The optimized interpolation of fish positions and speeds in an array of fixed acoustic receivers

Richard D. Hedger, François Martin, Julian J. Dodson, Daniel Hatin, François Caron, and Fred G. Whoriskey

Hedger, R. D., Martin, F., Dodson, J, J., Hatin, D., Caron, F., and Whoriskey, F. G. 2008. The optimized interpolation of fish positions and speeds in an array of fixed acoustic receivers. – ICES Journal of Marine Science, 65: 1248–1259.

The principal method for interpolating the positions and speeds of tagged fish within an array of fixed acoustic receivers is the weighted-mean method, which uses a box-kernel estimator, one of the simplest smoothing options available. This study aimed to determine the relative error of alternative, non-parametric regression methods for estimating these parameters. It was achieved by predicting the positions and speeds of three paths made through a dense array of fixed acoustic receivers within a coastal embayment (Gaspé Bay, Québec, Canada) by a boat with a GPS trailing an ultrasonic transmitter. Transmitter positions and speeds were estimated from the receiver data using kernel estimators, with box and normal kernels and the kernel size determined arbitrarily, and by several non-parametric methods, i.e. a kernel estimator, a smoothing spline, and local polynomial regression, with the kernel size or smoothing span determined by cross-validation. Prediction error of the kernel estimator was highly dependent upon kernel size, and a normal kernel produced less error than the box kernel. Of the methods using cross-validation, local polynomial regression produced least error, suggesting it as the optimal method for interpolation. Prediction error was also strongly dependent on array density. The local polynomial regression method was used to determine the movement patterns of a sample of tagged Atlantic salmon (*Salmo salar*) smolt and kelt, and American eel (*Anguilla rostrata*). Analysis of the estimates from local polynomial regression suggested that this was a suitable method for monitoring patterns of fish movement.

Keywords: acoustic telemetry, American eel, Atlantic salmon, optimized interpolation, smolt and kelt.

Received 21 November 2007; accepted 11 May 2008; advance access publication 30 June 2008.

R. D. Hedger, F. Martin, and J. J. Dodson: Département de Biologie, Université Laval, QC, Canada, G1K 7P4. D. Hatin: Ministère des Ressources naturelles et de la Faune du Québec, Direction de l'aménagement de la faune de l'Estrie, de Montréal et de la Montérégie, 201 Place Charles Le Moyne, Longueuil, QC, Canada, J4K 2T5. F. Caron: Ministère des Ressources naturelles et de la Faune, Direction de l'aménagement de la faune du Saguenay-Lac-Saint-Jean, 3950 Boul. Harvey, Jonquière, QC, Canada, G7X 8L6. F. G. Whoriskey: Atlantic Salmon Federation, PO Box 5200, St Andrews, NB, Canada, E5B 3S8. Correspondence to R. D. Hedger: tel: +1 418 656 2681; fax:+1 418 656 2043; e-mail: richard.hedger.1@ulaval.ca.

Introduction

The principal method for tracking individual fish within estuaries and coastal zones is through the use of attached ultrasonic tags which transmit signals that can be detected by receivers, usually positioned within a fixed array moored in the water body (Moore *et al.*, 1998; Finstad *et al.*, 2005; Gudjonsson *et al.*, 2005; Lacroix *et al.*, 2005; Whoriskey *et al.*, 2006). The advantage of using a fixed receiver array, as opposed to tracking individuals from a boat, is that the temporal consistency of the array configuration maintains a temporally constant error of the estimate of this position. The ideal array configuration depends on the type of study, but most have focused on large-scale movements so have positioned receivers at relatively large distances from one another. For example, Moore *et al.* (1998) used nine receivers positioned longitudinally along an estuary, separated by an interval of ~500 m, but with tags that had a maximum range of 100 m; Lacroix *et al.* (2005) used 32 receivers positioned in three groups separated by >10 km; Thorstad *et al.* (2007) used 19 receivers, positioned in three groups separated by an interval of ~27 km.

If the receivers are placed in proximity to one another, such that they provide continuous coverage of the region of interest, it may be possible to resolve more of the fish movements at short spatial and temporal scales (Klimley *et al.*, 2001; Heupel and Simpfendorfer, 2002). It may even be possible to interpolate the fish position and to decrease the error of the estimated position if they are placed close enough for the transmitted signals to be detected by more than one receiver (Simpfendorfer *et al.*, 2002). In this case, the objective is to estimate the centre of activity of the fish for a given time. This centre of activity is defined as the fish locus rather than the exact position of the fish, with the implicit assumption that the animal position will deviate from this locus over short time-scales. Two approaches are used, depending on the transmitter type. First, it may be possible to triangulate the transmitter position by comparing the time of arrival of the signal at different receivers, assuming that the shorter the time of arrival, the nearer will be the transmitter to the receiver (Klimley *et al.*, 2001; Ehrenberg and Steig, 2002). This is only possible if the transmitter provides a separate code for each emitted signal.

Alternatively, if it is not possible to link the same transmitted signal to different detections, where, for example, the transmitter is single-coded so that there is not a separate code for each signal, it may be possible to estimate the transmitter position from the relative number of detections at different receivers over a period of time, assuming in this case that the probability of detection is an inverse function of distance. The weighted-mean method (Simpfendorfer *et al.*, 2002; Egli and Babcock, 2004; Giacalone *et al.*, 2005) is the principal method used in this approach, which involves a form of non-parametric regression, equivalent to the Nadaraya–Watson box-kernel estimator, to estimate local means separately along two dimensions, typically longitude and latitude or easting and northing. Here, the spatial position estimated for any given time instant is the mean of the receiver positions, weighted according to the number of detections at each receiver, within a given time interval around that time instant, determined by the kernel size. The greater the kernel size, the greater the amount of smoothing. The Nadaraya–Watson box-kernel estimator is relatively simple in terms of the large number of alternative non-parametric regression methods that exist. For example, it gives equal weight to individual observations within the kernel, whereas it may be more intuitive to give greater weight to observations nearer the kernel centre, i.e. a normal kernel. Given that establishing a fixed receiver array is expensive, it may not be possible to ensure a large area-of-detection overlap between neighbouring receiver ranges, and the spatial arrangement of detections as a function of time may not be ideal for the box-kernel estimator.

The objective of this study was to determine the optimal method, in terms of minimizing error for a given sample size, for estimating the positions and speeds of fish tagged with single-coded ultrasonic transmitters within a fixed receiver array. To do this, we tested a variety of non-parametric regression models, e.g. kernel estimators, smoothing splines, and local polynomial regression, for estimating paths made through a fixed-receiver array using boat-mounted transmitters. We also tested the effect of array resolution. The optimal method was then applied to a sample of tagged Atlantic salmon (*Salmo salar*) smolt and kelt, and American eel (*Anguilla rostrata*) to evaluate its use for monitoring patterns of fish movement.

Material and methods

Study area

Gaspé Bay is a coastal embayment located on the northeastern coast of the Gaspé Peninsula in Québec, Canada (48.85°N 64.45°W; Figure 1). It has three distinct regions: the York Estuary, an inner, partially enclosed bay, and an outer bay which connects with the Gulf of St Laurence. The York Estuary is ~1 km wide, ~10 km long, and has a mean depth of several metres. The inner bay is >4 km wide and shallow, with a maximum depth of ~25 m, and a length of ~10 km. It is separated from the outer bay by a sand-wedge which becomes submerged at high tide, but there is a relatively deep channel in the north, ~1 km wide and 20 m deep, in which the vast bulk of the interchange of water between the inner and outer bays occurs. Gaspé Bay receives freshwater input from the River Dartmouth and the York River, a salmon river that discharges into the York Estuary.

The bay is typically vertically stratified in summer, with a surface layer of warm, low-salinity water of riverine origin, overlying a layer of colder, saline water of maritime origin (Carrière, 1973; Koutitonsky *et al.*, 2001). Typical of all estuaries of this size, the circulation is driven by tides, wind, pressure gradients, and riverine input (Koutitonsky and Bugden, 1991). Circulation patterns in Gaspé Bay are complex (Pettigrew *et al.*, 1991; Hedger *et al.*, 2008), which may complicate spatial and temporal patterns of migrating fish. Sampling in 2005 showed a gradient of surface salinity, increasing across the inner bay. Mean salinity was 10.80 near the mouth of the York Estuary, and 16.68 PSU near the deep channel separating the inner and the outer bays.

