Novel, continuous monitoring of fine-scale movement using fixed-position radiotelemetry arrays and random forest location fingerprinting

Andrew B. Harbicht*1, Theodore Castro-Santos2, William R. Ardren3, Dimitry Gorsky4 and Dylan J. Fraser1

© Concordia University, Montreal, QC, Canada;2U.S. Geological Survey, Turner Falls, MA, USA;3U.S. Fish and Wildlife Service, Essex Junction, VT, USA; and4U.S. Fish and Wildlife Service, Basom, NY, USA

Summary

1. Radio-tag signals from fixed-position antennas are most often used to indicate presence or absence of individuals, or to estimate individual activity levels from signal strength variation within an antenna’s detection zone. The potential of such systems to provide more precise information on tag location and movement has not been explored in great detail in an ecological setting.

2. By reversing the roles that transmitters and receivers play in localization methods common to the telecommunications industry, we present a new telemetric tool for accurately estimating the location of tagged individuals from received signal strength values. The methods used to characterize the study area in terms of received signal strength are described, as is the random forest model used for localization. The resulting method is then validated using test data before being applied to true data collected from tagged individuals in the study site.

3. Application of the localization method to test data withheld from the learning dataset indicated a low average error over the entire study area (<1 m), whereas application of the localization method to real data produced highly probable results consistent with field observations.

4. This telemetric approach provided detailed movement data for tagged fish along a single axis (a migratory path) and is particularly useful for monitoring passage along migratory routes. The new methods applied in this study can also be expanded to include multiple axes (x, y, z) and multiple environments (aquatic and terrestrial) for remotely monitoring wildlife movement.

Key-words: animal tracking, continuous remote monitoring, fixed position antenna arrays, location fingerprinting, radiotelemetry, random forests

Introduction

Numerous aspects of a species’ life history, habitat requirements, dispersal, migratory routes, foraging and home-range characteristics can be studied by monitoring individual movement over time (White & Garrott 1990; Cooke et al. 2013). In many environments, however, this presents a challenge due to difficulties associated with visibility limitations (e.g. aquatic environments) or manpower (e.g. around the clock observations). Hence, several methods have evolved to monitoring movement, the most fruitful of which make use of electronic tags; particularly passive integrated transponder (PIT) tags, acoustic transmitters (in aquatic environments), GPS tags and radio (VHF) transmitters.

Fixed-position PIT tag detection arrays (Armstrong, Braithwaite & Rycroft 1996; Lucas et al. 1999) provide location information via proximity detection (within 2 m of the detector) and therefore provide accurate estimates of location at a specific time. A drawback of PIT arrays is their limited detection range which requires a fairly dense array of detectors to provide fine-scale movement data (Castro-Santos, Haro & Walk 1996; Zydlewski et al. 2006). In contrast, acoustic telemetry arrays have much larger detection areas and, depending on the array deployment and tag type, can provide highly accurate estimates of location in two or three dimensions (Heupel, Semmens & Hobday 2006; Hanson et al. 2007). Yet the susceptibility of acoustic arrays to interference from turbulent waters and heterogeneous substrates (Bergé et al. 2012) limits their applicability in fast-flowing or shallow aquatic environments. GPS tags can conversely provide fairly accurate continuous movement data in many environments, but their high cost and size place limitations on study designs (Whitehouse, Karlof & Culler 2007; Hebblewhite & Haydon 2010). In many situations, radiotelemetry is often the preferred choice. Fixed-position radio antenna arrays detect and record the presence of tagged individuals over a much larger detection area than PIT tags, while functioning well in habitats where acoustic systems cannot. Modern VHF tags are also smaller and cheaper than GPS tags. In large environments, radio antennas can act as...
checkpoints, monitoring movement along a migratory path such as a river, whereas in smaller environments, where a single antenna’s detection area may include most of the area of interest, fluctuations in signal strength can be used to estimate activity levels (David & Closs 2001). Currently, the use of radiotelemetry to estimate location on a fine scale is an active area of research (Kays et al. 2011; Ward & Raim 2011; Ward, Sperry & Weatherhead 2013).

Herein, we describe a new method of continuously estimating fine-scale location of tagged individuals using VHF radiotransmitter tags and fixed-position antenna arrays. The method provides a means of remotely monitoring individual movement through an environment with overlapping detection areas that differs considerably from previous proximity detection and angulation methods. Location fingerprinting (LF) describes localization techniques that compare measured system characteristics to a pre-recorded database of system characteristics at known locations (Kjærgaard 2007). They are commonly employed in non-wildlife fields for providing location-aware services where standard localization techniques (i.e. GPS) are ineffective (i.e. indoors). LF methods are typically applied to estimate the location mobile computing devices (e.g. cellular phones) based on the signal strength they receive from signal emitting base-stations with known, fixed locations, often IEEE 802.11b based WLAN signals (Bahl & Padmanabhan 2000; Li et al. 2015). The recorded system characteristics (e.g. received signal strength) are then compared, using various estimation methods, to a pre-recorded database of system characteristics and locations, often referred to as a radio map. By reversing the roles of the transmitters and receivers (mobile transmitters and stationary receivers), we provide researchers with a new telemetric tool to allow continuous monitoring of location and movement using relatively inexpensive equipment.

