2023-05-29

Trade-offs in telemetry tag programming for deep-diving cetaceans: data longevity, resolution, and continuity

Cioffi, WR

https://pearl.plymouth.ac.uk/handle/10026.1/21450

10.1186/s40317-023-00334-1
Animal Biotelemetry
Springer Science and Business Media LLC

Cioffi *et al. Animal Biotelemetry* (2023) 11:23
https://doi.org/10.1186/s40317-023-00334-1

Animal Biotelemetry

METHODOLOGY

Open Access

Trade-offs in telemetry tag programming for deep-diving cetaceans: data longevity, resolution, and continuity

William R. Cioffi[1,2*], Nicola J. Quick[1,3], Zachary T. Swaim[1], Heather J. Foley[1], Danielle M. Waples[1], Daniel L. Webster[4,5], Robin W. Baird[5], Brandon L. Southall[1,2], Douglas P. Nowacek[1,6] and Andrew J. Read[1]

Abstract

Background Animal-borne telemetry instruments (tags) have greatly advanced our understanding of species that are challenging to observe. Recently, non-recoverable instruments attached to cetaceans have increased in use, but these devices have limitations in data transmission bandwidth. We analyze trade-offs in the longevity, resolution, and continuity of data records from non-recoverable satellite-linked tags on deep-diving *Ziphius cavirostris* in the context of a behavioral response study of acute noise exposure. We present one data collection programming scheme that balances resolution and continuity against longevity to address specific questions about the behavioral responses of animals to noise exposure in experimental contexts. We compare outputs between two programming regimes on a commercially available satellite-linked tag: (1) dive behavior summary defined by conductivity thresholds and (2) depth time-series at various temporal resolutions.

Results We found that time-series data vary from the more precisely defined dives from a dive summary record data stream by an acceptable error range for our application. We determined a 5-min time-series data stream collected for 14 days balanced resolution with longevity, achieving complete or nearly complete diving records in 6 out of 8 deployments. We increased our data message reception rate several fold by employing a boat based data capture system. Finally, a tag deployed in a group concurrently with a high-resolution depth recorder showed high depth concordance.

Conclusions We present the conceptual framework and iterative process for matching telemetry tag programming to research questions that we used and which should be applicable to a wide range of studies. Although designing new hardware for our specific questions was not feasible at the time, we were able to optimize the sampling regime of a commercially available instrument to meet the needs of our research questions and proposed analyses. Nevertheless, for other study species or designs, the complicated intersection between animal behavior and bandwidth of telemetry systems can often create a severe mismatch among research questions, data collection, and analysis tools. More flexible programming and purpose-built instruments will increase the efficacy of these studies and increase the scientific yield relative to the inherently higher risk of invasive studies.

Keywords Argos, Argos Goniometer, Satellite tag, Telemetry, Bio-logging, *Ziphius cavirostris*, Cuvier's beaked whales, Goose-beaked whales, Ziphiidae, Controlled exposure experiment

*Correspondence:
William R. Cioffi
wrc14@duke.edu
Full list of author information is available at the end of the article

Background

Satellite-linked bio-loggers (satellite tags) are an important component of the marine telemetry toolbox in part because they can uplink data to a satellite during deployments, and in some cases in near real time, rather than waiting for recovery of archived data. For tags with an archival capability, limited resources, uncertain detachment times, strong ocean currents, premature equipment failure, and battery life on radio beacons can all act to decrease the probability of device recovery. In fact, some tags deployed ballistically on cetaceans, such as those used in this study, lack any flotation to keep size and mass low, and therefore there is no expectation of recovery of archives and only satellite transmitted data are retained. Archived data, however, tends to allow higher resolution than satellite transmitted data because of the bandwidth limitations of many current marine satellite-linked tag systems, especially those that employ Argos [1, 2]. There are at least two important types of bandwidth limitation we consider here. First, the maximum bit rate that satellite-linked tag systems can support is constrained by hardware on the tag (transmitting device) and on satellites (receivers) and further the total throughput of the system is limited by orbiting characteristics, geographic coverage, and number of satellite receivers. Second, the behavior of the animal and device placement can be limiting factors [3]. In cetaceans, these limitations are often described by the frequency and duration of tag emergence into air, because successful uplinks occur only when the tag is out of water and a satellite is in radio range. Additionally, since some two-way communication satellite technologies are relatively slow compared to a surfacing event for a cetacean, many popular cetacean tags only have the ability for one-way communication meaning duplicate messages are often sent multiple times to ensure uplink since successful receipt is not determinable by the tag. This limitation also therefore decreases throughput.

Due to these constraints as well as other technical considerations in the manufacture of these devices, a variety of methods have been employed to maximize the amount of data obtained from non-recoverable satellite-linked tags. Data can be compressed or summarized, duty cycled, and sampling rate and/or resolution can be lowered to better accommodate bandwidth restrictions [4, 5]. Instruments can be designed to archive data, then release, and float to the surface where they transmit at a higher rate free floating and not constrained by the animals diving behavior [6]. Finally, stations with antennas and receivers affixed nearby to boats or on land can receive data directly from the instruments, ameliorating the limitations of poor satellite coverage by being available to collect data at times no satellites are available (for example, Argos Goniometer, Woods Hole Group, Bourne, Massachusetts, USA and Mote, Wildlife Computers, Redmond, Washington, USA, [7]). Specific solutions depend on the behavior of the species of interest, the logistics of deploying receiving stations, as well as the data needed to address research questions [8].

We have been using several types of Argos-linked satellite telemetry tags to study the movements and behavior of *Ziphius cavirostris* (family: Ziphiidae) off Cape Hatteras, North Carolina, USA [9, 10]. Here we focus on SPLASH10 tags (Wildlife Computers, Redmond, Washington, USA), which consist of a package of sensors including pressure, temperature, and conductivity, and an onboard computer and storage system that records, processes, and archives data. These tags are commonly attached to cetaceans in configurations that prevent recovery of the tag and the full archived record. In such cases, returned sensor data are entirely in the form of programmable data streams that are uplinked to the Argos satellite system, where they can be downloaded and decoded. Additionally, geographic position can be estimated from uplinks to the Argos system using a Doppler calculation. A very commonly utilized data stream consists of dive summary records (termed behavior log in the tag programming software and data portal) for any dives which meet a predetermined threshold based on pressure, duration, and conductivity (user-definable within a certain range). Other gross metrics are available in data streams, for example depth histograms over a given time span (for example, daily). Finally, a true time-series of depths or temperatures can be recorded at one of five supported sampling periods and dynamically calculated data resolution. Multiple data streams can be collected and transmitted concurrently with some limitations.

In our applications on *Z. cavirostris*, the surfacing behavior of the animals themselves creates a tremendous bandwidth bottleneck. The Argos system in use by these tags is limited to a maximum 32 bytes per data message (including any metadata) and, although several messages can be sent per minute, during typical ventilation behavior generally only one message is sent each time the tag breaks the surface. While more messages can be sent if an animal is floating at the surface with the tag exposed, this is an uncommon behavior for this species in our location. In addition, given the polar orbit of Argos satellites, at the latitude of our study site off the coast of Cape Hatteras, North Carolina (approx. 35–36° N), there is only about 9% temporal coverage. *Z. cavirostris* exhibit extremely long foraging dives (median: 59 min), shorter non-foraging dives (median: 19 min, and very short periods of ventilation during which they break the surface. The median duration of each ventilation period is only

2.2 min [10], during which time the animal will break the water's surface multiple times. For these reasons, in our area on *Z. cavirostris*, we receive an average of only about 20–30 raw uplinks per day on Argos. Disregarding corrupted data and status messages, often less than 10 messages (including duplicates) are logged per day via satellite. Total number of transmits typically range in the hundreds per day and tend to stay below our user-defined daily limit of 450 or 470. These limits are imposed to avoid excessive battery usage if the animal's behavior happens to present a tag at the surface for an extended period of time.

