# 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

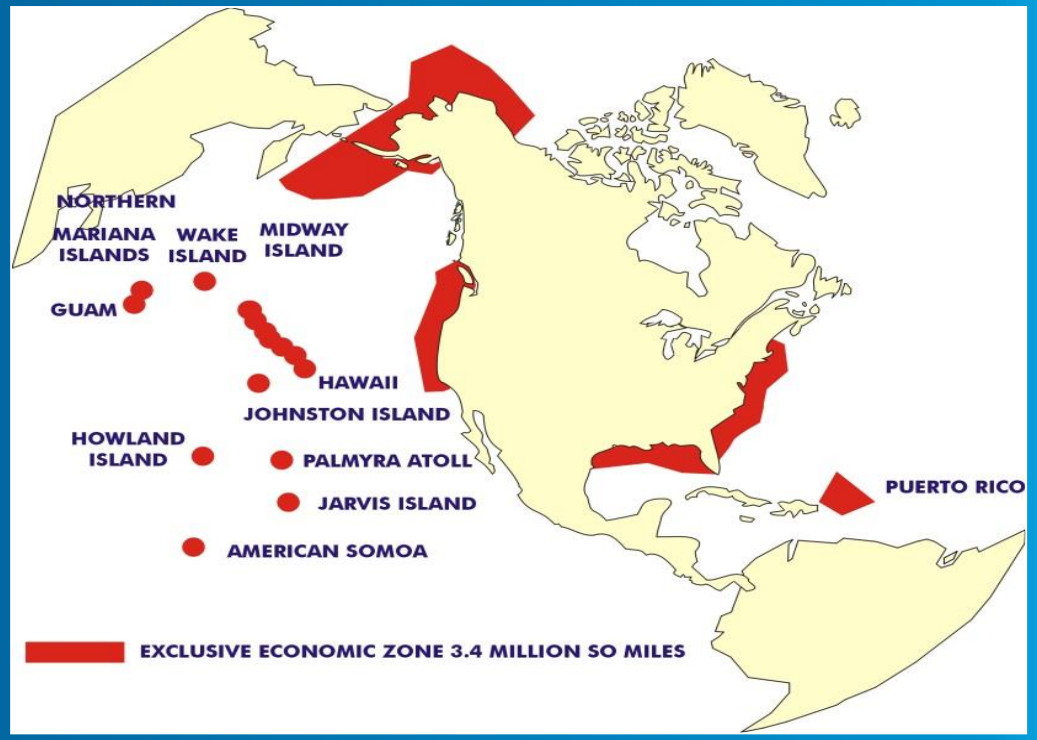
  } Hydrographic Gap
- GIS and the Hydrographic Health model
  - Big data, Big Problems
- Future Work

# Background – The Ocean is Big

- U.S. Exclusive Economic Zone (EEZ): 3,400,000 $nm^2$
- Surveyed to 'modern' standards: ~44,000 $nm^2$
- Average Hydrographic Acquisition: ~3,000 $nm^2$ / 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.

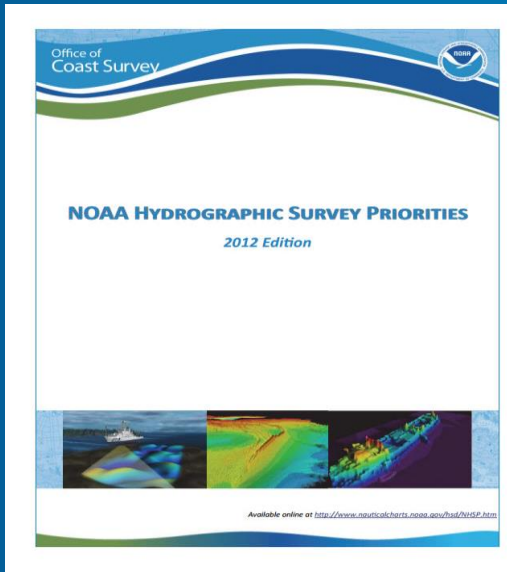


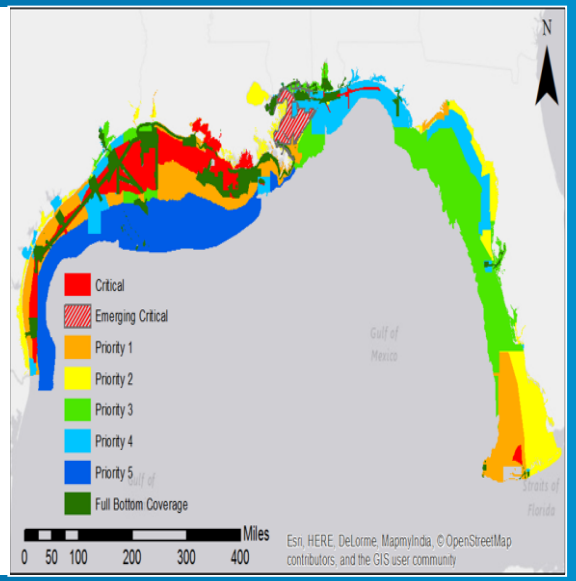
# National Hydrographic Survey Priorities (NHSP)

- NOAA produced National Hydrographic Survey Priorities (NHSP) document in 1994.
- 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

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

# A Model of Hydrographic Health

$$\text{Hydrographic Health} = \underbrace{\left(\text{Desired Survey Score} - \text{Present Survey Score}\right)^2}_{\textit{Hydrographic Gap}} \times \underbrace{\sum\left(\text{Consequence} \times \prod(\text{Likelihood})\right)}_{\textit{Hydrographic Risk}}$$

- 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.

$$\text{Hydrographic Health} = \left(\boxed{\text{Desired Survey Score}} - \text{Present Survey Score}\right)^2 \times \sum\left(\text{Consequence} \times \prod(\text{Likelihood})\right)$$

A combination of Seafloor Complexity and Under Keel Clearance will inform the desired Coverage Requirement

**2**

SEAFLOOR COMPLEXITY

Data: Standard Deviation of Seafloor Slope (within 5nm)

Std. Dev < 0.15 → "SIMPLE"

Std. Dev ≥ 0.15 → "COMPLEX"

**3**

DESIRED SURVEY SCORE

RANGE: 0 to 100

| | |
|---|---|
| DSS = 100 | Object Detection |
| DSS = 80 | Complete Coverage |
| DSS = 30 | Part. Bottom Coverage |
| DSS = 10 | Lesser Coverage |

**1**

DESIRED COVERAGE - DEFINED BY VESSEL TRAFFIC

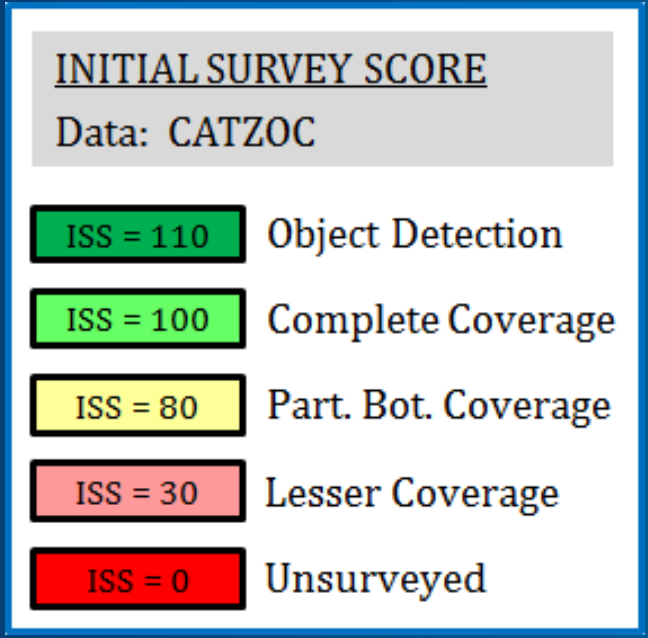
Data: For AIS records, compute UKC = Draft – Depth; if Draft is unknown, use Depth alone.

