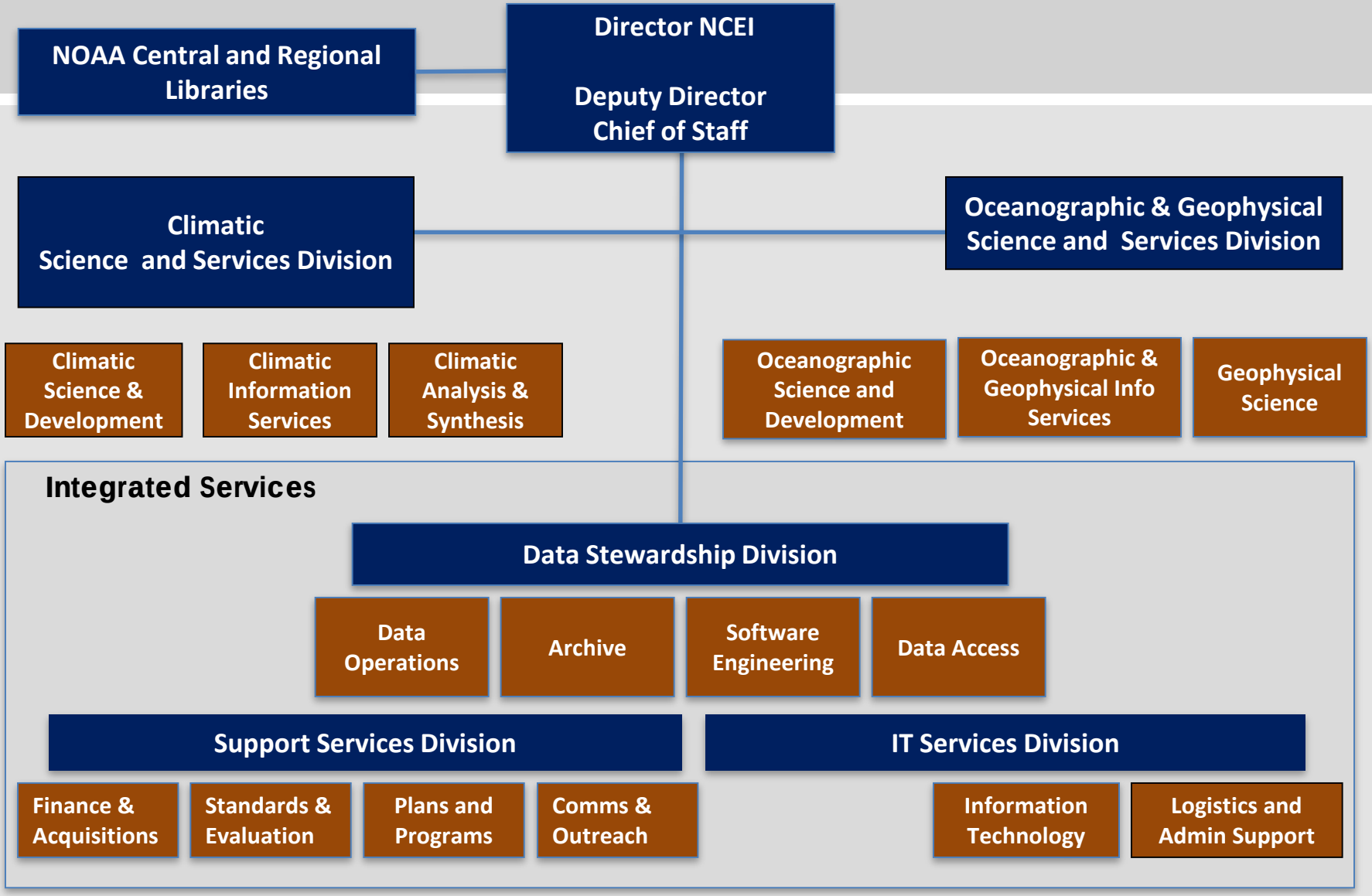
National Centers for Environmental Information Observational Data Workshop, May 22-24, 2018

Matthew Menne

NOAA National Centers for Environmental Information, Asheville, North Carolina, USA

National Center for Environmental Information (NCEI)



Data Ingest at NCEI

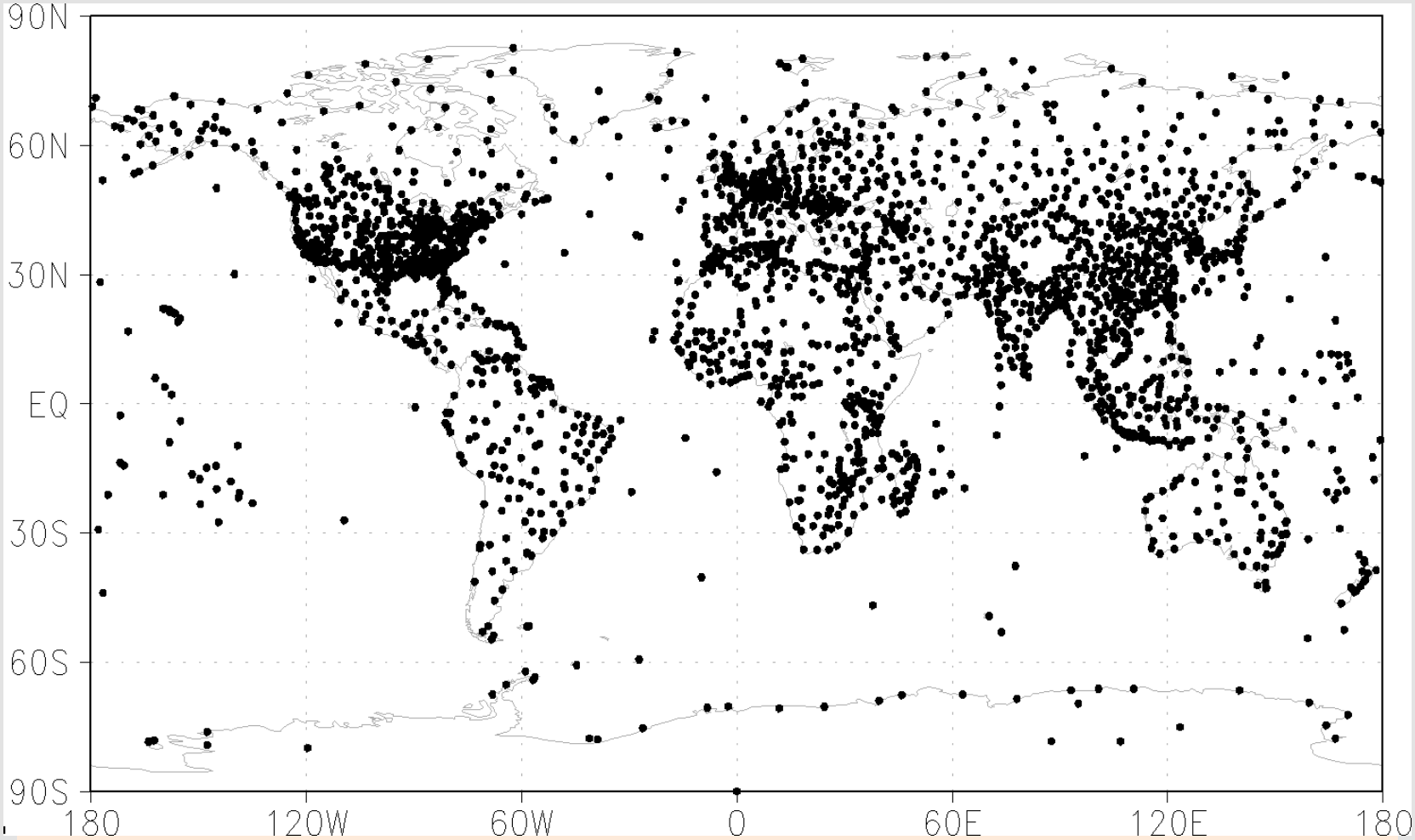
- ~250 data input streams archived according to specifications outlined in submission agreements (known as “Common Ingest’)
- Some are pushed to NCEI, others are pulled
- Some go straight into the archive (HDSS or CLASS)
- Others are put into the archive and then made available for further processing (as an “auth copy”)
- Tracking identifiers, locations, aliases etc. can be a big part of dataset management

