

Two Datasets with Long Names: ISPD and GHCN

Russell S. Vose
Product Development Branch
National Climatic Data Center



NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

A Few Governing Principles

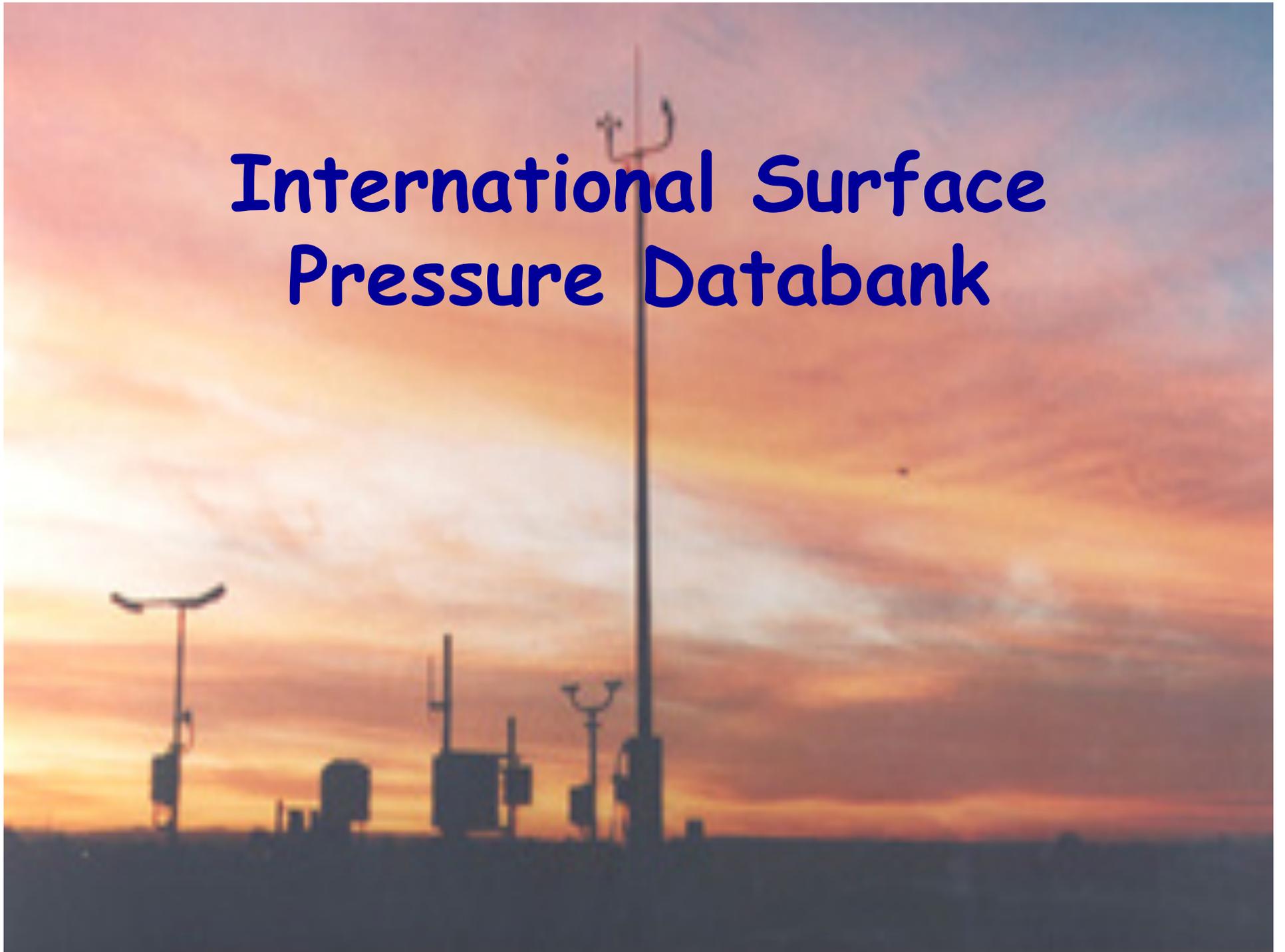
- **Integration**
 - Combine individual sources into a global dataset
 - Update regularly to the extent possible
- **Provenance**
 - Label source of each observation
 - Quality assure sources and data values
- **Openness**
 - Provide datasets to everyone free of charge
 - Incorporate user feedback



NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

International Surface Pressure Databank



ISPD Also Means

International Society for Peritoneal Dialysis
Institute of School and Parish Development
International Society for Prenatal Diagnosis
Indian Society for Pediatric Dermatology
International School of Protocol & Diplomacy
Illinois Society of Pediatric Dentistry



NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

Background

- GCOS/AOPC/OOPC Surface Pressure Working Group (SPWG)
- Identify, retrieve, digitize, integrate, and Q/A subdaily surface pressure data
- Basis for improved circulation fields and new reanalysis products (C2OR, ERA-CLIM)
- X. Yin as the major player



NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

ISPD v2.2.4 Sources

- **Global**
 - NCDC Integrated Surface Dataset (ISD)
 - NCEP-NCAR BUFR archive
- **National**
 - Australian Bureau of Meteorology
 - South African Weather Service
- **Miscellaneous**
 - Climate Database Modernization Program (U.S. and abroad)
 - Hadley Centre (old records for key stations)
- **33 sources**



NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

A Few ISPD Numbers

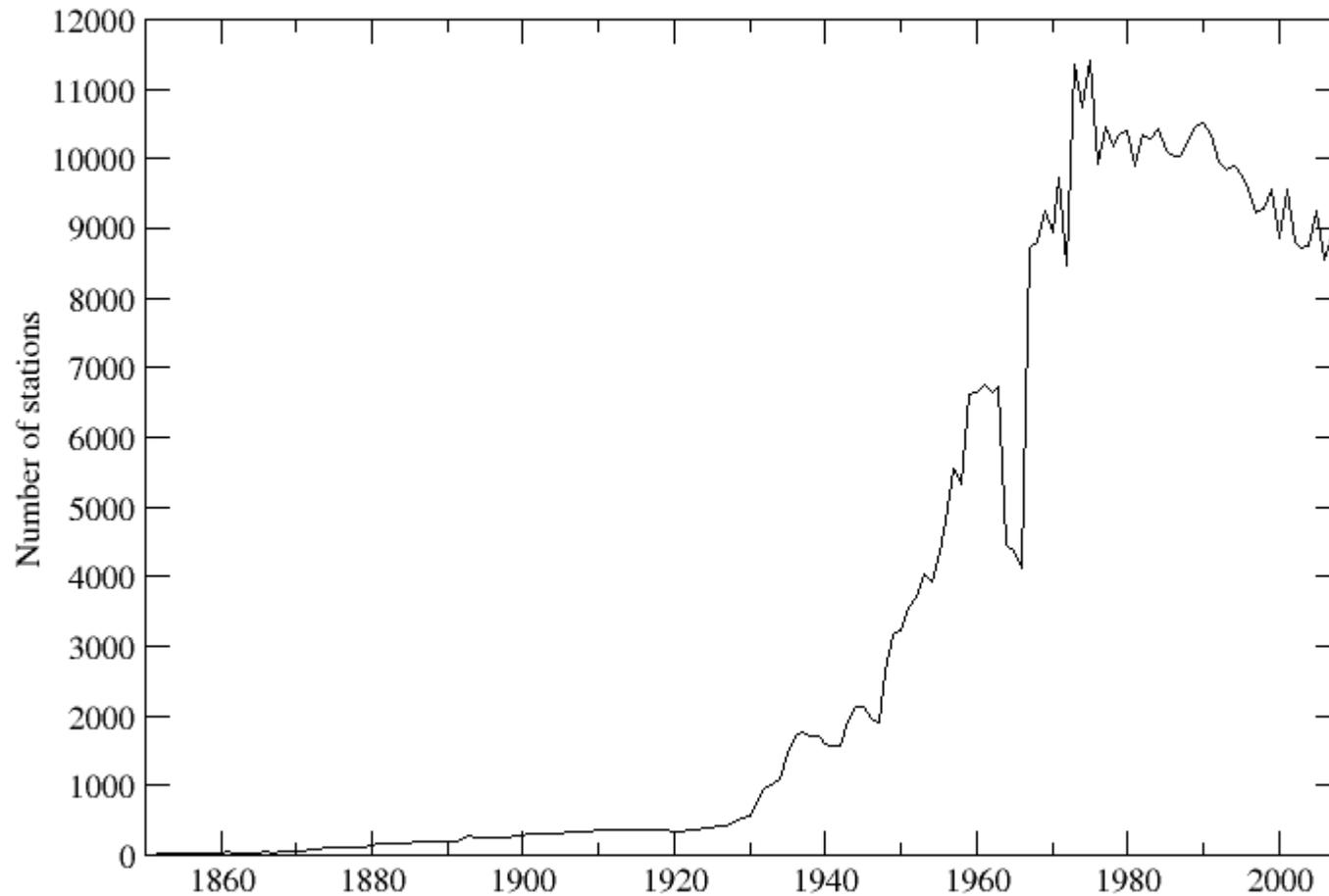
- Number of stations: 33,628
- Number of pressure observations: 1,751,291,205
- Period of record: 1768-2008