Transmitter and fixed receiver array

The transmitter model used was the V9-6L (Vemco Ltd). This has a diameter of 9 mm, a length of 20 mm, and weighs 3.3 g in air. It was set to emit an ultrasonic signal every 20–50 s at 69 kHz for a maximum of 53 d. Each transmitter has a separate single-coded ID: different emitted signals from the same transmitter have the same ID. The range of the transmitter varies between 200 and 700 m, depending on background noise and the parameters of the transmitter deployment (see Simpfendorfer *et al.*, 2008, for a discussion of the issues relating to transmitter range).

The acoustic receiver used was the VR2 (Vemco Ltd). This is a single-channel, omnidirectional acoustic receiver that records the time and signal ID. In all, 62 acoustic receivers were moored in a fixed array in York Estuary and Gaspé Bay in May and June 2006, to provide continuous coverage from a distance of ~5 km from the entrance of the York River into the York Estuary to the easternmost perimeter of the inner bay, including two receivers placed to the east of the sand-wedge (Figure 1). Receivers were not placed in the western part of the inner bay because of the presence there of aquaculture installations. However, the principal direction of salmon migration was in a seaward (eastward) direction (Hedger *et al.*, 2008), and eels mainly remained in York Estuary, so coverage of the main region occupied by fish was near complete. The array had a hexagonal configuration, with each non-boundary receiver surrounded by six adjacent near-equidistant receivers separated by a distance of ~507 m (332 m in the estuary and 611 m in the bay). Area-of-detection overlap, i.e. the area covered by more than one receiver, depended on the range of the transmitter, varying from zero overlap with a transmitter range of 200 m (Figure 1b) to >70% overlap with a transmitter range of 500 m (Figure 1c).

Determination of optimal interpolation method

The optimal interpolation method was determined by estimating independently known transmitter positions. A transmitter was attached to the outside of a boat below the water surface, and three boat paths were made through the inner bay: (i) Boat Path 1, 10 June 2006, from 09:51 to 14:14; (ii) Boat Path 2, 11 June 2006, from 10:21 to 14:29; and (iii) Boat Path 3, 12 June 2006, from 09:18 to 14:12 (Figure 2; left panels). Boat speed was maintained at 0.5–0.8 m s^{-1} to approximate fish ground speeds and to minimize the effect on transmitter range from boat-induced turbulence. The position of the transmitter was determined at intervals of ~20 m (approximately every 30 s) using a GPS. Boat-path locations were selected to be representative of areas occupied by migrating Atlantic salmon as determined by previous observations of migration in 2005. Boat paths were made across the entire bay to provide information on spatial variation in transmissibility.

The boat paths were estimated by four non-parametric regression methods available in the standard and *sm()* (Bowman and Azzalini, 1997) libraries of the statistics package R (Hornik,

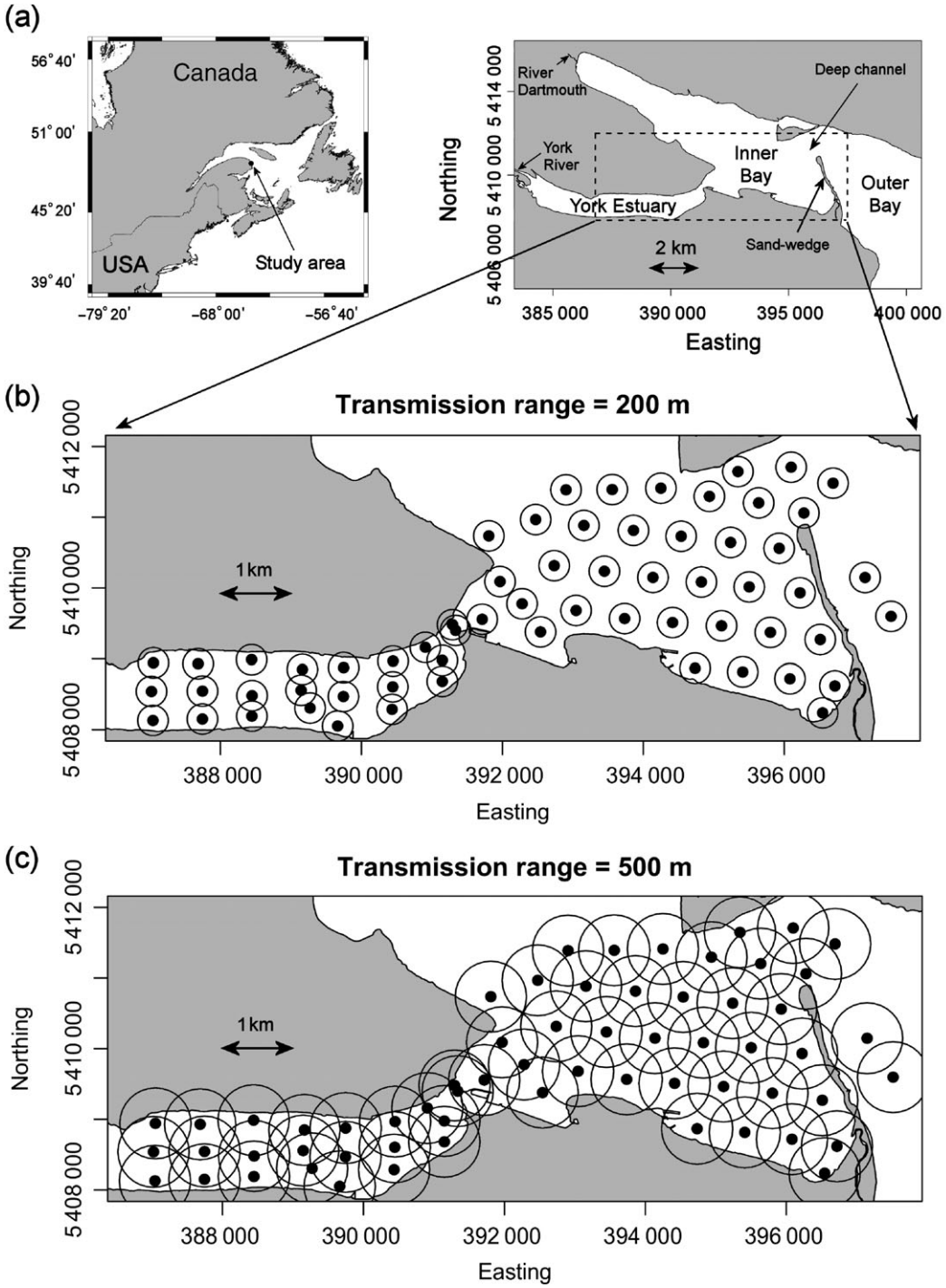

Figure 1. (a) Study area, showing the position of the VR2 receivers, along with the area of detection associated with transmission ranges of (b) 200 m and (c) 500 m.

2007): (i) the Nadaraya–Watson kernel estimator; (ii) a cross-validated kernel estimator; (iii) a cross-validated smoothing spline; and (iv) a cross-validated local polynomial. In all cases, relationships were established independently between position (easting or northing, the response variable) and time (the predictor variable). First, the Nadaraya–Watson box-kernel estimator was used to determine the error associated with the approach most commonly used; this estimator was also applied using a normal kernel. The kernel estimator was implemented using the *ksmooth()* function with arbitrarily determined bandwidths of 5, 10, 15, 30, and 60 min. This function scales the kernels so that their quartiles are at $\pm 0.25 \times$ bandwidth. The box kernel applies an equal weight to all observations within the kernel, so the kernel extends to $\pm 0.5 \times$ bandwidth. The normal kernel applies a weight directly proportional to a Gaussian "bell-shaped" curve, centred on the centre of the kernel; i.e. observations nearer

