Guidelines to merge stranding databases: the case of franciscana (*Pontoporia blainvillei*) in the extreme South of Brazil

Ingrid Schommer Penha1,2*, Danielle da Silveira Monteiro2,3, Eduardo Resende Secchi2, Sérgio Curi Estima3, and Paul Gerhard Kinas1

1Laboratório de Estatística Ambiental (LEA), Instituto de Matemática, Física e Estatística (IMEF), Universidade Federal do Rio Grande (FURG). Av. Itália, s/n - km 8, Carreiros, Anexo 2, Sala J07. Rio Grande, RS, Brazil
2Laboratório de Ecologia e Conservação da Megafauna Marinha (EcoMega), Instituto de Oceanografia (IO), Universidade Federal do Rio Grande (FURG). Av. Itália, s/n - km 8, Carreiros, 96201-900 Rio Grande, RS, Brazil
3Núcleo de Educação e Monitoramento Ambiental (NEMA). Rua Maria Araújo, 480, Cassino, 96207480 Rio Grande, RS, Brazil

*Corresponding author: ingridschommer.is@gmail.com

Abstract

Strandings of several species of marine mammals have been monitored over the years off the south coast of Rio Grande do Sul State, southern Brazil. A long time series of records implies extensive databases on strandings. In the region, these databases are maintained by two institutions: (i) Núcleo de Educação e Monitoramento Ambiental (NEMA) [Environmental Education and Monitoring Nucleus] and (ii) Laboratório de Ecologia e Conservação da Megafauna Marinha (ECOMEGA) [Ecology and Conservation of Marine Megafauna Laboratory] at the Federal University of Rio Grande – FURG. In order to make the time series as complete as possible and to make it a reliable source for research related to a particular species and region, the present work proposes a methodology for unifying both databases for franciscana (*Pontoporia blainvillei*) for the period 2000 - 2020 (NEMA, n = 3,029; FURG, n = 4,629). To build the unified database, specific metrics were outlined for the species and region, in order to confirm a ‘match’ of a record. The ‘match’ variable is a subjective value that classifies the resighting of a stranded animal as excellent (1), good (2), or regular (3). With the implementation of guidelines to merge stranding databases 1,812 ‘excellent’ and 97 ‘good’ combinations were recorded. Sixty records classified as 30 ‘regular’ matches were kept in the database, as they are possibly not recounts due to large differences in our primary and secondary metrics. To characterize a reliable match between distinct databases, the general guidelines outlined here need to be adapted according to the species of interest and the specificities of each monitoring program. The methodology developed to unify the databases had as its main objective the identification of matches.

Keywords: Franciscana, beach monitoring, strandings, southern Brazil, cetaceans, marine mammals.

ARTICLE HISTORY

Manuscript type: Article

Article History

Received: 20 June 2022
Received in revised form: 29 September 2022
Accepted: 10 October 2022
Available online: 31 October 2022

Handling Editor: Federico Riet Sapriza


Introduction

Databases of aquatic mammal strandings are important sources of information to assess the status of the species in a given habitat. The continuous monitoring of strandings provides the constitution of an expressive and reliable time series, with high scientific potential, which allows studies on stranding patterns, outlining possible external pressures, estimating mortality, and making predictions associated with the impact of fishing on mortality (Kinas, 2002; Geraci and Lounsbury, 2005; Pyenson, 2010; Prado et al., 2013; 2016). Hence, time series could aid in decision-making for the conservation of species (e.g. Prado et al., 2013; 2016).

The franciscana, *Pontoporia blainvillei* (Gervais and d’Orbigny, 1844), is considered the most threatened dolphin species in South America (Secchi et al., 2021), and is classified as vulnerable (VU), with decreasing population trend by the International Union for Conservation of Nature (IUCN, 2022). Franciscanas, the only living representative of the Pontoporiidae family, are endemic to the waters of the Southwest Atlantic Ocean (Secchi, 2010), and...
occurs from the coast of Espírito Santo (ES) (~18° S), in Brazil, to Chubut (~42° S), province of Argentine Patagonia (Bastida et al., 2007). The species has also been documented to occur in estuarine waters (e.g. Praderi, 1986).

