A Geoinformatics Approach to LiDAR / ALSM Data Distribution, Interpolation, and Analysis

Christopher J. Crosby[1], Jeffrey Conner[1], Efrat Frank[2], J Ramón Arrowsmith[1], Ashraf Memon[2], Viswanath Nandigam[2], Gilead Wurman[1] and Chaitan Baru[2]
[1]Department of Geological Sciences, Arizona State University, Tempe, AZ 85287
[2]San Diego Supercomputer Center, University of California, San Diego, La Jolla, CA 92093
chris.crosby@asu.edu

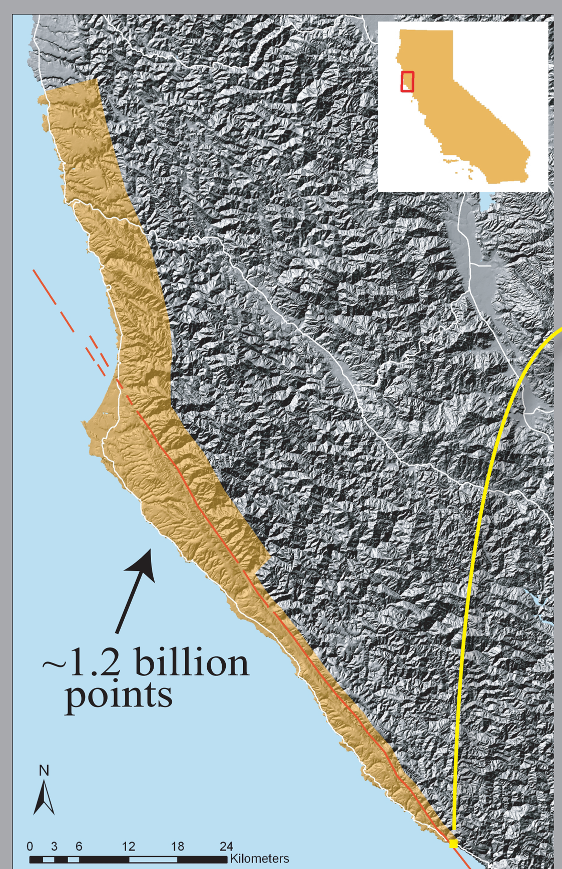
Arizona State University
http://activetectonics.asu.edu

GEON CYBERINFRASTRUCTURE FOR THE GEOSCIENCES

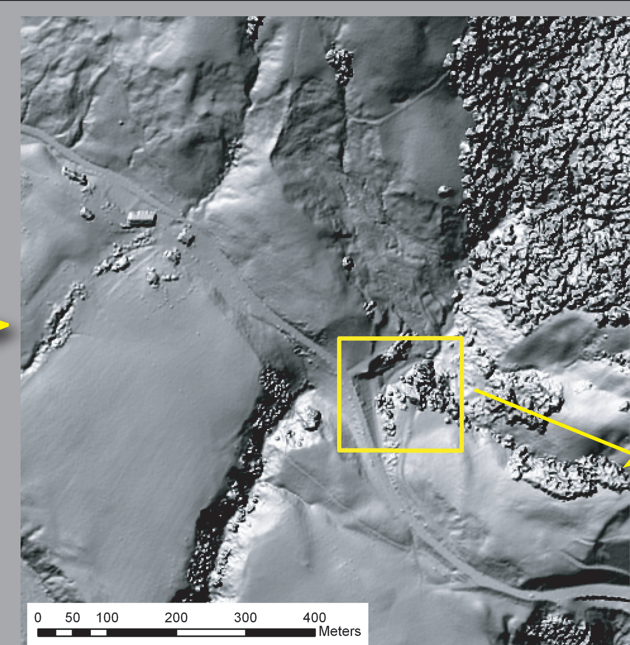
I. THE COMPUTATIONAL CHALLENGE

EXAMPLE DATASET: NASA / USGS Northern San Andreas Fault and associated marine terraces (flown February, 2003)

- LiDAR/ALSM generates massive data volumes - billions of returns are not uncommon in these data sets.
- Processing and analysis of these data requires significant computing resources not available to most geoscientists.
- Interpolation of these data challenges typical GIS / interpolation software.
  - our tests indicate that ArcGIS, Matlab and similar software packages struggle to interpolate even a small portion of these data.
- Ongoing advances in LiDAR technology, namely higher pulse rate instruments, means forthcoming data sets will be even larger due to higher point density.