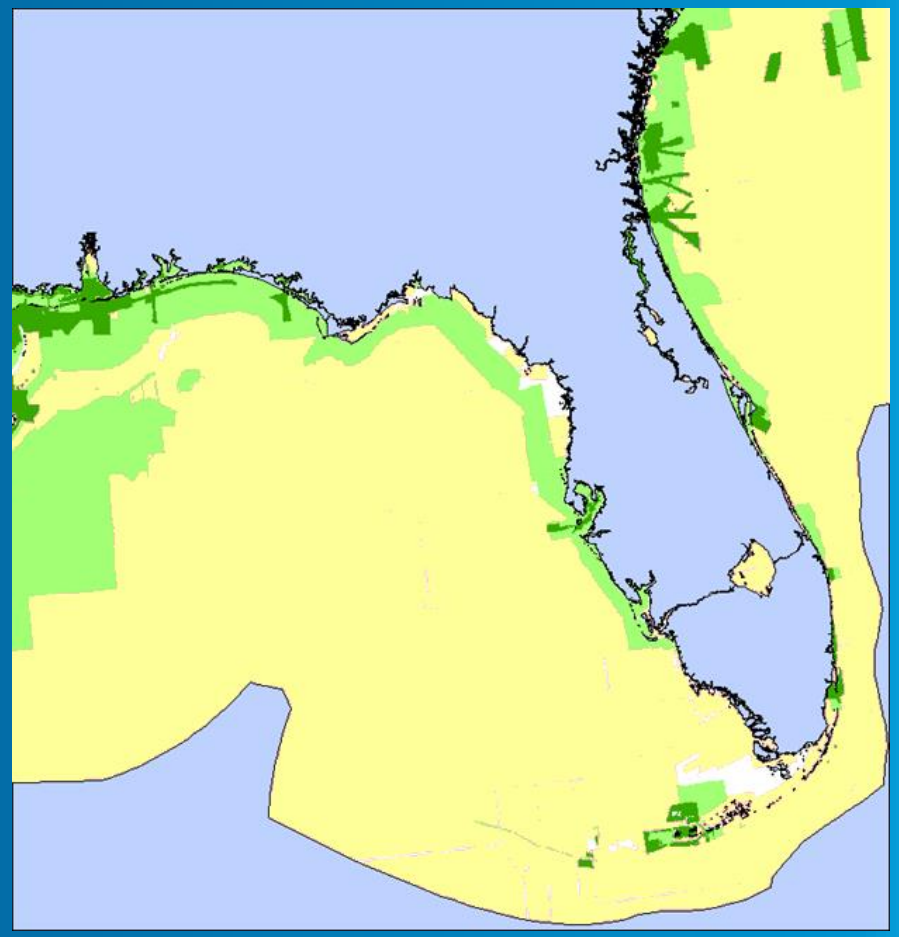
| If SIMPLE seafloor: | | If COMPLEX seafloor: | |
|---|---|---|---|
| DSS = 100 | UKC = [-1,1] | DSS = 100 | UKC = [-1,1] |
| DSS = 80 | UKC = [1,20m] | DSS = 80 | UKC = [1,40m] |
| | or Depth w/ AIS = [2,20m] | | or Depth w/ AIS = [2,40m] |
| DSS = 30 | UKC = [20,50m] | DSS = 30 | UKC = [20,100m] |
| | or Depth w/o AIS = [2,50m] | | or Depth w/o AIS = [2,100m] |
| DSS = 10 | UKC > 50m | DSS = 10 | UKC > 100m |
| | or Depth > 50m | | or Depth > 100m |

$$\text{Hydrographic Health} = \left(\text{Desired Survey Score} - \boxed{\text{Present Survey Score}}\right)^2 \times \sum\left(\text{Consequence} \times \prod(\text{Likelihood})\right)$$

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.



$$\text{Hydrographic Health} = \left(\text{Desired Survey Score} - \boxed{\text{Present Survey Score}}\right)^2 \times \sum\left(\text{Consequence} \times \prod(\text{Likelihood})\right)$$

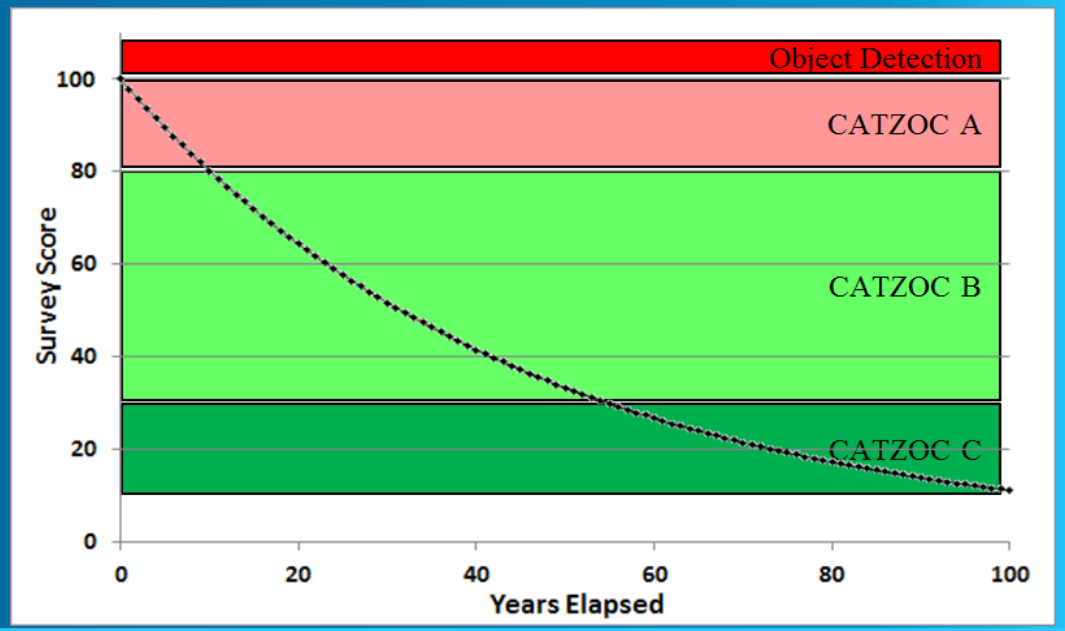
The Depreciation is modeled through exponential decay, where the Decay Coefficient, is the sum of several change parameters.

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

$$\text{Decay Coefficient} = \left(\Delta\text{Storms} + \Delta\text{Currents} + \Delta\text{Human Debris}\right) \times {}^{0.022}/_{4}$$

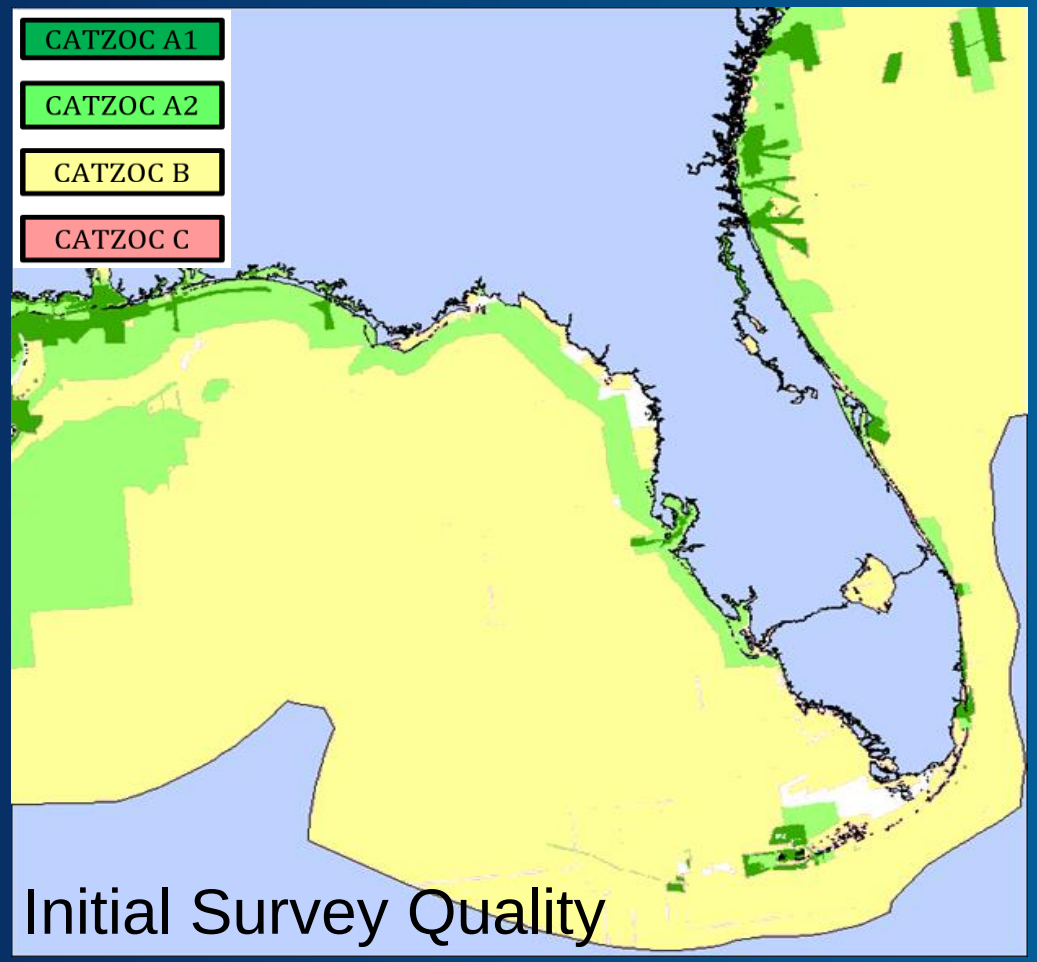
Where having one change term set to maximum decay will yield:

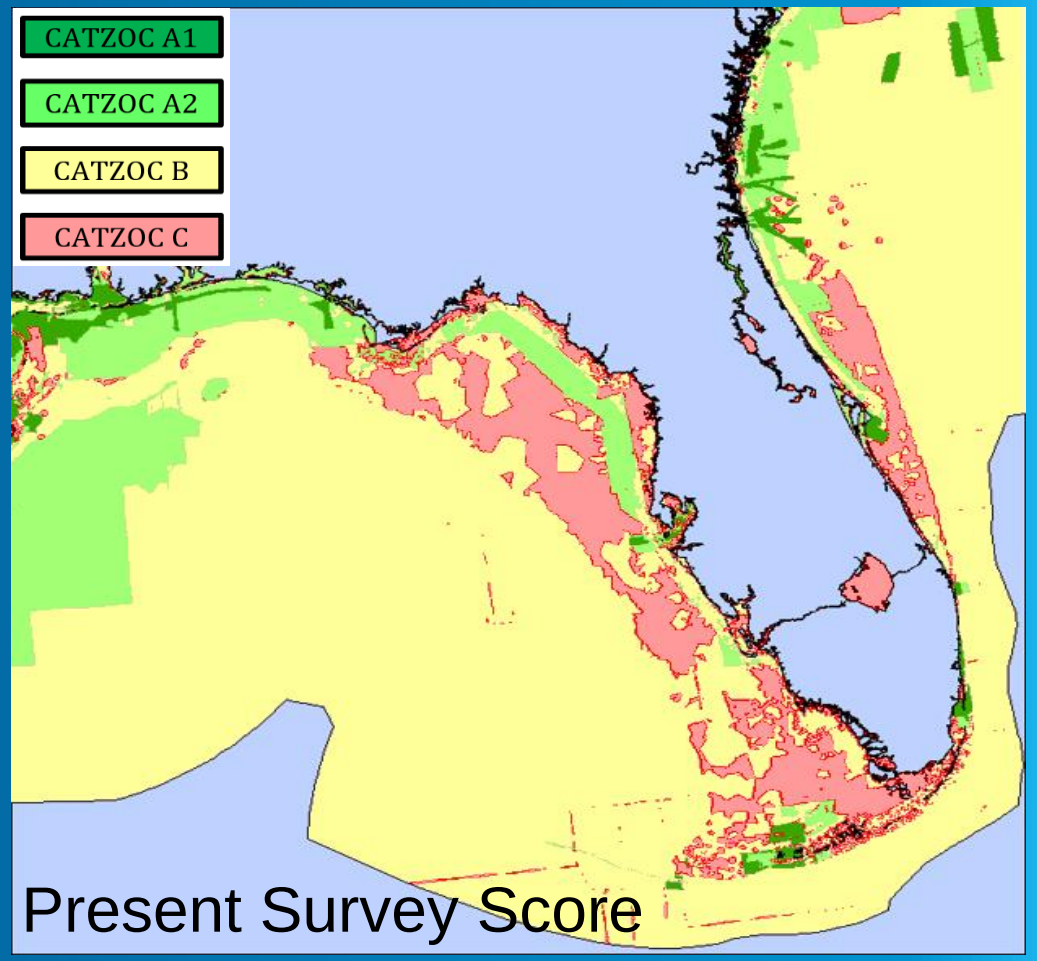
- CATZOC A → B in 10 years.
- CATZOC A → C in 55 years.
- CATZOC A → D in 100+ years.



$$\text{Hydrographic Health} = \left(\text{Desired Survey Score} - \boxed{\text{Present Survey Score}}\right)^2 \times \sum\left(\text{Consequence} \times \prod(\text{Likelihood})\right)$$

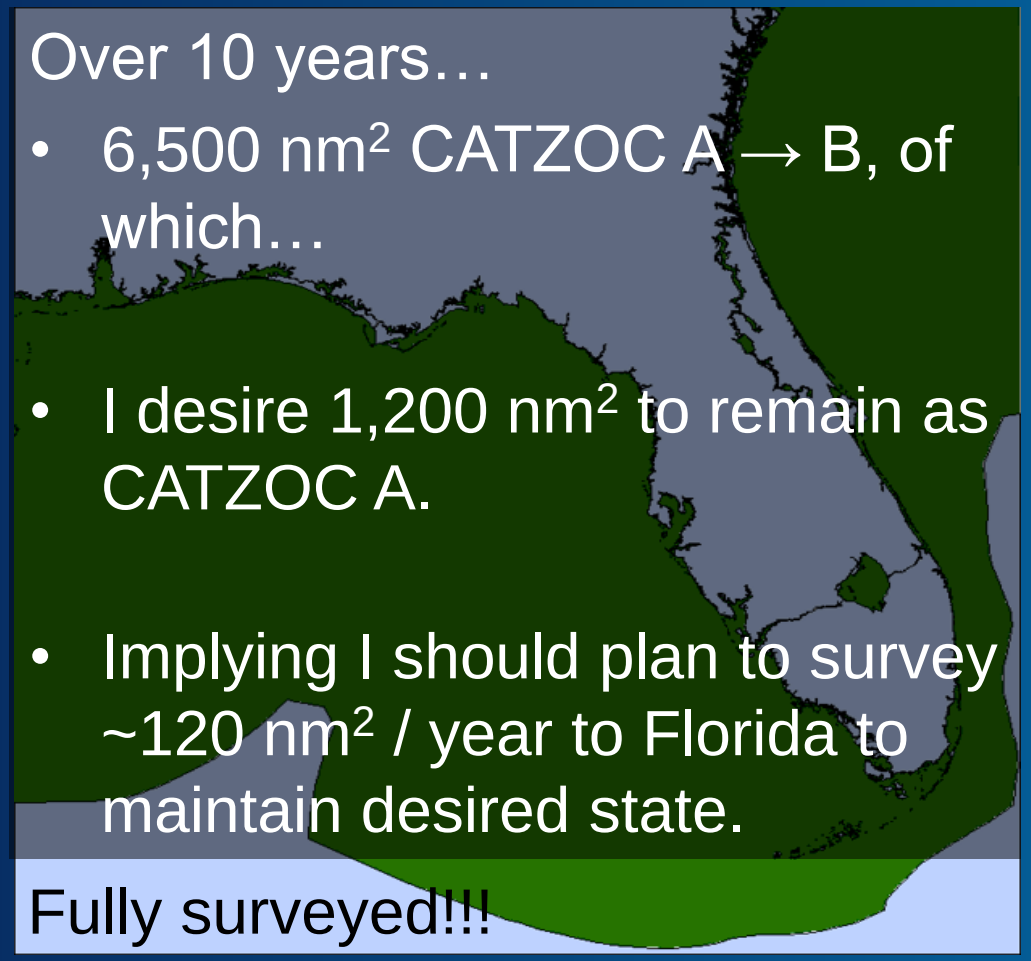
# All Depreciation Terms: Past to Present

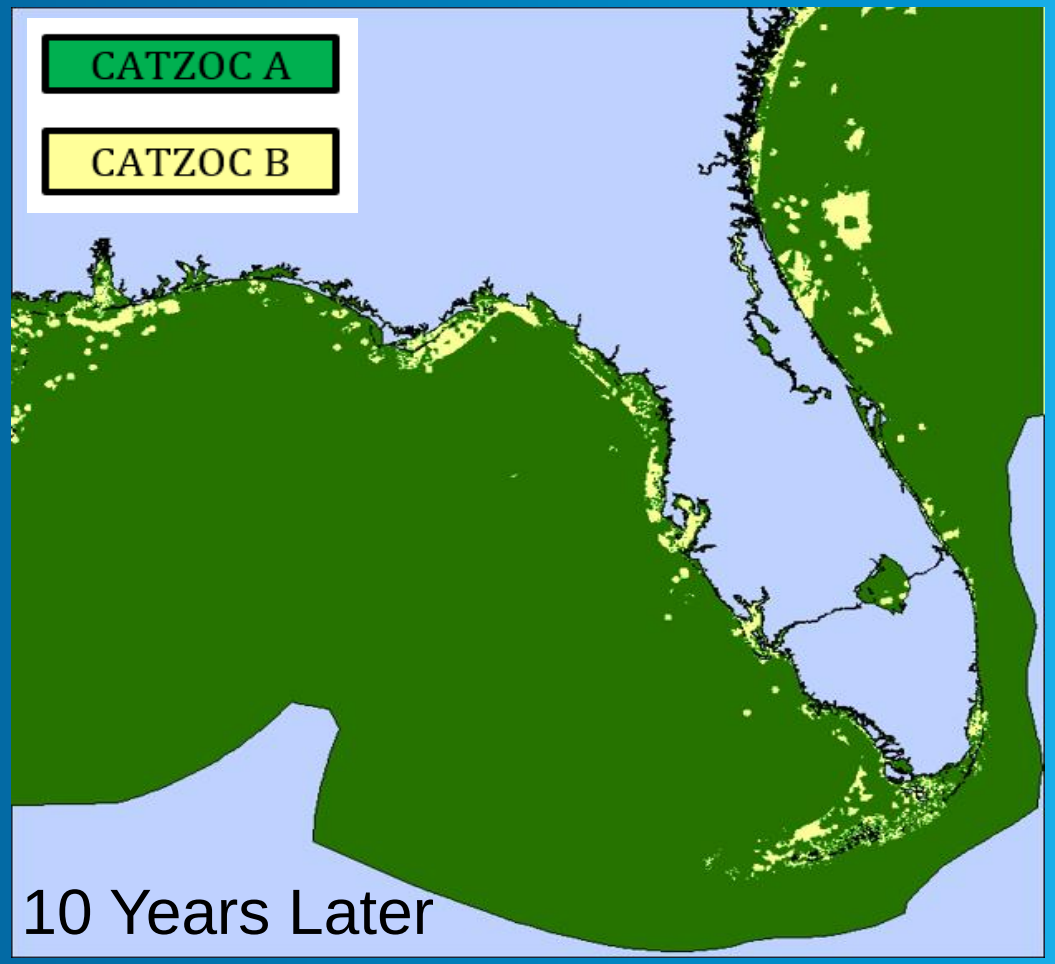


$$\text{Hydrographic Health} = \left(\text{Desired Survey Score} - \boxed{\text{Present Survey Score}}\right)^2 \times \sum\left(\text{Consequence} \times \prod(\text{Likelihood})\right)$$

