The generalized data management and collection protocol for Conductivity-Temperature-Depth Satellite Relay Data Loggers

Photopoulou *et al.*

Animal Biotelemetry

Open Access

The generalized data management and collection protocol for Conductivity-Temperature-Depth Satellite Relay Data Loggers

Theoni Photopoulou[1*], Michael A. Fedak[2], Jason Matthiopoulos[3], Bernie McConnell[2] and Phil Lovell[2]

Abstract

The software routines for data sampling and processing that are implemented on-board telemetry devices (tags) called Conductivity-Temperature-Depth Satellite Relay Data Loggers (CTD-SRDLs) enable the simultaneous collection of biological and *in-situ* environmental data by animal-platforms over periods of weeks to months, despite severe energy and bandwidth limitations imposed by their relatively small size. This extended operational lifetime is made possible by the use of software protocols on-board the tags that manage sensors, data collection, storage, compression and transmission to ensure that the most useful data are sent at appropriate resolution while minimizing redundancy. While tag software is tailored to the particular species under study and the questions being addressed with a given field deployment, the philosophy behind Sea Mammal Research Unit Instrumentation Group (SMRU-IG) software protocols is to adopt a general set of principles to achieve the best results within the energy and bandwidth constraints. Here, we discuss these and review the general protocol that is used to simultaneously collect information on geographical movements, diving behaviour and in-situ oceanographic information from marine mammals.

Keywords: CTD-SRDL, SMRU-IG, Software protocol, CLS-Argos, Marine mammals, Diving, Oceanography, Constraints

Background

The development and manufacture of animal telemetry devices that exploit the Argos satellite data collection and transmission system (CLS-Argos, http://www.argos-system.org/) began at the Sea Mammal Research Unit (SMRU) in the late 1980s [1]. Research and development are on-going by a dedicated group within SMRU (SMRU Instrumentation Group; SMRU-IG) using a package of sensors produced by Valeport Ltd (UK), but a stable version of these tags that has been routinely deployed since 2003 on wide-ranging polar marine mammals, is the Conductivity-Temperature-Depth Satellite Data Relay Logger, or CTD-SRDL (Fig. 1).

A precursor of the CTD-SRDL was deployed on beluga whales in 2001 to study the hydrographic conditions in an Arctic fjord [2], in Svalbard, and on Antarctic

fur seals at Bird Island, South Georgia [3]. Some of the first studies to use CTD-SRDLs involved the instrumentation of southern elephant seals from South Georgia and Kerguelen [4, 5] illustrating the usefulness of the devices for jointly studying mobile top predator behaviour and ecology and the ocean environment they inhabit [6]. Data from CTD-SRDLs are used not only to learn about the instrumented animals themselves [4, 7, 8], but also to update bathymetric information in poorly surveyed areas [9], collect extensive and sometimes otherwise inaccessible oceanographic data [10–15], and have become substantial contributors to the global ocean observing system [16–19].

The purpose of this instrument and its data collection protocol is to remotely and autonomously collect and transmit simultaneous hydrographic and behavioural data from air-breathing marine vertebrates. Hence, the protocol is designed for the collection of time-series data over long periods, while being constrained by the physical size of the device and bandwidth for data transfer; both those imposed by the limit time animals are at the surface

* Correspondence: theoni.photopoulou@uct.ac.za
[1]Centre for Statistics in Ecology, Environment and Conservation, Department of Statistical Sciences, University of Cape Town, Rondebosch 7701, Cape Town, South Africa
Full list of author information is available at the end of the article

Photopoulou *et al. Animal Biotelemetry* (2015) 3:21

Page 2 of 10
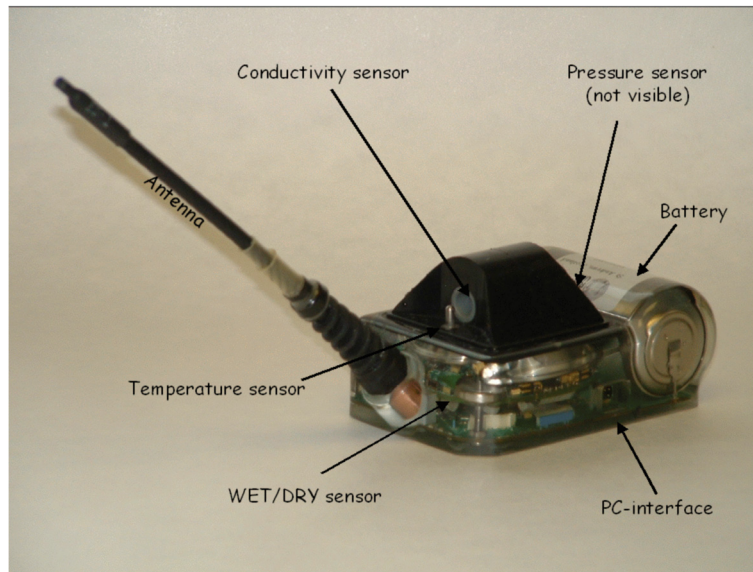
Fig. 1 Photograph of a CTD-SRDL, with visible hardware components labelled (photograph by Lars Boehme, SMRU). The tag is housed in normal solid epoxy rated either to 500 m or 2000 m depth. Standard sensors include pressure, a wet/dry saltwater switch, temperature, conductivity. The tag has a PC interface, is powered by a primary cell (battery) and the standard version includes an antenna

and by CLS-Argos technology. These constraints necessitate the use of efficient algorithms to produce small parcels of information that have a chance of being received by the ground station once they are transmitted, such as data summaries and abstracted data [20–22]. Ultimately, the objective is to achieve the most effective representation of animal behaviour and the *in-situ* environment, given the constraints on data processing, data storage, and data transmission. A thorough appreciation of the way they are collected and the algorithms by which they are compressed [20, 23] is critical for the meaningful scientific interpretation of the data that are finally received.

