Reduction of location error in GPS collar tracking data of bovine cattle by using data screening

Katja Polojärvi, *City of Ylivieska, Finland*
Alfred Colpaert, *University of Eastern Finland,*
Kenneth Matengu, *University of Namibia*

Abstract

Global Positioning System (GPS) telemetry enables tracking of an individual animal over a long period of time and can provide researches with accurate information on animal movement. However, various environmental factors influence the satellite signals received by a GPS collar, potentially deteriorating accuracy of position. We tracked Caprivi Sanga cattle using Televilt Tellus Basic GPS collars in North-East Namibia in 2006 and 2007. Locations obtained during the night, when the cattle are gathered inside livestock enclosures, revealed that a significant proportion of the locations are inaccurate. We used data of seven GPS collars for testing different data screening options as a way to reduce location error. Basic analysis showed that simple measures of accuracy like dilution of precision (DOP) and figure of merit (FOM) are not sufficient to remove erroneous locations from the data. We removed the inaccurate locations with the following condition: 2D location with DOP >6 or 0< altitude <850 m or altitude >1050 m or DOP ≥10 or FOM ≥10 or walking speed of the animal over 4.5 km/h. This data screening option eliminated 75% of the most erroneous locations (>300 m from the livestock enclosures) retaining 97.2% of the locations correctly located inside the livestock enclosures. Before data screening, 95% (1372) of the night-time locations that were located outside the livestock enclosures were located 71-406 m from the enclosures. The maximum error was over 10 km. After data screening, 95% (485) of the locations were 54-298 m from the enclosures; the maximum error was 4.4 km.

Introduction

Global Positioning System (GPS) telemetry is a widely used method in studies of animal movement, habitat use and resource selection. GPS collars have many benefits in studies of animal movement: the collars enable tracking of an individual animal over a long period of time and automatically record geographical position at predefined time intervals. The

---

Katja Polojärvi is an Environmental Officer at the City of Ylivieska, Finland. She has M.Sc. and Phil. Lic. degrees in geography from the Universities of Oulu and Eastern Finland. She has worked as a researcher in many GIS projects related to air quality, precision farming and cattle grazing at the Universities of Jyväskylä and Eastern Finland and at the Oulu University of Applied Sciences. E-mail: katjia.polojarvi@oamk.fi

Alfred Colpaert was born in the Netherlands and studied physical geography in Utrecht. He has a Masters in Geography from the University of Oulu, Finland. His PhD thesis was on the subject of integration of remote sensing and geographical information systems. He held several posts at the Department of Geography of the University of Oulu. In 2004 he was appointed Professor in Physical Geography at the University of Eastern Finland. Present research topics include geoinformatics, range land and environmental studies.E-mail: alfred.colpaert@uef.fi

Kenneth Matengu is a Senior Researcher at the Multidisciplinary Research Centre, Social Sciences Division and Director of External and International Relations at the University of Namibia. His research interests includes applications of GIS in indigenous environmental knowledge research, tourism, community-based management and has expertise in decentralisation and local government as well as in innovation diffusion. E-mail: kmatengu@unam.na

Volume 1, Number 2, September 2012 - ISSN 2026-7215

method has been used in wildlife studies, but also in monitoring semi-domestic reindeer (Kumpula & Colpaert, 2007) and domestic cattle grazing (e.g., Turner, Udal, Larson, & Shearer, 2000; Agouridis et al., 2004; Bailey, Keil, & Rittenhouse, 2004). GPS measurement has been usually assumed to be reliable and accurate in comparison to conventional radio-telemetry. Recent studies indicate that GPS error can be substantial and need to be accounted for when location accuracy is required.

Reliability of GPS collar data depends on the fix rate (i.e., rate of GPS measurement success) and accuracy of obtained fixes. Availability, quality and strength of satellite signals received by a GPS collar can be influenced by various environmental factors, all potentially leading to unsuccessful fix attempts and increased positional errors. Terrain obstructions (D'Eon, Serrouya, Smith, & Kochanny, 2002; Cain III, Krausman, Jansen, & Morgart, 2005; Lewis, Rachlow, Garton, & Vierling, 2007) and vegetation characteristics such as canopy cover and height (Di Orio, Callas, & Schaefer, 2003; Agouridis et al., 2004; Frair et al., 2004; DeCesare, Squires, & Kolbe, 2005; Lewis et al., 2007; Hansen & Riggs, 2008) are examples of environmental factors that may interfere with connection between satellites and GPS receivers. Error in GPS data will have influence on analyses of animal locations and movements (e.g., Jerde & Visscher, 2005). For example, systematically failing fix attempts cause a loss of information that can influence assessments of resource selection by animals. Position inaccuracy can lead to misclassification of habitat use depending on the magnitude of location error and landscape heterogeneity (Frair et al., 2004, p.202). Animal activity also causes variation both in fix rate and location error (D'Eon & Delparte, 2005; Lewis et al., 2007).

During 2006 and 2007, we collected GPS data using GPS collars to study grazing and movement patterns of Caprivi Sanga cattle on the floodplains of the Zambezi River in East Caprivi, North-East Namibia. In East Caprivi, the cattle graze under supervision of herdsmen during daytime and spend the night inside livestock enclosures (locally called kraal). Examination of the downloaded data revealed that there are obvious inaccuracies in the data since some night-time positions were located outside the enclosures at distances exceeding the nominal accuracy of the GPS-module (+/- 15 m). The quantity and magnitude of the location errors demanded elimination of at least the largest errors from the data using data screening techniques. Aim of the data screening is to eliminate inaccurate locations from the data while retaining the maximum amount of accurate positions. Caution is necessary, as unwanted data reduction resulting from too rigorous data screening may introduce additional biases (e.g., Frair et al., 2004; Lewis et al., 2007). However, the acceptable level of data accuracy and data reduction depends on the research goals and methods.

In studies carried out with GPS collars, different data screening options have been tested usually based on 2D/3D fix and DOP values (dilution of precision) (e.g., D'Eon et al., 2002; D'Eon & Delparte, 2005; Lewis et al., 2007). Two-dimensional (2D) and three-dimensional (3D) fixes are obtained when three or more satellites, respectively, are available; the latter usually are more accurate than the former (e.g., Di Orio et al., 2003, p. 373). Dilution of precision (DOP) is a mathematical representation for the quality of the GPS position solution that is affected mainly by the configuration of (distance and angle between) the satellites used to obtain the position. As a general rule, positions that are obtained with many satellites and have a low DOP value are accurate. Positions that have either low number of satellites, or have a high DOP value may not be as accurate (Televilt, 2006, p. 27).

