ELICITING EXPERT JUDGEMENT FOR THE PROBABILITY OF AUV LOSS IN CONTRASTING OPERATIONAL ENVIRONMENTS

G Griffiths[1] and A. Trembanis[2]

[1]National Oceanography Centre, Southampton, UK
[2] Department of Geological Sciences, University of Delaware, USA

[1] Empress Dock, Southampton SO14 3ZH UK T: +44 (0) 23 80596004 E: gxg@noc.soton.ac.uk
[2]Newark, Delaware 19716 USA T: +1 302 831 2569 E: art@udel.edu

Abstract

Each time an autonomous underwater vehicle (AUV) is used in the sea there is a non-zero probability of loss. Quantifying probability of loss is not an exact science; therefore much depends on the fault history of the vehicle, the operational environment and the complex relationships between the consequences of faults or incidents and the environment. While this problem may be stated in scientific terms, in practice, there is no solution through scientific means alone. This is an example of 'trans-science'. We suggest that an approach based on the formal process of eliciting expert judgement may be an effective means of approaching this problem, as the process has been used successfully for other trans-scientific questions. The paper provides an introduction to the process of eliciting expert judgement, outlines four exemplar environments: coastal, open water, under sea ice and under shelf ice, and gives a worked example of one expert's judgement on the probability of loss in the four environments arising from a real fault with the Autosub1 AUV. Using the fault history of the Autosub3 AUV, included in the Annex, we ask experts from among UUST attendees (and others) to take part in this expert judgement elicitation. Based on the results of this elicitation we aim to publish a paper in the peer-reviewed literature.

1. Introduction

On 16 February 2005 the Autosub2 AUV was lost beneath 250 m of ice, some 15 km from the seaward edge of the Fimbulisen ice shelf in Antarctica (Nicholls et al., 2006). The subsequent formal Loss Inquiry made a series of recommendations (Strutt, 2006). Two key recommendations were that a risk management strategy, tuned to the needs of an AUV group, should be developed, and that reliability analyses should be undertaken prior to future Autosub campaigns. With Autosub3 now in service, and with scientists funded for research that would take the vehicle to Antarctica - under an ice shelf - it is imperative that a Risk Management Process (RMP) suitable for AUV operations in hazardous environments be developed, tested, and applied. Moreover, we recognise that under ice is not the only hazardous environment, for example, operations in coastal waters, with attendant shipping traffic, shoals, the shoreline and environmental factors such as fog, also present challenging hazards for safe and successful AUV operation. Therefore for an RMP to be useful it should be capable of addressing risk in varied operating environments.

A draft RMP-AUV has been designed along the lines outlined in Strutt (2006). Embedded within the RMP-AUV is a mechanism for the responsible owner to make decisions on the acceptable risk of proposed AUV campaigns (Trembanis and Griffiths, 2006; Griffiths and Trembanis, 2007). The process assists the owner in reaching a decision by deriving a quantitative estimate of the acceptable risk of loss based purely on financial considerations that include: vehicle capital and operating costs, use to date (depreciation) and appetite for risk. Against this *acceptable* risk of loss the owner needs to compare the *likely* risk of loss for the vehicle on the proposed campaign. This risk comes from the operating environment (e.g. open water, coastal, under sea ice, under shelf ice) and from the consequence of faults/incidents with the vehicle.

There are difficulties in providing quantitative estimates for both these factors. Arriving at quantitative estimates of these factors is an example of problems termed 'trans-scientific' by Weinberg (1972)[1], in that the problems can be stated in scientific terms but, in practice, they cannot be solved through scientific means alone. Risk from the operating environment is generally poorly understood and rarely quantified, although a ranked list of risk from the environment could well be agreed – under

[1] This is not an easy text to find in the original, see http://garfield.library.upenn.edu/classics1991/A1991GB06700001.pdf for a commentary.

shelf ice is likely to be the highest risk and open water the least. With the vehicle, there are two major issues. First, determining the reliability of the vehicle and second translating reliability into risk of loss. Stokey et al. (1999) tackled the reliability of early Remus vehicles in a qualitative manner. Griffiths et al. (2003a, 2003b) and Podder et al. (2004) used a statistical approach for Autosub2 and Dorado respectively, while Chance (2003) plotted the pragmatic quantity 'availability' against time for their HUGIN AUV. A formalized statistical approach for reliability analysis of the 'Fetch' class DOERRI AUV has also recently been adopted and implemented (Trembanis and Griffiths, 2006). Despite these analyses, the statistical approach used with Autosub2 has been criticised by some for being based on only one vehicle. However, we contend that there is merit in a thorough statistical analysis of the through-life reliability and fault history of a single entity such as Autosub.

The second major issue, translating reliability estimates into risk of loss, has also been contentious. It relies on expert judgement on the significance and potential impact of faults or incidents, taking into account the operating environment, an assessment that may be rife with speculative interpretation. In Griffiths (2003a) for an under sea ice campaign, and in Griffiths and Trembanis (2007) for an under shelf ice campaign, the assessment expert forming the judgement (Griffiths) was closely associated with the Autosub development team. This is clearly open to criticism. However, it should be noted that the engineering trials and the reliability assessment were undertaken as distinct and separate tasks. The technical team prepared and conducted the trials campaigns and provided a written report on each mission. Griffiths performed an initial analysis alone, but refined his initial estimates through back and forth communication with the technical team. Trembanis considers that this approach is a strength and not a liability. However, the expert judgement process itself *was* only semi-formal, certainly when compared to the process characteristics described by Otway and von Winterfeldt (1992) and outlined in sections 2 and 3 below.

