Increasing global accessibility and understanding of water column sonar data

Carrie C. Wall

Charles Anderson, Sharon Mesick, Rost Parsons, Tim Boyer, and Susan J. McLean

www.ngdc.noaa.gov/mgg/wcd

NOAA Satellite and Information Service | National Centers for Environmental Information (NCEI) | Center for Coasts, Oceans, and Geophysics (CCOG)
NCEI...never heard of it!

- **National Centers for Environmental Information** (NCEI) formed from the merging of NOAA’s Data Centers:
  - National Geophysical Data Center (NGDC), Boulder, CO
  - National Oceanographic Data Center (NODC), Silver Spring, MD
  - National Climatic Data Center (NCDC), Asheville, NC

- Integrated ocean data stewardship
  - *Easier* to contribute data
  - Integrated discovery of and access to *all* data from a survey
  - Covering the gaps: assuring a proper home for all ocean data
Water Column Sonar Data

• Acoustic backscatter from the ocean surface to the seafloor

• Sonar Technology
  – Singlebeam echosounder
    • Single and multiple frequency
  – Multibeam echosounder
    • Single and multiple frequency
Fisheries Acoustics

- Why use multiple frequencies?
  - Depth range varies
  - **Signal size varies**
    - Large, swimbladder organisms reflect strongest at low frequencies (18 kHz)
    - Small organisms reflect strongest at high frequencies (200 kHz)

### Echosounder Frequency (kHz)

<table>
<thead>
<tr>
<th>Type</th>
<th>Volume backscattering strength $S_v$ (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copepods</td>
<td>-40</td>
</tr>
<tr>
<td>Small euphausiids</td>
<td>-40</td>
</tr>
<tr>
<td>Mackerel</td>
<td>-40</td>
</tr>
<tr>
<td>Small resonant bubbles</td>
<td>-40</td>
</tr>
<tr>
<td>Large bubbles</td>
<td>-40</td>
</tr>
<tr>
<td>Deep physoctists</td>
<td>-40</td>
</tr>
<tr>
<td>Deep phystostome</td>
<td>-40</td>
</tr>
</tbody>
</table>

**Fish with a swimbladder (hake)**

**Fish without a swimbladder (mackerel)**

**Small euphausiids (krill)**

**Copepod (small crustaceans)**

© Micropolitan.org

Photo: Stephen Brookes

Trenkel & Berger (2013)
Applications of Fisheries Acoustic Data

• Uses
  – Identify fish school morphology and behavior, estimate biomass, determine species and
tropic level
    • Primary means to non-invasively monitor fish and their environment
    • Valuable for fisheries management
  – Characterize habitat
    • Typically for commercially important species
    • Valuable for understanding and delineating essential fish habitat
  – Map methane seeps and undersea plumes (bubbles)
    • Valuable for environmental monitoring and oil and gas exploration
Water Column Sonar Data Archive

- National Marine Fisheries Service (NMFS) collects a lot of data
  - 78+ TB backlogged data, 12+ TB annually
  - **Increasingly large volume not stored sustainably, publicly funded data now need to be accessible**

- Partnership between NMFS and NCEI initiated in 2011
  - 2013: Water column sonar data project fully staffed and operational
  - **Goal: Stewardship, discovery, delivery and development of products related to NOAA’s water column sonar data at NCEI**

- 21.45 TB of data archived to date
  - Data received from NOAA, international and academic institutions
  - **12.74 TB of data requested and delivered**

---

**Data by Source (GB)**
- NMFS: 61.3%
- Other NOAA: 15.4%
- Other: 14.5%
- UNOLS: 9.2%
- OER: 8.1%
- NCCOS: 0.5%

**Data by Instrument (GB)**
- 40.5% Single frequency multibeam
- 31% Multiple frequency multibeam
- 15.1% Single and multiple frequency singlebeam
Water Column Sonar Data Archive
From the vessel to the web

Water Column Sonar Data Pipeline

Data Provider

- WCSD Packager
  - Create bag structure
  - Copy files to external drive
  - Generate cruise and series-level xml metadata files
  - Process instrument files for configuration parameters and trackpoints
  - Generate file-level xml metadata files
  - Update MD5 checksum manifest file

Instrument files

User entered cruise and instrument details

Data Package

- MD5 checksum manifest file
  - Data directory
    - Instrument data
      - File-level metadata xml(s)
    - Ancillary data directory
      - Calibration
      - CTD
      - XBT
      - SCS
      - Cruise-level metadata xml
      - Series-level metadata xml

