

Interpolation of animal tracking data in a fluid environment

Yann Tremblay^{1,*}, Scott A. Shaffer¹, Shannon L. Fowler¹, Carey E. Kuhn¹, Birgitte I. McDonald¹, Michael J. Weise¹, Charle-André Bost², Henri Weimerskirch², Daniel E. Crocker³, Michael E. Goebel⁴ and Daniel P. Costa¹

¹University of California, Santa Cruz, Long Marine Laboratory, Center for Ocean Health, 100 Shaffer Road, Santa Cruz, CA 95060, USA, ²Centre d'Etude Biologiques de Chizé, 79360 Villiers en Bois, France, ³Department of Biology, Sonoma State University, Rohnert Park, CA 94928, USA and ⁴NOAA, National Marine Fisheries, 8604 La Jolla Shores Drive, La Jolla, CA 92038, USA

*Author for correspondence (e-mail: tremblay@biology.ucsc.edu)

Accepted 7 November 2005

Summary

Interpolation of geolocation or Argos tracking data is a necessity for habitat use analyses of marine vertebrates. In a fluid marine environment, characterized by curvilinear structures, linearly interpolated track data are not realistic. Based on these two facts, we interpolated tracking data from albatrosses, penguins, boobies, sea lions, fur seals and elephant seals using six mathematical algorithms. Given their popularity in mathematical computing, we chose Bézier, hermite and cubic splines, in addition to a commonly used linear algorithm to interpolate data. Performance of interpolation methods was compared with different temporal resolutions representative of the less-precise geolocation and the more-precise Argos tracking techniques. Parameters from interpolated sub-sampled tracks were compared with those obtained from intact tracks. Average accuracy of the interpolated location was not affected by the interpolation method and was always within the precision of the tracking technique used. However, depending on the species tested, some curvilinear interpolation algorithms

produced greater occurrences of more accurate locations, compared with the linear interpolation method. Total track lengths were consistently underestimated but were always more accurate using curvilinear interpolation than linear interpolation. Curvilinear algorithms are safe to use because accuracy, shape and length of the tracks are either not different or are slightly enhanced and because analyses always remain conservative. The choice of the curvilinear algorithm does not affect the resulting track dramatically so it should not preclude their use. We thus recommend using curvilinear interpolation techniques because of the more realistic fluid movements of animals. We also provide some guidelines for choosing an algorithm that is most likely to maximize track quality for different types of marine vertebrates.

Key words: tracking, telemetry, Argos, geolocation, GPS, Bézier, cubic, hermite, spline, albatross, penguin, sea lion, fur seal, elephant seal, booby, seabird, marine mammal.

Introduction

Ecological studies of marine vertebrates have proven challenging due to our inability to observe individuals for long periods. Our understanding of what marine vertebrates do when they are out of sight thus relies almost exclusively on recording and/or transmitting electronic devices. Such devices can provide information on the geo-position of the tracked animal for a given time, as well as other behavioral, physiological or environmental information (Kooyman et al., 1992; Weimerskirch et al., 1997; Wilson et al., 1995). Satellite telemetry (using the Argos system), geolocation (by recording day length), and GPS (Global Positioning System) telemetry are the main tracking techniques, with satellite telemetry being used most commonly. These techniques differ with respect to

two fundamental characteristics: (1) the location accuracy and (2) the frequency at which locations are obtained. These two characteristics determine track quality and generally imply two levels of post processing: filtering and interpolating. Filtering of tracking data (by removing unlikely locations) addresses the problem of location inaccuracy and has received more attention than interpolation (Austin et al., 2003; McConnell et al., 1992; Sibert et al., 2003).

Interpolation of tracking data addresses the problem of uneven sampling. Animals are often equipped with instruments that record environmental and/or behavioral parameters in addition to a tracking device. These instruments generally have sampling rates that differ from the sampling rate of the tracking device. Therefore, by interpolating tracking data, each

measured parameter can be matched to an estimated location. Interpolation is also important because it provides locations that are equally spaced in time, which is necessary for further evaluation of habitat use (BirdLife International, 2004).

Either by choice or lack of an alternative, most authors represent their tracking data as straight lines between recorded points and do not interpolate their data (Block et al., 2005; Folkow et al., 2004; Pütz et al., 2000). The advantages of linear interpolation are its simplicity and that it represents the most conservative path an animal transits between two consecutive locations. However, straight lines are not consistent with fluid dynamics in which subjects moving in a fluid environment (air or water) probably do not follow straight lines. Fluid media are kingdoms of curves, being described by flows, vortices, turbulences and gradients (Vogel, 1994). For example, acoustic tracking of both oceanographic floats (Fratantoni and Richardson, 1999) and any seabird observed for a short time at sea classically shows a sinuous path (Alerstam et al., 1993; Weimerskirch et al., 2000). Additionally, navigators attempting to travel in a straight trajectory need to constantly correct vessel orientation to maintain the bearing. The corollary with tracking data is that a linearly interpolated track between relatively spaced locations (in time) is unrealistic, because sinuous movements are collapsed into single positions, which are not necessarily obtained when the animal actually turns. Because of the fluid properties, particles in the atmosphere or the oceans move in a curvilinear manner in relation to forces from density gradients and to attraction and Coriolis forces (Vogel, 1994). Most tracked animals are not passive bodies in fluids, but rather their movements are affected by these forces, either directly (e.g. wind, current) or indirectly (e.g. eddy targeted by a predator as a foraging zone).

Curvilinear tracks are consistent with marine animals moving along oceanic features such as eddies, sea-surface height anomalies, fronts or weather systems, which are all fluid, curvilinear structures (Ferraro et al., 2004; Murray et al., 2002; Polovina et al., 2000; Ream et al., 2005; Weimerskirch et al., 2002). Curve interpolation does not conflict with a straight path, because a straight line can be mathematically conceived as a particular curvilinear function.

Historically, the intuitive logic in using curves can be seen in the very first study, 15 years ago, describing satellite-tracked flying seabirds (Jouventin and Weimerskirch, 1990). The authors presented two figures of tracks: one using straight lines and the other using an undefined curve. More recently, use of Bézier curves and splines has been suggested as another way of representing paths (Turchin, 1998). Curvilinear interpolation thus appears to be a more natural way of interpolating marine animal tracks, especially in a fluid environment. However, to our knowledge, no study has ever attempted to use curvilinear interpolation for animal tracking data.

The difficulty in using curvilinear interpolation is that, unlike a straight line, an infinite number of curves can be mathematically calculated between two recorded locations. Consequently, the choice of a mathematical algorithm used to

interpolate along curves can modify the resulting interpolated tracks, thus emphasizing the need to evaluate the effects of different algorithms and to assess the risk of introducing errors to the track data.

This paper is the first to interpolate tracking data of several marine animals using various mathematical algorithms. Our goals were to propose alternatives to the linear method for interpolating tracking data in fluid media and to evaluate the potential pitfalls and benefits associated with curvilinear interpolation methods.

Materials and methods

Datasets

Tracking data from 10 species were chosen to represent four major groups of marine vertebrates: flying seabirds (males and females) (Laysan albatross, *Phoebastria immutabilis* Rothschild 1893; black-footed albatross, *P. nigripes* Audubon 1839; black-browed albatross, *Thalassarche melanophrys* Temminck 1828; red-footed booby, *Sula sula* Linnaeus 1766), penguins (males and females) (king penguin, *Aptenodytes patagonicus* Miller 1778; macaroni penguin, *Eudyptes chrysolophus* Brandt 1837), otariids [California sea lion (males only), *Zalophus californianus* Lesson 1828; Australian sea lion (females only), *Neophoca cinerea* Péron 1816; Antarctic fur seal (females only), *Arctocephalus gazella* Peters 1875] and phocids (males and females) (northern elephant seals, *Mirounga angustirostris* Gill 1966). Given the extreme differences in size and foraging ecology, male and female northern elephant seals were considered as two separate species for the purpose of this study (Le Boeuf et al., 2000).

