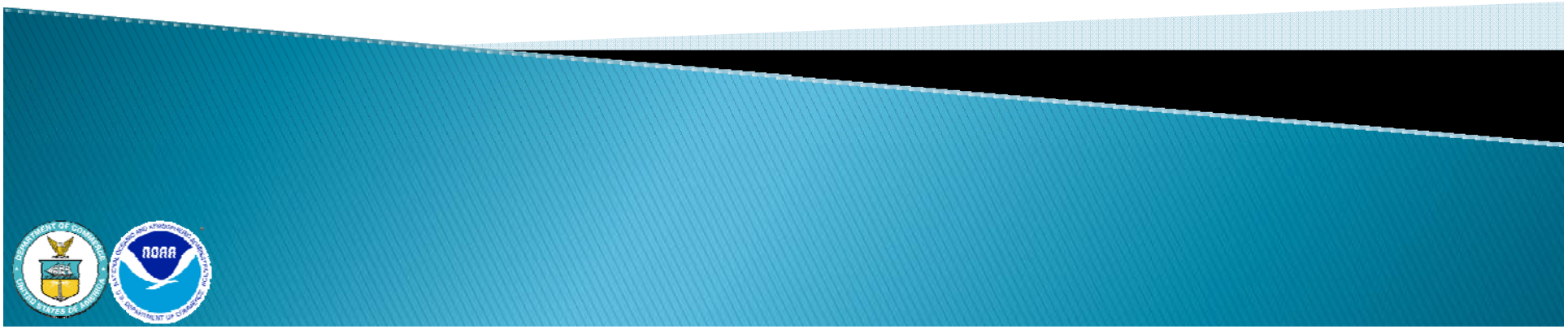
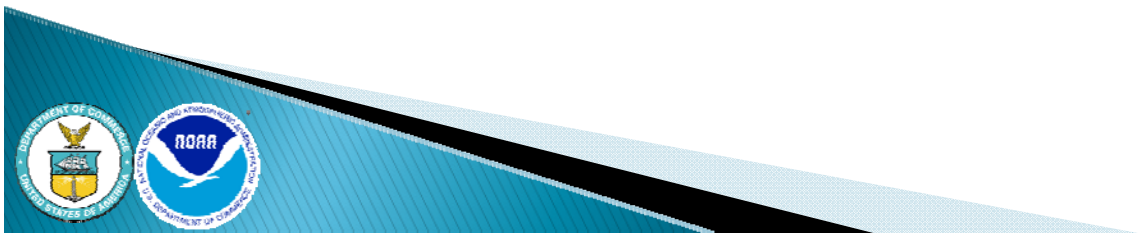
REPORT FROM THE GCOS ARCHIVE CENTRE

Jay Lawrimore
NOAA National Climatic Data Center
9 October 2013

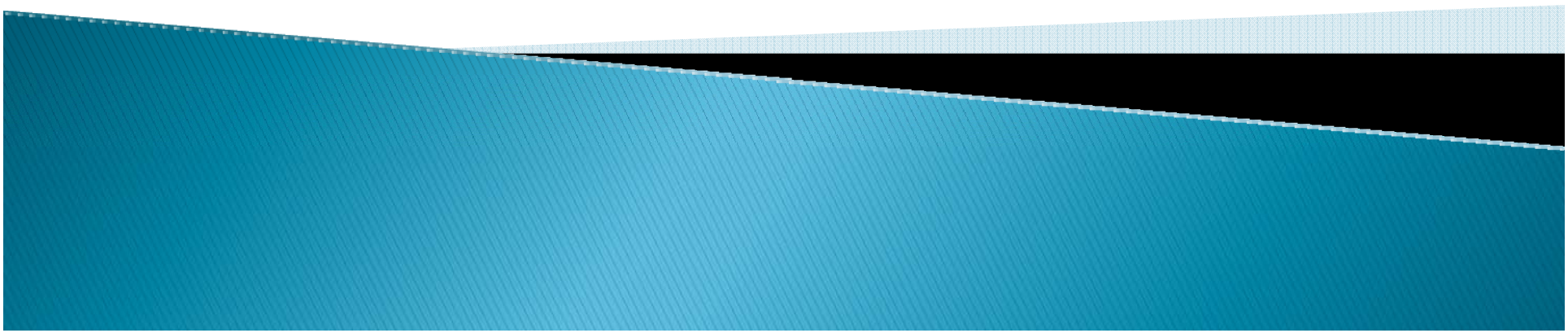


Overview

- Daily data
  - GHCN-Daily as the GSN Archive
- Monthly data
  - GHCN-Monthly and CLIMAT messages
- International Surface Temperature Initiative
  - Global Databank
- Need for Improved Metadata