**Materials and methods**

**STUDY SITE AND SPECIES**

The Boquet River is a tributary of Lake Champlain, New York, USA. Each fall, landlocked Atlantic salmon (Salmo salar) migrate up the river towards spawning habitat located in the river’s upper reaches. Four kilometres upstream from the river mouth, migrating salmon encounter their first barrier, a series of rapids followed by a dam equipped with a fish ladder. These rapids were chosen as our study site because they represent a typical environment in which information on the fine-scale movement patterns of a migratory species would be valuable to researchers. In the present setting, such information would previously only have been possible using an extensive short-ranged PIT tag array due to the irregular substrate and high degree of turbulence. A total of 24 adult salmon were intercepted along their upstream migration and fitted with radiotransmitters (Pisces model; Sigma Eight, Newmarket, ON, Canada) placed inside their abdominal cavity. Each radiotransmitter broadcasts a unique code at regular intervals. To minimize signal cancellation at the receivers due to simultaneous signal reception, transmission rates were staggered over 4–6 s per transmission and over four separate frequencies (164/C1 31, 164/C1 48, 164/C1 34 and 164/C1 33 MHz).

Tagged fish were then allowed to continue upstream into the ~295 m long study reach containing three resting/staging pools and two sections of rapids (Fig. 1). The lower pool (90 × 45 m, 3 m max. depth) often contains many salmon in the fall that appear to use it as a resting/staging area before attempting to climb the first section of rapids. These first rapids span roughly 100 m in length and have an average width of 24 m, with a depth varying from 20 to 70 cm. Immediately above the first set of rapids fish encounter the middle pool, measuring 47 × 28 m with a maximum depth of 2.5 m and is divided into two equally sized sub-basins. Above the middle pool is a second set of rapids that contains a natural ridge that cross-cuts the river’s flow and provides migrating salmon with the only suitable means of achieving the third and final pool (10 × 4 m, 1.5 m max. depth), which contains three small sub-basins. Once in the final pool, salmon must enter and pass a fish ladder to continue upstream.

![Fig. 1. The study reach within the Boquet River, NY, USA, along with locations of four fixed-position aerial radio antennas, one submerged radio antenna, and landmark locations at which the received signal strength from mobile very high frequency (VHF) radiotransmitters were characterized for each antenna.](https://example.com/fig1.jpg)
Sediment mixed with water may cause the sound velocity to change, reducing the ability of long-range LF detection. In this study, we assumed that the sound velocity was constant, to simplify the analysis. However, future studies could investigate how changes in sound velocity affect LF detection. Additionally, while we focused only on signals from the target species, other species could potentially interfere with detection. Future studies could explore how different species impact LF detection performance.

Acknowledgments

This study was supported by the National Sciences and Engineering Research Council of Canada (NSERC) and the University of Toronto. We thank the staff at the Experimental Lakes Area for their assistance in conducting the experiments. We also acknowledge the support of the Ontario Ministry of Natural Resources and Forestry (OMNRF) and the Canadian Wildlife Service (CWS) for their contributions to this project.

References

Baker, R. J., & McFarlane, D. (2005). The influence of water depth on the propagation of low-frequency acoustic signals. Marine Technology, 42(3), 275-290.

Crawford, T. J., & McFarlane, D. (2008). The effect of water temperature and salinity on the propagation of low-frequency acoustic signals. Journal of Experimental Marine Biology and Ecology, 357(1-2), 82-91.

Finney, D. J. (1962). Probit analysis. Cambridge University Press.

Harm, J. A., & McFarlane, D. (2012). The impact of water temperature and salinity on the propagation of low-frequency acoustic signals. Journal of Experimental Marine Biology and Ecology, 417, 16-23.

Winfield, I. F. K., & McFarlane, D. (2007). The influence of water depth on the propagation of low-frequency acoustic signals. Journal of Experimental Marine Biology and Ecology, 357(1-2), 82-91.
error with an antenna-specific SD) 10 times for every 1 m section of the study reach. Finally, the RSS patterns from the original learning data were added to the modelled RSS values ± their variance to form the final extrapolated radio map.

