Received: 11 November 2020 | Accepted: 16 February 2021

DOI: 10.1111/2041-210X.13604

RESEARCH ARTICLE

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

Merging computational fluid dynamics and machine learning to reveal animal migration strategies

Simone Olivetti[1] | Michael A. Gil[1,2] | Vamsi K. Sridharan[1] | Andrew M. Hein[1,3]

[1]Southwest Fisheries Science Center, National Oceanic and Atmospheric Administration, Institute of Marine Sciences, University of California, Santa Cruz, CA, USA

[2]Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO, USA

[3]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA, USA

Correspondence
Simone Olivetti
Email: simoneolivetti.eng@gmail.com

Andrew M. Hein
Email: andrew.hein@noaa.gov

Funding information
California Department of Fish and Wildlife, Grant/Award Number: P1896007-01; National Oceanic and Atmospheric Administration, Grant/Award Number: HPCC Incubator Program; National Science Foundaion, Grant/Award Number: IOS Grant 1855956

Handling Editor: Emily Shepard

Abstract

1. Understanding how migratory animals interact with dynamic physical environments remains a major challenge in migration biology. Interactions between migrants and wind and water currents are often poorly resolved in migration models due to both the lack of high-resolution environmental data, and a lack of understanding of how migrants respond to fine-scale structure in the physical environment.
2. Here we develop a generalizable, data-driven methodology to study the migration of animals through complex physical environments. Our approach combines validated computational fluid dynamic (CFD) modelling with animal tracking data to decompose migratory movements into two components, namely movement caused by physical forcing and movement due to active locomotion. We then use a flexible recurrent neural network model to relate local environmental conditions to locomotion behaviour of the migrating animal, allowing us to predict a migrant's force production, velocity and trajectory over time.
3. We apply this framework to a large dataset containing measured trajectories of migrating Chinook salmon through a section of river in California's Sacramento-San Joaquin Delta. We show that the model is capable of describing fish migratory movements as a function of local flow variables, and that it is possible to accurately forecast migratory movements on which the model was not trained.
4. After validating our model, we show how our framework can be used to understand how migrants respond to local-flow conditions, how migratory behaviour changes as overall conditions in the system change and how the energetic cost of migratory movements depends on environmental conditions in space and time. Our framework is flexible and can readily be applied to other species and systems.

KEYWORDS

bionergetics, computational fluid dynamics, machine learning, migration

1 | INTRODUCTION

Migration is an essential part of many animal life cycles (Dingle, 2015). For animals that swim and fly, migration often involves not only long-distance navigation and ecological interactions with conspecifics and predators, but also complex interactions with the physical environment in the form of air and water currents (Dingle, 2015; Flack et al., 2018; Smith, 2012). The way migratory animals interact with abiotic currents can determine the energetic cost of migration (Pennycuick, 2008) and even whether migration is feasible

at all (Alexander, 1998; Pennycuick, 2003). Because climate change and anthropogenic habitat alteration are modifying air and water currents at both small and large scales (Boning et al., 2008; Kling & Ackerly, 2020; Silva et al., 2018), management plans must increasingly consider how human activities influence the physical environment through which migrants travel (Thorstad et al., 2008). There is a growing recognition that managing migratory species must involve managing landscapes to facilitate successful migration (De Lucas et al., 2004; Silva et al., 2018). However, to make informed decisions about how changes to the environment will alter the ability of animals to migrate, we need a deeper understanding of how air and water currents influence migratory physiology and also migratory behaviour.

In the past, efforts to understand how migrants interact with abiotic forcing have tended to take a migration physiology perspective, where the emphasis has been on combining biomechanical models with physiological data to understand the cost of migration in flows (e.g. Martin et al., 2015). For example, classic work on animal migration considered the energetic costs of large-scale mean wind or water currents on the cost of a migratory journey and on the fuel loads required at stopovers, as well as the ranges migrants could achieve under favourable and unfavourable currents (Pennycuick, 2003, 2008). More recently, several studies have analysed physical data or models of wind or hydrodynamics in the context of animal migration (Arenas et al., 2015; Gao et al., 2015; North et al., 2008; Reddy et al., 2016; Weber et al., 2006). Nevertheless, a major outstanding challenge in migration biology is understanding how migrant behaviour and physical forcing by wind and water currents interact to determine how migrants move across a landscape, and the costs they incur when doing so.

One of the limitations of many animal tracking datasets is that only the positions and movements (e.g. via animal-borne accelerometers) of the animal are recorded, and details of the physical environment through which the animal moves are unknown. Because of this, movements must often be studied and interpreted without knowledge of the physical forces and sensory cues that influenced the observed motion of the animal. This severely limits the types of questions about migration behaviour that can be answered with movement data. While modern animal-borne sensors can aid in this problem (Hughey et al., 2018), at present, such sensors are often expensive and too heavy to be carried by small animals. Moreover, animal-borne sensors have the added limitation that they record conditions only in the vicinity of the sensor, leaving the range of conditions available to the animal elsewhere in the environment unknown.

Here, we present an alternative approach to the problem of inferring the physical variables an animal experiences as it moves. This approach combines animal tracking data with high-resolution physical models of the region through which the tracked animal moves. The essential data requirements are (a) animal tracking data describing the physical position of an animal or animals over time, (b) measurements of the structure of the physical environment (e.g. river bathymetry, local landscape topography) and (c) a collection of sample measurements of the physical variables one wishes to model (e.g. local water or wind velocity), preferably collected from the study region over the same range of conditions as those experienced by tracked animals. The latter two data sources are used to build a dynamic model of the physical environment that can then be used to infer the physical forces a tracked animal experienced at each location in the tracking dataset. The end result of fusing animal tracking data with the physical model is a dataset containing positions, velocities and accelerations of each tracked animal (inferred from the tracking data), as well as estimates of the physical forces experienced by the animal at each point in time. Such data can then be used to infer how physical forces influence movement behaviour, and to address a suite of questions related to the energetic output required to produce observed movements.

In what follows, we illustrate how to fuse animal tracking data and physical variables using, as an example, migratory juvenile Chinook salmon migrating through a section of river in the Sacramento-San Joaquin Delta in California. Tracking data consist of high spatial and temporal resolution tracks from salmon as they move through a key segment of the migration route. To model the flow environment these animals experience, we combine river bathymetry data with flow measurements taken in several places throughout the study region to develop a computational fluid dynamics (CFD) model of water flow through the entire study domain. We use the CFD model to estimate the dynamic fluid environment experienced by each individual along its migratory trajectory. We show how this dataset can then be used to estimate the force exerted on the animal by moving water as well as the force produced by the animal through locomotion. Finally, to explore how cues from the physical environment—in this case the flow cues experienced by fish—influence active swimming behaviour, we develop a recurrent neural network model to predict active locomotion as a function of flow cues, and to forecast fish movement trajectories over the near term. Taken together, the elements of our methodology allow one to explore a broad suite of questions about how migrants interact with environmental flows that have been challenging to address in past studies of animal migration. We illustrate several applications of our approach by applying it to questions about navigation behaviour and migratory energetics over a wide dynamic range of flow conditions.

2 | MATERIALS AND METHODS

The methodology we use to integrate tracking data with estimates of the flows animals experience is illustrated in Figure 1. In addition to estimating physical variables at each point in time, the framework includes a step to predict movement behaviour of animals as a function of these physical variables (Figure 1e,f) to determine the extent to which physical variables affect movement decisions. The data inputs to the modelling framework are animal trajectories and the bathymetry and hydrodynamic data needed to build the CFD model (Figure 1a,b). The hydrodynamic data consist of two-dimensional (along-stream and lateral) near-surface river water velocity

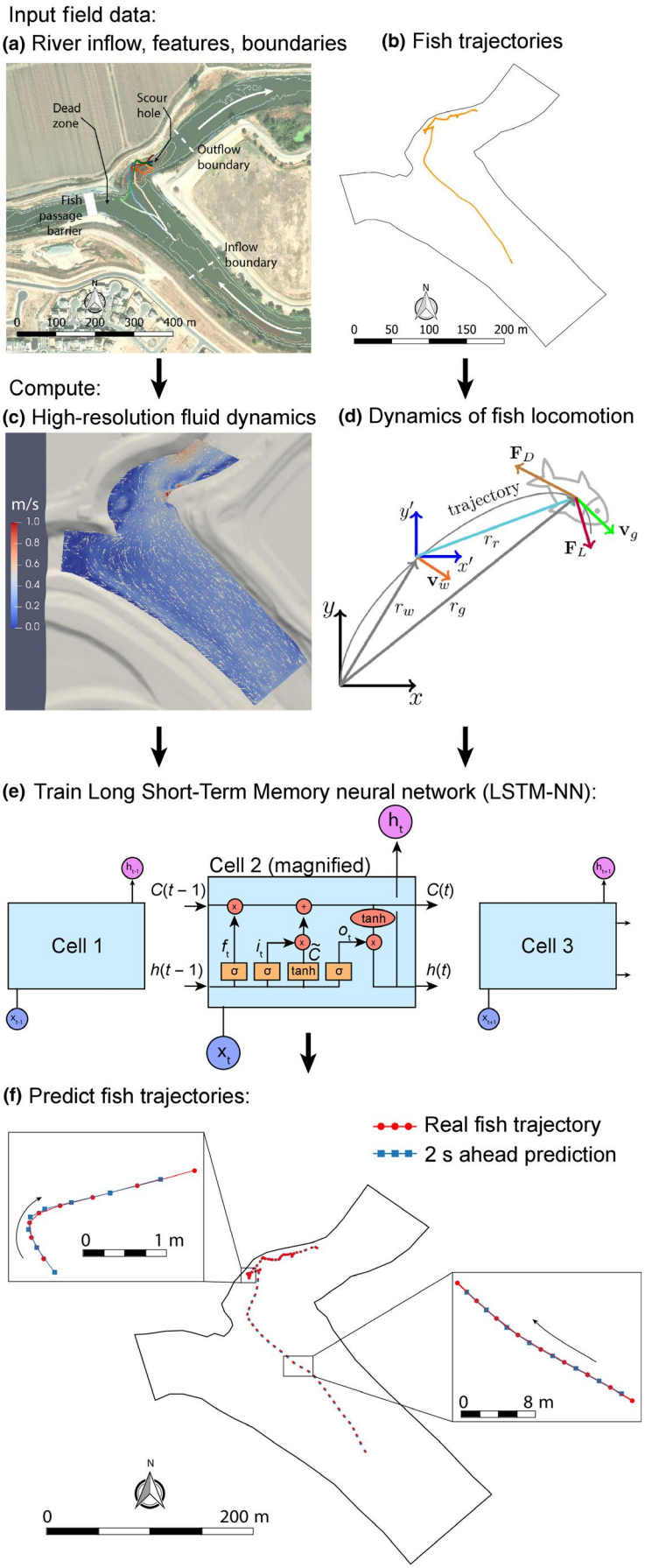

FIGURE 1 Modelling framework. (a) Physical features of the environment and inflow data are collected along with (b) migrant movement trajectories. (c) Physical data are used to build computational fluid dynamic (CFD) modelling of water flow. (d) CFD predictions are combined with observed fish movements used to decompose motion into drag-induced forcing by the flow and active locomotion. (e) The Long Short-Term Memory Neural Network (LSTM-NN) model is developed to forecast locomotion. (f) Locomotion predictions and flow are combined to forecast movement trajectories and predictions are compared to out-of-sample data

measurements collected with four acoustic Doppler current profilers (ADCPs; see Section 2.1 below), and river bathymetry obtained from the 2010 California Department of Water Resources and the US Geological Survey's 2-m-resolution multibeam sonar survey (Wang et al., 2018). Fish trajectories consisted of two-dimensional (alongstream and lateral) tracks obtained from the California Department of Water Resources (see Section 2.1 below).

