AGREEMENT ON THE INTERNATIONAL DOLPHIN CONSERVATION PROGRAM

37TH MEETING OF THE PARTIES

San Diego, California (USA)
17 August 2018

DOCUMENT MOP-37-02

DESIGN OF A SURVEY FOR EASTERN TROPICAL PACIFIC DOLPHIN STOCKS

Cornelia S. Oedekoven[1], Stephen T. Buckland[1], Laura Marshall[1] and Cleridy E. Lennert-Cody[2]

CONTENTS

1. EXECUTIVE SUMMARY

From the early 1980's, the management of eastern tropical Pacific Ocean (ETP) dolphin species and stocks has relied heavily on estimates of abundance from research vessel surveys. The current 12-year hiatus in these surveys, and the problematic nature of monitoring stock status from fishery-dependent data, means that there are presently no reliable indicators with which to monitor the status of ETP dolphin populations. A new ETP survey would be a step towards mitigating this situation.

The three primary considerations with respect to planning a new survey are: the objective(s) of the survey, which affect(s) the survey methodology; the species/stocks of interest, which may affect the definition of the survey area; and the available budget, which may affect both the survey methodology and the survey area. For the next ETP survey, two issues require additional considerations: recent evidence that trackline probability may be less than one for ETP dolphins; and the proposed use of tuna vessels. Hence, we considered two distinct objectives for the next survey: design it to give 1) estimates of relative abundance that are comparable with previous surveys; and 2) estimates of absolute abundance. We outline options

[1] Centre for Research into Ecological and Environmental Modelling, University of St Andrews, The Observatory, Buchanan Gardens, KY16 9LZ, UK
[2] Inter-American Tropical Tuna Commission, 8901 La Jolla Shores Drive, La Jolla CA 92037-1509, USA

for accomplishing either objective 1 or both.

For objective 1 the implications are that we need a set-up as close to that of earlier surveys as possible, which entails using the same or equivalent survey vessels as during previous ETP surveys as well as the same survey protocol of closing mode effort for the primary observer team on the flying bridge. If at least one of former ETP observers is included on each observer team, this will also improve comparability. To meet Objective 2, methods will be needed to estimate the probability of detection of a school that is on the trackline, and how it varies by sea state. Given that some schools that were initially close to the line may be evading detection, these methods should accommodate responsive movement and behavioral responses.

If tuna vessels were to be used in the new survey, there would need to be a calibration exercise to assess whether biases for tuna vessels are comparable to those for research vessels, and if they are not, to estimate a correction factor for tuna vessel estimates (vessel calibration). Considering tuna vessels for the main survey implies that objective 2 (for which we must estimate trackline detection probability) needs to be addressed for both research and tuna vessels to ensure that objective 1 can be met and estimates from research and tuna vessels are comparable.

Thus, there are several potential sources of bias that need to be evaluated and mitigated through the survey design, for example, vessel calibration, estimation of g(0), and school size calibration. The anticipated use of the new estimates of abundance will in part determine whether it is sufficient to estimate relative abundance, or if absolute abundance estimates are also needed. To this end, we propose to conduct a trial survey, and use long range drones which can operate from a vessel, such as the Flexrotor (http://aerovel.com/flexrotor/) or the HQ-55 (https://latitudeengineering.com/2017/12/hq-55-first-flight/), as the platform from which to collect data for evaluation of biases and estimation of correction factors. Among the various platform options, drones have the potential to be safer and more cost-effective than previously used methods. They may also be able to fly at higher Beaufort sea states compared to helicopters or fixed-wing aircraft, and, in comparison to the latter, operate from the vessel, and hence can be used throughout the survey area.

In developing these survey design options, there are three considerations that we carry forward from previous surveys. First, we assume that the species and stocks of interest are all or a subset of the 10 stocks listed in Gerrodette *et al*. (2008). Second, we assume that the species and stock definitions put forward previously are appropriate and scientifically sound. Finally, for all of the 10 stocks, we assume that the overall survey area, as defined in previous surveys is adequate. As the history of species-specific involvement of dolphins in the purse-seine fishery for tunas differs by species, priority stocks, and hence the survey area, might be defined in two ways:

A. If we take the priority stocks to be the 10 stocks for which Gerrodette *et al*. (2008) gave abundance estimates, all of which have suffered at least some mortality in the ETP purse-seine fishery for tunas, then the logical area to survey is the area for the 2006 survey.

B. If we take the priority stocks to be the northeastern offshore spotted dolphin and eastern spinner dolphin, which are listed as depleted by the MMPA, then the survey area could be taken to be just the strata where these stocks primarily occur: CORE, CORE2 and the NORTH COASTAL.

We consider three possible survey designs. The first, Design 1, is a design that we believe would be most likely to yield results similar in characteristics to previous NMFS surveys for priority stock group A. We also present two other options, Designs 2 and 3, that might be considered if practical constraints and resource limitations prevent implementation of the first survey design. Each of the three survey designs includes the following components: a trial survey, a post-trial assessment, and a main survey.

Design 1 addresses objectives 1 and 2 and includes priority species A. For Design 2, we only aim to address objective 1, i.e. ensure that new estimates will be comparable with previous estimates for priority stocks A. For both Design 1 and 2, the main survey includes two vessels with 120 sea-days each, one or two drones, and the same strata and proportional effort/area allocation as during the 2006 NMFS survey (STAR06). For Design 3, we address objectives 1 and 2, but only for the northeastern offshore spotted dolphin and the eastern spinner dolphin (priority stocks B) so that the survey region will only include the CORE, CORE2 and NORTH COASTAL strata. We assume that one vessel (plus drone) with 120 sea-days (option 1) could complete an equivalent amount of effort for these strata as during STAR06. Using two vessels (plus drones) with 120 sea-days each (option 2), would improve precision on the estimates compared to using one vessel.

Regardless of which survey design is selected, we recommend that the trial survey should be conducted in the traditional survey season, in the late summer or fall (August through November) of 2019, carried out in a relatively small area that is anticipated to have high densities of northeastern offshore spotted and eastern spinner dolphins. The main survey should be conducted in 2020. Should the purpose of the trials include vessel calibration, we propose to conduct a 30-day trial survey with two ships, one research and one tuna vessel. If the trials do not include a vessel calibration, i.e. if only research vessels are considered for the main survey, we propose to conduct a 14-day trial survey which mainly serves to test the ship and equipment, in particular the utility of the drones, and train personnel in the survey protocol for the main survey.

If the drones prove to be unsuitable for collecting the required data, we will need to assess other options. School size calibration data could, for example, be collected using a fixed-wing aircraft as during the STAR06 survey. However, we believe that alternatives such as fixed-wing aircraft or helicopter would not be suitable for collecting the required data for evaluating g(0), which is expected to vary by sea state and location. Consequently, if drones prove unsuitable, objective 2 will have to be dropped. It also follows that the vessel calibration could not be done in which case we recommend that tuna vessels should not be used.

We recommend that either Design 1 or Design 3 with two vessels be adopted. Design 1 has the advantage that it allows estimation of absolute abundance of the ten stocks in Gerrodette *et al*. (2008). In principle, it also allows assessment of whether there is movement of the stocks of offshore spotted dolphins across nominal stock boundaries, but in practice, such inference is hampered by poor precision. We might anticipate that, under Design 1, the coefficient of variation for the estimate of the north-eastern stock of offshore spotted dolphin will be of the order of 20%, while that of the western/southern stock might be around 30%. Given these levels of precision, only extremely large movements across the boundaries could be detected with any confidence. Inference may be further compromised by the inability to implement a randomized even-coverage design in the OUTER region, given the constraints imposed by the need to access a port at regular intervals, and by the limited effort relative to the size of the region achievable in a two-vessel survey.

Under Design 3, we anticipate that, with two vessels operating in the CORE, CORE2 and NORTH COASTAL, we might achieve a CV of around 14%. This gives a better chance of detecting any change since 2006, but of course any change detected could potentially arise from either changes in population abundance or large-scale movement across the stock boundaries or both. Although in principle the eastern spinner dolphin distribution may extend slightly beyond the CORE + CORE2 + NORTH COASTAL areas, previous surveys have reported almost no eastern spinner sightings in the OUTER area, and therefore inference for eastern spinner dolphins may be more secure.

Poor precision under Design 1 will be less problematic if the proposed survey is the first of a series of planned surveys, perhaps conducted once every three or four years, so that ongoing information is gathered, to allow more reliable determination of population trends. In the absence of such an approach, only very large changes since 2006 will be likely to be detected.

1. BACKGROUND

From the early 1980's, the management of eastern tropical Pacific Ocean (ETP) dolphin species and stocks has relied heavily on estimates of abundance from research vessel surveys. Population dynamics models, which take as input estimates of abundance, have been used to estimate population growth rates and other parameters for comparison to expectations (e.g., Smith 1983; Gerrodette & Wade 1991; Reilly *et al*. 2002; Hoyle and Maunder 2004; Wade *et al*. 2007) and to set limits on stock-specific incidental mortality in the tuna purse-seine fishery (IATTC 2006). In addition, abundance estimates from research vessel surveys have been used to estimate population trends (e.g. Gerrodette & Wade 1991, Gerrodette & Forcada 2002, Gerrodette *et al*. 2008). The research vessel surveys that have formed the basis for recent population dynamics modelling and management decisions were conducted by the National Marine Fisheries Service (NMFS) in 1986-1990, 1998-2000, 2003 and 2006 (Gerrodette *et al*. 2008). The current 12-year hiatus in these surveys, and the problematic nature of monitoring stock status from fishery-dependent data (Lennert-Cody *et al*. 2001; 2016), means that there are presently no reliable indicators with which to monitor the status of ETP dolphin populations. A new ETP survey would be a step towards mitigating this situation.

Despite the fact that many details of the design of the NMFS surveys used since 1986 have been carefully vetted, new research, and the passage of time, require that a new survey design be developed before a new survey can be conducted. There are two issues of primary concern. First, previous survey estimates have been assumed to be unbiased estimates of absolute abundance (e.g. Gerrodette & Forcada 2005). Recent research (Barlow 2015) suggests, however, that not all dolphin schools may have been detected on the trackline during previous surveys, even in the excellent survey conditions of Beaufort sea state 1. The implication of this research is that previous abundance estimates may be biased. If estimates of absolute abundance are desired, field methods of previous surveys need to be modified to allow estimation of trackline detection probability. Second, now 12 years on since the last survey in 2006, previous research vessels may not be available for a new survey. Since 2006, both research ships used during STAR06 and STAR03, the *David Starr Jordan* and *McArthur II*, have been decommissioned by NOAA; however, the David Starr Jordan is now owned by Stabbert Maritime (www.stabbertmaritime.com/) and, renamed as *Ocean Starr* and *Ocean Titan*, could potentially be chartered for the survey. Also, Stabbert Maritime owns the Ocean Titan which is very similar to the McArthur II in the technical specifications (see below). Furthermore, two tuna purse-seine vessels, or one tuna purse-seine vessel to be used in conjunction with a research vessel, have been offered for a new survey (MSC 2015). Therefore, new survey vessels, or perhaps tuna purse-seine vessels (hereafter referred to as tuna vessel for brevity), might be used to conduct a new survey. Switching from research vessels to tuna vessels, in particular, has the potential to introduce bias in the abundance estimates for the new survey, which would be problematic for trend estimation. Thus, there are several potential sources of bias that need to be evaluated and potentially mitigated with the design of a new survey.

The extent to which different sources of bias need to be addressed in the new survey design depends in part on the anticipated use of the new estimates of abundance, as this will determine whether it is preferable to obtain estimates of relative abundance, absolute abundance, or both. If the abundance estimates are to be used for trend estimation, the new survey need only produce estimates of relative abundance, provided biases are comparable with past surveys. In this case, it will be important to link to the historical time series of estimates that are available from 1986, and therefore introduction of time-varying biases with the advent of a new survey is of primary concern. If the abundance estimates are to be used to establish limits on incidental mortality, possibly in addition to estimating population trends, estimates of absolute abundance would be necessary, and therefore the survey design must mitigate to the extent possible all potential sources of bias, both time-varying and constant. Independent of the anticipated use of the abundance estimates, it is noted that if there are to be future surveys (beyond the one proposed for 2019/2020), producing estimates that are as unbiased as possible will help to ensure that the next survey estimates are more likely to be comparable to future survey estimates.

In this document we present several design options for a new ETP dolphin survey and the anticipated costs associated with each of those options. In developing these survey design options there are three considerations that we carry forward from previous surveys. First, we assume that the species and stocks of interest are all or a subset of the 10 stocks listed in Gerrodette *et al*. (2008), shown in Table 1 below.

TABLE 1. Dolphin stocks for which Gerrodette *et al*. (2008) presented abundance estimates.

Species	Scientific Name	Stock
Spotted dolphin	*Stenella attenuata*	northeastern offshore
Spotted dolphin	*Stenella attenuata*	western/southern offshore
Spotted dolphin	*Stenella attenuata graffmani*	coastal
Spinner dolphin	*Stenella longirostris orientalis*	eastern
Spinner dolphin	*Stenella longirostris*	whitebelly
Striped dolphin	*Stenella coeruleoalba*	
Rough-toothed dolphin	*Steno bredanensis*	
Short-beaked common dolphin	*Delphinus delphis*	northern, central and southern combined
Bottlenose dolphin	*Tursiops truncatus*	
Risso's dolphin	*Grampus griseus*	

Second, we assume that the species and stock definitions put forward previously (e.g., Dizon *et al*. 1994; Leslie & Morin 2016) are appropriate and scientifically sound. Finally, for all of the 10 stocks, we assume that the overall survey area, as defined in previous surveys (Figure 1; Jackson *et al*. 2004, 2008) is adequate, despite indications that the distribution of some species (e.g., common dolphin) may change geographically on an inter-annual basis due to ENSO events (Reilly and Fiedler 1994). In Section 2 we present the details of the survey design options. The budget for these options is presented in Section 3, and advantages and disadvantages are discussed in Section 4. We present conclusions in Section 5.

2. SURVEY DESIGN

There are three primary considerations with respect to planning a new survey. The first is the objective(s) of the survey, which will affect the survey methodology. The second is the species/stocks of interest, which may affect the definition of the survey area. The third is the available budget for the survey, which may affect both the survey methodology and the survey area. The first two considerations are discussed in more detail immediately below, and for the third consideration, the budget is detailed in Section 3.

2.1. Objectives

We consider two possible objectives of the survey:

1. Estimate relative abundance of priority stocks such that the estimates are comparable as far as possible with past estimates from NMFS surveys.
2. Estimate absolute abundance of the priority stocks.

