CADB
version 2

Melissa Ou

Team: Melissa Ou, Tom Collow, Donald Garrett, Adam Hartman

What is the CADB?

Climate Assessment Database (CADB) - Suite of of global station summary analyses

- Input: Global Telecommunications System (GTS), which is a global network of transmitted meteorological data from various sources of *hourly* data.
- Output: Estimated summary values for daily, weekly, monthly, and seasonal timescales.
- Data is considered **preliminary** (**QC lite**).
  - Other datasets, such as those from NCEI or CPC unified gauge dataset utilizes more complex post-processing techniques (the CADB is one of the primary input sources for the CPC unified gauge dataset).
  - This is **not intended to be a high quality controlled dataset** with complex post-processing, such as neighboring check techniques.
- **Benefit**: Data is available in a timely manner for real-time operational needs because of the lighter QC-ing. NCEI data takes more time to process because it utilizes more complex QC techniques.

Users: Joint Agricultural-Weather Facility (JAWF), USDA, internal CPC users (e.g. international desk, verification), private companies, etc.

Related Upgrades and changes (v2 versus v1)

- Version 2 implemented Jan 2020, v2 data starts Jan 1, 2020
- Expanded public data:
  - Global coverage (previously U.S.), ~12,000 stations currently
  - More parameters
- Well documented - Documentation is maintained with changes. Public technical document notifies of changes.
- Daily summary code in version control (Git) and internal Github
- Better output format:
  - Comma-separated values (v1 was fixed width, harder to read in)
  - Values have decimals! No need for users to determine decimals
  - Easier to read in output CSV files
  - Daily, weekly, and monthly CSV files have similar formats and header names
- Use of better station metadata
- Postgres relational database replaces legacy direct access file

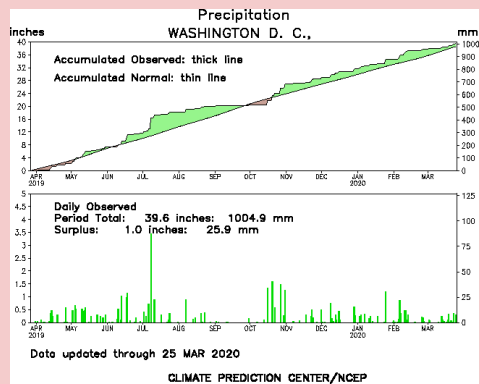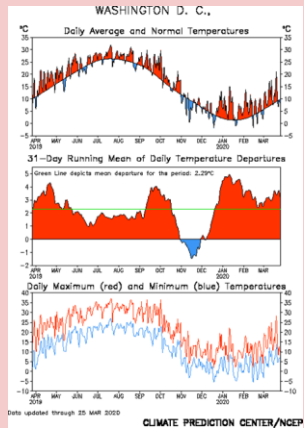# Station list metadata upgrade

- Overall challenge - many sources have conflicting metadata, reuse of a station ID number for different station locations (difficult for normals)
- New Python software that pulls in multiple metadata sources, using complex QC to provide an optimal station list metadata file (NCEI, aviation weather, CPC)
  - Previously more ad hoc scripts used, fewer input sources, most changes were sporadic and manual.
- More methodic way to upgrade, maintain, and add new stations from GTS data
- New format with more metadata including matched GHCND ID
- Increased spatial coverage for normals, better quality summary data (correctly matched station call/WMO ID streams of data), and improved quality for downstream apps (map plotting, timeseries information, etc.)

Coming soon: 1991-2020 Station normals

- Currently in review by management
- Upgraded python software to create monthly and daily station normals
- Software merges various sources of data for an optimized station normals dataset
  - Derived from CADB historical daily
  - Derived from GHCND monthly averages (NCEI)
  - NCEI "official" normals (currently only U.S.)
- QC-ing: compares normals to previous legacy 30yr normals, throws out values that are significantly different
- Previous (legacy) station normals little to no QC, normals used years outside of the expected period

New CADB Timeseries Web Application

Legacy static maps

Interactive web app - data served by Postgres database

Thanks to Ada Uzoma, David Stroud, and Tom Collow!

Customizable date

Sample temperature station coverage (many more stations now)

Hover for values

Download graphic options

https://www.cpc.ncep.noaa.gov/products/timeseries/
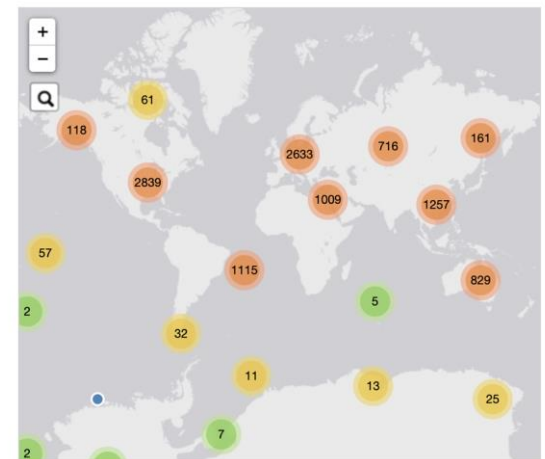
Thank you!

How to access data

Public data download page (including link to technical notes for changes): https://www.cpc.ncep.noaa.gov/products/cadb/

Public FTP access: https://ftp.cpc.ncep.noaa.gov/cadb_v2/

Contacts: (Daily CADB, normals, meta-data, timeseries app) melissa.ou@noaa.gov, adam.hartman@noaa.gov

Weekly, monthly, city data thomas.collow@noaa.gov,