



# Improving data retention and home range estimates by data-driven screening

Takahiro Shimada<sup>1,\*</sup>, Rhondda Jones<sup>2</sup>, Colin Limpus<sup>3</sup>, Mark Hamann<sup>1</sup>

<sup>1</sup>School of Earth and Environmental Sciences, and <sup>2</sup>School of Marine and Tropical Biology, James Cook University, Townsville, Queensland 4811, Australia

<sup>3</sup>Queensland Department of Environment and Resource Management, Brisbane, Queensland, Australia

**ABSTRACT:** Fast acquisition GPS technologies such as Fastloc GPS have been commonly used in recent years to study fine scale spatio-temporal ecology of marine vertebrates. While Fastloc GPS gives more accurate locations than earlier methods, it remains important to identify and remove locations with high error because some location fixes are much less accurate than others. Increasing the number of source satellites required for a valid fix is a simple filter method but it comes at the cost of great data loss. Using data sets acquired from loggerhead turtles *Caretta caretta*, we explored an alternative filtering approach, based on speed between successive locations, angles created by 3 consecutive locations, manufacturer's quality index, and number of satellites used for location calculation. The performance of the proposed filter method was evaluated by conducting terrestrial, mobile tests. When our filter method was used, the linear error (mean  $\pm$  SD) of Fastloc GPS data decreased from  $2645.5 \pm 29458.2$  m ( $n = 1328$ ) to  $47.1 \pm 61.0$  m ( $n = 1246$ ), while retaining more than 94 % of data. Our filter method also led to more accurate home range estimates than the simple filter method. Improvements in data retention and home range estimates will give more reliable information for marine spatial planning and habitat protection.

**KEY WORDS:** Fastloc GPS · Screening · Accuracy · Data retention · Home range estimation · *Caretta caretta*

Resale or republication not permitted without written consent of the publisher

## INTRODUCTION

Home range analysis can provide fundamental insights into species ecology and has obvious implications for wildlife management (Peckham et al. 2007, Hays 2008, Hamann et al. 2010). Although home range investigations have been carried out for decades on a multitude of taxa, the quantification of home ranges remains challenging. Home range studies have difficulties everywhere, but are more advanced for terrestrial than for marine vertebrates because of the relative ease of animal capture and tracking, especially via GPS tags. Marine studies have progressed more slowly; they tend to be more difficult to execute, and tagging hardware has the additional problem of either transmitting signals

through water or relying on short exposure times as animals surface.

Since the 1980s, satellite-linked platform terminal transmitters (PTTs) have been used to study animal movement, and their use in the marine environment has increased rapidly (Stoneburner 1982, Timko & Kolz 1982, Godley et al. 2008). However, the use of PTTs to understand habitat use in marine animals is often limited because signals can only be transmitted when the animal surfaces, and the surface time must be long enough for transmissions to be completed. As a result, sample sizes for marine fauna tend to be lower than those for terrestrial fauna. Moreover, since the accuracy of an acquired location depends on the number of Argos satellites and the number of messages which reach them, accuracy may also be

\*Email: takahiro.shimada@my.jcu.edu.au

low (Goulet et al. 1999, Eckert & Stewart 2001, Godley et al. 2002). Hence, although PTT technology and methods for data analysis have improved over time (Jonsen et al. 2005, Patterson et al. 2010), PTTs are best applied to studies investigating larger scale habitat use, such as long-distance migration.

Standard GPS tags offer better accuracy but still require enough time on the surface to complete a fix, so relatively few attempts to acquire location are successful in the marine environment (Jay & Garner 2002). Some studies increased the frequency of successful location fixes by keeping the power of the tags on so that receivers acquire a GPS fix whenever they happen to surface, but battery life is then greatly reduced (Ryan et al. 2004, Elkaim et al. 2006). More recently, the situation has been improved by the use of fast acquisition GPS technologies such as Fastloc GPS and quick fix pseudorangeing (Tomkiewicz et al. 2010). Consequently, in recent years, these alternative GPS tags have been used on several marine species and age classes to investigate fine scale habitat use (Sheppard et al. 2006, Sims et al. 2009, Kuhn et al. 2010), and their results are thus useful for marine spatial planning and habitat protection (Sheppard 2008).

The standard GPS technology is dependent on the number of source satellites for the accuracy of an initial fix but can increase the accuracy of location estimates by correcting the locations from successive GPS fixes. In stationary land-based tests, the accuracy of Fastloc GPS also improved as the number of source satellites increased (Bryant 2007a, Hazel 2009). However, unlike standard GPS, Fastloc GPS records and uses only one set of raw data at each fixing attempt and thus there is no correction from subsequently obtained data. That is, while Fastloc GPS technology has achieved the ability to capture the necessary data in less than 0.1 s, its accuracy does not improve with successive location fixes (Bryant 2007b, Rutz & Hays 2009, Sirtrack 2010). As a consequence of the trade-off, some Fastloc GPS locations are much less accurate than others if fewer satellites have been involved in fix acquisition. The occurrence of location fixes with high error leads to several data analysis challenges with Fastloc GPS data, especially with the analysis of tracking data and estimation of home ranges. Hence, it is important to develop and use methods that can identify and account for locations with high error.

There are many published methods to screen location data when it is obtained with conventional Argos PTT tags (McConnell et al. 1992, Tremblay et al. 2006, Freitas et al. 2008) but fewer exist for data acquired

from Fastloc GPS tags, despite the increase in their use. The simplest screening method for Fastloc GPS data involves the use of residual errors. Residual error is a quality index that represents the level of accuracy of a location fix (Sirtrack 2010). This method alone may not remove all locations with high error but it can be supplemented by taking into account the number of satellites used for calculations—e.g. by rejecting locations involving fewer than a given number of satellites. However in some studies, this may result in a majority of the Fastloc GPS data being discarded (Lonergan et al. 2009). An approach which provides a more reliable filter while retaining more of the data is therefore highly desirable.