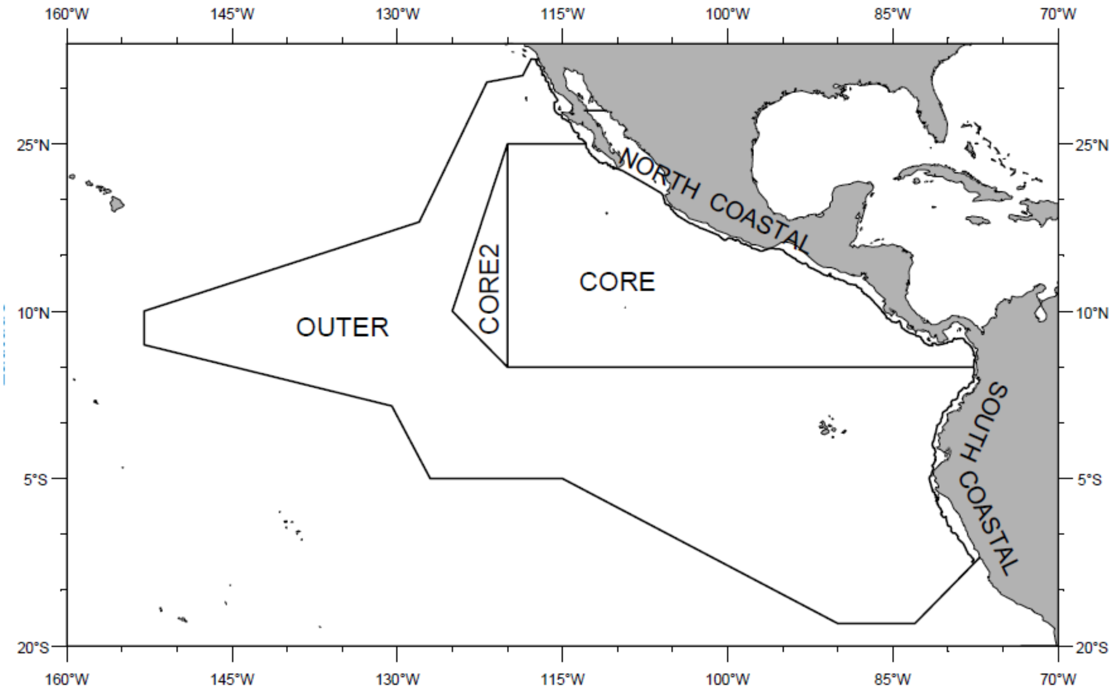

FIGURE 1. Strata for the STAR06 cruise (Gerrodette *et al*. 2008). The CORE area was expanded to include the CORE2 area during the 2003 and 2006 surveys.

2.1.1. Objective 1

The implications are that we need a set-up as close to that of earlier surveys as possible. This entails using the same survey protocol of closing mode effort for the primary observer team on the flying bridge (Kinzey *et al*. 2000b). Using exactly the same survey protocol implies conducting the survey during the same time of year with the same marine mammal observers, cruise leaders, survey vessels, and survey effort. However, because the most recent survey was conducted 12 years ago, using exactly the same protocol may be challenging. For example, some observers and cruise leaders are no longer available and therefore a new survey must involve at least some new personnel. Nonetheless, to the extent possible, previous ETP marine mammal observers and cruise leaders should be used as field personnel as they are experienced in implementing this protocol, as well as in identifying ETP cetaceans in the field to species and stock level. In addition, previous ETP observers have also been calibrated for school size estimation during previous ETP surveys and hence, their calibration factors can be estimated with better precision than those of new observers.

The vessels to be used in the survey are also a very important consideration. If the same vessels used for the 2006 survey (former NOAA research vessels David Starr Jordan and McArthur II) are no longer available, we would like to use research vessels with specifications that match closely with those vessels, e.g. the Ocean Titan instead of the McArthur II. Specifications for the STAR06 research vessels are given in Table 2.

Similarly, if tuna vessels are to be used in the new survey, they should be chosen to have specifications as close as possible to the research vessels used in the past (Table 2). Tuna vessel specifications would need to be obtained on an individual-vessel basis. However, for comparison to the information in Table 2, a few generalizations can be made based on the most recent information of vessels in the IATTC data base with a length ≥ 52.1m (the length of the smaller research vessel) and a fish-carrying capacity of > 363 t (IATTC Class-6 vessels) (information on 120 vessels are summarized here). The median length of these vessels was 63m (inter-quartile range (IQR): 60-70m), and thus closer in length to the McArthur II than to the

David Starr Jordan. The vast majority of these vessels have only one propeller, compared to the research vessels with two propellers. These tuna vessels have a faster cruising speed than the research vessels. The median cruising speed was 14 knots (IQR: 13-14.5 knots). The minimum cruising speed was 10 knots, which is about the cruising speed of the two research vessels. The median number of days at sea per trip was 60 (IQR: 46-72d), which is longer than the days at sea in previous survey legs, but the actual range of the tuna vessels is not known. The median fish-carrying capacity of this group of tuna vessels was 1330 cubic meters (IQR: 1161 – 1585 cubic meters), but a comparison of overall vessel size is not possible without information on the total cubic meter volume. Finally, assuming that roughly half the personnel aboard a tuna vessel during normal fishing operations would not be needed during a dedicated dolphin survey, we believe that vessels in this group would be able to accommodate about the same number of scientists as the research vessels.

Table 2. Specifications of the two research vessels for STAR06 and STAR03 (Jackson *et al*. 2008, https://en.wikipedia.org/wiki/NOAAS_David_Starr_Jordan_(R_444) and https://en.wikipedia.org/wiki/USNS_Indomitable_(T-AGOS-7)).

	David Starr Jordan	McArthur II
Total binocular height above sea level	10.7 m	15.2 m
Ship length	52.1 m	68.3 m
Gross Tonnage	873 GT	1,486 GT
Propulsion	Two 534-hp (398-kW) White Superior diesel engines, variable-pitch propellers	Diesel-electric; two shafts, fixed pitch propellers
Speed	10 knots (19 km/h) (cruising)	10.5 to 11 knots (19.4 to 20.4 km/h) (sustained)
Range:	7,500 nautical miles (13,900 km)	8,000 nautical miles (15,000 km)
Complement (scientists)	13	15
Wind protected flying bridge where observation gear can be installed	4 sets of bigeyes + 3 recorder stations each including mounted chair and computer box	4 sets of bigeyes + 3 recorder stations each including mounted chair and computer box

If tuna vessels were to be used in the new survey, there would need to be a calibration exercise to assess whether biases for tuna vessels are comparable to those for research vessels, and if they are not, to estimate a correction factor for tuna vessel estimates (vessel calibration). This vessel calibration is necessary because of the history of tuna vessel involvement in the purse-seine fishery on tunas associated with dolphins (e.g. NRC 1992), which may elicit behavioral responses from dolphins different from their behavioral responses to research vessels (see Section 2.4). The vessel calibration would need to be done in a trial survey that would take place prior to the main survey. Details of this trial survey are presented below in Section 2.4. Vessel calibration is assumed not to be necessary if only research vessels are used because, although there are differences among research vessels, they are assumed to be more homogeneous in their characteristics and are assumed unlikely to have been involved in the fishery at all or at least not in the recent fishery.

2.1.2. Objective 2

One of the critical assumptions for estimating absolute abundance using the conventional distance sampling methods that were used previously (e.g. as in Gerrodette *et al*. 2008) is that all schools on the transect line are detected. However, questions have been raised concerning whether probability of detection of schools on the transect line – often referred to as g(0) – is close to one in all sea states up to Beaufort 5 (Table 3; Barlow 2015). Therefore, to meet Objective 2, methods will be needed to estimate this probability, and how it varies by sea state. Given that some schools that were initially close to the line may be evading detection, these methods should accommodate responsive movement and behavioral responses. If Objective 2 is selected, a trial survey would be necessary to determine the efficacy of new survey methods for estimation of g(0).

TABLE 3. Estimated values of trackline detection probability g(0) for sightings conditions in Beaufort states 1–6 relative to Beaufort zero and total number of sightings used for these estimates from Barlow (2015). Coefficients of variation (CV) from jackknife method are in italics, and g(0) values significantly different from 1.0 (z-test) are in bold.

	Number of	Beaufort Sea State						
Species	Sightings	0	1	2	3	4	5	6
Stenella longirostris	969	1	**0.733**	**0.537**	**0.394**	**0.289**	**0.212**	**0.155**
			0.03	*0.06*	*0.09*	*0.13*	*0.16*	*0.19*
Stenella attenuata	1,653	1	**0.728**	**0.531**	**0.386**	**0.282**	**0.205**	**0.149**
			0.03	*0.06*	*0.09*	*0.12*	*0.15*	*0.18*

2.2. Priority stocks

The history of species-specific involvement of dolphins in the purse-seine fishery for tunas differs by species (e.g., NRC 1992; Scott *et al*. 2012). Therefore, priority stocks, and hence the survey area, might be defined in two ways, labelled as A and B below.

A. If we take the priority stocks to be the 10 stocks for which Gerrodette *et al*. (2008) gave abundance estimates, all of which have suffered at least some mortality in the ETP purse-seine fishery for tunas (e.g., Wahlen 1986; Hall and Lennert 1994; IATTC 2015), then the logical area to survey is the area for the 2006 survey (Figure 1).

B. If we take the priority stocks to be the northeastern offshore spotted dolphin and eastern spinner dolphin, which are listed as depleted by the U.S. Marine Mammal Protection Act (MMPA) (https://www.mmc.gov/priority-topics/species-of-concern/status-of-marine-mammal-species-and-populations/), then the survey area could be taken to be just the strata where these stocks primarily occur (Dizon *et al*. 1994): CORE, CORE2 and the NORTH COASTAL stratum (Figure 1). The northeastern offshore spotted dolphin stock is defined geographically and represents the offshore spotted dolphin stock that is found east of 120° W and north of 5° N (equivalent to the survey areas of CORE and NORTH COASTAL). The eastern spinner dolphin appears largely contained within the CORE + CORE2 + NORTH COASTAL survey areas (Perrin *et al*. 1991).

2.3. Overview of Survey Designs

We consider three possible survey designs. The first is a design that we believe would be most likely to yield results similar in characteristics to previous NMFS surveys for priority stock group A. We also present two other options that might be considered if practical constraints and resource limitations prevent implementation of the first survey design. Each of the three survey designs includes the following components: a trial survey, a post-trial assessment, and a main survey.

In each design, drones are proposed as the platform from which to collect data for evaluation of biases and estimation of correction factors (e.g., vessel calibration, evaluation of g(0) and school size calibration). Among the various platform options, drones have the potential to be safer and more cost-effective than previously used methods. They may also be able to fly at higher Beaufort sea states compared to helicopters or fixed-wing aircraft, and, in comparison to the latter, operate from the vessel and, hence, can be used throughout the survey area. Details of the drone procedures are presented in Section 2.6.

One purpose of the trial survey is to serve as a pilot study before the main survey using the same set of observers, protocol and field equipment as during the main survey. Pilot surveys generally serve the purpose of training scientific personnel, testing the survey protocol, daily operations and equipment for data collection as well as to allow calculations on how much effort is required for a given level of precision. This is particularly important if ships or observers other than previous ETP survey vessels will be used. In addition, the trial survey will serve to evaluate potential sources of bias. The post-trial assessment is for analysis of the trial survey data, and if necessary, re-evaluation of the design of the main survey. There are two different trial survey periods proposed below, depending on whether tuna vessels are to be used in the main survey. Details of the trial survey are presented in Section 2.4 and Appendix 1.

Absent information on the total funding available for a survey, in developing Designs 1-2 we have assumed that the resources to increase the level of survey effort in the OUTER area (Figure 1) will not be available. Were resources to be available, this would be a consideration for Designs 1-2 as the effort in the OUTER area in previous surveys (e.g., Gerrodette *et al*. 2018) was low relative to the size of the area, undoubtedly contributing to poor precision for those stocks for which the OUTER area plays a prominent role in their distribution (e.g., the western/southern stock of offshore spotted dolphin).

2.3.1. Design 1: address Objectives 1 and 2, priority stocks A

In Design 1 we aim to obtain estimates of both relative and absolute abundance for priority stocks A (see Section 2.2).

Trial survey

IF TUNA VESSELS ARE INVOLVED IN THE MAIN SURVEY

A 30-day trial survey with one research and one tuna vessel in a small study area with expected high abundance for at least the two main stocks, the northeastern offshore spotted dolphin and the eastern spinner dolphin (see Section 2.4), each vessel equipped with a drone:

- Test drone operations for school size calibration
- Test drone operations in terms of their ability to collect data to allow estimation of g(0) (as a function of various covariates)
- Collect g(0) and school size calibration data
- Collect line transect data to compare density, detection functions, encounter rate, school size and trackline detection probability estimates between the two vessels (vessel calibration)

Post-trial assessments:

- Assess utility of drones for collecting school size calibration data
- Assess potential differences in line transect estimates between the two vessels
- If possible, assess g(0) estimates between vessels for Beaufort sea states 0-5
- Assess if both vessels need to be outfitted with a drone for the full duration of the main survey

IF TUNA VESSELS ARE NOT INVOLVED IN THE MAIN SURVEY

A 14-day trial survey with one research vessel in the same study area as would be used above (see Section 2.4):

- Test drone operations for school size calibration
- Test drone operations in terms of their ability to collect data to allow estimation of g(0) (as a

function of various covariates)
- Collect g(0) and school size calibration data

Post-trial assessments:

- Assess utility of drones for collecting school size calibration data
- If possible assess g(0) estimates for both vessels for Beaufort sea states 0-5
- Assess if both vessels need to be outfitted with a drone for the full duration of the main survey or if we can reduce the time each vessel uses a drone

Main survey

- Two vessels, 120 sea-days each
- One or two drones (if one, switch vessels e.g. half way).
- Same strata (Figure 1) and proportional effort allocation per stratum as during STAR06
- Transect line placement with proportional effort/area allocation per stratum as STAR06

2.3.2. Design 2: address Objective 1, priority species A

In Design 2 we only aim to ensure that new estimates will be comparable with previous estimates.

Trial survey

IF TUNA VESSELS ARE INVOLVED IN THE MAIN SURVEY

This does not change from Design 1 if tuna vessels will be involved as we need to assess how density estimates compare between the vessels as well as assessing g(0) for the priority species for both research and tuna vessels to make estimates comparable.

IF TUNA VESSELS ARE NOT INVOLVED IN THE MAIN SURVEY

A 14-day long survey with one research vessel in a study area with expected high abundance of above two stocks:

- Test drone operations for school size calibration
- Collect school size calibration data

Post-trial assessments:

- Assess if school size calibration data is sufficient to remove the need for a drone for the main survey

Main survey

- Two vessels, 120 sea-days each
- One or two drones (if one, switch vessels e.g. half way).
- Same strata and proportional effort allocation per stratum as during STAR06
- Transect line placement with proportional effort/area allocation per stratum as STAR06

2.3.3. Design 3: address Objectives 1 and 2, priority species B

In Design 3 we assume that only the northeastern offshore spotted dolphin and the eastern spinner dolphin are of interest so that the western and southern regions within the ETP need not be surveyed. This could allow precision similar to previous surveys to be achieved with fewer resources. We will assume that the survey region will be the CORE, CORE2 and NORTH COASTAL strata of the last two surveys (Figure 1). During previous ETP surveys, effort within these three strata combined (~12,100km for STAR06, Table 4) was similar compared to the effort in the OUTER and SOUTH COASTAL strata combined (~9,200km for

STAR06).

TABLE 4. Area, effort, number of transects, and number of dolphin sightings in 2006 used to estimate abundance, by stratum. Strata are shown in Fig. 1 (Gerrodette *et al*. 2008).

Stratum	CORE	CORE2	OUTER	N. COASTAL	S. COASTAL
Area (10^6 km^2)	5.869	0.592	14.186	0.535	0.171
Effort (km)	10,268	768	9,131	1,027	35
Number of transects	98	5	68	22	1
Offshore spotted	102	7	21	4	0
Eastern spinner	63	4	0	1	0

The trial survey and post-trial assessment options are the same as those of Design 1, and therefore are not repeated here. The main survey, which is different from that of Design 1, is outlined below.

