A Risk-Based Approach to Determine Hydrographic Survey Priorities Using GIS

Patrick Keown (Presenter)
Gonsalves, M., Allen, C., Fandel, C. Gallagher, B., Hick, L.

2016 Esri Ocean GIS Forum, Redlands, CA
Presentation Overview

• Background
• Hydrographic Health Model Overview
  • Data Sources & Methods
  • Desired Survey Score
  • Present Survey Score
  • Hydrographic Risk
  • Results
• GIS and the Hydrographic Health model
  • Big data, Big Problems
• Future Work
Background – The Ocean is Big

- U.S. Exclusive Economic Zone (EEZ): 3,400,000 nm²
- Surveyed to ‘modern’ standards: ~44,000 nm²
- Average Hydrographic Acquisition: ~3,000 nm² / year

- Need to be smart about where we choose to deploy our limited resources.

- The ultimate goal being to maximize the hydrographic return on investment.
National Hydrographic Survey Priorities (NHSP)

- Ranked areas from “Critical” to “Priority 5”, based on Depth, Survey Quality and vessel traffic.
- Original data was stored in MapInfo.
Evolution of NHSP

• In 2012, an update was released.
• New needs were captured as “Emerging Critical”

Overall Limitations with NHSP

• Technology: lacked some geospatial data.
• Agnostic to consequence
• Individual Coverage requirements were not specifically addressed.
• No mechanism for managing re-survey work, or changes in requirements.
Hydrographic Health Model

• Benefits
  • Repeatable
  • Scalable
  • Analytical
  • Authoritative (whenever possible)
  • Modular / Adaptive
  • Increased transparency and results are easier to communicate to the public (ArcGIS Services, Story Maps, etc.)
# Hydrographic Health - Data Inputs

<table>
<thead>
<tr>
<th>Data type:</th>
<th>Purpose:</th>
<th>Source:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Where are the ships going?</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIS</td>
<td>Analyze traffic patterns</td>
<td>USCG, Marine Cadastre</td>
</tr>
<tr>
<td>AVIS database</td>
<td>Distinguish vessel types; calculate under-keel clearance</td>
<td>USCG</td>
</tr>
<tr>
<td>Active Captain reports</td>
<td>Proxy for recreational boaters</td>
<td>Active Captain</td>
</tr>
<tr>
<td><strong>What's the present state of hydrography?</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CATZOC</td>
<td>Initial data quality</td>
<td>NOAA</td>
</tr>
<tr>
<td>Survey Age</td>
<td>Present data quality</td>
<td>NOAA</td>
</tr>
<tr>
<td><strong>Is there a reason to be concerned with the state of hydrography?</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth</td>
<td>Identify areas of concern; calculate seafloor complexity</td>
<td>NOAA Charts</td>
</tr>
<tr>
<td>Grounding reports</td>
<td>Identify areas of concern</td>
<td>USCG</td>
</tr>
<tr>
<td>Charted Hazards</td>
<td>Identify areas of concern</td>
<td>NOAA Charts</td>
</tr>
<tr>
<td>Reported Hazards (PA/PD)</td>
<td>Identify areas of concern</td>
<td>NOAA Charts</td>
</tr>
<tr>
<td>Natural/Artificial Reefs</td>
<td>Identify areas of concern</td>
<td>NOAA Charts</td>
</tr>
<tr>
<td>Sanctuaries</td>
<td>Identify areas of concern</td>
<td>NOAA</td>
</tr>
<tr>
<td>Ports</td>
<td>Identify areas of concern</td>
<td>NOAA, USACE</td>
</tr>
<tr>
<td>Search &amp; Rescue Stations</td>
<td>Identify areas of concern</td>
<td>USCG</td>
</tr>
<tr>
<td>Bottom Type</td>
<td>Identify areas of concern</td>
<td>NOAA Charts</td>
</tr>
<tr>
<td><strong>Is there a reason to believe the seafloor is changing?</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hurricanes</td>
<td>Identify areas of change</td>
<td>NOAA</td>
</tr>
<tr>
<td>Tidal Currents</td>
<td>Identify areas of change</td>
<td>Georgia Tech</td>
</tr>
<tr>
<td>Marine Debris</td>
<td>Identify areas of change</td>
<td>NOAA Charts</td>
</tr>
</tbody>
</table>
A Model of Hydrographic Health

- **Hydrographic Gap**
  - The difference between the desired and present survey score; the larger the gap, the worse the hydrographic health.