A screening method often used for PTT locations is the identification of behaviour that is biologically or ecologically unrealistic for the study species. This approach can also be used for Fastloc GPS data. Speed is the popular limiting factor, and recent studies use the maximum speeds that were estimated from the conventional tracking methods, such as acoustic or PTT telemetry, to screen GPS data (Schofield et al. 2007, Preston et al. 2010, Witt et al. 2010). However, because Fastloc GPS data are generally more abundant and more accurate than Argos PTT data, maximum speeds estimated from Fastloc GPS data are likely to be more realistic than those obtained from Argos PTT data. Further, the angle between 3 consecutive locations is also commonly used to filter location data, but the choice of angle is often arbitrary (Costa et al. 2010, Witt et al. 2010). Hence, there is a need for an improved approach to define the limiting speeds and angles for Fastloc GPS data.

Once a data set has been obtained and appropriately filtered, estimation of home range has a choice of possible approaches. Some commonly used methods for home range estimation are minimum convex polygon, kernel density estimation, harmonic mean, linear home range, and grid cell count (Laver & Kelly 2008). Among those methods, kernel density estimator is currently the most frequently used and least biased home range estimator (Kernohan et al. 2001). Although the kernel method measures intensity of use by estimating the probability density along both *x* and *y* coordinates (Silverman 1986, Worton 1989), and therefore the influence of locations with high error or over-screening is presumed to be decreased, accuracy and precision of kernel estimates will be affected by the accuracy of location data and sample size.

Given the promising application of GPS satellite telemetry in wildlife studies, it is important to use appropriate filtering methods to increase data accuracy while minimizing unnecessary data loss, and to

allow accurate quantification of home range estimates. This paper explores filtering approaches for Fastloc GPS data obtained from loggerhead turtles *Caretta caretta*, and by conducting terrestrial tests, we evaluate the performance of the proposed filtering methods and determine which filter leads to the most accurate home range estimates.

## MATERIALS AND METHODS

### Turtle tracking

In 2009 and 2010, 4 adult female loggerhead turtles (T93038, T81920, T54430, K22217) were captured while nesting at Mon Repos and 4 adult male loggerhead turtles (K24365, T53800, T74361, QA14215) were captured at the Moreton Bay foraging ground by a rodeo method (Limpus 1978). The sex and maturity of male turtles were identified either using laparoscopy or by determining whether the tail length from the carapace was greater than 19.0 cm (Limpus & Limpus 2003). We mounted a Fastloc GPS unit on the first through third vertebral scutes of the carapace using Sika Anchor fix-3+ epoxy glue with fibreglass for extra strength. The tracking period ranged from 14 to 153 d and the data include the location fixes that were obtained during the foraging, migrating and inter-nesting periods. T54430 was not tracked during the foraging period due to device failure. The location data acquired during unit deployment, nesting and post-release activities were excluded from the data analysis.

### Data retrieval and conversion

Data were transmitted via the ARGOS network (K24365, QA14215, T53800, T54430, T81920 and T93038) or via a USB link directly from device to computer (K22217, T74361 and terrestrial tests). We downloaded the data from each unit using manufacturer-supplied software (Sirtrack Fastloc software) and then decoded data into GPS locations. Finally the location data were converted to the Universal Transverse Mercator coordinate system.

### Data screening

The first filtering method we used involved the use of the manufacturer's quality index (residual error) and the number of satellites used in each location

calculation (hereafter referred to as the generic filter in this paper). Following the manufacturer's instructions, we excluded locations from the analysis when residual error values were greater than 30 or fewer than 4 satellites were used for location calculation (Sirtrack 2010). Data sets after additional filtering according to available satellite number will be referred to as generic># where # represents the number of source satellites (e.g. generic>4). Consequently, the generic>3 data set contains the maximum number of locations and generic>7 limits the data set to locations acquired with 8 satellites (the maximum possible with Fastloc GPS).

The second filtering method, a data-driven filter, extends the generic filter. The data-driven filter screens the data according to the speed between successive locations, and the angle created by 3 consecutive locations, as well as the number of satellites used for location calculation. Speed and turning angles may vary with turtle behaviour (e.g. foraging, migrating, escaping from predators), but there are currently few studies of behavioural effects on swimming speeds and turning patterns of loggerhead turtles. Therefore, we derived the limiting speeds and angles as objectively as possible from the patterns of Fastloc GPS locations obtained from the turtles, informed by the accuracy associated with the number of source satellites. As errors may increase substantially when the number of source satellites is limited to 4 (Bryant 2007a, Hazel 2009), the data-driven filter is designed to remove erroneous locations derived from 4 satellites. The data group that has been screened by a data-driven filter will be referred to as data-driven>3.

To determine the limiting speed for the data-driven filter, and because high error may occur when only 4 satellites were used to estimate locations, we used data acquired from more than 4 satellites (i.e. generic>4) to estimate the maximum linear speed ( $V_{\max}$ ) a loggerhead turtle was observed to swim between 2 consecutive locations. We then extracted the locations that were preceded and followed by unrealistically fast speeds (i.e.  $>V_{\max}$ ) and labelled them over-speed error points (OSEPs). While this identified some erroneous locations (i.e. OSEPs), other obvious errors (e.g. locations on land) remained unscreened (e.g. Fig. 1B). Therefore, we developed additional limiting criteria to identify and screen other locations with high error.

We calculated the angles between the bearings of lines joining successive location points. This represents  $180^\circ$  minus the animal's turning angle; we will call this the inner angle in this paper. We then identi-

