ARPHA Preprints

PENSOFT

PREPRINT

Author-formatted, not peer-reviewed document posted on 13/07/2022

DOI: https://doi.org/10.3897/arphapreprints.e90210

A dataset of sea turtle occurrences around the Taiwan coast

Daphne Z. Hoh, Chia-Ling Fong, Huai Su, Pengyu Chen, Chia-Chen Tsai, Kelly W. H. Tseng, Melissa J. Y. Liu

ARPHA Preprints *Author-formatted, not peer-reviewed document posted on 13/07/2022*. DOI: https://doi.org/10.3897/arphapreprints.e90210

A dataset of sea turtle occurrences around the Taiwan coast

Daphne Z Hoh[‡,§], Chia-Ling Fong[‡,|,¶], Huai Su[‡], Pengyu Chen[‡], Chia-Chen Tsai[‡], Kelly WH Tseng[#], Melissa JY Liu[§]

‡ TurtleSpot Taiwan, Pingtung, Taiwan
§ Taiwan Biodiversity Information Facility (TaiBIF), Biodiversity Research Centre, Academia Sinica, Taipei, Taiwan
| Biodiversity Research Centre, Academia Sinica, Taipei, Taiwan
¶ Biodiversity Program, Taiwan International Graduate Program, Academia Sinica and National Taiwan Normal University, Taipei, Taiwan
# Independent Researcher, Taipei, Taiwan

Corresponding author: Daphne Z Hoh (daphnehohzhiwei@gmail.com)

Abstract

Background

We describe a dataset of sea turtle sightings around the coast of Taiwan and its islands. The data was collected by citizen scientists and reported to TurtleSpot Taiwan, a citizen science initiative that collects sea turtle sighting data mainly through the Facebook group. This dataset includes 3,515 sighting data dated from March 2010, except most of the data (n=3,128; 89%) falls were between June 2017 to December 2021. A standardized format of basic sighting information was suggested to anyone who wishes to report the occurrence. We also request photos and videos for turtle species identification and to record any physical abnormality. In addition to the basic data often associated with an occurrence, TurtleSpot aims to identify each sea turtle up to the individual level using the Photo ID method. Hence, if a good quality photo of left- and right-facial scutes were available, the sighted individual can be identified and given a unique Turtle ID. In total, 762 turtle individuals were assigned a turtle ID, comprising 723 green, 38 hawksbill, and 1 olive ridley turtles. The data of TurtleSpot Taiwan has allowed some ecological observations of sea turtles in the wild, such as witnessing the recovery of some injured turtles, behaviours, intra- and inter-species interactions, and physical abnormality. These data will offer essential information that helps to understand the foraging ecology of sea turtles and assists in the development of conservation measures.

New information

This dataset contains 3,515 occurrence records of sea turtles and is currently the largest public dataset of sea turtle sighting records in Taiwan. Post-publication of this dataset to

the Global Biodiversity Information Facility (GBIF) demonstrated that the number of green sea turtle *Chelonia mydas* occurrence records is few of the top in the world. The data served as the foundation for understanding sea turtle ecology in Taiwan's coastal waters.

Keywords

occurrence; sighting data; citizen science, Facebook group; coastal waters

Introduction

Sea turtle is a migratory species that travel between the nesting and foraging ground. In Taiwan, much is known about the nesting ecology (Cheng et al. 2009, King et al. 2013, Cheng et al. 2015, Cheng et al. 2018), probably owing to the easier accessibility of the nesting sites. Nonetheless, understanding of the foraging population and its ecology is still limited. Recognising the gap, TurtleSpot Taiwan - a community-led-citizen science initiative that collects sea turtle sighting reports was founded. It was co-founded by a group of marine enthusiasts made up of scientists, underwater photographers, and marine awareness educators in June 2017. Since then, we receive sighting reports via our Facebook group and our community and sighting reports are actively growing. Up to July 2022, our community grew up to over 18,900 members and on average, 57 sighting reports were received monthly. In addition to sighting reports, TurtleSpot aims to develop a database of turtle profiles in Taiwan by identifying each individual turtle using the Photo ID method (Dunbar et al. 2021). We identified sea turtle individuals through their unique facial-scutes patterns and record any distinct characteristics of their physical appearances such as caparace or limb injury is available. To encourage continuous reports of the citizen scientists, we allow the sighting reporter to name the turtle if the individual is a new record in our database.