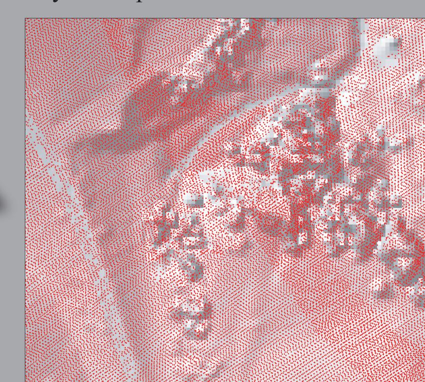


Hillshade of 30m USGS digital elevation model (DEM) with the extent of the Northern San Andreas and associated marine terraces LiDAR data set shown in orange. This data set contains ~1.2 billion data points. The trace of the 1906 earthquake is shown in red (Jennings, 1994). Yellow box shows location of detail image.

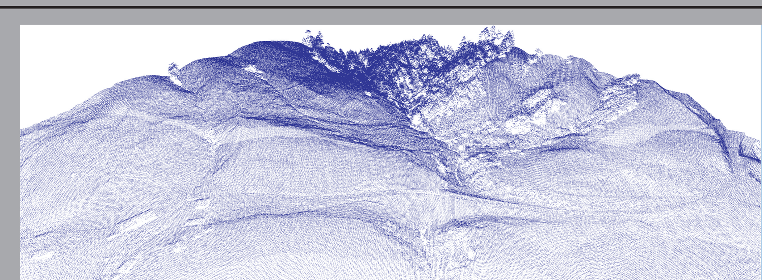


(left) Hillshade of 1.8 m DEM created from the NSAF LiDAR data. This DEM was created by interpolating 1.1 million data points.

(below) Detail view of a portion of the LiDAR point cloud that was interpolated to produce DEM shown at left. Note the density of the point cloud and swath overlap.