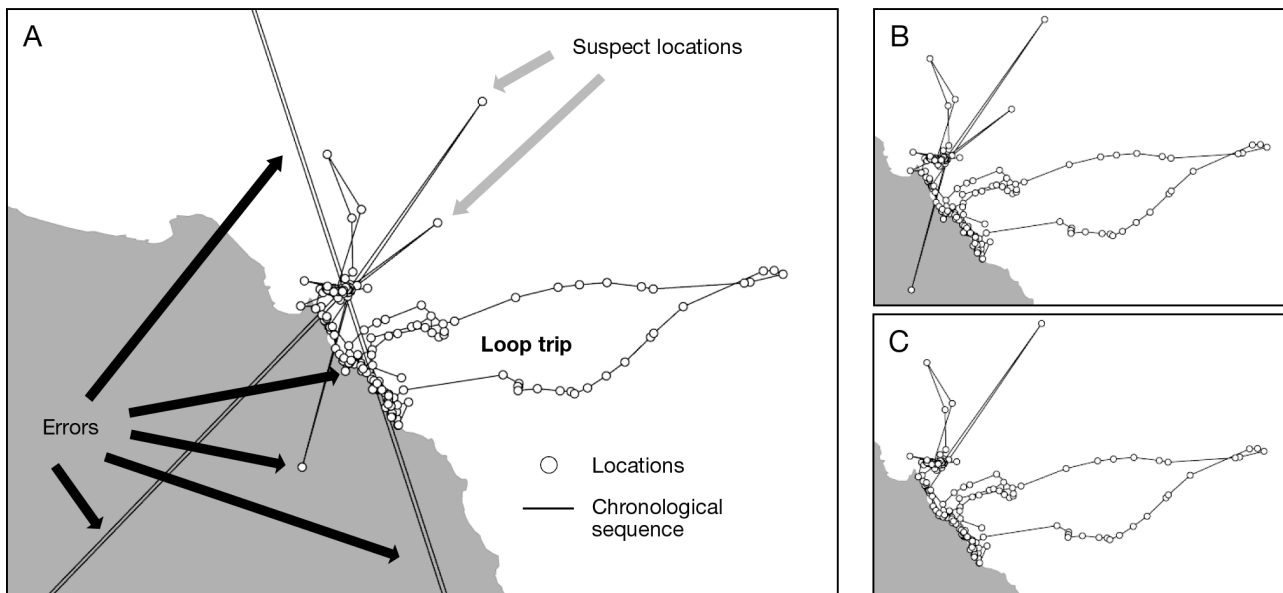


Fig. 1. An example of errors, suspect locations and a loop trip in a data set (A) after screening using the generic >3 filter (i.e. requiring >3 source satellites), (B) after removing locations that were preceded and followed by unrealistically fast speeds, and (C) after screening using the data-driven >3 filter (see 'Materials and methods: Data screening')

fied the inner angles that were associated with OSEPs for comparison with the angles associated with the remainder of the locations (non-OSEPs). Differences between the ranges of angles observed for OSEPs vs. non-OSEPs then guided selection of a limiting angle as described in 'Results: Data-driven filter'. The limiting angle was used to provide a preliminary identification of locations that may be erroneous (suspect locations).

Once suspect locations were extracted by the limiting angle, we calculated the linear speed for the turtle between the suspect location and both its previous location and its following location. The speeds were then compared to the speed of a 'loop trip' for an additional check of the legitimacy of the suspect locations. Loop trip behaviour is represented by spatial departure and return with more than 3 consecutive locations (Fig. 1A). Therefore, a loop trip was considered as an outlier if only one location was obtained during the movement. In our study, we considered continuous straight and fast movement to represent the departure to, and return from, a loop trip. Using the generic >3 data set, we calculated the net (i.e. straight-line) distance between the departure and turning location as well as the turning and return location of the loop trip, and from that calculated the net speed in and out ( $V_{lp}$ ). We considered the fastest  $V_{lp}$  observed in this study as the maximum inward or outward speed of loggerhead turtles during a loop trip. The suspect locations were removed

when the speed between the suspect location and either its previous or subsequent location exceeded the maximum  $V_{lp}$ .

### Kernel estimation

We used a Gaussian kernel density estimator for home range estimation. We chose fixed over adaptive kernel estimators with least squares cross validation (LSCV) as an automated bandwidth selector, because adaptive kernel was not available in the software we used, and because the fixed kernel methods with LSCV to select smoothing parameters appear to produce more accurate and precise estimates of home range areas and are less sensitive to autocorrelation within data sets than adaptive kernel methods (Seaman & Powell 1996, Seaman et al. 1999). While several studies have demonstrated that LSCV can be highly variable (Park & Marron 1990, Gitzen & Millsbaugh 2003) and fails to select appropriate bandwidth values when a sample size is too small (e.g. <30 locations; Seaman et al. 1999) or too large (e.g. >100 locations; Hemson et al. 2005), there have been attempts to improve the LSCV method by choosing more appropriate bandwidth values (Duong 2007, 2011). We used the improved LSCV bandwidth selector that is implemented in the R package 'ks' (Duong 2011, R Development Core Team 2011). The 'ks' package was also used to esti-

mate continuous utilization distribution. We used Geospatial Modelling Environment, an extension to ArcGIS, to generate volume contours from the utilization distribution (R Development Core Team 2011, Beyer 2012). A resolution of 50 m was used for the kernel grid because the mean linear error for properly filtered Fastloc GPS data was consistently within 50 m in our terrestrial tests (see Table 1).

### Terrestrial mobile test

We recorded walking and biking movements using both a handheld standard GPS receiver and a Fastloc GPS receiver in Townsville, Australia (terrestrial mobile test). The tests were conducted in open areas without overhead cover so as to mirror the environment of the turtles at sea. The tests were commenced after the standard GPS receiver (Garmin GPS60) gained an accuracy of <5 m according to the indication integrated in the unit. Once the standard GPS receiver attains the location estimates with high accuracy, the expected error of the successive locations will remain <15 m, according to the manufacturer (Garmin 2006). The Fastloc GPS unit was set to record a position every 4 min while the standard GPS receiver recorded a position every second. The clocks of the Fastloc GPS receiver and the standard GPS receiver were synchronized before the test began. Following the last test, the time difference between 2 units was recorded. Consequently, using the following equation, the time of each Fastloc GPS fix was adjusted according to the clock difference with the standard GPS so that the time of each standard GPS fix corresponded to the time of each Fastloc GPS fix:

$$\text{Adjusted } Ft_i = Ft_i \frac{\Delta t(Ft_i - t_{\text{sync}})}{t_{\text{end}} - t_{\text{sync}}} \quad (1)$$

where  $Ft_i$  is the time when the  $i$ th Fastloc GPS fix was acquired,  $\Delta t$  is the final clock difference between the 2 units,  $t_{\text{sync}}$  is the time at synchronization and  $t_{\text{end}}$  is the time on the Garmin unit when  $\Delta t$  was recorded. The range of  $\Delta t$  was 1.6 to 4.6 s.

