

A standardisation framework for bio-logging data to advance ecological research and conservation

Ana M. M. Sequeira¹  | Malcolm O'Toole¹ | Theresa R. Keates² | Laura H. McDonnell³ | Camrin D. Braun^{4,5} | Xavier Hoenner⁶  | Fabrice R. A. Jaïne^{7,8}  | Ian D. Jonsen⁸ | Peggy Newman⁹  | Jonathan Pye¹⁰ | Steven J. Bograd¹¹ | Graeme C. Hays¹²  | Elliott L. Hazen¹¹  | Melinda Holland¹³ | Vardis M. Tsonos¹⁴  | Clint Blight¹⁵  | Francesca Cagnacci¹⁶ | Sarah C. Davidson^{17,18}  | Holger Dettki¹⁹ | Carlos M. Duarte²⁰ | Daniel C. Dunn²¹ | Victor M. Eguíluz²²  | Michael Fedak¹⁵  | Adrian C. Gleiss²³ | Neil Hammerschlag^{3,24} | Mark A. Hindell²⁵  | Kim Holland²⁶ | Ivica Janekovic²⁷  | Megan K. McKinzie^{28,29} | Mônica M. C. Muelbert^{25,30}  | Chari Pattiaratchi²⁷ | Christian Rutz³¹  | David W. Sims^{32,33,34}  | Samantha E. Simmons³⁵  | Brendal Townsend¹⁰ | Frederick Whoriskey¹⁰ | Bill Woodward²⁹ | Daniel P. Costa³⁶ | Michelle R. Heupel³⁷ | Clive R. McMahon^{7,25}  | Rob Harcourt⁸  | Michael Weise³⁸

¹Oceans Institute and School of Biological Sciences, University of Western Australia, Crawley, WA, Australia; ²Department of Ocean Sciences, University of California Santa Cruz, Santa Cruz, CA, USA; ³Leonard and Jayne Abess Center for Ecosystem Science and Policy, University of Miami, Coral Gables, FL, USA; ⁴School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA, USA; ⁵Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, USA; ⁶CSIRO Oceans and Atmosphere, Hobart, TAS, Australia; ⁷Integrated Marine Observing System (IMOS) Animal Tracking Facility, Sydney Institute of Marine Science, Mosman, NSW, Australia; ⁸Department of Biological Sciences, Macquarie University, Sydney, NSW, Australia; ⁹Atlas of Living Australia, Melbourne Museum, Carlton, VIC, Australia; ¹⁰Ocean Tracking Network, Dalhousie University, Halifax, NS, Canada; ¹¹NOAA Environmental Research Division, Southwest Fisheries Science Center, Monterey, CA, USA; ¹²School of Life and Environmental Sciences, Deakin University, Geelong, VIC, Australia; ¹³Wildlife Computers, Redmond, WA, USA; ¹⁴NASA Jet Propulsion Laboratory, Pasadena, CA, USA; ¹⁵SMRU Instrumentation, Scottish Oceans Institute, St Andrews, UK; ¹⁶Department of Biodiversity and Molecular Ecology, Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, Trento, Italy; ¹⁷Department of Migration, Max Planck Institute of Animal Behavior, Radolfzell, Germany; ¹⁸Centre for the Advanced Study of Collective Behaviour, University of Konstanz, Konstanz, Germany; ¹⁹Swedish University of Agricultural Sciences, SLU Swedish Species Information Centre, Uppsala, Sweden; ²⁰Red Sea Research Centre (RSRC) and Computational Bioscience Research Center (CBRC), King Abdullah University of Science and Technology, Thuwal, Saudi Arabia; ²¹School of Earth and Environmental Sciences, University of Queensland, St Lucia, QLD, Australia; ²²Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), Palma de Mallorca, Spain; ²³Centre for Sustainable Aquatic Ecosystems, Harry Butler Institute, Murdoch University, Murdoch, WA, Australia; ²⁴Rosenstiel School of Marine & Atmospheric Science, Miami, FL, USA; ²⁵Institute for Antarctic and Marine Studies, University of Tasmania, Hobart, TAS, Australia; ²⁶Hawaii Institute of Marine Biology, University of Hawaii, Manoa, HI, USA; ²⁷Oceans Graduate School and the UWA Oceans Institute, The University of Western Australia, Crawley, WA, Australia; ²⁸Monterey Bay Aquarium Research Institute (MBARI), Moss Landing, CA, USA; ²⁹U.S. Animal Telemetry Network (ATN), NOAA Integrated Ocean Observing System, Silver Spring, MD, USA; ³⁰Institute of Marine Science, Federal University of São Paulo (IMar/UNIFESP), Santos, Brazil; ³¹Centre for Biological Diversity, School of Biology, University of St Andrews, St Andrews, UK; ³²Marine Biological Association of the United Kingdom, The Laboratory, Plymouth, UK; ³³Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, Southampton, UK; ³⁴Centre for Biological Sciences, University of Southampton, Southampton, UK; ³⁵U.S. Marine Mammal Commission, Bethesda, MD, USA; ³⁶Institute of Marine Sciences, Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA, USA; ³⁷Integrated Marine Observing System, University of Tasmania, Hobart, TAS, Australia and ³⁸Office of Naval Research, Arlington, VA, USA

This is an open access article under the terms of the Creative Commons Attribution NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society. This article has been contributed to by US Government employees and their work is in the public domain in the USA.

Correspondence

Ana M. M. Sequeira
 Email: ana.sequeira@uwa.edu.au

Funding information

The Pew Charitable Trusts; Office of Naval Research Global, Grant/Award Number: N00014-19-1-2573; Australian Research Council, Grant/Award Number: DP210103091

Handling Editor: Edward Codling

Abstract

1. Bio-logging data obtained by tagging animals are key to addressing global conservation challenges. However, the many thousands of existing bio-logging datasets are not easily discoverable, universally comparable, nor readily accessible through existing repositories and across platforms, slowing down ecological research and effective management. A set of universal standards is needed to ensure discoverability, interoperability and effective translation of bio-logging data into research and management recommendations.
2. We propose a standardisation framework adhering to existing data principles (FAIR: Findable, Accessible, Interoperable and Reusable; and TRUST: Transparency, Responsibility, User focus, Sustainability and Technology) and involving the use of simple templates to create a data flow from manufacturers and researchers to compliant repositories, where automated procedures should be in place to prepare data availability into four standardised levels: (a) decoded raw data, (b) curated data, (c) interpolated data and (d) gridded data. Our framework allows for integration of simple tabular arrays (e.g. csv files) and creation of sharable and interoperable network Common Data Form (netCDF) files containing all the needed information for accuracy-of-use, rightful attribution (ensuring data providers keep ownership through the entire process) and data preservation security.
3. We show the standardisation benefits for all stakeholders involved, and illustrate the application of our framework by focusing on marine animals and by providing examples of the workflow across all data levels, including filled templates and code to process data between levels, as well as templates to prepare netCDF files ready for sharing.
4. Adoption of our framework will facilitate collection of Essential Ocean Variables (EOVs) in support of the Global Ocean Observing System (GOOS) and inter-governmental assessments (e.g. the World Ocean Assessment), and will provide a starting point for broader efforts to establish interoperable bio-logging data formats across all fields in animal ecology.

KEYWORDS

bio-logging template, data accessibility and interoperability, data standards, metadata templates, movement ecology, sensors, telemetry, tracking

1 | INTRODUCTION

Bio-logging is a powerful set of methods that enables the collection of data about animal movement, behaviour, physiology and the physical environment (Hussey et al., 2015; Kays et al., 2015; Rutz & Hays, 2009). The rapid development and use of devices (hereafter 'tags') to collect, store and transmit bio-logging data began following the launch of the Argos satellite data collection and location system in the 1970s (Thums et al., 2018). Over the subsequent 50 years, the use of acoustic telemetry, light-based geolocation, and other forms of data logging and transmission have matured and become standard methods to understand animal distributions, habitat use, and population connectivity. Data are being generated at unprecedented rates, providing opportunities to conduct synthetic studies (Figure 1;

Block et al., 2011; Davidson et al., 2020; Hindell et al., 2020; Queiroz et al., 2019; Sequeira et al., 2018; Tucker et al., 2018) and address conservation challenges, such as those resulting from global environmental change (Brett et al., 2020; Hays et al., 2019; McGowan et al., 2017; Sequeira et al., 2019) as well as from extreme events (e.g. a global pandemic; Bates et al., 2020; Rutz et al., 2020). However, managing these data is challenging. Despite the growing number of collaborative regional and global initiatives launched to compile existing bio-logging data (Harcourt et al., 2019), there are no widely adopted data and metadata standards, and most existing bio-logging data remain undiscoverable and inaccessible. The lack of universal standards for bio-logging datasets hampers progress in ecological research, burdening researchers with technical and administrative hurdles each time data are shared and re-used (Campbell et al., 2016). Problems range

