Risk Models and Survey Completeness (or, "Are We There Yet?")

Brian R. Calder

Center for Coastal and Ocean Mapping and NOAA/UNH Joint Hydrographic Center,
University of New Hampshire, Durham, NH 03824, USA

Abstract

Hydrography has no well-established, mathematically rigorous, objective model for data quality. Survey effort can therefore sometimes be applied inefficiently, in that there can be as much effort expended on 1m high rocks in 30m depth as in 5m (which should be of lesser concern), or on areas where there is light to non-existent traffic. A computational model of return-on-investment associated with survey effort might allow assessment of where to survey first in a given area, and to determine when extant survey effort was "good enough" to meet survey specification, leading to survey efficiencies.

Mathematical models of risk associated with ship passage have previously been demonstrated as models for generalized end-user chart uncertainty, and re-survey priority estimation. Here, they are proposed as a model for survey completion which can be applied incrementally in order to rationalize the effort being applied in each area of the working grounds, and to determine when the area is sufficiently well surveyed to be considered complete.

This paper demonstrates the requirements for, and implications of, adopting such an approach for survey completeness prediction, focusing particularly on the data dependencies, and model calibration. The methods are illustrated with historical survey data from Hampton Roads, VA.

Motivation

The lack of a formal, mathematically rigorous, objective model for data quality means that it can be difficult for hydrographers to allocate resources (e.g., survey ships, or survey launches deployed from a mothership) efficiently before and during a survey. Often, this can mean that there is a blanket approach to survey, where each area gets the same amount of effort, whether that effort is warranted or not. For example, in a deep area where the surface traffic is shallow draft, is it always necessary to consider a full-coverage bottom survey with multibeam echosounders, or would some other survey methodology suffice? It might be possible to, say, conduct a singlebeam survey with associated imaging (e.g., a surface-mounted sidescan stave to verify lack of targets to suitable depth between singlebeam lines) and still achieve the same ends in the sense that the risk of grounding for the surface ships in the area would not be materially affected by the selection of survey technology.

Similarly, the natural response of most hydrographers is to attempt to identify as well as possible every hazard detected in the course of a survey. In many cases, however, this can result in resources being expended on investigation of features that are not hydrographically significant (e.g., a 1m rock in 30m of water where the average shipping draft is 10m) simply because the objects appear in the data and, as good hydrographers, there is an inbred existential dread of missing something that might just be important. Taken to the extreme this could entail significant expenditure of time and personnel, so some hydrographic common sense may be applied to the situation, attempting to rationalise when to pursue targets, and when to accept the potential risk. Given no formal model for this, however, are those decisions defensible?

A third aspect of essentially the same problem occurs when considering the question of whether the survey is complete. Depending on the jurisdiction, the definition for this might be whether the soundings themselves meet particular uncertainty requirements [1], or if the nodes in a grid are based on a sufficient number of soundings to ensure statistical stability [2], whether a certain percentage of grid nodes are occupied, or something more vague, such as that all significant objects have been detected. Each attempts to model some aspect of the data being collected as a proxy for completeness, but does not really address the ultimate use of the survey: the potential for a surface ship to encounter difficulty when traversing the area. If it were possible to estimate the potential for difficulty (at least with respect to the bathymetry, the only feature that can be controlled with a survey), then it might be possible to reconsider the criteria for "done" to encompass the case where it is demonstrated that the residual risk to surface shipping has been reduced sufficiently that the certifying agency is willing to accept it.

It seems logical, therefore, that there may be some achievable efficiencies if there were some mechanism to address the issue of whether the data collected was sufficient for use, or not. In the worst case, this would at least allow decisions taken to allocate resources to be made in a more principled manner.

Previous research [3-5] has considered model of mathematical risk associated with the underkeel clearance (UKC) for ships transiting through a given area. Initially proposed as a means to summarise the state of knowledge of an area for user applications [3-4] and later extended to questions of resurvey priority [5], the principle is to determine, at each point in time, the statistical distribution of the UKC of a ship [6], given the stochastic nature of all of the variables that make up this computation (e.g., [7-8]). Once achieved, it is possible to address questions such as the probability of grounding at any time, or over the duration of a transit, and even a more general sense of 'risk' as the expected loss (or cost) of the results of a particular ship having a particular UKC in any given location. (So that, for example, the potential cost attributed to the same event could be significantly higher over areas with rocky bottom than over mud flats, or for a LPG tanker compared to a Boston Whaler.) In the sense that this model addresses directly the question of the user's risk associated with transiting an area, given all that is known about the area, it can be seen as a suitable summary of the potential for difficulty associated with the transit, and therefore as an objective means to assess the state of knowledge of the area for a given ship, or ship population. Consequently, it is proposed that this model can also be used to assist in questions of survey completeness, and resource allocation.

