Taking stock of land-based citizen science cetacean data sets: regional integration and comparative analysis to deliver policy-relevant results

An independent workshop held alongside the
2019 World Marine Mammal Conference at
Four Points by Sheraton Barcelona Diagonal Hotel, Spain
Sunday, 8th December 2019

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

INSTITUTO DE INVESTIGACIONES MARINAS

WDC WHALE AND DOLPHIN CONSERVATION

Report editors and authors:

Editors / workshop coordinators: Alice E. M. Walters (WDC) & Graham J. Pierce (IIM CSIC)

Authors: Alice E. M. Walters (WDC), Jeroen Creuwels (Dutch Marine Mammal Workgroup), Katie Dyke (WDC), Peter Evans (Seawatch Foundation), Karen Hall (NatureScot), Conor Ryan (MRC – Song of the Whale), Nikki Taylor (JNCC), Gemma Veneruso (Bangor University), James Waggitt (Bangor University), Graham J. Pierce (IIM/CSIC)

Citation: Walters, A.E.M., Creuwels, J., Dyke, K., Evans, P.G.H., Hall, K., Ryan, C., Taylor, N., Veneruso, G., Waggitt, J. & Pierce, G.J., 2022. Taking stock of land-based citizen science cetacean data sets: regional integration and comparative analysis to deliver policy-relevant results. An independent workshop held alongside the 2019 World Marine Mammal Conference at Four Points by Sheraton Barcelona Diagonal Hotel, Spain. 8th December 2019. Whale and Dolphin Conservation. https://doi.org/10.31219/osf.io/rqudk

Acknowledgements

The editors give heartfelt thanks to all the session chairs, speakers, and other attendees who contributed to the workshop.

Contents

1. Introduction and background

Cetacean data sets collected through citizen science programmes are increasing in number, size, diversity and geographical coverage. They can make important contributions to monitoring, scientific research and public education, and hence contribute to achieving conservation objectives. Such programmes are often able to survey at spatial and temporal scales unachievable by professional research projects (e.g. with more complete temporal coverage). Thus, they can help to inform us about patterns and trends in the distribution and local abundance of different cetacean species, identify areas used for calving and feeding, and alert us to threats and unexpected changes in population status.

For effort-linked cetacean sightings data (i.e. sightings which can be associated with quantifiable survey effort), especially those arising from dedicated at-sea surveys, the Joint Cetacean Protocol (JCP) established standards which should allow robust estimates of distribution, abundance and population trends of cetacean species, focused on the North Sea and adjacent sea regions (Paxton et al., 2016).

However, citizen science data may not always meet the standards expected of traditional scientific research, and there is therefore naturally some resistance, among both management authorities and scientists, to integration and use of such data in official monitoring and research programmes. In the case of citizen science data on sightings of cetaceans from land, the exact nature of data collected often varies geographically, with volunteers supporting locally based programmes. Survey methodology may vary and may differ from what might be scientifically ideal, for example to meet differing volunteer requirements, in terms of time available, age, number of people involved, equipment available, training received, etc. In addition, land-based observations are not easily comparable to at-sea sightings.

The main purpose of this workshop was to look at ways to make the best use of coastal cetacean sightings data in monitoring and research, for example to improve our knowledge about diversity, distribution, local abundance and habitat use of cetaceans, including the detection of temporal and spatial trends. This included consideration of ways to integrate different data sets to provide information at a larger scale, including possible common protocols and data formats to facilitate future integration.

The workshop brought together data providers, analysts and data users, to share their knowledge and experience, and to discuss best practice in training, data collection, quality control and analysis. Discussions included consideration of standards and mechanisms to permit combination of data from different sources, and ways in which citizen science data on cetacean sightings could support conservation and research.

Topics for discussion at the workshop thus included:

- the nature of indicators and targets used in cetacean conservation and the legal basis of monitoring requirements;
- the types of information that can be obtained from sightings data and their utility to detect changes in distribution and abundance;
- approaches to training and motivation of volunteers;
- requirements for logistic support;
- detailed data collection protocols (e.g. number of observers, duration and frequency of watches, designation of observation sites);
- accuracy and consistency of species identification, counts of numbers of animals and recording of the presence of calves;
- the spatial and temporal coverage (extent and resolution) of datasets;
- how to account for factors which influence detectability of cetaceans (e.g. observer experience, fatigue, visibility, sea state, height of the viewing point, topography of the coastline, etc);
- anticipating issues that may arise during statistical analysis, e.g., spatial and temporal autocorrelation (non-independence of different observations that are close together in space and time);
- considerations of precision, accuracy and statistical power to detect differences and changes;
- promotion of the value of citizen science data alongside awareness of the properties and limitations of all datasets.

This workshop report summarises the content of the workshop, including presentations and discussions, and proposes some steps towards standardization, integration, combined analysis and formulation of advice based on such data sets. These outcomes / conclusions are directed at: (a) land-based citizen scientist groups with an interest in recording cetaceans, to recommend core data requirements for combined analysis and a schedule for when analysis should occur, and (b) scientists and policymakers, to highlight how citizen science could assist them.

2. Overview of the sessions

Session 1. Take stock of the various cetacean data sets resulting from land-based citizen science programmes (and wider data sets of interest for purposes of comparison and integration).

Chaired by Graham Pierce (CSIC / IIM) and Alice Walters (WDC)

Presentations:

Presentation: *1.1. Peter Evans: Utilising Land Watch Data in Species Monitoring and Conservation*

Summary of the session:

Participants introduced themselves and a list of participants, including contact details, was collated. The participant list and contact details appear in section 5 of this report. To kick off the day, Peter Evans presented an overview of the long-running Seawatch programme, giving examples of how land-watch data can be used in species monitoring and conservation. This was followed by a group discussion about other land-watch datasets as well as those from other sources. Datasets of interest to the immediate group were submitted through the pre-workshop questionnaire. These were reviewed and additions were made. *A list of these datasets is available in section 8 of this report.*

Further discussion touched on the difficulties of integrating full datasets into a single, overarching database and considered instead the merits of collating *metadata* for all types of cetacean data within a single searchable and accessible database. Rather than compiling the primary data (which was generally thought to be impractical and undesirable), this database would hold metadata and would allow relevant datasets to be easily identified using simple filters (e.g., species, data type, and perhaps geolocation information) and then enable users to request the data in their most up-to-date form through the metadata contact. It was recognised that, in particular, student projects and conference posters / presentations (such as those from SMM and ECS conferences) are not used to their full potential if they are accessible only through a specific institution or conference abstract book and / or if the authors move away from the institution where the work was done. The participants discussed several existing options to house this type of metadata: OBIS Seamap, NBN Gateway, EMODnet, Marine Recorder, MEDIN, Imardis (Bangor University) and potential difficulties (e.g. related to the primary data holder, focus on a specific data source, institution-specific design).

Conclusions and recommendations:

The workshop attendees proposed that both SMM and ECS should request that metadata are submitted when authors submit abstracts to present as a talk or poster.

To be realistic, i.e. recognising that not all abstracts may be accepted and that some abstracts refer to incomplete work, submission of metadata could be requested at the time of delivery rather than at the time of abstract submission. Further, it was suggested that both SMM and ECS should develop a searchable database of all posters, talks and metadata so that they can be more easily found and used.

The participants agreed that it would be useful to identify a host for metadata from all UK (and possibly wider) cetacean data sets. Metadata are less problematic to host and easier to maintain than primary data. It remains necessary to determine the appropriate regional scale, identify a suitable host, and agree on a format for metadata. Suggested metadata fields are presented in section 7 of this report within the *Draft* Best Practice Protocol and Data Standards for Land Watch Data Collection. Once these fields confirmed and a suitable host is found, metadata from all the datasets of interest provided toward this workshop could form the start of this collection of metadata. Recommendations for additional datasets of interest can be emailed to the workshop organisers.

Session 2. Critically assess the potential contribution of such data sets to cetacean monitoring and research needs, including the extent to which they complement other types of monitoring (e.g. boat surveys, acoustic monitoring, strandings monitoring), and their limitations.