CTD-SRDLs are particularly relevant for studies in which long-term datasets are required, and in cases where archival tags can only infrequently be recovered. These instruments have been deployed on a wide range of animals and have been particularly successful in providing oceanographic information from under-sampled locations, such as remote, ice-covered or near-shore regions in polar oceans [19]. The oceanographic data produced by these tags is widely used and are increasingly being made freely available [24]. For example, there is now a web portal [25] for accessing much of the data collected by these tags in a consistent format and standardized accuracy [10]. Here, we document the general structure of the protocol for data collection by CTD-SRDLs and, where useful, give example parameter values. We present a condensed overview of the types of data that are collected, in Table 1. We present schematic representations of the relationships between the different behavioural states (Fig. 3), the protocol for collection of

Table 1 Data collection and transmission summaries (CTD_GEN_07B)

Note: All values that can be changed to accommodate individual experimental needs are enclosed in [square brackets].

Behavioural state definitions

Enter "hauled out" if dry for [10] min, exit if wet for [40] sec.

Enter "diving" if wet and depth is greater than [6] m for [8] sec, exit if depth is less than [6] m for any length of time ([0] sec), or dry at any time.

"At the surface" is the complement of "hauled out" and "diving". If this state persists, and there is no dive for [9] min, it is called "cruising" and is treated as a separate state. Entry and exit information is stored and transmitted.

Data types

i. Dive record: [30 k] bytes are made available for dives, which are stored in groups of [3]. If we assume average dive duration of [10] min this results in a [6.25] day buffer with capacity for [300] dives.

ii. Haulout and cruise record: [480] bytes each are available for haulouts and cruises. Since only the entry and exit times are stored for haulouts and cruises this means that [30] haulouts and [30] cruises can be stored in the buffer.

iii. Summary record: [680] bytes are made available for [6] hour summaries, which are stored in groups of [3]. This results in a [7.5] day buffer.

iv. CTD or temperature cast record: [5 000] bytes are available for CTD casts which results in a [7.5] day buffer.

Photopoulou *et al. Animal Biotelemetry* (2015) 3:21

Page 3 of 10

CTD profiles (Fig. 4), and an overview of the possible outcomes as one parcel of information passes through the algorithm (Fig. 5).

Review
Aim
In this section we provide a detailed description of the order and hierarchy of events that take place on-board a CTD-SRDL when collecting, processing and sending data, so as to highlight the strengths, limitations and potential applications of the data returned by the instruments. The CTD-SRDL software specification CTD_GEN_07B, developed in 2007 for deployments on southern elephant seals *Mirounga leonina* is used here as a working example to illustrate the data collection, abstraction and compression routines carried out on-board a CTD-SRDL. Variants of this basic program have been developed since then, but this represents the most generally used protocol for CTD-SRDLs.

To illustrate the problem the protocol tries to solve, consider this challenge: about 100 kB of information are collected each day at 1-4 sec intervals. No more than 1 kB is likely to be received by the ground station. Therefore, data on the parameters of interest need to be fitted into several messages, of fixed size, that are not interdependent. These messages are required by the CLS-Argos to be of fixed size and no acknowledgement of safe receipt is returned to the CTD-SRDL. Although we present here a generalized version of the software protocol, it is one that illustrates the principles, strategies and trade-offs that form the common backbone of the software.

Hardware components
The tag consists of a set of sensors, a microcontroller (programmed mainly in C), a real-time clock, and data storage, linked to a transmitter and antenna to send the information to a satellite. A primary lithium cell to provides operating energy (Fig. 1). The microcontroller manages the overall operation of the tag including the sensors' operation and duty cycle, collection of data, and data compression, selection and analysis. Monitoring all of these is critical to making the best use of its battery. The standard sensors measure pressure, the temperature and conductivity of seawater. A wet-dry sensor acts as a surface detector to establish when to attempt to transmit [22] (Fig. 1). Variants have included other sensors for measuring environmental variables, such as swimming speed, fluorescence, and oxygen.

Software systems
The software design for CTD-SRDLs strives to overcome a set of constraints imposed primarily by three components of the system: 1) the hardware: constrained by size and limited energy of the battery; 2) the communication system: constrained by limited availability and bandwidth of the satellite system used to relay data to the ground station; and 3) the behaviour of marine animals: adds constraints to other elements by brief and infrequent surfacing, together with extended periods at sea. Details of the interaction of these three factors can be found in Table 2 (also Fig. 2). Aspects of the collection and compression routines for the behavioural and hydrographic data have already been described in [21, 23] and [22], respectively. These are presented succinctly in Tables 3 and 4.

Protocol overview and rationale
The tag can be activated prior to deployment by simulating submergence (allowing a current to flow between the two copper lugs in the front of the tag). Otherwise, the tag is activated automatically after deployment as soon as the wet-dry sensor is submerged, and does not switch off until it runs out of energy. The basic sampling regime generally occurs at a 4 sec interval and consumes little energy, but some sensors, such as conductivity, and depth to a certain extent, cannot be operated continuously and are switched on when needed according to specific parameters (Fig. 4, Additional file 1: Supplementary material S4).
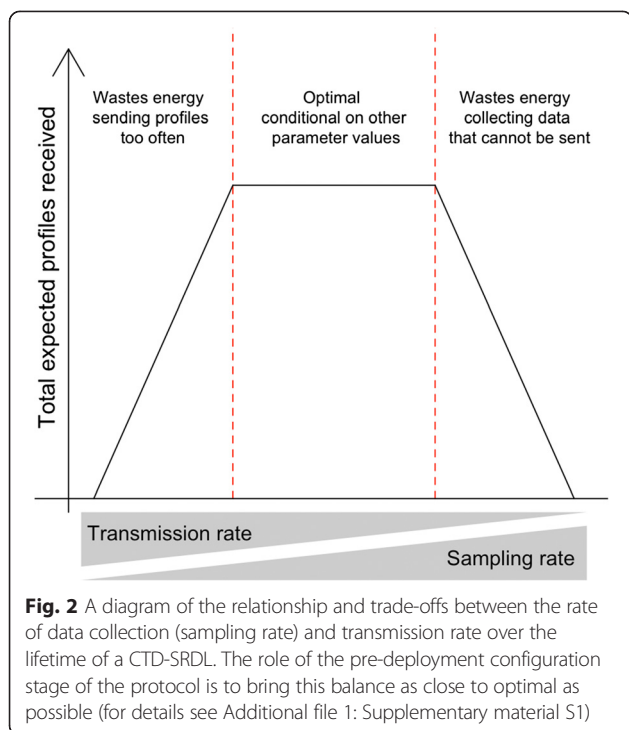
The protocol overview is as follows. First, temperature, pressure and wet/dry sensors are sampled at maximum temporal resolution (i.e., every 4 sec, or every 0.5 sec if within 6 m of the surface), and at maximum precision. Functional models of behaviour are used to structure

Table 2 Data relay

The tags are certified to communicate with six polar orbiting satellites operated by Argos [27]. Communications between the tag and the satellites are subject to strict rules. Information is sent in units, or messages (termed "uplinks"), which may be up to 960 msec in duration and require a minimum of 40 sec between attempted transmissions. Two or more complete uplinks are required for the system to compute a location. Uplinks have a fixed length: 28 bits are used to identify the instrument. A further 248 bits are available to represent the behavioural and hydrographic data.