# All Depreciation Terms: Past to Present.... (or Present to Future)



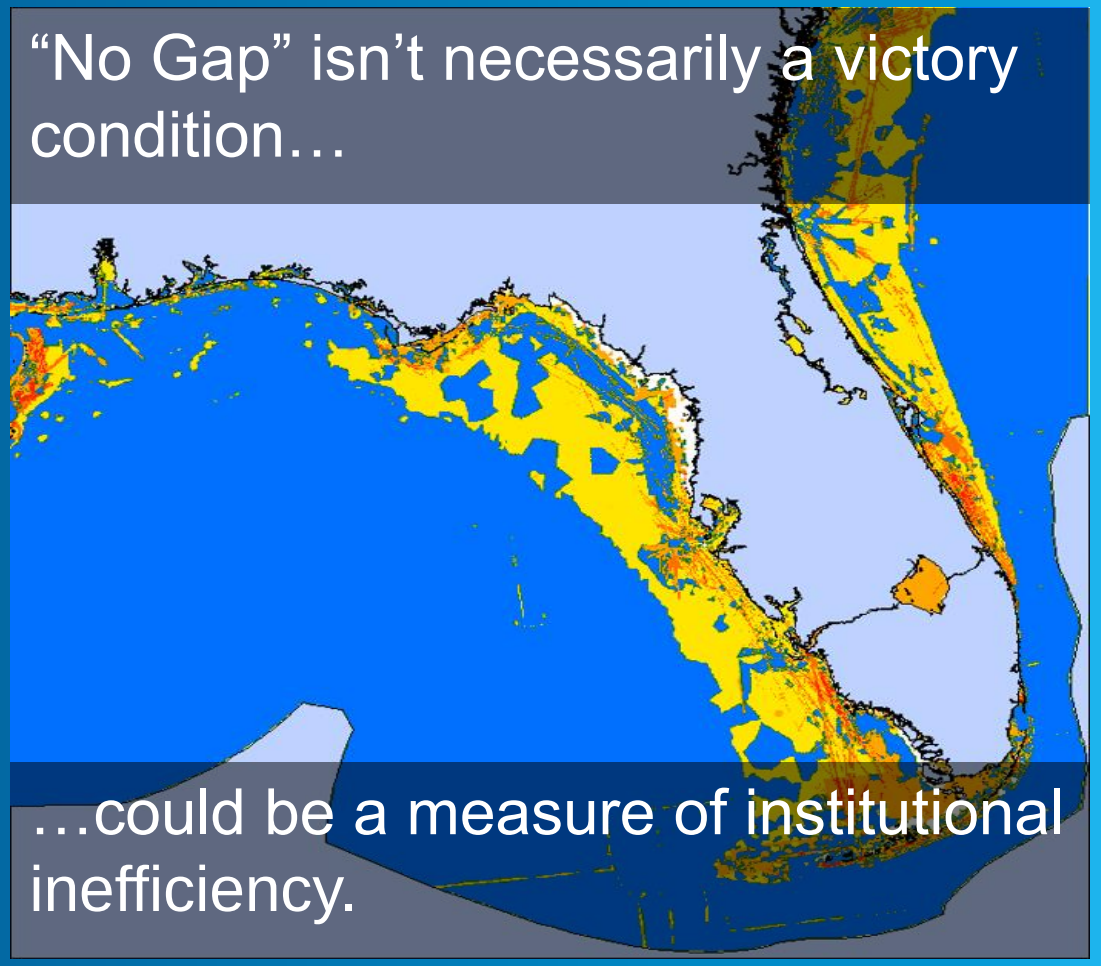
$$\text{Hydrographic Health} = \left(\text{Desired Survey Score} - \text{Present Survey Score}\right)^2 \times \sum\left(\text{Consequence} \times \prod(\text{Likelihood})\right)$$

# Hydrographic Gap: Desired – Present Survey Score

HYDROGRAPHIC GAP

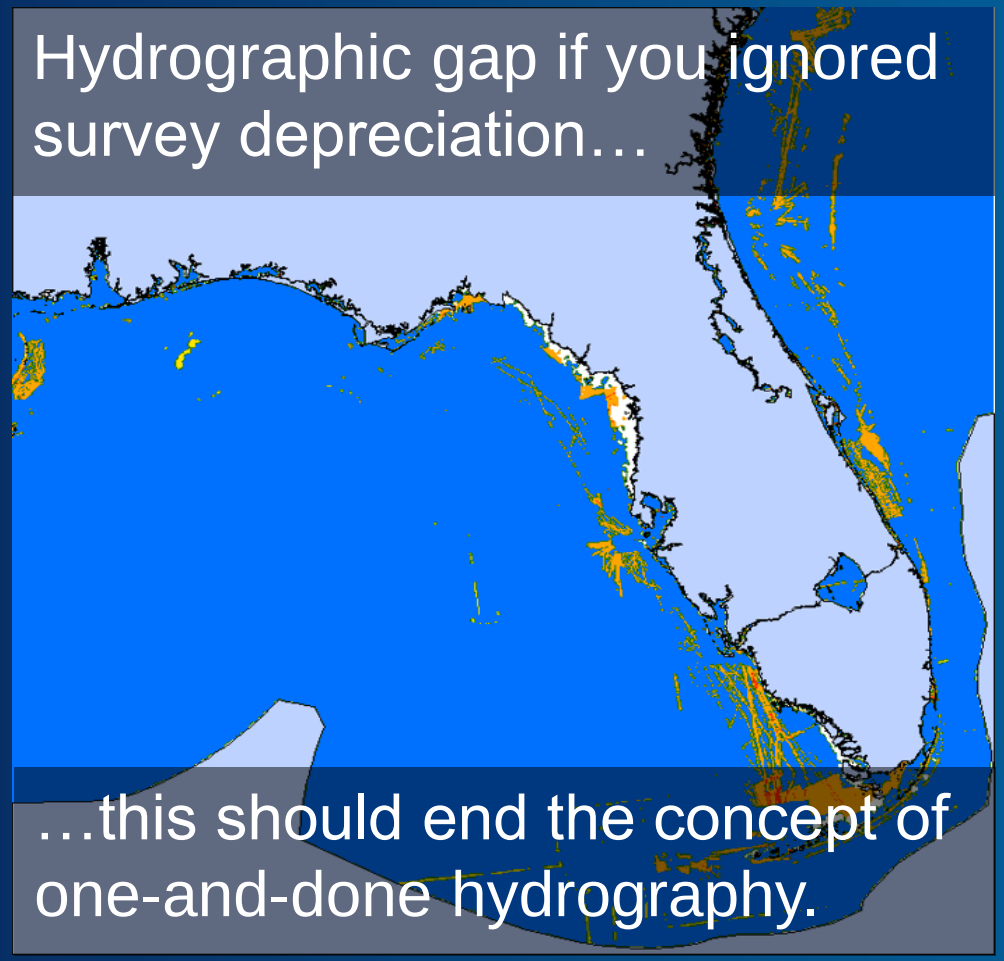
- HG < 0 — No Gap!
- (0,25) — Small Gap
- (25,50)
- (50,75)
- HG > 75 — Large Gap!