The aim of this study is to evaluate different data screening options and their effect on the magnitude of positional error, as a purpose to reduce the amount of inaccurate positions and thus increase accuracy of the data. We estimated: (i) effectiveness of a DOP limit of 10, recommended e.g. by D'Eon and Delparte (2005), and (ii) effectiveness of other variables recorded by the GPS collars (2D/3D, FOM and altitude) in elimination of inaccurate positions. We also estimated (iii) which critical variable values and combinations of variables reduce most effectively the magnitude of location error without causing unacceptable data reduction.

Study area

The GPS collar data were collected on grazing areas of seven villages in East Caprivi, North-East Namibia. Mutikitila, Limai, Isuswa, Ioma, Mubbu and Lyalumba are located in the Salambala Conservancy and Ivilivinzi on the floodplains east from Salambala. The annually flooding area is located between the Zambezi River and its tributary, the Chobe River (Figure 1). The topography of the area is very flat ranging between 926 to 937 m above sea level. The climate has a dry and a rainy season (summer), at the end of which the area is affected by sometimes severe flooding of the Zambezi River. Open grasslands and wetlands on the floodplains form a significant type of vegetation and landscape. On higher ground the vegetation is characterized by open savanna forest and shrubland.

During the tracking period, the cattle were kept overnight inside livestock enclosures (Figure 2). In total 12 livestock enclosures were used during this time. In Ioma and Mubbu, only the enclosure located in the village was used during the tracking period. In other villages, besides the enclosures in the villages the cattle were also moved to another grazing area where new enclosures were constructed for part of the year (Figure 1, Table 1).

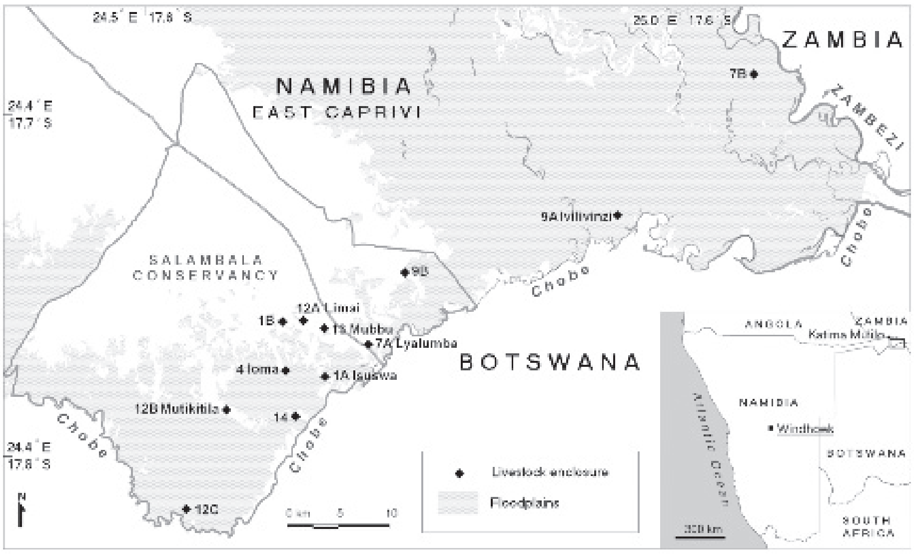

Figure 1. Location of the livestock enclosures in East Caprivi, North-East Namibia. Annually flooding areas are based on GIS data of Atlas of Namibia Project (2002).

Figure 2. The livestock enclosure, locally called “kraal”, in Lyalumba, East Caprivi (Photo: Katja Polojärvi ,4 July 2006).

Materials and methods

During 2006 and 2007, 14 Televilt Tellus Basic 5H2D v2.0 (store onboard, nominal accuracy +/- 15 m) GPS collars (Televilt/Followit Lindesberg Ab, Sweden) were deployed on Caprivi Sanga cattle in ten villages of East Caprivi. The collars were programmed using Tellus Project Manager (TPM) to record one location per hour every day of the year. GPS Positioning Time was set to 90 s and the sensitivity for the activity sensor was set to five (medium sensitivity). In addition to the date (yyyy-mm-dd), time (hh:mm, Greenwich Mean Time) and geographic coordinates (Latitude/Longitude in decimal degrees), the collars also recorded the following information (Televilt, 2006, p. 27):

- Time (s) the GPS receiver has used to obtain the fix.
- SV: number of the satellites used to obtain the fix.
- Altitude (m) when at least four satellites are available.
- 2D/3D: the obtained fixes are three-dimensional when the collar has contact with four or more satellites. Otherwise the obtained fixes are two-dimensional.
- DOP: dilution of precision is a measure of the quality of the GPS data being received from the satellites. DOP is a mathematical representation for the quality of the GPS position solution that is affected mainly by the configuration of (distance and angle between) the satellites used to obtain the position. DOP values are between 0.0 and 25.0.
- FOM: figure of merit values indicates the best accuracy achievable from the satellites being tracked; the lower the value, the more accurate position. The

calculation of FOM values is manufacturer specific and varies between different GPS module brands.

- Temperature (C°) inside the main housing at the time the position was obtained.
- X,Y: The activity level that is measured as a certain change in collar position during the time the collar has been used to obtain the fix.

Tracking period and amount of data varied due to several reasons. 10 GPS collars (collars 1-10), which were deployed on bulls in July 2006, encountered severe problems; one collar (collar 2) was lost immediately and six collars (collars 3, 4, 5, 6, 8 and 10) were disabled as a result of serious damage. Physical damage was most certainly caused by fighting between bulls in the same herd. During the tracking period, the collars were retrieved from the bulls, were downloaded and the functioning collars were redeployed on cows of the same herds. Four new collars (collars 11-14) made of reinforced material were also deployed in November 2006 and June 2007 to replace some of the disabled collars. One of the reinforced collars (collar 11) had serious malfunctions. In total, data of eight collars were lost or incomplete. The seven best functioning collars operating on the grazing areas of seven villages and in 12 livestock enclosures were chosen to study data accuracy and screening options (Table 1).

Most obvious errors in the GPS data are recorded fixes without geographical coordinates or with coordinates where latitude and/or longitude are saved incorrectly, these errors were deleted from the data. Fix rates (%), or proportions of the successfully obtained fixes to fix attempts, were calculated for the collar data. The sequential recording of locations at hourly intervals enables calculation of hourly walking distances of the tracked animals. The distances between sequentially recorded locations, also called step lengths, were calculated using Hawth's Analysis Tools for ArcGIS (version 3.26) (Beyer 2004).
Table 1. Operation periods of seven GPS collars in 12 livestock enclosures in East Caprivi, North-East Namibia in 2006 and 2007.