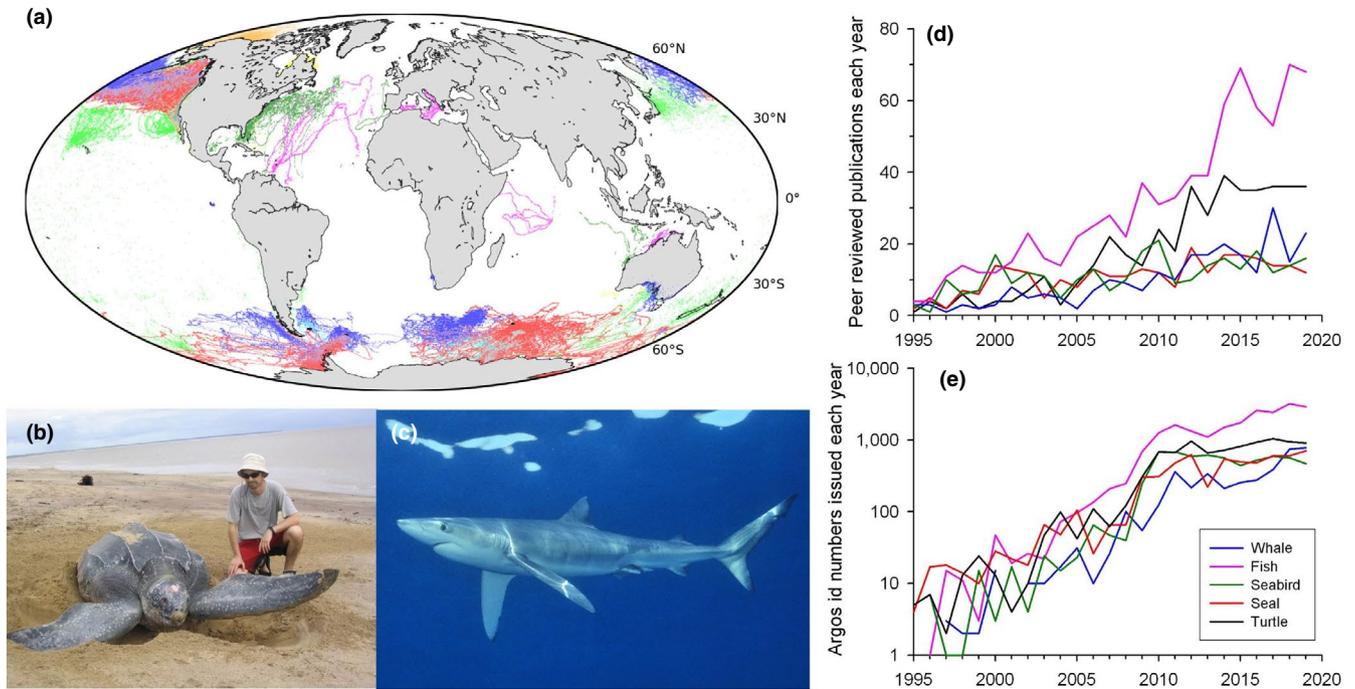


FIGURE 1 Value of synthesising bio-logging data. Efforts to integrate animal tracking results from across multiple studies can deliver fundamental insights into the ecology of diverse species as well as providing important information to help conservation. (a) Tracking results for >2,600 individuals across 50 marine species at global scale have shown similarities in the global movement patterns across taxa linked to habitat. Each colour represents different taxa: blue = seals, pink = sea turtles, light green = sea birds, dark green = sharks; re-drawn from Sequeira et al. (2018) by Dr Jorge Rodríguez. (b) A nesting leatherback turtle *Dermochelys coriacea*. Collated tracking results for adult leatherback turtles from tracking studies across the Atlantic and Pacific have identified overlap hotspots between pelagic longline fishing intensity and turtle foraging and have also revealed how foraging success varies between ocean basins and is linked to reproductive output and conservation status (Bailey et al., 2012; Fossette et al., 2014; Roe et al., 2014). Photo courtesy of Tom Doyle. (c) A blue shark *Prionace glauca*. Tracking thousands of pelagic sharks has revealed high overlap between fisheries and shark space use in the global ocean and has highlighted the importance of marine-protected areas for this group (Queiroz et al., 2019). Photo courtesy of Jeremy Stafford-Deitsch. However, even the largest animal tracking studies still only use a small fraction of the tracking data that have been collected (Hays & Hawkes, 2018). (d) The increase in the annual number of published satellite tracking studies across various taxa. The number of publications each year was obtained from Web of Science using the search terms ‘sea turtle satellite tracking’, ‘seal satellite tracking’, ‘whale satellite tracking’, ‘seabird satellite tracking’ and ‘fish satellite tracking’. The plot conveys the ever-increasing number of published satellite tracking studies. For legend to colours, see panel (e). (e) The number of Argos ids issued each year for satellite tracking studies with different marine taxa. Each satellite tag is programmed with an Argos id number. Although some Argos ids are reused while others may be unused, the number of Argos ids issued will broadly reflect the number of satellite tags deployed. As of May 2020, circa 50,000 Argos ids have been issued for marine animal tracking, including around 10,000 for sea turtle tracking, 4,500 for cetaceans, 6,000 for seabirds, 6,000 for pinnipeds and >20,000 for fish. Data on the number of Argos ids are supplied by CLS (<https://www.argos-system.org/>)

from acute issues with merging disparate datasets, through to the lack of an overarching framework that ensures (a) accuracy-of-use, (b) rightful attribution and ownership, and (c) data preservation security. The latter is especially relevant for older data not currently in use, but potentially invaluable as baseline for future work. Adoption of a framework to standardise bio-logging data will promote efficient data collation, usage and sharing consistent with FAIR (Findable, Accessible, Interoperable and Reusable) (Wilkinson et al., 2019) and TRUST (Transparency, Responsibility, User focus, Sustainability and Technology; Lin et al., 2020) data principles, and enable compliance with requirements of publishers and funding agencies.

Bio-logging is used for a broad range of taxa across terrestrial and marine ecosystems (Hussey et al., 2015; Kays et al., 2015). The high diversity of marine animals, ranging from small seabirds and fishes to the giant blue whale *Balaenoptera musculus*, and with high

mobility in three-dimensional space, has sparked a wide variety of engineering solutions, sensors and approaches to enable attachment of instruments and recovery of data. These include ‘store-on-board’ tags that need to be recovered for data retrieval (Glenn et al., 2009; Watanabe & Sato, 2008), and data-relay technologies (Hussey et al., 2015) for radio-transmitting or pop-up archival tags (Block et al., 1998). The data obtained can range from coarse temporal and spatial resolution (e.g. light-level-based geolocation), to precise location data in space and time (e.g. GPS; Global Positioning System), to very high-resolution pseudo-tracks from daily diary instruments (Wilson et al., 2008). Moreover, marine bio-logging datasets can include concurrent data on horizontal and vertical movements (i.e. in depth; similar to altitude in terrestrial bio-logging data), as well as physical measurements from ancillary sensors (Williams et al., 2020). The latter include detailed oceanographic

conductivity, temperature and depth (CTD) data that can be used to improve the outputs of ocean models (Moore et al., 2011; Roquet et al., 2013). Size constraints specific to CTD packages currently restrict their use to marine megafauna (i.e. the larger marine vertebrates). Such marine megafauna play an important role in collecting relevant data for a range of Essential Ocean Variables (EOVs) (Miloslavich et al., 2018; Muller-Karger et al., 2018), including temperature, salinity, fluorescence (proxy for chlorophyll-*a*) and dissolved oxygen, across a range of ecosystems. These ecosystems range from shallow coastal areas to the deep open ocean, and from the tropics to the poles, including ice-covered areas that are otherwise inaccessible to humans (Harcourt et al., 2019; Moore et al., 2011; Treasure et al., 2017). An example of the latter includes the near real-time temperature and salinity profile data collected by elephant seals that is made freely available daily via the Global Telecommunication System (GTS) of the World Meteorological Organization (wmo.int) for immediate use by weather forecasters and ship operators (Roquet et al., 2014). Marine megafauna are therefore strong candidates to become a key data contributor to the Global Ocean Observing System (GOOS), and indeed recently the GOOS-Steering Committee has endorsed and included AniBOS (Animal Borne Ocean Sensors) as one of its global networks that will provide a cost-effective and complementary observing capability through using animals as 'ocean samplers' (Harcourt et al., 2019). However, successful integration of datasets is strongly dependent on improving data standardisation.

Here, we provide a framework designed to facilitate standardisation of bio-logging data, including three data and metadata templates that can readily be used by manufacturers and researchers to upload data to compliant repositories. We propose that compliant repositories automate processing bio-logging data into four levels (described below) compiled to maximise interoperability and facilitate scientific discovery. Such outcomes will be key to improve conservation management and lead to policy development. Although our focus here is on marine bio-logging data, our objective is to contribute to standardising bio-logging datasets across all taxa and ecosystems, which is also one of the stated goals of the International Bio-Logging Society (bio-logging.net).

2 | MATERIALS AND METHODS

We hosted a workshop at the OceanObs'19 conference in Honolulu, Hawaii (oceanobs19.net), to develop a plan for global standardisation of marine bio-logging datasets. The workshop was attended by 28 representatives from national and regional tagging networks, manufacturing companies, and intergovernmental bodies, and the group was subsequently extended to include other key members from the bio-logging community. We recognised the common goal to improve the quality and consistency of processes, measurements, data, and applications through agreed procedures, evolving into and contributing to best practices (cf. Pearlman et al., 2019; Tanhua et al., 2019).