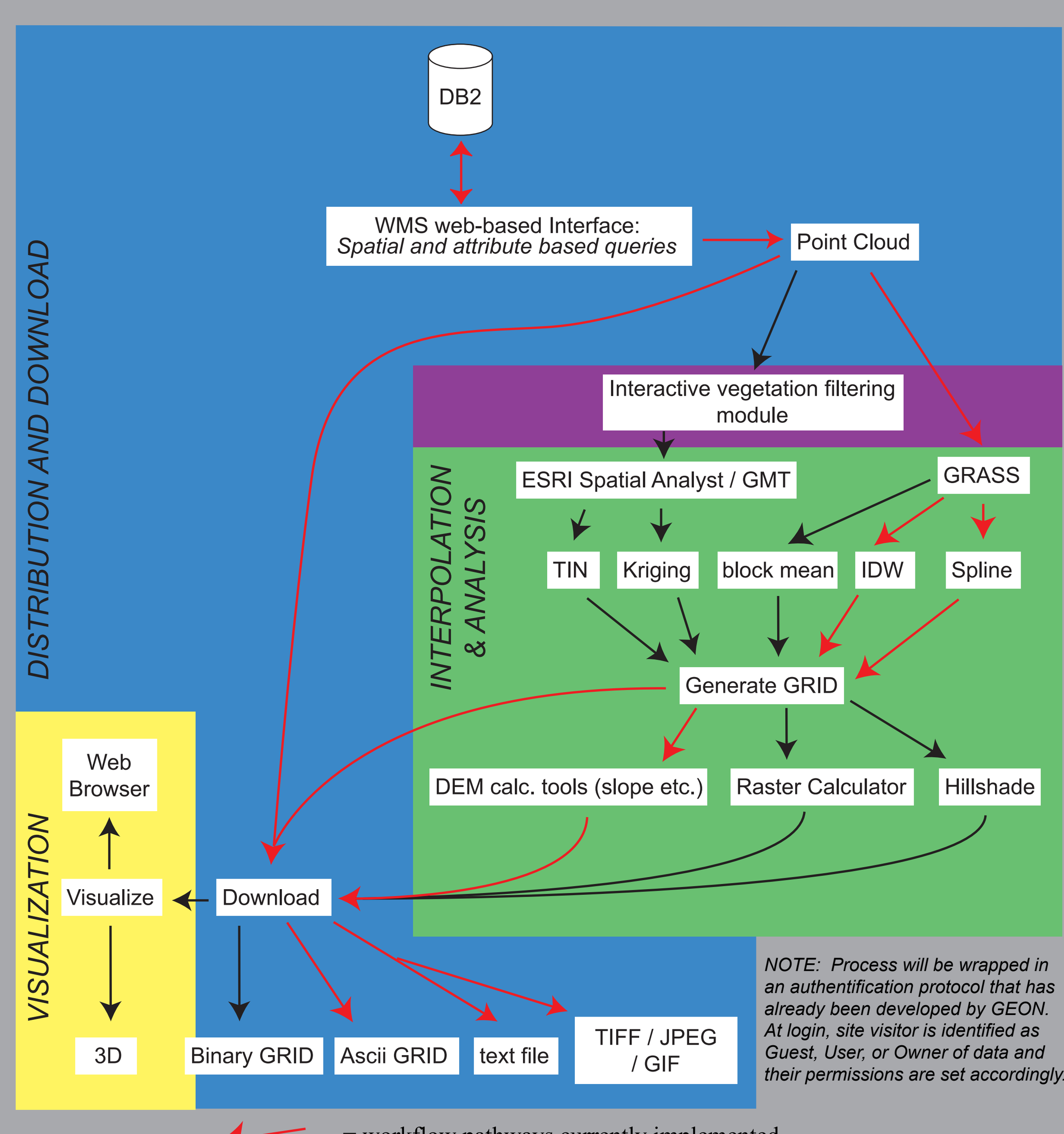


3D rendering of the ~1.1 million LiDAR point returns used to produce the DEM shown above. Note the detail with which trees, landforms and buildings are represented.



3D visualization produced using LViz, a free LiDAR point cloud and interpolated surface visualization tool developed in the Active Tectonics Research Group at Arizona State University. LViz can be downloaded at: http://activetectonics.la.asu.edu/GEONatASU/LViz.html

II. CONCEPTUAL GEON LiDAR WORKFLOW



- We propose utilizing the cyberinfrastructure developed by GEON to offer online data distribution, interpolation to grid, and analysis of large LiDAR datasets.
- GEON's grid computing infrastructure, and partnership with the TeraGRID (www.teragrid.org), offer users the ability to run computationally intensive interpolations and analyses of large LiDAR/ALSM datasets without having the computing resources locally.
- We have developed a workflow that runs on the GEON grid and utilizes modular web services to complete a variety of processing and analysis tasks.
- Although conceived for interpolation and analysis of LiDAR data, our workflow could be applied to any data set where interpolation of point data to a grid is computationally intensive (i.e. gravity data).