The combination of this transmission regime, the brief and infrequent surfacing behaviour of air-breathing marine animals, the fact that satellites will not always be available, the lack of acknowledgement of receipt of a message and the tags' energy constraints all demand complex data collection software, extreme data compression and transmission strategies that maximise the information that is sent using a small amount of energy. As a result, the software and processing routines implemented on-board CTD-SRDLs have been developed over many years to try to maximise the amount of biologically and hydrographically relevant information that can be sent, while using the lowest possible bandwidth.

The critical factor in terms of effective use of energy is the balance between the rate of data collection and the transmission of the data (Fig. 2). Routine sampling of the instrument's sensors requires little energy. It should be noted however, that even when longevity of the tag is forfeited for higher resolution data, there is still an upper limit to the throughput and rate of information transfer this tag is capable of.

Photopoulou *et al. Animal Biotelemetry* (2015) 3:21

Page 4 of 10

Fig. 2 A diagram of the relationship and trade-offs between the rate of data collection (sampling rate) and transmission rate over the lifetime of a CTD-SRDL. The role of the pre-deployment configuration stage of the protocol is to bring this balance as close to optimal as possible (for details see Additional file 1: Supplementary material S1)

the data in a biologically intuitive way. Standard behavioural categories are "diving", "hauled out" and "at the surface" (Table 1 and Fig. 3 for overview, Additional file 1: Supplementary material S2 for details). Data selection and compaction techniques are applied to each data type to reduce the physical memory required to store the data [23] (Table 3). A memory store of appropriate size is created, according to the deployment's data collection priorities, to hold the data for later transmission. The tag continually

Table 3 Data abstraction with the broken-stick model

The broken-stick model (BSM) [26, 28] is a piecewise linear approximation method, used here to represent time-depth dive and hydrographc profiles with a smaller number of points than their full resolution. This has been used as the abstraction algorithm on-board SRDLs for dive data and hydrographic data on CLS-Argos tags from 2007 onwards.

The number of iterations of the algorithm that are carried out on board the tag can be adjusted. The principle of the algorithm is to use line segments to approximate a shape, adding points to the piece-wise linear representation where it differs most from the original shape.

In the case of dive profiles, the differences between the BSM dive profile and the original profile recorded by the tag are measured vertically, i.e. metres. At each iteration of the algorithm, the point in the original profile with the biggest vertical distance from the abstracted profile is identified and added to the abstracted profile. This process is repeated until the desired number of points is reached [23]. The depth points chosen by the BSM are stored at the tag's internal resolution of 10 cm and the time resolution (4 sec or multiples) is known implicitly since it is linked to dive duration.

For CTD upcasts, which consist of fixed points as well as points chosen by the BSM, the first iteration of the algorithm is taken as the piecewise linear profile made up of the deepest point, the surface point and the intermediate fixed depth points.

Table 4 Data collection by hydrographic sensors

For a summary of CTD data collection please see the Supplementary material: Temperature and salinity cast records.

i. Pressure (Depth) and Temperature

The nominal accuracy of the pressure transducer is approximately 20 m for the entire range of temperature and pressure it is designed to sample (from 0 m to 2000 m and from -40 to +80 °C). In practice the temperature range is much less than these limits, and it has been found that the error in surface pressure following a deep dive can be about 0.5-1.5 m rather than zero. To minimize the pressure measurement error as much as possible, the instrument updates the zero-offset, whenever the wet/dry sensor reads "dry". The exact value of the zero-offset will vary according to the hydrographic conditions, which is why the instrument is programmed to update this value whenever the sensor reads "dry". The pressure reading is deliberately shifted away from zero to avoid interference from electrical noise generated by the tag.

Depth within a dive is calculated as follows: 1) pressure is collected, 2) the zero-offset is subtracted from the pressure measurement, 3) the resulting value is converted to depth, and 4) pressure is rounded to the nearest dbar (or nearest 2 for deep diving species) for the purpose of constructing a temperature vs depth or salinity vs depth profile. The resulting depth measurements are processed by the instrument at its internal resolution of 10 cm for behavioural samples such as the dive shape buffer.

Upcasts performed by the current generation of CTD-SRDLs yield 17 values each for temperature and conductivity. Temperature is collected at a resolution of 0.001 K within a valid range of -3 to 40 °C (see Supplementary material: Temperature and salinity cast records).

The minimum and maximum temperature within an upcast are sent with high precision, and each of the 17 temperature values is expressed as a percentage of the range (0 = Tmin, 100 = Tmax), in units of 1/1000. The resolution of the temperature measurements will be 0.001 K if the temperature range of the profile is less than 1 K, or (Tmax-Tmin)/1000 if the range is greater than 1 K. For CTD-SRDLs accuracy of the temperature measurement is at least 0.01 K.

ii. Salinity

Salinity is treated in the same way as temperature (see Supplementary material: Temperature and salinity cast records). Conductivity and temperature are used to calculate salinity, then conductivity is discarded. Salinity is stored at a resolution of 0.001 within a valid range of 10 to 50 psu. Only minimum and maximum salinity within an upcast are sent and each of the 17 salinity values is expressed as a percentage of the range (0 = Smin, 100 = Smax), in units of 1/250. Minimum and maximum salinities are sent with resolution 0.02 psu, which will limit the resolution of the salinity measurements when the salinity range of the profile is less than 5 psu. If the range is greater than 5 psu the resolution is (Smax-Smin)/250.