We use the hydrodynamic data as inputs to simulate the flow field in the section of river system with submeter spatial resolution and 1-s temporal resolution using an unsteady Reynolds-averaged Navier–Stokes (URANS) CFD model (Figure 1a–c). We use the fish trajectories to first quantify the kinematics of motion (i.e. the velocities and accelerations of the fish) and, subsequently, the hydrodynamic information to quantify the dynamics of motion, that is, the drag forces experienced by the fish and the locomotion forces exerted by the fish (Figure 1d). We then model the locomotion force of each individual using the information from the fish trajectories and local hydrodynamic forces by training the neural network model describing fish locomotion behaviour. Subsequently, we employ the trained neural network for multivariate time-series prediction of locomotion forces as a function of the time series of hydrodynamic forces and behavioural responses (Figure 1e). After producing predictions of locomotory behaviour, we used the drag force and the locomotion force predicted by the neural network to predict each individual fish's trajectory (Figure 1f).

2.1 | Field data

Flow and animal tracking data were provided by the California Department of Water Resources. These data were collected through a large collaborative study of a segment of the San Joaquin River within an agricultural and urban watershed in the California Central Valley (study details provided in McQuirk et al., 2015). The spatial locations of fish implanted with acoustic transmitters were inferred using tag detections by a hydrophone array extending over roughly 1 km of the San Joaquin River at the junction with Old River—a tributary—and immediately downstream of the southernmost extent of the Sacramento-San Joaquin Delta. The Delta is an inverted alluvial fan estuary formed at the confluence of the Sacramento River from the North and the San Joaquin River from the South, as well as numerous tributaries. This watershed is used by several species of salmonids of high conservation concern. Subpopulations of Chinook salmon *Oncorhynchus tshawytscha* and steelhead *Oncorhynchus mykiss* traverse portions of the San Joaquin River and the Delta during their juvenile migration to the Pacific Ocean (Williams, 2006), where they mature before returning as adults (see Sridharan et al., 2006 for a detailed description of the hydrometeorology and hydrodynamics in the Delta).

Our study domain includes distinct regions as shown in Figure 1a: (a) a 500 m long reasonably straight prismatic section of the mainstem San Joaquin River about 150 m downstream of a meandering section where the flow is southeast to northwest; (b) a junction at the northwestern region of the straight section where the Old River bifurcates to the west; and (c) a sharp 90° bend eastward in the mainstem San Joaquin River. During the period when the study was conducted, the bifurcation into Old River was blocked by a temporary earthen barrier (white box in Figure 1a). The eastward bend at the northern end of the domain is characterized by an approximately 10-m deep scour hole along the north bank where the flow separates and strongly recirculates before rejoining the freestream along the San Joaquin River (see Appendix D for the bathymetry of study domain).

Two-dimensional near-surface velocity fields were acquired by AECOM Technical Services between 23 April and 30 May 2012 using moored RDI Channel Master side-looking broadband ADCPs operating at 600 kHz. Each cross-section was comprised of 2-m bins, over which point velocity measurements were averaged over several minutes. A 5-m-resolution flow field was reconstructed at 15-min intervals throughout the study domain by first numerically computing the streamlines from the southermost ADCP cross-section and performing an inverse distance weighting interpolation using the velocity vectors obtained from the instrumented cross-sections (Stumpner, 2013a, 2013b). Fish trajectories were obtained from 424 Fall-run Chinook salmon implanted with injectable HTI hydroacoustic tags (M800 and 795Lm models) which were detected at 13 HTI hydroacoustic detectors (model 590) deployed in a two-dimensional array throughout the system. By co-locating fish position using a minimum of four detectors, fish positions were typically estimated at a precision of within 1 m every 2 s (McQuirk et al., 2015). In the present study, we used 184 of these tracks that were sufficiently long to be included in the neural network analysis. We applied our own post-processing pipeline to raw tag detections. This consisted of breaking tracks from each fish into subsegments if subsequent locations were separated by more than 30 s in time. Within each subsegment, we smoothed tracks using a third-order Savitzky–Golay filter with filter length of 22 s. Positions were also interpolated to a regular time interval of 2 s between subsequent locations.

2.2 | Fish behaviour

2.2.1 | Movement kinematics

The first step in our workflow is describing the kinematics of fish movement. The accuracy of position data in the depth dimension was poor, likely due to constraints on the positioning of hydrophones determined by the relatively shallow average depth of the study region (McQuirk et al., 2015). As a result, we were unable to study the movements of fish in the depth dimension, and we retained only the horizontal coordinates of the position of each fish. Accordingly, tracks are represented as two-dimensional trajectories through the river section, and we consider only horizontal components of the fish kinematics and dynamics. Henceforth, we assign the east–west direction as the *x*-dimension and the north–south direction as the *y*-dimension. To keep track of the relative motion of fish and flowing

water, we define two reference frames, namely an inertial frame (x, y) fixed at a point on the river bank and a relative frame (x', y') moving along the fish trajectory with water velocity $\mathbf{v}_w$ (Figure 1d). Given these reference frames, the position of a fish can be defined as follows:

$$\mathbf{r}_g = \mathbf{r}_r + \mathbf{r}_w. \quad (1)$$

Here $\mathbf{r}_g$ is the fish position with respect to the inertial frame (x, y), $\mathbf{r}_r$ is the fish position with respect to the relative frame (x', y') and $\mathbf{r}_w$ is the position of the relative frame with respect to the inertial frame. By recursively differentiating Equation 1 with respect to time we obtain the velocity $\mathbf{v}_g$ and acceleration $\mathbf{a}_g$ of each fish as follows:

$$\mathbf{v}_g = \mathbf{v}_r + \mathbf{v}_w, \quad (2)$$

$$\mathbf{a}_g = \mathbf{a}_r + \mathbf{a}_w, \quad (3)$$

$\mathbf{v}_g$ and $\mathbf{a}_g$ are the velocity (or overground velocity) and acceleration of fish with respect to the inertial frame; $\mathbf{v}_r$ and $\mathbf{a}_r$ are the relative velocity and acceleration of fish with respect to the relative frame; and $\mathbf{v}_w$ and $\mathbf{a}_w$ are the velocity and acceleration of the relative frame with respect to the inertial frame. The latter quantities can also be interpreted as velocity and acceleration of a water parcel along the fish's trajectory. Equations 2 and 3 are useful to decompose the fish motion (see Section 2.2.2 below).

2.2.2 | Movement dynamics

Once the kinematics are defined, we subsequently apply the momentum equation (i.e. Newton's second law of motion) to each fish to quantify its movement dynamics. In the horizontal plane, we identify two forces for each fish, namely locomotion force $\mathbf{F}_L$ and drag force $\mathbf{F}_D$ (Figure 1d). We assumed that vertical forces such as gravitational force and buoyancy balance each other resulting in null vertical acceleration. Defining the mass of fish as m_{fish}, the fish dynamics can be summarized as

$$m_{\text{fish}}\mathbf{a}_g = \mathbf{F}_L + \mathbf{F}_D. \quad (4)$$

The drag force acts opposite to the relative motion of the fish moving with respect to the surrounding flow and it can be defined (Hoerner, 1965) as

$$\mathbf{F}_D = -\frac{1}{2}\rho_w A_f C_d \,||\mathbf{v}_g - \mathbf{v}_w||(\mathbf{v}_g - \mathbf{v}_w), \quad (5)$$

where ρ_w is the water density, A_f is the wetted area of fish and C_d is the drag coefficient, see Appendix A for how we calculate C_d and A_f. The term $\mathbf{v}_g - \mathbf{v}_w$ is the fish relative velocity $\mathbf{v}_r$ with respect to the relative frame (see Equation 2). The locomotion force can then be calculated by inverting the momentum equation (see Equation 4). For this approach to be useful for understanding how instantaneous fish behaviours contribute to their overall migration trajectories, we need information on the drag force at a spatial and temporal resolution commensurate with the tracking data. While m_{fish} can be obtained from the metadata associated with the tracking experiments and $\mathbf{a}_g$ can be directly obtained from the tracking data, $\mathbf{F}_D$ cannot be calculated at the desired resolution from the 15-min 5-m-resolution interpolated ADCP $\mathbf{v}_w$ fields. We therefore developed the CFD model of the river system to estimate $\mathbf{v}_w$, and used this estimate to infer $\mathbf{F}_D$ and compute $\mathbf{F}_L$. The details of the CFD modelling are described in the following sections.

