

## Can AI predict animal movements? Filling gaps in animal trajectories using inverse reinforcement learning

TSUBASA HIRAKAWA,<sup>1,†</sup> TAKAYOSHI YAMASHITA,<sup>1</sup> TORU TAMAKI,<sup>2</sup> HIRONOBU FUJIYOSHI,<sup>1</sup> YUTA UMEZU,<sup>3</sup> ICHIRO TAKEUCHI,<sup>3,4,5</sup> SAKIKO MATSUMOTO,<sup>6</sup> AND KEN YODA<sup>6</sup>

<sup>1</sup>Department of Computer Science, Chubu University, 1200 Matsumoto, Kasugai, Aichi 487-0027 Japan

<sup>2</sup>Graduate School of Engineering, Hiroshima University, 1-4-1 Kagamiyama, Higashi Hiroshima, Hiroshima 739-8527 Japan

<sup>3</sup>Department of Computer Science, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466-8555 Japan

<sup>4</sup>RIKEN Center for Advanced Intelligence Project, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027 Japan

<sup>5</sup>Center for Materials Research by Information Integration, National Institute for Materials Science, 1-2-1 Sengen, Tsukuba 305-0047 Japan

<sup>6</sup>Graduate School of Environmental Studies, Nagoya University, Furo, Chikusa, Nagoya 464-8601 Japan

**Citation:** Hirakawa, T., T. Yamashita, T. Tamaki, H. Fujiyoshi, Y. Umezu, I. Takeuchi, S. Matsumoto, and K. Yoda. 2018. Can AI predict animal movements? Filling gaps in animal trajectories using inverse reinforcement learning. *Ecosphere* 9(10):e02447. 10.1002/ecs2.2447

**Abstract.** Focal animal sampling and continuous recording of behavior *in situ* are essential in the study of ecology. However, observation gaps and missing records are unavoidable because the focal individual can move out of sight and recording devices do not always work properly. Using an inverse reinforcement learning (IRL) framework, we have developed a novel gap-filling method to predict the most likely route that an animal would have traveled; within this framework, an algorithm learns a reward function from animal trajectories to find the environmental features preferred by the animal. We applied this approach to GPS trajectories obtained from streaked shearwaters (*Calonectris leucomelas*) and provide evidence of the advantages of the IRL approach over previously used interpolation methods. These advantages are as follows: (1) No assumptions about the parametric distribution governing movements are needed, (2) no assumptions regarding landscape preferences and restrictions are needed, and (3) large spatiotemporal gaps can be filled. This work demonstrates how IRL can enhance the ability to fill gaps in animal trajectories and construct reward-space maps in heterogeneous environments. The proposed methodology can assist movement research, which seeks to understand phenomena that are ecologically and evolutionarily significant, such as habitat selection and migration.

**Key words:** animal movement; behavioral monitoring; bio-logging; biotelemetry; *Calonectris leucomelas*; habitat selection; interpolation; inverse reinforcement learning; machine learning; reward map; tracking data.

Received 17 May 2018; revised 10 August 2018; accepted 23 August 2018. Corresponding Editor: Fie Fang.

**Copyright:** © 2018 The Authors. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

† E-mail: hirakawa@mprg.cs.chubu.ac.jp

### INTRODUCTION

Focal animal sampling (i.e., observation of a target individual) and the continuous recording of behavior *in situ* are essential in the study of ecology and evolution (Martin and Bateson 2017). Focal sampling can reveal the true properties of

behavior (e.g., behavior duration) and spatial location in the environment (Turchin 1998). Animal tracking technology, which is now being used in focal sampling, enables us to track the movement of animals at various spatial and temporal scales (Hussey et al. 2015, Kays et al. 2015, Wilmers et al. 2015). The derived animal-borne data can be

used to understand ecologically significant phenomenon, such as habitat selection, migration, dispersion, and foraging strategies, across multiple spatiotemporal scales (Nathan et al. 2008).

Animal locations are coarsely sampled or can be missed entirely in large sections by commonly used devices such as conventional very high frequency (VHF) triangulation techniques, Argos satellite tracking, and light-level geolocation sensors (Ryan et al. 2004). Even the advanced Global Positioning System (GPS), which can collect accurate positions at frequent sampling intervals, is often scheduled to coarsely collect locations at even intervals (Tomkiewicz et al. 2010). This is because GPS sensors consume a large amount of power, and the size of an animal-borne device is limited by the animal's carrying capacity. In addition, the operational sampling rate can vary because of satellite signal communication failures (Graves and Waller 2006, Frair et al. 2010, Focardi and Cecere 2014), which leads to missed data over periods of minutes or hours. The signal communication failures are not improved by higher sampling rates, especially for species traveling underwater (Pelletier et al. 2014), moving under forest canopies or within canyons, or staying inside a burrow (Cain et al. 2005), because GPS requires a clear line of sight to satellites. Such observation gaps and uneven sampling introduce non-negligible spatial and temporal biases in behavioral analysis, such as when determining home ranges and behavioral time budgets. For example, the most commonly used path-interpolation method is straight-line interpolation between each pair of observed locations (e.g., Lonergan et al. 2009), which underestimates the length and tortuosity of actual tracks (Freitas et al. 2008, Rowcliffe et al. 2012, Battaile et al. 2015) and the anti-predator responses of prey species (Creel et al. 2013).

These gaps are filled by implementing multiple observation methods that complement each other (Dewhurst et al. 2016, Godwin et al. 2016) and/or spatial interpolation methods (Tremblay et al. 2006, Russo et al. 2011, Winship et al. 2012, Technikis et al. 2015, Hooten et al. 2017). Such methods are undoubtedly useful for increasing confidence in movement data to understand behavior and habitat use in animals; however, these methods often have three general limitations. First, some methods assume specific movement properties and/or constraints regarding animal behavior

between the observation records. For example, an animal movement trajectory is often modeled as a random process (Codling et al. 2008), such as correlated random walks (RWs; Benhamou 2006, Rowcliffe et al. 2012), mixtures of RWs (Morales et al. 2004, Russo et al. 2011), Brownian bridges (Horne et al. 2007), and Kriging (Fleming et al. 2016). Other frequently assumed animal movement behaviors are that animals moving in fluid (i.e., water and air) follow curvilinear rather than linear paths due to flows, vortices, and turbulence (Tremblay et al. 2006). These mechanistic models should include realistic movement properties and certain parametric assumptions (Royer and Luttavague 2008), which depend on species and environmental conditions (Humphries et al. 2010, de Jager et al. 2011, Browning et al. 2018). In addition, the models should be considered hypotheses under examination, but intense controversy still exists over which model can be fit to the movement patterns of animals (e.g., Lévy walk vs. composite correlated RW; Edwards et al. 2007, Auger-Méthé et al. 2015).

Second, animals prefer or avoid specific environmental factors, which are not well captured by previous methods. These factors include obstacles, physical boundaries, commuting corridors, neighbors, predator habitats, territories, stopover sites, and energy-consuming paths (e.g., Wall et al. 2006, Shepard et al. 2013, Northrup et al. 2016, Yoda et al. 2017); thus, movement costs for animals are highly likely to be variable in space. Many previous gap-filling methods easily cross these boundaries. For example, straight-line interpolation (the most likely path based on RWs) neglects boundaries at which animal movements are physically restricted. Previous studies addressed this important problem by manually removing locations (Freitas et al. 2008) or by incorporating environmental information in the modeling process (Tremblay et al. 2009a, Benhamou and Cornélis 2010, Pedersen et al. 2011, Buchin et al. 2015, McClintock et al. 2017). These methods must explicitly formulate boundaries and interactions based on expert experience and prior knowledge (e.g., Brost et al. 2015), although no simple description of animal movement landscape constraints may be possible (Wilson et al. 2012, Gallagher et al. 2017). This could result in subjective models that fail to capture interactions between animals and the

environment, because these interactions are often unclear. For example, it is not obvious whether the movements of forest birds are constrained by open areas (Bélisle et al. 2001).

Third, previous interpolation methods are not sufficient to deal with large spatial gaps. Any interpolation method might be acceptable over a short time-scale and small space-scale gap. However, the longer the gap duration, the greater the discrepancy between the linear interpolation and real animal track (Rowcliffe et al. 2012), and the midpoint between the observed animal locations has the greatest uncertainty (Torres et al. 2011). Likewise, as the time and space intervals between pairs of locations increase, the assumption of an RW is likely to be violated (Horne et al. 2007) and the accuracy of interpolation declines (Bailey et al. 2008). Dead-reckoning technologies (i.e., reconstructed trajectories using speed and heading) might be used to fill gaps between known locations (Wilson et al. 2007, Bidder et al. 2015), but estimation errors accumulate as the distance traveled increases. In addition, some interpolation methods are susceptible to look-ahead bias in forecasting, in which the models use information just prior to a gap (e.g., movement angle) to estimate the next position. This leads to earlier movements being repeated as the gap period increases (Bailey et al. 2008).

Here, we introduce a new method to fill the gaps between observed locations based on inverse reinforcement learning (IRL; Russell 1998, Ng and Russell 2000), which is a machine learning approach. In the field of artificial intelligence, human trajectory prediction (path prediction) has many applications (Murino et al. 2017). A recent network analysis has shown that trajectories have a higher predictability than expected; the location of a person can be predicted with 93% accuracy (Song et al. 2010), indicating that machine learning, which infers knowledge from trajectory data, is a promising approach to the prediction of animal movements. The use of machine learning has become more common in ecological, behavioral, and environmental research over the last decade (Guilford et al. 2009, Resheff et al. 2014, Appelhans et al. 2015, Valletta et al. 2017, Weinstein 2017, Browning et al. 2018), but reinforcement learning (RL; Fig. 3; Kaelbling et al. 1996, Sutton and Barto 1998), which is a major machine learning paradigm along with supervised and

unsupervised learning, has not been applied to animal behavioral data in the field to date. RL has been widely used in the path planning of robots, allowing them to navigate from a starting point to a destination in a maze (Singh et al. 1994, Smart and Kaelbling 2002, Jaradat et al. 2011, Chen et al. 2017), or for the motion planning of robot arms, allowing an arm to move from one point to another (Stulp 2012, Levine et al. 2016). Given a reward function, an RL algorithm aims to generate a sequence of actions (or movements) for an agent to maximize the expected total reward. The action of the next time step is generated based on the internal state of the current time step. This process is called the Markov decision process (MDP). In this process, actions depend only on the states of the preceding time step, and the mapping from the state to the action is called the policy (Kaelbling et al. 1996, Sutton and Barto 1998). In this study, agents and action sequences correspond to animals and their trajectories, respectively, and an RL algorithm provides a long-term movement prediction (i.e., filling in gaps caused by unobserved movement). It is important to design an appropriate reward function for successful prediction; however, defining reward functions manually and a priori is difficult or impossible in many cases (Russell 1998, Ng and Russell 2000). IRL solves the inverse RL problem by learning the reward function from observed behaviors (trajectories of animals), and information about the environment in which the agent moves.

We provide an example of streaked shearwater (*Calonectris leucomelas*) trajectories derived from animal-borne GPS loggers. Our IRL method learned a reward function from their trajectories (approximately 80% of the entire dataset) to find the trajectories the shearwaters prefer to follow (Fig. 1). We then artificially introduced missing data into the middle of the trajectories in the remaining 20% of data and completed the gaps based on the learned reward function. Finally, we compared our prediction result with the traditional method (i.e., linear interpolation).

## METHODS

### Study site and species

This study was conducted between 2009 and 2015 at Awashima Island (38°28' N, 139°14' E),

