Are we done yet? An empirical estimator of level of effort for seafloor surveys - including an estimate for the full survey of US Waters

Samuel F. Greenaway[1], Amber Batts[1,2], and Jack Riley[1]

1. *NOAA Office of Coast Survey Development Lab, Hydrographic Systems & Technology Branch, Silver Spring, Maryland, USA, samuel.greenaway@noaa.gov*
2. *Earth Resource Technologies, Silver Spring, Maryland, USA*

Abstract

An estimate of the effort needed to survey some arbitrary area is a critical part of the planning efforts of any hydrographic office. We develop a simple, analytic model to estimate full coverage of an arbitrary area based on a fixed angular swath system such as a multibeam echosounder. This model incorporates one tunable parameter to account for the overall efficiency of survey execution. We used over 500 recent hydrographic surveys to validate this model and develop an estimate for this tunable parameter. We had expected this parameter to be strongly tied to seafloor complexity and thus regionally consistent; it was not. In fact, we could discern no strong relationship between this parameter and any variable investigated, including region, roughness, variability, depth, or survey size. We use this tuned model, including an estimate of uncertainty, to develop a model for survey effort, and apply the model to all of the U.S. waters. We find that a complete re-survey of U.S. waters of would take 14 M LNM (linear nautical miles) of survey, or approximately 216 years for a single platform running continuously at typical survey speeds. Accounting for areas already surveyed to modern standards, we calculate that we have surveyed 44% of the U.S. waters to modern standards by area, but only 18% by level of effort. To survey the remaining area to modern standards would take 12M LNM of survey, or approximately 177 years of a single platform running continuously at typical survey speeds.

Introduction

NOAA's Office of Coast Survey has been diligently surveying U.S. waters since originally directed by President Thomas Jefferson in 1807. A reasonable question is "When are we going to be done?" To develop an estimate towards this end, we approach the problem in stages: (1) create a simple model for modern full bottom coverage survey, (2) use over 500 recent surveys to empirically tune this model, and (3) develop an associated uncertainty model, and then (4) run the model for the entire waters of the U.S. EEZ.

We had in mind a few uses for this model. The estimation task was primary, but we also looked towards using it to both measure both the output and efficiency of an executed survey. For an output metric, we would like a measure that reasonably scales to effort but does not depend on the specifics of how a given survey was actually executed. For efficiency, we thought that if we had an effective model for expected level of effort for a given survey area, comparison with the model might lead to a metric of survey efficiency. Besides the obvious performance-monitoring functions of such a metric, this might also allow us to compare different approaches to a survey, e.g. line planning vs. free-form coverage, without having to resurvey a particular area to investigate the impact the tested parameter. We found that the underlying variability of a given survey from the model was likely too large to make the efficiency metric much use, though we do set clear thresholds to establish efficacy of any such treatment. On the other hand, using the ideas developed here for a better metric for tracking survey work is compelling; this theoretical linear distance metric scales appropriately with survey effort (like actual LNM), but is also insensitive to the actual survey execution (like the total area covered).

Section 1 - A Model

A fundamental feature of many modern multibeam echosounders is the effective angular swath width fixed for a given depth, typically to near 60 degrees from nadir on both sides. For a flat seafloor, the area surveyed, or covered, by one line of is according to the simple trigonometric equation of length (*l*) multiplied by width (w):

$$A = l * w = l * 2d \tan\theta \qquad (1)$$

Variable *A* defines the area surveyed, *l* is the length of the survey line, ***d*** is the water depth, and *Θ* is the effective swath angle. Rearranging and writing in a differential form:

$$\Delta l = \frac{\Delta A}{2d \tan\theta} \qquad (2)$$

This gives the length of survey line segment (Δl) required to survey some parcel (ΔA) of constant depth (d). This formulation is well suited to the raster model of bathymetry, where we define the depth at discrete node points, and associate each node (i) with a defined area per the resolution of the raster model. Summing across the surface nodes yields an estimate of the linear survey length required to cover the area:

$$L_t = \sum \Delta l = \sum_i \frac{\Delta A_i}{2d_i \tan\theta} \qquad (3)$$

Equation (3) represents the theoretical minimum linear survey distance required to survey some area with a system of fixed angular swath. To accommodate the practical considerations of executing a survey, we introduce an efficiency parameter (α) as the ratio of the actual total length of survey lines necessary to our modeled theoretical minimum value:

$$\alpha \equiv \frac{L_a}{L_t} \qquad (4)$$

So our model becomes:

$$L_a = \alpha \sum \Delta l = \alpha \sum_i \frac{\Delta A_i}{2d_i \tan\theta} \qquad (5)$$

Higher values of α reflect a less efficient survey, either due to the constraints of the survey areas, particular equipment in use, or operational inefficiency. For our model, we have assumed a fixed swath width of 60 degrees from nadir. We could equivalently incorporate this notion of efficiency into an idea of an effective swath width where:

$$\frac{\alpha}{\tan(60°)} = \frac{1}{\tan(\theta_{\text{eff}})} \qquad (6)$$

An Estimate for the Efficiency Parameter

To validate our model and derive an estimate for the efficiency parameter, we looked at 547 recent multibeam surveys submitted to NOAA's Hydrographic Survey Division as full 'H' surveys intended to comply with the NOAA specifications and deliverables. Each of these surveys has an associated raster BAG file (Calder, et al. 2005) with final depth estimates, as well as a Descriptive Report that includes survey effort linear metrics. These surveys were conducted by a variety of field units, both government and contract, in a variety of regions, across a range of bottom types, and in a variety water depths. We excluded surveys from consideration that used side-scan or single-beam methods for coverage, focusing entirely on full-bottom coverage multibeam surveys for this work. We also excluded surveys from unique platforms like the NOAA Ship *Ferdinand Hassler*. The *Hassler's* widely spaced, dual-head configuration results in uniquely high efficiency by this metric.

For each survey, we extracted the total reported linear distance of echosounding for each survey from the Descriptive Report. We also calculated the theoretical total linear survey length by calculating [eq. (3)] across the BAG and calculated the efficiency parameter by dividing the two.