2.1 | Progress to date and lessons learned on bio-logging data standardisation

Varying levels of data standardisation have been achieved by existing repositories storing spatially discrete acoustic telemetry data (e.g. OTN—Ocean Tracking Network, oceantrackingnetwork.org; AODN—Australian Ocean Data Network portal, portal.aodn.org.au). Such standardisation is crucial for acoustic data (resulting from detection of animal-borne transmitters through static receiver stations) to match detections across acoustic networks around the world that are managed by different user groups. Although these repositories are not yet fully interoperable, templates for reporting acoustic tracking data enable integration and rapid data sharing among researchers and existing networks (Bangley et al., 2020). For satellite and archival telemetry data, standardisation is more challenging given the many heterogeneous data file formats that result from the large number of sensors used, existing manufacturers, as well as settings and applications for different tags.

Several biogeographic data aggregators, such as the Global Biodiversity Information Facility (GBIF, gbif.org), the Ocean Biogeographic Information System (OBIS, obis.org) and the Atlas of Living Australia (ALA, ala.org.au), use the Darwin Core body of standards for data interoperability. Darwin Core is a glossary of terms well-suited for spatiotemporal biodiversity data maintained by the Biodiversity Information Standards Group (tdwg.org). However, it has limited capacity for capturing instrument metadata, and does not easily accommodate the multiple different intraspecific and interspecific behaviours (often expressed by metrics recorded by multiple sensors) that occur in a bio-logging study. To address this issue, OBIS has developed a schema (OBIS-Event-Data schema) relevant to acoustic and satellite telemetry data (De Pooter et al., 2017; github.com/tdwg/dwc-for-biologging). Another more recent development is the nc-eTAG, a file format and metadata specification for the production of archive quality, standards-based netCDF (network Common Data Form) data files for different types of electronic tags (Tsontos et al., 2020).

The bio-logging community can leverage these standardisation efforts as well as from learning the standardisation methods already achieved by other established networks (e.g. the Argo floats and Lagrangian drifters) to fast-track bio-logging data standardisation consistent with the GOOS' Framework for Ocean Observation (FOO) (Lindstrom et al., 2012). For example, physical oceanographers have (a) established a permanent Data Management and Communications (DMAC) Centre (ioos.noaa.gov/project/dmac/) providing free access to surface drifter data (aoml.noaa.gov/phod/gdp/), (b) developed a full set of universal data standards for Argo floats (argodatamgt.org/Documentation) and (c) defined procedures to access data (argo.ucsd.edu). There are also meta-repositories, such as Coastwatch (coastwatch.noaa.gov), that serve as meta-hosts by linking and translating oceanographic data files from other repositories, and which could be used as an example for a meta-repository of bio-logging data, particularly relevant for linking telemetry and oceanographic data collected by marine megafauna acting as ocean samplers (Harcourt et al., 2019; Treasure et al., 2017).

2.2 | Standardisation of bio-logging data is needed at multiple levels

Our proposed workflow (Figure 2) aims to advance the standardisation of bio-logging data, using as a starting point three simple templates (in comma-separated values; i.e. .csv format), which are fully described in the *Templates* section in Supporting Information. First, a Device Metadata template (Table S1) should be completed

by manufacturers or companies supporting tag data acquisition and decoding. This template, which comprises information pertaining to the instrument used, will be essential to complete the upload of original, decoded bio-logging data to repositories with relevant metadata about the device. Second, a Deployment Metadata template (Table S2) should be completed by the researchers deploying the tag devices, to encapsulate information about the animal tagged, tagging protocols followed and tag settings. This template provides

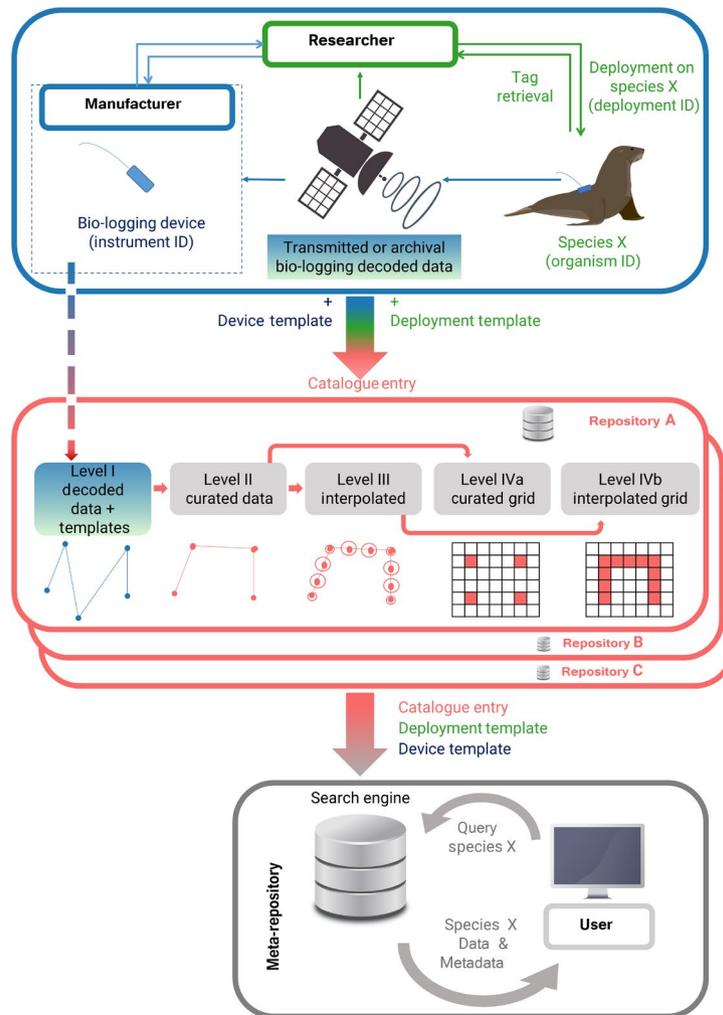


FIGURE 2 Flow for standardisation of bio-logging data from tag to search engine. Standardisation of bio-logging data will need a concerted and coordinated effort across manufacturers, researchers and repositories. It is crucial that the standardisation procedure starts as close as possible to the time of data production. Manufacturers will need to provide a Device metadata template (Table S1) to researchers, and will have the crucial role of creating a data file output option in their data processing software that allows export in a compliant standardised format as we specify here for upload to repositories. This step will be vital, as the current heterogeneity of files provided by the many existing manufacturers presents a major bottleneck for standardisation. Researchers will then have a central role in starting and maintaining data flow after deployment of bio-logging devices, by engaging in the data uploading process and providing the Deployment metadata template (Table S2) where specification of ‘permission-to-use’ (e.g. acknowledgement, consultation or co-authorship) is to be included. Despite the central role of researchers in establishing the data flow, the framework we propose is also prepared for direct upload of data by the manufacturer (indicated by the dashed arrow on the left) as it would be required for near real-time data availability at the repositories. Data are to be uploaded in a standardised format (Table S3) to facilitate data ingestion by repositories, and once at the repositories, bio-logging data and metadata are to be used and kept together during translation into data products (Levels I through to IV; refer to Figure 3 for details) that are to be easily discoverable through a global search engine acting as a meta-repository. Independent users will be able to use this meta-repository to search data and obtain specific data-level products accompanied by the respective metadata to translate it into synthesis products useful for management and conservation while abiding to the ‘permission-to-use’ specifications made by the researcher at the beginning of the process

essential information and context for translation of data into derived products and to enable assessment of possible biases during analysis (Kilkenny et al., 2010; Webster & Rutz, 2020). Clear description of conditions for data usage should be specified by the researchers in this template (e.g. including specific requirements for acknowledgement, attribution of ownership or need for co-authorship in resulting outcomes) or alternatively, default to existing licensing types (e.g. creativecommons.org). These two metadata templates include a range of metadata fields common to all types of bio-logging devices and resulting data, but are flexible enough to accommodate specific subsets unique to each data type (see Supporting Information). A third Input Data template including all data fields needed when using different tag types should also be filled to ensure datasets are standardised and to facilitate data ingestion by repositories (Table S3). This template should be filled by researchers collecting the data, or directly by those acting as the first contact point for data, which depending on the data type could include manufacturers or raw data decoders (e.g. for data collected by satellite).