Given a risk model for the shipping in a given area, many of the motivating problems are relatively simple to address. For example, areas that were assessed as high risk *a priori* would make suitable early targets for survey, since if the asset were to run out of time (e.g., due to weather or mechanical difficulties) at least the most significant areas would already have been covered. Of course, these areas need not necessarily be the oldest survey areas, or the shallowest, since risk is relative to the type of ships that transit the area. The ability of the available survey tools to effectively reduce the risk would also be included in the selection criteria: high risk areas might remain (relatively) high risk even after new surveys, leading to low return-on-investment.

Given that observation, conducting a survey in general reduces the risk in an area, although no survey tool can completely eliminate it. If ideal circumstances are assumed, however, a risk model can predict the minimum residual risk achievable, which then provides a calibration point against which the survey effort can be assessed. Different survey methods will achieve reduction in risk to different degrees, and therefore assessment of risk reduction can potentially be used as a means to select particular methods for a given area. Reduction of risk to the ideal minimal level might not be possible – but might also not be required. So long as the risk is reduced below a level acceptable to the agency involved, it might be possible to select cheaper, faster, or more approximate methods in a principled manner, and still meet requirements. Selection of a minimal acceptable risk might also, potentially, be a more natural way of specifying a survey requirement than more abstract measures associated with the measurements themselves.

The potential for risk reduction with a given tool provides a natural way to assess whether a survey area can be considered "done". Once a tool provides no further risk reduction, any further survey is wasted whether the residual risk meets requirements or not. In addition to providing an assessment of when survey operations should be curtailed, this also provides means for the field unit to demonstrate either that the survey meets requirements, or that the requirements cannot be achieved with the tools available.

An immediate corollary to this is that, given an acceptable minimum risk, a survey area could be considered "complete" when the risk falls below the minimum, even if it is not at the theoretical minimum for the selected survey tool applied to that area, or there is no other technology available that would improve on the situation. The obvious extension to economic arguments with respect to available technologies are here ignored, although they might be considered in a larger planning context.

Therefore it is proposed that it might be possible to address many of the questions raised by assessments of survey completeness though use of a suitable risk model, which also provides a better means to express the state of knowledge of the data that is available in the aftermath of the survey. Potential benefits of this approach include providing guidance on where the survey should start; allowing for continuous monitoring of survey progress; providing guidance on the selection of tools and verifying that the tools can actually achieve the stated aids; tying completeness of survey to the state of knowledge of the area with respect to usage (rather than some other more arbitrary assessment criterion); and assisting in ensuring that the survey data collected is sufficient for use before the assets leave the working grounds.

Implementation

Basic Structure of the Risk Model

The risk model used here [4] represents the statistical distribution of the underkeel clearance (UKC), given all of the measurements and effects that are incorporated into this measure (Figure 1). In addition to the obvious effects of waterlevel, draft and dynamic draft, and observed depth, the model includes effects such as the influence of wind and waves on the motion of the ship (which can cause significant pitch and roll which can affect the UKC dramatically for large ships, Figure 2), the size and shape of the ship hull, and the local current climate. The behaviour of the UKC can be summarised either through grounding probability (i.e., the probability mass below the zero UKC mark), or through a generalised risk function (Figure 3), which can be spatially adaptive.

