ZoaTrack - an online tool to analyse and share animal location data: User engagement and future perspectives

Peggy Newman[1], Ross G. Dwyer[2], Lee Belbin[1], Hamish A. Campbell[3]

[1] Atlas of Living Australia, CSIRO National Collections and Marine Infrastructure, Canberra, ACT, Australia, 2601

[2] School of Biological Sciences, University of Queensland, St Lucia, QLD, Australia, 4072

[3] Research Institute for Environment and Livelihoods, Charles Darwin University, NT, Darwin, Australia, 0909

ABSTRACT

ZoaTrack.org is a web application for visualising and analysing animal biotelemetry data. This online facility offers a robust set of free and highly accessible tools, which enables the non-specialist and specialist to better manage, store, visualise and analyse animal location data. The majority of the data stored in the facility has been collected by animal-borne satellite biotelemetry devices (i.e. GPS and ARGOS), but other types of individual-based location data (mark-recapture, VHF, Acoustic) have also been uploaded to Zoatrack.org. The platform is in its 8th year and this paper presents a review of the platform's user metrics and uptake by the research community. We also discuss the challenges in trying to sustain this type of research support facility and opportunities for the future.

Key words: animal tracking, bio-logging, biotelemetry, data visualisation, biodiversity repository

DOI: https://doi.org/10.7882/AZ.2019.031

Introduction

The technological advancement of animal-borne biotelemetry devices has broadened the scope of their zoological research applications. As a result, there has been a substantial increase in the number of biotelemetry devices deployed around the globe to address questions related to animal movement and habitat usage (Campbell *et al.* 2015). Not only have the number of deployments increased, but the volume and types of data that these devices collect have also increased significantly. Researchers are compiling ever larger datasets across ever greater spatial and temporal scales leading to challenges in analysing the data, as well as storing the data collections in perpetuity. To support this demand, a number of e-infrastructures have appeared in recent years to host animal telemetry datasets (Campbell *et al.* 2016). These e-infrastructures assure the permanence and integrity of these data for future use, and enable scientists to share data within and between disciplines. They also assist to satisfy journals and funding agencies who are increasingly encouraging or requiring the raw dataset from research projects to be archived in a discoverable and reusable manner.

In 2011, the ZoaTrack platform (formerly OzTrack) was created to facilitate the Australian scientific and natural resource management community to manage their animal biotelemetry data (Dwyer *et al.* 2015). ZoaTrack was built on open source software and the code is freely available for anyone to access on Github (https://github.com/AtlasOfLivingAustralia/oztrack). It is a Java based web application using a PostGIS-enabled database, Open Layers to display the maps in the browser, R to run various algorithms and GeoServer to generate and cache map tiles for efficiency (Hunter *et al.* 2013). Registered users of ZoaTrack are able to create new projects, upload datasets into projects, add new users to projects and specify access controls and embargo periods. A ZoaTrack project contains metadata pertaining to the project including information about related publications, and a data file containing an animal track data series. Once uploaded, data is immediately visualised on a configurable map. This map can be shared whilst keeping the raw animal location fixes embargoed (Figure 1). Once the data are open access, they are protected under a Creative Commons Attributions Licence agreement (www.creativecommons.org/licenses/by/3.0).

When a project has been created, ZoaTrack's suite of tools becomes available. The first step for the user is to apply some of the data cleansing and filtering tools (e.g. point deletion, speed filters, satellite dilution of precision class, data and time restrictions) to ensure the location fixes actually represent the animal's movement. These locations are then fitted with the animal's movement trajectory and overlaid onto a satellite map. Other environmental layers are optional. The user can generate

