

Leveraging location data of variable quality to reconstruct animal movements: application to a reintroduced island fox (*Urocyon littoralis*) population

Riley O. Mummah

University of California, Los Angeles

Angela H. Guglielmino

University of California, Los Angeles

Benny Borremans

University of California, Los Angeles

Timothy J. Coonan

National Park Service, Channel Islands National Park

K. C. Prager

University of California, Los Angeles

James O. Lloyd-Smith (✉ jlloydsmith@ucla.edu)

University of California, Los Angeles

Method Article

Keywords:

Posted Date: May 11th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1627609/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Despite significant advances in statistical approaches and data collection for analyzing wildlife movements over the last 50 years, there are limited analytical frameworks to be applied when spatial data are collected for purposes other than analyzing movement. Data collected for other purposes (e.g., sporadic captures or survival checks via telemetry) generally have lower temporal frequency or spatial precision than data collected to analyze fine-scale animal movement. The coarseness of the former renders them poorly suited for analysis using existing statistical tools.

Methods: We propose a new way of estimating animal movement trajectories by integrating variable quality location data – including frequent but spatially coarse, irregularly-shaped polygon location data arising from VHF telemetry as well as less frequent, more spatially precise location data – using functional data analysis combined with a spatial resampling algorithm. We apply this method to analyze location data from the reintroduced Channel Island fox (*Urocyon littoralis*) subspecies population on Santa Rosa Island, California, which were collected from 2003-2012 for purposes other than movement analyses but provide an ideal case study to develop and test these novel methods.

Results: By combining coarse-grained location knowledge, obtained through field notes and expert interpretation, with more precise location data, we reconstructed individual animal movement trajectories and demonstrated the utility of combining such data. Through the population ensemble of reconstructed trajectories, we learned that captive-born Channel Island foxes exhibited stronger seasonal movements than wild-born foxes, and most long-range movements occurred within the first two years of a fox's time on the island.

Conclusions: This methodology capitalizes on frequently overlooked but valuable spatial data, often found in field notes and expert knowledge, to reconstruct animal movements. Our approach has wide application to systems in which data of variable quality are collected for purposes other than studying movement and could have beneficial applications in species conservation, landscape ecology, disease management, and population monitoring.

Background

With the advancements in tracking technology, statistical approaches have evolved in parallel to analyze the generated datasets (1–3). The first formal method to analyze animal location data estimated home ranges using minimum convex polygons (4). Kernel-based methods, which use the density and associated error of location observations, followed and were some of the first to enable parametric estimation of space use, home ranges, and utilization distributions (4–7). Work by Getz et al. (8, 9) expanded these kernel-based methods into a nonparametric framework using local nearest-neighbor convex hulls (LoCoH), which more readily account for hard landscape boundaries common in real-world systems. They additionally expanded the approach to account for measurements through time, allowing the estimation of trajectories, or the path of an animal's movements. Additional methods such as

Brownian bridge modeling or continuous-time discrete-space (CTDS) models have been developed in recent years to capitalize on the availability of high-resolution data to estimate animal movement trajectories (10, 11).

Despite the many advances in animal tracking technology and methods for analysis of the resulting datasets, how to incorporate variable quality location data into movement analyses has been left unexplored (12). The spatial and temporal resolutions of location data collected in field studies vary significantly across species, technology used, and purpose of study. When research studies are focused specifically on animal movements, precise and frequent observations are typically collected, often leveraging high-resolution technologies such as GPS telemetry. However, when location data are collected more opportunistically or without the central purpose of studying animal movements (e.g., for demographic studies or population monitoring, or via trapping for health checks), the data often have lower temporal and/or spatial resolution (13–15). Furthermore, limited resources and logistical constraints can reduce the frequency and regularity of location data collection. Such datasets, with less precise and less frequent location data on individual animals, are often disregarded in formal movement analyses because they are not well suited for analysis via existing methodological approaches.

Low-resolution data can present further unappreciated opportunities to augment higher-resolution location datasets, extending the timeframe of an analysis or filling in data gaps, with imprecise data that are often deemed unsuitable for spatial analysis. For instance, field notes often contain a wealth of unused information on the locations of animals that only an expert on the system or field site would be able to interpret. These may be verbal descriptions of locations or of regions identified using a directional VHF antenna. By harnessing expert knowledge on the system and landscape characteristics, such notes could be converted into discrete areas within which the animal was located (i.e., not specific point locations or circular buffers, but “polygon” areas defined by landscape features), and such data could be useful in reconstructing animal movement patterns when higher-resolution data are sparse or unavailable.

The integration of these unconventional “polygons” with traditional (but perhaps temporally infrequent) location data presents a new frontier in estimating movement trajectories. Most methods used to estimate animal movement trajectories or home ranges rely on frequent, relatively precise location data in the form of specific coordinates collected via GPS or telemetry triangulation. There are currently no methods to accommodate imprecise “polygon” data or a mix of the two data types (point data and “polygon” data). Further challenges arise in estimating animal movement trajectories from such integrated data while simultaneously addressing irregular or sparse observations.

Recent work by Buderman et al. (16, 17) addressed a component of the challenge of analyzing telemetry datasets collected for purposes other than fine-scale movement analyses by applying a functional data analysis approach to estimate properties of movement behavior of reintroduced Canada lynx (*Lynx canadensis*) in Colorado. Functional data analysis uses basis functions, such as splines, to interpolate the locations between observation times, smoothing through the imprecise location data to estimate and

visualize the approximate trajectories of the animal. Buderman et al. (16, 17) presented two types of spatial data with differing error structures: Argos and GPS. Argos data have a range of seven error categories, in which GPS data correspond to the most precise Argos category. Argos data are effectively point locations with varying errors (i.e., the locations with uncertainties are circular or elliptical, not irregular shapes). To incorporate these data types, they parametrically modelled the error structures of each data type and incorporated them into a Markov Chain Monte Carlo (MCMC) spline fitting algorithm, which allowed them to infer movement behaviors from their estimated trajectories. Although Buderman et al. (16) utilized temporally suboptimal data that were collected for purposes other than movement analyses, their study did not address the challenges presented by much sparser temporal sampling common in many opportunistic location data sets, nor did it consider “polygon” data or the irregular error structure associated with these data.

We propose a new way of estimating animal movement trajectories using irregularly shaped polygon location data as well as more spatially precise location data, both characterized by infrequent and irregular temporal sampling, by combining functional data analysis with a spatial resampling algorithm. We apply this approach to location data from the reintroduced Channel Island fox (*Urocyon littoralis*) population on Santa Rosa Island, California, which were collected at great effort and expense for the purpose of population monitoring and provide an opportunity to demonstrate these novel methods for movement analysis in a unique ecological context.

Case study: Island fox reintroduction on Santa Rosa Island, California Channel Islands

The Channel Island fox is the smallest north American canid and is endemic to six of the eight Channel Islands off the coast of Southern California (18). They are crepuscular and omnivorous, mainly subsisting on invertebrates, fruits, and rodents. They lack natural predators and therefore typically have high survival and a docile nature (19). Island foxes generally form monogamous pairs and establish territories that they inhabit year-round, with little overlap with territories of adjacent pairs (20). Females typically give birth between April and May to litters of two to three kits on average. The offspring usually disperse in the fall to find a mate and establish their own ranges (19–21).

In the 1990s, the Channel Island fox population on Santa Rosa Island (SRI) and two neighboring islands experienced severe population declines due to a cascade of anthropogenic factors and were recognized federally as critically endangered (22). In 2000, the National Park Service (NPS) initiated an on-island captive breeding program on SRI, bringing the surviving 14 foxes into captivity which left the population temporarily extinct in the wild (19). Reintroduction efforts began in 2003, and at the conclusion of the breeding program in 2009, 96 captive foxes had been released into the wild. The reintroduced population continued to recover, and the fox population was delisted as an endangered species in 2016 (23). This study analyzes location data spanning 2003–2012, enabling us to characterize the movement behaviors during the initial growth phase of the reintroduced population.