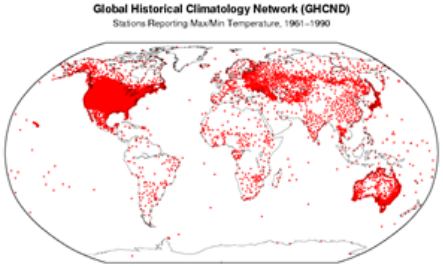


Daily Data



GHCN-Daily

The GSN Daily database



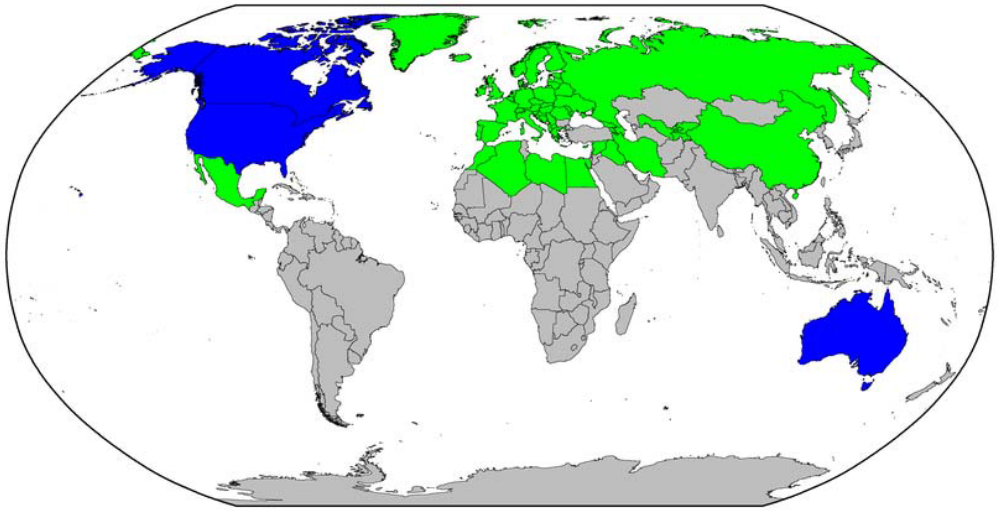
- The GSN daily database is managed as a subset of the Global Historical Climatology Network-Daily (GHCN-D) data set
  - http://www.ncdc.noaa.gov/oa/climate/ghcn-daily
- Global daily *in situ* dataset derived from multiple sources
- >25,000 temperature stns
- >80,000 precipitation stns
- ~25,000 snowfall/snowdepth

- More than 2 billion daily observations
  - Earliest value from January 2, 1833
  - Latest value from yesterday

- For ease of access, a separate directory of GSN station files is maintained at ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily/gsn

Updates to GHCN-D

Approximate distribution of areas with near-real-time or periodic updates in GHCN-Daily



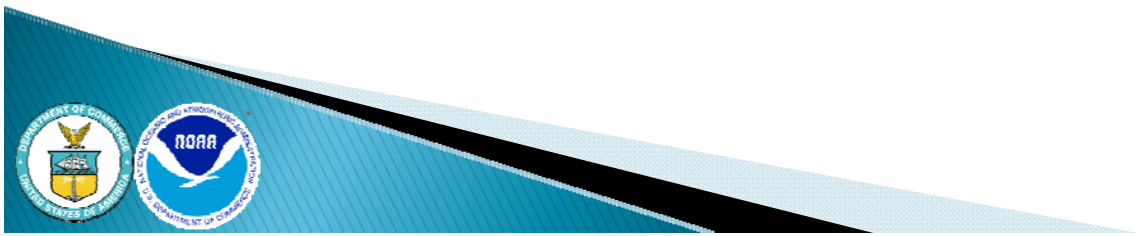
- Approximately 850 out of the 1019 GSN stations have daily data of some kind in the GHCN-Daily/GSN archive.
- A small number of the GSN sites are updated regularly through arrangements made by the GCOS Lead Centres (Iran, Estonia, Uzbekistan, Estonia, Cyprus) while other National Meteorological Services provide occasional updates via email.
- A growing number of stations are updated through bilateral agreements between the Lead Centre at NCDC (USA, Australia, Canada, Russia).

Ongoing Global Additions

- Within the past 2 years a mechanism to establish wider ongoing updates has been initiated with the European Climate Assessment activity led by KNMI
  - This program provides monthly updates of daily observations for stations throughout Europe, parts of North Africa, and the Middle East
- The NMS in Uzbekistan consistently provides data within the first two to five days each month
- The NMS of Cyprus provides data for Larnaca on a periodic basis via e-mail
- Data from the U.S., Canada and Australia are received via various mechanisms including ftp and web services.

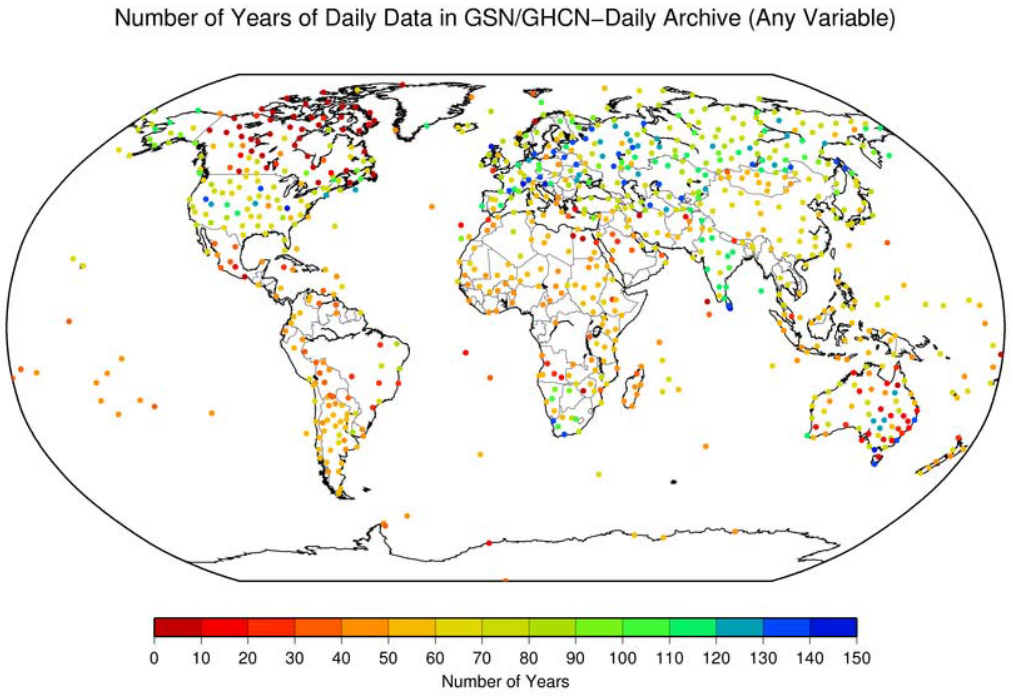
Ongoing Global Additions

- NCDC prefers to establish regular data updates of full period of record datasets through ftp or web services to ensure the most consistent and complete historical and current data
- Several countries have provided complete copies of their daily climate databases for inclusion in GHCN-D
- Updates for last month are generally available by the 15th of the current month