Recently, we have successfully instrumented several dozen *Z. cavirostris* as part of the Atlantic behavioral response study, an experiment to quantify behavioral responses to naval sonar signals. The aim of this behavioral response study is to collect dive data before, during, and after known exposures to mid-frequency (3–4 kHz) active sonar signals using controlled exposure experiments either from operational Navy vessel-based sources or a simulated source [11–13]. These exposures are acute, up to 1 h in duration, in select discrete time periods. We identified three key axes which represent trade-offs in satellite tag configuration when collecting dive data: *continuity* or the completeness or the number of gaps in the data record; data *resolution* as determined by the temporal and spatial sampling scheme; and *longevity* or the overall data record length. We note that each of these axes can have multiple dimensions, but suggest this paradigm is useful for appreciating that the relative importance of these three general trade-offs depends on the research question(s) being addressed, and an equal maximization function may not always be desirable. For example, prior to the start of the behavioral response study off Cape Hatteras, we programmed satellite tags to prioritize longevity and data resolution for the purpose of exploratory data collection (Fig. 1) [10]. Specifically, we collected multiple data streams at relatively fine sampling rates, and duty-cycled data collection to increase battery life and overall transmission length, which necessarily introduced data gaps (decreasing data continuity). In 2017, when the experimental phase of the study began, we wanted to determine whether mid-frequency active sonar disrupted deep foraging dives in known discrete exposure conditions, so we chose settings that prioritized longevity and continuity by collecting data only on foraging dives, at the cost of resolution. Thus, we employed a dive summary record only configuration, in which shallow dives were not recorded [5]. Later in 2018, we implemented a new programming scheme to address questions concerning potential behavioral responses across dive types and within dives using the time-series data stream. In this scheme, we sacrificed overall longevity to produce a continuous record and a true time-series, centered around a known disturbance event of interest, over which response could be assessed. In this paper, we describe specific outcomes of implementing this new programming scheme to increase temporal resolution of depth data collection and data continuity at the cost of longevity, given bandwidth limitations of Argos services, our field location, and *Z. cavirostris* behavior.

While, this new programming scheme used the existing onboard functionality of the tags, we were not aware of previous deployments in this specific configuration and so we deployed a provisional set of assessment tags that employed only time-series data to determine feasibility of this approach. In this paper, we report our programming optimization process within the context of a larger experimental study on behavior responses. We also discuss the efficacy of increasing bandwidth, a key limitation of our instruments, by employing an Argos Goniometer, a vessel-based UHF antenna and receiver system that intercepts radio transmissions from the tags. While many of our specific requirements and decisions were

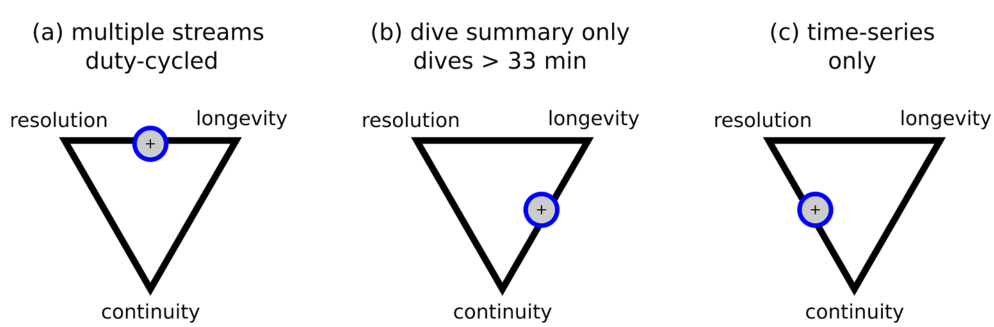

Fig. 1 A conceptual framework of the trade-offs in bandwidth and battery limited bio-logging instrument programming. We highlight three axes: temporal and spatial resolution of the data, longevity of the data stream (and battery), and continuity or completeness of the data record. Position of the icon indicates roughly the priority balance for each of three setting regimes discussed here: **a** *exploratory* settings with multiple data streams duty-cycled [10], **b** continuous dive summary records only [5], and **c** continuous time-series only

particular to our project, we suggest that the general approach we took provides practical insights for matching tag programming schema to specific research questions by considering general trade-offs, especially those driven primarily by bandwidth limitations.

Methods

Tag deployment and programming overview

The data used in this analysis were a subset of the satellite tags deployed from 2014 to 2019 on *Z. cavirostris*. These instruments were satellite-linked depth-recording SPLASH10-292 tags with the extended depth range option in the LIMPET configuration [14] deployed using a DAN-INJECT JM 25 pneumatic projector (DanWild LLC, Austin, Texas, USA). Tags ($n=16$) were attached with two 6.8-cm surgical grade titanium darts with backward-facing petals to the dorsal fin ($n=12$), base of the dorsal fin ($n=2$) or below the dorsal fin ($n=2$).

We used two primary datasets in this analysis. The first dataset consists of 8 *exploratory* tags (01–08) deployed in 2014–2016 programmed to collect a variety of data streams including both dive summary records (termed behavior log in the tag programming) and time-series data at a 2.5-min sampling period (Table 1). These tags were also configured to duty cycle for maximum longevity (see Additional File 1: Table S1 for details). Finally, data collected more than 48 h prior were deleted from the data message queue to prevent backlogs. The second dataset consists of 8 *assessment* tags (09–16) deployed in 2018 with all optional data streams disabled, except for time-series depth measurements collected at a 5-min sampling period. These samples are packaged in to discrete data messages containing information to decode 48 depth samples (4 h worth). We programmed these tags to collect data only for the first 14 days of deployment, but to transmit data continuously for a total of 100 days (the maximum programmable setting which in our case was always longer than the life of the tag deployment; Table 1). As the Argos system also generates position estimates from these transmissions, location data are not interrupted. We chose 14 days because we estimated it would take approximately another 14 days to uplink a complete data record for a total of approximately 28 days. This total duration is similar both to the estimated battery life and our observed median instrument survival for SPLASH10-292 tags on this species [5, 10]. We deployed these tags as a specific test of the efficacy of this sampling scheme. We also used data from a single tag in the same configuration deployed in 2019 for the purposes of comparison with a higher resolution bio-logger deployed in the same group of animals (see below).

Assessment of exploratory tags

We compared metrics which we calculated from time-series to those reported by the dive summary record data streams on our 8 *exploratory* tags where both of these

Table 1 Telemetry tag deployment summary

ID	Date	Longitude	Latitude	Tag type	Programming	Purpose	Lifetime
Tag01	2014-05-15	− 74.78	35.55	SPLASH10	Dive summary and time-series	Exploratory	60.0
Tag02	2014-09-16	− 74.71	35.66	SPLASH10	Dive summary and time-series	Exploratory	40.2
Tag03	2015-06-14	− 74.74	35.60	SPLASH10	Dive summary and time-series	Exploratory	56.3
Tag04	2015-06-14	− 74.69	35.63	SPLASH10	Dive summary and time-series	Exploratory	1.9
Tag05	2015-10-15	− 74.77	35.61	SPLASH10	Dive summary and time-series	Exploratory	34.4
Tag06	2015-10-21	− 74.75	35.62	SPLASH10	Dive summary and time-series	Exploratory	59.2
Tag07	2016-05-27	− 74.74	35.59	SPLASH10	Dive summary and time-series	Exploratory	36.3
Tag08	2016-08-21	− 74.69	35.61	SPLASH10	Dive summary and time-series	Exploratory	11.3
Tag09	2018-05-24	− 74.78	35.69	SPLASH10	Time-series only	Assessment	38.6
Tag10	2018-08-05	− 74.78	35.73	SPLASH10	Time-series only	Assessment	34.2
Tag11	2018-08-05	− 74.78	35.72	SPLASH10	Time-series only	Assessment	42.8
Tag12	2018-08-05	− 74.75	35.55	SPLASH10	Time-series only	Assessment	43.4
Tag13	2018-08-06	− 74.78	35.48	SPLASH10	Time-series only	Assessment	41.3
Tag14	2018-08-07	− 74.78	35.57	SPLASH10	Time-series only	Assessment	43.3
Tag15	2018-08-07	− 74.78	35.56	SPLASH10	Time-series only	Assessment	43.6
Tag16	2018-08-07	− 74.75	35.59	SPLASH10	Time-series only	Assessment	57.2
Tag17	2019-07-30	− 74.73	35.54	SPLASH10	Time-series only	DTAG comparison	24.9
Tag18	2019-08-06	− 74.75	35.58	DTAG	n/a	DTAG comparison	0.23

