Hierarchical, Memory-Based Movement Models for Translocated Elk (*Cervus canadensis*)

Andrea Falcón-Cortés[1]*, Denis Boyer[2], Evelyn Merrill[3], Jacqueline L. Frair[3] and Juan Manuel Morales[4]

[1] Instituto de Ciencias Físicas, Universidad Nacional Autónoma de México, Cuernavaca, Mexico, [2] Instituto de Física, Universidad Nacional Autónoma de México, Ciudad de México, Mexico, [3] Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada, [4] Grupo de Ecología Cuantitativa, INIBIOMA-CONICET, Universidad Nacional del Comahue, San Carlos de Bariloche, Argentina

The use of spatial memory is well-documented in many animal species and has been shown to be critical for the emergence of spatial learning. Adaptive behaviors based on learning can emerge thanks to an interdependence between the acquisition of information over time and movement decisions. The study of how spatio-ecological knowledge is constructed throughout the life of an individual has not been carried out in a quantitative and comprehensive way, hindered by the lack of knowledge of the information an animal already has of its environment at the time monitoring begins. Identifying how animals use memory to make beneficial decisions is fundamental to developing a general theory of animal movement and space use. Here we propose several mobility models based on memory and perform hierarchical Bayesian inference on 11-month trajectories of 21 elk after they were released in a completely new environment. Almost all the observed animals exhibited preferential returns to previously visited patches, such that memory and random exploration phases occurred. Memory decay was mild or negligible over the study period. The fact that individual elk rapidly become used to a relatively small number of patches was consistent with the hypothesis that they seek places with predictable resources and reduced mortality risks such as predation.

Keywords: memory-based movement models, spatial memory, attribute memory, animal learning, translocated elk, hierarchal Bayesian inference

1. INTRODUCTION

The use of spatial memory is well-documented in many animal species. For example, humans, non-human primates and other large-brained vertebrates make movement decisions based on spatial representations of their environments (Wills et al., 2010). These representations may allow animals to move directly to important sites in their environment that lie outside of their perceptual range (Normand and Boesch, 2009; Presotto and Izar, 2010), such as resource patches, sites that connect with other high quality sites in space, or safe spots to avoid predators, and may also allow them to estimate the travel cost to reach a particular place (Lanner, 1996; Janson, 2007; Janson and Byrne, 2007; Noser and Byrne, 2007). Another type of memory, described for the first time by Schacter (1992) and retaken by Fagan et al. (2013), encodes the attributes of landscape features under the name of *attribute memory*. While spatial memory allows animals to reduce uncertainty about

the location of geographical features, attribute memory reduces uncertainty concerning location-independent features of objects (Fagan et al., 2013). The information stored as attribute memory may be the abundance or types of food, and can be linked to spatial information. For example, food patch quality can be spatially encoded: patch quality is an attribute and its location is spatial information (Fagan et al., 2013). The combination of these two types of information allows animals to choose among alternative movement paths as has been observed in bumblebees (Lihoreau et al., 2011) or large herbivores (Avgar et al., 2013; Merkle et al., 2014). Identifying how animals use memory to make decisions is fundamental to developing a general theory of animal movement and space use (Gautestad and Mysterud, 2005; Morales et al., 2010; Spencer, 2012).

Memory is also critical in the emergence of spatial learning, which results from interactions with the environment and can be detected through changes in movement patterns (Mueller and Fagan, 2008). Adaptive behaviors based on learning can occur thanks to an interdependence between the acquisition of information over time and movement decisions (Falcón-Cortés et al., 2017, 2019). For instance, an animal can make decisions based on past successful experiences, resulting in a change of behavior and improved resource exploitation (Leonard, 1990; Bracis and Mueller, 2017; Jesmer et al., 2018; Merkle et al., 2019). Learning is consistent, for example, with frequent visits to certain locations, or site fidelity (Bonnell et al., 2013; Falcón-Cortés et al., 2017), and with the emergence of home range behavior or preferential travel routes (Van Moorter et al., 2009; Boyer and Walsh, 2010). The capability of learning can also bring other benefits beyond improved foraging; e.g., providing advantage in territorial defense (Potts and Lewis, 2014; Schlägel and Lewis, 2014; Schlägel et al., 2017), more effective escape from predators (Brown, 2001), and improving the route choice in migration (Bischof et al., 2012; Poor et al., 2012). Nevertheless, the connections between memory and spatial learning is not well understood. Theoretical models bring useful insights by predicting, for instance, how often memory should be used for the emergence of recurrent movements to a particular resource patch (Falcón-Cortés et al., 2017; Boyer et al., 2019).

Several theoretical studies have highlighted the role played by memory and cognitive abilities for foraging success (Boyer and Walsh, 2010), home range formation (Börger et al., 2008; Van Moorter et al., 2009; Berger-Tal and Avgar, 2012), and paved the way for inferring individual memory capacities from movement and environmental data (Avgar et al., 2013). The applications of these theoretical approaches to free-ranging animals are varied. For example, predictions of a simple memory model based on linear reinforcement through preferential revisits have been compared with the movements of capuchin monkeys, revealing movement rules found to generate very slow diffusion and heterogeneous space use (Boyer and Solis-Salas, 2014). On the other hand, Merkle et al. (2014) applied a patch-to-patch model to ranging data of American bison, finding that these animals remember valuable information about the location and quality of meadows (spatial and attribute memory) and use this information to revisit profitable locations.

The study of how spatio-ecological knowledge is constructed throughout the life of an individual has not been developed thoroughly. Data analyses that employ memory based models are promising but are often difficult to implement due to the short observation periods available, and the fact that the animals are observed in an environment already familiar to them. If memory is long-ranged, the above limitations may affect the results. To avoid these shortcomings, we used data from relocated animals. This means that the observed animals explored an unknown landscape at the start of their movement trajectories. In this new environment the spatial locations of different environmental features and patches were initially unknown to them. We analyzed the movement data from 21 relocated elk (*Cervus canadensis*) as described in Frair et al. (2007) and Wolf et al. (2009). We expected elk to show an initial exploratory phase in which the animals were getting familiarized with their new environment and collecting information about the location and quality of different habitat patches. We then expected an exploitation phase showing less random space use, eventually leading to the formation of home ranges. Furthermore, as the relocated animals came from three different sources with different degrees of similarity with the release site (see below), it is possible that some animals would show different strategies.

In a recent study, a memory-based movement model similar to the ones that we propose below was fitted to roe deer reintroduced into a novel environment, showing that home ranges in the absence of territoriality could emerge from the benefits of using memory during foraging (Ranc et al., 2020). Here we followed a similar approach, but placed emphasis on comparisons among alternative movement models. This allowed us to reveal possible differences in behaviors across individuals. We also paid special attention to the estimates of certain key parameters characterizing informed movement, such as the rate at which an animal used memory, and whether memory decayed over time and how.

We present four simple patch-to-patch movement models, defined through the probabilities of transiting from one patch to another. The simplest model is memoryless as it assumes that the transition probabilities only depend on the distance between the two patches and on the size of the target patch. For simplicity, we do not consider other patch variables such as patch quality. The remaining three models consider the role of memory. The manner in which we introduce memory in the dynamics is similar to that of Boyer and Solis-Salas (2014) and Falcón-Cortés et al. (2017): the probability to revisit a particular patch is modified by a factor which depends on the accumulated number of past visits to this patch, such that the most visited patches have a higher probability to be revisited. In these memory-based models we assume that animals remember patch locations (spatial memory) and the number of past visits to each patch (attribute memory). The main difference between these three models is the way in which animals use their memory. In the simplest case we suppose that animals have infinite memory, i.e., they can remember all the patches previously visited, and they use their memory at a constant rate. In another model we assume infinite memory but the rate at which the animal decides to use its experience increases with the number of explored patches. In the last model

we relax the assumption of infinite memory by introducing a memory decay associated to each patch visit (McNamara and Houston, 1985), whereas the rate of memory use increases as in the previous model.

2. METHODS

2.1. Ranging Data

We used data collected and presented by Frair et al. (2007); see also Wolf et al. (2009). The study area consisted of 15,800 km^2 along the eastern slopes of the Rocky Mountains in central Alberta, Canada. Approximately 2,000 elk inhabited the area during the study period, from December 2000 to September 2002 (Frair et al., 2007). Elevation was 500–1,500 m and the area was largely forested (68.7% of the total area). Dominant tree species included lodgepole pine *Pinus Contorta*, white spruce *Picea Glauca*, and aspen/poplar *Populus Tremuloides* and *P. Balsamea*. Interspersed throughout the forested matrix were wet and dry meadows (7.1%), cutover forest following timber harvest (4.3%), bare soil/rock outcrops (12.3%), rivers and lakes (2.1%), and areas regenerating from wildfire or site reclamation (<1%) (Frair et al., 2007; Wolf et al., 2009).