Aerial surveys and data on fisheries bycatch indicate that franciscanas are more common in shallow waters, up to the isobaths of 30 or 50 m depending on the region, and may occur less frequently beyond these depths (Danilewicz et al., 2010; Secchi et al., 2021). The franciscana’s habitat and diet overlap with the coastal gillnet fishing area in the southern coast of Rio Grande do Sul state (RS), Southern Brazil, and with the commercially-valued species such as whitemouth croaker (Micropogonias furnieri) and striped weakfish (Cynoscion guatucupa) (Secchi et al., 1997; Rodriguez et al., 2002; Danilewicz et al., 2009; Haimovici & Cardoso, 2016; Bassoi et al., 2021), respectively. The spatial overlap results in a very high mortality rate of franciscanas in gillnet fisheries (Prado et al., 2021), and many of these dead dolphins wash ashore on the adjacent coast (Prado et al., 2013). Thus, to create strategies to help franciscana conservation, four Franciscana Management Areas (FMAs) were proposed and refined (Secchi et al., 2003; Di Benedetto et al., 2010; Cunha et al., 2014).

For decades, NEMA and ECOMEGA have conducted beach monitoring efforts (BM) off the south coast of RS and these records allow a better understanding of the conservation status of franciscanas in this Brazilian state. In order to obtain more reliable results on stranded franciscana records, a protocol was established for the unification of stranding databases to prevent gaps in beach monitoring coverage, for example when one of the institutions was experiencing difficulties associated with logistics to carry out or complete the BM, and avoid recounts.

Materials and methods

The monitoring of strandings was routinely carried out along 355 km of coastline, subdivided into northern and southern sections. The northern section is approximately 135 km long and ranged from São José do Norte municipality to Peixe Lagoon. The southern section covers 220 km of beach from the breakwater of the Rio Grande Mouth Bar to Chuí municipality, in the extreme south of Brazil (Fig. 1). Both sections are part of the Franciscana Management Area III (FMA III – sensu Secchi et al., 2003), which extends from southern Santa Catarina, through RS and into Uruguay.

For the consolidation of a Unified (stranding) Database (UD), the databases generated between the years 2000-2020 by NEMA and ECOMEGA were used. Over these 21 years, 3,029 franciscana carcasses were registered by NEMA (with monthly BM), and 4,629 by FURG (with biweekly BM) to the north and south of the Rio Grande Mouth Bar. These counts include duplicates as they refer to the same carcass recorded by both institutions. This we called ‘a match’. In building the UD, the match has to be identified. Over the years, gaps in monitoring were experienced by both institutions. Nevertheless, at least one of the teams was in the field over the entire study period, and in this sense, one complemented the other’ data.

To identify matches and avoid double counts, we developed a protocol that can be outlined in three steps. The first step consisted in selecting variables to integrate the UD. To be a candidate, a variable had to be present in both databases and follow the same criteria (or be comparable), be simple and informative. In step two, a history of fieldwork was consolidated, organizing it according to its chronology, and grouping it in accordance with the institution that carried out the beach monitoring (FURG or NEMA) with the direction of monitoring (north or south), and the respective number of stranding records. BMs that occurred on the same day by different institutions were indicated. From BMs that occurred on the same day by both institutions, direct comparisons gave rise to a classification of a good match based only on our primary metric, or spatial proximity. All remaining cases needed a secondary (supporting) metric. These metrics, which will be outlined next, were used to check if two recorded carcasses could be classified as a match. The primary metric referred to spatial proximity, and was established according to variations between geographic coordinates (geo), latitudes (ΔLAT = latitude.nema - latitude.furg), and longitudes (ΔLONG = longitude.nema - longitude.furg), considering two carcasses sighted on the same day by the two institutions. The metric was elaborated due to the large amounts of BMs that occurred on the same day, especially in the first decade of these records. A total of 372 pairs of strandings (thus, 744 franciscanas in this context) recorded on the same day by both institutions were analyzed in order to delimit reliable intervals for possible matches. An exploratory analysis of the resulting differences was performed visually via scatterplots, histograms, and ordered cumulative percentiles (ascending). To calculate the geodetic distances (geo dist, the shortest distances between two points), the Software for Data Analysis (SoDA) package was used (Borcard et al., 2011). Those 372 pairs of stranding records were used to determine spatial and temporal ranges to be considered in our flowcharts shown in step three. These analyses were performed in the R software (R Development Core Team, 2017). To evaluate possible temporal matches between records, a temporal window of at most 30 days between successive BMs was considered. Beyond this time lag a carcass was assumed to have suffered...
decomposition (Prado et al., 2016). To establish subjective values that classifies the resighting of a stranded animal, or ‘match’, two schematic flowcharts were prepared to guide the comparisons, considering three periods within the 30 days: the first considers short intervals between surveys (up to three days) (Fig. 2), and the second is represented by two longer periods between BMs: (i) from four to 11 days after the first sighting, and (ii) from 12 to 30 days after the first sighting (Fig. 3). The three periods were proposed to facilitate the classification and to attribute greater precision in the analysis between two carcasses, because there is a variation between metrics according to the period between two BMs.