Methods

Spatial data collection

Fox trapping

The NPS has conducted annual fox trapping since 2004 to monitor the reintroduced population on SRI. From 2004 to 2008 while the wild population was still small, target trapping was conducted on an as-needed basis in locations known to be occupied by foxes. In 2009, the NPS began a structured 18-ladder grid trapping program on SRI to estimate the annual fox population size and density (Figure S1). Each of the ladder grids was run for 6 consecutive nights in July or August each year. Additional target trapping was conducted from July to January to place radio collars, for targeted sample collection, and to achieve a target vaccination goal not achieved through the grids. Trapping during the rest of the year was not possible due to the rainy season (winter) and birthing season (spring).

Fox telemetry

From 2003 to 2006, all reintroduced and wild-born foxes were collared, PIT-tagged, and tracked with VHF radio telemetry collars (Table S1). Locations were collected weekly via VHF, although the resolution on these locations could be quite poor at times since the primary purpose of these weekly fixes was to detect mortalities. Additional locations were obtained through opportunistic detections. From 2007 onward, only a subset of the wild population (40–50 individuals) was collared at any time due to the growing population size (24). After the conclusion of the captive breeding program in 2009, telemetry efforts shifted from weekly to every other week observations, coinciding with the shift to the grid trapping system.

Precise GPS coordinates were obtained for collared foxes at the site of reintroduction, trap capture, carcass collection, and via direct visual confirmation. Location data with GPS-like precision were also generated via triangulation telemetry and estimated using the program Locate II (25). For the ensuing analyses, triangulation telemetry and precise GPS coordinates will both be considered “GPS data.” The difference in precision between these two data types is negligible since we apply buffers to account for behavioral blurring below.

The less precise telemetry data were generated via single measurements with directional “Yagi” antennas or nondirectional “Omni” antennas, which were both used for routine telemetry detections. During telemetry surveillance when a collar signal was detected but the fox was not seen or triangulated, the fox PIT tag, date, and a verbal description of the general area of the signal were recorded. If the received radio signal was very strong or came from a well-defined landscape feature such as a canyon, the name of the feature was recorded. If the signal was heard with the directional antenna, the location at which the signal was detected and the direction from which it came were recorded. However, when the nondirectional antenna was used, only the location at which the signal was detected was recorded. These

methods resulted in a written list of general locations, landscape features, and directions for individual foxes, rather than a set of GPS coordinates.

Digitizing and filtering field data

To convert the written telemetry locations into a format compatible with geospatial analysis, the recorded descriptions were translated by a team member (A.H.G.) who worked as a fox biologist on SRI for more than 10 years and collected much of the raw data. Each unique place description ($n = 3,812$) was converted to a spatial polygon, mapped using the 'add polygon' tool in Google Earth Pro (26), and saved as a .kml file (Fig. 1A). For subsequent analyses, each polygon shapefile (.kml) was imported and processed in R version 4.1.0 using the *rgdal* package (27, 28).

Directional field notes were mapped in Google Earth Pro and converted to spatial polygons in R (A). These spatial polygons were integrated into each fox's sequence of location observations, resulting in a mix of GPS point and polygon data (B). To convert all data forms to point data, bootstrapped datasets were created by randomly sampling each polygon area and sampling GPS datapoints from a 1km^2 buffer area (shown here in green) on each observation date (C). For each animal, smoothing splines were fit to the easting and northing data separately (D). The median estimates in each coordinate direction were then paired by date to estimate an individual's movement through time on the island landscape as shown by the solid black lines (E). The light grey lines in panels D and E illustrate the spline fits to each resampled dataset, which are additionally summarized by a central 95% envelope (dashed black lines). This schematic was created with simulated data and does not represent the true movements of any individual.

Dataset filtering

To investigate the fundamental challenge of studying movement from a mixed spatial resolution dataset, we focused on data in the early period of fox reintroduction between November 2003 and March 2007. During this period, both GPS and polygon location data were collected intensively on a high proportion of individuals, making this dataset ideal for the application and testing of the method. Sixty-one (61) individual foxes had location data prior to March 2007. All location data for these 61 individuals prior to 2012 were included in the dataset. Any individual time series of locations that contained a gap of more than 365 days was split into two time series, linked by the individual PIT tag, to exclude the gap from trajectory estimation. Furthermore, each individual fox (or segment of the time series if it was split) was required to have more than 10 observations total to be included. Seven foxes had multiple captures and re-releases to the island. For these, the longest contiguous set of observation dates in the wild was selected for analysis.

Polygon filtering

Among the 2,130 unique polygon locations in our dataset, the area of fox location polygons varies drastically from less than 1km^2 to the entire area of the island (214km^2 ; Figure S2A). To incorporate these polygons into a movement analysis, they must be at a resolution that yields useful information about the location of the fox, while not adding too much noise. For example, polygons that are the same size as the

island add no additional information on location, and only serve as confirmation that the fox was still alive. We excluded polygons with areas over 2km^2 (Figure S2B), which is reported as the size of the fox home range (19). The length-to-width ratio was also used to filter out imprecise and uninformative polygons (Figure S2C), applying a maximum length-to-width ratio of 7.29 (a long, thin polygon) across all polygons. To filter the dataset to more elliptical polygons, which would mirror the typical error structure of location data, and exclude only the most extreme shapes, we only included polygons with a length-to-width ratio less than 5.

Resampling algorithm to integrate spatial data types

Our dataset contained both precise GPS or triangulated locations (termed 'GPS data') and less precise and irregularly shaped polygon location data (Fig. 1B). We needed to translate all the data into a single standardized type that was compatible with appropriate statistical approaches while accounting for the shapes and spatial uncertainties of the original data. To do this, we converted all GPS data points into circular 1km^2 buffers (clipped to the island coastline) centered on the original GPS locations to transform these data into areas like the polygons. These converted GPS areas are greater than the spatial uncertainty of the instrument but were intended to represent the spatial scale of daily fox movements for foraging and other purposes (29, 30).

We used a resampling approach to convert these polygons and buffers into an ensemble of synthetic point datasets. A location was uniformly, randomly sampled from within the area of the polygon or buffer, which allowed the full extent of the area to be incorporated into the movement estimation. Each iteration of this bootstrap (for every individual fox) leads to an ensemble of datasets with resampled point locations from within each observed location area on its observation date (Fig. 1C).

Estimate movement through time

Functional data analysis

Functional data analysis uses smoothing functions called splines to interpolate data. There are many varieties of splines. Buderman et al. (16) used a set of B-splines, which forms a linear combination of basis functions. However, this type of spline cannot be fit over large gaps between observations, so it is poorly suited to our data due to the irregularity of our sampling. Instead, we employed a more flexible approach using a smoothing spline (R package *npreg*) with the smoothing parameter (λ) set to 0.1, to ensure consistent curvature across individuals.

We fit a smoothing spline to each direction (easting and northing) of the resampled location data for each date within the interval of time an individual was located on the island (Fig. 1D). For each fox, 100 resampled datasets were created and fitted, giving rise to an ensemble of fitted splines. The fits were then summarized by taking the daily median location with a 95th percentile envelope of uncertainty computed in each coordinate direction (easting and northing). The ensemble of fitted splines and the estimated daily median locations were then paired based on date and visualized as a trajectory (Fig. 1E). We show the easting and northing plots to visualize uncertainties but do not illustrate the 95% uncertainty

envelope on the trajectory map for visual clarity. All data processing and analysis was performed in R version 4.1.0 (28). The fitting of the model to each resampled dataset was performed using the R package *npreg*.