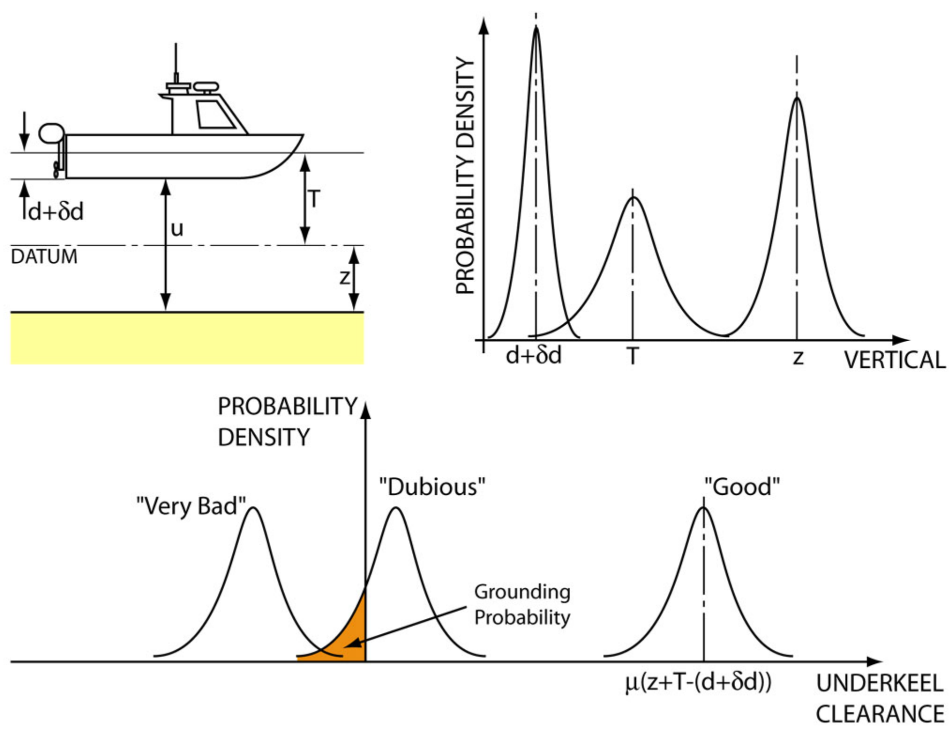

Figure 1: Basic structure of the underkeel clearance model. The uncertainty associated with the draft, dynamic draft, waterlevel, and observed depth influence the overall uncertainty of the underkeel clearance *u*. The behaviour of the UKC with respect to the available water determines the probability of grounding or, more generally, the risk.

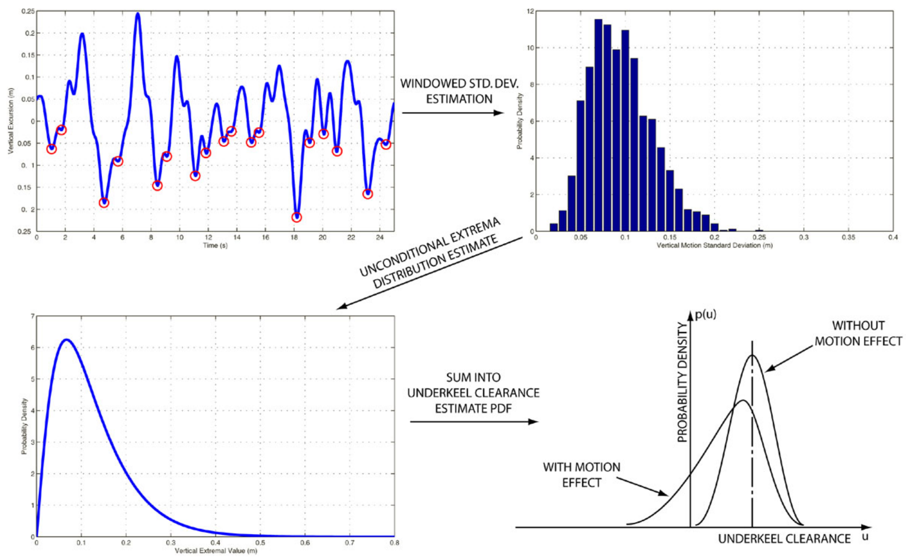

Figure 2: Effects of motion extrema on UKC. The negative extrema of the ship's vertical excursions can be analysed to determine their distribution, which is then mixed with the other UKC components, resulting in more left-tail probability mass. Where motion extrema are not measured, modeled results can be used instead.