In addition to the primary metrics of spatial distance and time lag between registered strandings among BMs, secondary biological metrics were determined to help compare observations. These biological metrics (bio) refer to information on the degree of decomposition (DD), total length (TL), and approximate total length; both teams used the same protocols for assigning DD and measuring TL. The secondary metrics depend on the species for which matching is being sought. To determine DDs, a decomposition category was used as indicated by Geraci and Lounsbury (2005), where: 1 = alive; 2 = freshly dead; 3 = moderate decomposition; 4 = advanced decomposition; and 5 = mummified.

Information on the degree of decomposition at first sighting as soon as a stranding occurred helps to project the decomposition into the future. Observing the chronology of the state of decomposition informs if there was an evolution or not in the comparison between two carcasses that represent a potential match (chronologically, a carcass cannot improve its decay status). When this does not occur in geographically close records, it characterizes that they are different animals. On the other hand, when the decomposition evolves as expected, it needs to be compatible with the time interval between successive BM records. For intervals longer than 10 days, degrees of decomposition 1 or 2 are disregarded as a resighting. As for TL, it was stipulated that a maximum variation of 10% was acceptable (as margin of error) concerning the first measure of the suspicious carcass. Suspicious carcass is one that is suspected to be a resighting, which has no geographical difference from one observed in the sighting or is in a close geographic range (primary metrics), however, it is necessary to verify the secondary metrics to confirm the ‘match’. Finally, and to aid in the comparisons, field notes taken by members of the research teams, if any, were indicated.

Figure 2. Flowchart of how a match of franciscana (Pontoporia blainvillei) carcasses was consolidated in the unified strandings database (UD) and classified. The proposed methodology is for beach surveys separated in time for up to three days. BM = beach monitoring effort; DD = degree of decomposition; TL = total length; NEMA = Núcleo de Educação e Monitoramento Ambiental; FURG = Universidade Federal do Rio Grande. Geographical differences (geo diff) result from variations between pairs of coordinates (ΔLAT; ΔLONG), measured in decimal degrees.
Figure 3. Flowchart of how a match of franciscana (*Pontoporia blainvillei*) carcasses was consolidated in the unified strandings database (UD) and classified. The proposed methodology is for beach surveys separated in time for up to thirty days. BM = beach monitoring effort; DD = degree of decomposition; TL = total length; NEMA = Núcleo de Educação e Monitoramento Ambiental; FURG = Universidade Federal do Rio Grande. Geographical differences (geo diff) result from variations between pairs of coordinates (ΔLAT; ΔLONG), measured in decimal degrees.

Usually, the notes were related to anthropic interactions (e.g. net filament brand, type of fishing net, visual description of injuries), and to re-sightings (e.g. spray marks on the carcass to inform that it had already been recorded by one of the institutions).

In step three, once the primary and secondary metrics had been defined, it was necessary to logically concatenate them for decision-making about the occurrence of a match. This was organized in the form of a flowchart, with the help of Lucidchart software, a free online tool for diagrams and visual communication (Lucidchart, 2022).