Effects of temporal frequency of location data

To investigate the effects of sampling frequency, we used a fox with data collected twice per week on average (fox 53313, $n = 187$ location data points; Fig. 2) and subsampled its full dataset to target four mean frequency levels: once per week ($n = 94$), every other week ($n = 44$), once per month ($n = 22$), and every other month ($n = 10$). The spline-based interpolation was applied to each subsampled dataset for comparison.

Population-level movement

To demonstrate the potential for our reconstructed trajectories to provide critical insights into population movement ecology studies, we calculated the median location (easting and northing pair) for every month each individual fox was located on the island. Then we took the difference between the median estimates for consecutive months to estimate the distance traveled through time. These estimates were summarized across the population to explore broad patterns in movement behavior during the early years of the reintroduction program.

Results

Integrated spatial dataset

Inclusion of the polygon data increased our dataset by roughly 250%, giving rise to a total dataset comprised of 1,700 GPS locations (Table S2) and 2,775 additional locations corresponding to 493 unique polygons (Figure S3) for 61 foxes between November 2003 and October 2012. The contribution of polygon data, as a proportion of all available data for each individual, ranged from 0–88% (Figure S4A); only five individuals had no polygon data available during this period. With the inclusion of polygon data, the median interval between observations of an individual's time series was 15 days, and there were two observations per month on average (Figure S4B, C). Without the polygon data, the median interval between observations and the average number of observations per month would have been 31 days and one observation, respectively. Additionally, using GPS data alone, thirteen individuals would have been excluded by our threshold of requiring at least 10 observations.

Spline-based movement trajectory estimation

For each fox, we estimated movement trajectories by fitting smoothing splines independently to both easting (Fig. 2A) and northing (Fig. 2B) coordinate directions of resampled locations, resulting in two-dimensional (spatial) trajectories after combining the two by observation date. These trajectories are shown on a map of the island (Fig. 2C), where the line represents the median movement trajectory of an individual through time. The 95% envelope represents the bounds of the central 95% of trajectories (Fig. 2;

dashed lines) and should be viewed more like a standard error than a standard deviation as it depicts uncertainty in the estimate of the median. The bounds of the 95% envelope are well within a reasonable range ($\sim 1 \text{ km}^2$) for typical daily movements, but many data points appear to fall outside these bounds. One factor contributing to this effect is that the daily movements of foxes can occasionally span a region greater than the uncertainty on these reconstructed median locations. Another factor is that the polygon data points are resampled in the algorithm using the full polygon area but are visualized in Fig. 2 using the coordinates of the polygon centroid, which means they may not fall within the envelope. It is also possible that data points were inaccurately recorded in the field. By its nature the spline fitting smooths through the locations and balances the residuals, which causes the median and center 95% envelope to appear as an average, running between the datapoints.

Over the 18-month period described in Fig. 2, fox 53313 started in the northwest area of the island near Arlington Canyon, then moved northeast to the ranch area of the island near the NPS headquarters. Ten months later the fox shifted into Verde Canyon, spent four months there, and then moved into the Cherry Canyon/Lower Water Canyon region. Other case study foxes (A7954 and 73D0D; Figure S5-6) are shown in the supplement and reveal similar resolution of fox movements over periods of one to two years.

Smoothing splines were fit to the easting (A) and northing (B) coordinates separately and then coordinates were paired by date to construct the movement trajectory on the island (C). Each plot shows the GPS (blue) and polygon (pink) locations on the date they were recorded. Polygon data points represent the coordinates of the polygon centroids, but the fitting was done with resampled locations (not shown). Individual spline fits are shown by the light grey lines. The solid black line denotes the median value across the 100 bootstrap fits, and the central 95% interval is illustrated by the dashed black lines.

Contributions of the different spatial data types

To explore the impact of our method's ability to include polygon data in our movement trajectory estimates, we examined three examples that illustrate when incorporating polygon data may add meaningful information. These three cases are illustrative of two key dimensions of the analysis: the proportion of polygon data to GPS data in a given dataset and the temporal resolution of high-quality location data (Table S3). For the first dimension, individual foxes may have relatively few or many polygon-based observations, and the abundance of polygons compared to GPS points is likely to influence the relative value of these data. The second dimension (temporal resolution) represents how often a fox is being observed through time. Higher temporal resolution means the fox was observed more often, which will improve the quality of movement trajectory estimation. We expected the inclusion of polygon data to matter most when the proportion of polygons was high or when the temporal frequency of GPS data was low.

The inclusion of polygon data has a major effect on trajectory reconstruction when a fox has few GPS points, as demonstrated by fox F3D2F (Fig. 3). Over an eleven-month span, this individual only has twelve GPS locations, which are highly concentrated in the first two months (Fig. 3D-F). If we only used GPS

data, this fox would have low temporal frequency and hence uninformed and inaccurate location estimates for much of the time period, as seen in Fig. 3D-F. By including polygon data, the number of observations increases four-fold, and the inclusion stabilizes the estimate of the trajectory in the second half of this individual's observation interval (Fig. 3A-C). The inclusion of polygon data in this case refines the estimates to a smaller area of the island and fills out the temporal resolution of the data. Indeed, using only the polygon data enables a well-defined reconstruction of the trajectory over much of this interval (Fig. 3G-I). Conversely, including the GPS data extends the temporal range of the data and reveals an earlier period when the individual was in a different region of the island.

The previous example demonstrated the utility of polygon data in enhancing the temporal frequency and range of an individual's dataset. However, when there are large gaps (relative to the full interval of observation time), the smoothing algorithm is forced to fit a curve to the gap with no information on where the fox is located. A minor case of this is seen in the GPS-only spline fits in Fig. 3F, where there is clearly too much curvature added to the trajectory across temporal gaps in the dataset. In extreme cases, this situation can lead to a breakdown of the method and artefacts arising from the spline fitting. A more extreme example is shown for fox 73D0D, where even in the best-case scenario when both data types are combined, there is a 6-month gap in the time series of locations (Figure S6A-C). The spline fits the existing data with a big quadratic curve over this period, and the trajectory of the fox is projected into the ocean. Across all three divisions of data type for fox 73D0D, the temporal irregularity of the data drives large variation in the fits, showing that, in some cases, smaller datasets spanning shorter time periods can have fewer problematic outcomes since the spline has so little to work with (e.g., Figure S6F is a better fit than Figure S6C). This example illustrates the general importance of temporal regularity in the observations, independent of spatial resolution.

The integration of polygon data to a GPS data-poor dataset has the greatest impact in narrowing the trajectory estimates. The first column (A, B, C) shows the trajectory estimate using all available data. The second (D, E, F) and third (G, H, I) columns show the estimated trajectories with GPS and polygon data, respectively. The first row (A, D, G) shows the fit of the easting coordinate. The second row (B, E, H) shows the spline fits for the northing coordinate, and the bottom row (C, F, I) shows the trajectory on the map of Santa Rosa Island. GPS data and the centroids of the polygons are illustrated by the blue and pink points, respectively. The grey lines signify the spline fits for the 100 resampled datasets, which are summarized by the median trajectory (solid black line) and the 95% uncertainty envelope (dashed black lines).

Effects of temporal frequency of location data

To further investigate the effects of data quantity and temporal frequency, we used our model fox from Fig. 2 (fox 53313) and subsampled its full dataset (which had an average of about 2 observations per week) to target four mean frequencies of location data: once per week, every other week, once per month, and every other month (Fig. 4). As the observations become less frequent, the level of variation in the trajectories increases conspicuously. The trajectories for observations once per week look quite similar to the trajectory estimated from the full dataset (Fig. 4; compare colored lines to black line). Trajectories estimated from observations taken every other week, once per month, or every other month show

increasing variation and some extreme extrapolations due to the smoothing spline's estimation over more gaps with less data (Fig. 4). These impressions from the mapped reconstructions are borne out by analyzing residuals of the trajectories relative to the median trajectory estimated from the full dataset (Figure S7). Importantly, though, even these trajectories built from less frequent observations still identify the correct region of the location of the fox. When the data is sampled down to every other month ($n = 10$ observations), the model fit yields limited information on the precise trajectory of the fox but may be more useful in estimating its home range.