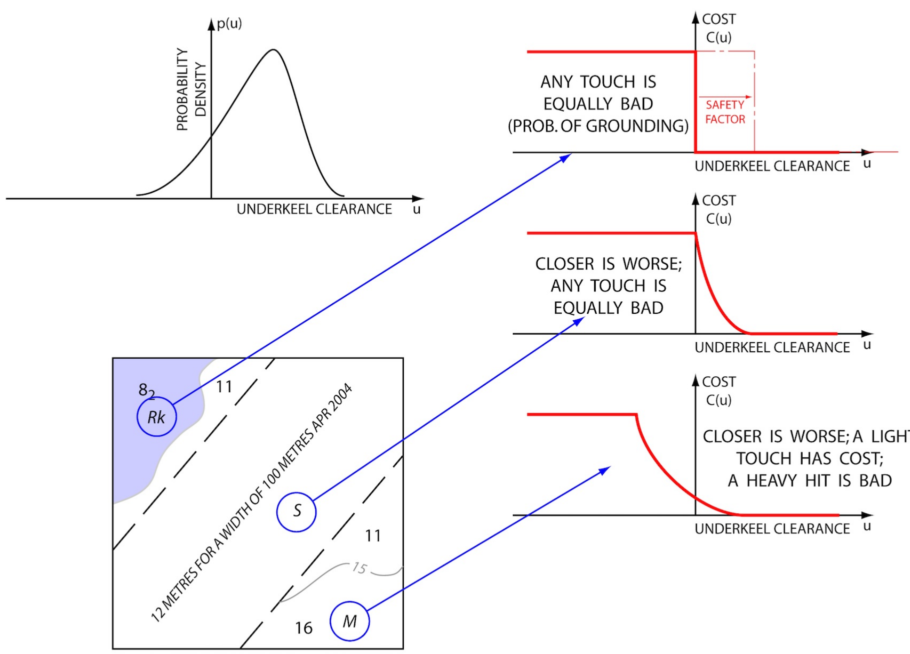

Figure 3: Risk cost functions with respect to the UKC. The risk is the product of UKC probability density and cost function, integrated over the whole UKC domain. The cost functions can be spatial variant to reflect the different costs in (for example) different regimes of rocky, sandy, or muddy seafloor.

The most basic form of the risk model predicts the UKC distribution at a point in space-time. This model can then be usefully applied in a number of different scenarios, for example by considering the risks associated with a given range of different manoeuvring scenarios that a ship might consider at a given moment in time, or by integrating the risk along a given trajectory. In the survey completeness case, the area of interest is split into a regular grid of analysis cells, and the model is applied to an ensemble of trajectory, environmental, and ship types within the cell; summary statistics are generated from a Monte Carlo assessment of risk along each trajectory within each cell. Full details of the basic risk model are given in [4].

Bathymetric Uncertainty and Potentially Unobserved Objects

A primary difficulty, and a big part of the UKC risk model, is that it is difficult, if not impossible, to assess a reliable bathymetric uncertainty from archive hydrographic data. That is, given data that was archived at survey scale (i.e., singlebeam, or leadline data), it is possible to generate a surface through an appropriate interpolation method, and even to predict a second-order uncertainty using methods such as kriging [9]. However, since the uncertainty estimation methods are based on estimates of spatial autocorrelation derived from the data, if the data are undersampled (as they almost always are), the uncertainty estimates will be unrealistic. Methods to partially accommodate this have been proposed [10, 11], but are difficult to manage and subjective.

As an alternative, the risk model uses a "best estimate" uncertainty for the bathymetric model, assuming that the model did capture everything that was important, and then incorporates a separate statistical model for those things which might have gone unobserved (e.g., glacial erratics that might have been between survey lines in a fixed-spacing singlebeam survey and undetected by the hydrographic practice of the time). In essence, the model gives up on a deterministic estimate of bathymetry and its uncertainty in favour of a model for what could potentially be there, given the information from the surveys available.

The extension to the basic UKC risk model (Figure 4) is based on the concept of marked spatial point processes [9]. In this model, a theoretical underlying point process indicates the rate at which objects might occur, and then for each object there is an associated distribution of potential heights above the seafloor and object horizontal dimension. Different types of objects can have different dimensions, and multiple overlying, spatially-varying models are possible.

The estimates of object height distributions are *a priori* estimated, and not fixed. In effect, each time a ship sweeps out new area with its hull and does not run aground, new information is provided as to the potential for there to be objects of a given height: the ship's hull acts as a type of wire-drag. Consequently, ships must maintain a constantly updated "personal" view of the possible height distributions based on their own experience, and AIS traffic data can be used to update object distributions (Figure 5) to reflect the local state of knowledge, even in the absence of any formal survey effort.

New surveys effectively constrain the potential for unobserved objects, and are a significant source of potential risk reduction. Object detection requirements from survey specifications are one source of calibration information for the magnitude of this effect.