III. GEON WORKFLOW IMPLEMENTATION

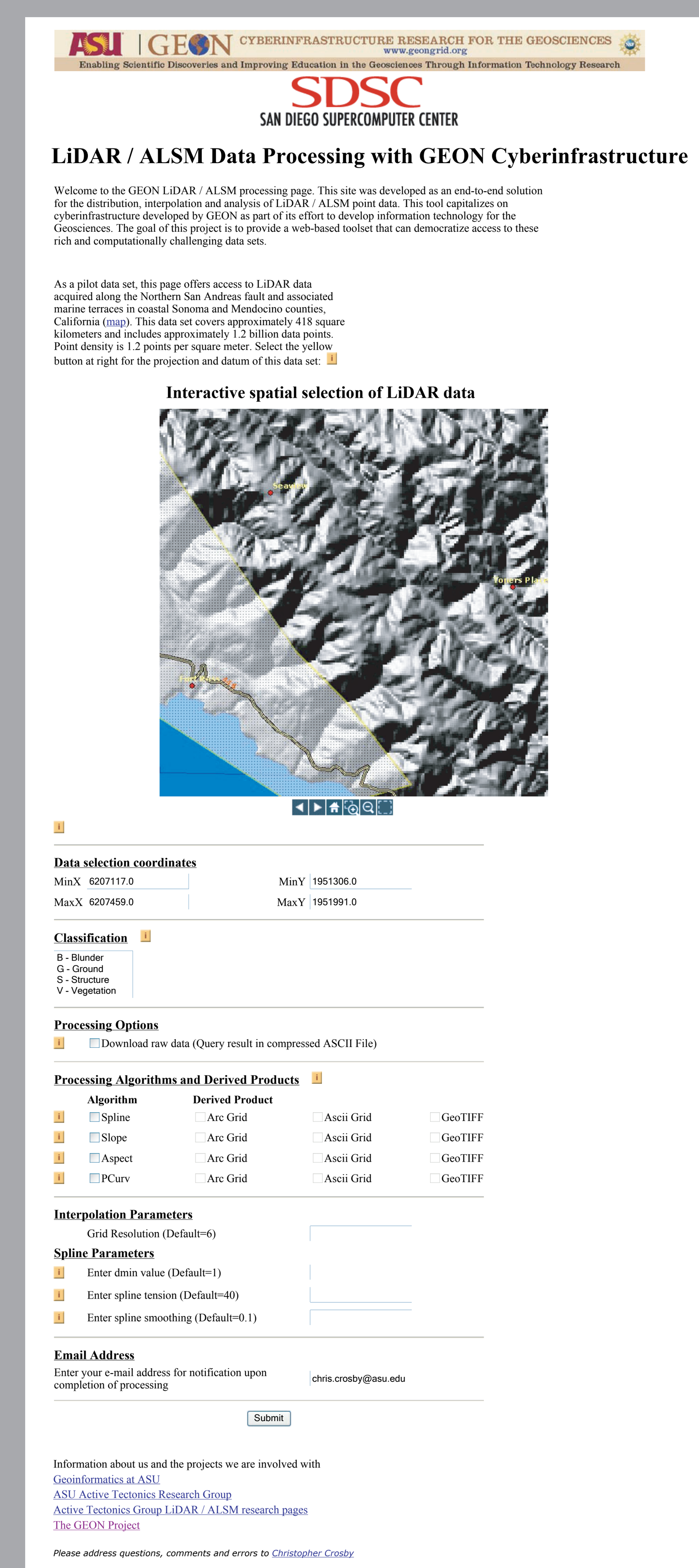
OVERVIEW:

- Using GEON cyberinfrastructure we have developed an internet-based LiDAR processing workflow that utilizes advanced spatial databases, GRASS Open Source GIS and web service technology to distribute, interpolate, analyze and visualize LiDAR data.

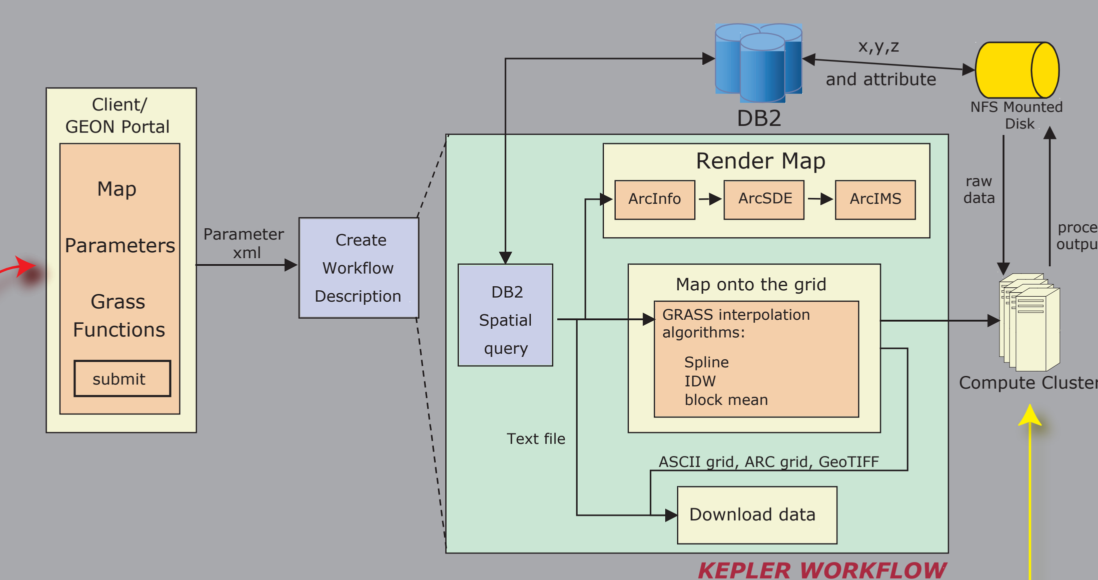
Features:
- Spatial and attribute based queries on raw LiDAR point cloud data.
- Spline interpolation to Digital Elevation Model (DEM).
- User control over interpolation parameters.
- Slope, aspect and profile curvature (pcurv) derived products.
- Download of of products in TIFF (with world file), ASCII and ESRI Arc ASCII (Arc GRID) formats.
- Visualization of data products via web browser window or in 3D via Fledermaus free viewer iView3D and our own OpenGL-based tool "LVIZ".