Main survey: option 1

- One vessel, 120 sea-days with one drone for the duration.
- Only CORE, CORE2 and NORTH COASTAL strata, with same amount of effort for these strata as during STAR06 (Figure 1).
- New transect line placement is required (see Section 2.5).

Main survey: option 2

- Two vessels, 120 sea-days each with a drone each for the duration.
- Only CORE, CORE2 and NORTH COASTAL strata, with same amount of effort as was achieved in all five strata during STAR06 (Table 4), hence, improving precision compared to Design 1 for the CORE, CORE2 and NORTH COASTAL strata, as compared to estimates from previous surveys for the same strata.
- New transect line placement is required (see Section 2.5).

2.3.4. General survey timeline

Presented in Tables 5 and 6 below is a general timeline for the trial and main surveys, which is largely independent of which survey design is selected.

TABLE 5. Milestones for the trial survey. Some dates are subject to ship availability. Shown survey end dates reflect a five day transit to and from the study area for both the 14-day and the 30-day trial survey and a two day port call for the latter.

Milestone	Deadline
Submit research clearance requests	September 2018
Draft tracklines & other survey design specifics	January 2019
Rent or purchase equipment; **Hire sea-going scientists**	February 2019
Pre-survey meetings with ship; **Draft survey instructions**	April 2019
Confirm foreign government observers; complete foreign travel paperwork	May 2019
Draft loading plan, pre-survey scientific meeting agenda; **Finalize track lines, ports of call, survey instructions**	June 2019
Load ship(s); **Conduct pre-survey scientific meeting**	July 2019
Survey begins	28 July 2019
Survey ends	14-day trial: 22 August 2019 30-day trial: 9 September 2019
Draft survey report including post-trial assessment	January 2020
Final survey report	April 2020

TABLE 6. Milestones for the main survey. Some dates are subject to ship availability.

Milestone	Deadline
Submit research clearance requests	September 2019
Draft tracklines & other survey design specifics	January 2020
Rent or purchase equipment; **Hire sea-going scientists**	February 2020
Pre-survey meetings with ship; **Draft survey instructions**	April 2020
Confirm foreign government observers; complete foreign travel paperwork	May 2020
Draft loading plan, pre-survey scientific meeting agenda; **Finalize track lines, ports of call, survey instructions**	June 2020
Load ships; **Conduct pre-survey scientific meeting**	July 2020
Survey begins	28 July 2020
Survey ends	7 December 2020
Draft survey report	May 2021
Final survey report	November 2021

2.4. Trial survey details

During the trial survey, each vessel will be set up in the same way as previous research vessels, with (at least) two 25x pedestal-mounted binoculars. In addition, each vessel will have a long-range drone which can operate from a vessel, such as the Flexrotor (http://aerovel.com/flexrotor/) or the HQ-55 (https://latitudeengineering.com/2017/12/hq-55-first-flight/), together with two pilots.

School size calibration

During previous ETP surveys, calibration of school size estimates for ETP observers was done by comparing estimates to counts from aerial photographs taken from manned helicopter or fixed-wing aircraft (e.g. Gerrodette *et al*. 2018). For this survey, we propose to replace the helicopter or fixed-wing aircraft with a drone to collect equivalent still photographs or video of the dolphin groups. During the trial, we need to assess the practicality of using such a drone as the aerial platform for collecting suitable high-resolution imagery. Automated camera equipment aboard the drone will record the high-resolution imagery, which will allow the drone to operate at a height where disturbance of the dolphins is highly unlikely. This is not possible with observers on an aircraft.

Trackline detection probability g(0)

Should Objective 2 be selected, we propose to use the drone to collect data that will allow us to estimate g(0) (the probability that schools on the trackline are detected). The preferred method for addressing the g(0) issue for the ETP survey is mark-recapture distance sampling (MRDS, e.g. Borchers 2012). In comparison to conventional distance sampling where, e.g., line-transect data are collected from a single platform, MRDS methods require double-observer platform data. Here, detections made from one platform, say platform 2, represent trials for the other platform, say platform 1. In this context, trial outcomes refer to whether or not platform 1 detects a group of dolphins initially detected by platform 2. Here it is crucial that the two observation platforms are such that platform 2 does not influence the observers on platform 1. For this survey, a drone would survey the area in front of the ship and serve as platform 2. The sightings made via the camera equipment aboard the drone would represent the trials for the flying bridge observers on platform 1. The drone can survey the area beyond the maximum sighting range of the observers and might simply record high-resolution images for later analysis. If an operator monitors images in real time, this might help to identify duplicate detections – those detected by both drone and observers. See Section 2.6 below for more details.

Vessel calibration

Considering tuna vessels for the main survey requires that data be collected to evaluate if estimates from research and tuna vessels are comparable. Potential reasons for incompatibility pertain mostly to the evasive behavior of the dolphins towards the purse-seine vessels (e.g., Pryor and Norris 1978; Lennert-Cody and Scott 2005) but also to the physical configuration of these vessels with respect to suitability as an observation platform to collect line transect data. If dolphin schools respond to tuna vessels differently than to research vessels, not correcting for these differences would mean that data from a survey conducted by one tuna vessel and one research vessel would not produce comparable estimates and thus, it would not be possible to combine estimates from the two survey vessels in a way that allows comparison with estimates from past research vessel surveys. If both survey vessels were tuna vessels, estimates would again not be comparable to previous research vessel surveys and lack of correction would introduce time-varying bias into the time series of abundance estimates.

Hence, if tuna vessels are considered for the main survey – either one tuna vessel and one research vessel or two tuna vessels – we propose, to conduct the trials simultaneously with two vessels, one research vessel and one tuna vessel. This configuration for the trials is required to calibrate tuna vessels against research vessels (vessel calibration; Objectives 1 & 2) and to allow a new abundance estimate to be made that is comparable with previous estimates (Objective 1). This vessel calibration will likely be the most complex and time-consuming part of the trial as sufficient data need to be collected to estimate potential differences with suitable precision between the two vessel types.

POST-TRIAL ASSESSMENT

School size calibration

Part of assessing suitability of the drone for the task of school size calibration is to determine whether the photographs or video are of equivalent or even improved quality compared to those taken during previous ETP surveys from helicopters or fixed-wing aircraft. They need to allow for obtaining true school sizes with a negligibly small error (e.g. Gerrodette *et al*. 2018) and species composition. Further we will assess for how long we will need the drone during the main survey. If school size calibration is the only reason for the drone, it may be sufficient to operate it on just one survey leg for each vessel.

Trackline detection probability g(0)

Suitability of the drone for collecting data for estimating g(0) will be determined during the trial survey. After the trial, we need to assess whether g(0) can be assumed to be one for Beaufort sea states 0-5, as for previous surveys. If not, we would want the drone throughout the survey on each vessel as g(0) is likely to vary through the region, as well as by sea state, school size and other factors.

Vessel calibration

After the trial, we will compare the estimates of animal density, g(0), detection functions and average detection probabilities within the search area, encounter rate and school sizes between the two vessels, taking into account the observation conditions (e.g., Beaufort sea state, visibility, swell height) encountered by each vessel. These detailed analyses will provide insight into potential differences in estimated animal densities between ships.

Utility of the drone and alternatives

The general utility of the drone for the required operations will likely become evident during the trial survey. The final decision on whether to use a drone for the main survey may depend, however, on a detailed assessment of the data collected by the drone (as described above), under what conditions it can operate and whether we can expect to be able to collect sufficient data for the desired objectives.

If the drones prove to be unsuitable for collecting the required data, aspects of the survey that were to involve evaluation of bias in the probability of detection the trackline would no longer be possible. Specifically, we believe that alternatives, i.e. fixed-wing aircraft or helicopter, would not be suitable for collecting the required data for evaluating g(0). Consequently, if drones prove unsuitable, Objective 2 will have to be dropped. Barlow (2015) provides estimates of g(0) for ETP dolphins for Beaufort sea states 0-6. However, his methods do not include investigating potential interactions between Beaufort sea state and spatially varying factors such as evasive response ("The methods used here cannot truly distinguish between bias due to differences in trackline detection probability and bias caused by responsive movement.", Barlow 2015). Hence, given the implications for absolute abundance estimation of the corrections to g(0) as a function of Beaufort sea state estimated by Barlow (2015), we believe that use of this method to adjust estimates of g(0) would need to be supported by a field study. Moreover, it follows that the vessel calibration could not be done if drones proved unsuitable, in which case we recommend that tuna vessels should not be used.

School size calibration is the only component for which it might be possible to collect data from alternative platforms. School size calibration data could, for example, be collected in the same manner as during the STAR06 survey where a fixed-wing aircraft operated out of airports along the west coast of Mexico (mainly Acapulco) to obtain aerial photography for a scheduled number of consecutives days with each of the two ships (Jackson *et al*. 2008; Gerrodette *et al*. 2008). During the scheduled days, if Beaufort sea state was ≤2, the aircraft flew out to the area of the ship and took photographs of the schools detected by the ship-board observers. The disadvantages of using fixed-wing aircraft for obtaining photographs for school size calibration include that the calibration schools are all from a restricted nearshore area, and all correspond to low sea states. Also, on the days with combined ship/airplane operations, no line transect data can be collected. Thus, the amount of calibration data that can be collected is highly dependent on excellent weather conditions during the days scheduled for these operations.

Another alternative to drones for collecting school size calibration data would be to use a helicopter operating from one of the survey vessels as was done for ETP surveys preceding STAR06. During STAR03, for example, the former David Starr Jordan carried a Hughes 500D helicopter equipped with two large-format military reconnaissance cameras mounted below the fuselage (e.g. Gerrodette *et al*. 2018). A helicopter might be carried aboard one or both of the survey vessels, particularly if the vessels were tuna vessels, and therefore school size calibration data might be collected in a broader region of the ETP than would take place with a fixed-wing aircraft. The disadvantages of using a helicopter are that the collection of data would be limited to low sea states, and placing observers in helicopters comes with an added safety risk that may not be acceptable.

Study area and timing of the trial survey

We propose that the trial survey should take place in the late summer or fall (August through November) of 2019, carried out in a relatively small area that is anticipated to have high densities of northeastern offshore spotted and eastern spinner dolphin stocks. Using a high density area provides the largest sample size of dolphin school encounters for the trials which is needed, in particular, for the vessel calibration. The areas south of the coastline between Manzanillo and Acapulco, Mexico, roughly between 15-18˚N and 100-104˚W have shown consistently highest densities of these stocks regardless of season (Figure 2).

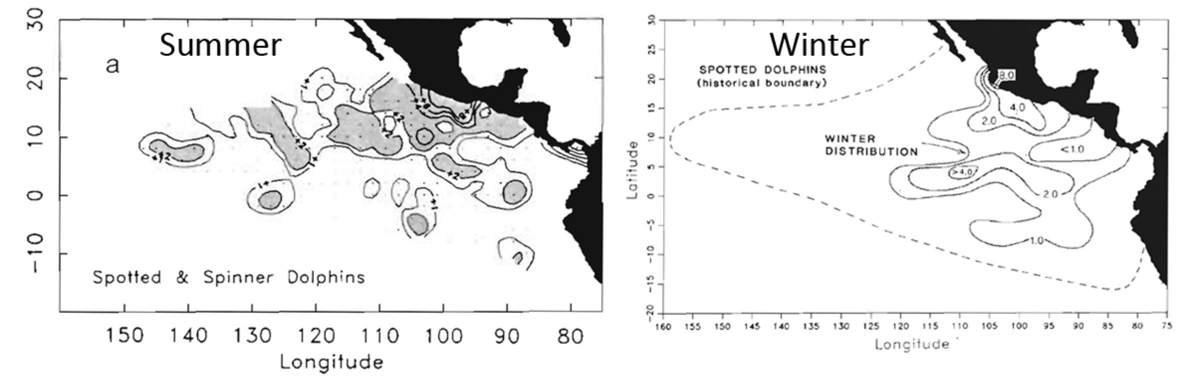

FIGURE 2. Summer and winter distributions of spotted and spinner dolphins in the ETP (contours of encounter rates per 185 km searched). Dots in left plot represent centers of 2˚ squares in which there was at least 185 km search effort. Left Fig. 5 and right Fig. 1 from Reilly (1990). Right plot shows spotted and spinner dolphin historical distribution boundary (shown dashed; from Perrin *et al*. 1983) and winter relative densities (shown as continuous contours of no. schools sighted per 185 km searched; from Au & Perryman (1985). Note that here, summer refers to July-December and winter to January-June.

A further benefit of this particular area is that it is located nearshore within the CORE stratum. ETP dolphins, e.g. spotted dolphins, have been reported to show evasive responses to tuna vessels, the level of this response being stronger in nearshore areas with a longer history of fishing compared to further offshore areas (e.g., Lennert-Cody and Scott 2005). ETP dolphins may also show evasive responses to the research vessels (CSO pers. observation during ETP surveys). During the trials we wish to determine whether the level of evasive response is similar between the two vessel types. Hence, using the suggested area with high densities and expected strong evasive responses will likely provide the largest sample size of evasive schools encountered during the trial survey.

As indicated in Figure 1 of Barlow (2015), the average Beaufort sea state in the proposed trial study area is lower than in much of the main survey area. The average Beaufort sea state in the trial survey study area was about 2.5 -3, compared to an average of about 3-4 over much of the CORE and CORE2 areas, and greater than 4 in the OUTER area. Therefore, during the trial survey, collecting sufficient data at high Beaufort sea states (Beaufort ≥ 4) may not be possible. Despite this, given the fact that g(0) was estimated by Barlow (2015) to be considerably less than 1.0 even at Beaufort sea states 1-2 (see Table 3 above from Barlow (2015)), we anticipate that if the drones prove suitable for collecting the type of data necessary to estimate g(0), even if only in low Beaufort conditions and without regard for species, those data will prove useful for evaluating bias.

Length of the trials

Should the purpose of the trials include vessel calibration, we propose to conduct a 30-day trial survey with two ships, one research and one tuna vessel. We quantitatively assessed what we expect a trial period of 30 days would allow us to determine with regards to differences in the parameter estimates – effective strip width (ESW), trackline detection probability, encounter rate and expected school size – between the two vessels (where we regard 30 days as the practical upper limit of a trial period). Details of this assessment can be found in Appendix 1. This assessment was based on estimates and their precision from previous ETP surveys. We expect that a 30-day trial survey would allow us to detect differences in the ESW of >20% for spotted dolphins and >30% for spinner dolphins, and that it would allow us to detect differences in school sizes of >50% for spotted and >60% for spinner dolphins. We expect that evasive behavior would result in negative bias in encounter rate and trackline detection probability and, hence, we only consider a decrease for these two parameters. We expect that a 30-day trial period would allow us to detect a decrease in encounter rate of >55% for spotted dolphins and >70% for spinner dolphins. We also expect that a 30-day trial period would allow us to detect a decrease in trackline detection probability of >30% for spotted dolphins and >45% for spinner dolphins. To be able to detect smaller changes, we would need to extend the trial survey beyond 30 days.

The results of the vessel calibration will be used to account for potential biases in estimates from the research and tuna vessels and make the estimates from the two platforms comparable. Here we would like to emphasize that this does not simply consist of multiplying the tuna vessel abundance estimate by a constant. It also involves adding the uncertainty associated with the calibration, i.e. the coefficient of variation (CV), to the CV of the estimate. We expect that this will drastically decrease the precision in the final abundance estimates.

