



Article (refereed) – Published version

Brito, M.P.; Smeed, D.A.; Griffiths, G. 2014 Underwater glider reliability and implications for survey design. *Journal of Atmospheric and Oceanic Technology*, 31 (12). 2858-2870. [10.1175/JTECH-D-13-00138.1](https://doi.org/10.1175/JTECH-D-13-00138.1)

This version available at <http://nora.nerc.ac.uk/502654/>

NERC has developed NORA to enable users to access research outputs wholly or partially funded by NERC. Copyright and other rights for material on this site are retained by the rights owners. Users should read the terms and conditions of use of this material at

<http://nora.nerc.ac.uk/policies.html#access>

© Copyright 2014 American Meteorological Society (AMS).
Permission to use figures, tables, and brief excerpts from this work in scientific and educational works is hereby granted provided that the source is acknowledged. Any use of material in this work that is determined to be “fair use” under Section 107 of the U.S. Copyright Act September 2010 Page 2 or that satisfies the conditions specified in Section 108 of the U.S. Copyright Act (17 USC §108, as revised by P.L. 94-553) does not require the AMS’s permission. Republication, systematic reproduction, posting in electronic form, such as on a web site or in a searchable database, or other uses of this material, except as exempted by the above statement, requires written permission or a license from the AMS. Additional details are provided in the AMS Copyright Policy, available on the AMS Web site located at (<http://www.ametsoc.org/>) or from the AMS at 617-227-2425 or copyrights@ametsoc.org.

Contact NOC NORA team at
publications@noc.soton.ac.uk

Underwater Glider Reliability and Implications for Survey Design

MARIO BRITO AND DAVID SMEED

National Oceanography Centre, Southampton, Hampshire, United Kingdom

GWYN GRIFFITHS

Autonomous Analytics, Southampton, Hampshire, United Kingdom

(Manuscript received 27 June 2013, in final form 25 April 2014)

ABSTRACT

It has been 20 years since the concept of the Autonomous Ocean Sampling Network (AOSN) was first introduced. This vision has been brought closer to reality with the introduction of underwater gliders. While in terms of functionality the underwater glider has shown to be capable of meeting the AOSN vision, in terms of reliability there is no communitywide hard evidence on whether persistent presence is currently being achieved. This paper studies the reliability of underwater gliders in order to assess the feasibility of using these platforms for future AOSN. The data used are taken from nonunderwater glider developers, which consisted of 205 glider deployments by 12 European laboratories between 2008 and 2012. Risk profiles were calculated for two makes of deep underwater gliders; there is no statistically significant difference between them. Regardless of the make, the probability of a deep underwater glider surviving a 90-day mission without a premature mission end is approximately 0.5. The probability of a shallow underwater glider surviving a 30-day mission without a premature mission end is 0.59. This implies that to date factors other than the energy available are preventing underwater gliders from achieving their maximum capability. This reliability information was used to quantify the likelihood of two reported underwater glider surveys meeting the observation needs for a period of 6 months and to quantify the level of redundancy needed in order to increase the likelihood of meeting the observation needs.

1. Introduction

There has been a significant increase in the use of autonomous underwater vehicles (AUVs) during the last decade and this trend seems set to continue (e.g., see the summary of a market report at <http://www.douglas-westwood.com/shop/shop-infopage.php?longref=902~0#.UVIK8BIBH8g>). We suggest there are two key reasons: first, ship-based field work is very expensive; and second, these vehicles, whether in industry or scientific research, have been shown to be capable of obtaining valuable data that augment preexisting means, such as moorings, towed systems, and profiling floats (see

special issue of *Limnology and Oceanography*, 2008, Vol. 53, No. 5, part 2). Nevertheless, it is appropriate to ask if they have yet fulfilled their true potential. Twenty years ago, Curtin et al. (1993) presented the requirements for an Autonomous Ocean Sampling Network (AOSN) that comprised several autonomous underwater vehicles and a distributed set of acoustic and point sensors to enable four-dimensional ocean sampling. Since then, several concerted efforts have developed and tested technology to implement this vision. One of the most significant developments is a class of autonomous underwater vehicles denoted as underwater gliders (Stommel 1989). These slow-moving, long-endurance, compact, affordable, buoyancy-driven vehicles can be used for monitoring large and mesoscale processes that are currently impossible to do using conventional, propeller-driven AUVs or moorings, and expensive if using research

 Denotes Open Access content.

Corresponding author address: Mario Brito, National Oceanography Centre, University of Southampton, Waterfront Campus, European Way, Southampton, Hampshire SO14 3ZH, United Kingdom.
E-mail: mario.brito@noc.ac.uk



This article is licensed under a [Creative Commons Attribution 4.0 license](https://creativecommons.org/licenses/by/4.0/).

DOI: 10.1175/JTECH-D-13-00138.1

ships and towed vehicles. A number of marine science programs have demonstrated the benefits of underwater gliders. For example, [Perry et al. \(2008\)](#) discussed using an underwater glider to gain deep understanding of blooms located off the coast of Washington State. Perry et al. gathered evidence to conclude that what satellite imagery led scientists to believe was an autumn bloom caused by destratification was instead a vertical redistribution of phytoplankton. Furthermore, the authors concluded that the concentration at the chlorophyll maxima was 3 times that predicted using only satellite imagery. [Todd et al. \(2011\)](#) used underwater gliders to assess the underwater effects of El Niño on the Californian Current System (CCS). The authors concluded that while the CCS was unusually warm and isopycnals unusually deep, there were no anomalous water masses in the region. [Hátún et al. \(2007\)](#) used two underwater gliders to sample eddies in order to understand how these contribute to the rapid restratification of the Labrador Sea interior following wintertime convection.

Developments in communications, intelligent mission planning, and sampling methods devised over the years have brought the AOSN vision closer to reality ([Curtin et al. 2005](#); [Leonard et al. 2010](#); [Alvarez and Mourre 2012a,b](#); [L'Hévéder et al. 2013](#)). Are we now at the stage when users can plan glider missions in full expectation of being able to achieve missions only limited by the sensors and the stored energy? To answer that question requires a study of how well gliders have performed on actual missions. However, over the years little or no study has been made on the reliability of underwater gliders. Results of risk, reliability, and availability analyses conducted for propeller-driven autonomous underwater vehicles ([Brito et al. 2010](#); [Brito and Griffiths 2011](#); [Brito et al. 2012](#); [Podder et al. 2004](#); [Griffiths et al. 2003](#)) cannot be used to infer the reliability of underwater gliders because details of implementation in hardware and software matter.

This paper investigates the reliability of underwater gliders, resulting in a risk profile as a function of mission endurance based on the operational history of 56 underwater gliders during the period January 2008–May 2012. The success of glider missions is dependent upon a number of factors: the inherent reliability of the component parts, the service history of the vehicles, the environment in which they operate, and the practices and procedures of the vehicle operators. We have not attempted to separate all of these factors. The focus here is to create a risk profile based on user experiences using commercially available off-the-shelf (COTS) gliders. Following the creation of the risk profile, we study the effect of some potential covariates in the risk profile. We close with a probabilistic method for quantifying the likelihood of a set of underwater gliders providing

coverage for a predefined observation time. This process allows the user to estimate how many underwater gliders are required in order to meet a given ocean coverage.