Purpose refers to the main use of the tags in the present analysis. DTAG pressure sensor data were decimated to 25 Hz before analysis. Lifetime refers to the total transmission lifetime of the tag in days

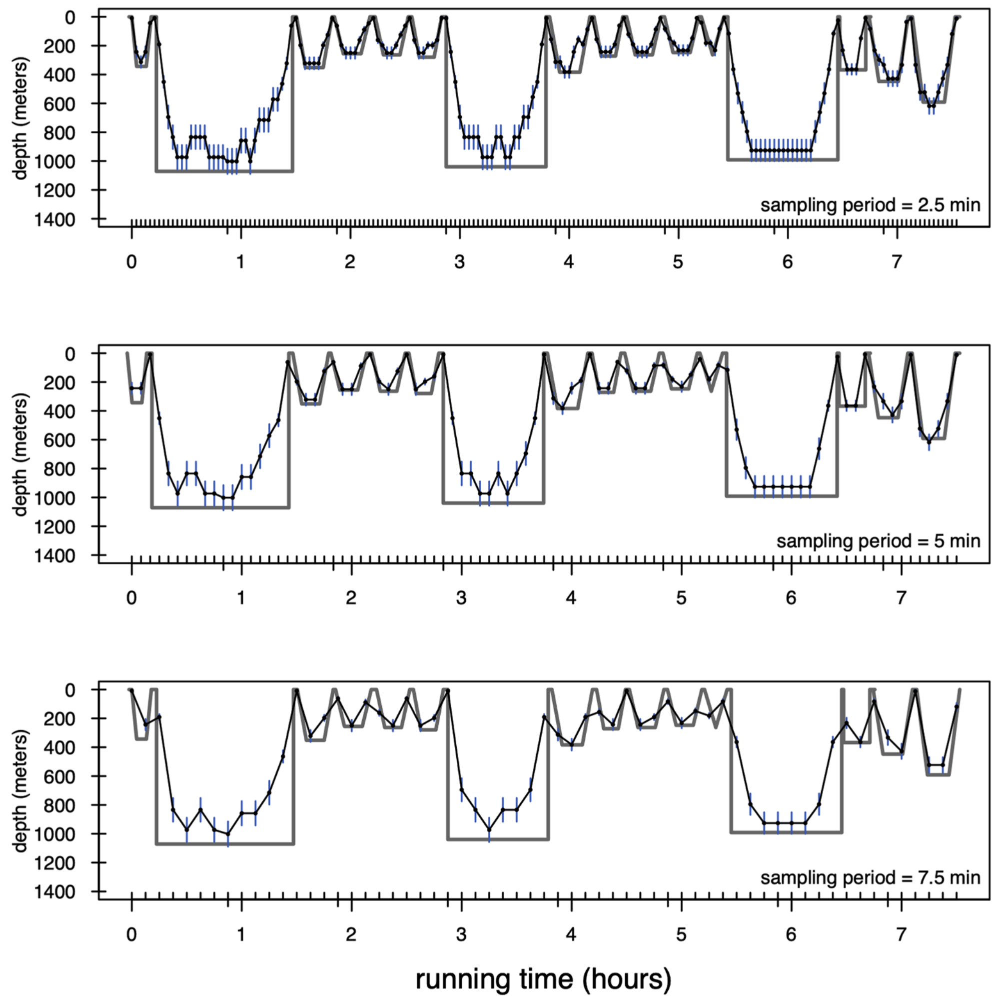

Fig. 2 A representation of two types of dive data streams collected concurrently on a single tag deployed on a *Ziphius cavirostris*. In gray is a pseudo-dive profile based on a dive summary record data stream which provides a maximum depth, start time, and end time for each dive. Note this is not a true dive profile as only maximum depth and general shape of the dive are indicated (see Methods for details). In black are time-series depths with reported error bands indicated by blue segments. Top panel shows the native resolution of time-series data for this deployment (period = 2.5 min), while the middle and bottom panel show resampled time-series data (period = 5.0, 7.5)

data streams were at times running concurrently (Fig. 2, Table 1). This was to assess the degree to which these data streams were comparable. The dive summary record data stream reports dive duration, maximum depth, and dive shape using a conductivity sensor to define the beginning and end of dives when a candidate dive passes a minimum duration (30 s) and depth (50 m) threshold. Both maximum depth and a dive shape metric are calculated from a 1-Hz dive record with reported tolerance of 1 m that is stored onboard the tag, but not transmitted to satellite. Three dive shapes are defined: square shaped dives were scored if greater than 50% of the dive was within 80% of the maximum depth, U-shaped dives were scored if between 20 and 50% of the dive was within 80% of the maximum depth, and V-shaped dives were scored if less than 20% of the dive was within 80% of the maximum depth [15]. Since the dive summary metrics are calculated from much higher resolution input data than the series data we considered the dive summary record data stream as truth and compared it to calculated metrics from the time-series data stream.

We employed a simple algorithm to define dives and convert depth time-series data into a similar format as the dive summary record data. In short, we flagged points where the first derivative of the time-series was crossing zero to determine points nearest a given surfacing. We used a depth filter of 50 m to exclude local maxima. While a filter of 50 m was mostly effective, we visually inspected the resultant dives against the original time-series to correct any false negatives or positives.

To estimate dive durations, we interpolated to the surface from the points nearest the surface using a vertical velocity of 1.4 $m \cdot s^{-1}$ based on the findings of Tyack and coauthors [16], who found that ascent and descent rates varied little for *Z. cavirostris* within several hundred meters of the surface. Maximum depth was estimated as the largest value recorded between the start and end of the dive. We calculated dive shape using the same categories as the dive summary record data stream with the time-series depth samples as input. In addition, we resampled the time-series data from a 2.5-min period to a 5.0- and 7.5-min period to investigate the impact of sampling frequency (Fig. 2). We compared these converted time-series data sets to the dive summary record data on duration, depth, and shape that were recorded simultaneously on the tag.

We also compared depths from two different tag types deployed on different individuals in the same group in 2019, as animals in the same group at the surface are known to maintain high levels of synchrony [17–20]. One of these tags was identical in type and programming to our 2018 time-series only satellite tags; the other was a shorter-term bio-logger attached by suction cups (DTAG) archiving pressure at 250 Hz, processed and decimated to 25 Hz [21].

Additional data collected by UHF antenna

To aid in tracking and data collection we used an Argos Goniometer (henceforth Goniometer) to localize tagged whales and receive data from their transmitters. We developed a visualization software in the form of an R package to assist in real-time tracking of individual tags [22]. In addition, we used data messages downloaded by the Goniometer to supplement those data messages received only via satellite. We converted Goniometer-received hexadecimal data into a format that could be inputted into Wildlife Computer's message decoding utilities using a custom R function [23]. Goniometer effort was approximated using the time difference between the first and last reception of a tag on the instrument per field day. Goniometers were affixed to 1 or 2 vessels per day which may have been engaged in a variety of activities including dedicated searching for previously tagged individuals.

Assessment of time-series only configured tag deployments

Time-series only *assessment* tags ($n=8$) were programmed with a 5.0-min sampling period which was chosen to achieve the highest temporal resolution possible with the most completeness, while sacrificing some longevity of dive data compared to our *exploratory* tags. We measured the overall life of the tag from deployment to the final uplink, and the number of data messages successfully received by satellite from the 14 days of time-series collection, and the number of consecutive messages without a data gap.

