Open Data, Data Services, and Cross-Disciplinary Collaboration in Geophysics

Jeff Freymueller

University of Alaska Fairbanks

Outline

- Data: past present and future
- Open data: EarthScope defined the new normal
- Future integrated IRIS/UNAVCO data flows
- Seismogeodesy: An example of integrating cross-disciplinary types of data

We've Come a Long Way, Baby

Now: Online, Download

- Today, all data are online, but transfer of large data sets remains much slower than local access
  - Most people download once, use locally repeatedly (local archive)
- Repeated download beginning to replace local archive
- A challenge: How to deal with meta-data updates?
  - “Oops, the sensor was not installed pointing north”, “Sorry, we reported the wrong type of antenna.”

Future: Will we access on demand?

- Increasing network speeds follow a power law
- How will you work if networking speeds are ~100x faster?

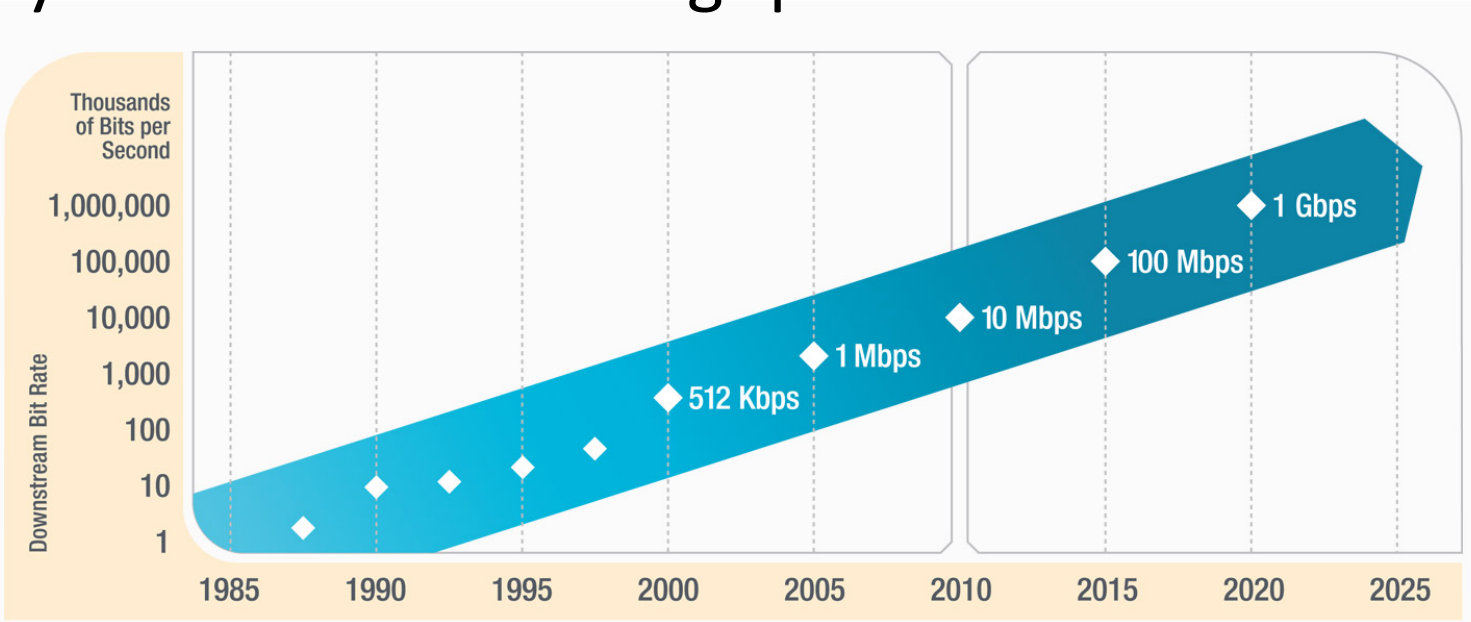
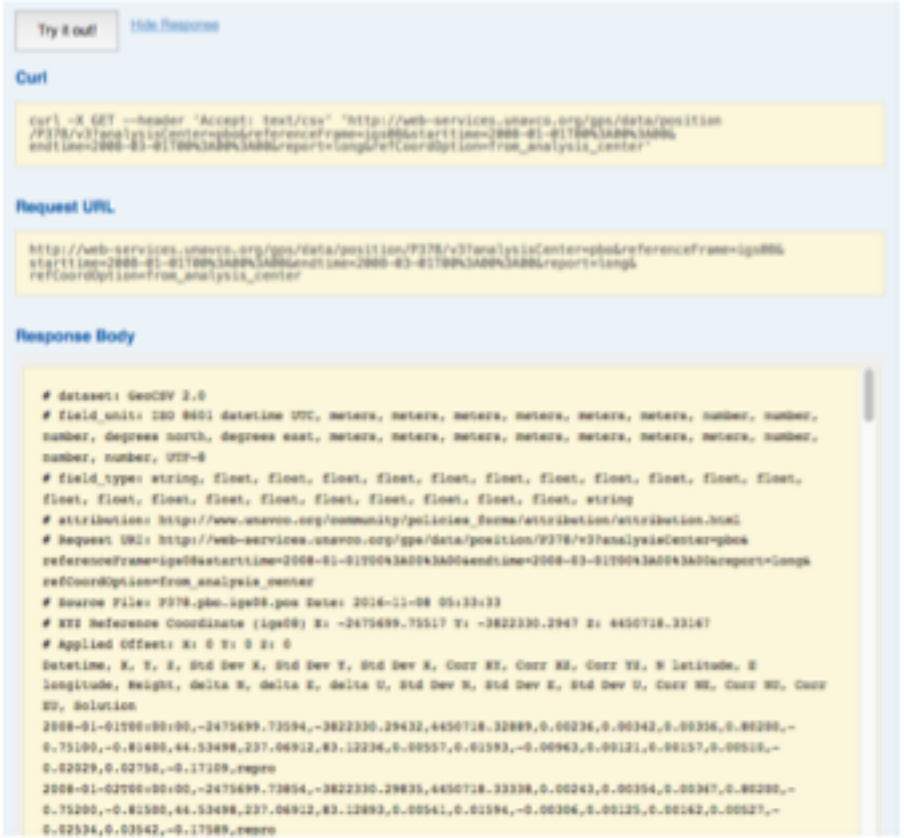