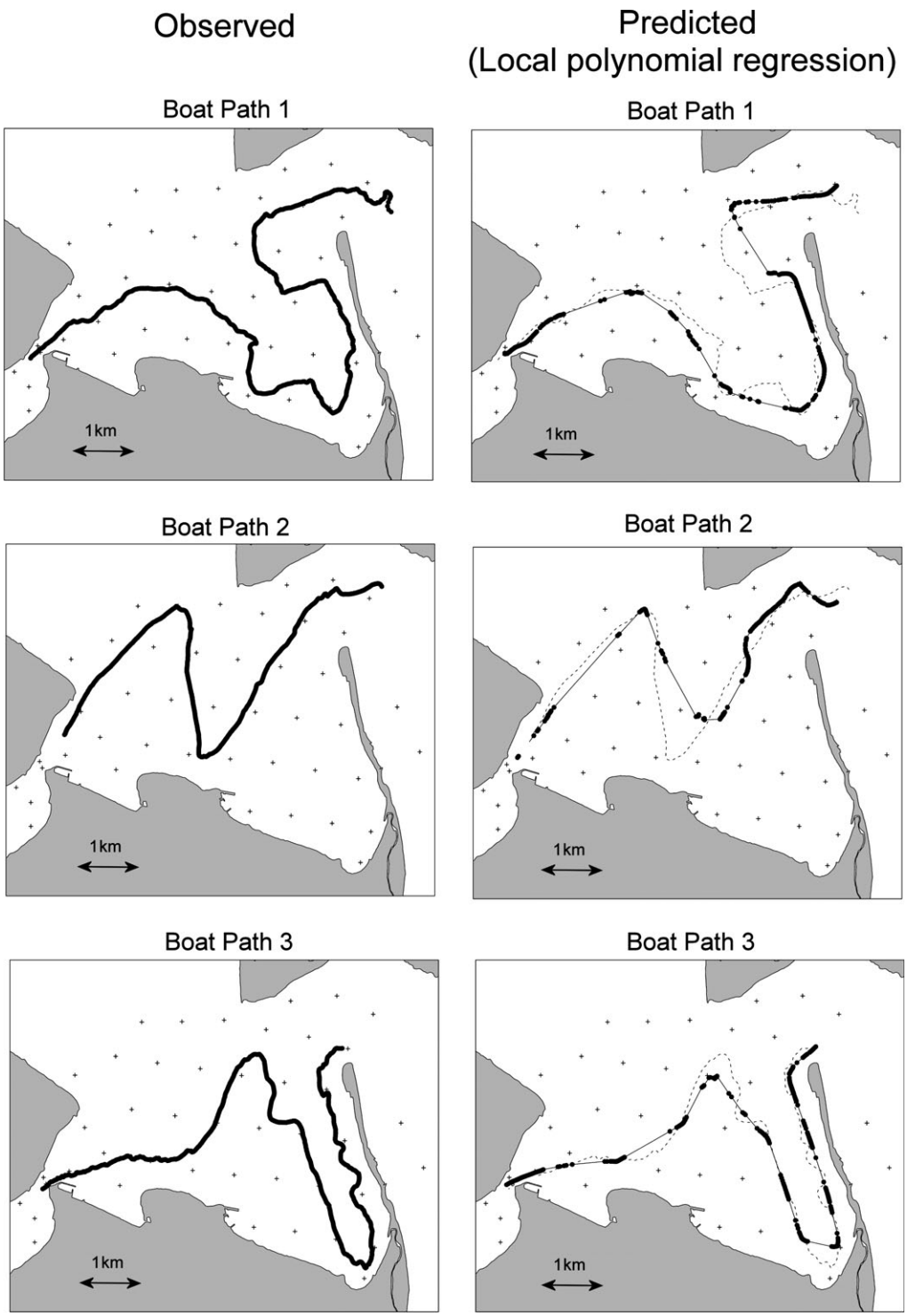

Figure 2. Boat paths observed (left panels) and estimated (right panels). Observed paths were determined by the onboard GPS; estimated paths were determined by local polynomial regression using cross-validation. Local polynomial estimates are shown by dots, circles, and the linear interpolation between these is shown by a continuous line. To aid comparison, observed paths are also shown in the right panels by a dashed line.

the centre of the kernel have more weight than those nearer the edge. Second, a cross-validated kernel estimator was applied, cross-validation being used to determined the optimal bandwidth iteratively. The algorithm used was originally developed by Bowman and Azzalini (1997) and implemented in R by the function *sm.regression()*, with its smoothing parameter determined using the *hcv()* cross-validation function. This estimator uses a normal kernel. Third, a cross-validated smoothing spline was applied. Spline fitting involves selecting a series of knots throughout the data, then using cubic regression to estimate points in between the

knots. The *smooth.spline()* function implemented by B. D. Ripley and M. Maechler, based on code by Hastie and Tibshirani (1990) was applied. Finally, a cross-validated, local polynomial regression was applied. The local polynomial regression uses weighted-least squares to fit a dth degree polynomial to the data. Freidman's SuperSmoother (Friedman, 1984), which has a variable bandwidth and is implemented in R with the *supsmu()* function, was used. The SuperSmoother algorithm applies three symmetrical smoothers using the nearest $k/2$ observations on each side of the observation to be predicted: (i) $k = 0.5n$, (ii) $k = 0.2n$, and (iii) $k = 0.05n$, where n is the number of observations. The optimal smoother is then chosen by cross-validation.

Errors, the mean absolute distance between observed and estimated position and the mean absolute time difference between observed and estimated speed, were determined for each method. All the interpolation methods estimated the boat positions at the times when the transmitter signals were received, and given that these times were not synchronous with those determined when the GPS measurements were made, the GPS measurement that was nearest in time to that of each estimated position was used in estimating the error.

Estimates from the weighted-mean and three cross-validated, non-parametric regressions were only possible in areas where there were high densities of signal detections. Examination of the VR2 data revealed sections of the boat paths that were not recorded by the VR2s. Transmitter positions in these sections were estimated by linear interpolation using the function *approx()* (Becker *et al.*, 1988). In our study, this was merely a post-processing procedure for aid in plotting the paths.

Effect of receiver-array density

The effect of receiver-array density and the inverse, the distance between neighbouring receivers, on the error of the estimate was determined by estimating the boat paths with two of the smoothing methods, the box-kernel estimator and local polynomial regression, using subsamples of the receiver array, by removing detections from selected receivers, and retaining the rest. Subsample sizes ranged from 25 to 43 receivers, the total number of receivers in the inner bay. The procedure was:

(i) select a subsample;

(ii) for the given subsample, select receivers randomly, estimate boat paths, then determine the errors of the estimate and the mean distance between neighbouring receivers;

(iii) repeat Step 2 99 times with the configuration of receivers changing at each iteration;

(iv) determine the mean error and mean distance between neighbouring receivers from the respective means of the 250 iterations from Steps 2 and 3;

(v) repeat Steps 1–4 with a different subsample size.

Application to migrating fish

The optimal interpolation method was then used to estimate the positions and ground speeds of a sample of Atlantic salmon (both smolt and kelt) and American eel. In all, 30 smolt, 24 kelt, and 20 eels were acquired in May and June from the York River. The transmitter-implantation procedure followed the method of Summerfelt *et al.* (1990). Each fish was anaesthetized by immersion in a 40 mg l^{-1} clove-oil solution for 5–10 min (Chanseau *et al.*, 2002), then a longitudinal incision was made on the ventral side, the transmitter inserted into the body cavity, and 3–4 silk points made to close the incision. All instruments were sterilized with a chlorhexidine gluconate 0.05% solution (Baxedin, Omega Laboratories Ltd) prior to use. The gills were irrigated throughout the surgery. After surgery, fish were placed in a holding tank with flowing water (7–10°C), then released into the York River.

The optimal interpolation method was applied separately to detections from the estuary and the bay. This was because a model fitted to the entire dataset for each fish often caused spurious predictions in the vicinity of the intersection between the York Estuary and the inner bay, where the estimated fish path did not pass through the narrow intersection, but crossed over the land surface ~100 m to the south and east. Additionally, a model fitted to the entire dataset for each fish occasionally caused spurious predictions in the inner bay, because of the influence of the large number of data in the York Estuary. Given that a transition from the estuary to the bay represented a major change in habitat, in terms of depth, salinity, and circulation, it can be expected that the behaviour of the fish changed, so the path predicted for the inner bay should be independent of that predicted for the estuary.

