From Prehistoric Ice Cores to Modern Satellite Observations – From the Depths of the Ocean to the Surface of the Sun

Environmental data and services from NOAA National Centers for Environmental Information

Margarita Gregg, Ph.D.
Deputy Director, NOAA’s National Centers for Environmental Information

December 2015
National Centers for Environmental Information

• Responsible for hosting and providing access to one of the most significant archives on Earth, with comprehensive oceanic, atmospheric, and geophysical data

• From the depths of the ocean to the surface of the sun and from million-year-old sediment records to near real-time satellite images

• Nation’s leading authority for environmental information
NCEI’s Strategic Vision

MISSION

Steward of the Nation’s Environmental Information

NCEI is responsible for preserving, monitoring, assessing, and providing public access to the Nation’s treasure of environmental data and information. NCEI data and information spans the breadth of weather (atmospheric and space), climate, oceans, geophysical, and coastal disciplines.

VISION

Be the Nation’s Trusted Authority on Environmental Data and Information.

NCEI will be the most comprehensive, accessible, and trusted source of state-of-the-art environmental data, information, and monitoring.
NCEI Slogan

Revealing the Past, Interpreting the Present, and Informing the Future
History of NCEI

• National Climatic Data Center, Asheville, North Carolina
  – Established in 1951 as the Weather Bureau Records Center
  – Renamed the National Climatic Center in 1970 and the National Climatic Data Center in 1984 to reflect its large collection of weather and climate data

• National Geophysical Data Center, Boulder, Colorado
  – Established in 1965 from the Coast and Geodetic Survey and Central Radio Propagation Laboratory

• National Oceanographic Data Center, Silver Spring, Maryland
  – Established in 1961 as an interagency facility administered by the U.S. Naval Hydrographic Office
  – Included Coastal Data Development, Stennis Space Center, Mississippi
    • Established in 2000 by NOAA as the coastal division of the National Oceanographic Data Center
Merger of NOAA Data Centers

- Part of Strengthening NESDIS
- Approved in the Consolidated and Further Continuing Appropriations Act, 2015, Public Law 113-235
- Build upon the full spectrum of weather/climatic, oceanographic, coastal, and geophysical products and services that the Data Centers previously delivered
- Provide consistent data management capability for all of NOAA
Geophysics

- Nation’s geophysical data archive ranging from the surface of the sun to the Earth’s seafloor and from the solid earth environment to weather in space
- Data are from satellites, space observations, ships, and models provides information on tsunamis, the U.S. Extended Continental Shelf, coastal Digital Elevation Models, geomagnetism, solar, and terrestrial
- Products and data support safe navigation both in space and on our oceans and coasts including the Arctic and the information needed in tsunami forecasting, sea level rise prediction, and storm surge inundation
Oceans and Coasts

• World’s largest archive of oceanographic and coastal data, ranging from water temperatures dating to the late 1700s to present day ocean salinity, nutrients, waves and currents

• Data are from ocean and coast observations, ships, buoys, satellites, remote sensing, and ocean model simulations

• Products and data are used to answer questions about ocean and coast phenomena, management of coastal and marine resources, natural disasters, and marine transportation
Weather and Climate

- World’s largest weather and climate data archive with records ranging from paleoclimatology data to centuries-old journals to data less than an hour old
- Data are from land-based weather and climate stations, ships, buoys, weather balloons, radar, satellites, and comprehensive weather and climate models
- A suite of information products is available to describe the national and global climate and to monitor the state of the weather and climate variations, extremes, and trends
National Centers for Environmental Information (NCEI) Functional Organization
NCEI has a Nationwide Presence

National Centers for Environmental Information (NCEI)
- Asheville, North Carolina
- Boulder, Colorado
- Silver Spring, Maryland
- Stennis, Mississippi

NCEI Field Locations
- Regional Climate Services Directors (RCSD)
- Cooperative Institutes
- Regional Climate Centers (RCC)
NCEI Ingests and Archives Environmental Data from U.S. and International Sources

Data spans stone-age to space-age... from the depths of the ocean to the sun ... and across the globe
NCEI products span from local to global, and weekly to decadal scales

- **Daily/Weekly**
  - Snowfall Impact Index
  - FEMA, disaster response
  - Hurricane Tracks
  - Emergency Planners
  - Tsunami Warning
  - Emergency Managers

- **Monthly**
  - Heating & Cooling Degree Days
  - Energy Sector
  - Solar Activity/Sun Spots
  - Power Distribution

- **Seasonal – Annual**
  - Temperature & Precipitation Outlooks
  - Agriculture
  - Billion $ Disasters, Climate Extremes Index
  - Insurance

- **Annual to Decadal**
  - Coastal Digital Elevation Models (DEM)
  - Hazard Mitigation
  - Climate Normals
  - Construction, Infrastructure, Agriculture
  - IPCC & National Climate Assessments
  - Gov’t Policymakers

- **Sectors**
  - Agriculture
  - Energy Sector
  - Insurance
  - Power Distribution
  - Solar Activity/Sun Spots
  - Hurricane Tracks
  - FEMA, disaster response
  - Snowfall Impact Index
  - Tsunami Warning
  - Emergency Managers