Chaired by Nikki Taylor (JNCC)

Presentations:

2.1 Conor Ryan: Merits of Land-Based Cetacean Monitoring: Examples from Ireland

2.2 Karen Hall: How social media is changing the way we learn about cetaceans in remote locations: bringing 60^{o} N sightings to the masses

Summary of the session:

This session assessed the potential of cetacean sighting data from land-based observations to contribute to research and conservation initiatives. Two presentations introduced the session, one giving examples of how effort-based monitoring of cetaceans from land-based watches has been used in Ireland and the other focusing on the contributions of casual sightings as recorded through social media. This second example concerned an area (Shetland) where there is little research on cetaceans and showed how opportunistic reports on social media can demonstrate cetacean presence, seasonality, real-time movements and even links to other regions. This was followed by a group discussion, which considered the potential applications of land-based data, the gaps and limitations associated with this data source, and the potential value of comparison with other datasets. Bullet points from this session (Table 1) were

collected on a white board remained on display throughout the day for consideration during later discussions.

Table 1: Bullet points from the section 2 discussion

Heading	Suggestions
Potential applications	● Seasonal temporal trends ● behavioural trends ● hotspots of feeding ● oceanographic features – interactions with tidal cycles ● monitoring disturbance ● relative abundance (not absolute abundance) ● calving areas ● human impacts ● pre-cursor to more targeted studies ● stewardship – involve communities ● impact assessment and planning ● persistence of species at a site – habitat use ● minimal disturbance to the animals; natural behaviour ● assess population / ecosystem health as top predators ● demonstrating individual ranges – for example: sightings of 'John Coe' as an indicator for the wider Orca population around Scotland ● migratory routes ● daily movements – behavioural states / travelling ● low-cost option to permit collection of land-based sighting data with improved spatial and/or temporal coverage and resolution (e.g. all-year round, annually, at regional scale).
Limitations of land-based data	● Are the watch sites representative of the wider area? ● animals can readily move offshore / away from the observer ● knowledge of observer variability ● inability to watch in bad weather ● seasonal influence on watcher ability and willingness to watch ● daylight hours only ● access / danger of coastline – only access to certain sites ● familiarity with species can influence sightings and create bias ● animals are moving in and out of such small areas – hard or impossible to do absolute abundance ● can be difficult to combine with other data sets ● cannot determine densities - but is this important?
Value of comparison with other data sets	● Coordination / collaboration to see movements at a finer scale ● comparison / integration with opportunistic sightings to fill gaps in time and space ● facilitate targeted photo-ID or acoustic research ● combine with adjoining vessel-based research to include coastal zone

Conclusions and recommendations:

It was generally agreed that data from land-based and, in particular, citizen science data collection programmes can often survey at spatial and temporal scales unachievable through other means. This makes them suitable to contribute to long-term and widespread monitoring initiatives where appropriate and/or as a precursor to more focused studies. Furthermore, they have minimal impact on the study animal and thus are ideal for research looking at natural behaviour or disturbance responses.

There are several factors limiting data collected through land watch studies including: whether the viewable survey area(s) are representative, if the animals can easily move out of sight, the difficulties of watching in bad weather and at distance, and the fact that the data collected are unsuitable for estimation of absolute abundance or densities. However, these studies have the potential to make contributions to understanding trends (seasonal, tidal, behavioural, etc), to identify persistently important 'hot-spot' areas for feeding or calving, and possibly even to calculate relative abundance.

Citizen science programmes can also be limited by the need to consider the ability of a range of observer(s): watch sites must be chosen with an eye to safety and accessibility for a wide range of individuals and the seasonal influence on observer willingness to watch must be considered. However, such initiatives can make important contributions in terms of public education and stewardship by involving local communities.

Land watch data are not always considered for integration with other data sources, but our discussion highlighted that it can be used to fill in gaps in time or space, for example enabling offshore vessel-based data to include coastal zones. As noted above, the wider net of land-based watching can highlight the need more a more focal study and / or facilitate targeted research (photo-ID, acoustic) to achieve better results.

Session 3. Critically assess how data from different surveys and/or different data sources could be brought together for combined analysis as well as possible applications of such analyses.

Chaired by James Waggitt (Bangor University)

Presentations:

3.1 Graham Pierce: Review of analysis of Shorewatch and other land-watch data-sets: techniques, successes and problems

3.2: Gemma Veneruso: The use of acoustics to inform and complement land-based surveys

3.3: Jeroen Creuwels: Citizen science observation on living and stranded cetaceans in the Netherlands

Summary of the session:

This session focussed on the information provided by different data sources/methods, their strengths and weakness, and identifying commonalities which could promote data collation and large-scale analysis. To facilitate these discussions, presentations introduced existing collations of three sources/methods (shore-based surveys, acoustic surveys, strandings), and explained how these collations could be or have been used in large-scale analyses. During the subsequent discussion, workshops delegates were divided into three groups, with each group assigned one source/method. These groups were asked to assess the information which can be collected from their source/method. In these assessments, groups were asked to consider everything provided by each source/method – ranging from relatively simple information such as date and location to more detailed information unique to that source/method. In turn, each group then entered this information on a whiteboard, allowing commonalities and differences to be identified. The information entered on the whiteboard is summarised in Table 2 below. After whole group-review, these details were used through the rest of the day to consider the potential contributions of each source/methods and whether they can be used in comparison with land-based sighting data to fill gaps.

Summary of the presentations

The first presentation described previous analyses of land-based cetacean sightings data from citizen science (or citizen-science assisted) programmes, specifically those run by the Sea Watch Foundation (SWF; UK-wide), Whale and Dolphin Conservation (WDC; Scotland) (“Shorewatch”) and Coordinadora para o Estudo dos Mamíferos Mariños (CEMMA; Galicia, Spain)).

All these programmes generate effort-based sightings data, i.e. we know the duration of watches and the number of observers present and it is usually possible to account in some way (not necessarily in a fully quantitative manner) for the nature of the observation site (e.g. height, field-of-view), observation conditions (e.g. sea state) and the experience of the observers. All three programmes used a series of identified sites distributed around the coast, although the sets of sites used evolved over time (e.g. with some consolidation of adjacent/overlapping observation points, and expansion or contraction of the network of locations over time). In principle, all these programmes operate throughout the year during daylight hours but observations tend to be sparser in the winter months, early morning and late evening. Unlike the other two programmes, Shorewatch is based on observations with a fixed duration (10 minutes) and a minimum time elapsed between the start of successive watches (1 hour).

All three datasets have yielded quantitative descriptions of how presence and/or sightings rate vary within a day, seasonally, from year to year, between sites and in relation to weather conditions. In the case of the two UK-based programmes, the coverage in space and time is somewhat unbalanced, limiting our ability to separate spatial and temporal components of variation in presence and sightings rates. The Galician dataset, although now discontinued, benefitted from the active participation of dedicated (employed) observers who were able to ensure that each site was visited monthly every year.

Using the Shorewatch dataset, seasonal and year-to-year patterns in average sightings rate were described for each species-site combination. However, even if it could be assumed that all watches produced independent data (i.e., ignoring autocorrelation), confidence limits for sightings rates would be wide, and many apparent differences (e.g., between sites) would not be statistically significant. On the other hand, after 6 or 7 watches at a site during a single day or after more than around 70 watches per month, it is apparent that the amount of additional information yielded by further watches is considerably reduced. Given the current protocol, at least for the more commonly sighted species, population declines of around 10% over 6 years (the reporting period for important EU conservation directives) could be detected statistically given a plausible number of watches (say 1000) per year. Technical details will appear in a forthcoming paper but it should be noted that apparent population declines which are evident at only a few sites could be indicative of a distribution shift rather than a fall in abundance. Results of analyses of CEMMA and SWF datasets are available in Pierce et al. (2010) and Evans et al. (2014).

The second presentation discussed the pros and cons of acoustic studies for cetacean monitoring and provided two examples in Wales where acoustic and visual studies overlap. One case study off Bardsey Island, presented preliminary findings of moored SoundTrap data that was collected in conjunction with WDC Risso's dolphin land-based watches. Bad weather limited opportunities to watch from land, with three sightings detected over the two-week period. In contrast the acoustic data recorded high encounter rates of Risso's dolphin click trains, at least once daily over the 12-day

recording period. The vast majority of encounters occurred at night. This short study highlighted that acoustics can reveal critical information such as encounter rates during periods when land-watches are not possible, i.e. in bad weather and at night.

The third presentation reviewed portals / databases for land-based sightings and strandings information in the Netherlands and noted the need for coordination of strandings databases across countries.

