Introduction to Lidar Technology and Data Collection

Christopher Crosby

San Diego Supercomputer Center / OpenTopography

(with content adapted from NCALM, David Phillips (UNVACO), Ian Madin (DOGAMI), Ralph Hagerud (USGS), and Dave Harding (NASA))

2009 SCEC LiDAR Short Course:

Using GeoEarthScope and B4 LiDAR data to analyze Southern California's active faults

December 3-4, 2009

LIDAR / LiDAR / lidar / ALSM… = light detection and ranging

• Billions of of accurate distance measurements with a laser rangefinder

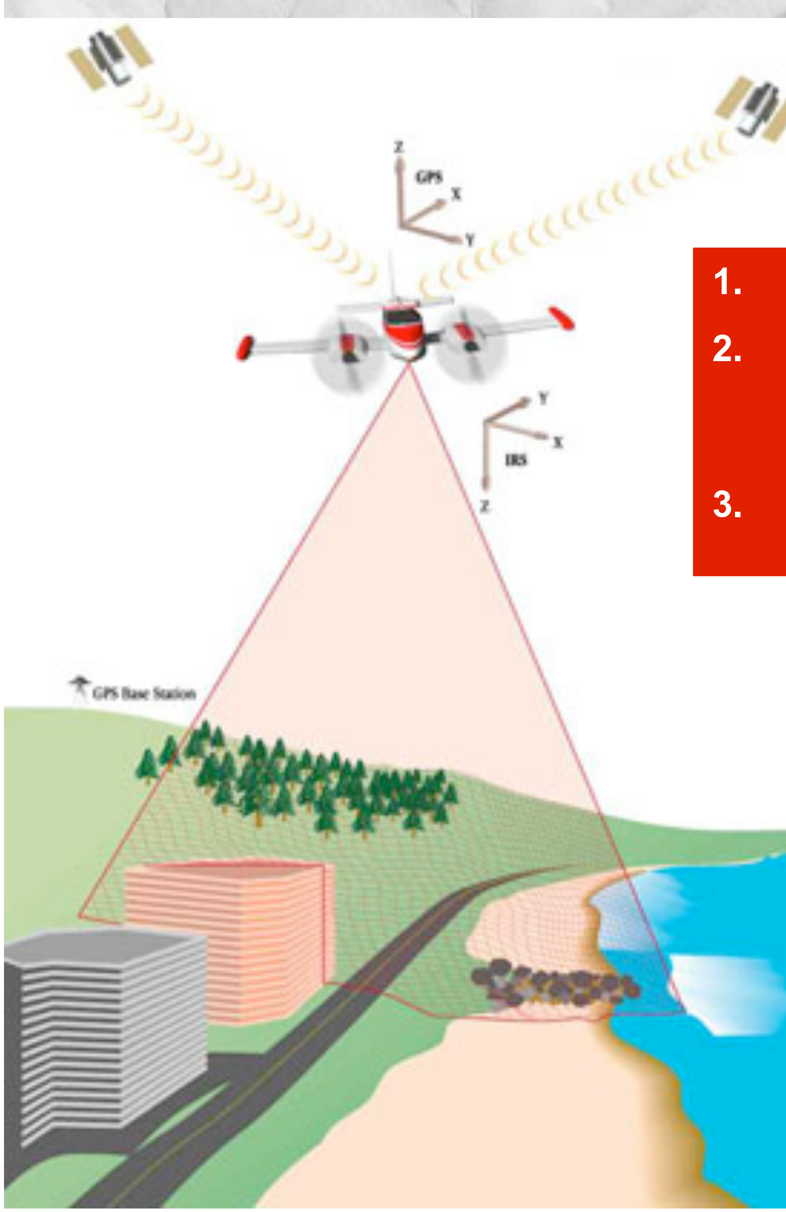
• Distance is calculated by measuring the time that a laser pulse takes to travel to and from an object.