Results

The variables (and their standardizations) selected for the composition of the UD were: date (year-month-day), geographic coordinates (latitude and longitude, in decimal degrees), direction of beach survey (north or south), km from stranding (distance traveled from the initial reference point of the BM to the carcass), total and approximate body length (in centimeters), sex, degree of decomposition and observations from field monitors. This set of variables was designed to provide as much detail as possible about each stranding.

From the variables chosen to compose the UD, metrics used in the matching process were elaborated, among which it is possible to emphasize: the intervals between monitoring: i) 0-3 days, ii) 4-11 days, and iii) 12- 30 days; the distance between the carcasses, extracted from the difference between geographic coordinates; DD intervals, associated with the carcass’ time on the beach; flexibility between TL measurements, value of the second sighting lower or higher than 10% in relation to the first; other field notes, which indicates the comparison between the notes of the beach observers.

Match flowcharts

The match between stranding records from different BMs is a subjective value to inform whether they belong to the same carcass. This value – which classifies a record as excellent (1), good (2), or regular (3) candidates for being a match – considers spatial and temporal variations between strandings within a maximum period of 30 days.

When the monitoring carried out by the groups took place on the same day, or at short intervals (up to three days), the proposed comparison takes the number of stranded individuals into account and the difference between geographic positions of the carcasses as main indicators of a match. For confirmation...
purposes, biological information, whenever available, was also compared, i.e. carcasses with little or no change, associated with TL and the evolution of DD, in relation to the first sighting. In this short period of time (0-3 days), when the total number of strandings is different between the records by institutions, a step prior to geographic and biological comparison, based on the records ‘km of stranding’ is indicated. This is the length of beach traveled, from km 0 to the carcass; extra efforts at the output may introduce biases, potentially altering the real value of the stranding location. Therefore, it is not a reliable metric for confirming a resight against geographic coordinates, but can represent a useful and quick step. The objective is a quick verification regarding the spatiality of the record on the beach; when there is equality in the records, it is easier in the later stage, otherwise a verification of the records via geographic and biological comparison is indicated to confirm a match. This is an optional step, in contrast to other steps.

When there is a long-time lag between successive BMs (Fig. 3), the proposed flowchart encompasses more details for decision-making, which emphasizes the need to include secondary metrics as supporting evidence for the match. These periods - that configure one week after the short period (4-11 days between successive BMs), and 18 days (12-30 days between successive BMs, to complete the decomposition interval described in the literature) - include larger changes associated with carcasses and advanced decomposition states, especially between 12 to 30 days. Between 4 and 11 days, the decomposition states were not indicated in Fig. 3, because that is considered a transition period, which makes it difficult to assign more assertive metrics.

**Primary Metric**

The differences obtained between geographic coordinates and the corresponding distance (in meters) for the 372 pairs of strandings recorded on the same day by both institutions were plotted via scatter plots and histograms (Fig. 4A-E). To make interpretations easy, the percentiles of the ordered distribution of the differences and of the respective values in meters are presented in Table 1.

It was possible to observe that most of the differences (80%) are between zero, or no distance, and approximately 900 m, which indicates that there are differences between the values of records of two carcasses that represent a potential match (Fig. 4 and Table 1), where the ideal is that this difference is null.

**Table 1.** Percentiles of variation between latitudes and longitudes, and their respective distances (geo dist) in meters, of two suspicious stranded franciscana (*Pontoporia blainvillei*) recorded on the same day by both intuitions (FURG and NEMA), between the years 2000 and 2020 on the south coast of Rio Grande do Sul, southern Brazil.

<table>
<thead>
<tr>
<th>Minimum</th>
<th>80%</th>
<th>90%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ LON</td>
<td>0.0</td>
<td>0.006445</td>
<td>0.020646</td>
</tr>
<tr>
<td>Δ LAT</td>
<td>0.0</td>
<td>0.006169</td>
<td>0.025390</td>
</tr>
<tr>
<td>geo dist</td>
<td>0</td>
<td>993.03</td>
<td>3810.93</td>
</tr>
</tbody>
</table>

However, it is considered a good result, since in the greatest variations (between 3.8 km and 11.9 km) occurred in only 10% of the differences. There is reasonable plausibility that it may have been the same carcass in these cases given that TL and DD were checked. The percentiles associated with the 90, 95 and 100% (maximum) cuts (which are equivalent to 10% of the variations) are highlighted, as encompass the possibility of large distances between two compatible carcasses (observed ‘dist geo’), so all resulting values are important in the context of the match, however, other metrics are needed for confirmation of the match. With the check between the carcasses, and confirmation of the compatibility of the other information collected (i.e., they are the same, or possibly the same), and to establish intervals that encompass even the greatest variability (in terms of large extensions), matches and their respective geographic ranges were proposed (see Table 2).