### Evaluation of filter performance

We used the terrestrial data to evaluate how data screening improves the accuracy of Fastloc GPS data sets and to determine which screening method leads to the most accurate home range estimates. Unlike for standard GPS, the accuracy of Fastloc GPS locations is

not improved by successive fixes, therefore the differences in the interval of location fixing between the turtle tracking and the terrestrial test will not affect the accuracy of Fastloc GPS fixes. Thus, the accuracy of Fastloc GPS locations in the terrestrial test should be comparable to their accuracy in the turtle data.

We screened the data acquired from the terrestrial mobile tests using both the generic and the data-driven filters, following the same procedures used for the turtle data (but with a variation of the loop criterion): the maximum  $V_{\text{lp}}$  of the terrestrial mobile test was estimated by multiplying the terrestrial  $V_{\text{max}}$  by the ratio of maximum  $V_{\text{lp}}/V_{\text{max}}$  from the turtle data. We then calculated the linear distances between the standard GPS locations and simultaneous Fastloc-derived GPS locations; we interpreted these values as the error distances of Fastloc GPS locations from true locations. After normalizing the data by log-transformation, we compared the means of the linear errors of the generic>3 and data-driven>3 data sets using Welch's 2-sample  $t$ -test to assess the performance of the data-driven filter.

We subsampled the Fastloc GPS locations of the terrestrial data using each of the screening methods (i.e. generic>3, data-driven>3, generic>4, generic>5, generic>6 and generic>7). We then calculated the kernel estimates from each of the subsampled Fastloc GPS data and the standard GPS data. We considered the kernel estimates derived from the standard GPS locations to represent the most accurate kernel home range estimates. To measure the differences in the kernel estimates derived from the standard GPS and each of the filtered Fastloc GPS data, we calculated the mean integrated squared error (MISE) as:

$$\text{MISE} = \frac{1}{n} \sum_{i=1}^n [f_{\text{Fastloc}}(x_i, y_i) - f_{\text{standard}}(x_i, y_i)]^2 \quad (2)$$

where  $n$  is the number of grid points,  $x$  and  $y$  are the grid coordinates,  $f_{\text{Fastloc}}(x_i, y_i)$  is the estimated density derived from Fastloc GPS data at the  $i$ th grid point, and  $f_{\text{standard}}(x_i, y_i)$  is the accurate density derived from standard GPS data. MISE is widely used as a global measure of differences among different kernel estimates (Seaman & Powell 1996, Horne & Garton 2006, Fieberg 2007b). As a smaller MISE indicates more similarity between estimates, we used MISE as an inverse index of accuracy for kernel estimates.

### Home ranges of turtles

We subsampled each turtle data set using each of the alternative screening methods (i.e. generic>3,

data-driven>3, generic>4, generic>5, generic>6 and generic>7) and computed home range estimates from the subsampled data sets. We aimed to examine the effect of data screening on the estimates of home range areas of the loggerhead turtles. Since the true home ranges of these animals are not known, the accuracy of the estimates could not be evaluated for the turtle data.

## RESULTS

### Filter performance in terrestrial tests

The duration of the terrestrial mobile tests ranged from 44 to 136 min and a total of 1583 Fastloc GPS locations were recorded. However, we excluded 255 of them from analysis because the corresponding standard GPS locations were not acquired. The  $V_{\max}$  derived from the terrestrial mobile test was 11.9 km h<sup>-1</sup> (n = 1017). The maximum  $V_{ip}$  of the terrestrial data was then calculated as 2.4 km h<sup>-1</sup> using the ratio of maximum  $V_{ip}$  to  $V_{\max}$  from the turtle data described in the following subsection (1.8 km h<sup>-1</sup>:8.9 km h<sup>-1</sup>). The linear distance between standard GPS and Fastloc GPS locations ( $\Delta_{loc}$ ) decreased as the number of source satellites increased or the data-driven filter was applied (Table 1). In particular, when locations derived from >3 satellites (i.e. generic>3) were used, the linear error was decreased significantly by using the data-driven filter ( $t_{(2)} 2239.6 = 2.43$ ,  $p < 0.05$ ; Table 1). Increasing the number of source satellites required for a valid fix offered further reductions in the linear error but it came at the cost of much greater data loss (Table 1).

We also investigated how data screening influences the accuracy of the home range estimates in the terrestrial test. The kernel estimates derived from data-driven>3 data had the lowest MISE, followed by those derived from generic>4. The remaining

Table 1. Linear differences between standard GPS locations and screened Fastloc GPS locations ( $\Delta_{loc}$ , in m). N: no. of locations; numbers in brackets are the percentage of each N when compared to that of the generic>3 data set

No. satellites	Filter	Mean $\Delta_{loc}$	SD	Max. $\Delta_{loc}$	N (%)
>3	Generic	2645.5	29458.2	484640.1	1328 (100)
>3	Data-driven	47.1	61.0	699.5	1246 (94)
>4	Generic	33.1	35.4	328.7	866 (65)
>5	Generic	24.8	22.5	319.5	491 (37)
>6	Generic	19.6	10.8	64.5	228 (17)
>7	Generic	18.8	9.2	40.1	79 (6)

generic filters led to higher MISE in the ascending order of generic>5, generic>6, generic>7, and the kernel estimates derived from generic>3 had the highest MISE (Fig. 2).

### Data-driven filter

We estimated the  $V_{\max}$  of the loggerhead turtles as 8.9 km h<sup>-1</sup> using the location data derived from >4 satellites (n = 3921 fixes). A total of 23 OSEPs was then identified in the data from 8 turtles (n = 6154 fixes) for which the speed both from a previous and to a subsequent location exceeded 8.9 km h<sup>-1</sup>. We found that all the inner angles at OSEPs were <59° while the inner angles at non-OSEPs were generally higher, ranging up to 180°. The proportion of inner angles ≤59° was significantly different between OSEPs and non-OSEPs (binomial test,  $p < 0.0001$ ). However, using 59° as a limiting angle may be too conservative as a filter since OSEPs are the locations with the most extreme error. Moreover, a previous study using Argos satellite telemetry found that most errors were associated with acute inner angles (i.e. <90°) (Keating 1994). Therefore, we used acute inner angles to make a preliminary identification of suspect locations. We then compared the linear speed immediately prior to and following the suspect location to the maximum  $V_{ip}$  to check the legitimacy of each suspect location. The maximum  $V_{ip}$  was calculated as 1.8 km h<sup>-1</sup> from 57 loop trips; each loop trip has 2  $V_{ip}$  (mean ± SD: 0.4 ± 0.3, n = 114).