Results

Optimal interpolation method

In all, 293, 287, and 253 detections were made by the VR2 receivers during Boat Paths 1, 2, and 3, respectively, representing a signal being detected every 53, 54, and 70 s. The maximum transmitter ranges were 522, 656, and 620 m for Boat Paths 1, 2, and 3, respectively: this was not necessarily indicative of a temporal change in transmissibility because the spatial configuration of the boat paths differed. The total number of detections, and therefore the detection rate, decreased with increasing distance between the transmitter and the nearest receiver: 52.9% of detections were within a distance of <200 m, 36.5% within a distance of 200–400 m, 9.4% within a distance of 400–600 m, and 1.0% with a distance of >600 m. At distances more than ~500 m and usually when the boat was at a midpoint between two receivers, the detection rate was so low that there was not a continuous coverage of the boat track by the VR2 array. A spatial trend in transmission range was apparent: mean distance between observed transmitter position and the position of VR2 signal detections was 250 m in the eastern (seaward) part of the array, and 131 m in the western (landward) part. An example of the estimated boat paths using a cross-validated local polynomial regression is given in Figure 2 (right panels).

The estimated paths from the kernel-estimator method were dependent on kernel size and type (Figure 3). With a decrease in kernel size, there was a decrease in the smoothing of the changes in boat-path orientation and a decrease in the truncation of the boat paths, but the predictions became more clustered towards the positions of the VR2 receivers, so that successive groups of estimates were sometimes separated by distances of several hundreds of metres. For a given kernel size, the estimated positions from the box kernel were more clustered towards the positions of the VR2 receivers than those from the normal kernel. The boat paths estimated from the cross-validated methods are shown in Figure 4. Of the three methods, local polynomial regression tended to produce less disjointed (i.e. least clustered) estimated positions along the path.

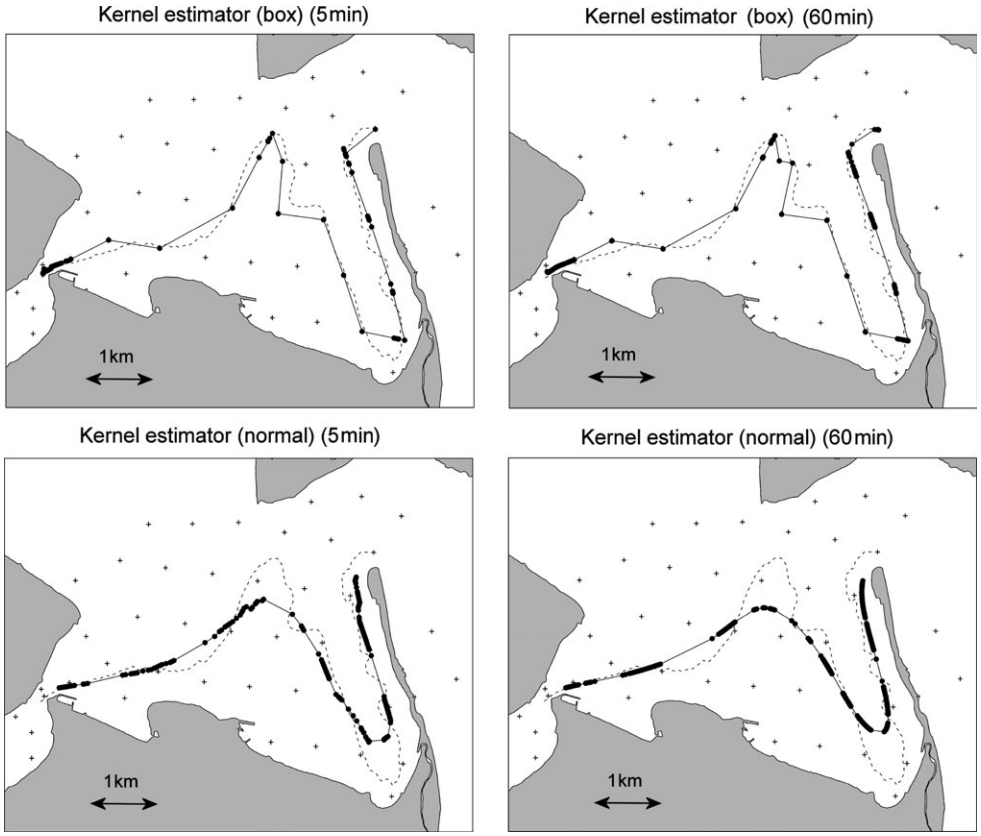

Figure 3. Effect of kernel size on the estimate of Boat Path 3 positions for the kernel-estimator method with box and normal kernels. Estimates are shown by dots, and the linear interpolation between these is shown by a continuous line. Observed paths are also shown by a dashed line.

The optimal kernel size for predicting position lay between the two extremes at 15 min, when considering all paths together for Boat Paths 1 and 3, and at 30 min for Boat Path 2 (Table 1; Figure 5). The mean error of estimated speeds increased with decreasing kernel size. A normal kernel produced smoother estimated boat paths, and a smaller positional and speed error than the box kernel. Of the three cross-validated methods, the local polynomial regression method produced least positional or speed error (Table 1; Figure 5).

Effect of receiver array density

By subsampling the receiver array to simulate the effect of using a less dense array, it was evident that the mean absolute distance error of the estimated boat-path position increased with an increase in the mean distance between neighbouring receivers in a near-linear fashion, for both the box-kernel estimator, i.e. the weighted-mean method, and the local polynomial regression (Figure 6). The mean distance had a large effect: an increase of ~5% from 590 to 620 m caused the mean absolute distance error to increase from 112 to 132 m when using local polynomial regression, and from 129 to 147 m when using the box-kernel method with a bandwidth of 15 min, increases of ~17 and ~13%, respectively. The method of local polynomial regression always produced less error than the box-kernel estimator.

Application to migrating fish

In total, 162 482 signals were detected by the VR2 receivers: the mean numbers of detections per smolt, kelt, and eel were, respectively, 3150, 2940, and 739, and the overall mean number of detections per fish was 2708. For the total time from first to last detection, some 466 808 signals would have occurred, given a mean signal interval of 35 s, so a large proportion of signals went undetected: >65% in the entire estuary/bay system, >64% in the estuary, and >72% in the bay. There were some instances of successive detections of fish at non-neighbouring receivers, and there were also indications of some fish leaving the array on the western margin, to return later. In one case, the time interval between two successive detections at the same receiver was 3.59 d, and given that this took place in one of the boundary receivers on the western margin, it is likely that this result was caused by a fish temporarily leaving the array.

Fish positions estimated by local polynomial regression were significantly correlated with the positions of VR2 detections: bay easting ($r = 0.99$), bay northing ($r = 0.96$); $p < 0.001$. These correlations were of a similar order to those determined between boat positions estimated from local polynomial regression and the positions of VR2 detections: bay easting ($r = 0.99$), bay northing ($r = 0.98$); $p < 0.001$. Predictions were therefore following the paths of the detections, suggesting that local polynomial regression was a valid method for predicting fish positions, although there were occasional gaps in the fish paths estimated from it at the midpoints between receivers where signals were not being detected. Visual inspection of the predicted fish paths and the predicted eastings and northings as a function of time did not reveal any major problems with the model fit, other than occasional poor fits at either end of a fish path resulting from situations where