The full location dataset for fox 53313 was subsampled to produce four observation frequencies: once per week, every other week, once per month, and every other month (as shown by the four panels). Each down-sampling was performed 10 times for each observation frequency; the resulting median trajectories from the full algorithm are shown by different colored lines in each panel. The median trajectory for the full dataset is shown in black.

Movement patterns in the reintroduced fox population

To demonstrate the research opportunities created by our new method, we examined the population-wide movement patterns in the island fox population in the early years of the reintroduction program. For each fox, we calculated the difference in the median location estimate between consecutive months to capture trends in the distance traveled through time. The majority of movements between consecutive months is short-range ($< 0.5\text{km}$; Fig. 5A) but movements up to 10km are observed. Movements $> 2\text{km}$ are seen in all months of the year. The captive-born individuals exhibit stronger seasonal movements, with longer-range movements occurring in the fall and winter, corresponding with the time of release to captivity, and shorter-range movements in the spring and summer (Fig. 5A). Wild-born individuals exhibit near constant short-range movements ($\sim 0.5\text{km}$) on average (Fig. 5A), but when we zoom in on wild-born individuals only and stratify by their time on the island (12 months and > 12 months), we see that wild-born individuals move nearly three times further in their first year on the landscape than in subsequent years, with longer-range movements more prevalent in their first fall and winter (Fig. 5C).

The distance traveled by month was calculated as the difference between median locations of consecutive months for each fox. In Panel A, the box and whisker plots illustrate the distribution of movements for each month and highlight the long-range movements (points) which are greater than the sum of the 75th percentile and 1.5 times the interquartile range. The lines denote the average distance traveled by month for captive-born ($n = 33$) and wild-born ($n = 28$) foxes. Panel B depicts the median distance traveled from time of first location on the island (either release from captivity or first capture in the wild). The semi-transparent bands denote the central 95% interval for the distribution of distance traveled. Outliers beyond the 95% interval are shown by the points. For panels A and B, captive-born foxes are shown in brown, and wild-born foxes in teal. Panel C, like Panel A, illustrates the distribution of movements for each month and highlights long-range movements (points). This panel only shows wild-born fox movements and is stratified by time on the island. The first 12 months of wild-born individuals is shown in orange. Time past 12 months is shown in purple. The average distance traveled per month for this time division is illustrated by the lines and the left-side axis.

We also examined the timing of movement behaviors relative to when foxes were first located (where the first location is defined by either release from captivity or a fox's first capture in the wild; Fig. 5B). When we align the data based on each fox's first location and follow their movement through time, we see that the median distance traveled is highest within each individual's first 24 months on the island (Fig. 5B). After two years, the movements for these foxes dwindle to less than 0.5 km.

Discussion

Motivated by a wealth of unconventional location data from reintroduced Channel Island foxes, we developed a novel approach to integrating spatial data types of varying quality to estimate animal movement through time. We showed that incorporating imprecise polygon data, derived from field notes and expert interpretation, can spatially enhance and temporally extend animal location datasets and improve estimation of movement trajectories. In our island fox case study, the inclusion of polygon data more than doubled the number of location points available for the 61 study individuals and reduced the variation in the fits of the trajectories, especially for individuals that had few precise location estimates. This method has broad potential applications to long-term monitoring datasets that were not collected with the explicit aim of studying movement.

With our newly developed algorithm, we reconstructed population-level movement summaries to better understand the initial movement of foxes after reintroduction. By stratifying by birth status (captive vs. wild), we examined the effects of reintroduction on broad behavioral movement patterns in the fox population. Captive-born foxes, which by definition had never roamed the island, exhibited a strong seasonal signal, moving longer distances in fall and winter and moving shorter distances in spring and summer. One direct explanation for this pattern is that most captive-born foxes were released in the fall (Sept-Nov), and it is natural to see long-range movements by individuals recently introduced to a novel landscape. Captive-born foxes may have explored the island more widely after release from captivity before establishing a home range, since there was little competition for space. In contrast, wild-born foxes had more stable movement behavior and showed a much more subtle seasonal signal. When stratified by time on the island, we found that wild-born individuals move longer distances in their first year and then move much less in subsequent years. There is a seasonal signal in the first twelve months of wild-born movement, similar to the captive-born movement pattern, but more subtle. The wild-born seasonal pattern aligns broadly with known island fox ecology: fox births occur in dens in the mid-to-late spring, leading to a period of limited long-range movement while pups are reared; juvenile dispersal occurs in the fall, consistent with the longer-range movement pattern we see at that time of year (20, 22, 30).

Our analyses demonstrate the interacting roles of spatial and temporal resolution in enabling movement reconstruction and highlight that the quality of the reconstruction depends primarily on temporal frequency and regularity of location data. Even when an individual's time series consists primarily of relatively imprecise polygon data (i.e., Fox F3D2F; Fig. 3), the algorithm can yield a similar median trajectory if the data are abundant, with no large temporal gaps. In our case study, polygon data filled large gaps in the time series of GPS data, delivering major improvements to the reconstruction. Data

quantity and temporal regularity are related, but data quantity does not imply temporal regularity as is seen with Fox 73D0D (Figure S6). When incorporating polygon data increases the amount of data available but concentrates the data in time, leaving long gaps, the smoothing spline algorithm can give problematic results, as seen by the fox movement trajectory projected into the ocean (Figure S6). In sum, the quality of the movement reconstruction is determined by both spatial and temporal resolution of the data, but good temporal coverage is essential.

We have demonstrated the considerable benefit of including polygon data in trajectory estimation, but initial translation of the recorded location descriptions into digital polygons was onerous and time-consuming. In future studies, polygon creation could be streamlined by having pre-defined areas on the island that the field notes map onto, rather than translating each phrase into its own precise polygon. These pre-defined areas could be delineated by topographical features much as our polygons were. Some degree of spatial resolution could be lost by this method, but the tradeoff of greater data availability could compensate for this loss. There is also scope to apply artificial intelligence methods, including natural language processing, whether trained by an expert or untrained and working from landscape features, to identify polygons. Further, although the inclusion of polygons added crucial information to our analysis, our uniform sampling of the polygon areas did not always match our beliefs and our expert knowledge of the fox's position. If the location of a fox was concentrated at one end of a long canyon, this positioning would not be captured by our current algorithm formulation due to uniform sampling of the area. Filtering the polygon areas and length-to-width ratios was a first step to managing uninformative polygons, but weighting the area of the polygon based on landscape characteristics or nearby GPS location data has the potential to further improve positional estimates.

There is great opportunity for data imputation methods to be built into modern spatial statistical techniques. Spatial data resampling methods could lend great strength and flexibility to statistical analysis by accommodating nearly any data type and facilitating the link of polygon locations to all spatial methods. A useful extension of this study would be to integrate the resampling of the polygons (to account for their shape/error) into the Bayesian framework developed by Buderman et al. (16). Their analysis assumed that the spatial dataset is fixed and described the error distributions of their observations parametrically. To integrate the polygon data into their framework, the spatial data could be resampled during MCMC estimation of the spline. Our resampling framework could be incorporated more broadly into other spatial and movement analyses and provides an avenue to integrate polygon-type location data into other spatial methods, as long as the data have sufficient temporal frequency.