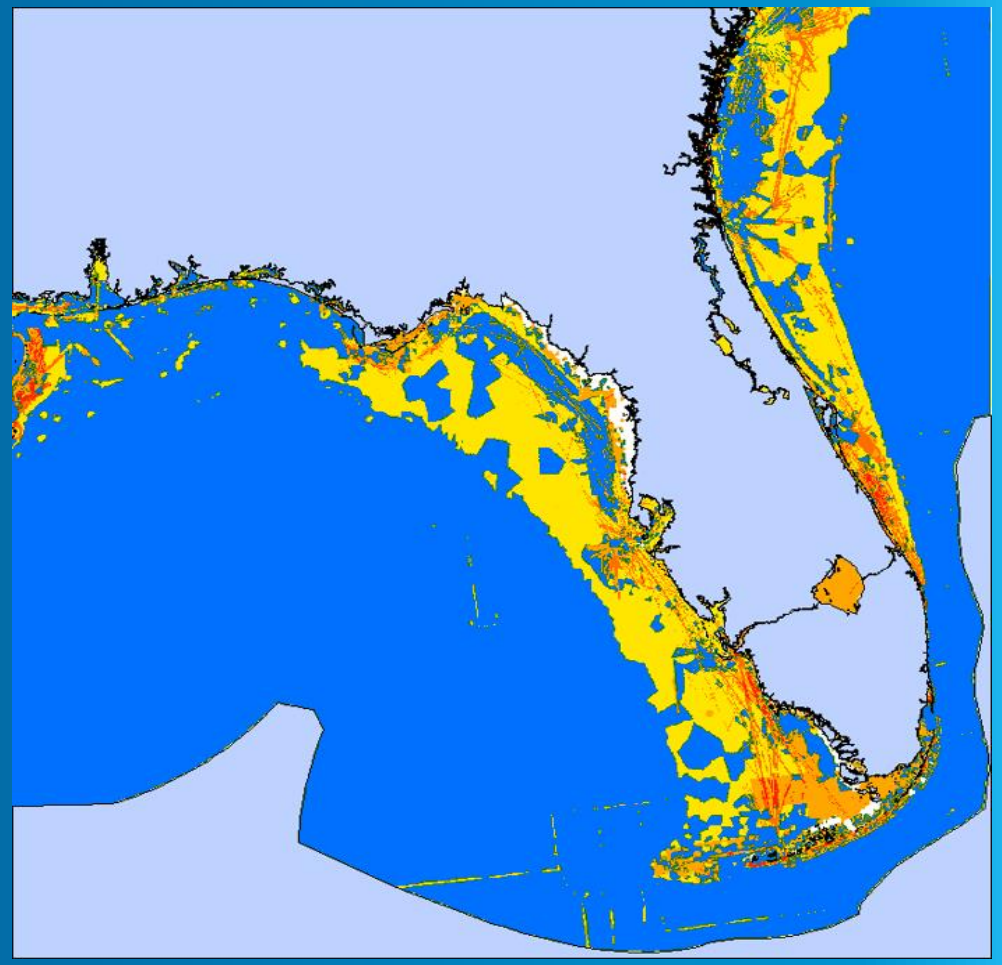
| Hydrographic Gap (Desire - Present) | |
|---|---|
| Magnitude of Gap | Sq. NM |
| No Gap! | 130,000 |
| 0 - 25 (small gap) | 19,000 |
| 25 - 50 | 6,300 |
| 50 - 75 | 2,500 |
| 75 - 100 (big gap) | 800 |



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

# Hydrographic Gap: Desired – Present Survey Score



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

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.

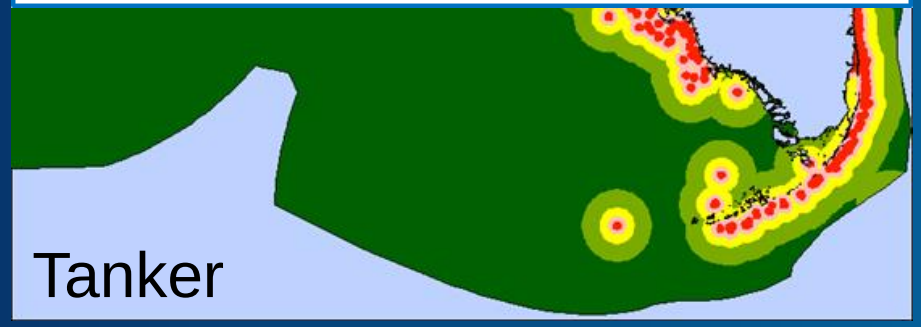
$$\text{Hydrographic Health} = \left(\text{Desired Survey Score} - \text{Present Survey Score}\right)^2 \times \sum \left(\text{Consequence} \times \prod(\text{Likelihood})\right)$$

# Consequence to Reefs (dependent on vessel type):

CONSEQUENCE: IMPACT TO REEFS & SANCTUARIES

Data: Distance to Reefs for Tankers versus other ships.

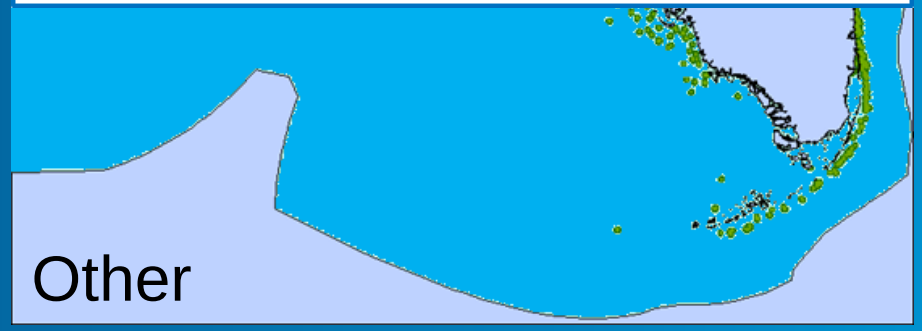
| If TANKER passes: | | If OTHER VESSEL passes: | |
|---|---|---|---|
| 5 | < 2.5 nm to reef | 5 | - |
| 4 | 2.5 – 5 nm | 4 | - |
| 3 | 5 – 10 nm | 3 | - |
| 2 | 10 – 20 nm | 2 | < 1.0 nm to reef |
| 1 | > 20 nm from reef | 1 | 1 – 2 nm |
| 0 | - | 0 | > 2 nm from reef |




CONSEQUENCE: IMPACT TO REEFS & SANCTUARIES

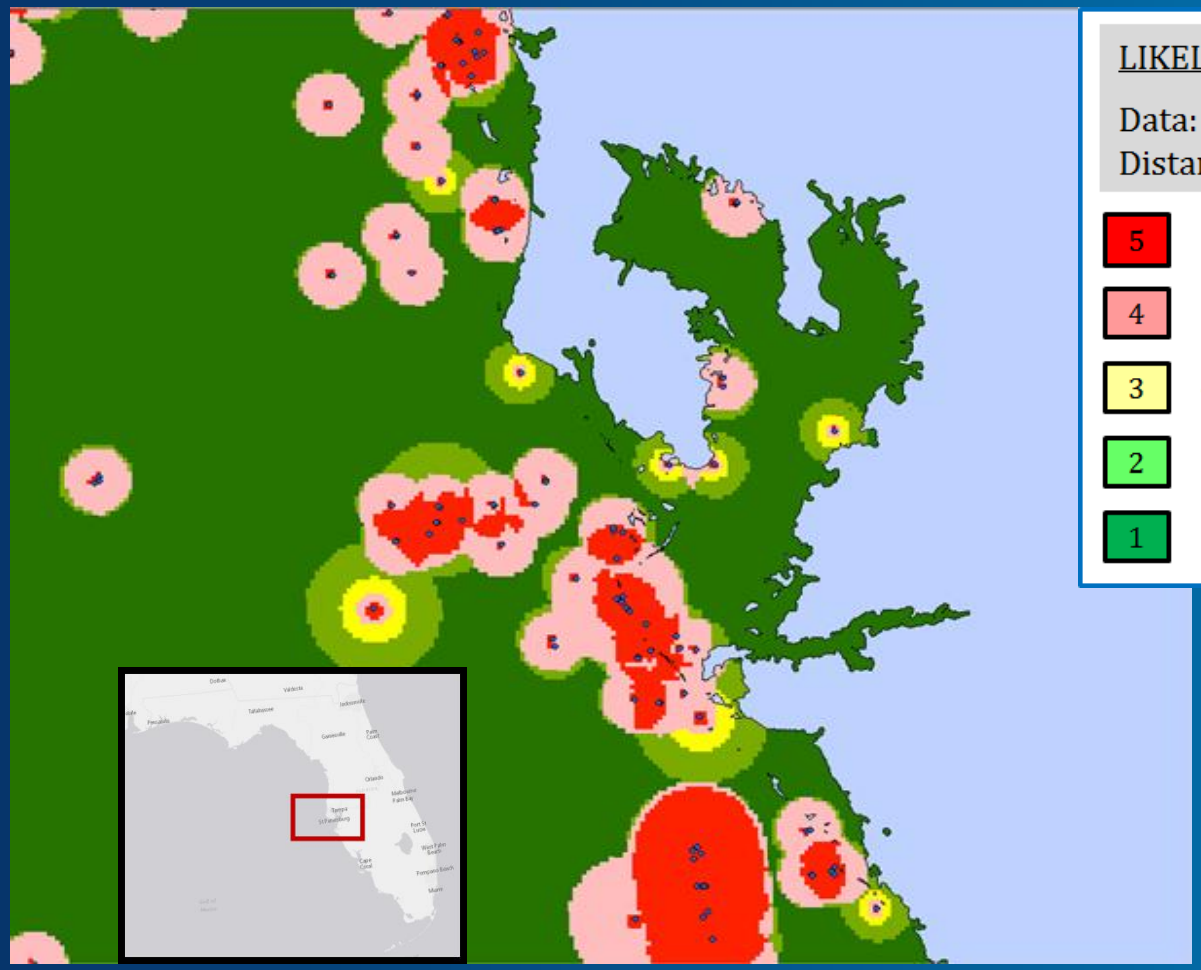
Data: Distance to Reefs for Tankers versus other ships.