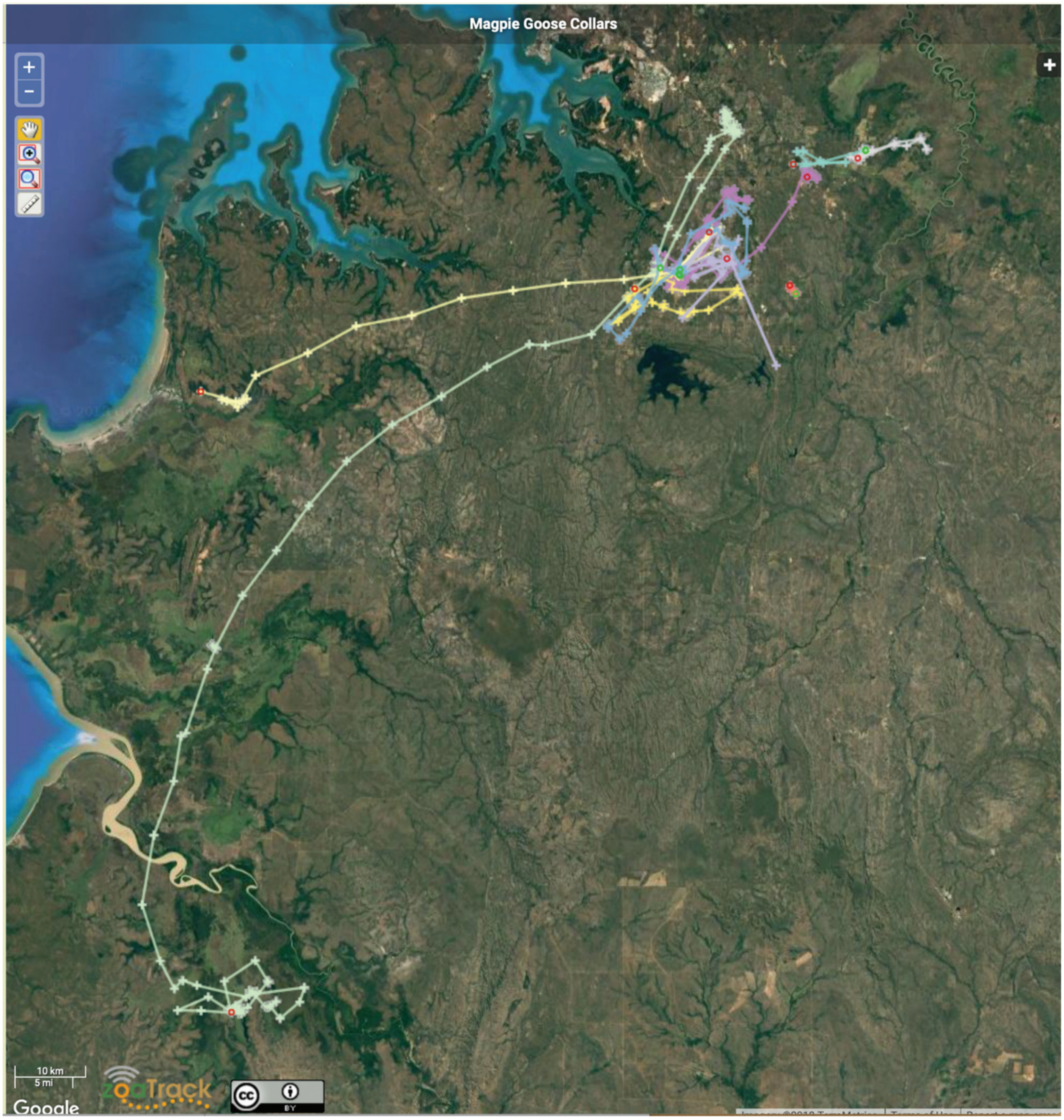

Figure 1. The ZoaTrack map screen shot with animal location fixes and fitted trajectories. The green circle denotes the start of the track and the red circle the end of the track. The data are satellite GPS-based location fixes from twelve Magpie Geese *(Anseranas semipalmata)* in Australia (URL:http://zoatrack.org/537). The image is generated by ZoaTrack so that a user may share the animal movement information without providing the raw location fixes for download.

an image of this map and data for sharing, whilst keeping the raw animal location data embargoed (Figure 1). The purpose of this is so researchers can provide project updates to partners, stakeholders and the general public whilst still keeping the raw location data under embargo until the study has been published.

Next these data undergo a series of temporal analysis functions (Step speed, step distance, cumulative distance moved, displacement distance), that are automatically calculated and plotted through time (Figure 2). This feature not only allows users to view the spatial extent of their tagged animal's movements, but it enables the user to quickly visualise temporal patterns in animal movements during the tracking period and the distance the animal has moved from the release site. ZoaTrack also offers a variety of home-range and space-use estimators (Maximum Convex Polygon, Kernel Utilisation Distribution, Hot spots; Figure 3). Full details of the ZoaTrack infrastructure (Hunter *et al.* 2013) and its ecological functionality (Dwyer *et al.* 2015) have previously been published. Following 6 years of operation, we examine how ZoaTrack is being utilised by the research and natural resource management communities and evaluate the opportunities and challenges for the platform.