Over the study period, female elk were translocated to the study area from three source sites within Alberta: (1) Banff and Jasper National Parks, mountainous areas with the full suite of predators present in the study area but protected from hunting, (2) Cross Ranch Conservation Area (ca 20 km southwest of Calgary), a hunted area of foothills and agricultural lands largely without predators, and (3) Elk Island National Park, a flat aspen parkland without predators or hunters, see Frair et al. (2007) for more details about these three sites. Collared animals included six females from the town site of Banff released in February 2001. Nine females were released from the Cross Area, six during December 2000 and three in December 2001, and six females were released from Elk Island between January and February 2002. The animals were captured primarily using corral traps baited with hay. These animals were transported to release areas in livestock trailers that held between 9 and 16 animals depending on the sex and age class composition. Elk were released directly from the trailers into the study area. The animals were released in a number of separate locations to increase independence between results from different individuals (Frair et al., 2007; Wolf et al., 2009).

Prior to release, translocated elk were fitted with GPS collars (LMRT4 and GPS2200, Lotek Wireless, ON, Canada) that collected locations every 2 h for up to 11 months. We used all locations of each collared animal during a season or until radio-contact was lost, the animal died, or GPS collars were retrieved via breakaway device (11 months post-release). All collars were equipped with mortality sensors that activated after 7 h of immobility. Collar tests across the range of cover and terrain conditions encountered within the study area indicated a high fix rate and positional accuracy of ≤50 m 80% of the time (Frair et al., 2007; Wolf et al., 2009).

Foraging patches were defined based on a 27-class landcover grid developed for this region (see Frair et al., 2005). The grid had a 28.5 m cell size, and an overall classification accuracy of 82.7%. Using ArcGIS (Environmental Systems Research Institute, Redlands, California), we combined those classes where elk can find forage [dry/mesic and wet meadows, shrubland, clearcuts, and reclaimed herbaceous (pipeline)] into a single foraging habitat class. Then, we converted the grid to a polygon layer without simplifying lines, which is equivalent to an 8-cell neighborhood rule for patch definition. We eliminated polygons <0.27 ha in size (essentially <3 contiguous pixels), and retained 16,782 patches for analysis. The resulting foraging patches averaged 6.93 ± 29.4 ha in size. For each elk GPS location occurring within a patch, we recorded the unique number for that patch, which allowed us to derive information on the time spent moving between foraging patches, the residency time within patches, and the return time to previously visited patches. Thus, we transformed the original GPS trajectories into a time series of patch to patch visits which included the time spent in each patch and the time traveling between patches. We assumed that most foraging occurred in these high biomass patches. **Figure 1** shows the map of the study area with the distribution of the foraging patches, as well as four representative trajectories during summer and winter for two elk.

2.2. Models

For each model below, we made the following assumptions:

- The animals were moving in a stationary 2d environment which consisted of a set of N available patches (resource sites), N is obtained from environmental data as detailed in the previous subsection. Patches were characterized by their area a_n, with n in $\{1, ..., N\}$. The Euclidean distance between the centroids of the patches n and m is denoted by $d_{n,m}$.
- We modeled discrete movement events: at each time step $t \rightarrow t + 1$ an animal decides to move to another patch (patch-to-patch movement) following a set of rules that we will explain below. The model does not take into account the actual time spent in a patch or between patches, and consider each trajectory as a whole without making distinction between seasons.
- An animal will go from patch n to patch m with probability $P_{n,m}$. This probability were computed in different ways for each model.
- All the parameters to estimate were positive numbers.

2.2.1. Model I

The first model is Markovian as it assumes that the forager chooses to visit a patch (m) in the environment by considering the distance ($d_{n,m}$) from its current patch (n) and the area (a_m) of the patch m. We define a probability vector $\mathbf{k} = (k_1, ..., k_N)$ whose m-th entry denotes the probability that the animal goes to patch m from patch n. Each entry is defined by:

$$k_m = d_m * c_m / \sum_r d_r * c_r, \quad r = 1, ..., N \quad (1)$$

with,

$$d_m = \exp(-(d_{n,m}/\alpha)^\beta)$$

FIGURE 1 | Map of the study area with the distribution of the foraging patches. As an illustrative example, the trajectories between patches of two representative elk are displayed, in summer (white circles) and in winter (black circles).
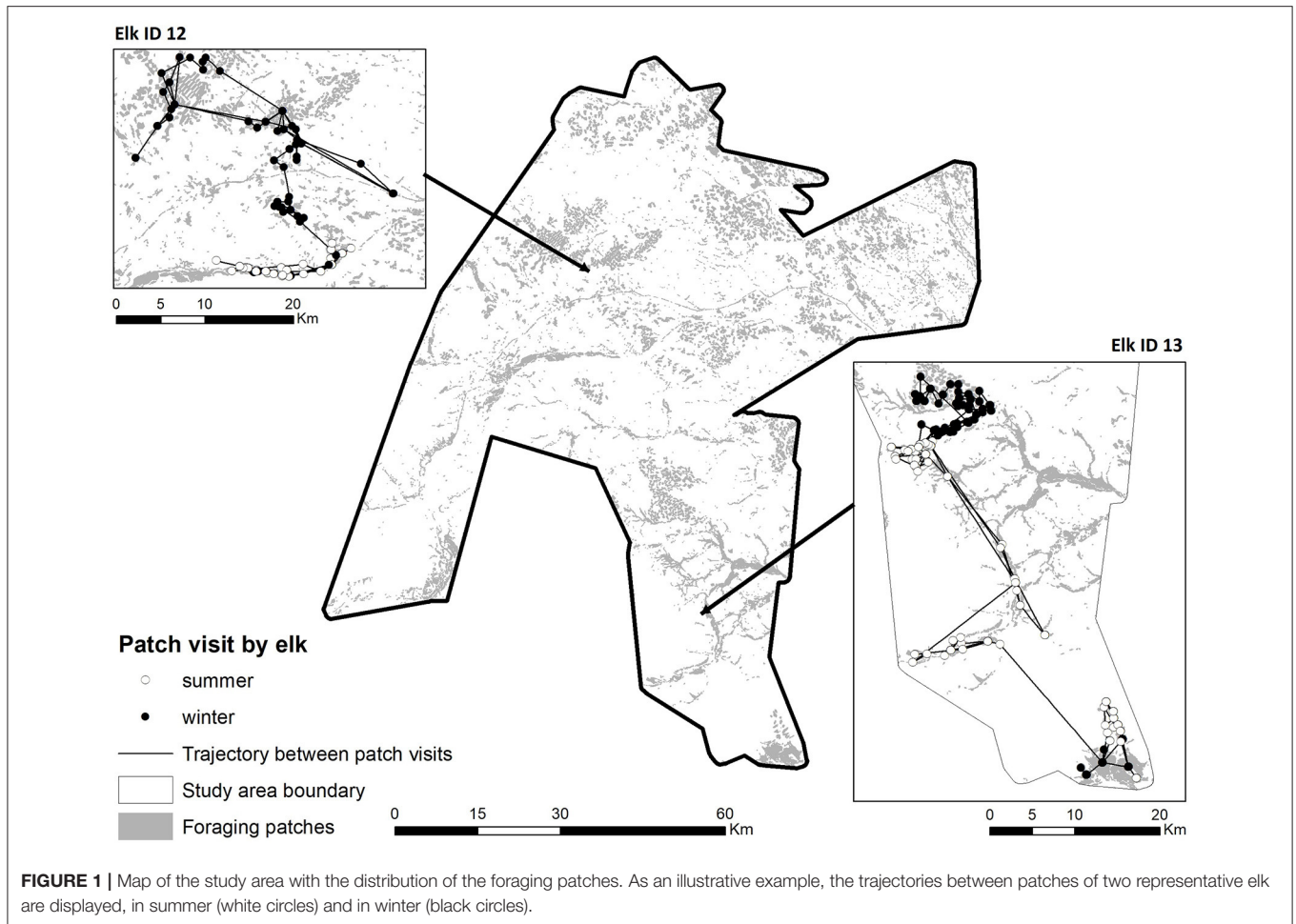
and

$$c_m = \exp(x_m)/(1 + \exp(x_m))$$

where $x_m = \lambda a_m + \kappa$, i.e., we assume that the probability to visit patch m decays exponentially with the distance to patch n ($d_{n,m}$) and increases with the area of patch m (a_m). This model (as well as the others below) does not specifically consider a variable for the patch quality, but assumes that the animals have a higher probability to cross a large patch than a small one. We aim at obtaining a hierarchical estimation for the parameters α, β, λ, and κ (see **Table 1**).