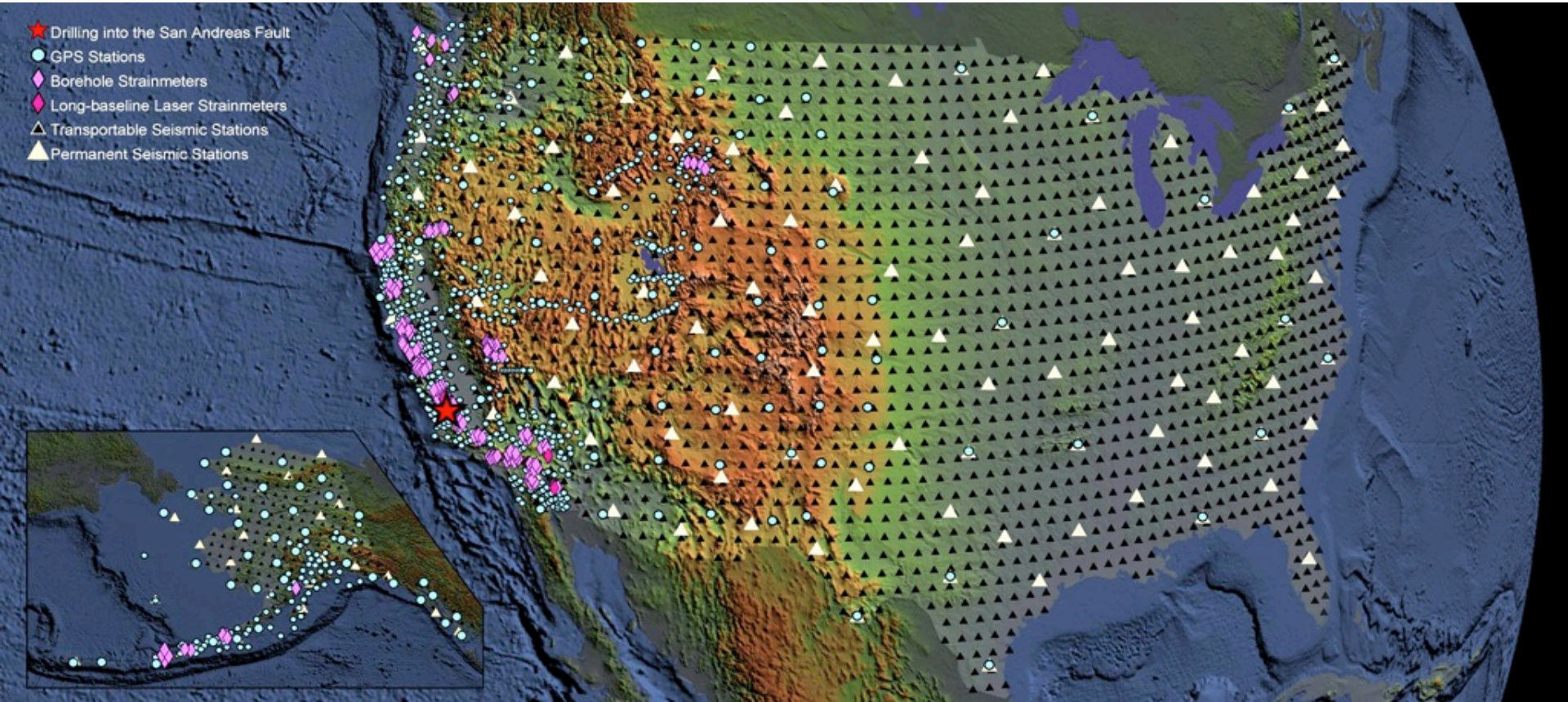
Image: Jacob Pedersen

Move to Web Services

- IRIS and UNAVCO are moving away from older transfer protocols to **web services**
- Efficiency, scalability
- Send a URL string, get a data set/product back