Conclusions and Recommendations:

The broad differences in the information provided and in the associated strengths and weaknesses, amongst sources/methods, mean that collation and analysis combining all data from all sources/methods would be difficult. Nevertheless, commonalities in the core-information (date, species, general location, and presence) provided by each method means that similar questions can be addressed using data from each source/method separately. Therefore, within data types (e.g. sightings data), collation and joint analysis of information from all sources is desirable. Results from such analyses can be compared / combined in meta-analyses to identify key changes in relative abundance and distribution. Thus, for example, analysis of acoustic surveys and strandings data can support analysis of shore-based sightings surveys, with both similarities and differences being potentially informative.

Table 2: Bullet points from the section 3 discussion

Data type	At- sea / vessel-based	Strandings	Acoustic
Details collected	• Date • Time • Location • Species • ID confidence • Number of animals. • Behaviour • Group composition • Environmental conditions (sea state, swell, visibility, prevailing weather) • Observer experience/ reliability • Platform height • Search status	• Date • Time • Location • Species • Number of animals • Size • Sex • Maturity / age • Body condition • Cause of death • Photos • Contaminants, pathogens • Genetics • Reproductive status (e.g. pregnant) • Diet • Individual animal ID • Decomposition state	• Date • Time • Presence / absence • Count detections per minute • Start time (encounter /click event) • End time (encounter/click event) • Call type • Towed array (lat-long, speed) • Target distance (if localised) • Depth • Sea temperature • Salinity • Acceleration
Caveats / confounding factors	• Subjectivity of scale/distances • Detection function • Strip width • Response of animals • Sea state	• Effort • Drift • Body condition	• Often unknown and changeable detection area • No knowledge of group size unless an array is used • Variable call rates – absence not always an absence • Classification of obscure species • Feeding events (buzzes)

Session 4. Synthesise recommendations for 'core' data requirements to ensure robust data sets for which future combined analysis is both achievable and useful. Land-based datasets and other datasets.

Chaired by Conor Ryan (Song of the Whale)

Presentations:

4.1: Nikki Taylor: An Overview of the Development of the Joint Cetacean Data Programme (JCDP) and Discussions on Data Standards

4.2: James Waggitt: Data Collection in Shore-Based Surveys

Summary of the session:

The fourth session covered different watch methodologies and the data they provide before considering whether it is possible to describe an 'ideal' land-watch protocol and identify 'core' data that should form part of all data sets and thus can form the basis for strong comparative analysis. The first presentation described the JCDP development process to understand core data and need for data standards when integrating different (ship-based and aerial) data sources. The second presentation introduced the strengths and weaknesses for different land-watch methodologies (continuous watch, scan watch, casual watch), both in practice when conducting watches and when doing analysis (detectability and density from shore-based surveys). Participants were again divided into three groups for discussion. These groups were asked to consider the range of land watch methodologies, identify commonalities and considering previous discussion about where land watch data could be most useful, to see if they could develop 'best practice' protocol recommendations, core data requirements to ensure suitability for data analysis, and suggested data standards to smooth integrating data between projects. Each group reported back their recommendations on a white board, as summarised below (Table 3), and these were then reviewed by the whole group.

Table 3: Bullet points from the section 4 discussion

Ideal Protocol	Core Data	Data Standards
● 10-30 mins ● Once per hour ● Naked eye / binoculars and telescope to ID ● Photographs if possible ● Trained observer (s) ● Scan by segment <= 180º ● Encourage even cover of daylight hours, months, sites ● Designated sites ● Ideally sea state <=3 and visibility >= 1km Data collected ● Species ● No. animals ● Groups ● Distance ● Bearing ● Simple behaviours ● Calves ● Photos ● Confidence of ID ● Record by Dictaphone if busy! ● Sea state (per segment?)	● Date/time for start and end of watch ● Location for watch ● Elevation and eye height ● Environmental variables (visibility and sea state) ● Observer name ● Location of animals ● Species and confidence ● Date & time seen ● Group size ● Group composition (adults and calves) ● Equipment used (binoculars? Scope, magnification?)	● Metadata required and consistent (platform, view, location details, height, etc) ● Effort data ○ Start/end time ○ GMT/BST ○ Equipment used ○ Observer name ○ Environment data ○ Vessel info ○ Feeding birds ● Sighting data ○ Species ○ ID confidence ○ No animals – range (min-max) ○ Age: calf and adult, maybe newborn – more categories might be too difficult ○ Behaviour: need to standardise categories and then include this in metadata

Conclusions and recommendations:

Variations in land watch methodologies result in different strengths and weaknesses but the different approaches all have their value; in practice, these may have developed to support different citizen scientist contributors, regional variations in species or watch platform, or distinct requirements from the data end-users. As was found when considering integrating varied types of data, however, commonalities across all land watch methodologies suggest that there are some core data which should be collected within any methodology to enable robust analysis and support the potential for comparison between different sources of land watch data. Towards describing an 'ideal protocol,' the discussion focused largely on a few different aspects. Firstly, effort-based data collection is preferred, also working to ensure even coverage across daylight hours, seasons, tidal states, and sites with appropriate consideration of where/when environmental conditions (sea state, visibility) might be unsuitable for watching. Secondly, it was recognised that the most "important" sites

are often those surveyed the most frequently, which might artificially increase their apparent importance. Finally, there was considerable discussion about the need to consider observer training when developing a methodology. Particularly if a study includes data collected by numerous different people and/or citizen scientists, the methodology should seek to ensure observer comparability, and analysis should include a review of observer performance – which in turn implies a need to quantify observer performance in terms of ability (e.g. to spot animals and identify species), experience and/or training received.

There was general consensus from the (largely UK-based) land-watch programmes present on the day that it would be beneficial to develop the minutes from this session into a document outlining best practice for citizen science land-watch data collection and data standards. A draft document is presented in section 7 of this report and a subsequent workshop will be organised to review this document. When complete, it is anticipated that this Best Practice document will support existing programmes to help ensure that their data is of a high standard and easily useable for comparative analysis, but it will also inform the development of any new or specifically targeted land watch programmes.

Session 5. Develop a roadmap to achieve better harmonised data collection and combined analysis

Chaired by Karen Hall (NatureScot)

Presentations:

A. Case Study by Dr Conor Ryan: Humpback whales in UK waters

Summary of the session:

(A – Humpback case study) The final session of the day started with a case study considering humpback whale sightings in the UK and how combining datasets has been used to investigate seasonal movements, residency time, entanglement risk and drivers and distribution. This was followed by a discussion about increasing integration of data to improve our understanding of humpback whales in UK waters, discussing who might have data to contribute, and considering whether programmes could set targets to support assessment of humpbacks. Sightings programmes (land-based, vessel-based, opportunistic/social media) can also be used to inform photo-ID efforts and thus inform seasonal movements and residency time as well as linking sightings to wider populations (Cape Verde and West Indies) to better understand potential impacts. Some attendees reflected that they share photo-ID around the relevant groups but there is no formal or generally accepted process. It was noted that humpbacks rarely fluke in the waters around Shetland and so there is value in taking dorsal fin images as well as fluke images.

(B – Roadmap to UK integration) The second half of the final session moved into a wider discussion about better integration of data, with a focus on raising the profile and usefulness of data collected by land watch programmes and setting out a road map to combined analysis in Scotland and/or the wider UK.

Discussion, conclusions and recommendations:

(A – Humpback case study) Happywhale.com is a shared database, already in operation, where photo ID images can be uploaded, identified, and then tracked around the globe. The database contains few data recorded from UK waters although multiple groups have sightings records and photos of humpbacks. It was agreed that it would be valuable to upload humpback photo-ID images into Happywhale to increase awareness of humpback movements worldwide and accessibility of records.

Furthermore, HWDT will be producing some work on humpbacks and will be approaching various groups to integrate their effort-based humpback sighting data.

(B – Roadmap to UK integration) Based on a wider discussion about better integration of data, there was unanimous support for increasing availability of metadata so that all relevant data for any project can be easily identified. It was also agreed that combined analysis of UK data (as a starting point) would raise the profile of land-based data, demonstrate the contributions that it can make and how it can be best utilised, and inform best practice by harmonising data collection and analysis by land watch programmes.