Conventional Data

Data streams that get further processing fall more and more into several big, integrated data pots

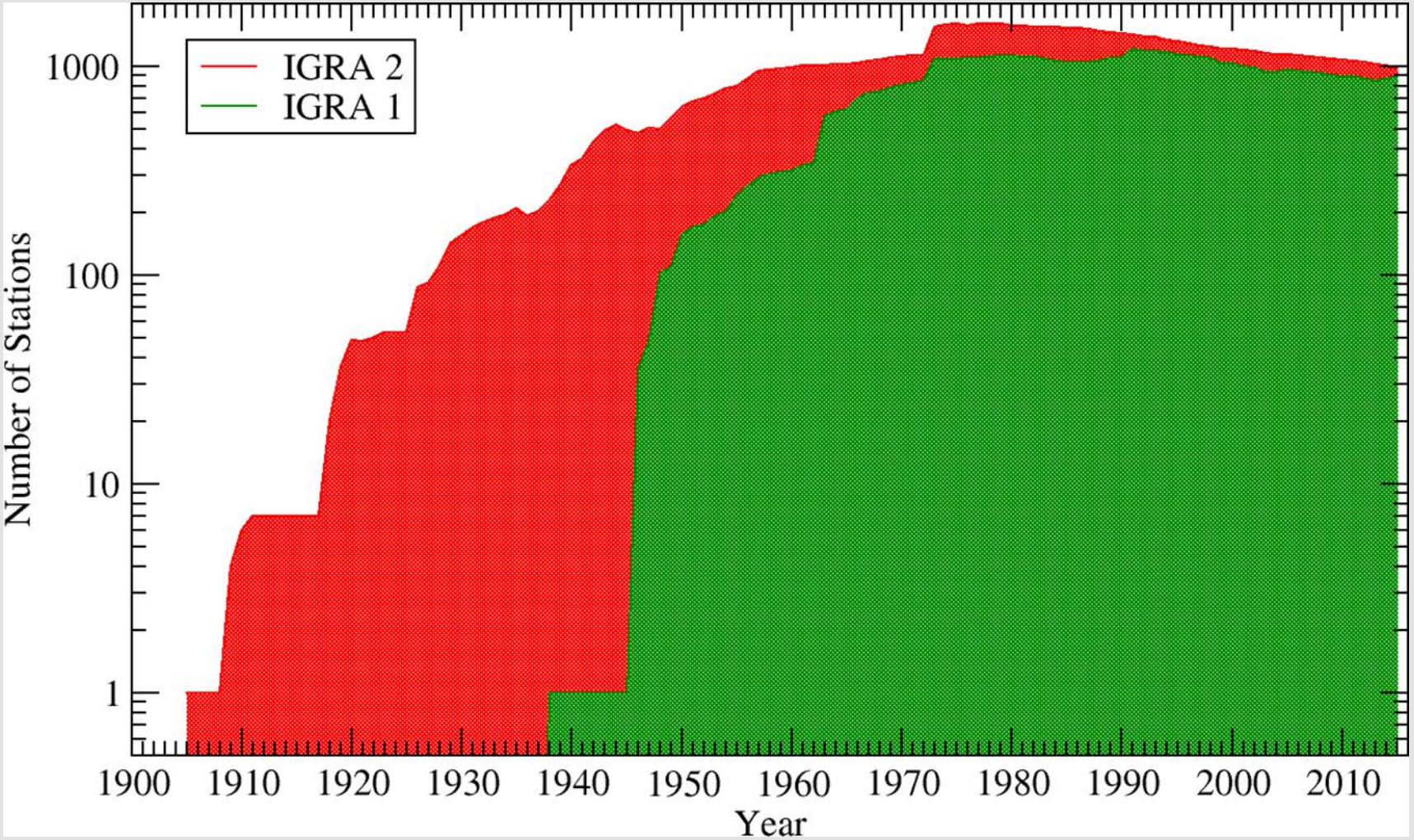
- –Integrated Global Radiosonde Archive (IGRA; Radiosonde)
- –Global Historical Climatology Network (GHCN; Land Station Data)
- –International Comprehensive Ocean-Atmosphere Data Set (ICAODS; Marine Surface)

IGRA 2 Station Map



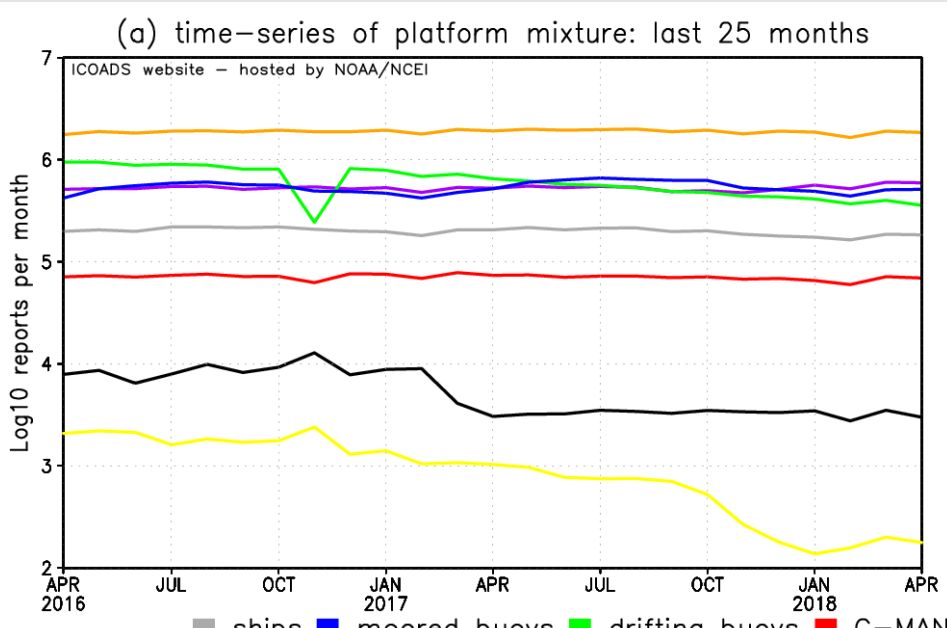
https://www.ncdc.noaa.gov/data-access/weather-balloon/integrated-global-radiosonde-archive/