The tracking data consist of 184 fish tracks for a total of 129,830 location points with a standardized temporal resolution of 2 s. We show several example tracks in Figure 1b. Given the fish position $\mathbf{x}_g(t_n)$ from each track, the fish velocity with respect to the inertial frame (see Section 2.2.1) is $\mathbf{v}_g(t_n) \approx (\mathbf{x}_g(t_{n+1}) - \mathbf{x}_g(t_n))/\Delta t$, where $t_n = [2, 4, 6, \ldots]$ and $\Delta t = 2s$. The fish velocity with respect to the relative frame is obtained by reversing Equation 2 such that $\mathbf{v}_r(t_n) = \mathbf{v}_g(t_n) - \mathbf{v}_w(t_n)$. $\mathbf{v}_w(t_n)$ is computed from the CFD results for each fish track (see Section 2.3). Consequently $\mathbf{a}_r(t_n) \approx (\mathbf{v}_r(t_n) - \mathbf{v}_r(t_{n-1}))/\Delta t$. With the kinematics defined thus, it is now possible to calculate the locomotion force for each fish by combining Equations 4 and 5 such that

$$\mathbf{F}_L(t_n) = \frac{1}{2}\rho_w A_f C_d \,||\mathbf{v}_r(t_n)||\mathbf{v}_r(t_n) + m_{\text{fish}}\mathbf{a}_g(t_n) = f(\mathbf{v}_g(t_n), \mathbf{v}_w(t_n)). \quad (6)$$

It is important to notice that $\mathbf{F}_L(t_n)$ is a function of $\mathbf{v}_g(t_n)$ and $\mathbf{v}_w(t_n)$ as shown in Equation 6.

2.3 | Hydrodynamic variables

The next step in our workflow is to compute the drag force $\mathbf{F}_D$ on the fish. Since $\mathbf{F}_D$ is a function of $\mathbf{v}_w$ (see Equation 5), we simulated the flow dynamics of the river using a three-dimensional CFD model based on URANS equations. The river flow is considered incompressible and isothermal with the deflection of the water surface being represented by a two-phase water-air volume of fluid (VOF) model. We used the openFOAM solver interFoam (Deshpande et al., 2012) to develop this model. Although the tracking data we used are two-dimensional, we constructed a three-dimensional CFD model to realistically represent the statistics of the turbulence and the flow dynamics at the scour hole and in regions near the channel banks. We assumed tracks were located within the uppermost cell of the CFD volume corresponding to approximately 0.3 m below the water surface.

2.3.1 | Solver and model parameters

The interFoam solver in openFoam implements the continuity and momentum equations for isothermal and incompressible flows along with an additional equation tracking the fraction of air within each parcel of water. The URANS models require turbulence closure equations in order to be a well-posed PDE system (Menter, 1994). We used the $k-\omega$ equations to represent the statistics of the

unresolved turbulence. The boundary conditions for the velocity and the water elevation are based on field measurements (see Section 2.1). The empirical flow velocity time series is available at the inlet section for the 3-month period from March to May with a time resolution of 15 min supersampled linearly at 2-s intervals.

2.3.2 | Modelling active locomotion: A neural network approach

The final step in our workflow (Figure 1e,f) is to develop a model describing how fish locomotion depends on features of the environment, including the hydrodynamic forces the animal experiences as it moves through the water. The details of sensory integration, processing and decision-making during navigation are poorly understood for most migratory species, including migratory fishes. To avoid making ad hoc assumptions that might arbitrarily restrict the form of the relationship between physical variables and movement behaviour, we modelled the effects of flow on movement behaviour using a flexible approach for time-series prediction, the Long Short-Term Memory Neural Network (LSTM-NN).

We selected the LSTM-NN as a reasonable model of movement behaviour for the following two reasons: first, in the past, LSTMs have been used successfully to model movements of vehicles and pedestrians (e.g.; Altché & De La Fortelle, 2017; Xue et al., 2018). Second, there is detailed documentation in the literature (Kang & Choi, 2005) on how LSTMs are implemented in TensorFlow (Abadi et al., 2015). This existing software implementation makes LSTMs a convenient modelling tool for describing the relationship between physical variables and migrant behaviour when no a priori model exists. Details of the underlying structure of the LSTM and how it maps inputs to outputs are given in the Appendix D. In the Discussion, we further elaborate on the *pros* and *cons* of LSTM and the situations in which it is likely to provide a good model of navigation behaviour.

In the current application, we use the LSTM to predict the locomotory force produced by migrating fish at each time step. We take, as input to the network, the overground velocity of fish, $\mathbf{v}_g$, and the water velocity, $\mathbf{v}_w$, because $\mathbf{F}_L = f(\mathbf{v}_g, \mathbf{v}_w)$ as shown in Equation 6. This assumes the fish could measure overground velocity, which could be accomplished, for example, through visual means, by estimating the optic flow of visual features on the benthos (e.g. the river bed itself, submerged debris or aquatic vegetation). In the past, environmental variables such as water acceleration, hydrostatic pressure (Goodwin et al., 2014), turbulent structures (Lacey et al., 2012), turbulent kinetic energy intensity (Gao et al., 2015) and circulation around the fish (Oteiza et al., 2017) have been used to explain fish movement behaviours. We decided to use the water velocity experienced by the fish because the river system under consideration is characterized by a relatively low turbulent kinetic energy content, and because other mechanisms of behavioural response to variables such as the local shear or circulation are not understood in complex environmental flows. Moreover, exploratory analyses including other variables in LSTM-NN training did not indicate improved performance.

The resulting trained LSTM-NN is a function that relates the overground velocity and water velocity experienced by a migrating fish at some time t_{n-1} to the locomotion force produced by that fish at time t_n:

$$\mathbf{F}_L(t_n) = \mathrm{LSTM}(\mathbf{v}_g(t_{n-1}), \mathbf{v}_w(t_{n-1})), \tag{7}$$

where t_n is the discrete time step with $n = [0, 1, \ldots, N-1, N]$. We note that the use of $\mathbf{v}_g$ in this formulation allows us to explicitly model the locomotion of the fish as a function of its memory of its response to the local environment, as well as its current sensory experience. Details of LSTM-NN structure and how inputs map to predictions are given in Appendix C.

2.4 | LSTM-NN fitting, predictions and out-of-sample testing

We used the LSTM-NN module available in TensorFlow (Abadi et al., 2015) for predicting $\mathbf{F}_L$. The training dataset consisted of the time series of overground velocities of fish and water velocities along the fish tracks. Furthermore, we used the time series related to the observed components of the locomotion force $F_{L_x}(t_n)$ and $F_{L_y}(t_n)$ computed with the field data, see Equation 6, as reference output for the LSTM-NN training. We optimized the LSTM-NN settings to minimize the average error of ΔF_{L_x} and ΔF_{L_y}, where Δ is the difference between the predicted and actual value. We tested a number of LSTMs-NNs with an increasing number of cells and used the k-nearest neighbour method (Arya et al., 1998) to select the architecture with the optimal number of cells (see Appendix C). We found an LSTM-NN with 112 cells to be the optimal configuration, because it produced ΔF_{L_x} and ΔF_{L_y} with minimal average error. After the end of the cascade of LSTM-NN cells, we included a dense layer of two rectified linear activation functions, *ReLU*, to output the model results (Abadiet al., 2015). The length of the training dataset was 60% of the original dataset subdivided in 72 batches; the total length of the dataset consists of 129,830 data points. We trained the LSTM over 30 epochs.

2.5 | Forecasting fish movements

Once the LSTM-NN model of $\mathbf{F}_L(t_n)$ is fitted to training data, it can be used to predict migrant trajectories by applying the forward Euler method to Equation 4 as follows:

$$m_{\mathrm{fish}} \frac{\mathbf{v}_g(t_n) - \mathbf{v}_g(t_{n-1})}{\Delta t} \approx m_{\mathrm{fish}} \mathbf{a}_g(t_n) = \mathbf{F}_L(t_n) + \mathbf{F}_D(t_n), \tag{8}$$

Hence, considering Equations 2 and 5

$$\begin{cases} \mathbf{F}_L(t_n) = \mathrm{LSTM}(\mathbf{v}_g(t_{n-1}), \mathbf{v}_w(t_{n-1})) \\ \mathbf{v}_g(t_n) = \mathbf{v}_g(t_{n-1}) + (\mathbf{F}_L(t_n) + \mathbf{F}_D(t_n)) \dfrac{\Delta t}{m_{\mathrm{fish}}} \\ \mathbf{x}_g(t_n) = \mathbf{x}_g(t_{n-1}) + \mathbf{v}_g(t_n) \Delta t \end{cases}. \tag{9}$$

The initial conditions $\mathbf{v}_g(t_0)$, $\mathbf{v}_w(t_0)$ and $\mathbf{x}_g(t_0)$ are determined from the corresponding field data. This scheme can be used both to predict fish velocities and trajectories in-sample, and to predict entirely new trajectories, given the appropriate input data.

3 | RESULTS

3.1 | CFD results

We used the CFD model of the study domain to compute flows over the duration that fish were present. In Figure 2a, we show a snapshot of the water velocity field in the horizontal section near the water surface (where the fish trajectories are assumed to be contained). The contour colours represent the water velocity magnitude, while the vectors represent the direction of local flow. The southeast region close to the inlet is characterized by a flow that tends to be uniform. In contrast, the northwest region close to the barrier shows a large area of flow recirculation; two counter-rotating vortexes appear along the barrier (Figure 2b). A vortex rotating in the anticlockwise direction on the northern bank is visible in Figure 2c. The formation of this vortex is due to the sharp bend of the river course and associated scour hole, causing the flow to recirculate along the north bank.

We validated the CFD model by comparing the velocity profiles from the numerical simulation against the velocity profiles from the field measurement; we show in Figure 2d,e that the CFD results (lines) are in good agreement with the ADCP measurements from two cross-sections which include a typical variation of ±5.8 cm/s within each velocity bin (dots; McQuirk et al., 2015).