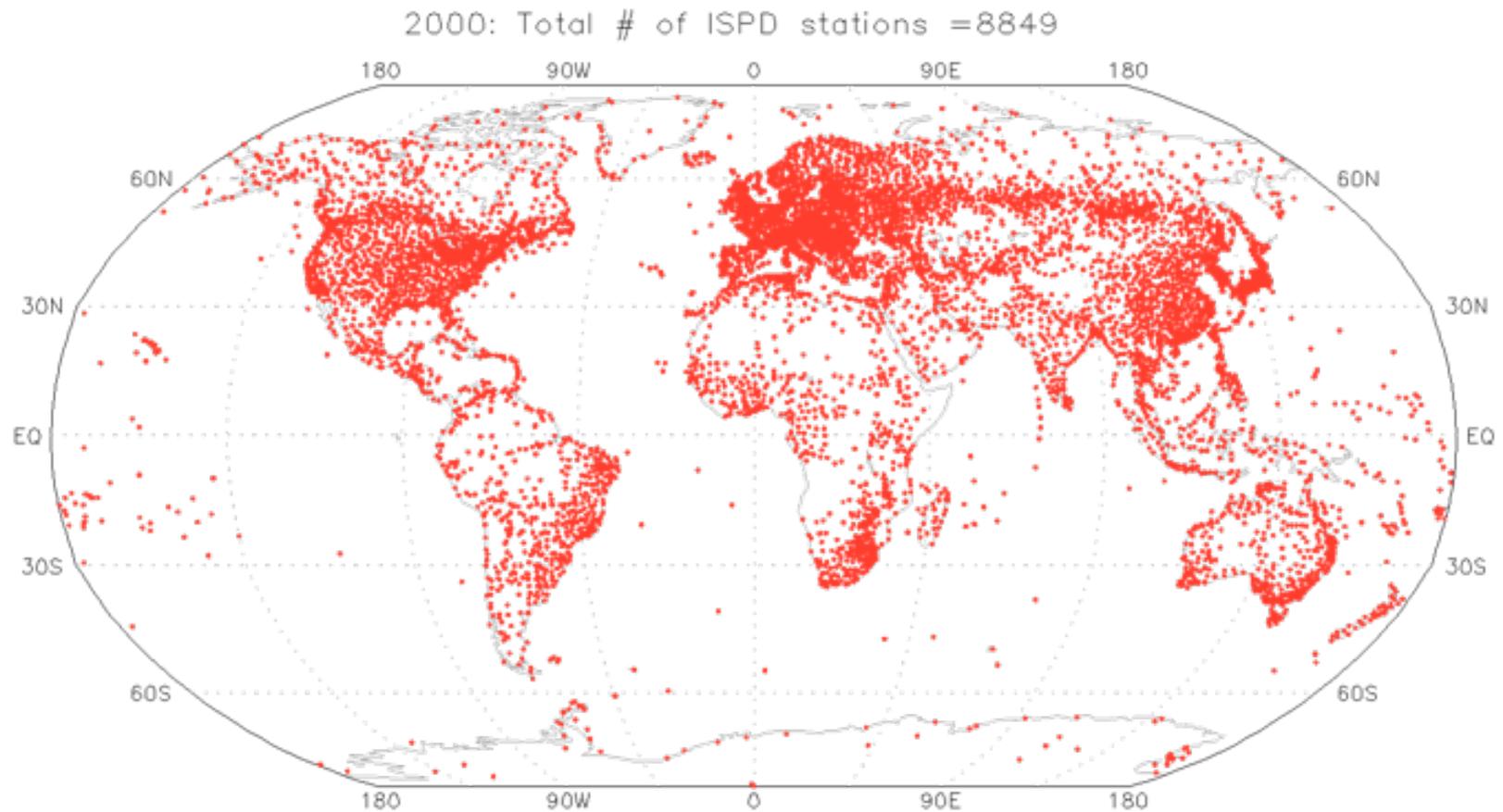
NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