Our long-term vision for standardising bio-logging data is the development of a suite of dynamic repositories with identical

protocols for data archiving and processing, resulting in interoperable data and metadata. Such interoperable formats will maintain standardisation and data flow as new data are collected (Figure 2). We note that much of the infrastructure needed for implementation already exists, including procedures, standardised vocabularies and formats. Therefore, standardisation could be achieved by improving the uptake of existing infrastructure, and by implementing processes and procedures similar to those used in other fields where data are constantly being updated. A relevant example of the latter is the product levels used by the remote sensing community (e.g. the US National Aeronautics and Space Administration Ocean Biology Processing Group; NASA OceanColor—oceancolor.gsfc.nasa.gov/products). They provide a framework for organising data at various levels, ranging from raw unprocessed instrumental data files (Level 0) to gridded data products with different levels of processing (up to Levels 3 and 4). Such data organisation is directly relevant to bio-logging and we have identified four equivalent levels at which bio-logging data could be standardised in repositories to satisfy most user needs (detailed in Figure 3). Our levels of standardisation start with already decoded bio-logging data (Level 1), instead of raw,

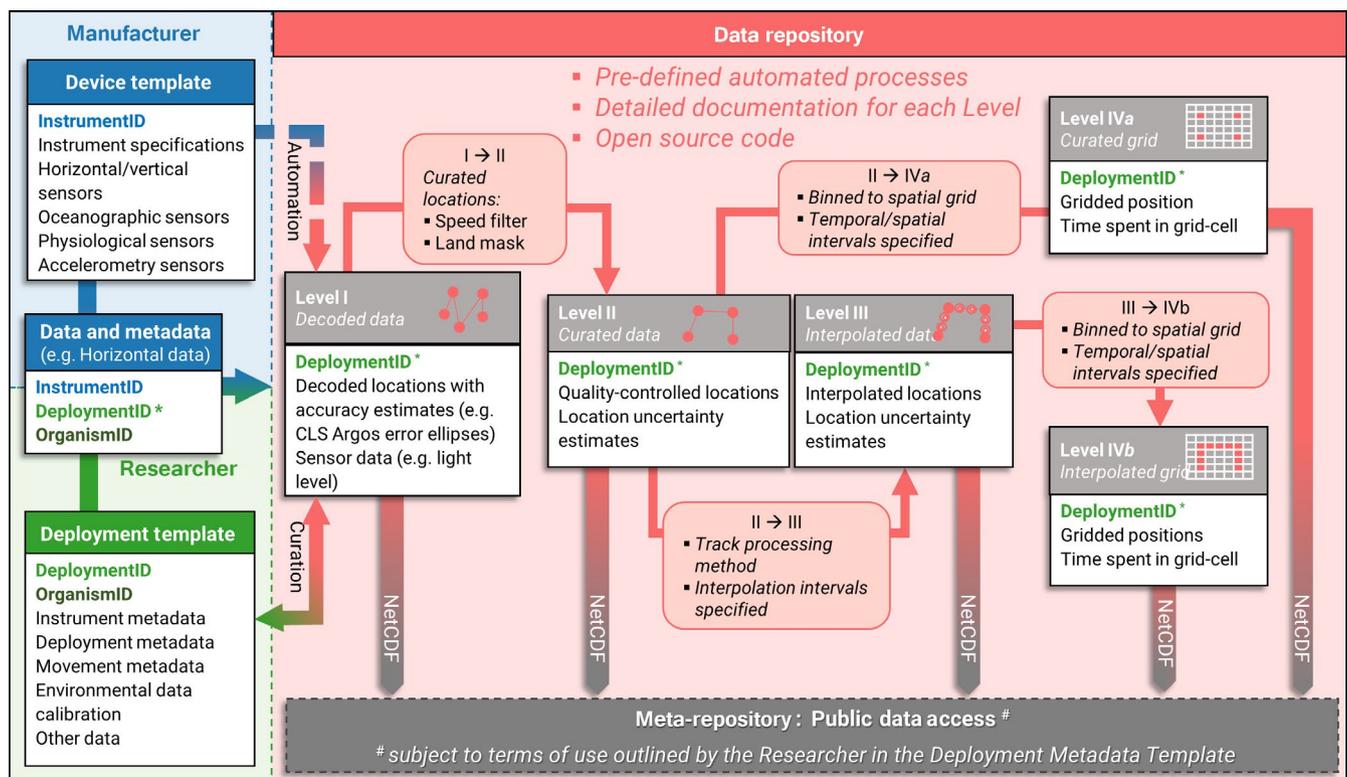


FIGURE 3 Diagram of data processing from Level I through to Level IV at the repositories. Example of data flow for horizontal bio-logging movement datasets. The translation of uploaded data into data products (Levels I–IV) should occur in a reproducible manner across all existing repositories to facilitate integration and interoperability of Level I–IV datasets across repositories. We therefore suggest that this be an automated and standardised process across repositories, where specific processing scripts and definitions for filters, interpolation intervals, and gridding are adopted across repositories (refer to the example we provide in github.com/ocean-tracking-network/biologging_standardization). Full documentation for the data processing settings used should be made available by repositories, including description of the filters used (e.g. speed filter), uncertainty associated with locations provided (e.g. error ellipses), track processing method, interpolation time interval, location uncertainty post-processing, temporal and spatial resolution for gridding. At each level, all metadata attributes should be retained to allow tracing of the same datasets in different formats, with DeploymentID being the key to match data with metadata. The data should be downloadable (where permissions allow) through netCDF files built using standardised CDL files and standardised controlled vocabularies compliant with the Climate Forecast (CF) metadata convention (see example provided on github.com/ocean-tracking-network/biologging_standardization)

unprocessed data files obtained from tags (equivalent to Level 0 in oceancolor products). This is because the Level 0 data are often subject to proprietary rights from manufacturers, and standardisation could become an impediment to innovation of protocols for data storage and transmission.

2.2.1 | Level I—Decoded sensor data

Decoded sensor data, that is, decrypted low-level information obtained directly from sensors after decoding Level 0 data, are critical to ensuring original and complete bio-logging datasets remain archived for future analysis and processing, particularly as downstream methods evolve. Researchers should transfer transmitted and archival data to repositories that share standardised procedures to receive individual datasets. This procedure should involve a step where the researcher assists in flagging (but not removing) meaningful versus erroneous or irrelevant data (e.g. measurements representing the tag deployment vs. pre-deployment). Level I data should include all data provided by the tag, with the relevant data flags. It is desirable that such data are made available immediately at the repository for visualisation in near real time (Sequeira et al., 2019), which can be made possible if upload is made directly by the first point of contact for the data (i.e. manufacturers). This visualisation should be made possible even if data access needs to follow a predefined embargo period, as is already practiced in some existing repositories (e.g. AODN, where some data can have a 2-year embargo despite most data being made open access immediately). Indeed, aggregation or delayed release of bio-logging data might be needed to protect endangered species, and also to allow researchers the opportunity to first publish their findings. Organisation of Level 1 data will also offer a straightforward option for users who are unable to process their data further (e.g. due to time constraints), but want to securely archive their data. Once at the repositories, we suggest that the Level I data and metadata be translated to processed products (Levels II through to IV) in an automated, standardised way as described below, with clear documentation provided at each step.

2.2.2 | Level II—Curated data

Curated bio-logging data, that is, a quality-controlled dataset after removal of invalid, inconsistent or erroneous data points, are a resource for any analyses and further processing ensuring original, unprocessed data are available. Erroneous data include all records that are not representative of an animal's behaviour, such as location points obtained before the tag is deployed or after tag detachment (e.g. a drifting tag), or other obviously impossible locations, such as those inland for animals that are exclusively marine (Freitas et al., 2008; Hoenner et al., 2018). These erroneous positions should be flagged by the researcher during the processing organisation of Level I data, and relevant information (e.g. date for the start of the track as opposed to deployment data) should be provided through a complete Deployment Metadata template. This template will include information to assist

in removing data that do not belong to the tracked animal (e.g. data transmitted by a tag floating after detachment). Production of Level II data can then be automated at the repository by applying relevant filters (e.g. land filter, speed filter), addressing the details provided in the Deployment Metadata template, and clearing or removing the data points flagged in Level I data. A clear log for all the steps employed should be documented by the repository (Figure 3), ensuring a clean and usable version of the original decoded data is available without manipulation or processing for any subsequent analyses.

2.2.3 | Level III—Interpolated data

Interpolated data, that is, processed bio-logging data that include smoothed and interpolated locations, are a resource needed often for analyses involving bio-logging datasets. Processing data in this way is commonly done by applying a state-space model. These types of models are used to filter the data and estimate the animal's most probable path (Braun et al., 2018; Johnson et al., 2008; Jonsen et al., 2005, 2020) or to infer behavioural states (Michelot et al., 2016), which can be used to generate area-use and network models. The processing of Level II data in this way leads to manipulation of the original positions so they are interpolated in equal time intervals to display the most likely track, which does not necessarily include all original positions and is why storage of Level II data is important. There are many different ways to apply state-space models to data. To facilitate integration into large-scale meta-analyses and global datasets, we suggest that repositories include automated processing to produce standardised Level III data while also providing alternatives for user-selected interpolation parameters. Again, the respective documentation detailing the processing used should accompany all resulting products.

2.2.4 | Level IV—Gridded data

Gridded data, that is, bio-logging data presented in a grid format with a specific grid-cell size and temporal resolution, are commonly used to harmonise behavioural data with environmental information from other sources. This procedure has been used in recent synthesis studies (Hindell et al., 2020; Queiroz et al., 2019; Sequeira et al., 2018) and will be needed to address key global challenges associated with human-induced stressors (Sequeira et al., 2019). For this step, a common temporal resolution and grid-cell size should be defined aiming to have standardised Level IV products readily available. This common spatiotemporal resolution could be monthly at 1 degree x 1 degree grid-cell sizes to reduce data gaps in environmental data collected by satellites, such as chlorophyll-*a* (Scales et al., 2017), and following results from other recent literature (Amoroso et al., 2018; Kroodsma et al., 2018a, 2018b; O'Toole et al., 2020). This gridding step should be applied to data Levels II and III to, respectively, produce Levels IVa (gridded curated data) and IVb (gridded interpolated data). In addition to these standardised procedures, options for the

user to select specific spatial and temporal resolutions to grid data. Levels II and III should also be provided by the repository.