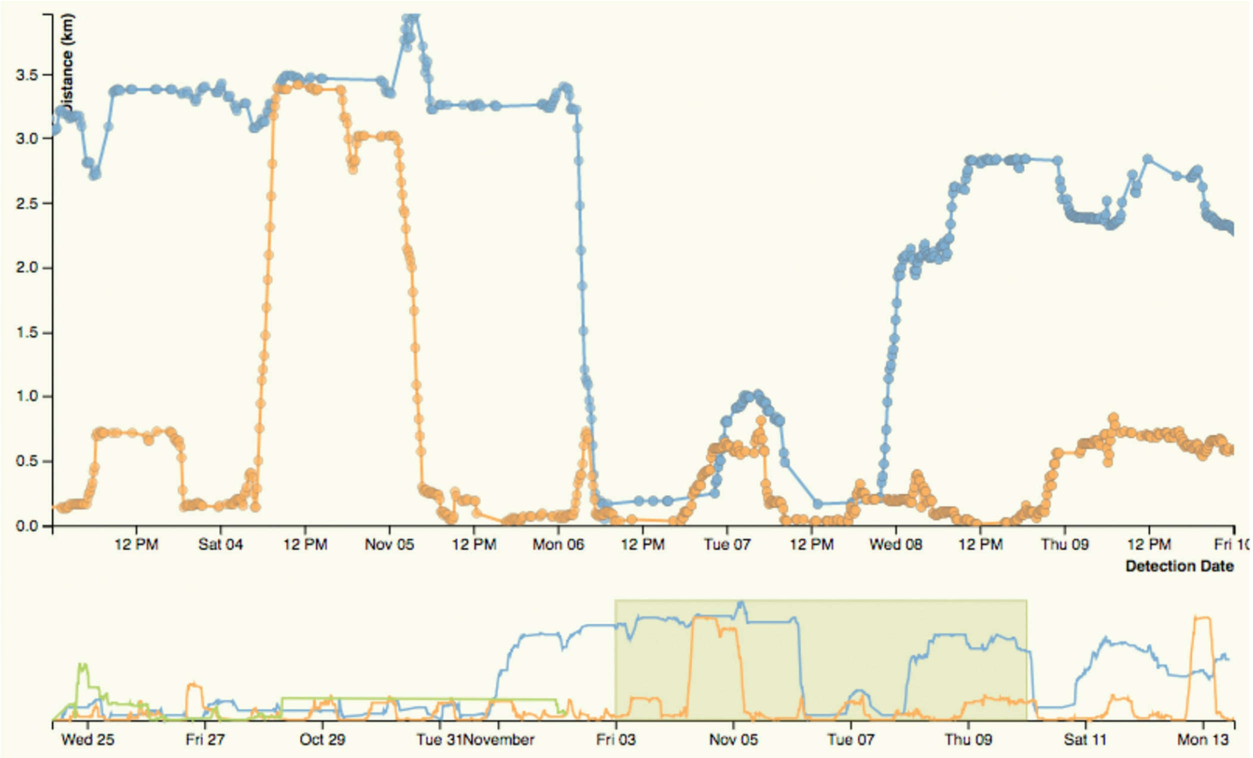

Figure 2. Screen shot of ZoaTrack temporal movement analysis graphical outputs. The bottom graph displays the entire timeline of the location data-set. The upper graph displays a zoomed in sub-set as defined by the shaded box, and shows the displacement distance of the animal from the release location. The data is from satellite GPS-based location fixes from three feral Pigs *(Sus Scrofa)* in Northern Australia (URL:http://zoatrack.org/projects/528).

Zoatrack for Data Storage

The ZoaTrack data repository has been steadily growing since inception in 2012. At the time of December 2018, the ZoaTrack repository contained 11,507 animal tracks (that is, sets of data representing animal movement through space and time) collected from 253 unique species (Figure 4). There were 866 registered users, comprising 512 projects, with 75% of these being projects having their data freely available. This high number of projects linked to multiple user accounts reflects how ZoaTrack is serving as a common analysis platform for teams of researchers and stakeholders to share, view, and analyse data.

Mammals were the most represented group within ZoaTrack (Figure 5a). This is not surprising because mammals are also the most represented animal group within the biotelemetry literature (Campbell *et al.* 2015). Birds and fish (including elasmobranches) did not feature highly despite this group being well represented in the biotelemetry literature. This may be the case as birds and fish do have taxa specific biotelemetry data repositories, and because the data structure for underwater acoustic telemetry is not currently supported within ZoaTrack. It was interesting to observe that ZoaTrack was also being heavily used to analyse tracking data from collared domestic cats, dogs and livestock animals, accounting for 14% of all projects, and 30% of animal tracks.