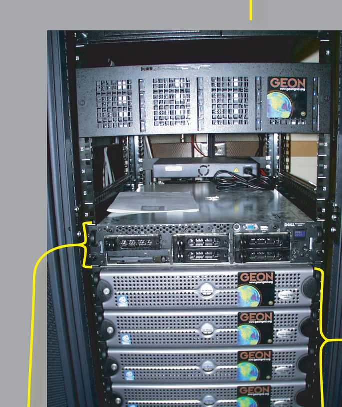
IMPLEMENTATION:

- The GEON LiDAR workflow utilizes the Kepler System (kepler-project.org) to integrate heterogeneous local and remote tools in a single interface:
  - Web and Grid services
  - GIS services (GRASS Open Source GIS)
  - Remote tools via SSH, SCP and GridFTP
  - Relational and spatial databases (Datastar Terascale Machine at San Diego Supercomputer Center)



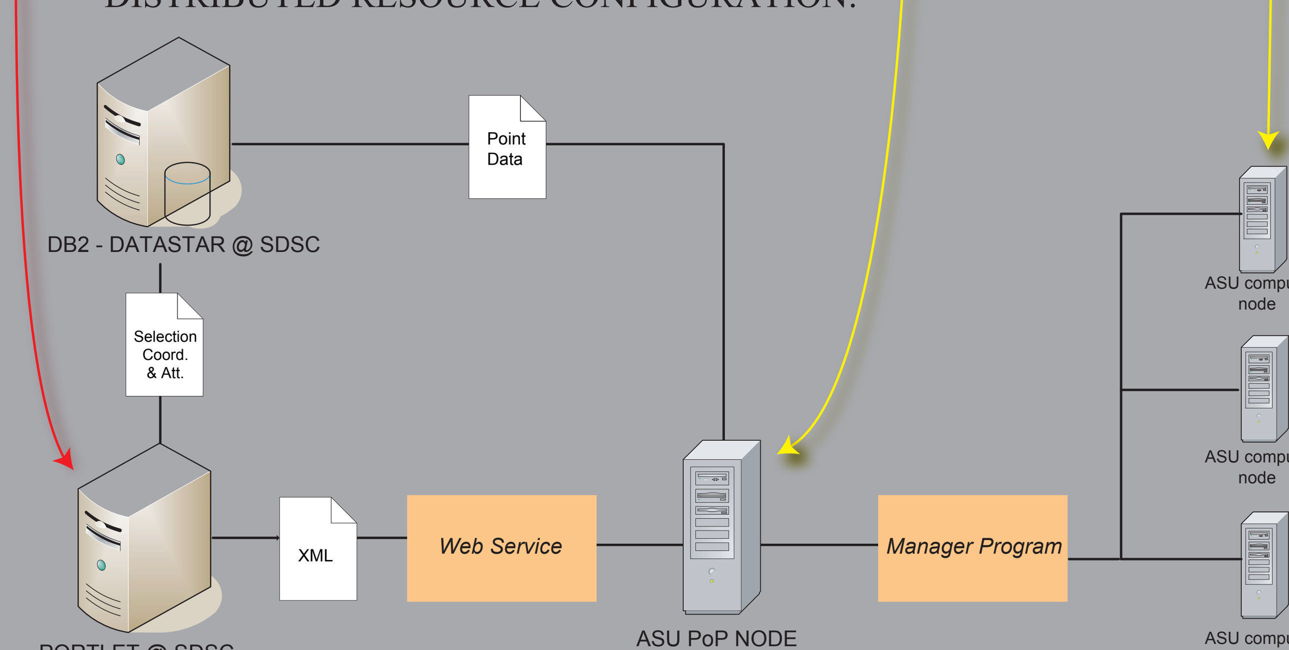
PORTLET ARCHITECTURE:





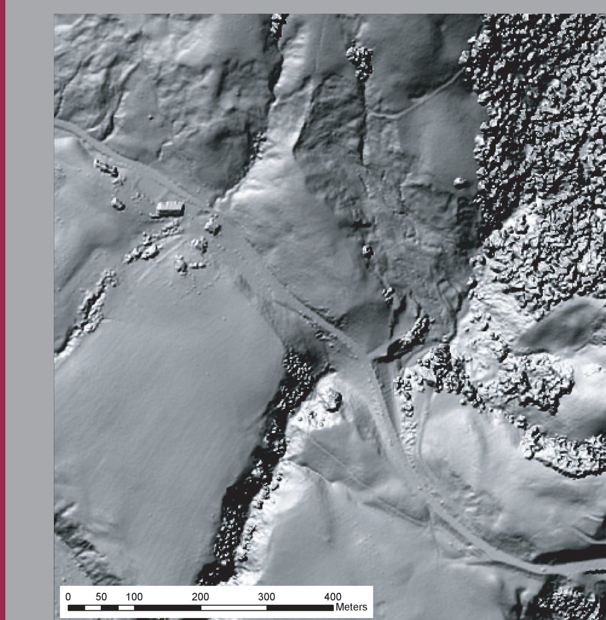
ASU GEON node "Agassiz" - six 2.8 Ghz Intel Xeon processors w/ 2 GB RAM per CPU. Running Linux ("Rocks" based on Red Hat 7.3). 2.5 TB of storage. All GEON participating institutions have similar resources that are made available via the GEON Grid.

DISTRIBUTED RESOURCE CONFIGURATION:



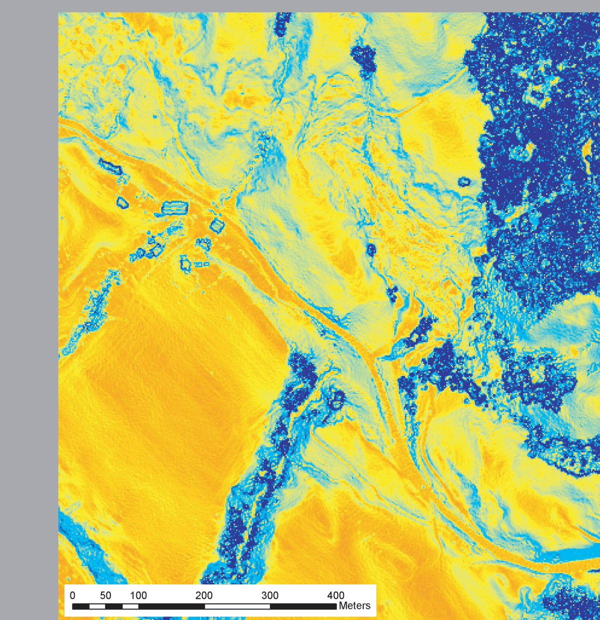
IV. PRODUCTS

Interpolation to DEM:

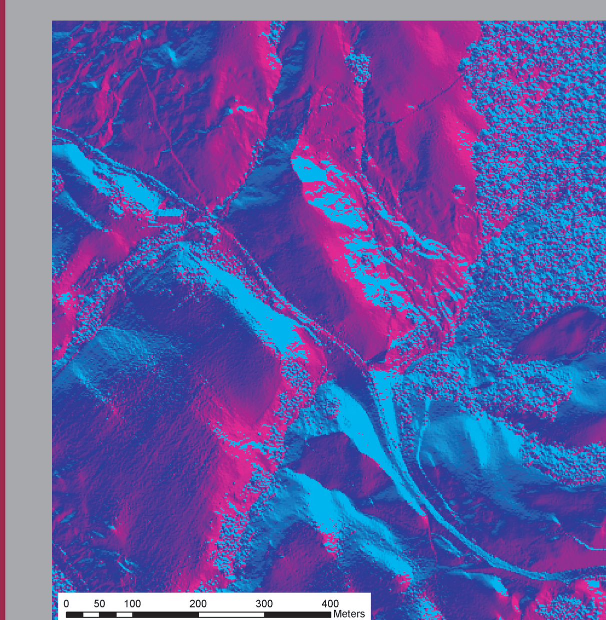


Hillshade of Digital Elevation Model (DEM) generated by GRASS spline interpolation algorithm.