| If TANKER passes: | | If OTHER VESSEL passes: | |
|---|---|---|---|
| 5 | < 2.5 nm to reef | 5 | - |
| 4 | 2.5 – 5 nm | 4 | - |
| 3 | 5 – 10 nm | 3 | - |
| 2 | 10 – 20 nm | 2 | < 1.0 nm to reef |
| 1 | > 20 nm from reef | 1 | 1 – 2 nm |
| 0 | - | 0 | > 2 nm from reef |



$$\text{Hydrographic Health} = \left(\text{Desired Survey Score} - \text{Present Survey Score}\right)^2 \times \sum\left(\text{Consequence} \times \prod \boxed{(\text{Likelihood})}\right)$$

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



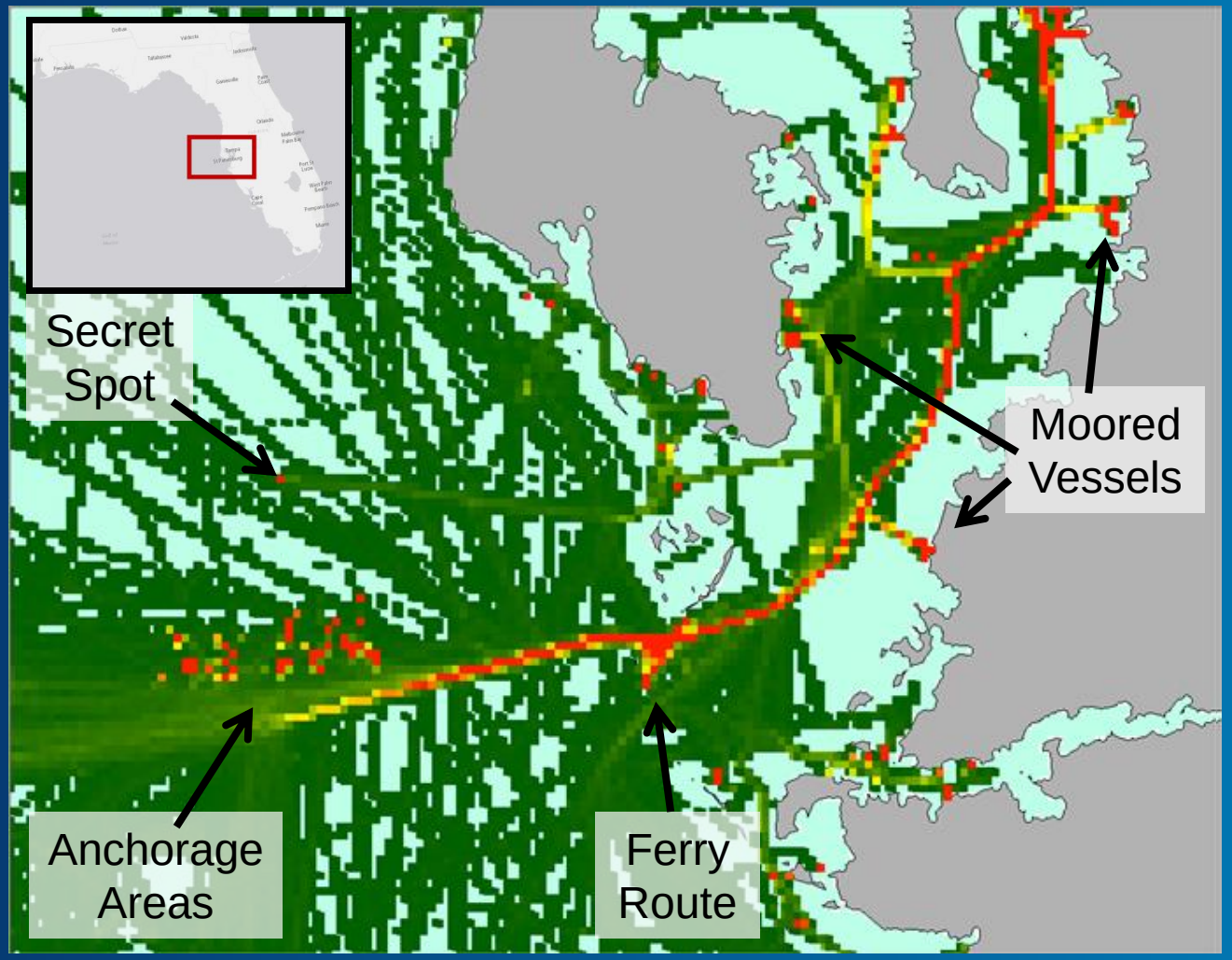
LIKELIHOOD: REPORTED ERRORS

Data: **MAX** of PA/PD Count (within 2nm) plus "Rep" (within 4nm), or Distance to nearest PA/PD, or Distance to nearest "Rep."

| | | | | | |
|---|---|---|---|---|---|
| 5 | Count of 4+ | 5 | <0.25 nm to PA/PD | 5 | <0.50 nm to "Rep" |
| 4 | Count of 2-4 | 4 | 0.25 – 0.50 nm | 4 | 0.50 – 1.0 nm |
| 3 | - | 3 | 0.50 – 1.0 nm | 3 | 1 - 2 nm |
| 2 | Count of 1 | 2 | 1 – 2 nm | 2 | 2 - 4 nm |
| 1 | No Errors | 1 | >2 nm to PA/PD | 1 | >4 nm to "Rep" |

$$\text{Hydrographic Health} = \left(\text{Desired Survey Score} - \text{Present Survey Score}\right)^2 \times \sum\left(\text{Consequence} \times \prod \boxed{(\text{Likelihood})}\right)$$

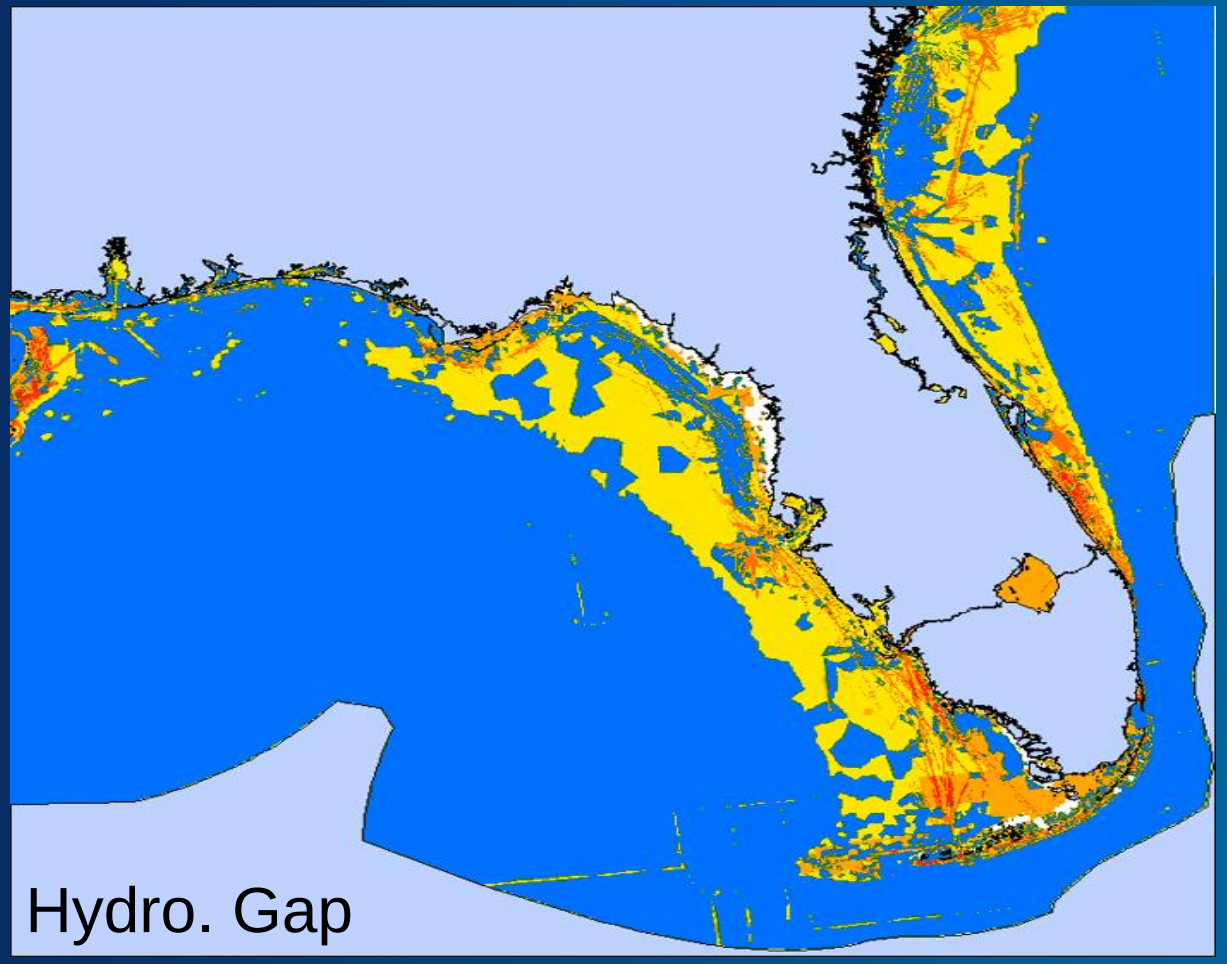
# Likelihood due to Traffic density (Unique vessel count):