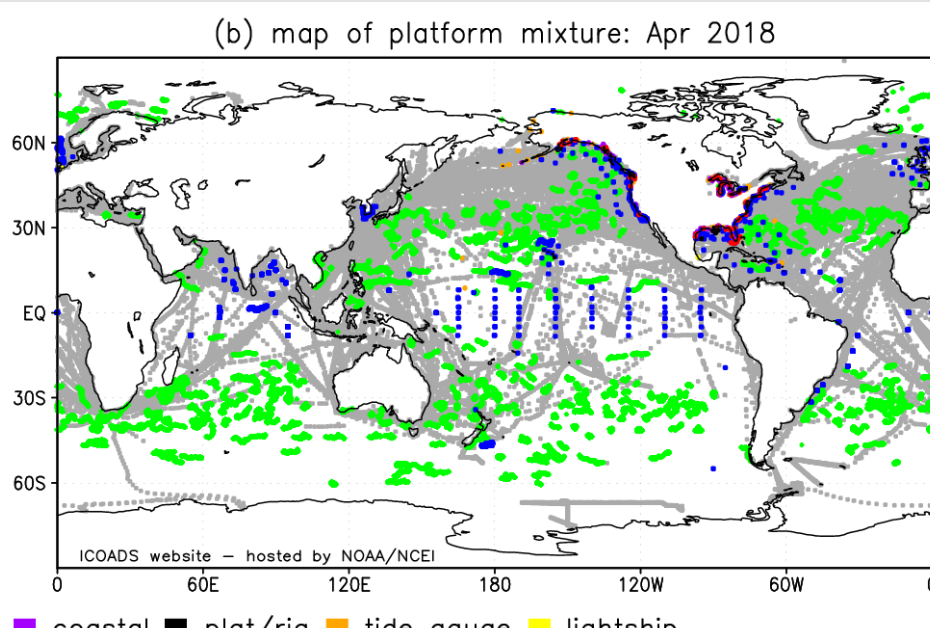
Number of Stations by Year in IGRA 1 and IGRA 2



Marine Observations At NCEI

- ICOADS R3.0.1
  - –Monthly updates, 2015-present
  - –Blend of Near-Real-Time GTS data from NCEI and NCEP
    - •Ingesting available platform types: ships, moored buoys, drifters, NDBC CMAN stations, coastal stations, platforms/oil rigs, tide gauges and lightships
    - •And all available parameters, e.g Air temp, pressure, SST, winds, etc.