- **Regions**
  - Local
  - Regional
  - National & Global
NCEI Environmental Data Archive Volume
Increasing Data Volumes from Station, Model, Radar, and Satellite Sources

2015 Total: 24 Petabytes

24 PB = Storage in a stack of iPhones 16 Eiffel Towers high (1,063 feet)

Due to increase in satellite and model data
NCEI Supports the Full Information Lifecycle

- Makes foundational investments in environmental information production and preservation
- Supports others’ application development and policy/decision-making
## Who Are Our Users? NCEI’s General User Profile

<table>
<thead>
<tr>
<th>Fraction (%)</th>
<th>Typical User</th>
<th>Data or Info Need</th>
<th>Preferred Format</th>
<th>Access Volume</th>
<th>Access Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>~70</td>
<td>General business, media, public</td>
<td>Qualitative</td>
<td>Point-and-click, graphics, assessments</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>~15</td>
<td>Researchers, business consultants</td>
<td>Quantitative</td>
<td>Digital downloads</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>~15</td>
<td>Value-added Providers (database scrapers)</td>
<td>Quantitative</td>
<td>Digital downloads machine to machine</td>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>
NCEI Environmental Data User-Requested Volume
Increasing Data Requests from Station, Model, Radar, and Satellite Sources

Types of Data Delivered:
- Satellite
- Radar
- Model
- Ground Based and Ocean Based

Volume (Petabytes)

<table>
<thead>
<tr>
<th>FY08</th>
<th>FY09</th>
<th>FY10</th>
<th>FY11</th>
<th>FY12</th>
<th>FY13</th>
<th>FY14</th>
<th>FY15</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>1.0</td>
<td>1.5</td>
<td>2.0</td>
<td>5.0</td>
<td>6.0</td>
<td>6.5</td>
</tr>
</tbody>
</table>
NCEI is Responding to a Broad Spectrum of Users
Users Increasingly Want Expert Interpretations

Percent changes in the amount of precipitation falling in very heavy events (the heaviest 1%) from 1958 to 2012 for each region. There is a clear national trend toward a greater amount of precipitation being concentrated in very heavy events, particularly in the Northeast and Midwest. (Figure source: updated from Karl et al. 2009).

**Data Documentation Division Precipitation**

**LONGITUDE:** Optional field given as part of geographic location output option. This value is given in decimated degrees to 4 decimal places. Western hemisphere values are less than 0. The maximum number of characters for this field is 10.

**DATE:** This is the year of the record (4 digits), followed by month (2 digits), followed by day of the month (2 digits), followed by a space and ending with a time of observation that is a two digit indication of the local time hour, followed by a colon (:) followed by a two digit indication of the minute which for this dataset will always be 00. Note: The subsequent data value will be for the hour ending at the time specified here. Hour 00:00 will be listed as the first hour of each date, however since this data is by definition an accumulation of the previous 60 minutes, it actually occurred on the previous day. 14 characters.

**HPCP:** The amount of precipitation recorded at the station for the hour ending at the time specified for DATE above given in hundredths of inches. The values 99999 means the data value is missing. The maximum number of characters for this field is 8. Hours with no precipitation are not shown. Note that observational elements listed above that are identified as optional (station name, geographic location and flags) must be added to the data output by means of selecting the appropriate checkbox for them (seen when using the Climate Data Online interface).
Assessing the Earth
International, National, Annual Assessments

International Assessments
- IPCC Fifth Assessment
- 2 NCEI lead authors on Fifth Assessment Report

National Assessments
- National Climate Assessments
- National Tsunami Hazard Assessment

Annual Assessments
- Bulletin American Meteorological Society (BAMs) State of the Climate Report
Tiers of Stewardship

6: National Services and International Leadership
- Lead, coordinate, or implement scientific stewardship activities for a community or across disciplines
- Establish highly specialized levels of data services and product assessments

5: Authoritative Records
- Combine multiple time series into a single, inter-calibrated product
- Establish authoritative quality, uncertainties, and provenance
- Ensure products are fully documented and reproducible

4: Derived Products
- Build upon archived data to create new products that are more broadly useful
- Distill, combine, or analyze products and data to create new or blended scientific data products

3: Scientific Improvements
- Improve data quality or accuracy with scientific quality assessments, controls, warning flags, and corrections
- Reprocess data sets to new, improved versions and distribute to users

2: Enhanced Access and Basic Quality Assurance
- Create complete metadata to enable automated quality assurance and statistic collection
- Provide enhanced data access through specialized software services for users and applications

1: Long Term preservation and Basic Access
- Preserve original data with metadata for discovery and access
- Serve as expert advisors on standards for data providers
- Archive only necessary data using appropriate retention schedules
- Safeguard data over its entire life-cycle
- Coordinate support agreements for sustainable data archiving
- Provide data citation services by mining DOIs
Examples of Reference Products

Climate Normals
Reference Environmental Data Records

World Ocean Atlas
Gulf of Mexico Data

Earth’s Magnetic Field
Coastal/Ocean Depths
Value of NCEI
Science, Service, Stewardship