As a response to these criticisms, in this paper, our aims are:

(a) To use a more formal approach to eliciting expert opinion on the risk of loss of Autosub3 in different operating environments.
(b) To widen the pool of expert opinion brought to bear on this subject.
(c) To engage with the AUV community on the usefulness of eliciting expert judgement in risk management.

There is an extensive literature on eliciting expert judgement (e.g. O'Hagan et al. (2006) and its 39 page bibliography). A substantial body of work on eliciting expert judgement arose from major studies on nuclear reactor safety after the Three Mile Island accident and other events in the United States in the 1980s (Keeney and von Winterfeldt, 1991; Otway and von Winterfeldt, 1992). More recently, O'Hagan et al. (2006) give examples of expert elicitation in medicine (e.g. diagnosis and treatment decisions, clinical trials, survival analysis), veterinary science, agriculture (e.g. crop yields), meteorology (e.g. severe weather conditions), business studies, economics and finance (e.g. outcomes of organisational change, error rate in auditing), engineering (e.g. structural safety). One particularly good engineering example is that of eliciting beliefs on the maintenance needs and costs associated with water treatment plants.

In this paper we will take ideas from the expert judgement literature and adapt them to use with AUVs. Our approach is to ask a wide range of experts from the AUV community (at UUST 2007 and elsewhere), with a diversity of backgrounds and opinions, to receive a degree of training in expert elicitation (this paper, its references and the associated presentation) and to complete a *pro forma* questionnaire (Annex A). This questionnaire includes a list of all faults and incidents with Autosub3 to date as the input data and asks the experts to (a) assess probabilities of the faults/incidents leading to loss in different environments and (b) to assess their own level of confidence in making each assessment. It is our intention to use the experts' assessments, anonymously and aggregated, as the basis for a journal paper. We welcome feedback on the usefulness of this approach and on what should be included for the journal paper.

2. Expert Judgement

2.1 Background/Types of approach

"Engineering judgement is often applied to bridge the gap between hard technical evidence … and unknown characteristics of a technical system"
Cooke and Goossens (2004).

Intuition and judgement permeate all scientific and engineering analysis from very basic decisions such as what to study and what techniques to adopt to more complicated assessments such as safety and forecasting (see Otway and von Winterfeldt, 1992 and O'Hagan et al., 2006). One approach to dealing with inherent complexity and uncertainty is through the utilization of expert judgement. Expert judgement is a process by which the opinions of experts are brought to bear on issues that involve some measure of science/engineering and policy. In basic terms, expert judgement is any process in which one undertakes consultation with one or more experts that have experience with similar projects to your own.

Expert judgements can, and routinely are, employed in a host of varying manners, from round table discussions (Sachman, 1974) to more formalised forecast assessments such as the Delphi Method (Linstone and Turoff, 2002). There is an extensive body of literature regarding expert judgement and the curious reader is directed to the work by Otway and von Winterfeldt (1992) (and references therein) as a well-written introduction to the topic.

Expert judgement requires the synthesis of opinions of experts in a subject where there is uncertainty due to insufficient data, when such data is unattainable because of physical constraints or lack of resources (Otway and von Winterfeldt, 1992; O'Hagan et al., 2006). Expert judgement or expert elicitation is essentially a scientific consensus methodology. Expert judgement is often used in the study of rare and/or highly controversial events. Expert elicitation allows one to parameterize and quantify the uncertainty as an 'educated guess', for the topic under study.

2.2 Process- steps 1-7

In their 1992 paper, Otway and von Winterfeldt enumerate seven stages to the process of expert judgement. In the following section we outline briefly each step of the process we intend to follow in application of this project. Note that although each stage is described, by the very nature of our paper, not all of the stages can be completed until after the completion and subsequent analysis of the *pro forma* questionnaires, in other words the completion and success of this venture depends on the participation of UUST attendees.

3. Applying Expert Judgement

Elicitation of expert opinion through a questionnaire is acknowledged to be more difficult than through a face-to-face, one-on-one interview, O'Hagan et al. (2006: 26). However, by keeping the question to be asked simple, we intend to avoid the pitfalls of potential misunderstanding. Asking for self-assessment on the level of confidence for each estimate also reduces the aggregated effect of those judgements where the expert feels less certain. We have also tested the questionnaire and the description of faults and incidents on three graduate students with some knowledge of AUVs (one in the UK and two in the US). We have incorporated their feedback on the questionnaire.

3.1 The Issues

Given the set of facts on faults and incidents with Autosub3 throughout its life to date, described in Annex A, we seek to predict the probability of loss of the vehicle in different operating environments. At issue is how likely is it that each fault or incident, taken in isolation, but with the expert's knowledge of the wider issues, could lead to loss in the four example environments. The actual question to be asked of each fault or incident is set out formally in section 3.3.