Our algorithm utilized smoothing splines to interpolate location data through space and time. Smoothing splines offer a flexible approach to data interpolation via tuning a single smoothing parameter either manually or through a generalized cross-validated procedure. For our analysis, we fixed the smoothing parameter to ensure that the curvature was consistent across individuals, rather than fitting it by various procedures (e.g., generalized cross-validation, maximum likelihood, Akaike information criterion). We selected a smoothing parameter value of 0.1 (on a scale of 0 to 1), balancing several considerations. If we had chosen a larger smoothing parameter, the trajectories would have become less curved and more

like connected lines between the observation points. This would lead to a lower degree of averaging across the locations, which could capture the precision of the high-resolution locations more confidently but would also put more confidence in the lower-quality polygon locations than desired, given that our uniform resampling algorithm can generate apparent locations anywhere within a polygon. Our choice of a low smoothing parameter enables the method to smooth out these errant points, but at the cost of occasionally failing to capture real, short-term, long-distance movements, plus the risk of overshooting the data when there are temporal gaps. Depending on the nature of the system and dataset and the goals of the analysis (e.g., home range estimation or construction of a contact network), more or less curvature may be optimal, but further study is required to uncover how the temporal and spatial resolution of the dataset influences the optimal choice of smoothing parameter.

The choice of spline can also impact the results and interpretation of the trajectories. We opted to use a smoothing spline to fit the coordinates since it could robustly accommodate the patterns in our data, but a more complex spline could add more flexibility to the fitting and may be appropriate for other systems where individuals make frequent, meaningful movements that are important to capture. An alternative B-spline framework would require the data to be more regular through time and is much more sensitive to temporal gaps in the data. Thus, the algorithm would fail for individuals with insufficient temporal regularity or data quantity. The smoothing spline we chose converges for all individuals but can give biologically dubious results when data are insufficient. Additional study is needed on the influence of data quantity and resolution on the choice of spline and associated parameters to improve the robustness of our approach to non-ideal datasets.

In our study of a terrestrial species on an oceanic island, the island boundary introduces a further challenge to movement estimation. The island coastline is a hard ecosystem boundary for the foxes, but such boundaries have only been addressed in a few estimation techniques (e.g., LoCoH and special implementations of kernel density estimation) (5, 8). Most other studies address this challenge by clipping the estimated home ranges or trajectories at the boundary, which could bias the estimates because it does not redistribute the density appropriately. In our method, the polygon resampling approach restricts point observations to the allowed region (i.e., the island's terrestrial surface). However, the spline can fit curves that extend beyond the island boundary (as in Figure S6C). A more restrictive spline fitting or explicit approach to reject any splines which project movement into the excluded area could address this shortcoming.

This novel estimation of animal movement trajectories has wide-ranging applicability. It is essential to recognize that, under our current parameterization, the reconstructed trajectories are intended to capture the central tendency of an animal's location in a given time period and will not reconstruct all short-term excursions from the broader path. Nevertheless, our approach generates systematic trajectories that align with expert knowledge, even if some quantitative details are imperfect – and our analysis of 61 foxes shows the potential to recover valuable population-scale insights into movement patterns. Such information, at individual and population scales, has important application to support studies of behavioral ecology, conservation, and disease spread (29–36).

Crucially, our approach enables such insights to be derived from existing and currently unused location data for a wide range of wildlife systems. We have detailed several clear steps that can be taken to further fine-tune our algorithm and improve the precision and accuracy of the estimated trajectories. We hope this work can be a useful first step towards analyzing these valuable data to gain important insights into animal movement ecology, as well as the spatiotemporal aspects of related fields.

Conclusions

Many studies of wildlife populations collect rich, long-term datasets that include location information of varying quality, for instance via monitoring programs with regular trapping or collared sentinel populations. However, to date these valuable datasets have not been usable for movement analyses due to a lack of appropriate, systematic methods to incorporate the irregular data that can occur. Our work opens new opportunities to use these unconventional data and to integrate them with other location data at any level of spatial or temporal resolution to reconstruct animal movement trajectories. The addition of these new data sources, arising from studies that were not originally designed to study movement, will expand the number of systems and time periods for which animal movement can be studied, with benefits for habitat use modelling, social network analyses or analyzing infectious disease transmission.

Abbreviations

GPS - Global positioning system

NPS – National Park Service

SRI – Santa Rosa Island

VHF – Very high frequency

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Strategic Environmental Research and Development Program (SERDP, RC-2635) of the U.S. Department of Defense and the U.S. National Science Foundation (DEB-1557022). The content of the information does not necessarily reflect the position or the policy of the U.S. government, and no official endorsement should be inferred.

Authors' contributions

JOL-S and KCP acquired funding for the study. TJC and AHG conducted the fieldwork. AHG digitized the field notes into polygons. ROM and JOL-S designed the study and developed the algorithm. AHG, BB, and KCP provided feedback on the algorithm. ROM conducted the coding and data analyses and wrote the first draft of the manuscript. All authors contributed to revisions. All authors read and approved the final manuscript.

Acknowledgements

We would sincerely like to thank the entire Santa Rosa Island fox team of the National Park Service for generating the wildlife data for this study. We would also like to thank Ottar Bjørnstad for his useful early suggestions and the Lloyd-Smith lab for their feedback through the development of this project.

References

1. Kie JG, Matthiopoulos J, Fieberg J, Powell RA, Cagnacci F, Mitchell MS, et al. The home-range concept: are traditional estimators still relevant with modern telemetry technology? *Philos Trans R Soc B Biol Sci*. 2010 Jul 27;365(1550):2221–31.
2. Tomkiewicz SM, Fuller MR, Kie JG, Bates KK. Global positioning system and associated technologies in animal behaviour and ecological research. *Philos Trans R Soc B Biol Sci* [Internet]. 2010 Jul 27 [cited 2021 Aug 16]; Available from: <https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2010.0090>
3. Walter WD, Onorato DP, Fischer JW. Is there a single best estimator? Selection of home range estimators using area-under-the-curve. *Mov Ecol*. 2015 Dec;3(1):10.
4. Mohr CO. Table of Equivalent Populations of North American Small Mammals. *Am Midl Nat*. 1947;37(1):223–49.
5. Benhamou S, Cornélis D. Incorporating Movement Behavior and Barriers to Improve Kernel Home Range Space Use Estimates. *J Wildl Manag*. 2010;74(6):1353–60.
6. Bullard F. Estimating the home range of an animal: a Brownian bridge approach. Johns Hopkins University. Master thesis; 1999.