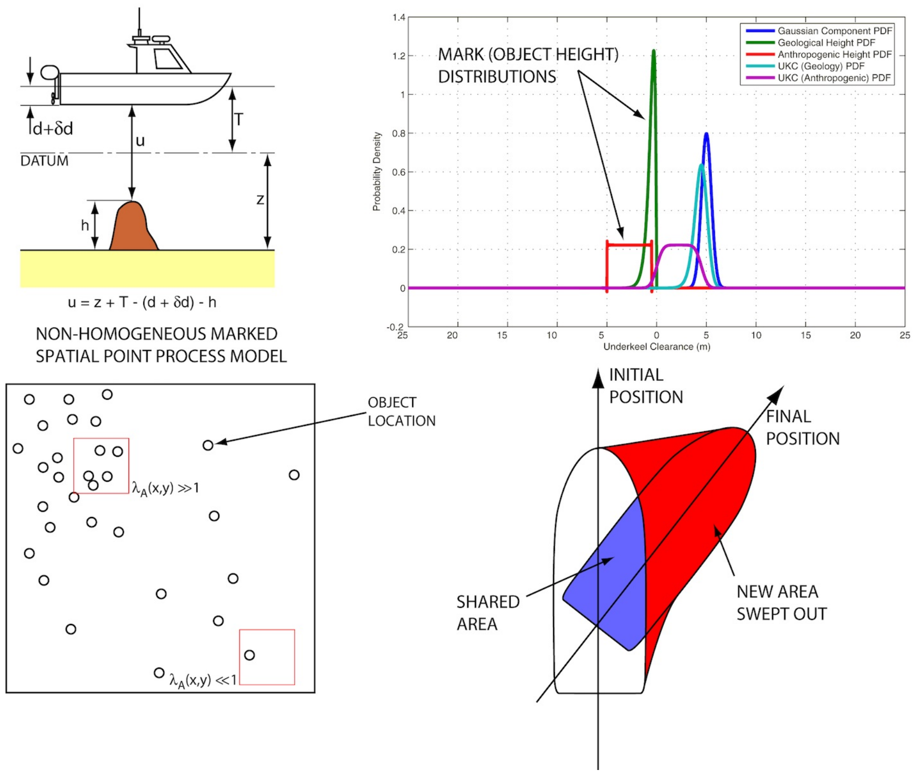

Figure 4: Marked spatial-point process object model. The potential presence of unobserved objects (of any kind) is represented by a spatial point process which measures the rate of object occurrence, and any clustering, within a given area; each potential object also has a height distribution that affects the UKC. Different types of objects (e.g., anthropogenic as opposed to geological) can have different rates, and height distributions. Using appropriate statistical techniques, the stochastic effects of these objects on UKC can be assessed.

Calibration from Source Data

The UKC risk model is strongly data driven, and therefore relies on the accuracy of the component models for plausibility of results. Estimates of bathymetric uncertainty are relatively well understood [12-14], and for most of the environmental data direct measurements can be used to form a climatology although real-time observations or an appropriate model could also be used. Ship traffic density, types, and dimensions can be derived from AIS data, with due attention to the quality assurance of the data (e.g., misreporting of draft and dimensions in feet rather than metres, [15]). Some care is required in assessing the traffic density and ship type distributions, however, to avoid undersampling and sampling bias.

The primary remaining calibration point is the unobserved object models. In the current model, these are estimated heuristically and then updated based on a year of AIS observations to constrain the potential distributions (Figure 5); additional constraints on height with respect to assumed water clarity are also added based on the assumption that an object in visible distance would probably have been reported. Alternative methods include using a small survey to calibrate for object abundance and height distribution.