If the trials do not include a vessel calibration, i.e. if only research vessels are considered for the main survey, we propose to conduct a 14-day trial survey which mainly serves to test the ship, equipment – in particular the utility of the drones – and train personnel in the survey protocol for the main survey.

Design of the trial survey

For a 30-day trial survey we propose a design consisting of 30 parallel lines in the study area oriented perpendicular to the coastline. Each line will be of the same length of approximately 120nm where the exact length depends on the number of daylight hours. An alternative design that may be considered would be to use fewer lines, but place the lines in pairs, sufficiently close to each other to minimize differences in spatial effects between the two lines, but not so close as to have data collection by one vessel influence the data collected by the other vessel. The two vessels would conduct repeated surveys along these pairs of lines to produce a paired survey design, the data from which could allow effects of Beaufort on g(0) to be more precisely evaluated because any spatial effects on g(0) would be minimized by the transect pairing.

For a 14-day trial survey, we propose a similar design with 14 parallel lines perpendicular to the coastline.

2.5. Main survey details

The main survey will be carried out in the year following the year of the trial survey. This will give sufficient time for the post-trial assessments and to change survey vessels, if necessary, in light of analyses of the trial data. It will also allow time to book the drones and pilots for the main survey, depending on the outcome of the trial survey, and, it will ensure that both the trial survey and the main survey take place during the same season as previous surveys. Depending on the trial results and the objectives, the main survey might be conducted with one or two research vessels, one research and one tuna vessel, or two tuna vessels, as well as with a drone and pilots on each vessel, assuming the trial demonstrated the utility of the drones.

For the design of the main survey, the following considerations were taken into account. We split the total number of survey days into five or six legs, depending on the capacity of the ship, where we took guidance on previously-used research ships for ETP surveys, the former NOAA vessels David Starr Jordan and McArthur II. We assumed that a ship like the David Starr Jordan could complete six legs of ~20 days at sea (maximum 24 days at sea) while a ship like the McArthur II could complete five legs of ~23 days at sea (maximum 29 days at sea). Note that here we define 'days at sea' as a 24 hour period since leaving port. These values were obtained from previous cruise reports (Kinzey *et al*. 1999, 2000a, 2001, Jackson *et al*. 2008, 2004).

When estimating the line length for a given leg between port calls we assumed that the ships will travel at night and that the total length travelled along a line in a given sea day would follow a simple equation:

$$(24 - x)\ hours \times 10nm$$

where the x represents the time spent closing on schools for species identification and school size estimation and varied between high and low density areas as well as considerations pertaining to bad weather that may delay the progression along the lines. Combining weather and closing mode, we assumed that x = 6 for high, x = 4 for medium and x = 2 for low density areas.

We used the same boundary points for strata as for STAR03 and STAR06, summarized in Tables 5 and 6 and illustrated in Figure 1. The Sea of Cortez was truncated at 27.975˚ North. Following previous ETP survey designs, for Designs 1 and 2 we allocated approximately three times the amount of effort / area for the CORE, CORE2 and NORTH COASTAL strata compared to the OUTER and SOUTH COASTAL strata. Each leg begins and ends at a port. Ports that were considered here included San Diego and Honolulu in the USA, Mazatlan, Manzanillo and Acapulco in Mexico, Punta Arenas in Costa Rica, Puerto Quetzal in Guatemala, one of Rodman, Balboa or Panama City in Panama, Manta in Ecuador and Callao in Peru. We note that this list is subject to review in the future.

TABLE 7. Study area boundary points for the STAR06 and STAR03 surveys. The eastern boundary was defined by the coastline of the Americas (Jackson *et al*. 2004, 2008).

Latitude	Longitude
32° 32.12' N	117° 7.34' W
32° 35.37' N	117° 27.82' W
32° 37.61' N	117° 49.52' W
31° 7.97' N	118° 36.30' W
30° 32.52' N	121° 52.00' W
18° 0.00' N	128° 0.00' W
10° 0.00' N	153° 0.00' W
7° 0.00' N	153° 0.00' W
1° 30.00' N	130° 30.00' W
5° 0.00' S	127° 0.00' W
5° 0.00' S	115° 0.00' W
18° 0.00' S	90° 0.00' W
18° 0.00' S	83° 0.00' W
12° 0.00' S	77° 0.00' W

TABLE 8. Strata Boundaries for the STAR06 and STAR03 surveys (Jackson *et al*. 2004, 2008): The coastal strata were inshore of the 1000 m depth contour. The CORE stratum was defined by the following points:

Latitude	Longitude
25° 0.00' N	112° 51.60' W
25° 0.00' N	120° 0.00' W
10° 0.00' N	125° 0.00' W
5° 0.00' N	120° 0.00' W
5° 0.00' N	77° 38.04' W

Example design for Survey Designs 1 & 2

For Design 1 and 2, the study area consists of a combination of five strata: OUTER, SOUTH COASTAL, CORE, CORE2 and the NORTH COASTAL. We provide examples of a two ship survey in Figure 3.

Example designs for Design 3:

For Design 3, the study area consists of a combination of three strata: CORE, CORE2 and the NORTH COASTAL. We provide examples of the single ship and two ship options in Figures 4 and 5.

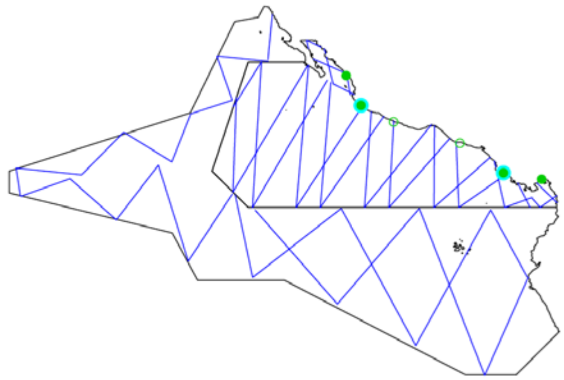

FIGURE 3: Example of a two vessel survey in the all strata including OUTER, CORE, CORE2, NORTH and SOUTH COASTAL strata. Ports are shown in green (Manzanillo, Mexico and Punta Arenas, Costa Rica highlighted in cyan). The actual realised total transect length for this survey is approximately 61,000km. We note that this includes travel time at night.

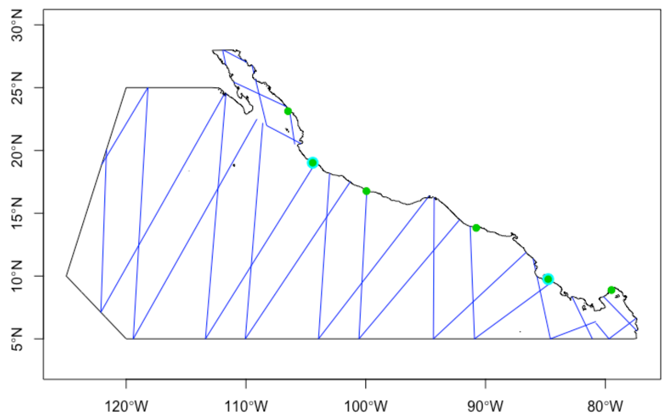

FIGURE 4: Example of a single vessel survey in the CORE, CORE2 and NORTH COASTAL strata using a 500km equal spaced zigzag design. Ports are shown in green (Manzanillo, Mexico and Punta Arenas, Costa Rica highlighted in cyan). The actual realised total transect length for this survey is approximately 31,000km. We note that this includes travel time at night.

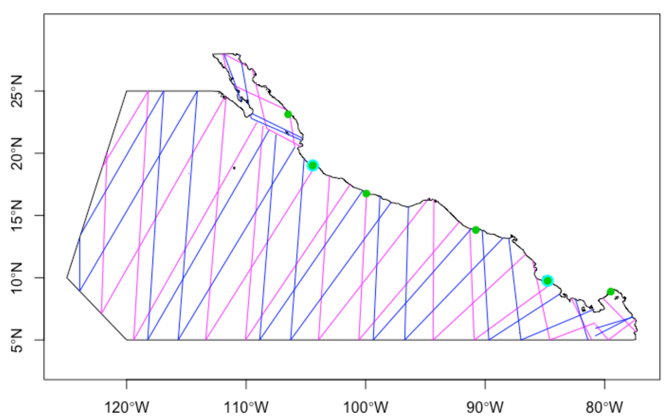

FIGURE 5: Example two vessel survey in the CORE, CORE2 and NORTH COASTAL using a 500km equal spaced zigzag design. Ports are shown in green (Manzanillo, Mexico and Punta Arenas, Costa Rica highlighted in cyan). The actual realised total transect length for this survey is approximately 62,000km. We note that this includes travel time at night.

2.6. Drone operations

The drones will be operated from the vessel, as far away as possible, whenever the primary observer teams on the flying bridge of the respective vessels are on-effort. Generally this entails that the ship moves along the transect lines at a constant speed of 10 nm. However, when the primary team detects a school of cetaceans, the ship closes on the school for obtaining school size estimates and species id which generally requires course and speed changes of the ship. As g(0) is likely to be lower in poor conditions, the drone would ideally need to operate up to Beaufort 5, and it would need to be able to stay aloft for extended periods, to ensure that it is searching for most of the time that the observers are on effort. Weather permitting, the drones will be launched at sunrise just before flying bridge observers begin daily effort and retrieved at sunset at the conclusion of effort. This will provide maximum time use of the drone and minimal interruptions for the survey effort of the primary observer team which requires constant speed of at least 8 knots. It is at the discretion of the drone pilots to request changes in speed of the ship for launching and landing of the drone.

The drones will operate several km ahead of the vessels, flying a zig-zag pattern across the trackline out to ~4km (exact distance to be determined during the trials) either side of the line, with the angle of the zig-zag determined to allow the drone to maintain station ahead of the ship. They will fly at an altitude that is unlikely to generate a response from the dolphins. It will be part of the trials to determine the best zig-zag pattern and the altitude at which the drones will be flying. High-resolution video will be recorded, to allow examining the footage for species identification, school composition and obtaining (at least approximate) true school size counts used for observer calibration at the analysis stage. This imagery will also be transmitted to the ship in real-time and monitored by drone staff on the vessel. If a school is detected by the staff monitoring the drone footage, they will alert the cruise leader (but not the flying bridge observers), and if the flying bridge observers subsequently detect the same school, both the drone and the ship will close on the school to secure better data on school size and species present, and to record any movement of the school towards or away from the line following initial detection by the drone. If it is considered feasible for the drone to monitor a school without alerting shipboard observers, it will do so for schools detected from the air, until either the shipboard observers detect the school or it passes abeam; however, to avoid cueing observers, it is likely that the drone will need to remain some distance ahead of the vessel unless the shipboard observers detect the school, as any change in flight pattern when a school is detected by the drone might cue observers to the presence of a school.

Feasibility of drone operation protocol

The general feasibility of the drone operations as described above was assessed using the expert opinion of two representatives from companies that operate suitable drones for our purposes, Mike Gomez from Precision Aviation and Dr. Aaron Farber from Latitude Engineering. Precision Aviation have the Flexrotor (https://www.flyprecision.com/unmanned/#1492191574538-466f6771-005d) and Latitude Engineering the HQ-55 (https://latitudeengineering.com/2017/12/hq-55-first-flight/). Both experts agreed that the survey protocol sounded generally feasible but that details need to be tested in the field. Further details can be found in Appendix 2.

We also did a preliminary evaluation of what the coverage probability of the drone (the proportion of the transect line that would be captured by the drone footage) would be based on a zigzag flight pattern of the drone intersecting the transect line at a constant distance ahead of the ship. This constant distance depends on the behavior of the dolphin schools and will be determined during the trials. The aim is to capture the dolphin schools with the drone footage before they reacted to the presence of the ship.

Here we used an endurance speed of the Flexrotor drone of 85km/hr and the ship survey speed of 18.5km/hr (10knots). If the drone is to survey in a zigzag pattern out to 4km to either side of the transect line, the drone is expected to cross the transect line at approximately 1.78km intervals. If the drone flies at 300m above sea level where the strip width in wide-angle mode is 368m (Table 9), the resulting coverage probability of the transect line is 0.21 (0.368km/1.78km). If the drone flies at 500m, the resulting coverage probability of the transect line is 0.34 (0.613km/1.78km). It is important, however, that the drone footage will provide enough ground resolution to detect dolphin schools.

TABLE 9. Drone flight height, resulting survey strip width covered by the video footage as well as ground resolution (cm per pixel) using 720 x 1280 video (GR_{1280px}) and 1080 x 1920 video (GR_{1920px}) shown for the maximum and minimum angle of the Trillium camera Orion HD50 (http://w3.trilliumeng.com).

	Wide angle: 63˚			Narrow angle: 2.2˚		
Height (m)	**Strip width (m)**	**GR_{1280px} (cm)**	**GR_{1920px} (cm)**	**Strip width (m)**	**GR_{1280px} (cm)**	**GR_{1920px} (cm)**
100	123	9.6	6.4	4	0.3	0.2
200	245	19.2	12.8	8	0.6	0.4
300	368	28.7	19.2	12	0.9	0.6
400	490	38.3	25.5	15	1.2	0.8
500	613	47.9	31.9	19	1.5	1.0
600	735	57.5	38.3	23	1.8	1.2
700	858	67	44.7	27	2.1	1.4
800	980	76.6	51.1	31	2.4	1.6
900	1103	86.2	57.5	35	2.7	1.8
1000	1226	95.8	63.8	38	3	2.0

Capturing footage for school size calibration will bring further challenges in that these images require both a ground resolution good enough for species or stock identification as well as a wide enough strip to cover the school with as few as possible passes. We acknowledge that using the drone for our purposes will be a challenge, in particular with regards to school size calibration. However, it is the purpose of the trial to determine the feasibility of these operations.

3. SURVEY BUDGET

We present four examples for the survey budget, each including the use of a research vessel (the Ocean Starr or an "in-kind" research vessel [see below]) and the Flexrotor drones for the full duration. The Ocean Starr has been included in each example budget because it is the least costly of the two formerly used survey vessels and it has the longest history of involvement in the previous surveys. Given the survey design considerations noted in the sections above, it is assumed that the vessels that participate in the trial survey will also participate in the main survey. The budget for Design 1 presented in Table 10 includes a 14-day trial survey with a single research vessel and a second research vessel for the main survey. The budget for Design 1 presented in Table 11 includes a 30-day trial survey and the main survey with a research vessel and a tuna vessel. We present two budgets for Design 3 based a 14-day trial survey and either a single research vessel in the main survey (option 1, Table 12) or two research vessels in the main survey (option 2, Table 13). The cost of the drone rental shown in Tables 10-13 includes the cost of two pilots. We assume that equipment (computing, flying bridge and communications) purchased for the trial will be brought forward to the main survey.

Complete cost estimates were obtained from several research vessels. Complete quotes were received for the Ocean Starr and the Ocean Titan (both currently of Stabbert Maritime, where the former was previously used for NMFS ETP surveys and the latter is very similar to the McArthur II, the other survey vessel used for STAR03 and STAR06), the Sally Ride (Scripps Institution of Oceanography, La Jolla CA) and the Nautilus (Ocean Exploration Trust, New London CT). The total cost of the Sally Ride and the Nautilus exceeded those of the Ocean Starr and the Ocean Titan, and thus were not used in developing the example survey budgets shown below in Tables 10-13.

