Preserving the integrity of the Canadian northern ecosystems through insights provided by reinforcement learning-based Arctic fox movement models

Catherine Villeneuve [1 2 3] **Frédéric Dulude-de Broin** [1 2 4] **Pierre Legagneux** [1 5 4 6] **Dominique Berteaux** [7 4 6]
Audrey Durand [8 1 2 3]

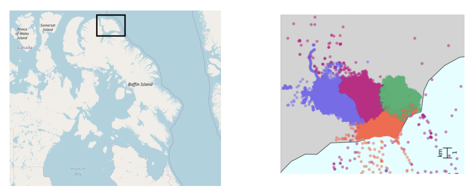
Abstract

Realistic modeling of the movement of the Arctic fox, one of the main predators of the circumpolar world, is crucial to understand the processes governing the distribution of the Canadian Arctic biodiversity. Current methods, however, are unable to adequately account for complex behaviors as well as intra- and interspecific relationships. We propose to harness the potential of reinforcement learning to develop innovative models that will address these shortcomings and provide the backbone to predict how vertebrate communities may be affected by environmental changes in the Arctic, an essential step towards the elaboration of rational conservation actions.

1. Introduction

The Canadian Arctic Archipelago is one of the most vulnerable places in the world to climate change. Its ecosystems, closely tied to the extreme conditions of the North, are particularly fragile and their integrity is threatened by a warming rate faster than lower-latitude parts of the globe (Bush & Lemmen, 2019). Because Arctic food webs are composed of a few highly interconnected species, negative impacts affecting one component of the system are likely to have cascading effects on all of the others. Understanding how disruptions of specific trophic interactions scale up to the entire system is key to anticipate the consequences of global environmental changes, but large scale modeling of ecological interactions is challenging and the ecosystem monitoring required to parametrize them are rare (Godsoe et al., 2017).

One of the flagship species for the study of climate change in the Arctic is the Arctic fox (*Vulpes lagopus*) (Ehrich et al., 2015). It is its only endemic terrestrial predator (Fugley & Ims, 2008) and its populations are currently monitored at 34 sites throughout the circumpolar North (Berteaux et al., 2017). In Canada, they are notably studied through a research station on *Bylot Island*, an 11 100 km^2 island located at the northern tip of Baffin Island, Nunavut (Figure 1a). Indeed, the movements of Bylot foxes are actively monitored with GPS tracking collars since 2007 (see Figure 1b).



(a) (b)

Figure 1. (a) Location of Bylot Island (box) on a map of Canada. (b) GPS location every 4 minutes of four foxes over a three month period in 2018 and 2019 on a map of Bylot.

The study of predators is particularly important in the Canadian Arctic, as they are the dominant force controlling its food web (Krebs et al., 2003; Gauthier et al., 2011). A change in predator-prey interactions can significantly disrupt the ecosystem dynamics, and understanding how predation affects the distribution of Arctic biodiversity is critical to preserve its integrity in the face of environmental changes. Such an understanding can only be achieved with the help of *mathematical models* able to transform field data into powerful explanatory and predictive tools.

Bylot's Arctic foxes tracking data, coupled with satellite maps and prey abundance information, can be used to model their movements at the individual level. Realistic modeling of animal movement is a crucial first step to understand the processes regulating species distribution and abundance. Indeed, in heterogeneous landscapes, movement ties ecological processes together and dictates

[1]Université Laval [2]Institut intelligence et données (IID) [3]Mila [4]Centre d'Études Nordiques (CEN) [5]CNRS — Centre d'Études Biologiques de Chizé (CEBC) [6]Centre de la science de la biodiversité du Québec (CSBQ) [7]Université du Québec à Rimouski (UQAR) [8]Canada CIFAR AI Chair. Correspondence to: Catherine Villeneuve <catherine.villeneuve.9@ulaval.ca>.