Tracking data for black-browed albatrosses and red-footed boobies were obtained using GPS tags (we used two GPS tracks per species, obtained from two different individuals), whereas all other tracks were obtained by satellite telemetry (Argos) using platform terminal transmitter (PTT) and appropriate attachment methodology (we used three Argos tracks per species, obtained from three different individuals). Information related to device characteristics, study sites and periods are given in Table 1.

Argos data were filtered using the filtering algorithm of the IKNOS (Greek for step, track, tracking, footprint...) toolkit (Y. Tremblay, unpublished). This algorithm uses several criteria in order to remove unlikely location: (1) realistic travel speeds of a subject between two fixes, (2) the change in azimuth of a subject between successive fixes, (3) the Argos location class, (4) the time elapsed between two consecutive fixes and (5) whether a location was on land or at sea. The IKNOS Argos filtering program allows the user to set limits for some of these criteria. These limits were selected and kept consistent within each species.

Filtered data were thereafter referred to as 'tracks' and were considered, by default, as being of the best quality that a tracking method permits. Although different filtering techniques can lead to slightly different tracks, there is no way to verify the accuracy of a given filtering process. The output

Table 1. *Technical information regarding methods for collecting the tracks used in this study*

Species	Tracking device	PTT power (W)	Transmission cycle (repetition rate or sampling interval)	Study location	Study period
Albatrosses					
<i>P. immutabilis</i>	Microwave Pico-100	0.1	Continuous (90 s)	Tern Island, USA	Incubation
<i>P. nigripes</i>	Microwave Pico-100	0.1	Continuous (90 s)	Tern Island, USA	Incubation
<i>T. melanophrys</i>	Newbehaviour GPS	–	(10 s)	Kerguelen Islands, France	Brooding
Boobies					
<i>Sula sula</i>	Newbehaviour GPS	–	(10 s)	Europa Island, France	Incubation
Penguins					
<i>A. patagonicus</i>	Sirtrack-Kiwisat	0.5	Duty cycled 6h on, 6h off (45 s)	Crozet, France	Incubation
<i>E. chrysolophus</i>	Sirtrack-Kiwisat	0.5	Continuous	Kerguelen Islands, France	Incubation
Otariids					
<i>Z. californianus</i> males	SMRU-SRDL		Continuous	California coast, USA	Wintering migration
<i>N. cinerea</i> females	Telonics		Continuous	Kangaroo Island, Australia	Pup rearing
<i>A. gazella</i> females	Wildlife Computers Spot 2	0.5	Continuous (45 s)	Livingstone Island, Antarctica	Pup rearing
Phocids					
<i>M. angustirostris</i> males	Telonics		Continuous	Guadalupe Island, Mexico	Post-molt migration
<i>M. angustirostris</i> females	Telonics		Continuous	Ano Nuevo, USA	Post-molt migration

from the filtering process is always considered satisfactory, on a more or less arbitrary basis (generally by visual inspection). Interpolation was done for a given set of filtered locations, independent of their actual accuracy. Filtering method had no ultimate impact on the interpolation calculations, so filtering parameters are not shown.

Interpolation algorithms

Six different mathematical algorithms (hereafter called curves, even when linear) were selected to interpolate tracks. The choice for these curves was mainly driven by their popularity in mathematical computing (Angel, 2003; Mortenson, 1997) and ease of implementation using Matlab v. 7.0 (The MathWorks, Natick, MA, USA).

Linear algorithm

Linear interpolation was computed to provide a comparison with other curves. This is the easiest, most conservative and most common interpolation method used to date.

Bézier curves

Since their formulation in the 1970s, Bézier curves have obtained dominance in the typesetting and design software industry (Bartels et al., 1998; Piegl, 1993). Currently, Bézier curves are found nearly everywhere in our everyday life, and web resources for equations, codes, courses and representations are plethora (see, for example, http://en.wikipedia.org/wiki/Bezier_curve). We used piecewise cubic Bézier curves along recorded tracks. The angle at which

the curve hits each point was controlled by the tangent vector of the angle defined by three consecutive points. A detailed explanation of the algorithm can be found at <http://astronomy.swin.edu.au/~pbourke/curves/bezier/cubicbezier.html>. Piecewise cubic Bézier curve computation allows definition of a parameter (μ) controlling elasticity of the curve. Since different choices for μ give different curves, we ran three versions of Bézier curves, with $\mu=0.1$ (straighter), 0.2 and 0.3 (more curved). The choice of these three values resulted from preliminary tests, which are explained in the Results.

Hermite splines and cubic splines

Piecewise cubic hermite interpolating polynomials were computed using the 'pchip' function in Matlab, following Fritsch and Carlson (1980) and Kahaner et al. (1988). Cubic spline interpolation was computed using the 'spline' function in Matlab, following de Boor (1978). Built-in functions of Matlab were run unmodified.

Strategy used to compare curve performances

Because we do not truly know where an animal is located between two recorded locations, it is impossible to compare any interpolated location to a reference location. Therefore, we extracted (i.e. sub-sampled) a set of locations from each track and used these locations as references. This resulted in tracks with fewer locations than the original tracks. The tracks were then interpolated using the different algorithms. For each curve, the set of interpolated locations corresponding (in time)

to the extracted set of reference locations was selected. The corresponding distance between them was calculated and further compared among algorithms. The process is illustrated in Fig. 1.

For Argos tracks, reference positions were extracted in two ways in order to investigate the effects of temporal resolution of tracks. First, we extracted the number of reference positions so that only one location per day in the track was left (closest location to local midday of each day). This process allowed us to interpolate tracks with a similar temporal resolution to tracks obtained using the geolocation positioning technique (Hill, 1994; Shaffer et al., 2005; Teo et al., 2004; Wilson et al., 1992). These tracks will hereafter be referred to as geolocation-like Argos tracks (Fig. 2). Australian sea lion tracks were too short (around two days) to be processed this way and were discarded from this part of the analysis.

Second, we randomly extracted a maximum of 10% of the total number of locations. This process resulted in tracks that were only slightly modified and thus they were similar to the original tracks. These tracks will hereafter be referred to as Argos tracks. Because curve calculation is sensitive to angles between locations, we did not allow reference locations to be selected consecutively. Extracted reference locations had to be separated by at least two locations so that angles on each side of a removed location would not be affected by the removal of another location. Because some tracks were short and had few locations, this process did not allow for the extraction of more than a couple of points at a time in those tracks. Then, we iterated the process several times until we obtained at least eight distinct extracted points (this was the maximum we could obtain given the total number of locations in these tracks).

Due to high spatial and temporal resolutions, GPS tracks do not need to be interpolated. The number of reference locations extracted from GPS data was thus calculated to provide tracks with a temporal resolution similar to or slightly better than that of the best Argos tracks [one location per hour, randomly spaced by at least 100 s (10×10 s sampling interval, arbitrarily)]. These tracks are referred to as Argos-like GPS tracks. Because GPS tracks had one position every 10 s, the number of reference locations was high, and consecutive locations in the track were thus highly auto-correlated (not estimated). For this reason, and in order to reduce the effects associated with pseudo-replication (Hurlbert, 1984), only 30 of the reference locations (randomly selected for each track) were used in the analysis.