Partial cost information was obtained from one other research vessel. An estimate of the 2018 daily operational cost of the El Puma (Universidad Nacional Autónoma de México; UNAM) was available: US $18K per day. While this only represents a quote for 2018, the daily rates for 2019 or 2020 may vary. It is noted that the El Puma also has a sister ship at the UNAM, the Justo Sierra, that currently operates in the Gulf of Mexico. The suitability of the El Puma (and Justo Sierra) for dolphin surveys would need to be determined via an inspection. Their online profiles[3] suggest they may be of sufficient size (50m total length), albeit shorter than the Ocean Starr, and sufficient cruising speed (12.5 knots). While the daily charter rate for the El Puma is higher than for the Ocean Starr, the latter was provided without including fuel costs or other consumables (lube oil and accommodation for scientists). After including these costs (with expected fuel costs estimated assuming a constant cruising speed of 10knots while at sea and using US $4 per gallon), the total costs for the Ocean Starr was higher than for the El Puma (averaging ~US $19.4K per day). However, if fuel can be purchased at a lower rate, this daily rate for the Ocean Starr with consumables will drop substantially. Costs aside, if the drones prove not to be useful for evaluation of g(0), then use of the Ocean Starr would be preferable from the point of view of improving comparability to previous surveys. In addition, compared to the Ocean Starr the El Puma is smaller in length and breadth. As a result, we expect that the El Puma provides a less stable platform with negative effects on observation conditions for flying bridge observers and a reduction in the amount of suitable conditions for drone operations as launching and retrieving the drone requires a stable platform.

The Alpha Helix (Centro de Investigación Científica y de Educación Superior de Ensenada, México) was also considered for the survey. However, its maximum speed (9 knots) is less than that necessary for closing mode in dolphin surveys (a minimum of 10 knots).

The use of an "in-kind" research vessel and a tuna vessel are also considered in the budget examples shown below. The report of the Marine Stewardship Council (MSC 2015) states that an in-kind contribution of the cost of a tuna vessel and of a research vessel (or of two tuna vessels) will be made for the survey by the Pacific Alliance for Sustainable Tuna ("Alliance")[4]. The dollar value of these in-kind vessel donations is not known to us. To construct example survey budgets for options that would use either the in-kind research vessel and/or a tuna vessel, we have assumed that all daily operating costs will be paid by the vessel donor, and therefore in the Tables 10-13, we show a value of US$ 0 for the survey vessel costs associated with these vessels. As part of this, we have further assumed that all costs related to outfitting any of these vessels for the survey, including the cost of construction and installation of a flying bridge platform for marine mammal observers and binocular mounts, a trial run (~half day) at sea, and transit costs to/from San Diego for loading/unloading the vessel, will be paid by the vessel donor. We note that obtaining estimates of such additional costs would require the specific vessels that are being considered to be inspected.

---

[3] http://www.buques.unam.mx/el-puma/descripcion-puma/ and http://www.buques.unam.mx/justo-sierra/descripcion-js/

[4] From page 352 of MSC (2015), item M-I DOL 2d: "By the first annual audit, the Alliance will provide a copy of the notarized binding agreement that was made with IATTC upon certification that outlines an in-kind contribution in the form of two tuna vessels or one tuna vessel and one oceanographic vessel, including all supplies, crew time, and fuel needed to conduct line transect studies that are designed by lead scientists managed by the IATTC. …"

TABLE 10. Budget outline for a 14-day trial with the Ocean Starr or an in-kind research vessel, and the main survey with two research vessels (one the Ocean Starr and the other the in-kind research vessel), for priority stocks A (Design 1). Square brackets indicate the in-kind survey option. [a]: The dollar value of the donation of the in-kind research vessel is not known to us. To construct an example survey budget for the in-kind option, we have assumed that all daily operating costs will be paid by the vessel donor, and therefore in the table below we show a value of US$ 0 for the in-kind survey vessel cost. We have further assumed that all costs related to outfitting this vessel for the survey, including the cost of construction and installation of a flying bridge platform for marine mammal observers and binocular mounts, and transit costs to/from San Diego for loading/unloading the vessel, as well as ship inspection by the chief scientist in the ship's home port including a trial run (~half day) at sea, will be paid by the vessel donor. [b]: A 5% contingency computed on non-salary costs, and to be used, if necessary, to cover unanticipated expenses. Numbers shown in bold font indicate section totals. The cost of project support from existing IATTC staff is not included in the budget.

	14-day trial Ocean Starr [In-kind research vessel]	Main survey 2 ships Ocean Starr + Ocean Titan [Ocean Starr + In-kind research vessel]
	(US$ 1000)	(US$ 1000)
Cetacean abundance		
Chief Scientist	99.41	281.51
Scientific support to Chief Scientist	87.82	197.41
Cruise leaders	22.29	335.20
Marine Mammal Observers	98.10	955.53
Survey coordinator	52.004	143.93
	359.62	**1,913.57**
Vessels and associated costs		
Survey vessels	720.32[0 [a]]	7,248.40[2,705.84[a]]
Ship agent + Port fees	20.25	290.50
	740.57 [20.25[a]]	**7,538.90[2,996.34[a]]**
Foreign observers		
Travel	**2.00**	**22.00**
Drones		
Equipment rental	420.00	4080.00
Observers	32.70	322.28
Fuel	0.26	3.96
	452.96	**4,406.24**
School size calibration		
Image analyses	58.01	97.64
Data cleaning and analyses	13.82	55.28
	71.8366	**152.92**
Trackline detection probability		
Image analyses	38.34	106.59
Data cleaning & analysis	20.73	69.10
	59.07	**175.68**
Equipment		
Computers	17.38	17.38
Flying bridge	104.36	104.36
Communications	17.77	164.17
Research permits	3.00	15.00
Miscellaneous travel	6.00	15.00
IT specialist	10.34	4.63
	158.84	**320.54**

IATTC headquarters-based contractors		
Accountant	130.57	124.04
Logistical coordinator	112.01	106.41
	242.58	**230.45**
Ship loading	**3.00**	**6.00**
Contingency [b]	**67.35 [31.34 [a]]**	**614.45 [387.32 [a]]**
Total	**2,157.82[1,401.48 [a]]**	**15,380.76 [10,611.07 [a]]**

TABLE 11. Budget outline for a 30-day trial with the Ocean Starr or an in-kind research vessel, and a tuna vessel, and the main survey with two vessels, a research vessel (either the Ocean Starr or the in-kind research vessel) and a tuna vessel, for priority stocks A (Design 1). Square brackets indicate the in-kind survey option. [a]: The dollar value of the donation of the in-kind research vessel and tuna vessel is not known to us. To construct an example survey budget for the in-kind option, we have assumed that all daily operating costs will be paid by the vessel donor, and therefore in the table below we show a value of US$ 0 for the in-kind survey vessel cost. We have further assumed that all costs related to outfitting these vessels for the survey, including the cost of construction and installation of a flying bridge platform for marine mammal observers and binocular mounts, and transit costs to/from San Diego for loading/unloading the vessels, as well as ship inspection by the chief scientist in the ship's home port including a trial run (~half day) at sea, will be paid by the vessel donor. [b]: A 5% contingency computed on non-salary costs, and to be used, if necessary, to cover unanticipated expenses. Numbers shown in bold font indicate section totals. The cost of project support from existing IATTC staff is not included in the budget.

	30-day trial Ocean Starr + tuna vessel [In-kind research vessel + tuna vessel]	Main survey 2 ships Ocean Starr + tuna vessel [In-kind research vessel + tuna vessel]
	(US$ 1000)	(US$ 1000)
Cetacean abundance		
Chief Scientist	145.81	281.51
Scientific support to Chief Scientist	129.43	197.41
Cruise leaders	82.87	335.20
Marine Mammal Observers	297.20	955.53
Survey coordinator	76.06	143.93
	731.37	**1,913.57**
Vessels and associated costs		
Survey vessels	1,103.34 [0[a]]	2,705.84 [0[a]]
Ship agent + Port fees	84.00	290.50
	1,187.34[84.00]	**2,996.34[290.50]**
Foreign observers		
Travel	**8.00**	**22.00**
Drones		
Equipment rental	1,380.00	4,080.00
Observers	109.22	322.28
Fuel	1.04	3.96
	1,490.24	**4,406.24**
School size calibration		
Image analyses	71.33	97.64
Data cleaning and analyses	27.64	55.28
	98.97	**152.92**
Trackline detection probability		
Image analyses	56.66	106.59
Data cleaning and analyses	41.46	69.10
	98.01	**175.68**
Equipment		
Computers	26.76	0.00
Flying bridge	208.72	0.00
Communications	57.14	163.20
Research permits	6.00	15.00
Miscellaneous travel	10.00	15.00
IT specialist	11.68	4.63

	320.30	**197.83**
IATTC headquarters-based contractors		
Accountant	130.57	124.04
Logistical coordinator	112.01	106.41
	242.58	**230.45**
Ship loading	**6.00**	**6.00**
Contingency [b]	**150.01 [94.84 [a]]**	**381.19 [245.90 [a]]**
Total	**4,332.82[3,174.31[a]]**	**10,482.22[7,641.09 [a]]**

TABLE 12. Budget outline for a 14-day trial and the main survey with one research vessel the Ocean Starr or an in-kind research vessel), for priority stocks B (Design 3, main survey option 1). Square brackets indicate the in-kind survey option. [a]: The dollar value of the donation of the in-kind research vessel is not known to us. To construct an example survey budget for the in-kind option, we have assumed that all daily operating costs will be paid by the vessel donor, and therefore in the table below we show a value of US$ 0 for the in-kind survey vessel cost. We have further assumed that all costs related to outfitting this vessel for the survey, including the cost of construction and installation of a flying bridge platform for marine mammal observers and binocular mounts, and transit costs to/from San Diego for loading/unloading the vessel, as well as ship inspection by the chief scientist in the ship's home port including a trial run (~half day) at sea, will be paid by the vessel donor. [b]: A 5% contingency computed on non-salary costs, and to be used, if necessary, to cover unanticipated expenses. Numbers shown in bold font indicate section totals. The cost of project support from existing IATTC staff is not included in the budget.

	14-day trial Ocean Starr [In-kind research vessel]	**Main survey 1 ship Ocean Starr [In-kind research vessel]**
	(US$ 1000)	(US$ 1000)
Cetacean abundance		
Chief Scientist	99.41	280.79
Scientific support to Chief Scientist	87.82	181.50
Cruise leaders	22.29	123.66
Marine Mammal Observers	98.10	477.77
Survey coordinator	52.00	88.75
	359.62	**1,152.47**
Vessels and associated costs		
Survey vessels	720.32[0[a]]	2,665.84 [0[a]]
Ship agent + Port fees	20.25	145.25
	740.57 [20.25[a]]	**2,811.09 [145.25[a]]**
Foreign observers		
Travel	**2.00**	**12.00**
Drones		
Equipment rental	420.00	2040.00
Observers	32.70	161.14
Fuel	0.26	1.98
	452.96	**2,203.12**
School size calibration		
Image analyses	58.01	56.57
Data cleaning and analyses	13.82	41.46
	71.83	**98.03**
Trackline detection probability		
Image analyses	38.34	61.05
Data cleaning and analyses	20.73	55.28
	59.07	**116.32**
Equipment		
Computers	17.38	0.00
Flying bridge	104.36	0.00
Communications	17.77	164.17
Research permits	3.00	7.50
Miscellaneous travel	6.00	15.00
IT specialist	10.34	2.32
	158.84	**188.89**
IATTC headquarters-based contractors		

Accountant	130.57	124.04
Logistical coordinator	112.01	106.41
	242.58	**230.45**
Ship loading	**3.00**	**6.00**
Contingency [b]	**67.35 [31.34 [a]]**	**260.94 [127.65 [a]]**
Total	**2,157.82 [1,401.49 [a]]**	**7,079.31[4,280.18 [a]]**

TABLE 13. Budget outline for a 14-day trial with the Ocean Starr or an in-kind research vessel, and the main survey with two research vessels (the Ocean Starr and the in-kind research vessel), for priority stocks B (Designs 3, main survey option 2). Square brackets indicate the in-kind survey option. [a]: The dollar value of the donation of the in-kind research vessel is not known to us. To construct an example survey budget for the in-kind option, we have assumed that all daily operating costs will be paid by the vessel donor, and therefore in the table below we show a value of US$ 0 for the in-kind survey vessel cost. We have further assumed that all costs related to outfitting this vessel for the survey, including the cost of construction and installation of a flying bridge platform for marine mammal observers and binocular mounts, and transit costs to/from San Diego for loading/unloading the vessel, as well as ship inspection by the chief scientist in the ship's home port including a trial run (~half day) at sea, will be paid by the vessel donor. [b]: A 5% contingency computed on non-salary costs, and to be used, if necessary, to cover unanticipated expenses. Numbers shown in bold font indicate section totals. The cost of project support from existing IATTC staff is not included in the budget.

	14-day trial Ocean Starr [In-kind research vessel]	**Main survey 2 ships: Ocean Starr + Ocean Titan [Ocean Starr + In-kind research vessel]**
	(US$ 1000)	(US$ 1000)
Cetacean abundance		
Chief Scientist	99.41	281.51
Scientific support to Chief Scientist	87.82	197.41
Cruise leaders	22.29	335.20
Marine Mammal Observers	98.10	955.53
Survey coordinator	52.00	143.93
	359.62	**1,913.57**
Vessels and associated costs		
Survey vessels	720.32[0 [a]]	7,248.40 [2,705.84 [a]]
Ship agent + Port fees	20.25	290.50
	740.57 [20.25[a]]	**7,538.90 [2,996.34 [a]]**
Foreign observers		
Travel	**2.00**	**22.00**
Drones		
Equipment rental	420.00	4,080.00
Observers	32.70	322.28
Fuel	0.26	3.96
	452.96	**4,406.24**
School size calibration		
Image analyses	58.01	97.64
Data cleaning and analyses	13.82	55.28
	71.83	**152.92**
Trackline detection probability		
Image analyses	38.34	106.59
Data cleaning and analyses	20.73	69.10
	59.07	**175.68**
Equipment		
Computers	17.38	17.38
Flying bridge	104.36	104.36
Communications	17.77	164.17
Research permits	3.00	15.00
Miscellaneous travel	6.00	15.00
IT specialist	10.34	4.63

	158.84	**320.54**
IATTC headquarters-based contractors		
Accountant	130.57	124.04
Logistical coordinator	112.01	106.41
	242.58	**230.45**
Ship loading	**3.00**	**6.00**
Contingency [b]	**67.35 [31.34 [a]]**	**614.45 [387.32 [a]]**
Total	**2,157.82 [1,401.49[a]]**	**15,380.75 [10,611.06 [a]]**

4. DISCUSSION

The series of dolphin surveys in the ETP conducted by the SWFSC is one of the most effort intensive worldwide, with a well-developed and consistent survey protocol. Hence, the question arises why not simply conduct another ETP survey following the same protocol and design as previous surveys? There are two main reasons that we considered here: recent evidence that trackline probability may be less than one for ETP dolphins (Barlow 2015) and the proposed use of tuna vessels (MSC 2015). Hence, we considered two distinct objectives for the next survey: 1) make it comparable with previous surveys and 2) estimate absolute abundances; and outline several options for accomplishing Objective 1 or both. Considering tuna vessels for the main survey implies that Objective 2 (estimating trackline detection probability) needs to be addressed for both research and tuna vessels to ensure that Objective 1 can be met and estimates from research and tuna vessels are comparable.