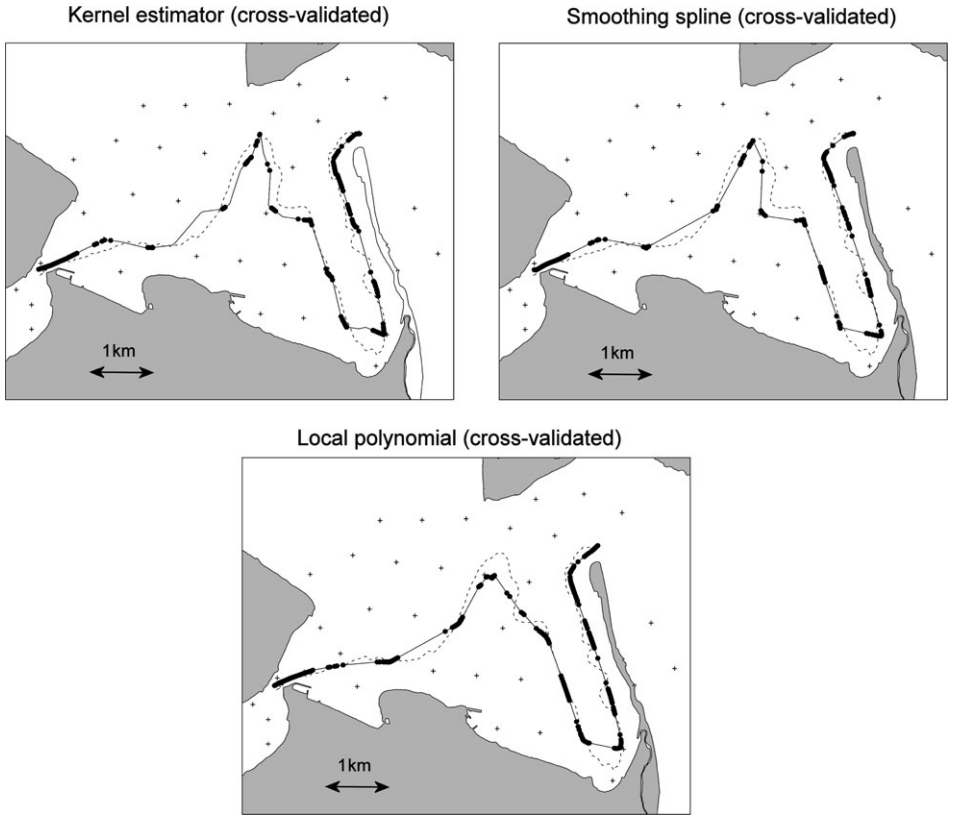

Figure 4. Boat Path 3 positions estimated from the cross-validated kernel estimator (normal kernel), smoothing spline, and local polynomial methods. Estimates are shown by dots, and the linear interpolation between these is shown by a continuous line. Observed paths are also shown by a dashed line.

Table 1. Non-parametric regression methods and the errors of their estimates of boat paths.

Non-parametric model	Kernel type	Bandwidth	Mean absolute distance (m)	Mean absolute speed difference (m s^{-1})
Kernel estimator	Box	User-determined ($t = 5$ min)	141	0.79
Kernel estimator	Normal	User-determined ($t = 5$ min)	114	0.32
Kernel estimator	Box	User-determined ($t = 10$ min)	137	0.48
Kernel estimator	Normal	User-determined ($t = 10$ min)	129	0.49
Kernel estimator	Box	User-determined ($t = 15$ min)	131	0.63
Kernel estimator	Normal	User-determined ($t = 15$ min)	132	0.28
Kernel estimator	Box	User-determined ($t = 30$ min)	124	0.37
Kernel estimator	Normal	User-determined ($t = 30$ min)	249	0.49
Kernel estimator	Box	User-determined ($t = 60$ min)	123	0.66
Kernel estimator	Normal	User-determined ($t = 60$ min)	270	0.36
Kernel estimator	Normal	Cross-validated	116	0.31
Smoothing spline	NA	Cross-validated	128	0.34
Local polynomial regression	NA	Cross-validated	107	0.24

the initial or final observation was an outlier. Fish paths, regardless of species (salmon or eel) or life-stage (smolt or kelt), were complex; selected paths are shown in Figure 7. Several patterns were found for salmon, including (i) near-linear direct migration to the open channel separating the inner bay from the outer bay (e.g. Smolt 3 and Kelt 5); (ii) repeated changes in direction (e.g. Smolt 11 and Kelt 2); (iii) initial migration towards the sand-wedge, followed by a northward migration, often at low tide (e.g. Smolt 9 and Kelt 7); and (iv) migration across the sand-wedge at high tide (e.g. Smolt 6). Eel, being largely resident in the estuary, did not display the same across-bay migration, but did occupy the longitudinal range of York Estuary (e.g. Eel 3), with one eel penetrating into the bay (Eel 4).

Eels exhibited the slowest predicted ground speed, with mean speeds of 0.15 and 0.14 m s^{-1} in the estuary and inner bay, respectively (note that here, the term "mean" refers to the population

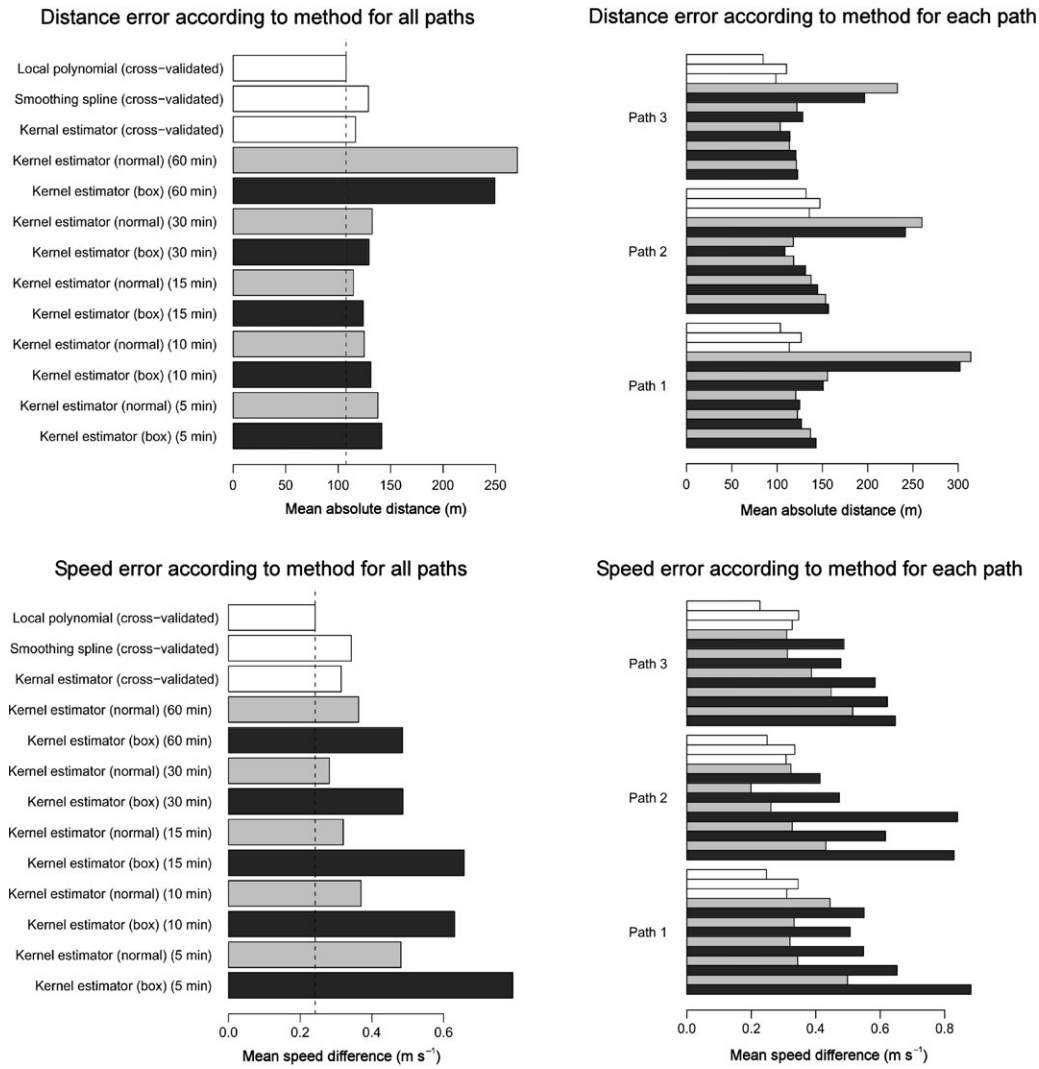

Figure 5. Mean absolute distance and mean absolute difference in speed between observed and estimated boat paths according to non-parametric regression method.

mean of the mean speeds of each fish). Smolt had a faster predicted ground speed, with a mean speed of 0.16 and 0.21 m s^{-1} in the estuary and inner bay, respectively, and kelt the fastest ground speed, mean speeds of 0.27 and 0.35 m s^{-1} in the estuary and bay, respectively.

Discussion and conclusions

Although the methods, kernel estimator, smoothing spline, and local polynomial regression we used differed, their implementation was based on the same rationale: that a series of detections at distinct locations can be smoothed to provide a local mean, or centre of activity, given that the probability of a signal being detected at a receiver is an inverse function of the distance between receiver and transmitter. The prediction error of all methods is dependent on (i) the density of the receiver array, (ii) the transmitter range, and (iii) the repeat interval of the transmitter.