Tackling Climate Change with Machine Learning workshop at ICML 2021

how species interact with each other (Cagnacci et al., 2017). Current methods in ecology are however insufficient to capture complex movement decisions and their sequential nature adequately. On the other hand, *reinforcement learning* (RL) (Sutton & Barto, 2018) methods, where an agent tries to optimize its decisions through trial and error, could be well suited to solve such tasks. They are still fairly unexploited in ecology (Frankenhuis et al., 2019), even if there is a growing interest in the application of machine learning to this field (Hirakawa et al., 2018; Lucas, 2020; Wijeyakulasuriya et al., 2020).

Our goal is to bridge this gap and harness the potential of RL to model Arctic biodiversity. Using extensive ecological data (Section 2) from the long-term monitoring program of Bylot Island, we aim to build models to predict movement decisions of Arctic foxes in a terrestrial community. This will provide the backbone for mechanistic models integrating species interactions to predict how vertebrate communities may be affected by environmental changes in the Arctic.

In this work, we present to which extent current methods fail to accurately model the movement of Arctic foxes (Section 3), and how we propose to use RL to address this issue (Section 4). We explain how we plan to demonstrate that RL approaches can do at least as well as the current methods used in ecology, and how we intend to use this technology to build significantly more realistic movement models, both in single-agent and multi-agent settings. Computer scientists and ecologists involved in the project will work in close collaboration to develop relevant models with strong generalization capabilities, so that the insights they provides can be reliably used to further our understanding of the mechanisms shaping Arctic biodiversity and help policymakers to guide management decisions in a changing environment.

2. Data

We will use the 4-minute interval GPS tracking data from 21 foxes during the summer of 2018 and 2019 ($687,129$ records), available in Movebank[1]. In addition, it is planned that ecologists will collect additional tracking data every summer for the 2021-2024 period. We also have access to the geographical location of 110 Bylot dens. Habitat and topography maps will be extracted from 0.3-m WorldView-3[2] (WV3) satellite images of Bylot Island that will be acquired during Summer 2021. A habitat map associates each pixel of a satellite image to an element of a discrete set of habitat types such as ice, complex wetlands, gravel beds, or wet meadows. We will generate a habitat map following the approach of Chen et al. (2017) with a training set that will be obtained from photointerpretations and manual classification of land sites in the field. We also have lower precision tracking data of 66 foxes (12-hour interval, $229,657$ records) from Lai et al. (2017) for the 2007-2013 period, which are nonetheless sufficiently accurate to help evaluate the realism of our models outside the areas covered by the 4-minute interval tracking data. Finally, as for the prey abundance information, we will use a map of snow goose nest density obtained through intensive field surveys (described in Grenier-Potvin et al. (2021)) and nest counts from the WV3 images. Lemming abundance will be determined based on their habitat preferences, as it is not possible yet to obtain a density map for this animal. Lemmings' habitat preferences are documented in Fauteux et al. (2015).

3. A fundamentally limited baseline

The most common way to model the movement of an animal at the individual level from the tracking data of N individuals in ecology is through *Step Selection Function* (SSF) (Fortin et al., 2005; Forester et al., 2009; Thurfjell et al., 2014; Avgar et al., 2016). Let $S_{i,1}, S_{i,2}, ..., S_{i,T}$ denote the recorded locations for an animal $i \in \{1, ..., N\}$, where each location $S_{i,t}$ is characterized by d explanatory variables (e.g. habitat type, prey density, ground elevation, distance to den) denoted as $\mathbf{s}_{i,t} \in \mathbb{R}^d$. The SSF is a *discrete choice model* that compares the recorded locations to J *control locations* based on their explanatory variables. Control locations for a given animal correspond to locations that could have been visited by this individual. They are usually obtained from the empirical distributions of *movements metrics* such as step lengths (the Euclidean distance between two consecutive locations) and turning angles (Fortin et al., 2005). Let $\mathbf{Y}_{i,t} \in \mathbb{R}^{J \times d}$ denote the matrix containing the explanatory variables $\mathbf{y}_{i,1t}, ..., \mathbf{y}_{i,Jt}$ of the J control locations of the animal i at step t. The data for an SSF analysis is therefore $\mathcal{D} = \{(\mathbf{s}_{i,t}, \mathbf{s}_{i,t+1}, \mathbf{Y}_{i,t}) \,|\, 1 \leq t \leq T-1 \wedge 1 \leq i \leq N\}$ and the probability that animal i moves to the location $S_{i,t+1}$ instead of any control location is defined as