Estimating trackline detection probability using MRDS methods requires that two observation platforms are set up that monitor the search area either simultaneously or where one platform is slightly ahead of the other. It is critical, however, that, at minimum, one platform is independent from the other, i.e. that detections made from this platform were not influenced by the other. For this survey, the primary (and independent) observation platform is the team of three observers on the flying bridge scanning the forward 180 degrees, two observers using big-eye binoculars and one observer naked eye/hand-held binoculars. Our preferred method for the secondary platform is a drone that flies several km ahead of the ship. This would allow investigating at a larger range (e.g., compared to a tracker on the top mast of a survey vessel) at what distance dolphins started reacting to the ship, which is critically important for MRDS methods (Burt *et al*. 2014). Compared to a helicopter operated from the survey vessel or fixed-wing aircraft operated from shore, the drone has the potential to work longer hours in any given day, in higher sea states, with much reduced safety risks to personnel on board, and with less disturbance to the dolphins. Fixed-wing aircraft would also be limited to nearshore areas. Furthermore, while tracking a dolphin school, a drone is less likely to cue in flying bridge observers.

An alternative to drone or helicopter would be to install a second observation platform on board the survey vessel. This was attempted during the ETP survey in 1998 on board the UNOLS research vessel Endeavor (Kinzey *et al*. 1999). An additional set of big-eyes was set up as the tracker position on the upper mast platform with a binocular height of 17.9m above sea level providing a maximum sighting distance of 15.2km. In comparison, the big-eyes on the flying bridge of the Endeavor were at 10.4km providing a ship-to-horizon sighting distance of 11.5km. Here, the tracker observed the area from 45 left to 45 right, as far as possible in front of the ship, and tracked each group of dolphins not yet detected by the flying bridge team until it was either detected by them or had passed the beam of the ship. Trackers were aware of the flying bridge detections but not vice versa. During the four-month survey, all tracker sightings for offshore spotted and spinner dolphins were also detected by the flying bridge team. However, with respect to spotted and spinner dolphins, the tracker had only first detected ten spotted and nine spinner dolphin groups throughout the entire survey (Table 6 in Kinzey *et al*. 1999). This does not constitute enough data to estimate g(0). Furthermore, if dolphins started evading the ship at distances greater than 15.2km, results from an MRDS analysis based on data collected in this manner would likely be biased.

It follows that drones are the only viable option for estimating trackline detection probability. Furthermore, as estimating trackline detection probability is necessary for calibrating tuna vessels against research vessels, the use of tuna vessels for the next survey relies on the utility of the drones. Should it be decided to not use tuna vessels for the next survey we still propose to use drones for collecting data for school size calibration, and to allow estimation of g(0), so that the concerns raised by Barlow (2015) can be assessed more generally. We would need a drone with a similar range, as dolphin schools need to be captured by video before reacting to the ship. We may, however, be able to reduce the time we require a drone to be used onboard each vessel.

These issues, i.e. vessel calibration and testing the utility of a drone, strongly highlight the need of a trial survey. However, we further emphasize that a trial survey should be conducted even if only research vessels were used for the main survey and even if we only aim to complete Objective 1. In this situation, the trial survey will serve as a pilot study before the main survey and the utility of a drone needs to be assessed. Testing the drone from platforms of opportunity, instead of during a trial survey, is not an option as the schools that the drones are tested on should be the same or at least similar species in terms of school size, group cohesion and evasive reactions to the ships. Hence, it is not simply a test of whether the drone is capable of taking adequate imagery but also whether this operation is feasible for the species of interest. These operations require that the ship closes on the dolphin schools, which is not generally possible using platforms of opportunity.

We recommend that either Design 1, or Design 3 with two vessels, be adopted. Design 1 has the advantage that it allows estimation of absolute abundance of all the stocks of Table 1. In principle, it also allows assessment of whether there is movement of the stocks of offshore spotted dolphins across nominal stock boundaries, but in practice, such inference is hampered by poor precision. We might anticipate that, under Design 1, the coefficient of variation for the estimate of the north-eastern stock of offshore spotted dolphin will be of the order of 20%, while that of the western/southern stock might be around 30%. Given these levels of precision, only extremely large movements across the boundaries could be detected with any confidence. Inference may be further compromised by the inability to implement a randomized even-coverage design in the OUTER region, given the constraints imposed by the need to access a port at regular intervals, and by the limited effort relative to the size of the region achievable in a two-vessel survey.

Under Design 3, we might anticipate that, with two vessels operating in the CORE, CORE2 and NORTH COASTAL strata, we might achieve a CV of around 14%. This gives a better chance of detecting any change since 2006, but of course any change detected could potentially arise from either changes in population abundance or large-scale movement across the stock boundaries or both. Although in principle the eastern spinner dolphin distribution may extend slightly beyond the CORE + CORE2 + NORTH COASTAL strata (Dizon *et al*. 1994), previous surveys have reported almost no eastern spinner sightings in the OUTER strata (Gerrodette and Forcada 2005; Gerrodette *et al*. 2005, 2008), and therefore inference for eastern spinner dolphin may be more secure.

Poor precision under Design 1 will be less problematic if the proposed survey is the first of a series of planned surveys, perhaps conducted once every three or four years, so that ongoing information is gathered, to allow more reliable determination of population trends. In the absence of such an approach, only very large changes since 2006 will be likely to be detected. The optimal frequency for a series of planned surveys could be estimated from existing survey data, and the final decision could take into consideration cost-benefit trade-offs.

In contrast to previous surveys, in the proposed survey designs we have not considered collection of ancillary data, such as oceanographic data. The reason for this is that we believe that with only a one-off survey requested, it is of the utmost importance to address what would appear to be the potentially greatest challenges: imperfect detection on the trackline and differences among survey vessels (assuming tuna vessels are being considered for the survey). If a series of planned surveys were to be entertained, it may be worthwhile to incorporate collection of some types of ancillary data into the survey design. For example, incorporation of oceanographic data collection into the survey design could contribute to future efforts of spatiotemporal modelling of encounter rate or abundance.

In order to finalize the designs of the trial and main surveys, and to determine which objectives are potentially feasible, several pieces of additional information are necessary. First, we need to be provided with total amount of fund available for the survey. Second, we need to be informed of which ports of call will be made accessible to the survey vessels for both the trial and main surveys. Finally, if one or two tuna vessels are to be provided for the survey, we need to know exactly which vessels those are so that their characteristics can be evaluated.

5. CONCLUSIONS

1. If a decision is made not to use tuna vessels, the trial would only need to use one research vessel and one drone, as vessel calibration would not be necessary.
2. If the drones are not (or cannot be) used, then Objective 2 may need to be dropped.
3. If the main survey were to follow immediately after the trial survey, a less ambitious (and shorter) trial survey would have to be implemented, mostly limited to ensuring that the drones are performing as required. In this case, a decision not to use tuna vessels should be made, as the trial could not deliver adequate precision to calibrate a tuna vessel against a research vessel, if there is evidence of different biases. Moreover, the trial survey would have to take place outside of the historical survey season (end July – early December) in order that the main survey would take place within the historical survey season.
4. Limiting the drones to the trial survey would be unsatisfactory because, if probability of detection on the trackline can be well below one for schools of the priority species, it is likely that it will vary by location, as animals may respond differently to survey vessels in different parts of the ETP. Further, estimation of this probability from a limited trial will add substantial imprecision to an abundance estimate. Hence the drone(s) should operate for both the trial and the main surveys.
5. Survey Design 1 allows Objectives 1 and 2 to be addressed. However, precision of estimates for the OUTER strata would be poor, as has been the case in previous surveys. If abundance estimates are required just for the northeastern offshore spotted and eastern spinner dolphin stocks, then having both survey vessels only operating in the CORE + CORE2 + NORTH COASTAL strata should give good precision, although movement of animals across boundaries may be an issue. If only one research vessel is funded for the main survey, then Design 3 is preferred to try to maintain precision at the level of previous surveys for priority stocks B.
6. Given the difficulty of estimating stock sizes without adequate precision, and given the future need for stock abundance estimates, we recommend that a series of surveys, perhaps with three years between each survey, be planned, rather than one-off surveys conducted only as the need arises.

ACKNOWLEDGEMENTS

We extend our sincere appreciation to the many people who took the time to answer questions, provide information and suggestions, and/or comment on an early draft of this document: Ernesto Altamirano, Lisa Ballance, Jay Barlow, David Borchers, Alejandro Cab, Michel Dreyfus, Aaron Farber, Megan Ferguson, Tim Gerrodette, Mike Gomez, Annette Henry, Toshihide Kitakado, Jeff Laake, Erick Largacha, Manuel Lutt, Sarah Mesnick , Erin Oleson, Hector Perez, Wayne Perryman, Shannon Rankin, Enrique Ureña, Nickolas Vogel, Andy Webb, Brad Wiley, and Sheyna Wisdom.

REFERENCES

Au DWK & WL Perryman (1985) Dolphin habitats in the eastern tropical Pacific. Fisheries Bulletin U.S. 83(4): 623-643.

Barlow J (2015) Inferring trackline detection probabilities, g(0), for cetaceans from apparent densities in different survey conditions. Marine Mammal Sciences. 31(3):923-943.

Borchers D (2012) A non-technical overview of spatially explicit capture-recapture models. Journal of Ornithology.152, (2), 435-444.

Buckland ST, DR Anderson, KP Burnham & JL Laake (2001) Introduction to Distance Sampling Estimating Abundance of Biological Populations. Oxford University Press.

Buckland ST, DR Anderson, KP Burnham, JL Laake, DL Borchers & L Thomas (2004) Advanced Distance Sampling. Oxford University Press.

Buckland ST, E Rexstad, TA Marques & CS Oedekoven (2015) Distance Sampling: Methods and Applications. Methods in Statistical Ecology, Springer.

Buckland ST, CE Lennert-Cody, T Gerrodette, J Barlow, JE Moore, A Webb, PT Fretwell, HJ Skaug, WL Perryman (2016) Review of potential methodologies for estimating abundance of dolphin stocks in the Eastern Tropical Pacific. Workshop on Methods for Monitoring the Status of Eastern Tropical Pacific Ocean Dolphin Populations 18-20 October 2016, La Jolla, California.

Gerrodette T & PR Wade (1991) Monitoring trends in dolphin abundance in the eastern tropical Pacific: Analysis of 1989 data. Rep. Int. Whal. Commn. 41: 511-515.

Gerrodette T & J Forcada (2002) Estimates of abundance of northeastern offshore spotted, coastal spotted, and eastern spinner dolphins in the eastern tropical Pacific Ocean. Admin Rep LJ-02-06, Southwest Fisheries Science Center, La Jolla (available at: swfsc.ucsd.edu/IDCPA/TunaDol_rep).

Gerrodette T & J Forcada (2005) Non-recovery of two spotted and spinner dolphin populations in the eastern tropical Pacific Ocean. Marine Ecology Progress Series 291: 1-21.

Gerrodette T, G Watters, W Perryman & L Balance (2008) Estimates of 2006 dolphin abundance in the eastern tropical Pacific with revised estimates from 1986-2003. NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-422.

Gerrodette T, WL Perryman & CS Oedekoven (2018) Accuracy and precision of dolphin group size estimates. Marine Mammal Science, in print. https://doi.org/10.1111/mms.12506.

Hall, M.A. and Lennert, C. (1994) Incidental mortality of dolphins in the eastern Pacific Ocean tuna fishery in 1992. Rep. Int. Whal. Commn. 44: 349-351.

Hoyle SD & MN Maunder (2004) A Bayesian integrated population dynamics model to analyse data for protected species. Animal Biodiversity and Conservation 27: 247-266.

IATTC (2006) Technical workshop on calculating Nmin for the dolphin stocks of the eastern Pacific Ocean. Special Report 14 of the Inter-American Tropical Tuna Commission, La Jolla, California. http://www.iattc.org/PDFFiles/SpecialReports/_English/Special-Report-14ENG.pdf

IATTC (2015) Annual Report of the Inter-American Tropical Tuna Commission, 2010. http://www.iattc.org/PDFFiles/AnnualReports/_English/IATTC-Annual-Report-2010.pdf

Jackson A, T Gerrodette, S Chivers, M Lynn, S Rankin & S Mesnick (2008) Marine mammal data collected during a survey in the eastern tropical Pacific ocean aboard NOAA ships David Starr Jordan and McArthur II, July 28 – December 7, 2006. NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-421.

Jackson A, T Gerrodette, S Chivers, M Lynn, P Olson & S Rankin (2004) Marine mammal data collected during a survey in the eastern tropical Pacific ocean aboard NOAA ships McArthur II and David Starr Jordan, July 29 – December 10, 2003 NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-366.

Kinzey D, T Gerrodette, J Barlow, A Dizon, WL Perryman, P Olson, A Von Saunder (1999) Marine mammal data collected during a survey in the eastern tropical Pacific aboard the NOAA ships McArthur and David Start Jordan and the UNOLS ship Endeavor JULY 31 - DECEMBER 9, 1998. NOAA Technical Memorandum NOM-TM-N M FS-SWFSC-283.

Kinzey D, T Gerrodette, J Barlow, A Dizon, WL Perryman & P Olson (2000a) Marine mammal data collected during a survey in the eastern tropical Pacific Ocean aboard the NOAA ships McArthur and David Starr Jordan, July 28 – December 9, 1999 NOAA Technical Memorandum NOM-TM-N MFS-SWFSC-293.

Kinzey D, P Olson & T Gerrodette (2000b) Marine mammal data collection procedures on research ship line-transect surveys by the Southwest Fisheries Science Center. Southwest Fisheries Science Center, Administrative Report LJ-00-08, 32 p.

Kinzey D, T Gerrodette, A Dizon, WL Perryman, P Olson & S Rankin (2001) Marine mammal data collected during a survey in the eastern tropical Pacific Ocean aboard the NOAA ships McArthur and David Starr Jordan, July 28 - December 9, 2000. NOAA Technical Memorandum NOM-TM-NMFS-SWFSC-303.

Lennert-Cody CE, ST Buckland & FC Marques (2001) Trends in dolphin abundance estimated from fisheries data: A cautionary note. Journal of Cetacean Research and Management 3:305–319.

Lennert-Cody CE & MD Scott (2005) Spotted dolphin evasive response in relation to fishing effort. Marine Mammal Science 21: 13–28.

Lennert-Cody CE, MN Maunder, PC Fiedler, M Minami, T Gerrodette, J Rusin, CV Minte-Vera, M Scott, & ST Buckland (2016) Purse-seine vessels as platforms for monitoring the population status of dolphin species in the eastern tropical Pacific Ocean: the use of fishing vessels as scientific platforms. Fisheries Research 178: 101-113.

Leslie MS & PA Morin (2016) Using Genome-Wide SNPs to Detect Structure in High-Diversity and Low-Divergence Populations of Severely Impacted Eastern Tropical Pacific Spinner (Stenella longirostris) and Pantropical Spotted Dolphins (S. attenuata). Frontiers in Marine Science, 3: 253.

MSC (2015) MSC Full-Assessment Public Certification Report: The northeastern tropical Pacific purse seine yellowfin and skipjack tuna fishery. Available at: https://fisheries.msc.org/en/fisheries/northeastern-tropical-pacific-purse-seine-yellowfin-and-skipjack-tuna-fishery/@@assessment-documentsets?documentset_name=Public+certification+report&phase_name=Public+certification+report+and+certificate+issue&start_date=2014-10-31&title=Initial+Assessment

NRC (1992) Dolphins and the Tuna Industry. National Academy Press, Washington, D.C., pp. 176.