2.2.2. Model II

We next incorporate memory effects through a parameter $q \in (0, 1)$ that defines the probability with which an animal decides to use its experience to revisit a patch. In this Model II, we assume that the forager has infinite memory, i.e., is capable of remembering all previously visited sites. Linear reinforcement is implemented by setting that the probability to choose a particular site for revisit is proportional to the accumulated number of visits to that site. This model has two types of movement decisions:

- With probability q the forager moves from patch n to patch m considering, besides the distance and area, the number of visits that patch m has received in the past. The entry m of the probability vector k is now defined by:

$$k_m = d_m * c_m * m_m / \sum_r d_r * c_r * m_r, \quad r = 1, ..., N \quad (2)$$

with d_m and c_m defined as in (1) and $m_m = n_m$, where n_m is the number of visits at site m until the present time t. Hence, $m_m = 0$ if the animal has never visited m.

- With probability $1 - q$ the forager does not use its memory and will choose a patch m using the probability vector k defined in (1). Hence the forager performs an exploratory movement.

2.2.3. Model III

Given that the data trajectories belong to animals that were released in an unfamiliar environment, it is reasonable to hypothesize that movements were dominated by exploration at early times and by memory at later times. In such case, one may allow the memory parameter q to vary with time.

In this model, the memory parameter depends on the number of unique visited sites (UVS) of the forager up to time t. To this end, we define $\mathbf{u} = (u_1, ..., u_T)$ as a vector of length T, with T the

TABLE 1 | Prior distributions.

Parameter	Prior distribution	Interpretation
α	Normal (0, 10)	Scale parameter for the exponential function that defines the probability decay with distance.
β	Normal (0, 1)	Shape parameter for the exponential function that defines the probability decay with distance.
λ	Normal (0, 1)	Slope parameter for the logit function that defines the probability increase as function of patch area.
κ	Normal (0, 1)	Intercept parameter for the logit function that defines the probability increase as function of patch area.
q	Beta (1, 1)	Parameter that defines the memory use frequency.
ρ	Normal (0, 10)	Scale parameter for the exponential function that defines the increase of probability memory use as function of the number of unique visited sites
ϵ	Normal (0, 1)	Shape parameter for the exponential function that defines the increase of probability memory use as function of the number of unique visited sites.
ν	Normal (0, 10)	Scale parameter for the exponential function that defines memory decay as function of time since last visit for each patch.
θ	Normal (0, 10)	Shape parameter for the exponential function that defines memory decay as function of time since last visit for each patch.

trajectory length and u_T the number of distinct patches visited by the forager up to time T ($u_1 = 1$). This vector is an observed data and q will depend on it as follows:

$$q(u_t) = 1 - \exp(-(u_t/\rho)^\epsilon) \tag{3}$$

In this model the total number of parameters to estimate is six, four of them already considered in Model I, plus two parameters for the increase of memory use as function of the UVS (ρ and ϵ, see **Table 1**).

2.2.4. Model IV

So far we have considered in Model II and III that foragers possess infinite memory. Besides, we have considered that reinforcement is linear, i.e., that an animal chooses a site for revisit with probability proportional to the total number of visits to that site. To incorporate memory decay, we assume in Model IV that the weight of any visit decays exponentially in time, from the value unity. Hence, the animal will forget those visits that are far away in the past and will remember very well those that are recent. Therefore, the recently visited sites have a larger probability to be visited again.

The memory factor defined in Model II now takes the form:

$$\mathrm{m}_m(t) = \sum_{i=1}^{n_m} \exp\{-[(t - t_i)/\nu]^\theta\} \tag{4}$$

with n_m the number of visits to patch m until time t, and t_i the time at which the i-th visit to this patch occurred. It is important to note that m_m defined in Equation (4) will be characterized by an exponential memory decay for $\theta = 1$, a stretched exponential decay for $\theta < 1$, and a super-exponential decay for $\theta > 1$. In this model, one needs to estimate eight parameters. The six parameters already considered in Model III and two more describing memory decay (ν and θ, see **Table 1**).

We fitted these four models to the data and then we performed a model comparison. We used two different tools to perform this comparison: a Posterior Predictive Check (PPC) to asses the model's ability to "predict" the data used to parameterize it, and the Watanabe-Akaike Information Criterion (WAIC) (Watanabe and Opper, 2010) as an approximation for out of sample predicting capacity of each model. These two tools help us to compare the four models above. Specifications about fitting and comparison are shown in the next sections.

2.3. Model Fitting

For some parameters such as q, the frequency of memory use, we used non-informative priors while for other parameters we used weakly informative priors (**Table 1**). All priors were truncated to take only non-negative values.

The models were fitted by using a two-stage approach as proposed by Hooten and Hefley (2019). Such fitting procedure was necessary because fitting the hierarchical level in only one stage would have been intractable computationally, regarding both memory and execution time. The first stage involves fitting the set of individual-level models independently using placeholder priors for all model parameters. Each individual has its own set of parameters for each model. This first-stage was achieved using Hamiltonian Monte Carlo (HMC) techniques implemented within the software Stan (Carpenter et al., 2017) and accessed via RStan (Stan Development Team, 2018). For all models we ran three HMC chains with 5,000 iterations for Model I and II, 10,000 iterations for Model III and IV. We discarded the first half of the iterations for warm-up, and obtained a Rhat < 1.1 and a reasonable number of effective samples (n_eff), from which the posterior distribution of all parameters were obtained. For each animal, the starting point of the fitting simulation was taken as the first visited patch observed. More details about how we performed the simulations are given in **Supplementary Material** (A Guide Example).

The second stage involved a simple MCMC algorithm to fit the full hierarchical Gaussian model using the posteriors from the first stage as priors (Hooten and Hefley, 2019). This second stage ran only one chain with 7,500 (15,000) iterations (the union of the three chains from the first stage) with 3,750

(7,500) iterations for warmup, a p−value p_v > 0.05 for the Geweke's statistic and a reasonable n_eff for all the relevant parameters in the different model dynamics. With this second step, we obtained the posterior distributions at the individual-level for the parameters of each animal (this fit takes into account the variability between individuals) as well as the posterior distributions of the parameters at the population-level.

2.4. Model Assessment and Comparison

In order to assess and compare the descriptive and predictive capacity of the different models, we use two kinds of tools: one qualitative and the other quantitative.

As qualitative assessments, we performed PPC on the number of unique patches visited by the animals through time. That is, for each animal we determined the number of unique patches visited (or UVS) as a function of the number of between-patch movements and compared this quantity with the predictions of simulated trajectories from the different models. For each simulated trajectory, we used parameter combinations sampled from the joint posterior of each of the corresponding model. For each model and individual animal, we simulated 1,000 trajectories, with initial position as same as the observed one, and we checked whether the observed change in number of UVS fell within the credible interval of the simulated ones. We thus could asses whether the observed pattern was consistent with the parameterized model.

As a quantitative assessment of model predictive capacity we used WAIC (Watanabe and Opper, 2010). This quantity is computed from the log-pointwise-predictive-density of each model, which was calculated from the posterior distributions obtained from the second-stage algorithm. This quantity helped us to suggest the best model for each individual: we say that a model is the best when it obtained the lowest WAIC and when the difference between this and other model's WAIC were >2.

3. RESULTS

3.1. Model Comparison

Considering the PPC for all individuals and models (**Supplementary Figure 9**), we found that six trajectories (out of the 21 individuals) were contained within the 95% credible interval (CI) of Model I, while 17 did so for Model II, 10 for Model III and 15 for Model IV.

The WAIC comparisons displayed in **Table 2** show us that Model I was not the best model for any individual, i.e., the calculated WAIC for Model I was never the smallest one for any animal. Model II had the smallest WAIC for 12 individuals. Model III was the best for nine animals, and Model IV was not the best for any individual. Therefore, in most cases, a constant rate of memory use and a linear reinforcement without memory decay provided a good description of their trajectories. These results agree qualitatively with those of the PPC.