The predominant country of origin for ZoaTrack projects was Australia (Figure 6). This finding is not surprising since ZoaTrack was built to assist the Australian biotelemetry community and many of the GIS layers are Australia-specific. Detections (individual locations), trajectories (lines drawn between consecutive locations) and space-use density estimators generated within ZoaTrack can also be exported to other Australian National Coordinated Research Infrastructure, such as the Atlas of Living Australia, and the Biodiversity and Climate Change Virtual Laboratory. The second largest user of ZoaTrack was the U.S.A, which is also not surprising because the U.S.A is the largest user of biotelemetry equipment. Canada comes in third, with very few European users. The lack of users within Europe may be because 'Movebank', is located at the Max Plank Institute in Germany, and it is likely that European researchers use this as their preferred data-repository for their animal biotelemetry data. There were also a handful of users from developing countries.

After a steady increase in the number of users, the number of projects, tracked animals and detections held in the database have shown a sharp increase within the platform since 2017 (Figure 4). This marked increase may have been due to a greater number of animals being tagged with biotelemetry devices, but also the forced 3-year open access data policy was revoked in ZoaTrack in 2017. Thus, users can now embargo their data indefinitely, using a 12-month roll-over cycle. The hope is that the rolling embargo encourages more users to let their data be open access when they no longer require an embargo.