2. Underwater gliders

Glider propel themselves by use of a buoyancy engine and thus must follow a sawtooth trajectory through the water. The horizontal speed is typically about 1 km h^{-1} . By traveling slowly and using minimal power, they are able to achieve endurance of several months. The first scientific missions with gliders were undertaken by the teams who developed the vehicles and their collaborators (e.g., [Rudnick et al. 2004](#)). But since about 2005, gliders have become available to the wider scientific community and COTS underwater gliders are being increasingly used by a growing number of institutions. At the time of this study, there were three main COTS underwater gliders: the Slocum ([Webb et al. 2001](#)), the Seaglider 1000 ([Eriksen et al. 2001](#)), and the Spray ([Sherman et al. 2001](#)).¹ These are typically equipped with conductivity, temperature, depth, fluorescence, and optical backscatter sensors, but many other sensors have been used.

3. Ocean coverage estimation

The term ocean coverage is used here to describe the likelihood that a target area will be observed for a required period of time. The ocean coverage is therefore inherently dependent on the glider reliability. Thus, in this section we will first address the problem of estimating underwater glider survival with mission endurance. Two approaches are presented. First, we present a nonparametric method for estimating the probability of survival of an underwater glider. In this paper we consider two different consequences of failure and therefore the term survival has two meanings. We use probability of survival to denote the likelihood of an underwater glider surviving a mission without premature end. The term probability of survival is also used to capture the likelihood of an underwater glider surviving a mission without loss. The Kaplan–Meier survival estimator is used for modeling both scenarios. The probability of survival varies with travel time. A second method, the Cox proportional-hazards model, is then presented to assess the impact of covariates.

a. Survival estimation

Methods for estimating the probability of survival based on historic data can be parametric and

¹ More recent gliders such as the Exocetus coastal glider and the SeaExplorer were not generally available during the study period.

nonparametric (Kalbfleisch and Prentice 1980). Parametric models assume that the probability of failure follows a particular trend, such as linear increasing failure rate or constant rate over time. Nonparametric models make no assumption with regard to the failure distribution. The survival dataset consists of two types of data: failure data and censored data. Failure data consists of the recorded time at which failure took place. In statistical survival modeling, a censored entry is an observation where failure was not observed. In our analysis when a glider is known to have survived a given mission time, this is denoted a right censored data entry. The Kaplan–Meier estimator [Eq. (1)] is a typical method for estimating the probability of survival based on the failure history. It estimates the probability of failure in a given interval, from the fraction of the number of entries that failed at that interval over the number of entries that have not failed. The number of entries that have not failed prior to interval i are denoted by n_i . The number of entries that have failed during interval i are denoted as d_i . The estimator uses the product rule for calculating the probability $\hat{S}(k)$ of a system surviving a sequence of k intervals,

$$\hat{S}(k) = \prod_{i=0}^{i=k} \left(\frac{n_i - d_i}{n_i} \right). \quad (1)$$

The variance for the original Kaplan–Meier estimator is typically computed using the “exponential” Greenwood formula (Prentice and Kalbfleisch 2002), defined as

$$\hat{V}(k) = \frac{1}{[\log \hat{S}(k)]^2} \sum_{n_i < k} \frac{d_i}{n_i(n_i - d_i)}. \quad (2)$$

The variance is a function of the number of entries at any given interval. It is very important to take into account the variance, as this has a direct effect on the confidence limits for the survival estimates as presented in Eqs. (3) and (4):

$$\exp\{-\exp[c_+(k)]\} < \hat{S}(k) < \exp\{-\exp[c_-(k)]\}, \quad (3)$$

where

$$c_{\pm}(k) = \log[-\log \hat{S}(k)] \pm z_{\alpha/2} \sqrt{\hat{V}}, \quad (4)$$

and where $z_{\alpha/2}$ is the upper $\alpha/2$ point of the standard normal distribution; the 5% point is used in this paper, which is 1.96. If Z is a random variable that has a normal distribution, then the upper $\alpha/2$ point of this distribution is that value of $z_{\alpha/2}$, which is such that $P(Z > z_{\alpha/2}) = \alpha/2$. This probability is the area under the standard normal curve to the right of $z_{\alpha/2}$.

b. Proportional hazards analysis

Estimating whether other variables have an influence on risk is almost as important as the estimation of risk

itself. The basic Cox proportional hazards model (Cox 1972) attempts to fit survival data with covariates \mathbf{z} to a hazard function, where \mathbf{z} is a vector of covariates: $\mathbf{z} = (z_1, z_2, \dots, z_n)$. In the literature, covariates are also denoted as explanatory variables; these are variables that may or may not have an influence on the hazard rate. The hazard function is the ratio between the failure density distribution and the cumulative survival function. The Cox proportional hazards model assumes that the baseline hazard $h_0(t)$ is proportional to the explanatory variables by a constant coefficient β . Here β is a vector of proportional coefficients, $\beta = (\beta_1, \beta_2, \dots, \beta_n)$. The formulation for the Cox proportional hazards model is presented in Eq. (5):

$$h(t | z) = h_0(t) e^{\beta_1 z_1 + \beta_2 z_2 + \dots + \beta_n z_n}. \quad (5)$$

The β coefficients are the unknown parameters in the model. These can be estimated using the method of maximum likelihood estimation (MLE). Using this technique, β is obtained by maximizing the partial likelihood given by the product over the r events. Cox (1972) showed that the likelihood function for the proportional for the proportional hazards model [Eq. (5)] is given by

$$L(\beta) = \prod_{j=1}^r \frac{\exp(\beta_j z_j)}{\sum_{l=1}^{n_j} \exp(\beta_l z_l)}. \quad (6)$$

If the β coefficient is greater than 0, an increase in the respective explanatory variable causes an increase in the hazard rate. If the β coefficient is lower than 0, it means that an increase in the respective explanatory variable causes a decrease in the hazard. A proportional coefficient of 0 means that the explanatory variable has no effect on the hazard rate.

The Cox proportional hazards model is very useful for comparing two groups of survival times, corresponding to, for example, two different vehicle makes for the same operational conditions or two different operating conditions for the same vehicle make. This hypothesis test is a procedure that enables us to assess the extent to which an observed set of data is consistent with a null hypothesis, where the null hypothesis represents the view there is no difference between two groups of survival data. In section 6, we use this approach for quantifying the effect of different conditions on the hazard rate.

c. Coverage estimation

Having presented means for quantifying the likelihood of a glider surviving a given time, this section addresses the question of how many gliders are needed in

order to meet a given coverage. We start with the known probability of failure and then use probability rules for deriving the formulation used for calculating the coverage.