Additional information on the post processed accuracy of the CTD data can be found in [10] and at [25].

monitors the timing (when) and frequency (how many attempts) of data transmission to control energy costs and reach its targeted lifetime. Due to unpredictability in the behaviour of the tagged animals, transmission is to some degree opportunistic. A number of transmission targets can be set during pre-deployment configuration. The tag monitors progress towards the next target and adjusts an hourly transmission quota accordingly (Table 2).

Data are organised into separate data templates called "pages", each of which must fit into the data space allowed for a single transmission. The way the information

Photopoulou *et al. Animal Biotelemetry* (2015) 3:21

Page 5 of 10
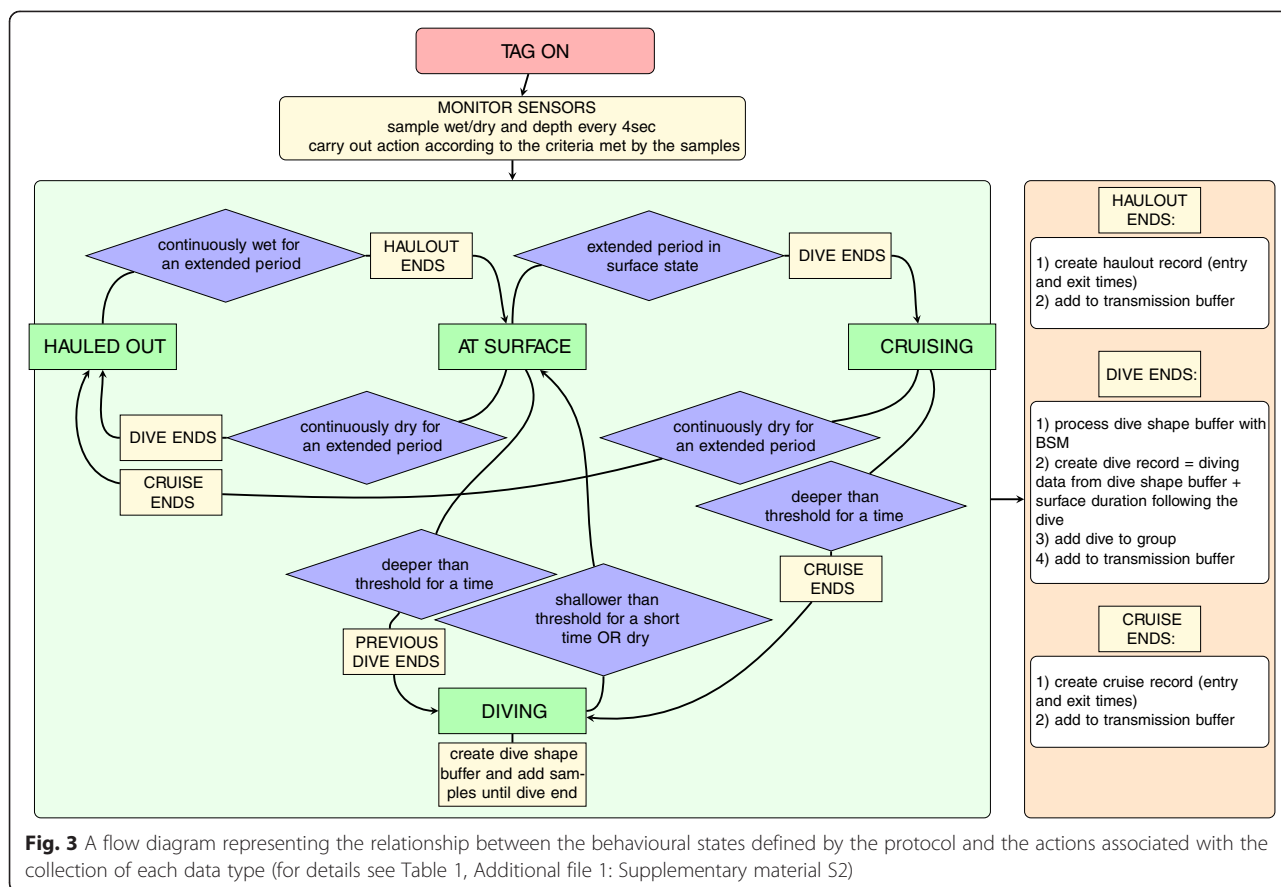
Fig. 3 A flow diagram representing the relationship between the behavioural states defined by the protocol and the actions associated with the collection of each data type (for details see Table 1, Additional file 1: Supplementary material S2)

is organised on each page is flexible and is tailored to user requirements. Specific parameter values are discussed in the Additional file 1: Supplementary material. Throughout the operation of the instrument, calculations are performed on the information received from the sensors to summarize data using a predefined structure for each type (Table 1 and Additional file 1: Supplementary material). The lifetime of the tag is largely determined by the use of these sensors and by the number of attempted transmissions.

Pre-deployment software configuration
Careful consideration of the data collection priorities prior to tag deployment is vital for getting the most useful information out of each deployment. Different studies may address a wide variety of questions from a broad range of species, so the controlling software must be easily adapted to change the balance of data sent and the rate of transmission. This need for flexibility demands complex data collection software. Sensor activation, sampling rates and duty cycles are controlled by parameters that can be modified by fixed or dynamic schedules. Program specifications for each deployment are adjusted specifically in accordance with user requirements based on the data collection priorities, the

behaviour of the target species, the questions of interest, and the range of hydrographic conditions the animal is expected to encounter. Only one data type can be maximised in each deployment. The amount of all other information likely to be received is then adjusted according to the constraints put in place by having to maximise the chosen data type.