Top priorities were considered to be:

- Agreement of first step analysis of land-watch data at a UK scale (inclusion of casual data as well as effort-based data)
- Identifying / developing a platform to hold metadata for citizen science data sets (and wider) which allows integrated searching and current contact details to enquire about using the base data. (ASCOBANS, ICES, MEDIN all suggested)
- Encouraging ECS / SMM to request inclusion of metadata as part of abstract submission and to develop a searchable database of all metadata, posters, and talks. Workshop representatives from SMM and ECS respectively to approach each society.
- Develop a communication group / forum between data providers, users, analysts, etc., to provide a centralised and transparent location to discuss and support best practice availability, access, and use of citizen science datasets, and to keep access and availability of datasets transparent.
- Increase the focus on proactive communication and feedback to citizen scientists, ensuring transparency about where data are being used and promoting opportunities for input into citizen science data collection programmes.

3. Outputs

Outputs from the workshop and proposed further actions are listed below. It should be noted that some activities originally planned for 2020 did not take place d ue to the pandemic and these need to be rescheduled.

I. Participant list and contact details (section 5 in this report)

II. Spreadsheet listing datasets considered to be of interest and submitted by attendees prior to and during the workshop (section 8 of this report). This list will form the basis of a metadata database once both the metadata proforma and a suitable host are identified. In the interim, suggestions for other datasets of interest can be made to the workshop organisers.

III. *Draft* Best Practice Protocol and Data Standards for Land Watch Data Collection (section 7 of this report) as developed out of discussion in Session 4. There was agreement that this should be reviewed, updated, and agreed at a subsequent workshop / discussion.

IV. Attending organisations support for combined analysis of UK land watch citizen science cetacean data to be reviewed, expanded and agreed in 2020. Once agreed, funding to complete this project will be sought and a timeline for work agreed.

V. Agreement by two workshop attendees to approach SMM and ECS to request that metadata be a requirement when submitting a talk or poster and that each society should maintain a searchable database of all posters, talks, and metadata.

VI. Agreement by attendees to identify and review potential hosts for a searchable metadata database for cetacean data to be discussed when reviewing the *draft* Best Practice Protocol and Data Standards document. (JCDP, ICES, ASCOBANS, MEDIN)

4. Abstracts

(1.1) UTILISING LAND-WATCH DATA IN SPECIES MONITORING & CONSERVATION

Peter G.H. Evans[1,2], Graham J. Pierce[1,3], Aleksandra Koroza[3], Gemma Veneruso[3], Caroline R. Weir[4], & James Waggitt[3]

[1]Sea Watch Foundation, Ewun y Don, Bull Bay, Anglesey, Wales
[2]School of Ocean Sciences, Bangor University, Menai Bridge, Anglesey, Wales
[3]Departamento de Ecología y Recursos Marinos, Instituto de Investigaciones Marinas (CSIC), Spain
[4]Ketos Ecology, Kingsbridge, Devon, UK

Staff from Sea Watch Foundation have been undertaking land watches to monitor cetaceans since the mid-1960s, developing a national network of volunteer observers that persists to the present day. There has been good overall coverage around the British Isles except in SE England, West Scotland and the Borders. Most effort has been between May and September, and since the year 2000.

Land watches can serve a useful function to investigate the relative abundance and distributions of cetacean species in coastal waters. Analyses of c. 84,000 hours of effort from 732 sites spanning a fifty-year time period were undertaken on c. 20,000 harbour porpoise and c. 27,000 bottlenose dolphin records to establish to what extent they can reflect distributions, as well as seasonal and annual trends in population numbers obtained over a wider area from dedicated vessel surveys.

The coastal ecotype of bottlenose dolphin typically occurs within a few kilometres of the shore whereas the harbour porpoise is more widely distributed across shelf seas. Nevertheless, in both cases, longer term trends from land watches mirrored those obtained from wider vessel surveys. Bottlenose dolphins exhibited a general increase, spreading southwards between East Grampian region and Northumberland (n = 56 sites) from 1990 to 2014. Photo-ID mark-recapture studies of the Moray Firth population also showed an expansion from this region southwards towards NE England during this time period. Harbour porpoise numbers per unit effort increased markedly in the region between Northumberland and South Devon (n = 112 sites) over the same time period. The wide-scale line-transect abundance surveys of SCANS (1994) and SCANS II (2005) indicated a similar range shift over the same time period.

Hotspots of coastal occurrence for both species from land watch data highlight important areas requiring protection. They also indicate possible displacement of porpoises by bottlenose dolphins at a national scale with stretches of coastline with concentrations of bottlenose dolphin coinciding with low numbers of harbour porpoise.

Land watches can serve other useful functions. They can help explain ecological relationships with oceanographic features such as tidal currents, provide year-round monitoring of cetacean activities and behaviour, and address a number of issues relating to conservation and management. One such example of the latter can be found in New Quay Bay within the Cardigan Bay Special Area of Conservation established for the important semi-resident bottlenose dolphin population that inhabits the area.

Around 9,000 hours of land watch effort collected between 2006 and 2018 were analysed. The number of boat passages, presence & number of dolphins seen per unit effort were recorded along with type of boat and whether they complied with the local code of conduct introduced in the early 1990s. Coverage was fairly evenly spread between years, and, in 2018, individual boat names were recorded. There were no significant trends in the number of boat passages per hour. The number of dolphins per hour was significantly lower between 2006-10 compared with 2011-18, but has not changed significantly in the last 8 years. The number of boat passages per hour was highest in the morning whereas the number of dolphins per hour showed no significant diurnal trend although highest in early morning and from late afternoon. Dolphins showed different responses to different individual vessels. Compliance to the local boating code of conduct was high (0.85-0.95) and showed no long-term trend. However, the probability of compliance was highest for Visitor Passenger Boats and lowest for Speed Boats. Non-compliance to the code of conduct had a negative effect on responses by dolphins. Both dolphin presence and number of dolphins were significantly negatively related to the number of boat passages per two-hour watch. These findings demonstrate how low-cost land watches, when conducted in a systematic manner over the long-term, can have major benefits to conservation management in terms of increasing our knowledge and understanding of impacts of recreational activities.

(2.1) MERITS OF LAND-BASED CETACEAN MONITORING: EXAMPLES FROM IRELAND

Conor Ryan[1] & Simon Berrow[2]

[1]Song of the Whale team, Marine Conservation Research, 1 Main Street, Kelvedon, United Kingdom

[2]Irish Whale and Dolphin Group, Merchant's Quay, Kilrush, Co. Clare, Ireland

We present an overview of several studies from Irish waters which demonstrate the conservation value of land-based cetacean monitoring. The first case study showed how the occurrence of a small, resident pod of bottlenose dolphins in Cork Harbour was influenced by tidal state (Ryan et al, 2010). A long-term monitoring project, initially under the auspice of ISCOPE, a volunteer-run land-based observer scheme (run by

IWDG) challenged the conventional understanding of the seasonal movements of harbour porpoises[2]. They do not necessarily go offshore during winter months. Despite poorer sea conditions and shorter day length, winter ranks as the second highest season for sighting-per-unit-effort. The ISCOPE data also revealed a sudden decline in local relative abundance of fin whales in inshore waters of the Celtic Sea during autumn months from 2004 – 2009 (Berrow et al., 2010). A nationwide but slow reduction in common dolphin relative abundance was also found. A case study was presented whereby land-based effort data were instrumental in informing more targeted research efforts, which would otherwise have been too costly and time-consuming to conduct. This included a stable isotope study on the diet of fin whales and a photo-identification study on the movements of humpback whales which appear to track the easterly progression of spawning events of herring in inshore waters of the Celtic Sea (Ryan et al., 2016). We emphasise the importance of interpreting shore-based sighting data together with photo-identification data to account for the effect of site fidelity in species such as humpback whales. Although a highly conspicuous species, they appear to occur in low abundance still, which may not be apparent from land-based data alone until site fidelity and a discovery curve can be constructed from photo-identification data. Finally, the question of "are humpback whales increasing in numbers?" was considered. For a recovering population on a feeding ground, one would expect group sizes to increase as cooperative foraging is a well-documented characteristic of humpback whales. IWDG data show that mean group size has remained constant over two decades (mean = 1.5, 1995 to 2015). Furthermore, the proportion of all cetacean sightings that are humpback whales has not increased significantly over the same time period ($R^2 = 0.29$).

Berrow, S., Whooley, P., Firth, L., Knights, A. 2010. *Review of inshore cetacean monitoring programme*. Final Report of ISCOPE II. National Parks and Wildlife Service.

Ryan, C., Rogan, E., & Cross, T. 2010. *The use of Cork Harbour by bottlenose dolphins (Tursiops truncates (Montagu, 1821)).* The Irish Naturalists' Journal, 1-9.