GPS collar	Enclo-sure	Location of the enclosure:	Latitude, Longitude (DMS)	Operation period of the GPS collar
1	1A	Isuswa:	17°55′55″S, 24°39′55″E	1.7.2006 – 20.8.2006, 29.11.2006 – 7.3.2007[1])
	1B	Grazing area (Isuswa):	17°53′0″S, 24°37′33″E	7.3.2007 – 4.6.2007
4	4	Ioma:	17°55′36″S, 24°37′42″E	2.7.2006 – 30.9.2006
7	7A	Lyalumba:	17°54′13″S, 24°42′17″E	4.7.2006 – 23.8.2006
	7B	Grazing area (Lyalumba):	17°39′58″S, 25°3′41″E	23.8.2006 – 1.2.2007
9	9A	Ivilivinzi:	17°47′23″S, 24°56′8″E	6.7.2006 – 18.2.2007
	9B	Grazing area (Ivilivinzi):	17°50′25″S, 24°44′20″E	18.2.2007 – 5.6.2007
12	12A	Limai:	17°52′57″S, 24°38′41″E	7.1.2007 – 1.6.2007[2])
	12B	Mutikitila:	17°57′40″S, 24°34′26″E	7.6.2007 – 15.8.2007
	12C	Grazing area (Mutikitila):	18°2′54″S, 24°32′11″E	15.8.2007 – 4.12.2007
13	13	Mubbu:	17°53′22″S, 24°39′52″E	5.6.2007 – 7.12.2007
14	14	Grazing area (Limai):	17°52′57″S, 24°38′41″E	4.6.2007 – 7.12.2007

[1]) The collar fell off on 21 August 2006, but was retrieved and redeployed on 29 November 2006.
[2]) The animal carrying the collar 12 used two other livestock enclosures on the grazing area of Limai from November 2006 to January 2007. The data are patchy as a result of constant movement between the enclosures and thus excluded from the study.

Data classification

The collar data were imported into the ArcMap 9.1 (ArcGIS 9/ESRI Inc., USA) GIS software for reviewing and analysing. Firstly, the data of the seven collars were divided into 12 data sets according to a livestock enclosure (Table 1). Date and time when the enclosures and grazing areas were used are easily recognised in ArcMap. Secondly, the collar data were classified into the daytime grazing locations and the night-time enclosure locations. The evaluation of different data screening options requires that accurate and inaccurate locations are known. Thus, thirdly, the enclosure locations were again classified into "accurate" inside and "inaccurate" outside locations.

In autumn 2008, we were able to record perimeters of the four livestock enclosures located on the grazing areas of Isuswa and Ivilivinzi (collars 1 and 9) using GPS. Perimeters of the other enclosures were not recorded in the field and thus a rigorous MCP (minimum convex polygon) approach was used in data classification.

Collars 1 and 9

During the field trip in September 2008, the perimeters of the four livestock enclosures of Isuswa and Ivilivinzi were recorded using a Fortuna Slim Bluetooth GPS (Fortuna Electronic, Taiwan) connected to a Toshiba ToughBook laptop with GeoMedia Professional 6.0 (Intergraph, USA) GIS software. The digitized perimeters of the enclosures were imported into the ArcMap 9.1 (ArcGIS 9/ESRI Inc., USA). Because of the nominal GPS position error (+/- 15 m), a buffer of 15 m was set around the perimeters of the enclosures. This buffer was used to remove positional ambiguity, e.g. ensuring that observations were correctly classified into inside and outside locations. The classification of the data into grazing and enclosure locations is also based on the boundary formed by the outer edge of the buffer.

In ArcMap, all locations were checked and classified according to arrival in and departure from the buffered enclosures. Movement is easy to review in ArcMap, and based on that the animals graze as late as possible and then walk directly into the enclosure before sunset. It must also be noted that any damage to crops by cattle carries a fine, as a consequence of which people ensure that their cattle are indeed in the enclosures during the growing season.

Several presumptions were needed when classifying the data: (1) When the animal arrives in the enclosure in the afternoon (between 15:00 and 18:00 GMT), it stays there until the following morning. (2) In case there are many error locations during the afternoon (i.e., many locations outside but very near the enclosure), the first location near the enclosure is the last grazing location of the day and the animal goes thereafter directly into the enclosure. Thus, the other locations outside but very near the enclosure are classified as inaccurate. (3) After leaving the enclosure in the morning (between 5:00 and 7:00 GMT) the animal does not go back into the enclosure before the afternoon. (4) In case there are many error locations during the morning, the last location outside but very near the enclosure is assumed to be the first grazing location. To avoid possible misclassification, all locations were carefully checked hour by hour in ArcMap revealing possible escapes or other reasons why tracked animals moved outside the enclosures during the night. In case the animal had escaped, was lost or migrating, the movements outside the enclosure are systematic and logical and thus easily recognised. However, movements of this kind were very exceptional. In total 226 locations in data of the collars 7, 9, 12, 13 and 14 were excluded from the study due to this kind of night-time movements.

The evening and night-time locations, when the cattle are supposed to be inside the enclosure, were divided in "accurate" locations inside and "inaccurate" locations outside the buffered enclosures. In ArcMap, the distances between the inaccurate outside locations and the enclosure buffers were calculated. It must be noticed that the diameters of the buffered enclosures varied between 67 and 102 m and thus all locations falling inside the buffered enclosures are not necessarily accurate. However, accurate locations are concentrated inside the buffered enclosures, and locations outside the enclosures are mainly inaccurate. Distribution of the night-time enclosure locations inside and outside the four buffered enclosures is illustrated in Figure 3.