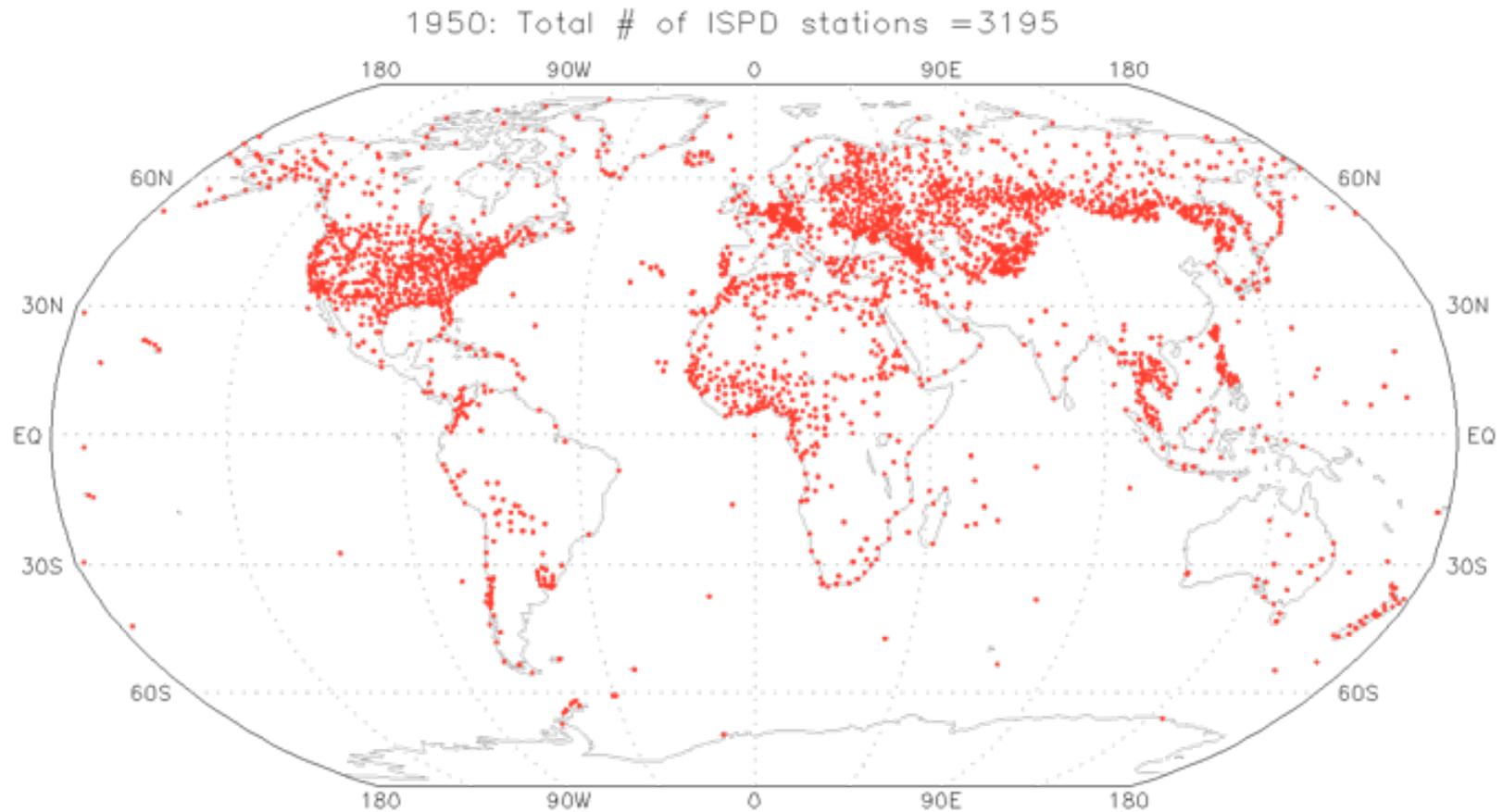
ISPD Network Through Time



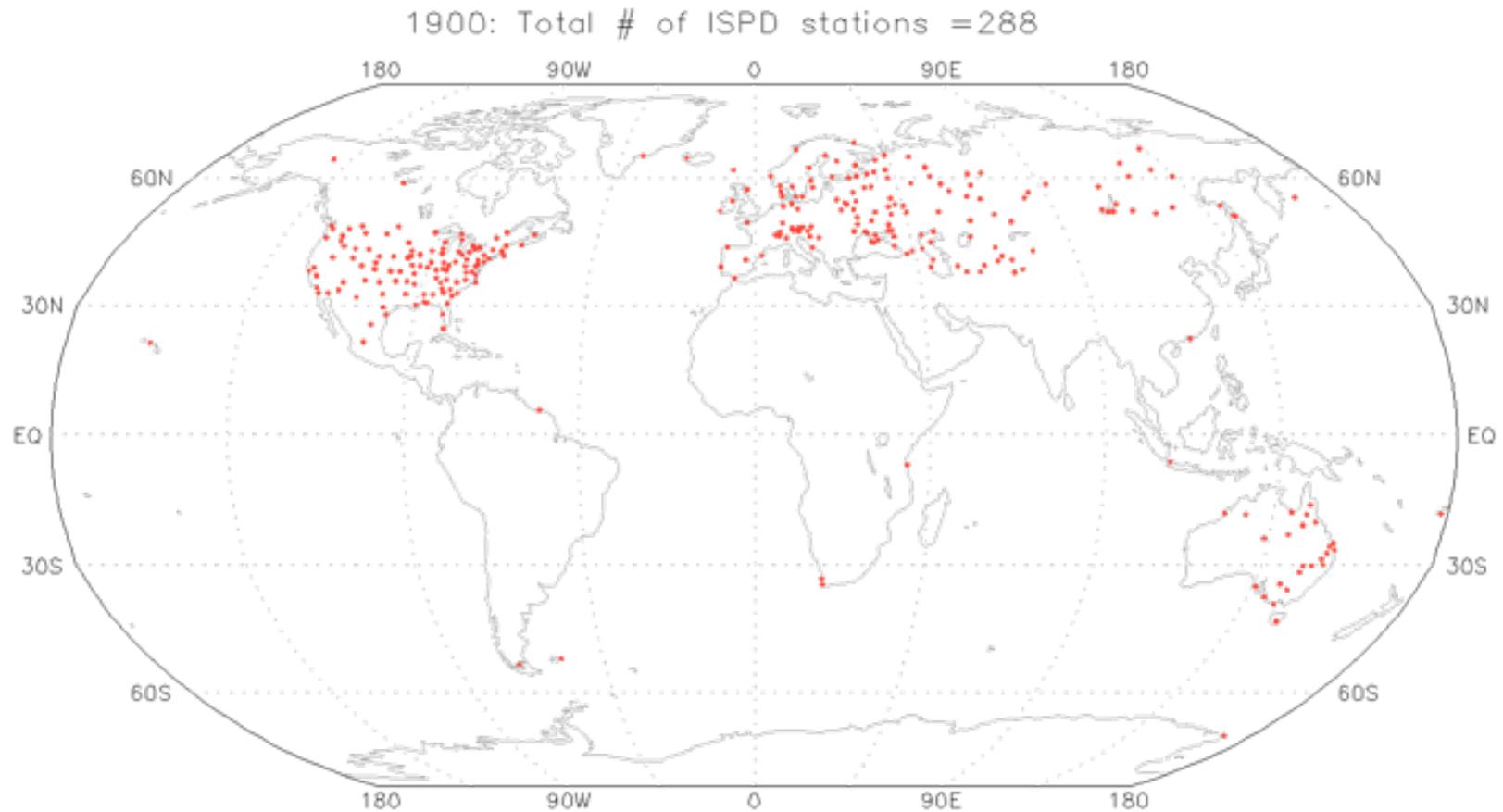
ISPD Stations in 2000



ISPD Stations in 1950



ISPD Stations in 1900



New Data Sets for ISPD v3+

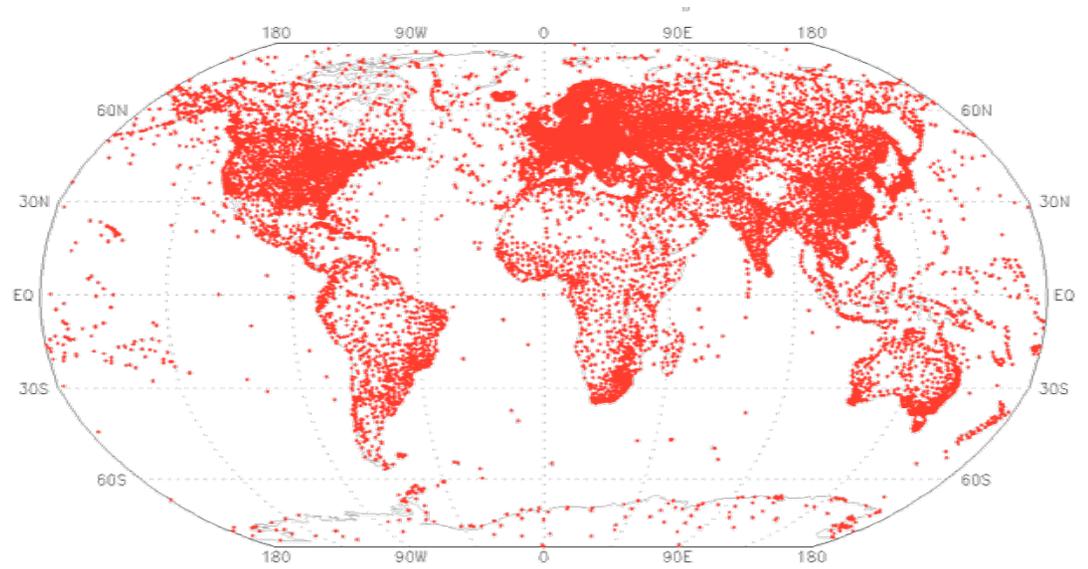
- CDMP data (Mexico, Uruguay, U.S., Vietnam)
- GLOBE student data
- Hadley Centre collections
- KNMI station collection
- MeteoFrance National Data Rescue Activity
- Seiki University collection for Asia
- And many others - all coordinated by ACRE!