Ryan, C., Whooley, P., Berrow, S.D., Barnes, C., Massett, N., Strietman, W.J., Broms, F., Stevick, P.T., Fernald, T.W. and Schmidt, C., 2016. A longitudinal study of humpback whales in Irish waters. Journal of the Marine Biological Association of the United Kingdom, 96(4), pp.877-883.

(2.2) HOW SOCIAL MEDIA IS CHANGING THE WAY WE LEARN ABOUT CETACEANS IN REMOTE LOCATIONS: BRINGING 60ºN SIGHITNGS TO THE MASSES

Karen Hall [1], Hugh R. Harrop [2], Paul V. Harvey [3], Peter G.H. Evans [4] & Chiara G. Bertulli [4]

[1] Scottish Natural Heritage, [2] Shetland Wildlife, [3] Shetland Amenity Trust, [4] Sea Watch Foundation

Shetland has long been an important area for cetaceans and seals due to its proximity to the European continental shelf. However, the isolated and exposed position at 60o N also makes it difficult for surveying with most data collection restricted to land-based surveys by citizen scientists. The last few years have seen substantial increases in sightings around Shetland. Since 1970, there have been over 4500 cetacean sightings with 1600 (36%) in the last decade. The question arises whether this is a genuine increase or a consequence of improved communication networks. Are we simply getting better at reporting their presence around our coastal waters? Here we assess the sighting trends of five regular species: harbour porpoise, Risso's dolphin, killer whale, minke whale, and humpback whale. For the larger species, there are sighting increases that correspond with when particular social media platforms came on stream. However, for harbour porpoise there is an opposite trend with citizen scientists only reporting unusual group sizes. A comparative analysis with effort corrected density data found that these large increases were unsupported. Sightings data were also examined to better understand site fidelity and movements. Photo ID was formerly restricted to scientists but through camera improvements and the power of social media, the quality and availability of photos has improved dramatically. Increasing social media reporting (Shetland Orca Sightings Facebook 17,000 followers and the WhatsApp group 256 users) means it is now possible to track killer whale pod movements in real time, or determine whether sightings of humpback whales or Risso's dolphins, days, or even years apart, are the same individuals. These findings show that social media can play an important role in both future conservation and public awareness of cetaceans but that data collection methods need to evolve to work with these platforms.

(3.1) REVIEW OF ANALYSIS OF SHOREWATCH AND OTHER LAND-WATCH DATA-SETS: TECHNIQUES, SUCCESSES, AND PROBLEMS

Graham J. Pierce[1,2,3], Paula Gutierrez[1], Alice E.M. Walters[4], Sarah J. Dolman[4], Peter G.H. Evans[5], Pia Anderwald[5], M. Begoña Santos[6], Angela Llavona[7], Jose Cedeira[7], Alfredo Lopez[7] & Mara Caldas[7].

[1]Instituto de Investigaciones Marinas (Consejo Superior de Investigaciones Científicas), Spain [2]OceanLab, University of Aberdeen, UK
[3]CESAM & Departamento de Biologia, Universidade de Aveiro, Portugal
[4]Whale and Dolphin Conservation, UK
[5]Sea Watch Foundation, UK
[6]Instituto Español de Oceanografía, Spain
[7]Coordinadora para o Estudio dos Mamíferos Mariños, Spain

Land-based cetacean sightings data collected through citizen science projects offer a valuable data source for research and environmental monitoring. One such dataset arises from the WDC Shorewatch programme which has been running at an increasing number of Scottish coastal sites over the last decade and uses standardized 10-minute watches separated in time by at least 1 hour. It provides data on species present and spatiotemporal patterns and trends in their distribution and local abundance. Nevertheless, the spatial distribution of survey effort is uneven and more observations are carried out in the summer months and in the middle of the day than at other times. Another important issue is temporal autocorrelation. At any one site, sightings rates obtained during different watches on the same day tend to be corrected with each other. This can be managed by combining data for each site-date combination. These site-date data are more consistent when more watches have been joined together for each site date. Over the six-year reporting cycle of the EU MSFD, the smallest amount of decline that is statistically detectable depends on the underlying sightings rate, such that the data will be of most use for monitoring bottlenose dolphins. Two other data sets examined are characterized by variable watch duration, which permits exploration of the value of watches of different durations. In the case of a Galician data set, sightings rate tends to be lower for short watches and to reach an asymptote after 50-60minutes of observation. The optimal duration of a watch depends on the travel time between watch sites.

(3.2) THE USE OF ACOUSTICS TO INFORM AND COMPLEMENT LAND-BASED SURVEYS

Veneruso, G.J. & Risch, D.

The COMPASS (Collaborative Oceanography and Monitoring for Protected Areas and SpecieS) project will deliver the first fully coherent network of passive acoustic monitoring buoys across the regional seas of the Republic of Ireland, Northern Ireland and West Scotland. Integrating the longest continuously maintained oceanographic monitoring stations in Europe within a network of new buoys equipped with oceanographic sensors, acoustic recorders and advanced fish tracking technology, this exciting and innovative project will build cross-border capacity for effective monitoring and management of Marine Protected Areas (MPAs). The project will develop long-term monitoring strategies for highly mobile protected species such as marine mammals and salmonids, and provide essential infrastructure for baseline oceanographic and ambient noise monitoring. The development of observational and data management capacity across the region will be complemented by the delivery of three truly regional scale environmental models, designed to support the management of a cross-border MPA network.

The combination of visual and acoustic monitoring is powerful: visual data can give confidence in acoustic identifications, informing about missed acoustic detections, behavioural context, numbers of animals in relation to number of calls, and on human behaviours. Acoustic data can extend the time and seasonal coverage of monitoring.

The second presentation discussed the pros and cons of acoustic studies for cetacean monitoring and provided two examples in Wales where acoustic and visual studies overlap. A case study off Bardsey Island has provided information on functional use of the Bardsey Island area by Risso's Dolphins, based on moored SoundTrap data that were collected in conjunction with WDC's land-based watching. Bad weather limited the land-based watching opportunities, with three sightings detected over the two-week period. In contrast the acoustic data recorded high encounter rates of Risso's dolphin click trains, at least once daily over the 12-day recording period. The vast majority of encounters occurred at night. This short study highlighted that acoustics can reveal critical information such as encounter rates during periods when land-watches are not possible, i.e. in bad weather and at night. Shore-based watches are potentially massively underestimating encounter rates for nocturnal species. Further information related to foraging can be obtained using an echosounder to detect prey species - although this does not work well for squid, krill and other cephalopods are detected. The second example concerned work off Anglesey in tidal energy development areas used a combination of sound traps, vessel-based visual and acoustic observations, and land-based observations.

(3.3) OVERVIEW OF LAND-BASED OBSERVATIONS ON LIVING AND STRANDED CETACEANS IN THE NETHERLANDS

Jeroen C.S. Creuwels[1,2] , Guido O. Keijl[2]

[1]Workgroup Marine Mammals, Dutch Mammal Association, Nijmegen, The Netherlands, email: *zee@zoogdiervereniging.nl*
[2]Naturalis Biodiversity Center, Leiden, The Netherlands

In the Netherlands there are three portals where most land-based cetacean sightings are recorded. First, many observations are recorded on the portal *waarneming.nl* (Dutch version of observation.org). This portal is very popular and used by a wide range of amateurs and professionals for recording nature observations, among which cetacean and seal sightings. Second, the portal *walvisstrandingen.nl* which only records cetacean strandings on the Dutch coast. It is mostly used by dedicated amateurs patrolling the Dutch shoreline for living and dead cetaceans on the beach. Third, the portal *trektellen.nl*, which is used by dedicated and experienced amateurs who monitor coastal and seabirds in standardised one-hour counts from fixed observation points along the coast. Seals and cetaceans are also included in these counts.

The portal *walvisstrandingen.nl* is maintained and validated by an expert at Naturalis Biodiversity Center, who checks each entry. Also, all available information on Dutch strandings is collated from other sources such as observation.org, Facebook and other social media. In this way almost all cetacean strandings are recorded along the Dutch shoreline, at least since 1920! The type of information that comes from strandings data (and post-mortem investigations) is very different from but complimentary to observational data. Especially when strandings data (such as gender, size, phase and condition) over a wide geographical range are added to data on observed numbers, new insights about population and distributional patterns can be obtained. An example is given for the North Sea, with stranding data from Scotland, England, Belgium, The Netherlands, Germany and Denmark. It is recommended that various national strandings programs collaborate and standardise their databases. The value of a validated strandings database is recognised by the government of the Netherlands, which regularly publishes factsheets on a wide variety of environmental indicators (which are compiled in a national Environmental Data Compendium). For the indicator species the harbour porpoise, the factsheet includes data from *walvisstrandingen.nl,* as well as the standard annual aerial surveys over the Dutch part of the North Sea and coastal monitoring data by *trektellen.nl*.