$$p_{i,t} := \frac{\exp(\mathbf{s}_{i,t+1}^{\intercal}\beta)}{\exp(\mathbf{s}_{i,t+1}^{\intercal}\beta) + \sum_{j=0}^{J}\exp(\mathbf{y}_{i,jt}^{\intercal}\beta)}, \tag{1}$$

where $\beta \in \mathbb{R}^d$ is a vector of unknown *selection parameters* estimated by maximizing the conditional regression likelihood given by $L(\beta) = \prod_{i=1}^{N}\prod_{t=1}^{T-1} p_{i,t}$ (Nicosia et al., 2017). The vector β can then be used within Eq. 1 in movement simulations to study the animal's behavior and preferences.

The SSF is simple and easy to implement, but it is a

[1] https://www.movebank.org/cms/webapp?gwt_fragment=page=studies,path=study1241071371

[2] http://worldview3.digitalglobe.com/

fundamentally limited approach. Indeed, it is based on an estimator assuming that $\mathcal{D}$ is made of *independent samples*. That is to say that the sequential nature of the data is ignored, even if it is inconsistent with the fact that the animal's space-use behavior is memory-based (Spencer, 2012). Furthermore, Arctic foxes concentrate their activities within areas called *home ranges* (see Figure 1b) and biased correlated random walk models, for which the SSF can be shown to be equivalent (Duchesne et al., 2015), are unable to faithfully recreate these patterns (Borger et al., 2008). The movement of the Arctic fox therefore cannot be realistically modeled using an SSF. As movement is a critical process affecting population numbers and the outcome of species interactions (Turchin, 2015), more realistic movement models of key predators must be developed in order to provide insights about how this important aspect of ecological dynamics will be affected by climate change in the Arctic.

4. From simple to realistic movement models

Aim 1 : Demonstrating the potential of RL

The assumptions made by the SSF are reminiscent of the ones defining the *contextual bandit* setting (see Lattimore & Szepesvári (2020)). In this simplified RL setting, a learning agent iteratively observes a context and is presented a set of available arms, among which it has to decide which one to pull. After pulling an arm, it receives a reward associated to the selected arm. The goal of the agent is to pull the arm leading to the highest reward given the context, assuming that its decision at a given time does not influence the future (arrival of contexts, available arms). We can formulate our problem under this setting by sampling at each time t a tuple $(\mathbf{s}_{i,t'}, \mathbf{s}_{i,t'+1}, \mathbf{Y}_{i,t'}) \sim \mathcal{D}$, define $\mathbf{s}_{i,t'}$ as the context, and consider that the set of available arms contains $\mathbf{s}_{i,t'+1}$ and the rows of $\mathbf{Y}_{i,t'}$. The agent is rewarded if it chooses $\mathbf{s}_{i,t'+1}$ and penalized otherwise. We will assume that the expected reward of a location is *linear with respect to its explanatory variables* (Chu et al., 2011; Abbasi-Yadkori et al., 2011; Agrawal & Goyal, 2013), so as to obtain a discrete choice model that should be very similar to the SSF. This initial model will allow us to demonstrate that RL approaches can educate us *at least as well* as the SSF on the basic aspects of the Arctic fox's space-use behavior.

Aim 2 : Enable the emergence of home range patterns and spatial memory management

The core of our project is to develop realistic models able to adequately capture the sequential nature of an Arctic fox's movement decisions. RL approaches are specifically designed to handle sequential data and have been applied in a variety of challenging tasks, like robotics (Kober et al., 2013) or autonomous driving (Sallab et al., 2017). We therefore strongly believe that they could be able to overcome the limitations of the SSF by replicating complex behaviors such as home range patterns or spatial memory management. Unlike contextual bandits, a typical RL agent assumes that each action may influence the whole trajectory and that a wrong choice may penalize it for all subsequent time steps.