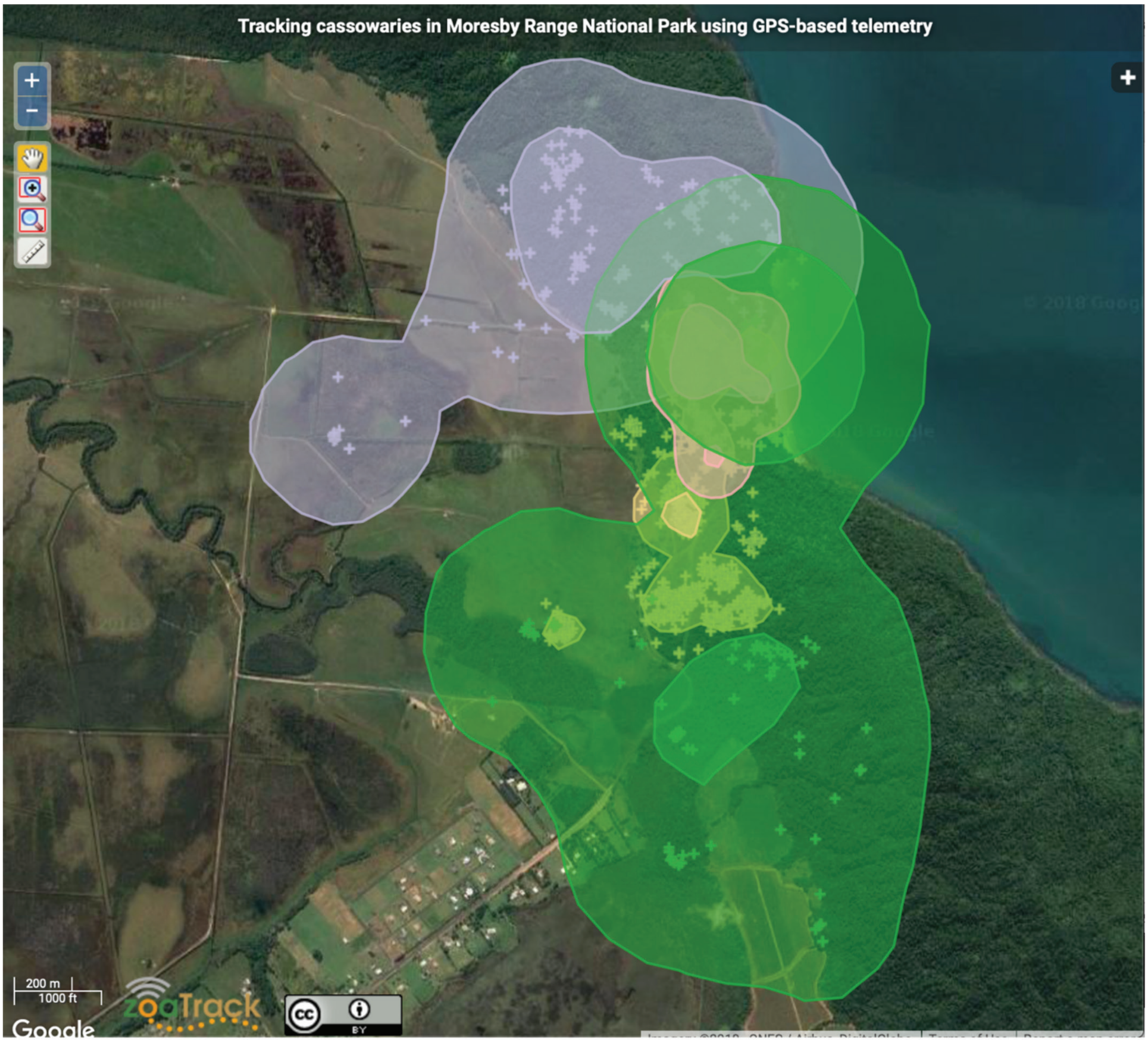

Figure 3. ZoaTrack screen shot of Kernel Utilisation Distribution plots for four Southern cassowaries *(Casuarius casuarius johnsonii)* at the 95 % and 50% level (URL: http://zoatrack.org/projects/1).

Ecological researchers often cite fear of data misuse and not obtaining appropriate credit for their work as the primary reason for not publicly archiving research data (Mills *et al.* 2015). In attempt to mitigate for this, and at the request of a user feedback survey, we installed a facility within ZoaTrack in 2016 to mint a Digital Object Identifier (DOI) for any uploaded dataset, thus enabling any user of ZoaTrack the ability to make their data a citable resource. However, the DOI functionality has rarely been used and at the time of writing only 13 manuscripts had DOIs minted by ZoaTrack. It appears that citation of the dataset is not sufficient credit for many biotelemetry researchers to share their data, even after the data has been published. We hope that the sharing of biotelemetry data through e-research infrastructure increases as more journals and funding bodies require the public archiving of research data.

Zoatrack for Data Analysis

The minimum convex polygon (MCP) and Kernel Utilisation Distribution (KUD) were the most frequently used functions within ZoaTrack to estimate a tagged animal's home-range (Figure 5b). These are probably the most widely used home-range estimator functions, and their frequent use demonstrates that ZoaTrack is providing an important function for the ecological research community.

A search within Google Scholar at the time of writing for the term "ZoaTrack" returned 58 publications that have used the facility to undertake scientific research. These studies encompass a diversity of animal species within terrestrial, marine and freshwater ecosystems. It has been estimated, that approximately 50% of all biotelemetry research projects are published within the scientific literature (Campbell *et al.* 2015). However, the proportion of projects within ZoaTrack that have been published is only around 11%. This short-fall maybe due to a lack of appropriate citing of ZoaTrack or because some studies that have yet to be published. It has been estimated that the average time period for a biotelemetry project to be published after the

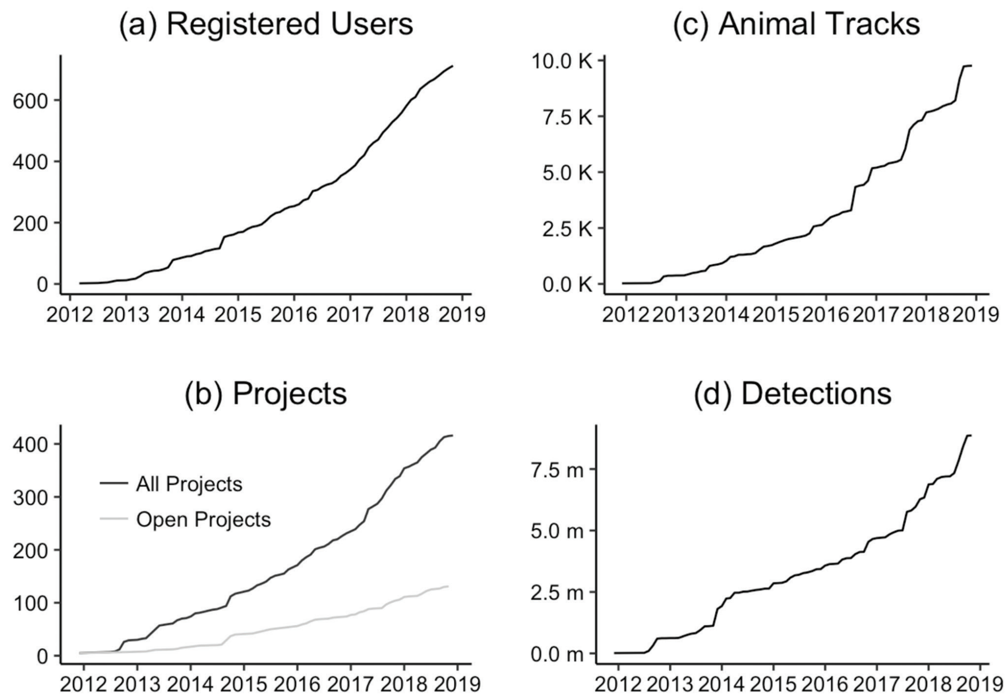

Figure 4. Line graphs illustrating the cumulative increase in ZoaTrack usage and data between 2012 and 2019.