2.2.5 | Additional compliance needed at the repositories

At the repository level, an automated mechanism should be used to create a unique catalogue entry ('EntrySourceID') when ingesting the standardised Level I data and metadata supplied by researchers or manufacturers (Tables S1–S3; Figure 2). Each entry will store data corresponding to one deployment from one device and will include global level metadata attributes relating to the Device and Deployment templates, including Organism details, and consistent with existing standards. The 'EntrySourceID' should couple the name of the repository ingesting data, the ownerName or projectName (provided in the Deployment template), and three key IDs contributed in the templates (InstrumentID, DeploymentID and OrganismID), using the following format: `urn:catalog:[repository]:[ownerName/projectName]:[InstrumentID]:[DeploymentID]:[OrganismID]`.

All entries should include a 'quality flag' describing the quality of the data (e.g. one of five levels: `no_data`, `bad_data`, `worst_quality`, `low_quality`, `acceptable_quality` and `best_quality`), which can be used to distinguish datasets with different data quality levels (e.g. geolocation vs. GPS data) and among those, the ones that have undergone quality control (QC) by researchers through a curation step. For acoustic telemetry data, where QC of the detection data is required, the 'Detection_QC' flags introduced by Hoenner et al. (2018), should be used where similar QC tests are implemented. These include 'FDA_QC', 'Distance_QC', 'Velocity_QC', 'DetectionDistribution_QC' and 'DistanceRelease_QC' (for details and definitions, refer to table 1 of that publication).

2.3 | Data format for interoperability

We suggest that all the data levels are made available at compliant repositories (Figure 3) and formatted to ensure data and metadata are kept together during all data exchanges. For this, a network Common Data Form (netCDF) format combined with standardised controlled vocabularies compliant with the Climate Forecast (CF) metadata convention could be most useful (see netCDF section in Supporting Information). NetCDF is a self-describing, machine-independent data format and associated set of software libraries, which supports the creation, access, and sharing of scientific data. Such an interoperable data file format would facilitate exchange of bio-logging data with associated metadata templates, and there are existing tools to facilitate conversion from netCDF to a range of output formats, including commonly used tabular text formats (e.g. .csv). Indeed, adoption of such standard formats by existing consortia such as the Marine Mammals Exploring the Oceans Pole to Pole (MEOP; meop.net) has increased data uptake by the oceanographic community, consolidating animal-collected data as a source to GOOS networks such as AniBOS and other end-users (Treasure et al., 2017). Recent developments,

including the nc-eTAG format (Tsonetos et al., 2020), which hierarchically stores blocks of attributes by tag or feature and allows specification of metadata consistent with the latest standards and next generation CF enhancements (github.com/Unidata/EC-netCDF-CF), provide a standards-based specification to store a range of bio-logging data, including satellite, archival and retrieved pop-up archival (PSAT) data types. Storing data as netCDF using standardised Common DATA Language (CDL) files (see netCDF section in supplementary information) will allow integration of tag instrument data file collections in web server technologies such as THREDDS Data Server (TDS; unidata.ucar.edu/software/tds), ERDDAP and OPenDAP for subsetting, aggregation and distribution of data to the community. Repositories should include information on how to use netCDF files and how to convert them to other formats as needed for input to other software programs for visualisation and analysis.

2.4 | Challenges for achieving standardisation

Standardisation of bio-logging data is needed to manage incoming data and to retrospectively compile the thousands of bio-logging datasets already in existence (Block et al., 2011; Queiroz et al., 2019; Ropert-Coudert et al., 2020; Sequeira et al., 2018; Thums et al., 2018). Infrastructure support and developments will be needed to keep pace with technological advances, including provisions for near real-time data, mobile receivers and novel tag types. Indeed, the need for standardisation across platforms will be exacerbated as sensor technology develops. Defining the metadata profiles for each of the existing and new sensors will also be necessary, and mapping common elements across metadata schemas will be needed to enable integration across at least a minimal subset of required attributes.

The setup of a workflow for production of archive-quality data files at all levels is also a challenge. Although the most familiar output format options that are widely used as input for analyses should continue to be available (e.g. .csv), capacity building to train the ecology community in the use of netCDF data formats will be needed. Specifically, technology and infrastructure gaps in least-developed countries need to be addressed, for example, by engaging networks of researchers and manufacturers in the creation of translation tools, that is, tools allowing translation between data types (e.g. Rosetta; unidata.ucar.edu/software/rosetta; a UNIDATA tool to convert tabular .csv files to standards compliant netCDF files) and 'software carpentry' courses (e.g. software-carpentry.org) to deliver training in data management and analysis.

The need for automation of data processing highlights the need to incorporate data science in ecology and to strengthen engagement between scientists from different disciplines (e.g. computer science and engineering with ecology). Machine-to-machine readability is important for effective standardisation, as is the ability to quickly visualise and analyse data across large and disparate datasets. For this, the coupling of metadata with different levels of processed tracking data and environmental and oceanographic data will need to be streamlined.

2.5 | Advantages of standardising bio-logging data

Standardised bio-logging data will lead to major advances in (a) understanding the distribution, movement and behaviour of species, (b) improving our capacity to make comparisons across regions and taxa and (c) providing concomitant environmental data that place animal behaviour information into its ecological context contributing to global observation. Importantly, these advances will, in turn, provide information needed for improving conservation outcomes for species at risk from human activities. Standardisation will facilitate a broad and effective use of bio-logging data to understand ecosystem dynamics and to establish collaborative networks of ecologists, environmental and data scientists, and ecosystem managers. Researchers contributing data will benefit from an effective framework for data storage and retrieval, allowing added value to all datasets collected while ensuring rightful attribution and accuracy-of-use. Additionally, if existing repositories provide harmonised, archive-quality netCDF files with consistent and well-structured metadata, data exchange can be streamlined through the creation of a global 'meta-repository' as a search engine (i.e. a discovery tool similar to datasetsearch.research.google.com).

Standardisation of bio-logging data will also facilitate standardisation and integration of other datasets, including relevant ancillary data that can improve our understanding about ecological and evolutionary responses of animals to environmental change. These might include data associated with the individual's origin, physiological state or movements prior to the tagging period, as well as dietary habits, growth rates and breeding behaviour, and could include datasets derived from tissue sampling, such as muscle plugs, fin clips, hairs, whiskers or feathers. Standardised vocabularies and data formatting options (including the nc-eTAG format described above) can be extended to deal with diverse ancillary information, in coordination with relevant data platforms and standards from other disciplines. Additionally, standards for netCDF-Linked Data (LD; <https://binary-array-ld.github.io/netcdf-ld>) that will enable automated cross-referencing of metadata within data files are now emerging. Moreover, formatting data as netCDF consistent with the CF standards will provide compatibility with the global observing communities and likely facilitate integration with a range of diverse environmental and oceanographic data products such as bathymetry, satellite-derived and modelled temperatures, winds and currents, and chlorophyll *a*.

Specifically for marine bio-logging data, standardisation will represent a step towards further integration of observations into GOOS, following similar procedures already used for a broad array of ocean sensor platforms, including gliders (Rudnick, 2016) and acoustic platforms. Delivering standardised data streams will provide the broader ocean community with a more efficient way to assess the state of the world's oceans as they change and inform national and international assessments, including the Regular Process, the World Ocean Assessment, assessments undertaken by the Intergovernmental Science Policy Platform on Biodiversity and Ecosystem Services, and Conventions on Biological Diversity and on Migratory Species. Bio-logging provides data on multidisciplinary EOVs that may act as 'indicators' to be used in national reporting

to biodiversity conventions and internationally to monitor progress towards the UN Sustainable Development Goal 14 (SDG14; un.org/sustainabledevelopment/sustainable-development-goals/) and the new targets under the Post-2020 Global Biodiversity Framework. Current developments associated with the blue economy agenda (Eikeset et al., 2018), the global aim to achieve SDG14, and the requirement to provide key observations in support of the UN Decade of the Ocean Science (Ryabinin et al., 2019), emphasise the need for marine bio-logging data to be made readily available. Appropriate information on movements and ecology is urgently needed to inform conservation of species at risk of extinction (Estes et al., 2016; McCauley et al., 2015).

ACKNOWLEDGEMENTS

We are thankful to ONR and UWA OI for funding the workshop, and to ARC for DP210103091. A.M.M.S. was funded by a 2020 Pew Fellowship in Marine Conservation, and also supported by AIMS. C.R. was the recipient of a Radcliffe Fellowship at the Radcliffe Institute for Advanced Study, Harvard University. We thank Suzi Kohin and Matthew Ruthishauser from Wildlife Computers for earlier discussions and feedback on the manuscript.