Since the agent must learn from existing trajectories, we will leverage *offline RL strategies* (see Levine et al. (2020)), which can learn a behavior policy from demonstrations associated with a reward signal. The challenging task of shaping a reward from the tracking data will be conducted in close collaboration with ecologists, allowing to integrate expert knowledge into the RL models. Finally, we intend to leverage the approximation power of *neural networks* (LeCun et al., 2015) and their ability to handle non-linear relationships for learning rich representations that can allow precise behavior modeling. Home range patterns and spatial memory management are expected to emerge from these more complex models. A successful integration of these two key elements would represent a major advance for the study of the Arctic biodiversity in a context of global change.

Aim 3 : Reveal the impact of territorial dynamics on the repartition of Arctic biodiversity

Arctic foxes are territorial animals and their intraspecific relationships greatly influence their space use (Grenier-Potvin et al., 2021). Limiting our approach to a single-agent perspective will prevent us from considering the true nature of these competitive interactions. Inspired by a game theoretical analysis of models of interacting wolf packs, where the Nash-equilibrium resulting from the tradeoff between expanding a territory and avoiding conflict gave rise to no-wolf's lands where prey's are empirically known to find refuge (Hamelin & Lewis, 2010), we aim to extend our model to the multi-agent setting in order to have a better understanding of the impact of the complex dynamics between neighboring foxes on their ecological community. Modeling foxes territorial dynamics as an *n-player adversarial game* (Song et al., 2018) could help reveal the influence of these relationships on the repartition of Arctic biodiversity.

Metrics

We will split the tracking data into train and test sets in order to evaluate the quality of the proposed models based on their accuracy score when predicting unseen fox trajectories. Models will be re-evaluated and improved each year during the 2021-2024 period using new data collected through the monitoring program at Bylot. Using realism metrics defined by ecologists on specific key elements such as the home range size of a simulated fox, we will also evaluate the quality of complete trajectories generated by the proposed models. This will be essential in order for them to be considered and used within the community to help understand the role of predation in the regulation of Arctic ecosystems as well as to anticipate how they will be affected by climate change.

5. Towards mutual enrichment

The first step towards anticipating the impact of climate change on the Canadian northern ecosystems is to *understand* their key determinants. Our movement models will meet this need through simulations that will help ecologists to explain how variations in predation risk contributes to shape the local biodiversity. They could also help predict Arctic foxes' response to specific environmental disturbances. Moreover, this novel application of RL to a real-world problem should raise interesting challenges for the RL research community. For example, assumptions made by most theoretical work for providing *generalization guarantees* are often considered too restrictive (Dulac-Arnold et al., 2021). We strongly believe that this interdisciplinary project will contribute to bridging the gap between RL theory and practice while opening new research directions for supporting the development of rational conservation actions.

Acknowledgements

This work was supported by the Sentinelle Nord program from the Canada First Research Excellence Fund. We thank Mathieu Godbout for his insightful comments and everyone involved in the monitoring program at Bylot.

References

Abbasi-Yadkori, Y., Pál, D., and Szepesvári, C. Improved algorithms for linear stochastic bandits. In *Advances in Neural Information Processing Systems*, 2011.

Agrawal, S. and Goyal, N. Thompson sampling for contextual bandits with linear payoffs. In *Proceedings of the 30th International Conference on Machine Learning*, 2013.

Avgar, T., Potts, J. R., Lewis, M. A., and Boyce, M. S. Integrated step selection analysis: Bridging the gap between resource selection and animal movement. *Methods in Ecology and Evolution*, 7(5):619–630, 2016.