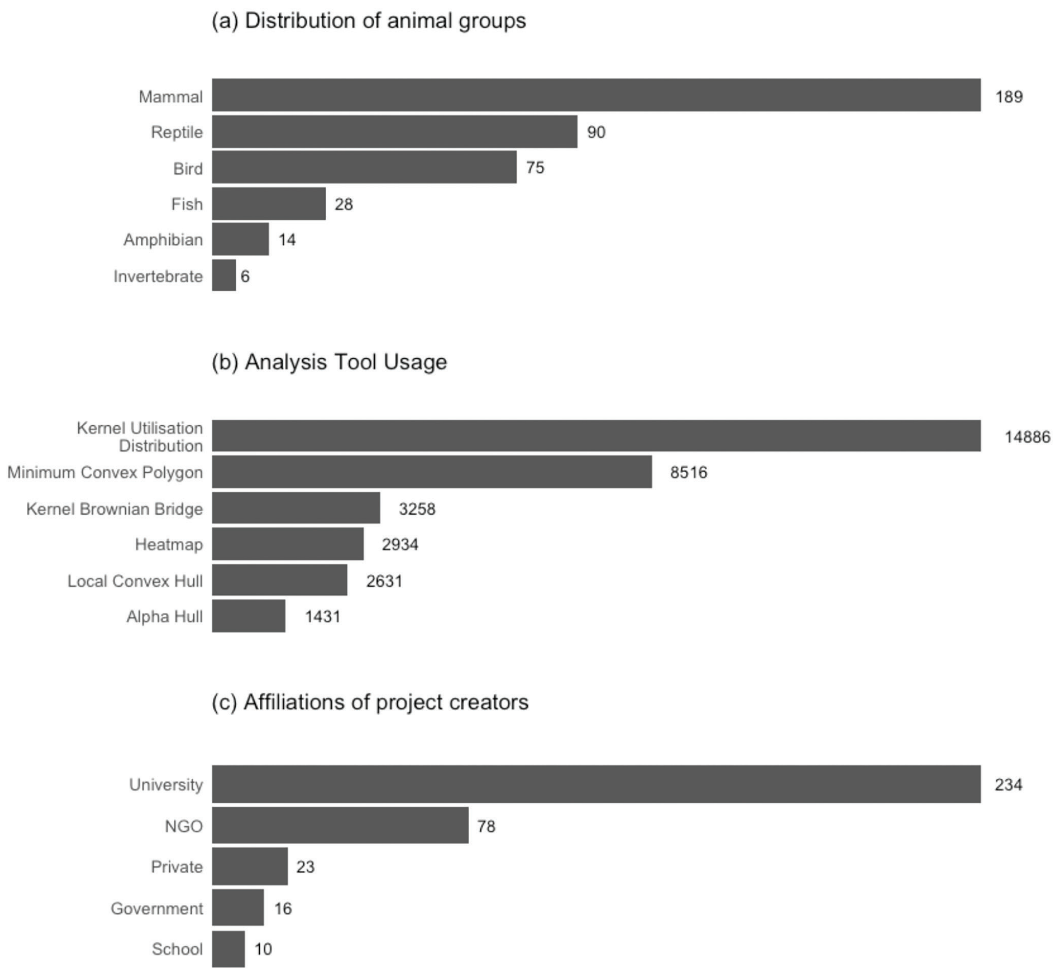

Figure 5. Bar graph illustrating how ZoaTrack is being used for the study of animal movement: (a) total number of projects across the animal groups, (b) total number of times each of the spatial analysis functions has been used, and (c) number of different affiliated organisations of the project creators.