3.2 | Fish migration behaviour and LSTM model predictions

The tracking data provided are an extensive collection of fish velocity and trajectory estimates from across the study domain. By applying the velocity decomposition introduced above to the fish trajectory data and CFD-generated flow velocity predictions, we were able to estimate the distinct contributions of water flow and migrant locomotion to the observed overground velocity of each migrating animal. In Figure 3b, we show the probability density function (*pdf*s) of the magnitudes of the fish overground velocity, $||\mathbf{v}_g||$, and the fish relative velocity $||\mathbf{v}_r||$ (i.e. the animal's velocity relative the the moving water), as well as the magnitude of water velocity at observed fish locations $||\mathbf{v}_w||$. Note that the overall magnitude of relative velocity of the fish—the component of velocity due to active locomotion—often exceeds the magnitude of water velocity, indicating that fish regularly swim at speeds that are higher than the speeds of the flows in which they are swimming. This can be seen more directly in the distribution of the ratio of relative velocity magnitude

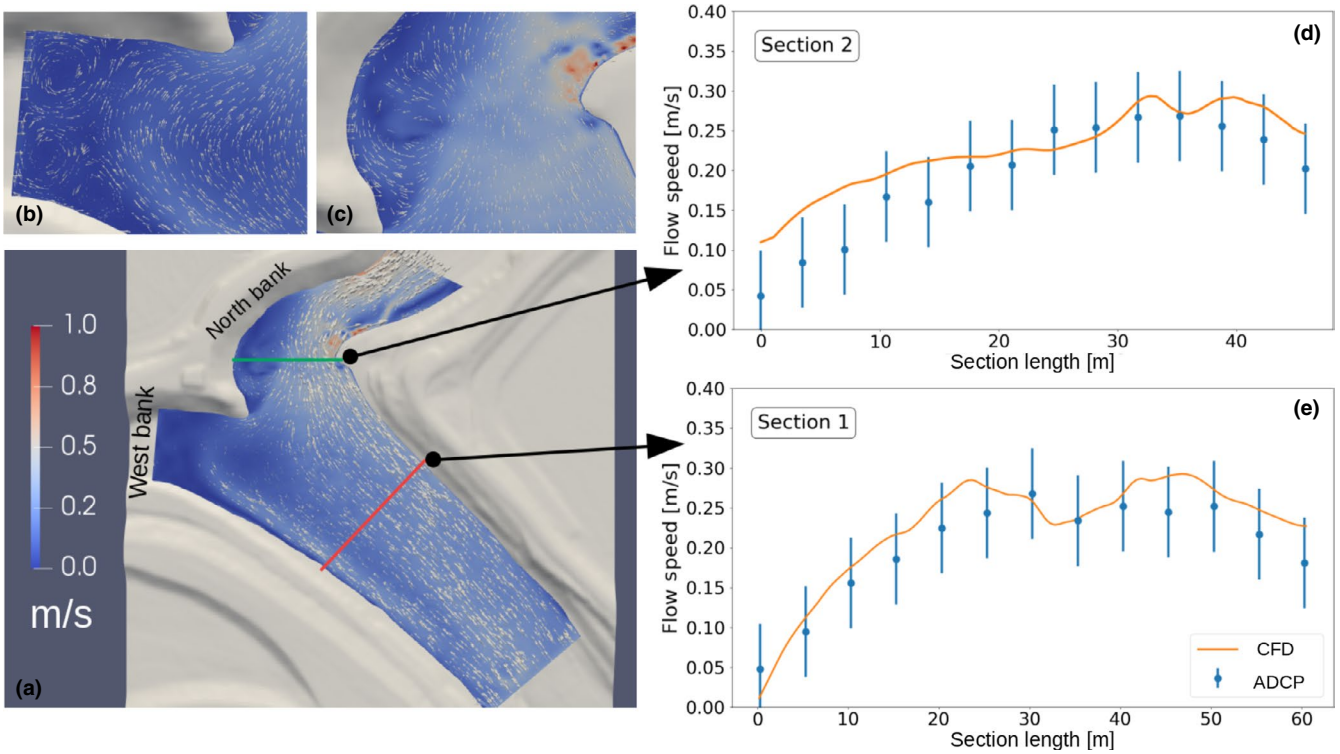

FIGURE 2 (a) Snapshots of the velocity field magnitude at one point in time. The colour bar indicates flow magnitude in units of m/s. Lines through the domain show cross-sections used for model validation. Red line: Section 1 (shown in panel e). Green line: Section 2 (shown in panel d). (b) Zoomed in view of the western bank showing regions of weak recirculation flow. (c) Zoomed in view of the northern bank showing a vortex. (d) Comparison between CFD flow predictions (line) and water velocity magnitude measured by ADCP (blue dots and error bars) in Section 2. Profiles averaged over 30 min. (e) Comparison between CFD predictions and data

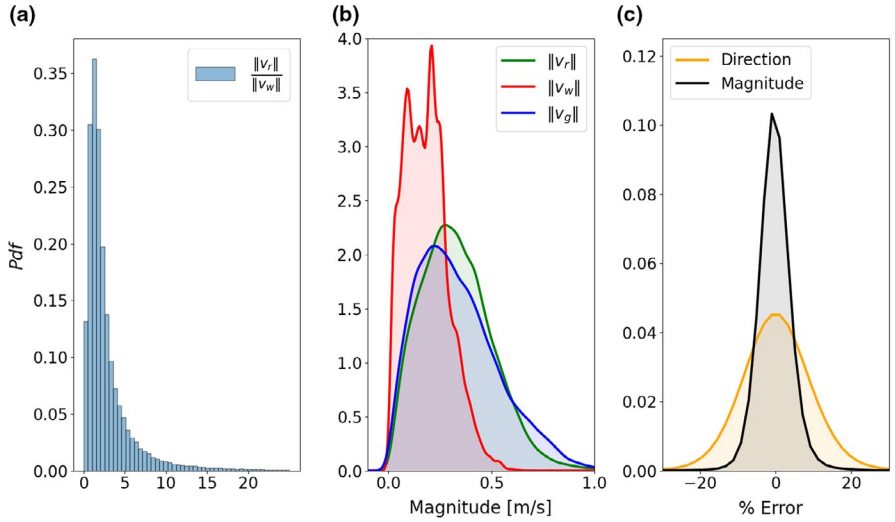

FIGURE 3 Empirical velocity data and LSTM-NN prediction performance. (a) Probability density function (*PDF*) of the ratio of the magnitudes of the fish relative velocity to the water velocity. (b) PDFs of magnitude of the fish relative and overground velocity (green and blue distributions respectively) and PDF of water speed at fish locations computed from CFD model (red distribution). (c) PDFs of prediction errors from the LSTM model shown as percentage error in predicted direction (orange distribution) and magnitude (grey distribution) of locomotion force

to the overground velocity magnitude, Figure 3a. The right tail of this distribution shows cases where fish are swimming at speeds that far exceed the speed of local water movement.

Employing Equation 6, velocity estimates can be used to estimate the locomotory force produced by each fish to achieve its observed motion. The LSTM-NN model of locomotion accurately predicted this locomotory force in the 51,932 data points (40% of the original dataset) that were held out during training (Figure 3c). Typical errors for direction are within 20% of observed values, and magnitude estimates are typically accurate to well within 10% of observed values (Figure 3c). Our results indicate that our model of fish swimming behaviour is able to predict this behaviour for times and locations on which the model was not trained (i.e. on the out-of-sample data).

Given a prediction for the locomotory force, the equation system in Equation 9 can be used to predict a fish's trajectory, $\mathbf{x}_g(t_n)$, in addition to the locomotion forces, accelerations and velocities.

We show the distributions of error in predicting position prediction measured in body length for several time ahead predictions, for example, from 2 s up to 30 s (predictions shown are for 51,932 data points held out-of-sample during training) in Figure 4a; while the tail of the error distribution includes significantly larger errors as the prediction horizon increases, the mode of the error distribution only grows by roughly one body length when moving from a prediction horizon of 2 s to a horizon of 30 s. Red and blue distributions in Figure 4a show 2 s-ahead and 6 s-ahead predictions, illustrating that increasing the forecast horizon from 2 to 6 s does not result in a dramatic decrease in the quality of predictions. Nevertheless, the discrepancies between the observed and predicted trajectories do continue to grow as the prediction horizon is increased as one would expect. In Figure 4b, we show the dependence of the mean and standard deviation of the error in predicting position on the forecast horizon. Up to forecast horizons of 30 s, the mean prediction error remains below four fish body lengths. It is worth noting that the mean and standard deviation of prediction error represent a small fraction of the typical travel distance during any given forecast horizon. For example, in 30 s the average travel distance is 91.1 body lengths while the mean error is about 3.5 body lengths (green scale in Figure 4b). In Figure 4c, we show a sequence of predictions along the length of a long trajectory. In Figure 4c, the blue line is the actual trajectory of a tagged fish while the red dots are the 2 s-ahead predictions; this fish trajectory consists of 455 points corresponding to 906 s of the fish's trajectory through our study region. We zoom into two parts of the trajectory which are structurally different from each other, namely a relatively straight section in Figure 4d and a sharply curving section in Figure 4e. In these plots, the black dots are initial locations to initialize the model in Equation 9. In both sections of the track, there is a close alignment between the observed and predicted trajectory points. In Figure 4f–g, we show the same sections of the track for 6 s-ahead predictions (three time steps ahead); while the accuracy tends to decrease as the prediction horizon increases, errors remain reasonably bounded, even in the highly curved region of the trajectory.

By combining the CFD model to predict flow and the LSTM-NN to predict fish locomotion in response to flows, one can explore a wide range of questions about how flow and locomotory behaviour of animals interact under different conditions. For example, by exploiting the first two equations in Equation 9, it is possible to estimate the movements of fish across the entire flow domain for different environmental conditions of interest. In Figure 5a, we show snapshots of the water velocity vector field in the river system during the period of lowest and highest outgoing flows respectively. In Figure 5b, we show the 'relative swimming velocity', R_{fw}, defined as the ratio of the magnitudes of relative velocity of fish and the water velocity, for the same flow conditions shown in Figure 5a. Two key patterns are immediately evident. First, the relative swimming velocities of the fish regularly exceed water velocity throughout much of the domain. This predicted spatial pattern is consistent with the empirical observation shown in Figure 3a,b that the observed relative swimming velocities regularly exceed water velocity. Moreover, this demonstrates the degree to which fish movements appear to be driven by active swimming behaviour rather than simple passive forcing by the flows. The second pattern evident in Figure 5b is that there is a strong spatial heterogeneity in the ratio of fish to water velocities, and these spatial patterns change as the overall flow transitions from weak to strong. For example, during low flows, the relative