We originally thought that the efficiency parameter would be strongly influenced by the characteristics of the local bathymetry, with high efficiency (α close to 1) in areas with relatively flat, low relief seafloor, and low efficiency (α>>1) in areas with steep bathymetry and indented seashore. We had thought that in the former case, a survey crew could plan and execute an orderly series of survey lines with optimal overlap along each line. Conversely, in areas with complex seafloor and complicated shoreline, survey lines would necessarily overlap in an inefficient way as the crew maneuvered to achieve full coverage. Because the character of the bathymetry and shoreline are, broadly speaking, regionally consistent, we figured that the efficiency parameter would be as well. It was not.

RESULTS

The calculated efficiency parameter for each survey is plotted in Figure 1.

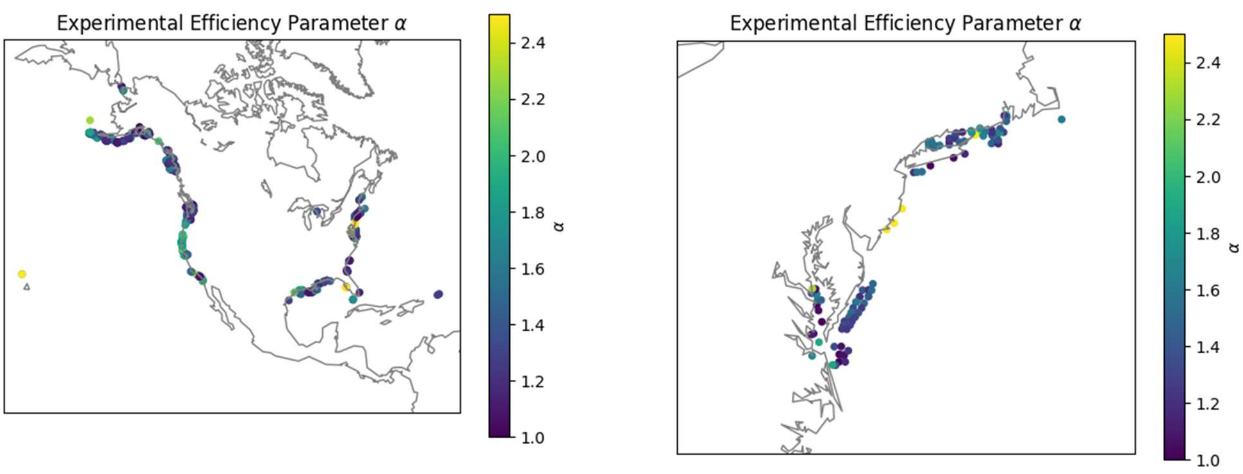

Figure 1: Efficiency parameter for all surveys (left) and detail for Mid-Atlantic Coast (right)

Effect of Region

To test our hypothesis that the survey efficiency would be strongly dependent on locale, we divided the results into eight regions (Table 1). For all regions, the efficiency parameter ranges from 1.29 in the Great Lakes to 2.83 in Hawaii. Excluding regions with less than ten surveys, the range is much tighter, 1.39 in the Atlantic to 1.59 in New England. A one way analysis of variance (ANOVA) showed a significant effect of region [$F(7, 546) = 7.48$, $p = 1.3\times10^{-8}$]. We then used the *post hoc* Tukey-Kramer method to evaluate the statistical significance of the region effect (Table 2). This showed only Hawaii and the West Coast belong to significantly different groups ($p<0.05$). Hawaii is significantly different from all regions, though the sample size was very small (N=2). The West Coast is significantly different from the Atlantic, the Gulf, and Alaska, as well as Hawaii. No other regions are statistically distinct from each other; 18 of the 28 regional-pairs are statistically indistinguishable.

Table 1: Efficiency parameter by Region

Region	Count	Average α	Variance
New England	27	1.59	0.11
Atlantic	79	1.39	0.11
Caribbean	6	1.36	0.05
Gulf	71	1.41	0.13
Great Lakes	2	1.29	0.01
West Coast	125	1.58	0.17
Alaska	235	1.45	0.10
Hawaii	2	2.83	1.84

Table 2: Results from Tukey-Kramer Test for the effect of region on the efficiency parameter. 'T' indicates the pairwise comparison are significantly (p<0.05) different, F indicates they are not. Hawaii and West Coast are the only regions with significant difference from the rest.

	New England	Atlantic	Caribbean	Gulf	Great Lakes	West Coast	Alaska	Hawaii
New England		F	F	F	F	F	F	T
Atlantic			F	F	F	T	F	T
Caribbean				F	F	F	F	T
Gulf					F	T	F	T
Great Lakes						F	F	T
West Coast							T	T
Alaska								T
Hawaii								

Effect of Unit

We also looked at the effect of the survey unit on the efficiency parameter. For this analysis, we treat a particular government survey ship (e.g. the NOAA Ship *Rainier*) or a contracted survey firm as a unit. There were 23 different units in the overall sample, but only eight with more than 10 surveys (Table 3). For the units with more than ten surveys, the average efficiency parameter ranged from 1.31 to 1.71. A one-way analysis of variance showed a significant effect of unit [$F(7, 488) = 14.6$, $p = 2.6 \times 10^{-8}$]. We then used the *post hoc* Tukey-Kramer method to evaluate the statistical significance of the unit effect. The two units with the highest α were not statistically distinct from each other. These two units, in turn, were statistically distinct ($p<0.05$) from all but one other unit. The remaining units were not statistically distinct from each other.

Table 3: Efficiency parameter by unit, only units with more than 10 surveys included. Unit name anonymized

Unit	Count	Average α	Variance
D	40	1.31	0.05
R	24	1.34	0.03
V	88	1.37	0.09
U	12	1.40	0.11
Q	137	1.40	0.07
G	59	1.45	0.09
F	34	1.59	0.16
S	95	1.71	0.12

Table 4: Results from Tukey-Kramer Test for the effect of unit on the efficiency parameter, only units with more than ten surveys shown. 'T' indicates the pairwise comparison are significantly ($p<0.05$) different, F indicates they are not. Only units 'F' and 'S' are significantly different from the rest.