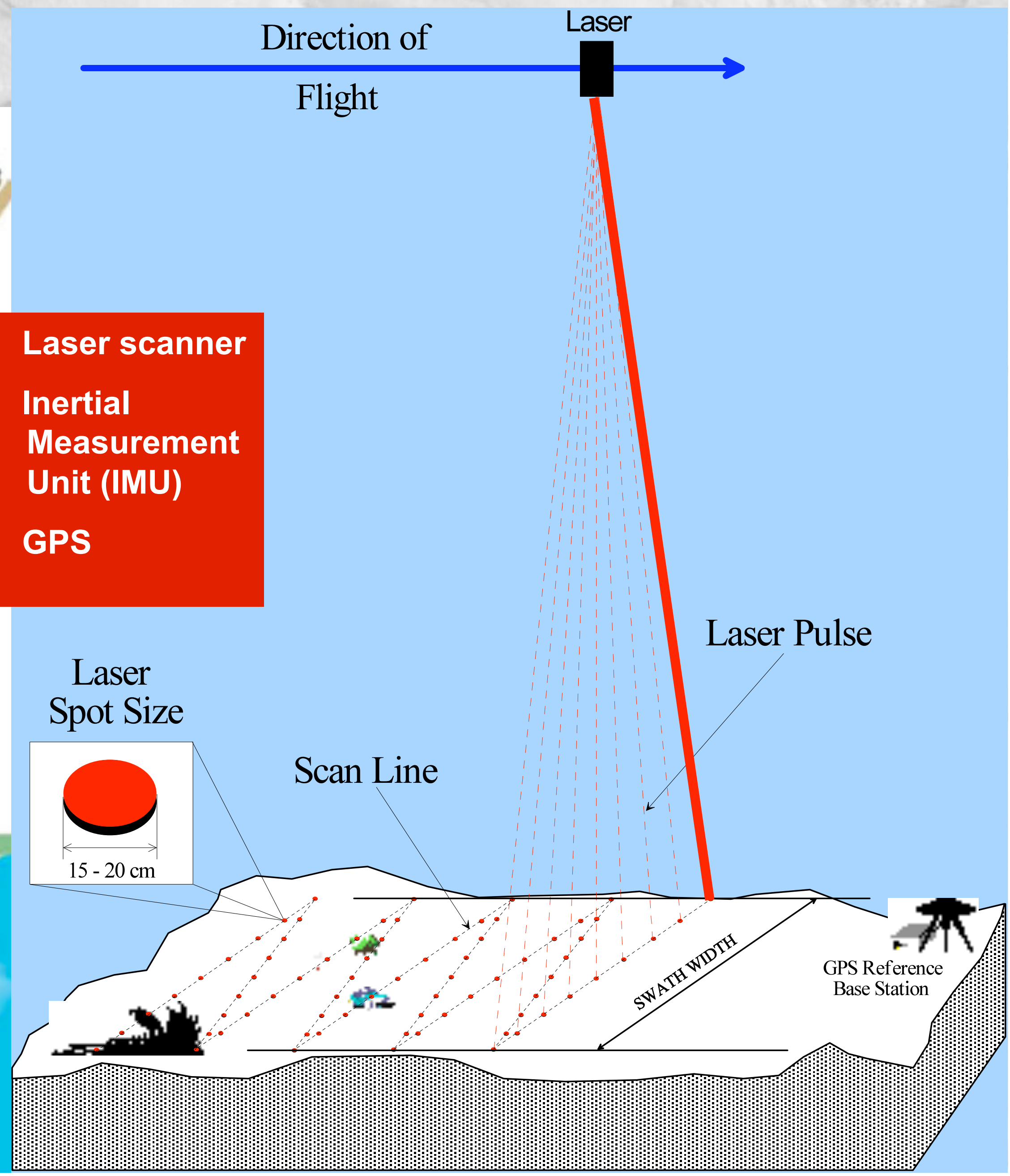
Ian Madin, DOGAMI

NCALM The National Center for Airborne Laser Mapping

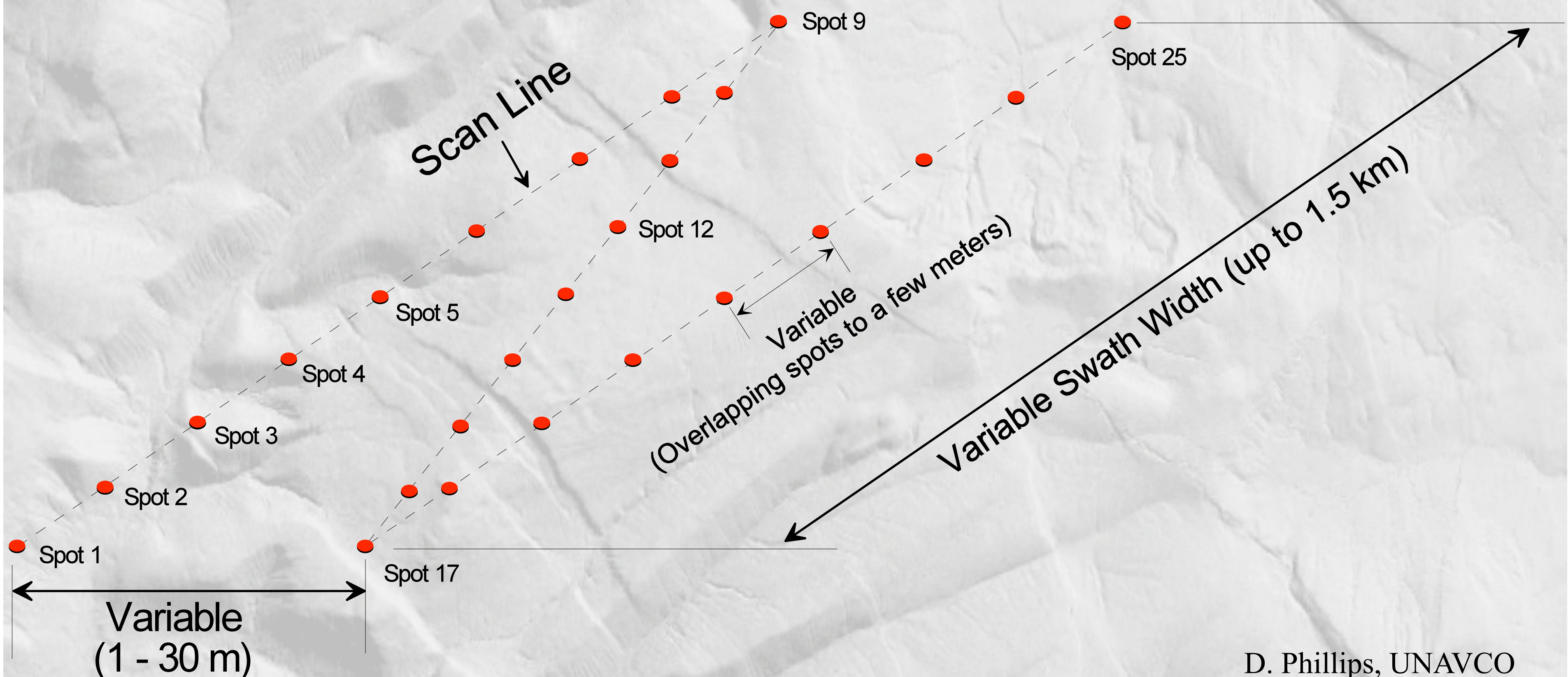
Lidar data collection

1. Laser scanner
2. Inertial Measurement Unit (IMU)
3. GPS





Surface Point Spacing

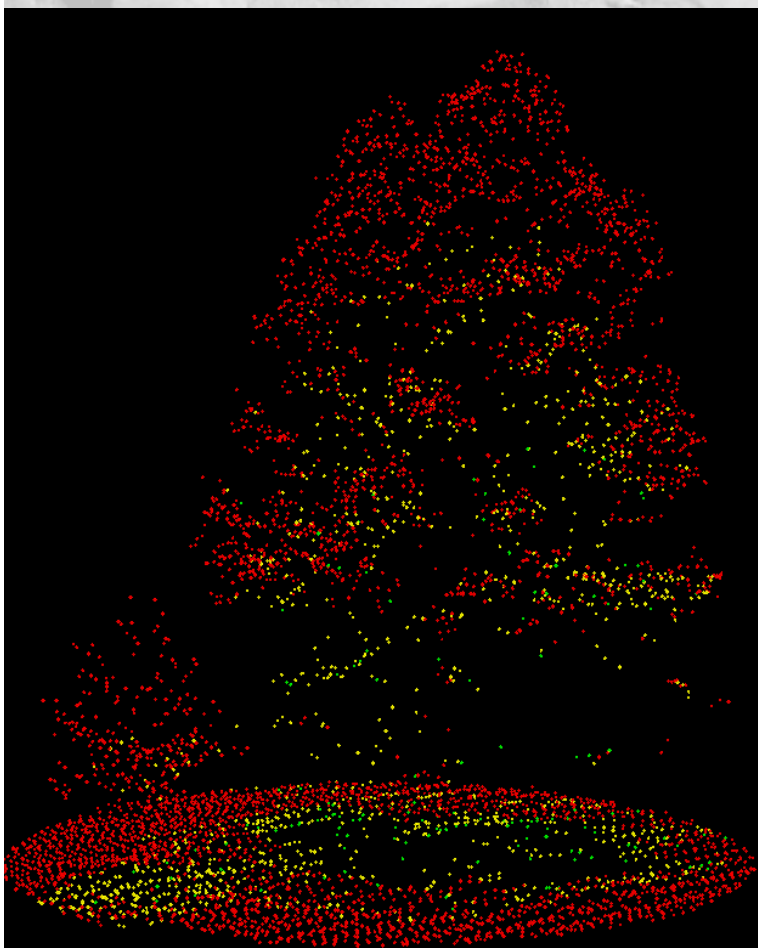


D. Phillips, UNAVCO

Scan line spacing, swath width, spot size and overlap can all be defined as necessary to achieve target data to specification

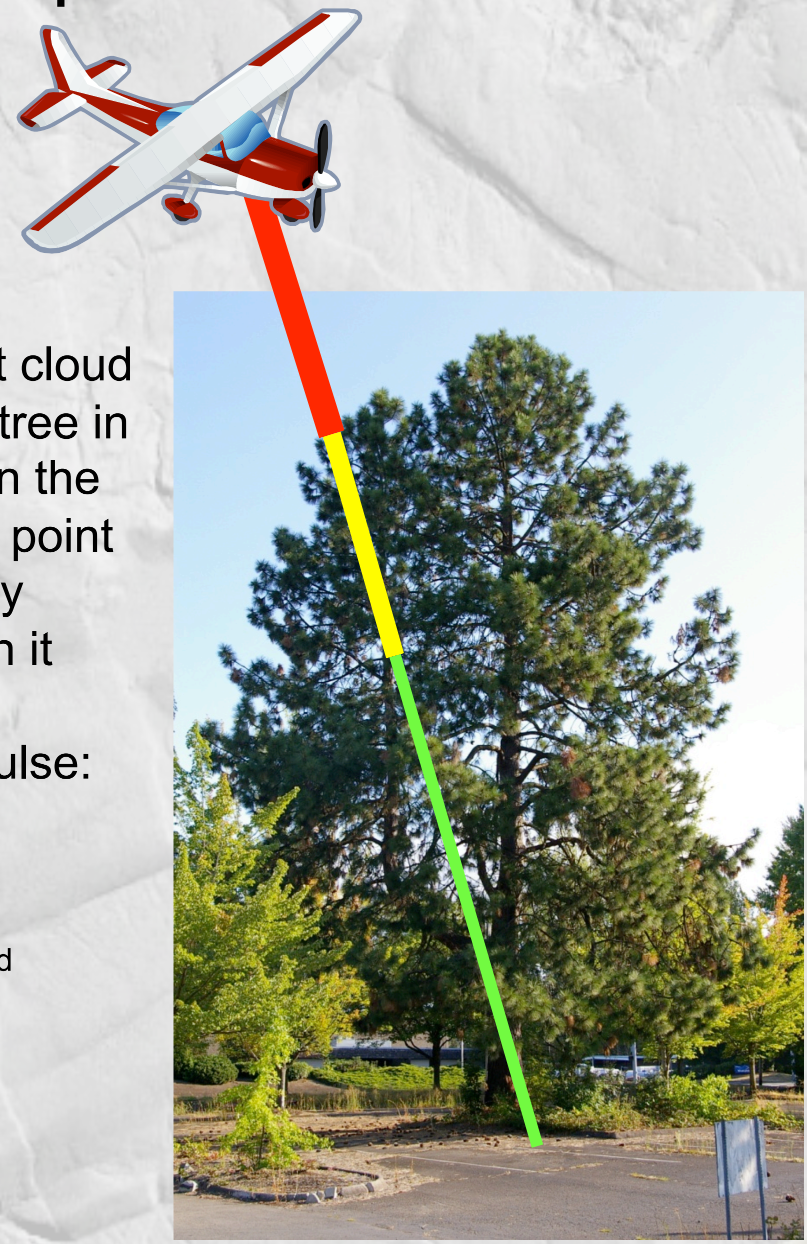
Each laser pulse can produce multiple consecutive measurements from reflections off several surfaces in its path

Ian Madin, DOGAMI

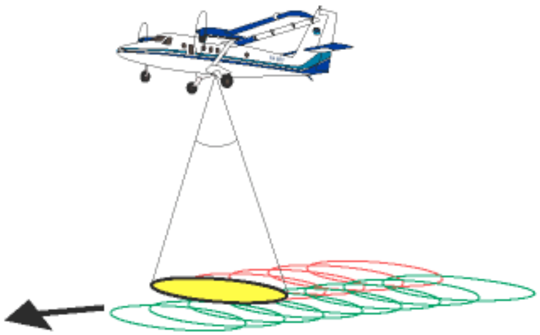


- Left = point cloud view of the tree in the photo on the right. Each point is colored by which return it was from a particular pulse:

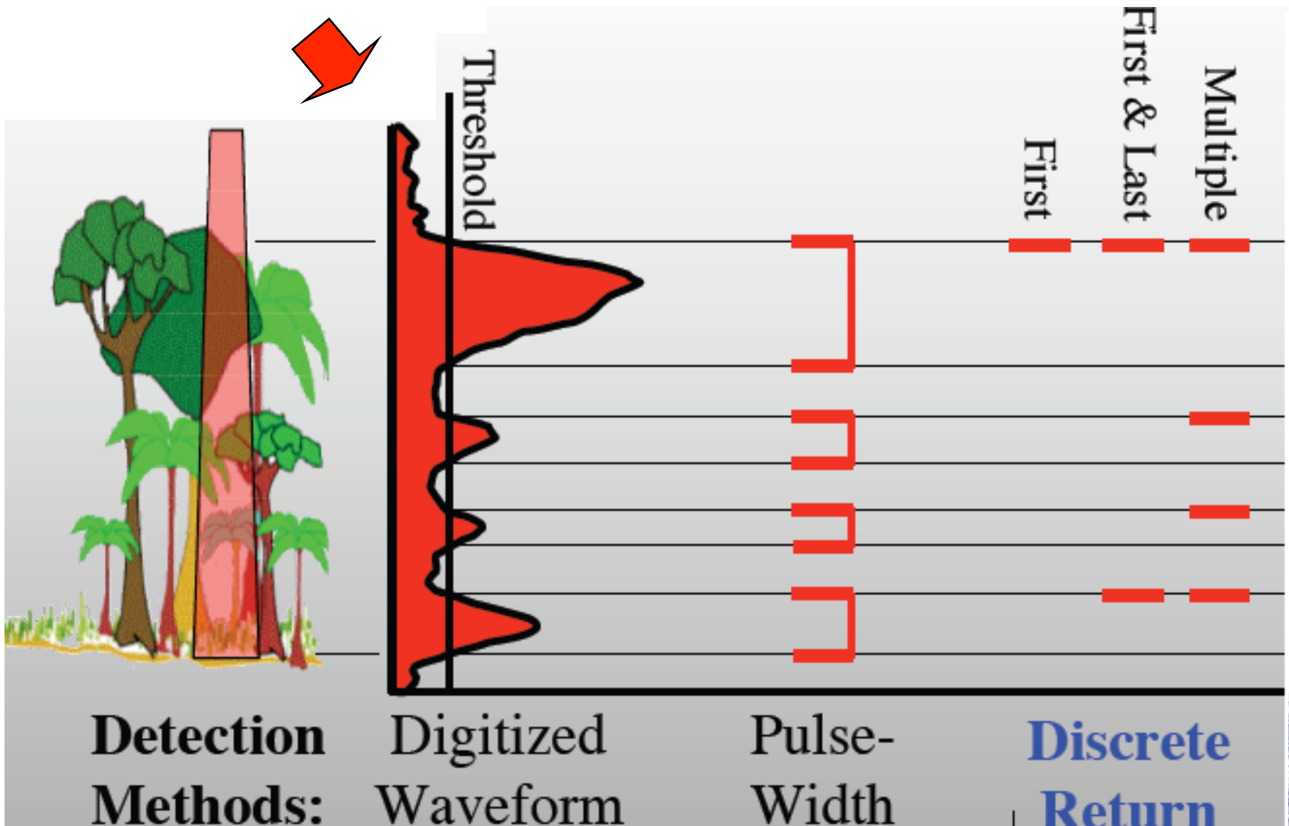
- Red= 1st
- Yellow = 2nd
- Green = 3rd

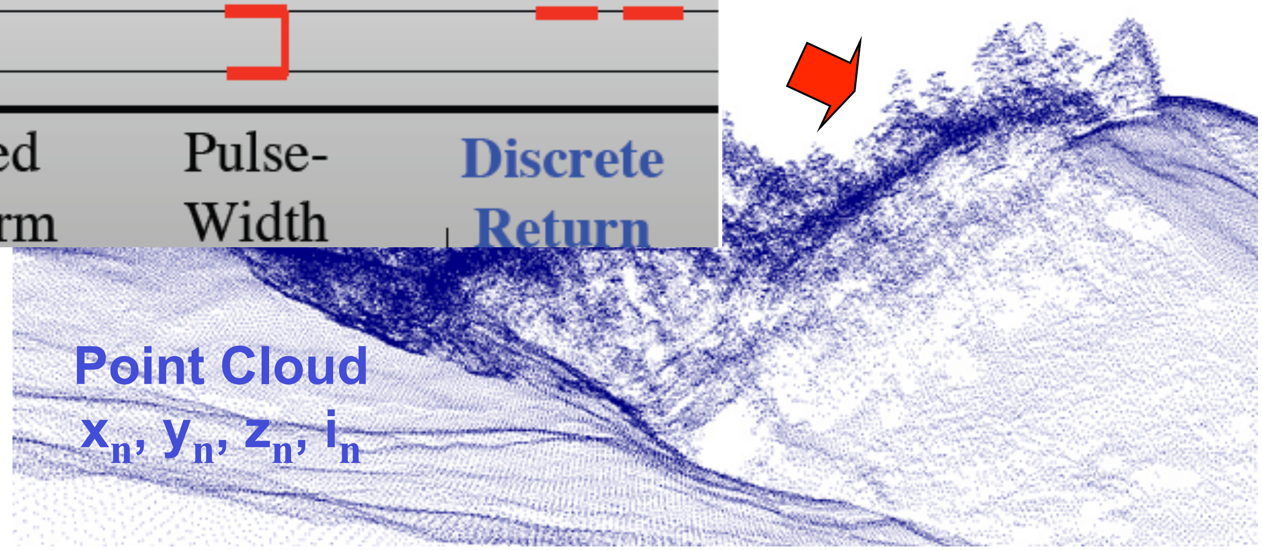


Lidar = Geodesy and signal processing



D. Harding, NASA





Typical Lidar Data Collection Parameters

- Aircraft: Cessna 337 Skymaster
- Personnel
  - One pilot, one operator in plane
  - GPS ground crew (2 to 10+ people)
- Scanner: Optech near-IR (Gemini)
- PRF: 33-125 KHz
- Flying height: 600 – 1,000m AGL
- Flying speed: 120 mph
- Swath overlap: 50% nominal
- Ground truthing: GPS (campaign & CORS)
- Navigation solution: KARS
- Point spacing: sub-meter
- Nominal Accuracy (on open hard and flat surface)
  - Vertical: 3 – 6 cm.
  - Horizontal: 20 – 30 cm.

NCALM
The National Center for Airborne Laser Mapping

Overall collection and deliverables workflow – example from EarthScope Lidar

1. **Planning** (client / community)
   – Target identification & prioritization
   – Definition of data specification & deliverables

2. **Collection** (vendor)
   – Additional GPS ground control?

3. **Processing** (vendor)
   – Scanning laser, IMU, GPS solutions
   – Point cloud generation
   – Data classification
   – Deliverable production

Overall collection and deliverables workflow II

3. **Qa / Qc** (Client)
   - Visual inspection of data
   - Ground control pts comparison
   - Swath to swath consistency checks

4. **Data delivery** (vendor)
   - Data arrives via hard drive or FTP
   - *This is the end of the line in many cases…*

5. **Data distribution** (client / 3rd party)
   - Data access levels vary dramatically
   - More on online data access to lidar tomorrow…

1. Lidar Acquisition Considerations

UNAVCO

- Target identification and prioritization
- Defining collection scheme and data product requirements
  - Tradeoffs concerning resolution vs. coverage
  - GPS ground control requirements
  - End use: geomorphology, geodesy, etc.
  - Cost (B4 ~$500/sq.km., NoCal ~$400/sq.km., DV ~$300/sq.km.)
- Seasonal constraints
  - “Leaf off”, snow, heat, etc.
- Data volume…lots of TB’s…yikes!
- Standard data products?
- Distribution scheme?

2. Collection

- Generally discussed in previous slides

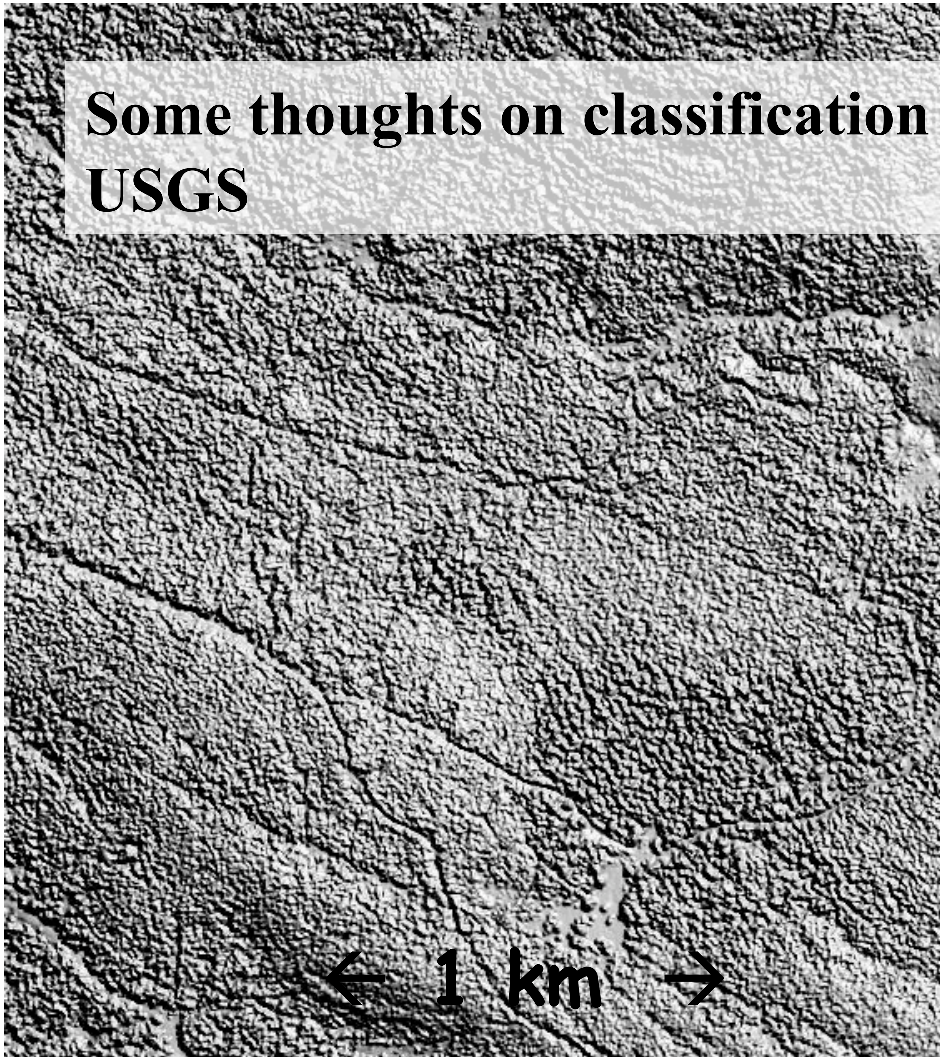
3. Processing

- GPS data processing and trajectory generation
  - Kinematic software (KARS, TRACK, etc.)