Data validation

We checked for mechanical or software failures in our data streams. Status messages periodically report the pressure transducer reading at a presumed zero depth (when the conductivity sensor reads dry). We used these readings and manual inspections of the dive record to identify periods of excessive pressure transducer drift or failure. We defined unacceptably high pressure transducer drift as two or more consecutive absolute value zero depth readings of greater than 10 m [24]. When data were flagged for drift, dive or depth data were not analyzed past the last known good status message, but other types of information such as locations and number of receptions were retained.

All analyses were carried out in the R programming language version 3.6.2 [25]. R packages colorspace 1.4–1, ggplot2 3.2.1, reshape2 1.4.3 and R.matlab 3.6.2 were used in visualizations [26–29].

Results

Assessment of exploratory tags

We extracted a total of 645 dives from the 2.5-min sampling period time-series data, compared to 598 and 457 for the 5.0- and 7.5-min sampling periods, respectively ($n=8$ tags in all cases). Mean difference between time-series extracted dive duration and the dive summary derived dive duration (error) increased as sampling period increased, although the maximum error was similar (Table 2). Most dive duration errors were within the theoretical maximum of twice the sampling period (Fig. 3). Time-series data tended to underestimate

Table 2 Mean absolute value differences among time-series data and concurrently collected dive summary records derived from the 8 *exploratory* tags

	2.5 min period	5.0 min period	7.5 min period
n dives extracted	645	598	457
Duration (s)	69 (0–1338)	182 (0–1589)	293 (0–1589)
Depth (m), all dives	30 (1–147)	34 (1–147)	43 (1–234)
Depth (m), dives < 33 min	18 (1–95)	21 (1–101)	27 (1–152)
Depth (m), dives > 33 min	81 (2–147)	82 (16–147)	83 (16–234)

Dive summary records include a dive duration calculated from submergence to emergence in air (as measured by a conductivity sensor) and maximum depth of each dive (as measured from an onboard pressure transducer sampled at 1 Hz). Time-series data were recorded at a 2.5-min sampling period and resampled to 5.0- and 7.5-min periods. Values in brackets indicate ranges

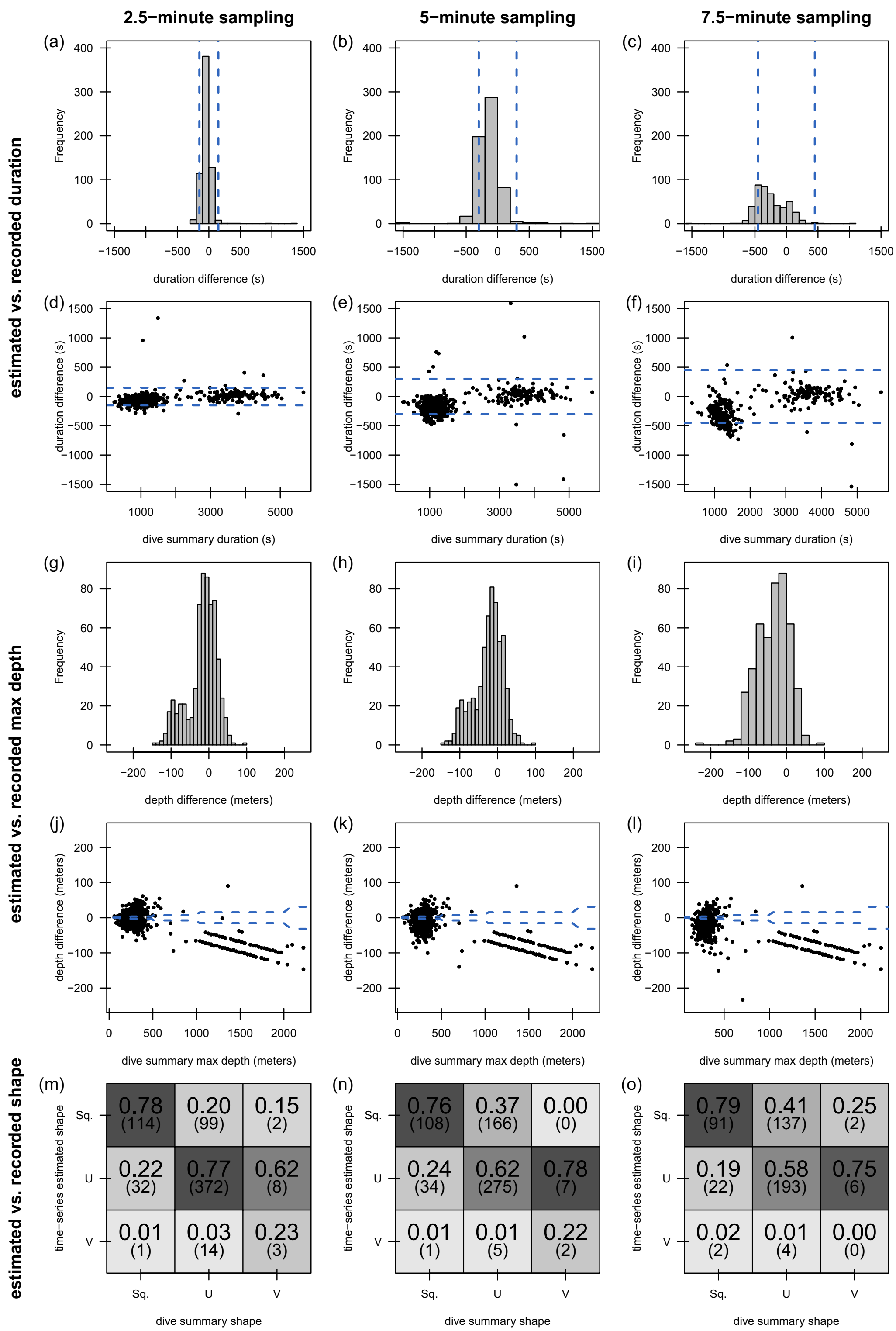

Fig. 3 Dive metrics calculated from time-series depths recorded at a native sampling period of 2.5 min and resampled to 5.0 and 7.5 min compared to concurrently collected dive summary records derived from conductivity sensor detected dives. **a**–**f** Show the distributions of the difference in dive duration between the two data collection methods for each sampling period. Blue broken lines indicate a theoretical error bound based on sampling period. **g**-**l** Show the distributions of difference in dive depth between the two data collection methods for each sampling period. Blue broken lines indicate the recorded error bounds for the dive summary record data maximum depth. **m**–**o** Show a confusion matrix of dive shapes calculated on board the tag (see Methods for details) and transmitted in the dive summary records, and estimated post hoc from the time-series

maximum depth, probably due to short forays to depths missed by the relatively coarse sampling. The linear effect seen at deeper depths is probably due to the encoding algorithm the tag uses (see a more detailed description of this algorithm in the discussion). Mean depth error increased only very slightly with increasing sampling period, but the maximum depth error increased more substantially (Table 2). Correct assignment of dive shape also decreased with increasing sampling period from approximately 76% correct at 2.5 min to 65% and 63% at 5.0 and 7.5 min, respectively. V-shaped dives were the most often miscategorized by the time-series data, but this type of dive was also the rarest. Square-shaped dives were the best identified at 76–79% correct for all sampling periods (Fig. 3).

We also compared dive depths between the two whales tagged in the same group, which we expected to be highly synchronous. One of the pair was tagged with a high-resolution DTAG and the other with a time-series programmed (5-min period) SPLASH10 tag (Fig. 4). DTAG depth calibration error was 2.3 m. Depths were highly correlated between the two instruments ($n = 56$ samples, $R^2 = 0.99$) with a mean depth difference of 30 m. Note that this difference includes both measurement error (from both tags), as well as any difference in the behavior of the two animals.