Public

WCSD Map Viewer

WCSD Data Manager

WCSD Ingest

- MD5 / manifest validation
- Parse cruise-level xml
- Parse series and file-level xml
- Gzip and prestige files

Archive DB

Archive Tape System

WCSD Data Manager
From the vessel to NCEI

• Easy data packaging interface
  – Standalone executable
  – Creates ISO compliant metadata
  – Hands-free once initiated

• Service Level Agreement with NMFS

• Submission Agreements
  – Completed:
    • AFSC, NEFSC, NWFSC, SEFSC & PIFSC
    • NCCOS, several academic institutions
  – To come:
    • SWFSC
From the vessel to NCEI

![WCSD Packager interface](image)

- **Calibration State**: Calibrated w/ calibration data
- **Calibration Date**: 6/14/2011
- **Calibration Location Lon/Lat**: 57.1768, -153.5019

**Calibration Location Name**: Three Saints Bay, Kodiak AK

**Path to calibration files**: /media/NGDC_1/DY1103/data/ancillary_data/calibration

**Calibration State Options**:
- Unknown: Calibration state not known and no calibration data available
- Uncalibrated: The instrument is known not to be calibrated and no calibration data is being submitted
- Uncalibrated w/ calibration data: Calibration was performed and data is being submitted, but instrument parameters were not updated
- Calibrated, w/o calibration data: Instrument calibrated and instrument parameters were updated, but no calibration data is being submitted

[Package Data]
From the archive to the web

Water Column Sonar Data

Abstract
Water column sonar data are collected by active acoustic (or sonar) technology and contain acoustic scattering information from the near surface to the seafloor. Scientists use these data to assess physical and biological characteristics of the ocean, including the spatial distribution of plankton, fish, methane seeps, and underwater oil plumes.

Data Providers: NOAA National Marine Fisheries Service (NMFS), NOAA Office of Ocean Exploration and Research (OER), University-National Oceanographic Laboratory System (UNOLS), NOAA offices outside of NMFS and OER (other NOAA), U.S. academic or private institutions (Other), and international groups (Non-U.S.).

Data Collection: A variety of sonars are used by the data providers, with Kongsberg’s EM 122 (12 kHz) and EM 302 (30 kHz), Simrad’s EK60 (18-710 kHz, split beam), and ME 70 (70-120 kHz, can be split beam) being most common. The configuration of each cruise’s sonar system (e.g., beam type and angle) can be found in the file metadata.

Data Citation: Cruises have specific citations that should be used when incorporating these data into your research, etc. Please see the More Information Help.

John Cartwright, NCEI
Jesse Varner, CIRES/NCEI
From the archive to the web
Current and Future Initiatives
Increasing the Impact of the Archive

• Provide direct link to NCEI CTD, bathymetric, biological and other oceanographic data

• Archive all NMFS backlogged and current data from all sonar systems
  – Completed for 3 of 6 NMFS FSC

• Incorporate data products to help users understand what’s in the data
  – Current efforts are focused on singlebeam data
  – Future work will expand to multibeam data

What’s in the data?
Processing algorithms transform these complex data into a single image that illustrates the quality and content of the acoustic data.
Users will be able to understand the quality and content of the raw data files prior to submitting data requests:

- More informed on whether the data/area of interest is in fact relevant to their research question
- Creates a more efficient process for both the user and the archive staff
- Users of various backgrounds can easily understand these complex data

Visualizing the Archived Data

- Noise in high frequency transducers
- Patches of zooplankton and fish
Benefits of the Archive

• Easy to use packaging interface to assist in submitting data to the archive
  – Help data providers meet their data management directives
  – Create ISO compliant metadata, mint digital object identifiers

• Easy access to the archived data through online web map
  – Central repository for global access to an unprecedented dataset

• Increase the potential for these data to be used beyond their original collection purpose
  – Identify areas data gaps or repeated surveys
  – Aim to promote collaboration and research that advances the field of marine ecosystem acoustics
Acknowledgements

• Special thanks to
  – *Data Providers & Partners*: Michael Jech, Rick Towler, Scott Furnish, Warren Mitchell, Dezhang Chu, Réka Domokos, Joe Warren, OER (Sharon Mesick and Susan Gottfried), and UNOLS
  – *Support*: Dan Kowal, John Cartwright, Jesse Varner, Dani Austin, Anna Milan, John LaRocque, Michael Gallagher, David Detlor

• Funding provided by

---

16