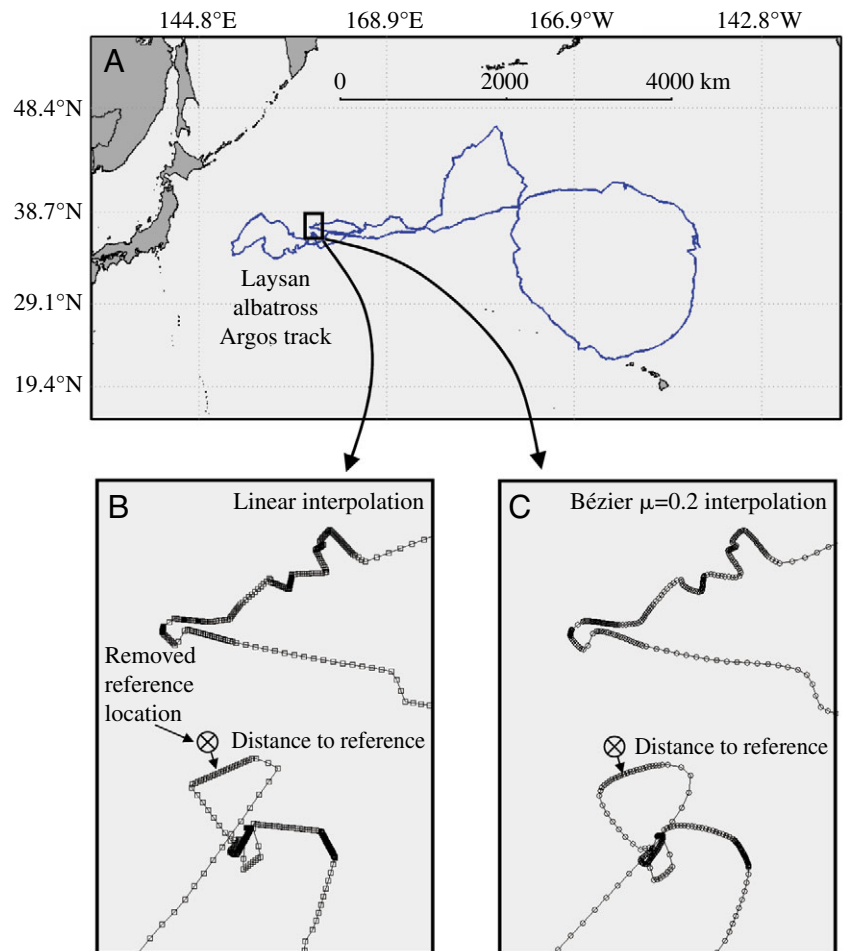


Fig. 1. Laysan albatross Argos track (A), and selected examples of linear (B) and Bézier (C) interpolation of this track (every 10 min). In B and C, the circled cross represents an Argos position that was removed to use as a reference position. The distance between this position and the corresponding interpolated location was calculated for each mathematical algorithm that we used (see Materials and methods). Note the possibility to visualize transit speed in interpolated tracks.

The start and end locations of each track were never removed, nor were they used as a reference location. Geolocation-like Argos, Argos and Argos-like GPS tracks were analyzed separately.

Data processing

Since tracks were recorded in an unprojected Greenwich coordinate system (latitude–longitude coordinates refer to a spherical coordinate system), they were first transformed (i.e. flattened) to a projected Cartesian coordinate system, and then interpolated data were transformed back for distance calculations. All calculations of distance were done following the great circle distance on the Earth geoid, thus taking into account the Earth's curvature. For the purpose of this study, the time of each location of the tracks was rounded to the nearest minute, and interpolated locations were also calculated for each minute.

In a Cartesian coordinate system, piecewise curves are

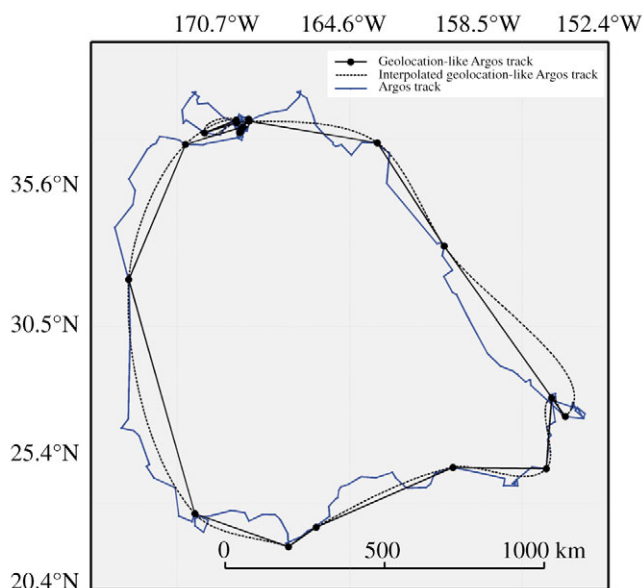


Fig. 2. Selected example of a black-footed albatross track, illustrating some of the various versions of the track. In this example, the geolocation-like Argos track was interpolated using the Bézier algorithm with $\mu=0.3$ (dashed line).

computed for equally spaced values on the x and y axes (corresponding here to each time unit). Consequently, interpolated locations were not equally spaced in the plane. This resulted in artificial non-linear speed between two consecutive interpolated locations. To overcome this problem, we over-sampled our interpolated data and then used a subset of these points (equally spaced locations by distance). The precision of this process was not mathematically exact, so interpolated locations were almost equally spaced. The level of over-sampling (50 times, i.e. one location every 1.2 s) was calculated to ensure that this approximation could be neglected.

Statistics

General linear models were computed using SYSTAT 10 (SPSS Inc, Chicago, IL, USA). Distribution of distances between reference locations and interpolated locations was skewed to the left. The average of such a distribution is off-centered proportionally to the extent of the tail. For this reason, the median, minimum and maximum (instead of the mean \pm s.d.) were used to describe the results (unless stated differently). Distances between reference locations and interpolated locations were \log_{10} transformed before performing inferential statistics. Statistical significance was considered at the $P<0.05$ level.

Results

Argos track characteristics

Our initial dataset of Argos tracks showed extensive differences between tracks. While albatrosses and phocids

traveled at scales of several thousand kilometers, most otariids and penguins traveled at scales of only several hundred kilometers (Table 2). With the Argos system, spatial accuracy is approximated by a scale of quality classes, 3, 2, 1, 0, corresponding to accuracies of <150 m, 150 to <350 m, 350 to <1000 m, >1000 m, and quality classes A, B, Z with no associated accuracies. Empirical studies have shown that these accuracies should not be taken in *stricto sensu*, and that average accuracy is commonly in the order of several kilometers (Fernández et al., 2001; Le Boeuf et al., 2000). Thus, it is complicated to establish an average accuracy for a given track. However, the difference in proportions of each quality class in a track revealed that phocids had relatively higher proportions of low quality locations than most other groups (Table 2). Species with the highest proportions of more accurate quality classes included *Z. californianus*, *A. gazella* and *E. chrysolophus* (Table 2).

Temporal resolutions were also extensively different between tracks, with tracks lasting from 2.3 to 226 days (Table 2), and with tracks composed of 1–17 locations per day on average. Because the transmitter's signal does not pass through water, non-diving species (i.e. albatrosses) had a higher number of locations per day than other species (Table 2).

Tracking data obtained through the Argos system can be affected by a high number of variables, such as quality and power of the transmitter, transmitter attachment location, satellite coverage and animal behavior. We therefore also obtained differences between track characteristics within species. Both proportions of quality classes and number of locations per day showed relatively large differences between different individuals from a given species (Table 2).