The purpose for preparing the current dataset was for a few reasons: (1) a give back to our community since most of these data were their contributions, (2) for an ongoing research on sea turtle foraging ecology in Taiwan (Fong et al., in prep), and (3) publicly open the data for advancement especially in the scientific and conservation communities.

Sampling methods

Study extent: We collect reports of all sea turtle sightings from any region, but our current focus is on the sighting data around the coasts of Taiwan and its islands.

Sampling description: More than half (n=2,235; 63.6%) of the data was provided by the citizen scientists. The remaining data (n=1,280) was personal records contributed by two of the co-authors.

Quality control: Each sighting reported in TurtleSpot's Facebook group was first checked by the group admin for the required information pending approval. The required information

includes sighting date, time, location (and dive site), depth, method, and for species and individual identification - high-quality photos and videos containing the whole body, left, and right faces of the sighted turtle (Fig. 1). The sighting provider will be requested to provide any of the missing information unless unavailable.

Step description:

1. Receive data: We receive sea turtle sighting reports contributed by citizen scientists via our Facebook group.
2. Approve report: Facebook post of a sighting report which followed our reporting format (Fig. 1) and passed the quality check by the group admin was approved to be visible in the Facebook group, allowing community interaction within the post.
3. Record data: Sighting information contained in the report was recorded in Google Sheets as raw data. These raw data (mostly in Chinese) were translated into English.
4. Determine additional information from the sighting report: We identified and recorded the additional information about the sighted individual turtle based on the photos and videos provided in each sighting report and additional notes provided by the citizen scientists. This information was mostly identified through observation on-site and through photo and videos. We observed the biological characteristics of the occurrence (sex, life stage, behaviour, associated taxa) and abnormal conditions of the turtle (e.g. fishing line entanglement, tumour and others).
5. Sea turtle individual identification: If high-quality photos of the left and right faces of the sighted sea turtle were available, we use the Photo ID method to identify the turtle individual. Currently, we use two methods to perform photo identification: (1) compare the facial scute pattern manually, and (2) HotSpotter (Crall et al. 2013), an open-source software. Each sea turtle individual was assigned a unique turtle ID.
6. Open data preparation: We first converted the data into Darwin Core Archive formatted occurrence data in Google Sheet using the Darwin Core Archive Assitant Add-on (Salim and Saraiva 2020) where applicable. Refer Data resources section for a detailed description of each column. We then validated the occurrence data using the GBIF Data Validator (Global Biodiversity Information Facility 2017) and edited the error. Lastly, we uploaded and published the data via Taiwan Biodiversity Information Facility IPT.

Geographic coverage

Description: Most of the sighting data were from Taiwan and its islands (Fig. 2), and only a few (n=35) were from other countries which include Indonesia, Philippines, Malaysia, Palau, Mariana Islands, Japan, Maldives, and United States.

 Author-formatted, not peer-reviewed document posted on 13/07/2022. DOI: https://doi.org/10.3897/arphapreprints.e90210

Taxonomic coverage

Description: Four species of sea turtles were recorded in the dataset (Hoh and Fong 2022), including green turtle (*Chelonia mydas*), hawksbill (*Eretmochelys imbricata*), olive ridley (*Lepidochelys olivacea*), and kemp's ridley (*Lepidochelys kempii*). 97.3% and 2.4% of the sightings from Taiwan are from the green and hawksbill turtles, respectively. Occurrences that failed to assign species (n=11) were recorded as Cheloniidae.

Taxa included:

Rank	Scientific Name	Common Name
kingdom	Animalia	Animal
phylum	Chordata	
subphylum	Vertebrata	
superclass	Reptilia	
order	Testudines	
suborder	Cryptodira	
superfamily	Chelonioidea	Marine turtle
family	Cheloniidae	
genus	Chelonia	
genus	Eretmochelys	
genus	Lepidochelys	
species	mydas	Green
species	imbricata	Hawksbill
species	olivacea	Olive Ridley
species	kempii	Kemp's Ridley

Temporal coverage

Data range: 2010-3-23 - 2021-12-29.