**Table 2.** Classification of the ‘match’ interval according to the variations in geographic position generated between two suspected strandings of franciscana (*Pontoporia blainvillei*) observed on the same day, by different institutions (FURG and NEMA) to the north and to the south breakwater of the Rio Grande Mouth Bar (RS, Brazil), 2000-2020. The ‘geographic range of the differences between two strandings’ is the result of the value of the differences between latitudes and between longitudes, therefore, they are differences in decimal degrees.

<table>
<thead>
<tr>
<th>Match</th>
<th>Subjective Value</th>
<th>Geographic range of differences between two strandings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>1</td>
<td>0 ↔ ± 0.025</td>
</tr>
<tr>
<td>Good</td>
<td>2</td>
<td>±0.025 → ± 0.05</td>
</tr>
<tr>
<td>Regular</td>
<td>3</td>
<td>±0.05 → ± 0.1</td>
</tr>
</tbody>
</table>

**Amount of data and resulting strandings**

Overall, 7,658 lines of BM data in the study area, in which each data line corresponds to a stranded franciscana (2000-2020), were compared (NEMA, with n = 3,029; FURG, with n = 4,629). With the application of the method established for the construction of the UD, 5,467 stranded franciscanas were counted – among which 1,812 matches were classified as ‘excellent’, and 97 were ‘good’. Duplicates identified in these classifications were removed. As for the ‘regular’ matches, 30 pairs were counted - equivalent to 60 franciscanas - and maintained in the UD. Due to the discrepancy in relation to the proposed metrics, ‘regular’ combinations were not considered as recounts, and each registry was kept in the database as an individual. The low number of ‘regular’ combinations shows that the proposed method encompasses most situations, thus, it was effective for finding matches.

**Discussion**

Strandings of marine mammals have been documented for decades along the southern Brazilian coast (e.g. Prado et al., 2016). They provide valuable information which helps to
understand the stranding patterns in a certain location, and to investigate their possible causes (Geraci and Lounsbury, 2005; Pyenson, 2010; Prado et al., 2013; 2016). According to studies by Prado et al. (2013) involving mark-recapture methods applied to bycaught franciscana carcasses in southern Brazil, a small proportion of the individuals marked that die at sea end up washed ashore as a stranding record.

The current work is presented in the form of a guide, outlining steps on how to proceed to elaborate a single stranding database from the comparison and merging of different records of beach monitoring efforts. A step-by-step outline is used so that other researchers of aquatic mammals – or marine turtles, or sea birds – may adapt it when constructing their own databases, taking the metrics presented here as a basis, expanding, and improving them according to the investigated species and the region of study. To unify stranding databases while reducing or avoiding organism recounts, it is necessary to investigate each database to consolidate trustworthy metrics based on survey data description. In this case, it is important to understand which variables are recorded, and which similarities the databases share. Among what they share in common, it is important to capture what is different between the collections (annotations) carried out, in order to delineate reliable limits (when surveys take place on the same day). Regarding the UD of franciscanas, the creation of a margin of error within a comparison was carried out using the geographic variations found in the data and bibliographic values associated with time and decomposition states (Geraci and Lounsbury 2005; Prado et al., 2013) and total length measured of a specimen. The choice for these variables was due to the possibility of being extracted from the values of the database, in the case of geographic coordinates, and by the available local literature, in the case of the decomposition time.