Perrin WF, MD Scott, GJ Walker, FM Ralston, & DWK Au (1983) Distribution of four dolphins (Stenella spp. and Delphinus delphis) in the eastern tropical Pacific, with an annotated catalog of data sources. NOAA techn. Memo NOAA-TM-NMFS-SWFC-38, p. 1-65.

Perrin, W.F., Akin, P.A., Kashiwada, J.V. (1991) Geographic variation in external morphology of the spinner dolphin *Stenella longirostris* in the eastern Pacific and implications for conservations. Fishery Bulletin 89: 411-428.

Pryor K & KS Norris (1978) The tuna-porpoise problem: Behavioral aspects. Ocean 21, 31–37.

Reilly SB (1990) Seasonal changes in distribution and habitat differences among dolphins in the eastern tropical Pacific. Marine Ecology Progress Series 66:1-11.

Reilly SB & PC Fiedler (1994) Interannual variability of dolphin habitats in the eastern tropical Pacific. I: Research vessel surveys, 1986-1990. Fishery Bulletin 92: 434-450.

Reilly SB, MA Donahue, T Gerrodette, K Forney, P Wade, L Balance, J Forcada, P Fiedler A Dizon, W Perryman, FA Archer & LF Edwards (2002) Report of the scientific research program under the international dolphin conservation program act. NOAA techn. Memo NOAA-TM-NMFS-SWFC-372, p. 1-100.

Scott MD, SJ Chivers, RJ Olson, PC Fiedler & K Holland (2012) Pelagic predator associations: tuna and dolphins in the eastern tropical Pacific Ocean. Marine Ecology Progress Series. 458: 283–302.

Smith TD (1983) Changes in size of three dolphin (Stenella spp.) populations in the eastern tropical Pacific. Fishery Bulletin 81: 1-13.

Wade PR & T Gerrodette (1993) Estimates of cetacean abundance and distribution in the eastern tropical Pacific. Rept. Int. Whal. Commn. 43:477-493.

Wade PR, GM Watters, T Gerrodette & SB Reilly (2007) Depletion of spotted and spinner dolphins in the eastern tropical Pacific: modelling hypotheses for their lack of recovery. Marine Ecology Progress Series 343: 1-14.

Wahlen B. (1986) Incidental dolphin mortality in the eastern tropical Pacific tuna fishery, 1973 through 1978. Fishery Bulletin 84: 559-569.

Appendix 1: Length of trial period

1. Introduction

Here we estimate the length of the trial period should the purpose of the trials include vessel calibration in addition to testing the use of drones for collecting data for *g(0)* estimation and school size calibration. The vessel calibration is the reason for involving two ships, i.e. one research and one tuna vessel (as opposed to only a research vessel), in the trials. It is also the reason for extending the length of the trial period beyond two weeks.

The aim of the vessel calibration is to determine whether data collected from tuna vessels is comparable to data collected from research vessels. Possible reasons for differences include evasive behavior of the dolphins towards tuna vessels, which may still occur when such a vessel is used as an observation platform. Hence, we need to determine whether the different components required for estimating density, i.e. the effective strip width, expected school size, encounter rate and trackline detection probability, differ between vessels. In order to determine, for example, a 20% decrease in encounter rate from the tuna vessel that is significant at the 5% level we need a large enough sample size to allow us to estimate this difference with sufficient precision. What 'sufficient' corresponds to is described quantitatively in the section Significant differences in point estimates below. Here we assess the size of the sample required to detect a significant difference. We use point estimates, coefficients of variation (CV) and sample sizes from existing literature for our calculations. We present the sample sizes we would need to detect a significant increase or decrease of 5, 10, 15, ..., 100% between vessels for each of the four components required for density estimation.

2. Methods

2.1. Significant differences in point estimates

To test for a significant difference between two point estimates $\widehat{PE}_r$ and $\widehat{PE}_t$ (where subscripts *r* and *t* refers to a research and tuna vessel, respectively) at the α = 5% error level we estimate the difference using:

$$\widehat{Diff}_{t,r} = \widehat{PE}_t - \widehat{PE}_r$$

Assuming independence, the standard error of the difference is estimated using:

$$\widehat{se}\left(\widehat{Diff}_{t,r}\right) = \sqrt{\widehat{var}\left(\widehat{PE}_r\right) + \widehat{var}\left(\widehat{PE}_t\right)}$$

We then use the estimated difference and standard error to build 95% confidence intervals:

$$\widehat{Diff}_{t,r} + z_{\alpha/2} \times \widehat{se}\left(\widehat{Diff}_{t,r}\right)$$

Should this confidence interval not contain zero, we reject the null hypothesis of no-difference between the two surveys. In practice, we can design the trial to ensure a positive correlation between estimates, in which case precision will be better than estimated here. However, we expect any correlation to be relatively weak, and so our calculations should give a good guide to precision.

2.1.1. Minimum change for detecting a significant difference in point estimates of the parameters

For the following we use the point estimates for the parameters *f(0)*, encounter rate 100*n/L (100 * number of detected schools / total line length) and expected school size *E(s)* from the STAR06 survey presented by Gerrodette *et al*. (2008) as the point estimates from the research vessel $\widehat{PE}_r$ and estimate the minimum change required to detect a difference in the point estimates for the tuna vessel $\widehat{PE}_t$. As we have no information on precision of these estimates from tuna vessels, we assume the same CVs for the research and tuna vessel.

As we expect the size of the CV to change for, e.g. a short trial period compared to the STAR06 survey, we need to consider the following. In general, the standard error $se(\widehat{PE})$ of the point estimate of a sample mean is defined as $se(\widehat{PE}) = sd/\sqrt{n}$ where sd is the sample standard deviation and n is the sample size (Davison 2003). It follows that if the sample size doubles, the standard error decreases by a factor of $1/\sqrt{2}$ or if the sample size halves the standard error increases by a factor of $\sqrt{2}$. Hence, we can only use the same standard errors as presented in Gerrodette *et al*. (2008) for calculations where the sample size is the same or, if not, adjust the standard error accordingly. For parameters pertaining to the detection function *f(0)* and to the expected school size *E(s)*, the number of samples is the number of detected schools used to estimate the parameters. For the encounter rate *100*n/L* the expected variance is proportional to the inverse of total line length *1/L* (Fewster *et al*. 2005) and, similarly, if the total line length halves, we expect the standard error to increase by a factor of $\sqrt{2}$. The same applies to the coefficient of variation (CV).

2.1.2. Minimum increase

To obtain an estimate of the minimum increase required to detect a significant change between the point estimates $\widehat{PE}_r$ and $\widehat{PE}_t$ using the test described above, we need to find an estimate $\widehat{PE}_t$ for which the lower boundary of the 95%CI of the difference $\widehat{Diff}_{t,r} = \widehat{PE}_t - \widehat{PE}_r$ equals zero:

$$\widehat{PE}_t - \widehat{PE}_r - z_{\alpha/2} \times \widehat{se}\big(\widehat{Diff}_{t,r}\big) = 0\ .$$

The minimum percent increase in $\widehat{PE}_t$ compared to $\widehat{PE}_r$ its then calculated as $(\widehat{PE}_t\ /\ \widehat{PE}_r - 1) \times 100.$

2.1.3. Minimum decrease

Equivalently, when estimating the minimum decrease we need to find an estimate $\widehat{PE}_t$ such that the upper boundary of the difference is zero:

$$\widehat{PE}_t - \widehat{PE}_r + z_{\alpha/2} \times \widehat{se}\big(\widehat{Diff}_{t,r}\big) = 0\ .$$

The minimum percent decrease in $\widehat{PE}_t$ compared to $\widehat{PE}_r$is then calculated as $(\widehat{PE}_t\ /\ \widehat{PE}_r - 1) \times 100.$

2.2. Point estimates, CVs and sample sizes used from existing literature

All estimates, CVs and sample sizes that we used for our calculations are listed in the Tables below. When possible, we drew these from the STAR06 reports (Table A1, Jackson *et al*. 2008; Gerrodette *et al*. 2008). However, these reports do not include estimates of trackline probability *g(0)* and none are available for ETP dolphins from other sources that were estimated using mark-recapture distance sampling methods (MRDS, e.g. Borchers 2012). Hence, we used the estimates of *g(0)* from Barlow (2015, A2) for spotted and spinner dolphins which are relative estimates of *g(0)* for Beaufort 1-6 compared to *g(0)* at Beaufort 0 (which he assumed to be one). In addition, we needed CVs of *g(0)* that were estimated using MRDS methods which were not available from Barlow (2015). Therefore , as a precision estimate for both spotted and spinner dolphins, we used the CV for estimating *g(0)* from the SCANS III survey (Hammond *et al*. 2017) that was estimated using MRDS. We used the CV for striped and common dolphins combined as these are the closest species to spotted and spinner dolphins from that report. This relies on the assumption that the CV for spotted or spinner dolphins would be equivalent. Furthermore, we make the assumption here that the same number of detections made by the primary observers will yield the same CV although in reality this depends at minimum also on the number of detections made by the tracker and how many of these are duplicates (although our tracker will be the drone instead of a ship-based tracker).

For the number of detections by Beaufort sea state from STAR06 we scaled the total number of detections by the proportion of effort conducted during the respective sea state (Table A5).

TABLE A1. Estimates, coefficients of variation (CV, given as percentage) and sample sizes (number of dolphin schools) or total line length from Gerrodette *et al*. (2008). Sample sizes were larger for estimating detection functions compared to school sizes as data were pooled across species for estimating *f(0)*. For encounter rate, only the effort from the CORE stratum was included for NE offshore spotted dolphins while for eastern spinner dolphins the effort from CORE and CORE2 strata were included.

	Detection function		
Species	*f(0)*	CV	Sample size
NE offshore spotted	0.255	6.2	151
Eastern spinner	0.255	7.3	84
	School size		
	E(s)	CV	Sample size
NE offshore spotted	118.2	17.3	102
Eastern spinner	196.3	14.9	68
	Encounter rate		
	*100*n/L*	CV	Line length (km)
NE offshore spotted	0.861	14.9	10268
Eastern spinner	0.305	18.1	11036

TABLE A2. Estimated values of *g(0)* for sightings conditions in Beaufort states 1–5 relative to Beaufort 0 (Barlow 2015). All *g(0)* values shown here were significantly different from 1.0 determined by z-tests.

	Beaufort					
	b0	b1	b2	b3	b4	b5
Spotted	1	0.733	0.537	0.394	0.289	0.212
Spinner	1	0.728	0.531	0.386	0.282	0.205

TABLE A3. Number of common and striped dolphin sightings made by the primary observer team, the tracker and duplicates between the two platforms of from the SCANS III ship survey detected during Beaufort 0-4 (from Table 4 in Hammond *et al*. 2017). Duplicates shown are Definite and Probable duplicates, as used in analysis presented in Hammond *et al*. (2017). The bold number is the sample size we used for our calculations.

Species	Total sightings	Primary sightings	Tracker sightings	Duplicates
Common	106	82	52	28
Striped	104	56	69	21
Total	**210**	**138**	**121**	**49**

TABLE A4. Estimated detection probabilities within the truncation distance of 2000m for common and striped dolphins combined during the ship survey of SCANS III (from Table 18 in Hammond *et al*. 2017). Coefficients of variation (CV) are given as percentages. The bold number is the CV used for our calculations.

	Average probability of detection assuming *g(0)*=1		Probability of detection on the transect line, *g(0)*		Overall average probability of detection	
Species	**Estimate**	**CV**	**Estimate**	**CV**	**Estimate**	**CV**
Common + striped dolphin	**0.131**	**11.7**	**0.421**	**11.5**	**0.055**	**16.4**

TABLE A5. Effort and sighting rates (all species) during STAR06, shown by sea state (Jackson *et al*. 2008).

Beaufort	Effort	Sightings	Sightings/1000km
0	100.1	33	329.55
1	375.4	78	207.8
2	1729.8	221	127.76
3	3212.2	261	81.25
4	9375.5	336	35.84
5	6952.1	199	28.62
6	492.1	7	14.23
7	0	0	0
Total	**22237.3**	**1135**	**51.04**

3. Converting the required sample sizes into required survey days

To convert required sample sizes into required survey days we need to know how many detections of spotted and spinner dolphins we can expect per 100km of effort in the area for the trial survey and how many km of effort we can expect to complete in closing mode in a day. There are no published records of encounter rates for spotted or spinner dolphins separately for the particular area of our trial survey. Hence, we combined the information from multiple studies. We used estimated encounter rates from previous ETP surveys presented in Gerrodette *et al*. (2008), Reilly (1990) and Schwarz *et al*. (2010). Reilly's estimated encounter rates for spotted and spinner dolphins combined in the area ranged between approximately 4 – 8 detections per 185km of survey effort. Here we use an average of 6 detections per 185km (or 3.24 detections per 100km). Encounter rates for NE offshore spotted and eastern spinner dolphins for STAR06 within were 0.861 and 0.305 per 100km, respectively (this includes the CORE stratum for spotted and CORE+CORE2 for eastern spinner dolphins). Assuming that the ratio spotted vs spinner detections would be the same in the trial area, i.e. 0.861 spotted / 0.305 spinner or 2.82 spotted detections for each spinner, and that we would detect, on average, six spotted and spinner schools combined per 100km, we expect to detect 4.43 spotted and 1.57 spinner schools per 100km.

During a study conducted in the same general area as we propose for the trial survey (Schwarz *et al*. 2010), 186 delphinid schools (including spotted, spinner and other dolphin species) were detected in closing mode during 34 survey days or 3,832.5km of survey effort (note this takes into account off-effort periods due to bad weather). This implies that the daily average of completed effort in closing mode was 112.72km. Based on this we assume that in closing mode we can expect to complete 112.72km of effort per day and that we can expect to detect 4.99 spotted dolphin schools and 1.77 eastern spinner dolphin schools per day.

In the following, we present the estimated sample sizes or line lengths as well as survey days required at minimum to detect an increase or decrease in the effective strip width *ESW*, school size and encounter rate between the two vessels. We use the *ESW* instead of *f(0)* as we find it easier to conceptualize a percent change for this parameter compared to *f(0)*. For a fixed truncation distance (here 5.5km) the conversion is done by $\widehat{ESW} = 1/\hat{f}(0)$. The CV for $\widehat{ESW}$ is the same as for $\hat{f}(0)$.

4. Results

Figure A1.1 depicts the estimated sample sizes or line lengths required to detect differences in the parameter estimates for the effective strip half-width, expected school size, encounter rate and trackline detection probability between a research and tuna vessel. The samples sizes and line lengths resulting from STAR06 are shown in the same Figure. The main conclusions are that it will be nearly impossible to detect significant increases or decreases <15% for any of the parameters even if we had the same sample sizes in detections and line length as STAR06. Differences between the ships will be hardest to detect for encounter rate. This is due to the fact that for encounter rate we expect the largest CVs.