Notes: TurtleSpot was officially founded in June 2017. Hence, most sighting records range from June 2017 to December 2021, comprising about 89% (n=3,128) of the dataset (Fig. 3). Occasionally, we receive sighting reports dated before June 2017, with the earliest dated in March 23rd, 2010.

 Author-formatted, not peer-reviewed document posted on 13/07/2022. DOI: https://doi.org/10.3897/arphapreprints.e90210

Usage licence

Usage licence: Other

IP rights notes: The dataset in the current work is licensed under a Creative Commons Attribution (CC-BY) 4.0 License. Any image and video accessed through the URL from the dataset are licensed under the Creative Commons Attribution (CC-BY-NC) 4.0 License.

Data resources

Data package title: Sea turtle sightings in Taiwan

Resource link: https://www.gbif.org/dataset/336b6790-062f-407f-a783-2f1d8874e6c3

Number of data sets: 1

Data set name: Sea turtle sightings in Taiwan

Data format: Darwin Core Archive

Data format version: 2021-07-15

Description: Data fields were standardised to 46 Darwin Core terms as listed in the following table. The column label and some of the description are written as listed in the List of Darwin Core terms (accessed June 2022; created by the TDWG Darwin Core Maintenance Group). A more specific description about the column used in the current dataset was also added if applicable.

Column label	Column description
occurrenceID	An identifier for the Occurrence (as opposed to a particular digital record of the occurrence).
catalogNumber	An identifier unique for the record within the dataset.
rightsHolder	A person or organization owning or managing rights over the resource.
recordedBy	Names of the sighting reporter/citizen scientist.
year	The four-digit year in which the Event occurred, according to the Common Era Calendar. Year of sighting.
month	The integer month in which the Event occurred. Month of sighting.
day	The integer day of the month on which the Event occurred. Day of sighting.
eventDate	Sighting date.
eventTime	The time or interval during which an Event occurred.
country	The name of the country in which the Location occurs.

countryCode	The standard code for the country in which the Location occurs.
higherGeography	A list of geographic names less specific than the information captured in the locality term.
locality	Name of the sighting location or dive site.
locationRemarks	More specific location compared to locality, usually the name of the dive site.
decimalLatitude	The geographic latitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location. Positive values are north of the Equator, negative values are south of it. Legal values lie between -90 and 90, inclusive.
decimalLongitude	The geographic longitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location. Positive values are east of the Greenwich Meridian, negative values are west of it. Legal values lie between -180 and 180, inclusive.
coordinateUncertaintyInMeters	The horizontal distance (in meters) from the given decimalLatitude and decimalLongitude describing the smallest circle containing the whole of the Location.
georeferenceRemarks	A note stating that our GPS coordinates were estimated from the dive site or sighting location.
geodeticDatum	The ellipsoid, geodetic datum, or spatial reference system (SRS) upon which the geographic coordinates given in decimalLatitude and decimalLongitude as based.
verbatimDepth	The original description of the depth below the local surface. This is an estimation provided by the sighting reporter.
samplingProtocol	Sighting method of the Occurrence.
associatedReferences	An URL links to the Facebook post from the sighting reporter in our Facebook Group, which we define as a single Occurrence event. The link may be broken if the sighting reporter decided to delete the post.
basisOfRecord	The specific nature of the data record.
individualCount	Our purpose of preparing this occurrence dataset is to identify each turtle individual. Hence, if a sighting report contains more than one sea turtle, this occurrence record will be duplicated as a new row. Hence, the individual count of each data is only '1'.
kingdom	The full scientific name of the kingdom in which the taxon is classified.
taxonRank	The taxonomic rank of the most specific name in the scientificName.
vernacularName	A common or vernacular name.
scientificName	The full scientific name.
taxonID	An identifier for the set of taxon information (data associated with the Taxon class). We use the URL of species in GBIF Backbone Taxonomy checklist.