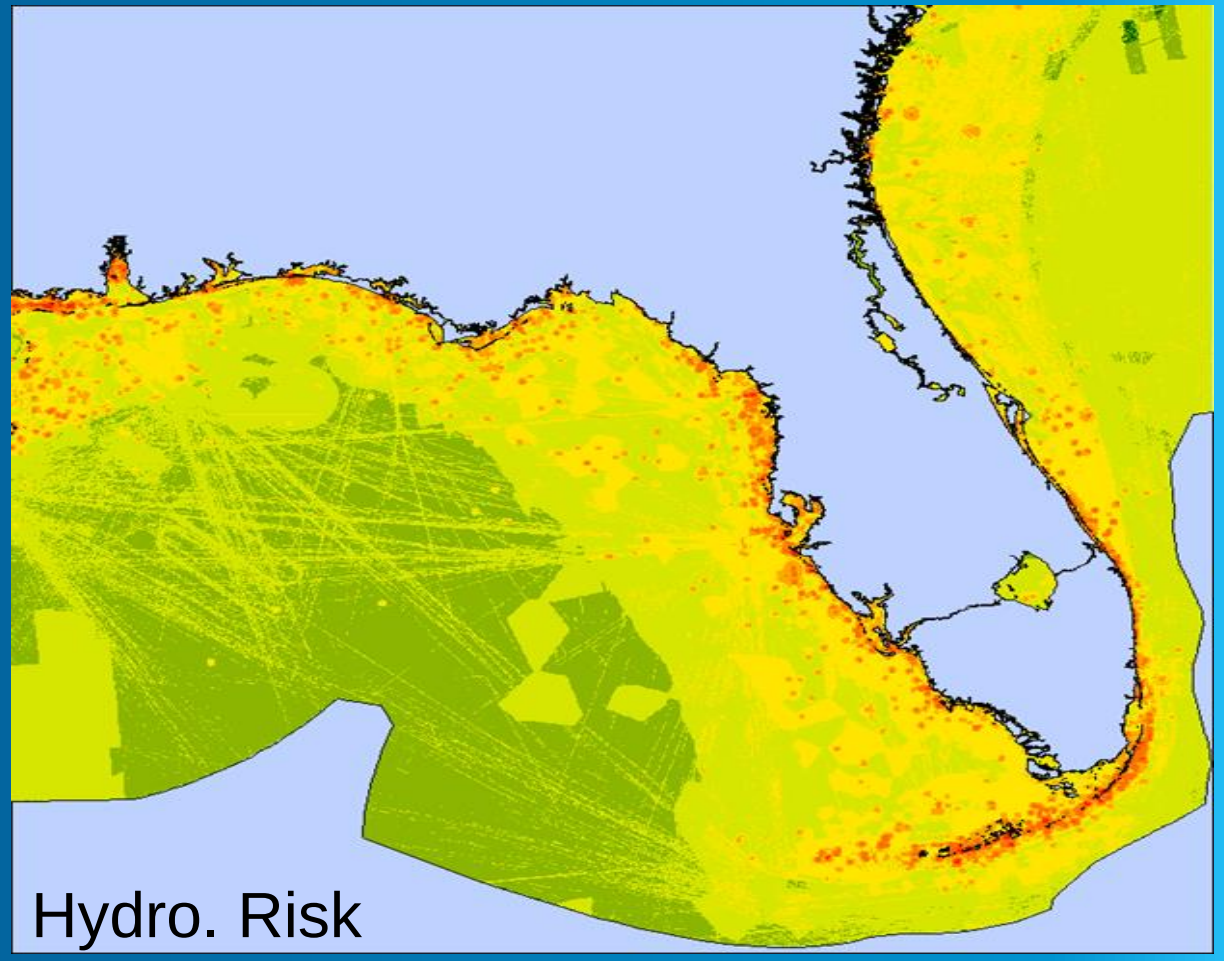



**LIKELIHOOD: TRAFFIC DENSITY**

| | |
|---|---|
| 5 | 100+ Boats |
| 4 | 11 – 100 Boats |
| 3 | 2 – 10 Boats |
| 2 | 1 Boat |
| 1 | No AIS |

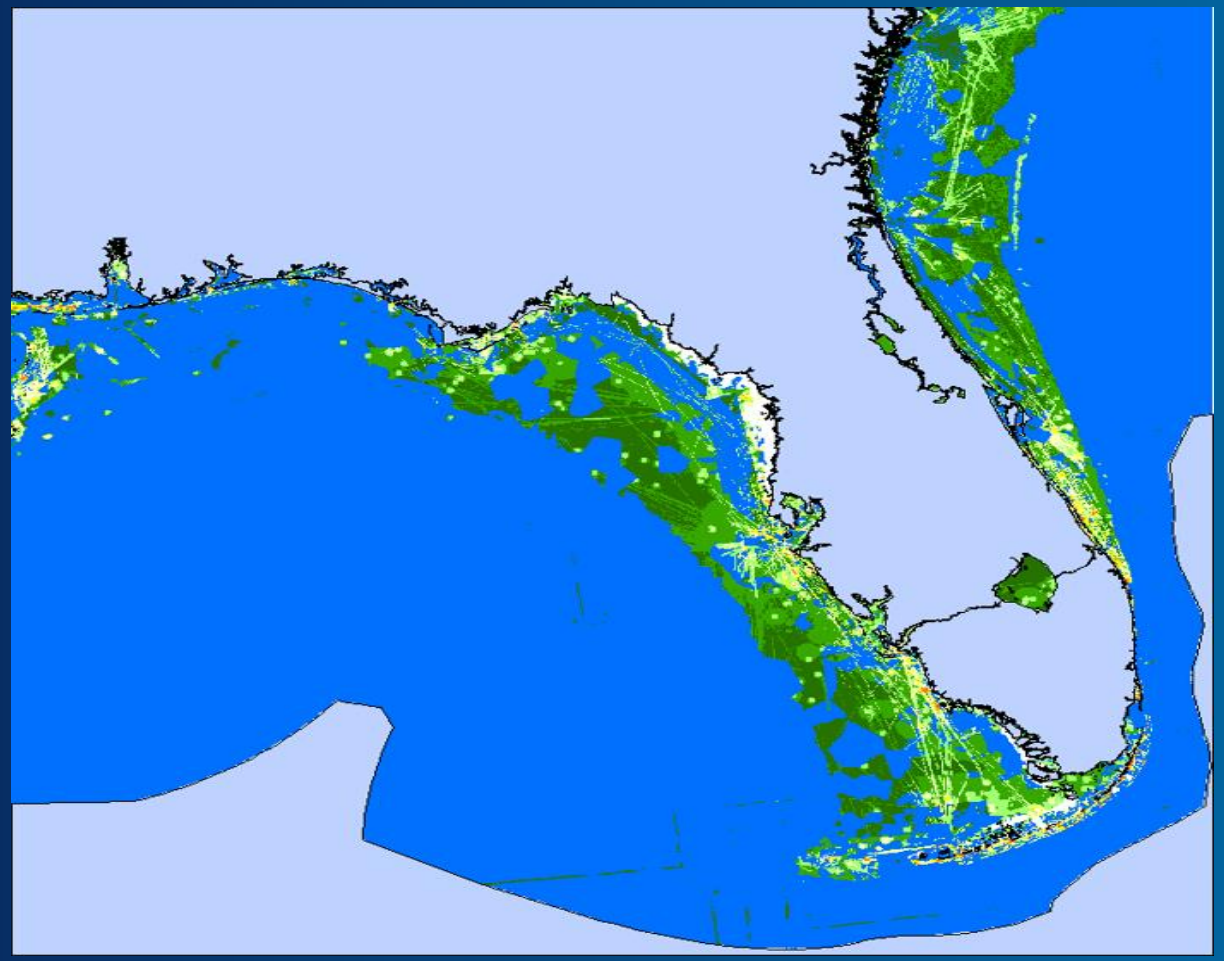
$$\text{Hydrographic Health} = \underbrace{\left(\text{Desired Survey Score} - \text{Present Survey Score}\right)^2}_{\text{Hydrographic Gap}} \times \underbrace{\sum\left(\text{Consequence} \times \prod(\text{Likelihood})\right)}_{\text{Hydrographic Risk}}$$