7. Worton BJ. Kernel Methods for Estimating the Utilization Distribution in Home-Range Studies. *Ecology*. 1989;70(1):164–8.
8. Getz WM, Fortmann-Roe S, Cross PC, Lyons AJ, Ryan SJ, Wilmers CC. LoCoH: Nonparametric Kernel Methods for Constructing Home Ranges and Utilization Distributions. Coulson T, editor. *PLoS ONE*. 2007 Feb 14;2(2):e207.
9. Getz WM, Wilmers CC. A local nearest-neighbor convex-hull construction of home ranges and utilization distributions. *Ecography*. 2004;27(4):489–505.
10. Hanks EM, Hooten MB, Alldredge MW. Continuous-time discrete-space models for animal movement. *Ann Appl Stat* [Internet]. 2015 Mar 1 [cited 2021 Aug 18];9(1). Available from: <https://projecteuclid.org/journals/annals-of-applied-statistics/volume-9/issue-1/Continuous-time-discrete-space-models-for-animal-movement/10.1214/14-AOAS803.full>
11. Kranstauber B, Kays R, LaPoint SD, Wikelski M, Safi K. A dynamic Brownian bridge movement model to estimate utilization distributions for heterogeneous animal movement. *J Anim Ecol*. 2012;81(4):738–46.
12. Cagnacci F, Boitani L, Powell RA, Boyce MS. Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges. *Philos Trans R Soc B Biol Sci*. 2010 Jul 27;365(1550):2157–62.
13. Mereu M, Agus B, Addis P, Cabiddu S, Cau A, Follesa MC, et al. Movement estimation of *Octopus vulgaris* Cuvier, 1797 from mark recapture experiment. *J Exp Mar Biol Ecol*. 2015 Sep 1;470:64–9.
14. White GC, Shenk TM. Population Estimation with Radio-Marked Animals. In: *Radio Tracking and Animal Populations* [Internet]. Elsevier; 2001 [cited 2021 Aug 18]. p. 329–50. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B9780124977815500141>
15. Winterstein SR, Pollock KH, Bunck CM. Analysis of Survival Data from Radiotelemetry Studies. In: *Radio Tracking and Animal Populations* [Internet]. Elsevier; 2001 [cited 2021 Aug 18]. p. 351–80. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B9780124977815500153>
16. Buderman FE, Hooten MB, Ivan JS, Shenk TM. A functional model for characterizing long-distance movement behaviour. Matthiopoulos J, editor. *Methods Ecol Evol*. 2016 Mar;7(3):264–73.
17. Buderman FE, Hooten MB, Ivan JS, Shenk TM. Large-scale movement behavior in a reintroduced predator population. *Ecography*. 2018 Jan;41(1):126–39.
18. Moore CM, Collins PW. *Urocyon littoralis*. *Mamm Species*. 1995 Jun 23;(489):1.
19. Coonan TJ, Schwemm CA, Garcelon DK. Decline and recovery of the Island Fox: a case study for population recovery. Cambridge; New York: Cambridge University Press; 2010. 212 p. (Ecology, biodiversity, and conservation).
20. Roemer GW, Smith DA, Garcelon DK, Wayne RK. The behavioural ecology of the island fox (*Urocyon littoralis*). *J Zool*. 2001;255(1):1–14.
21. Crooks KR, van Vuren D. Spatial Organization of the Island Fox (*Urocyon littoralis*) on Santa Cruz Island, California. *J Mammal*. 1996 Aug;77(3):801.

22. Coonan TJ, Schwemm CA, Roemer GW, Garcelon DK, Munson L. Decline of an Island Fox Subspecies to near Extinction. *Southwest Nat.* 2005;50(1):32–41.
23. US Fish and Wildlife Service. Endangered and Threatened Wildlife and Plants; Removing the San Miguel Island Fox, Santa Rosa Island Fox, and Santa Cruz Island Fox from the Federal List of Endangered and Threatened Wildlife, and Reclassifying the Santa Catalina Island Fox from Endangered to Threatened. *Fed Regist.* 2016 Aug 12;81(156):53315–33.
24. Coonan TJ, Bakker V, Hudgens B, Boser CL, Garcelon DK, Morrison SA. ON THE FAST TRACK TO RECOVERY: ISLAND FOXES ON THE NORTHERN CHANNEL ISLANDS. *Monogr West North Am Nat.* 2014;7:373–81.
25. Nams VO. *Locate II*. Truro, Nova Scotia, Canada; 1990.
26. Google Earth Pro. Map showing location of Santa Rosa Island [Internet]. 2018. Available from: earth.google.com/web/
27. Bivand R, Keitt T, Rowlingson B, Pebesma E, Sumner M, Hijmans R, et al. rgdal: Bindings for the “Geospatial” Data Abstraction Library [Internet]. 2021 [cited 2021 Aug 21]. Available from: <https://CRAN.R-project.org/package=rgdal>
28. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2021. Available from: <https://www.R-project.org/>
29. Resnik JR, Andelt WF, Stanley TR, Snow NP. Island fox spatial ecology and implications for management of disease. *J Wildl Manag.* 2018 Aug;82(6):1185–98.
30. Sanchez JN, Hudgens BR. Interactions between density, home range behaviors, and contact rates in the Channel Island fox (*Urocyon littoralis*). *Ecol Evol.* 2015 Jun;5(12):2466–77.
31. Almberg ES, Cross PC, Dobson AP, Smith DW, Hudson PJ. Parasite invasion following host reintroduction: a case study of Yellowstone’s wolves. *Philos Trans R Soc B Biol Sci.* 2012 Oct 19;367(1604):2840–51.
32. Benson JF, Abernathy HN, Sikich JA, Riley SPD. Mountain lions reduce movement, increase efficiency during the Covid-19 shutdown. *Ecol Solut Evid* [Internet]. 2021 Jul [cited 2022 May 4];2(3). Available from: <https://onlinelibrary.wiley.com/doi/10.1002/2688-8319.12093>
33. Haydon DT. Cross-disciplinary demands of multihost pathogens. *J Anim Ecol.* 2008 Nov;77(6):1079–81.
34. Mancy R, Rajeev M, Lugelo A, Bruncker K, Cleaveland S, Ferguson EA, et al. Rabies shows how scale of transmission can enable acute infections to persist at low prevalence. *Science.* 2022 Apr 29;376(6592):512–6.
35. Wagner A. BEHAVIORAL ECOLOGY OF THE STRIPED HYENA (*Hyaena hyaena*). 2022 May;
36. Wheeler DC, Waller LA, Biek R. Spatial analysis of feline immunodeficiency virus infection in cougars. *Spat Spatio-Temporal Epidemiol.* 2010 Jul;1(2–3):151–61.

Figures

Figure 1

Schematic of the data conversion and movement trajectory reconstruction.

Directional field notes were mapped in Google Earth Pro and converted to spatial polygons in R (A). These spatial polygons were integrated into each fox's sequence of location observations, resulting in a mix of GPS point and polygon data (B). To convert all data forms to point data, bootstrapped datasets were created by randomly sampling each polygon area and sampling GPS datapoints from a 1km² buffer area (shown here in green) on each observation date (C). For each animal, smoothing splines were fit to the easting and northing data separately (D). The median estimates in each coordinate direction were then paired by date to estimate an individual's movement through time on the island landscape as shown by the solid black lines (E). The light grey lines in panels D and E illustrate the spline fits to each resampled dataset, which are additionally summarized by a central 95% envelope (dashed black lines). This schematic was created with simulated data and does not represent the true movements of any individual.