Several steps are involved in balancing the data collection priorities with what is feasible, given the species and the likely conditions the animal platform will encounter. First, the number of events of each type (CTD record, dive record, haulout record, cruise record, behavioural data summary record, Table 1, Figs. 3, 4 and 5) that will occur over the deployment is estimated according to the biology of the study species. The probability of receiving any one transmission varies geographically, decreasing with proximity to competing transmitters and, because CLS-Argos satellites are polar orbiting, distance from the poles. Reception probability is estimated for each deployment based on its expected geographical location and extent. To configure the software to collect the desired data, the user needs to specify the minimum proportion of events (e.g. 10 % of dives, 90 % of summaries) that it is acceptable to receive during the deployment. This sets a lower bound on the resolution at which

Photopoulou *et al. Animal Biotelemetry* (2015) 3:21

Page 6 of 10
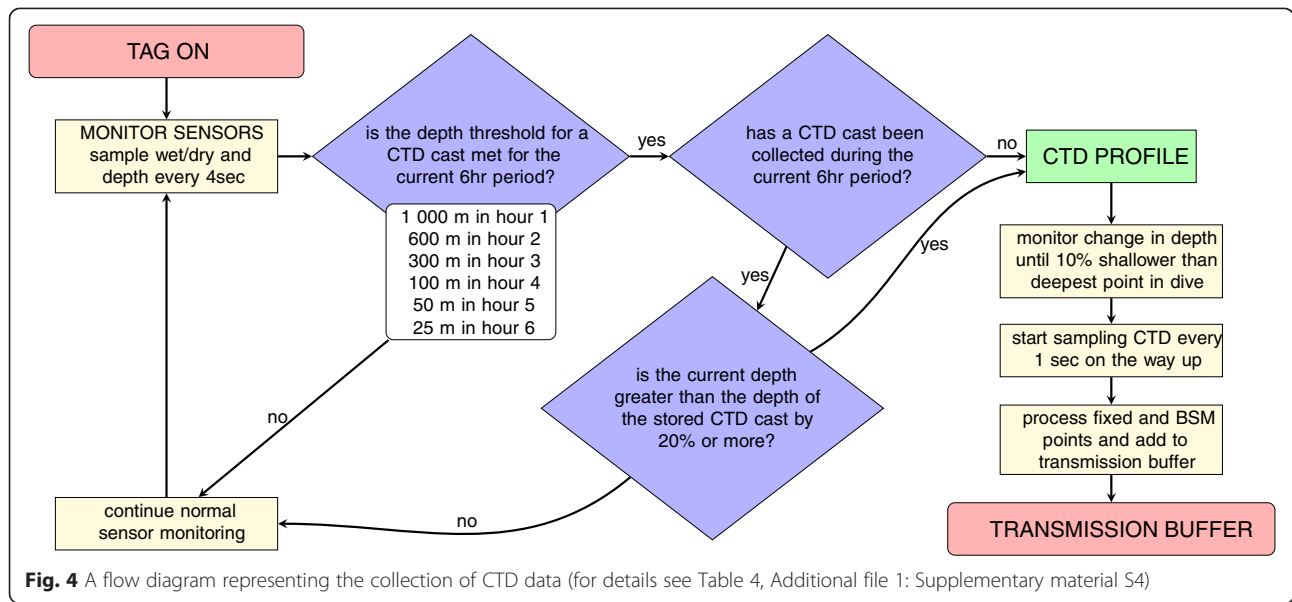
Fig. 4 A flow diagram representing the collection of CTD data (for details see Table 4, Additional file 1: Supplementary material S4)

different data types will be collected. It is then possible to decide which quantity to maximise (e.g. number of CTD profiles expected to be received). Next, template pages are created for the dive, summary, and CTD data types, specifying the number of bits needed to achieve the required resolution. The required resolution of each data type is based on the study's priorities, which dictate the proportion

of all transmissions that should be devoted to each data type. Each page may contain several data items within the maximum of 247 bits available for a single transmission. The proportion of transmissions assigned to each template and the CTD sampling rate are modified such that the target quantity is maximised and other data types each achieve the minimum number of receptions. Note that changing
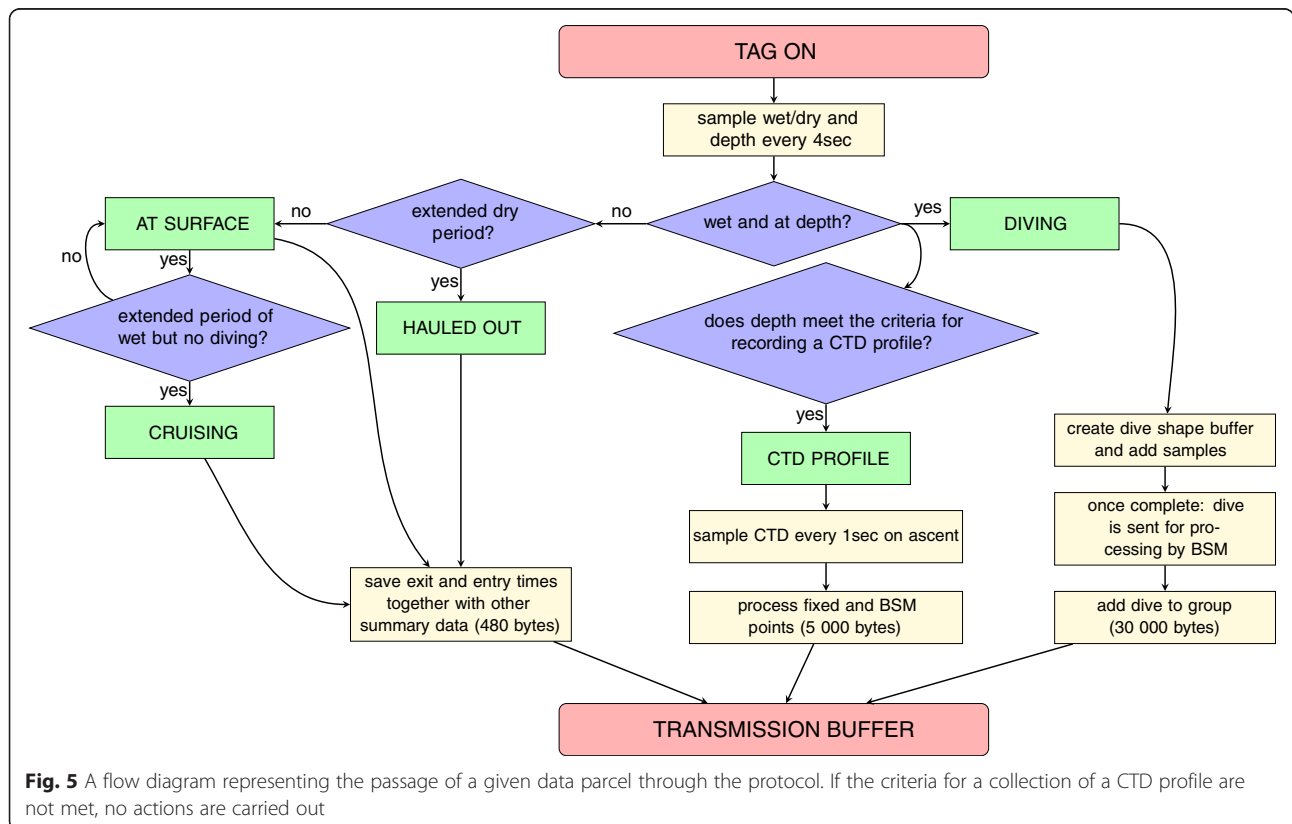
Fig. 5 A flow diagram representing the passage of a given data parcel through the protocol. If the criteria for a collection of a CTD profile are not met, no actions are carried out