Given that we know an instantiation of the probability of failure $p(T)$, at or before time T , which we will denote as p . If we wish to improve the probability of continuous monitoring during the time T , then we could deploy more than one glider. The probability of r failures among a group of N gliders can be computed using the binomial distribution

$$P_N(r, 1) = \binom{N}{r} p^r (1-p)^{N-r}, \quad (7)$$

where $\binom{N}{r}$, also denoted as r choose N , rC_N is $N!/r!(N-r)!$. The probability that at least one glider survives for the time T is given by

$$1 - P_N(N, 1) = 1 - p^N. \quad (8)$$

If the period of observations required exceeds the total endurance, then multiple missions are required for M sequential mission each with N gliders. The probability of at least one glider surviving each mission can be calculated as

$$P_{\text{surv}} = [1 - P_N(N, 1)]^M = (1 - p^N)^M. \quad (9)$$

4. GROOM underwater glider operational history

The task of gathering a broad representative sample of operational histories of glider deployments was undertaken as part of a European Union (EU) Seventh Framework Programme for Research and Technological Development project, Gliders for Research, Ocean Observation and Management (GROOM). The GROOM project (see <http://www.groom-fp7.eu/doku.php>) has 18 European partners, of which 12 operate gliders, working together to “design a new European Research Infrastructure that uses underwater gliders for collecting oceanographic data.” The participants were encouraged to provide operational data representative of a period of 2 years of operation. An online survey prompted the user to enter 1) the organization name; 2) the point of contact; 3) vehicle identifier; 4) start of mission date; 5) vehicle type (Slocum G1 shallow, Slocum G1 Deep, Slocum G2 Shallow, Slocum G2 Deep, Seaglider1000, Spray); 6) mission type (shelf deployment, shelf-edge deployment, deep-ocean deployment); 7) mission length in days; 8) mission maximum depth in meters; 9) did the mission end in failure (yes or no); 10) was the premission test successful? and 11) was the vehicle recovered at the end of the mission?

If the mission had ended in failure, the user was prompted to select from 15 primary causes: 1) collision with vessel; 2) collision with seabed; 3) collision with nets or other obstacle; 4) Iridium communications failure; 5) leak; 6) buoyancy pump failure; 7) power/battery failure; 8) command/control software failure (includes BaseStation); 9) onboard software failure; 10) data-logging failure; 11) navigation sensor failure—GPS; 12) attitude sensor failure (heading, pitch, or roll); 13) sensor failure; 14) altitude control; and 15) other failure (in this case, the user was encourage to write more details). If the mission had ended in failure, the user was prompted to answer what was the status of the altimeter at the time of the fault (bottom within range, bottom outside range).

In practice there may be a causation relation between some of the 15 primary causes. For example, a connector leak may cause a power failure. For this study we asked the users to specify the root cause as the primary cause for the incident. For the particular example, where a leak causes a power failure, which then causes a premature end of the mission, we expect the user to diagnose the leak as the primary cause for the premature mission end.

If the hypothesis is not thoroughly tested and evidence is not considered, then the task of fault diagnosis can be subject to epistemic uncertainty. In this study we assume that this has not been the case and that all primary faults are correctly diagnosed.

Among underwater glider operators the words *abort* and *mission* can have different technical meanings, so we clarify here the meaning as used in this paper. We use the word *mission* to refer to a single glider operation from the time of deployment to the time of recovery. A successful mission is one where operation continued until the planned recovery. During a mission there may be a technical problem, but if these are resolved without having to end the mission prematurely, we class the mission as successful. If, however, the mission is terminated prematurely because of technical issues, then we class this as an aborted mission.

a. Mission statistics

Reports were received on 205 missions carried out by 56 underwater gliders. The number of missions and the number of gliders used varied significantly from one institution to another (Table 1). The statistics for different vehicle makes are presented below (see Table 2).

As noted previously the success of glider missions is dependent not just on the reliability of the particular type of glider used but also on the service history of the vehicle, the environment in which it operates, and the practices and procedures of the operators. Furthermore, whether failure leads to loss of the vehicle is very much

TABLE 1. Summary of the data provided by the participating oceanographic laboratories.

Centre name	No. of missions	Start date–end date	No. of vehicles	Median endurance (days)	Operation environment		
					Shelf	Shelf edge (%)	Deep ocean (%)
Centre National de la Recherche Scientifique (CNRS)	56	Apr 2010–Mar 2012	14	32	12	84	4
Consejo Superior de Investigaciones Científicas (CSIC)	14	Apr 2011–May 2011	5	19	—	—	100
Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (OGS)	4	Jun 2008–Jul 2010	2	13	75	—	25
Alfred-Wegener-Institut für Polar- und Meeresforschung (AWI)	8	Jul 2008–Jul 2012	4	69	—	—	100
Consortio para el Diseño, Construcción, Equipamiento y Explotación de la Plataforma Oceánica de Canarias (PLOCAN)	14	Dec 2010–Oct 2012	4	13	7	22	71
University of East Anglia (UEA)	12	Mar 2010–Jan 2012	8	49	42	50	8
Oceanography Centre, University of Cyprus (OC-UCY)	7	Mar 2009–Dec 2011	2	80	71	29	—
Institut für Meereswissenschaften–Research Center for Marine Geosciences (IFM-GEOMAR)	3	Jun 2011–Nov 2012	1	14	—	—	100
Helmholtz-Zentrum Geesthacht–Zentrum für Material- und Küstenforschung GMBH (HZG)	12	Dec 2010–Jun 2012	2	23	100	—	—
North Atlantic Treaty Organization (NATO) Undersea Research Centre (NURC)	63	Jun 2010–Sep 2012	8	4	33	30	37
Scottish Association for Marine Science (SAMS)	3	Oct 2009–Sep 2012	2	123	—	33	67
Natural Environment Research Council (NERC)	9	Jul 2010–Aug 2012	4	39	—	33	67

dependent upon the available options for recovery. In our survey, five of the vehicles that were lost were Seagliders deployed in Arctic or Antarctic waters, where there are very limited opportunities for emergency recovery. Thus, although more Seagliders were lost, it is not possible to conclude that they are inherently more likely to be lost than Slocums.

b. Failure modes

A total of 63 mission aborts were recorded during the 205 missions. Seventeen specific failure modes have

been identified (Fig. 1). For four failures the root cause remains unknown. In general, there are a small number of observations for each failure mode; therefore, it is not possible to infer if a particular vehicle make is more prone to a particular failure mode than another make. However, for the three most common failure modes, we have compared the failure rate for Slocum gliders (deep, shallow, G1, and G2) with that for Seagliders. The results are shown in Table 3.

The hypotheses test is a procedure that enables us to assess the extent to which an observed set of data is

TABLE 2. Glider operation statistics.

	Seaglider 1000 m	Slocum G1 shallow	Slocum G1 deep	Slocum G2 shallow	Slocum G2 deep
No. of missions	42	68	72	9	14
Total endurance (days)	2514.5	772.05	1728	188	550.1
Median endurance (days)	64	7.65	19.5	18	12
Upper quartile (days)	80	15	37	25	25.8
Max endurance (days)	169	56	105	48	184
No. of aborts due to failures	19	13	23	3	5
Abort rate (per day)	0.00756	0.0168	0.0133	0.0159	0.00909
No. of losses	7	2	1	0	0

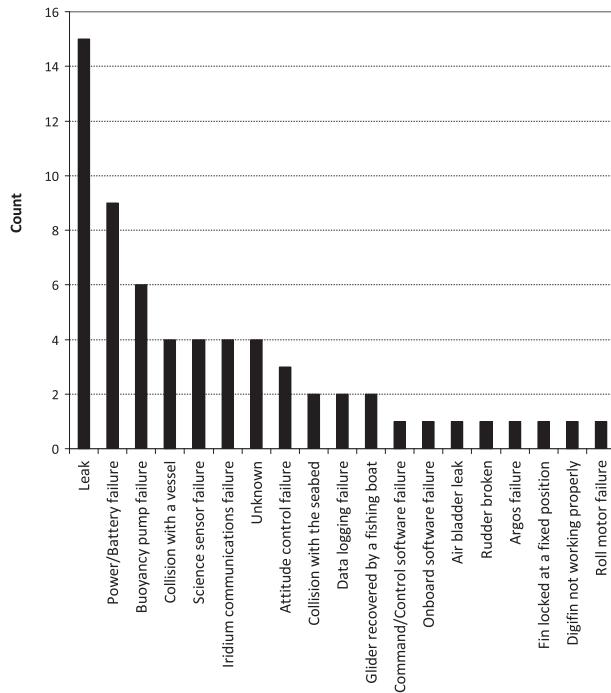


FIG. 1. Failure modes for all underwater gliders, shallow and deep.

consistent with a particular hypothesis, known as the null hypothesis. A null hypothesis represents a simplified view that specifies that there is no difference between the two groups of survival data. For each failure mode, we tested the null hypothesis that the failure rate for Slocums and Seaglidgers was indistinguishable. By comparing the actual difference in failure rates with the standard error of the difference, the probability of true failure rates being different can be calculated using the two proportion z test. The z test is a common statistical significance test that can be used for testing the hypothesis that differences in proportion of two sets of data are not statistical significant (O'Connor 2002).