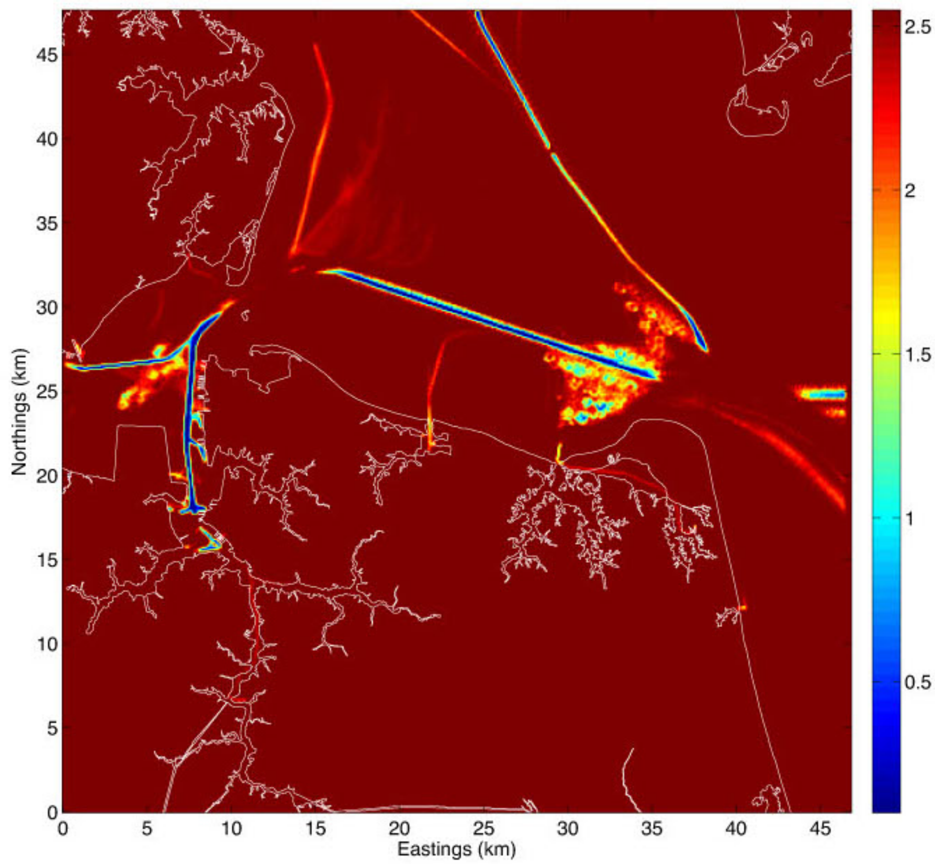

Figure 5: Updated object height distribution mean (in metres) based on a year of AIS traffic data in the Port of Hampton Roads, Chesapeake Bay, VA. The initially uniform distribution on the range [0, 5]m has been updated according to estimated ship drafts, showing where deep draft vessels have effectively constrained the potential for any objects since they did not hit them during numerous transits. Other areas with unchanged mean height of 2.5m saw no traffic, or only traffic with shallow draft relative to the depth.

Processes for Survey Management

Use of the UKC risk model for survey management revolves around continuous updating of the underlying model of the state of knowledge of the area, and recomputation of the risk factors associated with each area. Thus, for example, a risk estimate before the survey effort commences (Figure 6) can be used to identify target areas for early prosecution, while establishing the baseline improvement (i.e., the maximum improvement that could be achieved with perfect survey and all objects detected) is conducted by comparing the risk factors with and without objects and auxiliary bathymetric uncertainty estimates included (Figure 7). If required, the baseline could also be combined with the agency-derived maximum acceptable risk level so that improvements beyond this level would not be given higher priority, for example if the resource allocation algorithm was to select the simplest tool commensurate with reducing the estimated risk to, or below, the baseline.

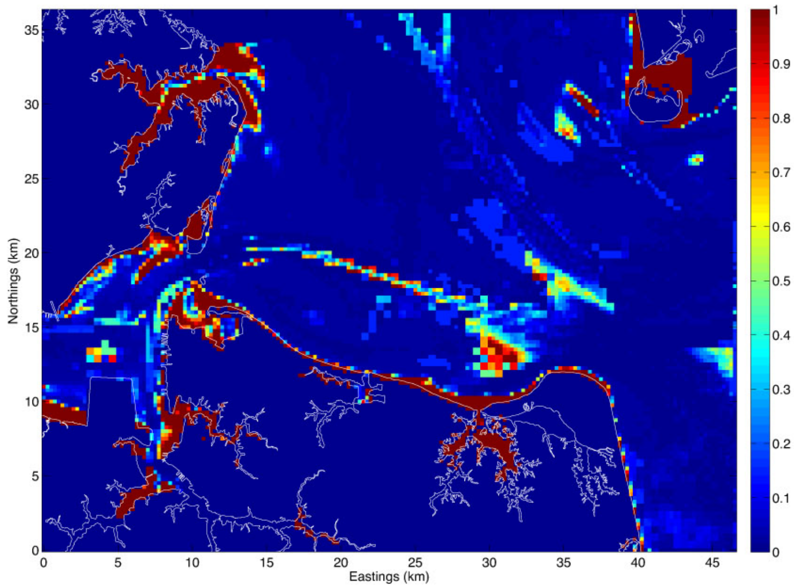

Figure 6: Estimate of risk associated with the approaches to the Port of Hampton Roads/Norfolk, VA based on historical data and nominal risk profiles, calibrated by AIS traffic data from 2013. Shallow areas are based mostly on a priori estimates of traffic, since the traffic density pattern is mostly in the deep draft channels.