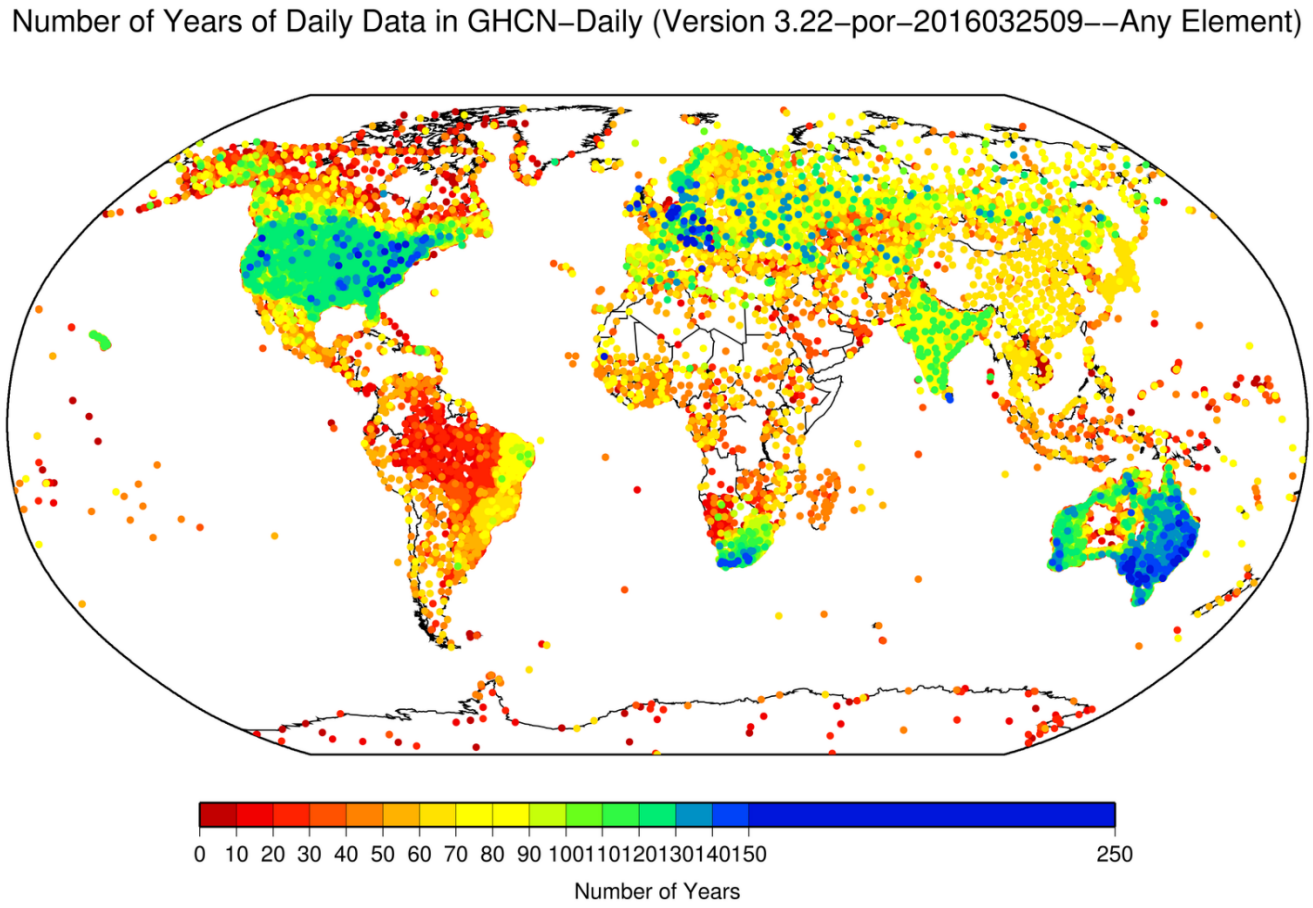




Land Surface Station Data

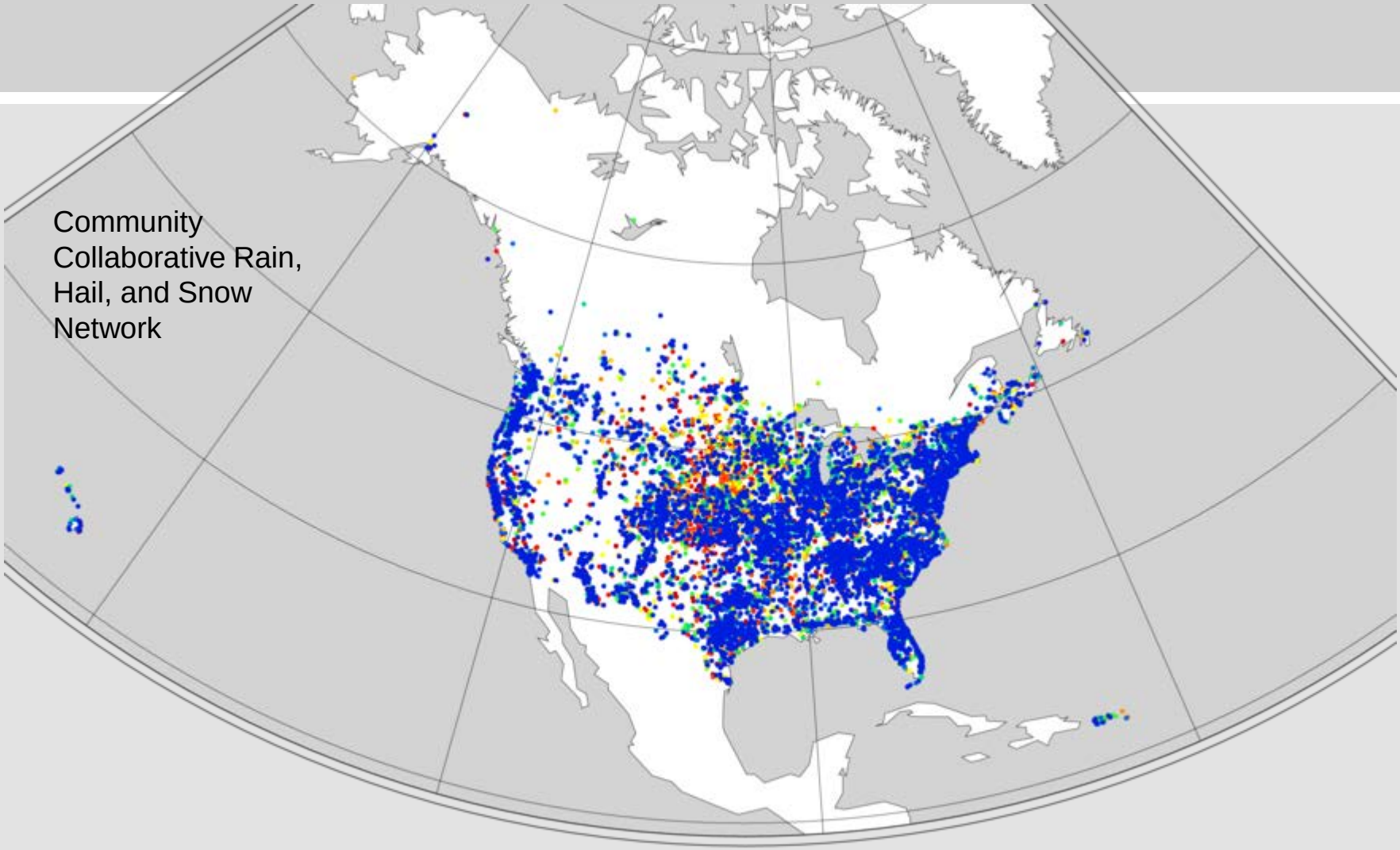
- Hourly Data
  - –GTS data exchange
  - –Data rescue
- Daily Data (no international exchange format until recently)
  - –National/Regional Archives on ftp
  - –Web services
  - –Data Rescue
  - –GTS SYNOP messages
- Monthly Data
  - –National Archives
  - –Data Rescue
  - –CLIMAT message exchange over the GTS (FM-71)