	D	R	V	U	Q	G	F	S
D		F	F	F	F	F	T	T
R			F	F	F	F	T	T
V				F	F	F	T	T
U					F	F	F	T
Q						F	T	T
G							F	T
F								F
S								

This analysis of the effect of unit and region seems to suggest that these categories are not very useful in a model of efficiency. While two of the eight regions were distinct, one of those regions (Hawaii) had very few surveys. Most of the units were also statistically indistinguishable from each other. Because of the concentration of individual units to particular areas, these two parameters are also not fully independent. For example, 47% of the work of unit 'S' was performed in the West Coast region, and 76% of all work done in the West Coast region was done by unit 'S'. This suggests the apparent distinctness of the West Coast region is a reflection of the distinctness of the unit that performed most of the work rather than anything particular to that region itself.

<u>Other Parameters</u>

With not much power in our initial idea of a regional dependence, we looked for additional parameters that might add to the skill of our model. These parameter were: average depth, depth standard deviation, and a measure of seafloor roughness or rugosity. Our particular interest in bathymetric measures reflects our intuition that deep, flat, low relief areas would naturally lead to higher coverage efficiency and thus a lower alpha parameter (closer to unity).

Figure 2 shows plots of the efficiency (alpha) parameter as a function of these other variables. While the efficiency parameter is associated with depth, depth standard deviation, survey size (all $p<0.001$), and rugosity ($p<0.060$), these factors explain little of the observed variability (R^2 = 0.048, 0.039, 0.036, and 0.006 respectively).

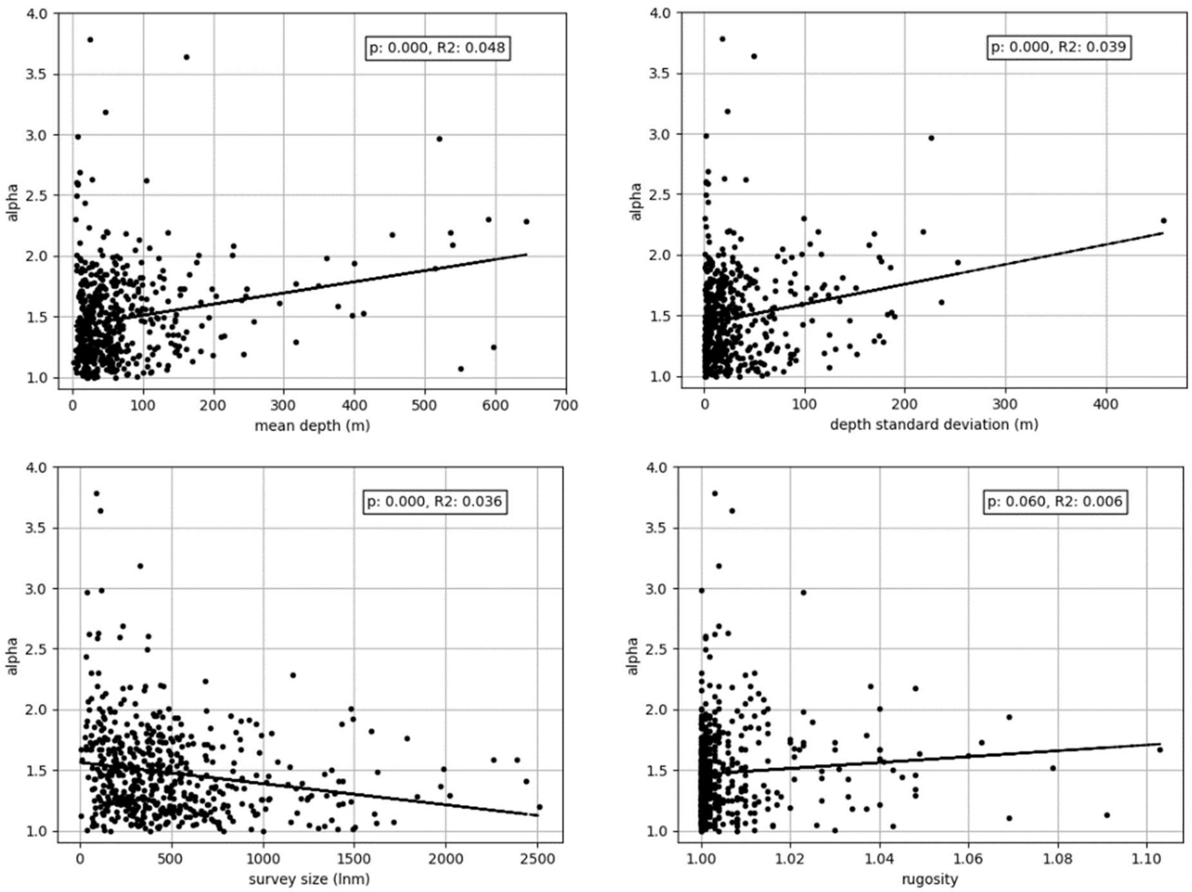

Figure 2: Efficiency parameter as a function of average depth (top left), depth standard deviation (top right), survey size (bottom left), and a measure of rugosity (bottom right). The efficiency parameter is associated with parameters depth, depth standard deviation, and survey size (all p<0.001), as well as the rugosity (p<0.060) parameter, but these factors explain little of the observed variability.

<u>Fitted Distribution and Modeled Uncertainty</u>

Unable to improve our model in a meaningful way with the inclusion of information on region or other physical parameters, we treated all of the study area as one group. We plotted a histogram of the efficiency parameter across all surveys and fit four continuous statistical distributions, Rayleigh, Chi, Gamma, and Inverse Gaussian, to the empirical results (Figure 3). The RMS residuals for each fit are Rayleigh: $2x10^{-4}$, Chi: $9x10^{-3}$, Gamma: $4x10^{-4}$, and Inverse Gauss: $4x10^{-6}$. Based on the low residual we selected the fitted inverse Gaussian to develop estimates of uncertainty.