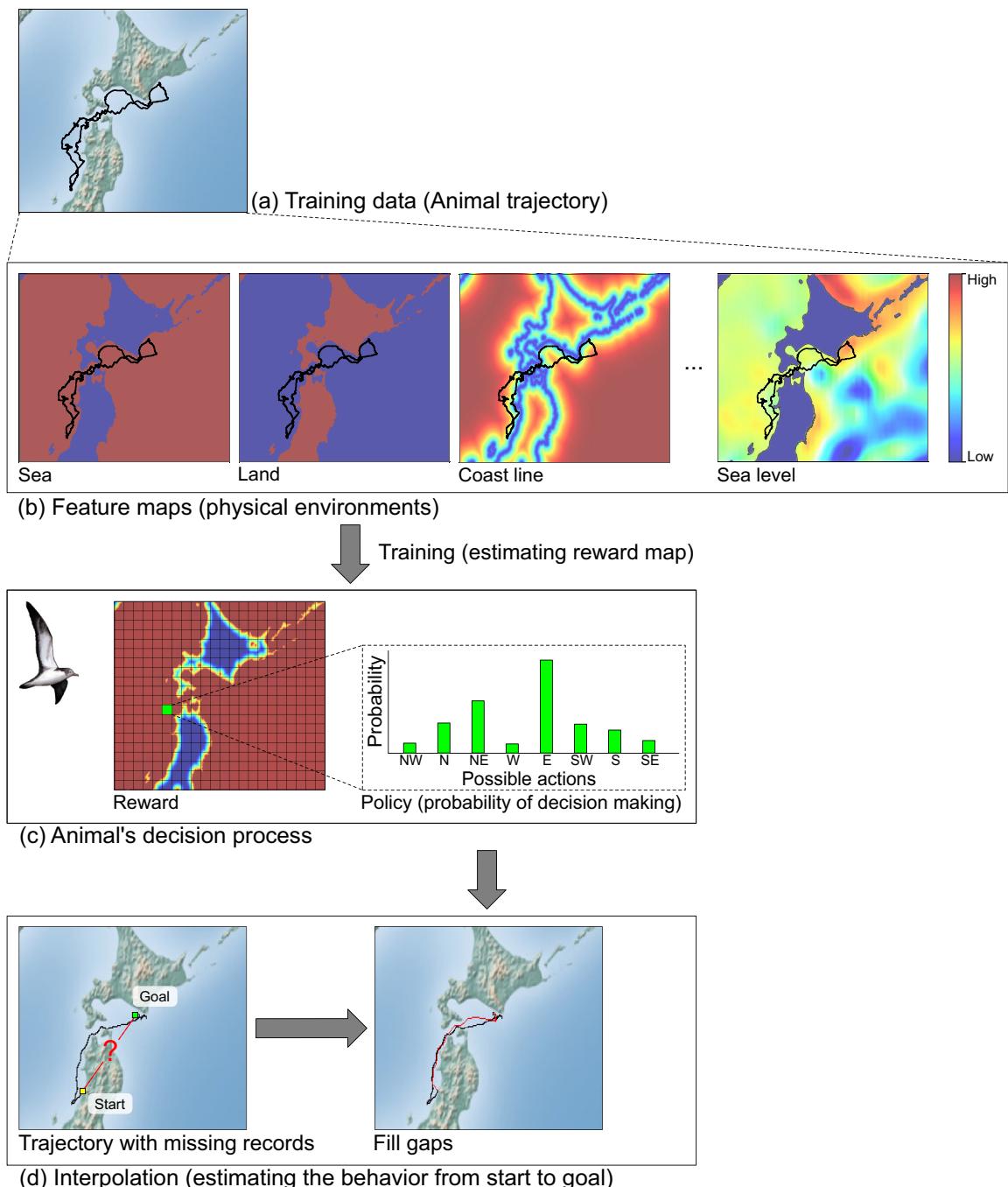


Fig. 1. Overview of the inverse reinforcement learning (IRL)-based trajectory interpolation method. Our IRL-based interpolation method estimates a reward that reproduces observed behaviors, after which the missing parts in a trajectory are interpolated by estimating the behaviors between both ends of the missing parts based on the estimated reward. (a) Given animal movement data are collected (training data) and (b) feature maps that might affect the movements are prepared, (c) a reward map and policy that generate an action at each location are estimated in the training phase. (d) In the test phase, the beginning and end of a missing trajectory are set as start and goal positions, respectively, and the behavior from start to goal is estimated (i.e., the trajectory is interpolated).

Japan. We attached GPS loggers (GiPSy or Axy-Trek, developed by TechnoSmArt, Roma, Italy) to the back feathers of 317 chick-rearing streaked shearwaters using Tesa tape. Adult shearwaters conduct round-trip foraging expeditions in areas up to 1000 km from the colony during the chick-rearing season (Matsumoto et al. 2017). The GPS loggers were programmed to record positional data every minute as trajectories in the format of latitude, longitude, and timestamp. However, the loggers did not always record a location every one min because of a variety of reasons, including the diving activities of the shearwaters and GPS antenna orientation during low flights close to waves. The overall logger mass was approximately 25 g, which corresponded to <5% of the body mass of each bird. Foraging trip duration and chick survival rate are not affected by the logger (Yoda et al. 2014). The procedures used in this study adhered to the guidelines of the Animal Experimental Committee of Nagoya University. All protocols were approved by the Ministry of the Environment, Japan.

#### Data treatment

Details of the data treatment before the following analyses are given in Appendix S1. We used 106 trajectories (53 males and 53 females; Fig. 2a, b) that do not contain large missing parts because they were used for training and evaluating the interpolation performance of the methods. The training trajectories are long enough for analysis, as the histograms of duration of training trajectories in Fig. 2a, b and statistics in Table 1 show. The average of durations is about 5000 min, or three and a half days.

After the data selection, we converted latitude ( $37^{\circ}30' \text{ N}$ – $45^{\circ}30' \text{ N}$ ) and longitude ( $136^{\circ}30' \text{ E}$ – $148^{\circ}30' \text{ E}$ ) into 200- and 300-cell grids with equal intervals, respectively. Each resultant grid cell corresponded to a square approximately 3 km wide. If we had used the continuous geographical coordinates directly or finer grid sizes, the size of the state space would have been extremely large, leading to massive computational costs. This is a common issue shared by other IRL methods (Ziebart et al. 2008, 2009, Kitani et al. 2012).

We also converted the timestamp into time elapsed since the start of the trajectory. Using the grid, the data points of a trajectory are in the

form of  $x$  and  $y$  grid coordinates with elapsed time. If successive data points have the same grid coordinates, then the first one is kept by removing others to ensure that a trajectory consists of a sequence of different cell positions. The elapsed time is then replaced with the number of movements in the grid from the beginning. This is a discrete time step corresponding to approximately fifteen minutes on average. Although we selected trajectories that had missing records of short durations (see Appendix S1), the missing records lead to consecutive points that are non-adjacent with each other in the converted grid coordinates. The corresponding missing parts were small (a few cells out of  $200 \times 300$  grid cells); therefore, we linearly interpolated the missing parts between these cells. As a result, the  $t$ -th data point of a trajectory consists of  $x_t$  and  $y_t$  grid coordinates, and a discrete time step  $z_t$ .

#### Modeling approach

**Reinforcement learning.**—RL is a type of machine learning framework that obtains optimal behavior to achieve a certain goal. The RL framework consists of an agent and an environment and contains the following elements (Fig. 3): State  $s$  represents the environment in which an agent is located, action  $a$  represents the behaviors that an agent performs at a certain state, immediate reward  $r$  represents the value that can be obtained at each state, and policy  $\pi$  represents a strategy for deciding an action. In the RL framework, an agent interacts with an environment by taking an action according to a policy. The action affects the state of the agent leading to a state transition with an immediate reward. In the case of a maze navigation task, where a robot moves toward the goal in a maze (Dayan and Hinton 1992, Kaelbling et al. 1996), the agent is the robot and the environment is the maze in which the robot behaves. The environment contains information affecting action decision of the robot, such as wall or passageway, in addition to locations (coordinates). When the robot moves around by following actions, the state (location) of the robot changes. The next action is decided based on the location and the information of the state. Similar to the maze navigation task, in our trajectory interpolation problem, the agent is a shearwater, and the environment is represented by a static geographical map that includes factors affecting action decision

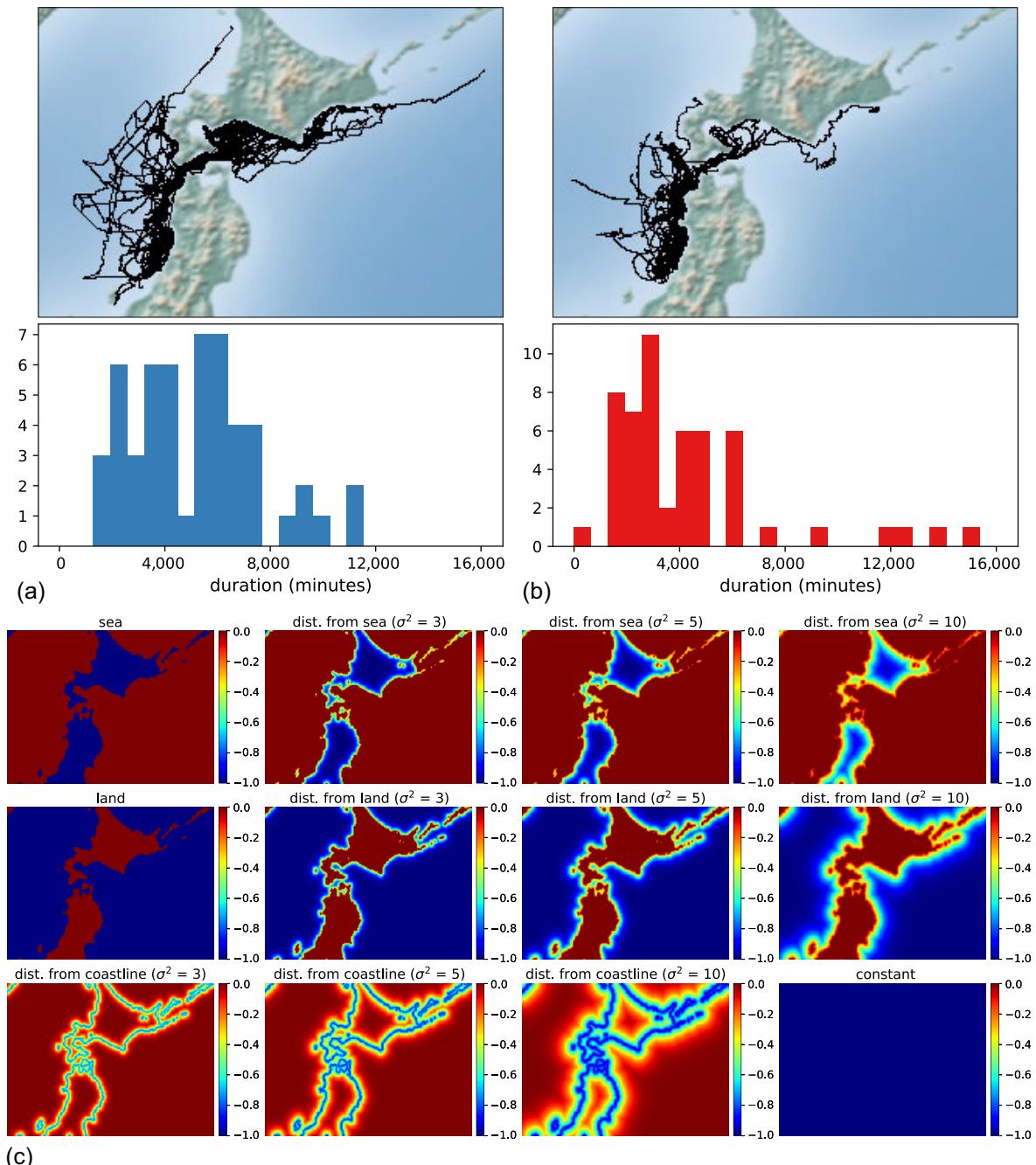


Fig. 2. Shearwater trajectories and feature maps. (a) Male shearwater trajectories and the histogram of trajectory duration ( $n = 53$ ). (b) Female shearwater trajectories and the histogram of trajectory duration ( $n = 53$ ). (c) Feature maps used in our experiments. Each feature map represents the degree to which a location belongs to the physical environment. For example, the sea map shows whether the location belongs to sea or not; the higher valued (red-colored) region indicates that the area is sea, whereas the lower valued (blue-colored) region indicates that the area is not sea. The distance from the sea map indicates how far the location is from the sea area (the details of the mathematical definition are described in Appendix S2). The constant map is also prepared, in which an agent follows the shortest path, because negative rewards are constantly accumulated irrespective of the routes.

of the shearwater, for example, whether the location is over sea or land.

The static geographical map is a continuous two-dimensional space. As mentioned before, the space is discretized into  $200 \times 300$  grid cells (Fig. 3, right). These cells are possible locations at which a shearwater can exist. In studies on two-dimensional path prediction (Ziebart et al. 2009, Kitani et al. 2012), the state of an agent is defined as a vector  $s_t = (x_t, y_t)$ , where  $x_t$  and  $y_t$  are two-dimensional coordinates. In contrast, we define the state as  $s_t = (x_t, y_t, z_t)$ , where  $x_t$  and  $y_t$  are discrete coordinates of the two-dimensional grid and  $z_t$  is the discrete time step.

The reason we add the discretized elapsed time  $z_t$  in state  $s_t$  is that shearwaters may take indirect routes. Path prediction methods for pedestrians (Ziebart et al. 2009, Kitani et al. 2012) aim to predict rather smooth and straight trajectories toward a goal that are as short as possible. Basically, path prediction methods within the RL framework define negative rewards (the

details are described in the inverse RL section), resulting in the predictions of relatively shorter trajectories with higher rewards. This is similar to the shortest path problem. These methods provide practical prediction results because pedestrians tend to take such direct trajectories. However, this is not suitable for our task because the round-trip foraging expeditions of shearwaters from the colony are not direct or smooth. Such indirect trajectories usually are longer than a direct flight to the goal. If we interpolate such indirect trajectories using the above path prediction methods, it would be difficult to interpolate them because a longer trajectory incurs a large negative reward. By introducing the discretized elapsed time as a part of the state, we can implicitly consider the length of trajectories; longer elapsed time indicates longer (i.e., indirect) trajectories. With this state space definition, the agent needs to repeat action selection  $z_t$  times to reach the goal regardless of the reward value, and indirect routes can be considered.

Actions are possible moving directions of the shearwater from the current state. The transition between grid cells is described by  $a_{xy} \in \{\text{north, south, east, west, northeast, northwest, southeast, southwest}\}$ . These transitions are for spatial states  $x_t$  and  $y_t$  only. For the state transition of  $z_t$ , we increment it by one irrespective of the spatial actions (Fig. 4a). Therefore, an action of our

Table 1. Statistics of the duration of shearwater trajectories (in min).