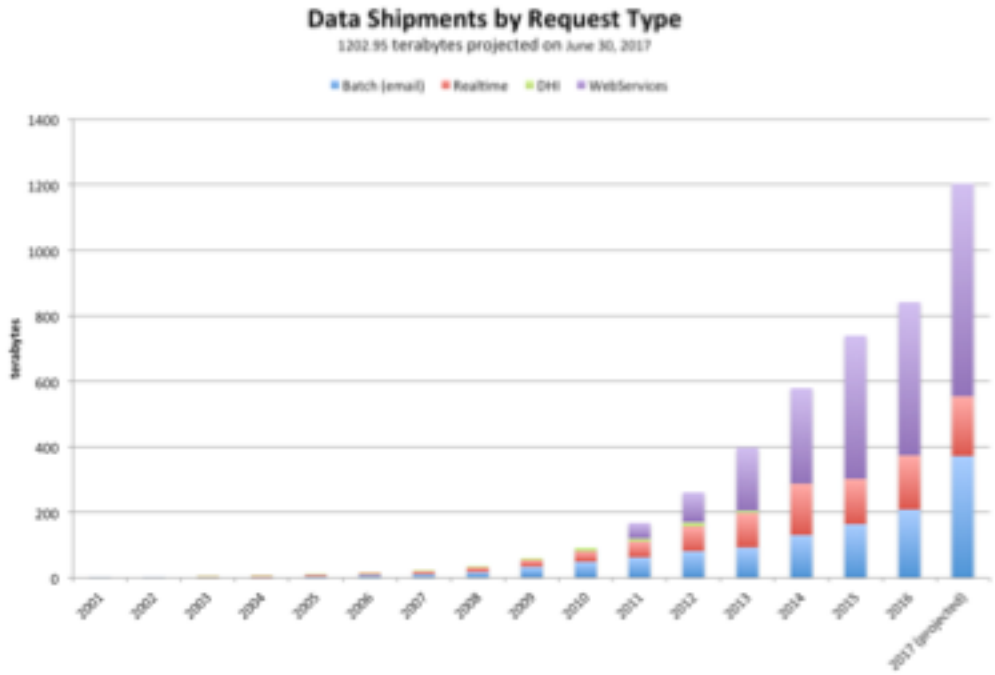


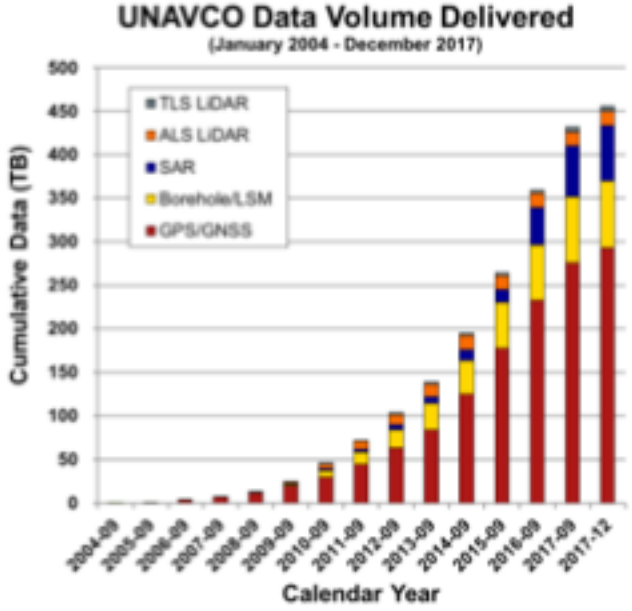
EarthScope and Open Data



EarthScope Data

The flood of open data from EarthScope changed the way we do science



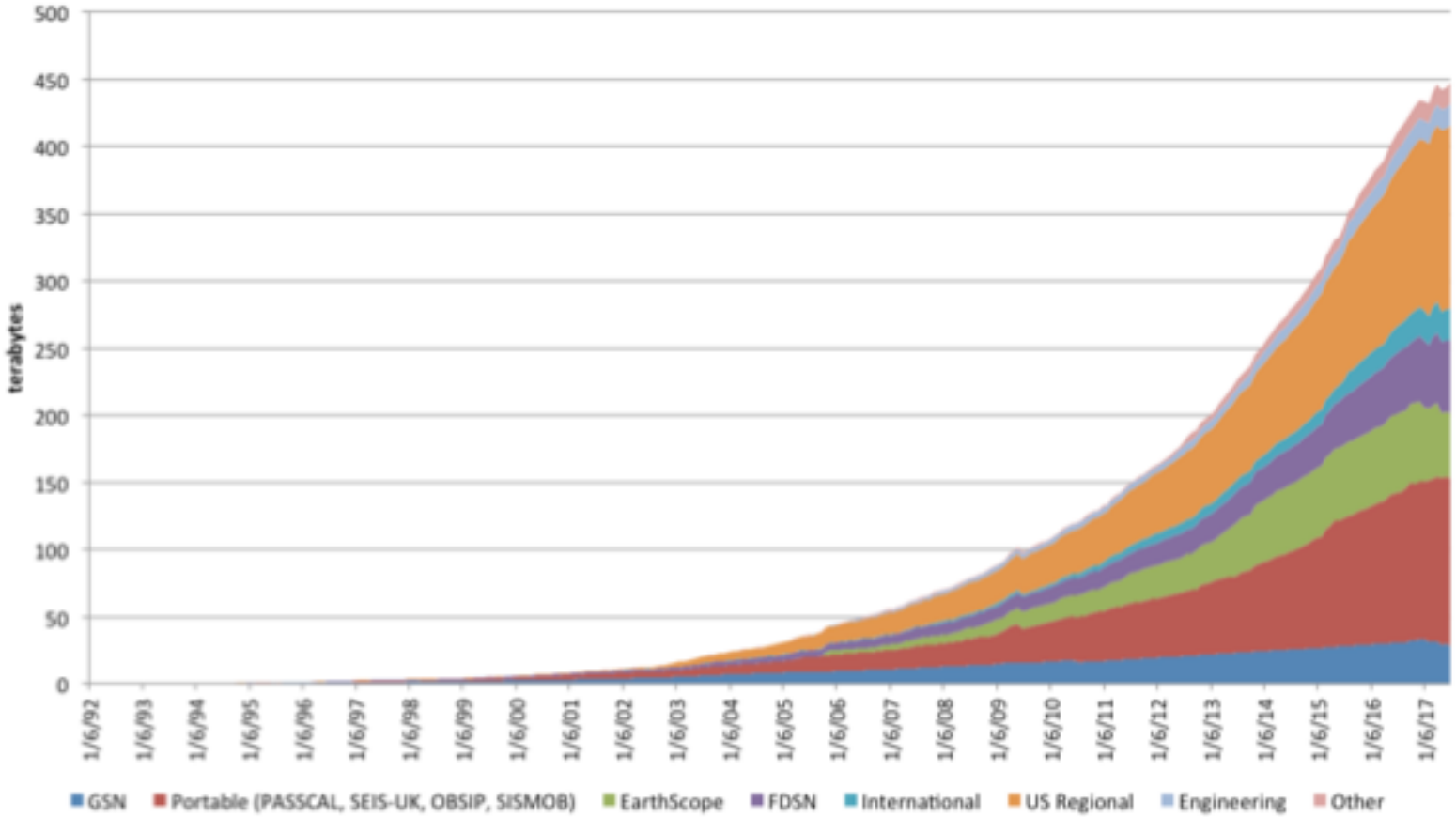


IRIS DMC Archive

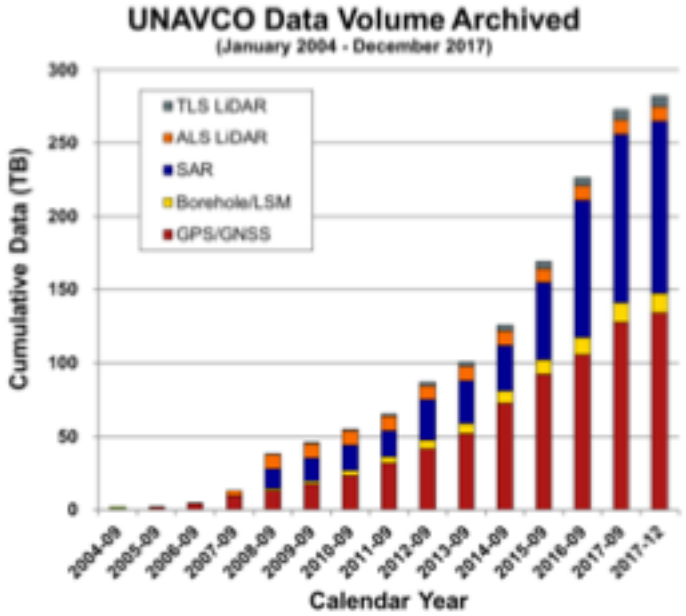
as of 1 July 2017