behavior	The behavior shown by the subject at the time the Occurrence was recorded.
occurrenceRemarks	Condition of the turtle during the sighting (e.g. alive, dead, stranding).
dynamicProperties	Any physical abnormality that was observed (e.g. injury, tumor, debris entanglement)
associatedTaxa	A simple description of association and vernacular name of taxa in which this Occurrence is to each of them.
lifeStage	The age class or life stage of the organism at the time the Occurrence was recorded. Estimated via physical appearence of the sighted turtle
sex	The sex of the biological individual represented in the Occurrence. Determination of sex is only applicable to adult sea turtles through the size of their tail. Sex determination is successful only when photos/videos of the tail are available.
organismName	A textual name or label assigned to an Organism instance. Mostly named by the citizen scientists.
license	A legal document giving official permission to do something with the resource. The license in this column applied to the text data of this dataset only.
identificationID	Turtle ID. Every identifiable turtle individual has a unique ID.
associatedMedia	An URL links to the website of TurtleSpot Turtle Photo ID database, which show media and information about this particular individual.
identifiedBy	Names of people who identified the turtle individual via Photo-ID method.
informationWithheld	Additional information that exists, but that has not been shared in the given record.
occurrenceStatus	A statement about the presence or absence of a Taxon at a Location. All value is 'present' in the current dataset.
eventRemarks	Notes about the incomplete sighting eventDate.
continent	The name of the continent in which the Location occurs.
county	The full, unabbreviated name of the next smaller administrative region than stateProvince (county, shire, department, etc.) in which the Location occurs.
island	The name of the island on or near which the Location occurs.

Acknowledgements

The authors thank Dr Yoko Nozawa and his laboratory members of Biodiversity Research Centre, Academia Sinica, Taiwan, and Taiwan Ocean Conservation Administration for supporting this community project. We thank Hao-Chia Huang, Marble Lo, Anita W, Liu Shih, Lyvia Chong, Aiden Lo, Pei-Shan Zhuang, and Te Hsiang Wong for their assistance in data filling. We thank the Taiwan Keep Walking Fund for funding the website construction of Taiwan Turtle Photo ID database. We thank slowIsFast Ltd. for advising the data management process. Taiwan Biodiversity Information Facility (TaiBIF) is funded by

the Ministry of Science and Technology, Taiwan (Grant ID: 108-2621-B-001-009-MY3), which supported the publication of this data paper.

Author contributions

DZH, CLF, HS, and PC conceived the study; DZH and CLF designed the scientific protocol; CLF, HS, and CCT performed most of the individual turtle identification; HS and CCT contributed a substantial amount of data; DZH wrote the paper and prepared figures; DZH, CLF, HS, PC, CCT and KWHT performed data curation; MJYL assisted throughout the data management process; all authors approved the manuscript.

References

- Cheng I, Huang C, Hung P, Ke B, Kuo C, Fong C (2009) Ten years of monitoring the nesting ecology of the green turtle, *Chelonia mydas*, on Lanyu (Orchid Island), Taiwan. Zoological Studies 48 (1): 83-94.
- Cheng I, Lin C, Tseng C (2015) Factors influencing variations of oxygen content in nests of green sea turtles during egg incubation with a comparison of two nesting environments. Journal of Experimental Marine Biology and Ecology 471: 104-111. https://doi.org/10.1016/j.jembe.2015.05.013
- Cheng I, Cheng W, Chan YT (2018) Geographically closed, yet so different: Contrasting long-term trends at two adjacent sea turtle nesting populations in Taiwan due to different anthropogenic effects. PLOS One 13 (7): e0200063. https://doi.org/10.1371/journal.pone.0200063
- Crall JP, Stewart CV, Berger-Wolf TY, Rubenstein DI, Sundaresan SR (2013) Hotspotter—patterned species instance recognition. 2013 IEEE workshop on applications of computer vision (WACV). IEEE. https://doi.org/10.1109/WACV.2013.6475023
- Dunbar S, Anger E, Parham J, Kingen C, Wright M, Hayes C, Safi S, Holmberg J, Salinas L, Baumbach D (2021) HotSpotter: Using a computer-driven photo-id application to identify sea turtles. Journal of Experimental Marine Biology and Ecology 535 https://doi.org/10.1016/j.jembe.2020.151490
- Global Biodiversity Information Facility (2017) GBIF data validator. https://www.gbif.org/tools/data-validator. Accessed on: 2022-6-28.
- Hoh D, Fong C (2022) Sea turtle sightings in Taiwan. Occurrence dataset. TurtleSpot Taiwan. URL: https://doi.org/10.15468/43z4mj
- King R, Cheng W, Tseng C, Chen H, Cheng I (2013) Estimating the sex ratio of green sea turtles (*Chelonia mydas*) in Taiwan by the nest temperature and histological methods. Journal of Experimental Marine Biology and Ecology 445: 140-147. https://doi.org/10.1016/j.jembe.2013.03.016
- Salim JA, Saraiva AM (2020) A Google Sheet Add-on for Biodiversity Data Standardization and Sharing. Biodiversity Information Science and Standards 4:e59228 https://doi.org/10.3897/biss.4.59228
- South A (2017) rnaturalearth: World Map Data from Natural Earth. R package version 0.1.0. URL: https://CRAN.R-project.org/package=rnaturalearth