Photopoulou *et al. Animal Biotelemetry* (2015) 3:21

Page 7 of 10

the CTD sampling rate affects the energy available for transmissions and hence the expected number of data items received. The sampling rate is tuned until an optimal balance between sampling and transmission is achieved.

Sometimes there is no satisfactory solution to the trade-off between the CTD sensors' sampling rate and the expected number of data items received, e.g. if battery capacity is inadequate (for long deployments), or the additive effect of the animal's surfacing behaviour and CLS-Argos transmission frequency restrictions do not allow enough transmissions (for short deployments). In these cases, the constraint on the number of receptions can be relaxed to reduce the required battery power, or two CLS-Argos transmitter id's can be used to increase transmission frequency.

An important assumption is made with regard to data relay by CTD-SRDLs; namely that all pages are equally likely to be transmitted and received and that there is no spatial or temporal bias in the type of data that are transmitted and subsequently received. Spatial and temporal bias can arise when one type of behaviour is carried out intensively (for example, if a long bout of diving is common prior to a long surface period), and when satellite coverage is particularly good or bad, in one area or for a time. Data are held in memory "buffers" to try to make this assumption approximately valid (Tables 2 and 5, Additional file 1: Supplementary material S1). The use of memory buffers makes the received data more representative of behaviour across the whole range of spatial and temporal locations visited during a deployment.

Software procedures during deployment

During operation, the information received from the tag's sensors triggers processes that extract and summarize data appropriately for each data type (Table 1, Fig. 3 and Additional file 1: Supplementary material S2, S3 and S4), and regulate the transmission scheduling of the tag, so as to take account of remaining battery power (Table 2). The behavioural model, parameterised appropriately for the study species (for parameter values routinely used for elephant seals see Table 1, Fig. 3; also Additional file 1: Supplementary material S2), is used to partition behaviour into three mutually exclusive and exhaustive states of haulout, dive, and surfacing behaviour. Behaviour is summarised over blocks of time, typically 4 or 6 hrs, to allow diurnal patterns to emerge. The summary statistics for each time block represent the entirety of the data recorded during that period, not a statistical sample. A subset of dive records is transmitted with the summaries, providing a more detailed representation of behaviour in relation to environmental information, within the context of the summary information. The received dive records are a subset because not all records that are collected and transmitted are received. It is

Table 5 Dive records

Dive records resulting from most telemetry tags using the CLS-Argos system will almost always be incomplete, i.e. have some missing dives, locations and CTD upcasts. This is exacerbated by the lack of acknowledgement when transmitted information is received, or "handshake" between the platform and the satellite. This lack of handshake creates the need for a stochastic transmission strategy of the information accumulated in the transmission buffer.

The proportion of dive data that is lost will depend on a number of factors, including the geographic region in which the instrumented animal operates (e.g. polar regions have better coverage by Argos satellites than equatorial regions because they are polar orbiting), the animal's diving ecology, how long the animal spends at the surface where transmission is possible, and the frequency with which that transmission window occurs.

For elephant seals, there are two main reasons why data are lost. Firstly, as for all species, there will be periods during deployment when conditions for satellite transmission are poor, and secondly, there may be periods when data accumulate at a higher rate than the tag can transmit them, for example periods of high diving activity that exceed the size (in terms of time) of the transmission buffer. Southern elephant seals are a species that spend very little time at the surface (mean surface time 2.18 min ± 2.74 standard deviation, range 0.05-9.48 min, n = 187 832 dives, SMRU-IG unpublished data), providing a narrow window for transmission. The consequences of elephant seal surfacing behaviour on the number of messages sent and the quality of the locations are shown in Fig. 3 of [22]. The transmission buffer in CTD-SRDLs designed for deployment on southern elephant seals is made large enough to hold many days worth of dives, to minimize the proportion of information lost during low Argos coverage, and to avoid data return being biased in favour of data collected during high Argos coverage (Table 2). In addition, as mentioned above, stored data are selected randomly from the transmission buffer. Information remains in the transmission buffer either until its "sell-by" date expires or it becomes displaced by newer information being added to the buffer.

assumed that the dives in the transmission buffer provide an unbiased sample of dives, from which dives are missing at random (Table 5, Additional file 1: Supplementary material S3).

When dives and hydrographic profiles are collected the raw data are submitted for further processing using a broken-stick model (BSM) [23, 26] (Table 3, Fig. 3). For dives, four iterations of the algorithm are carried out resulting in six times and six depths, including two surface points at the beginning and the end of the dive (Table 5, Additional file 1: Supplementary material S3). For hydrographic data, oceanographic requirements can, in some cases, determine specific depths at which temperature and salinity points must be collected, otherwise data are processed by BSM the same way as XBT casts [22, 26] (Additional file 1: Supplementary material S4). All numeric samples are coded as a proportion of each variable's range and stored together with the minimum and maximum values. Digitising information in this way reduces the memory required to store the data records and allows them to be fitted as closely as possible into the prescribed message size for transmission.