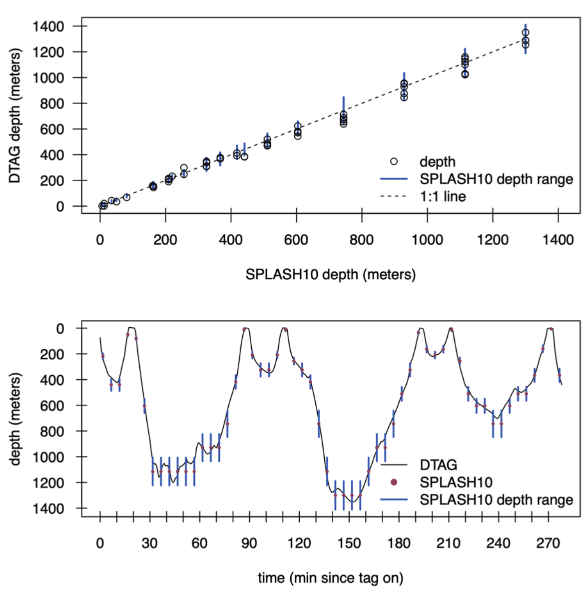

Fig. 4 Comparison of dive depths between a satellite tag configured to collect time-series depth data and a DTAG deployed on different individuals in the same group. Top panel shows depth comparison with the broken line indicating 1:1 and gray segments indicating the reported error range for time-series reported depths. Bottom panels show the DTAG dive profile (Tag18) overlaid with 5.0-min time-series depth sampling from the satellite tag (Tag17)

Assessment deployments of time-series tags

We deployed eight time-series only tags to assess the performance of this setting regime. We programmed these tags to collect 14 days of dive data, before transitioning to a transmit-only phase. One tag was deployed below the dorsal fin and never successfully transmitted any data (Tag10). A second tag suffered an unknown malfunction, apparently restarting at random intervals, which diminished the amount of transmitted data (Tag14). A third tag (Tag13) apparently experienced a reset about 37 days after deployment which had little impact on the tag performance. Of the 6 tags which functioned nominally for an extended period of time, 2 experienced significant pressure transducer drift, resulting in truncation of the reliable data. These tags could still be assessed for data completion and transmission statistics because these aspects of tag performance were unaffected by the pressure transducer malfunction.

From the 6 tags that uplinked data, there were 3 data gaps across 3 different tags. Two of these gaps were 8 h (2 data messages), while the last was only 4 h (1 data message). Most data messages were received by satellite or by both satellite and Goniometer. There was a general pattern that the very first messages were almost always received successfully as the tag's message queue was relatively empty at the beginning of the deployment, but data gaps otherwise occurred throughout the rest of the record. Most individual data messages were received several times, although all but one tag had at least one message only received once (see Additional File 2: Fig. S1 for details). Goniometer effort (time between reception of first and last Goniometer message) totaled approximately 212 h over 31 days for our first vessel and approximately 178 h over 26 days for our second vessel. Seven dive data messages (across 3 tags) were only received via Goniometer in the field (Fig. 5). The Goniometer also received additional status messages which did not reach satellites, although there was a higher rate of corrupt messages received in the field via the Goniometer than via satellite (Fig. 6). Daily rate of successfully decoded Goniometer messages was approximately 5 to 25 times greater than from satellite, while the corrupt rate ranged from 9 to 36 times greater.

Discussion

Skin-piercing LIMPET configured telemetry tags are more invasive than many other possible data collection techniques and their deployment involves risk both to animal subjects and researchers [8]. Further, experimental behavioral response studies are by design invasive as a potentially harmful stimulus is introduced

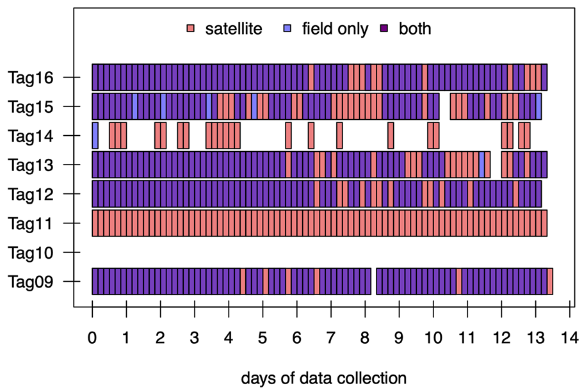

Fig. 5 Time-series message reception for 8 *assessment* tags in the time-series only programming configuration. Each block represents 48 time-series data points (= 4 h of data at our sampling period of 5 min). Colors denote if a message was received in the field only (via Argos Goniometer), from the Argos satellite system only or from both sources. Only successfully decoded messages are included in this plot. Tag10 never transmitted a successfully decoded time-series message. Note that total length of record varies as tags are programmed to record for 14 calendar days as opposed to exactly 336 h. Malfunctioning tags (Tag12, 14) were truncated to 14 days for comparison purposes

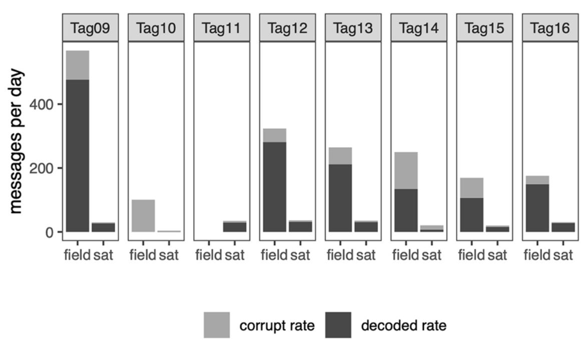

Fig. 6 Comparison of tag message capture rate between the Argos satellite system (sat) and Argos Goniometer (field) for 8 *assessment* tags in a time-series programming configuration. The proportion of messages successfully decoded indicated by darker bars. Rates were calculated using the time from the first message received from a particular tag on a given day to the last message received that day. Tag11 was omitted from the field calculation, since it was only in reception range for a short period during deployment and not revisited

into the animals' environment. Therefore every effort must be made to maximize the quantity and quality of data while minimizing impact and using the smallest possible number of animals. To those ends, this paper details the framework we used to match data collection to study questions, including when those questions evolved during the course of a single project.

Sampling rate and data record length

We chose a 5-min sampling period for time-series to maximize the temporal length of data, while reducing depth aliasing effects and minimizing data gaps due to limited uplink bandwidth. At the 5-min sampling period, we estimate about twice as many data messages are generated per day than can be uplinked to satellite. We solved this problem by truncating data collection to 14 days, and then transitioning the tags to cease dive data collection but continue transmitting for up to 100 days (the maximum), enabling the tag to uplink the backlog of messages.

Relying on a post-data collection period to complete data transmissions also creates the risk that a tag could fail before the data uploading is complete. Our sampling length of 14 days was largely dictated by this concern. This total duration fell within the calculated battery life of the instrument, given our programming, and commensurate with average deployment life observed in our study area [10]. In addition, 14 days was sufficient to conduct experimental treatments including adequate baseline and post-exposure periods given that experimental treatments were targeted for the middle of this period.

We were able to show high depth concordance in a single instance where one of our tags programmed in this manner was deployed in a group with another animal instrumented with a higher-resolution archival tag (Fig. 4). Though diving behavior can be very synchronous in *Z. cavirostris*, there is evidence that groups split at the bottom of long dives, so data from separately tagged individuals must be considered cautiously [17–20]. Additional examples and doubly tagged individuals would provide more reliable data on this point, but our current evidence is consistent with a fair representation of overall dive patterns.

This sampling scheme was in part only viable for our experimental questions because the study species, *Z. cavirostris*, perform long deep dives, thus alleviating some of the depth and temporal resolution limitations of this tagging system and programming regime. To answer similar questions about shallower, shorter or faster diving species or to answer questions about very small changes in depth, higher sampling rates and depth resolution would have been necessary. For some species and questions, an entirely different instrument may be necessary. Nevertheless, the iterative approach to considering trade-offs we outline is applicable to those situations and would provide guidance on selecting the best data streams and programming details fit for a particular experiment or inquiry.