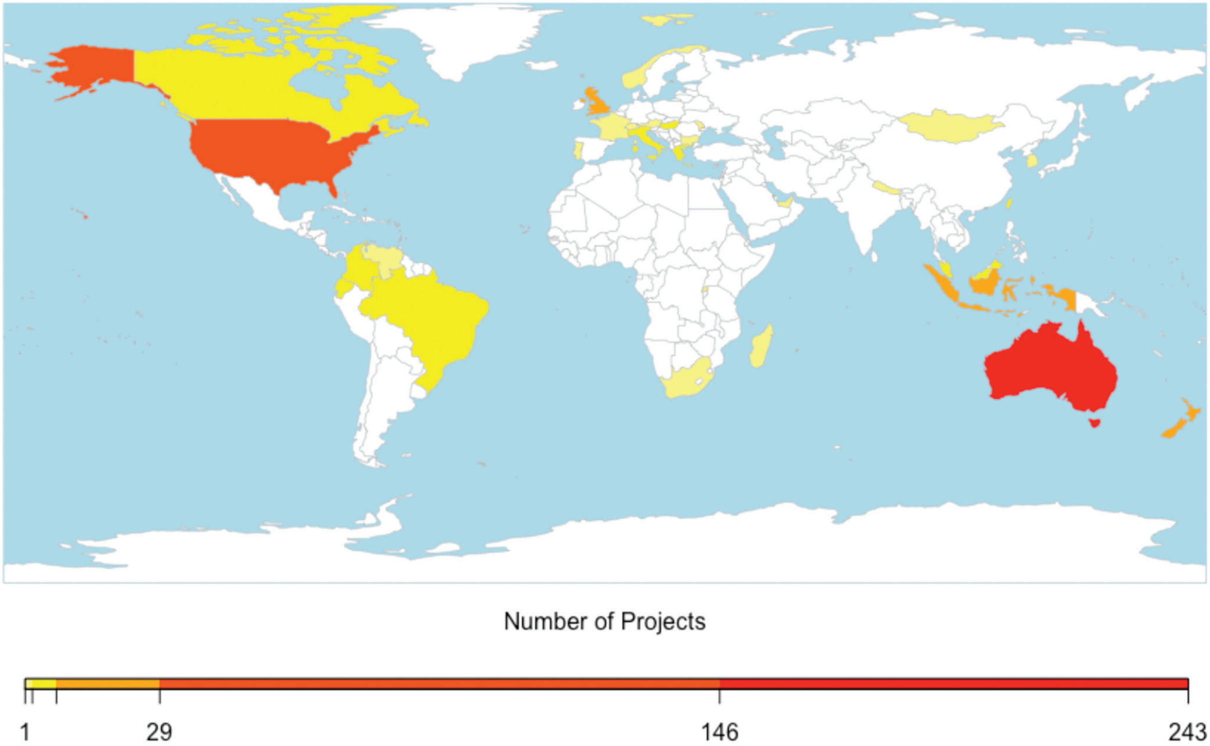

Figure 6. Map illustrating the global usage of ZoaTrack grouped by the number of projects per country.

completion of data collection is around 3 years (Campbell *et al.* 2015).

Most users of ZoaTrack were from the tertiary education sector, suggesting that research was the primary use of the platform (Figure 5c). A much smaller proportion of users were from governmental and non-governmental organisations, as well as private companies. Most of the private company users were environmental consultants. We know from personal communication that users from these three groups were deterred from entering data when there was only a 3-year limited embargo period. The unlimited embargo period, may therefore, result in an increase in these types of users. A small proportion of users were from high schools, and these projects had used biotelemetry devices to track the movements of marine turtles, whale sharks and livestock. It is encouraging that ZoaTrack was being used as an education tool to assist budding zoologists to understand animal movement and environmental interactions. There were at least four universities that were using ZoaTrack as a teaching tool within the undergraduate curriculum.

Future Opportunities and Challenges

The ultimate goal of ZoaTrack is that large amounts of data about animal occurrence and movement are made discoverable and available in perpetuity. Ensured discoverability of these data-collections would benefit the ecosystem research and ecological management communities through the provision of baseline values by which future measurements can be compared. However, a major barrier to realising this potential is that users are not providing high quality metadata about the animals, devices and the tagging regime used to enable the data to be shared and analysed into the future. Some users even deliberately provided misinformation about the species under study by not providing resolvable species scientific names. Achieving high quality metadata is a problem not limited to biotelemetry data. Ensuring that data collections have the appropriate metadata requires significant resourcing. A better approach may be to encourage users to provide high quality metadata by better understanding the benefits to themselves of a higher profile via data sharing through increased citations (Reichman *et al.* 2011). This may improve in the future, as it was recently demonstrated that the sharing of biotelemetry data shows a negative correlation with career progression (Campbell et al. 2019). In other words, early career scientists are much more open to data sharing practices than more senior scientists, and the hope is that this practice will continue as these scientists' careers progress.