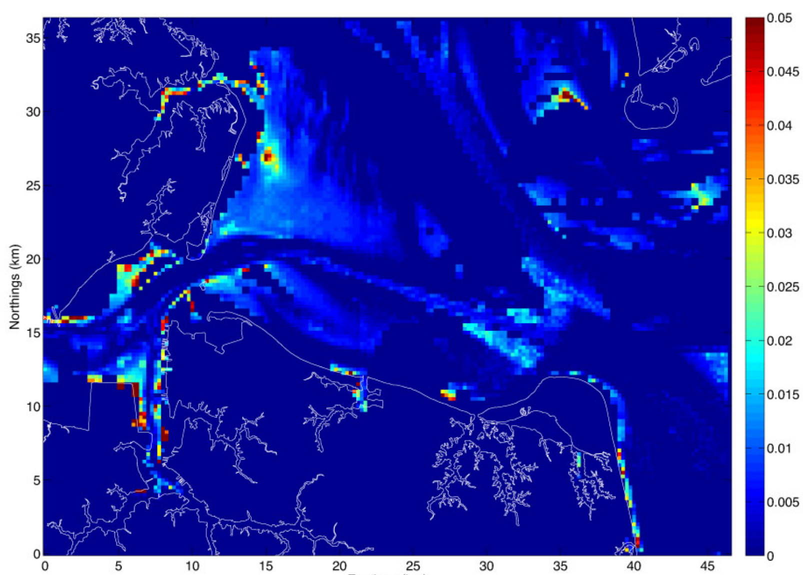

Figure 7: Difference in risk associated with ideal survey conditions across the entire region. Since this does not include traffic density weighting, survey effort might be directed to shallower areas. Note however that shallow areas which will not benefit from extra survey are not highlighted.

Similarly, assessment of newly collected survey data can be conducted by updating the bathymetric model appropriately, and then adjusting the unobserved object models to reflect the level of object detection achieved with the new survey. For multibeam survey, for example, it might be assumed that all objects above the minimum requirement are detected, and therefore that only smaller objects are expected; areas of partial survey could be accommodated by obvious means. For singlebeam survey with a surface imaging system, it might be assumed that any object within the extinction range of the imaging system would be detected, thereby constraining the potential height distributions. Conventional singlebeam surveys would continue to reflect the potential for unobserved objects. In any of these cases, improvements in risk assessed indicates the degree to which the measurements benefit the state of knowledge, and closeness to the baseline estimate (i.e., with ideal circumstances) measures the completeness of the survey, at least within the limitations of the tool selected.

Finally, tool selection can be addressed by risk reduction computations based on the assumed performance profile of the individual tools, and then selecting according to the required risk reduction to achieve the agency-derived maximum acceptable risk, with preference for the simpler tool when possible.

Discussion

The proposed model provides a flexible framework in which to build assessment methods for data. The modular structure of the model allows for substitution of a number of different components, for example to allow for real-time environmental observations, or statistics on the dynamics of particular ships, or ship classes, if they are available. The primary constraint on this data-driven flexibility, however, is that calibration of the models is an important consideration if the results are to be plausible. Use of AIS data to assess the shipping in the area, and to update the assumed object models based on the observed draft of ships does much to reduce the model's dependence on *a priori* assumptions, but does not remove them completely. Reliability of AIS data is also notoriously poor. Calibration, therefore, is expected to be the biggest constraint on applicability of these models in practice, although computational cost is also a consideration. The algorithm is very readily parallelisable, however, and therefore suited to distributed computation if there are significant speed constraints.

Apart from practical issues of implementation, this approach to survey assessment requires a reassessment of how survey results are reported. In particular, it requires the sponsoring agency to accept that there is a level of risk associated with the result of any survey, no matter how good, and that there may be areas that remain high risk even with fresh surveys. This has been problematic in the past, but better understanding of uncertainty models in general may have moved this idea towards greater acceptance. Pragmatically, but possibly most problematically, the method also requires that there is an agency-derived maximum allowable risk, or, equivalently, a minimum risk that is considered unimportant, so that there is a finite stopping point for the survey effort. Although this in many ways can make more sense than defining survey adequacy with respect to behaviours of the individual measurements, or statistics of the derived products, it is also potentially an assumption of liability that might be hard to swallow. Of course, that liability is already present, implicit in the survey practices that are currently used. Assuming that it is not there because it is not visible does not make it so.

Methods of this type, if adopted, have other benefits to practice. An immediate product of these types of methods is an assessment of the risks associated with transit through a given area of the chart; this was, in fact, the original design goal for the methods described. As an assessment of the (generalised) uncertainty of the area for communication to the charting end-user, this has much to recommend it. It is relatively simple to explain (at least in interpretation, if not in details); it is scalar and therefore readily provided to the user as, for example, colour overlays or contoured vector areas; and it represents not an assessment of what was done by the surveyor (which is of limited utility to the user) but an integrated assessment of the state of knowledge of the area, which more readily addresses the sorts of questions that a user might legitimately pose. Habilitating the term “risk” in the context of charts might, however, be an obstacle in adoption.