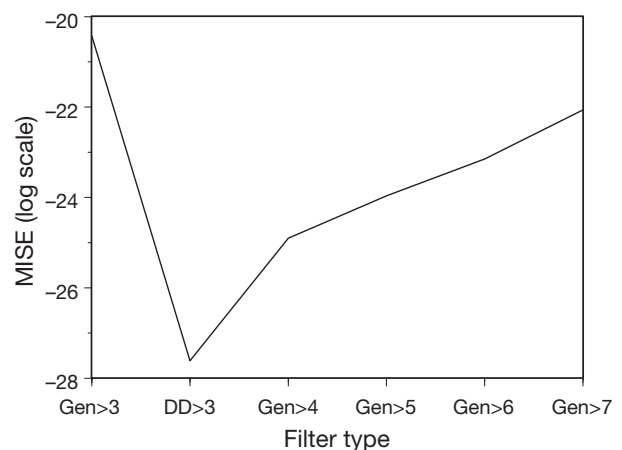


Fig. 2. Mean integrated squared error (MISE) of kernel estimates derived from the Fastloc GPS locations in the terrestrial mobile test. The Fastloc GPS data were screened by the data-driven filter (DD>3) and the generic filters that limited the use of locations by requiring the number of source satellites to be more than 3 (Gen>3), 4 (Gen>4), 5 (Gen>5), 6 (Gen>6), and 7 (Gen>7)

This resulted in a final filtering rule as follows: locations were removed if

- the speed both from a previous and to a subsequent location exceeded  $8.9 \text{ km h}^{-1}$ ,
- or if all of the following criteria applied:
- the number of source satellites was limited to 4,
  - the inner angle was acute and
  - the speed either from a previous or to a subsequent location exceeded  $1.8 \text{ km h}^{-1}$ .

### Effect of screening on turtle data

The data-driven filter successfully removed all obvious locations with high error when applied to the generic>3 data sets (e.g. Figs. 1C & 3) while keeping more than 95% of the locations (Table 2). All the locations with high error identified by the data-driven filter were also removed when the generic filters limited the use of locations to those made using >4 satellites (i.e. generic>4) because the obvious errors were all derived from 4 satellites. However, moving from generic>3 to generic>4 (i.e. requiring at least 5 satellites for a valid fix) removed more than a quarter of the location data (Table 2).

The shape and sizes of the home range areas of each turtle varied subject to varying proportions of high-error locations and sample size resulting from the use of different filters (e.g. Fig. 3). The home range boundaries estimated from the generic>3 data were expanded greatly due to the presence of outlying fixes. The data of T53800 did not retain enough locations for home range estimation when it was screened by the generic>7 filter.

## DISCUSSION

The results of our terrestrial test and turtle tracking showed that only a small proportion of Fastloc GPS data was highly erroneous but this led to inaccurate quantification of home range estimates (Figs. 2 & 3). We found that a screening method for Fastloc GPS data can be developed by identifying and checking suspect locations characterized by unlikely turning angles and speeds, informed by the location accuracy associated with the number of source GPS satellites.

Loggerhead turtle  $V_{\max}$  (the preliminary limiting factor used by the data-driven filter) is estimated as  $8.9 \text{ km h}^{-1}$  in our study. This is similar to the near-sprint linear speed of the same species measured during boat chases (Heithaus et al. 2002), but it is faster than the commonly used limiting speed of

$5 \text{ km h}^{-1}$ , which was estimated using PTT Argos data obtained from post-nesting green turtles *Chelonia mydas* (Luschi et al. 1998). The swimming speed assessed in our study using Fastloc GPS confirms that the linear speed of the loggerhead turtles is likely to be variable.

The variation in swimming speed is less critical in the filtering process for PTT data because the long interval between consecutive locations would have made short periods at high speed less obvious. Now that more accurate and shorter intervals between locations can be expected using Fastloc GPS, short periods of rapid movement may be mistakenly removed if conventional limiting speeds are used. Through using the faster limiting speed quantified in this study ( $V_{\max}$ :  $8.9 \text{ km h}^{-1}$ ), we found that the risk of over-screening was minimized, but many locations with high error were not identified because so few data exceeded  $V_{\max}$ . The locations identified by  $V_{\max}$  (i.e. OSEPs) can be regarded as the most extreme errors, and additional screening is needed. The combination of acute inner angles and a lower value for limiting speed (i.e. maximum  $V_{ip}$ ) provided that additional screen.

Our turtle tracking data indicate that OSEPs were associated with inner angles  $<59^\circ$ , suggesting that the  $20^\circ$  used by Witt et al. (2010) as a limiting angle may be too conservative when used as a filter for Fastloc GPS data, at least for loggerhead turtles. We used  $90^\circ$  as the limiting angle, which resulted in removing as much as 34% of the original location data in our data set if it was used as a standalone filter. Removing this proportion of data can lead to over-screening: 95% of Fastloc GPS locations had errors  $\leq 219 \text{ m}$  in the terrestrial mobile test (this study) or  $\leq 101 \text{ m}$  in a stationary test (Hazel 2009), which will be acceptable for many wildlife studies. We overcame the issue of over-screening by supplementing the use of the angle criterion with a maximum loop speed estimate and consequently managed to retain more than 95% of turtle tracking data. Note that the data-driven filter still has the potential to screen real but short sprinting out-and-back movements because animals are capable of short bursts of very fast speed which cannot be maintained for the long periods between most telemetry locations. Although we did not encounter this with our study, researchers should be aware of this possibility when using the data-driven approach.