Effect of interpolation method on the accuracy of the estimated location

Accuracy of the interpolated locations was different between species and between individuals within species but not between interpolation methods. No interaction between species and interpolation method was found. This was true in interpolated geolocation-like Argos, Argos and Argos-like GPS tracks. Statistical data are given in Table 3, and median values are summarized in Table 4, by species. Errors of the interpolated locations were greater in the geolocation-like Argos tracks than in the Argos tracks. Errors were also greater in fast-flying albatrosses (medians: 56.4–65.4 km and 10.6–12.8 km in geolocation-like Argos and Argos tracks, respectively) than in non-flying animals (medians: 4.8–10.4 km and 1.5–6.8 km in geolocation-like Argos and Argos tracks, respectively).

Comparison of curve interpolations versus linear interpolation

Accuracy of the estimated position

The interpolation methods used in this study had no impact on the accuracy of the estimated locations, as curve interpolation methods did not produce larger errors than the linear interpolation method.

Table 3. *Statistical results of General Linear Model computed with $\log_{10}(\text{distance to reference})$ as a dependent factor for each type of track (see Materials and methods for details)*

Source	Sum of squares	d.f.	Mean square	F-ratio	P
Geolocation-like Argos tracks					
Method	0.423	5	0.085	0.510	0.769
Species	5141.110	7	734.444	4433.847	<0.001
Method \times species	2.881	35	0.082	0.497	0.994
Individual (species)	86.522	16	5.408	32.643	<0.001
Error	4890.566	29522	0.166		
Argos tracks					
Method	0.222	5	0.044	0.273	0.928
Species	375.934	8	46.992	289.942	<0.001
Method \times species	0.176	40	0.004	0.027	1.000
Individual (species)	14.978	18	0.832	5.134	<0.001
Error	500.805	3090	0.162		
Argos-like GPS tracks					
Method	1.203	5	0.241	0.229	0.950
Species	31.194	1	31.194	29.648	<0.001
Method \times species	0.282	5	0.056	0.054	0.998
Individual (species)	189.636	2	94.818	90.119	<0.001
Error	742.813	706	1.052		

Occurrence of more accurate locations

We verified if the curve-interpolation methods produced a higher or lower occurrence of more accurate locations (i.e. closer to reference) than the linear interpolation method. For each track, and for each of the five non-linear curves, the percentage of interpolated locations closer to the reference than locations obtained with the linear method was calculated (Fig. 3). Percentages over 50% indicated that the curvilinear method resulted in a higher number of more accurate locations than the linear interpolation method, and *vice versa*. Overall, the occurrence of more accurate locations using curves was between 40 and 60% in geolocation-like tracks, and between 30 and 70% in Argos and Argos-like GPS tracks (Fig. 3). For 21 of the 24 geolocation-like tracks (87.5%), 19 of the 27 Argos tracks (70.4%) and three of the four Argos-like GPS tracks (75%), at least one curvilinear interpolation method provided a higher number of more accurate locations than the linear method.

Closer examination reveals that some curves gave a higher occurrence of more accurate locations than others, depending on the species (Fig. 3). For example, all six geolocation-like Argos tracks of albatrosses were improved by using the hermite curve, with an average of 58% better locations in the tracks (Fig. 3). Similarly, for the Argos tracks, four of the six albatross tracks had a higher or equal number of more accurate locations when using the Bézier curve with $\mu=0.1$, 0.2 or 0.3, but with only a mean value of 52% better locations. We compiled a list of the best algorithms for each group of species based on the proportion of tracks improved (Table 5). At least one curve interpolation method could be identified as being better than or equal to the linear interpolation method for each of the studied vertebrate groups. Given the differences between tracks within a species, and the fact that we only analyzed three

tracks per species, this table could not be created accurately at the species level.

The tracks we used for elephant seals were not greatly improved, if at all, by using curvilinear algorithms. Essentially, those tracks were particularly linear (see Fig. 4A). The same observation was made in black-browed albatross Argos-like tracks. When the number of more accurate locations was reduced, it was generally reduced by only 10–15% (Fig. 3).

Total length of tracks

The lengths of interpolated geolocation-like Argos tracks were compared with the lengths of the corresponding original intact Argos tracks, and the lengths of the interpolated Argos-like GPS tracks were compared with the lengths of the original intact GPS tracks. Without exception, all curvilinear algorithms produced better estimates of the length of the tracks (i.e. closer to original track length) than the linear interpolation method. Also, without exception, either the Bézier curve with $\mu=0.3$ or the cubic curve (the most relaxed curves in our study) systematically produced the best estimate of track length. Track length estimates were less than the original track lengths for 99% of all tracks.

In geolocation-like Argos tracks, estimated track length proportions were consistent across species and were $81.3 \pm 10.5\%$ of original track lengths (mean \pm s.d., range = 56.2–102.5%). The most relaxed cubic splines produced track lengths, on average, 15.8% shorter, whereas the straighter linear algorithm produced track lengths 20.0% shorter. Compared with linear interpolation, and depending on the algorithm chosen, curvilinear interpolation increased the estimated track length by 0.3–4.2%.

In Argos-like GPS tracks of red-footed boobies and black-browed albatrosses, estimated track lengths ranged from 63.5 to

Table 4. Median (minimum–maximum) distances (km) between reference locations and interpolated locations for the six interpolation algorithms and for each species