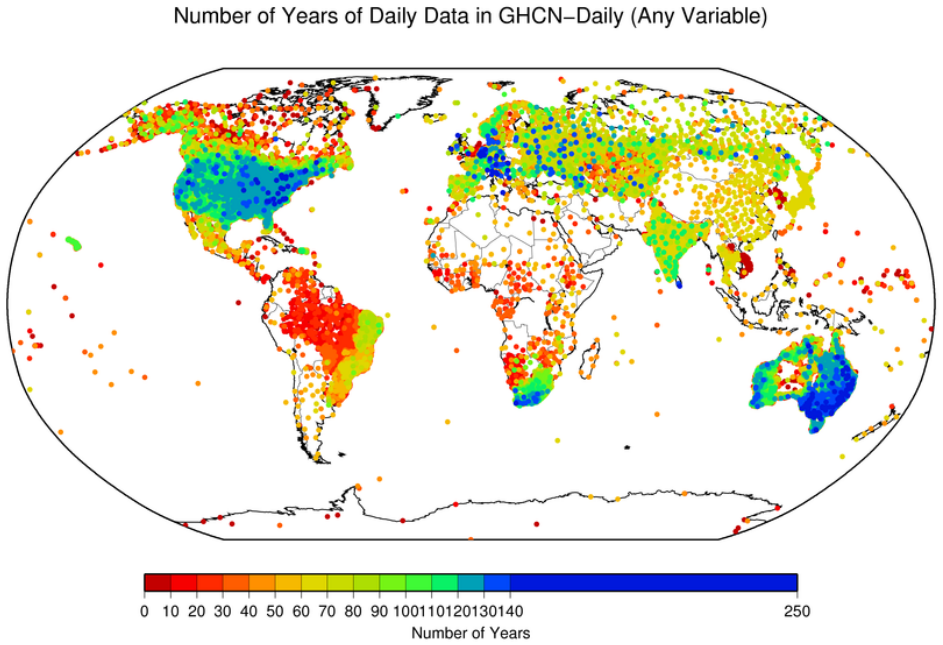
GSN Lengths of Record

- This figure shows the length of daily data records in the GHCN-Daily database for GSN stations as of March 2013
- Some GSN stations represent relatively newly established locations

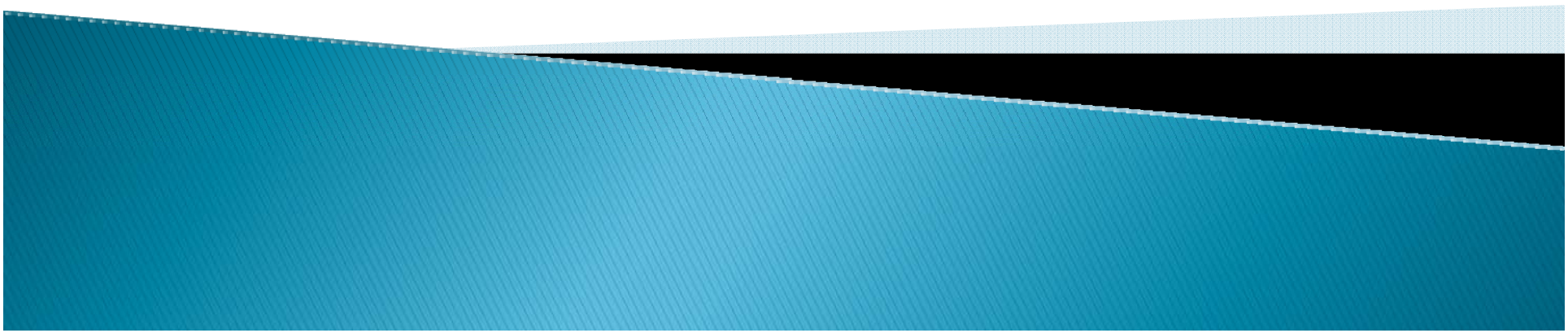


All GHCN-Daily Stations

- The number of stations with reasonably long periods of records available in GHCN-Daily is much larger than for the GSN subset of stations.