To illustrate these general results, we present a closer analysis of the PPC and WAIC for four representative individuals that portray different kinds of behaviors on a trajectory. **Figure 2** displays the PPC for each model and animals 1, 7, 11, and 17. **Table 2** shows WAIC for all models and the same four representative animals in gray. The lowest WAIC between

models for each individual is indicated in bold. We denoted as δ the difference between the WAIC of each model and the lowest one, and P_W as the effective number of parameters.

Figure 2-first Row shows the PPC results for individual 1 from Banff. Model I fitted well only the first steps of the trajectory, indicating that the animal was probably in exploration phase. Later on, the trajectory is no longer contained within Model I credible interval. Model II fitted well the final steps of the trajectory from this animal, suggesting that it followed an exploitation phase with q = 0.68 (from here on, all reported parameters values are the mean from their correspond posterior distribution). However, like Model I, neither Model II OR IV described the entire time series. Thus, Model III was the only acceptable model for animal 1, indicating that this particular individual increased its memory use as it explored the environment. In agreement with this finding, Model III had the lowest WAIC for this animal (**Table 2**).

Figure 2-second Row displays the PPCs for animal 7 from Cross Ranch. Here, Models I and III fitted well just the first trajectory steps, indicating a exploration phase, but overall, they were not acceptable for animal 7. In contrast, Model IV contained all the observed trajectory within its CI, suggesting that this particular individual increased its memory use as it explored the space and its memory decayed over time. Model II was also acceptable for animal 7, with a constant rate of memory use of q = 0.38. Therefore, animal 7 had two possible acceptable models. However, the lowest WAIC for individual 7 was for Model II, and the δ for Model IV was quite large (**Table 2**).

Figure 2-third Row corresponds to animal 11, also from Cross Ranch. Models I, III, and IV fitted well only the first steps of the observed trajectory. Model II contained within its CI the entire observed data, indicating that this particular animal used its memory at a constant and very high rate (q = 0.80), being most of the time visiting known patches. **Table 2** indicated the lowest WAIC for Model II, confirming the conclusion drawn from the PPC.

Figure 2-fourth Row corresponds to animal 17 from Elk Island. Model III fitted well just the first steps of the trajectory and it was not acceptable for this individual. Otherwise Models I, II, and IV contained within their respective CI all the observed trajectory. This give us three possible interpretations for animal 17: (i) The animal was always in exploratory phase. (ii) The individual used its memory at constant rate q = 0.25. (iii) The animal increased its memory use with time and its memory decayed over the time. **Table 2** shows that Model II actually had the lowest WAIC. Therefore, Model II can be considered as fairly good to describe and predict the trajectory of animal 17.

Table 2 summarizes models fit to each elk by their source population. Elk from 1 to 6 belong to the Banff and Jasper Source, animals from 7 to 15 to Cross Ranch, and elk from 16 to 21 to Elk Island. Models having the lowest WAIC are bolded. We can see that for all animals from Banff and Jasper, Model III was the best according to WAIC. For animals from Cross Ranch, Model II was the best for most of them. And for 66% of elk from Elk Island, Model II was the best. This suggests that animals from different source populations reacted differently to the new environment.

We discuss in the following the different parameters obtained from the fits.

TABLE 2 | WAIC from the pointwise log-likelihood for each model and each individual.

Source	ID	Model I WAIC Est	δ	Pw Est	Model II WAIC Est	δ	Pw Est	Model III WAIC Est	δ	Pw Est	Model IV WAIC Est	δ	Pw Est
	1	916.51	148.28	12.06	788.51	20.28	7.27	**768.23**	**0.00**	**7.05**	801.51	33.28	11.23
	2	1864.50	221.87	8.50	1645.72	3.09	8.46	**1642.63**	**0.00**	**9.50**	1724.00	81.37	13.08
	3	2751.13	342.30	5.60	2410.23	1.40	5.73	**2408.83**	**0.00**	**10.16**	2499.04	94.21	14.99
Banff and Jasper	4	295.62	38.37	13.92	263.21	6.32	7.25	**256.89**	**0.00**	**5.80**	267.09	10.20	6.98
	5	611.20	57.11	21.75	561.77	7.68	33.43	**554.09**	**0.00**	**29.81**	653.32	99.23	76.39
	6	1161.61	169.74	21.04	994.70	2.83	31.52	**991.87**	**0.00**	**29.74**	1031.30	39.43	33.59
	7	1670.40	70.08	8.93	**1600.32**	**0.00**	**11.42**	1654.73	54.41	18.34	1689.76	89.44	24.60
	8	398.77	33.54	8.78	370.41	5.18	2.83	**365.23**	**0.00**	**1.70**	372.14	6.91	2.32
	9	846.68	54.29	11.59	**792.39**	**0.00**	**9.40**	812.27	19.88	9.68	834.13	41.74	11.40
	10	793.78	21.14	10.00	**772.64**	**0.00**	**5.81**	857.71	85.07	16.11	846.62	73.98	9.43
Cross Ranch	11	3406.23	677.88	6.71	**2728.35**	**0.00**	**7.34**	2755.67	27.32	5.99	2868.43	140.08	18.11
	12	1259.08	98.78	7.88	**1160.30**	**0.00**	**11.72**	1193.49	33.19	18.01	1223.01	62.71	15.93
	13	1732.02	160.37	7.32	**1571.65**	**0.00**	**10.58**	1577.72	6.07	9.88	1627.82	56.17	14.03
	14	1020.30	102.70	7.39	**917.60**	**0.00**	**7.29**	921.85	4.25	7.55	951.79	34.19	9.45
	15	2561.35	213.99	7.37	**2347.36**	**0.00**	**5.49**	2394.94	47.58	18.72	2467.82	120.46	23.22
	16	289.98	3.03	13.39	**286.95**	**0.00**	**15.68**	295.28	8.33	16.24	296.51	9.56	18.12
	17	499.80	34.65	27.65	**465.15**	**0.00**	**10.51**	468.45	3.30	8.85	469.82	4.67	10.55
Elk Island	18	1604.57	118.75	6.80	**1485.82**	**0.00**	**7.30**	1485.86	0.04	5.80	1540.36	54.54	10.08
National Park	19	130.98	106.86	53.14	27.41	3.29	2.84	**24.12**	**0.00**	**1.40**	26.20	1.21	1.81
	20	23.02	5.18	0.78	18.05	0.21	0.69	**17.84**	**0.00**	**0.24**	18.92	1.08	0.27
	21	241.79	72.34	50.91	**169.45**	**0.00**	**12.35**	176.70	7.25	13.92	178.27	8.82	15.73

Table shows point estimates (Est) for information criterion WAIC, the effective number of parameters (P_W), and difference between WAIC's models as δ. In bold the lowest WAIC for each individual.

3.2. Spatial Parameters

The spatial parameters α, β, λ, and κ are present in all models. The estimated values for these parameters do not vary too much between the four different models. We present here a common interpretation for these parameters. From now on the analysis focuses on the individual-level estimate of each parameter [say p_j ($j = 1 : 21$)] as well as on the population-level parameter p.