Land Station Observations

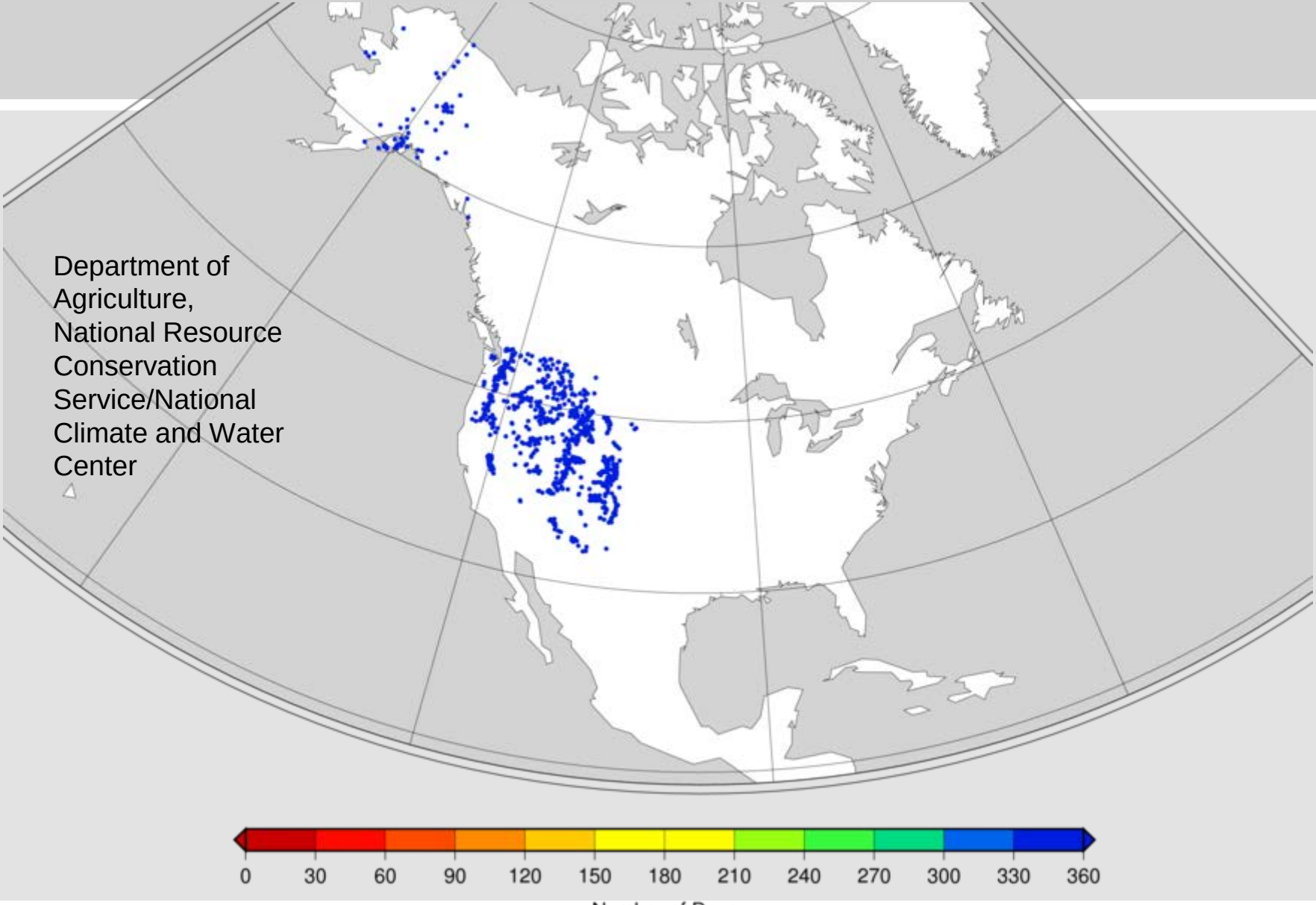


CoCoRaHS in 2017

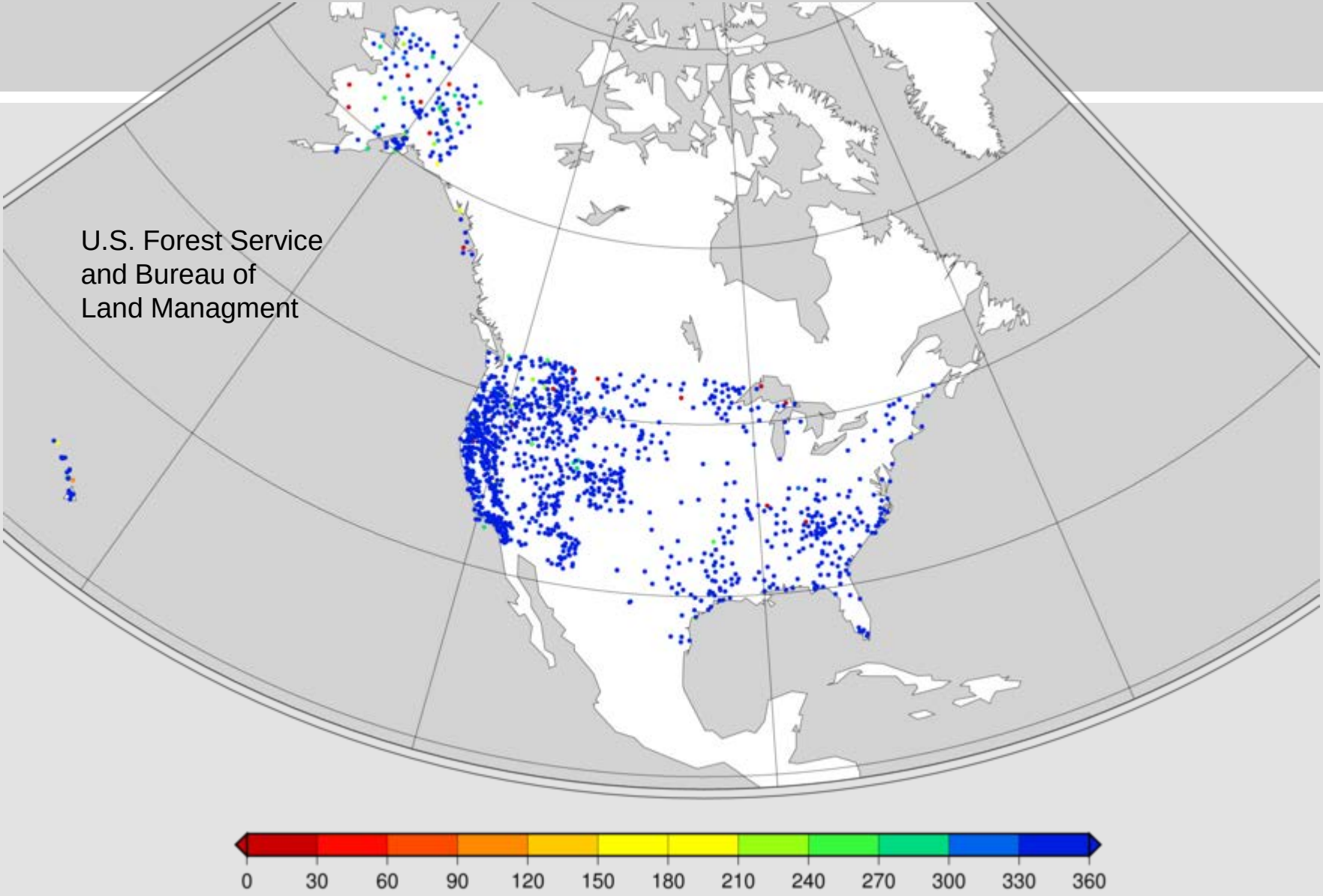
Community Collaborative Rain, Hail, and Snow Network