AUTHORS' CONTRIBUTIONS

A.M.M.S., M.O., D.P.C., M.R.H., C.R.M., R.H. and M.W. conceived the study and organised the workshop; A.M.M.S., M.O., T.R.K., L.H.M., I.D.J., J.P., S.J.B., E.L.H., K.H., M.H., C.B., D.C.D., M.F., M.A.H., M.K.M., M.M.C.M., S.E.S., B.T., F.W., B.W., D.P.C., M.R.H., C.R.M., R.H., M.W. and F.W. attended the workshop and prepared the first draft; A.M.M.S., M.O., T.R.K., L.H.M., C.D.B., X.H., F.R.A.J., P.N., J.P., S.J.B. and V.T. compiled information and prepared the templates; J.P., P.N., T.R.K., L.H.M., C.D.B., F.R.A.J., I.J., V.T. and A.M.M.S. prepared GitHub content; A.M.M.S., M.O., F.R.A.J., J.P., G.C.H., E.L.H., S.J.B. and M.H. prepared the figures; A.M.M.S. led the writing. All authors contributed and edited the manuscript.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/2041-210X.13593>.

DATA AVAILABILITY STATEMENT

All data used in the manuscript, including 'templates' and associated definition of terms, example data showing the format to be used for data upload, code to convert between standardised data levels, CDL and netCDF examples are available from github.com/ocean-tracking-network/biologging_standardization and Sequeira et al. (2021).

ORCID

Ana M. M. Sequeira  <https://orcid.org/0000-0001-6906-799X>
 Xavier Hoenner  <https://orcid.org/0000-0001-5811-1166>
 Fabrice R. A. Jaine  <https://orcid.org/0000-0002-9304-5034>
 Peggy Newman  <https://orcid.org/0000-0002-9084-5992>
 Graeme C. Hays  <https://orcid.org/0000-0002-3314-8189>
 Elliott L. Hazen  <https://orcid.org/0000-0002-0412-7178>

Vardis M. Tsontos  <https://orcid.org/0000-0002-1723-0860>
 Clint Blight  <https://orcid.org/0000-0002-3481-7428>
 Sarah C. Davidson  <https://orcid.org/0000-0002-2766-9201>
 Victor M. Eguiluz  <https://orcid.org/0000-0003-1133-1289>
 Michael Fedak  <https://orcid.org/0000-0002-9569-1128>
 Mark A. Hindell  <https://orcid.org/0000-0002-7823-7185>
 Ivica Janekovic  <https://orcid.org/0000-0001-8388-0848>
 Mônica M. C. Muelbert  <https://orcid.org/0000-0002-5992-5994>
 Christian Rutz  <https://orcid.org/0000-0001-5187-7417>
 David W. Sims  <https://orcid.org/0000-0002-0916-7363>
 Samantha E. Simmons  <https://orcid.org/0000-0003-0326-6990>
 Clive R. McMahon  <https://orcid.org/0000-0001-5241-8917>
 Rob Harcourt  <https://orcid.org/0000-0003-4666-2934>

REFERENCES

- Amoroso, R. O., Parma, A. M., Pitcher, C. R., McConnaughey, R. A., & Jennings, S. (2018). Comment on 'Tracking the global footprint of fisheries'. *Science*, 361, eaat6713. <https://doi.org/10.1126/science.aat6713>
- Bailey, H., Fossette, S., Bograd, S. J., Shillinger, G. L., Swithenbank, A. M., Georges, J. Y., Gaspar, P., Stromberg, K. H. P., Paladino, F. V., Spotila, J. R., Block, B. A., & Hays, G. C. (2012). Movement patterns for a critically endangered species, the leatherback turtle (*Dermochelys coriacea*), linked to foraging success and population status. *PLoS ONE*, 7, e36401. <https://doi.org/10.1371/journal.pone.0036401>
- Bangley, C. W., Whoriskey, F. G., Young, J. M., & Ogburn, M. B. (2020). Networked animal telemetry in the northwest Atlantic and Caribbean waters. *Marine and Coastal Fisheries*, 12, 339–347. <https://doi.org/10.1002/mcf2.10128>
- Bates, A. E., Primack, R. B., Moraga, P., & Duarte, C. M. (2020). COVID-19 pandemic and associated lockdown as a 'Global Human Confinement Experiment' to investigate biodiversity conservation. *Biological Conservation*, 248, 108665. <https://doi.org/10.1016/j.biocon.2020.108665>
- Block, B. A., Dewar, H., Farwell, C., & Prince, E. D. (1998). A new satellite technology for tracking the movements of Atlantic bluefin tuna. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 9384–9389. <https://doi.org/10.1073/pnas.95.16.9384>
- Block, B. A., Jonsen, D., Jorgensen, S. J., Winship, A. J., Shaffer, S. A., Bograd, S. J., Hazen, E. L., Foley, D. G., Breed, G. A., Harrison, A.-L., Ganong, J. E., Swithenbank, A., Castleton, M., Dewar, H., Mate, B. R., Shillinger, G. L., Schaefer, K. M., Benson, S. R., Weise, M. J., ... Costa, D. P. (2011). Tracking apex marine predator movements in a dynamic ocean. *Nature*, 475, 86–90. <https://doi.org/10.1038/nature10082>
- Braun, C. D., Galuardi, B., & Thorrold, S. R. (2018). HMMoce: An R package for improved geolocation of archival-tagged fishes using a hidden Markov method. *Methods in Ecology and Evolution*, 9, 1212–1220. <https://doi.org/10.1111/2041-210X.12959>
- Brett, A., Leape, J., Abbott, M., Sakaguchi, H., Cao, L., Chand, K., Golbuu, Y., Martin, T. J., Mayorga, J., & Myksovoll, M. S. (2020). Ocean data need a sea change to help navigate the warming world. *Nature*, 582, 181–183. <https://doi.org/10.1038/d41586-020-01668-z>
- Campbell, H. A., Urbano, F., Davidson, S., Dettki, H., & Cagnacci, F. (2016). A plea for standards in reporting data collected by animal-borne electronic devices. *Animal Biotelemetry*, 4. <https://doi.org/10.1186/s40317-015-0096-x>
- Davidson, S. C., Bohrer, G., Gurarie, E., LaPoint, S., Mahoney, P. J., Boelman, N. T., Eitel, J. U. H., Prugh, L. R., Vierling, L. A., Jennewein, J., Grier, E., Couriot, O., Kelly, A. P., Meddens, A. J. H., Oliver, R. Y., Kays, R., Wikelski, M., Aarvak, T., Ackerman, J. T., ... Hebblewhite, M. (2020). Ecological insights from three decades of animal movement tracking across a changing Arctic. *Science*, 370, 712.
- De Pooter, D., Appeltans, W., Bailly, N., Bristol, S., Deneudt, K., Eliezer, M., Fujioka, E., Giorgetti, A., Goldstein, P., Lewis, M., Lipizer, M., Mackay, K., Marin, M., Moncoiff, G., Nikolopoulou, S., Provoost, P., Rauch, S., Roubicek, A., Torres, C., ... Hernandez, F. (2017). Toward a new data standard for combined marine biological and environmental datasets – Expanding OBIS beyond species occurrences. *Biodiversity Data Journal*, 5, e10989.
- Eikeset, A. M., Mazzarella, A. B., Daviosdottir, B., Klinger, D. H., Levin, S. A., Rovenskaya, E., & Stenseth, N. C. (2018). What is blue growth? The semantics of 'Sustainable Development' of marine environments. *Marine Policy*, 87, 177–179.
- Estes, J. A., Heithaus, M., McCauley, D. J., Rasher, D. B., & Worm, B. (2016). Megafaunal impacts on structure and function of ocean ecosystems. *Annual Review of Environment and Resources*, 41(41), 83–116.
- Fossette, S., Witt, M. J., Miller, P., Nalovic, M. A., Albareda, D., Almeida, A. P., Broderick, A. C., Chacon-Chaverri, D., Coyne, M. S., Domingo, A., Eckert, S., Evans, D., Fallabrino, A., Ferraroli, S., Formia, A., Giffoni, B., Hays, G. C., Hughes, G., Kelle, L., ... Godley, B. J. (2014). Pan-Atlantic analysis of the overlap of a highly migratory species, the leatherback turtle, with pelagic longline fisheries. *Proceedings of the Royal Society B-Biological Sciences*, 281, 20133065.
- Freitas, C., Lydersen, C., Fedak, M. A., & Kovacs, K. M. (2008). A simple new algorithm to filter marine mammal Argos locations. *Marine Mammal Science*, 24, 315–325.
- Gleiss, A. C., Norman, B., Liebsch, N., Francis, C., & Wilson, R. P. (2009). A new prospect for tagging large free-swimming sharks with motion-sensitive data-loggers. *Fisheries Research*, 97, 11–16.
- Harcourt, R., Sequeira, A. M. M., Zhang, X. L., Roquet, F., Komatsu, K., Heupel, M., McMahon, C., Whoriskey, F., Meehan, M., Carroll, G., Brodie, S., Simpfendorfer, C., Hindell, M., Jonsen, I., Costa, D. P., Block, B., Muelbert, M., Woodward, B., Weise, M., ... Fedak, M. A. (2019). Animal-borne telemetry: An integral component of the ocean observing toolkit. *Frontiers in Marine Science*, 6, 326.
- Hays, G., Bailey, H., Bograd, S. J., Bowe, W. D., Campagna, C., Carmichael, R. H., Casale, P., Chiaradia, A., Costa, D. P., Cuevas, E., Nico de Bruyn, P. J., Dias, M. P., Duarte, C. M., Dunn, D. C., Dutton, P. H., Esteban, N., Friedlaender, A., Goetz, K. T., Godley, B. J., ... Sequeira, A. (2019). Translating marine animal tracking data into conservation policy and management. *Trends in Ecology & Evolution*, 34, 459–473.
- Hays, G. C., & Hawkes, L. A. (2018). Satellite tracking sea turtles: Opportunities and challenges to address key questions. *Frontiers in Marine Science*, 5, 432.
- Hindell, M. A., Reisinger, R. R., Ropert-Coudert, Y., Huckstadt, L. A., Trathan, P. N., Bornemann, H., Charrassin, J. B., Chown, S. L., Costa, D. P., Danis, B., Lea, M. A., Thompson, D., Torres, L. G., Van de Putte, A. P., Alderman, R., Andrews-Goff, V., Arthur, B., Ballard, G., Bengtson, J., ... Raymond, B. (2020). Tracking of marine predators to protect Southern Ocean ecosystems. *Nature*, 87–92.
- Hoenner, X., Huvneers, C., Steckenreuter, A., Simpfendorfer, C., Tattersall, K., Jaine, F., Atkins, N., Babcock, R., Brodie, S., Burgess, J., Campbell, H., Heupel, M., Pasquer, B., Proctor, R., Taylor, M. D., Udyawer, V., & Harcourt, R. (2018). Australia's continental-scale acoustic tracking database and its automated quality control process. *Scientific Data*, 5, 170206.
- Hussey, N. E., Kessel, S. T., Aarestrup, K., Cooke, S. J., Cowley, P. D., Fisk, A. T., Harcourt, R. G., Holland, K. N., Iverson, S. J., Kocik, J. F., Flemming, J. E. M., & Whoriskey, F. G. (2015). Aquatic animal telemetry: A panoramic window into the underwater world. *Science*, 348, 1221–1231.
- Johnson, D. S., London, J. M., Lea, M. A., & Durban, J. W. (2008). Continuous-time correlated random walk model for animal telemetry data. *Ecology*, 89, 1208–1215.