(4.1) AN OVERVIEW OF THE DEVELOPMENT OF THE JOINT CETACEAN DATA PROGRAMME (JCDP) AND DISCUSSIONS ON DATA STANDARDS

Nikki Taylor[1] and Kelly Macleod[1]

[1]Joint Nature Conservation Committee (jncc.gov.uk).

Aim of the Joint Cetacean Data Programme

Mobile marine species such as cetaceans, inhabit an environment which is challenging to survey and observe, which limits understanding of their complex ecology and biology. We rely on evidence to inform decisions on sustainable sea use; welfare; population status and emerging and changing pressures such as climate change. Decisions still need to be made even in the absence of a good evidence base, but a lack of understanding has an impact on the effectiveness of decisions on how we manage our seas for the benefit of all ecosystem functions. The Joint Cetacean Data Programme (JCDP) aims to provide a growing, collated dataset of at-sea cetacean survey data to serve as a resource to inform conservation, research and policy needs.

Previous work such as the Joint Cetacean Protocol (JCP) and Marine Ecosystems Research Programme (MERP) has demonstrated the value of collating these data, with outputs contributing to, for example, the identification and designation of Marine Protected Areas for harbour porpoise, and in Scotland for minke whale and Risso's dolphin.

The need for a data standard

For datasets to be collated into a single database, there needs to be commonality at least for some fields within the data. Having a defined standard for data collection will ease that process in the future. Standardising data offers advantages in four main areas: development of expertise and data quality, suitability of data for analyses, ease of data ingestion, and compatibility. These components all contribute towards maximising the use of independent datasets.

Development of expertise and data quality

Where field data collection is standardised, surveyor expertise increases due to familiarity with an agreed, robust protocol. As a result, data quality improves through coordinated use of an appropriate survey methodology, as well as improved ability of observers to carry out accurate and effective surveys.

Suitability of data for analyses

Standardisation of data at collection stage ensures that the requisite variables are collected using appropriate methods. It also ensures that the data are recorded using the correct naming convention, taxonomy, and other associated coding. Recording supporting information e.g. spatial and environmental data, will also be controlled, ensuring these data are suitable for use either independently, or collaboratively alongside other datasets.

Data ingestion process

Standardisation of datasets allows for automated upload and validation of data, saving time for both data owners and those hosting the data. The validation ensures that only those data that meet the standard will be stored, helping to maintain the quality of the data within the database.

Compatibility of datasets for combined use

Standardised data ease the burden of standardisation on those using the data. Data may need to be adapted to be used in analytical packages such as 'R' or 'Distance' but applying these adaptations from collection rather than to individual datasets further along the line, speeds up the process, eases data processing congestions and reduces opportunity for error.

https://jncc.gov.uk/our-work/joint-cetacean-data-programme/

(4.2) DATA COLLECTION IN SHORE-BASED SURVEYS

J.J Waggitt[1]

[1]School of Ocean Sciences, Bangor University, Menai Bridge, UK

Although sharing broadly similar methodology, a variety of shore-based survey designs are used in research and monitoring. This presentation summarises these variations and highlights potential issues when combining data in large-scale analyses investigating distributions or numbers. Observations can be either continuous watches, scan watches, or casual watches. Measurements of animals could be either counts, distributions or focal followers. It is suggested that continuous and scan watches recording counts and distributions can be combined if the area and time covered can be calculated. However, issues in detectability and spatiotemporal coverage need overcoming if these calculations are to be made. This is challenging in surveys covering coastal environments where the distribution and detectability of animals is particularly dynamic. Establishing the limitations of shore-based surveys are important when conceiving aims of large-scale analyses. However, the ability to collect data across different habitats, regions, seasons and years makes these approaches suitable for addressing important research and monitoring questions.

(5) HUMPBACK WHALES IN THE UK: VAGRANT OR NATIVE?

Conor Ryan[1]

[1]Song of the Whale team, Marine Conservation Research, 1 Main Street, Kelvedon, United Kingdom

Knowledge on the ecology of humpback whales in the eastern North Atlantic is lacking by comparison with their western counterparts. Humpback whales are considered ‘vagrant’ in UK waters and yet regular sightings suggest that they are more than just wanderers. While the rate of humpback whale sightings per unit time is increasing in UK waters, this seems to reflect an increase in sighting effort and ease of reporting due to smart phone ownership, social media and more dedicated survey effort (e.g. Shorewatch). I explore a number of ways to integrate different sources to address the question: are humpback whales vagrant in the UK? The seasonal movements of humpback whales in the UK are similar to those in Irish waters with a marked reduction or absence in March and April (photo ID + land-based sightings). Entanglement is the greatest known threat to humpback whales in Scottish waters. Entanglement risk is so high in the west of Scotland that the area represents a population sink for humpback whales (fishing effort + land-based sightings). Finally, I presented data on some apparent drivers of whale distribution in the Celtic Sea, using humpback whales and herring as an example (fisheries data + land-based sightings). Considering all available data and published findings, I conclude that humpbacks ought to be considered native rather than vagrant. With the available information, it is not yet apparent if they are recovering in the region, as in the western North Atlantic. Considering the significant threat of entanglement (12% in inshore waters around the Hebrides), the precautionary principle ought to be applied until such time as reliable abundance estimates can be made for humpback whales in this region.

5. Participants List

Tom Bean	N/A
Jeroen Creuwels	Dutch Marine Mammal Workgroup
Becky Dudley	HWDT
Katie Dyke	WDC
Sonja Eisfeld-Pierantonio	WDC
Peter Evans	Seawatch Foundation
Karen Hall	NatureScot
Lauren Hartny-Mills	HDWT
Nicola Hodgins	WDC
Stefanie Krafft	Bangor University
Alison Lomax	HWDT
David Pfender	WDC
Graham Pierce	IIM (CSIC)
Susanna Quer	N/A
James Robbins	University of Portsmouth
Conor Ryan	MCR (Song of the Whale)
Anne Simonis	Contractor to NOAA Fisheries
Nikki Taylor	JNCC
Gemma Veneruso	Bangor University
James Waggitt	Bangor University
Alice Walters	WDC

Unable to attend but contributed presentation slides: Denise Risch, SAMS

6. Agenda for the day

8.30am – 8.50am	Arrive and register
9.00am – 9.55am	Session 1: Taking stock of land-based citizen science cetacean data 9.00 Introduction to venue 9.05 Round table & introductions 9.30 **UK's longest running cetacean citizen science programme**, Peter Evans 9.40 Full group discussion
9.55am – 10.45am	Session 2: Assessing potential / priority contributions of land-based data 9.55 **The merits of land-based effort and sighting schemes**, Conor Ryan 10.05 **Potential data applications. SNH / Northern Isles**, Karen Hall 10.10 Small group discussions: best applications, gaps, limits, comparisons 10.30 Bring together ideas, highlight priorities for focus in later sections
10.45am	*Coffee break – provided*
11.15am – 12.45pm	Session 3: How to achieve combined & comparative analysis 11.15 **Shorewatch report, considerations for combined analysis**, Graham Pierce 11.25am **Short talks about combining with other data sources** Various 11.45am Small groups: analysis techniques by data source, region, application 12.25pm Full group discussion: review suggested approaches
12.45pm – 1.45pm	*Lunch – provided* *Potential lunchtime chat on science learning in citizen science with Anne Simonis, NOAA*
1.45pm – 3.10pm	Session 4: Protocols and data standards 1.45 **Review JCDP development & data standards**, Nikki Taylor 1.55pm **Protocols, standardisation, detection and more**, James Waggitt 2.05pm Small group discussions: ideal protocol, core data, data standards 2.40 Full group discussion: mandatory core data, additional requirements for comparison with other data types
3.10pm	*Comfort break*
3.20pm – 4.45pm	Session 5: Roadmap toward combined analysis 3.20pm **Humpbacks, a case study to road map:** discussion led by Conor Ryan 3.50pm Small group discussions to develop roadmaps based on priorities from this morning 4.25pm Review with the wider group, confirm next steps
4.45pm	Wrap-up, final thoughts, goodbye

7. *Draft* Best Practice Protocol and Data Standards for Land Watch Data Collection

As discussed during the *Taking stock of land-based citizen science cetacean data sets: regional integration and comparative analysis to deliver policy-relevant results* workshop held alongside the 2019 World Marine Mammal Conference.

