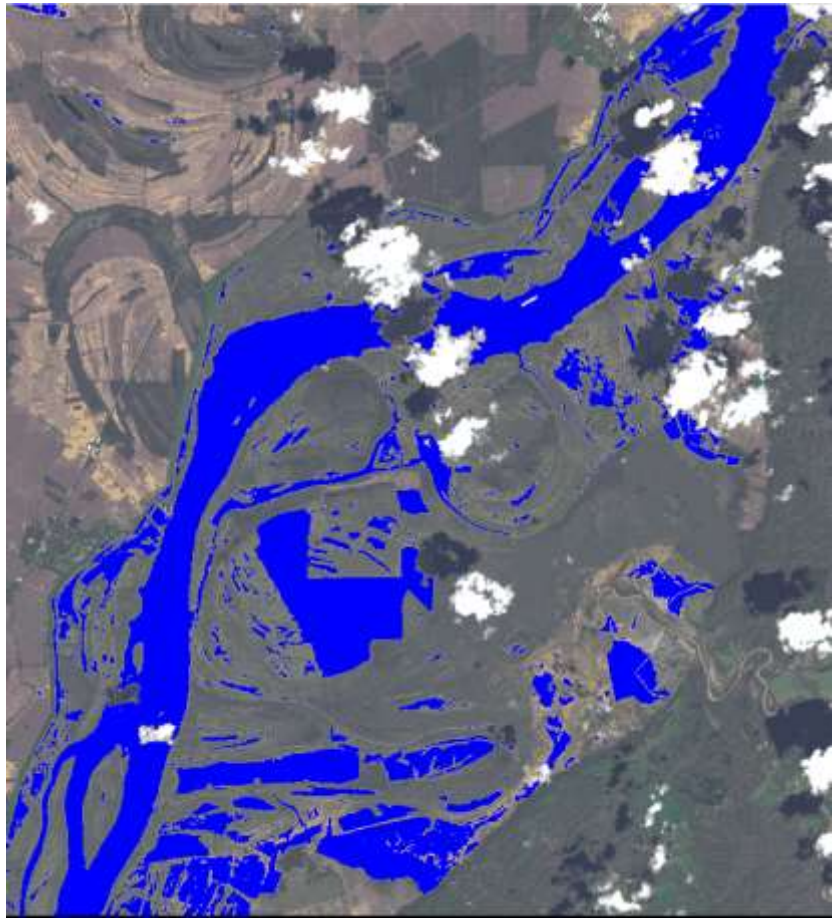
Landsat 8 image

SB generated 150 clusters and assigned to water and non-water classes

JW generated 30 clusters and assigned them to water, cloud, and non-water classes



JW (most)



SB (least)

Maximum difference

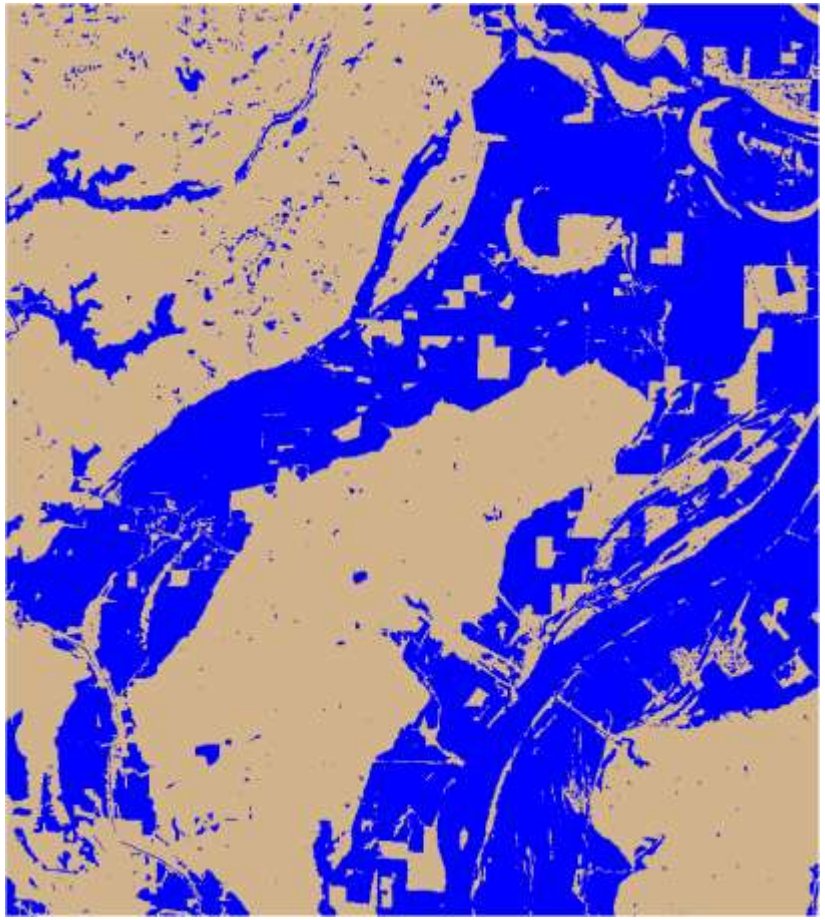
Level of detail (thematic classes)