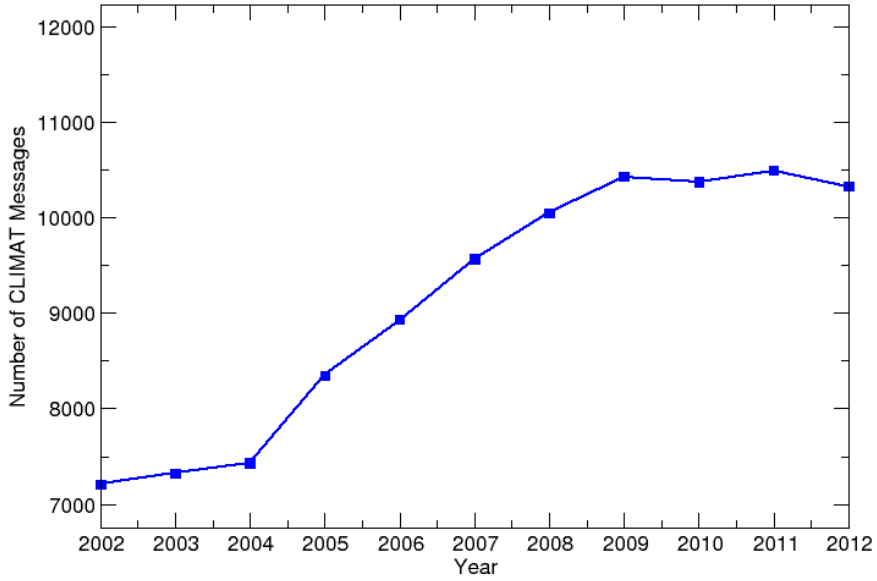


Monthly Mean Surface Data



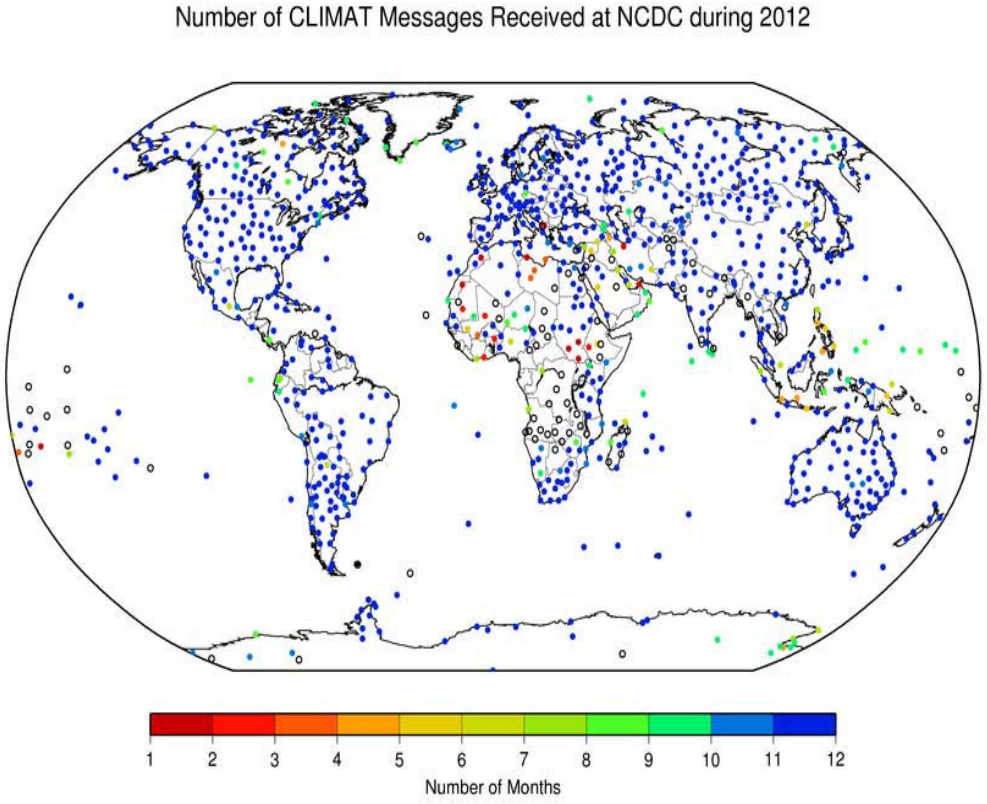
CLIMAT messages received

- The number of CLIMAT messages received at the Archive Centre from GSN sites has remained nearly constant over the last several years
- In addition to CLIMAT messages received over the GTS at NCDC, the UK Met Office provides 100 to 200 reports each month



CLIMAT messages received in 2012

- Most CLIMAT stations provide data all 12 months of the year

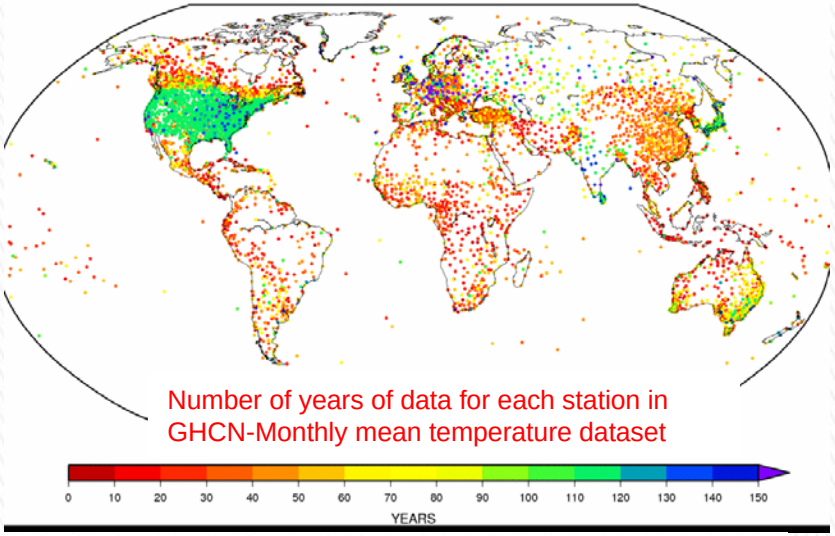


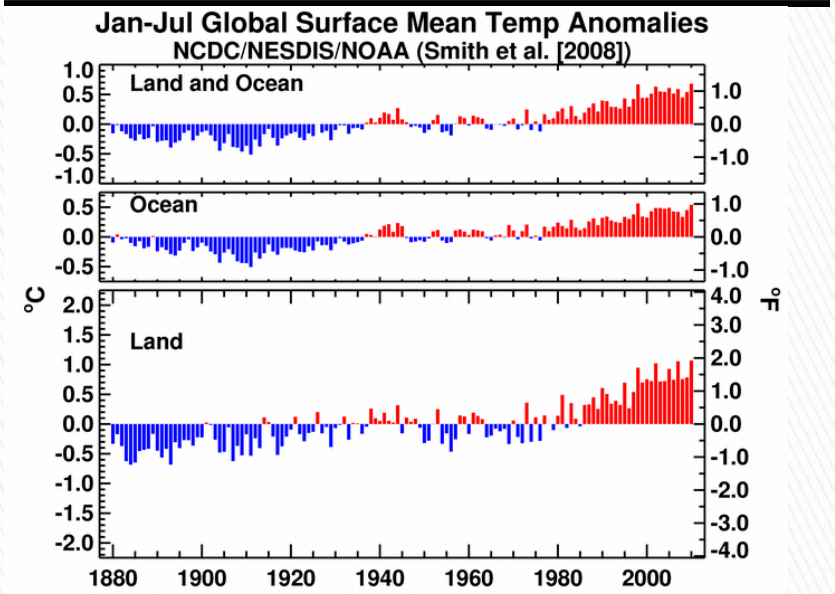
Monthly CLIMAT messages

- Monthly CLIMAT messages serve as the basis for operational climate monitoring activities and are used to update NCDC's GHCN-Monthly dataset
- The latest version of GHCN-M (version 3.2.0) was completed and released in 2011
  - Lawrimore, J.H., M.J. Menne, B.E. Gleason, C.N. Williams, D.B. Wuertz, R.S. Vose, and J. Rennie (2011). An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3, *J. Geophys. Res*.