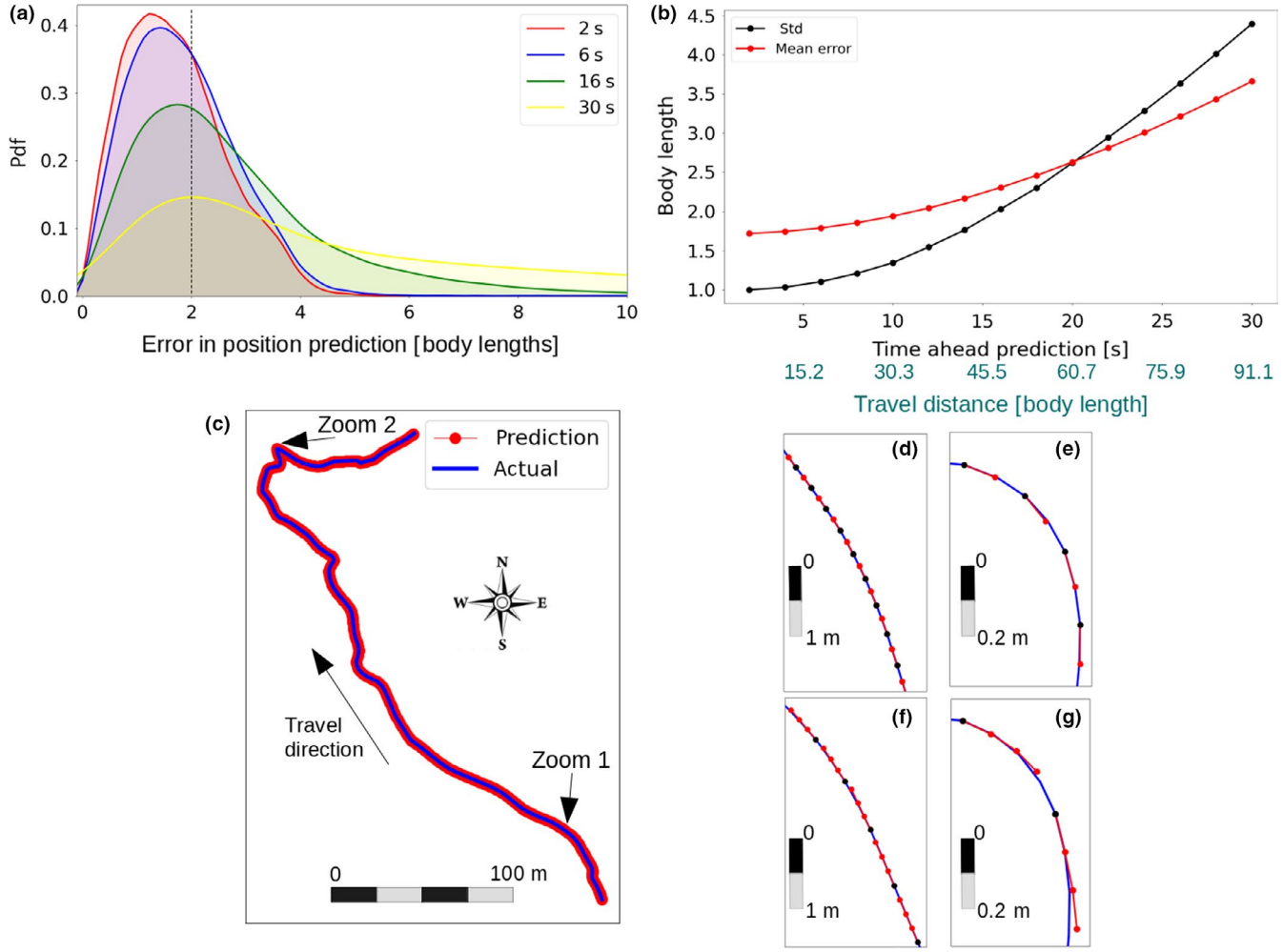

FIGURE 4 Trajectory prediction performance over different forecast time horizons. (a) *Pdf*s of the error in predicting the position of the fish [in body lengths] for several of the prediction horizons. The vertical line at two body lengths indicates that the mode of the error in predicting fish positions is well-contained for even large forecast time horizons. (b) Mean error (red) and standard deviation (black) in predicting the position over forecast horizon (time in seconds) and average distance travelled by the fish (scale in green [in body lengths]). (c) Example of a single trajectory prediction. Blue line shows observed trajectory; red points show predicted trajectory for 2 s-ahead predictions. (d) Zoomed into a straight section of the track (zoom 1) for 2 s-ahead prediction. Black points show an initial location of the fish from which the Equation 9 is initialized. Red points show predicted location 2 s-later. (e) Zoomed into a curved section of the track (zoom 2) for 2 s-ahead prediction. (f and g) 6 s-ahead predictions zoomed into the sections shown in (d and e) respectively. Colours and symbols in (e, f and g) are as in (d)

swimming velocity is greatest in the open channel, upstream of the bend. During high flows, relative swimming velocity is much slower in this same region. During high flow, the relative velocities of fish are smaller than the water velocity in the regions of high circulation near the channel bifurcation, whereas this pattern is not evident at low flows.

To determine the impact this spatial and flow-dependent variation in behaviour has on migration energetics, we can estimate the rate of power output required to achieve predicted movements across the domain. We do this by defining a quantity we will call 'locomotory scope', $S_L = 1 + \mathbf{F}_L \cdot \mathbf{v}_r/\mathrm{RMR}$, which characterizes the power output required to fuel resting metabolism and locomotion, normalized by the resting metabolic rate (RMR; see Appendix A for RMR calculation). The locomotory scope is a measure of the power output of an animal measured in units of resting metabolic rate. Thus, a locomotory scope of 1 corresponds to a case where an individual devotes no power to locomotion, whereas a value of 5 corresponds to a case where the total rate of power output (including resting metabolism) is five times resting metabolic rate. Note that locomotory scope as it is defined here is not the same as aerobic scope because we neglect any power loss due to inefficiencies in force production, and we do not consider other sources of power consumption (e.g. specific dynamic action) that could be relevant during migration. Thus, locomotory scope should be taken as a lower bound on the relative power output required during movements.

Our analysis of predicted locomotory scope reveals strong spatial patterns in power output as well as strong differences in patterns across specific instances of low and high-flow conditions (Figure 5c). Under low-flow conditions, locomotory scope was generally below 1.5, indicating that the power required to produce predicted

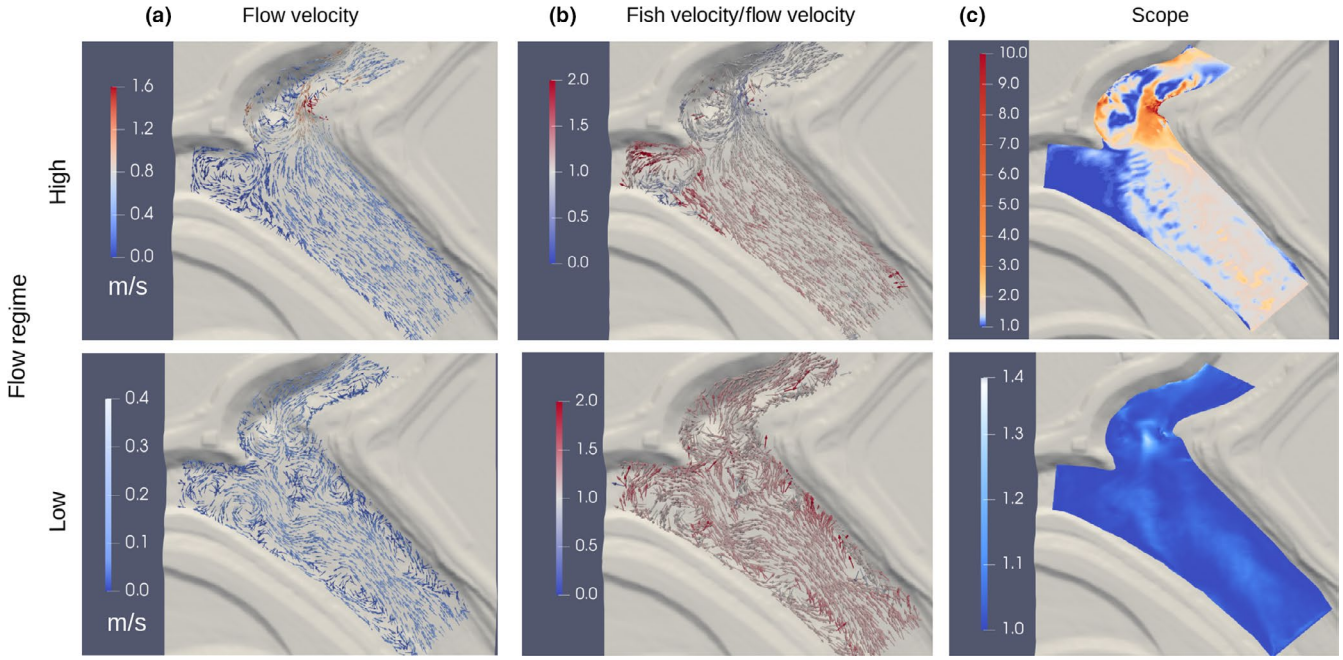

FIGURE 5 Predicted water flow, velocity ratio and locomotory scope under high (upper panels) and low (lower panels) net flow conditions. (a) Flow velocity field. (b) Ratio of the relative velocity of the fish to the water velocity. (c) Locomotory scope, S_L. Note large differences in range of predicted scope between high (upper) and low (lower) flow conditions

migratory movements across the domain was generally less than half of the resting metabolic rate of migrants. The highest relative power outputs were predicted to be along the centre of the main channel, upstream of the channel bend and in a region of relatively high circulation near the eastward channel bend (Figure 5c). In strong contrast to these patterns, locomotory scope during high flows was as large as 10 in some regions of the domain, indicating that predicted movements in those regions required total power outputs 10 times higher than resting metabolic rate. The highest rates of power output were predicted to be in the region of strong flow recirculation near the eastward channel bend and along the eastern bank in the same region. However, even outside of these regions the locomotory scope exceeded a value of 1.5 throughout much of the domain. One might expect migrants to take advantage of strong oceanward flows in the high-flow conditions by drifting passively rather than swimming actively. In contrast, our model suggests that migrants generally use far more power under high-flow scenarios, particularly in local regions of strong unsteady flow. This is due, at least in part, to the fact that in weak flows, the locomotion force tends to be biased in alignment with the direction of local water flow (Appendix A), whereas in more powerful flows, fish locomotion is aligned opposite or orthogonal to the direction of local water flow.