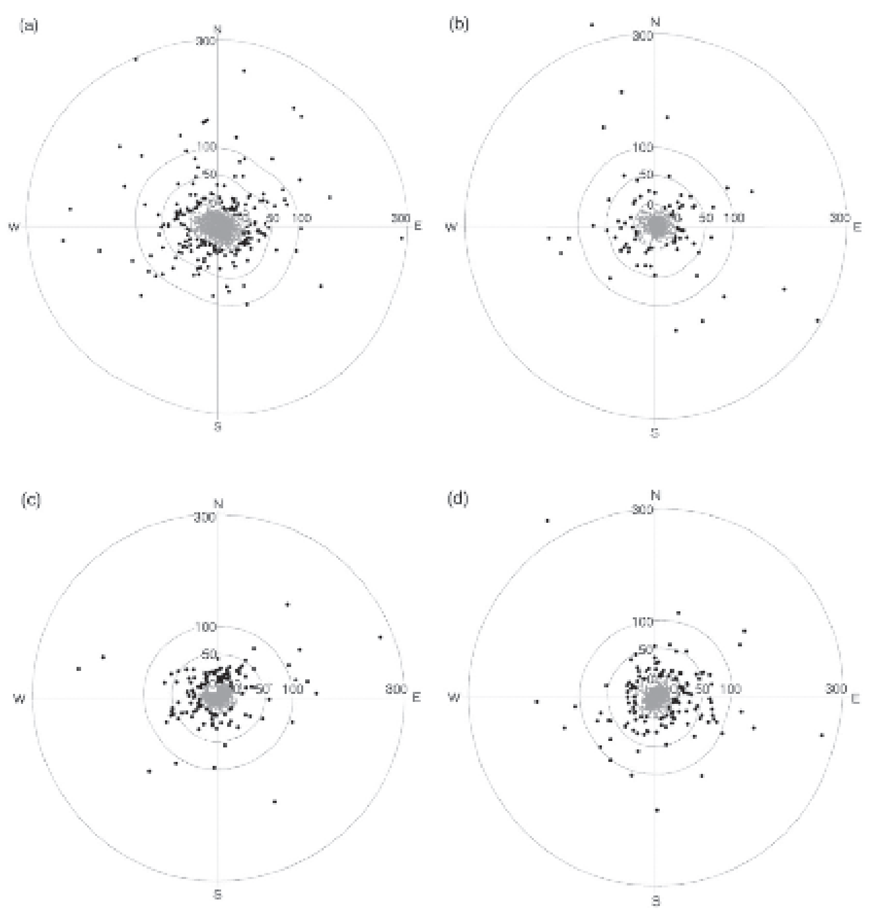

Figure 3. Distribution of the night-time enclosure locations inside (grey dots) and outside (black dots) the four livestock enclosures: (a) 1A, (b) 1B, (c) 9A and (d) 9B. Circles (0, 50, 100 and 300) illustrate the distance (m) from the outer edge of the enclosure buffer. Largest outliers are not shown in the figure.

Collars 4, 7, 12, 13 and 14

The perimeters of the livestock enclosures used by the animals carrying the collars 4, 7, 12, 13 and 14 were not digitized with GPS in the field. These collars were used to test and refine the data screening strategies derived from analysing collars 1 and 9. To obtain the size and form of the enclosures we generated polygons for the enclosures based on the fact that the night-time observations were predominantly correctly located inside the enclosures forming dense clusters that reveal the location and even the shape of the enclosures. To simplify the classification of very large amount of locations into the daytime grazing locations and the night-time enclosure locations, the arrival in and departure from the enclosures in the morning and in the afternoon were first manually checked in ArcMap. Because the exact perimeters of the enclosures were not known, only the locations between 20:00 and 4:00 (GMT) were chosen for the study. For the classification into the "accurate" locations inside and "inaccurate" locations outside the enclosures we created minimum convex polygons (MCP) of 90% and 40% by using the Home Range Tools for ArcGIS (version 1.1.) (Rodgers, Carr, Beyer, Smith, & Kie, 2007). The floating mean algorithm calculates the arithmetic mean of all points, and then drops the farthest single point. The mean is recalculated from the subset of points and another point is dropped. This continues until the requested percentage of points remains selected (Rodgers & Carr, 2002, p. 17).

MCP 90% forms a polygon clearly outside the location clusters leaving the farthest and most erroneous locations outside the polygon. The "accurate" locations inside the enclosures were chosen by creating MCP 40%. The result of the calculation is a smaller polygon clearly inside the dense location clusters. The limit of 40% was not calculable in data of enclosures 7B (Lyalumba) and 14 (Limai) and thus MCP 41% and MCP 50%, respectively, were used. The real perimeter of the enclosure is located approximately between the edges of MCP 90% and inner MCP 40% (41%, 50%). All the locations in this zone were excluded from the study. Only the farthest locations outside the MCP 90% and the locations inside the MCP 40% core of the enclosures were included in the study. An example of data classification is illustrated in Figure 4.

Location error means the distance between a location recorded by GPS receiver and a true location of the tracked object at the moment when the GPS fix is obtained. Because the livestock enclosures are areas, exact location errors in metres are not calculable. However, understanding about the magnitude of location error is necessary. Because of this, the geometric centroid (i.e., centre of area) of the MCPs 90% and distances between the GPS locations and the centroid were calculated in ArcMap. The centroids do not represent the real geometric centres of the enclosures but are close approximations of them.

Evaluation of data screening options

The simple method to remove erroneous locations from GPS tracking data is to remove all locations above or below a certain DOP, FOM or 2D/3D limit. However, additional measures can be calculated from the GPS data. We have utilized the step length between hourly observations as a measure of accuracy. The maximum walking distance per hour of the cattle was estimated from hourly observations during the migration from one grazing area to another. The maximum hourly walking distances were 4019 m/h (collar 1, Ioma), 4396 m/h (collar 7, Lyalumba) and 4042 m/h (collar 9, Ivilivinzi). For example, the tracked animal of Lyalumba migrated nearly 50 km from the eastern flooded area to the hinterland during one day, which means that the walking had to be straightforward. The step length limit was set to 4500 m. Before removing these cases from the data, time between two sequential locations was ensured to be one hour by manual checking in ArcMap. We also applied a measure to verify 3D accuracy, as the area is very flat, we could identify 3D observations with exceptional large deviations from the local base level.