Sex	Minimum	Maximum	Mean $\pm$ Std.
Males	1469.9	11,171.1	5138.1 $\pm$ 2370.5
Females	524.9	15,274.2	4279.3 $\pm$ 3108.3

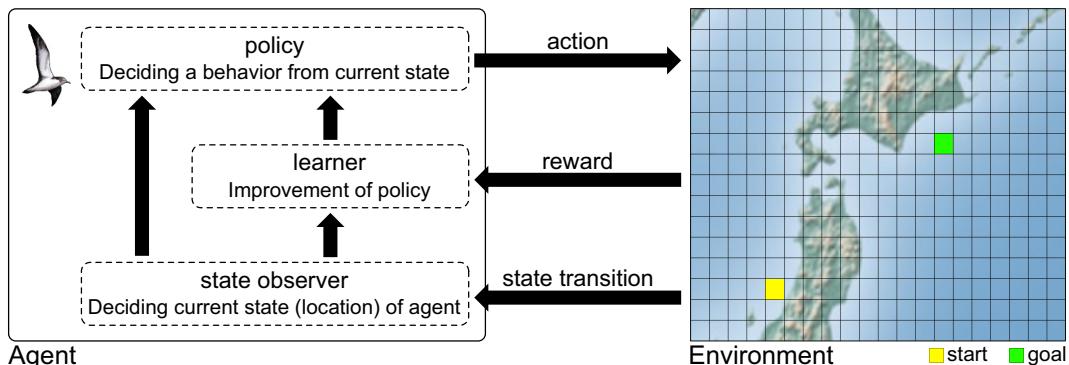


Fig. 3. Reinforcement learning (RL) framework in a shearwater trajectory interpolation problem. In our trajectory interpolation problem, the agent and environment correspond to a shearwater and a two-dimensional static geographical map where the shearwaters behave, respectively. Initially, the agent is located at the start and then moves toward the goal. When the agent takes an action based on the policy, the state (current location of the agent) is transferred and the agent obtains a reward value. The agent aims to find a policy to obtain a higher expected total reward. In the RL framework, the policy is repeatedly updated and gradually improved.

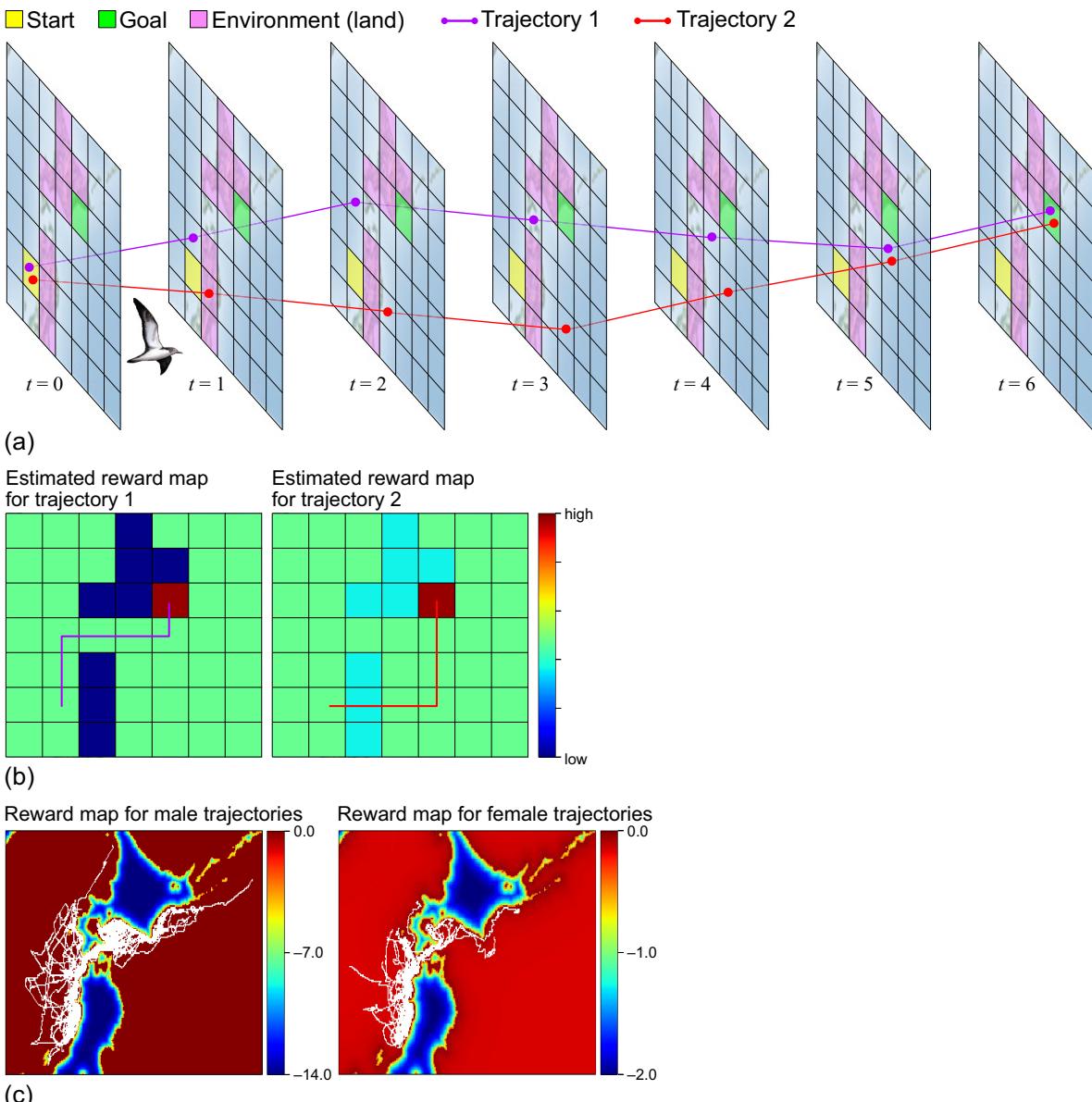


Fig. 4. Concept of inverse reinforcement learning (IRL). (a) Behavior of agents in our IRL-based interpolation method. Because our proposed method considers the state in the form of  $s_t = (x_t, y_t, z_t)$ , the movement of the agent is defined not only by spatial transitions between cells in the grid, but also by temporal transitions as increments of discretized elapsed time. (b) Estimated reward values are dependent on the environment and observed behaviors. For example, if we estimate the reward map with trajectory 1, in which the agent does not cross the pink-colored cells (as in (a)), this type of cell is assumed to influence the agent behavior negatively and has a lower reward value. In contrast, using trajectory 2, in which the agent moves across the cells, yields the inference that these cells have less influence on the agent's behavior and a higher reward value. Note that the reward of the goal state increases in general. (c) Estimated reward values (reward maps) for males and females. Red-colored regions have higher rewards, and blue-colored regions have lower rewards.

method is defined as  $a = (a_{xy}, 1)$ . In the following, a trajectory is denoted by  $\zeta = \{(s_0, a_0), (s_1, a_1), \dots\}$ , which is a sequence of state-action pairs  $(s_t, a_t)$ , and a set of trajectories by  $Z = \{\zeta_1, \dots, \zeta_n\}$ .

To interpolate a trajectory (Fig. 1d) based on the RL framework, we specify an initial state (start) where a missing record begins, and a terminal state (goal) where the missing record ends. An agent repeatedly takes actions according to a policy, during which the agent obtains and accumulates immediate reward values  $r$ . This accumulated reward  $R$  indicates the quality of the behaviors (the set of the state-action pairs). RL estimates the optimal policy by which the agent obtains the highest accumulated reward; clearly, the optimal policy is dependent on the definition of the appropriate reward values. However, immediate or accumulated reward values are not always given explicitly, or they can be difficult to define, especially for ecology experiments. Therefore, we use IRL, which is the problem of estimating reward values from behavior data (Fig. 4).

*Inverse reinforcement learning.*—Of the IRL methods, we take a maximum entropy (MaxEnt) approach (Ziebart et al. 2008). MaxEnt IRL is modeled under the assumption that the observed actions of the agent are samples from a near-optimal stochastic policy. Based on this assumption, it learns the reward function that matches the observed actions. MaxEnt IRL is a probabilistic approach, and the Bellman equations of value function  $V$  and action-value function  $Q$  are extended probabilistically as follows:

$$\begin{aligned} Q(s, a) &= r(s; \theta) + E_{p(s'|s, a)}[V(s')] \\ V(s) &= \text{softmax}_a Q(s, a) = \max Q(s, a) \\ &\quad + \log [1 + \exp \{\min Q(s, a) - \max Q(s, a)\}], \end{aligned}$$

where  $p(s'|s, a)$  is a state transition probability, and  $\text{softmax}_a$  is a soft version of the maximum (Kitani et al. 2012). Using these probabilistic Bellman equations, the policy  $\pi$  of the MaxEnt IRL framework is defined by

$$\pi(a|s; \theta) \propto \exp(Q(s, a) - V(s)).$$

Unlike a deterministic policy, which takes a single action at every state,  $\pi(a|s; \theta)$  is a probability distribution of actions (Fig. 1c). With this stochastic policy, actions with higher rewards are more

likely to be selected, and states of trajectories are considered to be sampled from the policy.

In IRL methods, reward values are usually defined by a reward function linear in features:

$$R(\zeta; \theta) = \sum_t r(s_t; \theta) = \sum_t \theta^T f(s_t),$$

where  $\theta$  is a weight vector, and  $f(s_t)$  is a feature vector obtained at state  $s_t$  along  $\zeta$ . In our problem, we used 12 feature maps (Fig. 2c). Each feature map represents the degree to which a location belongs to the physical environments and consists of one of three types: (1) binary maps for indicating sea and land, (2) exponentiated distance maps from the sea, land, and coastline, and (3) a constant map. The details of the definitions of these feature maps are described in Appendix S2.

Given a stochastic policy  $\pi(a|s; \theta)$ , the expected accumulated reward is calculated by taking the expectation with respect to  $p_\pi(\zeta)$ , which is the probability that trajectory  $\zeta$  is generated based on the policy (Abbeel and Ng 2004, Ziebart et al. 2008, Kitani et al. 2012), as follows:

$$\begin{aligned} E_{p_\pi(\zeta)}[R(\zeta; \theta)] &= E_{p_\pi(\zeta)} \left[ \sum_t \theta^T f(s_t) \right] \\ &= \theta^T E_{p_\pi(\zeta)} \left[ \sum_t f(s_t) \right]. \end{aligned}$$

The aim of MaxEnt IRL is to find a reward function that matches the observed shearwaters' actions. More formally,  $\bar{\pi}$  is an unknown policy of observed shearwater trajectories  $Z$ , and MaxEnt IRL tries to find this policy. This means that the expected feature vector  $E_{p_\pi(\zeta)} \left[ \sum_t f(s_t) \right]$  matches the feature vector obtained from the observed shearwater trajectories  $E_{p_\pi(\zeta)} \left[ \sum_t f(s_t) \right] = \frac{1}{|Z|} \sum_{\zeta \in Z} \left[ \sum_t f(s_t) \right] = \bar{f}$ .

This feature match is, however, ill-posed (Abbeel and Ng 2004, Ziebart et al. 2008). Hence, MaxEnt IRL chooses the policy  $\pi$  that maximizes the entropy

$$H(p_\pi) = - \int_{\zeta} p_\pi(\zeta) \log p_\pi(\zeta) d\zeta,$$

with the constraints of  $\int_{\zeta} p_\pi(\zeta) d\zeta = 1$ ,  $p_\pi(\zeta) \geq 0$ , and feature match  $E_{p_\pi(\zeta)} \left[ \sum_t f(s_t) \right] = \bar{f}$ . Using

Lagrange multipliers  $\theta$ , which is the weight parameter to be estimated, the probability distribution with maximum entropy along with the constraints is given by (Jaynes 1957, Berger et al. 1996)

$$p_\pi(\zeta|\theta) = \frac{\exp(\sum_t \theta^T f(s_t))}{Z(\theta)} = \frac{\exp(\sum_t \theta^T f(s_t))}{\int_{\zeta'} \exp(\sum_t \theta^T f(s'_t)) d\zeta'},$$

where  $Z(\theta)$  is a partition function defined in the denominator of the second equation (note that  $Z(\theta)$  is different from  $Z$ , which is a set of trajectories). Parameter  $\theta$  is estimated by maximizing the log-likelihood

$$\hat{\theta} = \operatorname{argmax}_\theta L(\theta) = \operatorname{argmax}_\theta \frac{1}{|Z|} \sum_{\zeta \in Z} \log p_\pi(\zeta|\theta),$$

and the exponentiated gradient method (Kivinen and Warmuth 1997) is used to update the weight by

$$\theta \leftarrow \theta e^{\lambda \nabla L(\theta)},$$

until convergence with an appropriate initial value and learning rate  $\lambda$ . The gradient (Ziebart et al. 2008) is given by

$$\begin{aligned} \nabla L(\theta) &= \frac{1}{|Z|} \sum_{\zeta \in Z} \nabla \log p_\pi(\zeta|\theta) \\ &= \frac{1}{|Z|} \sum_{\zeta \in Z} \left[ \sum_t f(s_t) - \right. \\ &\quad \left. \nabla \log \int_{\zeta'} \exp \left( \theta^T \sum_t f(s'_t) \right) d\zeta' \right] \\ &= \bar{f} - \frac{1}{|Z|} \sum_{\zeta \in Z} \nabla \log \int_{\zeta'} \exp \left( \theta^T \sum_t f(s'_t) \right) d\zeta' \\ &= \bar{f} - \nabla \log \int_{\zeta'} \exp \left( \theta^T \sum_t f(s'_t) \right) d\zeta' \\ &= \bar{f} - \int_{\zeta'} \sum_t f(s'_t) \frac{\exp(\sum_t \theta^T f(s'_t))}{\int_{\zeta''} \exp(\sum_t \theta^T f(s''_t)) d\zeta''} d\zeta' \\ &= \bar{f} - \int_{\zeta'} \sum_t f(s'_t) p_\pi(\zeta|\theta) d\zeta' \\ &= \bar{f} - E_{p_\pi(\zeta'|\theta)} \left[ \sum_t f(s'_t) \right]. \end{aligned}$$

In the last line, the second term is the expected feature vector of policy  $\pi$  and is approximated by

$$E_{p_\pi(\zeta|\theta)} \left[ \sum_t f(s_t) \right] = \frac{1}{|Z|} \sum_{\zeta \in Z} \sum_t f(s_t) D_\zeta(s_t),$$

where  $D_\zeta(s_t)$  is an expected state visitation count that expresses the probability of being in a certain state  $s_t$  of trajectory  $\zeta$  and is efficiently computed by backward and forward passes (Ziebart et al. 2008, Kitani et al. 2012). In the backward pass, the probabilistic Bellman equations are applied repeatedly until convergence to obtain policy  $\pi(a|s; \theta)$ . In the forward pass,  $D_\zeta(s)$  is propagated by

$$D_\zeta^{k+1}(s) = \sum_{s', a} D_\zeta^k(s') \pi(a|s; \theta) p(s'|s, a).$$

After the convergence,  $D_\zeta(s)$  is computed by summing over the iterations as  $D_\zeta(s) = \sum_k D_\zeta^k(s)$ .