The generic filters have an advantage in their simplicity, and our terrestrial study found a positive relationship between the number of source satellites and the accuracy of locations for mobile transmitters (Table 1), as found in stationary land-based tests

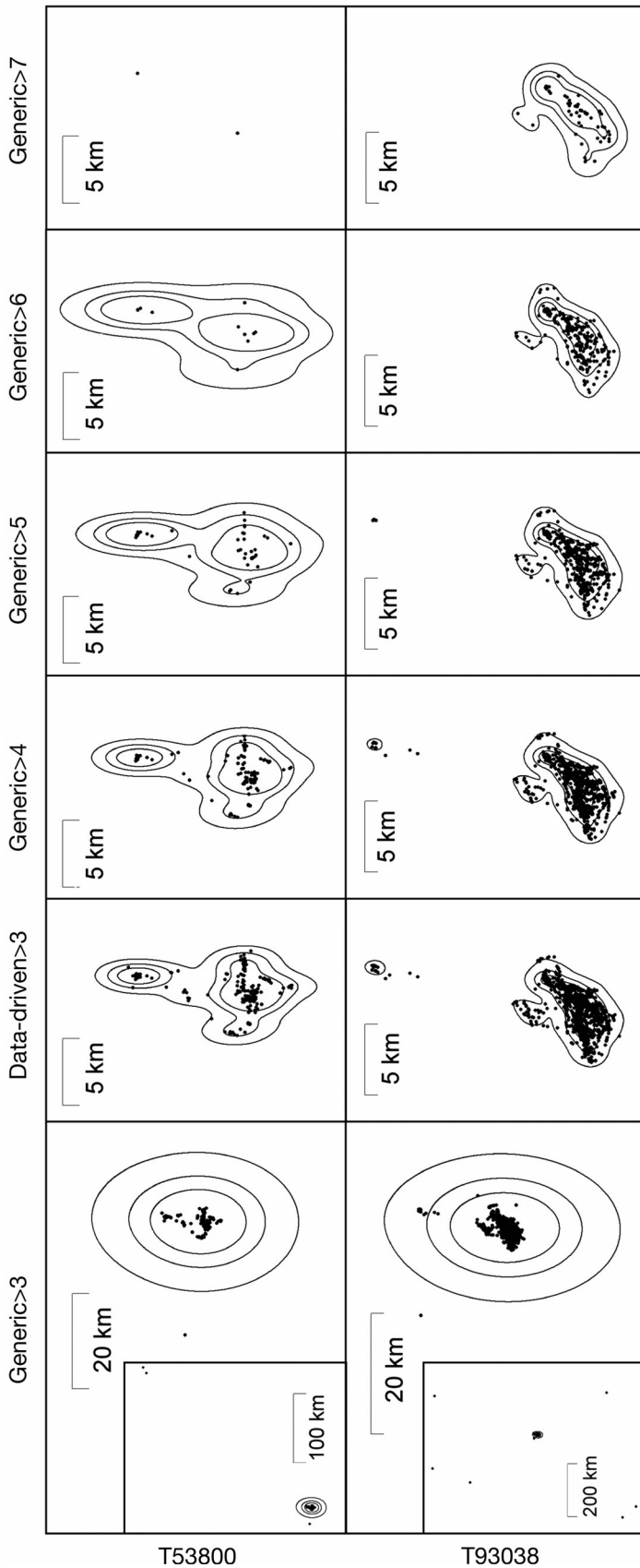


Fig. 3. *Caretta caretta*. Density contours of home range estimates for 2 turtles (T53800 and T93038) during their foraging period. Data were screened by the data-driven filter (data-driven>3) and the generic filters that limited the use of locations by requiring the number of source satellites to be more than 3 (generic>3), 4 (generic>4), 5 (generic>5), 6 (generic>6), and 7 (generic>7). Lines represent 95, 75 and 50% volume contours of the home range estimate. Dots are Fastloc GPS locations

(Bryant 2007a, Hazel 2009). Therefore, using only the location data derived from larger numbers of source satellites can be a straightforward screening method. However, we found that there is a risk of excluding smaller scale but important habitats by the generic filters, even when original sample size is large. For example, the sample size obtained from T93038 was the largest of the 8 turtles, but when the use of locations was limited by requiring the number of source satellites to be more than 6 (i.e. generic>6, generic>7), the northern-most aggregation disappeared because the patch consisted of locations derived from 4, 5 and 6 satellites (Fig. 3). Although the proportion of locations that generated the northern-most patch is small (1.5%), the habitat may be important to the animal during particular times of the year or with variation in availability of food or environmental preferences. In this case, if the use of locations had been limited to those derived from >6 satellites by the generic filters, the habitat distribution would have been underestimated and may not have picked up patches of peripheral habitat that are important to the animal. Moreover, in smaller data sets (e.g. T53800) over-screening may not retain enough locations for a home range analysis (Fig. 3).

For very large data sets (e.g. T93038), requiring locations to be derived from >4 satellites (i.e. generic>4) may be an adequate filtering method because locations with high error were associated with 4 satellites in our data set, and expected errors (mean ± SD) for generic>4 data sets were low 33.1 ± 35.4 m (Table 1). The drawback of the generic>4 filter is that it screens more locations than necessary: moving from generic>3 to generic>4 removed up to 53% of locations in our turtle data set (Table 2). A large reduction in sample size will result in decrease in precision of kernel estimates (Fieberg 2007a), loss of valuable ecological information such as detailed animal movements (Mills et al. 2006) and reduction in the power of statistical comparison (Whitlock & Schluter 2009).

For constructing home range boundaries, removing a large proportion of data by the generic>4 filter may not be so critical since the home range areas derived from the data-driven>3 and generic>4 data sets are very



Table 2. Remaining number of Fastloc GPS locations after each filter was applied to turtle location data sets (turtle IDs given above columns). Only the data acquired during the foraging period are shown. The numbers in the brackets are the percentage of each sample size (i.e. number of locations) when compared to that of the generic>3 data set from the same turtle

No. satellites	Filter	Number of locations (percentage against generic>3)						
		K22217	K24365	QA14215	T53800	T74361	T81920	T93038
>3	Generic	475 (100)	504 (100)	226 (100)	219 (100)	964 (100)	1058 (100)	1174 (100)
>3	Data-driven	469 (99)	500 (99)	223 (99)	215 (98)	951 (91)	1009 (96)	1123 (96)
>4	Generic	334 (70)	309 (62)	106 (47)	119 (54)	705 (73)	565 (53)	823 (70)
>5	Generic	242 (51)	167 (33)	40 (18)	45 (21)	459 (48)	299 (28)	490 (42)
>6	Generic	127 (27)	73 (15)	16 (7)	11 (5)	250 (26)	121 (11)	261 (22)
>7	Generic	46 (10)	16 (3)	3 (1)	2 (1)	92 (10)	38 (4)	64 (6)

similar (e.g. Fig. 3). Based on the result of our terrestrial tests, we considered kernel estimates derived from data-driven>3 data sets to represent better home range models than those derived from generic-filtered data. The similarity in the home range areas derived from data-driven>3 and generic>4 data sets indicates that the fixed kernel estimator with the improved LSCV as a bandwidth selector is robust against reduction in sample size as long as high-error locations are properly screened, at least in the volume contours of home ranges.