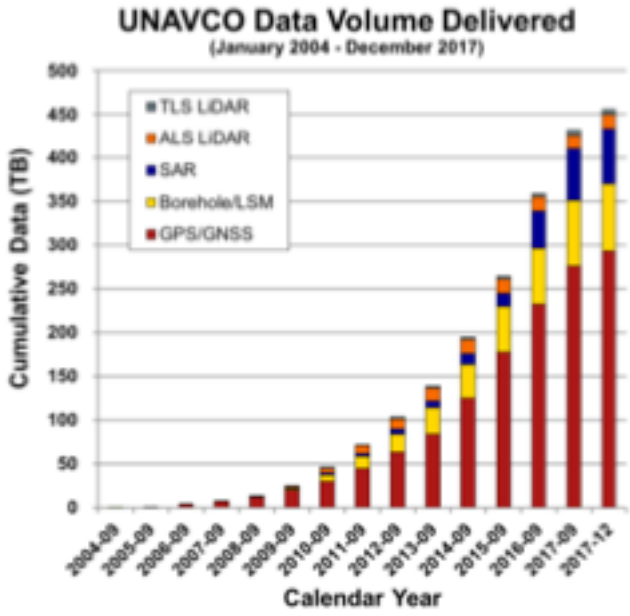
447.0 terabytes

reflects removal of replicated data



GAGE Data Metrics





- Cumulative Data Archived through Mar 2018 ~306 Tb
- Cumulative Data and Derived Data Products Delivered through Mar 2018 ~474 Tb
- Total EarthScope PBO Data Volume Archived (all products) = 145 Tb
- Total EarthScope PBO Data Volume Delivered (all products) = 349 Tb