3.2 Selecting Experts

For many (but by no means all) of the judgements asked for on individual faults or incidents there is likely to be a degree of uncertainty over the response. It is here that the experience, background, and insight of the individual expert are most important. As a consequence, the success (or not) of the elicitation process is strongly dependent on the knowledge of the experts. Ideally, according to O'Hagan et al. (2006:27), each expert (a) has specific technical and domain knowledge (e.g. closely involved in AUV design or operations), (b) is able to approach a problem via formal principles (e.g. through causal reasoning – the analysis of cause and effect), (c) uses established strategies (e.g. questioning/reviewing first assessments) and (d) relies more on procedural knowledge (e.g. relationships and an appreciation of what is important) and less on declarative knowledge (e.g. facts and simple rules). At the highest level of expertise, there is agreement in the literature that judgement is intuitive, with "an automaticity of action deriving from a wealth of knowledge and experience", that may typically take ten years to gather (O'Hagan et al., 2006:54). Experts should also have a realistic view of their competence for each particular problem.

Clemen and Winkler (1985) examined the precision and value of information elicited from dependent and independent sources. If the experts within a pool have limited diversity or a strong dependence (e.g. from one organisation, or all academics), they concluded that this would "have a serious detrimental effect on the precision and value of the information". Our aim is to maximise independence, with experts from different backgrounds, areas of expertise, nationality etc.

The preceding paragraphs well describe many participants at UUST, and it is from such a cohort that we seek volunteer experts to take part in this study.

3.3 Clearly define issues

One of the key stages in the expert elicitation process is the definition of the problem or issue to be judged.

For the purposes of this paper we wish to make the stated issue as clear and concise as possible. In the course of evaluating each fault log entry, the expert respondee is asked to assess the following question, "*What is the probability of loss of the vehicle in the given environment X given fault/incident Y*?"

This question is the key yardstick for the evaluation process and a strict and consistent adherence to this question will help to maintain a level of consistency between responses and respondees. It is important also to note that our interest in this matter is with respect to the impact of the fault on loss of the vehicle, not, for instance, on the impact that the given fault might have on science delivery, but rather, will this fault lead to the loss of the vehicle as a complete system given the environmental information and one's own expert opinion.

3.4 – 3.5 Training the Experts and Eliciting Judgements

One of the main focus points of the symposium presentation will be to provide brief training to the attendees in the completion of the *pro forma* questionnaire in Appendix A.

The literature of expert elicitation acknowledges that the precision of estimates is improved if experts have access to independent information, to allow a degree of calibration. We have sought, with limited success, such independent information. First, for open water and coastal environments Leviathan, a leading marine insurance binding authority, have stated that they have not paid out on an AUV loss in the last two years[2]. Second, out of some 150 vehicles produced by Hydroid, and used in open and coastal waters, and under sea ice, we believe that none have been lost[3]. Third, through the early stages of Seaglider development and operations, eight out of the first ten vehicles were lost, in environments that ranged from open water to areas infested with sea ice; of the next twelve built, two were lost as of September 2005[4].

In order to be of most use to the process it is important that those who graciously agree to conduct the expert assessment complete the entire questionnaire. The questionnaire is envisioned, and has been tested, to take approximately 3-4 hours in total although it should be stressed that there is no time constraint for its completion. The fault/incident descriptions, it should be noted, are the distillation of trials and science missions reports (by Griffiths) and thus are by nature concise. Where our students felt the initial draft was too terse, we have expanded the fault/incident descriptions. It is therefore left to the expert assessment of the respondee to determine the impact of the given fault. If for some reason one does not feel that sufficient description is available then this can be reflected both the confidence level of the probability assessment and the comments after each assessment.

It is also important for the elicitation process that we have a clear sense of the backgrounds and expertise of the respondees, therefore the favour of a reply to the section in Appendix A entitled "*Expert Details*" is appreciated. It is not necessary, although welcome, for respondees to include their name and contact details, but anonymous responses are in order. Those wishing to make electronic submissions are invited to download the MS Word file from www.noc.soton.ac.uk/OED/gxg/UUSTRiskPaper.html or to contact either of the authors.

3.6 Analyzing and Aggregating

Research has shown that many experts, when asked to use the full probability range, tend too often to opt for values close to 1 or 0[5], O'Hagan et al. (2006: 68). Furthermore, there is evidence that an expert's ability to provide unbiased estimates shows no correlation with the expert's technical or domain expertise. However, if experts are aware that particular types of faults or incidents have led to loss, or not, their subjective judgements may be less biased. This outcome feedback is clearly important, and it argues for open dissemination of faults and loss within the AUV community.

Handling differences of opinion and a range of subjective probabilities is easier than identifying bias. In their review of combining probabilities from experts, Clemen and Winkler (1999) describe mathematical and behavioural combination techniques. The Autosub Loss Inquiry used a behavioural approach, requiring the experts gathered together to interact and produce a single, agreed, group judgement (Strutt, 2006). This approach is not without its problems, including group polarisation (or 'group-think').

Where experts do not exchange information, mathematical combining techniques are appropriate. While current research considers Bayesian belief nets to provide a mathematically defensible, rigorous and

---

[2] Personal communication, Keith Broughton of Leviathan with Griffiths, June 2007.