Although unnecessary data loss with the generic>4 filter did not appear to have a large influence on the construction of home range boundaries, there is little reason to use the generic>4 filter over the data-driven filter when the latter method is available because the data-driven filter (1) will lead to more accurate home range estimates than the generic filters, and (2) will satisfactorily remove locations with high error while retaining more data than the generic filters, as shown in our terrestrial test. Unlike the generic filters, the criteria used in the data-driven filter would need to be determined for each species, and may need to be evaluated separately for different habitats or populations. When the filtering criteria used for a data-driven filter cannot be determined, the generic>4 filter may be a useful alternative screening method.

**Acknowledgements.** This was a collaborative research project of the Queensland Department of Environment and Resource Management (DERM) and James Cook University (JCU) as part of a Queensland Turtle Conservation Project (QTCP). Financial and logistic support was provided by DERM and JCU. We thank the QTCP volunteers for their help in capturing and handling turtles, C. Marshall for the collection of some terrestrial mobile test data, J. Hazel for providing constructive comments on data analysis, and J. Moloney for his advice on GIS analysis. We are truly grateful to R. Powell and 2 anonymous reviewers for thorough reviews that greatly improved an earlier version of this paper. K. Lay of Sirtrack provided explanations of location fixing by Fastloc GPS tech-

nology. We also thank T. Duong, the author of R package 'ks', for explanations of the scripts and his expert knowledge of kernel smoothing. T.S. was supported by a NERP scholarship. This research was conducted under the ethics permits WISP 05967409 of DERM and A1372 of JCU.

#### LITERATURE CITED

- Beyer HL (2012) Geospatial Modelling Environment (Version 0.6.0.0). [www.spatialecology.com/gme](http://www.spatialecology.com/gme)
- Bryant E (2007a) 2D location accuracy statistics for Fastloc® cores running firmware versions 2.2 & 2.3. In: Technical Report TR01. Wildtrack Telemetry Systems, Leeds
- Bryant E (2007b) Fastloc™ technology. Accessed 10 Oct 2011. [www.wildtracker.com/fastloc.htm](http://www.wildtracker.com/fastloc.htm)
- Costa DP, Robinson PW, Arnould JPY, Harrison AL and others (2010) Accuracy of ARGOS locations of pinnipeds at-sea estimated using Fastloc GPS. *PLoS ONE* 5:e8677
- Duong T (2007) ks: kernel density estimation and kernel discriminant analysis for multivariate data in R. *J Stat Softw* 21:1–16
- Duong T (2011) ks: kernel smoothing. R package version 1.8.4. <http://CRAN.R-project.org/package=ks>
- Eckert SA, Stewart BS (2001) Telemetry and satellite tracking of whale sharks, *Rhincodon typus*, in the Sea of Cortez, Mexico, and the North Pacific Ocean. *Environ Biol Fishes* 60:299–308
- Elkaim GH, Decker EB, Oliver G, Wright B (2006) Go deep marine mammal marker for at-sea monitoring. *GPS World* 17:30–33
- Fieberg J (2007a) Kernel density estimators of home range: smoothing and the autocorrelation red herring. *Ecology* 88:1059–1066
- Fieberg J (2007b) Utilization distribution estimation using weighted kernel density estimators. *J Wildl Manag* 71:1669–1675
- Freitas C, Lydersen C, Fedak MA, Kovacs KM (2008) A simple new algorithm to filter marine mammal Argos locations. *Mar Mamm Sci* 24:315–325
- Garmin (2006) GPS 60 owner's manual. Appendix B: specifications. Olathe, KS
- Gitzen RA, Millsbaugh JJ (2003) Comparison of least-squares cross-validation bandwidth options for kernel home-range estimation. *Wildl Soc Bull* 31:823–831
- Godley BJ, Richardson S, Broderick AC, Coyne MS, Glen F, Hays GC (2002) Long-term satellite telemetry of the movements and habitat utilisation by green turtles in the Mediterranean. *Ecography* 25:352–362
- Godley BJ, Blumenthal JM, Broderick AC, Coyne MS, Godfrey MH, Hawkes LA, Witt MJ (2008) Satellite tracking of