Additional data collected by UHF antenna

To decrease the risks of not receiving data packages via satellites associated with this programming regime, we utilized a Goniometer to download additional data. Fortunately, by the end of our test deployments almost all messages that were obtained with the Goniometer had also been successfully received via satellite transmission, but the Goniometer provided additional security by collecting the most crucial messages (concurrent in time to the experimental exposures) before they were received by satellite and therefore guarding against potential future tag failure. Near real-time monitoring of received messages from experimental animals was possible in the field, which enabled strategic sampling to fill prioritized gaps and increased our ability to capture complete records by staying with an animal longer or moving on to other priority animals after highest priority messages were captured. Vessels with extended endurance (overnight capabilities) can greatly increase the potential data reception bandwidth, allowing for finer scale sampling or longer duration of sampling. These types of benefits have also been demonstrated in field sites with suitable land stations nearby [7].

Iterative approach

Although we have focused on data resolution, longevity, and continuity, there are many other important factors to consider when deploying tags. These considerations include: weighing the risk of harm to the animal with the value of data collected [8, 30, 31]; the cost and time expenditure in deployment and analysis; how sample size is affected by programming regimes [32]; the appropriateness of data to biological questions [33]; species behavior; and the probability of success in achieving the experimental objectives during critical data collection periods.

One possible downside to tailoring tag programming regimes to each question or experiment is the complication of creating non-comparable datasets. For instance, if tags are deployed in the context of a long-term study, year-to-year comparisons may be of interest. For that reason, it is often more beneficial to collect data in a fashion such that it can be compared to historical samples, even as new questions and protocols are added to a project. In our case, the exploratory data collection paradigm was not suitable to meet the specific experimental objectives from the Atlantic behavioral response study, given the short temporal (up to 1 h) nature of experimental treatments [5].

Assessment time-series data

Onboard data processing can increase the efficiency of bio-logging devices, especially on those which transmit data with bandwidth limited systems. For instance, when using the dive summary records to capture only long foraging dives, each data message comprises approximately 9 h of *Z. cavirostris* behavior depending on the diving rate. In contrast, a time-series data stream set to a 5.0-min sampling period only comprises 4 h of data in a message and is dive rate independent. For species or applications where finer sampling is needed, this would be further reduced. In return, however, a true time-series even at relatively coarse depth resolution allows the calculation of activity budgets and summary statistics based on depth, spectral densities, custom shape parameters, and other vertical movement parameters. These data are also well suited for more sophisticated continuous time behavioral modeling (for example, [34, 35]). Again, our ability to recover this type of information from a relatively coarse diving time-series depends on the long deep dives of *Z. cavirostris* and sampling rate and depth resolution would need to be considered for other species and applications and of course can be incorporated into models in a straight forward manner. Importantly, even in this species, shorter dives under this sampling scheme can be extremely aliased or completely obliterated and so careful consideration must be made when attempting to back out dive-by-dive metrics. If absolute dive durations are paramount for an analysis then another type of data stream or instrument would be needed. Additionally, shape as calculated in the dive summary record data stream seems to be unreliable for some classes. Nevertheless, we note that the time-series allows for different types of shape analysis, which may be as relevant or more to diving behavior.

A further benefit of the time-series data in our case was that depth measurements were closely linked to a real-time clock, which was in contrast to the more temporally imprecise records in the dive summary record data stream (start time is only measured to the minute). In the time-series, any concurrent tags sample almost simultaneously, allowing for direct comparison of the diving behavior of animals tagged within and between groups or, as we have shown here, even between different instrument types.

Another major consideration in our experiment was the depth resolution loss in the time-series only configuration. The lower depth resolution in the time-series was partially compensated by the fact that multiple depths were sampled during each dive, as opposed to a single depth in the dive summary record data stream (maximum depth) albeit at a finer resolution, but this could be an important consideration depending on the application.

Species-specific behavior

The sampling resolution and depth accuracy to resolve, for example, individual dives are highly taxon-dependent, as is the degree to which the animal's diving behavior creates bandwidth bottlenecks. *Z. cavirostris* create a significant bandwidth bottleneck by virtue of the small amount of time they spend at the surface, but the fact that their dives tend to be long and deep offsets this challenge by permitting the use of coarser sampling resolutions. Even the shorter dives of *Z. cavirostris* average 19 min [10], so a sampling period of 5 min does not typically cause aliasing, which could obscure dive events in the time-series record. Considering the sample period alone, it should be possible to detect any dive of 10 min or greater. Due to the limitations in depth accuracy and variation in dive shape, however, short and shallow dives may sometimes be unobserved. As the shorter dives of *Z. cavirostris* also tend to be relatively deep (> 100 m), this is not typically a problem for this species. In comparison, in the sympatric population of short-finned pilot whales (*Globicephala macrorhynchus*), the maximum recorded dive duration is 26 min and dives are typically shallower than for *Z. cavirostris* with a maximum recorded depth of 1360 m [9, 36]. Therefore, a 5-min sampling period would be insufficient to capture the same percentage of dives for this population. In fact, for some applications this type of tag may not deliver suitable data at all for more shallowly diving species.

Tag failure and limitations

We experienced multiple instrument failures during the deployments of the time-series *assessment* tags. Three of the 8 tags suffered catastrophic failures rendering most of the return data unusable (a fourth tag was deployed too low on the animal to break the surface and transmit data messages). Such equipment failures are unavoidable in small-run electronics, especially when exposed to extreme conditions at or beyond their tolerances such as those deployed on deep-diving cetaceans, but failure rate must also be incorporated into the risk assessment of any programming scheme and, indeed, any tagging program [8]. In this case, early failure could lead to dramatic reductions in completeness of the data record, so we took steps to mitigate this using the Goniometer.

Design limitations in our chosen instruments also impacted our data even when tags were functioning to specification. For example, the depth resolution of the time-series data in SPLASH10 tags is dynamically calculated from the maximum recorded depth (transmitted at some resolution itself) for a 1, 2, 4, or 8 h data block (corresponding to the different sampling period options). Depths are split into 16 bins, which are narrower at shallow depths and wider near the maximum depth. This encoding is convenient since each depth point can be stored as just 4 bits, but can also cause complications in modeling as the resolution is constantly changing. In addition, the manufacturer declined to share the exact encoding algorithm, which further hampers efforts to produce consistent and reproducible analyses. There are other drawbacks in the dive summary record data stream, such as a lack of precision in the recorded timestamp of data messages (presumably to save bandwidth). The limitations mentioned here are device specific, but all instruments involve trade-offs in data collection choices, and these examples serve to highlight the general need to consider downstream data analysis before data collection especially in high risk projects and/or invasive protocols.

Future development

There is a clear need for more flexible and/or purpose-built bio-logging instruments to answer many of the current and pressing questions in large marine vertebrate research, especially within the context of experimental behavioral response studies (for example, [2]). Newer satellite systems such as the upcoming Kinéis constellation (CLS and the French Space Agency) or the currently stalled Icarus Initiative's space instrument may also help alleviate some of the current problems but will of course require parallel hardware development. In addition, there are specific requirements for instruments that are pushed to extreme environments, such as the significant pressure ranges visited by beaked whales. Hardware development is very expensive and, therefore, not always feasible, although in the case of deep-diving cetaceans at conservation risk, such research would be advantageous.

In our study, new hardware development was not possible, but we were able to tailor the sampling regimes of existing instruments available to us to better fit our requirements. Through this process it became clear more flexible and transparent hardware and software are needed. Additional control to set sampling rates and regimes could lead to more creative solutions in difficult bio-logging problems that would in turn enable data collection for a greater array of biological and applied conservation questions. Open source instruments could be a solution to creating accessible, flexible platforms for asking these questions consistently, transparently, and reproducibly and indeed these types of devices are on the rise (for example, [37]). This route will require strong partnerships between engineers and biologists (for example, [38]), and significant and ongoing commitment from funders.