A further implication of the methods outlined here is that it is possible that areas of a survey might either not be covered (because no tool available will improve the situation), or that some areas might receive at best “good enough” survey, with the potential for residual undetected objects to be present (albeit ones that are extremely unlikely ever to interact with surface shipping). While this seems intellectually allowable, and even in some senses just a codified expression of current practice, it could be argued to be contrary to the fundamental principles of hydrographic survey, or at least to the surveyor’s soul. It remains to be seen whether the community is ready to accept methods that maybe meet the letter, but possibly not the spirit, of the requirements that they themselves set.

One possible objection to relying on AIS traffic to assess the types of ships in the region is that this is not static. That is, if the conditions for completion of survey include the idea that detailed investigation of objects that pose no risk to current traffic, what does that say about the survey if the traffic changes? While a valid concern, this does not seem to offer any more difficulty than is currently faced when the usage of a port or approach changes and the current charting information is dated. That is, if a new container port were completed, and bigger ships were expected, it would be usual to consider whether the hydrographic data were sufficient to support the ships, and commission a new survey if not. The approach using the proposed methods would not differ significantly, except that it might be considered to be more objective: at least, given these methods, it is known what the conditions for completion were, making it simpler to identify areas which might be of concern.

References

[1] Int. Hydro. Org., IHO Standards for Hydrographic Surveys, 5ed. IHO, Monaco, 2008.

[2] National Oceanic and Atmospheric Administration, Hydrographic Survey Specifications and Deliverables. NOAA, Silver Spring, MD, March 2016.
[3] Calder, B. R., Uncertainty Representation in Hydrographic Surveys and Products. Proc. 5th Int. Conf. on High Resolution Surveys in Shallow Water, Portsmouth, NH, 2008.
[4] Calder, B. R., On Risk-Based Expression of Hydrographic Uncertainty. Marine Geodesy, DOI 10.1080/01490419.2014.933141, 2014.
[5] Calder, B. R., Assessing Resurvey Priority for a Chart Portfolio. Proc. 7th Int. Conf. on High Resolution Surveys in Shallow Water, Plymouth, England, 2015.
[6] Silver, A. L. and J. F. Dalzell, Risk-based Decisions for Entrance Channel Operation and Design. Int. J. Offshore and Polar Eng., 8(3):200-206, 1998.
[7] Cartwright, D. E. and M. S. Longuet-Higgins, The Statistical Distribution of the Maxima of a Random Function. Proc. Roy. Soc. Lond. A, 237(1209):212-232, 1956.
[8] Ochi, M. K., On Prediction of Extreme Values. J. Ship Res., 17:29-37, 1973.
[9] Cressie, N. and C. Wikle, Statistics for Spatio-Temporal Data. Wiley Series in Probability and Statistics, Hoboken, NJ, 2011.
[10] Calder, B. R., On the Uncertainty of Archive Hydrographic Data Sets. IEEE J. Oceanic Eng., 31(2):249-265, 2006.
[11] Zambo, S. J., P. A. Elmore, A. L. Perkins, and B. S. Burgeois, Uncertainty Estimation for Sparse Data Gridding Algorithms. Proc. U.S. Hydro. Conf., National Harbor, MD, 2015.
[12] Hare, R., A. Godin, and L. A. Mayer, Accuracy Estimation of Canadian Swath (Multibeam) and Sweep (Multi-transducer) Sounding Systems. Tech. Rep., Canadian Hydro. Service, Ottawa, ON, 1995.
[13] Lurton, X. and J.-M. Augustin, A Measurement Quality Factor for Swath Bathymetry Sounders. IEEE J. Ocean. Eng., 35(4):852-862, 2010.
[14] Calder, B. R. and L. A. Mayer, Automatic Processing of High-rate, High-density Multibeam Echosounder Data. Geochem., Geophy., Geosyst., 4(6), DOI 10.1029/2002GC000486, 2003.
[15] Calder, B. R. and K. Schwehr, Traffic Analysis for the Calibration of Risk Assessment Methods. Proc. U.S. Hydro. Conf., Norfolk, VA, 2009.

Author Biography

Brian Calder (brc@ccom.unh.edu, +1 (603) 862-0526) is an Associate Research Professor and Associate Director at CCOM/JHC (UNH, USA). He has a Ph.D. in Electrical and Electronic Engineering, completing his thesis on Bayesian methods in SSS processing (1997). He is currently focusing on statistically robust automated data processing approaches and tracing uncertainty in hydrographic data.

Aknowledgements

This work was sponsored by NOAA grants NA10NOS4000073, and NA15NOS4000200, which are gratefully acknowledged.