Species	Interpolation method						N
	Linear	Bézier $\mu=0.1$	Bézier $\mu=0.2$	Bézier $\mu=0.3$	Hermite	Cubic	
Geolocation-like Argos tracks							
Albatrosses							
<i>P. immutabilis</i>	60.8 0.7–354.9	60.1 0.5–351.3	59.1 0.3–350.6	58.8 0.1–353.7	59.6 0.3–352.8	61.8 0.1–352.0	1147
<i>P. nigripes</i>	65.4 0.8–243.9	62.0 1.3–240.0	59.0 1.3–236.4	56.5 1.3–233.5	56.9 0.8–231.9	56.4 2.0–256.9	513
Penguins							
<i>A. patagonicus</i>	6.6 0.1–22.0	6.6 0.1–22.3	6.5 0.1–22.5	6.5 0.1–22.7	6.6 0.1–22.5	6.1 0.2–22.2	434
<i>E. chrysolophus</i>	5.0 0.7–60	5.0 0.7–60.1	4.9 0.4–60.1	4.8 0.2–60.1	4.9 0.9–60.1	5.3 0.2–60.2	143
Otariids							
<i>Z. californianus</i>	4.8 0–63.1	4.8 0–63.6	5.1 0–64.1	5.1 0–64.5	4.8 0–62.9	5.0 0.1–62.3	851
<i>N. cinerea</i>	–	–	–	–	–	–	–
<i>A. gazella</i>	10.4 0.3–51.3	9.8 0.6–51.2	9.4 0.8–51.0	9.2 0.7–50.9	9.4 0.5–50.6	9.2 0.4–49.4	251
Phocids							
<i>M. angustirostris</i> males	7.1 0.3–64.5	7.3 0.4–65.4	7.4 0.1–66.3	7.6 0.2–67.2	7.3 0.4–69.5	7.6 0.4–68.7	272
<i>M. angustirostris</i> females	8.1 0.2–78.6	8.2 0.3–78.6	8.2 0.3–78.6	8.2 0.3–78.6	8.1 0.2–78.6	8.3 0.3–78.6	1320
Argos tracks							
Albatrosses							
<i>P. immutabilis</i>	12.3 1.0–64.0	12.5 1.0–66.7	12.3 1.0–69.3	12.6 1.0–71.8	12.7 1.0–67.1	12.8 0.2–67.5	124
<i>P. nigripes</i>	11.5 1.6–58.7	10.8 0.8–53.7	11.4 0.6–48.5	11.7 0.7–43.7	10.6 0.5–58.2	10.6 1.0–57.8	56
Penguins							
<i>A. patagonicus</i>	1.9 0.3–10.7	1.8 0.1–10.8	1.9 0.2–10.8	1.9 0.3–10.9	1.9 0.2–10.7	2.1 0.3–10.9	48
<i>E. chrysolophus</i>	3.2 0.7–19.1	3.0 0.3–20.7	3.0 0.3–22.9	2.9 0.6–25.3	2.9 0.4–19.0	2.9 0.2–19.0	32
Otariids							
<i>Z. californianus</i>	1.5 0.1–28.5	1.5 0.1–28.4	1.5 0.1–28.3	1.5 0.1–28.1	1.6 0.1–28.2	1.5 0.1–28.1	77
<i>N. cinerea</i>	7.0 0.1–50.1	6.9 0.2–50.4	6.9 0.3–50.7	6.8 0.5–51.0	7.0 0.6–50.1	7.0 1.2–50.8	27
<i>A. gazella</i>	1.9 0.4–4.5	2.0 0.4–4.8	2.0 0.4–5.1	2.1 0.4–5.3	1.9 0.3–4.8	2.0 0.3–4.8	33
Phocids							
<i>M. angustirostris</i> males	6.1 1.2–160.4	6.0 0.9–161	6.0 0.4–160.8	5.9 0.5–160.6	5.9 0.6–160.9	6.8 1.1–160.6	51
<i>M. angustirostris</i> females	7.0 0.7–48.0	6.6 0.7–48.5	6.5 0.7–49.1	6.6 0.7–49.7	7.1 0.8–50.0	7.0 0.8–50.4	79
Argos-like GPS track							
<i>S. sula</i>	0.8 0.004–22.2	0.8 0.002–21.5	0.8 0.001–20.9	0.8 0.001–20.9	0.8 0.002–22.2	1.4 0.001–20.9	60
<i>T. melanophrys</i>	2.2 0.001–21.8	2.1 0.001–21.8	2.0 0.001–21.8	1.9 0.001–20.9	2.2 0.001–21.8	2.2 0.001–21.8	60

83.8% and from 28.3 to 43.2% of the original GPS track lengths, respectively. These estimates were therefore, on average, 33.5 and 64.8% shorter than the original track length in red-footed booby and black-browed albatross tracks, respectively.

Shape and plausibility of the curves

By visually inspecting the interpolated tracks, we noticed

that cubic splines produce oscillations and overshoots that are not consistent with original track data (Fig. 4B). This artifact is problematic because the original shape of the track was modified. By contrast, all the other algorithms we used produced turns that were tangential to each recorded location, thus giving conformal interpolated tracks without unexpected oscillations.

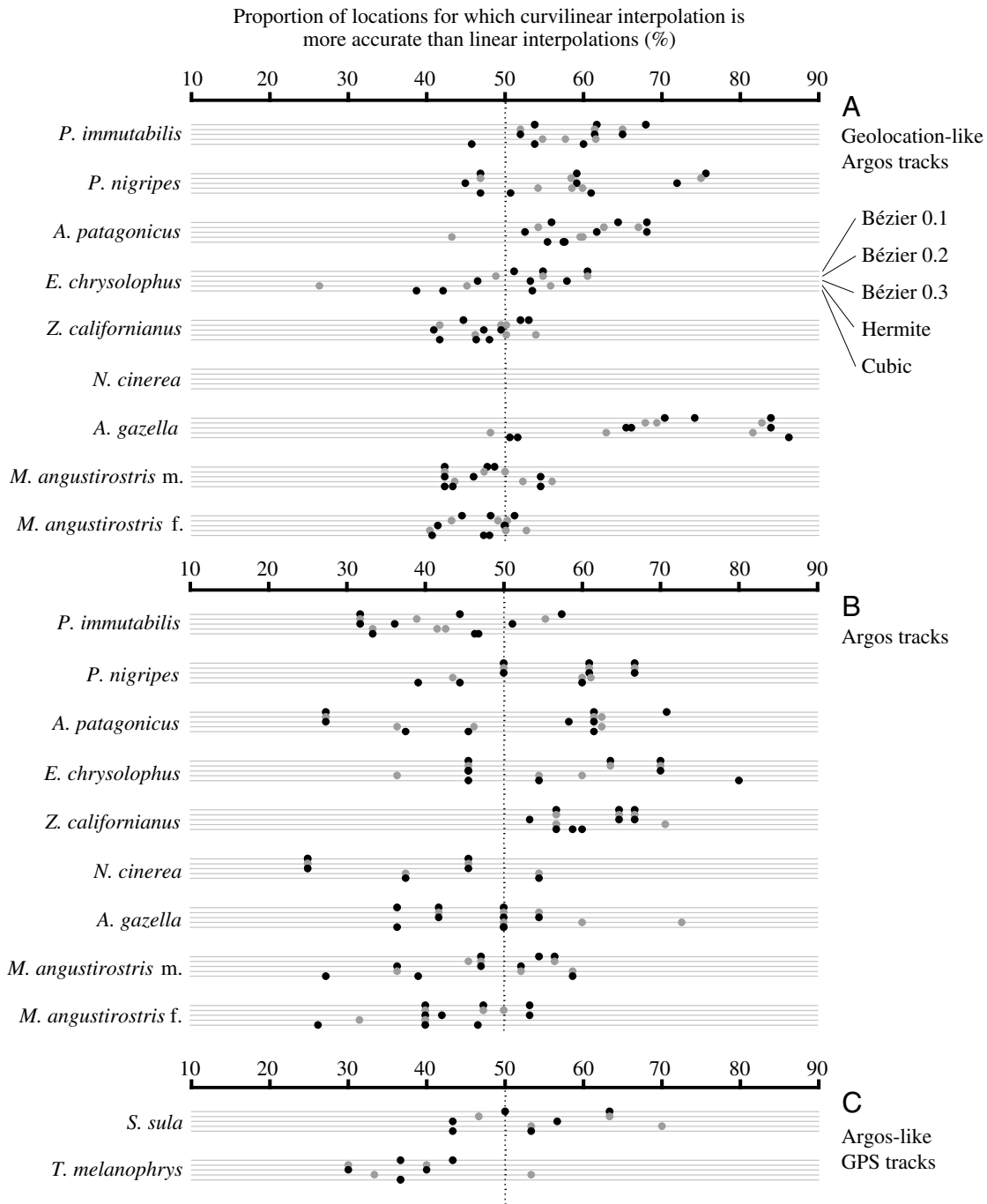


Fig. 3. Occurrence of more accurate locations when using curvilinear algorithm compared with a linear interpolation method for each species, each curvilinear algorithm and each track (i.e. each individual). Dots (representing individuals/tracks) are alternately shown in black and grey for clarity. Dots to the right of the 50% line represent tracks in which the curvilinear method yields more accurate locations than the linear interpolation method.

Table 5. Interpolation algorithms that maximize the probability of obtaining higher occurrence of more accurate interpolated locations (closer to the reference position)

Predator type	Type of tracks	
	Geolocation	Argos
Albatrosses	Hermite	Bézier 0.1/0.2/ 0.3
Penguins	Bézier 0.1	Bézier 0.1/ 0.2
Otariids	Bézier 0.1	Hermite
Phocids (linear tracks)	Hermite	Bézier 0.1

When several algorithms are given, we advise using the most relaxed algorithm (shown in bold) in order to optimize the estimation of track length. The numbers following Bézier are the values for μ (see Materials and methods for details).