Berteaux, D., Thierry, A. M., Alisauskas, R., Angerbjörn, A., Buchel, E., Doronina, L., Ehrich, D., Eide, N. E., Erlandsson, R., Flagstad, O., Fuglei, E., Gilg, O., Goltsman, M., Henttonen, H., Ims, R. A., Killengreen, S. T., Kondratyev, A., Kruchenkova, E., Kruckenberg, H., Kulikova, O., Landa, A., Lang, J., Menyushina, I., Mikhnevich, J., Niemimaa, J., Norén, K., Ollila, T., Ovsyanikov, N., Pokrovskaya, L., Pokrovsky, I., Rodnikova, A., Roth, J. D., Sabard, B., Samelius, G., Schmidt, N. M., Sittler, B., Sokolov, A. A., Sokolova, N. A., Stickney, A., Unnsteinsdóttir, E. R., and White, P. A. Harmonizing circumpolar monitoring of arctic fox: benefits, opportunities, challenges and recommendations. *Polar Research*, 36(sup1):2, 2017.

Borger, L., Dalziel, B. D., and Fryxell, J. M. Are there general mechanisms of animal home range behaviour? a review and prospects for future research: Home range modelling. *Ecology Letters*, 11(6):637–650, 2008.

Bush, E. and Lemmen, D. S. Canada's changing climate report. Technical report, Government of Canada, Ontario, ON, 2019.

Cagnacci, F., Boitani, L., Powell, R. A., and Boyce, M. S. Offline reinforcement learning:tutorial, review, and perspectives on open problems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1550): 2157–2162, 2017.

Chen, Z., Pasher, J., Duffe, J., and Behnamian, A. Mapping arctic coastal ecosystems with high resolution optical satellite imagery using a hybrid classification approach. *Canadian Journal of Remote Sensing*, 43(6):513–527, 2017.

Chu, W., Li, L., Reyzin, L., and Schapire, R. E. Contextual bandits with linear payoff functions. In *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics*, pp. 208–214, 2011.

Duchesne, T., Fortin, D., and Rivest, L.-P. Equivalence between step selection functions and biased correlated random walks for statistical inference on animal movement. *PLOS One*, 10(4), 2015.

Dulac-Arnold, G., Levine, N., Mankowitz, D. J., Li, J., Paduraru, C., Gowal, S., and Hester, T. Challenges of real-world reinforcement learning: Definitions, benchmarks and analysis. *Machine Learning*, 2021.

Ehrich, D., Ims, R. A., Yoccoz, N. G., Lecomte, N., Killengreen, S. T., Fuglei, E., Rodnikova, A. Y., Ebbinge, B. S., Menyushina, I. E., Nolet, B. A., Pokrovsky, I. G., Popov, I. Y., Schmidt, N. M., Sokolov, A. A., Sokolova, N. A., and Sokolov, V. A. What can stable isotope analysis of top predator tissues contribute to monitoring of tundra ecosystems? *Ecosystems*, 18(3):404–416, 2015.

Fauteux, D., Gauthier, G., and Berteaux, D. Seasonal demography of a cyclic lemming population in the canadian arctic. *Journal of Animal Ecology*, 84(5):1412–1422, 2015.

Forester, J. D., Im, H. K., and Rathouz, P. J. Accounting for animal movement in estimation of resource selection functions: sampling and data analysis. *Ecology*, 90(12): 3554–3565, 2009.

Fortin, D., Beyer, H. L., Boyce, M. S., Smith, D. W., Duchesne, T., and Mao, J. S. Wolves influence elk movements: Behavior shapes a trophic cascade in yellowstone national park. *Ecology*, 86(5):1320–1330, 2005.

Frankenhuis, W. E., Panchanathan, K., and Barto, A. G. Enriching behavioral ecology with reinforcement learning methods. *Behavioural Processes*, 161:94–100, 2019.

Fugley, E. and Ims, R. A. Global warming and effects on the arctic fox. *Science Progress*, 91(2):175–191, 2008.