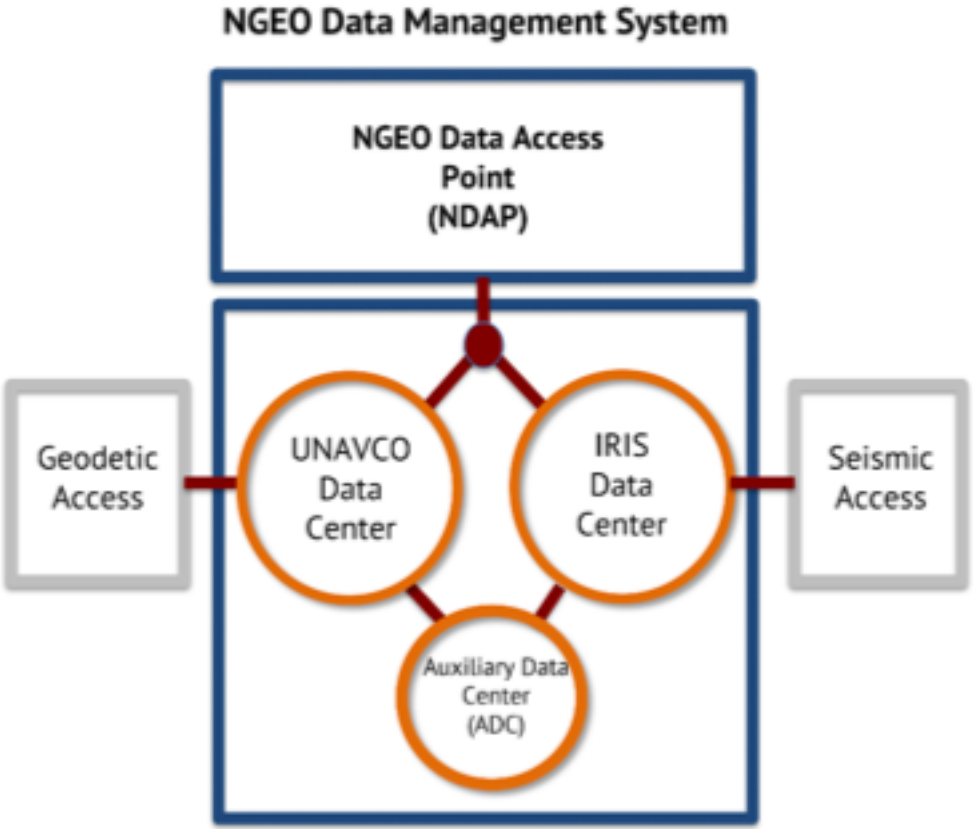
Impacts of Open Data

- The vast holdings of open data have accelerated the pace of our science
  - Analyses of data that used to be research are not done automatically
  - The research frontier is now more in synthesizing and integrating
- Cultural shifts are in progress…