To estimate the optimal weight vector  $\hat{\theta}$ , computation of  $D_\zeta(s)$  and updating  $\theta$  are iterated.

*Interpolation.*—Given a missing part, we set the beginning and end states of the missing part as  $s_0$  and  $s_g$ . Then, we estimate policy  $\pi(a|s; \hat{\theta})$  using a backward pass with estimated weight  $\hat{\theta}$ . Given  $s_g$ , the backward pass computes the probabilistic Bellman equations,  $Q(s, a)$  and  $V(s)$ , from  $s_g$  (Kitani et al. 2012). Then, we compute a stochastic policy  $\pi(a|s; \theta)$  from the estimated  $Q(s, a)$  and  $V(s)$ . The deterministic trajectory for the missing part is obtained by repeatedly selecting actions. At each state, we select the action  $\operatorname{argmax}_a \pi(a|s; \hat{\theta})$  that has maximum probability under the estimated policy at the state. This action selection and the state transition are repeated from start  $s_0$  until the agent reaches goal  $s_g$ . The resultant set of state-action pairs is the interpolated trajectory. Note that this procedure is deterministic and hence the same interpolation is obtained for the same pair of  $s_0$  and  $s_g$ ; however, it is possible to probabilistically interpolate trajectories by sampling actions according to the estimated policy.

Our implementation is available on GitHub: [https://github.com/thirakawa/MaxEnt\\_IRL\\_trajectory\\_interpolation](https://github.com/thirakawa/MaxEnt_IRL_trajectory_interpolation) (<https://doi.org/10.5281/zenodo.1402941>).

*Model evaluation.*—To evaluate the accuracy of our IRL interpolation method, we utilized the holdout method for male and female datasets. We randomly divided the trajectory dataset into two subsets: One was used for training, that is, estimating the optimal weight vector, and the other was

used for the evaluation of interpolation accuracy. Therefore, we randomly selected approximately 80% (43 trajectories for each sex) for training and used the remaining 10 trajectories for testing. For these test trajectories, we artificially created missing parts. Specifically, we randomly selected two points from the former and latter parts of a trajectory, respectively. The selected two points were set as the beginning and end of a missing part of the trajectory and we interpolated the missing part. To confirm the effect of our interpolation method on large missing parts, we used only artificial missing parts that were at least 50% as long as the whole trajectory.

For quantitative evaluation, we compared our interpolation method with the linear interpolation method, which is very common and widely used in ecology (Tancell et al. 2013, Carneiro et al. 2014, Granadeiro et al. 2014, Camprasse et al. 2017, Mendez et al. 2017, Poli et al. 2017, Yoda et al. 2017, Lone et al. 2018, Nishizawa et al. 2018). As an evaluation metric, we used the modified Hausdorff distance (MHD; Dubuisson and Jain 1994). The MHD is a type of distance metric that is used to compute the similarity between several object shapes. Because our method deals with trajectories, we computed the similarity of the ground-truth and interpolated trajectories as two-dimensional lines (curves). A trajectory can be expressed as a set of points. Given two trajectories,  $A = \{a_1, \dots, a_{N_a}\}$  and  $B = \{b_1, \dots, b_{N_b}\}$ , the MHD between these two trajectories  $MHD(A, B)$  is defined as

$$MHD(A, B) = \max \left\{ \frac{1}{N_a} \sum_{a \in A} d(a, B), \frac{1}{N_b} \sum_{b \in B} d(b, A) \right\}.$$

Here,  $d(a, B)$ , the distance between a point  $a$  and a set of points  $B$ , is defined as

$$d(a, B) = \min_{b \in B} \|a - b\|,$$

where  $\|a - b\|$  is Euclidean distance. We used the Mann–Whitney  $U$  test to compare the performance of each interpolation.

In our experiments, we set the initial value of the weight vector  $\theta$  to 1.0. The learning rate  $\lambda$  was experimentally determined to be 0.01. We defined the state transition probability as

$$p(s'|s, a) = \begin{cases} 1 & \text{if } s' \text{ is the next state from } s \text{ by } a \\ 0 & \text{otherwise.} \end{cases}$$

This means that once an action is decided, the agent state uniquely transitions toward the move direction.

Finally, we applied our method to interpolate actual trajectories with large missing parts to demonstrate a practical application of the IRL interpolation method.

## RESULTS

Our IRL-based interpolation method constructed trajectories that enabled the shearwaters to avoid flying over land (Figs. 5 and 6), while the linear interpolation method constructed trajectories that required flying over land. Several trajectory predictions for the artificially created missing paths matched almost perfectly with the original ones (e.g., Fig. 5c, d, f, and i for males and Fig. 6a, c, f, h for females). Other results provided different trajectories when compared with the ground truth (e.g., Fig. 5a, b, g, h, j for males and Fig. 6b, g, i, j for females). However, these results were still better than the results of the linear interpolation methods. The limitations of our IRL-based interpolation method are illustrated in Figs. 5e and 6g, in which the beginning and end of a missing part were rather close. When interpolating small missing parts, the reward map does not affect the action decision. Hence, actions moving directly toward the goal have higher probabilities. Therefore, the interpolation results become close to a straight line and are almost the same as the results of the linear interpolation method.

The average MHD showed that our IRL-based interpolation outperformed the linear interpolation method by a factor of 2.3 for males and 2.7 for females (Table 2). Similar to the results obtained using the test data, our method filled the gaps in actual trajectories with very large missing parts (Fig. 7), which are not included in the dataset.

Our IRL-based interpolation method learns the optimal rewards from training data; thus, we estimated optimal weight vector  $\theta$ , whose elements indicate the degree of influence of each feature map (Fig. 8). Of all the feature maps, sea maps had relatively larger weights, while land maps had smaller weights. This indicates that the shearwaters were affected by factors related to the sea, and flight routes over the sea are given

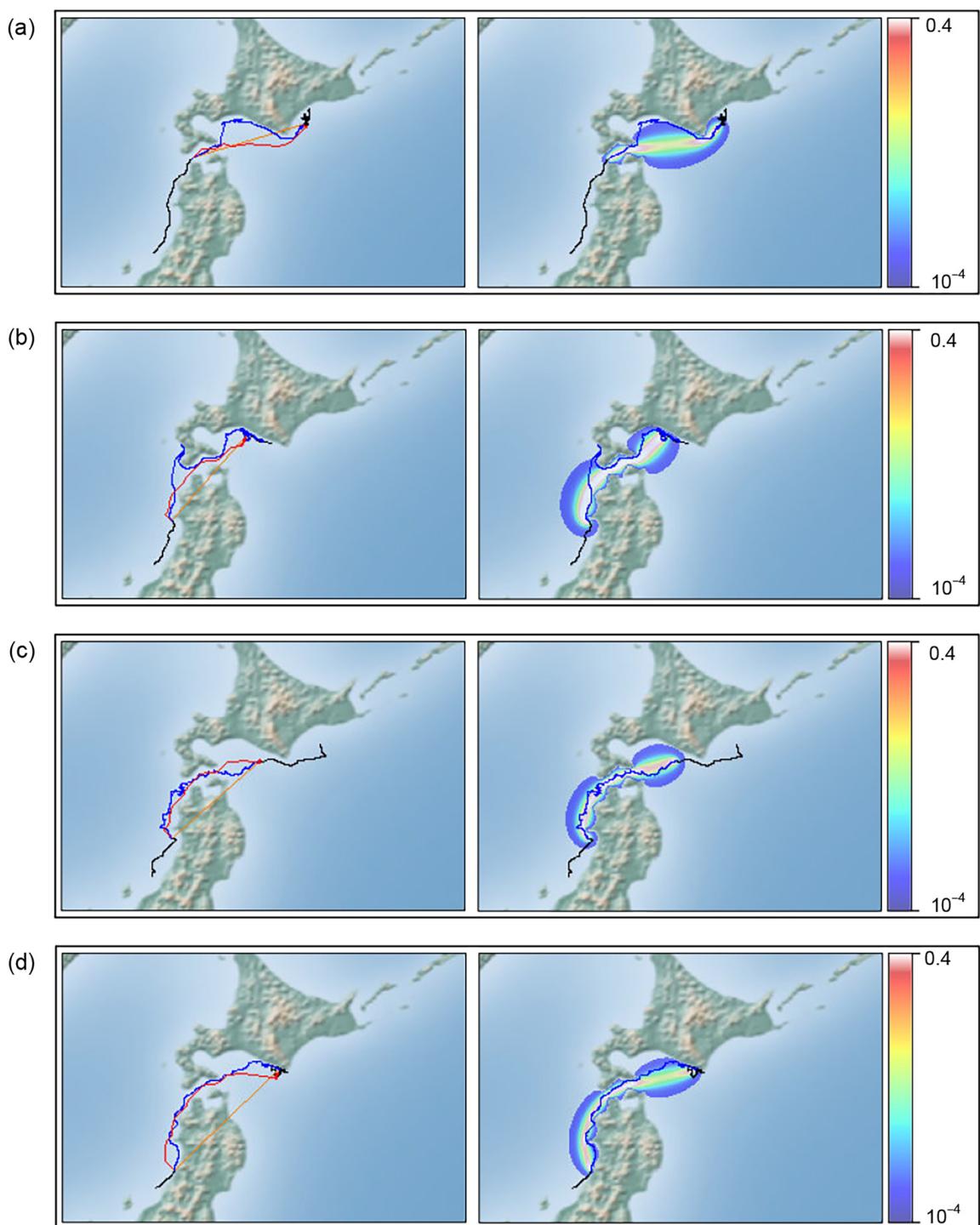
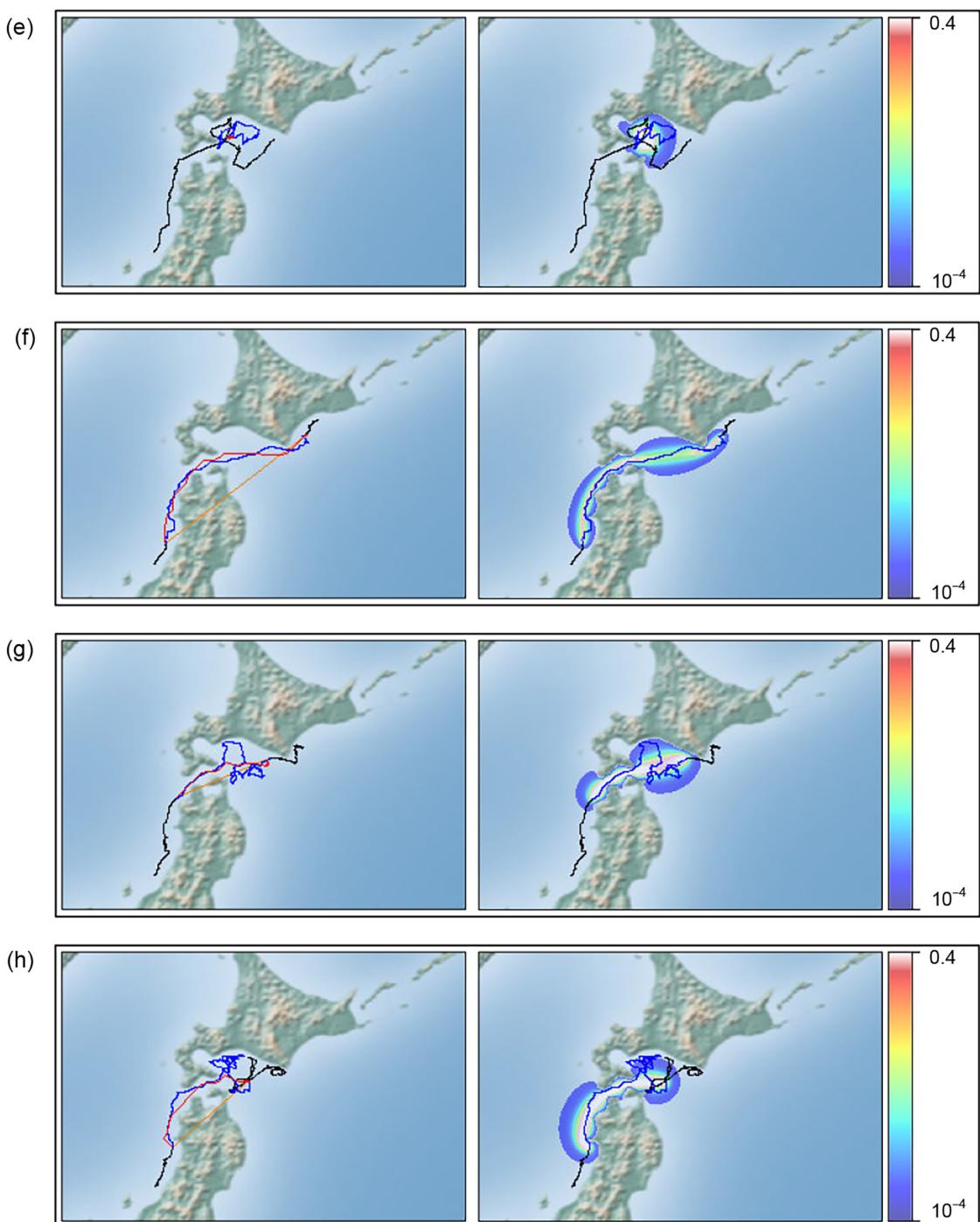
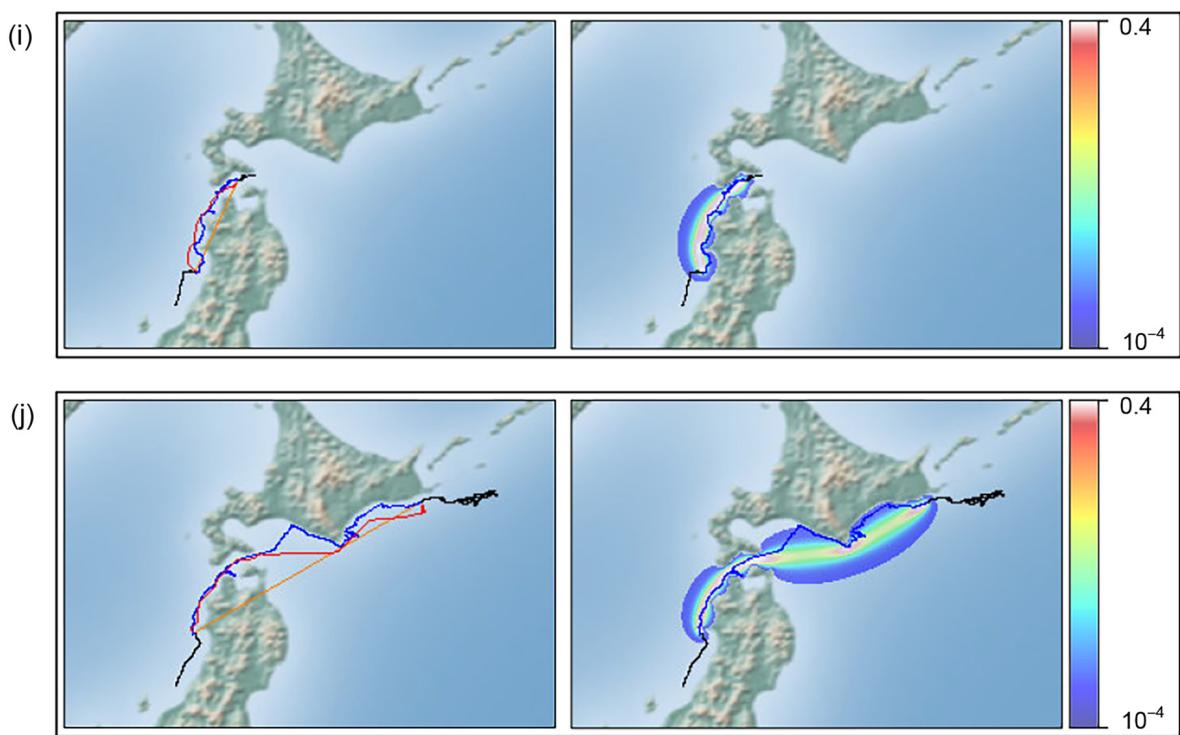