Bézier curves varied depending on parameter μ , being straighter with small values and more relaxed (more curvilinear) for higher values. Bézier curves with $\mu=0.2$ were relatively similar to the hermite curves, except for the more linear parts of the tracks for which hermite curves were straighter (data not shown). In Bézier curves, we used 0.1, 0.2 and 0.3 for μ , because some preliminary tests showed that high values (above 0.5, and particularly over 1) tended to produce very sinuous paths, sometimes with loops that were non-existent in the original track data. For μ values below 0.5, the track shape always conformed to original track data (data not shown).

Discussion

Use of several algorithms to interpolate tracks of various marine vertebrates allowed us to propose alternative and more realistic ways of interpolating tracking data in fluid media. Several curvilinear algorithms resulted in interpolated tracks that had a greater number of more accurate locations, which produced a better estimate of track lengths, and still led to conservative analyses of tracks.

Accuracy of interpolated locations and factors affecting it

Accuracy of interpolated locations was always within the accuracy of the tracking method used. Geolocation tracking technique typically provides one to two locations per day, with an accuracy of 100–400 km (Phillips et al., 2004; Shaffer et al., 2005; Teo et al., 2004). Accuracy of Argos data is between ~0.8 km and 50 km (Fernández et al., 2001; Le Boeuf et al., 2000). The errors of interpolated locations in geolocation-like tracks or Argos tracks fell within or below those respective ranges in all of our trials (Table 4). Additionally, errors of interpolated locations were always smaller than the distance that the animals were potentially able to travel during the average time elapsed between recorded locations (Tables 2, 4). For example, in 24 h (temporal resolution of geolocation-Argos tracks), albatrosses were able to travel 561–828 km (Table 2), yet median errors were only 60 km (Table 4). It is noteworthy that, in Argos tracks, we calculated errors using

tracks of diminished quality (some locations were removed), so the true error may have been even smaller, and our estimates may be higher than actual range of errors.

The interpolation errors showed significant differences between species and, to a lesser extent, between individuals within species (Table 3). In particular, albatrosses have typically larger errors than all other species, either considering geolocation-like or Argos-like tracks. Compared with elephant seals, albatross tracks were of similar length, had better overall spatial accuracy, were of greater temporal resolution (Table 2) and yet had larger errors in the interpolated locations. Therefore, neither spatial scale/accuracy nor temporal resolution of the tracks can be a factor explaining the greater errors in interpolated locations of albatrosses. Between two recorded locations, albatrosses are able to fly larger distances because of their higher traveling speeds (Table 2). Traveling speed is therefore logically a crucial factor affecting errors of interpolated locations. Compared with other fast-flying seabirds [*S. sula* and *T. melanophrys* (tracks obtained from GPS, errors between 0.8 and 2.2 km; Table 4)], interpolated distances in tracks of both *P. immutabilis* and *P. nigripes* were still larger (tracks obtained with Argos, errors between 10.6 and 12.8 km; Table 4). The differences in accuracies between the two tracking techniques (several kilometers) were therefore most likely to explain the majority of errors between these species. Obviously, the overall shape of a track is also important in affecting the extent of the errors in interpolated locations. For example, interpolation of tracks of both male and female elephant seals was not dramatically enhanced using curvilinear interpolation. As a matter of fact, these tracks were extremely linear for long periods of time (see Fig. 4A for female tracks). In cases of more rounded tracks for phocids, it is likely that the hermite curve performs better than the Bézier curve with $\mu=0.1$, as indicated in Table 5 for Argos tracks of otariids. It is also likely that a Bézier curve with $\mu=0.08$ (the smaller μ , the straighter the track) would work better with the elephant seal data we used in the present study.

The factors affecting interpolation errors are multiple and interrelated, including (non-exhaustively) traveling speed of the animal, spatial accuracy of the locations, temporal resolution and shape of the track. It is important to note that spatial scale of tracks could potentially have a great impact on the interpolation errors, especially if scale of movements approaches the tracking method's spatial accuracy. In our case, all tracks were, by far, larger (Table 2) than the estimated accuracy of 'several kilometers' as described earlier. Also, animals may behave differently at different spatial scales, exhibiting, for example, more curvilinear movements at small spatial scales (e.g. when searching for food in a patch) and more straight movements at large spatial scales (e.g. when migrating or changing foraging zone). The effects of these factors are difficult to separate, and they probably differ between species and between individuals within species. We suggest that these factors most likely explain observed

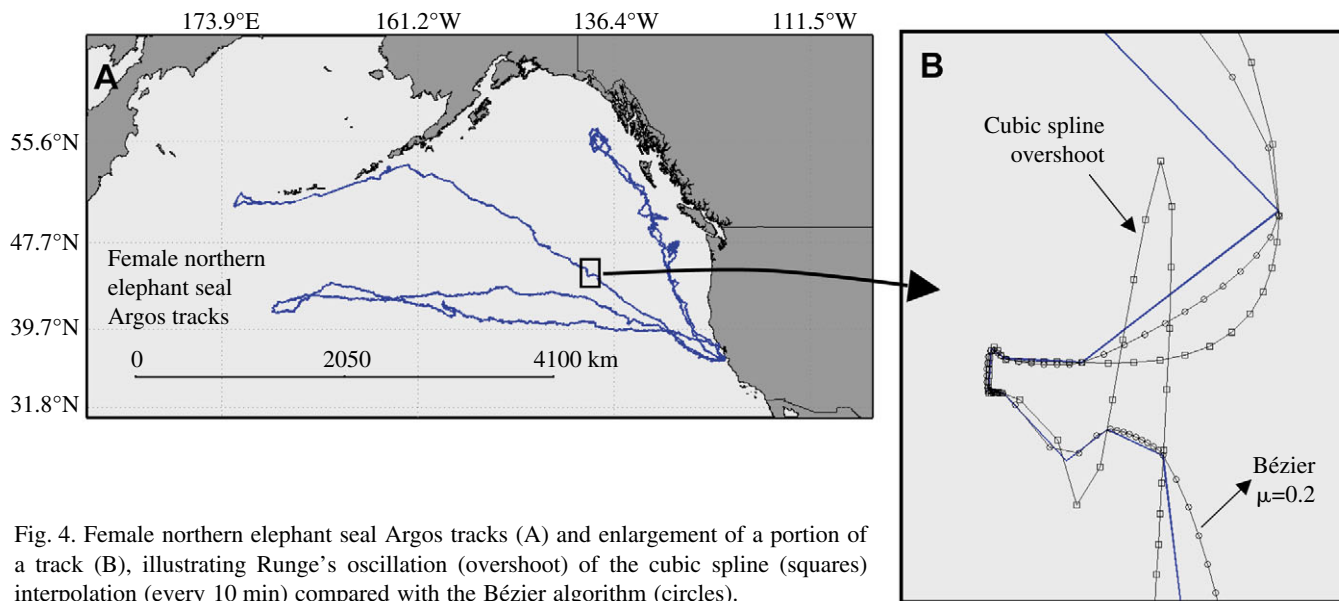


Fig. 4. Female northern elephant seal Argos tracks (A) and enlargement of a portion of a track (B), illustrating Runge's oscillation (overshoot) of the cubic spline (squares) interpolation (every 10 min) compared with the Bézier algorithm (circles).

differences in interpolation errors, both between and within species.

Curvilinear versus linear algorithms

We show that the choice of a curvilinear algorithm that produces less accurate locations would not dramatically affect the data because differences between algorithms were within the precision of the tracking method. Thus, the choice of an interpolation algorithm is not a major obstacle to the use of curvilinear algorithms for interpolating data.