- sea turtles: Where have we been and where do we go next? *Endang Species Res* 4:3–22
- Goulet AM, Hammill MO, Barrette C (1999) Quality of satellite telemetry locations of gray seals (*Halichoerus grypus*). *Mar Mamm Sci* 15:589–594
- Hamann M, Godfrey MH, Seminoff JA, Arthur K and others (2010) Global research priorities for sea turtles: informing management and conservation in the 21st century. *Endang Species Res* 11:245–269
- Hays GC (2008) Sea turtles: a review of some key recent discoveries and remaining questions. *J Exp Mar Biol Ecol* 356:1–7
- Hazel J (2009) Evaluation of fast-acquisition GPS in stationary tests and fine-scale tracking of green turtles. *J Exp Mar Biol Ecol* 374:58–68
- Heithaus MR, Frid A, Dill LM (2002) Shark-inflicted injury frequencies, escape ability, and habitat use of green and loggerhead turtles. *Mar Biol* 140:229–236
- Hemson G, Johnson P, South A, Kenward R, Ripley R, Macdonald D (2005) Are kernels the mustard? Data from global positioning system (GPS) collars suggests problems for kernel home-range analyses with least-squares cross-validation. *J Anim Ecol* 74:455–463
- Horne JS, Garton EO (2006) Likelihood cross-validation versus least squares cross-validation for choosing the smoothing parameter in kernel home-range analysis. *J Wildl Manag* 70:641–648
- Jay C, Garner G (2002) Performance of a satellite-linked GPS on Pacific walrus (*Odobenus rosmarus divergens*). *Polar Biol* 25:235–237
- Jonsen ID, Flemming JM, Myers RA (2005) Robust state-space modeling of animal movement data. *Ecology* 86:2874–2880
- Keating KA (1994) An alternative index of satellite telemetry location error. *J Wildl Manag* 58:414–421
- Kernohan BJ, Gitzen RA, Millsaugh JJ (2001) Analysis of animal space use and movements. In: Millsaugh JJ, Marzluff JM (eds) *Radiotracking and animal populations*. Academic Press, San Diego, CA, p 125–166
- Kuhn C, Tremblay Y, Ream R, Gelatt T (2010) Coupling GPS tracking with dive behavior to examine the relationship between foraging strategy and fine-scale movements of northern fur seals. *Endang Species Res* 12:125–139
- Laver PN, Kelly MJ (2008) A critical review of home range studies. *J Wildl Manag* 72:290–298
- Limpus CJ (1978) The reef. In: Lavery HJ (ed) *Exploration north: Australia's wildlife from desert to reef*. Richmond Hill Press, Richmond, p 187–222
- Limpus CJ, Limpus DJ (2003) Biology of the loggerhead turtle in western South Pacific Ocean foraging areas. In: Bolten AB, Witherington BE (eds) *Loggerhead sea turtles*. Smithsonian Institution, Washington, DC, p 93–113
- Loneragan M, Fedak M, McConnell B (2009) The effects of interpolation error and location quality on animal track reconstruction. *Mar Mamm Sci* 25:275–282
- Luschi P, Hays GC, Del Seppia C, Marsh R, Papi F (1998) The navigational feats of green sea turtles migrating from Ascension Island investigated by satellite telemetry. *Proc R Soc Lond B* 265:2279–2284
- McConnell BJ, Chambers C, Fedak MA (1992) Foraging ecology of southern elephant seals in relation to the bathymetry and productivity of the Southern Ocean. *Antarct Sci* 4:393–398
- Mills KJ, Patterson BR, Murray DL (2006) Effects of variable sampling frequencies on GPS transmitter efficiency and estimated wolf home range size and movement distance. *Wildl Soc Bull* 34:1463–1469
- Park BU, Marron JS (1990) Comparison of data-driven bandwidth selectors. *J Am Stat Assoc* 85:66–72
- Patterson TA, McConnell BJ, Fedak MA, Bravington MV, Hindell MA (2010) Using GPS data to evaluate the accuracy of state-space methods for correction of Argos satellite telemetry error. *Ecology* 91:273–285
- Peckham SH, Diaz DM, Walli A, Ruiz G, Crowder LB, Nichols WJ (2007) Small-scale fisheries bycatch jeopardizes endangered pacific loggerhead turtles. *PLoS ONE* 2:e1041
- Preston TJ, Chiaradia A, Caarels SA, Reina RD (2010) Fine scale biologging of an inshore marine animal. *J Exp Mar Biol Ecol* 390:196–202
- R Development Core Team (2011) R: a language and environment for statistical computing. [www.R-project.org](http://www.R-project.org)
- Rutz C, Hays GC (2009) New frontiers in biologging science. *Biol Lett* 5:289–292
- Ryan PG, Petersen SL, Peters G, Grémillet D (2004) GPS tracking a marine predator: the effects of precision, resolution and sampling rate on foraging tracks of African penguins. *Mar Biol* 145:215–223
- Schofield G, Bishop CM, MacLean G, Brown P and others (2007) Novel GPS tracking of sea turtles as a tool for conservation management. *J Exp Mar Biol Ecol* 347:58–68
- Seaman DE, Powell RA (1996) An evaluation of the accuracy of kernel density estimators for home range analysis. *Ecology* 77:2075–2085
- Seaman DE, Millsaugh JJ, Kernohan BJ, Brundige GC, Raedeke KJ, Gitzen RA (1999) Effects of sample size on kernel home range estimates. *J Wildl Manag* 63:739–747
- Sheppard JK (2008) *The spatial ecology of dugongs: applications to conservation management*. PhD Thesis, James Cook University, Townsville
- Sheppard JK, Preen AR, Marsh H, Lawler IR, Whiting SD, Jones RE (2006) Movement heterogeneity of dugongs, *Dugong dugon* (Muller), over large spatial scales. *J Exp Mar Biol Ecol* 334:64–83
- Silverman BW (1986) *Density estimation for statistics and data analysis*. Chapman & Hall, London
- Sims DW, Queiroz N, Humphries NE, Lima FP, Hays GC (2009) Long-term GPS tracking of ocean sunfish *Mola mola* offers a new direction in fish monitoring. *PLoS ONE* 4:e7351
- Sirtrack (2010) *Fastloc™ user guide version: 1.3*. Sirtrack, Havelock North
- Stoneburner DL (1982) Satellite telemetry of loggerhead sea turtle movement in the Georgia Bight. *Copeia* 1982:400–408
- Timko RE, Kolz AL (1982) Satellite sea turtle trackings. *Mar Fish Rev* 44:19–24
- Tomkiewicz SM, Fuller MR, Kie JG, Bates KK (2010) Global positioning system and associated technologies in animal behaviour and ecological research. *Phil Trans R Soc B* 365:2163–2176
- Tremblay Y, Shaffer SA, Fowler SL, Kuhn CE and others (2006) Interpolation of animal tracking data in a fluid environment. *J Exp Biol* 209:128–140
- Whitlock MC, Schluter D (2009) *The analysis of biological data*. Roberts & Company Publishers, Greenwood Village, CO
- Witt MJ, Åkesson S, Broderick AC, Coyne MS and others (2010) Assessing accuracy and utility of satellite-tracking data using Argos-linked Fastloc-GPS. *Anim Behav* 80:571–581
- Worton BJ (1989) Kernel methods for estimating the utilization distribution in home-range studies. *Ecology* 70:164–168