**Estimating Tag Location**

Along with the extrapolated radio map created above, a random forest model was used to estimate tag location. Random forests are an ensemble machine learning method that combine multiple models based on ‘weak’ subsets of the learning data to create a single model with high accuracy (Breiman 2001). When applied to LF techniques, random forests outperform other machine learning methods such as artificial neural nets and support vector machines while requiring less training time complexity (Mo et al. 2014). Random forest models can also provide the user with the distribution of conditional predictions produced by the trees, rather than only a single conditional mean for the forest. These conditional prediction distributions were used to assess confidence in each prediction by calculating the 0.025 and 0.975 quantiles for each estimate (the prediction interval, Meinshausen 2006). The random forest model used to estimate tag distance was created using the **randomForest** package in R (Liaw & Wiener 2002) and used RSS values from the five antennas as explanatory variables, along with a variable indicating which antenna had the maximum RSS value per tag transmission. The tendency of RSS values to decrease proportionally when tags were lower in the water column necessitated the inclusion of relative RSS values (RSS values divided by the sum of RSS values for each transmission) as explanatory variables as well. A total of 1000 trees (ntree = 1000 in R) were created within the model and three explanatory variables were assessed at each branching event (mtry = 3). The model was allowed to potentially over-fit its predictions by allowing a minimum terminal node size of one (nodesize = 1) as this assisted in the estimation of the prediction intervals. To reduce computation time, predictions were made on a subset of the original data containing only one copy of each unique RSS pattern observed from tagged fish. The resulting distance estimates were then joined to the original data by matching the RSS patterns.

**Localization Trend**

Since RSS values received by the stationary antennas represented the true signal strength from tags at a given location plus some error, the resulting raw distance estimates also reflected this error. Thus, a Kalman smoother was applied to each fish’s time series of raw distance estimates to approximate true distance trends. Such smoothers are used with LF methods to increase overall accuracy of location estimates (Guvenc et al. 2003; Chan, Baciu & Mak 2009) by iteratively estimating the true location at a given time based on the previous location, Bayesian inference and an estimation of its joint probability distribution for each time step. The iterative application of prediction and measurement update steps are what allow smoothers to quickly move through a time series of noisy location data and identify the most likely position at each time step. When dealing with complete time series (i.e. all location estimates have been made), this filtering process can be taken one step further by incorporating information from both past and future observations when estimating the true location at any given time. This is the basis of the Kalman smoother implemented in R (tsSmooth command, ‘stats’ package; R Core Team 2016). Smoother parameters were estimated separately for each fish using local level state space models constructed with the **StructTS** command (also in the R ‘stats’ package). A new variable, Estimate Difference, was calculated for each tag transmission as the difference (in metres) between raw and smoothed distance estimates.

**Model Calibration**

Once the distance trend was identified for each tagged fish, certain irregularities presented themselves as periods of a fish’s time series where Estimate difference was above average. These irregularities were handled one of two ways according to their duration. For short irregular periods (<10 transmissions or 50 s) we omitted raw distance estimates with prediction intervals exceeding 100 m prior to applying the Kalman smoother. Long irregularities (>10 transmissions) were thought to arise from tagged fish deviating from the expected migratory path and were anticipated as our extrapolated radio map is a one-dimensional representation of a three-dimensional environment. These deviations produced sequential raw distance estimates with larger Estimate differences and wider prediction intervals. To account for such deviations, the extrapolated radio map was expanded to include additional RSS patterns at various distances from the river mouth. These additional RSS patterns were calculated from the mean RSS values recorded during longer irregular period, whereas the associated distances corresponded to a distance halfway between the preceding and following distance estimates.

**Model Performance**

Model performance was assessed in three different ways by: (i) using the test data removed from the learning data prior to construction of the extrapolated radio map, (ii) comparing the extrapolated radio map constructed above to ‘estimated’ radio maps constructed by calculating the average RSS patterns produced by the 24 tagged fish at estimated distances from the river mouth and (iii) examining the realism and reliability of the approach using two case studies.

**Assessment (i)**

Cross-validation of the random forest model was conducted using the hold-out method by estimating distances for the 114 RSS patterns recorded at known locations and withheld from the learning data. These 114 randomly selected RSS patterns originated from 41 of 49 landmarks and spanned 87-6% of the study reach’s length. Once distance estimates were produced, method accuracy was assessed by calculating the differences between the predicted and actual distances, the prediction error, and comparing this to zero using a t-test. Trends in estimation accuracy and precision were also evaluated by regressing prediction error against the known distances using a linear model; a slope of zero indicated consistent error throughout the study reach.