We have shown, first, that the density of the receiver array is important, a small increase in distance between receivers causing a large increase in error. It was also evident that even with the comparatively short distances between different receivers, most transmitter signals were not detected. This problem was particularly acute for the midpoints between receivers, where there were sometimes insufficient detections to interpolate the transmitter position using one of the non-parametric regression methods, so necessitating linear interpolation. Although a dense grid was used for this study, an even denser grid is clearly required. Our second main finding was that the detection rate depended on how far a signal can travel. Even the strongest transmitters of the size that can be used in fish such as salmon or eel, for example, tend to have a maximum range of ~700 m, so this range may be reduced by background noise generated by hydrological conditions (tidal current, river flow, water turbulence, and waves), meteorological conditions (wind and rain), water turbidity, and water-column depth (Voegeli *et al.*, 1998; Finstad *et al.*, 2005; Heupel *el al.* 2006). Spatial variation in noise, and hence transmissibility, as shown here, will bias the estimated position. Additionally, the maximum range may be affected by parameters of the receiver deployment, such as receiver depth (Lacroix and Voegeli, 2000) and receiver orientation (Clement *et al.*, 2005). For example, currents may displace the buoy to which the receiver

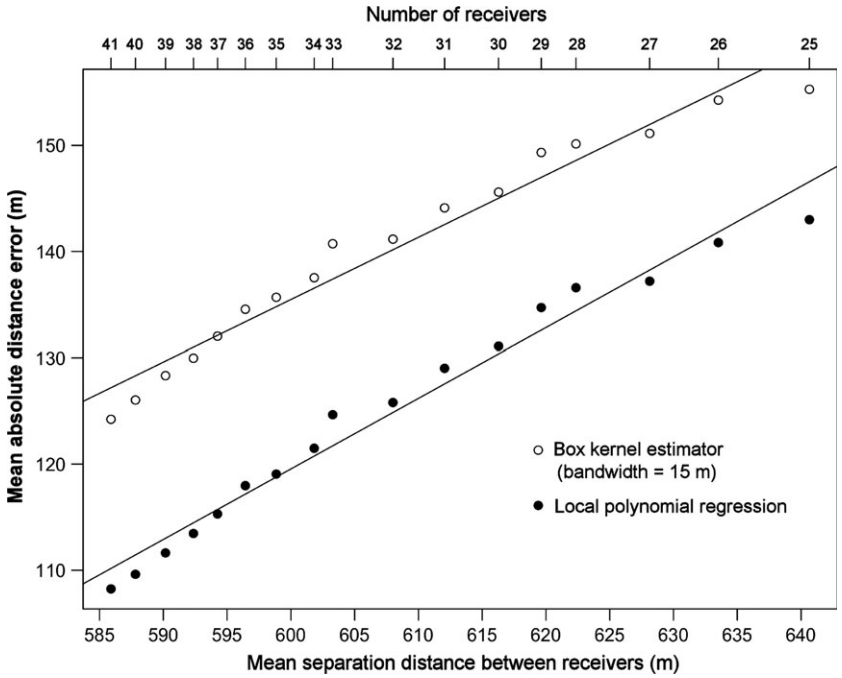

Figure 6. Effect of sample size and distance between neighbouring receivers on the error of the estimated boat paths.

is attached, changing the longitudinal axis of the receiver so that it is no longer perpendicular to the water surface, resulting in reduced range (A. P. Klimley, pers. comm.). These factors may result in a large number of signals remaining undetected. In this case, >65% of emitted signals were not detected; by way of comparison, Simpfendorfer *et al.* (2008) found that almost 60% of signals were not detected in their study of the performance of a receiver array mounted within a river. Finally, our results indicate that the repeat interval of the transmitter will have a great impact on the detection rate. This interval is variable, according to transmitter configuration. In this instance, the transmitter (V9-6L) emitted one signal on average every 35 s, but other authors have used transmitters with much shorter repeat intervals. For example, Moore *et al.* (1998) used transmitters emitting an average of 60 signals per min; Økland *et al.* (2001) used transmitters (16M, Advanced Telemetry Systems) emitting an average of 50–80 signals per min. The interaction between these three aspects—resolution of the receiver array, transmitter range, and the repeat interval of the transmitter—should be considered before implementing any fish-tracking study. In particular, it may be advisable to use boat transects to determine the transmitter range under different environmental conditions. The receiver array used here, with a mean separation distance of ~507 m, provided near-complete coverage, but the many undetected signals, and the inability to use non-parametric regression to interpolate the transmitter position throughout the entire time the transmitter was within the embayment, suggests that a denser receiver array, greater transmitter range, or higher repeat interval would have been better.

Despite the constraints imposed by any given transmitter and receiver configuration, it is possible to optimize the estimate of the position and speed of the transmitter by selecting the most appropriate method of interpolation. This study has shown that for the transmitter and array configuration implemented:

(i) the box-kernel estimator, the most commonly used method, tended to produce larger errors than the normal kernel estimator;

(ii) the local polynomial regression was generally the optimal method, producing smaller errors than the kernel estimator or spline methods;

(iii) the amount of smoothing, determined by the bandwidth time interval, was crucial in determining the error.

In estimating the transmitter position, the box-kernel estimator either included or excluded detections, depending on whether or not they occurred within the window specified by the bandwidth, fragmenting the estimated transmitter positions into clusters around the VR2 positions. The normal-kernel estimator, in contrast, assigned a weight to each detection according to its temporal distance from the detection time for which a transmitter position was being estimated, so resulted in a less fragmented distribution of the estimated transmitter positions. It can therefore be inferred that if a kernel estimator is to be used, the box kernel should be rejected in favour of the normal kernel. Of the other two methods, smoothing spline and local polynomial, the smoothing spline overfitted the data when using a cross-validated smoothing parameter. A smoother model could have been fitted, but this would have entailed setting the smoothing parameter subjectively. Of all the methods tried, the cross-validated, local polynomial regression worked best. This approach had the advantage of a variable bandwidth, with a span comprising a specific proportion of the data rather than using a fixed-width kernel, so it provided a flexible adjustment to spatial changes in inflexion of the path of the transmitter. For example, with a span of 0.2, the local polynomial regression used the nearest 20% of the data ($\pm$10%) around any detection time to predict the transmitter position, regardless of whether there had been a small or large change in