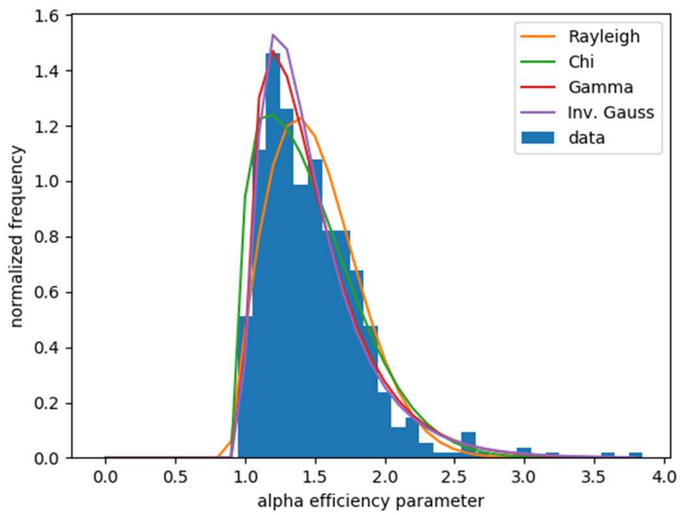

Figure 3: Histogram of efficiency parameter (all surveys) including four distribution fits

The selection of the inverse Gaussian is motivated entirely by the closeness of fit, we have no *a priori* good reason to believe that this parameter should be governed by such a distribution.

For any one survey, we use our inverse Gaussian fit to estimate the average efficiency parameter to be 1.48 with a 95% confidence interval of 1.03 to 2.45 (using the fitted cumulative distribution function per the percent point range [0.025,0.975]). Thus, for any one survey there is considerable uncertainty in our estimate. But if we consider a campaign of many independent surveys, the estimate of the *overall* efficiency parameter becomes significantly more precise (on order $N^{-½}$) due to the Central Limit theorem (CLT). We verify this notion empirically, using repeated drawings ($N_d = 1000$) from the fitted inverse Gaussian distribution for number of simulated surveys (N_s = 1 to 100). We then calculate the mean and standard deviation across those sampled surveys to model the uncertainty of an arbitrary campaign of N surveys. Figure 4 shows the overall deviation model as a function of N independent surveys:

$$\sigma_{\text{campaign}} = \frac{0.426}{\sqrt{\text{N}}} \tag{7}$$

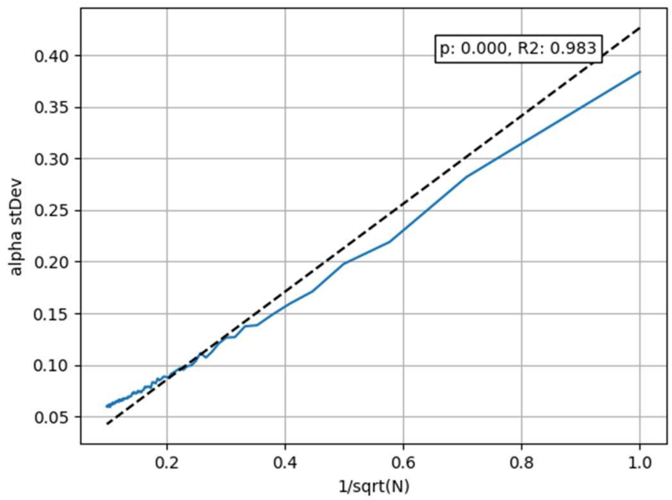

Figure 4: modeled overall deviation from campaign of N independent surveys

For a campaign of multiple surveys, the uncertainty of the effective α for a collection of surveys due to the variance in the estimate for α for a single survey will reduce with an increase in N. The uncertainty contribution due to the original estimate of the mean value (computed from the 547 recent surveys) will not. Estimating the uncertainty of the mean value with the standard error of the mean (recognizing that this does not strictly apply to a non-Gaussian distribution), we get

$$\sigma_{\text{mean}} = \frac{\sigma_{547}}{\sqrt{547}} = \frac{0.367}{\sqrt{547}} = 0.016 \tag{8}$$

Combining in a root-mean-square sense the campaign variance uncertainty, which scales with N, with the uncertainty in the mean, which does not, we arrive at an overall estimate of uncertainty for the mean α from an ensemble of N surveys, in Equation 9:

$$\hat{\mu}_{\alpha}\ (95\%\ \text{CI}) = \pm 1.96\sqrt{\left(\frac{0.426}{\sqrt{\text{N}}}\right)^2 + 0.016^2} \cong \pm\frac{0.83}{\sqrt{\text{N}}} \tag{9}$$

This Central Limit Theorem approach to a campaign α uncertainty is similar to the score interval approach as discussed in the paper by Arefi et al (Arefi 2008). In addition to the score interval, they examine a Wald interval and a likelihood ratio interval for inverse-Gaussian mean confidence estimation. Arefi et al prefer the likelihood ratio interval-expression, adapted here in Equation 10:

$$\frac{\text{N}\lambda\hat{\mu}_{\alpha}}{\text{N}\lambda + 1.96\sqrt{\text{N}\lambda\hat{\mu}_{\alpha}}} \le \hat{\mu}_{\alpha}\ (95\%\ \text{CI}) \le \frac{\text{N}\lambda\hat{\mu}_{\alpha}}{\text{N}\lambda - 1.96\sqrt{\text{N}\lambda\hat{\mu}_{\alpha}}} \tag{10}$$

In Equation 10, λ is the inverse Gaussian distribution shape parameter, and point estimated from our fit to the 547 recent surveys data – the maximum likelihood estimate (MLE) expression in equation 11:

$$\text{MLE}\ \hat{\lambda} = \text{N}\left[\sum_{i=1}^{\text{N}} \frac{(\alpha_i - \bar{\alpha})^2}{\alpha_i \bar{\alpha}^2}\right]^{-1} \tag{11}$$

A comparison of this likelihood ratio approach compared to the central-limit based approach is shown in Figure 5. The results are comparable with the exception that the likelihood ratio approach asymptotically approaches zero for large N because it does not include a notion of the uncertainty of the inverse Gaussian distribution shape parameter.