- Jonsen, I. D., Flemming, J. M., & Myers, R. A. (2005). Robust state-space modelling of animal movement data. *Ecology*, *86*, 2874–2880.
- Jonsen, I. D., Patterson, T. A., Costa, D. P., Doherty, P. D., Godley, B. J., Grecian, W. J., Guinet, C., Hoenner, X., Kienle, S. S., Robinson, P. W., Votier, S. C., Whiting, S., Witt, M. J., Hindell, M. A., Harcourt, R. G., & McMahon, C. R. (2020). A continuous-time state-space model for rapid quality control of argos locations from animal-borne tags. *Movement Ecology*, *8*. <https://doi.org/10.1186/s40462-020-00217-7>
- Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, *348*, 1222–1231.
- Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M., & Altman, D. G. (2010). Improving Bioscience Research Reporting: The ARRIVE Guidelines for Reporting Animal Research. *PLoS Biology*, *8*, e1000412. <https://doi.org/10.1371/journal.pbio.1000412>
- Kroodsma, D. A., Mayorga, J., Hochberg, T., Miller, N. A., Boerder, K., Ferretti, F., Wilson, A., Bergman, B., White, T. D., Block, B. A., Woods, P., Sullivan, B., Costello, C., & Worm, B. (2018a). Response to comment on 'Tracking the global footprint of fisheries'. *Science*, *361*, eaat7789. <https://doi.org/10.1126/science.aat7789>
- Kroodsma, D. A., Mayorga, J., Hochberg, T., Miller, N. A., Boerder, K., Ferretti, F., Wilson, A., Bergman, B., White, T. D., Block, B. A., Woods, P., Sullivan, B., Costello, C., & Worm, B. (2018b). Tracking the global footprint of fisheries. *Science*, *359*, 904–908. <https://doi.org/10.1126/science.aao5646>
- Lin, D., Crabtree, J., Dillo, I., Downs, R. R., Edmunds, R., Giaretta, D., De Giusti, M., L'Hours, H., Hugo, W., Jenkyns, R., Khodiyar, V., Martone, M. E., Mokrane, M., Navale, V., Petters, J., Sierman, B., Sokolova, D. V., Stockhause, M., & Westbrook, J. (2020). The TRUST Principles for digital repositories. *Scientific Data*, *7*. <https://doi.org/10.1038/s41597-020-0486-7>
- Lindstrom, E., Gunn, J., Fischer, A., McCurdy, A., Glover, L. K., Alverson, K., Bex, B., Burkill, P., Chavez, F., Checkley, D., Clark, C., Fabry, V., Hall, J., Masumoto, Y., Meldrum, D., Meredith, M., Monteiro, P., Mulbert, J., Pouliquen, S., ... Wilson, S. (2012). *A framework for ocean observing (by the Task Team for the Integrated Framework for Sustained Ocean Observing)*. UNESCO.
- McCauley, D. J., Pinsky, M. L., Palumbi, S. R., Estes, J. A., Joyce, F. H., & Warner, R. R. (2015). Marine defaunation: Animal loss in the global ocean. *Science*, *347*, 1255641–1255647. <https://doi.org/10.1126/science.1255641>
- McGowan, J., Beger, M., Lewison, R. L., Harcourt, R., Campbell, H., Priest, M., Dwyer, R. G., Lin, H. Y., Lentini, P., Dudgeon, C., McMahon, C., Watts, M., & Possingham, H. P. (2017). Integrating research using animal-borne telemetry with the needs of conservation management. *Journal of Applied Ecology*, *54*, 423–429. <https://doi.org/10.1111/1365-2664.12755>
- Michelot, T., Langrock, R., & Patterson, T. A. (2016). moveHMM: An R package for the statistical modelling of animal movement data using hidden Markov models. *Methods in Ecology and Evolution*, *7*, 1308–1315.
- Miloslavich, P., Bax, N. J., Simmons, S. E., Klein, E., Appeltans, W., Aburto-Oropeza, O., Garcia, M. A., Batten, S. D., Benedetti-Cecchi, L., Checkley, D. M., Chiba, S., Duffy, J. E., Dunn, D. C., Fischer, A., Gunn, J., Kudela, R., Marsac, F., Muller-Karger, F. E., Obura, D., & Shin, Y. J. (2018). Essential ocean variables for global sustained observations of biodiversity and ecosystem changes. *Global Change Biology*, *24*, 2416–2433.
- Moore, A. M., Arango, H. G., Broquet, G., Edwards, C., Veneziani, M., Powell, B., Foley, D., Doyle, J. D., Costa, D., & Robinson, P. (2011). The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems Part II – Performance and application to the California Current System. *Progress in Oceanography*, *91*, 50–73.
- Muller-Karger, F. E., Miloslavich, P., Bax, N. J., Simmons, S., Costello, M. J., Pinto, I. S., Canonico, G., Turner, W., Gill, M., Montes, E., Best, B. D., Pearlman, J., Halpin, P., Dunn, D., Benson, A., Martin, C. S., Weatherdon, L. V., Appeltans, W., Provoost, P., ... Geller, G. (2018). Advancing marine essential observations and data requirements of the complementary essential ocean variables (EOVs) and essential biodiversity variables (EBVs) frameworks. *Frontiers in Marine Science*, *5*, 211.
- O'Toole, M., Queiroz, N., Humphries, N. E., Sims, D. W., & Sequeira, A. M. M. (2020). Quantifying effects of tracking data bias on species distribution models. *Methods in Ecology and Evolution*, *12*, 170–181.
- Pearlman, J., Bushnell, M., Coppola, L., Karstensen, J., Buttigieg, P. L., Pearlman, F., Simpsons, P., Barbier, M., Muller-Karger, F. E., Munoz-Mas, C., Pissierssens, P., Chandler, C., Hermes, J., Heslop, E., Jenkyns, R., Achterberg, E. P., Bensi, M., Bittig, H. C., Blandin, J., ... Whoriskey, F. (2019). Evolving and sustaining ocean best practices and standards for the next decade. *Frontiers in Marine Science*, *6*, 277.
- Queiroz, N., Humphries, N. E., Couto, A., Vedor, M., da Costa, I., Sequeira, A. M. M., Mucientes, G., Santos, A. M., Abascal, F. J., Abercrombie, D. L., Abrantes, K., Acuña-Marrero, D., Afonso, A. S., Afonso, P., Anders, D., Araujo, G., Arauz, R., Bach, P., Barnett, A., ... Sims, D. W. (2019). Global spatial risk assessment of sharks under the footprint of fisheries. *Nature*, *572*, 461–466.
- Roe, J. H., Morreale, S. J., Paladino, F. V., Shillinger, G. L., Benson, S. R., Eckert, S. A., Bailey, H., Tomillo, P. S., Bograd, S. J., Eguchi, T., Dutton, P. H., Seminoff, J. A., Block, B. A., & Spotila, J. R. (2014). Predicting bycatch hotspots for endangered leatherback turtles on longlines in the Pacific Ocean. *Proceedings of the Royal Society B-Biological Sciences*, *281*, 20132559.
- Ropert-Coudert, Y., Van de Putte, A. P., Reisinger, R. R., Bornemann, H., Charrassin, J. B., Costa, D. P., Danis, B., Huckstadt, L. A., Jonsen, I. D., Lea, M. A., Thompson, D., Torres, L. G., Trathan, P. N., Wotherspoon, S., Ainley, D. G., Alderman, R., Andrews-Goff, V., Arthur, B., Ballard, G., ... Hindell, M. A. (2020). The retrospective analysis of Antarctic tracking data project. *Scientific Data*, *7*, 94. <https://doi.org/10.1038/s41597-020-0406-x>
- Roquet, F., Williams, G., Hindell, M. A., Harcourt, R., McMahon, C., Guinet, C., Charrassin, J.-B., Reverdin, G., Boehme, L., Lovell, P., & Fedak, M. (2014). A Southern Indian Ocean database of hydrographic profiles obtained with instrumented elephant seals. *Nature Science Data*, *1*, 140028. <https://doi.org/10.1038/sdata.2014.28>
- Roquet, F., Wunsch, C., Forget, G., Heimbach, P., Guinet, C., Reverdin, G., Charrassin, J. B., Bailleul, F., Costa, D. P., Huckstadt, L. A., Goetz, K. T., Kovacs, K. M., Lydersen, C., Biuw, M., Nost, O. A., Bornemann, H., Ploetz, J., Bester, M. N., McIntyre, T., ... Fedak, M. A. (2013). Estimates of the Southern Ocean general circulation improved by animal-borne instruments. *Geophysical Research Letters*, *40*, 6176–6180. <https://doi.org/10.1002/2013GL058304>
- Rudnick, D. L. (2016). Ocean research enabled by underwater gliders. *Annual Review of Marine Science*, *8*, 519–541. <https://doi.org/10.1146/annurev-marine-122414-033913>
- Rutz, C., & Hays, G. C. (2009). New frontiers in biologging science. *Biology Letters*, *5*, 289–292. <https://doi.org/10.1098/rsbl.2009.0089>
- Rutz, C., Loretto, M. C., Bates, A. E., Davidson, S. C., Duarte, C. M., Jetz, W., Johnson, M., Kato, A., Kays, R., Mueller, T., Primack, R. B., Ropert-Coudert, Y., Tucker, M. A., Wikelski, M., & Cagnacci, F. (2020). COVID-19 lockdown allows researchers to quantify the effects of human activity on wildlife. *Nature Ecology & Evolution*, *4*, 1156–1159. <https://doi.org/10.1038/s41559-020-1237-z>
- Ryabinin, V., Barbieri, J., Haugan, P., Kullenberg, G., Smith, N., McLean, C., Troisi, A., Fischer, A., Arico, S., Aarup, T., Pissierssens, P., Visbeck, M., Enevoldsen, H. O., & Rigaud, J. (2019). The UN decade of ocean science for sustainable development. *Frontiers in Marine Science*, *6*. <https://doi.org/10.3389/fmars.2019.00470>
- Scales, K. L., Hazen, E. L., Jacox, M. G., Edwards, C. A., Boustany, A. M., Oliver, M. J., & Bograd, S. J. (2017). Scale of inference: On the sensitivity of habitat models for wide-ranging marine predators to the resolution of environmental data. *Ecography*, *40*, 210–220. <https://doi.org/10.1111/ecog.02272>