[3] Personal communication, Graham Lester of Hydroid with Griffiths, July 2007. This was the case after a REMUS 100 AUV 'lost' for 10 months was recovered recently, essentially intact.

[4] Persoanl communication Charles Eriksen with Griffiths, September 2005.

[5] Indeed Griffiths et al. (2003a) *only* considered 1 or 0 as possible outcomes, while Griffiths and Trembanis (2007) considered only 0, 0.25 or 1.

effective way of combining judgements (O'Hagan, 1998; Sigurdsson et al., 2001), they are challenging to implement. As a consequence, our initial approach will be to use a simple linear or logarithmic opinion pool, Clemen and Winkler (1999):

$$p(\theta) = \sum_{i=1}^{n} w_i p_i(\theta) \qquad p(\theta) = k\prod_{i=1}^{n} p_i(\theta)^{w_i}$$

where w_i is the weight given by expert i (of n) for the probability $p_i(\theta)$.

3.7 Complete analysis and write up

The complete analysis will document the aggregated experts' judgements on the probabilities of leading to loss for each fault and each environment as opinion pool means and a measure of spread. Importantly, the reasons why experts arrived at their judgements will be summarised. Using these sets of probabilities, and example AUV campaigns for each environment, we will model the overall probability of losing a vehicle in each campaign using Kaplan Meier and Weibull methods as used in Griffiths et al. (2003a). Probabilities of loss will be compared with data from independent sources (if available) for coastal/open water environments, and with earlier single-expert predictions in Griffiths and Trembanis (2007) for under ice.

At this level of detail, which we suggest is necessary for this first analysis of AUV faults using formal expert elicitation by questionnaire, the results will be published as a National Oceanography Centre research report and made freely available[6]. A journal paper will be written using distilled information, describing the method and the results.

4. Environments

We have chosen four contrasting environments as examples for this study. They were chosen because they are well known to us and they represent both common and challenging AUV operating environments. Clearly the method can be applied in other settings, such as near the seabed, in complex terrain, or within enclosed environments such as pipes, cenotes, or lakes. In the following sub-sections are concise notes on key factors from each environment that may effect experts' judgements on probability of faults or incidents leading to loss.

There are some factors that are common to one or more environments. Perhaps the most significant is the process of launch and recovery, frequently from a ship. Incidents during launch and recovery are not uncommon; they can, and have, led to loss or write-off. The occurrence and impact of such incidents has been sufficiently high that some insurance providers have suggested co-insurance, or risk sharing, during these specific parts of a mission (Griffiths et al., 2007).

4.1 Open Water

Open water, away from the coast and traffic lanes, where the water depth is less than the crush depth of the vehicle, forms a relatively benign operating environment. An emergency response of rising to the surface, or descending to the seabed, is feasible, and from either location telemetry of data and position is possible. Clearly the risks are higher if the water depth exceeds the crush depth. While hazards mid-water are few, on the surface high winds and/or waves, fog and other vessels may increase risk and the consequences of technical failures in navigation or communication systems. Operating close to the seabed can be hazardous, placing reliance on collision avoidance or altitude-sensing hardware, algorithms and software.

4.2 Coastal

Coastal settings, defined as waters from the shelf edge (150-200 m water depth) and landward towards the shore, and including inland waters, can be challenging locations for AUV operations. While well below crush depth, many challenges remain. This setting includes shipping lanes and bay mouths as well as the near-shore (just outside of the surf zone), and estuaries. Physical hazards in this setting include high density ship traffic comprising, among others, commercial, military, and personal watercraft; divers (recreational and commercial) (Patterson, Sias, and Gouge, 2001); engineering structures (e.g. bridges, breakwaters, piers, jetties, groins, etc.); fishing gear (e.g. pound nets, lobster/crab pots. Environmental hazards include turbid waters and strong fluid flows (currents and waves) that make search and recovery problematic. Coastal settings do, however, afford a host of launch/recovery options including ships, boat ramps, docks and piers, which can be used in tandem or switched to mid-mission as conditions require.

Shallow depths and strong hydrodynamic flow present increased risk for collision and thus place added importance on collision avoidance systems. The rapid spatial and temporal changes to environmental conditions in coastal settings also place a premium on navigation and communication systems. The proximity to logistical centres, however, does provide advantages for operational adjustments (e.g. operations can be moved to more benign locations and additional support supplies can be more readily acquired).

[6] A pdf will be available via http://eprints.soton.ac.uk

For the purposes of the questionnaire, we ask experts to consider a semi-open, highly developed coastal embayment with depths of 40 m maximally, relatively sheltered from waves but subject to tidal currents of ~1-1.5 $m.s^{-1}$. Vessel traffic includes commercial and recreational vessels and occasional personal watercraft.

4.3 Sea ice and icebergs

Sea ice and icebergs pose a wide spectrum of risk that merits an expert elicitation study in its own right. There are numerous classes or types of sea ice, and each may pose a threat of some magnitude to AUV operations. Ice types are described by Wadhams (2000), and MacDonald (1969) described how ice affects vessel operations. More specific information on ice types and their effect on AUV operations is available on the Polar AUV Guide website[7].