GHCN-M Land Surface Stations

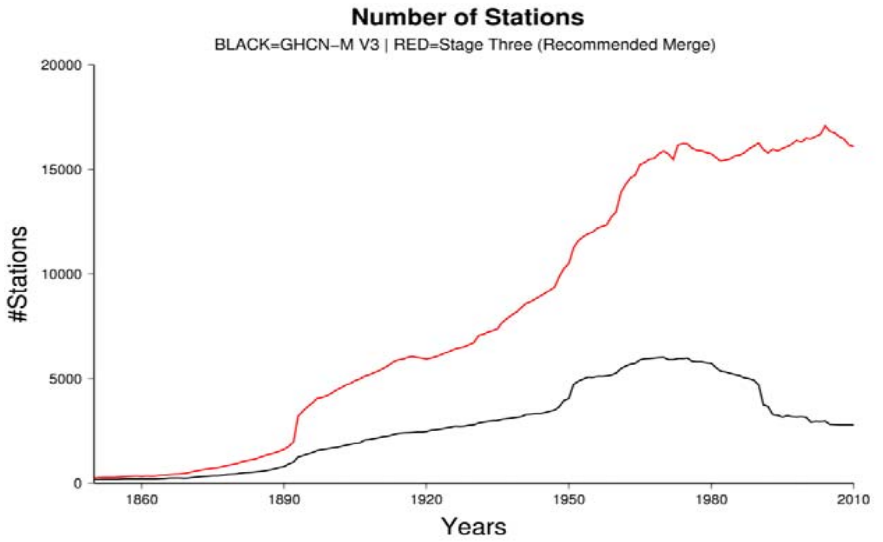
- NCDC monitors mean changes in global land surface temperature using the **Global Historical Climatology Network (GHCN)** dataset
  - Composed of **7,280 stations** that record monthly mean temperatures
- In use since the early 1990s. GHCN-M Version 3 was released in May 2011

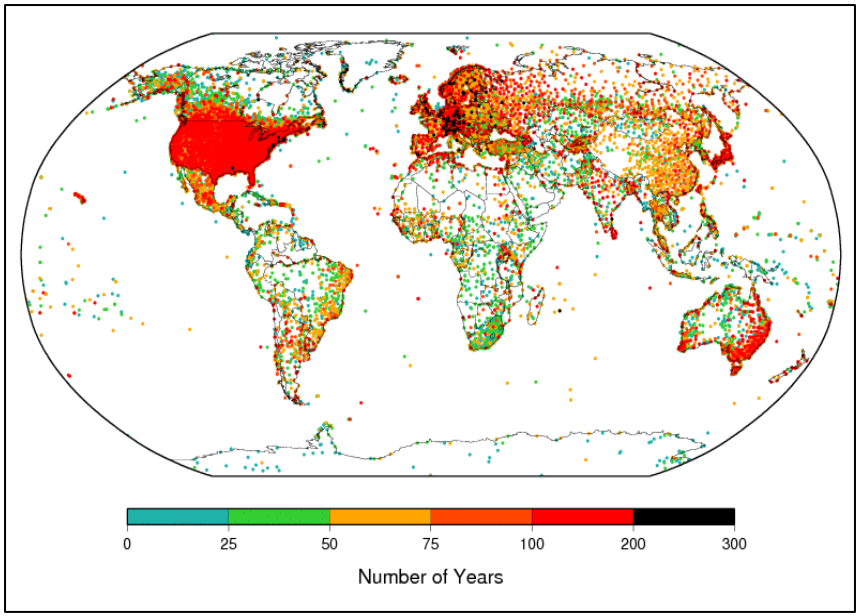




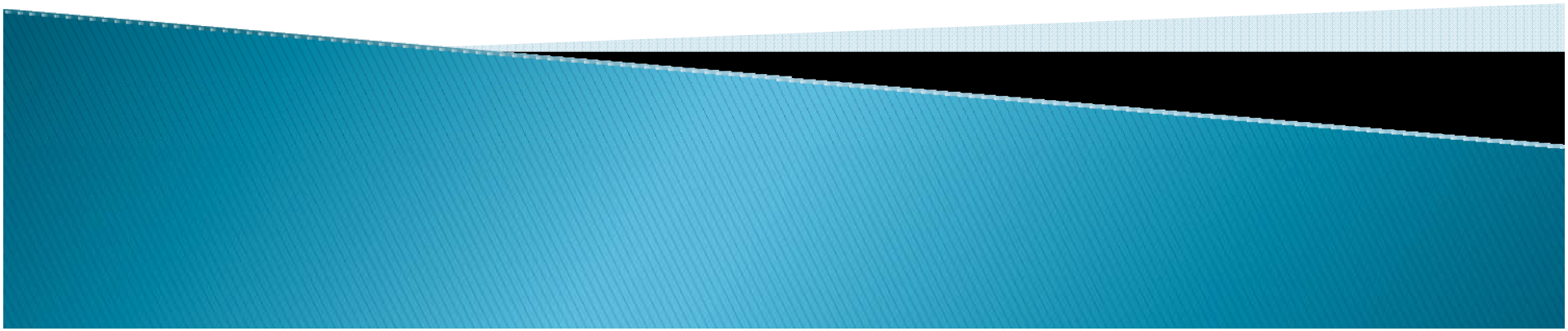
GHCN-Monthly version 4

- A new version of GHCN-Monthly (version 4) will be released in late 2014
- It will contain approximately 4 times as many stations as version 3
- This new version will ensure consistency between GHCN-Monthly and GHCN-Daily datasets