- Sequeira, A. M. M., Hays, G. C., Sims, D. W., Eguiluz, V. M., Rodríguez, J. P., Heupel, M. R., Harcourt, R., Calich, H., Queiroz, N., Costa, D. P., Fernandez-Gracia, J., Ferreira, L. C., Goldsworthy, S. D., Hindell, M. A., Lea, M. A., Meekan, M. G., Pagano, A. M., Shaffer, S. A., Reisser, J., ... Duarte, C. M. (2019). Overhauling ocean spatial planning to improve marine megafauna conservation. *Frontiers in Marine Science*, 6. <https://doi.org/10.3389/fmars.2019.00639>
- Sequeira, A. M. M., Keates, T. R., McDonnell, L. H., Braun, C. D., Newman, P., & Pye, J. (2021). Ocean-tracking-network/biologging_standardization: Zenodo Archival Release (Version 0.2). *Zenodo*, <https://zenodo.org/record/4560722#YDcJV-gzaUk>
- Sequeira, A. M. M., Rodríguez, J. P., Eguiluz, V. M., Harcourt, R., Hindell, M., Sims, D. W., Duarte, C. M., Costa, D. P., Fernandez-Gracia, J., Ferreira, L. C., Hays, G. C., Heupel, M. R., Meekan, M. G., Avenn, A., Bailleul, F., Baylis, A. M. M., Berumen, M. L., Braun, C. D., Burns, J., ... Thums, M. (2018). Convergence of marine megafauna movement patterns in coastal and open oceans. *Proceedings of the National Academy of Sciences of the United States of America*, 115, 3072–3077. <https://doi.org/10.1073/pnas.1716137115>
- Tanhua, T., McCurdy, A., Fischer, A., Appeltans, W., Bax, N., Currie, K., DeYoung, B., Dunn, D., Heslop, E., Glover, L. K., Gunn, J., Hill, K., Ishii, M., Legler, D., Lindstrom, E., Miloslavich, P., Moltmann, T., Nolan, G., Palacz, A., ... Wilkin, J. (2019). What we have learned from the framework for ocean observing: Evolution of the global ocean observing system. *Frontiers in Marine Science*, 6. <https://doi.org/10.3389/fmars.2019.00471>
- Thums, M., Fernandez-Gracia, J., Sequeira, A. M. M., Eguiluz, V. M., Duarte, C. M., & Meekan, M. G. (2018). How big data fast tracked human mobility research and the lessons for animal movement ecology. *Frontiers in Marine Science*, 5. <https://doi.org/10.3389/fmars.2018.00021>
- Treasure, A. M., Roquet, F., Ansorge, I. J., Bester, M. N., Boehme, L., Bornemann, H., Charrassin, J. B., Chevallier, D., Costa, D. P., Fedak, M. A., Guinet, C., Hammill, M. O., Harcourt, R. G., Hindell, M. A., Kovacs, K. M., Lea, M. A., Lovell, P., Lowther, A. D., Lydersen, C., ... de Bruyn, P. J. N. (2017). Marine mammals exploring the oceans pole to pole: A review of the MEOP consortium. *Oceanography*, 30, 132–138. <https://doi.org/10.5670/oceanog.2017.234>
- Tsontos, V. M., Arms, S. A., & Lam, C. H. (2020). *The nc-eTAG specification and netCDF templates for electronic tagging data*. (NASA, Ed.). 45.
- Tucker, M. A., Bohning-Gaese, K., Fagan, W. F., Fryxell, J. M., Van Moorter, B., Alberts, S. C., Ali, A. H., Allen, A. M., Attias, N., Avgar, T., Bartlam-Brooks, H., Bayarbaatar, B., Belant, J. L., Bertassoni, A., Beyer, D., Bidner, L., van Beest, F. M., Blake, S., Blaum, N., ... Mueller, T. (2018). Moving in the Anthropocene: Global reductions in terrestrial mammalian movements. *Science*, 359, 466–469. <https://doi.org/10.1126/science.aam9712>
- Watanabe, Y., & Sato, K. (2008). Functional dorsoventral symmetry in relation to lift-based swimming in the ocean sunfish *Mola mola*. *PLoS ONE*, 3. <https://doi.org/10.1371/journal.pone.0003446>
- Webster, M., & Rutz, C. (2020). How STRANGE are your study animals? *Nature*, 582, 337–340. <https://doi.org/10.1038/d41586-020-01751-5>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., Santos, L. B. D., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2019). The FAIR Guiding Principles for scientific data management and stewardship (vol. 15, 160018, 2016). *Scientific Data*, 6, 160018.
- Williams, H. J., Taylor, L. A., Benhamou, S., Bijleveld, A. I., Clay, T. A., Grissac, S., Demšar, U., English, H. M., Franconi, N., Gómez-Laich, A., Griffiths, R. C., Kay, W. P., Morales, J. M., Potts, J. R., Rogerson, K. F., Rutz, C., Spelt, A., Trevail, A. M., Wilson, R. P., & Börger, L. (2020). Optimizing the use of biologgers for movement ecology research. *Journal of Animal Ecology*, 89, 186–206. <https://doi.org/10.1111/1365-2656.13094>
- Wilson, R. P., Shepard, E. L. C., & Liebsch, N. (2008). Prying into the intimate details of animal lives: Use of a daily diary on animals. *Endangered Species Research*, 4, 123–137. <https://doi.org/10.3354/esr00064>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Sequeira AMM, O'Toole M, Keates TR, et al. A standardisation framework for bio-logging data to advance ecological research and conservation. *Methods Ecol Evol.* 2021;00:1–12. <https://doi.org/10.1111/2041-210X.13593>