4 | DISCUSSION

Here, we have developed a general methodology to combine quantitative estimates of a turbulent environment with measurements of the movements of animals to better understand migratory behaviour in the wild. Our methodology combines animal tracking data with high-resolution physical modelling of environmental flows—here achieved using CFD—to estimate the dynamic flow environment migrants experience and determine the component of force of migratory movements due to active locomotion by the animal. Finally, we employ recurrent neural network methods to relate the physical conditions experienced by the migrant to locomotion behaviour, and use this model to forecast movements over times and conditions outside those included in the training data.

Reconstruction of the local physical environment and decomposition of active and passive components of movement have the potential to offer new insights into the processes that influence the movements of migrating animals in complex environments. This approach extends recent work to characterize how migrants move in relation to coarser scale environmental flows such as water currents and regional wind patterns (e.g. the use of favourable prevailing winds and fast air streams by migrating insects, Alerstam et al., 2011). Our approach also holds significant promise as a tool for the management of migratory species because, after careful testing on out-of-sample data, our framework allows one to make predictions about both the physical *and* behavioural consequences of modifying the migratory environment, for example, by raising or lowering flow, altering the bathymetry or course of the river, or installing equipment such as water diversion facilities along the migration route (Silva et al., 2018; Thorstad et al., 2008). Although we have applied our framework to migratory fish in a river system, the same methods could be used to understand migratory strategies of flying species by combining high-fidelity tracking during flight (Ling et al., 2018) with CFD modelling of environmental features of interest (e.g. wind turbines, Martin et al., 2017) or physical modelling of turbulent convective flow in the atmospheric boundary

layer (Reddy et al., 2016). It is well known that flying animals also respond to local air flows (Scacco et al., 2019, Shepard et al., 2016 and Dabiri, 1993); however, constructing and validating models of air flow poses some unique challenges. For example, it is often challenging to collect high-resolution time-varying data on air currents in the atmosphere that can be used during model validation. For small-scale flow phenomena such as boundary layer flows over localized topography and built-up areas on the order of a few hundred square-metres to a few square kilometres, wind-tunnel experiments over downscaled models can provide validation datasets for high-resolution URANS and LES models of atmospheric flow (e.g. Jimenez & Moser, 1998; Kellnerová et al., 2018). For flows distributed over larger open areas, on the order of tens of square kilometres, a combination of wind vane, flux tower and LiDAR and Radar measurements may be used to produce reliable estimates of the air currents (Friedrich et al., 2012; Madala et al., 2015). A more recent alternate strategy has been to dynamically downscale global circulation models to spatial–temporal resolutions required for regional-scale analysis and validate these downscaled models using a regional network of weather stations (Wagenbrenner et al., 2016; Winstral et al., 2017). For many physical modelling methods, open source software packages are readily available (e.g. openFoam), as are packages for constructing statistical models (e.g. R, TensorFlow) of migration behaviour once the locomotion component of migratory movements has been computed (Figure 1e,f).

Computational fluid dynamics, and computational modelling of the flow environment more generally, have already proven to be useful for studying environmental flows in the context of animal migrations. For example, Gisen et al. (2016) developed a 3D CFD model of a hydropower dam tailrace using a Detached-Eddy Simulation is the conventional case to evaluate impacts on migrants. Reddy et al. (2016) developed a computational model of thermals in the atmospheric boundary layer to study how soaring birds navigate complex turbulent motion of air. Gualtieri et al. (2019) modelled fish migration through a river system as particles characterized by two bioenergetic parameters, one related to the drag force a fish experienced and one related to the energy needed by a fish to remain in a specific location. Similar assumptions were adopted by Ramón et al. (2018) who studied the hydrodynamic drivers of juvenile salmon movements using CFD to compute the flow field across a river system. Although Gualtieri et al. (2019) and Ramón et al. (2018) modelled fish as passive particles dragged by the river flow, as we show here, even small migratory fish can swim very actively, and in many cases, their locomotion force production is significant. Indeed, our analysis of relative velocity of fish and water (Figure 5) shows that the component of ground speed due to active locomotion is often greater in magnitude than the water speed, even in relatively fast flows. Our findings corroborate results from other systems (e.g. Arenas et al., 2015), and suggest more generally that even small migratory animals such as the juvenile salmon considered here (mean length 112 mm) spend significant amounts of energy on locomotion, even when the net direction of environmental flow aligns with the direction of migration.

Several researchers have begun using CFD models to attempt to understand how migrants navigate complex physical environments at spatial and temporal resolutions similar to those considered in our study. For instance, Goodwin et al. (2014) used a steady-state RANS CFD model to compute water field velocity in combination with an ad hoc fish behavioural model to represent fish movements in the vicinity of hydropower facilities. Gao et al. (2015) used a similar approach for a slot fishway, applying a parametric model of fish movement. Martin et al. (2017) combined a CFD model of a wind turbine and aerodynamic modelling of bat flight to understand how flying bats might interact with the forces produced by wind turbines. In the present study, we extended the approaches of these past models by developing a URANS CFD model to compute time-dependent flow variables. We employed a time-dependent CFD model because the flow field through complex channel morphologies like the one studied here can be extremely dynamic, particularly in river and estuary systems where flows can change due to a variety of reasons including precipitation, effects of tides, sudden storms and floods, and local water diversions and runoff. A dynamic, time-varying CFD model allows us to model changes in flows that occur as inflows change. In general, a dynamic model will be necessary to correctly decompose drag and locomotion forces when the flow field changes appreciably over time. Not accounting for changes in flow will lead to biased estimates of these components.

After we validated that model against empirical flow measurements, we used flow estimates, along with observed fish migration trajectories, to infer the drag and locomotory forces that produced observed fish accelerations. Rather than prescribing an ad hoc model of locomotion behaviour, we used a flexible recurrent neural network model (the LSTM-NN) to describe how flow cues and past behaviour influence locomotion behaviour in the near future. Importantly, this approach provides accurate near-term forecasts of migrant behaviour on out-of-sample data. Thus, our model both captures observed patterns of locomotion in complex flows, and is capable of making accurate out-of-sample predictions to evaluate hypothesis about the implications of migratory behaviour across space and over ranges of environmental conditions (e.g. Figure 5).

To predict swimming behaviour, we relied on a flexible multivariate time-series method. Multivariate time-series analysis methods such as the LSTM-NN have become popular in many fields including health care (Kang & Choi, 2014), phoneme classification (Kang & Choi, 2005), and activity and action recognition (Fu, 2015; Geurts, 2001; Pavlovic et al., 1999; Yu & Lee, 2015). In our analysis, the LSTM-NN model of swimming behaviour revealed that knowledge of the flow environment the animal experiences as it moves can allow one to make accurate out-of-sample forecasts of a fish's future movements, at least over short time-scales (e.g. 2 s–30 s). This suggests not only that features of the flow influence the movement decisions animals make as they migrate (Liao, 2007; Oteiza et al., 2017), but also that the behavioural rules or 'behavioural algorithms' (Hein et al., 2020) that relate flow to locomotion behaviour are at least reasonably similar, both across individual animals, and over the range of time periods included in

our study. We believe this workflow of building data-driven models of behaviour and validating predictions of those models on out-of-sample data is crucial, given that our understanding of how animals perceive and respond to sensory cues during migration is still far from complete. The flexibility of recurrent neural networks frees our approach, at least to some extent, from assumptions about the precise functional form relating flow variables to the swimming behaviour of migrants. However, one disadvantage of using a highly flexible framework like LSTM-NN to relate environmental variables to fish behaviour is that, due to the complexity of the neural network model structure, there is no compact symbolic representation of the functional relationships between input and output variables (Martin et al., 2018). We expect future studies will unpack the patterns described phenomenologically by our LSTM-NN model of movement behaviour. In particular, it will be insightful to determine whether migration behaviour, like some other animal behaviours including predator evasion (Hein et al., 2018) and prey interception (Brighton et al., 2017), can be described accurately by a set of relatively simple control algorithms (Hein et al., 2020). Future work could apply other modelling paradigms (e.g. control theory, neuro-ecological modelling; Bar et al., 2015; Brighton et al., 2017) to address this and other fundamental questions, including (a) which variables most influence locomotion, (b) whether migratory behaviour varies appreciably over time and (c) the extent to which different individuals respond to environmental variables in different ways. Notably, all of these questions require estimates of both the behavioural actions taken by individual migrants and the environmental variables experienced by those individuals. Our methodology provides a way to acquire such estimates.

While the overall methodology presented here holds much promise, it nevertheless has important limitations. Firstly, due to computational limitations on the simulation of turbulent flow, the spatial and temporal resolution of our CFD model is limited. This means that we cannot resolve fine-scale flow at the scale of the migrating fish's body, nor can we fully resolve temporal fluctuations in flow due to turbulence. This makes it challenging to directly link our model of locomotion behaviour with biomechanical (e.g. Bandyopadhyay, 2002; Cui et al., 2017; Lighthill, 1971) or behavioural models (e.g. Oteiza et al., 2017) that describe movement of the migrant's body. Nevertheless, our model does have the ability to resolve larger features in the flow on the spatial scale of tens of body lengths. Such features include gradients in water velocity near channel banks and zones of strong recirculation (e.g. see Figure 2). This allowed us to conclude, for example, that effects of these features on migratory behaviour can be significant (Figure 5). A second limitation of our approach is due to the tracking data themselves. Tracking data were acquired through hydrophone detections of animals implanted with acoustic transmitters. These data therefore have limited spatial resolution and the status of tagged animals are unknown (e.g. tags from fish consumed by larger predatory fish can still be detected by the hydrophone array). Such limitations are worth considering when choosing a tag technology to use for studies that will combine tracking and physical modelling to study migratory movement behaviour. Another important consideration is that our framework cannot fully address the question of whether high- or low-flow conditions are more favourable for migration because it does not consider how energy use trades off with other potentially important quantities related to migration success such as the travel time through regions of high predation risk (Anderson et al., 2005). The times taken by fish to traverse our entire study region were longer, on average, when overall flow was weak (mean of 63 min for trajectories experiencing the weakest 10% of flows) than when overall flow was strong (mean of 51 min for trajectories experiencing strongest 10% of flows). However, variability in this trend was significant. Nevertheless, travel time and other trade-offs could be included in our framework by integrating additional data sources (e.g. predation risk data).