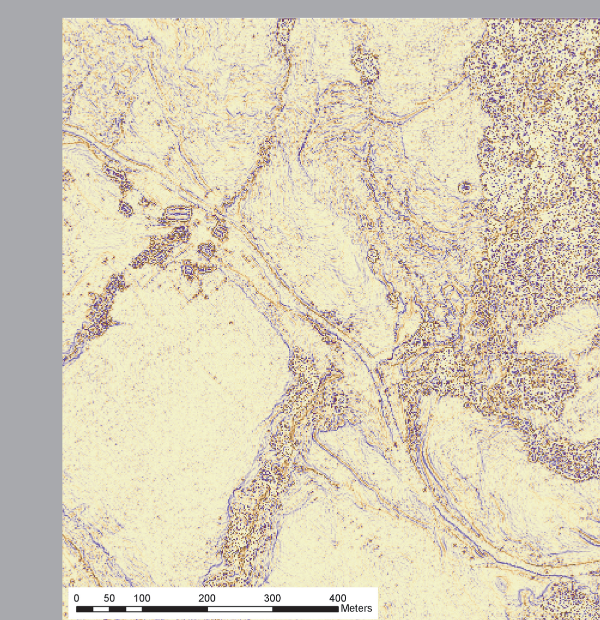
Derived products:



Slope map derived from spline DEM.

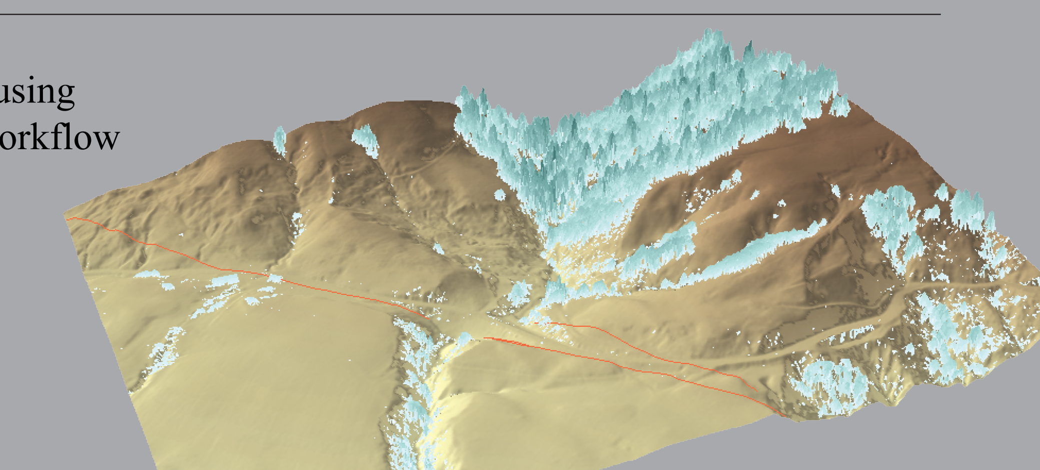


Aspect map derived from spline DEM shown above.



Profile curvature (PCurv) map derived from spline DEM shown above.

Figure produced using GEON LiDAR workflow outputs:



Right-lateral offset of Mill Gulch along the northern San Andreas fault. Trace of the 1906 earthquake rupture as mapped by F. Matthes (Lawson, 1908) shown in red.

V. COMING SOON

- Data sets:
  - NASA/USGS: W. Mt. Rainier
  - NCALM: Napa, E. CA. Shear Zone, Southern San Andreas Laser Scan
  - Hector Mine Earthquake
  - Geo-EarthScope?
- Additional interpolation algorithms and analysis tools:
  - Block mean
  - IDW
  - Hillshade
  - Watershed analysis
- Recruitment of additional GEON Grid clusters to increase computing resources available for LiDAR interpolation and analysis.
- Internet-based 2 and 3D visualization of LiDAR workflow outputs via ArcIMS and Fledermaus.

REFERENCES:

Neteler, M. and H. Mitasova, 2002, Open Source GIS: A GRASS GIS Approach, Kluwer Academic Publishers, 434 p.

This material is based upon work supported by the National Science Foundation under Grant No. 0225543 (GEON)