Fig. 5. Interpolation results of 10 test trajectories and probability distributions for male streaked shearwaters. In each panel (a–j), the left figures show the interpolated trajectories, and the right figures show the probability distributions estimated by our method. In the interpolated trajectories, the black lines show the trajectories that were recorded by GPS loggers, and the blue lines show trajectories in which missing parts were artificially created in the black lines.

(Fig. 5. *Continued*)

The red and orange lines show the trajectories interpolated by our method and linear interpolation, respectively. In the probability distributions, the distributions are shown as a heat map: Higher probabilities are shown as warmer colors, and lower probabilities are shown as cooler colors. Very low probabilities ( $\leq 10^{-4}$ ) are not shown in this figure.



(Fig. 5. *Continued*)

a higher reward. In contrast, coastline maps and constant maps did not affect behavior. Males and females showed a generally similar but slightly different pattern in the weights (Fig. 8).

## DISCUSSION

The recent explosion of movement data means that a methodology that can take full advantage of the data volume is required (Demšar et al. 2015, Kays et al. 2015). In particular, data preprocessing methods such as spatiotemporal interpolation have become important to an unprecedented degree (Boyd et al. 2004). In this study, we introduced IRL, which allows us to predict realistic paths and enhances our ability to fill gaps in animal trajectories. Our method has three advantages when compared to previously used interpolation methods: (1) Assumptions about the parametric distribution of animal movement are not required, (2) assumptions about animals' preferences for and restrictions regarding landscape are not required, and (3) the ability to fill large spatiotemporal gaps. Moreover, it is able to avoid look-ahead bias and does not repeat earlier

trajectory patterns. The IRL predicted that shearwaters avoid flying over land, which is consistent with a previous report (Yoda et al. 2017), without needing the condition to be explicitly added to the model. In addition, our result shows that IRL helps bridge the gaps in GPS locations much more accurately over a scale of tens of kilometers than the most widely used linear interpolation method.

Machine learning methods including RL and IRL are concerned with making more accurate predictions at the expense of interpretability (Valletta et al. 2017). This contrasts with mechanistic models (e.g., curvilinear and RW models) to make assumptions and inferences on movement parameters, which results in more explainable models. However, our knowledge about animal movements is typically incomplete, and therefore, mechanistic models might consider only several aspects of the mechanisms in complex real-world systems. Our approach overcomes this limitation by focusing on predictions given an amount of data at the expense of explainability for the mechanisms underlying animal movements. The IRL framework enables us to correct movement data and further help mechanistic models reveal more

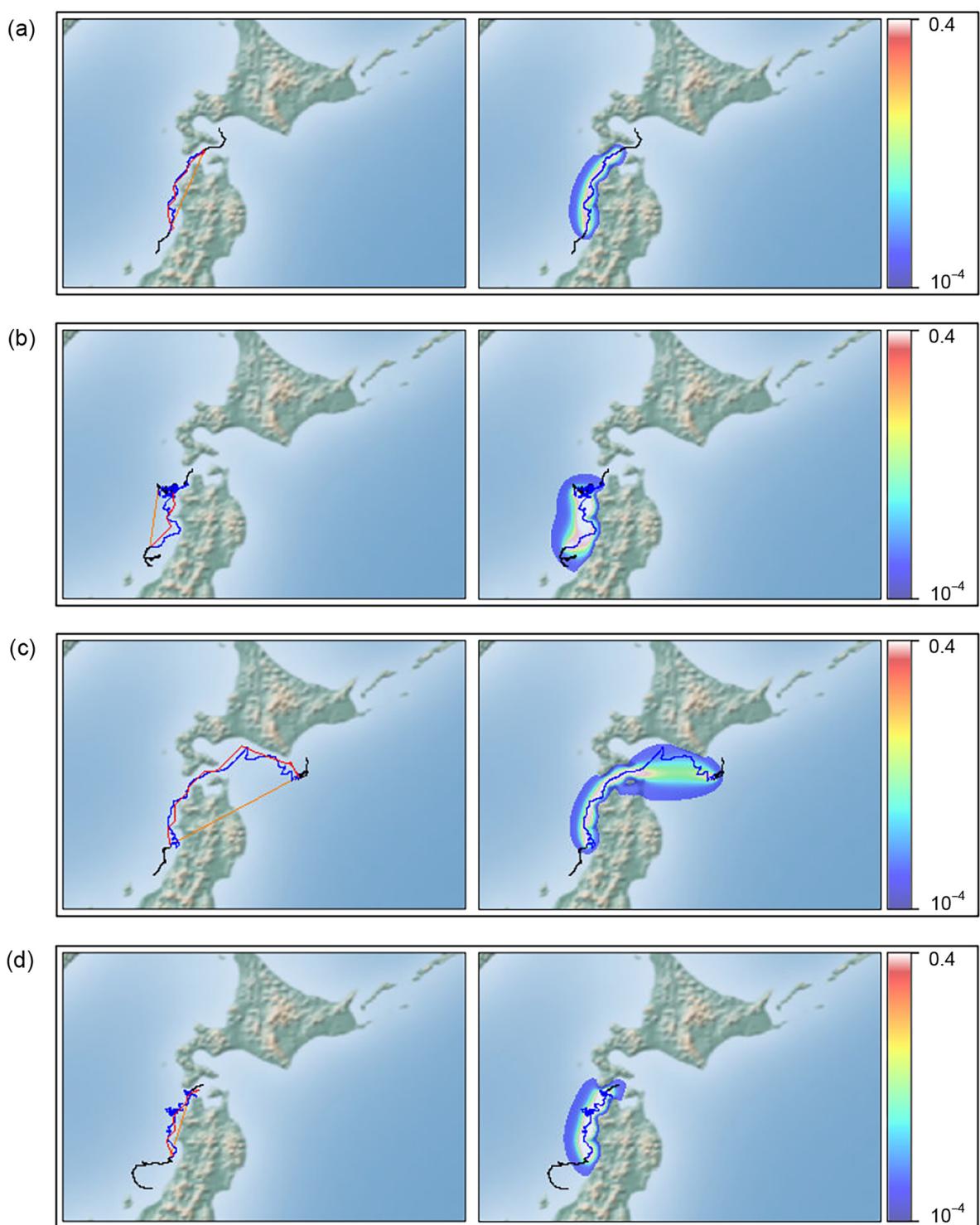
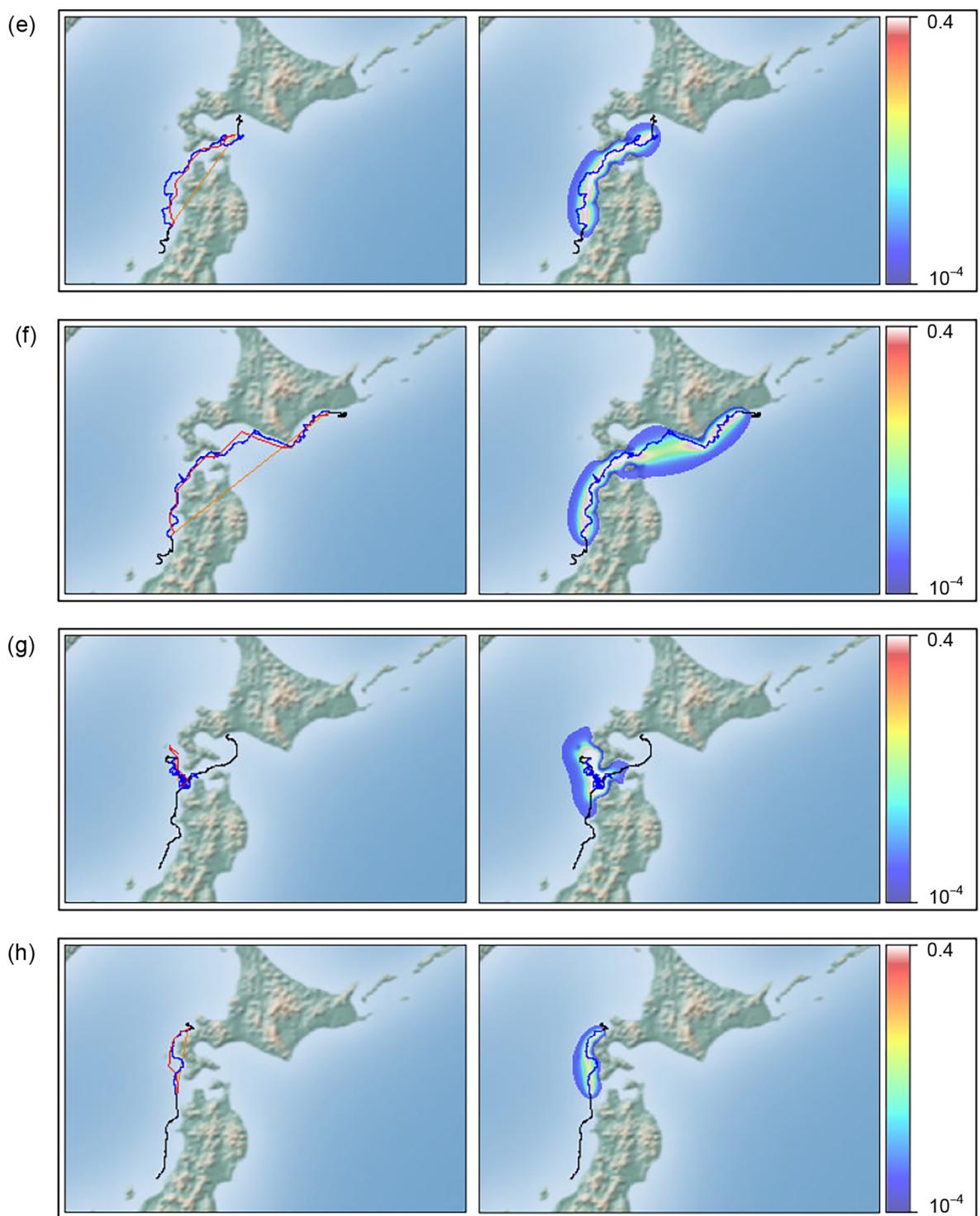


Fig. 6. Interpolation results of 10 test trajectories and probability distributions for female streaked shearwaters. Plots (a–j) are as in Fig. 5.