Despite its limitations, our framework can be used to gain traction on questions that have fascinated migration biologists for many years. Many such questions relate to how migrants use energy as they move through a landscape. As demonstrated in Figure 5, our methodology has much potential to address these types of questions. For example, when applied to distinct environmental conditions observed in our dataset, locomotion force predictions revealed that fish generally spend far more energy moving through the landscape when the overall rate of flow is high than when the rate of flow is low, despite the fact that the net flow direction is aligned with the direction of migration. Our analysis provides additional insights into the cause of this pattern; when fish swim in slow currents, their movements are generally oriented uniformly relative to the flow with a slight bias towards alignment in the direction of flow (see Appendix A). On the other hand, when migrants move through high-speed currents, their movements are primarily oriented against the flow or laterally relative to the direction of flow. These lateral and opposing movements require greater power output. It is also important to note our methodology is in no way limited to the study of migratory movements. Both swimming and flying animals modulate short-term movement behaviour in response to local environmental flows (James, 2007; Scacco et al., 2019, Shepard et al., 2016). The same methodology presented here can be applied to study animal movement behaviours beyond the context of migration.

In this work, we have presented a general methodology for merging data and modelling of environmental currents with tracking data to understand animal migratory behaviour. Our approach extends more traditional methods in migration biology, which have often either ignored interactions with wind and water currents, or modelled these interactions in simple ways that are not fully informed by physical data (e.g. Alexander, 1998; Hein et al., 2012; Pennycuick, 2003; Stier et al., 2014). We believe our framework has the potential to shed new light on how migrants interact with wind and water currents and how behaviour and biophysics interact to determine the costs and benefits of different migratory strategies and environmental conditions.

ACKNOWLEDGEMENTS

The authors thank the California Department of Water Resources (Jacob McQuirk) for providing flow and tracking data, US Geological Survey California Water Science Center (Jon Burau) for technical information related to ADCP data and HTI Sonar (Sam Johnston) for technical information related to fish tracking data. They also thank M. Celis and A. Fahimipour for technical support with computing and for helpful suggestions on the manuscript, and participants in 'Data-theory seminar' for suggestions that improved this work. This project was supported, in part, through the California Department of Fish and Wildlife from the Water Quality, Supply and Infrastructure Improvement Act of 2014 (CA Department of Fish and Wildlife grant No. P1896007-01), the National Oceanic and Atmospheric Administration High Performance Computing and Communications Program and the National Science Foundation (IOS Grant 1855956).

AUTHORS' CONTRIBUTIONS

S.O., M.A.G., V.K.S. and A.M.H. conceived the ideas and designed the methodology; V.K.S. and A.M.H. pre-processed movement and hydrologic datasets; S.O. constructed CFD and LSTM models and carried out all analyses with input from all the authors; S.O. and A.M.H. led the writing of the manuscript. All the authors contributed critically to the drafts and gave final approval for publication.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.13604.

DATA AVAILABILITY STATEMENT

Data and Python scripts used for processing and post-processing are available at the Dryad Digital Repository (Olivetti, 2021).

ORCID

Simone Olivetti https://orcid.org/0000-0003-1460-2806
Vamsi K. Sridharan https://orcid.org/0000-0003-1457-6900
Andrew M. Hein https://orcid.org/0000-0001-7217-6185

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jozefowicz, R., Jia, Y., Kaiser, L., Monga, R., … Zheng, X. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems*. Retrieved from https://www.tensorflow.org/

Alerstam, T., Chapman, J. W., Bäckman, J., Smith, A. D., Karlsson, H., Nilsson, C., Reynolds, D. R., Klaassen, R. H. G., & Hill, J. K. (2011). Convergent patterns of long-distance nocturnal migration in noctuid moths and passerine birds. *Proceedings of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rspb.2011.0058

Alexander, R. M. (1998). When is migration worthwhile for animals that walk, swim or fly? *Avian Biology*, *29*, 387–394. https://doi.org/10.1242/jeb.015024

Altché, F., & De La Fortelle, A. (2017). An LSTM network for highway trajectory prediction. *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. https://doi.org/10.1109/ITSC.2017.8317913

Anderson, J. J., Gurarie, E., & Zabel, R. W. (2005). Mean free-path length theory of predator–prey interactions: Application to juvenile salmon migration. *Ecological Modelling*, *186*(2), 196–211. https://doi.org/10.1016/j.ecolmodel.2005.01.014

Arenas, A., Politano, M., Weber, L., & Timko, M. (2015). Analysis of movements and behavior of smolts swimming in hydropower reservoirs. *Ecological Modelling*, *312*, 292–307. https://doi.org/10.1016/j.ecolmodel.2015.05.015

Arya, S., Mount, D. M., Netanyahu, N. S., Silverman, R., Wu, A., & Wu, A. Y. (1998). An optimal algorithm for approximate nearest neighbor searching in fixed dimensions. *ACM*. https://doi.org/10.1145/293347.293348

Bandyopadhyay, P. R. (2002). Maneuvering hydrodynamics of fish and small underwater vehicles. *Integrative & Comparative Biology*, *42*(1), 102–117. https://doi.org/10.1093/icb/42.1.102

Bar, N. S., Skogestad, S., Marcal, J. M., Ulanovsky, N., & Yovel, Y. (2015). A sensory-motor control model of animal flight explains why bats fly differently in light versus dark. *PLoS Biology*, *13*. https://doi.org/10.1371/journal.pbio.1002046

Boning, C. W., Dispert, A., Visbeck, M., Rintoul, S. R., & Schwarzkopf, F. U. (2008). The response of the Antarctic Circumpolar Current to recent climate change. *Nature Geoscience*, *1*, 864–869. https://doi.org/10.1038/ngeo362

Brighton, C. H., Thomas, A. L. R., & Taylor, G. K. (2017). Terminal attack trajectories of peregrine falcons are described by the proportional navigation guidance law of missiles. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(51), 13495–13500. https://doi.org/10.1073/pnas.1714532114

Cui, Z., Gu, X., Li, K., & Jiang, H. (2017). CFD studies of the effects of waveform on swimming performance of carangiform fish. *Applied Science*, *7*(2), 149. https://doi.org/10.3390/app7020149

Dabiri, J. O. (1993). Migration by soaring or flapping flight in birds: The relative importance of energy cost and speed. *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.1993.0164

De Lucas, M., Janss, G. F., & Ferrer, M. (2004). The effects of a wind farm on birds in a migration point: The Strait of Gibraltar. *Biodiversity & Conservation*, *13*, 395–407. https://doi.org/10.1111/1365-2664.13107

Deshpande, S. S., Anumolu, L., & Trujillo, M. F. (2012). Evaluating the performance of the two-phase flow solver interFoam. *Computational Science & Discovery*. https://doi.org/10.1088/1749-4699/5/1/014016

Dingle, H. (2015). *Migration: The biology of life on the move*. Oxford Scholarship Online. ISBN 9780199640386.

Flack, A., Nagy, M., Fiedler, W., Couzin, I. D., & Wikelski, M. (2018). From local collective behavior to global migratory patterns in white storks. *Science*, *360*, 911–914. https://doi.org/10.1126/science.aap7781

Friedrich, K., Lundquist, J. K., Aitken, M., Kalina, E. A., & Marshall, R. F. (2012). Stability and turbulence in the atmospheric boundary layer: A comparison of remote sensing and tower observations. *Geophysical Research Letters*. https://doi.org/10.1029/2011GL050413

Fu, Y. (2015). *Human activity recognition and prediction*. Springer.

Gao, Z., Andersson, H. I., Dai, H., Jiang, F., & Zhao, L. (2015). A new Eulerian-Lagrangian agent method to model fish paths in a vertical slot fishway. *Ecological Engineering*. https://doi.org/10.1016/j.ecoleng.2015.12.038

Geurts, P. (2001). *Pattern extraction for time series classification* (pp. 115–127). Springer.

Gisen, D. C., Weichert, R. B., & Nestler, J. M. (2016). Optimizing attraction flow for upstream fish passage at a hydropower dam employing 3D Detached-Eddy Simulation. *Ecological Engineering*. https://doi.org/10.1016/j.ecoleng.2016.10.065

Goodwin, R. A., Politano, M., Garvin, J. W., Nestler, J. M., Hay, D., Anderson, J. J., Weber, L. J., Dimperio, E., Smith, D. L., & Timko, M. (2014). Fish navigation of large dams emerges from their modulation

of flow field experience. *Proceedings of the National Academy of Sciences of United States of America*, *111*(14), 5277–5282. https://doi.org/10.1073/pnas.1311874111

Gualtieri, C., Ianniruberto, M., Filizola, N., Santos, R., & Endreny, T. (2019). A 3D analysis of spatial habitat metrics about the confluence of Negro and Solimões rivers, Brazil. *Ecohydrology*, *13*(1). https://doi.org/10.1002/eco.2166

Hein, A. M., Altshuler, D. L., Cade, D. E., Liao, J. C., Martin, B. T., & Taylo, G. K. (2020). An algorithmic approach to natural behavior. *Current Biology*, *30*(11), R663–R675. https://doi.org/10.1016/j.cub.2020.04.018

Hein, A. M., Gil, M. A., Twomey, C. R., Couzin, I. D., & Levin, S. A. (2018). Conserved behavioral circuits govern high-speed decision-making in wild fish shoals. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(48), 12224–12228. https://doi.org/10.1073/pnas.1809140115

Hein, A. M., Hou, C., & Gillooly, J. F. (2012). Energetic and biomechanical constraints on animal migration distance. *Ecology Letters*, *15*(2), 104–110. https://doi.org/10.1111/j.1461-0248.2011.01714.x

Hoerner, S. F. (1965). *Fluid-dynamic drag: Theoretical, experimental and statistical information*. Published by the Author. ISBN 9780123742995.

Hughey, L., Hein, A. M., Strandburg-Peshkin, A., & Jensen, F. (2018). Challenges and solutions for studying collective animal behavior in the wild. *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.2017.0005

James, C. L. (2007). A review of fish swimming mechanics and behaviour in altered flows. *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.2007.2082

Jimenez, J., & Moser, R. D. (1998). *A selection of test cases for the validation of large-eddy simulations of turbulent flows*. North Atlantic Treaty Organization, Advisory Group for Aerospace Research Development.