Sea ice and icebergs pose a hazard to AUV operations for several reasons:

- Ice can form a rigid lid to the ocean, hampering or even preventing recovery after a technical failure or incident.
- Afloat, deep ice keels and icebergs pose collision hazards. If in shallow water, especially if they are grounded, ice keels and icebergs may test severely the collision avoidance and path planning systems within an AUV.
- Thin ice may pose different hazards: semi-transparent grease ice may be sufficient to hamper visual sighting on recovery; nilas, up to 10 cm thick, may damage appendages such as antennas.
- Continuous multiyear ice, such as fast ice or sikussak can form a barrier as effective as an ice shelf (see section 4.4) should an AUV become stranded, especially if the support vessel has limited icebreaking capability.
- Ice need not be continuous to pose a threat; brash ice can be a hazard during launch and recovery, especially to appendages and propeller blades.

An important factor affecting the level of risk posed by sea ice is the icebreaking capability of the support vessel as this affects the likelihood of success or failure should recovery from under ice become necessary. The risk appetite and time allocated for search and recovery are also factors, as are the availability of supporting tools such as an emergency location beacon on the vehicle and whether an ROV is on board the vessel to aid recovery.

Because of the wide range of risks, for the purpose of this study, we ask that experts focus on a scenario where first year ice dominates (0.3–2.0m thick), with ice keels to 15m, and sporadic icebergs and a support vessel able to break 2 m ice at 2kt.

4.4. Shelf Ice

Ice shelves are the floating edges of continental ice sheets, and, with a typical thickness of 180 m at the seaward edge, form an impenetrable barrier. If an AUV becomes stranded under an ice shelf through a fault or incident, the chance of recovery must be almost zero. An ROV recovery might be possible if the stranding was no more than a few hundred metres from the ice front. Further in, it is possible to drill through the ice (e.g. using hot water), and if the AUV position is known accurately a recovery might be possible. However, such operations are very costly and involve complex logistics.

Experts should bear in mind, as outcome feedback, that only two AUVs have ever attempted under ice shelf missions, and both were lost, one on its first such mission, the other (Autosub2) on its second.

5. Completing the Questionnaire

5.1 Guidelines to completing the questionnaire

Experts are asked to adhere to the following guidelines in completing the *pro forma* questionnaire:

- Plan to allot between 2-4 hours for completion of the questionnaire. There is no time limit; this is merely a suggestion for planning purposes.
- Please remember to include a confidence index on your response to each fault/incident. Confidence indices range between 1-5 with 1 being a low level of confidence in the assessment and 5 being a high level of confidence.
- Remember to assess the fault/incident with respect to probability of loss of the asset (vehicle) not simply as a subsystem fault or lack of data delivery etc.

Note that the probability estimate is left to the discretion of the respondee with the caveat that values are within the natural range of zero and unity. Examples of probability responses are: 1/10 (e.g. the given fault is likely to lead to the loss of the vehicle in 1 out of every 10 missions); 1/100 (e.g. the given fault is likely to lead to the loss of the vehicle in 1 out of every 100 missions) and so on. Either fractional or decimal probabilities are acceptable responses.

[7] www.srcf.ucam.org/polarauvguide/environment/icetypes.php

			Estimated probability of leading to loss (on scale of 0 to 1), with confidence level (1 to 5) for each estimate in the grey boxes.				
No.	**Dist (km)**	**Fault/incident description**	**Open**	**Coast**	**Sea Ice**	**Shelf Ice**	**Reasons**
186	34	A software bug that manifested itself seven hours after launch (and 34km travelled) meant that the vehicle was stuck in an oval pattern (125 by 75m) at its correct operating depth. The course angle relative to a cardinal direction affected the time taken for the fault to emerge. 7hr was shortest time.	0.001	0.003	0.1	0.7	Ice Shelf: if occurs, a sure loss, but reduced from 1 as required mission length may be less than time to fault emergence. Under sea ice, rescue very possible, but not certain. For loss in open water, fault would need to be compounded by failure of one or more of acoustic beacon-emergency release-ARGOS on surface. Coastal assessed higher as extra time on surface exposes to higher hazard.
			4	3	3	4	

Table 1. An example of a completed fault/incident entry from mission 186 of Autosub1 during a campaign in the open waters of the North Sea in 1999. The vehicle was easily found on this occasion by listening on the RV *Scotia* for its acoustic pinger.

5.2 Example Assessment

Table 1 shows what we would consider to be a well-completed expert assessment of a real fault that happened on Autosub1 in the North Sea in 1999. The reasons for the assessments call for knowledge of the usual systems to be found on an AUV rather than necessarily requiring detailed knowledge of the particular sub-systems on Autosub. If general knowledge is not adequate in particular cases, then an assessment should be attempted, but a lower confidence level assigned.

6. Conclusions

We consider that recording fault histories for AUVs is important. It is part of good practice in providing immediate feedback to the operating teams on performance and reliability and it also leads to an ability to model statistically the reliability of one or more vehicles.

In previous papers, we have shown how informal expert judgement can be used to estimate probability of loss from knowledge of the vehicle fault history. However, such ideas have not been without controversy; as a consequence, in this paper, we have set out a more formal approach to eliciting expert judgement based on widely accepted practices following a substantial review of the literature.