- Comprehensive environmental information across economic/environmental sectors and science disciplines for decision making, planning, and operations
- Timely and authoritative data and information with well described origin and certainty
- Service to a wide-range of customers across government, academia and industry
- Preservation of data and information for future generations
NCEI Dependencies and Requirements

• NCEI cost model for data
• Developing partnerships with OSGS and other NESDIS offices
• Modernizing NCEI’s Data Stewardship and IT Infrastructure
• Developing the next generation of data access
  – As the three websites from the former NO, NG, and NC are brought together, NCEI has an opportunity to improve how our customers access NOAA data. We must ensure that the billions spent on observations are fully leveraged by getting info into the hands of users. NCEI’s website is not currently designed to deliver data effectively
  – Linkages to the possible NOAA portal TBD
• Meeting new data requirements
  – NCEI should be ready to assist other NOAA offices as they work towards meeting the requirements under PAR, and the President’s Open Data Initiative. Resources are not currently in the NCEI budget
Examples of “Gold” standard NCEI Products (Data Set Maturity Matrix Model Level 6)

- Climatological Atlas of the Nordic Seas and Northern North Atlantic
- World Ocean Atlas 2013
- National Climate Assessment
- BAMS State of the Climate in 2013
- Explaining Extreme Events of 2013 from a Climate Perspective
- Extended Continental Shelf (ECS) Project
- Post-Sandy Digital Elevation Model
- World Magnetic Model for 2015-2020
The FY 2016 NCEI Priorities’ guiding principles are:
- Ensure data security.
- Provide users with more effective and integrated ways of obtaining weather, climatic, coastal, oceanographic, and geophysical data and information.

The FY 2016 NCEI Priorities’ key themes are:
- Build on our strengths.
- Advance development and distribution of integrated datasets, products, and services.
- Further organizational excellence.
NCEI FY16 Focus Areas

• Enhance NCEI’s Suite of Use-Inspired Data Products
• Develop Integrated Data Products to Monitor and Assess the Changing Environment
• Improve the Collection and Validation of User Requirements
• Support Data Services
• Build Common Stewardship Processes to Increase User Confidence
• Build Common Interfaces to Improve Data Access
• Implement a Common NCEI Project Management Process
• Invest in the NCEI Workforce
• Integrate with NOAA and NESDIS Enterprise Initiatives
NCEI: Acorns that Seed the Future

Data & Information Leading to Future Knowledge

Observing Systems & Model Outputs
BACK UP SLIDES
NCEI Data Relevant to Energy-Land-Water

Key to understand the interplay of physical, biological and social processes at the intersection of the energy-water-food nexus
# How do we Develop & Maintain “Gold Standard” Reference Datasets?

- Apply a maturity model similar to engineering readiness levels - a practical approach
- Goes beyond measurements of quality to *usability*, *transparency*, and *understanding*
- Assess products using a “Maturity Matrix” covering seven aspects of data maturity

<table>
<thead>
<tr>
<th>Level</th>
<th>Data Use</th>
<th>Code Stability</th>
<th>Metadata &amp; QA</th>
<th>Documentation</th>
<th>Validation</th>
<th>Public Release</th>
<th>Science &amp; Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Unified and coherent record; considered scientifically irrefutable following extensive scrutiny</td>
<td>Homogeneous and published error budget</td>
<td>Provenance tracking and reproducibility; meets international standards</td>
<td>Peer-reviewed product algorithm, validation, processing and metadata</td>
<td>Validated independent cross-checks, open inspection, and continuous interrogation</td>
<td>Publicly available from long-term archives</td>
<td>Used in various published applications and assessments</td>
</tr>
<tr>
<td>5</td>
<td>Unified and coherent record</td>
<td>Stable and reproducible</td>
<td>See below</td>
<td>See below</td>
<td>See below</td>
<td>Source code portable and released; uncertainty estimate</td>
<td>See below</td>
</tr>
<tr>
<td>4</td>
<td>Research and Operations</td>
<td>See above</td>
<td>Provenance tracking and reproducibility; meets international standards</td>
<td>Draft Operational Algorithm Description</td>
<td>See below</td>
<td>Source code released; Data available but of unknown accuracy</td>
<td>See below</td>
</tr>
<tr>
<td>3</td>
<td>Research</td>
<td>Minimal changes expected</td>
<td>See below</td>
<td>Peer-reviewed algorithm and product descriptions</td>
<td>Uncertainty estimated over widely distribute times/location; differences understood</td>
<td>See below</td>
<td>Provisionally used in applications and assessments demonstrating positive value</td>
</tr>
<tr>
<td>2</td>
<td>Research</td>
<td>Some changes expected</td>
<td>Research grade (extensive)</td>
<td>CTD</td>
<td>Uncertainty estimated for select locations/time</td>
<td>Data available but of unknown accuracy</td>
<td>Limited or ongoing</td>
</tr>
<tr>
<td>1</td>
<td>Research</td>
<td>Significant changes likely</td>
<td>Incomplete</td>
<td>Draft concept of Theory Document (CTD)</td>
<td>Minimal</td>
<td>Limited data availability</td>
<td>Little or none</td>
</tr>
</tbody>
</table>