The Future: IRIS and UNAVCO Together

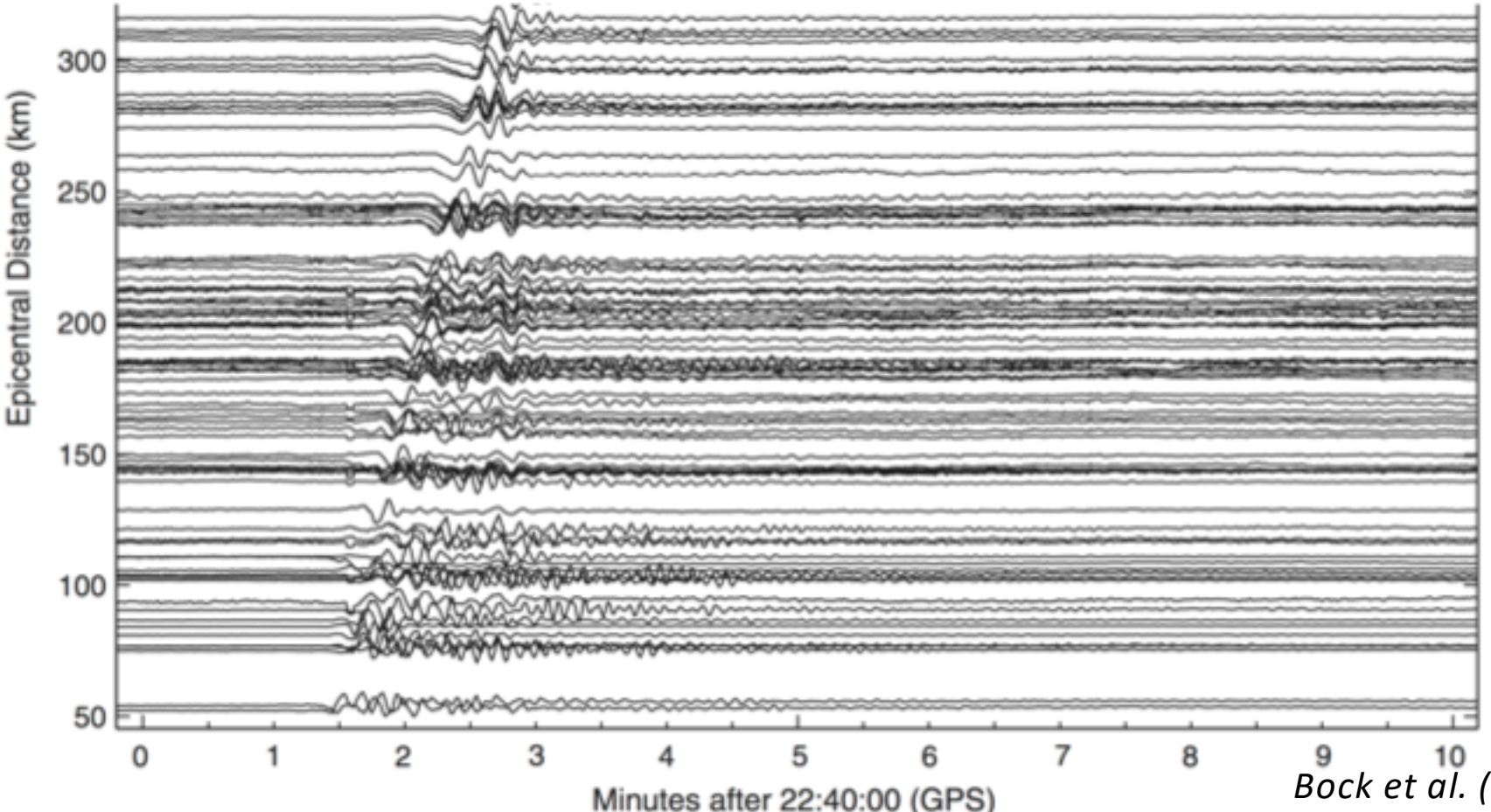
- IRIS and UNAVCO have proposed to work together more closely on the future NGEO facility.
  - NSF is still working on the NGEO decision…
- Seismologists and geodesists commonly work on different things, but we can make use of each others' products
- Seismogeodesy holds great promise for integrating the fields, at least for large earthquakes