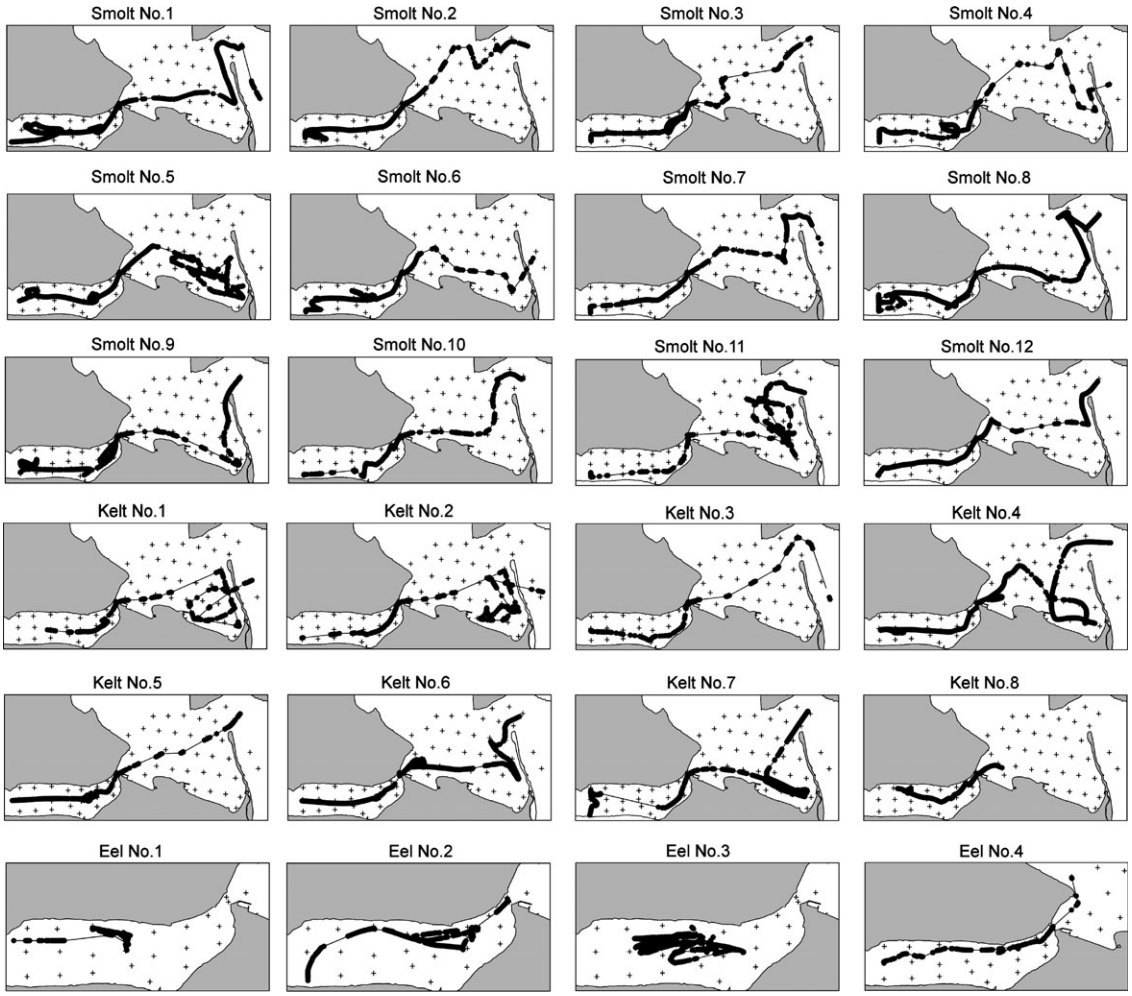

Figure 7. Fish paths estimated using local polynomial regression. Estimates are shown by dots, and the linear interpolation between these is shown by a continuous line. The observed paths are also shown by a dashed line.

the detection position. Therefore, in regions where there was greater change, it was able to increase the inflexion of its estimates. Finally, the amount of smoothing had a crucial effect on the error of the estimate. For example, the distance error of the kernel estimator increased by a factor of two with an increase in bandwidth from 15 to 60 min. A cross-validation approach for determining the optimal bandwidth was effective and is suggested for future studies.

The local polynomial regression was a valid method for estimating fish paths, with a similar correlation between estimated positions and the VR2 positions to that found when using boat paths, the prediction of which were independently validated by reference to a GPS. Although validation of the method was only undertaken with boat paths made in the bay, fish paths were predicted within the estuary as well as the bay. Not only did the configuration of the receiver array vary between the two regions, with a denser array in the estuary, but the environmental characteristics varied between the two regions, so the ambient noise levels probably also varied. Although only validated within the bay, the technique was transferable to the estuary because detection rates in the estuary (35.6%) and the bay (27.8%) were relatively similar. The slightly greater rate of detection in the estuary can be attributed to the greater density of the receivers there.

The dense receiver array exposed complex patterns of fish movement, ranging from residence in the estuary for eel, to the across-bay migration of salmon smolt and kelt. Ground speeds were of a similar order to those reported previously. For example, the mean ground speeds for smolt of 0.16 and 0.21 m s^{-1} in the estuary and inner bay, respectively, were similar to those calculated by Moore *et al.* (1998; 0.14–0.29 m s^{-1}) and Finstad *et al.* (2005; 0.15–0.23 m s^{-1}). Mean eel ground speeds (0.15 and 0.14 m s^{-1} in the estuary and inner bay, respectively) were slower than those identified by McCleave and Arnold (1999; 0.35–0.58 m s^{-1}) for European eel: the differences may be attributed to differences in study area or behaviour of individual eels.

To conclude, the principal method used for interpolating fish positions and speeds within a fixed receiver array—the weighted-mean, box-kernel estimator—does not produce the minimum possible error. A normal kernel tends to produce less error, and alternative methods such as local polynomial regression may further reduce the error. The bandwidth is crucial; cross-validation may be a suitable method for determining optimal bandwidths. It is suggested that future studies should predict fish positions by more than one method, so that the method-specific variation among the predictions can be used to indicate the relative strengths and weaknesses of each method as well as demonstrating

the uncertainty implicit in interpolating fish positions. Some form of method validation using empirical data is suggested. Here, this was achieved by trailing a boat-mounted transmitter through the study area. A better evaluation could be attained by making multiple transects throughout the entire region of interest under differing environmental conditions, giving more information on how spatial and temporal changes in ambient noise could affect the transmitter range, and hence the error of the interpolation.

Acknowledgements

The study was funded by contributions from Geoide (GEOmatics for Informed DEcisions, a Canadian Centre of Excellence) the Ministère des Ressources Naturelles et de la Faune du Québec, the Atlantic Salmon Federation, the Centre collégial de transfert de technologie des pêches (CCTTP-CSP, Grande Rivière), the Fondation pour le saumon du Grand Gaspé, ALCAN Inc., and Québec-Océan. We thank P. Brooking and G. Doucette of the Atlantic Salmon Federation for providing equipment and assistance in the realization of this project, and J. Roy of the Société de gestion des rivières du Grand Gaspé for kindly familiarizing us with the York River and giving us access to fishing pools. Finally, we thank those who assisted in fieldwork: D. Fournier, V. Cauchon, J-F. Bourque, O. Deshaies, T. Garneau, N. Harnois, M. Lalonde, J. Lapointe, J. Leclère, A. Remy, A. Richard, and E. Valiquette. This study is a contribution to the scientific programme of CIRSA (Centre Interuniversitaire de Recherche sur le Saumon Atlantique) and Québec-Océan.

References

Becker, R. A., Chambers, J. M., and Wilks, A. R. 1988. The New S Language. Wadsworth and Brooks/Cole Advanced Books and Software, Monterey. 702 pp.

Bowman, A. W., and Azzalini, A. 1997. Applied Smoothing Techniques for Data Analysis: the Kernel Approach with S-Plus Illustrations. Oxford University Press, Oxford. 208 pp.

Carrière, J. B. 1973. A physical oceanographic study of the Havre and Baie de Gaspé. MSc thesis, Dalhouse University, Halifax. 76 pp.

Chanseau, M., Bosc, S., Galiay, E., and Oules, G. 2002. L'utilisation de l'huile de clou de girofle comme anesthésique pour les smolts de saumon atlantique (*Salmo salar* L.) et comparaison de ses effets avec ceux du 2-phenoxyethanol. Bulletin Français de la Pêche et de la Pisciculture, 365/366: 579–589.

Clement, S., Jepsen, D., Karnowski, M., and Schreck, C. B. 2005. Optimization of an acoustic telemetry array for detecting transmitter-implanted fish. North American Journal of Fisheries Management, 25: 429–436.

Egli, D. P., and Babcock, R. C. 2004. Ultrasonic tracking reveals multiple behavioural modes of snapper (*Pagrus auratus*) in a temperate, no-take marine reserve. ICES Journal of Marine Science, 61: 1137–1143.

Ehrenberg, J. E., and Steig, T. W. 2002. A method for estimating the "positional accuracy" of acoustic tags. ICES Journal of Marine Science, 59: 140–149.

Finstad, B., Okland, F., Thorstad, E. B., Bjorn, P. A., and McKinley, R. S. 2005. Migration of hatchery-reared Atlantic salmon and wild anadromous brown trout post-smolts in a Norwegian fjord system. Journal of Fish Biology, 66: 86–96.

Friedman, J. H. 1984. A variable span scatterplot smoother. Laboratory for Computational Statistics, Stanford University Technical Report 5.

Giacalone, V. M., D'Anna, G., Garofalo, G., Collins, K., and Badalamenti, F. 2005. Estimation of positioning error from an array of automated direction receivers in an artificial reef area. *In* Aquatic Telemetry: Advances and Applications. Proceedings of the Fifth Conference on Fish Telemetry held in Europe, Ustica, Italy, 9–13 June 2003, pp. 245–253. Ed. by G. Marmulla, G. Lembo, and M. T. Spedicato. FAO/COISPA, Rome. 295 pp.