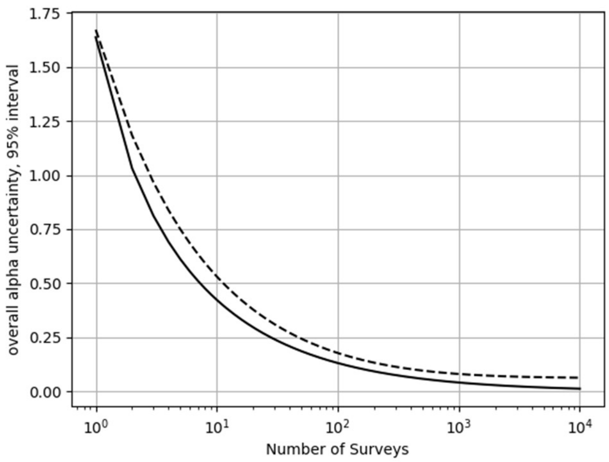

Figure 5: Ensemble estimate for uncertainty of α for multiple survey campaign (note logarithmic x-axis). Dashed line is Central Limit Theorem-based interval; solid line is from the likelihood ratio interval (Arefi 2008)

Thus while our estimate of the level of effort to complete one survey is rather large (± 70% at 95% CI), for a campaign of one-hundred surveys, we expect our estimates to be much more precise (under ± 10% at 95% CI). The uncertainty estimate asymptotically approaches 3% (at 95% CI) for large N.

<u>Section II - An Estimate of Survey Effort</u>

With an empirically tuned model and a sense of the model uncertainty, we can now turn to the task of estimating how much effort it will take to survey any given area. We need some concept of the depths across our area, and then simply apply equation 5 across the area. Our initial interest concerns two specific questions: How much effort would it be to resurvey the entire U.S. Exclusive Economic Zone (EEZ, Figure 6), and given that we have adequate surveys for some of the total area already, how much effort would it be to complete mapping the U.S. EEZ to modern standards? For both of these questions, we are also interested characterization of the effort; i.e., is the bulk of the remaining effort in deep water? In Alaska?

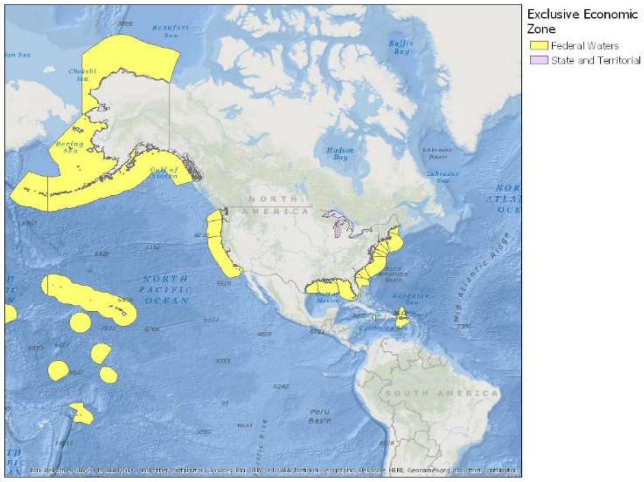

Figure 6: The US EEZ.

In order to ascertain the nature and location of the effort, we broke the U.S. EEZ into state and federal waters (associating the federal waters offshore of a state with that state) and further divided each into a series of depth bands. For our input depth layer, we used the GEBCO 2019 bathymetric compendium (REF).

We calculated a layer of linear survey effort per resolution cell (using equation 5) and then summed across the particular area and depth band of interest. For coverage of areas already mapped to modern standards, we considered areas mapped with complete MBES, 200% side scan with concurrent MBES, 100% side scan with complete MBES, or side scan with no MBES as mapped to modern standards. This includes surveys conducted by NOAA Coast Survey specifically for charting, as well as surveys conducted under the extended continental shelf program, and exploration surveys conducted by the NOAA Office of Ocean Exploration.

While we calculated results for the entire area, we present results only for waters deeper than four meters. Four meters is often (though not always) the approximate inshore limit of current Coast Survey hydrographic surveys (Coast Survey 2018), and because of the inverse relationship between area covered and depth, the estimates for full coverage of very shallow water are enormous.

RESULTS

Estimate of Effort for Full Resurvey of the US EEZ

To survey, completely, the entire 3.6 M SNM (square nautical miles) of the US EEZ would require a survey ship to acquire data for 14.2 ± 0.5 M (at 95% CI) linear nautical miles (LNM). To estimate uncertainty, we need an idea of how many independent surveys would be required, so used an estimate of 500 LNM per survey and then use equation (9) to calculate the uncertainty. The average value for all the surveys evaluated in the first part of this paper was 480. For large regions, these values are close to the asymptotic value of 3% (at 95% CI).

To present the survey effort more accessibly, we mirror the approach of (Mayer, et al. 2018) and define a ship-year as the distance travelled by a platform running continuously at a survey speed of 7.5 knots (14 km/h). Thus, one-ship year is 65,745 LNM. It is important to recognize that an actual survey ship can rarely approach a ship-year, thus defined, of data acquisition in any given year due to resupply and refit time in port, transit time to project area, as well as time off-line due to turns, CTD casts, and other operational aspects of conducting an actual survey. Other factors, such as achieved survey speeds, deployment of multiple platforms from a mother ship, operational efficiency will also be highly variable across different survey approaches. These are all aspects worthy of detailed study, but not specifically addressed in the present model; here, we address the on-line, on-project LNM required to accomplish a survey task. While changing the rate of acquisition changes the total figures, it obviously does not change the relative distribution of effort. Using 65,745 LNM/ship-year, we estimate it would take 216 ± 7 ship years to completely map the entire US EEZ with full bottom coverage (Table 5).

While 77% of the area of the EEZ is waters deeper than 1,000 m, this area is relatively quick to map. At nine ship-years, this constitutes only 4% of the total effort. Conversely, waters less than 200 m make up 18% of the total area of the EEZ, but constitute 94% of the total effort. Nor is the effort divided equally on a regional basis, the waters less than 200 meters in Alaska and Florida constitute nearly 63% of the total effort.

Table 5: Area and estimated survey effort for complete resurvey in ship-years; as defined here, a ship-year equals 65,745 LNM (linear nautical miles) of survey. State area values include the portion of federal waters contained within the offshore extended-boundaries.