**Assessment (ii)**

Here the extrapolated radio map was compared to three ‘estimated’ radio maps. Estimated radio maps were created by averaging antenna-specific RSS values across the 24 tagged fish for each 1 m section of the fish path using (i) raw distance estimates produced by the random forest model, (ii) Kalman-smoothed raw distance estimates and (iii) a subset of Kalman-smoothed distance estimates with prediction intervals below a threshold of 100 m. Our aim was to identify distances along the fish path where measured RSS values differed consistently from extrapolated RSS values due to mischaracterization of the fish path (irregular differences) or the effect of depth (proportional differences),
and to demonstrate the effect of applying Kalman smoothers to raw and filtered distance estimates.

Assessment (iii)

We assessed model plausibility by evaluating movements of two tagged individuals. We selected one salmon known to have successfully climbed the ladder and another, tagged on the same day, which displayed considerably different behaviour.

Results

Radio Map

The loess models used to extrapolate mean RSS values across the study reach produced curves that decreased with distance in front of each antenna in a manner similar to the standard negative power attenuation curves typically observed with radio signals (Whitehouse, Karlof & Culler 2007), but with additional undulations. These undulations likely reflect changes in signal attenuation produced by differences in depth and elevation along the irregular terrain of the migration path (Fig. 2). A radio tag moving in an upstream direction (left to right in Fig. 2) along the expected migratory path, moved sequentially into, then out of each antenna’s primary detection zone. At the downstream limit of antenna coverage, RSS values received by each of the five antennas are at the noise floor. As the tag enters the lower pool, however, 3720–3800 m, mean RSS values detected by antenna 1 (situated at 3720 m, Fig. 2) increase rapidly until they reach their maximum of 29. They then gradually decrease over the entire study reach, with the exception of a slight increase at the downstream limit of the two sets of rapids, 3800 m and 3950 m, as the tag re-enters shallower water. Mean RSS values at antenna 2 (situated at 3829 m, Fig. 2) remain near the noise floor throughout most of the lower pool but begin to climb near its upstream limit.

Antenna 2’s RSS values plateau at around 23 until the tag passes under the antenna, when they climb rapidly to a maximum of 45 at 3857 m. Beyond this maximum, RSS values received by antenna 2 decrease until the tag reaches the downstream limit of the second set of rapids where they begin to increase again before falling once more as the tag exits the second rapids 3985 m and enters the final pool. Antenna 3 (situated at 3921 m, Fig. 2) was positioned to overlook the middle pool and exhibited its maximum RSS values directly in front of the antenna, with RSS values tapering off rapidly in both the upstream and downstream directions. An exception to this is a levelling off of RSS values as the upstream limit of the first rapids. Antenna 4 (situated at 4018 m, Fig. 2) expressed RSS values that began rising above the noise floor at the downstream limit of lower pool and which increased throughout the entire study reach until the tag passes directly under the antenna. Noticeable periods of steep increases in RSS for antenna 4 occur at the lower and upper limits of the first rapids, and the lower limit of the second rapids, all coinciding with shallow water depth. The dropper antenna (antenna 5, 4010 m, Fig. 2) in the upper pool only began receiving tag transmissions at the upper limit of the second rapids. RSS values at this antenna climbed in a stepwise manner until reaching a maximum immediately in front of the entry way to the fish ladder, 4015 m, before dropping again as the tag enters the concrete fish ladder at 4020 m.

Model Assessment

Test Data

To test the accuracy of our localization method over the entire study reach, we estimated distances for withheld RSS data and calculated the prediction error for these estimates (the difference between predicted and true distances). Predictions tended
to be slightly upstream from the true distances and had a mean error of 0.6 m and a 95% confidence interval of −1.48 to 2.70, which did not differ significantly from zero ($P = 0.56$, $t = 0.57$, d.f. = 113). When regressed against the known distances, prediction errors produced a non-significant positive slope of $0.017 \pm 0.015$ (±SE, Fig. 3a), with the greatest errors occurring primarily at either end of the first rapids (3800–3885 m). There was no apparent trend between prediction errors and prediction intervals (confidence) for the withheld data, though there was a tendency for the dispersal of the prediction errors around its mean to be greater at greater prediction intervals (Fig. 3b). Interestingly, however, even the most ‘uncertain’ estimates differed from the true distance by <10 m, as the greatest errors were associated with predictions with intermediate prediction intervals spanning 30–60 m.