Gudjonsson, S., Jonsson, I. R., and Antonsson, T. 2005. Migration of Atlantic salmon, *Salmo salar*, smolt through the estuary Ellidaar in Iceland. Environmental Biology of Fishes, 74: 291–296.

Hastie, T., and Tibshirani, R. 1990. Generalized Additive Models. Chapman and Hall, London. 262 pp.

Hedger, R. D., Martin, F., Hatin, D., Caron, F., Whoriskey, F. G., and Dodson, J. J. 2008. Active migration of wild Atlantic salmon *Salmo salar* smolt through a coastal embayment. Marine Ecology Progress Series, 355: 235–246.

Heupel, M. R., Semmens, J. M., and 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.

Heupel, M. R., and Simpfendorfer, C. A. 2002. Estimation of mortality of juvenile blacktip sharks, *Carcharhinus limbatus*, within a nursery area using telemetry data. Canadian Journal of Fisheries and Aquatic Sciences, 59: 624–632.

Hornik 2007. The R FAQ. ISBN 3-900051-08-9. http://CRAN.R-project.org/doc/FAQ/R-FAQ.html

Klimley, A. P., Le Boeuf, B. J., Cantara, K. M., Richert, J. E., Davis, S. F., and Van Sommeran, S. 2001. Radio acoustic positioning as a tool for studying site-specific behavior of the white shark and other large marine species. Marine Biology, 138: 429–446.

Koutitonsky, V. G., and Bugden, G. L. 1991. The physical oceanography of the Gulf of St Lawrence: a review on the synoptic variability of motion. Canadian Special Publication in Fisheries and Aquatic Sciences, 113: 57–90.

Koutitonsky, V. G., Desrosiers, G., Pelletier, E., Zakardjian, B., Ouellet, D., de Monty, L., Guyondet, T., *et al.* 2001. Études hydrodynamique, sédimentologique et benthique pour le choix de sites de mariculture d'omble de fontaine dans la baie de Gaspé. Institut des sciences de la mer de Rimouski, Rimouski. 207 pp.

Lacroix, G. L., Knox, D., and Stokesbury, M. J. W. 2005. Survival and behaviour of post-smolt Atlantic salmon in coastal habitat with extreme tides. Journal of Fish Biology, 66: 485–498.

Lacroix, G. L., and Voegeli, F. A. 2000. Development of automated monitoring systems for ultrasonic transmitters. *In* Advances in Fish Telemetry, pp. 37–50. Ed. by A. Moore, and I. Russell. CEFAS, Lowestoft, Suffolk, UK.

McCleave, J. D., and Arnold, G. P. 1999. Movement of yellow- and silver phase European eels (*Anguilla anguilla* L.) tracked in the western North Sea. ICES Journal of Marine Science, 56: 510–536.

Moore, A., Ives, S., Mead, T. A., and Talks, L. 1998. The migratory behaviour of wild Atlantic salmon (*Salmo salar* L.) smelts in the River Test and Southampton Water, southern England. Hydrobiologia, 372: 295–304.

Økland, F., Erkinaro, J., Moen, K., Niemelä, E., Fiske, P., McKinley, R. S., and Thorstad, E. B. 2001. Return migration of Atlantic salmon in the River Tana: phases of migratory behaviour. Journal of Fish Biology, 59: 862–874.

Pettigrew, B., Booth, D. A., and Pigeon, R. 1991. Oceanographic observations in Havre de Gaspé during the summer 1990. Canadian Data Report of Hydrography and Ocean Sciences, 100. 94 pp.

Simpfendorfer, C. A., Heupel, M. R., and Collins, A. B. 2008. Variation in the performance of acoustic receivers and its implication for positioning algorithms in a riverine setting. Canadian Journal of Fisheries and Aquatic Sciences, 65: 482–492.

Simpfendorfer, C. A., Heupel, M. R., and Hueter, R. E. 2002. Estimation of short-term centers of activity from an array of omnidirectional hydrophones and its use in studying animal movements. Canadian Journal of Fisheries and Aquatic Sciences, 59: 23–32.

Summerfelt, R. C., and Smith, L. S. 1990. Anesthesia, surgery, and related techniques. *In* Methods for Fish Biology, pp. 213–272. Ed. by C. B. Schreck, and P. B. Moyle. American Fisheries Society, Bethesda, MD.

Thorstad, E. B., Økland, F., Finstad, B., Sivertsgard, R., Plantalech, N., Arne Bjørn, P., and McKinley, R. S. 2007. Fjord migration and survival of wild and hatchery-reared Atlantic salmon and wild brown trout post-smolts. Hydrobiologia, 582: 99–107.

Voegeli, F. A., Lacroix, G. L., and Anderson, J. M. 1998. Development of miniature pingers for tracking Atlantic salmon smolts at sea. Hydrobiologia., 3171/372: 35–46.

Whoriskey, F. G., Brooking, P., Doucette, G., Tinker, S., and Carr, J. W. 2006. Movements and survival of sonically tagged farmed Atlantic salmon released in Cobscook Bay, Maine, USA. ICES Journal of Marine Science, 63: 1218–1223.

Appendix

Here, we list the R-code used in the non-parametric regression methods: (i) kernel estimator (box kernel) with the kernel size determined arbitrarily; (ii) kernel estimator (normal kernel) with the kernel size determined arbitrarily; (iii) cross-validated kernel estimator (normal kernel); (iv) smoothing spline; and (v) polynomial regression. The sm package is required for the cross-validated kernel estimator. The input data are the positions of the VR2 detections as a function of time (Table A1 shows 20 rows of sample data).

Table A1. Sample acoustic-telemetry data.

VR2 Station	VR2 Easting	VR2 Northing	VR2 detection time (minutes past a reference time)
101	391714	5409556	0
101	391714	5409556	0.45
101	391714	5409556	0.98
81	391337	5409399	1.63
101	391714	5409556	1.65
82	391293	5409480	1.78
101	391714	5409556	2.12
82	391293	5409480	2.32
101	391714	5409556	2.67
81	391337	5409399	2.88
82	391293	5409480	3.03
101	391714	5409556	3.38
82	391293	5409480	3.57
101	391714	5409556	3.9
82	391293	5409480	4.25
81	391337	5409399	4.58
82	391293	5409480	4.73
101	391714	5409556	5.07
81	391337	5409399	5.17
82	391293	5409480	5.32

A two-stage procedure is used, implemented separately for easting and northing domains: (i) the positions of the transmitter are estimated at the times of the detections; and (ii) the positions of the transmitter are estimated for a time sequence supplied by the user.

(i) The object X.obs is a vector of the positions (in easting or northing domains) of the VR2 detections. This is the second or third column in Table A1. The object *Time* is a vector of the time of these detections. This is the fourth column in Table A1. The object *X.pred* contains the predicted positions of the transmitter (in either easting or northing domains).

(ii) The object *X.pred.approx* contains the predicted positions of the transmitter at *T.seq*, a sequence of times supplied by the user, and is determined using linear approximation. This procedure may be required if interpolation of the transmitter positions is desired at times other than those of the detections. For example, it may be desirable to interpolate all positions at intervals of 1 min, but the detections may not have coincided with these intervals.

```
library(sm)
# Kernel estimator (box kernel)
bw=5                                  # bandwidth in minutes
X.pred <- ksmooth( Time , X.obs , kernel=c("box") , bandwidth=bw )
X.pred.approx <- approx( X.pred, xout=T.seq)

# Kernel estimator (normal kernel)
X.pred <- ksmooth(Time , X.obs , kernel=c("normal") , bandwidth=bw)
X.pred.approx <- approx( X.pred , xout=T.seq )

# Cross-validated kernel estimator (normal kernel)
hs <- 1.5                     # determined arbitrarily by the user
hcvx <- hcv( Time , X.obs , hstart=hs )
temp <- sm.regression( Time , X.obs , h=hcvx )
X.pred <- data.frame( temp$eval.points , temp$estimate)
colnames(X.pred) <- c("x","y")
X.pred.approx <- approx( X.pred , xout=T.seq )

# Smoothing spline
X.pred <- smooth.spline( Time, X.obs )
X.pred.approx <- approx( X.pred, xout=T.seq )

# Local polynomial
X.pred <- supsmu( Time, X.obs )
X.pred.approx <- approx( X.pred, xout=T.seq )
```

doi:10.1093/icesjms/fsn109