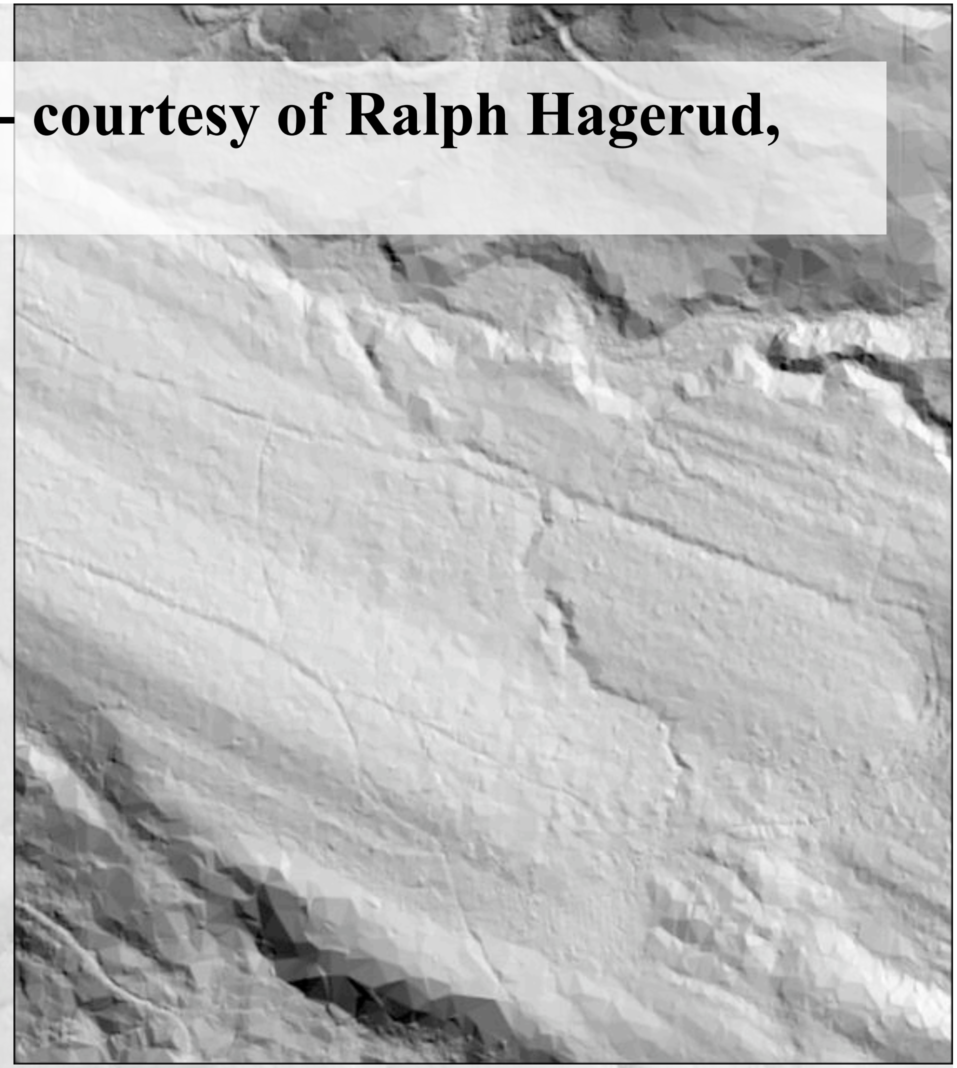
- LiDAR range processing and XYZ point cloud generation
  - Proprietary software (Terascan at present)

- Point cloud classification:
  - Typically completed with proprietary software.
  - Limited open source / free software available to “do it yourself”.
  - Not fully automated – significant manual intervention necessary.

Some thoughts on classification - courtesy of Ralph Hagerud, USGS



all surveyed points



ground points identified by semi-automatic processing

Nookachamps Creek, east of Mount Vernon, Washington

What is ground?

- Ground is smooth
  - despiking, iterative linear interpretation algorithms
- Ground is continuous (single-valued)
  - No-multiples algorithm
- Ground is lowest surface in vicinity
  - Block-minimum algorithms

R. Hagerud, USGS

Ground is smooth

⇒ despike algorithm

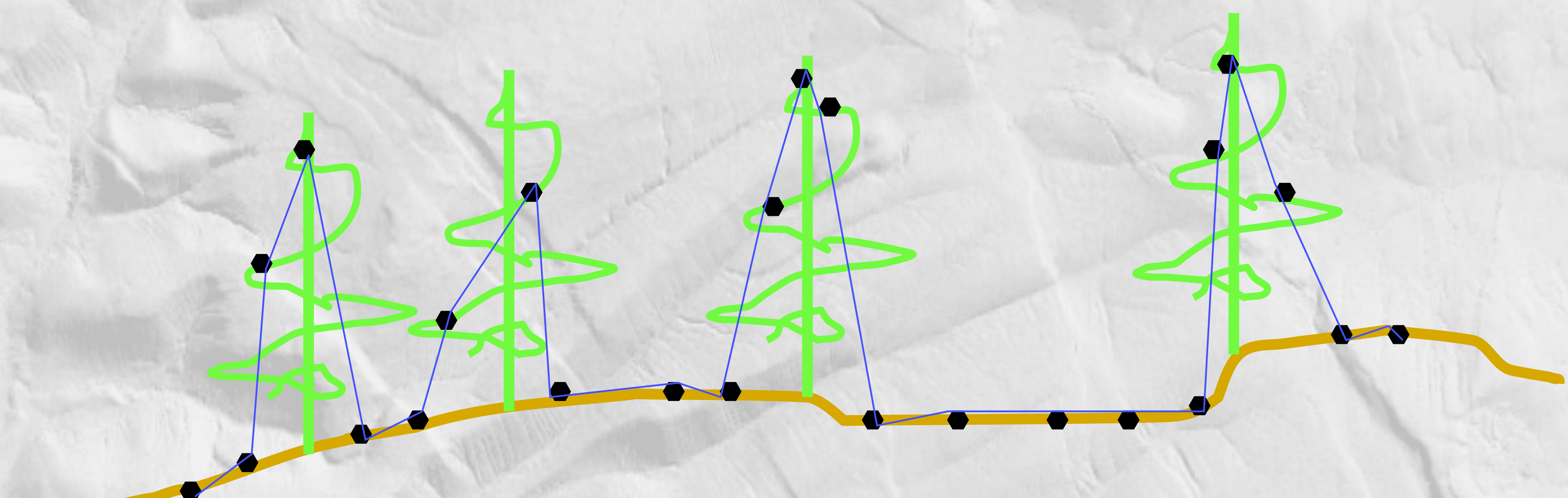
1. flag all points as ground
2. repeat:

    build TIN (triangulated irregular network) of ground points
    identify points that define strong positive curvatures
    flag identified points as not-ground