Use of specific curves to interpolate tracks of marine vertebrates can, however, improve the probability of obtaining more accurate locations, depending on the species tracked. Algorithms shown in Table 5 are provided as guidelines for other researchers to use when selecting algorithms to analyze tracking data. We obtained a higher number of more accurate locations with some specific algorithms, but this was not reflected in the median distances of errors. This suggests that, even if they were more accurate in occurrence, distances were still very close to each other, and the effect of more accurate locations was possibly compensated by other locations that were of poorer accuracy. The consistent improvement of interpolated locations in tracks using curvilinear *vs* linear algorithms indicates that curves correspond more closely to the way marine vertebrates actually move. Further, it is interesting to note that the gain in using curvilinear interpolation was more obvious in geolocation-like tracks than in Argos tracks (Fig. 3). This is logical, because large-scale curvilinear movements were less visible in a geolocation-like track than in an Argos track (Fig. 2). In the same way, a highly accurate GPS track sampled every 10 s clearly shows curvilinear movements even with linear interpolation. Consequently, the lower the temporal resolution of a track, the higher is the gain in using curvilinear algorithms to interpolate the data.

Changes in the track length, when using a curvilinear

interpolation method compared with a linear interpolation method, indicated that track lengths were almost always considerably underestimated and that curvilinear interpolation algorithms more closely approximated actual track lengths. The linear interpolation method always resulted in the absolute minimum distance that an animal transited along the track. Similarly, track length estimated with curvilinear interpolation also underrepresented the distance an animal transited. Consequently, there was no risk of overestimating track length using curvilinear interpolation methods. In our study, even the most relaxed algorithms underestimated track lengths by at least 15%. It is important to note that interpolated geolocation-like Argos tracks were underestimated by a similar value, regardless of species. This contrasted with estimates of interpolated Argos-like GPS track lengths, which were very different between red-footed boobies and black-browed albatrosses. This indicates that the underestimation of track length is mostly due to fine-scale movements that cannot be recorded using either the geolocation or the Argos tracking technique.

Our results show that geolocation tracks are 15–20% shorter than the length of tracks measured using the Argos tracking technique. By contrast, it is harder to make such a generalization with Argos tracks. However, we show that Argos track lengths can be 40–70% shorter than actual track lengths (obtained using GPS). This probably depends on the activity of the animals at small spatial/temporal scales, *i.e.* below the resolution of the Argos tracking technique. Track lengths of migrating animals that engage in straighter movements should be estimated fairly well, but track lengths of foraging animals that exhibit small-scale convoluted movements should be poorly estimated.

An improvement of 0.3–4.2% in track length between a curvilinear and linear interpolation method seems like a small improvement. However, given the length of some tracks, those percentages can represent several hundreds of kilometers,

which may be substantial in terms of calculations for animal energetics and behavior. The fact that track distances were more accurate implies that estimated traveling speeds between fixes would also be more accurate when using curves as opposed to straight lines.

Sinuosity of animal tracks is an important parameter because it is used as a descriptor of animal activity, especially to distinguish between transiting and foraging phases or identifying operational spatial scales (Fauchald and Tveraa, 2003; Nams, 1996; Weimerskirch et al., 2002). The temporal resolution of a track is therefore a crucial factor in determining sinuosity, because the angles and their frequency depend directly on it. Another improvement from curvilinear interpolation is that the frequency distribution of angles is changed in a way that takes into account the number of interpolated segments for a given angle. This is equivalent to accounting for the time needed to turn, which is not possible to do with linear interpolation because angles are never changed.

Remarks and conclusion

The cubic spline was the most relaxed curve we used. It was also the only algorithm that was non-conformal. Cubic spline interpolation, as we applied it, had overshoots and large oscillations (Runge's oscillation), resulting in the interpolation of track locations that were not always induced by the recorded track data (Fig. 4B). We therefore suggest that conformal curvilinear interpolation algorithms be used, meaning those that create turns tangential to each recorded location.

Although curvilinear interpolation algorithms are advantageous to use, they do not solve the problem of low sampling interval and/or low spatial accuracy of initial data. The question of sampling resolution is crucial in quantitative analysis of animal tracks (Turchin, 1998). The fact that interpolation methods produce locations equally spaced in time does not imply that they are accurate, especially when temporal resolution of initial data is low. The accuracy of interpolated data is ultimately a function of initial temporal and spatial resolutions. Consequently, interpolation should not be misused, as for example a way of correcting poor quality tracks. Also, if the time interval chosen for interpolating tracking data does not allow the animal to travel more than the spatial accuracy of the tracking technique used (given its traveling speed), it is obvious that interpolated data are over-sampled and cannot represent accurately fine-scale movements of the animal. This emphasizes, on the one hand, the relationships between spatial accuracy and sampling interval of a track and, on the other hand, spatial scale of movement and traveling speed of the animal. These relationships must be known and understood prior to use and interpretation of interpolated data.

However, interpolation is important to apply to understand habitat use, as having location equally spaced in time is a way to account for time spent in a given zone.

Argos tracks are obtained with an estimated accuracy for each location. In this study, we did not take the accuracy of

each location into account in order to interpolate the tracks. Rather, we considered every non-filtered location as an actual position of the animal, and our interpolated tracks passed through each location. An alternative method of processing would take the accuracy of locations into account and would calculate a curve that does not necessarily pass through each location. In this case, the interpolation might pass by a certain distance, which would be proportional to the accuracy of the location. We believe this method could give satisfactory results; however, it has two major drawbacks. First, the resulting track is almost entirely made up, with almost no actual measured locations in the new track. Second, the resulting track length would be further underestimated than the previously shown.

In conclusion, we propose that curvilinear interpolation should be used instead of linear interpolation for animal tracking data obtained in fluid media, and this should be done only with conformal algorithms. Except for this particular restriction, curvilinear algorithms provide conservative analyses of track data, with no risk of considerably reducing track quality. Furthermore, curvilinear interpolation methods can ameliorate the track quality (see Table 5 for guidelines) and allow researchers to obtain tracks that are more likely to represent animal movement in a fluid medium.

The programming codes that we used to interpolate our tracking data are easy to implement and can be obtained directly from the corresponding author.

We thank all the people who helped collect data in the field, as well as Jesse Hiat and various mathematicians all over the world (particularly Paul Bourke and André Jaun) for inspirational conversations and web sites. We are also grateful to Pete Raimondi for statistical advice. This research was part of the Tagging of Pacific Pelagics (TOPP) program, funded in part by the National Ocean Partnership Program (N00014-02-1-1012) and the Office of Naval Research (N00014-00-1-0880 & N00014-03-1-0651).

References

- Alerstam, T., Gudmundsson, G. A. and Larsson, B.** (1993). Flight tracks and speed of Antarctic and Atlantic seabirds: radar and optical measurements. *Philos. Trans. R. Soc. Lond. B* **340**, 55-67.
- Angel, E.** (2003). *Interactive Computer Graphics: A Top-Down Approach Using OpenGL*. Boston: Addison Wesley Professional.
- Austin, D., McMillan, J. I. and Bowen, W. D.** (2003). A three-stage algorithm for filtering erroneous Argos satellite locations. *Mar. Mamm. Sci.* **19**, 371-383.
- Bartels, R. H., Beatty, J. C. and Barsky, B. A.** (1998). Bézier Curves. In *An Introduction to Splines for Use in Computer Graphics and Geometric Modelling* (ed. M. B. Morgan), pp. 211-245. San Francisco: Morgan Kaufmann.
- BirdLife International** (2004). *Tracking Ocean Wanderers: The Global Distribution of Albatrosses and Petrels. Results from the Global Procellariiform Tracking Workshop, 1-5 September, 2003, Gordon's Bay, South Africa*. Cambridge, UK: BirdLife International.
- Block, B. A., Teo, S. L. H., Walli, A., Boustany, A., Stokesbury, M. J. W., Farwell, C. J., Weng, K. C., Dewar, H. and Williams, T. D.** (2005). Electronic tagging and population structure of Atlantic bluefin tuna. *Nature* **434**, 1121-1127.
- de Boor, C.** (1978). *A Practical Guide to Splines*. Berlin: Springer-Verlag.