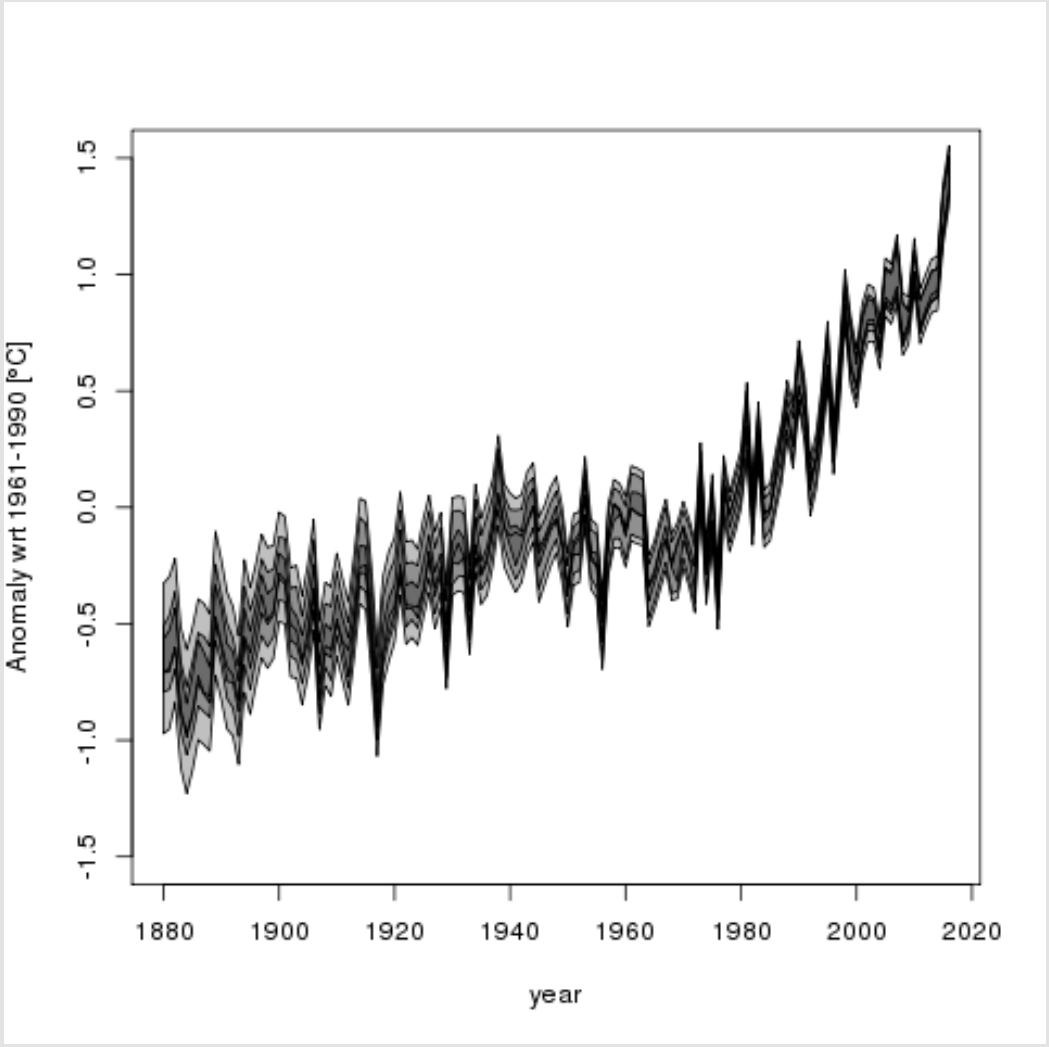
SNOTEL (SNOw TELemetry) in 2017



RAWS (Remote Automated Weather Stations) in 2017



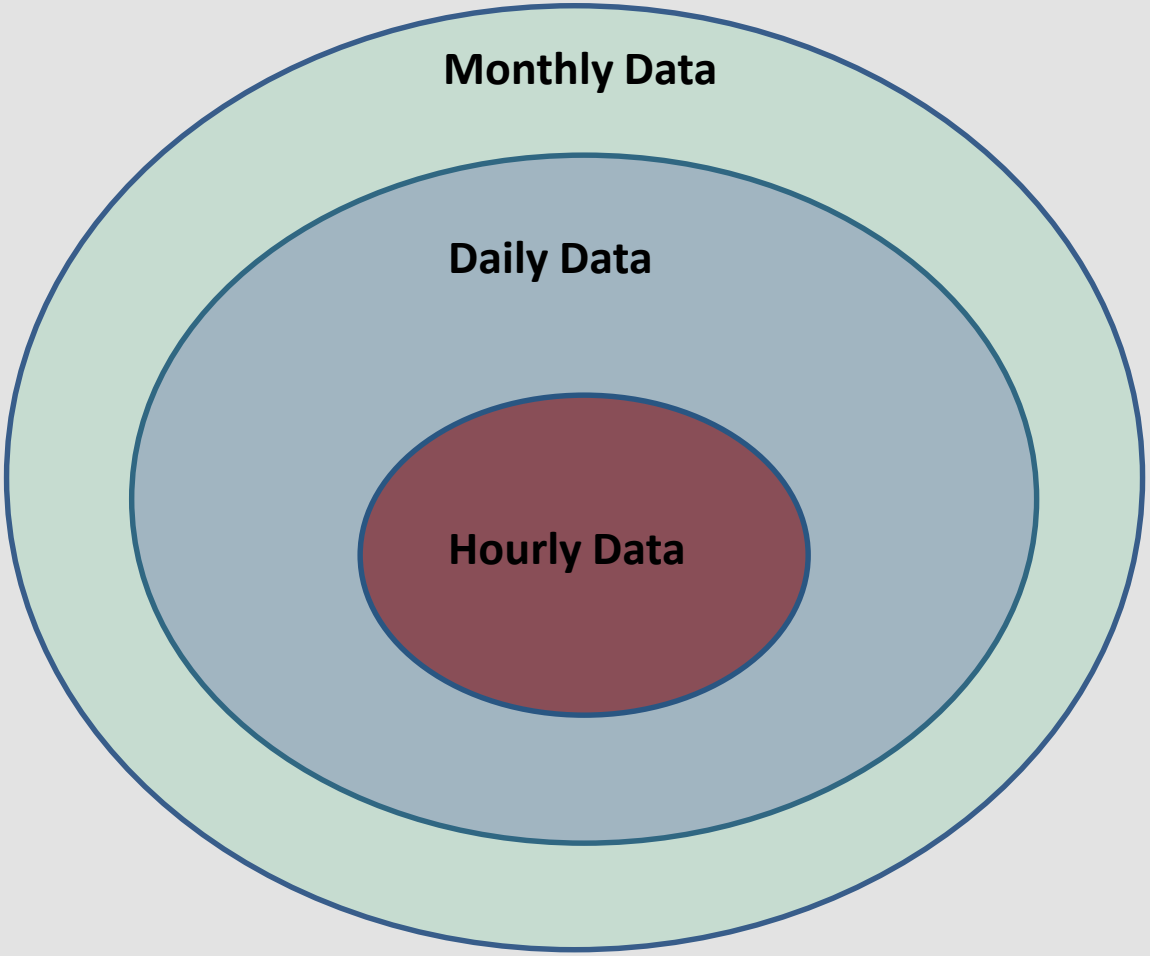
Global land surface air temperature anomalies for GHCN monthly v4 with uncertainties



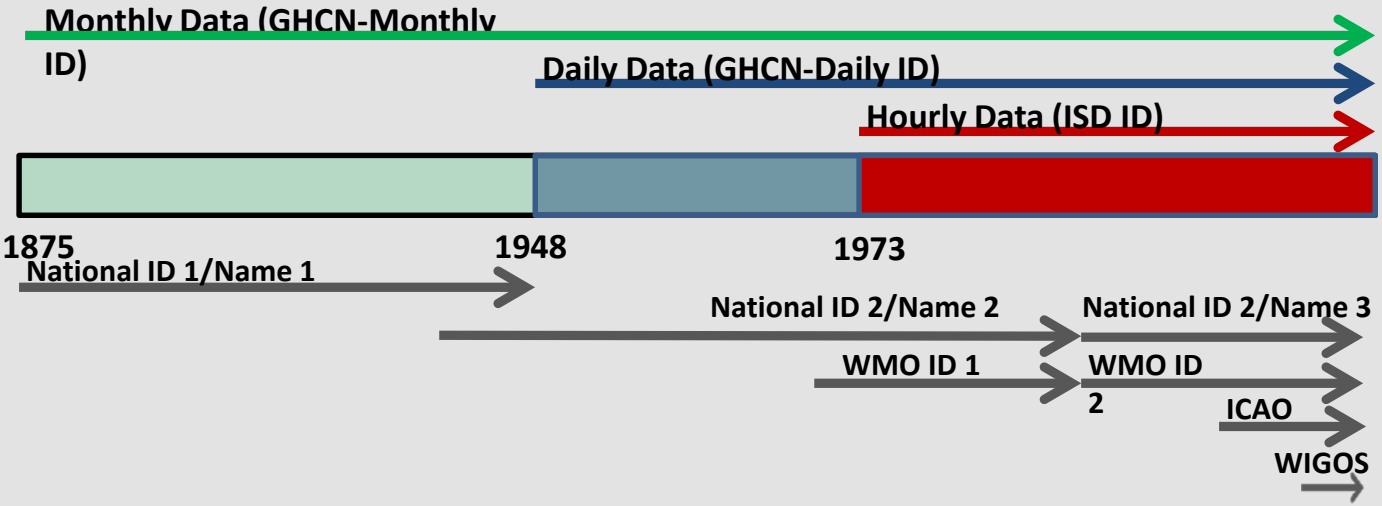
Digital Archives

Datasets covering three different time resolutions for land stations were developed and have evolved independently (and are among the most popular of all NCEI products)

- **Hourly** (Integrated Surface Daily) and **Daily** (GHCN-Daily) **datasets** account for about **two-thirds** of data orders handled by User Engagement Branch **and have many tens of thousands of unique users online** each month from all sectors (agriculture, energy, insurance, legal, engineering, education, science etc.)
- Monthly data support climate monitoring (e.g., drought), normals and other areas