	Areas - (Square Nautical Miles, thousands)									Effort - (ship years)								
	5-20 m	20-40 m	40-200 m	200-1,000 m	1,000 - 1500 m	1,500 to 3,000 m	3,000-5,750 m	5750 +	Total	5-20 m	20-40 m	40-200 m	200-1,000 m	1,000 - 1500 m	1,500 to 3,000 m	3,000-5,750 m	5750 +	Total
Alabama	0	1	1	1	0	0	-	-	3	0.4	0.4	0.1	0.0	0.0	0.0	-	-	1.0
Alaska	32	70	285	60	21	100	460	35	1,064	34.5	28.0	46.4	1.9	0.2	0.5	1.3	0.1	112.8
American Samoa	0	0	0	0	0	2	114	1	118	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.3
California	1	1	6	15	9	17	115	-	165	0.9	0.4	0.9	0.3	0.1	0.1	0.3	-	3.1
Caribbean	0	0	0	3	2	6	39	9	60	0.1	0.1	0.1	0.1	0.0	0.0	0.1	0.0	0.5
Connecticut	0	0	0	-	-	-	-	-	0	0.3	0.1	0.0	-	-	-	-	-	0.3
Delaware	0	0	0	-	-	-	-	-	1	0.4	0.1	0.0	-	-	-	-	-	0.5
Florida	14	18	26	38	6	2	16	-	121	15.0	7.5	4.1	0.9	0.1	0.0	0.1	-	27.6
Georgia	2	2	1	1	-	-	-	-	6	1.7	0.8	0.2	0.0	-	-	-	-	2.7
Guam	0	0	0	0	0	0	-	-	0	0.0	0.0	0.0	0.0	0.0	0.0	-	-	0.0
Hawaii	1	2	3	11	6	28	671	0	721	0.9	0.7	0.5	0.2	0.1	0.1	1.7	0.0	4.2
Illinois	-	-	1	-	-	-	-	-	1	-	-	0.1	-	-	-	-	-	0.1
Indiana	-	-	0	-	-	-	-	-	0	-	-	0.0	-	-	-	-	-	0.0
Louisiana	6	5	8	8	7	30	7	-	72	7.6	1.9	1.4	0.2	0.1	0.2	0.0	-	11.4
Maine	1	1	9	2	-	-	-	-	12	0.7	0.3	0.9	0.1	-	-	-	-	2.0
Maryland	1	1	1	0	0	4	5	-	12	1.5	0.3	0.2	0.0	0.0	0.0	0.0	-	2.1
Massachusetts	2	4	19	4	1	9	14	-	52	1.9	1.4	2.9	0.2	0.0	0.0	0.0	-	6.4
Michigan	-	-	29	0	-	-	-	-	29	-	-	2.0	0.0	-	-	-	-	2.0
Minnesota	-	-	2	0	-	-	-	-	2	-	-	0.1	0.0	-	-	-	-	0.1
Mississippi	0	0	0	-	-	-	-	-	0	0.4	0.1	0.0	-	-	-	-	-	0.4
New Hampshire	0	0	0	-	-	-	-	-	0	0.0	0.0	0.0	-	-	-	-	-	0.1
New Jersey	1	3	4	0	0	4	3	-	16	1.2	1.1	0.8	0.0	0.0	0.0	0.0	-	3.1
New York	1	2	8	1	0	3	1	-	16	0.7	0.6	1.4	0.0	0.0	0.0	0.0	-	2.8
North Carolina	4	6	3	7	1	11	44	-	76	3.9	2.7	0.5	0.2	0.0	0.1	0.1	-	7.5
Northern Mariana Islands	0	0	0	0	0	0	-	-	1	0.0	0.0	0.0	0.0	0.0	0.0	-	-	0.1
Ohio	-	-	3	-	-	-	-	-	3	-	-	0.2	-	-	-	-	-	0.2
Oregon	0	0	4	4	1	27	12	-	49	0.3	0.1	0.5	0.1	0.0	0.1	0.0	-	1.2
Pacific Remote Islands	0	0	0	2	6	71	714	67	860	0.1	0.0	0.0	0.0	0.1	0.4	1.8	0.1	2.5
Pennsylvania	0	-	1	-	-	-	-	-	1	0.0	-	0.0	-	-	-	-	-	0.0
Puerto Rico	0	0	0	0	0	0	-	-	1	0.6	0.1	0.0	0.0	0.0	0.0	-	-	0.7
Rhode Island	0	0	1	0	0	0	-	-	2	0.2	0.2	0.2	0.0	0.0	0.0	-	-	0.6
South Carolina	3	3	2	12	1	2	0	-	23	3.1	1.3	0.4	0.3	0.0	0.0	0.0	-	5.0
Texas	5	6	9	6	6	5	0	-	36	4.2	2.4	1.5	0.1	0.1	0.0	0.0	-	8.4
Virgin Islands	0	0	0	0	0	0	0	-	0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	-	0.2
Virginia	2	2	1	0	0	2	1	-	9	2.5	0.9	0.2	0.0	0.0	0.0	0.0	-	3.7
Washington	1	1	4	2	1	16	-	-	24	0.8	0.3	0.5	0.0	0.0	0.1	-	-	1.8
Wisconsin	-	-	7	0	-	-	-	-	7	-	-	0.5	0.0	-	-	-	-	0.5
Total	79	128	439	177	72	341	2,216	112	3,565	84.0	51.9	66.7	4.7	0.7	1.8	5.9	0.2	215.9
% of total	2%	4%	12%	5%	2%	10%	62%	3%	100%	39%	24%	31%	2%	0%	1%	3%	0%	100%

Estimate of Effort to Complete Mapping EEZ to Modern Standards

Excluding all areas already mapped to modern standards, we estimate that 56% of area of the US EEZ remains unmapped to modern standards, but 82% is unmapped from a level of effort perspective. To complete this work, we estimate 177 ± 6 ship-years is required. Again, while the bulk of the unmapped area is in deep waters (68% of the unmapped area is deeper than 1,000 m), the vast majority of the remaining effort, 96%, is in water less than 200 m, with 39% of the remaining effort in waters less than 20 m.

Table 6: Area and estimated survey effort to map areas not currently mapped to modern standards. Area and estimated survey effort to map those areas not already mapped to modern standards; as defined here, a ship-year equals 65,745 LNM (linear nautical miles) of survey. State area values include the portion of federal waters contained within the offshore extended-boundaries.