Failure due to a leak was the most observed failure mode. Out of the 15 failures, 14 occurred on Slocum vehicles; so, the rate of occurrence was about 3 times greater for Slocum gliders. One possible explanation is that Slocum vehicles are opened by users more often than Seaglider1000s, as Seaglidgers have generally been

serviced by the makers. Therefore, the O-rings of the Slocum vehicles tend to be more disturbed than the O-rings on the Seaglider1000.

The second most common failure mode was power/battery issues; these occurred 7 times more frequently for Seaglidgers than for Slocums. Seaglidgers employ lithium batteries, while Slocums can be set to employ either lithium or alkaline batteries. This may be the root of the observed differences in failure rate. However, we did not collect data concerning the type of battery employed on all Slocums in the dataset. Therefore, without further information we cannot test whether battery chemistry was a contributory factor in the different failure rates.

The third most common failure mode was buoyancy pump failure. The failure rates for the two different underwater glider makes are indistinguishable for this failure mode.

c. Failure analysis

Underlying the procedure used in this section is the assumption that there are no significant differences between survival times of each group; that is, the difference that has been observed is due to chance variation. Two well-known tests used for comparing survival distributions are the *log-rank* and the *Wilcoxon* tests (Collett 2003).

When we applied these tests to compare the survival distributions of the Seaglider 1000 and Slocum G1 deep, we concluded that there is no evidence against the null hypothesis. As presented in column 3 of Table 4, the values of P from both tests suggest that differences between the survival of the Seaglider 1000 and that of a Slocum G1 deep are not statistically significant. When we compared the Slocum G2 deep survival distribution with the distribution of the aggregated dataset of Slocum G1 deep and Seaglider 1000, we concluded that there is no evidence against the null hypothesis. Since the differences between the failure distributions of Seaglider1000, Slocum G1 deep, and Slocum G2 deep are not statistically significant, we can aggregate the mission data of these three vehicles to make a unique dataset that represents the operational history of deep underwater gliders. If we consider the shallow gliders—Slocum G1 shallow and Slocum G2 shallow—the

TABLE 3. Breakdown of the top three failure modes and the confidence that the failure mode is vehicle dependent.

Failure mode	Slocums		Seaglider1000s		Confidence (%)
	No. of failures	Failure rate	No. of failures	Failure rate	
Leaks	14	0.0859	1	0.0238	91
Power/battery failure	3	0.0184	6	0.143	97
Buoyancy pump failure	5	0.0307	1	0.0238	57

TABLE 4. Comparison of the survival of different underwater glider makes. Deep gliders include Seaglider1000, Slocum G1 deep, and Slocum G2 deep. Shallow gliders include Slocum G1 shallow and Slocum G2 shallow.

Statistical test	Deep gliders				Shallow gliders	
	Seaglider1000 compared to Slocum G1		Seaglider1000 and Slocum deep G1 compared to Slocum deep G2		Slocum G1 shallow compared to Slocum G2 shallow	
	χ^2	P	χ^2	P	χ^2	P
Wilcoxon test	1.4966	0.2212	2.2008	0.3327	0.4679	0.4938
Log-rank test	1.3823	0.2697	1.3787	0.5019	0.0003	0.9872

large values of P indicate that the difference in the failure distribution for these two types of vehicles is not statistically significant. Therefore, the operational history for both vehicles can be aggregated to form a unique dataset corresponding to shallow underwater gliders.

5. Underwater gliders survival

Here we present the results of survival analysis carried out on the glider mission data collected from the GROOM project participants. A mission risk profile is created for each type of underwater glider and the Cox proportional hazards method, presented in section 3b, is used for estimating the effect of different covariates. The survival analyses are carried for two scenarios: abort and loss.

a. Likelihood of underwater glider abort

The analyses were carried out for the deep underwater gliders and for the shallow underwater gliders separately. The deep underwater gliders' dataset consisted of 128 missions; the shallow of 77 missions. The probability of a vehicle completing its planned mission without aborting is presented in Figs. 2a and 2b. The 95% confidence limits are presented in gray; they increase with endurance, as there are fewer missions of longer endurance.

The probability of the mission not ending in abort decreases relatively rapidly during the first 15 days for deep gliders and the first 5 days for shallow gliders, but after this time the probability appears to decrease at a constant rate. This suggests that the failure rate is almost constant with time beyond the first few days; that is, for example, failures are just as likely to emerge in the sixth week of deployment as during the fourth week. This contrasts with the profiles for the Autosub and International Submarine Engineering Limited (ISE) Explorer propeller-driven AUVs, where the risk profile reduces more significantly in the first tens of kilometers, allowing risk mitigation by monitoring the vehicle for this distance before committing to the mission (Brito

et al. 2010, 2012). Because of the almost constant failure rate, a monitoring distance would not be as effective as a risk mitigation strategy for gliders.

b. Likelihood of underwater glider loss

In the 205 missions considered in this study, 10 of them resulted in vehicle loss, 8 of the losses were for deep underwater gliders, and 2 of the losses were for shallow underwater gliders. The probability of a shallow underwater glider surviving a deployment is presented in Fig. 2c, and the probability of survival of a deep underwater glider is presented in Fig. 2d. For shallow underwater gliders, the survival distribution shows that the probability of a glider surviving a 30-day mission is 0.9. The probability of a deep underwater glider vehicle surviving a 30-day mission is 0.97. For both distributions the 95% confidence interval is quite large.

6. Proportional hazards

The proportional hazards method (section 4) was used to examine evidence on whether operational factors influenced the abort and loss outcomes. The analyses were carried out using the JMP 7 statistical analysis tool from the Statistical Analysis System Institute, Inc. (SAS). The estimates for the hazard analysis are presented in Table 5.

a. Effect of operational depth

For the abort scenario, results show that there is a high confidence in the proportional hazards estimates for both shallow and deep gliders. Negative values for β_1 indicate that the probability of an abort reduces with increasing operational depth; see Eq. (5). For the loss scenario, results show that there is no dependency between the probability of loss and the glider operational depth. The P values for both shallow and deep gliders are large.

b. Effect of altimeter status

One of the "autonomous" behaviors of the glider is in its ability to detect and react to the presence of the