3. until no or few points are flagged

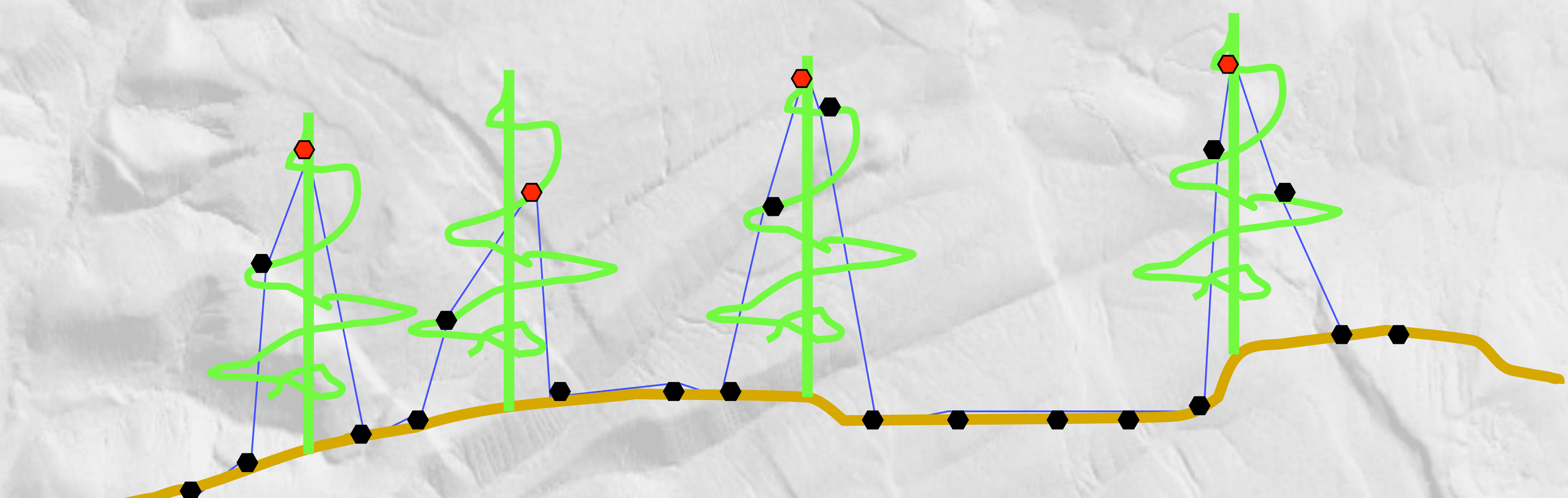
R. Hagerud, USGS

Start with mixed ground and canopy returns (e.g. last-return data), build TIN



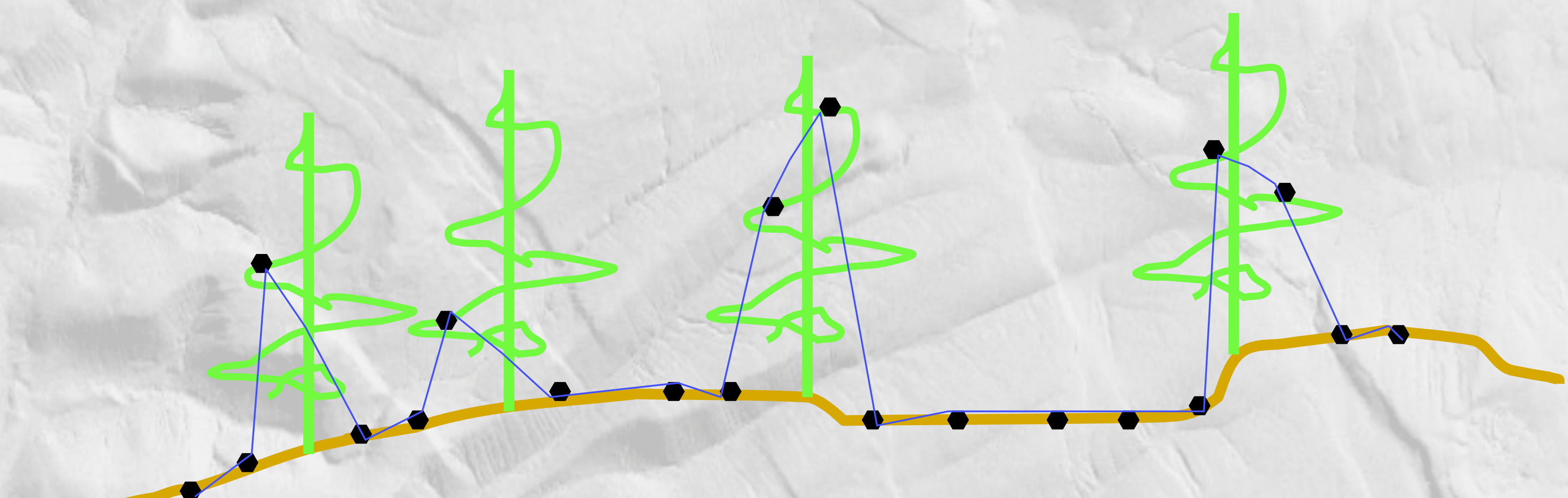
R. Hagerud, USGS

Flag points that define spikes
(strong convexities)



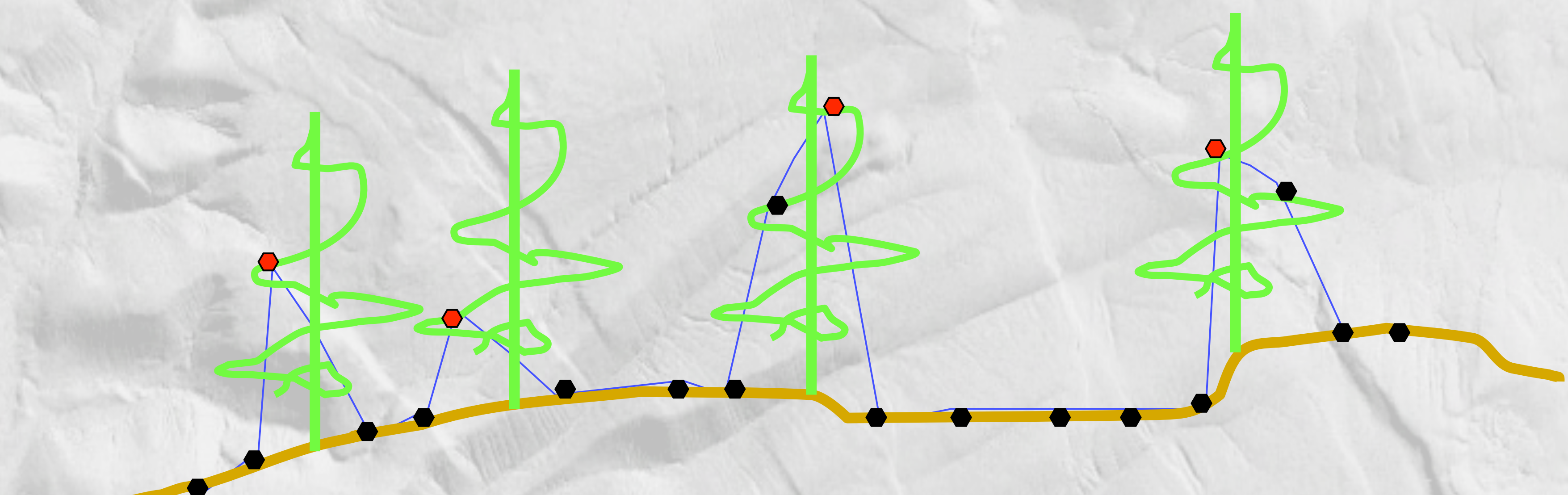
R. Hagerud, USGS

Rebuild TIN



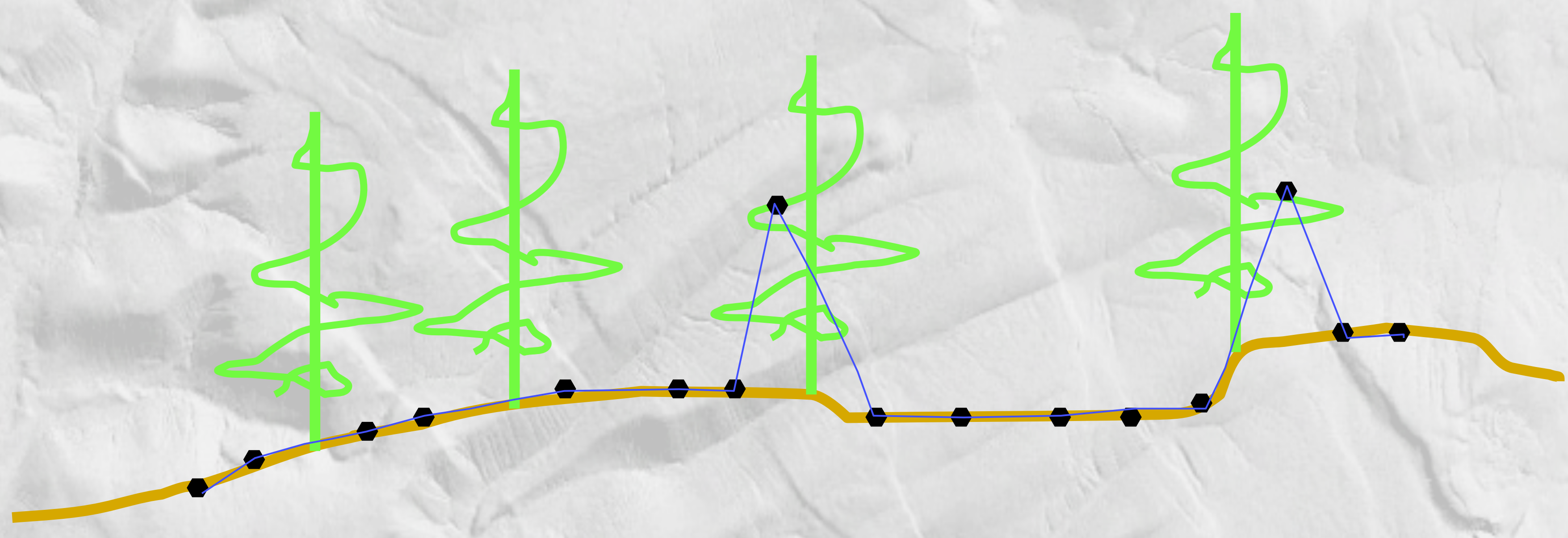
R. Hagerud, USGS

Flag points that define spikes
(strong convexities)