Proposed NGEO Data Access Point



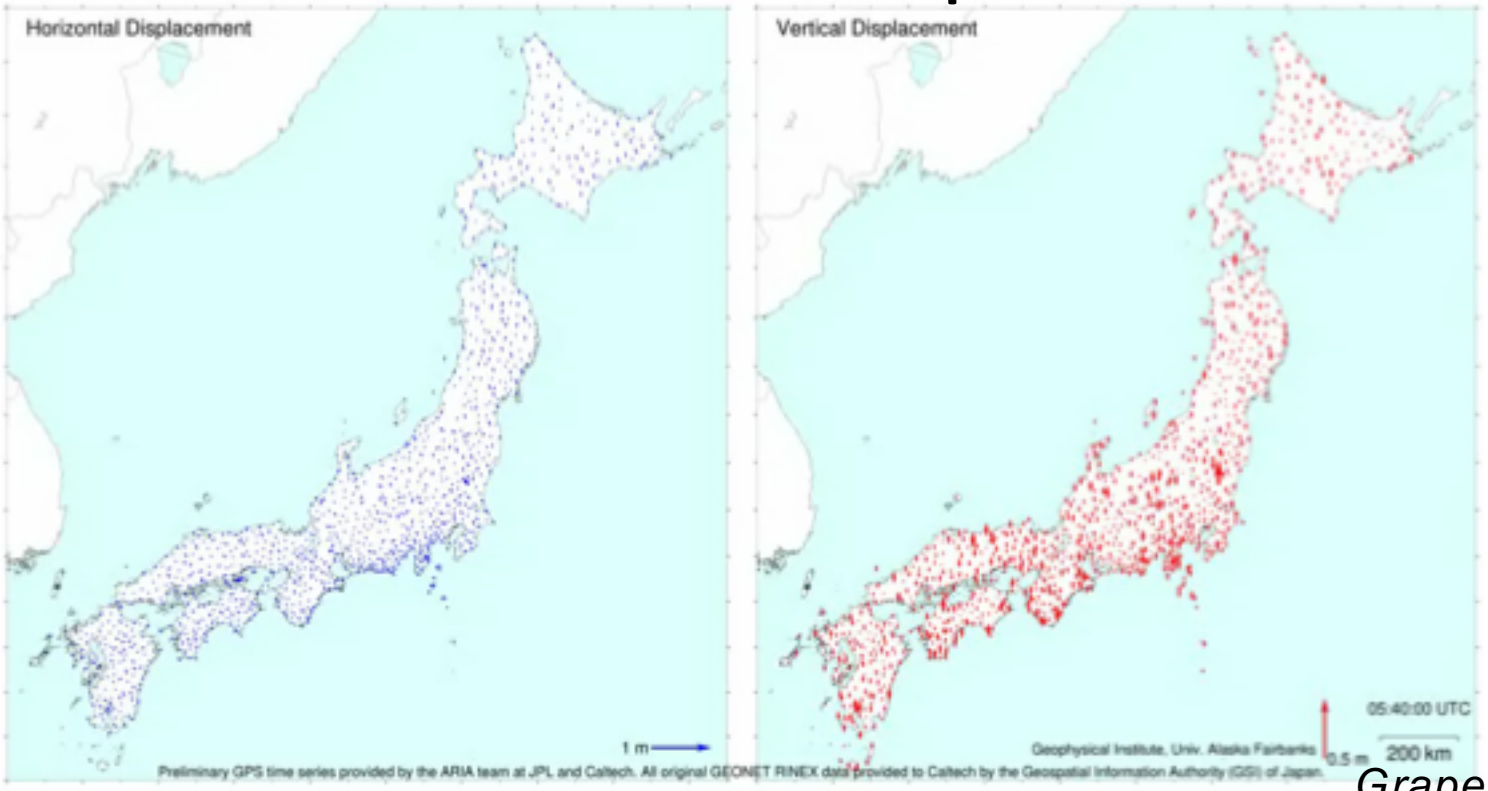
- Building on EarthCube GeoWS project, IRIS and UNAVCO proposed a unified NGEO data access point
  - Cross-disciplinary users are probably most likely to access products rather than raw data
  - Will work for raw data, too!
  - Future: GPS waveform data
- Easier to implement with modern protocols

Seismogeodesy



Bock et al. (2011)

Seismic Wavefield from 2011 Great East Japan Earthquake



Grapenthin et al. (2011)
Time series from JPL/ARIA

Seismogeodesy

- Looser definition:
  - Use of high rate GPS positions to study dynamic ground motions
  - Using geodetic data the way a seismologist would
- Tighter definition:
  - Combination of co-located geodetic and seismic instruments to produce a single time series of ground motions

Advantages and Applications

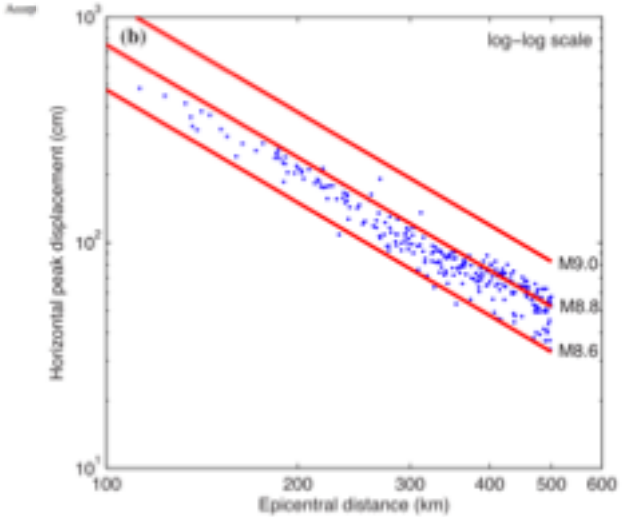
- Kinematic GPS position time series combine information about static and dynamic offsets
- GPS position/velocity records can be filtered to be used as seismograms
  - It is even possible to do tomography with GPS seismograms!
    - Why? For some places and times, there are more GPS than seismometers
- No instrumental saturation with GPS
  - No problems with clipping, ground tilts, etc

Ground Motion/Magnitude Scaling holds to M~9

Geophysical Journal International

Determination of earthquake magnitude using GPS displacement waveforms from real-time precise point positioning

Rongxin Fang, Chuang Shi, Weiwei Song, Guangxing Wang and Jingnan Liu



Back to Gutenberg (1945)!