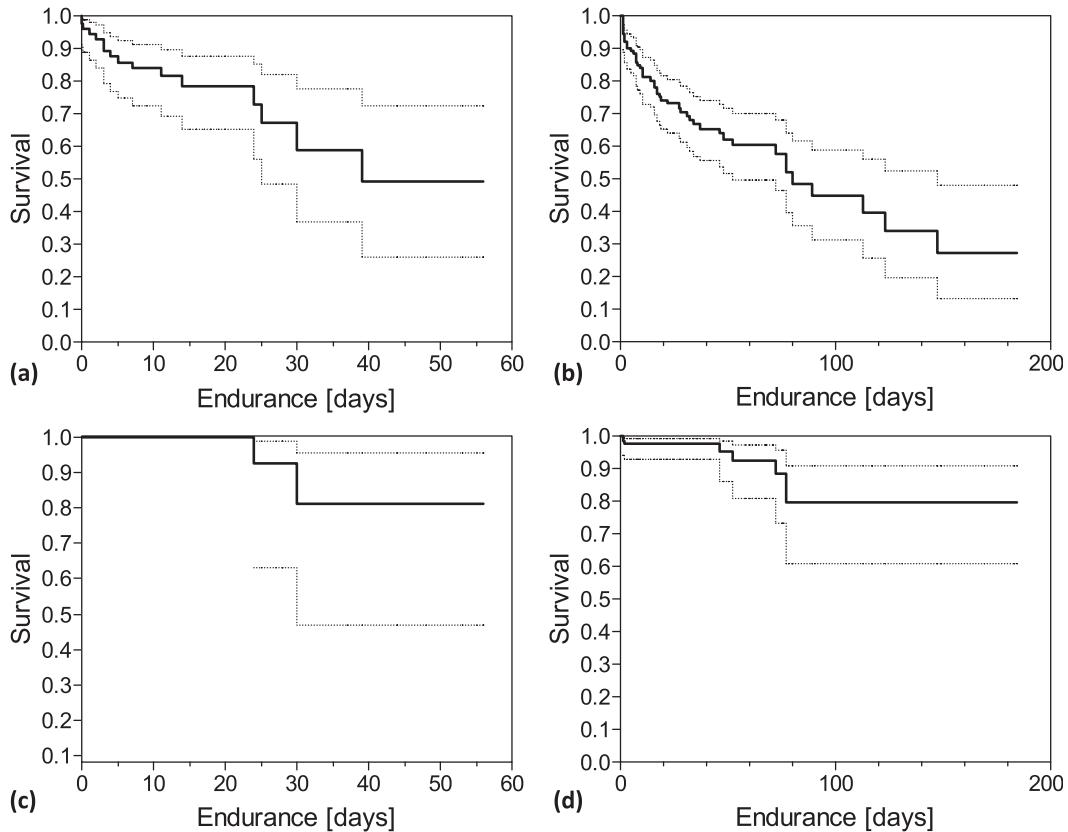


FIG. 2. Probability of survival distributions: (a) probability of surviving without aborting for shallow gliders; (b) probability of surviving without aborting for deep gliders; (c) probability of surviving, without loss, for shallow gliders; and (d) probability of surviving, without loss, for a deep glider.

seafloor; that is, it can determine when to inflect if the bottom depth is less than the commanded inflexion. Getting this wrong could lead to collision with the seafloor and possible consequential damage, for example, to the hull (leaks), the sensors, possibly the communications antenna, and possibly the external bladder, giving rise to a buoyancy engine problem. In this section we attempt to establish whether there is a correlation between the vehicle loss and the status of the altimeter.

For 7 of the 16 aborts that occurred on shallow vehicles, the bottom was outside the range of the altimeter. Of the 47 failures that occurred on deep underwater gliders, for 16 of them the vehicle was within altimeter range of the bottom. For both shallow and deep gliders, the proportional hazards analyses confirm that there is no correlation between the status of the altimeter and the probability of the glider being lost. A possible explanation for this is that gliders move very slowly. Thus, unless the environment is energetic (e.g., fast near-bottom currents) and/or strewn with obstacles, such as fishing gear or large rocks, collision with the seafloor is not typically traumatic. This perhaps explains the lack of relationship between altimeter status and vehicle loss.

7. Example of coverage estimation

The previous two sections have used reliability modeling methods for quantifying the reliability of underwater gliders as a function of mission endurance and the effect of explanatory variables on the risk profile. The explanatory variables considered were the maximum operating depth of the altimeter status. In this section we move from a problem of a single-vehicle deployment to a problem of multivehicle deployment. In this section we assess the impact of underwater glider reliability on

TABLE 5. Statistics on the proportional hazards analysis: β_1 is the proportional constant, L is the lower confidence limit, and U is the upper confidence limit for the proportional constant.

Covariate	Scenario	Statistics	Shallow gliders	Deep gliders
Operational depth	Abort	β_1	-0.009 679	-0.001 615
		P	0.0042	0.003
	Loss	β_1	0.019 691 64	-0.000 882 7
		P	0.2526	0.5687
Altimeter status	Loss	β_1	9.5424	0.1235
		P	0.0959	0.7379

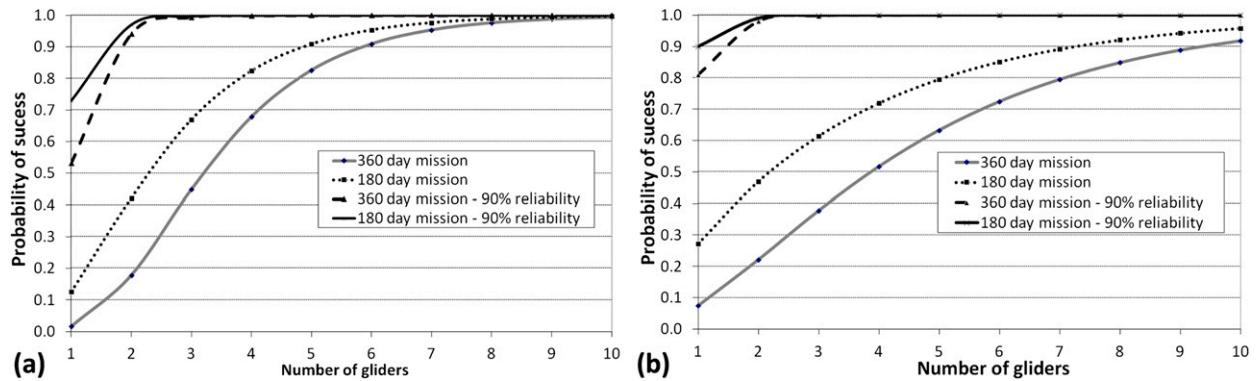


FIG. 3. Probability of successfully completing the two observation targets for a fleet of 1–10 gliders: (a) shallow gliders and (b) deep gliders.

mission planning. In subsection 7a, we consider the situation where a single glider is required to survey an area for a long period of time. Two case studies are considered, a 180-day mission and a 360-day mission. We estimate the likelihood of this survey being successful. Then we consider the impact of adding redundancy, that is, using multiple gliders to improve reliability.

A small number of glider fleet configurations have been tested in recent years, with different degrees of success (Rudnick et al. 2004; Hodges and Fratantoni 2009; Testor et al. 2007). In this section we consider the impact of underwater glider reliability on the risk of a glider network design and operation. In subsection 7b we conduct reliability analysis of the “virtual” mooring array proposed by Hodges and Fratantoni (2009). In subsection 7c we considered the network design proposed by L’Hévéder et al. (2013).

a. Single measurement location

The huge benefit of underwater gliders is the ability for long endurance missions. However, from the dataset available to us, few missions make use of the full endurance. In this case study we consider that the aim is to have at least one glider in continuous operation for a given period of time, in a situation where replacement gliders cannot be deployed to cover failures. The probability of achieving this, as a function of the number of gliders, can be calculated using Eq. (9). Taking the deep glider example, based on Fig. 2 we assume the practical upper limit of endurance is 180 days. For the 180-day coverage, the minimum number of missions (M) equals 1, while for the 360-day coverage $M = 2$. Figure 3 presents the probability of providing continuous coverage with one or more deep gliders, for the two periods of interest: 180 days and 360 days.