(Fig. 6. *Continued*)

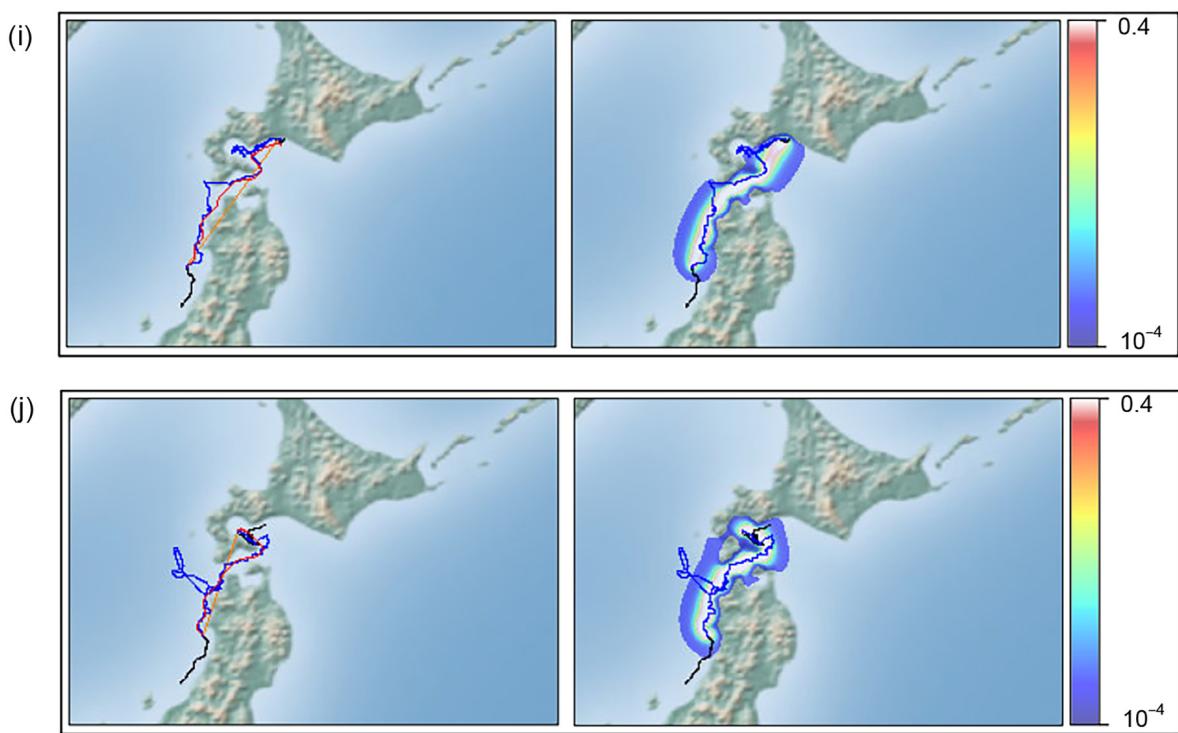
(Fig. 6. *Continued*)

Table 2. Mean MHD (modified Hausdorff distance) of interpolated results for males and females based on linear and IRL interpolations.

Sex	Linear interpolation	IRL interpolation	P-value
Males	$12.204 \pm 4.483$	$5.331 \pm 2.608$	0.001
Females	$9.367 \pm 5.271$	$3.443 \pm 1.654$	0.0009

Notes: A smaller value indicates better accuracy. Statistical comparisons of the performances of each interpolation use the Mann–Whitney  $U$  tests.

about animal movements as well as the evaluation of habitat use in animals.

Our model produced adequate results for most ecological research purposes (e.g., home range estimation) by filling the gaps in GPS locations, but led to some interpolation errors, especially for convoluted paths (e.g., Fig. 5e, h). These paths are often interpreted as animal foraging or searching for prey (area restricted search, ARS; Kareiva and Odell 1987), and they might be difficult to interpolate using either our method or previous methods. However, if feature maps related to ARS are included, this problem could be resolved. For

example, seabird foraging behavior is associated with biotic (e.g., chlorophyll *a*) and abiotic (e.g., sea surface temperature) environments (Tremblay et al. 2009b); therefore, ARS might be predictable if additional feature maps are included. Ecologists now have access to various environmental data obtained from satellites (Running et al. 2004, Neumann et al. 2015) and drones (Turner 2014, Zhang et al. 2016, Strandburg-Peshkin et al. 2017), which could be incorporated as feature maps. Further, because animal trajectories might be affected by spatiotemporally changing landscape features that are both environmental (e.g., weather, wind, time of day, lunar rhythm; Baigas et al. 2017) and social factors (e.g., the presence of congeners; Yoda et al. 2011), IRL has the potential for increasing interpolation precision by incorporating these dynamically changing factors into the feature space (Lee and Kitani 2016). Because we extended the state space to three-dimensional grid cells, our interpolation method has much potential for considering dynamically changing factors. Therefore, our future work includes introducing dynamic feature maps and interpolating

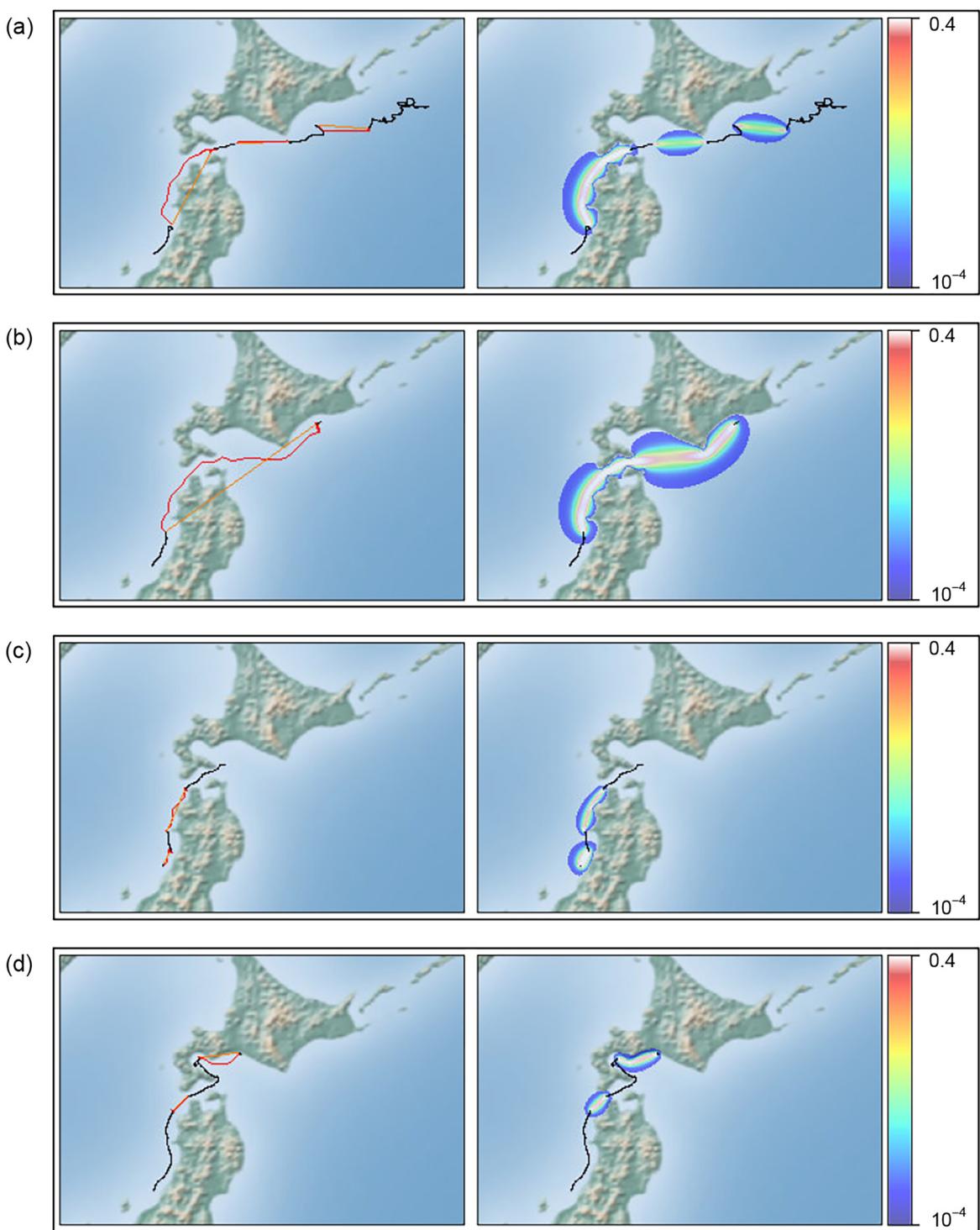


Fig. 7. Interpolation results of four actual male and female trajectories and probability distributions. Our IRL-based method was applied to actual trajectories with large missing parts that were obtained by a shearwater-borne GPS logger. In each of the panels (a–d), the plots are the same as those of Figs. 5 and 6, but without blue lines because of the real missing parts. (a–c) Interpolation results of male trajectories and (d) interpolation results of female trajectories.

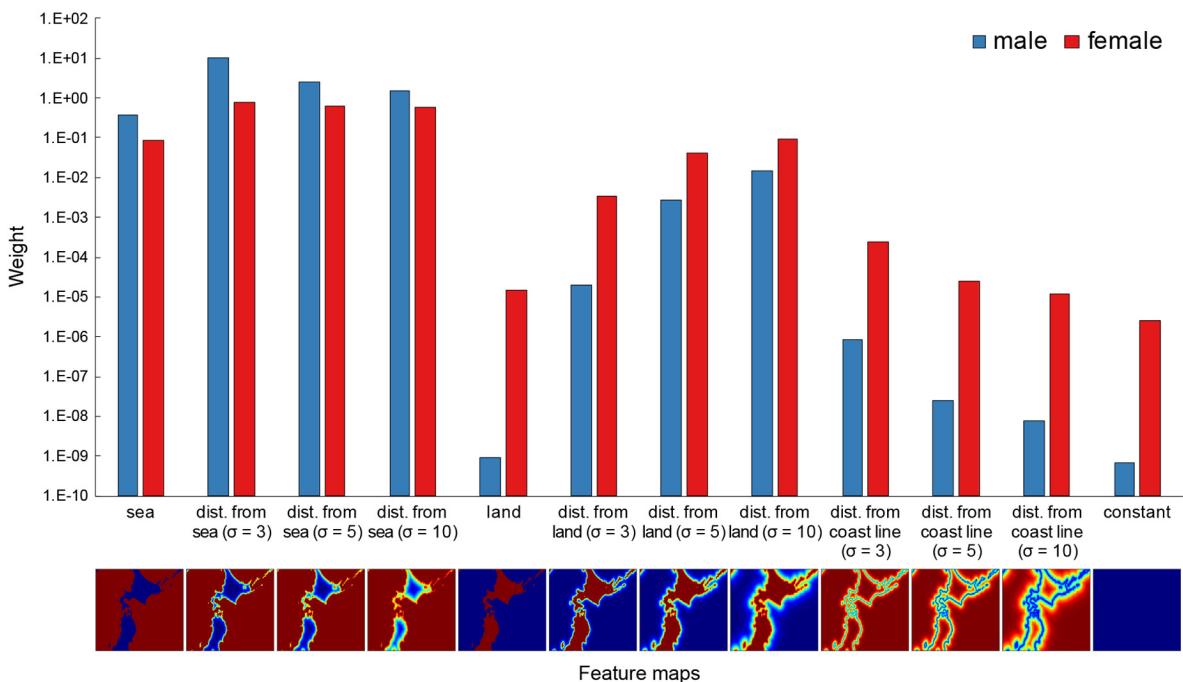


Fig. 8. Relationship between feature maps and learned weight vector. Learned weights indicate the degree to which each feature map (i.e., the physical environment) contributes to the estimation of the reward for shearwaters' behavior. For example, sea map and distance from sea maps have higher reward for both sexes; this indicates that the shearwaters prefer moving over the sea and avoiding moving over land. The weight of the constant map is low, indicating that the shearwaters do not follow the shortest path to reach a goal.

trajectories with convoluted animal foraging paths. Finally, our IRL method computes the weight of each feature map and determines which environmental predictors are actually important for predicting trajectories (Fig. 8). This might help environmental variable selections in models that include mechanistic interactions between animals and their environments (Kearney 2006, Lonergan et al. 2009). In addition, different feature map weights were obtained for males and females; females were more likely to be affected by distance from the coastline. This supports the notion of habitat differences between sexes in this species (Matsumoto et al. 2017) and indicates the importance of the inclusion of individual attributes for movement prediction.