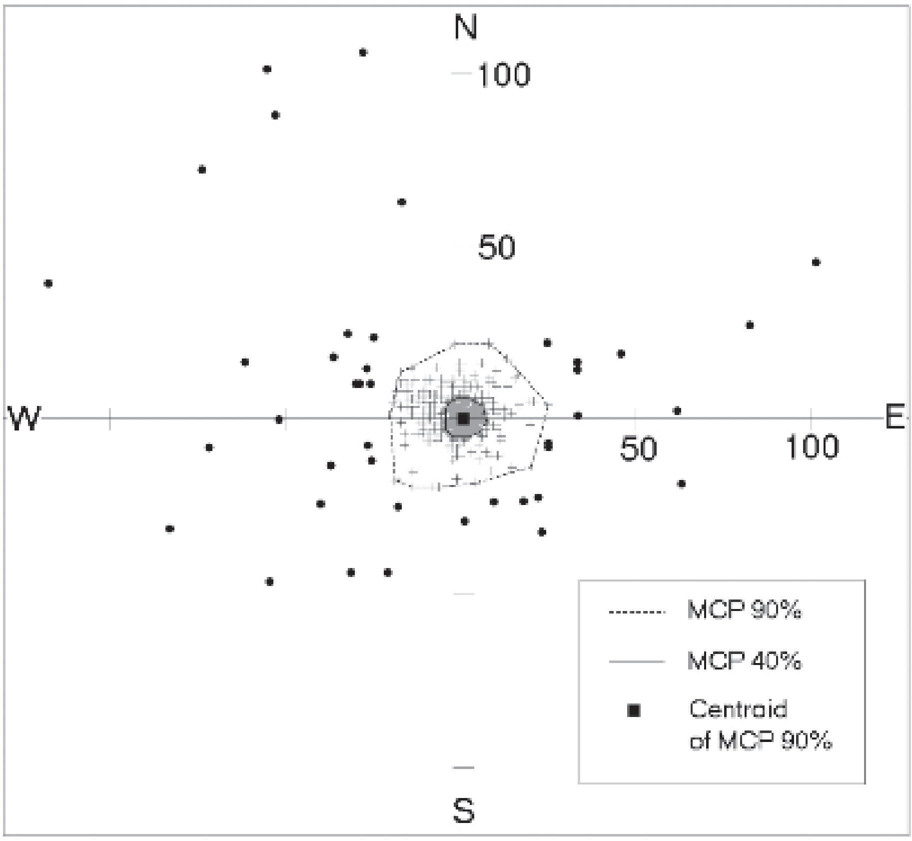

Figure 4. An example of data classification: The night-time enclosure locations (recorded 20:00-4:00 GMT) in data 7A (Lyalumba) classified into the accurate locations inside the livestock enclosure (grey dots inside MCP 40%), the inaccurate locations outside the enclosure (black dots outside MCP 90%) and the locations excluded from the study (black crosses) based on the minimum convex polygons. The distances (50 and 100 m) are calculated from the centroid of the MCP 90% (black square).

DOP and FOM values and altitudes with different limits, step lengths over 4500 m and 2D fixes were tested as data screening options, first each variable separately, and thereafter using different combinations and limits of the variables. First, proportions of these variables among the "inaccurate" outside locations and "accurate" inside locations were calculated with SPSS 16.0 (SPSS Inc., USA) statistical software. 3D locations, low DOP and FOM values and altitudes near the base altitude level (926-937 m.a.s.l.) are expected to be prevalent among the "accurate" inside locations. On the contrary, 2D locations, high DOP and FOM values as well as altitudes clearly differing from the base altitude level should be prevalent among the "inaccurate" outside locations.

Different data screening options were evaluated using the enclosure locations of 1A, 1B, 9A and 9B. In SPSS, the ability of different options to eliminate largest location errors and resulting elimination of accurate locations were compared by calculating proportions of eliminated locations among the "inaccurate" outside locations for different distances

(<=100 m, >100 m, > 300 m) and likewise among the "accurate" locations inside the buffered enclosures. Effects on the magnitude of location error were estimated by calculating mean, standard error of mean, standard deviation and frequency percentiles (50%, 95%, 99% and 100%) for the distances the "inaccurate" locations are from the buffered enclosures, and by comparing these calculations before and after data screening. It must be stressed that the calculated distances only approximate the magnitude of location error and are not absolute values expressed as radial distance of error location from true location.

The results of the first phase were used to distinguish the variables and combinations that were most effective in data screening. These data screening options were again tested using the data of collars 4, 7, 12, 13, and 14. In SPSS, mean, standard error of mean, standard deviation and frequency percentiles (50%, 95%, 99% and 100%) were calculated for the distances between the "inaccurate" locations and the centroids of the MCPs 90%. The calculations were made before and after data screening with different options. Proportions of eliminated locations among the "inaccurate" locations on different distances (<=100 m, >100 m, > 300 m) and among the "accurate" locations inside the MCPs 40% (41%, 50%) were also calculated.

Results

The fix rates of the collars are mainly over 95%, collars 12 and 13 have the lowest fix rates 90.3% and 86.4% respectively. In data of collars 1 and 9, the night-time enclosure locations comprise over 50% of all the obtained fixes, and 8.5 and 6.4% of them, respectively, were "inaccurate" locations outside the enclosures. In data of the other collars, only enclosure locations between 20:00 and 4:00 (GMT) were chosen for the study decreasing the number of observations in the test. Because data of these collars were classified using minimum convex polygons of 40% and 90%, proportions of the locations classified as "accurate" inside and "inaccurate" outside the enclosures were naturally approximately 40% and 10%, except in data 7B and 14 where MCP40% was replaced with MCP41% and MCP50% (Table 2).

Table 2. Fix rates (%) and successfully obtained fixes for seven Televilt Tellus Basic GPS collars in East Caprivi, North-East Namibia as classified into the night-time enclosure locations (inside and outside) and daytime grazing locations.

	Collar						
	1	4	7	9	12[1])	13	14
Fixes	5528	2098	4868	7778	7053	3848	4344
Fix rate %	96.8	96.8	95.9	97.0	90.3	86.4	97.1
Enclosure locations[2])	3529	805	1791	4648	2784	1548	1574
Classified inside[3])	3228	323	731	4352	1115	620	791
%	91.5	40.1	40.8	93.6	40.1	40.1	50.3
Classified outside[4])	301	80	179	296	277	154	157
%	8.5	9.9	10.0	6.4	9.9	9.9	10.0
Grazing locations	1999	817	1835	3090	2384	1201	1657
Other locations[5])	0	476	1242	40	1885	1099	1113

[1]) Tracking period when the livestock enclosures 12A-C were used by the collared animal.
[2]) Collars 4, 7, 12, 13 and 14: the enclosure locations recorded between 20:00 and 4:00 GMT.
[3]) Collars 1 and 9: inside the buffered enclosures. Collars 4, 7, 12 and 13: inside MCP40% (data 7B: MCP41%). Collar 14: inside MCP50%.
[4]) Collars 1 and 9: outside the buffered enclosures. Collars 4, 7, 12, 13 and 14: outside MCP90%.
[5]) Locations when the tracked animal migrated or otherwise stayed overnight somewhere else than in the livestock enclosure. In data of collars 4, 7, 12, 13 and 14, the enclosure locations recorded before 20:00 and after 4:00 (GMT) are also included in these locations.

Obtained fixes are mainly three-dimensional (3D), proportions of which are higher among the "accurate" locations inside the livestock enclosures. Two-dimensional (2D) fixes are more prevalent among the "inaccurate" locations outside the enclosures. The proportion of 2D fixes is as high as 34% in data 12A, but mostly the proportions are less than 10%. DOP <3 values are clearly more prevalent among the "accurate" locations than among the "inaccurate" locations. Proportions of DOP >10 values among the "inaccurate" locations vary between 0.0 and 8.8% indicating that a DOP limit of 10 is ineffective. Differences in proportions of FOM >10 values among the "accurate" and "inaccurate" locations are instead much larger being higher among the "inaccurate" locations varying between 9.4 and 22.1%. Calculations made with different altitude limits revealed that altitudes over 0 m but under 850 m or altitudes over 1050 m eliminated most effectively inaccurate locations. These altitudes are more prevalent among the "inaccurate" locations and are actually very infrequent among the "accurate" enclosures locations. This indicates that the altitude limits are able to remove erroneous locations from the GPS data, while retaining accurate locations. Only a few outliers are located over 4500 m from the enclosures and thus very few locations have a step length value exceeding 4500 m (Table 3).