Figure 3 shows that for deep underwater gliders we would need to deploy 10 gliders in order to achieve 0.95

probability of successfully providing continuous coverage for 180 days without replacement. A fleet of 20 gliders would be required to have a probability of 0.92 for continuous coverage over 360 days.

However, if we were to consider a shorter mission length, where possible, it would yield a different requirement in terms of the number of gliders. For example, for deep underwater gliders, a mission length of 25 days would have a probability of premature end of 0.27. If we were to run four gliders per a 25-day mission over a period of 360 days, the probability of providing coverage is 0.93. At least eight gliders would be needed, and this would imply that we would have to rotate the gliders 14 times during the year. The fleet size and mission length combinations can be selected to meet a desirable, or at least acceptable, coverage target.

The above-mentioned results indicate that shorter deployments will achieve the same level of confidence while having fewer gliders deployed at any one time. However, in practice the choice of strategy would also have to take into account the cost of different scenarios and other factors.

The calculations given above are very conservative and are intended to give an indication of the number of vehicles required. In practice it will not always be necessary to recover all gliders at the same time. A more efficient strategy would be to recover each glider only when necessary; however, it is important to take into consideration the time required to organize a new deployment and for a glider to navigate to the operational area.

b. Virtual mooring array case study

If currents are not stronger than the glider’s speed, then a glider can be programmed to perform repeated profiles while holding a horizontal position nearly constant. This mode of sampling is known as a *virtual mooring*. An example of this is the virtual mooring array

deployed for 10 days in the Philippine Sea, east of Luzon Strait (Hodges and Fratantoni 2009). During this experiment five Slocum shallow gliders were deployed in five different positions.

Using the analysis in the previous section, we can consider the likely success of this array if it were continued for a period of 6 months. The probability of at least one virtual mooring failure is calculated as follows: $p_{fm} = 1 - (1 - pf_1)(1 - pf_2)(1 - pf_3)(1 - pf_4)(1 - pf_5)$, where pf_i is the probability of the glider holding station i failing to maintain the station for 6 months. We will assume that shallow gliders are used. Given that the probability of failure for a single 60-day mission is 0.49, and that three sequential 60-day missions are required, pf_i is easily calculated as 0.867. Therefore, the probability of at least one failure over three deployments at each of the five sites of the proposed network is 0.99996. This puts a requirement for adding underwater glider redundancy at each station. If each station comprises four underwater gliders, then the probability that at least one of the four gliders at each site will complete a single 60-day mission (p_{fm1}) is $1 - 0.49^4 = 0.9424$. Thus, the probability of continuous monitoring of three sequential deployments of four gliders at one site is 0.837, and the probability of all five sites having continuous records is 0.411. Using the binomial distribution, we evaluate the probability of at least four, three, and two sites being successful is 0.811, 0.967, and 0.997.

c. Glider network for a synoptic view of the oceanic mesoscale variability case study

L'Hévéder et al. (2013) considered the minimum number of gliders needed to sample mesoscale variability. The authors propose to deploy an array of gliders in a comb structure. The optimal number of gliders was selected to maximize the analysis skill evaluation objective and to minimize the objective analysis error. The analysis skill evaluation objective was quantified by a combination of the root-mean-square error and the spatial pattern correlation between the glider network's simulated data and the controlled field data. The objective analysis error objective was calculated based on the minimum error variance. The authors demonstrated that the optimum number of gliders necessary to sample mesoscale variability was 10. We will assume that all the gliders are deep gliders. In the previous section, we considered the number of gliders that needed to be deployed, so that there would be a high probability of one or more of them completing the mission. An alternative strategy would be, if it is possible, to replace each glider that fails during the mission. When large numbers of gliders are needed, such as in this example, we expect this strategy to require fewer resources. However,

coverage will not be as complete when we take into account the time required to replace a vehicle. Here we consider the number of gliders required for this strategy. To do so we make the assumption that the failure rate is constant in time. If the probability of a glider failing in a given time interval ΔT is $p(\Delta T)$, given a batch of N gliders, the probability of x of them failing is

$$q(x, \Delta T) = B(x, N, p), \quad (10)$$

where B is the binomial distribution. In the limit of ΔT being very small, we can ignore the possibility of multiple failures in any one time interval and so the probability of there being L failures during a time $T = M\Delta T$ is approximately

$$Q(L, T, N) = B(L, M, q), \quad (11)$$

where

$$q = q(1, \Delta T) = Np(1 - p)^{N-1}. \quad (12)$$

The probability distribution of the number of replacements for a 10-glider deployment is shown in Fig. 4. For a typical survivability of 0.5 for a 90-day deep glider deployment, the expected number of replacements is 7 and there is a 16% chance that 10 or more gliders will be needed.

d. Improving reliability

Reliability improvement of underwater gliders will be possible if communication between users and manufacturers is proactive, enabling the discussion of failure modes and potential mitigation activities. Such reliability improvement has occurred on profiling floats. Profiling floats had a target life expectancy of 4 years, performing 150 cycles during this period. However, in 2001 only 20% of the Autonomous Profiling Explorer (APEX) floats could meet this requirement (Kobayashi et al. 2009). The fact that faulty floats could not be recovered made it difficult to identify the root causes for failures. Nevertheless, research institutes and the manufacturer engaged in fault investigations and a number of improvements were made as a result. For example, the batteries of the early floats had a design vulnerability that meant that every time a battery cell was damaged, it caused a chain reaction, in which other battery cells in the same pack would also fail. The battery circuit design was changed; a diode was introduced between cells, so that if one cell is damaged, it will not damage the cell next to it. Another improvement was made with regard to the piston used to control the buoyancy. The pump used in the early APEX float allowed small sediments to mix with the oil. This would eventually cause the piston

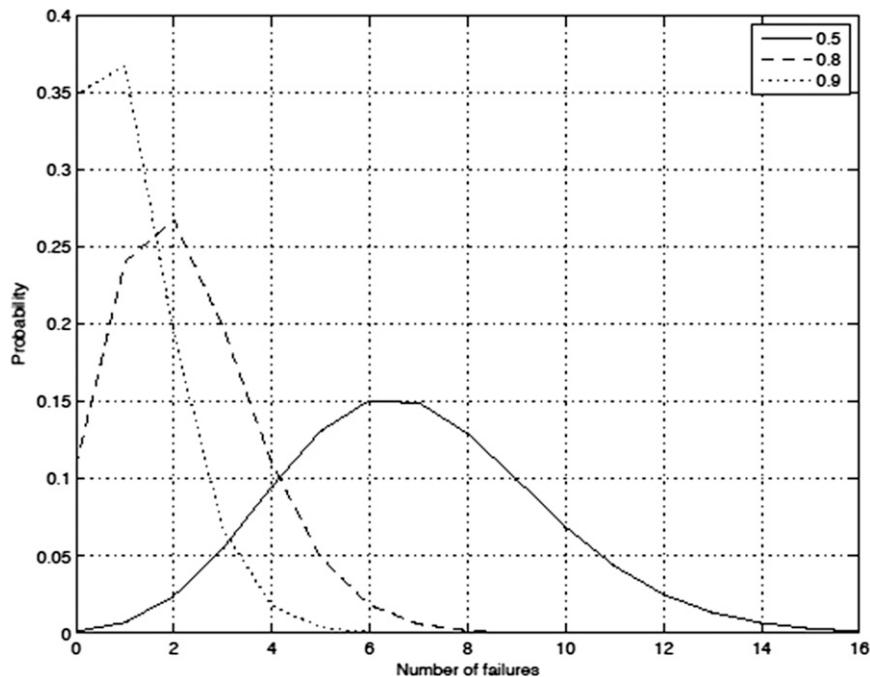


FIG. 4. Probability density function for the number of failures in a 90-day deployment of 10 gliders. The probability of a single glider surviving a 90-day mission is assumed to be 0.5 (continuous line), 0.8 (dashed line), and 0.9 (dotted line).

to get stuck in a fixed position. A new pump was designed by the manufacturer that did not have this failure mode.