**Extrapolated vs. estimated radio maps**

Several interesting results were produced by comparing estimated and extrapolated radio maps (Fig. 4a). The distribution of mean RSS values at estimated distances were generally below those modelled from learning data throughout the entire study reach and at each of the antennas; two exceptions were antenna 1 (purple, Fig. 4) at the upstream limit of the lower pool (3780–3800 m) and antenna 4 (red, Fig. 4) in the upper pool (3995–4018 m), where RSS values from tagged fish at estimated locations were higher. These trends were even more apparent for Kalman-smoothed estimates (Fig. 4b). Interestingly, despite the Kalman filter being applied to the distance estimates and not the RSS values, the overall effect was a reduction in variation in RSS trends throughout the study reach. Unexpected fluctuations in mean, Kalman-smoothed RSS values were apparent, such as the local maximum exhibited by the first (purple), third (blue) and fourth (red) antennas at around 3845 m. This overall increase in signal strength across three of the aerial antennas might reflect tagged fish preferring to pass through a particularly shallow section in the first rapids situated at ~3845 m. In addition, antennas 2 (orange), 3 (blue) and 4 (red) produced a local minimum at 3900 m. This drop coincided closely with the downstream limit of the second pool, a particularly deep sub-basin that would have been the first suitable resting area for fish that successfully climbed the first rapids. Filtering out uncertain observations prior to applying the Kalman smoother did not result in a noticeable difference with the raw Kalman-smoothed data (Fig 4c).

**Case example**

The following two case studies demonstrate the ability of LF methods to produce continuous fine-scale movement data for tagged individuals over extended time periods. Both individuals, named ‘Bill and Ted’ for simplicity, were captured, tagged and released into the lower pool on 29 September 2014. Bill was released at 15.30 h, and remained in the lower pool for 71.2 h before attempting to climb the first set of rapids. This attempt lasted only 1.7 min and was ultimately unsuccessful, though it was followed by two other attempts over a span of 4 h. The third attempt was successful and Bill arrived in the middle pool on 2 October at 19.55 h, where he remained for the next 48 h. Bill’s first attempt at the second rapids occurred after 15.7 h and lasted 3.5 min, but ultimately resulted in a further 47.9 h being spent in the middle pool. During this time, Bill shifted towards the downstream edge of the middle pool and moved into and out of the first rapids a total of 14 times, averaging 7 min each. On 5 October, at 11.41.35 h, Bill reentered the second rapids and successfully climbed to the third pool in 2.8 min. Once in the upper pool, Bill remained relatively immobile for another 10 days, dropping back into the

---

© 2017 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society, *Methods in Ecology and Evolution*, 8, 850–859
upper section of the second rapids several times throughout. Finally, on 16 October at 3.28.23 h the signal dropped out as Bill moved through the fish ladder and into the fish trap operated by the NY Department of Environmental Conservation (NYDEC). Bill remained in the trap until the next day when, at 11.00 h the trap was tended by NYDEC staff and he was passed above the dam and out of the study reach. In general, the mean RSS values received from Bill (Fig. 5, upper two panels) at Kalman-smoothed distance estimates match closely with the estimated RSS patterns produced by all 24 tagged fish (Fig. 4) with the exception of a spike in RSS values for antenna one at the upstream limit of the bottom pool as well as generally higher RSS values received by antenna two in the lower sections of the first set of rapids. Bill spent relatively little time in these two sections, preferring to spend most time in the middle and upper pools.

In contrast to Bill’s successful climb up the rapids, Ted did not reach the final pool. Tagged and released into the lower pool only 30 min after Bill (16.00 h), Ted remained in the lower pool for 76.5 h before exiting the study reach in a downward direction on 2 October at 20.31.10 h. Ted then returned to the study reach after 12.3 h, always remaining near the downstream limit of the lower pool and exiting the range a total of 15 times in the first 10.4 days. Ted then attempted to climb the first set of rapids four times before again exiting the study reach in a downstream direction. Each of these attempts to climb the first rapids lasted <3 min. Ted then spent the remainder of the study period either in the lower pool or below the study reach, in the lower section of the river, though he was never detected at the river mouth. Most of Ted’s time was in the downstream half of the lower pool (Fig. 5, lower two panels), or at the downstream limit of the study reach, likely at greater depths than originally modelled in the extrapolated radio map based on relatively lower RSS values within the lower pool.

Fig. 4. The mean received signal strengths at known locations as modelled using loess curves (coloured lines), and the RSS values averaged over 24 tagged Atlantic salmon at distances estimated by random forest localization model (points, panel a). The remaining two panels display the mean RSS values for those same 24 salmon after application of a Kalman smoother to (b) the raw distance estimates, and (c) only those distance estimates with confidence intervals <100 m.
Discussion

We provide a new telemetric tool with which to remotely study animal movement in their natural environment. Our results support that LF techniques employing random forest models are a practical and accurate means of observing and recording the movement patterns of individuals over long periods, while requiring a relatively limited amount of equipment and time. The successful application of LF methods to ecological telemetry greatly expands the uses of fixed-position antenna arrays from simple estimates of activity levels over time (David & Closs 2001; Tucker et al. 2014) and opens the door to many different possibilities. The ability of LF methods to adjust and expand the level of detail required allow such methods to be applied equally well in a variety of aquatic and terrestrial environments wherever overlapping antenna coverage is possible.