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

# Hydrographic Health:



| HYDROGRAPHIC HEALTH ($100^2 \times \ldots$) | | | |
|---|---|---|---|
| HH $>5^7$ | Highest Hydro Need | 1 $nm^2$ | 0.0% |
| $(5^6,5^7)$ | | 20 $nm^2$ | 0.0% |
| $(5^5,5^6)$ | | 230 $nm^2$ | 0.1% |
| $(5^4,5^5)$ | | 700 $nm^2$ | 0.4% |
| $(5^3,5^4)$ | | 2,300 $nm^2$ | 1.4% |
| $(5^2,5^3)$ | | 4,900 $nm^2$ | 3.1% |
| $(5^1,5^2)$ | | 10,700 $nm^2$ | 6.8% |
| $(0, 5^1)$ | Lowest Hydro Need | 9,600 $nm^2$ | 6.1% |
| HH $\leq 0$ | Optimal Hydro Health | 130,000 $nm^2$ | 82.0% |

# 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

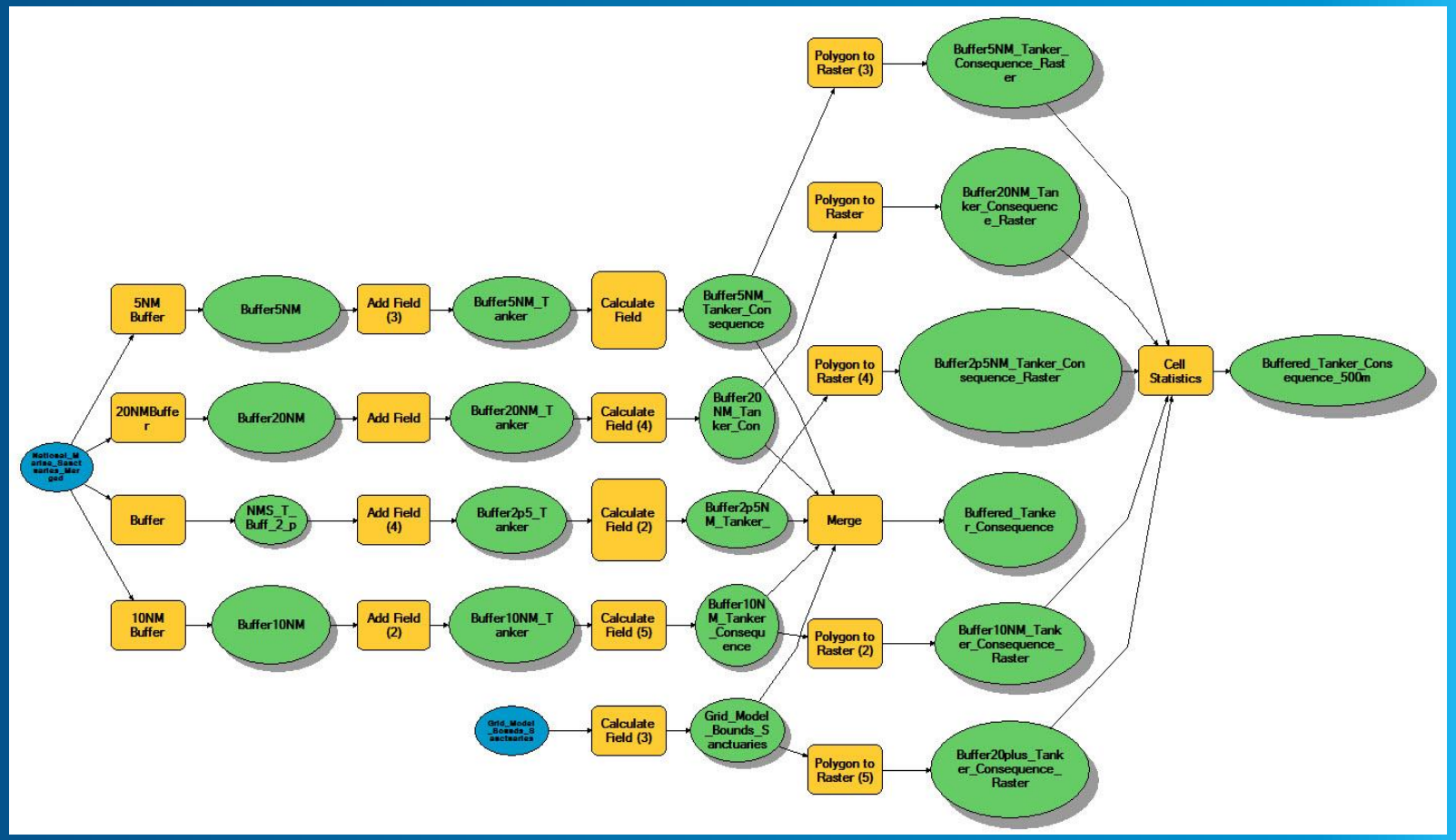
# Hydro Health Data Inputs

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

# 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

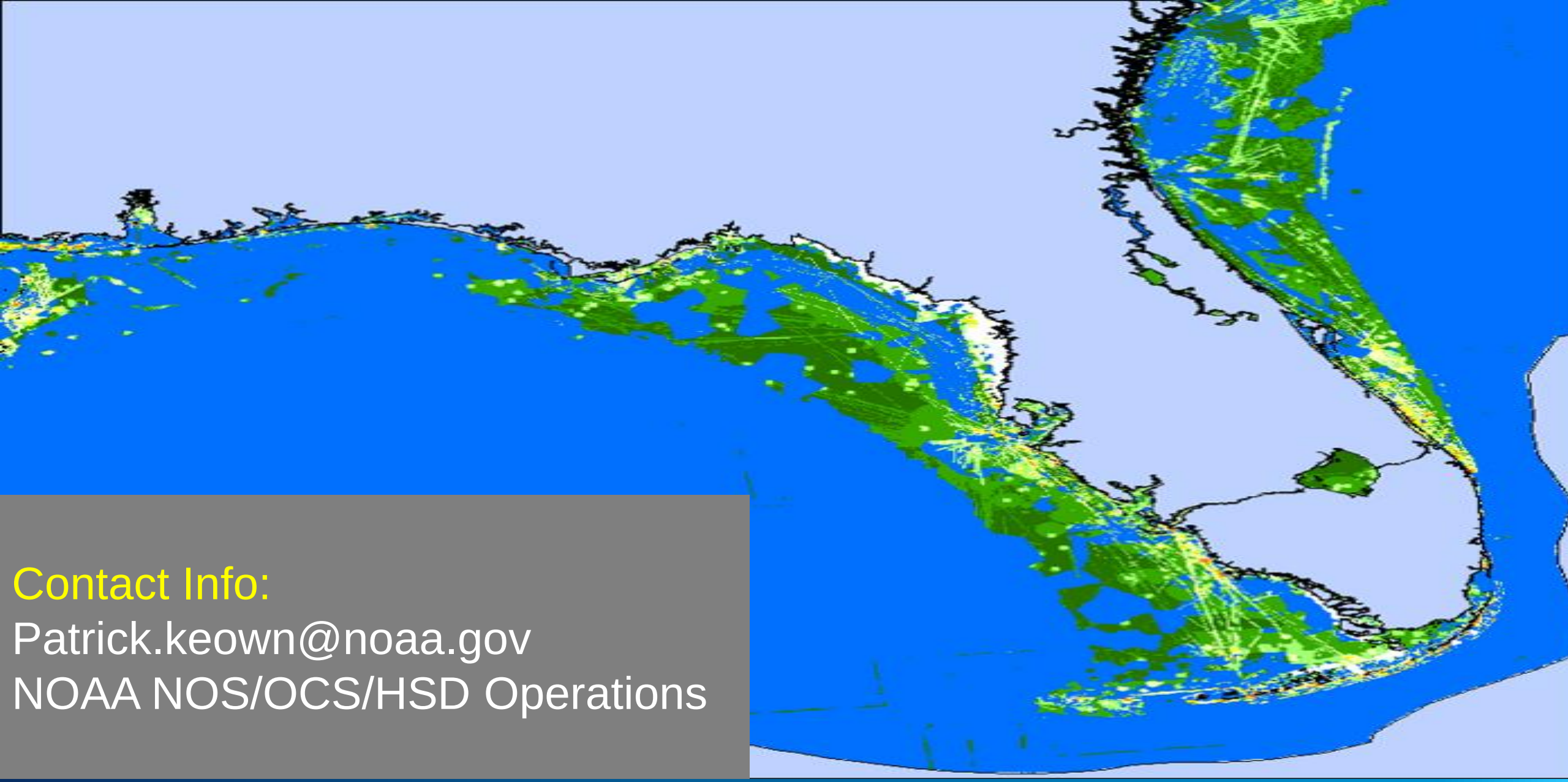
- ArcGIS Pro with Python
  - Able to quickly run scripts in the Python window
  - Utilizing Pros multithreading while running Python (when calling Pro tools from Python)

# 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