Background and aims

In December 2019, a workshop was held alongside the World Marine Mammal Conference in Barcelona on the topic: Taking stock of land-based citizen science cetacean data sets: regional integration and comparative analysis to deliver policy-relevant results. The main purpose of this workshop was to look at ways to make the best use of coastal sightings data in monitoring and research, for example to improve our knowledge about distribution and local abundance of cetaceans and to allow us to detect temporal / spatial trends in cetacean distribution, habitat use, abundance and diversity. This included discussing ways to integrate different data sets to provide information at a larger scale and hence consideration of possible common protocols and data formats to facilitate future integration.

The workshop aimed to achieve this through the sharing of knowledge and experience, bringing together data providers, analysts and data users to consider best practice in training, data collection, quality control and analysis, including how and when combination and comparison with other data sources would be useful, and how to use such data to support conservation and research.

Cetacean data sets collected through citizen science programmes are increasing in number, size, diversity and geographical coverage. They can make important contributions to monitoring (in support of conservation), scientific research, and public education. Such programmes are often able to survey at spatial and temporal scales unachievable by professional research projects and they can inform us about patterns and trends in the distribution and local abundance of different cetacean species, help identify areas used for calving and feeding, and alert us to threats and unexpected changes in population status.

Nevertheless, the resulting data may not always meet the standards expected of ‘proper’ scientific research, and there is therefore some resistance, among both management authorities and scientists, to integration and use of such data in monitoring and research programmes. In the case of citizen science data on sightings of cetaceans from land, the exact nature of data collected often varies geographically, with volunteers supporting locally-based programmes. Survey methodology can also vary, for example to meet differing volunteer requirements, e.g., in terms of time available, age, group vs individual, equipment, training, etc.

This draft document proposes some steps towards standardization and best practice for land watch protocol and data requirements as discussed during the workshop. It forms the basis for subsequent discussion and agreement. The layout of essential/recommended data fields is based on a combination of NBN, JCDP and GIgateway to encourage consistency between programmes and data compatibility.

Basic protocol requirements as suggested during the workshop

This forms a starting point for considering land-based data collection

- Data collection programmes should ensure that all observers should receive consistent training around protocol, equipment, and species ID.
- Observer training, particularly when using citizen scientists, should be followed up with support materials and subsequent refresher trainings where possible.
- Effort-based observations where possible but suggested fields and vocabulary can also apply to opportunistic sightings where possible.
- Consider whether focusing on designated sites might increase effort in those areas.
- Scans with naked eye / binoculars. Access to a telescope to assist with species ID
- Scan in segments where the field of view is large to encourage consistency between observers
- Even coverage of daylight hours, seasons, tides, sites.
- Ideally sea state <=3, visibility allows entire survey area to be viewed.
- Duration of the watch. Minimum survey length? Maximum length? Consider sufficient effort against observer fatigue. Multiple efforts over a given period?
- Photographs where possible

Proposed essential data fields and vocabulary

Suggested core data fields which are essential to enable effective use of the data for many applications (e.g. species, locations), and others which are not necessarily essential to answer key questions and complete primary analyses (e.g. observer name) are listed below under four headings: metadata, effort, environmental data and observations – species, number, behaviour.

Metadata - Providing essential data to identify the dataset

	Field name	Description	Required?
About the data			
1	Data set Title	Name by which the dataset is known	
2	Data source	Organisation with primary responsibility for the data	
3	About the data provider	Brief summary about the data provider	
4	Purpose of data collection	The purpose for which the records were originally collected	
5	Data collection methodology	A concise description of the methods and techniques used in the creation of the dataset	
6	Observer training	Are observers trained? If so, how?	
7	Data quality control	An assessment of the quality of data capture and/or the collated dataset as a whole	
Temporal Extent			
8	Start date	Date on which data was first collected	
9	End date	Date on which data was last collected	
10	Data collection status	Status of data collection since last data collection point (ongoing, complete, etc.)	
11	Frequency of data collection	Frequency at which data are collected (year-round, seasonal, etc.)	
12	Frequency of data update	Frequency at which data holdings are updated	
Spatial Extent			
13	National extent	Geographic extent of dataset - UK or beyond	
14	Regional extent	Geographic extent or occurrence of dataset by administrative area, repeated sites can be listed here	
15	Spatial referencing system	Nation of the spatial referencing system used in the dataset (e.g. Latitude and longitude)	
16	Coordinate reference system	Name of the coordinate referencing system used in the dataset (e.g., WGS 84)	
17	West Bounding Coordinate	Westernmost coordinate of the survey area	
18	East Bounding Coordinate	Easternmost coordinate of the survey area	
19	North Bounding Coordinate	Northernmost coordinate of the survey area	
20	South Bounding Coordinate	Southernmost coordinate of the survey area	
Data Provider – *where possible, use contact details for a role rather than an individual person*			
21	Title	Data providers nominated contact (role / title)	
22	Address	Data providers full address	
23	Telephone number	Data providers telephone number	
24	Email address	Data providers email address	
25	Web address	Data providers web address	
Data Use			
26	Terms of use	Data providers terms of use for this dataset	
27	License	The licence that should be applied to all records in the dataset (CC0, CC-BY, CC-BY-NC or OGL)	
28	Restrictions	Restrictions for accessing the dataset	
29	Citation	Citation required for the dataset	
30	Data format	Format in which data can be provided	

31	Additional Information	Further information about the dataset i.e. publications or online resources	
32	Associated datasets	Related datasets or description of data resources which may be used in association with the subject dataset	
33	Keywords	Keywords summarising the subject of the dataset	
34	Metadata updates	Date on which the Metadata was last changed, and by whom	

Effort - Effort is essential for certain types of analysis and recording effort is therefore recommended as part of a survey protocol where possible to enable maximum potential use of the data. However, data without effort records may still be used in basic analysis of e.g., distribution.

	Field	**Description**	**Required?**
Effort data			
	Site name	Name of the location from where the watch is occurring	
	Site height	The (eye/platform) height of the observer above sea level (metres)	
	Site location	Latitude / longitude (decimal degrees)	
	Date	Date of the survey	
	Start time	Start time of survey effort (hh:mm)	
	End time	End time of survey effort (hh:mm)	
	Primary search method	How was the search primarily completed? E.g. naked eye; binoculars	
	Equipment used	Binoculars? Scope? Magnification? Naked eye?	
	Observer name	Name or other reference	
	Observer experience	If applicable, define the level of observer performance	
	No. of observers		
	Vessels observed	This varies greatly between programmes (number, type, distance, location, activity, etc)	
	Bird / other species	Could be included here or in sightings	
	Comments		

Environmental data - The environmental conditions whilst surveying have an impact on probability of observing animals and is therefore essential information to record alongside observations.

	Field	**Description**	**Required?**
Environmental data			
	Sea state	Sea state using the Beaufort Sea State scale – should / could this vary across survey area or be averaged?	yes
	Swell/wave height	In metres – noted during workshop: this can be inconsistent between observers (optional)?	
	Visibility	Greatest distance at which objects can be distinguished. Is the survey area impacted by fog or other	yes
	Weather	Predominant weather conditions	
	Glare	Location of search area affected by glare (optional?)	

Observation – species, number, behaviour - Accurate recording of a sighting is essential to the dataset.