A potential weakness of IRL might come from its need for a moderate to large amount of training data; the greater the number of movement tracks used, the more accurate the interpolation becomes. However, we used 43 trajectories for each sex as training datasets and obtained

reasonable interpolation results; this amount of data has been obtained in recent tracking studies (Bélisle et al. 2001, Carey et al. 2009). In actuality, tracking datasets are becoming larger as the era of big data is coming to ecology (Block et al. 2011), driven by the increasing availability of low-cost loggers and real-time transmitters (Wilmers et al. 2015).

We used the outward parts of the round-trips in the evaluations; however, we can also interpolate the homeward parts as well. The difference between the outward and homeward parts is merely the start and end locations, and the proposed method can interpolate gaps in the homeward parts without any modification. This is because the backward pass computes information from any goal state using the state transition probability  $p(s'|s, a)$ , which is symmetric with respect to  $s$  and  $s'$ .

In the state representation, we used the discrete time step  $z_t$ , which ignores the actual timing information, that is, how long the shearwater

stayed in the grid cell. The timing information is an important factor, and our future work will include how to manage this information. Nevertheless, the proposed method is still useful in terms of location information because ecological studies often estimate and analyze home ranges without exact timing information (e.g., Robinson et al. 2012, Yoda et al. 2012).

We applied our method to the GPS tracking data of seabirds, but our approach has broad applicability to other species, including aquatic/terrestrial and large/small animals. It can also be applied to other tracking methods such as the visual tracking of focal animals (e.g., primates; Schliehe-Diecks et al. 2012) and automated image-based tracking (Dell et al. 2014). Some extensively used tracking methods, such as VHF, Argos transmitters, and light-level geolocation sensors, provide several locations per hour or day. Studies using these intermittent locations, which are not sufficient to reconstruct time-series animal movements, might significantly benefit from our method, although the imprecise estimates in locations measured by Argos or geolocation sensors should be pre-processed (e.g., Kearney 2006, Winship et al. 2012). In addition, interpolation by IRL can be applied to a variety of fields such as fishery management (Russo et al. 2011) and livestock management (Pandey et al. 2009) to interpolate the trajectories of vessels and livestock, respectively.

Another potential use of IRL is the construction of cost- and risk-space maps in heterogeneous environments that are consistent with animal behavior (Fig. 4). The locations of suitable and non-suitable habitats facilitate or constrain animal movements, affecting species distribution and migration corridors (Barton et al. 2015). Environmentally dependent suitability and costs have been modeled as niche models (Kearney 2006), energy landscapes (Wilson et al. 2012), and landscapes of fear (Gallagher et al. 2017), all of which are likely to predict route choice and species distribution. Previous methods are based on simplistic and mechanistic calculations of costs and risks induced by environmental gradients such as slope inclines and flow (Wall et al. 2006, Shepard et al. 2013), because it is difficult to manually specify the parameters for complex real-world movement trajectories. A big advantage of IRL is its ability to construct reward maps for movement in

heterogeneous environments by observing demonstrators (i.e., animals) rather than by manually specifying a reward function. Thus, IRL can integrate energy landscapes, landscapes of fear, and any other drivers of animal distribution in space to predict how animals are distributed (Strandburg-Peshkin et al. 2017) and to provide information for conservation applications (Barton et al. 2015).

In conclusion, we demonstrated that IRL can fill the gaps between observed GPS points obtained from animals. Originally, RL was inspired by reinforcement in the context of the operant behavior of animals (Skinner 1938). Thus, one of the key contributions of our research is to reimport this concept into ecology and open the possibility of applying the framework of RL and IRL in this field. Because the framework is general enough to handle non-trajectory representations, such as sequences of discrete states of behavior (Kitani et al. 2012), it is possible to estimate the behavior between each pair of observed behaviors. Therefore, with the advent of high-resolution observation data, IRL has the potential to serve as a fundamental tool for ecological research on animals.

## ACKNOWLEDGMENTS

This research was funded by the Grants-in-Aid for Scientific Research from the Japan Society of the Promotion of Science (JP16H01769, JP16H06538, JP16H06540, JP16H06541, JP16K21735).

## LITERATURE CITED

- Abbeel, P., and A. Y. Ng. 2004. Apprenticeship Learning via Inverse Reinforcement Learning. International Conference on Machine Learning (ICML) 1–8.
- Appelhans, T., E. Mwangomo, D. R. Hardy, A. Hemp, and T. Nauss. 2015. Evaluating machine learning approaches for the interpolation of monthly air temperature at Mt. Kilimanjaro, Tanzania. *Spatial Statistics* 14:91–113.
- Auger-Méthé, M., A. E. Derocher, M. J. Plank, E. A. Codling, and M. A. Lewis. 2015. Differentiating the Lévy walk from a composite correlated random walk. *Methods in Ecology and Evolution* 6:1179–1189.
- Baigas, P. E., J. R. Squires, L. E. Olson, J. S. Ivan, and E. K. Roberts. 2017. Using environmental features to model highway crossing behavior of Canada lynx in the Southern Rocky Mountains. *Landscape and Urban Planning* 157:200–213.

- Bailey, H., G. Shillinger, D. Palacios, S. Bograd, J. Spotila, F. Paladino, and B. Block. 2008. Identifying and comparing phases of movement by leatherback turtles using state-space models. *Journal of Experimental Marine Biology and Ecology* 356:128–135.
- Barton, P. S., et al. 2015. Guidelines for using movement science to inform biodiversity policy. *Environmental Management* 56:791–801.
- Battaile, B. C., C. A. Nordstrom, N. Liebsch, and A. W. Trites. 2015. Foraging a new trail with northern fur seals (*Callorhinus ursinus*): Lactating seals from islands with contrasting population dynamics have different foraging strategies, and forage at scales previously unrecognized by GPS interpolated dive data. *Marine Mammal Science* 31:1494–1520.
- Béliste, M., A. Desrochers, and M. J. Fortin. 2001. Influence of forest cover on the movements of forest birds: a homing experiment. *Ecology* 82:1893–1904.
- Benhamou, S. 2006. Detecting an orientation component in animal paths when the preferred direction is individual-dependent. *Ecology* 87:518–528.
- Benhamou, S., and D. Cornélis. 2010. Incorporating movement behavior and barriers to improve kernel home range space use estimates. *Journal of Wildlife Management* 74:1353–1360.
- Berger, A. L., S. A. Della Pietra, and V. J. Della Pietra. 1996. A maximum entropy approach to natural language processing. *Computational Linguistics* 22:39–71.
- Bidder, O. R., J. S. Walker, M. W. Jones, M. D. Holton, P. Urge, D. M. Scantlebury, N. J. Marks, E. A. Magowan, I. E. Maguire, and R. P. Wilson. 2015. Step by step: reconstruction of terrestrial animal movement paths by dead-reckoning. *Movement Ecology* 3:23.
- Block, B. A., et al. 2011. Tracking apex marine predator movements in a dynamic ocean. *Nature* 475:86–90.
- Boyd, I. L., A. Kato, and Y. Ropert-Coudert. 2004. Bio-logging science: sensing beyond the boundaries. *Memories of National Institute of Polar Research* 58:1–14.
- Brost, B. M., M. B. Hooten, E. M. Hanks, and R. J. Small. 2015. Animal movement constraints improve resource selection inference in the presence of telemetry error. *Ecology* 96:2590–2597.
- Browning, E., M. Bolton, E. Owen, A. Shoji, T. Guilford, and R. Freeman. 2018. Predicting animal behaviour using deep learning: GPS data alone accurately predict diving in seabirds. *Methods in Ecology and Evolution* 9:681–691.
- Buchin, K., S. Sijben, E. E. van Loon, N. Sapir, S. Mercier, T. Arseneau, and E. P. Willems. 2015. Deriving movement properties and the effect of the environment from the Brownian bridge movement model in monkeys and birds. *Movement Ecology* 3:18.
- Cain, J. W., P. R. Krausman, B. D. Jansen, and J. R. Moggart. 2005. Influence of topography and GPS fix interval on GPS collar performance. *Wildlife Society Bulletin* 33:926–934.
- Camprasse, E. C. M., G. J. Sutton, M. Berlincourt, and J. P. Y. Arnould. 2017. Changing with the times: Little penguins exhibit flexibility in foraging behaviour and low behavioural consistency. *Marine Biology* 164:169.
- Carey, M. J., C. E. Meathrel, and N. A. May. 2009. A new method for the long-term attachment of data-loggers to shearwaters (Procellariidae). *Emu* 109:310–315.
- Carneiro, A. P. B., A. Manica, and R. A. Phillips. 2014. Foraging behaviour and habitat use by brown skuas *Stercorarius linnbergi* breeding at South Georgia. *Marine Biology* 161:1755–1764.
- Chen, Y. F., M. Everett, M. Liu, and J. P. How. 2017. Socially aware motion planning with deep reinforcement learning. *International Conference on Intelligent Robots on Systems (IROS)*:1343–1350.
- Codling, E. A., M. J. Plank, and S. Benhamou. 2008. Random walk models in biology. *Journal of the Royal Society Interface* 5:813–834.
- Creel, S., J. A. Winnie, and D. Christianson. 2013. Underestimating the frequency, strength and cost of antipredator responses with data from GPS collars: an example with wolves and elk. *Ecology and Evolution* 3:5189–5200.
- Dayan, P., and G. E. Hinton. 1992. Feudal reinforcement learning. *Advances in Neural Information Processing (NIPS)* 1:271–278.
- de Jager, M., F. J. Weissling, P. M. J. Herman, B. A. Nolet, and J. van de Koppel. 2011. Lévy walks evolve through interaction between movement and environmental complexity. *Science* 332:1551–1553.
- Dell, A. I., et al. 2014. Automated image-based tracking and its application in ecology. *Trends in Ecology & Evolution* 29:417–428.
- Demšar, U., K. Buchin, F. Cagnacci, K. Safi, B. Speckmann, N. van de Weghe, D. Weiskopf, and R. Weibel. 2015. Analysis and visualisation of movement: an interdisciplinary review. *Movement Ecology* 3:5.
- Dewhurst, O. P., H. K. Evans, K. Roskilly, R. J. Harvey, T. Y. Hubel, and A. M. Wilson. 2016. Improving the accuracy of estimates of animal path and travel distance using GPS drift-corrected dead reckoning. *Ecology and Evolution* 6:6210–6222.
- Dubuisson, M.-P., and A. K. Jain. 1994. A modified Hausdorff distance for object matching. *International Conference on Pattern Recognition (ICPR)*: 566–568.
- Edwards, A. M., et al. 2007. Revisiting Lévy flight search patterns of wandering albatrosses, bumblebees and deer. *Nature* 449:1044–1049.