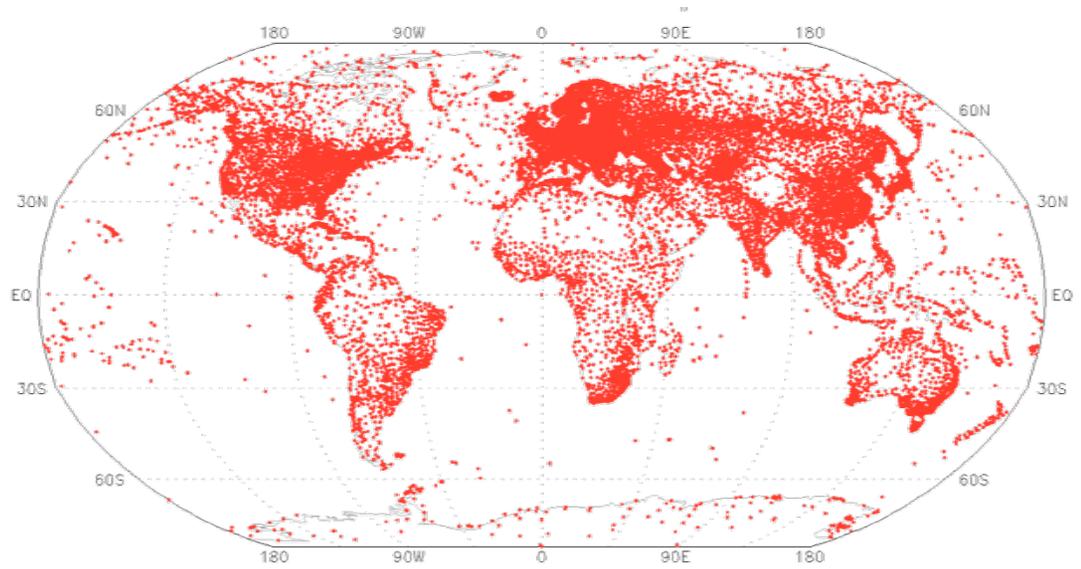
NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