	Field	Description	Required?
Observation data			
	Effort-based?	Is this sighting linked to an effort? If no, define observer, date, location, equipment.	Yes
	Start time of sighting	The time the sighting was first made	Yes
	End time of sighting	The time when the sighting ended	
	Species	Pre-defined species codes	Yes
	ID confidence		Yes
	Number	Number of individuals of a species in each sighting – This should at least include a minimum count but could also include a range: min, max, best?	Yes
	Group composition	Consider what is possible for observers. Suggest adult, calf, (maybe) newborn as basic categories which have value for defining site use. Where possible, juveniles and sub-adults may also be identified but this is less likely to be consistent between observers. All unknown considered adult?	
	Range	Distance of animal(s) in metres from the observer ait first sighting	
	Bearing	Position of the animals at first sighting, recorded in relative degrees	
	Behaviours	Standardised categories work best for comparison across datasets.	
	Direction of travel	If appropriate	
	Comments	About the sighting	
	Photos?	If any were taken, location, etc.	
	ID confirmation		

8. Datasets of interest suggested by attendees

Data set title	Organisation	Contact details	Temporal coverage	Geographic coverage	Data source type(s)	Data format
ADRIFT in the California Current	NOAA Southwest Fisheries Science Center		2020 - ongoing	*US West coast*	*drifting acoustic recording buoys*	
Orca	*ORCA*		*2006 onwards March – October for boat-based. 2019 onwards for stationary platforms (at sea, rather than land-based but we will be using similar techniques and analyses)*	*Focus around UK and surrounding waters, with additional boat-based data in Arctic, Antarctic, North & South Atlantic, North Pacific*	*Cetacean sightings and survey effort from ships of opportunity and fixed platforms*	
CEMMA land watches	*CEMMA*		2003 to 2011, possibly more recent too	Coast of Galicia	CEMMA volunteers	
Happywhale.com	Happywhale.com			World wide	photo-ID	
Allied Whale	Allied Whale					
IWDG	IWDG				varied	
Walvisstrandingen.nl	*Naturalis*		*1920 - present*	*Coasts of The Netherlands*	*observations*	
Bardsey Island Soundtrap data	*SEACAMS, Bangor University*		*12 days, continuous recording*	*Two recorders, NW and SE of Bardsey Island, Wales*	*Acoustic*	
South Stack porpoise tracking study	*SEACAMS, Bangor University*		*Approx. 1 month 2018*	*West Anglesey, Wales, up to 1.5km offshore*	*Visual, video records*	
West Anglesey cetacean surveys	*SEACAMS, Bangor University*		*2 years Jan. 2015- Dec. 2016. Monthly surveys*	*West Anglesey, Wales*	*Visual, towed acoustic*	

Data set title	Organisation	Contact details	Temporal coverage	Geographic coverage	Data source type(s)	Data format
South Stack Soundtrap data	*SEACAMS, Bangor University*		*2 X 1month, summers 2017 and 2018*	*West Anglesey, Wales, 7 X recorders*	*Acoustic*	
Bardsey Island	*WDC*	data@shorewatch.org	*1999 – present (with some gaps)*	*The waters around Bardsey Island and the Llŷn peninsula*	*Effort-based, land watch data*	Excel
Land-based data	BIAZA and/ or Sea Watch		*2018-2019*	*Point Lynas, Anglesey*		*Currently pdf files*
Boat-based data	*BIAZA and/ or Sea Watch*		*2018-2019*	*North Wales (Rhos-on-Sea to Bull Bay, Anglesey)*		*Currently pdf files*
NORCET project		*Sea Watch Foundation, Colin D. MacLeod (GISinEcology), Ian Hay (East Grampian Coastal Partnership), Graham Pierce (IIM), Susanna Quer, and Kevin Hepworth (South Grampian Seawatch Group).*	2004 to 2019 (latest data available from 2015)	NE Scotland to Northern Islands (Shetlands and Orkney)	Ferry-based surveys	
UK Land-based effort & sightings	*Sea Watch Foundation*		*1965-present*	*UK-wide*		*Access & Excel*
WDC Shorewatch Programme	WDC	data@shorewatch.org	*2005-present*	*Specified sites around northern scotland*	*effort-based, land watch sightings*	Bespoke database
Risso's photo-ID on Lewis	*WDC*	data@shorewatch.org		*Northwest Scotland, 'The Minch'*	*Photo-ID, visual, acoustic (PODs & Soundtraps)*	Excel

9. Bibliography of publications related to this discussion as suggested by participants

de Boer MN, Simmonds MP, Reijnders PJH, Aarts G., 2014. The Influence of Topographic and Dynamic Cyclic Variables on the Distribution of Small Cetaceans in a Shallow Coastal System. PLoS ONE 9:e86331. https://doi.org/10.1371/journal.pone.0086331

Embling, C.B., Walters, A.E.M., Dolman, S.J., 2015.How much effort is enough? The power of citizen science to monitor trends in coastal cetacean species. Global Ecology and Conservation, Volume 3, pp867-877.

Evans, P.G.H., Pierce, G.J., Veneruso, G., Weir, C.R., Gibas, D., Anderwald, P. & Begoña Santos, M. (2015) Analysis of long-term effort-related land-based observations to identify whether coastal areas of harbour porpoise and bottlenose dolphin have persistent high occurrence and abundance (revised June 2015), JNCC Report No. 543, JNCC, Peterborough, ISSN 0963-8091.

Goetz, S., Read, F.L., Ferreira, F., Martínez Portela, J., Santos, M.B., Vingada, J.V., Siebert, U., Marçalo, A., Santos, J., Araújo, H., Monteiro, S., Caldas, M., Riera, M. & Pierce, G.J., 2015. Cetacean occurrence, habitat preferences and potential for cetacean-fishery interactions in Iberian Atlantic waters: results from cooperative research involving local stakeholders. Aquatic Conservation: Marine and Freshwater Ecosystems 25, 138-154.

Llavona, A., 2018. Population parameters and genetic structure of the harbour porpoise (*Phocoena phocoena*, L. 1758) in the northwest the Iberian peninsula. PhD thesis, Universidade de Aveiro.

National Academies of Sciences, Engineering, and Medicine 2018. Learning Through Citizen Science: Enhancing Opportunities by Design. Washington, DC: The National Academies Press. https://doi.org/10.17226/25183.

Paxton, C.G.M., Scott-Hayward, L., Mackenzie, M., Rexstad, E. & Thomas, L., 2016. Revised Phase III Data Analysis of Joint Cetacean Protocol Data Resource, JNCC Report No. 517, JNCC, Peterborough, ISSN 0963-8091.

Piatt, J.F. and Methven, D.A., 1992. Threshold foraging behavior of baleen whales. Marine Ecology Progress Series, 84, pp.205-210.

Pierce, G.J., Caldas, M., Cedeira, J., Santos, M.B., Llavona, A., Covelo, P., Martinez, G., Torres, J., Sacau, M. & López, A., 2010. Trends in cetacean sightings along the Galician coast, north-western Spain, 2003–2007, and inferences about cetacean habitat preferences. Journal of the Marine Biological Association of the United Kingdom 90, 1547-1560.

Quer Imedio, S, 2014. Identification of ecogeographical (EGV) variables associated with occurrence of Harbour Porpoise (*Phocoena phocoena*) off the NorthEast coast of Scotland. BSc thesis, Universidad de Barcelona.

Robertson, F., Wood, J., Joslin, J., Joy, R., and Polagye, B., 2018. Marine mammal behavioral response to tidal turbine sound, Final technical report for DE-EE0006385.

Ryan, C., Rogan, E. and Cross, T., 2010. The use of Cork Harbour by bottlenose dolphins (Tursiops truncatus (Montagu, 1821)). The Irish Naturalists' Journal, pp.1-9.

Ryan, C., Leaper, R., Evans, P.G., Dyke, K., Robinson, K.P., Haskins, G.N., Calderan, S., van Geel, N., Harries, O., Froud, K. and Brownlow, A., 2016. Entanglement: an emerging threat to humpback whales in Scottish waters. In Scientific Committee Meeting of the International Whaling Commission.

Ryan, C., Whooley, P., Berrow, S.D., Barnes, C., Massett, N., Strietman, W.J., Broms, F., Stevick, P.T., Fernald, T.W. and Schmidt, C., 2016. A longitudinal study of humpback whales in Irish waters. Journal of the Marine Biological Association of the United Kingdom, 96(4), pp.877-883.

Weir, C., Hodgins, N., Dolman, S., & Walters, A. (2019). Risso's dolphins (Grampus griseus) in a proposed Marine Protected Area off east Lewis (Scotland, UK), 2010–2017. *Journal of the Marine Biological Association of the United Kingdom, 99*(3), 703-714. doi:10.1017/S0025315418000516

All Seawatch publications:
https://drive.google.com/file/d/1Eo7is3d4OXwwOL_wyvWQ1f7dl7Np9Juk/view?usp=sharing

https://www.uu.nl/sites/default/files/ijsseldijk_et_al._2018._spatiotemporal_analysis_of_white-beaked_dolphin_strandings._lutra_61_002.pdf