In this section we study the impact of reliability improvement in the confidence that the glider network design will meet the observation target.

First, we consider the case of a single measurement location as in section 7a. Figure 3 shows that if the reliability of an underwater glider is increased to 90%, this results in a high confidence that the coverage target will be met with fewer gliders. For deep gliders, a deployment of two gliders would give a confidence of 98% that the target measurements would be obtained with one or more gliders for a period of 360 days.

For the multiple glider deployment considered in section 7c, Fig. 4 shows that if the success rate for 90-day missions can be increased from 0.5 to 0.9, then the number of gliders required to ensure a 10-glider deployment is likely to be greatly reduced, with the expected number of replacements being one and the chance of needing three or more being only 9%.

If the top three failure modes identified in this study—leaks, battery failure, and buoyancy pump failure—are mitigated, this would lead to a change in the risk profile. Figure 5 presents the survival distribution for considering that the top three failure modes were completely mitigated. Each failure was replaced with a censored

entry, since each failure resulted in an early mission termination. Thus, replacing the failure flag with a censored flag would not result in a risk improvement. Therefore, we assume that the endurance for each one of these failure entries equals the average of the endurance of all missions that were successful. For shallow underwater gliders, the average endurance of all successful missions was 13 days, while for deep underwater gliders this was 43 days.

The probability of shallow glider surviving, with no premature mission end, a 60-day mission is 0.58, a 0.09 increase from the unmitigated case. The probability of a deep glider surviving a 180-day mission is 0.48, a 0.2 increase from the unmitigated case.

8. Discussion and conclusions

The paper presents a probabilistic framework for calculating the coverage that can be achieved by a fleet of underwater gliders. We showed how the probability of successfully meeting a required coverage can be calculated using the survival estimate for a given mission. We use this framework for estimating the coverage for two proposed ocean sampling networks.

In the examples provided, we considered that all missions were of equal endurance. In practical terms missions can be of different endurance. In addition, in

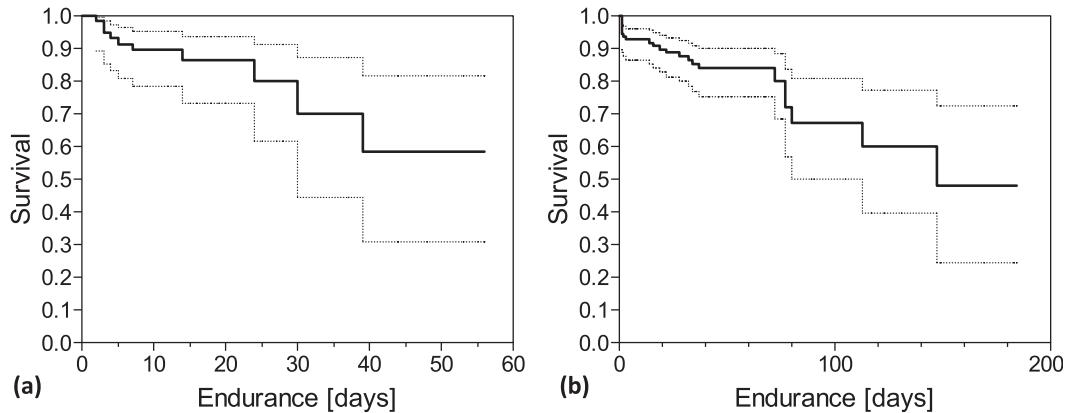


FIG. 5. Underwater glider survival taking into account mitigation of the top three failure modes: leaks, battery failures, and buoyancy pump failure: (left) shallow gliders and (right) deep gliders.

the examples, we considered that the vehicles were deployed concurrently. Again, this may not be the case in practice. However, despite these assumptions, the probabilistic formalism presented in this paper still applies to the different scenarios mentioned.

A crucial aspect of this analysis is that the survival profiles for each vehicle type were created based on unbiased data. The users consisted of people that were not involved in the development life cycle of underwater gliders. For these users the underwater gliders come as a COTS product. The framework proposed here for estimating the coverage obtained by a fleet of underwater gliders was derived from first principles of probability theory. This framework can be used to support the survey design of any organization, including organizations that have been involved in the development life cycle of underwater gliders, which comprise more experienced users. This is possible provided that the organization uses their own operation data to generate the risk profiles.

We used fault history data from 205 glider missions, provided by nondevelopers, to build risk profiles with endurance. We concluded that the risk profiles of different underwater glider makes are not statistically significant. Therefore, our analysis focused on two classes of gliders: deep and shallow. For shallow gliders we concluded that the probability of not aborting a 30-day mission is approximately 0.5. For deep underwater gliders, the probability of not aborting a 90-day mission is approximately 0.59. A key observation is that successful glider deployments with vehicles available today imply conducting missions that are well below the maximum endurance of the vehicle.

These glider failure profiles have a similarity in form to those of relatively early APEX floats (2000–01 and 2003) in the analysis of Kobayashi et al. (2009). APEX

floats deployed in subsequent years generally showed a growth in reliability, such that by 2006 the probability of completing 100 cycles was over 90%, compared with ~20% in 2000–01. The challenge for manufacturers is to achieve the same reliability growth for gliders.

In targeting this reliability issue, user feedback to manufacturers to inform ongoing developments by manufacturers is important. By doing so the glider reliability can increase, as has shown to be possible from the APEX float experience reviewed in this paper.

If, for example, the probability of failure for a 180-day deep glider mission could be reduced from 0.73 to 0.25, then the number of gliders needed to be deployed simultaneously to achieve 95% coverage would be reduced from 10 to 3.

Underwater gliders are arguably one of the most significant technology developments in autonomous underwater sampling. They provide an effective way to conduct marine science surveys and, for some missions, they completely eliminate the costs associated with surface vessels. The perception is that underwater gliders are a relatively cheap alternative to measurements made from ships and moorings. However, in this paper we have shown that in order to achieve a high level of confidence in obtaining data, multiple underwater gliders are required. Therefore, when evaluating the cost of underwater glider observations, the number of vehicles required to meet the necessary level of confidence needs to be considered. Using current technology and practices, a high level of confidence may require a costly operation. However, if glider manufacturers and operators can achieve a similar improvement in reliability as was made for Argo floats, then the costs will fall significantly.