	Areas - (Square Nautical Miles, thousands)									Effort - (ship years)								
	5-20 m	20-40 m	40-200 m	200-1,000 m	1,000 - 1500 m	1,500 to 3,000 m	3,000-5,750 m	5750 +	Total	5-20 m	20-40 m	40-200 m	200-1,000 m	1,000 - 1500 m	1,500 to 3,000 m	3,000-5,750 m	5750 +	Total
Alabama	0	1	0	0	0	-	-	-	2	0.12	0.4	0.1	0.0	0.0	-	-	-	0.6
Alaska	31	63	243	41	12	41	304	24	759	32.5	25.3	40.1	1.3	0.1	0.2	0.9	0.0	100.4
American Samoa	0	0	0	0	0	0	73	1	74	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.2
California	0	0	3	2	1	1	32	-	39	0.5	0.2	0.5	0.0	0.0	0.0	0.1	-	1.3
	0	0	0	0	0	1	22	0	24	0.1	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.3
	0	0	0	0	1	3	0	-	4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-	0.0
Caribbean	0	0	0	1	1	3	22	0	28	0.1	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.3
Connecticut	0	0	-	-	-	-	-	-	0	0.1	0.0	-	-	-	-	-	-	0.1
Delaware	0	0	0	-	-	-	-	-	0	0.1	0.0	0.0	-	-	-	-	-	0.1
Florida	13	17	23	26	4	0	6	-	89	13.9	7.0	3.7	0.6	0.0	0.0	0.0	-	25.3
Georgia	1	2	1	0	-	-	-	-	4	1.1	0.7	0.1	0.0	-	-	-	-	2.0
Guam	0	0	0	0	0	0	-	-	0	0.0	0.0	0.0	0.0	0.0	0.0	-	-	0.0
Hawaii	1	1	1	1	1	3	314	0	322	0.9	0.6	0.2	0.0	0.0	0.0	0.8	0.0	2.4
Illinois	-	-	1	-	-	-	-	-	1	-	-	0.1	-	-	-	-	-	0.1
Indiana	-	-	0	-	-	-	-	-	0	-	-	0.0	-	-	-	-	-	0.0
Louisiana	3	4	6	0	0	9	4	-	27	3.7	1.7	1.2	0.0	0.0	0.0	0.0	-	6.7
Maine	1	0	8	2	-	-	-	-	10	0.7	0.2	0.7	0.1	-	-	-	-	1.7
Maryland	0	1	1	0	-	-	-	-	2	0.7	0.2	0.2	0.0	-	-	-	-	1.1
	1	0	0	-	-	-	-	-	1	0.9	0.1	0.0	-	-	-	-	-	1.0
	1	3	16	2	0	0	0	-	21	0.7	1.1	2.4	0.1	0.0	0.0	0.0	-	4.3
Massachusetts	1	3	16	2	0	0	0	-	22	1.5	1.2	2.5	0.1	0.0	0.0	0.0	-	5.3
Michigan	-	-	29	0	-	-	-	-	29	-	-	2.0	0.0	-	-	-	-	2.0
Minnesota	-	-	2	0	-	-	-	-	2	-	-	0.1	0.0	-	-	-	-	0.1
Mississippi	0	0	0	-	-	-	-	-	0	0.1	0.0	0.0	-	-	-	-	-	0.2
New Hampshire	0	0	0	-	-	-	-	-	0	0.0	0.0	0.0	-	-	-	-	-	0.0
New Jersey	1	2	3	0	-	-	0	-	6	0.4	0.9	0.7	0.0	-	-	0.0	-	2.0
New York	0	1	8	0	-	-	0	-	9	0.3	0.4	1.3	0.0	-	-	0.0	-	2.1
North Carolina	3	5	2	4	0	2	1	-	17	3.4	2.2	0.4	0.1	0.0	0.0	0.0	-	6.2
Northern Mariana Islands	0	0	0	0	0	0	-	-	0	0.0	0.0	0.0	0.0	0.0	0.0	-	-	0.0
Ohio	-	-	3	-	-	-	-	-	3	-	-	0.2	-	-	-	-	-	0.2
Oregon	0	0	2	0	0	3	3	-	9	0.2	0.1	0.3	0.0	0.0	0.0	0.0	-	0.6
Pacific Remote Islands	0	0	0	0	2	17	435	28	482	0.0	0.0	0.0	0.0	0.0	0.1	1.0	0.1	1.3
Pennsylvania	0	-	1	-	-	-	-	-	1	0.0	-	0.0	-	-	-	-	-	0.0
Puerto Rico	0	0	0	0	-	-	-	-	1	0.5	0.1	0.0	0.0	-	-	-	-	0.7
Rhode Island	0	0	1	0	-	-	-	-	1	0.1	0.1	0.2	0.0	-	-	-	-	0.3
South Carolina	3	3	2	9	1	1	0	-	18	2.7	1.2	0.3	0.2	0.0	0.0	0.0	-	4.4
Texas	4	5	7	1	3	1	0	-	20	3.3	2.0	1.4	0.0	0.0	0.0	0.0	-	6.7
Virgin Islands	0	0	0	0	0	0	-	-	0	0.1	0.0	0.0	0.0	0.0	0.0	-	-	0.1
Virginia	0	1	1	-	-	-	-	-	3	0.7	0.5	0.2	-	-	-	-	-	1.4
Washington	0	0	2	0	0	2	-	-	5	0.6	0.2	0.3	0.0	0.0	0.0	-	-	1.0
Wisconsin	-	-	7	0	-	-	-	-	7	-	-	0.5	0.0	-	-	-	-	0.5
Total	64	113	373	89	24	84	1,194	53	1,992	68.4	45.5	57.2	2.5	0.2	0.4	3.0	0.1	177.4
% of total	3%	6%	19%	4%	1%	4%	60%	3%	100%	39%	26%	32%	1%	0%	0%	2%	0%	100%

Conclusion

We developed a simple model to estimate survey effort for any given area. This model is driven by the fact that for many modern mapping systems, width of survey coverage is proportional to depth. We introduced an efficiency parameter and tuned this parameter using data from over 500 actual surveys. We had expected this parameter to be strongly tied to seafloor complexity and thus regionally consistent; it was not. We investigated a number of additional variables to add skill to the model. We could discern no strong relationship between this parameter and any variable investigated, including region, survey unit, seafloor roughness, variability, average survey depth, or survey extents. This was a surprising result. Perhaps the nature of the survey areas in different regions tend to offset. For example, the relatively flat areas off the mid-Atlantic coast might allow for consistent line spacing and well planned coverage (which would tend to improve the survey efficiency), but those same areas tend to be shallow, and necessitate a higher percentage overlap to accommodate a given steering error (which would tend to degrade survey efficiency). Other parameters that we did not consider here, such as weather or oceanographic conditions, may also have a significant influence on survey efficiency. We hope the methods outlined here establish a good foundation for future study and facilitate others to explore additional parameters of interest.