Hypothetical Land Station Record
"Station X"



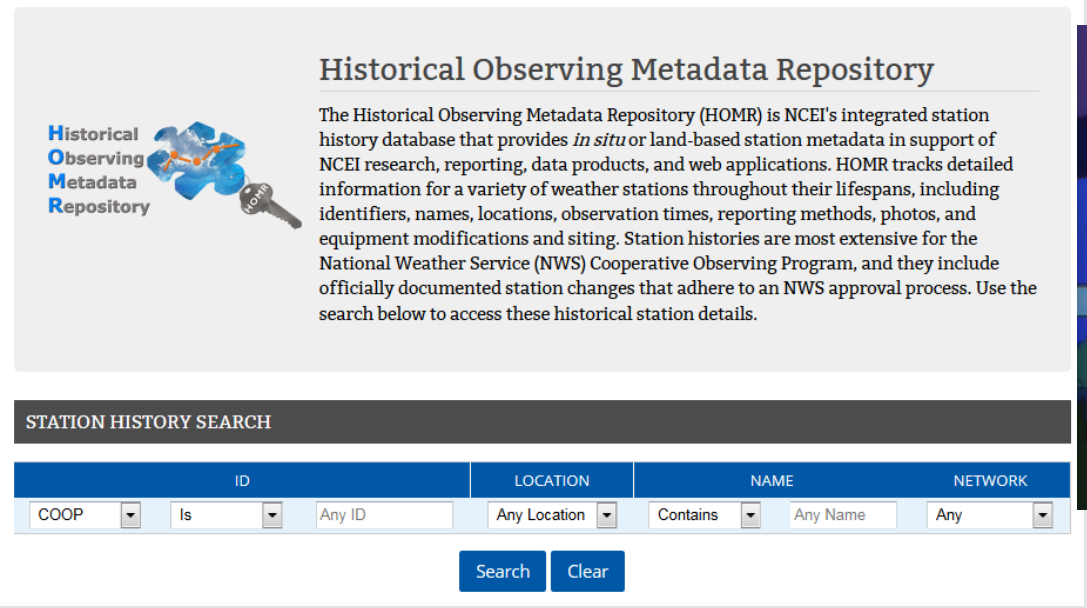
Identifiers	Consolidation of IDs over time (ICAO, WBAN, FAA, WMO, COOP, WIGOS…)
Names	Stations can have many aliases
Locations	Latitude/longitude, elevations, topography, obstructions, relocations
Elements	Observation times, reporting methods
Equipment	Types, modifications and siting

Characteristics of Latest Conventional Dataset Curation

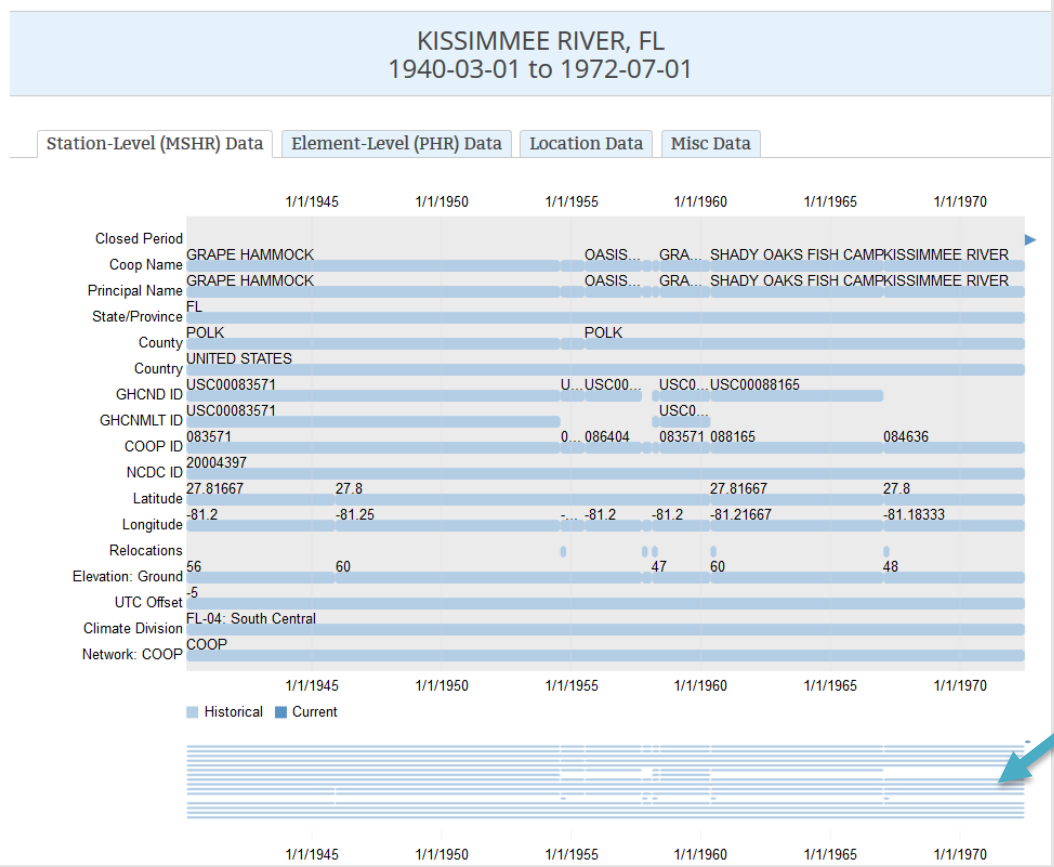
- Centralized collection of data archives (under World Data Center system)
- Reformatting native digital formats from numerous sources to a common standard
- Short time delay updates for weather and climate monitoring
  - primarily via GTS and web services or simple ftp protocols
- Regular resynchronization of source databases with integrated dataset
- Mechanism for ongoing integration of newly available data sources
  - via data and metadata matching algorithms
- A system for documenting, tracking, and resolving errors
  - This is done for GHCN-Daily and some other datasets at NCEI via the Datzilla system
- Management of station histories and other metadata that are consistent with data records (e.g., aliases, location and instrument changes, etc.)
  - Accomplished by the NCEI Historical Observing Metadata Repository (HOMR)

NCEI's Station History Database / Web Service

https://www.ncdc.noaa.gov/homr



HOMR serves up station chronologies from database using D3 Javascript visualization library