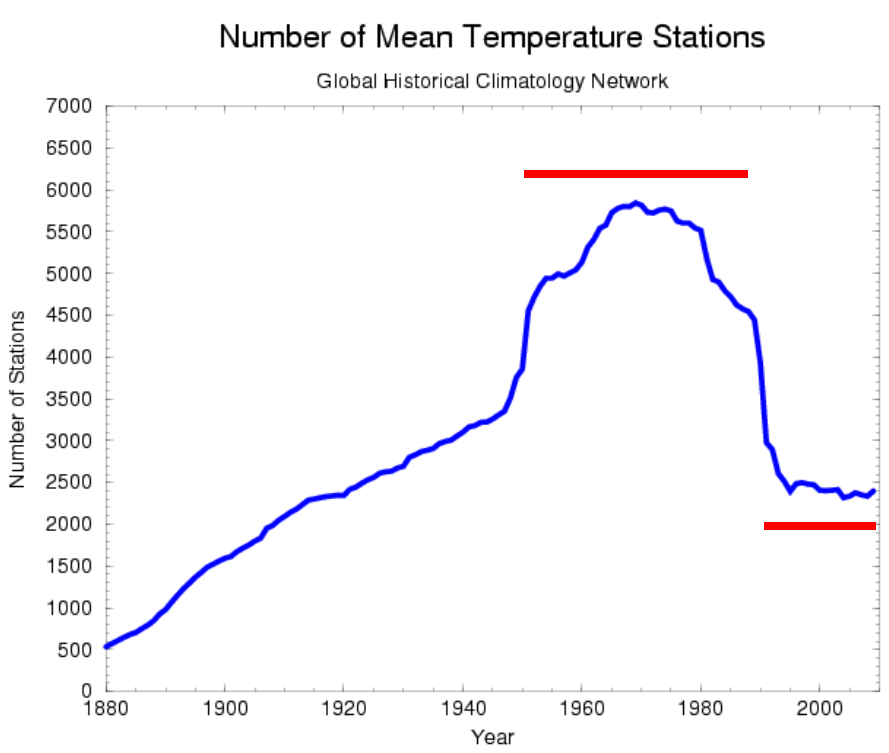


International Surface Temperature Initiative



Building and maintaining datasets remains a high priority

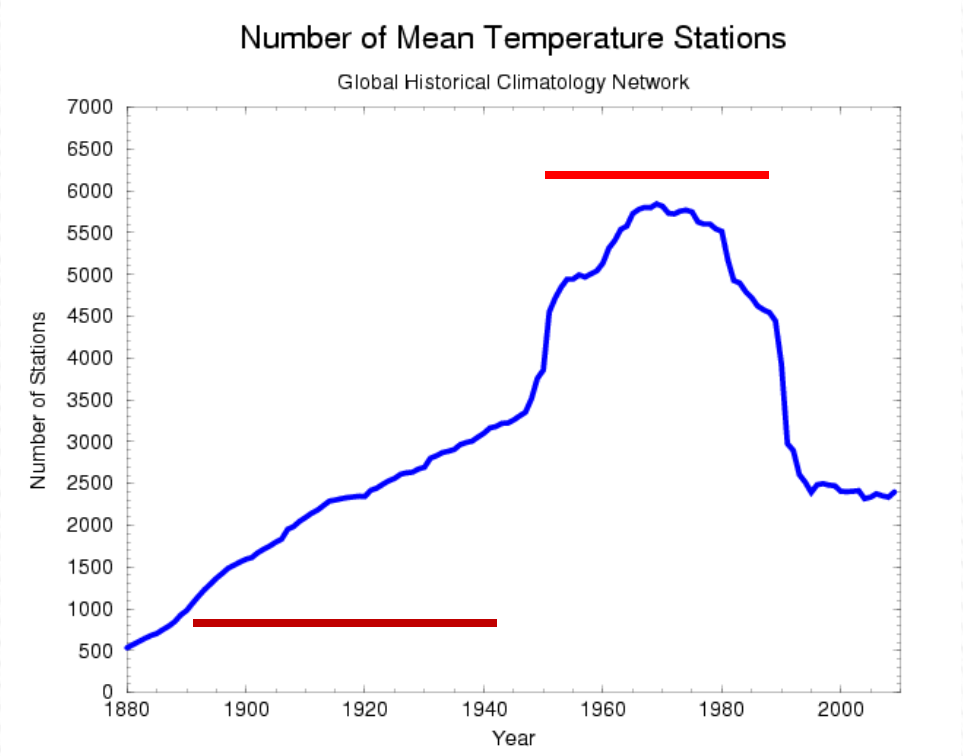
- Operational Ingest and Acquisition alone provides only a portion of all data
- Without continued attention to acquire and incorporate data from other sources, the number of stations in the datasets suffer decline
  - Many stations do not report in near real-time
  - Addition of other stations requires years of effort to locate, collect, process, and incorporate



Number of stations in the GHCN-Monthly mean temperature dataset from 1880 through 2009.

Building and maintaining datasets remains a high priority

- There are also major gaps in data before the 1950s
  - Daily datasets such as GHCN-Daily can help fill in the most recent decades
  - But do little to help before 1950
- Data rescue is essential to building up the dataset before 1950
  - Millions of images preserved in the past 10 years but little effort to digitize from the images



Number of stations in the GHCN-Monthly mean temperature dataset from 1880 through 2009.

International Surface Temperature Initiative

- To address these and other issues associated with global dataset development, transparency, quality control, and benchmarking
- A new international effort initiated in September 2010 and led by NCDC, the UK Met Office, and other international partners

