Surface Water and Ocean Topography (SWOT) Mission

Lake Data Products
Tamlin Pavelsky and Jean-Francois Cretaux

SWOT Science Definition Team Meeting, January 14, 2015
Endmembers of SWOT Lake Hydrologists

Endmember 1: The Local Dynamicist
- Interested in detailed spatial variations in water surface height.
- Wants to see local dynamics of lake shore variations over time.
- Wants all available observations, whether complete or not.
- Doesn’t mind working with large and complex datasets over a relatively small geographic area.

Endmember 2: The Global Modeler
- Interested in well-organized information on lake height, water storage, and inundation extent.
- Prioritizes data consistency: the same type of data for all lakes, in a format that’s easy to interface with a model.
- Doesn’t care about detailed patterns of inundation extent.
- Does not want to deal with datasets showing incomplete data for lakes that are only partially observed.

We want to serve both of these endmember hydrologists and everyone in between.
Proposed Vector Data Products

- **Pass-based product:**  
  - Produced separately for every SWOT overpass

- **Cycle-averaged product:**  
  - Produced from all overpasses for one 21-day cycle

- For both datasets, there needs to be an online tool allowing extraction of attribute data in tabular form.

- Both of these products will depend on an *a priori* lake mask that must be generated before launch.
A Priori Lake Information

• We already have a great deal of information about where lakes are, globally [Talk by Yongwei Sheng tomorrow].
  – Global Lakes and Wetlands Database (Lehner and Doll, 2004)
  – Landsat-derive databases (Sheng, in prep; Verpoorter et al, 2014)

• Before launch, we should develop an *a priori* lakes database that contains:
  – All lake features likely to be detectable by SWOT, with each lake having a unique identification code
  – A flag for whether each feature is likely to experience ice cover.

• This *a priori* mask should be updated during the mission using SWOT data.
Example Lakes for Illustration

Lake A: 24 km²
Lake B: 7.5 km²
Lake C: 0.8 km²

Lakes A & B in a priori mask
Lake C not in a priori mask
1. From the SWOT pixel cloud product, identify discrete, connected regions of inundation with substantially identifiable boundaries (regions 1-4). Intersect these SWOT-observed regions with the a priori lake boundaries.
2. For SWOT-observed regions that intersect a priori lakes, assign the corresponding a priori lake code. If more than one a priori lake is intersected, assign the code for the larger intersected area.
3. For SWOT-observed inundated regions that do not intersect an existing lake, assign a general code indicating this fact.
Pass-Based Lake Vector Product II

Attributes:

- Height based on the average height of inundated pixels (not including laid-over pixels), relative to best known geoid.
- Area value defined as the total observed area for the SWOT-observed region.
- A no-layover area value for the total area used to calculate the height, and various flags (ice, rain, etc.)
- Flag indicating whether the a priori lake area falls entirely within the swath.
- Flag indicating if the SWOT-observed region intersects multiple a priori lakes.

**Attribute Table:**

<table>
<thead>
<tr>
<th>SWOT ID</th>
<th>Apriori ID</th>
<th>Obs. Area</th>
<th>Error</th>
<th>Height</th>
<th>Height Error</th>
<th>Ice Flag</th>
<th>Multi Flag</th>
<th>Partial Flag</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>22</td>
<td>0.7</td>
<td>144.32</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>1.9</td>
<td>0.3</td>
<td>143.11</td>
<td>0.05</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>B</td>
<td>3.5</td>
<td>0.4</td>
<td>143.25</td>
<td>0.03</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>No</td>
<td>0.8</td>
<td>0.2</td>
<td>150.65</td>
<td>0.12</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
• Lake boundaries will vary from pass to pass and will be restricted to match the calculated area value for each pass.
• Lake heights will be calculated using whatever portion of the lake is observed, and an associated error will be calculated.
• Ice flags will be calculated using SWOT data informed by a priori information (e.g. ice will never be assigned to lakes in the central Amazon basin).
• Initially, no storage change value will be produced for the swath-based product. We propose that after at least a year, the science team produce a series of height-storage anomaly rating curves based on the cycle-based product described below.
  – With these rating curves, it will be possible to estimate storage change in the swath-based product even for lakes that are not completely observed.
Cycle Lake Vector Product I

- All water pixels identified by any overpass during the 21-day cycle will be overlaid to form a continuous, composite pixel cloud product.
- For each lake in the a priori mask, area, height, and storage change (along with associated errors and flags) will be calculated. All values will be calculated as the average or sum of all water pixels connected to the a priori lake.
- Heights will be calculated relative to the best available geoid.

NOTE: there is some disagreement on whether the final product should reflect variations in lake boundaries.
Cycle Lake Vector Product II

- Storage change will be computed relative to a baseline height-area pairing defined by the first SWOT cycle, at least until the first reprocessing.
- How to deal with lakes that are separate in the a priori mask but that become connected remains TBD.
- For lakes with partially diffuse or indeterminate boundaries (e.g. patchy wetlands), a best effort will be made to determine a suitable area based on the contour of 50% inundation. Efforts will be made to factor the resulting uncertainty into the error accounting.
- All calculated values will be assigned to a database associated with the a priori mask. In other words, every 21 days, each lake in the mask will have a new height, area, and storage change value (with associated errors and flags) added to its associated database.
- The a priori mask will be updated regularly (where the meaning of regularly is TBD), and upon update storage change values for any added lakes will be computed for all prior observations.
Query-able Lakes Databases

- It is viewed as essential that, through some mechanism, it be possible for scientists to obtain multitemporal SWOT height, area, and storage change values (along with associated flags) on all lakes in the a priori mask WITHOUT downloading many separate vector files showing the explicit lake boundaries.
- This goal is likely to be best achieved using a web-based tool that allows users to specify the region to be queried, the length of time, and which lake attributes should be output.
- This tool should be applicable to both the pass-based product and the cycle-based product.
Final Notes

- We propose that there should be no wetland vector product produced by the project, beyond the generic pixel cloud.
- We propose that reservoirs should be treated the same way as natural lakes and that they should be included in the vector product.