Author-formatted, not peer-reviewed document posted on 13/07/2022. DOI:
https://doi.org/10.3897/arphapreprints.e90210

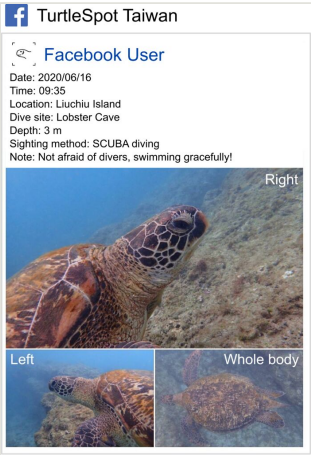

Figure 1.

Format of reporting sea turtle sighting to TurtleSpot Taiwan. Facebook icon made by Freepik via Flaticon.

Author-formatted, not peer-reviewed document posted on 13/07/2022. DOI:
https://doi.org/10.3897/arphapreprints.e90210

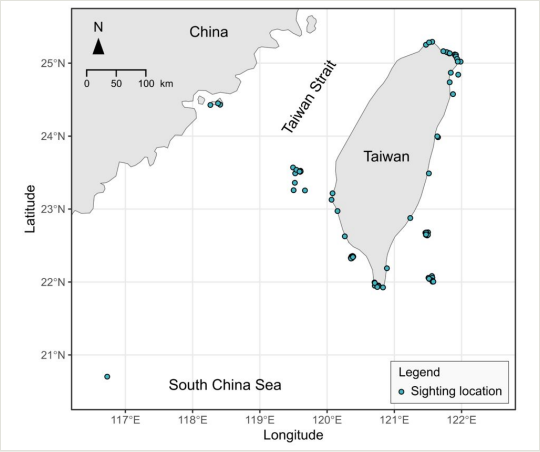

Figure 2.

Location of sea turtle sightings in Taiwan. Data outside of Taiwan is not shown. Map was plotted using the R package 'rnaturalearth' (South 2017).

Author-formatted, not peer-reviewed document posted on 13/07/2022. DOI:
https://doi.org/10.3897/arphapreprints.e90210

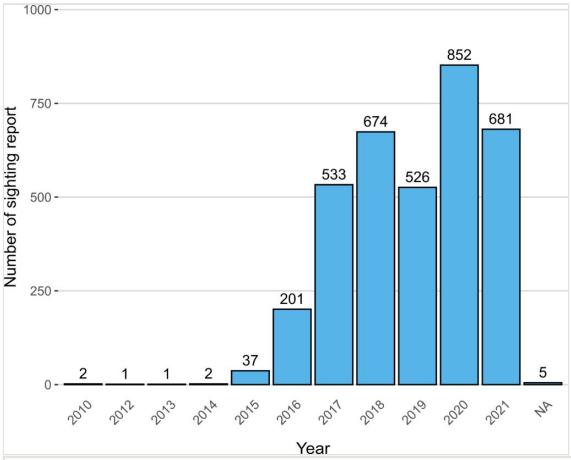

Figure 3.

Number of sighting records across the year. 'NA' indicates no sighting year was given.