Parameter α, which controls the scale of the exponential decay with distance between patches (see **Supplementary Tables 6–9**) fluctuated little among individuals and across the four models [$0.60 \leq \alpha_j \leq 2.57$ (km)], with a population average between models of $\alpha = 1.71$. Parameter β, which controls the shape of the exponential decay, varied between 0.64 and 1.50 among individuals, with a population average of $\beta = 1.09$, i.e., close to the exponential shape. These values mean that distance played an important role in patch selection; the animals did not choose patches beyond one or two kilometers from their actual positions (maybe due to the patchiness of the environment) as shown by the posterior curve in **Figure 3** (Top). These results highlight the importance of "distance discounting" in movement choices, even when memory was involved.
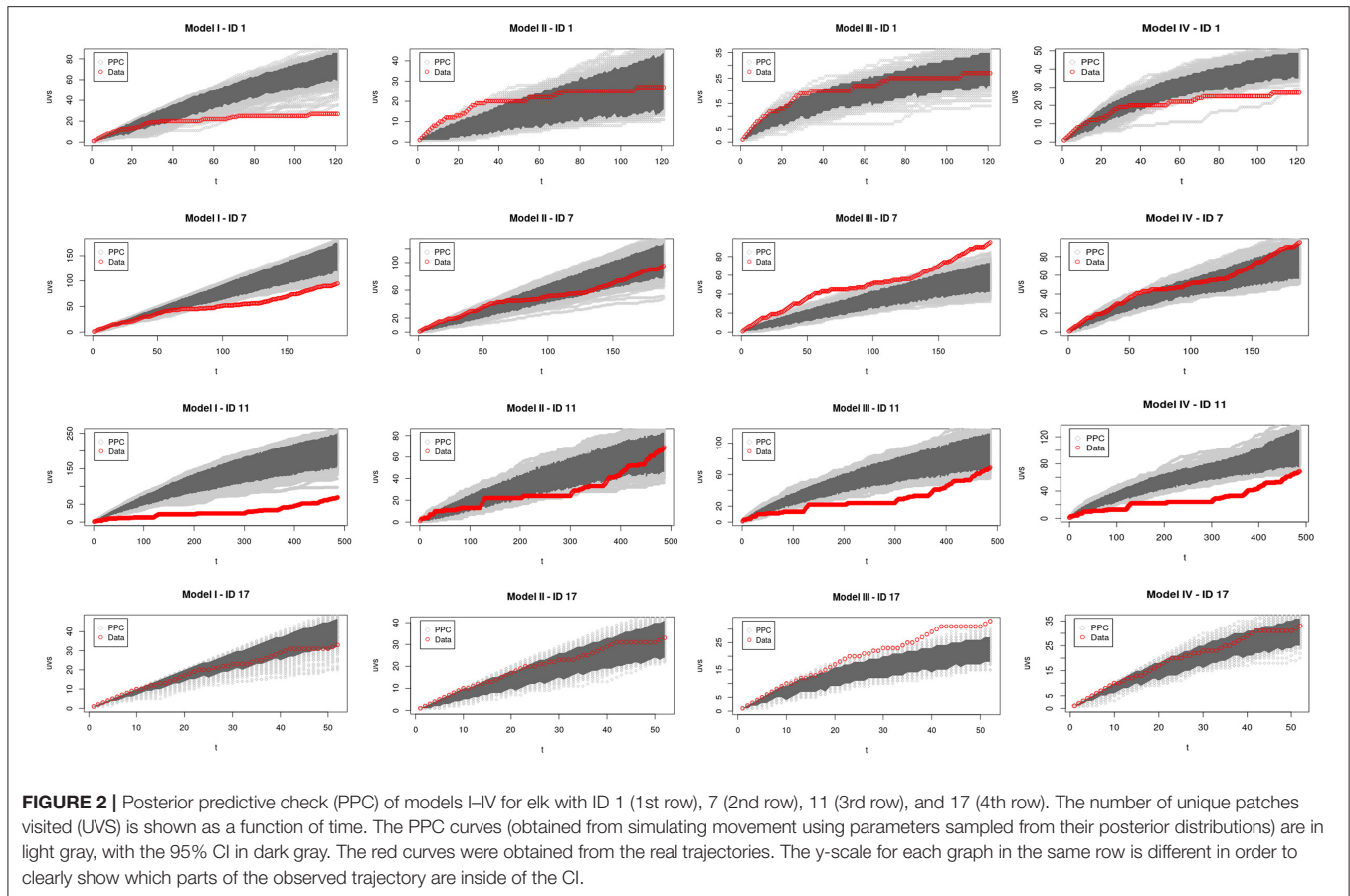
Parameter λ, which controls the slope of the logit increase with patch area, also fluctuated little among individuals and models [$2.34 \leq \lambda_j \leq 3.70$ (ha)], with a population average of $\lambda = 2.90$. Whereas, parameter κ, which controls the intercept of the

logit increase, had fluctuations between 0.02 and 0.42 among individuals, and a population average of $\kappa = 0.14$. We conclude that patch area played a significant role during patch use: the probability increased rapidly for patches of area around 1 ha, and saturated for patches with area >2 ha as shown by the posterior curve in **Figure 3** (Bottom).

3.3. Memory Use

Figure 4 displays the marginal posterior distributions of the parameter q, that defines the probability of memory use in Model II. As mentioned earlier, this model was considered the best for 12 individuals. For these individuals q had a minimum value of 0.18 and a maximum of 0.80, but most of them had a $q \approx 0.5$. Hence, according to this model, roughly half of the moves from patch to patch performed by most of the animals are informed by memory, while the other half can be considered as exploratory. For those animals with values of q far from 0.5, the trajectories are either dominated by memory (e.g., ID 11) or by exploratory movements (e.g., ID 10 and 21).

Model III assumes that q grows from zero with the number of UVS at time t (u_t), as defined by Equation (3). **Figure 5** (Top) displays the marginal posterior distributions of the parameter ρ. This parameter defines the number of visited sites needed for the onset of important memory effects. For those individuals for which this Model III was considered
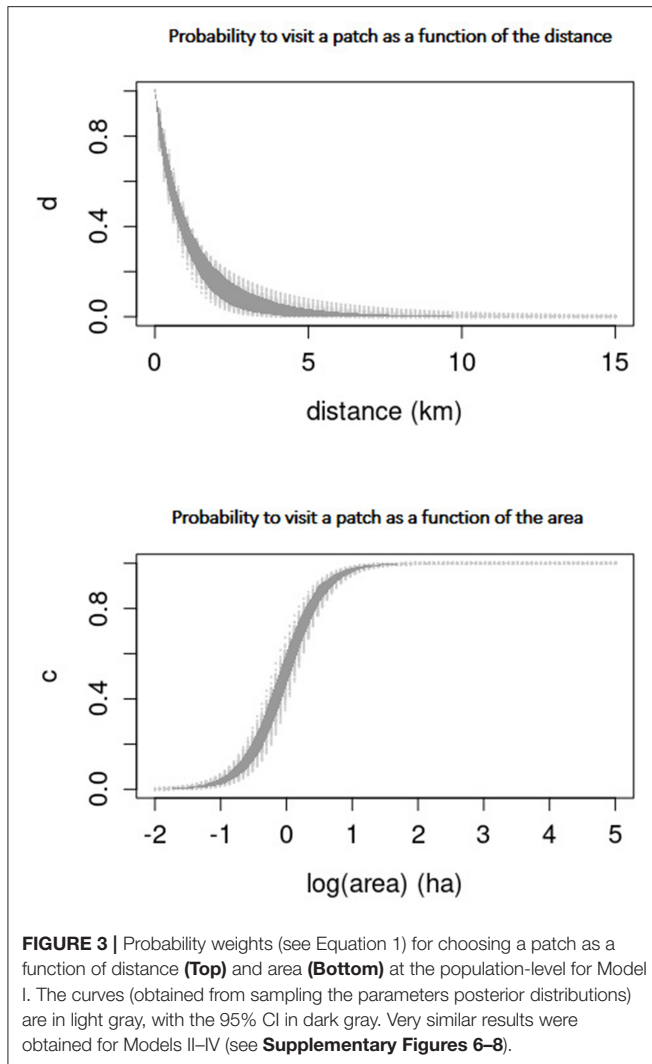
FIGURE 2 | Posterior predictive check (PPC) of models I–IV for elk with ID 1 (1st row), 7 (2nd row), 11 (3rd row), and 17 (4th row). The number of unique patches visited (UVS) is shown as a function of time. The PPC curves (obtained from simulating movement using parameters sampled from their posterior distributions) are in light gray, with the 95% CI in dark gray. The red curves were obtained from the real trajectories. The y-scale for each graph in the same row is different in order to clearly show which parts of the observed trajectory are inside of the CI.

the best ρ had values between 12.87 and 12.93. Likewise, the shape parameter ϵ **Figure 5** (Center) of the exponential ranged between 0.03 and 0.86. **Figure 5** (Bottom) displays the growth of memory use as a function of **u** at the population-level. Memory use increased rapidly to 0.5 when the unique visited sites were between 5 and 10, before slowly tending to its asymptotic value.