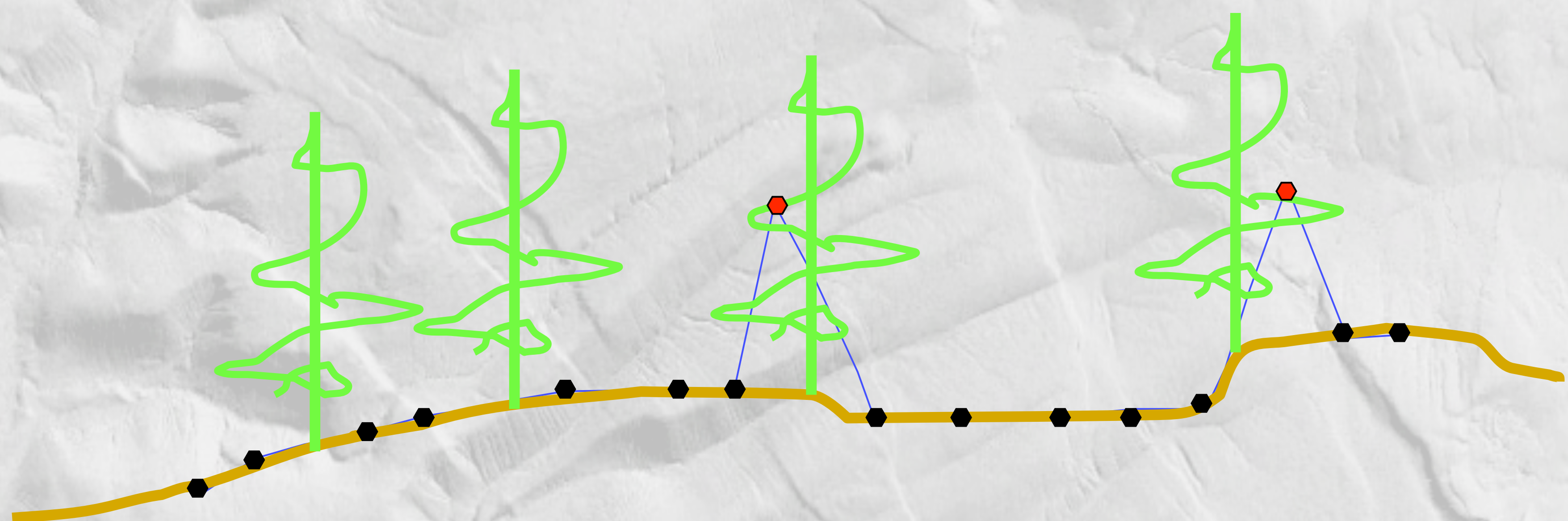
R. Hagerud, USGS

Rebuild TIN



R. Hagerud, USGS

Flag points that define spikes (strong convexities)



R. Hagerud, USGS

Rebuild TIN

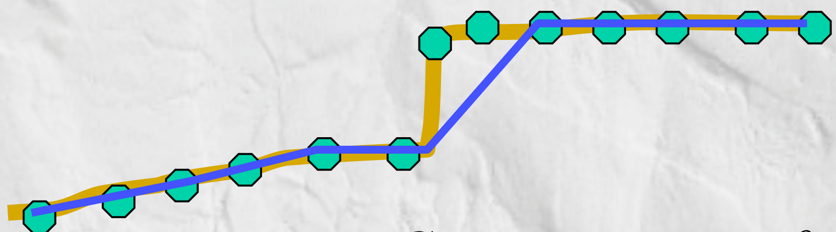
R. Hagerud, USGS

Despike algorithm

- It works
- It's automatic
  - Cheap(!)
  - All assumptions explicit
- It can preserve breaklines
- It appears to retain more ground points than other algorithms

R. Hagerud, USGS

Despike algorithm



Cross-section of highway cut

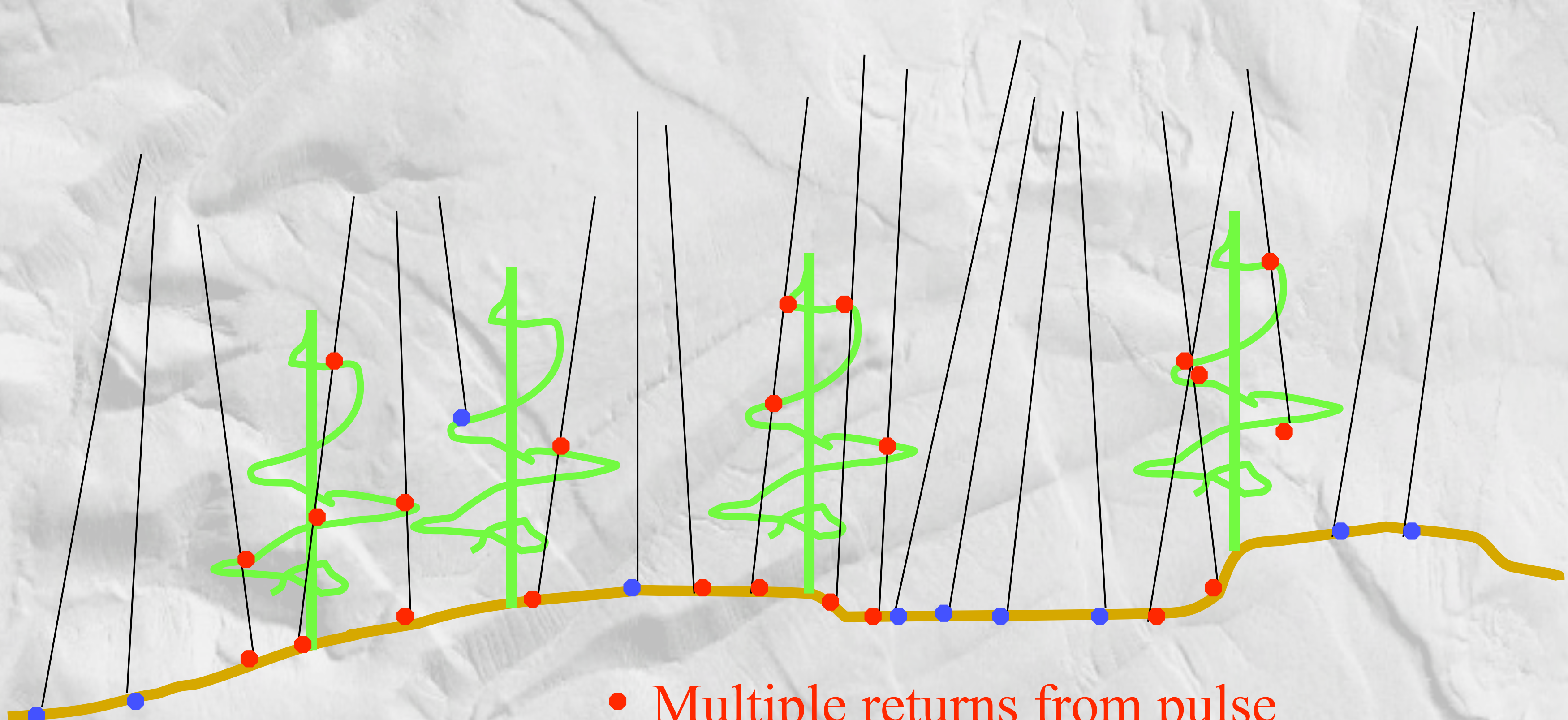
Problems:

- **Removes some corners**
- **Sensitive to negative blunders**
- **Computationally intensive**
- **Makes rough surfaces**
  - **Real? Measurement error? Misclassified vegetation?**

R. Hagerud, USGS

Ground is continuous (i.e., single-valued)

⇒ No-multiples algorithm



Multiple returns from pulse

Single return from pulse

R. Hagerud, USGS

No-multiples algorithms

- Fast
- Identify open areas
- Hopeless in woods

R. Hagerud, USGS

Ground is lowest surface in vicinity
⇒ block minimum algorithms

- **Computationally rapid with raster processing**
  - **Tweedy texture**
  - **Biased low on slopes**
- **Appropriate block size is inversely proportional to penetration rate**
  - **Requires human intervention to adjust block size**
- **Implicit assumption that ground is horizontal**

  (Successful users of block-minimum algorithms work in flat places)

R. Hagerud, USGS

In the real world…

- **Almost all return classification is done with proprietary codes**

- **Successful classification uses a mix of**
  - **Sophisticated code**
  - **Skilled human**
    - **To adjust code parameters**
    - **To identify and remedy problems**