Kang, H., & Choi, S. (2005). Framewise phoneme classification with bi-directional LSTM and other neural network architectures. *Neural Networks*, 602–610. https://doi.org/10.1016/j.neunet.2005.06.042

Kang, H., & Choi, S. (2014). Bayesian common spatial patterns for multi-subject EEG classification. *Neural Networks*, 39–50. https://doi.org/10.1016/j.neunet.2014.05.012

Kellnerová, R., Fuka, V., Uruba, V., Jučáková, K., Nosek, Š., Chaloupecká, H., & Jaňour, Z. (2018). On street-canyon flow dynamics: Advanced validation of LES by time-resolved PIV. *Atmosphere*, https://doi.org/10.3390/atmos9050161

Killen, S. S., Glazier, D. S., Rezende, E. L., Clark, T. D., Atkinson, D., Willener, A. S. T., & Halsey, L. G. (2016). Ecological influences and morphological correlates of resting and maximal metabolic rates across teleost fish species. *The American Naturalist*, *187*(5), 592–606. https://doi.org/10.1086/685893

Kimmerer, W., Avent, S. R., Bollens, S. M., Feyrer, F., Grimaldo, L. F., Moyle, P. B., Nobriga, M., & Visintainer, T. (2005). Variability in length–weight relationships used to estimate biomass of Estuarine fish from survey data. *Transactions of the American Fisheries Society*, https://doi.org/10.1577/T04-042.1

Kling, M. M., & Ackerly, D. D. (2020). Global wind patterns and the vulnerability of wind-dispersed species to climate change. *Nature Climate Change*, *10*, 868–875. https://doi.org/10.1038/s41558-020-0848-3

Lacey, R. W. J., Neary, V. S., Liao, J. C., Enders, E. C., & Tritico, H. M. (2012). The IPOS framework: Linking fish swimming performance in altered flows from laboratory experiments to rivers. *River Research and Applications*. 429–443. https://doi.org/10.1002/rra.1584

Liao, J. C. (2007). A review of fish swimming mechanics and behaviour in altered flows. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1484), 1973–1993. https://doi.org/10.1098/rstb.2007.2082

Lighthill, M. J. (1971). Large-amplitude elongated-body theory of fish locomotion. *Proceedings of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rspb.1971.0085

Ling, H., McIvor, G. E., Nagy, G., MohaimenianPour, S., Vaughan, R. T., Thornton, A., & Ouellette, N. T. (2018). Simultaneous measurements of three-dimensional trajectories and wingbeat frequencies of birds in the field. *Journal of The Royal Society Interface*. https://doi.org/10.1098/rsif.2018.0653

Madala, S., Satyanarayana, A. N. V., Srinivas, C. V., Kumar, M. E. A., & Marshall, R. F. (2015). Stability and turbulence in the atmospheric boundary layer: A comparison of remote sensing and tower observations. *Geophysical Research Letters*. https://doi.org/10.1016/j.atmosenv.2015.02.059

Martin, B. T., Munch, S. B., & Hein, A. M. (2018). Reverse-engineering ecological theory from data. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1878). https://doi.org/10.1098/rspb.2018.0422

Martin, B. T., Nisbet, R. M., Pike, A., Michel, C. J., & Danner, E. M. (2015). Sport science for salmon and other species: Ecological consequences of metabolic power constraints. *Ecology Letters*, *18*, 535–544. https://doi.org/10.1111/ele.12433

Martin, J. E., Politano, M. S., Prakash, S., Carrica, P. M., & Markfort, C. D. (2017). Optimizing bat carcass search areas using a CFD-Lagrangian Modeling Approach. Technical report to Mid American Energy Company 414.

McQuirk, J., Simon, K., & Cane, M. (2015). AECOM. In *An evaluation of juvenile salmonid routing and barrier effectiveness, predation, and predatory fishes at the Head of Old River*. ICF International, and Turnpenny Horsefield Associates.

Menter, F. R. (1994). Two-equation eddy-viscosity turbulence models for engineering applications. *AIAA Journal*, *32*(8), 1598–1605. https://doi.org/10.2514/3.12149

North, E. W., Schlag, Z., Hood, R. R., Li, M., Zhong, L., Gross, T., & Kennedy, V. S. (2008). Vertical swimming behavior influences the dispersal of simulated oyster larvae in a coupled particle-tracking and hydrodynamic model of Chesapeake Bay. *Marine Ecology Progress*, *359*, 99–115. https://doi.org/10.3354/meps07317

Olivetti, S. (2021). Python scripts and database. *Dryad Digital Repository*, https://datadryad.org/stash/dataset/doi:10.5061/dryad.2547d7wq4

Oteiza, P., Odstrcil, I., Lauder, G., Portugues, R., & Engert, F. (2017). A novel mechanism for mechanosensory-based rheotaxis in larval zebrafish. *Nature*, *547*(7664), 445–448. https://doi.org/10.1038/nature23014

Pavlovic, V., Frey, B. J., & Huang, T. S. (1999). Time-series classification using mixed-state dynamic Bayesian networks. *IEEE Computer Society Conference*, 609–615. https://doi.org/10.1016/j.neunet.2005.06.042

Pennycuick, C. J. (2003). The concept of energy height in animal locomotion: Separating mechanics from physiology. *Theoretical Biology*, *224*, 189–203. https://doi.org/10.1016/S0022-5193(03)00157-7

Pennycuick, C. J. (2008). *Modelling the flying bird*. Theoretical Ecology. Academic Press. ISBN 9780123742995.

Ramón, C. L., Acosta, M., & Rueda, F. J. (2018). Hydrodynamic drivers of Juvenile-Salmon out-migration in the Sacramento river: Secondary circulation. *Journal of Hydraulic Engineering Division of the American Society of Civil Engineers*. https://doi.org/10.1061/%28ASCE%29HY.1943-7900.0001484

Reddy, G., Celani, A., Sejnowski, T. J., & Vergassola, M. (2016). Learning to soar in turbulent environments. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(33), E4877–E4884. https://doi.org/10.1073/pnas.1606075113

Scacco, M., Flack, A., Duriez, O., Wikelski, M., & Safi, K. (2019). Static landscape features predict uplift locations for soaring birds across Europe. *Royal Society Open Science*. https://doi.org/10.1098/rsos.181440

Shepard, E. L. C., Ross, A. N., & Portugal, S. J. (2016). Moving in a moving medium: New perspectives on flight. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1704). https://doi.org/10.1098/rstb.2015.0382

Silva, A. T., Lucas, M. C., Castro-Santos, T., Katopodis, C., Baumgartner, L. J., Thiem, J. D., Aarestrup, K., Pompeu, P. S., O'Brien, G. C., Braun, D. C., Burnett, N. J., Zhu, D. Z., Fjeldstad, H.-P., Forseth, T., Rajaratnam, N., Williams, J. G., & Cooke, S. J. (2018). The future of fish passage science, engineering, and practice. *Fish and Fisheries*, *19*, 340–362. https://doi.org/10.1111/faf.12258

Smith, R. J. F. (2012). *The control of fish migration*. Zoophysiology. Springer. ISBN 9783642823480.

Sridharan, V. K., Monismith, S. G., Fringer, O. B., & Fong, D. A. (2018). Evaluation of the delta simulation model-2 in computing tidally driven flows in the Sacramento-San Joaquin Delta. *San Francisco Estuary and watershed Science*, *16*(2). https://doi.org/10.15447/sfews.2018v16iss2art6

Stier, A. C., Hein, A. M., Parravicini, V., & Kulbicki, M. (2014). Larval dispersal drives trophic structure across Pacific coral reefs. *Nature Communications*. https://doi.org/10.1038/ncomms6575

Stumpner, P. (2013a). *Memo on HORB – 2009 and 2011 hydrodynamic data processing and interpolation*. Task No. TEMBAR-07. AECOM Technical Services Inc.

Stumpner, P. (2013b). *Memo on head of old river deployment 2012*. AECOM Technical Services Inc.

Thorstad, E. B., Okland, F., Aarestrup, K., & Heggberget, T. G. (2008). Factors affecting the within-river spawning migration of Atlantic salmon, with emphasis on human impacts. *Reviews in Fish Biology and Fisheries*, *18*, 345–371. https://doi.org/10.1007/s11160-007-9076-4

Wagenbrenner, N. S., Forthofer, J. M., Lamb, B. K., Shannon, K. S., & Butler, B. W. (2016). Downscaling surface wind predictions from numerical weather prediction models in complex terrain with WindNinja. *Atmospheric Chemistry and Physics*, *16*(8), 5229–5241. https://doi.org/10.5194/acp-16-5229-2016

Wang, R. F., Ateljevich, E., Fregoso, T. A., & Jaffe, B. E. (2018). A revised continuous surface elevation model for modeling. *Methodology for Flow and Salinity Estimates in the Sacramento-San Joaquin Delta and Suisun Marsh: 39th Annual Progress Report*. Bay-Delta Office, California Department of Water Resources and Pacific Coastal and Marine Science Center, United States Geological Survey. pp. 125–170.

Weber, L. J., Goodwin, R. A., Li, S., Nestler, J. M., & Anderson, J. J. (2006). Application of an Eulerian-Lagrangian-Agent method (ELAM) to rank alternative designs of a juvenile fish passage facility. *Hydroinformatics*, *8*, 271–295. https://doi.org/10.2166/hydro.2006.006

Weller, H. G., Tabor, G., Jasak, H., & Fureby, C. (1998). A tensorial approach to computational continuum mechanics using object-oriented techniques. *Computers in Physics*, *12*(6), 620–https://doi.org/10.1063/1.168744

Williams, J. G. (2006). Central Valley salmon: A perspective on Chinook and steelhead in the Central Valley of California. *San Francisco Estuary and Watershed Science*, *4*, 3. https://doi.org/10.15447/sfews.2006v4iss3art2

Winstral, A., Jonas, T., & Helbig, N. (2017). Statistical downscaling of gridded wind speed data using local topography. *Journal of Hydrometeorology*, *18*(2), 335–348. https://doi.org/10.1175/JHM-D-16-0054.1

Xue, H., Huynh, D. Q., & Reynolds, M. (2018). An LSTM network for highway trajectory prediction. *SS-LSTM: A hierarchical LSTM model for pedestrian trajectory prediction*. https://doi.org/10.1109/WACV.2018.00135

Yu, Z., & Lee, M. (2015). Real-time human action classification using a dynamic neural model. *Neural Networks*, *69*, 29–43. https://doi.org/10.1016/j.neunet.2015.04.013

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Olivetti S, Gil MA, Sridharan VK, Hein AM. Merging computational fluid dynamics and machine learning to reveal animal migration strategies. *Methods Ecol Evol*. 2021;00:1–15. https://doi.org/10.1111/2041-210X.13604