3.4. Memory Decay

Model IV takes into account all the assumptions of Model III, with the addition of a decay in memory. **Figure 6** (Top) displays the marginal posterior distribution of the scale parameter ν that defines the time scale of memory decay. For the population-level this parameter was estimated as $\nu = 10.78$. The mean shape parameter θ was estimated as $\theta = 0.30$, thus, memory decayed on average more slowly than exponentially, namely, as a stretched exponential [**Figure 6** (Center)]. Actually, **Figure 6** (Bottom) reveals that the weight of a visit to a patch (whose initial value is 1) often decayed very slowly in time. In many realizations, this weight remained significant (> 0.2) even after a number of steps ($t \sim 100$) much larger than the mean half-time (estimated as ~ 5 steps, or ~ 30 h in physical time). Therefore, the half-time of memory decay is not very meaningful here.
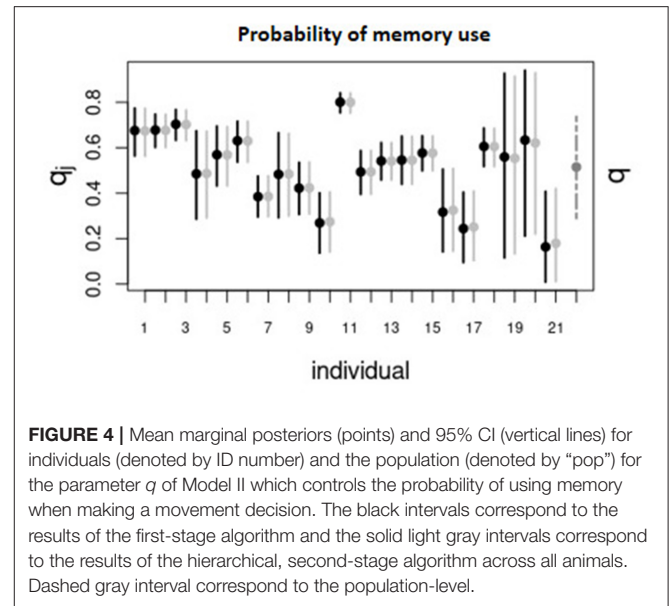
4. DISCUSSION

We have presented four simple models to fit a set of movement data collected in western Canada for 21 elk relocated into a new environment. In a first stage, Bayesian estimates were carried out at the individual-level using Hamiltonian Monte Carlo sampling. A hierarchical analysis was next implemented following the algorithm proposed in Hooten and Hefley (2019), allowing us to infer how the population as a whole is adapting to a new environment. All the results obtained at the individual-level can be found in the **Supplementary Material**. To compare and evaluate these four different models we used two tools, one quantitative the other qualitative. We used, on the one hand the Watanabe Akaike Information Criterion (WAIC) and on the other hand, a Posterior Predictive Check (PPC) based on the number of unique patches visited by the animals. The results obtained by these two tests were in agreement. We found that the trajectories of all animals were far from being described by a memoryless random walk and rather exhibited patterns of recurrent revisits to patches. Although it is possible that some unquantified patch feature makes them more attractive to the animals and hence more likely to be revisited, it is unlikely that the pattern of patch use described by the time series of unique patches visited can result from memoryless movement. In other words, unobserved patch features will have to have a

FIGURE 3 | Probability weights (see Equation 1) for choosing a patch as a function of distance **(Top)** and area **(Bottom)** at the population-level for Model I. The curves (obtained from sampling the parameters posterior distributions) are in light gray, with the 95% CI in dark gray. Very similar results were obtained for Models II–IV (see **Supplementary Figures 6–8**).

FIGURE 4 | Mean marginal posteriors (points) and 95% CI (vertical lines) for individuals (denoted by ID number) and the population (denoted by "pop") for the parameter q of Model II which controls the probability of using memory when making a movement decision. The black intervals correspond to the results of the first-stage algorithm and the solid light gray intervals correspond to the results of the hierarchical, second-stage algorithm across all animals. Dashed gray interval correspond to the population-level.

very particular distribution in order to match observed changes in patch use with time.

Our Models II and III, that consider an infinite memory capability (with constant and dynamic rate of use, respectively) combined with a linear reinforcement of the visited patches, fitted and predicted well all the trajectories. This is consistent with the results exposed by Wolf et al. (2009) in which a thorough statistical study of habitat selection found that elk had a strong tendency to select the most recently visited locations to forage instead of selecting locations only by their quality. Moreover, the values of our spatial parameters, and the dispersion curves that they defined, corresponded well with resident elk movement scales reported in Frair et al. (2005). Foraging movements were of the order of hundreds of meters and relocating moves of the order of 1.6 km. Our fourth model, that considered a dynamic use of memory as well as memory decay, was not considered as the best model for any individual. That model therefore seems too sophisticated for this population over the observation period.

The exploitation-exploration paradigm is a well known concept in ecology. There are several models that have focused on identifying and predicting these two phases from single animal trajectories (Morales et al., 2004; Jonsen et al., 2007) but they are often based on memoryless dynamics and the exploitation-exploration phases are the result of different types of random walks movements. Our Model II is memory-based and the use of memory is governed by a constant parameter q. While the exploration phase corresponds to random decisions unrelated with experience, the exploitation phase is ruled by the use of memory and the reinforcement learning acquired by experience. These simple assumptions were enough to adequately represent the temporal changes in the number of unique patches visited (**Supplementary Figure 9**) by twelve animals and therefore to identify the presence of these two phases. It is important to note that those twelve individuals for which Model II was considered the best model, as well as the nine animals for which Model III gave the best results, had a high value of q (near 1/2, typically). This suggests that these animals used memory intensively, instead of performing pure random walks (which correspond to the limit $q \to 0$). A previous study on capuchin monkeys that used a similar model found a value of q near 0.12 over a 6-month period (Boyer and Solis-Salas, 2014). In that model the environment was represented as a regular discrete lattice in which each point was a site to visit. The high values of q observed in our study could be explained by frequent decisions to return to high-resource patches or safe places, for instance those where the predation risk (by wolves or humans) is lower. The movement patterns produced by these high values of q is also consistent with the scale movement results exposed in Frair et al. (2005) that shows that elk make use of certain patches and do not explore beyond them, possibly to reduce their mortality rate and predation risk.

We also found that animals from the same source population tended to behave similarly: for most of the animals relocated from Banff and Jasper, Model III was considered the best model, whereas most of the elk coming from Cross Ranch and Elk Island

FIGURE 5 | (Top and middle) Same as Figure 4 but for parameters ρ and ϵ (resp.) of Model III. (Bottom) Increase of the probability of memory use q as a function of the number of unique visited sites u at the population-level. The curves (obtained from sampling parameters from the joint posterior distributions) are in light gray, with the 95% CI in dark gray.

FIGURE 6 | Same as Figure 4 for the parameters ν (Top) and θ (middle) of Model IV. (Bottom) Decay of the weight of a visit to a patch, $\exp\{-[(t - t_i)/\nu]^\theta\}$ vs. t, with , $t_i = 1$, sampled from the joint posterior distribution of ν and θ.

were best described by Model II. These results might be explained by the experience animals had before translocation: we speculate that if the original environment was similar to the new one or the animal was not naive to predators, the animal relied more heavily on memory as they visited new patches (Model III). Conversely, if the original environment was very different or the animals naive to predators, then the they kept high rates of exploration (Model
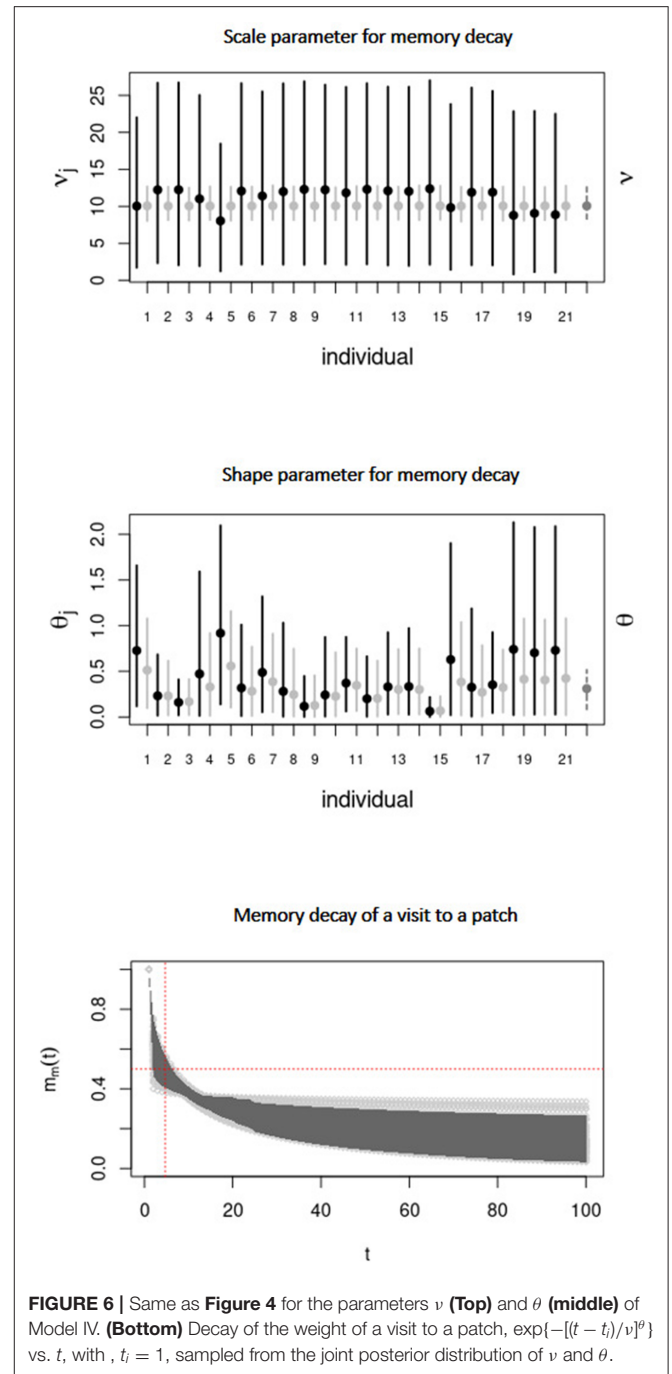
II). This hypothesis stems from the fact that Banff and Jasper are mountainous with similar kinds of valley meadows as the new habitat, and that the animals were familiar with predators, while Cross Ranch and Elk Island have quite different habitat backgrounds, mostly wide-open areas dominated by agriculture and flatland, respectively, and with animals naive to predators.