Through presenting a full summary of the fault history of the Autosub3 AUV, and an introduction to expert judgement elicitation, our aim is to obtain a broad-based expert assessment of the probability of loss of the vehicle in different operating environments. As an academic exercise, the process and our findings will form the basis of a journal paper. More practically, returns from experts will be used within the existing Risk Management Process – AUV to better inform the Director NOCS as to the likely risk of loss in sending Autosub3 beneath ice. Results will also inform the technical team as to which faults/incidents a wide cohort of experts consider most likely to lead to loss, and hence which areas need to be given priority.

Acknowledgments:

We are grateful to the members of the Autosub technical team for the source data from their fault and incident logs. Our students Kieran Rutherford, Adam Skarke and Robert Bogucki provided valuable comment on the draft Annex.

7. References

Chance, T. S., 2003. AUV surveys – extending our reach, 24000 km later. *Proc 11th International Symposium on Unmanned Untethered Submersible Technology*, AUSI, New Hampshire.

Clemen, R. T. and Winkler, R. L., 1985. Limits for the precision and value of information from dependent sources. *Operations Research*, 33(2): 427–422.

Clemen, R. T. and Winkler, R. L., 1999. Combining probability distributions from experts in risk analysis. *Risk Analysis*, 19(2): 187–203.

Cooke, R. M. and Goossens, L. H. J., 2004. Expert judgement elicitation for risk assessments of critical infrastructures. *Journal of Risk Research*, 7(6): 643–656.

Griffiths, G., Millard, N. W, McPhail, S. D., Stevenson, P. and Challenor, P. G., 2003a. On the Reliability of the Autosub Autonomous Underwater Vehicle. *Underwater Technology,* 25(4): 175-184.

Griffiths, G., Millard, N. W., McPhail, S. D. and Riggs, J., 2003b. Effect of upgrades on the reliability of the Autosub AUV. *Proc 13th International Symposium on Unmanned Untethered Submersible Technology*, AUSI, New Hampshire.

Griffiths, G. and Trembanis, A., 2007. Towards a risk management process for Autonomous Underwater Vehicles. Pp. 103–118 in: Griffiths, G. and Collins, K., *Proceedings of a Masterclass on AUV Technology for Polar Science*. Society for Underwater Technology, London.

Griffiths, G., Bose, N., Ferguson, J. and Blidberg, R., 2007. Insurance for autonomous underwater vehicles. *Underwater Technology*, 27(2): 43–48.

Keeney, R. L. and Winterfeldt, D. von, 1991. Eliciting probabilities from experts in complex technical problems. *IEEE Transactions on Engineering Management*, 38(3): 191–201.

Linstone, H.A. and Turrof, M. eds., 2002. *The Delphi Method: Techniques and Applications,* Addison-Wesley, Englewood Cliffs.

MacDonald, E.A., 1969. *Polar operations*. US Naval Institute Press, Annapolis. 242pp.

Nicholls, K. W. et al., 2006. Measurements beneath an Antarctic ice shelf using an autonomous underwater vehicle. Geophysical Research Letters, 33, 2006, doi:10.1029/2006GL025998 4pp.

O'Hagan, A., 1998. Eliciting expert beliefs in substantial practical applications. *The Statistician*, 47(1): 21–35.

O'Hagan, A. et al., 2006. *Uncertain Judgements: Eliciting Experts' Probabilities*. John Wiley & Sons, Chicester. 338pp.

Otway, H. and Winterfeldt, D. von, 192. Expert judgement in risk analysis and management: Process, context, and pitfalls. *Risk Analysis*, 12(1): 83–93.

Patterson, M.R., Sias, J.H. and Gouge, D.V., 2001. AUVs and scientific diving: a looming conflict? *Journal of the Marine Technology Society,* 34: 75-81.

Podder, K., Sibenac, M., Thomas, H., Kirkwood, W. and Bellingham, J., 2004. Reliability growth of Autonomous Underwater Vehicle – Dorado. *Proc. Oceans 2004*, Kobe, Japan. MTS/IEEE. Pp. 856-862.

Sackman, H., 1974. *Delphi Assessment: Expert Opinion, Forecasting and Group Process*. Report R-1283-PR, RAND Corporation.

Sigurdsson, J. H., Walls, L. A. and Quigley, J. L., 2001. Bayesian belief nets for managing expert judgement and modelling reliability. *Quality and Reliability Engineering International*, 17: 181–190.

Stokey, R., Austin, T., von Alt, C., Purcell, M., Goldsborough, R., Forrester, N. and Allen, B., 1999. AUV bloopers or why Murphy must have been an optimist. *Proc 11th International Symposium on Unmanned Untethered Submersible Technology*, New Hampshire, AUSI, pp. 32-40.

Strut, J. (editor), 2006. *Report of the inquiry into the loss of Autosub2 under the Fimbulisen*. NOCS Research and Consultancy Report available as pdf at http://eprints.soton.ac.uk/41098/

Trembanis, A. and Griffiths, G., 2006. Extreme Environments and Extreme Science: Reliability and Risk Assessment for Autonomous Systems with Application to Polar Campaign, *Eos Trans. AGU, 87*(52), Fall Meet. Suppl., Abstract C41B-0325

Wadhams, P,, 2000. *Ice in the ocean*. Taylor and Francis, London. 364pp.