When tested using withheld data, the LF method produced distance estimates with very little error on average (within 2 m). This was despite our random forest model using an extrapolated radio map which modelled RSS patterns between landmarks in order to provide continuous RSS estimates. This result highlights the robustness and flexibility of random forests as an estimation algorithm for LF methods. In addition, use of a radio map allows for flexibility when choosing the desired level of precision by adjusting the number of landmarks, artificially binning RSS patterns, or extrapolating RSS patterns into unsampled sections of the study reach (Krumm & Platt 2003; Lee & Han 2012).

Differences between the extrapolated and estimated radio maps provide insights into various features of the study environment and the behaviour of tagged individuals. For instance areas of the study reach where RSS values differed

© 2017 The Authors. Methods in Ecology and Evolution published by John Wiley & Sons Ltd on behalf of British Ecological Society, Methods in Ecology and Evolution, 8, 850–859

Fig. 5. The mean received signal strength at estimated distances from the river mouth for two radio tagged Atlantic salmon captured and subsequently released into the lower pool of the study reach on 29 September and monitored until 19 November 2014. Histograms show the number of tag transmissions (1 every 5 s) received from each fish at each 1 m section of the study reach.
proportionally relative to the modelled RSS values suggest that tagged individuals spent more time deeper in the water column than expected. Alternatively, such proportional differences may have resulted from the positioning of the VHF tag antennas. Tag antennas exited the fish on their ventral surface, placing them in close proximity to the rocky substrate in shallower sections of the study reach, potentially causing signal interference that would not have been detected when the landmarks were characterized using a tag attached to a long wooden dowel. Finally, non-proportional differences between observed and extrapolated radio maps could arise when individuals depart from the modelled migratory path, though these departures were fairly evident upon inspection of the Estimate differences, and prediction intervals and could be incorporated into the radio map before re-estimating locations. In fact, such differences may even provide more information about the study range. For example unnoticed landscape features may have caused the increase in mean RSS values, relative to the expected values, at ~3845 m by the third and fourth antennas or the decrease in RSS values at ~3890 m by antenna 2. Rerunning the random forest model using this estimated radio map may therefore increase prediction consistency or reduce prediction intervals in these areas. This should be done with care, however, and the resulting model should be cross-validated using more withheld data.

From our case examples, it is clear that LF methods can provided detailed estimates of movement patterns for tagged individuals beyond the standard presence or absence data typically collected with VHF tags, while providing greater detail about fine-scale habitat use (Fig. 5). This level of accuracy, combined with the flexibility of programmable VHF tags, makes LF techniques well suited to the study of movement patterns for a wide variety of species. For example LF methods are highly applicable to studies of fish passage where it is important to understand how individuals approach a particular passage structure and how long they remain in each stage of passage (the rates of passage, Castro-Santos & Perry 2012).

Such path selection and passage rates can also be of value in terrestrial environments at highway crossings or when studying movement patterns among meta-populations. In such situations, LF methods may be expanded to collect two- or three-dimensional movement data by altering three aspects of the above methods: first, by constructing a multi-dimensional radio map; second, by using a categorical dummy response variable in the random forest model; and third, by using multiple or omni-directional antennas with the receivers. With any LF exercise, accurate location estimation requires the study environment to be thoroughly characterized in terms of RSS patterns. As with this study, however, environments may contain inaccessible areas in which RSS values may need to be modelled as a function of their x and y (and z) coordinates. The resulting radio map, in table form, would require an additional categorical variable summarizing the location (e.g. ‘B1’ corresponds to $x=2, y=1$) as random forest models only predict single response variables. The resulting predictions would then be back-transformed to provide location estimates. A limitation of LF methods, however, is that each tag transmission must be received by multiple antennas in order for location to be estimated. Larger study environments will therefore require a greater number of antennas, while receivers may need to be paired with a single omni-directional, or multiple directional antennas to avoid ‘blind spots’ within the environment.

Such multi-antenna stations have already proven useful when combined with automated radiotelemetry (ART) methods for studying small mammals, reptiles and birds (Kays et al. 2011; Ward & Raim 2011; Ward, Sperry & Weatherhead 2013). For the moment, however, the accuracy of the LF methods developed in our study (within 2 m) exceeds the accuracy of such methods (typically within 50 m), though whether this will hold when LF methods are applied to multi-dimensional environments has yet to be seen. One advantage of LF methods over ART methods is that they do not rely on theoretical models of radio signal attenuation curves as do ART systems, so accuracy is less affected by obstacles in non-ideal (i.e. no clear line-of-sight) environments. LF methods can simply account for the effect of such obstacles in their radio maps either by characterizing the RSS values surrounding the obstacle or by applying a correction factor to modelled RSS values (Lee & Han 2012). Methods also exist within the LF literature to account for temporal changes in radio mapped environments (Yin, Yang & Ni 2008) allowing for accuracy to be maintained despite temporal variation in signal attenuation.

