
Christian Schwatke, Denise Dettmering and Wolfgang Bosch

Deutsches Geodätisches Forschungsinstitut (DGFI)
email: schwatke@dgfi.badw.de
DAHITI

- New “Database for Hydrological Time Series of Inland Water” (DAHITI)
- Currently, it contains about 180 water level time series over lakes, rivers, reservoirs, and wetlands
- Free data access via Open Altimeter Database (OpenADB) (http://openadb.dgfi.badw.de)
Rationale for DAHITI

What is new in DAHITI?

The DAHITI processing is based on:

- Using **multi-mission** altimeter data whenever available
- Transform ellipsoidal heights to **physical heights** using EIGEN6C2 geoid model (where water will flow)
- Assume water bodies to form an equipotential surface (equal physical heights)
- Waveform classification by „**Support Vector Machine (SVM)““
- Outlier rejection by „**Support Vector Regression (SVR)““
- Building a smooth space-time series by a **Kalman Filter** approach
# Data Holding

<table>
<thead>
<tr>
<th>Mission</th>
<th>1Hz</th>
<th>High-frequent</th>
<th>Retracking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cryosat-2 (20Hz)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Envisat (20Hz)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ERS-1 (20Hz)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ERS-2 (20Hz)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Geosat (10Hz)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>GFO (10Hz)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>HY-2A (20Hz)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>IceSat (40Hz)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Jason-1 (20Hz)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Jason-2 (20Hz)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Saral/Altika (40Hz)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Topex/Poseidon (10Hz)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Methodology: Preprocessing

1. Extraction of raw data from OpenADB

2. Classification (optional)
   - Using method of „Support Vector Machine (SVM)“
   - Three Classes (Brown-Linear, Brown-Exponential, Single Peak)
   - Two Features (Kurtosis, Skewness)
   - SVM model includes the decision function to assign every waveform to one class

3. Retracking (optional)
   - Each waveform is assigned to a retracking algorithm

<table>
<thead>
<tr>
<th>SVM-Class</th>
<th>Retracker</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown-Linear</td>
<td>5-Parameter β-Retracker</td>
<td>Zwally and Brenner, 2001</td>
</tr>
<tr>
<td></td>
<td>(linear trailing edge)</td>
<td></td>
</tr>
<tr>
<td>Brown-Exponential</td>
<td>5-Parameter β-Retracker</td>
<td>Deng and Featherstone, 2006</td>
</tr>
<tr>
<td></td>
<td>(exponential trailing edge)</td>
<td></td>
</tr>
<tr>
<td>Single-Peak</td>
<td>Improved Threshold Retracker</td>
<td>Hwang et al., 2006</td>
</tr>
</tbody>
</table>
Methodology: Preprocessing

4. Physical heights estimation using original or retracted ranges

5. Calculation of standard deviations of heights along pass

6. Reject outliers
   - Min/max height
   - Limit of standard deviation
   - Classes from classification
   - „Support Vector Regression“ along pass
   - „Support Vector Regression“ along mission
Kalman Filter Approach
- Input data: Physical heights and STD
- Noise and errors of data are considered
- Individual time steps (e.g. 1d, 10d, 30d)

Processing
- Create hexagonal grid for target
- Interpolate heights to nearest node
- STD used for weighting the data
- New heights are estimated by weighting heights from an earlier time step and new data at current time step
- Forward and backward filtering
- Mean height of nodes is estimated considering error limit (e.g. 0.1m)
Results

Lake Mweru (5,120 km²)

Lake Michigan (58,016 km²)

- Correlation with gauge (red): 0.95
- Very good absolute agreement due to same height reference (WGS84)
Results

Lake Manitoba (4,700 km²)

- Correlation with gauge (red): 0.81
- Offset due to unknown height reference of gauge
- Ice coverage in altimeter data visible by negative peaks

Lake of the Woods (4,349 km²)

- Correlation with gauge (red): 0.82
- Offset due to unknown height reference of gauge
Results

Lake Chiuta (199 km²)

- Even very good results in very small lakes and rivers

Amazon River

- Correlation with gauge (red): 0.98
- Offset due to unknown location and height reference of gauge
Lake Chad (1,500 km²)

Improvements of time series after retracking Jason-1 and Jason-2

<table>
<thead>
<tr>
<th>Mission</th>
<th>Passes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jason-1 (20Hz)</td>
<td>248</td>
</tr>
<tr>
<td>Jason-2 (20Hz)</td>
<td>248</td>
</tr>
<tr>
<td>Topex (10Hz)</td>
<td>248</td>
</tr>
</tbody>
</table>

*retracked
Discussion

• DAHITI provides promising time series of inland water for hydrological applications

• Kalman filter approach for height estimation and SVR for outlier detection leads to smooth reliable time series with high correlations with gauges

• Additional classification and retracking leads to improved results in small water bodies

Outlook

• Error assessment for the lake level heights

• Expansion of the altimeter data base with high-frequent and SGDR data

• Expansion of DAHITI by new targets

• Use more classes in the classification step (e.g. brown+peak)

• Use other/improved retracking algorithms
See you on OpenADB

http://openadb.dgfi.badw.de

Poster No. 13 (CAW):

DAHITI: A New Database of Water Level Time Series for Lakes, Rivers and Wetlands from Multi-Mission Satellite Altimetry

(Schwatke C., Dettmering D., Bosch W.)

Poster No. 28 (OSTST):

OpenADB: An Open Altimeter Database providing high-quality altimeter data and products

(Schwatke C., Dettmering D., Bosch W.)

Acknowledgement:

DAHITI holdings are based on altimetry missions operated by CNES/NASA (TOPEX, Jason-1), ESA (Envisat, Cryosat-2), USNavy/NOAA (GFO), CNES/NASA/Eumetsat/NOAA (Jason-2), and ISRO/CNES (Saral). The mission data are disseminated by AVISO, ESA, NOAA, and PODAAC. The time series of gauges are provided by NOAA Tides and Currents, Lake of the Woods Control Board, and Water Survey of Canada.