Midwest

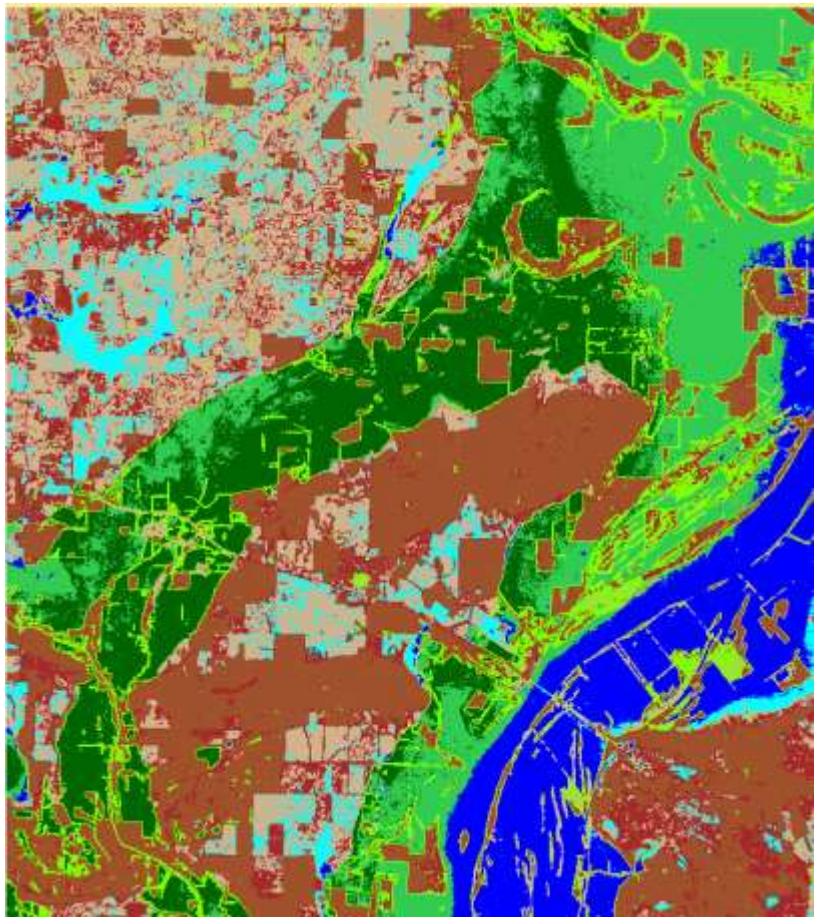
Landsat 5 image

CM generated 100 clusters and assigned to water and non-water classes

JL generated 200 clusters and assigned them to 8 types of water and other non-water classes



CM - 3 classes, 2 colors for water and land



JL – 8 classes, 8 colors of water and land

Results

- Variation in naming convention
  - High, low or medium confidence on identifying water
  - Clear, shallow and turbid water
  - Land-water mixed class
  - One analyst used symbols (colors) but did not name the clusters

Observations

- Analysts did not classify the 4 images the same way
  - Difference in number of clusters
    - One analyst started with 200 and ended with 50
  - Difference in colors and classes
    - One analyst started with 8 colors and ended with 3
  - Difference in naming
    - One analyst forgot or neglected to name the water classes preferring to use colors
    - Only happened in one of their images
  - Fatigue factor
    - How much time of day and number of images classified in one sitting make a difference

Conclusions and recommendations

- Medium to high variability was noticed
  - Area of the flood in the image was not a factor
- Naming convention varied among analysts
  - This could be a problem for large floods involving multiple analysts
- International Charter on Disasters has to develop some protocols
  - Minimum of number of clusters generated (too few are not enough)
  - Name of the classes have standardized (clear water, turbid water, …)
    - Color schemes can be standardized
  - Cloud masks can be applied to exclude areas
  - Degree of confidence could be a different attribute (high, medium, low)

Thanks

Questions