In summary, the fingerprint localization methods employed in this study, using a random forest estimation method, offer a novel approach to continuous automated wildlife telemetry. The level of precision that fingerprint localization can offer in fluvial environments, from tens of metres to within a metre, was formally available only through the use of dense arrays of PIT antennas in shallow waters (e.g. Sullivan 2004) or acoustic telemetry arrays in deep, open water (e.g. Daniel Deng et al. 2011). In addition, the ease with which received signal strength throughout an environment can be characterized (both for aquatic and/or terrestrial environments), and the ability of such systems to be expanded to include 2D or 3D environments should make them an ideal choice in many research situations.

Authors’ contributions

W.A., T.C.-S. and A.H. conceived the ideas and methodology, whereas D.G. assisted with equipment and data collection. A.H. and T.C.-S. analysed the data. All authors participated in draft revisions, particularly D.F., and gave final approval for publication.

Acknowledgements

For study set-up, data collection and/or comments on the manuscript, we thank D. Hand, M. Little, M. Bonamy, Z. Eisenhower, the town of Willsboro, NY, USFWS, NYDEC, USGS and two anonymous reviewers. The findings and conclusions in the article are those of the authors and do not necessarily represent the views of the USFWS. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.
Data accessibility

The radio map and imputed telemetry data used in this study are available on the Digital Data Repository DRYAD, https://doi.org/10.5061/dryad.5rp65 (Harbicht et al. 2017).

References

Armstrong, I.D., Braithwaite, V.A. & Rycorel, P. (1996) A flat-bed passive integrated transponder antenna array for monitoring behaviour of Atlantic salmon parr and other fish. *Journal of Fish Biology*, 48, 539–541.

Bahl, P. & Padmanabhan, V.N. (2000) RADAR: an in-building RF-based user location and tracking system. Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064), pp. 775–784. IEEE, Piscataway, NJ, USA.

Bergé, J., Capra, H., Pella, H., Steig, T., Ovidio, M., Bultel, E. & Lamouroux, N. (2012) Probability of detection and positioning error of a hydro acoustic telemetry system in a first-flowing river: intrinsic and environmental determinants. *Fisheries Research*, 125–126, 1–13.

Breiman, L. (2001) Random forests. *Machine Learning*, 45, 5–32.

Castro-Santos, T., Haro, A. & Walk, S. (1996) A passive integrated transponder (PIT) tag system for monitoring fishways. *Fisheseries Research*, 28, 253–261.

Castro-Santos, T. & Perry, R.W. (2012) Time-to-event analysis as a framework for quantifying fish passage performance. *Telemetry Techniques: A User Guide for Fisheries Research* (eds N.S. Adams, J.W. Beam & J.H. Eiler), pp. 427–452. American Fisheries Society, Bethesda, MD, USA.

Chan, E.C.L., Baci, G. & Mak, S.C. (2009) Using Wi-fi signal strength to localize in Wireless sensor networks. 2009 WRI International Conference on Communications and Mobile Computing, pp. 538–542. IEEE, Yunnan, China.

Cooke, S.J., Midwood, J.D., Thiem, J.D., Klimley, P., Lucas, M.C., Thorstad, E.B., Eiler, J., Holbrook, C. & Ebner, B.C. (2013) Tracking animals in freshwater with electronic tags: past, present and future. *Animal Biotelemetry*, 1, 1–19.

Daniel Deng, Z., Wailand, M.A., Fu, T., Seim, T.A., LaMarche, B.L., Choi, E.Y., Carlson, T.J. & Brad Eppard, M. (2011) A cabled acoustic telemetry system for detecting and tracking juvenile salmon: Part 2. three-dimensional tracking and passage outcomes. *Sensors*, 11, 5661–5676.

David, B.O. & Closs, G.P. (2001) Continuous remote monitoring of fish activity with restricted home ranges using radiotelemetry. *Journal of Fish Biology*, 59, 705–715.

Guvenç, I., Abdallah, C.T., Jordan, R. & Dedegolu, O. (2003) Enhancements to RSS based indoor tracking systems using Kalman filters. *GPSs & International Signal Processing Conference*, Dallas, TX, USA, pp. 91–102.

Hanson, K.C., Cooke, S.J., Suski, C.D., Nizergoda, G., Phelan, F.J.S., Tinline, R. & Philipp, D.P. (2007) Assessment of largemouth bass (*Micropterus salmoides*) behaviour and activity at multiple spatial and temporal scales utilizing a whole-lake telemetry array. *Hydrobiologia*, 582, 243–256.