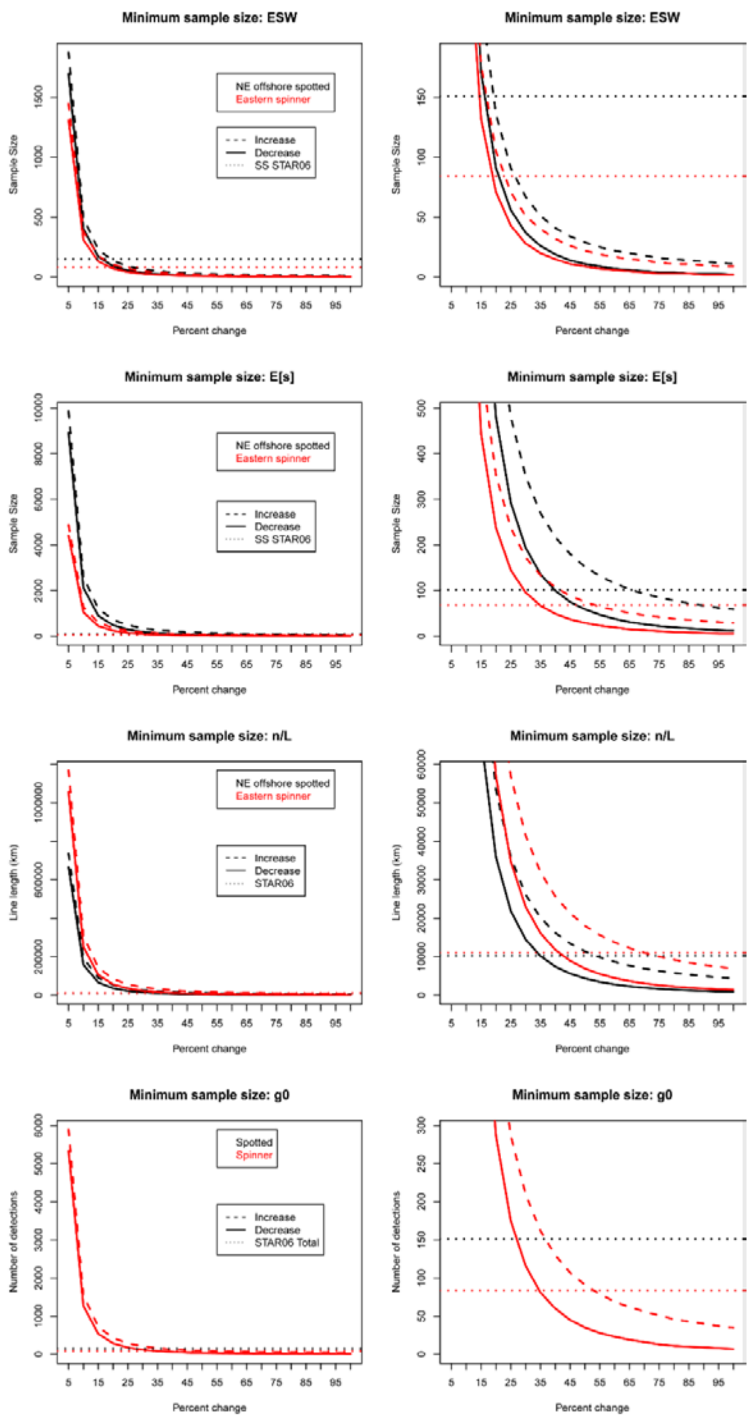

FIGURE A1.1. Expected minimum sample size (number of detected schools or line length) required to detect a significant percent change in the effective strip width (ESW), expected school size *E[s]*, encounter rate *n/L* and trackline detection probability *g(0)* between a research and tuna vessel (right plots are zoomed in to reveal more detail). For both vessels we assumed the same CV. See Tables A2-A5 point estimates, sample sizes and CVs used for the calculations. Note that for *g(0)* we used the same CV for both spotted and spinner resulting in the same sample sizes for both species (hence, only red lines show in the above plots).

5. Required survey days

Here we assess what we expect a trial period of 30-days would allow us to determine. We expect that a 30-day trial survey would allow us to detect differences in the ESW >20% for spotted dolphins and >30% for spinner dolphins and that it would allow us to detect differences in school sizes of >50% for spotted and >60% for spinner dolphins (Figure A1.2).

We expect that evasive behavior would result in negative bias in encounter rate and *g(0)* and, hence, we only consider decrease for these two parameters. We expect that a 30-day trial period would allow us to detect a decrease in encounter rate of >55% for spotted dolphins and >70% for spinner dolphins. We also expect that a 30-day trial period would allow us to detect a decrease in trackline detection probability of >30% for spotted dolphins and >45% for spinner dolphins.

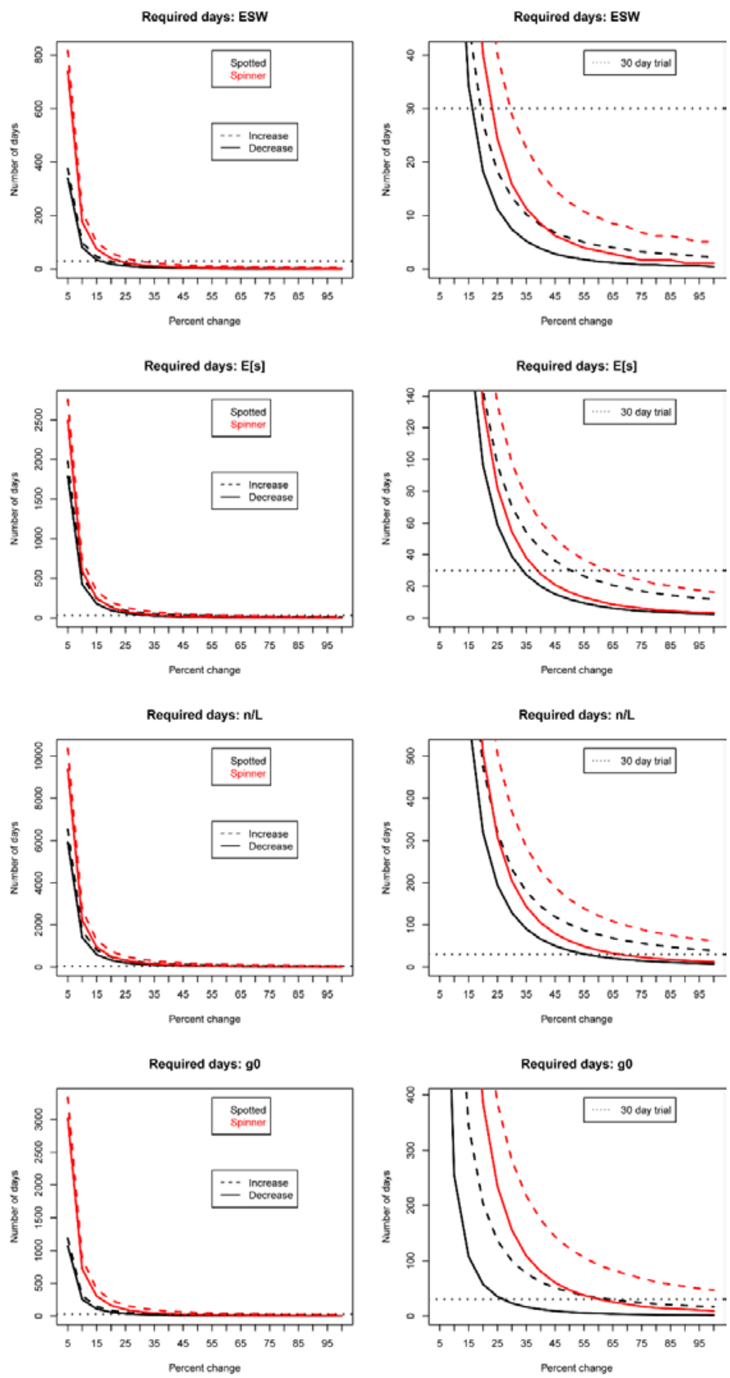

FIGURE A1.2. Estimated survey days required to detect a significant percent change in the effective strip width (ESW), expected school size *E[s]*, encounter rate *n/L* and trackline detection probability *g(0)* between a research and tuna vessel (right plots are zoomed in to reveal more detail). For both vessels we assumed the same CV. See Tables A2-A5 point estimates, sample sizes and CVs used for the calculations.

Appendix 2: Initial assessment of feasibility for drone operations via questionnaire

REPLIES FROM MIKE GOMEZ
Precision Aviation
Chehalem Airpark
17770 NE Aviation Way
Newberg, OR 97132

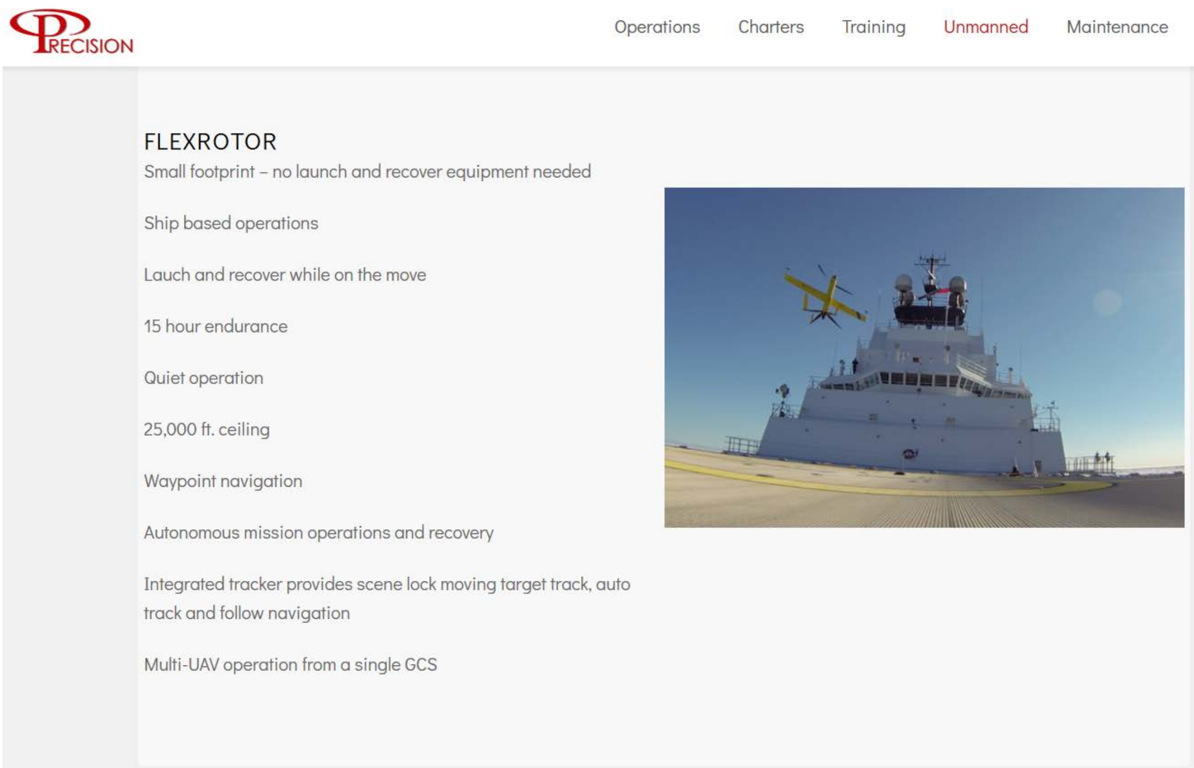

Figure A2.1 Screenshot from Precision Aviation website showcasing their Flexrotor. Source: https://www.flyprecision.com/unmanned/#1492191574538-466f6771-005d.

How high above the water would the drone fly keeping in mind that it should not affect dolphin behavior?
3 - 5000 msl
How wide of a strip does the footage provide given the height of the drone?
Depends on zoom level. Detection based on tuna searches is very easy with follow up zoom to get details. Suggest less than .5 mile strip.
Does the drone record its flight path, i.e. its gps locations at regular intervals?
Yes, 1 pulse per second.
Can the drone operate in Beaufort sea states 0 – 5 following the above protocol? [you already said it can operate up to Beaufort 2-3 only, correct?]
Yes, must be able to land on a fairly calm deck so depending on ship recommendation is 2-3 only pending survey.
Can the drone be launched / land from the ship while it is cruising at 10 knots (alternatively at 8 knots)?
Yes, depends on deck movement vertically
Does it need a helipad?

No, but unobstructed deck space is a must

Would the drone pilots be monitoring the real time footage for detecting dolphins? [I believe you mentioned that you guys do this?]

Yes

How far can the footage be transmitted back to the ship in real-time for monitoring?

Up to 100km

REPLIES FROM DR AARON FARBER

UAV Science Program Manager
Latitude Engineering
744 S. Euclid Ave., Tucson, AZ 85719

Figure A2.2. HQ-55 operated by Latitude Engineering. Source: https://latitudeengineering.com/2017/12/hq-55-first-flight/.

How high above the water would the drone fly keeping in mind that it should not affect dolphin behavior?

We don't know that yet. However, based on our historical acoustic testing, I imagine we are looking at somewhere between 1000 ft and 3000 ft. Obviously, this may depend on wind and sea conditions, and the biggest unknown, at what volume and frequency will the dolphins respond to vs be unaffected. This may be something to investigate during initial operations or pre deployment testing.

How wide of a strip does the footage provide given the height of the drone?

Again, this will depend on how much resolution we need to identify species. If you can provide some details on what you need to see to identify species, we can work backwards. However, Since we don't need IR (right?), I am imagining I would recommend using the Trillium HD50 system, which has a 720p, global shutter camera with a zoomable lens and field of view of 63-2.2 degrees. At 2000 ft altitude, that would correspond to a swath of 747 meters at the wide end and 23 meters at the narrow end. I would also open up the discussion with them to look at upgrading it to a 1080p camera, if that isn't already on their development path.

Does the drone record its flight path, i.e. its GPS locations at regular intervals?

Yes, the telemetry file is recorded on the ground station at an interval set by the pilot. This can be as high as 25 hz, but is often set to lower. If we are recording video on board for post processing, it is also possible to record the telemetry on the aircraft as well, which will ensure the telemetry is recorded at the full 25 hz and with no potential gaps due to momentary comms dropouts due to antenna masking or interference. If you don’t need that kind of frequent position, then the ground station recording is more than sufficient.

Can the drone operate in Beaufort sea states 0 – 5 following the above protocol?

We have not tested it on a ship above Beaufort sea states of approximately 3 or 4. However, that was done before we had the autonomous launch and recovery capability, which will significantly enhance the aircraft’s ability to handle higher sea states. Based on what we have seen and what we expect with the performance, I am relatively confident that we can operate in sea state 5 conditions. We will be able to answer this much more confidently after ship testing next spring.

Can the drone be launched / land from the ship while it is cruising at 10 knots (alternatively at 8 knots)?

In simple terms, yes, that should be fine. However, it’s a bit more complicated on a ship. It is less a matter of the speed of the vessel than the speed of the prevailing winds and the relative motion of the ship vs the aircraft. For land based operations, we have a max wind limit of 30 kts for launch and recovery. However, that is required to ensure we can position hold to the ground during VTOL mode. If the ship can be heading in any direction while underway during launch and recovery, then 10 kts can be handled almost certainly.

Does it need a helipad?

It doesn’t explicitly need a helipad, but that certainly reduces complexity and increases safety of the operation. Without one, we might have to build platforms to provide more clearance from railings and other hardware that cannot be removed.

Would the drone pilots be monitoring the real time footage for detecting dolphins?

This can be done by the pilots or by a separate crew member. The software and controls are pretty straightforward and easy to get up to speed on, if it makes sense to have someone with more knowledge of the dolphins at the controls of the camera.

How far can the footage be transmitted back to the ship in real-time for monitoring?

I need to confirm numbers with the radio guys, but I would say that if you want to go out more than 10 miles, we should consider a tracking antenna solution. If we go with that, you can expect 50+ miles of range, but the exact range will be dependent on the altitude the aircraft is operating at.