After compression, related data types are grouped into pages according to the page templates (e.g., dives are stored in groups of three) and stored in the transmission

Photopoulou *et al. Animal Biotelemetry* (2015) 3:21

Page 8 of 10

buffer (Additional file 1: Supplementary material S1). In addition, nine bits of error checking code are interleaved in each page to allow most errors in received transmission to be detected. Each page must fit into the bit limits (239 + 9 = 248 bits) of a single CLS-Argos message. Pages of data are drawn from the buffer for transmission in proportion to the desired likelihood of reception specified in the pre-deployment configuration. This is necessary because there is no two-way communication to indicate whether a given page of data has been successfully received by the satellite.

Avoiding transmission bias

An additional technique is used to avoid transmissions being biased towards periods of better satellite coverage or changes in the animal's behaviour that favour transmission. Typically, five different pages are produced and stored in the transmission buffer, ready for transmission by CLS-Argos. Two types of page (labelled here: 1a and 1b) contain detailed information about dives, one page contains the summary data (2) including behavioural information regarding any haulouts and cruises for one summary period, and the two last pages each contain part of a CTD profile (3 and 4) because the hydrographic data are spread over two separate messages (fixed points and broken-stick points, when applicable). Pages are selected for transmission in a prescribed sequence according to their type (for example: 1a 1b 2 3 4 3 4) in order to achieve the correct proportion of each type of information as described above. The sequence of selection repeats indefinitely, cycling through the information in the transmission buffer.

When it is time for a transmission, the page template is applied to the next unexpired item in the appropriate circular buffer. Sending pages more often increases their probability of reception. An item may remain in the buffer for many of these cycles before being discarded when it reaches its "sell-by" date (e.g. each summary message is transmitted 40 times on average to achieve a 95 % probability of being received, for more detail see Additional file 1: Supplementary material S1). The independence of the insertion process (as the data are collected and processed) and the extraction process (as the data are selected from the buffer and transmitted) ensures that the likelihood of data being received is not biased by the animal's behaviour at the time of transmissions. The effect is that of randomizing the information in the buffer so that each data item is equally likely to be received.

Conclusions

The CTD-SRDL is an animal telemetry device whose hardware and software have been developed over the course of more than two decades. Its soft- and hardware systems are purpose-built for the remote collection of data on the biology of marine animals and their environment, and this is the first time they have been holistically reviewed. The evolution of the instrument is on-going, driven both by biological and oceanographic questions. The field of animal telemetry has expanded dramatically since the first prototype predecessor of a CTD-SRDL was deployed and transmission technologies have diversified substantially. Nonetheless, the CTD-SRDL is still the most commonly used instrument for collecting information about widely ranging species in high latitude areas. Almost 500 000 CTD and 5 million dive profiles respectively, have been collected to date by CTD-SRDLs since 2003 (SMRU-IG unpublished data, June 2015). These CTD profiles are not only numerous but they are also of global importance. They have contributed substantially to the World Ocean Database [29] by providing hydrographic information from areas never directly surveyed [16, 24] and now comprise 70 % of all oceanographic profiles ever collected south of 60 °S, with increasing contribution in the Arctic [19]. At the poles, the CLS-Argos system continues to be the most effective satellite transmission platform for animal data, particularly when a) the animals do not spend long enough periods at the surface or hauled out to allow the use of the Iridium communications system, or b) when they do not haul out in areas that allow the use of other transmission technologies such as GSM. The species that have been instrumented with CTD-SRDLs most often include southern and northern elephant seals (*M. leonina* and *M. angustirostris*), Weddell seals (*Leptonychotes weddellii*), grey seals (*Halichoerus grypus*), hooded seals (*Cystophora cristata*) and several species of sea turtle.

The outlook on future development of CTD-SRDLs, similarly to other animal telemetry devices, includes reducing the size and increasing the battery life of the device. Neither of these is a trivial challenge. Reduction in size is severely limited by the size of the conductivity and temperature sensors, while making the tag more energy efficient depends on how much energy the sensors are using as well as the overall energy availability provided by the battery. The addition of novel sensors, such as oxygen and pH, which would deliver highly desirable and previously unobserved datasets, is most often at odds with reducing the size and energy requirements of the device. These sorts of developments are constantly being researched but involve implicit challenges imposed by the existing framework of limited memory and energy as well as stringent bandwidth constraints.

The trade-off between precision and bias, highlighted above, is equally relevant here as it is in other applications of optimisation. Maximising timing and frequency of data transmission is important, but not if the most informative bits of data (i.e., precision) are lost in the process. The key objective for this protocol is to collect

Photopoulou *et al. Animal Biotelemetry* (2015) 3:21

Page 9 of 10

the most relevant data while dealing with the sometimes strict constraints of technology and animal behaviour. The dynamic modification of the transmission rates on-board telemetry devices, either due to extrinsic or intrinsic factors affecting transmission could result in bias. The use of memory buffers, in the case of CTD-SRDLs, is intended to ameliorate this but may, as a consequence, mask the degree of intensity of behaviours, which is in turn important information regarding the animal platform. Due to multifaceted objectives and trade-offs, the usefulness of any device for the remote collection of animal observations will only be as good as the overlap between the capabilities of the device, the specific behaviour of the animals and the scientific questions sought to be answered using the resulting data. We hope that this review will promote appropriate usage of both historical CTD-SRDL data and data that continue to be collected, as well as contribute to the on-going development of animal telemetry devices.

Additional file

Additional file 1: Supplementary material.

Author details
[1]Centre for Statistics in Ecology, Environment and Conservation, Department of Statistical Sciences, University of Cape Town, Rondebosch 7701, Cape Town, South Africa. [2]Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews, Scotland KY16 8LB, UK. [3]Institute of Biodiversity, Animal Health and Comparative Medicine, Graham Kerr Building University of Glasgow, Glasgow, Scotland G12 8QQ, UK.