It is noteworthy that the model in which memory decays with time (Model IV) was not supported as the best model for any of the animals during the period of this study. This suggests that elk remember very well the places they have visited at least within

1 year. Similar findings have been reported for other species such as American bison (Merkle et al., 2014), sheep (Gautestad and Mysterud, 2005), woodland caribou (Avgar et al., 2015), or chimpanzees (Janmaat et al., 2013). These works reported evidence of long-term or very slowly decaying memory, with individuals having the ability to return sometimes to locations which had not been visited for months, or even years.

Studying the movement trajectories of translocated animals provides a promising way to understand how animals use memory. Our findings are qualitatively consistent with those recently reported by Ranc et al. (2020) on reintroduced roe deer. The movements of those animals were described by a model including both memory and resource preferences, somehow similarly to ours in the memory mode, with a reinforcement that saturated to a limiting value instead of growing linearly as here. Their fitted model was able to predict the dynamics of home range formation observed in roe deer, thus bringing support to the hypothesis that memory constitutes an important mechanism for home range emergence (Börger et al., 2008; Van Moorter et al., 2009). Although not analyzed in detail here, it is very likely that the models that we have fitted would also predict several movement properties indicative of limited space use and home range behavior in elk, but it would be important to have longer observation periods to verify this.

Our models and data analysis show a clear effect of distance and patch area on the probability of a patch being used in the next move. Thus, the configuration of patches in the landscape will affect how space is used and how memory is built. Several extensions would make these models more realistic and complex. For example, the probability of moving from one patch to another could be affected not only by distance and patch area but also by more realistic estimates of movement costs due to topography and other landscape variables such as different habitat types and predation risk between patches. Furthermore, it would be interesting to include properties of patches that would make them more or less attractive, and also to consider potential seasonal changes in these attributes.

Compared to previous work that studied habitat selection in these animals (Frair et al., 2007), our model is quite coarse as we are only considering moves from patch to patch without taking into account how animals go from one patch to another or how they move inside patches. It would also be possible to consider continuous time modeling taking into account the time that an animal spend going from one patch to another, as well as the residence time within patches. The residence time is a key movement component, which can exhibit high variations within home ranges due to a higher selectivity among habitat types (Van Moorter et al., 2016). Our modeling approach also ignored the fact that in a network of patches, nearby patches can compete as possible destinations due to their spatial configuration (Ovaskainen and Cornell, 2003). This effect can be approximated by considered all possible ways in which an individual leaving a particular patch can eventually reach another patch in the network, although the computational costs are substantial (Morales et al., 2017).

Our models could capture features of the movement patterns of the study animals with a minimum number of parameters and rather simple dynamical rules. Such simplicity is advantageous if one wishes to apply the same models to other data sets. Particularly, a single parameter q quantifies the behavior of an animal memory-wise, and can serve as a basis for comparisons between individuals or between species. Substantial variations of this parameter among individuals of a same species and in a same environment, as observed here, indicate that the movement strategies employed are quite flexible.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: please contact Prof. Evelyn Merrill (emerrill@ualberta.ca) for dataset details. Requests to access these datasets should be directed to Prof. Evelyn Merrill.

ETHICS STATEMENT

The animal study was reviewed and approved by Animal capture, handling, and transportation procedures were approved by Province of Alberta (permit No. 1432GP) and University of Alberta (permit No. 300401).

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.702925/full#supplementary-material

REFERENCES

Avgar, T., Baker, J. A., Brown, G. S., Hagens, J. S., Kittle, A. M., Mallon, E. E., et al. (2015). Space-use behaviour of woodland caribou based on a cognitive movement model. *J. Anim. Ecol.* 84, 1059–1070. doi: 10.1111/1365-2656.12357

Avgar, T., Deardon, R., and Fryxell, J. M. (2013). An empirically parameterized individual based model of animal movement, perception, and memory. *Ecol. Model.* 251, 158–172. doi: 10.1016/j.ecolmodel.2012.12.002

Berger-Tal, O., and Avgar, T. (2012). The glass is half-full: overestimating the quality of a novel environment is advantageous. *PLoS ONE* 7:e34578. doi: 10.1371/journal.pone.0034578

Bischof, R., Loe, L. E., Meisingset, E. L., Zimmermann, B., Van Moorter, B., and Mysterud, A. (2012). A migratory northern ungulate in the pursuit of spring: jumping or surfing the green wave? *Am. Nat.* 180, 407–424. doi: 10.1086/667590

Bonnell, T. R., Campenní, M., Chapman, C. A., Gogarten, J. F., Reyna-Hurtado, R. A., Teichroeb, J. A., et al. (2013). Emergent group level navigation: an agent-based evaluation of movement patterns in a folivorous primate. *PLoS ONE* 8:e78264. doi: 10.1371/journal.pone.0078264

Börger, L., Dalziel, B. D., and Fryxell, J. M. (2008). Are there general mechanisms of animal home range behaviour? A review and prospects for future research. *Ecol. Lett.* 11, 637–650. doi: 10.1111/j.1461-0248.2008.01182.x

Boyer, D., Falcón-Cortés, A., Giuggioli, L., and Majumdar, S. N. (2019). Anderson-like localization transition of random walks with resetting. *J. Stat. Mech.* 2019:053204. doi: 10.1088/1742-5468/ab16c2

Boyer, D., and Solis-Salas, C. (2014). Random walks with preferential relocations to places visited in the past and their application to biology. *Phys. Rev. Lett.* 112:240601. doi: 10.1103/PhysRevLett.112.240601

Boyer, D., and Walsh, P. D. (2010). Modelling the mobility of living organisms in heterogeneous landscapes: does memory improve foraging success? *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 368, 5645–5659. doi: 10.1098/rsta.2010.0275

Bracis, C., and Mueller, T. (2017). Memory, not just perception, plays an important role in terrestrial mammalian migration. *Proc. R. Soc. B Biol. Sci.* 284:20170449. doi: 10.1098/rspb.2017.0449

Brown, C. (2001). Familiarity with the test environment improves escape responses in the crimson spotted rainbowfish, melanotaenia duboulayi. *Anim. Cogn.* 4, 109–113. doi: 10.1007/s100710100105

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., et al. (2017). Stan: A probabilistic programming language. *J. Stat. Softw.* 76, 1–31. doi: 10.18637/jss.v076.i01

Fagan, W. F., Lewis, M. A., Auger-Méthé, M., Avgar, T., Benhamou, S., Breed, G., et al. (2013). Spatial memory and animal movement. *Ecol. Lett.* 16, 1316–1329. doi: 10.1111/ele.12165

Falcón-Cortés, A., Boyer, D., Giuggioli, L., and Majumdar, S. N. (2017). Localization transition induced by learning in random searches. *Phys. Rev. Lett.* 119:140603. doi: 10.1103/PhysRevLett.119.140603

Falcón-Cortés, A., Boyer, D., and Ramos-Fernández, G. (2019). Collective learning from individual experiences and information transfer during group foraging. *J. R. Soc. Interface* 16:20180803. doi: 10.1098/rsif.2018.0803