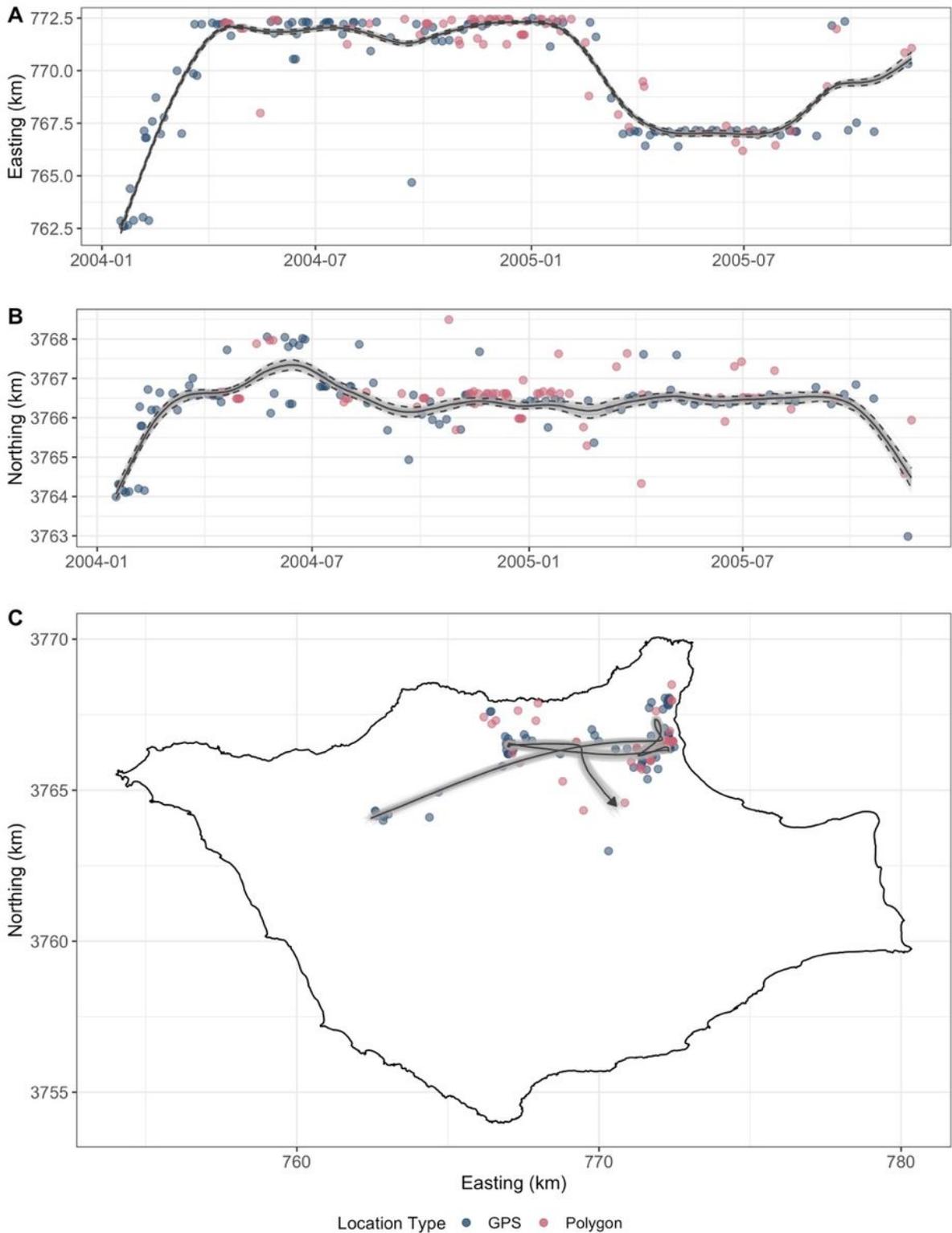


Figure 2

Movement trajectories fit via smoothing splines for fox 53313.

Smoothing splines were fit to the easting (A) and northing (B) coordinates separately and then coordinates were paired by date to construct the movement trajectory on the island (C). Each plot shows the GPS (blue) and polygon (pink) locations on the date they were recorded. Polygon data points

represent the coordinates of the polygon centroids, but the fitting was done with resampled locations (not shown). Individual spline fits are shown by the light grey lines. The solid black line denotes the median value across the 100 bootstrap fits, and the central 95% interval is illustrated by the dashed black lines.

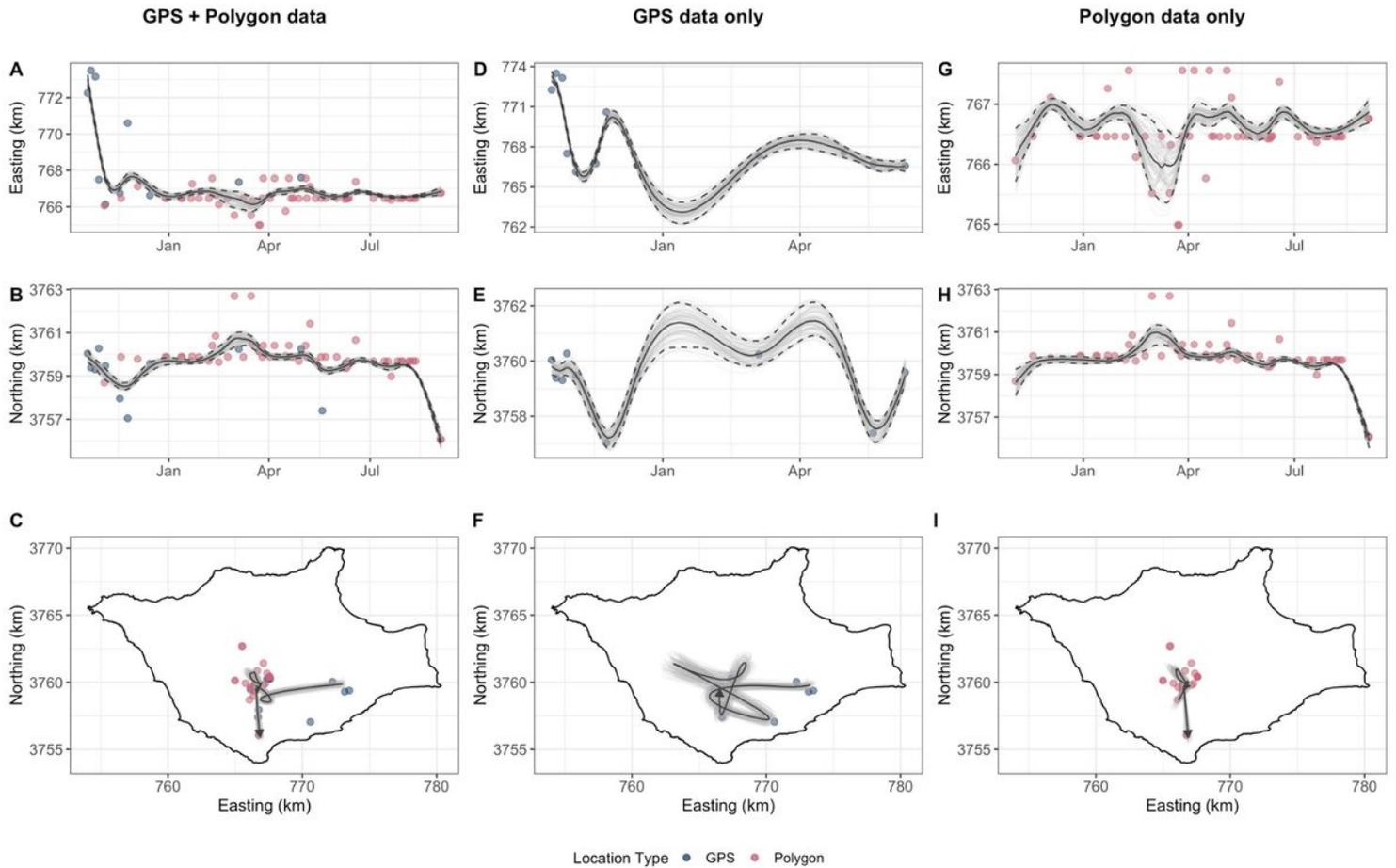


Figure 3

Fitted movement trajectories of fox F3D2F by data type.

The integration of polygon data to a GPS data-poor dataset has the greatest impact in narrowing the trajectory estimates. The first column (A, B, C) shows the trajectory estimate using all available data. The second (D, E, F) and third (G, H, I) columns show the estimated trajectories with GPS and polygon data, respectively. The first row (A, D, G) shows the fit of the easting coordinate. The second row (B, E, H) shows the spline fits for the northing coordinate, and the bottom row (C, F, I) shows the trajectory on the map of Santa Rosa Island. GPS data and the centroids of the polygons are illustrated by the blue and pink points, respectively. The grey lines signify the spline fits for the 100 resampled datasets, which are summarized by the median trajectory (solid black line) and the 95% uncertainty envelope (dashed black lines).