Table 3. Variety of data screening options and proportions (%) of fixes inside and outside the livestock enclosures fulfilling the conditions.

		Fixes (%)							
		Data screening option:							
Enclosure		3D	2D	DOP<3	DOP>10	FOM>10	0<alt <850 m	alt>1050 m	step length >4500 m
Inside									
1A	2321	96.7	3.3	84.4	0.4	2.0	0.9	0.5	0.0
1B	907	93.5	6.5	81.8	0.9	1.8	1.0	0.7	0.0
4	323	97.5	2.5	86.1	0.6	0.0	0.6	0.0	0.0
7A	168	99.4	0.6	88.1	0.0	0.6	0.0	0.6	0.0
7B	563	99.6	0.4	94.7	0.0	0.4	0.0	0.0	0.0
9A	2902	99.1	0.9	87.8	0.3	0.8	0.3	0.1	0.0
9B	1450	97.7	2.3	83.1	0.5	1.6	1.0	0.6	0.0
12A	490	94.7	5.3	79.2	0.8	1.2	0.2	0.0	0.0
12B	230	95.2	4.8	84.8	0.0	1.8	0.0	0.4	0.0
12C	395	96.5	3.5	85.3	0.0	1.5	0.0	0.3	0.0
13	620	100.0	0.0	91.5	0.0	1.1	0.0	0.0	0.0
14	791	100.0	0.0	87.4	0.3	0.9	0.0	0.0	0.0
Outside									
1A	214	85.5	14.5	58.9	5.1	17.3	20.1	10.7	0.0
1B	87	87.4	12.6	60.9	0.0	18.4	21.8	10.3	0.0
4	80	92.5	7.5	56.3	3.8	15.0	8.8	18.8	0.0
7A	42	95.2	4.8	69.0	2.4	14.3	7.1	14.3	2.4
7B	137	92.0	8.0	82.5	0.7	10.2	13.9	7.3	0.0
9A	149	94.0	6.0	68.5	5.4	9.4	8.1	8.1	2.0
9B	147	90.5	9.5	63.9	2.0	13.6	12.2	16.3	0.7
12A	122	65.6	34.4	47.5	8.2	22.1	14.8	14.8	0.0
12B	57	84.2	15.8	50.9	8.8	21.1	17.5	14.0	0.0
12C	98	90.8	9.2	64.3	4.1	14.3	15.3	5.1	0.0
13	154	100.0	0.0	60.4	1.9	12.3	1.9	0.6	0.0
14	157	100.0	0.0	55.4	1.9	12.1	4.5	0.6	0.0

In the combined data of the seven GPS collars, removal of all locations outside the enclosures that have DOP ≥10 eliminated 10% of the locations at a distance of more than 300 m from the livestock enclosures (distance from the enclosure buffers in data of collars 1 and 9, and from the centroids of MCP90% polygons in data of collars 4, 7, 12, 13 and 14). FOM ≥10 as a data screening option was instead able to eliminate 25% of the outside locations on the same distance. Lower limits for DOP and FOM values were also tested but they increased elimination of accurate locations inside the enclosures. Compared to the DOP and FOM values, data screening option 0< altitude <850 m or altitude >1050 m was more effective in elimination of the erroneous locations and in retaining accurate locations (Table 4). However, all these three options were not able to eliminate location errors satisfactorily by themselves.

Before data screening, the magnitude of location error expressed as mean distance between the “inaccurate” locations and the livestock enclosures varied between 31.5 and 159.6 m in data of seven collars. 95% of the “inaccurate” locations outside the enclosures were at a distance from 71.1 to 405.9 m. Maximum distance was as high as 10327 m in data of collar 9 and only 150 m in data of collar 13 (Table 5).

The best performing screening option for the combined data was as follows:
2D location with DOP >6 or 0< altitude <850 m or altitude >1050 m or step length >4500 m or DOP ≥10 or FOM ≥10

This combination of variables was able to eliminate 75% of the most erroneous locations at a distance of more than 300 m (30 locations eliminated from 40 most erroneous locations) and removed only 2.8 % of the “accurate” locations (315 eliminated from 11160 “accurate” locations) (Table 4). After data screening with this option, the mean distance of the remaining “inaccurate” locations vary between 23.5 and 133.0 m. 95% of these remained locations were at a distance of 54.5 to 297.6 m. In data of collar 12, one location with a distance of 4423 m remained. In data of the other collars, maximum distances were from 130.5 to 363.6 m after data screening (Table 5).

Table 4. Variety of data screening options and elimination % of the locations outside and inside the livestock enclosures in individual and combined data of seven GPS collars classified according to the distance from the enclosures.

Data screening option	Distance from enclosure	Elimination % of the locations							
		Collar							
		1	4	7	9	12	13	14	ALL
DOP ≥10									
	> 300 m	0.0	0.0	0.0	14.3	14.3	-	0.0	10.0
	> 100 m	2.1	7.1	0.0	8.8	12.5	0.0	0.0	7.4
	≤ 100 m	3.9	3.0	1.3	3.1	3.9	2.0	2.1	2.9
	Inside	0.6	0.6	0.0	0.4	0.4	0.0	0.3	0.4
FOM ≥10									
	> 300 m	16.7	0.0	0.0	28.6	28.6	-	50.0	25.0
	> 100 m	29.8	42.9	28.6	32.4	31.3	60.0	14.3	31.2
	≤ 100 m	18.1	13.6	11.4	10.3	13.3	13.4	11.9	13.3
	Inside	2.4	0.3	0.4	1.4	2.1	1.3	1.0	1.6
0< altitude <850 m or altitude >1050 m									
	> 300 m	66.7	66.7	0.0	28.6	23.8	-	0.0	32.5
	> 100 m	53.2	57.1	52.4	47.1	37.5	20.0	0.0	42.0
	≤ 100 m	19.5	21.2	17.1	19.1	21.0	2.0	5.6	17.2
	Inside	1.4	0.6	0.1	0.8	0.3	0.0	0.0	0.8
2D location with DOP >6 or 0< altitude <850 m or altitude >1050 m or step length >4500 m or DOP ≥10 or FOM ≥10									
	> 300 m	83.3	100.0	100.0	100.0	66.7	-	0.0	75.0
	> 100 m	68.1	85.7	66.7	82.4	63.5	60.0	7.1	65.4
	≤ 100 m	43.3	31.8	63.6	28.6	34.3	16.1	17.5	29.6
	Inside	4.4	2.2	0.5	2.6	3.0	1.3	1.3	2.8

Table 5. The magnitude of location error expressed as distances (m) the inaccurate locations are located outside the livestock enclosures1) in data of seven GPS collars before and after data screening with the option: 2D locations with DOP >6 or 0< altitude <850 m or altitude >1050 m or step length >4500 m or DOP ≥10 or FOM ≥10.