Final recommendations

Lessons from our deployment and programming strategies should be generalizable to similar problems in other taxa and contribute to a growing literature on best practices in bio-telemetry. Our recommendations are to follow the logical thought process of any complex field experiment with specific objectives and constraints: start with the research questions, design analyses to address specific components, and optimize data collection for those analyses and questions. Trialing data collection methods with pilot data or real deployments provides added value and allows for protocol refinement. Extensive testing, as presented here, is expensive and sometimes infeasible given the constraints of research budgets and the objectives of applied studies. Our funders allowed us to strategically and systematically evaluate tag settings to determine optimal solutions to best meet the specific research objectives of long-term studies of baseline behavior and behavioral responses of whales to sonar in our study site. The level of testing described here may not always be desirable and must be weighed with the potential impacts of an invasive instrumentation and the overall risk of a project. Computer simulations and bench tests are viable alternatives, but the intersection of animal behavior, weather, deployment location, and satellite coverage can be difficult to model or reproduce in the lab. A hybrid approach using simulation or modeling based on similar species and deployments can also increase the likelihood of success in field tests. Together these suggestions can serve to maximize scientific yield while seeking to minimize risk and impact to the study subjects.

Abbreviations

UHF	Ultra-high frequency
LIMPET	Low impact minimally percutaneous electronic transmitter

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40317-023-00334-1.

Additional file 1: Table S1. settings.csv. Detailed settings parameters for SPLASH10 tags.

Additional file 2: Figure S1: Time-series message reception count via satellite only for 8 *assessment* tags in the time-series only programming configuration.

Acknowledgements

We thank Joel Bell of Naval Facilities Engineering Command Atlantic for sustained project support and the latitude to experiment with methodology. Jessica Aschettino of HDR, Inc. provided assistance with tag deployments in 2018 as well as continued logistical support along with Daniel Engelhaupt of HDR, Inc. Jeanne Shearer prepared the DTAG data. We thank all members of the field team including Rafaella Lobo, Andrew Westgate, Jillian Wisse, Eleanor Heywood, Captain Reed Meredith of the F/V Kahuna, and Captain Jimmy Horning Jr. of the F/V Hog Wild. Greg Schorr and Dave Haas provided helpful input on tag programming and strategy. Matthew Rutishauser and Kenady Wilson of Wildlife Computers assisted during discussions on technical aspects of the tags. We thank the two anonymous reviewers for their thoughtful and helpful comments.

Author contributions

AR, BS, DN, RB initially conceived the satellite tag deployment design (the behavioral response study) which generated these data and DLW, DMW, HF, NQ, WC, ZS provided additional input. DLW, DMW, HF, WC, ZS led the field work and data collection; AR, BS, DN, NQ, provided additional field assistance. WC conceived of the current study, analyzed the data, and wrote the first draft of the manuscript with input and revisions from all authors. All authors read and approved the final manuscript.

Funding

This work was supported by the US Fleet Forces Command Marine Species Monitoring Program through the Naval Facilities Engineering Command Atlantic under contract nos. N62470-10-D-3011 (Task Orders 14, 21), N62470-10-D-3011 (Task Order 57), and N62470-15-D-8006 (Task Orders 07, 28, 18F4036, 19F4029) issued to HDR, Inc.

Availability of data and materials

All scripts and data used to produce these analyses and figures are available at https://github.com/williamcioffi/zc_series (https://doi.org/10.5281/zenodo.6589596).

Declarations

Ethics approval and consent to participate

Satellite tags were deployed under National Marine Fisheries Service scientific research permit numbers 17086 and 20605 to Robin W. Baird, 14809-03 to Douglas P. Nowacek, and 16239 to Daniel T. Engelhaupt. Photo identification was conducted under National Marine Fisheries Service general authorization letter of confirmation number 19903 to Andrew J. Read. All activities were approved by Institutional Animal Care and Use Committees at the respective institutions.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Duke University Marine Laboratory, Nicholas School of the Environment, Duke University, Beaufort, NC, USA. [2]Southall Environmental Associates, Inc., Aptos, CA, USA. [3]School of Biological and Marine Science, University of Plymouth, Drake's Circus, Plymouth, UK. [4]Bridger Consulting Group, Bozeman, MT, USA. [5]Cascadia Research Collective, Olympia, WA, USA. [6]Electrical and Computer Engineering, Pratt School of Engineering, Duke University, Durham, NC, USA.

Received: 28 May 2022 Accepted: 16 May 2023
Published online: 29 May 2023

References

1. Photopoulou T, Fedak MA, Matthiopoulos J, McConnell B, Lovell P. The generalized data management and collection protocol for conductivity-temperature-depth satellite relay data loggers. Anim Biotelem. 2015. https://doi.org/10.1186/s40317-015-0053-8.
2. Schorr GS, Rone BA, Falcone EA. Integrated measurement of naval sonar operations and precise cetacean locations: integration of Fastloc GPS into a LIMPET tag. Final Report for Task C, contract Number: N66604–14-C-2438. Marine Ecology and Telemetry Research. 2017. https://apps.dtic.mil/sti/trecms/pdf/AD1073443.pdf