- **Hydrographic Risk**
  - Modeled as the risk to surface navigation due to inaccurate depths or unknown hazards; the greater the risk, the worse the hydrographic health.

\[
\text{Hydrographic Health} = \left( \frac{\text{Desired Survey Score} - \text{Present Survey Score}}{\text{Hydrographic Gap}} \right)^2 \times \sum \left( \text{Consequence} \times \prod \text{Likelihood} \right) \]

\[
\begin{align*}
\text{Hydrographic Gap} & \quad \text{Hydrographic Risk}
\end{align*}
\]
A combination of Seafloor Complexity and Under Keel Clearance will inform the desired Coverage Requirement.
The Present Survey Score is combination of the Initial Survey Quality, the Changeability of the area, and the Time Elapsed since the last survey.

Initial Survey Quality is strictly a function of CATZOC.

\[
\text{Hydrographic Health} = (\text{Desired Survey Score} - \text{Present Survey Score})^2 \times \sum \left( \text{Consequence} \times \prod \text{(Likelihood)} \right)
\]
The Depreciation is modeled through exponential decay, where the **Decay Coefficient**, is the sum of several change parameters.

\[
\text{Present Survey Score} = \text{Initial Survey Score} \times e^{-\text{(Decay Coefficient)} \times \text{(Age of Survey)}}
\]

\[
\text{Decay Coefficient} = (\Delta \text{Storms} + \Delta \text{Currents} + \Delta \text{Human Debris}) \times 0.022 / 4
\]

Where having one change term set to maximum decay will yield:

- CATZOC A → B in 10 years.
- CATZOC A → C in 55 years.
- CATZOC A → D in 100+ years.
Hydrographic Health $= \left( \text{Desired Survey Score} - \text{Present Survey Score} \right)^2 \times \sum \left( \text{Consequence} \times \prod \text{(Likelihood)} \right)$

All Depreciation Terms: Past to Present

Initial Survey Quality

Present Survey Score
Over 10 years...
- 6,500 nm² CATZOC A → B, of which...
  - I desire 1,200 nm² to remain as CATZOC A.
  - Implying I should plan to survey ~120 nm²/year to Florida to maintain desired state.

10 Years Later

Fully surveyed!!!
No Gap isn’t necessarily a victory condition…

…could be a measure of institutional inefficiency.
Hydrographic Gap: Desired – Present Survey Score

Hydrographic gap if you ignored survey depreciation...

...this should end the concept of one-and-done hydrography.
Total Risk = Risk to Ship + Risk to Commerce + Risk to Environment

**CONSEQUENCES**

- Risk to SHIP
  - Proximity to Search and Rescue
  - Bottom Type

- Risk to COMMERCE
  - Proximity to Port
  - Impact to Tourism
  - Impact to Fisheries

- Risk to ENVIRONMENT
  - Proximity to Reefs
  - Proximity to Sanctuaries
  - Proximity to Oil Response

**LIKELIHOODS**

- Depth
- Traffic Density
- Survey Quality (CATZOC)
- Reported Hazards (PA/PD)
- Known Hazards
- Reported Groundings

* Terms in gray not presently incorporated into model.

Model designed to be modular in nature… include what you can.
Consequence to Reefs (dependent on vessel type):

**CONSEQUENCE: IMPACT TO REEFS & SANCTUARIES**
Data: Distance to Reefs for Tankers versus other ships.

If **TANKER** passes:
- 5 < 2.5 nm to reef
- 4 2.5 - 5 nm
- 3 5 - 10 nm
- 2 10 - 20 nm
- 1 > 20 nm from reef
- 0 -

If **OTHER VESSEL** passes:
- 5 < 2.5 nm to reef
- 4 2.5 - 5 nm
- 3 5 - 10 nm
- 2 10 - 20 nm
- 1 > 20 nm from reef
- 0 -

**Tanker**

**Other**
Likelihood due to Reported Hazards (distance & density driven):

LIKELIHOOD: REPORTED ERRORS
Data: MAX of PA/PD Count (within 2nm) plus “Rep” (within 4nm), or Distance to nearest PA/PD, or Distance to nearest “Rep.”