Collar (enclosures)	Inaccurate locations	Eliminated locations	Distance (m)[1])					
					Percentiles			
	n	%	Mean (S.E.)	SD	50	95	99	100
Before data screening:								
1 (A, B)	301		56.7 (6.4)	111.3	23.8	213.3	551.7	1300.2
4	80		95.2 (15.0)	134.2	63.3	284.6	1171.0	1171.0
7 (A, B)	179		96.2 (43.9)	587.2	38.0	141.3	1788.5	7892.2
9 (A, B)	296		159.6 (60.0)	1032.6	18.4	166.0	8849.3	10326.5
12 (A, B, C)	277		156.9 (23.1)	385.3	80.3	405.9	3043.8	4423.2
13	154		31.5 (1.6)	19.4	24.9	71.1	132.4	150.3
14	157		49.7 (4.3)	54.3	30.5	186.6	329.8	363.6
After data screening:								
1 (A, B)	159	47.2	35.2 (4.2)	53.2	15.2	153.0	307.6	329.9
4	47	41.3	61.7 (2.3)	19.0	60.7	98.9	130.5	130.5
7 (A, B)	123	31.3	44.8 (2.4)	27.1	35.6	112.9	179.7	191.2
9 (A, B)	193	34.8	23.5 (2.1)	29.8	12.8	74.0	142.0	221.4
12 (A, B, C)	154	44.4	133.0 (30.6)	380.2	67.2	297.6	2890.1	4423.2
13	127	17.5	29.1 (1.4)	16.1	24.4	54.5	136.8	150.3
14	131	16.6	48.1 (4.7)	53.3	30.0	187.0	342.3	363.6

[1]) Collars 1 and 9: Distance from the outer edge of the buffered perimeter of the livestock enclosure.
Collars 4, 7, 12, 13 and 14: Distance from the centroid of minimum convex polygon of 90%.

Discussion

The GPS collar data were collected in a true pasture environment and natural circumstances, thus representing in practise the problems and challenges that are related to this method. The study revealed that animal tracking by GPS collars effectively produces long-term data about the animal's movements. However, the method also suffers from position inaccuracies, malfunctions and even loss or destruction of some devices. Stronger belt materials reduced at least part of these problems. Selecting of tranquil animal individuals for tracking is also recommended. A possible solution to GPS accuracy problems could be collars allowing differential correction of the downloaded data. However, the seven best functioning GPS collars produced data of animal movements from three months to almost one year. Two collars had a fix rate of 90% or less, but otherwise the fix rates are high (96-97%).

The GPS collar data consist of both failed fix attempts and successfully obtained fixes with different accuracy depending on the prevailing conditions. Fix success and positional accuracy are dependent on the number and configuration of the available GPS satellites, atmospheric conditions and topography. Also, local factors, such as canopy cover and incorrect position of a collar (i.e., GPS antenna not facing the sky), interfere with reception of the GPS signal, thus diluting the fix success and accuracy.

The locations obtained during the night-time were mainly correctly located inside the livestock enclosures, but 6.4-10.0% of the locations were inaccurate outside the enclosures and the maximum errors were several kilometres (150-10327 m). Actually, the proportions of inaccurate locations are higher because of the use of a MCP 90% in testing inaccuracy of collar data 4, 7, 12, 13 and 14. However, less than 3% of these inaccurate locations were located further than 300 m from the livestock enclosures.

When animal tracking is implemented during a long period of time, such as one year, it is not possible to observe tracked animals continuously and examine all constantly changing factors that possibly are affecting GPS performance. Since these factors in the research area were not rigorously studied, we did not estimate causes for the inaccuracies of the data but rather tried to achieve a reasonable method to eliminate largest errors and thus increase the accuracy of the data. This method, also called data screening, lacks general rules how GPS collar data should be processed to eliminate inaccurate locations and retain accurate ones. Dilution of precision (DOP) values, especially with limit of 10, and two-dimensional (2D) fixes are examples of data screening options used in GPS collar studies, but commonly accepted methods are lacking. D'Eon & Delparte (2005, pp. 387-388) have concluded that even the establishment of a rigorous PDOP based (positional dilution of precision) screening method would be difficult because PDOP values are not necessarily related to the size of the location error. According to them, a PDOP limit of 10 removes effectively major outliers but does not significantly change an average location error. We concluded that DOP values are very ineffective in data screening. If the DOP value is ≥10, the location is most likely inaccurate, and removal of these cases from the data can be recommended. But our study also revealed that only very few cases of the inaccurate locations actually have high DOP values. Only four cases (10%) of 40 most erroneous locations (distance > 300 m) have DOP ≥10 and any of the largest outliers are not included them. The "accurate" locations inside the livestock enclosures have mainly DOP <3 (79-95% of the locations), but the low DOP values are also fairly common among the "inaccurate" locations outside the enclosures (48-83% of the locations).

2D fixes are obtained when only three satellites are available. Small number of the satellites may reduce accuracy of the obtained fixes, but not necessarily. Based on the data screening evaluations, we do not recommend removal of all 2D locations automatically. Removal of 2D location when DOP value is also higher retains more effectively accurate locations. We used 2D location with DOP >6 as one data screening option. The result supports the view of Lewis et al. (2007, p. 670) who also recommended to screen out 2D locations at a specific PDOP cut-off (PDOP >5).

Figure of merit (FOM) values ≥10 and altitudes differing from the base altitude level (0< altitude <850 m or altitude >1050 m) were more effective in data screening than DOP values. Altitudes offered the best single variable recorded by the GPS collars that was able to eliminate 32.5% of the most erroneous locations (distance > 300 m) retaining 99.2% of the "accurate" locations inside the livestock enclosures. The altitude given by the GPS is dependent upon not only on the number of satellites, but also on satellite constellation, atmospheric conditions and terrain, thus partially indicating circumstances during a fix attempt. However, any of these single variables were not able to eliminate the largest location errors by themselves.