This model works at all scales- from a small, single vessel project to global scales, though the uncertainty of the estimate (in fractional terms) improves with larger areas. We used this model to develop an estimate for the level of effort to either complete or resurvey U.S. waters. While we figure we have 56% of the area of the U.S. EEZ remaining to map to modern standards, we have 82% from a level of effort perspective. The disparity between these numbers reflects a fundamental, and somewhat counterintuitive, reality of modern mapping systems: because coverage is proportional to depth, it is much, much slower to map in shallow water than it is to map in deep water. Surveys in deep water make rapid progress on adding *area* to the mapped tally, but most of the *work* is in shallow water. To finish mapping to modern standards, we estimate 177 ± 6 (at 95% CI) ship-years of work. The uncertainty in this figure reflects that of the model, but we again stress that we define a ship-year to mean one ship steaming continuously for one year- a level of actual effort unlikely to be achieved in practice. We also recognise that implicit in the idea that we can map a deep area to modern standards quickly is an acceptance of lower resolution data in deeper water. If we were to map all areas to the same resolution, we would likely have to maintain a constant distance of the sensor system to the seafloor and the coverage rate would be largely independent of depth.

The other potential application for this model, the efficiency evaluation goal, is perhaps less useful. We had thought that if we had an effective model for survey effort, we might be able to compare an executed survey to the model to judge efficiency or evaluate the impact of some tested parameter. The uncertainty model sets clear thresholds for such a signal to rise above the observed variability. For small numbers of surveys, this threshold is very high (± 70% at 95% CI), so this model is probably not an effective tool to evaluate individual surveys or individual treatments.

This model does, when run on the raster results from a survey, directly give a metric for survey output that appropriately scales with survey effort, but is independent of the actual effort expended on that survey. We think this theoretical nautical mile metric may usefully supplement current metrics linked to actual linear nautical miles of effort, which is not independent of the particulars of the execution of the survey, and area surveyed, which is independent of the survey approach, but badly tracks actual progress towards the goal.

Even with a reasonably large survey fleet of government owned and contracted vessels, and leaving aside the resurvey requirements of rapidly changing areas, the level of effort to complete the full survey of U.S. waters to modern standards is likely prohibitive to contemplate achieving over a scale of decades. This does not however mean that the task is hopeless. Prioritization efforts (e.g. (Riding, Roberts and Priovolos 2016), (Chenier, Abado and Martin, CHS Priority Planning Tool (CPPT) - A GIS model for Defining Hydrographic Survey and Charting Priorities 2018)) are obviously critical to achieving the highest impact with available resources; we can clearly not simply commit to getting it all done in the next five to ten years.

On the other hand, this study highlights that major regional campaigns are entirely feasible. We could, for instance complete the survey of New England Waters in 7.4 ship-years, or all the Pacific Islands in 3.7 ship-years. It also suggests that we may need to think of alternative survey approaches or technology such as satellite-derived bathymetry (e.g. (Chenier, Faucher and Ahola, Satellite-Derived Bathymetry for Improving Canadian Hydrographic Service Charts 2018)) for some of the hardest area to survey (e.g. 4-20 meters depths off Florida and Alaska). We hope this study provides a useful framework for planning the survey fleet, instrumentation needs, and prioritization of survey efforts to best achieve the goal President Thomas Jefferson originally set out in 1807, that Coast Survey be "provided for surveying the coasts of the United States." We hope to soon extend this analysis to the globe.

References

Arefi, M., Borzadaran, G.R. Mohtashami and Vaghei Y. 2008. "A note on interval estimation for the mean of inverse Gaussian distribution." *SORT 32 (1)*, January-June: 49-56. http://www.idescat.cat/sort/sort321/32.1.2.arefi-etal.pdf.

Calder, Brian, Shannon Byrne, Bill Lamey, Richard T. Brennan, James D. Case, David Fabre, Barry Gallagher, et al. 2005. "The Open Navigation Surface Project." *International Hydrographic Review* 6 (2): 9-18.

Chenier, Rene, Loretta Abado, and Heather Martin. 2018. "CHS Priority Planning Tool (CPPT) - A GIS model for Defining Hydrographic Survey and Charting Priorities." *International Journal of Geo-Information.* doi:10.3390.

Chenier, Rene, Marc-Andre Faucher, and Ryan Ahola. 2018. "Satellite-Derived Bathymetry for Improving Canadian Hydrographic Service Charts." *International Journal of Geo-Information* 7 (8): 306.

Coast Survey. 2018. *Hydrographic Surveys Specifications and Deliverables.* Silver Spring, MD: DOC/NOAA/NOS, April. https://nauticalcharts.noaa.gov/publications/docs/standards-and-requirements/specs/hssd-2018.pdf.

Mayer, Larry, Martin Jakobsson, Graham Allen, Boris Dorschel, Robin Falconer, Vicki Ferrini, Geoffroy Lamarche, Helen Snaith, and Pauline Wetherall. 2018. "The Nippon Foundation—GEBCO Seabed 2030 Project: The Quest to See the World's Oceans Completely Mapped by 2030." *Geosciences* 8 (2): 63. doi:10.3390/geosciences8020063.

Riding, John, Jennifer Roberts, and Gianis Priovolos. 2016. *New Zealand Hydrographic Risk Assessment.* Wellington, NZ: Land Information New Zealand.