- Count of 4+:
  - 5 <0.25 nm to PA/PD
  - 5 <0.50 nm to “Rep”

- Count of 2-4:
  - 4 0.25 – 0.50 nm
  - 4 0.50 – 1.0 nm

- Count of 1:
  - 3 0.50 – 1.0 nm
  - 3 1 – 2 nm

- No Errors:
  - 2 1 – 2 nm
  - 2 2 – 4 nm

- >2 nm to PA/PD:
  - 1 >4 nm to “Rep”
Likelihood due to Traffic density (Unique vessel count):

- **Anchorage Areas**
- **Ferry Route**
- **Moored Vessels**

### LIKELIHOOD: TRAFFIC DENSITY

- **5**: 100+ Boats
- **4**: 11 – 100 Boats
- **3**: 2 – 10 Boats
- **2**: 1 Boat
- **1**: No AIS
\[ \text{Hydrographic Health} = \left( \frac{\text{Desired Survey Score}}{\text{Present Survey Score}} \right)^2 \times \sum \left( \text{Consequence} \times \prod \text{Likelihood} \right) \]
Hydrographic Health:

\[
\text{Hydrographic Health} = \left( \frac{\text{Desired Survey Score}}{\text{Present Survey Score}} \right)^2 \times \sum \left( \text{Consequence} \times \prod \text{(Likelihood)} \right)
\]

<table>
<thead>
<tr>
<th>HYDROGRAPHIC HEALTH (100^2\times...)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HH &gt;5'</td>
<td>Highest Hydro Need</td>
</tr>
<tr>
<td>(5’,5’)</td>
<td>20 nm²</td>
</tr>
<tr>
<td>(5’,5’)</td>
<td>230 nm²</td>
</tr>
<tr>
<td>(5’,5’)</td>
<td>700 nm²</td>
</tr>
<tr>
<td>(5’,5’)</td>
<td>2,300 nm²</td>
</tr>
<tr>
<td>(5’,5’)</td>
<td>4,900 nm²</td>
</tr>
<tr>
<td>(5’,5’)</td>
<td>10,700 nm²</td>
</tr>
<tr>
<td>(0, 5’)</td>
<td>Lowest Hydro Need</td>
</tr>
<tr>
<td>HH ≤ 0</td>
<td>Optimal Hydro Health</td>
</tr>
</tbody>
</table>
Caveats

• Not GIS Programmers
  • Backgrounds mainly in Hydrography and some basic GIS Analytics.

• Office transitioned to ESRI 3-4 years ago.

• Limited to what new technologies we could access with limited notice.
  • Implementation of new technologies can take years.

• We don’t know everything…..maybe people in this room have solutions that we did not come across
# Hydro Health Data Inputs

<table>
<thead>
<tr>
<th>Data type:</th>
<th>Purpose:</th>
<th>Source:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Where are the ships going?</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIS</td>
<td>Analyze traffic patterns</td>
<td>USCG, Marine Cadastre</td>
</tr>
<tr>
<td>AVIS database</td>
<td>Distinguish vessel types;</td>
<td>USCG</td>
</tr>
<tr>
<td></td>
<td>calculate under-keel clearance</td>
<td></td>
</tr>
<tr>
<td>Active Captain reports</td>
<td>Proxy for recreational boaters</td>
<td>Active Captain</td>
</tr>
<tr>
<td><strong>What’s the present state of hydrography?</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CATZOC</td>
<td>Initial data quality</td>
<td>NOAA</td>
</tr>
<tr>
<td>Survey Age</td>
<td>Present data quality</td>
<td>NOAA</td>
</tr>
<tr>
<td><strong>Is there a reason to be concerned with the state of hydrography?</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth</td>
<td>Identify areas of concern;</td>
<td>NOAA Charts</td>
</tr>
<tr>
<td></td>
<td>calculate seafloor complexity</td>
<td></td>
</tr>
<tr>
<td>Grounding reports</td>
<td>Identify areas of concern</td>
<td>USCG</td>
</tr>
<tr>
<td>Charted Hazards</td>
<td>Identify areas of concern</td>
<td>NOAA Charts</td>
</tr>
<tr>
<td>Reported Hazards (PA/PD)</td>
<td>Identify areas of concern</td>
<td>NOAA Charts</td>
</tr>
<tr>
<td>Natural/Artificial Reefs</td>
<td>Identify areas of concern</td>
<td>NOAA Charts</td>
</tr>
<tr>
<td>Sanctuaries</td>
<td>Identify areas of concern</td>
<td>NOAA</td>
</tr>
<tr>
<td>Ports</td>
<td>Identify areas of concern</td>
<td>NOAA, USACE</td>
</tr>
<tr>
<td>Search &amp; Rescue Stations</td>
<td>Identify areas of concern</td>
<td>USCG</td>
</tr>
<tr>
<td>Bottom Type</td>
<td>Identify areas of concern</td>
<td>NOAA Charts</td>
</tr>
<tr>
<td><strong>Is there a reason to believe the seafloor is changing?</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hurricanes</td>
<td>Identify areas of change</td>
<td>NOAA</td>
</tr>
<tr>
<td>Tidal Currents</td>
<td>Identify areas of change</td>
<td>Georgia Tech</td>
</tr>
<tr>
<td>Marine Debris</td>
<td>Identify areas of change</td>
<td>NOAA Charts</td>
</tr>
</tbody>
</table>
Model Builder / ArcGIS Processes