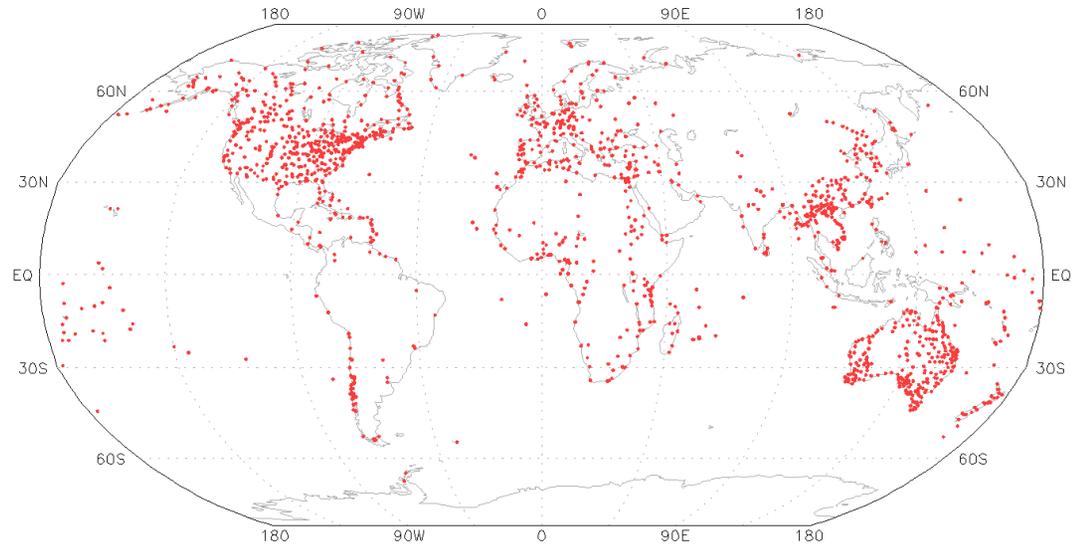
ISPD 2



ISPD 2

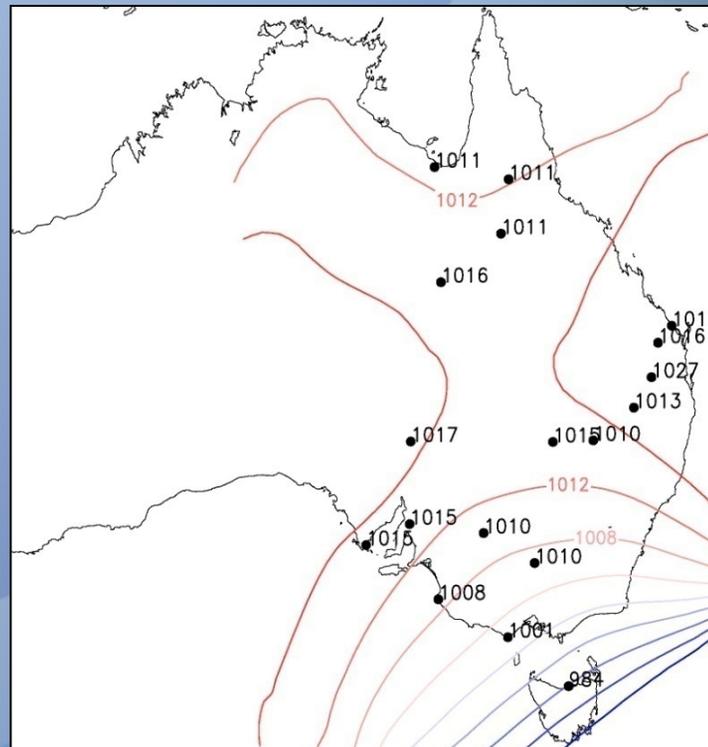


ISPD 3+

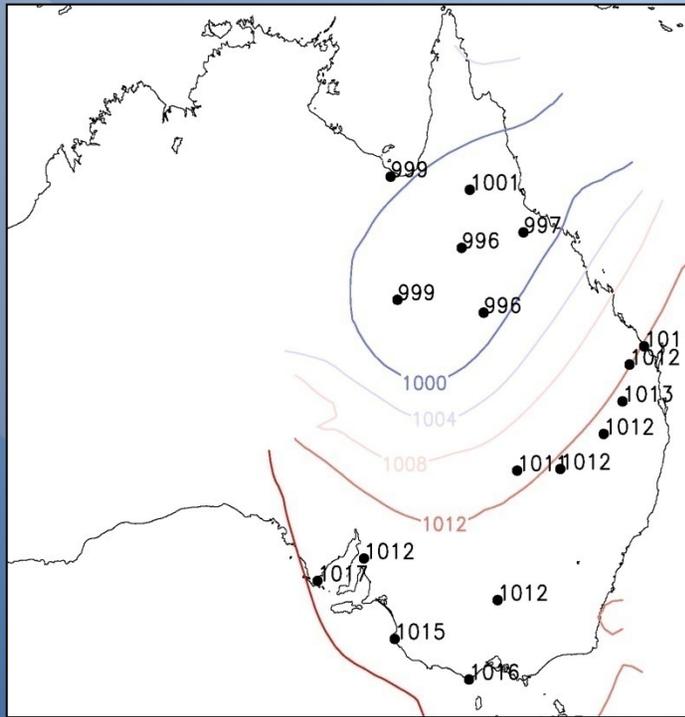


Quality Assurance

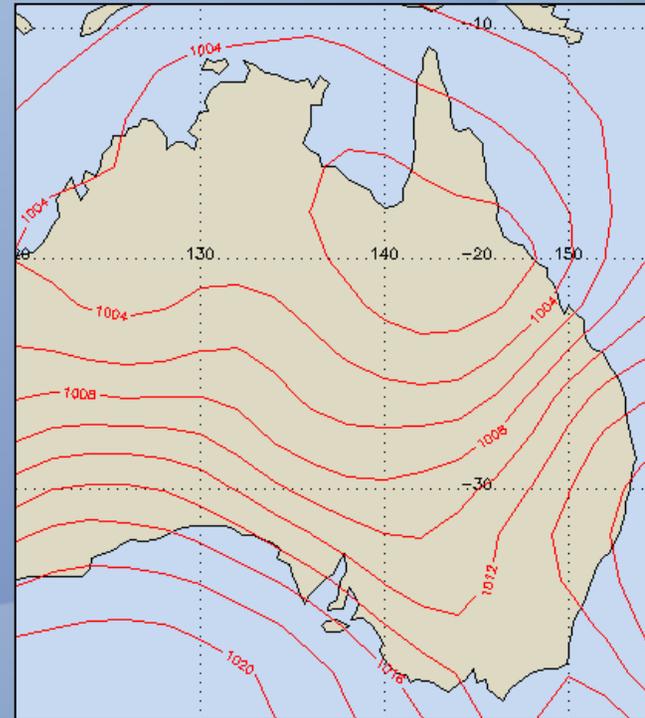
- **ISPD v2.2.4**
 - Removal of duplicates
 - Format compliance
- **ISPD v3**
 - Some checks from (Durre et al., 2010)
 - Develop new checks for known ISPD problems
 - C2OR feedback and comparisons



C20R Comparisons



~0500 UTC Jan. 22, 1918
(ISPD)



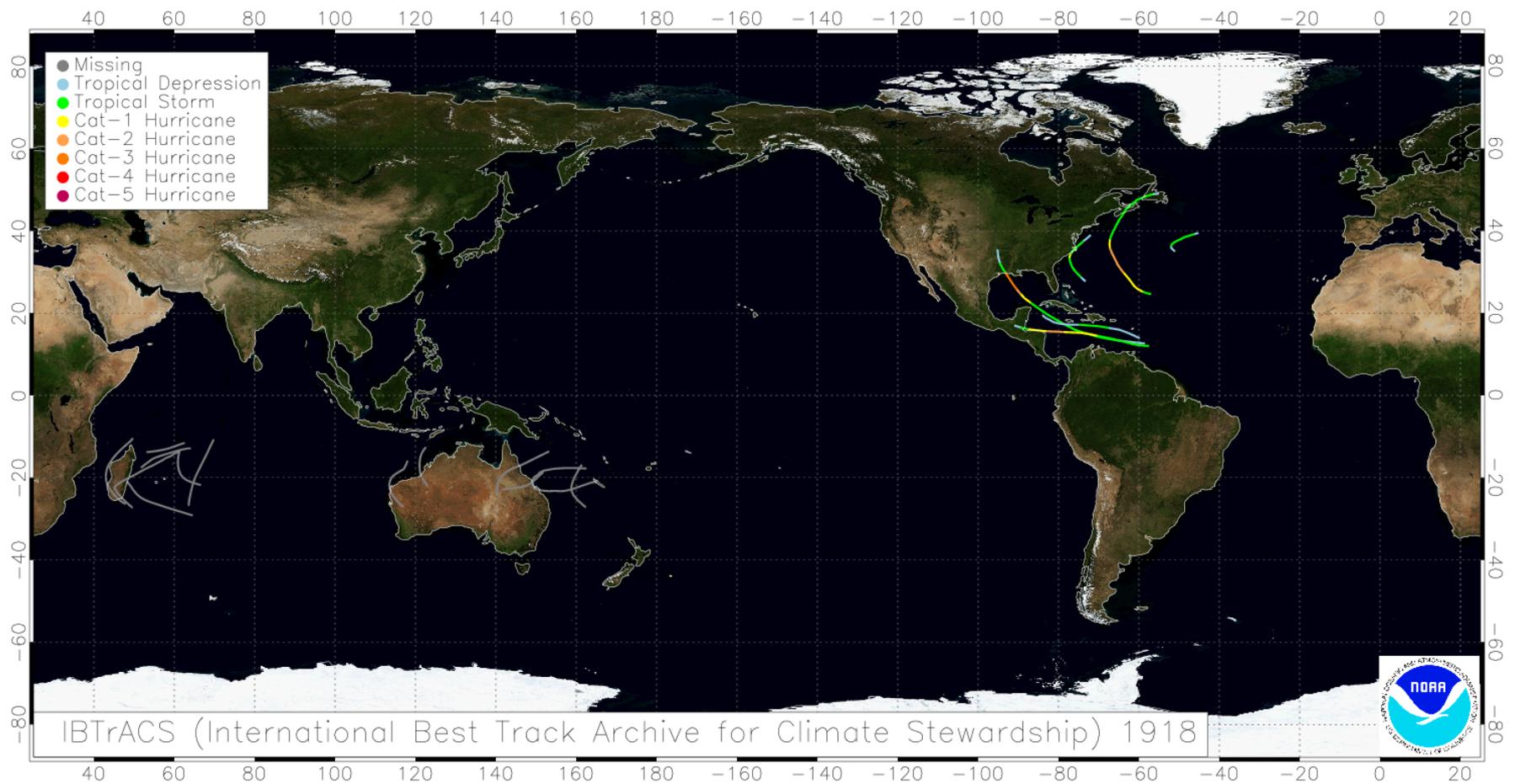
~0000 UTC Jan. 22, 1918
(C2OR)



NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

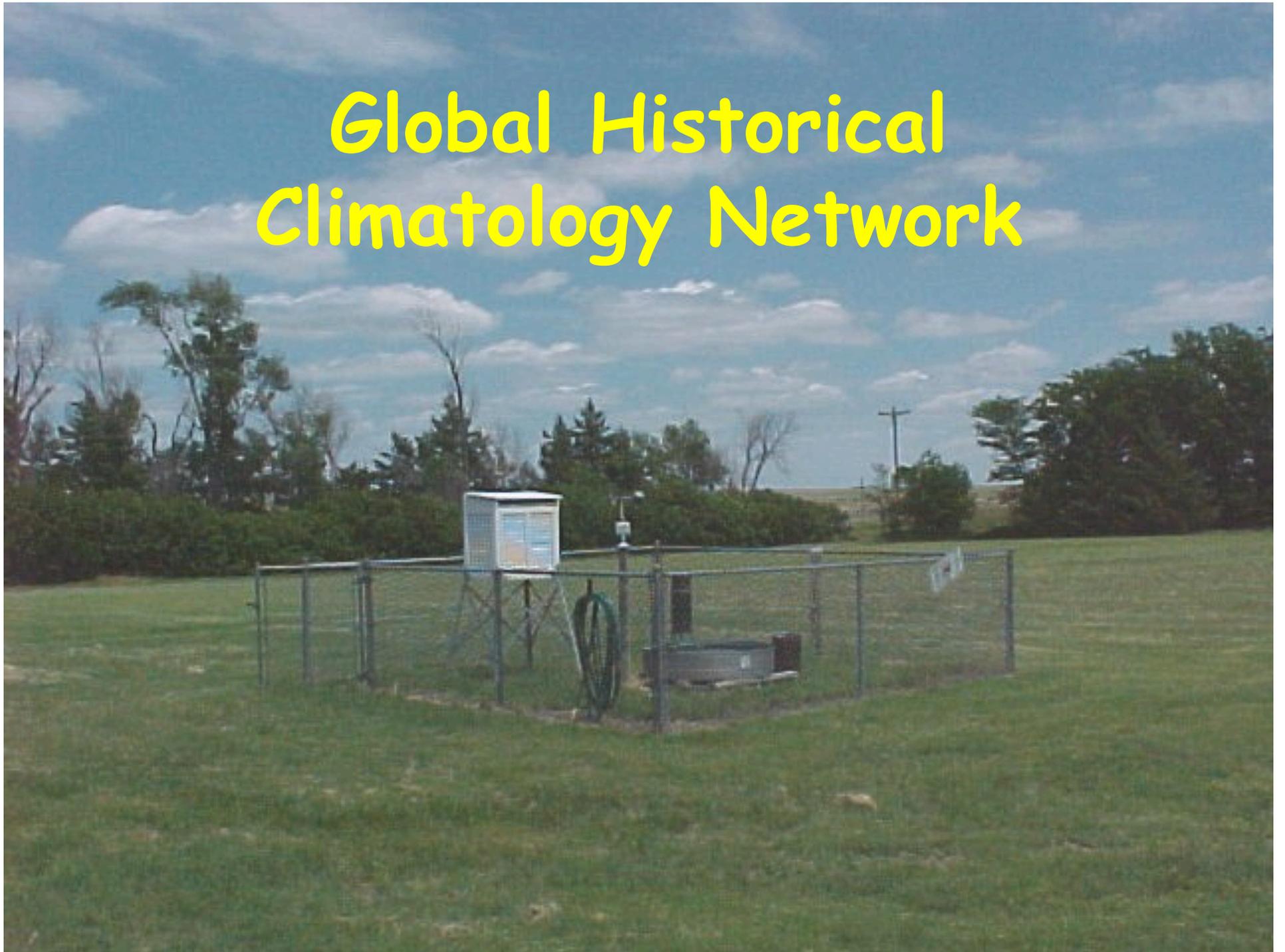
Tropical Cyclone Tracks



NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

Global Historical Climatology Network



Also Associated with GHCN

Growing Healthy Churches Network

Does "homogenized" mean cooked?

NOAA/GHCN are crooks, too.

The dog ate my data.

Feed your dog the food he needs and wants.



NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

GHCN Vital Statistics

Variables:	Temperature and precipitation (et al.)
Stations:	~22,000 (T) and 75,000 (P)
Coverage:	Global land areas
Resolution:	Daily (length of record varies)
Updates:	Daily (via multiple data feeds)
Uses:	Research, monitoring, assessments
Access:	NCDC web site
Players:	M. Menne, I. Durre, R. Vose



NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

Acquisition and Integration

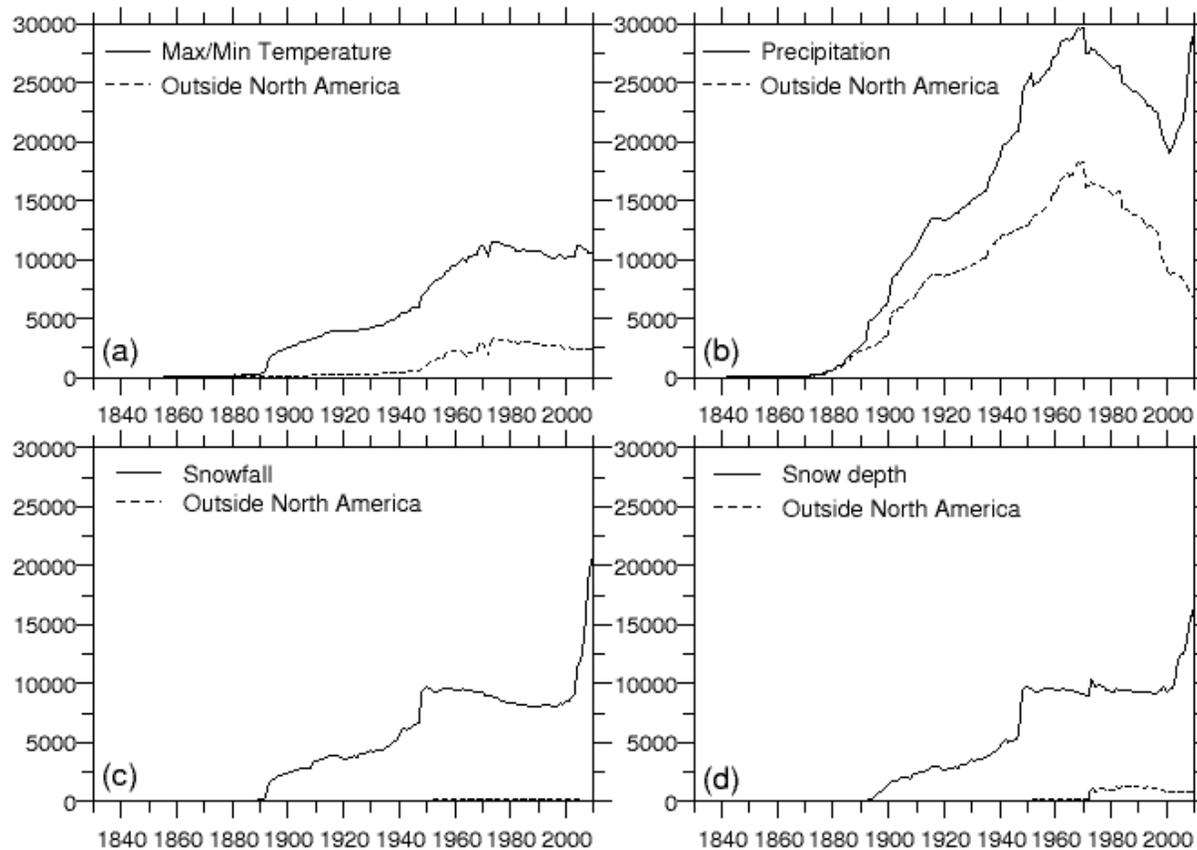
- **No central repository**
 - GTS helps
 - But things are scattered around quite a bit
- **Maximize spatial coverage**
 - Exploit NMC contacts and tap into GCOS
 - Use synoptic reports if no other alternative
- **20 major source datasets (and many others)**
 - International collection (personal contacts)
 - U.S. collection (NCDC archives)
 - Government exchange data (GCOS and/or bilaterals)
 - Global Summary of the Day (synoptic data)



NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

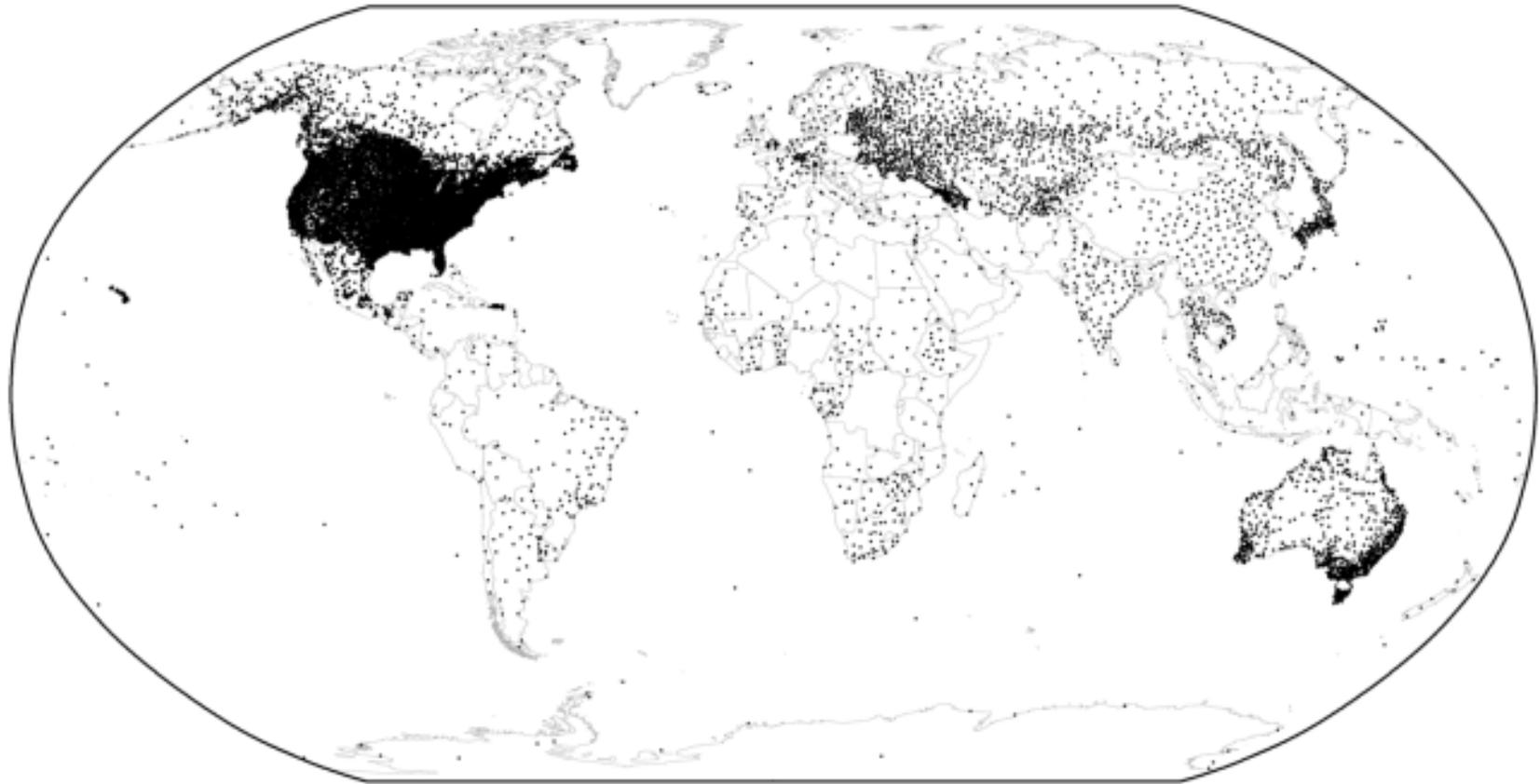
Network through Time



Max/Min Stations: 1991-2010



Max/Min Stations: 1961-1990



Max/Min Stations: 1931-1960



Max/Min Stations: 1901-1930



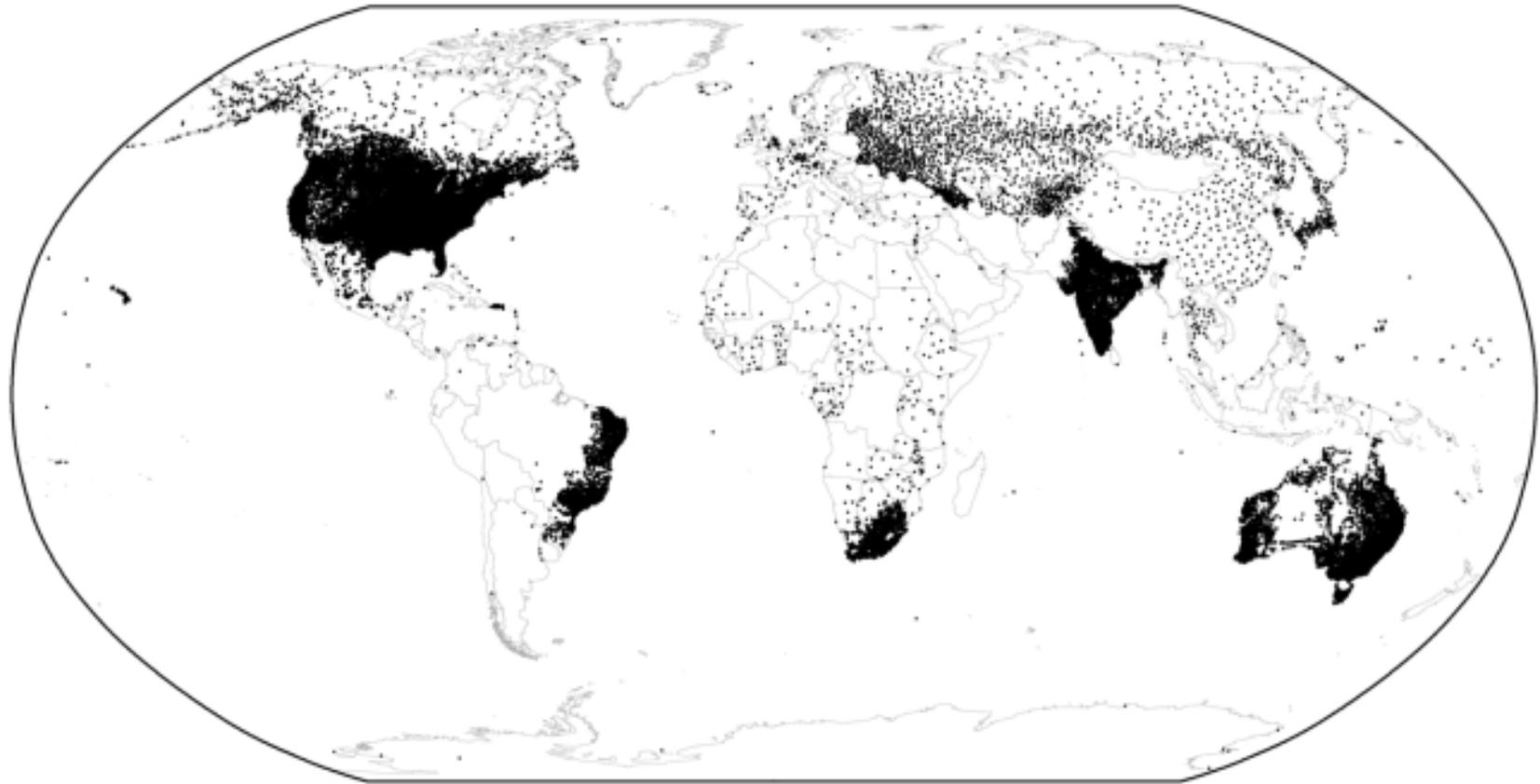
Prcp. Stations: 1991-2010



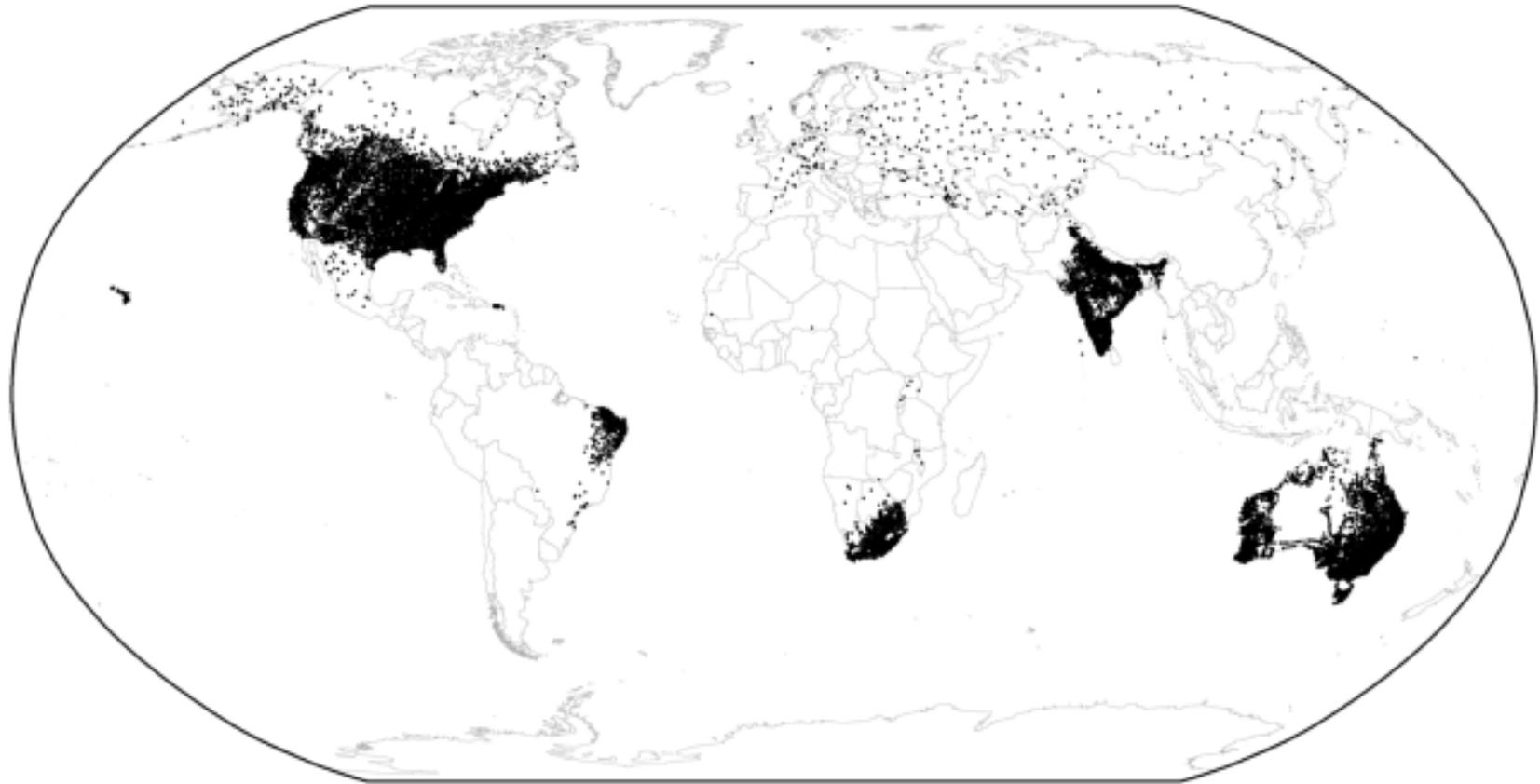
Prcp. Stations: 1961-1990



Prcp. Stations: 1931-1960



Prcp. Stations: 1901-1930



Quality Assurance

- **For more information**
 - Evaluation approach: Durre et al. (2008) *JAMC*
 - List of QA reviews: Durre et al. (2010) *JAMC*
- **General principles**
 - Fully automated approach that runs operationally
 - Occasional application of source-specific integrity checks
- **Major features**
 - 19 separate tests (excessive duplication, exceedance of physical and statistical limits, excessive persistence, large distributional gaps, internal and spatial inconsistencies)
 - Roughly 0.3% of values get flagged, 1-2% false positives