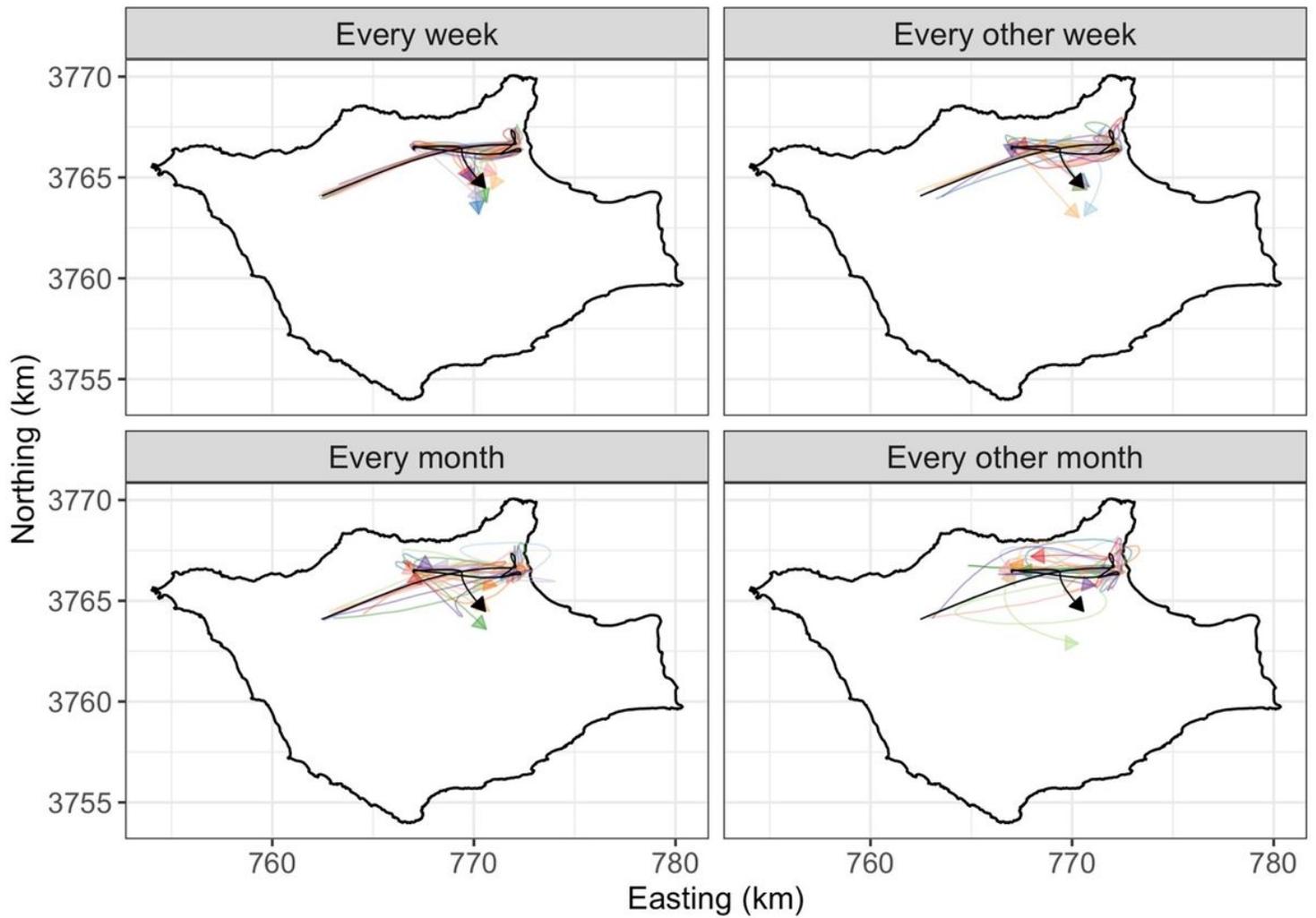


Figure 4

Estimated movement trajectories for down-sampled location datasets of fox 53313.

The full location dataset for fox 53313 was subsampled to produce four observation frequencies: once per week, every other week, once per month, and every other month (as shown by the four panels). Each down-sampling was performed 10 times for each observation frequency; the resulting median trajectories from the full algorithm are shown by different colored lines in each panel. The median trajectory for the full dataset is shown in black.

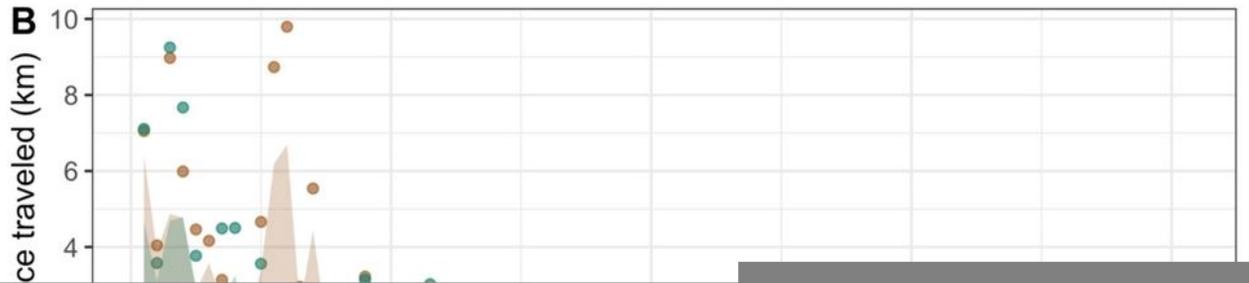
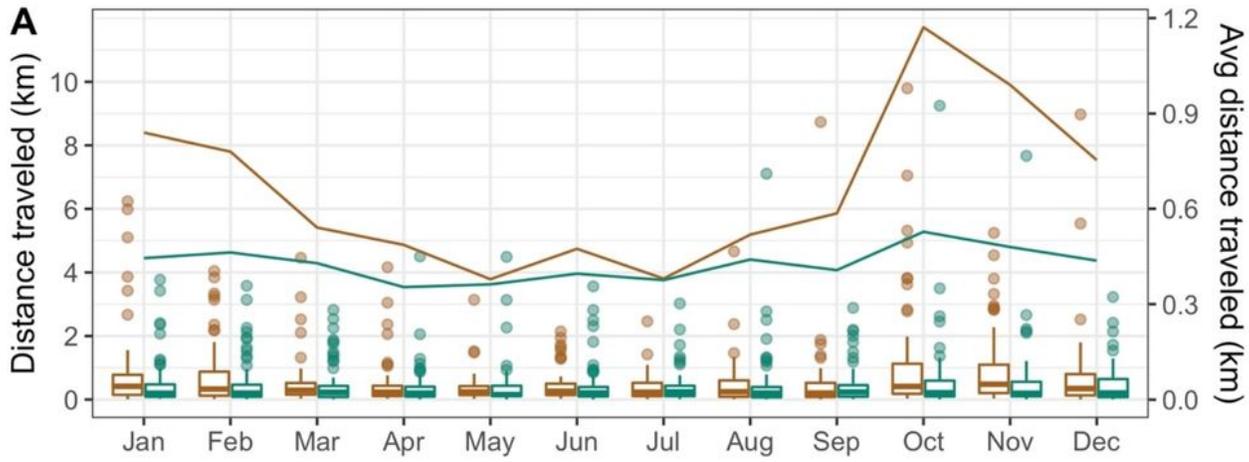


Figure 5

The distribution of distance traveled stratified by birth status.

The distance traveled by month was calculated as the difference between median locations of consecutive months for each fox. In Panel A, the box and whisker plots illustrate the distribution of movements for each month and highlight the long-range movements (points) which are greater than the

sum of the 75th percentile and 1.5 times the interquartile range. The lines denote the average distance traveled by month for captive-born (n=33) and wild-born (n=28) foxes. Panel B depicts the median distance traveled from time of first location on the island (either release from captivity or first capture in the wild). The semi-transparent bands denote the central 95% interval for the distribution of distance traveled. Outliers beyond the 95% interval are shown by the points. For panels A and B, captive-born foxes are shown in brown, and wild-born foxes in teal. Panel C, like Panel A, illustrates the distribution of movements for each month and highlights long-range movements (points). This panel only shows wild-born fox movements and is stratified by time on the island. The first 12 months of wild-born individuals is shown in orange. Time past 12 months is shown in purple. The average distance traveled per month for this time division is illustrated by the lines and the left-side axis.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SUPPLEMENT.docx](#)