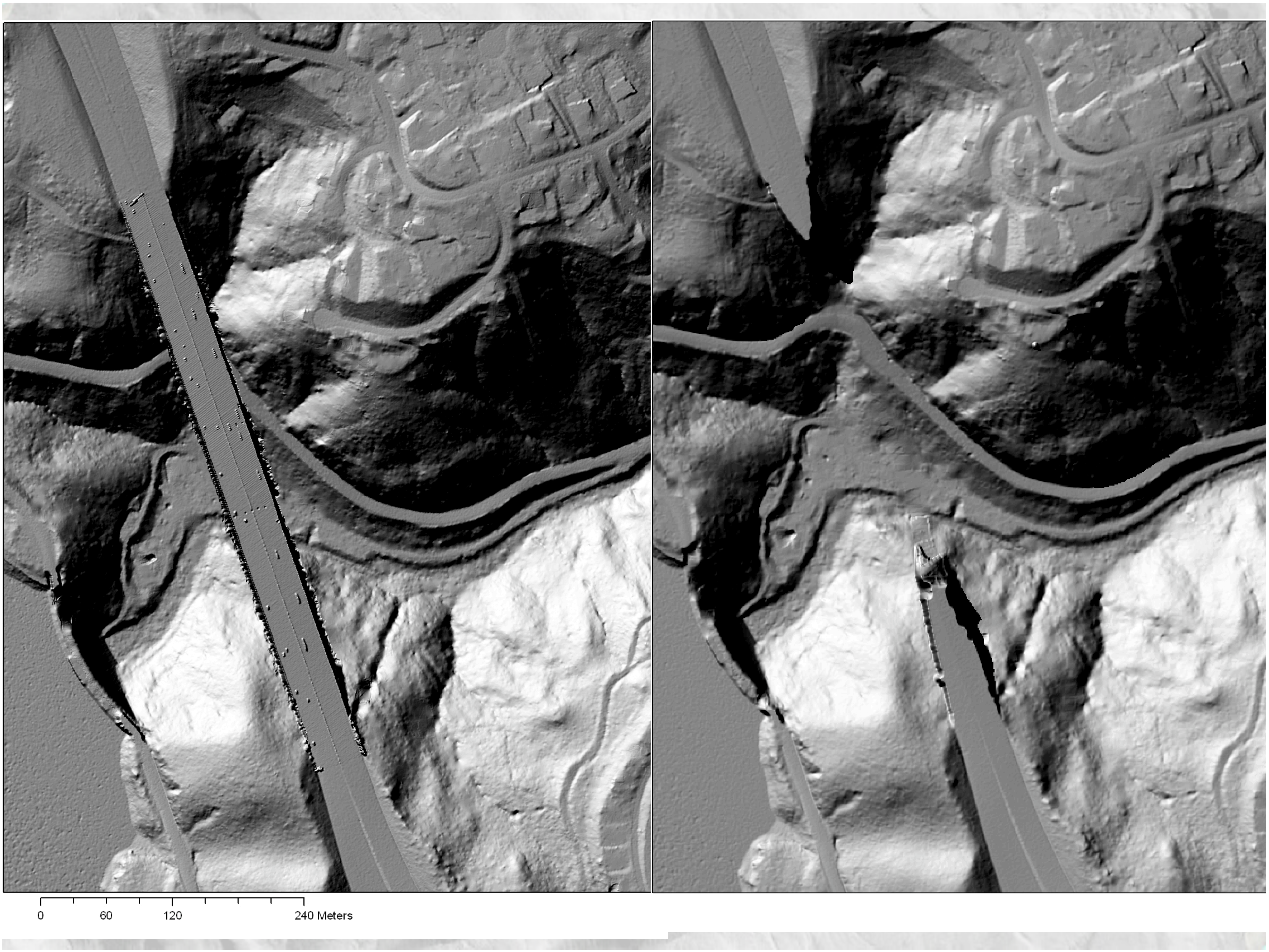
- **Let somebody else do it!** ***and then carefully check their work***

- **We have no useful metrics for accuracy of return classification**

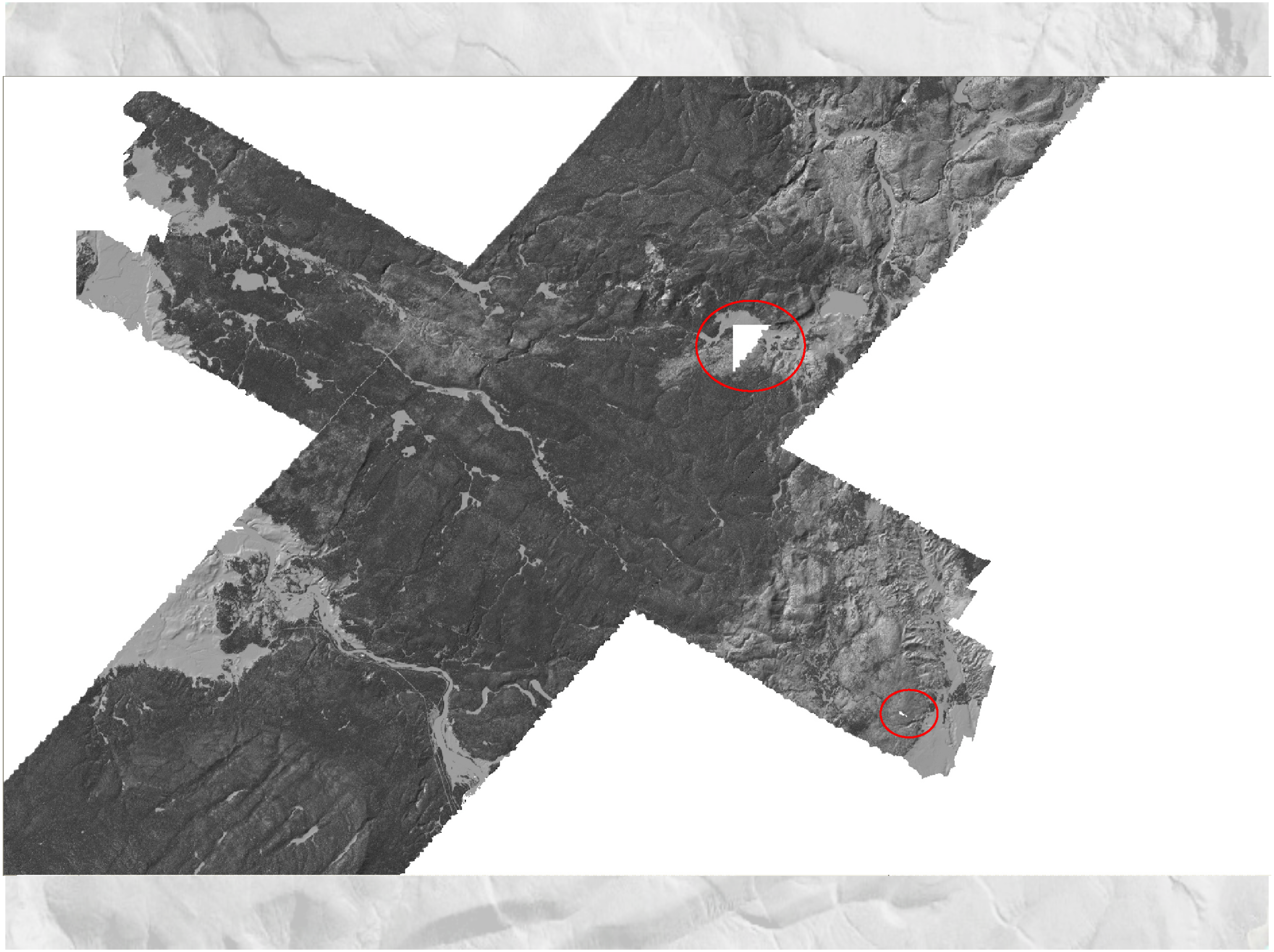
R. Hagerud, USGS

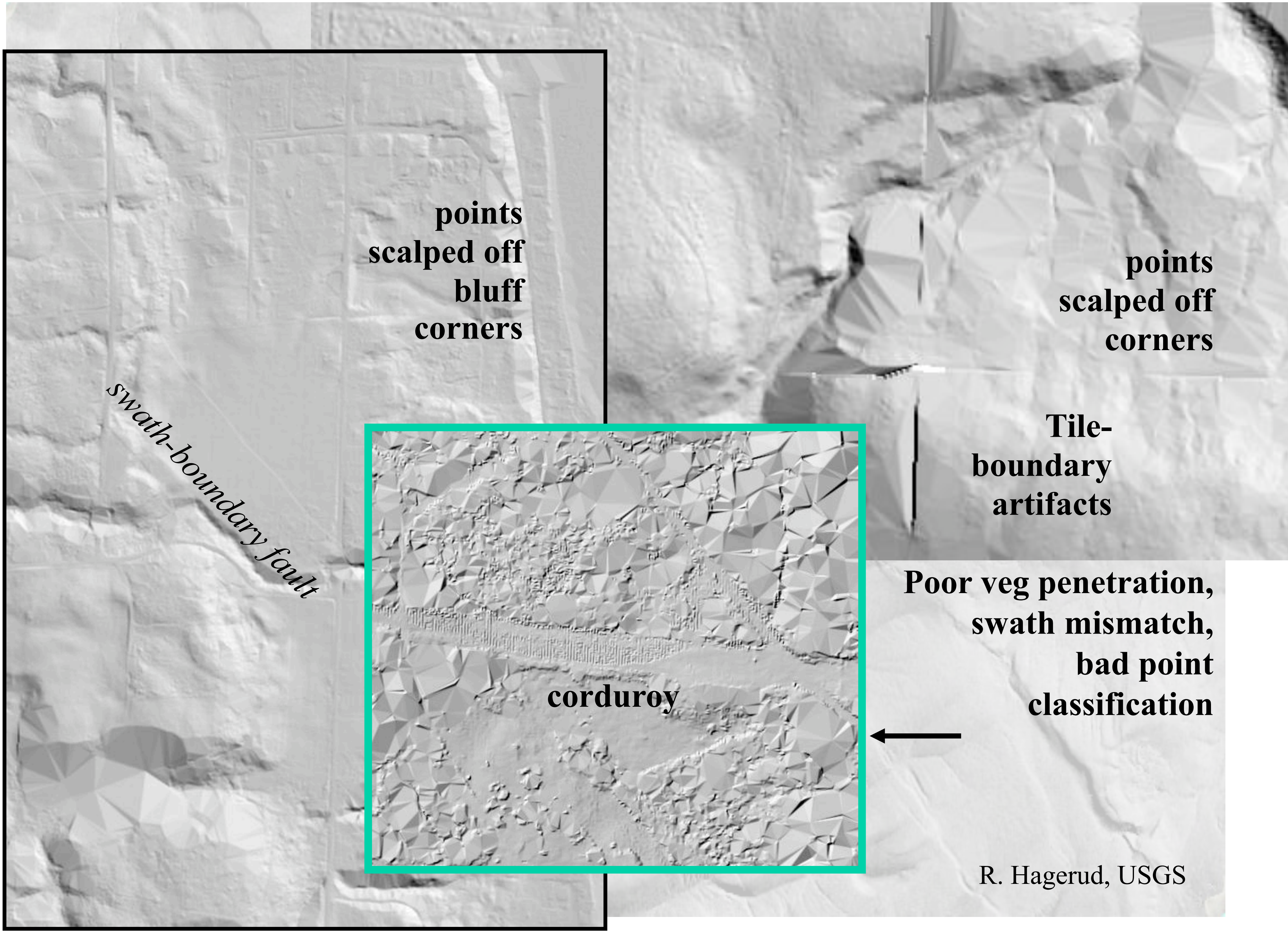
4. Qa / Qc

- **A QA protocol: 3 analyses**
  - **Test against ground control**
  - **Examine images of bare-earth surface model**
  - **Evaluate internal consistency**