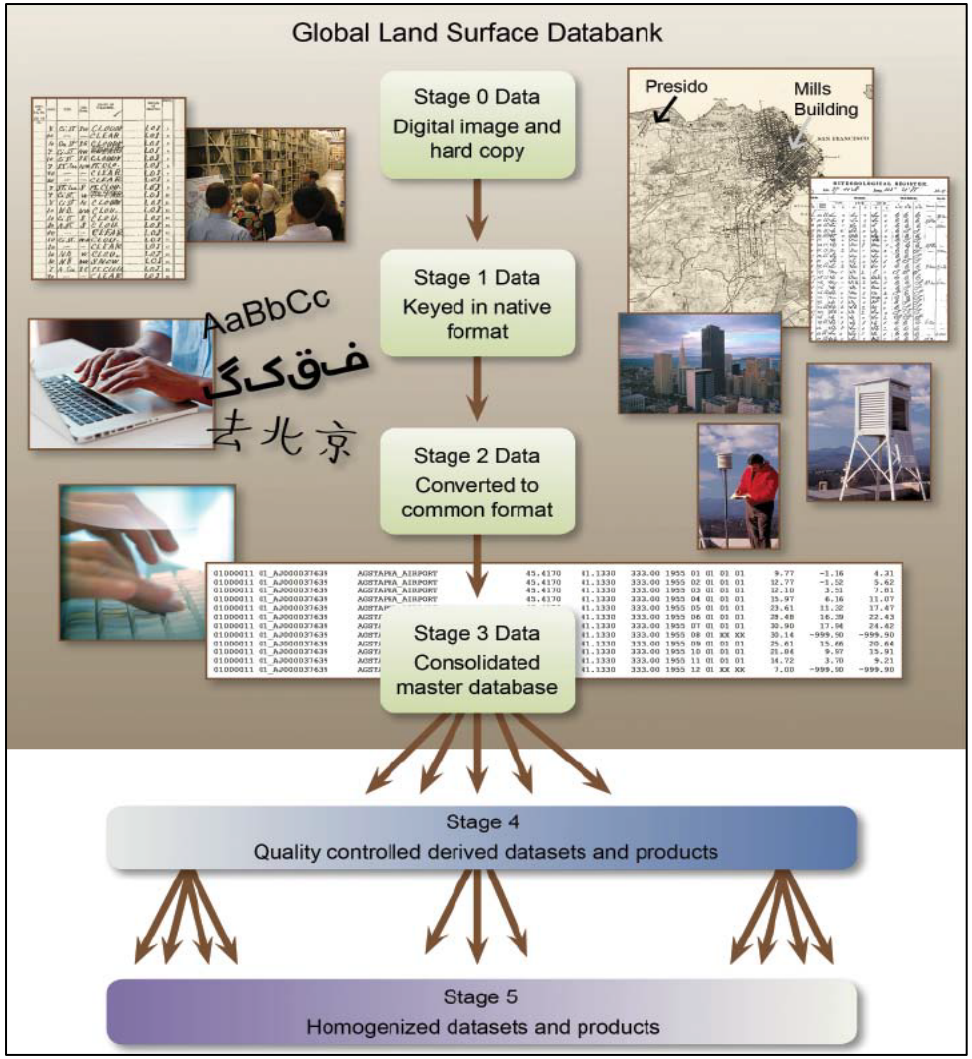
International Surface Temperature Initiative

- Has the potential to greatly benefit efforts to improve global collections of surface observations
  - Initially focusing solely on temperature observations
  - Objective is to develop a comprehensive global repository of land surface observations that are spatially and temporally complete, fully traceable from the point of observation, freely available, easily accessible, and with quantified uncertainties.

Global Land Surface Databank

- Of initial priority is the establishment of a Databank of land surface temperature existing in various “Stages”
  - Stage 0: paper record
  - Stage 1: Native format
  - Stage 2: Common format
  - Stage 3: Merged dataset
- There also will be quality controlled (Stage 4) and bias corrected (Stage 5) data.
- Lawrimore, J., J. Rennie, and P. Thorne (2013), Responding to the Need for Better Global Temperature Data, Eos Trans. AGU, 94(6), 61.

Global Land Surface Databank



Global Land Surface Databank

- Contributions to the Databank include new data sources from Brazil, Uruguay, the Netherlands, Japan, Spain, Antarctica, and many other countries
- Data provided by NMSs in their native format are considered Stage 1
- As part of initial processing they are converted to a common format and have data provenance tracking flags added to each data value
  - These flags provide traceability back to the original source and help guarantee data authenticity, integrity, and quality

Global Land Surface Databank

- A databank of more than 32,000 stations with monthly mean, maximum, and minimum temperature is now available:

- ftp://ftp.ncdc.noaa.gov/pub/data/globaldatabank/

and

- ftp://ftp.meteo.ru/pub/data/globaldatabank/

Metadata Needs

- Metadata collection at NCDC for stations outside U.S. networks has received little attention in comparison to U.S. networks
- Detailed station histories have yet to be fully exchanged internationally
  - Is a standard station history format needed? or more attention needed?
- Most available global metadata at NCDC is outdated
- Little to no information on observing instruments, practices, and surrounding areas
- In the last two decades greater capabilities emerged such as higher density gridded topographical data and mapping capabilities such as Google Maps

Metadata

- Metadata for stations outside the U.S. are contained primarily in ASCII flat files of the Global Historical Climatology Network–Monthly and –Daily datasets and the newly established International Surface Temperature Databank
- Metadata often consists of no more than station location and elevation

Metadata

- The GHCN-Monthly version 2 dataset development process included development of additional metadata using other sources such as
  - Digital Elevation Models
  - Night-Light data from Satellites to determine urban versus rural areas
  - Operational Navigation Charts to assess distance to airports, oceans, lakes, and other topographical features.

Metadata Options

- Some potential options to consider:
  - Photographs of present conditions, coordinates, station contacts, instrument type, etc.
  - General description of surroundings (in addition to option (a))
  - Assign a code of 0-9 which best describes the current predominant land use within circles around each station at radii of 100 m, 1 km, and 10 km
  - The intent should be to code the predominant types of land use; not all types.