Gauthier, G., Bertaux, D., Bêty, J., Tarroux, A., Therrien, J. F., McKinnon, L., Legagneux, P., and Cadieux, M. C. The tundra food web of bylot island in a changing climate and the role of exchanges between ecosystems. *Écoscience*, 18(3):223–235, 2011.

Godsoe, W., Jankowski, J., Holt, R. D., and Gravel, D. Integrating biogeography with contemporary niche theory. *Trends in Ecology & Evolution*, 32(7):488–499, 2017.

Grenier-Potvin, A., Clermont, J., Gauthier, G., and Berteaux, D. Prey and habitat distribution are not enough to explain predator habitat selection: addressing intraspecific interactions, behavioural state and time. *Movement Ecology*, 9(1):12, 2021.

Hamelin, F. M. and Lewis, M. A. A differential game theoretical analysis of mechanistic models for territoriality. *Journal of Mathematical Biology*, 61(5):665–694, 2010.

Hirakawa, T., Yamashita, T., Tamaki, T., Fujiyoshi, H., Umezu, Y., Takeuchi, I., Matsumoto, S., and Yoda, K. Can ai predict animal movements? filling gaps in animal trajectories using inverse reinforcement learning. *Ecosphere*, 9(10), 2018.

Kober, J., Bagnel, J. A., and Peters, J. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, 32(11):1238–1274, 2013.

Krebs, C. J., Danell, K., Angerbjörn, A., Agrell, J., Berteaux, D., Brathen, K. A., Danell, O., Erlinge, S., Fedorov, V., Fredga, K., Hjältén, J., Högstedt, G., Jónsdóttir, I. S., Kenney, A. J., Kjellén, N., Nordin, T., Roininen, H., Svensson, M., Tannerfeldt, M., and Wiklund, C. Terrestrial trophic dynamics in the canadian arctic. *Canadian Journal of Zoology*, 81(5):827–843, 2003.

Lai, S., Bêty, J., and Berteaux, D. Movement tactics of a mobile predator in a meta-ecosystem with fluctuating resources: The arctic fox in the high arctic. *Oikos*, 126 (7):937–947, 2017.

Lattimore, T. and Szepesvári, C. *Bandit algorithms*. Cambridge University Press, 2020.

LeCun, Y., Bengio, Y., and Hinton, G. Deep learning. *Nature*, 521:436–444, 2015.

Levine, S., Kumar, A., Tucker, G., and Fu, J. Offline reinforcement learning:tutorial, review, and perspectives on open problems. *arXiv preprint arXiv:2005.01643*, 2020.

Lucas, T. C. D. A translucent box: Interpretable machine learning in ecology. *Ecological Monographs*, 90(4), 2020.

Nicosia, A., Duchesne, T., Rivest, L.-P., and Fortin, D. A multi-state conditional logistic regression model for the analysis of animal movement. *The Annals of Applied Statistics*, 11(3), 2017.

Sallab, A. E., Abdou, M., Perot, E., and Yogamani, S. Deep reinforcement learning framework for autonomous driving. *Electronic Imaging*, 2017(1):70–76, 2017.

Song, J., Ren, H., Sadigh, D., and Ermon, S. Multi-agent generative adversarial imitation learning. In *Advances in Neural Information Processing Systems*, 2018.

Spencer, W. D. Animal ecology meets gps-based radiotelemetry: a perfect storm of opportunities and challenges. *Journal of Mammalogy*, 93(4):929–947, 2012.

Sutton, R. S. and Barto, A. G. *Reinforcement Learning : An Introduction*. MIT Press, 2nd edition, 2018.

Thurfjell, H., Ciuti, S., and Boyce, M. S. Applications of step-selection functions in ecology and conservation. *Movement Ecology*, 2(1):4, 2014.

Turchin, P. *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants.* Beresta Books, 2015.

Wijeyakulasuriya, D. A., Eisenhauer, E. W., Shaby, B. A., and Hanks, E. M. Machine learning for modeling animal movement. *PLOS One*, 15(7), 2020.