Frair, J. L., Merrill, E. H., Allen, J. R., and Boyce, M. S. (2007). Know thy enemy: experience affects elk translocation success in risky landscapes. *J. Wildl. Manage.* 71, 541–554. doi: 10.2193/2006-141

Frair, J. L., Merrill, E. H., Visscher, D. R., Fortin, D., Beyer, H. L., and Morales, J. M. (2005). Scales of movement by elk (*Cervus elaphus*) in response to heterogeneity in forage resources and predation risk. *Landsc. Ecol.* 20, 273–287. doi: 10.1007/s10980-005-2075-8

Gautestad, A. O., and Mysterud, I. (2005). Intrinsic scaling complexity in animal dispersion and abundance. *Am. Nat.* 165, 44–55. doi: 10.1086/426673

Hooten, M. B., and Hefley, T. J. (2019). *Bringing Bayesian Models to Life*. Boca Raton, FL: CRC Press. doi: 10.1201/9780429243653

Janmaat, K. R., Ban, S. D., and Boesch, C. (2013). Chimpanzees use long-term spatial memory to monitor large fruit trees and remember feeding experiences across seasons. *Anim. Behav.* 86, 1183–1205. doi: 10.1016/j.anbehav.2013.09.021

Janson, C. H. (2007). Experimental evidence for route integration and strategic planning in wild capuchin monkeys. *Anim. Cogn.* 10, 341–356. doi: 10.1007/s10071-007-0079-2

Janson, C. H., and Byrne, R. (2007). What wild primates know about resources: opening up the black box. *Anim. Cogn.* 10, 357–367. doi: 10.1007/s10071-007-0080-9

Jesmer, B. R., Merkle, J. A., Goheen, J. R., Aikens, E. O., Beck, J. L., Courtemanch, A. B., et al. (2018). Is ungulate migration culturally transmitted? Evidence of social learning from translocated animals. *Science* 361, 1023–1025. doi: 10.1126/science.aat0985

Jonsen, I. D., Myers, R. A., and James, M. C. (2007). Identifying leatherback turtle foraging behaviour from satellite telemetry using a switching state-space model. *Mar. Ecol. Prog. Ser.* 337, 255–264. doi: 10.3354/meps337255

Lanner, R. M. (1996). *Made for Each Other: A Symbiosis of Birds and Pines*. Oxford; New York, NY: Oxford University Press.

Leonard, B. (1990). "Spatial representation in the rat: conceptual, behavioral," in *Neurobiology of Comparative Cognition*, eds R.P. Kesner and D. S. Olton (New York, NY: Psychology Press; Taylor & Francis Group), 363.

Lihoreau, M., Chittka, L., and Raine, N. E. (2011). Trade-off between travel distance and prioritization of high-reward sites in traplining bumblebees. *Funct. Ecol.* 25, 1284–1292. doi: 10.1111/j.1365-2435.2011.01881.x

McNamara, J. M., and Houston, A. I. (1985). Optimal foraging and learning. *J. Theor. Biol.* 117, 231–249. doi: 10.1016/S0022-5193(85)80219-8

Merkle, J., Fortin, D., and Morales, J. M. (2014). A memory-based foraging tactic reveals an adaptive mechanism for restricted space use. *Ecol. Lett.* 17, 924–931. doi: 10.1111/ele.12294

Merkle, J. A., Sawyer, H., Monteith, K. L., Dwinnell, S. P., Fralick, G. L., and Kauffman, M. J. (2019). Spatial memory shapes migration and its benefits: evidence from a large herbivore. *Ecol. Lett.* 22, 1797–1805. doi: 10.1111/ele.13362

Morales, J. M., di Virgilio, A., del Mar Delgado, M., and Ovaskainen, O. (2017). A general approach to model movement in (highly) fragmented patch networks. *J. Agric. Biol. Environ. Stat.* 22, 393–412. doi: 10.1007/s13253-017-0298-1

Morales, J. M., Haydon, D. T., Frair, J., Holsinger, K. E., and Fryxell, J. M. (2004). Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology* 85, 2436–2445. doi: 10.1890/03-0269

Morales, J. M., Moorcroft, P. R., Matthiopoulos, J., Frair, J. L., Kie, J. G., Powell, R. A., et al. (2010). Building the bridge between animal movement and population dynamics. *Philos. Trans. R. Soc. B Biol. Sci.* 365, 2289–2301. doi: 10.1098/rstb.2010.0082

Mueller, T., and Fagan, W. F. (2008). Search and navigation in dynamic environments-from individual behaviors to population distributions. *Oikos* 117, 654–664. doi: 10.1111/j.0030-1299.2008.16291.x

Normand, E., and Boesch, C. (2009). Sophisticated euclidean maps in forest chimpanzees. *Anim. Behav.* 77, 1195–1201. doi: 10.1016/j.anbehav.2009.01.025

Noser, R., and Byrne, R. W. (2007). Travel routes and planning of visits to out-of-sight resources in wild chacma baboons, papio ursinus. *Anim. Behav.* 73, 257–266. doi: 10.1016/j.anbehav.2006.04.012

Ovaskainen, O., and Cornell, S. (2003). Biased movement at a boundary and conditional occupancy times for diffusion processes. *J. Appl. Probabil.* 2003, 557–580. doi: 10.1239/jap/1059060888

Poor, E. E., Loucks, C., Jakes, A., and Urban, D. L. (2012). Comparing habitat suitability and connectivity modeling methods for conserving pronghorn migrations. *PLoS ONE* 7:e49390. doi: 10.1371/journal.pone.0049390

Potts, J. R., and Lewis, M. A. (2014). How do animal territories form and change? Lessons from 20 years of mechanistic modelling. *Proc. R. Soc. B Biol. Sci.* 281:20140231. doi: 10.1098/rspb.2014.0231

Presotto, A., and Izar, P. (2010). Spatial reference of black capuchin monkeys in Brazilian Atlantic forest: egocentric or allocentric? *Anim. Behav.* 80, 125–132. doi: 10.1016/j.anbehav.2010.04.009

Ranc, N., Cagnacci, F., and Moorcroft, P. R. (2020). Memory drives the formation of animal home ranges: evidence from a reintroduction. *BioRxiv [Preprint].* 1–42. doi: 10.1101/2020.07.30.229880

Schacter, D. L. (1992). Priming and multiple memory systems: Perceptual mechanisms of implicit memory. *J. Cogn. Neurosci.* 4, 244–256. doi: 10.1162/jocn.1992.4.3.244

Schlägel, U. E., and Lewis, M. A. (2014). Detecting effects of spatial memory and dynamic information on animal movement decisions. *Methods Ecol. Evol.* 5, 1236–1246. doi: 10.1111/2041-210X.12284

Schlägel, U. E., Merrill, E. H., and Lewis, M. A. (2017). Territory surveillance and prey management: Wolves keep track of space and time. *Ecol. Evol.* 7, 8388–8405. doi: 10.1002/ece3.3176

Spencer, W. D. (2012). Home ranges and the value of spatial information. *J. Mammal.* 93, 929–947. doi: 10.1644/12-MAMM-S-061.1

Stan Development Team (2018). *Rstan: The R Interface to Stan. R Package Version 2.17. 3.*

Van Moorter, B., Rolandsen, C. M., Basille, M., and Gaillard, J.-M. (2016). Movement is the glue connecting home ranges and habitat selection. *J. Anim. Ecol.* 85, 21–31. doi: 10.1111/1365-2656.12394

Van Moorter, B., Visscher, D., Benhamou, S., Börger, L., Boyce, M. S., and Gaillard, J.-M. (2009). Memory keeps you at home: a mechanistic model for home range emergence. *Oikos* 118, 641–652. doi: 10.1111/j.1600-0706.2008.17003.x

Watanabe, S., and Opper, M. (2010). Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *J. Mach. Learn. Res.* 11, 3571–3594. Available online at: http://jmlr.org/papers/v11/watanabe10a.html

Wills, T. J., Cacucci, F., Burgess, N., and O'Keefe, J. (2010). Development of the hippocampal cognitive map in preweanling rats. *Science* 328, 1573–1576. doi: 10.1126/science.1188224

Wolf, M., Frair, J., Merrill, E., and Turchin, P. (2009). The attraction of the known: the importance of spatial familiarity in habitat selection in wapiti *Cervus elaphus. Ecography* 32, 401–410. doi: 10.1111/j.1600-0587.2008.05626.x