- Fauchald, P. and Tveraa, T.** (2003). Using first-passage time in the analysis of area-restricted search and habitat selection. *Ecology* **84**, 282-288.
- Fernández, P., Anderson, D. J., Sievert, P. R. and Huyvaert, K. P.** (2001). Foraging destinations of three low-latitude albatross (*Phoebastria*) species. *J. Zool.* **254**, 391-404.
- Ferraro, S., Georges, J. Y., Gaspar, P. and Le Maho, Y.** (2004). Where leatherback turtles meet fisheries. *Nature* **429**, 521-522.
- Folkow, L. P., Nordoy, E. S. and Blix, A. S.** (2004). Distribution and diving behaviour of harp seals (*Pagophilus groenlandicus*) from the Greenland Sea stock. *Polar Biol.* **27**, 281-298.
- Fratantoni, D. M. and Richardson, P. L.** (1999). SOFAR float observations of an intermediate-depth Eastern boundary current and mesoscale variability in the Eastern Tropical Atlantic Ocean. *J. Phys. Oceanogr.* **29**, 1265-1278.
- Fritsch, F. N. and Carlson, R. E.** (1980). Monotone piecewise cubic interpolation. *SIAM J. Numerical Anal.* **17**, 238-246.
- Hill, R. D.** (1994). Theory of geolocation by light levels. In *Elephant Seals. Population Ecology, Behavior, and Physiology* (ed. B. J. Le Boeuf and R. M. Laws), pp. 227-236. Berkeley: University of California Press.
- Hurlbert, S. H.** (1984). Pseudoreplication and the design of ecological field experiments. *Ecol. Monogr.* **54**, 187-211.
- Jouventin, P. and Weimerskirch, H.** (1990). Satellite tracking of wandering albatrosses. *Nature* **343**, 746-748.
- Kahaner, D., Moler, C. and Nash, S.** (1988). *Numerical Methods and Software*. Upper Saddle River, NJ: Prentice Hall.
- Kooyman, G. L., Ponganis, P. J., Castellini, M. A., Ponganis, E. P., Ponganis, K. V., Thorson, P. H., Eckert, S. A. and Le Maho, Y.** (1992). Heart rates and swim speeds of Emperor penguins diving under sea ice. *J. Exp. Zool.* **165**, 161-180.
- Le Boeuf, B. J., Crocker, D. E., Costa, D. P., Blackwell, S. B., Webb, P. M. and Houser, D. S.** (2000). Foraging ecology of northern elephant seals. *Ecol. Monogr.* **70**, 353-382.
- McConnell, B. J., Chambers, C. and Fedak, M. A.** (1992). Foraging ecology of southern elephant seals in relation to the bathymetry and productivity of the Southern Ocean. *Antarctic Sci.* **4**, 393-398.
- Mortenson, M. E.** (1997). *Geometric Modeling*. New York: John Wiley Press.
- Murray, M. D., Nicholls, D. G., Butcher, E. and Moors, P. J.** (2002). How wandering albatrosses use weather systems to fly long distances. 1. An analytical method and its application to flights in the Tasman Sea. *Emu* **102**, 377-385.
- Nams, V. O.** (1996). The VFractal: a new estimator for fractal dimension of animal movement paths. *Landscape Ecol.* **11**, 289-297.
- Phillips, R. A., Silk, J. R. D., Croxall, J. P., Afanasyev, V. and Briggs, D. R.** (2004). Accuracy of geolocation estimates for flying seabirds. *Mar. Ecol. Progr. Ser.* **266**, 265-272.
- Piegl, L.** (1993). *Fundamental Developments of Computer Aided Geometric Design*. San Diego: Academic Press.
- Polovina, J. J., Kobayashi, D. R., Parker, D. M., Seki, M. P. and Balazs, G. H.** (2000). Turtles on the edge: movement of loggerhead turtles (*Caretta caretta*) along oceanic fronts, spanning longline fishing grounds in the central North Pacific, 1997-1998. *Fish. Oceanogr.* **9**, 71-82.
- Pütz, K., Ingham, R. J. and Smith, J. G.** (2000). Satellite tracking of the winter migration of Magellanic penguins *Spheniscus magellanicus* breeding in the Falkland Islands. *Ibis* **142**, 614-622.
- Ream, R. R., Sterling, J. T. and Loughlin, T. R.** (2005). Oceanographic features related to northern fur seal migratory movements. *Deep-sea Res. II. Top. Stud. Oceanogr.* **52**, 823-843.
- Shaffer, S. A., Tremblay, Y., Awkerman, J. A., Henry, R. W., Teo, S. L. H., Anderson, D. A., Croll, D. A., Block, B. A. and Costa, D. P.** (2005). Comparison of light- and SST-based geolocation with satellite telemetry in free-ranging albatrosses. *Mar. Biol.* **147**, 833-843.
- Sibert, J. R., Musyl, M. K. and Brill, R. W.** (2003). Horizontal movements of bigeye tuna (*Thunnus obesus*) near Hawaii determined by Kalman filter analysis of archival tagging data. *Fish. Oceanogr.* **12**, 141-151.
- Teo, S. L. H., Boustany, A., Blackwell, S. B., Walli, A., Weng, K. C. and Block, B. A.** (2004). Validation of geolocation estimates based on light level and sea surface temperature from electronic tags. *Mar. Ecol. Progr. Ser.* **283**, 81-98.
- Turchin, P.** (1998). *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants*. Sunderland, MA: Sinauer Associates.
- Vogel, S.** (1994). *Life in Moving Fluids. The Physical Biology of Flow*. Princeton: Princeton University Press.
- Weimerskirch, H., Wilson, R. P. and Lys, P.** (1997). Activity pattern of foraging in the wandering albatross: a marine predator with two modes of prey searching. *Mar. Ecol. Progr. Ser.* **151**, 245-254.
- Weimerskirch, H., Guionnet, T., Martin, J., Shaffer, S. A. and Costa, D. P.** (2000). Fast and fuel-efficient? Optimal use of wind by flying albatrosses. *Proc. R. Soc. Lond. B* **267**, 1869-1874.
- Weimerskirch, H., Bonadonna, F., Bailleul, F., Mabile, G., Dell'Omo, G. and Lipp, H. P.** (2002). GPS tracking of foraging albatrosses. *Science* **295**, 1259.
- Wilson, R. P., Ducamp, J. J., Rees, W. G., Culik, B. M. and Nickamp, K.** (1992). Estimation of location: global coverage using light intensity. In *Wildlife Telemetry: Remote Monitoring and Tracking of Animals* (ed. I. G. Priede and S. M. Swift), pp. 131-134. New York: Ellis Horwood.
- Wilson, R. P., Pütz, K., Grémillet, D., Culik, B. M., Kierspel, M., Regel, J., Bost, C. A. and Lage, J.** (1995). Reliability of stomach temperature changes in determining feeding characteristics of seabirds. *J. Exp. Biol.* **198**, 1115-1135.