Updates and Reprocessing

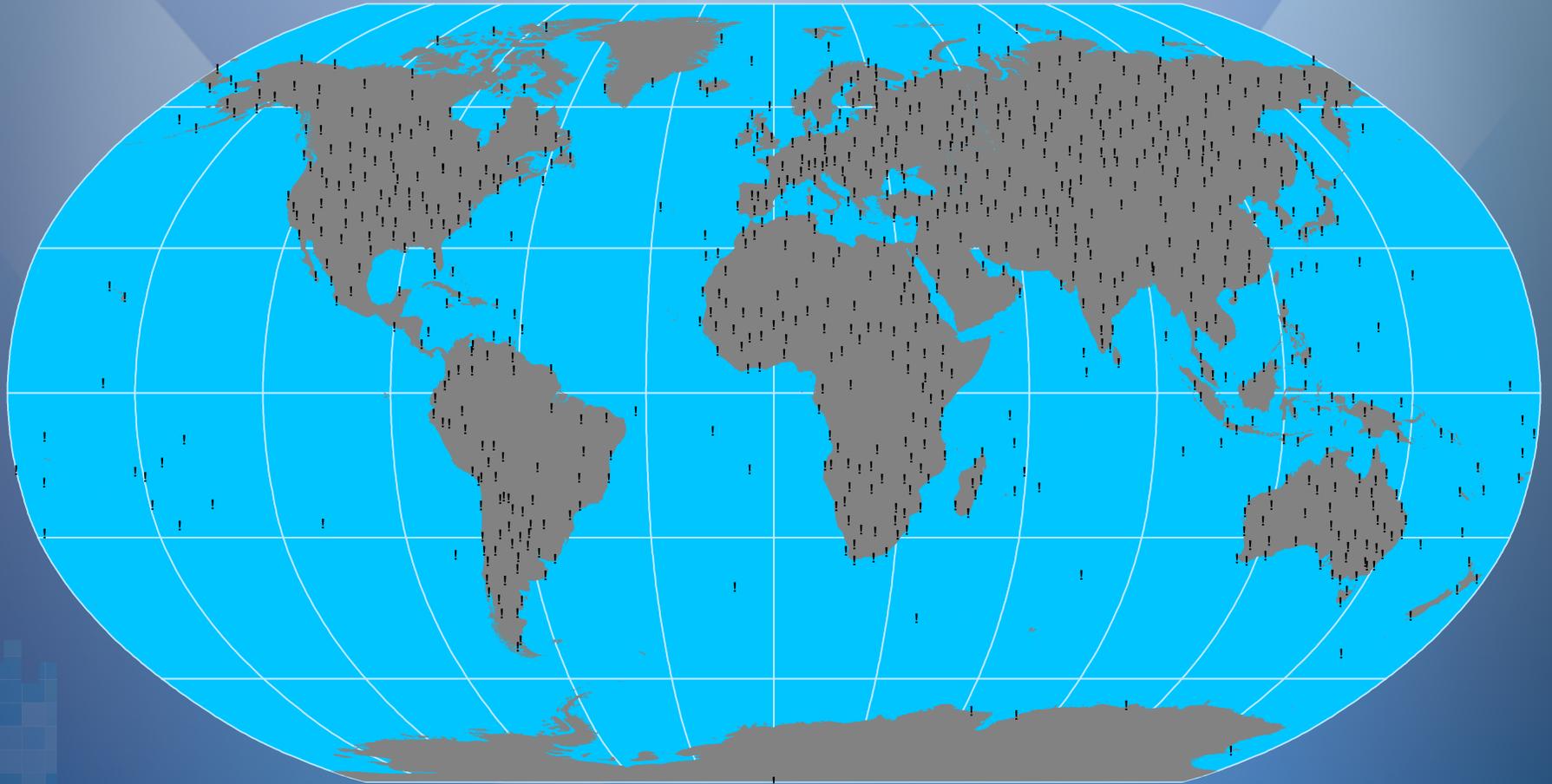
- Updated twice per day
- Reprocessed once per week from scratch
- Each updated and reprocessed version is archived at NCDC (e.g., POR2010100813)



NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

GHCN-Daily: The GSN Archive



NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

World Extreme Weather Archive

At a task meeting of the WMO OPAG 2 group, members unanimously agreed that the creation of a world archive for verifying, certifying and storing world weather extremes was useful and necessary. GHCN-Daily served as a baseline in the initial set of world records assembled by the WMO rapporteur for climate extremes.



NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010

Biased Take-Away Messages

- These are the biggest datasets of their kind
- They're also the best



NOAA's National Climatic Data Center

3rd ACRE Workshop
Baltimore, MD, November 3-5, 2010