- Fleming, C. H., W. F. Fagan, T. Mueller, K. A. Olson, P. Leimgruber, and J. M. Calabrese. 2016. Estimating where and how animals travel: an optimal framework for path reconstruction from autocorrelated tracking data. *Ecology* 97:576–582.
- Focardi, S., and J. G. Cecere. 2014. The Lévy flight foraging hypothesis in a pelagic seabird. *Journal of Animal Ecology* 83:353–364.
- Frair, J. L., J. Fieberg, M. Hebblewhite, F. Cagnacci, N. J. DeCesare, and L. Pedrotti. 2010. Resolving issues of imprecise and habitat-biased locations in ecological analyses using GPS telemetry data. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365:2187–2200.
- Freitas, C., C. Lydersen, M. A. Fedak, and K. M. Kovacs. 2008. A simple new algorithm to filter marine mammal Argos locations. *Marine Mammal Science* 24:315–325.
- Gallagher, A. J., S. Creel, R. P. Wilson, and S. J. Cooke. 2017. Energy landscapes and the landscape of fear. *Trends in Ecology & Evolution* 32:88–96.
- Godwin, E. M., M. J. Noad, E. Kniest, and R. A. Dunlop. 2016. Comparing multiple sampling platforms for measuring the behavior of humpback whales (*Megaptera novaeangliae*). *Marine Mammal Science* 32:268–286.
- Granadeiro, J. P., P. Brickle, and P. Catry. 2014. Do individual seabirds specialize in fisheries' waste? The case of black-browed albatrosses foraging over the Patagonian Shelf. *Animal Conservation* 17:19–26.
- Graves, T. A., and J. S. Waller. 2006. Understanding the causes of missed global positioning system telemetry fixes. *Journal of Wildlife Management* 70:844–851.
- Guilford, T., J. Meade, J. Willis, R. Phillips, D. Boyle, S. Roberts, M. Collett, R. Freeman, and C. M. Perrins. 2009. Migration and stopover in a small pelagic seabird, the Manx shearwater *Puffinus puffinus*: insights from machine learning. *Proceedings of the Royal Society B: Biological Sciences* 276:1215–1223.
- Hooten, M. B., D. S. Johnson, B. T. McClintock, and J. M. Morales. 2017. Animal movement: statistical models for telemetry data. CRC Press, Boca Raton, Florida, USA.
- Horne, J. S., E. O. Garton, S. M. Krone, and J. S. Lewis. 2007. Analyzing animal movements using Brownian bridges. *Ecology* 88:2354–2363.
- Humphries, N. E., et al. 2010. Environmental context explains Lévy and Brownian movement patterns of marine predators. *Nature* 465:1066–1069.
- Hussey, N. E., et al. 2015. Aquatic animal telemetry: a panoramic window into the underwater world. *Science* 348:1255642.
- Jaradat, M., M. Al-Rousan, and L. Quadan. 2011. Reinforcement based mobile robot navigation in dynamic environment. *Robotics and Computer-Integrated Manufacturing* 27:135–149.
- Jaynes, E. T. 1957. Information theory and statistical mechanics. *Physical Review* 106:620–630.
- Kaelbling, L. P., M. L. Littman, and A. W. Moore. 1996. Reinforcement learning: a survey. *Journal of Artificial Intelligence Research* 4:237–285.
- Kareiva, P., and G. Odell. 1987. Swarms of predators exhibit “preytaxis” if individual predators use area-restricted search. *American Naturalist* 130: 233–270.
- Kays, R., M. C. Crofoot, W. Jetz, and M. Wikelski. 2015. Terrestrial animal tracking as an eye on life and planet. *Science* 348:aaa2478.
- Kearney, M. 2006. Habitat, environment and niche: What are we modelling? *Oikos* 115:186–191.
- Kitani, K. M., B. D. Ziebart, J. A. Bagnell, and M. Hebert. 2012. Activity forecasting. *European Conference on Computer Vision (ECCV)*:201–214.
- Kivinen, J., and M. K. Warmuth. 1997. Exponentiated gradient versus gradient descent for linear predictors. *Information and Computation* 132:1–63.
- Lee, N., and K. M. Kitani. 2016. Predicting wide receiver trajectories in American football. *The IEEE Winter Conference on Applications of Computer Vision (WACV)*:1–9.
- Levine, S., P. Pastor, A. Krizhevsky, and D. Quillen. 2016. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *International Symposium on Experimental Robotics (ISER)*:173–184.
- Lone, K., K. M. Kovacs, C. Lydersen, M. Fedak, M. Andersen, P. Lovell, and J. Aars. 2018. Aquatic behaviour of polar bears (*Ursus maritimus*) in an increasingly ice-free Arctic. *Scientific Reports* 8:9677.
- Lonergan, M., M. Fedak, and B. McConnell. 2009. The effects of interpolation error and location quality on animal track reconstruction. *Marine Mammal Science* 25:275–282.
- Martin, P., and P. Bateson. 2017. Measuring behavior. Cambridge University Press, Cambridge, UK.
- Matsumoto, S., T. Yamamoto, M. Yamamoto, C. Zavalaga, and K. Yoda. 2017. Sex-related differences in the foraging movement of streaked shearwaters *Calonectris leucomelas* breeding on Awashima Island in the Sea of Japan. *Ornithological Science* 16:23–32.
- McClintock, B. T., J. M. London, M. F. Cameron, and P. L. Boveng. 2017. Bridging the gaps in animal movement: Hidden behaviors and ecological relationships revealed by integrated data streams. *Ecosphere* 8:e01751.
- Mendez, L., A. Prudor, and H. Weimerskirch. 2017. Ontogeny of foraging behaviour in juvenile red-footed boobies (*Sula sula*). *Scientific Reports* 7:13886.

- Morales, J. M., D. T. Haydon, J. Frair, K. E. Holsinger, and J. Fryxell. 2004. Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology* 85:2436–2445.
- Murino, V., M. Cristani, S. Shah, and S. Savarese. 2017. Group and crowd behavior for computer vision. Academic Press, Cambridge, Massachusetts, USA.
- Nathan, R., W. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz, and P. Smouse. 2008. A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences of the United States of America* 105:19052–19059.
- Neumann, W., S. Martinuzzi, A. B. Estes, A. M. Pidgeon, H. Dettki, G. Ericsson, and V. C. Radeloff. 2015. Opportunities for the application of advanced remotely-sensed data in ecological studies of terrestrial animal movement. *Movement Ecology* 3:8.
- Ng, A. Y., and S. Russell. 2000. Algorithms for inverse reinforcement learning. *International Conference on Machine Learning (ICML)*: 663–670.
- Nishizawa, B., T. Sugawara, L. C. Young, E. A. Vanderwerf, K. Yoda, and Y. Watanuki. 2018. Albatross-borne loggers show feeding on deep-sea squids: implications for the study of squid distributions. *Marine Ecology Progress Series* 592:257–265.
- Northrup, J. M., C. R. Anderson, M. B. Hooten, and G. Wittemyer. 2016. Movement reveals scale dependence in habitat selection of a large ungulate. *Ecological Applications* 26:2746–2757.
- Pandey, V., G. A. Kiker, K. L. Campbell, M. J. Williams, and S. W. Coleman. 2009. GPS monitoring of cattle location near water features in south Florida. *Applied Engineering in Agriculture* 25:551–562.
- Pedersen, M. W., T. A. Patterson, U. H. Thygesen, and H. Madsen. 2011. Estimating animal behavior and residency from movement data. *Oikos* 120:1281–1290.
- Pelletier, L., A. Chiaradia, A. Kato, and Y. Ropert-Coudert. 2014. Fine-scale spatial age segregation in the limited foraging area of an inshore seabird species, the little penguin. *Oecologia* 176:399–408.
- Poli, C. L., A.-L. Harrison, A. Vallarino, P. D. Gerard, and P. G. R. Jodice. 2017. Dynamic oceanography determines fine scale foraging behavior of masked boobies in the Gulf of Mexico. *PLoS ONE* 12: e0178318.
- Resheff, Y. S., S. Rotics, R. Harel, O. Spiegel, and R. Nathan. 2014. AcceleRater: a web application for supervised learning of behavioral modes from acceleration measurements. *Movement Ecology* 2:27.
- Robinson, P. W., et al. 2012. Foraging behavior and success of a mesopelagic predator in the northeast Pacific Ocean: insights from a data-rich species, the northern elephant seal. *PLoS ONE* 7:e36728.
- Rowcliffe, J. M., C. Carbone, R. Kays, B. Kranstauber, and P. A. Jansen. 2012. Bias in estimating animal travel distance: the effect of sampling frequency. *Methods in Ecology and Evolution* 3:653–662.
- Royer, F., and M. Lutcavage. 2008. Filtering and interpreting location errors in satellite telemetry of marine animals. *Journal of Experimental Marine Biology and Ecology* 359:1–10.
- Running, S. W., R. R. Nemani, F. A. Heinsch, M. Zhao, M. Reeves, and H. Hashimoto. 2004. A continuous satellite-derived measure of global terrestrial primary production. *BioScience* 54:547–560.
- Russell, S. 1998. Learning agents for uncertain environments (extended abstract). *International Conference on Machine Learning (ICML)*:278–287.
- Russo, T., A. Parisi, and S. Cataudella. 2011. New insights in interpolating fishing tracks from VMS data for different métiers. *Fisheries Research* 108:184–194.
- Ryan, P. G., S. Petersen, G. Peters, and D. Grémillet. 2004. GPS tracking a marine predator: the effects of precision, resolution and sampling rate on foraging tracks of African penguins. *Marine Biology* 145:215–223.
- Schliebe-Diecks, S., M. Eberle, and P. M. Kappeler. 2012. Walk the line-dispersal movements of gray mouse lemurs (*Microcebus murinus*). *Behavioral Ecology and Sociobiology* 66:1175–1185.
- Shepard, E., R. P. Wilson, W. G. Rees, E. Grundy, S. A. Lambertucci, and S. B. Vosper. 2013. Energy landscapes shape animal movement ecology. *American Naturalist* 182:298–312.
- Singh, S. P., A. G. Barto, R. Grupen, and C. Connolly. 1994. Robust reinforcement learning in motion planning. *Neural Information Processing Systems (NIPS)* 1:655–662.
- Skinner, B. F. 1938. The behavior of organisms. Appleton-Century-Crofts, New York, New York, USA.
- Smart, W. D., and L. P. Kaelbling. 2002. Effective reinforcement learning for mobile robots. *International Conference on Robotics and Automation (ICRA)* 4:3404–3410.
- Song, C., Z. Qu, N. Blumm, and A. L. Barabasi. 2010. Limits of predictability in human mobility. *Science* 327:1018–1021.
- Strandburg-Peshkin, A., D. R. Farine, M. C. Crofoot, and I. D. Couzin. 2017. Habitat and social factors shape individual decisions and emergent group structure during baboon collective movement. *eLife* 6:e19505.
- Stulp, F. 2012. Reinforcement learning with sequences of motion primitives for robust manipulation. *IEEE Transactions on Robotics* 28:1360–1370.
- Sutton, R. S., and A. G. Barto. 1998. Reinforcement learning: an introduction. MIT Press, Cambridge, Massachusetts, USA.

- Tancell, C., R. A. Phillips, J. C. Xavier, G. A. Tarling, and W. J. Sutherland. 2013. Comparison of methods for determining key marine areas from tracking data. *Marine Biology* 160:15–26.
- Technitis, G., W. Othman, K. Safi, and R. Weibel. 2015. From A to B, randomly: a point-to-point random trajectory generator for animal movement. *International Journal of Geographical Information Science* 29:912–934.
- Tomkiewicz, S. M., M. R. Fuller, J. G. Kie, and K. K. Bates. 2010. Global positioning system and associated technologies in animal behaviour and ecological research. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365:2163–2176.
- Torres, L. G., D. R. Thompson, S. Bearhop, S. Votier, G. A. Taylor, P. M. Sagar, and B. C. Robertson. 2011. White-capped albatrosses alter fine-scale foraging behavior patterns when associated with fishing vessels. *Marine Ecology Progress Series* 428:289–301.
- Tremblay, Y., P. W. Robinson, and D. P. Costa. 2009a. A parsimonious approach to modeling animal movement data. *PLoS ONE* 4:e4711.
- Tremblay, Y., S. Bertrand, R. W. Henry, M. A. Kappes, D. P. Costa, and S. A. Shaffer. 2009b. Analytical approaches to investigating seabird-environment interactions: a review. *Marine Ecology Progress Series* 391:153–163.
- Tremblay, Y., et al. 2006. Interpolation of animal tracking data in a fluid environment. *Journal of Experimental Biology* 209:128–140.
- Turchin, P. 1998. Quantitative analysis of movement. Sinauer Associates, Sunderland, Massachusetts, USA.
- Turner, W. 2014. Sensing biodiversity. *Science* 346:301–302.
- Valletta, J. J., C. Torney, M. Kings, A. Thornton, and J. Madden. 2017. Applications of machine learning in animal behaviour studies. *Animal Behaviour* 124:203–220.
- Wall, J., I. Douglas-Hamilton, and F. Vollrath. 2006. Elephants avoid costly mountaineering. *Current Biology* 16:R527–R529.
- Weinstein, B. G. 2017. A computer vision for animal ecology. *Journal of Animal Ecology* 87:533–545.
- Wilmers, C. C., B. Nickel, C. M. Bryce, J. A. Smith, R. E. Wheat, and V. Yovovich. 2015. The golden age of bio-logging: How animal-borne sensors are advancing the frontiers of ecology. *Ecology* 96: 1741–1753.
- Wilson, R. P., F. Quintana, and V. J. Hobson. 2012. Construction of energy landscapes can clarify the movement and distribution of foraging animals. *Proceedings of the Royal Society B: Biological Sciences* 279:975–980.
- Wilson, R., et al. 2007. All at sea with animal tracks; methodological and analytical solutions for the resolution of movement. *Deep Sea Research Part II: Topical Studies in Oceanography* 54:193–210.
- Winship, A. J., S. J. Jorgensen, and S. A. Shaffer. 2012. State-space framework for estimating measurement error from double-tagging telemetry experiments. *Methods in Ecology and Evolution* 3:291–302.
- Yoda, K., M. Murakoshi, K. Tsutsui, and H. Kohno. 2011. Social interactions of juvenile Brown Boobies at sea as observed with animal-borne video cameras. *PLoS ONE* 6:e19602.
- Yoda, K., K. Shiomi, and K. Sato. 2014. Foraging spots of streaked shearwaters in relation to ocean surface currents as identified using their drift movements. *Progress in Oceanography* 122:54–64.
- Yoda, K., T. Tajima, S. Sasaki, K. Sato, and N. Niizuma. 2012. Influence of local wind conditions on the flight speed of the great cormorant *Phalacrocorax carbo*. *International Journal of Zoology* 2012: 187102.
- Yoda, K., T. Yamamoto, H. Suzuki, S. Matsumoto, M. Müller, and M. Yamamoto. 2017. Compass orientation drives naïve pelagic seabirds to cross mountain ranges. *Current Biology* 27:R1152–R1153.
- Zhang, J., J. Hu, J. Lian, Z. Fan, X. Ouyang, and W. Ye. 2016. Seeing the forest from drones: testing the potential of lightweight drones as a tool for long-term forest monitoring. *Biological Conservation* 198:60–69.
- Ziebart, B. D., A. Maas, J. A. Bagnell, and A. K. Dey. 2008. Maximum entropy inverse reinforcement learning. *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence* 8:1433–1438.
- Ziebart, B. D., N. Ratliff, G. Gallagher, C. Mertz, K. Peterson, J. A. Bagnell, M. Hebert, A. K. Dey, and S. Srinivasa. 2009. Planning-based prediction for pedestrians. *International Conference on Intelligent Robots (IROS)*:3931–3936.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online at: <http://onlinelibrary.wiley.com/doi/10.1002/ecs2.2447/full>