Considering the franciscana and the study region, seasonality with well-defined periods can affect the speed of decomposition of the carcass, thus affecting the time of its availability on the beach; at colder temperatures the disappearance of the carcass can be delayed (Geraci and Lounsbury, 2005). In southern Brazil, seasonality has important effects on the entire length of the beach, and consequently, on the activities that take place on it (Nimer, 1989). According to Calliari and Klein (1993), when analyzing the 215 km extension of beach from Cassino to Chui (south direction of the BM), interactions between the coastal zone, the oceanographic dynamics and the present climatology provide significant spatial variations in beach characteristics. This creates different morphodynamic profiles on the same beach – which in turn generates variations in currents, presence of sandbanks, differences in wave energy, among others. These variations cause changes in physical, chemical, and biological characteristics and affect carcasses beyond the speed of decomposition, which can force the mobility of the body on the beach, for example, in periods of surf. In the context of beach dynamics, it is essential to emphasize the importance of comparison metrics that go beyond geographic distance, referred to as secondary metrics, to validate a match between two carcasses that are in different positions but

Figure 4. Scatter plots and density histograms resulting from differences between longitude (A and B) and latitude (C and D) data and the respective geographic distances between records of franciscana (Pontoporia blainvillei) strandings, in meters (E). Positive and negative values are due to differences between measurements.
suspected to be the same. In addition, considering that there may be carcass mobility, associated with oceanographic dynamics, efforts are recommended in research with the applicability of methods and equipment that allow monitoring of this mobility, to trace more precise quantitative geographic metrics, and even understand more about the decomposition time in a specific region.

Thus, values associated with decomposition time and region studied must also be adapted by organism and location. For example, in southern Brazil (specifically in the study area), summers with average daily temperatures of 28°C and winters with average daily temperatures of 13°C have been recorded (Nimer, 1989), or a range of 15°C between the two seasons, which can affect the speed of decomposition of franciscanas, and TL. In the case of franciscanas, considering the FMAs (Secchi et al., 2003; Di Beneditto et al., 2010; Cunha et al., 2014), and possible seasonal differences according to geographic specificities, a time series with daily or weekly collection of TL and DDs for a period of time long enough to cover the carcass decomposition time, would be ideal for confirmation and understanding of the evolution of carcasses. Despite the emphasis on method flexibilization according to location and the handling organization, there is a need to standardize the annotation format and measurement units (e.g. TL in meters or centimeters), simplifying the number of variables when redundant information is present (e.g., choose between decimal degrees or degrees, minutes and seconds or UTM coordinates for geographic information). Standardization and choice of specific variables help the field observer to avoid errors during the data digitization phase and can facilitate the understanding by researchers and analysts unfamiliar with the original database. Another point is to value the use of marks (e.g. color spray) to communicate between institutions whether a carcass has been counted already, avoiding recounts. However, mechanisms that optimize the identification of the animal – such as an alternative when no spray is available on site, or the paint completely disappeared in a period shorter than expected - are recommended. Furthermore, to reduce errors associated with recounts and delineate more reliable geographic intervals for searching for carcasses, considering beach mobility due to coastal dynamics, we emphasize the need for more mark-recapture studies applied to bycaught carcasses at sea, and focus on beach decomposition times (as carried out by Prado et al., 2013) applied to already stranded carcasses. In addition to monitoring the decomposition times of the beached carcasses, the same type of study could focus on TL measurements between short periods, and what influences decomposition.

Finally, this study pioneers in its attempt to generalize procedures to merge stranding databases produced by different institutions and within the same area. The proposed method aimed to create a step-by-step guide to help researchers working on marine mammal conservation to merge long-term stranding databases, which is an important basis for conservation research.

Acknowledgments

This study was carried out in partnership between the Laboratório de Estatística Ambiental, of the Instituto de Matemática, Física e Estatística (IMEF), FURG, ECOMEGA/IO, and NEMA. There were 21 years of biweekly beach monitoring, so without the collaboration of field monitors, teachers, drivers, and volunteers, as well as logistical support to carry out the fieldwork, this work would not have been possible. This research was partially supported by an environmental offset measure established through a Consent Decree/Conduct Adjustment Agreement between Petrorio and the Brazilian Ministry for the Environment, with the Brazilian Biodiversity Fund – FUNBIO as implemener, to the RS Ports – Port Authority of the Ports of Rio Grande do Sul and Petrobras, Yaqu Pacha and by CAPES, through a research grant.

References