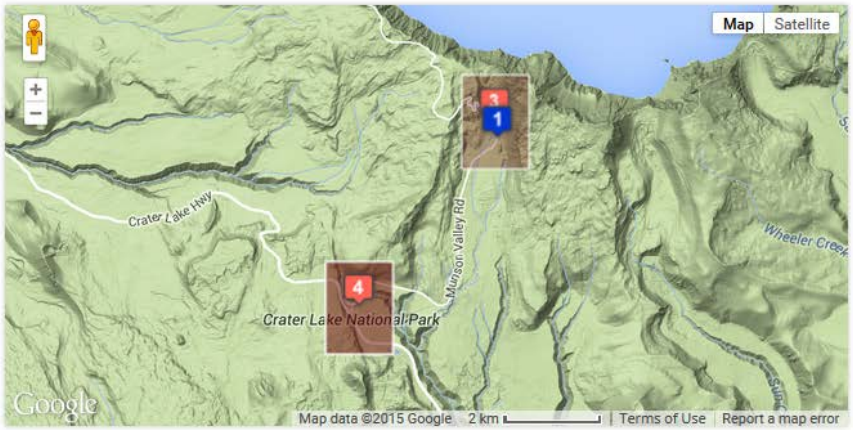


6 COOP ID changes, 6 name changes, same station. Old policy was new ID required for name change.

Using date Focus Chart, can dynamically zoom in/out for more detail

CRATER LAKE NPS HQ, OR
1919-10-06 to Present

Station-Level (MSHR) Data | Element-Level (PHR) Data | Location Data | Misc Data



Low-Precision Map Locations

Descriptions | Topography | Relocations | Obstructions

Date Ranges (Show All / Hide All) — Toggle lat/lon format

2001-08-14 to Present	42.8966, -122.1327 (from DDMMSS) 1
1997-08-04 to 2001-08-14	42.8967, -122.1328 (from DDMMSS) 2
1930-06-18 to 1997-08-04	42.9, -122.13333 (from DDMM) 3
1926-11-01 to 1930-06-18	42.866667, -122.166667 (from DDMM) 4

ID was kept between locations, but very different locations

WILMINGTON WB CITY, NC
1870-12-12 to 1951-10-02

Station-Level (MSHR) Data | Element-Level (PHR) Data | Location Data | Misc Data

✔ Low-Precision Map Locations

✔ Descriptions ✔ Relocations

Date Ranges (Show All / Hide All) Toggle lat/lon format

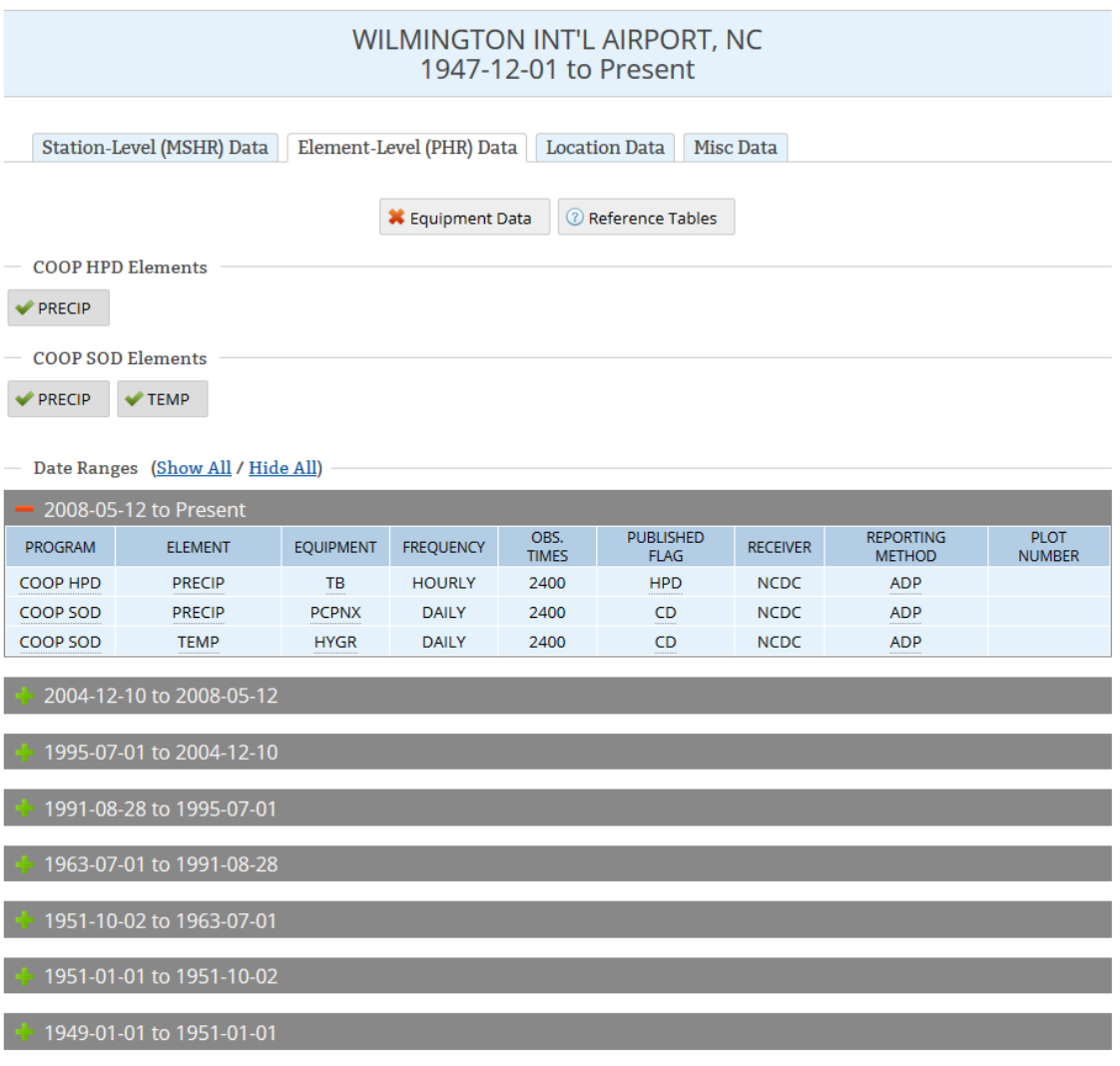
1870-12-12 to 1951-10-02 34°14'N, 077°57'W 1

Description	
1931-11-24 to 1951-10-02	U.S. CUSTOM HOUSE, WATER ST. BETWEEN MKT AND PRINCESS STS.
1890-07-01 to 1931-11-24	U.S. POST OFFICE BUILDING FRONT AND CHESTNUT STREETS
1881-01-01 to 1890-07-01	FIRST NATIONAL BANK 19 NORTH FRONT STREET
1879-11-27 to 1881-01-01	FRONT STREET, BETWEEN MARKET AND PRINCESS STREETS
1874-12-01 to 1879-11-27	BANK OF NEW HANOVER (4TH FLOOR) FRONT STREET, BETWEEN MKT & PRINCESS
1870-12-12 to 1874-12-01	MARKET AND WATER STREETS (SOUTHEAST CORNER)

Relocations	
1931-11-24	600 ft SW
1890-07-01	500 ft N
1881-01-01	100 ft N
1874-12-01	300 ft NE

Station locations are reported in varying degrees of precision and accuracy. This is especially a problem with older stations, which were located using now outdated techniques and technology. A location with low precision cannot be accurately displayed as a point on a map. In these cases we show the smallest area that the precision can contain, i.e. the station was located somewhere in the bounding box. The marker represents the center of the box, and is for reference only.

Location descriptions and relocations are valuable with older stations where the less precise latitude/longitude did not change, but the station did move. Several NCEI data products use this supplemental info in their automated QA.



Shows historical changes in equipment, observation times by element

Thank You!