Weinberg, A., 1972. Science and trans-science. *Minerva*, 10: 209–222.

WMO, 1970. *World Meteorological Organisation sea-ice nomenclature, terminology, codes and illustrated glossary*. WMO/OMM/BMO 259, TP 145. Geneva, World Meteorological Organisation.

ANNEX A – THE FULL RANGE APPROACH

A Microsoft Word version of this questionnaire may be obtained from http://www.noc.soton.ac.uk/OED/gxg/UUSTRiskPaper.html

Expert Details	
Name (if willing):	Contact email (if willing):
Nationality or domicile:	Organisation type:
Area of Expertise:	Years of experience of AUVs:

Table 1 *Discovery* 295T July 2005 – AUV and other trials in the SW Approaches.

			Estimated probability of leading to loss (on scale of 0 to 1), with confidence level 1 (lo) to 5 (hi) for each estimate in the grey boxes.				
No.	**Distance (km)**	**Fault/incident description**	**Open**	**Coast**	**Sea Ice**	**Shelf Ice**	**Reasons**
384	1.5	Mission aborted (to surface) due to network failure. (Much) later tests showed general problem with the harnesses (bad crimp joints).					
		Loop of recovery line came out from storage slot, long enough to tangle propeller.					
385	15.2	Autosub headed off in an uncontrolled way, due to a side effect of the removal of the upwards-looking ADCP.					
386	26	GPS antenna failed at end of mission.					
387	27.2	Homing failed, and the vehicle headed off in an uncontrolled direction. Mission was stopped by acoustic command. Problem was due to (a) the uncalibrated receiver array, and (b) a network message (“homing lost”) being lost on the network.					

At this point, you may want to consider what your estimated probabilities of loss arising from each fault or incident would mean for this campaign, which was in open water. For each entry, subtract your probability of loss from one, to form probability of survival, then, multiply each together to give the campaign probability of survival. Is this overall probability of survival for these faults, in open water, sensible, in your expert judgement? If not, you may want to ‘recalibrate’ your judgement on individual faults or incidents and reassess the probabilities.

Table 2 *Terschelling* May 2006 – AUV trials in the SW Approaches.

No.	Distance (km)	Fault/incident description	Estimated probability of leading to loss (on scale of 0 to 1), with confidence level 1 (lo) to 5 (hi) for each estimate in the grey boxes. Open	Coast	Sea Ice	Shelf Ice	Reasons
388	0.5	Aborted after 4 minutes post dive, due to network failure. Logger data showed long gaps, up to 60s, across all data from all nodes, suggesting logger problem.					
		Depth control showed instability. +/- 1m oscillation due to incorrect configuration gain setting.					
389	3	Vehicle went into homing mode, just before dive and headed north. Vehicle mission stopped by acoustic command. It was fortunate that the ship-side acoustics configuration allowed the ship to steam at 9kt (faster rather than 6kt with the towfish) and catch the AUV.					
		Separately, homing mode not exited after 2 minutes, as expected. It will continue on last-determined heading indefinitely – a Mission Control configuration error.					
		Problem with deck side of acoustic telemetry receiver front end, unrelated to vehicle systems.					
391	31	ADCP down range limited to 360m, reduced accuracy of navigation.					
		GPS antenna flooded. No fix at end point of mission.					
		EM2000 swath sonar stopped logging during mission.					
392	32	As consequence of GPS failure on M391, AUV ended up 700m N and 250m E of expected end position.					
393	5	Acoustic telemetry giving poor ranges and no acoustic telemetry.					
394	3	Jack-in-the-box recovery <u>*float*</u> came out, wrapping its line around the propeller, jamming it, and stopping the mission. Caused severe problems in recovery, some damage to upper rudder frame, sub-frame and GPS antenna. Required boat to be launched.					

395	8	Jack-in-the-box *line* came out, wrapped around the propulsion motor and jammed.					
396	4	Current estimation did not work, because minimum time between fixes for current to be estimated had been set to 15min; leg time was only 10min. Mission stopped and restarted with configurable time set to 5min.					
397	4	Main lifting lines became loose, could have jammed motor.					
398	8	Operators ended mission prematurely, they believed the AUV was missing waypoints. In fact, a couple of waypoints had been positioned incorrectly.					

Table 3 *Discovery* June-July 2006 – Biological measurements in the NE Atlantic

No.	Distance (km)	Fault/incident description	Estimated probability of leading to loss (on scale of 0 to 1), with confidence level 1 (lo) to 5 (hi) for each estimate in the grey boxes.				Reasons
			Open	Coast	Sea Ice	Shelf Ice	
401	7.5	Configuration mistake; ADCP up configured as down-looking ADCP causing navigation problems through tracking sea surface as reference. This data was very noisy and put vehicle navigation out by a factor of 1.5.					
		Damaged on recovery, "moderately serious" to sternplane, shaft bent.					
402	274	Stern Plane stuck up during attempt to dive, 2d 20h into mission. Stern plane actuator had flooded.					
		Abort due to network failure. Abort release could not communicate with depth control node for 403s. Possibly side-effect of actuator or motor problems.					
		Motor windings had resistance of 330 ohm to case. Propeller speed dropping off gradually during a dive					
		Only one position fix from tail mounted ARGOS transmitter.					
		GPS antenna damaged on recovery.					
403	140	Recovery light line was wrapped around the propeller on surface. Flaps covering the main recovery lines (and where the light line was towed) were open.					
		Took over 1 hour to get GPS fix at final waypoint.					
		Propeller speed showed same problem as m402. Subsequent testing of motor with Megger showed resistance of a few kohm between windings.					