Harbicht, A.B., Castro-Santos, T., Ardren, W., Gorsky, D. & Fraser, D. (2017) Data SI, supporting information from: Novel, continuous monitoring of fine-scale movement using fixed-position radio telemetry arrays and random forest location fingerprinting. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.5rp65

Hebblewhite, M. & Hayden, D.T. (2010) Distinguishing technology from biology: a critical review of the use of GPS telemetry data in ecology. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 365, 2303–2312.

Heupel, M.R., Semmens, J.M. & Hobday, A.J. (2006) Automated acoustic tracking of aquatic animals: scales, design and deployment of listening station arrays. *Marine and Freshwater Research*, 57, 1–13.

Kays, R., Tilak, S., Crofoot, M. et al. (2011) Tracking animal location and activity with an automated radio telemetry system in a tropical rainforest. *Computer Journal*, 54, 1931–1948.

Kjærgaard, M.B. (2007) A taxonomy for radio location fingerprinting. *Location-and Context-Awareness* (eds J. Hightower, B. Schiele & T. Strang), pp. 139–156. Springer, Berlin, Heidelberg.

Krumm, J. & Platt, J. (2003) Minimizing Calibration Efforts for an Indoor 802.11 Device Location Measurement System. Technical Report. Microsoft Research, Seattle, WA.

Kuechle, V.B. & Kuechle, P.J. (2012) Radio telemetry in freshwater: the basics. *Telemetry Techniques: A User Guide for Fisheries Research* (eds N.S. Adams, J.W. Beam & J.H. Eiler), pp. 91–139. American Fisheries Society, Bathesda, MD, USA.

Lee, M. & Han, D. (2012) Voronoi tessellation based interpolation method for Wi-Fi radio map construction. *IEEE Communications Letters*, 16, 404–407.

Li, H.L., Quan, W., Ji, G. & Qian, Z.H. (2015) Wireless indoor positioning algorithm based on PCA. *Proceedings of the 2015 International Conference on Artificial Intelligence and Industrial Engineering*, pp. 8–9. Atlantis Press, Paris, France.

Lian, A. & Wiener, M. (2002) Classification and regression by randomForest. *R News*, 2, 18–22.

Lucas, M.C., Mercer, T., Armstrong, J.D., McGinty, S. & Rycorel, P. (1999) Use of a flat-bed passive integrated transponder antenna array to study the migration and behaviour of lowland river fishes at a fish pass. *Fisheries Research*, 44, 183–191.

Mönnichsen, N. (2006) Quantile regression forests. *Journal of Machine Learning Research*, 7, 983–999.

Mo, Y., Zhang, Z., Lu, Y., Meng, W. & Agha, G. (2014) Random forest based coarse locating and KPICA feature extraction for indoor positioning system. *Mathematical Problems in Engineering*, 2014, 1–8.

R Core Team (2016) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Sisson, T.J. (2004) Evaluation of the Turners Falls Fishway Complex and Potential Improvements for Passing Adult American Shad. University of Massachusetts Amherst, Amherst, MA, USA.

Tucker, C.R., Radzio, T.A., Strickland, J.T., Britton, E., Delaney, D.K. & Ligon, D.B. (2014) Use of automated radio telemetry to detect nesting activity in ornate box turtles, *Terrapene ornata*. The American Midland Naturalist, 171, 78–89.

Ward, M.P. & Raim, A. (2011) The fly-and-social foraging hypothesis for diurnal migration: Why American crows migrate during the day. *Behavioral Ecology and Sociobiology*, 65, 1411–1418.

Ward, M.P., Sperry, J.H. & Weatherhead, P.J. (2013) Evaluation of automated radio telemetry for quantifying movements and home ranges of snakes. *Journal of Herpetology*, 47, 337–345.

White, G.C. & Garrett, R.A. (1990) *Analysis of Wildlife Radio-Tracking Data*. 3rd edn. Academic Press Inc, San Diego, CA, USA.

Whitehouse, K., Karlof, C. & Culler, D. (2007) A practical evaluation of radio signal strength for ranging-based localization. *ACM SIGMOBILE Mobile Computing and Communications Review*, 11, 41.

Yin, J., Yang, Q. & Ni, L.M. (2008) Learning adaptive temporal radio maps for signal-strength-based location estimation. *IEEE Transactions on Mobile Computing*, 7, 869–883.

Zydelowski, G.B., Horton, G., Dubreuil, T., Letcher, B., Casey, S. & Zydlewski, J. (2006) Remote monitoring of fish in small streams. *Fisheries*, 31, 492–502.

Received 16 September 2016; accepted 17 January 2017

Handling Editor: Oscar Gaggioiotti

Supporting Information

Details of electronic Supporting Information are provided below.

Data SI. R script used to prepare telemetry data from fixed position antennas, create a one-dimensional radio map of the study environment and estimate tag location using random forest location fingerprinting method.