Validation of risk estimation, in general, is a difficult task. We believe that for underwater gliders this is

a particularly difficult task because these platforms are constantly evolving. The faults presented in this study generated lessons learned that were disseminated by all partners. It is possible to validate the estimated risk profiles for the underwater gliders by comparing the risk profiles presented in this paper with the risk profiles generated using data collected since this study. In making this comparison we must take into account the impact of fault mitigation. This can be achieved using the approach presented in Brito et al. (2012). The statistical tests discussed in this paper—the Wilcoxon and the log-rank tests—can be applied to compare the two risk profiles, prior to 2012—presented in this paper—and after 2012, the results of a future study.

Acknowledgments. The authors acknowledge the support of the European Seventh Framework Programme Grant 284321. The authors also thank the researchers who took their time to complete the survey of the glider mission history. The authors are extremely grateful to the three anonymous reviewers of this article for taking their time to review this article; in our view their comments and challenges helped improve significantly the quality of this paper. The authors are also extremely grateful to Clayton Jones and Ben Allsup from Teledyne Research Webb for reviewing the early versions of this manuscript.

REFERENCES

- Alvarez, A., and B. Mourre, 2012a: Oceanographic field estimates from remote sensing and glider fleets. *J. Atmos. Oceanic Technol.*, **29**, 1657–1662, doi:10.1175/JTECH-D-12-00015.1.
- , and —, 2012b: Optimum sampling designs for a glider-mooring observing network. *J. Atmos. Oceanic Technol.*, **29**, 601–612, doi:10.1175/JTECH-D-11-00105.1.
- Brito, M. P., and G. Griffiths, 2011: A Markov chain state transition approach to establishing critical phases for AUV reliability. *IEEE J. Oceanic Eng.*, **36**, 139–149, doi:10.1109/JOE.2010.2083070.
- , —, and P. Challenor, 2010: Risk analysis for autonomous underwater vehicle operations in extreme environments. *Risk Anal.*, **30**, 1771–1788, doi:10.1111/j.1539-6924.2010.01476.x.
- , —, J. Ferguson, D. Hopkin, R. Mills, R. Pederson, and E. MacNeil, 2012: A behavioral probabilistic risk assessment framework for managing autonomous underwater vehicle deployments. *J. Atmos. Oceanic Technol.*, **29**, 1689–1703, doi:10.1175/JTECH-D-12-00005.1.
- Collett, D., 2003: *Modelling Survival Data in Medical Data*. Chapman & Hall/CRC, 391 pp.
- Cox, D. R., 1972: Regression models and life-tables. *J. Roy. Stat. Soc.*, **34B**, 187–220.
- Curtin, T. B., J. G. Bellingham, J. Catipovic, and D. Webb, 1993: Autonomous oceanographic sampling networks. *Oceanography*, **6**, 86–94.
- , D. M. Crimmins, J. Curcio, M. Benjamin, and C. Roper, 2005: Autonomous underwater vehicles: Trends and transformations. *Mar. Technol. Soc. J.*, **39**, 65–75, doi:10.4031/002533205787442521.
- Eriksen, C. C., T. J. Osse, R. D. Light, T. Wen, T. W. Lehman, P. L. Sabin, J. W. Ballard, and A. M. Chiodi, 2001: Seaglider: A long-range autonomous underwater vehicle for oceanographic research. *IEEE J. Oceanic Eng.*, **26**, 424–436, doi:10.1109/48.972073.
- Griffiths, G., N. W. Millard, S. D. Mcphail, S. D. P. Stevenson, and P. G. Challenor, 2003: On the reliability of the Autosub autonomous underwater vehicle. *Underwater Technol.*, **25**, 175–184, doi:10.3723/175605403783101612.
- Hátún, H. C., C. C. Eriksen, and P. B. Rhines, 2007: Buoyant eddies entering the Labrador Sea observed with gliders and altimetry. *J. Phys. Oceanogr.*, **37**, 2838–2854, doi:10.1175/2007JPO3567.1.
- Hodges, B. A., and D. M. Fratantoni, 2009: A thin layer of phytoplankton observed in the Philippine Sea with a synthetic moored array of autonomous gliders. *J. Geophys. Res.*, **114**, C10020, doi:10.1029/2009JC005317.
- Kalbfleisch, J. D., and R. L. Prentice, 1980: *The Statistical Analysis of Failure Time Data*. Wiley, 321 pp.
- Kobayashi, T., B. King, and N. Shikama, 2009: An estimation of the average lifetime of the latest model of APEX floats. *J. Oceanogr.*, **65**, 81–89, doi:10.1007/s10872-009-0008-x.
- Leonard, N. E., D. A. Paley, R. E. Davis, D. M. Fratantoni, F. Lekien, and F. Zhang, 2010: Coordinated control of an underwater glider fleet in an adaptive ocean sampling field experiment in Monterey Bay. *J. Field Rob.*, **27**, 718–740, doi:10.1002/rob.20366.
- L'Hévéder, B., L. Mortier, and P. Testor, 2013: A glider network design study for a synoptic view of the oceanic mesoscale variability. *J. Atmos. Oceanic Technol.*, **30**, 1472–1493, doi:10.1175/JTECH-D-12-00053.1.
- O'Connor, P. D. T., 2002: *Practical Reliability Engineering*. 4th ed. John Wiley & Sons Ltd., 513 pp.
- Perry, M. J., B. S. Sackmann, C. C. Eriksen, and C. M. Lee, 2008: Seaglider observations of blooms and subsurface chlorophyll maxima off the Washington coast. *Limnol. Oceanogr.*, **53**, 2169–2179, doi:10.4319/lo.2008.53.5_part_2.2169.
- Podder, T. K., M. Sibenac, H. Thomas, W. Kirkwood, and J. G. Bellingham, 2004: Reliability growth of autonomous underwater vehicle-Dorado. *Oceans '04 MTS/IEEE: Techno-Ocean '04; Bridges across the Oceans*, Vol. 2, IEEE, 856–862.
- Prentice, R. L., and J. D. Kalbfleisch, 2002: *The Statistical Analysis of Failure Time Data*. Wiley, 432 pp.
- Rudnick, D. L., R. E. Davis, C. C. Eriksen, D. M. Fratantoni, and M. J. Perry, 2004: Underwater gliders for ocean research. *Mar. Technol. Soc. J.*, **38**, 73–84, doi:10.4031/002533204787522703.
- Sherman, J., R. E. Davis, W. B. Owens, and J. Valdes, 2001: The autonomous underwater glider “Spray.” *IEEE J. Oceanic Eng.*, **26**, 437–446, doi:10.1109/48.972076.
- Stommel, H., 1989: The Slocum mission. *Oceanography*, **2**, 22–25, doi:10.5670/oceanog.1989.26.
- Testor, P., and Coauthors, 2007: European gliding observatories (EGO). *The Coriolis Newsletter*, No. 4, Coriolis/Mercator, 11–12. [Available online at www.coriolis.eu.org/News-Events/Newsletters/Coriolis-4/.]
- Todd, R. E., D. L. Rudnick, R. E. Davis, and M. D. Ohman, 2011: Underwater gliders reveal rapid arrival of El Niño effects off California's coast. *Geophys. Res. Lett.*, **38**, L03609, doi:10.1029/2010GL046376.
- Webb, D. C., P. J. Simonetti, and P. J. C. P. Jones, 2001: SLOCUM: An underwater glider propelled by environmental energy. *IEEE J. Oceanic Eng.*, **26**, 447–452, doi:10.1109/48.972077.