3. Mul E, Blanchet M-A, Biuw M, Rikardsen A. Implications of tag positioning and performance on the analysis of cetacean movement. Anim Biotelem. 2019. https://doi.org/10.1186/s40317-019-0173-7.
4. Fedak MA, Lovell P, Grant SM. Two approaches to compressing and interpreting time-depth information as collected by time-depth recorders and satellite-linked data records. Marine Mammal Sci. 2001. https://doi.org/10.1111/j.1748-7692.2001.tb00982.x.
5. Quick NJ, Cioffi WR, Shearer J, Read AJ. Mind the gap—optimizing satellite tag settings for time series analysis of foraging dives in Cuvier's beaked whales (*Ziphius cavirostris*). Animal Biotelemetry. 2019;7(1):5. https://doi.org/10.1186/s40317-019-0167-5.
6. Musyl MK, Domeier ML, Nasby-Lucas N, Brill RW, McNaughton LM, Swimmer JY, Lutcavage MS, Wilson SG, Galuardi B, Liddle JB. Performance of pop-up satellite archival tags. Mar Ecol Prog Ser. 2011;433:1–28. https://doi.org/10.3354/meps09202.
7. Jeanniard-du-Dot T, Holland K, Schorr GS, Vo D. Motes enhance data recovery from satellite-relayed biologgers and can facilitate collaborative research into marine habitat utilization. Anim Biotelem. 2017;5(1):17. https://doi.org/10.1186/s40317-017-0132-0.
8. Andrews RD, Baird RW, Calambokidis J, Goertz CEC, Gulland FMD, Heide-Jorgensen MP, Hooker SK, Johnson M, Mate B, Mitani Y, Nowacek DP, Owen K, Quakenbush LT, Raverty S, Robbins J, Schorr GS, Shpak OV, Townsend FI Jr, Uhart M, Wells RS, Zerbini AN. Best practice guidelines for cetacean tagging. IWC J Cetacean Res Manag. 2019;20(1):27–66. https://doi.org/10.47536/jcrm.v20i1.237.
9. Baird RW, Webster DL, Swaim ZT, Foley HJ, Anderson DB, Read AJ. Spatial Use by Cuvier's Beaked Whales, Short-finned Pilot Whales, Common Bottlenose Dolphins, and Short-beaked Common Dolphins Satellite Tagged off Cape Hatteras, North Carolina, in 2014. Prepared for U.S. Fleet Forces Command. Submitted to Naval Facilities Engineering Command Atlantic, Norfolk, Virginia, under Contract No. N62470–10-D-3011, Task Orders 14 and 21, Issued to HDR, Inc. 2015. https://www.navymarinespeciesmonitoring.us/files/7814/3750/5412/Baird_et_al_2015_Hatteras_Odontocete_Tagging_-_FINAL.pdf
10. Shearer JM, Quick NJ, Cioffi WR, Baird RW, Webster DL, Foley HJ, Swaim ZT, Waples DM, Bell JT, Read AJ. Diving behaviour of Cuvier's beaked whales (*Ziphius cavirostris*) off Cape Hatteras. North Carolina Royal Soc Open Sci. 2019;6(2):181728. https://doi.org/10.1098/rsos.181728.
11. Southall BL, Moretti D, Abraham B, Calambokidis J, DeRuiter SL, Tyack PL. Marine mammal behavioral response studies in southern california: advances in technology and experimental methods. Mar Technol Soc J. 2012;46(4):48–59. https://doi.org/10.4031/MTSJ.46.4.1.
12. Southall BL, Nowacek DP, Miller PJO, Tyack PL. Experimental field studies to measure behavioral responses of cetaceans to sonar. Endanger Species Res. 2016;31:293–315. https://doi.org/10.3354/esr00764.
13. Schick RS, Bowers M, DeRuiter S, Friedlaender A, Joseph J, Margolina T, Nowacek DP, Southall BL. Accounting for positional uncertainty when modeling received levels for tagged cetaceans exposed to sonar. Aquat Mamm. 2019;45(6):675–90. https://doi.org/10.1578/AM.45.6.2019.675. Accessed2019-11-20.
14. Andrews RD, Pitman RL, Ballance LT. Satellite tracking reveals distinct movement patterns for Type B and Type C killer whales in the southern Ross sea. Antarct Polar Biol. 2008;31(12):1461–8. https://doi.org/10.1007/s00300-008-0487-z.
15. Wildlife Computers: SPLASH10 (-F, -BF, -FL, -X, -L, -LX, -FX) TDR10 (-DD, -F, -BF, -X, -L, FL, -FX, -LX, -BX) With host version 1.26.3002 User Guide. 2019. https://static.wildlifecomputers.com/SPLASH10-TDR10-User-Guide.pdf
16. Tyack PL, Johnson M, de Aguilar Soto N, Sturlese A, Madsen PT. Extreme diving of beaked whales. J Exp Biol. 2006;209(21):4238–53. https://doi.org/10.1242/jeb.02505.
17. Alcázar-Treviño J, Johnson M, Arranz P, Warren VE, Pérez-González CJ, Marques T, Madsen PT, de Aguilar Soto N. Deep-diving beaked whales dive together but forage apart. Proc Royal Soc B. 2021;288(1942):20201905. https://doi.org/10.1098/rspb.2020.1905.
18. Aguilar de Soto N, Visser F, Tyack PL, Alcazar J, Ruxton G, Arranz P, Madsen PT, Johnson M. Fear of killer whales drives extreme synchrony in deep diving beaked whales. Sci Rep. 2020. https://doi.org/10.1038/s41598-019-55911-3.
19. Aguilar de Soto N, Visser F, Madsen PT, Tyack P, Ruxton G, Alcazar J, Arranz P, Johnson M. Beaked and killer whales show how collective prey behaviour foils acoustic predators. bioRxiv. 2018. https://doi.org/10.1101/303743.
20. Cioffi WR, Quick NJ, Foley HJ, Waples DM, Swaim ZT, Shearer JM, Webster DL, Friedlaender AS, Southall BL, Baird RW, Nowacek DP, Read AJ. Adult male Cuvier's beaked whales (*Ziphius cavirostris*) engage in prolonged bouts of synchronous diving. Mar Mamm Sci. 2021;37(3):1085–100. https://doi.org/10.1111/mms.12799.
21. Johnson MP, Tyack PL. A digital acoustic recording tag for measuring the response of wild marine mammals to sound. IEEE J Oceanic Eng. 2003;28(1):3–12. https://doi.org/10.1109/JOE.2002.808212.
22. Cioffi WR. Monitorgonio: visualize Argos goniometer output in the field. (2020). 10.5281/zenodo.3647687. https://www.github.com/williamcioffi/monitorgonio
23. Cioffi WR. Parsegonio: parse Argos Goniometer log data and convert into Prv. (2020). 10.5281/zenodo.3698261. https://www.github.com/williamcioffi/parsegonio
24. Baird RW, Webster DL, Swaim ZT, Foley HJ, Anderson DB, Read AJ. Spatial Use by Cuvier's Beaked Whales and Short-finned Pilot Whales Satellite Tagged off Cape Hatteras, North Carolina in 2017. Prepared for U.S. Fleet Forces Command. Submitted to Naval Facilities Engineering Command Atlantic, Norfolk, Virginia, under Contract No. N62470–15-D-8006, Task Order 50, Issued to HDR, Inc. 2018. https://www.navymarinespeciesmonitoring.us/files/3415/2105/6871/Baird_et_al._2018_-_Hatteras_Odontocete_Tagging_2017_-_FINAL.pdf
25. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. R Foundation for Statistical Computing. 2019. https://www.R-project.org/.
26. Zeileis A, Fisher JC, Hornik K, Ihaka R, McWhite CD, Murrell P, Stauffer R, Wilke CO. colorspace: a toolbox for manipulating and assessing colors and palettes. J Stat Softw. 2020;96(1):1–49. https://doi.org/10.18637/jss.v096.i01.
27. Zeileis A, Hornik K, Murrell P. Escaping RGBland: selecting colors for statistical graphics. Comput Stat Data Anal. 2009;53(9):3259–70. https://doi.org/10.1016/j.csda.2008.11.033.
28. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2016.
29. Wickham H. Reshaping data with the reshape package. J Stat Softw. 2007;21(12):1–20. https://doi.org/10.18637/jss.v021.i12.
30. Bodey TW, Cleasby IR, Bell F, Parr N, Schultz A, Votier SC, Bearhop S. A phylogenetically controlled meta-analysis of biologging device effects on birds: deleterious effects and a call for more standardized reporting of study data. Methods Ecol Evol. 2018;9(4):946–55. https://doi.org/10.1111/2041-210X.12934.
31. Horning M, Andrews RD, Bishop AM, Boveng PL, Costa DP, Crocker DE, Haulena M, Hindell M, Hindle AG, Holser RR, Hooker SK, Hückstädt LA, Johnson S, Lea M-A, McDonald BI, McMahon CR, Robinson PW, Sattler RL, Shuert CR, Steingass SM, Thompson D, Tuomi PA, Williams CL, Womble JN. Best practice recommendations for the use of external telemetry devices on pinnipeds. Anim Biotelem. 2019. https://doi.org/10.1186/s40317-019-0182-6.
32. Pollock K. Experimental design of telemetry projects. J Raptor Res. 1987;21(4):129–31.
33. Hebblewhite M, Haydon DT. Distinguishing technology from biology: a critical review of the use of gps telemetry data in ecology. Philos Trans Royal Soc B. 2010;365(1550):2303–12. https://doi.org/10.1098/rstb.2010.0087.
34. Hewitt J, Schick RS, Gelfand AE. Continuous-time discrete-state modeling for deep whale dives. J Agric Biol Environ Stat. 2021;26(2):180–99. https://doi.org/10.1007/s13253-020-00422-2.
35. Hewitt J, Gelfand AE, Quick NJ, Cioffi WR, Southall BL, DeRuiter SL, Schick RS. Kernel density estimation of conditional distributions to detect responses in satellite tag data. Anim Biotelem. 2022;10(1):28. https://doi.org/10.1186/s40317-022-00299-7.
36. Quick NJ, Isojunno S, Sadykova D, Bowers M, Nowacek DP, Read AJ. Hidden Markov models reveal complexity in the diving behaviour of short-finned pilot whales. Sci Rep. 2017;7:45765. https://doi.org/10.1038/srep45765.
37. Fahlbusch JA, Harrington KJ. A low-cost, open-source inertial movement GPS logger for eco-physiology applications. J Exp Biol. 2019. https://doi.org/10.1242/jeb.211136.

38. Holton MD, Wilson RP, Teilmann J, Siebert U. Animal tag technology keeps coming of age: an engineering perspective. Philos Trans Royal Soc B. 2021;376(1831):20200229. https://doi.org/10.1098/rstb.2020.0229.

Publisher's Note