We concluded that the best and the only way to remove the largest location errors from the data is usage of a combination of several variables. We achieved the best result by using the data screening option; 2D location with DOP >6 or 0< altitude <850 m or altitude >1050 m or step length >4500 m or DOP ≥10 or FOM ≥10 which removed effectively largest location errors retaining accurate locations and significantly decreased the magnitude of location error. However, this data screening option cannot be fully generalized to other data and research areas. Data screening options presented in GPS collar studies, in general, seem to be difficult to generalise to other data. Based on our results, significance of DOP limit of 10 in data screening is very small, but on the other hand, there is not any harm to use it as a data screening option. FOM values are instead manufacturer specific calculations and vary between GPS module brands. Thus, FOM limit of 10 is probably applicable only if the tracking data are collected using the same manufacturer and GPS module than we did. The altitude proved to be the most effective single variable in data screening, which shows that other options than DOP values and 2D locations should also be considered and tested as data screening options. In our case, the study area is topographically very flat and thus altitudes that were remarkably different from the base altitude level of 926-937 m above sea level were used as a data screening option. However, altitudes are difficult to apply in data screening if the research area is topographically variable. A digital elevation model could be very useful for this option. In addition, abnormally high hourly walking distances of tracked animals are also very useful in data screening. In our study, a step length limit of 4.5 km/h removed effectively the largest outliers from the data. This requires measurement or estimation of the maximum speed of the tracked animal species.

It must be accepted in data screening that all errors cannot be removed from data, because at some point it will result in an unacceptable elimination of accurate locations. What criteria should be used in data screening, how effectively error locations should be eliminated and how high amount of data reduction is acceptable depends on the specific research goals. This requires profound understanding of both data and analysis methods.

Conclusions

We conclude that GPS collar tracking data contain location errors that must be taken into account before further data analyses. A DOP limit of 10 is not able to remove largest location errors from the data and thus insignificantly increases data accuracy. If DOP value is >10, the location is most likely inaccurate and can be removed from the data, but its significance in data screening is very weak. The altitude is the most effective single variable in data screening. However, any of the single variables recorded by GPS collars are not able to eliminate the largest location errors by themselves. Data screening option; 2D location with DOP >6 or 0< altitude < 850 m or altitude >1050 m or step length >4500 m or DOP ≥10 or FOM ≥10 offered the most effective solution to eliminate the inaccurate locations from the data, while retaining accurate locations and thus significantly reducing the magnitude of location error.

Acknowledgements

We want to express our gratitude and thanks to Mr. Shadreck Siloka, Mr. David Matengu, Mr. Kamwi Kamwi, Mr. Mathew Jankie, Mr. Morgan Sinkolela, Mr. Richard Nyambe, Ms. Bena-Mukela Mutemwa, Mr. Mbukusa, Mr. Daniel Mbala and Mr. Simata in Katima Mulilo and the rural villages who allowed us to track their cattle and made the effort to help us. We warmly thank Mr. David Matengu who troubled himself with solving out many practical problems and making sure the success of the GPS tracking. The project was funded by the Finnish Academy of Sciences, project number 110573, and the Ministry of Foreign Affairs.

References

Agouridis, C. T., Stombaugh, T. S., Workman, S. R., Koostra, B. K., Edwards, D. R., & Vanzant, E. S. (2004). Suitability of a GPS collar for grazing studies. *American Society of Agricultural Engineers*, 47, 1321–1329.

Atlas of Namibia Project (2002). Directorate of Environmental Affairs, Ministry of Environment and Tourism. Retrieved from Digital Atlas of Namibia website: http://www.uni-koeln.de/sfb389/e/e1/download/atlas_namibia/main_namibia_atlas.html.

Bailey, D. W., Keil, M. R., & Rittenhouse, L. R. (2004). Research observation: Daily movement patterns of hill climbing and bottom dwelling cows. *Journal of Range Management*, 57, 20–28.

Beyer, H. L. (2004). Hawth's Analysis Tools for ArcGIS. Retrieved November 4, 2006 from http://www.spatialecology.com/htools.

Cain III, J. W., Krausman, P. R., Jansen, B. D., & Morgart, J. R. (2005). Influence of topography and GPS fix interval on GPS collar performance. *Wildlife Society Bulletin*, 33, 926–934.

DeCesare, N. J., Squires, J. R., & Kolbe, J. A. (2005). Effect of forest canopy on GPS-based movement data. *Wildlife Society Bulletin*, 33, 935–941.

D'Eon, R. G., & Delparte, D. (2005). Effects of radio-collar position and orientation on GPS radio-collar performance, and the implications of PDOP in data screening. *Journal of Applied Ecology*, 42, 383–388.

D'Eon, R. G., Serrouya, R., Smith, G., & Kochanny, C. O. (2002). GPS radiotelemetry error and bias in mountainous terrain. *Wildlife Society Bulletin*, 30, 430–439.

Di Orio, A. P., Callas, R., & Schaefer, R. J. (2003). Performance of two GPS telemetry collars under different habitat conditions. *Wildlife Society Bulletin*, 31, 372–379.

Frair, J. L., Nielsen, S. E., Merrill, E. H., Lele, S. R., Boyce, M. S., Munro, R. H. M., Stenhouse, G. B., & Beyer, H. L. (2004). Removing GPS collar bias in habitat selection studies. *Journal of Applied Ecology*, 41, 201–212.

Hansen, M. C., & Riggs, R. A. (2008). Accuracy, precision, and observation rates of global positioning system telemetry collars. *The Journal of Wildlife Management*, 72, 518–526.

Jerde, C. L., & Visscher, D. R. (2005). GPS measurement error influences on movement model parameterization. *Ecological Applications*, 15, 806–810.

Kumpula, J., & Colpaert, A. (2007). Snow conditions and usability value of pasture land for semi-domesticated reindeer (*Rangifer tarandus tarandus*) in northern boreal forest area. *Rangifer*, 27, 25–39.

Lewis, J. S., Rachlow, J. L., Garton, E. O., & Vierling, L. A. (2007). Effects of habitat on GPS collar performance: using data screening to reduce location error. *Journal of Applied Ecology*, 44, 663–671.

Rodgers, A. R., & Carr, A. P. (2002). HRE: The Home Range Extension for ArcView™ (Beta Test Version 0.9, July 1998) – User's Manual. Retrieved April 15, 2009 from http://blue.lakeheadu.ca/hre/. The manual is also available at http://www.blueskytelemetry.com/downloads.asp.

Rodgers, A. R., Carr, A. P., Beyer, H. L., Smith, L., & J. G. Kie, J. G. (2007). HRT: Home Range Tools for ArcGIS. Retrieved April 15, 2009 from http://blue.lakeheadu.ca/hre/. The tools are also available at http://www.blueskytelemetry.com/downloads.asp.

Televilt (2006). *Tellus GPS System - User Manual*. Followit Lindesberg AB (former Televilt, TVP Positioning AB), Sweden.

Turner, L. W., Udal, M. C., Larson, B. T., & Shearer, S. A. (2000). Monitoring cattle behaviour and pasture use with GPS and GIS. *Canadian Journal of Animal Science*, 80, 405–413.