A possible avenue to enhance data sharing and the metadata provided could be for ZoaTrack to extract the data straight from the devices themselves. The platform already does this for a handful of device manufacturers, but to enable this to be achieved across the high and growing number of biotelemetry manufacturers, improved standards in data format and structure are first required (Campbell *et al.* 2016). Work is progressing internationally towards this goal with the creation of a Data Standardisation Working Group within the recently formed International Biologging Society and a new Machine Observations Interest Group with Biodiversity Information Standards (TDWG: http://

tdwg.org). Both groups recognize a need for at least two standardisation efforts; one standard for the exchange of raw bio-logging data and another for exposing bio-logging data through the international network of repositories for biological observations or species occurrence records such as the Atlas of Living Australia (ALA, http://ala.org.au) and the Global Biodiversity Information Facility (GBIF, http://gbif.org).

Complete datasets need to be supported and exchanged through repositories such as ZoaTrack and Movebank, but should a 'ping' of a species at a place and time be equivalent to a human observation? Currently we would suggest no as the bio-logging data would increasingly overwhelm human observation data. Biodiversity data repositories such as the ALA and GBIF provide an effective role for bio-logging data discovery but we believe specialist repositories such as ZoaTrack and Movebank need to be promoted as the primary location for bio-logging data management, visualisation and analysis. We would recommend that a representative record for each animal track be submitted to biodiversity repositories using their internationally agreed Darwin Core Standard (https://dwc.tdwg.org/). This record would be in effect, a metadata record that stored the species name, date and other basic geographical information and would have a direct link back to the bio-logging repository such as ZoaTrack that can display full details and provide specialist analyses. With this strategy, bio-logging data are discoverable by species name, location and investigator through repositories like the ALA and GBIF.

In conclusion, ZoaTrack is assisting a diversity of researchers and natural resource managers from across the globe to better visualise and analyse their biotelemetry data. The significance of ZoaTrack as a data repository of animal location and movement data to draw upon for meta-analyses studies is beginning to be realised (Sequeira *et al.* 2018). Significant opportunities and challenges remain in regards to metadata standards and data repository interoperability.

Acknowledgements

We would like to thank the Atlas of Living Australia Team, in particular Adam Collins and we acknowledge Professor Jane Hunter, Wilfred Brimblecombe, and Charles Brooking for the early development of the platform as Oztrack under the NeCTAR eResearch tools program.

References

Campbell, H.A., Micheli-Campbell, M., Udywer, V. 2019. Early Career researchers embrace data sharing. *Trends in Ecology& Evolution* **34:** 95-98. https://doi.org/10.1016/j.tree.2018.11.010

Campbell, H.Al, 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:** 1. https://doi.org/10.1186/s40317-015-0096-x

Campbell, H.A., Beyer, H.L., Dennis, T.E., Dwyer, R.G., Forester, J.D., Fukuda, Y., *et al*. 2015. Finding our way: on the sharing and reuse of animal telemetry data in Australasia. *Science of the Total Environment* **534:** 79-84. https://doi.org/10.1016/j.scitotenv.2015.01.089

Dwyer, R.G., Brooking, C., Brimblecombe, W., Campbell, H.A., Hunter, J., Franklin, C.E. 2015. An open Web-based system for the analysis and sharing of animal tracking data. *Animal Biotelemetry* **3:** 1. https://doi.org/10.1186/s40317-014-0021-8

Hunter, J., Brooking, C., Brimblecombe, W., Dwyer, R.G., Campbell, H.A., Franklin, C.E. 2013. OzTrack--E-Infrastructure to Support the Management, Analysis and Sharing of Animal Tracking Data. *eScience (eScience), 2013 IEEE 9th International Conference on*, 140-147. https://doi.org/10.1109/eScience.2013.38

Mills, J., Teplitsky, C., Arroyo, B., Charmantier, A., Becker, P.H., Birkhead, T.R., Bize, P., *et al*. 2015. Archiving Primary Data: Solutions for Long-Term Studies. *Trends in Ecology& Evolution* **30:** 581-589. https://doi.org/10.1016/j.tree.2015.07.006

Reichman, O.J., Jones, M.B., Schildhauer, M.P. 2011. Challenges and opportunities of open data in Ecology. *Science* **331:** 703-705. https://doi.org/10.1126/science.1197962

Sequeira, A.M.M., Rodríguez, J.P., Eguíluzc, V.M., Harcourt, R., Hindelle, M., Sims, D.W., Duartea, C.M., *et al*., 2018. Convergence of marine megafauna movement patterns in coastal and open oceans. *P.N.A.S.* **115:** 3072-3077. https://doi.org/10.1073/pnas.1716137115