0
60
120
240 Meters

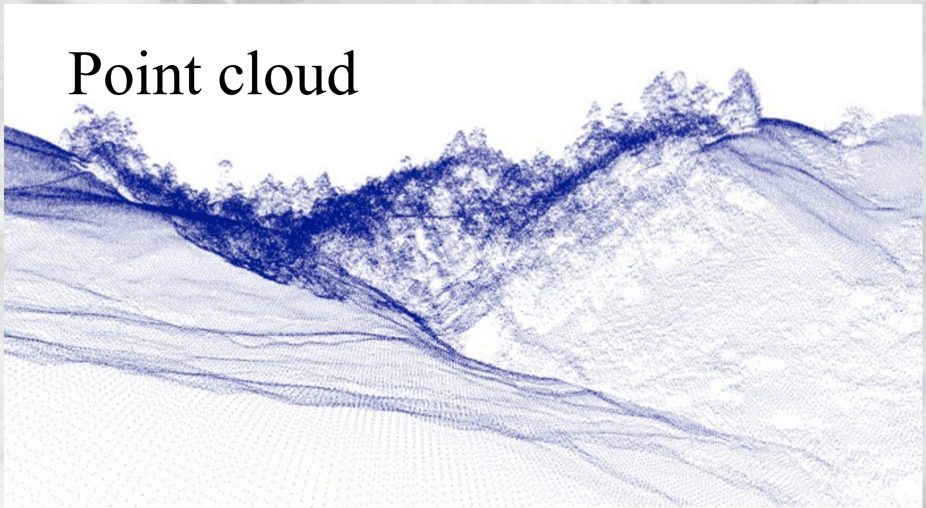



points
scalped off
bluff
corners
swath-boundary fault
points
scalped off
corners
Tile-
boundary
artifacts
Poor veg penetration,
swath mismatch,
bad point
classification
corduroy
R. Hagerud, USGS

5. Data Delivery

- Data typically arrives on HD from vendor.

- Deliverables:
  - Point cloud (ascii or LAS)
  - Bare earth and first return DEMs
  - Data mosaics at lower resolution (e.g. 1 m vs 0.5 m)
  - Metadata (FGDC if lucky) – XML, machine readable
  - Report of the survey – PDF, human readable



Deliverables - DEMs

- **DEM Data**:
  - Bare earth and first return DEMs in tiles (1 km x 1 km, USGS ¼ quad)
  - Hillshades of above DEMs (?)
  - Mosaics at lower resolution (?)
  - Intensity images (?)

- **File Formats**:
  - No standards
  - Common: Arc ESRI binary grid, ERDAS .IMG, GeoTiff, ascii grids, Surfer .grd, etc.

Deliverables – Point Cloud

- **X,Y,Z + attributes**:
  - Attributes: GPS time, Intensity / RGB, return #, classification (ground, vegetation, other), swath ID
  - All return files:
    - Organized into tiles (1 km x 1 km, subset of USGS ¼ quad) or by swath (USGS advocating this)

- **File Formats**:
  - ASCII (.txt, .xyz)
    - Easily parsed (linux – painful on Windows), portable, HUGE, need to move to another format for on-the-fly analysis.

```
x,y,z,gpstime,intensity,classification,flight_line
560149.82,4108410.91,-14.54,331709.549800,5,2,9
560149.54,4108410.78,-14.04,331709.549800,5,1,9
```

Deliverables – Point Cloud II

File Formats:

- LAS (.las)
  - Standard format (at v. 1.3) defined by ASPRS (American Society for Photography and Remote Sensing).
  - Binary – smaller, easily parsed and indexed with correct libraries (libLAS)
  - Standard…
  - Robust header
    - Scanner info, processing software, spatial coordinates, bounding box, # of points in file
  - Requires software that can read and write LAS
  - More restrictive in terms of what attributes you can add
  - LAS vs. fully populated LAS – still need to output all the attribution
  - Version 1.3 supports waveforms!

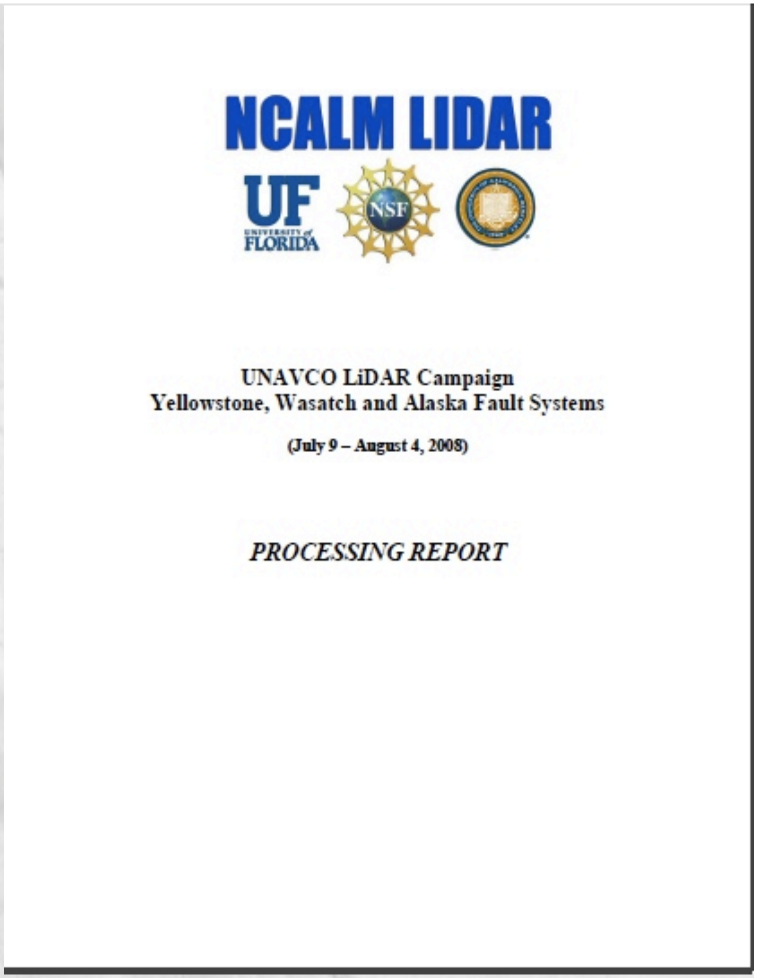
Item	**Format**	**Size**	**Required**
X	long	4 bytes	*
Y	long	4 bytes	*
Z	long	4 bytes	*
Intensity	unsigned short	2 bytes	
Return Number	3 bits	3 bits	*
Number of Returns (given pulse)	3 bits	3 bits	*
Scan Direction Flag	1 bit	1 bit	*
Edge of Flight Line	1 bit	1 bit	*
Classification	unsigned char	1 byte	
Scan Angle Rank (-90 to +90) – Left side	char	1 byte	*
File Marker	unsigned char	1 byte	
User Bit Field	unsigned short	2 bytes	

ASPRS Standard LIDAR Point Classes

Classification Value (bits 0:4)	***Meaning***
0	Created, never classified
1	Unclassified[1]
2	Ground
3	Low Vegetation
4	Medium Vegetation
5	High Vegetation
6	Building
7	Low Point (noise)
8	Model Key-point (mass point)
9	Water
10	*Reserved for ASPRS Definition*
11	*Reserved for ASPRS Definition*
12	Overlap Points[2]
13-31	*Reserved for ASPRS Definition*

Deliverables – Metadata

- **Report of the Survey**:
  - PDF format (human readable)
    - Data provider, area surveyed, when surveyed, instrument used, processing software and methods, spatial coordinates and datums, know issues, etc.
    - Spatial reference framework
    - Data provider's report on data quality
    - Naming, formats, spatial organization of data files



NCALM LIDAR

UF
UNIVERSITY of FLORIDA

NSF

UNAVCO LiDAR Campaign
Yellowstone, Wasatch and Alaska Fault Systems

(July 9 – August 4, 2008)

PROCESSING REPORT

- **FGDC (or similar) metadata**:
  - XML (machine readable)
  - Ideally populated by vendor and client
  - Not delivered by NCALM…

5. Data distribution

- Very little data makes it online
- Access mechanisms vary
- Who funds hosting of multi-TB datasets?

More tomorrow AM…