- Benefits
  - Repeatable
  - Transferable
  - Easy to learn

- Sanctuaries
- Ports
- Bottom Type
- AIS – Unique Count
Model Builder / ArcGIS Process Limitations

- Lack of one step processes (i.e. Unique Count inside a grid cell).
- Inability to iterate / slice quickly without caching and filling memory.
- Can’t see up front if physical memory will be maxed then determine a way forward.
- Put simply the needed process doesn’t exist (that we know of).
- Unable to troubleshoot geoprocessing failures more in depth.
Big Data – Big Problems

- AIS
  - UKC
    - Extract Values to Points*
      - Tolerable on small areas and limited features didn’t expand well
  - Unique Count
    - Spatial One-to-Many join
      - Segmented grids elevate a lot of the problems.

- Storms
  - Large Dataset
  - Lines → Polygons
    - Overlapping polygons as a function of another dataset
Missing tools

- Catzoc / Surdex
  - No supersession routine
- AIS
  - No one step unique counting
  - Method used may not even be intended for this purpose.
- PA/PD
  - Possibly distance vs density
Solution - Scripting via Python

• Under-Keel Clearance
  • UKC Extract Values to Points would not run on Cargo vessels in our limited FL area. This was foreshadowing our future scaling problems.
  • By using Python scripting, all UKC data successfully completed. Largest dataset (50 million+ points) ran through the UKC process in under two hours.

• Known Hazards
  • This dataset has multiple geometries that need to be analyzed and that was not possible with any tool we knew of. Developed new Python tool to work through this.
Solution - Scripting via Python

- Storms
  - Converted each storm (buffered polygon) to raster and then imported to numpy array
  - Gave ability to compute number of storms per grid cell and perform if/then statements to hash out latest storm. (If year of storm is greater than x then include, if not exclude)

- PA/PD
  - Esri solutions were originally built in Model Builder but scaling up to the entire US posed issues. Resulted to Python scripting to avoid small clipping and iterating manually in ArcGIS / ArcPro
Solution – ArcGIS Pro

• ArcGIS Pro drastically improved processing speeds and tool functionality
  • Merging large datasets
  • Clipping
  • Deleting columns

• ArcGIS Pro with Python
  • Able to quickly run scripts in the Python window
  • Utilizing Pros multithreading while running Python (when calling Pro tools from Python)
Scripting via Python Limitations

• One expert Python Programmer on site
  • Programmer Is not on our staff and often busy on other projects.
  • Although through this project one employee has become more skilled in Python and we are sending more to training.

• Major changes or troubleshooting also have to fall on programmer

• Development can take hours, days, or even weeks.
Future Work

• Currently expanding to the entire U.S.
• Build and implement communication plan
• In the process of submitting paper to International Hydrographic Review (IHR) for review and publication.
• Ground truth applicable datasets and equations
• Act on Results
• Refine the model based on results
• Build maintenance plan
Contact Info:
Patrick.keown@noaa.gov
NOAA NOS/OCS/HSD Operations