Metadata Examples

0. UNKNOWN
1. NON-VEGETATED (barren, desert)
2. COASTAL OR ISLAND
3. FOREST
4. OPEN FARMLAND, GRASSLAND OR TUNDRA
5. SMALL TOWN, LESS THAN 1000 POPULATION
6. TOWN 1000 TO 10,000 POPULATION
7. CITY AREA WITH BUILDINGS LESS THAN 10 METERS DISTANT*
8. CITY AREA WITH BUILDINGS GREATER THAN 10 METERS DISTANT*
9. AIRPORT

*of three stories or greater

Metadata

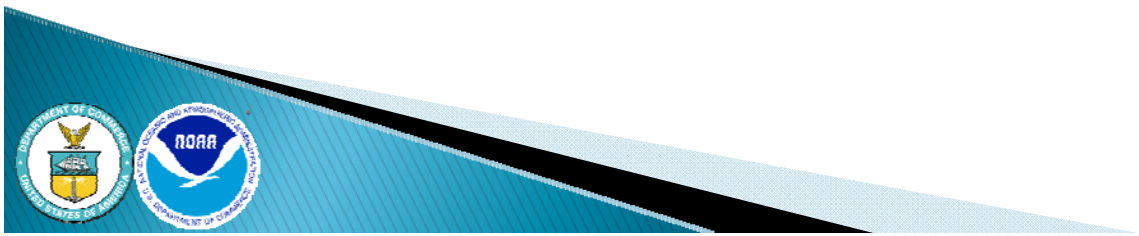
- Little metadata for stations outside the U.S. reside in NCDC's Oracle-based metadata information systems
- Given the necessary resources, NCDC could develop the capacity to store, manage, and provide access to new sources of metadata from the international community. One option is to provide in the form of web services

Metadata

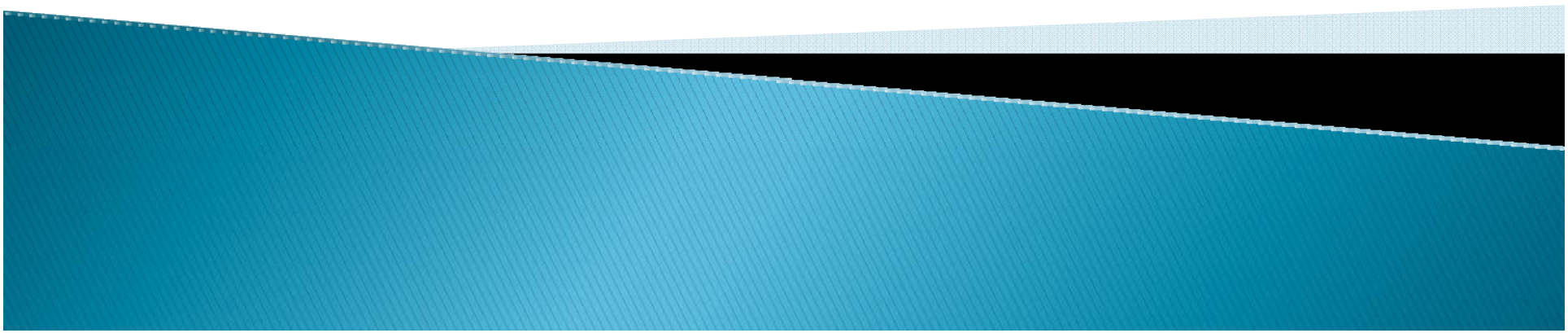
- NCDC can accept, as a minimum, photographs of station siting, as well as information regarding the type of instrumentation and the timing of station location and instrument changes.

Summary

- Daily data
  - GHCN-Daily as the GSN Archive
- Monthly data
  - GHCN-Monthly and CLIMAT messages
- International Surface Temperature Initiative
  - Global Databank
- Need for Improved Metadata

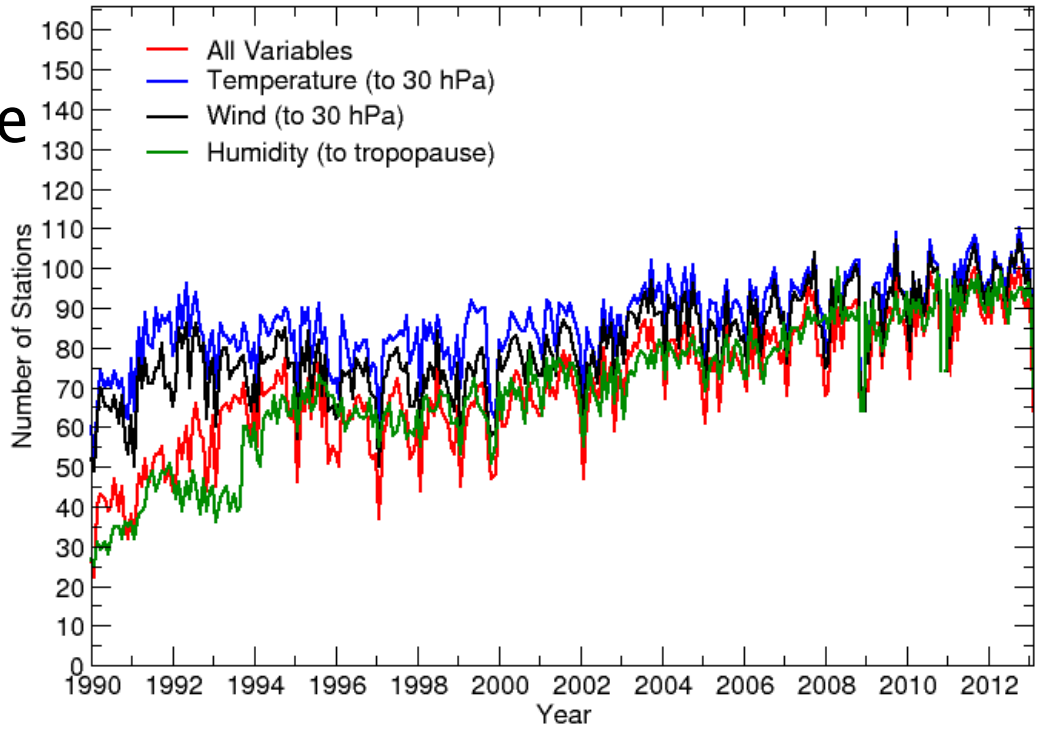


Thank you!



GUAN reports

- The number of GUAN stations meeting the minimum performance requirements
  - Temperature to 30 hPa
  - Wind to 30 hPa
  - Humidity to Tropopause
- Reports have been steady to slightly increasing during the last few years.



Currently 171 stations in the network