References
1. McConnell B, Chambers C, Fedak MA. Foraging ecology of southern elephant seals in relation to the bathymetry and productivity of the Southern Ocean. Antarct Sci. 1992;4:393–8.
2. Lydersen C, Nøst O, Lovell P, McConnell B, Gammelsrød T, Hunter C, et al. Salinity and temperature structure of a freezing Arctic fjord—monitored by white whales (Delphinapterus leucas). Geophys Res Lett. 2002;29:2119.
3. Hooker SK, Boyd IL. Salinity sensors on seals: use of marine predators to carry CTD data loggers. Deep Sea Res Part I Oceanogr Res Pap. 2003;50:927–39.
4. Biuw M, Boehme L, Guinet C, Hindell M, Costa D, Charrassin J-B, et al. Variations in behavior and condition of a Southern Ocean top predator in relation to in situ oceanographic conditions. Proc Natl Acad Sci U S A. 2007;104:13705–10.
5. Charrassin J-B, Hindell M, Rintoul SR, Roquet F, Sokolov S, Biuw M, et al. Southern Ocean frontal structure and sea-ice formation rates revealed by elephant seals. Proc Natl Acad Sci U S A. 2008;105:11634–9.
6. Fedak M. Marine animals as platforms for oceanographic sampling: a " win/ win " situation for biology and operational oceanography. Mem Natl Inst Polar Res Spec Issue. 2004;133–147.
7. Biuw M, Nøst O, Stien A, Zhou Q, Lydersen C, Kovacs K. Effects of hydrographic variability on the spatial, seasonal and diel diving patterns of southern elephant seals in the eastern Weddell Sea. PLoS One. 2010;5:1–14.
8. Heerah K, Andrews-Goff V, Williams G, Sultan E, Hindell M, Patterson T, et al. Ecology of Weddell seals during winter: Influence of environmental parameters on their foraging behaviour. Deep Sea Res Part II Top Stud Oceanogr. 2013;88–89:23–33.
9. Padman L, Costa DP, Bolmer ST, Goebel ME, Huckstadt L, Jenkins A, et al. Seals map bathymetry of the Antarctic continental shelf. Geophys Res Lett. 2010;37:1–5.
10. Roquet F, Charrassin J-B, Marchand S, Boehme L, Fedak M, Reverdin G, et al. Delayed-Mode Calibration of Hydrographic Data Obtained from Animal-Borne Satellite Relay Data Loggers. J Atmos Ocean Technol. 2011;28:787–801.
11. Meredith MP, Nicholls KW, Renfrew IA, Boehme L, Biuw M, Fedak M. Seasonal evolution of the upper-ocean adjacent to the South Orkney Islands, Southern Ocean: Results from a "lazy biological mooring.". Deep Sea Res Part II Top Stud Oceanogr. 2011;58:1569–79.
12. Arthun M, Nicholls KW, Makinson K, Fedak MA, Boehme L. Seasonal inflow of warm water onto the southern Weddell Sea continental shelf, Antarctica. Geophys Res Lett. 2012;39:n/a.
13. Nicholls KW, Boehme L, Biuw M, Fedak MA. Wintertime ocean conditions over the southern Weddell Sea continental shelf, Antarctica. Geophys Res Lett. 2008;35, L21605.
14. Costa D, Klinck JM, Hofmann EE, Dinniman MS, Burns JM. Upper ocean variability in west Antarctic Peninsula continental shelf waters as measured using instrumented seals. Deep Sea Res Part II Top Stud Oceanogr. 2008;55:323–37.
15. Århun M, Nicholls KW, Boehme L. Wintertime Water Mass Modification near an Antarctic Ice Front. J Phys Oceanogr. 2013;43:359–65.
16. Roquet F, Wunsch C, Forget G, Heimbach P, Guinet C, Reverdin G, et al. Estimates of the Southern Ocean general circulation improved by animal-borne instruments. Geophys Res Lett. 2013;40:6176–80.
17. Boehme L, Meredith M, Thorpe SE, Biuw M, Fedak MA. Antarctic Circumpolar Current frontal system in the South Atlantic: Monitoring using merged Argo and animal-borne sensor data. J Geophys Res. 2008;113, C09012.
18. Boehme L, Thorpe SE, Biuw M, Fedak MA, Meredith M. Monitoring Drake Passage with elephant seals: Frontal structures and snapshots of transport. Limnol Oceanogr. 2008;53:2350–60.
19. Fedak MA. The impact of animal platforms on polar ocean observation. Deep-Sea Res II Top Stud Oceanogr. 2013;7–13.
20. Fedak MA, Lovell P, Grant S. Two approaches to compressing and interpreting time-depth information as collected by time-depth recorders and satellite-linked data recorders. Mar Mammal Sci. 2001;17:94–110.
21. Fedak MA, Lovell P, McConnell B, Hunter C. Overcoming the constraints of long range radio telemetry from animals: getting more useful data from smaller packages. Integr Comp Biol. 2002;42:3–10.
22. Boehme L, Lovell P, Biuw M, Roquet F, Nicholson J, Thorpe SE, et al. Technical Note: Animal-borne CTD-Satellite Relay Data Loggers for real-time oceanographic data collection. Ocean Sci Discuss. 2009;6:1261–87.

Photopoulou *et al. Animal Biotelemetry* (2015) 3:21

Page 10 of 10

23. Photopoulou T, Lovell P, Fedak MA, Thomas L, Matthiopoulos J. Efficient abstracting of dive profiles using a broken-stick model. Methods Ecol Evol. 2015;6:278–88.

24. Roquet F, Williams G, Hindell MA, Harcourt R, McMahon C, Guinet C, et al. A Southern Indian Ocean database of hydrographic profiles obtained with instrumented elephant seals. Sci Data. 2014;1:140028.

25. Marine Mammals Exploring the Oceans Pole to Pole [http://www.meop.net/]

26. Rual P. For a better XBT bathy-message on board quality control, plus a new data reduction method. In: Picaut J, Lukas R, Delcroix T, editors. Western Pacific International Meeting and Workshop on Toga Coare. New Caledonia: Noumea; 1989. p. 823–33.

27. Argos User's Manual [http://www.argos-system.org/web/en/381-supportand-help.php]

28. Toms J, Lesperance M. Piecewise regression: a tool for identifying ecological thresholds. Ecology. 2003;84:2034–41.

29. World Ocean Database [https://www.nodc.noaa.gov/OC5/WOD/pr_wod.html]