GEOPHYSICAL RESEARCH LETTERS, VOL. 40, 6089–6094, doi:10.1002/2013GL058391, 2013

Earthquake magnitude scaling using seismogeodetic data

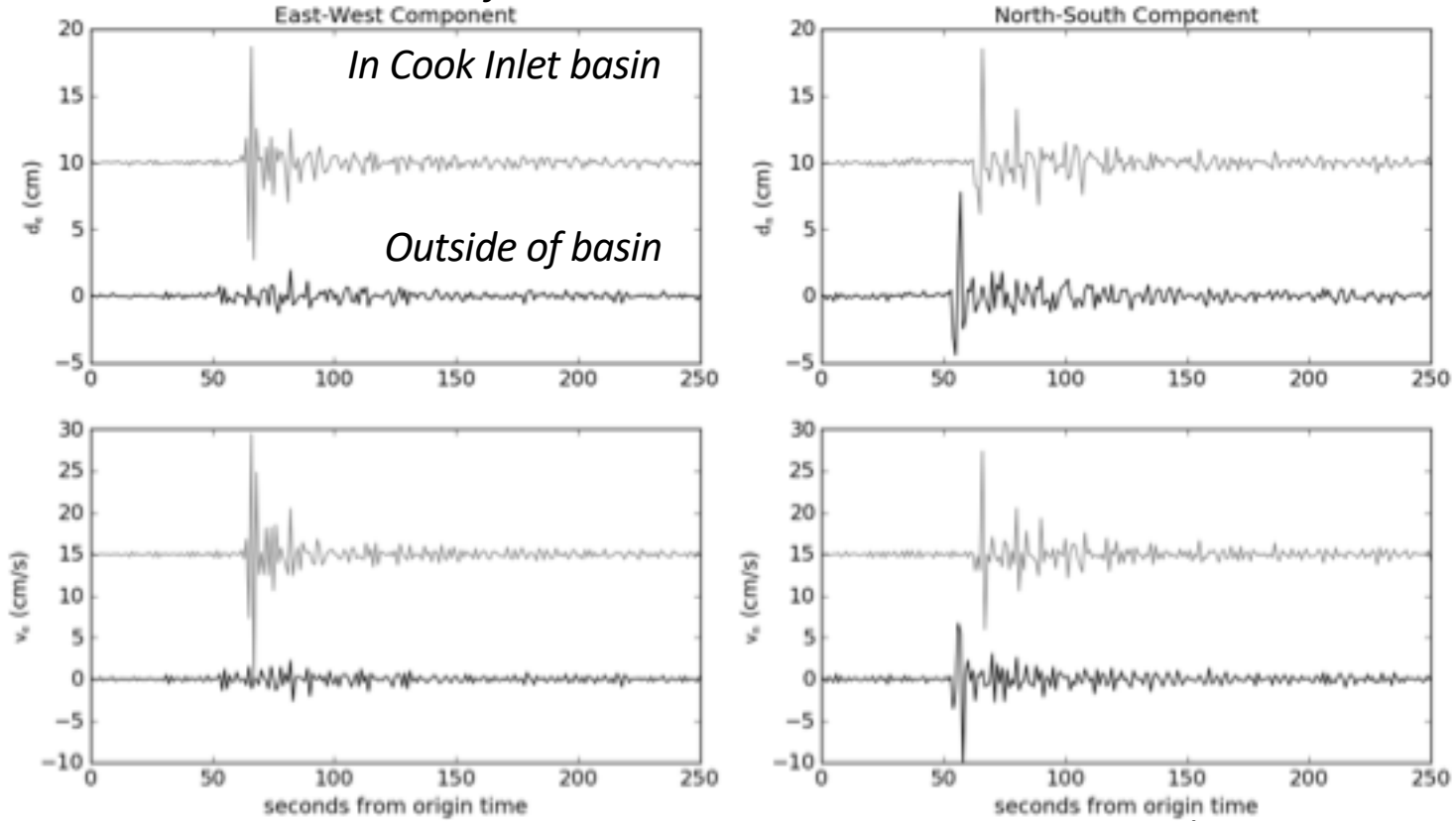
Brendan W. Crowell,[1,2] Diego Melgar,[2] Yehuda Bock,[2] Jennifer S. Haase,[2] and Jianghui Geng[2]

Received 21 October 2013; revised 26 November 2013; accepted 27 November 2013; published 10 December 2013.

$$\log(P_d) = -0.893 + 0.562M_w - 1.731\log(R) \qquad (2)$$

GPS Displacements/Velocities

filtered between 0.25-0.5Hz



2016 Iniskin earthquake, Alaska (Grapenthin et al., 2018)

Conclusions

- Open data and fast network speeds have changed/are changing the way we work
- Future scientific frontiers will increasingly depend on synthesis and integration of large data sets
- IRIS and UNAVCO are poised to help all of us exploit these opportunities
- There are exciting scientific opportunities in seismogeodesy, the fusion of seismology and geodesy