404	75	Pre-launch, abort weight could not be loaded successfully due to distorted keeper. "If not spotted, could have dropped out during mission", considered low probability of distortion ***and*** not checked.					
		Pre-launch, potential short circuit in motor controller that could stop motor.					
		Propeller speed showed same problem as on m402 and 403.					
		CTD drop-out of 1 hour (shorter drop-outs noted in previous missions).					
		M404 recovery was complicated when lifting lines and streaming line became trapped on the rudder (probably stuck on the Bolen where the two were attached). Recovery from the situation required the trapped lifting lines grappled astern of the ship, attached to the gantry lines, and the caught end cut.					
		The forward sternplane was lost due to lifting line trapping between the fin and its flap on recovery.					
		The acoustic telemetry nose transducer was damaged due to collision with the ship.					

Table 4 *Terschelling* July 2006 – Turbulence studies in the Irish Sea

			Estimated probability of leading to loss (on scale of 0 to 1), with confidence level 1 (lo) to 5 (hi) for each estimate in the grey boxes.				
No.	**Distance (km)**	**Fault/incident description**	**Open**	**Coast**	**Sea Ice**	**Shelf Ice**	**Reasons**
405	2.5	Fault found pre-launch, LXT tracking transducer had leaked water – replaced.					
		Fault found pre-launch, starboard lower rudder and sternplane loose.					
406	104	AUV ran slower than expected and speed dropped off during mission, due to motor problem.					
		Current spikes of 3A and voltage drops in first part of mission.					
		Propulsion motor failed 500V Megger on recovery on windings to case.					
		One battery pack out of four showed intermittent connection.					
		Acosutci telemetry unit gave no replies.					
		On surfacing first GPS fix was 1.2km out.					
		Spikes in indicated motor rpm					
407	204	Acoustic telemetry unit gave no replies at all – no tracking or telemetry.					
		Noise spikes on both channels of turbulence probe data.					

408	302.5	Propulsion motor felt rough when turned by hand – bearings replaced before deployment.					
		Aborted at 50m due to overdepth as no depth mode commanded. Unless compounded by another problem, this would show itself immediately on first dive.					
		No telemetry from Acoustic telemetry unit.					
		Difficulty stopping Autosub on surface via radio command. Separate problems with the two WiFi access points.					
		Still spikes on motor rpm that need investigating.					

Table 5 ***Terschelling*** **March 2007 – Deep water AUV reliability proving trials in Norway.**

No.	Distance (km)	Fault/incident description	Estimated probability of leading to loss (on scale of 0 to 1), with confidence level 1 (lo) to 5 (hi) for each estimate in the grey boxes. Open	Coast	Sea Ice	Shelf Ice	Reasons
409	1.5	No acoustic telemetry or transponding. LXT ship side USBL receiver had leaked during mission giving poor bearings to sub, replaced with spare.					
410	9	No acoustic telemetry or transponding.					
411	128	No GPS fix at the end of the mission. GPS antenna bulkhead had water inside and had flooded.					
412	270	No GPS fix at end of mission. After next mission, GPS fixes started coming in after vehicle power up/power down; perhaps problem was due to initialisation with receiver – and not this time the antenna.					
		Problem at start for holding pattern. Holding pattern timed out due to programming mistake.					
415	6	Prior to dive, checks showed reduced torque on rudder actuator. Actuator replaced with new one - first use for this new design of actuator motor and gearbox. However, AUV spent most of mission "stuck" going around in circles at depth due to rudder actuator fault. The new actuator overheated, melting wires internally, the motor seized, and internal to the main pressure case, the power filter overheated. Some of the damage may have been caused by an excessive current limit (3A); correct setting was 0.3A. But this does not explain high motor current. Possible damage during testing when motor stalled on end stop? Compounded by wiring to motor held tightly to case with cable ties, and worse, covered with tape (acting as an insulator). Wires were not high temperature rated.					

415	6	Three harness connectors failed due to leakage, affecting payload systems: EM2000 tube, ADCP_down, and Seabird CTD. Despite connector problems the system worked without glitches and failed only when the power pins had burned completely through on the connector feeding power to the abort system					
		Although it worked properly at the start of the mission at a range of 1200m, the acoustic telemetry stopped working at the end of mission. Hence could not stop the mission acoustically when needed.					
416	18	Not possible to communicate with vehicle at 1180m depth; holding pattern caused a timeout, and AUV surfaced. Acoustic telemetry max range was 500m for digital data.					
418	15	When homing was stopped deliberately after 10 min, the AUV did not go into a “stay here” mode. Rather it continued on the same heading; stopped by acoustic command 500m from shore. Cause was incorrect configuration of mission exception for homing. Default in campaign configuration script was not set due to inexperience with new configuration tools.					